A combinatorial toolbox to explore the usage of sesquiterpene synthases:

structure - function studies, substrate promiscuity and directed evolution

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

Víctor González Requena

Supervisor: Rudolf K. Allemann

2019

Abstract

Terpenoids have a wide range of applications in the pharmaceutical, agrochemical, perfume, pigment and biofuel industries. The generation of novel terpenoids and increased understanding of the mechanisms of their formation would be of great value in both academic and industrial settings.

Terpene synthases catalyse the metal-dependent conversion of linear isoprenyl diphosphates into a myriad of cyclic or acyclic structures through a series of complex reaction cascades involving high-energy carbocationic intermediates. The final intermediate is quenched through proton loss or nucleophilic attack of water to generate pure hydrocarbon or alcohol compounds.

The ability of terpene synthases to generate hydroxylated products through regio- and stereoselective 'water capture' represents a potential short-cut to the manufacture of oxidised terpenes, which is yet to be fully explored.

The first part of this project focused on two sesquiterpene synthases, (-)-germacradien-4-ol synthase (GdolS) and (+)-epicubenol synthase, which catalyse water quench to generate hydroxylated products. These enzymes were characterised, and their water management mechanisms explored through the use of site-directed mutagenesis (SDM) and product formation analysis of the obtained mutants. This work has revealed a delicate structural task for A176 in GdolS that governs product selection, among other studies.

The second part of this project describes the synthesis of substrate analogues for their use with a selection of sesquiterpene synthases. These analogues were designed to intercept reaction pathways or trap alternative products. The insertion of a C6-fluorine atom can electronically influence the neighbouring carbocation after 1,6-cyclisation. C6-methyl group incorporation affect the enzyme-substrate complex geometry through steric effected and conformational changes. The absence of a methyl group at C7 gives novel insights into the requirement needed for catalysis in sesquiterpene synthases because it is a potentially more flexible substrate and reduces the stability of the forming carbocation intermediate after 1,6-cyclisation. Thiirane analogues were also synthesised to explore their acceptance by the sesquiterpene synthases.

The last section describes the developmental work for a directed evolution approach to increase the catalytic performance of sesquiterpene synthases. This method combines the use of error-prone PCR to diversify the sesquiterpene synthases encoding genes and a colorimetric assay to screen sesquiterpene synthase activities.

Acknowledgements

This work was supported by Cardiff University and the UK's Biotechnology and Biological Sciences Research Council (BBSRC).

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. It has been a pleasure to work under your supervision.

Thanks to Dr. David Miller for his help proofreading this work. Extended thanks to Dr. Verónica González and Dr. Katherine Williams, for teaching me the molecular biology techniques during the first year. Also, thanks to Dr. Rob Jenkins and Thomas Williams for their help with NMR and MS spectroscopy techniques.

Thanks to all the members from the Allemann group, you have contributed to make my studies more enjoyable. Thanks to Dr. Chris Jones, for his assistance in the organic lab. Thanks to Dr. Aduragbemi Adesina, Dr. Alan Scott, Dr. Luke Johnson, Dr. Robert Mart, Dr. Antonio Angelastro, Dr. Alice Dunbabin, Dr. Juan Faraldos and Dr. Louis Luk for their scientific advice and assistance when it was needed. Also, thanks to my colleagues Martin, Dr. Marianna Loizzi, Florence, Gwawr, Huw, Jenny and Gareth.

Special thanks to Dr. Prabhakar Srivastava for his priceless support in this project, and for always having a positive mind.

I would also like to give a special thanks to my friends Raquel and Ed, for their constant help since the beginning of my Ph.D.

I would like to give special thanks to Dr. Alexander Nödling, for his generous support during these years. Also special thanks to Nicolò and Davide for the long talks and the good moments. Thank you guys for your friendship.

Many other people from Cardiff have contributed to make my days better, it has been nice to meet you all. I would also like to thank my friends in Spain, for the funny moments when I have been back.

My sincere gratitude to my family. Special thanks to my sister Cristina and to Mayed, for their constant belief that I can be the best. Also, thanks to my nephews Áxel and Sergi, you have inspired me.

Thank you, Mum and Dad, for everything. Without your support, life lessons and love I would not be where I am today.

Finally, a dearly thanks to Mireya. You have been always holding my hand during this adventure and I would have never done this without your unconditional love.

Table of contents

List of abbreviations

ADS	Amorphadiene synthase
APS	Ammonium persulfate
AT-AS	Aristolochene synthase from *Aspergillus terreus*
ATP	Adenonsine triphosphate
BL21(DE3)-RP	BL21(DE3)-Codon Plus RP cells
βME	2-Mercaptoethanol
BTAC	Benzyltriethylammonium cation
CD	Circular dichroism
CHES	N-Cyclohexyl-2-aminoethanesulfonic acid
CMP	Cytidyl monophosphate
COSY	Correlation spectroscopy
DCM	Dichloromethane
DCS	δ-Cadinene synthase
DHFDP	2,3-Dihydro farnesyl diphosphate
DIBALH	Diisobutyl aluminium hydride
DMADP	Dimethylallyl diphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DOX	Deoxyxylulose phosphate
EBFS	(*E*)-β-Farnesene synthase
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EpicS	Epicubenol synthase
EP-PCR	Error-prone polymerase chain reaction
FDP	Farnesyl diphosphate
FDPS	Farnesyl diphosphate synthase
FGDP	Farnesylgeranyl diphosphate
FGDPS	Farnesylgeranyl diphosphate synthase
FHP	Farnesyl hydroxyphosphonate
FSDP	Farnesyl-*S*-thiolodiphosphate
GAS	Germacrene A synthase
GC-MS	Gas chromatography mass spectrometry
GdolS	Germacradien-4-ol synthase
Gd11olS	Germacradien-11-ol-synthase

GDP	Geranyl diphosphate
GDPS	Geranyl diphosphate synthase
GDS	Germacrene D synthase
GGDP	Geranylgeranyl diphosphate
GGDPS	Geranylgeranyl diphosphate synthase
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	Heteronuclear multiple-bond correlation
HMG-CoA	3-Hydroxy-3-methylglutaryl coenzyme A
HPLC	High pressure liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear Single-Quantum Correlation
HWE	Horner-Wadsworth-Emmons reaction
IDP	Isopentenyl diphosphate
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
HcS	Hedycaryol synthase
MEP	Methylerythritol phosphate
MES	2-(N-Morpholino)ethanesulfonic acid
MVA	Mevalonate pathway
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidised form
NBS	N-Bromosuccinimide
NDP	Nerolidyl diphosphate
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
NTA	Nitrilotriacetic acid
OD	Optical density
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PP_i	Inorganic diphosphate
PR-AS	Aristolochene synthase from *Penicillium roqueforti*
PS	Pentalenene synthase
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SdS	Selinadiene synthase
Sf-CinS1	Cineole synthase from *Salvia fruticose*
So-BPPS	Bornyl diphosphate synthase from *S. officinalis*
So-CinS1	Cineole synthase from *S. officinalis*

So-SabS1	Sabinene synthase from *S. officinalis*
Sp-SabS1	Sabinene synthase from *S. pomifera*
TDP	Thiamine diphosphate
TEAP	Bis-triethylammonium hydrogen monophosphate
TEAS	5-epi-aristolochene synthase from *Nicotiana tabacum*
TEV	Tobacco etch virus
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Tris	Tris-(hydroxymethyl)-amino-methane
UV	Ultraviolet
VT	Variable Temperature
WT	Wild type

CHAPTER 1

INTRODUCTION

Chapter 1. Introduction

1.1. Terpenes. Biological roles and applications

Terpenes and terpenoids [the terpenome] constitute the largest and most diverse family of natural products, accounting for above 80.000 structurally and stereochemical different compounds to date.[1] This chemodiversity is achieved from a small pool of natural occurring metabolites and is fundamentally constituted of isoprene units, as first observed by Otto Wallach.[2] Isoprene units (**1**) are generally linked through head-to-tail (regular), head-to-head (irregular) or tail-to-tail (irregular) elongations, as stated by Leopold Ruzicka in the pioneering `Biogenetic isoprene rule´.[3] Accordingly, terpenes can be categorised depending on the number of C_5 units: Hemiterpenes (C_5), monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), sesterterpenes (C_{25}), triterpenes (C_{30}), sesquaterpenes (C_{35}), tetraterpenes (C_{40}), and higher polyterpenes ($C_{>40}$), Figure 1.

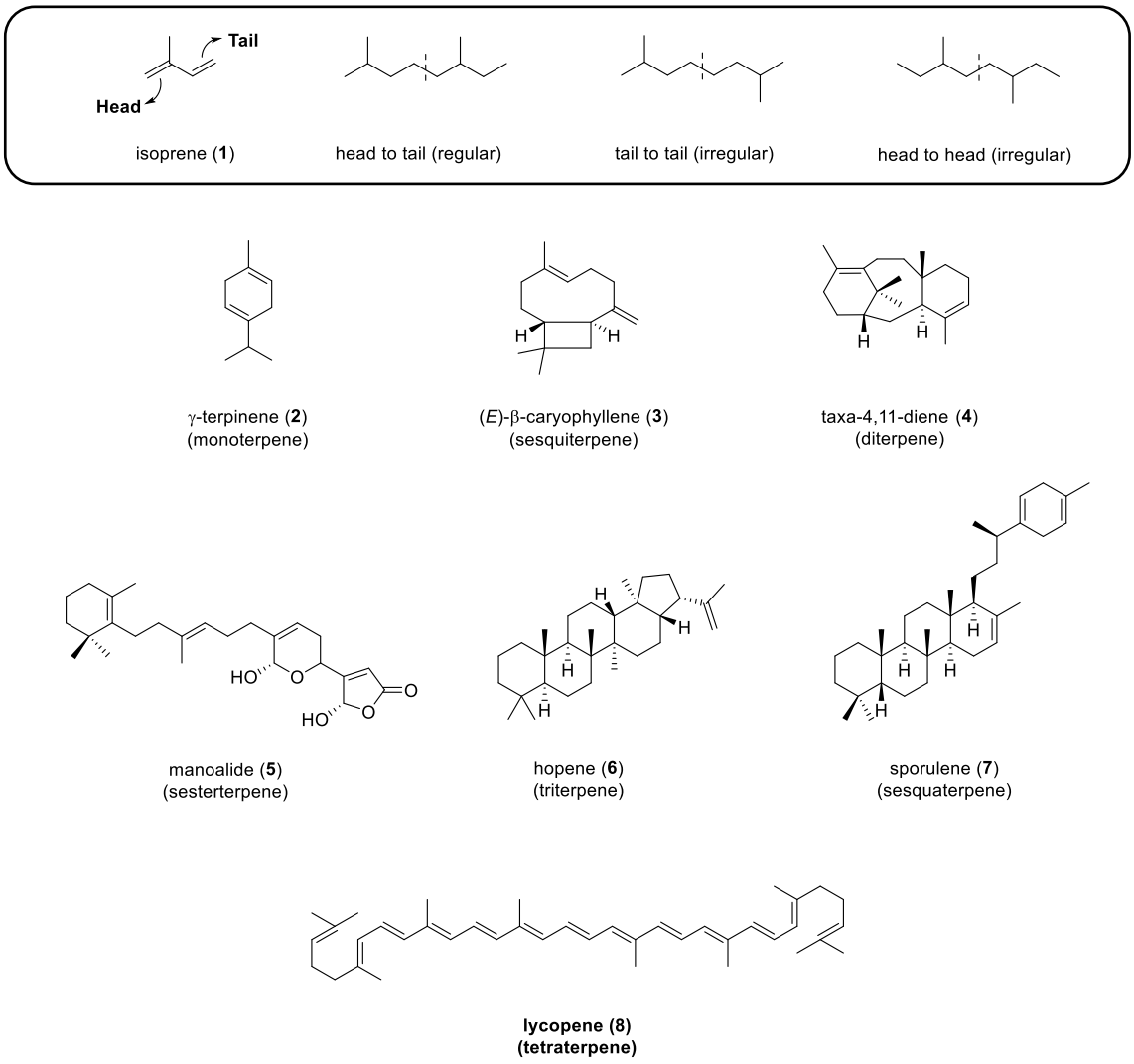


Figure 1. Examples of terpenes including their structural classification. Upper box: Types of isoprene elongations.

"The investigation of the ethereal oils, of both the new and the known constituents, by one single man can embrace only an exceedingly modest territory in proportion to the inexhaustible material which the plant world affords us."

– OTTO WALLACH–

Terpenes play many essential roles in nature, with a wide profile of biological activities in all forms of life.[4,5] Terpenes can serve as *primary metabolites*, involved in growth and development biological processes. Sterols are present in most eukaryotes fulfilling maintenance of membrane physiological properties such as fluidity and permeability and acting as hormones and bile acids.[6,7] While cholesterol (**9**) and ergosterol (**10**) are the major sterols in vertebrates and fungi respectively, plants synthesise a complex mixture of sterols that we refer as phytosterols.[8–10] Hopanoids can be found in primitive bacteria (archae, cyanobacteria, etc) that develop in very extreme conditions.[11] Carotenoids, which are synthesised by plants and microorganisms but not animals, are necessary for photosynthetic processes and function as antioxidants.[12–15] Quinones play a vital role in numerous electrochemical reactions for energy transduction and storage, such as respiration and photosynthesis.[16,17]

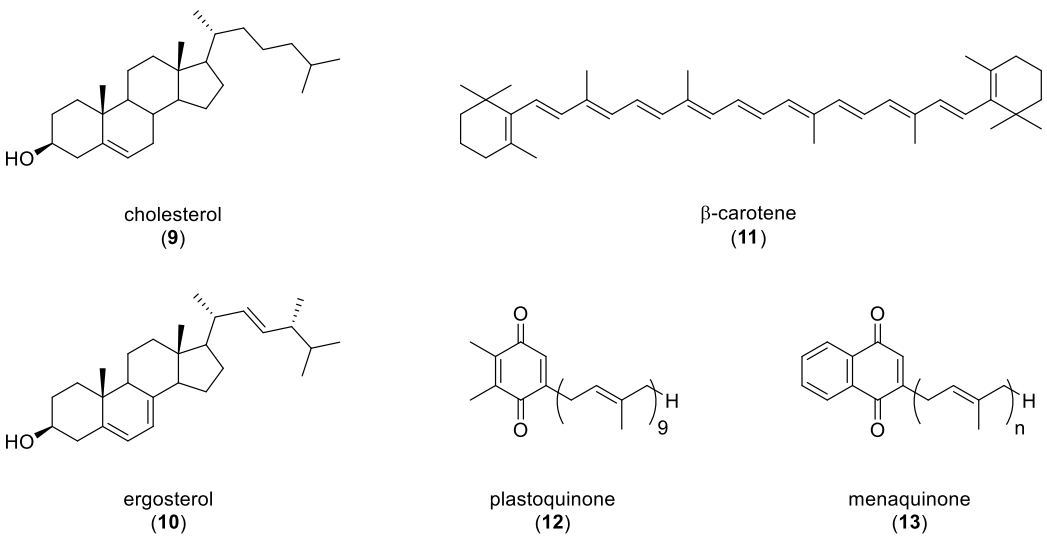


Figure 2. Examples of primary metabolites.

Secondary metabolites are a major role for terpenes in all organisms, mediating antagonistic and mutualistic interactions among them.[4,18,19] Many terpenes have been reported to act as toxins, growth inhibitors or deterrents to microorganisms and animals. Plants make use of terpenes for defensive roles against insect pests, and in resistance to pathogenic fungi and bacteria.[20,21] Insects spray terpenoids at their enemies for defence[22] and, in marine organisms, terpenes are implicated not only in defensive roles against predators and pathogens, but they also serve to prevent their surfaces from being colonised by other organisms.[23,24] On the other hand, terpene-related secondary metabolites are also used for communication with partners. Terpenes are low molecular weight, lipophilic molecules with high vapor

pressures at standard temperatures, and their structural variety make them perfect candidates for very specific messages, specifically, insect-insect, plant-animal/insect and plant-plant interactions, such as signalling for attracting pollinators and pheromones.[25,26]

The vast number of biological interactions in which terpenoids are involved, only few of which have been referred to here, make terpenes of great interest for industrial and medicinal purposes. Since ancient times, terpenoids were used in complex mixtures as medicine, flavouring, scents and recreational drugs. In recent times, the improvements on identification, together with extraction and purification techniques and the access to engineered host organisms for higher production to match society´s demands, make terpenoids industrially relevant chemicals, including pharmaceuticals, flavours, fragrances, cosmetics, food supplements, pesticides and disinfectants.[27–30]

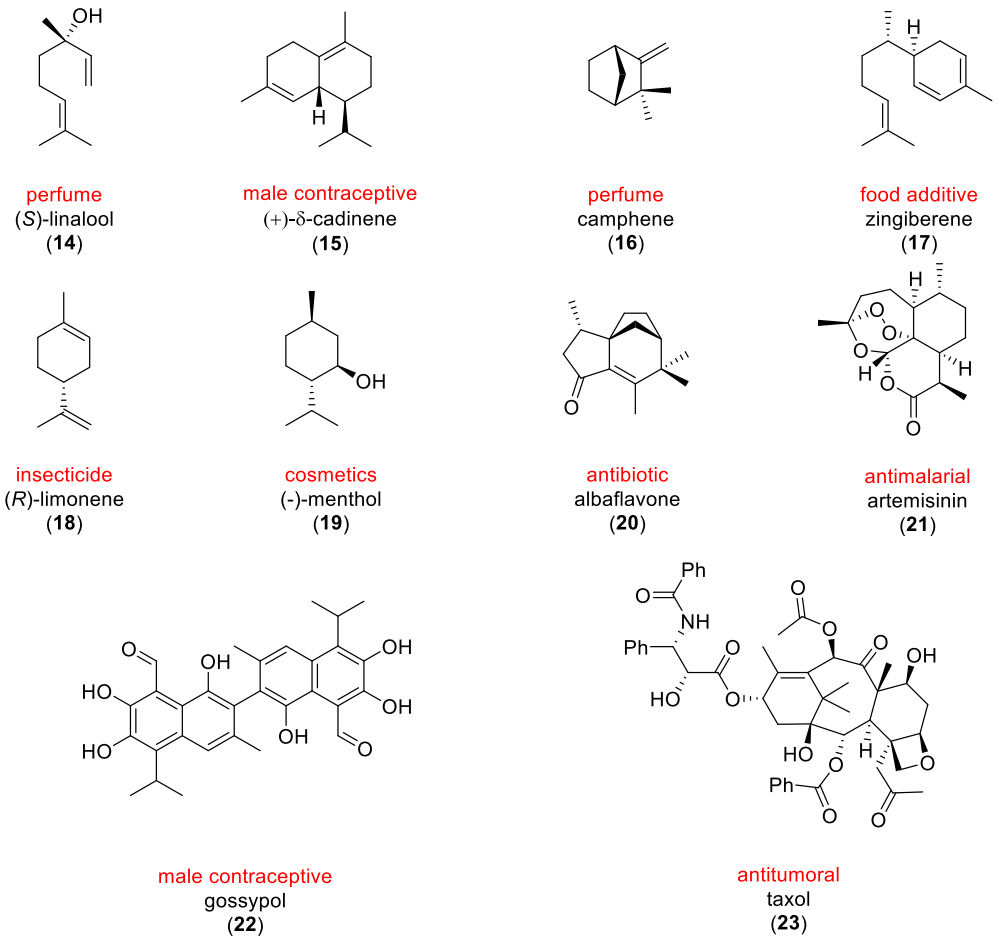
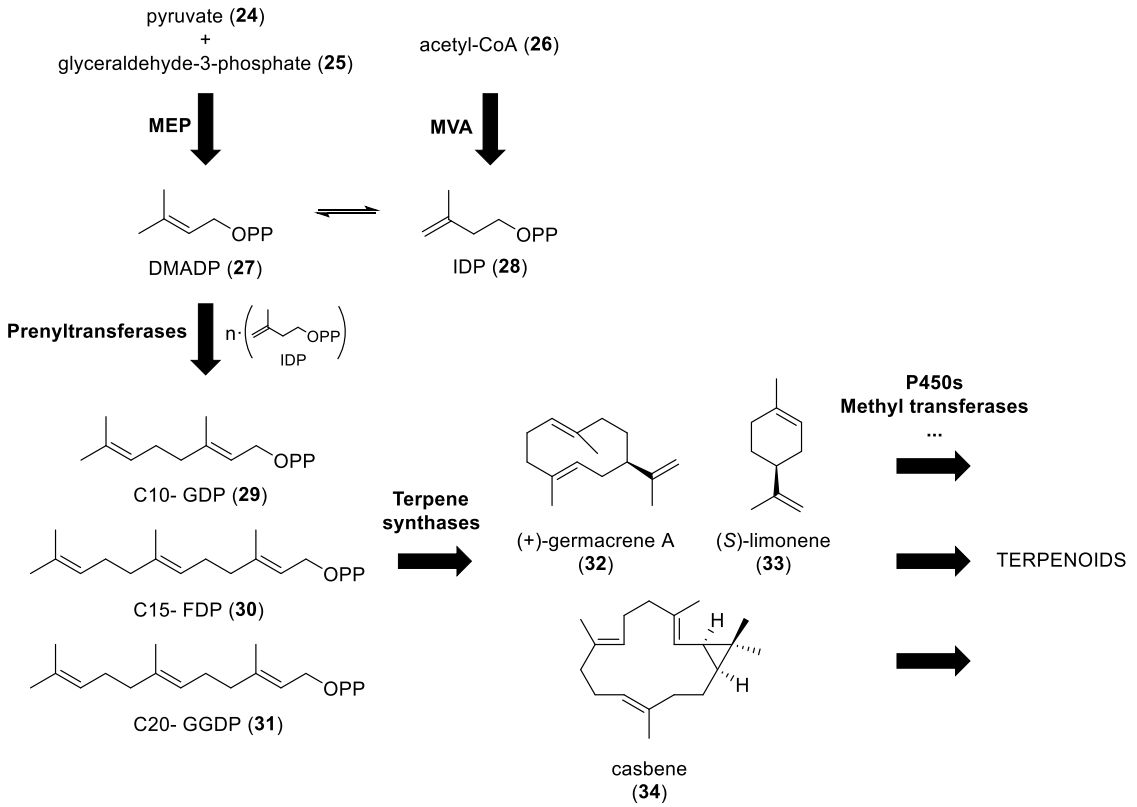


Figure 3. Example of terpenoids with industrial applications.

1.2. Terpenes biosynthesis

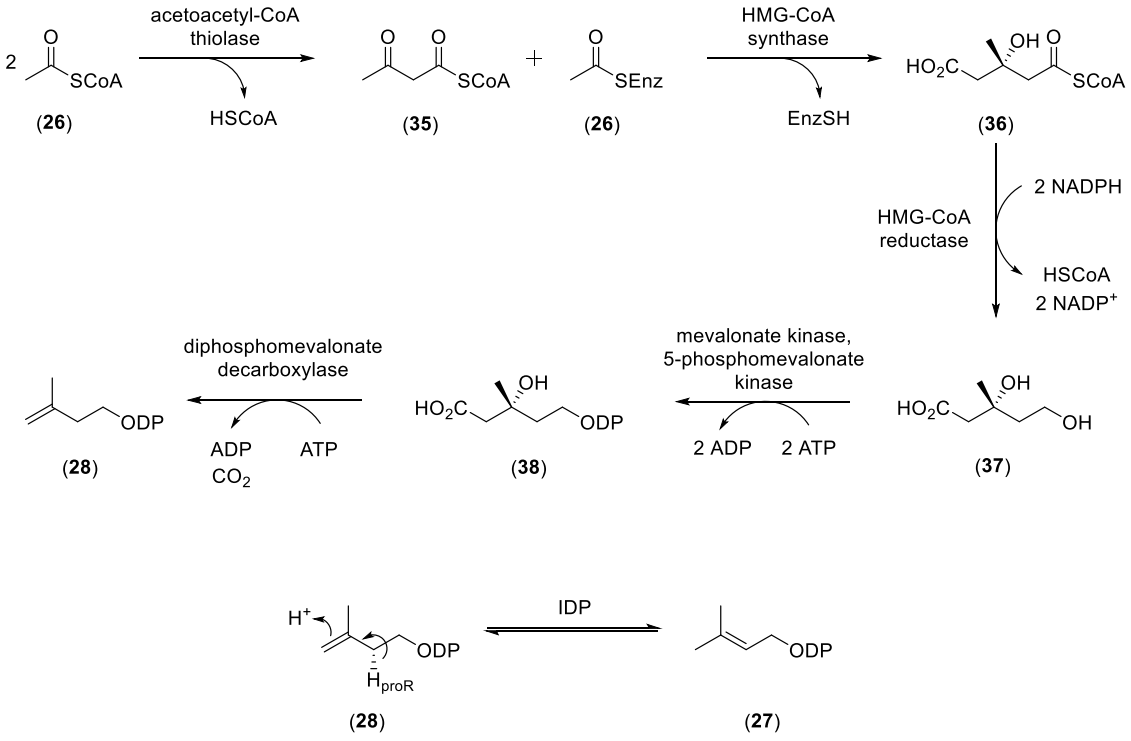
In nature, the extensive terpenome is achieved through an intricate network of biological pathways that can generally be divided into three different stages.[31–34] Initially, all terpenoids are derived from dimethylallyl diphosphate (DMADP, **27**) and isopentenyl diphosphate (IDP, **28**), the activated isoprene counterparts. The biosynthesis of IDP and DMADP takes part in the mevalonate pathway (MVA) and the so-called non-mevalonate pathway (Non-MVA, methylerythritol phosphate (MEP) or deoxyxylulose phosphate (DOX) pathway). Then, the prenyltransferases are responsible for the elongation of the isoprenyl diphosphate units to build higher molecular weight polyprenyl diphosphate precursors. These enzymes generate the 10-, 15- and 20- carbon compounds geranyl diphosphate (GDP, **29**), farnesyl diphosphate (FDP, **30**) and geranylgeranyl diphosphate (GGDP, **31**), respectively. Larger terpene precursors can be assembled through combination of these polyprenyl diphosphates, as shown in Section 1.2.3. Lastly, terpene synthases catalyse the conversion of polyprenyl diphosphate precursors to a myriad number of carbon skeletons that can be further modified by P450s, methyl transferases and others to complete the terpenome, Scheme 1.



Scheme 1. Representation of the biological synthesis of terpenoids.

1.2.1. Mevalonate pathway (MVA)

The mevalonate pathway (Scheme 2) takes place in eukaryotes, the cytosol of plants and in some bacteria.[35–38] The MVA is initiated with the acetyl-coenzyme A thiolase guided Claisen condensation[39] of two molecules of acetyl-coA (**26**) to give acetoacetyl-CoA (**35**). Through the HMG-CoA synthase,[40] a third molecule of acetyl-CoA(**26**) is added *via* an aldol condensation with **35** to give 3-hydroxy-3-methylglutaryl-CoA (**36**). Then, HMG-CoA reductase[41] makes use of two equivalents of NADPH for the synthesis of mevalonic acid (**37**), which is sequentially diphosphorylated by mevalonate kinase[42] and 5-phosphomevalonate kinase[43] to afford 5-diphosphomevalonic acid (**38**) using two molecules of ATP. An ATP-dependant decarboxylation and water departure, which is catalysed by diphosphomevalonate decarboxylase,[44] give isopentenyl diphosphate (IDP, **28**). IDP isomerase catalyses the reversible isomerisation of IDP (**28**) to DMADP (**27**), which proceeds *via* [1,3]-antarafacial proton addition to the *Re* face of the IDP double bond and the removal of the 2-pro-*R*-proton.[45,46]



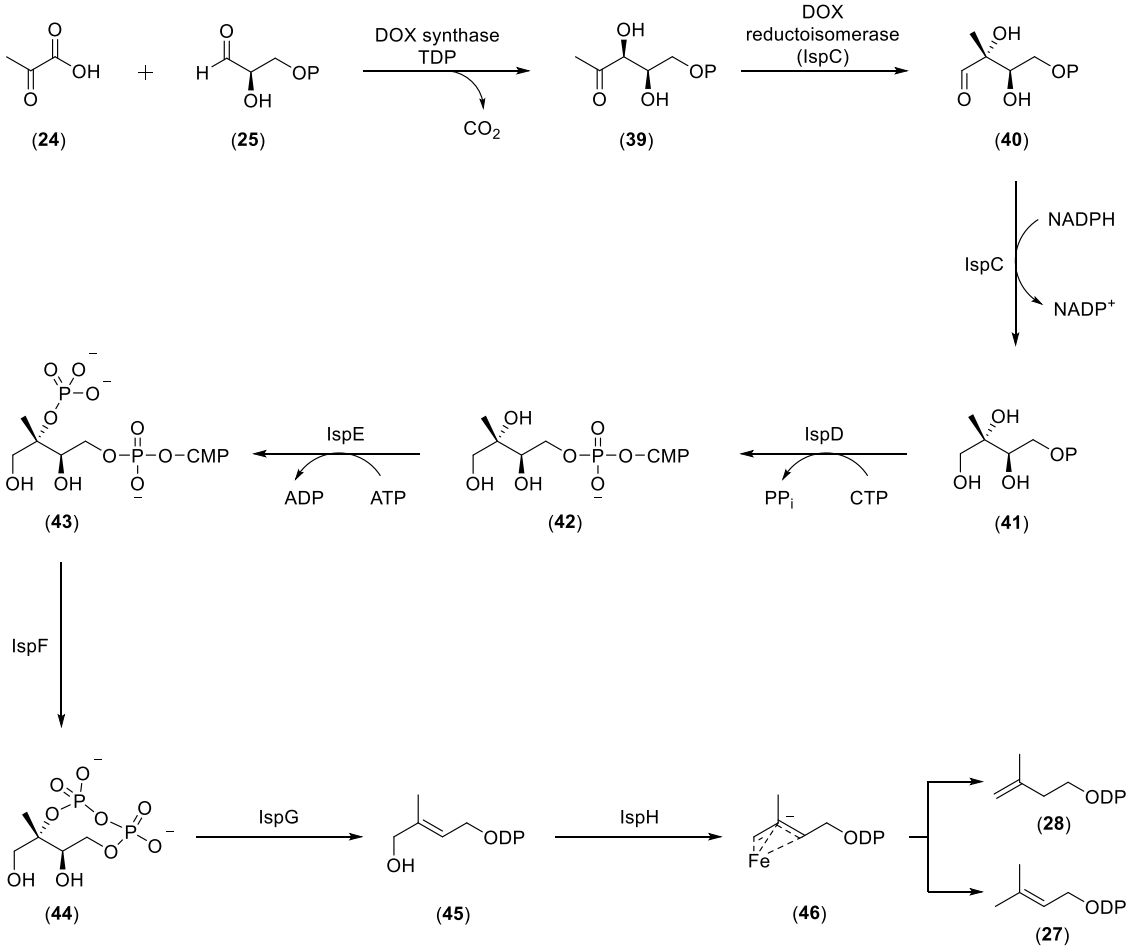
Scheme 2. MVA pathway. Bottom, IDP-DMADP isomerisation.

The discovery of the mevalonate pathway provided a powerful tool for many studies.[47] For example, regulation of elevated cholesterol levels, which is associated to cardiovascular mortality and morbidity, can be achieved through the manipulation of the mevalonate pathway *via* addition of statins, which are HMG-CoA reductase inhibitors.

Experimental inconsistencies in the study of some plastid terpenoids, led to the belief that an alternative pool of substrate could exist. Later, feeding experiments with ^{13}C-labelled acetate and glucose led the research groups of Rohmer and Arrigoni to probe the co-existence of the non-mevalonate pathway.[48,49]

1.2.2. Non-mevalonate pathway

The non-mevalonate pathway (Scheme 3) occurs in the chloroplasts of green algae and plants, cyanobacteria, eubacteria and apicomplexan parasites. This alternative pathway starts with the condensation of pyruvate (**24**) and glyceraldehyde-3-phosphate(**25**)[50] to form 1-deoxy-*D*-xylulose-5-phosphate (**39**), which is catalysed by 1-deoxy-*D*-xylulose-5-phosphate synthase (DXS) making use of thiamine diphosphate (TDP) cofactor.[51] **39** is then isomerised to the corresponding aldehyde 2-*C*-methyl-*D*-erythrose-4-phosphate (**40**) by 1-deoxy-*D*-xylulose 5-phosphate reductoisomerase (IspC), which also catalyses the NADPH-dependant reduction of **40** to 2-*C*-methyl-*D*-erythritol-4-phosphate (**41**).[52] Subsequently, cytidine monophosphate (CMP) is added to **41** by IspD to give diphosphocytidyl-2-*C*-methyl-*D*-erythriol (**42**),[53] followed by ATP-dependant IspE catalysed phosphorylation to form the respective 2-phosphate (**43**).[54] This intermediate is converted to 2-C-methyl-*D*-erythriol-2,4-cyclodiphosphate (**44**) by IspF.[55] 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate (**45**) is afforded *via* a reductive ring opening of the cyclic diphosphate catalysed by IspG.[56,57] In the final step, IspH[58,59] catalyses the two stepwise single-electron transfers from an iron sulfur cluster and the cleavage of the hydroxyl moiety to give the allylic carbanion (**46**). Protonation of the carbanion gives **46** either IDP (**28**) or DMADP (**27**).

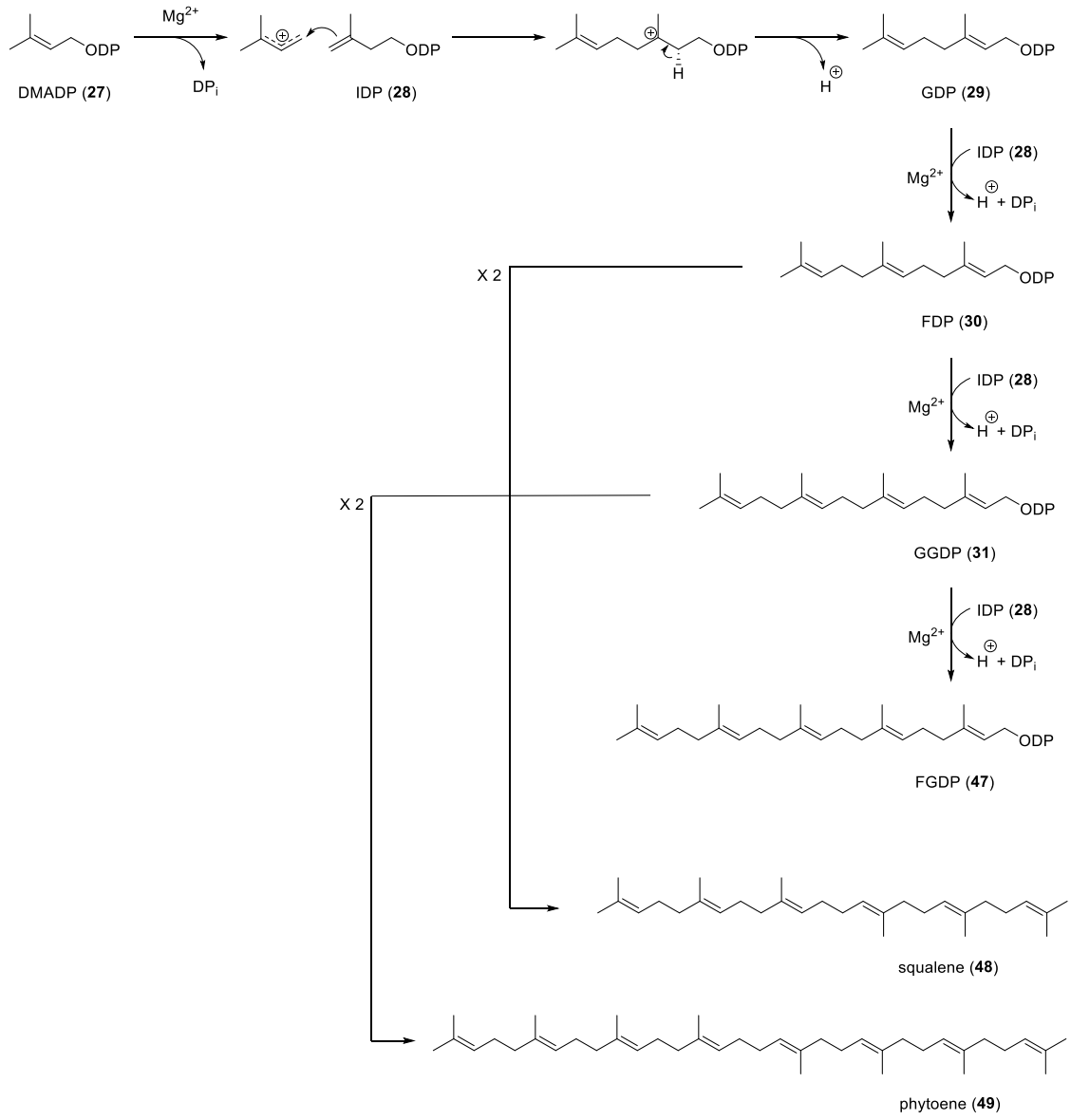


Scheme 3. Non-MVA pathway.

1.2.3. Isoprenyl diphosphate elongation

Prenyl transferases (also called polyisoprenyl diphosphate synthases) catalyse the chain elongation reactions of the two monomeric isoprene building blocks, DMADP (**27**) and IDP (**28**), to build larger linear polyisoprene units (Scheme 4). DMADP (**27**), which is synthesised *via* the non-mevalonate pathway or by isomerisation of IDP (**28**), is the starter unit used by geranyl diphosphate synthase (GDPS)[60,61] for the generation of geranyl diphosphate (GDP, C10, **29**) through nucleophilic addition of IDP (**28**) and proton abstraction. This reaction starts with a Mg^{2+}- dependent dissociation of the diphosphate group of DMADP (**27**) to form an allylic cation. This is followed by nucleophilic attack of the electron-rich double bond of IDP, and finally by stereoselective deprotonation at the C2 carbon of the newly formed isoprenoid to give GDP (**29**). In the same fashion, sequential additions of IDP catalysed by farnesyl diphosphate synthase (FDPS),[62,63] geranylgeranyl diphosphate synthase (GGDPS)[64,65] and farnesylgeranyl diphosphate synthase (FGDPS) give farnesyl diphosphate (FDP, C15, **30**), geranylgeranyl diphosphate (GGDP, C20, **31**) and farnesylgeranyl diphosphate (FGDP, C25, **47**), respectively. FDP (**30**) and GGDP (**31**) are also used as substrates for the generation of squalene

(C30, **48**) and carotenoids (e.g. phytoene, C40, **49**) respectively, and higher lengths can be achieved (e.g. natural rubbers).



Scheme 4. Isoprenyl diphosphate elongations.

The production of *cis*- isoprenyl diphosphates is also possible, this however requires changes of the orientation of IDP relative to DMADP. *Trans*- and *cis*- prenyltransferases contain distinct tertiary structures. Many studies have focused on the stereochemistry of these elongations and the structural determinants for the production of different chain- length products in these enzymes, which are not covered in this work.[66,67]

These linear polyisoprene units serve as the precursors for terpene synthases, which catalyse stereospecific reactions to produce mature terpenes. The terpenome is further expanded through sequential modifications in the carbon skeletons, for example P450-oxidised terpenoids.

1.3. Terpene synthases

Terpene synthases make the most remarkable contribution to the chemodiversity of terpenoids.[31,33] These enzymes catalyse the conversion of simple acyclic polyisoprene diphosphates into a multitude of cyclic or acyclic hydrocarbon or alcohol structures with high chemical precision, in what are considered some of the most complex reactions found in nature. The mechanism of action of these enzymes can be generally divided in three phases; 1) generation of a reactive carbocation, 2) guidance of the carbocation intermediates through a series of ring closures and/or rearrangements and 3) reaction termination, through proton loss or water capture.

As mentioned earlier, terpene synthases can be classified depending on the length of the linear diphosphate precursor they use. Monoterpene synthases use GDP (C10, **29**), sesquiterpene synthases use FDP (C15, **30**), diterpene synthases use GGDP (C20, **31**), sesterterpene synthases use GFDP (C25, **47**) as substrate, and so on.

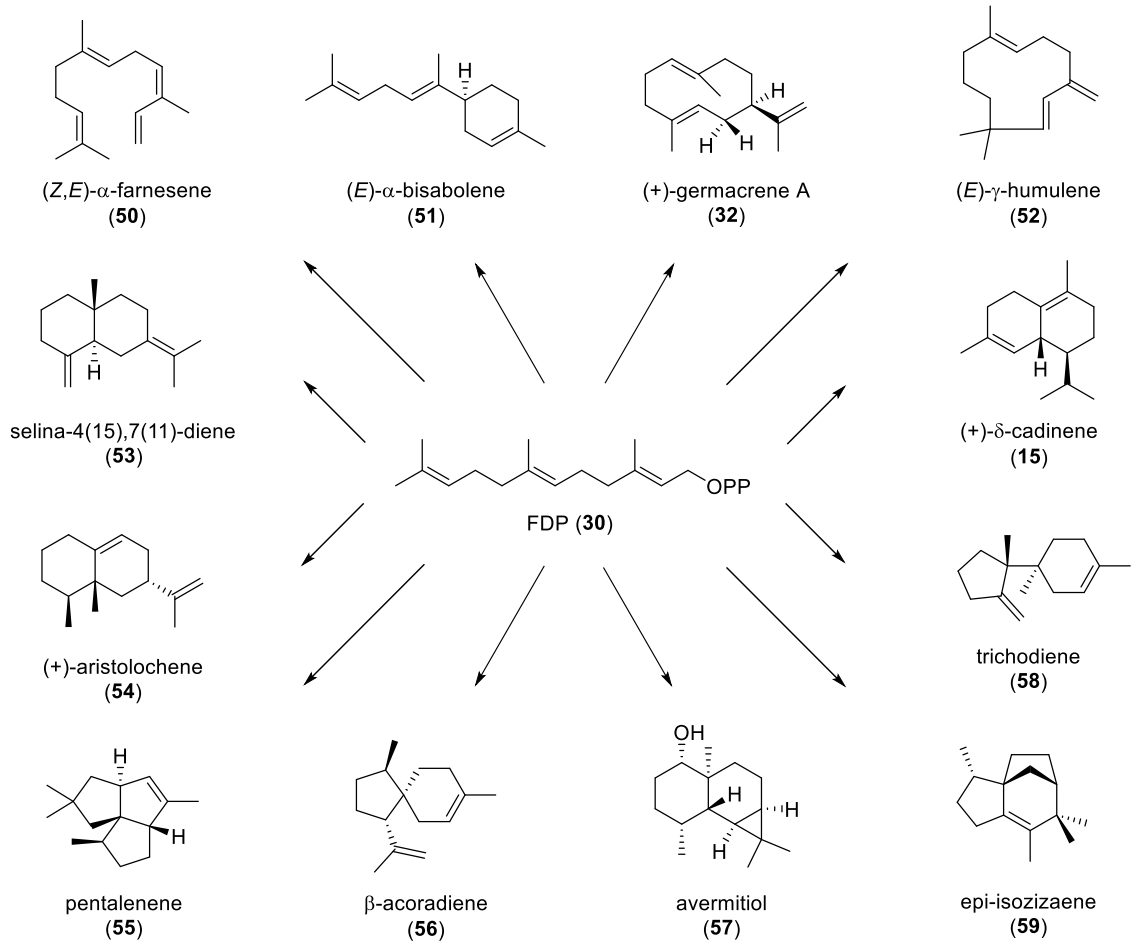
In addition, terpene synthases are divided in two classes depending on the nature of the initial ionisation and their tertiary structure. Class I terpene synthases use a trinuclear metal cluster to assist the cleavage of the diphosphate ester bond, whereas in class II terpene synthases, catalysis begins by protonation of the distal C-C double bond or an epoxide derivate thereof (their mechanism of action is not covered here).[33] The tertiary structure of terpene synthases is comprised of up to three domains, designated α, β and γ, each adopting an α helical barrel fold. In general, the domain architecture for class I terpene synthases is found as α, αβ, αβγ, wherein α is the active domain. As for class II terpene synthases, the domain architecture can be found as βγ or αβγ. In these two architectures, the catalytic region is located at the interface between domains βγ. There are also bifunctional terpene synthases, which can be composed of αβγ or αα architectures, and catalyse tandem reactions involving class I and class II catalysis (Table 1).

Carbon atoms	Terpene prefix	Linear precursors	Synthase class	Domain architecture
5	hemi-	IDP (**28**)	I	αβ
10	mono-	GDP (**29**)	I	αβ
15	sesqui-	FDP (**30**)	I	α, αα, αβ, αβγ
20	di-	GGDP (**31**)	I, II	α, αα, αβγ
25	sester-	GFDP (**47**)	I, II	α, αα
30	tri-	squalene (**48**)	II	βγ, αβγ

Table 1. General scheme of terpene nomenclature, linear precursors (DP = diphosphate), synthase classification, and commonly observed domain architectures. Adapted from reference [33].

1.4. Sesquiterpene synthases

Sesquiterpene synthases are class I terpene synthases that catalyse the conversion of FDP (**30**) to hundreds of different linear and cyclic sesquiterpene structures with high regio- and stereoselectivity (Scheme 5). These chemical reactions involve (on average) bonding, hybridisation or configuration adjustments in two thirds of the carbon atoms of FDP.[68]



Scheme 5. Representation of sesquiterpene natural product diversity.

The remainder of this section outlines structure-mechanism relationships in sesquiterpene synthases, which have been elucidated through the combinatorial use of X-ray crystallographic structures, mutagenesis studies and the use of substrate/intermediate analogues.

1.4.1. Structural features and their catalytic implications

In 1997, the elucidation of the X-ray crystal structures of tobacco 5-epi-aristolochene synthase from *Nicotiana tabacum*[69] (TEAS, Figure 4) and pentalenene synthase from *Streptomyces UC5319*[70] (PS) provided the first structural fundamentals for the understanding of sesquiterpene synthases mode of action. A few years later, the crystal structure of fungal aristolochene synthase from *Penicillium roqueforti*[71] (PRAS) was also obtained. These structures have served as the archetypes for understanding structure-function relationships in sesquiterpene synthases.

Sesquiterpenes synthases are widespread in plants, fungi and bacteria. Although they share very low amino acid sequence identity, they all contain a conserved tertiary structure consisting of α-helices in antiparallel fashion connected by short interconnecting loops and turns, called class I terpene synthase α-fold (Figures 4 and 5).

This α-fold is also shared by avian farnesyl diphosphate synthase[62] and human squalene synthase,[72] which indicates divergence from a common ancestor early in the evolution of terpene biosynthesis. Bacterial and fungal sesquiterpene synthases, such as PS and PRAS, are usually functional as single α-domain monomers, although homodimeric enzymes can be also found, such as trichodiene synthase from *Fusarium sporotrichioides*.[73] Plant sesquiterpene synthases such as TEAS may contain an additional N-terminal domain, called the β-domain (Figures 4 and 5). The catalytic contribution of the β-domain in plants is not fully understood yet, although it has been proposed to ensure the correct folding of the C-terminal α-domain (active domain) and assist in desolvation events.[74] In addition, plant sesquiterpene synthases can also have αβγ domain architectures , in which the catalytic contribution of the third domain is unknown (e.g. bisabolene synthase[75]), Figure 5.

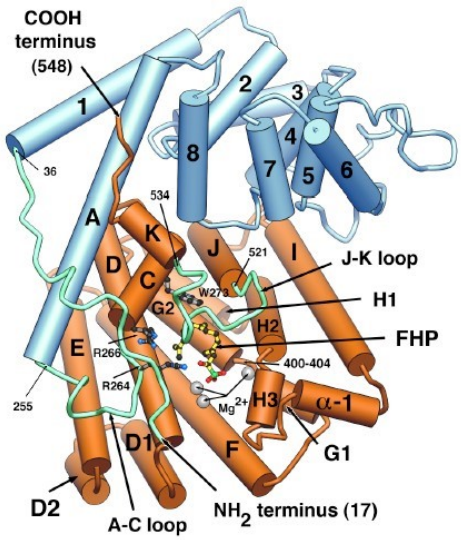


Figure 4. Cartoon representation of the X-ray crystal structure of TEAS in complex with farnesyl hydroxyphosphonate (FHP). Picture taken from reference [69]. In blue, the β-domain. In orange, the α- domain. Helices and the α- domain loops are labelled, including A-C, F-G1, J-K. Helix H is disrupted between segments H1 and H2 by P_{430}. In helix G there is a 'kink' between residues 400 and 404 that forms G1 and G2 segments.

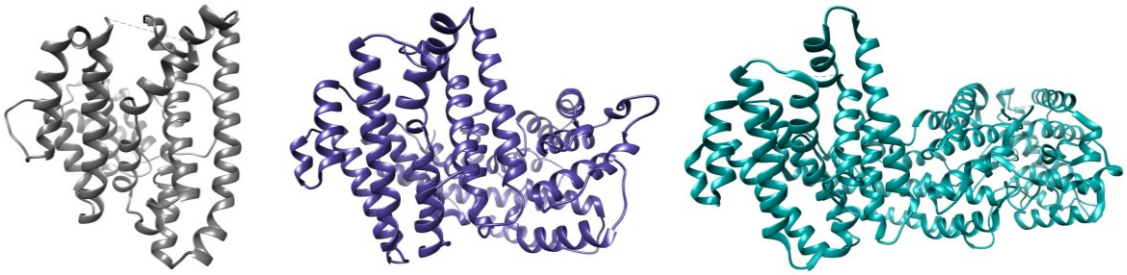


Figure 5. From left to right, ribbon diagram representations of the X-ray crystal structures of pentalenene synthase [pdb:1PS1][70], 5-epi-aristolochene synthase [pdb:5EAS][69] and bisabolene synthase [pdb:3SDQ][75].

The active site of sesquiterpene synthases contains two highly conserved metal-binding motifs on helices D and H, which are the 'DDXXD/E' motif (aspartate-rich motif) and the 'NSE/DTE' motif. They face one another across the active site entrance and guide the substrate through Mg^{2+}-PP_i-enzyme cluster formation, generating the Michaelis complex needed for catalysis (Figure 6).[76,77] Although these two regions are very much conserved, some exceptions have been observed. For instance, (+)-δ-cadinene synthase from *Gossypium arboreum*[78] possesses a second aspartate rich motif (DDVAE) in place of the usual 'NSE/DTE' motif and (+)-germacrene D synthase from *Solidago canadensis*[79] possesses an aspartate rich motif comprising NDTYD.

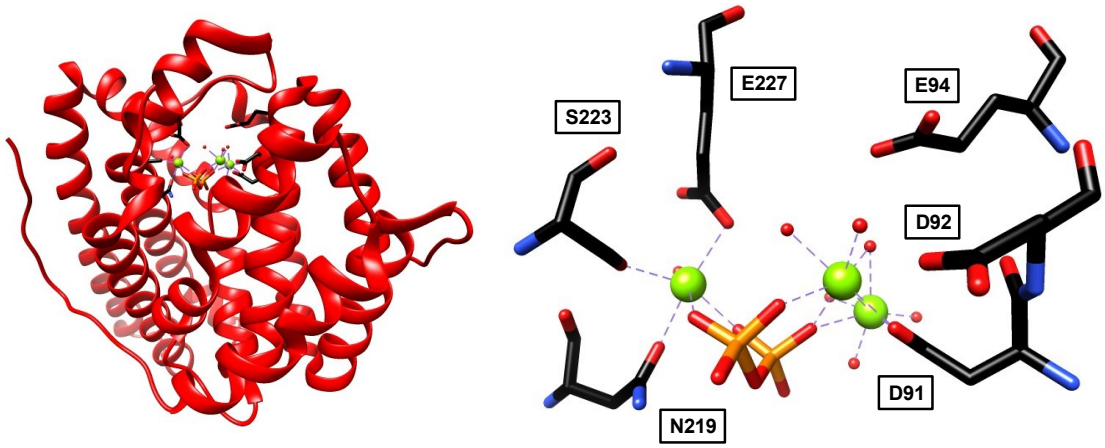
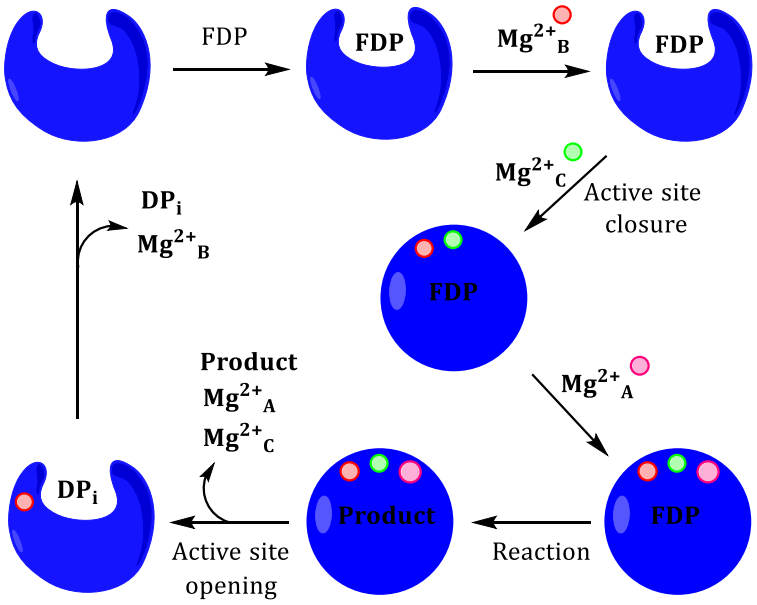


Figure 6. Left, ribbon diagram representation of the X-ray crystal structure of ATAS [pdb: 2OA6][80]. Right, cartoon representation of the X-ray crystal structure of ATAS including the conserved motifs interacting with PP_i and Mg^{2+} ions.

The substrate binding cycle in sesquiterpene synthases was studied in detail with aristolochene synthase from *Aspergillus terreus* (ATAS). A series of X-ray co-crystal structures of ATAS cocrystallised with various combinations of Mg^{2+}, PP_i and FDP (**30**) revealed for the first time the Mg^{2+}-PP_i-enzyme cluster in its closed conformation (Figure 6).[80] Later crystallographic studies making use of fluorinated FDP analogues revealed different binding interactions, which led Shishova *et al.* to propose a model for substrate and metal binding.[81]

Initially, the substrate is recognised by hydrogen bonding interactions with the diphosphate tail (Scheme 6); presumably, then the binding of Mg^{2+}_B takes place, causing a reorientation of the diphosphate group. Following this, Mg^{2+}_C binds to the first aspartate in the DDxxD motif and to the diphosphate group triggering active site closure. Binding of Mg^{2+}_A completes the magnesium cluster, displacing Mg^{2+}_C and provoking a coordination of Mg^{2+}_A and Mg^{2+}_C to the first aspartate in the DDxxD motif and completing the conformational changes for the Michaelis-complex to be formed. Opening of the active site and product release is proposed to occur in a reverse way. Van der Kamp *et al.*[82] gave further evidence to this proposal using molecular dynamics simulations on ATAS. In this study, they observed that the apo-enzyme adopts different conformations from open to near-closed, and that the binding of FDP and Mg^{2+}_B shift the equilibrium to 'distinctly open' conformations, which is followed by the binding of the rest of the magnesium ions.



Scheme 6. Representation of the proposed catalytic cycle of sesquiterpene synthases.

This event, along with structural changes in the H-α1, J-K, F-G1 loops and the G1/2 helix break motifs (Figures 4 and 7) trigger active site closure and hence substrate ionization, which is then sheltered from premature solvent quench. These conformational changes vary between terpene synthases, which again is indicative of divergent evolution from a common template. [73,74,83,84]

Dickchat *et al.*[84,85] have identified significant structural shifts in two bacterial terpenoid synthases, selina-4(15),7(11)-diene synthase (SdS) and hedycaryol synthase (HcS). X-ray crystal structures of SdS in the apo-state, in coordination with $(Mg^{2+})_3$-PP_i and with $(Mg^{2+})_3$-DHFDP (2,3-dihydro-FDP, inhibitor) were used to identify movements of the G1/2-helix break motif between the open and closed conformations (Figure 7).[85] The data shows a shift of arginine-178 towards the diphosphate ligand upon substrate binding. Concurrently, aspartate-181 is orientated to form a hydrogen bond with

arginine-178. This shift leads to a conformational change of the G1/2-helix break motif, assisting the carbonyl oxygen of glycine-182 to point towards C3 of DHFDP and facilitating PP_i departure.[85] Similar results were observed for HcS in complex with nerolidol, in which the carbonyl of V179 points towards C1 of nerolidol.[84] The structural importance of the G1/2 helix break in sesquiterpene catalytic fidelity will be further discussed in Section 2.2.

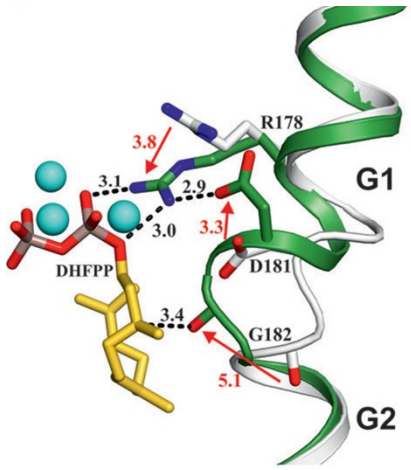


Figure 7. Cartoon representation. Superposition of the G1/2 of SdS helix in the open (grey) and closed (bounded with DHFDP, green) conformations. Picture taken from reference [85]. Interactions are shown as dashed lines, selected distances are labelled, and red arrows show the arrangement of amino acids from open to closed conformations.[85]

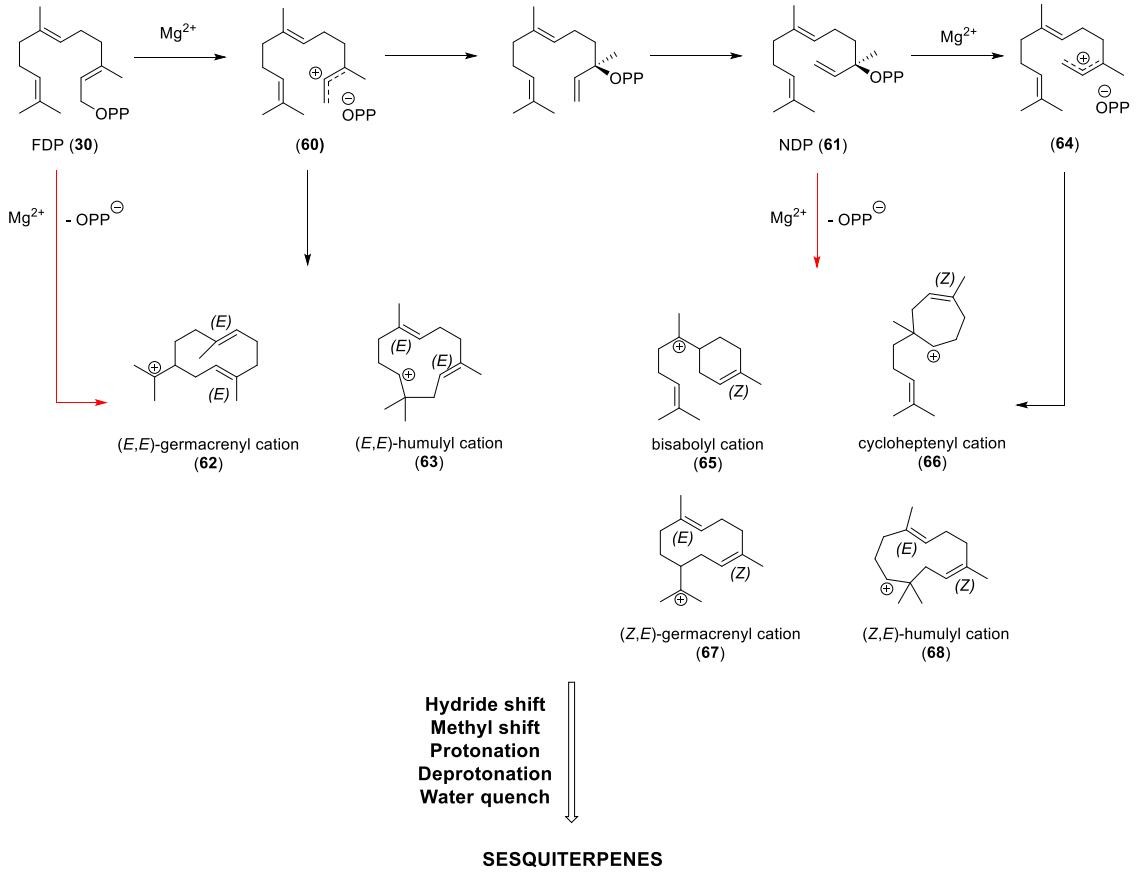
Once the template for FDP (**30**) cyclization is fully formed in the closed active site conformation, the hydrophobic active site composition allows carbocation stabilisation through cation-π interactions *via* aromatic side chains guiding the flexible substrate and intermediates during catalysis.[33,86,87] Aliphatic residues are also involved in forming the active site template, creating an ensemble governing substrate and intermediate conformations.[88]

1.4.2. Reaction mechanisms

The biosynthetic cascade orchestrated by sesquiterpene synthases is initiated by metal dependent cleavage of the diphosphate anion of FDP (**30**) to form a transoid farnesyl diphosphate carbocation (**60**). The next catalytic events rely on the ability of some synthases to generate the cisoid counterpart of FDP, nerolidyl diphosphate (NDP, **61**) (Scheme 7).

The 'transoid' synthases, ionise FDP(**30**) to generate a 2,3-transoid farnesyl cation (**60**), which can only react with the distal C10–C11 double bond *via* 1,10- and 1,11- intramolecular cyclisation to generate (*E*,*E*)-germacrene-11-yl (**62**) and (*E*,*E*)-humul-10-yl cation (**63**), respectively. By contrast, the 'cisoid' synthases catalyse first the *trans-cis* isomerisation of (2*E*,6*E*)-FDP (**30**) to form (3*R*)-nerolidyl diphosphate (NDP, **61**). Rotation of the C2-C3 bond and reionisation of NDP give the 2,3-cisoid farnesyl cation (**64**). This event provides other catalytic opportunities such as, a 1,6-cyclisation to the bisabolyl cation (**65**), a 1,7-cyclisation to the cycloheptenyl cation (**66**), a 1,10-cyclisation to the (*E*,*Z*)-germacrenyl cation (**67**) or a 1,11-cyclisation to the (*E*,*Z*)-humulyl cation (**68**).[89] After cyclisation, a further cationic cascade occurs, and the reaction is finished by either water quench or deprotonation. Few sesquiterpenes are acyclic and are quenched at early stages of the mechanism.

Note that ionisation and cyclisation events can occur *via* stepwise or concerted mechanism.



Scheme 7. Initial steps in the reaction mechanisms leading to the formation of sesquiterpenes.

1.5. Methods for the study of sesquiterpene synthases

The art of forming the terpenome relies upon a perfect communion between enzyme and substrates. Terpene synthases are specialists in the guidance of carbocation intermediates to generate a myriad of cyclic and acyclic products. To make this happen, the active site machinery should contain the right chemical tools. Substrate recognition and C-OPP_i cleavage precede a cascade of C-C, C-H and C-O interactions, in which aromatic and aliphatic amino acids serve as the enzymatic template for catalysis. Many of the terpene synthases focus all the catalytic requirements onto the formation of a single product, while others can generate multiple products. For example, γ-humulene synthase from *A. grandis*[90] generates 52 different metabolites. Multiproduct formation could be attributed to an incomplete evolution of a concrete template, however, the presence of highly stereospecific products in promiscuous terpene synthases, for instance in MtTPS5 from *Medicago truncula*[91] shows a pivotal evolved role of the active site to tightly control the substrates and reactive intermediates in multiproduct terpene synthase, thus maximising the diversification of products.[92]

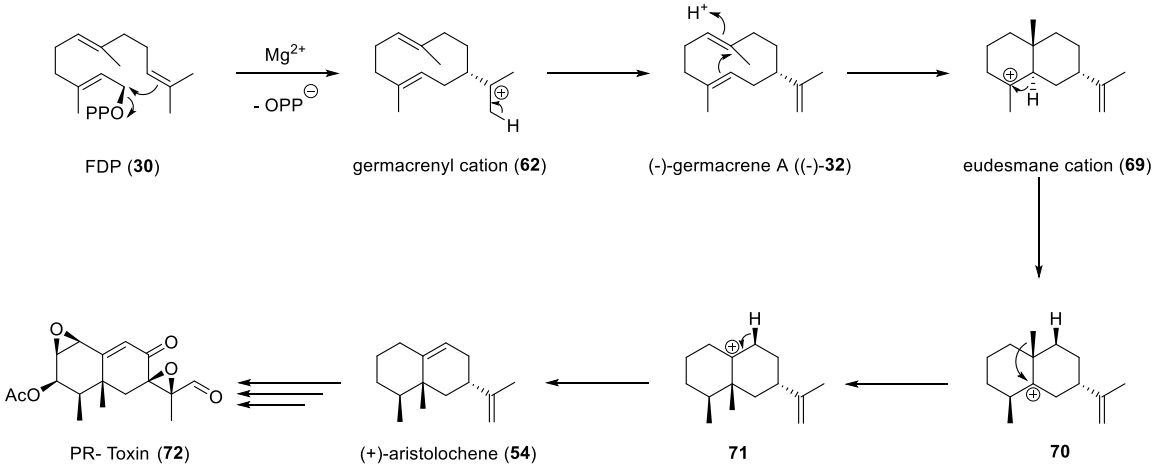
The enzymatic mechanism of terpene synthases is hard to decipher. Experimental methods have been used to elucidate the structural and mechanistic details of specific enzymes. First and foremost, site-directed mutagenesis studies (SDM) are used to probe the contribution of selected amino acids to the catalytic activity and product profile. Alternatively, the course of reactions can be followed using; 1) aza-analogues that mimic the carbocation intermediates, 2) fluorinated substrates that stabilise/ destabilise putative carbocations and interrupt the reaction cascade; or 3) substrate isotopologues as a probe to validate the nature and stereochemistry of the hydride shifts. In addition, substrate isomers and further alterations in the FDP chain provide (if accepted) access to new products using the same template architecture. These products can also give away mechanistic intricacies. The exchange of metal cofactors and pH assays can also give access to alternative mechanistic pathways.[92]

These approaches are covered in the remainder of this section, including their applicability on specific enzymes.

1.5.1. Site-directed mutagenesis

Site-directed mutagenesis is one of the most widely used tools for the investigation of protein structure and function. In some cases, alteration of selected amino acid side chains leads to the production of slower or non-functional enzymes. However, in many cases these mutations can lead to changes in substrate specificity and/or product alterations. Overall, through the use of SDM, it is possible to address fundamental questions regarding molecular recognition and catalytic control in terpene synthases and it offers the possibility to access new products by rational engineering. SDM has been used to investigate the metal-binding motifs, active-site aromatic and aliphatic residues, as well as putative active-site acid/bases in many sesquiterpene synthases.[68]

Aristolochene synthase from *P. roqueforti* (PRAS)[71] is a monomeric sesquiterpene synthase that catalyses the metal-dependent formation of (+)-aristolochene (**54**), which is the precursor of several fungal toxins such as PR-toxin (**72**). The active site of PRAS has been extensively targeted for SDM, providing relevant structural and mechanistic information about its catalytic mechanism (Scheme 8).



Scheme 8. Proposed reaction mechanism of the conversion of FDP (**30**) to (+)-aristolochene (**54**) catalysed by PRAS, and its transformation to PR-Toxin (**72**).

The role of the two conserved metal-binding motifs in the recognition and ionisation of the diphosphate motif of the FDP by PRAS has been addressed.[93] The various generated mutants in the aspartate-rich region (**D**^{115}**D**VL**E**) and the second motif (**N**^{244}DIY**S**YDK**E**) led to increasing amounts of the intermediate (-)-germacrene A ((-)-**32**), reaching 100% in the E252Q mutant, and was generally accompany with decreased catalytic efficiency. These results, which are comparable to other sesquiterpene synthases,[76] were indicative of perturbations in the metal cluster coordination, thus leading to conformational changes of the substrate and intermediates (Table 2).

PRAS-	k_{cat} (10^{-3} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})	54 (%)	32 (%)	Other (%)
WT	43 ± 2	0.6 ± 0.1	7.2×10^4	94	4	2
D115E	16.2 ± 0.4	2.7 ± 0.2	6.0×10^3	75	19	6
D115N			inactive			
D116E	1.30 ± 0.05	1.10 ± 0.13	1.2×10^3	62	35	3
D116N	2.2 ± 0.1	1.5 ± 0.1	1.5×10^3	63	35	2
E119D	9.7 ± 0.2	0.32 ± 0.02	3.0×10^4	94	4	2
E119Q	4.8 ± 0.1	0.15 ± 0.02	3.2×10^4	84	14	2
N244L			inactive			
N244D	0.120 ± 0.003	3.3 ± 0.2	36.4	19	81	0
S248A	1.30 ± 0.04	5.5 ± 0.4	2.4×10^2	21	79	0
E252D	2.20 ± 0.09	3.4 ± 0.4	6.5×10^2	19	81	0
E252Q	0.22 ± 0.01	10.1 ± 0.7	21.8	0	100	0

Table 2. Catalytic parameters and product distribution of PRAS and binding motifs related mutants.[93]

In addition, the role of the two arginine residues R200 and R340 (which are conserved among sesquiterpene synthases) in the stabilization of the diphosphate leaving group during catalysis has been probed with SDM.[94] Replacements of R200 and R340 by amino acids with decreasing hydrogen bonding capabilities (K>Q>E>M) resulted in a progressive deactivation of the enzyme and accumulation of (-)-germacrene A ((-)-**32**), which is consistent with the proposal that PP_i acts as the active site acid that protonates (-)-germacrene A ((-)-**32**)during catalysis in PRAS (Table 3).

PRAS-	k_{cat} (10^{-3} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})	54 (%)	32 (%)	Other (%)
WT	84	0.53	15.8×10^4	91.5	7.5	<1
R200K	0.6	1.04	620	19.2	78.9	1.8
R200Q			inactive			
R200E			inactive			
R340K	0.8	1.59	510	54.4	42.8	2.8
R340M			inactive			

Table 3. Catalytic parameters and product distribution of PRAS and binding motifs related mutants. [93]

The contribution of Y92 and W334 active site aromatic residues in the enzymatic conversion catalysed by PRAS has also been object of SDM studies. Y92 of PRAS is postulated to sterically assist the correct

folding of FDP (**30**) for 1,10 cyclisation. Acyclic products (mainly (*E*)-β-farnesene, **3**) are observed upon reduction of the size of the side chain of Y92 by mutation to valine, cysteine and alanine.[95–97] W334 facilitates the generation of the eudesmane cation (**69**) through stabilising cation-π interactions. The substitution of W334 with valine and leucine resulted in the accumulation of (-)-germacrene A ((-)-**32**), while the production of W334F had less impactful effects on the product distribution (Table 4).[98,99]

PRAS-	k_{cat} (10^{-3} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})	54 (%)	32 (%)	Acyclic (%)	Other (%)
WT	30 ± 10	2.3 ± 0.5	(1.3 ± 0.3) × 10^4	91.5	7.5	0	<1
Y92V	0.25	4.86	51.65	28.2	11.2	42.6	18
Y92C	0.49 ± 0.13	50.27 ± 9.11	10.2 ± 4.3	6.8	0	61.3	31.9
Y92A	1.37 ± 0.14	83.40 ± 10.53	16.5 ± 2.4	0	0	72.6	27.4
W334F	2.65 ± 0.7	66.8 ± 5.8	51.65	85.5	9.4	0	5.1
W334V	0.09 ± 0.01	33.6 ± 5.2	10.2 ± 4.3	4.7	95.3	0	0
W334L	0.023 ± 0.002	74.8 ± 17	16.5 ± 2.4	0	100	0	0

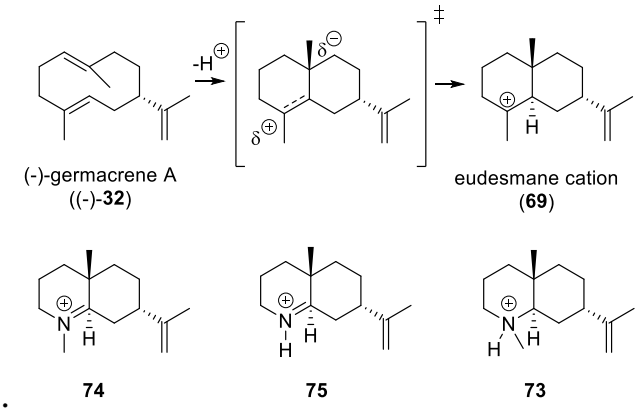
Table 4. Catalytic parameters and product distribution of PR-AS and Y92 and W334 mutants.[95–99]

Allemann *et al.*[88] further demonstrated the contribution of the active site template provided by PRAS for the reaction outcome with mutations to the aliphatic amino acids L108, V88 and T89.

1.5.2. Aza-analogues.

Mechanistic studies of reactions catalysed by terpene synthases prove difficult due to the carbocationic nature of most reaction intermediates. These are quite transient due to rapid deprotonation or rearrangement steps, which does not allow their isolation or observation. Aza- analogues are nitrogen-containing analogues that mimic the stereochemical and electrostatic properties of the putative carbocation intermediates. The sp^2 hybridised carbocation is substituted with a sp^3 hybridised nitrogen or sp^2 hybridised iminium ion. These aza-analogues cannot be consumed by the enzyme and have been used in kinetic studies, often acting as competitive inhibitors.[100,101] In addition, they can be used for co-crystallization in order to gain protein- substrate mimic structural details of the enzyme at different stages of the reaction (snapshots of the different stages of the mechanism).[83,102,103]

For example, the aza-analogue **73** has been shown to act as a potent competitive inhibitor of PRAS, thus giving further evidence for the eudesmane cation (**69**) formation.[104] In addition, the aza-analogues **74** and **75**, which were synthesised to mimic the possible transition state between (-)-germacrene A (**32**) and the eudesmane cation (**69**), act as PP_i-dependent inhibitor, supporting a possible role for PP_i as the active site acid in charge of carbocation formation (Scheme 9).[105]

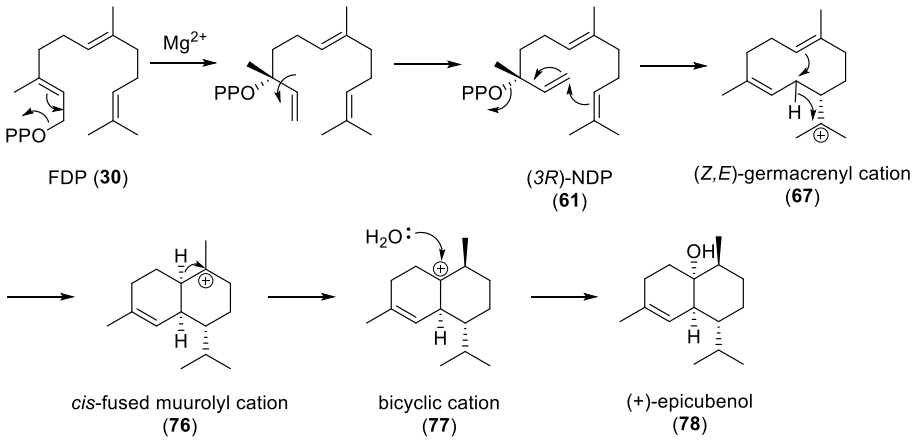


Scheme 9. Aza-analogues and the carbocations they mimic.

1.5.3 Substrate isotopologues

Substrate isotopologues are homologous to the natural substrate in which one or more atoms have been substituted with atoms of different mass number. These has been extensively used for uncovering stereo- and regiochemical details of the reaction catalysed by terpene synthases, including those for the production of (+)-aristolochene (**54**),[106] pentalenene (**55**),[107–109] (+)-δ-cadinene (**3**),[110] (-)-germacrene D (**91**, Scheme 14),[111] trichodiene (**58**),[112] and amorpha-4,11-diene (**104**, Scheme 16),[113,114] amongst others.

The enzymatic conversion of FDP (**30**) to (+)-epicubenol (**78**) by (+)-epicubenol synthase (EpicS) has been analysed using deuterated and tritiated substrates.[115–117] Cane *et al.* demonstrated that FDP (**30**) is isomerised to (3*R*)-NDP (**61**), which is then converted to the intermediate (*Z,E*)-germacrenyl cation **67** and reacts to the cis-fused muurolyl cation **76** by a 1,3-hydride shift and a ring closure. Following this, a 1,2-hydride shift to generate the bicyclic cation **77** and water capture gives (+)-epicubenol (**78**) (Scheme 10).

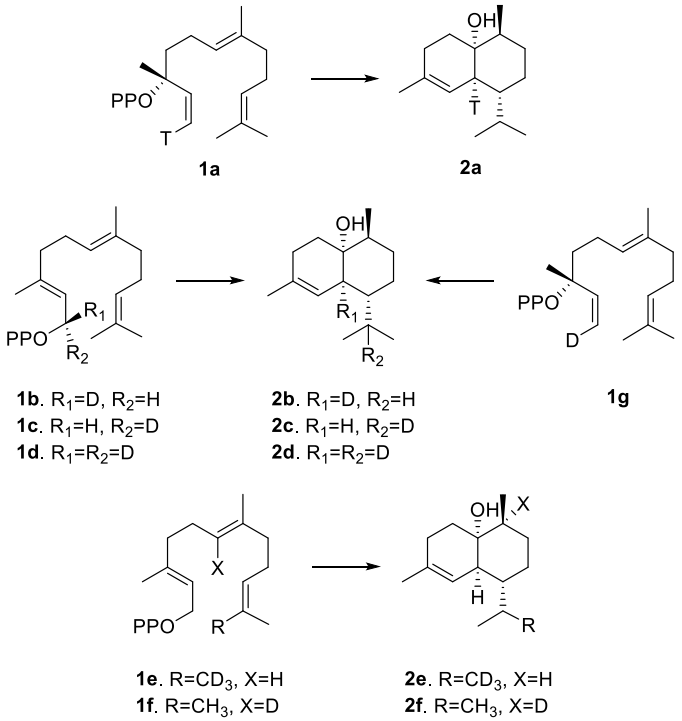


Scheme 10. Proposed reaction mechanism of the conversion of FDP (**30**) to (+)-epicubenol (**78**), catalysed by EpicS.

The specificity for (3*R*)-NDP (**61**) was examined through incubations with stereospecifically tritiated NDP (Scheme 11). While (3*R*)-1*Z*-[1-^{3}H]NDP (**1a**) was turned over to the corresponding tritiated epicubenol (**2a**), (3*S*)-1*Z*-[1-^{3}H]NDP was not accepted as a substrate.[117] Furthermore, incubations with both 1*R*-[1-^{2}H]FDP (**1b**) and (3*R*)-(1*Z*)-[1-^{2}H]NDP (**1g**) gave [5-^{2}H]-epicubenol (**2b**), which demonstrated that the conversion of FDP (**30**) to (3*R*)-NDP (**61**) proceeds with suprafacial stereochemistry. The 1,3-hydride shift was demonstrated using [1,1-^{2}H]-FDP (**1d**) and the stereochemistry elucidated by the incubations of 1*R*-[1-^{2}H]FDP (**1b**) and 1*S*-[1-^{2}H]FDP (**1c**) with EpicS.[115,117] Finally, the 1,2-hydride shift was further demonstrated with the use of [6-^{2}H]-FDP (**1f**),

which gave the deuterated epicubenol analogue **2f** after incubation with EpicS, as judged by NMR spectroscopy analysis (Scheme 11).[115]

Independent studies by Nabeta *et al.* using EpicS from a different organism confirmed that this mechanism is shared by these two enzymes.[118,119]



Scheme 11. FDP analogues used for probing the catalytic mechanism of EpicS and the resulting enzyme generated products.

1.5.4 Fluorinated analogues

The use of fluorinated analogues represents another approach to mechanistic investigations. Fluorine (r_F= 1.47 Å) is larger than hydrogen (r_H= 1.20 Å), but similar enough to not produce significant undesirable steric effects or affect binding affinity. Due to the high electronegativity, fluorine exerts strong electronic influence at the α, β and γ-positions of putative carbocations, acting both as a σ-withdrawing and π-donating substituent. Carbocations at the α-position are stabilised by π-donation of a lone pair of electrons into the empty p-orbital, while carbocations at positions β and γ are destabilised through inductive effects (Figure 8).

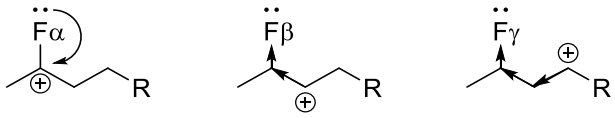
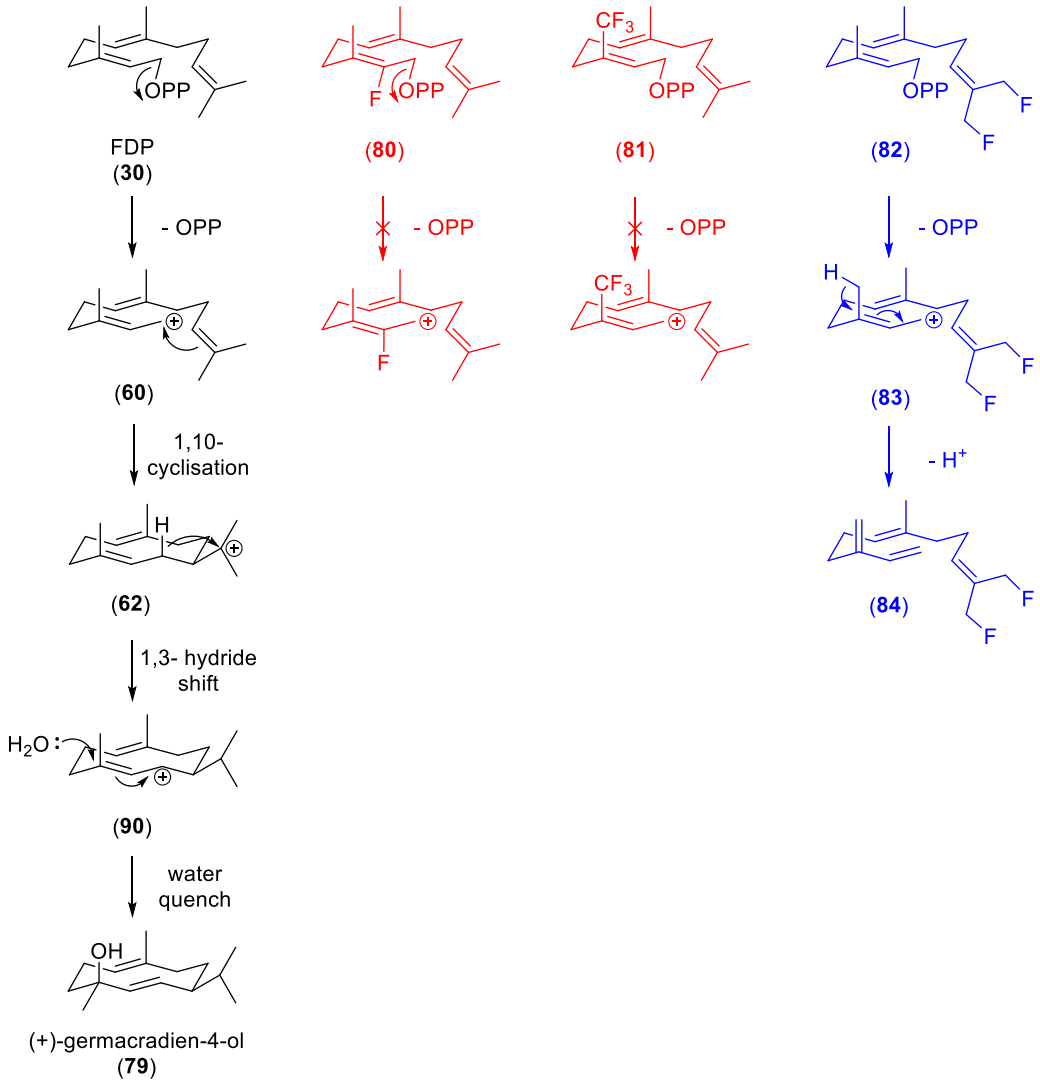


Figure 8. Fluorine electronic influence.

As a result, the strategic positioning of fluorine in the substrate analogues can be used for probing the existence of carbocation intermediates and/or guiding the reaction through alternative pathways.[89,120–124] Fluorinated analogues are also useful for X-ray crystal structure elucidations because they are often found to be potent competitive inhibitors.[81]

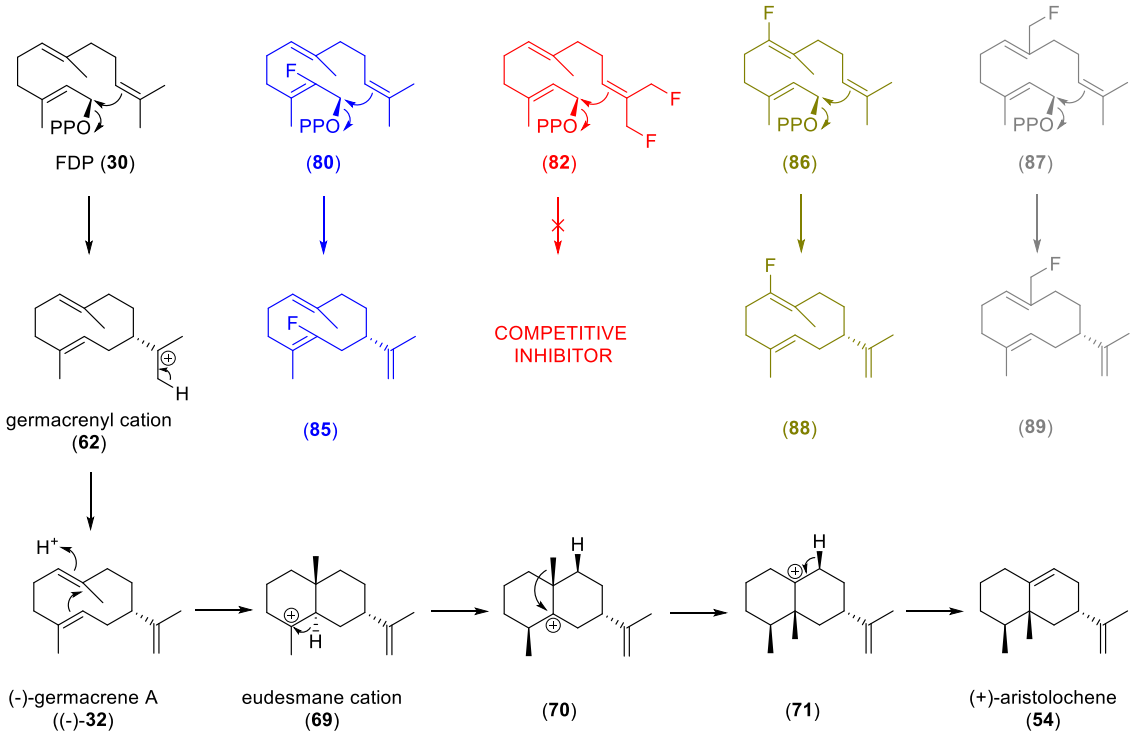
The mechanism of (-)-germacradien-4-ol synthase from *Streptomyces citricolor* (GdolS), in which FDP (**30**) is converted to (-)-germacradien-4-ol (**79**), has been interrogated using fluorinated analogues, namely 2F-FDP (**80**), 15,15,15-F_3-FDP(**81**) and 12,13-F_2-FDP(**82**) (Scheme 12).[125]

Incubations with 2F-FDP (**80**) and 15,15,15-F_3-FDP (**81**) generated no extractable products, and steady-state kinetic studies revealed that they are potent non-competitive inhibitors. These results might suggest that initial farnesyl carbocation formation and 1,10-cyclisation proceeds in a stepwise manner. Due to the position of the fluorine atoms at C2 and C15, the electron-withdrawing effect increases the energy of the emerging carbocation and thus formation is prevented. To further probe this result, Gd4olS was incubated with 12,13-F_2-FDP (**82**). The position of the fluorine atoms should not prevent the formation of farnesyl cation **83**, but instead prevent the following the 1,10-cyclisation because the π-electron density of C10-C11 double bound is reduced and the energy of the arising germacradienyl cation increased. Hence, 12,13-difluoro-(*E*)-β-farnesene (**84**) was the only product detected.



Scheme 12. In black, reaction catalysed by GdolS. In blue, conversion of 12,13-F_2-FDP to 12, 13-F_2-farnesese by GdolS. In red, abortive reactions using 2F-FDP and 15,15,15-F_3-FDP.

In contrast to GdolS, 2F-FDP (**80**) is a substrate for PRAS, generating the corresponding 2F-(-)-germacrene A (**85**) analogue and probing a concerted mechanism to access to the germacrenyl cation **62**. Consequently, 12,13-F_2-FDP (**82**) is an inhibitor for PRAS. Furthermore, (2*E*, 6*Z*)-6F-FDP (**86**) and 14F-FDP (**87**) are also substrates for PRAS and the related (-)-germacrene A analogues **88** and **89** were respectively obtained upon enzymatic incubation, which arise from preventing eudesmane cation formation and further evidence the existence of the (-)-germacrene A (**32**) as an neutral intermediate (Scheme 13).



Scheme 13. Reaction catalysed by PRAS (black) and products after incubation with analogues (blue (**80**), red (**82**), green (**86**) and gray (**87**)).

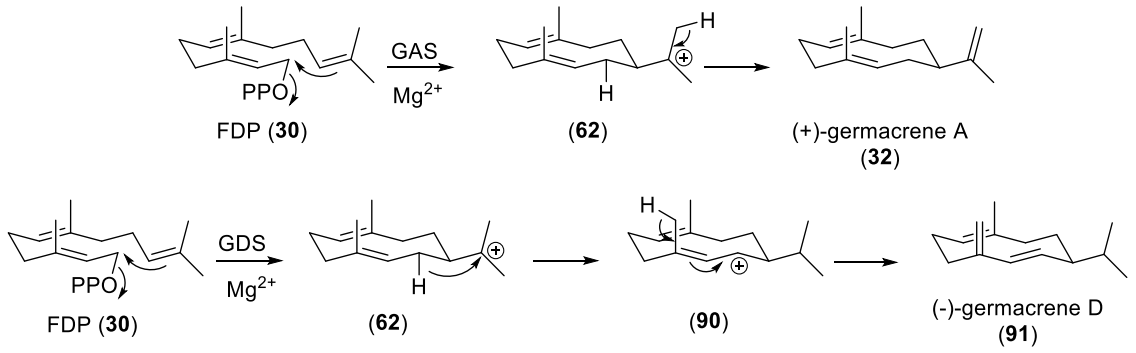
As exemplified with fluorinated analogues, the incubation of altered substrates with terpene synthases can be used for the generation of novel natural products, in which terpene synthases serve as the catalytic template for giving access to a new pool of chemicals with medicinal or agrochemical purposes.

1.6. Substrate promiscuity

The promiscuity of sesquiterpene synthases to accept novel analogues of (*E*,*E*)-FDP (**30**) has been reported for a range of them (See also Section 1.5.4). The synthesis of FDP analogues and their incubation with the synthases not only serves as mechanistic probe, but also provides an interesting area of research towards novel terpenoids.

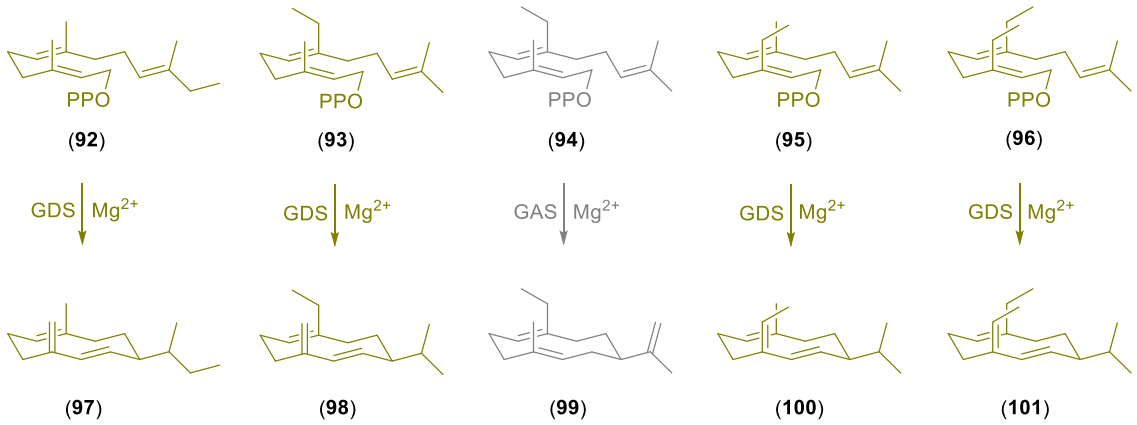
1.6.1. Methylated analogues

(+)-Germacrene A synthase (GAS) and (-)-germacrene D synthase (GDS) from *Solidago canadensis*[111,126,127] are two plant sesquiterpene synthases in charge of (+)-germacrene A (**32**) and (-)- germacrene D (**91**) formation respectively (Scheme 14). These sesquiterpenes act as semiochemicals, effecting an olfactory response in some insects, and are industrially used in perfumes, cosmetics, food and beverages. The practical application of these compounds in crop protection is limited because their rather chemical instability.[127–129] Nevertheless, design of analogues with improved activity and stability has been reported.[123,130]



Scheme 14. Top, reaction catalysed by GAS for the formation of (+)-germacrene A (**32**). Bottom, reaction catalysed by GDS for the formation of (-)-germacrene D (**91**).

In view of the biological and economic significance of these compounds, Allemann *et al.* investigated the capacity of GAS and GDS to produce non-natural germacrenes from methylated FDPs (**92**-**96**) (Scheme 15).[123,130] The enzymatic products were characterised by GC-MS and NMR spectroscopy and the biological activities of GAS and GDS towards the non-natural substrates were assessed. In Scheme 15 the products (**97**-**101**) obtained upon incubation of methylated analogues (**92**-**96**) with GDS and GAS are shown. Interestingly, 14,15-dimethyl-germacrene D (**101**) is a potent attractant of grain aphids and a potential crop protection agent.

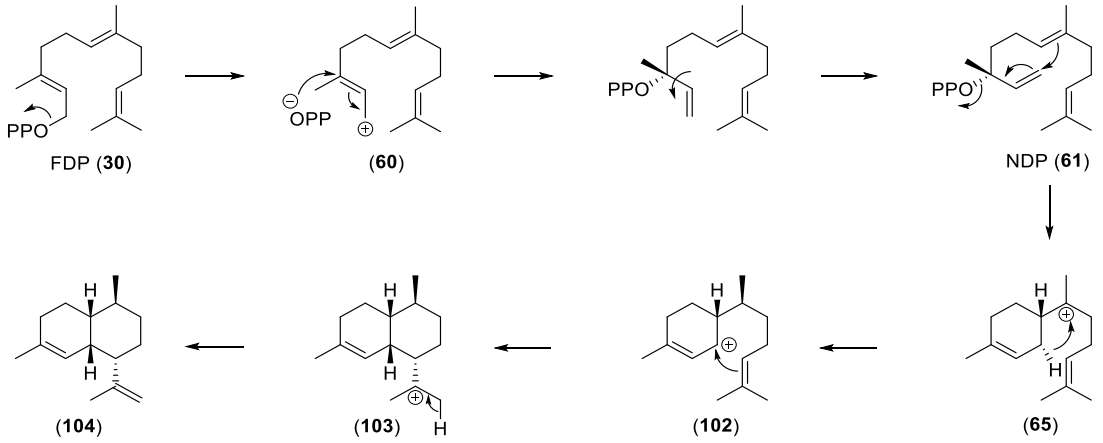


Scheme 15. Methylated analogues and the enzymatic products obtained after incubation with GDS and GAS.

1.6.2. Oxygenated analogues

The enzymatic conversion of FDP analogues containing nucleophilic functional groups by sesquiterpene synthases has also been investigated.[131,132]

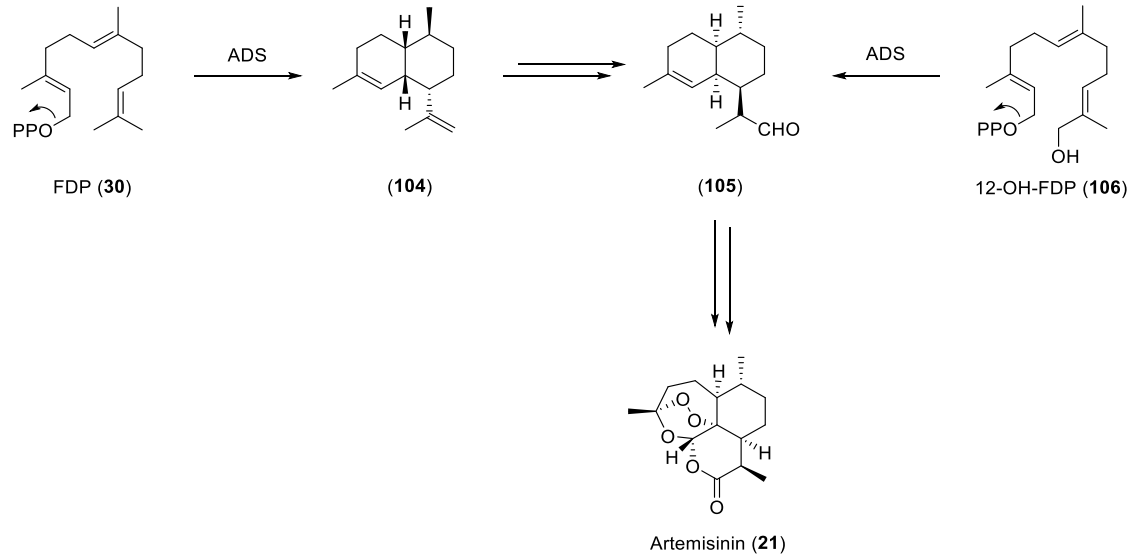
Amorphadiene synthase from *Artemisia annua* (ADS) catalyses the conversion of FDP (**30**) to amorpha-4,11-diene (**104**, Scheme 16)[133–135], which is the biological precursor of artemisinin (**21**, Scheme 17), a potent antimalarial drug.



Scheme 16. Reaction catalysed by ADS for the generation of amorpha-4,11-diene.

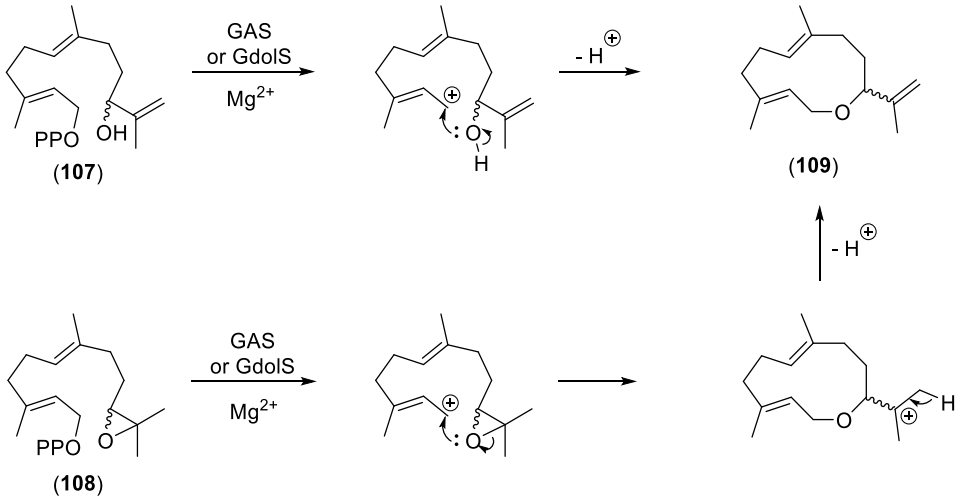
Several efforts have been directed to the production of artemisinin in a cost-effective manner. Currently, the most efficient way relies on the combination of biosynthesis and synthetic chemical steps. In general, ADS produces amorpha-4,11-diene (**104**), which can be converted to dihydroartemisinic aldehyde (DHAAl, **105**), an *en-route* intermediate during artemisinin production. DHAAl (**105**) can then be converted to artemisinin (**101**) following 4 chemical steps.[136,137] Recently, however, Allemann *et al.*[131] developed the synthesis of 12-OH-FDP (**106**) and exploited the substrate

promiscuity of ADS. 12-OH-FDP (**106**) is converted by ADS to the intermediate DHAAl (**105**). This approach shortens the synthesis towards artemisinin in that it only uses one enzyme and requires a single oxidation step (Scheme 17).



Scheme 17. Biosynthetic routes for artemisinin production.

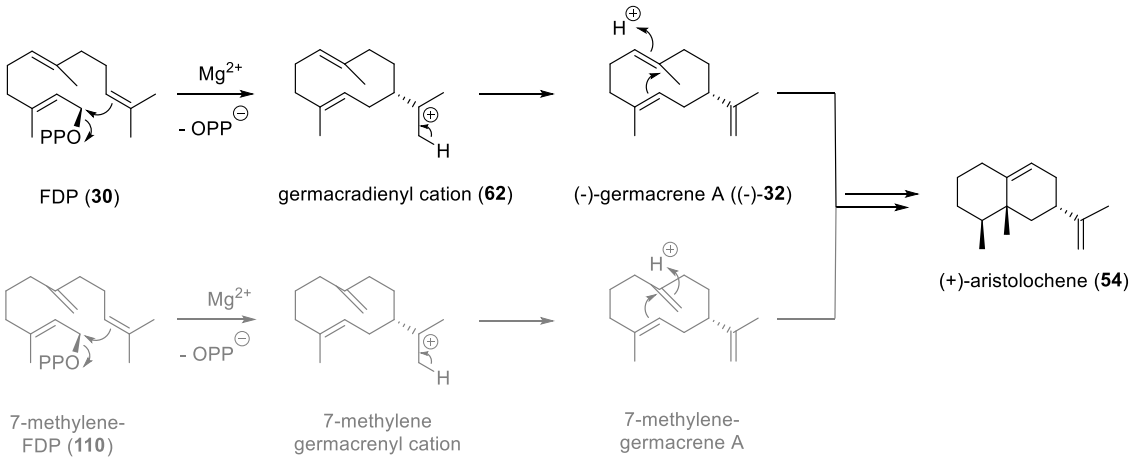
Nucleophiles can also be inserted into the FDP- carbon chain with the aim of trapping known carbocation intermediates, as shown by Alleman *et al.*[132] Allylic alcohol **107** and epoxide **108** are converted to medium size cyclic ether **109** (Scheme 18) by GAS and GdolS (other enzymes were also tested confirming the proposal).



Scheme 18. Proposed reaction mechanism for the GAS- and GdolS-catalysed conversions of alcohol **107** and epoxide **108** to the medium size cyclic ether **109**.

1.6.3. 7-methylene-FDP

7-methylene-FDP (**110**) is an isomer of FDP, in which the 6,7-double bond is shifted to the 7,14-position. Incubation of **110** with PRAS produces (+)-aristolochene (**54**) (50.1%), 7-methylene germacrene A (30.1%), valencene (10%) and 2 other minor products (no farnesenes observed). This work demonstrates the versatility of PRAS to convert an unnatural substrate to the natural products (+)-aristolochene and valencene and shows the possibility to access complex natural products by rational engineering of the substrate.[138]



Scheme 19. Proposed reaction mechanism for the PRAS-catalysed conversions of FDP (**30**, black) and 7Me-FDP (**110**, grey) to form (+)-aristolochene (**54**).

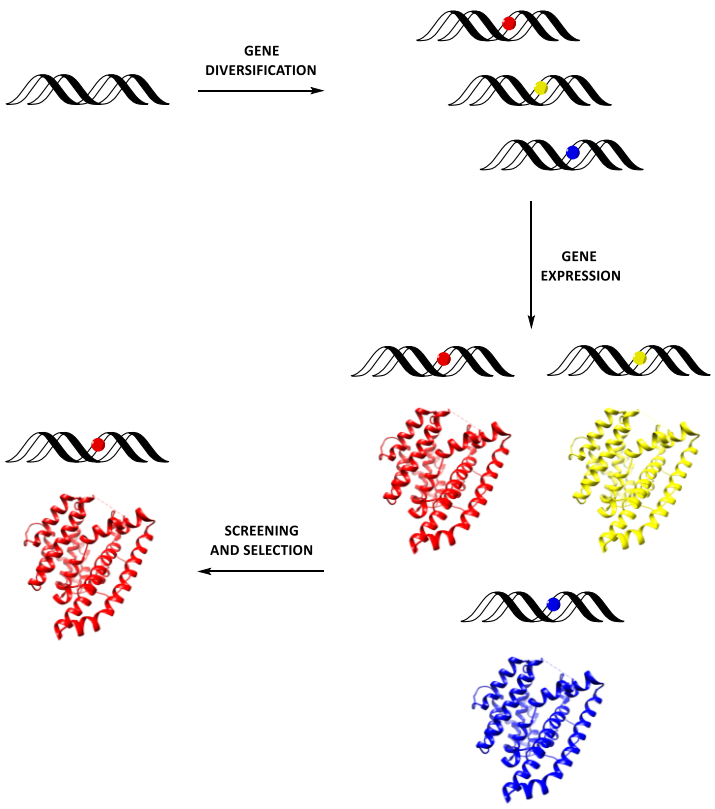
1.7. Terpenoid overproduction. Directed evolution on terpene synthases

Terpenoids are a highly diverse family of natural compounds from which many commercial chemicals are derived, including flavours, fragrances, agrochemicals and medicines. These valuable compounds are traditionally isolated from plants, microbes and marine organisms, but their access is limited, as purification from these natural sources is tedious because they are produced in small quantities and require huge amounts of natural resources. Their extraction needs to be performed with non-environmentally friendly organic solvents, and alternative chemical synthesis of these molecules represents a synthetic challenge for chemists and is frequently associated with low yields and high costs.[139,140]

The engineering of metabolic pathways to produce large quantities of terpenoids in a controllable biological host is an alternative to extractions from the natural source or chemical synthesis.[141–147] In fact, the reconstitution of heterologous pathways in microorganisms has been demonstrated to reach the desired metabolites in the g/L scale, including amorpha-4,11-diene (**104**),[148] taxa-4,11-diene (**4**)[149] and lycopene (**8**)[150] and others. These can be further improved by using fermentation processes.[151,152]

These approaches have served to increase the production of terpenoids, however, the lack of steady-state kinetic information and an incomplete understanding of the terpenoid biosynthetic pathways is a limitation for an efficient terpenoid overproduction platform.[153] Specifically, terpene synthases are generally slow enzymes and are considered the major limiting factor in the total production of terpenoids.[154]

Directed evolution represents an excellent 'nature-mimicking' approach for the engineering of enzymes with improved properties in comparison with that of the natural ones.[155,156] The evolution of protein in the laboratory relies on methods to generate genetic diversity in a short amount of time and to identify and isolate the desired protein variant by chemical or physical means.[157,158]



Scheme 20. Design for the directed evolution of enzymes.

In general, genetic diversity is achieved in the laboratory *via* two different approaches:

1) *Focused mutagenesis*, in which only the amino acids likely to influence the desired function (e.g. selected amino acids inside the active site) are exposed for mutations. Normally, these amino acids are individually replaced with all possible amino acids by saturation mutagenesis, which involves the use of degenerate codons.[159]
2) *Random mutagenesis*, which does not consider any structure-function information within the exposed gene. In the past, random mutagenesis has been achieved by chemical or physical stresses, such as alkylation, deamination or UV radiation. Nowadays, these mutations are normally generated by error-prone polymerase chain reaction (EP-PCR), which makes use of low fidelity DNA polymerases, unequal concentrations of NTPs and Mn^{2+} cofactor to increase the error rate during DNA synthesis.[160]

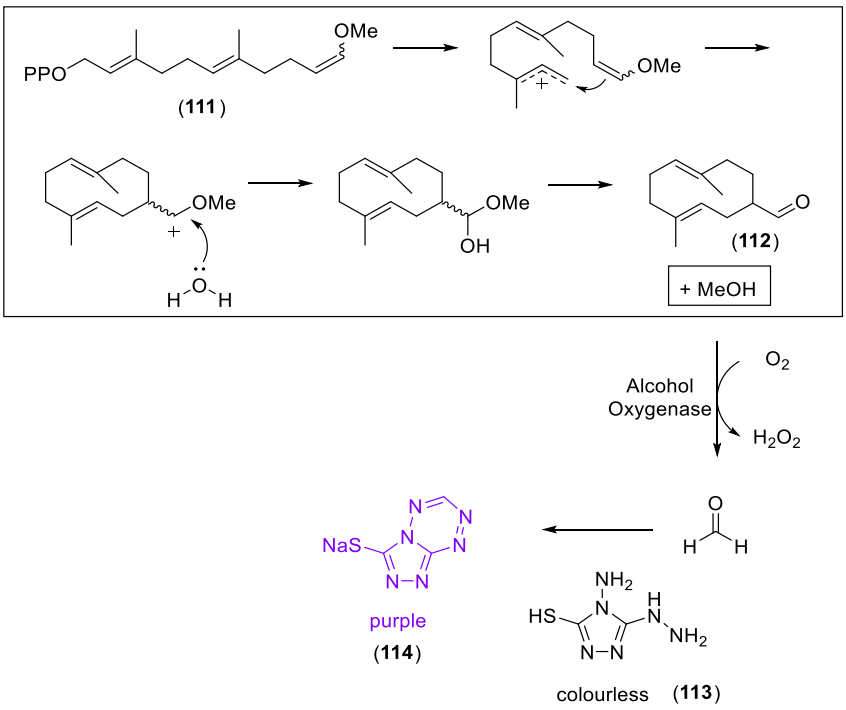
Following the generation of a library of mutants, it is desired to apply a screening method for the selection of the best evolutionary product, which must be adjusted to the nature of the enzymatic or product properties under investigation.

Regarding terpene synthases, the use of a directed evolution approach to improve their catalytic activity has been limited to the availability of high-throughput screening systems. This is not trivial, as their

substrates and products are colourless and are often non-essential secondary metabolites. However, new genes have been successfully discovered and improved using the following approaches:

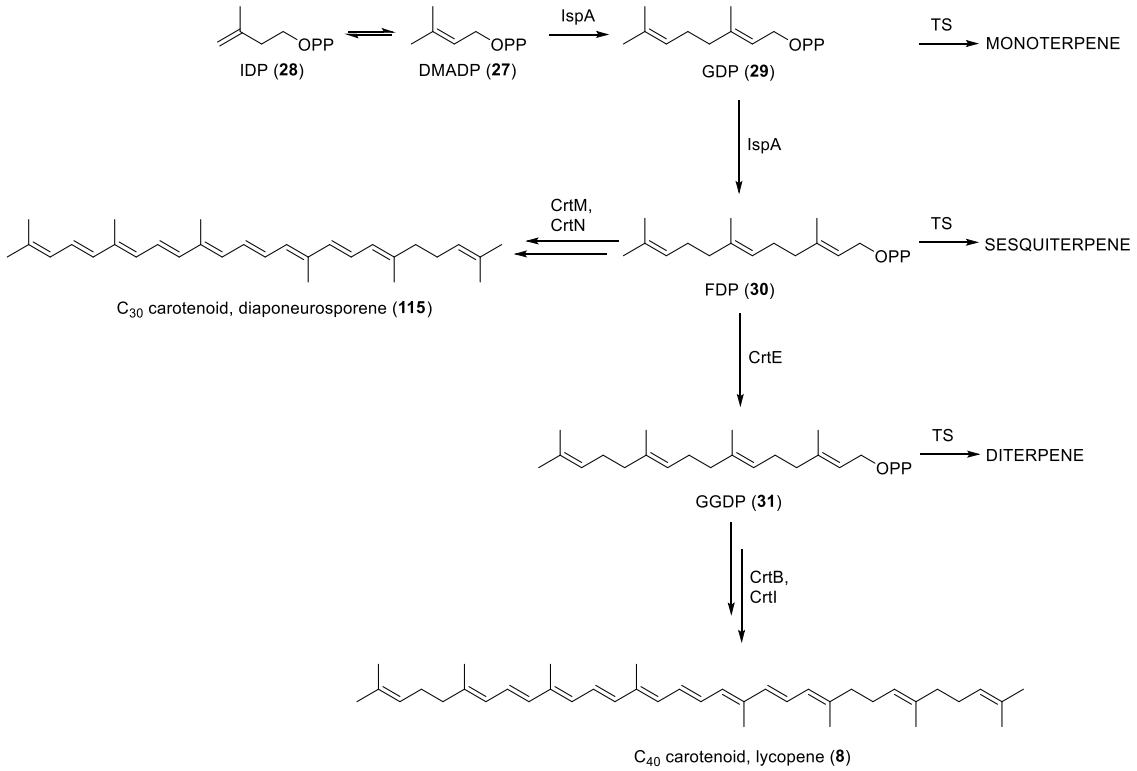
1) Withers *et al.* took advantage of the toxicity of prenyl diphosphates IDP (**28**) and DMADP (**27**) to isolate terpene synthase clones from a library of mutants (fragments of *Bacillus subtilis* genomic DNA). To screen hemiterpene synthases from a set of clones, they co-expressed the necessary genes to convert exogeneous mevalonate to IDP and DMADP (toxic for the cells) and the mutated genes. As a result, the consumption of IDP and DMADP by hemiterpene synthases showed no reduction in the population of cells and allowed the identification of new genes.[161]

2) Lauchi *et al.* used an assay based on product formation to optimise 1,10-sesquiterpene synthases, Scheme 21.[162] The consumption of FDP analogue **111** by 1,10- sesquiterpene synthases generates the aldehyde **112** and methanol, which can be converted enzymatically to formaldehyde in the presence of oxygen. In turn, formaldehyde was reacted with Purpald (**113**) to generate a purple compound (**114**) in the wells containing active enzymes, which absorbance could be measured by UV spectroscopy.



Scheme 21. Conversion of **111** to **112** and methanol by 1,10-sesquiterpene synthases and quantification of methanol by enzymatic-coupled assay.

3) Recently, Furubayashi *et al.* developed a high-throughput colorimetric assay for terpene synthase based on substrate consumption.[163] The assay relies on the competition of two biological pathways for the same substrate. Carotenoids- producing enzymes use FDP (**30**) and GGDPP (**31**) as substrates, which are also the substrates for terpene synthase (Scheme 22). The expression of active terpene synthases results in reduced availability of substrate to produce carotenoids and should attenuate the orange/yellow colour of the host cells. This system allows the discrimination between active or inactive genes and search for optimised activities.



Scheme 22. Substrate competition between biological pathways.

1.8. Aims

This project was divided in three parts, all designed to enhance the usage of sesquiterpene synthases.

The first part was focused on two bacterial sesquiterpene synthases, germacradien-4-ol (GdolS) and epicubenol (EpicS) synthases. Albeit with unequal fidelity, both enzymes generate alcohol products with high degree of stereo- and regio-specificity. These enzymes were to be characterised, and their structure-function relationships to be examined by means of single-mutations in specific active site regions and environmental modulation (pH, cofactor), with particular interest in the water management strategy.

In the second part, it was desired to obtain mechanistic information of terpene synthase reaction mechanisms and to produce novel terpenoids, which might be of potential industrial interests. For this purpose, 6F-FDP (two isomers), 6Me-FDP (two isomers), 7H-FDP, 2,3-Thii-FDP and 10,11-Thii-FDP were to be prepared and their turnover with a set of sesquiterpene synthases, including 1,6- and 1,10- sesquiterpene synthases was to be probed.

The last part covers the development of a directed evolution approach with the aim of improving the catalytic activity of sesquiterpene synthases and/or obtaining enzymes with different functionality. Accordingly, to diversify the genes encoding for 3 different sesquiterpene synthases (GdolS, EpicS and SdS) it was decided to run EP-PCR and subclone the obtained mutated genes by Golden Gate Assembly. The colorimetric selection assay was to be based on substrate consumption and relies on a direct competition between two isoprenoid-generating biological pathways in order to inspect the prior obtained active (from WT enzymes) and inactive (inactive mutant) genes.

CHAPTER 2

STRUCTURE - FUNCTION STUDIES: INVESTIGATION OF WATER MANAGEMENT MECHANISMS IN HYDROXYLATING SESQUITERPENE SYNTHASES

Chapter 2. Structure-function studies: Investigation of Water Management Mechanisms in Hydroxylating Sesquiterpene Synthases

2. 1 Preface

The intrinsic chemistry undertaken by terpene synthases attracts continued interest due to their ability to create a manifold of natural products, and especially derivatives thereof, that match the diverse requirements of our society in agriculture, medicine, or even the cosmetic industry.[164,165] Terpene synthases are of especial mechanistic importance because they generate most of the natural occurring products from simple linear isoprenyl diphosphates. As stated in Section 1.5, extensive studies have uncovered many of the catalytic strategies used by these enzymes, however a multitude of strategies still remains elusive. A small number of terpene synthases can generate hydroxylated products through regio- and stereospecific 'water capture', which represents a (not fully explored) biological short-cut to the production of oxidised terpenes and whose mechanistic intricacies are yet to be understood. These are of particular interest, as oxidation steps of pure hydrocarbon terpene intermediates are usually performed by P450 cytochromes, and water capture strategies by terpene synthases offer means to circumvent these additional functionalisation steps.[166,167]

The active site of terpene synthases includes hydrophobic aliphatic and aromatic residues, which act as a structural template for the binding of FDP (**30**) and a guide for the carbocation intermediates (Sections 1.4 and 1.5). The volume of a terpene synthase active site is slightly larger than the volume of FDP (**30**), which ensures a neat fit between enzyme and substrate for catalysis. With this, water molecules are sterically displaced from the active cavity upon FDP (**30**) binding, which provides chemical protection from an aberrant water molecule that could prematurely quench a carbocation intermediate.[33,67] Despite this, it is found that many hydroxylated terpenes arise from the quenching of carbocation intermediates by water molecules.[84,168–176] Even in non-hydroxylating enzymes water molecules have been shown trapped in closed active sites (e.g. as observed in bornyl diphosphate synthase from *Salvia officinalis*[177] and aristolochene synthase from *Aspergillus terreus*[178]).

Terpene synthases encompass a refined architecture in which water management proves crucial for selective product formation, however detailed mechanistic information of this strategy is limited.

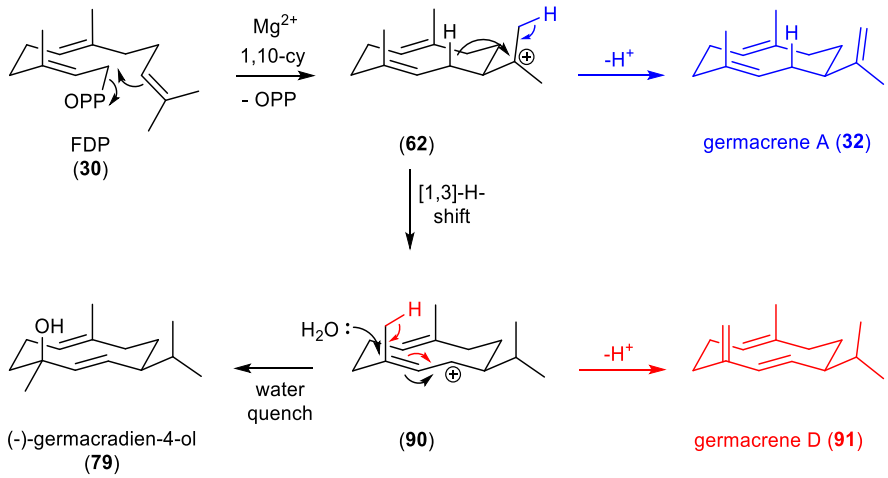
The work presented in this chapter describes the study of the water management strategy of two hydroxylating enzymes; (-)-germacradien-4-ol synthase from *Streptomyces citricolor*[125,169] (GdolS) and (+)-epicubenol synthase from *Streptomyces griseus*[170] (EpicS). While GdolS converts FDP (**30**) into the single alcohol product (-)-germacradien-4-ol (**79**), EpicS exerts lower fidelity and some coproducts (hydroxylated and non-hydroxylated) are formed in addition to (+)-epicubenol (**78**), the main product. The mechanisms of both catalytic reactions are well established; however, little is known about the enzymatic contribution to assist nucleophilic water addition. Here, the effect of amino acids and assay condition modulations on the reactions catalysed by GdolS and EpicS are analysed, and it is

endeavoured to learn about the water management mechanism. A particular focus will lie on their conversion into non-hydroxylating enzymes and the regulation of their water management strategies.

First, the production and characterization of GdolS and EpicS wild types were carried out. These enzymes were incubated with FDP (**30**) and their products were analysed by GC-MS, also the steady-state catalytic parameters of the reaction catalysed by these enzymes were measured.
For this study, analysis of the active sites of GdolS and EpicS was needed, but the absence of a closed crystal structures in complex with a ligand restricted the interpretation of the enzymatic water management strategies in GdolS and EpicS. Computational studies and homology modelling have been assessed to understand mechanistic details in terpenoid synthases.[82,179–183] In view of this, the active site of GdolS and EpicS were explored with the assistance of available X-ray crystal structures in closed conformations from related sesquiterpene synthases, i.e. selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis*[85] (SdS, [4OKZ]), aristolochene synthase from *Aspergillus terreus*[178,184] (ATAS, [4KUX] and [4KVD]) and aristolochene synthase from *Penicillium roqueforti*[99] (PRAS, [1F1P]). Upon this, selected amino acids were mutated to examine their potential contribution to the incorporation of water during catalysis. Lastly, the assay conditions, specifically the pH and metal ions, were altered to attempt to change the product profile and gain further details of the water usage in the investigated terpene synthases.

2.2. (-)-Germacradien-4-ol synthase (GdolS)

The bacterial sesquiterpene synthase (-)-germacradien-4-ol synthase from *Streptomyces citricolor* (GdolS) catalyses the Mg^{2+}- dependent formation of the alcohol (-)-germacradien-4-ol (**79**) from FDP (**30**, Scheme 23). Since it acts with high fidelity and has a relatively simple catalytic mechanism, it represents an ideal model for the examination of active site water control by terpene synthases. The catalytic mechanism of GdolS is proposed to proceed through Mg^{2+} mediated PP_i cleavage followed by a 1,10-ring closure to form the germacrenyl cation (**62**).[125] Gd4olS then drives a 1,3-hydride shift protecting intermediate **62** from proton loss that would give germacrene A (**32**) and, after formation of the allylic carbocation **90**, allows selective water quenching of intermediate **90** to form the final (-)-germacradien-4-ol (**79**). An alternative proton loss is avoided, which otherwise would generate the hydrocarbon product germacrene D (**91**, Scheme 23).[125,169]



Scheme 23. In black, proposed reaction mechanism for the conversion of FDP (**30**) to (-)-germacradien-4-ol (**79**), catalysed by GdolS. Blue and red are alternative pathways to generate germacrene A (**32**) and germacrene D (**91**), respectively.

2.2.1. Characterisation of the wild-type GdolS

The GdolS encoding expression vector (pET16b-GdolS) was a generous gift from Prof. Yasuo Ohnishi (University of Tokyo).[169] The pET16b expression vector contains a T7 promoter (IPTG inducible) for controlling expression and an antibiotic resistance marker (ampicillin) for cell selection. The GdolS gene is located downstream of an in-frame sequence encoding for an *N*-terminal deca-histidine tag, thus allowing purification by nickel-affinity chromatography. GdolS was produced and purified following the protocols described in the literature.[125,169]

Heterologous expression

Escherichia coli BL21(DE3)-RP cells were used for harbouring pET-GdolS under T7 promoter control. After gentle defrost of a 50 μL aliquot of calcium competent BL21(DE3)-RP cells, 50 ng of plasmid (pET-GdolS) were added and the mixture was kept on ice for 30 minutes. Following heat-shock (40 ºC, 45 s), the mixture was returned onto the ice. Then, 1 mL of LB medium was added to the mixture and incubated at 37 ºC for an hour with constant shaking (150 rpm). Cells were harvested by centrifugation (1 min, 3300 g), resuspended in 200 mL fresh LB medium and plated into LB-agar medium containing the appropriate antibiotics for selection (100 μg·mL^{-1}, ampicillin). The plate was incubated for 16 h at 37 ºC. After this time, a single colony was used to inoculate 100 mL of LB medium containing the appropriate antibiotics (100 μg mL^{-1} ampicillin). The culture was grown at 37 ºC overnight whilst shaking (150 rpm). The day after, 5 mL of this culture was transferred to 500 mL of LB medium containing the appropriate antibiotic (100 μ·mL^{-1} ampicillin). The culture was incubated at 37 ºC until OD_{600} = 0.6 was reached, was then induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG, 0.2 mM) and allowed to grow for additional 4 hours. Cells were harvested by centrifugation at 4 ºC (3400 g, 10 min). The supernatant was discarded, and the pellet stored at -20 ºC.

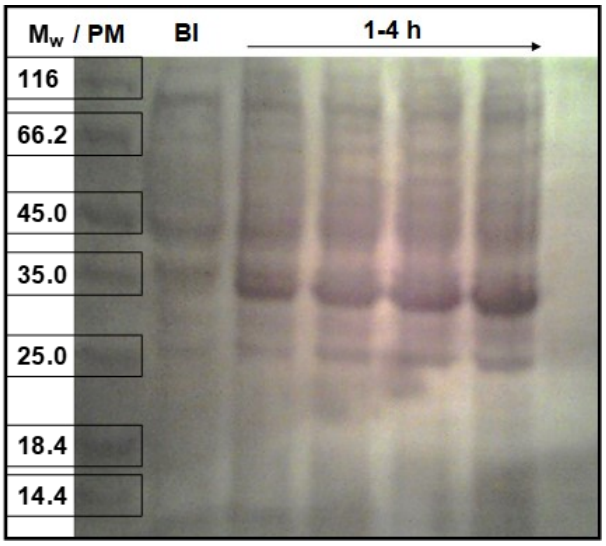


Figure 9. SDS-polyacrylamide gel showing the expression of GdolS, expected Mw = 38.687 kDa. PM: protein marker, BI: sample before induction.

Extraction and purification

The stored frozen pellet was thawed on ice and resuspended in cell lysis buffer (40 mL, 50 mM Tris, 100 mM NaCl, 5 mM imidazole, 10% glycerol (v/v), pH 8). Cells were disrupted by sonication (4 ºC, 5 min, pulse 5 s on/ 10 s off, 40% amplitude) and the resulting suspension was centrifuged. SDS-PAGE showed that the protein was present in the supernatant and that was applied to a Ni-NTA affinity column. The column was washed with a gradient of imidazole (from 10 mM to 500 mM), and the collected fractions were analysed by SDS-PAGE (Figure 10). The fractions containing >90 pure GdolS (molecular weight 38,687 Da) were combined, dialysed for 16 hours in dialysis buffer (10 mM Tris, 10% glycerol, pH 8) and concentrated (Amicon, YM 30). The protein concentration was measured by using the method of Bradford[185,186] and stored at 0 °C.

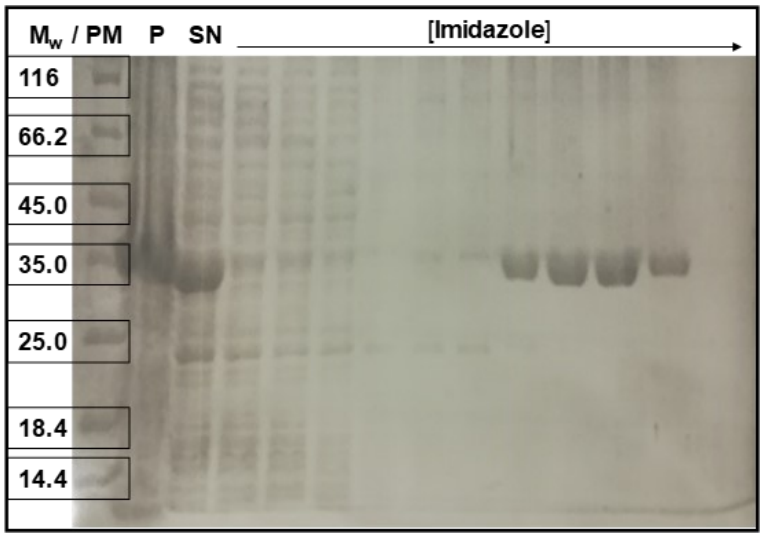


Figure 10. SDS-polyacrylamide gel showing the purification of GdolS. PM: protein marker, P: pellet, SN: supernatant.

Incubation with FDP and product analysis

Analytical incubations of GdolS with FDP (**30**) were carried out to test the activity of the purified enzyme, as well as to be able to compare the wild-type product profile with those of future mutants. The incubation was performed by overlaying GdolS-FDP mixture with pentane and shaking at 25°C.

GC-MS analysis of the organic phase after 24 h showed a single product, with a molecular ion of m/z = 222 ($[M]^+$, Figure 11), which corresponds to (-)-germacradien-4-ol (**79**). The mass spectrum is in good agreement with the literature.[125,169]

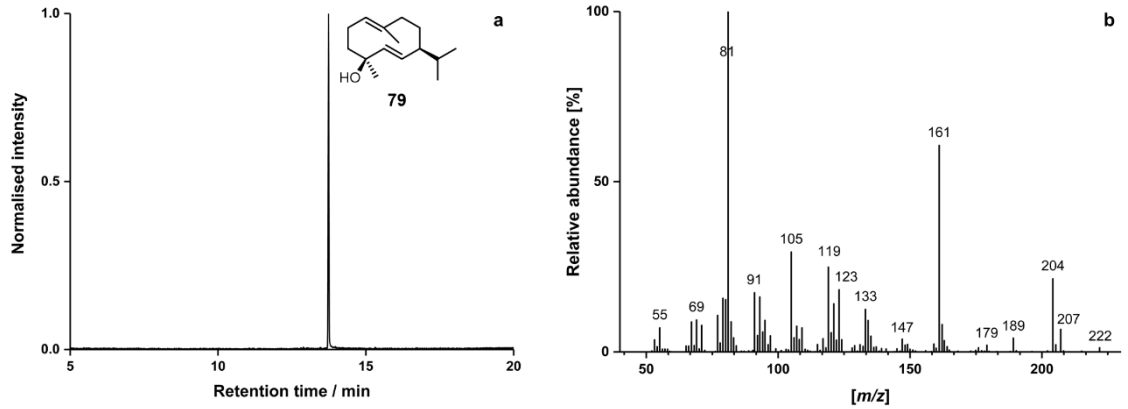


Figure 11. a) Total ion chromatogram of the pentane extractable products arising from the incubation of GdolS with FDP (**30**). b) Mass spectrum of (-)-germacradien-4-ol (**79**).

Steady-state kinetics

In order to determine the steady-state catalytic parameters of GdolS, kinetic assays were carried out according to the standard radiolabelled-substrate procedure commonly used with terpene synthases,[187] and as stated in materials and methods (Section 7.1.19). This technique involves the incubation of various concentrations of [1-^{3}H]-FDP with a fixed concentration of enzyme (GdolS), for a defined time and temperature. After the incubation time was completed, the reaction mixture was extracted in an organic solvent and passed through a small silica column. Due to the substrate's low solubility in the organic solvent, only the enzymatic radiolabelled products were collected. The organic extracts were pooled into Ecoscint fluid cocktail (National Diagnostics) and the radioactivity was measured using a liquid scintillation analyser (TRI-CARB 2900TR). The data collected (CPM) was converted to rate vs substrate concentration and fitted to a Michaelis-Menten curve, $V = (V_{max}[S])/(K_M+[S])$, to determine the kinetic parameters k_{cat} and K_M.

Prior to performing steady state kinetics, the optimal assay incubation conditions were determined, namely [Mg^{2+}], [enzyme], pH and time (Figure 12). Initially, the catalysis buffer for the incubation of (2*E*,6*E*)-[1-^{3}H]-FDP with GdolS was identical to that reported, which consists of 50 mM HEPES

(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 2.5 mM $MgCl_2$ and 5 mM βME (2-mercaptoethanol) at pH 8.0.[125,169]

To investigate the incubation time, analytical incubations with [1-^{3}H]-FDP (4 μM) and GdolS (30 nM) in incubation buffer (50 mM HEPES, 2.5 mM $MgCl_2$, pH 8.0) were performed between 0-20 min at 30 ºC. Activity vs time increased linearly up to 20 min. Therefore, 10 min was chosen as incubation time, which is in the linear conversion range and delivers enough activity to be measured adequately. Thereby it is important that time and activity show a linear relationship to ensure measurement of the initial rate of the reaction. For the optimisation of enzyme concentration, a [GdolS] range of 20-500 nM was incubated with [1-^{3}H]-FDP (4 μM) for 10 min at 30 ºC in the incubation buffer (50 mM HEPES, 2.5 mM $MgCl_2$, pH 8.0). The reaction rate increased linearly up to 200 nM and plateaued at higher concentrations, which is indicative of enzyme aggregation. In the linear range, the enzyme concentration is proportional to the rate, therefore a concentration of 100 nM was chosen, as it is within the linear range and provides satisfactory activity to be measured. As for the optimization of magnesium cofactor concentration, 100 nM GdolS was incubated with [1-^{3}H]-FDP (4 μM) in incubation buffer containing a range of 0.5-10 mM of [Mg^{2+}], for 10 min at 30 ºC. The maximum value observed correspond to 2.5 mM Mg^{2+}, and hence this concentration was selected. Finally, the pH for the incubations was varied between 5.0 and 10.0. The rate increased up to pH 8.0 and decrease from 8.5 to 10.0. Therefore, a pH of 8.0 was chosen for all further incubations.

Following this, the optimised conditions: 100 nM enzyme, 2.5 mM $MgCl_2$, pH 8.0 and incubation time of 10 min were used to measure the steady state kinetic parameters of the reaction catalysed by GdolS (Figure 13). For this, a range of [1-^{3}H]-FDP (0.1 μM- 40 μM) was used and the protocol was carried out as described in materials and methods (Section 7.1.19). The k_{cat} and K_M values were determined to be 0.068 ± 0.001 s^{-1} and 1.02 ± 0.09 μM respectively, which are in good agreement with those reported previously in our laboratory (0.079 ± 0.003 s^{-1} and 1.07 ± 0.13 μM).[125]

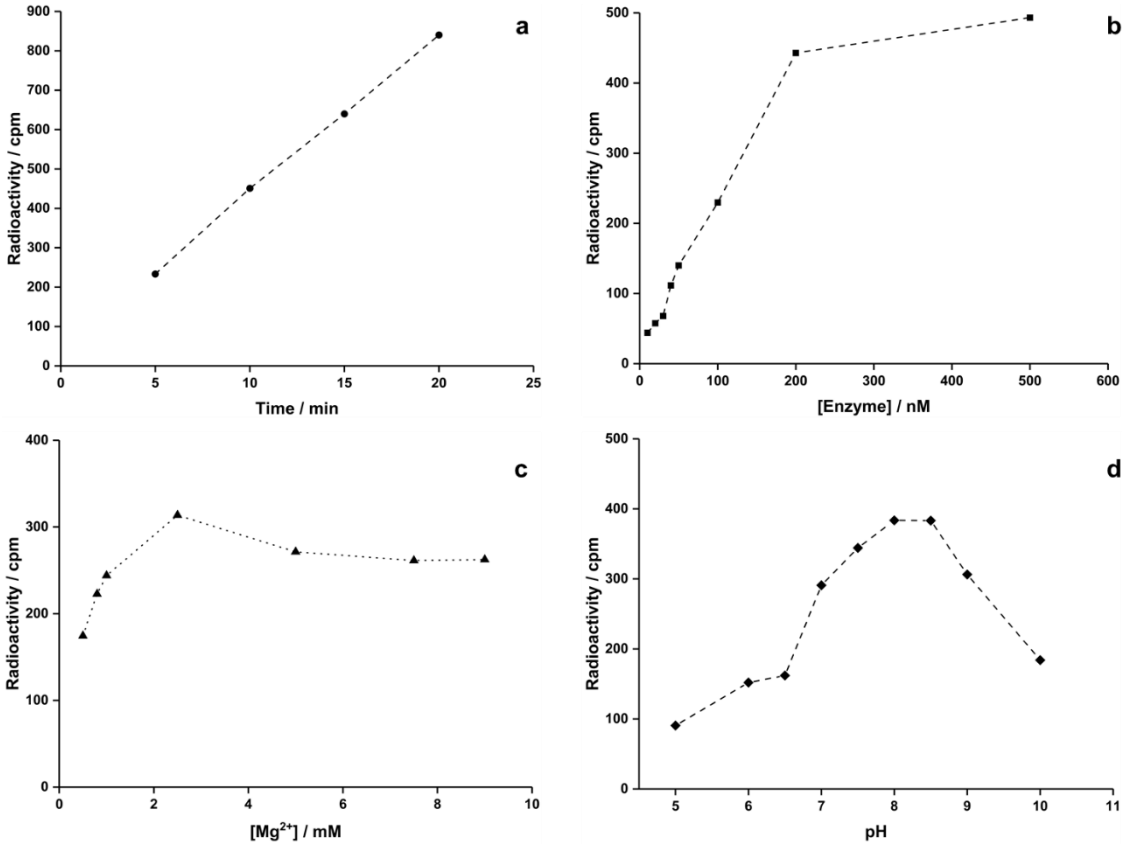


Figure 12. Optimization kinetics conditions for GdolS. Effects of variation of a) time, b) [enzyme], c) [Mg2+], and d) pH on the reaction rate.

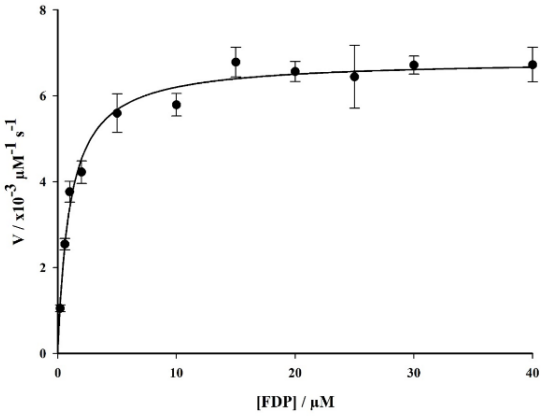


Figure 13. Representative Michaelis-Menten graph of steady-state kinetic parameters of GdolS.

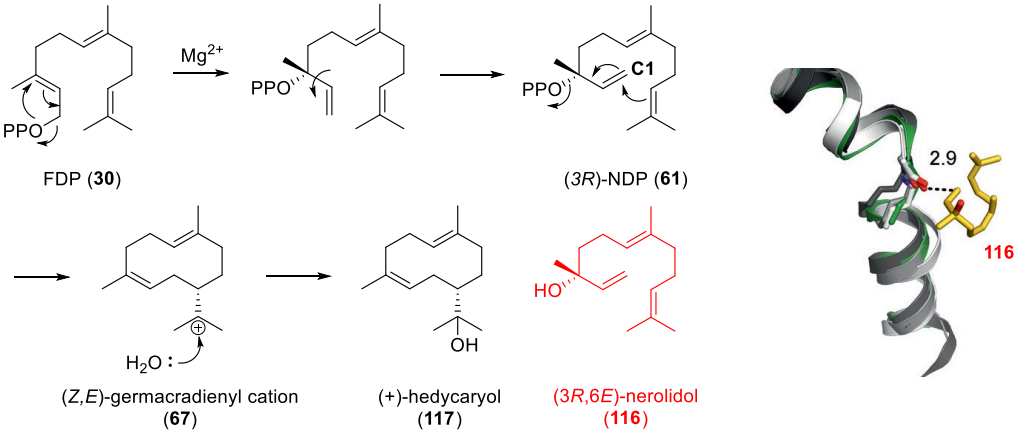
2.2.2. Water management study - Site directed mutagenesis

The crystal structure of GdolS was solved in its open conformation (no ligand bonded) in 2016,[125] and shown to consist of a single α-domain. The origin of water was elucidated using buffer prepared from isotope labelled $H_2{}^{18}O$. Several mutants of GdolS have been investigated, but most of them had minor effect on the wild-type product profile or led to catalytically inactive enzymes. The only interesting candidate, GdolS-N218Q, generated germacradien-4-ol (**79**) and germacrene A (**32**) in similar amounts, however, the catalytic activity was compromised as compared to the wild-type enzyme. N218 is part of the NSE motif (see Section 1.4.1) and the drastic reduction in catalytic efficiency may have arisen from disturbing the $Mg^{2+}{}_3$ cluster conformation. No suitable (polar) water binding residues were identified in the active site and it was postulated that loop movements facilitate entry of bulk water at the active site entrance quenching the final carbocation generated. In the following, several known mechanistic details are laid out to simplify the reasoning for the mutagenesis study carried out in this work.

G1/2 helix break

The G1/2 helix break motif, which is a small loop in the middle of the central G helix (Figure 4 and 7 at Section 1.4.1), has been shown to play a key structural task for catalysis among terpene synthases (see below).

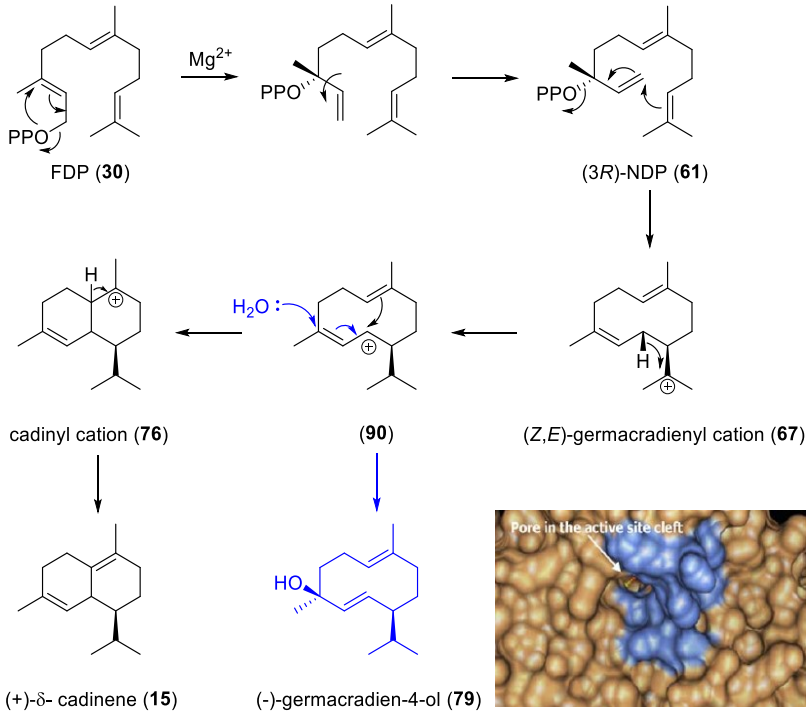
In (+)-hedycaryol synthase (HcS) from *Kitasatospora setae*,[84] the carbonyl group of V179 (located in the G1/2 helix break) assists reionization of NDP (**61**). The crystal structure of HcS in complex with (3*R*,6*E*)-nerolidol (**116**),[84] which resembles the intermediate 3*R*-NDP (**61**) shows that the G1/2 helix break dipole (and the carbonyl group of V179) points towards the C1 of **116** (Scheme 24).



Scheme 24. Proposed reaction mechanism for the formation of (+)-hedycaryol (**117**) catalysed by HcS. Right, cartoon representation of the superposition of X-ray crystal structures from HcS-**116** (green) and other enzymes (pentalenene, epi-isozizaene and taxadiene synthases), highlighting the distance between the carbonyl group of V179 and the C1 of **116**. Picture taken from reference [84].

Hence it is proposed to help the flux of electrons from the C1-C2 double bond to the C2-C3 position and migrate the emerging positive charge in C3 to C1 to enhance the 1,10 cyclisation that leads to (2*Z*,6*E*)-hedycaryol (**117**) after water quench (Scheme 24).

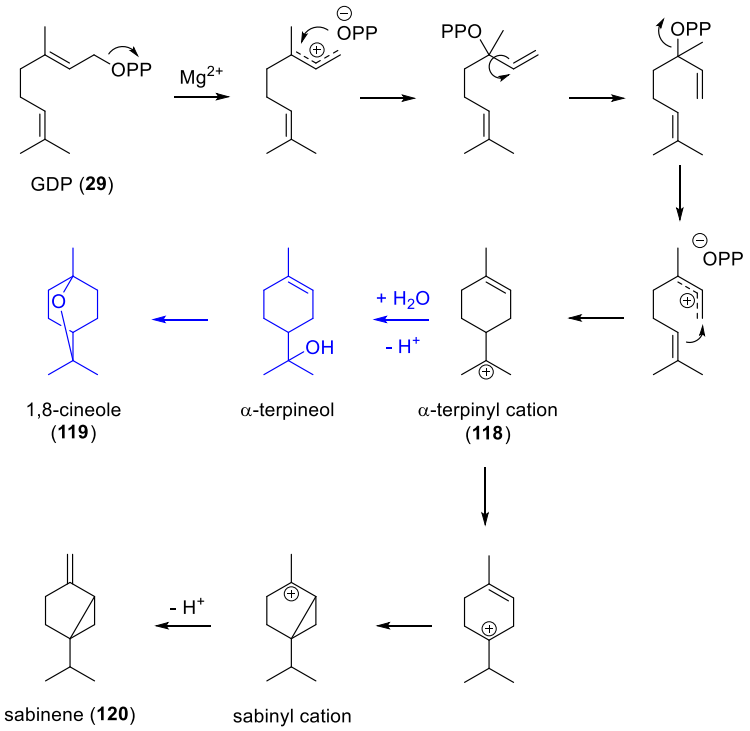
In (+)-δ-cadinene synthase (DCS) from *Gossypium arboreum,*[188], the G1/2 helix break motif plays a critical role in protecting the forming carbocations from the bulk solvent. DCS converts FDP (**30**) to (+)-δ-cadinene (**15**, Scheme 25). The N403P/L405H mutant,[189] which alters the composition of the G1/2 helix break, severely shifts the product profile and results in a functional germacradien-4-ol synthase (53% and 93% of germacradien-4-ol (**79**) production *in vitro* and *in vivo,* respectively). Computational modelling suggests that water could enter through a pore in the active site of the engineered DCS (located at the G1/2 helix break of the mutated enzyme).[189]



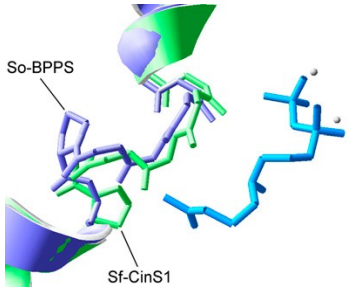
Scheme 25. Proposed reaction mechanisms for the formation of (+)-δ-cadinene (**15**) and (-)-germacradien-4-ol (**79**) by engineered DCS. The inset illustrates the pore observed in the DCS- mutated enzyme, which is located in the G1/2 helix break region. Picture taken from reference [189].

In cineole synthase (*Sf*-CinS1) from *Salvia fruticose,*[190] a monoterpene synthase that converts GDP (**29**) into 1,8-cineole (**119**, Scheme 26), the corresponding helix break was also observed to be of structural importance. This helix break comprises residues I446, G447, G448, I449 and P450. The introduction of a second Pro residue at I449 resulted in an inactive enzyme, whereas substitution of Thr for P450 led to roughly 87% production of sabinene (**120**), which is a non-hydroxylated product resulting from the abolishment of nucleophilic water capture in the terpinyl carbocation intermediate **118**. In fact, sabinene synthases from *S. pomifera* (*Sp*-SabS1) and *S. officinalis* (*So*-SabS1) contain Thr at this position,

but cineole synthase from *S. officinalis* (*So*-CinS1) conserves the Pro residue, thus suggesting that P450 is a structural determinant for cineole formation by hydroxylation in cineole synthase (Scheme 26).



Enzyme	Amino acid
So-CinS1	P450
Sf-CinS1	P
Sp-SabS1	T
So-SabS1	T
So-BPPS	P



Scheme 26. Proposed catalysed mechanism for the formation of 1,8-cineole (**119**) and sabinene (**120**) by *Sf*-CinS1-WT and the *Sf*-CinS1-P450T mutant. Bottom, table containing the corresponding residues to P450 in cineole-like and sabinene-like synthase enzymes, and a cartoon representation of the region under investigation, taken from reference [190] .

In SdS, this region comprises the R172 (sensor), D181 (linker) and G182 (effector) triad, which is proposed to assist substrate ionisation and active-site closure (Section 1.4.1).[85]

The overall composition and orientation of this helix break motif is very much conserved among terpene synthases, including TEAS, PRAS, pentalenene synthase and trichodiene synthase, to name few more examples.[70,71,80,85,191]

With this knowledge in mind, the active site of GdolS was re-examined for mutational studies in this work. The G1/2 helix of GdolS is flanked by R165 and G189. This helix is situated in the centre of the active site cleft and contains a helical disruption constituted by amino acids A176 and A177. The G1/2 helix break of GdolS contains R172, T175 and A176 triad, which corresponds to R178, D181 and G182 in SdS[85] (Section 1.4.1) and might possess an important unrevealed structural task in GdolS catalysis (figure 14).

G-helix: DR^{165}EGYLTLRRGTAAMESIFDMIERLG^{189}H

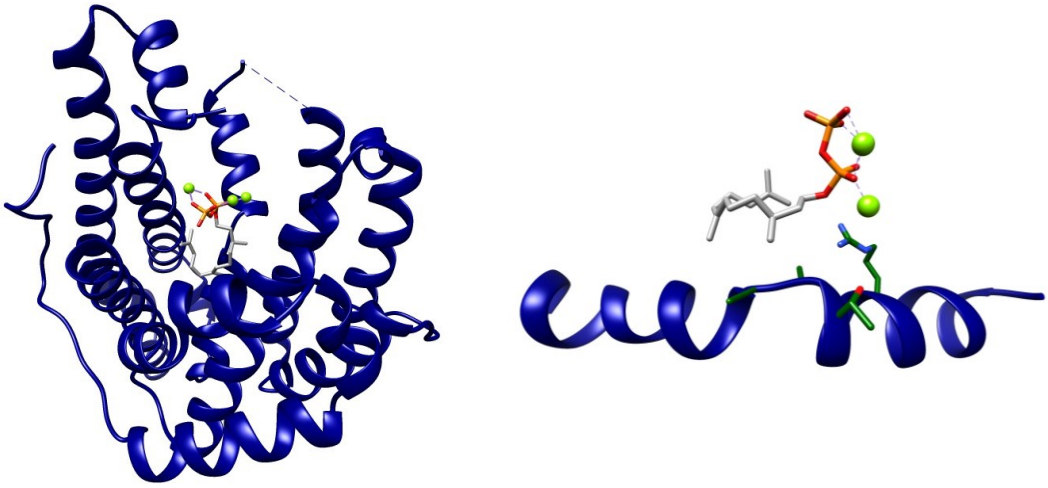


Figure 14. Left, cartoon representation of the crystal structure of GdolS-WT ([PDB:5I1U], blue) in complex with FHDP, docked from SdS [PDB: 4OKZ]. Right, cartoon representation of the G-helix of GdolS-WT (blue) in complex with FHDP, docked from SdS. The cited amino acids are in green.

From these, the important catalytic role of R172 in diphosphate binding was previously analysed by colleagues within this laboratory (unpublished), but the proximity of T175 and especially the "kink" A176 to the ephemerae remained elusive

Herein, we employ site directed mutagenesis (SDM) to elucidate the role of A176 in GdolS, which is proximal to the substrate and a potential contributor to loop movements that may open/ close a water channel through the active site. In addition, neighbouring amino acids T175 and A177 were targeted to further investigate the region of interest. The results show that the A176 side chain is essential for catalysis and product fidelity. Single A176I and A176M mutations drastically alter product distributions, switching GdolS into functional germacrene A synthases.

Expression and purification of selected G1/2 helix break mutants

To assess the importance of the G1/2 helix break composition in GdolS, the following mutants were generated by site-direct mutagenesis: A176G, A176V, A176L, A176I, A176M, A176F, A176D, A176T, A176Q, T175N, T175D, T175C, A177I. The designated primers are illustrated in materials and methods (Section 7.1.8), and the presence of the desired mutations was confirmed by DNA sequencing.

The heterologous expression and purification of GdolS mutants was carried out following the same protocol described for GdolS WT (Section 2.2.1) and as stated in materials and methods (Sections 7.1.13 and 7.1.14).

Analysis of the sesquiterpene products and kinetics studies.

The activity of GdolS mutants was determined by incubation of the resulting proteins with FDP (**30**) as described in materials and methods (Section 7.1.17). The GC-MS analysis of the pentane extractable products showed alternative products in comparison with the wild-type enzyme, which were identified by comparison of the retention time and the mass spectra of authentic enzymatic products of wild type GAS and GDS.

To examine the role of A176 in GdolS catalysis, alanine was replaced by hydrophobic and hydrophilic residues of varying size. In general, substitution of non-polar amino acids for A176 led to different product profiles, whereas replacement of polar amino acids for A176 did not affect product fidelity.

First, A176 was replaced by glycine, a more flexible amino acid with reduced size. Product distribution was not altered in comparison to the WT (Figure 15).

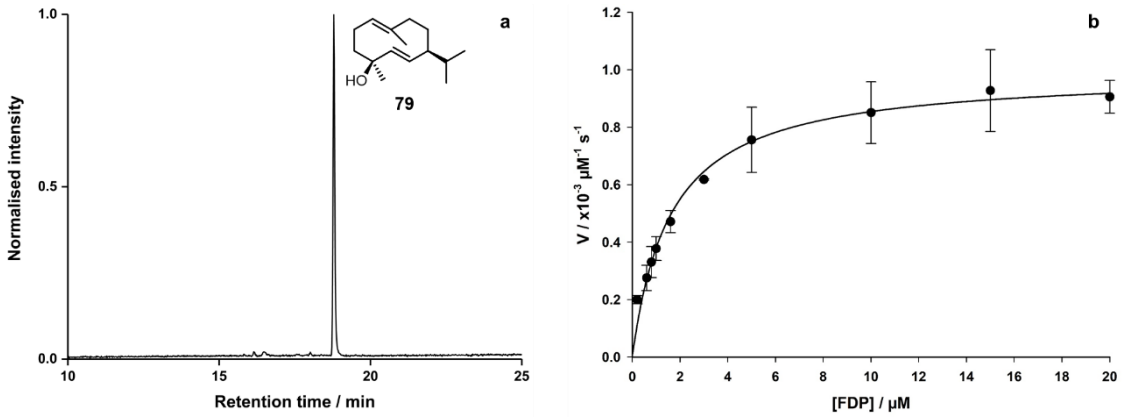


Figure 15. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176G. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176G.

In contrast, the replacement of A176 with valine, which has roughly twice the Van der Waals volume as alanine, resulted in a slightly shifted product distribution, producing 91% of germacradien-4-ol (**79**) and 9% of germacrene A (**32**) (Figure 16), a product that results from deprotonation of intermediate **62** (Scheme 23).

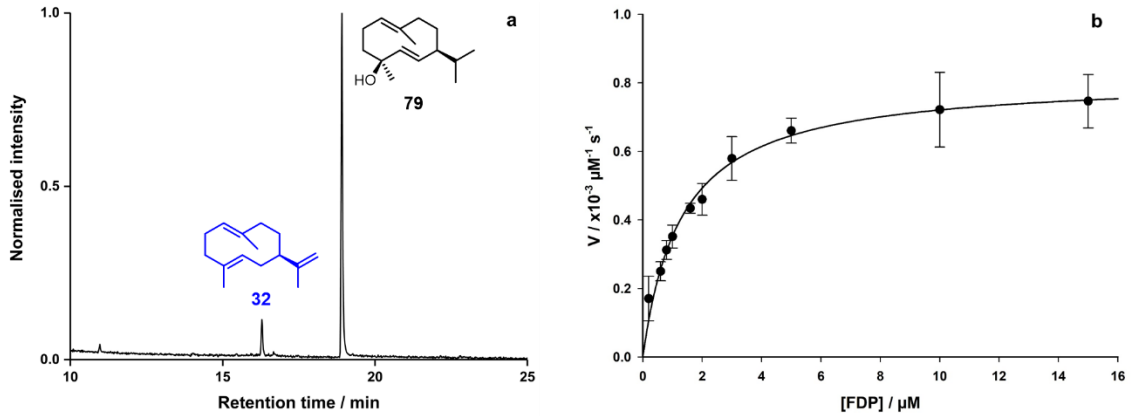


Figure 16. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176V. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176V.

Aliphatic residues have been shown to cooperate with aromatic counterparts to force FDP (**30**) into the correct conformation for cyclisation in PRAS.[88] In GdolS, the absence of farnesene products upon A176V substitution shows that replacement of the methyl group with a bulkier aliphatic side chain does not sterically affect the cyclisation reaction. This finding seems to rule out any contribution of A176 to positioning and orientation of the C10-C11 double bond prior to cyclisation. A homology model of GdolS with unreactive farnesyl-*S*-thiolodiphosphate (FSDP, docked from PRAS[88]) suggests that this might be the role of the bulky aliphatic M73 residue (L108 in PRAS[88]), which is situated in the lower hydrophobic active site region directly pointing towards the prenyl chain tail (Figure 17). This amino acid was not mutated during this study.

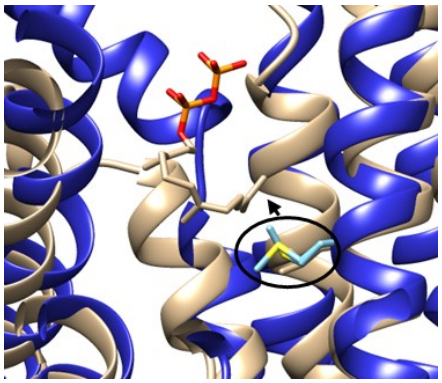


Figure 17. Cartoon representation of the crystal structure of GdolS-WT ([PDB: 5I1U], blue), superposed with the crystal structure of the complex PRAS-FSDP ([PDB: 1F1P], grey). Highlighting GdolS-M73.

Following these results, it was decided to produce A176T and A176L in GdolS. These functional alterations could provide more information regarding the steric and electrostatic contribution of A176 to catalysis. Incubation of A176L with **1** led to the accumulation of germacrene A (**32**, 20 %) and germacrene D (**91**, 6 %) (Figure 18), products that arise from premature deprotonation of intermediates **62** and **90**, respectively (Scheme 23). On the other hand, A176T displayed a product profile similar to that of GdolS-WT. (Figure 19).

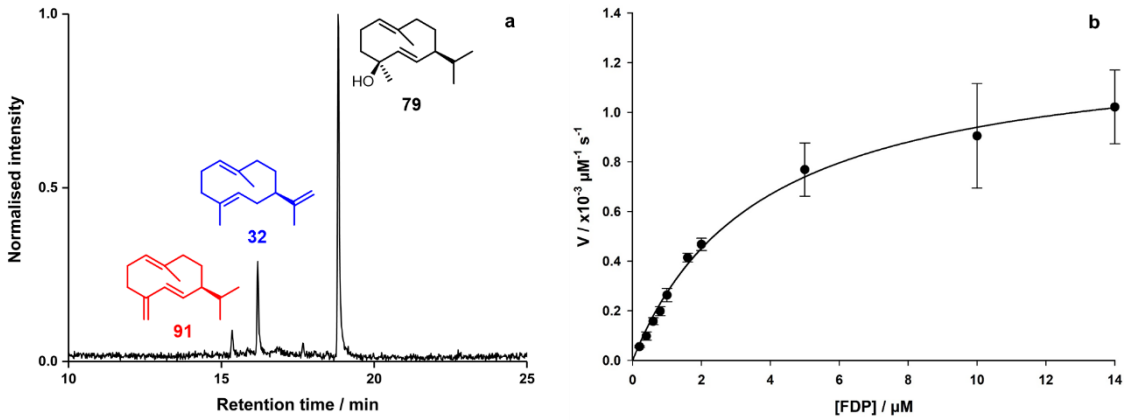


Figure 18. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176L. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176L.

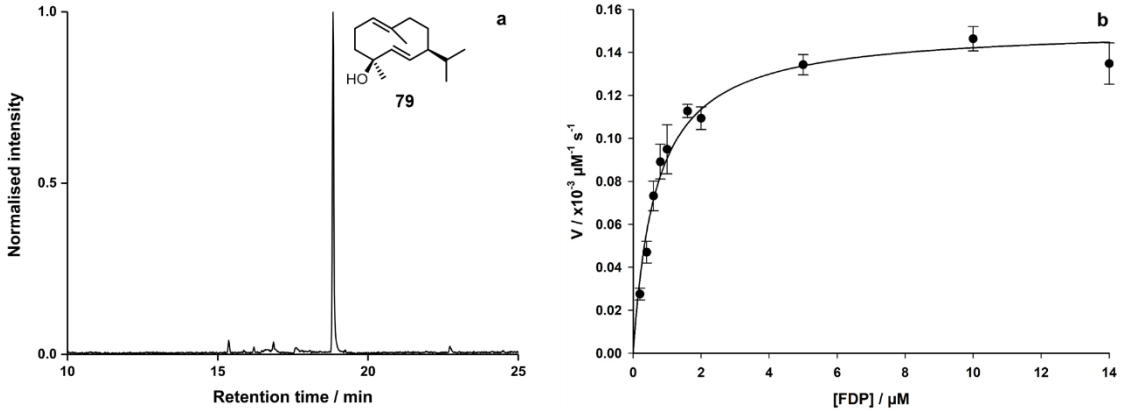


Figure 19. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176T. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176T.

The observed WT-like product profile upon substitution of A176 for a larger but polar side chain (A176T) seems to rule out water activation as the main role for A176. Otherwise a steric displacement of an activated water molecule and formation of aberrant products would be expected, as observed for A176V and A176L (non-polar switches).[192] Also, the possibility of water influx through a channel in

this region is discarded. Instead, A176 perhaps plays a passive catalytic role by comprising part of the template contour after FDP binding.

In terpene synthases, crystal structures often show hydrogen bonding networks, which might be constituted with organised water molecules interacting with the diphosphate motif, Mg^{2+} ions and/or polar side chains.[178] Also, A176 could be close to an aromatic side chain responsible for intermediate **90** cation-π stabilisation.[33,86,98,99] Larger aliphatic groups can disrupt these potential associations in GdolS and prompt a structural adjustment that results in an alternate template and hence 'naturally inhibited' deprotonation of carbocation intermediates **62** and **90** might happen**,** producing products **32** and **91** (Scheme 23), respectively. To probe the hypothesis, A176 was further modified to isoleucine, methionine, phenylalanine, glutamine and aspartic acid. GdolS-A176F was inactive, with phenylalanine presumably occupying too much volume in the active site cavity. The isoleucine variant generated **81** and **32** in similar quantities (41% and 59% respectively, Figure 20). Although leucine and isoleucine have a similar van der Waals volume their different conformations influence reaction termination unequally, highlighting a delicate structural role of A176.

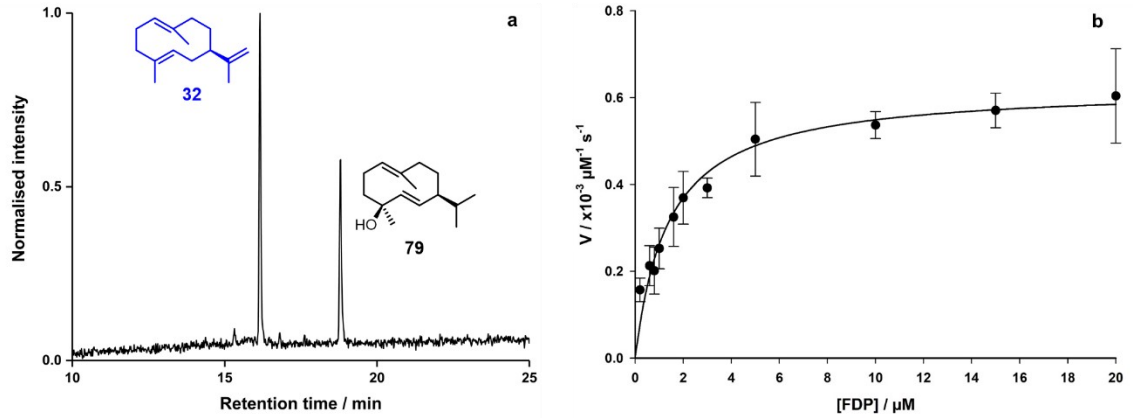


Figure 20. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (30) with GdolS-A176I. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176I.

Consequently, GdolS-A176M exhibits a drastic change in product distribution, with the formation of a 10:1 mixture of **32** and **79** as the predominant reaction products (Figure 21). While, for A176Q and A176D, only germacradien-4-ol (**79**) is observed after incubation with **30** (Figures 22 and 23, respectively).

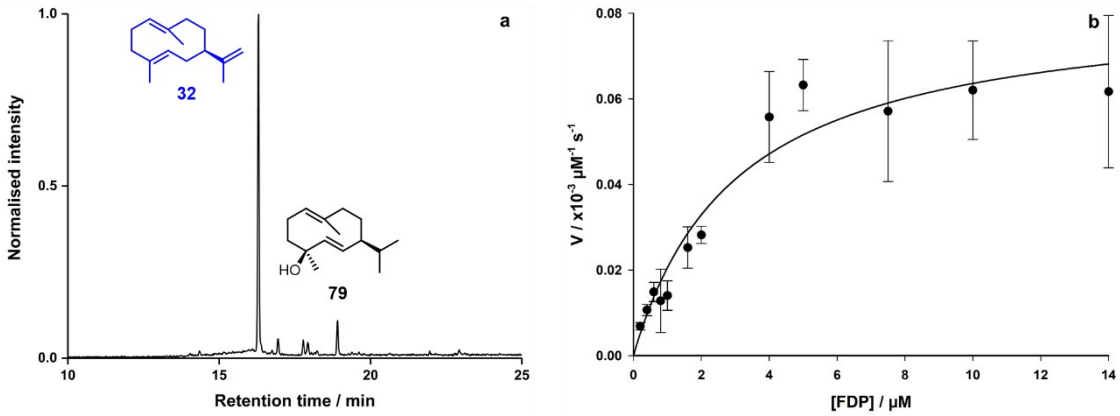


Figure 21. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176M. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176M.

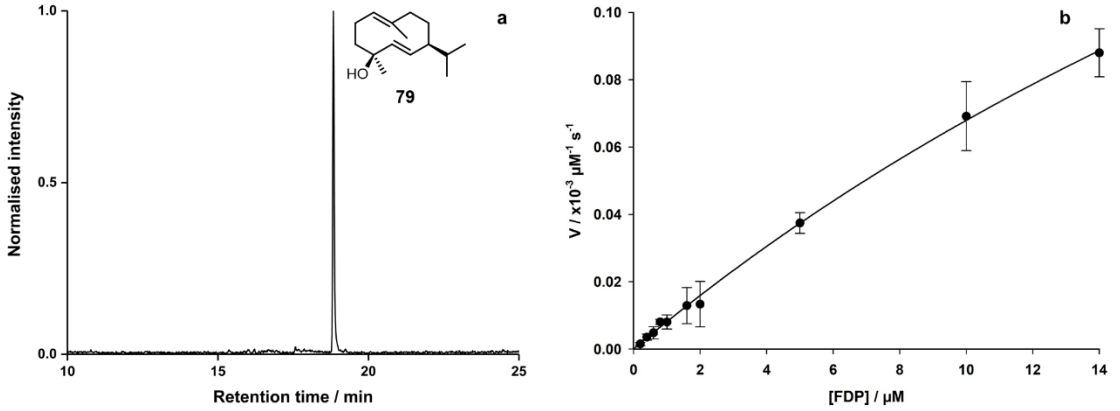


Figure 22. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176Q. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176Q.

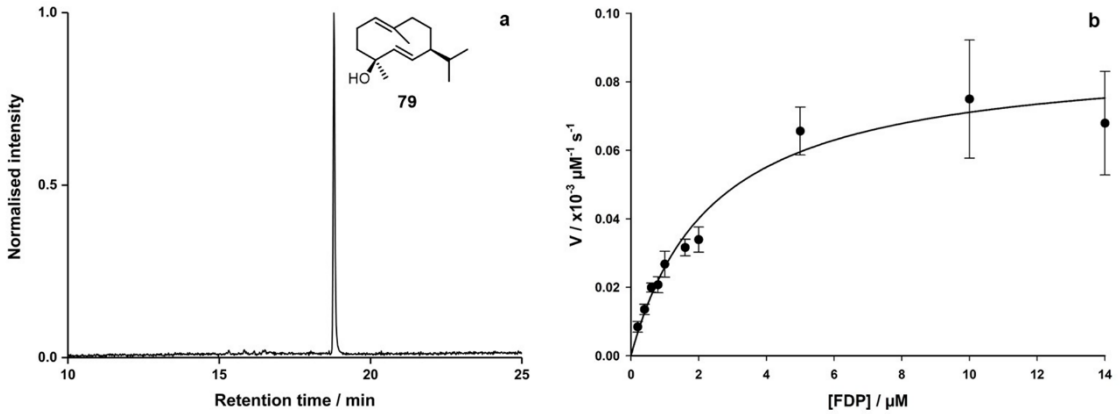


Figure 23. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-A176D. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-A176D.

Taken all results together, it is shown that upon substitution of GdolS-A176 with larger non-polar aliphatic side chains the product distribution shifts to germacrene A (**32**) (Figure 24). A176 plays a key structural role that governs product fidelity, non-polar aliphatic residues modulate the substrate template contour, resulting in a misfunction of water capture. Also, replacement of A176 with the smaller glycine or polar-aliphatic amino acids does not changes the product profile in comparison with the wild type (Figure 24). This might suggest that A176 is located nearby a hydrogen network, which can be altered with larger hydrophobic amino acids, however, A176G does not break these associations as it is smaller and polar amino acids can participate in these associations without disrupting the template for catalysis. The PP_i motif could be responsible for the deprotonation of intermediates **62** and **90** in the mutated enzymes, as proposed for other terpene synthases. [94,193]

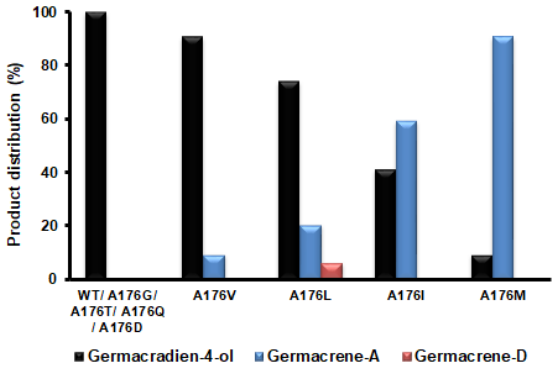


Figure 24. Representation of the product distribution (%) generated by Gd4olS-WT and Gd4olS-A176 mutants.

The k_{cat} and K_M parameters of the reactions catalysed by GdolS-A176 mutants were measured in order to get further catalytic information, Table 5. In general, K_M values for all mutants are similar to those measured for wild-type GdolS. The most significant difference is observed in A176Q, which it is 40-fold higher than that for the wild-type enzyme, implying a significant steric influence on substrate binding. This agrees with the inactivity observed upon A176F substitution. Catalytic efficiency changes were dominated by the k_{cat} values. This stands for a late catalytic influence of GdolS-176 side chains, that after substrate binding defines part of the enzymatic cavity for catalysis. A176G, A176V, A176L and A176I display a k_{cat} about 10-fold lower than that of GdolS-WT. A176M and A176D mutations however result in 100-fold lower k_{cat} in comparison with the wild-type enzyme (Table 5).

Enzyme	k_{cat} [x 10^{-3} s^{-1}]	K_M [μM]	k_{cat}/K_M [x 10^3 M^{-1} s^{-1}]
WT-GdolS	68 ± 1	1.02 ± 0.09	66.7 ± 5.9
GdolS-A176G	10 ± 0.3	1.61 ± 0.18	6.2 ± 0.7
GdolS-A176V	8 ± 0.2	1.31 ± 0.12	6.12 ± 0.60
GdolS-A176L	13 ± 0.6	3.70 ± 0.43	3.51 ± 0.44
GdolS-A176I	6 ± 0.3	1.39 ± 0.19	4.32 ± 0.63
GdolS-A176M	0.83 ± 0.07	3.03 ± 0.66	0.27 ± 0.06
GdolS-A176F	n/a[a]	n/a	n/a
GdolS-A176D	0.9 ± 0.05	2.41 ± 0.35	0.37 ± 0.06
GdolS-A176T	2 ± 0.03	0.67 ± 0.05	2.97 ± 0.22
GdolS-A176Q	4 ± 0.94	44.83 ± 14.15	0.09 ± 0.04

[a] Not applicable.

Table 5. Measured steady-state kinetic parameters for the reactions catalysed by GdolS-WT and GdolS-mutants.

After these findings, it was decided to further investigate the G-helix break region in GdolS.

In SdS,[85] D181 (T175 in GdolS) shifts upon FDP binding to interact with the guanidine group of R178 (R172 in GdolS) and this event triggers a structural rearrangement of G182 (A176 in Gd4olS). Several mutants at T175 were prepared, again using standard SDM techniques. T175C and T175N were able to functionally substitute threonine in the presumed hydrogen-bonding network, with no change in product distribution (Figures 25 and 26, respectively), albeit having smaller k_{cat} values (Table 6). In contrast, GdolS-T175D was inactive.

GdolS-A177 was also mutated to I177, which was obtained mostly insoluble under the expression conditions used. The small amount soluble protein obtained after basic extraction was inactive.

Mutations at the positions T175 an I177 are inconclusive so far and more SDM work would need to be carried out for proper assessment of their enzymatic functions.

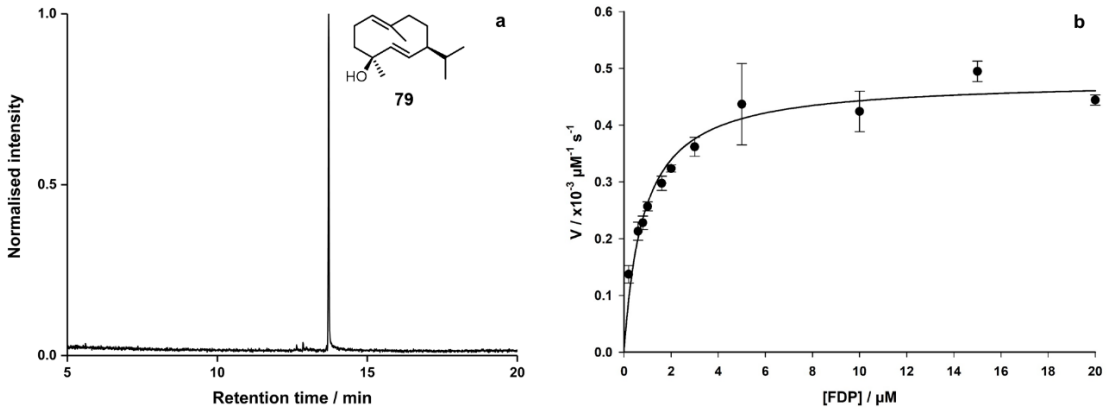


Figure 25. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-T175C. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-T175C.

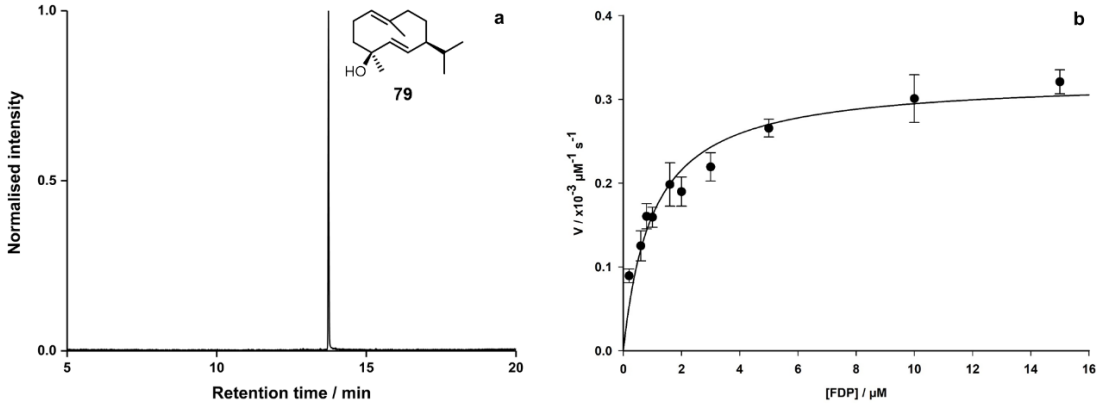


Figure 26. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-T175N. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-T175N.

Enzyme	k_{cat} [x 10^{-3} s^{-1}]	K_M [μM]	k_{cat}/K_M [x 10^3 M^{-1} s^{-1}]
WT-Gd4olS	68 ± 1	1.02 ± 0.09	66.7 ± 5.9
Gd4olS-T175D	n/a[a]	n/a[a]	n/a[a]
Gd4olS-T175C	5 ± 0.12	0.83 ± 0.08	6.03 ± 0.63
Gd4olS-T175N	3 ± 0.09	1.01 ± 0.10	2.97 ± 0.31
Gd4olS-A177I	n/a[a]	n/a[a]	n/a[a]

[a] Not applicable.

Table 6. Steady-state kinetic parameters for the reactions catalysed by GdolS-WT and GdolS-mutants.

Potential water binding residues

Polar residues can bind to water molecules in the active site of terpene synthases, and these might participate in catalysis or not. For example, the water management strategy in ATAS is particularly appealing. ATAS is a high-fidelity terpene synthase that converts FDP solely into (+)-aristolochene (**54**, Scheme 8 at Section 1.5.1).[194] There is no visible role for water in the catalytic mechanism, however, crystal structures of ATAS (and mutants)[184] in complex with the unreactive analogue farnesyl-*S*-thiolodiphosphate (FSDP) as well as aza-analogues showed trapped water molecules in the active site, which are hydrogen bonding to polar amino acids and/or Mg^{2+} ions (Figure 27).[184]

Because hydroxylated products are not observed, these water molecules must be chemically inert. Faraldos *et al.*[178] probed the role of these active site water molecules in ATAS by perturbing their hydrogen bond partners S303, N299 and Q151 (Figure 27).

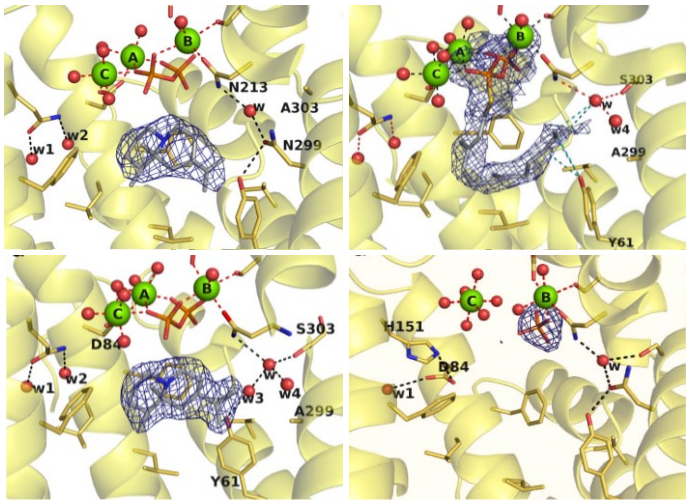


Figure 27. Set of ATAS X-ray crystal structure representations showing the hydrogen bonds and metal ion coordination interactions (black and red dashed lines respectively). Water molecules are labelled as w. Images taken from reference [178].

Replacements of S303 for alanine, histidine and aspartic acid in ATAS led to the accumulation of hydroxylated products ((*E*)-nerolidol (**116**) and (*E*,*E*)-farnesol (**121**)), together with germacrene A (**32**). These results are consistent with defects in the active site template that would otherwise hold the C1 and C10 close enough for the initial 1,10-cyclisation. Furthermore, it was found that there is a Mg^{2+}-bound water molecule close to the C1 and C3 atoms of the farnesyl chain in the ATAS-FSPP complex, which could be responsible for the generation of the linear alcohol products, quenching at either side of the allylic cation.[178]

Regarding N299, substitution for alanine led to the production of aristolochene (**54**) and germacrene A (**32**), which are cyclic products, thus indicative that N299A substitution does not affect the

conformational requirement for 1,10 cyclisation. N299L also produced close to equal quantities of germacrene A (**32**) and aristolochene (**54**), confirming the findings. Interestingly, the X-ray crystal structure of N299A in complex with aza-analogue reveals the binding of two new water molecules, which occupy the former positions of Oδ and Nδ atoms of N299, and Y61 flips 90º to make Van der Waals contact with the aza-analogue (Figure 27).

Lastly, Q151 was mutated for histidine and glutamate. The crystal structure of Q151H-ATAS with FSDP revealed unpredicted structural features in the active site. In here, water molecule "w2" is absent, and the D84 side chain points to an unusual orientation to form hydrogen bonds with "w1", and the Oδ atom of D84 forming a Van der Waals contact with Nε2 of H151. These interactions compromise the metal coordination environment and only $Mg^{2+}{}_B$ and Mg^{2+}c are observed, in which Mg^{2+}c is no longer bonded to the protein but changed the location to coordinate to six other water molecules. Due to this structural disorder, Q151H-ATAS generates a mixture of aristolochene (**54**), germacrene A (**32**), farnesol **121** and racemic nerolidol. The generation of hydroxylated products, as in S303A and N299A, is possibly attributed to the farnesyl/nerolidyl cation interception with $Mg^{2+}{}_C$-bound water molecules. This detailed study represents of how polar side chains can be hydrogen bonded with trapped solvent molecules in the enzyme-substrate complex, which in turn can be part of the active site contour and not necessarily act as a reagent in terpene synthase catalysed mechanism.[178]

Analysis of the crystal structure of GdolS in superposition with ATAS-FSDP (5I1U and 4KUX respectively, Figure 28) pointed to amino acids H150 (Q151 in ATAS), Y303 (N299 in ATAS) and E307 (S303 in ATAS) as potentially relevant water binding residues in GdolS. Indeed, Y303 and E307 were previously analysed in our laboratory as potential water binding residues. These side chains were substituted for amino acids with or without hydrogen bonding capabilities, but the results ruled out the hypothesis.[125] H150 is ideally located not only to form hydrophobic interactions that might help to control the conformation of FDP and active site closure, but could also be involved in water positioning for its use in catalysis. In view of this, H150 was replaced with Y, F, W, C, R. This data shows the importance of H150 for catalytic efficiency but rules out water activation as the main role in catalysis.

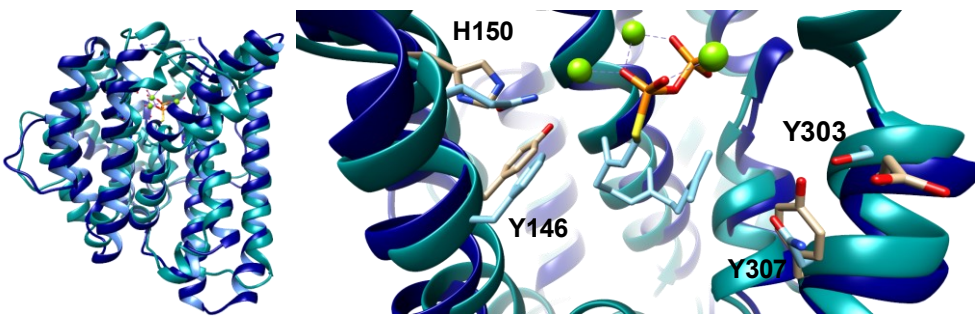


Figure 28. Cartoon representations of the crystal structure of GdolS ([PDB: 5I1U], blue) superposed with the crystal structure of ATAS- FSDP ([PDB: 4KUX], grey). Amino acids under study are labelled.

GdolS-H150Y and Gd4olS-H150C were generated to test the possibility that water binding can be restored with these alternative polar functional groups, also to test the influence of the aromatic chain in GdolS catalysis. Both H150Y and H150C mutants yielded functional GdolS, the only product detected was (-)-germacradien-4-ol (**79**, Figures 29 and 31). Interestingly, H150C had a dramatic loss of catalytic efficiency, while H150Y was only around 10-times less efficient than the wild type (Table 7). This implies that H150 may govern cation-π stabilization of a putative carbocation. To test this, H150R was generated, which resulted in a very inefficient enzyme (catalytic parameters could not be measured), Figure 32 and Table 7. H150F and H150W were also created. Both enzymes were functional GdolS (Figures 30 and 32), with H150W having a drastic detrimental effect on the catalytic efficiency, presumably being too large and disturbing the template required for catalysis (Table 7).

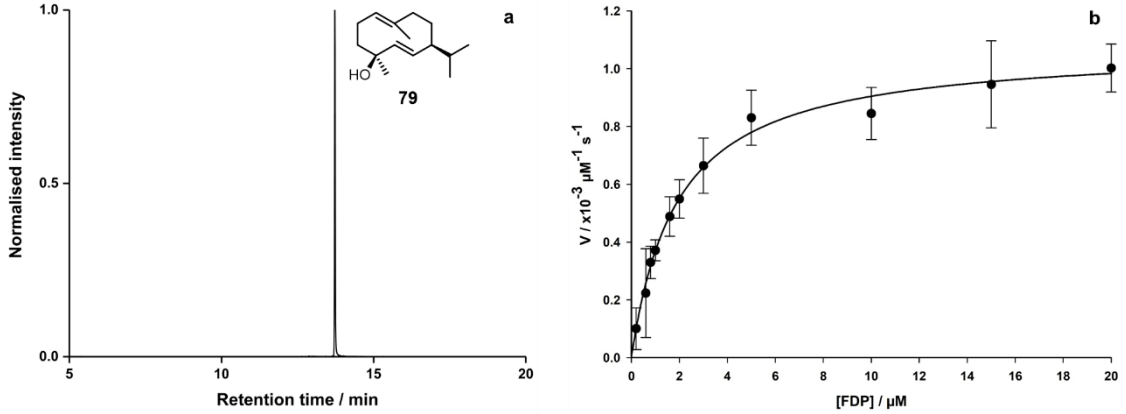


Figure 29. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-H150Y. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-H150Y.

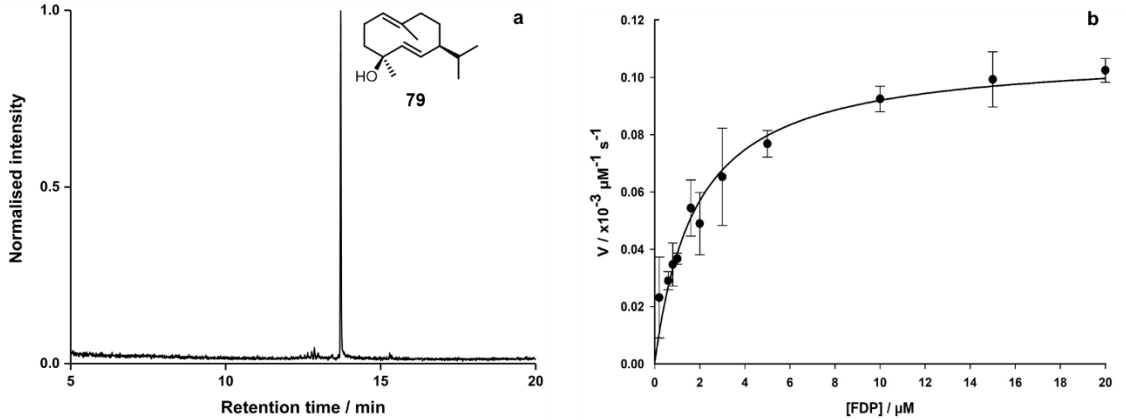


Figure 30. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-H150F. b) Michaelis-Menten graph for the calculation of steady-state kinetic parameters of GdolS-H150F.

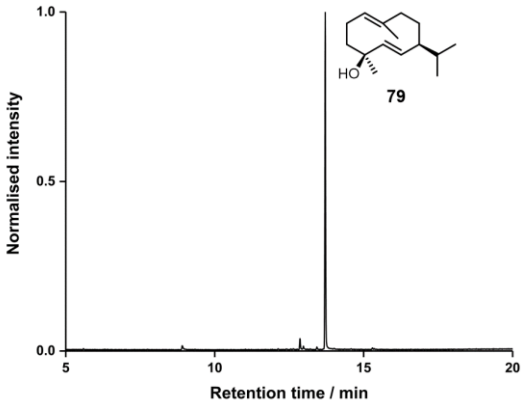


Figure 31. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-H150C.

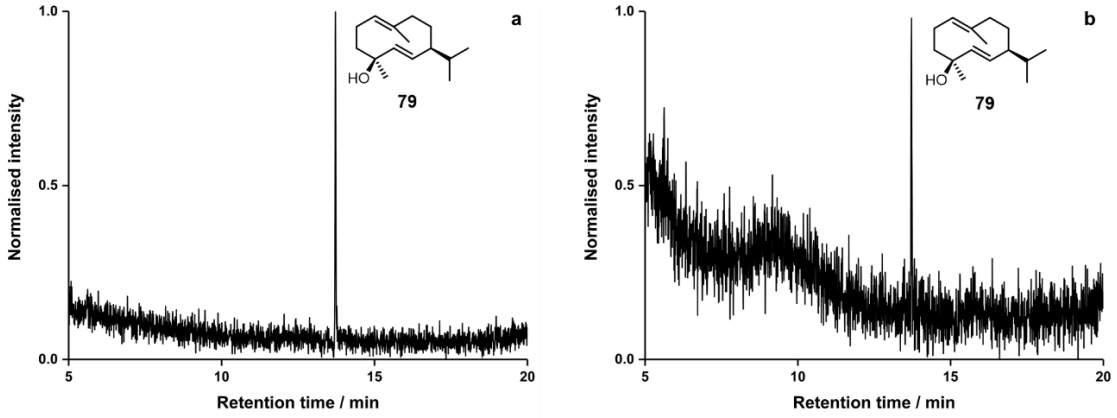


Figure 32. a) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-H150W. b) Total ion chromatogram of the pentane extractable products arising from incubation of FDP (**30**) with GdolS-H150R.

Enzyme	k_{cat} [x 10^{-3} s^{-1}]	K_M [μM]	k_{cat}/K_M [x 10^3 M^{-1} s^{-1}]
WT-Gd4olS	68 ± 1	1.02 ± 0.09	66.7 ± 5.9
Gd4olS-H150Y	11 ± 0.36	1.91 ± 0.20	5.77 ± 0.64
Gd4olS-H150F	1 ± 0.04	1.81 ± 0.21	0.55 ± 0.68
Gd4olSH150W	n/m[a]	n/m[a]	n/m[a]
Gd4olS-H150C	n/m[a]	n/m[a]	n/m[a]
Gd4olS-H150R	n/m[a]	n/m	n/m[a]

[a] No measurable activity; activity too low for kinetic parameters to be determined.

Table 7. Steady-state kinetic parameters for the reactions catalysed by GdolS-WT and GdolS-mutants.

In Summary, GdolS has been expressed and purified, it has been identified the product generated upon incubation with FDP (**30**), and its catalytic parameters has been measured. Apart from these, the active site of GdolS has been examined with the use of crystal structures of homologue enzymes and the influence of several mutations in GdolS catalytic mechanism has been investigated.

The G1/2 helix break of GdolS has been engineered to produce a non-hydroxylating enzyme. A176, which is located in the G1/2 helix break, plays a key structural role that governs product fidelity. Non-polar aliphatic residues modulate the substrate template contour, resulting in a malfunction of water capture. Replacement of A176 with glycine had no impact on the product profile. However, if A176 was replaced with larger non-polar aliphatic residues (valine, leucine, iIsoleucine and methionine) the outcome switched to the production of **32** rather than **79**. The lack of acyclic products suggested a late contribution of A176 to catalysis, which is also supported with kinetic studies. The G1/2 helix break has been postulated as a key region for water access through a channel in other sesquiterpene synthases. To judge whether this is a consequence of a steric barrier to water influx, A176 was also replaced with polar amino acids of varying sizes. Indeed, substitutions A176D, A176T and A176Q resulted in functional GdolS and ruled out this hypothesis. Also, these results differ with the possibility of a steric displacement of an activated water molecule, which otherwise should abolish the nucleophilic addition even if polar amino acids are introduced. The combination of these results suggest that A176 might be in a place near an active site water molecule, which presumably forms part of a hydrogen-network. The 'spectator' role of A176 is compromised upon substitution with larger non-polar amino acids, that break the WT associations and hence result in aberrant proton loss in the carbocation intermediates **62** and **90** (Scheme 23). Furthermore, these associations can be restored with amino acids having hydrogen bonding capabilities, as it is the case for threonine, aspartic acid and glutamine.

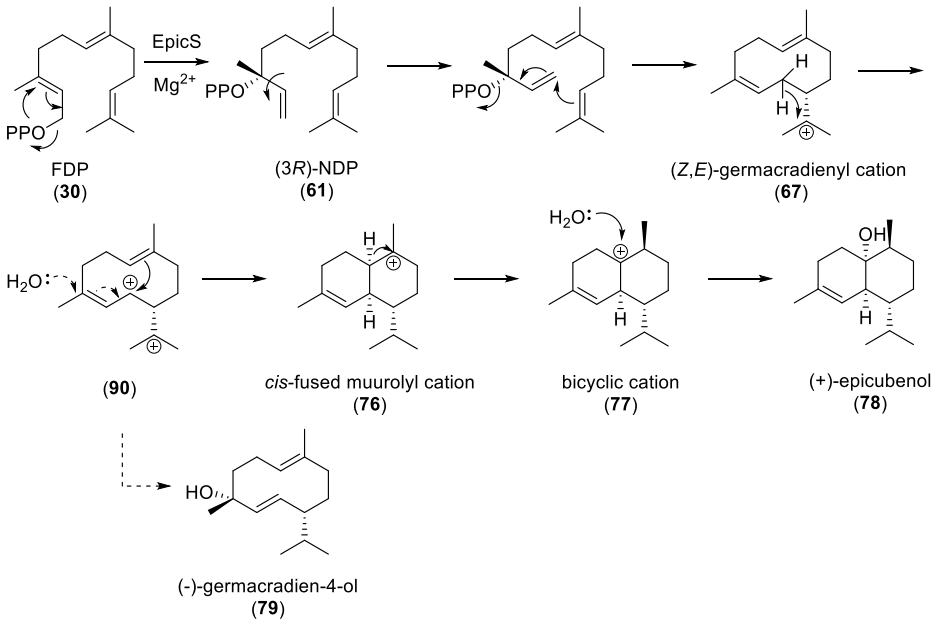
T175 and A177I were also mutated, but the results were not conclusive and further exploration is needed on these to assess their roles.

Lastly, the role of H150 side chain was investigated. The mutations generated here suggest that H150 might be involved in cation-π stabilization function.

2.3. (+)-Epicubenol synthase (EpicS)

The hydroxylated terpene (+)-epicubenol (**78**) has been isolated from plants[195,196] and *Streptomyces* organisms.[197–200] The mechanism and stereochemical course for the enzymatic production of (+)-epicubenol has been extensively studied making use of a purified enzyme, as well as cell-free extract, from the liverwort *Heteroscyphus planus* and *Streptomyces sp-LL-B7* (as detailed in section 1.5.3).[115–119] Nakano *et al.* revealed the gene encoding for (+)-epicubenol synthase (SGR6065) in the genome sequence of *Streptomyces griseus IFO13350*[201] and cloned for the first time.[170]

(+)-Epicubenol synthase (EpicS) from *Streptomyces griseus* is a more promiscuous sesquiterpene alcohol synthase in comparison to GdolS, and other coproducts have been reported, including (-)-germacradien-4-ol (**79**). The proposed cyclisation mechanism to form (+)-epicubenol (**78**) starts with the isomerisation of FDP (**30**) to (3*R*)-NDP (**61**), followed by a 1,10-cyclisation to give (*Z*,*E*)-germacradienyl cation (**67**). Following this, a 1,3-hydride shift takes place to form the allylic carbocation (**90**). Then, EpicS directs a 1,6-cyclisation to form the *cis*-fused muurolyl cation (**76**) and a 1,2-hydride shift to yield the bicyclic cation **77**, which in turn is quenched by water addition to generate (+)-epicubenol (**78**), Scheme 27. Alternative reaction pathways can arise from the generated intermediates (Scheme 28, below).



Scheme 27. Proposed reaction mechanism of the conversion of FDP to (+)-epicubenol (**78**) and germacradien-4-ol (**79**, dashed), catalysed by EpicS.

2.3.1. Characterisation of the wild-type EpiCS

The EpicS encoding gene was cloned into pET-16b vector from its microbial source, which was a generous gift from Prof. Yasuo Onhishi (University of Tokyo). The pET16b-EpicS plasmid was used for the production of EpicS in *E. coli* and the resulting enzyme was characterised according to the protocols described in literature.[170]

Cloning EpicS gene into pET-16b vector

The chromosomal DNA of *S. griseus IFO13350* was used as a template for the PCR amplification of EpicS encoding gene (insert, ~1100 bp) using the appropriate primers and the PCR products were digested with BamHI and NdeI to afford the flanked insert (Section 7.1.7). pET16b-GdolS was used as the source for the desired vector. This plasmid was double digested with BamHI and NdeI restriction endonucleases to yield the desired pET16b acceptor vector (~5600 bp). The products after PCR amplification and digestions were checked on an agarose gel (Figure 33) and the desired fragments were extracted using a QIAquick gel extraction kit protocol (Section 7.1.11). The two complementary linear fragments were ligated using T4 DNA ligase to give pET16b-EpiCS (Figure 33), as indicated in Section 7.1.7.

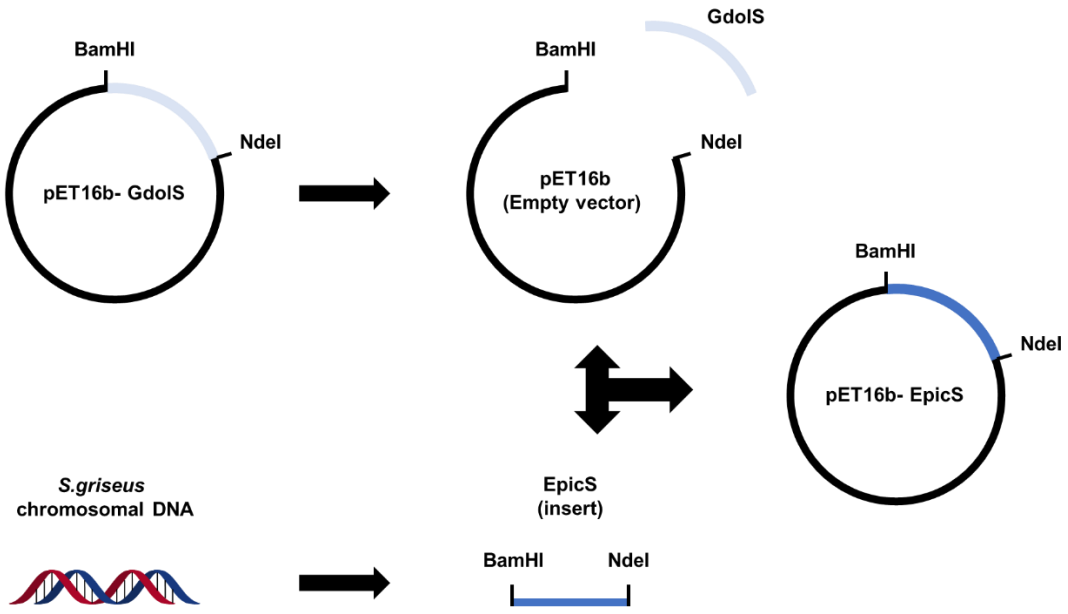


Figure 33. Cloning strategy for the generation of pET16b-EpicS plasmid.

The ligation products were directly transformed into competent *E. coli* XL1-Blue cells and the colonies were grown on agar plates containing ampicillin (100 μg mL^{-1}) for selection. After this, a single colony was used to inoculate LB medium overnight and the new plasmid was extracted using a QIAprep Spin Miniprep Kit protocol (Section 6.1.11). The construction of the new plasmid was first checked by re-digestion with SmaI, which is a restriction site only present in the new plasmid. This afforded a fragment with the expected size (6780 bp, Figure 34). DNA sequencing confirmed the presence of the desired sequence in the correct frame.

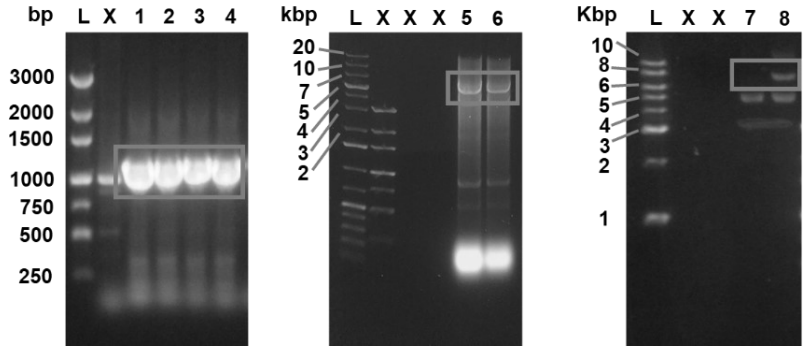


Figure 34. Agarose gels. Left, EpicS insert after PCR amplification (lanes 1-4). Middle, acceptor vector pET-16b after double digestion (lanes 5 and 6). Right, SmaI digestion control of the constructed plasmid (lanes 7 and 8).

Heterologous expression

To determine the optimal conditions for the expression of the EpicS gene, *E. coli* BL21(DE3)-RP cells were transformed with pET16b-EpicS plasmid and test expression were carried out changing the IPTG concentration, the time and the temperature.

Initially, expression was carried out using the same procedure as for GdolS (Section 2.2.1 and 7.1.13). However, after cell lysis we observed that this protocol resulted in an insoluble protein, which was only present in the pellet after cell lysis (Figure 35, lanes 1 and 2).

After this result, it was decided to test longer expression times at lower temperatures with the aim of obtaining a soluble protein. Hence, test expressions were carried out for 16 h at 15 ºC and 25 ºC using different amounts of IPTG (0.1 - 0.3 mM). SDS-PAGE was used to determine the levels of expression (Figure 35, lanes 3-14). All conditions generated soluble proteins after cell lysis, indicating that expressions at reduced temperatures and extended time were better tolerated. Regarding these results, we decided to carried out the expression of EpicS at 16 ºC for 16 h adding 0.1 mM IPTG, which is in good agreement with the conditions reported by Nakano *et al.*[170] After this period, cells were harvested by centrifugation (4 ºC, 3500 g, 10 min) and the pellet was stored at -20 ºC until further use.

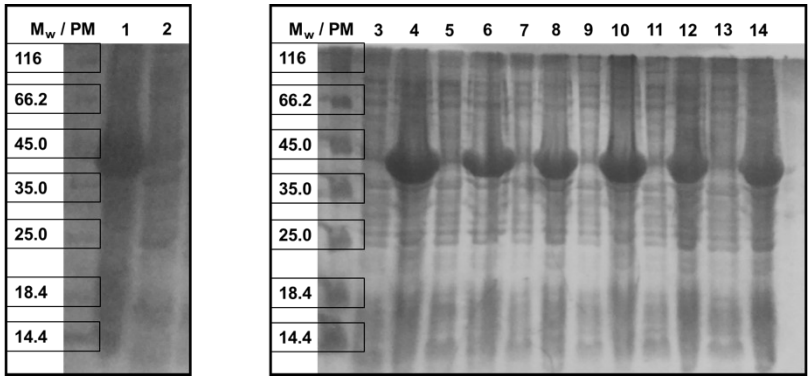


Figure 35. SDS- polyacrylamide gels. Expression test. M_W in kDa. PM: protein marker. 1. pellet after 37 ºC for 4 hours. 2. Supernatant after 37 ºC for 4 hours. 3-4. P and S at 25 ºC with 0.3 mM IPTG. 5-6. P and S at 25 ºC with 0.2 mM IPTG. 7-8. P and S at 25 ºC with 0.1 mM IPTG. 9-10. P and S at 15 ºC with 0.3 mM IPTG. 11-12. P and S at 15 ºC with 0.2 mM IPTG. 13-14. P and S at 15 ºC with 0.1 mM IPTG.

Extraction and purification

In order to purify EpicS, the stored pellet was thawed on ice and resuspended in cell lysis buffer (40 mL, 50 mM Tris, 100 mM NaCl, 5 mM imidazole, 10 % glycerol (v/v), pH 8). Cells were then disrupted by sonication and the resulting suspension was centrifuged. Normally, reducing agents or detergents are used in the cell lysis buffer to avoid aggregation caused by the formation of disulfide bonds between cysteine residues, and to help solubilise proteins. However, EpicS was found in the supernatant under these conditions as judged by SDS-PAGE (Figure 36) and thus the cell lysate supernatant was directly loaded onto a Ni-NTA affinity column. The column was washed with a gradient of imidazole (from 10 to 500 mM), and the fractions were collected and checked by SDS-PAGE (Figure 35). The fractions containing >90 % pure EpicS (molecular weight 43,150 Da) were combined, dialysed for 16 hours in dialysis buffer and concentrated. The colorimetric Bradford assay[185,186] was used to determine the protein concentration.

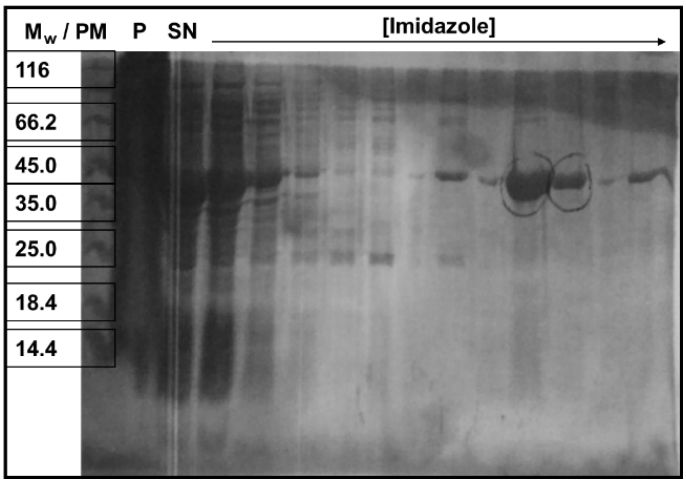


Figure 36.SDS-polyacrylamide gel of samples from different stages of EpicS purification.

In order to further verify the presence of EpicS, the purified was to be subjected to MS analysis. Hence, we buffer exchanged our protein using a VIVA spin column with a 30 kDa cut off into 10 mM Tris-HCl at pH 8.0 and diluted to 3 μM. The MS analysis confirmed the expected mass with the loss of one methionine: $(Mw - Met)_{calc}$ = 43018.92 Da, Mw_{obs} = 43018.5039 Da.

Incubation with FDP and product analysis

Analytical incubations of EpicS with FDP (**30**) were carried out to test the activity of the purified enzyme, as well as to compare the wild-type product profile with those of future mutants. A solution of 1 μM EpicS and 200 μM FDP in incubation buffer [250 μL, 50 mM Tris, 5 mM βME, 5 mM $MgCl_2$, pH 8.0] was prepared. The aqueous layer was overlaid with HPLC grade pentane (0.5 mL) and the resulting mixture was stirred overnight at 25 °C. The pentane extracts were subjected to GC-MS analysis according to materials and methods (Section 7.1.17). Negative controls (no enzyme) were also performed to ensure that analysed products arose from enzyme catalysis.

GC-MS analysis showed a major product, with a molecular ion of m/z = 222 ($[M]^+$), which corresponds to (+)-epicubenol (**78**, Figure 37). In addition, at least 4 more products were detected in the *in vitro* EpicS reaction. The total ion chromatogram (product profile) is in good agreement with that reported previously (Figure 38),[170] although a new sesquiterpene product has been observed (*****, unidentified compound, Figure 44), which may have arisen from an alternative deprotonation along the course of the reaction (Scheme 28). The mass spectrum of the main compound (**78**) is in good agreement with that previously published (Figure 40).[170] The second major product was identified as germacradien-4-ol (**79**) by comparison with an authentic enzymatic sample after incubation of FDP (**30**) with GdolS and compared with the literature (Figure 41).[170] Compounds **122** and **123** correspond to (+)-cubenene and β-gurjunene respectively, by comparison of their mass spectra with those published (Figures 42 and 43).[170]

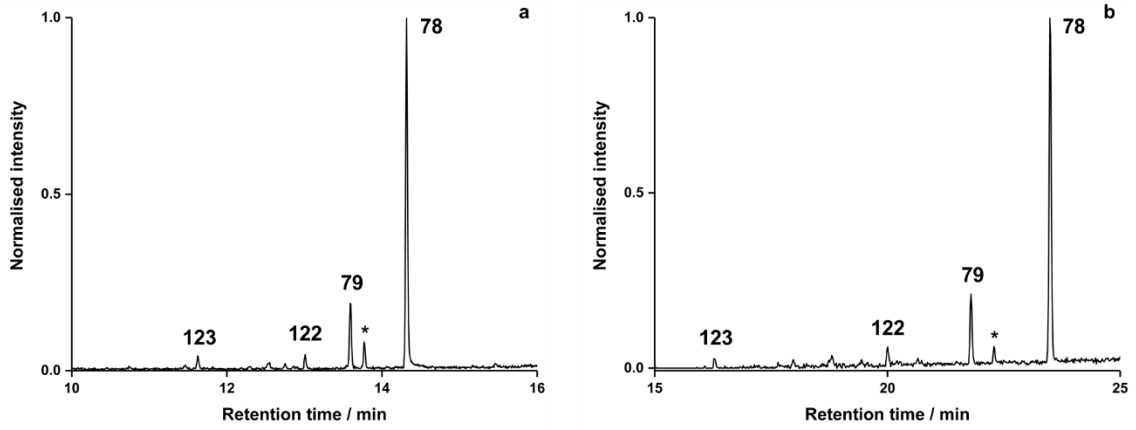


Figure 37. Total ion chromatograms from the extractable products arising from the incubation of FDP (**30**) with EpicS. a) GC-MS method 1, b) GC-MS method 2.

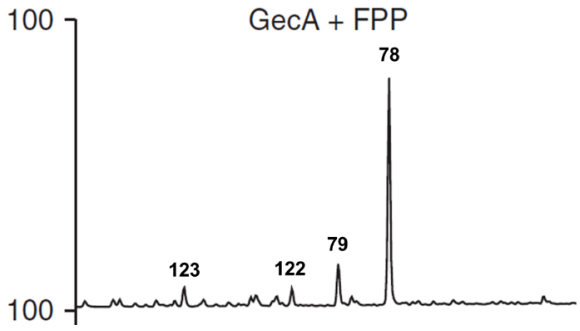
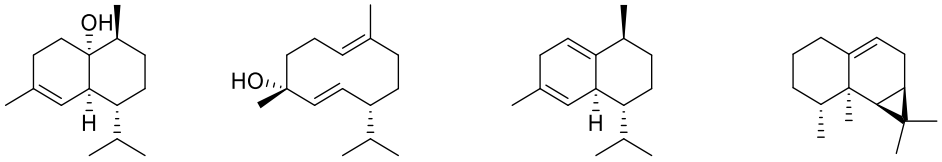


Figure 38. Reported total ion chromatograms from the extractable products arising from the incubation of FDP (**30**) with EpicS, from reference [170].



(**78**) (**79**) (**122**) (**123**)

Figure 39. Structures of the pentane extractable products obtained upon incubation of FDP (**30**) with EpicS.

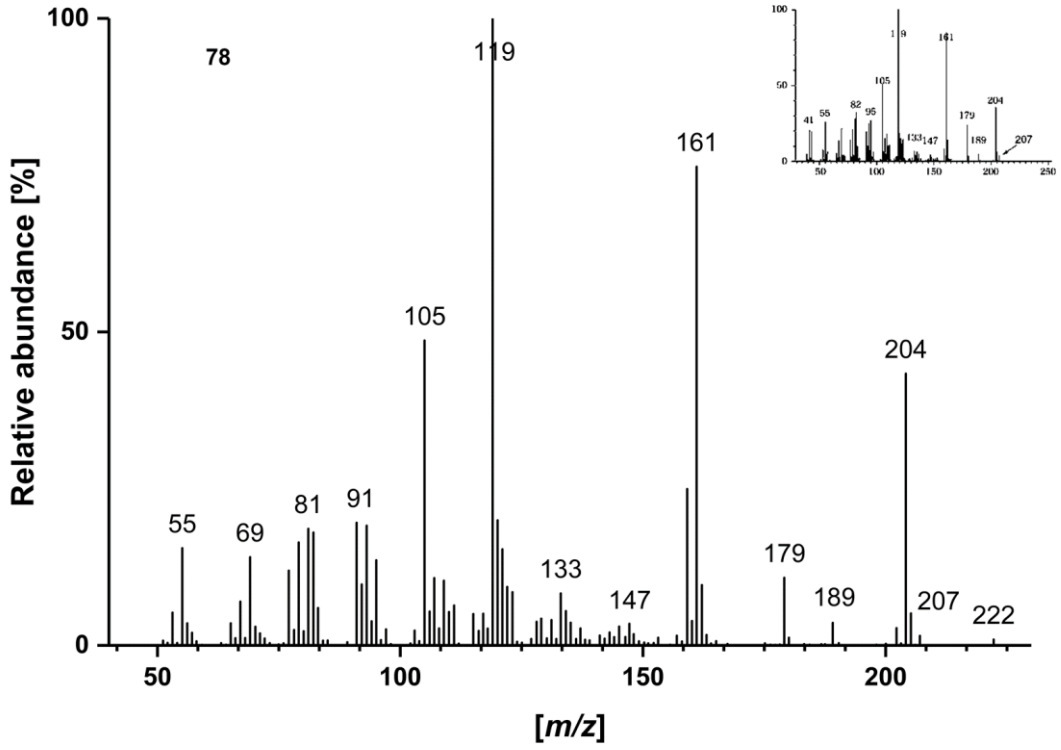


Figure 40. Mass spectrum of compound **78**. Inset, reported mass spectrum of (+)-epicubenol (**78**).[170]

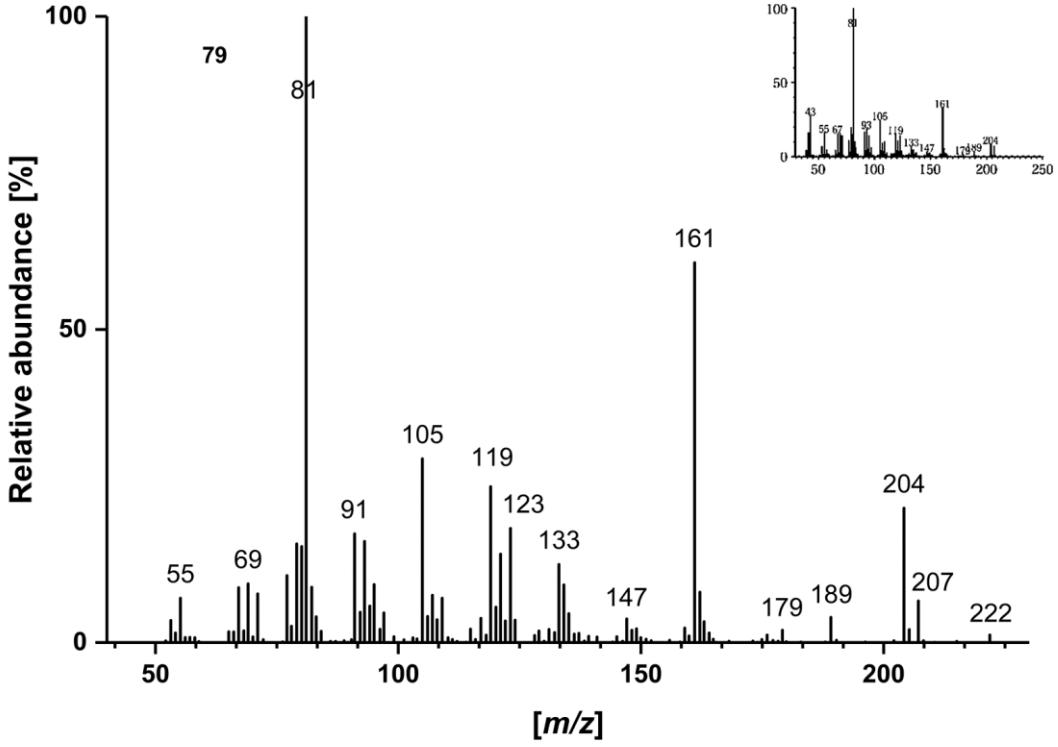


Figure 41. Mass spectrum of compound **79**. Inset, reported mass spectrum of germacradien-4-ol (**79**).[170]

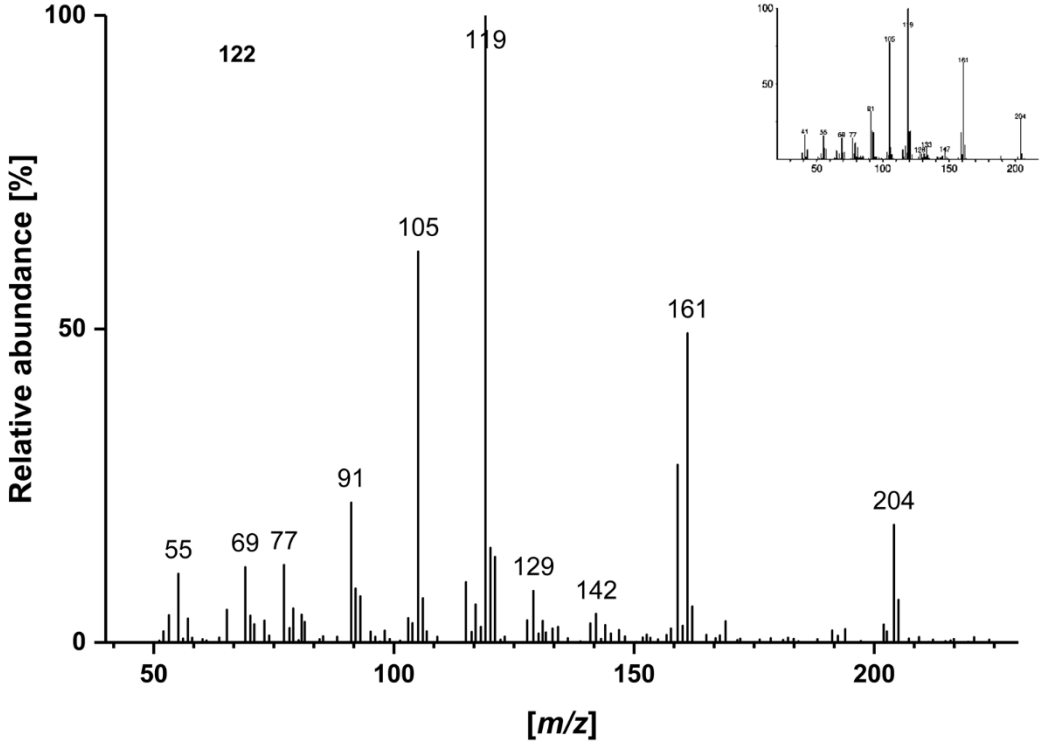


Figure 42. Mass spectrum compound **122**. Inset, reported mass spectrum of (+)-cubenene (**122**).[170]

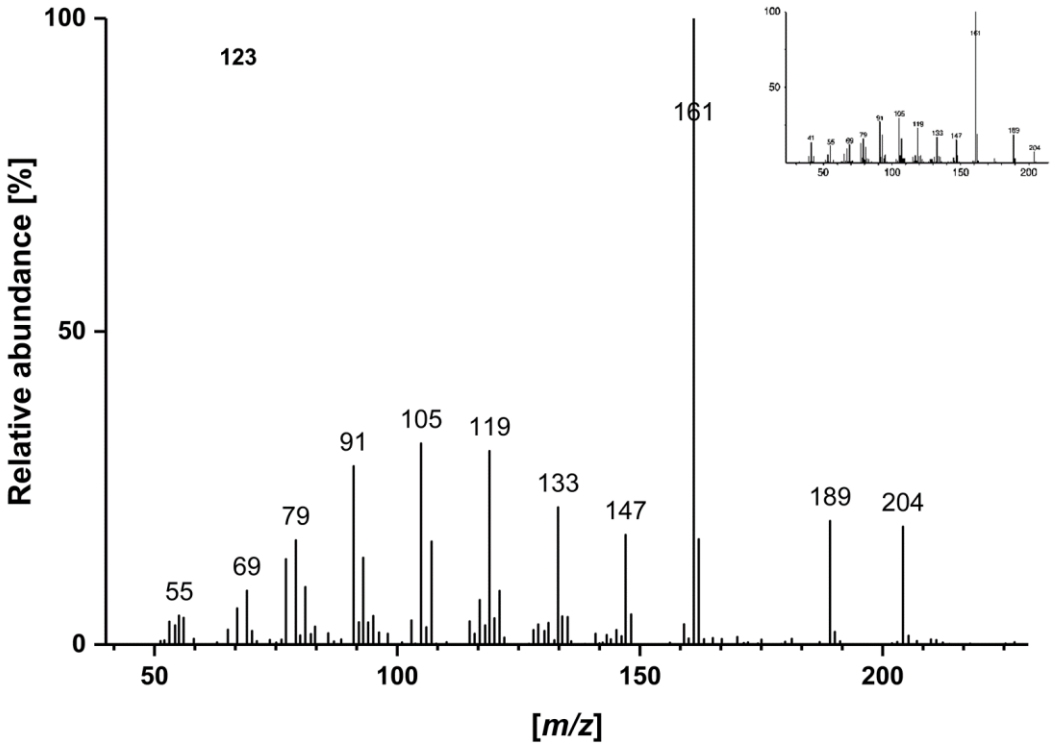


Figure 43. Mass spectrum of compound 123. Inset, reported mass spectrum of β-gurjunene (**123**)[170].

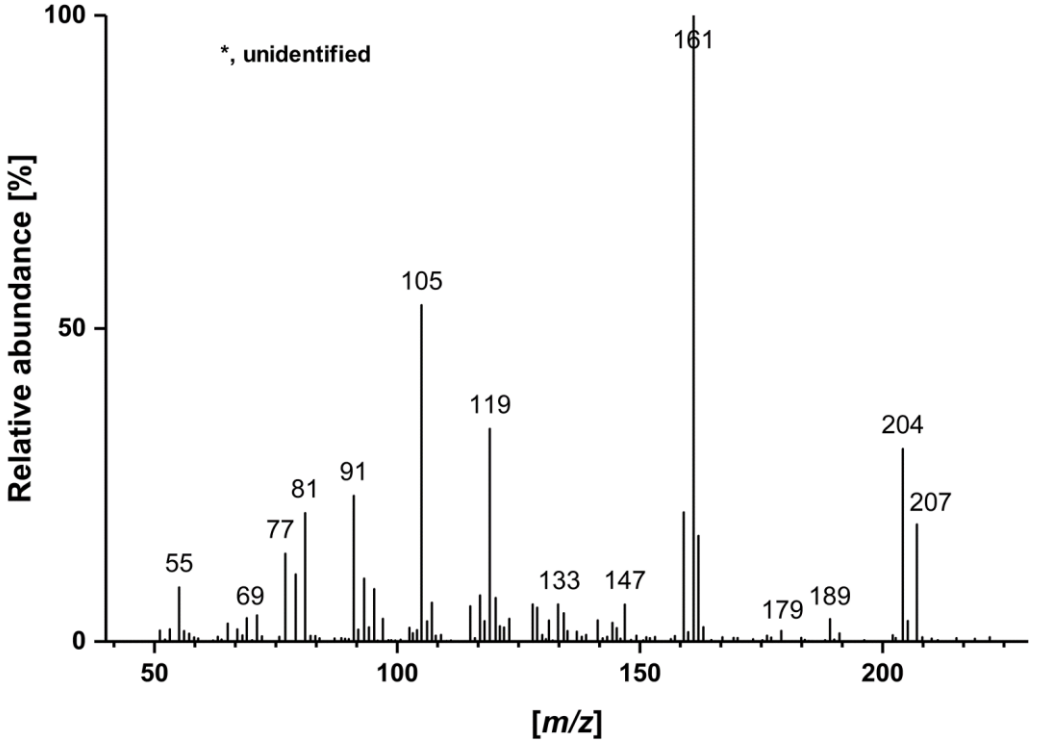
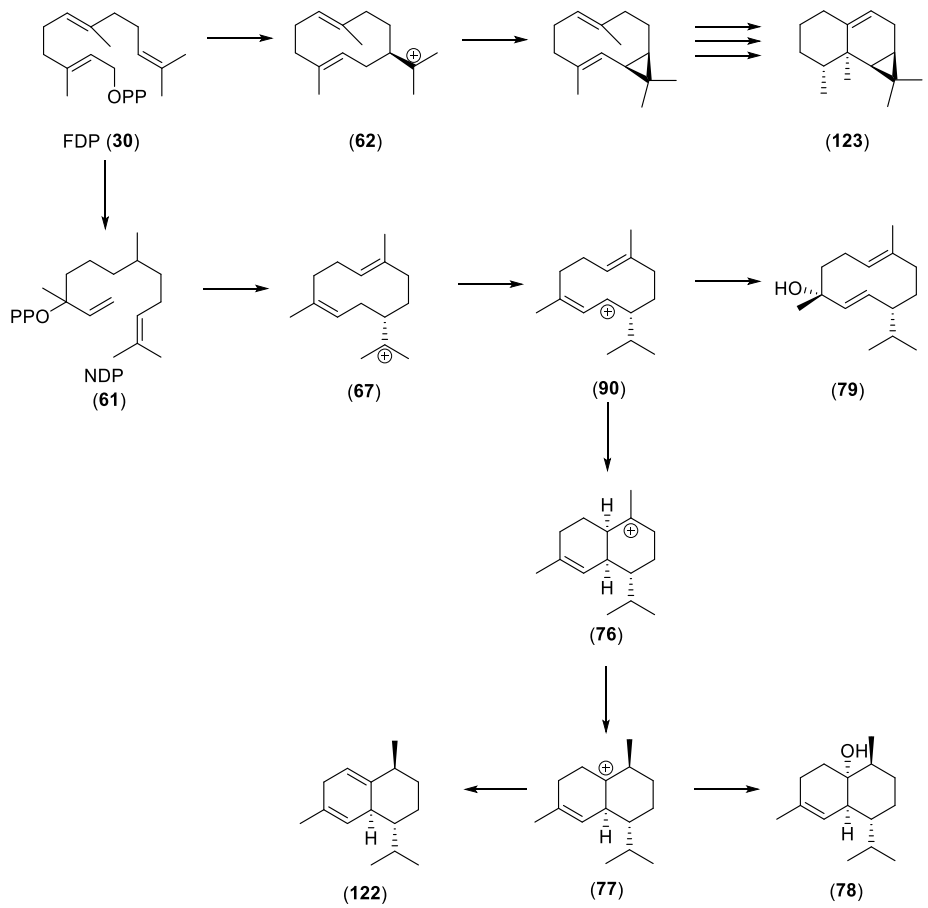


Figure 44. Mass spectrum of the unidentified pentane extractable compound (*****).



Scheme 28. Proposed reaction mechanisms catalysed by EpicS for the conversion of FDP (**30**) to **78**, **79**, **122** and **123**.

Steady-state kinetics

The steady-state catalytic parameters of EpicS were measured according to the standard radiolabelled-substrate procedure commonly used with terpene synthases,[187] similar to the procedure with GdolS (Section 2.2.1) and as stated in methods (section 7.1.19).

Prior to performing steady-state kinetics, the [Mg^{2+}], [enzyme], pH and time conditions were optimised (Figure 45). To investigate the magnesium cofactor concentration, 10 nM EpicS was incubated with 4 μM of [1-^{3}H]-FDP in incubation buffers [50 mM HEPES, pH 8.0] containing a range of [Mg^{2+}], from 0.2-9.0 mM, for 10 min at 30 ºC. The maximum value observed correspond to 1 mM Mg^{2+}, and hence this concentration was selected. For the optimisation of the employed enzyme concentration, a [EpicS] range of 5-200 nM was incubated with 4 μM [1-^{3}H]-FDP for 10 min at 30 ºC in the incubation buffer (50 mM HEPES, 1 mM $MgCl_2$, pH 8). The reaction rate increased linearly up to 150 nM, at which an activity plateau was reached, and which indicates enzyme aggregation at higher concentrations. Therefore, a concentration of 100 nM was chosen, as it is within the linear range and provides satisfactory activity to be measured. Regarding the incubation time, analytical incubations with [1-^{3}H]-FDP (4 μM) and EpicS (100 mM) in incubation buffer (50 mM HEPES, 1 mM $MgCl_2$, pH 8.0) were performed between 0-20 min at 30 ºC. The activity increased linearly up to 20 min; hence 10 min was chosen as general incubation time. As for the optimization of the pH, the pH value for the incubations was varied between 5.0 and 10.0. The rate increased up to pH 8.0 and decreased from 8.0 to 10.0. Consequently, the pH value of 8.0 was chosen for the steady-state kinetics assays.

As a result, the optimised conditions: 100 nM enzyme, 1 mM $MgCl_2$, pH 8.0 and an incubation time of 10 min were used to measure the steady state kinetic parameters of the reaction catalysed by EpicS (Figure 46). With this purpose, a range of [1-^{3}H]-FDP (0.2 μM- 20 μM) was used and the protocol was carried out as described in materials and methods (Section 7.1.19). The k_{cat} and K_M values were determined to be 0.0030 ± 0.0001 s^{-1} and 1.29 ± 0.10 μM respectively.

Nakano *et al.*[168] measured the steady state kinetic parameters using a GC-MS based method[202] and reported k_{cat} and K_M values as 0.026 ± 0.001 and 254 ± 7.1 nM respectively. The k_{cat} reported here is about 10-fold lower, while the K_M value is 5-fold higher. Differences in the steady state parameters based on different methods used have been reported previously.[203] The values measured here are comparable with those of other sesquiterpene synthases measured using the same method.

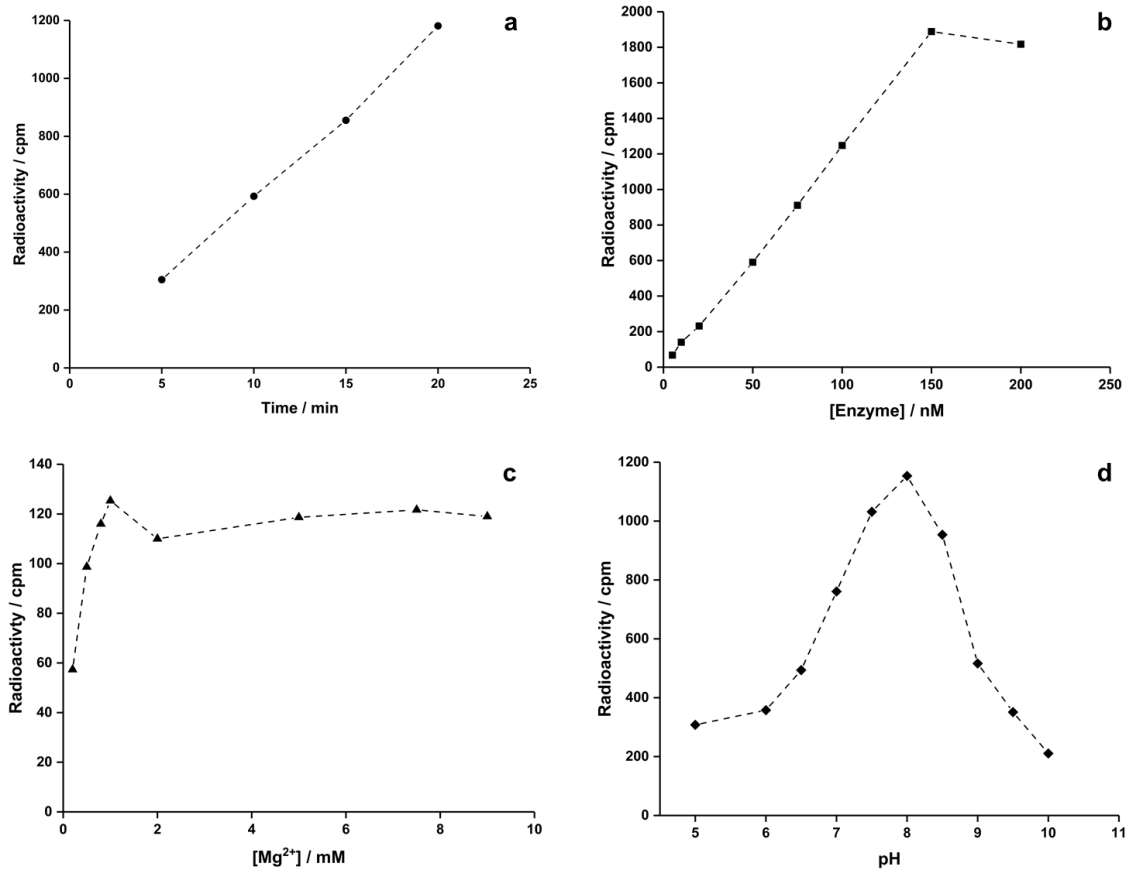


Figure 45. Optimization kinetics conditions for EpicS. Effects of variation of a) time, b) [enzyme], c) [Mg^{2+}], and d) pH on the reaction rate.

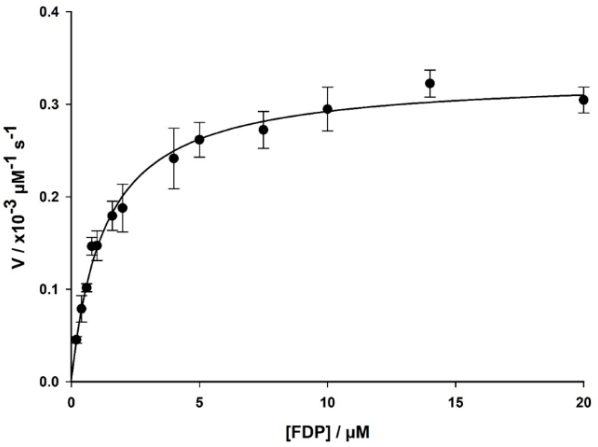


Figure 46. Michaelis-Menten graph for the calculation of steady-state kinetic parameters of EpicS.

Structure of EpicS

Circular dichroism (CD) spectroscopy was used to evaluate the secondary structure of EpicS. CD spectroscopy is a valuable technique for the elucidation of secondary structure elements, such as α-helices or β-sheets, and provides information about the folding properties of proteins.[204,205] The structures of sesquiterpene synthases contain α-helices in antiparallel fashion and thus display characteristic negative bands at 208 and 222 nm and a positive band at 193 nm.[33,69,70]

The CD of EpicS (3 nM) was recorded between 190 and 280 nm at 20 ºC as stated in materials and methods (Section 7.1.16) and showed the typical α-helical spectrum with minima at 208 and 222 nm, and a maximum at 193 nm (figure 47a). EpicS structure was also checked at different temperatures ramping from 5 to 90 ºC and the melting point was determined to be 39.9 ºC (Figure 47b).

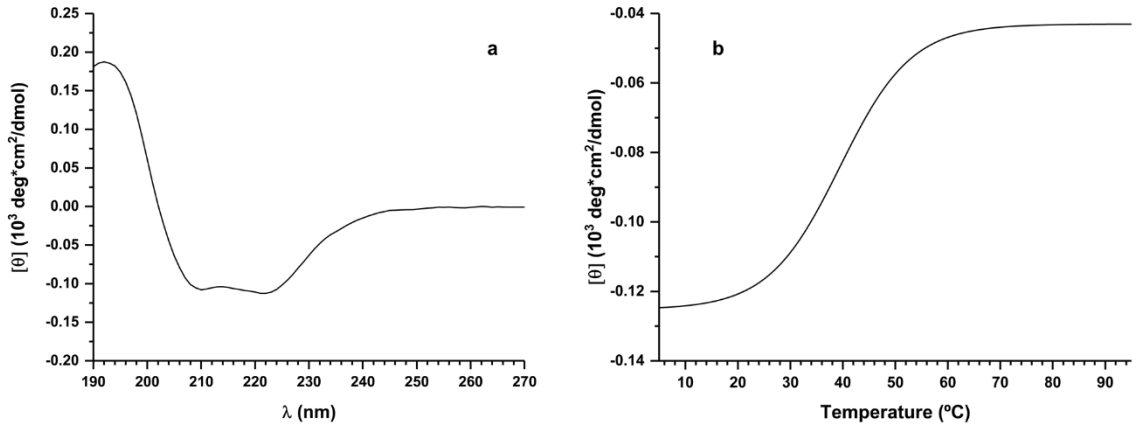


Figure 47. Circular dichroism spectrum of EpicS measured as described in 7.1.16. b) Mean residual ellipticity recorded at 222 nm for a range of temperatures (from 5 ºC to 95 ºC).

Crystal structure

A crystal structure of EpicS has not been solved yet. In fact, only few bacterial sesquiterpene synthases have been crystallised to date.[70,83–85,125]

The crystal structure of pentalene synthase (PS) from *Streptomyces UC5319*[70] was determined in 1997 in its apo state, but it was not until 2010 when the first bacterial sesquiterpene synthase structure in closed conformation was obtained. Epi-isozizaene (EIZS) from *Streptomyces coelicolor A3(2)*[83] was crystallised in complex with 3 Mg^{2+} ions, inorganic diphosphate (PP_i), and the benzyltriethylammonium cation (BTAC). In addition, a structure of EIZS-D99N was obtained in its apo conformation. The comparison of these two structures provided the first evidence of conformational changes required for substrate binding and catalysis in a bacterial terpene cyclase. A few years ago, Dickschat *et al.* solved the crystal structures of hedycaryol synthase (HcS) *from Kitasatospora setae*[84] in complex with nerolidol (**116**, Scheme 24), and selina-4(15),7(11)-diene synthase (SdS) from *Streptomyces pristinaespiralis*[85] in its open and closed (ligand-bound) conformations (Figure 7, Section 1.4.1). These served to elucidate

structural changes governing substrate ionisation and active site closure (Sections 1.4.1 and 2.2.2). Recently, GdolS from *Streptomyces citricolor*[125] was also solved in its apo conformation. In view of this, the aim of this project was to crystallise EpicS to obtain a picture of its tertiary structure and give further details of bacterial sesquiterpene synthases mode of action (if a closed, ligand bound conformation was achieved).

Initially, the crystallographic conditions reported for the mentioned enzymes were adapted and tested (summarised in table 8) for the crystallisation of EpicS. However, none of these conditions proved successful. The presence of his tags in recombinant proteins is a powerful tool for the purification because it allows the selective binding of the expressed protein to a nickel-affinity column. However, literature is inconsistent with the effects of these tags in the crystallisation processes and thus it is a factor to be considered.[206] There may be drawbacks to these tags. For example, they may alter the solubility or increase the aggregation of the purified protein and many crystallographers believe that this effect might hinder crystallisation and thus will cleave the tag before screening for crystallisation conditions and/or perform a second screen after tag removal in proteins that failed to be crystallised at the first attempt.[207,208] In contrast, other surveys have shown that his tags are not necessarily hazardous for the protein structure and might be actually helpful.[209]

Reference	[EpicS]	Protein buffer (A)	Precipitant (B)	Reservoir	A:B
EIZS[83]	10 mg/mL	20 mM Tris-HCl 300 mM NaCl 10% glycerol pH 7.0	100 mM Bis-Tris 25-28% PEG 3350 0.2 M $(NH_4)_2SO_4$ pH 5.5	1 mL of B	1:1
HcS[84]	20 mg/mL	4 M sodium formiate pH 7.0	100 mM Tris 1.6 M $(NH_4)_2SO_4$ pH 8.0	1 mL of B	1:1
SdS[85]	5 mg/ mL	200 mM $MgCl_2$ 100 mM Tris-HCl 24 % PEG 3350 pH 7.8	100 mM Tris 1.6 M $(NH_4)_2SO_4$ pH 8.0	1 mL of B	1:1
GdolS[125]	5 mg/mL	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.8 mM KH_2PO_4 5% glycerol pH 7.4	0.1 M Tris 2.0 M $(NH_4)_2SO_4$ pH 8.5	80 µL of B	1:1

Table 8. Conditions tested for the crystallisation of apo-EpicS.[83–85,125]

In view of these results, it was decided to introduce a tobacco etch virus (TEV) protease cleavage site sequence in between his tag and the start codon of the EpicS encoding gene by Gibson assembly technique (Figure 48, and as stated in methods 7.1. 9.).[210] TEV protease is a highly specific cysteine protease that recognises the amino acid sequence Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser) and cleaves between the Gln and Gly/Ser residues. This insertion was confirmed by DNA sequencing and allowed us to remove his tag, which was thought to prevent artificial conformational changes in the protein.

For the cleavage, EpicS-TEV was expressed similar to EpicS WT. This was followed by Ni^{2+}- affinity column purification and protein dialysis, and the resulting protein was incubated with TEV protease (readily available in our laboratory) for 16 h at 4 °C. After this period, the protein was passed through the Ni^{2+}- affinity column for a second time. This time, the flow through was collected instead of the elution fractions with imidazole. Note that not all the protein was collected in the flow through, which is indicative that only a percentage of the gene was cleaved. The flow through was concentrated and the activity of the enzyme in reaction with FDP (**30**) was checked by GC-MS, which showed no differences with respect to the EpicS-WT.

Figure 48. Section of the DNA sequence of GdolS highlighting designed primers for the TEV cleavage site insertion via Gibson assembly.[210]

Following this, to crystallise EpicS was attended again. This time, it was not only repeated the trials with the conditions mentioned above (Table 8), but available multiscreen condition plates (MORFEO and JBScreen JCSG++) were also used. Among all the conditions tried only one condition gave crystals (Figure 49). Enzyme buffer: 100 mM NaCl and 10 mM HEPES, pH 8.0. Precipitant solution: sodium/phosphate buffer pH 7.2, 10% glycerol and 40% PEG 3350. This condition worked for different concentrations of protein (5, 10 and 15 mg/mL) and different ratios of solution: precipitant (1:1, 2:1, 3:1) for each set of concentrations.

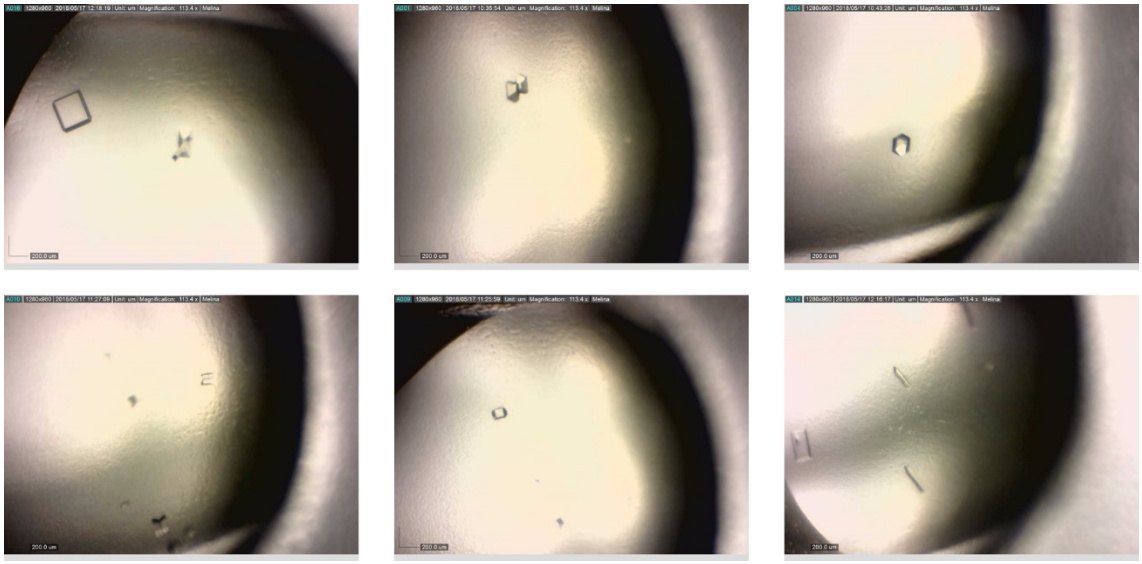


Figure 49. Crystals obtained using enzyme buffer: 100 mM NaCl and 10 mM HEPES, pH 8.0, and precipitant solution: sodium/phosphate buffer pH 7.2, 10% glycerol and 40% PEG 3350, using different concentrations of EpicS in the sample.

Crystals grew within a week and were collected by Dr. Pierre Rizkallah (School of Medicine, University Hospital Wales). The data was collected by Edvardas Kalvaitias (PhD student in the Allemann group, Cardiff University) using synchrotron radiation at the Diamond Light Source (London, UK). Unfortunately, the results obtained indicated that they were small molecules, not protein.

Further optimizations were done, namely an extra purification of the protein by anion exchange column to reach > 95 % purity (Figure 50). Every attempt, however, proved unsuccessful and it was decided to carry out mutagenesis studies in EpicS based on structure homology models.

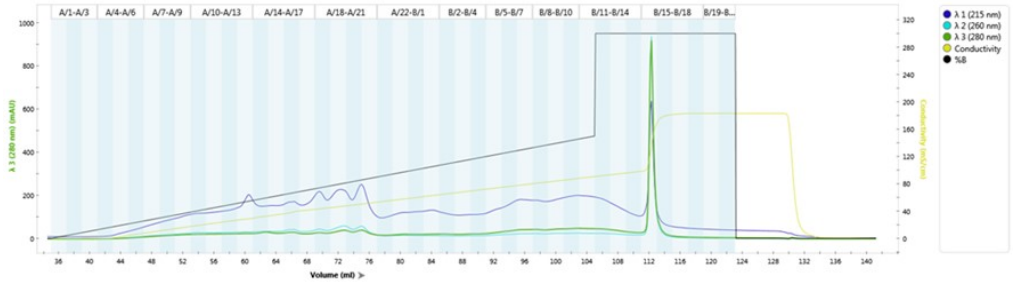


Figure 50. Elution profile of the anion exchange column used for the purification of EpicS. B15 containing purified protein.

2.3.2. Site directed mutagenesis in EpicS

As mentioned earlier, the crystal structure of EpicS has not been elucidated to date, and thus the identification of specific amino acid residues involved in the catalytic mechanism remains elusive. Also, there have not been reported any structure-function studies. Thus, it was decided to employ SDM studies to alter EpicS active site residues based on homology models from related sesquiterpene synthases with the aim of gaining structural details of EpicS mode of action, including (but not only) water management strategies. The expression and purification of EpicS mutants were identical to that of the wild type enzyme (Section 2.3.1), unless otherwise stated.

Metal binding motifs

EpicS contains both conserved metal binding regions, the aspartate rich motif (**D**^{81}**D**QL**D**) and the NSE motif (**N**^{226}DVY**S**FEK**E**), which are responsible for the Mg^{2+} diphosphate-enzyme cluster formation and the ionization of the substrate (Section 1.4.1, structural features). This cluster formation is normally accompanied with the binding of water molecules to the magnesium ions.[178] Alterations to these conserved residues should reveal their importance in catalysis, but could also lead to detrimental defects in the metal-water-enzyme network and result in changing of the plumbing of EpicS.

Initially, the aspartate residues of the **D**^{81}**D**XX**D** region were substituted with glutamate (E), maintaining the functionality but increasing the size of the side chain. EpicS-D81E and EpicS-D85E yielded inactive enzymes and no pentane extractable products were detected. However, EpicS-D82E resulted in a promiscuous germacradien-4-ol synthase (GdolS). The incubation of EpicS-D82E with FDP (**30**) gave germacradien-4-ol (**79**) as the major product, as well as 3 minor sesquiterpene products (Figure 51). One minor product could be identified as (*E,E*)-farnesol (**121**) with an authentic synthetic sample, but the others remained unidentified. While functional, the steady-state kinetic parameters could not be measured under the assay conditions, indicating a severe effect on the catalytic activity.

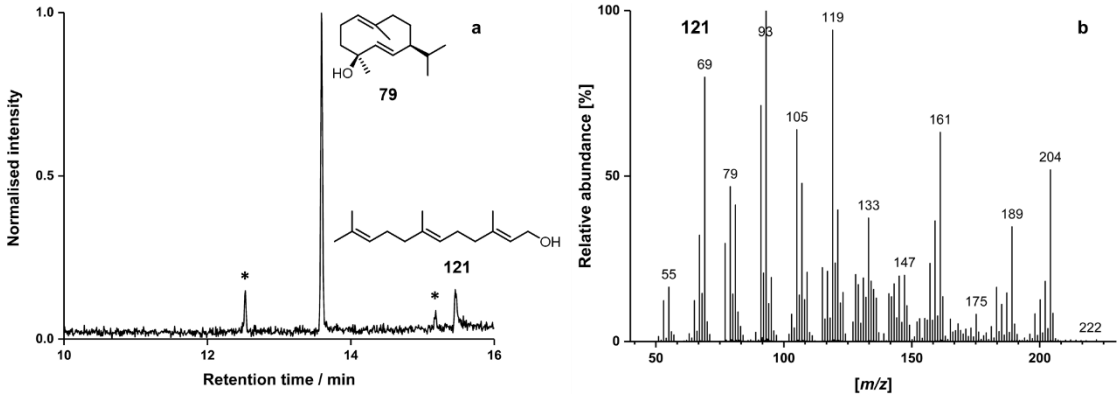


Figure 51. a) Total ion chromatogram of the pentane extractable products arising from the incubation of EpicS-D82E with FDP (**30**). b) Mass spectrum of (*E*,*E*)-farnesol (**121**).

Additional substitutions of aspartate residues to asparagine were made, generating EpicS D81N and D82N. The absence of negative charges resulted in inactive enzymes (Table 9), with no detectable products after incubation with FDP (**30**):

Enzyme	(+)-epicubenol (78)	germacradien-4-ol (79)	(+)-cubenene (122)	β-gurjunene (123)	(*E*,*E*)-farnesol (121)
EPICS-WT	73.8	12.8	3.3	5.5	0
D81E			Not active		
D81N			Not active		
D82E	0	69.3	0	0	16.7
D82N			Not active		
D85E			Not active		

Table 9. Product distribution (%) in the pentane extractable products obtained from incubation of Epic WT and DDXXD mutants with FDP (**30**). Unknown products are not included.

These results are in concordance with previous observations in related terpene synthases, as following. In ATAS, the first aspartate coordinates with two Mg^{2+} atoms (Mg^{2+}_A and Mg^{2+}_C) upon enzyme-substrate complex formation as shown by Shishova[81] and Van der Kamp[82] (Section 1.4.1). In fact, the coordination of this aspartate residue to two magnesium ions is well conserved among terpene synthases (Figure 52).[211] Thus, it is not surprising that the alteration of either the size (EpicS-D81E) or the functional group (EpicS-D81N) severely disturbs the trinuclear cluster formation in EpicS and results in inactive enzymes. As for the second aspartate residue, crystallographic data of related sesquiterpene synthases supports that it forms a salt bridge with a highly conserved active site arginine residue, which in turn forms hydrogen bonds with the diphosphate motif (Figure 52).[73,83,211] While EpicS-D82N is inactive, EpicS-D82E might evoke a defect in the hydrogen bonding network that results in the premature quenching of the intermediates farnesyl cation **60** and the allylic cation **90**, which in turn give alternative hydroxylated products (*E*,*E*)-farnesol (**121**) and (+)-germacradien-4-ol (**79**), respectively. The inactivity of EpicS-D85E is harder to explain. This third aspartate residue also coordinates with the two magnesium ions Mg^{2+}_A and Mg^{2+}_C in the plant sesquiterpene synthase 5-*epi*-aristolochene synthase.[69] However, in the bacterial sesquiterpene synthases *epi*-isozizaene[83] and trichodiene[73] this third aspartate is shown pointing away from the active site. In the bacterial pentalene synthase D84E mutant only a three times lower catalytic activity was observed, while the corresponding first and second aspartate residue mutants (D81E and D82E) resulted in severe reduction of the catalytic activity.[212] In EpicS, mutation D85E might result in a steric hinderance of the Michaelis complex formation.

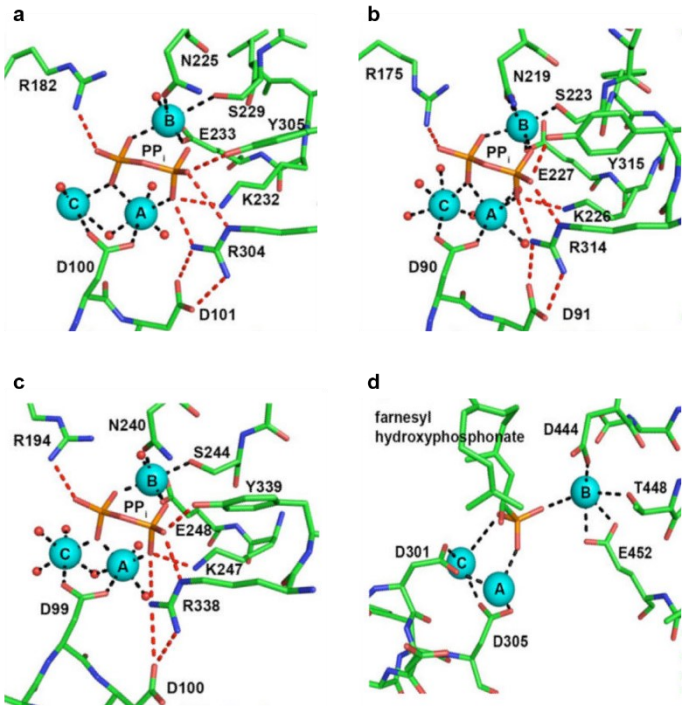


Figure 52. Representation of the trinuclear metal cluster in terpenoid synthases. a) trichodiene synthase from *F. sporotrichiodes*. b) aristolochene synthase from *A. terreus*. c) epi-isozizaene synthase from. *S. coelicolor*. d) 5-epi-aristolochene synthase from *N. tabacum*. Images from reference [211].

Mutations in the **N**^{226}DVY**S**FEK**E** motif of EpicS were also introduced. The residues mutated here are very well conserved among fungal and bacterial terpene synthase, while in plant terpene synthases the motif appears as **D**XX**T**XXX**E**, but in both scenarios they are found binding to one magnesium ion (Mg^{2+}_B). Replacement of asparagine with glutamine in GdolS (GdolS-N218Q) yielded a functional germacradien-4-ol/germacrene A synthase, producing a 1:1 mixture of germacradien-4-ol (**79**) and germacrene A (**32**).[125] In analogy, it was decided to generate EpicS-N226Q, as well as EpicS-N226L, which eradicates the hydrogen bonding capability. Both mutants resulted in inactive proteins. Mutation of S230 and E234 also yielded inactive enzymes. Even when small replacements were made (S230T and E234D), no pentane extractable products were observed by GC-MS analysis after incubations with FDP (**30**). In agreement, substitution of S230 with alanine did not produce any enzymatic product after incubation with FDP (**30**).

G1/2 helix break

Following the studies on the G1/2 helix break of GdolS (section 2.2.2), the influence on catalysis of the amino acids allocated on this region of EpicS was investigated.

The G1/2 helix of EpicS is framed between E173 and G198 amino acids and contains the mentioned helix disruption comprised by amino acids G184 and A185 (Figure 53). The amino acids R180, T183 and G184 are the corresponding so-called sensor, linker and effector in SdS (R178, D181 and G182).[85] It was hypothesised that these amino acids have a major role in EpicS catalysis.

G-helix: SE^{173}ETYIAKRRHTGAIHVCMDLIEIVA^{197}G

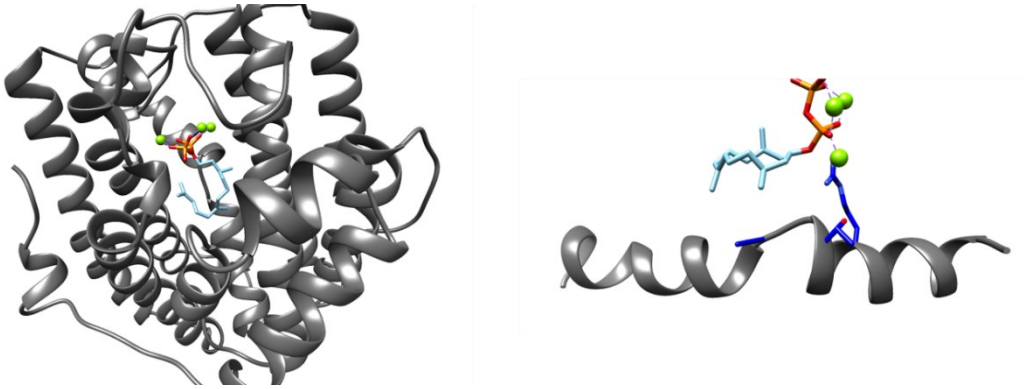


Figure 53. Left, cartoon representation of the homology model of EpicS based in SdS [PDB: 4OKZ][85] (dark grey) in complex with FHDP, docked from SdS [PDB: 4OKZ]. Right, cartoon representation of the G-Helix of EpicS-WT (dark grey) in complex with FHDP, docked from SdS [PDB: 4OKZ]. The cited amino acids are in blue.

Arginine-R180 is strictly conserved among bacterial and fungal terpene synthases and forms part of a highly conserved hydrogen- bonding network. The corresponding residues in EZIS,[83] ATAS,[80] TS[73] are R194, R175 and R182 respectively, and are shown pointing towards the PP_i to donate hydrogen bonds (Figure 52 above). Also, in PRAS,[94] SDM studies probed the role of the respective R200 in diphosphate activation, and presumed to participate in the protonation of germacrene A (Section 1.5.1, SDM studies). In GdolS, the corresponding R172 side chain also proved to be essential for catalysis (Daniel Grundy, Cardiff University).[213] Based on these studies, it was hypothesised that replacement of R180 in EpicS for lysine and histidine, which can be also involved in these proton transfer processes, could modify these interactions and turn to alternative products. Both mutants were prepared by standard SDM methods mentioned before and detailed in materials and methods (Section 7.1.8.). EpicS-R180K and EpicS-R180H were inactive enzymes under experimental conditions, with no

pentane extractable products observed after enzymatic incubations. These results suggest that R180 is essential for catalysis in EpicS.

As for threonine-T183 (linker, D181 in SdS and T175 in Gd4olS), it was mutated to serine with the aim of changing the PP_i-enzyme relationship, but without eradicating the hydrogen binding capability. This substitution was not tolerated by EpicS and led to an inactive enzyme.

Finally, G184 (D182 in SdS, effector; A176 in GdolS) was also mutated to G184I, because in GdolS the substitution at A176 with non-polar larger amino acids led to alternative product profiles. The resulting enzyme was mostly insoluble, nevertheless after basic extraction we obtained a portion of soluble protein that generated (*E,E*)-farnesol (**121**). The inability of EpicS-G184I to isomerise FDP (**30**) to NDP (**61**) and to force the 1,10-cyclisation suggests that this substitution might lead to a steric disruption of the enzymatic Michaelis-complex leading the initial farnesyl cation to be water quenched by water to form **121**. Due to these results, it was decided not to continue the SDM studies in this region.

2. 4. Environmental modulations (divalent cations and pH)

The product selectivity in terpene synthases is mainly determined by the conformational freedom and fit of the substrate inside the active site pocket. This can be modified by engineering the protein (Section 2.2 and 2.3), but also by changing the reaction conditions.[92]

Metal cofactors

Metal cofactors are fundamental for catalysis in terpene synthases because divalent ions are required to form salt bridges with the diphosphate motif and initiate the catalytic mechanism, as shown in the binding mechanism mode of these synthases described in Section 1.4.1. Experimental data shows that sesquiterpene synthases are generally more efficient when using Mg^{2+} but can also make use of Mn^{2+} at low concentrations.[92,214–216] In addition, several sesquiterpene synthases accept different divalent metal cations, such as Co^{2+}, Ca^{2+} or Ni^{2+}.[217–219] In fact, the use of noncanonical metal cations can alter the product specificity observed with Mg^{2+} and thus represents an approach to redirect reaction pathways.[91,218,220,221]

Regarding EpicS, the catalytic activity using different divalent metal ions to Mg^{2+} (1mM chloride salt) has been reported. The catalytic efficiency using Mn^{2+} cofactor was about 10 times slower than with Mg^{2+}.The presence of Fe^{2+}, Zn^{2+}, Cu^{2+} or Cu^{2+} inhibited the catalytic activity (1 mM chloride salt).[170] The product profile after incubation of EpicS with FDP in the presence of Mn^{2+} has not been reported. Thus, the incubation of EpicS with Mn^{2+} was tested (using 0.1 and 1 mM chloride salt). The extractable products were identical as compared with the use of Mg^{2+}.

pH- dependent product profile

The pH dependency of the reaction catalysed by EpicS and GdolS were also examined. The product profile of GdolS did not change within a pH range from 6.0 to 10.0, generating exclusively germacradien-4-ol (**79**). In contrast, with EpicS a dramatic effect was observed on the product spectrum when altering the pH of the reaction. Under alkaline conditions EpicS becomes a selective germacradien-4-ol synthase, thus the reaction ends with a nucleophilic attack of water at the intermediate **90** to form only germacradien-4-ol (**79,** Scheme 27) , being detrimental to a 1,6-cyclisation that would allow the production of (+)-epicubenol (**78**). The trend observed is consistent, with the quantity of (+)-epicubenol (**78**) and the quantity of germacradien-4-ol (**79**) decreasing and increasing respectively in a linear fashion from pH 8.0 up to pH 10.0 (Figure 54).

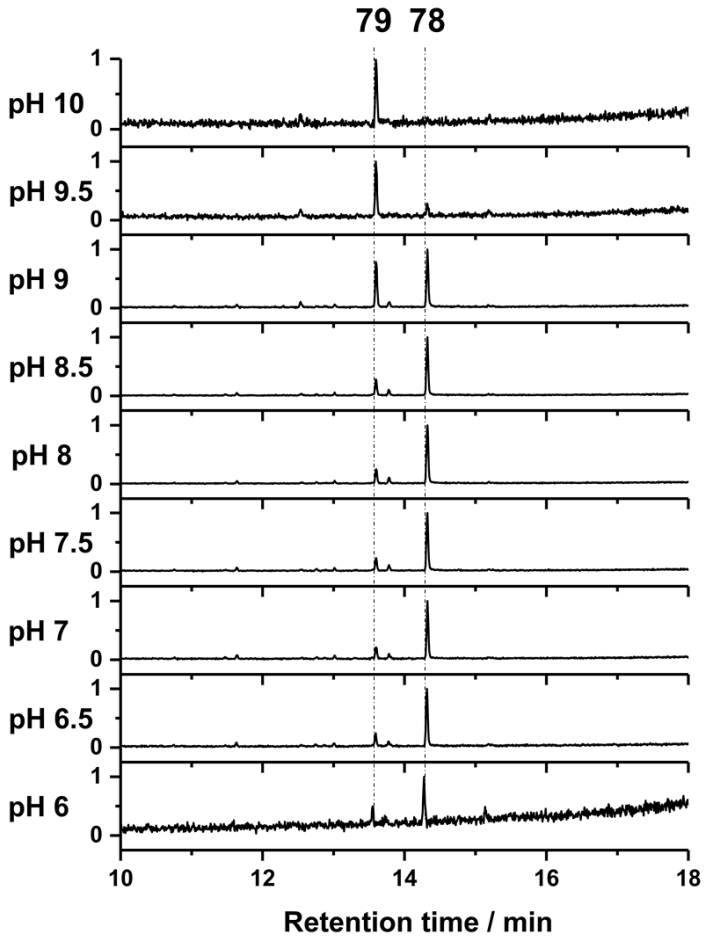


Figure 54. Total ion chromatograms of the pentane extractable products arising after incubation of FDP (**30**) with EpicS across a pH range 6-10.

To test whether these results are originated from a misfolding of EpicS under the buffer conditions, the CD spectra of EpicS were collected at pH range 6-10 (Figure 55). There was no significant change to the spectrum of EpicS at different pHs. In addition, the CD spectra of EpicS were collected different at times in pH 6, 9 and 10, and the protein showed to be stable over 60 min at these pHs. These results clearly rule out a loss of structure as the reason for the observed alteration of the product profile. Alternative, it may be that the stronger nucleophilicity of bulk water is competing with the nucleophilicity of the C6-C7 double bond that participates in the 1,6-cyclisation of the double ring, leading to premature quenching of the reaction.

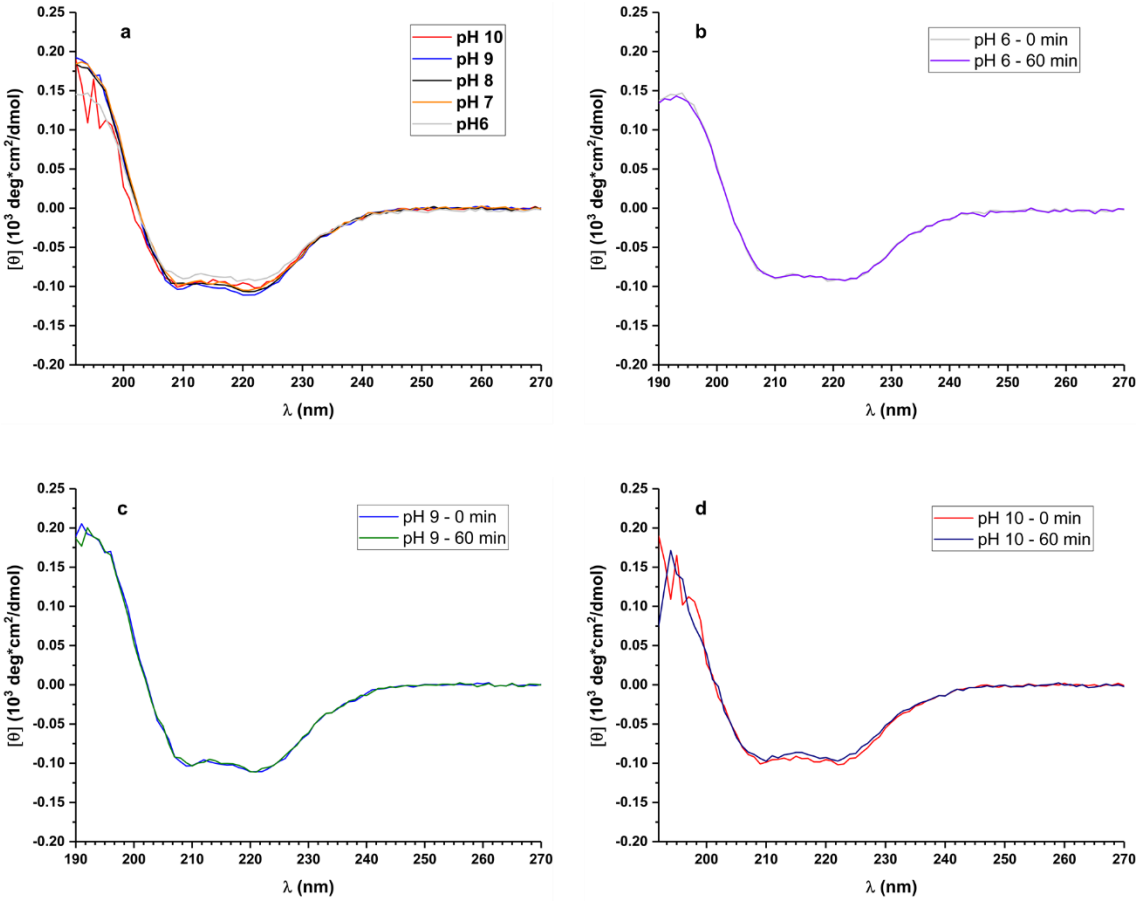


Figure 55. a) CD spectra of EpicS at different pH values (from pH 6.0 to 10.0). c) Spectra recorded at 0- and 10-min intervals in pH 6.0. c) Spectra recorded at 0- and 10-min intervals in pH 9. d) Spectra recorded at 0- and 10-min intervals in pH 10.0.

In summary, EpicS has been cloned and expressed successfully. The products obtained after incubations of EpicS with FDP (**30**) has been analysed, also the steady-state kinetic parameters have been determined. The secondary structure of EpicS has studied by means of circular dichroism, but the crystal structure could not be obtained. Apart from this, the active site of EpicS has been explored with the assistance of the crystal structures of related sesquiterpene synthases and functional mutations has been produced. From these, substitutions in the metal binding motifs severely affected the catalytic activity of EpicS. Apart from EpicS-D82E, all mutants resulted in inactive enzymes. EpicS-D82E converts EpicS into a functional GdolS, generating about 70% of germacradien-4-ol in addition to three minor products. One of the minor products is (*E*,*E*)-farnesol (**121**), and although the others could not be identified, one of them is predicted to be an extra hydroxylated compound because of the retention time displayed. The substitutions in the G1/2 helix break region also resulted in inactive enzymes (with the exception of EpicS-G184I that generated **121**), suggesting a crucial role for catalysis of the amino acids substituted. Lastly, the reactions catalysed by EpicS has been tested in different conditions. Notably, the product profile of EpicS is pH dependant, and the wild-type enzyme is converted to a high-fidelity GdolS at pH 10.0.

CHAPTER 3

RATIONAL DESIGN AND SYNTHESIS OF FDP ANALOGUES FOR THEIR USE WITH 1,6- AND 1,10-SESQUITERPENE SYNTHASES

Chapter 3. Rational design and synthesis of FDP analogues for their use with 1,6- and 1,10-sesquiterpene synthases

3.1. Preface

This chapter describes the design and synthesis of FDP analogues chosen to be tested as substrates for a range of 1,6- and 1,10-sesquiterpene synthases. The FDP analogues covered here are classified depending on the modification introduced into the backbone of natural (2*E*,6*E*)-FDP (**30**); including fluorinated, methylated, nor-methylated and sulfur-containing analogues (Figure 56). Overall, the aim was to manipulate the 1,6-cyclisation required for catalysis in 1,6- sesquiterpene synthases, generate new germacrene products or trap alternative carbocation intermediates along the course of the enzymatic reactions. The incubations of these analogues (Chapter 4) will empower the interrogation of the promiscuity exhibited by sesquiterpene synthases and the generation of alternative enzymatic products that might be advantageous for industry.

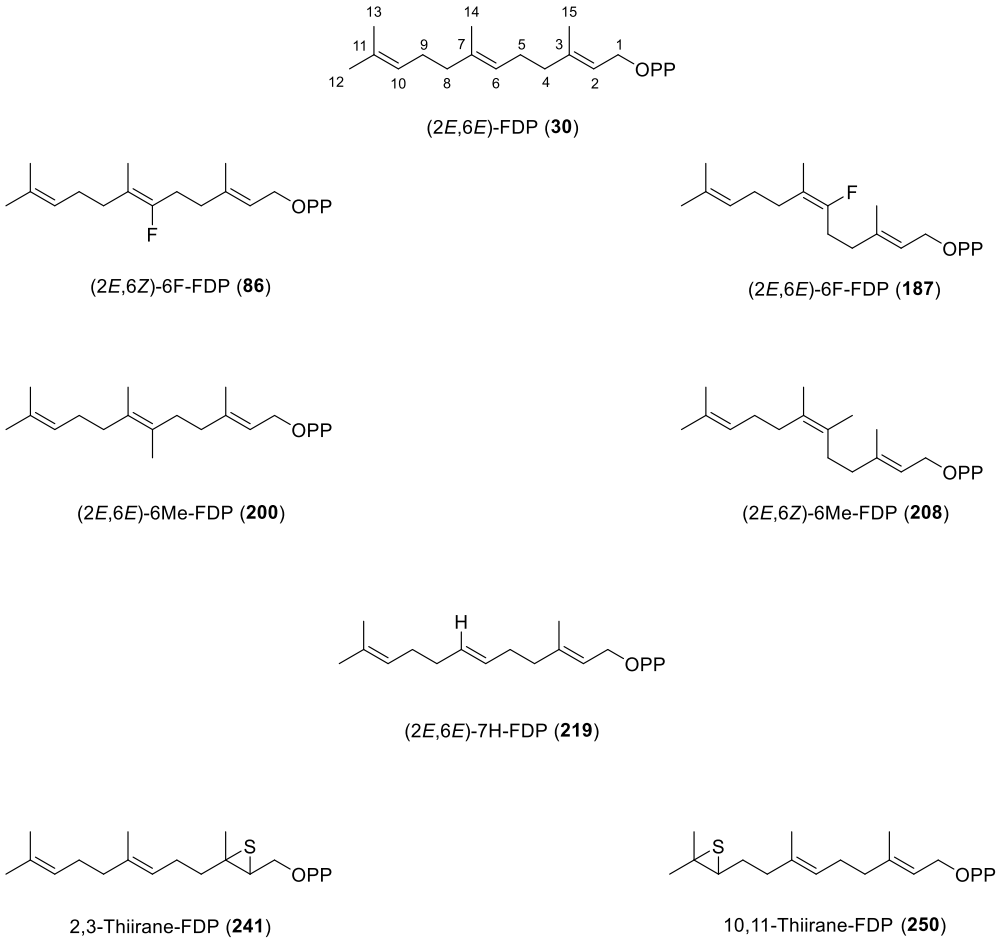
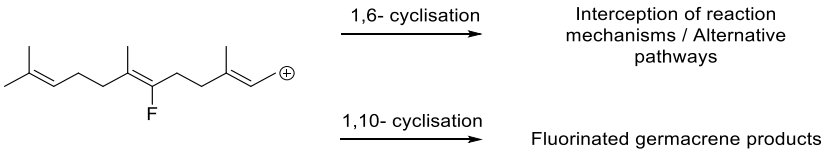


Figure 56. FDP and FDP-analogues targeted for probing sesquiterpene synthases promiscuity.

- The presence of highly reactive carbocation intermediates makes **fluorinated FDP analogues** powerful tools for the study of the enzymatic reactions catalysed by sesquiterpene synthases (Section 1.5.4). The highly electronegative fluorine atoms strongly influence the stability of carbocation intermediates, but have a little steric effect when substituted for hydrogen.[222–224] The rational positioning of fluorine atoms into the FDP backbone can abolish the formation of proposed intermediates and/or diversify the natural enzymatic pathways, thus giving evidence of the existence of key intermediates and sometimes access to new enzymatic products that can amplify or modify the biological activities of the unsubstituted parent natural products (Scheme 29).



Scheme 29. Strategy for the interrogation of sesquiterpene synthases with (2*E*,6*Z*)-6F-FDP (**86**).

- **Methylated FDP analogues** represent a new pool of substrates to access novel terpenoids upon incubation with sesquiterpene synthases. These can have altered biological properties that might otherwise not be easily accessible by chemical synthesis (Section 1.6.1).[123,130,223] The replacement of hydrogen with a bulkier methyl group is often tolerated by sesquiterpene synthases and rational engineering of the active site by site directed mutagenesis can help to improve the substrate specificity.[130]

The methylated FDP analogues previously reported, however, contain the methyl groups on "passive" positions of the FDP backbone, and do not influence the stability of developing carbocations along the cyclisation cascades. The introduction of a methyl group at the C6 position with (2*E*,6*E*)-6Me-FDP (**200**) can affect the enzyme-substrate complex geometry through steric effected and conformational changes, and the effect of changing the C6-C7 double bond conformation in the catalytic mechanism of sesquiterpene synthases remains unknown. For instance, a 1,6-cyclisation mechanism to form bisabolyl (**65**) and cadinyl (**76**) carbocations, Figure 57.

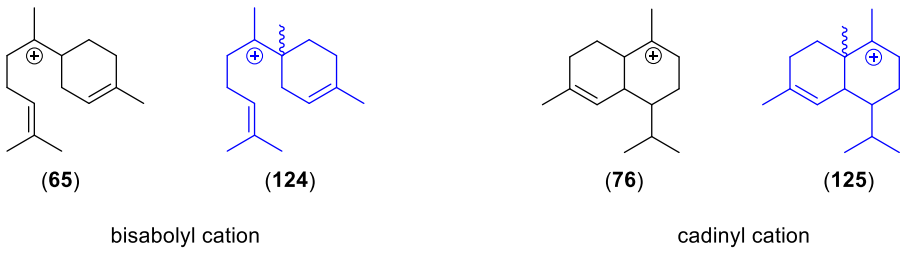


Figure 57. Bisabolyl and cadinyl cations (black), and their corresponding C6-methylated analogues (blue).

- The rationale behind employing **nor-methylated FDP analogue** (2*E*,6*E*)-7H-FDP (**219**) design was the replacement of the methyl group at the C7 position with hydrogen for the destabilisation of emerging carbocations after a 1,6-cyclisation, namely the formation of hypothetical bisabolyl (**65**) and cadinyl (**76**) cations (Figure 58). Decreasing the degree of substitution of a putative carbocation reduces the stability afforded by hyperconjugation, and therefore could alter the reaction cascade. *A priori*, germacrene-type carbocations (arising after 1,10-cyclisation) should not be electronically influenced by this modification, and therefore could be formed.

Another asset to the methyl/hydrogen replacement is the reduction of the steric bulk of the FDP substrate, which could result in a more flexible backbone chain within the active site and thus lead to alternative product formation. For example, the resulting analogue may not be able to adopt the conformation/orientation required to allow initial 1,10- or 1,6-cyclisations and be subjected to early quenching through water addition or deprotonation to the corresponding farnesol/nerolidol or farnesenes respectively. If accepted by sesquiterpene synthases, this novel analogue would lead to the production of new terpenoids.

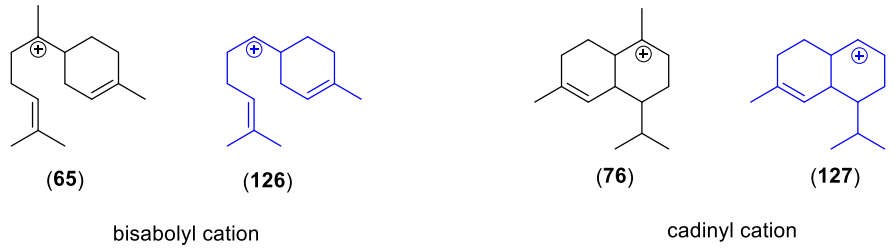


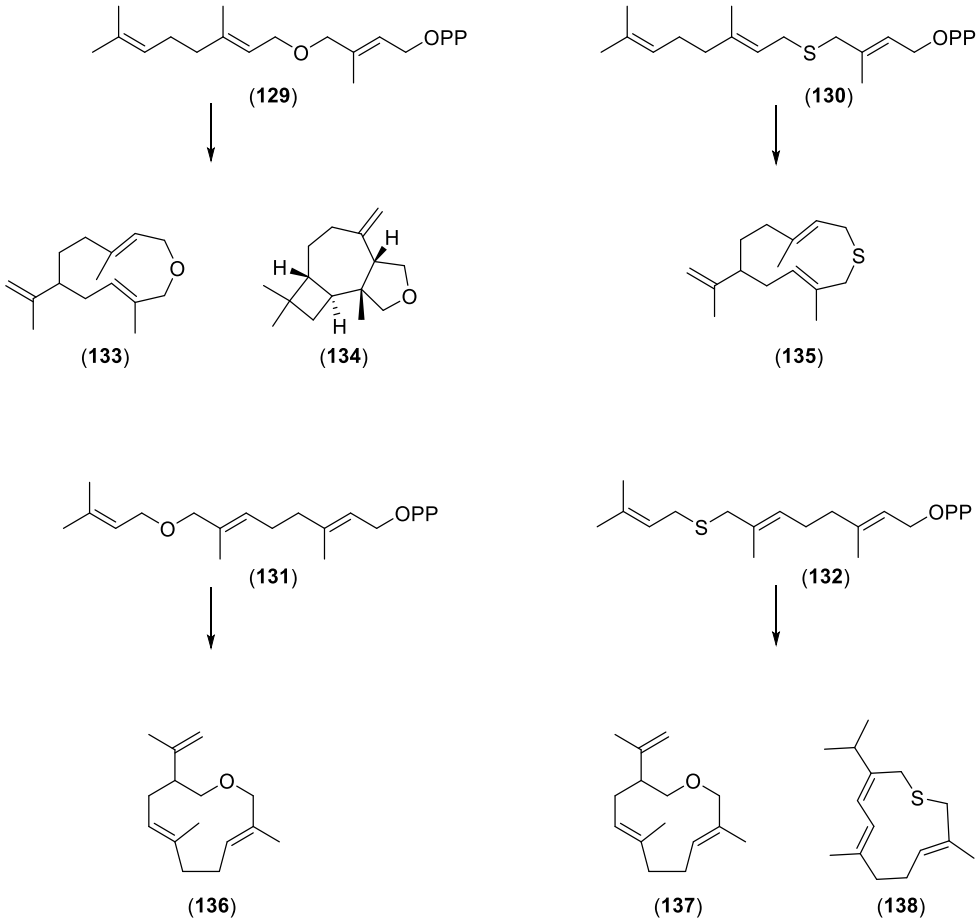
Figure 58. Bisabolyl and cadinyl cations (black), and their corresponding nor-methylated analogues (blue).

- Sesquiterpene synthases are Class I terpene synthases that generally convert (2*E*,6*E*)-FDP (**30**) to a myriad number of 15-carbon skeletons. The all-*trans* configuration of the substrate forces the first intramolecular cyclisation to procced through the remote double bond of **30** (C10-C11, Figure 56). In addition, isomerisation of the C2-C3 double bond can be achieved by migration of the diphosphate group in the substrate that leads to the enzyme-bound tertiary allylic intermediate (3*R*)-NDP (**61**), which allows the nucleophilic cyclisation attack of the neighbouring double bound (C6-C7) and increases the diversity of product formation (Section 1.4.2). This event can be analysed by using the isomeric (2*Z*,6*E*)-FDP (**128**) as a substrate, because it mimics the secondary *cisoid* nerolidyl cation intermediate, (**64**, Scheme 7 in Section 1.4.2).[225–227] In addition, (2*Z*,6*E*)-FDP (**128**) has been used to investigate the substrate promiscuity of various sesquiterpene synthases. While some enzymes generate different products upon incubations with the 2-*transoid* and 2-*cisoid* isomers,[228,229] others retain their catalytic fidelity.[226]

In view of this, it was decided to test for the first time the acceptance of alternate substrate isomers **187** and **208** (Figure 56) by the enzymes under investigation.

- **Heteroatoms** can be inserted into the FDP-carbon chain for the generation of novel terpenoids. Latest research in the Allemann group[132] has probed the versatility of 1,10-sesquiterpene cyclases to convert hydroxy- and epoxy-containing substrates into a novel 10-membered cyclic terpenoid ether by trapping carbocation intermediates (Section 1.6.2). Also, recently, the conversion of oxygen- and sulfur-containing unnatural FDPs (**129**-**132**, Scheme 30) to heteroatom-modified macrocyclic and tricyclic sesquiterpenoids by a range of 1,10- and 1,11-sesquiterpene synthases has been reported (**133**-**138**, Scheme 30).[230] Yet, in this report, the introduced heteroatoms are not actively involved in the cyclisation mechanism and only act as part of the unreactive backbone (Scheme 30).

These works have prompted the exploration of the ability of a range of sesquiterpene synthases to accept thiirane-containing FDP analogues. The aim was to synthesise 2,3-thiirane- and 10,11-thiirane modified FDPs (**241** and **250** respectively in Figure 56) and test the tolerance of 1,6- and 1,10-sesquiterpene synthases upon analytical incubation.



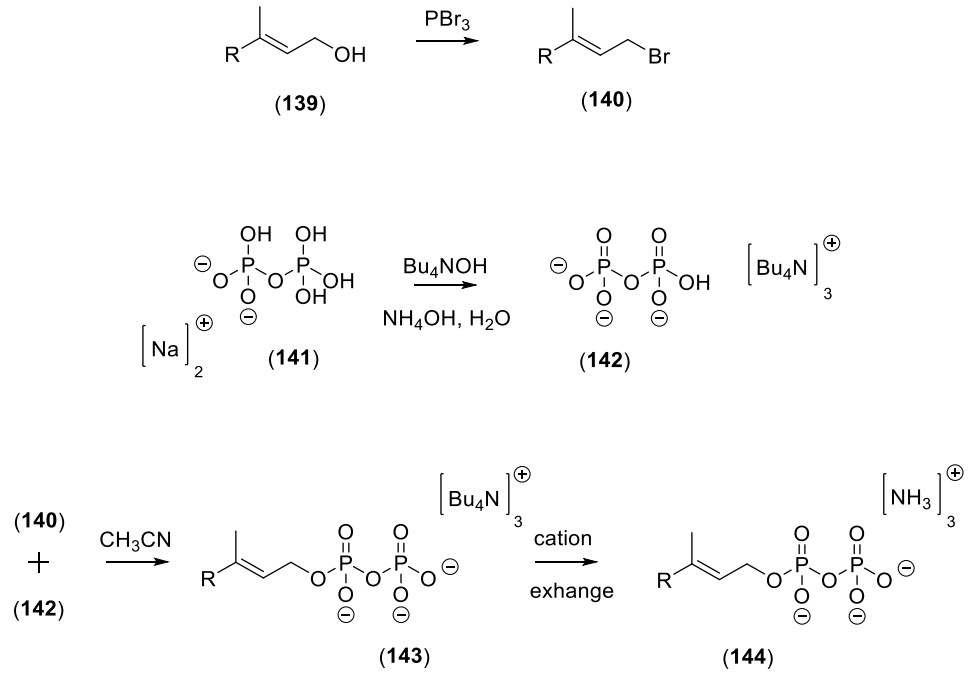
Scheme 30. Heteroatom-containing unnatural substrates and their conversion by range of 1,10- and 1,11-cyclases. Adapted from reference [230].

3.2. Diphosphorylation of FDP and analogues

The synthesis of FDP (**30**) and FDP-analogues (**30**, **86**, **187**, **200**, **208**, **219**, **241** and **250**) started with the preparation of the corresponding farnesol analogues followed by diphosphorylation (Schemes 31 and 32). Two diphosphorylation methods have been used. The first involves the halogenation of farmesol's terminal alcohol and the sequential direct displacement with a pre-made diphosphate salt (Scheme 31), while the second is a one-pot sequence that enables the condensation of the farnesol derivates and inorganic phosphate using trichloroacetonitrile (Scheme 32).

The displacement method used in this work was first described by Poulter *et al.*[231,232] (Scheme 31, method 1). The generated primary allylic alcohol (**139**) was converted into the corresponding bromide (**140**) *via* addition of phosphorous tribromide.[233] The resulting crude bromide was used without further purification and stirred with tris(tetrabutylammonium)hydrogen diphosphate salt (**142**) for 16-24 h in dry acetonitrile at room temperature. After this period, the crude farnesyl diphosphate (**143**) was passed through a column of DOWEX 50W-X8 cation exchange resin (ammonium form) to generate the trisammonium farnesyl diphosphate salt (**144**), which was purified by HPLC. The purified fractions were lyophilised to yield the desired product as a white powder.

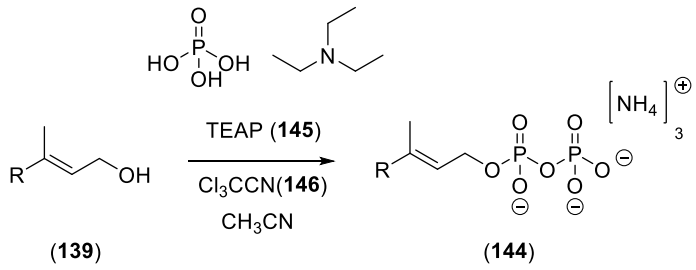
The tris(tetrabutylammonium) hydrogen diphosphate salt (**142**) was prepared in advance. For this, disodium dihydrogen diphosphate (**141**) was passed through a DOWEX AG-50W-X8 cation exchange resin (H^+ form) to form diphosphoric acid, which was immediately titrated to pH 7.3 with 40% (w/w) aqueous tetra-*n*-butylammonium hydroxide. The resulting solution was lyophilised to yield **142** as a white powder, which was recrystallised from ethyl acetate prior to use.



Scheme 31. Diphosphorylation of farnesol analogues, method 1.

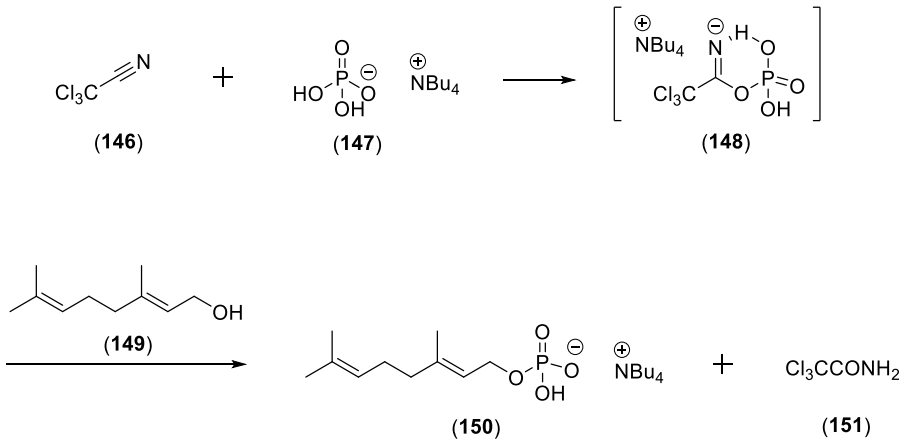
The second approach was first developed in 1959[234] and represents a more direct method for the generation of phosphate monoesters with an activated form of phosphoric acid under mild conditions (Scheme 32, method 2). This pioneering work, which involves the production of mono-, di- and triphosphates, has been optimised later.[235,236] In 1993, Keller *et al.*[237] improved the yield of diphosphate obtained and reported the separation of the phosphate species by flash column chromatography.

In this work, to a solution of the respective allylic alcohol (**139**) in acetonitrile at 37 ºC, was added an inorganic monophosphate source (**145**) and trichloroacetonitrile (**146**) and the resulting mixture was incubated for 15 min. The addition was repeated three times at five minute intervals and then the reaction mixture was directly passed through a silica column pre-equilibrated with running buffer (iPrOH:NH_4OH:H_2O, 6:2.5:0.5). The inorganic monophosphate source, bis-triethylammonium hydrogen monophosphate (TEAP, **145**), was prepared *in situ* prior to use as a solution in acetonitrile. The triethylammonium counter ions are readily exchanged during the silica column to give the trisammonium farnesyl diphosphate salt (**144**) as a white powder. The fractions were checked by TLC (mobile phase iPrOH: NH_4OH: H_2O, 6:3:1), and the desired diphosphate was lyophilised to yield the desired product. (Scheme 32).



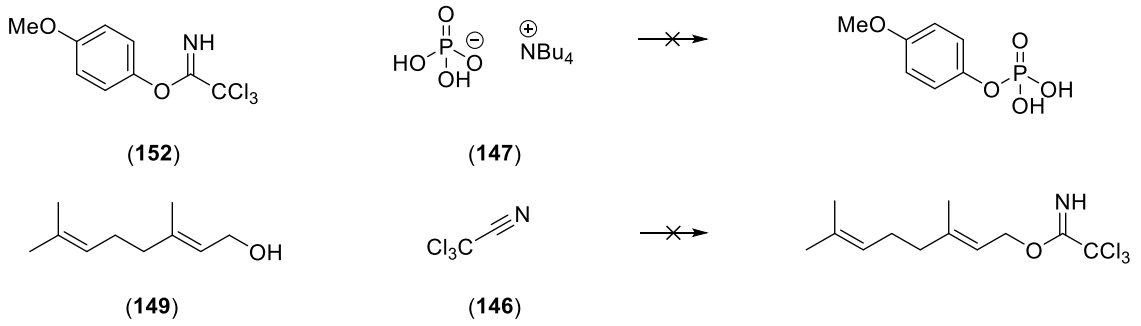
Scheme 32. Phosphorylation of farnesol analogues, method 2.

Danilov *et al.*[238] used commercially available tetrabutylammonium dihydrogen monophosphate (**147**, Scheme 33) as the inorganic phosphate source, however, this method requires an extra ion- exchange column to generate the ammonium salt as well as producing tetraphosphate species. Recently, Lira *et al.*[239] proposed a mechanism for this reaction. Trichloroacetonitrile (**146**) reacts with tetrabutylammonium dihydrogen phosphate (**147**) to give the corresponding reactive phosphorylated trichloroacetimidate (**148**), followed by nucleophilic attack by geraniol (**149**) to yield the desired phosphorylated alcohol (**150**) and trichloroacetamide (**151**, Scheme 33).



Scheme 33. Diphosphorylation mechanism proposed by Lira *et al.*[239]

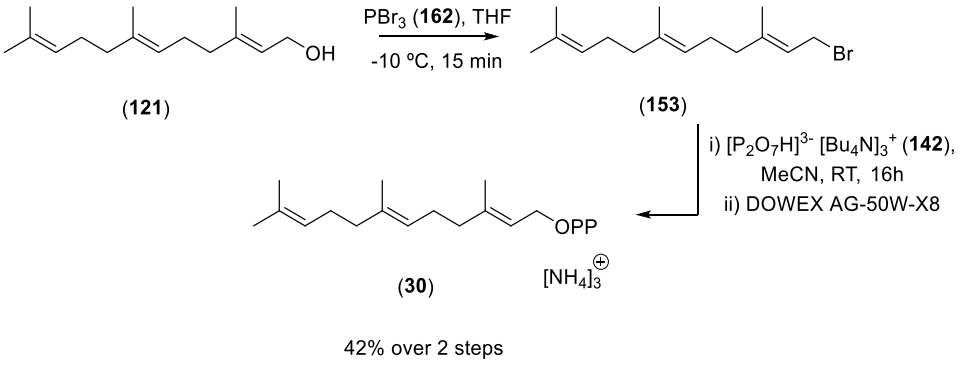
This proposal was supported by ^{13}C-NMR spectroscopy experiments. No reaction products were observed when 4-methoxybenzyl trichloroacetimidate (**152**) was mixed with tetrabutylammonium dihydrogen phosphate (**147**) or by pre-mixing trichloroacetonitrile (**146**) with geraniol (**149**, Scheme 34), which rules out the activation of the alcohol by trichloroacetonitrile (**146**) and displacement by tetrabutylammonium dihydrogen phosphate (**147**). In contrast, when geraniol (**149**), trichloroacetonitrile (**146**) and tetrabutylammonium dihydrogen phosphate (**147**) were successively mixed, it was observed the disappearance of the ^{13}C signals from trichloroacetonitrile (**146**) and the appearance of the corresponding signals from the phosphorylated alcohol (Scheme 34).



Scheme 34. Control reactions performed by Lira *et al.* to establish the mechanism shown in Scheme 33.[239]

3.3. Synthesis of ((2*E*,6*E*)-FDP, 30)

(2*E*,6*E*)-FDP (**30**) was synthesised as a control to check the catalytic activity of various enzymes by GC-MS, including GAS, GDS, GdolS, Gd11olS, ADS, AS, SdS, EpicS and DCS-W279Y (Chapter 4), and the generated mutants of GdolS and EpicS (Chapter 2). For this, the protocol described by Poulter *et al.*[231,232] was followed. The commercially available (*E*,*E*)-farnesol (**121**) was converted quantitatively to the corresponding farnesyl bromide[233] (**153**) and diphosphorylated using tris(tetrabutylammonium) hydrogen diphosphate (**142**). After 16 hours, FDP (**30**) was obtained as a salt containing tris(tetrabutylammonium) counter ions, which was exchanged to the desired trisammonium salt **30** by cation exchange chromatography. The overall yield was 42% after purification by reverse phase HPLC (Scheme 35).

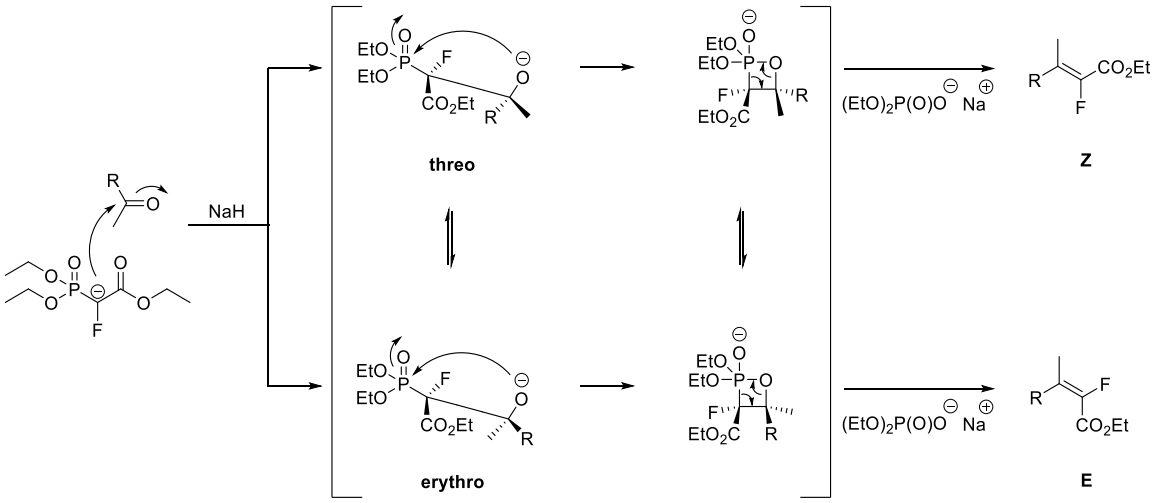


Scheme 35. Synthesis of FDP (**30**) from (2*E*,6*E*)-farnesol **121**.

3.4 Synthesis of (2*Z*,6*E*)-6-fluorofarnesyl diphosphate ((2*E*,6*Z*)-6F-FDP, 86)

While (2*E*,6*Z*)-6F-FDP (**86**) has been prepared previously in the Allemann group,[122,123] it was necessary to repeat the synthesis, and a novel more simplified procedure was developed to improve the yield obtained.

Initially, 6-methyl-5-hepten-2-one (**154**) was used in a Horner-Wadsworth-Emmons reaction (HWE, Scheme 36) with triethyl 2-fluoro-2-phosphonoacetate (**155**)[240] using NaH (**156**) as base in THF.[241]



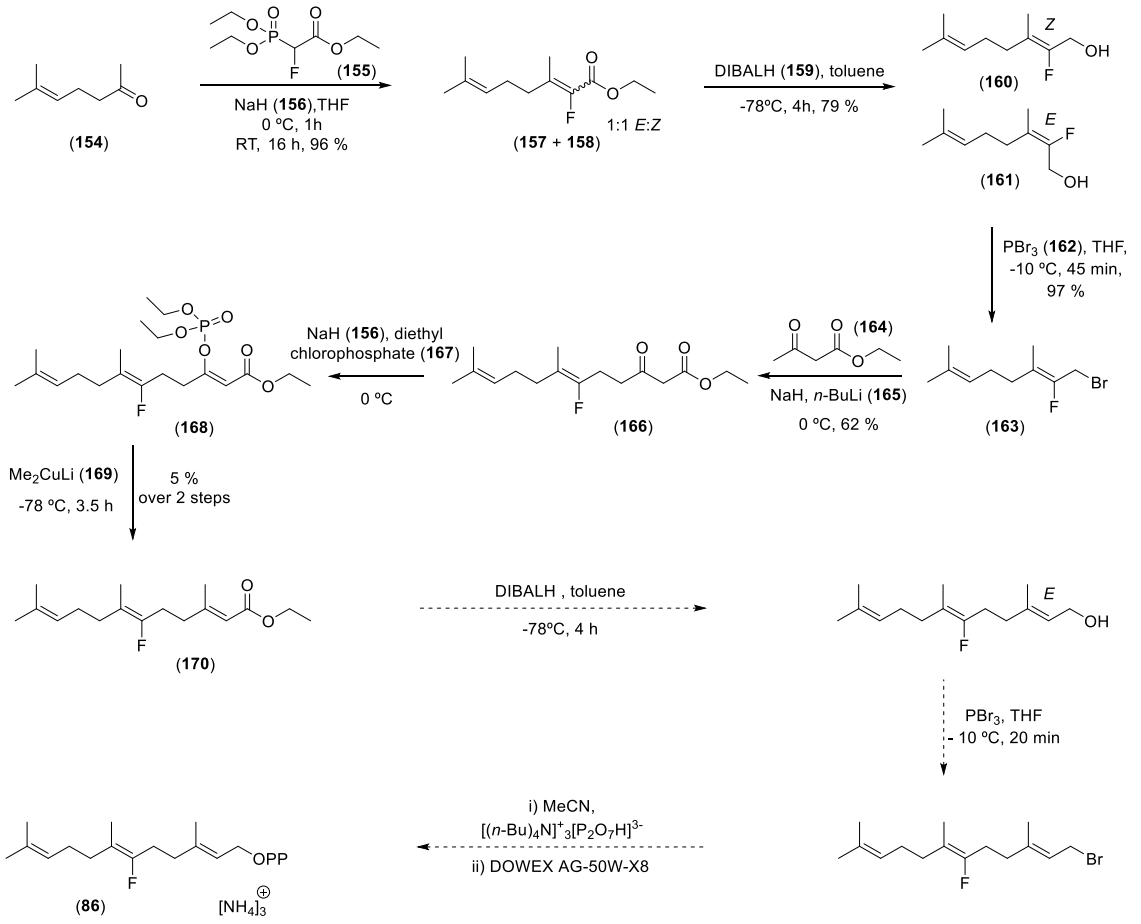
Scheme 36. HWE mechanism with betaine formation and driving force of phosphine oxide formation.

The resulting mixture of esters (**157** + **158**, Z/E 1:1) was reduced to 2F-geraniol (**160**) and 2F-nerol (**161**) with high stereoselectivity using DIBALH (**159**), which were then separated by flash column chromatography (Scheme 37).

During this study, the stereochemistry of the esters obtained after HWE reactions were investigated with the assistance of ^{1}H NMR spectrum, in which they were presented magnetic anisotropy effects due to the presence of a carbonyl group. These effects were generally observed as changes in the chemical shift of a key protons depending on its special relationship to a carbonyl group, being shielded or deshielded. Alternatively, the reduced alcohols were analysed with the assistance of ^{1}H- ^{1}H NOESY NMR experiments, which resulted valuable for making stereochemical assignments.

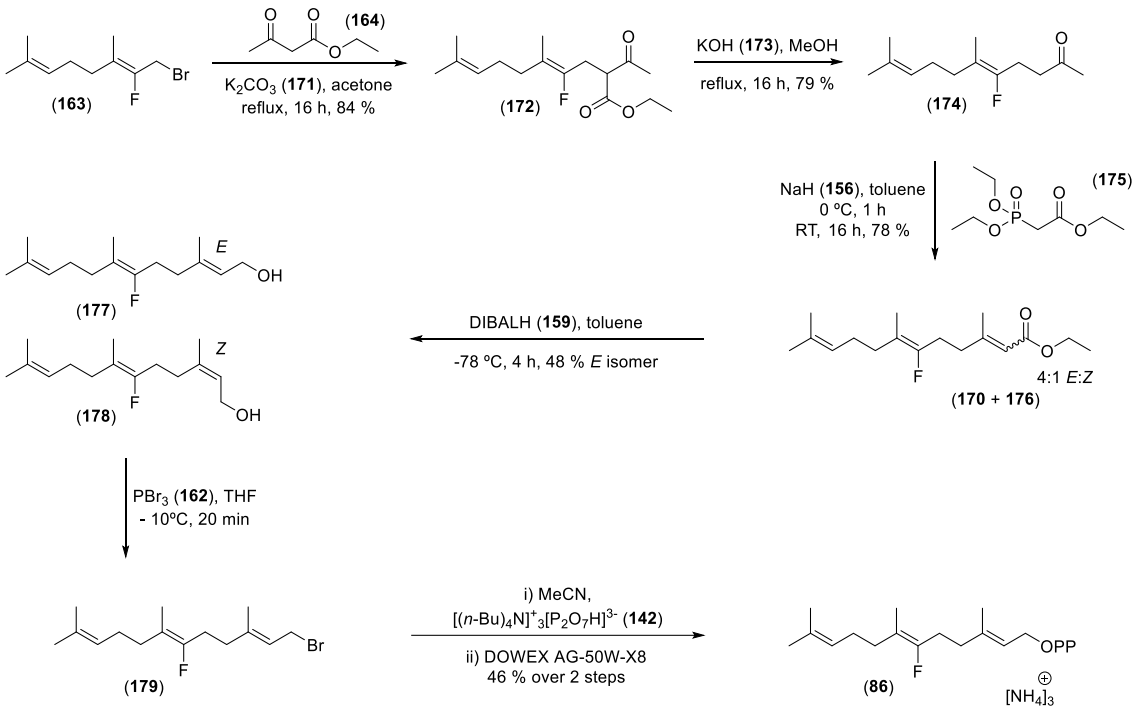
The bromination of 2F-geraniol (**160**) to give **163** was achieved *via* addition of phosphorous tribromide (**162**).

Incorporation of the second isoprene moiety was initially attempted following the sequence described by Weiler *et al*,[242,243] which comprises dianion alkylation to form **166**, enol phosphate formation (**168**) and lithium dimethylcuprate coupling to give ester **170** (Scheme 37). The formation of **170**, however, was achieved in low yield due to the unstable nature of the enol phosphate **168**.



Scheme 37. First synthetic route used to access (2*E*,6*Z*)-6F-FDP (**86**).

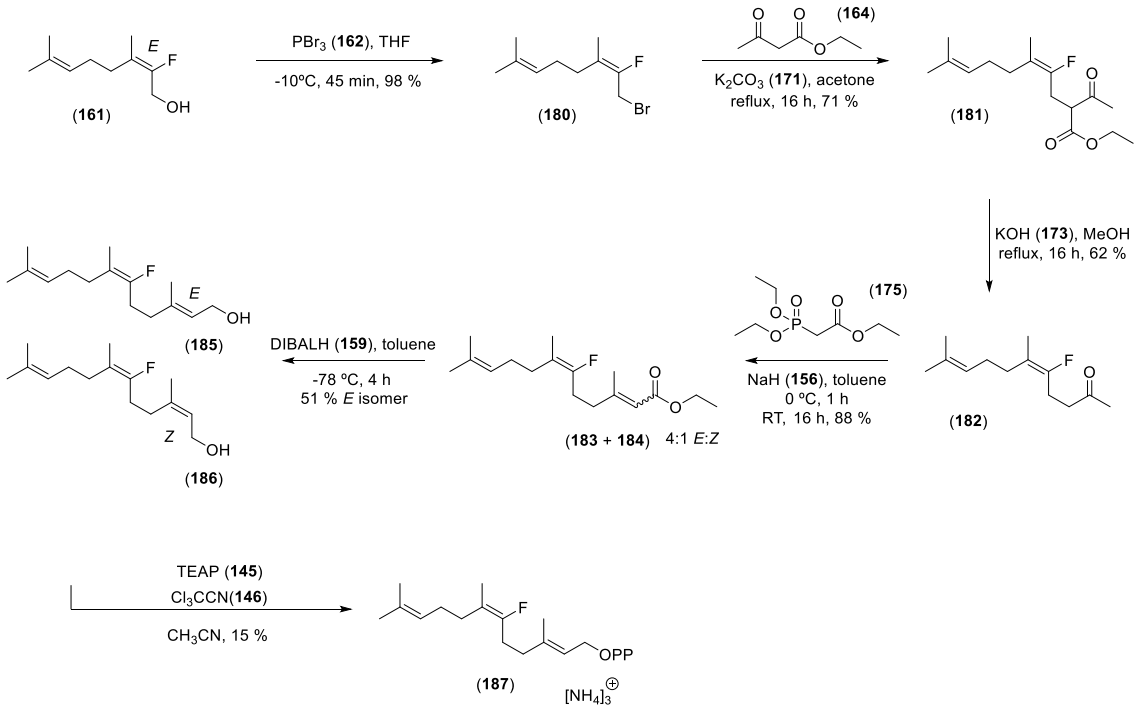
With the purpose of improving the yields obtained, but also to use compounds easier to handle, the route to access ester **170** was modified (Scheme 38). Thus, the carbon chain was extended by alkylating bromide **163** with ethyl acetoacetate (**164**) using potassium carbonate (**171**) as the base in dry acetone.[244,245] The resulting geranyl β-ketoester **172** was decarboxylated with potassium hydroxide (**173**) in methanol under reflux conditions to yield fluorinated geranyl acetone **174**. Following this, a second HWE reaction was carried out using triethyl phosphonoacetate (**175**),[246] to yield a mixture of 2*E* and 2*Z* esters (**170** + **176**) with good stereoselectivity (*E*/*Z* 4:1). The mixture of esters was reduced with DIBALH (**159**) to give alcohols **177** and **178**, which were separated by flash column chromatograph. Lastly, alcohol **177** was diphosphorylated using Poulter's methodology (*via* bromide **179** formation) to give **86**.[231,232]



Scheme 38. Second synthetic route to (2*E*,6*Z*)-6F-FDP (**86**).

3.5 Synthesis of (2*E*,6*E*)-6-fluorofarnesyl diphosphate ((2*E*,6*E*)-6F-FDP, 187)

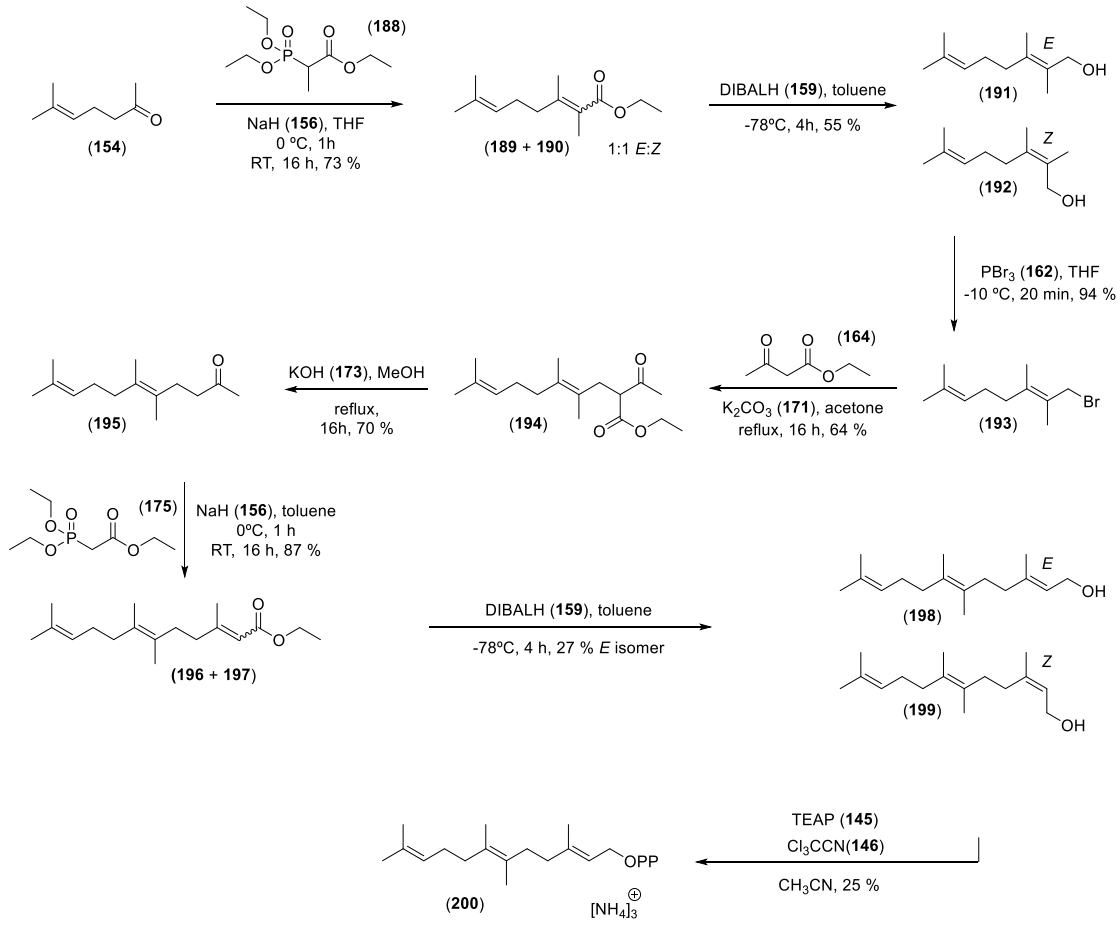
(2*E*,6*E*)-6F-FDP (**187**) was synthesised making use of the previously obtained 2F-nerol (**161**, Scheme 37) following the second route described above for the synthesis of (2*E*,6*Z*)-6F-FDP (**86**, Scheme 38), with the exception that diphosphorylation was carried out following Keller's instructions (Scheme 39).[237]



Scheme 39. Synthetic route to (2*E*,6*E*)-6F-FDP (**187**) from 6F-nerol (**161**).

3.6. Synthesis of (2*E*,6*E*)-6-methylfarnesyl diphosphate ((2*E*,6*E*)-6Me-FDP, 200)

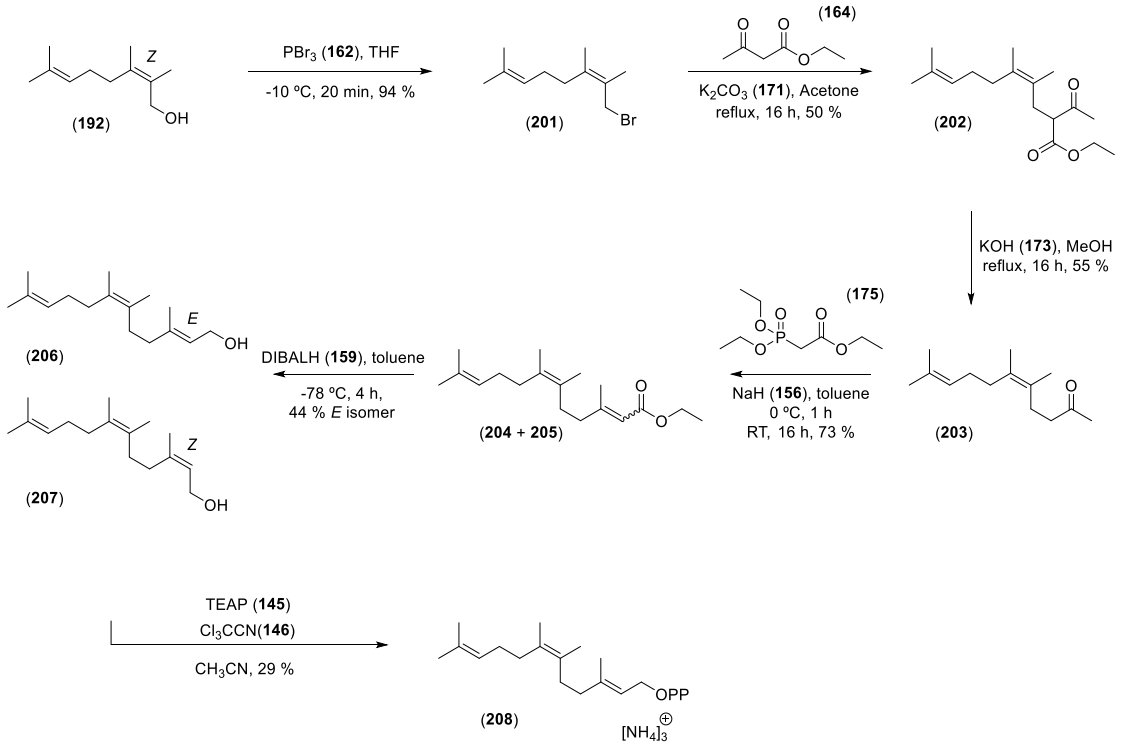
The novel compound (2*E*,6*E*)-6Me-FDP (**200**, Scheme 40) was synthesised following the route 2 previously developed for the synthesis of 6F-fluorinated analogues (Schemes 38 and 39). Triethyl 2-phosphonopropionate (**188**) was used for the HWE reaction with 6-methyl-5-hepten-2-one (**154**) to generate a 1:1 mixture of esters **189** and **190**, that were reduced using DIBALH (**159**) to produce 2-methyl geraniol (**191**, *E* isomer) and 2-methyl nerol (**192**, *Z* isomer). The *E* isomer **191** was brominated using phosphorous tribromide (**162**). The bromination proceeded faster than when fluorine was present, and 20 min were sufficient for reaction completion. Subsequently, the allylic halide **193** was alkylated with ethyl acetoacetate (**164**) using potassium carbonate (**171**) in dry acetone to yield the methylated geranyl β-ketoester **194**, which was decarboxylated with potassium hydroxide (**173**) in methanol under reflux overnight to afford the corresponding ketone **195**. The following HWE reaction gave *E* and *Z* esters (**196** + **197**), which was reduced to the alcohols using DIBALH (**159**). Separation by flash column chromatography afforded the all-*trans* 6-methyl farnesol **198**, separated from the *cis* isomer **199**. Diphosphorylation of farnesol analogue **198** was carried out following Keller's instructions (Scheme 40).[237]



Scheme 40. Synthetic route to (2*E*,6*E*)-6Me-FDP (**200**).

3.7. Synthesis of (2*E*,6*Z*)-6-methylfarnesyl diphosphate ((2*E*,6*Z*)-6Me-FDP, 208)

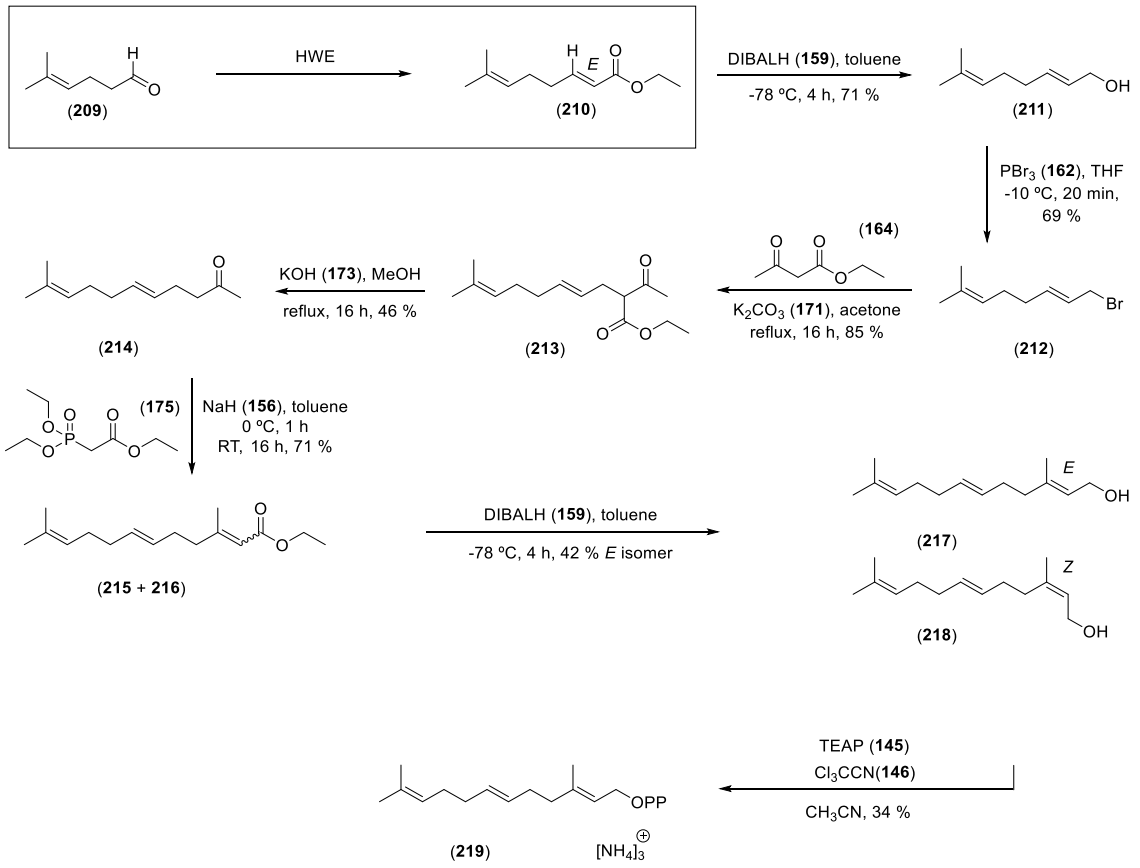
In analogy to the synthesis of (2*E*,6*E*)-6F-FDP (**187**, Scheme 39), (2*E*,6*Z*)-6-methyl-FDP (**208**) was synthesised from the previously obtained 2-methyl nerol (**192**, Scheme 40), Scheme 41.



Scheme 41. Synthetic route to (2*E*,6*Z*)-6Me-FDP (**208**) from 2-Me-nerol (**192**).

3.8. Synthesis of (2*E*,6*E*)-7-hydrogenfarnesyl diphosphate ((2*E*,6*E*)- 7H-FDP, 219)

The methodology followed for the synthesis of (2*E*,6*E*)-7H-FDP (**219**) was the same as for the above described analogues, however, it was necessary to prepare aldehyde **209** for the initial olefination using the HWE reaction (Scheme 42). Furthermore, two different phosphonate reagents (**175** and **231**, Scheme 45) were tested for the HWE reaction with aldehyde **209** (discussed below), which were completely selective for the *E* isomer of ester **210**.



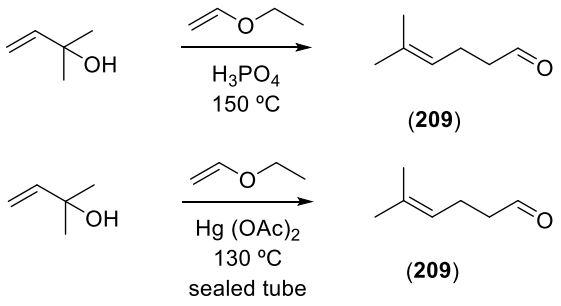
Scheme 42. Synthetic route to (2*E*,6*E*)-7H-FDP (**219**).

3.8.1. Preparation of 5-Methyl-4-hexenal (209)

One-pot synthesis

Initially, two different methods[247,248] were attempted to obtain aldehyde **209** in a one-step synthesis (Scheme 43) by MChem student Alexi T. Sedikides (Allemann group, Cardiff University). For both methods, the ^{1}H NMR spectra of the crude material indicated full conversion of the alcohol and the presence of an aldehyde peak. Purification by flash-column chromatography (both, regular and triethylamine-deactivated) resulted in product degradation. The use of the crude mixture directly in the olefination reactions was also attempted, however this two-step performance was very low yielding (<10 %).

A different route to **209** was explored successfully and high yielding (malonate route, Scheme 44).

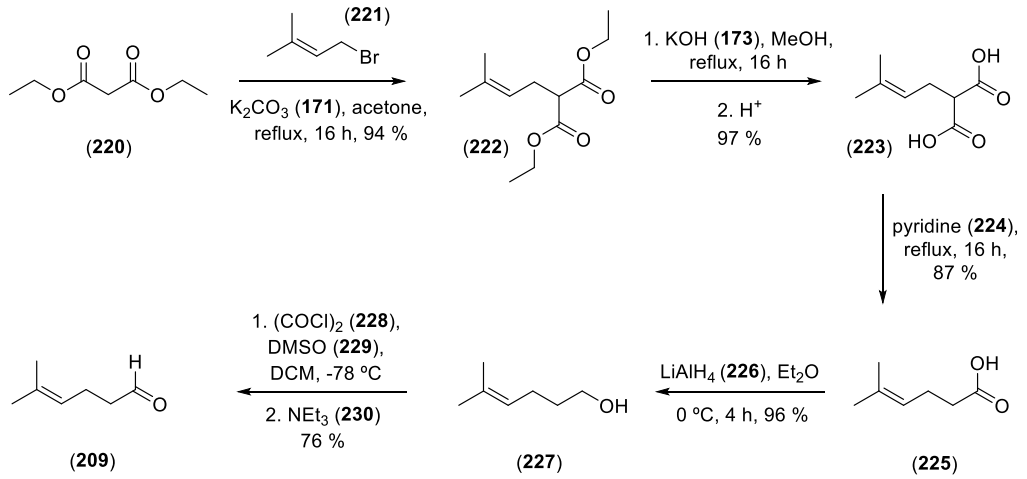


Scheme 43. One-pot synthesis of aldehyde **209** (synthesis performed by MChem student Alexi T. Sedikides under my supervision).

Malonate route

The synthetic route that was pursued is outlined in Scheme 44, which was developed in collaboration with MChem student Alexi T. Sedikides (Allemann group, Cardiff University). First, diethyl malonate (**220**) was treated with potassium carbonate (**171**) and allylic bromide **221** in dry acetone to yield the alkylated malonate **222**.[244] Following this, hydrolysis of **222** using potassium hydroxide (**173**) in methanol yielded the desired diacid **223**. The diacid **223** was then decarboxylated by refluxing in pyridine (**224**), followed by acidic work-up to afford the monoacid **225**. Reduction of the resulting monoacid **225** with an excess of lithium aluminium hydride (**226**) in dry diethyl ether gave the primary alcohol **227**, which was stable under atmospheric conditions for several weeks. A Swern oxidation[249,250] was used for the conversion of **227** to aldehyde **209** (Scheme 44). The reported boiling point of aldehyde **209** is 150 ºC and evaporation of DMSO (**229**, b.p. 189 ºC) by reduced pressure was not possible. Multiple washes with cold water of the organic phase during liquid-liquid extraction did not yield in complete removal of residual DMSO (**229**), as indicated by ^{1}H NMR spectroscopy. The use of the crude product **209** in the attempted olefination, Scheme 42, was unproductive because of decomposition of the materials. Purification by flash-column chromatography proved successful and aldehyde **209** was obtained pure in good yield. Oxidation of **227** with stoichiometric amounts of

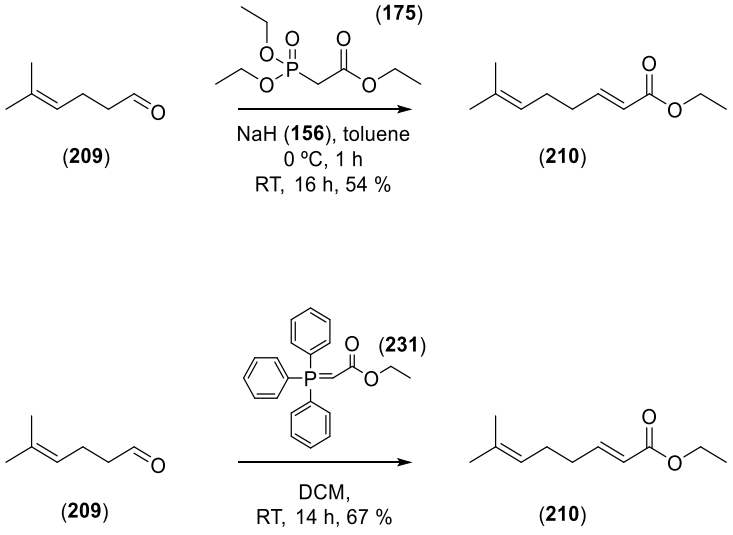
4-acetylamino-TEMPO oxidation[251] was also attempted several times by Alexi T. Sedikides (MChem student, Cardiff University), however full conversion was not achieved.



Scheme 44. Malonate route to access aldehyde **209**.

3.8.2. Preparation ethyl (*E*)-7-methylocta-2,6-dienoate (205)

Two stabilised phosphorous ylides were tested for the conversion of aldehyde **209** to the ester **210** (Scheme 45). First, the use of **175** with NaH (**156**) in toluene gave exclusively the *E* alkene ester **210**, 54% yield) from **209**. Similarly, the base-free reaction of triphenyl substituted phosphorane **231** with **209** yielded the *E* isomer of the forming ester **210** in DCM (67% yield).



Scheme 45. Conversion of aldehyde **209** to ester **210**.

3.9. Synthesis of 2,3- and 10,11-thiirane FDP analogues (2,3-thii-FDP (241) and 1,10-thii-FDP (250))

3.9.1. Preparation of 2,3-thiirane FDP (241)

2,3-Thiirane farnesol (**240**) was prepared *via* rearrangement of a dimethylamino thiocarbamate derivative of 2,3-epoxy farnesol (**237**) to the respective dimethylamino carbamate of 2,3-epithio alcohol (**239**, Scheme 46).[252]

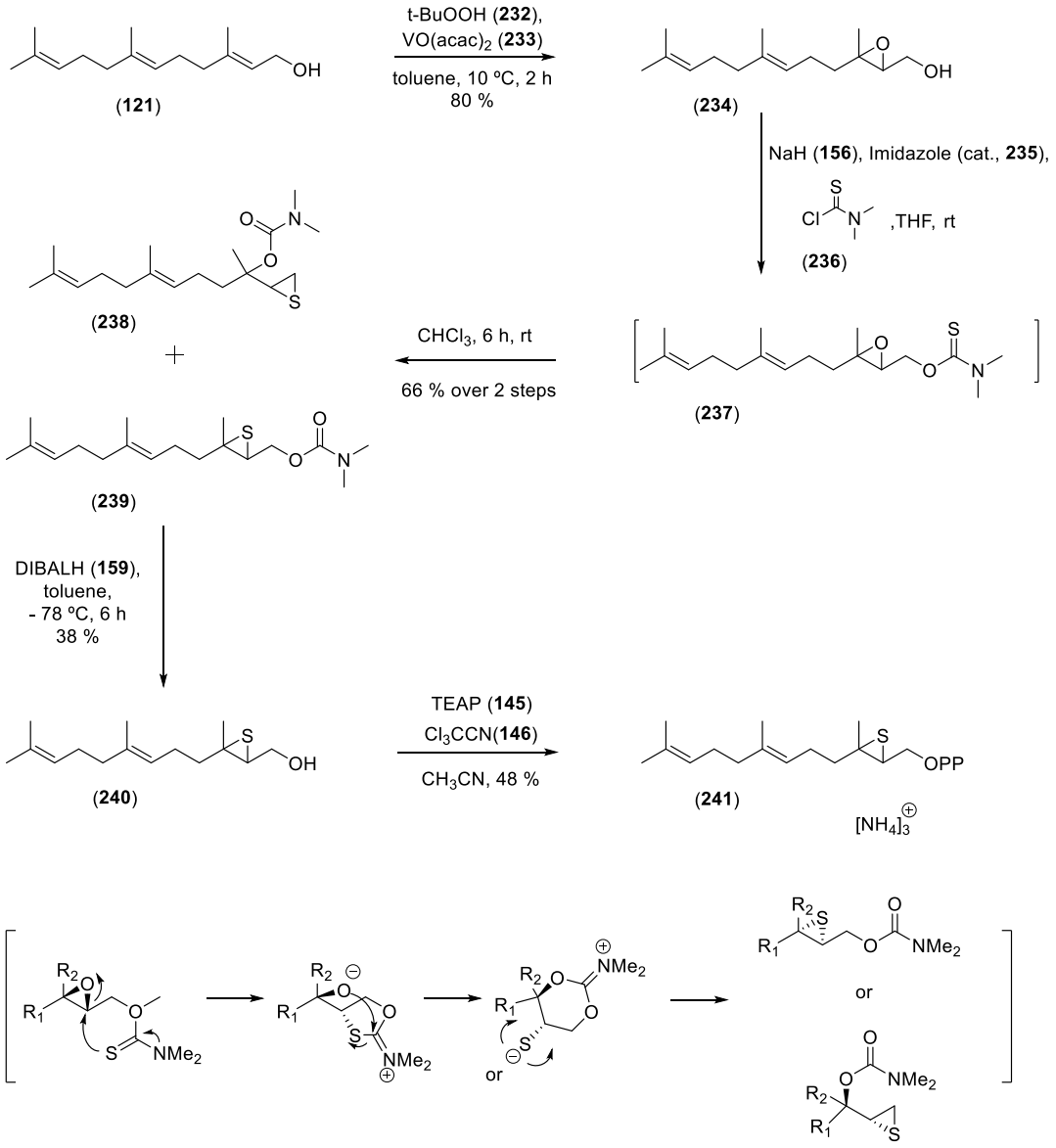
First, 2,3-epoxy farnesol (**234**) was prepared directly from farnesol (**121**) using an achiral epoxidation method established by Sharpless.[253] This method, which involves a vanadyl transition metal hydroperoxide reagent, exhibits remarkable reactivity towards allylic alcohols and resulted in a high yielding and regioselective formation of the 2,3-epoxy alcohol **234**.

The intramolecular epoxide-thiirane interchange reactions using thiourea or thiocyanates are well documented (see below). Here, following the instructions from Wicha *et al.*,[252] 2,3-epoxy farnesol (**234**) was treated with sodium hydride (**156**) in THF in the presence of a catalytic amount of imidazole (**235**). TLC analysis showed the presence of 3 products. Incubation of the crude product in chloroform gave compounds **238** and **239** (1:1), verified by NMR and MS spectroscopy analysis. A rearrangement mechanism has been proposed,[252] through inversion of configuration at both stereogenic centers of the epoxide **234** leading to **238** and **239** from the unstable compound 2,3-epoxy thiocarbamate **237** (Scheme 46).

The two obtained thiirane-carbamate species were separated by flash-column chromatography and 2,3-thiirane carbamate (**239**) was reduced to the desired 2,3-thiirane alcohol (**240**).

Different reducing agents were used for the reduction of the carbamate **239**. Sodium borohydride ($NaBH_4$) proved to be ineffective in both methanol and water, even after extended times of reaction (>48 h). Wicha *et al.*,[252] reported the use of lithium aluminium hydride ($LiAlH_4$, **226**), which was found to be inefficient in the reduction of **239** and only the starting material could be recovered. Reduction with DIBALH (**159**) was achieved at -78 ºC, however, the reaction did not arrive at full completion even after excess of DIBALH (**159**) was added and the reaction time was extended.

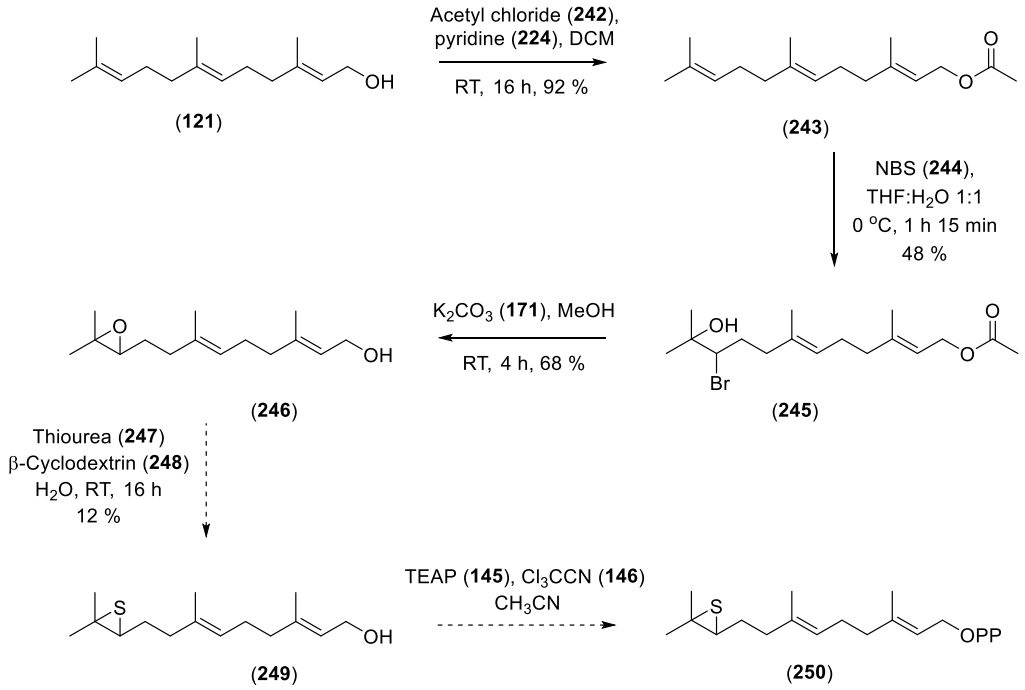
After DIBALH reduction, starting material (**239**) and product (**240**) were separated by flash- column chromatography to yield the 2,3-thiirane farnesol (**240**), which was diphosphorylated following Keller's instructions[237] to give 2,3-thiirane-FDP (**241**).



Scheme 46. Synthesis of 2,3-thiirane FDP (**241**). Bottom, the proposed mechanism for the formation of thiirane-carbamates, adapted from reference [252].

3.9.2. Preparation of 10,11-thiirane FDP (250)

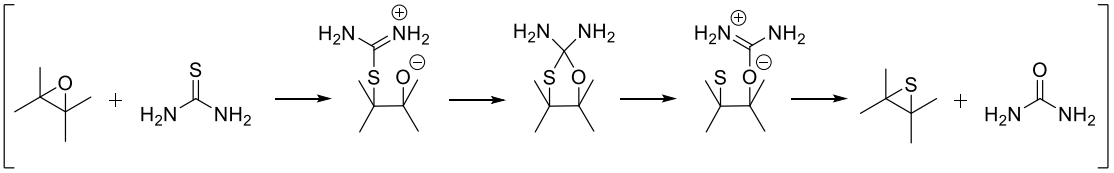
The synthesis of 10,11-thiirane FDP (**250**) was attempted from 10,11-epoxide farnesol (**246**) using a sulfur source, as shown below. First, farnesol (**121**) was protected to generate the corresponding farnesyl acetate **243**. Then, **243** was treated with *N*-bromosuccinimide (NBS, **244**) in THF: H_2O (1:1) to yield the bromohydrin **245**. Ring closure and acetyl deprotection were achieved using potassium carbonate (**171**) as the base in methanol, yielding 10,11-epoxy farnesol (**246**, Scheme 47).



Scheme 47. Synthetic route to access product **250**.

Conversion of 10,11-epoxide to 10,11- thiirane

A variety of methods have been developed for the preparation of thiiranes.[254,255] Amongst these, the conversion of epoxides to thiiranes by an oxygen-sulfur exchange reaction is the most common method (with thiourea in Scheme 48).



Scheme 48. Proposed mechanism for the conversion of oxiranes to thiiranes using thiourea.

Several sulfur sources have been used to produce thiiranes from epoxides, such as thiourea, [256–260] inorganic thiocyanates[261–263], phosphine sulfide,[264] silica gel supported potassium thiocyanate,[265] indium halides/KSCN[266] and polymeric cosolvent/NH_4SCN,[267] together with solvent-free conditions,[268] and ionic liquids.[269] However, most of these methods suffer from important limitations such us extended reaction times, elevated temperatures, hazardous reagents and solvents and undesirable side reactions due to rearrangements or polymerization of the oxiranes resulting in low yields of the thiiranes.

Here, several conditions were used to attempt to form 10,11-thiirane farnesol (**249**) from 10,11- epoxy farnesol (**246**), which are summarised in Table 10. None of these methods was successful even after extended reaction times, and the starting material was normally recovered. We hypothesised that the epoxide was hindered by its substituents making it difficult to perform the reaction.

Starting material (SM)	Sulfur source (equivalents)	Solvent	Time (hours)	Temperature (°C)	Observed product
246	thiourea (**247**) (1.5 eq.)	MeOH	>48 h	R.T.	SM
246	thiourea (**247**) (3 eq.)	MeOH	>48 h	Reflux	SM
246	KSCN (**251**) (1.5 eq.)	H_2O: EtOH (1:1)	>48 h	R.T.	SM
246	KSCN (**251**) (3 eq.)	80% H_3PO_4 H_2O: ether	>48 h	R.T.	SM
246	NH_4SCN (2 eq.)	H_2O	>48 h	R.T.	SM

Table 10. Reaction conditions attempted for the conversion of **132** to **135**.

Rama Rao *et al.*[270,271] have reported the use of β-cyclodextrin (**248**) for the synthesis of thiiranes from epoxides in water using both thiourea (**247**) and potassium thiocyanate (**251**). Cyclodextrins are cyclic oligosaccharides that possess hydrophobic cavities, which can bind substrates selectively and catalyse chemical reactions with high selectivity. Catalysis occurs through reversible formation of host-guest complexes by non-covalent interactions. Assisted orientation of the thiirane ring might be achieved for favourable attack.

Thus, the conditions stated by Rama Rao *et al.*[270,271] were used for the ring opening of 10,11-epoxy farnesol (**246**) (Scheme 47). First, β-cyclodextrin (**248**) was dissolved in water at 60 ºC and **246** was added to form a complex with **248**. Then, thiourea (**247**) was added to the mixture at room temperature and the reaction was left for 16 hours. Starting material (**246**) was always present even at long reaction times (>48 hours). Separation of the two observed compounds was feasible by flash column chromatography and **249** was obtained pure. The diphosphorylation of **249** to obtain 10,11- thiirane-FDP (**250**) was attempted by the method 1 (halogenation and diphosphorylation), but proved unsuccessful, with the formation of numerous side products. The method 2 (developed by Keller[237]) gave **250**, but also with various impurities that could not be identified. Due to time constrains and the nature of the assay, it was decided to test **250** with sesquiterpene synthases, even if the product could not be fully characterised or purified.

3.10. Synthetic results

In this chapter, the synthesis of a set of unnatural FDP analogues (**86**, **187**, **200**, **208**, **219**, **241** and **250**) has been described for testing the substrate behaviour with sesquiterpene synthases (Chapter 4).

A methodology previously developed for the synthesis of (2*E*,6*Z*)-6F-farnesol (**177**) was first attempted (Scheme 37), but ester **170** was achieved in low yield. This was later modified, and **177** was obtained in 18% yield over seven steps (Scheme 38).In the same fashion, but starting from 2F-nerol (**161**), the synthesis of a new (2*E*,6*E*)-6F-farnesol (**185**) was performed in 15% yield (Scheme 39).
This methodology was also followed for the synthesis of novel (2*E*,6*E*)-6Me-farnesol (**198**) and (2*E*,6*Z*)-6Me-farnesol (**206**), which were obtained in 4% and 3% yield, respectively (Schemes 40 and 41). The synthesis of (2*E*,6*E*)-7H-farnesol (**217**, Figure 42) was performed by following the same methodology used for the analogues described before. This time, however, the aldehyde **209** was needed, and the malonate route was followed for this purpose (Scheme 44). Overall, **217** was obtained in 2% yield over 12 steps. 2,3-Thiirane-farnesol (**240**) was prepared from the epoxide counterpart (**234**), following Wicha *et al.* [252] instructions with minor modifications, and was obtained in 10% yield over 4 steps (Scheme 46). For the synthesis of 10,11-thiirane-farnesol (**249**), it was first obtained the epoxide analogue **246** and then converted to **249** using Rama Rao *et al.*[270,271] approach (Scheme 47). Overall, **249** was obtained in 4% yield over 4 steps.

The obtained farnesol analogues were converted to the corresponding FDP analogues using the diphosphorylation methods 1 (Scheme 31) and 2 (Scheme 32), described in Section 3.2. These are summarised in table 11. Generally, method 1 resulted higher yielding, but method 2 was a faster approach that probe more useful for this exploratory study.

Farnesol	Method/ yield	FDP analogue
(*E*,*E*)-farnesol (**121**)	Method 1/ 42 %	(*E*,*E*)-FDP (**30**)
(*E*,*Z*)-6F-farnesol (**177**)	Method 1/ 46 %	(*E*,*Z*)-6F-FDP (**86**)
(*E*,*E*)-6F-farnesol (**185**)	Method 2/ 15 %	(*E*,*E*)-6F-FDP (**187**)
(*E*,*E*)-6Me-farnesol (**198**)	Method 2/ 25 %	(*E*,*E*)-6Me-FDP (**200**)
(*E*,*Z*)-6Me-farnesol (**206**)	Method 2/ 29 %	(*E*,*Z*)-6Me-FDP (**208**)
(*E*,*E*)-7H-farnesol (**217**)	Method 2/ 34 %	(*E*,*E*)-7H-FDP (**219**)
2,3-thiirane-farnesol (**240**)	Method 2/ 48 %	2,3-thiirane-FDP (**241**)
10,11-thiirane-farnesol (**249**)	Method 2/ Not applicable	10,11-thiirane-FDP (**250**)

Table 11. Summary of the diphosphorylation methods used for the production of FDP analogues.

CHAPTER 4

INCUBATION OF FDP ANALOGUES WITH SESQUITERPENE SYNTHASES

Chapter 4. Incubation of FDP analogues with sesquiterpene synthases

4.1. Preface

The work described in this chapter explores the substrate promiscuity of nine sesquiterpene synthases with respect to their ability to transform a pool of seven FDP-analogues. Synthesis of these compounds is described in Chapter 3. To simplify the narrative here, the numbering of FDP and FDP-analogues have been adapted as follow:

FDP and FDP analogues	Numbering in this chapter
(2*E*,6*E*)-FDP (**30**)	**1**
(2*E*,6*Z*)-6F-FDP (**86**)	**2**
(2*E*,6*E*)-6F-FDP (**187**)	**3**
(2*E*,6*E*)-6Me-FDP (**200**)	**4**
(2*E*,6*Z*)-6Me-FDP (**208**)	**5**
(2*E*,6*E*)-7H-FDP (**219**)	**6**
2,3-thiirane-FDP (**241**)	**7**
10,11-thiirane-FDP (**250**)	**8**

Table 12. FDP and FDP analogues used during this study.

In addition, the pentane extractable products after incubation of FDP (**1**) and FDP analogues (**2**-**8**) with the sesquiterpene synthases examined here have been labelled in relation/ concordance with the substrate used, as exemplified in Figure 59 with FDP (**1**) as the model. Due to this, the numberings of some previously described products have been adapted in this chapter as follow:

Generated products	Numbering in this chapter
(+)-germacrene A ((+)-**32**)	(+)-**1a**
(-)-germacrene A ((-)-**32**)	(-)-**1a**
(-)-germacrene D (**91**)	**1b**
(-)-germacradien-4-ol (**79**)	**1c**
amorpha-4,11-diene (**104**)	**1e**
(+)-aristolochene (**54**)	**1f**
(+)-epicubenol (**78**)	**1i**
(+)-δ-cadinene (**15**)	**1j**
(2*E*,6*Z*)-6F-germacrene A (**88**)	**2a**

Table 13. Renumbering of the terpene products obtained in this chapter.

For the sesquiterpene synthases under investigation, the initial cyclisation of FDP (**1**) proceeds between the positions 1 and 6 or 10, respectively. In addition, some 1,10-sesquiterpene synthases can undergo a second 1,6-cyclisation (Schemes 50 and 51). These were chosen as model enzymes because generally, the substrate modifications are made at the C6-C7 double bond of the FDP (**1**) structure. Thus, the influence of these modifications was investigated in enzymes with different cyclisation pathways.
Four of the chosen sesquiterpene synthases exclusively enable a 1,10-cyclisation; the plant-derived (+)-germacrene A synthase (GAS) and (-)-germacrene D synthase (GDS) from *Solidago canadensis*[111,123,126,127,272] produce (+)-germacrene A ((+)-**1a**, Figure 59) and (-)-germacrene D (**1b**) respectively. The bacterial (-)-germacradien-4-ol synthase from *Streptomyces citricolor* (GdolS)[125,169] converts FDP (**1**) to (-)-germacradien-4-ol (**1c**) and germacradien-11-ol-synthase (Gd11olS)[175,273,274] gives (4*S*,7*R*)-germacra-1(10)*E*,5*E*-diene-11-ol (**1d**) and traces of (-)-germacrene-D (**1b**).

Regarding selective 1,6- sesquiterpene synthases, amorpha-4-11-diene synthase (ADS) from *Artemisia annua*,[113,131,133,135] which produces (1*S*,6*R*,7*R*,10*R*)-amorpha-4,11-diene (**1e**) after two 1,6-cyclisations, was used.

Additionally, 1,10-sesquiterpene synthases that are able to catalyse a second 1,6-cyclisation were employed. (+)-Aristolochene synthase from *Penicillium roqueforti* (PRAS)[95,97,214] produces (+)-aristolochene (**1f**) together with small quantities of (-)-germacrene A ((-)-**1a**) and traces of valencene; selina-4(15),7(11)-diene synthase (SdS)[85] gives selina-4(15),7(11)-diene (**1g**) and small quantities of germacrene B (**1h**); (+)-epicubenol synthase (EpicS) from *Streptomyces coelicolor*[170] that produces (+)-epicubenol (**1i**) as the main product, as well as germacradien-4-ol (**1c**) and other minor coproducts (*****). The mutant (+)-δ-cadinene synthase-W279Y (DCS-W279Y, produced by Marianna Loizzi)[275] gives (+)-δ-cadinene (**1j**) and (-)-germacradiene-4-ol (**1c**) in 8:2 ratio. The mechanisms of the reactions catalysed by these synthases upon incubation with **1** are summarised in Scheme 50 and Scheme 51.

The rationale behind the use of this set of enzymes is diverse. In general, these were chosen as model enzymes to analyse the influence of the modifications introduced at the C6-C7 double bond of the FDP (**1**) structure on various enzymatic cyclisation mechanisms. Also, some enzymes can serve as controls to identify certain products, for example, EpicS and DCS- W279Y can serve as control enzymes to identify GdolS- like products, or PRAS for the identification of GAS- like products.

The enzymes used in this study were obtained as follows:

GdolS and EpicS were expressed and purified as stated in sections 2.2.1 and 2.3.1, respectively. GAS, GDS, ADS, PRAS and DCS-W279Y were expressed and purified following well established procedures in our laboratory.[88,123,131,272,275] Gd11olS and SdS were generously prepared for the experiments detailed in this chapter by Dr Prabhakar Srivastava (Allemann group, Cardiff University).

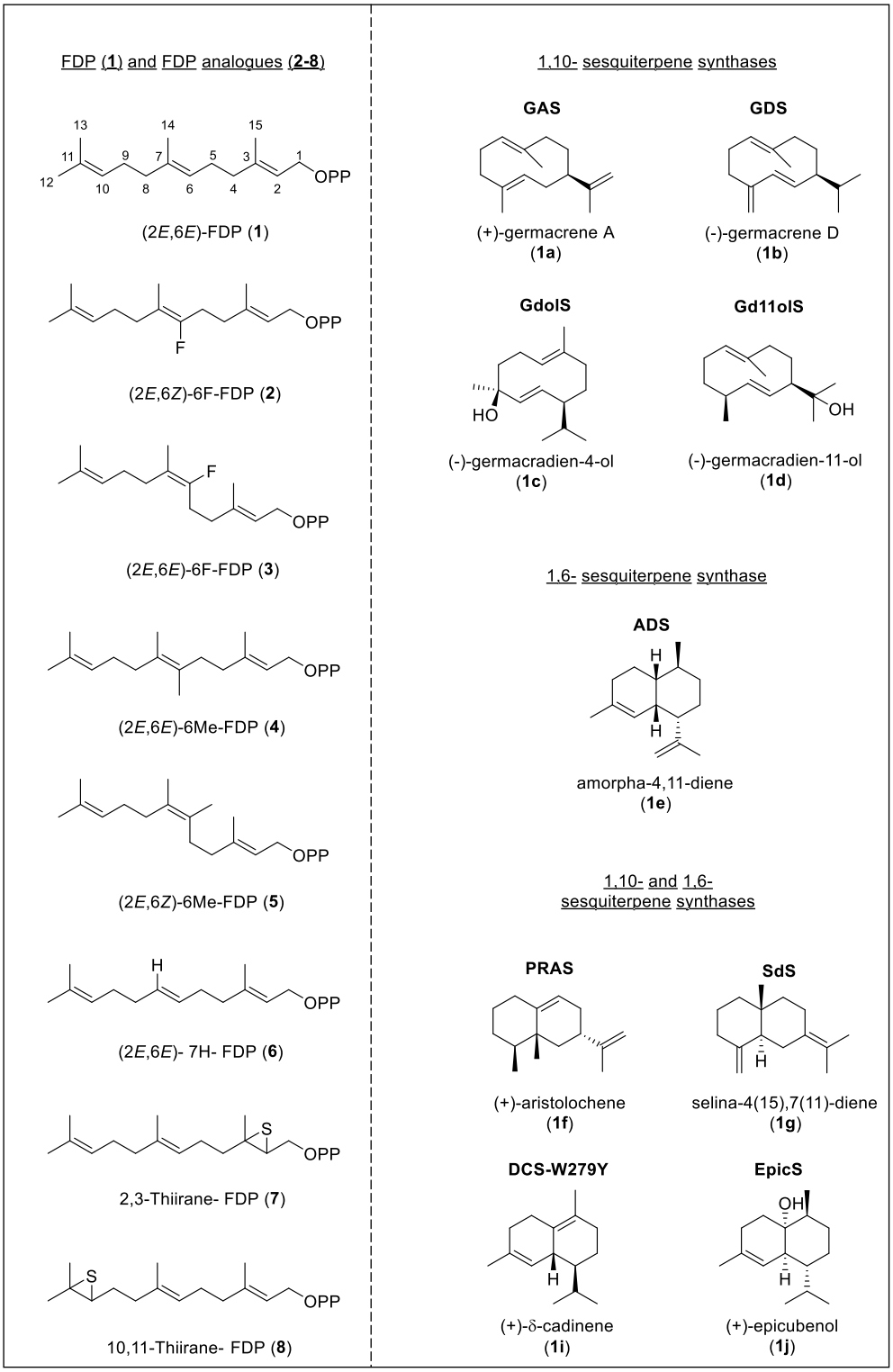


Figure 59. Left, FDP (**1**) and FDP- analogues (**2**-**8**). Right, sesquiterpene synthases used in this study and their main product upon incubation with FDP (**1**).

4.2. Substrate behaviour of (2*E*,6*E*)-FDP (1)

First, the enzymes covered in this study were incubated with FDP (**1**) and the pentane extractable products were analysed by GC-MS to ensure that their product profile was as described in the literature, Figure 60. Also, the mass spectra of the natural products were obtained to help with structural identification of future sesquiterpene analogues.

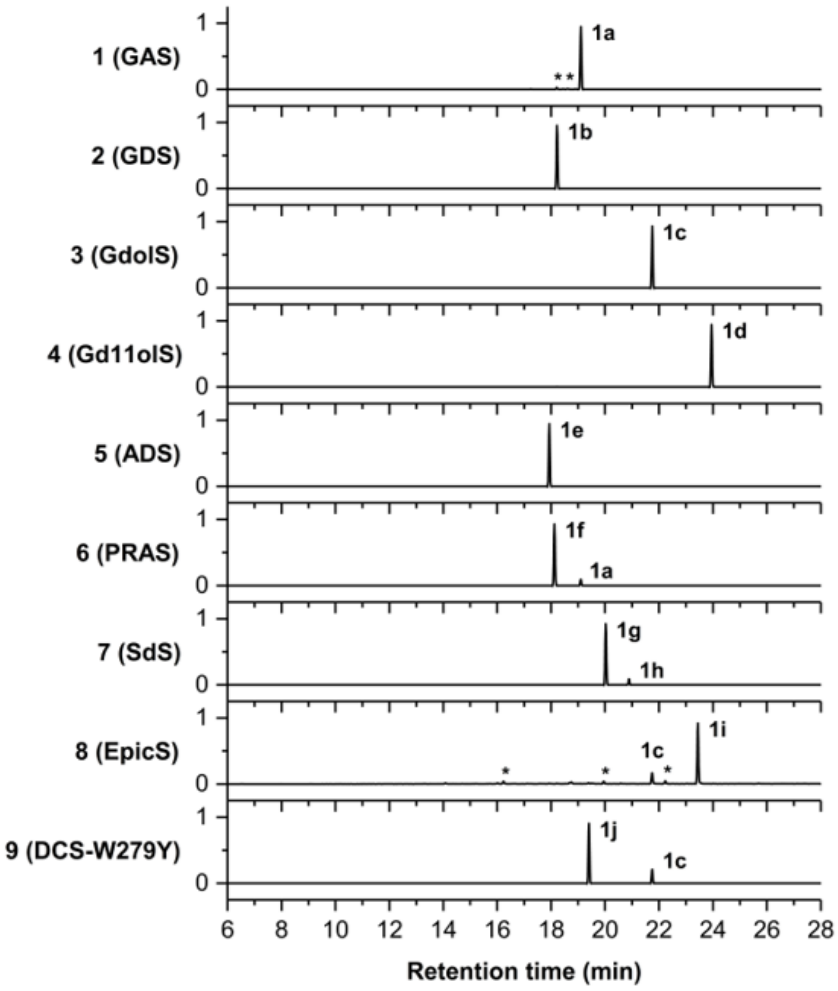


Figure 60. Total ion chromatograms of the pentane extractable products arising after incubations of FDP (1) with nine sesquiterpene synthases. Trace compounds (<5 %) are labelled with an asterisk.

Incubation 1. The incubation of FDP (**1**) with GAS yielded a predominant product in the pentane extracts eluting at 19.10 min, as judged by GC-MS analysis (Figure 61a). Its mass spectrum ionisation pattern is in agreement with that described in the literature for (+)-germacrene A ((+)-**1a**).[95] It shows the molecular ion with m/z = 204, $[M]^+$, Figure 61b. The base peak has m/z = 93 and displays fragments with m/z = 189 ($[M - CH_3]^+$) and 162 ($[M - C_3H_6]^+$), consistent with the parent structure of (+)-**1a**. Also, this characterisation was reinforced by comparison of the retention time and mass spectrum with the (-)-germacrene A product generated PRAS (incubation 6).[95] In addition, the well-known thermal Cope rearrangement of (+)-**1a** to (-)-β-elemene (Scheme 49) was observed through rising the injection temperature to 250 ºC at the GC injection port. A new product was observed in the chromatogram at a shorter retention time.[276] Traces of (-)-germacrene D and α-humulene were also detected.[272]

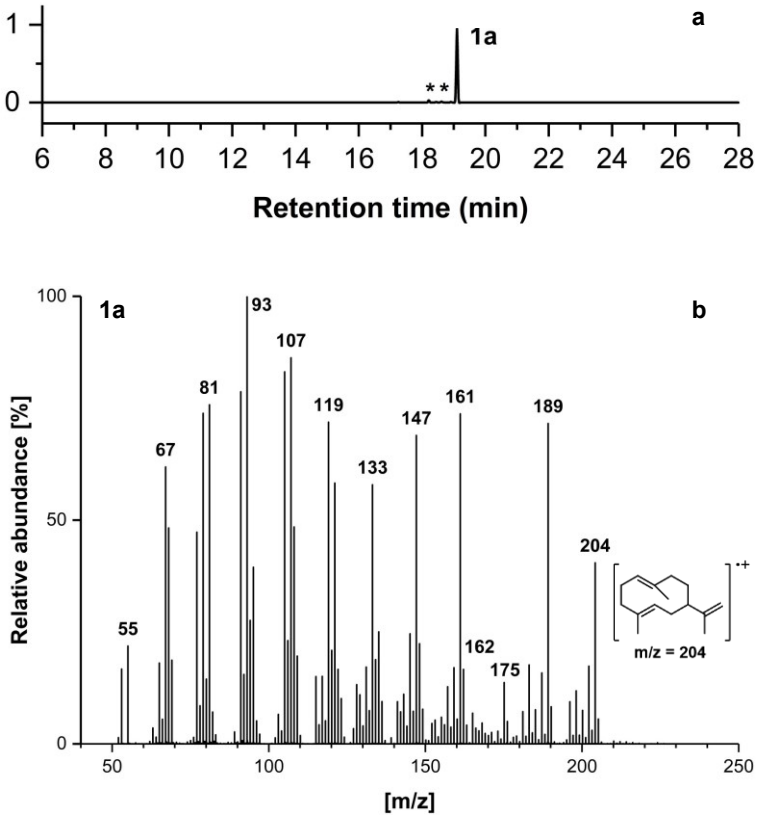
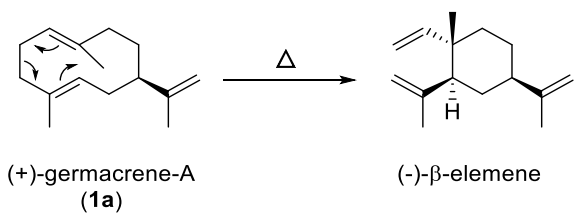


Figure 61. a) Total ion chromatogram of the pentane extractable products arising from incubation 1. b) Mass spectrum of the major product from incubation 1, showing the molecular ion of **1a**.



Scheme 49. Thermal Cope rearrangement of (+)-germacrene-A ((+)-**1a**) to (-)-β-elemene.[276]

Incubation 2. The incubation of GDS with FDP (**1**) gave rise to a single compound in the organic extracts eluting at 18.22 min, as judged by GC-MS (Figure 62a). The mass spectrum of this compound shows a molecular ion with m/z = 204 ($[M]^+$) and a base fragment peak with m/z = 161 ($[M - C_3H_7]^+$. This is consistent with the loss of 43 units of mass corresponding to loss of an isopropyl group from the parent ion (Figure 62b), and it is identical to that previously reported for (-)-germacrene D (**1b**).[123,272,277]

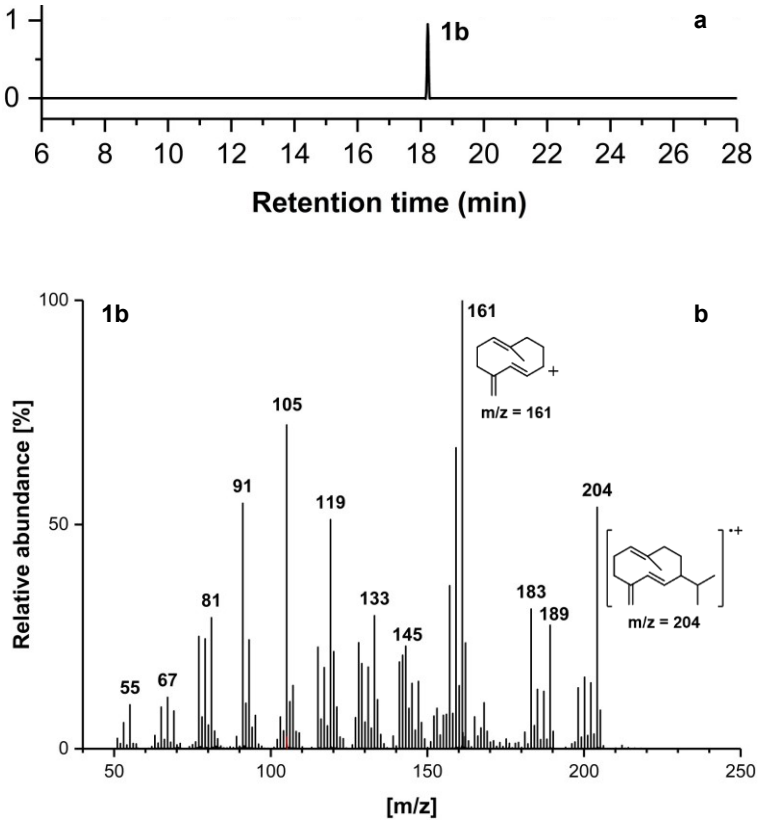


Figure 62. a) Total ion chromatogram of the pentane extractable products arising from incubation 2. b) Mass spectrum of the extracted product from incubation 2, showing the molecular ion of **1b** and possible fragment ion from **1b** with m/z = 161.

Incubation 3. The incubation of FDP (**1**) with GdolS produced (-)-germacradien-4-ol as the sole product, eluting at 21.75 min (**1c**, Figure 63a), as stated in Section 2.2.1 and previously reported. The observed mass spectrum matched those previously published.[125,169] Since the ionization mode is EI^+ in this analysis, the molecular ion could not be detected due to dehydration (m/z = 222, $[M]^+$), but the mass corresponding to the dehydrated structure (m/z = 204, $[M - H_2O]^+$) was observed, Figure 63b.

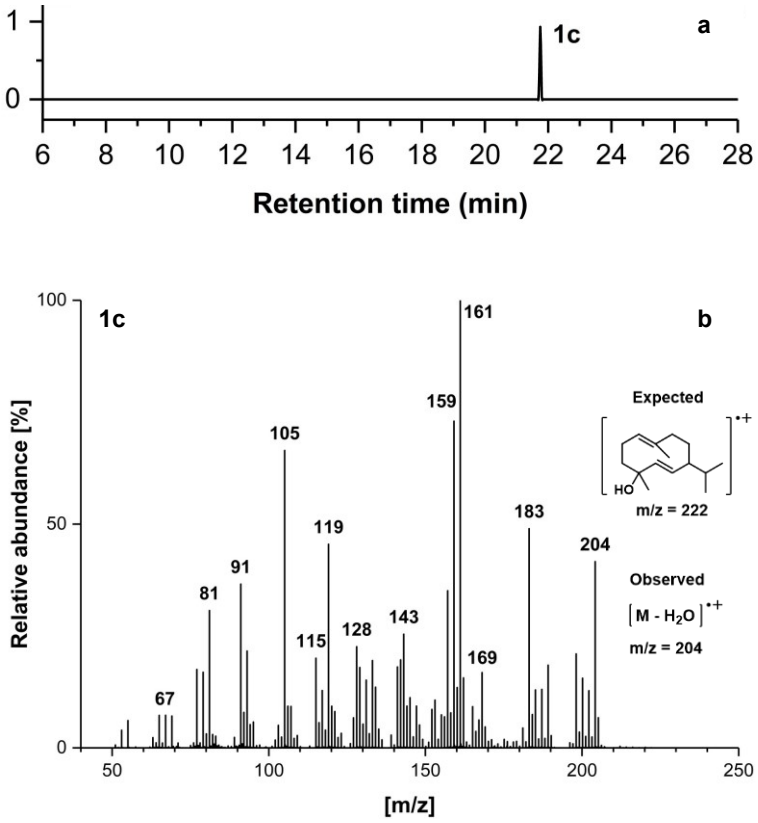


Figure 63. a) Total ion chromatogram of the pentane extractable products arising from incubation 3. b) Mass spectrum of the extracted product from incubation 3, showing the molecular ion of **1c** and highlighting the observed ion.

Incubation 4. The single product obtained in the organic extracts from the incubation of FDP (**1**) with Gd11olS eluted at 23.95 min, as judged by GC-MS. This compound was identified as germacradien-11-ol (**1d**, Figure 64a) because the mass spectrometry data shows a molecular ion peak with m/z = 222 ($[M]^+$), showing the presence of a hydroxyl group (Figure 64b) and is in accordance with the literature.[278] The base peak is with m/z = 161 $[M - H_2O - C_3H_7]^+$, and has major fragmentat ions with m/z = 204 $([M - H_2O]^+)$ and 189 $([M - H_2O - CH_3]^+)$. Traces of (-)-germacrene D (**1b**) are also reported in the literature, but were not observed in this work.[274]

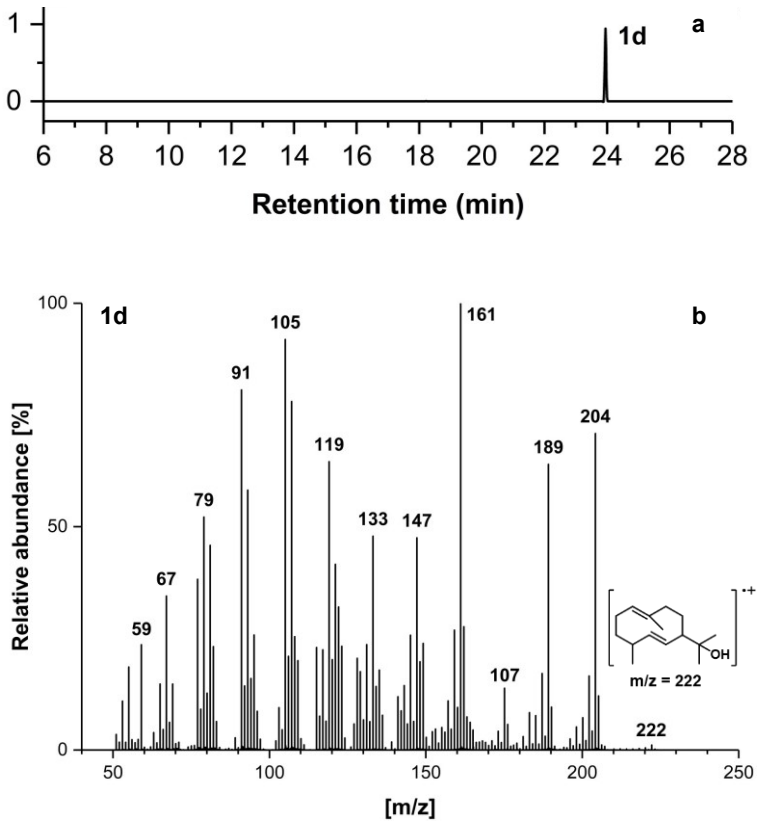


Figure 64. a) Total ion chromatogram of the pentane extractable products arising from incubation 4. b) Mass spectrum of the extracted product from incubation 4, showing the molecular ion of **1d**.

Incubation 5. The incubation of **1** with ADS gave exclusively one compound in the pentane extracts eluting at 19.93 min, as judged by GC-MS (Figure 65a). The mass spectrum of this compound completely matches that described in the literature for amorpha-4,11-diene (**1e**).[113,131,134] As in previous reports, it has a molecular ion with m/z = 204 ($[M]^+$) and the most abundant fragment ion is with m/z = 189 ($[M - CH_3]^+$), Figure 65b.

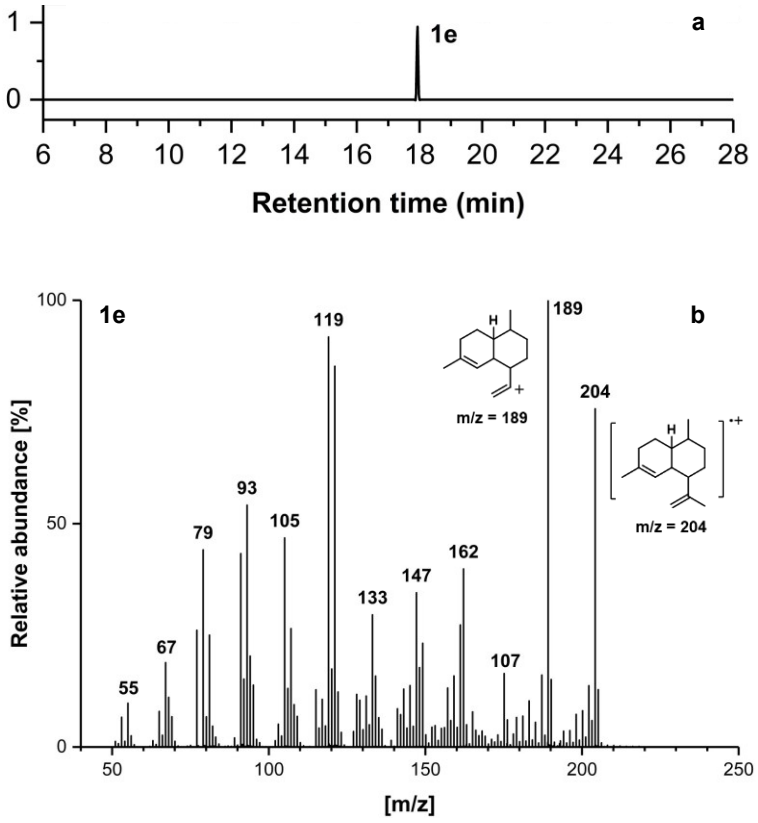


Figure 65. a) Total ion chromatogram of the pentane extractable products arising from incubation 5. b) Mass spectrum of the extracted product from incubation 5, showing the molecular ion of **1e** and a possible fragment ion of **1e** with m/z = 189.

Incubation 6. The major product in the pentane extracts after incubation of **1** with PRAS eluted at 18.12 min, as judged by GC-MS (Figure 66a), which corresponds to (+)-aristolochene (**1f**). Also, a minor product eluted at 19.10 min, which is (-)-germacrene A ((-)-**1a**). Their mass spectra are in accordance with the literature, with **1f** showing a molecular ion with m/z = 204 ($[M]^+$) and the major fragmentation peak with m/z = 189 ($[M - CH_3]^+$), Figure 66b.[93,95,97,138]

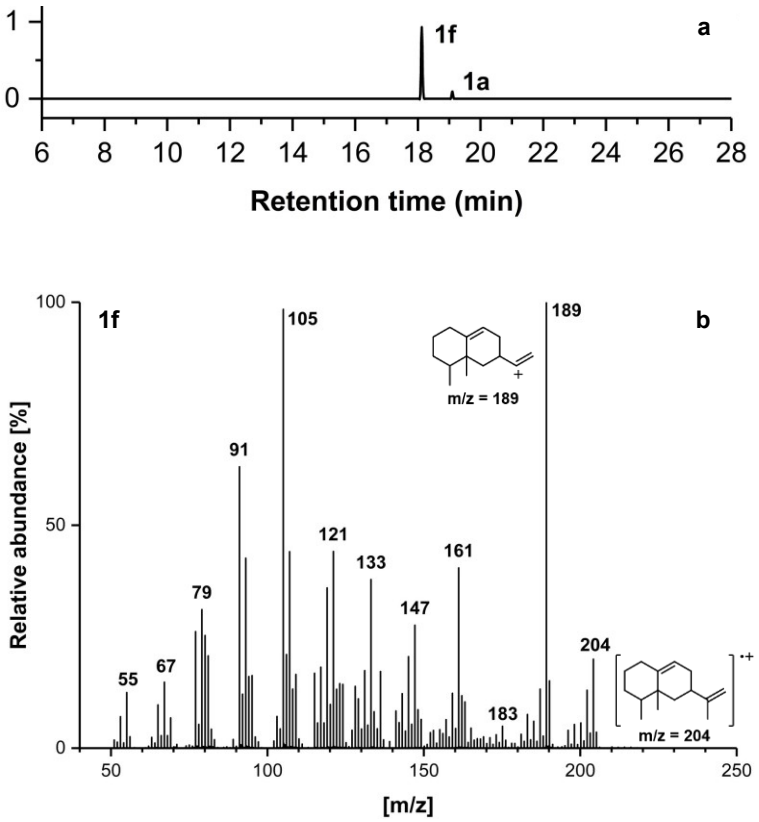


Figure 66. a) Total ion chromatogram of the pentane extractable products arising from incubation 6. b) Mass spectrum of the major product from incubation 6, showing the molecular ion of **1f** and a possible fragment ion of **1f** with m/z = 189.

Incubation 7. The incubation of SdS with **1** yielded a major product eluting at 20.02 min, and a minor product eluting at 20.88 min. The major compound corresponds to selina-4(15)-7(11)-diene (**1g**, Figure 67a), which has a mass spectrum showing a molecular ion with m/z = 204 ($[M]^+$), the base peak with m/z = 161 ($[M - C_3H_7]^+$), and a major fragment with m/z = 189 ($[M - CH_3]^+$). The minor compound corresponds to germacrene B (**1h**, Figure 67a). The mass spectrum of **1h** displays a base peak with m/z = 121. These match the published mass spectra for selina-4(15)-7(11)-diene and germacrene B, respectively, Figures 67b and 67c.[85,279]

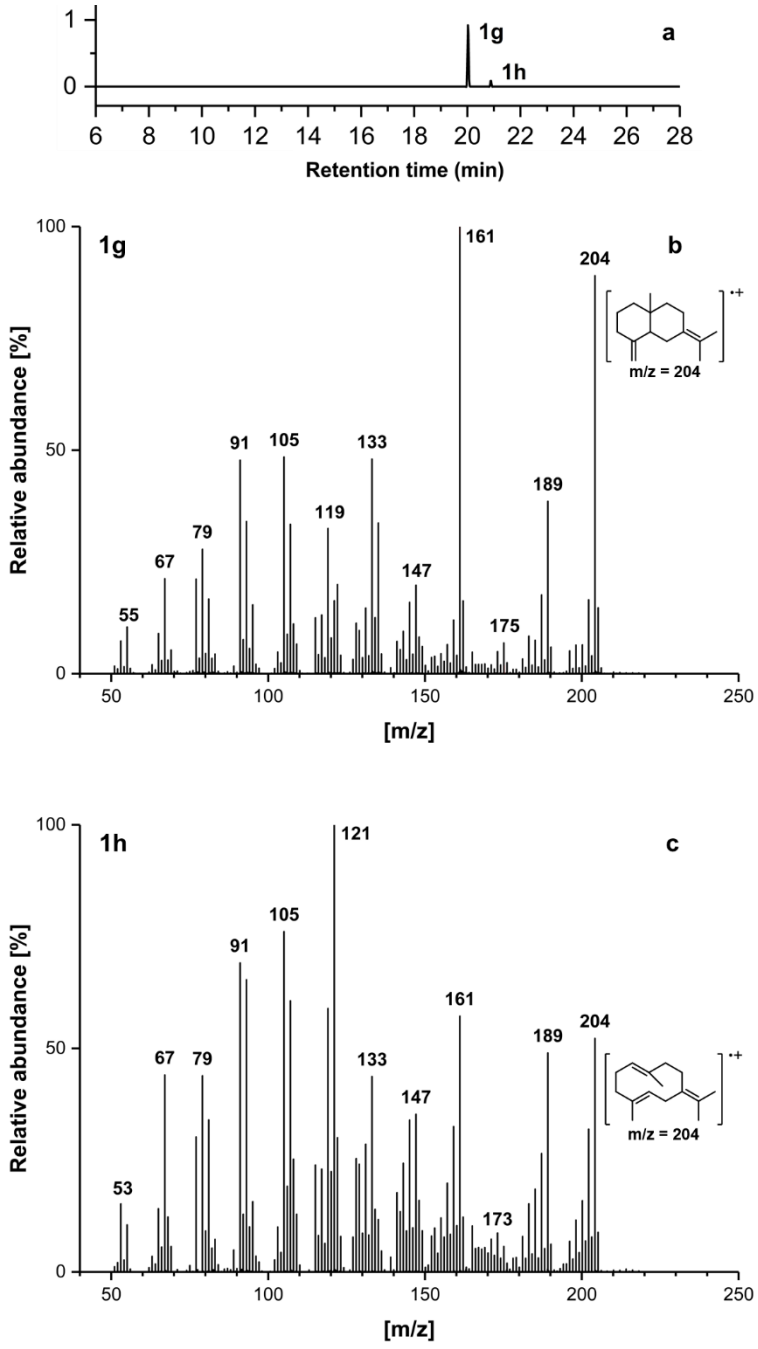


Figure 67. a) Total ion chromatogram of the pentane extractable products arising from incubation 7. b) and c) Mass spectra of the extracted products from incubation 7, showing the molecular ions of **1g** and **1h**, respectively.

Incubation 8. The incubation of FDP (**1**) with EpicS produced (+)-epicubenol, eluting at 23.44 min (**1i**, Figure 68a), and (-)-germacrene-4-ol, eluting at 21.75 min, (**1c**, Figures 68a and 63a) in a ratio 88:12, apart from other minor products (Figure 68a), as judged by GC-MS analysis of the obtained organic extracts. This is stated in section 2.3.1 and in accordance with the literature.[170]

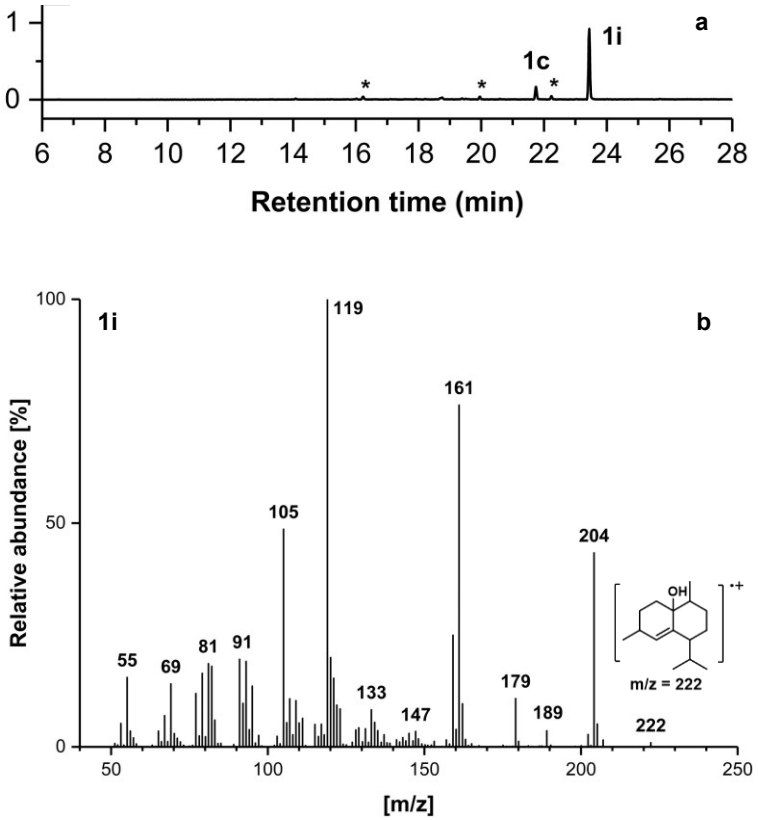


Figure 68. a) Total ion chromatogram of the pentane extractable products arising from incubation 8. b) Mass spectrum of the main product extracted from incubation 8, showing the molecular ion of **1i**.

Incubation 9. The incubation of FDP (**1**) with DCS-W279Y gave two products in the pentane extracts, (+)-δ-cadinene (compound eluting at 19.40 min, **1j**, Figure 69a), and (-)-germacradien-4-ol (product eluting at 21.75 min, **1c**, Figures 69a, 68a and 63a). All mass spectra observed were in fully agreement with those previously published.[74,125,275] The mass spectrum of **1j** displays a molecular ion peak with m/z = 204 ($[M]^+$) and the base peak is with m/z=161 ($[M - C_3H_7]^+$).

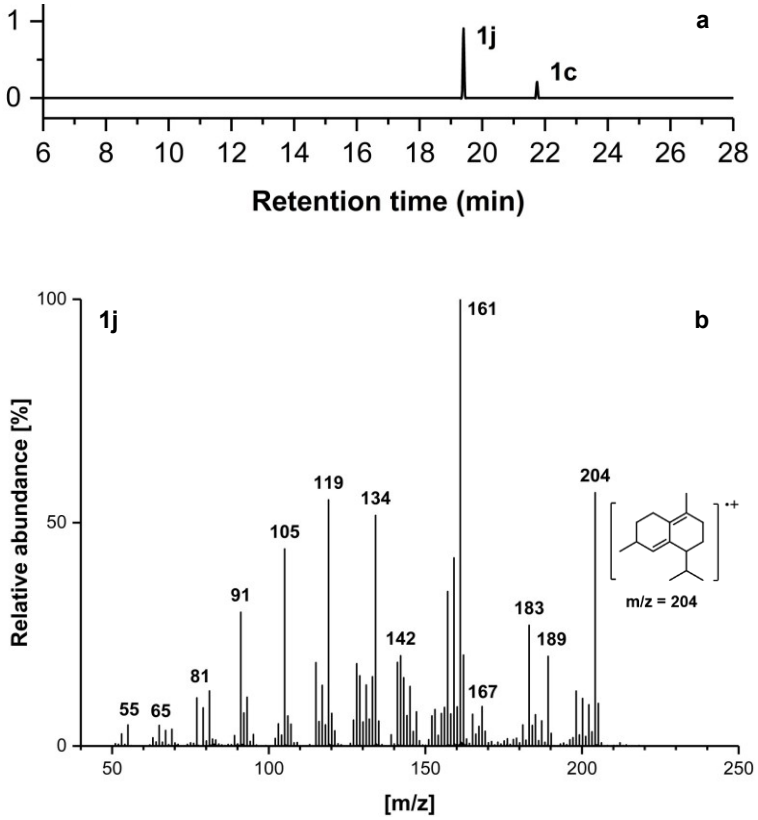
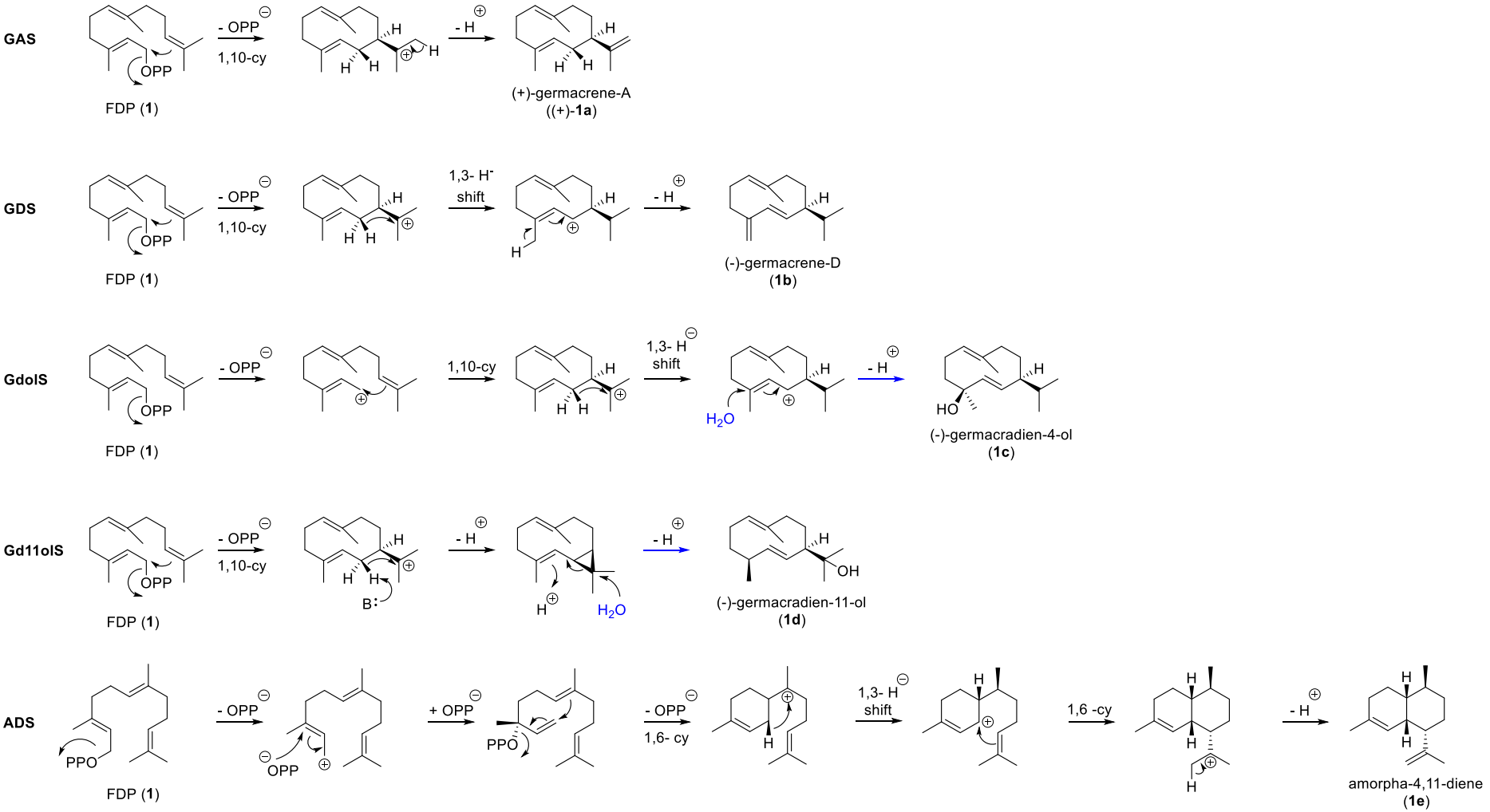


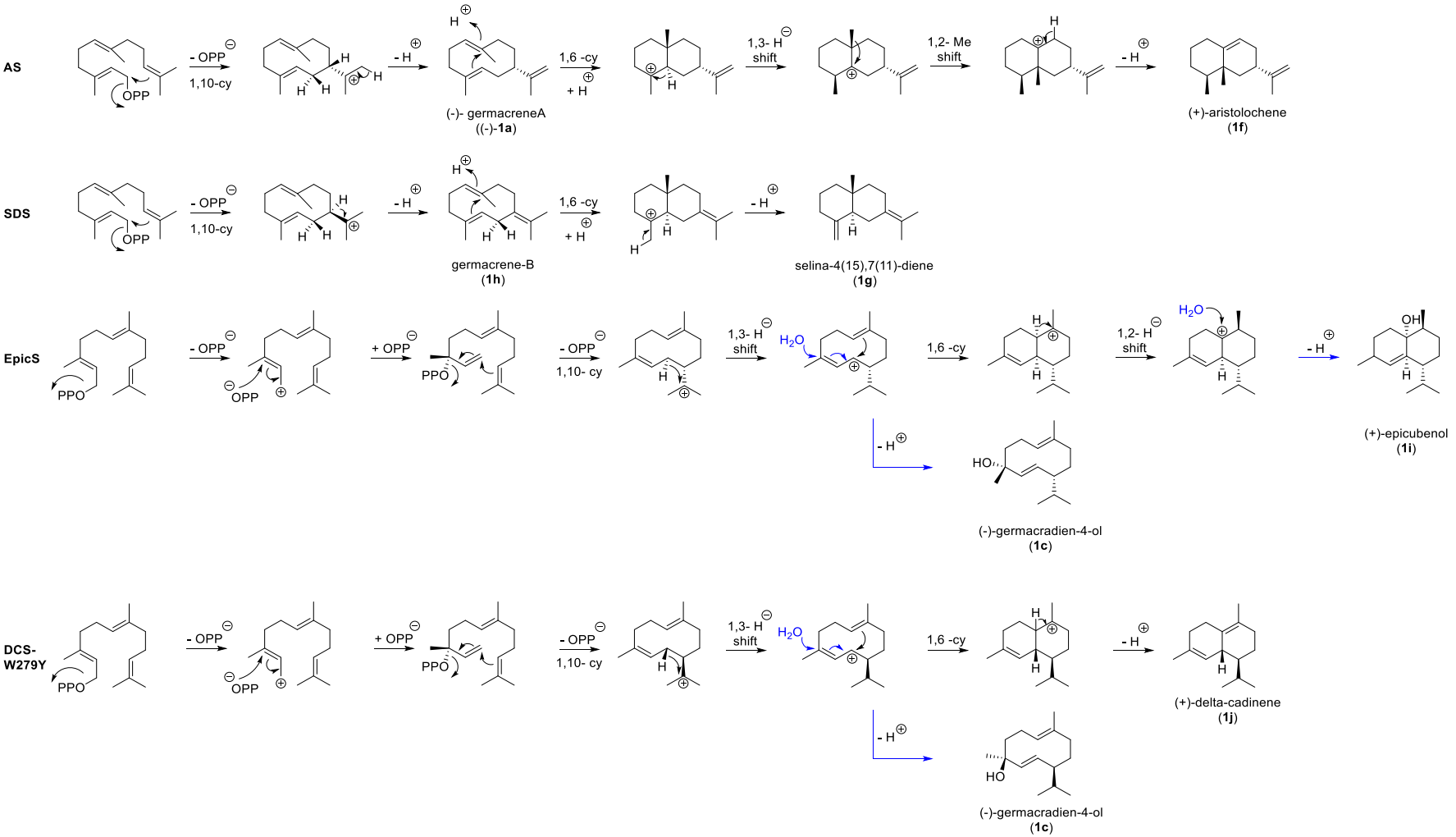
Figure 69. a) Total ion chromatogram of the pentane extractable products arising from incubation 9. b) Mass spectrum of the main extracted product from incubation 9, showing the molecular ion of **1j**.

Summary

The product profiles of these enzymes upon incubation with **1** were in excellent agreement with the literature under the incubation conditions used here (Section 7.1.17). This demonstrates the full functionality and purity of the prepared sesquiterpene synthases and allows for confident incubation studies with the FDP analogues (**2**-**8**). A summary of the different cyclisation pathways for FDP (**1**) with all nine sesquiterpene synthases as proposed in the literature is shown in Schemes 50 and 51.



Scheme 50. Proposed mechanisms for the reactions catalysed by GAS, GDS, GdolS, Gdd11olS (1,10-sesquiterpene synthases) and ADS (1,6-sesquiterpene synthase) upon incubation with **1**. Trace products (<5 %) are not included.



Scheme 51. Proposed mechanisms for the reactions catalysed by AS, SDS, EPICS and DCS-W279Y (1,10- + 1,6-sesquiterpene synthases) upon incubation with **1**. Trace products (<5 %) are not included.

4.3. Substrate behaviour of (2*E*,6*Z*)-6F-FDP (2)

The use of (2*E*,6*Z*)-6F-FDP (**2**) with sesquiterpene synthases has been previously reported. **2** has been converted by GAS and GDS to the corresponding (2*E*,6*Z*)-6F- germacrene A ((+)-**2a**) and (2*E*, 6*Z*)-6F-germacrene D (**2b**) modified sesquiterpene analogues, respectively (Figure 70).[123,130] **2** is also a substrate for aristolochene synthases TEAS[280] and PRAS[122], and (-)-**2a** was identified as the main enzymatic product. In addition, steady-state kinetic measurements with both (2*E*,6*E*)-FDP (**1**) and (2*E*,6*Z*)-6F-FDP (**2**) demonstrated that the fluorine replacement at position C6 of **1** has negligible effects on enzyme binding, substrate orientation and ionisation. The initial 1,10-cyclisation catalysed by TEAS maintained equal affinity for both, non-fluorinated and fluorinated substrates.[280] In contrast, (2*E*,6*Z*)-6F-FDP (**2**) was not a substrate for DCS, and kinetic measurements demonstrated that it acts as a potent competitive inhibitor, which agrees with the abolishment of an enzymatic 1,6-ring closure pathway occurring in DCS.[281] When fluorine is inserted at the C6 position, it behaves as a spectator in the cyclisation reaction catalysed by 1,10-cyclases, but is directly involved in the mechanisms of 1,6-cyclases, and might inhibit their natural cyclisation. Presumably, this happens due to destabilisation by C6-fluorine of the positive charge at C7 in 1,6-cyclisations, but it can also reduce the nucleophilicity of the C6-C7 double bond (Sections 1.5.4 and 3.1).

With these precedents, the incubations of GdolS, Gd11olS (1,10-sesquiterpene synthases), ADS (1,6-sesquiterpene synthase), EpicS, SdS and DCS-W279Y (1,10- & 1,6-sesquiterpene synthases) with (2*E*,6*Z*)-6F-FDP (**2**) are covered here for the first time, with the aim of expanding the pool of fluorinated analogues and to continue gaining insights about C6- fluorination influence in catalysis. In addition, the incubations of **2** with GAS, GDS (1,10- sesquiterpene synthases), and PRAS (1,10- & 1,6-sesquiterpene synthase) served as controls for product identification (Figure 70).

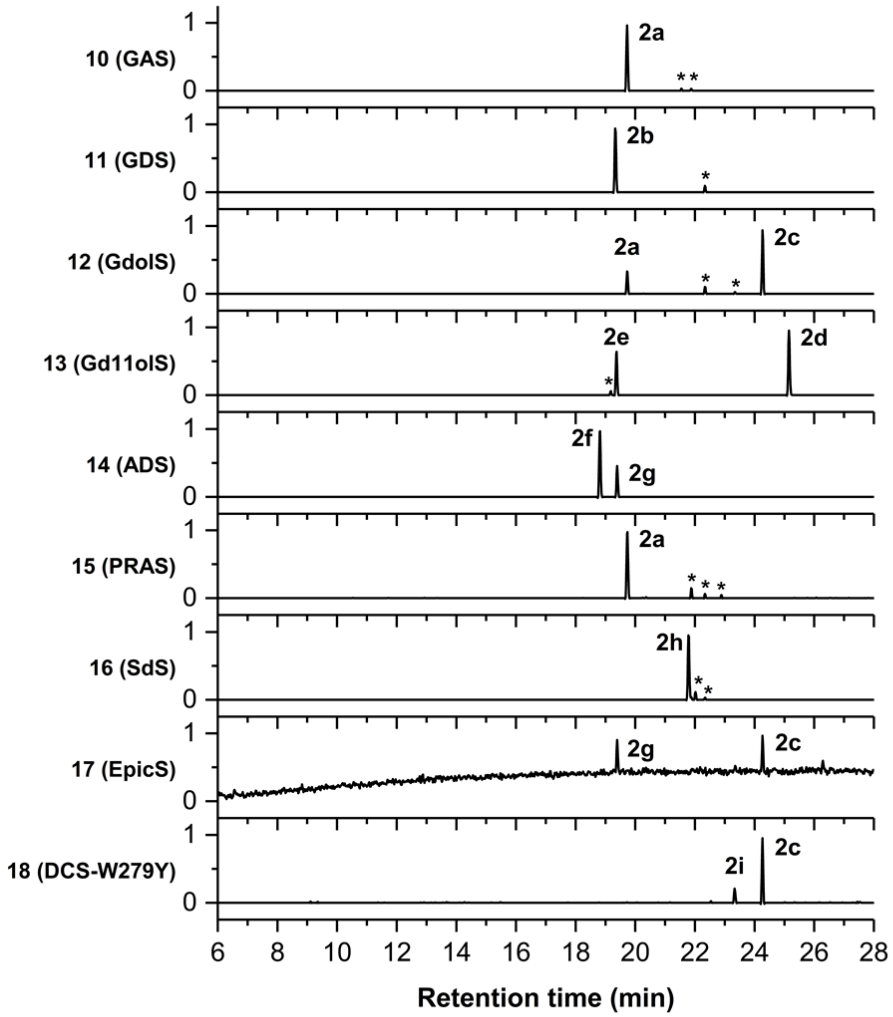


Figure 70. Total ion chromatograms of the pentane extractable products arising after incubations of (2*E*,6*Z*)-6F-FDP (**2**) with nine sesquiterpene synthases. Selected minor compounds are labelled with an asterisk.

1,10- Sesquiterpene synthases (GAS, GDS, GdolS and Gd11olS)

Incubation 10. The incubation of **2** with GAS yielded a single major product in the pentane extracts (eluting at 19.72 min, Figure 71a), as judged by GC-MS. Its mass spectrum shows a molecular ion with m/z= 222 ($[M]^+$), indicating the presence of fluorine in the structure, and fragment peaks with m/z= 207 ($[M - CH_3]^+$), and 202 ($[M - HF]^+$), consistent with the loss of CH_3 or HF respectively from the parent fluorinated compound, Figure 71b. The base peak has m/z= 68. This mass spectrum is fully in agreement with that previously reported for (2*E*,6*Z*)-6F-germacrene A, (+)-**2a**.[123]

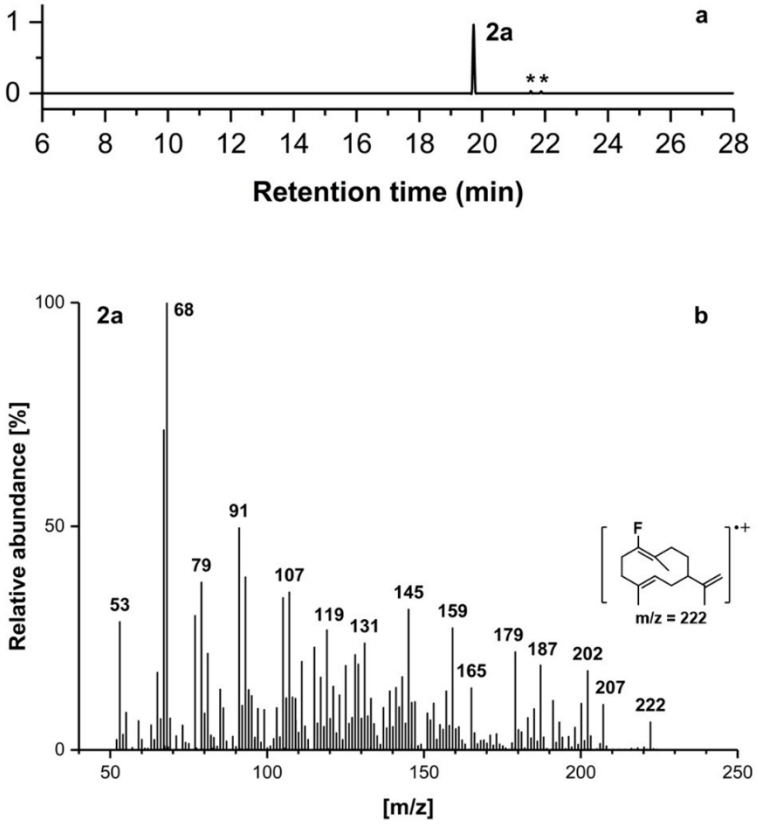
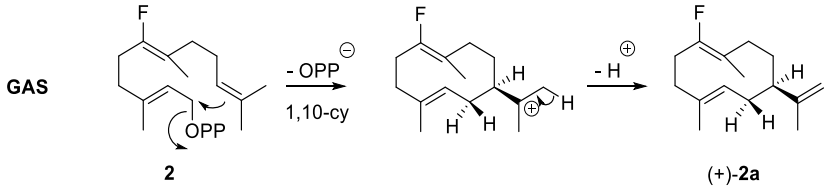


Figure 71. a) Total ion chromatogram of the pentane extractable products arising from incubation 10. b) Mass spectrum of the main extracted product from incubation 10, showing the molecular ion of (+)-**2a**.



Scheme 52. Proposed reaction mechanism catalysed by GAS upon incubation with **2**.[123]

Incubation 11. The incubation of GDS with **2** gave rise to a major compound in the pentane extracts, eluting at 19.33 min in the chromatogram (Figure 72a). The mass spectrum of this compound shows a molecular ion with m/z = 222 ($[M]^+$), Figure 72b, agreeing with the presence of a fluorinated sesquiterpene. In addition, it has a fragment peak with m/z = 179 ($[M - C_3H_7]^+$), and a base peak with m/z = 159 ($[M - HF - C_3H_7]^+$), which indicates the loss of an isopropyl group and HF from the observed molecular ion. The mass spectrum of this compound is in agreement with that of the previously reported (2*E*,6*Z*)-6F-germacrene D (**2b**).[123]

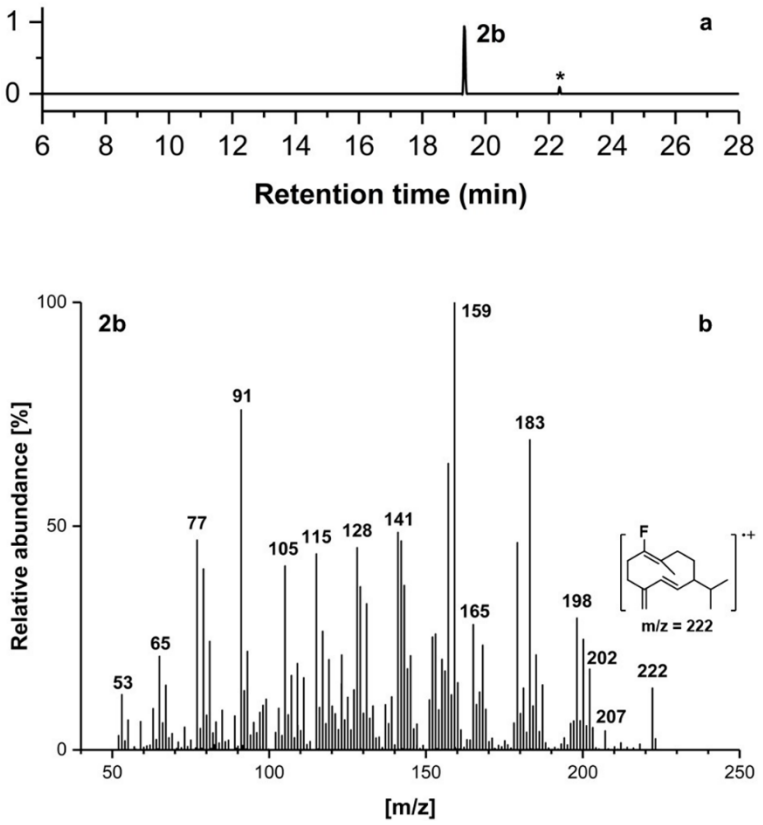
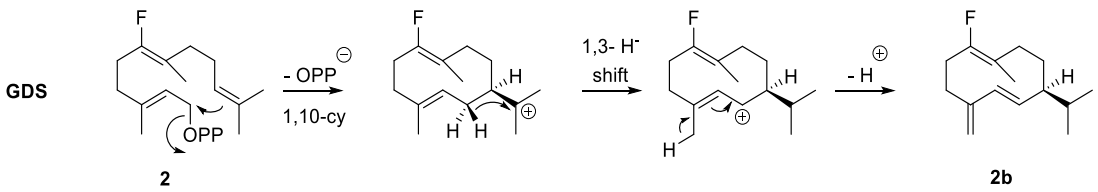


Figure 72. a) Total ion chromatogram of the pentane extractable products arising from incubation 11. b) Mass spectrum of the main extracted product from incubation 11, showing the molecular ion of **2b**.



Scheme 53. Proposed reaction mechanism catalysed by GDS upon incubation with **2**.[123]

Incubation 12. The incubation of **2** with GdolS gave a major organic extractable product eluting at 24.27 min, as judged by GC-MS analysis (Figure 73a). In addition, three other compounds were observed eluting at 19.72 min, 22.33 min and 23.35 min in the chromatogram, albeit in minor proportions. The major product was hypothesised to be 6F-germacradien-4-ol analogue (**2c**). As with GAS and GDS, GdolS should be able to direct the 1,10-cyclisation and generate a fluorinated germacrenyl ionic intermediate upon incubation with **2**. The fluorine at position C6 supposedly does not participate in the rest of the enzymatic 1,10- cyclisation mechanism (Scheme 54). The molecular ion detected is with m/z = 222, rather than the expected molecular ion for **2c** with m/z = 240 ($[M]^+$), which indicates that the hydroxyl group was not detected ($[M - H_2O]^+$). Nevertheless, a peak with m/z = 220 ($[M - HF]^+$) was observed, indicative of the loss of HF from the parent compound. The observed fragment ions with m/z = 207 ($[M - H_2O - CH_3]^+$) and 202 ($[M - H_2O - HF]^+$) are consistent with the loss of CH_3 or HF respectively from the observed dehydrated compound. The base peak with m/z = 159 ($[M - H_2O - HF - C_3H_7]^+$) can arise from further loss of an isopropyl group (Figure 73b). To probe this hypothesis, compound **2** was incubated with EpicS (incubation 17) and DCS-W279Y (incubation 18) because both enzymes produce germacradien-4-ol (**1c**) as a by-product upon incubation with FDP (**1**). In fact, **2d** was also obtained as the major product in these incubations, which further validated this proposal.

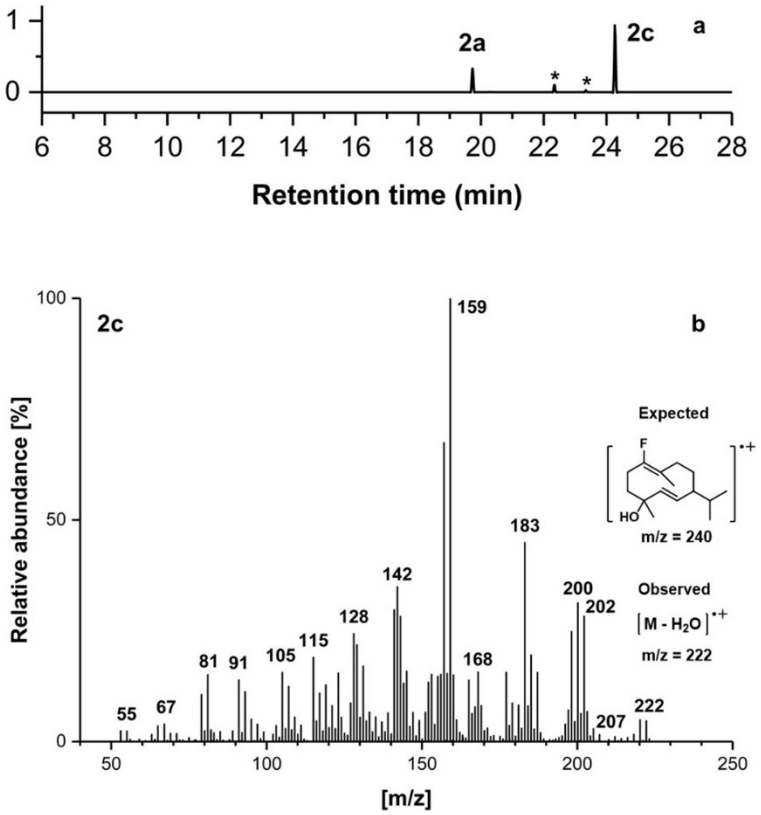
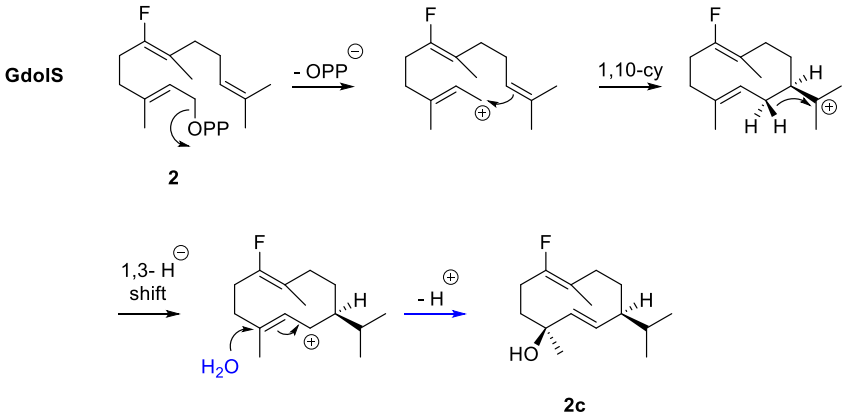


Figure 73. a) Total ion chromatogram of the pentane extractable products arising from incubation from incubation 12. b) Mass spectrum of the main extracted product from incubation 12, showing the molecular ion of **2c** (expected).



Scheme 54. Proposed reaction mechanism catalysed by GdolS upon incubation with **2**.

The compound eluting at 19.72 min has the same retention time and an identical mass spectrum as **2a** (Figure 71b). This is the same product observed upon the incubation of **2** with GAS and PRAS (incubations 10 and 15).

Incubation 13. Two compounds with similar ratio were detected in the organic extracts from the incubation of **2** with Gd11olS, as judged by GC-MS (Figure 74a). The mass spectrum of the product eluting at 25.15 min shows a molecular ion of m/z = 240 ($[M]^+$, indicating the presence of a hydroxyl group, and major fragment peaks with m/z = 222 ($[M - H_2O]^+$), 202 ($[M - H_2O - HF]^+$) and 187 ($[M - H_2O - HF - CH_3]^+$), agreeing with loss of H_2O from the parent compound followed by loss of HF from the dehydrated ion and a further loss of CH_3. The base peak is at m/z = 159 ($[M - H_2O - HF - C_3H_7]^+$), corresponding to the fragmentation of the isopropyl group from the observed fragment peak with m/z = 202 (Figure 74b). This product was proposed to be (2*E*,6*Z*)-6F-germacrene-11-ol (labelled as **2d**) because, as shown in incubations 10-12, fluorine at C6 does not affect the enzymatic mechanism of germacrene-like synthases (Scheme 55).

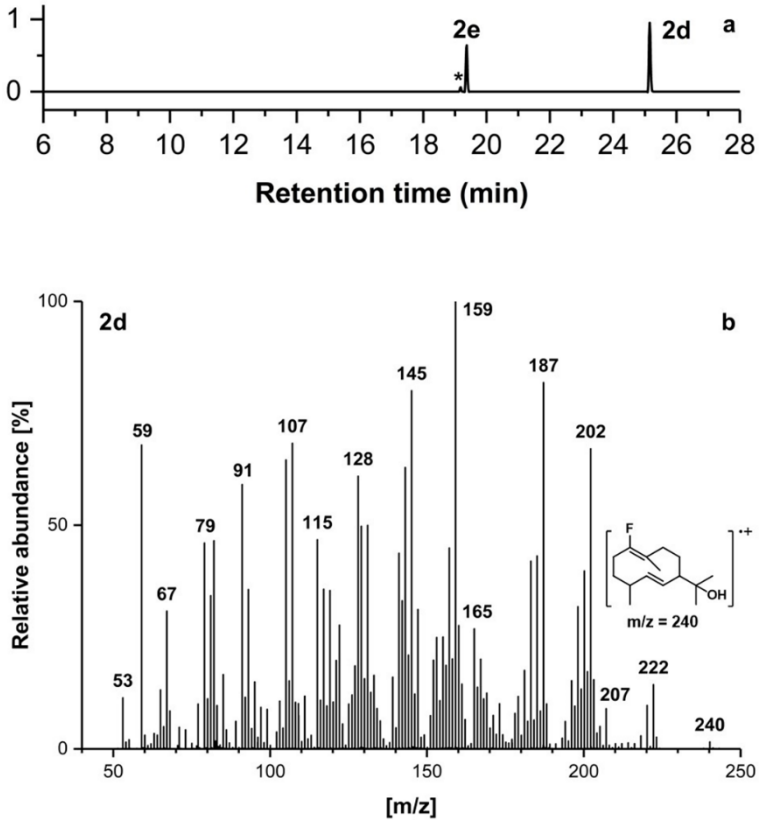
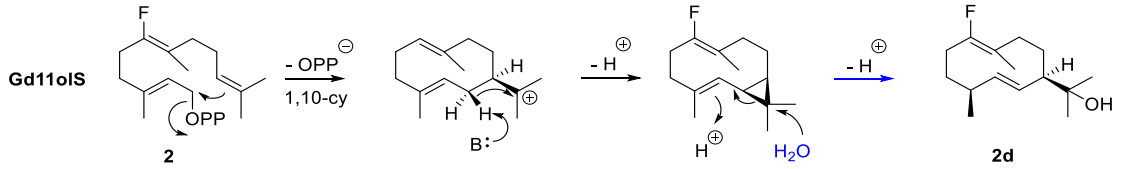


Figure 74. a) Total ion chromatogram of the pentane extractable products arising from incubation 13. b) Mass spectrum of the product eluting at 25.15 min obtained from incubation 13, showing the molecular ion of **2d**.



Scheme 55. Proposed reaction mechanism catalysed by Gd11olS to transform **2** into **2d**.

The compound eluting at 19.37 min (**2e**), is a fluorinated compound because its mass spectrum shows a molecular ion with m/z= 222 ($[M]^+$) and major fragment peaks with m/z= 207 ($[M - CH_3]^+$) and 202 ($[M - HF]^+$), Figure 75. The predominant peaks were at m/z = 121 and 107. In comparison with the previously reported incubation of Gd11olS with FDP (**1**), where germacrene D (**1b**) is generated as a coproduct,[274] it was speculated that this compound could be (2*E*,6*Z*)-6F-germacrene D (**2b**). However, **2e** and **2b** do not share their retention times and their mass spectra do not match. **2e** remains unidentified.

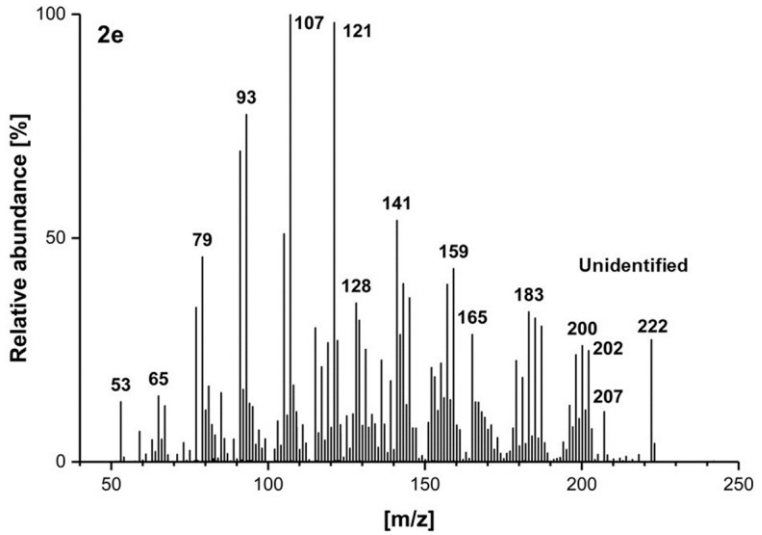
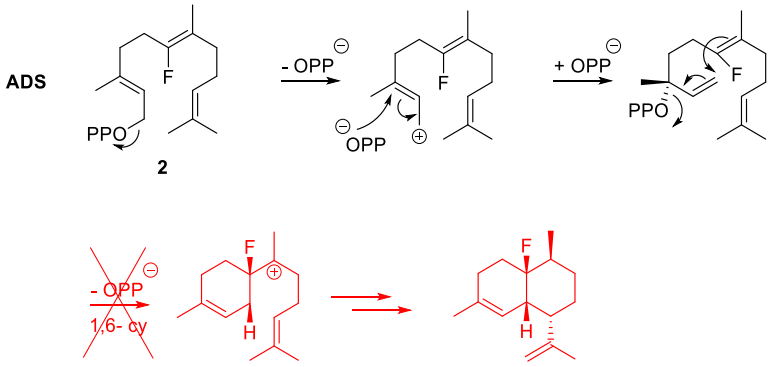


Figure 75. Mass spectrum of the product eluting at 19.37 min (**2e**) obtained from incubation 13, which remains unidentified.

In summary, the incubations of **2** with 1,10-sesquiterpene synthases generated the corresponding 6F-germacrene analogues (**2a** - **2d**) of the natural observed products (**1a** - **1d**). In addition to the previously obtained **2a** and **2b**, we have prepared **2c** and **2d**, as novel fluorinated sesquiterpene alcohol compounds to date. GAS, GDS and GdolS transform **2** with high fidelity into **2a**, **2b** and **2c**, respectively. As for Gd11olS, apart from **2d**, a second major product is observed (**2e**). The enzymatic response governing the conversion of **2** to **2e** by Gd11olS was not investigated, other than this, Gdol11S produces a (yet) unidentified fluorinated sesquiterpene in relatively good quantity to be scalable, which arises from an alternate pathway as compared with the natural conversion of FDP (**1**).

1,6- Sesquiterpene synthase (ADS)

Incubation 14. The incubation of **2** with ADS gave two products in the pentane extracts, as judged by GC-MS. The major product eluted at 18.82 min (**2f**, 69%), together with a minor product eluting at 19.39 min (**2g**, 31%), Figure 76a. The mass spectra of both compounds show their molecular ions with m/z = 222 ($[M]^+$) and fragment peaks with m/z = 207 ($[M - CH_3]^+$) and 202 ($[M - HF]^+$), consistent with the presence of fluorinated compounds, Figures 76b and 76c. ADS usually catalyses an initial 1,6-cyclisation of the farnesyl cation, and this event might be abolished by positioning fluorine at C6 as a result of the destabilisation of the forming carbocation and/or reducing the nucleophilicity of the C6-C7 double bond (Scheme 56), as observed with DCS-WT.[281] Because of this, it was expected to have no sesquiterpene production or the presence of aberrant farnesenes and/ or farnesol counterparts of **2**. Alternatively, ADS might catalyse different cyclisation pathways, and 1,7- , 1,10- or 1,11-cyclisations may occur.



Scheme 56. Proposed interception of the reaction mechanism catalysed by ADS upon incubation with **2**.

The presence of 6F-germacrenes (A, D or B) was ruled out through comparison of the retention time and mass spectra of the pentane extractable products with those of authentic standards obtained upon incubations of **2** with GAS, GDS and SdS (incubations 10, 11 and 16). Also, the presence of (*E*,*Z*)- 6F-farnesol was ruled out, as it was compared with injection of authentic synthetic standard. The products observed could be farnesenes. In fact, the mass spectrum of **2f** resembles that observed for 8-methoxy-E-β-farnesene,[282] which is a farnesene compound arising from the incubation of *E*-β-farnesene synthase (EBFS) with 8-methoxy-FDP.

The mass spectrum of **2g** was cryptic when analysed. It shows a base peak with m/z = 187 ([M - HF - CH_3], which might indicate the presence of a bicyclic structure. In fact, it was found in the literature that amorpha-4,11-diene aldehyde[131] has a similar mass spectrum if compared with **2g**, which means these could share similar structures. These compounds remain unidentified at the time of writing.

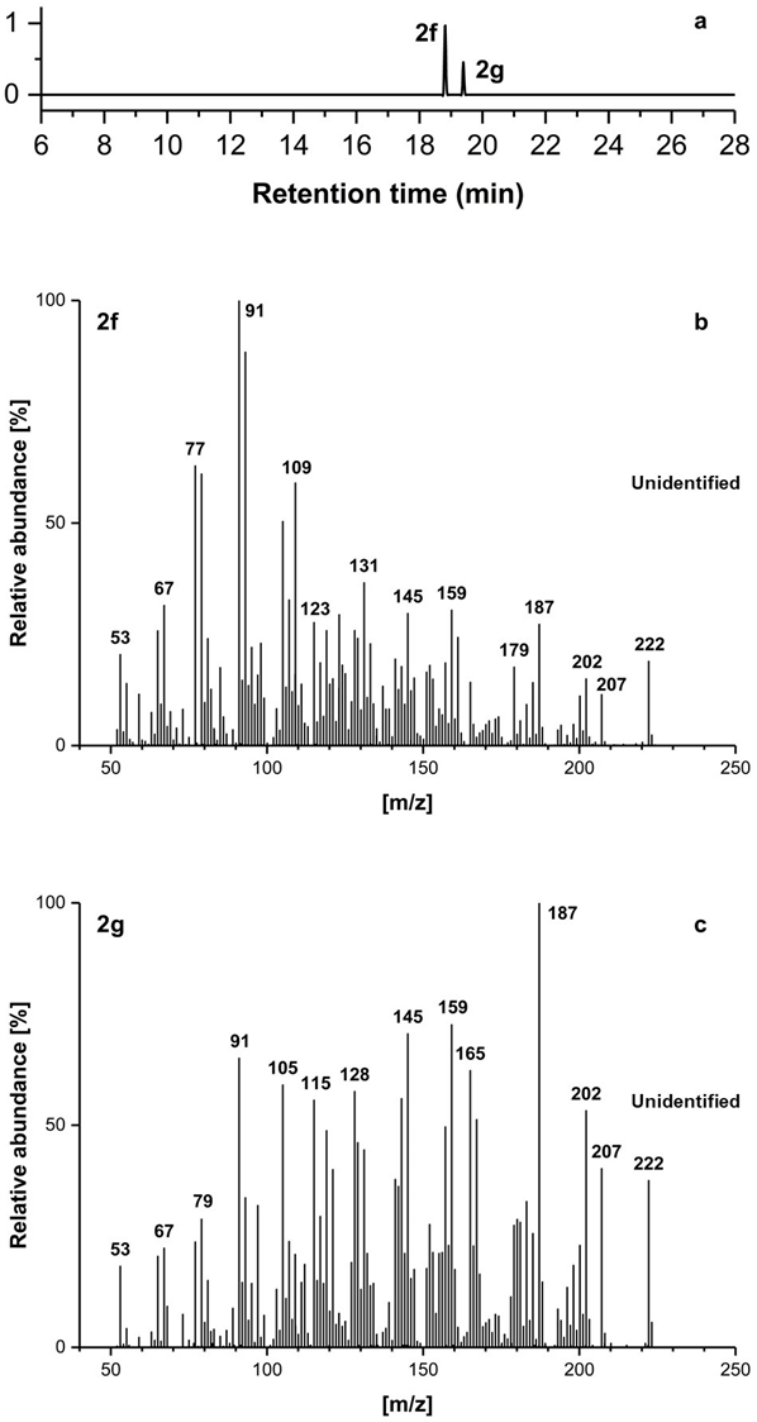


Figure 76. a) Total ion chromatogram of the pentane extractable products arising from incubation 14. b) Mass spectrum of the extracted product eluting at 18.32 min (labelled as **2f**) from incubation 14, which remains unidentified. c) Mass spectrum of the extracted product eluting at 19.39 min (labelled as **2g**) from incubation 14, which remains unidentified.

1,10- + 1,6- Sesquiterpene synthases (PRAS, SDS, EpicS, DCS-W279Y)

Incubation 15. The incubation of PRAS with **2** gave a major organic extractable product eluting at 19.72 min (Figure 77), which was identified as (2*E*,6*Z*)-6F-germacrene A ((-)-**2a**) in agreement with incubation 10 and previously published results.[122] This was shown to be in agreement with the generation of putative 6F-eudesmane cation, in that the fluorine at C6 can destabilise the forming carbocation at C7 (if a stepwise 1,6- cyclisation mechanism occurs) and/ or reduces the electron density of the C6-C7 double bond (if a concerted 1,6-cyclisation mechanism occurs) and prevents the 1,6-cyclisation reaction (Scheme 57).

Minor products eluted at 21.89 min, 22.33 min, and 22.89 min. These were not investigated here, but their mass spectra are in concordance with the presence of fluorinated sesquiterpenes, also a similar gas chromatogram has been described in the literature upon the incubation of **2** with PRAS.[122]

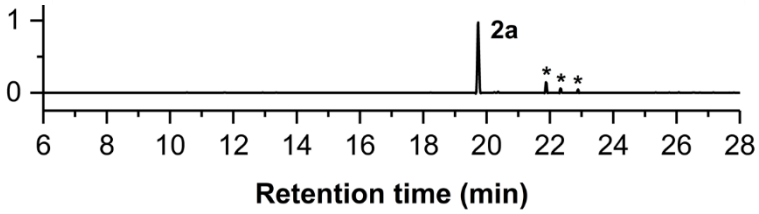
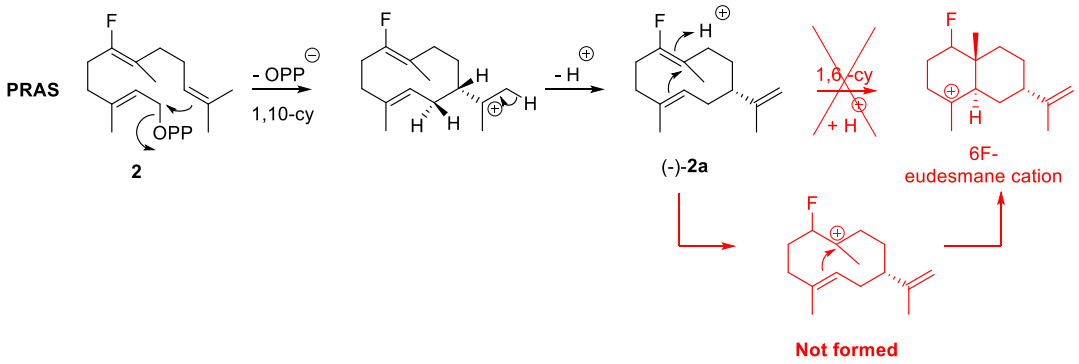


Figure 77. a) Total ion chromatogram of the pentane extractable products arising from incubation 15.



Scheme 57. Proposed interception of the reaction mechanism catalysed by PRAS (red), and production of (-)-**2a** (black) upon incubation with **2**.

Incubation 16. The incubation of SdS with **2** yielded a major extractable compound, which eluted at 21.78 min (labelled as **2h**, 85%, Figure 78a). This compound has a mass spectrum displaying a molecular ion with m/z = 222 ($[M]^+$) and a major fragment with m/z = 207 ($[M - CH_3]^+$), which are consistent with the presence of a monofluorinated sesquiterpene (Figure 78b). Based on previous results upon the incubation of **2** with PRAS (incubation 15), this compound is proposed to be (2*E*,6*Z*)-6F-germacrene B (**2h**). This is made because germacrene B (**1h**, incubation 7) is a putative neutral intermediate generated during the reaction catalysed by SdS upon incubation with **1** In turn, SdS catalyses the conversion of **2** to 6F- germacrene B (**2h**) but cannot proceed further in the reaction mechanism due to the position of the fluorine, which either can destabilise the forming carbocation (if a stepwise 1,6- cyclisation mechanism occurs), or reduce the electron density of the C6-C7 double bond (if a concerted 1,6-cyclisation mechanism occurs), leading to the accumulation of (2*E*,6*Z*)-6F-germacrene B (**2h**, Scheme 58).

In fact, the mass spectra of the postulated 6F-germacrene B (**2h**) and germacrene B (**1h**) contain similar fragmentation patterns when compared, sharing a base peak with m/z = 121. The fluorine atom situated at the C6 position hinders the second 1,6-cyclisation by SdS, leading to the accumulation of 6F-germacrene B (**2h**).

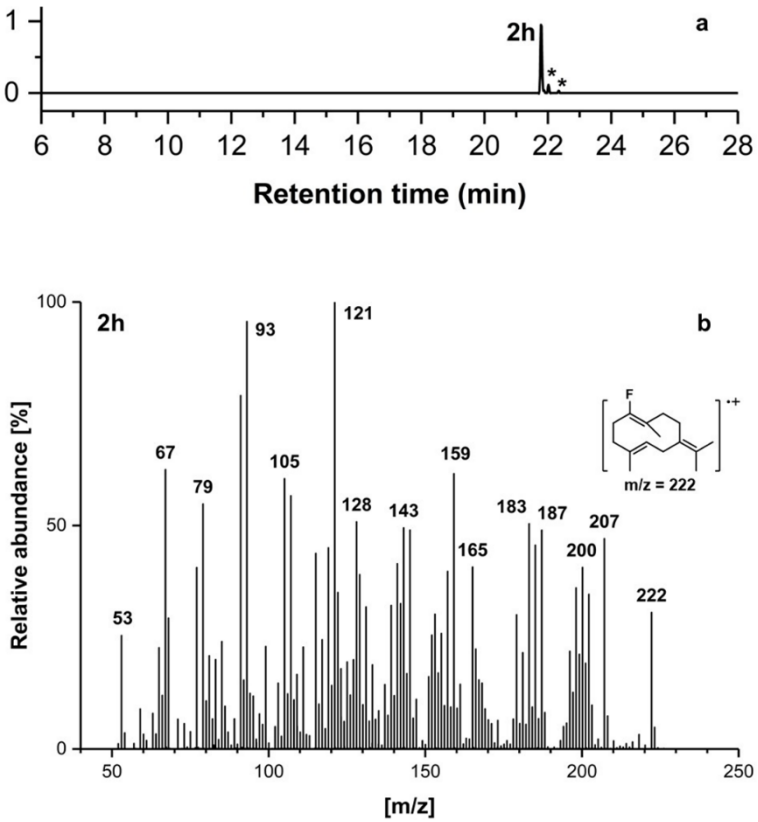
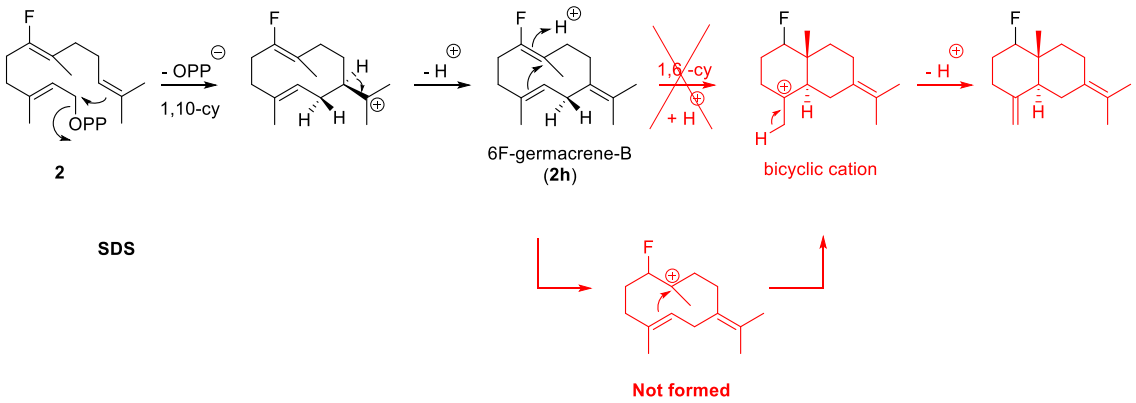


Figure 78. a) Total ion chromatogram of the pentane extractable products arising from incubation 16. b) Mass spectrum of the major extracted product eluting at 21.78 min (**2h**), obtained from incubation 16, showing the molecular ion of (2E,6Z)- 6F- germacrene B (**2h**).

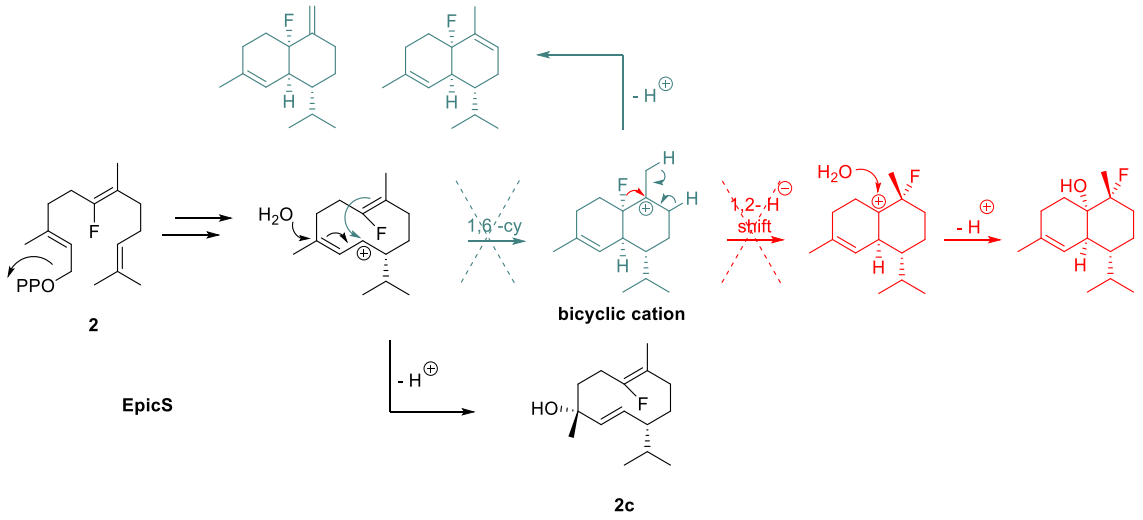


Scheme 58. Proposed interception of the reaction mechanism catalysed by SdS (red), and production of **2h** (black) upon incubation with **2**.

Incubation 17. The incubation of **2** with EpicS gave two compounds in the pentane extracts in low yield (Figure 79a). The compound eluting at 24.27 min is proposed to be 6F-germacradien-4-ol (**2c**), in concordance with incubations 12 and 18, where **2c** is the main product upon incubation of **2** with GdolS and DCS-W279Y respectively.

This result agrees with the postulated mechanism catalysed by EpicS, in that the 1,6-cyclisation might be hindered due to fluorine electronic destabilisation of the forming adjacent carbocationic intermediate and/or reduction of the nucleophilicity at the C6-C7 double bond needed for the 1,6-intramolecular attack, and thus 6F-germacradien-4-ol is accumulated by water quenching the 10- membered allylic intermediate, Scheme 59.

The second product eluted at 19.39 min (labelled as **2g**) and was difficult to unambiguously identify. The mass spectrum of this compound displays a molecular ion peak with m/z = 222 ($[M]^+$), and the characteristic fragment peaks with m/z = 207 ($[M - CH_3]^+$) and m/z = 202 ($[M - HF]^+$), suggesting it was a non-hydroxylated fluorinated sesquiterpene. The most stable fragment is with m/z = 187 ($[M - HF - CH_3]^+$), which might be indicative of the presence of a bicyclic product (figure 79b). Based on the proposed catalytic mechanism underwent by EpicS upon incubation with FDP (**1**), if EpicS was able to direct the 1,6-cyclisation, the subsequent 1,2-hydride shift is not possible because the hydrogen has been substituted with fluorine and the kinetic properties have changed. With this, the bicyclic carbocationic intermediate before the 1,2-hydride shift might loss a proton to generate the final product. A similar scenario was found at the incubation of **2** with ADS (incubation 13) and producing the same compound, **2g**. It is postulated that this may be novel pathway, which results in the production of 6F-bicyclic compounds. Nevertheless, the catalytic possibilities and the identity of this compound need to be further investigated.



Scheme 59. In red, proposed interception of the mechanisms catalysed by EpicS upon incubation with **2**. In black, proposed reaction mechanism for the conversion of **2** to **2c** by EpicS. In grey, alternative postulated mechanisms for the conversion of **2** into possible bicyclic products by EpicS.

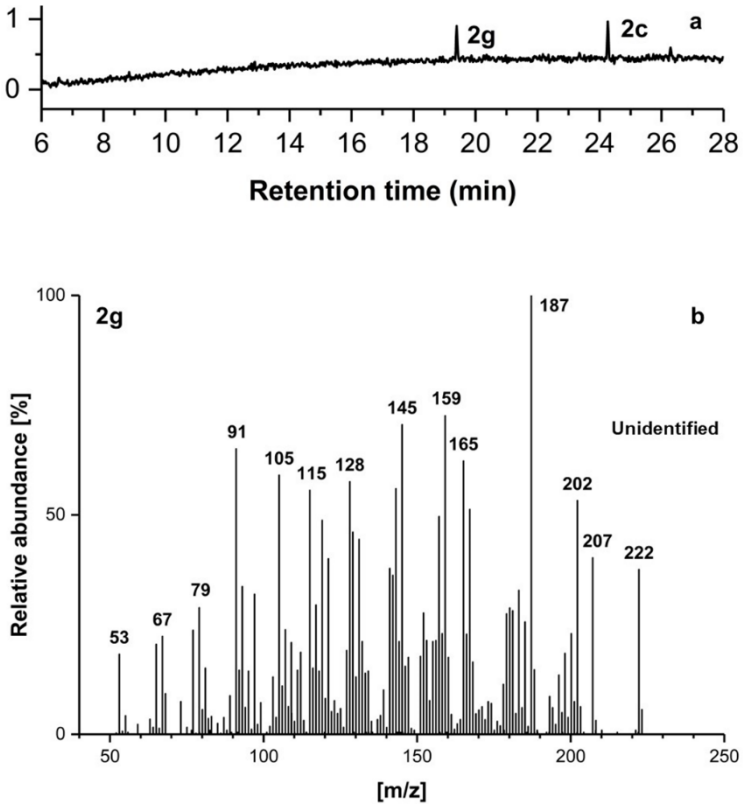


Figure 79. a) Total ion chromatogram of the pentane extractable products arising from incubation 17. b) Mass spectrum of the extracted product eluting at 19.39 min (labelled as **2g**) obtained from incubation 17, which remains unidentified.

Incubation 18. The incubation of **2** with DCS-W279Y gave two products in the pentane extracts (Figure 80a). The major product eluted at 24.27 min (72%), which was identical to that obtained after incubation of **2** with GdolS or EpicS and was identified as (2*E*,6*Z*)-6F-germacrene-4-ol (**2c**). The minor product eluted at 23.35 min (labelled as **2i**, 28%). The identity of this product remains unknown, however, a fluorinated δ-cadinene analogue has been ruled out (Scheme 60) because it has been already proven that **2** is a potent competitive inhibitor for DCS-WT catalysis.[281] The mass spectrum of this compound suggests that it is a monofluorinated sesquiterpene because it displays a molecular ion with m/z = 222 ($[M]^+$), and fragments with m/z = 207 ($[M - CH_3]^+$), and m/z = 202 ($[M - HF]^+$), Figure 80b. Beside this, it contains a major fragment peak with m/z = 220 that might result from the loss of HF from a fluorinated sesquiterpene alcohol (although the corresponding molecular ion with m/z = 240 is not observed). The fragment with m/z = 220 could also be a reduced form of the parental pure hydrocarbon structure. Either way, the base peak with m/z = 159 might indicate the presence of a germacrene analogue, which could arise from quench (proton loss or water quench) of a germacrene intermediate along the reaction cascade (Scheme 60). The identity of this product remains unknown by the time of this writing.

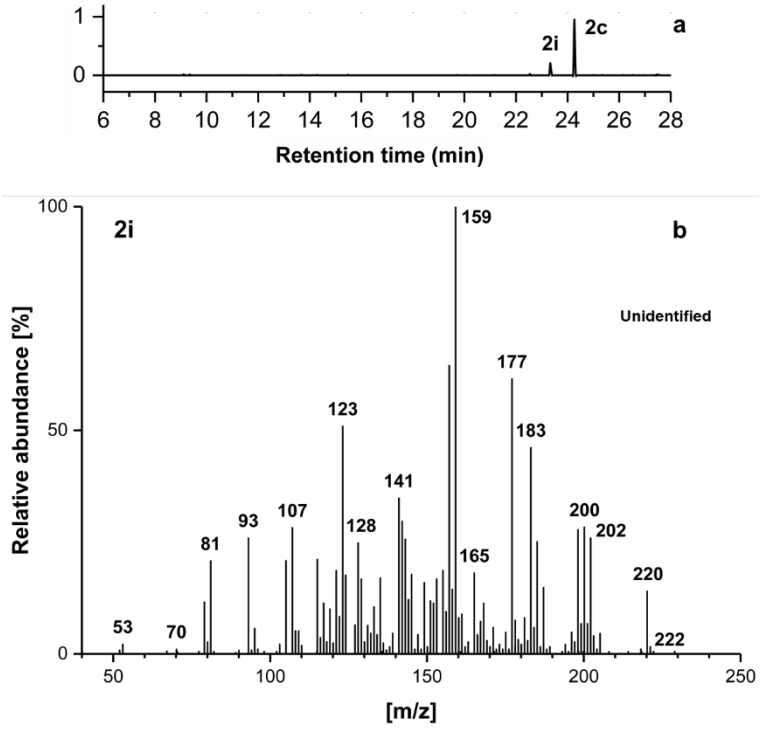
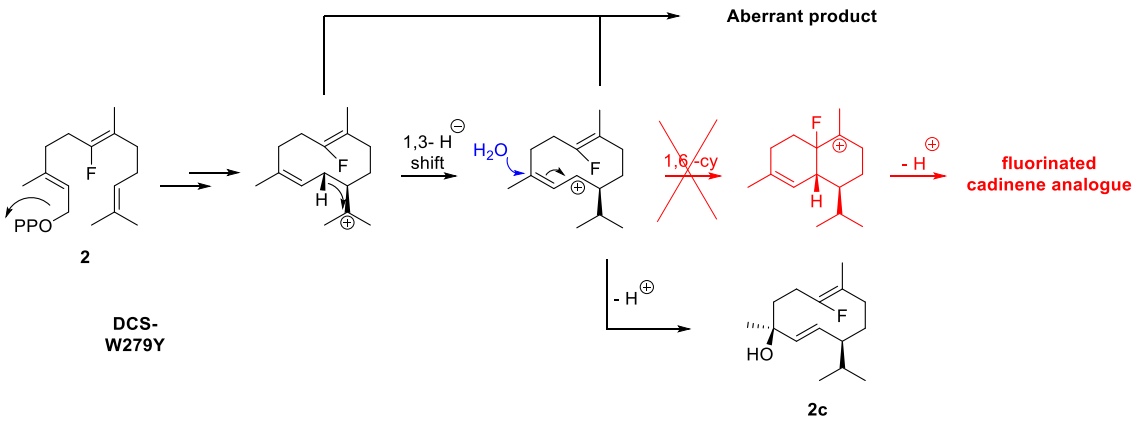


Figure 80. a) Total ion chromatogram of the pentane extractable products arising from incubation 18. b) Mass spectrum of the extracted product eluting at 23.35 min (labelled as **2i**) obtained from incubation 18, which remains unidentified.



Scheme 60. In red, proposed interception of the reaction mechanism catalysed by DCS-W279Y upon incubation with **2**. In black, proposed reaction mechanisms for the conversion of **2** to **2c** by DCS-W279Y, and other possible germacrene product.

In summary, it is proposed here that the major products arising after the incubation of these 1,10- + 1,6-sesquiterpene synthases with **2** are all germacrene analogues. These products seem to arise as abortive products due to the presence of the fluorine atom in the substrate, and in support of the proposed catalytic mechanisms of these enzymes. E.g., **2a** with PRAS, **2h** with SdS and **2c** with EpiCS and DCS-W279Y, this is due to the presence of fluorine adjacent to the putative forming carbocations and/or due to decrease in the electron density of the C6-C7 double bound. There appears to not be any major disruption to the enzyme: substrate ternary complex geometry that would lead to a drastic change in product outcome.

Controls were performed by incubation of **2** without enzyme under otherwise identical conditions to check that the formed products are of enzymatic origin. No products were detected in these control incubations. Also, (2*E*,6*Z*)-6F-farnesol (**177**) was analysed by GC-MS under similar conditions as the incubation extracts, but it does not correspond to the products obtained from incubations 10-18.

4.4. Substrate behaviour of (2*E*,6*E*)-6F-FDP (3)

The universality of (2*E*,6*E*)-FDP (**1**) as a substrate for sesquiterpene synthase is under debate (Section 3.1) and the acceptance of isomeric counterparts might be more prevalent than previously assumed.[92] A decade ago, santalene and sergamotene synthase (SBS) from the wild tomato plant *Solanum habrochaites* was characterised,[283] being the first sesquiterpene synthase discovered which uses (2*Z*,6*Z*)-FDP (**252**) exclusively as the substrate for its conversion into 5 different sesquiterpene products. *Solanum habrochiates* also produces 7-epizingiberene synthase (EZS), which converts (2*Z*,6*Z*)-FDP (**252**) into 7-epizingiberene.[284,285] Of special interest is the ability of santalene synthase from *Santalum album*[286] (SaSSy) to use accept both (2*E*,6*E*)-FDP (**1**) and (2*Z*,6*Z*)-FDP (**252**) isomers as substrates. Here, the incubation and compatibility of (2*E*,6*E*)-6F-FDP (**3**) with the enzymes under analysis was tested (Figure 81).

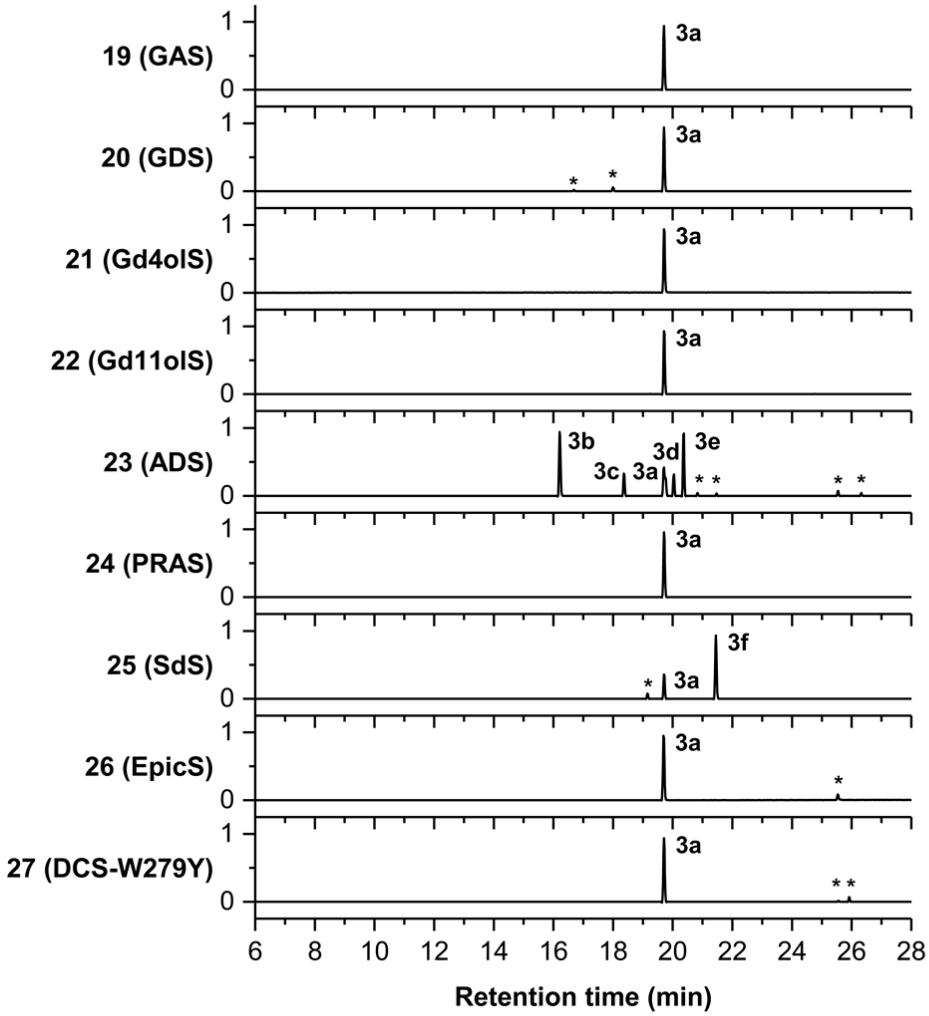


Figure 81. Total ion chromatograms of the pentane extractable products arising after incubations of (2*E*,6*E*)-6F-FDP (**3**) with nine sesquiterpene synthases. Trace compounds (< 5 %) are labelled with asterisk.

1,10- Sesquiterpene synthases

Incubation 19. The incubation of **3** with GAS gave exclusively one product in the pentane extracts, eluting at 19.71 min (labelled as **3a**, Figure 82a).The mass spectrum of this compound shows the molecular ion peak with m/z = 222 ($[M]^+$) and fragment peaks with m/z = 207 and 202, consistent with the loss of CH_3 or HF respectively from the parental compound, confirming it as a sesquiterpene fluorinated product (Figure 82b). The base peak appears with m/z = 68. **3a** has the same retention time as (2*E*,6*Z*)-6F-germacrene A (**2a**), which was obtained after incubation of (2*E*,6*Z*)-6F-FDP (**2**) with GAS (incubation 10). Their mass spectra are very similar. This is indicative that **3a** might be (*E*,*E*)-6F-germacrene A, an isomeric counterpart of **2a**, and that these are not separated in the column. Substrate **3** is a stereoisomer of **2** because the C6-C7 double bond of the FDP structure contains the same substituents but with different orientation. In **3**, the initial position of the C1 from C10 or C6 might be unfavourable for a ring closure to happen. Beside this, we propose that GAS is able to overcome this conformational barrier and formation of **3a** is allowed *via* a single-bond rotation of the substrate prior to ionisation and 1,10-cyclisation (Scheme 61).

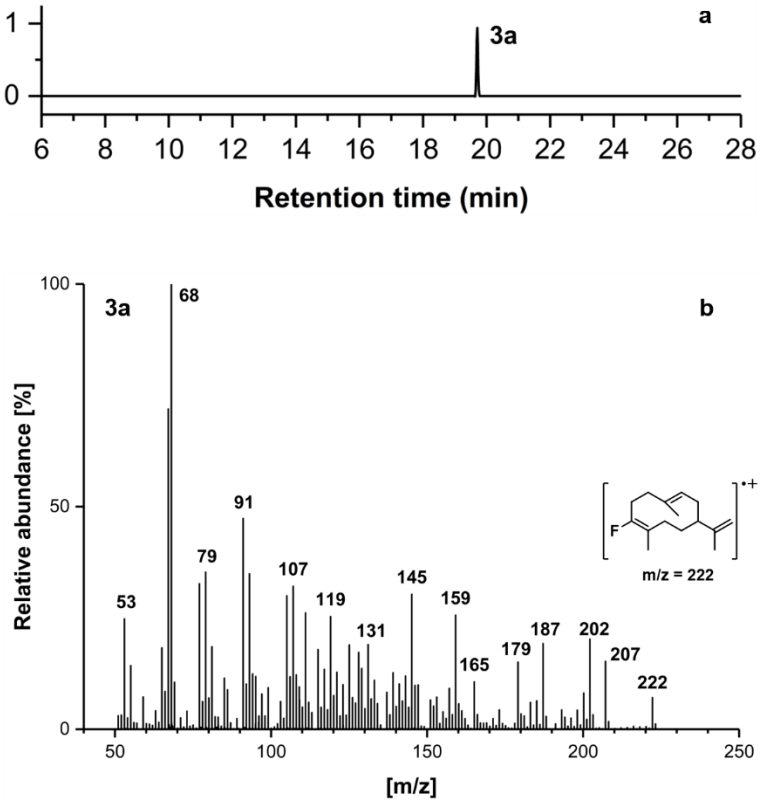
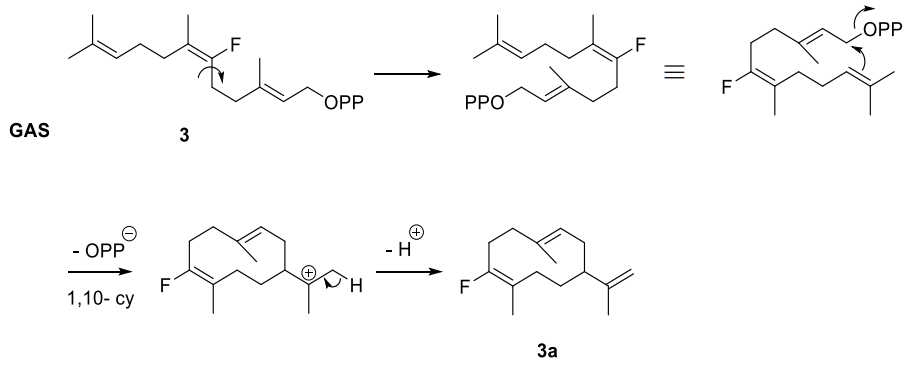


Figure 82. a) Total ion chromatogram of the pentane extractable products arising from incubation 19. b) Mass spectrum of the extracted product eluting at 19.71 min (**3a**) obtained from incubation 19, showing the molecular ion of **3a**.



Scheme 61. Proposed reaction mechanism catalysed by GAS for the enzymatic formation of **3a** from **3**.

Incubation 20. The incubation of GDS with **3** yielded also **3a** as the major extractable product, eluting at 19.71 min. In addition, traces of fluorinated sesquiterpene products were observed eluting at 17.99 min and 16.68 min, as judged by their mass spectra fragmentation pattern. The low yielding character of these products discarded further investigation of their identity (Figure 83).

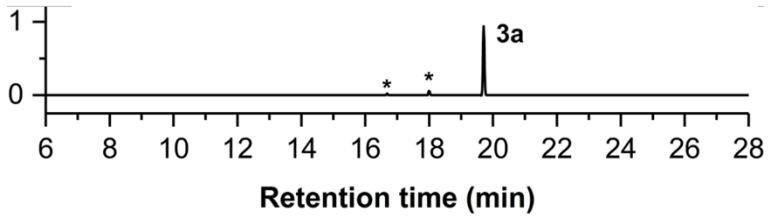


Figure 83. a) Total ion chromatogram of the pentane extractable products arising from incubation 20.

Incubation 21. The incubation of **3** with GdolS produced **3a** as the sole product in the organic extracts, eluting at 19.71 min (Figure 84).

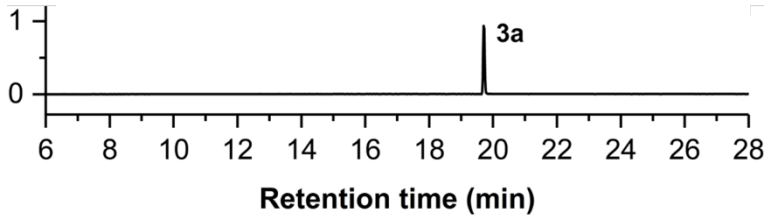
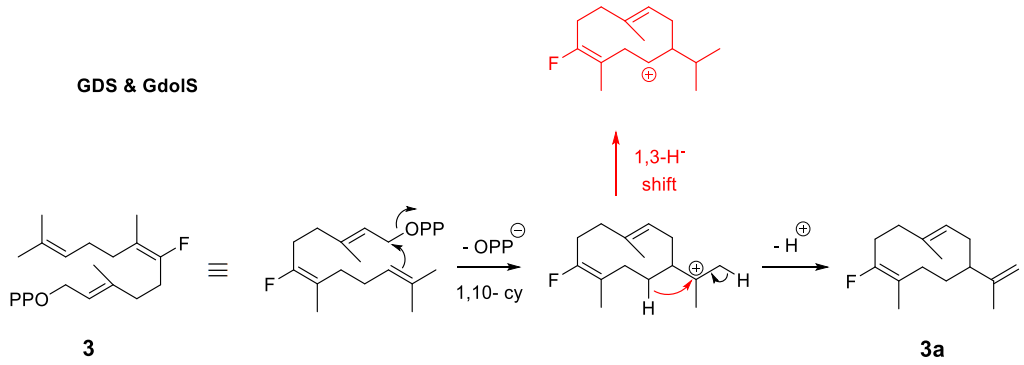


Figure 84. a) Total ion chromatogram of the pentane extractable products arising from incubation 21.

The natural reaction mechanisms exerted by GDS and GdolS is abolished upon incubation with **3**. This is because the formation of a new secondary carbocation after the 1,3-hydride shift is energetically restricted (Scheme 62).



Scheme 62. Proposed interception of the reaction mechanism catalysed by GDS and Gd4olS upon incubation with 3 (red), and catalysed conversion of **3** to **3a** by GDS and GdolS.

Incubation 22. The incubation of **3** with Gd11olS yielded exclusively a pentane extractable product eluting at 19.71, which is **3a** (Figure 85).

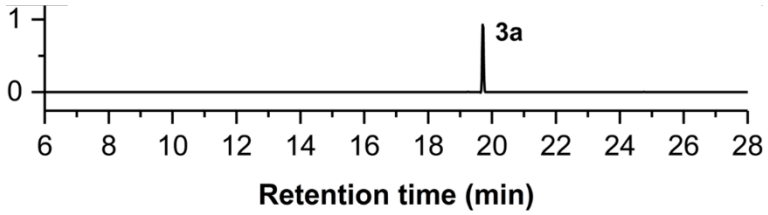


Figure 85. a) Total ion chromatogram of the pentane extractable products arising from incubation 22.

Overall, the four 1,10-sesquiterpene synthases used generate **3a** with high fidelity. These enzymes are able to accommodate **3** into the active site and use it as substrate for the generation of (*E*,*E*)-6F-germacrene A, **3a**. The presumed unproductive positioning of C1 with respect to C10 in the backbone of substrate **3** might force the enzyme to assist a single bond rotation that facilitates the spatial proximity between C1 and C10 for a 1,10-cyclisation to occurs. However, the conformation of the forming germacrenyl carbocationic intermediate is now unproductive to continue the naturally catalysed reaction mechanisms of these enzymes, thus leading to aberrant formation of **3a**.

1,6- Sesquiterpene synthase

Incubation 23. The incubation of **3** with ADS yielded a mixture of products in the organic extracts (Figure 86). The observed diversity on product formation, which is in contrast to the 1,10-sesquiterpene synthases, demonstrate that ADS produced an alternated intermediate which was not able to follow the natural cascade, hence aberrant products are formed (Scheme 63). Starting from **3**, ADS might catalyse the isomerisation to the NDP-like fluorinated analogue intermediate, but then, the presence of fluorine at C6 might reduce the nucleophilicity of the C6-C7 double bond needed to attack C1 and form a 6-membered ring intermediate. Also, this fluorine might destabilise the putative forming bisabolene cation, which would contain the fluorine adjacent to the positive charged carbon. With this scenario, the 1,6-cyclisation is sequestered, and alternate pathways are taken place (Scheme 63). The only generated compound that could be identified was **3a,** a presumed isomer of (6*E*,6*Z*)-6F-germacrene A (**2a**), eluting at 19.71 min. The other products resulted difficult to identify due to their novel character. This suggests that the incubation of **3** with ADS give rise to alternate pathways as compared with the use of FDP (**1**), and that ADS can be used for the preparation of novel fluorinated compounds, which are presumably not amorphadienes.

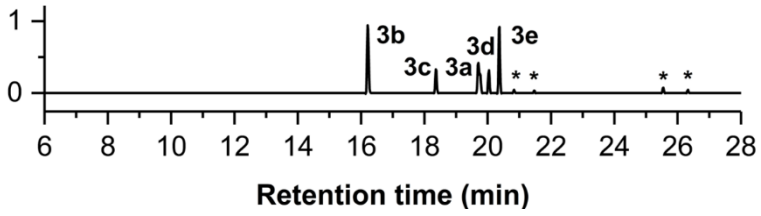
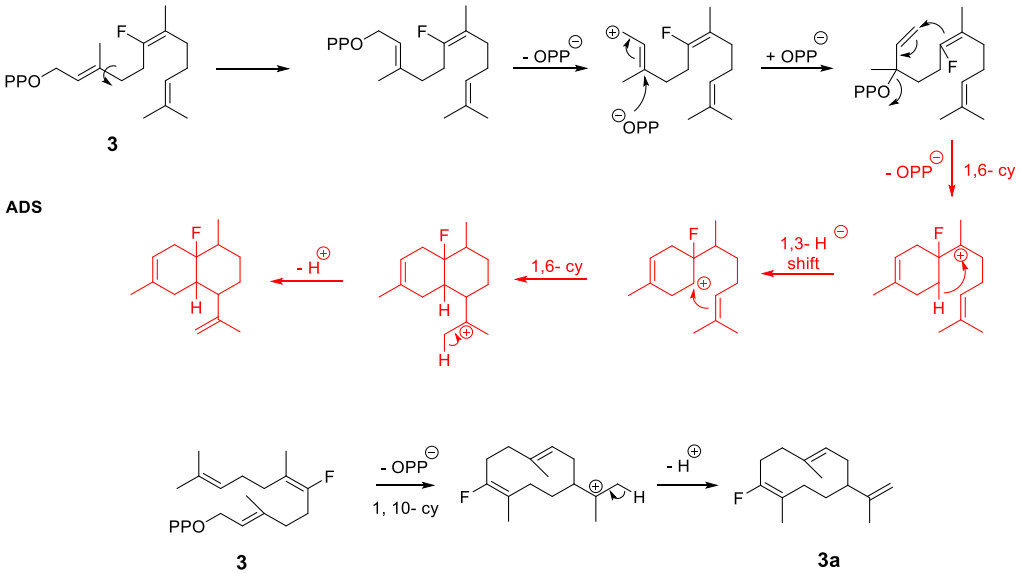


Figure 86. a) Total ion chromatogram of the pentane extractable products arising from incubation 23.



Scheme 63. In red, proposed interception of the mechanisms catalysed by ADS upon incubation with **3**. In black, proposed reaction mechanism for the conversion of **3** to **3a** by ADS.

1,10- + 1,6- Sesquiterpene synthases

Incubation 24. The incubation of **3** with PRAS gave one product in the organic extracts, **3a** (eluting at 19.71 min, Figure 87). This results can be rationalised as follows: PRAS directs the substrate **3** to (2*E*,6*E*)-6F-germacrene A (**3a**), but then is unable to catalyse the proton capture mediated 1,6-cyclisation, which seems conformationally unfeasible with the two double bounds implicated far away from each other (Scheme 64).

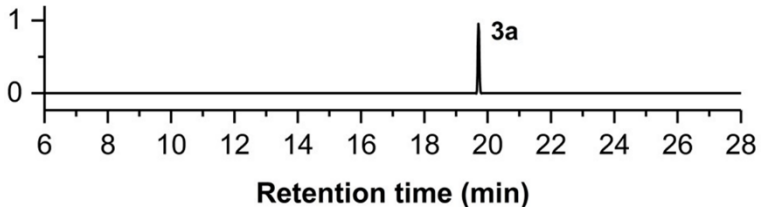
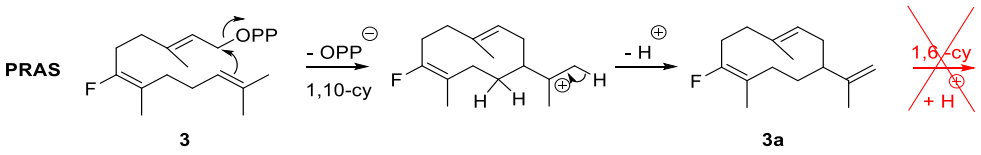


Figure 87. a) Total ion chromatogram of the pentane extractable products arising from incubation 24.



Scheme 64. In red, proposed interception of the mechanisms catalysed by PRAS upon incubation with **3**. In black, proposed reaction mechanism for the conversion of **3** to **3a** by PRAS.

Incubation 25. The organic extracts obtained from the incubation of **3** with SdS gave two major products, as judged by GC-MS (Figure 88a). Following the results obtained in the previous incubation, where **3a** is accumulated as an aberrant product because the 1,6-cyclisation is not performed in PRAS upon incubation with **3**, we speculated that the major product obtained in this incubation is (2*E*,6*E*)-6F-germacrene B (eluting at 21.45 min, labelled as **3f**, Figure 87a). Thus, SdS catalyses the formation of **3f**, but the conformation of this compound is not productive to afford a 1,6-cyclisation and it is accumulated as an aberrant product (Scheme 65). Its mass spectrum displays a molecular ion with m/z = 222 ($[M]^+$), and fragmentation peaks with m/z = 207 ($[M - CH_3]^+$), and 202 ($[M - HF]^+$), which are characteristic for 6F- sesquiterpene fluorinated compounds, Figure 88b. In addition to this, it contains a major fragmentation peak with m/z = 159 ($[M - HF - C_3H_7]^+$), which might suggest the presence of a germacrene analogue. Also, the base peak is with m/z = 121, which is a common feature compared with **1h** (germacrene B) and **2h** (postulated to be (*E*,*Z*)-6F-germacrene B). The overall mass spectrum assignment could validate the proposal that (*E*,*E*)-6F-germacrene B (**3f**) is the major product upon incubation of **3** with SdS. The other product observed was **3a** ((*E*,*E*)-6F- germacrene A), which eluted at 19.71 min and it is a common product for all the synthases used for this study. However, SdS is the first enzyme found to generate an alternative germacrene analogue to **3a** from its natural

1,10-cyclisation mechanism, in contrast with incubations 20-22, where GDS, GdolS and Gd11olS all generate exclusively **3a** upon incubation with **3**.

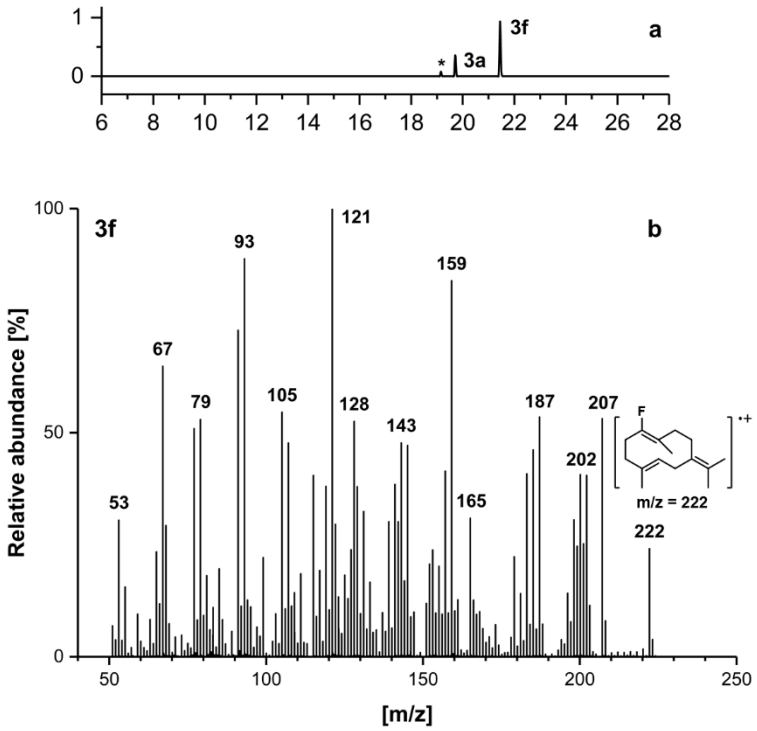
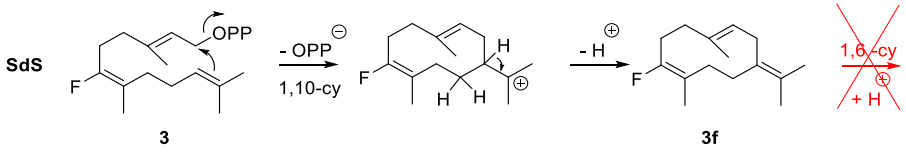


Figure 88. a) Total ion chromatogram of the pentane extractable products arising from incubation 25. b) Mass spectrum of the product eluting at 21.45 min (**3f**) obtained from incubation 25, showing the molecular ion of (2*E*,6*E*)-6F-germacrene B (**3f**).



Scheme 65. Reaction mechanism for the proposed conversion of **3** to **3f** by SdS. In red, proposed interception of the reaction mechanism catalysed by SdS upon incubation with **3**.

Incubation 26. The incubation of EpicS with **3** generated almost exclusively **3a**. (Figure 89)**.** EpicS produces (+)-epicubenol (**1j**, incubation 8) from FDP (**1**), *via* formation of NDP and a subsequent 1,10-cyclisation (Scheme 52). In this case, the isomerisation prerequisite is unlikely due to the conformation of the fluorinated NDP-analogue would be unproductive to proceed with the 1,10- cyclisation (Scheme 66). Then, in analogy with GdolS and others, EpicS leads to the production of the germacrene A analogue **3a**. The formation of 6F-epicubenol and 6F-germacradien-4-ol is prevented due to the putative secondary cation after the 1,3-hydride shift is no longer part of the allylic system observed for the FDP conversion (Scheme 66).

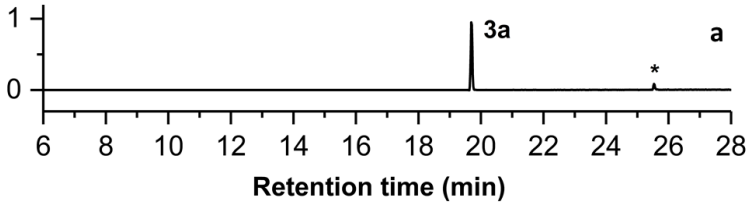
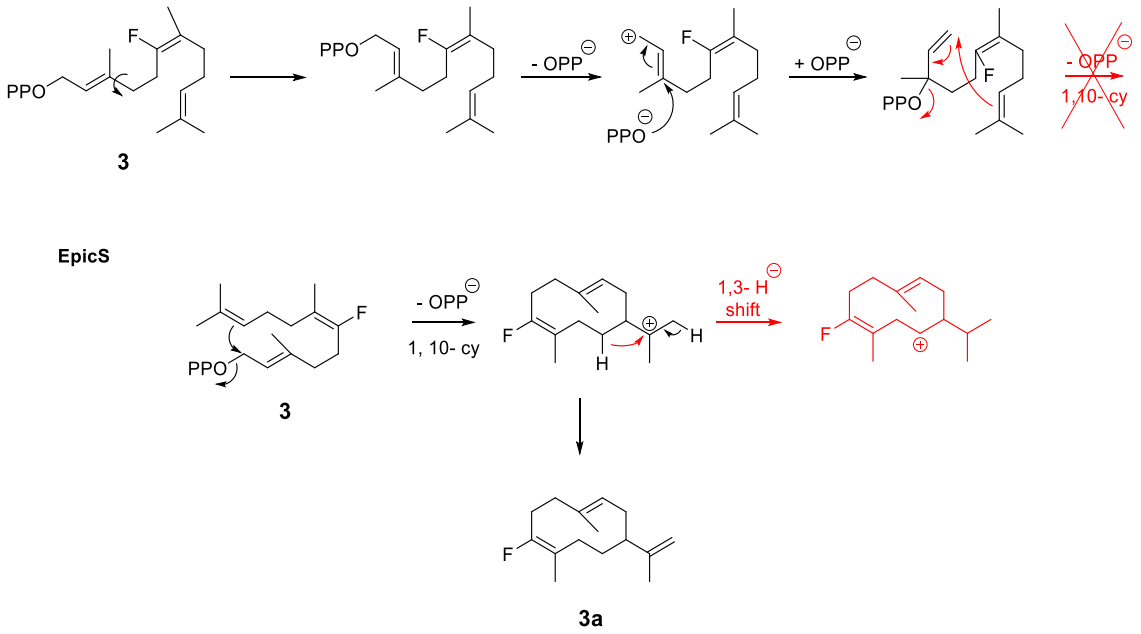


Figure 89. a) Total ion chromatogram of the pentane extractable products arising from incubation 26.



Scheme 66. In red, proposed interceptions of the reaction mechanism catalysed by SdS upon incubation with 3. In black, proposed reaction mechanism for the conversion of **3** to **3a** by EpicS.

Incubation 27. The incubation of **3** with DCS-W279Y yielded **3a** as the main product in the organic extracts, eluting at 19.71 min (Figure 90). This result can be easily rationalised if compared with the incubations of **3** with GdolS and EpicS (incubations 21 and 26).

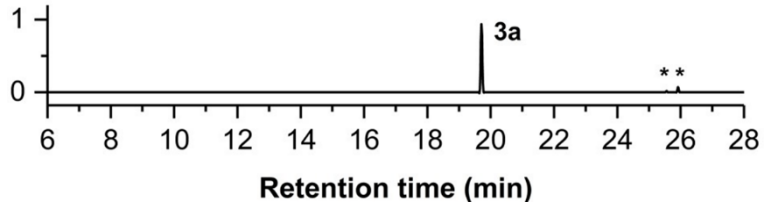


Figure 90. a) Total ion chromatogram of the pentane extractable products arising from incubation 27.

In summary, the 1,10- + 1,6-sesquiterpene synthases used generate **3a** as the main product, as observed with exclusively 1,10-sesquiterpene synthases. In addition, SdS generated a new fluorinated sesquiterpene in major quantities, which is proposed to be a (2*E*,6*E*)-6F-germacrene B analogue, **3f**.

Controls were performed by incubation of **3** without enzyme under otherwise identical conditions to check that the formed products are of enzymatic origin. No products were detected in these control incubations. Also, 2*E*,6*E*-6F-farnesol (**185**) was analysed by GC-MS under similar conditions as the incubation extracts and it does not correspond to the products obtained from incubations 19-27.

4.5. Substrate behaviour of (2*E*,6*E*)-6Me-FDP (4)

As stated in Sections 1.6.1 and 3.1, in our group, germacrene A (GAS) and germacrene D (GDS) synthases have been probed on their acceptance of different methylated FDP analogues as substrates to generate methylated germacrene analogues, namely 12Me-, 14Me-, 15Me- and 14,15-diMe-FDP.[123,130] Even though the substrates used in these studies (12Me, 14Me. 15Me and 14,15-diMe-FDPs) might be sterically challenging the enzyme, they contain the extra methyl group in positions not directly implicated in the catalytic mechanism (cyclisation, hydride shift, methyl shift, etc), thus only sterics decide if the mechanism works and the naturally proceeding pathways are not altered by means of stabilisation/ destabilisation of putative carbocations.

Here, the analytical incubations of the novel (2*E*,6*E*)-6Me-FDP (**4**) with the enzymes under analysis (Figure 91) are described. In 1,6-sesquiterpene synthases, the added methyl group at C6 might interfere with the natural mechanism and generate alternative products due to alteration of the reaction pathway. In germacrene-like synthases (1,10-sesquiterpene synthases), germacrene analogues of the natural counterparts are expected to be formed if the substrate binding conformation is adequate.

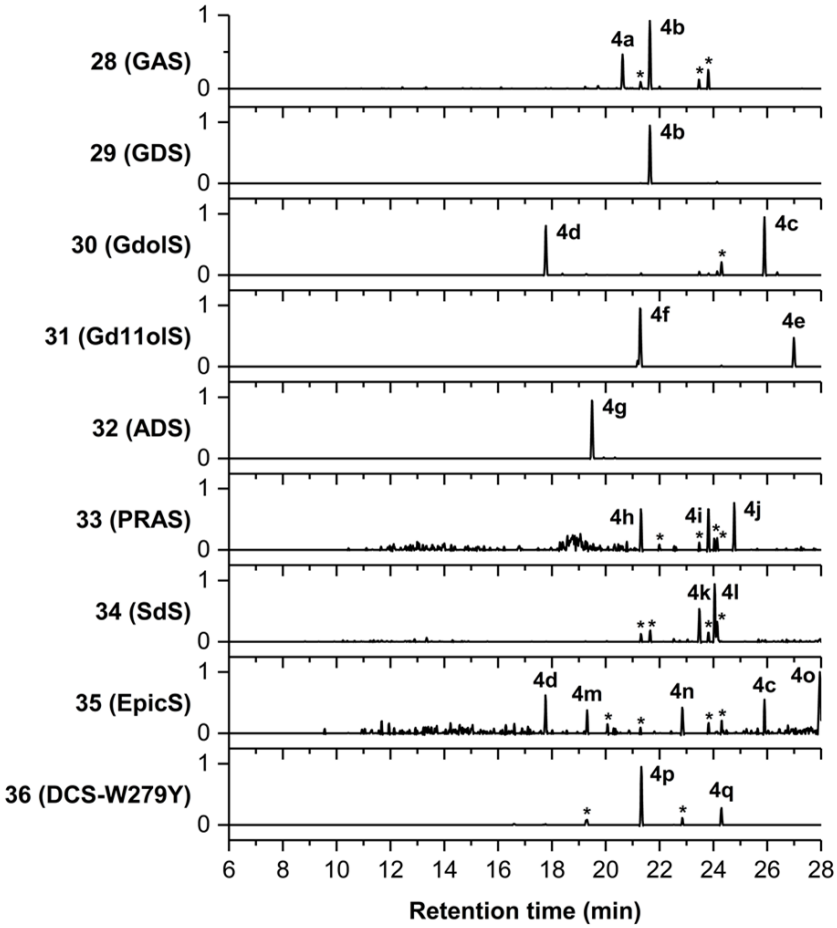


Figure 91. Total ion chromatograms of the pentane extractable products arising after incubations of (2*E*,6*E*)-6Me-FDP (**4**) with nine sesquiterpene synthases. Selected minor compounds are labelled with an asterisk to simplify the lecture.

1,10- Sesquiterpene synthases

Incubation 28. The organic extracts obtained from the incubation of **4** with GAS gave two major products eluting at 20.62 min (labelled as **4a**) and at 21.64 min (labelled as **4b**) in a mixture of 5 enzymatic products, as judged by GC-MS (Figure 92a). The mass spectra of all these compounds suggest they might be methylated sesquiterpene products because they show their molecular ions with m/z = 218 ($[M]^+$) and a major fragment peak with m/z = 203, which correspond to the loss of a methyl group from the parent methylated sesquiterpene, $[M - CH_3]^+$. The mass spectrum of the predominant product **4b** displays the most stable fragment with m/z = 175, corresponding to the loss of an isopropyl unit ($[M - 43]^+$) from the parental compound, Figure 93. Its fragmentation pattern is identical with that published of 14Me-germacrene D.[123] Also, it showed a similar fragmentation pattern to that of (+)-germacrene D (**1b,** incubation 2), but with fragment peaks 14 units bigger, coinciding with an extra methyl group. Due to this, it was hypothesised that **4b** is (2*E*,6*E*)-6Me-germacrene D. To probe this, we incubated **4** with GDS (incubation 29), which accordingly gave exclusively **4b**. The second main product, **4a**, was speculated to be (2*E*,6*E*)-6Me-germacrene A accordingly with the natural reaction mechanism catalysed by GAS. In fact, the mass spectrum shows the most stable fragment with m/z = 107 (Figure 92b), which is 14 units bigger than the base peak for germacrene A (**1a,** incubation 1), and their fragmentation pattern compare well.

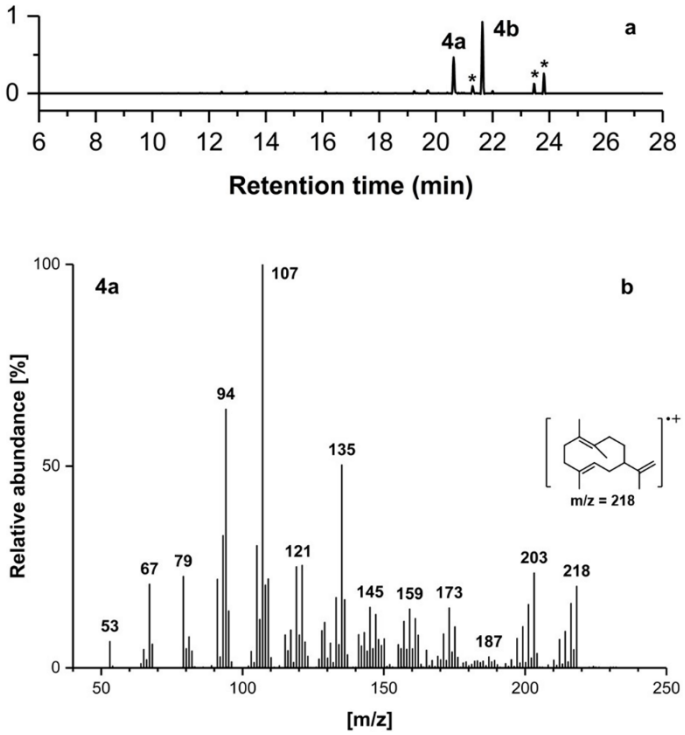


Figure 92.a) Total ion chromatogram of the pentane extractable products arising from incubation 28. b) Mass spectrum of the extracted product eluting at 20.62 min (**4a**) obtained from incubation 28.

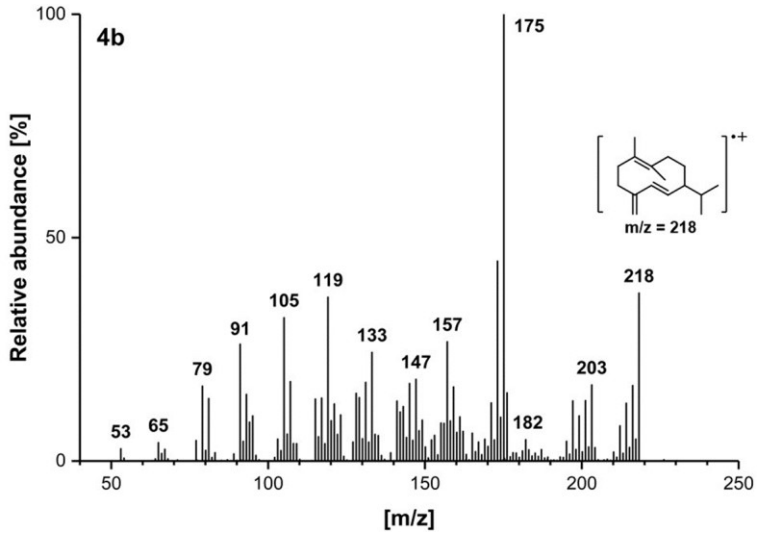
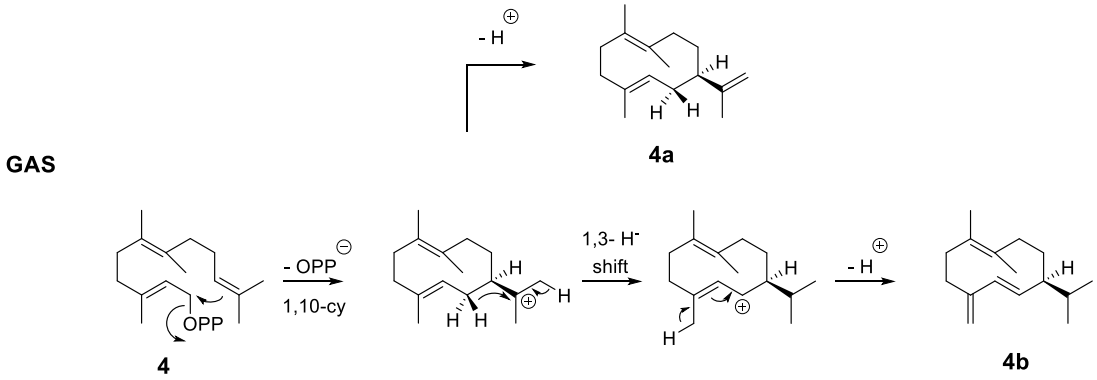


Figure 93. Mass spectrum of the pentane extracted product eluting at 21.64 min (labelled as **4b**) obtained from incubation 28.

That GAS generates 6Me-germacrene D (**4b**) over 6Me-germacrene A (**4a**) is not trivial. This might arise as a result of an altered fitting of **4** into the active site of GAS in comparison with the fitting of **1**, that leads to an alternative deprotonation, which is difficult to predict and/or analyse without using computational tools.



Scheme 67. Proposed reaction mechanisms for the conversions of **4** to **4a** and **4b** by GAS.

Incubation 29. The incubation of **4** with GDS gave a sole product in the pentane extracts (compound eluting at 21.64, **4b**, Figure 94). This compound was identified as 6Me-germacrene D, previously described in incubation 28.

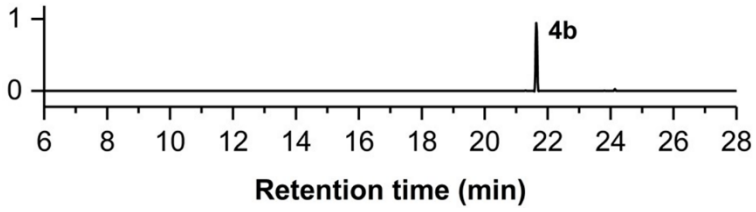
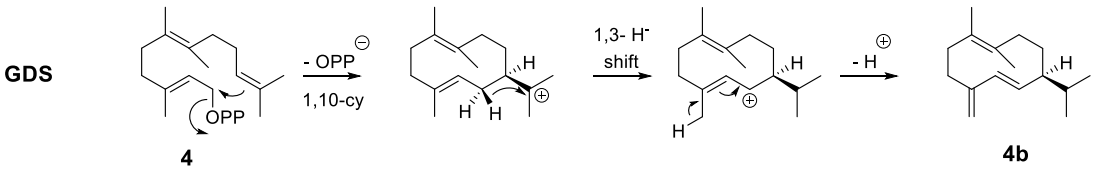


Figure 94. Total ion chromatogram of the pentane extractable products arising from incubation 29.



Scheme 68. Proposed reaction mechanism for the conversion of **4** to **4b** by GDS.

Incubation 30. The incubation of **4** with GdolS generated two major products, as judged by GC-MS analysis of the pentane extracts (Figure 95a). The mass spectrum of the compound eluting at 25.90 min (labelled as **4c**, Figure 93a) shows a molecular ion peak with m/z = 236 ($[M]^+$), suggesting the presence of a methylated sesquiterpene alcohol. Also, the presence of a fragment peak with m/z = 221 was detected, which indicates the loss of a methyl group from the parent alcohol, $[M - CH_3]^+$. The most stable fragment peak showed m/z = 175, corresponding to the loss of the isopropyl group from the dehydrated compound ($[M - H_2O - C_3H_7]^+$), Figure 95b. This product is proposed to be (2*E*, 6*E*)-6Me -germacradien-4-ol (**4c**), which is the product arising from the natural catalysed reactions by G4olS (Scheme 69). The second main product eluted at 17.77 min (labelled as **4d**, Figure 95a). The mass spectrum of **4d** shows similar fragmentation pattern to **4c** but presenting a molecular ion peak with m/z = 218 ($[M]^+$), which supports the presence of a pure hydrocarbon sesquiterpene with an extra methyl group. It also displays the most stable fragment with m/z = 175 ($[M - C_3H_7]^+$), which might reveals the presence of a germacrene-like structure, Figure 96. The identity of this compound is yet unknown.

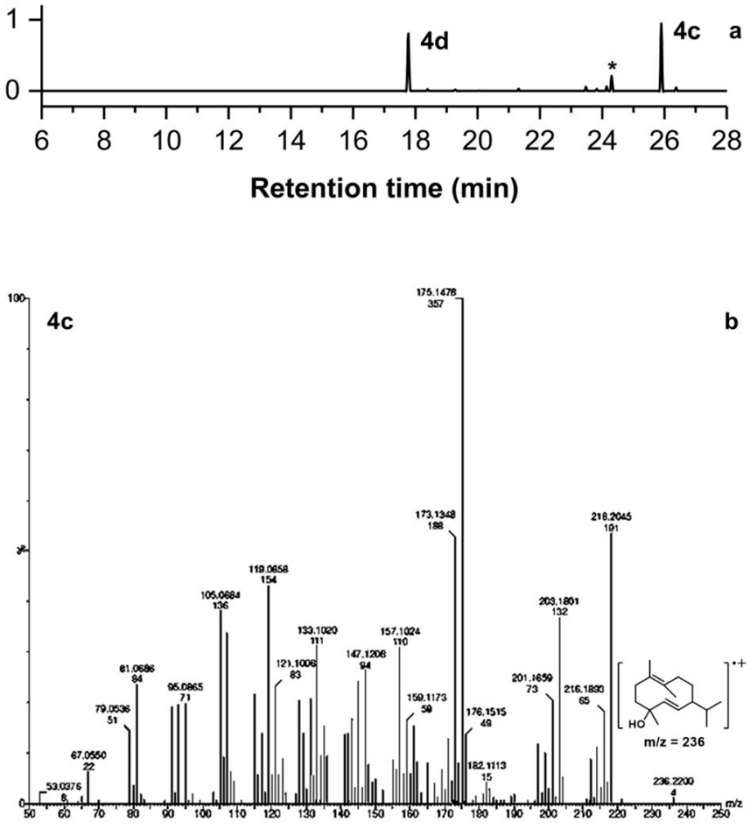


Figure 95. a) Total ion chromatogram of the pentane extractable products arising from incubation 30. b) Mass spectrum of the product eluting at 25.90 min (4c) obtained from incubation 30.

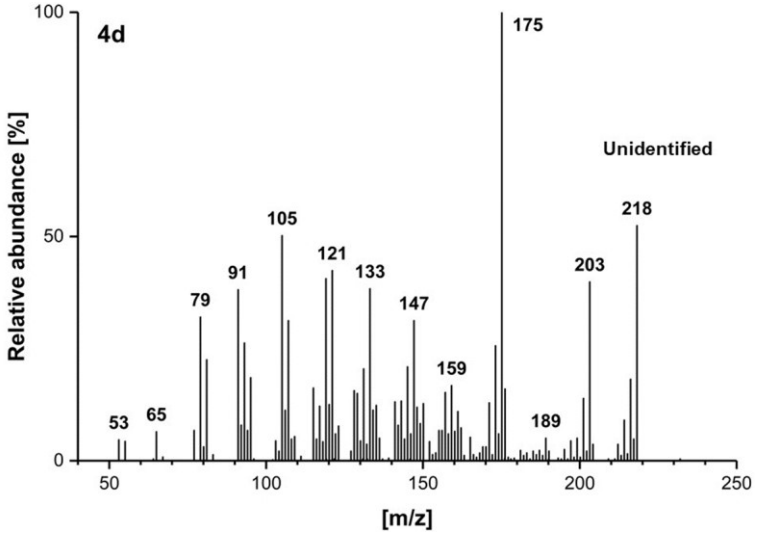
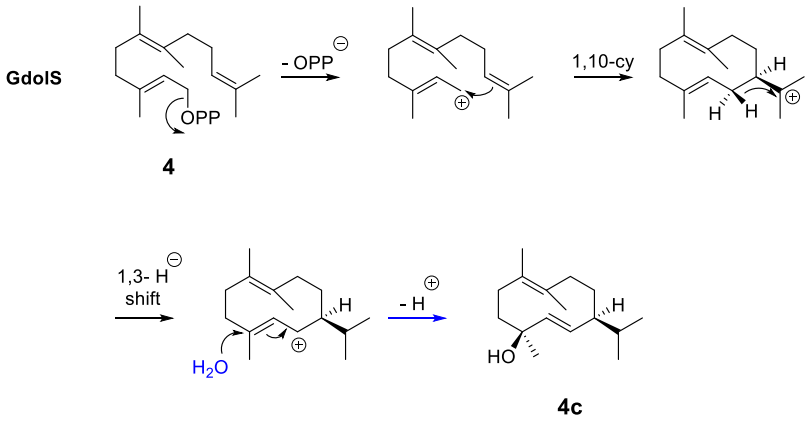


Figure 96. Mass spectrum of the product eluting at 17.77 min (**4d**) obtained from incubation 30, which remains unidentified.



Scheme 69. Reaction mechanism for the proposed conversion of **4** to **4c** by GdolS.

Incubation 31. The incubation of **4** with Gd11olS gave two major extractable products eluting at 21.27 min and 26.99 min (labelled as **4f** and **4e**, respectively, Figure 97a). The mass spectrum of **4e** shows a molecular ion with m/z = 236 ($[M]^+$), indicative of the presence of a methylated sesquiterpene alcohol (Figure 97b). This is proposed to be (2*E*,6*E*)-6Me-germacradien-11-ol as all other germacrene synthases produced their natural analogue upon incubation with **4** (incubations 28 to 31) and the mechanism underwent by Gd11olS upon incubation with **1** might be maintained with this methylated substrate, Scheme 70. **4f**, however, is a non-hydroxylated methylated sesquiterpene, showing a molecular ion with m/z = 218 ($[M]^+$). The most stable fragmentation peak, as for **4a**, arises with m/z = 107, Figure 97c. Due of this, the product may be a germacrene A derivative, as observed in other studies.[122]

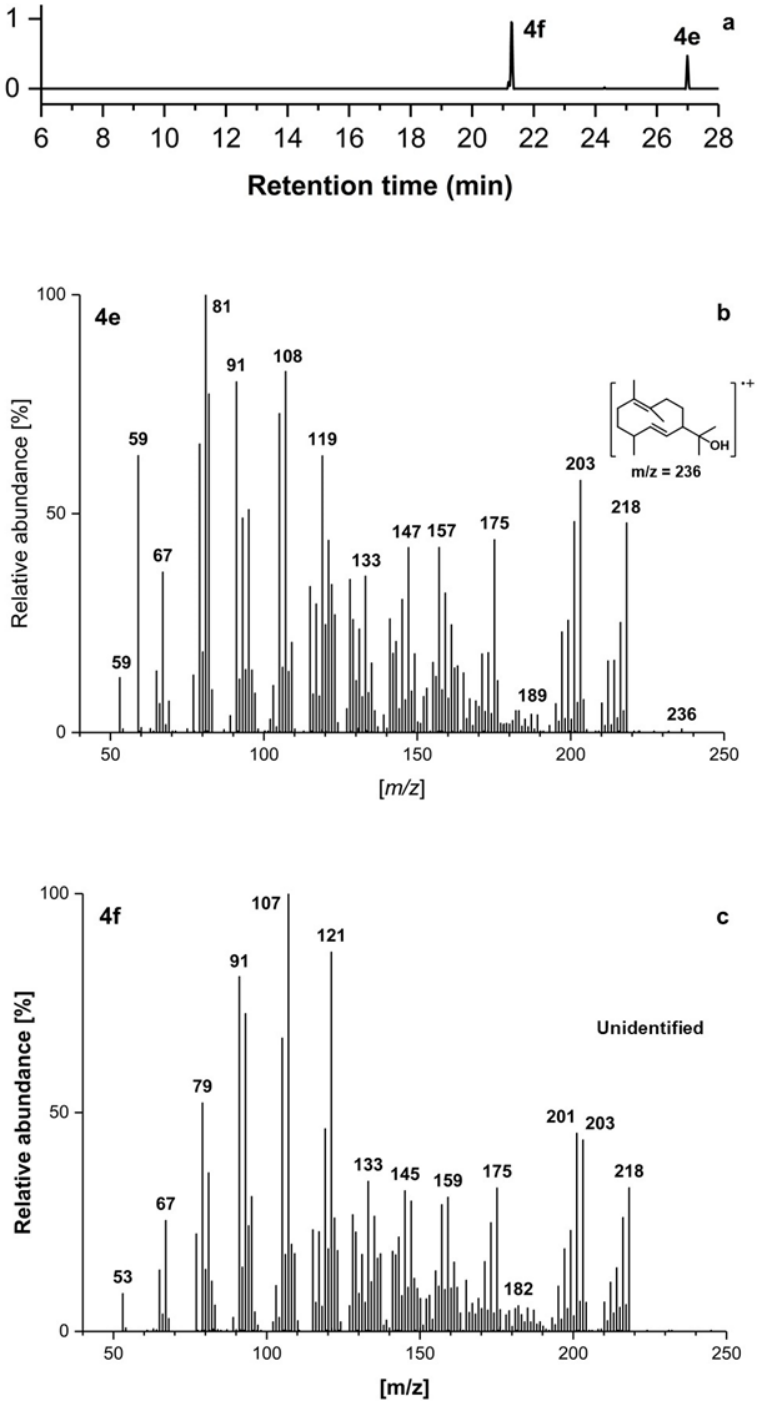
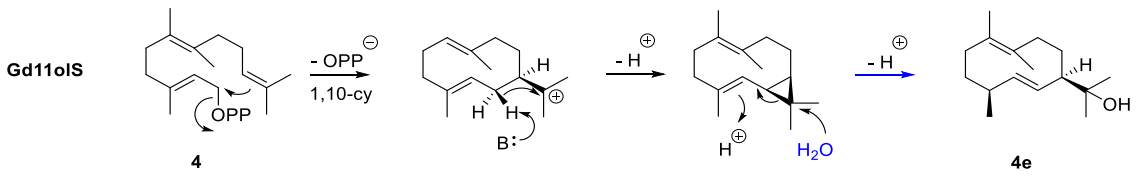


Figure 97. a) Total ion chromatogram of the pentane extractable products arising from incubation 31. b) Mass spectrum of the product eluting at 26.99 min (**4e**) obtained from incubation 31. c) Mass spectrum of the product eluting at 21.27 min (**4f**) obtained from incubation 31.



Scheme 70. Reaction mechanism for the proposed conversion of **4** to **4e** by Gd11olS.

Overall, all the 1,10-sesquiterpene synthases used here generate the methylated analogue of their natural products upon incubation with **4**, albeit with different fidelity. GDS is the only germacrene synthase able to exclusively generate (2*E*,6*E*)-6Me-germacrene D (**4b**). The others, GAS, GdolS and Gd11oS, produce varying quantities of an unidentified product. 1,10-sesquiterpene synthases are more promiscuous using **4** than when using **1** or **2**, and this is because C6-fluorine in **2** does not alter the natural reaction mechanisms neither by electronic or steric reasons in these enzymes, however, **4** contains an extra methyl substituent at C6 that, through its extra bulk, probably alters the overall geometry of the enzyme-substrate complex and changes the product distribution.

1,6- Sesquiterpene synthase

Incubation 32. The incubation of **4** with ADS yielded a single product eluting at 19.48 min, as judged by GC-MS analysis of the pentane extractable products (labelled as **4g**, Figure 98a). The mass spectrum of **4g** presents a molecular ion with m/z =218 ($[M]^+$), corresponding to the expected mass for a methylated sesquiterpene (Figure 98b). The most stable fragmentation peak is found with m/z=203, $[M - CH_3]^+$. This is comparable with amorphadiene (**1e**, Figure 65). The mass spectrum of **1e** displays the most stable fragment with m/z = 189 (14 units smaller in comparison with that of **4g**), and it is similar to that of **4g** overall. Because of this, we propose that **4g** is 6Me- amorphadiene, and therefore the first 1,6-cylcisation is not hindered in ADS upon incubation with **4** and the following steps are also proceeding as normal, including the second 1,6- cyclisation to produce a two membered ring structure (Scheme 71).

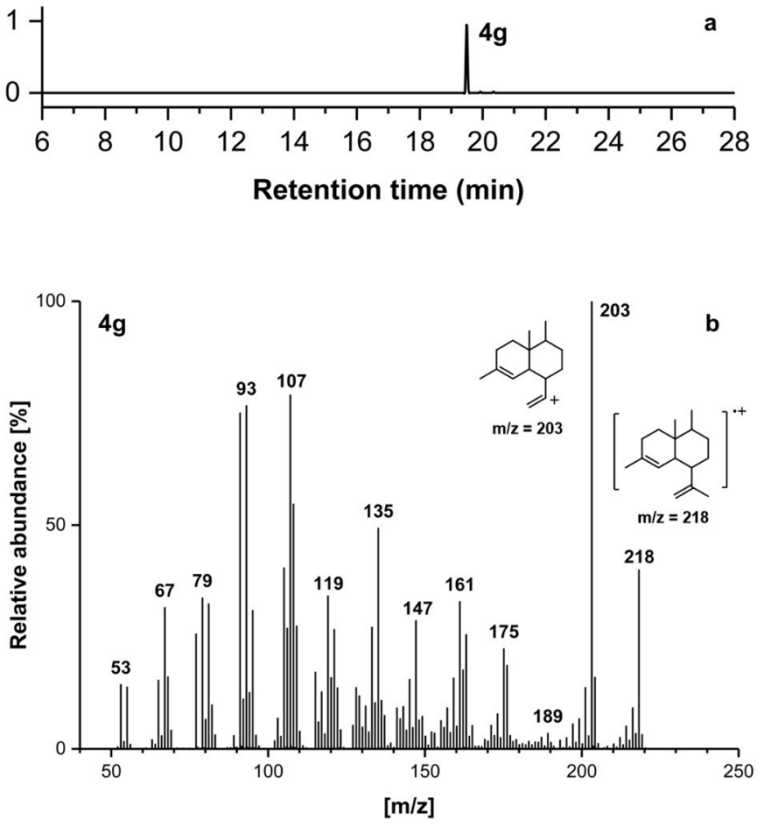
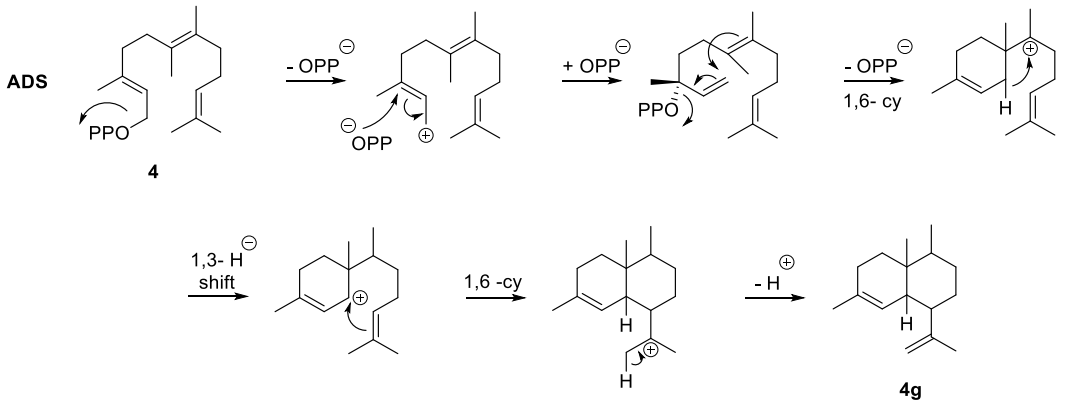


Figure 98. a) Total ion chromatogram of the pentane extractable products arising from incubation 32. b) Mass spectrum of the product eluting at 19.48 min (**4g**) obtained from incubation 32, showing the molecular ion of **4g** and its most stable fragment.



Scheme 71. Reaction mechanism for the proposed conversion of **4** to **4g** by ADS.

The high product fidelity observed upon incubation of **4** with ADS to produce the proposed 6Me-amorphadiene analogue **4f**, is indicative that **4** is well accommodated in the active site of ADS to proceed through a 1,6-cyclisation, and that the 6Me group does also not interfere in the mechanism. This is in contrast with the results after incubation of **2** with ADS (incubation 14), where the C6 fluorine destabilises the putative carbocation that forms after the 1,6-cyclisation. Due to the high impact that amorphadiene has in medicine and the need of alternative amorphadiene- related compounds, namely for the treatment of malaria,[133,287] this newly produced compound may be a good candidate for its use for large scale production and further conversion to methylated artemisinin analogue, as it is obtained without the presence of side products.

1,10- + 1,6- Sesquiterpene synthases

Incubation 33. The incubation of **4** with PRAS generated very low levels of multiple products (Figure 99a). The three major compounds eluted at 21.30 min (**4h**), 23.82 min (**4i**) and 24.77 min (**4j**). From these, the mass spectra from **4i** and **4j**, show a molecular ion with m/z = 236 ($[M]^+$), suggesting that these might be methylated sesquiterpene alcohols, however their retention time do not match with (2*E*,6*E*)-6Me-farnesol (**198**), which was injected as a control. Thus, these might be aberrant hydroxylated compounds formed by quench of the intermediates formed along the course of the reaction to 6Me- aristolochene (methylated aristolochene analogue, Scheme 72). The compound labelled as **4h** could be the methylated aristolochene shown in Scheme 72. There is no remarkable evidence for this, but the mechanism of aristolochene formation upon incubation of PRAS with **1** (Scheme 52) might accept this performance. Also, the mass spectrum of **4h** is similar to that described for (+)-aristolochene (**1f**, Figure 66), containing the most stable fragment peaks with m/z = 203 ([M - CH3]$^+$) and m/z = 119 (Figure 99b), which are 14 units bigger that the most stable fragment peaks found in the mass spectrum of **1f**. In addition to this, it is interesting that methylated germacrene A is not accumulated here, despairing with the results obtained upon incubations with fluorinated analogues. This might indicate that PRAS is able to perform the second 1,6- cyclisation upon incubation with **4**, as proposed for ADS (incubation 32).

Nevertheless, PRAS behaves promiscuously and low yielding upon incubation with **4**, thus the fitting of **4** in the active site of PRAS might be unproductive and this enzyme is not of interest for its use as platform for the generation of novel terpenoids, unless it is engineered by SDM studies.[130]

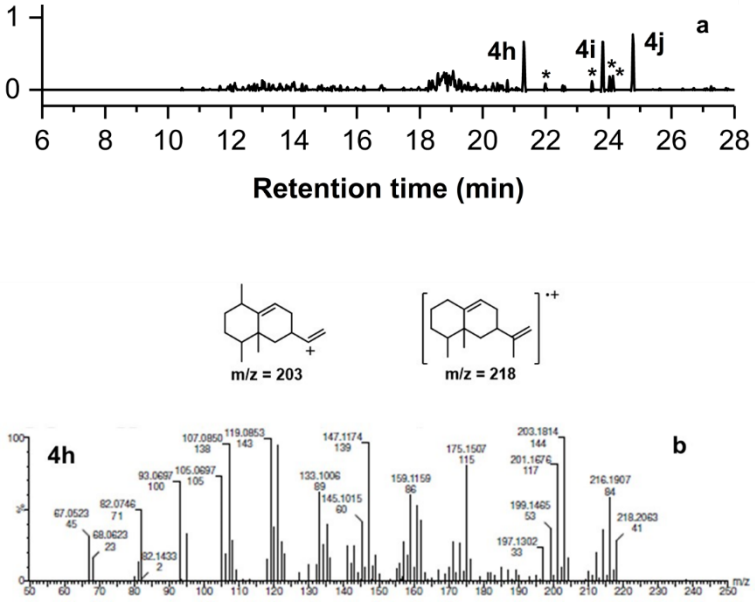
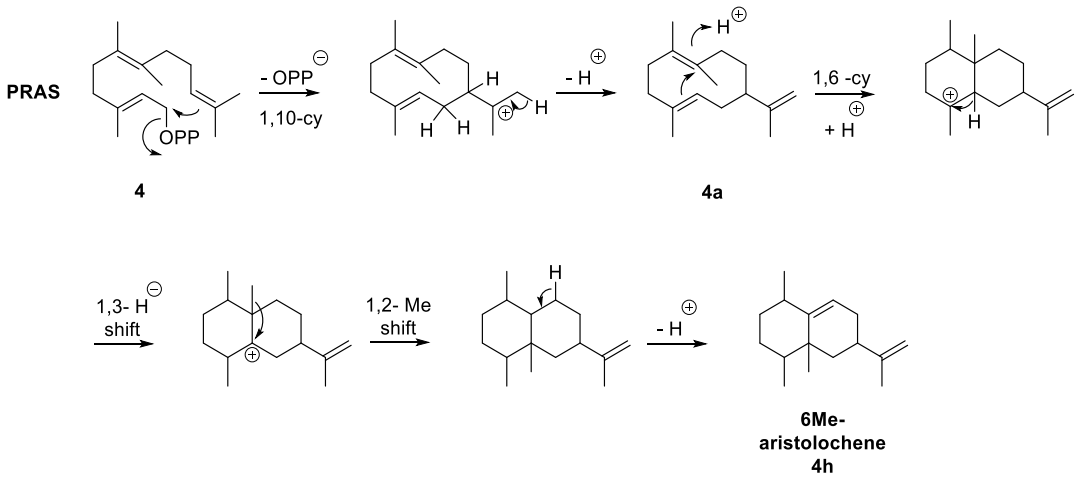


Figure 99. a) Total ion chromatogram of the pentane extractable products arising from incubation 33. b) Mass spectrum of the product eluting at 21.30 min (**4h**) obtained from incubation 33, showing the molecular ion of **4h** and its most stable fragment.



Scheme 72. Reaction mechanism for the proposed conversion of **4** to **4h** by PR-AS.

Incubation 34. The incubation of **4** with SdS resulted also in low product formation under the conditions used here. Two main compounds were detected in the pentane extracts, apart from other minor products. These major products eluted at 23.49 min and 24.04 min (labelled as **4k** and **4l** respectively, Figure 100a). These were proposed to be methylated sesquiterpenes because their mass spectra showed a molecular ion with m/z = 218 ([M]$^+$). Methylated germacrene B and selina-4(15),7(11)-diene could be formed by following the natural reaction cascades proceeding in SdS (Scheme 73). However, these were not further investigated due to the low-yielding character of product formation.

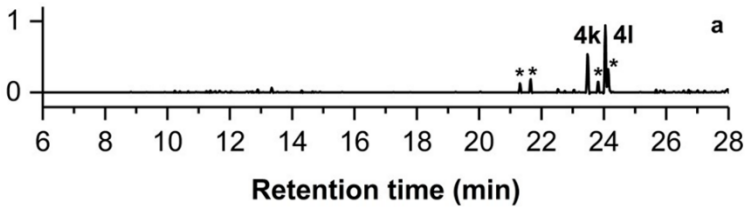
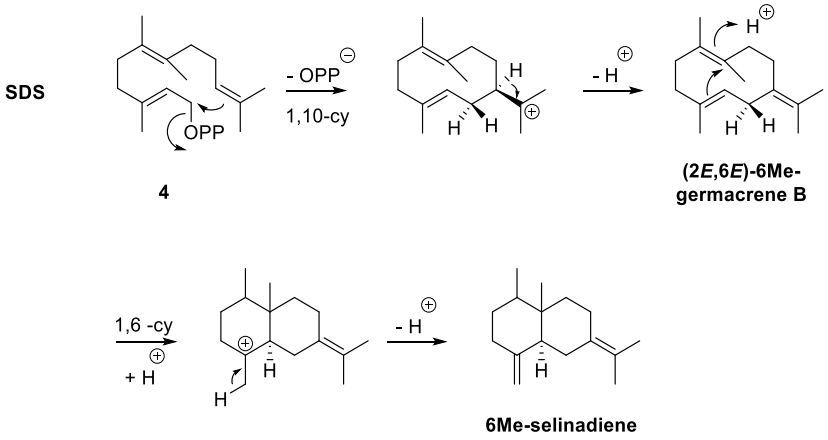


Figure 100. a) Total ion chromatogram of the pentane extractable products arising from incubation 34. b) Mass spectrum of the product eluting at 23.49 min (**4k**) obtained from incubation 34, showing the molecular ion of (2*E*,6*E*)-6Me-germacrene B (**4k**).



Scheme 73. Reaction mechanism for the hypothetical conversion of **4** to 6Me-germacrene B and 6Me-selina-4(15),7(11)-diene by SdS.

Incubation 35. The incubation of EpicS with **4** also gave a broad mixture of detectable methylated sesquiterpene in very low quantity (Figure 101). Among these products, we found (2*E*,6*E*)-6Me-germacradien-4-ol (**4c**) eluting at 25.90 min, which is a natural analogue product (Scheme 74), as EpicS produces germacradien-4-ol (**1c**) upon incubation with **1** (Scheme 50). Also, it was observed a common product with Gd4olS, which is labelled as **4d**. This compound remains unknown by the time of writing. The compound eluting at 27.95 min (labelled as **4o**) was identified as (2*E*,6*E*)-6Me-farnesol (**198**) upon comparison of the retention time and mass spectrum with that of authentic premade synthetic sample of **198**. The other compounds detected were not studied due to the diversity of product formation, but the presence of 7-diMe-epicubenol (Scheme 74) seems unfeasible because it requires a new 1,2- methyl shift.

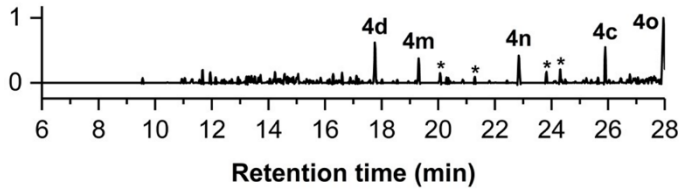
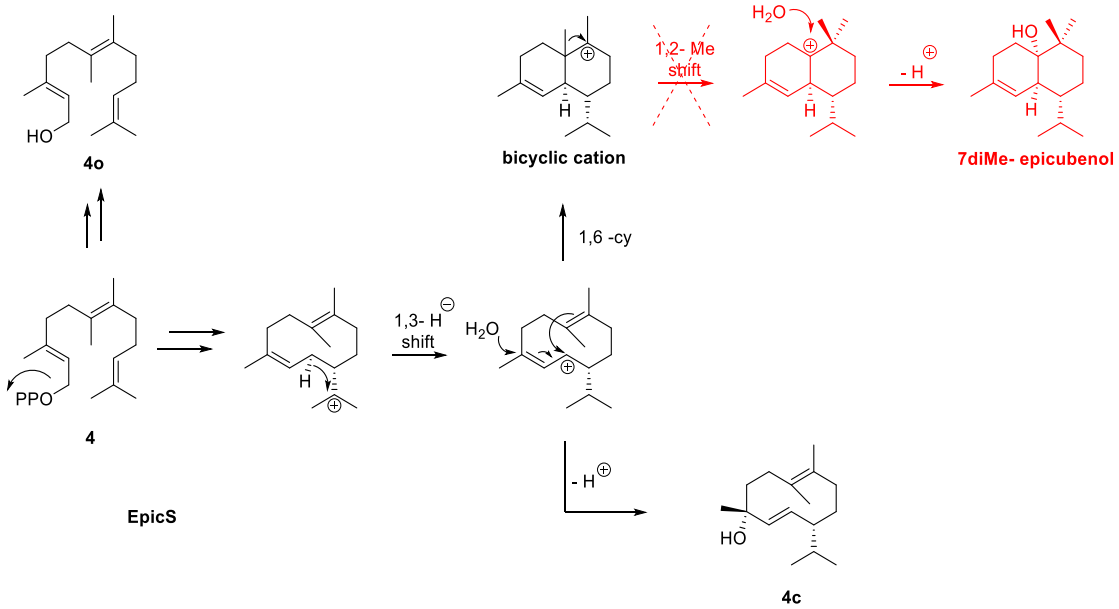


Figure 101. a) Total ion chromatogram of the pentane extractable products arising from incubation 35.



Scheme 74. In red, proposed interception of the reaction mechanism catalysed by EpicS upon incubation with **4**. In black, proposed reaction mechanism for the conversion of **4** to **4c** and **4o** by EpicS.

Incubation 36. The incubation of **4** with DCS-W279Y yielded two predominant products in the pentane extracts, eluting at 21.32 min (labelled as **4p**, Figure 102a) and 24.30 min (labelled as **4q**, Figure 102a). The mass spectra of these compounds are similar, however, the mass spectrum of **4q** contains a molecular ion peak with m/z = 236 ([M]$^+$), suggesting it is a methylated sesquiterpene alcohol (Figure 102b), while the mass spectrum of **4p** presents a molecular ion with m/z = 218 ([M]+), agreeing with the presence of a pure hydrocarbon methylated sesquiterpene (Figure 103). Remarkably, (2*E*,6*E*)-6Met-germacradien-4-ol is not formed, which would be a naturally expected product (Scheme 50). The path to methylated 6Me-δ-cadinene is hindered due to the presence of the C6-methyl instead of the proton required for natural deprotonation (Scheme 75), and thus alternative cadinenes/bicyclic products might be formed. The formation of germacrene-like products it is also feasible (Scheme 75), supported by stable fragmentation peaks with m/z = 175 in the mass spectra of these compounds.

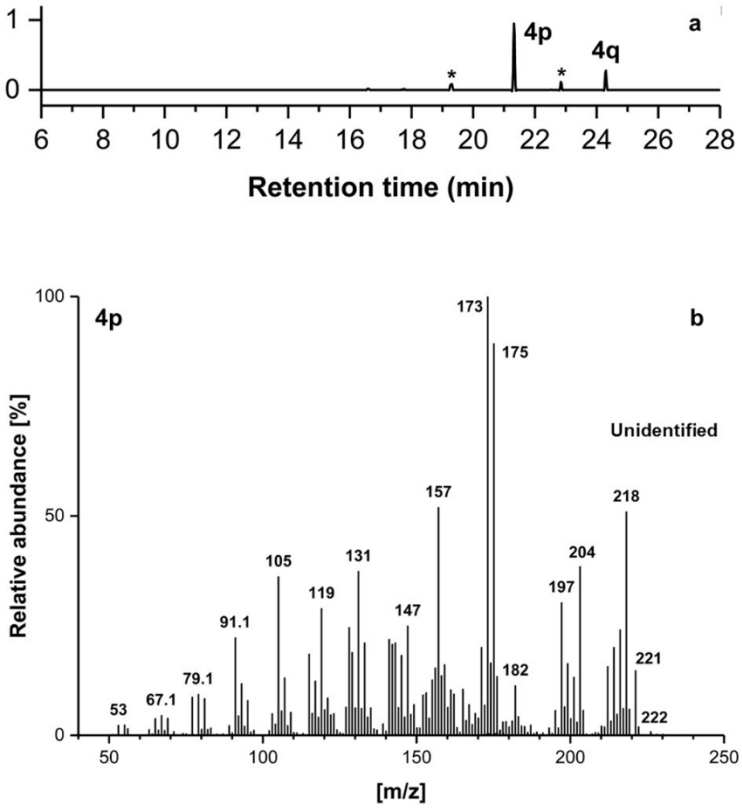


Figure 102. a) Gas chromatogram from incubation 36. b) Mass spectrum of the product eluting at 21.32 min (**4p**) obtained from incubation 36, which remains unidentified.

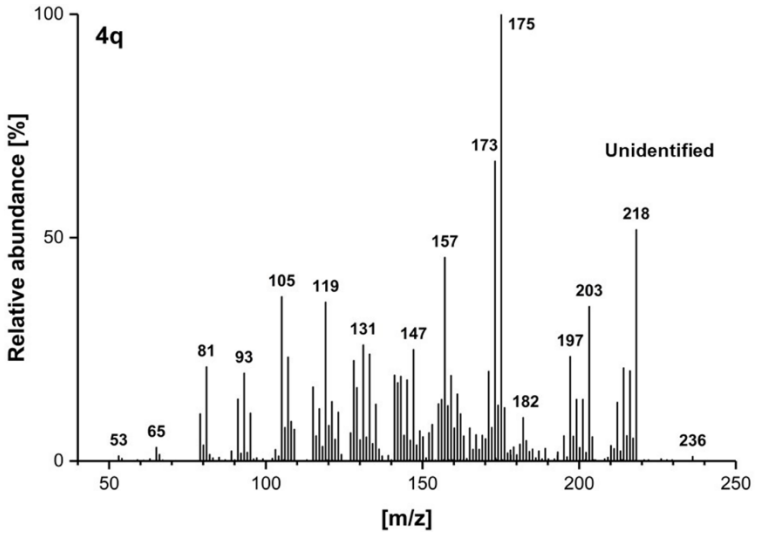
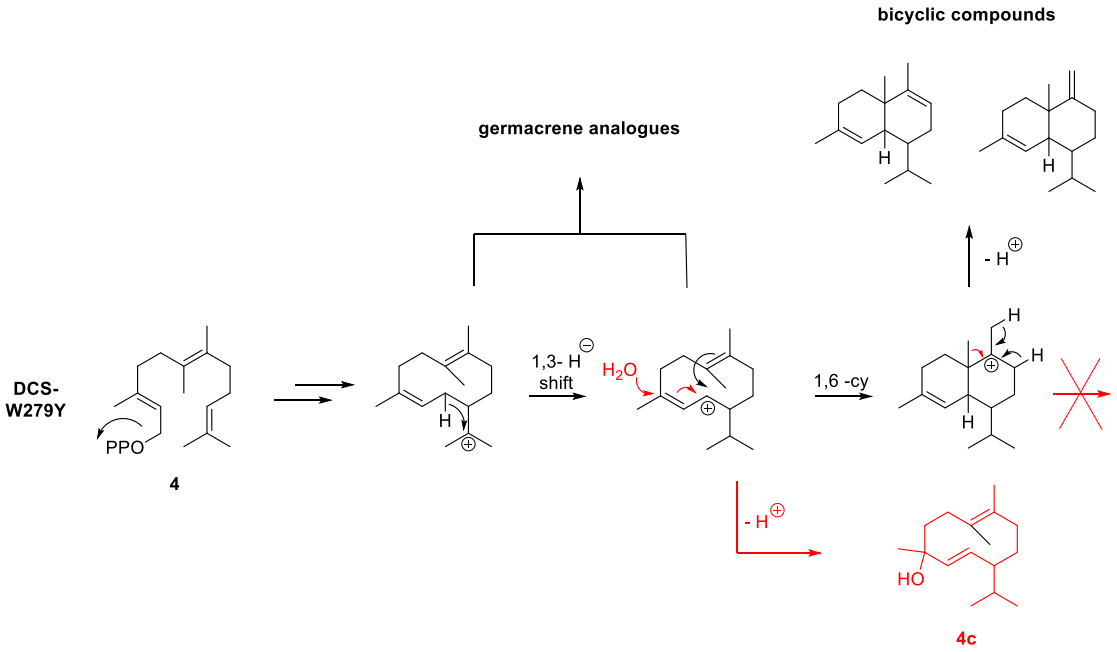


Figure 103. Mass spectrum of the product eluting at 24.30 min (**4q**) obtained from incubation 36, which remains identified.



Scheme 75. In red, proposed interception of the *reaction* mechanism*s* catalysed by DCS-W279Y upon incubation with **4**. In black, alternative products speculated to arise from aberrant deprotonations.

Overall, 1,10- + 1,6-sesquiterpene synthases behave promiscuously when catalysing the transformation of **4**. As observed the for conversion of this compound by ADS (See incubation 32). Presumably, he presence of the additional methyl group does not hinder the 1,6-cyclisation mechanism, however, it may disrupts the enzyme-substrate geometry in these enzymes (PRAS, SdS, EpicS and DCS-W279Y), that lead to a more complex product distribution. In addition, in EpicS the presence of the C6 methyl group intercepts the natural reaction mechanism as it is forbidden the 1,2-hydride shift in the process, and in DCS-W279Y the final deprotonation to form δ-cadinene product is also not possible. Most of the products obtained proved difficult to identify due to the low catalytic efficiency of their production and the low average yield obtained.

Controls were performed by incubation of **4** without enzyme under otherwise identical conditions to check that the formed products are of enzymatic origin. No products were detected in these control incubations. Also, (2*E*,6E*)*-6Me-farnesol (**198**) was analysed by GC-MS under similar conditions as the incubation extracts, which corresponded to **4o**, a product obtained after the incubation of **4** with EpicS.

4.6. Incubation of (2*E*,6*Z*)-6Me-FDP (5) with the enzymes under study

(2*E*,6*E*)-6F-FDP (**3**) resulted a substrate for all the sesquiterpene synthases used in this study (Section 4.5). In analogy, the use of (*E*,*Z*)-6Me-FDP (**5**) as a novel substrate for sesquiterpene synthases is covered here (Figure 104), and their product formation is analysed.

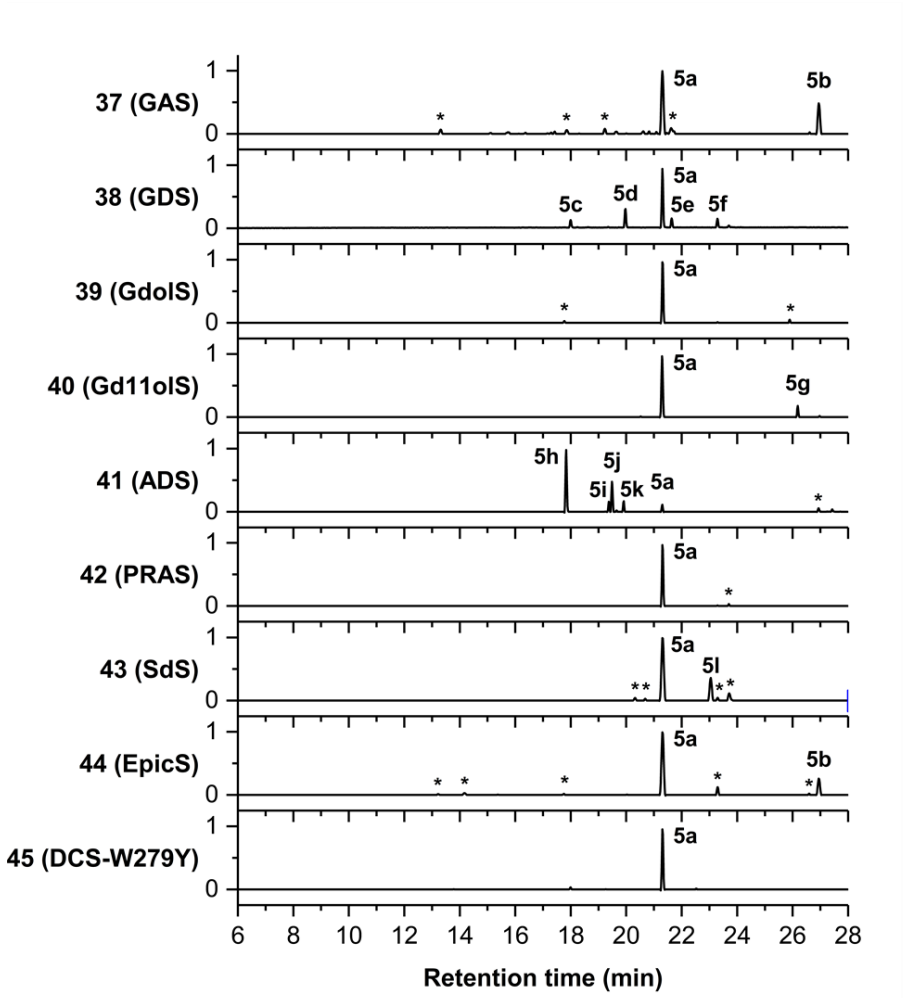


Figure 104. Total ion chromatograms of the pentane extractable products arising after incubations of (2*E*,6*Z*)-6Me-FDP (**5**) with nine sesquiterpene synthases. Selected minor compounds are labelled with an asterisk.

1,10- Sesquiterpene synthases

Incubation 37. The incubation of **5** with GAS gave two main products eluting at 21.30 min and 26.03 min (labelled as **5a** and **5b**, respectively, Figure 105). The mass spectra of **5a** displays a molecular ion peak with m/z = 218 $[M]^+$, and a major fragment with m/z = 203 $[M - CH_3]^+$, confirming the presence of a methylated sesquiterpene (Figure 106a). The overall fragmentation patterns observed at the mass spectra of (+)-germacrene A (**1a**) and **5a** are similar, but with 14 units bigger fragments for **5a**, which might indicate that **5a** is (*E*,*Z*)-6Me-germacrene A (**5a**). As observed for the conversion of **3** to **3a** (See section 4.4), GAS might assist to accommodate (*E*,*Z*)-6Me-FDP (**5**) in the active site with the correct conformation to perform an ionisation-mediated 1,10-cyclisation, and a deprotonation to yield (*E*,*Z*)-6Me-germacrene A (**5a**, Scheme 76). We observed that **4a** and **5a**, which are proposed to be (*E*,*E*)-6Me-germacrene A and (*E*,*Z*)-6Me-germacrene A, respectively, do not coelute from the GC column. This is in contrast with that observed for **2a** and **3a**, which are proposed to be (*E*,*Z*)-6F-germacrene A and (*E*,*E*)-6F-germacrene A, and do coelute in the GC column. The chromatographic properties of **4a** and **5a** methylated isomers differ due to the presence of the extra methyl substituent present in their C6- C7 double bound, and this is opposite when fluorine is present at this position.

The mass spectrum of **5b** shows a molecular ion peak with m/z = 236 ($[M]^+$), indicating the presence of a methylated sesquiterpene alcohol (Figure 106b). This compound was identified as (*E*,*Z*)-6Met-farnesol through comparison with authentic synthetised compound, **206**.The presence of farnesol suggest that GAS behaves less precise to perform the 1,10-cyclisation when using **5** (in comparison to the use of the natural substrate, **1**).

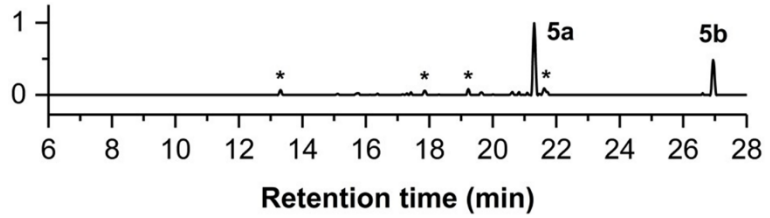


Figure 105. Total ion chromatogram of the pentane extractable products arising from incubation 37.

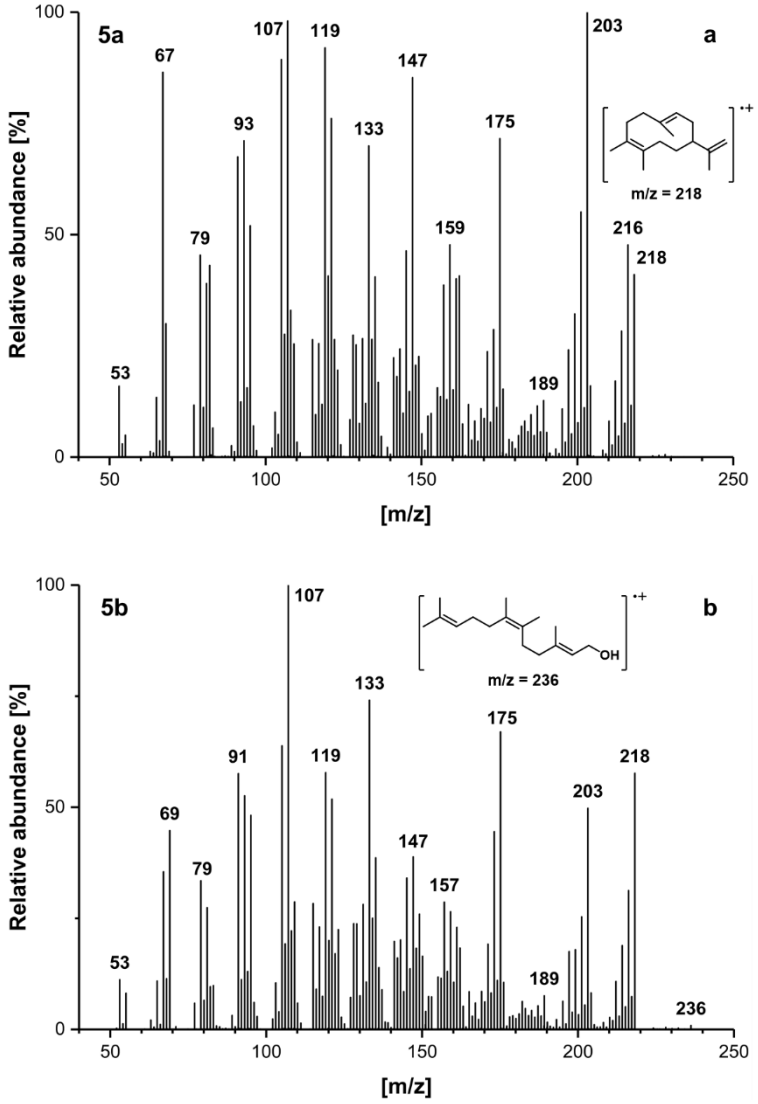
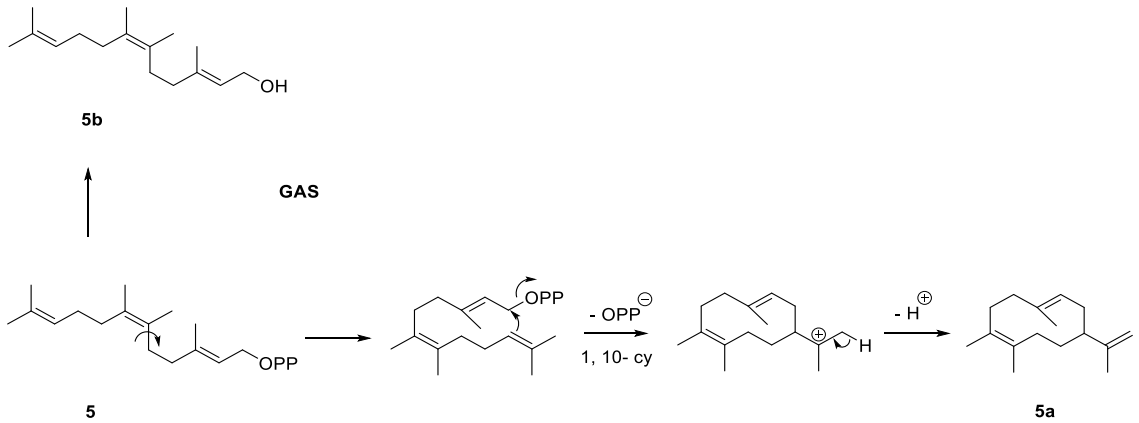


Figure 106. a) Mass spectrum of the product eluting at 21.30 min (**5a**) obtained from incubation 37, showing the molecular ion of (2*E*,6*Z*)-6Me-germacrene A (**5a**). b) Mass spectrum of the product eluting at 26.03 min (**5b**) obtained from incubation 37, showing the molecular ion of (2*E*,6*Z*)-6Me-farnesol (**5b**).



Scheme 76. Proposed catalytic mechanism by GAS to generate **5a** and **5b** from **5**.

Incubation 38. The incubation of GDS with **5** also yielded **5a** as the major product in the pentane extracts, eluting at 21.30 min in the chromatogram (Figure 107). In addition, 4 minor products were detected. This is in contrast as compared with the incubation of **3** with GDS (incubation 20), where GDS generates exclusively **3a**. The conformation of (2*E*,6*Z*)-6Me-FDP (**5**) hinders the formation of an allylic secondary carbocation (Scheme 77), and other reaction pathways arise.

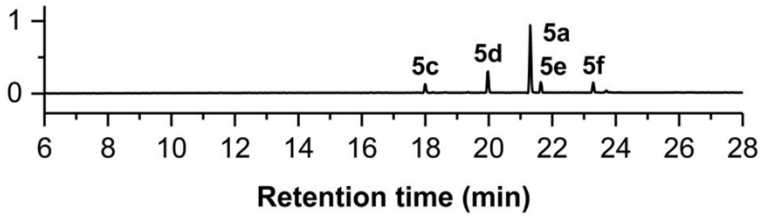


Figure 107. a) Total ion chromatogram of the pentane extractable products arising from incubation 38.

Incubation 39. The incubation of **5** with GdolS produced **5a** almost exclusively, as judged by GC-MS analysis of the pentane extracts (eluting at 21.30 min, Figure 108). The production of **5a** is not surprising, as it would form a non-stabilised secondary carbocation following the natural cyclisation (Scheme 77), which is a common feature with GDS. However, this incubation also reveals that GdolS is the most selective germacrene synthase towards the conversion of **5** to **5a**, including GAS, GDS and G11olS among these.

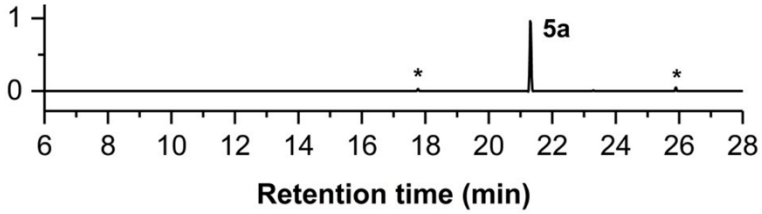
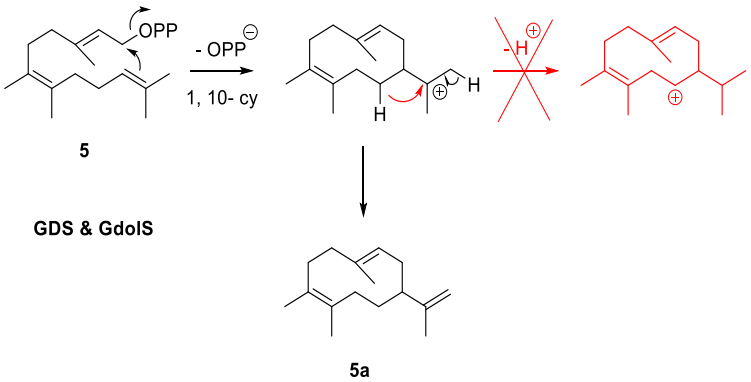


Figure 108. a) Total ion chromatogram of the pentane extractable products arising from incubation 39.



Scheme 77. Proposed interception of the reaction mechanisms catalysed by GDS and GdolS upon incubation with **5** (red), and formation of **5a** (black).

Incubation 40. The incubation of **5** with Gd11olS gave a predominant product in the organic extracts eluting at 21.30 min, which was identified as **5a** (Figure 109a). Another minor methylated sesquiterpene was found eluting at 26.20 min (labelled as **5g**). Its late retention time suggests it might be an alcohol compound. This compound does not coelute with (2*E*,6*Z*)-6Me-farnesol (previously identified, **5b**. The mass spectrum has a molecular ion peak with m/z = 218 ([M]$^+$), and there is no evidence to prove the presence of the hydroxyl group (Figure 109b). This minor product remains unidentified.

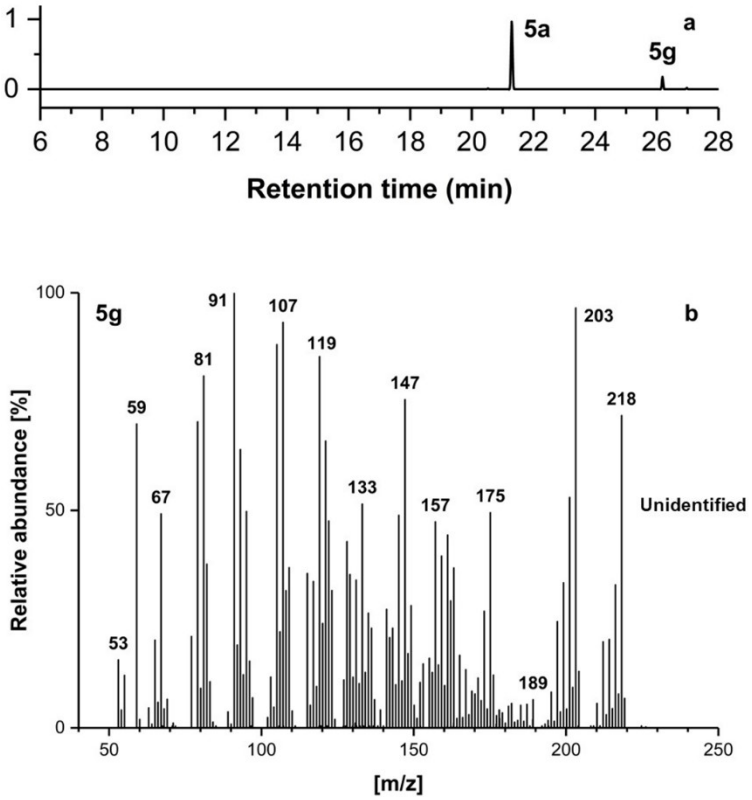


Figure 109. a) Total ion chromatogram of the pentane extractable products arising from incubation 40. b) Mass spectrum of **5g**, obtained from incubation 40, and remains unidentified.

In summary, all the 1,10-sesquiterpene synthases used here generated **5a**, which is postulated to be (2*E*,6*Z*)-6Me-germacrene A. GdolS is a high fidelity sesquiterpene synthase in the production of **5a** upon incubation with **5**. GAS and Gd11olS produced larger amounts of coproducts as compared with GdolS. GDS behaves promiscuously when it is incubated with **5**, and four other products are formed apart from **5a**. The results obtained upon the incubation of **5** with 1,10-sesquiterpene synthases are in concordance with the results from the incubation of **3** with these enzymes, where it was observed that **3a** was a high selective product. However, the presence of the methyl group in the position C6 increase the catalytic promiscuity of these enzymes in comparison with the presence of a fluorine atom, which is further indicative of a disrupted enzyme-substrate complex geometry.

1,6- Sesquiterpene synthase

Incubation 41. The incubation of **5** with ADS yielded a mixture of methylated products in the organic extracts (Figure 110). This promiscuity exerted by ADS upon incubation with **5** might be the result of unproductive conformations of **5** or the intermediates formed during the course of this reaction, that lead to aberrant structures. **5a** was observed as a product, which eluted at 21.30 min, but in very minor proportion in comparison with the incubation of **5** with 1,10-sesquiterpene synthases. ADS might follows the natural pathway (isomerisation through NDP) and access a conformational source structure for the 1,6-cyclisation (Scheme 78). At this point, the reaction might differ from incubation 23 (incubation of **3** with ADS), due to the lack of electronic destabilisation of the forming bisabolyl cation. However, the newly incorporated methyl group may influence the conformation of the substrate in the active site, and this may be the main factor for the observed promiscuity of ADS upon incubation with **5** (Scheme 78). Beside these mechanistic speculations, it was difficult to identify these products by mass spectrum analysis.

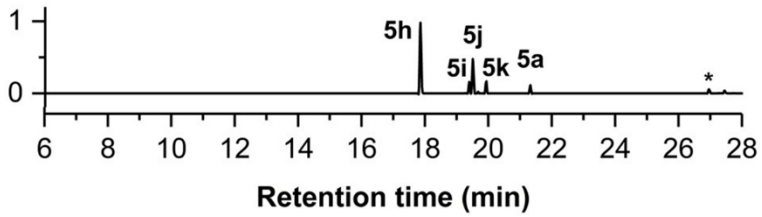
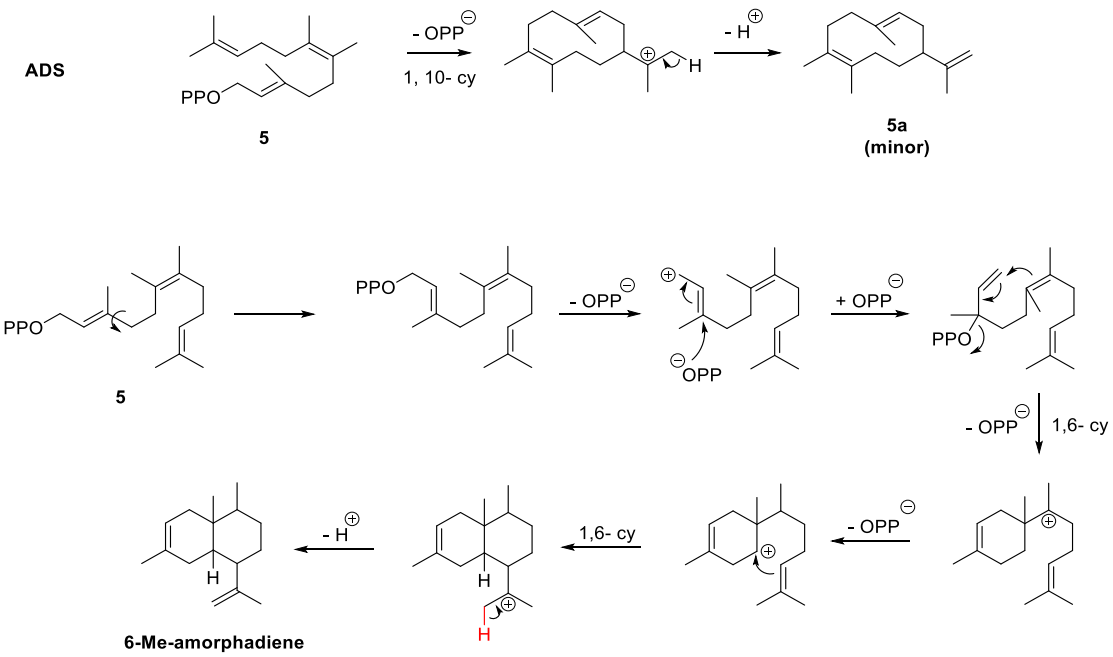


Figure 110. Total ion chromatogram of the pentane extractable products arising from incubation 41.



Scheme 78. Top, proposed reaction mechanism catalysed by ADS to generate **5a** from **5**. Bottom, possible reaction mechanism to access 6Me-amorphadiene.

1,10- + 1,6- Sesquiterpene synthases

Incubation 42. The incubation of **5** with PRAS gave almost exclusively **5a** in the organic extracts, eluting at 21. 30 min (Figure 111). This result fully agrees with that observed for the incubation of **3** with PRAS. PRAS catalysed to conversion of **5** to **5a**, which conformation is unproductive for an intramolecular 1,6-cyclisation (Scheme 79).

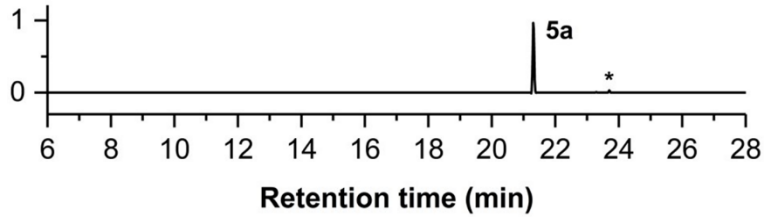
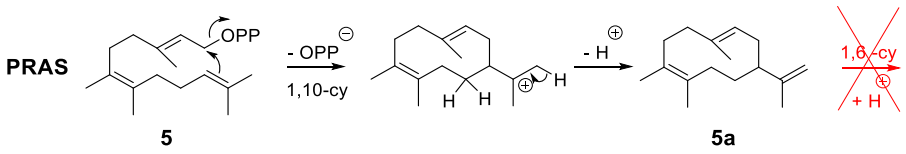


Figure 111. Total ion chromatogram of the pentane extractable products arising from incubation 42.



Scheme 79. In red, proposed interception of the reaction mechanism catalysed by PRAS upon incubation with **5**. In black, proposed reaction mechanism for the conversion of **5** into **5a** by PRAS.

Incubation 43. The incubation of **5** with SdS produced **5a** as the main product, as judged by GC-MS analysis of the pentane extracts (eluting at 21.30 min, Figure 112a). The second major product, which eluted at 23.05 min (labelled as **5l** in Figure 112a), could be anticipated to be (2*E*,6*Z*)-6Me-germacrene B. This is because SdS has been shown able to catalyse the conversion of **3** to **3f**, which was postulated to be (2*E*,6*E*)-6F-germacrene B (incubation 25, Scheme 65). In fact, the mass spectrum of **5l** contains the most stable fragment with m/z = 121 (Figure 112b), a shared feature with **1h** (germacrene B), **2h** (postulated as (2*E*,6*Z*)-6F-germacrene B) and **3f**.

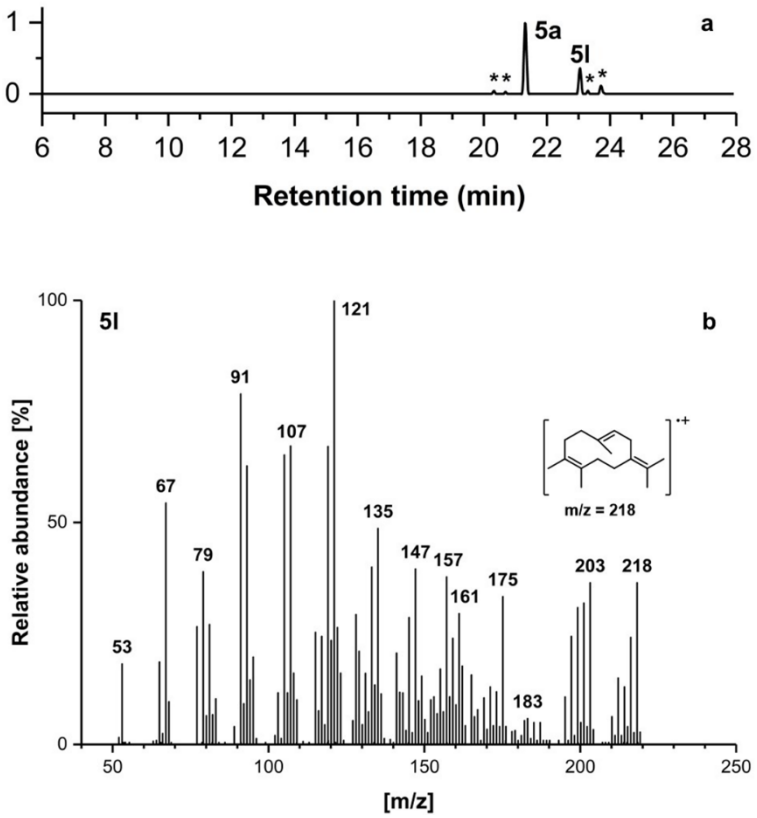
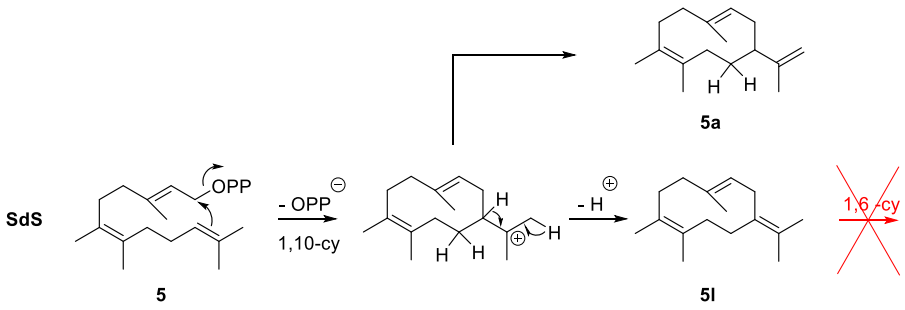


Figure 112. a) Total ion chromatogram of the pentane extractable products arising from incubation 43. b) Mass spectrum of **5l**, obtained from incubation 43, showing the molecular ion of (2*E*,6*Z*)-6Me-germacrene B (**5l**).



Scheme 80. In red, proposed interception of the reaction mechanism catalysed by SdS upon incubation with **5**. In black, proposed mechanism for the conversion of **5** into **5a** and **5l** by SdS.

Incubation 44. The incubation of EpicS with **5** gave **5a** as the major compound in the pentane extracts eluting at 21.30 min, Figure 113. The presence of (2*E*,6*Z*)-6Me-farnesol (**5b**), which eluted at 26.94 min was also observed. This was in concordance with incubation 23, in which **3a** was obtained as the main product upon the incubation of **3** with EpicS. The natural 1,10-cyclisation *via* NDP formation (isomerisation) is unlikely in this case (Scheme 81). Alternatively, EpicS catalyse the formation of **5a** *via* farnesyl cation formation-induced 1,10-cyclisation (Scheme 81). The presence of farnesol **5b** might be attributed to a low precision of EpicS to allocate the substrate with the C1 and C10 in reactive positions to form a C-C bond.

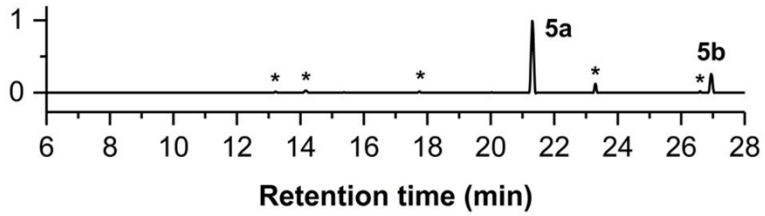
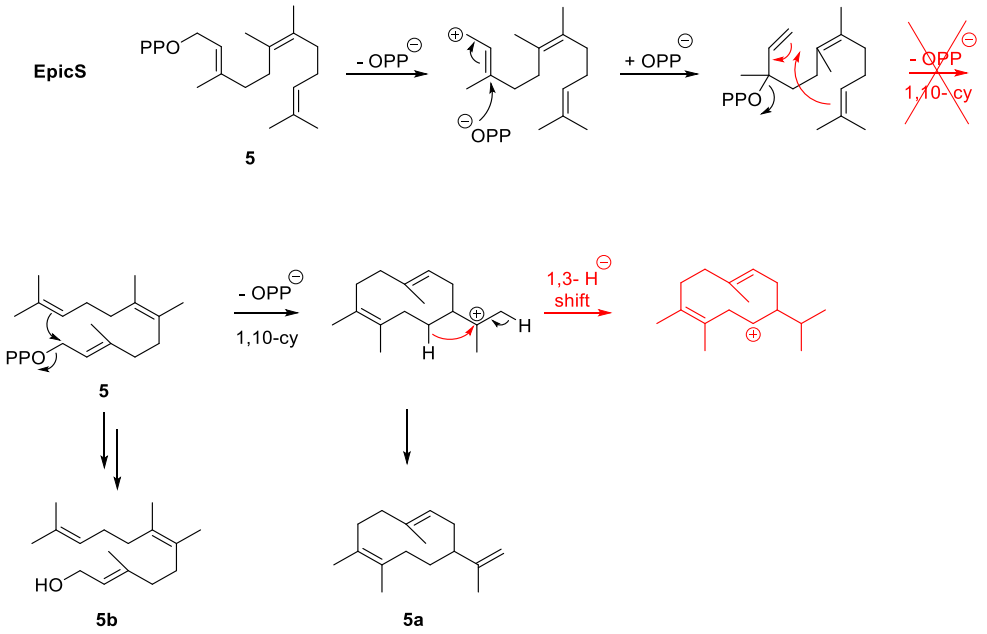


Figure 113. Total ion chromatogram of the pentane extractable products arising from incubation 44.



Scheme 81. In red, proposed interception of the reaction mechanisms catalysed by EpicS upon incubation with **5**. In black, proposed mechanism for the conversion of **5** into **5a** and **5b** by EpicS.

Incubation 45. The incubation of **5** with DCS-W279Y yielded **5a** as the sole organic extractable product, eluting at 21.30 min in the chromatogram (Figure 114). This agrees with incubations 18, 39 and 44 among others. DCS-W279Y catalyses a 1,10-cyclisation from **5** and generates **5a**, being unable to proceed with the second (natural) 1,6-cyclisation towards cadinene formation.

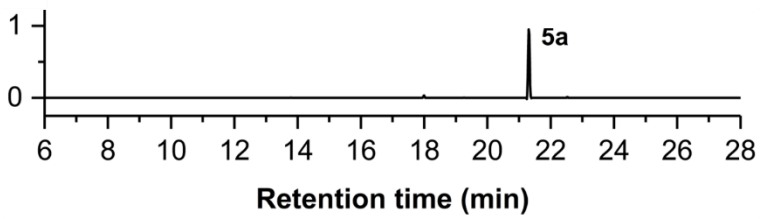


Figure 114. Total ion chromatogram of the pentane extractable products arising from incubation 45.

In summary, the incubations of **5** with 1,10- and 1,16- sesquiterpene synthases yielded **5a** as the main organic extractable product, with exception of SdS, which may be able to generate (*E*,*Z*)-6Me-germacrene B (**5l**).

Controls were performed by incubation of **4** without enzyme under otherwise identical conditions to check that the formed products are of enzymatic origin. No products were detected in these control incubations. Also, 2*E*,6*Z*-6Me-farnesol (**206**) was analysed by GC-MS under similar conditions as the incubation extracts, which it correspond to **5b**, a product obtained after the incubation of **5** with GAS and EpicS.

4.7. Substrate behaviour of (2*E*,6*E*)-7H-FDP (6)

The incubations of (*E*,*E*)-7H-FDP (**6**) with sesquiterpene synthases is analysed here for the first time (Figure 115).

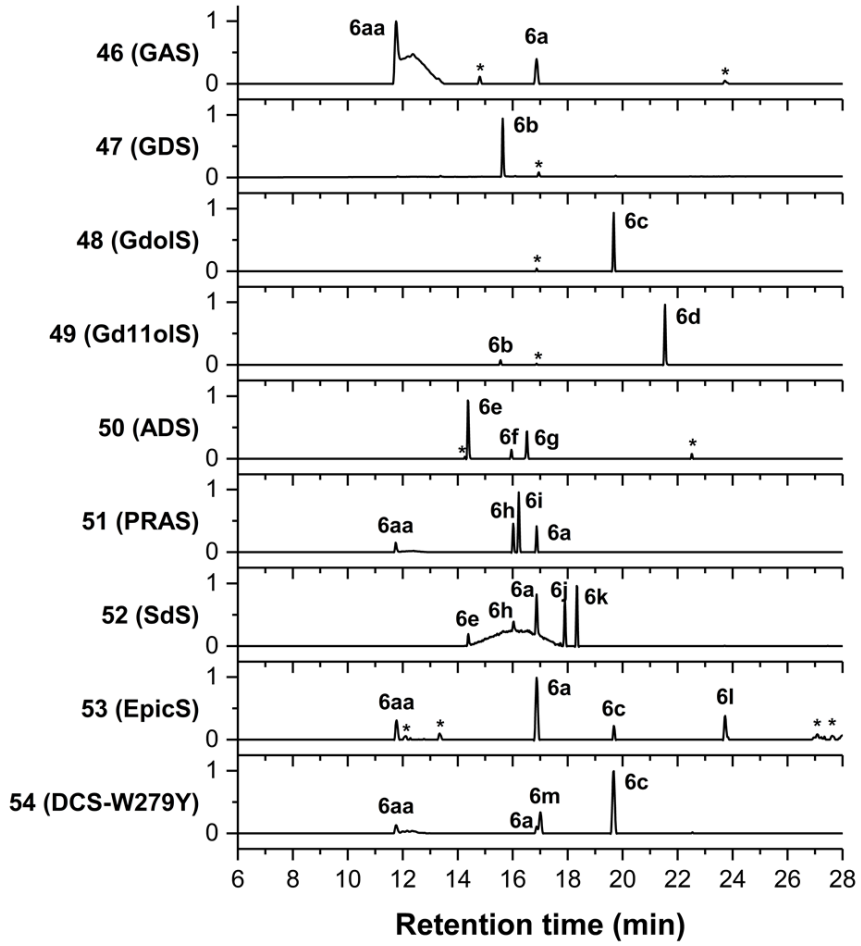


Figure 115. Gas chromatogram of the pentane extractable products after incubations of (2*E*,6*E*)-7H-FDP (**6**) with nine sesquiterpene synthases. Selected minor compounds are labelled with an asterisk.

1,10- Sesquiterpene synthases

Incubation 46. The incubation of (2*E*,6*E*)-7H-FDP (**6**) with GAS yielded two main detectable products, as judged by GC-MS analysis of the organic extracts (Figure 116a). These compounds eluted at 11.74 min and 16.87 min. The sharp peak observed for the compound eluting at 11.74 min (**6aa**, Figure 116a) was accompanied by a broad signal characteristically observed for germacrene A-like compounds as a result of their thermal instability.[122,276] This suggests that **6aa** is a germacrene A thermally-induced derivative. The compound eluting at 16.87 min (labelled as **6a**, Figure 116b) has a mass spectrum that displays a molecular ion with m/z =190 ($[M]^+$), indicating the absence of a methyl group from the natural FDP (**1**) backbone. Its most stable fragment appears with m/z = 147 ($[M - C_3H_7]^+$), indicating the loss of an isopropyl unit. In comparison, the mass spectra from **6a** and **6aa** present identical fragmentation pattern (Figures 116b and 117, respectively). With this, **6a** is hypothesised to be 7H- germacrene A and **6aa** is its elemene derivate, which is the main observed compound and has been probably generated during the analysis. To probe this, we induced the thermal cope rearrangement of **6a** by injecting the sample into the GC injection port at higher temperature (250 ºC), Figure 118 and Scheme 82.[276] This resulted in the disappearance of **6a** in the chromatogram, which was fully converted to **6aa** with these conditions (Figure 118). In addition to these, a minor product was observed eluting at 14.89 (labelled with an asterisk, Figure 116a), which also disappeared from the gas chromatogram upon the Cope-rearrangement (Figure 118), so this might also be a germacrene A derivative.

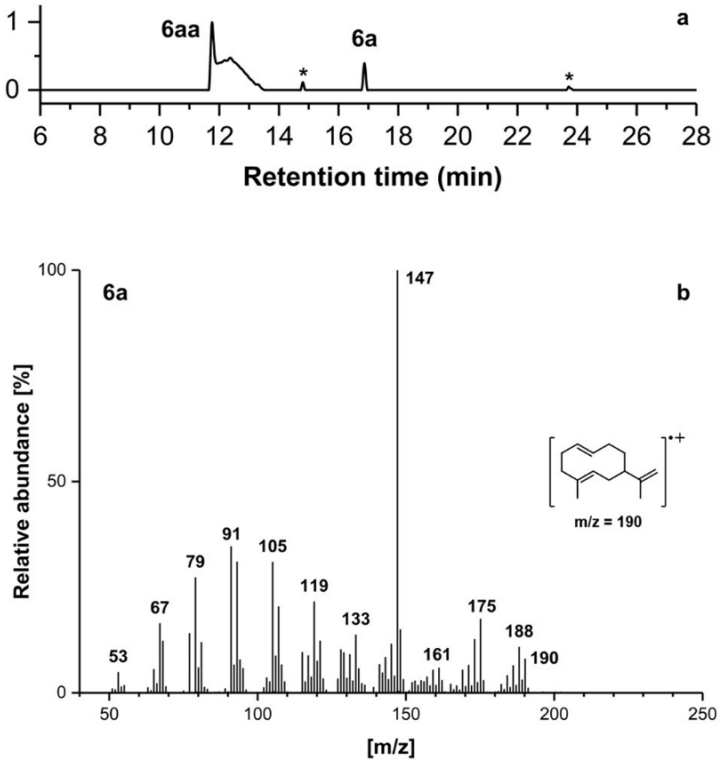


Figure 116. a) Total ion chromatogram of the pentane extractable products arising from incubation 46. b) Mass spectrum of **6a**, obtained from incubation 46, showing the molecular ion of (2*E*,6*E*)-7H-germacrene A.

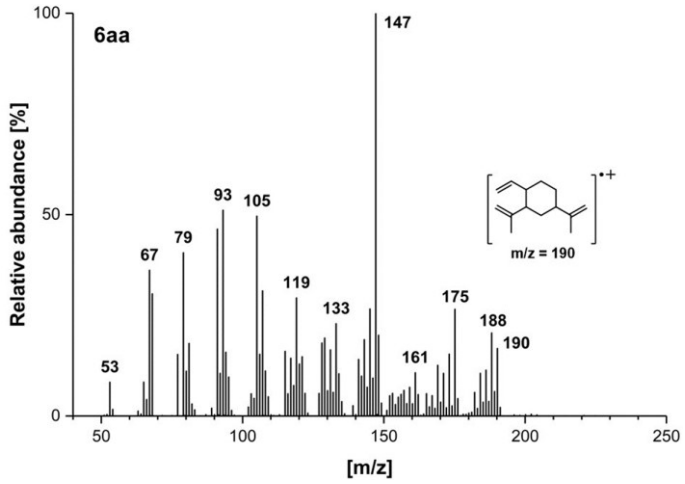


Figure 117. b) Mass spectrum of **6aa**, obtained from incubation 46, showing the molecular ion of β-7H-elemene (**6aa**).

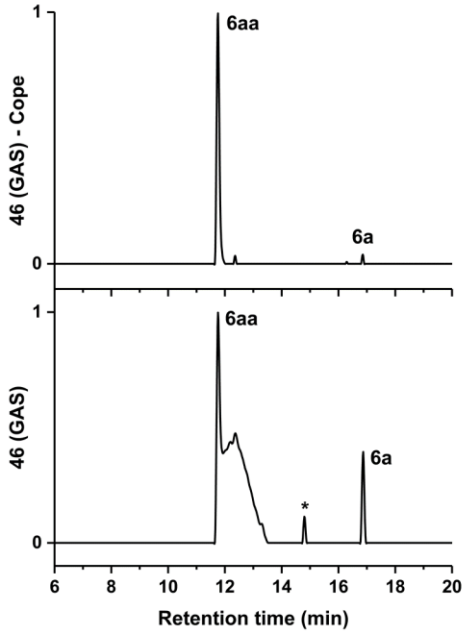
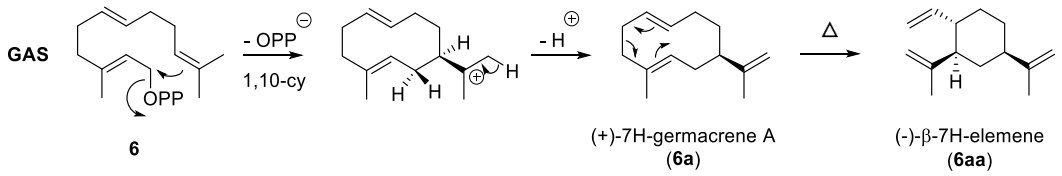


Figure 118. a) Total ion chromatogram of the pentane extractable products arising from incubation 46, with an injection port at 250 ºC (Cope rearrangement). b) Total ion chromatogram of the pentane extractable products arising from incubation 46.



Scheme 82. Proposed reaction mechanism catalysed by GAS to generate **6a** from **6**, and thermal Cope rearrangement of (2*E*,6*E*)-7H-germacrene A (**6a**) to β-7H-elemene (**6aa**).

Incubation 47. The incubation of GDS with **6** gave a predominant product in the organic extracts, as judged by GC-MS (eluting at 15.56 min, **6b**, Figure 119a). The mass spectrum of **6b**, showed a molecular ion with m/z = 190 ($[M]^+$), indicative that it is a 7H- sesquiterpene. Its most stable fragment peak appears with m/z = 147 ($[M - C_3H_7]^+$), which is consistent with the loss of an isopropyl group from the parent ion (Figure 119b). This fragment is also predominant in the mass spectra of (-)-germacrene D (**1b**) (m/z = 161, $[M - C_3H_7]^+$), (*E*,*Z*)-6F-germacrene D (**2b**) (m/z = 159, $[M - HF - C_3H_7]^+$) and (*E*,*E*)-6Met-germacrene D (**4b**). This data suggests that **6b** is (2*E*,6*E*)-7H-germacrene D. Thus, the reaction catalysed for GDS upon incubation with **6** proceeds following the natural pathway (Scheme 83).

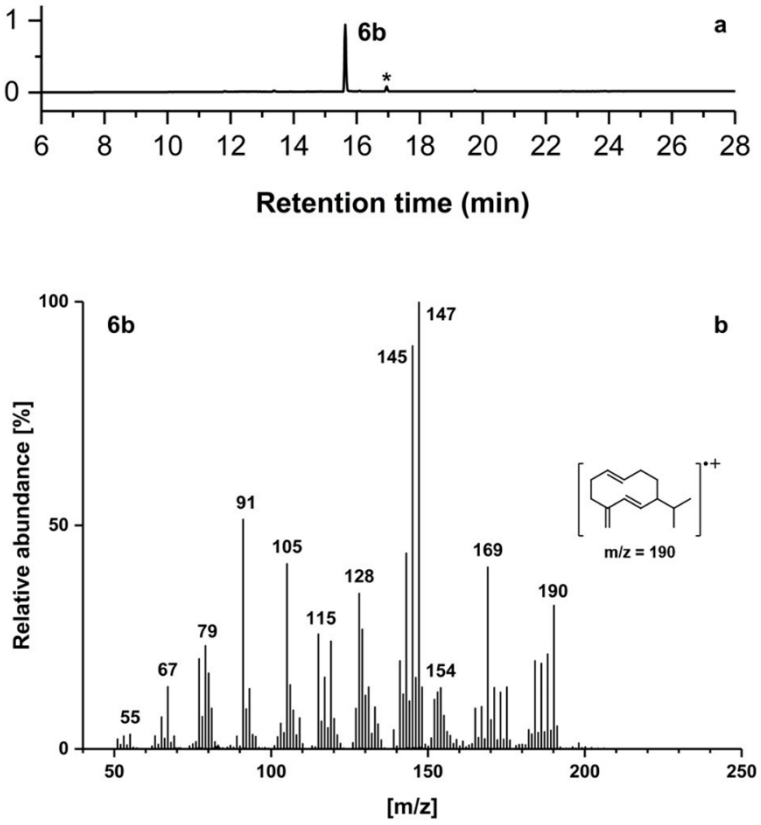
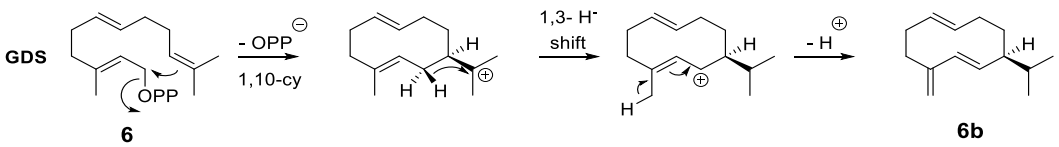


Figure 119. a) Total ion chromatogram of the pentane extractable products arising from incubation 47. b) Mass spectrum of **6b**, obtained from incubation 47, showing the molecular ion of (2*E*,6*E*)-7H-germacrene D (**6b**)



Scheme 83. Proposed catalytic mechanism by GDS to generate **6b** from **6**.

Incubation 48. The incubation of **6** with GdolS gave compound **6c** almost exclusively in the organic extracts, eluting at 19.67 min (Figure 120a). The mass spectrum of **6c** has a molecular ion with m/z = 208 ($[M]^+$), suggesting the presence of a 7H-sesquiterpene alcohol (Figure 120b). In addition, it shows a major peak with m/z = 190 ($[M - H_2O]^+$) and an abundant fragment peak with m/z = 147 ($[M - C_3H_7]^+$). This data suggests that **6c** is (2*E*,6*E*)- 7H- germacradien-4-ol (Scheme 84). This assumption is in concordance with the production of (2*E*,6*E*)-7H-germacrene-4-ol (**6c**) by EpicS and DCS-W279Y, which served as controls (see incubations 53 and 54). In fact, the mass spectrum of **6c** is comparable to those of (-)-germacradien-4-ol (**1c**), (2*E*,6*Z*)- 6F-germacradien-4-ol (**2c**), and (2*E*,6*E*)-6Me-germacradien-4-ol (**4c**).

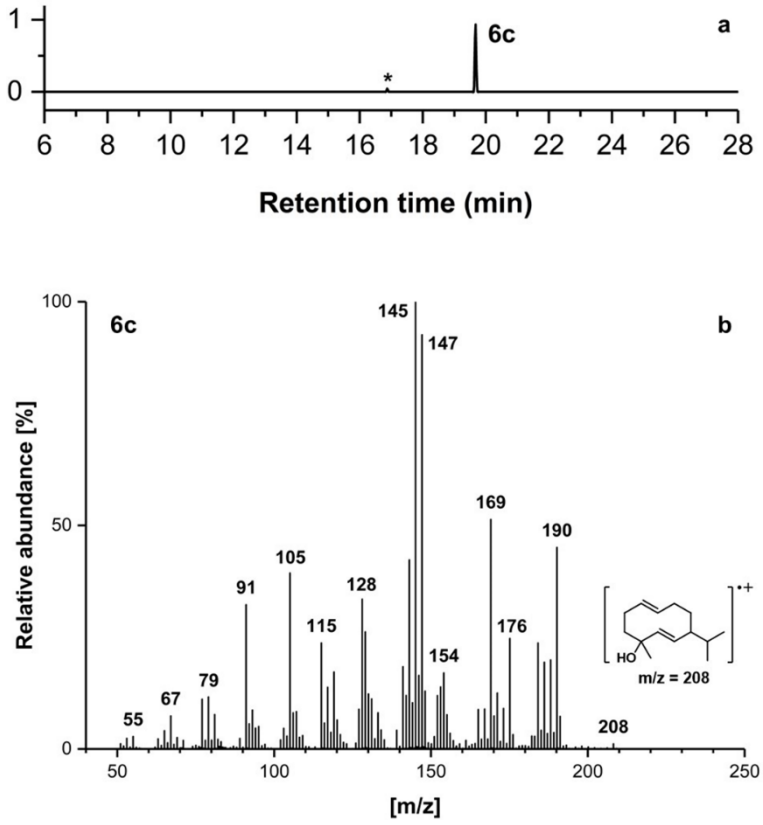
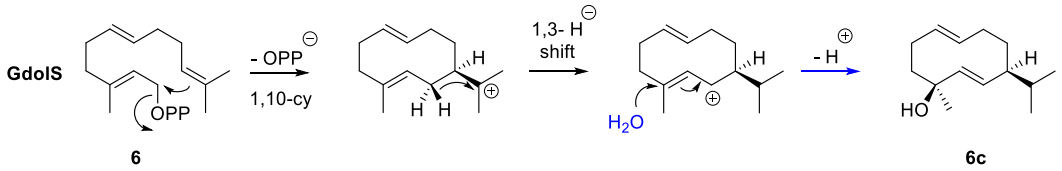


Figure 120. a) Total ion chromatogram of the pentane extractable products arising from incubation 48. b) Mass spectrum of **6c**, obtained from incubation 48, showing the molecular ion of (2*E*,6*E*)- 7H- germacradien-4-ol (**6c**)



Scheme 84. Proposed reaction mechanism catalysed by GdolS to generate **6c** from **6**.

Incubation 49. The incubation of **6** with Gd11olS gave a major product in the pentane extracts eluting at 21.54 min in the chromatogram (**6d**, Figure 121a). The mass spectrum of **6d** displays a molecular ion peak with m/z = 208 ($[M]^+$), indicating that it might be a 7H-sesquiterpene alcohol (Figure 121b). The most stable fragment appears with m/z= 147 ($M - H_2O - C_3H_7]^+$). In view of previous results with 1,10-sesquiterpene synthases GAS, GDS and GdolS, the proposed mechanism catalysed by G11olS should not be altered upon incubation with the modified analogue **6** in comparison with the natural proceeding pathway (Scheme 85). Hence, we propose that **6d** is (2*E*, 6*E*)-7H-germacradien-11-ol. 7H-Germacrene D (**6b**) is also observed as a minor product eluting at 15.56 min. This result agrees with the reported minor product observed for the incubation of **1** with Gd11olS, which it is germacrene D.[274]

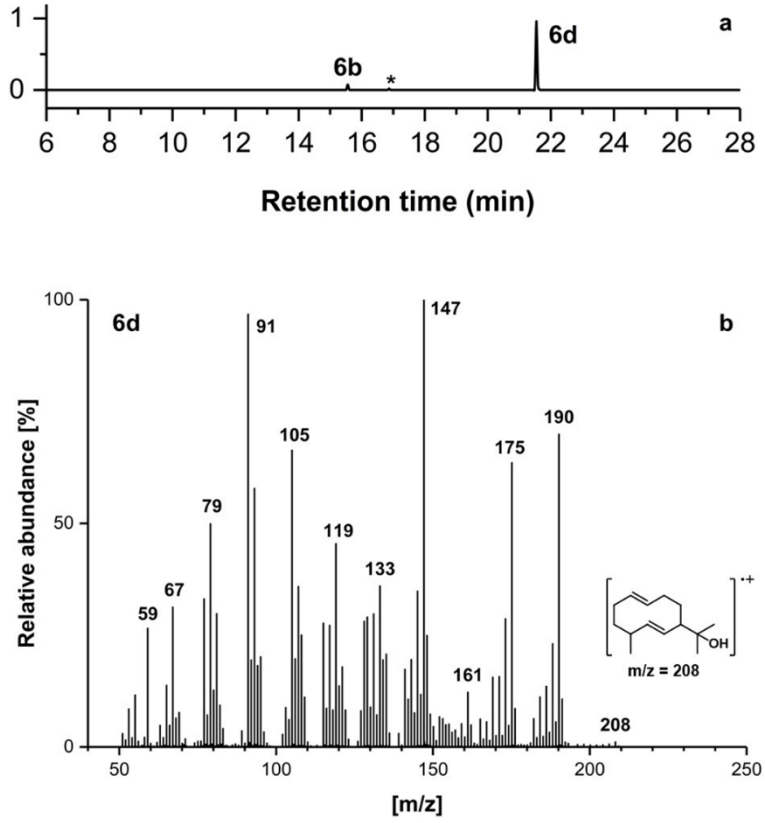
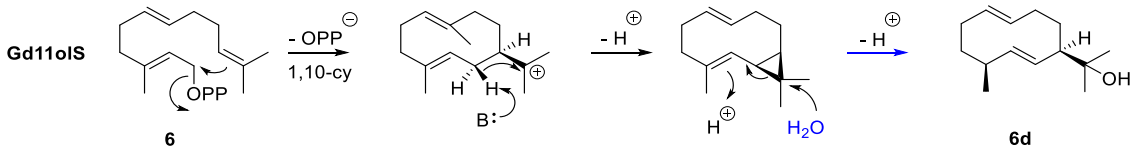


Figure 121. a) Total ion chromatogram of the pentane extractable products arising from incubation 49. b) Mass spectrum of **6d**, obtained from incubation 49, showing the molecular ion of (2*E*,6*E*)- 7H- germacradien-11-ol (**6d**).



Scheme 85. Proposed reaction mechanism catalysed by GdolS to generate **6d** from **6**.

In summary, the results obtained upon the incubation of **6** with 1,10- sesquiterpene synthases reveal that the removal of the methyl group at C6 does not result in a drastic change to the reactions catalysed. Thus, the presence of the methyl group at the C6 position of FDP is not a crucial requirement for generating a productive substrate and interestingly its absence does not result in a cavity in the enzyme-substrate complex that might lead to greater product diversity.

The appearance of **6aa** in the chromatogram arising from incubation of **6** with GAS is understood as a thermal rearrangement, which might have occurred during the analysis.

The germacrene synthases used here (GAS, GDS, Gd4olS and Gd4olS) generate the 7H- analogues of their corresponding natural products, and this might be extended to other similar 1,10-sesquiterpene synthases. The use of **6** with germacrene synthases serves as a platform for the production of a series of novel compounds, of which only a few have been described here.

1,6- Sesquiterpene synthase

Incubation 50. The incubation of **6** with ADS, in contrast with the use of 1,10-sesquiterpene synthases, gave a mixture of at least 3 pentane extractable products eluting at 14.38 min, 15.65 min and 16.52 min (labelled as **6e**, **6f** and **6g**, respectively, Figure 122). The absence of the methyl at C6 of FDP might affect the naturally proceeding 1,6-cyclisation in ADS (Scheme 50). This is in contrast with the results obtained upon the incubation of (2*E*,6*E*)-6Me-FDP (**3**) with ADS, which resulted in selective production of 6Me-amorphadiene (**4g**, See above in incubation 32). The incubation of ADS with **6**, as in the case of **1**, might start with the generation of NDP, however, the next step might be truncated with the absence of the methyl substituent in the C6-C7 double bond, probably because now a high energy secondary cation would be formed (Scheme 86), rather than the tertiary carbocation observed during the cyclisation of **1** with ADS (Scheme 50).

This could result in an alternative 1,7-cyclisation; however, this is unlikely as the forming carbocation would be also secondary (Scheme 86). 1,10- or 1,11-cyclisations might be alternative proceedings, but they are unlikely, as ADS is exclusively a 1,6-sesquiterpene synthase. Also, the enzyme might generate farnesenes or farnesol derivates of **6**, resulting from aberrant deprotonation or water quench of the linear substrate or intermediates (Scheme 86). The presence of (2*E*,6*E*)-7H-farnesol (**217**) was ruled out as the elution time and the fragmentation patterns of **6e**, **6f** or **6g** did not match with a synthetic standard of **217.** The identity of these compounds has not been further investigated due to time constraints and data availability. However, this might be farnesenes and/or other products arising from aberrant deprotonations along the course of the reactions.

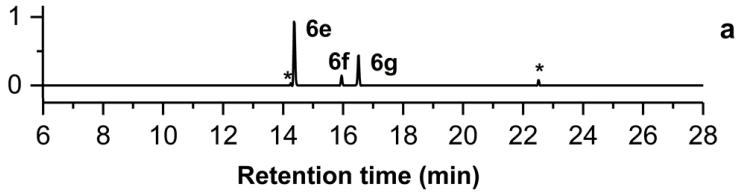
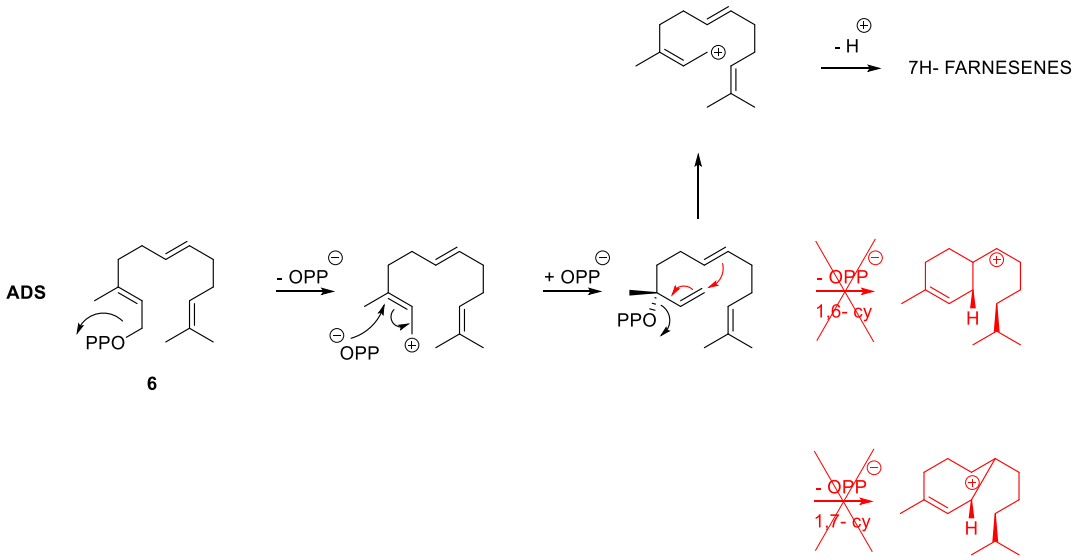


Figure 122. a) Total ion chromatogram of the pentane extractable products arising from incubation 50.

These results show that ADS generate a more diverse array of products, suggesting that in this case, the absence of the C7 methyl group allows more flexibility and motion of reaction intermediates during catalysis.

At least one of the products observed may be a farnesene, which would indicate the failure of the initial cyclisation in ADS. The formation of a secondary carbocation is energetically restricted, which might lead to aberrant products (farnesenes or others).



Scheme 86. In red, proposed interception of the reaction mechanism catalysed by ADS upon incubation with **6** (including an alternative 1,7-cyclisation). In black, proposed reaction mechanism for the conversion of **6** to possible 7H-farnesenes.

1,10- + 1,6- Sesquiterpene synthases

Incubation 51. The incubation of PRAS with **6** gave a mixture of four products in the organic extracts as judged by GC-MS analysis (Figure 123). **6a** and its elemene derivate **6aa** are observed eluting at 16.87 min and 11.74 min, respectively. This is in accordance with incubation 5, where it is observed the production of (+)-germacrene A (**1a**) as minor product upon incubation of **1** with PRAS. This also agrees with incubation 46, where both **6a** and **6aa** were observed upon incubation of **6** with GAS. Apart from these, two more products eluted at 16.02 min and 16.22 min (**6h** an **6i**, Figures 123). The mass spectra of these products display the corresponding molecular ions with m/z = 190 ($[M]^+$), indicating the presence of pure hydrocarbon 7H-sesquiterpenes (Figures 125a and 125b). They both contain a major fragment peak with m/z = 147 ($[M - C_3H_7]^+$), suggesting that they might be germacrene analogues. The Cope rearrangement of the observed mixture of products was tested and it was only observed **6aa** an minor quantity of **6a** in the gas chromatogram (Figure 124). Based on these results, it is proposed that **6h** and **6i** are germacrene A derivatives.[122]

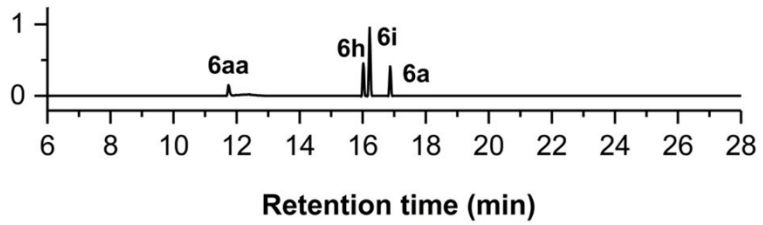


Figure 123. Total ion chromatogram of the pentane extractable products arising from incubation 51.

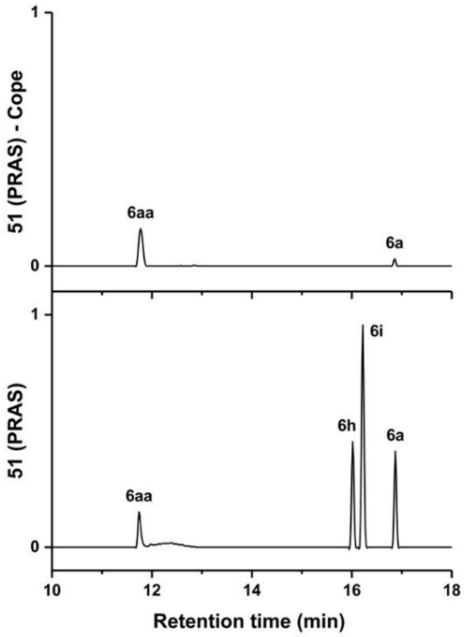


Figure 124. Top, total ion chromatogram of the pentane extractable products arising from incubation 51, with an injection port at 250 ºC (Cope rearrangement). Bottom, total ion chromatogram of the pentane extractable products arising from incubation 51.

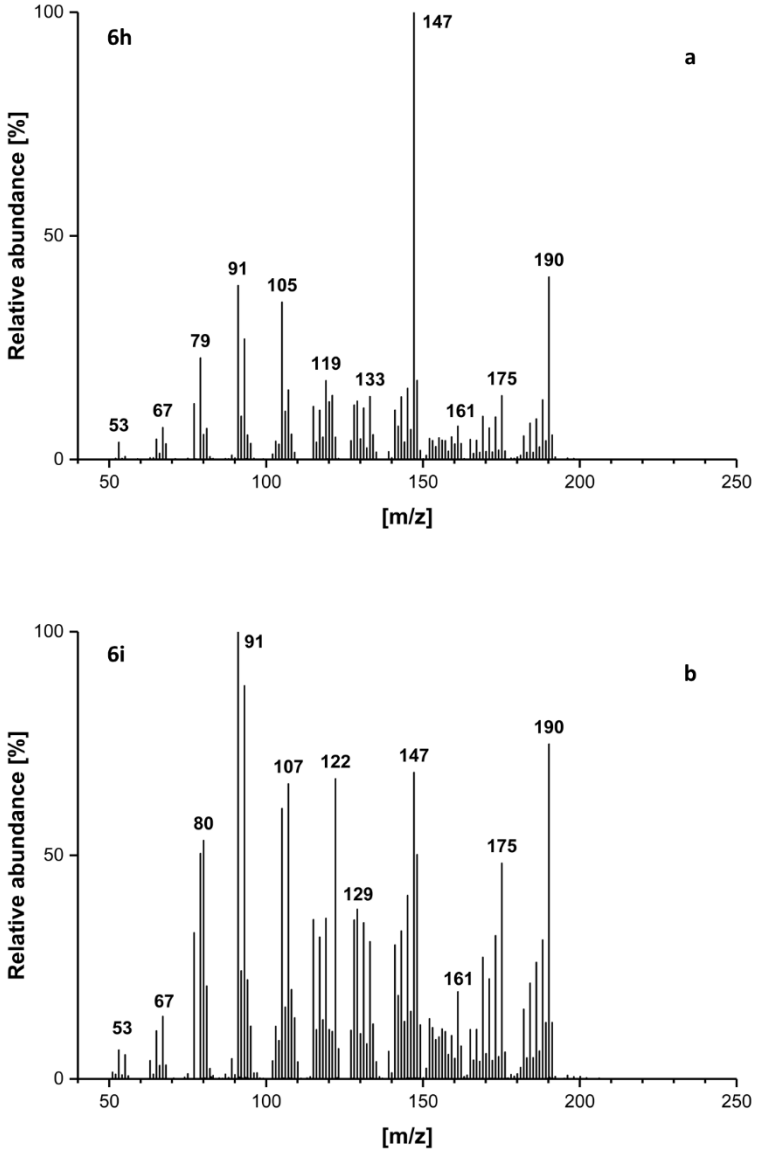


Figure 125. a) Mass spectrum of the product eluting at 16.02 min (**6h**) obtained from incubation 51, which remains unidentified. b) Mass spectrum of the product eluting at 16.22 min (**6i**) obtained from incubation 51, which remains unidentified.

Incubation 52. The incubation of SdS with **6** also yielded a mixture of 7H-sesquiterpenes in the organic extracts, as judged by GC-MS analysis (Figure 126). The mass spectra of these products have molecular ions with m/z = 190 ($[M]^+$). **6a** was present in the mixture accompanied with a broad signal, previously observed for germacrene A and derivatives (incubation 46, 51 and references[122,276]). In addition to this, **6e** was produced, which is a common product with ADS and was been postulated as a possible 7H-farnesene derivative (incubation 50). **6h**, which eluted at 16.02 min, is also a product formed upon incubation of **6** with AS (incubation 51). This is proposed to be a germacrene A derivative but is not conclusively identified. Apart from these, 2 new compounds were generated by SdS upon incubation with **6**, eluting at 17.90 min (**6j**) and 18.34 min (**6k**), which are 7H-sesquiterpenes according to their mass spectra fragmentation patterns. In analogy with the observed products upon incubation of **6** with 1,10-sesquiterpene synthases (incubations 46 - 49), where the natural-like analogues are formed (**6a** - **6d**), and also with PRAS, where 7H-germacrene A (**6a**) is produced (incubation 51), (2*E*,6*E*)-7H-germacrene B analogue is expected to be formed in the incubation of **6** with SdS. The presence of 7H-selinadiene may be also possible following the natural occurring pathway in SdS (Scheme 87). These, however, could not be identified with the data available.

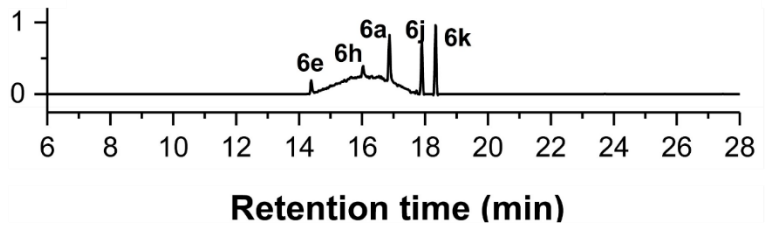
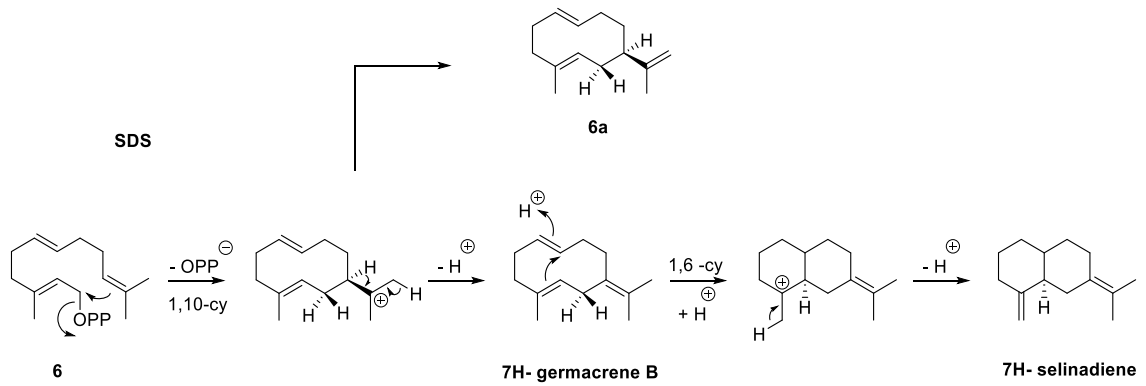


Figure 126. Total ion chromatogram of the pentane extractable products arising from incubation 52.



Scheme 87. In black, proposed reaction mechanisms for the conversion of **6** to **6a** by SdS, and possible formation of 7H-germacrene B and 7H-selinadiene, following the natural pathway exerted by SdS.

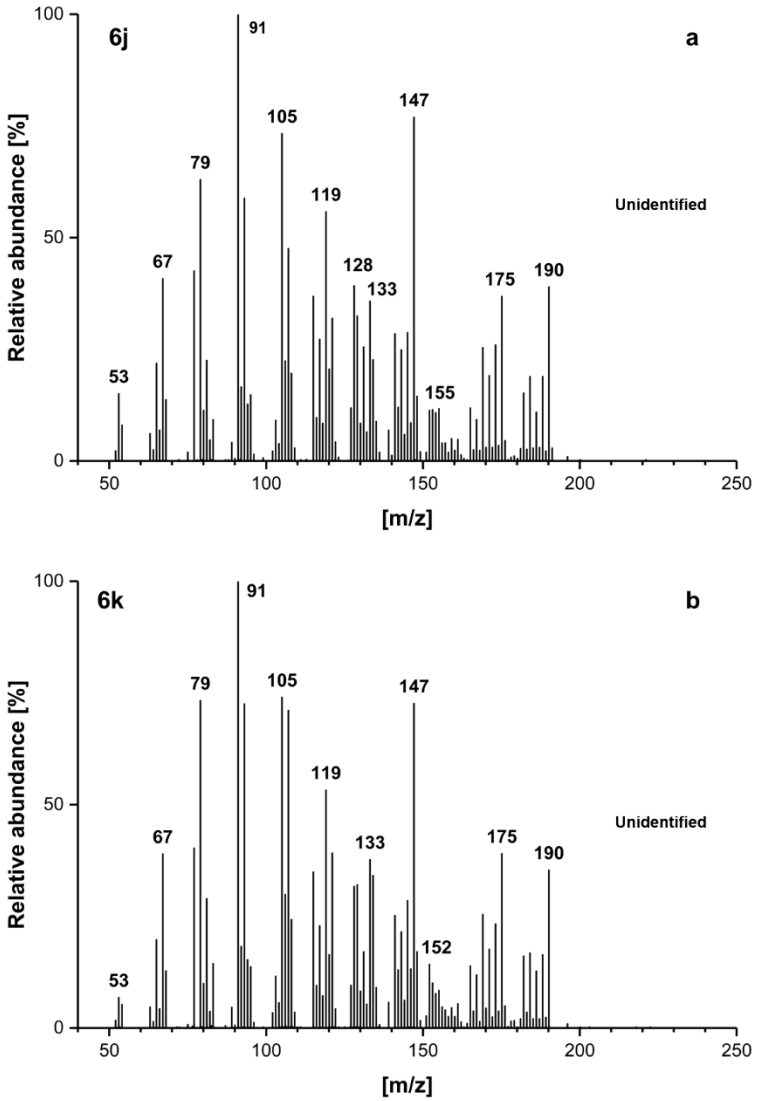


Figure 127. a) Mass spectrum of the product eluting at 17.90 min (**6j**) obtained from incubation 52, which remains unidentified. b) Mass spectrum of the product eluting at 18.34 min (**6k**) obtained from incubation 52, which remains unidentified.

Incubation 53. The incubation of **6** with EpicS produced **6a** as the mayor product in the organic extracts, together with its thermally induced derivate **6aa** (Figure 128a). As expected from previous analysis using EpicS (incubations 8 and 17), the production of (2*E*,6*E*)-7H-germacradien-4-ol (**6c**, eluting at 19.67 min) was detected, which is made by following the natural pathway catalysed by EpicS and water quenching at the allylic cation intermediate formed prior to 1,6- cyclisation (Scheme 88). An additional product eluted late at 23.74 min (**6l**) and its mass spectrum presents a molecular ion with m/z = 208 ([M]$^+$), Figure 128b, which matched the data of (2*E*,6*E*)-7H-farnesol (**212**), which was injected in the column for comparison. The absence of a 7H-epicubenol analogue is indicative that the 1,6-cyclisation is not proceeding upon incubation of **6** with EpicS, and this is because the nature of the forming carbocation, which would be secondary (Scheme 88).

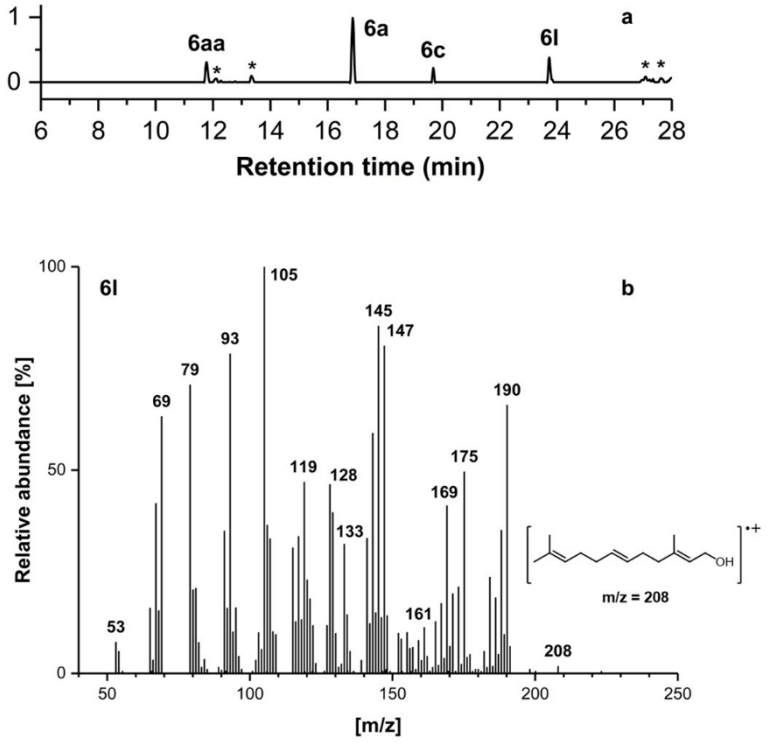
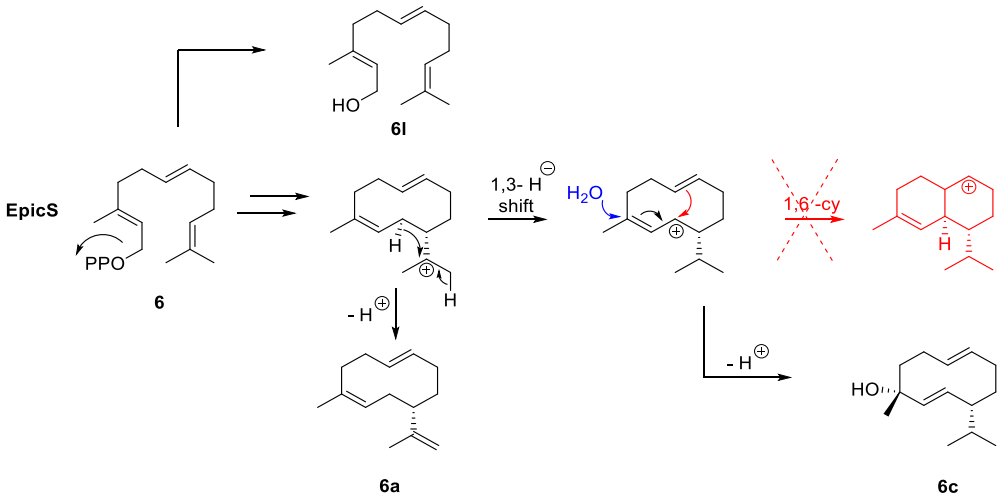


Figure 128. a) Total ion chromatogram of the pentane extractable products arising from incubation 53. b) Mass spectrum of **6l**, obtained from incubation 53, showing the molecular ion of (2*E*,6*E*)-7H-farnesol (**6l**)



Scheme 88. In red, proposed interception of the reaction mechanism catalysed by EpicS upon incubation with **6**. In black, proposed mechanisms for the conversion of 6 to (2*E*,6*E*)-7H-farnesol (**6l**), (2*E*,6*E*)-7H-germacrene A (**6a**) and (2*E*, 6*E*)-7H-germacradien-4-ol (**6c**),

Incubation 54. The incubation of **6** with DCS-W279Y gave **6c** as the main product, eluting at 19.67 min (Figure 129a). The products eluting at 11.74 min and 16.87 minutes correspond to **6aa** and **6a**, respectively. Apart from these, a novel 7H-sesquiterpene product eluted at 17.00 (**6m**). The mass spectrum of **6m**, which is a 7H-sesquiterpene, contains the most stable fragment with m/z = 147 ([M - C_3H_7]$^+$), which might indicate the loss of an isopropyl unit from a ten-membered ring structure, but this is unknown yet (Figure 129b). The presence of 7H-δ-cadinene, as for 7H-epicubenol (incubation 53), looks unfeasible due to the presumably formation of a secondary carbocation upon a putative 1,6-cyclisation (Scheme 89).

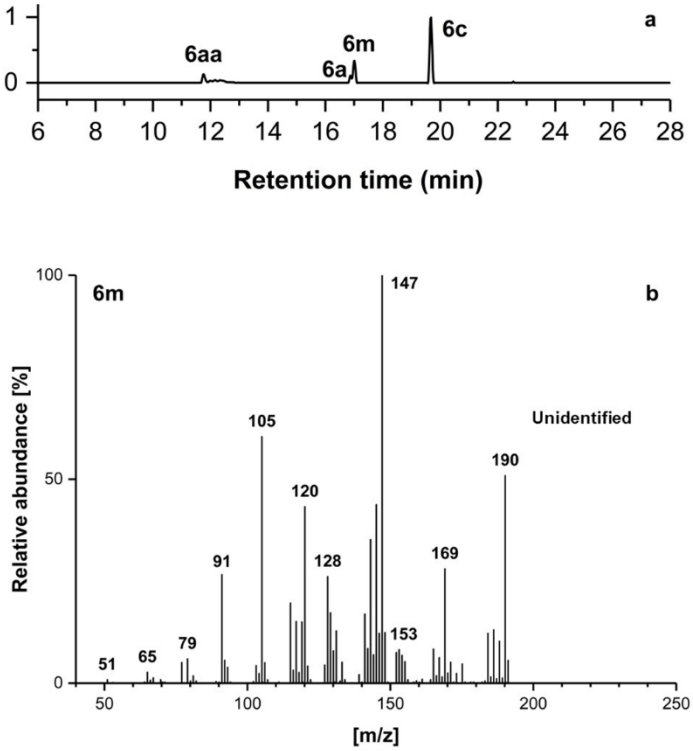
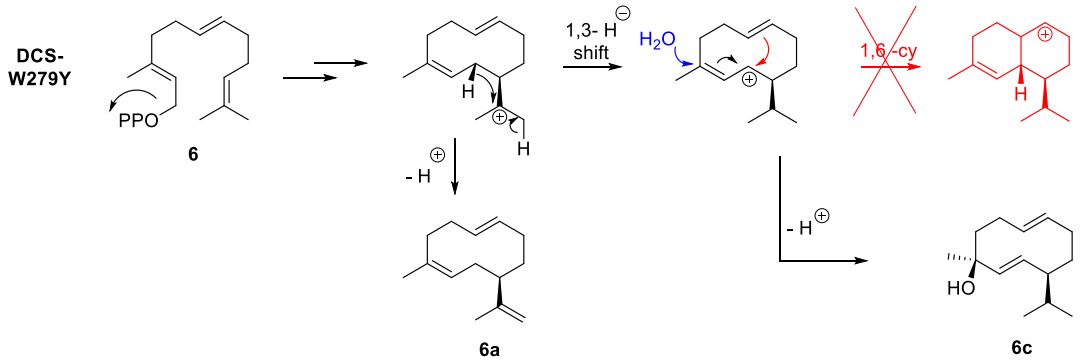


Figure 129. a) Total ion chromatogram of the pentane extractable products arising from incubation 54. b) Mass spectrum of the peak eluting at 17.00 min (**6m**), obtained from incubation 54, which remains unidentified.



Scheme 89. In red, proposed interception of the reaction mechanism catalysed by DCS- W279Y upon incubation with **6**. In black, proposed mechanisms for the conversion of **6** to **6a** and **6c** by DCS- W279Y.

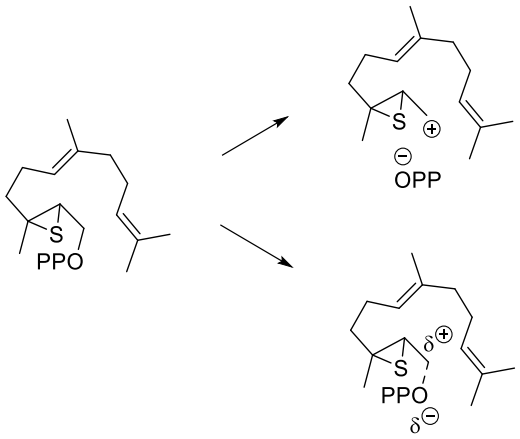
In summary, the 1,10- + 1,6-sesquiterpene synthases behave promiscuously upon incubation with **6** (incubations 51-54). This is in contrast to the incubation of **6** with 1,10-sesquiterpene synthases (incubations 46-49), but similar to the incubation of **6** with exclusive 1,6-sesquiterpene synthase (incubation 50).

PRAS produces **6a**, and other products postulated to be germacrene A-derivates, but there is no evidence for the production of byciclic structures arising from the subsequent 1,6-cyclisation that occurs with FDP as substrate. SdS produces a mixture of at least 5 7H-sesquiterpenes. The presence of 7H-selinadiene analogues has not been ruled out, however, the presence of **6a** implies a deficient/ alternate conformation of the substrate and/or intermediates during catalysis in SdS that lead to aberrant deprotonation. As for EpicS and DCS-W279Y, the 1,6-cyclisation might be prevented since there can no longer be a tertiary cation generated during this process.

4.8. Substrate behaviour of 2,3-thiirane-FDP (7)

Previously, 2,3-epoxy-FDP was incubated with sesquiterpene synthases in our laboratories.[213] This compound was not turned over by GAS, GDS, Gd4olS, ADS, PRAS and DCS-WT. For these enzymes, it was proposed that stabilisation of the 2,3-epoxyfarnesyl cation is not possible, as the formation of the primary carbocation is highly disfavoured. For PRAS, the 1,10- cyclisation is assumed to be concerted, so this epoxide should not influence the reaction. Yet, no reaction was observed, indicating that the SN_2 step might be not concerted but the phosphate leaves before the intramolecular nucleophilic attack occurs, which would be prevented by the missing stabilisation of the epoxide (Follow in Scheme 90, but with 2,3- thiirane as model).
Here, the incubations of 2,3-thiirane-FDP (**7**) with the 9 sesquiterpene synthases under study were tested. The C2 - C3 double bond of FDP (**1**) is not present in **7**, and this could inhibit the initial ionisation happening in these enzymes. However, sulphur is less electronegative than oxygen (χPauling 2,58 and 3.44, respectively), and so less destabilising to an adjacent positive charge.

The results obtained were difficult to interpret. The incubations of **7** with all sesquiterpene synthases gave the natural occurring products respectively (**1a** - **1j**), which would be expected if FDP (**1**) would have been used. This could only mean that either there are NMR spectroscopy- unobservable impurities in the thiirane product or that there is enzymatic decomposition of the substrate taking place. Due to this, we re-analysed the HR-MS data obtained from **7**, and we found that there is a fragment peak which mass could correspond to the molecular ion of **1**, albeit in very minor quantities. These observations would have to be re-examined for a complete interpretation of the results. However, the absence of any sulphur containing structures, suggests that **7** might not be a substrate for GAS, GDS, G4olS, Gd11olS, ADS, PRAS, SDS, EpicS and DCS-W279Y. In analogy with the previous study with 2,3-epoxyde-FDP, stabilisation of the 2,3-thiirane farnesyl cation might not be possible and the formation of the primary carbocation (or transient state in the case of a concerted cyclisation) is highly disfavoured (Scheme 90).



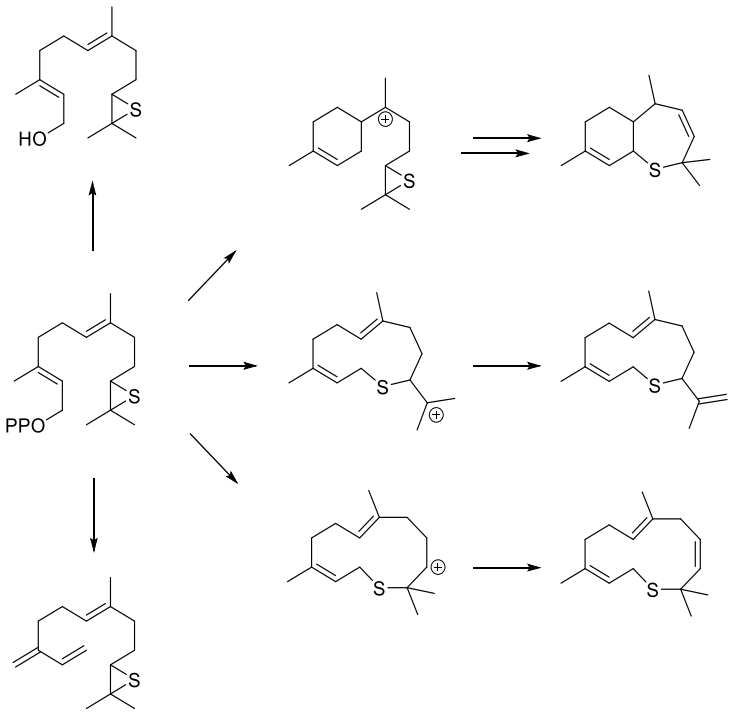
Scheme 90. Carbocation formation is inhibited with the presence of the sulphur at C2-C3 of FDP.

As also stated for the 2,3-epoxyde-FDP mentioned, another explanation might be that 2,3-thiirane FDP (**7**) has change the configuration at C2 and C3. While in FDP, C1 to C4 and C14 are all coplanar due to the sp^2 nature of C2 and C3, in **7**, C2 and C3 are sp^3 in nature and tetrahedral with the added complication of 2 new added stereocenters. This change in configuration may affect the substrate fitting inside the active site and provide an unproductive alignment between the C10 or C6 and C1 for cyclisation.[213]

4.9. Incubation of 10,11-Thiirane-FDP (8) with the enzymes under study

Previous studies in our group demonstrated that 10,11-epoxy-FDP is a substrate for 1,10-sesquiterpene synthases (Section 1.6.2 and 3.1).[132] In analogy to this, we tested 10,11- thiirane-FDP (**8**) with GAS, GDS, GdolS, G11olS, ADS, AS, SDS, EpicS and DCS-W279Y.

GC-MS analysis of the pentane extracts showed no detectable product formation, which indicates that **8** is not a substrate for these sesquiterpene synthases (Scheme 91). These results are not conclusive because the purification and characterisation of **8** was not fully covered during the current synthetic route, as stated in Section 3.9.2.



Scheme 91. Potential products that might arise from the incubations of **8** with the sesquiterpene synthases under investigation, but these were not observed.

4.10. Summary, preparative-scale incubation of 6 with GDS and Gd11olS and NMR spectroscopy characterisation of (2*E*,6*E*)-7H-germacrene D (6b) and (2*E*,6*E*)-7H-germacradien-11-ol (6d)

The incubation of FDP-analogues **2** - **6** (Figure 130) with the sesquiterpene synthases under investigation has given rise to a series of new sesquiterpene analogues, some of which sadly remain unidentified. Those that have been identified are summarised below.

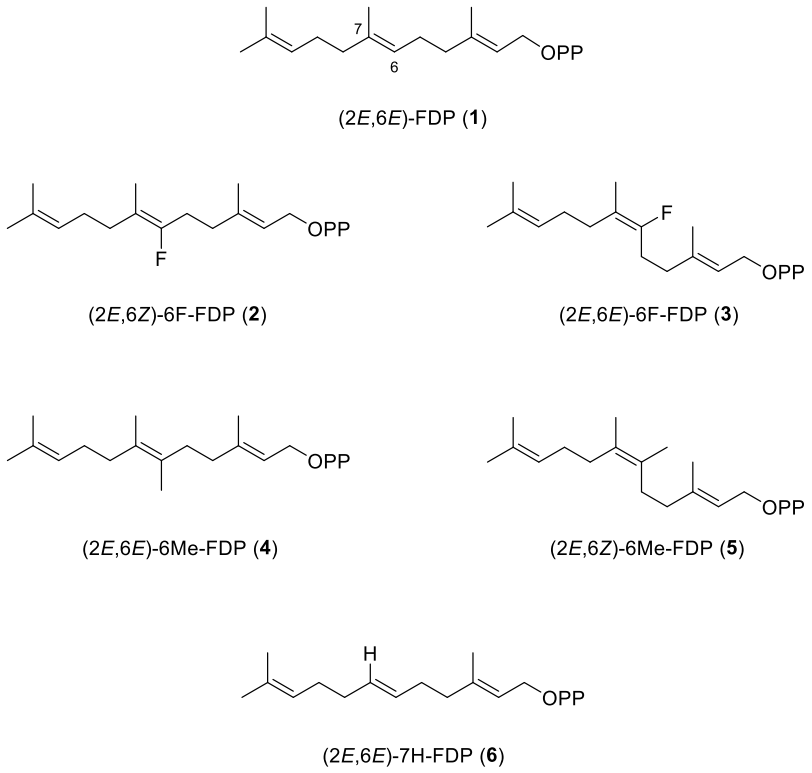


Figure 130. FDP (**1**) and FDP analogues (**2**-**6**).

1,10- Sesquiterpene synthases (GAS, GDS, GdolS and Gd11olS)

Enzyme	Substrate	Products
GAS	**1**	**1a**
	2	**2a**
	3	**3a**
	4	**4a, 4b**
	5	**5a, 5b**
	6	**6a**

Table 14. Observed products upon incubation of GAS with substrates **1**-**6**.

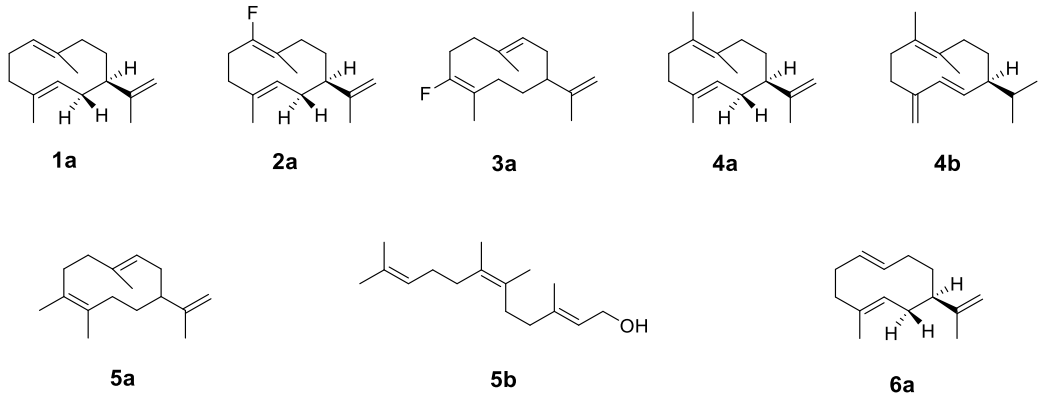


Figure 131. Structures of the identified products upon incubation of GAS with substrates **1**-**6**.

Enzyme	Substrate	Products
GDS	**1**	**1b**
	2	**2b**
	3	**3a**
	4	**4b**
	5	**5a, 5c - 5f (*p.* 175)**
	6	**6b**

Table 15. Observed products upon incubation of GDS with substrates **1**-**6**. Red, unidentified compounds.

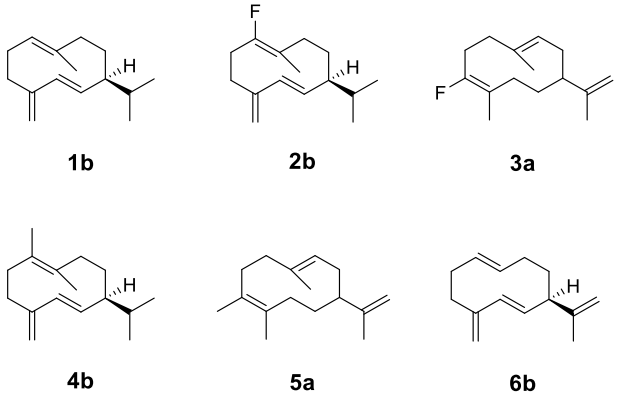


Figure 132. Structures of the identified products upon incubation of GDS with substrates **1**-**6**.

Enzyme	Substrate	Products
GdolS	1	1c
	2	2c, 2a
	3	3a
	4	4c, 4d (*p.* 159)
	5	5a
	6	6c

Table 16. Observed products upon incubation of GdolS with substrates **1**-**6**. Red, unidentified compound.

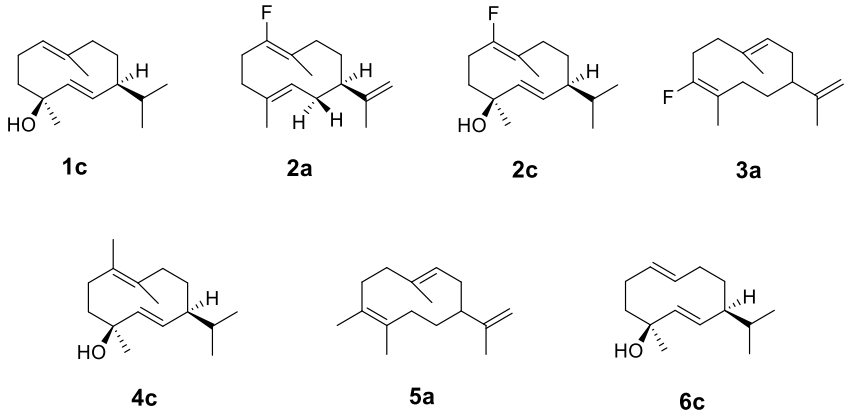


Figure 133. Structures of the identified products upon incubation of GdolS with substrates **1**-**6**.

Enzyme	Substrate	Products
Gd11olS	1	1d
	2	2d, 2e (*p.* 136)
	3	3a
	4	4e, 4f (*p.* 161)
	5	5a, 5g (*p.* 176)
	6	6d, 6b (minor)

Table 17. Observed products upon incubation of Gd11olS with substrates **1**-**6**. Red, unidentified compounds.

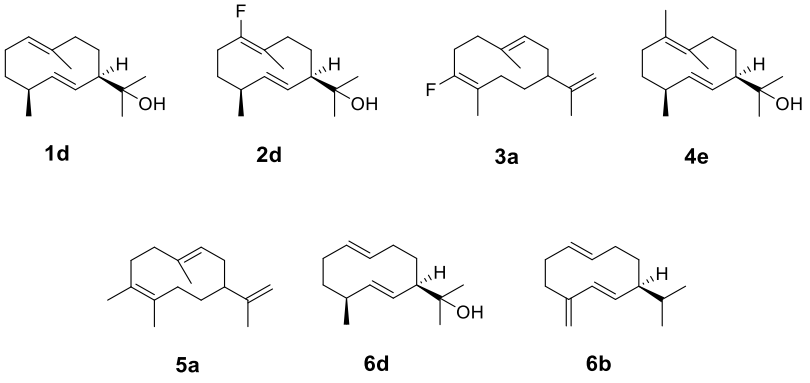
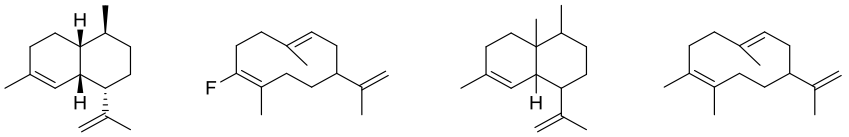


Figure 134. Structures of the identified products upon incubation of Gd11olS with substrates **1**-**6**.

1,6- Sesquiterpene synthase (ADS)

Enzyme	Substrate	Products
ADS	**1**	**1e**
	2	**2f - 2g (*p.* 139)**
	3	**3a, 3b - 3e (*p.* 150)**
	4	**4g**
	5	**5a, 5h - 5k (*p.* 177)**
	6	**6e - 6g (*p.* 189)**

Table 18. Observed products upon incubation of ADS with substrates **1**-**6**. Red, unidentified compounds.



1e **3a** **4g** **5a**

Figure 135. Structures of the identified products upon incubation of ADS with substrates **1**-**6**. Red, unidentified compounds.

1,10- + 1,6- Sesquiterpene synthases (PRAS, SdS, EpicS and DCS-W279Y)

Enzyme	Substrate	Products
PRAS	**1**	**1f, 1a (minor)**
	2	**2a**
	3	**3a**
	4	**4h, 4i - 4j (*p.* 165)**
	5	**5a**
	6	**6a, 6h - 6i (*p.* 191)**

Table 19. Observed products upon incubation of PRAS with substrates **1**-**6**. Red, unidentified compounds.

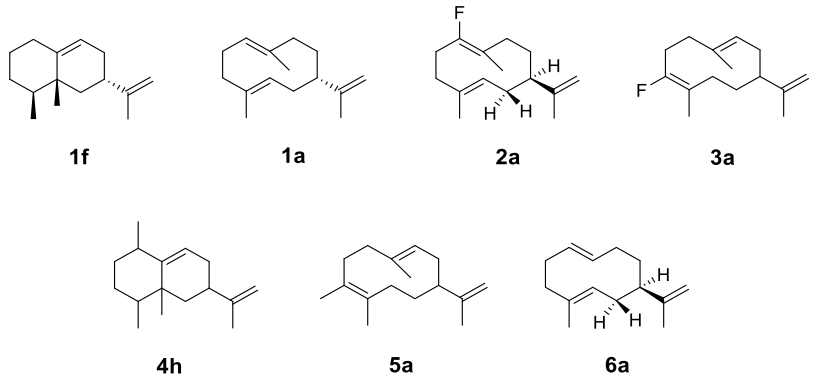


Figure 136. Structures of the observed products upon incubation of PRAS with substrates **1**-**6**.

Enzyme	Substrate	Products
SdS	**1**	**1g, 1h (minor)**
	2	**2h**
	3	**3a, 3f**
	4	**4k - 4l (*p.* 167)**
	5	**5a, 5l**
	6	**6a, 6e, 6h, 6j - 6k (*p.* 193)**

Table 20. Observed products upon incubation of SdS with substrates **1**-**6**. Red, unidentified compounds.

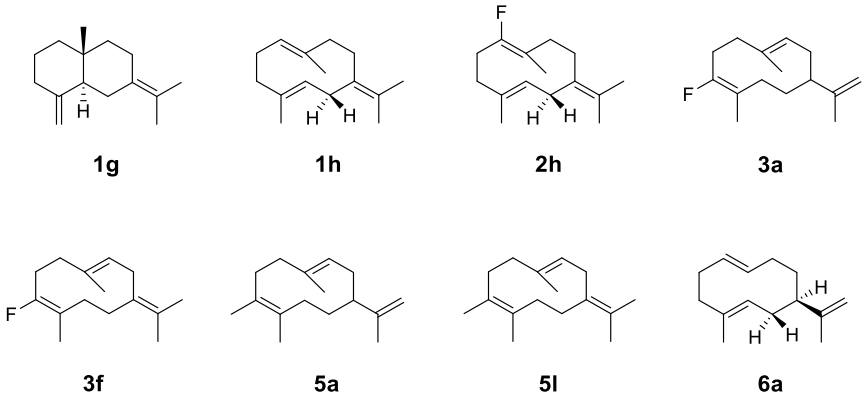


Figure 137. Structures of the observed products upon incubation of SdS with substrates **1**-**6**.

Enzyme	Substrate	Products
EpicS	**1**	**1i, 1c (minor)**
	2	**2c, 2g (*p.* 143)**
	3	**3a**
	4	**4c, 4o, 4d, 4m - 4n (*p.* 168)**
	5	**5a, 5b**
	6	**6a, 6c, 6l**

Table 19. Observed products upon incubation of EpicS with substrates **1-6**. Red, unidentified compounds.

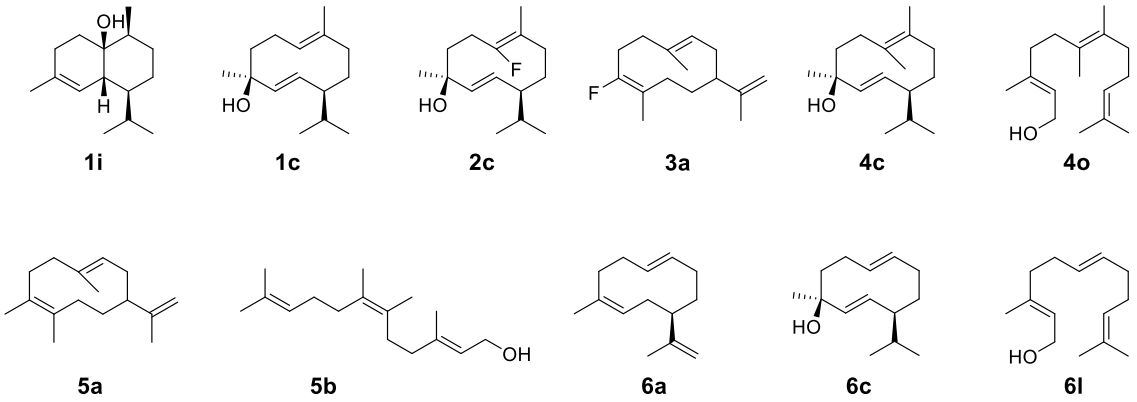


Figure 138. Structures of the identified products upon incubation of EpicS with substrates **1-6**.

Enzyme	Substrate	Products
DCS-W279Y	**1**	**1j, 1c (minor)**
	2	**2c, 2i (*p.* 144)**
	3	**3a**
	4	**4p - 4q (*p.* 169)**
	5	**5a**
	6	**6a, 6c, 6m (*p.* 197)**

Table 21. Observed products upon incubation of DCS-W279Y with substrates **1-6**. Red, unidentified compounds.

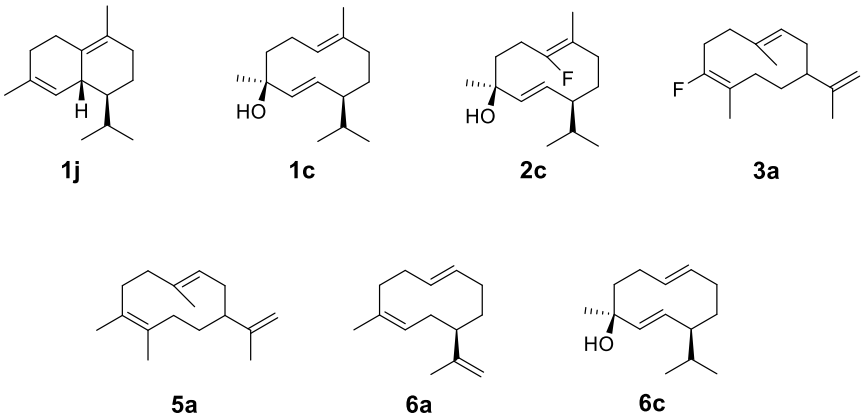


Figure 139. Structures of the observed products upon incubation of DCS- W279Y with substrates **1-6**.

The limited time available for this work prevented the production of all postulated compounds in quantities sufficient to be characterised by NMR spectroscopy. However, from the series of compounds generated, it was decided to produce milligram quantities of a couple of representative compounds for full characterisation, namely, (2*E*,6*E*)-7H-germacrene D (**6b**) and (2*E*,6*E*)-7H-germacradien-11-ol (**6d**).

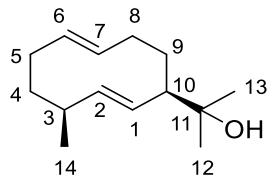
These compounds were chosen for several reasons. (-)-germacrene D is an olfactory ligand acting as insect repellent, however, the analogue 14,15-dimethyl-germacrene D acts as potent attractant (Section 1.6.1).[130] Changing the hydrocarbon composition of these compounds might generate further bioactive germacrene D analogues. On the other hand, (+)-germacradien-11-ol is an intermediate in the enzymatic synthesis of geosmin, which is well known for its unpleasant flavour and object of many studies,[175,273,274,288–294] also it is an alcohol compound, which might result more stable than pure hydrocarbon structures.

Characterisation of these compounds by NMR spectroscopy also serves as an experimental tool to fully validate the earlier tentative identification of 7H- germacrene compounds through analysis of their mass spectra and through prediction based upon the proposed catalytic mechanisms of the enzymes.

Preparative incubations of **6** with Gd11olS and GDS were used for production of milligram quantities of novel products **6d** and **6b**, respectively, as described in materials and methods (Section 7.1.18). The reactions were left a total time of 48 hours, and their product generation over this extended period was monitored by GC, confirming the production of the desired products.

After this time, the products were extracted and evaporated as stated in materials and methods (Section 7.1.18).

In the case of **6d**, minor products formed removed using a short column containing deactivated silica. The isolated compound was identified as (2*E*,6*E*)-7H-germacradien-11-ol (**6d**), by NMR spectroscopy.



6d

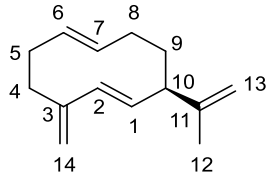
Figure 140. Structure of compound **6d** with numbering system used in assignment

Appendix - ^{1}H NMR, ^{13}C NMR, DEPT135, HSQC, COSY, NOESY, HMBC spectra and VT NMR EXPERIMENTS.

	Compound 6b
^{1}H NMR (500 MHz, $CDCl_3$)	δ 5.55 (dd, *J* = 16.1, 3.4 Hz, 1H, **H2**), 5.07 (dddd, *J* = 15.5, 10.7, 3.2, 1.6 Hz, 1H, **H6**), 4.96 (dddd, *J* = 15.4, 10.7, 3.1, 1.5 Hz, 1H, **H7**), 4.80 (ddd, *J* = 16.1, 9.8, 2.0 Hz, 1H, **H1**), 2.53 (m, 1H, **H3**), 2.48 – 2.37 (m, 2H, **H5b** & **H8b**), 2.36 – 2.29 (m, 1H, **H10**), 2.24 – 2.15 (m, 1H, **H8a**), 2.12 – 2.06 (m, 1H, **H5a**), 1.74 – 1.67 (m, 1H, **H9b**), 1.63 (ddd, *J* = 9.0, 6.1, 2.4 Hz, 1H, **H4b**), 1.56 – 1.48 (m, 1H, **H4a**), 1.24 (d, *J* = 6.9 Hz, 3H, **H14**), 1.19 – 1.11 (m, 1H, **H9a**), 1.17 (s, 3H, **H13**), 1.07 (s, 3H, **H12**).
^{13}C NMR (126 MHz, $CDCl_3$)	δ 144.6 (s, **C2**), 134.2 (s, **C6**), 134.0 (s, **C7**), 128.7 (s, **C1**), 71.9 (s, **C11**), 58.7 (s, **C10**), 34.44 (s, **C3**), 34.42 (s, **C8**), 31.6 (s, **C4**), 28.9 (s, **C5**), 27.0 (s, **C12**), 26.5 (s, **C13**), 26.3 (s, **C9**), 15.3 (s, **C14**).

Table 22. Complete assignment of compound **6b**. Chemical shifts are reported relative to TMS in ppm

The NMR spectroscopy data for compound **6b** collected from the crude sample confirmed the presence of characteristic peaks for (2*E*, 6*E*)-7H-germacrene-D (**6b**). However, this compound proved very unstable upon attempted isolation, and all the efforts to purify it, or even to concentrate it from pentane failed.



6d

Figure 141. Structure of compound **6b**.

Appendix- ^{1}H NMR spectrum.

CHAPTER 5

EXPLORING A DIRECTED EVOLUTION APPROACH TO IMPROVE THE CATALYTIC ACTIVITY OF SESQUITERPENE SYNTHASES

Chapter 5. Exploring a directed evolution approach to improve the catalytic activity of sesquiterpene synthases

5.1. Preface and general strategy

Academic and industrial interest in the unique catalytic mechanisms performed by terpene synthases and their usage to produce new natural products is driving the application of new methodologies to engineer new enzymes with better performance. In nature, enzymes are perfected by continuous evolution.[156–158]

One common approach to enhance or alter enzyme reaction pathways is site- directed mutagenesis (SDM, Section 1.5.1 and Chapter 2). Selectively changing single amino acid based on mechanistic proposals and structural knowledge is usually performed in the active site of an enzyme. Yet, a lot of details on the origin of the high activity of enzymes is still lacking, e.g. chain movements, allosteric regulations, etc. Hence, beneficial mutations outside of the active site are missed by site- directed mutagenesis.

Directed evolution methodology, however, represent a closer approach to nature's process, as it allows the production of a library of mutated enzymes without reduced selection bias in a short time, and has been used in the last decades in the area of enzyme engineering, as described in section 1.6.[156–158]

In these laboratories there has historically never been an evolutionary approach to improve sesquiterpene synthase catalytic activity due to the paucity of the screening method. Establishing the methodology to perform directed evolution of sesquiterpene synthases was the objective of the last project performed, and is outlined in this chapter.

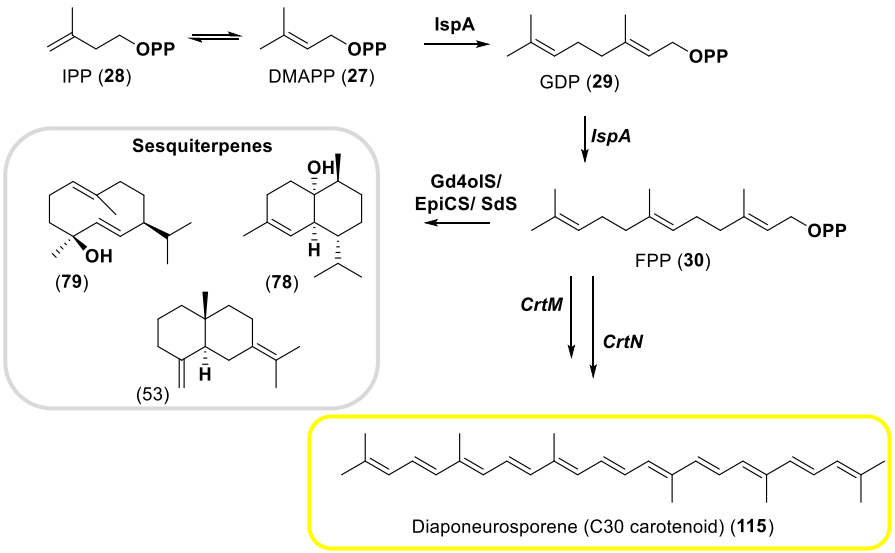
A prerequisite for a successful evolution experiment is an assay to address the requirements for screening of the mutated genes, so that we can select the better genes of our interest. Here, it is a colorimetric screening developed by Umeno *et al.*,[163,295] that relies on the competition between two biological pathways for the same substrate (section 1.6, and Scheme 92).

The principle used for the selection of active enzymes is based on the use of a screening plasmid containing *CrtM* and *CrtN* genes, which drive the production of diponeurosporene (**115**), a yellow carotenoid, from FDP (**30**). Sesquiterpene synthases use FDP (**30**) as substrate, hence, to couple these two biological pathways in the same cells allows for selection of the best sesquiterpene synthases attending on the pigmentation of the cells (Schemes 92 and 93).

To choose an appropriate starting enzyme, which has an amplifiable catalytic signal, is foundational. In this study we used three different sesquiterpene synthase, SdS, EpicS and GdolS (Scheme 92) for several reasons. These enzymes are available in the laboratory and their mechanisms of reaction are well stablished.[85,125,169,170,296,297] Also, they are used as models to extend the available knowledge of the enzymatic contribution of the G1/2- helix break in sesquiterpene synthases, which has been

demonstrated to be of catalytic importance in these enzymes (Section 1.4 and Chapter 2) and might be a hot spot to improve their activity by 'laboratory evolution'.[155] Using the genes of these enzymes as templates, it was aimed to generate a library of mutated genes by using EP-PCR.

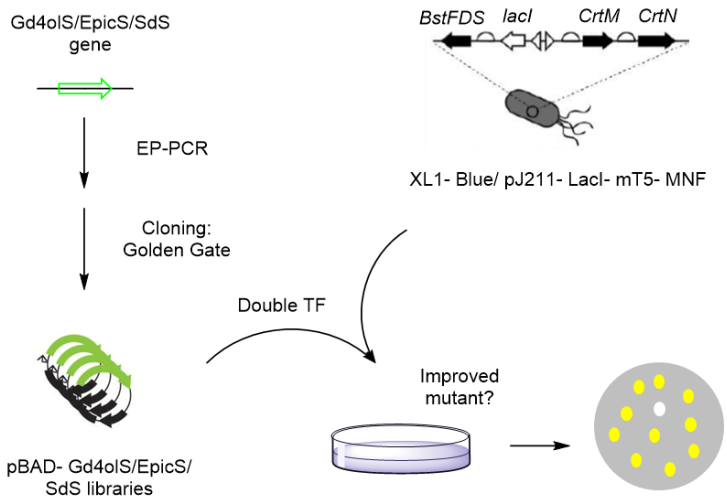
EP-PCR can be applied in the whole gene of these sesquiterpene synthases or in selected parts of the gene (e.g. the G1/2 helices) and coupled with the screening method to select better enzymatic variants (Scheme 93).



Scheme 92. Carotenoid and sesquiterpene producing pathways compete for FDP (**30**). Adapted from [163,295].

The work covered here comprises several steps of this project (Scheme 93):

First, the preparation of XL1-blue cells containing the **screening plasmid** (Plasmid 1), which contains the required genes for the production of diaponeurosporene (**115**, Scheme 92). Then, the construction of the plasmids containing sesquiterpene synthases gene sequences (plasmids 2-5). These plasmids contain a controllable **expression system** for the production of sesquiterpene synthases which is compatible with the screening system and allows an optimal substrate competition (pBAD). Also, the aim was to perform control tests of the screening method. This was to be done by co-expression of active or inactive sesquiterpene synthases with the carotenoid producing enzymes and visualising the cells varying in the levels of diaponeurosporene (control attenuation/not attenuation). Lastly, it was planned to run EP-PCR to generate a library of mutated genes and to subclone them in the control plasmids by Golden Gate methodology.[160,298,299]



Scheme 93. Cartoon depiction of the planned approach for directed evolution of sesquiterpene synthases. Adapted from [163,295].

If the prior objectives were to be reached, the following objectives would follow up:

1) To screen the library of mutants, and to visually select the best candidates.

2) *Second screening*. Repeat the screening of the candidates but with the incorporation of fluorometric techniques to measure more accurately the intensity of pigmentation at this second stage.[163,295]

3) To sequence and express the successful candidates.

4) To measure the catalytic parameters and compare with the natural ones.

5.2. Preparation of XL1-blue cells containing the screening plasmid and construction of sesquiterpene synthase-containing plasmids

5.2.1 Screening plasmid (plasmid 1)

The screening plasmid involved in this project, pJ211-lacI-mP$_{T5}$-MNF (Plasmid 1, Figure 142), was a generous gift from Prof. Daisuke Umeno. It contains the *CrtM* and the *CrtN* genes from *Staphylococcus aureus* which convert FDP (**30**) into a yellow carotenoid compound, diaponeurosporene (**115**). It also contains a FDP synthase from *Geobacillus Stearothermophilus* (BstFDPS) which is thought to increase the conversion of GPP (**29**) to FDP (**30**), and therefore can increase the amount of FDP (**30**) available in the cell. The origin of replication is p15A and it has a kanamycin resistance marker. [163,295]

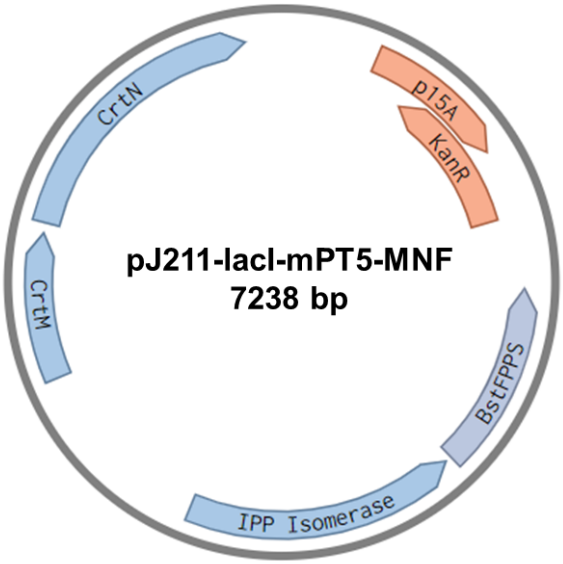


Figure 142. Map of plasmid 1.

XL1-blue cells were transformed with plasmid 1 following the protocol described in materials and methods, Section 7.1.5. These cells will be used for double transformations with plasmids 2-5 (below).

5.2.2. Generation of sesquiterpene synthase-containing plasmids (plasmids 2-5)

The plasmids for sesquiterpene synthases expressions were designed according to the experimental section. These plasmids would serve for screening controls (check active and inactive sesquiterpene synthases) and as the templates for EP-PCR and the subcloning of the randomised gene fragments. Four different sesquiterpene producing plasmids were created:

- Plasmid 2 contains the pBAD vector (acceptor vector 1) and GdolS-WT (insert 1).
- Plasmid 3 is composed of pBAD vector (acceptor vector 2) and EpicS-WT (insert 2).
- Plasmid 4 contains the pBAD vector (acceptor vector 3) and SdS-WT (insert 3).
- Plasmid 5 is composed of pBAD vector (acceptor vector 1) and GdolS-A176F (insert 4).

Plasmids 2-4 are positive controls, because these enzymes are active to catalyse FDP cyclisation. This should reduce the colour produced by the competitive carotenoid production. In contrast, plasmid 5 serves as negative control, as it is an inactive enzyme and does not compete for FDP consumption, thus it should not attenuate the pigmentation.

These plasmids were created using the Golden Gate cloning technique (conditions and reactants used are detailed in Section 7.1.10.), as illustrated in Figure 144.

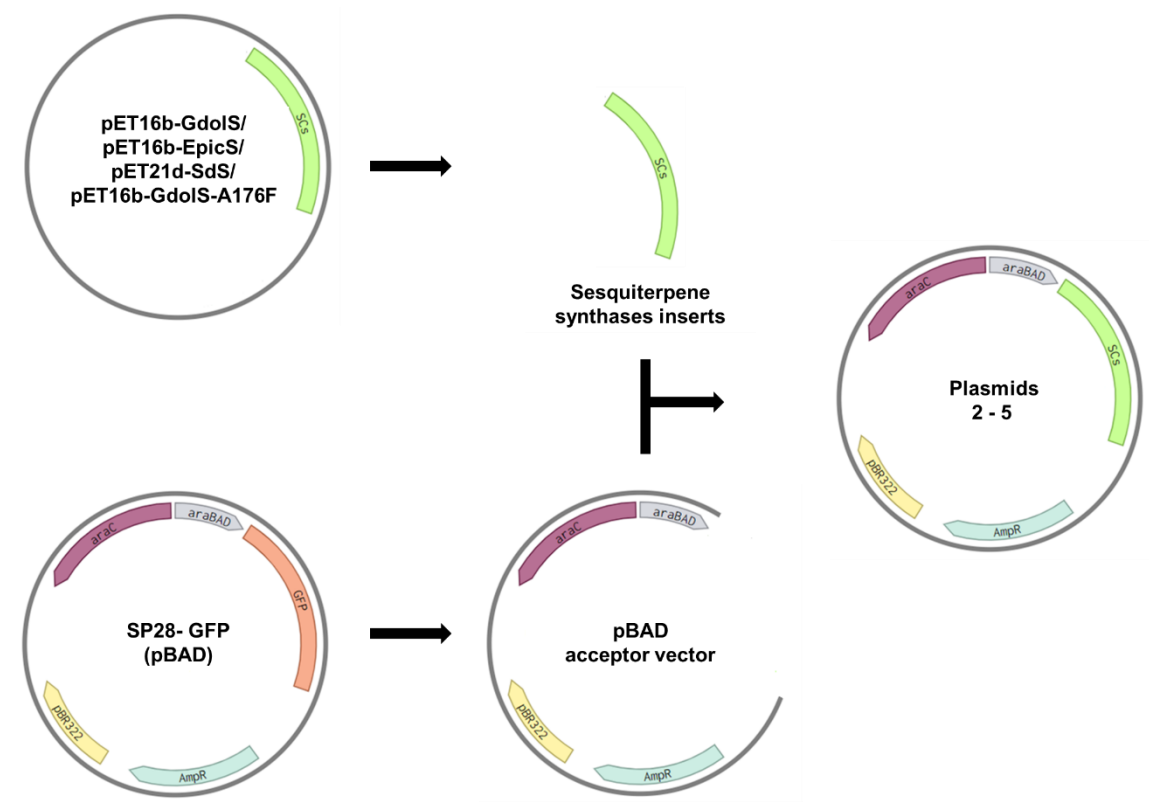


Figure 143. Cloning strategy for the construction of plasmids 2-5.

The primers used to amplify the inserts and cloning vectors were designed to place the Type IIS recognition site distal to the cleavage site, such that the Type IIS REase used can break apart the recognition sequence from the Golden Gate assembly.[298]

We found that BbSI Type IIS REase recognition site was not present in the original plasmids. Therefore, we introduced BbSI Type IIS in the primers used for the amplifications of the acceptor vectors and inserts and served as complementary overhangs for ligation (Figure 144).

After careful design of the primers, the desired acceptor vectors were amplified. These vectors were amplified from the SP28-pBAD plasmid, which was a generous gift from Sanjay Patel (Cardiff University, PhD student) (Figures 143 and 145). SP28-pBAD originally contained the GFP (green fluorescent protein) gene sequence, which was the replaced fragment in the construction of the new plasmids (2-5). Additionally, the four genes encoding for the sesquiterpene synthases (GdolS, EpicS, SdS and GdolS-A176F) were amplified from their original plasmids (Figure 143 and 145). Then, both fragments (acceptor vectors and inserts) were ligated according to the Golden Gate protocol.[298]

Figure 144. Designed primers for the amplification of inserts 1-4 (left) and acceptor vectors 1-3 (right).

The generated DNA fragments were analysed by agarose gel electrophoresis, and the sizes were matching the expected ones for acceptor vectors (lanes 1-4, ~4000 bp) and inserts 1-4 (lanes 5-13, ~1000 bp), as shown in Figure 145. After ligations, we run a control digestion using the restriction enzyme BamHI, which is only present in the acceptor vector (Figure 145). All the fragments (one per plasmid) showed the expected size for plasmids 2-5 (~5000 bp).

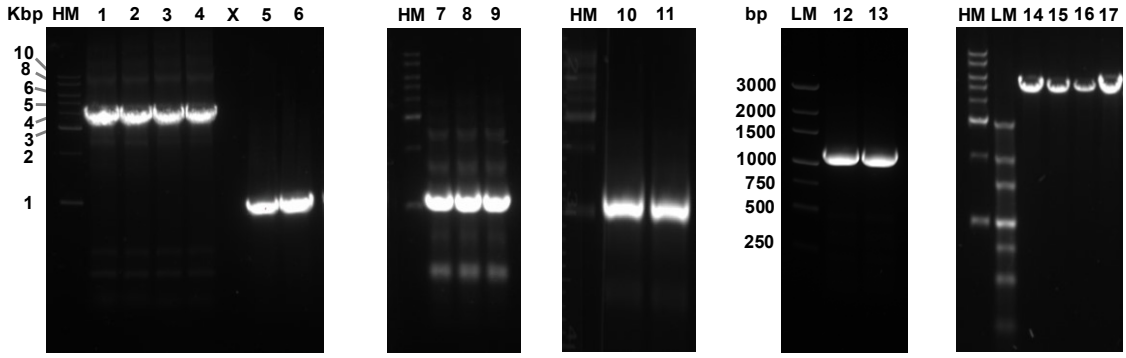


Figure 145. Agarose gels. HM: High marker, LM: Low marker; 1-4. Amplification of acceptor vector (one for each cloning). 5 and 6: Amplification insert 1; 7 – 9: Amplification insert 2; 10 and 11: insert 4; 12 and 13: insert 3; 14 - 17: restriction digest fragments of plasmids 2-5 with BamHI.

5. 3. Screening controls

The work presented here illustrates the controls designed and performed to check that the screening system used is able to differentiate between active and inactive sesquiterpene synthases. This method is performed by co-expressing sesquiterpene synthase genes in a carotenoid-producing cell.[163] Because both pathways compete for the same substrate (FDP, **30**), when the sesquiterpene synthases are sufficiently active then the quantity of carotenoid produced in the cell is lower. Plasmids 2-4 are hypothesised to compete actively and attenuate the colour of the colonies, whereas plasmid 5 is thought to be a negative control and should not reduce pigmentation intensity.

The sesquiterpene used in this study are built in a pBAD vector, which has a different inducible promoter to that of the screening plasmid. While the screening plasmid is expressed under the T7 promoter and requires IPTG, the sesquiterpene plasmids contain an L-arabinose-inducible promoter.

5.3.1. L-arabinose concentration (using plasmids 1 and 2)

First, GdolS-WT gene (plasmid 2) was used as a model to set the amount of arabinose used in the method. Thus, plasmid 2 was introduced into chemically competent XL1-blue cells containing plasmid 1, and the transformants were plated onto LB agar containing kanamycin and ampicillin for selection of double transformants and topped with nitrocellulose membrane. This membrane helps to see the colour of the colonies, as well as to transfer then to other media. After 16 hours at 37 ºC we observed no appreciable yellow pigmentation formed. Then, the nitrocellulose membrane containing the double transformed cells were transferred in pieces to 5 different LB agar plates containing a constant concentration of IPTG (100 μM, as previously published[295]) and 5 different concentrations of L-arabinose (Figure 146) and incubated at room temperature in darkness for 48 hours.

1. LB agar, ampicillin, kanamycin and 100 μM IPTG (no arabinose used).
2. LB agar, ampicillin, kanamycin, 100 μM IPTG and 0.1 % arabinose.
3. LB agar, ampicillin, kanamycin, 100 μM IPTG and 0.5 % arabinose.
4. LB agar, ampicillin, kanamycin, 100 μM IPTG and 1 % arabinose.
5. LB agar, ampicillin, kanamycin, 100 μM IPTG and 2 % arabinose.

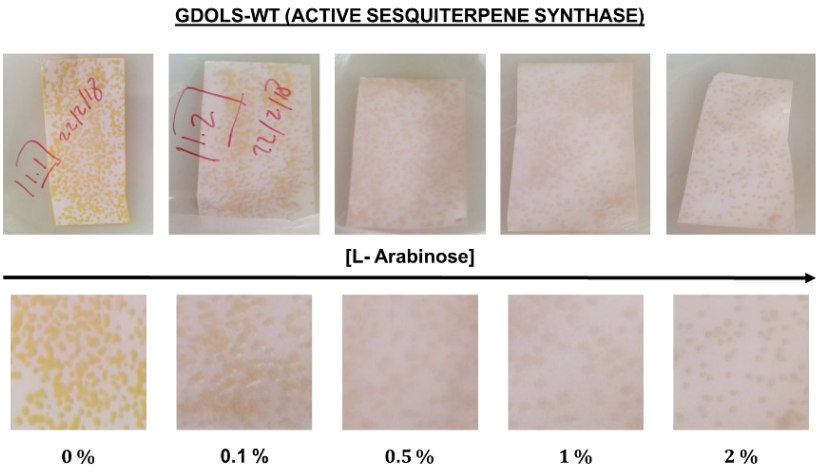


Figure 146. Screening control. From left to right, increasing amounts of L-arabinose in the media.

The results clearly showed a reduction in pigmentation upon the use of higher concentrations of L-arabinose. This is indicative that GdolS is being produced in higher yields. Regarding the visualised colour intensities in these cells, it was decided to set 1 % of arabinose in the media for later experiments. This concentration attenuates the pigmentation of the cell sufficiently to be analysed and differentiate them from colourful (no sesquiterpene production) colonies.

5.3.2. Activity vs inactivity (using plasmids 1 and 4, and 1 and 5)

Once the concentration of the L-arabinose inducer was determined using GdolS-WT as model, the ability of the system to differentiate active enzymes from inactive ones was controlled. For this part, we used the previously constructed sesquiterpene-producing plasmids 4 and 5. SdS (plasmid 4) is the sesquiterpene synthase with the lowest catalytic activity compared with Gd4olS and EpicS. GdolS-A176F (plasmid 5) is an inactive mutant. Thus, using these two variants we ensure to control if the experiment can differentiate activity from inactivity.

Plasmids 4 and 5 were independently introduced into XL1-blue cells containing the screening plasmid 1. As before, the transformants were spread onto LB agar (without inducer) topped with nitrocellulose membrane. After 16 hours incubation at 37 ºC, these colonies-containing membranes were transferred to LB agar containing 100 mM of IPTG and 0 % or 1 % of L-arabinose and left for further 48 hours at room temperature in the dark.

As illustrated in Figure 147, the pigmentation of the colonies was attenuated when plasmid 4 was used, however, the incorporation of plasmid 5 did not reduce the colour intensity. These results probed the ability of the screening system to classify terpene synthases according to their catalytic activity.

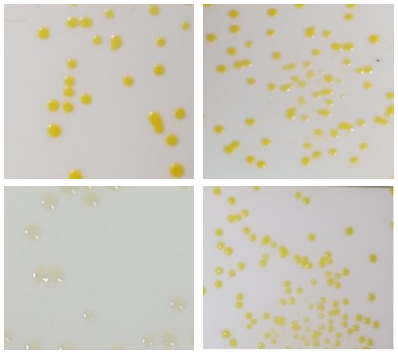


Figure 147. Screening control. Left, SdS using 0 % (top) an 1 % (bottom) of L-arabinose in the media. Right, Gd4olS-A176F. using 0 % (top) an 1 % (bottom) of L-arabinose in the media.

5.4. Random mutagenesis. Error-prone PCR (EP-PCR)

This part comprises a preliminary optimisation of the PCR conditions using *Taq* polymersase for the amplification of the whole gene encoding for SdS (used as a model, ~1000 bp). From these conditions, a pilot EP-PCR-based amplification was run with the genes encoding for the G1/2 helix present in GdolS, EpicS and SdS (~ 100 bp, from plasmids 2-4). Lastly, it was generated a library of mutated SdS-G1/2 helix encoding genes. This library was stored for future analysis and can be used for subcloning into the appropriate vector and be screened with the previously described controlled system.
The EP-PCR conditions used for this study were optimised following established instructions, with slight variations.[160,299]

5.4.1. Taq polymerase to work (amplification of the whole gene encoding SdS, plasmid 4)

Using the pBAD-SdS plasmid as template and the corresponding primers (See Figure 145 above), the ability of *Taq* polymerase to amplify the whole gene encoding SdS (~1000 bp) was tested under different reaction compositions and conditions, which are detailed in Tables 23 and 24.
Figure 148 indicates the size of the DNA products formed after these reactions. From the initial conditions (which were unproductive), three different extending temperatures were screenned (62 ºC lanes 1 and 2, 65 ºC lanes 3 and 4 and 68 ºC lanes 5 and 6 (Figure 148). Expected size amplicon (~ 1000 bp) was not observed, but undesired fragments with smaller size were detected. To optimise the amplification, three different primer concentrations were tested (using 0.5 μM in lanes 7 and 10, 1 μM in lanes 8 and 11 and 2 μM in lanes 9 and 12) with a fixed extention temperature at 68 ºC. In these reactions *Taq* polymerase was added using the hot start procedure in lanes 7 to 9, or adding the polymerase when reaching the annealing temperature in lanes 10 to 12, also adding DMSO to the mixture. It was observed the formation of desired DNA fragment (~ 1000 bp) in all these reactions, however, the addition of the polymerase at the annealing temperature resulted in more product formation. Furthemore, the use of 2 μM of primers achieved the best amplifications observed, albeit formation of other undesired fragments.

Then, from these conditions, the influence of two different concentrations of template were checked. Lanes 13 to 16 represent the use of 20 ng of template using 65 ºC (lanes 13 and 14) and 68 ºC (lanes 15 and 16) for annealing, respectively. Both temperatures showed similar product formation, in accordance with the observation from lanes 3 to 6, where there were similar results for both annealing temperatures. Lanes 17 and 18 respresent the use of 200 ng of template at 68 ºC. These conditions provided higher amplification of the desired sequence (~ 1000 bp), albeit the quantity of the undesired products were similar to the others. At this point, we though that this condition (experiment 18) was optimizated enough to continue with the next step of this the study, as we could obtain reasonable

amplification levels of the SdS gene and the undesired products could be easily separated with gel purification kits.

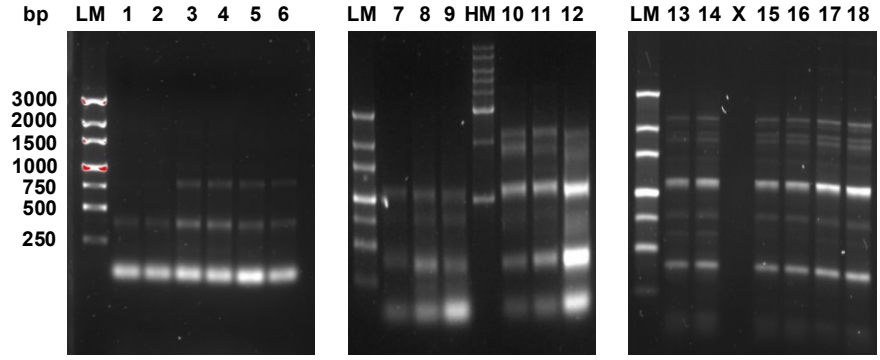


Figure 148. Agarose gels showing DNA size of the reaction products after PCR reactions. The reaction conditions used are shown in Tables 23 and 24.

Experiment	buffer	[$MgCl_2$] (mM)	[dNTPs] (µM each)	[primers] (µM each)	[Template] ng/ 50 µL	*Taq* pol. (U/µL)	DMSO
Start	1X	1.5	0.2	0.5	20	0.05	-
1	1X	1.5	0.2	0.5	20	0.05	-
2	1X	1.5	0.2	0.5	20	0.05	-
3	1X	1.5	0.2	0.5	20	0.05	-
4	1X	1.5	0.2	0.5	20	0.05	-
5	1X	1.5	0.2	0.5	20	0.05	-
6	1X	1.5	0.2	0.5	20	0.05	-
7	1X	1.5	0.2	0.5	20	0.05	4 %
8	1X	1.5	0.2	1	20	0.05	4 %
9	1X	1.5	0.2	2	20	0.05	4 %
10	1X	1.5	0.2	0.5	20	0.05	4 %
11	1X	1.5	0.2	1	20	0.05	4 %
12	1X	1.5	0.2	2	20	0.05	4 %
13	1X	1.5	0.2	2	20	0.05	4 %
14	1X	1.5	0.2	2	20	0.05	4 %
15	1X	1.5	0.2	2	20	0.05	4 %
16	1X	1.5	0.2	2	20	0.05	4 %
17	1X	1.5	0.2	2	200	0.05	4 %
18	1X	1.5	0.2	2	200	0.05	4 %

Table 23. Chemical components used for PCR experiments 1- 18

Experiment	Initial denaturation (ºC/ min)	Denaturation (ºC/ sec)	Annealing (ºC/ sec)	Extension (ºC/sec)	Final extension (ºC/ min)	Hot start
Start	94/ 3	94/ 45	55/ 30	62/ 99	62/ 10	YES
1	94/ 3	94/ 45	55/ 30	62/ 99	62/ 10	YES
2	94/ 3	94/ 45	55/ 30	62/ 99	62/ 10	YES
3	94/ 3	94/ 45	55/ 30	65/ 99	65/ 10	YES
4	94/ 3	94/ 45	55/ 30	65/ 99	65/ 10	YES
5	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	YES
6	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	YES
7	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	YES
8	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	YES
9	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	YES
10	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
11	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
12	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
13	94/ 3	94/ 45	55/ 30	65/ 99	65/ 10	NO
14	94/ 3	94/ 45	55/ 30	65/ 99	65/ 10	NO
15	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
16	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
17	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO
18	94/ 3	94/ 45	55/ 30	68/ 99	68/ 10	NO

Table 24. Conditions used for the PCR experiments 1 – 18.

5.4.2. EP-PCR conditions (amplify the gene encoding G1/2 helixes, plasmids 2-4)

After briefly optimising the PCR conditions using the whole gene encoding SdS (with *Taq* polymerase), the EP-PCR variations were applied for the amplification of a shorter secuence of the sesquiterpene genes. As mentioned earlier, it was decided to randomise the sequence of the sesquiterpene synthase genes encoding for the G1/2 helices of GdolS, EpicS and SdS. This was easily applicable, because plasmids 2 to 4 were already constructed and these were used as the template to amplify G1/2 helices. These also serve as the acceptor vectors for ligation of the library of mutated fragments obtained after EP-PCR amplification (Figure 149).

Thus, primers for amplification of the G1/2 helix secuences of GdolS, EpicS and SdS were designed according to Golden Gate technique requirements (previously shown in Figure 144 for the amplification of the entire sequences of these synthases and adapted to the new regions). After this, EP-PCR conditions were used to do pilot amplifications (only one round of cycles) of these sequences (experiment 19 for SdS, 20 for Gd4olS and 21 for EpicS in Tables 25 and 26). The conditions used for these pilot reactions were similar to the ones optimised before (Tables 23 and 24), however, this time the polymerase fidelity was artificially descreased (to increase the rate of mutations) by adding Mn^{2+} to the reaction, increasing the amount of Mg^{2+} and using unequal concentration of dNTPs (Table 25).[160,299]

Exp.	buffer	[$MnCl_2$] mM	[$MgCl_2$] mM	[dATP], [dGTP] µM each	[dCTP], [dTTP] µM each	[primers] µM each	[Template] ng/ 50 µL	*Taq* pol. U/µL	DMSO
19	1X	0.5	7	0.2	1	2	200	0.05	4 %
20	1X	0.5	7	0.2	1	2	200	0.05	4 %
21	1X	0.5	7	0.2	1	2	200	0.05	4 %

Table 25. Chemical components used for the EP-PCR experiments 19 – 21.

Exp.	I.denaturation ºC/ min	Denaturation ºC/ sec	Annealing ºC/ sec	Extension ºC/sec	F.extension ºC/ min	Hot start	-	-	-
19	94/ 3	94/ 45	55/ 30	68/ 20	68/ 10	NO	-	-	-
20	94/ 3	94/ 45	55/ 30	68/ 20	68/ 10	NO	-	-	-
21	94/ 3	94/ 45	55/ 30	68/ 20	68/ 10	NO	-	-	-

Table 26. Conditions used for the EP-PCR experiments 19 – 21.

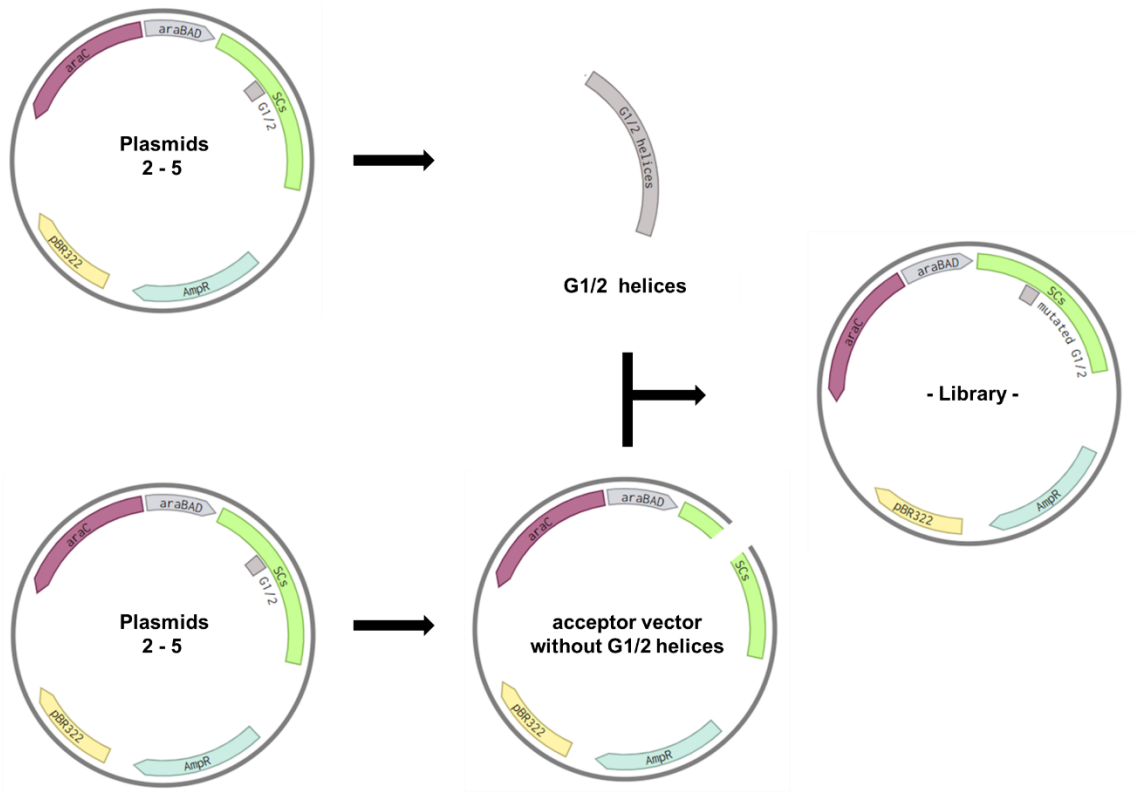


Figure 149. Cloning strategy for the generation of libraries of new genes encoding for sesquiterpene synthases with randomization of the sequences encoding for their G1/2 helices.

As shown in figure 150, these EP-PCR reactions (experiments 19-21) yielded the desired fragments in good levels of production and selectively (~ 100 bp). These fragments were stored at -20 ºC for future analysis.

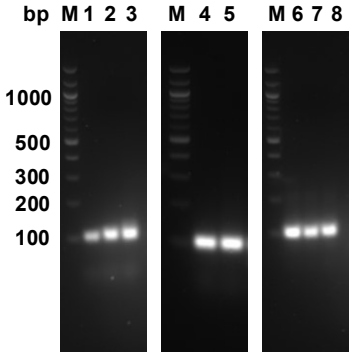


Figure 150. Agarose gels. Observed amplification products after reactions 19 (lanes 1-3), 20 (lanes 4-5) and-21 (lanes 6-8).

5.4.3. Library of mutated genes (multiple rounds of amplification of the gene encoding G1/2 helix of SdS, plasmid 4)

Finally, the aim was to obtain a library of mutated genes under 'controlled conditions'[160,299] (Figure 151) to subclone them in the appropiate acceptor vector (See Figure 149 above) and to lastly couple these new plasmids with the previously controlled screening system to select the best mutated genes. The library of mutated genes created here was made by using SdS encoding gene as a model (plasmid 4) and the primers designed to amplify its secuence encoding for G1/2 helix break. For this, the conditions used before in reactions 19-21 were used following established procedures.[160,299] Thus, after every 4 PCR-cycles 1/10 of the product mixture was taken and sequentially added to the next round until 32 cycles were run. After every 4 cyles a sample of the reaction mixture was saved to be analysed by agarose gel. As shown in Figure 151, the EP-PCR conditions used selectively amplified the region of interest (~100 bp, G1/2 helix of SdS). Also, it was observed an increment in the amount of product formation along the 8 rounds, which is in well agreement with the literature.[160] Following this, the eight fractions were pooled together, which is the library of mutated genes to be screened. Because it was observed undesired fragments at lower size (< 100 bp, Figure 151), the PCR products were purified by agarose gel extraction standarised procedure (Figure 151) and the DNA library was stored at -20 ºC.

Note that every 4 cyles it was saved an aliquot of the sample at -20 ºC for future analysis of the mutations introduced each round (this is not covered here).

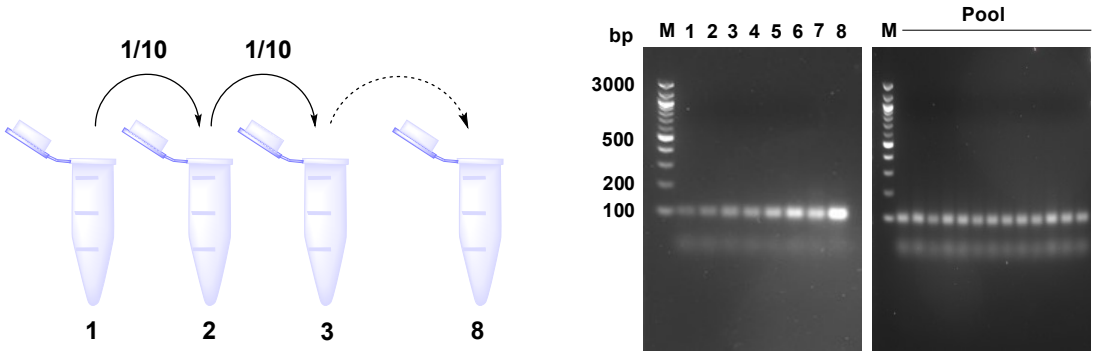


Figure 151. Left, representation of the serial dilutions used for the generation of a library of mutants (8 rounds). Right, agarose gels from the reactions 1 to 8, and the pool of fractions from the eight rounds (library).

The next step, which is not covered here, is to amplify the vector acceptor and used it for the cloning of the library of mutants by Golden Gate (as shown in Figure 150).

5.5. Summary and future work

The directed evolution approach developed here comprises two parts: the introduction of random mutations in selected sesquiterpene synthases encoding genes (or in the genes encoding for their G1/2 helices) and the screening of the generated library of mutants to select the more active enzymes. First, it was prepared XL1-blue cells containing the screening plasmid (plasmid 1). Also, 4 different plasmids have been constructed by using the golden gate technique, each containing a different sesquiterpene gene (GdolS, EpicS, SdS and GdolS-A176M; plasmids 2-5). These plasmids have been designed to co-express with plasmid 1 in XL1-blue cells, and thus they contain an L-arabinose inducible promoter (pBAD vector). Following this, the co-expression of GdolS (plasmid 2) with the carotenoid producing enzymes has been established, using a constant amount of IPTG but varying concentrations of L-arabinose. This control worked well, decreasing the attenuation of cell colour when increasing L-arabinose concentration. From this, 1 % was considered a suitable concentration to perform the assay. As a control to check the reliability of the system, plasmids 4 (encoding SdS- active) and 5 (encoding Gd4olS-A176F- inactive) were independently co-expressed with plasmid 1, and resulted successful, showing yellow colonies when the inactive plasmid was used and white colonies when the active plasmid was in use.

To introduce random mutations in the sesquiterpene synthase genes, EP-PCR and the Golden Gate technique has been employed. First, the conditions and the composition of the PCR using *Taq* polymerase were optimised for the amplification of the whole gene of SdS, and then, these conditions were adulterated to artificially introduce mutations in the G1/2 helices of GdolS, EpicS and SdS.

The library of mutated genes (amplified G1/2 helices fragments) contain the appropriate overhangs for their subcloning into acceptor vectors (in which the G1/2 helices fragments has been removed).

This work is yet to be done. In addition to this, other steps are to be reached, which include:

- To screen the library of mutants, and to visually select the best candidates.
- *Second screening*. Repeat the screening of the candidates but with the incorporation of fluorometric techniques to measure more accurately the intensity of pigmentation at this second stage.[163,295]
- To sequence and express the successful candidates.
- To measure the catalytic parameters and compare with the natural ones.

CHAPTER 6

GENERAL CONCLUSIONS AND FUTURE WORK

Chapter 6. General conclusions and future work

This thesis has described complementary studies to examine the usage of sesquiterpene synthases.

Figure 152. Route of investigation.

First, **structure-function studies** have been carried out to investigate the water management mechanisms of two hydroxylating sesquiterpene synthases, (+)-germacradien-4-ol synthase (GdolS) and (+)-epicubenol synthase (EpicS). These enzymes were overproduced in *E.coli*, extracted, purified and characterised. Selected active site amino acids residues were mutated to examine their role in enzyme catalysis.

In Gd4olS, the G1/2 helix break was altered because it can participate in substrate ionisation and has a potential role in water influx into the active site, as shown for other related sesquiterpene synthases. In this region, A176 was shown to be crucial for catalysis in Gd4olS, participating in product selection. The replacement of A176 with larger non-polar amino acids (valine, leucine, isoleucine and methionine) resulted in a switch on the product outcome from hydroxylated to non-hydroxylated products, namely germacrene A. However, substitution of A176 with the smaller amino acid glycine or with polar amino acids of different size (threonine, glutamic acid and glutamine) resulted in the exclusive production of (-)-germacradien-4-ol, the wild type product.

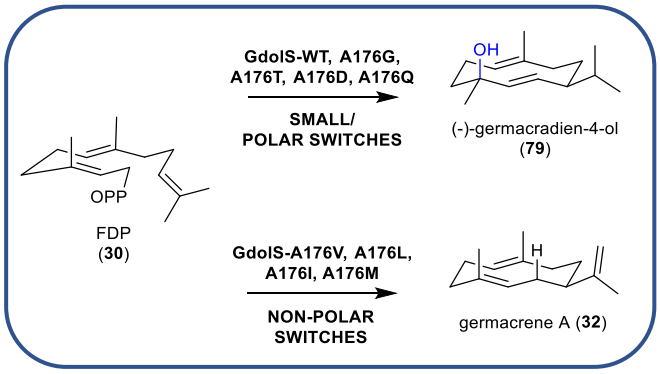


Figure 153. Representation of product selection for GdolS-WT and A176 mutants.

In addition, T175 and A177 were mutated. T175C and T175N mutations did not alter the product profile, albeit resulting in enzymes with lower catalytic values in comparison with the wild type enzyme, and T175D was inactive. T175 was hypothesised to participate in hydrogen-bonding interactions, and

these results indicate that cysteine and asparagine could functionally substitute T175 in the putative hydrogen-network. As for A177, it was substituted with isoleucine resulting in an inactive enzyme. The results of these substitutions were inconclusive and further mutations at positions T175 and A177 with amino acids of different size and functionality are needed for a better interpretation of their enzymatic role for catalysis.

Furthermore, the catalytic role of H150 in GdolS water capture was examined. H150Y and H150F were functional GdolS enzymes, albeit having 1 and 2 orders of magnitude lower catalytic efficiency as compared to the wild type enzyme, respectively. However, H150C showed a dramatic loss of catalytic efficiency and H150R was inactive under the experimental conditions used. These results suggest that H150 may govern cation-π stabilization of a putative carbocation. H150W resulted in a drastic drop of catalytic efficiency, presumably being too large and disturbing the template required for catalysis.

In EpicS, the metal binding motifs DDxxD and NSE were subjected to single point mutations to investigate their role in enzyme catalysis. These mutations reduced the catalytic efficiency of EpicS, showing their crucial role to catalysis. Remarkably, substitution of the second aspartate in the DDxxD motif to glutamate (D82E) resulted in a functional GdolS, highlighting the importance of this amino acid in water management.

EpicS mutants were also generated in the G1/2 helix break region. R180K and R180H were detrimental mutations for catalysis, also T183S resulted in an inactive enzyme. Interestingly, G184I, which was found in the insoluble fraction in the same expression conditions, generated farnesol after basic extraction of the protein into the soluble fraction. Other mutations will be required in this region to complete an investigation of their role in catalysis.

The broad applicability of natural terpenoids in industry often relies on the access to functionalised hydrocarbon structures, such as oxidised terpenes. The regio- and stereo-specific water capture mechanism in some terpene synthases provides a cost-effective approach for the production of oxygen-containing terpenes, which otherwise requires the use of P450s or complex non-biological chemical transformations. However, the presence of water in the active site of sesquiterpene synthases is not trivial and challenging to analyse. Despite the need for more accurate computational studies, site-directed mutagenesis is a powerful approach to show how the product specificity of sesquiterpene synthases can be rationally modified and facilitates the process of engineering the active site architecture concerning water capture.

Apart from these SDM studies, the reaction conditions for the incubations of GdolS and EpicS with FDP were modulated. Interestingly, the product outcome of the reaction catalysed by EpicS changed from pH 8.0 to pH 10.0, increasing and decreasing the production of (+)-germacradien-4-ol and (+)-epicubenol, respectively. Probably, there is a subtle change in the folding architecture between these

pHs that change the product outcome. Further experiments need to be done to investigate this more definitively.

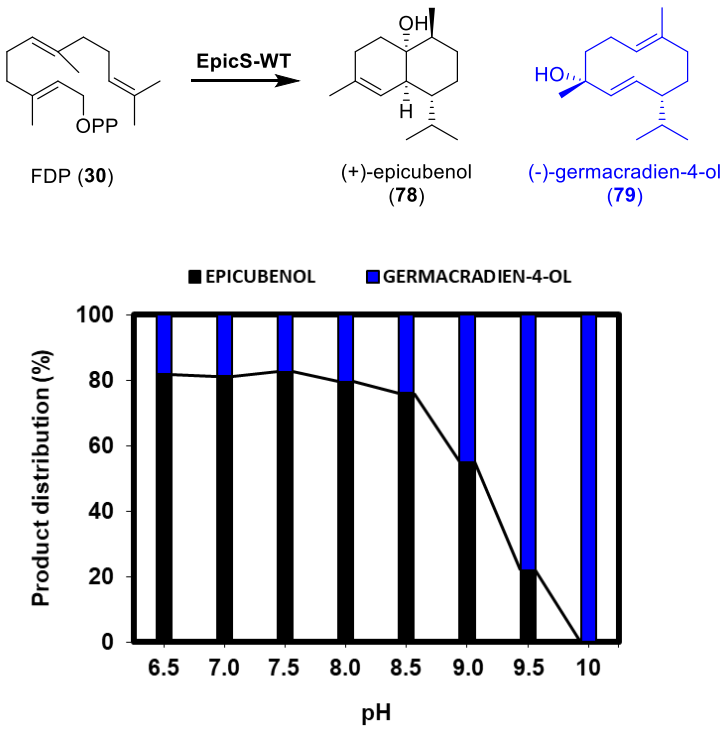


Figure 154. Representation of the product selection after the incubation of FDP (**30**) with EpicS across a pH range 6-10

The second part of the project was the rational design and synthesis of **FDP analogues**, which were incubated with a series of sesquiterpene synthases with the aim of accessing alternative pathways to generate novel terpenoids. The pentane extractable products after these incubations were analysed by GC-MS for determination of product identity.

The synthesis of (*E*,*Z*)-6F-FDP was attempted using a HWE reaction and the olefination was carried out following two different approaches. First, the sequence described by Weiler and Sum[242,243] was followed. This relies on dianion formation and phosphoenolate formation before methylation. However, this resulted in low yields and an alternative approach was developed, which relies on the formation of a β-ketoester, decarboxylation and a second HWE reaction. The later approach was also used for the synthesis of (*E*,*E*)-6F-FDP, (*E*,*E*)-6Me-FDP, (*E*,*Z*)-6Me-FDP and (*E*,*E*)-7H-FDP. In particular, to produce (*E*,*E*)-7H-FDP required the synthesis of the precursor aldehyde used in the first HWE reaction, which was obtained through a malonate-based route.

The incubations of (*E*,*Z*)-6F-FDP enabled further insights into how of sesquiterpene synthases discriminate between 1,6-, 1,10-cyclisation mechanisms and also to access novel fluorinated sesquiterpenes. In 1,10-sesquiterpene synthases (GAS, GDS, GdolS and Gd11olS), the natural analogues were obtained with high selectivity suggesting that the fluorine atom at C6 does not affect

the requirements needed for a 1,10-cyclisation, except for Gd11olS, which gave rise to an additional product in about 50:50 ratio. The identity of this product remains unknown, and a preparative-scale incubation of Gd11olS with (*E*,*Z*)-6F-FDP and characterisation of the obtained products (separated or as a mixture) by NMR spectroscopy work need to be done. The incubation of ADS, which is a 1,6-sesquiterpene synthase, gave rise to two pentane extractable products. No cyclic product formation was expected because the fluorine at C6 can reduce the nucleophilicity of the C6-C7 double bond and/or reduce the stability of the forming carbocation after 1,6-cyclisation. A potential non-cyclic product is 6F-farnesene, which could be probed by coelution with authentic samples. The mass spectrum of the other product remains cryptic but may indicate a bicyclic compound. This could be investigated by NMR spectroscopy analysis of the pentane extracts from a preparative-scale incubation. Lastly, (*E*,*Z*)-6F-FDP was incubated with 1,10-sesquiterpene synthases that are able to perform a second 1,6- cyclisation (PRAS, SdS, EpicS and DCS-W279Y) and the organic extractable products obtained were analysed by GC-MS. PRAS and SdS produced the 6F-analogues of their proposed neutral intermediates along the course of their catalysed reaction cascades, which are (*E*,*Z*)-6F-germacrene A and (*E*,*Z*)-6F-germacrene B, showing that the second cyclisation is hindered by either reduction of the electronic density of the C6-C7 double bond implicated in catalysis or due to destabilisation of a putative carbocation after the 1,6-ring closure.

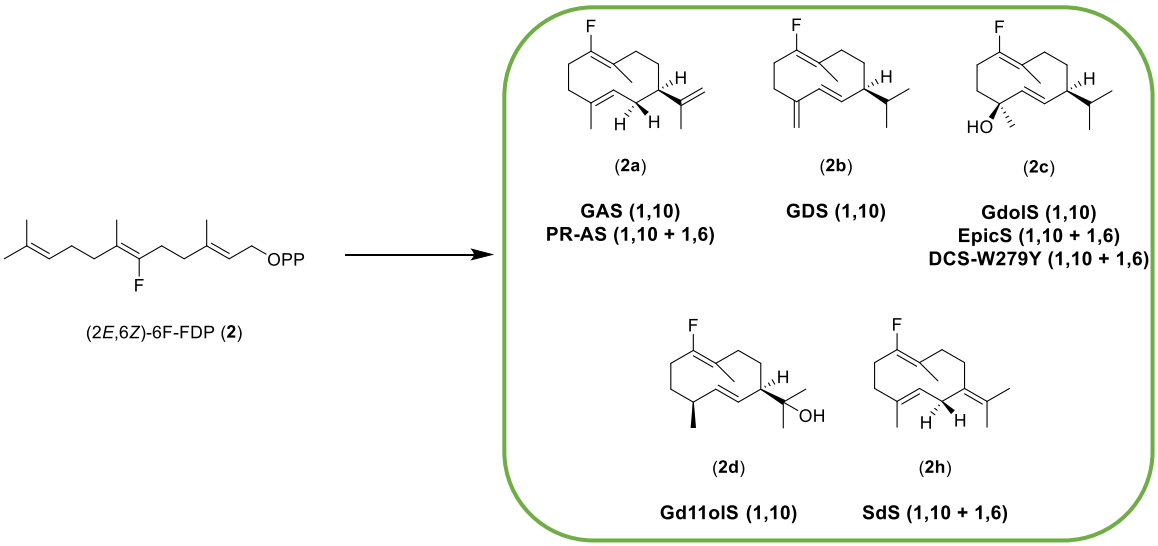


Figure 155. Structures of the observed enzymatic products upon incubation of (*E*,*Z*)-6F-FDP (**2**) with sesquiterpene synthases.

The incubations of (*E*,*E*)-6F-FDP with the sesquiterpene synthases used in this study were novel in nature and gave key lessons. Overall, the sesquiterpene synthases tested generated (*E*,*E*)-6F-germacrene A, which indicated that these enzymes are able to transform C6-C7 isomeric counterpart of the natural (*E*,*E*)-FDP. GDS and Gd4olS cannot access their natural products due to the impossible generation of an allylic cation after the 1,3-hydride shift, and thus generated abortive (*E*,*E*)-6F-germacrene A. ADS generated a mixture of products, giving further evidence of an altered 1,6-

cyclisation. PRAS produced (*E*,*E*)-6F-germacrene A with high selectivity, and SdS generated a major product that was proposed to be (*E*,*E*)-6F-germacrene B. This is because the mechanism to access (*E*,*E*)-6F-germacrene B by SdS should not be hindered with the presence of fluorine atom at C6 (in contrast with that observed with GDS and GdolS catalysis), however, the conformation of the 10-membered ring is not suitable for the proton-mediated 1,6-cyclisation. The incubation of EpicS and DCS-W279Y with (*E*,*E*)-6F-FDP yielded (*E*,*E*)-6F-germacrene A with high selectivity, with served as controls to validate the proposed mechanistic interceptions.

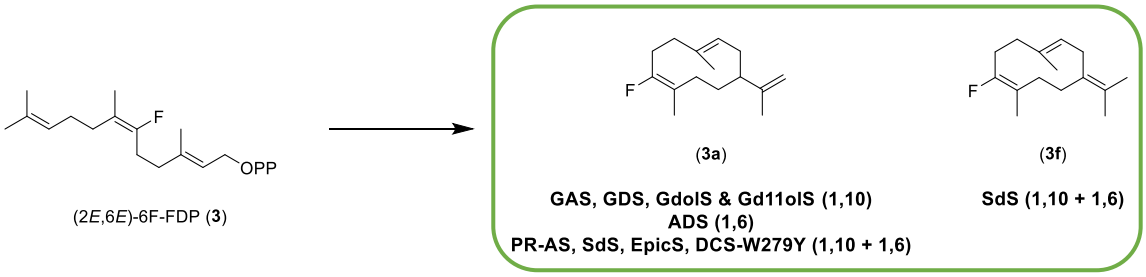


Figure 156. Structures of the observed enzymatic products upon incubation of (*E*,*E*)-6F-FDP (**3**) with sesquiterpene synthases.

(*E*,*E*)-6Me-FDP contains an extra methyl group at the C6-C7 double bond, which might alter the required conformation of the substrate for catalysis and result in alternative pathways. It can also increase the stabilisation of the adjacent putative carbocation after 1,6-ring closure. In general, all sesquiterpene synthases used in this study behaved more promiscuously when compared with their natural catalysed reactions upon incubation with FDP, except GDS and ADS which generate exclusively one product. In 1,10- sesquiterpene synthases, we presumably obtained the corresponding natural analogues, albeit with formation of side products. GAS generated (*E*,*E*)-6Me-germacrene A and (*E*,*E*)-6Me-germacrene D, and GdolS and Gd11ols generated (*E*,*E*)-6Me-germacradien-4-ol and (*E*,*E*)-6Me-germacradien-11-ol, respectively. Remarkably, ADS presumably produce 6Me-amorphadiene. The identities of these products should be confirmed by NMR spectroscopy characterisation, but this formation would represent a highly important contribution towards the generation of modified artemisinin analogues from amorphadiene, which is highly in demand these days. Lastly, the 1,10- + 1,6- sesquiterpene synthases generated a complex mixture of products. Although bicyclic products might have been formed, these results clearly show the reduced ability to perform a second 1,6- cyclisation from a germacrene intermediate in these synthases. The identities of these products need to be further analysed.

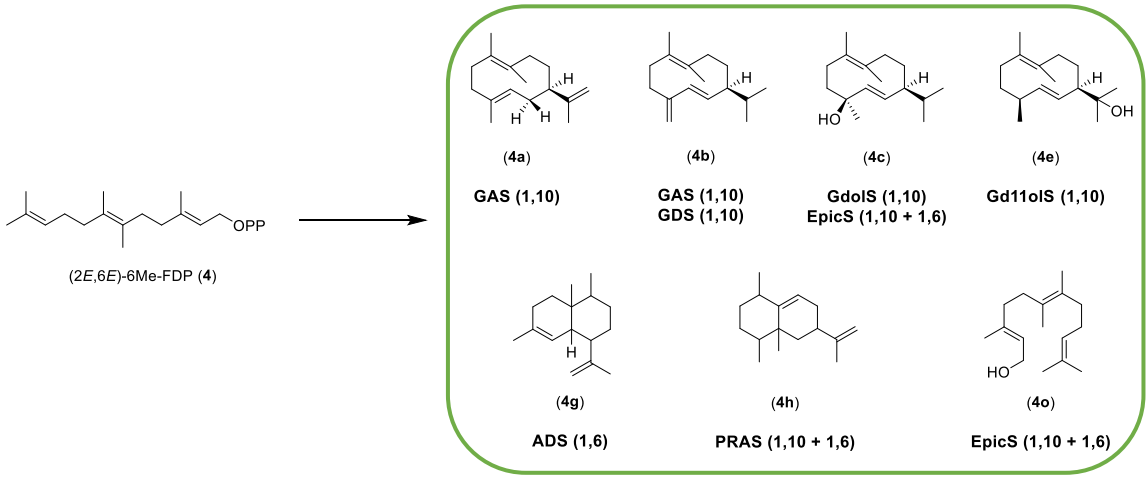


Figure 157. Structures of the observed enzymatic products upon incubation of (*E*,*E*)-6Me-FDP (**4**) with sesquiterpene synthases.

(*E*,*Z*)-6Me-FDP was also generated to continue exploring the ability of sesquiterpene synthases to accommodate and transform C6-C7 isomers of the natural (*E*,*E*)-FDP. A common feature was that all the sesquiterpene synthases generated the proposed product (*E*,*Z*)-6Me-germacrene A upon incubation with (*E*,*Z*)-6Me-FDP, which may validate some of the proposals mentioned above regarding the conformational changes arising from the use of an isomeric alternate substrate. For 1,10- sesquiterpene synthase, the main product observed was (*E*,*Z*)-6Me-germacrene A, albeit with some exceptions. GAS also generated (*E*,*Z*)-6Me-farnesol, which arises from aberrant water quench of the linear (*E*,*Z*)-6Me-farnesyl cation. Also, GDS generated a mixture of minor products, whose identities should be further investigated in the future. Incubation of ADS, a 1,6- sesquiterpene synthase, resulted in a mixture of extractable products upon incubation with (*E*,*Z*)-6Me-FDP. This is in agreement with a 1,6- cyclisation mechanisms, that might be perturbed due the presence of an additional methyl substituent in the C6-C7 double bond. The identities of these compounds also need to be further investigated. The 1,10- + 1,6- sesquiterpene synthase yielded (*E*,*Z*)-6Me-germacrene A as the predominant product in the organic extracts. PRAS and SdS do not perform a second cyclisation due to the conformation of their neutral intermediates (*E*,*Z*)-6Me-germacrene A and (*E*,*Z*)-6Me-germacrene B, respectively. Thus, PRAS generated exclusively (*E*,*Z*)-6Me-germacrene A and SdS resulted in the production of (E,Z)-6Me-germacrene A and (*E*,*Z*)-6Me-germacrene B. EpicS and SdS do not produce their natural product analogues upon incubation with (*E*,*Z*)-6Me-FDP, because the allylic cation system required for carbocation stabilisation after 1,3-hydride shift is not present in the catalytic process. Thus, they produce (*E*,*Z*)-6Me-germacrene A by abortive deprotonation of germacrenyl-like cation. EpicS also gave (*E*,*Z*)-6Me-farnesol, which indicated lower fidelity to perform the initial 1,10- cyclisation upon incubation of EpicS with (*E*,*Z*)-6Me-FDP as compared with the use of (*E*,*E*)-FDP.

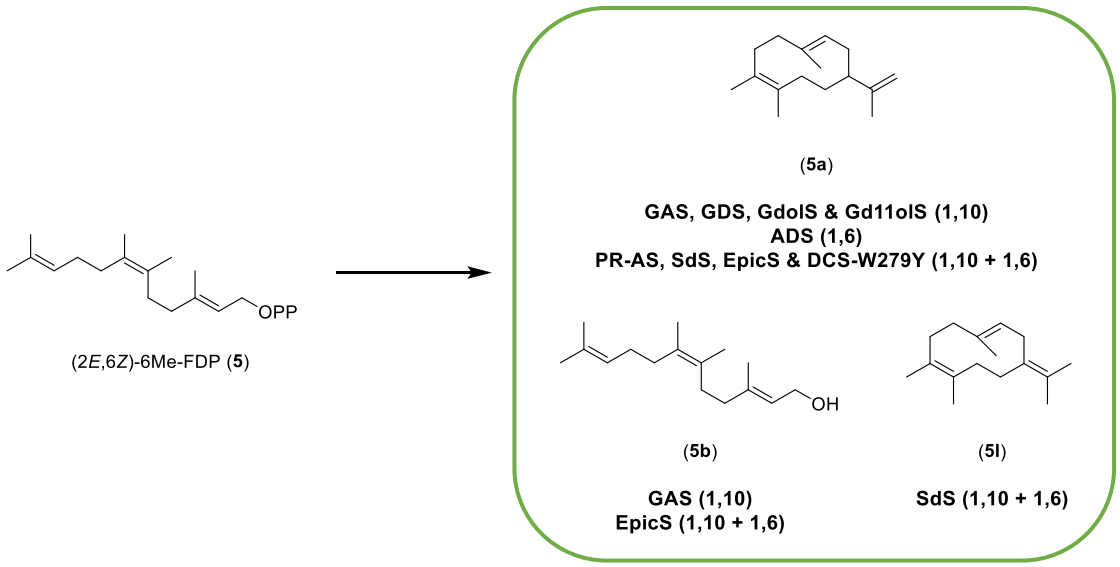


Figure 158. Structures of the observed enzymatic products upon incubation of (*E*,*Z*)-6Me-FDP (**5**) with sesquiterpene synthases.

The incubations of (*E*,*E*)-7H-FDP with these sesquiterpene synthases were performed for the first time and further highlighted the generation of novel sesquiterpene compounds by simple changes in the FDP backbone. The incubations of (*E*,*E*)-7H-FDP with 1,10-sesquiterpene synthases (germacrene synthases) gave rise to the natural analogues with high selectivity, which give evidence that their catalysed mechanisms remain unchanged with one less substituent in the C6-C7 double bond. In the case of ADS (1,6-sesquiterpene synthase), at least three 7H- sesquiterpene products were observed in the pentane extracts after incubation with (*E*,*E*)-7H-FDP. The identities of these products are unsolved and need to be investigated in the future. Here it is proposed that at least one product might be a 7H-farnesene analogue because the absence of the C6- methyl group can reduce electron density at C6-C7 double bond required for intramolecular nucleophilic attack. Also, if a 1,6-cyclisation were to occur, it would generate a high-energy secondary bisabolyl-like carbocation. This could be demonstrated by preparing 7H-farnesenes from the 7H-farnesol equivalent and comparison of their retention times and mass spectra by GC-MS, or by means of NMR spectroscopy. The 1,10- + 1,6-sesquiterpene synthases behaved more promiscuously upon incubation with (*E*,*E*)-7H- FDP compared to the products arising from incubation of natural (*E*,*E*)-FDP, which may suggest that there is more flexibility and motion of intermediates during catalysis in these enzymes. PRAS gave (*E*,*E*)-7H-germacrene A as a major product. The side products in this case could be germacrene A derivates because they disappeared upon induced Cope rearrangement. SdS also generated (*E*,*E*)-7H-germacrene A in a mixture of at least 4 more products. The identities of these are still unknown, but it is proposed that (*E*,*E*)-7H-germacrene B might be a product. 7H-selinadiene could be also formed, but it is still unknown whether the C6-C7

double bond with one less substituent can accept a proton to assist the 1,6- cyclisation needed in this case, which is of future interest. In EpicS and DCS-W279Y, the second 1,6-cyclisation did not occur when using (*E*,*E*)-7H-FDP either because of the reduction of the nucleophilicity in the C6-C7 double bond needed for intramolecular attack or due to the formation of a putative secondary carbocation after 1,6-cyslisation. As a result, (*E*,*E*)-7H-germacradien-4-ol is observed in their pentane extracts. In EpicS, (*E*,*E*)-7H-germacrene A and (*E*,*E*)-7H farnesol were also observed, indicating the interception of the natural mechanism at different stages of the reaction cascade catalysed by EpicS.

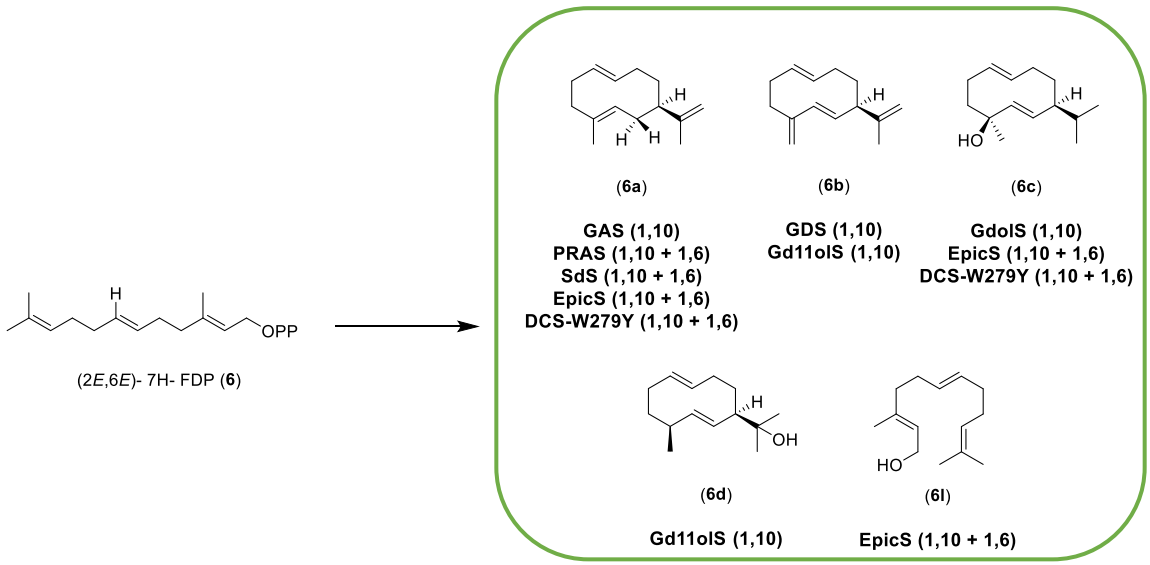


Figure 159. Structures of the observed enzymatic products upon incubation of (*E*,*E*)-7H-FDP (**6**) with sesquiterpene synthases.

The identification of the generated sesquiterpene analogues through analysis of their mass spectra and through prediction based upon the proposed catalytic mechanism of these enzymes will benefit from NMR spectroscopy interpretation of their structure. Here, (2*E*,6*E*)-7H-germacradienol (**6d**) has been produced in sufficient quantity for 1D- and 2D-NMR spectroscopy characterisation, confirming its identity.

In the context of this project, and to further exploit the substrate promiscuity of these enzymes, the following compound could be of interest:

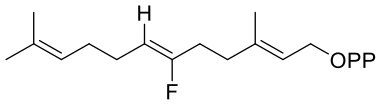


Figure 160. Structure of (2*E*,6*Z*)-6F,7H-FDP, a prospective future analogue.

The last part of the project was to develop a directed evolution approach to generate high yielding sesquiterpene synthases. This approach relies on the coupling of an EP-PCR technique to generate random mutations in the genes encoding for sesquiterpene synthases (or a region of these genes) with the Golden Gate technique for cloning of the amplified fragments, and a colorimetric screen that allows the selection of the improved enzymes based on substrate consumption. The validation of the screening system was done using active and inactive sesquiterpene enzymes. First, the screening plasmid encoding the genes for carotenoid production were transformed in *XL1*-blue cells and the sesquiterpene genes were independently introduced in pBAD expression vectors. To find out the optimal concentration of L-arabinose (pBAD inducer), different concentrations of L-arabinose were tested with a constant concentration of IPTG (carotenoid pathway inducer) and using GdolS encoding plasmid as a model. Once this was done, SdS (active) and GdolS-A176F (inactive) genes were independently cotransformed with the screening plasmid. The observed pigmentation was reduced when the active enzyme was used, but remained unchanged when the enzyme used was inactive. After these results, the EP-PCR conditions were optimised and the genes encoding for the G1/2 helices of SdS, GdolS and EpicS were randomised. Future work needs to be carried out to continue developing this approach. First, these libraries of mutated fragments need to be subcloned in the appropriate acceptor vectors (primers have been already designed to generate acceptor vectors only missing the G1/2 helix encoding genes). The screening of new genes generated should visually indicate active sesquiterpene synthases. This screening needs to be repeated to confirm the phenotype and the pigmentation needs be visualised using fluorometric technique. For example, the carotenoids can be extracted in acetone and the colour intensity can be evaluated by absorbance measurements. Other techniques such as FACS could be used for the analysis. Then, the best candidates should be expressed, characterised and their sequence must be confirmed to allow identification of the successful mutants. The best candidates will be subjected to further rounds of directed evolution to fully optimise for either productive or alternate product generation.

CHAPTER 7

MATERIALS AND METHODS

Chapter 7. Materials and methods

7.1. Biological methods

7.1.1. Materials and general methods

Oligonucleotides primers were purchased from Sigma Aldrich (UK). *Pfu* DNA polymerase, *Taq* DNA polymerase, PrimeSTAR® polymerase, T4 DNA ligase, dNTPs, restriction enzymes, safe green dye, and protein and DNA markers were purchased from Thermofisher scientific (UK) or New England BioLabs (UK). The spin columns for DNA mini preparations were purchased from QIAGEN (UK) and dialysis membrane was purchased from Medicell. DNA concentration was measured using a Nanodrop 3300 Fluorospectrometer from Thermo Fisher Scientific. All mutated and ligated constructs were confirmed by DNA sequence analysis using the internal Eurofins facilities (School of Bioscience, Cardiff University, UK). Nickel beads used for nickel affinity chromatography were purchased from Expedeon$_{LTD}$.

[1-^{3}H]-FDP was purchased from Fluorochem, 'Ecoscint™ O' and 'Ecostint™ A' scintillation fluids were purchased from National Diagnostics. [1-^{3}H]-FDP was diluted by addition of unlabelled FDP to give a final specific activity of 75 mCi/mmol. All other chemicals were from Sigma-Aldrich, Fisher or Melford.

Protein concentration was measured by the Bradford method[185] using bovine serum as the calibration standard. UV spectroscopy was performed using a Jasco V-660 spectrophotometer.

The expression vector encoding GdolS gene from *Streptomyces citricolor* (construct pET16b-SC1) and the chromosomal DNA of Streptomyces griseus IFO13350, which contains the gene encoding for EpicS, were received generously from Yasuo Ohnishi (University of Tokyo, Tokyo).

7.1.2. Bacterial strains

Two strains of *E. coli* were used in this investigation. BL21(DE3)-Codon Plus RP (BL21(DE3)-RP) were used for expression of all the wild-type proteins and mutants, while XL1-Blue were used for PCR cloning of all the genes and for co-expression of proteins in the colorimetric screening method.

7.1.3. Competent cells

Preparation

The desired cell strain (50 μL, BL21-RP or XL1-Blue) was incubated in LB medium overnight at 37 ºC whilst shaking (150 rpm). 1 mL of the overnight culture was used to inoculate 100 mL of LB medium, until the optical density at 600 nm reached 0.6 AU (OD_{600}= 0.6). The cells were then placed on ice for 15 minutes, harvested at 3300 g for 5 min, and the supernatant solution was discarded. The cells were then resuspended in calcium chloride buffer 1 (40 mL) and placed on ice for further 15 minutes. Then, the mixture was centrifuged again, and the pellet was resuspended in calcium chloride buffer 2 (5 mL). After this, the cells were divided into 50 μL aliquots, frozen immediately with liquid nitrogen and stored at -80 ºC.

Calcium chloride buffer 1

Calcium chloride (100 mM) was dissolved in deionised water (100 mL). Then, the solution was sterilised in an autoclave (20 min, 121 ºC).

Calcium chloride buffer 2

Calcium chloride (100 mM) was dissolved in deionised water and glycerol (15% v/v) was added. Then, the final volume was made up to 100 mL with deionised water and the solution was sterilised in an autoclave (20 min, 121 ºC).

7.1.4. Supercompetent cells

Preparation

Supercompetent cells were prepared as described above for competent cells. However, this time it was used rubidium chloride buffer 1 and rubidium chloride buffer 2 instead of calcium chloride buffers 1 and 2.

Rubidium chloride buffer 1

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

Rubidium chloride buffer 2

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

7.1.5. Preparation XL1-Blue containing pJ211-lacI-mP$_{T5}$-MNF

pJ211-lacI-mPT5-MNF (plasmid 1) was transformed into supercompetent XL1-blue cells. After gently defrost 50 μL aliquot of supercompetent XL1-Blue cells, 50 ng of plasmid 1 was added, and the mixture was kept on ice for 30 minutes. Following heat shock (40 ºC, 45 s), the mixture was returned to the ice. 1 mL of LB medium was added to the mixture and incubated at 37 ºC for one hour with shaking (150 rpm). Cells were harvested by centrifugation (1 min, 3300 g), resuspended in 200 μL of fresh LB medium and plated onto LB agar medium containing kanamycin for selection (50 μg mL^{-1}). The plate was incubated for 16 hours at 37 ºC. Then, a single transformed colony was used to inoculate 100 mL of LB medium containing kanamycin (50 μg mL^{-1}) at 37 ºC overnight. 1 mL of the overnight culture was used to inoculate 100 mL of LB medium containing kanamycin (50 μg mL^{-1}), and the same procedure described for the preparation of cells was followed (7.1.4., above). The cells were stored at -80 ºC.

7.1.6. Growth media and antibiotics

Luria-Bertani (LB) medium

Tryptone (10 g), yeast extract (5 g) and sodium chloride (1 g) were dissolved in 1 L of deionised water. The resulting solution was autoclaved (20 min, 121 ºC) and then cooled to room temperature for its use.

LB agar

Tryptone (10 g), yeast extract (5 g), sodium chloride (1 g) and agar (15 g) were dissolved in 1 L of deionised water, and the resulting solution was autoclaved (20 min, 121 ºC). After this, the solution was allowed to cool to about 40 ºC before the required antibiotic was added and the LB agar was poured into plates under aseptic conditions. The plates were allowed to slowly cool to room temperature and then stored at 4 ºC.

Ampicillin

Ampicillin was dissolved in deionised water to give a stock 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 stock 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.

7.1.7. Traditional cloning of EpicS (insert) into pET16b vector

Insert amplification

EpicS encoding gene was amplified from the chromosomal DNA of *S.griseus* IFO13350. The primers used contain NdeI and BamHI restriction sites (Table 27, in italics). The PCR components and protocol are detailed in Tables 28 and 29, respectively. The amplified fragment was purified by agarose gel electrophoresis.

EpicS-Fw	GGCCATATCGAAGGTCGT*CATATG*GACAGCGAACTGCCGGACATC
EpicS-Rv	CTTTGTTAGCAGCC*GGATCC*TCAGCGGGCCCGGGGAGCCTC

Table 27. Primers used for the amplification of the gene encoding for EpicS.

Component	Concentration
1. Pfu DNA polymerase buffer (5X)	1X
2. dNTPs	200 μM each
3. Fw primer	1 μM
4. Rv primer	1 μM
5. DNA template	0.1-0.2 μg/50μL
6. Primestar polymerase (2.5 U/μL)	1.25 U
7. Nuclease-free water to final volume of 50 μL	

Table 28. PCR chemical components for the amplification of the gene encoding for EpicS.

Step		Temperature (°C)	Time (min)
Initial denaturation		95	3
Denaturation	**x 33**	95	1
Annealing		55-62	2
Extension		72	12
Final extension		72	10

Table 29. PCR protocol for the amplification of the gene encoding for EpiCS.

Restriction site digestion

Both insert encoding EpicS and pET16b-GdolS plasmid were, separately, double digested with BamHI and NdeI restriction endonucleases following the manufacturer's guidelines. The mixtures were left to incubate for 1 h at 37 °C. After digestion, DNA fragments (insert and acceptor vector) were isolated by agarose gel electrophoresis purification.

Ligation of EpicS gene with pET16b vector

After restriction digest and agarose gel electrophoresis purification, the obtained insert and acceptor vector fragments were ligated to construct pET16b-EpicS. Ligation was carried out using T4 DNA ligase. Vector and insert were incubated in a 1:2 molar ratio, respectively, with T4 DNA ligase (1 μL), T4 ligase buffer (2 μL), and sterile deionised water to 20 μL for 16 hours at 16 ºC.

After this time, the mixture was cooled on ice and transformed into XL1-blue E. coli cells to isolate the DNA *via* DNA-miniprep. The obtained pET16b-EpicS plasmid was stored at -20 ºC (and the gene encoding for EpicS was sequenced).

7.1.8. Site directed mutagenesis (SDM)

Site directed mutagenesis was carried out using the Quickchange site directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. QIAprep® spin miniprep kit (QIAGEN, UK) was used for the purification of plasmids according to the manufacturer's instructions. All the mutated plasmids were confirmed by DNA sequence analysis.

The mutagenic primers for used in GdolS and EpicS are detailed in Table 30.

The PCR chemical composition and PCR protocol used for SDM was the same as stated in Tables 28 and 29. After PCR, the DNA mixture was digested with 0.5 μL of DpnI (10 U/ μL) for 1 hour at 37 °C to cleave the parental DNA template. Then, the mixture was cooled on ice and transformed into XL1-blue cells and it was isolated the plasmid DNA *via* DNA-miniprep Mutations were confirmed by DNA sequencing.

Name	5' – Sequence – 3'
GdolS – T175C	Forward: 5' GACGCTCAGACGCGGTTGCGCCGCGATGGAGAG 3' Reverse: 5' CTCTCCATCGCGGCGCAACCGCGTCTGAGCGTC 3'
GdolS – T175D	Forward: 5' GCTCAGACGCGGTGATGCCGCGATGGAGAG 3' Reverse: 5' CTCTCCATCGCGGCATCACCGCGTCTGAGC 3'
GdolS – T175N	Forward: 5' CTCAGACGCGGTAATGCCGCGATGGAGAG 3' Reverse: 5' CTCTCCATCGCGGCATTACCGCGTCTGAG 3'
GdolS – A176G	Forward: 5' CGCTCAGACGCGGTACCGGCGCGATGGAGAGCATCTTC 3' Reverse: 5' GAAGATGCTCTCCATCGCGCCGGTACCGCGTCTGAGCG 3'
GdolS – A176V	Forward: 5' GACGCGGTACCGTGGCGATGGAGAGC 3' Reverse: 5' GCTCTCCATCGCCACGGTACCGCGTC 3'
GdolS – A176L	Forward: 5' CTCAGACGCGGTACCCTGGCGATGGAGAGCATC 3' Reverse: 5' GATGCTCTCCATCGCCAGGGTACCGCGTCTGAG 3'
GdolS – A176I	Forward: 5' CTCAGACGCGGTACCATTGCGATGGAGAGCATC 3' Reverse: 5' GATGCTCTCCATCGCAATGGTACCGCGTCTGAG 3'
GdolS – A176M	Forward: 5' GACGCTCAGACGCGGTACCATGGCGATGGAGAGCATCTTC 3' Reverse: 5' GAAGATGCTCTCCATCGCCATGGTACCGCGTCTGAGCGTC 3'
GdolS – A176F	Forward: 5' CTCAGACGCGGTACCTTTGCGATGGAGAGCATC 3' Reverse: 5' GATGCTCTCCATCGCAAAGGTACCGCGTCTGAG 3'
GdolS – A176T	Forward: 5' CAGACGCGGTACCACCGCGATGGAGAG 3' Reverse: 5' CTCTCCATCGCGGTGGTACCGCGTCTG 3'
GdolS – A176D	Forward: 5' CAGACGCGGTACCGATGCGATGGAGAGCATC 3' Reverse: 5' GATGCTCTCCATCGCATCGGTACCGCGTCTG 3'
GdolS – A176Q	Forward: 5' CTCAGACGCGGTACCCAGGCGATGGAGAGCATC 3' Reverse: 5' GATGCTCTCCATCGCCTGGGTACCGCGTCTGAG 3'
GdolS – A177I	Forward: 5' CAGACGCGGTACCGCCATTATGGAGAGCATCTTCG 3' Reverse: 5' CGAAGATGCTCTCCATAATGGCGGTACCGCGTCTG 3'
GdolS – H150F	Forward: 5' CTACTTCGCCTGCTTTCCCGCGGAGGCCG 3' Reverse: 5' CGGCCTCCGCGGGAAAGCAGGCGAAGTAG 3'
GdolS – H150Y	Forward: 5' CTACTTCGCCTGCTACCCCGCGGAGG 3' Reverse: 5' CCTCCGCGGGGTAGCAGGCGAAGTAG 3'
GdolS – H150W	Forward: 5' CTACTTCGCCTGCTGGCCCGCGGAGGCC 3' Reverse: 5' GGCCTCCGCGGGCCAGCAGGCGAAGTAG 3'
GdolS – H150C	Forward: 5' CTACTTCGCCTGCTGCCCCGCGGAGGC 3' Reverse: 5' GCCTCCGCGGGGCAGCAGGCGAAGTAG 3'
GdolS – H150R	Forward: 5' GTACTACTTCGCCTGCCGTCCCGCGGAGGCCGCCG 3' Reverse: 5' CGGCGGCCTCCGCGGGACGGCAGGCGAAGTAGTAC 3'
EpicS – D81E	Forward: 5' GTTCCTCGTCGAAGACCAGCTCGAC 3' Reverse: 5' GTCGAGCTGGTCTTCGACGAGGAAC 3'
EpicS – D81N	Forward: 5' GGCTGTTCCTCGTCAATGACCAGCTCGACGAC 3' Reverse: 5' GTCGTCGAGCTGGTCATTGACGAGGAACAGCC 3'
EpicS – D82E	Forward: 5' CCTCGTCGACGAACAGCTCGACGAC 3' Reverse: 5' GTCGTCGAGCTGTTCGTCGACGAGG 3'
EpicS – D82N	Forward: 5' CTGTTCCTCGTCGACAATCAGCTCGACGACGG 3' Reverse: 5' CCGTCGTCGAGCTGATTGTCGACGAGGAACAG 3'
EpicS – D85E	Forward: 5' CGACCAGCTCGAAGACGGCCACCTC 3' Reverse: 5' GAGGTGGCCGTCTTCGAGCTGGTCG 3'
EpicS – N226Q	Forward: 5' CACGTGTGCTGGGCCCAGGACGTGTACTCCTTC 3' Reverse: 5' GAAGGAGTACACGTCCTGGGCCCAGCACACGTG 3'
EpicS – N226L	Forward: 5' CACGTGTGCTGGGCCCTGGACGTGTACTCCTTC 3' Reverse: 5' GAAGGAGTACACGTCCAGGGCCCAGCACACGTG 3'
EpicS – S230T	Forward: 5' CCAACGACGTGTACACCTTCGAGAAGGAG 3' Reverse: 5' CTCCTTCTCGAAGGTGTACACGTCGTTGG 3'
EpicS – S230A	Forward: 5' GCCAACGACGTGTACGCGTTCGAGAAGGAGCAG 3' Reverse: 5' CTGCTCCTTCTCGAACGCGTACACGTCGTTGGC 3'
EpicS – E234D	Forward: 5' CTTCGAGAAGGATCAGGTGCTCGGC 3' Reverse: 5' GCCGAGCACCTGATCCTTCTCGAAG 3'
EpicS – R180K	Forward: 5'GAGACGTACATCGCCAAGAAACGTCACACCGGGGCCATC 3' Reverse: 5' GATGGCCCCGGTGTGACGTTTCTTGGCGATGTACGTCTC 3'
EpicS – R180H	Forward: 5' GTACATCGCCAAGCATCGTCACACCGGGG 3' Reverse: 5' CCCCGGTGTGACGATGCTTGGCGATGTAC 3'
EpicS – T183S	Forward: 5' CAAGCGCCGTCACAGCGGGGCCATCCAC 3' Reverse: 5' GTGGATGGCCCCGCTGTGACGGCGCTTG 3'
EpicS – G184I	Forward: 5' CAAGCGCCGTCACACCATTGCCATCCACGTCTGC 3' Reverse: 5' GCAGACGTGGATGGCAATGGTGTGACGGCGCTTG 3'

Table 30. Mutagenic primers used for GdolS and EpicS.

7.1.9. Gibson assembly. Insertion of the TEV cleavage site into EpicS

The TEV cleavage site was inserted between the His-tag and the first codon in the gene that encodes for EpicS by using the Gibson assembly method.[210]

First, the primers were designed to amplify the desired fragment with appropiate overlaps for the following ligation and insertion of the TEV cleavage site encoding nucleotides (Table 31, in italics). Then, the desired DNA fragment was amplified by two step method PCR, using a high-fidelity DNA PrimeSTAR® polymerase (Tables 32 and 33).

The desired linearised gene (1 fragment) was added to Gibson Assembly Master Mix and incubated at 50 ºC for 1 hour. The resulting DNA plasmid was purified by agarose gel electrophoresis chromatography and stored at -20 ºC. Sequencing results confirmed the incorporation of the TEV cleavage site at the desired region of EpicS encoding gene.

Name	5' – Sequence – 3'
TEV-Fw	CATCATCAC*GAGAATCTTTATTTTCAGAGC*AGCGGCCATATCGAAGG
TEV-Rv	CCGGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTG*CTCTTAGAAATAAAAGTC*

Table 31. Primers used for amplification of the desired fragment of EpicS for its use in Gibson assembly.

Component	V (µL)
1. PrimeSTAR® HS (premix)	25
2. Primer Fw (100 µM)	0.5
3. Primer Rv (100 µM)	0.5
4. DNA template 50 ng μL^{-1}	1
5. DMSO	2
6. Nuclease-free water to final volume of 50 µL	

Table 32. PCR chemical components.

Step		Temperature (°C)	Time
Initial denaturation		98	3 minutes
Denaturation	**x 33**	98	15 seconds
Extension	**x 33**	72	7 minutes
Final extension		72	20 minutes

Table 33. PCR protocol.

7.1.10. Golden gate Assembly. Construction of plasmids 2-5 and error-prone PCR (EP-PCR)

Golden Gate Assembly was used for the constructions of plasmids 2-5.[298] The inserts encoding the sesquiterpene synthases (inserts 1-4) were cloned into the acceptor vectors (ac. vectors 1-3) by designing overlapping ends (containing BbSI Type IIS restriction enzyme sites) which can join the desired amplified DNA together after cleavage by the non-palindromic restriction enzyme (BbSI) followed by ligation to bind the DNA together. Note that the primers used for GdolS and GdolS-A176M were the same.

The primer used for the amplification of inserts and acceptor vectors in this study are detailed in tables 34 and 35.

Name	5' – Sequence – 3'
EPICS-INSERT	Forward: 5' GATGAAGACTAATGGACAGCGAACTGCCG 3' Reverse: 5' GATGAAGACTATCAGCGGGCCGG 3'
GDOLS-INSERT	Forward: 5' GATGAAGACTAATGTCCGACGACACCTCAC 3' Reverse: 5' GATGAAGACTATCAGCCGTGGTGGTAGC 3'
SDS-INSERT	Forward: 5' GATGAAGACTAATGGAGCCGGAACTGACC 3' Reverse: 5' GATGAAGACTATTACGCGCTACGTTGCG 3'
EPICS-ACVECTOR	Forward: 5' TATGAAGACATCTGAAGCTCGAGATCTGCAGCT 3' Reverse: 5' AATGAAGACTACCATGGTTAATTCCTCCTGTTAGCC 3'
GDOLS-ACVECTOR	Forward: 5' TATGAAGACATCTGAAGCTCGAGATCTGCAGCT 3' Reverse: 5' AATGAAGACTAACATGGTTAATTCCTCCTGTTAGCCC 3'
SDS-ACVECTOR	Forward: 5' TATGAAGACATGTAAAGCTCGAGATCTGCAGCT 3' Reverse: 5' AATGAAGACTACCATGGTTAATTCCTCCTGTTAGCC 3'

Table 34. Primers used for amplification of the genes encoding the whole sequence of GdolS, GdolS-A176F, EpicS and SdS (inserts 1-4). And primers used for amplification of the genes encoding the whole sequence of the pBAD vector (acceptor vectors 1-3).

Name	5' – Sequence – 3'
EPICS-G1/2-INSERT	Forward: 5' gatgaagactaGAGGAGACGTACATCGCC 3' Reverse: 5' gatgaagactaCGCGACGATCTCTATGAGG 3'
GDOLS-G1/2-INSERT	Forward: 5' gatgaagactaCGAGAGGGCTATCTGACG 3' Reverse: 5' gatgaagactaTCCGAGCCGCTCGATCATG 3'
SDS-G1/2-INSERT	Forward: 5' gatgaagactaCTTAACGATTACACCCTGATGCG 3' Reverse: 5' gatgaagactaGTGACCCATCTCCAGCATC 3'
EPICS-G1/2-ACVECTOR	Forward: 5' tatgaagacatCGCTggcATCGACGCACCG 3' Reverse: 5' aatgaagactaCCTCCGAGGGCACCACG 3'
GDOLS-G1/2-ACVECTOR	Forward: 5' tatgaagacatCGGAcacTTCGAGGTGCCTCAG 3' Reverse: 5' aatgaagactaCTCGgtcCGGTGGCTGCC 3'
SDS-G1/2-ACVECTOR	Forward: 5' tatgaagacatTCACggcTACGAGCTGCAG 3' Reverse: 5' aatgaagactaTAAGgtcCGGCACGGTACC 3'

Table 35. Primers used for amplification of the genes encoding the G1/2 helix of GdolS, EpicS and SdS (inserts 5-7). And primers used for amplification of the genes encoding the whole secuence of the pBAD vector and the sequence of the sesquiterpene synthases except the corresponding G1/2 helixes (acceptor vectors 4-6).

High fidelity amplification

The desired DNA fragments were amplified by PCR, using a high-fidelity DNA PrimeSTAR® polymerase (Tables 36 and 37). This was only applied for the amplification of insert 1-4 and acceptor vectors 1-3.

Component	V (µL)
1. PrimeSTAR® HS (premix)	25
2. Primer Fw (100 µM)	0.5
3. Primer Rv (100 µM)	0.5
4. DNA template 50 ng μL^{-1}	1
5. DMSO	2
6. Nuclease-free water to final volume of 50 µL	

Table 36. PCR chemical components.

Step		Temperature (°C)	Time (s)
Initial denaturation		98	180
Denaturation	**x 33**	98	10
Annealing		55-62	5
Extension		72	300 (vector)/ 60 (insert)
Final extension		72	600

Table 37. PCR protocol.

Golden Gate assembly

The amplified fragments were ligated by Golden Gate assembly. The components detailed in Table 38, were incubated at 37 ºC for 1 hour, followed by incubation at 55 ºC for 5 minutes. This was followed by a heat inactivation step at 85 ºC for 5 minutes. After this time, the mixtures were cooled on ice and transformed into XL1-blue E. coli cells to isolate the DNA plasmids *via* DNA-miniprep. The constructed plasmids 2-5 were stored at -20 ºC. Sequencing of the inserted region was carried out using the primers in Table 39.

Component	V (µL)
1. T4 DNA Ligase buffer (10X)	1
2. T4 DNA Ligase	0.5
3. BbSI	0.5
4. BSA	0.1
5. Vector (100 ng μL^{-1})	1
6. Insert (50 ng μL^{-1})	1
6. Nuclease-free water to final volume of 10 µL	

Table 38. Composition used for ligation using Golden Gate technique.

Name	5' – Sequence – 3'
PBAD-F_W	ATGCCATAGCATTTTTATCC
PBAD-R_V	GATTTAATCTGTATCAGG

Table 39. Standard primers used for sequencing of plasmids 2-5, from pBAD-arabinose promoter.

Low fidelity (EP-PCR) amplification

Alternatively, with the aim of introducing random mutations in the genes of GdolS, EpicS and SdS, it was carried out EP-PCR amplifications of inserts 5-7 using low fidelity *Taq* polymerase. First it was done a test amplification with insert 3, and the conditions with maximum amplification were used to amplify inserts 5-7. These conditions were adapted to increase the randomization rate, adding Mn^{2+}, increasing Mg^{2+} concentration and unbalancing the quantity of dNTPs added. The optimised reaction conditions are stated in tables 40 and 41.

Component	Concentration
1. Buffer (containing KCl)	1X
2. $MnCl_2$	0.5 mM
3. $MgCl_2$	7 mM
4. dATP & dGTP	0.2 μM (each)
5. dCTP & dTTP	1 μM (each)
6. Primers	2 μM (each)
7. Template	200 ng
8. Taq pol. (U μL^{-1})	0.05 U
9. DMSO	4 %
10. Nuclease-free water to final volume of 50 μL	

Table 40. EP-PCR chemical components.

Step		Temperature (°C)	Time (s)
Initial denaturation		94	180
Denaturation	**x 33**	94	45
Annealing		55	30
Extension		68	10
Final extension		68	600

Table 41. EP-PCR protocol. Polymerase was added at the annealing temperature.

7.1.11. DNA visualisation and purification

TAE buffer (50X)

Tris- base (2 M), glacial acetic acid (1 M) and EDTA (0.5 M) were dissolved in deionised water (900 mL). Then, the pH was adjusted to 8.0 and the total volume was made up to 1 L. It was diluted to 1X prior use.

Agarose gel electrophoresis

Agarose gel were preparing using TAE buffer (60 mL) with 1 % w/v ratio of agarose (0.6 g). This mixture was heated and once it was all dissolved was left cooling down. At this time, 5 μL of safe-green dye was added. DNA samples were mixed with green buffer (1:10 ratio buffer and DNA sample, respectively) and loaded into the gel wells, together with a DNA molecular weight ladder on another well to help identify the size of the fragments. The electrophoresis was run at a constant voltage of 100 V for 1 hour, and directly visualised by UV using a Syngene GeneFlash UV light box (Syngene, Cambridge, UK).

Miniprep of DNA

A single colony from an agar plate containing the required transformed cells was used to inoculate 10 mL of LB medium (containing the appropriate antibiotic) and incubated overnight at 37 ºC whilst shaking. After this time, the cells were harvested by centrifugation (5 min, 3300 g). Then, the QIAprep® spin miniprep kit (QIAGEN, UK) was used with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) to purify the plasmid following the manufacturer's instructions.

Agarose gel DNA purification

The required DNA samples were loaded into agarose gel electrophoresis, and the DNA was visualised using green-safe dye. The required DNA fragments were visualised and cut from the gel using a razor blade. Then, the DNA was extracted from the agarose gel using QIAquick gel extraction kit (QIAGEN, UK) with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) following the manufacturer's protocol.

QIAprep® spin miniprep buffers

P1: Tris-HCl (50 mM) and EDTA (10 mM) were dissolved in deionised water (450 mL) and the pH was adjusted to 8. Then, RNase A (1.82 μM) was added and the total volume was made up to 500 mL with deionised water. P1 was store at 4 ºC.

P2: NaoH (0.2 M) an SDS (1 % w/v, 34.7 mM) were dissolved in deionised water (100 mL).

N3: Guanidine hydrochloride (4 M) and potassium acetate (500 mM) were dissolved in deionised water (90 mL) and the pH was adjusted to 4.2. Then the total volume was made up to 100 mL with deionised water.

PB: Guanidine hydrochloride (5 M) and Tris-HCl (20 mM) were dissolved in deionised water (62 mL) and the pH was adjusted to 7.5. Lastly, ethanol (38 mL) was added to the mixture.

PE: NaCl (20 mM) and Tris-HCl (2 mM) were dissolved in deionised water (20 mL). Then, 80 mL of ethanol were added, and the mixture was adjusted to pH= 7.5.

EB: Tris-HCl (20 mM) was dissolved in deionised water (170 mL), and the pH was adjusted to 8.5. Then, the total volume was made up to 200 mL with deionised water.

QIAquick gel extraction buffers

QG: Guanidine thiocyanate (5.5 M) and Tris-HCl (2 mM) were dissolved in deionised water (100 mL), and the pH was adjusted to 6.6.

PE: Prepared as above.

EB: Tris-HCl (10 mM) was dissolve in deionised water (200 mL), and the pH was adjusted to 8.5.

7.1.12. Transformation of BL21-RP competent cells

As stated in 7.1.3, after gently defrost 50 μL aliquot of the appropriate competent cells, 50 ng of plasmid was added, and the mixture was kept on ice for 30 minutes. Following heat-shock (40 ºC, 45 s), mixture was returned to the ice. 1 mL of LB medium was added to the mixture and incubated at 37 ºC for an hour with constant shaking (150 rpm). Cells were harvested by centrifugation (1 min, 3300 g), resuspended in 200 μL of fresh LB medium and plated into LB-agar medium containing appropriate antibiotic for selection (100 μg mL^{-1} ampicillin). The plate was incubated at 37 ºC overnight.

7.1.13. Expression of proteins

A single transformed colony was used to inoculate 100 mL of LB medium containing antibiotic (100 μg mL^{-1} ampicillin). This culture was incubated at 37 ºC overnight whilst shaking (150 rpm). Then, 5 mL of this culture was transferred to 500 mL of LB medium containing 100 μg mL^{-1} of ampicillin. The 500 mL cultures were incubated at 37 ºC until the optical density at 600 nm reached 0.6 (OD_{600}=0.6), at which time gene expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG). From here, the expression protocols differ among enzymes:

Expression of GdolS and GdolS mutants

Gene expression was induced adding 0.2 mM of IPTG and allowed to grow for additional 4 hours at 37 ºC. Cells were harvested by centrifugation at 4 ºC (10 min, 3300 g). The supernatant was discarded, and the pellet stored at -20 ºC.

Expression of EpicS and EpicS mutants

The procedure of expression described above for GdolS resulted unsuccessful in EpicS, thus it was needed to optimise the overexpression of EpicS.

Test expressions

Test expressions in 500 mL cultures were carried out, investigating factors such as the temperature, the length of expression time and the concentration of IPTG used (Table 42). 1 mL sample was removed prior to induction and at defined time intervals after induction. The cells in these samples were harvested by centrifugation (1 minute, 3300 g, Eppendorf centrifuge 5415R) and the pellet resuspended in SDS sample buffer and analysed by SDS-PAGE (Section 7.1.15).

Entry	Temperature (ºC)	Time (h)	IPTG (mM)
1	37	4	0.2
2	25	16	0.3
3	25	16	0.2
4	25	16	0.1
5	15	16	0.3
6	15	16	0.2
7	15	16	0.1

Table 42. Test expression conditions for the optimisation of EpicS-WT overexpression.

Expression of EpicS and EpicS mutants

Gene expression was induced adding 0.1 mM of IPTG and allowed to grow for additional 16 hours at 15 ºC. Cells were harvested by centrifugation 4 ºC (10 min, 3300 g). The supernatant was discarded, and the pellet stored at -20 ºC.

7.1.14. Purification of proteins

Cell lysis

The harvested cells after expression were re-suspended in cell lysis buffer and sonicated. Cells were disrupted by sonication (4 ºC, 5 min, pulse 5 s on/ 10 s off, 40% amplitude) and the resulting suspension was centrifuged at 4 ºC (45 min, 38000 g). SDS-PAGE showed that protein was present in the supernatant fraction.

Cell lysis buffer

Tris-base (50 mM), NaCl (100 mM), Imidazole (5 mM) and glycerol (10 % v/v) were dissolved in deionised water (1 L). Then, the pH was adjusted to 8.

Ni-NTA affinity chromatography

Protocol

The protein supernatant was applied to a Ni-NTA affinity column (Expedeon, 5 cm) equilibrated with Ni-NTA buffer I and the flow-through was collected. Then, the column was washed with 5 column volumes (CV) of Ni-NTA buffer I. After this, the protein was eluted using a gradient of imidazole (50 mM (2 CV), 100 mM (2 CV), 250 mM (2 CV) and 500 mM (2 CV)), and the fractions were analysed by SDS-PAGE. The fractions containing >90 pure protein (M_w GdolS= 38,687 Da or M_w EpicS = 43,150 Da) were combined and dialysed.

The gradient of imidazole concentrations was made using Ni-NTA buffer I and 2.

Ni-NTA buffer I

Tris-base (50 mM), NaCl (100 mM), imidazole (10 mM) and glycerol (10 % v/v) were dissolved in 900 mL deionised water and the pH adjusted to 8.0. The total volume was made up to 1 L with deionised water.

Ni-NTA buffer II

Tris-base (50 mM), NaCl (100 mM), Imidazole (500 mM), and glycerol (10 % v/v) were dissolved in 900 mL deionised water and the pH adjusted to 8.0. The total volume was made up to 1 L with deionised water.

Anion exchange chromatography

EpicS was also purified by anion exchange chromatography. Protein in solution after previous purification step was dialysed into FPLC-Q-sepharose low salt buffer for 2 hours three times. The resulting solution was injected and eluted through a resource Q-sepharose column with a gradient flow of FPLC-Q-sepharose high salt buffer (70 mL, 0-50 % gradient) at a flow rate of 3 mL min^{-1} followed by an isocratic flow of FPLC Q-sepharose high salt buffer (18.1 mL, 100 %) at 3 mL min^{-1}.The UV absorbance was monitored at 280 nm.

FPLC Q-sepharose low salt buffer

Tris-Base (20 mM), β-mercaptoethanol (5 mM) and sodium chloride (20 mM) were dissolved in deionised water (900 mL). Then, the pH was adjusted to 8, and the final volume was made up to 1 L. The resulting buffer was filtered and degassed immediately prior to use.

FPLC Q-sepharose low salt buffer

Tris-Base (20 mM), β-mercaptoethanol (5 mM) and sodium chloride (1 M) were dissolved in deionised water (900 mL). Then, the pH was adjusted to 8, and the final volume was made up to 1 L. The resulting buffer was filtered and degassed immediately prior to use.

Dialysis of proteins

The combined fractions from purification were dialysed using a dialysis tubing with 14 kDa membrane for 16 h against dialysis buffer to remove the excess of imidazole. After this period, the dialysed protein was concentrated using AMICON system (YM 30) and the resulting solution was stored at 4 ºC (or -20 ºC as protein stocks).

Dialysis buffer

Tris-base (10 mM) and glycerol (10 %, v/v) were dissolved in deionised water to a final volume of 3 L. Then, the pH was adjusted to 8.0.

Bradford assay

The concentration of the resulting protein was determined using the method of Bradford.[185,300] Thus, a range of concentrations of bovine serum albumin protein (BSA, from 0.01 μg mL^{-1} to 0.1 μg mL^{-1}) in deionised water (200 μL) was used for the calibration standard curve and the Bradford reagent was used for detection (Table 43).

1 mg/mL stock BSA (µL)	Deionised water (µL)	Bradford reagent (µL)	Sample concentration (µg mL^{-1})
2	198	800	0.01
4	196	800	0.02
6	194	800	0.03
8	192	800	0.04
10	190	800	0.05
12	188	800	0.06
14	186	800	0.07
16	184	800	0.08
18	182	800	0.09
20	180	800	0.1

Table 43. Composition of the dilutions prepared for the Bradford assay.

The absorbance was measured from 430 - 600 nm using SHIMADZU BioSpec UV spectrophotometer. The ratios of the absorbances at 590 nm and 450 nm were calculated and plotted against the BSA concentrations to give a standard curve. Then, the same procedure was repeated using different unknown concentrations of purified protein and the ratios of absorbances were compared to the BSA standard curve to calculate the concentration of the samples.

Bradford reagent

Brilliant blue G250 (0.24 mM) was dissolved in ethanol (4 mL) and phosphoric acid (40 mL) was added. Then, the final volume was made up to 400 mL with deionised water and the solution was stored in dark at 4 ºC.

7.1.15. SDS-PAGE

The protein samples were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The gels were prepared and run using the mini-PROTEAN system from Bio-Rad.

SDS resolving buffer

Tris-Base (1.5 M) was dissolved in deionised water (90 mL) and the pH was adjusted to 8.8. Then, the final volume was taken to 100 mL.

SDS stacking buffer

Tris-Base (0.5 M) was dissolved in deionised water (90 mL) and the pH was adjusted to 6.8. Then, the final volume was taken to 100 mL.

Electrode running buffer 10X

Tris-Base (0.25 M) and glycine (2M) were dissolved in 600 mL deionised water. Then, SDS (10 % w/v) was added to the solution. Once the SDS had also dissolved, the final volume was made up to 1 L with deionised water. Electrode running buffer 10X was diluted to 1X prior to use.

SDS sample buffer

Tris-HCl (0.5 M, 1.25 mL), bromophenol blue (0.6 %, 200 μL), glycerol (2.5 mL), SDS (10 %, 2 mL) and β-mercaptoethanol (500 μL) were dissolved in deionised water (3.55 mL).

Resolving gel (12 %)

Resolving buffer (2.5 mL), 30% acrylamide/ bis-acrylamide (4 mL), SDS (10%, 0.1 mL) and deionised water (3.4 mL) were all mixed. Then, immediately prior to use, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution was mixed.

Stacking gel (4 %)

Stacking buffer (2.5 mL), 30% acrylamide/ bis-acrylamide (1.7 mL), SDS (10%, 0.1 mL) and deionised water (5.7 mL) were all mixed. Then, immediately prior to use, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution was mixed.

SDS gel stain

Coomassie brilliant blue R-250 (60 mg) was dissolved in ethanol. The mixture was made up to 1 L with deionised water and acidified with concentrated hydrochloric acid (2 mL).

SDS-PAGE protocol

Resolving buffer was poured into the gel template and covered with a layer of isopropanol. Once this had polymerised, isopropanol was removed, and the stacking buffer was poured. A gel comb to create the wells was immediately inserted and the solution was allowed to polymerise. After polymerisation, the comb was removed, and gel was set. Protein samples were mixed with SDS sample buffer (1:1 v/v) and incubated 5 minutes at 85 ºC. Then, the samples were loaded into the gel and electrode running buffer was added to the tank. A protein marker was also loaded for identification of the desired proteins. The gel was allowed to run at 150 V for 1 hour.

SDS-PAGE visualisation

Once the gel had run, it was removed from the plates, submerged in water and heated in the microwave for 1 min. This process was repeated twice more. Then, water was replaced with SDS gel stain, heated in the microwave for 2 min and left to cool to room temperature whilst shaking (100 rpm). Then, SDS gel stain was removed and the gel was submerged in water and heated in the microwave for 2 min. This process was repeated until bands were visible.

7.1.16. Circular dichroism spectroscopy (CD)

CD spectrum

Circular dichroism experiments were performed on an Applied PhotoPhysics Chirascan spectrometer. Spectra of EpicS (3 μM) in HEPES (20 mM) buffer at pH 8 was measured between 190 nm and 300 nm using a 10 mm quart cuvette under N_2 with 50 nm/min scan speed, 0.5 nm data pitch, 1 nm bandwidth and 0.5 s response time.

CD melting curves experiments

Circular dichroism melting curve experiments were performed measuring at 222 nm between 5 ºC and 95 ºC, with a temperature ramp of 2 ºC min^{-1}, step of 0.5 ºC and tolerance of 0.2 ºC with measurements taken for 12 s per point.

CD spectra at different pHs

For the spectra measured at different pH values (6-10), the protein was dissolved to the required concentration (3 μM) in the following buffers: 20 mM MES buffer at pH 6, 20 mM HEPES buffer at pH 7 and 8, and 20 mM CHES buffer at pH 9 and 10. The pH was adjusted with the addition of HCl (0.1 M) and NaOH (0.1 M). For pH 6, 9 and 10, measurements were carried out at 0 and 60 minutes to analyse possible changes to the structure of the enzyme over time.

Mean residue ellipticity (MRE)

The spectra obtained from the CD spectrometer were converted to MRE using the equation 1:

Equation 1

$$[\theta]_{MRE}\frac{[\theta]}{10 \cdot n \cdot c \cdot l}$$

Where,

θ = molar ellipticity in millidegrees.

n = number of peptide bonds in the protein.

c = molar concentration of protein in the sample.

l = pathlength in cm.

7.1.17. GC-MS analysis of products

The pentane extractable products after incubations of enzymes with FDP and FDP analogues were analysed by GC-MS.

Incubation buffer

Tris-base (50 mM) and magnesium chloride (5 mM) were dissolved in deionised water (150 mL), and βME (5 mM) was added. The pH was adjusted to 8, and the final volume was made up to 200 mL. This buffer was stored at 4 ºC

Protocol

A solution of 1 μM enzyme and 200 μM FDP in incubation buffer (250 μL) was prepared. The aqueous layer was overlaid with HPLC grade pentane (0.5 mL) and the resulting mixture was incubated (16 h) at 25 °C, whilst shaking (200 rpm). The incubations were repeated without enzyme as negative controls. The reaction was quenched by vortexing for 30 seconds and the products were extracted into the pentane layer twice. The incubations with different metals ions or pHs were run preparing the buffers accordingly. Thus, different divalent cations were used at 0.1 and 1 mM concentration and different buffers were used to adjust the pH range: 20 mM MES buffer at pH 6, 20 mM HEPES buffer at pH 7 and 8, and 20 mM CHES buffer at pH 9 and 10 (as for the CD experiments).

GC-MS methods

GC-MS analysis of incubation products was performed on a Hewlett Packard 6890 GC apparatus fitted with: column= 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 the EI^+ mode with scanning once a second with a scan time of 0.95 s. Method 1: injection port 100 °C; split ratio 5:1; initial pressure 1 kPa; initial temperature 80 °C (1 min hold), ramp of 4 °C/min to 180 °C (2 min hold), flow 1 mL/min; Method 2: injection port 100 °C; split ratio 5:1; initial pressure 1 kPa; initial temperature 80 °C (2 min hold), ramp of 8 °C/min to 280 °C (3 min hold), flow 1 mL/min.

7.1.18. Preparative-scale enzymatic incubations

The preparative incubations of FDP analogue **6** with GDS and Gd11olS were carried out to produce milligram quantities of novel **6b** and **6d** sesquiterpene products. The incubations were scaled to 200 mL with the following buffer: 50 mM Tris, 5 mM $MgCl_2$, 5 mM βME, pH 8, with an enzyme concentration of 5 μM and adding 0.5 mM of substrate. The incubation was overlaid with pentane.
The reaction mixtures were left for 24 hours at room temperature whilst stirring. After this period, the pentane was carefully removed, and fresh pentane was added on top of the mixtures before leaving the reactions for 24 more hours. After 48 hours, the pentane extracts were combined, and the aqueous

layers were further extracted with pentane (3 x 20 mL). The combined extracts were washed with brine and the solvent was removed carefully under reduced pressure (500 mbar minimum pressure at 25 ºC) to about 5 mL. Then, the remaining solvent was distilled from the sample at 40 ºC

7.1.19. Steady state kinetics

General protocol

Kinetics catalysis buffer 1, 2, or 3 (125 μL), extra buffer (40- 110 μL) and (2*E*,6*E*)-[1- ^{3}H]-FDP (0- 40 μM) were placed in an Eppendorf tube and cooled to 0 ºC. The volume of extra buffer varied with the concentration of (2*E*,6*E*)-[1-^{3}H]-FDP, with a total volume of 235 μL. Then, the studied enzyme (15 μL, 0.1 μM) was added to the solution at 0 ºC and overlaid with 1 mL of hexane. The reactions were incubated using a top shaker (10 min, 350 rpm, 30 ºC). After this, the reactions were quenched by addition of EDTA (50 μL, 0.1 M) and vortexed for 30 s. The hexane layer was decanted, and the sample extracted with hexane and Et_2O (11:1) in the same way (2 × 750 μL). The pooled organic extracts were passed through a short column of silica (~500 mg) into 15 mL of Ecoscint fluid (National Diagnostics), and the silica was then washed with a further portion of 11:1 hexane and Et_2O (750 μL). The radioactive mixture was analysed using a TRI-CARB 2900TR Liquid Scintillation Analyzer. Measurements were done in triplicate. In addition, background reactions (no enzyme) were subtracted. The data collected (CPM) was converted to rate vs substrate concentration and fitted to a Michaelis-Menten curve (equation 2), to determine the kinetic parameters k_{cat} and K_M using the commercial SigmaPlot package (Systat Software).

Equation 2

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

The turnover number k_{cat} was calculated from V_{max} using equation 3:

Equation 3

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

Where, V_{max} 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 catalysis buffer 1 (Gd4olS)

Magnesium chloride (5 mM), HEPES (50 mM) and β-mercaptoethanol (5 mM) were dissolved in deionised water (90 mL). Then, the pH was adjusted to 8.0 and the solution diluted to 100 mL with deionised water.

Kinetics catalysis buffer 2 (EpicS)

Magnesium chloride (2 mM), HEPES (50 mM) and β-mercaptoethanol (5 mM) were dissolved in deionised water (90 mL). Then, the pH was adjusted to 8.0 and the solution diluted to 100 mL with deionised water.

Kinetics catalysis buffer 1 (SdS)

Magnesium chloride (10 mM), HEPES (50 mM) and β-mercaptoethanol (5 mM) were dissolved in deionised water (90 mL). Then, the pH was adjusted to 8.0 and the solution diluted to 100 mL with deionised water.

Extra buffer

HEPES (50 mM) and β-mercaptoethanol (5 mM) were dissolved in deionised water (90 mL). Then, the pH was adjusted to 8.0 and the solution diluted to 100 mL with deionised water.

7.1.20. Calculation of errors and normalisation

Standard Errors of the weighted mean

The estimate of the true value for K_M and k_{cat} was calculated as the weighted average. This is need because each set of measurements for the same experiment was carried out separately and independently. The weighted average mean (X_{wav}) was calculated using the following equation 4:

Equation 4

$$X_{WAV} = \frac{\sum w_i x_i}{\sum w_i}$$

Where x_i is each value in the sample and w_i are the reciprocal squares of each measurement's individual uncertainty, calculated with the equation 5:

Equation 5

$$w_i = \frac{1}{\sigma_i{}^2}$$

The errors in this work are expressed as the standard error of the weighted mean (σ_{wav}) that is defined by the equation 6:

Equation 6

$$\sigma_{wav} = \frac{1}{\sqrt{\sum w_i}}$$

Propagation of the errors

Propagation of errors was performed where appropriate using the following equation 7:

Equation 7

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

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

Normalisation

When data was normalised, the equation 8 was used:

Equation 8

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} x\ 100$$

Where X' is the normal value, X is the original value and X_{max} and X_{min} are the maximum and the minimum values of the data set, respectively.

7.2. Synthetic procedures

7.2.1. General synthetic procedures

All chemicals were purchased from Sigma-Aldrich, Alfa-Aesar, Fisher Scientific, Fisher or Fluorochem and used without further purification, unless otherwise stated. Anhydrous tetrahydrofuran (THF), diethyl ether (Et_2O), toluene (tol.), dichloromethane (DCM) and acetonitrile (CH_3CN) were obtained from a MBraun SP800 solvent purification system.

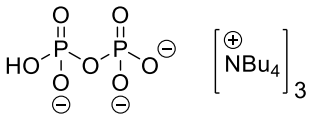
^{1}H NMR, ^{13}C NMR, ^{31}P NMR and ^{19}F NMR spectra were measured on a Bruker Avance III 600, Avance 500, Avance III HD 400 or a Bruker Fourier 300 NMR spectrometer. The spectra are reported as chemical shifts in parts per million, downfield from tetramethylsilane (^{1}H and ^{13}C), trichlorofluoromethane (^{19}F) and phosphoric acid (^{31}P), multiplicity (s-singlet, d-doublet, t-triplet, q-quartet, m-multiplet, dd-doublet of doublet, dt-doublet of triplets, dq-doublet of quartets), integral, coupling and assignment, respectively. Assignments are made to the limitations of COSY, NOESY, DEPT 90/135, gradient HSQC, and gradient HMBC spectra.

Mass spectra were measured on a Waters GCT Premier time of flight mass spectrometer and a Waters LCT time of flight mass spectrometer. The methods used are generally the same as stated in Section 7.1.17.

Thin layer chromatography (TLC) was performed on pre-coated aluminium plates of silica G/UV_{254}. TLC visualisations were performed with 4.2 % ammonium molybdate and 0.2 % ceric sulphate in 5 % H_2SO_4 (Hanessian's stain) or UV light. Ion-exchange chromatography was performed using ion-exchange resin (Amberlyst 131 wet, H^+ form) pre-equilibrated with ion-exchange buffer (25 mM NH_4CO_3 containing 2% isopropanol (iPrOH), 2CV). Reverse phase HPLC was performed on a system consisting of a Dionex P680 pump and a Dionex UVD170u detector unit. The column used was a 150 x 21.2 mm Phenomenex Luna C-18 column. Crude diphosphates were eluted under isocratic conditions with 10 % B for 20 min, a linear gradient to 60 % B over 25 min, then a linear gradient to 100% B over 5 min and lastly with 100% B for 10 min; solvent A: acetonitrile (HPLC grade), solvent B: 25 mM NH_4HCO_3, 2% iPrOH flow rate of 5.0 cm^3 min^{-1}, detection at 220 nm.

7.2.2. Preparation of (2*E*,6*E*)-FDP (30)[231,232]

Tris(tetra-n-butyl ammonium)hydrogen diphosphate (142)[231,232]

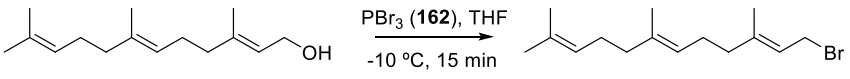


(**142**)

A solution of disodium dihydrogen diphosphate (**141**, 3.3 g, 15 mmol) in aqueous ammonium hydroxide (15 mL, 10% v/v) was passed through a column of Amberlist 131 cation- exchange resin (H^+ form). The free acid was eluted with deionised water (110 mL) and the resulting solution was immediately titrated with aqueous tetra-n-butyl ammonium hydroxide (40% w/w) to pH 7.3. The solution was then lyophilised, and the obtained residue was recrystallised from ethyl acetate. For this, the white solid obtained was dried by azeotropic removal of water with acetonitrile (100 mL) under reduced pressure and the resulting solid was warmed to 40 ºC in ethyl acetate before being filtered under vacuum to remove the insoluble contaminants. The filtrate was concentrated to half the volume and cooled to -20 ºC to give the final compound as white crystals (**142**, 5.5 g, 41%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 3.39 - 3.31 (24 H, m, 12 x NC*H*$_2$CH$_2$CH$_2$CH$_3$), 1.70 - 1.60 (24H, m, 12 x NCH$_2$C*H*$_2$CH$_2$CH$_3$), 1.49 - 1.40 (24 H, m, 12 x NCH$_2$CH$_2$C*H*$_2$CH$_3$), 0.97 (36 H, t, J = 7.3 Hz, 12 x NCH$_2$CH$_2$CH$_2$C*H*$_3$).
^{13}C NMR (126 MHz, $CDCl_3$) δ 58.86 (s, N*C*H$_2$CH$_2$CH$_2$CH$_3$), 24.25 (s, NCH$_2$*C*H$_2$CH$_2$CH$_3$), 19.82 (s, NCH$_2$CH$_2$*C*H$_2$CH$_3$), 13.88 (s, NCH$_2$CH$_2$CH$_2$*C*H$_3$).
^{31}P NMR (202 MHz, $CDCl_3$) δ –5.77 (s).
HRMS (ES^-, [M – H]): calculated for $[C_{16}H_{39}NO_7P_2]^-$: 418.2124, found: 418.2134.

(2*E*,6*E*)-1-bromo-3,7,11-trimethyldodeca-2,6,10-triene (153)[231]



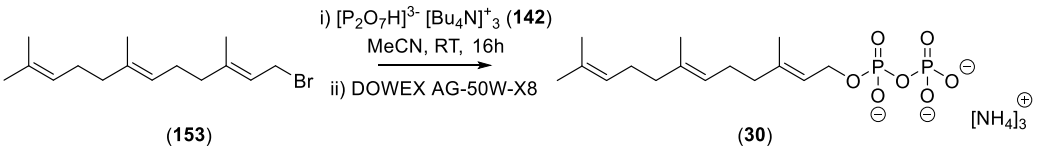
(**121**) (**153**)

To a solution of (*E*,*E*)-farnesol (**121**, 105 mg, 0.47 mmol, 1 eq.) in anhydrous THF (5 mL) was added dropwise a solution of PBr_3 (**162**, 22.5 µL, 0.24 mmol, 0.5 eq.) in THF (5 mL) at -10 °C under argon atmosphere. The reaction mixture was stirred for 15 min and quenched with *sat.* sodium bicarbonate solution (10 mL). Organic products were extracted with diethyl ether (3 x 10 mL). The organic extracts were washed with water (3 x 10 mL) and brine (3 x 10 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **153** as a pale-yellow oil (135 mg), which was used in the next reaction without further purification.

^{1}H NMR (300 MHz, $CDCl_3$) δ 5.53 (1 H, t, *J* = 8.5 Hz, C*H*CH_2Br), 5.14 – 5.03 (2 H, m, 2 x CC*H*CH_2), 4.02 (2 H, d, *J* = 8.5 Hz, C*H*$_2$Br), 2.13 – 1.93 (8 H, m, 4 x allylic C*H*$_2$), 1.73 (3 H, s, C*H*$_3$), 1.68 (3 H, s, C*H*$_3$), 1.60 (6 H, s, 2 x C*H*$_3$).

TLC (Hexane: EtOAc = 4: 1) R_f= 0.95.

(2*E*, 6*E*)-3,7,11-Trimethyl dodeca-2,6,10-trien-1-yl diphosphate (FDP, 30)[231]



(**153**) (**30**)

To a stirred solution of (2*E*,6*E*)-farnesyl bromide (**153**, 135 mg, 0.47 mmol, 1 eq.) in anhydrous acetonitrile (5 mL) was added a solution of tris(tetra-*n*-butyl ammonium)hydrogen diphosphate (**142,** 835 mg, 0.94 mmol, 2 eq.) in acetonitrile (10 mL) and stirred at room temperature for 16 hours. The reaction mixture was concentrated under reduced pressure and the resulting residue was dissolved in ion- exchange buffer solution (25 mM NH_4HCO_3, 2% iPrOH) (5 mL). The solution was passed through a NH_4^+- form cation exchange column using ion- exchange buffer solution as eluent. The eluted fractions were checked by TLC and compared with an authentic FDP standard (6:3:1, iPrOH: H_2O: NH_4OH). The fractions containing the desired compound were collected and lyophilised. The white powder was dissolved in ion- exchange buffer (4 mL) and purified by reverse phase HPLC. The appropriate samples were collected and lyophilised to yield (2*E*,6*E*)- farnesyl diphosphate **30** as a white powder (87 mg, 42%).

^{1}H NMR (400 MHz, D_2O) δ 5.45 (1H, t, *J* = 6.7 Hz, C*H*CH_2O), 5.24 – 5.13 (2H, m, 2 x CC*H*CH_2), 4.45 (2H, t, J_{H-P} = 6.5 Hz, CHC*H*$_2$O), 2.18 – 1.95 (8H, m, 4 x allylic C*H*$_2$), 1.70 (3H, s, C*H*$_3$), 1.67 (3H, s, C*H*$_3$), 1.60 (6H, s, 2 x C*H*$_3$).

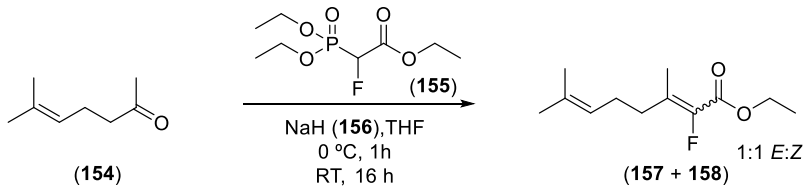
^{13}C NMR (101 MHz, D_2O) δ 142.73 (s, C_q), 136.65 (s, C_q), 133.50 (s, C_q), 124.42 (s, $CH_2(CH_3)CCHCH_2$), 124.25 (s, $(CH_3)_2CCHCH_2$), 119.87 (d, *J* = 8.8 Hz, *C*HCH_2O), 62.42 (d, *J* = 5.2 Hz, *C*H_2O), 38.80 (s, *C*H_2), 38.74 (s, *C*H_2), 25.72 (s, *C*H_2), 25.60 (s, *C*H_2), 24.81 (s, *C*H_3), 16.92 (s, *C*H_3), 15.60 (s, *C*H_3), 15.20 (s, *C*H_3).

^{31}P NMR (162 MHz, D_2O) δ –6.67 (d, *J* = 21.8 Hz), –10.42 (d, *J* = 22.2 Hz).

HRMS (ES^-, [M – H]): calculated for $[C_{15}H_{28}O_7P_2]^-$: 381.1232, found: 381.1238.

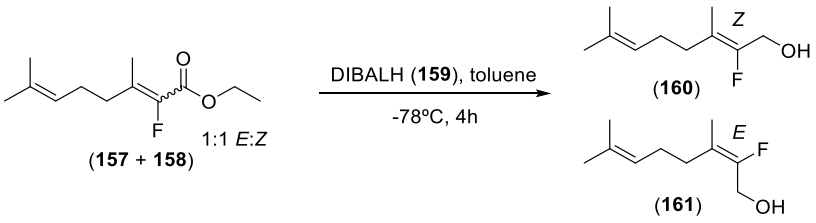
7.2.3. Synthesis of (2*E*,6*Z*)- and (2*E*, 6*E*)-6- fuorofarnesyl diphosphates [(2*E*,6*Z*)- and (2*E*,6*E*)-6F- FDP], (86)[122,123] and (187)

Ethyl (*E*) and (*Z*)-2-fluoro-3,7-dimethylocta-2,6-dienoate (157 and 158)[122,123]



A solution of triethyl 2-fluoro-2-phosphonoacetate (**155**, 14.46 mL, 71.31 mmol, 1.5 eq.) in anhydrous THF (20 mL) was added dropwise to a stirred suspension of sodium hydride (**156**, 2.85 g, 71.31 mmol, 1.5 eq.) in dry THF (20 mL) at 0 °C under argon atmosphere. The mixture was stirred for 1 hour at room temperature (colour changed from clear yellow to orange) and cooled to 0 °C prior to the addition of 6-methyl-5-hepten-2-one (**154**, 6 g, 47.54 mmol, 1 eq.) in anhydrous THF (20 mL). After addition, ice bath was removed, and the mixture was stirred at room temperature for 16 hours. Cold water was added slowly to the reaction mixture (40 mL) and organic products extracted with diethyl ether (3 x 100 mL). The organic extracts were washed with water (3 x 100 mL) and brine (3 x 100 mL), dried over anhydrous Na_2SO_4 and filtered. Evaporation under reduced pressure afforded the fluoro esters intermediates as a 1:1 mixture of cis and trans (**157** and **158**) isomers (as judged by GC and ^{19}F NMR spectroscopy).The mixture was passed through a silica plug (20% ethyl acetate: hexane as eluent) and reduced without further separation (9.81 g, 96%).

Ethyl (*Z*) and (*E*)-2-fluoro-3,7-dimethylocta-2,6-dien-1-ol (160 and 161)[122,123]



A hexane solution of DIBALH (**159**, 90 mL, 91.6 mmol, 1 M, 2 eq.) was added dropwise to the mixture of fluorinated esters **157** and **158** (9.81 g, 45.8 mmol, 1 eq.) in anhydrous toluene (20 mL) at -78 °C under argon atmosphere. The reaction was followed via TLC and Rochelle salts were added dropwise at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 100 mL). The organic extracts were washed with water (3 x 100 mL) and brine (3 x 100 mL), dried over anydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography (5% ethyl acetate: hexane as eluent) gave the 2*E* isomer **161** (3.6 g, 45.7%) followed by the 2*Z* isomer **160** (2.6 g, 33.1%) fluorinated alcohols (2-fluoro-nerol and 2-fluoro-geraniol, respectively) as clear oils.

(*Z*)-2-fluoro-3,7-dimethylocta-2,6-dien-1-ol (160)

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.11 (1H, t, *J* = 5.9 Hz CC*H*), 4.24 (2H , d, $J_{H–F}$ = 22.3 Hz, C*H*$_2$OH), 2.06–2.16 (4H, m, 2 x allylic C*H*$_2$), 1.69 (3H, s, C*H*$_3$), 1.67 (3H, d, $J_{H–F}$= 2.9 Hz, C*H*$_3$CCF), 1.61 (3H, s, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 152.85 (d, $J_{C–F}$ = 241.7 Hz, CH_3CCF), 132.26 (s, CH_3CCH), 123.80 (s, CH_3CCH), 116.24 (d, $J_{C–F}$ = 15.7 Hz, CH_3CCF), 58.25 (d, $J_{C–F}$ = 31.7 Hz, *C*H$_2$OH), 29.95 (d, $J_{C–F}$ = 6.7 Hz, *C*H$_2$(CH$_3$)CCFCH$_2$OH), 26.12 (d, *J* = 1.8 Hz, allylic *C*H$_2$), 25.82 (s, *C*H$_3$), 17.79 (s, *C*H$_3$), 15.53 (d, $J_{C–F}$ = 5.3 Hz, *C*H$_3$CCF).

^{19}F NMR (376 MHz, $CDCl_3$) δ –121.42.

HRMS (EI^+, [M – H_2O]): calculated for $[C_{10}H_{15}F]^+$; 154.1158, found 154.1160.

(*E*)-2-fluoro-3,7-dimethylocta-2,6-dien-1-ol (161)

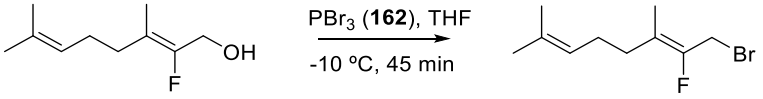
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.08 (1H, t, *J* = 7.0 Hz CH_3CCH), 4.19 (2H , d, $J_{H–F}$ = 22.8 Hz, C*H*$_2$OH), 2.14 – 2.02 (4H, m, 2 x allylic C*H*$_2$), 1.70 and 1.69 (6H, terminal C*H*$_3$ and C*H*$_3$CCF, overlap), 1.60 (3H, s, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 153.75 (d, $J_{C–F}$ = 243.9 Hz, CH_3CCF), 133.21 (s, CH_3CCH), 123.40 (s, CH_3CCH), 115.87 (d, $J_{C–F}$ = 14.0 Hz, CH_3CCF), 57.97 (d, $J_{C–F}$ = 31.6 Hz, *C*H$_2$OH), 31.93 (d, $J_{C–F}$ = 4.9 Hz, *C*H$_2$(CH$_3$)CCFCH$_2$OH), 26.52 (d, $J_{C–F}$ = 3.2 Hz, allylic *C*H$_2$), 25.80 (s, *C*H$_3$), 17.81 (s, *C*H$_3$), 13.67 (d, $J_{C–F}$ = 8.9 Hz, *C*H$_3$CCF).

^{19}F NMR (376 MHz, $CDCl_3$) δ –119.42.

HRMS (EI^+, [M – H_2O]): calculated for $[C_{10}H_{15}F]^+$; 154.1158, found 154.1157.

(*Z*)-1-bromo-2-fluoro-3,7-dimethylocta-2,6-diene (163)



(**160**) (**163**)

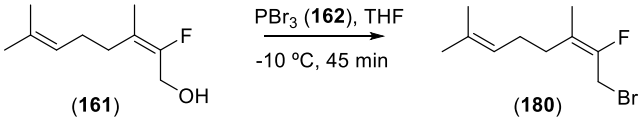
To a solution of fluorinated alcohol **160** (1.60 g, 9.3 mmol, 1 eq.) in anhydrous THF (10 mL) was added dropwise a solution of PBr_3 (**162**, 0.44 mL, 4.7 mmol, 0.5 eq.) in THF (5 mL) at -10 °C under argon atmosphere. The reaction mixture was stirred for 30 min and quenched with sat. *sat.* sodium bicarbonate sol. (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **163** as a pale-yellow oil (2.12 g, 97%), which was used in the next reaction without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.07 (1H, t, *J* = 6.7 Hz CC*H*), 4.06 (2H , d, $J_{H–F}$ = 23 Hz, CH_2Br), 2.17–2.06 (4H, m, 2 x allylic CH_2), 1.74 (3H, s, CH_3), 1.66 (3H, s, CH_3), 1.59 (3H, s, CH_3).

^{19}F NMR (376 MHz, $CDCl_3$) δ –115.87.

TLC (Hexane: EtOAc = 4: 1) R_f 0.95.

(*E*)-1-bromo-2-fluoro-3,7-dimethylocta-2,6-diene (180)



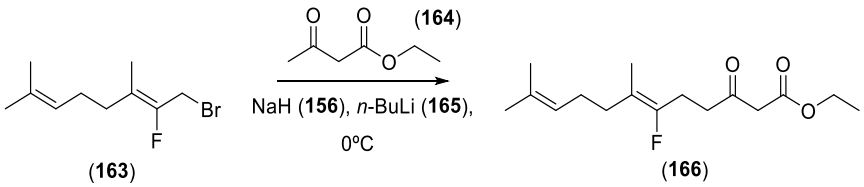
To a solution of fluorinated alcohol **161** (3 g, 17.4 mmol, 1 eq.) in anhydrous THF (20 mL) was added dropwise a solution of PBr_3 (**162**, 0.82 mL, 8.7 mmol, 0.5 eq.) in anhydrous THF (10 mL) at -10 °C under argon atmosphere. The reaction mixture was stirred for 30 min and quenched with *sat.* sodium bicarbonate sol. (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **180** as a pale-yellow oil (4.02 g, 98%), which was used in the next reaction without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.09 (1H, t, *J* = 7.1 Hz, (CH_3)CC*H*), 4.07 (2H , d, *J* = 23.0 Hz, CH_2Br), 2.17 – 2.01 (4H, m, 2 x allylic CH_2), 1.70 (3H, s, CH_3), 1.69 (3H, s, CH_3), 1.61 (3H, s, CH_3).

^{19}F NMR (376 MHz, $CDCl_3$) δ –114.56.

TLC (Hexane: EtOAc = 4: 1) R_f 0.95.

Ethyl (*Z*)-6-fluoro-7,11-dimethyl-3-oxododeca-6,10-dienoate (166)[242,243]

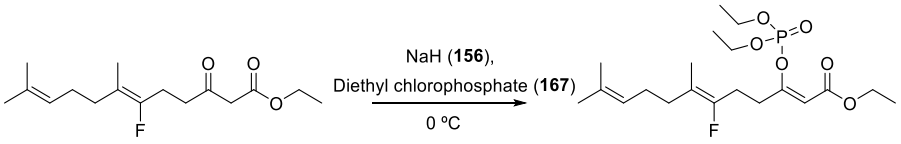


To a suspension of NaH (**156**, 1.2 g, 20.2 mmol, 60% in mineral oil, 3.1 eq) in anhydrous THF (50 mL) was added ethyl acetoacetate (**164**, 2.48 mL, 19.6 mmol, 3 eq) dropwise at 0 °C under argon atmosphere. After 10 min, *n*-BuLi (**165**, 8 mL, 20.2 mmol, 2.5 M in hexane, 3.1 eq) was added slowly. The mixture was stirred for further 30 min before a solution of bromide **163** (1.5 g, 6.52 mmol, 1 eq) in THF (20 mL) was added dropwise. After 30 min at room temperature TLC analysis showed that the reaction completion and was quenched with addition slow addition of HCl (10%, 20 mL). The organic products were extracted into diethyl ether (3 x 50 mL). The combined ethereal extracts were washed with water

(3 x 50 mL) and brine (3 x 50 mL) and dried over anhydrous sodium sulphate. Removal of the solvent gave the crude product, and purification by silica column flash chromatography afforded ester **166** as a yellow oil (1.14 g, 62%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.14 – 5.02 (1H, m, C*H*CH_2CH_2), 4.20 (2H, q, *J* = 7.1 Hz, OC*H*$_2$$CH_3$), 3.45 (2H, s, COC*H*$_2$$CO_2$), 2.74 (2H, t, *J* = 7.2 Hz, $CFCH_2CH_2$), 2.53 (2H, dt, *J* = 22.2, 7.2 Hz, $CFCH_2CH_2$), 2.12 – 1.95 (4H, m, 2 x allylic C*H*$_2$), 1.68 (3H, s, C*H*$_3$), 1.59 and 1.57 (6H, m , C*H*$_3$CCF and C*H*$_3$CCH), 1.28 (3H, t, *J* = 7.1 Hz, OCH_2CH_3).
^{13}C NMR (101 MHz, $CDCl_3$) δ 201.69 (s, $CH_3C(O)CH$), 167.03 (s, $CH_3CH_2OC(O)$), 152.40 (d, $J_{C–F}$ = 241.2 Hz, CH_3CCFCH_2), 131.79 (s, CH_3CCH), 123.90 (s, CH_3CCH), 112.50 (d, $J_{C–F}$ = 16.6 Hz, CH_3CCFCH_2), 61.42 (s, CH_2CH_3), 49.37 (s, $COCH_2CO_2$), 39.64 (s, CH_2CH_2CCF), 29.66 and 26.14 (d, $J_{C–F}$ = 7.2 Hz, CH_2CH_2CCF), 25.68 (s, CH_3), 22.65 (d, $J_{C–F}$ = 29.9 Hz $CFCH_2CH_2$), 17.62 (s, CH_3), 15.64 (d, $J_{C–F}$ = 5.8 Hz CH_3CCF), 14.07 (s, CH_2CH_3).
^{19}F NMR (376 MHz, $CDCl_3$) δ –114.78.

Ethyl (*Z*)-3-((diethoxyphosphoryl)oxy)-6-fluoro-7,11-dimethyldodeca-6,10-dienoate (168)[242,243]



(**166**) (**168**)

A solution of β-ketoester **166** (0.48 g, 1.7 mmol, 1 eq.) in diethyl ether (10 mL) was added dropwise to a solution of NaH (**156,** 0.2 g, 5 mmol, 60% in mineral oil, 3 eq.) in diethyl ether (10 mL) at 0 °C under argon. The mixture was then stirred at 0 °C for 1 hour and diethyl chlorophosphite (**167**, 0.48 mL, 3.35 mmol, 2 eq.) was added slowly. The resulting solution was stirred for an additional 2 hours at 0 °C. The reaction was quenched by addition of *sat.* ammonium chloride solution (10 mL). The organic products were extracted with diethyl ether (3 x 50 mL) and washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate and filtered. Concentration under reduced pressure gave the crude product **168** as a pale-yellow oil, which was used without further purification. (0.71 g).

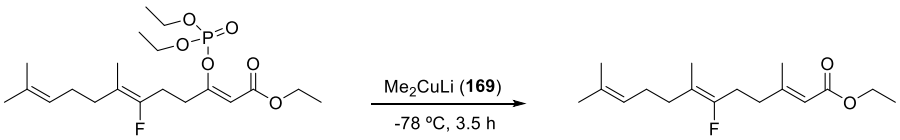
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.37 (1H, s, CC*H*CO_2), 5.13 – 5.05 (1H, m, CC*H*CH_2), 4.26 (4H, quintet, *J* = 7.0, 2 x POC*H*$_2$$CH_3$), 4.15 (2H, q, *J* = 7.0 $CO_2CH_2CH_3$), 2.74 – 2.41 (4H, m, $CFCH_2CH_2$), 2.11 – 1.95 (4H, m,2 x allylic C*H*$_2$), 1.67 (3H, s, C*H*$_3$CCH), 1.59 (3H, d, *J* = 1.9, C*H*$_3$CCF), 1.58 (3H, s, C*H*$_3$CCH), 1.36 (6H, t, *J* = 6.8 Hz, 2x $POCH_2CH_3$), 1.26 (3H,t , *J* = 6.9 Hz, $CO_2CH_2CH_3$).
^{13}C NMR (101 MHz, $CDCl_3$) δ 163.60 (s, C_q), 160.11(d, *J* = 7.2 Hz, C_q), 152.02 (d, $J_{C–F}$ = 241.8, CH3*C*CF), 131.80 (s, C_q), 123.92 (s, $(CH_3)_2CHCH_2$), 113.26 (d, $J_{C–F}$ = 16.5 Hz, CH_3CCFCH_2), 106.12

(d, J_{C-P} = 7.6 Hz, CO*C*HCO$_2$), 64.87 and 64.81(s, 2 x PO*C*H$_2$CH$_3$), 59.92 (s, CO$_2$*C*H$_2$CH$_3$), 32.49 (s, CFCH$_2$*C*H$_2$), 29.75 (d, J_{C-P} = 7.2 Hz, CCHCH$_2$*C*H$_2$), 26.24 (d, J_{C-F} = 1.7 Hz, CCH*C*H$_2$CH$_2$), 26.07 (d, J_{C-F} = 29.9, CF*C*H$_2$CH$_2$), 25.67 (s, *C*H$_3$CCH), 17.62 (s, *C*H$_3$CCH), 16.11 (s, POCH$_2$*C*H$_3$), 16.04 (s, POCH$_2$*C*H$_3$), 15.49 (d, J_{C-F} = 5.8 Hz, *C*H$_3$CCF), 14.22 (s, CO$_2$CH$_2$*C*H$_3$).

^{19}F NMR (376 MHz, CDCl$_3$) δ –114.35.

^{31}P NMR (162 MHz, CDCl$_3$) δ –8.61.

(2*E*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trienoate (170)[242,243]



(**168**) (**170**)

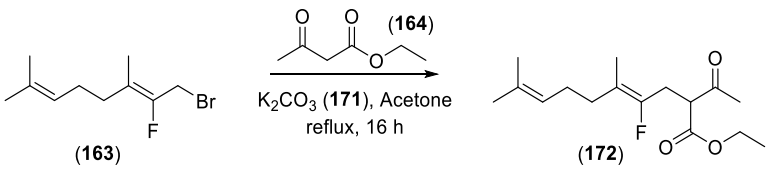
A stirred suspension of CuCN (0.305 g, 3.4 mmol, 2 eq.) in anhydrous diethyl ether (10 mL) under argon was cooled to 0 °C and MeLi (4.25 mL, 6.8 mmol, 1.6 M in diethyl ether, 4 eq.) was added dropwise. The mixture was stirred for 30 min at 0 °C. This solution was then cooled to -78 °C and a solution of **168** (0.71 g, 1.7 mmol, 1 eq.) in diethyl ether (10 mL) was added slowly over 5 min. The resulting solution was stirred for 3 h at -78 °C, iodomethane was (0.27 mL, 4.25 mmol, 2.5 eq.) added and the mixture was stirred for further 15 min at -78 °C. The yellow-orange mixture was put in an ice-bath and quenched with *sat.* ammonium chloride solution and concentrated ammonium solution (1:1 15 mL). The organic products were extracted with diethyl ether (3 x 50 mL) and washed with water (3 x 50 mL) and brine (3 x 50 mL). The organic extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification of the crude by silica gel flash chromatography (10% ethyl acetate in hexane as eluent) gave **170** as a clear oil (24 mg, 5%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 5.68 (1H, s, CC*H*CO$_2$), 5.13 – 5.06 (1H, m, CC*H*CH$_2$), 4.14 (2H, q, *J*= 7.1, CO$_2$C*H*$_2$CH$_3$), 2.45 –2.28 (4H, m, (CFC*H*$_2$C*H*$_2$), 2.17 (3H, d, *J*= 1.1 Hz C*H*$_3$CCHCO$_2$), 2.08–2.01 (4H, m, (CH$_3$)$_2$CCHC*H*$_2$C*H*$_2$), 1.68 (3H, s, C*H*$_3$CCHCH$_2$CH$_2$), 1.60 (3H, s, C*H*$_3$CCHCH$_2$CH$_2$), 1.56 (3H, t, *J* = 2.6 Hz, C*H*$_3$CCF), 1.27 (3H,t , *J* = 7.0 Hz, CO$_2$CH$_2$C*H*$_3$).

^{13}C NMR (101 MHz, CDCl$_3$) δ 166.73 (s, C$_q$), 158.46 (s, C$_q$), 152.14 (d, J_{C-F} = 240, CH$_3$C*C*F), 131.78 (s, C$_q$), 123.99 (s, ((CH$_3$)$_2$C*C*HCH$_2$), 116.21 (s, C*C*HCO$_2$), 112.31 (d, J_{C-F} = 17.1 Hz, CH$_3$*C*CFCH$_2$), 59.55 (s, O*C*H$_2$CH$_3$), 37.77 (s, CFCH$_2$*C*H$_2$), 29.71 (d, J_{C-F} = 5.5 Hz, ((CH$_3$)$_2$CCHCH$_2$*C*H$_2$), 27.00 (d, J_{C-F} = 30.2 Hz, CF*C*H$_2$CH$_2$), 26.25 (d, J_{C-F} = 2.5, ((CH$_3$)$_2$CCH*C*H$_2$CH$_3$), 25.70 (s, *C*H$_3$CCHCH$_2$), 18.70 (s, *C*H$_3$CCHCO$_2$), 17.65 (s, *C*H$_3$CCHCH$_2$), 15.54 (d, J_{C-F} = 6.1 Hz, *C*H$_3$CCF), 14.33 (s, OCH$_2$*C*H$_3$).

^{19}F NMR (376 MHz, CDCl$_3$) δ –113.48 (s).

Ethyl (Z)-2-acetyl-4-fluoro-5,9-dimethyldeca-4,8-dienoate (172)[245]



Ethyl acetoacetate (**164**, 1.11 mL, 8.7 mmol, 1.5 eq.) and potassium carbonate (**171**, 1.6 g, 11.6 mmol, 2 eq.) were added to a solution of bromide (**163**, 1.36 g, 5.8 mmol, 1 eq.) in dry acetone (30 mL) and refluxed for 16 hours. Volatile solvents were removed under reduced pressure and the resulting organic mixture extracted into diethyl ether (3 x 50 mL). Organic extracts were washed with water (3 x 50 mL), brine (3 x 50 mL) dried over anhydrous Na_2SO_4 and filtered. Products were concentrated under reduced pressure and purified by silica gel flash chromatography (20% ethyl acetate in hexane as eluent) to give **172** as a colourless oil (1.35 g, 84%).

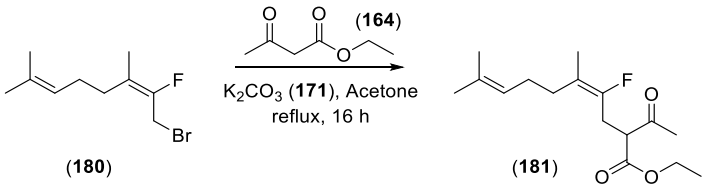
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.05 (1H, t, *J* = 6.8 Hz, CH_3CCH), 4.18 (2H, q, *J* = 7.1 Hz, CH_3CH_2O), 3.76 (1H, t, *J*= 7.4 Hz, $COCHCO$), 2.78 (2H, dd, *J* = 21.8, 7.4 Hz, $CCFCH_2CH$), 2.25 (3H, s, $COCH_3$), 2.12 – 1.91 (4H, m, 2 x allylic CH_2), 1.66 (3H, s, CH_3), 1.58 (6H, s, 2 x CH_3), 1.26 (3H, t, *J*= 7.1 Hz, OCH_2CH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 202.34 (s, $CH_3C(O)CH$), 169.06 (s, $CH_3CH_2OC(O)$), 151.54 (d, *J* = 396.9 Hz, $CH_3CC(F)CH_2$), 132.04 (s, CH_3CCH), 123.91 (s, CH_3CCH), 114.46 (d, *J* = 16.3 Hz, CH_3CCFCH_2), 61.75 (s, CH_3CH_2O), 56.43 (d, , *J* = 1.3 Hz, $COCHCO$), 30.10 (d, *J* = 40.0 Hz, $CH_2CH_3CCFCH_2$), 29.84 (s, CH_3CO), 27.42 (d, *J* = 29.0 Hz, CH_3CCFCH_2), 26.23 (d, *J* = 1.9 Hz, CH_3CCHCH_2), 25.82 (s, CH_3), 17.77 (s, CH_3), 15.64 (d, *J* = 5.6 Hz OCH_2CH_3), 14.20 (s, CH_3).

^{19}F NMR (376 MHz, $CDCl_3$) δ –114.69.

HRMS (EI^+, [M]): calculated for $[C_{16}H_{25}O_3F]^+$; 284.1788, found 284.1792.

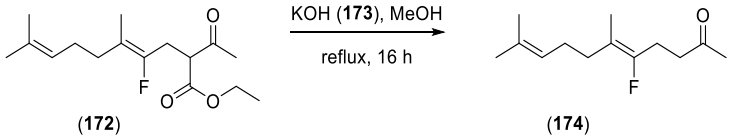
Ethyl (*E*)-2-acetyl-4-fluoro-5,9-dimethyldeca-4,8-dienoate (181)[245]



Ethyl acetoacetate (**164**, 4.1 mL, 32.2 mmol, 1.5 eq.) and potassium carbonate (**171**, 4.4 g, 32.2 mmol, 2 eq.) were added to a solution of bromide (**180**, 3.78 g, 16.1 mmol, 1 eq.) in dry acetone (30 mL and refluxed for 16 hours. Volatile solvents were removed under reduced pressure and the resulting organic mixture extracted into diethyl ether (3 x 100 mL). Organic extracts were washed with water (3 x 50 mL), brine (3 x 50 mL) dried over anhydrous Na_2SO_4 and filtered. Products were concentrated under reduced pressure and purified by silica gel flash chromatography (20% ethyl acetate in hexane as eluent) to give **181** as a colourless oil (3.24 g, 71%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.07 (1H, t, *J* = 7.0 Hz, $(CH_3)CCH$), 4.19 (2H, q, *J* = 7.1 Hz, CH_3CH_2O), 3.74 (1H, t, *J*= 7.3 Hz, $COCHCO$), 2.80 (2H, dd, *J* = 21.0, 6.9 Hz, $CCFCH_2CH$), 2.26 (3H, s, $COCH_3$), 2.09 – 1.91 (4H, m, 2 x allylic CH_2), 1.68 (3H, s, CH_3), 1.61 (3H, d, *J* = 3.4 Hz CH_3CCF), 1.60 (3H, s, CH_3), 1.26 (3H, t, *J*= 7.1 Hz,CH_3CH_2O).
^{13}C NMR (101 MHz, $CDCl_3$) δ 202.29 (s, $CH_3C(O)CH$), 169.10 (s, $CH_3CH_2OC(O)$), 151.29 (d, *J* = 241.7 Hz, $CH_3CC(F)CH_2$), 132.43 (s, $(CH_3)CCH$), 123.61 (s, $(CH_3)CCH$), 114.13 (d, *J* = 14.4 Hz, $(CH_3)CC(F)CH_2$), 61.74 (s, CH_3CH_2O), 56.60 (d, *J*= 1.1 Hz, $C(O)CHC(O)$), 31.92 (d, *J* = 5.4 Hz, $CH_2(CH_3)CC(F)CH_2$), 29.75 ($CH_3C(O)$), 27.16 (d, *J* = 28.6 Hz, $(CH_3)CC(F)CH_2$), 26.51 (d, *J* = 3.0 Hz, $(CH_3)CCHCH_2$), 25.85 (s, CH_3), 17.78 (s, CH_3), 14.18 (s, OCH_2CH_3), 13.48 (d, *J* = 9.6 Hz, CH_3CCF).
^{19}F NMR (376 MHz, $CDCl_3$) δ –112.94.
HRMS (EI^+, [M]): calculated for $[C_{16}H_{25}O_3F]^+$; 284.1788, found 284.1765.

(*Z*)-5-fluoro-6,10-dimethylundeca-5,9-dien-2-one (174)[245]

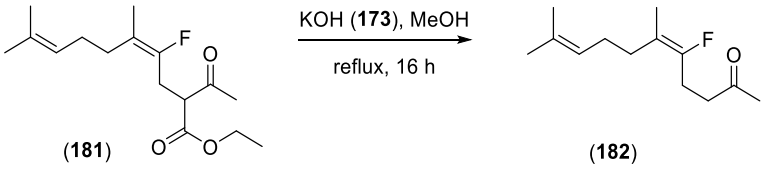


Potassium hydroxide (**173**, 0.336 g, 5.8 mmol, 1.5 eq.) was added to a solution of **172** (1.1 g, 4 mmol, 1 eq.) in methanol (20 mL) and refluxed for 16 hours. The reaction mixture was cooled to room temperature and acidified with 1M HCl solution until pH 2-3. Organic products were extracted with diethyl ether (3 x 50 mL) and washed with water (3 x 20 mL) and brine (3 x 20 mL). Organic layers

were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification *via* silica column flash chromatography gave the desired ketone **174** as a pale-yellow oil (0.67 g, 79%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.08 (1H, t, *J* = 6.5 Hz, CH_3CCH), 2.65 – 2.59 (2H, m, CH_2CO), 2.55 – 2.43 (2H, m, CH_2CH_2CO), 2.16 (3H, s, $CH_2C(O)CH_3$), 2.08 – 2.01 (4H, m, 2 x allylic CH_2), 1.67 (3H, s, CH_3), 1.59 (3H, s, CH_3), 1.58 (3H, d, *J* = 3.5 Hz CH_3CCF).
^{13}C NMR (101 MHz, $CDCl_3$) δ 207.71 (s, CH_3CO), 152.90 (d, *J* = 241.1 Hz, CH_3CCF), 131.78 (CH_3CCH), 123.97 (CH_3CCH), 112.17 (d, *J* = 16.8 Hz, CH_3CCF), 40.28 (s, CH_2CH_2CO), 30.04 (s, CH_3CO), 29.68 (d, *J* = 7.3 Hz, CH_2CH_3CCF), 26.19 (d, *J* = 1.6 Hz, allylic CH_2), 25.71 (s, CH_3), 22.88 (d, *J* = 29.7 Hz, $((CH_3)CCFCH_2)$, 17.65 (s, CH_3), 15.47 (d, *J* = 5.9 Hz, $(CH_3)CCF$).
^{19}F NMR (376 MHz, $CDCl_3$) δ –114.58.
HRMS (EI^+, [M]): calculated for $[C_{13}H_{21}OF]^+$; 212.1576, found 212.1578.

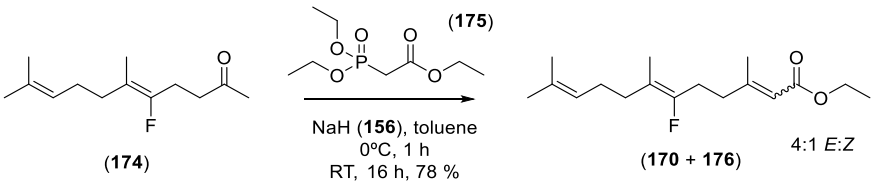
(*E*)-5-fluoro-6,10-dimethylundeca-5,9-dien-2-one (182)[245]



Potassium hydroxide (**173**, 1.3 g, 24 mmol, 3 eq.) was added to a solution of **181** (2.3 g, 8 mmol, 1 eq.) in methanol (40 mL) and refluxed for 16 hours. The reaction mixture was cooled to room temperature and acidified with 1M HCl solution. Organic products were extracted with diethyl ether (3 x 100 mL) and washed with water (3 x 50 mL) and brine (3 x 50 mL). Organic layers were dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification *via* silica column flash chromatography gave the desired ketone **182** as a pale-yellow oil (1.04 g, 62%).

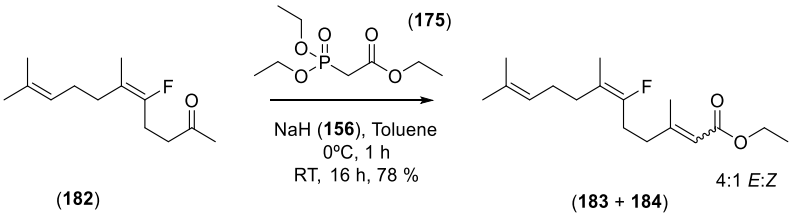
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.07 (1H, t, *J* = 7.0 Hz, $(CH_3)CCH$), 2.62 (2H, dd, *J* = 9.6, 5.0 Hz, $CFCH_2CH_2C(O)$), 2.56 – 2.42 (2H, m, CH_2CH_2CO), 2.15 (3H, s, $CH_2C(O)CH_3$), 2.09 – 1.93 (4H, m, 2 x allylic CH_2), 1.68 (3H, s, CH_3), 1.61 (3H, d, *J* = 3.2 Hz, CH_3CCF), 1.59 (3H, s, CH_3).
^{13}C NMR (101 MHz, $CDCl_3$) δ 207.79 (s, CH_3CO), 153.60 (d, *J* = 242.5 Hz, $(CH_3)CCF$), 132.33 (s, $(CH_3)CCH$), 123.73 (s, $(CH_3)CCH$), 112.06 (d, *J* = 15.0 Hz, $(CH_3)CCF$), 40.68 (CH_2CH_2CO), 32.04 (d, *J* = 5.7 Hz, $CH_2(CH_3)CCF$), 30.11 (s, CH_3CO), 26.57 (d, *J* = 3.1 Hz, allylic CH_2), 25.88 (s, CH_3), 22.79 (d, *J* = 29.6 Hz, $((CH_3)CCFCH_2)$, 17.78 (s, CH_3), 13.45 (d, *J* = 9.7 Hz, $(CH_3)CCF$).
^{19}F NMR (376 MHz, $CDCl_3$) δ –113.02.
HRMS (EI^+, [M]): calculated for $[C_{13}H_{21}OF]^+$; 212.1576, found 212.1573.

(2*E*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trienoate (170)[246]



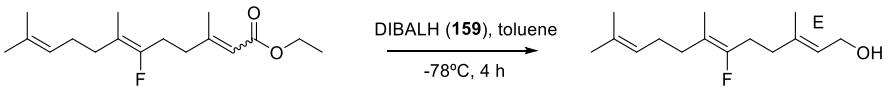
A suspension of sodium hydride (**156**, 70 mg, 1.3 mmol, 60% dispersion in mineral oil, 1.1 eq.) in anhydrous THF (10 mL) was cooled to 0 ºC. Ethyl 2-(diethoxyphosphoryl)acetate (**175**, 260 µL ,1.3 mmol, 1.1 eq.) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and **174** (250 mg, 1.2 mmol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched with water (20 mL) and reaction products extracted with diethyl ether (3 x 50 mL). Combined organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL). Extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave a mixture of the fluorinated esters **170** and **176** (4: 1 of trans and cis isomers, as judged by GC and ^{19}F NMR spectroscopy) as a clear oil (263 mg, 78%) that were used in the next reaction without separation.

Ethyl (2*E*,6*E*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trienoate (183)[246]



A suspension of sodium hydride (**156**, 0.2 g, 4.62 mmol, 1,1 eq.) in anhydrous THF (30 mL) was cooled to 0 ºC. Ethyl 2-(diethoxyphosphoryl)acetate (**175**, 1.03 g, 4.62 mmol, 1.1 eq.) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and **182** (0.9 g, 4.2 mmol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched with water (50 mL) and reaction products extracted with diethyl ether (3 x 100 mL). Combined organic extracts were washed with water (3 x 100 mL) and brine (3 x 100 mL). Extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave a mixture of fluorinated esters **183** and **184** (4: 1 mixture of trans and cis isomers, as judged by GC and ^{19}F NMR spectroscopy) as a clear oil (1.04 g, 88%) that were used in the next reaction without separation.

(2*E*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (177)[122,123]



(**170** + **176**) (**177**)

A hexane solution of DIBALH (**159**, 1.15 mL, 1.14 mmol, 1 M, 2 eq.) was added dropwise to a mixture of ester **170** and **176** (160 mg, 0.57 mmol, 1 eq.) fluorinated esters in anhydrous toluene (20 mL) at -78 ºC under argon. Reaction was followed *via* TLC and Rochelle salts were added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 50 mL). The organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 5% ethyl acetate/hexane as eluent separated the 2*E*,6*Z*-fluorinated alcohol **177** (77 mg, 56%) and the 2*Z*,6*Z*-fluorinated alcohol **178** (16 mg, 12%) as clear oils.

(2*E*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (177)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.43 (1H, t, *J* = 6.7 Hz, CC*H*CH_2OH), 5.13 – 5.07 (1H, m, $(CH_3)CCH$), 4.15 (2H, d, *J* = 6.9, $CHCH_2OH$), 2.39 – 2.28 (2H, m, $CFCH_2CH_2$), 2.19 (2H,t , *J* = 7.5Hz, $CFCH_2CH_2$), 2.10 – 2.01 (4H, m, allylic CH_2), 1.70 (3H, s, CH_3), 1.68 (3H, s, CH_3), 1.60 (3H, s, CH_3CCHCH_2OH), 1.56 (3H, d, *J* = 1.8, CH_3CF).

^{13}C NMR (126 MHz, $CDCl_3$) δ 154.04 (d, *J* = 242.4 Hz, CH_3CCF), 138.96 & 131.87 (s, 2 x CH_3CCH), 124.20 ($(CH_3)_2CCH$), 124.10 (s, $CCHCH_2O$), 111.72 (d, *J* = 17.0 Hz, CH_3CCF), 59.50 (s, $CHCH_2OH$), 36.58 (s, $CFCH_2CH_2$), 29.82 (d, *J*= 7.4 Hz, $(CH_3)_2CCHCH_2CH_2$), 27.51 (d, *J* = 29.7 Hz, $(CH_3)CCFCH_2$), 26.42 (d, *J* = 1.5 Hz, $(CH_3)_2CCHCH_2CH_2$), 25.84 (s, $CH_3CCHCH_2CH_2$), 17.79 (s, CH_3CCHCH_2OH), 16.36 (s, $CH_3CCHCH_2CH_2$), 15.66 (d, *J* = 6.1 Hz, $(CH_3)CCF$).

^{19}F NMR (471 MHz, $CDCl_3$) δ –112.79 (t, *J* = 23.0 Hz).

HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{15}H_{23}F]^+$; 222.1784, found 222.1780.

(2*Z*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (178) - not used

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.47 (1H, t, *J* = 7.2 Hz, CC*H*CH_2OH), 5.13 – 5.07 (1H, m, $(CH_3)CCH$), 4.11 (2H, d, *J* = 7.2, $CHCH_2OH$), 2.39 – 2.22 (4H, m, $CFCH_2CH_2$), 2.07 – 2.01 (4H, m, allylic CH_2), 1.76 (3H, s, CH_3), 1.68 (3H, s, CH_3), 1.60 (3H, s, CH_3CCHCH_2OH), 1.56 (3H, d, *J* = 2.6, CH_3CCF).

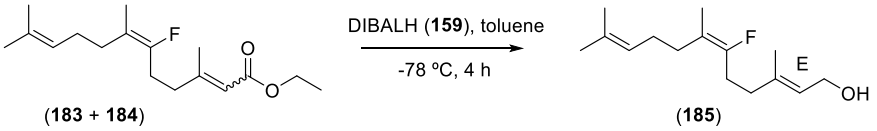
^{13}C NMR (126 MHz, $CDCl_3$) δ 153.87 (d, *J* = 241.9 Hz, CH_3CCF), 138.89 & 131.94 (s, 2 x CH_3CCH), 125.36 ($(CH_3)_2CCH$), 124.12 (s, $CCHCH_2O$), 112.25 (d, *J* = 16.9 Hz, CH_3CCF), 59.13 (s, $CHCH_2OH$), 29.80 (d, *J*= 7.5 Hz, $(CH_3)_2CCHCH_2CH_2$), 29.02 (d, *J*= 0.9 Hz, $CFCH_2CH_2$), 27.48 (d, *J* = 29.8 Hz,

(CH_3)CCFCH_2), 26.37 (d, J = 1.6 Hz, allylic CH_2), 25.84 (s, CH_3CCHCH$_2$CH$_2$),23.47 (s, CH_3CCHCH$_2$CH$_2$), 17.79 (s, CH_3CCHCH$_2$OH), 15.67 (d, J = 6.0 Hz, CH_3CCF).

^{19}F NMR (471 MHz, $CDCl_3$) δ –112.97.

HRMS (EI$^+$, [M – H_2O]): calculated for [$C_{15}H_{23}F$]$^+$; 222.1784, found 222.1786.

(2*E*,6*E*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (185)[122,123]



(**183** + **184**) (**185**)

A hexane solution of DIBALH (**159**, 4.6 mL, 4.6 mmol, 1 M, 2 eq.) was added dropwise to a mixture of esters **185** and **186** (0. 66 g, 2.3 mmol, 1eq) fluorinated esters in anhydrous toluene (50 mL) at -78 ^{o}C under argon. Reaction was followed *via* TLC and Rochelle salts were slowly added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 100 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50mL), dried over sodium anhydrous sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 5% ethyl acetate/hexane as eluent gave 2*E*,6*E*-fluorinated alcohol **185** (284 mg, 51%) and 2*Z*,6*E*-fluorinated alcohol **186** (49 mg, 9%) as clear oils.

(2*E*,6*E*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (185)

^{1}H NMR (400 MHz, D_2O) δ 5.46 – 5.40 (1H, m, (CH_3)CC*H*CH$_2$OH), 5.11 – 5.05 (1H, m, (CH_3)$_2$CC*H*CH$_2$), 4.15 (2H, d, J = 6.9, CHCH_2OH), 2.39 – 2.27 (2H, m, (CH_3)CCFCH_2), 2.19 (2H,dd , J = 9.0, 6.6 Hz, CFCH$_2$CH_2), 2.05 (2H, dd, J = 14.9, 7.3 Hz, allylic CH_2), 1.97– 1.89 (2H, m , (CH_3)$_2$CCHCH$_2$CH_2), 1.69 (3H, s, C*H*3), 1.68 (3H, s, CH_3), 1.62 (3H, d, J = 3.3 Hz, (CH_3)CCF), 1.60 (3H, s, CH_3CCHCH$_2$OH).

^{13}C NMR (101 MHz, D_2O) δ 154.62 (d, J = 243.6 Hz, CH_3C*C*F), 138.97 & 132.26 (s, 2 x CH_3*C*CH), 124.04 (s, C*C*HCH$_2$OH), 123.80 (s, (CH_3)$_2$C*C*H), 111.49 (d, J = 15.3 Hz, CH_3*C*CF), 59.47 (s, CH*C*H$_2$OH), 36.88 (d, J = 0.7 Hz CFCH$_2$*C*H$_2$), 32.14 (d, J= 5.8 Hz, (CH_3)$_2$CCHCH$_2$*C*H$_2$), 27.41 (d, J = 29.6 Hz, (CH_3)CCF*C*H2), 26.72 (d, J = 3.1 Hz, allylic *C*H$_2$), 25.86 (s, *C*H$_3$), 17.78 (s, *C*H$_3$CCHCH$_2$OH), 16.37 (s, *C*H$_3$), 13.35 (d, J = 9.7 Hz, (*C*H$_3$)CCF).

^{19}F NMR (376 MHz, D_2O) δ –111.47.

HRMS (EI$^+$, [M – H_2O]): calculated for [$C_{15}H_{23}F$]$^+$; 222.1784, found 222.1784.

(2*Z*,6*E*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (186) - not used

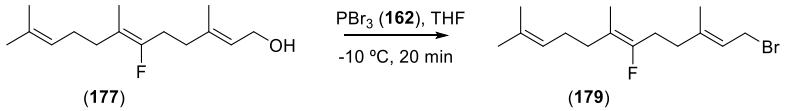
1H NMR (500 MHz, $CDCl_3$) δ 5.46 (1H, t, *J* = 6.9 Hz, CC*H*CH_2O), 5.07 (1H, t, *J* = 7.0 Hz, (CH_3)CC*H*), 4.10 (2H, d, *J* = 7.1, CHC*H*$_2$OH), 2.40 – 2.14 (4H, m, 2 x allylic C*H*$_2$), 2.04 (2H, dd, *J* = 14.3, 7.2 Hz, (CH_3)CCFC*H*$_2$), 1.96 – 1.87 (2H, m, allylic C*H*$_2$), 1.75 (3H, s, C*H*$_3$), 1.68 (3H, s, C*H*$_3$), 1.62 (3H, d, *J* = 3.2 Hz, (C*H*$_3$)CCF), 1.60 (3H, s, C*H*$_3$).

^{13}C NMR (126 MHz, $CDCl_3$) δ 154.22 (d, *J* = 243.1 Hz, (CH_3)C*C*F), 138.65 & 132.16 ((2 x (CH_3)*C*CH), 125.28 (C*C*HCH_2O), 123.61 ((CH_3)C*C*H), 112.00 (d, *J* = 15.1 Hz, (CH_3)*C*CF), 58.97 (CH*C*H_2O), 32.02 (d, *J* = 5.7 Hz, *C*H_2(CH_3)CCF), 29.14 (allylic *C*H_2), 27.20 (d, *J* = 29.7 Hz, (CH_3)CCF*C*H_2), 26.65 (d, *J* = 2.9 Hz, allylic *C*H_2), 25.72, 23.41 & 17.65 (3 x *C*H_3), 13.17 (d, *J* = 9.8 Hz, (*C*H_3)CCF).

^{19}F NMR (471 MHz, $CDCl_3$) δ –111.58. (t, *J* = 23.1 Hz).

HRMS (EI$^+$, [M – H_2O]): calculated for [$C_{15}H_{23}F$]$^+$; 222.1784, found 222.1785.

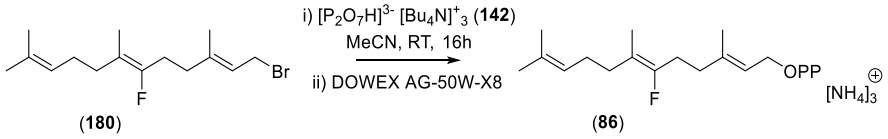
(2*E*,6*Z*)-1-bromo-6-fluoro-3,7,11-trimethyldodeca-2,6,10-triene (179)[122,123]



To a solution of fluorinated alcohol **177** (70 mg, 0.3 mmol, 1 eq.) in anhydrous THF (5 mL) was added dropwise a solution of PBr_3 (**162**, 14 μL, 0.15 mmol, 0.5 eq.) in anhydrous THF (5 mL) at -10 °C under argon. The reaction mixture was stirred for 30 min and quenched with sat. sodium bicarbonate sol. (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure to give crude **179** as a pale-yellow oil (assumed 100%), which was used in the next reaction without further purification.

TLC (Hexane: EtOAc = 4: 1) R_f 0.95.

(2*E*,6*Z*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate ([2*E*,6*Z*]-6F-FDP, 86)
[231,232]

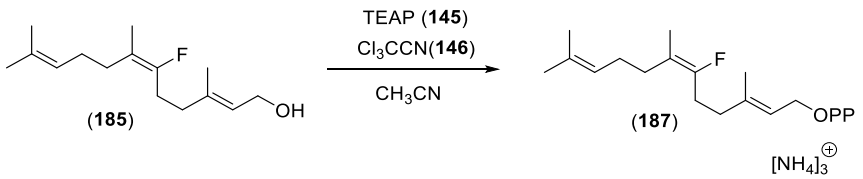


To a stirred solution of bromide **179** (91 mg, 0.3 mmol, 1 eq.) in anhydrous acetonitrile (5 mL) under argon was added a solution of tris tetra butyl ammonium diphosphate (**142**, 540 mg, 0.6 mmol, 2 eq.) in THF (10 mL). The reaction mixture was stirred for 16 h and then the solvent was removed under pressure. The resulting residue was dissolved in buffer solution (25 mM NH_4HCO_3, 2% iPrOH) (10 mL). The solution was slowly passed through an ion exchange column containing Amberlyst 131 (wet H^+ form) mesh cation exchange resin pre-equilibrated with ion-exchange buffer (25 mM NH_4HCO_3, 2% isopropanol). The appropriate fractions were collected and lyophilised to yield the desired

diphosphate as a white powder. The powder was dissolved in buffer solution (25 mM NH_4HCO_3, 2% iPrOH) (10 mL) and purified by reverse phase HPLC (Solvent A: acetonitrile, solvent B: 2% iPrOH). The appropriate fractions were collected and lyophilised to give **86** as a white powder. (62 mg, 46%).

^{1}H NMR (500 MHz, D_2O) δ 5.46 (1H, t, *J* = 7.3 Hz, CC*H*CH_2O), 5.18 (1H, t, *J* = 7.0 Hz, $(CH_3)_2$CC*H*CH_2), 4.47 (2H, t, *J* = 5.9 Hz CHC*H*$_2$O), 2.41 (2H, dt, *J* = 24.1, 7.1 Hz, CFC*H*$_2$$CH_2$), 2.23 (2H, t, *J* = 7.1 CFCH$_2$C*H*$_2$), 2.17 – 1.97 (4H, m, (C*H*$_3$)$_2$CCHC*H*$_2$C*H*$_2$), 1.73 (3H, s, C*H*$_3$CCHCH$_2$CH$_2$), 1.69 (3H, s, C*H*$_3$CCHCH$_2$O), 1.62 (3H, s, C*H*$_3$CCHCH$_2$CH$_2$), 1.58 (3H, d, *J* = 1.3 Hz, C*H*$_3$CCFCH$_2$)
^{13}C NMR (126 MHz, D_2O) δ 154.26 (d, *J* = 238.0 Hz, CH$_3$C*C*F), 141.88 and 133.72 (C_q), 124.09 (s, $(CH_3)_2$C*C*HCH$_2$), 120.12 (d, *J* = 8.2 Hz, *C*HCH$_2$O), 112.27 (d, *J* = 16.6 Hz, CH$_3$*C*CFCH$_2$), 62.84 (d, *J* = 5.4 Hz, *C*H$_2$O), 35.83 (s, CFCH$_2$*C*H$_2$), 28.79 (d, *J* = 7.3 Hz, $(CH_3)_2$CCHCH$_2$*C*H$_2$), 26.59 (d, *J* = 29.4 Hz, CF*C*H$_2$CH$_2$), 25.36 (s, $(CH_3)_2$CCH*C*H$_2$CH$_2$), 24.83 (s, *C*H$_3$CCHCH$_2$O), 16.88 (s, *C*H$_3$CCHCH$_2$CH$_2$), 15.49 (s, *C*H$_3$CCHCH$_2$CH$_2$), 14.52 (d, *J* = 6.4 Hz, *C*H$_3$CCF).
^{19}F NMR (471 MHz, D_2O) δ –113.91 (t, *J* = 24.5 Hz).
^{31}P NMR (202 MHz, D_2O) δ –10.52 (d, *J* = 20.7 Hz), –10.84 (d, *J* = 20.6 Hz).
HRMS (ES^-, [M]): calculated for $[C_{15}H_{26}FO_7P_2]^-$; 399.1138, found 399.1147.

(2*E*,6*E*)-6-fluoro-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate ([2*E*,6*E*]-6F-FDP, 187)
[237]



Bis-triethylammonium phosphate (**145**, 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 5 mL) at 5-minutes intervals to a solution of **185** (93 mg, 0.4 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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 10 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound **187** as a light fluffy powder (27 mg, 15%).

^{1}H NMR (400 MHz, D_2O) δ 5.51 – 5.44 (1H, m, CC*H*CH_2OH), 5.22 – 5.15 (1H, m, $(CH_3)_2$CC*H*), 4.53 – 4.43 (2H, m, CHC*H*$_2$OH), 2.49 – 2.35 (2H, m, (CH_3)CCFC*H*$_2$), 2.27 – 2.19 (2H, m, CFCH$_2$C*H*$_2$),

2.17 – 2.07 (2H, m $(CH_3)_2CCHCH_2CH_2$), 2.06 – 1.98 (2H, m, $(CH_3)_2CCHCH_2CH_2$),1.74 (3H, s, CH_3), 1.70 (3H, s, CH_3), 1.52 (6H, s, 2 x CH_3).

^{13}C NMR (101 MHz, D_2O) δ 154.82 (d, *J* = 239.3 Hz, CH_3CCF), 141.93 (s, C_q), 133.93 (s, C_q), 123.85 (s, $CCHCH_2OH$), 120.06 (s, $(CH_3)_2CCH$), 112.13 (d, *J* = 14.8 Hz, CH_3CCF), 62.87 (s, $CHCH_2OH$), 36.12 (s, $CFCH_2CH_2$), 31.00 (d, J= 6.1 Hz, $(CH_3)_2CCHCH_2CH_2$), 26.50 (d, *J* = 29.1 Hz, $(CH_3)CCFCH_2$), 25.67 (d, *J* = 2.6 Hz, allylic CH_2), 24.89 (s, CH_3), 16.88 (s, CH_3CCHCH_2OH), 15.48 (s, CH_3), 12.30 (d, *J* = 9.9 Hz, $(CH_3)CCF$).

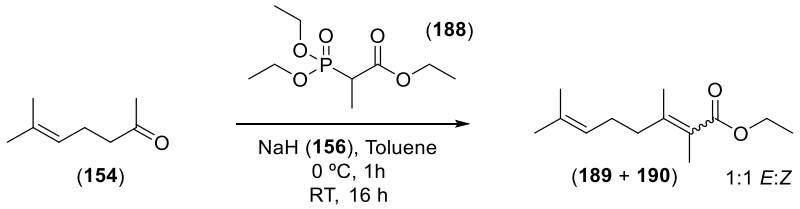
^{31}P NMR (202 MHz, D_2O) δ –8.33 (d, *J* = 15.1 Hz), –10.56 (d, *J* = 18.0 Hz).

^{19}F NMR (471 MHz, D_2O) δ –112.38 (t, *J* = 24.4 Hz).

HRMS (ES$^-$, [M]): calculated for $[C_{15}H_{26}FO_7P_2]^-$; 399.1138, found 399.1151.

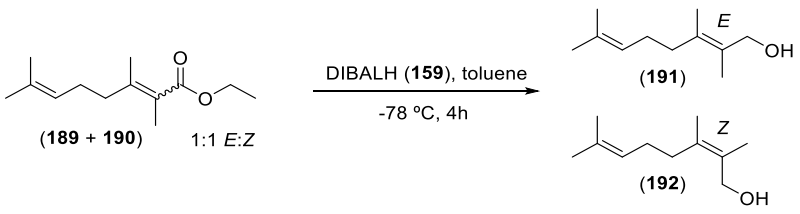
7.2.4. Synthesis of (2*E*,6*E*)- and (2*E*, 6*Z*)-6-methylfarnesyl diphosphate [(2*E*,6*E*)- and (2*E*,6*Z*)-6Met-FDP], (189) and (202)

Preparation of ethyl (*E*) and (*Z*)-2,3,7-trimethylocta-2,6-dienoate (189 and 190)



A solution of triethyl 2-phosphonopropionate (**188**, 13.5 mL, 63 mmol, 1.5 eq.) in anhydrous THF (20 mL) was added dropwise to a stirred suspension of sodium hydride (**156**, 2.5 g, 63 mmol, 1.5 eq) in dry THF (20 mL) at 0 ºC under argon atmosphere. The mixture was stirred for 1 hour at room temperature (colour changed from clear yellow to orange) and cooled to 0 ºC prior to the addition of 6-methyl-5-hepten-2-one (**154**, 5.3 g, 42 mmol, 1 eq.) in anhydrous THF (20 mL). After addition, ice bath was removed, and the mixture was stirred at room temperature for 16 hours. Cold water was added slowly to the reaction mixture (40 mL) and organic products extracted with diethyl ether (3 x 100 mL). The organic extracts were washed with water (3 x 100 mL) and brine (3 x 100 mL), dried over anhydrous Na_2SO_4 and filtered. Evaporation under reduced pressure afforded the fluoroester intermediate as a 1:1 mixture of **189** and **190** isomers (as judged by GC analysis). The mixture was passed through a silica plug (20% ethyl acetate: hexane as eluent) and reduced without further separation (6.42 g, 73%).

Preparation of (*E*)- and (*Z*)-2,3,7-trimethylocta-2,6-dien-1-ol (191 and 192)



A hexane solution of DIBALH (**159**, 60 mL, 60 mmol, 1 M, 2 eq.) was added dropwise to a mixture of fluorinated esters **189** and **190** (6.4 g, 30 mmol, 1 eq.) in anhydrous toluene (40 mL) at -78 °C under argon atmosphere. Reaction was followed via TLC and Rochelle salts were added dropwise at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 4 hours at room temperature and organic products extracted with DCM (3 x 100 mL). The organic extracts were washed with water (3 x 100 mL) and brine (3 x 100 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography (5% ethyl acetate:

hexane as eluent) gave the *E* isomer (**193**) (1.76 g, 32%) followed by the *Z* isomer (**194**) (1.52 g, 27%) alcohols as clear oils.

(2*E*)-2,3,7-trimethylocta-2,6-dien-1-ol (191)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.15 – 5.09 (1H, m, $(CH_3)CCH$), 4.12 (2H, s, $CC(CH_3)CH_2OH$), 2.09 – 2.00 (4H, m, 2 x allylic CH_2), 1.75 (3H, s, $CC(CH_3)$ CH_2OH), 1.73 (3H, s, CH_3), 1.68 (6H, s, CH_3), 1.61 (3H, s, CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 133.18 (C_q), 131.86 (C_q), 128.03 ($(CH_3)CC(CH_3)$ CH_2OH), 124.27 ($(CH_3)CCH$), 64.24 ($C(CH_3)$ CH_2OH), 35.07 & 26.52 (2 x allylic CH_2), 25.83 & 18.05 (2 x $(CH_3)CCH$), 17.72 (s, CH_3), 16.34 (s, $C(CH_3)$ CH_2OH).

HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{11}H_{18}]^+$; 150.1409, found 150.1409.

(2*Z*)-2,3,7-trimethylocta-2,6-dien-1-ol (192)

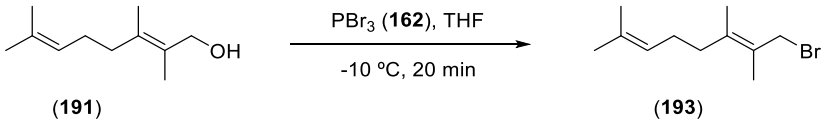
^{1}H NMR (500 MHz, $CDCl_3$) δ 5.10 (1H, t, J = 7.2 Hz, $(CH_3)CCH$), 4.07 (2H, d, J = 0.6 Hz, $C(CH_3)CH_2OH$), 2.15 – 2.11 (2H, m, $CHCH_2CH_2$), 2.08 - 2.02 (2H, m, $CHCH_2CH_2$), 1.74 (3H, s, $C(CH_3)CH_2OH$), 1.69 (6H, s, 2 x $(CH_3)CCH$), 1.59 (3H, s, CH_3).

13C NMR (126 MHz, $CDCl_3$) δ 133.01(s, C_q), 132.57 (s, C_q), 128.58 (s, $(CH_3)CC(CH_3)CH_2OH$), 124.25 (s, $(CH_3)CCH$), 63.67 (s, $C(CH_3)CH_2OH$), 34.20 (s, $CHCH_2CH_2$), 27.21 (s, $CHCH_2CH_2$), 25.78 & 18.87 (2 x $(CH_3)CCH$), 17.73 (CH_3), 16.94 ($C(CH_3)$ CH_2OH).

TLC (Hexane: EtOAc = 4: 1) R_f 0.10.

HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{11}H_{18}]^+$; 150.1409, found 150.1409.

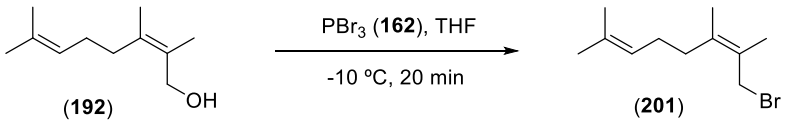
(2*E*)-1-bromo-2,3,7-trimethylocta-2,6-diene (193)



To a solution of fluorinated alcohol **191** (0.9 g, 5.4 mmol, 1 eq.) in anhydrous THF (10 mL) was added dropwise a solution of PBr_3 (**33**, 254 µL, 2.7 mmol, 0.5 eq.) in THF (5 mL) at -10 °C under argon atmosphere. The reaction mixture was stirred for 30 min and quenched with sat. $NaHCO_3$ sol. (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **193** as a pale-yellow oil (1.18 g, 94%), which was used in the next reaction without further purification.

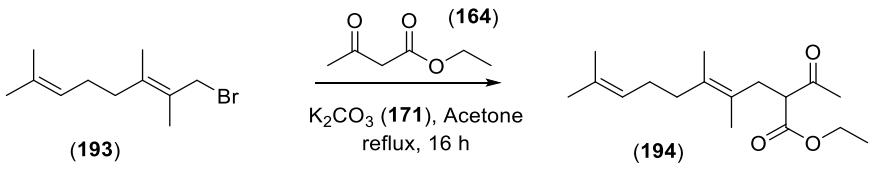
^{1}H NMR (500 MHz, $CDCl_3$) δ 5.13 – 5.07 (1H, m, $(CH_3)CCH$), 4.06 (2H, s, $CC(CH_3)$ CH_2Br), 2.13 – 2.07 (4H, m, 2 x allylic CH_2), 1.76 (3H, s, CH_3), 1.75(3H, s, CH_3), 1.68 (3H, s, CH_3), 1.60 (3H, s, CH_3).
TLC (Hexane: EtOAc = 4: 1) R_f 0.95.

(2*Z*)-1-bromo-2,3,7-trimethylocta-2,6-diene (201)



To a solution of fluorinated alcohol **192** (0.84 g, 5 mmol, 1 eq.) in anhydrous THF (20 mL) was added dropwise a solution of PBr_3 (**162**, 235 µL, 2.5 mmol, 0.5 eq.) in THF (10 mL) at -10 °C under argon atmosphere. The reaction mixture was stirred for 30 min and quenched with *sat.* $NaHCO_3$ solution (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **201** as a pale-yellow oil (1.09 g, 94%), which was used in the next reaction without further purification.
^{1}H NMR (300 MHz, $CDCl_3$) δ 5.17 – 5.01 (1H, m, $(CH_3)CCH$), 4.07 (2H, s, $CC(CH_3)$ CH_2Br), 2.22 – 1.94 (4H, m, 2 x allylic CH_2), 1.77 (3H, s, CH_3), 1.76 (3H, s, CH_3), 1.69 (3H, s, CH_3), 1.61 (3H, s, CH_3).
TLC (Hexane: EtOAc = 4: 1) R_f 0.95.

Ethyl (*E*)-2-acetyl-4,5,9-trimethyldeca-4,8-dienoate (194)



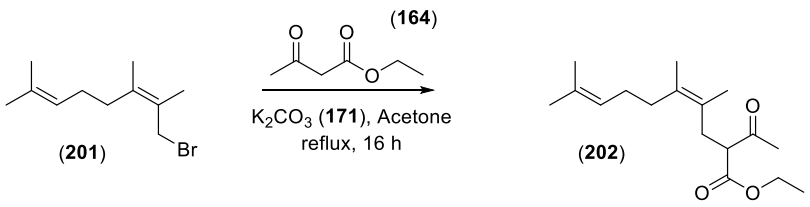
Ethyl acetoacetate (**164**, 0.99 mL, 7.7 mmol, 1.5 eq.) and potassium carbonate (**171**, 1.4 g, 10.2 mmol, 2 eq.) were added to a solution of bromide **193** (1.18 g, 5.1 mmol, 1 eq.) in dry acetone (30 mL) and refluxed for 16 hours. Volatile solvents were removed under reduced pressure and the resulting organic mixture extracted into diethyl ether (3 x 50 mL). Organic extracts were washed with water (3 x 50 mL), brine (3 x 50 mL) dried over anhydrous Na_2SO_4 and filtered. Products were concentrated under reduced pressure and purified by silica gel flash chromatography (20% ethyl acetate in hexane as eluent) to give **194** as a colourless oil (0.91 g, 64%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.13 – 5.03 (1H, m, $(CH_3)_2CCH$), 4.22 – 4.09 (2H, m, CH_3CH_2O), 3.58 (1H, dd, *J* = 8.1 Hz, 7.3 Hz, $COCHCO$)., 2.69 – 2.52 (2H, m, $CC(CH_3)CH_2CH$), 2.21 (3H, s, $COCH_3$), 2.04 – 1.94 (4H, m, 2 x allylic CH_2), 1.67 (s, 3H), 1.66 (3H, d, *J* = 1.3 Hz, CH_3), 1.63 (3H, d, *J* = 1.4 Hz, CH_3), 1.59 (3H, s, CH_3), 1.25 (1H, t, *J* = 7.1 Hz, CH_3CH_2O).

^{13}C NMR (126 MHz, $CDCl_3$) δ 203.46 ($CH_3C(O)CH$), 170.03 ($CH_3CH_2OC(O)$), 132.07 ($(CH_3)CCH$), 131.72 ($(CH_3)CC(CH_3)$), 124.33 (CH_3CCH), 124.08 ($(CH_3)CC(CH_3)$), 61.40 (CH_3CH_2O), 58.40 ($C(O)CHC(O)$), 35.05 (allylic CH_2), 33.30 ($CC(CH_3)CH_2$), 29.31 ($CH_3C(O)$), 26.64 (allylic CH_2), 25.83 (CH_3), 18.55 (CH_3), 17.80 (CH_3), 17.71 (CH_3), 14.29 (OCH_2CH_3).

HRMS (ES^+, [M]): calculated for $[C_{17}H_{28}O_3]^+$; 280.2038, found 280.2038.

Ethyl (*Z*)-2-acetyl-4,5,9-trimethyldeca-4,8-dienoate (202)



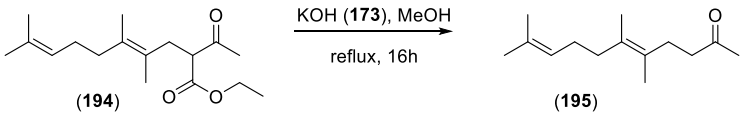
Ethyl acetoacetate (**164**, 0.92 mL, 7.1 mmol, 1.5 eq.) and potassium carbonate (**171**, 1.22 g, 9.4 mmol, 2 eq.) were added to a solution of bromide **201** (1.09 g, 4.7 mmol, 1 eq.) in dry acetone (30 mL) and refluxed for 16 hours. Volatile solvents were removed under reduced pressure and the resulting organic mixture extracted into diethyl ether (3 x 50 mL). Organic extracts were washed with water (3 x 50 mL), brine (3 x 50 mL) dried over anhydrous Na_2SO_4 and filtered. Products were concentrated under reduced pressure and purified by silica gel flash chromatography (20% ethyl acetate in hexane as eluent) to give **202** as a colourless oil (0.66 g, 50%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.13 – 5.07 (1H, m, $(CH_3)CCH$), 4.20 – 4.12 (2H, m, CH_3CH_2O), 3.57 (1H, dd, *J* = 8.1 Hz, 7.0 Hz, $COCHCO$)., 2.69 – 2.55 (2H, m, $CC(CH_3)CH_2CH$), 2.21 (3H, s, $COCH_3$), 2.14 – 1.94 (4H, m, 2 x allylic CH_2), 1.68 (s, 3H), 1.63 (3H, s, CH_3), 1.61 (3H, s, CH_3), 1.56 (3H, s, CH_3), 1.25 (1H, t, *J* = 7.1 Hz, CH_3CH_2O).

^{13}C NMR (126 MHz, $CDCl_3$) δ 203.44 ($CH_3C(O)CH$), 170.09 ($CH_3CH_2OC(O)$), 131.92 ($(CH_3)CCH$), 131.79 ($(CH_3)CC(CH_3)$), 124.37 (CH_3CCH), 124.03 ($(CH_3)CC(CH_3)$), 61.40 (CH_3CH_2O), 58.80 ($C(O)CHC(O)$), 34.49 (allylic CH_2), 32.70 ($CC(CH_3)CH_2$), 29.27 ($CH_3C(O)$), 27.09(allylic CH_2), 25.85 (CH_3), 18.76 (CH_3), 18.27 (CH_3), 17.76 (CH_3), 14.22 (OCH_2CH_3).

HRMS (ES^+, [M]): calculated for $[C_{17}H_{28}O_3]^+$; 280.2038, found 280.2041.

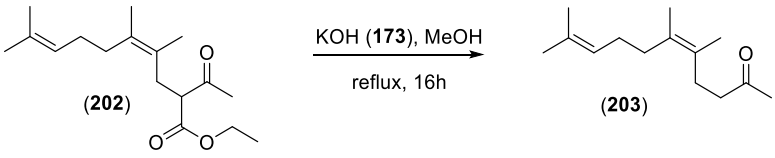
(*E*)-5,6,10-trimethylundeca-5,9-dien-2-one (195)



Potassium hydroxide (**173**, 0.55 g, 9.6 mmol, 3 eq.) was added to a solution of **194** (0.91 g, 3.2 mmol, 1 eq.) in methanol (20 mL) and refluxed for 16 hours. The reaction mixture was cooled to room temperature and acidified with 1M HCl solution. Organic products were extracted with diethyl ether (3 x 50 mL) and washed with water (3 x 20 mL) and brine (3 x 20 mL). Organic layers were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification *via* silica column flash chromatography gave the desired ketone **195** as a pale-yellow oil (0.46 g, 70 %).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.15 – 5.06 (1H, m, $(CH_3)CCH$), 2.48 – 2.41 (2H, m, $CH_2C(O)$), 2.31 – 2.23 (2H, m, $CH_2CH_2C(O)$), 2.14 (3H, s, $CH_2C(O)CH_3$), 2.03 – 1.94 (4H, m, 2 x allylic CH_2), 1.69 (3H, s, CH_3), 1.64 (3H, s, CH_3), 1.62 (3H, s, CH_3), 1.61 (3H, d, J = 0.7 Hz, CH_3).
^{13}C NMR (75 MHz, $CDCl_3$) δ 209.41 ($CH_3C(O)$), 131.58 ($(CH_3)CCH$), 129.53 ($(CH_3)CC(CH_3)$), 126.75 ($(CH_3)CC(CH_3)$), 124.47 ($(CH_3)CCH$), 42.40 ($CH_2CH_2C(O)$), 30.06 ($CH_3C(O)$), 29.06 ($CH_2CH_2C(O)$), 34.88 & 26.77(2 x allylic CH_2), 25.85 & 18.2 & 17.96 & 17.71 (4 x CH_3).
HRMS (EI^+, [M]): calculated for $[C_{14}H_{24}O]^+$; 208.1827, found 208.1829.

(*Z*)-5,6,10-trimethylundeca-5,9-dien-2-one (203)



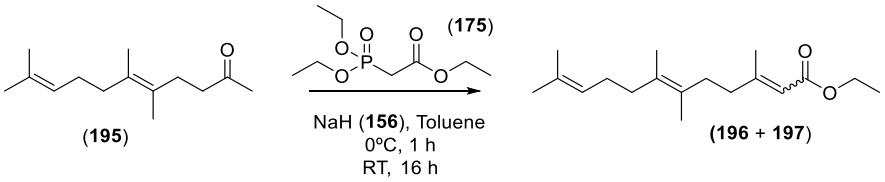
Potassium hydroxide (**173**, 0.40 g, 7.05 mmol, 1.5 eq.) was added to a solution of **202** (0.66 g, 2.35 mmol, 1 eq.) in methanol (20 mL) and refluxed for 16 hours. The reaction mixture was cooled to room temperature and acidified with 1M HCL solution. Organic products were extracted with diethyl ether (3 x 50 mL) and washed with water (3 x 20 mL) and brine (3 x 20 mL). Organic layers were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification *via* silica column flash chromatography gave the desired ketone **203** as a pale-yellow oil (0.27 g, 55%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.13 – 5.05 (1H, m, $(CH_3)CCH$), 2.49 – 2.40 (2H, m, $CH_2C(O)$), 2.31 – 2.24 (2H, m, $CH_2CH_2C(O)$), 2.13 (3H, s, $CH_2C(O)CH_3$), 2.03 – 1.99 (4H, m, 2 x allylic CH_2), 1.67 (3H, s, CH_3), 1.62 (3H, d, J = 0.8 Hz, CH_3)., 1.61 (3H, s, CH_3), 1.59 (3H, d, J = 0.7 Hz, CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 209.24 (CH_3C(O)), 131.66 ((CH_3)CCH), 129.45 ((CH_3)CC(CH_3)), 126.73 ((CH_3)CC(CH_3)), 124.50 ((CH_3)CCH), 42.99 (CH_2CH_2C(O)), 30.01 (CH_3C(O)), 28.48 (CH_2CH_2C(O)), 34.33 & 27.22(2 x allylic CH_2), 25.84 & 18.56 & 18.32 & 17.73 (4 x CH_3).

HRMS (EI$^+$, [M]): calculated for $[C_{14}H_{24}O]^+$; 208.1827, found 208.1833.

(2*E*,6*E*)- and (2*Z*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (196 and 197)



A suspension of sodium hydride (**156**, 132 mg, 3.3 mmol, 60% dispersion in mineral oil, 1.5 eq.) in anhydrous THF (10 mL) was cooled to 0 ºC. Ethyl 2-(diethoxyphosphoryl)acetate (**175**, 680 μL , 3.3 mmol, 1.5 eq) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and **195** (460 mg, 2.2 mmol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched with water (20 mL) and reaction products extracted with diethyl ether (3 x 50 mL). Combined organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL). Extracts were dried over sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave the fluorinated esters **196** and **197** as a clear oil (530 mg, 87%) that were used in the next reaction without separation. Esters were partially separated for analysis.

(2*E*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (196)

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.68 – 5.64 (1H, m, CCHC(O)OCH_2CH_3), 5.14 – 5.08 (1H, m, (CH_3)CCH), 4.14 (2H, q, J = 7.1 Hz, OCH_2CH_3), 2.18 (3H, d, J = 1.3 Hz, CH_2(CH_3)CCHC(O)), 2.17 – 2.13 (4H, m, CH_2CH_2(CH_3)CCHC(O)), 2.07 – 1.95 (4H, m, 2 x allylic CH_2), 1.68 (3H, s, CH_3), 1.64 (6H, s, 2 x CH_3), 1.60 (3H, s, CH_3), 1.57 (3H, s, CH_3) 1.20 (3H, t, J = 7.1 Hz, OCH_2CH_3).

^{13}C NMR (75 MHz, $CDCl_3$) δ 167.07 (CHC(O)O), 160.42 ((CH_3)CCHC(O)O), 131.57 & 129.42 & 127.25 ((CH_3)CC), 124.54 ((CH_3)CCH), 115.58 (CCHC(O)O), 59.60 (OCH_2CH_3), 39.61 (CH_2(CH_3)CCHC(O)), 33.16 (CH_2CH_2(CH_3)CCHC(O)), 34.93 & 26.83 (allylic CH_2), 25.87, 19.09, 18.22, 18.21 & 17.74 (5 x CH_3), 14.48 (OCH_2CH_3).

HRMS (EI$^+$, [M]): calculated for $[C_{18}H_{30}O_2]^+$; 278.2246, found 278.2238.

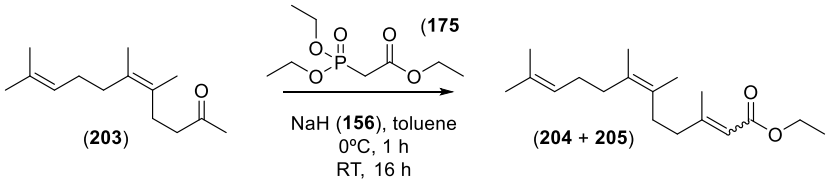
(2*Z*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (197)

^{1}H NMR (300 MHz, $CDCl_3$) δ 5.66 – 5.63 (1H, m, CC*H*C(O)OCH_2CH_3), 5.15 – 5.09 (1H, m, (CH_3)CC*H*), 4.14 (2H, q, *J* = 7.1 Hz, OC*H*$_2$$CH_3$), 2.63 (2H, dd, *J* = 9.5, 6.8 Hz, C*H*$_2$(CH_3)CCHC(O)), 2.16 (2H, dd, *J* = 9.5, 6.9 Hz, C*H*$_2$$CH_2$($CH_3$)CCHC(O)), 2.06 – 1.96 (4H, m, 2 x allylic C*H*$_2$), 1.90 (3H, d, *J* = 1.3 Hz, CH_2(C*H*$_3$)CCHC(O)), 1.69 (3H, s, C*H*$_3$), 1.68 (3H, s, C*H*$_3$), 1.61 (3H, s, C*H*$_3$), 1.43 (3H, s, C*H*$_3$), 1.27 (3H, t, *J* = 7.1 Hz, OCH_2C*H*$_3$).

^{13}C NMR (75 MHz, CDCl3) δ 166.31 (CH*C*(O)O), 160.37 ((CH_3)*C*CHC(O)O), 131.36 & 128.96 & 127.81 ((CH_3)*C*C), 124.55 ((CH_3)C*C*H), 116.03 (C*C*HC(O)O), 59.42 (O*C*H$_2$$CH_3$), 34.85 (allylic *C*H$_2$), 33.41 (*C*H$_2$($CH_3$)CCHC(O)), 32.11 (*C*H$_2$$CH_2$($CH_3$)CCHC(O)), 26.70 (allylic *C*H$_2$), 30.33 & 25.73 & 25.47, 17.99 & 17.60 (5 x *C*H$_3$), 14.37 (OCH_2*C*H$_3$).

HRMS (EI$^+$, [M]): calculated for [$C_{18}H_{30}O_2$]$^+$; 278.2246, found 278.2230.

(2*E*,6*Z*)- and (2*Z*,6*Z*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (204 and 205)



A suspension of sodium hydride (**156**, 68 mg, 1.7 mmol, 60% dispersion in mineral oil, 1.5 eq.) in anhydrous THF (10 mL) was cooled to 0 ºC. Ethyl 2-(diethoxyphosphoryl)acetate (**175**, 350 μL , 1.7 mmol, 1.5 eq.) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and **203** (235 mg, 1.1 mmol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched with water (20 mL) and reaction products extracted with diethyl ether (3 x 50 mL). Combined organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL). Extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave the fluorinated esters **204** and **205** as a clear oil (220 mg, 73%) that were used in the next reaction without separation. Esters were partially separated for analysis.

(2*E*,6*Z*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (204)

^{1}H NMR (300 MHz, $CDCl_3$) δ 5.66 (1H, s, CC*H*C(O)OCH_2CH_3), 5.13 – 5.08 (1H, m, (CH_3)CC*H*), 4.14 (2H, q, *J* = 7.1 Hz, OC*H*$_2$$CH_3$), 2.17 (3H, s, CH_2(C*H*$_3$)CCHC(O)), 2.17 – 2.11 (4H, m, C*H*$_2$$CH_2$(CH$_3$)CCHC(O)), 2.05 – 1.96 (4H, m, 2 x allylic C*H*$_2$), 1.68 (3H, s, C*H*$_3$), 1.63 (6H, s, 2 x C*H*$_3$), 1.60 (3H, s, C*H*$_3$), 1.28 (3H, t, *J* = 7.1 Hz, OCH_2C*H*$_3$).

^{13}C NMR (75 MHz, $CDCl_3$) δ 167.06 (CH*C*(O)O), 160.41 ((CH$_3$)*C*CHC(O)O), 131.68, 129.31 & 127.23 ((CH$_3$)*C*C), 124.53 ((CH$_3$)C*C*H), 115.51 (C*C*HC(O)O), 59.60 (O*C*H$_2$CH$_3$), 34.47 (allylic *C*H$_2$), 40.26 (*C*H$_2$(CH$_3$)CCHC(O)), 32.75 (*C*H$_2$CH$_2$(CH$_3$)CCHC(O)), 27.28 (allylic *C*H$_2$), 25.85, 19.08,18.56, 18.55 & 17.75 (5 x *C*H$_3$), 14.48 (OCH$_2$*C*H$_3$).

HRMS (EI$^+$, [M]): calculated for $[C_{18}H_{30}O_2]^+$; 278.2246, found 278.2236.

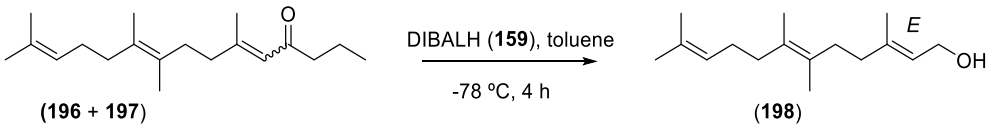
(2*Z*,6*Z*)- 3,6,7,11-tetramethyldodeca-2,6,10-trienoate (205)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.64 (1H, s, CC*H*C(O)OCH$_2$CH$_3$), 5.17 – 5.08 (1H, m, (CH$_3$)CC*H*), 4.14 (2H, q, *J* = 7.1 Hz, OC*H*$_2$CH$_3$), 2.64 (2H, dd, *J* = 9.6, 6.8 Hz, CH$_2$C*H*$_2$(CH$_3$)CCHC(O)), 2.10 (2H, dd, *J* = 9.5, 6.7 Hz, C*H*$_2$CH$_2$(CH$_3$)CCHC(O)), 2.09 – 1.98 (4H, m, 2 x allylic C*H*$_2$), 1.89 (3H, s, CH$_2$(C*H*$_3$)CCHC(O)), 1.69 (3H, s, C*H*$_3$), 1.68 (3H, s, C*H*$_3$), 1.64 (3H, s, C*H*$_3$), 1.60 (3H, s, C*H*$_3$), 1.27 (3H, t, *J* = 7.1 Hz, OCH$_2$C*H*$_3$).

^{13}C NMR (126 MHz, $CDCl_3$) δ 166.25 (CH*C*(O)O), 160.31 ((CH$_3$)*C*CHC(O)O), 131.32, 128.83 & 127.80 ((CH3)*C*C), 124.62 ((CH$_3$)C*C*H), 115.98 (C*C*HC(O)O), 59.41 (O*C*H$_2$CH$_3$), 34.30 (allylic *C*H$_2$), 32.89 (*C*H$_2$(CH$_3$)CCHC(O)), 32.67 (*C*H$_2$CH$_2$(CH$_3$)CCHC(O)), 27.25 (allylic *C*H$_2$), 25.73, 25.48, 18.45, 18.3 & 17.61 (5 x *C*H$_3$), 14.38 (OCH$_2$*C*H$_3$).

HRMS (EI$^+$, [M]): calculated for $[C_{18}H_{30}O_2]^+$; 278.2246, found 278.2236.

(2*E*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trien-1-ol (198)



A hexane solution of DIBALH (**159**, 5.7 mL, 5.7 mmol, 1 M, 3 eq.) was added dropwise to a mixture of fluorinated esters **196** and **197** (530 mg, 1.9 mmol, 1eq.) in anhydrous toluene (20 mL) at -78 ºC under argon. Reaction was followed *via* TLC and Rochelle salts were added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 50 mL). The organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 5% ethyl acetate/hexane as eluent separated the 2*E*,6*E*-fluorinated alcohol **198** (123 mg, 27%) and the 2*Z*,6*E*-fluorinated alcohol **199** (46 mg, 10%) as clear oils.

(2*E*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trien-1-ol (198)

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.42 (1H, t, *J* = 7.0 Hz, $(CH_3)CCHCH_2OH$), 5.15 – 5.08 (1H, m, $(CH_3)CCH$), 4.15 (2H, d, *J* = 6.9 Hz, $(CCHCH_2OH)$, 2.15 – 2.08 (2H, m, allylic CH_2), 2.06 – 1.95 (6H, m 3 x allylic CH_2, overlap), 1.70 (3H, d, *J* = 0.6 Hz, CH_3), 1.68 (3H, s, CH_3), 1.64 (6H, s, 2 x CH_3), 1.60 (3H, d , *J* = 0.6 Hz, CH_3).

^{13}C NMR (101 MHz, CDC_{l3}) δ 140.44 ($(CH_3)CC(CH_3)CH_2O$), 131.52, 128.65 & 128.06 (3 x$(CH_3)CC$), 124.52 & 123.20(2 x $(CH_3)CCH$), 59.59 ($CHCH_2O$), 38.11, 34.92, 33.48, 26.88 (4 x allylic CH_2), 25.87 (CH_3), 18.25 (CH_3), 18.19 (CH_3), 17.74 (CH_3), 16.52 (CH_3).

HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{16}H_{26}]^+$; 218.2035, found 218.2036.

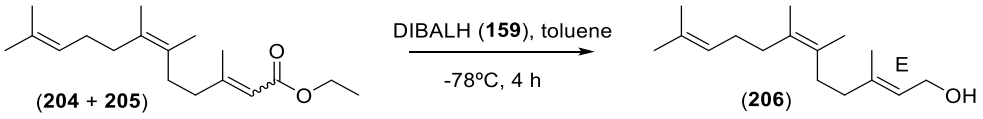
(2*Z*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trien-1-ol (199) - not used

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.42 (1H, t, *J* = 7.1 Hz, $(CH_3)CCHCH_2OH$), 5.12 (1H, b, $(CH_3)CCH$), 4.10 (2H, d, *J* = 7.2 Hz, $(CCHCH_2OH)$, 2.15 – 2.05 (4H, s, 2 x allylic CH_2), 2.05 – 1.96 (4H, m, 2 x allylic CH_2), 1.78 (3H, s, CH_3), 1.68 (3H, s, CH_3), 1.66 (3H, s, CH_3), 1.65 (3H, s, CH_3), 1.60 (3H, s, CH_3).

^{13}C NMR (75 MHz, $CDCl_3$) δ 140.63 ($(CH_3)CC(CH_3)CH_2O$), 131.59, 129.16 & 128.21 (3 x$(CH_3)CC$), 124.56 & 124.33 (2 x $(CH_3)CCH$), 59.21 ($CHCH_2O$), 34.91, 33.73, 30.72 & 26.82 (4 x allylic CH_2), 25.86 (CH_3), 23.80 (CH_3), 18.44 (CH_3), 18.23 (CH_3), 17.74 (CH_3).

HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{16}H_{26}]^+$; 218.2035, found 218.2035.

(2*E*,6*Z*)-3,6,7,11-tetramethyldodeca-2,6,10-trien-1-ol (206)



(**204** + **205**) (**206**)

A hexane solution of DIBALH (**159**, 2.2 mL, 2.2 mmol, 1 M, 3 eq.) was added dropwise to a mixture of fluorinated ester **204** and **205** (200 mg, 0.7 mmol, 1eq.) in anhydrous toluene (10 mL) at -78 ºC under argon. Reaction was followed *via* TLC and Rochelle salts were added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 50 mL). The organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 5% ethyl acetate/hexane as eluent separated the 2*E*,6*Z*-fluorinated alcohol **206** (72 mg, 44%) and the 2*Z*,6*Z*-fluorinated alcohol **207** (14 mg, 9%).as clear oils.

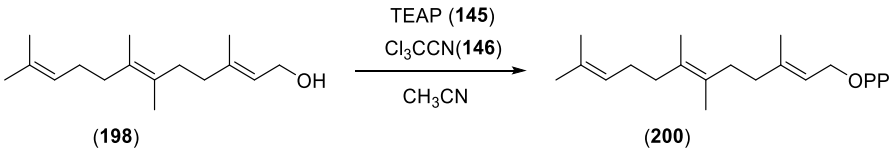
(2*E*,6*Z*)-3,6,7,11-tetramethyldodeca-2,6,10-trien-1-ol (206)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.35 (1H, t, *J* = 6.8 Hz, $(CH_3)CCHCH_2OH$), 5.05 (1H, s, $(CH_3)CCH$), 4.08 (2H, d, *J* = 6.9 Hz, $CCHCH_2OH$), 2.09 – 2.03 (2H, m, allylic CH_2), 2.00 – 1.90 (6H, m, 3 x allylic CH_2), 1.63 (3H, s, CH_3), 1.62 (3H, s, CH_3), 1.57 (6H, s, 2 x CH_3), 1.54 (3H, d , *J* = 0.6 Hz, CH_3).
^{13}C NMR (126 MHz, CDC_{13}) δ 140.50 ($(CH_3)CC(CH_3)CH_2O$), 131.58, 128.63 & 128.04 (3 x$(CH_3)CC$), 124.68 & 123.18 (2 x $(CH_3)CCH$), 59.60 ($CHCH_2O$), 38.79, 34.50, 33.09, 27.39 (4 x allylic CH_2), 25.87 (CH_3), 18.62 (CH_3), 18.55 (CH_3), 17.77 (CH_3), 16.54(CH_3).
HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{16}H_{26}]^+$; 218.2035, found 218.2035.

(2Z,6Z)-6-methyl-farnesol (207) - not used

$_{1}$H NMR (500 MHz, $CDCl_3$) δ 5.43 (1H, t, *J* = 7.0 Hz, $(CH_3)CCHCH_2OH$), 5.14 – 5.09 (1H, m, $(CH_3)CCH$), 4.09 (2H, d, *J* = 6.9 Hz, $CCHCH_2OH$), 2.15 – 2.06 (4H, m, 2 x allylic CH_2), 2.06 – 1.96 (4H, m 2 x allylic CH_2), 1.77 (3H, d, *J* = 0.6 Hz, CH_3), 1.68 (3H, s, CH_3), 1.65 (3H, s, CH_3), 1.64 (3H, s, CH_3), 1.61 (3H, d , *J* = 0.6 Hz, CH_3).
$_{13}$C NMR (126 MHz, CDC_{13}) δ 140.60 ($(CH_3)CC(CH_3)CH_2O$), 131.65, 129.22 & 128.14 (3 x$(CH_3)CC$), 124.58 & 124.39 (2 x $(CH_3)CCH$), 59.19 ($CHCH_2O$), 34.55, 33.22, 31.26 & 27.37 (4 x allylic CH_2), 25.86 (CH_3), 23.83 (CH_3), 18.79 (CH_3), 18.48(CH_3), 17.78 (CH_3).
HRMS (EI^+, $[M - H_2O]$): calculated for $[C_{16}H_{26}]^+$; 218.2035, found 218.2033.

(2*E*,6*E*)-6methyl-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate ([2*E*,6*E*]-6-methyl-FDP, 200)



(**198**) (**200**)

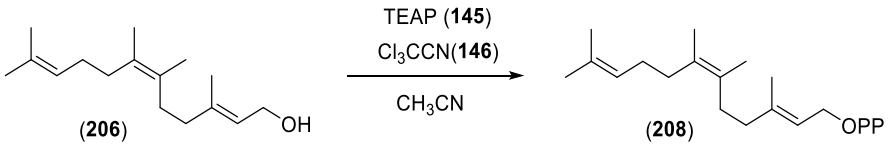
Bis-triethylammonium phosphate (**145**, TEAP) was prepared immediately prior to use by the dropwise addition of a solution of phosphoric acid in acetonitrile to a stirring solution of triethylamine in acetonitrile. TEAP was then added in three portions (3 x 5 mL) at 5-minutes intervals to a solution of **198** (20 mg, 0.08 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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 desired diphosphate were combined, ammonia and isopropanol removed under vacuum and the resulting aqueous solution diluted to 10 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound **200** as a light fluffy powder (9 mg, 25%).

^{1}H NMR (500 MHz, D_2O) δ 5.40 (1H, t, *J* = 7.1 Hz, $(CH_3)CCHCH_2O$), 5.18 – 5.11 (1H, m, $(CH_3)CCH$), 4.42 – 4.35 (2H, m, $CCHCH_2O$), 2.16 – 1.90 (8 H, m, 2 x allylic CH_2), 1.67 (3H, s, CH_3), 1.60 (9 H, s, 3 x CH_3), 1.54 (3H, s, CH_3).

^{31}P NMR (202 MHz, D_2O) δ –6.94 (d, *J* = 22.4 Hz), –10.51 (d, *J* = 21.7 Hz).

HRMS (ES$^-$, [M]): calculated for $[C_{16}H_{29}O_7P_2]^-$; 395.1389, found 395.1405.

(2*E*,6*Z*)-6methyl-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate ([2E,6Z]-6-methyl-FDP, 208)

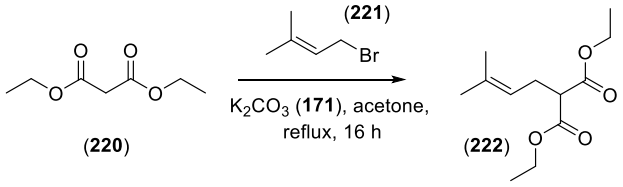


Bis-triethylammonium phosphate (**145**, TEAP) was prepared immediately prior to use by the dropwise addition of a solution of phosphoric acid in acetonitrile to a stirring solution of triethylamine in acetonitrile. TEAP was then added in three portions (3 x 5 mL) at 5 minute intervals to a solution of **206** (29 mg, 0.12 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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 10 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound as a light fluffy powder (16 mg, 29%).

^{1}H NMR (400 MHz, D_2O) δ 5.49 (1H, t, *J* = 6.5 Hz, $(CH_3)CCHCH_2O$), 5.27 – 5.21 (1H, m, $(CH_3)CCH$), 4.49 (2H, t, *J* = 6.5 Hz, $CCHCH_2O$), 2.27 – 2.02 (8H, m, 4 x allylic CH_2), 1.77 (3H, s, $CH3$), 1.71 (3H, s, CH_3), 1.68 (6H, s, 2 x CH_3), 1.64 (3H, s, CH_3).
^{31}P NMR (162 MHz, D_2O) δ –6.65 (d, *J* = 23.1 Hz), –10.35 (d, *J* = 22.2 Hz).
HRMS (ES^-, [M]): calculated for $[C_{16}H_{29}O_7P_2]^-$; 395.1407, found 395.1390.

7.2.5. Synthesis of (2*E*,6*E*)-7-hydrogenfarnesyl diphosphate ([(2*E*,6*E*)-7H-FDP], 203)

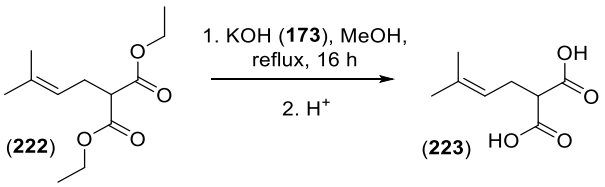
Diethyl 2-(3-methylbut-2-en-1-yl)malonate (222)[244]



To a solution of diethyl malonate (**220**, 22.8 mL, 0.15 mol, 1.5 eq.) in dry acetone (150 mL), was added 1-bromo-3-methylbut-2-ene (**221**, 11.6 mL, 0.1 mol, 1 eq.) and potassium carbonate (**171**, 27.6 g, 0.2 mol, 2 eq.). The resulting mixture was refluxed for 16 h. After this time, the solvent was removed under reduced pressure and water was added (50 mL). The organic mixture was extracted into diethyl ether (3 x 100 mL). Organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 10 % ethyl acetate/hexane as eluent afforded **216** as a clear oil (21.5 g, 94%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.08 – 5.03 (1H, m, $(CH_3)_2CC\mathit{H}CH_2$), 4.17 (4H, q, J = 7.1, 2 x $OC\mathit{H}_2CH_3$), 3.31 (1H, t, *J* = 7. 7, C(O)C*H*C(O)), 2.57 (2H, t, *J* =7.5, CHC*H*$_2$CH), 1.67 (3H, d, *J* = 1.1, C*H*$_3$), 1.65 (3H, s, C*H*$_3$), 1.24 (6H, t, *J* = 7.1, 2 x $OCH_2C\mathit{H}_3$).
^{13}C NMR (126 MHz, $CDCl_3$) δ 169.39 (s , *C*O), 134.93 (s, C_q), 119.83 (s, $(CH_3)_2C\mathit{C}HCH_2$), 61.38 (s, O*C*H$_2$CH$_3$), 52.36 (s, C(O)*C*HC(O)), 27.68 (s, CH*C*H$_2$CH), 25.86 (s, *C*H$_3$), 17.88 (s, *C*H$_3$), 14.19 (s, $OCH_2\mathit{C}H_3$).
HRMS (EI$^+$, [M]): calculated for $[C_{12}H_{20}O_4]^+$; 228.1362, found 228.1358.

2-(3-methylbut-2-en-1-yl) malonic acid (223)[301]

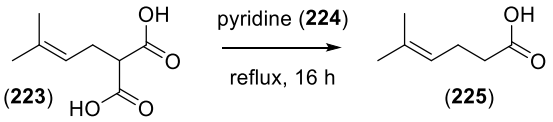


To a stirred solution of KOH (**173**, 15.9 g, 0.28 mmol, 3 eq.) in MeOH (150 mL), was added diethyl ester **222** (21.5 g, 0.09 mol, 1 eq.). The resulting solution was refluxed overnight. The solvent was removed under reduced pressure. The resulting crude solid was dissolved in distilled water, acidified with 2 M HCl to pH 2, and extracted with diethyl ether (3 x 100 mL Organic extracts were washed with

water (3 x 50 mL) and brine (3 x 50 mL), dried over sodium sulphate, filtered and concentrated under reduced pressure to afford diacid **223** as a pale yellow solid (15 g, 97%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.18 (2H, b, 2 x O*H*), 5.10 (1H, t, *J* = 7.3 Hz, $(CH_3)_2CCHCH_2$), 3.44 (1H, t, *J* = 7. 5, C(O)C*H*C(O)), 2.64 (2H, t, *J* =7.4, CHC*H*$_2$CH), 1.70 (3H, s, C*H*$_3$), 1.64 (3H, s, C*H*$_3$),
^{13}C NMR (126 MHz, $CDCl_3$) δ 175.00 (s , *C*O), 136.10 (s, C_q), 118.92 (s, $(CH_3)_2CCHCH_2$), 51.97 (s, C(O)*C*HC(O)), 27.64 (s, CH*C*H$_2$CH), 25.89 (s, *C*H$_3$), 17.89 (s, *C*H$_3$).
HRMS (ES^-, [M – H]): calculated for $[C_8H_{12}O_4]^-$; 171.0657, found 171.0654.

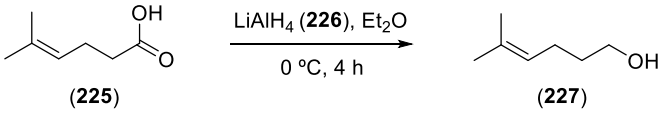
5-methylhex-4-enoic acid (225)[301]



A stirred solution of dicarboxylic acid **223** (15 g, 0.09 mol) in pyridine (**224**, 50 mL) was heated at reflux overnight. The reaction mixture then cooled to room temperature, diluted with distilled water (50 mL), acidified with 2 N HCl solution to pH 2, and extracted with dichloromethane (3 x 50 mL). The combined organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20% ethyl acetate/hexane as eluent afforded **225** as an orange oil (9.95 g, 87%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.10 (1H, t, *J* = 6.8 Hz, $(CH_3)_2CCHCH_2$), 2.40 – 2.28 (4H, t, *J* =7.4, 2 x allylic C*H*$_2$), 1.69 (3H, s, C*H*$_3$), 1.62 (3H, s, C*H*$_3$).
^{13}C NMR (126 MHz, $CDCl_3$) δ 180.13 (s, *C*O), 133.54 (s, C_q),122.20 (s, $(CH_3)_2CCHCH_2$), 34.43 & 25.80 (s, 2 x allylic C*H*$_2$), 23.46 (s, *C*H$_3$), 17.78 (s, *C*H$_3$).
HRMS (EI^+, [M]): calculated for $[C_7H_{12}O_2]^+$; 128.0837, found 128.0834.

5-methylhex-4-en-1-ol (227)[301]

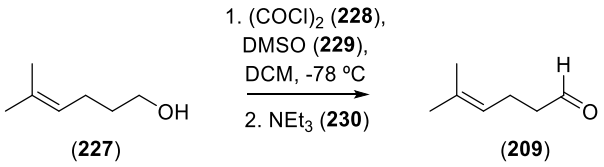


A solution of the carboxylic acid **225** (9.95 g, 0.08 mol, 1 eq.) in dry Et_2O (117 mL) was added dropwise to a solution of LiAlH4 (**226**, 14.8 g, 0.4 mol, 5 eq.) in dry Et_2O (100 mL) at 0 °C under argon atmosphere. After addition, the reaction flask was allowed to cool to room temperature. The reaction

was followed *via* TLC and Rochelle salts were added at 0 °C after complete consumption of starting material. Then, the organic products were extracted with DCM (3 x 100 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20% ethyl acetate/hexane as eluent gave the alcohol **227** as an orange oil. (7.9 g, 96%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.12 (1H, t, *J* = 7.1 Hz, $(CH_3)_2CCHCH_2$), 3.63 (2H, t, *J* = 6.5 Hz, CH_2OH), 2.05 (2H, q, *J* =7.2, $CHCH_2CH_2$), 1.68 (3H, s, CH_3), 1.61 (3H, s, CH_3), 1.63 – 1.56 (2H, m, $CHCH_2CH_2$).
^{13}C NMR (126 MHz, $CDCl_3$) δ 132.29 (s, C_q),123.99 (s, $(CH_3)_2CCHCH_2$), 62.82 (s, CH_2OH), 32.47 (s, $CHCH_2CH_2$), 25.81 (s, CH_3), 24.48 (s, $CHCH_2CH_2$), 17.76 (s, CH_3).
HRMS (EI$^+$, [M]): calculated for $[C_7H_{14}O]^+$; 114.1045, found 114.1046.

5-methylhex-4-enal (204)[249,250]

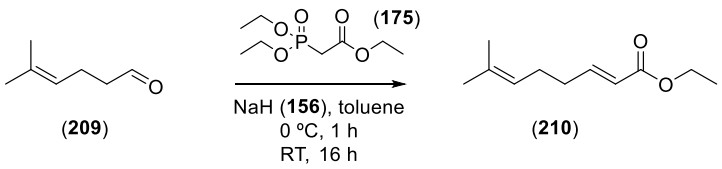


(**227**) (**209**)

To a stirred solution of oxalyl chloride (**228**, 9 mL, 0.1 mol, 2 eq.) in DCM (60 mL) at -78 ºC, was added DMSO (**229**, 15 mL 0.2 mol, 4 eq.) in DCM (60 mL) and stirred for 45 minutes. Then, a solution of the alcohol **227** (6 g, 0.05 mol, 1 eq.) in DCM (60 ml) was added slowly over 5 min and the reaction mixture was stirred for an hour. After this period, triethylamine (**230**, 81 mL, 0.55 mol, 11 eq.) in DCM (60 mL) was added over 5 min and the reaction mixture allowed to warm slowly to room temperature. The reaction was quenched by the addition of water (100 mL). The organic layer was separated, and the aqueous layer was extracted with DCM (3 x 100 mL). The organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20% ethyl acetate/hexane as eluent gave the aldehyde **209** (4.3 g, 76%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.74 (1H, s, CHO), 5.07 (1H, t, *J* = 6.9 Hz, $(CH_3)_2CCHCH_2$), 2.44 (2H, t, *J* = 7.3 Hz, CH_2CHO), 2.30 (2H, q, *J* =7.2, $CHCH_2CH_2$), 1.67 (3H, s, CH_3), 1.61 (3H, s, CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 202.69 (s, CHO), 133.30 (s, C_q),122.25 (s, $(CH_3)_2CCHCH_2$), 44.06 (s, CH_2CHO), 25.73 (s, CH_3), 21.01 (s, $CHCH_2CH_2$), 17.76 (s, CH_3).
HRMS (EI$^+$, [M]): calculated for $[C_7H_{12}O]^+$; 112.0888, found 112.0885.

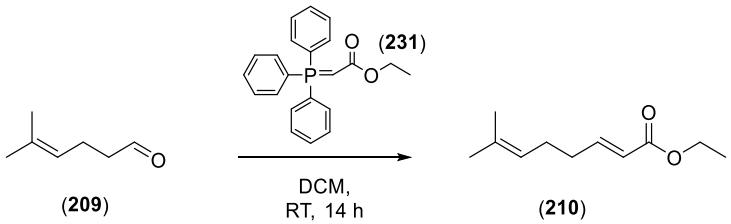
Ethyl (*E*)-7-methylocta-2,6-dienoate (210)[302]



A suspension of sodium hydride (**156**, 2.28 g, 0.06 mol, 60% dispersion in mineral oil, 1.5 eq.) in anhydrous THF (50 mL) was cooled to 0 ºC. Ethyl 2–(diethoxyphosphoryl)acetate (**175**, 11.4 mL, 0.06 mol, 1.5 eq.) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and aldehyde **209** (4.25 g, 0.04 mol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched slowly adding water (20 mL) and reaction products extracted with diethyl ether (3 x 100 mL). Combined organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL). Extracts were dried over sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave the ester **210** (only trans isomer) as a clear oil (3.9 g, 54%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 6.96 (1H, dt, *J* = 15.6, 6.8 Hz, $CH_2CHCHCO$), 5.82 (1H, d, *J* = 15.6 Hz, $CH_2CHCHCO$), 5.10 (1H, t, *J* = 6.5 Hz, $(CH_3)_2CCHCH_2$), 4.18 (2H, q, *J* = 7.1 Hz, OCH_2CH_3), 2.26 – 2.19 (2H, m, $CHCH_2CH_2CH$), 2.18 – 2.10 (2H, m, $CHCH_2CH_2$), 1.69 (s, 3H, CH_3), 1.60 (s, 3H, CH_3), 1.28 (3H, t, *J* = 7.1 Hz, OCH_2CH_3).
^{13}C NMR (500 MHz, $CDCl_3$) δ 166.90 (s, CO), 149.11(s, $CH_2CHCHCO$), 132.90 (s, C_q), 123.02 (s, $(CH_3)_2CCHCH_2$), 121.55 (s, $CH_2CHCHCO$), 60.27 (s, OCH_2CH_3), 32.58 (s, $CHCH_2CH_2CH$), 26.73 (s, $CHCH_2CH_2$), 25.81 (s, CH_3), 17.87 (s, CH_3), 14.42 (s, OCH_2CH_3).
HRMS (EI^+, $[M]^+$): calculated for $[C_{11}H_{19}O_2]^+$; 183.1385, found 183.1386.

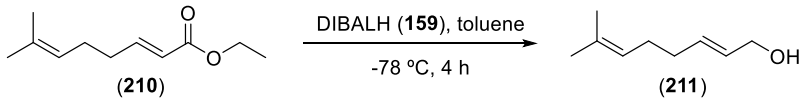
Second alternative



To a stirred solution of aldehyde **209** (1.6 g, 14 mmol, 1 eq.) in DCM (30 mL) at room temperature, was added (carbetoxymethylene)triphenylphosphorane (**231**, 5 g, 14.3 mol, 1.01 eq.), portion wise over 10 min. The mixture was stirred for 14 hours and concentrated under reduced pressure. The residue was triturated with pentane and the resulting white solid was removed by filtration. Concentration of

the filtrate and purification via silica gel flash chromatography (10% ethyl acetate in hexane) gave the ester **210** (only trans isomer) as a clear oil (1.7 g, 67%) (^{1}NMR, ^{13}CNMR and MS analysis is in full agreement with previous reaction).

(*E*)-7-methylocta-2,6-dien-1-ol (211)[302]



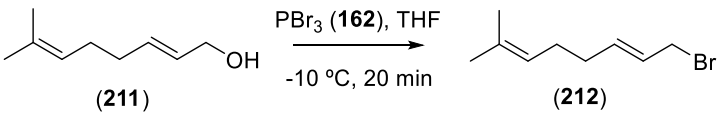
A hexane solution of DIBALH (**159**, 46 mL,46 mmol, 1 M, 3 eq.) was added dropwise to ester **210** (3 g, 15 mmol, 1eq.) in anhydrous toluene (50 mL) at –78 ºC under argon. Reaction was followed *via* TLC and K/Na tartrate salts were added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 50 mL). The organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 10 % ethyl acetate/hexane as eluent gave alcohol **211** as a clear oil. (1.5 g, 71%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.74 – 5.60 (2H, m, $CH_2CHCHCH_2$), 5.14 – 5.07 (1H, m, $(CH_3)_2CCHCH_2$), 4.09 (2H, q, J = 4.4 Hz, $CHCH_2OH$), 2.12 – 2.01 (4H, m, 2 x allylic CH_2), 1.69 (s, 3H, CH_3), 1.60 (s, 3H, CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 133.25 & 129.20 (s, $CH_2CHCHCH_2$), 132.11 (s, C_q), 123.88 (s, $(CH_3)_2CCHCH_2$), 63.99 (s, $CHCH_2OH$), 32.57 (s, $CHCH_2CH_2CH$), 27.83 (s, $CHCH_2CH_2$), 25.83 (s, CH_3), 17.88 (s, CH_3).

HRMS (EI$^+$, [M]): calculated for $[C_9H_{16}O]^+$; 140.1201, found 140.1203.

(*E*)-1-bromo-7-methylocta-2,6-diene (212)

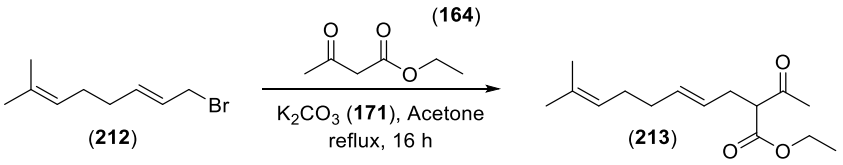


To a solution of alcohol **206** (1.5 g, 10 mmol, 1 eq.) in anhydrous THF (20 mL) was added dropwise a solution of PBr_3 (**162**, 0.5 mL, 5 mmol, 0.5 eq.) in THF (10 mL) at –10 °C under argon atmosphere. The reaction mixture was stirred for 30 min and quenched with *sat.* $NaHCO_3$ solution (10 mL). Organic products were extracted with diethyl ether (3 x 50 mL). The organic extracts were washed with water

(3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give crude **207** as a pale-yellow oil (1.4 g, 69%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.82 – 5.65 (2H, m, $CH_2CHCHCH_2$), 5.12 – 5.06 (1H, m, $(CH_3)_2CCHCH_2$), 3.95 (2H, q, J = 7.3 Hz, $CHCH_2Br$), 2.13 – 2.03 (4H, m, 2 x allylic CH_2), 1.69 (s, 3H, CH_3), 1.60 (s, 3H, CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 136.44 & 132.41 (s, $CH_2CHCHCH_2$), 126.62 (s, C_q), 123.53 (s, $(CH_3)_2CCHCH_2$), 33.74 (s, $CHCH_2Br$), 32.42 & 27.52 (s, 2 x allylic CH_2), 25.83 (s, CH_3), 17.90 (s, CH_3).

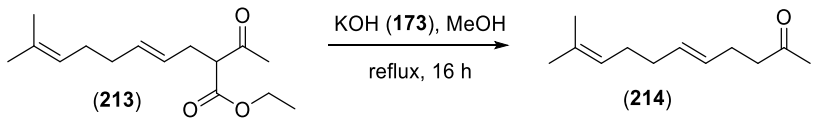
Ethyl (*E*)-2-acetyl-9-methyldeca-4,8-dienoate (213)



Ethyl acetoacetate (**164**, 1.3 mL, 10 mmol, 1.5 eq.) and potassium carbonate (**171**, 1.9 g, 14 mmol, 2 eq.) were added to a solution of bromide **212** (1.4 g, 7 mmol, 1 eq.) in dry acetone (30 mL) and refluxed for 16 hours. Volatile solvents were removed under reduced pressure and the resulting organic mixture extracted into diethyl ether (3 x 50 mL). Organic extracts were washed with water (3 x 50 mL), brine (3 x 50 mL) dried over anhydrous Na_2SO_4 and filtered. Products were concentrated under reduced pressure and purified by silica gel flash chromatography (20% ethyl acetate in hexane as eluent) to give **213** as a colourless oil (1.5 g, 85%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.54 – 5.47 (1H, m, $CH_2CHCHCH_2$), 5.38 – 5.29 (1H, m, $CH_2CHCHCH_2$), 5.10 – 5.03 (1H, m, $(CH_3)_2CCHCH_2$), 4.18 (2H, q, J = 7.1 Hz, OCH_2CH_3), 3.45 (1H, t, J = 7.5 Hz, $COCHCO$), 2.61 – 2.58 (2H, m, $CHCHCH_2CH$), 2.21 (3H, s, $COCH_3$), 2.03 – 1.95 (m, 4H, 2 x allylic CH_2), 1.67 (3H, s, CH_3), 1.58 (3H, s, CH_3), 1.26 (t, J = 7.1 Hz, OCH_2CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 202.98 (s, $CH_3C(O)CH$), 169.54 (s, $CH_3CH_2OC(O)$), 133.62 (s, $CH_2CHCHCH_2$), 131.92 (s, C_q), 125.69 (s, $CH_2CHCHCH_2$), 123.95 (s, $(CH_3)_2CCHCH_2$), 61.44 (s, OCH_2CH_3), 60.03 (s, $COCHCO$), 32.81 (s, allylic CH_2), 31.43 (s, $CHCHCH_2$), 29.22 (s, $COCH_3$), 27.99 (s, allylic CH_2), 25.80 (s, CH_3), 17.85 (s, CH_3), 14.25 (s, OCH_2CH_3).
HRMS (EI^+, [M]): calculated for $[C_{15}H_{24}O_3]^+$; 252.1725, found 252.1727.

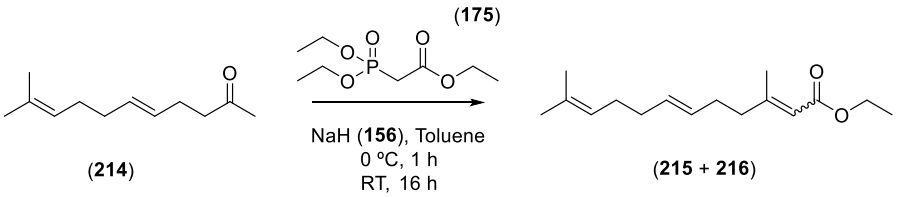
(*E*)-10-methylundeca-5,9-dien-2-one (214)



Potassium hydroxide (**173**, 1 g, 18 mmol, 3 eq.) was added to a solution of **213** (1.5 g, 6 mmol, 1 eq.) in methanol (40 mL) and refluxed for 16 hours. The reaction mixture was cooled to room temperature and acidified with 1M HCl. Organic products were extracted with diethyl ether (3 x 100 mL) and washed with water (3 x 50 mL) and brine (3 x 50 mL). Organic layers were dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification *via* silica column flash chromatography gave the desired ketone **214** as a pale-yellow oil (0.5 g, 46%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.47 – 5.33 (2H, m $CH_2CHCHCH_2$), 5.11 – 5.05 (1H, m, $(CH_3)_2CCH$), 2.47 (2H, t, J = 7.4 Hz, $CHCHCH_2CH_2CO$), 2.24 (2H dt, J = 13.8, 6.6 Hz, $CHCHCH_2CH_2$), 2.12 (3H, s, CH_2COCH_3), 2.03 – 1.94 (4H, m, 2 x allylic CH_2), 1.67 (3H, s, CH_3), 1.58 (3H, s, CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 208.65 (CO), 131.76 (C_q), 131.29 & 128.53 (s, $CH_2CHCHCH_2$), 124.10 (s, $(CH_3)CCH$), 43.67 (s, $CHCHCH_2CH_2CO$), 32.83 (s, allylic CH_2), 30.04 (s, CH_2COCH_3), 28.15 (s, allylic CH_2), 26.97 (s, $CHCHCH_2CH_2$), 25.79 (s, CH_3), 17.83 (s, CH_3).
HRMS (EI^+, $[M]^+$): calculated for $[C_{12}H_{20}O]^+$; 180.1514, found 180.1514.

Ethyl (6*E*)-3,11-dimethyldodeca-2,6,10-trienoate (215)



A suspension of sodium hydride (**156**, 170 mg, 4.2 mmol, 60% dispersion in mineral oil, 1.5 eq.) in anhydrous THF (10 mL) was cooled to 0 ºC. Ethyl 2-(diethoxyphosphoryl)acetate (**175**, 870 μL, 4.2 mmol, 1.5 eq.) was added to the stirred suspension and the reaction stirred at room temperature for 1 hour. Then, the reaction mixture was cooled to 0 ºC and **214** (0.5 g, 2.8 mmol, 1 eq.) added and reaction stirred at room temperature for 16 hours. Reaction was quenched with water (10 mL) and reaction products extracted with diethyl ether (3 x 50 mL). Combined organic extracts were washed with water (3 x 50 mL) and brine (3 x 50 mL). Extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification via silica gel flash chromatography (10% ethyl

acetate in hexane) gave a mixture of esters **215** and **216** as a clear oil (0.5 g, 71%) that was used in the next reaction without separation. Esters were partially separated for analysis.

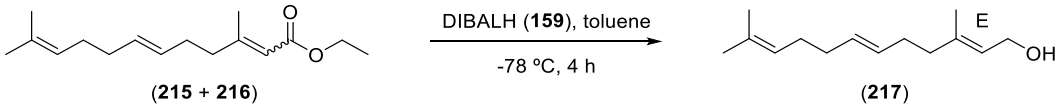
(2*E*,6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (215)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.67 – 5.64 (1H, m, CC*H*C(O)OCH_2CH_3), 5.48 – 5.33 (2H, m $CH_2CHCHCH_2$), 5.12 – 5.07 (1H, m, $(CH_3)CCH$), 4.14 (2H, q, *J* = 7.1 Hz, OCH_2CH_3), 2.22 – 2.13 (4H, m, $CHCHCH_2CH_2CH$), 2.15 (3H, d, *J* = 1.3 Hz, $CHC(CH_3)CHCO$), 2.06 – 1.97 (m, 2 x allylic CH_2), 1.68 (3H, s, CH_3), 1.59 (3H, s, CH_3), 1.27 (3H, t, *J* = 7.1 Hz, OCH_2CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 167.01 (s, $CHC(O)O$), 159.62 (s, $(CH_3)CCHC(O)O$), 131.79 (s, $(CH_3)CC$), 131.26 & 128.92 (s, $CH_2CHCHCH_2$), 124.16 (s, $(CH_3)CCH$), 115.90 (s, $CCHC(O)OCH_2CH_3$), 59.60 (s, OCH_2CH_3), 41.08 & 30.64 (s, $CHCHCH_2CH_2CH$) 32.90 & 28.23 (s, 2 x allylic CH_2), 25.82 (s, CH_3), 18.93 (s, $CHC(CH_3)CHCO$), 17.87 (s, CH_3), 14.47 (s, OCH_2CH_3).

(2*Z*, 6*E*)-3,6,7,11-tetramethyldodeca-2,6,10-trienoate (216)

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.67 – 5.64 (1H, m, CC*H*C(O)OCH_2CH_3), 5.47 – 5.43 (2H, m $CH_2CHCHCH_2$), 5.13 – 5.08 (1H, m, $(CH_3)CCH$), 4.13 (2H, q, *J* = 7.1 Hz, OCH_2CH_3), 2.70–2.64 (2H, m, $CHCHCH_2CH_2$), 2.20 – 2.13 (2H, m, $CHCHCH_2CH_2C$), 2.06 – 1.96 (m, 2 x allylic CH_2), 1.87 (3H, d, *J* = 1.3 Hz, $CHC(CH_3)CHCO$), 1.68 (3H, s, CH_3), 1.59 (3H, s, CH_3), 1.27 (3H, t, *J* = 7.1 Hz, OCH_2CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 166.49 (s, $CHC(O)O$), 160.09 (s, $(CH_3)CCHC(O)O$), 131.69 (s, $(CH_3)CC$), 130.79 & 129.57 (s, $CH_2CHCHCH_2$), 124.30 (s, $(CH_3)CCH$), 116.49 (s, $CCHC(O)OCH_2CH_3$), 59.57 (s, OCH_2CH_3), 33.52 (s, $CHCHCH_2CH_2C$), 31.34 (s, $CHCHCH_2CH_2$), 32.91 & 28.27 (s, 2 x allylic CH_2), 25.83 (s, CH_3), 25.47 (s, $CHC(CH_3)CHCO$), 17.86 (s, CH_3), 14.47 (s, OCH_2CH_3).

(2*E*,6*E*)-3,11-dimethyldodeca-2,6,10-trien-1-ol (217)



(**215** + **216**) (**217**)

A hexane solution of DIBALH (**159**, 4.5 mL, 4.5 mmol, 1 M, 3 eq.) was added dropwise to a mixture of esters **215** and **216** (380 mg, 1.5 mmol, 1eq.) esters in anhydrous toluene (20 mL) at –78 °C under argon. Reaction was followed *via* TLC and Rochelle salts were added at 0 °C after complete consumption of starting material. The resulting suspension was stirred for 16 hours at room temperature and organic products extracted with DCM (3 x 50 mL). The organic extracts were washed

with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 5% ethyl acetate/hexane as eluent separated the 2E,6E- (142 mg, 42%) and 2Z,6E- (29 mg, 9%) alcohols **217** and **218** as clear oils.

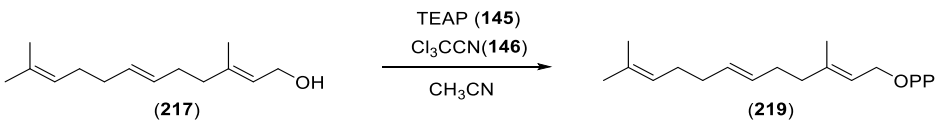
(2*E*,6*E*)-3,11-dimethyldodeca-2,6,10-trien-1-ol (217)

$_1$H NMR (500 MHz, $CDCl_3$) δ 5.47 5.34 (3H, m, 3 x C*H*), 5.14 – 5.07 (1H, m, $(CH_3)_2CCH$), 4.15 (2H, d, *J* = 6.9 Hz, $CHCH_2OH$), 2.14 – 1.95 (8H, m, 4 x allylic CH_2), 1.69 (3H, s, CH_3), 1.67 (3H, s, CH_3), 1.60 (6H, s, CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 139.79 (s, *C*q), 131.78 (s, *C*q), 130.55, 129.79 & 123.51 (3 x *C*H), 124.21 (s, $(CH_3)_2CCH$), 59.52 (s, $CHCH_2OH$), 39.67, 32.93, 31.01 & 28.30 (s, 4 x allylic CH_2), 25.86 (CH_3), 17.89 (CH_3), 16.42(CH_3).
TLC (Hexane: EtOAc = 4: 1) R_f = 0.10.
HRMS (EI^+, [M]): calculated for $[C_{14}H_{22}]^+$; 190.1722, found 190.1722.

(2Z,6E)-3,11-dimethyldodeca-2,6,10-trien-1-ol (218) - not used

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.48 5.33 (3H, m, 3 x C*H*), 5.14 – 5.07 (1H, m, $(CH_3)_2CCH$), 4.10 (2H, d, *J* = 7.1 Hz, $CHCH_2OH$), 2.16 – 1.97 (8H, m, 4 x allylic CH_2), 1.74 (3H, s, CH_3), 1.69 (3H, s, CH_3), 1.60 (6H, s, CH_3).
^{13}C NMR (126 MHz, $CDCl_3$) δ 139.91 (s, C_q), 131.83 (s, C_q), 131.08, 129.62 & 124.55 (3 x *C*H), 124.17 (s, $(CH_3)_2CCH$), 59.18 (s, $CHCH_2OH$), 32.91, 32.18, 31.26 & 28.25 (s, x allylic CH_2), 25.84 (CH_3), 23.57 (CH_3), 17.87 (CH_3).
TLC (Hexane: EtOAc = 4: 1) R_f 0.15.
HRMS (EI^+, [M]): calculated for $[C_{14}H_{22}]^+$; 190.1722, found 190.1721.

(2*E*,6*E*)-3,11-dimethyldodeca-2,6,10-trien-1-yl diphosphate, ([(2*E*, 6*E*)-7-hydrogen-FDP], 219)



(**217**) (**219**)

Bis-triethylammonium phosphate (**145**, 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 5 mL) at 5 minutes intervals to a solution of farnesol derivate **212** (60 mg, 0.3 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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 xx mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound **203** as a light fluffy powder (41 mg, 34%).

$_1$H NMR (500 MHz, $CDCl_3$) δ 5.49 5.33 (3H, m, 3 x C*H*), 5.16 – 5.11 (1H, m, $(CH_3)_2CCH$), 4.39 (2H, t, *J* = 6.5 Hz, $CHCH_2O$), 2.17 – 1.88 (8H, m, 4 x allylic CH_2), 1.63 (3H, s, CH_3), 1.61 (3H, s, CH_3), 1.53 (3H, s, CH_3).

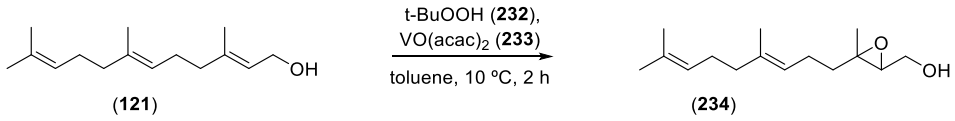
^{13}C NMR (126 MHz, $CDCl_3$) δ 143.21 (s, C_q), 134.19 (s, C_q), 131.27, 130.97 & 124.92 (s, 3 x *C*H), 120.41 (d, *J* = 8.5 Hz, $CHCH_2O$), 39.34, 32.56, 30.68 & 27.81 (s, 4 x allylic CH_2), 25.36 (CH_3), 17.53 (CH_3), 16.15 (CH_3).

^{31}P NMR (162 MHz, D_2O) δ –6.67 (d, *J* = 21.8 Hz), –10.42 (d, *J* = 22.2 Hz).

HRMS (ES^-, [M]): calculated for $[C_{14}H_{25}O_7P_2]^-$; 367.1076, found 367.1090.

7.2.6. Synthesis of 2,3-thiirane farnesyl diphosphate [2,3-Thii-FDP] (241)

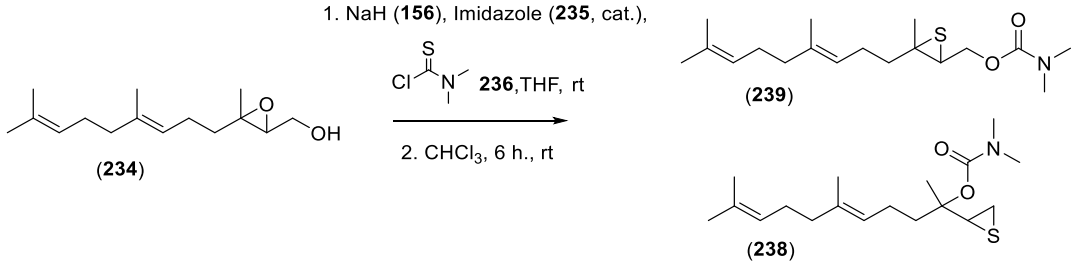
(*E*)-(3-(4,8-dimethylnona-3,7-dien-1-yl)-3-methyloxiran-2-yl)methanol (234)[253]



To a stirred solution of *E*,*E*-farnesol (**121**, 2.4 g, 10.8 mmol, 1 eq.) in anhydrous toluene (40 mL) at 10 ^{0}C was added a catalytic amount of vanadyl acetoacetate (**233**, 144 mg, 0.54 mmol, 0.05 eq.). Then, tert-butyl-hydroperoxide (**232**, 1.9 mL, 11.9 mmol, 70% in H_2O, 1.1 eq.) was added dropwise over 10 min. The reaction was stirred at 10 ºC for 1 h 45 min and then was quenched with NaOH/brine (20 mL, 1:1). The aqueous layer was washed with EtOAc (3 x 20 mL) and the organic extracts were washed with water (3 x 20 mL) and brine (3 x 20 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20% ethyl acetate/hexane yielded epoxide **234** (2.06 g, 80%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.06 – 4.98 (2H, m, 2 x C*H*CH_2), 3.76 (1 H, ddd, *J* = 11.6, 7.1, 4.2 Hz, C(O)HC*H*$_2$OH), 3.61 (1 H, ddd, *J* = 12.1, 6.9, 4.4 Hz, C(O)HC*H*$_2$OH), 2.92 (1H, dd, *J* = 6.8, 4.2, C(CH_3)C(O)*H*CH_2), 2.09 – 1.87 (6H, m, 2 x allylic C*H*$_2$), 1.61 (3H, s, C*H*$_3$), 1.67 – 1.36 (2H, m, CH_2C*H*$_2$C(CH_3)C(O)), 1.54 (3H, s, C*H*$_3$), 1.53 (3H, s, C*H*$_3$), 1.24 (3H, s, C*H*$_3$).
^{13}C NMR (101 MHz, $CDCl_3$) δ 135.83 (s, *C*$_q$), 131.49 (s, *C*$_q$), 124.20 & 123.19 (s, , 2 x C*C*HCH_2), 62.98 (s, C(CH_3)*C*(O)HCH_2), 61.48 (s, C(O)H*C*H_2OH), 61.24 (s, CH_2*C*(CH_3)C(O)H), 39.67, 26.63 & 23.61 (s, 3 x allylic *C*H_2), 38.52 (s, CH_2*C*H_2C(CH_3)C(O)), 25.72 (s, *C*H_3),17.71 (s, *C*H_3), 16.80 (s, *C*H_3), 16.00 (s, *C*H_3).
HRMS (EI$^+$, [M]): calculated for $[C_{15}H_{26}O_2]^+$; 238.1933, found 238.1931.

(*E*)-(3-(4,8-dimethylnona-3,7-dien-1-yl)-3-methylthiiran-2-yl)methyl dimethylcarbamate (239) [252]



To a stirred solution sodium hydride (**156**, 335 mg, 8.4 mmol, 2 eq.) in THF (20 mL) was added a solution of epoxide **234**(1 g, 4.2 mmol, 1 eq.) and imidazole (**235**, 29 mg, 0.24 mmol, 0.1 eq.) in THF (20 mL). After 15 min, a solution of dimethylcarbamoyl chloride (**236**, 0.57 g, 4.62 mmol, 11 eq.) in THF (10 mL) was added and stirring was continued for 1.5 h. The resulting mixture was diluted with hexane (30 mL) and washed with water (3 x 20 mL) and brine (3 x 20mL). Extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. The crude obtained was diluted with $CHCl_3$ (20 mL) and stirred for 6h (TLC indicated complete disappearance of the initial product). The solvent was removed under reduced pressure and the residue obtained was purified by silica gel flash chromatography using 20% ethyl acetate/hexane to yield a 1:1 mix (GC and ^{1}H NMR) of thiiranes **232** and **233** (904 mg, 66%). Silica gel flash chromatography using 2% ethyl acetate/hexane partially separated compounds **239** (329 mg, 24%) and **238** (275 mg, 20%).

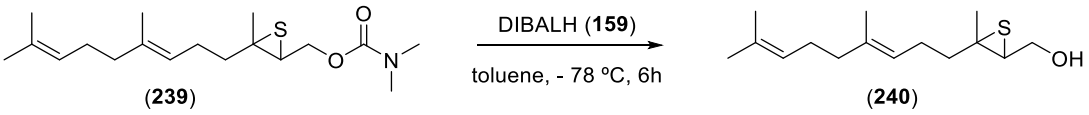
(*E*)-(3-(4,8-dimethylnona-3,7-dien-1-yl)-3-methylthiiran-2-yl)methyl dimethylcarbamate (239)

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.15 – 5.04 (2H, m, 2 x CC*H*CH_2), 4.47 (1H, dd, *J* = 11.7, 6.0 Hz, C(S)HC*H*$_2$O), 4.10 (1H, dd, *J* = 11.7, 8.1 Hz, C(S)HC*H*$_2$O), 3.04 (1H, dd, *J* = 8.1, 6.1 Hz, C(CH_3)C(S)*H*CH_2), 2.93 (6H, s, C(O)N(C*H*$_3$)$_2$), 2.22 – 1.92 (6H, m, 3 x allylic C*H*$_2$), 1.76 – 1.68 (2H, m, CH_2C*H*$_2$C(CH_3)C(S)), 1.67 (3H, d, *J* = 1.0 Hz, C*H*$_3$), 1.60 (3H, s, C*H*$_3$), 1.59 (6H, s, 2 x C*H*$_3$).
^{13}C NMR (101 MHz, $CDCl_3$) δ 156.21 (s, *C*O), 136.00 (s, *C*$_q$), 131.57 (s, *C*$_q$), 124.33 & 123.18 (s, 2 x C*C*HCH_2), 66.05 (s, C(S)H*C*H_2O), 50.60 (s, CH_2*C*(CH_3)C(S)H), 44.54 (s, C(CH_3)*C*(S)HCH_2), 43.85 (s, CH_2*C*H_2C(CH_3)C(S)), 39.78, 26.76 & 26.42 (s, 3 x allylic *C*H_2), 36.61 & 36.06 (s,2 x C(O)N(*C*H_3)$_2$), 25.83 (s, *C*H_3), 20.36 (s, *C*H_3), 17.82 (s, *C*H_3), 16.10 (s, *C*H_3).
HRMS (EI$^+$, [M]): calculated for [C_{18}H32NO_2S]$^+$; 326.2154, found 326.2155.

(*E*)-6,10-dimethyl-2-(thiiran-2-yl)undeca-5,9-dien-2-yl dimethylcarbamate (238) - not used

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.14 (1H, t, *J* = 7.0 Hz, CC*H*CH_2), 5.09 (1H, t, *J* = 6.9 Hz, CC*H*CH_2), 3.27 (1H, t, *J* = 6.1 Hz CC(S)*H*), 2.88 (6H, s, C(O)N(C*H*$_3$)$_2$), 2.60 (1H, dd, *J* = 5.7, 1.4 Hz, 1H, CC(S)HC*H*$_2$), 2.51 (1H, dd, *J* = 6.5, 1.3 Hz, CC(S)HC*H*$_2$), 2.24 – 1.84 (8H, m, 4 x allylic C*H*$_2$), 1.68 (3H, s, C*H*$_3$), 1.61 (3H, s, C*H*$_3$), 1.60 (3H, s, C*H*$_3$), 1.36 (3H, s, C*H*$_3$).
^{13}C NMR (126 MHz, $CDCl_3$) δ 155.53 (s, *C*O), 135.70 (s, *C*$_q$), 131.55 (s, *C*$_q$), 124.41 & 123.75 (s, 2 x C*C*HCH_2), 82.16 (s, CH_2*C*(CH_3)C(S)H), 41.03, 39.82, 26.82 & 22.20 (s, 4 x allylic C*H*$_2$), 40.92 (s, CC(S)*H*), 36.31 (s, C(O)N(*C*H_3)$_2$), 25.83 (s, *C*H_3), 23.20 (s, CC(S)HC*H*$_2$), 19.85 (s), 17.83 (s, *C*H_3), 16.07 (s, *C*H_3).
HRMS (EI$^+$, [M]): calculated for [C_{18}H32NO_2S]$^+$; 326.2154, found 326.2149.

(*E*)-(3-(4,8-dimethylnona-3,7-dien-1-yl)-3-methylthiiran-2-yl)methanol (240)



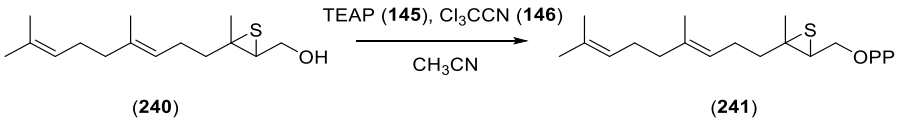
A hexane solution of DIBALH (**159**, 5.4 mL,5.4 mmol, 1 M, 5 eq.) was added dropwise to carbamate **232** (350 mg, 1.1 mmol, 1eq.) in anhydrous toluene (20 mL) at -78 ºC under argon. The mixture was stirred for 6 hours and quenched *via* addition of Rochelle salts at 0 °C. Organic products extracted with DCM (3 x 20 mL). The organic extracts were washed with water (3 x 10 mL) and brine (3 x 10 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 10 % ethyl acetate/hexane as eluent gave alcohol **234** as a colourless oil. (105 mg, 38%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.16 – 5.04 (2H, m, 2 x CC*H*CH_2), 3.86 (2H, d, *J* = 5.8 Hz, C(S)HC*H*$_2$O), 3.05 (1H, t, *J* = 6.5 Hz, C(CH_3)C(S)*H*CH_2), 2.27 – 1.94 (6H, m, 3 x allylic C*H*$_2$), 1.83 – 1.69 (2H, m, CH_2C*H*$_2$C(CH_3)C(S)), 1.68 (3H, d, *J* = 1.1 Hz, C*H*$_3$), 1.61 (3H, s, *J* = 0.6 Hz C*H*$_3$), 1.60 (6H, s, 2 x C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 136.09 (s, C_q), 131.63 (s, C_q), 124.34 & 123.23 (s, 2 x C*C*HCH_2), 63.07 (s, C(S)H*C*H$_2$O), 51.41 (s, CH_2*C*(CH_3)C(S)H), 49.57 (s, C(CH_3)*C*(S)HCH_2), 43.84 (s, CH_2*C*H$_2$C(CH_3)C(S)), 39.80, 26.74 & 26.33 (s, 3 x allylic *C*H$_2$), 25.84 (s, *C*H$_3$), 20.24 (s, *C*H$_3$), 17.83 (s, *C*H$_3$), 16.17 (s, *C*H$_3$).

HRMS (ES^+): calculated for $[C_{15}H_{27}OS]^+$; 255.1783, found 255.1788.

(*E*)-1-((3-(4,8-dimethylnona-3,7-dien-1-yl)-3-methylthiiran-2-yl)methoxy) diphosphate [(2,3-thiirane-FDP], 241)



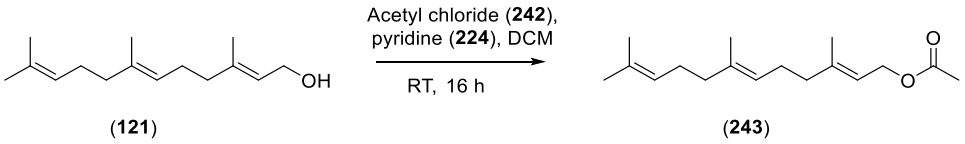
Bis-triethylammonium phosphate (TEAP, **145**) was prepared immediately prior to use by the dropwise addition of a solution of phosphoric acid in acetonitrile to a stirring solution of triethylamine in acetonitrile. TEAP was then added in three portions (3 x 5 mL) at 5 minutes intervals to a solution of **240** (50 mg, 0.2 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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 : H2O, 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 30 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound as a light fluffy powder (45 mg, 48%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.23 – 5.11 (2H, m, 2 x CC*H*CH_2), 4.26 – 4.17 (1H, m, C(S)HC*H*$_2$O), 4.08 – 3.98 (1H, m, C(S)HC*H*$_2$O), 3.17 (1H, t, *J* = 6.8 Hz, C(CH_3)C(S)*H*CH_2), 2.29 – 1.85 (8H, m, 4 x allylic C*H*$_2$), 1.63 (3H, s, C*H*$_3$), 1.59 (3H, s, C*H*$_3$), 1.56 (6H, s, 2 x C*H*$_3$).
^{13}C NMR (101 MHz, $CDCl_3$) δ 136.78 (s, *C*$_q$), 131.89 (s, *C*$_q$), 124.47 & 123.74 (s, 2 x C*C*HCH_2), 66.59 (s, C(S)H*C*H_2O), 52.57 (s, CH_2*C*(CH_3)C(S)H), 45.96 (s, C(CH_3)*C*(S)HCH_2), 42.62 (s, CH_2*C*H_2C(CH_3)C(S)), 38.74, 25.65 & 25.63 (s, 3 x allylic *C*H_2), 24.82 (s, *C*H_3), 19.01 (s, *C*H_3), 16.93 (s, *C*H_3), 15.16 (s, *C*H_3).
^{31}P NMR (202 MHz, D_2O) δ –10.73 (d, *J* = 17.1 Hz), –11.54 (d, *J* = 17.2 Hz).
HRMS (ES$^-$, [M]): calculated for [$C_{15}H_{27}O_7P_2S$]$^-$; 413.0953, found 413.0960.

7.2.7. Synthesis of 10,11-thiiranefarnesyl diphosphate ([10,11-Thii-FDP], 250).

(2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl acetate (243)[213]

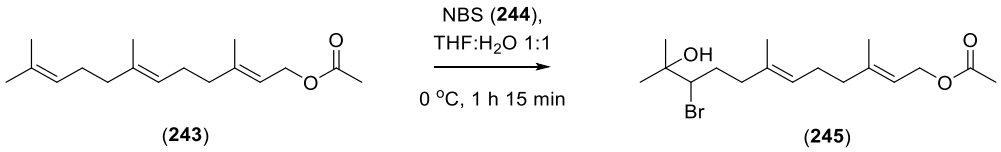


(**121**) (**243**)

(2*E*,6*E*)-farnesol (**121**, 2.4 g, 11 mmol, 1eq.), pyridine (**224**, 1.1 mL, 13 mmol, 1.2 eq.), acetyl chloride (**242**, 0.93 mL, 13 mmol, 1.2 eq.) in DCM (50 mL) were stirred at room temperature under argon for 16 h. The resulting mixture was washed with saturated $NaHCO_3$ solution (3 x 20 mL), then H_2O (3 x 10 mL) and the combined aqueous extracts re-extracted with diethyl ether (3 x 25 mL). The combined organic extracts were dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20% ethyl acetate/hexane as eluent gave farnesyl acetate **243** as a colourless oil. (2.68 g, 92%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.37 – 5.31 (1H, m, CC*H*CH_2O), 5.13 – 5.05 (2H, m, 2 x CC*H*CH_2), 4.58 (2H, d, *J* = 7.1 Hz, CCHC*H*$_2$O), 2.15 – 1.94 (8H, m, 4 x allylic C*H*$_2$), 2.05 (3H, s, COC*H*$_3$), 1.70 (3H, d, *J* = 0.6 Hz, C*H*$_3$), 1.67 (3H, d, *J* = 1.0 Hz, C*H*$_3$), 1.59 (6H, 2 x C*H*$_3$).
^{13}C NMR (101 MHz, $CDCl_3$) δ 171.29 (*C*O), 142.44 (*C*$_q$), 135.60 (*C*$_q$), 131.47 (*C*$_q$), 124.43 & 123.73 (s, 2 x C*C*HCH$_2$) 118.36 (s, C*C*HCH$_2$O), 61.54 (s, CCH*C*H$_2$O), 39.82, 39.65, 26.83 & 26.30 (s, 4 x allylic *C*H$_2$), 25.83 (s, *C*H$_3$), 21.20 (s, CO*C*H$_3$), 17.82 (s, *C*H$_3$), 16.60 (s, *C*H$_3$), 16.14 (s, *C*H$_3$).
HRMS (EI$^+$, [M]): calculated for $[C_{17}H_{28}O_2]^+$; 264.2089, found 264.2097.

(2*E*,6*E*)-10-bromo-11-hydroxy-3,7,11-trimethyldodeca-2,6-dien-1-yl acetate (245)[213]



(**243**) (**245**)

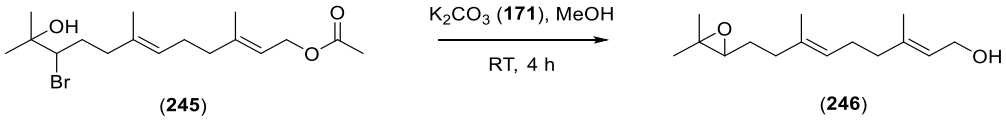
Farnesyl acetate **243** (2.4 g, 9.1 mmol, 1 eq.) was dissolved in THF: H2O (50 mL, 1:1) and cooled to 0 °C. Then, N-bromosuccinimide (**244**, 1.96 g, 10.9 mmol, 1.2 eq.) was added portion wise over 15 min at 0 °C. Once the addition was complete, the suspension was stirred for 1 h at 0 °C. After this time, the mixture was concentrated under reduced pressure and the resulting residue was taken into Et_2O (100 mL). The organic mixture was washed with water (3 x 50 mL) and brine (3 x 50 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel

flash chromatography using 20% ethyl acetate/hexane afforded the bromohydrin **245** (1.9 g, 48%) as a pale-yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.36 – 5.30 (1H, m, CC*H*CH_2O), 5.17 (1H, t, *J* = 6.8 Hz, CC*H*CH_2), 4.58 (2H, d, *J* = 7.1 Hz, CCHC*H*$_2$O), 3.95 (1H, dd, *J* = 11.4, 1.9 Hz, C(Br)C*H*(OH)), 2.35 – 2.03 (6H, m, 3 x C*H*$_2$), 2.04 (3H, s, COC*H*$_3$), 2.02 – 1.91 (1H, m, C(Br)CH(OH)C*H*$_2$), 1.82 – 1.71 (1H, m, C(Br)CH(OH)C*H*$_2$), 1.69 (3H, d, *J* = 0.6 Hz, C*H*$_3$), 1.58 (3H, d, *J* = 0.6 Hz, C*H*$_3$), 1.33 (3H, C*H*$_3$), 1.32 (3H, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.30 (*C*O), 142.09 (C_q), 135.65 (C_q), 125.47 (s, C*C*HCH$_2$) 118.54 (s, C*C*HCH$_2$O), 72.55 (s, *C*Br) 70.82 (s, C(Br)*C*H(OH)), 61.53 (s, CCH*C*H$_2$O), 32.10 (s, C(Br)CH(OH)*C*H$_2$) 39.47, 38.23 & 26.23 (s, 3 x *C*H$_2$), 26.65(s, *C*H$_3$), 21.20 (s, CO*C*H$_3$), 26.00 (s, *C*H$_3$), 16.58 (s, *C*H$_3$), 15.95 (s, *C*H$_3$).

(2*E*,6*E*)-9-(3,3-dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-ol (246)[213]



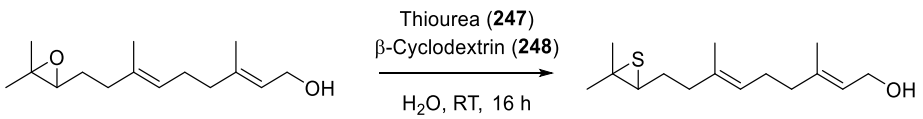
Bromohydrin **245** (1.9 g, 5.26 mmol, 1 eq.) was taken up in MeOH (50 mL) and K_2CO_3 (**171**, 1.45 g, 10.52 mmol, 2 eq.) was added to the solution. The reaction mixture was stirred for 4 h at room temperature and then, the solvent was removed under reduced pressure. The resulting residue was poured into H_2O (100 mL) and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine (3 x 30 mL), dried over anhydrous sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 20 % ethyl acetate/hexane gave 10,11- epoxy farnesol **246** (850 mg, 68%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.43 – 5.37 (1H, m, CC*H*CH_2O), 5.18 – 5.12 (1H, m, CC*H*CH_2), 4.13 (2H, d, J = 6.8 Hz, CHC*H*$_2$O), 2.69 (1H, t, J = 6.2 Hz, CC(O)*H*CH_2), 2.19 – 2.00 (6H, m, 3 x C*H*$_2$), 1.63 (3H, d, *J* = 0.6 Hz, C*H*$_3$), 1.65 – 1.59 (2H, m, C(O)C*H*$_2$CH$_2$), 1.61 (3H, s, C*H*$_3$), 1.29 (3H, s, C*H*$_3$), 1.25 (3H, s, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 139.45 (s, C_q), 134.46 (s, C_q), 124.64 (s, C*C*HCH$_2$), 123.73 (s, C*C*HCH$_2$O), 64.30 (s, C*C*(O)HCH$_2$), 59.46 (s, CH*C*H$_2$O), 58.55 (s, (CH$_3$)$_2$*C*C(O)), 39.51, 36.45 & 26.24 (s, 3 x *C*H$_2$), 27.36 (s, C(O)*C*H$_2$CH$_2$), 24.99 (s, *C*H$_3$), 18.90 (s, *C*H$_3$), 16.35 (s, *C*H$_3$), 16.09 (s, *C*H$_3$).

HRMS (EI$^+$, [M – H_2O]): calculated for $[C_{15}H_{24}O]^+$; 220.1827, found 220.1827.

(2*E*,6*E*)-9-(3,3-dimethylthiiran-2-yl)-3,7-dimethylnona-2,6-dien-1-ol (249)

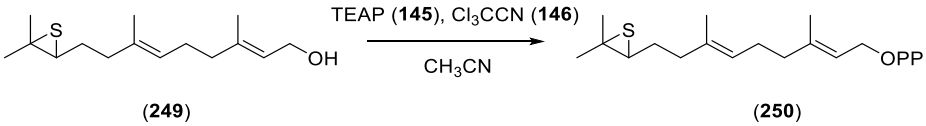


(**246**) (**249**)

To a stirred solution of β-cyclodextrin (**248**, 3.84 g, 3.4 mmol, 1 eq.) in water (50 mL) at 60 ºC was slowly added a solution of epoxide **246** (800 mg, 3.4 mmol, 1 eq.) in acetone (5 mL). After addition, the mixture was cooled to room temperature, thiourea (**247**, 516 mg, 6.8 mmol, 2 eq.) was added and the reaction was stirred for 16 h at room temperature. The resulting mixture was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with water (3 x 10 mL), brine (3 x 10 mL), dried over sodium sulphate, filtered and concentrated under reduced pressure. Purification by silica gel flash chromatography using 2% ethyl acetate/hexane gave 10,11-thiirane farnesol **249** (104 mg, 12%).

^{1}H NMR (500 MHz, CDCl3) δ 5.40 – 5.35 (1H, m, CC*H*CH$_2$O), 5.19 – 5.13 (1H, m, CC*H*CH$_2$), 4.13 (2H, d, *J* = 6.8 Hz, CHC*H*$_2$O), 3.36 (1H, dd, *J* = 10.6, 1.8 Hz, CC(S)*H*CH$_2$), 2.22 – 2.05 (6H, m, 3 x C*H*$_2$), 1.64 (3H, s, C*H*$_3$), 1.61 (3H, s, C*H*$_3$), 1.60 – 1.54 (1H, m, C(S)C*H*$_2$CH$_2$), 1.44 – 1.38 (1H, m, C(S)C*H*$_2$CH$_2$), 1.17 (3H, s, C*H*$_3$), 1.14 (3H, s, C*H*$_3$).
^{13}C NMR (101 MHz, CDCl$_3$) δ 138.83 (s, *C*$_q$), 135.12 (s, *C*$_q$), 125.26 (s, C*C*HCH$_2$), 124.25 (s, C*C*HCH$_2$O), 77.73 (s, C*C*(S)HCH$_2$), 73.21 (s, (CH$_3$)$_2$*C*C(S)), 59.38 (s, CH*C*H$_2$O), 39.37, 36.65 & 25.65 (s, 3 x *C*H$_2$), 29.01 (s, C(S)*C*H$_2$CH$_2$), 26.49 (s, *C*H$_3$), 23.30 (s, *C*H$_3$), 15.99 (s, *C*H$_3$), 15.85 (s, *C*H$_3$).
HRMS (EI$^+$, [M]): calculated for [C$_{15}$H$_{26}$SO]$^+$; 254.1704, found 254.1698.

1-(((2*E*,6*E*)-9-(3,3-dimethylthiiran-2-yl)-3,7-dimethylnona-2,6-dien-1-yl)oxy) diphosphate (250)



(**249**) (**250**)

Bis-triethylammonium phosphate (**145**, 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 5 mL) at 5 minutes intervals to a solution of **249** (50 mg, 0.2 mmol) in trichloroacetonitrile (**146**, 5 mL) at 37 °C. The reaction was incubated at 37 ºC 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$_4$OH: H$_2$O, 6 : 2.5 : 0.5) was then begun and fractions collected and analysed by TLC (isopropanol : conc. NH$_4$OH : H$_2$O, 6 : 3 : 1, visualised with basic KMnO$_4$), and those containing the diphosphate were

combined, ammonia and isopropanol removed under vacuum and the resulting aqueous solution diluted to 10 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound **250** as a light fluffy powder. The yield is not reported due to the presence of impurities that could not be removed under the conditions used.

^{1}H NMR (500 MHz, D_2O) δ 5.46 (1H, t, *J*= 6.8 Hz, CC*H*CH_2O), 5.30 (1H, t, *J* = 7.0 Hz, CC*H*CH_2), 4.42 (2H, t, *J* = 6.5 Hz, CHC*H*$_2$O), 4.15 (1H, d, *J* = 10.8, CC(S)*H*CH_2), 2.28 – 2.03 (6H, m, 3 x C*H*$_2$), 1.73 (3H, s, C*H*$_3$), 1.66 (3H, s, C*H*$_3$), 1.74 – 1.44 (2H, m, C(S)C*H*$_2$$CH_2$), 1.38 (3H, s, C*H*$_3$), 1.33 (3H, s, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 143.17 (s, C_q), 135.58 (s, C_q), 124.93 (s, C*C*HCH_2), 119.42 (s, C*C*HCH_2O), 84.97 (d, *J*=25.95, C*C*(S)HCH_2 62.76 (s, CH*C*H_2O), 38.77, 35.55 & 28.78 (s, 3 x *C*H$_2$), 28.09 (d, C(S)*C*$H_2$$CH_2$), 25.19 (d, *C*H$_3$), 21.31 (s, *C*H$_3$), 15.67 (s, *C*H$_3$), 15.14 (s, *C*H$_3$).

CHAPTER 8

REFERENCES

[1] Connolly, J. D.; Hill, R. A. *Dictionary of Terpenoids*; Chapman and Hall, London, **1991**.

[2] *Nobel Lectures*, *Chemistry 1901-1921*, Elsevier Publishing Company, Amsterdam, **1966**.

[3] L. Ruzicka, *Experientia* **1953**, *9*, 357–367.

[4] J. Gershenzon, N. Dudareva, *Nat. Chem. Biol.* **2007**, *3*, 408–414.

[5] W. Brandt, L. Bräuer, N. Günnewich, J. Kufka, F. Rausch, D. Schulze, E. Schulze, R. Weber, S. Zakharova, L. Wessjohann, *Phytochemistry* **2009**, *70*, 1758–1775.

[6] R. A. Demel, B. de Kruyff, *Biochim. Biophys. Acta* **1976**, *457*, 109–132.

[7] I. Schuler, A. Milont, Y. Nakatanif, G. Ourissono, A. Albrecht, P. Benveniste, M. Hartmann, *Proc. Nati. Acad. Sci.* **1991**, *88*, 6926–6930.

[8] P. Benveniste, *Annu. Rev. Plant Biol.* **2004**, *55*, 429–457.

[9] S. Deng, T. Wei, K. Tan, M. Hu, F. Li, Y. Zhai, S. Ye, *Sci China Life Sci* **2016**, *59*, 183–193.

[10] H. Schaller, *Progress in Lipid Research* **2003**, *42*, 163–175.

[11] E. J. Dufourc, *J. Chem. Biol.* **2008**, *1*, 63–77.

[12] G. E. Bartley, P. A. Scolnik, *The Plant Cell* **1995**, *7*, 1027–1038.

[13] A. V. Rao, L. G. Rao, *Pharmacol. Res.* **2007**, *55*, 207–216.

[14] J. Fiedor, K. Burda, *Nutrients* **2014**, *6*, 466–488.

[15] G. Sandmann, M. Albrecht, G. Schnurr, O. Knörzer, P. Böger, *Trends Biotechnol.* **1999**, *17*, 233–237.

[16] H. Nohl, W. Jordan, R. J. Youngman, *Adv. Free Radical Biol. Med.* **1986**, *2*, 211–279.

[17] E. J. Son, J. H. Kim, K. Kim, C. B. Park, *J. Mater. Chem.* **2016**, *4*, 11179–11202.

[18] R. Croteau, T. M. Kutchan, N. G. Lewis, *American Society of Plant Physiologists* **2000**.

[19] J. H. Langenheim, *J. Chem. Ecol.* **1994**, *20*, 1223–1280.

[20] F. Chen, D. Tholl, J. C. D'Auria, A. Farooq, E. Pichersky, J. Gershenzon, *The Plant Cell* **2003**, *15*, 481–494.

[21] J. P. Morrissey, A. E. Osbourn, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 708–724.

[22] A. Quintana, J. Reinhard, R. Faure, P. Uva, A. Eve, G. Massiot, *J. Chem. Ecol.* **2003**, *29*, 639–652.

[23] V. J. Paul, M. P. Puglisi, R. Ritson-Williams, *Nat. Prod. Rep.* **2006**, *23*, 153–180.

[24] P. Bhadury, P. C. Wright, *Planta* **2004**, 561–578.

[25] N. Dudareva, F. Negre, D. A. Nagegowda, I. Orlova, *Crit. Rev. Plant Sci.* **2006**, *25*, 417–440.

[26] J. A. Pickett, R. K. Allemann, M. A. Birkett, *Nat. Prod. Rep.* **2013**, *30*, 1277–1283.

[27] G. M. Cragg, D. J. Newman, *Journal of Ethnopharmacology* **2005**, *100*, 72–79.

[28] R. B. Croteau, E. M. Davis, K. L. Ringer, M. R. Wildung, *Naturwissenschaften* **2005**, *92*, 562–577.

[29] V. Srivastava, A. S. Negi, J. K. Kumar, M. M. Gupta, S. P. S. Khanuja, *Bioorg. Med. Chem.* **2005**, *13*, 5892–5908.

[30] W. Schwab, R. Davidovich-Rikanati, E. Lewinsohn, *The Plant Journal* **2008**, *54*, 712–732.

[31] D. W. Christianson, *Science* **2007**, *316*, 60–61.

[32] D. Tholl, *Curr. Opin. Plant Biol.* **2006**, *9*, 297–304.

[33] D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.

[34] J. C. Sacchettini, C. D. Poulter, *Science* **1997**, *277*, 1788–1789.

[35] H. M. Miziorko, *Arch. Biochem. Biophys.* **2011**, *505*, 131–143.

[36] J. Lombard, D. Moreira, *Mol. Biol. Evol.* **2011**, *28*, 87–99.

[37] D. J. Mcgarvey, R. Croteau, *The Plant Cell* **1995**, *7*, 1015–1026.

[38] E. I. Wilding, J. R. Brown, A. P. Bryant, A. F. Chalker, D. J. Holmes, K. A. Ingraham, S. Iordanescu, C. Y. So, M. Rosenberg, M. N. Gwynn, *J. Bacteriol.* **2000**, *182*, 4319–4327.

[39] R. K. Wierenga, A. M. Haapalainen, G. Merila, *Trends Biochem. Sci.* **2006**, *31*, 64–71.

[40] B. J. Bahnson, *Proc. Nati. Acad. Sci.* **2004**, *101*, 16399–16400.

[41] M. Hedl, L. Tabernero, C. V. Stauffacher, V. W. Rodwell, *J. Bacteriol.* **2004**, *186*, 1927–1932.

[42] P. Nyati, C. Rivera-Perez, F. G. Noriega, *PloS One* **2015**, *10*, 1–14.

[43] T. J. Herdendorf, H. M. Miziorko, *Biochemistry* **2006**, *45*, 3235–3242.

[44] K. Bloch, S. Chaykin, A. H. Phillips, A. de Waard, *J. Biol. Chem.* **1959**, *234*, 2595–2604.

[45] K. Clifford, J. W. Cornforth, R. Mallaby and G. T. Phillips, *Chem. Commun.* **1971**, 1599–1600.

[46] I. P. Street, D. J. Christensen, C. D. Poulter, *J. Am. Chem. Soc.* **1990**, *112*, 8577–8578.

[47] P. Liao, A. Hemmerlin, T. J. Bach, M. Chye, *Biotechnol. Adv.* **2016**, *34*, 697–713.

[48] M. Rohmer, M. Knani, P. Simonin, B. Sutter, H. Sahm, *Biochem J.* **1993**, *295*, 517–524.

[49] D. Arigoni, S. Sagner, C. Latzel, W. Eisenreich, A. Bacher, M. H. Zenk, *Proc. Nati. Acad. Sci.* **1997**, *94*, 10600–10605.

[50] M. Rohmer, M. Seemann, S. Horbach, S. Bringer-Meyer, H. Sahm, L. P. Cnrs, I. Le Bel, B. Pascal, D. C. De Mulhouse, D. Ju, *J. Am. Chem. Soc.* **1996**, *118*, 2564–2566.

[51] L. M. Lois, N. Campos, S. R. Putra, K. Danielsen, M. Rohmer, A. Boronat, *Proc. Natl. Acad. Sci.* **1998**, *95*, 2105–2110.

[52] T. Kuzuyama, S. Takahashi, H. Watanabe, *Tetrahedron Lett.* **1998**, *39*, 4509–4512.

[53] T. Kuzuyama, M. Takagi, K. Kaneda, H. Seto, *Tetrahedron Lett.* **2000**, *41*, 703–706.

[54] T. Kuzuyama, M. Takagi, K. Kaneda, H. Watanabe, *Tetrahedron Lett.* **2000**, *41*, 2925–2928.

[55] M. Takagi, T. Kuzuyama, K. Kaneda, H. Watanabe, *Tetrahedron Lett.* **2000**, *41*, 3395–3398.

[56] S. Hecht, W. Eisenreich, P. Adam, S. Amslinger, K. Kis, A. Bacher, D. Arigoni, F. Rohdich, *Proc. Natl. Acad. Sci.* **2001**, *98*, 14837–14842.

[57] A. Kollas, E. C. Duin, M. Eberl, B. Altincicek, M. Hintz, A. Reichenberg, D. Henschker, A. Henne, I. Steinbrecher, D. N. Ostrovsky, *FEBS Lett.* **2002**, *532*, 432–436.

[58] P. Adam, C. Krieger, S. Amslinger, D. Arigoni, F. Rohdich, S. Hecht, K. Ga, A. Bacher, W. Eisenreich, *Proc. Natl. Acad. Sci.* **2001**, *99*, 1158–1163.

[59] M. Wol, M. Seemann, B. Tse, S. Bui, Y. Frapart, D. Tritsch, A. Boronat, A. Garcia, M. Rodr, M. Rohmer, *FEBS Lett.* **2003**, *541*, 115–120.

[60] C. C. Burke, M. R. Wildung, R. Croteau, *Proc. Natl. Acad. Sci.* **1999**, *96*, 13062–13067.

[61] T. Chang, F. Hsieh, T. Ko, K. Teng, P. Liang, *The Plant Cell* **2010**, *22*, 454–467.

[62] L. C. Tarshis, M. Yan, C. D. Poulter, J. C. Sacchettini, *Biochemistry* **1994**, *33*, 10871–10877.

[63] K. L. Kavanagh, K. Guo, J. E. Dunford, X. Wu, S. Knapp, F. H. Ebetino, M. J. Rogers, R. G. G. Russell, U. Oppermann, *Proc. Natl. Acad. Sci.* **2006**, *103*, 7829–7834.

[64] D. P. Kloer, R. Welsch, P. Beyer, G. E. Schulz, *Biochemistry* **2006**, *45*, 15197–15204.

[65] T. Chang, R. Guo, T. Ko, P. Liang, *J. Biol. Chem.* **2006**, *281*, 14991–15000.

[66] J. S. Dickschat, *Nat. Prod. Rep.* **2011**, *28*, 1917–1936.

[67] Y. Gao, R. B. Honzatko, R. J. Peters, *Nat. Prod. Rep.* **2012**, *29*, 1153–1175.

[68] D. J. Miller, R. K. Allemann, *Nat. Prod. Rep.* **2012**, *29*, 60–71.

[69] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **1997**, *277*, 1815–1820.

[70] C. A. Lesburg, G. Zhai, D. E. Cane, D. W. Christianson, *Science* **1997**, *277*, 1820–1824.

[71] J. M. Caruthers, I. Kang, M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *J. Biol. Chem.* **2000**, *275*, 25533–25539.

[72] J. Pandit, D. E. Danley, G. K. Schulte, S. Mazzalupo, T. A. Pauly, C. M. Hayward, E. S. Hamanaka, J. F. Thompson, H. J. Harwood, J. M. Caruthers, I. Kang, M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *J. Biol. Chem.* **2000**, *275*, 25533–25539.

[73] M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Proc. Natl. Acad. Sci.* **2001**, *98*, 13543–13548.

[74] V. González, D. J. Grundy, J. A. Faraldos, R. K. Allemann, *Org. Biomol. Chem.* **2016**, *14*, 7451–7454.

[75] R. P. Mcandrew, P. P. Peralta-Yahya, A. Degiovanni, J. H. Pereira, M. Z. Hadi, J. D. Keasling, P. D. Adams, *Structure* **2011**, *19*, 1876–1884.

[76] D. E. Cane, Q. Xue, B. C. Fitzsimons, *Biochemistry* **1996**, *35*, 12369–12376.

[77] D. W. Christianson, *Chem. Rev.* **2006**, *106*, 3412–3442.

[78] H. A. Gennadios, V. Gonzalez, L. Di Costanzo, A. Li, F. Yu, D. J. Miller, R. K. Allemann, D. W. Christianson, *Biochemistry* **2009**, *48*, 6175–6183.

[79] I. Prosser, I. G. Altug, A. L. Phillips, W. A. König, H. J. Bouwmeester, M. H. Beale, *Arch. Biochem. Biophys.* **2004**, *432*, 136–144.

[80] E. Y. Shishova, L. Di Costanzo, D. E. Cane, D. W. Christianson, *Biochemistry* **2007**, *46*, 1941–1951.

[81] E. Y. Shishova, F. Yu, D. J. Miller, J. A. Faraldos, Y. Zhao, R. M. Coates, R. K. Allemann, D. E. Cane, D. W. Christianson, *J. Biol. Chem.* **2008**, *283*, 15431–15439.

[82] M. W. Van Der Kamp, J. Sirirak, J. Zìurek, R. K. Allemann, A. J. Mulholland, *Biochemistry* **2013**, *52*, 8094–8105.

[83] J. A. Aaron, X. Lin, D. E. Cane, D. W. Christianson, *Biochemistry* **2010**, *49*, 1787–1797.

[84] P. Baer, P. Rabe, C. A. Citron, C. C. De Oliveira Mann, N. Kaufmann, M. Groll, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 213–216.

[85] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652–7656.

[86] D. A. Dougherty, *Science* **1996**, *271*, 163–168.

[87] R. D. Kersten, J. K. Diedrich, J. R. Yates, J. P. Noel, *ACS Chem. Biol.* **2015**, *10*, 2501–2511.

[88] J. A. Faraldos, V. González, M. Senske, R. K. Allemann, *Org. Biomol. Chem.* **2011**, *9*, 6920–6923.

[89] J. P. Noel, N. Dellas, J. A. Faraldos, M. Zhao, B. A. Hess, L. Smentek, R. M. Coates, P. E. O'Maille, *ACS Chem. Biol.* **2010**, *5*, 377–392.

[90] C. L. Steele, J. Crock, J. Bohlmann, R. Croteau, *J. Biol. Chem.* **1998**, *273*, 2078–2089.

[91] W. Boland, S. Garms, T. G. Köllner, *J. Org. Chem.* **2010**, *75*, 5590–5600.

[92] A. Vattekkatte, S. Garms, W. Brandt, W. Boland, *Org. Biomol. Chem.* **2018**, *16*, 348–362.

[93] B. Felicetti, D. E. Cane, *J. Am. Chem. Soc.* **2004**, *126*, 7212–7221.

[94] J. A. Faraldos, V. Gonzalez, R. K. Allemann, *Chem. Commun.* **2012**, *48*, 3230–3232.

[95] M. J. Calvert, P. R. Ashton, R. K. Allemann, *J. Am. Chem. Soc.* **2002**, *124*, 11636–11641.

[96] M. J. Calvert, S. E. Taylor, R. K. Allemann, *Chem. Commun.* **2002**, *2*, 2384–2385.

[97] A. Deligeorgopoulou, R. K. Allemann, *Biochemistry* **2003**, *42*, 7741–7747.

[98] A. Deligeorgopoulou, S. E. Taylor, S. Forcat, R. K. Allemann, *Chem. Commun.* **2003**, *9*, 2162–2163.

[99] J. A. Faraldos, A. K. Antonczak, V. González, R. Fullerton, E. M. Tippmann, R. K. Allemann, *J. Am. Chem. Soc.* **2011**, *133*, 13906–13909.

[100] D. E. Cane, G. Yang, R. M. Coates, Hyung-Jung Pyun, Thomas M. Hohn, *J. Org. Chem.* **1992**, *57*, 3454–3462.

[101] M. Loizzi, D. J. Miller, R. K. Allemann, *Org. Biomol. Chem.* **2019**, *17*, 1206–1214.

[102] L. S. Vedula, Y. Zhao, R. M. Coates, T. Koyama, D. E. Cane, D. W. Christianson, *Arch. Biochem. Biophys.* **2007**, *466*, 260–266.

[103] L. S. Vedula, M. J. Rynkiewicz, H. J. Pyun, R. M. Coates, D. E. Cane, D. W. Christianson, *Biochemistry* **2005**, *44*, 6153–6163.

[104] J. A. Faraldos, B. Kariuki, R. K. Allemann, *J. Org. Chem.* **2010**, *75*, 1119–1125.

[105] J. A. Faraldos, R. K. Allemann, *Org. Lett.* **2011**, *13*, 1202–1205.

[106] D. E. Cane, P. C. Prabhakaran, J. S. Oliver, D. B. Mcllwaine, *J. Am. Chem. Soc.* **1990**, *112*, 3209–3210.

[107] P. M. Harrison, J. S. Oliver, D. E. Cane, *J. Am. Chem. Soc.* **1988**, *110*, 5922–5923.

[108] D. E. Cane, J. S. Oliver, P. M. Harrison, C. Abell, B. R. Hubbard, C. T. Kane, R. Lattman, *J. Am. Chem. Soc.* **1990**, *112*, 4513–4524.

[109] W. Weiner, E. Cane, *Can. J. Chem.* **1994**, *72*, 118–127.

[110] C. R. Benedict, J. Lu, D. W. Pettigrew, J. Liu, R. D. Stipanovic, H. J. Williams, *Plant Physiol.* **2001**, *125*, 1754–1765.

[111] C. O. Schmidt, H. J. Bouwmeester, S. Franke, W. A. König, *Chirality* **1999**, *11*, 353–362.

[112] D. E. Cane, H. Chiu, P. Liang, K. S. Anderson, *Biochemistry* **1997**, *36*, 8332–8339.

[113] S. Picaud, P. Mercke, X. He, O. Sterner, M. Brodelius, D. E. Cane, P. E. Brodelius, *Arch. Biochem. Biophys.* **2006**, *448*, 150–155.

[114] S. Kim, K. Heo, Y. Chang, S. Park, S. Rhee, S. Kim, *J. Nat. Prod.* **2006**, *69*, 758–762.

[115] D. E. Cane, M. Tandon, *Tetrahedron Lett.* **1994**, *35*, 5355–5358.

[116] D. E. Cane, M. Tandon, P. C. Prabhakaran, *J. Am. Chem. Soc.* **1993**, *115*, 8103–8106.

[117] D. E. Cane, M. Tandon, *J. Am. Chem. Soc.* **1995**, *117*, 5602–5603.

[118] K. Nabeta, K. Kigure, H. Okuyama, T. Takasawa, *J. Chem. Soc., Perkin Trans. 1* **1995**, 1935–1939.

[119] K. Nabeta, M. Fujita, K. Komuro, K. Katayama, T. Takasawa, *J. Chem. Soc., Perkin Trans. 1* **1997**, 2065–2070.

[120] R. Pongdee, H. W. Liu, *Bioinorg. Chem.* **2004**, *32*, 393–437.

[121] F. Yu, D. J. Miller, R. K. Allemann, *Chem. Commun.* **2007**, 4155–4157.

[122] D. J. Miller, F. Yu, W. Knight, R. K. Allemann, *Org. Biomol. Chem.* **2009**, *7*, 962–975.

[123] O. Cascón, S. Touchet, D. J. Miller, V. Gonzalez, J. A. Faraldos, R. K. Allemann, *Chem. Commun.* **2012**, *48*, 9702–9704.

[124] D. J. Miller, F. Yu, R. K. Allemann, *ChemBioChem* **2007**, *8*, 1819–1825.

[125] D. J. Grundy, M. Chen, V. González, S. Leoni, D. J. Miller, D. W. Christianson, R. K. Allemann, *Biochemistry* **2016**, *55*, 2112–2121.

[126] I. Prosser, A. L. Phillips, S. Gittings, M. J. Lewis, A. M. Hooper, J. A. Pickett, M. H. Beale, *Phytochemistry* **2002**, *60*, 691–702.

[127] R. Mozuraitis, M. Stranden, M. I. Ramirez, H. Mustaparta, *Chem. Senses* **2002**, *27*, 505–509.

[128] W. S. Bowers, C. Nishino, M. E. Montgomery, L. R. Nault, M. W. Nielson, *Science* **1977**, *196*, 680–681.

[129] T. J. A. Bruce, M. A. Birkett, J. Blande, A. M. Hooper, J. L. Martin, B. Khambay, I. Prosser, L. E. Smart, L. J. Wadhams, *Pest Manage. Sci.* **2005**, *61*, 1115–1121.

[130] S. Touchet, K. Chamberlain, C. M. Woodcock, D. J. Miller, M. A. Birkett, J. A. Pickett, R. K. Allemann, *Chem. Commun.* **2015**, *51*, 7550–7553.

[131] M. Demiray, X. Tang, T. Wirth, J. A. Faraldos, R. K. Allemann, *Angew. Chem. Int. Ed.* **2017**, *56*, 1–5.

[132] F. Huynh, D. J. Grundy, R. L. Jenkins, D. J. Miller, R. K. Allemann, *ChemBioChem* **2018**, *19*, 1834–1838.

[133] H. J. Bouwmeester, T. E. Wallaart, M. H. A. Janssen, B. Van Loo, B. J. M. Jansen, M. A. Posthumus, C. O. Schmidt, J. W. De Kraker, W. A. König, M. C. R. Franssen, *Phytochemistry* **1999**, *52*, 843–854.

[134] Y. J. Chang, S. H. Song, S. H. Park, S. U. Kim, *Arch. Biochem. Biophys.* **2000**, *383*, 178–184.

[135] P. Mercke, M. Bengtsson, H. J. Bouwmeester, M. A. Posthumus, P. E. Brodelius, *Arch. Biochem. Biophys.* **2000**, *381*, 173–180.

[136] F. Lévesque, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2012**, *51*, 1706–1709.

[137] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, et al., *Nature* **2013**, *496*, 528–532.

[138] J. A. Faraldos, D. J. Grundy, O. Cascon, S. Leoni, M. W. Van Der Kamp, R. K. Allemann, *Chem. Commun.* **2016**, *52*, 14027–14030.

[139] S. B. Horwitz, *Nature* **1994**, *367*, 593–594.

[140] M. Marine, N. Product, J. M. Whitney, J. S. Parnes, K. J. Shea, *J. Org. Chem.* **1997**, *3263*, 8962–

8963.

[141] V. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **2003**, *21*, 796–802.

[142] J. D. Keasling, T. T. Diagana, V. Hale, N. Renninger, *The American Journal of Tropical Medicine and Hygiene* **2018**, *77*, 198–202.

[143] J. R. Anthony, L. C. Anthony, F. Nowroozi, G. Kwon, J. D. Newman, J. D. Keasling, *Metab. Eng.* **2009**, *11*, 13–19.

[144] J. Alonso-Gutierrez, R. Chan, T. S. Batth, P. D. Adams, J. D. Keasling, C. J. Petzold, T. S. Lee, *Metab. Eng.* **2013**, *19*, 33–41.

[145] O. A. Carter, R. J. Peters, R. Croteau, *Phytochemistry* **2003**, *64*, 425–433.

[146] K. K. Reiling, Y. Yoshikuni, V. J. J. Martin, J. Newman, J. Bohlmann, J. D. Keasling, *Biotechnol. Bioeng.* **2004**, *87*, 200–212.

[147] P. P. Peralta-Yahya, M. Ouellet, R. Chan, A. Mukhopadhyay, J. D. Keasling, T. S. Lee, *Nat. Commun.* **2011**, *2*, 483.

[148] J. D. Newman, J. Marshall, M. Chang, F. Nowroozi, E. Paradise, D. Pitera, K. L. Newman, J. D. Keasling, *Bioresour. and Bioeng.* **2006**, *95*, 684–691.

[149] P. K. Ajikumar, W. H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, *Science* **2010**, *330*, 70–74.

[150] H. Alper, K. Miyaoku, G. Stephanopoulos, *Nat. Biotechnol.* **2005**, *23*, 612–616.

[151] H. Tsuruta, C. J. Paddon, D. Eng, J. R. Lenihan, T. Horning, L. C. Anthony, R. Regentin, J. D. Keasling, N. S. Renninger, J. D. Newman, *PloS One* **2009**, *4*, 1–12.

[152] P. J. Westfall, D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, et al., *Proc. Natl. Acad. Sci.* **2012**, *109*, E111–E118.

[153] G. Bian, Z. Deng, T. Liu, *Curr. Opin. Biotechnol.* **2017**, *48*, 234–241.

[154] L. J. Weaver, M. M. L. Sousa, G. Wang, E. Baidoo, C. J. Petzold, J. D. Keasling, *Biotechnol. Bioeng.* **2015**, *112*, 111–119.

[155] F. H. Arnold, *Chem. Eng. Sci.* **1996**, *51*.

[156] O. Kuchner, F. H. Arnold, *Trends Biotechnol.* **1997**, *15*, 523–530.

[157] M. S. Packer, D. R. Liu, *Nat. Rev. Genet.* **2015**, *16*, 379–394.

[158] C. Zeymer, D. Hilvert, *Annu. Rev. Biochem.* **2018**, *87*, 131–157.

[159] J. A. Wells, M. Vasser, D. B. Powers, *Gene* **1985**, *34*, 315–323.

[160] D. S. Wilson, A. D. Keefe, *Current protocols in molecular biology*, **2001**.

[161] S. T. Withers, S. S. Gottlieb, B. Lieu, J. D. Newman, J. D. Keasling, *Appl. Environ. Microbiol.* **2007**, *73*, 6277–6283.

[162] R. Lauchli, K. S. Rabe, K. Z. Kalbarczyk, A. Tata, T. Heel, R. Z. Kitto, F. H. Arnold, *Angew. Chem. Int. Ed.* **2013**, *52*, 5571–5574.

[163] M. Furubayashi, M. Ikezumi, J. Kajiwara, M. Iwasaki, A. Fujii, N. Li, K. Saito, D. Umeno, *PloS One* **2014**, *9*, 1–11.

[164] J. Bohlmann, C. I. Keeling, *The Plant Journal* **2008**, *54*, 656–669.

[165] J. Bohlmann, *Structure* **2011**, *19*, 1730–1731.

[166] D. K. Ro, E. M. Paradise, M. Quellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, et al., *Nature* **2006**, *440*, 940–943.

[167] B. Zhao, X. Lin, L. Lei, D. C. Lamb, S. L. Kelly, M. R. Waterman, D. E. Cane, *J. Biol. Chem* **2008**, *283*, 8183–8189.

[168] C. Nakano, S. Horinouchi, Y. Ohnishi, *J. Biol. Chem* **2011**, *286*, 27980–27987.

[169] C. Nakano, F. Kudo, T. Eguchi, Y. Ohnishi, *ChemBioChem* **2011**, *12*, 2271–2275.

[170] C. Nakano, T. Tezuka, S. Horinouchi, Y. Ohnishi, *J. Antibiot.* **2012**, *65*, 551–558.

[171] P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 13448–13451.

[172] W. K. W. Chou, I. Fanizza, T. Uchiyama, M. Komatsu, H. Ikeda, D. E. Cane, *J. Am. Chem. Soc.* **2010**, *132*, 8850–8851.

[173] C.-M. Wang, R. Hopson, X. Lin, D. E. Cane, *J. Am. Chem. Soc.* **2009**, *131*, 8360–8361.

[174] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *J. Am. Chem. Soc.* **2010**, *132*, 2998–3008.

[175] D. E. Cane, R. M. Watt, *Proc. Natl. Acad. Sci.* **2003**, *100*, 1547–51.

[176] Y. Hu, W. K. W. Chou, R. Hopson, D. E. Cane, *Chem. Biol.* **2011**, *18*, 32–37.

[177] D. A. Whittington, M. L. Wise, M. Urbansky, R. M. Coates, R. B. Croteau, D. W. Christianson, *Proc. Natl. Acad. Sci.* **2002**, *99*, 15375–15380.

[178] M. Chen, W. K. W. Chou, N. Al-Lami, J. A. Faraldos, R. K. Allemann, D. E. Cane, D. W. Christianson, *Biochemistry* **2016**, *55*, 2864–2874.

[179] D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1023–1194.

[180] M. Isegawa, S. Maeda, D. J. Tantillo, K. Morokuma, *Chem. Sci.* **2014**, *5*, 1555–1560.

[181] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli, et al., *Nucleic Acids Res.* **2014**, *42*, 1–7.

[182] S. R. Hare, D. J. Tantillo, *Beilstein J. Org. Chem.* **2016**, *12*, 377–390.

[183] T. E. O'Brien, S. J. Bertolani, D. J. Tantillo, J. B. Siegel, *Chem. Sci.* **2016**, *7*, 4009–4015.

[184] M. Chen, N. Al-Lami, M. Janvier, E. L. D'Antonio, J. A. Faraldos, D. E. Cane, R. K. Allemann, D. W. Christianson, *Biochemistry* **2013**, *52*, 5441–5453.

[185] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248–254.

[186] T. Zor, Z. Selinger, *Anal. Biochem.* **1996**, *236*, 302–308.

[187] F. Karp, Y. Zhao, B. Santhamma, B. Assink, R. M. Coates, R. B. Croteau, *Arch. Biochem. Biophys.* **2007**, *468*, 140–146.

[188] C. H. Benedict, I. Alchanati, P. J. Harvey, J. Liu, R. D. Stipanovic, A. A. Bellt, *Phytochemistry* **1995**, *39*, 327–331.

[189] Y. Yoshikuni, V. J. J. Martin, T. E. Ferrin, J. D. Keasling, *Chem. Biol.* **2006**, *13*, 91–98.

[190] S. C. Kampranis, D. Ioannidis, A. Purvis, W. Mahrez, E. Ninga, N. A. Katerelos, S. Anssour, J. M. Dunwell, J. Degenhardt, A. M. Makris, et al., *The Plant Cell* **2007**, *19*, 1994–2005.

[191] M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Biochemistry* **2002**, *41*, 1732–1741.

[192] H. Kawaide, K. I. Hayashi, R. Kawanabe, Y. Sakigi, A. Matsuo, M. Natsume, H. Nozaki, *FEBS J.* **2011**, *278*, 123–133.

[193] T. A. Pemberton, D. W. Christianson, *J. Antibiot.* **2016**, *69*, 486–493.

[194] D. E. Cane, I. Kang, *Arch. Biochem. Biophys.* **2000**, *376*, 354–364.

[195] J. D. Connolly, W. R. Phillips, S. Huneck, *Phytochemistry* **1982**, *21*, 233–234.

[196] K. Nabeta, K. Katayam, S. Naragawara, *Phytochemistry* **1993**, *32*, 117–122.

[197] N. N. Gerber, *Phytochemistry* **1971**, *10*, 185–189.

[198] C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall, J. S. Dickschat, *ChemBioChem* **2012**, *13*, 202–214.

[199] P. Rabe, C. A. Citron, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 2345–2354.

[200] C. A. Citron, L. Barra, J. Wink, J. S. Dickschat, *Org. Biomol. Chem.* **2015**, *13*, 2673–2683.

[201] Y. Ohnishi, J. Ishikawa, H. Hara, H. Suzuki, M. Ikenoya, H. Ikeda, A. Yamashita, M. Hattori, S. Horinouchi, *J. Bacteriol.* **2008**, *190*, 4050–4060.

[202] P. E. O. Maille, J. Chappell, J. P. Noel, *Anal. Biochem.* **2004**, *335*, 210–217.

[203] M. Vardakou, M. Salmon, J. A. Faraldos, P. E. O'Maille, *MethodsX* **2014**, *1*, 187–196.

[204] G. Holzwarth, P. Doty, *J. Am. Chem. Soc.* **1965**, *87*, 218–228.

[205] N. J. Greenfield, *Nat. Protocols* **2007**, *1*, 2876–2890.

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

[207] Z. S. Derewenda, *Methods* **2004**, *34*, 354–363.

[208] D. S. Waugh, *Trends Biotechnol.* **2005**, *23*, 316–320.

[209] M. Carson, D. H. Johnson, H. Mcdonald, J. Lawrence, *Acta Cryst.* **2007**, *D63*, 295–301.

[210] N. E. Biolabs, *Manual* **2012**, 1–16.

[211] J. A. Aaron, D. W. Christianson, *Pure Appl. Chem.* **2010**, *82*, 1585–1597.

[212] M. Seemann, G. Zhai, J. De Kraker, C. M. Paschall, D. W. Christianson, D. E. Cane, *J. Am. Chem. Soc.* **2002**, *124*, 7681–7689.

[213] D. J. Grundy, R. K. Allemann, Expanding the Terpenome : Complementary Approaches to Novel Terpenoids, **2015**.

[214] D. Plattner, *Arch. Biochem. Biophys.* **1989**, *272*, 137–143.

[215] S. S. Dehal, R. Croteau, *Arch. Biochem. Biophys.* **1988**, *261*, 346–356.

[216] T. M. Hohn, F. Vanmiddlesworth, *Arch. Biochem. Biophys.* **1986**, *251*, 756–761.

[217] S. Picaud, M. Brodelius, P. E. Brodelius, *Phytochemistry* **2005**, *66*, 961–967.

[218] M. B. Quin, S. N. Michel, C. Schmidt-Dannert, *ChemBioChem* **2015**, *16*, 2191–2199.

[219] T. G. Ko, C. Schnee, J. Gershenzon, J. Degenhardt, *The Plant Cell* **2004**, *16*, 1115–1131.

[220] J. Crock, M. Wildung, R. Croteau, *Proc. Natl. Acad. Sci.* **1997**, *94*, 12833–12838.

[221] S. Picaud, L. Olofsson, M. Brodelius, P. E. Brodelius, *Arch. Biochem. Biophys.* **2005**, *436*, 215–226.

[222] D. O'Hagan, *Chem. Soc. Rev.* **2008**, *37*, 308–319.

[223] M. Schlosger, D. Michel, *Tetrahedron* **1996**, *52*, 99–108.

[224] D. O'Hagan, H. S. Rzepa, *Chem. Commun.* **1997**, 645–652.

[225] D. E. Cane, G. Yang, Q. Xue, J. H. Shim, *Biochemistry* **1995**, *34*, 2471–2479.

[226] T. G. Köllner, P. E. O'Maille, N. Gatto, W. Boland, J. Gershenzon, J. Degenhardt, *Arch. Biochem. Biophys.* **2006**, *448*, 83–92.

[227] A. Vattekkatte, N. Gatto, T. G. Köllner, J. Degenhardt, *Org. Biomol. Chem.* **2015**, *13*, 6021–6030.

[228] F. Lopez-Gallego, S. A. Agger, D. Abate-Pella, M. D. Distefano, C. Schmidt-Dannert, *ChemBioChem* **2010**, *11*, 1093–1106.

[229] A. Vattekkatte, S. Garms, W. Boland, *J. Org. Chem.* **2017**, *82*, 2855–2861.

[230] C. Oberhauser, V. Harms, K. Seidel, B. Schröder, K. Ekramzadeh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat, A. Kirschning, *Angew. Chem. Int. Ed.* **2018**, *57*, 11802–11806.

[231] V. M. Dixit, F. M. Laskovics, W. I. Noall, C. D. Poulter, *J. Org. Chem.* **1981**, *46*, 1967–1969.

[232] V. J. Davisson, A. B. Woodside, T. R. Neal, K. E. Stremler, M. Muehlbacher, C. D. Poulter, *J. Org. Chem.* **1986**, *51*, 4768–4779.

[233] R. M. Coates, D. A. Ley, P. L. Cavender, *Journal of Organic Chemistry* **1978**, *43*, 4915-4922.

[234] F. Cramer, W. Bohm, *Angew. Chem.* **1959**, *71*, 775.

[235] R. H. Cornforth, G. Popják, *Methods Enzymol.* **1969**, *15*, 359–390.

[236] R. S. Cahn, C. K. Ingold, V. Prelog, *Experientia* **1956**, *12*, 81–94.

[237] R. K. Keller, R. Thompson, *J. Chromatogr.* **1993**, *645*, 161–167.

[238] L. L. Danilov, T. N. Druzhinina, N. A. Kalinchuk, S. D. Maltsev, V. N. Shibaev, *Chemistry and Chem. Phys. Lipids* **1989**, *51*, 191–203.

[239] L. M. Lira, D. Vasilev, R. A. Pilli, L. A. Wessjohann, *Tetrahedron Lett.* **2013**, *54*, 1690–1692.

[240] A. Thenappan, D. J. Burton, *J. Org. Chem.* **1990**, *55*, 4639–4642.

[241] Y. Komatsu, T. Kitazume, *J. Fluorine Chem.* **2000**, *102*, 61–67.

[242] S. N. Huckin, L. Weiler, *J. Am. Chem. Soc.* **1974**, *96*, 1082–1087.

[243] L. Weiler, *Tetrahedron* **1981**, *37*, 303–317.

[244] Y. Qiu, X. Zhu, Y. Li, Y. He, F. Yang, J. Wang, H. Hua, L. Zheng, L. Wang, X. Liu, *Org. Lett.* **2015**, *17*, 3694–3697.

[245] B. Yang, Z. Lu, *J. Org. Chem.* **2016**, *81*, 7288–7300.

[246] W. S. Wadsworth, W. D. Emmons, *J. Am. Chem. Soc.* **1961**, *83*, 1733–1738.

[247] A. R. Jadhav, R. S. Thombal, P. Nigam, V. H. Jadhav, *Tetrahedron Lett.* **2015**, *56*, 5235–5237.

[248] J.-P. Dulcère, J. Rodriguez, *Synthesis* **1993**, *4*, 399–405.

[249] K. Omura, D. Swern, *Tetrahedron* **1978**, *34*, 1651–1660.

[250] A. J. Mancuso, S. Huang, D. Swern, *J. Org. Chem.* **1978**, *43*, 2480–2482.

[251] J. M. Bobbitt, *J. Org. Chem.* **1998**, *63*, 9367–9374.

[252] P. Kalicki, M. Karchier, K. Michalak, J. Wicha, *J. Org. Chem.* **2010**, *75*, 5388–5391.

[253] K. B. Sharpless, R. C. Michaelson, *J. Am. Chem. Soc.* **1973**, *95*, 6136–6137.

[254] M. Sander, *Chem. Rev.* **1966**, *66*, 297–339.

[255] E. Vedejs, G. A. Krafft, *Tetrahedron* **1982**, *38*, 2857–2881.

[256] E. P. Adam, K. N. Ayad, F. P. Doyl, D. O. Hollan, W. H. Hunter, J. H. C. Nayler, A. Queen, *J. Chem. Soc.* **1960**, *532*, 2660–2665.

[257] R. Ketcham, V. P. Shah, *J. Org. Chem.* **1963**, *28*, 229–230.

[258] Y. Gao, K. B. Sharpless, *J. Org. Chem.* **1988**, *53*, 4114–4116.

[259] H. Bouda, M. E. Borredon, M. Delmas, A. Gaset, *Synth. commun.* **1989**, *19*, 491–500.

[260] F. Kazemi, A. R. Kiasat, S. Ebrahimi, *Synth. Commun.* **2003**, *33*, 595–600.

[261] H. Sharghi, M. A. Nasseri, K. Niknam, *J. Org. Chem.* **2001**, *66*, 7287–7293.

[262] E. E. Tamelen, *J. Am. Chem. Soc.* **1951**, *73*, 3444–3448.

[263] B. Tamami, A. R. Kiasat, *Synth. Commun.* **1996**, *26*, 3953–3958.

[264] T. H. Chan, J. R. Finkenbin, *J. Am. Chem. Soc.* **1972**, *94*, 2880–2882.

[265] M. O. Brimeyer, A. Mehrota, S. Quici, A. Nigam, S. L. Regen, *J. Org. Chem.* **1980**, *45*, 4254–4255.

[266] I. T. S. Yadav, B. V. S. Reddy, G. Baishya, *Synlett* **2003**, *3*, 396–398.

[267] B. Tamami, M. Kolahdoozan, *Tetrahedron Lett.* **2004**, *45*, 1535–1537.

[268] B. Kaboudin, H. Norouzi, *Tetrahedron Lett.* **2004**, *45*, 1283–1285.

[269] J. S. Yadav, B. V. S. Reddy, C. S. Reddy, K. Rajasekhar, *J. Org. Chem.* **2003**, *68*, 2525–2527.

[270] N. S. Krishnaveni, K. Surendra, M. S. Reddy, Y. V. D. Nageswar, K. R. Rao, *Adv. Synth. Catal.* **2004**, *346*, 395–397.

[271] K. Surendra, N. S. Krishnaveni, K. R. Rao, *Tetrahedron Lett.* **2004**, *45*, 6523–6526.

[272] V. Gonzalez, S. Touchet, D. J. Grundy, J. A. Faraldos, R. K. Allemann, *J. Am. Chem. Soc.* **2014**, *136*, 14505–14512.

[273] D. E. Cane, H. Ikeda, *Acc. Chem. Res.* **2012**, *45*, 463–472.

[274] J. Jiang, X. He, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 8128–8129.

[275] M. Loizzi, V. González, D. J. Miller, R. K. Allemann, *ChemBioChem* **2018**, *19*, 100–105.

[276] J. W. De Kraker, M. C. R. Franssen, A. De Groot, W. A. König, H. J. Bouwmeester, *Plant Physiol.* **1998**, *117*, 1381–1392.

[277] P. Steliopoulos, M. Wu, K. Adam, A. Mosandl, *Phytochemistry* **2002**, *60*, 13–20.

[278] C. P. Cornwell, N. Reddy, D. N. Leach, S. Grant Wyllie, *Flavour Fragr. J.* **2001**, *16*, 263–273.

[279] R. D. Hartle, C. H. Fawcet, *Phytochemistry* **1969**, *8*, 637–643.

[280] J. A. Faraldos, Y. Zhao, P. E. O'Maille, J. P. Noel, R. M. Coates, *ChemBioChem* **2007**, *8*, 1826–1833.

[281] J. A. Faraldos, D. J. Miller, V. González, Z. Yoosuf-Aly, O. Cascón, A. Li, R. K. Allemann, *J. Am. Chem. Soc.* **2012**, *134*, 5900–5908.

[282] M. Demiray, Exploring and Mapping the Functional Chemical Space of Amorphadiene Synthase with Non-Canonical Farnesyl Diphosphate Analogues, **2016**.

[283] C. Sallaud, D. Rontein, S. Onillon, F. Jabes, P. Duffe, C. Giacalone, S. Thoraval, C. Escoffier, G. Herbette, N. Leonhardt, et al., *The Plant Cell* **2009**, *21*, 301–317.

[284] P. M. Bleeker, P. J. Diergaarde, K. Ament, S. Schütz, B. Johne, J. Dijkink, H. Hiemstra, R. De Gelder, M. T. J. De Both, M. W. Sabelis, *Phytochemistry* **2011**, *72*, 68–73.

[285] P. M. Bleeker, R. Mirabella, P. J. Diergaarde, A. VanDoorn, A. Tissier, M. R. Kant, M. Prins, M. De Vos, M. A. Haring, R. C. Schuurink, *Proc. Natl. Acad. Sci.* **2012**, *109*, 20124–20129.

[286] C. G. Jones, J. Moniodis, K. G. Zulak, A. Scaffidi, J. A. Plummer, E. L. Ghisalberti, E. L. Barbour, J. Bohlmann, *J. Biol. Chem.* **2011**, *286*, 17445–17454.

[287] N. J. White, *Science* **2008**, *320*, 330–334.

[288] N. Gerber, N. Brunswick, *Appl. Microbiol.* **1965**, *13*, 935–938.

[289] G. G. Harris, P. M. Lombardi, T. A. Pemberton, T. Matsui, T. M. Weiss, K. E. Cole, D. E. Cane D. W. Christianson, *Biochemistry* **2016**, *54*, 7142–7155.

[290] N. N. Gerber, *Crit. Rev. Microbiol.* **1979**, *7*, 191–214.

[291] A. J. Hassett, E. R. Rohwer, *J. Chromatogr.* **1999**, *849*, 521–528.

[292] P. Darriet, M. Pons, S. Lamy, D. Dubourdieu, *J. Agric. Food Chem.* **2000**, *48*, 4835–4838.

[293] T. P. Heil, R. C. Lindsay, *J. Environ. Sci. Health* **1988**, *B23 (5)*, 489–512.

[294] C. Agric, E. Stn, B. No, G. H. Green, *J. Agric. Food Chem.* **1976**, *24*, 1246–1247.

[295] M. Tashiro, H. Kiyota, S. Kawai-Noma, K. Saito, M. Ikeuchi, Y. Iijima, D. Umeno, *ACS Synth. Biol.* **2016**, *5*, 1011–1020.

[296] D. E. Cane, Manish Tandon, *Tetrahedron Lett.* **1994**, *35*, 5355–5358.

[297] S. Das, M. Dixit, D. T. Major, *Bioorg. Med. Chem.* **2016**, *24*, 4867–4870.

[298] C. Engler, R. Gruetzner, R. Kandzia, S. Marillonnet, *PloS One* **2009**, *4*, e5553.

[299] R. C. Cadwell, G. F. Joyce, *Genome Res.* **1992**, *2*, 28–33.

[300] O. Ernst, T. Zor, *J. Vis. Exp.* **2010**, *38*, 1–6.

[301] W. Cocker, N. W. A. Geraghty, T. B. H. Mcmurry, P. V. R. Shannon, *J. Chem. Soc., Perkin Trans. 1* **1984**, 2245–2254.

[302] T. Takigawa, K. Mori, M. Matsui, *Agr. Biol. Chem.* **1975**, *39*, 249–258.

CHAPTER 9

APPENDIX

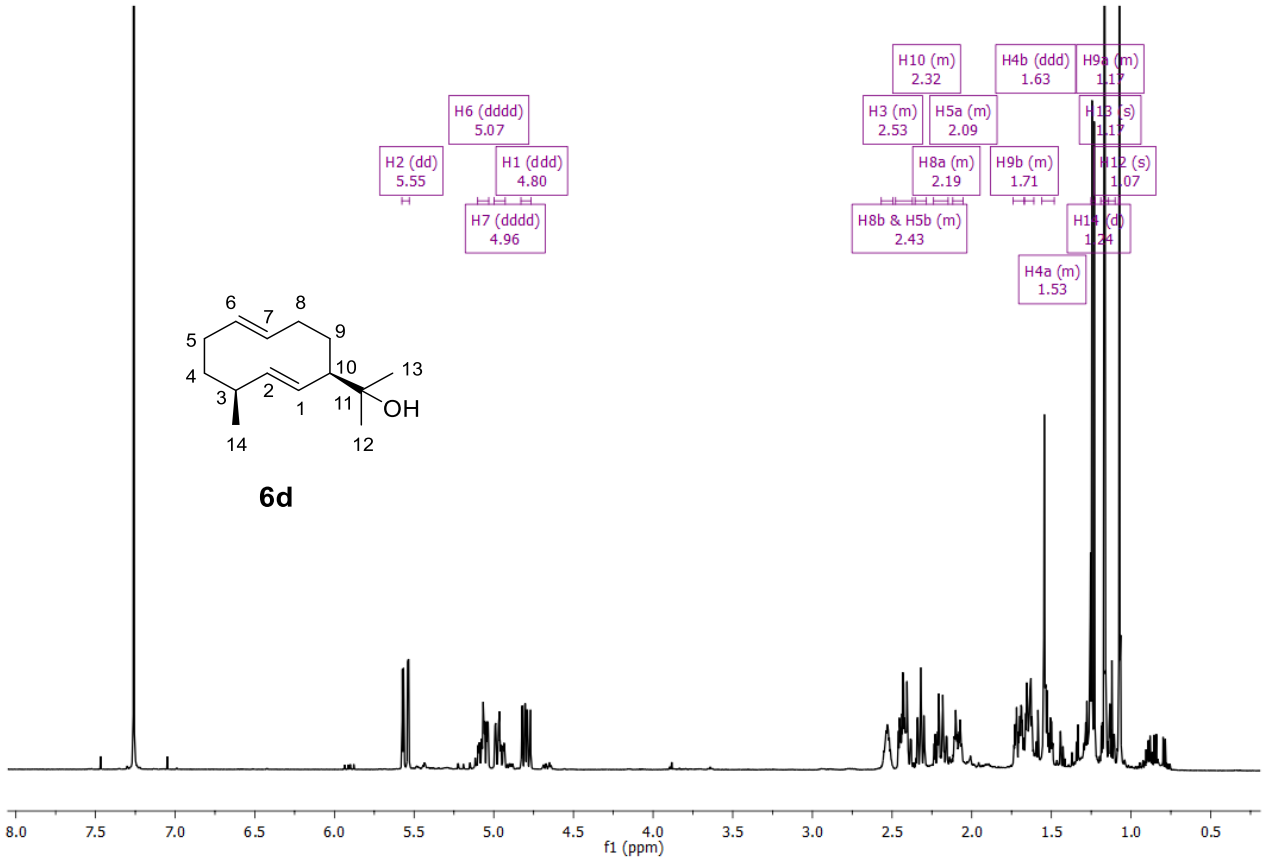


Figure 161. ^{1}H-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

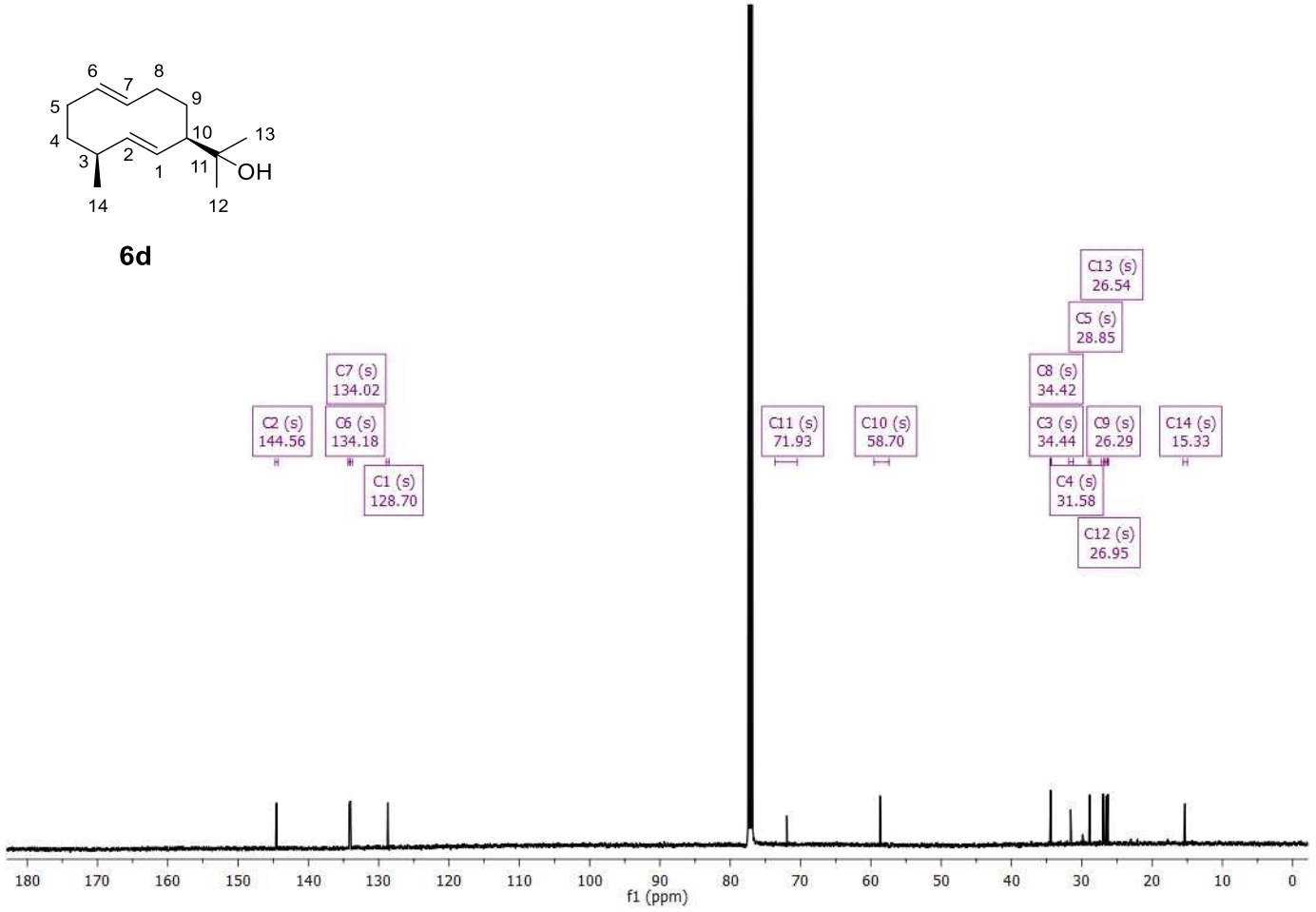


Figure 162. ^{13}C-NMR spectrum (126 MHz, $CDCl_3$, 298 K) of compound **6d**.

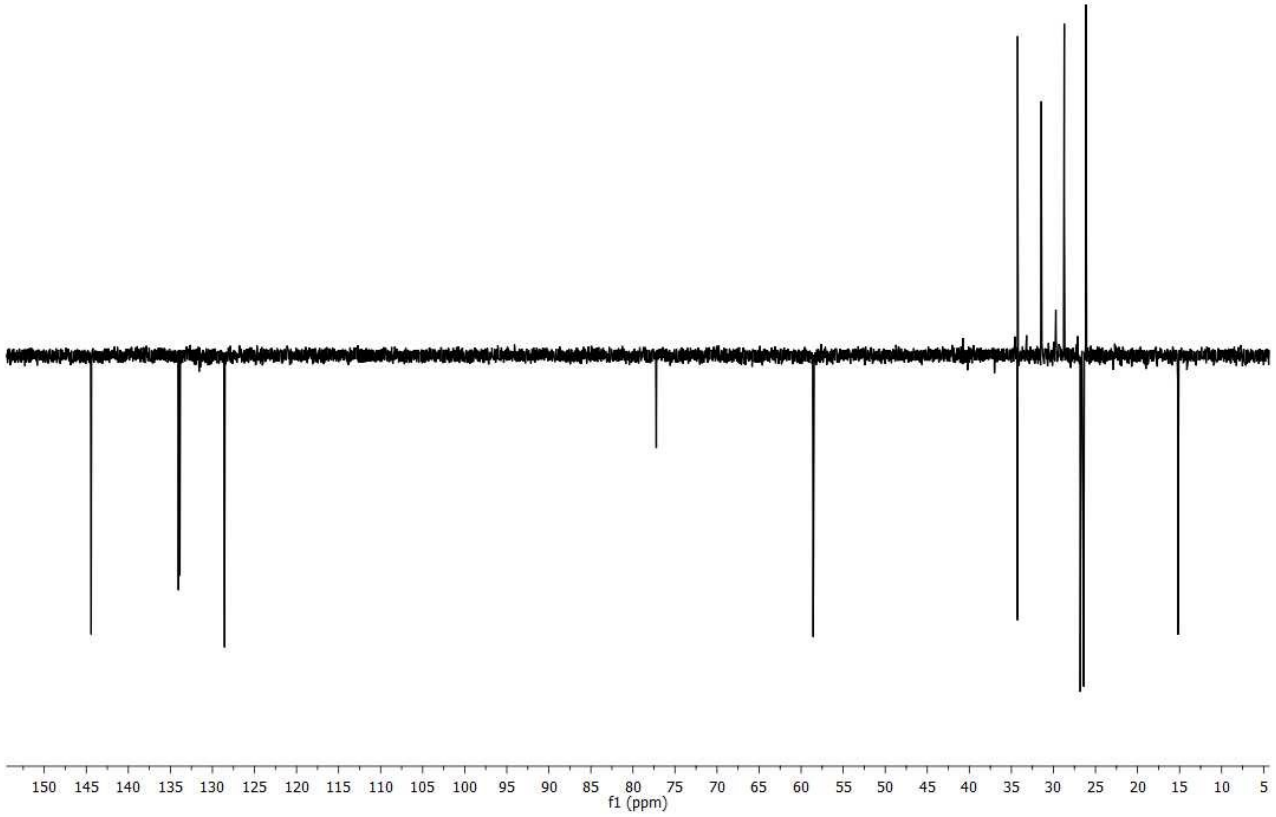


Figure 163. DEPT135-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

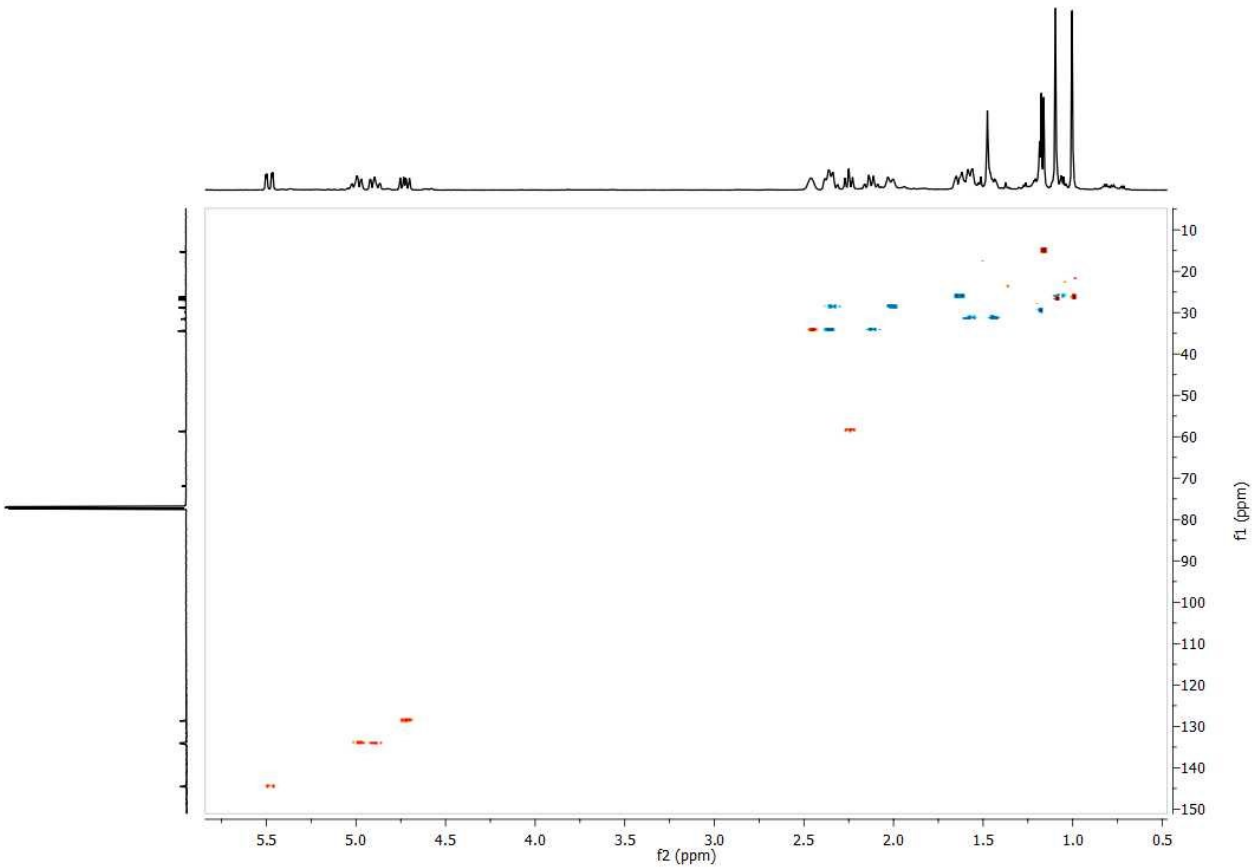


Figure 164. ^{1}H-^{13}C HSQC-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

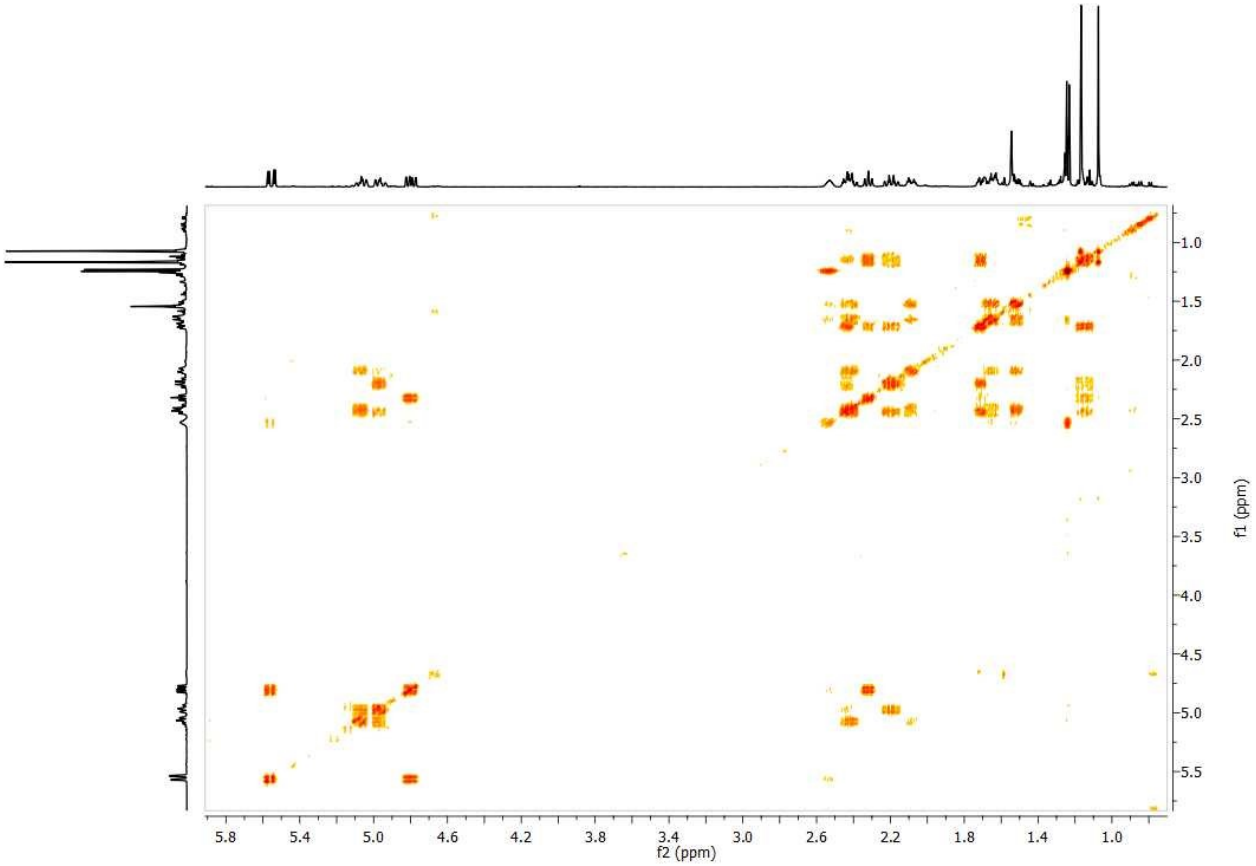


Figure 165. ^{1}H-^{1}H COSY-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

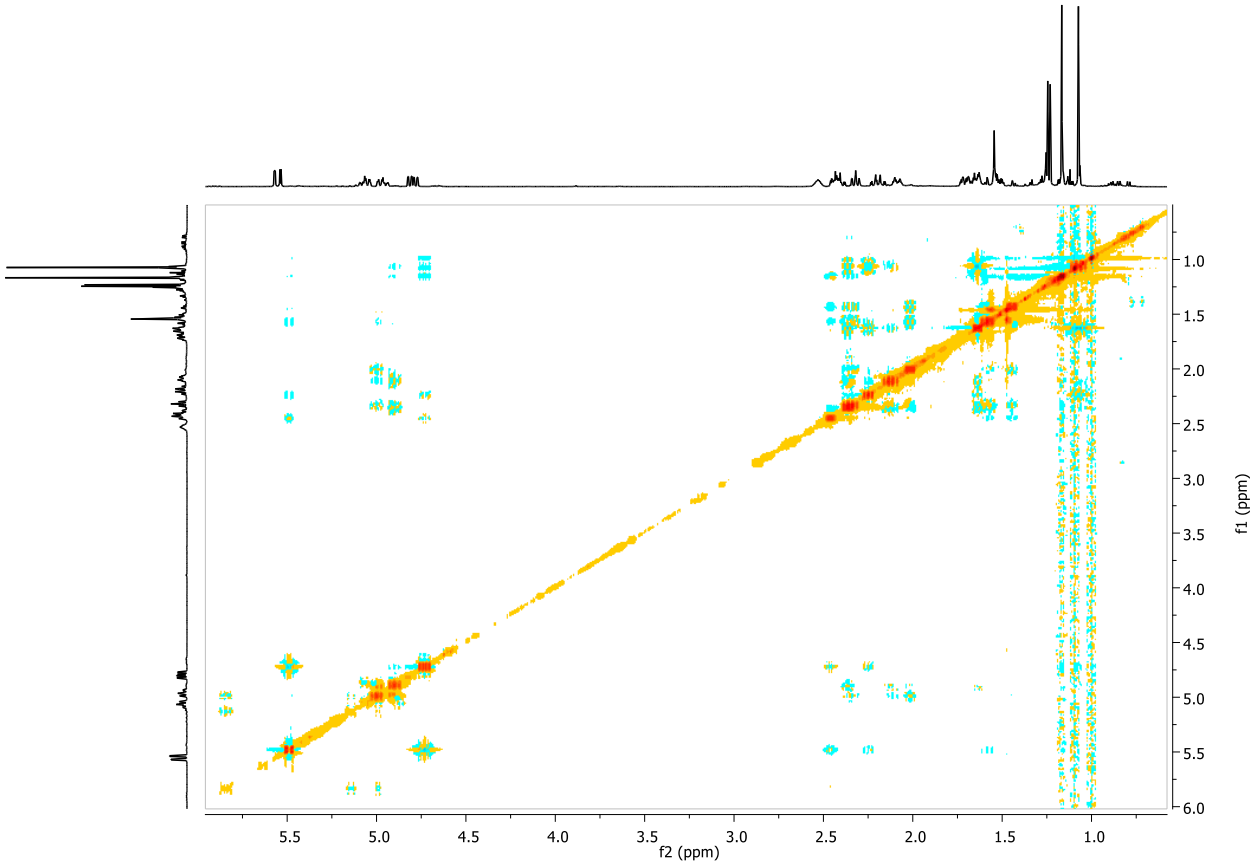


Figure 166. ^{1}H-^{1}H NOESY-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

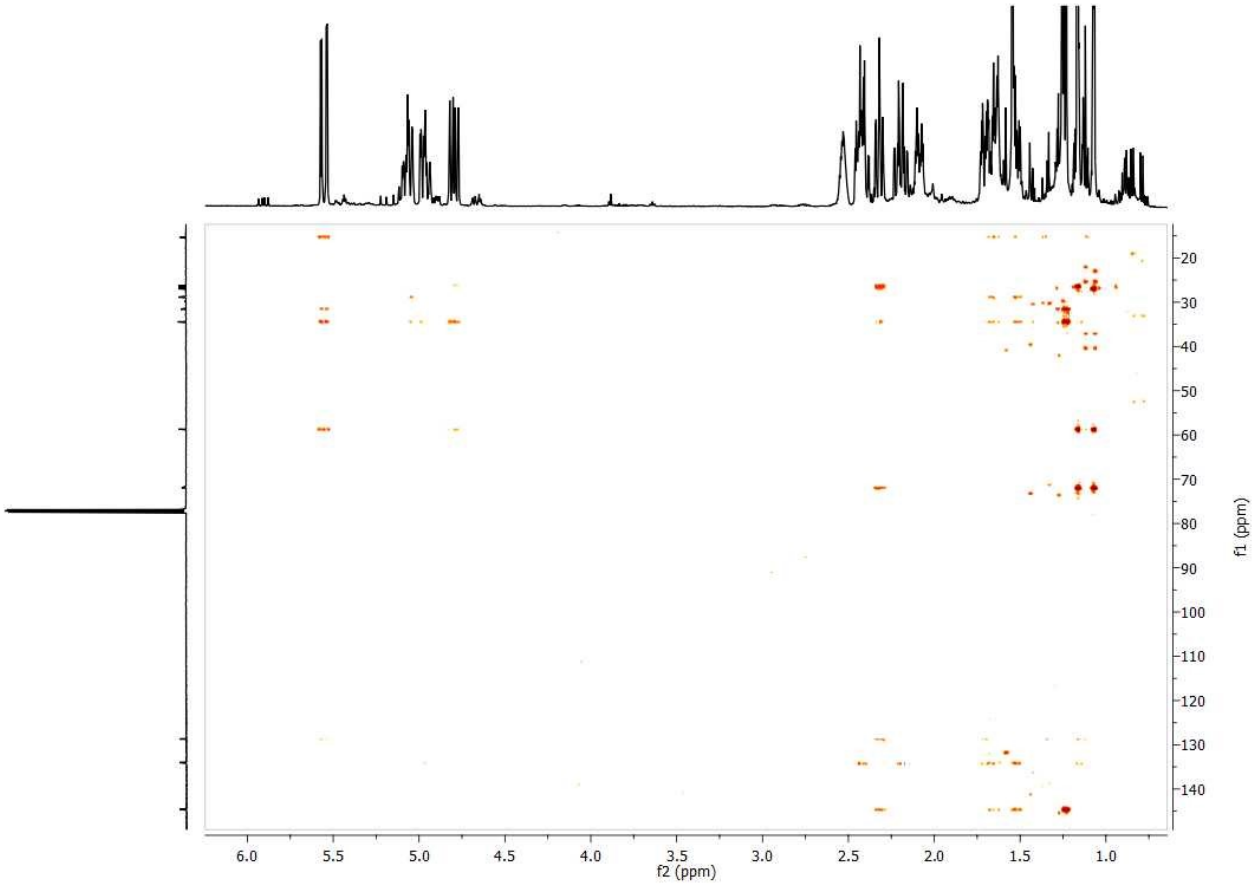


Figure 167. ^{1}H-^{13}C HMBC-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6d**.

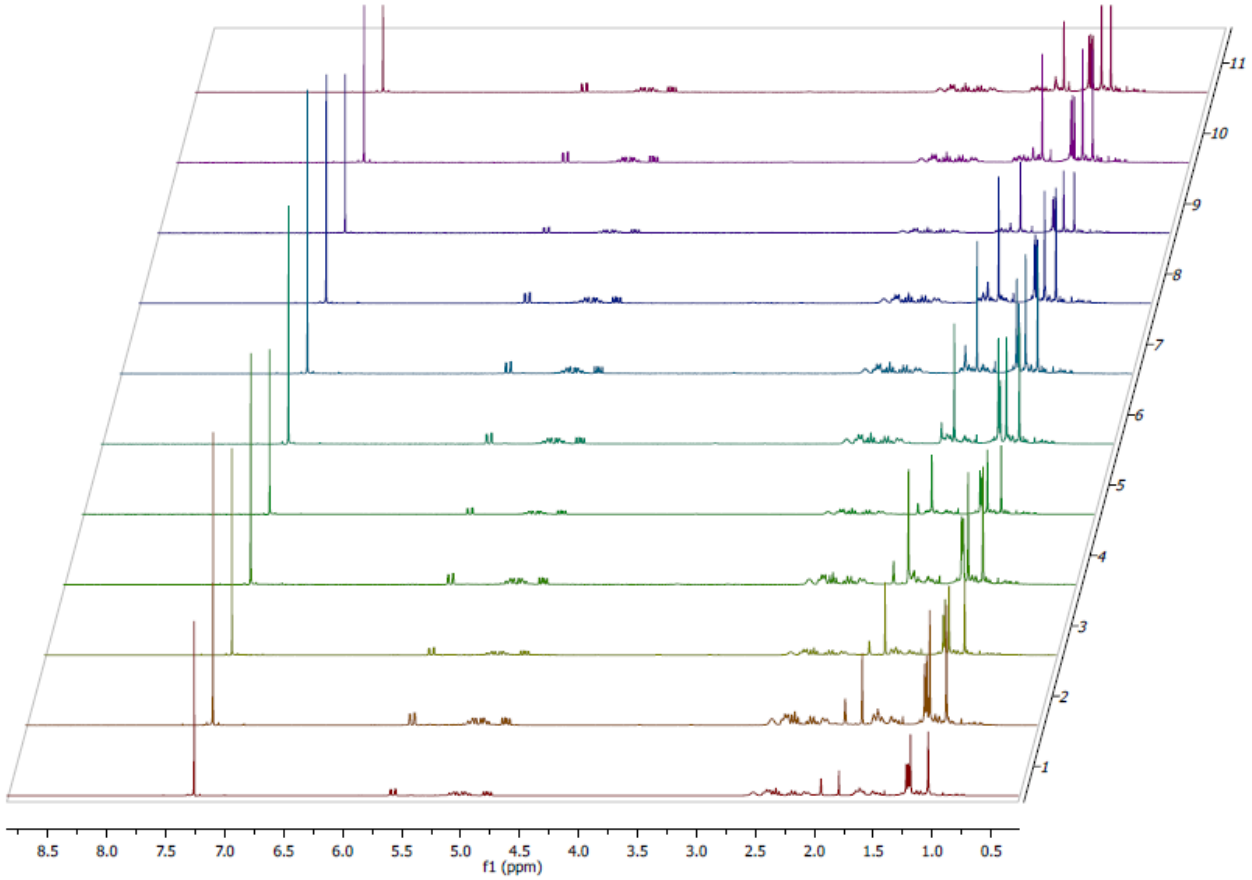


Figure 168. VT-NMR experiments, compound **6d**.

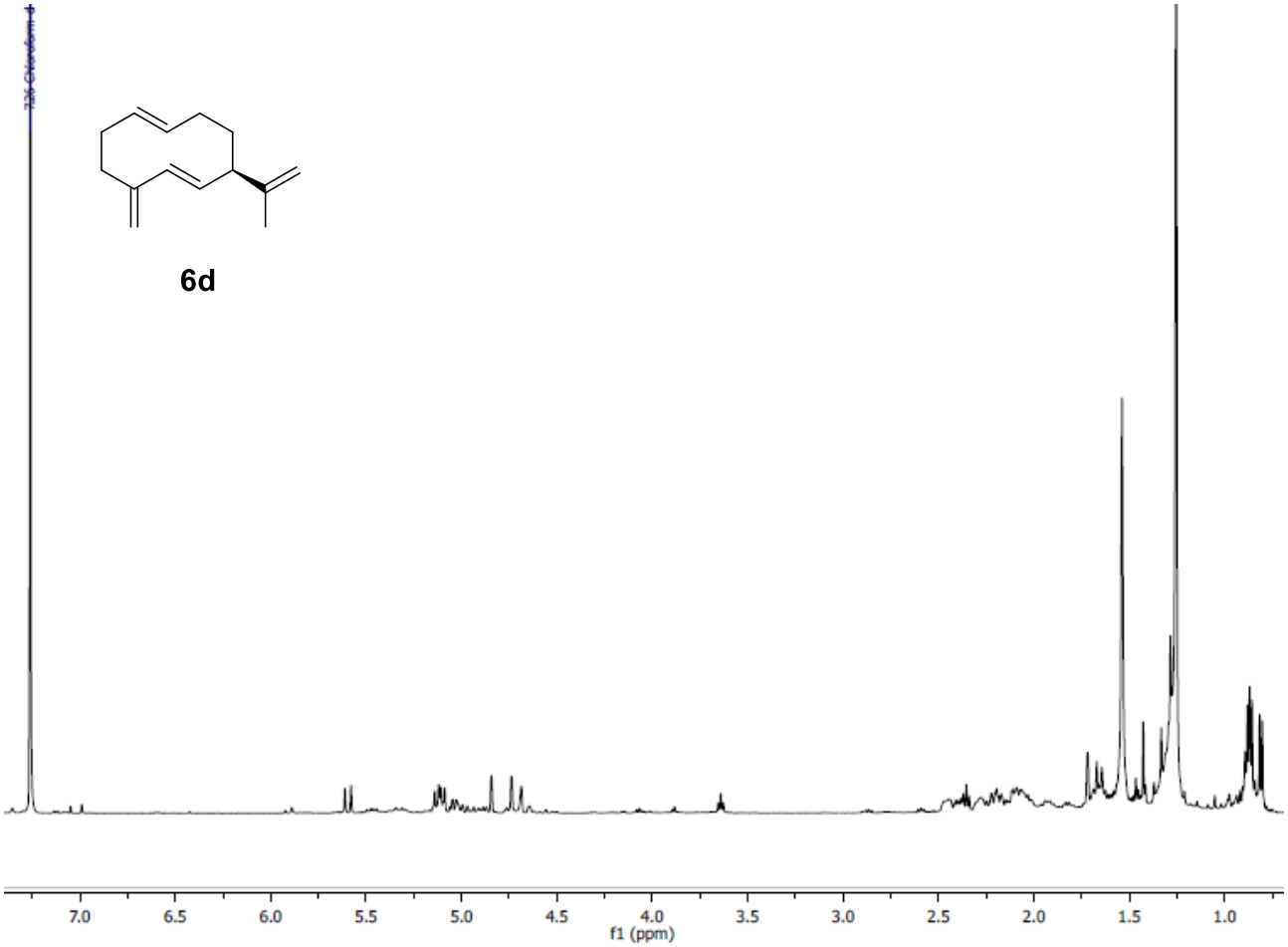


Figure 169. ^{1}H-NMR spectrum (500 MHz, $CDCl_3$, 298 K) of compound **6b**.