

Efficient chemoenzymatic synthesis of terpenes

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

Florence Huynh

Supervisor: Rudolf K. Allemann



Cardiff University
School of Chemistry
September 2020

blank

ABSTRACT

Terpenes and terpenoids are natural products produced by terpene synthases from linear polyprenyl diphosphates. These enzymes catalyse a complex cascade of reactions involving carbocations, modification in bonding and hybridisation in a regio- and stereoselective manner. A great number of them are high-value compounds as they display relevant biological activities and their supply is limited. The terpenoid artemisinin is, for example, the most used treatment against malaria and currently extracted from human cropped *Artemisia annua*.

Recent studies have promoted the *in vitro* use of natural or unnatural substrates with terpene synthases to generate existing or novel terpenoids. Chemically synthesised farnesyl diphosphates containing heteroatoms, hydroxy, fluoride or alkyl groups have been used with sesquiterpene synthases (STS) to generate new sesquiterpene analogues. This chemoenzymatic approach has allowed the discovery of compounds showing potential application in the agriculture, flavour, or fragrance industry. This thesis focuses on the design and exploration of methodologies to address the challenges of the chemoenzymatic production of natural and novel terpenoids. The first part of the thesis focuses on improving the diphosphorylation procedure of farnesol and its analogues. A one-pot enzymatic diphosphorylation involving undecaprenol kinase and a variant of isopentenyl phosphate kinase was proven to convert farnesol successfully to farnesyl diphosphate (FDP) in 95% yield. In comparison to the traditional method, a chemical diphosphorylation, which gives FDP in 40% yield on average. Several farnesol analogues were also diphosphorylated using this method showing its potential versatility, namely: 6,15-dimethylfarnesol, 14,15-dimethylfarnesol, 10,11-epoxyfarnesol and 8-methoxyfarnesol. The sesquiterpene product arising from the incubation of 10,11-epoxyfarnesyl diphosphate with germacradien-4-ol synthase was isolated and fully characterised for the first time. Ultimately, this work illustrated how the high-performance counter current chromatography (HPCCC), a technique traditionally used for chromatography, could be used as a liquid-liquid reactor for biocatalysis to improve sesquiterpene yields. Several sesquiterpenes were synthesised in a high yield using the efficient mixing created by the centrifugal forces in the HPCCC apparatus. Carrying out the reaction in this counter current chromatography apparatus gave >70% yields while being almost 10 times faster than the segmented flow system and 70 times faster than batch. This methodology also found applications in stereoselective esterification with lipases, cutting the reaction time from one hour in batch to six minutes using the HPCCC apparatus.

LIST OF PUBLICATIONS

The following publications have evolved from the work of the candidate:

F. Huynh, D. J. Grundy, R. L. Jenkins, D. J. Miller, R. K. Allemann, *ChemBioChem* **2018**, *19*, 1834–1838.

F. Huynh, D. J. Miller, R. K. Allemann, in *Methods Enzymol.*, Elsevier **2018**, *608*, 83–95.

M. Elsherbini, F. Huynh, A. Dunbabin, R. Allemann, T. Wirth, *Chem. A Eur. J.* **2020**, *26*, 11423–11425.

F. Huynh, M. Tailby, A. Finniear, K. Stephens, R. K. Allemann, T. Wirth, *Angew. Chemie Int. Ed.* **2020**, *59*, 16490–16495.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, Prof. Rudolf K. Allemann, for giving me the opportunity to do a PhD in his group, for providing the excellent research facilities that made this work possible and for his continued supervision.

I thank the School of Chemistry, Cardiff University for funding my PhD, BioExtractions (Wales) for allowing me to perform the first-ever sesquiterpene synthase biocatalysis in a HPLC and Colin Bright for all the technical help with the HPLC.

Special thanks go to Professor Thomas Wirth, for all the helpful scientific discussions we had on Fridays, for his support throughout my PhD and his enthusiasm for chemistry and biochemistry.

I would also like to thank several current and past members of staff at Cardiff University. Dr Veronica Gonzalez for teaching me all the in and out of protein production in less than a month, Dr Prabhakar Srivastava, Dr Aduragbemi Adesina, Dr Alan Scott and Dr Luke Johnson for transferring their expertise in molecular biology to me and for all the helpful scientific discussions. Dr Robert Mart for his help and advice in the lab as well as his tireless efforts to make sure all the lab equipment was always in working order. Further thanks go to Dr David Miller for proofreading this thesis as well as teaching me how to improve my scientific writing in the last four years. I would also like to thank Dr Daniel Grundy for pushing me towards the PhD pathway and for proofreading my thesis. A big thank to Dr Alice Dunbabin, for everything she has done for me: bringing me back to sanity in 2018, proofreading my thesis and most importantly making sure the office and lab were an enjoyable place to work in. Lastly, I would like to give a very special thanks to Tom Williams, Dr Robert Jenkins and the rest of the technical staff for their support and advice when I had troubles with the NMR or GC machines. Tom, you have always taken the time out to help me or answer my questions about GC and for that, I am incredibly grateful.

I would also like to thank all the members of the Allemann group, in particular, Ed and Raquel for their invaluable friendship, I am very lucky to have met both of you. Many thanks to the single-cell organism (Gwawr, Jenny, Alice, Raquel) for always having my back and being themselves. And finally, I would like to thank all the friends that I made at Cardiff: Martin, Chris, Donya, Mohammed, Alex, Max.

Lastly, I would like to thank my friends and family for their love and support. Specifically, my partner, Ronan, for his incredible support, his help with some of my figures, for trying to teach me a million new different shortcuts on the computer and making me laugh every day. I would also like to thank Linda for being constantly positive and supportive of me even though we have not lived in the same city for nearly ten years. A big thank you to my friends back in France, Laetitia, Gorka, Jeehuyn, Julie, Margot, Sebastien for giving me motivational speeches when I needed and for our regular catch up to remind me that not everything revolved around a PhD.

TABLE OF CONTENTS

ABSTRACT	III
LIST OF PUBLICATIONS.....	V
ACKNOWLEDGMENTS	VII
TABLE OF CONTENTS.....	IX
LIST OF FIGURES.....	XIII
LIST OF TABLES.....	XXV
LIST OF SCHEMES	XXVII
LIST OF ABBREVIATIONS	XXIX
CHAPTER 1.....	1
1.1 TERPENES.....	3
1.1.1 <i>Terpene biosynthesis</i>	4
1.1.2 <i>Mevalonate pathway</i>	7
1.1.3 <i>Deoxyxylulose 5-phosphate pathway</i>	7
1.1.4 <i>The lost alternate mevalonate pathway in archaea</i>	9
1.1.5 <i>Isoprenyl diphosphate elongation</i>	10
1.1.6 <i>Terpene synthases</i>	10
1.2 SESQUITERPENES.....	11
1.2.1 <i>Structure</i>	12
1.2.2 <i>Origin of sesquiterpene diversity and its mechanism</i>	13
1.2.3 <i>Kinetics of sesquiterpene synthases</i>	16
1.2.4 <i>Current production of sesquiterpenes and recent developments</i>	17
1.3 SESQUITERPENE SYNTHASES USED IN THIS STUDY	21
1.3.1 <i>Amorpha-4,11-diene synthase</i>	21
1.3.2 <i>Aristolochene synthase</i>	22
1.3.3 <i>Germacradiene-4-ol synthase</i>	25
1.3.4 <i>Germacrene D synthase</i>	25
1.4 APPROACHES TO NOVEL TERPENOIDS	26
1.4.1 <i>Unnatural substrates</i>	27
1.4.2 <i>Engineering sesquiterpene synthases</i>	33
1.4.3 <i>Limitations</i>	34

1.5	BIOCATALYTIC OPTIMISATION OF SESQUITERPENE SYNTHASES	34
1.5.1	<i>Batch process</i>	35
1.5.2	<i>Flow systems</i>	35
	<i>Segmented flow systems</i>	36
1.5.3	<i>Site directed mutagenesis for improved catalytic efficiency</i>	38
1.6	AIM OF THE PROJECT	39
1.6.1	<i>Context</i>	39
1.6.2	<i>Aims</i>	40
CHAPTER 2	41
2.1.	PREFACE	43
2.2.	UNDECAPRENOL KINASE	47
2.2.1	<i>Expression</i>	47
2.2.2	<i>Incubation of farnesol with UK</i>	53
2.3.	ISOPENTENYL PHOSPHATE KINASE	54
2.4.	ENZYMATIC SYNTHESIS OF FDP FROM FARNESOL.....	61
2.5.	SUMMARY	68
CHAPTER 3	71
3.1.	PREFACE	73
3.2.	DIMETHYLATED ANALOGUES FOR NOVEL SEMIOCHEMICALS	73
3.2.1.	<i>6,15-Dimethyl FDP synthesis</i>	75
3.2.2.	<i>12,15-Dimethyl FDP synthesis</i>	77
3.2.3.	<i>Incubation with GDS and variants</i>	78
3.3.	TEMPLATING GDOLS FOR THE SYNTHESIS OF A CYCLIC ETHER	81
3.4.	VARIOUS FDP ANALOGUES SYNTHESISED IN THIS WORK	88
3.4.1.	<i>14,15-dimethyl FDP synthesis</i>	88
3.4.1.	<i>12-Hydroxy FDP synthesis</i>	89
3.4.2.	<i>8-Methoxy FDP and 12-methoxy FDP synthesis</i>	90
3.5.	SUMMARY	91
CHAPTER 4	93
4.1.	PREFACE.....	95
4.2.	EXPRESSION AND PURIFICATION OF SESQUITERPENES SYNTHASES.....	95
4.3.	(S)-14,15-DIMETHYLGERMACRENE D SYNTHESIS USING A SEGMENTED FLOW SYSTEM	99
4.4.	OSCILLATORY SEGMENTED FLOW SYSTEM.....	103

4.5.	HPCCC	105
4.6.	TRANSESTERIFICATION BY LIPASES	116
4.7.	SUMMARY	125
CHAPTER 5.....		127
5.1.	DEVELOPMENT OF AN EFFICIENT DIPHOSPHORYLATION PROCESS.....	128
5.2.	EXPANDING THE TERPENOME.....	129
5.3.	NOVEL METHODOLOGIES FOR SESQUITERPENE CATALYSIS	129
5.4.	FUTURE WORK.....	130
CHAPTER 6.....		133
6.1.	BIOLOGICAL METHODS	135
6.1.1.	<i>Materials</i>	135
6.1.2.	<i>Media</i>	135
6.1.3.	<i>Antibiotics stocks</i>	135
6.1.4.	<i>Agar plate preparation</i>	136
6.1.5.	<i>Bacterial strains</i>	136
6.1.6.	<i>Competent cells</i>	136
6.1.7.	<i>Supercompetent cells</i>	137
6.1.8.	<i>Gene sequence</i>	138
6.1.9.	<i>Cloning</i>	143
6.1.10.	<i>Agarose gel</i>	146
6.1.11.	<i>Gel extraction</i>	147
6.1.12.	<i>Gibson assembly</i>	147
6.1.13.	<i>Golden gate</i>	149
6.1.14.	<i>Transformation of competent cells</i>	150
6.1.15.	<i>Plasmid preparation (MiniPrep of DNA)</i>	151
6.1.16.	<i>Protein expression</i>	151
6.1.17.	<i>Protein purification</i>	152
6.1.18.	<i>Ni-NTA affinity chromatography</i>	153
6.1.19.	<i>Purification of AS using anion exchange</i>	154
6.1.20.	<i>Dialysis</i>	154
6.1.21.	<i>Desalting</i>	154
6.1.22.	<i>SDS-PAGE electrophoresis</i>	155
6.1.23.	<i>Western blotting</i>	157

6.1.24.	<i>Protein concentration – Bradford assay</i>	158
6.1.25.	<i>GC-MS analysis of products</i>	158
6.1.26.	<i>GC-FID analysis</i>	159
6.1.27.	<i>UK and IPK phosphorylation incubation</i>	159
6.1.28.	<i>Preparative scale enzymatic incubation</i>	160
6.2.	SYNTHESIS	161
6.2.1.	<i>General experimental</i>	161
6.2.2.	<i>Experimental</i>	162
6.3.	FLOW METHODS	187
6.3.1.	<i>General experimental</i>	188
6.3.2.	<i>GC-MS analysis</i>	189
6.3.3.	<i>GC FID analysis</i>	189
6.3.4.	<i>Sesquiterpene synthases, HPLCCC, flow, batch protocols</i>	189
	<i>Incubation buffers and enzyme concentrations</i>	189
	<i>HPLCCC protocols</i>	190
	<i>Segmented flow protocol</i>	194
	<i>Batch protocol</i>	194
6.3.5.	<i>Lipase activity</i>	195
6.3.6.	<i>Lipase HPLCCC and batch protocols</i>	196
6.3.7.	<i>Data analysis</i>	198
REFERENCES		199
APPENDIX		I
I.	GAS CHROMATOGRAMS AND CALIBRATIONS CURVES	II
II.	NMR SPECTROSCOPIC ANALYSIS OF CYCLIC ETHER	XIX
III.	NMR SPECTROSCOPIC ANALYSIS OF ARTEMISINIC-11S,12-EPOXIDE	XXVIII

LIST OF FIGURES

Figure 1-1 Examples of commercially available terpenes with the terpene skeleton highlighted in red (Aldrich prices)	4
Figure 1-2 Structures of a) isoprene b) regular c) irregular condensation of isoprene units. ^[16]	5
Figure 1-3 MVA and DXP pathways.	8
Figure 1-4 Alternative mevalonate pathway found in archaea.	9
Figure 1-5 Mechanistic representation of trans prenyl diphosphate biosynthesis.	10
Figure 1-6 Cartoon image of crystal structures of two sesquiterpene synthases 1a. Aristolochene synthase (ATAS) from <i>Aspergillus terreus</i> complexed with three Mg ²⁺ and diphosphate (PPi) 1b. Epi-isozizaene synthase (EIZS) from <i>Streptomyces coelicolor</i> complexed with [Mg ²⁺] ₃ -PPi. (PDB: 2OA6 and 3KB9). Magnesium ions are shown in blue balls, PPi as a stick structure where oxygen atoms are red and phosphorus is orange. 1c. Close up on ATAS active site highlighting the NSE and DDXXE motif 1d. Close up on EIZS highlighting the NSE and DDXXE motif. ^[33,34]	12
Figure 1-7 Scheme representing first steps of reaction mechanism of all sesquiterpene synthases. Loss of diphosphate leads to the formation of a highly reactive carbocation. The resulting farnesyl cation can undergo a large range of cyclisation bringing diversity to sesquiterpenes backbones.	14
Figure 1-8 Model of the substrate, metal-binding and conformational changes in ATAS (adapted from Shishova et al. 2008). ^[39]	15
Figure 1-9 Scheme for the biosynthesis of sesquiterpene through MVA pathway in a sugar fed engineered host. ^[14]	18
Figure 1-10 Resorcinarene chemical structure and structure of the hexameric resorcinarene capsule (modified from CIF CCDC 1590340).	19
Figure 1-11 Products formed from the resorcinarene capsule. *Isolated yield. ^[53-56,58]	20
Figure 1-12 Incorporation of deuterium into (1R,4R,4aS,8aR)-amorpha-4,11-diene from deuterated FDPs identified by NMR spectroscopy.	22
Figure 1-13 Cyclisation outcome of fluorinated FDPs by PRAS.	24
Figure 1-14 Examples of analogues of FDP turned over by sesquiterpene synthases and their products published by Allemann R. K. and co-workers.	28

Figure 1-15 a. Product generated by TmS from FDP b. C6 methylated analogues of FDP turned over by TmS published by Dickschat J. S. and co-workers.	29
Figure 1-16 a. Product generated by Bot2 from FDP. b-d. examples of methyl-shifted analogues of FDP turned over by Bot2 published by Kirschning A. and co-workers.....	30
Figure 1-17 Examples of heteroatom containing analogues of FDP from Kirschning and co-workers turned over by sesquiterpene synthases and their product.	31
Figure 1-18 Carbonyl containing sesquiterpenes from oxygenated analogues of FDP with ADS.	32
Figure 1-19 Synthesis of aldehyde (116) from vinyl ether containing analogue of FDP coupled with assay for methanol detection for high-throughput screening of suitable mutant.	32
Figure 1-20 a. Batch procedure for terpene synthases b. Segmented flow set up for terpene synthesis. Black double ended arrows represent mass transfer between the two phases. Grey circular arrow represents the convective flow inside the segment.	37
Figure 1-21 a. (1R,7R,8aS)-Aristolochene synthesis in flow b. (1R,4R,4aS,8aR)-Amorpha-4,11-diene synthesis in flow.	37
Figure 2-1 Various procedure for the chemical diphosphorylation of polyprenol. Prenol is used as example. a. Cramer method n = 0, 1, 2 b. Davisson method c. Wessjohann method.	45
Figure 2-2 a. UK-catalysed conversion of undecaprenol to undecaprenol phosphate. b. IPK-catalysed conversion of IP to IDP.....	46
Figure 2-3 Agarose gel showing the results from digestion of Plasmid 1 pet28a-UK. Lane 1) uncut pet28a-UK, Lane 2) digestion with XbaI, Lane 3) digestion with NcoI and XhoI.	48
Figure 2-4 Agarose gel showing the results from digestion of Plasmid 2 pet32 Xa/Lic-UK. Lane 1) Ladder, Lane 2) PCR products from the amplification of plasmid A1, Lane 3) PCR products from the amplification of pet32 Xa/Lic vector containing Z-FDPS Lane 4) uncut plasmid 2. Lane 6-10) Different Gibson assembly products digested with HindIII and XbaI, lane 7 contains the 2 expecting bands.	50
Figure 2-5 Left: Western blot gel of the test expression of the UK gene in pet28a (Plasmid 1), pet 32 Xa/Lic (Plasmid 2) and IPK wild type as positive reference for expression. (1-4) E. coli BL21-(DE3), BL21-CodonPlus(DE3)-RIL, BL21-CodonPlus(DE3)-RP and BL21-CodonPlus(DE3)-RIL and C43(DE3) pLysS; (5) E. coli BL21-AI cells; Two proteins of 45 kDa (+) and 35 kDa (-) size were used as positive and negative control respectively for the western blot. Left: SDS-PAGE: analysis of the pellet after expression at various temperature. +: after induction and -: before induction	52

Figure 2-6 Purification of undecaprenol kinase. a. SDS-PAGE: analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography. B. SDS-PAGE: analysis of the pellet (PE) and supernatant (SN) after basic extraction of the pellet after sonication followed by purification by affinity chromatography of the supernatant.52

Figure 2-7 FPLC chromatogram of gel filtration chromatography using Sephadex G-25 resin for desalting the pooled fractions containing UK after nickel affinity chromatography.53

Figure 2-8 UK-catalysed synthesis of FP from farnesol. a. TLC plate with FP and FDP chemically synthesised, lane 1: precipitate from the incubation of farnesol with UK. b. ³¹P NMR spectra recorded over the time of the incubation of farnesol with UK. FP is formed after only 17 min and the intensity of the peak decreases when FP precipitates.54

Figure 2-9 Cartoon representations of an X-ray crystal structure of Mj IPK in complex with IP(PDB: 3K52) a. Full view b. Close view of the active site with residues around the IP tail displayed.55

Figure 2-10 Purification of IPK and variant 3. a. IPK WT: SDS PAGE analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography. b. IPK variant 3: SDS PAGE analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography.56

Figure 2-11 GC-MS analysis of the pentane extractable compounds arising from the incubation of FP with IPK variant, Gd4oIS and analysis of the white precipitate by TLC a. Total ion chromatogram highlighting the formation of one major product. b. Corresponding mass spectrum of the corresponding compound eluting at 13.58 min. c. TLC of the redissolved precipitate in 1% H₃PO₄ in butanol (TLC solvent system NH₄OH:iPrOH:H₂O 6:3:1).....57

Figure 2-12 Cyclodextrins structures. a. 3D structure of β-CD b. Chemical structure of alpha-cyclodextrin (α-CD), beta-cyclodextrin (β-CD), and gamma-cyclodextrin (γ-CD) and 2-hydroxypropyl-β-CD (HP-β-CD).....58

Figure 2-13 Solubility test of FP and FDP at 5 mM. 1&3) FP in buffer. 2&4) FDP in buffer. 5-10) FP in buffer and various CD (5 mM). 11-14) FDP in buffer and various CD (5 mM). 15) FP + FDP in buffer containing MgCl₂ (20 mM) and β-CD (5 mM) HP-β-CD.....59

Figure 2-14 Coupled IPK-sesquiterpene synthase-alkaline phosphatase assay to test for FP activity of variants.....60

Figure 2-15 GC-MS analysis of the heptane extractable compounds arising from one of the coupled assay IPK-GdoIS-AP a. Total ion chromatogram highlighting the formation of two products with a retention time of 13.58 min (gd4ol) and 15.44 min (farnesol). b. Corresponding

mass spectrum of the corresponding compound eluting at 15.44 min, characteristic of farnesol.....	60
Figure 2-16 Results from the coupled assay IPK-Gd4oIS-AP for FP activity with the 10 IPK variants.	61
Figure 2-17 a. One pot enzymatic synthesis of (S)-germacrene D from farnesol catalysed by UK, IPK 3 and GDS. b. Recycling scheme of ATP with PEP.	62
Figure 2-18 Optimisation of the one pot enzymatic synthesis of (S)-germacrene D from FDP. Lane 1: UK (2 μ M), IPK 3 (2 μ M), MgCl ₂ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), GDS (40 μ M), farnesol (5 mM), 37 °C, 20 hours. Lane 2: same as lane 1 with farnesol (2 mM), β -CD (2 mM), HP- β -CD (2 mM). Lane 3: same as lane 2 with, β -CD (4 mM), HP- β -CD (4 mM). Lane 4: same as lane 2 with UK (0.4 μ M), IPK 3 (2 μ M). Lane 5: same as lane 2 with UK (0.2 μ M), IPK 3 (2 μ M).....	63
Figure 2-19 a. Total ion chromatogram of the pentane extractable compounds from the one pot enzymatic synthesis of (S)-germacrene D from farnesol. (S)-germacrene D eluting at 12.26 min and farnesol eluting at 15.44 min. Conditions: UK (0.2 μ M), IPK 3 (2 μ M), MgCl ₂ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), GDS (40 μ M), farnesol (2 mM), 37 °C, 20 hours b. Mass spectrum of compound eluting at 12.26 min corresponding to (S)-germacrene D.	64
Figure 2-20 Analogues of FDP tested with the design one pot enzymatic diphosphorylation.	65
Figure 2-21 a. Total ion chromatogram of 10,11-epoxy-farnesol b. Total ion chromatogram of heptane extractable compounds from the incubation of 10,11-epoxyfarnesol with IPK 3, UK, Gd4oIS and subsequently AP.	66
Figure 2-22 a. Total ion chromatogram of heptane extractable compounds from the incubation of 14,15-dimethylfarnesol with IPK 3, UK, Gd4oIS and subsequently AP. 14,15-dimethylfarnesol eluting at 17.33 min. 14,15-dimethylgermacrene D eluting at 15.62 min. b. TLC of the white precipitate formed at the bottom of the incubation of 14,15-dimethylfarnesol with IPK 3, UK, Gd4oIS.	66
Figure 2-23 a. Total ion chromatogram of heptane extractable compounds from the incubation of 8-methoxy-farnesol with IPK 3, UK, ADS and subsequently AP. 8-methoxy- γ -humulene eluting at 16.85 min. 8-methoxy-farnesol eluting at 17.00 min. b. Mass spectrum of compound eluting at 16.85 min corresponding to 8-methoxy- γ -humulene. The mass spectrum of the product corresponds well with reported literature. ^[78]	67

Figure 2-24 Total ion chromatogram of heptane extractable compounds from the incubation of 12-methoxy-farnesol with IPK 3, UK, ADS and subsequently AP. 12-methoxy-farnesol eluting at 18.28 min.	68
Figure 3-1 Second generation design of (S)-dimethylated germacrene D from dimethylated FDP. a. Top: (S)-14,15-dimethylgermacrene D was found to be an aphid attractant. Bottom: (S)-15-dimethylgermacrene D was found to be an aphid repellent. b. Dimethylated analogues synthesised in this chapter with both C15 methylated.	74
Figure 3-2 Total ion chromatograms of the pentane-extractable products arising from incubation of FDP and analogues with GDS. a. FDP with GDS b. 12,15-dimethyl FDP with GDS c. 6,15-dimethyl FDP with GDS.	79
Figure 3-3 Homology models of GDS with dimethylated analogues in the active site based on 5-epi AS from <i>Nicotiana tabacum</i> complexed with farnesyl hydroxyphosphonate (FHP) (PDB-5EAT) ^[139] . Red arrows indicate the additional unnatural methyl groups.	80
Figure 3-4 Purification of germacradien-4-ol synthase. a. SDS-PAGE: analysis of the supernatant (SN) and pellet (PE) after sonication of the pellet after expression and purification by nickel affinity chromatography. Pierce™ unstained protein MW marker was used as reference of protein size (L).	82
Figure 3-5 Total ion chromatogram of the pentane extractable compounds from the incubation of 10,11-epoxy-FDP with Gdol4S. Major compound eluting at 15.52 min, side products at 15.83 min and 20.16 min (compound eluting at 15.83 min identified to be 10,11-epoxyfarnesene by Dr. Grundy, unknown compound at 20.16 min). b. Mass spectrum of compound eluting at 15.52 min. c. Total ion chromatogram of purified oil from the incubation a.	83
Figure 3-6 Relative quantities of products in a competitive assay of 10,11-epoxy FDP (0.125 mM) and FDP (0.025 mM) with GdolS as judged by relative integration of product peaks in the total ion chromatogram of the pentane extractable products.	84
Figure 3-7 ¹ H-NMR spectra (400 MHz, CDCl ₃) of 146 from -50 °C to +50 °C (10 °C increment).	86
Figure 3-8 ¹ H-NMR spectra (400 MHz, CDCl ₃) in the alkyl region (δ = 1.3 - 1.8 ppm) of 146 from -50 °C to +50 °C (20 °C increment).	87
Figure 3-9 a. observed NOEs for x . b. Sketch of the proposed conformational equilibrium at -50 °C between the major down-down (CH ₃ -14 and 15 compared) conformer of (R)- 146 and the minor up-down form.	87

Figure 4-1 Purification of ADS and AS, red arrows indicating the enzyme. a. ADS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 66 kDa). b. AS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by anion exchange Q-sepharose of the RAW (expected MW: 39 kDa).96

Figure 4-2 Purification of GDS and GDS Y406F, red arrows indicating the enzyme. a. GDS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 63 kDa). b. GDS Y406F: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 63 kDa)97

Figure 4-3 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with ADS a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.53 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.53 min.97

Figure 4-4 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with AS. a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.68 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.68 min.98

Figure 4-5 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with GDS. a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.63 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.63 min.98

Figure 4-6 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with GDS Y406F. a. Total ion chromatogram highlighting the formation of two products, predominant one with a retention time of 15.06 min. b. Corresponding mass spectrum of the corresponding compound eluting at 15.06 min identified as the 14,15-dimethylgermacrene D.99

Figure 4-7 Half-Normal Plot of effects resulting from the fractional factorial two-level design of experiment. In blue negative effects and orange positive effects on the yield..... 102

Figure 4-8 a. Schematic representation of the oscillatory segmented system used. A. reservoir containing pentane B. reservoir containing the enzyme in buffer and FDP. The modular system allows drawing liquid from the reservoir when the valve is closed in position 3 and open in position 1 and 2. Liquid is injected in the flow reactor when valve is switched to position 2 closes and open for 1 and 3. Oscillatory motion in the reactor is created by both syringes simultaneously pushing and drawing liquid with position 2 closed and 1 and 3 open. b. Results from performing the biocatalysis in an oscillatory segment system. Yields were obtained using GC-MS and calculated by using α -humulene as internal standard.104

Figure 4-9 High-performance counter-current chromatography (HPCCC) device. a. Schematic view showing only one of the two bobbins for clarity. (Huynh F. et al., 2020) b. Showing the settling and mixing zone in the HPCCC column generated by variable centrifugal forces induced by the planetary motion of the bobbin. The mixing zone is situated toward the centre of the axis of revolution coinciding with a low acceleration field and the settling zone coinciding with the high acceleration field away from the centre of revolution. c. Photo of the HPCCC showing the two bobbins and its rotation in the apparatus. Photo of the HPCCC extracted and adapted from Dynamic extraction® presentation by Dr. David J. Rooke.....106

Figure 4-10 Comparison of the diffusion pattern of different biphasic processes. a. batch in a stirred tank. b. segmented flow. C. high-performance counter-current chromatography in normal phase, the organic phase is flowing, alternative segments are forming and mixing and settling zones are appearing.107

Figure 4-11 Setting up the HPCCC to perform a biphasic reaction using sesquiterpene synthases and pentane as extraction solvent. (Huynh F. et al., 2020)108

Figure 4-12 Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the HPCCC system (6 μ M AS, 0.35 mM FDP, 0.5 mL \cdot min $^{-1}$ pentane flow rate), determined by GC-FID. (Huynh F. et al., 2020)109

Figure 4-13 Left: Characterisation of the relationship between the pentane flow rate and glycerol content of the aqueous phase on the stationary phase retention S_f . The surface highlights that S_f is severely reduced with glycerol in the buffer and a high flow rate. Right: Characterisation of the relationship between the pentane flow rate and rotation speed of the bobbin on the stationary phase retention S_f . Surface highlights that S_f is severely reduced with low rotation speed and high flow rate. (Surface rendered using Matlab 2018b). S_f : Percentage retention of the stationary phase relative to total column volume. (Huynh F. et al., 2020)..110

Figure 4-14 Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the HPCCC system (6 μM AS, 0.7 mM FDP, 2 $\text{mL}\cdot\text{min}^{-1}$ pentane flow rate), determined by GC-FID. (Huynh F. et al., 2020)..... 111

Figure 4-15 Comparison of the batch, segmented flow and HPCCC method for different substrates by PR-aristolochene synthase (AS), amorphadiene synthase (ADS) and (S)-germacrene D synthase (GDS). Yields were determined by GC-FID and calculated by using a calibration curve plotted with α -humulene. (Huynh F. et al., 2020) 112

Figure 4-16 a. Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the preparative column of the HPCCC system (6 μM AS, 0.7 mM FDP, 5 $\text{mL}\cdot\text{min}^{-1}$ pentane flow rate), determined by GC-FID. b. Isolated yields for different substrates by PR-aristolochene synthase (AS), amorphadiene synthase (ADS) and (S)-germacrene D synthase (GDS). The reaction products are shown performed on a preparative scale with a 135 mL HPCCC column using the optimised conditions (pentane flow rate: 5 $\text{mL}\cdot\text{min}^{-1}$), each reaction was repeated four times to increase amount of product and reduce product loss when removing pentane under reduced pressure. (Huynh F. et al, 2020)..... 113

Figure 4-17 Total ion chromatograms and mass spectra of the pentane extracted products arising from the incubation of 12-OH FDP with ADS by HPCCC. a. Total ion chromatogram with three compounds eluting at 14.83 min (side product), 15.08 min (R-DHAA), 15.12 min (S-DHAA) b. Flame ionisation detection chromatogram (chiral column) with three compounds eluting at 15.03 min, 15.57 min, 15.66 min. c. Total ion chromatogram of purified aldehydes. d-e. Mass spectrum of compound eluting at 15.08 min and 15.12 min. f. Purified side product chromatogram. g. Mass spectrum of compound eluting at 14.83 min..... 114

Figure 4-18 Products arising from the incubation of 12-OH FDP with ADS. a. Products and mechanism proposed by Demiray et al. b. Products and mechanism suggested in this chapter. 116

Figure 4-19 Setting up the HPCCC to perform a biphasic reaction using CalB lipase and heptane as extraction solvent. (Huynh F. et al., 2020)..... 118

Figure 4-20 a. CalB-catalysed transesterification of octanol by vinyl acetate. b. Flame ionisation detection chromatogram of the collected heptane at the exit of the HPCCC with two compounds eluting at 7.74 min (octanol) and 8.81 min. c. Total ion chromatogram of purified compound eluting at 8.81 min in the GC FID. d. Mass spectrum of compound eluting at 7.66 min corresponding to octyl acetate after comparison to NIST library..... 119

Figure 4-21 Influence of the ratio alcohol:vinyl acetate on the conversion of octanol to octyl acetate. a. Batch reaction. b. Reaction performed in the HPCCC, residence time 6 min... 120

Figure 4-22 Influence of the temperature, enzyme concentration and substrate concentration on % conversion in HPCCC. a. Various temperature, reaction condition: 1600 rpm, 2 mL.min⁻¹, CalB (1 mg.mL⁻¹), octanol (30 mM), vinyl acetate (30 mM). b. Various enzyme concentration, reaction condition: 1600 rpm, 2 mL.min⁻¹, 20 °C, octanol (30 mM), vinyl acetate (30 mM). c. Various substrate concentration, reaction condition: 1600 rpm, 2 mL.min⁻¹, CalB (1 mg.mL⁻¹), octanol (30 mM), vinyl acetate (30 mM)..... 121

Figure 4-23 a. 3-Level face centered DoE performed. b. Results from the DoE, conversion depending on the temperature. c. Conversion depending on the rotation speed d. Conversion depending on the residence time. (Huynh F. et al., 2020) 122

Figure 4-24 Result from the DoE, 3D view, conversion depending on temperature, rotation speed and residence time. 122

Figure 4-25 Recycling efficiency of CalB for the transesterification of octanol with vinyl acetate in HPCCC (Octanol 100 mM (1:3 ratio, alcohol:acetate), CalB (1 mg.mL⁻¹), 20 °C, 1600 rpm, 6 min residence time). (Huynh F. et al, 2020) 124

Figure 4-26 a. Flame ionisation detection chromatogram of a racemic mixture of 2-pentyl acetate chemically synthesised 3.45 min (S)-2-pentyl acetate, 4.04 min (R)-2-pentyl acetate. b. Flame ionisation detection chromatogram of the pentane collected at the exit of the HPCCC from the incubation of 2-pentanol with vinyl acetate. (R)-2-pentanol eluted with the solvent. 125

Figure 6-1 Gibson assembly: 5' exonuclease cuts in the 5' to 3' direction creating a 20 bp-long sticky end. 5' end of one segment is complementary to the 3' end of the other for annealing. A polymerase fills in the gaps of the annealed single strands region. A ligase joins segments covalently and fixes DNA nicks..... 148

Figure 6-2 Golden Gate assembly: BsaI cuts DNA outside of its recognition site (shown in red) creating a 4 base pair long sticky end (shown in yellow and blue) then T4 Ligase connects complimentary DNA fragments and fixes DNA nick. 149

Figure A-1 Calibration curve for flame ionization detection response of α-humulene..... II

Figure A-2 Calibration curve for flame ionization detection response of farnesal. II

Figure A-3 Example of a calibration curve for flame ionization detection response of octanol. III

Figure A-4 Flame ionization detection response of AS with FDP by HPCCC using method 2. IV

Figure A-5 Flame ionization detection response of ADS with FDP by HPCCC using method 2. V

Figure A-6 Flame ionization detection response of GDS with FDP by HPCCC using method 2.....	VI
Figure A-7 Flame ionization detection response of ADS with 12-OH FDP by HPCCC using method 2. (11-R and 11-S epimer at 15.571 min 15.567 min).....	VII
Figure A-8 Total ion chromatogram of authentic sample of FDP with aristolochene synthase by HPCCC on the preparative scale method. Inset: mass spectrum of peak at 12.68 min. VIII	VIII
Figure A-9 Total ion chromatogram of authentic sample of FDP with amorphadiene synthase by HPCCC on the preparative scale method. Inset: mass spectrum of peak at 12.53 min. . IX	IX
Figure A-10 Total ion chromatogram of authentic sample of FDP with germacrene D synthase by HPCCC on the preparative scale method. Inset: mass spectrum of peak at 12.63 min. .. X	X
Figure A-11 Total ion chromatogram of authentic sample of 12-OH FDP with amorphadiene synthase by HPCCC on the preparative scale method.....	XI
Figure A-12 Mass spectrum of peak at 15.12 min of total ion chromatogram of authentic sample of 12-OH FDP with amorphadiene synthase by HPCCC on the preparative scale method.....	XII
Figure A-13 Mass spectrum of peak at 15.10 min of total ion chromatogram of authentic sample of 12-OH FDP with amorphadiene synthase by HPCCC on the preparative scale method.....	XIII
Figure A-14 Flame ionization detection response of octanol.	XIV
Figure A-15 Flame ionization detection response of octyl acetate.....	XV
Figure A-16 Example of a flame ionization detection response of the enzymatic transesterification of octanol with vinyl acetate by Cal B lipase (octanol at 7.744 min and octyl acetate at 8.812 min).	XVI
Figure A-17 Flame ionization detection response of chemically synthesised racemic mixture of 2-pentyl acetate.	XVII
Figure A-18 Flame ionization detection response of the enzymatic transesterification of rac-2-pentanol with vinyl acetate by Cal B lipase [(R)-2-pentylacetate at 4.131 min].	XVIII
Figure A-19 Full ¹ H NMR spectrum (400 MHz, CDCl ₃ , 298 K) of macrocyclic ether.....	XIX
Figure A-20 ¹ H NMR spectrum (400 MHz, CDCl ₃ , 298 K) of macrocyclic ether.	XX
Figure A-21 ¹ H NMR spectrum (400 MHz, CDCl ₃ , 223 K) of macrocyclic ether.	XXI
Figure A-22 ¹ H NMR spectrum (400 MHz, CDCl ₃ , 323 K) of macrocyclic ether.	XXII
Figure A-23 ¹³ C NMR spectrum (100 MHz, CDCl ₃ , 323 K) of macrocyclic ether.	XXIII
Figure A-24 COSY NMR spectrum (400 MHz, CDCl ₃ , 323K) of macrocyclic ether.	XXIV

Figure A-25 Edited HSQC NMR spectrum (400 MHz, CDCl ₃ , 323 K) of macrocyclic ether.	XXV
Figure A-26 HMBC NMR spectrum (400 MHz, CDCl ₃ , 323 K) of macrocyclic ether.....	XXVI
Figure A-27 NOESY NMR spectrum (400 MHz, CDCl ₃ , 223 K) of macrocyclic ether.....	XXVII
Figure A-28 ¹ H NMR spectrum (500 MHz, CDCl ₃ , 298 K) of artemisinic-11S,12-epoxide.	XXVIII
Figure A-29 ¹³ C NMR spectrum (125 MHz, CDCl ₃ , 298 K) of artemisinic-11S,12-epoxide.	XXIX
Figure A-30 COSY NMR spectrum (500 MHz, CDCl ₃ , 298K) of artemisinic-11S,12-epoxide.	XXX
Figure A-31 Edited HSQC NMR spectrum (500 MHz, CDCl ₃ , 298 K) of artemisinic-11S,12-epoxide.	XXXI
Figure A-32 HMBC NMR spectrum (500 MHz, CDCl ₃ , 298 K) of artemisinic-11S,12-epoxide.	XXXII
Figure A-33 DEPT 135 NMR spectrum (500 MHz, CDCl ₃ , 298 K) of artemisinic-11S,12-epoxide.	XXXIII

LIST OF TABLES

Table 1.1 General classification of terpenes. Prenyl transferases couple successively five-carbon building blocks to form linear terpenes which undergo cyclisation to form complex cyclic structures (examples are given in the last column).	6
Table 1.2 Comparison of the kinetics characteristics of ADS wild type and mutants.....	39
Table 2-1 Table of created and expressed variant of Mj IPK by SDM.....	56
Table 3-1 Summar of all GDS variant tested with 6,15-dimethyl FDP and 12,15-dimethyl FDP. + sign indicates product was generated. – sign indicates no product was generated.	80
Table 3-2 Optimisation of the incubation of 10,11-epoxyFDP with Gdols using flow segmented system. Yield measured by GC FID using farnesol as internal standard.....	85
Table 3-3 Proton NMR spectroscopic assignment of 146 at -50 °C and +50 °C. [a] All resonances are δ_H (400 MHz, CDCl ₃) in ppm. Entries are chemical shift followed by multiplicity and coupling constants (Hz).	88
Table 4-1 Results of DoE experiments for GDSY406F in flow. Yields were analysed by GC MS and calculated by using heptadecane as internal standard. *yield measured by GC FID. Experiments with (-) did not give any results due to technical issues.....	101
Table 4-2 Incubation conditions for batch, flow and HPLCC set-up.....	111
Table 6-1 List of plasmids used in this project.	138
Table 6-2 Gene sequences of protein used in this project. His Tag underlined. *Gene synthesised by Thermo Fisher including necessary Bsal recognition sequence (underlined), and overhangs (in bold) for Golden Gate cloning.	138
Table 6-3 6-4 Oligonucleotide primers for the insertion of 1 base pair into plasmid containing UK gene in pET28a (5' to 3'). Added base pair depicted in bold.	143
Table 6-5 Oligonucleotide primers for site-directed mutagenesis (5' to 3') of IPK WT with the changed codon in bold and underlined.	144
Table 6-6 List of all variant created using SDM from primers in Table 6-5.	144
Table 6-7 PCR mixture for SDM protocol	145
Table 6-8 PCR for SDM conditions.....	145
Table 6-9 PCR mixture	146
Table 6-10 PCR conditions.....	146
Table 6-11 Recipe of Qiagen DNA purification buffers.	147

Table 6-12 Gibson assembly reaction mixture.....	148
Table 6-13 Golden Gate reaction mixture.....	150
Table 6-14 PCR primer for cloning UK in pET32-Xa Lic	150
Table 6-15 Expression conditions for proteins used in this work	152
Table 6-16 List of cell lysis buffer used in this work.	153
Table 6-17 List of dialysis buffer used in this work.	154
Table 6-18 Gel casting buffers for SDS-PAGE (amount to make 2 gels).	155
Table 6-19 List of incubation buffer and enzyme concentration used in chapter 3.	190

LIST OF SCHEMES

Scheme 1-1 Simplistic model for the mechanism of biocatalysis of FDP by a sesquiterpene synthase to its product. E: enzyme, [E·FDP] enzyme FDP complex, [E·P] enzyme product complex, P: product.....	16
Scheme 1-2 Putative catalytic mechanism of amorpho-4,11-diene synthase.....	21
Scheme 1-3 Putative catalytic mechanism of PRAS.....	23
Scheme 1-4 Putative catalytic mechanism of GdoIS.....	25
Scheme 1-5 Putative catalytic mechanism of (S)-GDS.....	25
Scheme 1-6 Fragmentation of (S)-germacrene D from the incubation with (1R)-[1-2H]-FDP.	26
Scheme 1-7 Proposed mechanism for the conversion of 8-anilino GDP (117) by TEAS.....	33
Scheme 2-1 Potential two-step enzymatic synthesis of FDP analogues from farnesol analogues using UK and IPK.....	47
Scheme 2-2 Coupled IPK-sesquiterpene synthase assay to test for FP activity of variants .	57
Scheme 2-3 One pot enzymatic diphosphorylation scheme of farnesol.	64
Scheme 3-1 Possible products resulting from a Horner-Wadsworth-Emmons olefination. ...	76
Scheme 3-2 Synthesis of 6,15-dimethyl FDP Reagent and conditions i. Triethyl-2-phosphonopropionate (1.1 eq), n-BuLi (1.1 eq), THF, 0 °C, 4 h, ii. DIBAL-H (4 eq), Toluene, 78 °C, 3 h iii. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. iv. Ethyl acetoacetate (3 eq), NaH (3.5 eq), n-BuLi (3.1 eq), THF, 0 °C. v) LiOTf (3 eq), Tf ₂ O (1.3 eq), NEt ₃ (3 eq), DCM, 0 °C, 2 h. vi. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. vii. DIBAL-H (3 eq), Toluene, -78 °C, 1h. viii. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. xi. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method ^[109] see chapter 2).	76
Scheme 3-3 Synthesis of 12,15-dimethyl FDP Reagent and conditions i. SeO ₂ (0.1 eq), Salicylic acid (0.1 eq), tBuOOH (0.1 eq), DCM, 0 °C to R.T., 40 h then NaBH ₄ (2 eq), EtOH. ii. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. iii. MeMgBr (5 eq), Et ₂ O, 0 °C, 2h. iv. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. v. Ethyl acetoacetate (3 eq), NaH (3.5 eq), nBuLi (3.1 eq), THF, 0 °C. vi. LiOTf (3 eq), Tf ₂ O (1.3 eq), NEt ₃ (3 eq), DCM, 0 °C, 2 h. vii. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. viii. DIBAL-H (3 eq), Toluene, -78 °C, 1h. ix. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. x. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method ^[109] see chapter 2).....	77

Scheme 3-4 Proposed reaction mechanism for the GdIS catalysed conversion of epoxide x to the medium sized terpene ether 146 .	81
Scheme 3-5 i. NBS, THF/H ₂ O ii. K ₂ CO ₃ , MeOH. iii. LiCl (4 eq), S-collidine (6 eq), MsCl (2 eq), DMF, 3h, 0 °C iv. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method ^[109] see chapter 2).	83
Scheme 3-6 Synthesis of 14,15-dimethyl FDP Reagent and conditions. i. Ethyl acetoacetate (3 eq), NaH (3.5 eq), n-BuLi (3.1 eq), THF, 0 °C. ii. LiOTf (3 eq), Tf ₂ O (1.3 eq), NEt ₃ (3 eq), DCM, 0 °C, 2 h. iii. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. iv. DIBAL-H (3 eq), Toluene, -78 °C, 1h. v. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. vi. Ethyl acetoacetate (3 eq), NaH (3.5 eq), n-BuLi (3.1 eq), THF, 0 °C. ii. LiOTf (3 eq), Tf ₂ O (1.3 eq), NEt ₃ (3 eq), DCM, 0 °C, 2 h. vii. LiOTf (3 eq), Tf ₂ O (1.3 eq), NEt ₃ (3 eq), DCM, 0 °C, 2 h. viii. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. ix. DIBAL-H (3 eq), Toluene, -78 °C, 1h. x. PBr ₃ (0.5 eq), THF, -10 °C, 30 min. xi. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method ^[109] see chapter 2).	89
Scheme 3-7 Synthesis of 12-OH FDP Reagent and conditions. i. LiCl (4 eq), 2,4,6-collidine (6 eq), MsCl (2 eq), DMF, 3h, 0 °C ii. SeO ₂ (0.1 eq), Salicylic acid (0.1 eq), tBuOOH (0.1 eq), DCM, 0 °C to R.T, 20 h. iii. (Bu ₄ N) ₃ OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method ^[109] see chapter 2).	90
Scheme 3-8 Synthesis of 8-OMe FDP and 12-OMe FDP. Reagent and conditions. i. 3,4-dihydropyran (1.1 eq.), pyridinium p-toluenesulfonate (0.1 eq), DCM 1 h. ii. SeO ₂ (0.1 eq), salicylic acid (0.1 eq), tBuOOH (0.1 eq), DCM, 0 °C to R.T., 20 h. iii. NaH (1.5 eq), MeI (5 eq), THF, 0 °C, 24 h. iv. p-toluenesulfonic acid (0.1 eq), MeOH, 1 h.	91
Scheme 4-1 Mechanism of Cal B-catalysed transesterification of vinyl acetate with a primary alcohol.	117
Scheme 4-2 Stereoselective transesterification of 2-pentanol by CalB lipase using vinyl acetate.	124

LIST OF ABBREVIATIONS

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

Ac – Acetate

ADP – Adenine diphosphate

ADS – Amorphadi-4,11-diene Synthase

AS – Aristolochene Synthase

ATAS – Aristolochene synthase from
Aspergillus terreus

ATP – Adenine triphosphate

β ME – Beta mercaptoethanol

n-BuLi – *n*-Butyl lithium

CoA / HSCoA – Coenzyme A

CTP – Cytidyl phosphate

DCM – Dichloromethane

DCS – δ -Cadinene Synthase

DHAAI – Dihydroartemisinic aldehyde

DHP – 3,4-Dihydro-2H-pyran

DIBAL-H – Diisobutylaluminium hydride

DIPEA – Diisopropylethylamine

DMADP – Dimethyl allyl diphosphate

DMAP – Dimethyl amino pyridine

DMF – Dimethyl formamide

DMS – Dimethyl sulphide

DOE – Design of experiment

DOX – Deoxyxylulose phosphate

DPM – 5-diphosphomevalonic acid

DTT – Dithiothreitol

DXP – Deoxyxylulose 5-phosphate

DXPS – 1-Deoxy-D-xylulose-5-phosphate
synthase

E β -FS – (*E*)- β -farnesene Synthase

E. coli – *Escherichia coli*

ee – Enantiomeric excess

EI – Electron ionisation

EIZS – *Epi*-isozizaene synthase

ER – Endoplasmic reticulum

EZS – 7-Epizingiberene Synthase
 FDA – Food and Drug Administration
 FDP – (*E,E*)-Farnesyl diphosphate
 FPLC – Fast protein liquid chromatography
 GA-3P – D-glyceraldehyde-3-phosphate
 GAS – Germacrene A Synthase
 GC-MS – Gas chromatography – Mass spectrometry
 GdoIS – Germacradien-4-ol synthase
 GDP – Geranyl diphosphate
 GDS – (*S*)-Germacrene D Synthase
 GGDP – Geranylgeranyl diphosphate
 GFDP – Geranylfarnesyl diphosphate
 h – hour
 HEPES – 2[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
 His-tag – Histidine tag
 HMG – 3-Hydroxy-3-methylglutaryl
 HPLC – High performance liquid chromatography
 HSQC – Heteronuclear Single Quantum Coherence
 HWE – Horner-Wadsworth-Emmons
 IDP – Isopentenyl diphosphate
 IDPI – Isopentenyl diphosphate isomerase
 IP – Isopentenyl phosphate
 IPTG – Isopropyl- β -D-1-thiogalactopyranoside
 IspC/D/F/H/G – Reductoisomerase
 kDa - Kilo Dalton
 LB medium – Lysogeny broth medium
 MEP – 2-C-methyl-D-erythritol-4-phosphate
 MEcPP – 2-C-methyl-D-erythritol 2,4-cyclodiphosphate
 min – minute(s)
 MK – mevalonate kinase
 MVA – Mevalonate
 NADP+ – Nicotinamide adenine dinucleotide phosphate
 NADPH – Nicotinamide adenine dinucleotide phosphate hydride
 NBS – N-bromosuccinimide
 NCS - N-chlorosuccinimide
 NDP – Neryl diphosphate
 Ni-NTA – Nickel affinity column
 NIST – National Institute of Standards and Technology
 NMR – Nuclear magnetic resonance
 NOESY – Nuclear Overhauser Effect Spectroscopy
 OD600 – Optical density at 600 nm
 ODP – Diphosphate moiety
 pI – Isoelectric point
 PMK – phospho-MVA kinase
 PPMD – phosphomevalonate decarboxylase
 ppm – Parts per million
 PPTS – Pyridinium para-toluenesulfonate
 PRAS – Aristolochene synthase from *Penicillin roquefortii*
 pTSA – para-Toluene sulfonic acid
 Q-seph – Q-sepharose
 RNA – Ribonucleic acid
 r.t. – Room temperature

Sat. – Saturated
SBS – Santalene and Bergamotene
Synthase
SDM – Site directed mutagenesis
SDS PAGE - Sodium dodecyl sulfate
polyacrylamide gel
STS. – sesquiterpene synthase(s)
TB medium – Terrific-Broth medium
TEAS – 5-*epi*-aristolochene synthase from
Nicotiana tabacum
TDP – Thiamine diphosphate

THF – Tetrahydrofuran
TLC – Thin layer chromatography
TMSBr - Trimethyl silyl bromide
TMSCl – Trimethyl silyl chloride
Ts – Tosylate
TsCl – Tosyl chloride
UV-vis – Ultra violet-visible
WT – Wild type
Z-FDP – (2Z,6Z) farnesyl diphosphate
ZS – Zingiberene Synthase

CHAPTER 1

INTRODUCTION

1.1 TERPENES

Terpenes and terpenoids belong to a large family of natural products characterised by an impressive variety of chemical structures, functionalities, and bioactivities. More than 80 000 different terpenoids are known to date, all arising from successive exquisite regio- and stereo-specific cyclisation derived from the five-carbon isoprene unit to yield complex polycyclic structures with several stereocenters.^[1,2] These complex cyclisation reactions are catalysed by a class of enzymes called terpene synthases.^[2] Terpenes have been identified in all domains of life, from microorganisms to insects, but most terpenes are found in plants.^[3-6] Terpenes show a broad range of biological functions including roles in electron carriers (prenylquinones, coenzyme Q₁₀, **1**),^{[7][8]} photosynthesis (chlorophyll A, **2**), hormones ((*S*)-abscisic acid, **3**),^[9] and semiochemicals [(*E*)- β -farnesene], **4**).^[10,11] It is this range of bioactivities that has drawn significant interest from the medicinal, agricultural, fragrance and food industries with compounds worth several hundreds of pounds per gram (Figure 1-1). Examples of commercially utilised terpenoids include artemisinin (**6**), which is widely used as an antimalarial drug;^[12] paclitaxel (Taxol™, **5**) as a treatment for several cancers,^[13] and (*R*)-(+)-limonene (**7**) or (1*R*,2*S*,5*R*)-(-)-menthol (**8**) for flavouring. Until recently, production of these compounds was performed exclusively from plant extraction; for example, artemisinin is still extracted from *Artemisia annua*. Rising demands, limited long-term reserves and high costs have brought increased interest in using bio-synthetic tools as a sustainable alternative for the production of terpenes.^[14] The work presented in this thesis focuses on discovering new valuable terpene products and exploring novel procedures as well as optimising existing processes for the synthesis of terpenes through biological tools.

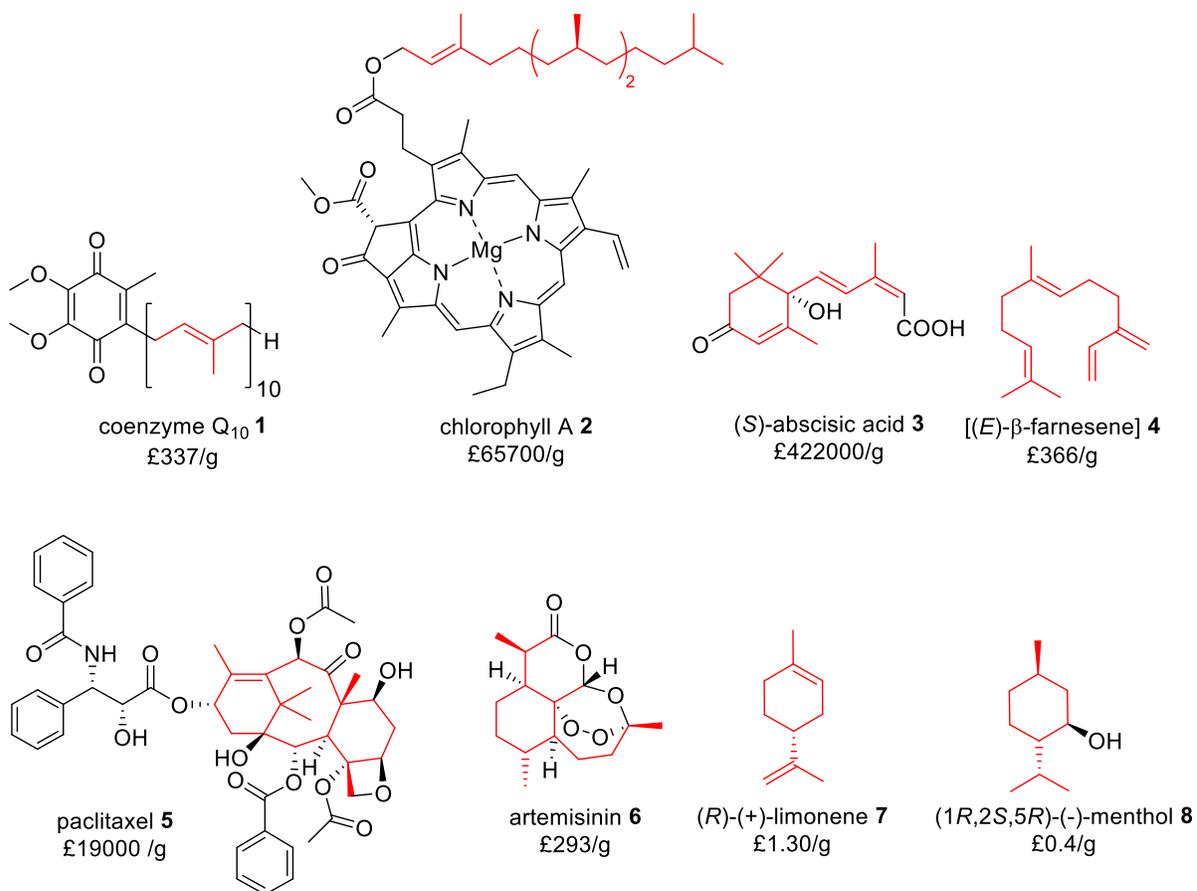


Figure 1-1 Examples of commercially available terpenes with the terpene skeleton highlighted in red (Aldrich prices)^a

1.1.1 Terpene biosynthesis

Virtually all terpenes originate from the same two simple universal building blocks, isopentenyl diphosphate (IDP, **9**) and its isomer dimethylallyl diphosphate (DMADP, **10**).^[2] For example, condensation of DMADP with IDP in a regular fashion (i.e. head-to-tail) forms a C₁₀ linear chain called geranyl diphosphate (GDP, **11**). Subsequent additions of IDP by prenyltransferases form longer, linear terpenes that are consecutively subjected to cyclisation by terpene synthases.^[2] Those cyclic terpenes often undergo further transformations such as oxidations by cytochrome P450s to form a wide variety of metabolites called terpenoids such as artemisinin (**6**).^[15]

^a Prices can be significantly lower from other suppliers, consulted on 30/09/2019

Due to their large structural diversity, a systematic nomenclature has been created. They are classified according to the “biogenetic isoprene rule” introduced by Ruzicka in 1953 which classifies each structure by the number of isoprene building blocks (**12**) they are composed of (commonly used acronyms are summarised in Table 1.1).^[16] On the other hand, the appellations “terpenoid” and “terpene” are not used consistently throughout the literature as there is currently no consensus on their definitions. In this work, a terpene is defined as the direct product from a terpene synthase, whereas terpenoids result from further biomodification of the terpene. Examples of terpenes and terpenoids are given in Table 1.1. GDP (**11**) can be cyclised to the monoterpene (*S*)-limonene (**13**) by limonene synthase, farnesyl diphosphate (FDP, **14**) is converted by trichodiene synthase to the sesquiterpene (1*S*,4*S*)-trichodiene (**15**), geranylgeranyl diphosphate (GGDP, **16**) to (4*aS*, 6*S*)-taxadiene (**17**) by taxadiene synthase and can be further oxidised to the diterpenoid paclitaxel (**5**). Alternatively, head-to-head coupling of farnesyl diphosphate (**12**) and geranylgeranyl diphosphate (**13**) form squalene (**15**) which then forms epoxy-squalene which can cyclise to the precursor of cholesterol (**17**).

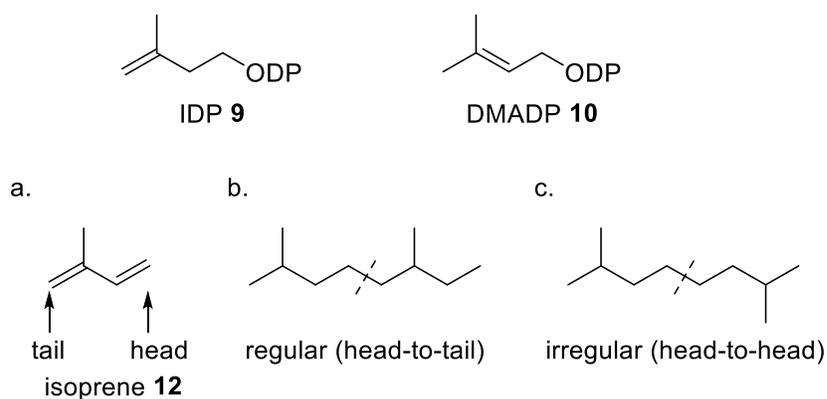
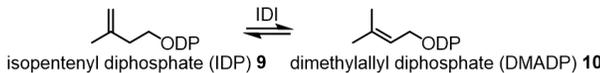
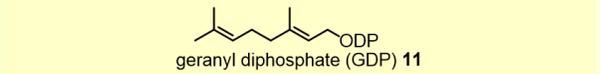
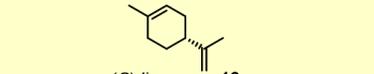
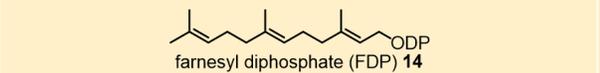
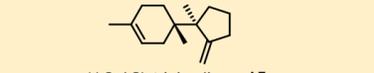
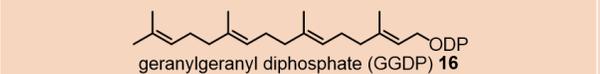
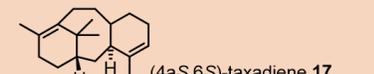
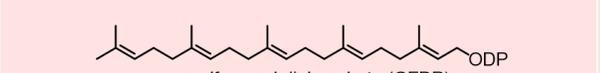
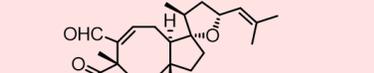
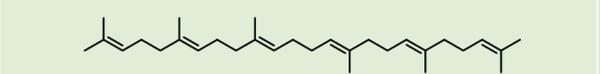
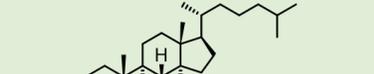
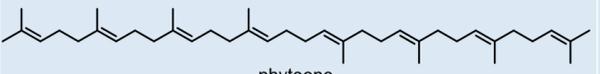
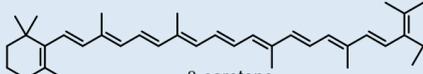


Figure 1-2 Structures of a) isoprene b) regular c) irregular condensation of isoprene units.^[16]

Table 1.1 General classification of terpenes. Prenyl transferases couple successively five-carbon building blocks to form linear terpenes which undergo cyclisation to form complex cyclic structures (examples are given in the last column).

Prefix	C _x	Linear isoprenoid precursor	Example of cyclic product
Hemi-	5	 isopentenyl diphosphate (IDP) 9 dimethylallyl diphosphate (DMADP) 10	
Mono-	10	 geranyl diphosphate (GDP) 11	 (S)-limonene 13
Sesqui-	15	 farnesyl diphosphate (FDP) 14	 (1S,4S)-trichodiene 15
Di-	20	 geranylgeranyl diphosphate (GGDP) 16	 (4aS,6S)-taxadiene 17
Sester-	25	 geranylfarnesyl diphosphate (GFDP)	 ophiobolin A
Tri-	30	 squalene 15	 cholesterol 17
Tetra-	40	 phytoene	 β-carotene

IDP (**9**) and DMADP (**10**) are produced along two currently known pathways depending on the organism and subcellular location they are produced in. The mevalonate pathway (MVA), found in all eukaryotes, plants and archaea is localised in the cytosol with recent studies suggesting certain enzymes of the MVA pathway are also located in the mitochondria, endoplasmic reticulum (ER) and peroxisomes. The deoxyxylulose 5-phosphate pathway (DXP) (also called the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway) is found in bacteria cytosol and plant plastids.^[17] Monoterpenes, diterpenes, and tetraterpenes are mostly produced in the plastids whereas sesquiterpenes and triterpenes are mostly synthesised in the cytosol.^[18] Interestingly, the two pathways are not equally found in all three domains of life (eukaryotes, bacteria and archaea). For example, archaea only possess the MVA pathway and two decades ago, Grochowski *et al.* discovered that archaea use different variation of the MVA pathway.^{[19][20]}

1.1.2 Mevalonate pathway

The MVA pathway is composed of the following chemical steps: first a Claisen condensation of two molecules of acetyl-CoA (**19**) to acetoacetyl-CoA (**20**) by acetyl-coenzyme A (CoA) C-acetyltransferase (or acetoacetyl-CoA thiolase). Acetoacetyl-CoA (**20**) then undergoes aldol condensation with another acetyl-CoA (**19**) to form (3*S*)-3-hydroxy-3-methylglutaryl-CoA (HMG-CoA, **21**) via HMG-CoA synthase. HMG-CoA (**21**) is reduced twice to (*R*)-mevalonic acid (MVA, **22**) by nicotinamide adenine dinucleotide phosphate (NADPH) dependent enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR). The resulting alcohol is successively phosphorylated twice by adenosine triphosphate (ATP)-dependent MVA kinase (MK) and phospho-MVA kinase (PMK) to give (*R*)-5-diphosphomevalonic acid (DPM, **23**). A last phosphorylation of DMP (**24**) causes decarboxylation by MVA diphosphate decarboxylase (MPDC) to generate IDP (**9**). IDP (**9**) can isomerise to DMADP (**10**) in an approximately 3:1 ratio (DMADP:IDP) via IDP isomerase (IDPI) (Figure 1-3).^[21]

1.1.3 Deoxyxylulose 5-phosphate pathway

The DXP pathway starts from the condensation of pyruvate (**25**) and D-glyceraldehyde-3-phosphate (GA-3P, **26**) catalysed by thiamine diphosphate (TDP) dependent 1-deoxy-D-xylulose 5-phosphate synthase (DXS) to give 1-deoxy-D-xylulose 5-phosphate (DXP, **27**). A subsequent intramolecular rearrangement and reduction by NADPH dependent DXP reductoisomerase (DXR or IspC) yields 2-*C*-methyl-D-erythritol-4-phosphate (methylerythritol phosphate, MEP, **28**). Transfer of cytidyl phosphate onto **28** via CDP dependent 2-*C*-methyl-D-erythritol 4-phosphate cytidyltransferase (MCT or IspD) results in 4-(cytidine 5' - diphospho)-2-*C*-methyl-D-erythritol (CDP-ME). The product is then phosphorylated on the 2-hydroxyl group of CDP-ME by 4-(cytidine 5' - diphospho)-2-*C*-methyl-D-erythritol kinase (CMK or IspE) and cyclisation of the product 2-phospho-4-(cytidine 5'-diphospho)-2-*C*-methyl-D-erythritol (CDP-ME2P, **29**) by 2-*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS or IspF) gives 2-*C*-methyl-D-erythritol 2,4-cyclodiphosphate (MEcPP, **30**). Finally, two consecutive reductions of MEcPP by 4-hydroxy-3-methylbut-2-enyldiphosphate (HMBPP) synthase to give 1-hydroxy-2-methyl-2-(*E*)-butenyl diphosphate (**31**) followed by HMBPP reductase result in formation of both DMADP (**10**) and IDP (**9**) in a ratio of 1:5. The mechanism of the last two reductive steps has still not been fully elucidated but is believed to involve two single-electron transfer in a [4Fe-4S] cluster (Figure 1-3).^[22]

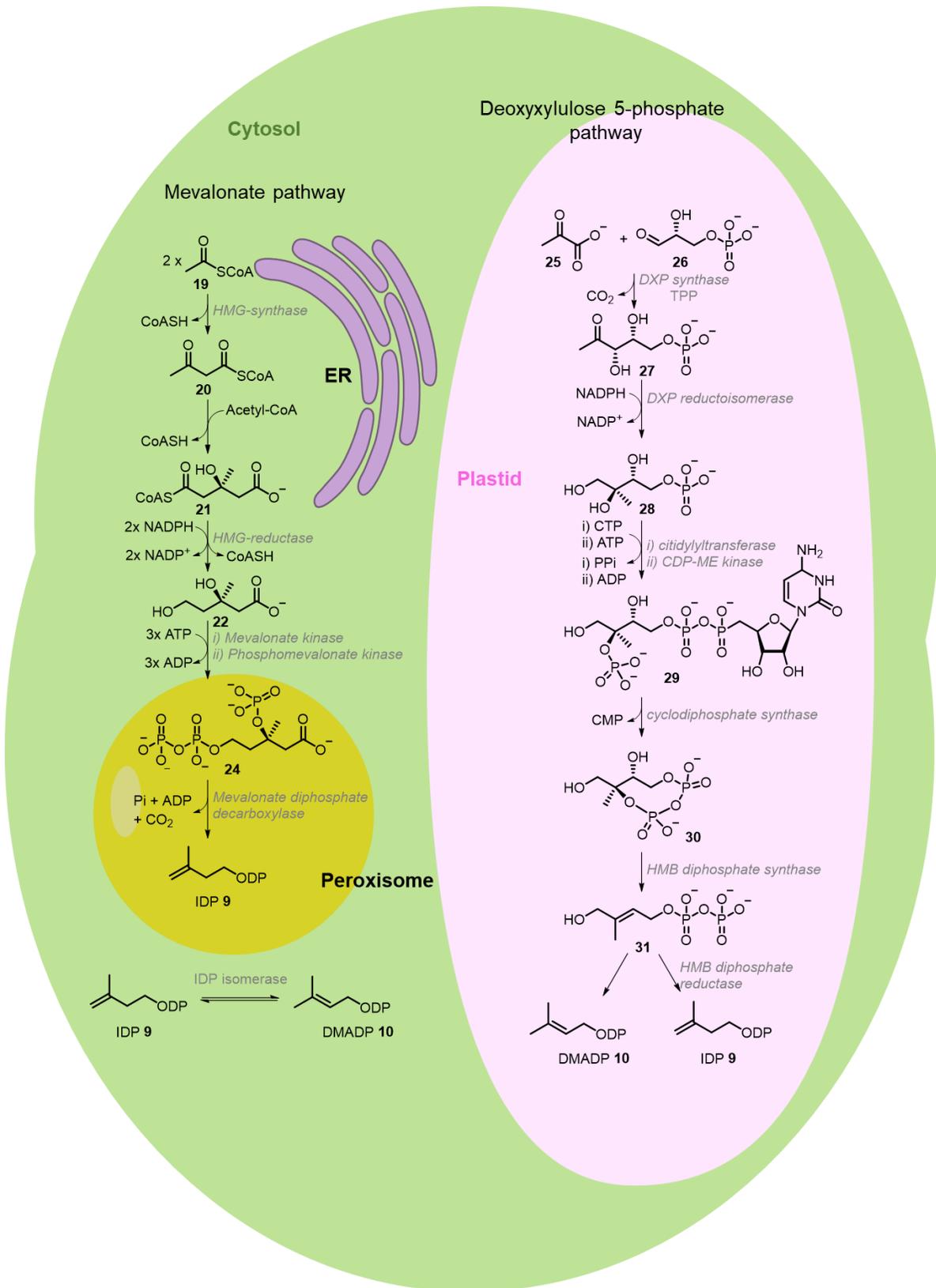


Figure 1-3 MVA and DXP pathways.

1.1.4 The lost alternate mevalonate pathway in archaea

Archaea have been shown to use the MVA pathway to produce IDP (**9**) and DMADP (**10**). However, characterisation of all the putative enzymes in the previously described MVA pathway was incomplete until recently. Enzymes involved in the formation of IDP (**9**) from phosphomevalonate remained unknown for a long time due to low sequence identity in genome analysis for PMK and PPMD.^[23–25] It is only since 2013 that an alternate pathway to the MVA pathway has been fully elucidated for archaea. In 2006, Grochowski *et al.* identified a gene encoding for an ATP-dependent kinase later called isopentenyl kinase (IPK) catalysing the conversion of isopentenyl phosphate (IP, **33**) to IDP (**9**) in *Methanocaldococcus jannashii* (Mj).^[17,20,26] In 2013, mevalonate-5-phosphate decarboxylase was identified to catalyse the reaction from (*R*)-phosphomevalonate (**34**) to IP (**33**) uncovering the mystery behind the hypothetical alternate mevalonate pathway.^[27] Since then, novel pieces of evidence of other alternative MVA pathways have been found.^[28]

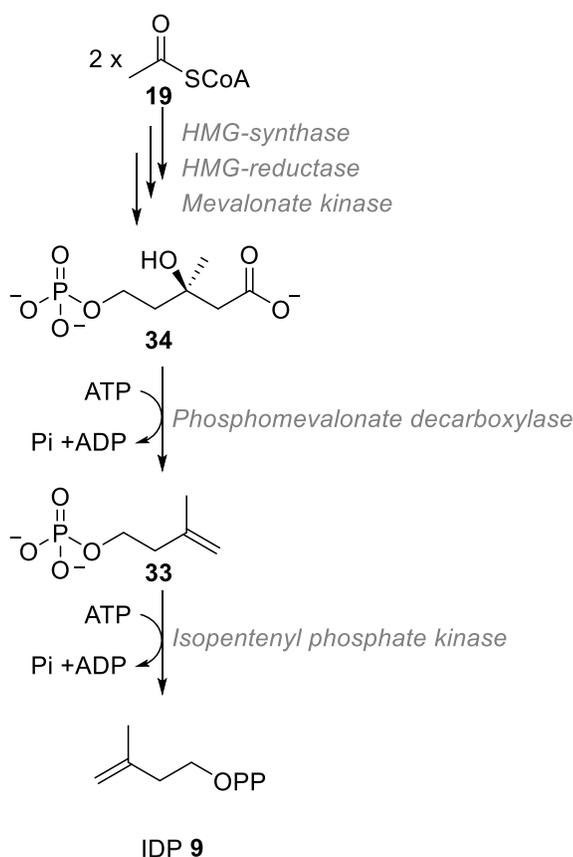


Figure 1-4 Alternative mevalonate pathway found in archaea.

1.1.5 Isoprenyl diphosphate elongation

Synthesis of longer linear isoprenyl diphosphates mentioned in 1.2 such as GDP (**11**), FDP (**14**) or GGDP (**17**) is catalysed by isoprenyl diphosphate synthases (IDPS) called GDP synthase (GDPS), FDP synthase (FDPS) or GGDP synthase (GGDPS) respectively. In the case of GDP (**11**), catalysis starts with ionisation of DMADP (**10**) favoured by a trinuclear magnesium coordination to the diphosphate group followed by nucleophilic attack of IDP (**9**) double bond onto the newly formed dimethylallyl cation.^[2] A final stereoselective deprotonation of the *pro-R*-hydrogen at C2 gives GDP (**11**).^[29] For FDPS or GGDPS, GDP (**11**) stays in the active site and additional IDPs are added in the same manner to generate C₁₅ and C₂₀ linear prenyl diphosphates.

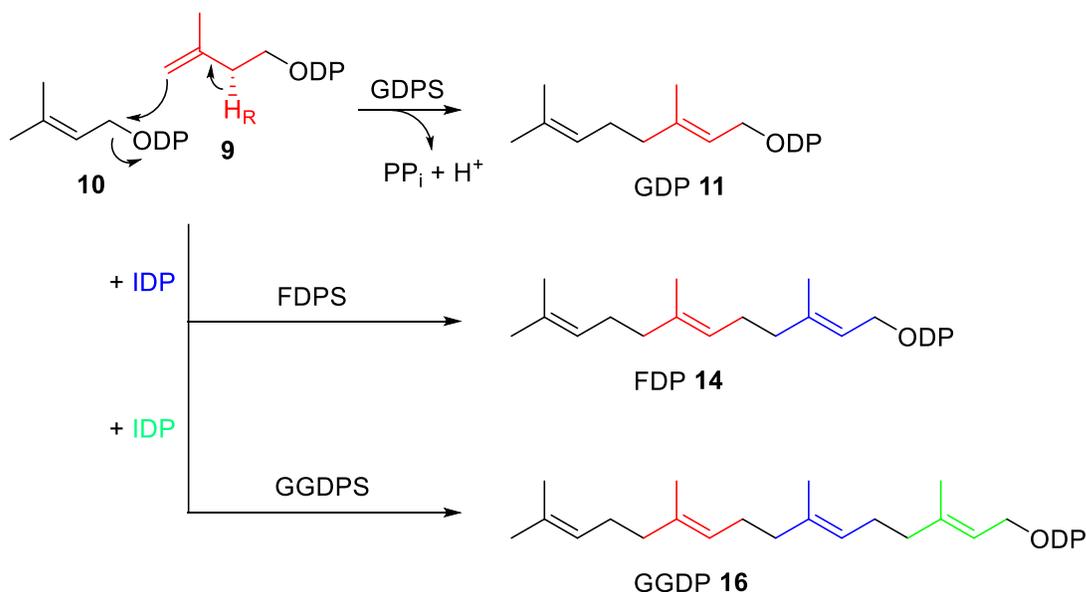


Figure 1-5 Mechanistic representation of *trans* prenyl diphosphate biosynthesis.

1.1.6 Terpene synthases

The conversion of linear isoprenyl diphosphates to cyclic products is catalysed by terpene synthases. Terpene cyclisations are one of the most complex reactions found in nature. From the small pool of acyclic substrates mentioned above (Figure 1-5) more than ten thousand cyclic terpene products can be produced. The active site of terpene synthases is host to complex multistep cascade reactions involving carbocation intermediates changing their

hybridization, bonding and stereochemistry.^[2] Such reactive species can only exist because terpene synthases have a nonpolar active site pocket with protected dipoles and aromatic residues that shelter and stabilise carbocations from other reactive species.^[2] Remarkably, some synthases have evolved to accommodate polar groups such as water in their active site to form specific hydroxylated products.

Terpene synthases can be divided into two major classes depending on how the reaction is initiated. Class I synthases use a Mg^{2+} binding motif to coordinate the diphosphate group of the polyprenyl diphosphate, promoting ionisation of the substrate through cleavage of the diphosphate moiety and formation of an allyl cation.^[2] In contrast, class II synthases form the initial carbocation by aspartic acid catalysed protonation of the distal double bond (or an epoxide derivative) to form a tertiary carbocation.^[2] Structurally, class I synthases usually have their active site in the middle of an α -helical domain containing an aspartate-rich motif DDXXD/E and a NSE/DTE (N,D)DXX(S,T)XXXE which binds three Mg^{2+} .^[30] In rare cases, two aspartate-rich motifs and no NSE/DTE motif are found.^[31] On the other hand, class II active sites are located between two α -helical domains and are characterised by an acid motif DDXD.

This work will focus on one specific sub-class of terpene synthases: the sesquiterpene synthases utilizing the 15-carbon chain FDP (**14**) as a substrate.

1.2 SESQUITERPENES

Sesquiterpenes are produced by sesquiterpene synthases (STS). Several hundred sesquiterpenes are known to date all resulting from stereo, chemo, regio- specific cyclisation of only one common precursor: FDP (**14**).^[2] A wide variety of sesquiterpene is found in nature from a-, mono-, bi- and tri-cyclic structures. Some cyclases are high fidelity, that is to say only a single product is generated but majority of cyclases are promiscuous and generate multiple products. An extreme example is γ -humulene synthase which produces 52 different sesquiterpene products.^[32] Over the last past four decades, research development has enabled us to gain a deeper understanding of terpene synthase chemistry to exploit these enzymes. Manipulating the active site of terpene synthases or feeding them non-natural substrates has led to the discovery of new terpenoids with potentially useful applications through their bioactivity.

1.2.1 Structure

Majority of plant sesquiterpene synthases belong to the class I and are composed of two α -helical domains while synthases from fungi and bacteria are single domain.^[2] Catalysis takes place within the C-terminal α -helical bundle, where FDP binds to a trinuclear metal cluster to then be converted to a sesquiterpene product (see Figure 1-6).^[2]

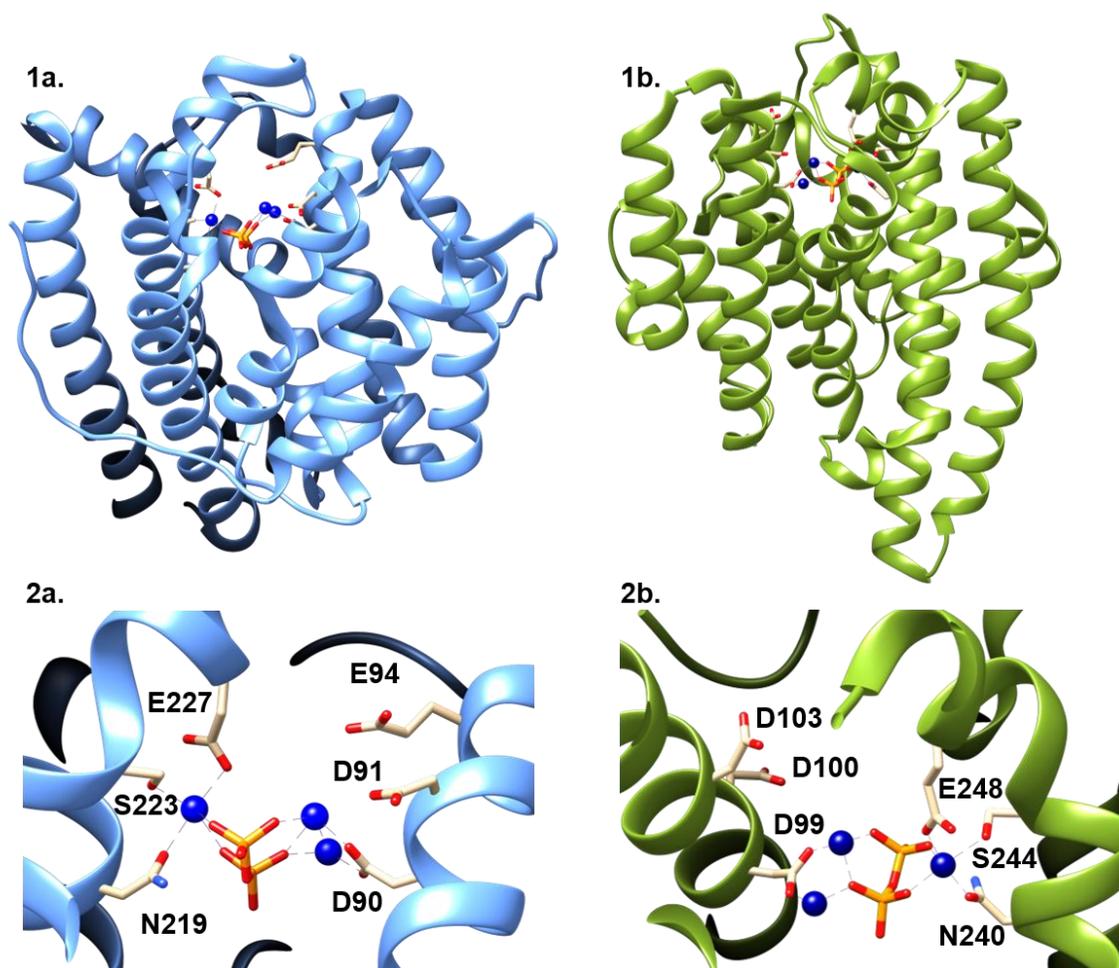


Figure 1-6 Cartoon image of crystal structures of two sesquiterpene synthases 1a. Aristolochene synthase (ATAS) from *Aspergillus terreus* complexed with three Mg^{2+} and diphosphate (PPI) 1b. Epi-isozizaene synthase (EIZS) from *Streptomyces coelicolor* complexed with $[Mg^{2+}]_3$ -PPI. (PDB: 2OA6 and 3KB9). Magnesium ions are shown in blue balls, PPI as a stick structure where oxygen atoms are red and phosphorus is orange. 1c. Close up on ATAS active site highlighting the NSE and DDXXE motif 1d. Close up on EIZS highlighting the NSE and DDXXE motif.^[33,34]

The function of the second domain, the N-terminal domain is still under investigation although it has been proposed it might assist the catalytic C-terminal domain by ensuring a correct folding, desolvating and protecting the active site. The metal cluster is made of three magnesium ions. Magnesium cations are pivotal for terpene catalysis and all sesquiterpene cyclases have in common two Mg^{2+} binding sites that coordinate three Mg^{2+} which in turn interact with the substrate FDP (**14**). The first region is typically an aspartate-rich motif “DDXXD/E” binding two Mg^{2+} and the second is a “NSE/DTE” motif binding a third Mg^{2+} .^[35] Some sesquiterpene synthases do not follow this rule including (*R*)-germacrene D synthase from *Solidago canadensis* that has an NDTYD motif instead of a DDXXD motif and δ -cadinene synthase from *Gossypium arboreum* that has a second DDXXD motif instead of a DTE motif.^[31] Nevertheless, these two aspartate-rich regions possess the same function, which is binding three magnesium ions to anchor FDP within the active site through the coordination of its diphosphate group. Catalysis is then initiated by the cleavage of the C-O bond to form the highly reactive farnesyl cation intermediate. Formation of the farnesyl cation (**35**) results in a series of carbocationic rearrangements chaperoned by the active site, it eventually ends by quenching of the final carbocation to form the sesquiterpene product.^[35]

1.2.2 Origin of sesquiterpene diversity and its mechanism

As explained above, the complex reaction cascade of sesquiterpenes is initiated by cleavage of the diphosphate group yielding (*2E,6E*)-farnesyl cation (**35**). This farnesyl cation (**35**) can only undergo [1,10] or [1,11] cyclisation due to structural constraint resulting in (*E,E*)-germacradienyl cation (**36**) and (*E,E*)-humulyl cation (**37**) respectively. This farnesyl cation can also isomerise through a proposed tertiary diphosphate specie leading to (*2Z, 6E*)-farnesyl cation (nerolidyl cation, **38**). **38** can in turn undergo, [1,10], [1,11] as well as [1,6] and [1,7] cyclisation resulting in (*Z,E*)-germacradienyl cation (**39**), (*Z,E*)-humulyl cation (**40**), bisabolyll cation (**41**) and cycloheptanyl cation (**42**). This myriad of possibilities together with additional hydride, methyl shifts, second cyclisation and rearrangements show how a single precursor can form several hundreds of different compounds. Alternatively, the farnesyl cation (**35**) can be quenched by water or lose a proton, terminating the reaction in early stages to form linear products such as nerolidol (**43**) and farnesenes.^[2,36,37]

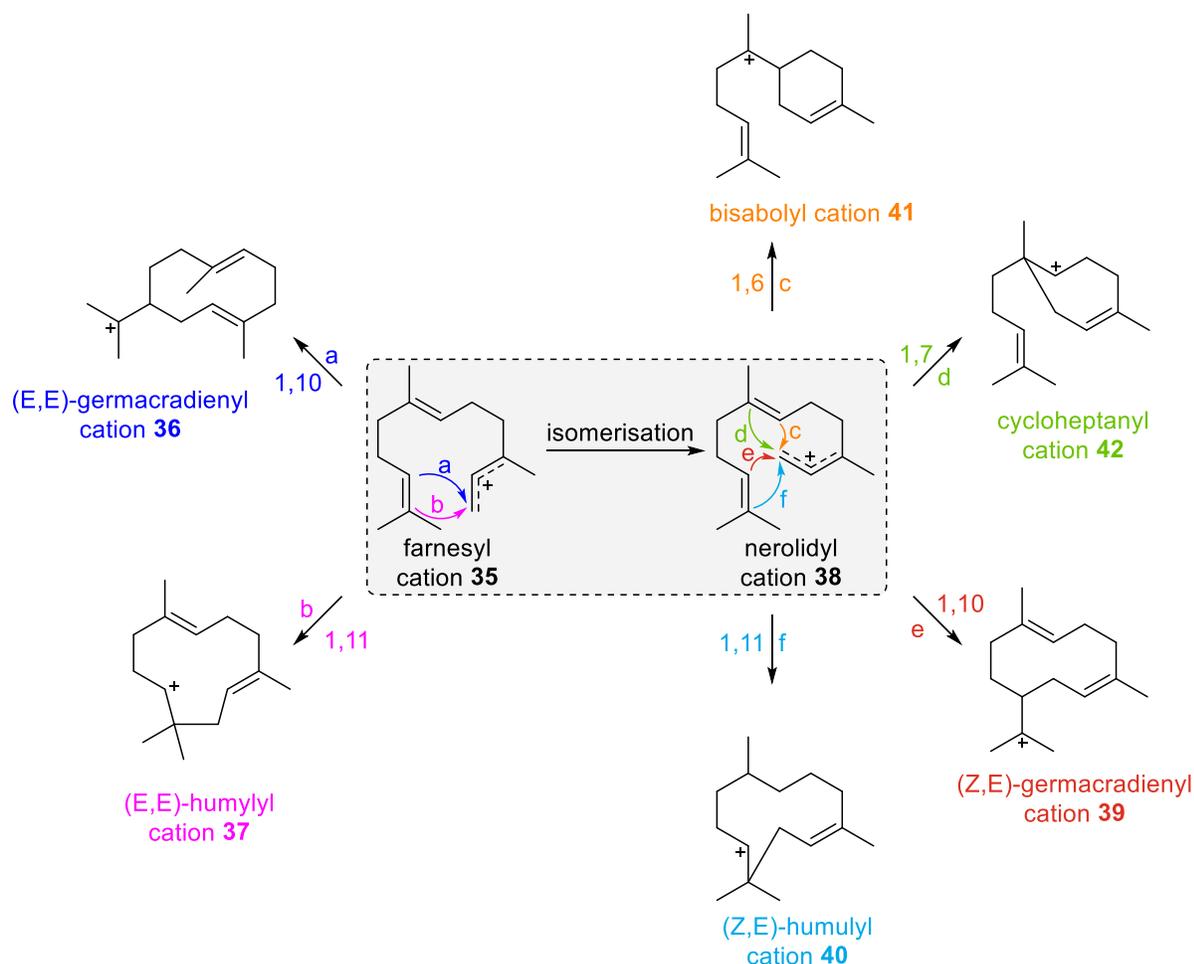


Figure 1-7 Scheme representing first steps of reaction mechanism of all sesquiterpene synthases. Loss of diphosphate leads to the formation of a highly reactive carbocation. The resulting farnesyl cation can undergo a large range of cyclisation bringing diversity to sesquiterpenes backbones.

Key understanding of how STS protect active species from early quenching was achieved two decades ago from the first few successful crystallised STS: 5-*epi*-aristolochene synthase (TEAS) from *Nicotiana tabacum* and (1R,7R,8aS)-aristolochene synthase from *Aspergillus terreus* (AT-AS).^[33,38] Together with later studies using 2-fluoro-farnesyl diphosphate (2F-FDP, **44**) as an inhibitor, researchers showed that exclusion of water from the active site is achieved by enzyme conformational change upon substrate binding.^[39] During catalysis, the active site is closed by electrostatic hydrogen bonds using the trinuclear magnesium cluster bound to the diphosphate group and DDXXD motif as well as Arg and Lys residues. Twelve different structures of ATAS were crystallised with varying numbers of magnesium in the active site (1 to 3), with and without the diphosphate group in a closed and open conformation of the

protein.^[33] Molecular dynamics simulations by Shishova *et al.* supported by Van der Kamp *et al.* proposed the following sequence for the STS catalysis cycle (see Figure 1-8).

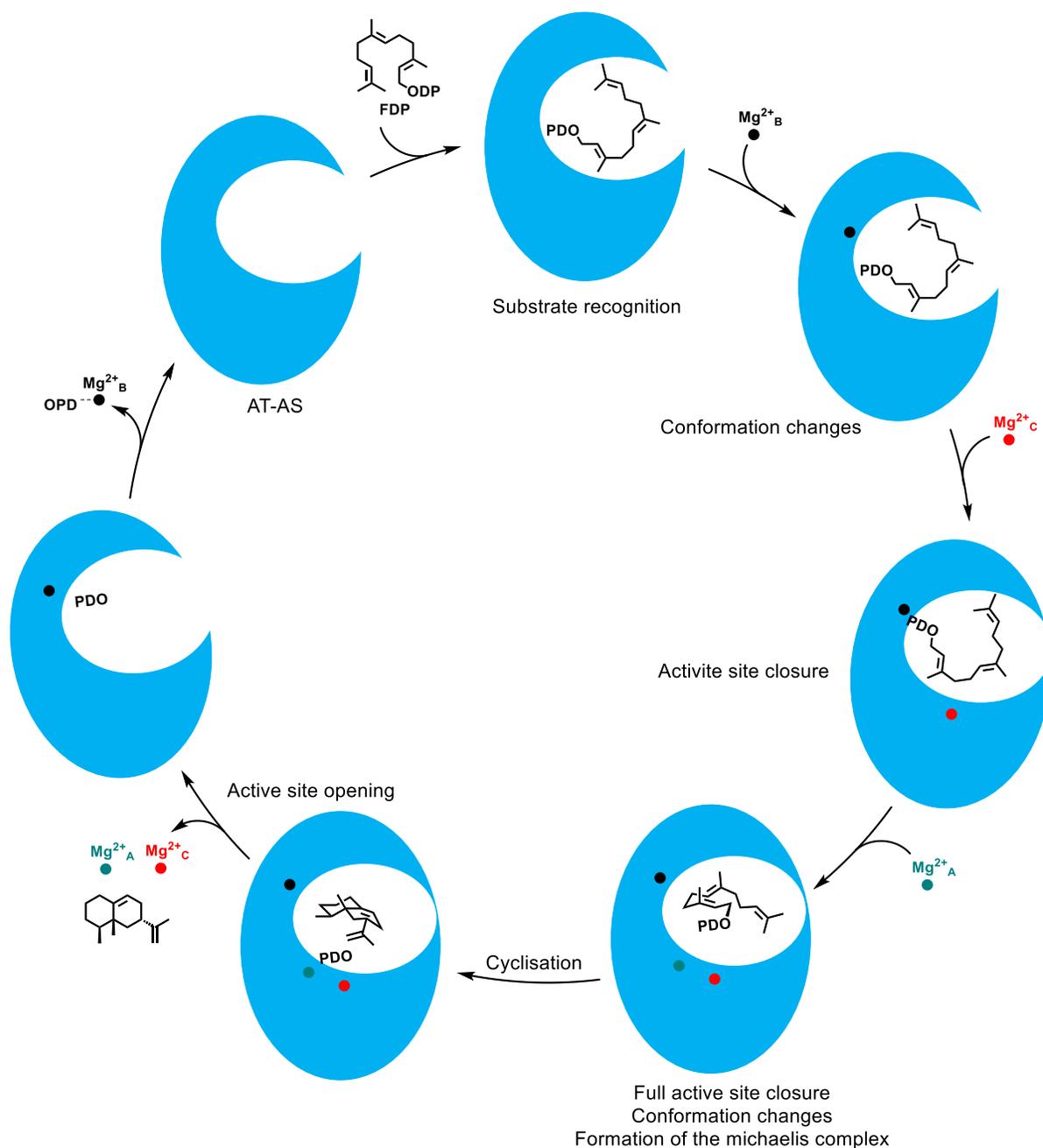


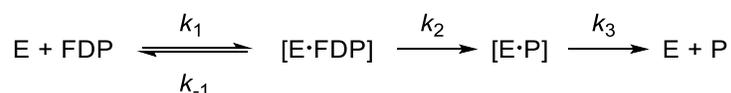
Figure 1-8 Model of the substrate, metal-binding and conformational changes in ATAS (adapted from Shishova *et al.* 2008).^[39]

First hydrogen bond donor residues such as Arg and Tyr lead the recognition of the diphosphate moiety of FDP (**14**), one magnesium Mg^{2+B} possibly binds at the same time or

subsequently to the NSE/DTE motif which triggers reposition of the substrate into the right orientation. A second Mg^{2+}_C binding to the DDXXD motif induces further conformational changes causing active site closure. A final third Mg^{2+}_A comes into DDXXD motif finalising the active site closure and formation of a Michaelis-complex. When the sesquiterpene product is formed, it is hypothesised that opening of the active site and release of the product occurs in the reverse sequence.^[39,40]

1.2.3 Kinetics of sesquiterpene synthases

Turnover numbers (k_{cat}) usually range between 1 to 1000 s^{-1} for most metabolic enzymes at their physiological pH.^[41] Sesquiterpene synthases are generally below average values making them slow enzymes. The average k_{cat} value of a sesquiterpene synthase is 2.5 s^{-1} , and the average value for an enzyme in central metabolism is 79 s^{-1} .^[42] Pre-steady state kinetics on several sesquiterpene synthases showed that the rate-limiting step of the catalysis is the product release, most likely due to the hydrophobic nature of the sesquiterpene product. Initially Mathis *et al.* performed pre-steady state kinetics using isotope traps ($[\text{}^3\text{H}]\text{-FDP}$, **45**) and rapid quenching of the reaction on vetispiradiene synthase from *Hyoscyamus muticus* (HVS) and two chimeric sesquiterpenes from *epi*-aristolochene and vetispiradiene.^[43] Results showed that enzyme-substrate complex formation and ionisation of FDP (**14**) was 10-70 times faster than the rest of chemical steps for all three enzymes. It was unclear if additional chemical steps to form the final product from the intermediate germacrene A or product release was the rate limiting step as experiments did not differentiate different sesquiterpene products.^[43] Additional studies from Cane *et al.* on trichodiene synthase from *Fusarium sporotrichioides* using rapid quenching showed that the release of the product was indeed the rate determining step in the catalytic cycle. The observed rate for the release of the product (0.086 s^{-1}), which is equivalent to the measured k_{cat} , proved to be 40 times slower than consumption of FDP.^[44]



Scheme 1-1 Simplistic model for the mechanism of biocatalysis of FDP by a sesquiterpene synthase to its product. E: enzyme, [E·FDP] enzyme FDP complex, [E·P] enzyme product complex, P: product.

1.2.4 Current production of sesquiterpenes and recent developments

Terpenes represent a 510 million US\$ worldwide market that is expected to grow steadily.^b Sesquiterpenes find their applications in fragrance, food industries, as building blocks for pharmaceutical synthesis but also increasingly as biofuel. Today, most of the known sesquiterpenoids are extracted from plants as the complexity of their structures is often a challenge for industrial scale production by chemical synthesis. Total synthesis involves multistep reactions often resulting in low yield, low enantiomeric purity and is overall an expensive process. Although significant advances have been made to solve those challenges,^[45–47] industries still rely on plant extraction. As we know, extraction from natural sources is limited, with highly variable quality and availability due to inherent problems from crop harvest. As an example, α -santalol and β -santalol priced at £74.10 per mL and widely used in perfumery for their woody scent are still currently obtained by distillation of heartwood of mature *Santalum album* trees (1 – 2% wt of oil),^[14] while chemical synthesis gives 8% overall yield. (-)-patchoulenol (**48**), the main component of the patchouli oil is still produced by distillation of *Pogostemon cablin* leaves for perfumery and fabric softener industries.^[14] Economic, environmental factors, and the search for a reliable supply chain have driven research to find more sustainable processes to access these high value compounds.

Production in vivo by engineered yeast

Recent advances in metabolic engineering have made possible the production of terpenes using large scale microbial fermentation of engineered organisms (Figure 1-9).^[14] *Escherichia coli* (*E. coli*) and *Saccharomyces cerevisiae* (*S. cerevisiae*) are the two most widely used hosts.^[48] Optimisation of sesquiterpenoid productivity was achieved through several mechanism such as increasing the expression of key enzymes in the MEV or DXP pathway by introducing additional promoters or restricting the gene encoding for squalene synthase, a competing reaction with STS using two FDPs to generate squalene (**15**).^[48] (1R,4R,4aS,8aR)-Amorpha-4,11-diene (**49**), precursor of artemisinin (**6**) was for example successfully produced at 40 g.L⁻¹ using engineered *S. cerevisiae* containing genes for the

^b https://www.theexpresswire.com/pressrelease/Terpenes-Market-2019-Globally-Market-Size-Analysis-Share-Research-Business-Growth-and-Forecast-to-2024-360-Research-Reports_10350212

overexpression of the mevalonate pathway enzymes and amorpha-4,11-diene synthase.^[49] *Trans*- β -farnesene (**50**), a linear sesquiterpene used in a wide range of industries such as polymers, lubricants and crop protection was very recently brought to industrial production by Amyris also using *S. cerevisiae* engineered organism.^[14]

Evova and Isobionics have also commercialised valencene (up to 80% purity), nootkatone (**51**) (98%), β -elemene (**52**) (98%), β -bisabolene (**53**) (90%), δ -cadinene (**54**) (95%), and germacrene D (**55**) (85%) from yeast production. However, significant efforts are still needed towards higher yield, purity and lower cost of production using the current *in vivo* methods. This method does also not allow the production of unnatural sesquiterpenes yet.

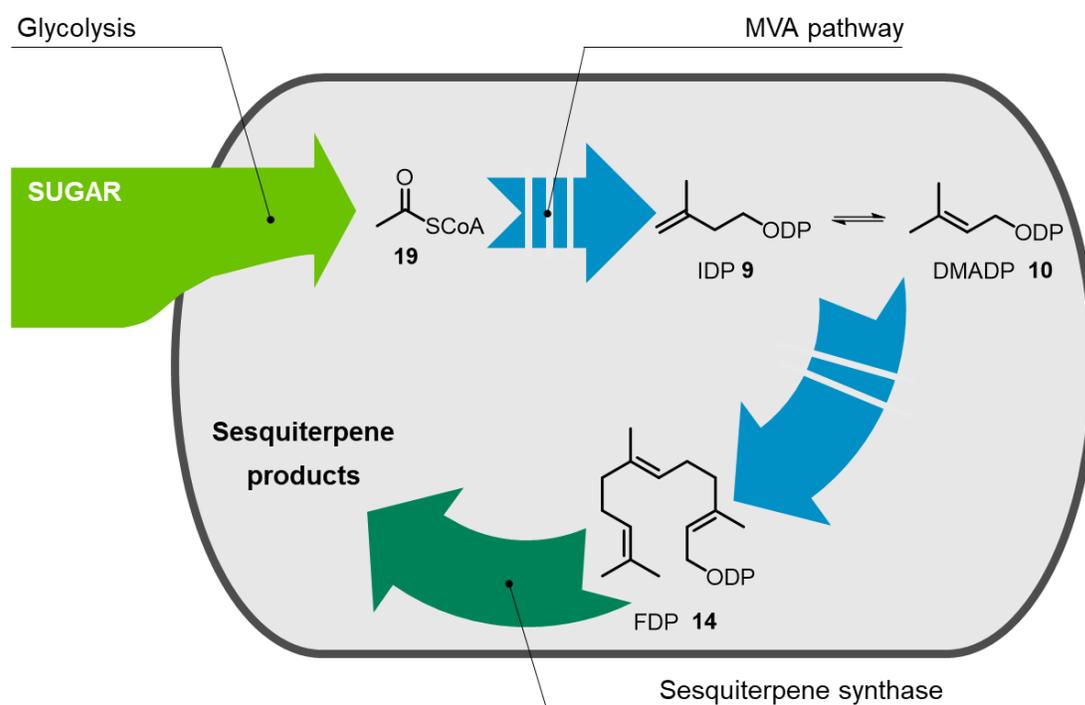


Figure 1-9 Scheme for the biosynthesis of sesquiterpene through MVA pathway in a sugar fed engineered host.^[14]

Production *in vitro*

In vivo production of terpenes has been a continuous challenge as careful design and balance of the biosynthetic pathways is needed for high fermentation efficiency. Cell free alternatives, *in vitro* solutions, have been developed recently as a potential answer to this issue. Whole biosynthetic pathways coupled with cofactor regeneration for sesquiterpenes have been

reconstituted and enabled the optimisation of each step (e.g. adjusting equivalents of substrates or enzyme to avoid deleterious accumulation of intermediates and side reactions). For example, an enzymatic synthesis of patchoulol has been reported from acetic acid, and (1R,4R,4aS,8aR)-amorpha-4,11-diene (**49**) from (*R*)-mevalonic acid (**22**) (MVA) or isoprenol (**56**) and prenol (**57**).^[50–52]

Supramolecular chemistry – an artificial sesquiterpene cyclase

Tiefenbacher *et al.* have reported the first synthesis of multiple sesquiterpenes using a supramolecular catalyst mimicking a sesquiterpene synthase active site.^[53–56] Using a hexameric resorcinarene capsule developed by Atwood *et al.* (Figure 1-10),^[57] isolongifolene (**58**) has been successfully produced from (2*Z*)-cyclofarnesyl acetate (**59**) in 11% yield, δ -selinene (**60**) in 18% yield from 2*E*,6*Z*-farnesyl acetate (**61**). They published the shortest synthesis of the tricyclic sesquiterpene presilphiperfolan-1 β -ol (**62**) to date from commercially available (-)-caryophellene oxide (**63**) (Figure 1-11).^[58]

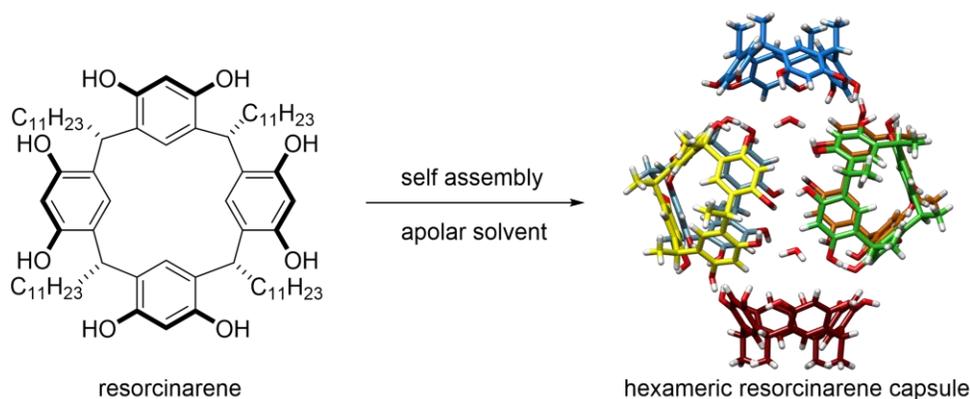


Figure 1-10 Resorcinarene chemical structure and structure of the hexameric resorcinarene capsule (modified from CIF CCDC 1590340).

Kinetic studies showed that encapsulation of the starting material is the rate limiting step.^[56] Interestingly, only 2*E* isomers of farnesol or farnesyl acetate afforded 1,10 cyclisation giving δ -selinene (**60**). Coupled together with 6*Z* bond conformation, higher product specificity was observed toward δ -selinene (**60**) and only four other side products were generated (Figure 1-11a).^[56] A better selectivity for isolongifolene (**58**) was only achieved when using an already conformationally restricted substrate, showing the capsule is not yet close to mimicking the

intricate three-dimensional active site of terpene synthases leading to precise substrate folding and intermediate quenching.^[55] Nevertheless, this promising approach offers a new “one pot” synthesis of high value compounds with potential to create novel analogues that the engineered host cannot yet achieve.

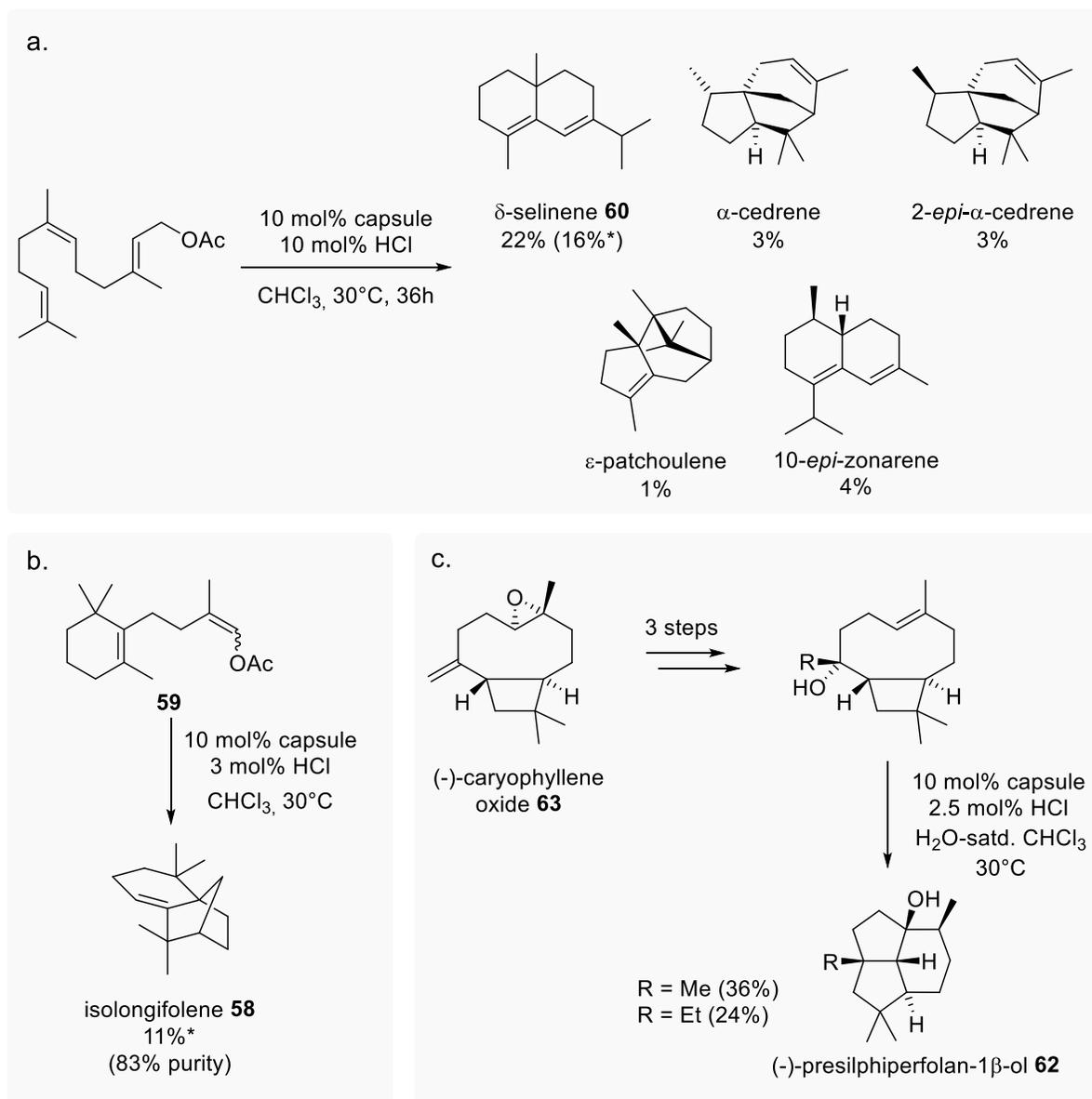
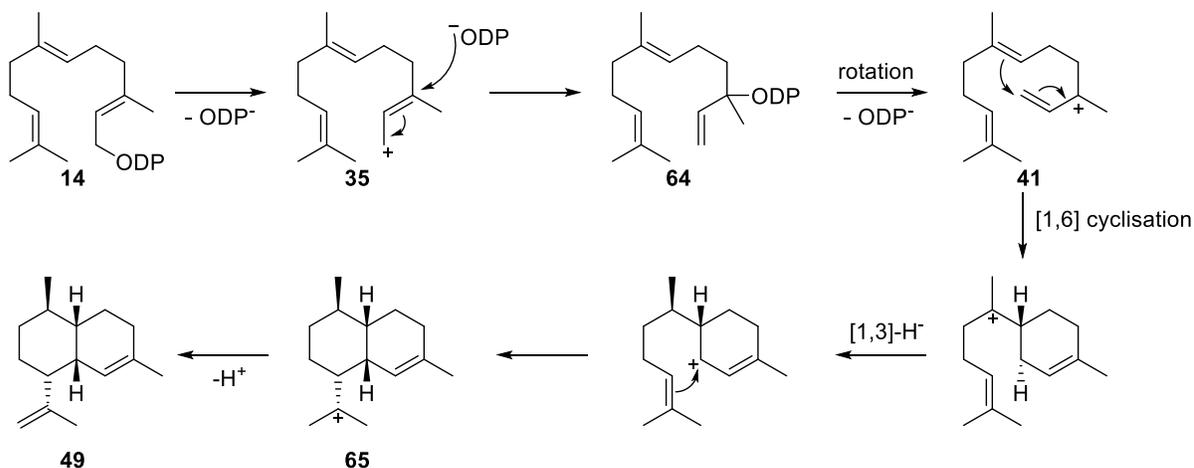


Figure 1-11 Products formed from the resorcinarene capsule. *Isolated yield.^[53–56,58]

1.3 SESQUITERPENE SYNTHASES USED IN THIS STUDY

1.3.1 Amorpha-4,11-diene synthase

Amorpha-4,11-diene (amorphadiene) synthase (ADS) from *Artemisia annua* catalyses the formation of (1R,4R,4aS,8aR)-amorphadi-4,11-diene (**49**), a key intermediate in the biosynthesis of artemisinin. The mechanism was investigated and fully established using deuterated FDPs coupled with GC-MS and NMR spectroscopy. Initial formation of a farnesyl cation (**35**) is followed by isomerisation to nerolidyl diphosphate (**64**) (NDP). Then a 1,6 cyclisation occurs to form bisabolyl cation (**41**) followed successively by a stereoselective [1,3]-hybride shift and a 1,10 cyclisation to give a bicyclic amorphyl cation (**65**). Final deprotonation from C12 or C13 generates (1R,4R,4aS,8aR)-amorpha-4,11-diene (**49**).^[59–62]



Scheme 1-2 Putative catalytic mechanism of amorpha-4,11-diene synthase.

Brodelius and co-workers suggested that the mechanism goes *via* the nerolidol cation (**38**) and a [1,6] cyclisation due to the presence of minor bisabolene side product in the incubation of FDP (**14**) with ADS. Moreover, ADS can catalyse the conversion of GDP (**11**) to form cyclic products such as (*R*)-limonene (**13**), α -terpinol (**66**) and terpen-4-ol (**67**) supporting an initial [1,6] cyclisation. GC-MS analysis of incubations of (1*S*)-[1-²H]-FDP (**68**), (1*R*)-[1-²H]-FDP (**69**) and [1,1-²H₂]-FDP (**70**) (double deuterated C1 FDP) with ADS showed that the pro-*R*-H-1 remains at the C1 position while pro-*S*-H-1 has migrated to the C7. Kim and co-workers studied incubation products of (1*S*)-[1-²H]-FDP (**68**), (1*R*)-[1-²H]-FDP (**69**) and [1,1-²H₂]-FDP (**70**) with ADS by ¹H and ²H NMR spectroscopy. Double deuterated FDP (**70**) formed

(1R,4R,4aS,8aR)-amorpha-4,11-diene with 2 signals on the ^2H NMR spectra at $\delta_{\text{H}} = 1.38$ and 2.60 while the ^1H NMR spectra showed no signals at $\delta_{\text{H}} = 1.40$ and 2.55 corresponding to H7 and H1 respectively (Figure 1-12). On the other hand, the product arising from (1R)-[1- ^2H]-FDP (**69**) only showed a single signal at $\delta_{\text{H}} = 2.60$ while from (1S)-[1- ^2H]-FDP (**68**) gave a signal at $\delta_{\text{H}} = 1.38$. These spectra showed a [1,6] cyclisation with pro-S-H-1 transferring to C7, pro-R-H-1 remaining at the C1, confirming the previous GC-MS findings (Figure 1-12).^[60]

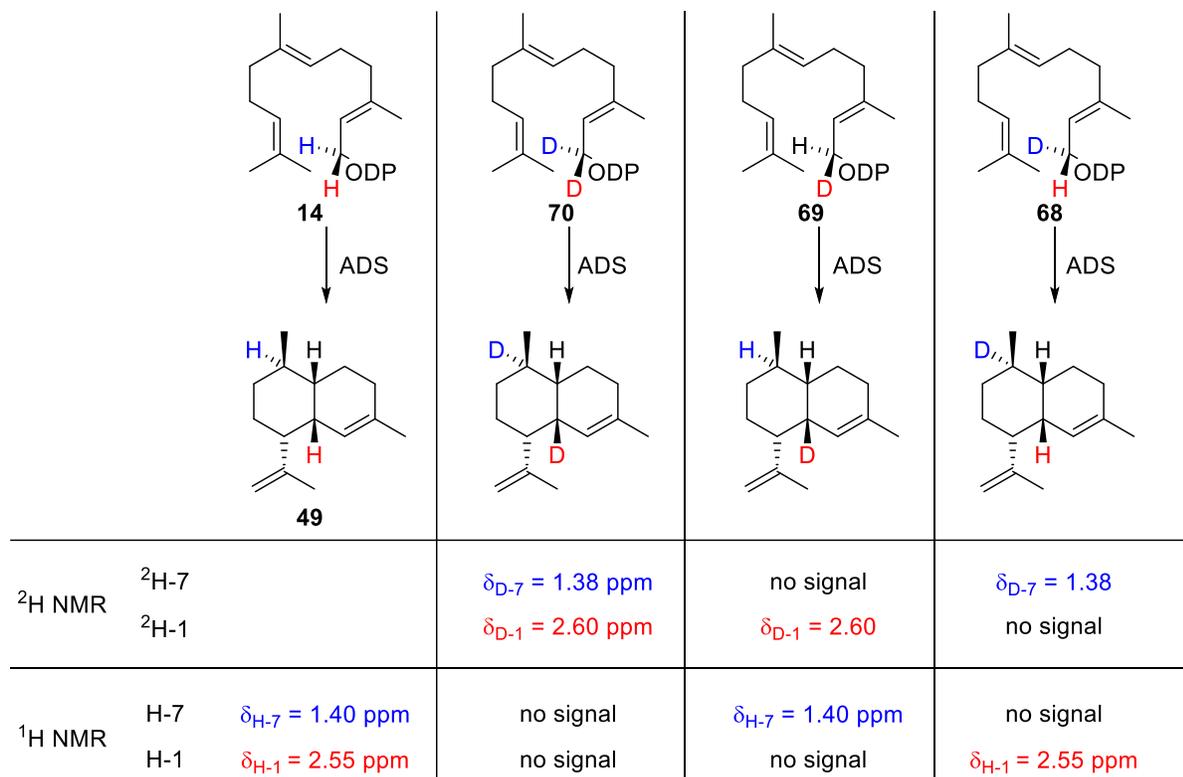
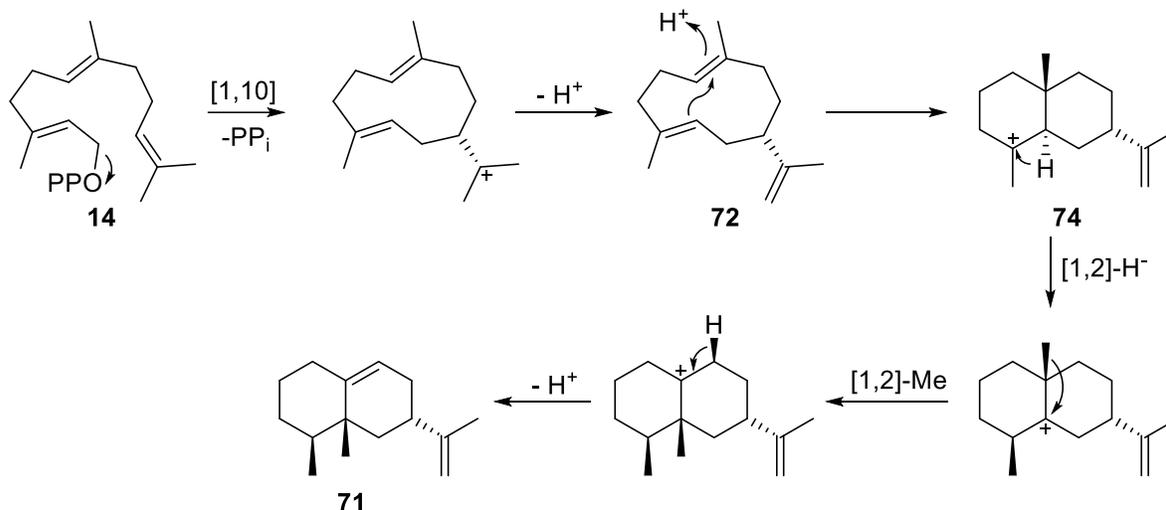


Figure 1-12 Incorporation of deuterium into (1R,4R,4aS,8aR)-amorpha-4,11-diene from deuterated FDPs identified by NMR spectroscopy.

1.3.2 Aristolochene synthase

Aristolochene synthase from *Penicillin roquefortii* (PRAS) catalyses the formation of the main product (1R,7R,8aS)-aristolochene (**71**) (92 %) as well as (S)-germacrene A (**72**) (7.5%) and valencene (**73**) (< 0.5%) from FDP (**14**). The enzyme has been extensively studied from this bacterial organism as well as from *Aspergillus terreus* (ATAS). The mechanism has been

demonstrated by site directed mutagenesis (SDM) and incubation with fluorinated FDP analogues. It was shown to go through a concerted [1,10] cyclisation to form (*S*)-germacrene A (**72**). Additional investigations using aza-analogues and SDM showed a subsequent protonation of the 6,7 double bond of germacrene A leading to a second cyclisation to give a bicyclic eudesmane cation (**74**). Then [1,2] hydride and methyl shifts occur and final deprotonation of H₅ on C8 gives (1*R*,7*R*,8*aS*)-aristolochene (**71**).^[63–66]



Scheme 1-3 Putative catalytic mechanism of PRAS.

It was first hypothesised that the mechanism goes through a germacrene A (**72**) intermediate from its high proportion of side products. They first showed that the Y29F PRAS mutant catalysed a higher proportion of germacrene A (**72**) compared to the wild type.^[67] To ascertain this hypothesis, fluorine analogues of FDP that would only lead to germacrene A products due to destabilisation of further carbocation formation after germacrene A were designed. (2*Z*, 6*E*)-6F-FDP (6F-FDP, **75**), 14-F-FDP (**76**) and 2-F-FDP (**77**) were incubated with PRAS and all generated corresponding fluorogermacrene A analogues, proving the role of germacrene A as an intermediate in the mechanism of PRAS (see Figure 1-13).^[68,69] Formation of 2-fluorogermacrene A (**78**) also indicated that [1,10] cyclisation and leaving of the diphosphate group is a concerted mechanism because if those two steps were to occur stepwise, the allylic farnesyl cation would be destabilised by the β -position of the fluorine in 2-F-FDP (**77**) (see Figure 1-13).^[70] This finding was further supported by verifying that [12,13]-F₂-FDP (**78**) acted as a competitive inhibitor of PRAS with a K_i comparable to the reported K_m ($K_i = 0.8 \mu\text{M}$, $K_m = 2.3 \mu\text{M}$).^[71] In a concerted mechanism, a carbocation would be formed on C11 however with

two β -fluorines the latter cannot develop resulting in competitive inhibition (see Figure 1-13). On the other hand, a stepwise mechanism would lead to an allylic farnesyl cation (**79**) which would be expected to end the carbocationic cascade by a final deprotonation to form farnesenes (**80** and **81**) as a ring closure is not favourable due to a formation of a carbocation with two β -fluorine atoms (see Figure 1-13).^[71]

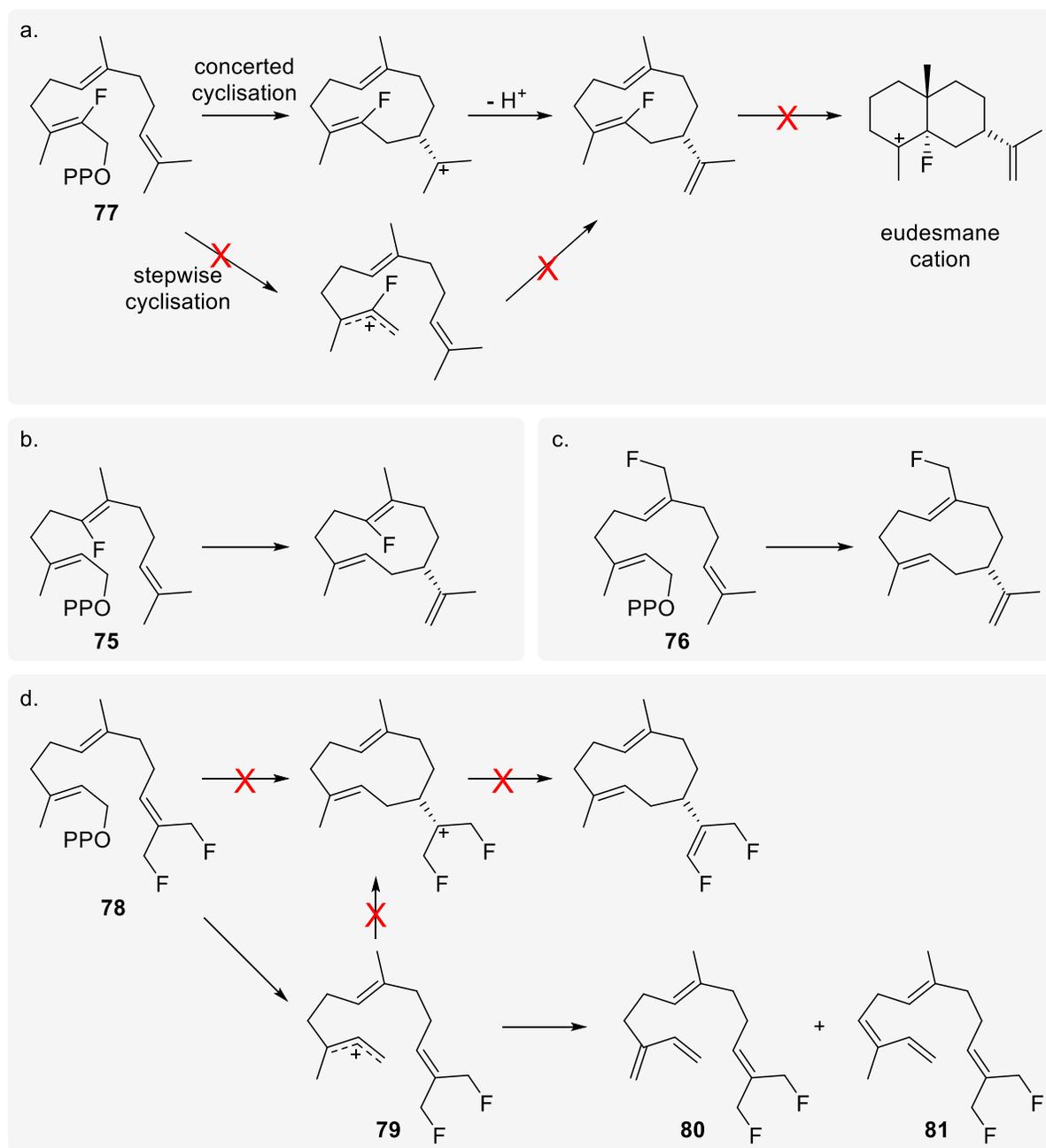
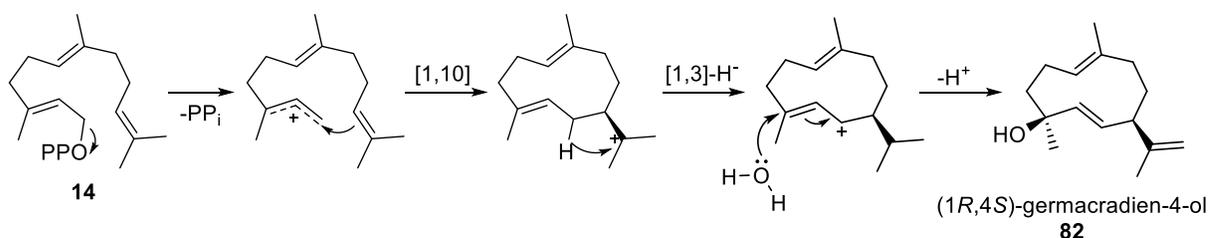


Figure 1-13 Cyclisation outcome of fluorinated FDPs by PRAS.

1.3.3 Germacradiene-4-ol synthase

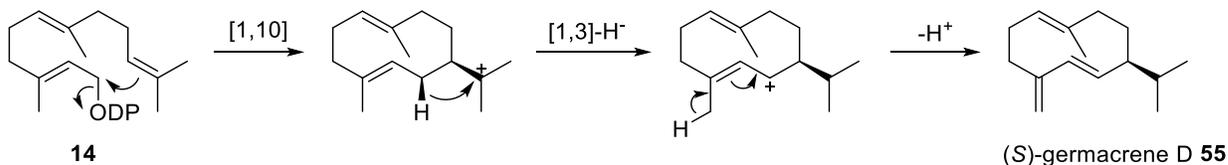
Germacradien-4-ol synthase (Gdols) from *Streptomyces citricolor* catalyses the conversion of FDP (**14**) to (1*R*,4*S*)-germacradien-4-ol (**82**) as the sole product. Grundy *et al.* demonstrated that 2-F-FDP (**77**) and 15-CF₃-FDP were inhibitors of Gdols while incubation of 12,13-F₂-FDP (**78**) with Gdols gave 12,13-difluorofarnesene (**80**) suggesting [1,10] ring closure to the germacryl cation is a stepwise mechanism.^[72] Incubation of 1-[²H₂]-FDP (**70**) with Gdols also showed a [1,3] hydride shift from following the cyclisation (see Scheme 1-4). Interestingly this hydride shift is not stereoredefined as incubation with 1-(*R*)-[²H₁]-FDP (**69**) gave products with fragments of *m/z* = 161 and *m/z* = 162. Alternatively, stereochemical information is lost through equilibration of the allylic cation. Final quenching is proposed to originate from bulk water in the solvent upon the product release step.^[72]



Scheme 1-4 Putative catalytic mechanism of Gdols.

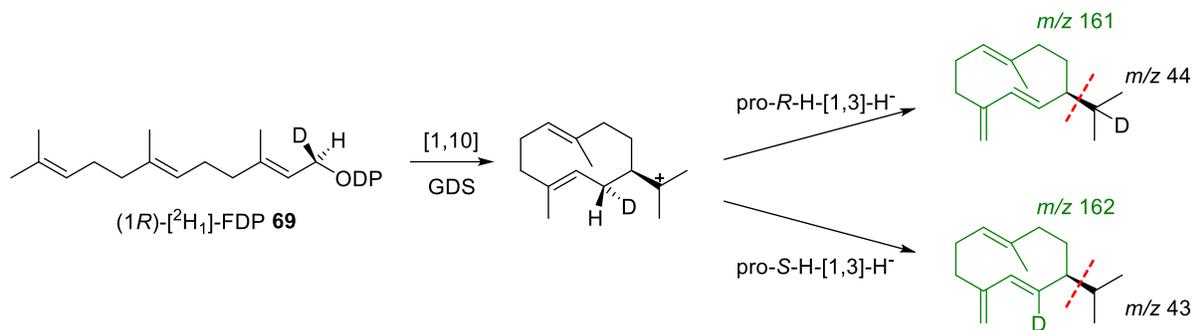
1.3.4 Germacrene D synthase

(*S*)-Germacrene D synthase (GDS) from *Solidago canadensis* catalyses the production of (*S*)-germacrene D (**55**) as the sole product, which acts as a repellent against aphids. It has been previously isolated and fully characterised with a $k_{\text{cat}} = 0.03 \text{ s}^{-1}$ and $K_{\text{M}} = 2.0 \pm 0.3 \mu\text{M}$. The mechanism was proposed to proceed through a [1,10] ring closure followed by a 1,3-hydride shift at a final deprotonation at C15 yield (*S*)-germacrene (**55**).^[73,74]



Scheme 1-5 Putative catalytic mechanism of (*S*)-GDS.

The mechanism was heavily studied using deuterated FDPs coupled with GC-MS. Incubation of (1*S*)-[1-²H]-FDP (**68**) and (1*R*)-[1-²H]-FDP (**69**) gave products with a molecular ion [M⁺] m/z = 205, revealing the presence of one deuterium in the final compound. The most abundant fragment ion in the incubation with (1*R*)-[1-²H]-FDP (**69**) was at m/z = 161 while (1*S*)-[1-²H]-FDP (**68**) gave m/z = 162. This fragment arises from the loss of the isopropyl group from M⁺. This result demonstrated that the pro-*R*-H-1 of FDP performs the 1,3-hydride shift.^[74]



Scheme 1-6 Fragmentation of (*S*)-germacrene D from the incubation with (1*R*)-[1-²H]-FDP.

1.4 APPROACHES TO NOVEL TERPENOIDS

Natural terpenoids and sesquiterpenes show a wide range of bioactivities. However, the final terpene product is usually difficult to functionalise or modify to create analogues in the search for better potency. A striking example is the antimalaria drug: artemisinin. The search for more potent and bioavailable compounds is difficult due to the structure being mostly hydrocarbon based. As a matter of fact, most analogues come from modification of the sole lactone group.^[75] Research has demonstrated that harnessing the synthetic potential of sesquiterpene synthases instead to produce novel or analogue compounds is possible. This enzymatic pathway is more appealing because cyclases are readily capable of exquisitely controlling stereo-, chemo- and regio- selectivity, whereas chemical synthesis is costly, cumbersome and will likely not achieve the same degree of selectivity. Over the past decade, deeper understanding has enabled exploitation of the promiscuity of sesquiterpene synthases. Enzyme engineering for unnatural substrate acceptance, control and modification of their product profiles enabled easy access to analogues of known terpenoids or completely novel compounds. This coupled with existing robust strategies to express these proteins make this method an appealing strategy to produce new high value sesquiterpenes compared to existing methods.

1.4.1 Unnatural substrates

The discovery of novel sesquiterpenes has been reported numerous in the literature through selectively feeding unnatural substrates to wild type enzymes or production as by-products from mechanistic investigations.^[76] It was shown by Allemann and co-workers that alterations (mainly additional methyl groups or fluorine) in the FDP structure are possible and lead to the production of analogues of the naturally occurring sesquiterpene. For example, a range of analogues were synthesised using (S)-GDS and germacrene A synthase (GAS) and tested for behavioural activity of aphids, a major world crop pest. Notably, FDP analogues containing an extra methyl group on C15 were not readily accepted by the natural enzyme and mutation of the active site was performed to facilitate substrate acceptance. (S)-12-Methylgermacrene D (**84**) and (S)-14-methylgermacrene D (**85**) were found to be repellent against aphids while (S)-14,15-dimethylgermacrene D (**86**) was found to have a high attractant semiochemical activity towards aphids (see Figure 1-14).^[77] Further studies are ongoing to commercialise this novel compound. Interestingly, feeding ADS with analogues of FDP with additional methoxy groups led to a structurally distinct hydrocarbon skeleton compared to the natural occurring product (see Figure 1-14). Instead of a [1,6] cyclisation through an isomerisation to NDP, 8-methoxy FDP (8-OMe FDP, **87**) resulted in a [1,11] cyclisation product (**88**).^[78] In contrast, 12-methoxy FDP (**89**) successfully ionised and isomerised but only went through the first [1,6] cyclisation giving two products with sesquiphellandrene (**90**) and zingiberene skeletons (**91**) (see Figure 1-14).^[78] These results from Demiray *et al.* suggested that this inductively withdrawing functional group with a π -acid functionality modify the outcome of carbocationic species. This demonstrates the limitations of the chemoenzymatic pathway for the synthesis of sesquiterpene analogues guided by corresponding non-natural FDP analogues. However, this has allowed generation of novel terpenoids.

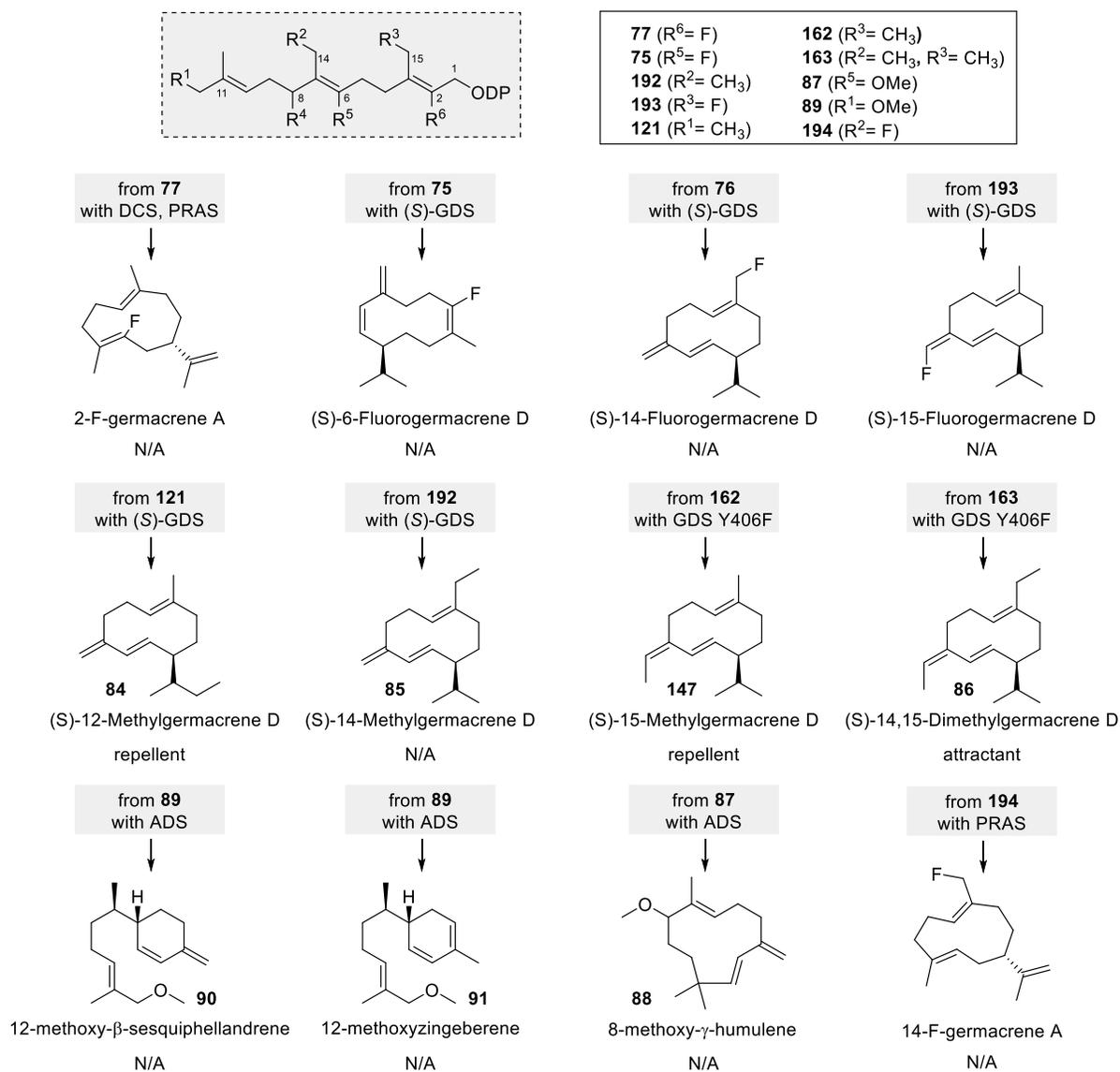


Figure 1-14 Examples of analogues of FDP turned over by sesquiterpene synthases and their products published by Allemann R. K. and co-workers.

Hou *et al.* also reported a C6 methylated FDP (**92**) being accepted by T-muurolool synthase from *Roseiflexus castenholzii* (TmS).^[79] Depending on the stereochemistry of the additional methyl group, products generated by TmS were different. The *R* isomer did not only produced the methylated analogue natural product (**93**), but also two other products (**94**, **95**) including one (**94**) where the normal second cyclisation to the two 6-membered ring did not happen (see Figure 1-15b).^[79]

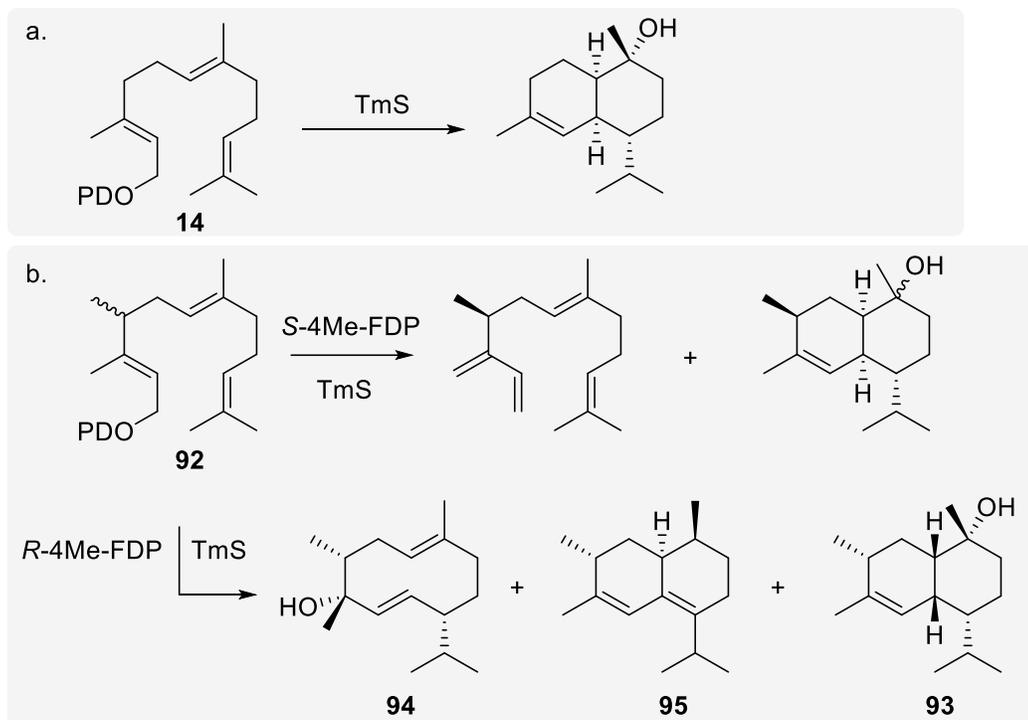


Figure 1-15 a. Product generated by TmS from FDP b. C6 methylated analogues of FDP turned over by TmS published by Dickschat J. S. and co-workers.

An unconventional approach was taken by Harms *et al.*, where novel sesquiterpenes were generated from FDP with methyl groups shifted by one position toward the diphosphate moiety from presilphiperfolan-8 β -ol synthase (Bot2) from *Botrytis cinerea*.^[80] Not only were these atypical FDP analogues substrates to Bot2, they apparently also modified Bot2's cyclisation pathway. When C11 methyl was shifted to C10 (**96**), Bot2 generated a humulene type structure (**97**) instead of the tricyclic presilphiperfolan-8 β -ol (**98**). Similar 11-membered ring structures (**99**, **100**) were produced when the three methyl group were shifted (**101**) (C3 to C2, C7 to C6 and C11 to C10) (see Figure 1-16). Surprisingly, when only two methyl group were shifted (C3 to C2 and C7 to C6) (**102**), a 10-membered ring sesquiterpene (**103**) similar to germacrene A was formed (see Figure 1-16).^[80]

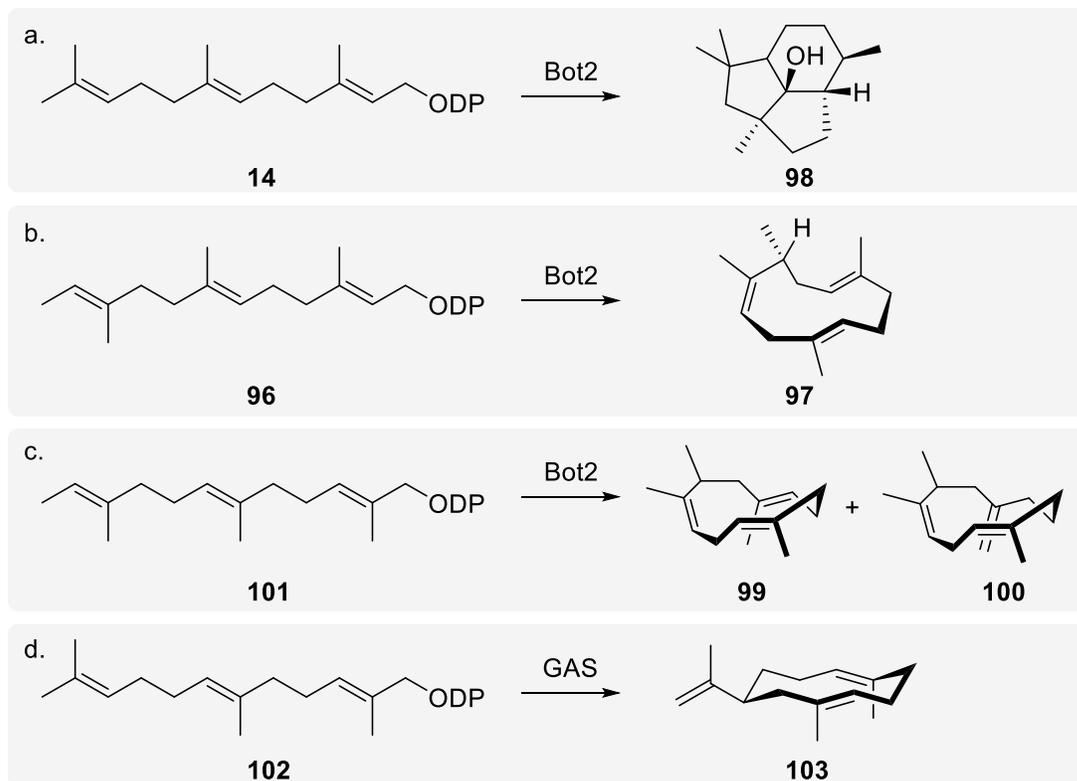


Figure 1-16 a. Product generated by Bot2 from FDP. b-d. examples of methyl-shifted analogues of FDP turned over by Bot2 published by Kirschning A. and co-workers.

Kirschning and co-workers introduced heteroatoms into the FDP structure and tested them on the following sesquiterpene synthases: [1,11] cyclisation synthase: presilphiperfolan-8-β-ol synthase (Bot2) from *Botrytis cinerea*, caryolan-1ol synthase (GcoA) from *Streptomyces griseus* cyclisation, and a [1,10] cyclisation synthases: viridiflorene synthase (Tps32) from *Solanum lycopersicum*, cubebol synthase (Cop4) from *Coprinus cinereus*, (+)-T-muurolol synthase (TmS) from *Roseiflexus castenholzii*.^[81] While nitrogen containing analogues did not give any products, oxygen and sulphur functionalised analogues (thioethers and ethers) were generally turned over by the cyclases yielding heteroatom-containing medium cycles. Interestingly, (X=O, m=1, n=2) with [1,10] or [1,11] cyclisation type cyclases (such as Tps32, GcoA, Cop4, GcoA) were all able to produce the macrocyclic ether (104). Similar behaviour (product 105) was also observed with the sulphur containing analogue.^[81] Substrate (X=S, m=2, n=1) was converted to structure (106) by again both [1,10] and [1,11] cyclisation type cyclases such as Pts, Tps32, PenA, Cop4, Bot2 or TmS via [1,10] cyclisation. But the oxygen analogue (X=O, m=2, n=1) was only turned over to the corresponding oxygen analogue (107) by TmS, all other enzymes produced the linear product (108) instead. The most notable product is the tricyclic terpene formed (109) from (X=O, m=1, n=2) by Bot2. It is hypothesised

that the oxygen stabilises one of the carbocation intermediates, allowing a different mechanism to occur.^[81]

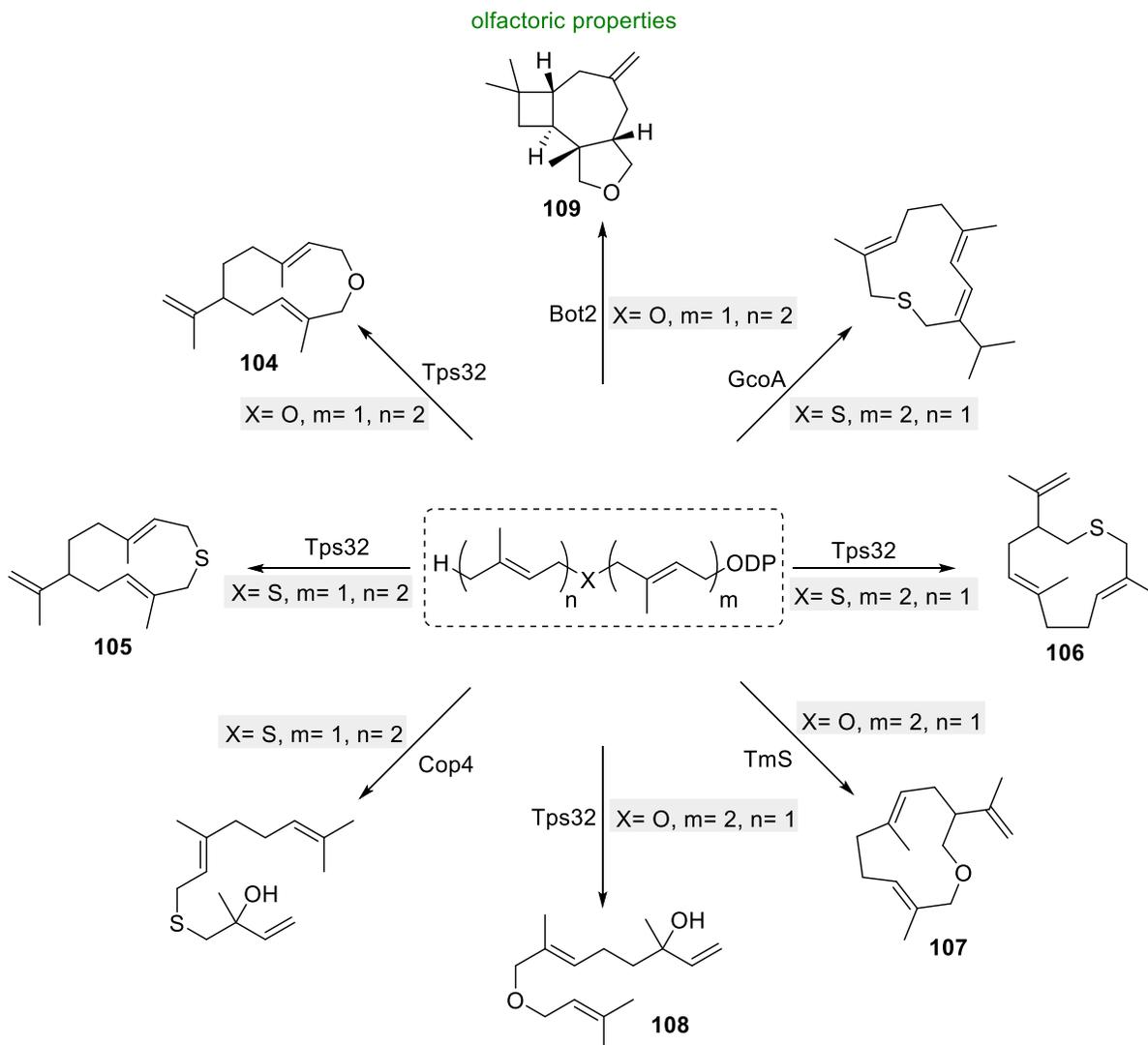


Figure 1-17 Examples of heteroatom containing analogues of FDP from Kirschning and co-workers turned over by sesquiterpene synthases and their product.

Sesquiterpenes with carbonyl functional groups were also successfully produced using non-natural FDPs. Allemann and co-workers reported the first synthesis of dihydroartemisinic aldehyde (DHAAI) by amorpho-4,11-diene synthase from several oxygenated FDPs.^[82] 12-Hydroxy FDP (12-OH FDP, **111**) gave a mixture of isomers with a ratio of 3:2 (11S : 11R) (**112** and **113**) and one minor product identified as the enol form of 11S (**114**) (Figure 1-18).^[82] Resulting DHAAI was later successfully converted to artemisinin (**6**) in four chemical steps making this chemoenzymatic route the most concise synthesis of this anti-malaria drug known to date.^[83]

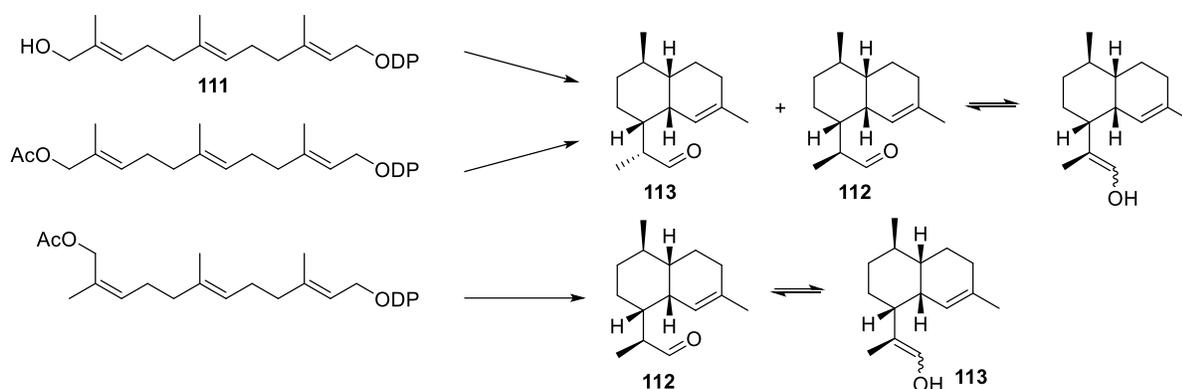


Figure 1-18 Carbonyl containing sesquiterpenes from oxygenated analogues of FDP with ADS.

Arnold and co-workers used a vinyl methyl ether FDP analogue (**115**) as substrate for BcBOT2 and SSCG_02150 synthases that naturally produce presilphiperfolan-8 β -ol (PSP, **98**) and (-)- δ -cadinene (**54**) from FDP (**14**) respectively.^[84] Both enzymes converted the non-natural FDP (**115**) to the aldehyde (**116**). It is believed to go through a [1,10] cyclisation, followed by water quenching of the resulting carbocation to generate a hemiacetal. A final release of methanol produces the aldehyde. By taking advantage of the fact that methanol is generated as side product, they successfully coupled the reaction to a colorimetric assay involving alcohol oxidase (AOX) converting methanol to formaldehyde and Purpald which turns purple in presence of an aldehyde. This was used with directed evolution to create a more thermostable synthase. BCBOT2 T_{50} was increased by 12 °C while retaining comparable enzyme activity and expression level (Figure 1-19).^[84]

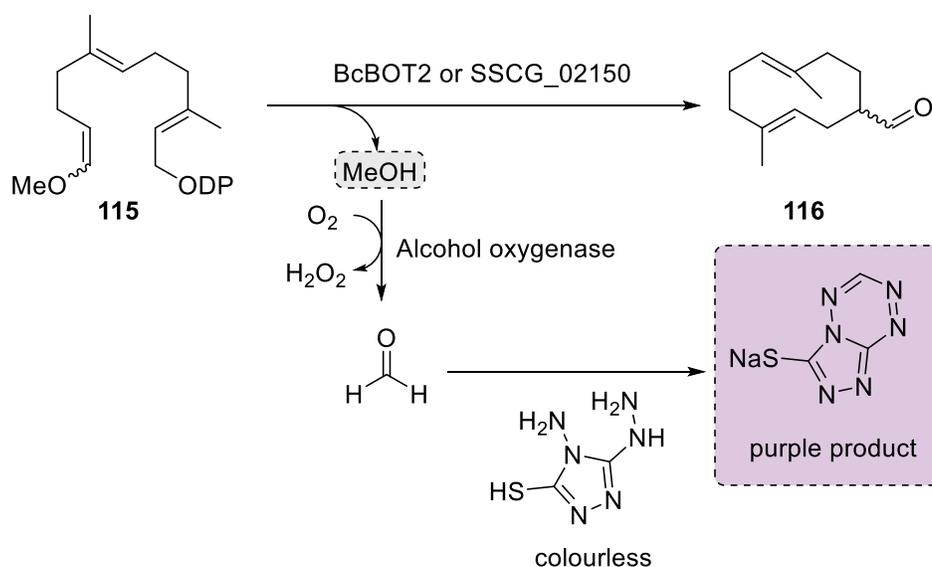
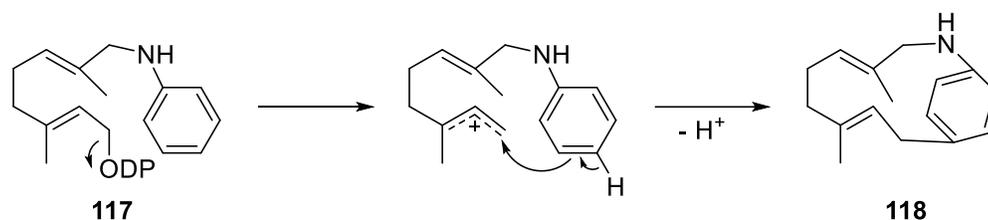


Figure 1-19 Synthesis of aldehyde (**116**) from vinyl ether containing analogue of FDP coupled with assay for methanol detection for high-throughput screening of suitable mutant.

Notably, only one example of an alkaloid synthesised by a sesquiterpene cyclase has been demonstrated to date. An aza-analogue of GDP, 8-anilino GDP (AGDP) (**117**) was originally designed to bind TEAS active site for crystallographic studies. However, upon studying the crystal structure of TEAS – AGDP complex Rising *et al.* discovered AGDP was a substrate for TEAS and underwent cyclisation generating a novel 13-membered paracyclophane alkaloid (**118**).^[85]



Scheme 1-7 Proposed mechanism for the conversion of 8-anilino GDP (**117**) by TEAS.

1.4.2 Engineering sesquiterpene synthases

Discoveries of new sesquiterpene products do not only arise from feeding non-natural substrates to cyclases. Researchers have also shown how enzymes can be engineered to form new products from natural FDP as well as turning promiscuous cyclases into high fidelity enzymes in a predictable manner.

Rational design and random mutagenesis have been used to convert a δ -cadinene synthase (DCS) into a functional GdoIS. Work from Yoshikuni and co-workers used directed evolution to search for potentially improved variants. A high throughput screening activity method was created with a chloramphenicol acetyltransferase (CAT) fused to the mutant DCS gene. Higher activity of CAT was correlated to high solubility of the protein variant because it was shown that the C-terminal protein is dependent on the solubility of the N-terminal fused protein.^[86] This fast pre-screening method was however not able to evaluate activity of the cyclase.^[86] Through this screening, two mutants N403P and L405H were identified that produce a higher proportion of (1*R*,4*S*)-germacradien-4-ol (gd4ol, **82**) with around 52% selectivity. L405H showed improved selectivity up to 93% when the incubation with FDP (**14**) was not overlaid with an organic solvent; this phenomenon is explained by faster evaporation of δ -cadinene (**54**) over gd4ol (**82**).^[86] Later Allemann and co-workers successfully converted DCS to a high fidelity GdoIS (up to 90% selectivity) by a single mutation using careful rational design and previous work reported by Yoshikuni. They identified the residue W279 to be responsible in shielding the active site from bulk water.^[87] Mutation to alanine reduced

hydrophobic contacts with the reactive carbocation intermediate allowing water quenching to form almost exclusively germacradien-4-ol corrected to (1*R*,4*S*)-germacradien-4-ol (**82**).

In 2006, Yoshikuni *et al.* also engineered the highly promiscuous γ -humulene synthase to several high fidelity sesquiterpene synthases.^[32] Initial side reactions were enhanced to achieve high selectivity through directed evolution. Mutation of 19 residues around the active site revealed four residues were responsible of substrate selectivity. Careful rational combination of mutations resulted in modified product profiles with higher selectivity toward one product. This method was successfully translated to multiple other cyclases demonstrating the power of enzyme engineering to create or modify terpenoid skeletons.

The entire active site of epi-isozizaene synthase was also remodelled by single point mutations turning the promiscuous cyclase into nine different high fidelity sesquiterpene synthases: (*E*)- β -farnesene synthase with F96A variant, β -curcumene synthase with F95H variant, sesquisabinene A synthase with F96S, F96M or F96Q, β -acoradiene synthase with F95M, β -cedrene synthase with F198L, zizaene synthase with F96W, (*Z*)- γ -bisabolene synthases with F96V and W203F variant.^[88]

Many more examples have been reported in the literature showing the potential to predict and direct the product outcome of a sesquiterpene synthase. New sesquiterpenes remain to be discovered and developed through protein engineering. Together with the use of non-natural FDP analogues, a new range of sesquiterpenes with potential bioactivities are to be discovered.

1.4.3 Limitations

The main limitations of this chemoenzymatic method is the production of FDP or its analogues. Currently FDP (**14**) can be potentially synthesised *in vivo* (see section 1.2.4) but analogues of FDP still need to be chemically synthesised. This process is presently inefficient and costly due to the last step of the synthesis where a diphosphate group needs to substitute the alcohol functional group.^[89] The incubation of FDP and its analogues with STS-s also require further optimisation due to their low yield. Current developments are presented in the section below.

1.5 BIOCATALYTIC OPTIMISATION OF SESQUITERPENE SYNTHASES

The discovery of a great number of high value-terpenoids has been reported though the conversion of analogues of FDP by recombinant terpene synthases. This chemoenzymatic

pathway is a convenient platform compared to the cumbersome and costly chemical synthesis to discover and produce, new compounds with potential bioactivities. The process however suffers from several weak points to be reasonably upscaled for research or commercial use. An efficient and up-scalable route to these compounds is essential for its use. Extensive efforts in the past have been therefore devoted to the optimisation of the biocatalysis of FDP by sesquiterpene synthases.

1.5.1 Batch process

The most widely used procedure for the catalysis of sesquiterpene synthases *in vitro* in batch uses a biphasic mixture comprising of an aqueous phase containing the enzyme, cofactor (Mg^{2+}) and substrate (FDP or analogues) in incubation buffer and an organic phase (usually hexane or pentane). Various enzyme and substrate concentrations are used across the literature as well as different methods (reaction vessels and mixing time) and quantity of solvent (ratio between aqueous/organic phase). This biphasic system has been rapidly introduced because the use of a single solvent (water) led to slow catalytic activity due to hydrophobicity of the product formed. Indeed, as described in section 1.2.3, pre steady state kinetics have shown the rate limiting step of sesquiterpene synthases is the release of the product. Introducing an organic solvent to extract the hydrophobic product prevents the enzymatic reaction from stalling. Reported yields in the literature range from 1% to 70% with isolated yields only given when product characterisation is needed.^[70,78,81,90] This procedure is therefore not predictable and not reliable for the production of existing and new sesquiterpenes in a cheap and efficient way. In the race to solve this issue, Brodelius and co-workers produced epi-aristolochene from GDP and IDP using a fused FDPS and *epi*-aristolochene synthase (*epi*AS).^[91] They observed great improvement of product formation from an equivalent amount of FDP using the fused enzymes compared to the two enzymes in one pot. This might be explained by improved substrate channelling due to the enzyme's proximity. However, they only conducted the reaction on a pmol scale.

1.5.2 Flow systems

Flow chemistry is a process involving channels or tubings to carry out a reaction. Reagents and reactants are pumped through a mixing device in a temperature-controlled tubing: the

reactor and final mixture is collected at the exit. Flow chemistry has, over the last two decades, become widely used in organic chemistry thanks to its easy scalability, safety in using hazardous chemicals or short lifetime chemicals as well as its process being continuous.^[92] Flow chemistry has numerous applications in multiphase reaction systems (liquid-liquid, gas-liquid, solid liquid or even triphasic solid-liquid-gas).^[92] In the case of liquid-liquid reactions, flow chemistry reactors can considerably improve efficiency of the reaction compared to batch methods thanks to enhanced heat transfer and efficient phase mixing from increased interfacial area. A half-filled 50 mL round bottom flask has for example a surface area of approximately $66 \text{ m}^2 \cdot \text{m}^{-3}$ while a coiled tubing allows between $830 - 3200 \text{ m}^2 \cdot \text{m}^{-3}$ interfacial area when the internal diameter size of the tubing ranges from $0.5 - 1.0 \text{ mm}$.^[92,93] This increased area leads to higher mass transfer, hence being the perfect reaction vessel for mass transfer limited reactions. Mass transfer can be increased using static mixers such as packed beds or active mixers such as pulsing or oscillations.^[93] Phase-transfer catalyst can also be used to facilitate exchange of one compound from one phase to another.^[93]

Segmented flow systems

A lot of liquid-liquid flow regimes used are segmented and defined as slug (i.e. a T-mixer allows the formation of droplets of the liquid 1 in liquid 2, this generates alternating segments of liquid 1 and 2) (see Figure 1-20).^[93] Typically, when this flow pattern is used, internal circular flow in each liquid further increase mass transfer by constantly renewing the interfacial area between the two liquids (see Figure 1-20).^[93] This type of system has been more recently applied to biocatalytic systems. In 2013, Wirth and co-workers demonstrated that the traditional low yielding biocatalytic batch reaction could be improved using segmented flow chemistry.^[83,94] When injecting two immiscible liquids through a T mixer, alternating small segments of the two liquids are created. This property was used to improve sesquiterpene biosynthesis using terpene cyclases. Segmented flow chemistry allowed a continuous extraction of the aqueous layer containing the enzyme and FDP by an organic solvent: pentane. The flow system improved surface-liquid volumetric ratio and provided better control over the mixing and interaction of the enzyme with the organic layer. By optimising the diameter of the tubing and the ratio between the aqueous and organic phase higher reaction rates and yields could be achieved. Higher mass transfer between the two phases is dramatically improved compared to batch synthesis. Constant renewal of the interfacial area

is generated through internal circulation in each segment due to the shear forces between the axis of a segment and the wall of the tubing.^[95,96]

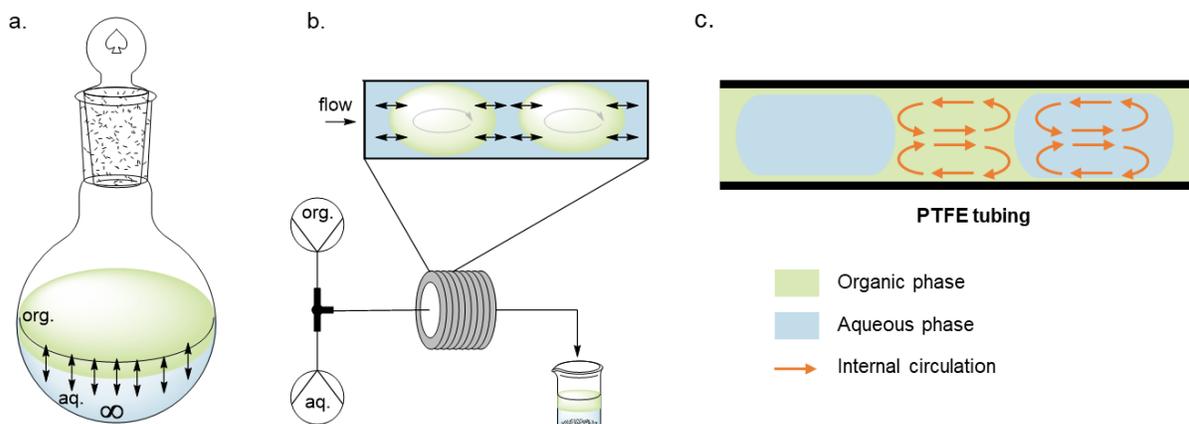


Figure 1-20 a. Batch procedure for terpene syntheses b. Segmented flow set up for terpene synthesis. Black double ended arrows represent mass transfer between the two phases. Grey circular arrow represents the convective flow inside the segment.

A design of experiment (DOE) was used to investigate the influences of the internal reactor diameter, volumetric ratio of aqueous:organic phase, and residence time.^[83] Bigger tubing resulted in lower yields due to insufficient interfacial area. Enzyme deactivation by precipitation was also observed due to excessive exposure to pentane. Similarly, a too high or low ratio of pentane led to lower yields. Time had a proportional influence on the yield and 96% yield (by GC MS) was obtained using 0.35 mM FDP (**14**) with 6 μ M AS in 90 min. (1R,4R,4aS,8aR)-Amorpha-4,11-diene (**49**) was synthesised in 69% yield from FDP(**14**) with ADS in 70 min.^[83]

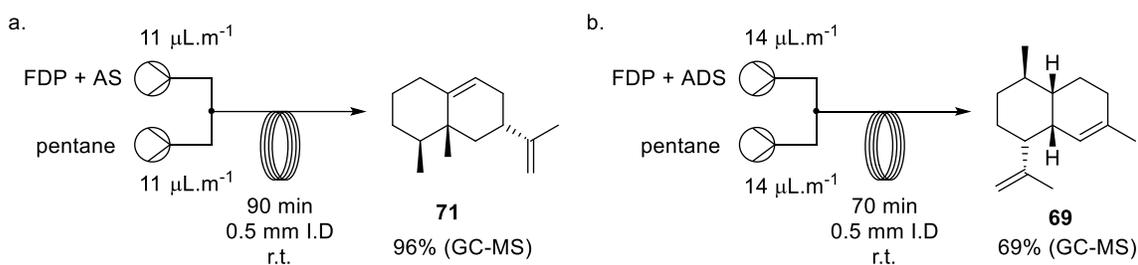


Figure 1-21 a. (1R,7R,8aS)-Aristolochene synthesis in flow b. (1R,4R,4aS,8aR)-Amorpha-4,11-diene synthesis in flow.

Packed bed

Enzymes can be used in a segmented flow system as mentioned above, but most of the research has focused on attaching enzymes to solid supports for intensification process and recycling purposes.^[97] Immobilisation on a carrier offers several attractive advantages; usually enzyme stability is enhanced enabling enzyme recycling, sometimes selectivity and reaction rates are improved.^[98] Growing interest in this method has driven research and nowadays many immobilisation supports exist with a broad range of enzymatic reactions performed using immobilised proteins.^[99–102] However only few examples can be found in the literature for the immobilisation of terpene synthases. Croteau and co-workers showed (-)-patchoulol cyclase could be immobilised by coupling the cyclase amino groups to N-hydroxysuccinimide activated 6-aminohexanoic acid Sepharose 4B. Storage stability at 4 °C was improved, however, the enzyme lost up to 80% of its activity when immobilised.^[103]

In 2020, Wirth and co-workers also successfully immobilised germacradien11-ol synthase (Gd11oIS) and germacradien-4-ol synthase (Gd4oIS) on a hybrid-controlled porous glass. This support was coated with an organic polymer and chelated Fe³⁺ allowing enzyme immobilisation through the fused histidine tag on heterologously expressed synthases. Both enzymes retained their activity for at least 50 cycles with a conversion of about 50%.^[104]

1.5.3 Site directed mutagenesis for improved catalytic efficiency

Sesquiterpene synthases have shown to have on average a lower catalytic efficiency than other enzymes. Few investigations have focused on optimising their catalytic efficiency for a better yield. Zhang and co-workers observed a dramatic improvement with a single amino acid substitution in the active site of ADS. The k_{cat} of of the T399S ADS variant was measured to be 83% higher than the wild type with k_{cat}/K_m increasing from 0.385 to 0.662 $\mu\text{M}^{-1}\cdot\text{s}^{-1}$.^[42] This could be explained by serine favouring the deprotonation step more than threonine, and possibly a more hydrophilic active site favouring product release. Further studies from Quax and co-workers identified a second key residue for catalytic improvement in ADS using a bioluminescent assay (malachite green) for a high throughput screening of the variant's catalytic activity.^[105] The H448A variant was shown to have a better catalytic efficiency by 4-fold and the double mutant T399S/H448A demonstrated to have a 5 times better turnover rate while conserving a similar product profile. When incorporated for (1R,4R,4aS,8aR)-amorpho-

4,11-diene production *in vivo* in *E. coli*, the double mutant produced 4 times more (2.7 mg/L) than the wild type (0.7 mg/L). Clearly these studies show the beneficial impact of a higher k_{cat} on the sesquiterpene yield.

Table 1.2 Comparison of the kinetics characteristics of ADS wild type and mutants.

Enzyme	k_{cat} (s^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \cdot \text{M}^{-1}$)
ADS WT	0.20±0.007	5.47±0.558	3.6×10 ⁴
ADS H448A	0.68±0.03	5.04±0.753	13.5×10 ⁴
ADS T399S	0.39±0.06	5.89±1.278	6.6×10 ⁴
ADS T399S/H448A	1.00±0.14	8.50±1.605	11.8×10 ⁴

1.6 AIM OF THE PROJECT

1.6.1 Context

Terpenoids and terpenes are high value compounds used in a broad range of industries and as intermediates for chemical synthesis of high value compounds. Due to their structural complexity, the majority of those compounds are still extracted from natural sources. More recently, through technology advances and growing understanding of their biosynthesis, terpenes such as artemisinin were successfully produced *in vivo* from sugars through an engineered host. However, a deeper understanding of the whole machinery is needed to produce non-natural terpenes through this process. The rapid rise of resistance against drugs and industry demand for new terpenoids drive a constant race for the development of new alternatives, either with analogues of existing compounds or completely novel compounds. The use of sesquiterpene cyclases *in vitro* combined with modified substrates proved to be a successful pathway to produce new compounds. While existing chemistry methods cannot compete with the high regio-, stereo- and chemo-selectivity of terpene synthases yet. Nevertheless, key milestones still need to be achieved to exploit this promising chemoenzymatic pathway. For example, further investigations into the ability of terpene synthases to accept non-natural FDP precursors should be accomplished. If for example other simple di-methylated terpenes can be generated from their corresponding FDP analogues such as the novel (S)-14,15-dimethylgermacrene D (**86**) and show enhanced or potential semiochemical activity, it would provide a predictable, reliable process to produce new

compounds for testing against aphids, a major crop pest. However, for this to be economically viable, two key challenges need to be overcome: an efficient procedure should be designed for the FDP precursor synthesis and an efficient incubation process with the cyclase need to be designed. Currently the limiting factor for FDP synthesis is the low yield and cumbersome diphosphorylation step of the farnesol analogue to its FDP analogue. Finally, despite the recent discovery of the high yield segmented flow process for the enzymatic conversion of FDP (**14**) by terpene synthases, improvements can be achieved in terms of reaction time and application with FDP analogues or other sesquiterpene synthases.

1.6.2 Aims

The project was divided into parts aiming to focus on the key issues in utilising terpene synthases *in vitro* for the synthesis of known or new terpenes. The first aim was to find a reliable and efficient alternative procedure to the traditional diphosphorylation of farnesol (**119**) which involved two time-consuming ion exchange procedures. Two approaches were investigated for this process; a chemical synthesis without the need for ion exchange involving a trimethylsilyl protected diphosphate (**120**), and an enzymatic approach using kinases to diphosphorylate farnesol (**119**). The second aim was to investigate the production of novel terpenes from FDP analogues. Building on the discovery of the novel aphid attractant (S)-14,15-dimethylgermacrene D (**86**), other sesquiterpene analogues were investigated using (S)-germacrene D synthase and germacradien-4-ol synthase, leading to the full characterisation of a novel cyclic ether. The final part of the project focused on exploring segmented flow chemistry, oscillatory flow chemistry and high-performance counter-current chromatography (HPCCC) as tools for efficient and high yielding *in vitro* synthesis of high value sesquiterpenes.

CHAPTER 2

DEVELOPMENT OF AN EFFICIENT DIPHOSPHORYLATION PROCESS

2.1. PREFACE

Terpene synthases produce a range of high value compounds. The use of these enzymes on an industrial scale requires an efficient, economic, and safe synthetic route for producing their substrates. This chapter outlines the chemical challenges in substrate production, namely: the diphosphorylation of farnesol to form farnesyl diphosphate (FDP) and proposes an enzymatic alternative to traditional approaches. Isoprenoid diphosphates are terpene synthases' natural substrate. The abstraction of the diphosphate group in the enzyme active site is the essential step initiating a carbocation cascade of reactions converting the linear isoprenoid diphosphate to a high-value terpene product. To generate the first carbocation and trigger subsequent reactions, the diphosphate group coordinates to the metal cluster in the enzyme catalytic site. This creates an electrophilic driving force and triggers ionisation of the substrate. In sesquiterpene synthases (STS), it starts from farnesyl diphosphate (FDP), its natural substrate; upon abstraction of the diphosphate group, a farnesyl cation is generated. It will then undergo a series of complex reactions leading to the final product. The first chemical synthesis of FDP (**14**) from farnesol (**119**) was reported by Cramer *et al.* in 1959.^[106] It involved the reaction of farnesol with phosphoric acid in trichloroacetonitrile, generating the corresponding diphosphate after two in-situ successive phosphorylations. (see Figure 2-1) This procedure was successively modified by Cornforth *et al.*, Danilov *et al.* and Keller *et al.* to improve the yield and selectivity towards the diphosphate and establish a time-efficient procedure for the preparation of diphosphorylated polyprenols.^[107-109] In practice, a solution of bis(triethylammonium) hydrogenphosphate (TEAP) in acetonitrile is added to farnesol (**119**) in trichloroacetonitrile over three separate additions (three additions at five-minute intervals at 37 °C). TEAP is easily prepared by mixing triethylamine and phosphoric acid in acetonitrile. The reaction mixture is then transferred to a silica column pre-equilibrated with the following buffer: *i*PrOH:NH₄OH:H₂O, 6:2.5:0.5. This allows the separation of phosphates species and exchange of the triethylammonium counter ion to readily generate the trisammonium salt. Optimised, this procedure only takes thirty minutes to run and is followed by a flash chromatography purification with yields varying from 20 – 40%. High purity is not easily achieved due to the formation of a mixture of mono-, di-, tri-, tetraphosphate species. In 1986, Davisson *et al.* developed a selective method.^[110,111] This process is now the prevailing method to synthesise FDP (**14**) (see Figure 2-1). It requires the conversion of farnesol into the corresponding halogenated derivative (preferably the chloride form) followed by its S_N2 displacement with tris(tetrabutylammonium) hydrogendiphosphate ion (**123**). The

tris(tetrabutylammonium) hydrogendiphosphate (**123**) ion must be prepared from the commercially available disodium dihydrogen diphosphate (**124**) through ion-exchange chromatography. It should be noted that poor handling of this highly hygroscopic compound yields unsuccessful diphosphorylation. After a usually overnight reaction of the farnesyl halide and diphosphorylating agent, the resulting product (**125**) must undergo a second ion exchange chromatography followed by lyophilisation and purification by HPLC or reverse phase flash chromatography. The second ion exchange is required to get a satisfactory level of purification because tetrabutylammonium counteractions leads to tailing during the chromatography step. Moreover, tri(tetrabutylammonium) FDP salt is a tacky solid difficult to handle, whereas ammonium salts of FDP is a fluffy white powder. Altogether the procedure from Davisson *et al.* is cumbersome to perform and scale up due to the two time-consuming ion-exchange chromatography processes. Moreover, although selective toward the diphosphate species, yield reproducibility is mediocre with reported values in the literature ranging from 10-70%.^[78,80,82,90]

All the procedures described above require either a lot of time or are not selective; in both cases yield is below 50% on average and both methods should not be run on more than one mmol scale. The high cost, low yield, and unscalable diphosphorylation process renders the *in vitro* chemoenzymatic synthesis of terpenoids industrially unattractive as well as making its use troublesome for research purposes. Interestingly, no new alternative in the literature has been proposed since the publication of the unselective method (2005) by Keller *et al.*

In 2005, Wessjohann *et al.* published a patent showing that the tetratrimethylsilyl ester of the diphosphoric acid (**120**) could be used to displace prenyl bromide (**126**) and generate prenyl diphosphate (DMADP, **10**) in excellent yield (83%).^[112] (see Figure 2-1) Synthesis of the diphosphorylating agent (**120**) can be performed on a gram-scale from disodium dihydrogen diphosphate (**124**) in under an hour.^[113] The TMS protected diphosphate (**120**) is reported to be stable for at least two months at -20 °C.^[113] Simple reaction with water leads to the hydrolysis of one tetratrimethylsilyl ester which allows subsequent deprotonation of the hydroxide by *N,N*-Diisopropylethylamine (DIPEA) to form an active species capable of displacing the bromide (see Figure 2-1). Upon the formation of DMADP, HBr-DIPEA crystals will precipitate driving the reaction towards the product. Final isolation of DMADP (**10**) is performed by deprotecting the diphosphate group using a basic aqueous solution containing a counter-ion of choice such as NH₄⁺ from 6M NH₄OH. This method appears to be selective,

rapid, scalable (up to grams) and no time-consuming ion exchange chromatography is involved.

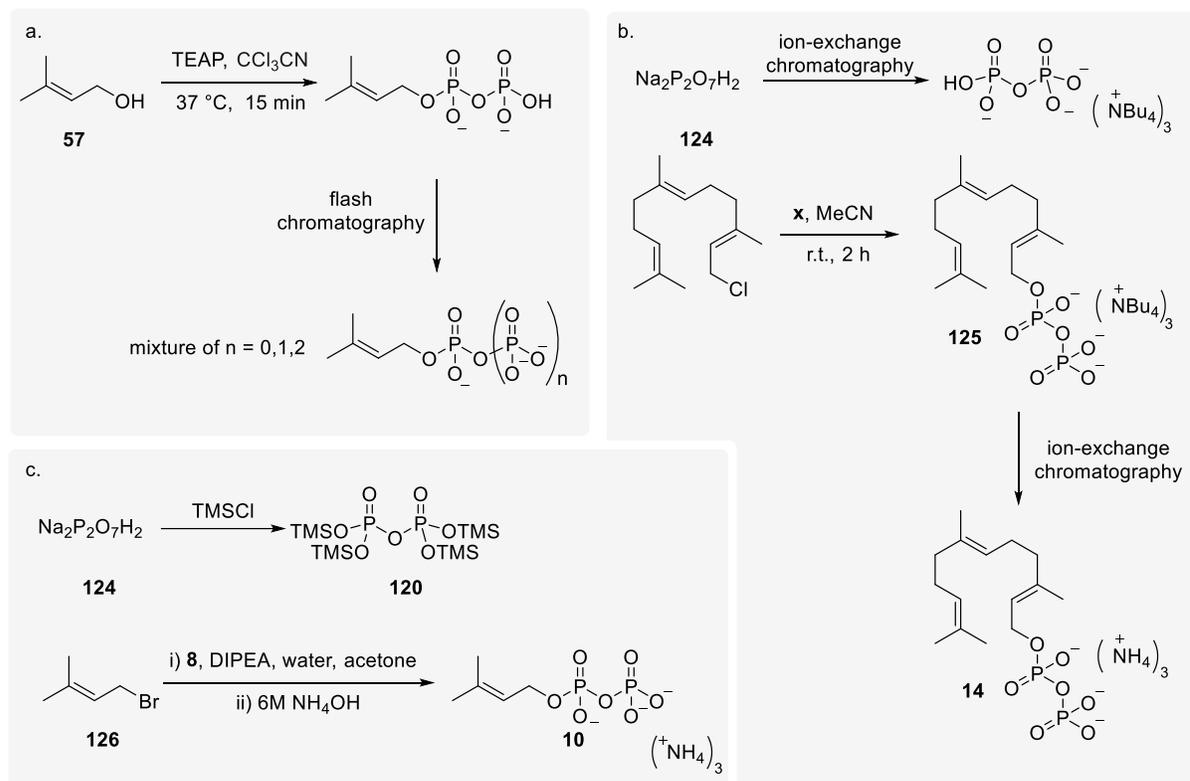


Figure 2-1 Various procedure for the chemical diphosphorylation of polyprenol. Prenol is used as example. a. Cramer method $n = 0, 1, 2$ b. Davisson method c. Wessjohann method.

This method was therefore investigated to synthesise FDP (**14**). Synthesis of the protected diphosphoric acid (**120**) is sensitive to acid impurities from hydrolysis of trimethylsilyl chloride, however, it was successfully synthesised in quantitative yields and characterised by NMR spectroscopy. Residual hydrolysis product of **120** could be seen in the ^1H NMR spectra when performing the characterisation in deuterated chloroform with signal at $\delta = 0.03$ ppm. The next step was performed on farnesyl bromide (**127**). Farnesyl bromide (**127**) was synthesised from farnesol using PBr_3 as brominating agent. Both steps combined resulted in 30-40% yield on 2 - 4 mmol scale. Using alternative bases such as 2,2,6,6-tetramethylpiperidine or 1,8-diazabicyclo[5.4.0]undec-7-ene resulted in poor yields. This process is overall less time consuming (three hours including both steps and all work-up and one-day waiting time), cheaper, and feasible on a bigger scale than the Davisson *et al.* method (here 4 mmol). However, this method showed poor reproducibility in terms of yield due to the second step

being difficult to repeat consistently. Success of the procedure was only as high as 43% and reasons why the procedure did not yield the intended product could not be clearly identified.

An alternative enzymatic pathway was therefore investigated to perform the diphosphorylation of farnesol (or analogues of farnesol) to give FDP (or analogues). Diacylglycerol kinase A from *Streptococcus mutans* was recently discovered to have an activity as an ATP-dependent undecaprenol kinase (UK) and turn undecaprenol (**128**) into undecaprenol phosphate (**129**).^[114–116] (see Figure 2-2) Varieties of polyprenol from C10 to C55 were also shown to be excellent substrates for this enzyme.^[117] In parallel, isopentenyl phosphate kinase (IPK) from *Methanocaldococcus jannaschii* has been observed by Grochowski *et al.* to catalyse the ATP-dependent phosphorylation of isopentenyl phosphate (IP, **33**) to isopentenyl diphosphate (IDP, **9**) in the archaea isoprenoid Lost pathway (see Chapter 1 and Figure 2-2).^[118]

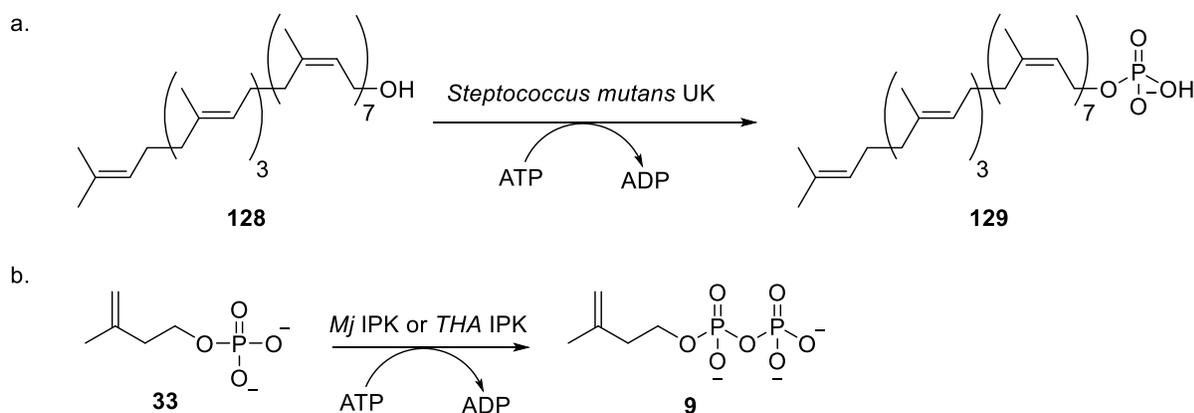
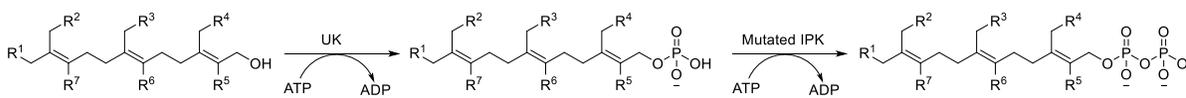


Figure 2-2 a. UK-catalysed conversion of undecaprenol to undecaprenol phosphate. b. IPK-catalysed conversion of IP to IDP.

The crystal structures of IPK from *Methanocaldococcus jannaschii* (*Mj*) and *Thermoplasma acidophilum* (*Tha*) in complex with IP (**33**) and IDP (**9**) show that the C5 substrate lies in a hydrophobic pocket.^[119,120] In *Mj* IPK, IP is surrounded by Ala63, Phe76, Met79, Phe83, Ile86, Met90, Ile146 and Ile156 around the tail.^[120] Rational engineering of *Mj* IPK and *Tha* IPK was successfully performed to direct phosphorylation of bigger substrates such as farnesyl phosphate (**130**) and geranyl phosphate (**131**) respectively.^[119,120] *Mj* IPK F83A I86A I146A was identified to convert 27% farnesyl phosphate (FP, **130**) to FDP (**14**).^[120] Coupled together with a terpene synthase they could potentially successively phosphorylate farnesol (or analogues of farnesol) to generate any sesquiterpenes. (see Scheme 2-1)



Scheme 2-1 Potential two-step enzymatic synthesis of FDP analogues from farnesol analogues using UK and IPK.

Thus far, there has been no enzymatic procedure to diphosphorylate farnesol (or analogues) reported in the literature; even less from farnesol to the sesquiterpene product.

Here, the aim was to first heterologously produce undecaprenol kinase (UK), then perform site-directed mutagenesis of isopentenyl phosphate kinase to rationally create variants capable of efficiently phosphorylating farnesyl phosphate (or analogues). Combined, they can be used as a tool for a rapid diphosphorylation of polyprenol and be potentially coupled with a sesquiterpene synthase for the rapid synthesis of sesquiterpenes from farnesol (or analogues).

2.2. UNDECAPRENOL KINASE

2.2.1 Expression

A codon-optimised gene encoding for undecaprenol kinase (UK) from *Streptococcus mutans* was purchased in a *pMARQ* plasmid from Thermofisher (see Table 6-2, Material and methods) and subcloned into a *pET28a* expression vector encoding an N-Terminal poly-histidine tag (His Tag) through golden gate assembly (**plasmid 1**). Protocols are described in detail in section 6.1 Materials and methods. Successful cloning was verified through digestion with two restriction enzymes. Restriction enzymes have unique cleavage sites in a DNA sequence and will result in a digestion pattern that allows identification of the correct clone. This digestion pattern was visualised by agarose gel electrophoresis which separates out DNA fragments by size. Visualisation of the fragments were obtained by pre staining the agarose gel with SybrSafe, a fluorescent dye that intercalates the DNA bases. **Plasmid 1** was digested with NcoI and XhoI, yielding theoretically two 5231 and 504 base pair fragments when digested (calculated using Benchling online software^c). Digested samples were analysed by agarose gel electrophoresis. The size of DNA fragments after digestion were an identical

^c <https://benchling.com>

match to what expected and therefore showed evidence of successful golden gate procedure. The newly formed plasmid was sequenced (Eurofins Genomics) to further validate the formation of the new construct pET28a-UK.

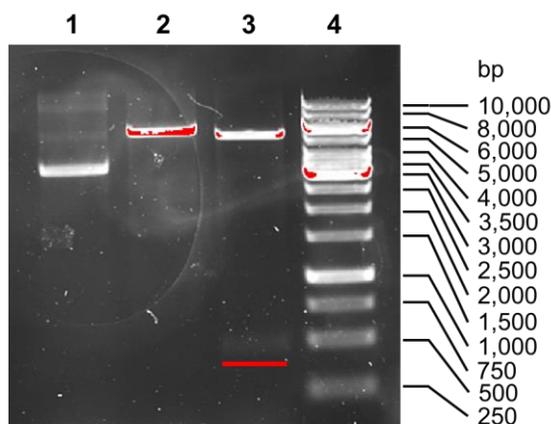


Figure 2-3 Agarose gel showing the results from digestion of **Plasmid 1** *pet28a-UK*. Lane 1) uncut *pet28a-UK*, Lane 2) digestion with *XbaI*, Lane 3) digestion with *NcoI* and *XhoI*.

Gene expression in the *pET28-a* vector is controlled by a T7 promoter and contains a gene encoding for the lac repressor protein (LacI). **Plasmid 1** was transformed in *E. coli* BL21 (DE3) cells for expression. *E. coli* strains such as BL21 (DE3) are ideal hosts for expressing these genes as they have been engineered to contain the gene for T7 RNA polymerase in their genome whose expression is under the control of the lacUV5 promoter. This allows control of the gene expression by IPTG (isopropyl- β -D-1-thiogalactopyranoside). If IPTG is added to the culture media, LacI unbinds the lac promoter leading to transcription of T7 polymerase into mRNA. Subsequently, the T7 polymerase can bind to the T7 promoter situated on the plasmid of the gene of interest. IPTG is used preferentially to allolactose because it cannot be degraded by bacterial metabolic enzymes. Expression of **Plasmid 1** in BL21 (DE3) cells were induced with IPTG (0.2 mM) when the cell culture achieved an optical density of 0.6 at 600 nm (OD_{600}). Cultures were then incubated at 16 °C overnight with shaking prior to harvesting the cells. Protein expression protocols are detailed in section 6.1 Materials and methods. SDS polyacrylamide gel electrophoresis (SDS PAGE) was used to verify protein overexpression by observing the presence of a band at the estimated molecular weight of UK: 18.5 kDa (estimated using from the amino acid sequence using Benchling online software). SDS PAGE was used to also verify if the protein was in a soluble or insoluble form by lysing a small

amount of cells through sonication. Unfortunately, no expression was observed. In the attempt to identify the issue, other cell lines were tested such as BL21-CodonPlus(DE3)-RP, BL21-CodonPlus-RIL, C41(DE3) pLysS. C41(DE3) strains are derived from BL21(DE3) and effective in expressing toxic and membrane proteins as they have an unknown mutation which prevents cell death.^[121] Moreover, toxic proteins are better expressed because the *pLysS* plasmid allows the production of T7 lysozyme to reduce basal level expression of the gene of interest.^[122] BL21-CodonPlus-RIL cells contain additional copies of tRNA gene which recognises AGA/AGG, AUA and CUA codons for expression of AT-rich genomes. BL21-CodonPlus-RP contains additional copies of tRNA gene to recognise AGA and AGG arginine codons and CCC proline codon for GC-rich genomes. Unfortunately, in all attempts using various media, induction temperature and strains, no expression was detected by SDS PAGE. Since various conditions and cell lines yielded no successful expression, the gene was subcloned into a different expression vector, *pET32-Xa/Lic*. The *pET32a-Xa/Lic* is an expression vector encoding a thioredoxin tag (Trx Tag) as well as a N-terminal His Tag and an S Tag, all cleavable using a Factor Xa protease, an optional C-terminal His Tag is also encoded. The Histidine Tag (His Tag) is usually composed of 6-10 consecutive histidines and allows the rapid and inexpensive purification of the protein of interest by immobilised metal affinity chromatography (usually nickel). The S Tag is a defined 15 amino acids sequence and allows an easy purification of the protein of interest because has a strong interaction with the 103 amino acid S-protein. The Trx Tag is generally recommended for improving the solubility of the protein of interest and the production of soluble proteins containing disulphide bonds as it has been shown to allow bond formation in *E. coli* cytoplasm.^[123]

The plasmid *pET32 Xa/Lic* was present in the Allemann group library and contained the gene for Z-farnesyl diphosphate synthase. This plasmid was used as a source of the required backbone *pET32 Xa/Lic* vector. DNA encoding for the UK gene from **Plasmid 1** and the *pET32 Xa/Lic* vector were amplified by polymerase chain reaction (PCR) and the UK gene was then cloned in the *pET32 Xa/Lic* vector using the Gibson Assembly[®] method (**Plasmid 2**). The primers used for these PCR reactions and protocols are described in detail in section 6.1 Materials and methods. Successful cloning was verified through digestion. **Plasmid 2** was digested with HindIII and XbaI, yielding two 5344 and 915 base pair fragments when digested while the original vector containing zFPS would yield two fragments of 1362 and 5344 base pairs. Digested samples were analysed by agarose gel electrophoresis. Samples showed

evidence of successful cloning procedure and were sequenced to further validate the formation of *pET32 Xa/Lic-UK*. (see Figure 2-4, lane 7)

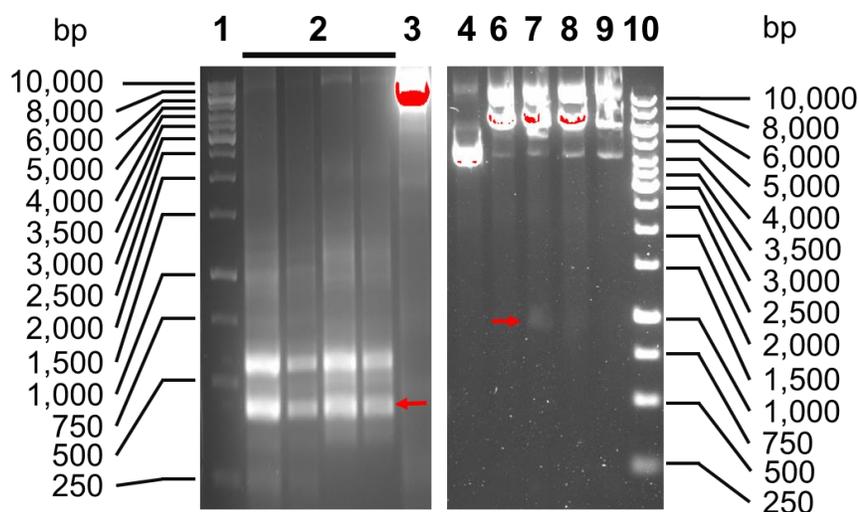


Figure 2-4 Agarose gel showing the results from digestion of **Plasmid 2** *pet32 Xa/Lic-UK*. Lane 1) Ladder, Lane 2) PCR products from the amplification of plasmid A1, Lane 3) PCR products from the amplification of *pet32 Xa/Lic* vector containing Z-FDPS Lane 4) uncut plasmid 2. Lane 6-10) Different Gibson assembly products digested with *HindIII* and *XbaI*, lane 7 contains the 2 expecting bands.

In order to determine if over expression from the new vector was occurring, *E. coli* BL21(DE3), BL21-CodonPlus(DE3)-RP, BL21-CodonPlus-RIL, C41(DE3) pLysS cells were transformed with the **Plasmid 2** and 2 mL test expressions were carried out to investigate which cell line was suitable for expression. Cultures were induced with IPTG (0.2 mM) at $OD_{600} = 0.6$ and incubated for a further 4 hours at 37 °C. Western blot was used to characterise expression using the His Probe HRP (horseradish peroxidase) as this method is more sensitive and specific for the detection of protein than SDS PAGE. (see Figure 2-5) Expressed UK in *pET32 Xa/Lic* is expected to have a relative molecular mass of 31,819 Da (calculated from the amino acid sequence using Benchling online software). The protein appeared to express well in *E. coli* BL21(DE3), BL21-CodonPlus(DE3)-RP, BL21-CodonPlus-RIL as it can be detected around the 30 kDa band in the protein ladder. **Plasmid 1** showed no expression of UK (see Figure 2-5). Because expression temperature has an impact on level and quality of expressed proteins, various temperatures (16 °C, 20 h; 20 °C, 20 h, 25 °C, 20 h, 37 °C, 4 h) were investigated using BL21 (DE3) strains. Samples of the culture were taken after and before induction and analysed by SDS PAGE. It appears only overnight expression at 16 °C yielded UK (see Figure 2-5). This can be explained by a reduced rate of overall bacteria

mechanism (translation, transcription, cell division) when expressing at a low temperature and potential reduced protein aggregation and degradation. Therefore, under the optimal conditions found, **Plasmid 2** was expressed in BL21 (DE3) cells at 37 °C until reaching $OD_{600} = 0.6$, the temperature was reduced to 16 °C, cells were induced with IPTG (0.2 mM) and expression was allowed to continue overnight (20 h).

An attempt to extract and purify the protein from the pellets was performed by sonication and purification by nickel affinity chromatography. Ni-NTA affinity column was used for the purification. The supernatant (40 mL) after sonication was applied to the resin, the His Tag motif from the protein coordinates to Ni^{2+} and other proteins are eluted in the flow through as they do not bind to the resin. The resin is then washed in fractions with an increasing amount of imidazole (20 mM – 500 mM in buffer) to elute the protein from the affinity column (Full procedure is described in section 6.1 Materials and methods). SDS PAGE was used to identify which fractions contained the protein of interest. UK started to elute in buffer containing 250 mM imidazole (see Figure 2-6). Fractions containing UK were pooled and loaded on a desalting column to remove the imidazole and concentrated by ultrafiltration (see Figure 2-7). Protein concentration was measured using the modified Bradford procedure. Despite presence of the Trx Tag, an only small amount of soluble protein was released from the supernatant after sonication (see Figure 2-6a). A basic extraction procedure was therefore subsequently performed to extract the protein from the inclusion bodies. This procedure involves unfolding and refolding the protein's tertiary structure through increasing the pH (around 12) of the solution containing the resuspended pellet (unfolding step) followed by decreasing it back to pH 8 (refolding step). The buffer contained Tween® 20 (1%) because such detergent helps solubilising membrane proteins, glycerol (10%) was also added to reduce protein aggregation and improve its stability^[124] and TCEP (0.1 mM) was added as reducing agent to further reduce protein aggregation from the formation of disulfide bonds between cysteine residues. The basic extraction method showed a slight improvement in protein extraction to the supernatant solution as determined by SDS-PAGE (see Figure 2-6b).

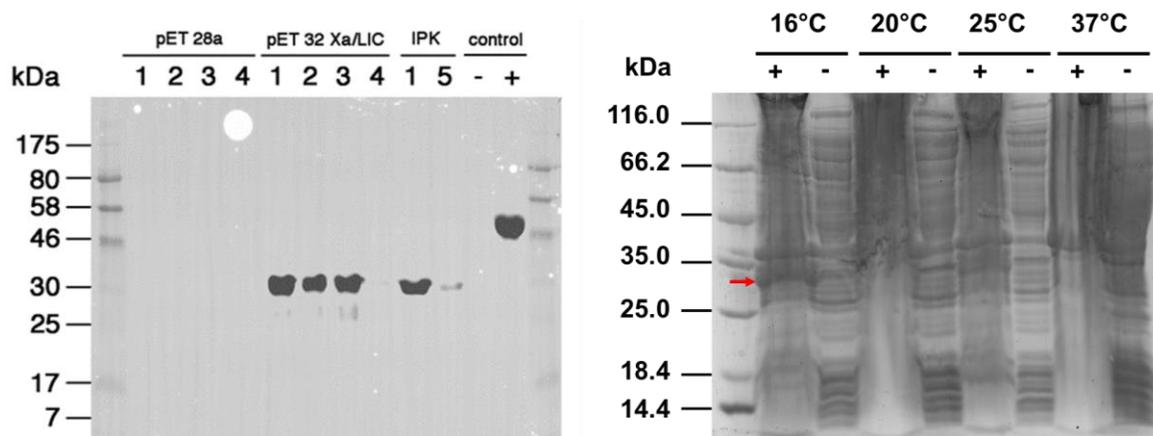


Figure 2-5 Left: Western blot gel of the test expression of the UK gene in *pet28a* (**Plasmid 1**), *pet 32 Xa/Lic* (**Plasmid 2**) and *IPK* wild type as positive reference for expression. (1-4) *E. coli* BL21-(DE3), BL21-CodonPlus(DE3)-RIL, BL21-CodonPlus(DE3)-RP and BL21-CodonPlus(DE3)-RIL and C43(DE3) pLysS; (5) *E. coli* BL21-AI cells; Two proteins of 45 kDa (+) and 35 kDa (-) size were used as positive and negative control respectively for the western blot. Left: SDS-PAGE: analysis of the pellet after expression at various temperature. +: after induction and -: before induction

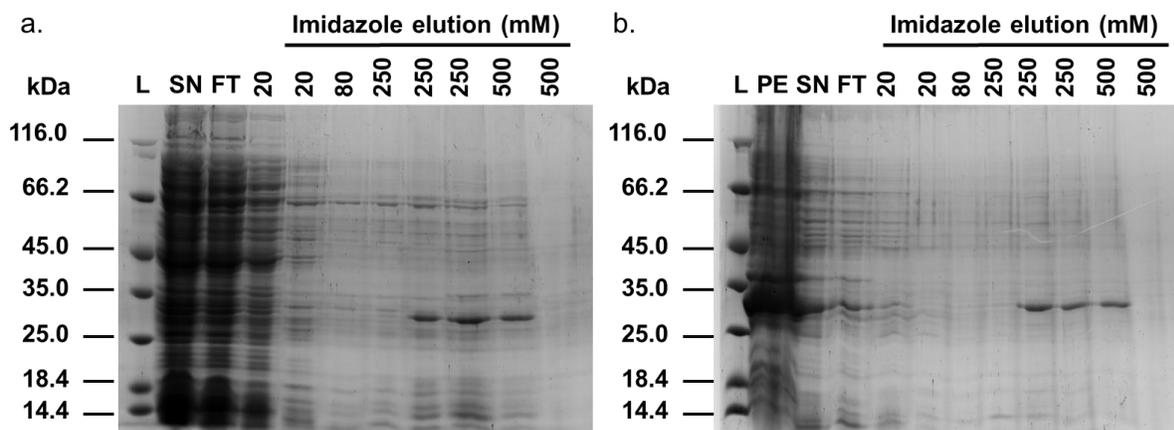


Figure 2-6 Purification of undecaprenol kinase. a. SDS-PAGE: analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography. B. SDS-PAGE: analysis of the pellet (PE) and supernatant (SN) after basic extraction of the pellet after sonication followed by purification by affinity chromatography of the supernatant.

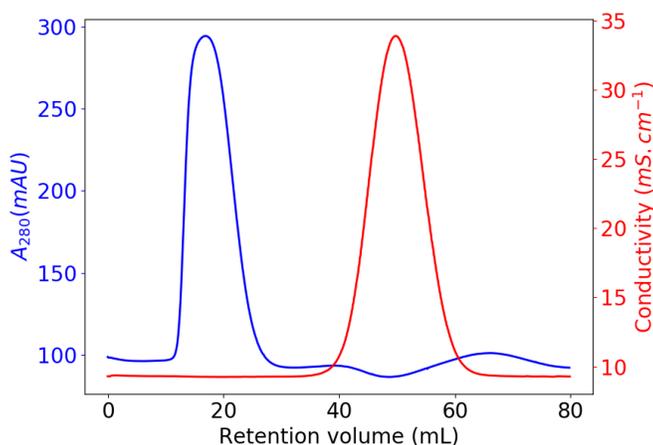


Figure 2-7 FPLC chromatogram of gel filtration chromatography using Sephadex G-25 resin for desalting the pooled fractions containing UK after nickel affinity chromatography.

2.2.2 Incubation of farnesol with UK

Analytical incubations of UK and farnesol were carried out to test the activity of the purified undecaprenol kinase. Farnesol (5 mM) together with UK (10 μ M) were put in incubation buffer (20 mM Tris, 20 mM MgCl₂, 1 mM ATP, 10 mM PEP, 20 U pyruvate kinase) and left for 2 hours at 37 °C with shaking (250 rpm). ATP (**132**) was regenerated using pyruvate kinase (PK) with phosphoenolpyruvic acid (PEP, **133**). (see Figure 2-17b for ATP-PEP recycling system). A white precipitate formed over time. Analysis of the precipitate by TLC against chemically synthesised FP (**130**) and FDP (**14**) using the method by Keller *et al.* showed the solid had the same R_f as FP. (see Figure 2-8) Interestingly, farnesyl phosphate (FP, **130**) precipitated upon formation allowing easy isolation for characterisation. The reaction was also followed by ³¹P NMR spectroscopy. Four signals could be seen in the control corresponding to the α , β and γ phosphorus of ATP (**132**) and phosphorus in PEP (**133**). Over time two new signals appeared at $\delta = 3.3$ and 2.2 ppm corresponding to the phosphorus from FP (**130**) and free phosphate respectively. The intensity of the signal for FP quickly decreased (see Figure 2-8), this can be explained by the precipitation of FP over time. No precipitate was observed after 17 and 31 min, and after 53 min the whole NMR tube was filled with white solid. This correlates with the NMR spectrum at 53 min, where the signal corresponding to FP had a very low intensity.

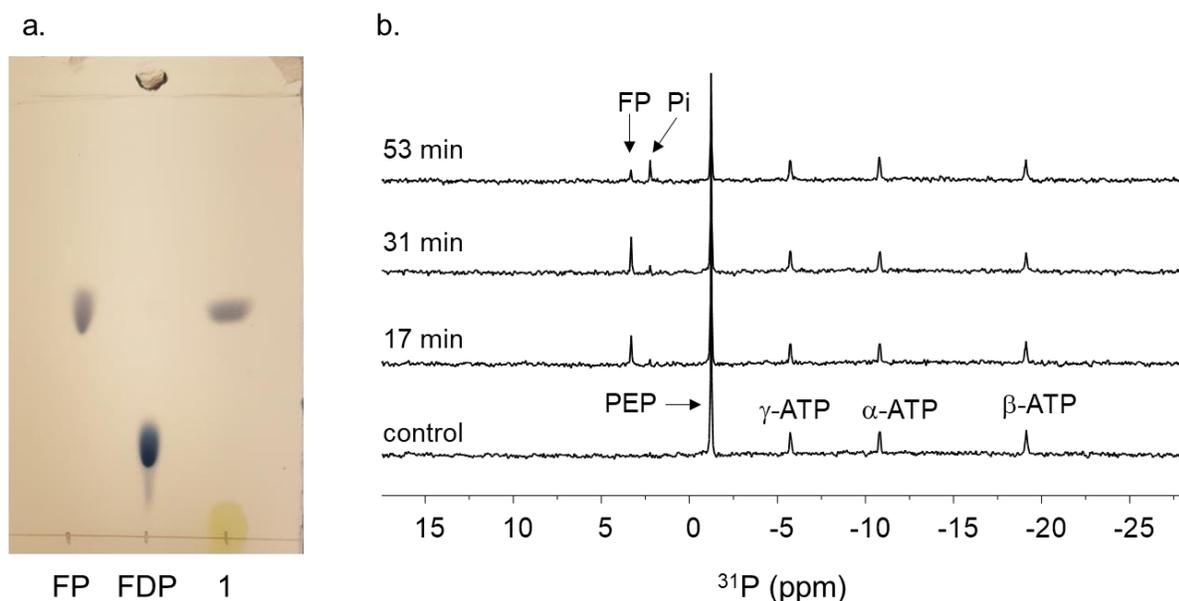


Figure 2-8 UK-catalysed synthesis of FP from farnesol. a. TLC plate with FP and FDP chemically synthesised, lane 1: precipitate from the incubation of farnesol with UK. b. ^{31}P NMR spectra recorded over the time of the incubation of farnesol with UK. FP is formed after only 17 min and the intensity of the peak decreases when FP precipitates.

2.3. ISOPENTENYL PHOSPHATE KINASE

The optimised gene encoding for the isopentenyl phosphate kinase from *Methanocaldococcus jannaschii* was purchased from Thermofisher and subcloned in a *pET28a* vector by Dr. Luke Johnson from the Allemann group. Wild type IPK was produced in *E. coli*. BL21 (DE3). As shown previously, the wild type IPK does not readily accept FP (**130**) as substrate. However, published crystal structures in the “apo” form and complexed with IP or IDP allow rational mutagenesis for larger isoprenoid acceptance. Several key residues in IPK were identified in the hydrophobic pocket near the IP (**33**) tail and mutated for FP (**130**) acceptance.

IPK is composed of two domains, the N-terminal domain is responsible for binding the phosphate group in IP while the C-terminal domain binds the phosphate group in ATP (**132**). In this hydrophobic pocket, five residues: at the back of the binding pocket F76, F83, at the front of the binding pocket I86, I146, I156 can be potentially mutated to accommodate

substrates with a longer tail than IP (**33**). (see Figure 2-9) Only variants containing mutations at the front and back of the pocket were designed as studies from Dellas *et al.*^[120] showed that mutations at the back were only effective combined with a mutation at the front. For this purpose, point mutations were carried out to create a library of 18 variants containing single, double, triple, quadruple, or quintuple substitutions. All residues were mutated to alanine. Alanine was chosen because it removes all carbons past the β -carbon, thus reducing steric effects to its minimum. Replacing residues with glycine (the smallest amino acid) would lead to backbone flexibility and potential conformational changes which would likely render the protein inactive. Moreover, substitution by alanine minimises alteration of the hydrophobic environment of the active site since all selected amino acids for mutation are also hydrophobic. The primers used for SDM and protocols are described in detail in section 6.1 Materials and methods. All variants were sent for sequencing to verify successful SDM. Single- and double-point variants were not expressed as literature shows they very weakly accept GP (**131**) or FP (**130**) as substrates. A total of 10 variants were expressed in *E. coli* BL 21 (DE3) at 37 °C for 4 hours and purified using nickel affinity chromatography. (see Table 2-1) Variants appeared to express qualitatively with the same yield as the wild type version. (see Figure 2-10).

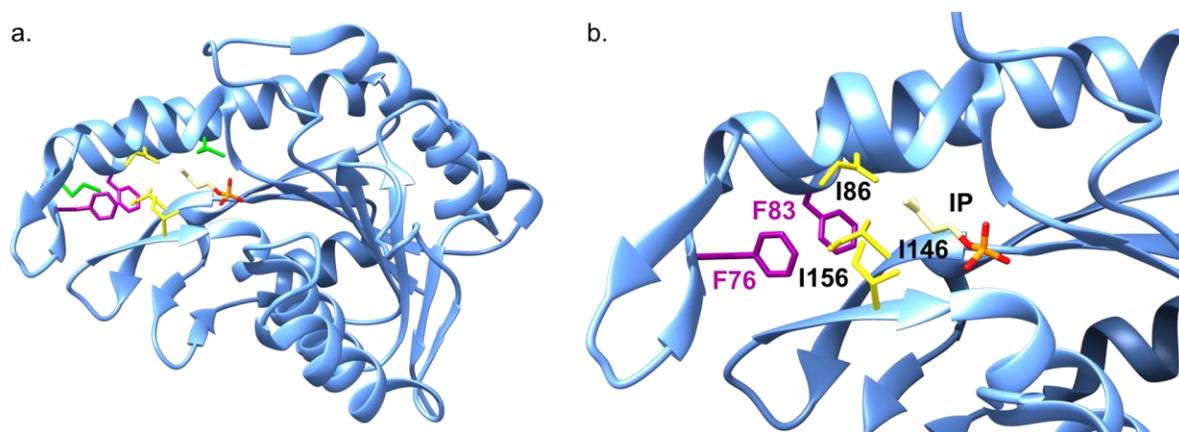


Figure 2-9 Cartoon representations of an X-ray crystal structure of Mj IPK in complex with IP (PDB: 3K52) a. Full view b. Close view of the active site with residues around the IP tail displayed.

Table 2-1 Table of created and expressed variant of Mj IPK by SDM.

IPK variant number	Mutated position				
	76	83	86	146	156
WT	Phe	Phe	Ile	Ile	Ile
1	Ala	Ala	Ala	Ile	Ile
2	Ala	Ala	Ile	Ala	Ile
3	Ala	Ala	Ile	Ile	Ala
4	Phe	Ala	Ala	Ala	Ile
5	Phe	Ala	Ala	Ile	Ala
6	Phe	Ala	Ile	Ala	Ala
7	Ala	Ala	Ala	Ile	Ala
8	Ala	Ala	Ile	Ala	Ala
9	Phe	Ala	Ala	Ala	Ala
10	Ala	Ala	Ala	Ala	Ala

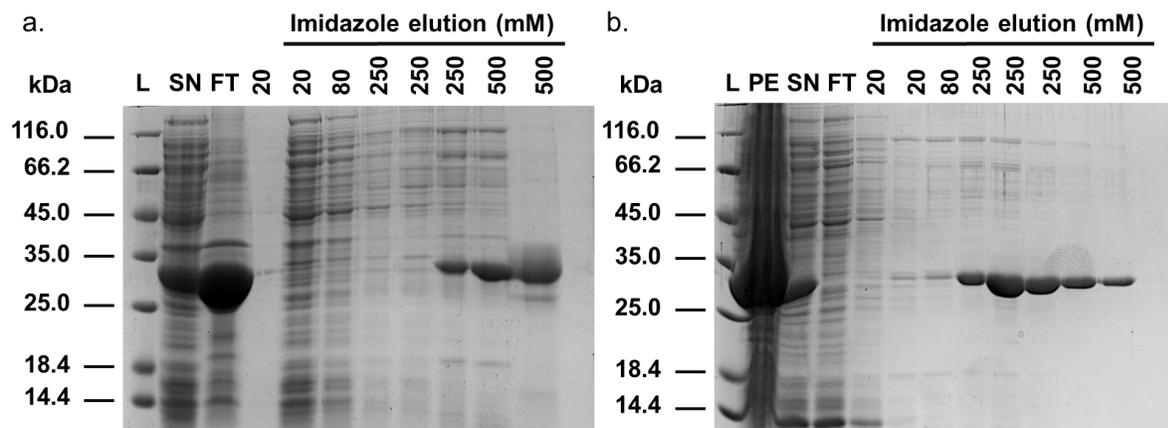
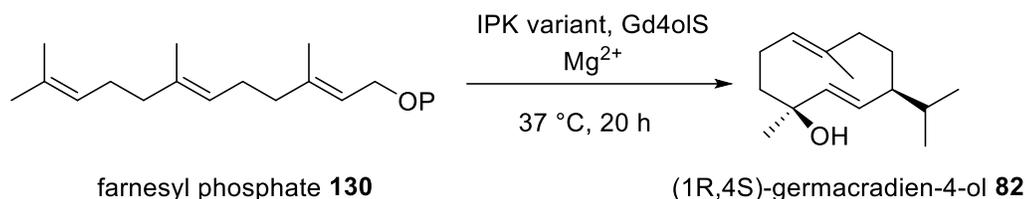


Figure 2-10 Purification of IPK and variant 3. a. IPK WT: SDS PAGE analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography. b. IPK variant 3: SDS PAGE analysis and supernatant (SN) after sonication of the pellet after expression and purification by nickel affinity chromatography.

To test the activity of each variant with FP (**130**) in a relatively high-throughput manner, the reaction was coupled with a sesquiterpene synthase in a concentration 20 times higher than the variant. If the variant was to accept FP (**130**), the FDP (**14**) produced will, in turn, be converted to the sesquiterpene product. The STS was put in large excess so that it is not rate-limiting. Gd4oIS was chosen as it had been previously fully characterised in the literature with a reported k_{cat} of 0.198 s^{-1} and K_M $0.115 \text{ }\mu\text{M}$. Gd4oIS was overexpressed from a plasmid in

the Allemann library and purified as reported in the literature (section 6.1 Materials and methods).^[72]



Scheme 2-2 Coupled IPK-sesquiterpene synthase assay to test for FP activity of variants

The reaction was carried out overnight at 37 °C with shaking (350 rpm) in 500 μ L buffer containing ATP (5 mM), FP (2 mM), $MgCl_2$ (20 mM), IPK variant (0.5 μ M), Gd4olS (10 μ M). The aqueous solution was extracted with pentane and the organic layer was analysed by GC-MS. As expected, when performing the assay, addition of FP in the incubation buffer lead to precipitation of a FP- Mg^{2+} complex as seen in the reaction of farnesol with UK. Despite this, FDP was produced because sesquiterpene product was detected by GC-MS. (see Figure 2-11) The total ion chromatogram showed a single product peak with a molecular ion of $m/z = 222$ corresponding to $[M(C_{15}H_{26}O)]^+$ and major fragments $m/z = 207$ ($[M - CH_3]^+$), 204 ($[M - H_2O]^+$), 161 ($[M - (H_2O + C_3H_7)]^+$), 81. (see Figure 2-11) The mass spectrum of the product, germacradien-4-ol (**82**) corresponds well with reported literature.^[125] Upon TLC analysis of the precipitate, FDP was also identified. (see Figure 2-11)

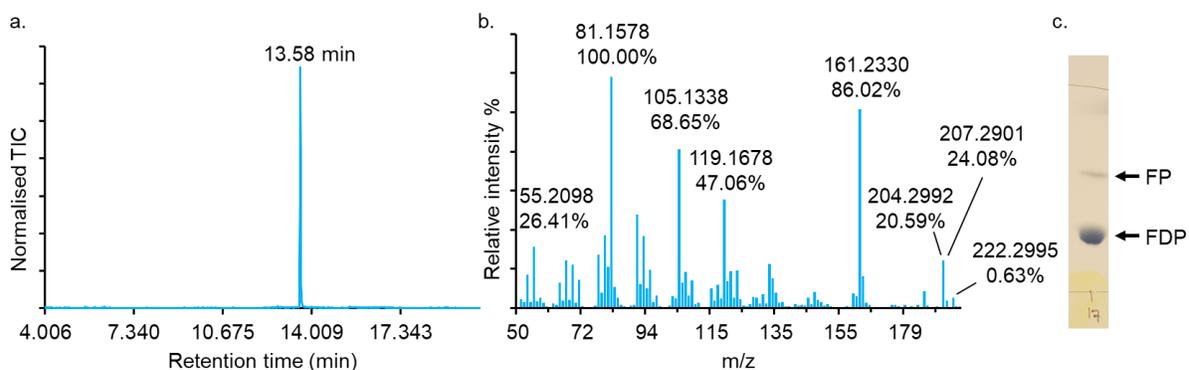


Figure 2-11 GC-MS analysis of the pentane extractable compounds arising from the incubation of FP with IPK variant, Gd4olS and analysis of the white precipitate by TLC a. Total ion chromatogram highlighting the formation of one major product. b. Corresponding mass spectrum of the corresponding compound eluting at 13.58 min. c. TLC of the redissolved precipitate in 1% H_3PO_4 in butanol (TLC solvent system $NH_4OH:iPrOH:H_2O$ 6:3:1)

In principle, the precipitation of polyprenophosphates affords easy isolation and characterisation of the FDP product (**14**). However, in this system, it stops the cascade reaction to the sesquiterpene product. The one pot-type procedure from farnesol to the sesquiterpene product can therefore not be achieved with a high yield due to the combined removal of FP (**130**) and FDP (**14**) complexed with magnesium. Addition of detergent would be detrimental to the overall reaction as tween 20 or glycerol are substrates for UK. Cyclodextrins (CD) have been reported for medicinal purposes to help the solubility of terpenes as they possess a hydrophobic cavity and a hydrophilic surface. (Figure 2-12a)

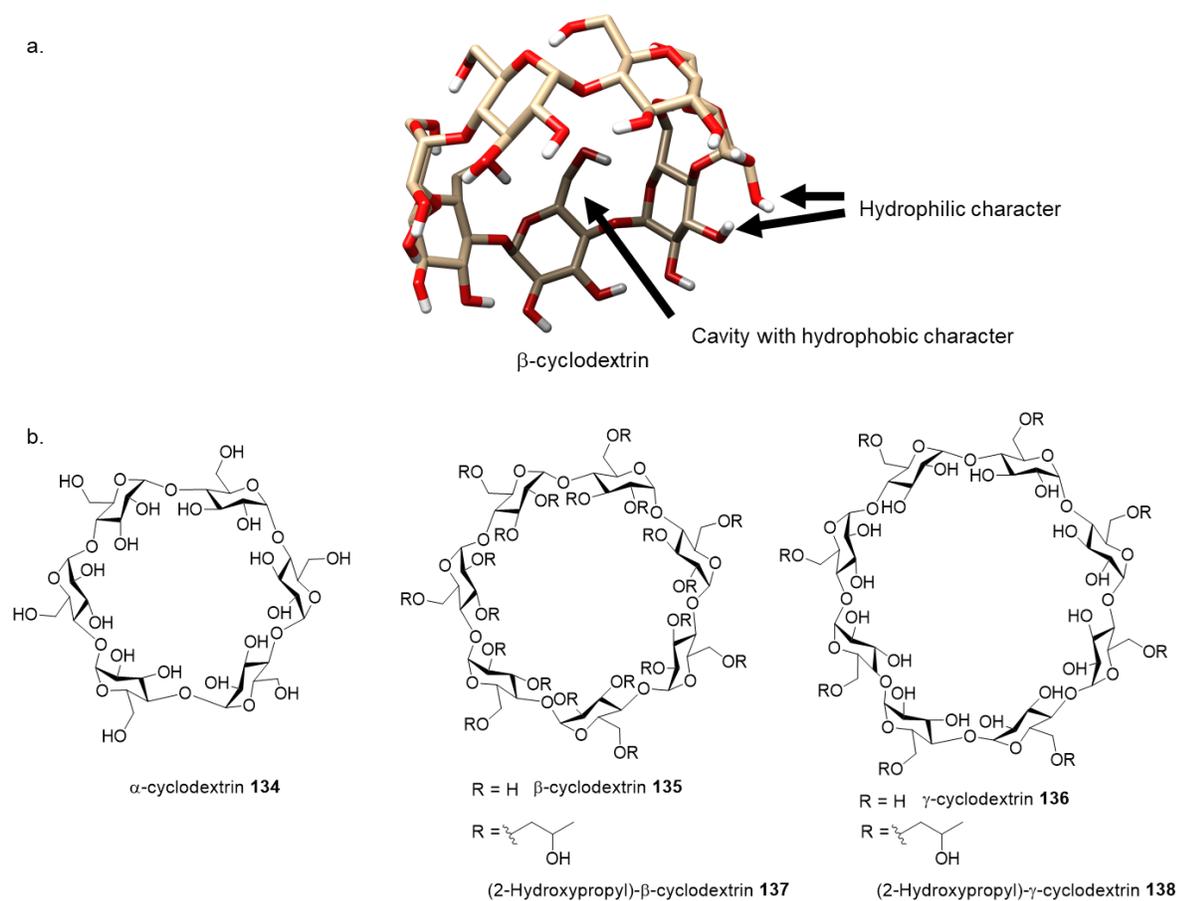


Figure 2-12 Cyclodextrins structures. a. 3D structure of β -CD b. Chemical structure of alpha-cyclodextrin (α -CD), beta-cyclodextrin (β -CD), and gamma-cyclodextrin (γ -CD) and 2-hydroxypropyl- β -CD (HP- β -CD).

FDP could potentially fit in the cavity of cyclodextrins as it contains a hydrophobic tail and prevents precipitation of FP and FDP with Mg^{2+} . Depending on the number of glucose units they are made from, cyclodextrins cavity sizes range from 0.57 nm to 0.95 nm. Several

cyclodextrins were tested with FP and FDP: alpha-cyclodextrin (α -CD, **134**), beta-cyclodextrin (β -CD, **135**), and gamma-cyclodextrin (γ -CD, **136**), 2-hydroxypropyl- β -CD (HP- β -CD, **137**) and 2-hydroxypropyl- γ -CD (HP- γ -CD, **138**). (Figure 2-12b) Substitution by hydroxypropyl group conferred higher hydrophilicity of CD. Surprisingly, Mg-FP and Mg-FDP complexes were not broken in presence of α -CD or γ -CD while showing improvements in β -CD/ HP- β -CD. (Figure 2-13, lane 5 - 14) This can be explained from the difference in the size of each CD, presumably, FP and FDP have a better fit in β -CD forms. Interestingly, the precipitate was not observed when having HP- β -CD and FP together while β -CD was observed to be better to keep FDP in solution (Figure 2-13, lane 8 and 11). When added together in equimolar quantities in buffer, no precipitate was observed for 10 consecutive hours. (Figure 2-13, lane 15) Further experiments were therefore carried out in presence of β -CD and HP- β -CD.

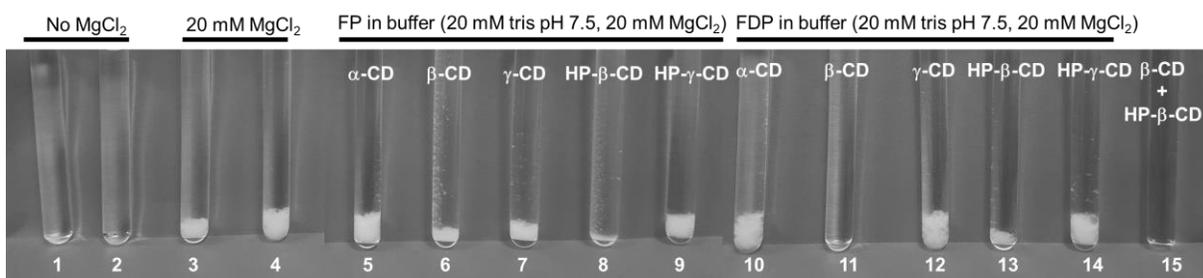


Figure 2-13 Solubility test of FP and FDP at 5 mM. 1&3) FP in buffer. 2&4) FDP in buffer. 5-10) FP in buffer and various CD (5 mM). 11-14) FDP in buffer and various CD (5 mM). 15) FP + FDP in buffer containing $MgCl_2$ (20 mM) and β -CD (5 mM) HP- β -CD.

The coupled assay designed in Scheme 2-2 does not allow direct quantification of the % conversion of FP to the sesquiterpene product through GC because unreacted FP will remain in the aqueous solution. Directly measuring the reaction yield through quantification by GC FID is a possibility but due to the small scale of the reaction (total volume 200 μ L) and the reproducibility of the results could be challenging. Another subsequent step was therefore added to hydrolyse the unreacted FP. (see Figure 2-14) Alkaline phosphatase (10 $mg \cdot mL^{-1}$) together with $MgCl_2$ (100 mM) was added after the 20-hour incubation and left at 37 $^{\circ}C$ with shaking (250 rpm). After 2 h, the solution was extracted with heptane (500 μ L) and the organic layer was analysed by GC-MS. This allowed direct comparison of each variant by calculating the ratio between farnesol and sesquiterpene product. Control experiments without a variant of IPK, without Gd4oIS (**82**) or without both enzymes were performed and resulted in only farnesol being observed in the heptane extractable fraction.

Chapter 2. Development of an efficient diphosphorylation process

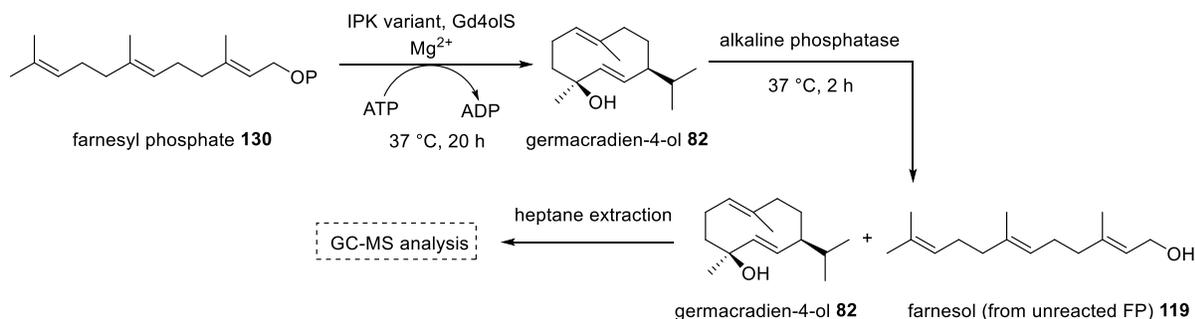


Figure 2-14 Coupled IPK-sesquiterpene synthase-alkaline phosphatase assay to test for FP activity of variants.

An example of the chromatogram obtained with IPK F83A I86A I156A (IPK variant 5) is displayed in Figure 2-15. Two peaks were observed eluting at 13.58 min and 15.44 min. Gd4ol (**82**) was identified above in Figure 2-11. The second compound was identified as farnesol through its mass spectrum. (see Figure 2-15) This assay was performed on all variants and wild type IPK in triplicates and results are shown in Figure 2-16.

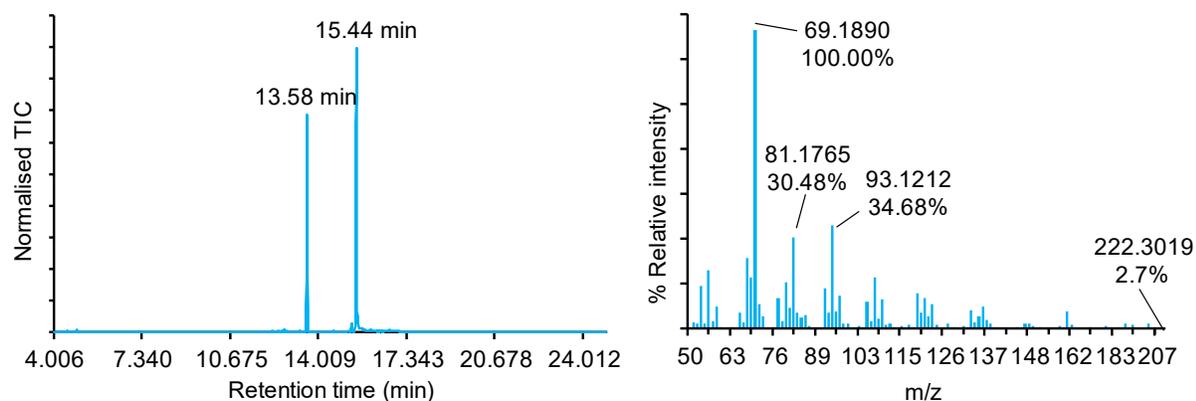


Figure 2-15 GC-MS analysis of the heptane extractable compounds arising from one of the coupled assay IPK-Gd4olS-AP. a. Total ion chromatogram highlighting the formation of two products with a retention time of 13.58 min (gd4ol) and 15.44 min (farnesol). b. Corresponding mass spectrum of the corresponding compound eluting at 15.44 min, characteristic of farnesol.

Results suggest that mutation on the top of the α B helix (I86A) has little effect on the activity towards FP (or even negative, IPK 3 versus IPK7), while a mutation on the β 9- β 10 hairpin (I146 or I156) is necessary for accommodating FP (IPK 1, 2 and 3). The mutation F76A on the β 3- β 4 hairpin appears to be critical for a high activity as the variants not containing this mutation do not perform as well (IPK 4 and 5 versus IPK 3).

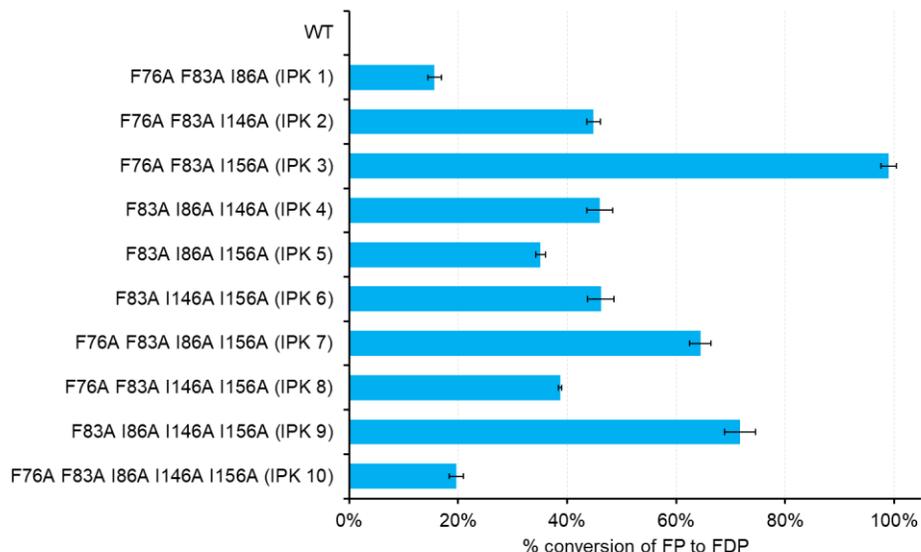


Figure 2-16 Results from the coupled assay IPK-GdoIS-AP for FP activity with the 10 IPK variants.

It was also found that increasing the number of mutations was detrimental to the activity of the enzyme; adding a fourth mutation to the highly potent variant 3 (> 99% conversion of FP to FDP) generated variants 7 and 8 that gave decreased FP conversion, showing only 64% and 39% conversion respectively. These kind of results are expected as the substitution of amino acids (even with conservative mutations) in an enzyme's catalytic site often have a profound impact on enzyme activity due to several reasons such as misfolding or decrease of affinity of the protein to the substrate. The most potent variant was found to be IPK 3 and contained I156A, F76A and F83A mutations. Phenylalanine is a bulky hydrophilic residue and when replaced with a small hydrophilic amino acid it predictably resulted in extremely good activity with > 99% conversion of FP to FDP. For further studies, variant 3 (IPK F76A, F83A, I156A) was selected as it had demonstrated high activity towards FP.

Kinetic characterisation of IPK variants was not performed because no assay (fluorescence, absorbance, or NMR spectroscopy) was reliable due to FP/FDP precipitation. The use of high cyclodextrins concentrations to prevent precipitation also proved to inhibit the reaction leading to inaccurate results.

2.4. ENZYMATIC SYNTHESIS OF FDP FROM FARNESOL

After showing that UK was able to turn farnesol (**119**) to FP (**130**) and re-engineering IPK to convert > 99% FP to FDP, both enzymes were combined to create the first *in vitro* enzymatic route to FDP (**14**) from farnesol (**119**). Both kinases are ATP-dependent enzymes. In previous

studies, ATP was used in equimolar amounts, however, due to its high cost, ATP (**132**) should be used in catalytic amounts and regenerated *in situ* using an enzyme catalysed reaction. Several enzymatic systems have been described in the past for the regeneration of ATP from ADP (**139**) *in situ*. Most used systems are acetyl phosphate/acetate kinase (AcP/AcK) and phosphoenolpyruvate/pyruvate kinase (PEP/PK).^[126] In this work, PEP/PK regeneration system was chosen as it was readily available in the Allemann laboratory. (see Figure 2-17) To efficiently screen optimum conditions for the synthesis of FDP, UK and IPK were coupled with a sesquiterpene. In this case (*S*)-germacrene D synthase was used as a stock that had been prepared fresh. (see Figure 2-17)

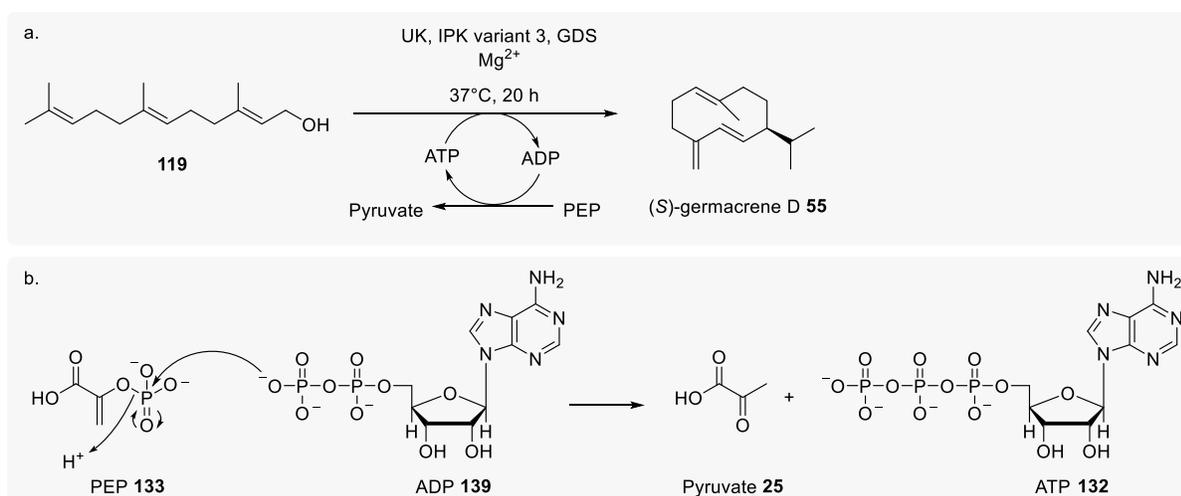


Figure 2-17 a. One pot enzymatic synthesis of (*S*)-germacrene D from farnesol catalysed by UK, IPK 3 and GDS. b. Recycling scheme of ATP with PEP.

Through comparison of reported k_{cat}/K_M value of the two kinases, IPK might have better catalytic efficiency for its natural substrate compared to UK for FP (reported values are 0.34 $\mu\text{M}\cdot\text{s}^{-1}$ and 0.17 $\mu\text{M}\cdot\text{s}^{-1}$ respectively). A few reported *Tha* IPK variants exhibited similar catalytic efficiencies for GP to the wild type *Mj* IPK enzyme for IP.^[119] Therefore, the first assays were carried out with equimolar quantities of kinases (2 μM). Unfortunately, the pentane extractable fraction was composed of more than 80% farnesol (see Figure 2-19 lane 1). Such results can be explained as heavy precipitation could be observed after 5 min incubation at 37 °C despite CD being present. The solid was identified to be FP; suggesting accumulation of a high concentration of FP in the solution due to a slower catalytic rate of IPK **variant 3** compared to UK. This likely led to the previously observed FP-Mg²⁺ complex. The concentration of farnesol was reduced from 5 mM to 2 mM, leading to a slight improvement

of farnesol (**119**) conversion to (S)-germacrene D (**55**) (see Figure 2-19a lane 2). Unfortunately, a small amount of precipitate was still observed, therefore cyclodextrin concentrations were increased to obtain a concentration twice as high as the starting material. Interestingly no precipitate was observed but lower conversion was observed. (see Figure 2-19 lane 3) This demonstrates again that entrapment of FP/FDP in CD has a significant impact on the kinases reaction rates. While cyclodextrins allowed the reaction to process through maintaining FP and FDP in solution, a too high concentration slows down significantly the conversion of FP to FDP. To overcome this key issue, the ratio between UK and IPK 3 was modified to 1:5 and 1:10 ratio. (see Figure 2-19 lane 5 and 4 respectively) As expected, the proportion of sesquiterpene product increased, with conversion of 95% of farnesol to (S)-germacrene D when using 1:10 ratio (UK:IPK, 0.2 μ M:2 μ M). (see Figure 2-19 lane 4 and Figure 2-19) Those results hint that IPK 3 exhibits slower catalytic efficiency than UK. However, a value could not be measured in the kinetics assays as heavy precipitation hindered the reliability of fluorescence and absorbance measures. An assay involving NMR spectroscopy was also attempted but same issue was encountered. However, slower catalytic rates can be reasonably expected when mutating the active site of an enzyme.

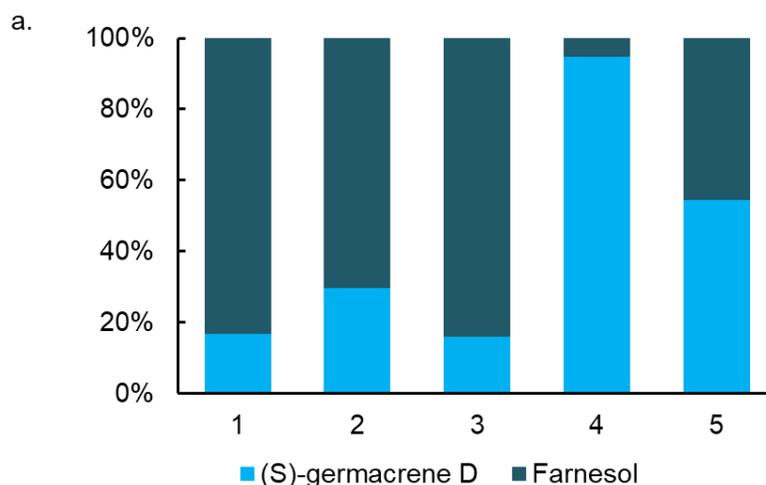


Figure 2-18 Optimisation of the one pot enzymatic synthesis of (S)-germacrene D from FDP. Lane 1: UK (2 μ M), IPK 3 (2 μ M), $MgCl_2$ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), GDS (40 μ M), farnesol (5 mM), 37 °C, 20 hours. Lane 2: same as lane 1 with farnesol (2 mM), β -CD (2 mM), HP- β -CD (2 mM). Lane 3: same as lane 2 with, β -CD (4 mM), HP- β -CD (4 mM). Lane 4: same as lane 2 with UK (0.4 μ M), IPK 3 (2 μ M). Lane 5: same as lane 2 with UK (0.2 μ M), IPK 3 (2 μ M).

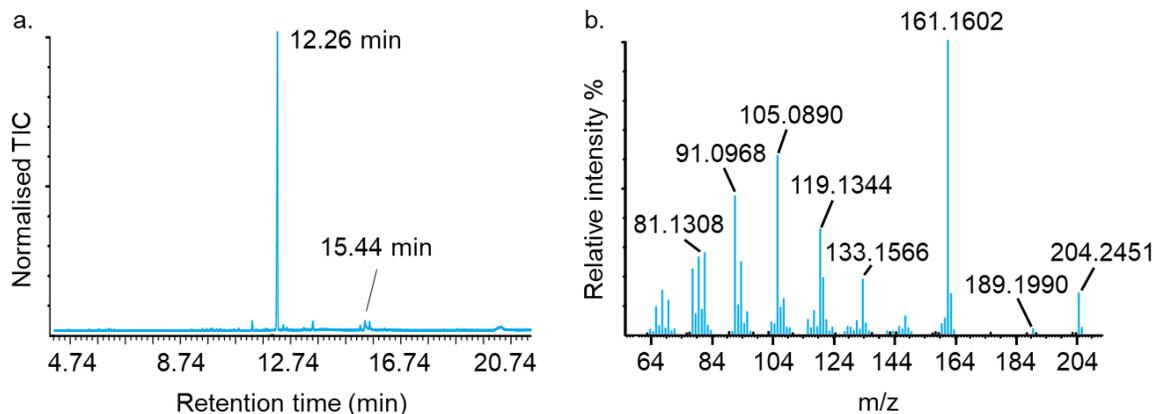
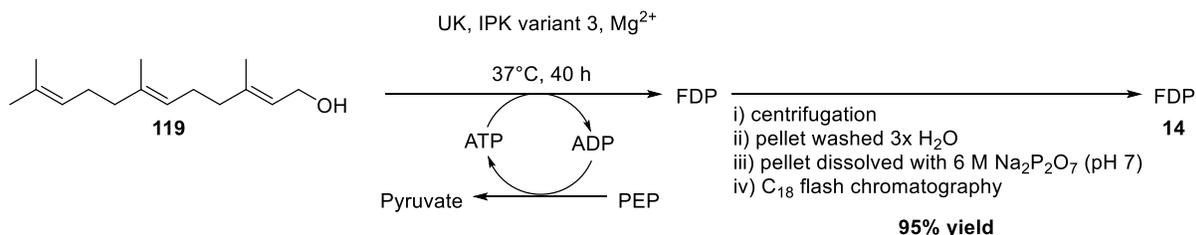


Figure 2-19 a. Total ion chromatogram of the pentane extractable compounds from the one pot enzymatic synthesis of (*S*)-germacrene D from farnesol. (*S*)-germacrene D eluting at 12.26 min and farnesol eluting at 15.44 min. Conditions: UK (0.2 μ M), IPK 3 (2 μ M), $MgCl_2$ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), GDS (40 μ M), farnesol (2 mM), 37 $^{\circ}$ C, 20 hours b. Mass spectrum of compound eluting at 12.26 min corresponding to (*S*)-germacrene D.

Optimised conditions were used to evaluate the efficiency of this novel enzymatic diphosphorylation method from farnesol. HP- β -Cyclodextrin (**137**) only was used to maintain FP in solution and allowed FDP to precipitate for easy and rapid isolation. Satisfactory 95% yield was achieved from farnesol in a 0.1 mmol scale.



Scheme 2-3 One pot enzymatic diphosphorylation scheme of farnesol.

Next, several FDP analogues were synthesised to test the promiscuity of the two kinases. Dimethylated analogues were chosen such as 14,15-dimethylfarnesol (**140**), precursor to (*S*)-14,15-dimethylgermacrene D (**86**), a novel potent semiochemical against aphids, a worldwide pest. 14,15-farnesol (**140**) was synthesised following a previously reported synthetic route by a Master' student (Clotilde Phillipe) who was supervised by the author of this work for 5 months.^[77] 6,15-dimethylfarnesol (**141**) synthesis is described in chapter 3. Oxygenated analogues were also tested as they are bulkier analogues. 12-methoxy farnesol

(**142**), 8-methoxy farnesol (**143**) and 10,11-epoxyfarnesol (**144**) were synthesised as previously described and synthesis route is described in chapter 3.^[78]

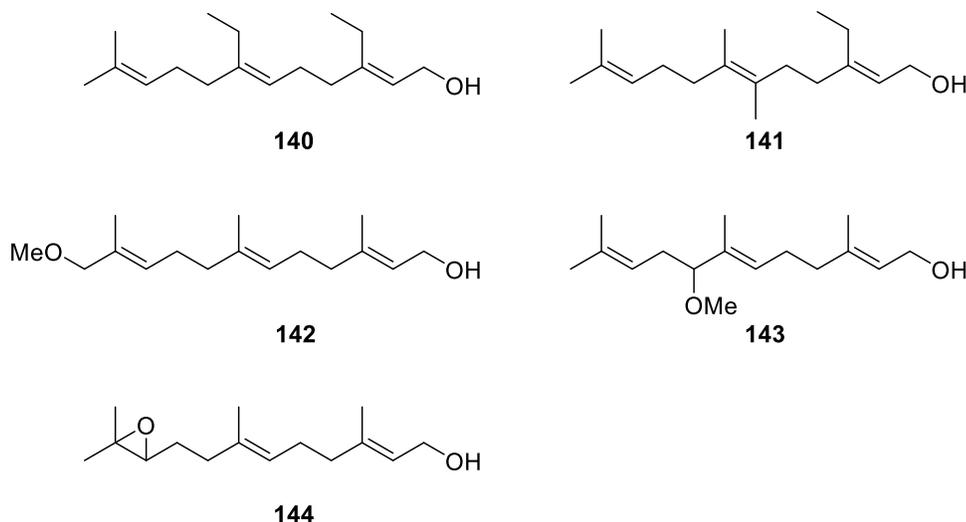


Figure 2-20 Analogues of FDP tested with the design one pot enzymatic diphosphorylation.

Enzymatic diphosphorylations were coupled with sesquiterpene synthases readily accepting the FDP analogues for a fast assessment of this new procedure with the exception of 6,15-dimethylfarnesol (**141**). GDS Y406F was used for 14,15-dimethylfarnesol (**140**); ADS for 12-methoxy farnesol (**142**) and 8-methoxy farnesol (**143**) and Gd4oIS for 10,11-epoxyfarnesol (**144**). All synthases were expressed as described in section 6.1 Materials and methods.

6,15-dimethyl FDP (**145**) was successfully isolated in 89% yield using the procedure in Scheme 2-3. 10,11-epoxyfarnesol (**144**) was found to be a good substrate for UK and IPK 3 because full conversion was observed in the total ion chromatogram (see Figure 2-21). The major product is a cyclic ether (**146**) eluting at 13.28 min and a full characterisation is given in Chapter 3. 14,15-dimethylfarnesol showed lower conversion. 14,15-dimethylgermacrene D (**86**) eluted at 15.62 min but starting material could be seen at 17.33 min (see Figure 2-22a). Heavy precipitation was observed in the reaction vessel before adding alkaline phosphatase. Upon analysis of the solid by TLC, 14,15-FDP was identified as the major compound as it eluted similarly to FDP. (see Figure 2-22b) It is possible that GDS Y406F was rate-limiting in this one pot synthesis. This would lead to an accumulation of 14,15-dimethylFDP in the buffer and results in its precipitation with magnesium as observed.

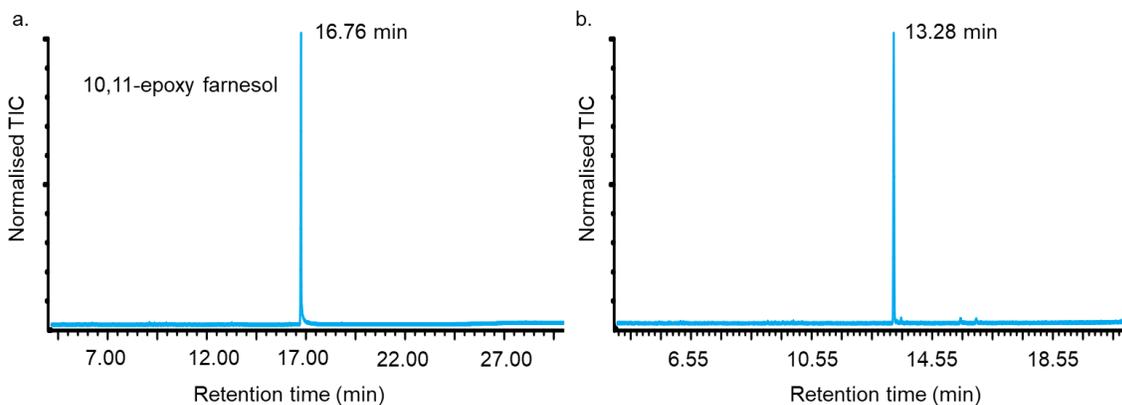


Figure 2-21 a. Total ion chromatogram of 10,11-epoxy-farnesol b. Total ion chromatogram of heptane extractable compounds from the incubation of 10,11-epoxyfarnesol with IPK 3, UK, Gd4oIS and subsequently AP.

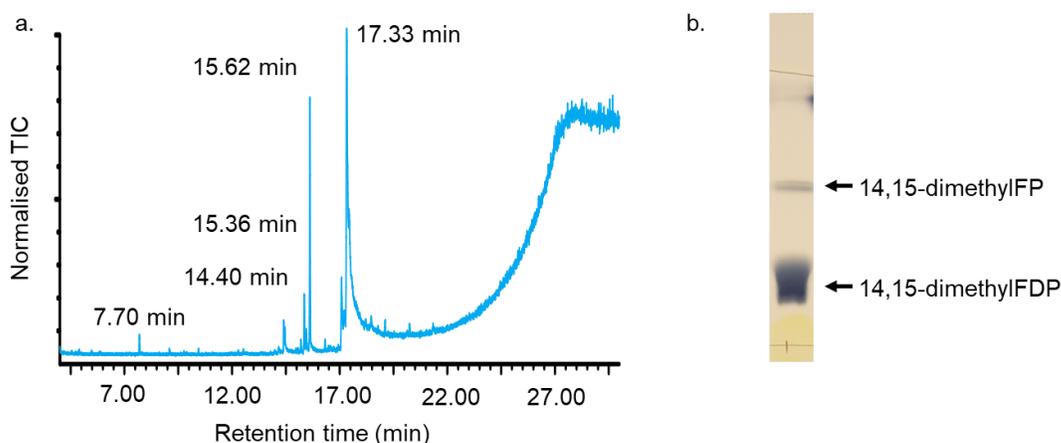


Figure 2-22 a. Total ion chromatogram of heptane extractable compounds from the incubation of 14,15-dimethylfarnesol with IPK 3, UK, Gd4oIS and subsequently AP. 14,15-dimethylfarnesol eluting at 17.33 min. 14,15-dimethylgermacrene D eluting at 15.62 min. b. TLC of the white precipitate formed at the bottom of the incubation of 14,15-dimethylfarnesol with IPK 3, UK, Gd4oIS.

Analogues containing methoxy functional group were found to not perform as well as FDP (**14**). Both analogues were coupled with ADS because Demiray *et al.* reported the 8-methoxy FDP (**87**) and 12-methoxy FDP (**89**) to be substrates of ADS. ADS generates the 8-methoxy- γ -humulene (**88**) as major product from 8-methoxy FDP (**87**) while 12-methoxy FDP (**89**) is converted to 12-methoxy- β -sesquiphellandrene (**90**) and 12-methoxy-zingiberene (**91**) by ADS. Incubation with 8-methoxy farnesol (**143**) showed low turnover to 8-methoxy- γ -

humulene (**88**) when using IPK 3. When using IPK 9 (IPK F83A I86A I146A I156A), 8-methoxy- γ -humulene (**88**) was found to be the major heptane extractable compound (see Figure 2-23) although some starting material was still observed. On the other hand, incubation of 12-methoxy farnesol (**142**) resulted in only starting material being observed in the organic extractable compounds after incubation (see Figure 2-24). This can be explained by the bulkiness of this additional group at the end of farnesol tail. It is possible that the active site of IPK needs to be further modified to accept substrates longer than the C15 farnesol.

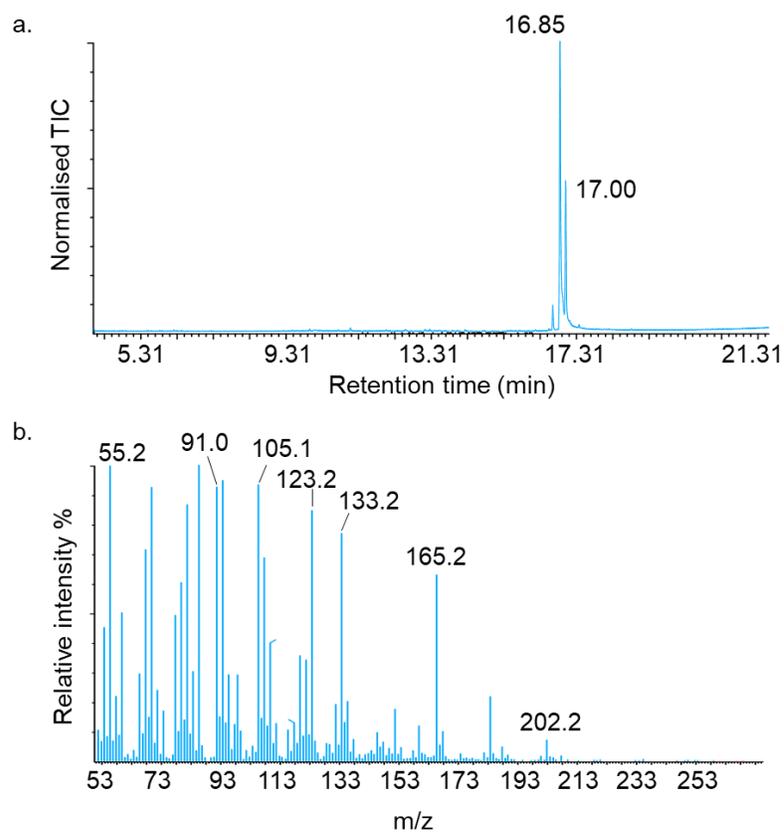


Figure 2-23 a. Total ion chromatogram of heptane extractable compounds from the incubation of 8-methoxy-farnesol with IPK 3, UK, ADS and subsequently AP. 8-methoxy- γ -humulene eluting at 16.85 min. 8-methoxy-farnesol eluting at 17.00 min. b. Mass spectrum of compound eluting at 16.85 min corresponding to 8-methoxy- γ -humulene. The mass spectrum of the product corresponds well with reported literature.^[78]

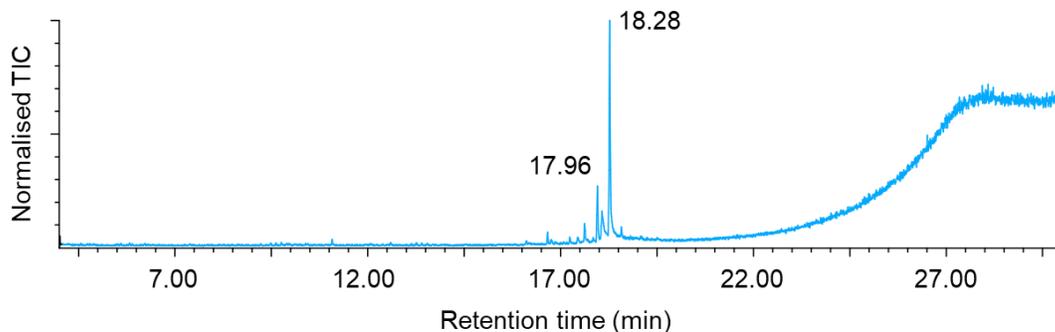


Figure 2-24 Total ion chromatogram of heptane extractable compounds from the incubation of 12-methoxy-farnesol with IPK 3, UK, ADS and subsequently AP. 12-methoxy-farnesol eluting at 18.28 min.

2.5. SUMMARY

This chapter describes the investigation of an alternative chemical pathway to the traditional Davisson *et al.* procedure for diphosphorylation of farnesyl bromide. Despite greatly shortening the length of the diphosphorylation procedure as well as drastically reducing the labour required, the method proved to not be highly reproducible. In parallel, a novel *in vitro* enzymatic synthesis route was developed for a one pot diphosphorylation of farnesol (**119**) using two kinases: undecaprenol kinase (UK) in its wild type form with an engineered isopentenyl phosphate kinase (IPK). UK was successfully expressed in *E. coli*, a total of eighteen variants of IPK containing one to five mutation in the active site were created through site-directed mutagenesis. Out of the eighteen variants, ten (containing three to five mutations) were expressed in *E. coli*. An assay involving a sesquiterpenes synthase (GDS or GdoS) was designed to access IPK variants' performance as well as optimising the conditions of the one pot diphosphorylation. This *in vitro* method can be used to diphosphorylate farnesol and allow easy isolation as FDP (**14**) precipitates upon formation. This novel enzymatic route can also be directly coupled with a sesquiterpene synthase to generate in a single reaction a high value sesquiterpene product from commercial farnesol (£1 /g). This is also a useful method to rapidly test the activity of sesquiterpene synthase variants produced for mechanistic investigations or when looking for novel sesquiterpenoids. Finally, this method was successfully applied to few farnesol analogues 10,11-epoxy farnesol (**144**), 6,15-dimethylfarnesol (**141**), 14,15-dimethylfarnesol (**140**), 8-methoxyfarnesol (**143**), 12-methoxyfarnesol (**142**). This shows that this enzymatic procedure coupled with a sesquiterpene synthase can rapidly demonstrates if a chemically synthesised farnesol analogue is a substrate for a cyclase. Up until the time of writing, after synthesising a farnesol analogue (usually over 10 steps, overall yield <10%), the procedure from Davisson *et al.* was

used to generate the corresponding FDP analogue. Because this last chemical step gives an inconsistent and often low yield, researchers would invariably lose over 50% of the starting material and were obliged to synthesise at least 0.5 g for biological testing. With this one pot enzymatic procedure, less than 1 mg is required to test for substrate acceptance by the sesquiterpene cyclase. Further optimisation of the IPK active site needs to be performed but this enzymatic route can be readily used in research laboratory scale for sesquiterpene synthase variants screening.

CHAPTER 3

EXPANDING THE TERPENOME

3.1. PREFACE

In this chapter, the development and synthesis of several FDP analogues will be described. These analogues were incubated with sesquiterpene synthases in an attempt to expand the library of sesquiterpenes. Firstly, analogues of FDP were synthesised to mimic the natural substrate to generate a second generation of semiochemicals based on the discovery of the aphid attractant: 14,15-dimethylgermacrene D (**86**). Then an analogue of FDP was synthesised with the aim of prematurely terminating the normal carbocationic cascade reaction of Gdols hence generating a new sesquiterpene product.

3.2. DIMETHYLATED ANALOGUES FOR NOVEL SEMIOCHEMICALS

Recently, collaborators from the Allemann group at Rothamsted Research have identified (*S*)-germacrene D (**55**) as a potent semiochemical that acts on major world crop pests such as aphids and other arthropods.^[11,127] However, due to the chemical instability of **55**, new semiochemicals have been investigated to overcome this issue.^[128] It has been hypothesised that if the enzyme (*S*)-GDS was able to convert a modified FDP to its analogue product germacrene D, it will probably be recognised by the aphid's receptor and act as a semiochemical. This theory has been explored by Touchet *et al.* A series of fluorinated and methylated analogues of FDP were synthesised and incubated them with (*S*)-GDS.^[129] Not all analogues were accepted to the same level by the wild type enzyme. To overcome this issue site-directed mutagenesis was used and several non-natural substrates were successfully converted by (*S*)-GDS Y406F variant. (*S*)-12-Methylgermacrene D (**84**) and (*S*)-15-methylgermacrene D (**147**) were found to act as repellents of *Sitobion avenae* by electrophysiological and behavioural assay. (*S*)-14,15-Dimethylgermacrene D (**86**) on the other hand was found to be a powerful attractant (see Figure 3-1).^[129] These findings supported the idea of rationally designing new aphid semiochemicals based on their acceptance by sesquiterpene synthase naturally biosynthesising semiochemicals. The surprising inverse activity of the dimethylated analogue opened a new research area in the development of dimethylated germacrenes in the search of new olfactory ligands with such attractant activity.

In this chapter, the biosynthesis of two closely related analogues of (*S*)-14,15-dimethylgermacrene D was therefore attempted (see Figure 3-1). To do so, dimethylated FDP

analogues: (6,15)-dimethyl FDP (**145**) and (12,15)-dimethyl FDP (**148**) were chosen as substrate for (*S*)-germacrene D synthase. These two analogues were chosen to demonstrate whether methylation on C15 and another methyl group might be the ground of the unexpected reversal bioactivity. Moreover, substrate with an additional methyl group on C15 or C12 were already observed to possess behavioural activity. Another crucial aspect to this method is to also find relatively short chemical synthesis of FDP analogues for an industrial viability prospect. The C6 methylation respect this condition, that is why it was chosen as a second position for methylation. The dimensions of the enzyme active site and its plasticity are still cryptic because no crystal structure has yet been published, these two relatively close analogues to the natural substrate FDP will also allow further characterisation of GDS.

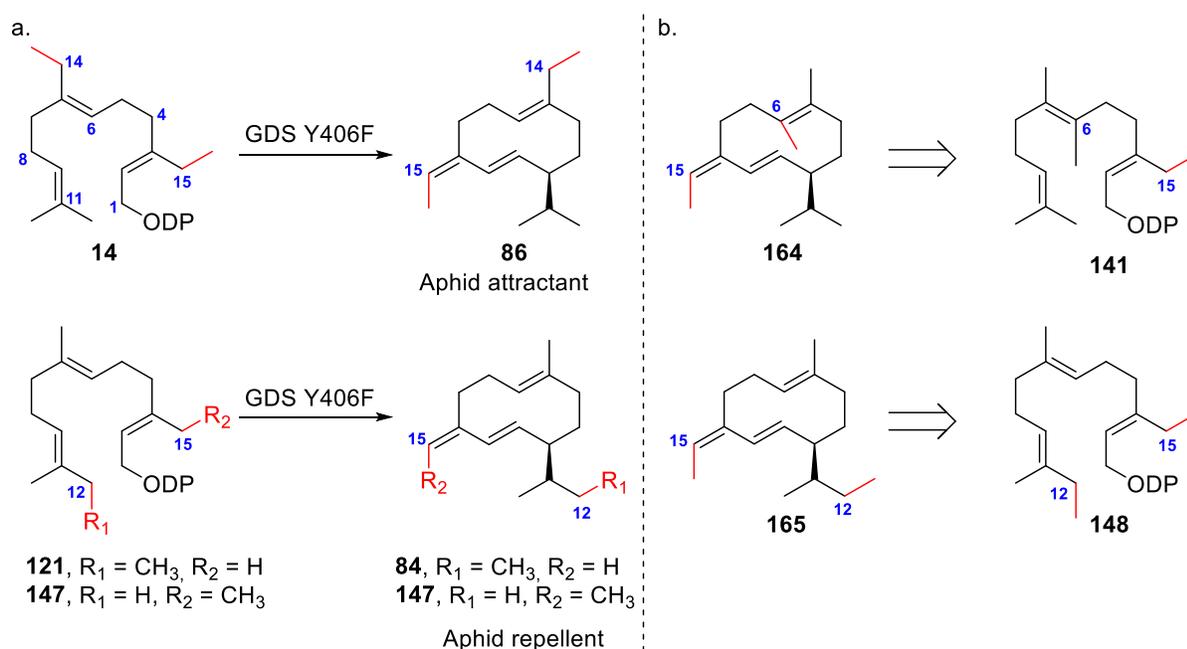
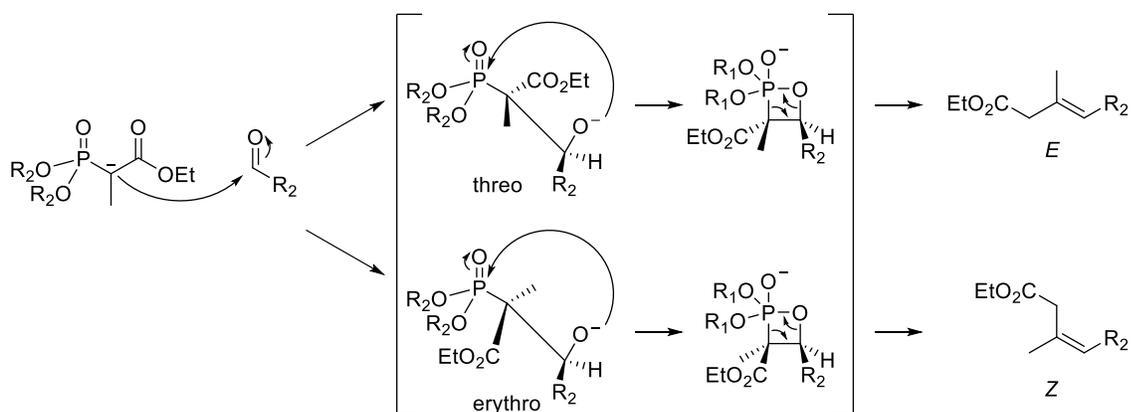


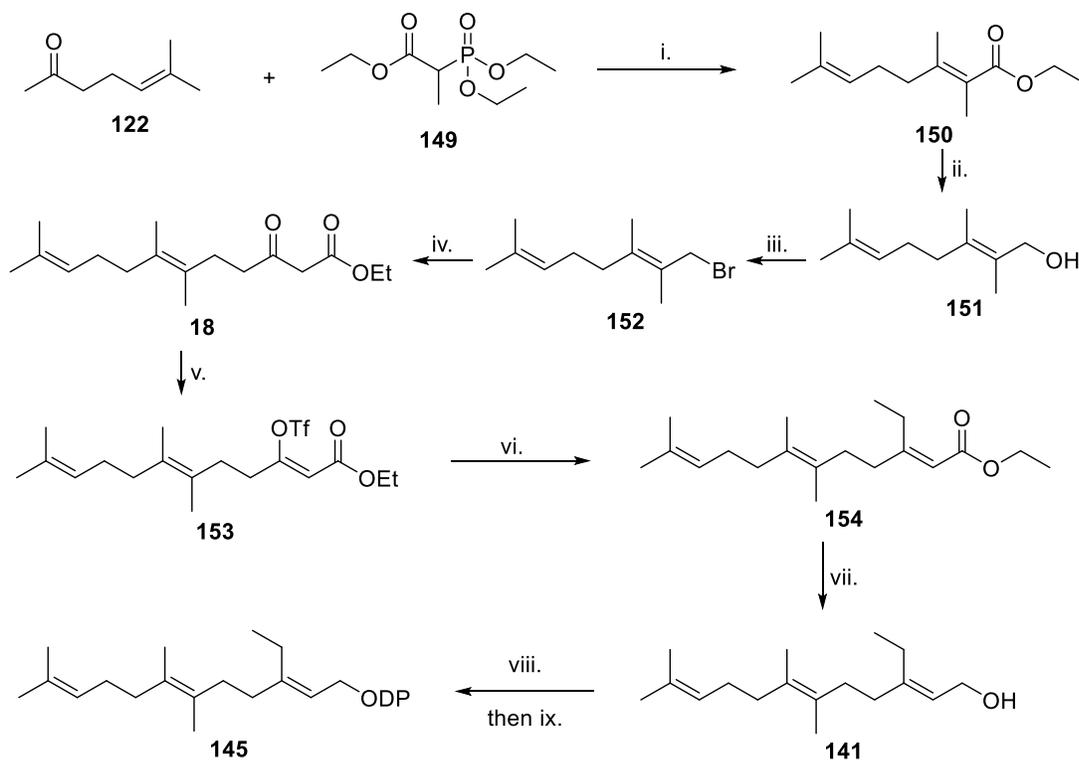
Figure 3-1 Second generation design of (*S*)-dimethylated germacrene D from dimethylated FDP. a. Top: (*S*)-14,15-dimethylgermacrene D was found to be an aphid attractant. Bottom: (*S*)-15-dimethylgermacrene D was found to be an aphid repellent. b. Dimethylated analogues synthesised in this chapter with both C15 methylated.

3.2.1. 6,15-Dimethyl FDP synthesis

(6,15)-Dimethyl FDP (**145**), a novel analogue was prepared via a nine steps synthesis (see Scheme 3-2). Firstly, a Horner-Wadsworth-Emmons reaction was performed on 6-methylhept-5-en-2-one (**122**) with ethyl 2-(diethoxyphosphoryl)propanoate (**149**) to form the α,β -unsaturated ester (**150**) in 67% yield (1:1, *Z:E*). The Horner-Wadsworth-Emmons is a modified version of the Wittig olefination.^[130,131] A phosphonoester substituted with electron-withdrawing group such as ethyl trifluoroethylphosphonopropionate could have been used to favour *Z*-olefin product because such group promote the formation of *Z*-olefin during the last elimination step by stabilising the formation of the erythro intermediate adduct as reported by Still and Gennari.^[132] Due to the explorative goal of this study, it was chosen to avoid supplementary steps in the synthesis. Synthesis of ethyl trifluoroethylphosphonopropionate would add either three more steps (overall yield 84%) to the synthesis or only one with a reported poor yield of 12%.^[133–135] Then reduction of α,β -unsaturated ester (**150**) with diisobutyl aluminium hydride (DIBAL-H) followed by flash chromatography to separate the *E* and *Z* isomer gave the alcohol (**151**) (36%). The latter was then brominated using phosphorus tribromide to form **152** (quantitative) and underwent a chain elongation with a dianion prepared by treating ethyl acetoacetate successively with sodium hydride and *n*-butyl lithium (63%). The resulting compound (**18**) was then transformed in a stereoselective manner to the (*Z*)-enol triflate (**153**) (quantitative) in order to yield **154** by a copper-mediated alkylation using a Grignard reagent (44% over 2 steps). Weiler and Sum have previously reported that this isoprenoid ester can be synthesised by coupling a vinyl phosphate analogue with Me_2CuLi , but 100% conversion was found to be successful only with enol triflate.^[136] Moreover, McMurry and Scott have used these vinyl triflates with a variety of substituents and showed their powerful qualities to give substituted alkenes in a stereo- and regioselective manner.^[137] Ester (**154**) was reduced with DIBAL-H to give the farnesol derivative (**141**) (76%). Finally, a successive bromination and diphosphorylation following Davisson *et al.* procedure yielded the final compound (**145**) which was used as the ammonium salt (33%).^[110]



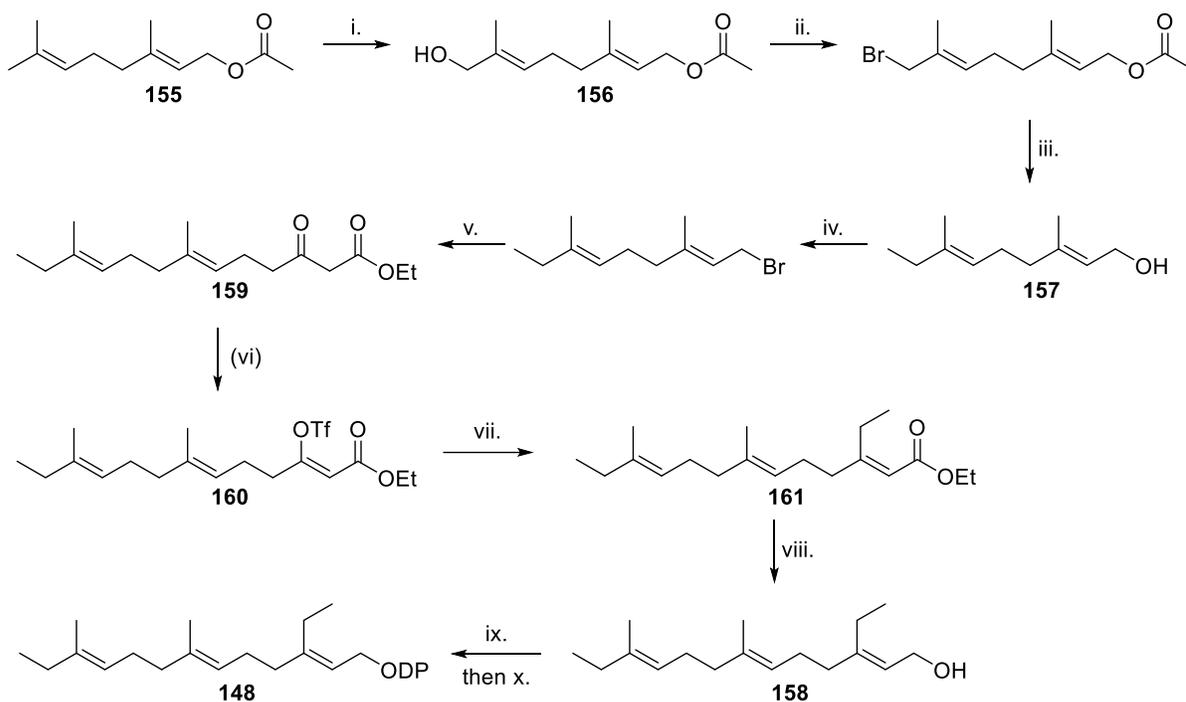
Scheme 3-1 Possible products resulting from a Horner-Wadsworth-Emmons olefination.



Scheme 3-2 Synthesis of 6,15-dimethyl FDP Reagent and conditions i. Triethyl-2-phosphonopropionate (1.1 eq), *n*-BuLi (1.1 eq), THF, 0 °C, 4 h, ii. DIBAL-H (4 eq), Toluene, 78 °C, 3 h iii. PBr₃ (0.5 eq), THF, -10 °C, 30 min. iv. Ethyl acetoacetate (3 eq), NaH (3.5 eq), *n*-BuLi (3.1 eq), THF, 0 °C. v. LiOTf (3 eq), Tf₂O (1.3 eq), NEt₃ (3 eq), DCM, 0 °C, 2 h. vi. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. vii. DIBAL-H (3 eq), Toluene, -78 °C, 1h. viii. PBr₃ (0.5 eq), THF, -10 °C, 30 min. ix. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method^[110] see chapter 2).

3.2.2. 12,15-Dimethyl FDP synthesis

(12,15)-Dimethyl FDP (**148**) was prepared via an 11-step synthesis (see Scheme 3-3). The first step was an allylic oxidation of commercially available geranyl acetate (**155**) using catalytic amount of selenium dioxide, salicylic acid and tert-butyl hydroperoxide.^[138] Tert-butyl hydroperoxide has been used as co-oxidant to reduce toxic by-products and reoxidizes the selenium (II) compounds *in situ*. Relatively average conversion of the starting material was observed when the reaction was left overnight (40%). If left for too long, SeO₂ can generate epoxides and diols hypothetically from perselenious acid species.^[138] The reaction was therefore monitored by TLC until the majority of the starting material converted to either the allylic alcohol or aldehyde.



Scheme 3-3 Synthesis of 12,15-dimethyl FDP Reagent and conditions i. SeO₂ (0.1 eq), Salicylic acid (0.1 eq), tBuOOH (0.1 eq), DCM, 0 °C to R.T., 40 h then NaBH₄ (2 eq), EtOH. ii. PBr₃ (0.5 eq), THF, -10 °C, 30 min. iii. MeMgBr (5 eq), Et₂O, 0 °C, 2h. iv. PBr₃ (0.5 eq), THF, -10 °C, 30 min. v. Ethyl acetoacetate (3 eq), NaH (3.5 eq), nBuLi (3.1 eq), THF, 0 °C. vi. LiOTf (3 eq), Tf₂O (1.3 eq), NEt₃ (3 eq), DCM, 0 °C, 2 h. vii. CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. viii. DIBAL-H (3 eq), Toluene, -78 °C, 1h. ix. PBr₃ (0.5 eq), THF, -10 °C, 30 min. x. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method^[110] see chapter 2).

The crude mixture was then subjected to treatment with sodium borohydride to yield the alcohol (**156**) (80%). The compound **156** was then brominated using phosphorus tribromide followed by a methylation of C12 using methylmagnesium bromide as well as deprotection of the alcohol on C1 in one pot to yield the alcohol **157** (29% over 2 steps). To form the final compound **159**, a chain elongation was necessary. Compound **157** was brominated using phosphorus tribromide and treated with the ethyl acetoacetate dianion as previously described above to yield the α,β -ketoester **159** (52%). The resulting product was converted to the enol triflate **160** to allow the conjugate addition of an additional ethyl group on the α -isoprene unit as described above (61% over 2 steps). Ester **161** was reduced using DIBAL-H (55%) to **158**, followed by bromination using phosphorus tribromide and using Davisson *et al.* procedure to give the final compound and was used as the ammonium salt (**148**) (37%).^[110]

3.2.3. Incubation with GDS and variants

An expression vector containing the gene for GDS (pET21d) was retrieved from the Allemann plasmid library and overexpressed in *E. coli*. BL21 DE3 at 37 °C for 4 hours. The resulting enzyme was purified by nickel affinity chromatography as previously described.^[90] SDS-page can be seen in section 4.2.

Because Touchet *et al.* reported that the 15-methyl FDP (**162**) and 14,15-dimethyl FDP (**163**) were very poor substrate for wild type GDS, it was anticipated that both 6,15-dimethyl FDP (**141**) and 12,15-dimethyl FDP (**148**) would either be poor substrate or simply not turned over by GDS.^[129] Nevertheless, both analogues were incubated with GDS. The hypothesis proved to be correct since both gas chromatograms from the incubations showed no signals (see Figure 3-2).

Since no X-ray single crystal structure of GDS has been published before, a homology model of GDS was generated using I-Tasser^[139] based on 5-*epi* AS from *Nicotiana tabacum* complexed with farnesyl hydroxyphosphonate (FHP) reported crystal structure (PDB-5EAT).^[140] 6,15-dimethyl FDP and 12,15-dimethyl FDP were then docked in the catalytic site in an attempt to identify critical residues hindering substrate folding (see Figure 3-3). Interestingly no immediate steric clashes were observed.

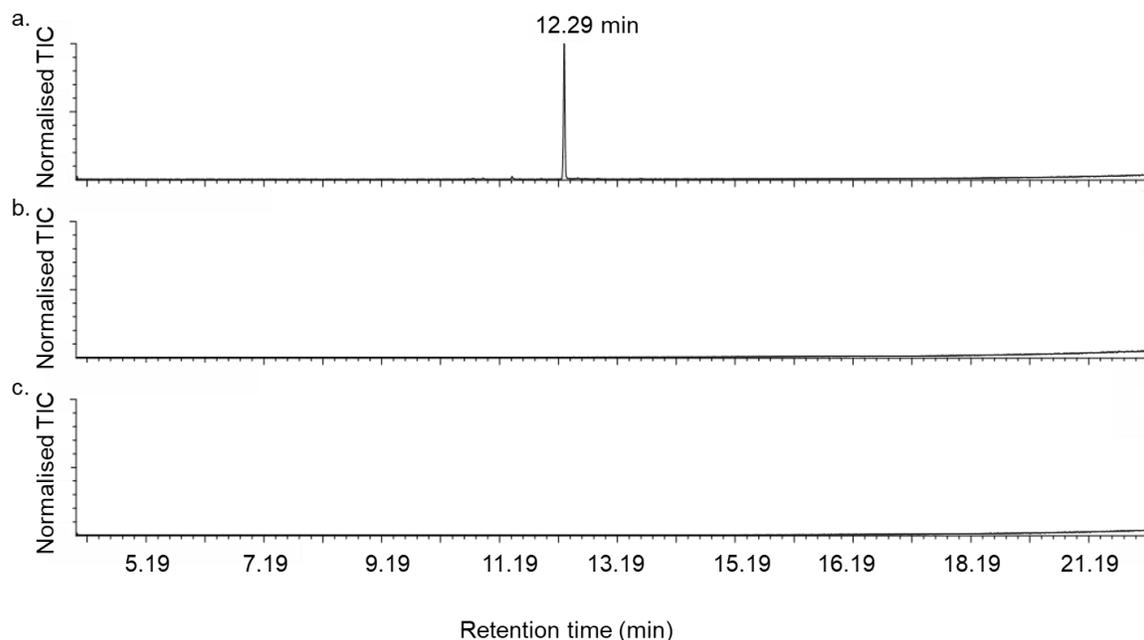


Figure 3-2 Total ion chromatograms of the pentane-extractable products arising from incubation of FDP and analogues with GDS. a. FDP with GDS b. 12,15-dimethyl FDP with GDS c. 6,15-dimethyl FDP with GDS.

Touchet *et al.* reported that GDS Y406F variant was effective in converting 15-methyl and 14,15-dimethyl FDP (**163**) with a higher catalytic efficiency (k_{cat}/K_M) than the wild type version (GDS Y406F: $K_M = 12.75 \mu\text{M}$, $k_{cat} = 0.00853 \text{ s}^{-1}$; $k_{cat}/K_M = 6.69 \text{ M}^{-1} \cdot \text{s}^{-1}$; GDS WT $K_M = 3.6 \mu\text{M}$, $k_{cat} = 0.0094 \text{ s}^{-1}$, $k_{cat}/K_M = 2.6 \text{ M}^{-1} \cdot \text{s}^{-1}$). On the other hand, Y524F variant showed reduced enzyme catalytic activity ($K_M = 5.34 \mu\text{M}$, $k_{cat} = 0.0143 \text{ s}^{-1}$; $k_{cat}/K_M = 2.669 \text{ M}^{-1} \cdot \text{s}^{-1}$), while the W275F variant was found to be less efficient than Y406F variant ($K_M = 2.06 \mu\text{M}$, $k_{cat} = 0.0082 \text{ s}^{-1}$; $k_{cat}/K_M = 3.971 \text{ M}^{-1} \cdot \text{s}^{-1}$).^[140] Touchet *et al.* also reported that 12-methyl FDP was naturally converted by the wild type version of GDS.^[140] This potentially suggest that the Y406F variant could convert 12-15-dimethyl FDP (**148**). In the 6,15-dimethyl analogue (**145**) case, no 6-methyl FDP analogue was reported to be converted by any sesquiterpene synthase. It was therefore chosen to primarily determine if the Y406F variant was turning over these two new dimethylated FDP analogues as it appears that the Y406 residue play a key role for any C15 methylated analogue.

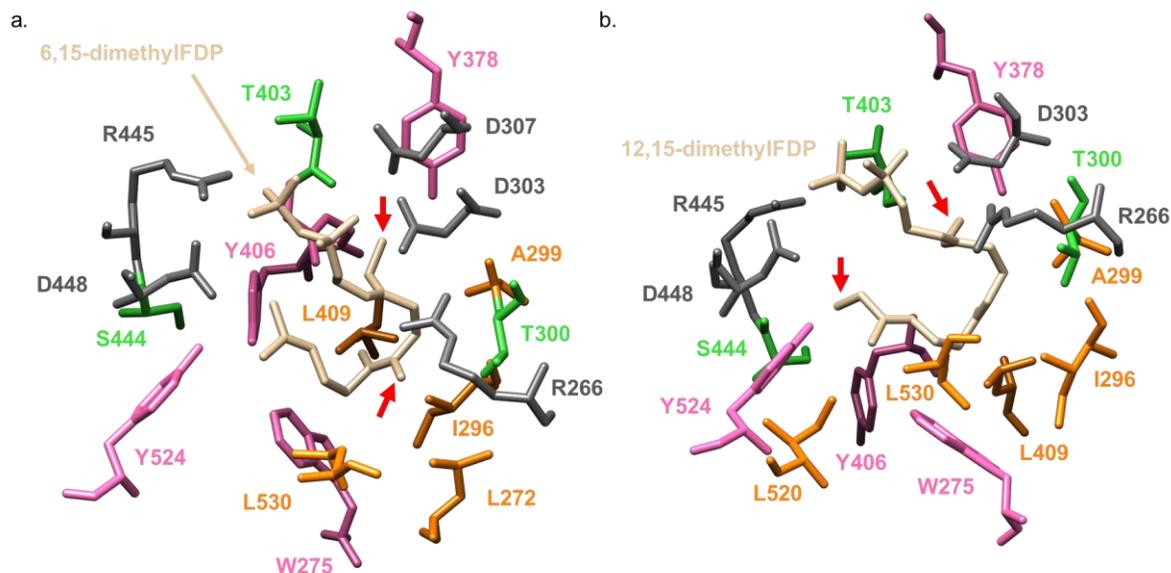


Figure 3-3 Homology models of GDS with dimethylated analogues in the active site based on 5-epi AS from *Nicotiana tabacum* complexed with farnesyl hydroxyphosphonate (FHP) (PDB-5EAT)^[140]. Red arrows indicate the additional unnatural methyl groups.

A range of GDS mutants were present in the Allemann library. They were created using site directed mutagenesis by Dr. Touchet. Therefore, the plasmid containing the GDS Y406F gene was directly expressed and purified as previously described.^[140] Both analogues and FDP as a positive control were incubated with GDS Y406F. Unfortunately, no activity was observed for both analogues. Then a variety of Y406 variants containing changes to smaller residue than phenylalanine to further decrease potential steric hindrance (valine, leucine, and isoleucine), they again showed no activity towards both new analogues. Lastly GDS W275F variant was expressed and incubated with again no activity seen (summary of variants tested in Table 3-1).

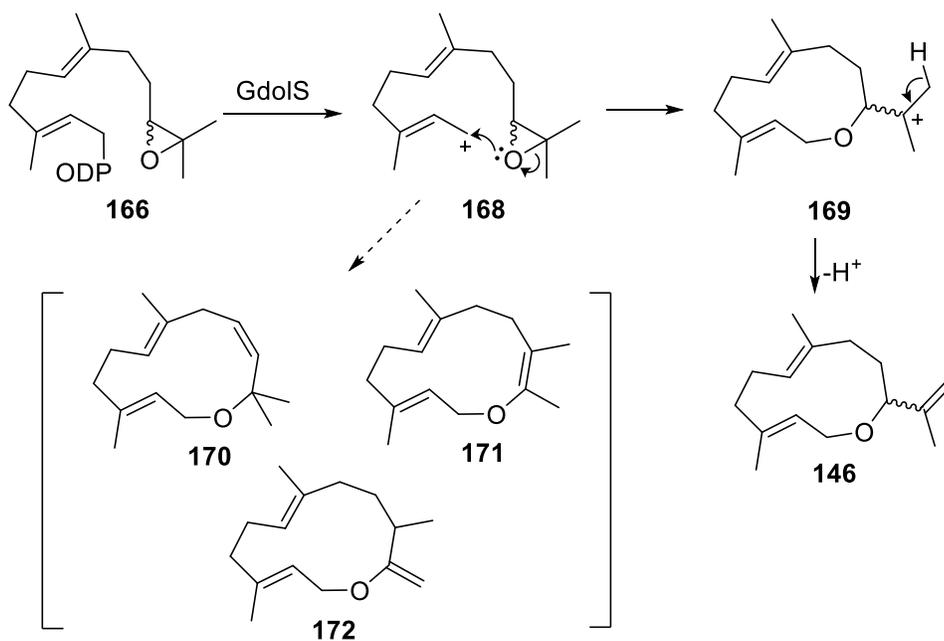
Table 3-1 Summar of all GDS variant tested with 6,15-dimethyl FDP and 12,15-dimethyl FDP. + sign indicates product was generated. – sign indicates no product was generated.

	GDS WT	GDS Y406F	GDS 406V	GDS W275F	GDS Y406I	GDS Y406L
FDP	+	+	+	+	+	+
12,15-dimethyl FDP	-	-	-	-	-	-
6,15-dimethyl FDP	-	-	-	-	-	-

In conclusion, both 6,15-dimethyl FDP (**145**) and 12,15-dimethyl FDP (**148**) were not converted by GDS wild type or various GDS variants. Because these two analogues are not readily accepted by GDS, further active site engineering was not pursued. This non acceptance hints that 6,15-dimethyl germacrene D (**164**) and 12,15-dimethyl germacrene D (**165**) have no semiochemical activity based on the conclusion from Touchet *et al.*^[140]

3.3. TEMPLATING GDOLS FOR THE SYNTHESIS OF A CYCLIC ETHER

Another approach to novel sesquiterpenes is to introduce a substituent or modify the FDP skeleton to divert the mechanism pathway. This has been illustrated in Chapter 1 with for example the destabilisation of intermediate carbocation through strategically placed fluorine in the FDP structure. Previous work in the group by Dr. Daniel J. Grundy explored the use of the 10,11-epoxide FDP (**166**) or 10-hydroxy-11-ene FDP (**167**) as a substrate for germacradien-4-ol synthase from *Streptomyces citricolor* (Gdols) to control the outcome of the cyclisation product.^[129] Gdols is a 1,10-cyclase producing germacradien-4-ol (**82**) from FDP where the final step of the mechanism is the quenching of the carbocation by water (see Chapter 1, section 1.3.3).^[141] Introducing the epoxide as a nucleophile would allow a possible capture of the farnesyl cation (**168**) and a final deprotonation of C12 of **169** would produce **146**.



Scheme 3-4 Proposed reaction mechanism for the Gdols catalysed conversion of epoxide *x* to the medium sized terpene ether **146**.

Alternatively, other products could be formed by cleavage of the C10-O bond followed by nucleophilic attack at C1 that might lead to a 12-membered ethereal carbocation (see Scheme 3-4). Proton loss would give **170**; a [1,2]-methyl shift from C11 followed by proton loss would result in **171** or **172** (see Scheme 3-4). Dr. Grundy showed that the 10,11-epoxide FDP (**166**) was a substrate for GdoIS and germacrene A synthase (GAS).^[142] However, full characterisation of the product was not fully performed. In this section, full characterisation of the product generated by GdoIS from 10,11-epoxide FDP will be performed. To do so, **166** was synthesised and GdoIS expressed in *E. coli* BL21 DE3.

Synthesis of 10,11-epoxy FDP (**166**) can be performed in 4 steps from farnesyl acetate.^[143] Stereospecific epoxidation of the 10,11 double bond using N-bromosuccinimide has been described before.^[144] It generates a bromohydrin which can undergo a base-catalysed ring closure, with these conditions, the acetate group is deprotected at the same time and produces the 10,11-epoxy farnesol (**144**) in 80% yield over 2 steps. The alcohol is finally chlorinated, diphosphorylated using Davison *et al.* procedure and used as the ammonium salt (**166**) (54%) (see Scheme 3-5).^[110]

An expression vector containing the gene for GdoIS (pET16b-SC1) (a gift from Professor Ohnishi, University Tokyo) was overexpressed in *E. coli*. BL21 DE3 at 16 °C for 20 hours. The resulting enzyme was purified as previously described by nickel affinity chromatography. SDS PAGE was used to identify which fractions contained the protein of interest. (see Figure 3-4).^[141]

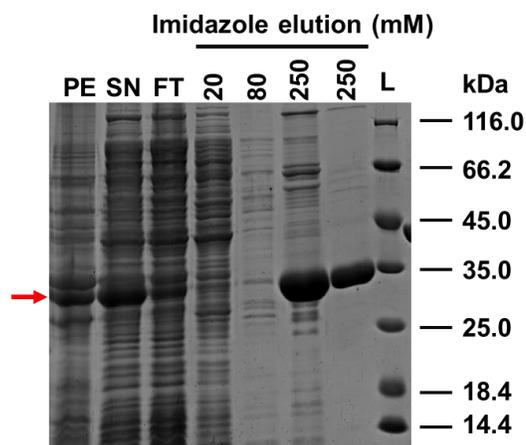
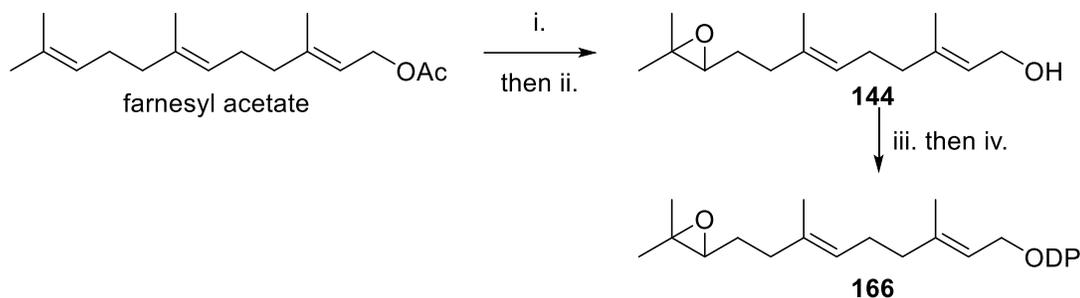


Figure 3-4 Purification of germacradien-4-ol synthase. a. SDS-PAGE: analysis of the supernatant (SN) and pellet (PE) after sonication of the pellet after expression and purification by nickel affinity chromatography. Pierce™ unstained protein MW marker was used as reference of protein size (L).



Scheme 3-5 i. NBS, THF/H₂O ii. K₂CO₃, MeOH. iii. LiCl (4 eq), S-collidine (6 eq), MsCl (2 eq), DMF, 3h, 0 °C iv. OPP (2 eq), MeCN followed by cation exchange (Davisson et al. method^[110] see chapter 2).

Then a preparative scale incubation in batch of **166** (60 mg) with GdoIS (10 μM in 50 mM Tris, 5 mM β-mercaptoethanol, 5 mM MgCl₂, pH 7.5) was performed to give a colourless oil purified by preparative thin layer chromatography (5% ether in pentane, 30% yield) (see Figure 3-5).

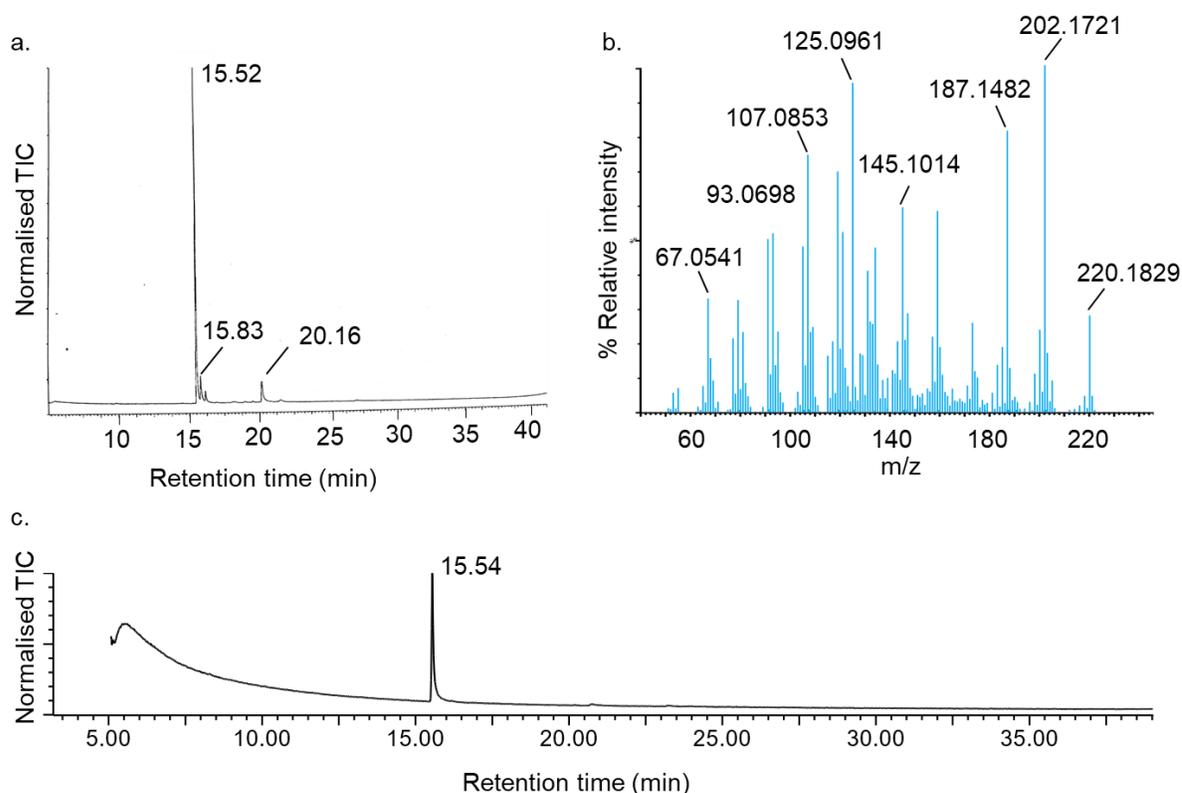


Figure 3-5 Total ion chromatogram of the pentane extractable compounds from the incubation of 10,11-epoxy-FDP with GdoIS. Major compound eluting at 15.52 min, side products at 15.83 min and 20.16 min (compound eluting at 15.83 min identified to be 10,11-epoxyfarnesene by Dr. Grundy, unknown compound at 20.16 min). b. Mass spectrum of compound eluting at 15.52 min. c. Total ion chromatogram of purified oil from the incubation a.

To estimate the efficiency, with which GdoIS turned over the 10,11-epoxy FDP, competitive incubations of **166** and FDP (**14**) with GdoIS were performed and analysed by GC-MS to follow the relative production of germacradien-4-ol (**82**) and **146**. Incubations of 3 μM enzyme with 0.125 mM **166** and 25 μM FDP (**14**) led to similar turnover of the two substrates. This result suggests that turnover for the conversion of 10,11 epoxy FDP (**166**) by GdoIS is about 20% of that of the natural substrate FDP.

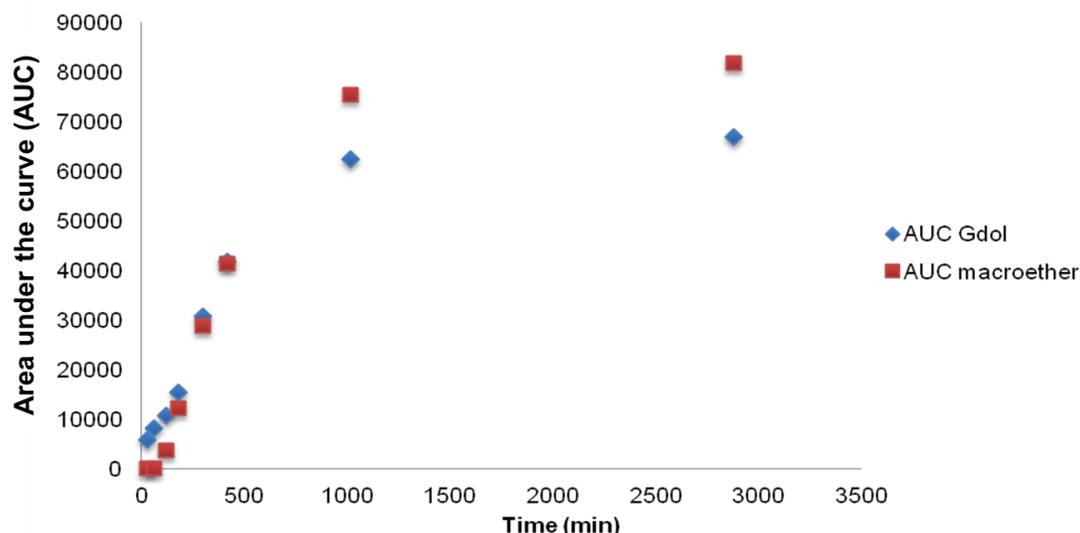


Figure 3-6 Relative quantities of products in a competitive assay of 10,11-epoxy FDP (0.125 mM) and FDP (0.025 mM) with GdoIS as judged by relative integration of product peaks in the total ion chromatogram of the pentane extractable products.

The incubation yield was then improved using the segment flow procedure developed for sesquiterpene synthases in the Allemann group by varying the residence time, ratio between organic and aqueous phase, concentration of GdoIS, Mg^{2+} or pH (see Figure 1-20).^[94,145] Tang *et al.* reported that reactors with 0.5 mm internal diameter (I.D.) yielded higher yields for aristolochene synthase (AS) and amorphadiene synthase (ADS) thanks to small segments of organic/aqueous phase. A increasing the ratio between the organic and aqueous phase reduces the size of segments and was shown to deactivate AS while optimal yields were obtained using 4:1 ratio with ADS. Similar behaviour to AS is observed with GdoIS (see Table 3-2, entry 4 versus 7, 6, 10, 11). Overall, yields were found to be lower when pH was at 7.5 instead of 8.0 and Mg^{2+} concentration at 5 mM instead of 7.5 mM. This supports the optimal condition results obtained by Dr. Grundy for GdoIS optimal pH for steady-state kinetic measurements (Mg^{2+} = 5 mM, pH 8.0, [enzyme] = 30 nM).^[142] Varying the enzyme

concentration did not greatly influence the yield while residence time from 60 min were observed to result in higher yields. The segmented flow procedure improved the yield to 75% (measured by GC) using 5 μM concentration GdoIS, 7.5 mM Mg^{2+} , at pH 8.0 with a residence time of 60 min, and a ratio organic:aqueous phase of 1:1.

Table 3-2 Optimisation of the incubation of 10,11-epoxyFDP with GdoIS using flow segmented system. Yield measured by GC FID using farnesol as internal standard.

Entry	Time (min)	Ratio (org:aq)	Enzyme (μM)	Mg^{2+} (mM)	Substrate (mM)	pH	Yield
1	30	1	3	5	0.35	7.5	0%
2	30	3	3	5	0.35	7.5	0%
3	45	2	5	5	0.35	7.5	8%
4	30	3	10	5	0.35	7.5	22%
5	60	0.5	5	5	0.35	7.5	58%
6	30	3	15	5	0.35	7.5	21%
7	30	1	10	5	0.35	7.5	32%
8	60	3	5	5	0.35	7.5	50%
9	30	0.5	15	5	0.35	8	41%
10	30	3	10	7.5	0.35	8	20%
11	30	3	5	7.5	0.35	8	31%
12	60	1	15	7.5	0.35	8	54%
13	60	1	10	7.5	0.35	8	60%
14	90	1	5	7.5	0.35	8	75%
15	120	1	15	7.5	0.35	8	70%

Initial $^1\text{H-NMR}$ spectroscopic analysis of the purified product at room temperature in CDCl_3 showed broad, poorly defined signals in some areas of the spectrum. This prevented a full assignment of the spectrum (see yellow spectrum on Figure 3-7) but allowed excluding several hypothetical structures (see Scheme 3-4). Four alkene protons were discernible in the region $\delta_{\text{H}} = 4.5\text{-}5.5$ ppm with a clear pair of singlets corresponding to an exo-methylene group. This ruled out structures **170** and **171** as reaction products (see Scheme 3-4). There were also three singlets corresponding to methyl groups between $\delta_{\text{H}} = 1.5$ and 2.0 ppm, ruling out structure **172**, which would give a doublet for one of these resonances (see Scheme 3-4). However, the presence of several broad signals prevented a full assignment of the spectrum. Slow exchange on the NMR timescale at room temperature has been observed previously for similar medium sized ring systems;^[70] hence variable temperature NMR spectra were measured between -50 °C and $+50$ °C. The structure was successfully elucidated at $+50$ °C. The NMR signals were much sharper at $+50$ °C and an ABX system corresponding to the

diastereotopic protons on C1 was identified between $\delta_{\text{H}} = 3.6$ and 4.0 ppm. The signals at $\delta_{\text{H}} = 4.98$ and 5.32 ppm also resolved into a triplet and a double doublet. They were assigned as the protons on C6 and C2 of **146** (Table 3-3 and Figure 3-9).

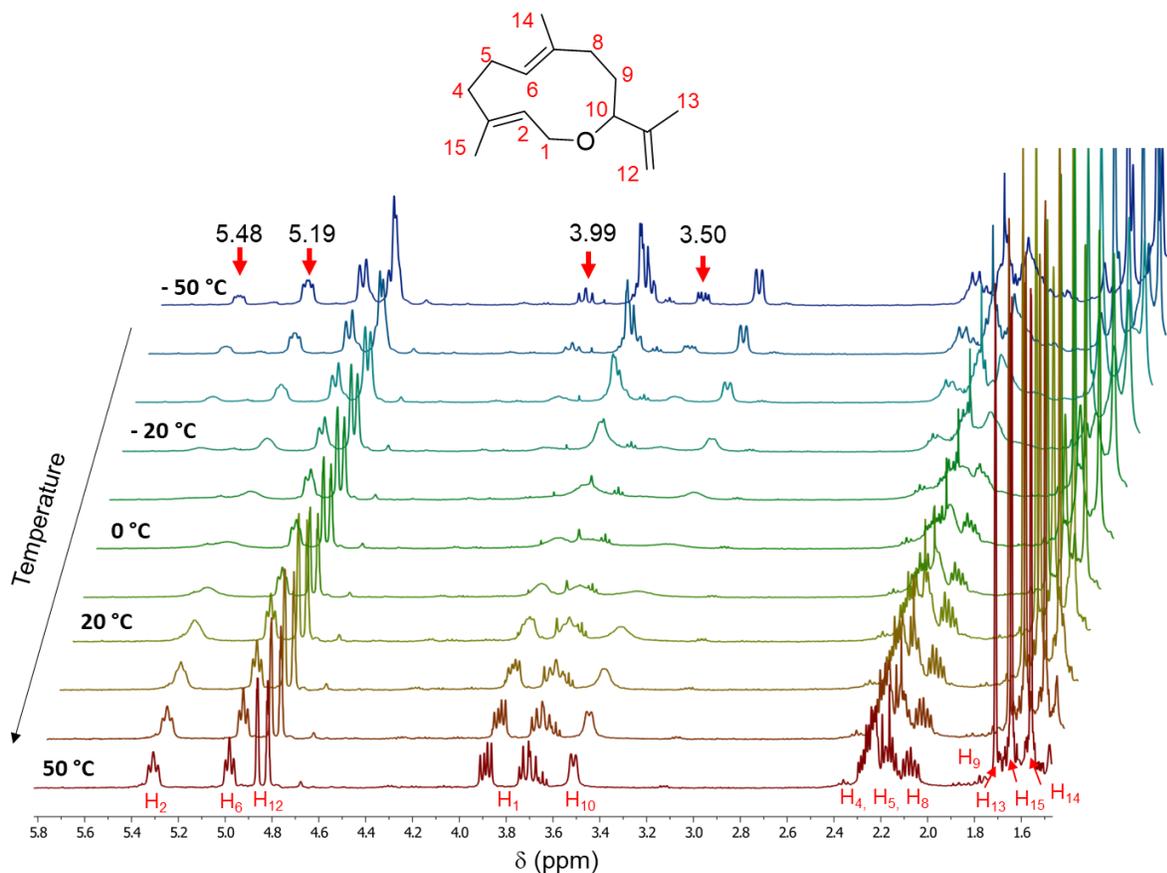


Figure 3-7 ^1H -NMR spectra (400 MHz, CDCl_3) of **146** from $-50\text{ }^\circ\text{C}$ to $+50\text{ }^\circ\text{C}$ ($10\text{ }^\circ\text{C}$ increment).

A full assignment of the ^1H and ^{13}C NMR spectra is given in the section 6.1.28 (Materials and methods) and confirmed **146** as the major product generated by GdolS from **166**. At $-50\text{ }^\circ\text{C}$ two conformations were apparent as indicated, for example, by the signal at $\delta_{\text{H}} = 5.32$ ppm corresponding to the proton on C2 split into two resonances at $\delta_{\text{H}} = 5.19$ and 5.48 ppm (Table 3-3, Figure 3-7 and Figure 3-9.) in $\sim 2:1$ ratio. Resonances for the minor conformation of **146** also appeared at $-50\text{ }^\circ\text{C}$ in all other regions of the spectrum, particularly clearly for the protons on C1 and C10 ($\delta_{\text{H}} = 3.2 - 4.0$ ppm) and in the alkyl region (see Figure 3-8). A NOESY spectrum at $-50\text{ }^\circ\text{C}$ showed several distinct NOEs (Figure 3-9) allowing some conformational restraints to be applied. Close proximity was apparent between the protons on C2 and C6, C13 and C15 and C10 and C12. Hence, the major conformation of (*R*)-**146** is the down-down

form, in which C14 and C15 are on the same side of the ring system (Figure 3-9) with (*S*)-**146** being the mirror image. NOEs were not detectable for the minor isomer, however, the most significant changes in the $^1\text{H-NMR}$ spectrum, as temperature decreased, corresponded to the protons on C1 and C2 suggesting that the minor conformer is the up-down conformation (Figure 3-9), where the methyl groups are on opposite sides of the 10-membered ring. These two alkyl groups change position most during the conformational transition.

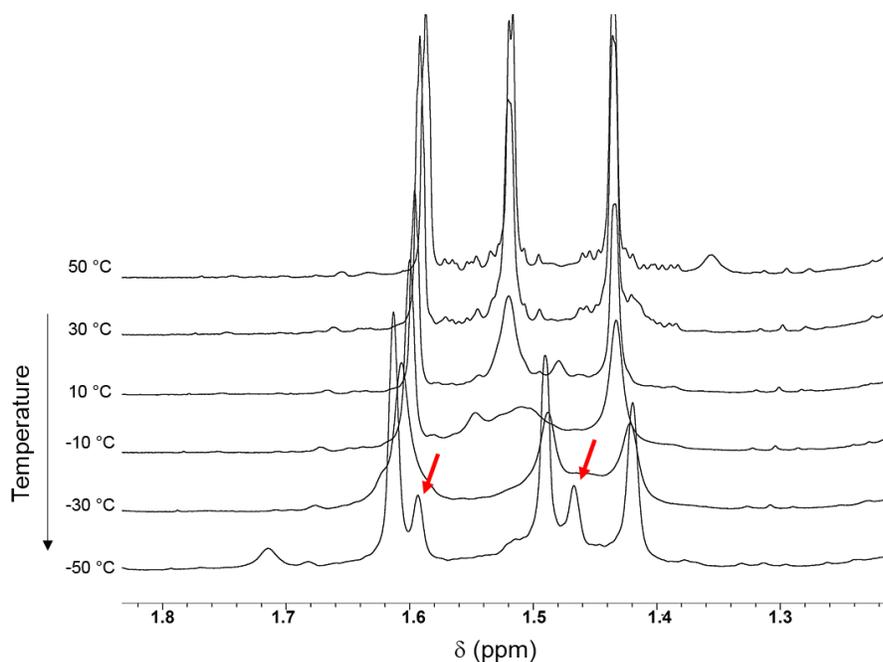


Figure 3-8 $^1\text{H-NMR}$ spectra (400 MHz, CDCl_3) in the alkyl region ($\delta = 1.3 - 1.8$ ppm) of **146** from -50 °C to $+50$ °C (20 °C increment).

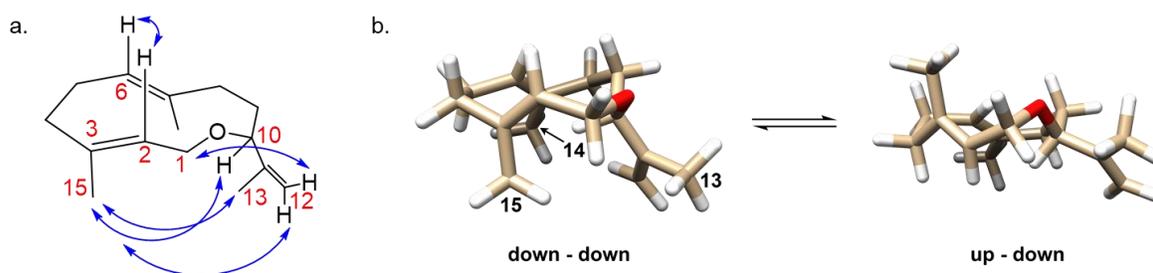


Figure 3-9 a. observed NOEs for **x**. b. Sketch of the proposed conformational equilibrium at -50 °C between the major down-down (CH_3 -14 and 15 compared) conformer of (*R*)-**146** and the minor up-down form.

Table 3-3 Proton NMR spectroscopic assignment of **146** at -50 °C and +50 °C. [a] All resonances are δ_H (400 MHz, CDCl₃) in ppm. Entries are chemical shift followed by multiplicity and coupling constants (Hz).

Assignment	-50 ° C major conformer ^[a]	-50 ° C minor conformer ^[a]	+50 ° C ^[a]
CH ₂ -1	3.81 – 3.68 (m)	AB part of ABX system 3.99 (dd, ² J _{HH} = 12.0, ³ J _{HH} = 10.5) and 3.50 (dd, ² J _{HH} = 12.5, ³ J _{HH} = 5.0)	AB part of ABX system 3.89 (dd, ² J _{HH} = 12.3, ³ J _{HH} = 7.0) and 3.70 (dd, ² J _{HH} = 9.0, ³ J _{HH} = 7.0)
CH-2	5.19 (dd, ³ J _{HH} = 10.0, ³ J _{HH} = 6.0)	X part of ABX system 5.48 (dd, ³ J _{HH} = 10.5, ³ J _{HH} = 5.0)	X part of ABX system 5.32 (dd, ³ J _{HH} = 9.0, ³ J _{HH} = 7.0)
CH ₂ -4	2.46 – 1.90 (m)	2.46 – 1.90 (m)	2.32 – 2.02 (m)
CH ₂ -4	2.46 – 1.90 (m)	2.46 – 1.90 (m)	2.32 – 2.02 (m)
CH ₂ -5	2.46 – 1.90 (m)	2.46 – 1.90 (m)	2.32 – 2.02 (m)
CH-6	4.99 – 4.89 (m)	4.99 – 4.89 (m)	4.98 (t, ³ J _{HH} = 7.0)
CH ₂ -8	2.46 – 1.90 (m)	2.46 – 1.90 (m)	2.32 – 2.02 (m)
CH ₂ -9	1.61 – 1.45 (m)	1.61 – 1.45 (m)	1.68 – 1.52 (m)
CH-10	3.26 (d, ³ J _{HH} = 9.5)	3.26 (d, ³ J _{HH} = 9.5)	3.51 (m)
CH ₂ -12	4.87 – 4.68 (m)	4.87 – 4.68 (m)	4.88 – 4.84 (m) and 4.84 – 4.79 (m)
CH ₃ -13	1.68 (s)	1.68 (s)	1.70 (s)
CH ₃ -14	1.49 (s)	1.54 (s)	1.54 (s)
CH ₃ -15	1.56 (s)	1.66 (s)	1.64(s)

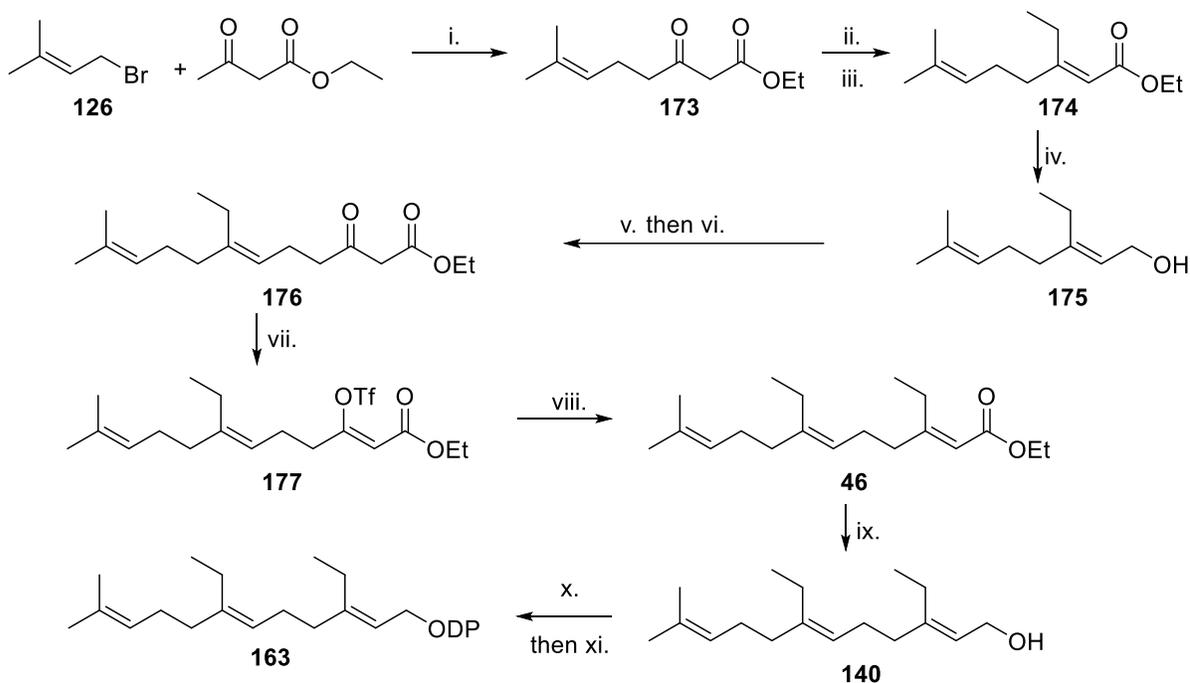
3.4. VARIOUS FDP ANALOGUES SYNTHESISED IN THIS WORK

Various analogues were synthesised for chapter 2 and chapter 4 studies. Their synthesis route is described below.

3.4.1. 14,15-dimethyl FDP synthesis

(14,15)-Dimethyl FDP (**163**) was prepared with the help of a project student, Clotilde Phillipe as part of the scope investigation of the one pot enzymatic diphosphorylation of polyprenol described in chapter 2 and the optimisation of the incubation conditions with the synthase is described in chapter 4 via an 11-step synthesis (see Scheme 3-6). Commercially available prenyl bromide (**126**) was subjected to chain elongation with a dianion prepared by treating ethyl acetoacetate successively with sodium hydride and *n*-butyl lithium (89%). The resulting compound (**173**) was then transformed in a stereoselective manner to the (*Z*)-enol triflate (quantitative) in order to yield **174** by a CuCN-mediated methylation using a Grignard reagent (61% over 2 steps). This placed the first methyl group on the future C14 (14,15)-dimethyl FDP (**163**). The α,β -unsaturated ester **174** was then reduced with di-isobutyl aluminum hydride

(DIBAL-H) to the alcohol (**175**, 88%). The latter was brominated using phosphorus tribromide (quantitative) and underwent a second identical chain elongation (**176**, 57%) followed by generation of the vinyl triflate and copper catalysed alkylation to produce **177** (60% over 2 steps). Reduction of the alcohol with DIBAL-H afforded (14,15)-dimethylfarnesol (**140**, 88%). Bromination using phosphorus tribromide followed by diphosphorylation using Davisson *et al.* procedure gave the final compound (**163**) and was used as the trisammonium salt (24%).^[110]

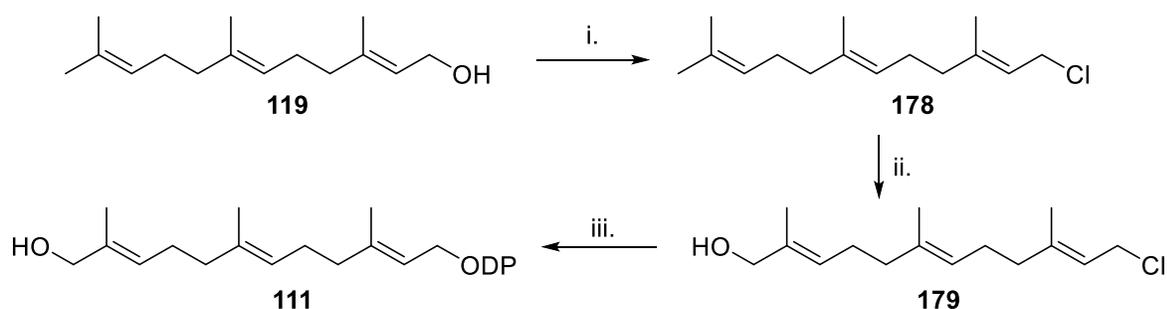


Scheme 3-6 Synthesis of 14,15-dimethyl FDP Reagent and conditions. *i.* Ethyl acetoacetate (3 eq), NaH (3.5 eq), *n*-BuLi (3.1 eq), THF, 0 °C. *ii.* LiOTf (3 eq), Tf₂O (1.3 eq), NEt₃ (3 eq), DCM, 0 °C, 2 h. *iii.* CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. *iv.* DIBAL-H (3 eq), Toluene, -78 °C, 1h. *v.* PBr₃ (0.5 eq), THF, -10 °C, 30 min. *vi.* Ethyl acetoacetate (3 eq), NaH (3.5 eq), *n*-BuLi (3.1 eq), THF, 0 °C. *vii.* LiOTf (3 eq), Tf₂O (1.3 eq), NEt₃ (3 eq), DCM, 0 °C, 2 h. *viii.* CuCN (2.5 eq), EtMgBr (1.5 eq), THF, -78 °C, 3h. *ix.* DIBAL-H (3 eq), Toluene, -78 °C, 1h. *x.* PBr₃ (0.5 eq), THF, -10 °C, 30 min. *xi.* OPP (2 eq), MeCN followed by cation exchange (Davisson *et al.* method^[110] see chapter 2).

3.4.1. 12-Hydroxy FDP synthesis

12-OH FDP was prepared as part of the scope investigation of using novel methodologies for the incubation of sesquiterpenes synthases described in Chapter 4. This analogue was

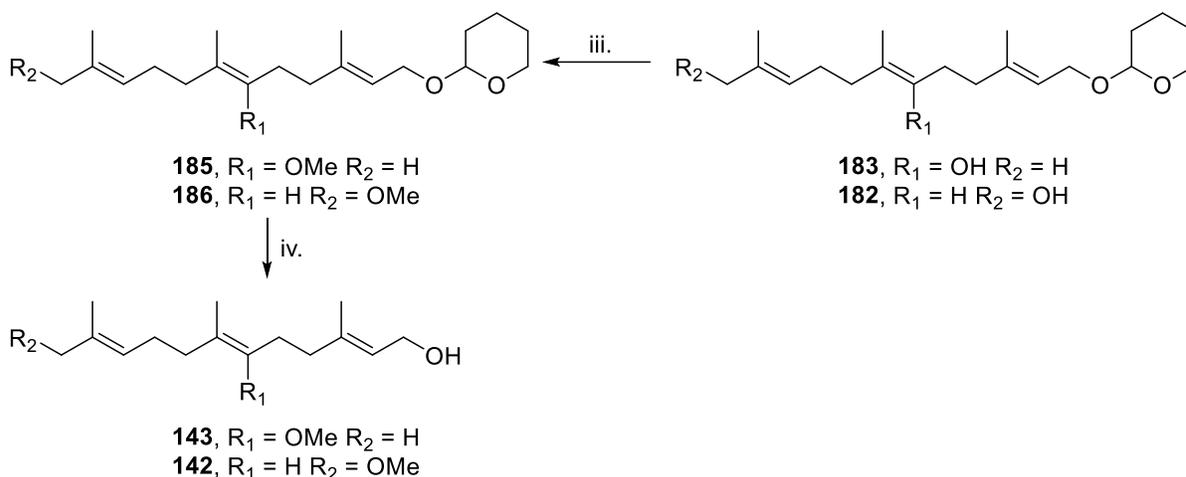
synthesised using a previously reported synthetic route from farnesol.^[146] The allylic alcohol farnesol (**119**) is first converted to the corresponding allylic chloride **178** using methane sulfonyl chloride with lithium chloride and 2,4,6-collidine. The resulting crude product is subjected to an allylic oxidation of its omega position using selenium dioxide in catalytic quantities together with salicylic acid and *tert*-butyl hydroperoxide giving **179** in 50% yield. It is finally diphosphorylated using the procedure of Davisson *et al.* yielding **111**, 46%) and used as its trisammonium salt (46%).^[110]



*Scheme 3-7 Synthesis of 12-OH FDP Reagent and conditions. i. LiCl (4 eq), 2,4,6-collidine (6 eq), MsCl (2 eq), DMF, 3h, 0 °C ii. SeO₂ (0.1 eq), Salicylic acid (0.1 eq), tBuOOH (0.1 eq), DCM, 0 °C to R.T, 20 h. iii. (Bu₄N)₃OPP (2 eq), MeCN followed by cation exchange (Davisson *et al.* method^[110] see chapter 2).*

3.4.2. 8-Methoxy FDP and 12-methoxy FDP synthesis

8-Methoxy FDP (8-OMe FDP) and 12-methoxy FDP (12-OMe FDP) were prepared to understand the promiscuity of the two enzymes used in the one pot enzymatic diphosphorylation of polyprenols described in Chapter 2. Synthetic routes to both analogues have been reported previously and they were prepared in the same manner.^[78] The five-step synthetic route is similar for both compounds. It starts from the protection of the hydroxy group of commercially available farnesol with 3,4-dihydropyran to give (**181**) in quantitative yield. It is followed by an allylic oxidation using catalytic amount of SeO₂ as previously described in section 3.2. Major products formed are from an allylic oxidation on C12 and C8. Compound can be purified at this stage to give **182** and **183** in 14%, and 46% yield, respectively. This is followed by methylation using sodium hydride and methyl iodide yielding **184** and **185** in 69% and 67% yield respectively. Both compounds are deprotected with catalytic amount of *p*-toluenesulfonic acid to yield 8-OMe farnesol (**143**, 86%) and 12-OMe farnesol (**142**, 73%).



*Scheme 3-8 Synthesis of 8-OMe FDP and 12-OMe FDP. Reagent and conditions. i. 3,4-dihydropyran (1.1 eq.), pyridinium *p*-toluenesulfonate (0.1 eq), DCM 1 h. ii. SeO_2 (0.1 eq), salicylic acid (0.1 eq), *t*BuOOH (0.1 eq), DCM, 0 °C to R.T., 20 h. iii. NaH (1.5 eq), MeI (5 eq), THF, 0 °C, 24 h. iv. *p*-toluenesulfonic acid (0.1 eq), MeOH, 1 h.*

3.5. SUMMARY

Several FDP analogues were synthesised and incubated with sesquiterpene synthases to create novel sesquiterpenes. Although a similar approach to Touchet *et al.* work was used, 6,15-dimethyl FDP (**145**) and 12,15-dimethyl FDP (**148**) proved to not be accepted by (*S*)-germacrene D synthase or published variants. This shows a limitation to the chemoenzymatic method to create novel sesquiterpenoids. Further knowledge needs to be gathered to engineer proteins to accept unnatural substrate. In this case, obtaining of an X-ray crystal structure of GDS would greatly facilitate the process. On the other hand, it was verified that having an in-built nucleophile on the starting substrate allowed trapping an intermediate cation in the mechanism, here the farnesyl cation. The 10,11-epoxy FDP (**166**) was turned over by GdoIS to a 11-membered cyclic ether and fully characterised using variable temperature NMR spectroscopy.

CHAPTER 4

NOVEL METHODOLOGIES FOR SESQUITERPENE CATALYSIS

4.1. PREFACE

This chapter will present both an optimisation of the synthesis of the high-value compound (S)-14,15-dimethylgermacrene through segmented flow chemistry, in addition to results of a novel application of an oscillatory segmented flow system and a high-performance counter-current chromatography system (HPCCC) for the enzyme-catalysed production of sesquiterpenes. As discussed in chapter 1, in nature sesquiterpenes are synthesised in a single step from FDP in exquisite regioselective and stereospecific pathways by sesquiterpene synthases. Researchers have used recombinant sesquiterpene synthases expressed in *E. coli* to produce and expand the terpenome in search of novel compounds with biological or synthetic uses as an alternative to low-yield plant extraction or costly total synthesis. The bottleneck of this approach is the typically high hydrophobicity of the sesquiterpene products causing catalysis of the enzyme to slow down or even stop, this can be overcome with the use of biphasic systems. In this work, a biphasic segmented flow system was used to improve the conversion of 14,15-dimethyl FDP (**163**) to (S)-14,15-dimethylgermacrene D (**86**) by (S)-GDS Y406F through design of experiment (DoE). Although being a relatively high yielding procedure, the production rate was not sufficient to deliver the compound within a reasonable time frame for external studies of the potency of the compound as a semiochemical. For this reason, an oscillatory segmented flow and a HPCCC system were alternatively considered with several STS to improve the yield and reduce the reaction time. The oscillatory system proved to deactivate the enzyme leading to low yields. On the other hand, excellent results with the HPCCC lead to further investigation of this novel method with other enzymes. The HPCCC system was successfully applied to sesquiterpene synthases and lipases. This work has demonstrated the versatility of the HPCCC as reaction vessel for enzyme-catalysed batch and continuous processes.

4.2. EXPRESSION AND PURIFICATION OF SESQUITERPENES SYNTHASES

Recombinant aristolochene synthase (AS) from *Penicillium roqueforti*, amorphadiene synthase (ADS) from *Artemisia annua* and (S)-germacrene D synthase (GDS) from *Solidago canadensis* were obtained from the Allemann group gene library, expressed in *E. coli* BL21 (DE3) RP, *E. coli* BL21 (DE3), *E. coli* BL21 (DE3) respectively and purified as previously

described.^[129,145] SDS polyacrylamide gel electrophoresis (SDS PAGE) was used to verify protein overexpression and purification (see Figure 4-1). Due to the presence of a histidine tag in the recombinant ADS, GDS and GDS Y406F enzymes, these enzymes were able to be purified by nickel affinity chromatography (see chapter 2, section 2.2.1). Purification by this method typically gave a purity of 95% as judged by SDS-PAGE (see Figure 4-1, Figure 4-2). AS on the other hand, was purified by anion exchange column (Quaternary (Q) Sepharose) (see Figure 4-1). This purification technique allows binding of the charged enzyme to the oppositely charged resin. An increasing gradient of sodium chloride buffer was then used to elute the protein with high purity from the column. Protein concentrations were estimated using the Bradford method with commercial bovine serum albumin as calibration standard and aliquots of enzyme were stored in 10% glycerol at $-20\text{ }^{\circ}\text{C}$ until further use.^[147]

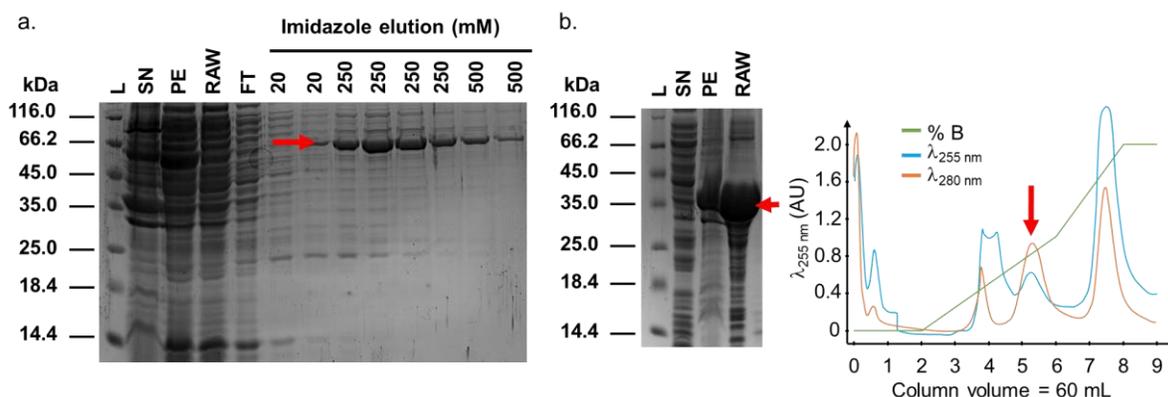


Figure 4-1 Purification of ADS and AS, red arrows indicating the enzyme. a. ADS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 66 kDa). b. AS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by anion exchange Q-sepharose of the RAW (expected MW: 39 kDa).

Products generated by sesquiterpene synthases in this chapter using various methods (segmented flow, oscillatory segment flow or HPLC) were analysed by GC-MS and compared to the literature data to identify the compounds. Gas chromatograms and mass spectra of the products are presented below (see Figure 4-3 to Figure 4-6). All techniques generated the expected product with similar amounts of side products compared to reported literature except for AS where surprisingly under 3% of germacrene A was observed as side product.^[148–151]

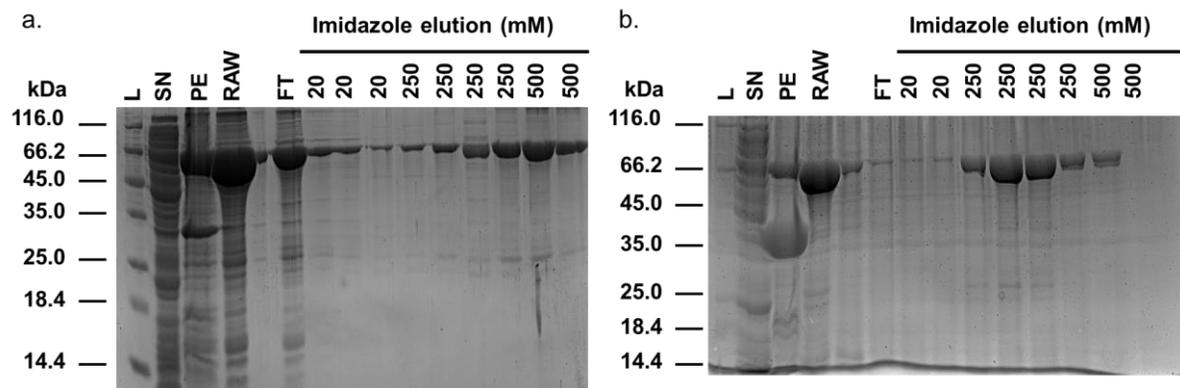


Figure 4-2 Purification of GDS and GDS Y406F, red arrows indicating the enzyme. a. GDS: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 63 kDa). b. GDS Y406F: SDS-PAGE analysis of the supernatant after sonication, pellet (PE) and supernatant (RAW) after basic extraction of the pellet after sonication followed by purification by nickel affinity chromatography of the RAW (expected MW: 63 kDa)

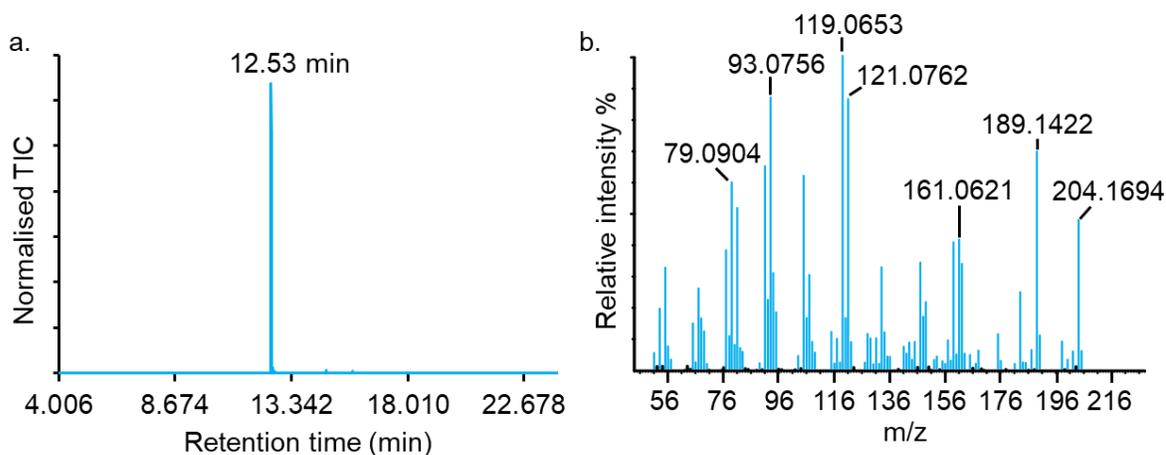


Figure 4-3 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with ADS a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.53 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.53 min.

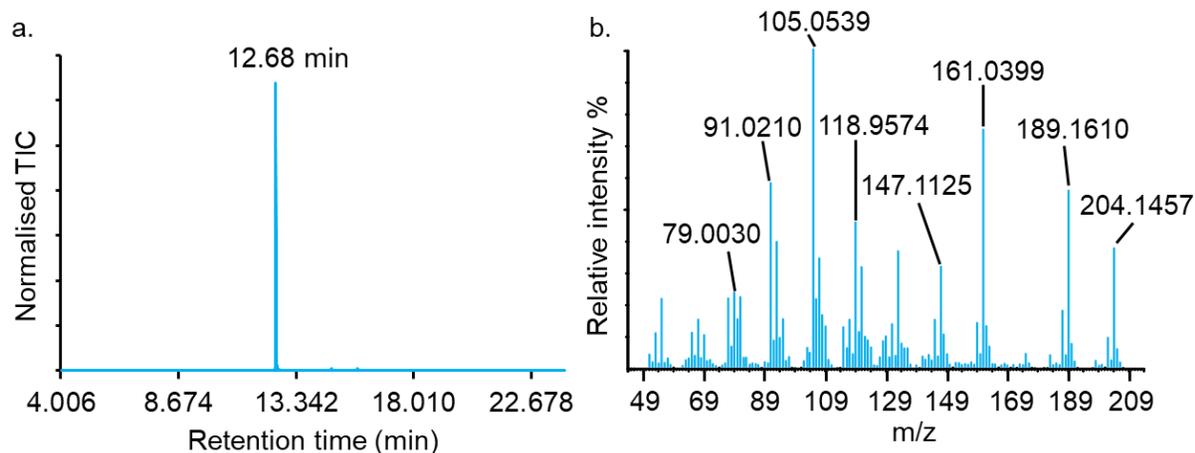


Figure 4-4 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with AS. a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.68 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.68 min.

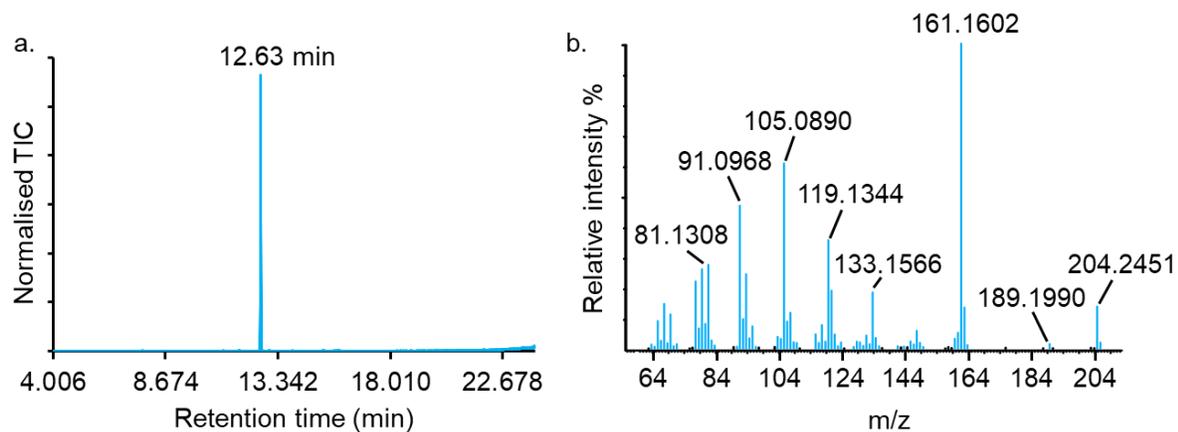


Figure 4-5 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with GDS. a. Total ion chromatogram highlighting the formation of one major product with a retention time of 12.63 min. b. Corresponding mass spectrum of the corresponding compound eluting at 12.63 min.

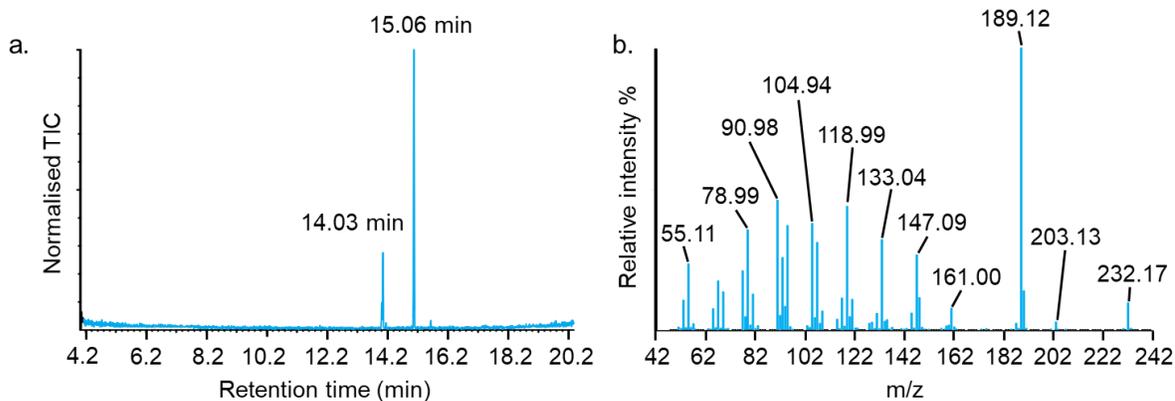


Figure 4-6 GC-MS analysis of the pentane extractable compounds arising from the incubation of FDP with GDS Y406F. a. Total ion chromatogram highlighting the formation of two products, predominant one with a retention time of 15.06 min. b. Corresponding mass spectrum of the corresponding compound eluting at 15.06 min identified as the 14,15-dimethylgermacrene D.

4.3. (S)-14,15-DIMETHYLGEMACRENE D SYNTHESIS USING A SEGMENTED FLOW SYSTEM

Owing to the promising semiochemical activity of the (S)-14,15-dimethylgermacrene D (**86**) (see Chapter 1, section 1.4.1), an efficient and up-scalable route to this compound is essential for its commercialisation.^[129] Extensive efforts were therefore devoted to the optimisation of the (S)-14,15-dimethylgermacrene D synthesis as the compound was needed for field trials (10 - 50 mg). Following previously reported work in the Allemann group,^[94,145] a segmented flow system was created by injecting an aqueous phase (a solution containing the enzyme and the substrate) and an organic phase (an organic solvent) through a T-mixer. The conversion of the substrate to the product by the enzyme naturally took place in the aqueous phase and the organic phase was continuously extracting the product to drive the reaction equilibrium toward the product and increase the reaction's yield. A DoE (design of experiment) approach was used in the past.^[145] However, only the one variable at a time method (OVAT) had been used to improve GDS conversion yield.^[94] Moreover, only a few factors were taken into account such as the internal diameter of the reactor tubing, the residence time, ratio between the organic and aqueous phase and enzyme concentration. Therefore, a screening design using a two-level fractional factorial design was carried out to characterise the major factors impacting on the yield, quantify their effects and interactions between each other. Eight parameters were chosen, namely:

- the sesquiterpene synthase concentration used in the incubation buffer (A),
- the substrate (FDP) concentration (B),
- the pH of the incubation buffer (C),
- the Mg^{2+} concentration in the buffer (D),
- the residence time (linear flow velocity) of the reaction (E),
- the inner diameter (I.D.) of the tubing reactor (F),
- the ratio between the organic and aqueous phase (G),
- the temperature (H).

Factorial designs are used to evaluate the importance of a factor to a process. In the case of 8 factors being considered, a full factorial design requires running 256 experiments.^[152] In practice, this is time consuming and costly to run such high number of reactions. A fractional factorial design was therefore used to reduce the number of runs. A 2^{8-6}_{IV} fractional factorial design with 4 additional central points was chosen because running 20 experiments allowed factors being differentiated from two-factor interactions hence reducing aliases (confounded effects).^[152-154] This type of DoE design is a resolution IV design meaning two-factor interactions can be confounded with other two-factor interactions.^[152,155] However, statistically only few factors will be found to be important among the 8 and two confounded interactions are most likely containing a significant factor.^[155] Pentane was chosen as the organic solvent because of its high log P. Moreover, no product was observed when extracting with toluene, acetonitrile, dichloromethane or ethyl acetate. Using hexane and heptane gave lower extraction yields (determined by GC-MS with α -humulene as internal standard), respectively 29% and 35% compared to 46% with pentane (results not shown here). Heptadecane was then chosen as the internal standard for further studies. The enzyme in buffer and the substrate were premixed, and the reaction was not quenched as previous studies have shown no impact on product yield.^[94] 2-Mercaptoethanol (5 mM) and bovine serum albumin (0.1%) were added to prevent enzyme aggregation and help improve enzyme stabilisation. Finally, the third reaction volume was collected and analysed by GC. Yields are reported in Table 4-1.

Chapter 4. Novel methodologies for sesquiterpene catalysis

*Table 4-1 Results of DoE experiments for GDSY406F in flow. Yields were analysed by GC MS and calculated by using heptadecane as internal standard. *yield measured by GC FID. Experiments with (-) did not give any results due to technical issues.*

A (mM)	B (mM)	C	D (mM)	E (min)	F (mm)	G(orga:aq)	H (°C)	Yield ^b
0.003	0.35	8.5	5	30	0.5	3	37	5%
0.015	0.35	8.5	15	60	0.8	3	37	7%
0.003	0.1	8.5	15	60	0.5	1	37	24%
0.003	0.35	7.5	15	60	0.5	3	20	5%
0.015	0.1	8.5	15	30	0.5	3	20	46%
0.015	0.35	7.5	15	30	0.5	1	37	18%
0.015	0.1	7.5	5	60	0.5	3	37	-
0.003	0.1	7.5	5	30	0.5	1	20	21%
0.003	0.35	8.5	15	30	0.8	1	20	10%
0.009	0.23	8	10	45	0.5	2	28.5	-
0.003	0.35	7.5	5	60	0.8	1	37	6%
0.015	0.1	7.5	15	60	0.8	1	20	66% (60%*)
0.009	0.23	8	10	45	0.8	2	28.5	41%
0.015	0.35	7.5	5	30	0.8	3	20	17%
0.015	0.1	8.5	5	30	0.8	1	37	38%
0.003	0.1	7.5	15	30	0.8	3	37	87%
0.003	0.1	8.5	5	60	0.8	3	20	34%
0.015	0.35	8.5	5	60	0.5	1	20	11%
0.009	0.23	8	10	45	0.5	2	28.5	35%
0.010	0.35	7.5	10	45	0.8	2	28.5	47%

Design Expert® Version 10 software^[156] was used to analyse the data. Figure 4-7 represents the half-normal probability plot which shows ordered estimated factor effects from less important (right) to most important (left). Points away from the straight red line are significant factors, blue represents a negative factor while orange a positive factor.

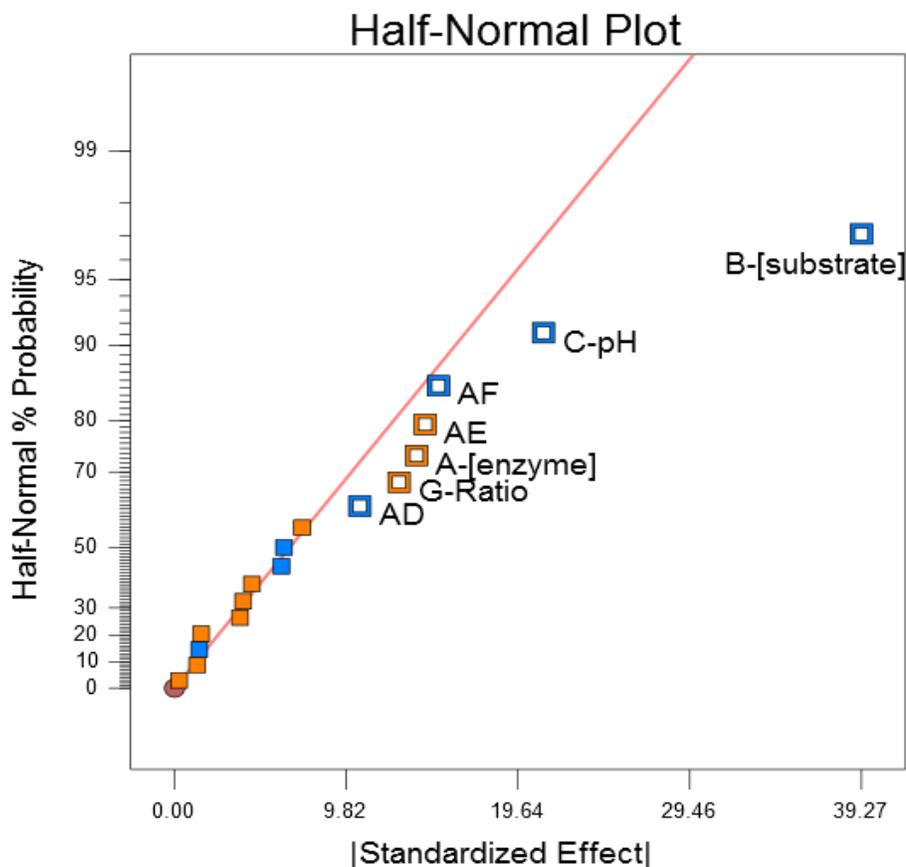


Figure 4-7 Half-Normal Plot of effects resulting from the fractional factorial two-level design of experiment. In blue negative effects and orange positive effects on the yield.

This study showed that two factors out of the eight were predominantly influencing the yield. First substrate concentration was found to have a highly negative impact the yield when increased (from 0.1 mM to 0.35 mM), while high pH is detrimental. It is known that enzymes have an optimal pH where activity is at its highest. The concentration, on the other hand, reveals that when using a segmented flow system, a low concentration is needed to maximise the yield of the reaction. Moreover, it is worth noting that the interaction of the enzyme concentration and the tubing diameter negatively influences the yield. This can be explained by the fact that increasing the diameter increases the segment size and leads to an insufficient interfacial area, hence giving a lower yield. Despite a high yield (87%) observed in one of the DoE experiments, this system quickly proved to be inefficient on lab scale. Using these conditions with a 2 mL reactor would only yield 1 mg in 56 hours, it was also not possible to leave the system unattended for long periods due to enzyme precipitation leading to blockage of the tubing. If the conditions giving 66% GC yield were used, the reaction would take less

time as the ratio between organic and aqueous phase was smaller, however due to the residence time being doubled, it would still take 74 hours to produce 1 mg. Increasing the flowrate for faster production was also impossible due to the length of tubing required, leading to an extremely high back pressure rendering the system unusable. Moreover, due to an excessive amount of emulsion at the exit of the reactor, loss of product occurred reducing the yield even further. Therefore, despite promising yields during the DoE, an alternative method was required for production of the desired sesquiterpene on a reasonable scale.

4.4. OSCILLATORY SEGMENTED FLOW SYSTEM

A potential improvement of the existing flow segmented system was investigated with the oscillatory multiphase flow method. This procedure is a variant of the segmented flow where an oscillation of the liquid/gas segments is induced to increase mixing and interfacial area.^[157] Oscillation of the droplets (*i.e.* segments of liquid) can be achieved by applying an alternative pressure gradient on one end of the tubing. In the literature, researchers have used this method to improve chemical as well as some biochemical reactions.^[157] A programmable syringe pump (Modular lab) with infuse/withdraw capabilities connected to a flow reactor similar to the one used in the paragraph above (see 4.3) was used to induce this oscillatory motion in the segments (Figure 4-8). This would enhance the mixing and mass transfer between the two immiscible liquids. Aristolochene synthase from *Penicillium roqueforti*, amorphadiene synthase from *Artemisia annua* and (S)-germacrene D synthase from *Solidago canadensis* were chosen as model enzymes for the study and preliminary experiments were performed at different oscillation rates to evaluate the efficiency of this methodology. A 4 mL reactor (I.D. 0.5 mm) was used and filled with 0.5 mL of water then a segmented system ($V = 2$ mL) was prepared by injecting at the same flowrate through a T mixer a solution of pentane containing α -humulene as internal standard and an aqueous solution containing the enzyme and FDP, resulting in segments of a similar size. Then 0.5 mL of water was injected to fill the rest of the reactor, and the oscillatory motion was then created by withdrawing and reinjecting the last 0.5 mL of water at a determined flow rate ($0.050 \text{ mL}\cdot\text{s}^{-1}$, $0.1 \text{ mL}\cdot\text{s}^{-1}$ or $0.2 \text{ mL}\cdot\text{s}^{-1}$). After 1 h at room temperature, the reactor full content was collected and the organic phase separated, reaction yield was calculated by GC MS using the internal standard.

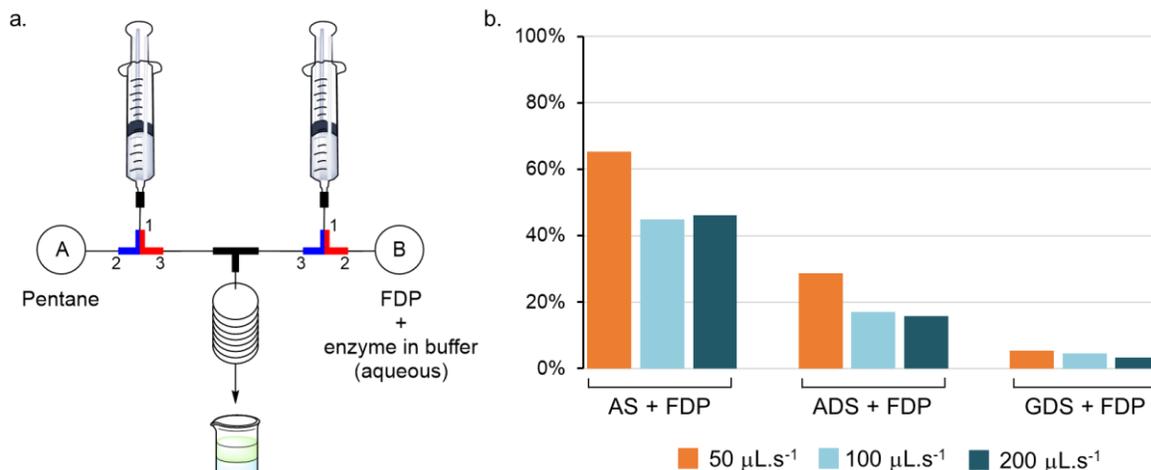


Figure 4-8 a. Schematic representation of the oscillatory segmented system used. A. reservoir containing pentane B. reservoir containing the enzyme in buffer and FDP. The modular system allows drawing liquid from the reservoir when the valve is closed in position 3 and open in position 1 and 2. Liquid is injected in the flow reactor when valve is switched to position 2 closes and open for 1 and 3. Oscillatory motion in the reactor is created by both syringes simultaneously pushing and drawing liquid with position 2 closed and 1 and 3 open. b. Results from performing the biocatalysis in an oscillatory segment system. Yields were obtained using GC-MS and calculated by using α -humulene as internal standard.

In almost all the cases, oscillatory segmented flow led to lower yields with increasing oscillation rate (see Figure 4-8). Upon collection of the reactor content, a significant amount of precipitated enzyme was also observed in all reaction conditions, this potentially indicates that contact between the enzyme and the organic phase was too high and the mixing too strong. As a result, generally low yields (45-65% for AS, 16-29% for ADS, 3-5% for GDS) were observed from this method which correlated with a low enzyme activity caused by the type of mixing the oscillatory motion created. This method do not perform better than traditional batch reaction (10-45% yield) or flow-segmented systems (50-97% yield).^[94,145] It is apparent that this method is unsuitable for such enzyme, GDS appears, in particular, more prone to lose activity in contact to pentane. Therefore, the oscillatory segmented flow system was not further investigated.

4.5. HPCCC

The high-performance counter-current chromatography (HPCCC) has been traditionally used as a liquid/liquid chromatography to separate mixtures of natural compounds or to purify pharmaceuticals on analytical or industrial scales.^[158–164] Recent developments in this field have made robust, reliable, high flow counter-current instruments commercially available. There are currently several types of CCC instruments but only the J-type machine has been used in this work.^[159] A J-type equipment is designed in a way in which the tubing do not get twisted while it is rotated, this allows a continuous use of the machine without the need of performing regular maintenance.

In contrast to high-performance liquid chromatography (HPLC), the stationary phase is a liquid phase in the HPCCC. The two liquids in the HPCCC are defined as follows: one liquid is stationary in the column (the stationary phase) and the other is pumped through it (the mobile phase). The HPCCC system uses a PTFE tubing wrapped in a coil onto a drum (bobbin) (see Figure 4-9). Each HPCCC has two bobbins. The bobbins (hence the coil) are rotated in a planetary motion that creates a radial centrifugal force. The immobile liquid phase is held stationary in the coiled tubing through those centrifugal forces. The second (mobile) liquid phase, immiscible with the stationary liquid phase, is at the same time pumped through the system. In CCC, two terms are important to define: the “head” which is where the liquid is pushed out of the coil and the “tail” in opposition to the “head”. Depending on the rotation direction, the head and tail position will change. During a run, the lighter phase (here the pentane) will always move toward the “head” while the heavy phase will remain at the “tail”. Therefore, when choosing the lighter phase as the mobile phase, the liquid needs to be pumped through the tail so it goes through the heavy phase, making the heavy phase unable to exit the column. This is how the heavy phase will stay stationary on the column. This is called tail-to-head direction or normal phase mode of the HPCCC. The HPCCC can also be run in reverse mode (head-to-tail). Experiments in this chapter were performed on a Dynamic Extractions Spectrum instrument (Slough, UK) which was fitted with an analytical scale column with a volume of 22 mL (0.8 mm I.D. PTFE tubing) and a semi-preparative scale column with a volume of 135 mL (1.6 mm I.D. of 1.6 mm PTFE tubing). However, a range of HPCCC instruments exist and reactors are available up to 18 L.

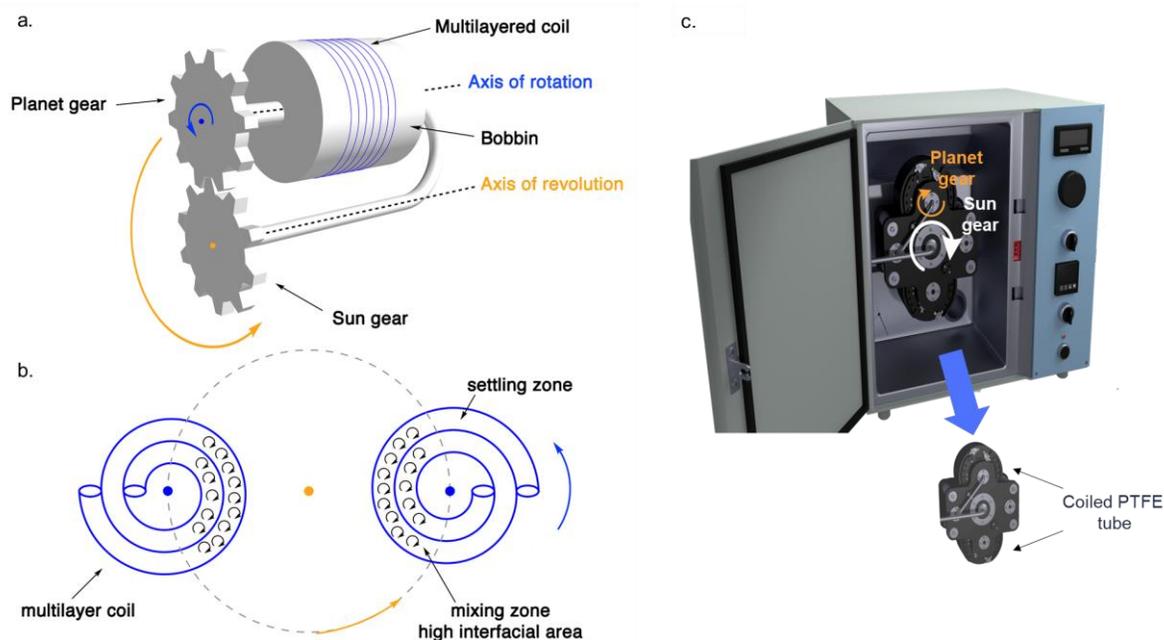


Figure 4-9 High-performance counter-current chromatography (HPCCC) device. a. Schematic view showing only one of the two bobbins for clarity. (Huynh F. et al., 2020) b. Showing the settling and mixing zone in the HPCCC column generated by variable centrifugal forces induced by the planetary motion of the bobbin. The mixing zone is situated toward the centre of the axis of revolution coinciding with a low acceleration field and the settling zone coinciding with the high acceleration field away from the centre of revolution. c. Photo of the HPCCC showing the two bobbins and its rotation in the apparatus. Photo of the HPCCC extracted and adapted from Dynamic extraction® presentation by Dr. David J. Rooke.

Because the bobbin revolves around the central axis while simultaneously rotating around its axis at the same velocity, it creates an effect similar to wave mixing with up to 4.8 million partitioning steps per hour.^[165] This results in a better surface area between the two phases and therefore a potential higher mass-transfer rate than the batch system and the segmented flow system (see Figure 4-10). Moreover, the constant change between settling and mixing constantly renews the interfacial area which should push the equilibrium of the reaction towards the sesquiterpene product. Thus, providing an ideal platform to improve the biosynthesis of product-inhibited enzymes such as sesquiterpene synthases. In this chapter, we proposed to use the HPCCC as a reactor for catalysis with the stationary phase containing the synthase and substrate while passing through an organic phase to continuously extract the product. In an early publication, the use of a counter current chromatography device for an enzymatic ester hydrolysis has been described. However, few details are given in the

seminal publication.^[166] A later report described the use of the centrifugal partition chromatography equipment (CPC) as a reactor for biocatalytic esterification.^[167]

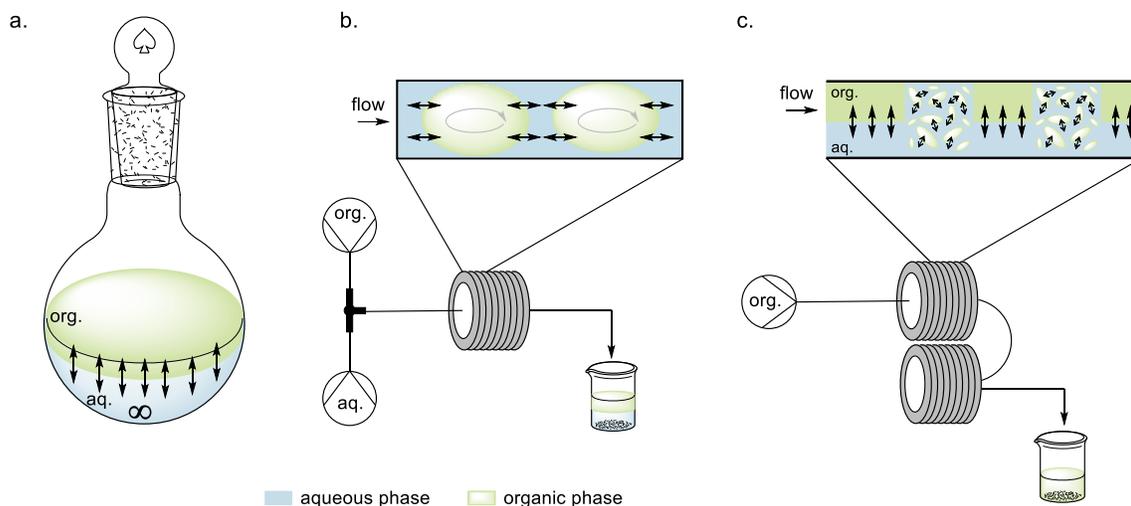


Figure 4-10 Comparison of the diffusion pattern of different biphasic processes. a. batch in a stirred tank. b. segmented flow. C. high-performance counter-current chromatography in normal phase, the organic phase is flowing, alternative segments are forming and mixing and settling zones are appearing.

First, a method was developed to set up the HPCCC as a reactor for biocatalysis. The HPCCC was used in normal phase therefore, with the stationary phase as the aqueous phase containing the enzyme and the substrate (FDP, **14**) in catalysis buffer, and the mobile phase as the organic solvent, pentane. To set up a reaction, the column was filled firstly with pentane ($V_{c,P}$) (Step 1) followed by V_{aq} mL of aqueous buffer containing enzyme (Step 2). At this stage, the tubing contains two segments as depicted in Figure 4-11. Then the rotation of the bobbin is turned on and pentane is pumped from tail-to-head to equilibrate the system, forming a true stationary phase and a mobile phase. $(V_{c,P} - V_{aq})$ mL of injected organic phase will exit the column before the system is equilibrated. We define this time as the equilibration phase. After this period any pentane that exits the column will have undergone intense mixing with the stationary phase and is therefore collected for analysis.

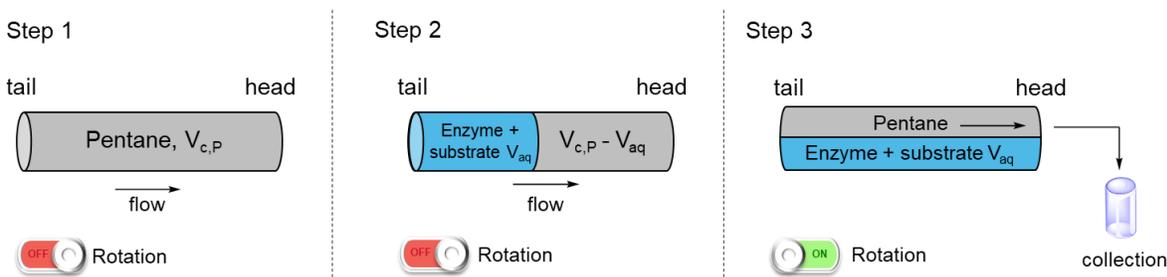


Figure 4-11 Setting up the HPLC to perform a biphasic reaction using sesquiterpene synthases and pentane as extraction solvent. (Huynh F. et al., 2020)

The conversion of FDP (**14**) to (1R,7R,8aS)-aristolochene (**72**) by AS was first examined. Purified enzyme (6 μM) and FDP (0.35 mM) in a buffer solution (20 mM Tris base, 5 mM 2-mercaptoethanol, 10% glycerol, 3 mM MgCl_2 , pH = 7.5) was loaded into a pentane-filled analytical column (22 mL volume) to form a stationary phase. Buffer, substrate and enzyme concentration were kept identical to those previously used for batch and segmented flow experiments. To assess the efficiency of HPLC, it was first determined how long the mobile phase needed to be circulated to reach a plateau in reaction yield. The pentane was collected in 5 mL fractions and analyzed by GC FID. The yield was calculated by comparing peak areas of the product to a calibration curve using α -humulene. Surprisingly, over 95% yield (measured by GC FID) was reached in less than one column volume (CV) at a flow rate of the mobile phase of $0.5 \text{ mL} \cdot \text{min}^{-1}$ (see Figure 4-12). The first phase depicted in blue on the graph is the equilibration of the biphasic system. It was collected and analysed to verify that no product was contained in those fractions. The residence time is defined as the ratio between the volume of the mobile phase in the reactor and the flow rate of the mobile phase. Control reactions (no substrate or no enzyme) were negative. To further investigate the potential of the HPLC, the concentration of the substrate was doubled (0.7 mM) to increase the scale of the reaction. This did not impact on the turnover as the same range of yield was achieved after 0.5 CV. Above 0.7 mM, precipitation of a white solid occurred making it impossible for higher FDP concentrations to be tested.

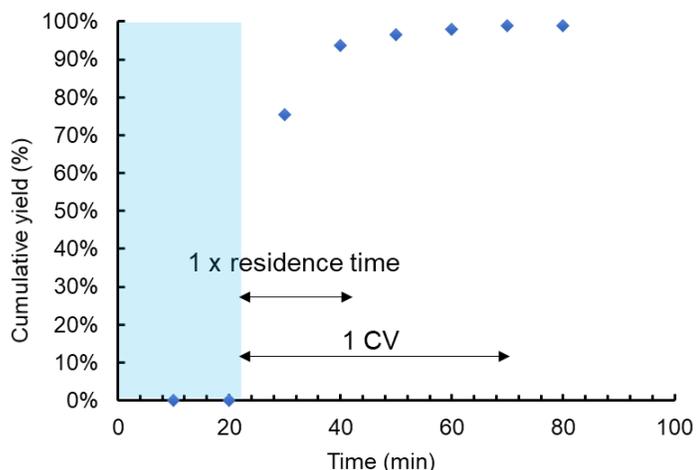


Figure 4-12 Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the HPLCC system (6 μM AS, 0.35 mM FDP, 0.5 $\text{mL}\cdot\text{min}^{-1}$ pentane flow rate), determined by GC-FID. (Huynh F. et al., 2020)

The flow rate was next increased from 0.5 $\text{mL}\cdot\text{min}^{-1}$ to 2 $\text{mL}\cdot\text{min}^{-1}$, to give a four-fold reduction in reaction time. The increased flow rate proved greater than the retaining centrifugal force exerted on the aqueous phase leading to loss of the stationary phase when using 2 $\text{mL}\cdot\text{min}^{-1}$ flowrate. This is due to the amount of retained stationary phase being closely correlated to its density, the revolving speed of the coil, and the flow rate of the organic phase. The relationship between the rotation speed, flow rate and composition of the incubation buffer was therefore characterised using the value S_f to determine the best parameters to use for faster reaction time (see Figure 4-13). The percentage retention of the stationary phase relative to the total column volume is defined as S_f using the following equation:

$$S_f (\%) = \frac{V_c - (V_d - V_{ec})}{V_c} \times 100$$

where V_c is the volume of the column, V_d is the measured displaced volume at the exit, V_{ec} is the volume resulting from the tubing linking the two bobbins, valves or tubing connecting the column to the pumps (for this system is was determined to 1 mL).

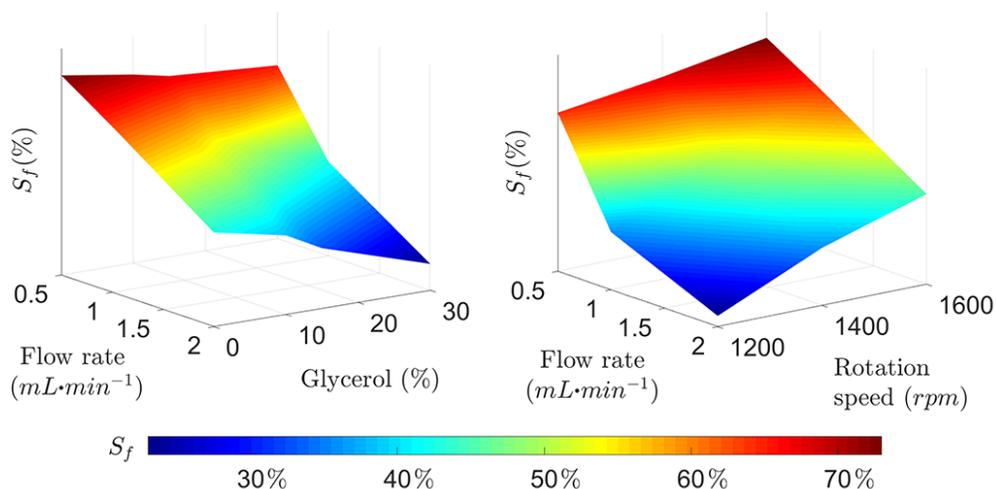


Figure 4-13 Left: Characterisation of the relationship between the pentane flow rate and glycerol content of the aqueous phase on the stationary phase retention S_f . The surface highlights that S_f is severely reduced with glycerol in the buffer and a high flow rate. Right: Characterisation of the relationship between the pentane flow rate and rotation speed of the bobbin on the stationary phase retention S_f . Surface highlights that S_f is severely reduced with low rotation speed and high flow rate. (Surface rendered using Matlab 2018b). S_f : Percentage retention of the stationary phase relative to total column volume. (Huynh F. et al., 2020)

S_f is observed to be severely reduced when increasing glycerol %, flow rate or decreasing the rotation speed. Glycerol is usually required to maintain the stability of the sesquiterpene cyclases in flow and batch, but the additional density it provides reduced S_f (see Figure 4-13). However, the reaction time when using HPCCC is so short it can be potentially omitted entirely. Therefore glycerol was removed from the incubation buffer and rotation speed set to 1600 rpm to retain as much aqueous phase as possible in the column which is approximately half of the column volume (10 mL) at a pentane flow rate of 2 $mL \cdot min^{-1}$. The reaction was performed again using AS (6 μM) and FDP (0.70 mM). This increased flow rate did not negatively impact the yield (see Figure 4-14) and reduced the reaction time from 44 to 11 minutes. Thus, HPCCC was demonstrated to be more efficient than segmented flow with a ~10-fold shorter reaction time.

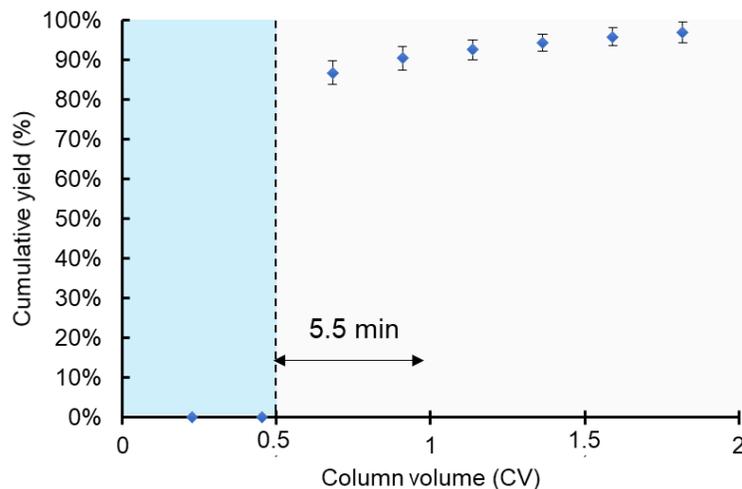


Figure 4-14 Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the HPCCC system (6 μ M AS, 0.7 mM FDP, 2 mL \cdot min $^{-1}$ pentane flow rate), determined by GC-FID. (Huynh F. et al., 2020)

Considering these results, these optimised conditions were applied to several different sesquiterpene synthases using FDP as well as a modified analogue, 12-OH FDP (**111**) with ADS. Batch, segmented flow and HPCCC reactions were performed to compare the three methods. Incubation buffers, enzyme and substrate concentration for each reaction were as follows:

Table 4-2 Incubation conditions for batch, flow and HPCCC set-up.

	Enzyme	Buffer	Glycerol (%)	Reducing agent	Enzyme (μ M)	Substrate (mM)
Batch	AS with FDP	Tris base (20 mM), MgCl ₂ (3 mM), pH = 7.5.	15	β -Me (5 mM)	6	0.35
Flow			15	β -Me (5 mM)		
HPCCC			-	-		
Batch	ADS with FDP	HEPES (25 mM), MgCl ₂ (5 mM), pH = 7.5.	10	DTT (1 mM)	10	0.35
Flow			10	DTT (1 mM)		
HPCCC			-	-		
Batch	GDS with FDP	Tris base (20 mM), MgCl ₂ (10 mM), pH = 7.5	10	β -Me (5 mM)	12	0.35
Flow			10	β -Me (5 mM)		
HPCCC			10	β -Me (5 mM)		
Batch	ADS with 12-OH FDP	Glycine (25 mM), NaOH (25 mM), MgCl ₂ (5 mM), pH = 9.4	-	DTT (1 mM)	10	0.35
Flow			-	DTT (1 mM)		
HPCCC			-	-		

Batch reactions were run for 24 hours with pentane (10 mL) overlaying the aqueous phase (10 mL). Yields were calculated by spiking a known volume of pentane with α -humulene or synthesised farnesal. The segmented flow reactions were run in 2 mL PTFE tubing (0.5 mm I.D.) using a ratio of 1 between organic:aqueous phase, except for ADS with 12-OH-FDP where 0.5 ratio (organic:aqueous) was used. Different residence times were used: 90, 70, 100 and 90 min for AS, ADS, GDS and ADS + 12-OH-FDP respectively. Vastly improved yields (up to 99% by GC FID) were observed compared to batch synthesis (21-32%, see Figure 4-15) or segmented flow synthesis (57-96%). GDS proved to be more difficult to use in the HPCCC without glycerol, therefore 10% was added and residence time increased to 14 min to retain the same amount of aqueous phase in the HPCCC.

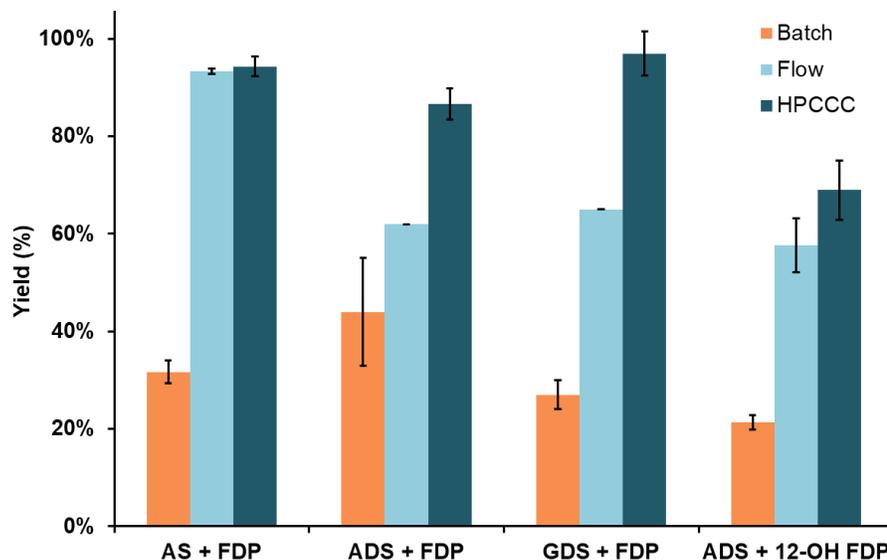


Figure 4-15 Comparison of the batch, segmented flow and HPCCC method for different substrates by *PR-aristolochene synthase* (AS), *amorphadiene synthase* (ADS) and *(S)-germacrene D synthase* (GDS). Yields were determined by GC-FID and calculated by using a calibration curve plotted with α -humulene. (Huynh F. et al., 2020)

HPCCC has been previously reported to scale well when used for purification.^[164] To verify if this statement applies to this new methodology, the optimized conditions were tested with AS using a larger HPCCC column (135 mL). In this case, 0.7 mM FDP was used, and 65 mL was retained on the column. A similar extraction profile was obtained to those for the smaller column (see Figure 4-16). Reactions were repeated four times to obtain a substantial amount of compound and avoid the loss of material due to evaporation. This resulted in excellent

isolated yields (see Figure 4-16) in good agreement with those measured by GC-FID for the smaller column.

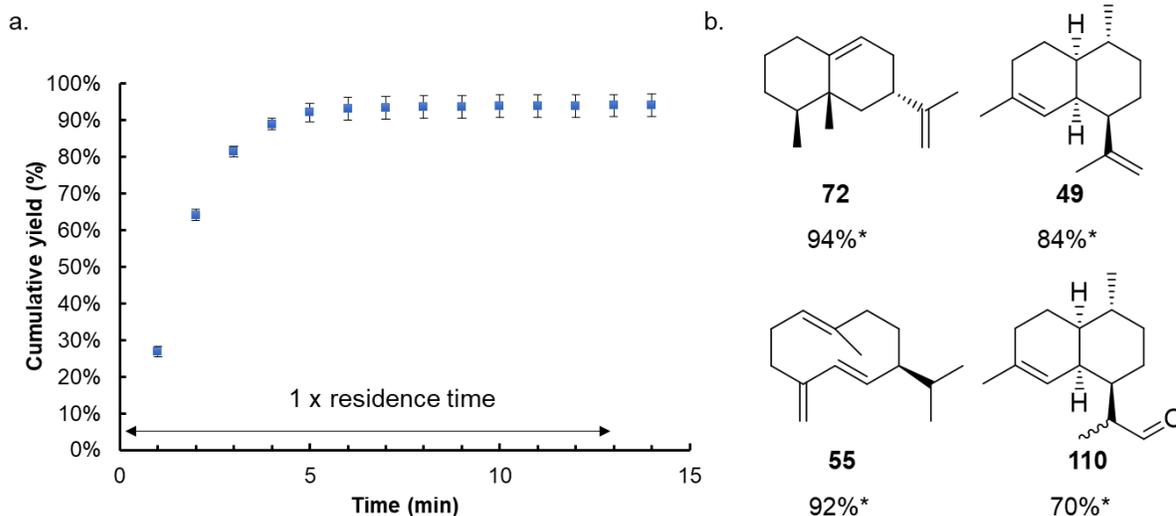


Figure 4-16 a. Time course for the AS-catalysed conversion of FDP to (1R,7R,8aS)-aristolochene in the preparative column of the HPLC system (6 μM AS, 0.7 mM FDP, 5 $\text{mL}\cdot\text{min}^{-1}$ pentane flow rate), determined by GC-FID. b. Isolated yields for different substrates by PR-aristolochene synthase (AS), amorphadiene synthase (ADS) and (S)-germacrene D synthase (GDS). The reaction products are shown performed on a preparative scale with a 135 mL HPLC column using the optimised conditions (pentane flow rate: 5 $\text{mL}\cdot\text{min}^{-1}$), each reaction was repeated four times to increase amount of product and reduce product loss when removing pentane under reduced pressure. (Huynh F. *et al*, 2020)

As expected, reactions of FDP with AS, ADS or GDS using the HPLC formed >90% of the major product, and 30-35 mg were isolated, highlighting the practicality of this method for producing meaningful quantities of sesquiterpene products. The reaction of 12-OH FDP with ADS generated three products in the pentane phase, as previously reported.^[151] Due to the relative ease of scale-up, enough material was produced to purify the products generated by ADS from 12-OH FDP, resulting in isolation of the previously uncharacterised side product (see Figure 4-17). In the original publication from Derimay *et al.*, aldehydes **112** and **113** could not be purified from the side product observed in the GC spectrum at 14.83 min and it was therefore tentatively proposed to be the enol form of aldehyde based from the compound's mass spectrum and experiments in deuterated buffer.^[151] Carrying out the incubation in deuterated buffer resulted in the incorporation of a deuterium in solely 11S-**112** epimer. It was therefore hypothesised that 11S-**112** could be deprotonated by the enzyme to form the enol

and then tautomerise to the 11*S*-**112** aldehyde while once 11*R*-**113** is produced, it remains as a stable conformer.

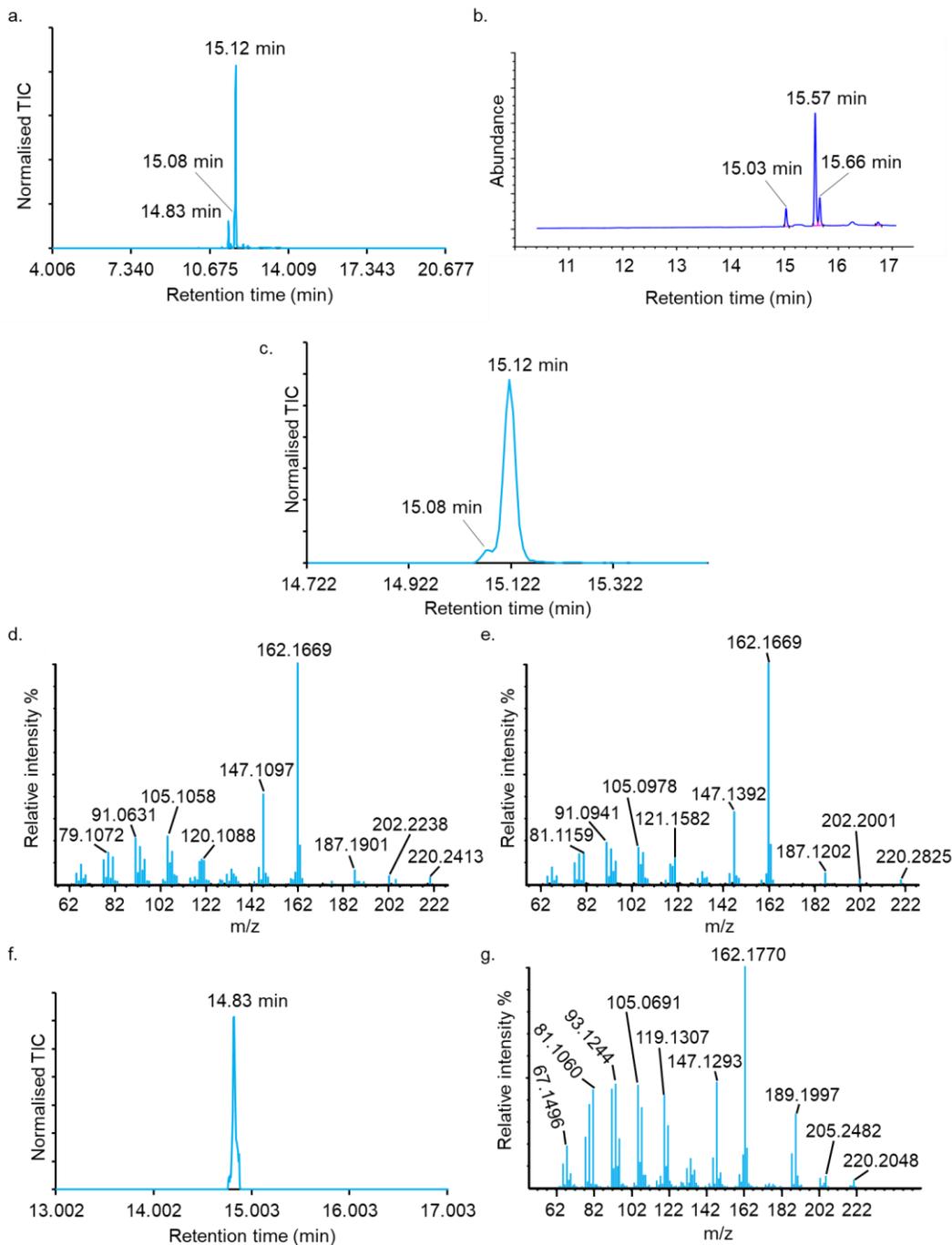


Figure 4-17 Total ion chromatograms and mass spectra of the pentane extracted products arising from the incubation of 12-OH FDP with ADS by HPLCC. a. Total ion chromatogram with three compounds eluting at 14.83 min (side product), 15.08 min (*R*-DHAA), 15.12 min (*S*-DHAA) b. Flame ionisation detection chromatogram (chiral column) with three compounds eluting at 15.03 min, 15.57 min, 15.66

min. c. Total ion chromatogram of purified aldehydes. d-e. Mass spectrum of compound eluting at 15.08 min and 15.12 min. f. Purified side product chromatogram. g. Mass spectrum of compound eluting at 14.83 min.

The purified side product with a retention time of 14.83 min was shown to have the same elementary composition as the aldehyde since the mass spectrum displayed a molecular ion peak with $m/z = 220$ (see Figure 4-17). A major fragmentation peak with $m/z = 162$ demonstrated a loss of a C_3H_6O fragment validating a potential enol structure but analysis of the purified compound NMR spectrum revealed the presence of only one alkene proton $\delta_H = 5.10$ ppm disfavouring the enol hypothesis. Moreover, ^{13}C and DEPT 135 spectra showed only two signals corresponding to the C2-C3 alkene bond, $\delta_C = 120.6$ and 139.7 ppm respectively. Instead, a clear AB system is observed with roofing doublet and dd signals at $\delta_H = 2.75$ and 2.50 ppm with $J = 4.7$ Hz and $J = 4.7$ and 0.5 Hz respectively. These two signals can be assigned to two deshielded diastereotopic protons such as a CH_2 in an epoxide group. This is further corroborated with a signal at $\delta_C = 58.2$ ppm corresponding to a quaternary carbon correlating with the diastereotopic CH_2 in a HMBC experiment and a fragment with $m/z = 205$ corresponding to the loss of a methyl group. Further elucidation of the structure using NMR spectroscopy confirmed the side product C5 generated is artemisinic-11S,12-epoxide (**186**) (see Figure 4-18). A 1H and ^{13}C NMR spectra is given in the section 6.3.4 Materials and methods and in the Annexes. This compound has only been reported to be produced by one engineered P450 from amorpho-4,11-diene in *E. coli* in a single publication.^[168]

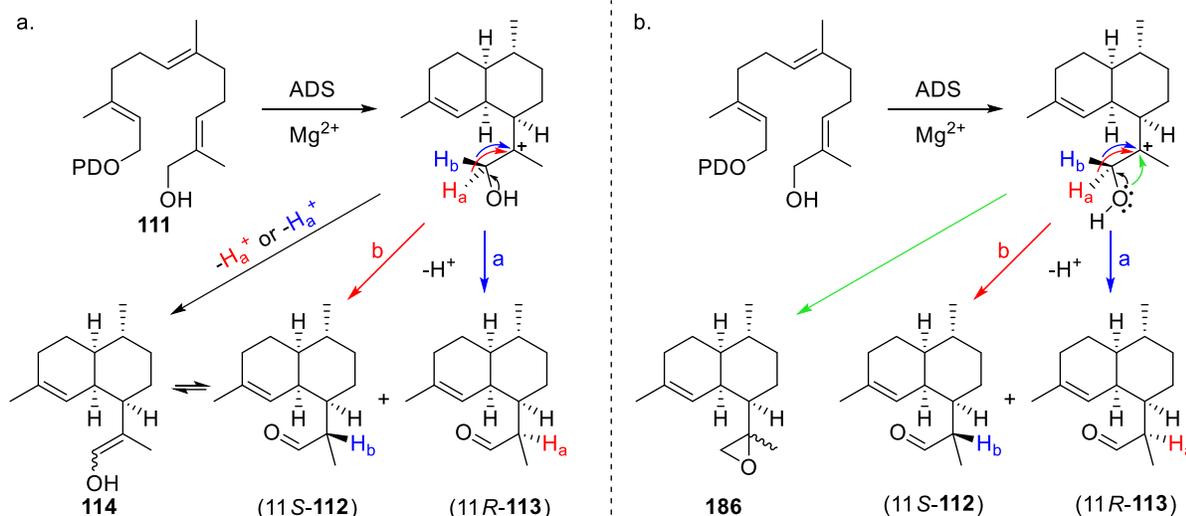
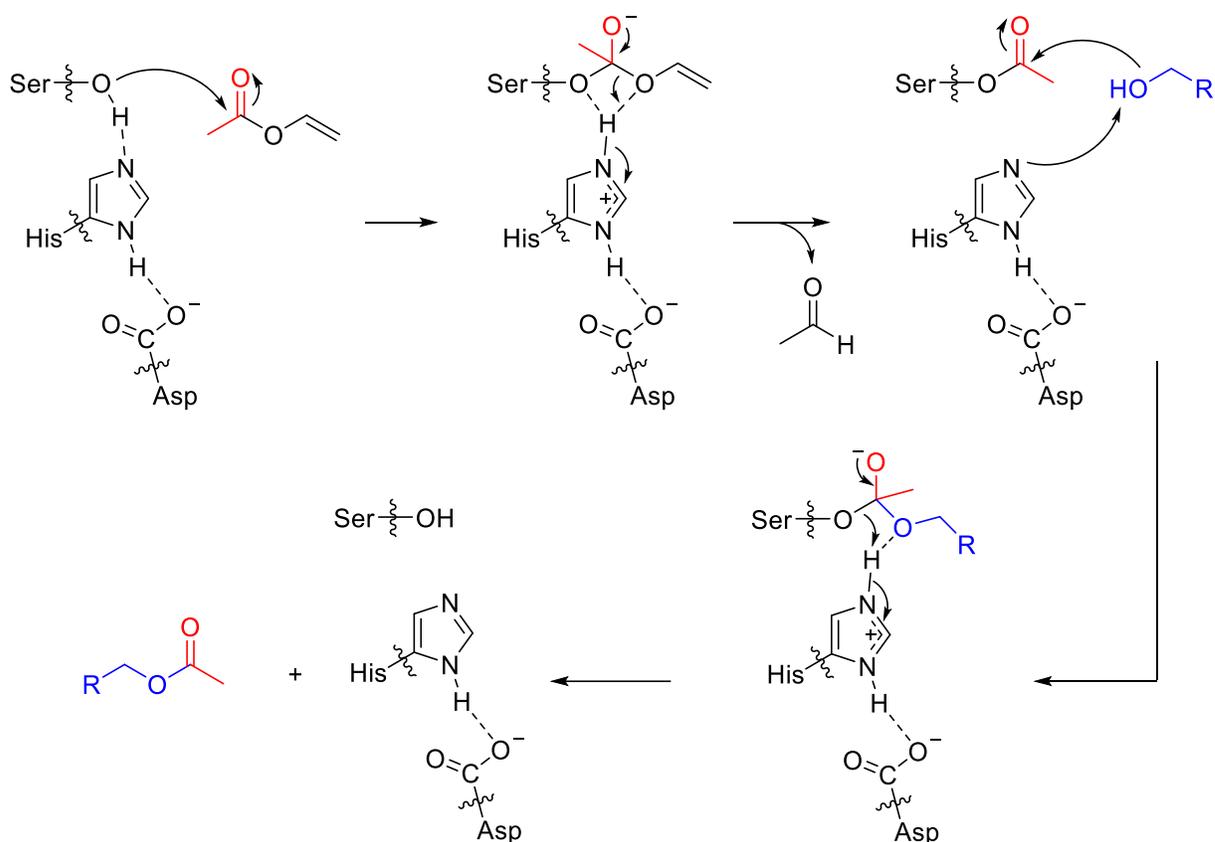


Figure 4-18 Products arising from the incubation of 12-OH FDP with ADS. a. Products and mechanism proposed by Demiray et al. b. Products and mechanism suggested in this chapter.

4.6. TRANSESTERIFICATION BY LIPASES

Lipases are hydrolases catalysing trans-, inter- and esterification. Of all enzymes, lipases have been the most studied enzyme and are nowadays used widely in many sectors such as pharmaceutical, textile, food industries or for biodiesel production. They are generally thermostable and compatible with organic solvents. They are often used immobilised to enable an easy workup, create a continuous process and recycle the enzyme.^[169] However, this also results in loss of activity as well as limitation in the mass transfer.^[170] One of the other key challenges of lipases is the formation of by-products which inhibit the enzyme, this is usually solved in the same way as with sesquiterpene synthases using an organic solvent for continuous extraction.^[171] This makes the HPCCC a potential solution for an efficient and fast method for lipase esterification. To further demonstrate the capabilities of the HPCCC system as a methodology for biocatalysis, lipase from *Candida Antarctica B* (CaLB) was used to produce commercially valuable octyl acetate (**187**) and (*R*)-2-pentyl acetate (**188**) from the corresponding alcohol (**189** and **190**) by transesterification with vinyl acetate (**191**) (see Figure 4-19).^[171,172] CaLB was selected for this study due to its high activity and stability.



Scheme 4-1 Mechanism of Cal B-catalysed transesterification of vinyl acetate with a primary alcohol.

For lipases, the HPCCC system was used in normal phase; that is to say, the lipase was held stationary in the aqueous phase and organic solvent was mobile. The HPCCC was set up slightly differently than for STS because both substrate and product were both soluble for the lipase transesterification are in the organic phase (see Figure 4-19). This differs to the application with sesquiterpene synthases, where only the product was soluble in the organic phase. In the case of lipases, it was possible to use the HPCCC in a continuous process similar to when using immobilised enzymes. The HPCCC was set up as followed: the analytical column was first filled with CalB in incubation buffer (20 mM phosphate buffer, pH 7.2), the rotation was then turned on to 1600 rpm and heptane was pumped through at 2 mL.min⁻¹. When hydrodynamic equilibrium was reached (volume of aqueous phase expelled is constant), the organic phase was switched to a heptane solution containing the substrates at a set concentration. The liquid at the exit of the reactor was collected after reaching steady-state (three times residence time, the residence time is defined as the ratio between the volume of mobile phase in the reactor and the mobile phase flow rate, here the residence time

is 6 min) and analysed by GC FID. The conversion was calculated using a calibration curve of octanol starting material.

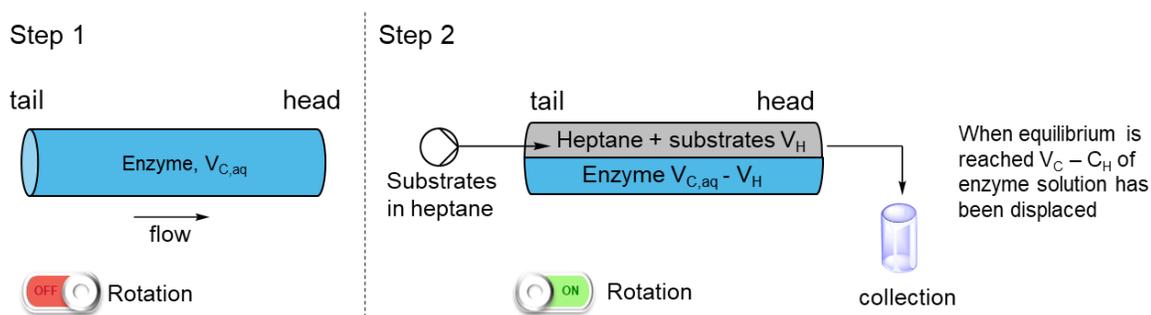


Figure 4-19 Setting up the HPCCC to perform a biphasic reaction using CalB lipase and heptane as extraction solvent. (Huynh F. et al., 2020)

Using these conditions, an initial experiment was performed with 10 mL of stationary phase at 25 °C containing 1 mg.mL⁻¹ of enzyme in the analytical column. Initial concentration of octanol was set to 30 mM to be able to compare future results to reported results when using immobilised CalB, and the molar ratio between octanol (**189**) and vinyl acetate (**191**) was set to one.^[172] Analysis by GC FID of the collected organic phase showed that a new product with a retention time of 8.81 min was formed, while octanol was not totally consumed as seen in Figure 4-20b. (retention time 7.74 min). The fraction was purified by flash chromatography in 10% EtOAc in hexane to isolate the new product and analysis by GC-MS and NMR spectroscopy showed that the product was octyl acetate (**187**) (see Figure 4-20c. and d.). This validated the potential of the HPCCC system to perform biocatalysis with lipases, and further optimisation was then performed to develop this application.

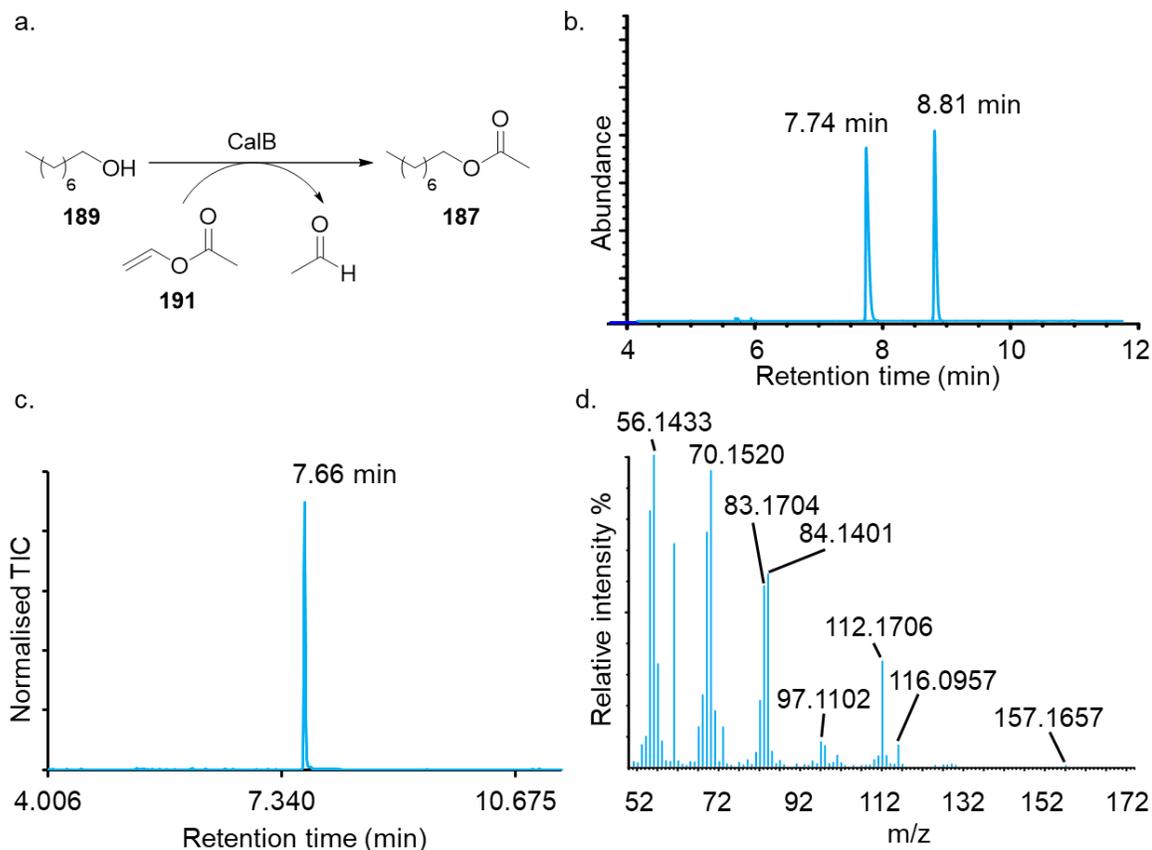


Figure 4-20 a. CalB-catalysed transesterification of octanol by vinyl acetate. b. Flame ionisation detection chromatogram of the collected heptane at the exit of the HPCCC with two compounds eluting at 7.74 min (octanol) and 8.81 min. c. Total ion chromatogram of purified compound eluting at 8.81 min in the GC FID. d. Mass spectrum of compound eluting at 7.66 min corresponding to octyl acetate after comparison to NIST library.

Many reports show the transesterification kinetics heavily depend on the molar ratio between the alcohol and acetylating agent due to the alcohol being a competitive inhibitor of the reaction.^[171] A higher concentration of vinyl acetate was expected to yield increased rate and conversion. Therefore, a set of experiments were performed with the same conditions as above with varying ratio of alcohol:acetate. As a comparison, the reactions were performed with the traditional batch method using a biphasic system in a round bottom flask. Similar conditions were used as for the HPCCC experiment: 10 mL incubation buffer (20 mM phosphate buffer pH 7.2, 1 mg.mL⁻¹ CalB) and 10 mL heptane containing the substrates. The reaction was stirred (800 rpm) at r.t. (25 °C) and the organic phase was regularly sampled to measure the conversion. As expected, a higher ratio of vinyl acetate yielded higher conversion in both batch and HPCCC methods. Conversion is observed to quickly plateau after 1.5 hours

in batch, with 16% at 1:1 ratio (alcohol:vinyl acetate) while reaching about 75% when using 1:9 ratio. In comparison, the HPCCC achieved 34% conversion in 6 minutes using 1:1 and 98% using 1:9 ratio. 98% conversion was also obtained when using 1:6 ratio. This shows the potential of this new methodology, showing qualitatively faster kinetics and almost total conversion of the substrate in only 6 min at 25 °C (see Figure 4-21).

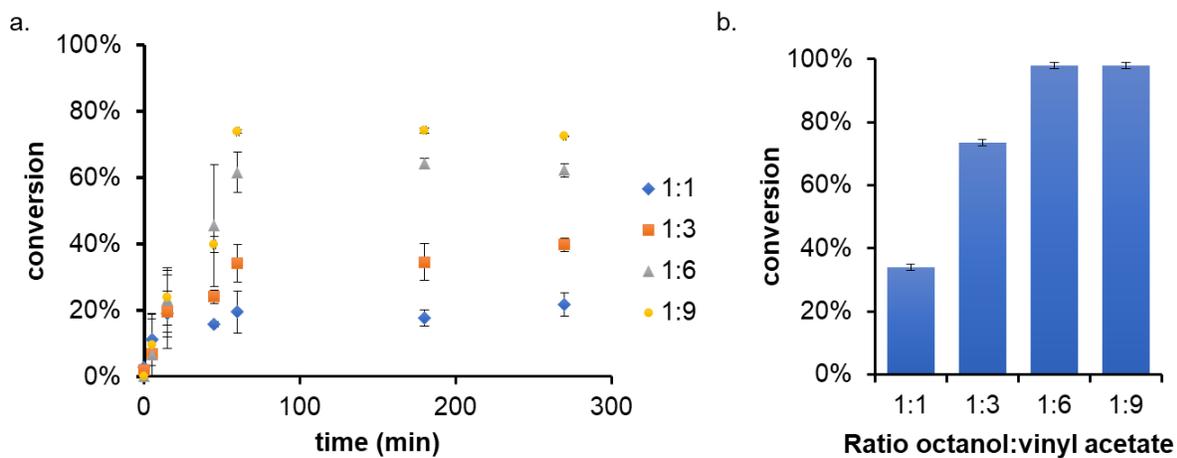


Figure 4-21 Influence of the ratio alcohol:vinyl acetate on the conversion of octanol to octyl acetate. a. Batch reaction. b. Reaction performed in the HPCCC, residence time 6 min.

Next, the influence of the temperature and enzyme concentration were studied using the HPCCC at a substrate ratio of 1:1 to reveal if these two parameters influenced the conversion. Interestingly, only a small variation in conversion was seen with no apparent trend (see Figure 4-22). Therefore, the concentration of alcohol was varied. The conversion was found to increase between 30 mM and 100 mM while decreasing beyond 100 mM octanol (see Figure 4-22). This type of behaviour was also observed in the literature and hypothesised as a dead-end inhibition complex formation by n-octanol with the lipase.^[171,172]

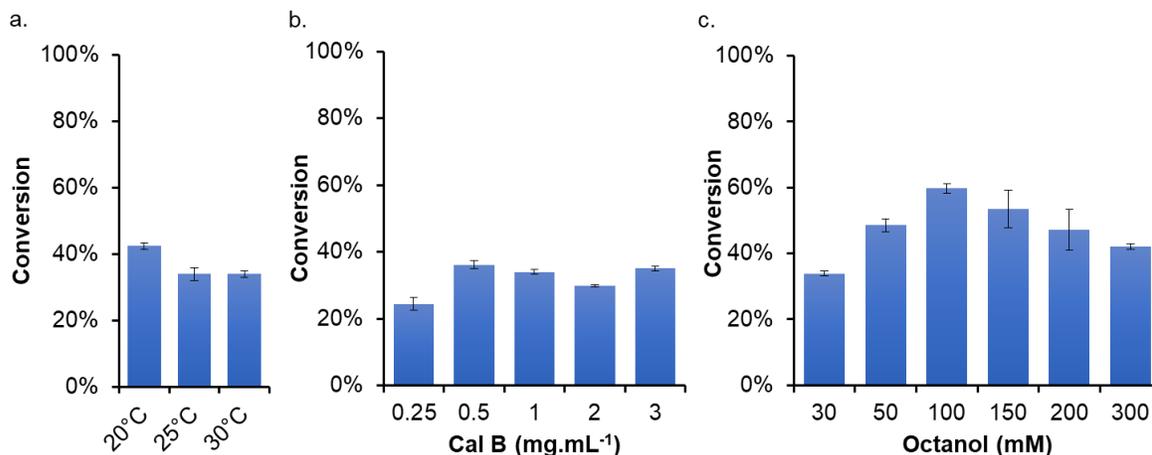


Figure 4-22 Influence of the temperature, enzyme concentration and substrate concentration on % conversion in HPCCC. a. Various temperature, reaction condition: 1600 rpm, 2 mL.min⁻¹, CalB (1 mg.mL⁻¹), octanol (30 mM), vinyl acetate (30 mM). b. Various enzyme concentration, reaction condition: 1600 rpm, 2 mL.min⁻¹, 20 °C, octanol (30 mM), vinyl acetate (30 mM). c. Various substrate concentration, reaction condition: 1600 rpm, 2 mL.min⁻¹, CalB (1 mg.mL⁻¹), octanol (30 mM), vinyl acetate (30 mM).

For further studies, 100 mM octanol was therefore chosen. Subsequently, the three main controllable parameters when using the HPCCC were studied using a DoE method, namely (i) temperature, (ii) residence time, and (iii) rotation speed (see Figure 4-23a). A face-centred design was used and three levels for each parameter were set with the 15 reactions screened in random order. Again 1:1 ratio was chosen to see a change in conversion and an enzyme concentration of 1 mg.mL⁻¹ was arbitrarily chosen. The temperature was varied from 17 °C to 37 °C, a lower temperature was not possible due to the chilling capacity required. The rotation speed of the bobbins was varied from 1200 rpm to 1600 rpm. Lower rotation speed would have resulted in a too low S_f value while 1600 rpm is the maximum rotation speed of the HPCCC. For a consistent study, all experiments were set up at the same S_f value (volume of stationary phase). Therefore, the system was set up at the highest flowrate and lowest rotation speed studied (2 mL.min⁻¹ and 1200 rpm). This resulted in only 7 mL of aqueous phase being retained on the column for each experiment. Lastly, the residence time was varied from 7.5 min to 30 min (see Figure 4-23a).

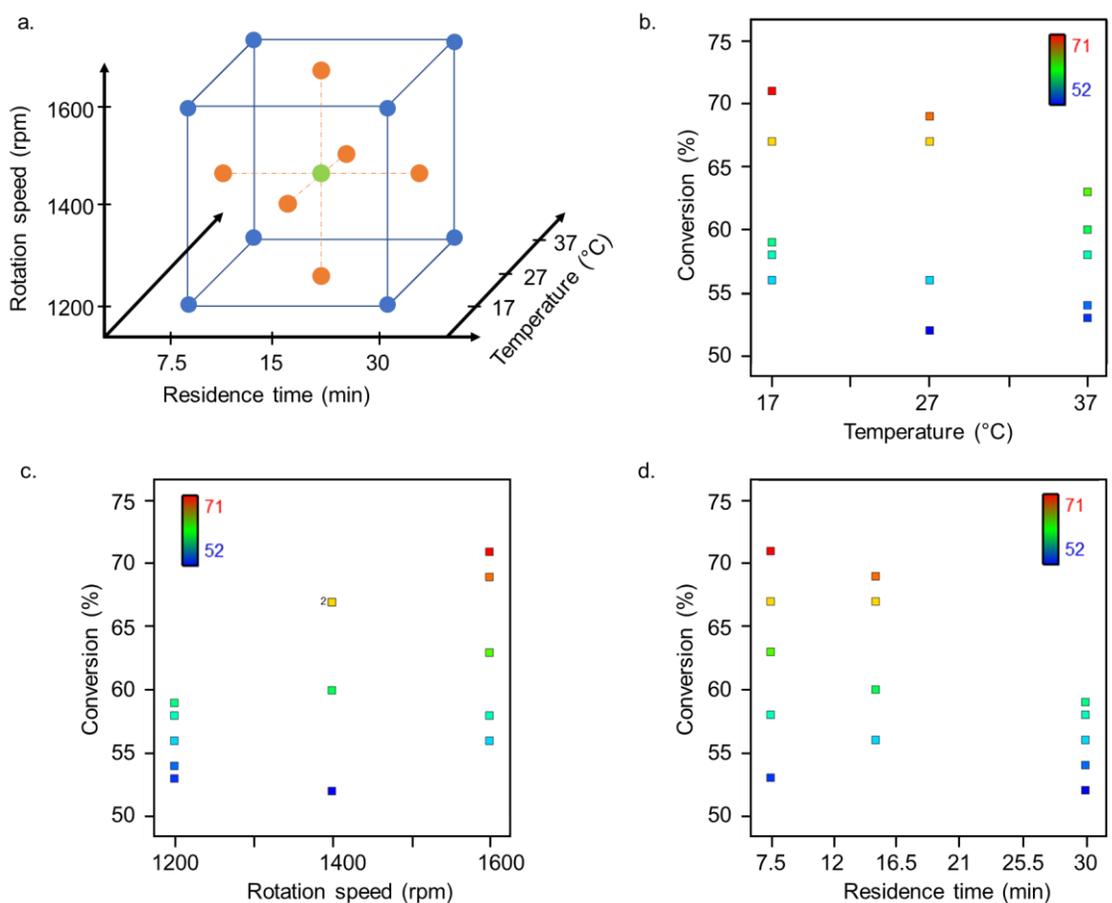


Figure 4-23 a. 3-Level face centered DoE performed. b. Results from the DoE, conversion depending on the temperature. c. Conversion depending on the rotation speed d. Conversion depending on the residence time. (Huyhn F. et al., 2020)

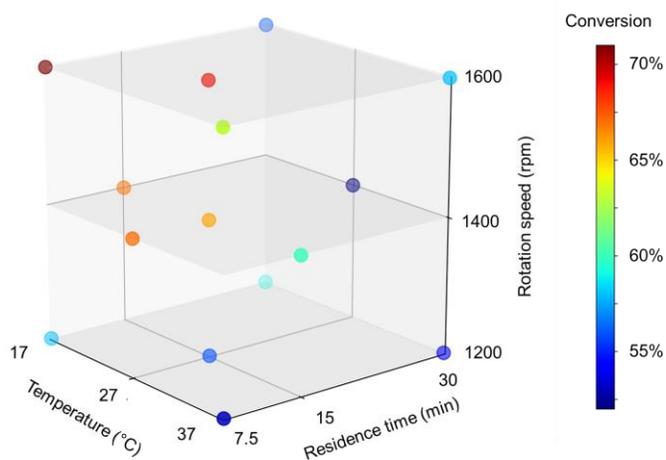


Figure 4-24 Result from the DoE, 3D view, conversion depending on temperature, rotation speed and residence time.

Results showed that temperatures towards 37 °C decreased conversion of octanol to octyl acetate. Longer residence time and low rotation speed of the bobbin sharply decreased the conversion (see Figure 4-23). Figure 4-24 shows the 3D representation of the results from the DoE and clearly indicates that the top left corner of the cube is where conditions are optimal for high conversion that is to say: high rotation speed, low temperature and short residence time. Theoretically, a lower rotation speed leads to weaker centrifugal forces hence higher partition between phases allowing higher interfacial area.^[159,173] A higher rotation speed will create stronger centrifugal forces leading to more separated phases (*i.e.* the upper phase will be located in the inner portion of the coil and lower phase will be in the outer zone). Mixing is, therefore, weaker on high rotation speed. This result indicates a fine balance is to be found to avoid inactivation of the enzyme by the interfacial area between the enzyme and the organic solvent being too high. This was for example observed with sesquiterpene synthases when using a high ratio of pentane:enzyme in the flow segmented system or when the oscillatory segmented flow was trialled on sesquiterpene synthases. The highest conversion (71%) was obtained using the following conditions: 17 °C, 1600 rpm and a residence time of 7.5 min. Compared to the batch reactor, conversion was doubled while time reduced by ten times using equimolar quantities of substrates.

In order to reach full conversion, the quantity of vinyl acetate was increased to obtain 1:3 ratio, residence time was decreased to 6 min by increasing the amount of stationary phase (aqueous phase), while keeping rotation speed at its maximum and temperature in the HPCCC at its lowest. Unfortunately, when running the HPCCC for a long time, the chiller used was only able to keep the temperature to a stable 20 °C due to the heat generated by the rotating bobbins. Therefore, the reaction was run at 20 °C, 1600 rpm with a residence time of 6 min. The organic layer was collected from the third residence time to the sixth and solvent removed under reduced pressure. The crude product was purified by flash chromatography (10% EtOAc in hexanes) to obtain a colourless oil in 97% yield.

Next, reusability of the enzyme was assessed by running the system for 22 cycles at 1600 rpm, 20 °C, 6 min residence time. Fractions were collected regularly, and yield was measured by GC FID. Figure 4-25 shows a decrease in yield between cycle 6 and 8 from >95% to 77% then yield slowly decreased to reach 33% by recycling number 22. Decreasing activity of CalB over time has been reported numerous times, in an immobilised system or free system.^[174-176] The loss in activity may be caused by repeated contact and shear stresses in the HPCCC as well as a natural loss of activity of CalB standing at room temperature. However, the use

of the HPCCC allows the use of CalB in a simple immobilised form without the cost of immobilising the enzyme on a resin while being highly efficient in terms of conversion and reaction time.

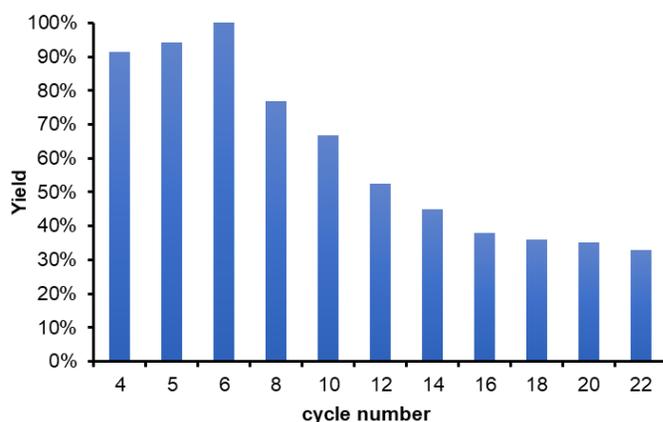
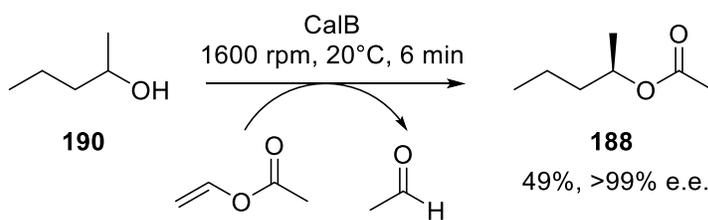


Figure 4-25 Recycling efficiency of CalB for the transesterification of octanol with vinyl acetate in HPCCC (Octanol 100 mM (1:3 ratio, alcohol:acetate), CalB (1 mg.mL⁻¹), 20 °C, 1600 rpm, 6 min residence time). (Huynh F. et al, 2020)

Finally, these optimised conditions were used to study if performing the reaction in the HPCCC system allowed stereoselectivity of CalB to be retained. 2-pentanol (**190**) was chosen as substrate to produce the valuable (*R*)-2-pentyl acetate (**188**) widely used in perfumery and the food industry. (see Figure 4-25). Performing the reaction in heptane resulted in the formation of the desired product acetylated, however, upon removal of the solvent, both (*S*)-pentanol and (*R*)-2-pentyl acetate were lost due to their low boiling point and volatility. The reaction was therefore performed again in pentane as it has a lower boiling point than heptane and pleasingly resulted in 49% isolated yield (> 99% e.e.). Enantiomeric excess was measured by GC FID using a racemic mixture of 2-pentyl acetate synthesised chemically from racemic 2-pentanol (see Figure 4-26). The product was presumed to be (*R*)-2-pentyl acetate (**188**) as CalB is reported to catalyse stereoselectivity the transesterification of (*R*)-alcohols.^[177,178]



Scheme 4-2 Stereoselective transesterification of 2-pentanol by CalB lipase using vinyl acetate

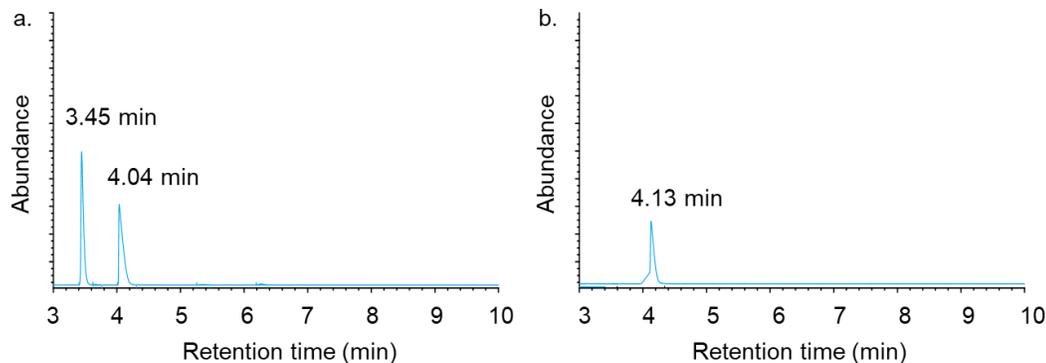


Figure 4-26 a. Flame ionisation detection chromatogram of a racemic mixture of 2-pentyl acetate chemically synthesised 3.45 min (S)-2-pentyl acetate, 4.04 min (R)-2-pentyl acetate. b. Flame ionisation detection chromatogram of the pentane collected at the exit of the HPCCC from the incubation of 2-pentanol with vinyl acetate. (R)-2-pentanol eluted with the solvent.

4.7. SUMMARY

In the chapter, important parameters influencing the conversion of 14,15-dimethyl FDP (**1-4**) by GDS Y406F using a flow segmented system were evaluated. A fractional factorial DoE design showed that a low concentration of FDP was critical for high yields. Upon this discovery, it was decided to pursue other methods as the results indicated an extremely low production per hour (about 1 mg in 56 hours) using the segmented flow approach. The oscillatory segmented flow system was shown to inactivate sesquiterpene synthases due to excessive mixing from the oscillating segments. Finally, the HPCCC was probed as a reaction vessel for sesquiterpene synthases and lipases. Positive results have demonstrated the potential of the high-performance counter-current chromatography system as a novel reactor for biphasic enzymatic reactions. More precisely, it dramatically improved the reaction rate of two classes of enzymes known for product inhibition by an optimised interfacial area between the aqueous and organic phase. Optimisation of both reactions' conditions proved to be possible in a minimum amount of experiments and relatively easy showing the HPCCC has potential for other applications. Sesquiterpene synthases successfully converted FDP (**14**) to high-value compounds in the HPCCC. All synthases retained their regio- and stereo-specificity. Three synthases were tested and yielded up to 99% in a minimum amount of time. On a small scale, the reaction was 10 times faster in the HPCCC than in a flow segmented system. This process was also successful when using unnatural substrates such as the 12-OH-FDP (**111**) with ADS. Production of about 35 mg crude products in less than 2 hours allowed characterisation of the

side product for the first time, identified to be the artemisinic-11S,12-epoxide (**186**). Although this process is not a continuous process like the segmented flow system, it can produce 30 mg of compound in one hour while the flow segmented system would take 350 hours (2 weeks)⁴ to theoretically produce an equivalent amount. The HPCCC system also showed to be as efficient with lipases. In this case, the reactor could operate in a continuous manner as the substrate and product were both soluble in the organic mobile phase. The influence of the rotation speed, temperature and residence time (flow rate) were studied. Interestingly, short residence time and high rotation speed led to higher yields even though a higher rotation speed meant less mixing thus, interfacial area. This proves that a high level of control of the interfacial area between the two immiscible phases is needed to allow efficient mass transfer while avoiding inactivation of the enzyme due to excessive contact to organic solvent. A high temperature led in the same manner to low conversion of the substrate expressing a lower enzyme activity. CalB was reused up to 8 times before a drop in yield below 70%. Similar results are reported in the literature with immobilised CalB. A better understanding of the hydrodynamic behaviour in the HPCCC could increase the recyclability of the enzyme. One other key advantage of the HPCCC over biphasic systems (immobilised CalB or liquid/liquid reaction) is the collection of only the organic phase at the exit of the reactor thus avoiding the separation work up step always generated when using immobilised enzymes or batch biphasic reaction. A separation is indeed sometimes complicated due to the presence of an emulsion, often leading to loss of product as seen previously with sesquiterpene synthases. The development of such a method offers an alternative approach to the traditional low yielding batch process and the flow segmented system developed recently in our group for the production of high-value terpenoids.

⁴ Value calculated from reported results from Tang *et al.*

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

This project aimed to develop new methodologies to improve the chemoenzymatic synthesis of sesquiterpenes from farnesyl diphosphate (FDP) or analogues of FDP and to further expand the sesquiterpene using unnatural FDP analogues. New methodologies were investigated as the chemoenzymatic synthesis of sesquiterpenes is currently not commercially relevant due to its overall low yield. There are two challenging steps in the chemoenzymatic synthesis of sesquiterpenes: the diphosphorylation of farnesol or its analogues, and the biocatalytic conversion of these diphosphate species to a sesquiterpene by a sesquiterpene synthase. Two different methods were investigated to improve the diphosphorylation yield, a chemical approach using a trimethylsilyl (TMS) protected diphosphate and an enzymatic approach using two different enzymes: undecaprenol kinase (UK) and isopentenyl phosphate kinase (IPK) in a one-pot reaction. The improvement of FDP conversion by sesquiterpene synthases was investigated using an oscillatory segmented flow system and by repurposing high-performance counter current chromatography (HPCCC) equipment as a liquid-liquid reactor.

5.1. DEVELOPMENT OF AN EFFICIENT DIPHOSPHORYLATION PROCESS

Chemical diphosphorylation using a trimethylsilyl protected diphosphate species proved to be feasible, however moderate yields were obtained (around 40%). One advantage was the easy scalability of the method, performed on a 4 mmol scale, whereas the Davisson *et al.* method could not be used above 1 mmol scale. Despite being less cumbersome, this alternative chemical procedure showed extremely poor reproducibility.

On the other hand, the development of a one-pot, two-step enzyme cascade was successful. Farnesol was converted to farnesyl phosphate by heterologously expressed UK in *E. coli*. Isopentenyl phosphate kinase was rationally engineered using site directed mutagenesis to accept farnesyl phosphate as substrate. The variant IPK F76A F83A I146A showed extremely good activity toward farnesyl phosphate with >95% conversion to FDP. Four or five residue changes in the IPK active site proved to be detrimental to the activity toward farnesyl phosphate. In a one-pot reaction containing UK and IPK F76A F83A I146A, FDP was synthesised from farnesol in 95% yield in a 0.1 mmol scale. This method was also applied to diphosphorylate 6,15-dimethyl farnesol in 89% yield. Other analogues of farnesol were tested in combination with sesquiterpene synthases to rapidly evaluate the promiscuity of both

kinases. 14,15-Dimethyl farnesol showed promising results but the conversion could not be quantified due to an apparent low activity of GDS Y406F. Pleasingly, 10,11-epoxy farnesol was fully converted to a cyclic sesquiterpene product with Gd4oIS. Less satisfying results were obtained with methoxy substituted farnesols. Since little activity was observed for 8-methoxy farnesol and 12-methoxy farnesol, other variants of IPK were tested. When using UK and IPK F76A F83A I146A with ADS, >50% of 8-methoxy farnesol was converted to 8-methoxy- γ -humulene.

5.2. EXPANDING THE TERPENOME

A library of FDP analogues was synthesised in this work. 6,15-Dimethyl FDP and 12,15-dimethyl FDP were incubated with GDS and variants known to accept 14,15-dimethyl FDP. Unfortunately, these two analogues were not accepted as substrates and were therefore not further investigated. However, the novel enzymatic product arising from the incubation of 10,11-epoxy FDP with Gd4oIS was successfully isolated and fully characterised for the first time. Variable temperature NMR spectroscopy coupled with 2D NMR experiments allowed the full characterisation of the structure; an 11-membered cyclic ether. This illustrates the potential of introducing a functional group to modify the course of the natural sesquiterpene mechanism allowing access to novel sesquiterpenes with potential bioactivities.

5.3. NOVEL METHODOLOGIES FOR SESQUITERPENE CATALYSIS

Segmented flow chemistry was used to improve the enzymatic synthesis of the aphid attractant (*S*)-14,15-dimethyl germacrene D from 14,15-dimethyl FDP using GDS. A 69% yield (measured using GC) was observed after optimisation of the reaction conditions using a design of experiment. Conversely, the oscillatory flow segmented system proved to not be suitable for sesquiterpene synthases due to heavy enzyme precipitation over the course of the incubation in the tubing; this is likely due to mixing being too vigorous leading to denaturing of the enzyme from excessive exposure to the organic phase.

The HPCCC has demonstrated it to be an excellent novel reaction vessel for sesquiterpene synthases by enabling efficient biocatalysis through continuous extraction of the sesquiterpene. Traditionally used for liquid-liquid chromatography, this system offers a

platform to immobilise the enzyme and the substrate together in an aqueous phase while running an organic solvent through it to continuously extract the product. GC yields of >80% were obtained when incubating FDP with ADS, AS or GDS in the analytical column of the HPCCC (22 mL reactor, 11 min reaction time). Subsequently, isolated yields were obtained by using the preparative column (135 mL, 20 min residence time), and around 30 mg of each sesquiterpene product was isolated (from the combination of four experiments in the preparative column), namely: (1R,7R,8aS)-aristolochene (94% isolated yield from FDP), (1R,4R,4aS,8aR)-amorpha-4,11-diene (84% from FDP), (S)-germacrene D (92% from FDP) and dihydroartemisinic aldehyde (70% from 12-OH FDP). Thanks to this efficient process, the side product arising from the incubation of 12-OH FDP with FDP was for the first time successfully isolated and characterised by NMR spectroscopy. It was identified to be the artemisinic-11S,12-epoxide instead of the previously hypothesised enol form of dihydroartemisinic aldehyde. This methodology was also successfully applied to lipases (Cal B in this work). Since both substrates and products are soluble in the organic phase, this enzymatic reaction was able to run as a continuous process offering a useful alternative to the traditional enzyme immobilisation on a resin. Octyl acetate was produced in 97% isolated yield from octanol and CalB could be reused for at least 8 cycles with >75% yield. The activity of the enzyme slowly decreased to reach 33% yield by cycle 22. The stereoselectivity of CalB was unaffected using this method as (*R*)-2-pentylacetate was generated from a racemic mixture of 2-pentanol in 49% yield (>99% e.e.).

5.4. FUTURE WORK

This work has demonstrated the possibility to create unnatural terpenoids from FDP analogues and has also identified and addressed the two key challenges of the chemoenzymatic approach for the synthesis of sesquiterpenes. Further research should be conducted to widen the scope of analogues accepted by IPK; obtaining a crystal structure of the variant IPK F76A F83A I146A with FDP bound to the active site would allow a better understanding of the folded conformation of FDP.

The advantages and potential of the HPCCC apparatus as a reactor for enzymatic reactions was clearly illustrated in this work. Further studies need to be undertaken to identify a way to use the HPCCC for sesquiterpene synthases in a continuous process similar to current immobilised enzymes. For this purpose, the HPCCC could be used in a co-current elution mode where the two phases are pumped in the reactor simultaneously (in a similar manner

as the segmented flow system). The use of “dual-mode” could also be explored (where mobile and stationary phases are switched during the experiment). This mode would allow feeding of a new solution containing new substrate and enzyme and is defined as intermittent countercurrent extraction (ICcE). A dual-flow mode is also possible where each phase is pumped in the system simultaneously from opposite directions inside the column. However, this mode requires modification of the HPCCC apparatus.

The possibility of combining the one-pot enzymatic diphosphorylation and the sesquiterpene synthase step in the HPCCC also warrants further investigation and could prove to be beneficial for upscaling the chemoenzymatic synthesis of natural and unnatural sesquiterpenes.

In addition, further research should examine the HPCCC as a potential vessel for organic synthesis. As an example, organic reactions involving two phases, such as a catalyst or an immiscible side product such as water in an esterification, could be studied.

CHAPTER 6

MATERIALS AND METHODS

6.1. BIOLOGICAL METHODS

6.1.1. Materials

All chemicals were purchased from Sigma-Aldrich, Apollo Scientific UK, Melford Laboratories Ltd or Fisher Scientific Ltd and were used without further purification. Restriction enzymes, T4 ligase, Golden gate assembly mix and Gibson Assembly Mix were purchased from New England Biolabs UK Ltd (Hitchin, United Kingdom). A prestained protein size marker (14.4-116.0) kDa was used to identify proteins by 12% SDS-gel. PrimeSTAR polymerase was purchased from Takara Bio Europe (Saint-Germain-en-Laye, France). Oligonucleotide primers for site-directed mutagenesis were purchased from Sigma-Aldrich, the pET16b-SC1 plasmid construct was a generous gift from Prof. Yasuo Ohnishi, University of Tokyo. All the mutated and ligated constructs were confirmed by DNA sequence analysis using Eurofins.

6.1.2. Media

Luria-Bertani (LB) media was prepared by dissolving tryptone (10 g), yeast extract (5 g) and sodium chloride (10 g) in deionised water (1 L) and pH was adjusted to 7.4 with sodium hydroxide (1 M). Media was autoclaved at 121 °C for 15 minutes before usage.

Terrific broth (TB) was prepared by dissolving tryptone (12 g), yeast extract (24 g) and glycerol (4 mL) in deionised water (900 mL). In a separate flask, a solution of monopotassium phosphate (0.17 M) and dipotassium phosphate (0.72 M) in deionised water (100 mL) was prepared. Both solutions were autoclaved at 121 °C for 15 minutes and the solutions were combined at the sterile bench immediately before usage.

Media was cooled to room temperature before use. Any required antibiotic was added immediately before use under aseptic conditions.

6.1.3. Antibiotics stocks

Antibiotics were prepared as sterile stock solutions (mg/mL). Solid antibiotics were weighted out and dissolved in deionised water (10 mL). Solutions were filter-sterilised (0.2 µm) under aseptic conditions and aliquoted out into sterile Eppendorfs (1 mL). Antibiotics were then

diluted 1:1000 into media to yield the correct concentrations: ampicillin (100 µg/mL), kanamycin (50 µg/mL), and tetracycline (15 µg/mL). Stock solutions were stored in -20 °C freezer and thawed on ice prior to use.

6.1.4. Agar plate preparation

Agar solution for plates was prepared by dissolving agar (7.5 g), tryptone (5 g), yeast extract (2.5 g) and sodium chloride (10 g) in deionised water (500 mL). The suspension was autoclaved at 121 °C for 15 minutes. The resulting solution was allowed to cool down to approximately 40 °C before adding the required antibiotics and pouring it into sterile plates under aseptic conditions. Plates were allowed to cool and set before storage in the cold room at 4 °C until required.

6.1.5. Bacterial strains

A number of *E. coli* strains were used in the course of this work. XL1-Blue cloning strain was employed for DNA amplification and cloning. *E. coli* BL21(DE3), BL21-CodonPlus(DE3)-RP, BL21-CodonPlus-RIL, C41(DE3) pLysS, BL21 AI, BL21 cell lines were used for test expression and protein expression.

6.1.6. Competent cells

Calcium chloride buffer I

Calcium chloride (100 mM) was dissolved in deionised water and autoclaved at 121 °C for 15 minutes. The resulting solution was stored at 4 °C prior to use.

Calcium chloride buffer II

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

Protocol

An aliquot of the cell strain (50 µL) was incubated in LB media (50 mL) overnight at 37 °C in the absence of antibiotics, except for XL1-Blue cells for which the media was supplemented

with tetracycline. An aliquot (1 mL) of the overnight culture was used to inoculate fresh LB media (100 mL) and the resulting culture was grown at 37 °C until the optical density at 600 nm (OD_{600}) reached 0.6 for the cloning strain (XL1-Blue), or 0.9-1.2 for the expression strains. Cells were chilled on ice for 15 to 30 minutes, and then distributed between two sterile 50 mL flacon tubes at the sterile bench.

Cells were harvested by centrifugation (3,400 g, 10 minutes). The supernatant was discarded, and the cells were gently resuspended in calcium chloride buffer I (5 mL). Cells were incubated on ice for an additional 15 minutes and then harvested by centrifugation (3,400 g, 10 minutes). The supernatant was discarded, the cell pellet was gently resuspended in calcium chloride buffer II (5 mL), and cells were incubated on ice for 10 minutes. The cells were divided into aliquots (50 μ L) into pre-chilled 1.5 mL Eppendorf tubes and flash-frozen in liquid nitrogen. Cells were stored at -80 °C until required.

6.1.7. Supercompetent cells

Rubidium chloride buffer I

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

Rubidium chloride buffer II

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

Protocol

Supercompetent cells were prepared in the same manner as competent cells (6.1.6) but with rubidium chloride buffers I and II in place of calcium chloride buffers I and II respectively.

6.1.8. Gene sequence

Plasmids used in this work were taken from the plasmid library of Prof. Allemann or purchased as gene sequences from ThermoFisher Scientific. A list of all plasmids used is found in Table 6-1 and sequence of the gene of interest used in Table 6-2.

Table 6-1 List of plasmids used in this project.

Plasmid reference	Gene	Vector	Antibiotic resistance
Plasmid 1	His _{x6} -Thrombin-UK	pET28a	Kanamycin
Plasmid 2	Thioredoxin- His _{x6} -Thrombin-STag-UK	pET-32 Xa/LIC	Ampicillin
-	His _{x6} -Thrombin-IPK	pET28a	Kanamycin
-	His _{x10} -GdoIS	pET16b	Ampicillin
-	ADS- His _{x6}	pET21d	Ampicillin
-	AS	pLM1	Ampicillin
-	GDS- His _{x6}	pET21d	Ampicillin

Table 6-2 Gene sequences of protein used in this project. His Tag underlined. *Gene synthesised by Thermo Fisher including necessary BsaI recognition sequence (underlined), and overhangs (in bold) for Golden Gate cloning.

Name	Sequence
UK*	<p>CATGGTCTCAAGTCATGCCGATGGATCTGCGTGATAATAAACAGAGCCAG AAGAAATGGAAAAATCGTACCCTGACCAGCAGCCTGGAATTTGCACTGAC CGGTATTTTTACCGCCTTTAAGAAGAACGCAACATGAAAAACATGCCGT TAGCGCACTGCTGGCAGTTATTGCAGGTCTGGTTTTTAAAGTTAGCGTGAT CGAATGGCTGTTTCTGCTGCTGAGCATTTTTCTGGTTATCACCTTTGAAATT GTGAACAGCGCCATTGAAAATGTTGTTGATCTGGCAAGCGATTACCATTTT AGCATGCTGGCAAAAAACGCCAAAGATATGGCAGCCGGTGCAGTTCTGGT</p>

GATTAGCGGTTTTGCAGCCCTGACCGGTCTGATTATCTTTGTTCCGAAAAT
 TTGGTTCCTGCTGTTTCACTAATGATTGAGACCATG

IPK ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCG
 CGGCAGCCATATGGAAAACCTGTATTTTCAGTCCATGCTGACCATTCTGAA
 ATTAGGTGGTAGCATTCTGAGCGATAAAAATGTGCCGTATAGCATCAAATG
 GGATAACCTGGAACGTATTGCCATGGAAATCAAAAATGCCCTGGACTACTA
 TAAAACCAGAACAAGAGATCAAGCTGATTCTGGTTCATGGTGGTGGTGC
 ATTTGGTCATCCGGTTGCCAAAAGTATCTGAAAATTGAGGACGGCAAAAA
 AATCTTCATCAACATGGAAAAGGCTTTTGGGAAATTCAGCGTGCAATGCG
 TCGTTTTAACAACATTATTATCGATACCCTGCAGAGCTATGATATTCCGGCA
 GTTAGCATTACGCCGAGCAGCTTTGTTGTTTTTGGTGATAAACTGATCTTTG
 ACACCAGCGCCATTAAGAAATGCTGAAACGTAATCTGGTTCCGGTGATTG
 ATGGTGATATTGTGATCGATGATAAAAACGGCTATCGCATTATTAGCGGTG
 ATGATATTGTTCCGTATCTGGCCAATGAACTGAAAGCAGATCTGATTCTGTA
 TGCCACCGATGTTGATGGTGTCTGATTGATAACAAACCGATCAAACGCAT
 CGACAAGAACAACATCTATAAAATCCTGAACTATCTGAGCGGCAGCAATAG
 CATTGATGTTACCGGTGGTATGAAATACAAGATCGATATGATCCGCAAAAA
 CAAATGCCGTGGCTTTGTGTTAATGGCAATAAAGCCAACAACATTTACAA
 AGCACTGCTGGGTGAAGTTGAAGGCACCGAAATTGATTTTAGCGAGTAA

GDS ATGGCAGCGAAACAGGGCGAGGTTGTGCGCCAGATGCAGATTATAGTTA
 TCATCCGAGTCTGTGGGGCGATCAATTCCTGCATTACGATGAACAAGAAGA
 CGACCAGGTGCGAGGTGGATCAGCAGATTGAAATTCTGAAAGAGGAGACTC
 GTAAAGAGATTCTGAGTAGCCTGGATGATCCGGCTAAACACACCAACCTG
 CTGAAACTGATTGATGTTATCCAGCGTCTGGGCATCGCTTATTACTTTGAA
 CACGAAATTACCCAGGCGCTGGGCCATATCTATAACGTTTACGGTGACGA
 ATGGAACGGCGGCTCGACCTCGCTGTGGTTTCGCCTGCTGCGTCAGCAG
 GGTTTCTATGTAAGCTGTGATATCTTTAACATCTACAACTGGATAATGGCA
 GCTTTAAGGATAGCCTGACCAAAGACATTGAGTGATGCTGAACTGTATG
 AAGCTGCGTATATGCGTGTTACAGGGCGAAATCATTCTGGATGAAGCGCTG
 GAATTTACCAAACCCACCTGGAGCAGATTGCAAAAAGATCCACTGCGTTGT
 AATAATACGCTGAGCCGTCACATTTACGAAGCCCTGAAACGCCCGATTGCG
 CAAACGTCTGCCGCGCGTGCATGCACTGCAATATATGCCGTTTTACGAGC
 AACAAGATTCGCACAATAAAAGTCTGCTGCGTCTGGCTAACTGGGTTTCA
 ACCGCCTGCAAAGTCTGCATAAGAAAGAACTGAGCCAGCTGTGCAAGTGG

TGGAAGGAATTTGACGCGCCGAAGAATCTGCGTTATGTGCGTGACCGTCT
 GGTAGAGCTGTACTTTTGGGTAAGTGGGCGTTTATTTTGAACCGCAATACAG
 CCGTAGTCGCATTTTCTGACCAAGGTAATCAAATGGCGACCATTCTGGA
 TGACACCTATGATATTCACGGGACGTATGAAGAAGTGGAGATCTTTACCAA
 GGCAGTGCAGCGCTGGAGCATTACCTGCATGGATACCCTGCCGGATTACA
 TGAAAATGATCTACAAGAGCCTGCTGGACGTTTACGAAGAGATGGAAGAAA
 TTATTGACAAAGATGGTAAAGCGTATCAGGTGCACTATGCGAAAGATAGCA
 TGATCGATCTGGTGACCAGCTATATGACTGAAGCCAAATGGCTGCATGAA
 GGTCATGTTCCAACGTTTGGAGGAATATAATAGCATCACGAATCTGACGGGT
 GGCTATAAAATGCTGACGACGAGCAGCTTCGTGGATATGCCGGGTGATAT
 TGTGACGCAGGAGAGTTTCAAATGGGCGCTGAACAATCCGCCTCTGATTA
 AAGCGTCGGCAGATGTCTCGCGCATTATGGACGACATCGTTGGGCATAAA
 GAAGAACAGCAACGCAAACATCTGCCTTCGCGTGTGGAATGTATATGAA
 GAAATACCATCTGGCCGAAGAGGATGTGTATGATCTGCTGAAACAGCGCG
 TCGAAGACGCCTGGAAAGACCTGAACCGCGAAACTCTGACTTGCAAGGAC
 ATCCATATGGCCCTGAAAATGCGCCCAATCAACCTGGCCCGCGTGATCGA
 CATGCTGTACAAGAATGACGACAACCTGAAGAACGTCGGTCAAGAAATCC
 AAGACTACATCAAGTCGTGTTTCATCAACGCCATCAGTGTCAAGGATCCGA
 ATTCGAGCTCCGTCGACAAGCTTGCGGGCCGCACTCGAGCACCACCACCAC
CACCACTGA

ADS ATGGCCTTGACTGAAGAGAAACCTATAAGGCCAATTGCAAATTTCCACCT
 TCTATTTGGGGCGATCAATTTTTGATTTATGAGAAACAAGTTGAACAGGGT
 GTGGAGCAAATAGTAAACGATCTAAAGAAGGAAGTAAGACAGTTGTTAAAG
 GAAGCATTGGATATTCCTATGAAACATGCAAATTTGTTGAAGCTGATTGAC
 GAGATTCAACGTTTAGGTATTCCGTATCATTTTGAACGTGAAATTGATCATG
 CATTGCAATGTATTTACGAGACCTATGGTGATAATTGGAATGGCGACAGGT
 CTAGCTTATGGTTTAGGCTGATGCGTAAACAAGGATACTATGTCACGTGTG
 ATGTGTTTAATAACTATAAAGACAAGAATGGTGCTTTTAAACAATCGTTAGC
 GAATGATGTTGAAGGATTGTTGGAATTATATGAGGCTACGTCCATGAGAGT
 TCCGGGCGAAATAATTCTTGAAGATGCCCTGGGATTCACAAGATCAAGGCT
 ATCGATTATGACAAAGGACGCGTTTAGTACAAACCCCGCTTTATTCACTGA
 AATCCAGAGAGCTTTAAAGCAACCATTGTGGAAGAGATTGCCAAGGATCGA
 GGCCGCCAGTACATACCCTTCTATCAGCAACAAGACTCCCATAATAAAAC
 TCTGCTAAAGTTAGCTAAACTGGAGTTCAATCTATTGCAGAGCCTACATAA

GGAAGAGTTGAGTCACGTATGCAAGTGGTGGGAAGGCATTTGATATTAAGAA
 AAATGCCCCGTGCCTGCGTGACCGTATCGTTGAATGTTACTTCTGGGGTCT
 GGGTTCTGGTTATGAACCACAGTACTCCCGTGACGTGTGTTCTTCACTAA
 AGCTGTAGCTGTTATCACCTGATCGATGACACTTACGATGCTTACGGCAC
 CTACGAAGAAGTGAAGATCTTTACTGAAGCTGTAGAACGCTGGTCTATCAC
 TTGCCTGGACACTCTGCCGGAGTACATGAAACCGATCTACAACTGTTTAT
 GGATACCTACACCGAAATGGAGGAATTCCTGGCAAAGAAGGCCGTACCG
 ACCTGTTCAACTGCGGTAAAGAGTTTGTAAAGAATTCGTACGTAACCTGA
 TGGTTGAAGCTAAATGGGCTAACGAAGGCCATATCCCGACTACCGAAGAA
 CATGACCCGGTTGTTATTATCACCGGTGGTGCAAACCTTGCTAACAACCTACC
 TGTATCTGGGTATGTCCGACATCTTTACCAAGGAATCTGTTGAATGGGCT
 GTTTCTGCACCGCCGCTGTTCCGTTACTCCGGTATTCTGGGTGTCGTCTG
 AACGACCTGATGACCCACAAAGCAGAGCAGGAACGTAAACACTCTTCCTC
 CTCTCTGGAATCCTACATGAAGGAATATAACGTTAACGAGGAGTACGCACA
 GACTCTGATCTATAAAGAAGTTGAAGACGTATGGAAAGACATCAACCGTGA
 ATACCTGACTACTAAAACATCCCGCGCCCGCTGCTGATGGCAGTAATCTA
 CCTGTGCCAGTTCCTGGAAGTACAGTACGCTGGTAAAGATAACTTCACTCG
 CATGGGCGACGAATACAAACACCTGATCAAATCCCTGCTGGTTTACCCGAT
 GTCCATCTGTCCCGGGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGG
 CCGCACTCGAGCACCACCACCACCACCCTGA

AS ATGGAAAACCTGTATTTTCAGTCCATGGCTACCTCAACAGAAACCATTTCTT
 CCCTGGCCCAACCGTTCGTGCACCTTGAAAACCCTATCAATAGCCCTCTG
 GTCAAAGAGACAATCAGGCCAGAAATGACACGACGATCACTCCGCCTCC
 TACTCAGTGGTCGTACCTTTGCCATCCACGAGTGAAGGAGGTACAGGACG
 AAGTCGATGGATACTTCTGGAGAACTGGAAATTTCCAGCTTCAAGGCTG
 TTCGCACCTTCTTGATGCCAAGTTCTCAGAGGTTACTTGTCTTTTCTTCCC
 TCTTGCACTGGACGATCGCATCCACTTTGCCTGCCGACTGCTGACCGTTCT
 CTTCTTATCGATGATGTTCTTGAGCATATGTCTTTCGCGGATGGAGAAGC
 CTACAACAACAGATTGATCCCGATATCGCGTGGAGACGTGCTCCCGGACC
 GAACCAAGCCAGAAGAGTTCATTCTCTATGACCTCTGGGAAAGCATGCGC
 GCCCATGATGCGGAGCTGGCCAACGAAGTTCTTGAGCCGACTTTCGTGTT
 CATGCGCGCGCAGACAGATCGAGCGGTTTGGAGCATCCATGAATTGGGG
 CATTATCTCGAGTACCGTGAGAAGGATGTAGGCAAGGCGTTGCTTTCGGC
 TCTAATGAGATTCTCGATGGGACTTAGACTCAGTGCAGATGAGCTTCAGGA

TATGAAAGCCCTTGAAGCCAACTGTGCCAAGCAGCTTTCTGTAGTCAATGA
 CATATACAGTTATGACAAGGAAGAGGAAGCGTCTCGGACTGGACACAAGG
 AGGGAGCCTTCCTTTGTTTCAGCTGTGAAGGTTCTGGCGCAGCTTTCTGTAG
 TCNATGACNTATACAGTTATGACAAGGAAGAGGAAGCGTCTCGGACTGGA
 CACAAGGAGGGAGCCTTCCTTTGTTTCAGCTGTGAAGGTTCTGGCGGAGGA
 GTCCAAGCTTGGTATACCCGCAACGAAACGCGTGCTTTGGTCTATGACTC
 GGGAGTGGGAGACTGTGCATGACGAGATCGTGGCAGAGAAGATCGCATC
 CCCAGACGGCTGCTCTGAAGCTGCCAAGGCGTATATGAAGGGCCTAGAGT
 ACCAGATGAGCGGCAATGAGCAGTGGAGCAAGACCACGCGTATGTACAAC
 TAA

GdoIS ATGGGCCATCATCATCATCATCATCATCATCACAGCAGCGGCCATATC
 GAAGGTCGTCATATGATGTCCGACGACACCTCACTTGAGCTTCCGTTACC
 CACCGCCGCAACCCCATCAGACCGAGGCCGCCGATCGCCATCTCGAAT
 GGCTCCAACGCCACCGCGAGCTGGCCGCCGTCGTCAGCGGATCAACCTA
 CACCGGTTGGGACATCACGGAACCTCGCCTCCCTGGTCTACCCCGAGAGCT
 CCGCCGAGGATCTGGCCCTGGCCGCCGACCTCATGGGGTTCTACTTCCTC
 TTCGACGACCAGTTTCGACAGCCCGCTCGGGCGCCGCCCGAGCAGGTGG
 CCCTGATCTGCGAACGGCTCTCCGCCATCGCGCACGGCACCCCTCACGGC
 TGTCACCTCACCTCCGAACGCGCCTTCGCCGACCTCTGGCGGCGCATCA
 CCCTTGGCATGACGGACCGTTGGCGCGCACGGGCCGCCTGCAACTGGGA
 GTACTACTTCGCCTGCCACCCCGCGGAGGCCGCCGGCCGGACCATCGGG
 CAGCCGCCGGACCGGGAGGGCTATCTGACGCTCAGACGCGGTACCGCCG
 CGATGGAGAGCATCTTCGACATGATCGAGCGGCTCGGCCACTTCGAGGTG
 CCCAGCACGTCATGCACCACCCGCTGTTCCGGCAGCTTCGCCAACTCGC
 GGCCGACATCCCGTCGTTACCAATGACGTGCGCTCCTTCGCACAGGAGT
 CCGAGCGCGGCGACGTGGCCAACCTGGTGTGATGATCGTCCGGCGGGACCG
 CTGCTGCTCCACCGAGGAGGCCTGCGCGGTGGTCTGGGACGAGGCCAG
 CGCATGGCCGACCGGTTCTGCGACCTGCGCGACCAACTCCCGGACGCCT
 GCCGCTCGATGTCGCTGGACCCGGCTCAGCGGCTGGCCGCCGAGCGCTA
 CGCCGACGGCATGGCGCTCTGGCTCGCGGGTTACCTCCACTGGGAGTCC
 CACACCCGCCGCTACCACCACGGCTGA

6.1.9. Cloning

Site directed mutagenesis

Primers design

Oligonucleotide primers were designed in the online software Benchling (<https://benchling.com>) following the criteria listed below when possible:

- The melting temperature (T_m) should be lower than 72 °C, ideally between 40 – 60 °C.
- Primers should begin and terminate with two consecutive G or C, and ideally have a GC content of 40 to 60 %. However, there should be no more than 3 G or C in the first 5 and last 5 bp.
- Primers should not be longer than 60 base pairs (bp). Ideally designed primers were 20-45 bp-long
- Primers should contain at least 15 consecutive base pairs matching the plasmid to be amplified in the first cycle. Primers for SDM should contain the mutation in the centre of the primer and contain ideally 2 bp mismatch.
- T_m of forward and reverse primer should be in the same range

Primers were analysed in the online software Benchling (<https://benchling.com>) to check T_m , primer length, mispriming and primer secondary structures.

Primers used for the insertion of one base pair into the **Plasmid 1** to obtain UK gene in frame are shown in Table 6-3. Primers used to perform the site-directed mutagenesis of IPK are shown in Table 6-5, variant containing 3 or more mutations were obtained from consecutive SDM. Each mutation was confirmed by sequencing using the forward (T7 promoter) or the reverse primer (T7 terminator).

Table 6-3 6-4 Oligonucleotide primers for the insertion of 1 base pair into plasmid containing UK gene in pET28a (5' to 3'). Added base pair depicted in bold.

Name	5' Sequence 3'
UK FWD	CAGTCTATGCC G ATGGATCTGCGTGATAATAAACAGAGCC
UK REV	CATCGGCAT A GAAGTAAAATACAGGTTTTCCATATGGCTGC

Table 6-5 Oligonucleotide primers for site-directed mutagenesis (5' to 3') of IPK WT with the changed codon in bold and underlined.

Name	5' Sequence 3'
F83A FWD	GGAAAAGGC <u>GCG</u> TGGGAAATTCAGCGTGCAATGCGTC
F83A REV	CC <u>CGC</u> GCCTTTTTCCATGTTGATGAAGATTTTTTGGCCGTCCTC
F76A F83A REV	GGAAAAGGC <u>GCG</u> TGGGAAATTCAGCGTGCAATGCGTC
F83A I86A FWD	GGAAAAGGC <u>GCG</u> TGGGAA <u>GCT</u> CAGCGTGCAATGCG
I146A I156A FWD	GTGAT <u>GCT</u> GTGATCGATGATAAAAACGGCTATCGC <u>GCG</u> ATTAGC
I146A REV	CATCGATCAC <u>AGC</u> ATCACCATGAATCACCGGAACCAGATTACG
I156A FWD	CTATCGC <u>GCG</u> ATTAGCGGTGATGATATTGTTCCGTATCTGG
I156A REV	GCTAAT <u>CGC</u> GCGATAGCCGTTTTTATCATCGATCACAGC

Table 6-6 List of all variant created using SDM from primers in Table 6-5.

IPK variant	
n°	n°
a F83A	2 F76A F83A I146A
b I146A	3 F76A F83A I156A
c I156A	4 F83A I86A I146A
d F76A F83A	5 F83A I86A I156A
e F83A I86A	6 F83A I146A I156A
f F83A I146A	7 F76A F83A I86A I146A
g F83A I156A	8 F76A F83A I86A I156A
h I146A I156A	9 F76A F83A I146A I156A
1 F76A F83A I86A	10 F83A I86A I146A I156A

Site directed mutagenesis protocol

Site-directed mutagenesis (SDM) was performed using the same cycle criteria as for PCR reactions (see Table 6-8). Biometra Thermocycler T-Gradient Thermoblock thermocycler was used. PrimeSTAR thermophilic polymerase was used. Two different reaction mixtures were performed in parallel (see Table 6-7).

Table 6-7 PCR mixture for SDM protocol

Reaction components	Protocol 1	Protocol 2
	V (μ L)	V (μ L)
DNA template (5 ng/ μ L)	1	2.5
Primer FWD ^a (125 ng/ μ L)	0.5	0.5
Primer REV ^b (125 ng/ μ L)	0.5	0.5
PrimeSTAR polymerase	12.5	12.5
Master mix ^c		
diH ₂ O	10.5	9.0

^a FWD = forward (5' to 3'). ^b REV = reverse (3' to 5'). ^c Master mix contains dNTPs (deoxyribose nucleotide triphosphate) and the polymerase in the appropriate buffer with MgSO₄.

Table 6-8 PCR for SDM conditions

Step	Temperature (°C)	Time (s)	Number of cycles
Initial denaturing	96	180	1
Denaturing	96	60	
Annealing	60	60	33
Extending	72	720	
Final extension	72	1800	1

The result of the PCR reaction was always digested with DpnI (0.5 μ L) for 15 minutes at 37 °C to eliminate the parental plasmid from the mixture. No further purification was performed for SDM applications, and 10 μ L of the reaction mixture was used to transform XL1-Blue cells.

PCR protocol

PCR was performed using the protocol described in Table 6-10. Thermocycler T-Gradient Thermoblock thermocycler was used. PrimeSTAR thermophilic polymerase was used.

Table 6-9 PCR mixture

Reaction components	V (μL)
DNA template (50 ng/ μL)	1
Primer FWD ^a (10 μM)	0.5
Primer REV ^b (10 μM)	0.5
PrimeSTAR polymerase	25
Master mix ^c	21
diH ₂ O	2
DMSO ^d	2

^a FWD = forward (5' to 3'). ^b REV = reverse (3' to 5'). ^c Master mix contains dNTPs (deoxyribose nucleotide triphosphate and the polymerase in the appropriate buffer with MgSO₄). ^d DMSO = dimethyl sulfoxide is added to avoid mispriming for long primers, volume added depends on the reaction.

Table 6-10 PCR conditions

Step	Temperature ($^{\circ}\text{C}$)	Time (s)	Number of cycles
Initial denaturing	96	120	1
Denaturing	96	30	
Annealing	60	30	33
Extending	72	60 per 1 kbp ^a	
Final extension	72	600	1

^a Extension time depends on the size of the desired PCR product, around 60 seconds per 1 kbp (kilo base pair).

The result of the PCR reaction was always digested with DpnI (0.5 μL) for 15 minutes at 37 $^{\circ}\text{C}$ to eliminate the parental plasmid from the mixture. No further purification was performed for SDM applications, and 5 μL of the reaction mixture was used to transform XL1-Blue cells. DNA was purified by gel extraction from a 1% agarose gel.

6.1.10. Agarose gel

Tris-acetate EDTA buffer (TAE, 50x) was prepared by adding Tris-base (242 g), glacial acetic acid (57.1 mL) and EDTA solution (10 mL, 0.5 mM, pH 8.0) in deionised water (diH₂O, 1 L). TAE buffer (1x) was prepared by diluting 50x of TAE (20 mL) in diH₂O (1 L). Agarose gels were prepared by suspending agarose (1 g) in TAE buffer (1x, 100 mL) in a conical flask and heating in a microwave until boiling. SYBR Safe dye (5 μL) was then added to the mixture to allow gel imaging and DNA extraction with blue light. The solution was then poured into a gel casting tray (Bio-Rad Sub Cell GT Mini) and was left to set for 30 minutes at room temperature. Gels were then submerged in 1x TAE buffer in the gel tank (Bio-Rad) before loading the samples. DNA ladder (Melford), 10 kbp was loaded in the gel as reference. FastDigest Green

buffer (x10, ThermoFisher Scientific) was added to the DNA samples (1 μ L per 10 μ L of sample) and they were loaded into the gel. Gels were run at 100 V for 40 minutes then imaged using a ChemiDoc Gel Imaging System (Bio-Rad).

6.1.11. Gel extraction

Bands of interest (visualised using ChemiDoc Gel Imaging System (Bio-Rad)) were cut from the gel with a scalpel and transferred to a weighted 1.5 mL Eppendorf microcentrifuge tube. Gel slices were dissolved in QG buffer (300 μ L/100 μ g gel) and incubated at 50 °C for 10 minutes. PE buffer (100 μ L/100 μ g gel) was added and mixed by inverting the tube several times. The mixture was applied to a Qiagen Spin Miniprep Column. Columns were spun down at 16,110 g for 60 s in an Eppendorf benchtop microcentrifuge. The flow-through (FT) was discarded and column was washed with 750 μ L PE buffer and spun down for 60 s. The FT was again discarded, and columns were spun down for 60 s to completely remove PE buffer. The column was loaded on top of a new 1.5 mL microcentrifuge tube and autoclaved diH₂O (50 μ L) was loaded on top of the column membrane. Columns were spun for 60 s to obtain the purified DNA. DNA concentrations were determined by absorbance at 260 nm using a Nanodrop 1000 UV Visible spectrophotometer (ThermoFisher Scientific).

Table 6-11 Recipe of Qiagen DNA purification buffers.

QG	P1	P2	N3	PE
20 mM Tris-HCl	50 mM Tris-HCl	200 mM NaOH	0.9 M KOAc	10 mM Tris-HCl
5.5 M guanidine thiocyanate	10 mM EDTA	1% SDS	4.2 M guanidine hydrochloride	80% ethanol
-	RNase A	-	-	-
-	100 μ g/ml	-	-	-
pH 6.6	pH 8.0	-	pH 4.8	pH 7.5

6.1.12. Gibson assembly

The Gibson assembly method (Figure 6-1) is a technique relying on three different enzymes. First, an exonuclease cuts approximately 20 bp (base pairs) in the 5' to 3' direction exposing the complementary sequence: a sticky end. 20 bp of the 5' end of one of the sections is required to be complementary to the 3' of the next section. This yields two complementary sticky ends that can anneal and be a substrate for a DNA ligase. Purified PCR products were mixed in 1:1 or 1:3 ratio to achieve a final volume of 10 μ L and were added to the Gibson

assembly master mix (10 μ L). Reactions were incubated at 50 $^{\circ}$ C for 1 hour, and then used to transform XL1-Blue cells (2 μ L).

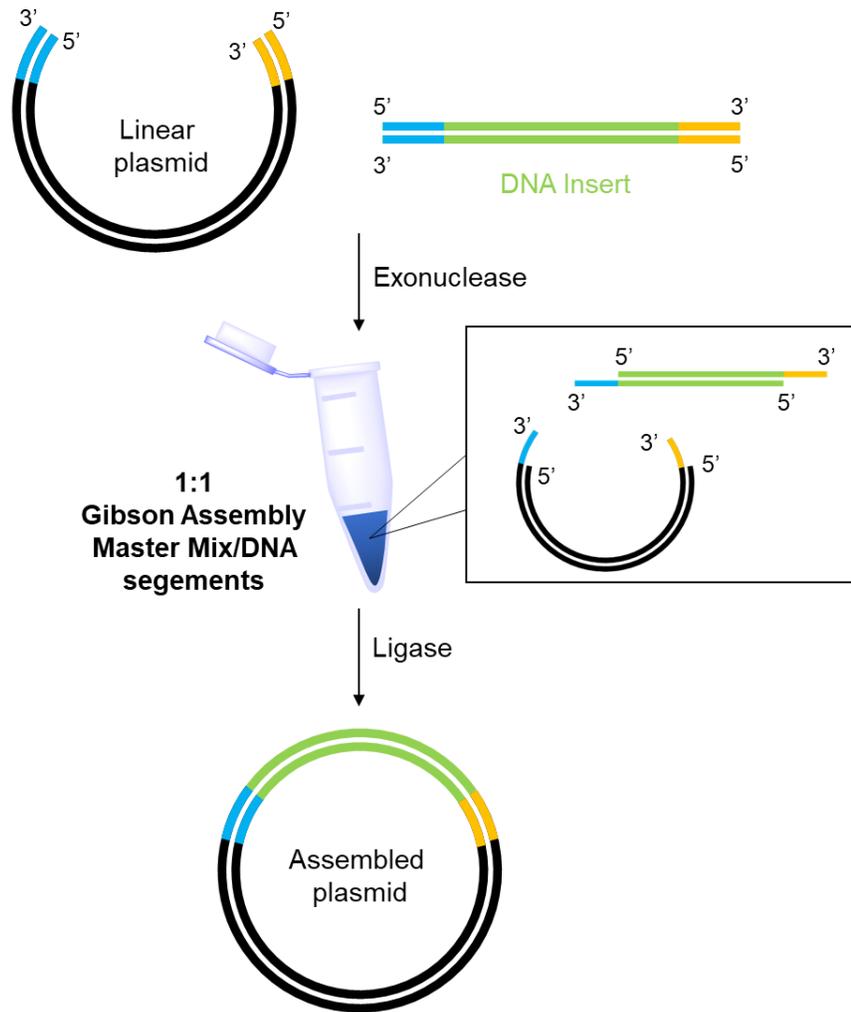


Figure 6-1 Gibson assembly: 5' exonuclease cuts in the 5' to 3' direction creating a 20 bp-long sticky end. 5' end of one segment is complementary to the 3' end of the other for annealing. A polymerase fills in the gaps of the annealed single strands region. A ligase joins segments covalently and fixes DNA nicks.

Table 6-12 Gibson assembly reaction mixture

Reaction components	V (μ l)
Total amount DNA fragments (0.02-0.05 pmols) ^a	10
Gibson Assembly Master mix ^b (x2)	10
diH ₂ O	Up to 20*

^apmols = (weight in ng) x 1,000 / (base pairs x 650 daltons) ^bMix contains buffer, optimised mix of exonuclease, DNA polymerase, Taq DNA ligase. *Final reaction volume of 20 μ L.

6.1.13. Golden gate

Golden Gate assembly (Figure 6-2) relies on a type IIS restriction enzyme *BsaI* able to cut DNA outside of the recognition site (upstream, shown in red). This leads to a 4 bp overhangs and allows T4 ligase to join these 4 base pairs with a complementary fragment between the 5' end of the first segment and the 3' end of the other. This method is interesting as the ligated product will not contain the original restriction site. This drives the ligation/digestion equilibrium towards the desired product. Reaction mixture (see Table 6-13) was incubated for 1 hour at 37 °C and then used to transform XL1-Blue cells (2 μ L).

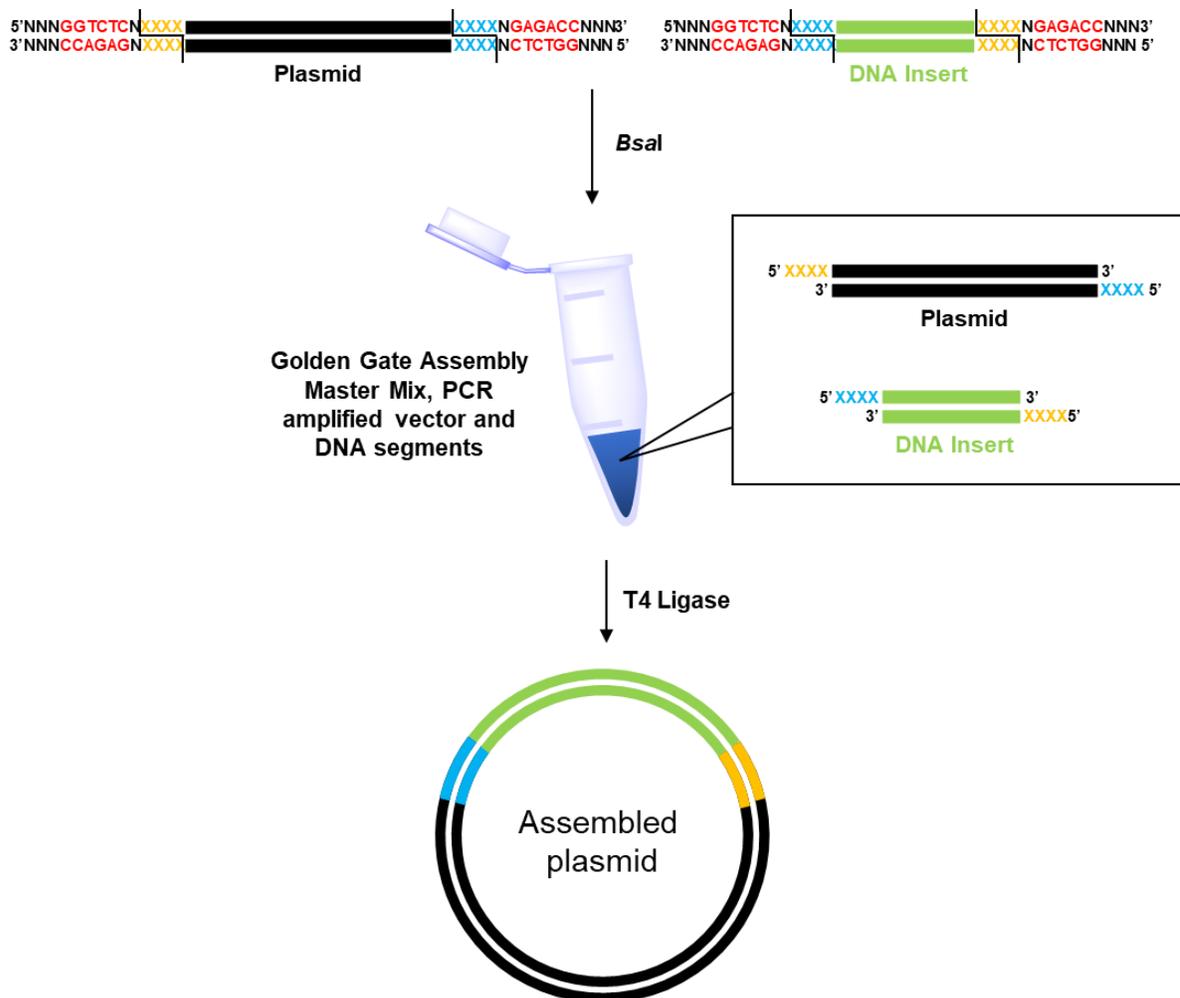


Figure 6-2 Golden Gate assembly: *BsaI* cuts DNA outside of its recognition site (shown in red) creating a 4 base pair long sticky end (shown in yellow and blue) then T4 Ligase connects complimentary DNA fragments and fixes DNA nick.

Table 6-13 Golden Gate reaction mixture

Reaction components	V (μ l)
Destination plasmid (75 ng/ μ L)	1
DNA segment 1 (75 ng/ μ L)	1
Golden Gate Buffer ^a (x10)	2
Golden Gate enzyme mix ^b	1
diH ₂ O	Up to 20*

^a Buffer including Mg²⁺ and Bovine serum albumin. ^bMix contains optimised mix of BsaI and T4 DNA ligase. *Final reaction volume of 10 μ L.

Table 6-14 PCR primer for cloning UK in pET32-Xa Lic

Name	5' Sequence 3'
PCR UK FWD	CCAGATCTGGGTACCATGCCGATGGATCTGCGTG
PCR UK REV	GGCCGCAAGCTTTTAGTGAAACAGCAGGAACC
PCR pET32-Xa Lic FWD	GCTGTTTCACTAAAAGCTTGCGGCCGCACTCG
PCR pET32-Xa Lic REV	GATCCATCGGCATGGTACCCAGATCTGGGCTGTCC

6.1.14. Transformation of competent cells

The appropriate frozen competent cells (50 μ L) were allowed to thaw on ice. Plasmids encoding the gene of interest were added to the cells (1 μ L for plasmids of 100 ng/ μ L DNA concentration) and mixture mixed gently. The mixture was then incubated on ice for at least 30 minutes and heat shocked at 42 °C (water bath) for 45 seconds before returning to ice. LB media (1 mL) was added to the mixture under aseptic conditions and incubated at 37 °C for 1 hour whilst shaking (250 rpm). Cells were harvested by centrifugation at 3,300 g for 60 seconds (Eppendorf centrifuge 5415R) and the supernatant was discarded. The cells were resuspended in fresh LB media (100 μ L) and suspension was spread on agar plates containing the necessary antibiotic(s). The plate was then incubated overnight at 37 °C.

6.1.15. Plasmid preparation (MiniPrep of DNA)

A single colony from a plate, or a spike from a glycerol stock, harbouring the required transformed cells was used to inoculate 10 mL of LB medium (containing the appropriate selective antibiotic(s)) and incubated overnight at 37 °C whilst shaking (250 rpm). Cells were harvested by centrifugation (3,400 g, 10 minutes) and the *Easy Pure*® Plasmid MiniPrep kit (TransGen Biotech, Beijing, China) was used to purify the plasmid following the manufacturer's protocol.

6.1.16. Protein expression

A single colony from a plate was picked with a sterile pipette tip and used to inoculate LB media (50 mL) containing the appropriate antibiotic. The culture was grown overnight in an incubator at 37 °C with shaking (250 rpm). The resulting culture was then used to inoculate fresh LB-media (5 ml of culture per 500 mL media). Culture was grown until reaching an OD₆₀₀ of 0.6. At this stage, gene transcription was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG, see Table 6-15 for concentration) and were incubated at the required temperature for the required time (see Table 6-15). Cells were harvested by centrifugation (3,400 g, 15 minutes) and pellets were stored at -20 °C.

Test expressions were used to optimise the protein production of UK, factors such as expression cells, and temperature of expression were investigated. They were typically carried out on a 1 mL scale in a 1.5 mL Eppendorf. Cells were induced when OD₆₀₀ reached 0.6 with the required compound (IPTG (0.2 mM) or L-arabinose (0.2 %)) and incubated at the required temperature during the required time (37 °C – 4 hours, 25°C – 20 hours, 20°C – 20 hours, 16°C – 20 hours). The cells were harvested by centrifugation, (3300 g, 1 minute, Eppendorf centrifuge 5415R) the pellets resuspended in SDS-sample buffer and analysed by SDS-PAGE.

Table 6-15 Expression conditions for proteins used in this work

Protein	Temperature (°C)	[IPTG] (mM)	Time (hours)
UK	16	0.2	20
IPK and variants	16	0.2	20
GdoIS	37	0.2	3
ADS	20	0.5	6
AS	37	0.5	4
GDS and variants	37	0.5	4

6.1.17. Protein purification

Cell lysis

Frozen induced cells harvested after expression were thawed on ice before being re-suspended in the appropriate buffer (see Table 6-16) containing phenylmethylsulphonyl fluoride (PMSF, 1 mM) to inhibit proteases and prevent degradation of the protein of interest. Cells were lysed by sonication (5 second sonication period (40% amplitude) followed by 10 second rest period for a total of 15 minutes, cells were kept in an ice-water bath). In some cases, lysozyme (1 mg.mL⁻¹) was added prior sonication, the suspension was stirred at 4°C for 30 min before the sonication step. Cell debris was then removed from the suspension by centrifugation at 4 °C, (38,000 g, 40 minutes) and the resulting supernatant and pellet analysed by SDS-PAGE.

Basic extraction

For proteins that remained as insoluble inclusion bodies (see Table 6-16), a basic extraction procedure was carried out to solubilise the protein after the cell lysis step. The cell pellet obtained after the centrifugation step was resuspended in the appropriate buffer (40 mL from 750 mL expression media, see Table 6-16) and the suspension was stirred at 4 °C for 30 min. The pH of the solution was subsequently adjusted to 12 using sodium hydroxide (6 M), at pH = 12 the milky solution becomes clearer. After 30 min stirring at 4 °C, 2-mercaptoethanol

is added to the solution to a concentration of 5 mM and the pH is slowly decreased to 8 by dropwise addition of 6 M hydrochloric acid. The solution is finally stirred for a further 30 min at 4 °C before centrifugation at 4 °C, (38,000 g, 40 minutes) and the resulting supernatant and pellet analysed by SDS-PAGE.

Table 6-16 List of cell lysis buffer used in this work.

Protein	Lysozyme	Lysis Buffer	Basic extraction
UK	1 mg.mL ⁻¹	50 mM Tris, 500 mM NaCl, 1% Tween 20, 10% Glycerol, pH 8.0	Yes
IPK and variants	1 mg.mL ⁻¹	50 mM Tris, 300 mM NaCl, 10% Glycerol, 0.5 mM TCEP ^a , pH 8.0	-
GdoIS	-	50 mM Tris, 100 mM NaCl and 10 mM imidazole, pH 8.0	Yes
ADS	1 mg.mL ⁻¹	50 mM Tris, 500 mM NaCl, 2 mM 2-mercaptoethanol, 10% Glycerol, pH 8.0	Yes
AS	-	20 mM Tris, 5 mM EDTA, 5 mM 2-mercaptoethanol, pH 8.0	Yes
GDS and variants	-	50 mM Tris, 50 mM NaCl, 0.2% Tween 20, 5 mM 2-mercaptoethanol, pH 8.0	Yes

^a tris(2-carboxyethyl)phosphine

6.1.18. Ni-NTA affinity chromatography

Harvested supernatant after centrifugation (IPK and variants) or basic extraction (UK, GdoIS, ADS, and GDS) was applied to an Ni-NTA column (5 mL Ni-NTA Affinity Resin - Amintra® , Expedeon, Cambridge, United Kingdom in an Econo-Column® Chromatography Column, Bio-Rad, 2.5 I.D. × 20 cm) pre-equilibrated with the buffer used for cell lysis, and flow-through collected. The column was then washed with 5 column volumes (CV) of lysis buffer and a gradient of imidazole (20, 60, 100, 250, 500 mM) applied. Solutions of imidazole consisted of lysis buffer containing 20, 60, 100, 250, and 500 mM concentrations of imidazole. The column was washed with 2 CV of each solution for a stepwise purification. Fractions were analysed by SDS-PAGE electrophoresis and those containing the protein of interest (>95 % pure) were combined for dialysis (ADS, GdoIS, GDS and variants) or desalting (UK, IPK and variants).

6.1.19. Purification of AS using anion exchange

Anion exchange chromatography was carried out with an AKTA FPLC system. Harvested supernatant after basic extraction was loaded to a pre-equilibrated Q-Sepharose™ Fast Flow column (Amersham Pharmacia Biotech™, Amersham, United Kingdom, 2.5 cm I.D. x 20 cm) with 20 mM Tris, 5 mM EDTA, 5 mM 2-mercaptoethanol, pH 8.0, and flow-through collected. The column was then washed with 5 column volumes (CV) of lysis buffer to remove any unbound protein. Protein was eluted using a linear gradient of 100 mM to 600 mM NaCl. Fractions absorbing at 280 nm were analysed by SDS-PAGE electrophoresis and those containing the protein of interest (>95 % pure) were combined for dialysis.

6.1.20. Dialysis

Fractions from purification found to contain the desired protein (>95 % pure) were dialysed to remove the residual imidazole. Fractions were combined in dialysis tubing (MWCO 12 - 14 kDa) and stirred overnight (16 h) in the appropriate buffer (see Table 6-17). After dialysis, the protein solution was concentrated by ultrafiltration (AMICON system, YM 30).

Table 6-17 List of dialysis buffer used in this work.

Protein	Dialysis Buffer	Volume
GdIS	10 mM Tris, pH 8.0	5 L
ADS	25 mM HEPES, 100 mM NaCl, 1 mM DTT, pH 7.5	5 L
AS	20 mM Tris, 2 mM 2-mercaptoethanol, pH 8.0	3 L
GDS and variants	10 mM Tris, 5 mM 2-mercaptoethanol, pH 7.5	3 L

6.1.21. Desalting

Fractions from purification found to contain the desired protein (>95 % pure) were desalted to remove the residual imidazole. Fractions were combined and loaded on a pre-equilibrated HiPrep™ 26/10 Desalting (GE Healthcare, Amersham United Kingdom, column volume = 60 mL, 2.6 mm I.D. 100 mm, 17 – 132 µm particle size) with desalting buffer (20 mM Tris, 100 mM NaCl, 10 % glycerol, pH 8.0) at 4 mL.min⁻¹. The protein was then eluted using 100%

desalting buffer at 4 mL.min⁻¹ and then the column was further washed with the desalting buffer (2 C.V.) to remove remaining salt. The presence of the protein was confirmed by SDS-PAGE electrophoresis.

6.1.22. SDS-PAGE electrophoresis

Protein samples were analysed by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were prepared, cast and run using the MiniPROTEAN system (Bio-Rad, California, USA)

Resolving and stacking gel mixture

The recipes for buffers used to cast the gels are found in Table 6-18. Each component was prepared in advance and stored at room temperature apart from 10% (w/v) APS which was freshly prepared. Resolving and stacking gel mixture were prepared by mixing all ingredient together in a Sterilin tube with APS and TEMED added only prior pouring the gel between the gel plates to initiate polymerisation.

Table 6-18 Gel casting buffers for SDS-PAGE (amount to make 2 gels).

Components	Resolving (12%)	Stacking (5%)
diH ₂ O	3.4 mL	2.85 mL
Resolving buffer: 1.5 M Tris-HCl (pH 8.0)	2.5 mL	-
Stacking buffer: 0.5 M Tris-HCl (pH 6.8)	-	1.25 mL
Acrylamide/bis-acrylamide 30% (8% w/v)	4 mL	0.85 mL
SDS ^a 10% (w/v)	100 µL	50 µL
APS ^b 10% (w/v)	100 µL	50 µL
TEMED ^c	20 µL	10 µL

^a SDS = sodium dodecylsulfate. ^b APS = ammonium persulfate. ^c TEMED = N,N,N',N'-tetramethylethylenediamine.

SDS sample buffer

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

Electrode running buffer (X10)

Tris-base (0.25 M, 30.3 g) and Glycine (2 M, 150.14 g) were dissolved in 600 mL deionised water, once dissolved SDS (10% w/v, 10 g) was added and left to stir slowly. After SDS dissolution, the final volume was brought up to 1 L with deionised water. Electrode running buffer X10 was diluted 10-fold prior to use.

SDS gel stain

Coomassie brilliant blue G-250 (50 mg) was dissolved in ethanol (50 mL), once dissolved, phosphoric acid (56 mL, 85 wt.% in H₂O) and 850 mL of water was added, and solution was stirred for 20 minutes. Dextrin-15 (2.5 g), α -cyclodextrin (2.5 g), β -cyclodextrin (2.5 g), γ -cyclodextrin (2.5 g) were then added and solution was for a further 60 minutes. The solution was stored at 4°C.

SDS Page protocol

Gel plates were held together in a casting gate to set the gels. Resolving gel mixture (approximately 3 mL) was poured between the gel plates, leaving around 2 cm space at the top, and covered with isopropanol (around 0.5 mL). The gel was left to polymerise for 20 minutes, isopropanol was removed and stacking gel mixture was poured on top of the resolving gel. A comb (for 10 or for 15 samples) was inserted in the stacking solution between the plates to create loading wells. After polymerisation, the comb was removed, and gels fitted into the running chamber filled with electrode running buffer. Protein sample (10 μ L) were incubated with SDS sample buffer (10 μ L) for 5 min at 85 °C and loaded into the gel's wells (for pellet samples, samples were centrifuged in a benchtop centrifuge (7,000 g, 5 minutes) prior loading). Gel was allowed to run at 150 V for 1 hour. Once run gel were stained for 5 minutes in the SDS gel stain to reveal the protein bands then imaged in ChemiDoc™ XRS+ gel imager (Bio-Rad). This stain does not require destaining.

6.1.23. Western blotting

Protein samples were analysed by western blot to identify protein of interest from the crude lysate. Western blotting was performed using the Trans-Blot® Turbo™ Transfer Starter System, Mini PVDF and Trans-Blot Turbo RTA Mini 0.2 µm PVDF Transfer Kit (Bio-Rad, California, USA)

Transfer buffer

Transfer buffer (1x) was prepared by 1:4 dilution of the commercial Bio-Rad 4x transfer buffer in deionised water.

Tris-buffered saline with Tween 20 (TBST) buffer

TBST buffer was prepared by dissolving Tris-base (25 mM, 3.0 g) and sodium chloride (150 mM, 8.8 g) in deionised water (800 mL) and adjusting the pH to 7.5 with diluted hydrochloric acid. Tween 20 (0.1% v/v) and deionised water were added to a final volume of 1 L. The buffer was stored at 4 °C.

Blocking solution

Bovine serum albumin (3 g) was dissolved in TBST buffer (100 mL). Solution was prepared just prior use.

Western blotting protocol

Proteins were first separated by SDS-PAGE electrophoresis and gel was immediately transferred to a pre-activated polyvinylidene difluoride (PVDF) membrane. The PVDF membrane was activated by submersion in methanol for 12 seconds. Two nitrocellulose membranes (stacks) were soaked in transfer buffer for three minutes, then a transfer sandwich was created by first placing one stack in the Trans-Blot Turbo Transfer System (Bio-Rad) then the PVDF membrane followed by the SDS-PAGE gel and finally the second stack. Excess transfer buffer and air bubbles were removed using a blotting roller before closing the transfer cell. The electrophoretic transfer was performed at 25 V, 1.3 A for 7 min. Once proteins were transferred onto the PVDF membrane, the membrane was left shaking in blocking buffer (20

mL) for 1 hour at room temperature. The membrane was then washed in TBST buffer (2 x 15 mL, 10 min) and incubated with His Probe HRP (horseradish peroxidase) Conjugate (0.8 mg/mL) overnight. The next morning, the PVDF membrane was washed in TBST buffer (4 x 15 mL, 10 min) and incubated with Clarity western enhanced chemiluminescence (ECL) substrate (Bio-Rad, 10 mL, 5 min) to reveal the proteins. The resulting membrane was analysed at 425 nm using a ChemiDoc™ XRS+ gel imager (Bio-Rad).

6.1.24. Protein concentration – Bradford assay

A modified version of the Bradford assay was used to determine the concentration of proteins.

Bradford reagent

Coomassie Brilliant Blue G-250 (20 mg) was dissolved in ethanol (2 mL), and phosphoric acid (20 mL, 80% wt. %) was added. Solution was brought to a final volume of 200 mL with deionised water. The solution was stored in the dark at 4 °C until use. Reagent was filtered through a 0.2 µm Millex syringe filter immediately prior to use.

Bradford assay

A sample of known concentration of bovine serum albumin (BSA, 1 mg.mL⁻¹) in deionised water was used to prepare a series of dilutions in deionised water from 2 to 100 µg.mL⁻¹ in a final volume of 200 µL. Each dilution was performed in triplicate. Bradford reagent (800 µL) was then added to each dilution and the absorbance from 400 to 600 nm was recorded using a Shimadzu spectrophotometer. The ratio of absorbances at 590 nm and 450 nm was calculated and the values for BSA were used to plot a standard curve. The same protocol was repeated to measure the concentration of an unknown protein sample. Three samples of different dilutions were prepared and recorded. Protein concentration was determined by comparing the absorbance ratios of unknown samples to the standard curve.

6.1.25. GC-MS analysis of products

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

General incubation buffer

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

GC-MS method

Gas chromatography coupled with mass spectrum (GC-MS) was performed on a Perkin Elmer Clarus 680 GC fitted with a Perkin Elmer Elite-1 column (30 m x 0.25 mm internal diameter) and a Perkin Elmer Clarus SQ 8 C mass spectrometer detecting in the range m/z 50-800 in EI+ mode with scanning once a second with a scan time of 0.9. The program uses an injection port temperature of 100 °C; split ratio 19:1; initial temperature 80 °C hold 2 min, ramp of 8 °C/min to 280 °C hold 3 min.

6.1.26. GC-FID analysis

Gas chromatography with flame ionisation detector (GC-FID) chiral analysis was performed on an Agilent 7890A GC system fitted with a Restek Rt-bDEXsm column (30 m x 0.32 mm internal diameter). GC FID method A: uses an injection port temperature of 200 °C, 5 μ L was injected with a 19:1 split. The oven temperature was held at 80 °C for 1 min and then rose at 8 °C/min to 200 °C and then held for 2 min.

6.1.27. UK and IPK phosphorylation incubation**Undecaprenol kinase incubation**

Farnesol was monophosphorylated in a solution (1 mL) containing farnesol (5 mM), UK (10 μ M), ATP (1 mM), PEP (10 mM), pyruvate kinase (20 U) in incubation buffer (20 mM Tris, 20 mM $MgCl_2$). The solution was left under shaking (250 rpm) at 37 °C for 2 hours before the solution was quenched with methanol (0.5 mL). The solution was centrifuged (3,400 g, 5 min) and supernatant was discarded. For the TLC analysis, pellet was redissolved using a solution of butanol and 1% phosphoric acid (1:1, 200 μ L).

Isopentenyl phosphate kinase incubation

Farnesyl phosphate (2 mM) was added to a solution (500 μ L) containing ATP (5 mM), $MgCl_2$ (20 mM), IPK variant (0.5 μ M), Gd4olS (10 μ M), β -CD (2 mM), and HP- β -CD (2 mM). The solution was left under shaking (350 rpm) at 37 °C overnight before alkaline phosphatase (10 mg.mL⁻¹) and $MgCl_2$ (100 mM) were added. The solution was left under shaking (250 rpm) at 37 °C for 2 hours before heptane (500 μ L) was added to extract the organic compounds. The organic layer was analysed by GC-MS.

One pot diphosphorylation of alcohols

Farnesol (7.2 μ L, 2 mM) was added to a solution ($V_{tot} = 25$ mL) containing UK (0.2 μ M), IPK 3 (2 μ M), $MgCl_2$ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), and the solution was left under shaking (250 rpm) at 37 °C for 20 hours before quenching with methanol (10 mL). The precipitate was isolated by centrifugation (3,400 g, 5 min), washed with water 3 times by repeated suspension and centrifugation and $Na_4P_2O_7$ (5 mL, 250 mM, pH 7) was added. The solution was loaded on to a SNAP Ultra C18 cartridge (Biotage®) to purify FDP by flash chromatography (a linear gradient from 0% to 2% A over 3 CV then 2% to 5% over 2 CV followed by 5% to 90% over 5CV and finally a linear gradient to 90% A over 2 CV.; solvent A: CH_3CN , solvent B: 25 mM NH_4HCO_3 , flow rate 25 mL.min⁻¹, detecting at 210 nm and 220 nm) to give the title compound **14** as a white solid in 95% yield. The same procedure was used to synthesise 6,15-dimethyl FDP (**145**) in 87% yield.

One pot synthesis of sesquiterpene from farnesol analogues

Farnesol analogue (2 mM) was added to a solution ($V_{tot} = 0.5$ mL) containing UK (0.2 μ M), IPK 3 (2 μ M), sesquiterpene cyclase (5-200 μ M), $MgCl_2$ (20 mM), ATP (0.1 mM), β -CD (5 mM), HP- β -CD (5 mM), PEP (10 mM), PK (50 U), and the solution was left under shaking (250 rpm) at 37 °C overnight before the solution was extracted with heptane (0.5 mL). Layers were separated and the organic layer was analysed by GC-MS.

6.1.28. Preparative scale enzymatic incubation

Preparative incubation of the 10,11-epoxy farnesyl diphosphate (**166**) was used for production of milligram quantities of product.

To a gently stirred solution of 10,11-epoxy farnesyl diphosphate (**166**, 30 mg, 67 μmol) in incubation buffer (50 mM Tris, 5 mM β -mercaptoethanol, 5 mM MgCl_2 , pH 8.0, 450 mL) was added GdoIS to a final concentration of 6 μM and the mixture was overlaid with pentane (50 mL). The flask was stoppered and stirred gently for 24 h at room temperature then a further portion of 10,11-epoxy farnesyl diphosphate (30 mg, 67 μmol) was added into the aqueous layer. After a further 24 h the solution was cooled to 4 $^\circ\text{C}$ then stirred for a further 64 h. The pentane layer was separated, and the aqueous layer was extracted with pentane (5 x 5 mL) by gentle swirling and slow separation. The combined pentane extracts were washed with brine, dried (Na_2SO_4), filtered and solvent was then removed carefully under reduced pressure (800 mbar minimum pressure at 30 $^\circ\text{C}$ water bath temperature to avoid loss of volatile product) to yield the terpenoid product as an oil. Purification by preparative thin layer chromatography on silica (5% ether in pentane) gave the title compound **146** as a colourless oil. (9 mg, 30%). δ_{H} (400 MHz, CDCl_3 , 323 K) ^1H NMR (400 MHz, CDCl_3 , 323 K) δ 5.32 (1 H, dd, $^3J_{\text{H,H}} = 9.0$ Hz and $^3J_{\text{H,H}} = 7.0$ Hz, $\text{C}=\text{CHCH}_2\text{O}$), 4.98 (1 H, t, $^3J_{\text{H,H}} = 7.0$ Hz, $\text{C}=\text{CHCH}_2\text{CH}_2$), 4.88 – 4.84 (1 H, s, $\text{C}=\text{CHH}$), 4.84 – 4.79 (1H, s, $\text{C}=\text{CHH}$), 3.89 (1 H, dd, $^2J_{\text{H,H}} = 12.0$ Hz, $^3J_{\text{H,H}} = 7.0$ Hz, CHCH_2O), 3.70 (1 H, dd, $^2J_{\text{H,H}} = 9.0$ Hz, $^3J_{\text{H,H}} = 7.0$ Hz, CHCH_2O), 3.51 (1 H, m, $\text{CH}_2\text{OCHC}=\text{C}$), 2.32 – 2.02 (6 H, m, 3 x CH_2), 1.70 (3 H, s, $\text{CH}_3\text{C}=\text{CH}_2$), 1.68 – 1.52 (2H, m, CH_2CHO), 1.64 (3H, s, $\text{CH}_3\text{CC}=\text{CHCH}_2\text{O}$), 1.54 (3H, s, $\text{CH}_3\text{C}=\text{CHCH}_2\text{CH}_2$). ^{13}C NMR (151 MHz, CDCl_3) δ (100 MHz, CDCl_3 , 323 K) 147.0 ($\text{C}=\text{HH}$), 139.3 ($\text{CH}_3\text{C}=\text{CHCH}_2\text{O}$), 134.6 ($\text{CH}_3\text{CCH}_2\text{CH}_2\text{CHO}$), 127.8 ($\text{CH}=\text{CCH}_2\text{CH}_2$), 124.7 ($\text{C}=\text{CHCH}_2\text{O}$), 110.7 ($\text{C}=\text{CH}_2$), 78.5 ($\text{CH}_2\text{OCHC}=\text{CH}_2$), 62.7 (CH_2OCH), 39.7 (CH_2CH_2), 39.3 (CH_2CH_2), 39.7 (CH_2CH_2), 25.1 (CH_2CH_2), 17.8 ($\text{CH}_3\text{C}=\text{CH}_2$), 16.1 ($\text{CH}_3\text{C}=\text{CHCH}_2\text{O}$), 15.1 ($\text{CH}_3\text{CCH}_2\text{CH}_2$). m/z (EI^+) 220.2 (6%, M^+), 202.2 (22), 187.2 (20), 159.1 (15), 147.1 (12), 145.1 (17), 134.1 (32), 125.1 (100), 121.1 (50), 119.1 (40), 93.1 (51), 91.1 (38), 81.1 (33), 79.1 (32), 68.1 (31), 67.1 (57), 55.1 (10). HRMS (EI^+ , M^+) found 220.1828, $\text{C}_{15}\text{H}_{24}\text{O}$ requires 220.1827.

6.2. SYNTHESIS

6.2.1. General experimental

All chemicals were purchased from Sigma-Aldrich, Alfa-Aesar or Fisher Scientific and used without further purification unless otherwise stated. Anhydrous tetrahydrofuran (THF), diethyl

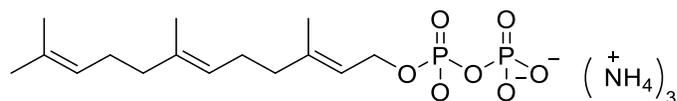
ether, toluene and acetonitrile were obtained from a MBraun SPS800 solvent purification system. Dichloromethane, and triethylamine were distilled from calcium hydride and KOH under nitrogen respectively. Room temperature (rt) refers to 20 - 25 °C. Temperatures of 0 °C and -10 °C were obtained using ice/water baths and ice/acetone baths respectively.

Thin layer chromatography (TLC) was performed on pre-coated aluminium plates of silica of Merck silica gel 60 F254 (0.20 mm). TLC visualizations were performed with 4.2% ammonium molybdate and 0.2% ceric sulfate in 5% H₂SO₄, or 0.1% berberine hydrochloride in EtOH or UV radiation (254 nm). Ion-exchange chromatography was performed using ion-exchange resin (Amberlyst 131 wet, H⁺ form and DOWEX 40 W, NH₄⁺ form) pre-equilibrated with ion-exchange buffer (25 mM NH₄HCO₃ containing 2% isopropanol). Manual column chromatography was performed using silica gel 60 (Merck, 230-400 mesh), while automated column chromatography and reverse phase purifications were performed on a Biotage® Isolera Four using Biotage® SNAP Ultra cartridges (Biotage, Uppsala, Sweden). The solvents used were laboratory grade. Lyophilisation process was carried out with a Christ Alpha 2-4 LD Plus Freeze Dryer fitted with a RZ6 vacuum pump (SciQuip, Newton, United Kingdom)

¹H, ¹³C NMR and ³¹P NMR spectra were measured on a Bruker DPX-600MHz spectrometer, Bruker Avance 500 NMR (¹H: 500 MHz, ¹³C: 125 MHz) spectrometer, Bruker Avance 400 (¹H: 400 MHz, ¹³C: 100 MHz) spectrometer or a Bruker Fourier 300 NMR spectrometer (¹H: 300 MHz, ¹³C: 76 MHz) at 25 °C unless otherwise stated and are reported as chemical shifts in parts per million downfield from tetramethylsilane and all coupling constants, J, (to the nearest 0.5 Hz) in Hz. Multiplicities are reported using the following notation: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High and low-resolution electrospray mass spectra of small molecules were determined by Waters GCT Premier time of flight mass spectrometer or a Waters LCT time of flight mass spectrometer.

6.2.2. Experimental

(2*E*,6*E*)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl tris-ammonium diphosphate **14**



Procedure 1

A stirring solution of N-chlorosuccinimide (1.12 g, 8.4 mmol) in anhydrous dichloromethane (50 mL) was cooled to -30 °C in an acetonitrile/dry ice bath and dimethyl sulfide (0.68 mL, 9.2 mmol) was added dropwise. The mixture was allowed to warm to 0 °C in an water/ice bath before being cooled down to -40 °C. A solution of farnesol (1.9 mL, 7.6 mmol) in dry dichloromethane (10 mL) was added dropwise to the milky mixture over 10 min. The reaction was allowed to warm to 0 °C slowly in an ice/water bath. After an additional hour of stirring, the now colourless solution was allowed to stir at rt for 15 min before quenching using cold brine (50 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 50 mL). The combined organic fractions were washed with sat. CuSO₄ (50 mL), sat. NaHCO₃ (50 mL) and brine (50 mL), dried (MgSO₄), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield the chloride, which was used without further purification.

The chloride intermediate was taken up in anhydrous acetonitrile (30 mL, 0.25 M) and tris(tetrabutylammonium)hydrogen diphosphate (13.5 g, 15.2 mmol) dissolved in anhydrous acetonitrile (15 mL) was added and stirred under N₂ for 16 h. The acetonitrile was removed under vacuum and the sticky brown residue dissolved in 5 mL of ion-exchange buffer (25 mM NH₄HCO₃, 2% iPrOH), and the resulting cloudy solution was passed slowly through a column of Dowex 50W-X8 cation-exchange (100 g) resin pre-equilibrated in ion-exchange buffer. Fractions were eluted with the same buffer and analysed by TLC (3:1:1, iPrOH:NH₄OH:buffer) and the fractions found to contain product were combined and lyophilised. The white powder was dissolved in 2 mL ion exchange buffer and purified by reverse-phase flash chromatography (SNAPCHAT Ultra C18 column (25 g), eluting a linear gradient from 0% to 2% A over 3 CV then 2% to 5% over 2 CV followed by 5% to 90% over 5CV and finally a linear gradient to 90% A over 2 CV.; solvent A: CH₃CN, solvent B: 25 mM NH₄HCO₃, flow rate 25 mL.min⁻¹, detecting at 210 nm and 220 nm) to give the title compound as a white solid in 13% yield (436 mg, 1.008 mmol)

¹H NMR (500 MHz, D₂O): δ 5.41 (1 H, t, *J* = 7.0, CHCH₂O), 5.18-5.10 (2 H, m, 2 x C=CHCH₂CH₂), 4.41 (2 H, t, *J* = 6.5, CHCH₂O), 2.11-1.94 (8 H, m, 2 x CHCH₂CH₂), 1.67, 1.63 and 1.57 (2 x 3 H and 1 x 6 H, 3 x s, 4 x CH₃).

³¹P NMR (121 MHz, D₂O): δ -6.49 (d, *J* = 22), -10.25 (d, *J* = 22).

HRMS (ES⁻, [M - H]⁻) found 381.1235, C₁₅H₂₇O₇P₂ requires 381.1232

Procedure 2

A stirring solution of farnesol (1.9 mL, 7.6 mmol) in dry THF (50 mL) was cooled to -10 °C in an acetone/ice bath and phosphorus tribromide (0.356 mL, 3.8 mmol) was added dropwise.

The mixture was stirred at $-10\text{ }^{\circ}\text{C}$ for 15 min and quenched using sat. NH_4Cl (50 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 50 mL). The combined organic fractions were washed with NaHCO_3 and brine, dried (MgSO_4), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield farnesyl bromide, which was used without further purification.

The bromide intermediate was taken up in anhydrous acetonitrile (30 mL, 0.25 M) and tris(tetrabutylammonium)hydrogen diphosphate (13.5 g, 15.2 mmol) dissolved in anhydrous acetonitrile (15 mL) was added and stirred under N_2 for 16 h. The acetonitrile was removed under vacuum and the sticky brown residue dissolved in 5 mL of ion-exchange buffer (25 mM NH_4HCO_3 , 2% *i*PrOH), and the resulting cloudy solution was passed slowly through a column of Dowex 50W-X8 cation-exchange (100 g) resin pre-equilibrated in ion-exchange buffer. Fractions were eluted with the same buffer and analysed by TLC (3:1:1, *i*PrOH: NH_4OH :buffer) and the fractions found to contain product were combined and lyophilised. The white powder was dissolved in 2 mL ion exchange buffer and purified by reverse-phase flash chromatography (SNAPCHAT Ultra C18 column (25 g), eluting a linear gradient from 0% to 2% A over 3 CV then 2% to 5% over 2 CV followed by 5% to 90% over 5CV and finally a linear gradient to 90% A over 2 CV.; solvent A: CH_3CN , solvent B: 25 mM NH_4HCO_3 , flow rate $25\text{ mL}\cdot\text{min}^{-1}$, detecting at 210 nm and 220 nm) to give the title compound as a white solid in 24% yield (784 mg, 1.811 mmol)

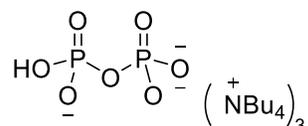
Procedure 3

A stirring solution of farnesol (0.5 mL, 2 mmol) in dry THF (20 mL) was cooled to $-10\text{ }^{\circ}\text{C}$ in an acetone/ice bath and phosphorus tribromide (0.094 mL, 1 mmol) was added dropwise. The mixture was stirred at $-10\text{ }^{\circ}\text{C}$ for 15 min and quenched using sat. NH_4Cl (20 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 20 mL). The combined organic fractions were washed with NaHCO_3 and brine, dried (MgSO_4), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield farnesyl bromide, which was used without further purification.

Dry *N,N*-diisopropylethylamine (2.3 mL, 13.2 mmol) was added dropwise to the trimethylsilyl protected pyrophosphate intermediate **120** under inert atmosphere at $0\text{ }^{\circ}\text{C}$. Water in acetone (0.115 mL, 6.4 mmol, 1.6 equiv, acetone/water 9:1) was then added dropwise. The suspension was warmed up to room temperature and the farnesyl bromide intermediate was added dropwise, after a brief stirring, the solution was left at r.t. until solid crystallised. The liquid phase was added dropwise to a stirred solution of ammonia (100 mL, 6 M in water) at $0\text{ }^{\circ}\text{C}$ (pH after addition was 11.47). The resulting cloudy solution was washed with ether (2 x

50 mL). Acetonitrile (250 mL) was added to the aqueous layer to precipitate FDP. The white solid (2.90 g) was filtered and purified by flash chromatography (25 mM NH_4HCO_3 , 2% iPrOH/MeCN, Snapchat C18 30 g, CV = 90 mL, $\text{H}_2\text{O}/\text{MeCN}$, 0% to 5% MeCN over 10 CV, 10% to 90% MeCN over 5 CV, 90% MeCN for 5 CV, UV collection 210 nm & 220 nm) to give a white fluffy solid in 43% yield (372 mg, 0.859 mmol).

Tris(tetra-n-butyl ammonium)hydrogen diphosphate **123**



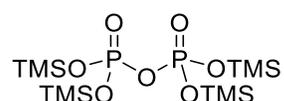
A solution of disodium dihydrogen diphosphate (3.3 g, 15 mmol) in ammonium hydroxide (25 mL, 10% v/v) was passed through a column of Amberlyst® 131 cation exchange resin (Hydrogen-form) and the free acid eluted with deionised water (110 mL). The flowthrough was collected and immediately titrated with aqueous tetra-n-butyl ammonium hydroxide (40% w/v) to pH 7.3. The resulting clear solution (~150 mL) was lyophilised to yield a hygroscopic white solid in quantitative yield (13.5 g, 15 mmol) and used without further purification.

^1H NMR (300 MHz, CDCl_3): δ 3.50 – 3.30 (24 H, m, 12 x $\text{NCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$), 1.75 – 1.55 (24 H, m, 12 x $\text{NCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$), 1.53 – 1.36 (24 H, m, 12 x $\text{NCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$), 0.92 (3 H, t, J 7.5, 12 x $\text{NCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$).

^{13}C (75 MHz, CDCl_3) δ 58.88, 24.30, 19.84, 13.92.

^{31}P (121 MHz, CDCl_3) δ -5.76.

Tetrakis(trimethylsilyl) diphosphate **120**

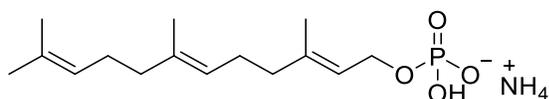


A stirring solution of finely powdered dihydrogen disodium diphosphate (1.1 g, 5 mmol), in dry formamide (5 mL) was cooled down to 0 °C and trimethylsilyl chloride (2.78 mL, 22 mmol) was added. After 1 h of stirring, petroleum ether (10 mL) was added and solution was stirred for an additional 5 min. The clear top layer was transferred into a dry 25 mL flask and the solvent as well as excess TMSCl were evaporated off under reduced pressure (ca. 2 mbar) at 40 °C to give a colorless liquid. The trimethylsilyl protected pyrophosphate intermediate was used without further purification.

^1H NMR (500 MHz, None) δ 0.55 (s, 36H, 12 x CH_3).

^{31}P NMR (202 MHz, None) δ -29.72 (s).

(2*E*,6*E*)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl ammonium phosphate **130**



A solution of bis-ethylammonium phosphate (TEAP) in acetonitrile was prepared by adding a solution of phosphoric acid in acetonitrile (18.2 mL, 21% v/v) to a solution of triethylamine in acetonitrile (30 mL, 52% v/v).

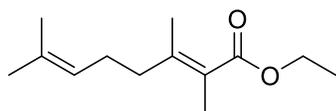
To a stirred solution of farnesol (1 mL, 4 mmol) in trichloroacetonitrile (10 mL, ca. 100 mmol) was added the TEAP solution in acetonitrile at 37°C for 15 min (3 successive additions of 16 mL at five-minute intervals). The reaction mixture was purified on silica column pre-equilibrated with the running buffer (*i*PrOH:NH₄OH:H₂O, 6:2.5:0.5) using an isocratic gradient of the running buffer. Fractions were analysed by TLC (*i*PrOH:NH₄OH:H₂O, 6:3:1) and fractions containing the product pooled together and solvent removed under reduced pressure. The final product was dissolved in ammonium hydrogen carbonate (25 mM) and lyophilised to yield a white solid in 22 % yield (280 mg, 0.9 mmol).

^1H NMR (400 MHz, D₂O) δ 5.36 (t, J = 7.0 Hz, 1H, CHCH_2O), 5.17 – 4.99 (m, 2H, 2 x $\text{C}=\text{CHCH}_2\text{CH}_2$), 4.36 (t, J = 6.2 Hz, 2H, CHCH_2O), 2.14 – 1.87 (m, 8H, 2 x CHCH_2CH_2), 1.66, 1.61, 1.56, 1.53 (4 x s, 4 x 3H, 4 x CH_3). 319.3818.

^{31}P NMR (162 MHz, D₂O) δ 0.42 (s).

HRMS (ES⁻): calculated for (C₁₅H₂₇O₄P · [H])⁻: 301.1569 found: 301.1573.

Ethyl (*E*)-2,3,7-trimethylocta-2,6-dienoate **150**



To a solution of triethyl 2-phosphonopropionate (21.7 mL, 101.1 mmol) in dry THF (100 mL) was added *n*-BuLi (40.5 mL, 101.1 mmol, 1.5 eq.) was added dropwise at 0 °C. The reaction was stirred for 2 hours at rt, cooled to 0 °C and 6-methyl-5-hepten-2-one (10 mL, 67.4 mmol) was added. The reaction was stirred at rt overnight and quenched with water (150 mL). The aqueous layer was separated and extracted with ethyl acetate (3 x 50 mL). The combined

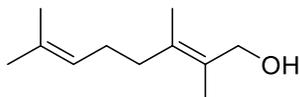
organic fractions were washed with water (100 mL), brine (100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (10% ethyl acetate in hexane) yielded a mixture of *E* and *Z* alkene in 73% yield as a clear oil (10.0 g, 47.6 mmol). Flash chromatography (5% ethyl acetate in hexane) partially separated *E* ester for analysis.

^1H NMR (300 MHz, CDCl_3): δ 5.19 – 5.07 (1H, m, $(\text{CH}_3)\text{CCH}$), 4.19 (2H, q, $J = 7.1$ Hz, OCH_2CH_3), 2.42 – 2.30 (2H, m, $\text{CH}_2(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 2.19 – 2.14 (2H, m, CH_2), 1.87 (3H, s, CH_3), 1.85 (3H, s, CH_3), 1.80 (3H, s, CH_3), 1.62 (3H, s, CH_3), 1.31 (3H, t, $J = 7.1$ Hz, OCH_2CH_3).

^{13}C NMR (75 MHz, CDCl_3): δ 169.9 ($(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 146.1 ($(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 145.6 ($(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 131.8 ($(\text{CH}_3)\text{CCH}$), 124.0 ($(\text{CH}_3)\text{CCH}$), 60.0 (OCH_2CH_3), 36.5 ($\text{CH}_2(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 29.7 ($\text{CH}_2\text{CH}_2(\text{CH}_3)\text{CC}(\text{CH}_3)\text{C}(\text{O})$), 20.9 (CH_3), 17.6 (CH_3), 15.9 (CH_3), 15.3 (CH_3), 14.3 (OCH_2CH_3).

HRMS (ES^+): calculated for $[\text{C}_{13}\text{H}_{23}\text{O}_2]^+$; 211.1698, found 211.1698.

(E)-2,3,7-Trimethylocta-2,6-dien-1-ol **151**



DIBAL-H (71.4 mL, 71.4 mmol, 1 M in THF) was added to a mixture of *E* and *Z* **150** (5 g, 23.8 mmol) in anhydrous toluene (100 mL) at -78 °C. Reaction was quenched with HCl (100 mL, 1 M) at 0 °C when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 2 hours). The aqueous layer was separated and extracted with dichloromethane (3 x 50 mL). The combined organic fractions were washed with sat. aq. NaHCO_3 (3 x 100 mL), brine (3 x 100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5% ethyl acetate in hexane) yielded the title compound in 36% yield as a clear oil (1.45 g, 8.6 mmol).

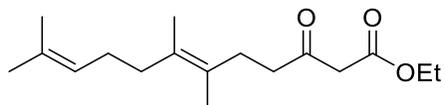
^1H NMR (300 MHz, CDCl_3) δ 5.12 (m, 1H, $\text{C}=\text{CH}$), 4.12 (s, 2H, CH_2OH), 2.05 (m, 4H, CH_2CH_2), 1.70 – 1.80 (m, 6H, 2 x CH_3), 1.68, 1.61 (2 x s, 2 x 3H, 2 x CH_3).

^{13}C NMR (75 MHz, CDCl_3) δ 133.15, 131.80 (2 x CH_3CC), 127.86 ($\text{C}=\text{CH}$), 124.13 ($\text{C}=\text{CH}$), 64.15 (CH_2OH), 34.95, 26.40 (2 x CH_2), 25.76, 17.98, 17.64, 16.26 (4 x CH_3).

HRMS (EI^+): calculated for $[\text{C}_{11}\text{H}_{20}\text{O}_2] - \text{H}_2\text{O}]^+$; 150.1409, found 150.1403.

NOESY ^1H - ^1H interactions of Z isomer: signal at 5.04 to 2.02 and 1.5, 4.00 to 2.02 and 1.68 to 1.52 ppm.

Ethyl (*E*)-6,7,11-trimethyl-3-oxododeca-6,10-dienoate **18**



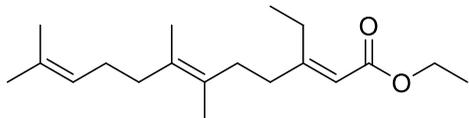
A stirring solution of (*E*)-2,3,7-trimethylocta-2,6-dien-1-ol **151** (1.52 g, 9.0 mmol) in dry THF (50 mL) was cooled to $-10\text{ }^\circ\text{C}$ in an acetone/ice bath and phosphorus tribromide (0.423 mL, 4.5 mmol) was added dropwise. The mixture was stirred at $-10\text{ }^\circ\text{C}$ for 15 min and quenched using sat. NH_4Cl (50 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 50 mL). The combined organic fractions were washed with NaHCO_3 and brine, dried (MgSO_4), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield the corresponding bromide, which was used without further purification.

Ethylacetoacetate (3.41 mL, 27 mmol) was added drop wise over 10 min into a stirring suspension of NaH (1.26 g in 60% oil, 31.5 mmol) in dry THF (100 mL) at $0\text{ }^\circ\text{C}$. After 30 min of stirring at rt, the solution was cooled down to $0\text{ }^\circ\text{C}$ and *n*-BuLi (11.1 mL, 2.5 M in hexanes, 27.9 mmol) was added drop wise over 10 min. The yellow solution was stirred 30 min at rt and then cooled down to $0\text{ }^\circ\text{C}$. The bromide intermediate in dry THF (5 mL) was added dropwise and the mixture was stirred 1 h at $0\text{ }^\circ\text{C}$ and over the weekend at rt. The reaction was quenched with HCl (50 mL, 1 M), diluted with water (50 mL). The aqueous layer was separated and extracted with ether (3 x 50 mL). The combined organic fractions were washed with water (2 x 100 mL), brine (2 x 100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 to 20% ethyl acetate in hexane) yielded the title compound in 63% yield as a clear oil (1.60 g, 5.7 mmol).

^1H NMR (300 MHz, CDCl_3) δ 5.26 – 5.05 (m, 1H, $\text{C}=\text{CH}$), 4.21 (q, $J = 7.1\text{ Hz}$, 2H, OCH_2CH_3), 3.46 (s, $J = 2.3\text{ Hz}$, 2H, COCH_2CO), 2.62 – 2.02 (m, 2 x 4H, 4 x CH_2), 1.69 (s, 3H, CH_3), 1.64 (m, 9H, 3 x CH_3), 1.30 (t, $J = 7.1\text{ Hz}$, 3H, OCH_2CH_3).

^{13}C NMR (75 MHz, CDCl_3) δ 203.12 (CO), 167.29 ($\text{OCOCH}_2\text{CH}_3$), 131.56 ($\text{C}=\text{C}$), 129.78 ($\text{C}=\text{C}$), 126.24 ($\text{C}=\text{CH}$), 124.32 ($\text{C}=\text{CH}$), 61.41 (OCH_2CH_3), 49.43 (OCCH_2CO), 41.55, 34.78, 28.61, 26.65 (4 x CH_2), 25.77, 18.12, 17.87, 17.63, 14.14 (5 x CH_3).

HRMS (AP^+): calculated for $(\text{C}_{17}\text{H}_{28}\text{O}_3 + [\text{H}])^+$: 281.2072, found: 281.2070.

Ethyl (2*E*,6*E*)-3-ethyl-6,7,11-trimethyldodeca-2,6,10-trienoate **154**

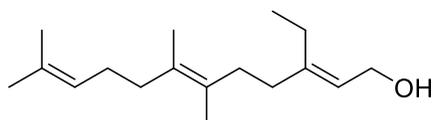
A stirring solution of ethyl (*E*)-6,7,11-trimethyl-3-oxododeca-6,10-dienoate **18** (500 mg, 1.8 mmol) in dry dichloromethane (25 mL) under inert atmosphere was cooled to 0 °C and lithium triflate (842 mg, 5.4 mmol) was added. After 10 min of stirring at 0 °C, dry triethylamine (0.752 mL, 5.4 mmol) was added dropwise and then stirred at 0°C for a further 5 min. Triflic anhydride (0.387 mL, 2.3 mmol) was then added and stirred for 1.5 h at 0 °C before it was quenched with sat. NH₄Cl (25 mL). The aqueous layer was separated and extracted with dichloromethane (3 x 10 mL). The combined organic fractions were washed with brine, dried over magnesium sulphate, filtered and solvent was removed under reduced pressure to give a yellow oil (709 mg). This triflate intermediate was then stored at -20 °C and used the following day without further purification.

A stirring suspension of copper cyanide (403 mg, 4.5 mmol) in dry THF (25 mL) was cooled down to -78 °C in an acetone/dry ice bath, and ethyl magnesium bromide (0.9 mL, 3 M in ether, 2.7 mmol) was added dropwise at -78 °C. The resulting suspension was stirred at -78 °C for a further 30 min before the dropwise addition of the triflate intermediate in dry THF (2 mL). After 2 h at -78 °C, the reaction mixture was quenched with sat. NH₄Cl (20 mL). The aqueous layer was separated and extracted with ether (3 x 10 mL), and the combined organic fractions were washed with brine, dried over magnesium sulphate, filtered and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 to 20% ethyl acetate in hexane) yielded the title compound in 44% yield over 2 steps as a clear oil (243 mg, 0.8 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.61 (s, 1H, C=CHCO), 5.11 (s, 1H, C=CH), 4.14 (q, *J* = 7.1 Hz, 2H, OCH₂CH₃), 2.63 (q, *J* = 7.5 Hz, 2H, CH=CCH₂CH₃), 2.15 - 2.01 (m, 8H, CH₂), 1.64 (t, *J* = 11.5 Hz, 14H), 1.27 (t, *J* = 7.1 Hz, 3H), 1.09 (t, *J* = 7.5 Hz, 3H).

HRMS (AP⁺): calculated for (C₁₉H₃₂O₂ + [H])⁺: 293.2481 found: 293.2485.

(2*E*,6*E*)-3-Ethyl-6,7,11-trimethyldodeca-2,6,10-trien-1-ol **141**



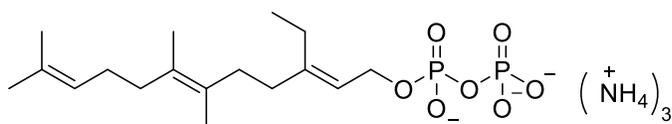
DIBAL-H (2.4 mL, 1 M in THF, 2.4 mmol) was added to ethyl (2*E*,6*E*)-3-ethyl-6,7,11-trimethyldodeca-2,6,10-trienoate **154** (243 mg, 0.8 mmol) in anhydrous toluene (15 mL) at -78 °C. Reaction was quenched with HCl (20 mL, 1 M) at 0 °C when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 1 hour). The aqueous layer was separated and extracted with dichloromethane (3 x 10 mL). The combined organic fractions were washed with sat. aq. NaHCO₃ (3 x 10 mL), brine (3 x 20 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5% ethyl acetate in hexane) yielded the title compound in 76% yield as a clear oil (156 mg, 0.6 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.38 (t, *J* = 7.0 Hz, 1H, C=CHCH₂OH), 5.30 (s, 1H, OH), 5.12 (s, 1H, C=CH), 4.16 (d, *J* = 7.0 Hz, 2H, CH₂OH), 2.22 – 1.92 (m, 8H, 4 x allylic CH₂), 1.68 (s, 3H, CH₃), 1.64 (s, 3H, CH₃), 1.60 (s, 3H, CH₃), 1.57 (s, 6H, 2 x CH₂), 1.00 (t, *J* = 7.6 Hz, 3H, CH₃CH₂).

¹³C NMR (75 MHz, CDCl₃) δ 142.85 (C=CH), 135.62 (C=CH), 131.40 (C=C), 124.31 (C=C), 123.45 (C=CH), 120.27 (C=CH), 66.10 (CH₂OH), 39.69, 39.45, 26.69, 26.10 (4 x allylic CH₂), 25.74 (CH₃), 25.67 (CH₂CH₃), 17.72, 16.14, 16.06 (3 x CH₂).

HRMS (ES⁺): calculated for (C₁₇H₃₀O - [H₂O])⁺: 232.2191 found: 232.2199.

(2*E*,6*E*)-3-Ethyl-6,7,11-trimethyldodeca-2,6,10-trien-1-yl tris-ammonium diphosphate **145**

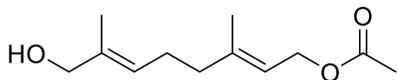


This compound was prepared and purified from alcohol **141** (78 mg, 0.3 mmol) in the same manner as FDP **14**, using protocol 3 to give the title compound as a white solid in 33% yield (49 mg, 0.11 mmol).

¹H NMR (500 MHz, D₂O) δ 5.48 (t, *J* = 7.0 Hz, 1H, C=CHCH₂O), 5.24 (m, 1H, C=CH), 4.50 (m, 2H, CH₂O), 2.29 – 2.05 (m, 8H, allylic CH₂), 1.69 (m, 9H, CH₃), 1.63 (s, 3H, CH₃), 1.02 (t, *J* = 7.5 Hz, 3H, CH₂CH₃).

³¹P NMR (202 MHz, D₂O) δ -6.29 (d, *J* = 22.3 Hz), -10.41 (d, *J* = 22.4 Hz).

HRMS (ES⁻, [M - H]⁻) found 409.1547, C₁₅H₂₇O₇P₂ requires 409.1545.

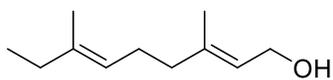
(2E,6E)-8-hydroxy-3,7-dimethylocta-2,6-dien-1-yl acetate 156

To a stirring suspension of selenium dioxide (0.5 g, 4.7 mmol) and salicylic acid (0.6 g, 4.7 mmol) in dichloromethane (100 mL) was added *t*BuOOH (19 mL, 70% in water, 140 mmol) and the mixture was stirred at rt for 10 min. The mixture was cooled to 0 °C, and geranyl acetate (10 mL, 46.7 mmol) in dichloromethane (20 mL) was added. The reaction mixture was stirred for a further 40 h at rt. The suspension was diluted with ether (100 mL) and successively washed with sat. aq. sodium hydrogen carbonate (200 mL), sat. aq. copper sulphate (200 mL), sat. aq. sodium sulphite (200 mL), water (200 mL), brine (200 mL), dried over magnesium sulfate and solvent was removed under reduced pressure. The crude mixture was dissolved in EtOH (100 mL) and NaBH₄ (3.5 g, 93.4) was added. The mixture was stirred for 2 h at rt before it was quenched with dropwise addition of HCl (1M, 50 mL). The aqueous layer was separated and extracted with dichloromethane (3 x 30 mL). The combined organic fractions were washed with water (100mL), brine (100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 – 30% ethyl acetate in hexane) yielded the title compound in 80% yield as a yellow oil (6.3 g, 29.9 mmol).

¹H NMR (400 MHz, CDCl₃) δ 5.48 – 5.25 (m, 2H, 2 x C=CH), 4.58 (d, *J* = 7.0 Hz, 1H, CH₂OCOCH₃), 3.99 (s, 1H, CH₂OH), 2.23 – 2.07 (m, 4H, CH₂CH₂), 2.06 (s, 3H, CH₃CO), 1.70 (s, 1H, CH₃), 1.66 (s, 1H, CH₃).

¹³C NMR (75 MHz, CDCl₃) δ 171.28 (CO), 141.78 (C=CH), 135.26 (C=CH), 125.31 (C=CH), 118.63 (C=CH), 68.91 (CH₂OH), 61.45 (CH₂OCO), 39.06, 25.64 (CH₂CH₂), 21.10 (CH₃CO), 16.43, 13.72 (2 x CH₃).

LRMS (ES⁺): calculated for (C₁₂H₂₀O₃[Na])⁺: 235.1310, found: 235.1299.

(2E,6E)-3,7-Dimethylnona-2,6-dien-1-ol 157

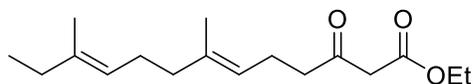
A stirring solution of (2*E*,6*E*)-8-hydroxy-3,7-dimethylocta-2,6-dien-1-yl acetate **156** (4.44 g, 20.9 mmol) in dry THF (100 mL) was cooled to -10 °C in an acetone/ice bath and phosphorus tribromide (0.987 mL, 10.5 mmol) was added dropwise. The mixture was stirred at -10 °C for 15 min and quenched using sat. NH₄Cl (50 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 50 mL). The combined organic fractions were washed with NaHCO₃ and brine, dried (MgSO₄), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield the corresponding bromide, which was used without further purification.

The bromide intermediate was taken up in dry THF (200 mL) and the stirred solution was cooled to 0 °C. Methyl magnesium bromide (35 mL, 3 M in ether, 105 mmol) was added dropwise through a syringe driver (25 mL/h) and the solution was left stirred for 2 h at 0 °C before it was quenched quenching with sat. NH₄Cl (100 mL). The aqueous layer was separated and extracted with ether (3 x 80 mL). The combined organic fractions were washed with brine (100 mL), dried over magnesium sulphate, filtered and solvent evaporated under reduced pressure. Purification by flash chromatography on silica (5 – 20% ethyl acetate in hexane) yielded the title compound in 29% yield over 2 steps as a clear oil (1.0 g, 6.1 mmol). ¹H NMR (300 MHz, CDCl₃) δ 5.42 (m, 1H, C=CHCH₂OH), 5.09 (m, 1H, C=CHCH₂CH₂), 4.16 (m, 2H, CH₂OH), 2.02 (m, 6H, allylic 3 x CH₂), 1.68 (s, 3H, CH₃), 1.59 (s, 3H, CH₃), 0.98 (t, *J* = 7.4 Hz, 3H, CH₃CH₂).

¹³C NMR (75 MHz, CDCl₃) δ 139.95 (C=CH), 137.33 (C=CH), 123.27 (C=CH), 122.29 (C=CH), 59.46 (CH₂OH), 39.60, 32.35, 26.24 (3 x CH₂), 22.11, 16.33, 12.83 (3 x CH₃).

LRMS (AP⁺): calculated for (2 x C₁₇H₂₈O₃ + 2[H])⁺: 338.3185, found: 338.3501.

Ethyl (6*E*,10*E*)-7,11-dimethyl-3-oxotrideca-6,10-dienoate **159**



A stirring solution of alcohol **157** (1.0 g, 6.1 mmol) in dry THF (50 mL) was cooled to -10 °C in an acetone/ice bath and phosphorus tribromide (0.291 mL, 3.1 mmol) was added dropwise. The mixture was stirred at -10 °C for 15 min and quenched using sat. NH₄Cl (50 mL) and layers were separated. The aqueous layer was extracted with hexane (3 x 50 mL). The combined organic fractions were washed with NaHCO₃ and brine, dried (MgSO₄), the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield the corresponding bromide, which was used without further purification.

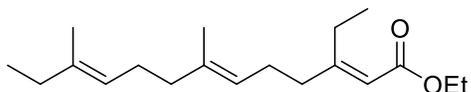
Ethylacetoacetate (2.31 mL, 18.3 mmol) was added drop wise over 10 min into a stirring suspension of NaH (0.856 g in 60% oil, 21.4 mmol) in dry THF (100 mL) at 0 °C. After 30 min of stirring at rt, the solution was cooled down to 0°C and *n*-BuLi (7.6 mL, 2.5 M in hexanes, 18.9 mmol) was added drop wise over 10 min. The yellow solution was stirred 30 min at rt and then cooled down to 0 °C. The bromide intermediate in dry THF (5 mL) was added dropwise and the mixture was stirred 1 h at 0 °C and over the weekend at rt. The reaction was quenched with HCl (50 mL, 1 M), diluted with water (50 mL). The aqueous layer was separated and extracted with ether (3 x 50 mL). The combined organic fractions were washed with water (2 x 100 mL), brine (2 x 100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 to 20% ethyl acetate in hexane) yielded the title compound in 52% yield over 2 steps as a clear oil (0.907 g, 3.2 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.10 – 5.08 (m, 2H, C=CH), 4.21 (q, *J* = 7.1 Hz, 2H, OCH₂CH₃), 3.45 (s, 2H, COCH₂CO), 2.58 (t, *J* = 7.4 Hz, 2H, CH₂CH₂CO), 2.45 – 2.18 (m, 2H, allylic CH₂), 2.18 – 1.90 (m, 6H, 3 x allylic CH₂), 1.62 (s, 3H, CH₃), 1.60 (s, 3H, CH₃), 1.29 (t, *J* = 7.1 Hz, 3H, OCH₂CH₃), 0.98 (t, *J* = 7.5 Hz, 3H, CH₃CH₂C).

¹³C NMR (75 MHz, CDCl₃) δ 202.76 (COCH₂), 167.28 (OCOCH₂), 137.04 (C=CH), 136.82 (C=CH), 122.54 (C=CH), 122.04 (C=CH), 61.40 (OCH₂CH₃), 49.43 (OCCH₂CO), 43.10 (CH₂COCH₂), 39.69, 32.35, 26.45, 22.14 (4 x allylic CH₂), 16.04, 15.95, 14.14, 12.84 (4 x CH₃).

LRMS (ES⁺): calculated for (C₁₇H₂₈O₃[Na])⁺: 303.1936, found: 303.1931.

Ethyl (2*E*,6*E*,10*E*)-3-ethyl-7,11-dimethyltrideca-2,6,10-trienoate **161**



A stirring solution of ethyl (6*E*,10*E*)-7,11-dimethyl-3-oxotrideca-6,10-dienoate **159** (500 mg, 1.8 mmol) in dry dichloromethane (25 mL) under inert atmosphere was cooled to 0 °C and lithium triflate (842 mg, 5.4 mmol) was added. After 10 min of stirring at 0 °C, dry triethylamine (0.752 mL, 5.4 mmol) was added dropwise and then stirred at 0°C for a further 5 min. Triflic anhydride (0.387 mL, 2.3 mmol) was then added and stirred for 1.5 h at 0 °C before it was quenched with sat. NH₄Cl (25 mL). The aqueous layer was separated and extracted with dichloromethane (3 x 10 mL). The combined organic fractions were washed with brine, dried over magnesium sulphate, filtered and solvent was removed under reduced pressure to give

a yellow oil (741 mg). This triflate intermediate was then stored at -20 °C and used the following day without further purification.

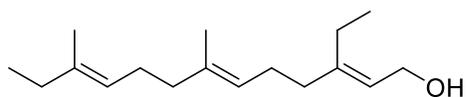
A stirring suspension of copper cyanide (403 mg, 4.5 mmol) in dry THF (25 mL) was cooled down to -78 °C in an acetone/dry ice bath, and ethyl magnesium bromide (0.9 mL, 3 M in ether, 2.7 mmol) was added dropwise at -78 °C. The resulting suspension was stirred at -78 °C for a further 30 min before the dropwise addition of the triflate intermediate in dry THF (2 mL). After 2 h at -78 °C, the reaction mixture was quenched with sat. NH₄Cl (20 mL). The aqueous layer was separated and extracted with ether (3 x 10 mL), and the combined organic fractions were washed with brine, dried over magnesium sulphate, filtered and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 to 20% ethyl acetate in hexane) yielded the title compound in 61% yield over 2 steps as a clear oil (335 mg, 1.1 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.61 (s, 1H, C=CHCO), 5.18 – 4.96 (m, 2H, C=CH), 4.14 (q, *J* = 7.1 Hz, 2H, OCH₂CH₃), 2.62 (q, *J* = 7.5 Hz, 2H, CH₃CH₂C=C), 2.17 (s, 4H, CH₃), 2.02 – 1.92 (m, 12H; allylic CH₂), 1.59 – 1.61 (m, 6H, CH₃), 1.27 (t, *J* = 7.1 Hz, 3H, OCH₂CH₃), 1.07 (t, *J* = 7.5 Hz, 3H, CH₃CH₂C), 0.97 (t, *J* = 7.5 Hz, 3H, CH₃CH₂C).

¹³C NMR (75 MHz, CDCl₃) δ 166.57 (OCOCH₂CH₃), 165.71 (C=CHCO), 136.99, 136.12 (2 x C=CH), 123.03, 122.62 (2 x C=CH), 114.80 (C=CHCO), 59.49 (OCH₂CH₃), 39.72, 37.99, 32.35, 26.51, 26.13, 25.37 (6 x allylic CH₂), 16.08, 15.94, 14.34, 13.04, 12.83 (5 x CH₃).

LRMS (AP⁺): calculated for (C₁₉H₃₂O₂ + [H])⁺: 293.2481 found: 293.2479.

(2*E*,6*E*,10*E*)-3-Ethyl-7,11-dimethyltrideca-2,6,10-trien-1-ol **158**



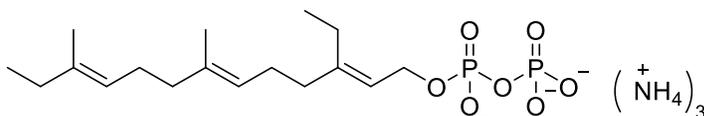
DIBAL-H (3.3 mL, 1 M in THF, 2.4 mmol) was added to **161** (335 mg, 1.1 mmol) in anhydrous toluene (15 mL) at -78 °C. Reaction was quenched with HCl (20 mL, 1M) at 0 °C when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 1 hour). The aqueous layer was separated and extracted with dichloromethane (3 x 10 mL). The combined organic fractions were washed with sat. aq. NaHCO₃ (3 x 10mL), brine (3 x 20 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5% ethyl acetate in hexane) yielded the title compound in 55% yield as a clear oil (141 mg, 0.6 mmol).

^1H NMR (300 MHz, CDCl_3) δ 5.38 (t, $J = 6.9$ Hz, 1H, $\text{C}=\text{CHCH}_2\text{OH}$), 5.20 – 4.99 (m, 2H, 2 x $\text{C}=\text{CH}$), 4.16 (d, $J = 7.0$ Hz, 2H, CH_2OH), 2.20 – 1.90 (m, 12H, 6 x allylic CH_2), 1.60 – 1.59 (m, 6H, $\text{CH}_3\text{C}=\text{CH}$), 0.99 (tt, $J = 7.5, 3.7$ Hz, 6H, 2 x CH_3CH_2).

^{13}C NMR (75 MHz, CDCl_3) δ 145.82 ($\text{C}=\text{CHCH}_2\text{OH}$), 136.94, 135.40 ($\text{C}=\text{CH}$), 123.89, 122.74, 122.71 (3 x $\text{C}=\text{CH}$), 61.72 (CH_2OH), 39.75, 36.42, 32.37, 26.51, 23.53, 16.08 (6 x allylic CH_2), 15.96, 14.71 (2 x $\text{CH}_3\text{C}=\text{CH}$), 13.77, 12.84 (2 x CH_3CH_2).

HRMS (ES^+): calculated for $(\text{C}_{17}\text{H}_{30}\text{O} - [\text{H}_2\text{O}])^+$: 232.2191 found: 232.2190.

(2*E*,6*E*,10*E*)-3-Ethyl-7,11-dimethyltrideca-2,6,10-trien-1-yl tris-ammonium diphosphate **148**



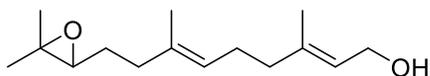
This compound was prepared and purified from alcohol **158** (141 mg, 0.6 mmol) in the same manner as FDP **14**, using protocol 3 to give the title compound as a white solid in 37% yield (55 mg, 0.12 mmol).

^1H NMR (400 MHz, D_2O) δ 5.27 (bs, 1H, $\text{C}=\text{CHCH}_2\text{OH}$), 5.07 (bs, 2H, 2 x $\text{C}=\text{CH}$), 4.31 (t, $J = 7.7$ Hz, 2H, CH_2O), 2.01 – 1.97 (m, 12H, 6 x allylic CH_2), 1.52 (s, 6H, $\text{CH}_3\text{C}=\text{CH}$), 0.3 – 0.85 (m, 6H, 6 x CH_3CH_2).

^{31}P NMR (162 MHz, D_2O) δ -6.72, -10.48.

HRMS (ES^-): calculated for $(\text{C}_{15}\text{H}_{27}\text{O}_7\text{P}_2 - [\text{H}])^-$, found 409.1540, requires 409.1545.

(2*E*,6*E*)-9-(3,3-dimethyloxiran-2-yl)-3,7-Dimethylnona-2,6-dien-1-ol **144**



A stirring solution of 2, 6-*E*, *E*-farnesyl acetate (1.1 g, 4.2 mmol) in THF (105 mL) was cooled to 0 °C and diluted with water (80 mL) to give a cloudy solution. *N*-Bromosuccinimide (0.82 g, 4.6 mmol) was added in 3 portions over the course of 1 h at 0 °C and the suspension was stirred for a further 1 h at 0 °C before solvent partially removed under reduced pressure in a 5 °C water bath. The solution was diluted with ether/hexane (1:1, 25 mL) and washed with water (2 x 25 mL), dried over sodium sulphate and solvent was removed under reduced pressure. Purification by flash chromatography on silica (20% ethyl acetate in hexane) yielded the bromohydrin intermediate as a pale-yellow oil and was taken up immediately in methanol (25 mL). Potassium carbonate (1.1 g, 8.0 mmol) was added to the solution and the reaction mixture was stirred for 4 h at rt. The mixture was concentrated under reduced pressure and

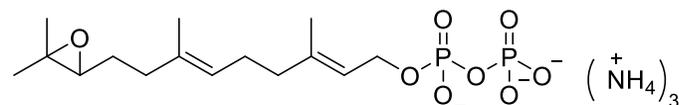
then diluted with EtOAc (50 mL) and water (50 mL). The aqueous layer was separated and extracted with EtOAc (3 x 20 mL). The combined organic fractions were washed with HCl (1 M, 15 mL), brine (50 mL), dried over sodium sulphate, and solvent removed under reduced pressure. Purification by flash column chromatography on silica (20% EtOAc in hexane) yielded the title compound as a clear oil in 80% yield over 2 steps (0.81 g, 3.4 mmol).

^1H NMR (500 MHz, CDCl_3) δ 5.41 (m, 1H, CCHCH₂OH), 5.25 – 5.06 (m, 1H, CCHCH₂CH₂), 4.14 (d, J = 6.9 Hz, 2H, CCHCH₂OH), 2.70 (t, J = 6.2 Hz, 1H, C(O)CH), 2.20 – 1.99 (m, 6H, 3 x CH₂), 1.69 – 1.66 (m, 3H, CH₃), 1.65 – 1.60 (m, 5H, CH₂, CH₃), 1.30, 1.26 (s, 3H, 2 x CH₃).

^{13}C NMR (126 MHz, CDCl_3) δ 139.41 (CH₃CCH), 134.37 (CH₃CCH), 124.52 (CCHCH₂CH₂), 123.62 (CCHCH₂OH), 64.17 (C(O)CH), 59.38 (CCHCH₂OH), 58.38 (C(O)CH), 39.40, 36.34, 27.28, 26.15 (4 x CH₂), 24.87, 18.79, 16.24, 15.98 (4 x CH₃).

HRMS (ES^+) calculated for ($\text{C}_{15}\text{H}_{26}\text{O}_2 + [\text{Na}]^+$), found 261.1838, requires 261.1831.

(2*E*,6*E*)-9-(3,3-dimethyloxiran-2-yl)-3,7-Dimethylnona-2,6-dien-1-yl tris-ammonium diphosphate **166**



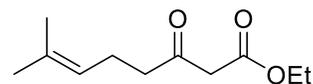
This compound was prepared and purified from alcohol **144** (200 mg, 0.8 mmol) in the same manner as FDP **14**, using protocol 3 to give the title compound as a white solid in 54% yield (194 mg, 0.4 mmol).

^1H NMR (500 MHz, D_2O) δ 5.15 – 5.07 (2 H, m, 2 x C=CH), 4.25 – 3.78 (2 H, m, C(O)CHCH₂O), 3.21 (1 H, dd, J 7.5, 3.5, C(O)CHCH₂O), 2.10 – 1.89 (6 H, m, 3 x CH₂), 1.59, 1.54, 1.53, 1.27 (4 x 3 H, 4 x s, 4 x CH₃).

^{13}C (126 MHz, D_2O) δ 136.90, 133.53 (2 x C=CH), 124.50 (C=CH), 123.71 (C=CH), 64.47 (C(O)CHCH₂O), 63.77 (C(O)CHCH₂O), 38.84, 37.67, 25.84 (3 x CH₂), 24.89 (CH₃), 23.02 (CH₂), 17.02, 15.72, 15.23 (3 x CH₃); ^{31}P (202 MHz, D_2O) -6.95 (m), -10.47 (m).

HRMS (ES^-) calculated for ($\text{C}_{15}\text{H}_{27}\text{O}_8\text{P}_2 - [\text{H}]^-$), found 397.1182, requires 397.1181.

Ethyl 7-methyl-3-oxooct-6-enoate **173**



Ethylacetoacetate (8.5 mL, 67.1 mmol) was added drop wise over 20 min into a stirring suspension of NaH (4.1 g in 60% oil, 100.7 mmol) in dry THF (250 mL) at 0 °C. After 30 min

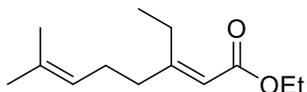
of stirring at rt, the solution was cooled down to 0 °C and *n*-BuLi (40.3 mL, 2.5 M in hexanes, 100.7 mmol) was added drop wise over 10 min. The yellow solution was stirred 30 min at rt and then cooled down to 0 °C. 3,3-dimethylallyl bromide in dry THF (10 mL) was added dropwise and the mixture was stirred 1 h at 0 °C and over the weekend at rt. The reaction was quenched with HCl (100 mL, 1 M), diluted with water (50 mL). The aqueous layer was separated and extracted with ether (3 x 100 mL). The combined organic fractions were washed with water (2 x 100 mL), brine (2 x 100 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5% ethyl acetate in hexane) yielded the title compound in 89% yield as a clear oil (5.95 g, 30.0 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.06 (1H, s, C=CHCH₂), 4.16-4.22 (2H, q, OCH₂CH₃, J = 7.2 Hz), 3.42 (2H, s, COCH₂CO), 2.53-2.58 (2H, t, CH₂CH₂CO, 7.5 Hz), 2.26-2.30 (2H, m, CH₂CH₂CO), 1.67 (3H, s, CH₃C=CH), 1.61 (3H, s, CH₃C=CH), 1.25-1.30 (3H, t, OCH₂CH₃, 7.2 Hz)

¹³C NMR (126 MHz, CDCl₃) δ 202.6 (COCH₂CO), 167.2 (COCH₂CH₃), 133.1 (C=CHCH₂), 122.5 (C=CHCH₂), 61.3 (OCH₂CH₃), 49.4 (COCH₂CO), 43.0 (CH₂CH₂CO), 25.6 (CH₃C=CH), 22.2 (CH₂CH₂CO), 17.6 (CH₃C=CH), 14.1 (OCH₂CH₃)

HRMS (EI⁺) calculated for (C₁₁H₁₈O₃)⁺, found: 198.1256, requires 198.1260.

Ethyl (*E*)-3-ethyl-7-methylocta-2,6-dienoate **174**



A stirring solution of **173** (3.0 g, 15.3 mmol) in dry dichloromethane (300 mL) under inert atmosphere was cooled to 0 °C and lithium triflate (8.3 g, 53.0 mmol) was added. After 10 min of stirring at 0 °C, dry triethylamine (10.2 mL, 53.0 mmol) was added dropwise and then stirred at 0 °C for a further 5 min. Triflic anhydride (3.8 mL, 22.7 mmol) was then added and stirred for 1.5 h at 0 °C before it was quenched with sat. NH₄Cl (100 mL). The aqueous layer was separated and extracted with dichloromethane (3 x 50 mL). The combined organic fractions were washed with brine, dried over magnesium sulphate, filtered and solvent was removed under reduced pressure to give a yellow oil. This triflate intermediate was then stored at -20 °C and used the following day without further purification.

A stirring suspension of copper cyanide (3.39 g, 37.8 mmol) in dry THF (190 mL) was cooled down to -78 °C in an acetone/dry ice bath, and ethyl magnesium bromide (7.6 mL, 3 M in ether, 23.0 mmol) was added dropwise at -78 °C. The resulting suspension was stirred

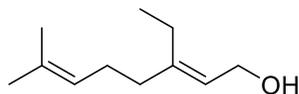
at -78 °C for a further 30 min before the dropwise addition of the triflate intermediate in dry THF (10 mL). After 2 h at -78 °C, the reaction mixture was quenched with sat. NH₄Cl (100 mL). The aqueous layer was separated and extracted with ether (3 x 50 mL), and the combined organic fractions were washed with brine (2 x 70 mL), dried over magnesium sulphate, filtered and solvent was removed under reduced pressure. Purification by flash chromatography on silica (5 % ethyl acetate in hexane) yielded the title compound in 61% yield over 2 steps as a clear oil (1.93 g, 9.2 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.60 (1H, s, C=CHCO), 5.08 (1H, m, C=CHCH₂), 4.10-4.17 (2H, q, OCH₂CH₃, 7.1 Hz), 2.57-2.64 (2H, q, CCH₂CH₃, 7.5 Hz), 2.14-2.15 (4H, m, C=CHCH₂CH₂ and C=CHCH₂CH₂), 1.68 (3H, s, CH₃C=CH), 1.61 (3H, s, CH₃C=CH), 1.24-1.29 (3H, t, OCH₂CH₃, 7.1 Hz), 1.03-1.09 (3H, t, CCH₂CH₃, 7.5 Hz)

¹³C NMR (126 MHz, CDCl₃) δ 166.5 (C=CHCO), 165.5 (C=CHCO), 132.4 (C=CHCH₂), 123.1 (C=CHCH₂), 114.8 (C=CHCO), 59.5 (OCH₂CH₃), 38.0 and 26.2 (CHCH₂CH₂ and CHCH₂CH₂), 25.7 (CH₃C=CH), 25.3 (CCH₂CH₃), 17.7 (CH₃C=CH), 14.3 (OCH₂CH₃), 13.0 (CCH₂CH₃)

HRMS (EI⁺): calculated for (C₁₃H₂₂O₂)⁺, found: 210.1621, requires 210.1620.

(*E*)-3-Ethyl-7-methylocta-2,6-dien-1-ol **175**



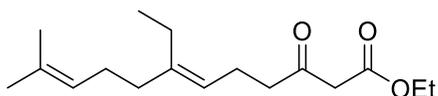
DIBAL-H (15 mL, 1 M in THF, 14.7 mmol) was added to **174** (1.55 g, 7.4 mmol) in anhydrous toluene (150 mL) at -78 °C. Reaction was quenched with HCl (30 mL, 1M) at 0 °C when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 1 hour). The aqueous layer was separated and extracted with dichloromethane (3 x 30 mL). The combined organic fractions were washed with sat. aq. NaHCO₃ (3 x 70 mL), brine (2 x 70 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (10% ethyl acetate in hexane) yielded the title compound in 88% yield as a clear oil (1.09 g, 6.5 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.35-5.40 (1H, t, C=CHCH₂OH, 7 Hz), 5.10-5.11 (1H, m, C=CHCH₂CH₂), 4.11-4.19 (2H, d, CHCH₂OH, 7 Hz), 2.05-2.13 (6H, m, CHCH₂CH₂, CHCH₂CH₂ and CCH₂CH₃), 1.69 (3H, s, CH₃C=CH), 1.60 (3H, s, CH₃C=CH), 0.96-1.02 (3H, t, CCH₂CH₃, 7.5 Hz)

^{13}C NMR (126 MHz, CDCl_3) δ 144.5 ($\text{C}=\text{CHCH}_2$), 131.1 ($\text{CH}_3\text{C}=\text{CH}$), 122.9 ($\text{CH}_3\text{C}=\text{CH}$), 121.8 ($\text{C}=\text{CHCH}_2$), 58.0 (OCH_2CH_3), 35.4 (CH_2CH_2), 25.6 ($\text{CH}_3\text{C}=\text{CH}$), 24.7 (CH_2CH_2), 16.7 ($\text{CH}_3\text{C}=\text{CH}$), 12.7 (OCH_2CH_3)

HRMS (EI^+): calculated for $(\text{C}_{11}\text{H}_{20}\text{O}_2)^+$, found: 168.1521, requires 168.1514.

Ethyl (*E*)-7-ethyl-11-methyl-3-oxododeca-6,10-dienoate **176**



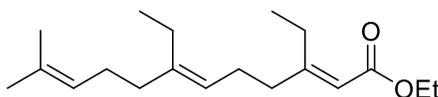
This compound was prepared and purified from alcohol **175** (1.24 g, 7.4 mmol) in the same manner as **173**, to give the title compound as a colourless oil in 57% yield (1.17 g, 4.2 mmol).

^1H NMR (300 MHz, CDCl_3) δ 4.98-5.09 (2x1H, m, $\text{C}=\text{CHCH}_2$ x2), 4.16-4.23 (2H, q, OCH_2CH_3 , 7.1 Hz), 3.43 (2H, s, COCH_2CO), 2.34-2.60 (2H, t, $\text{CH}_2\text{CH}_2\text{CO}$, 7.5 Hz), 2.24-2.31 (2H, m, $\text{CH}_2\text{CH}_2\text{CO}$), 1.96-2.07 (6H, m, $\text{CHCH}_2\text{CH}_2\text{C}$, $\text{CHCH}_2\text{CH}_2\text{C}$ and $\text{CH}_2\text{C}=\text{CH}$), 1.68 (3H, s, $\text{CH}_3\text{C}=\text{CH}$), 1.60 (3H, s, $\text{CH}_3\text{C}=\text{CH}$), 1.26-1.31 (3H, t, OCH_2CH_3 , 7.1 Hz), 0.92-0.98 (3H, t, CCH_2CH_3 , 7.5 Hz)

^{13}C NMR (126 MHz, CDCl_3) δ 202.7 (COCH_2COO), 167.3 (COCH_2COO), 142.7 ($\text{CH}_2\text{C}=\text{CH}$), 131.5 ($\text{CH}_3\text{C}=\text{CH}$), 124.3 ($\text{C}=\text{CH}$), 121.5 ($\text{C}=\text{CH}$), 61.4 (OCH_2CH_3), 49.4 (COCH_2CO), 43.4 ($\text{CH}_2\text{CH}_2\text{CO}$), 36.4 (CH_2), 26.8 (CH_2), 25.7 ($\text{CH}_3\text{C}=\text{CH}$), 23.1 (CH_2), 21.8 ($\text{CH}_2\text{CH}_2\text{CO}$), 17.7 ($\text{CH}_3\text{C}=\text{CH}$), 14.1 (OCH_2CH_3), 13.2 (CCH_2CH_3)

HRMS (EI^+): calculated for $(\text{C}_{17}\text{H}_{28}\text{O}_3)^+$, found: 279.1089, requires 280.2038.

Ethyl (*2E,6E*)-3,7-diethyl-11-methyldodeca-2,6,10-trienoate **46**



This compound was prepared and purified from **176** (4.2 mmol) in the same manner as **174**, to give the title compound as a colourless oil in 60% yield over two steps (0.74 g, 2.5 mmol).

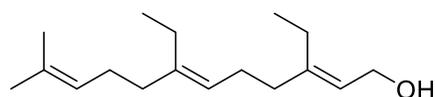
^1H NMR (300 MHz, CDCl_3) δ 5.60 (1H, s, $\text{C}=\text{CHCO}$), 5.04-5.08 (2x1H, m, $\text{CH}_3\text{C}=\text{CH}$ and $\text{C}=\text{CHCH}_2$), 4.12-4.14 (2H, q, OCH_2CH_3 , 7.1 Hz), 2.56-2.64 (2H, q, $\text{COCH}=\text{CCH}_2\text{CH}_3$, 7.5 Hz), 2.15-2.16 (2x2H, m, $\text{CH}_2\text{CH}_2\text{CCH}_2$ and $\text{CH}_2\text{CH}_2\text{CCH}_2$), 1.95-2.05 (3x2H, m, $\text{CH}_2\text{CH}_2\text{CCH}_2$, $\text{CH}_2\text{CH}_2\text{CCH}_2$ and $\text{CH}_2\text{CH}_2\text{CCH}_2$), 1.66 (3H, s, $\text{CH}_3\text{C}=\text{CH}$), 1.59 (3H, s, $\text{CH}_3\text{C}=\text{CH}$), 1.23-1.28

(3H, t, OCH₂CH₃, 7.1 Hz), 1.03-1.08 (3H, t, COCH=CCH₂CH₃, 7.5 Hz), 0.91-0.96 (3H, t, CH₃CH₂C=CHCH₂, 7.6 Hz)

¹³C NMR (126 MHz, CDCl₃) δ 166.5 (C=CHCO), 165.6 (C=CHCO), 142.0 (CH₂C=CHCH₂), 131.4 (CH₃C=CH), 124.4 (C=CHCH₂), 122.6 (C=CHCH₂), 114.8 (C=CHCO), 59.5 (OCH₂CH₃), 38.3 and 36.5 (CHCH₂CH₂C and CHCH₂CH₂C), 26.9, 25.8 and 25.7 (CHCH₂CH₂CCH₂, CHCH₂CH₂CCH₂ and CHCH₂CH₂CCH₂), 25.3 (CH₃CH₂C=CHCO), 23.2 (CH₂), 17.8 (CH₃C=CH), 14.3 (OCH₂CH₃), 13.2 (COCH=CCH₂CH₃), 13.0 (CH₃CH₂C=CHCH₂)

HRMS (EI⁺): calculated for (C₁₉H₃₂O₂)⁺, found: 292.2341, requires 292.2402.

(2*E*,6*E*)-3,7-diethyl-11-methyldodeca-2,6,10-trien-1-ol **140**



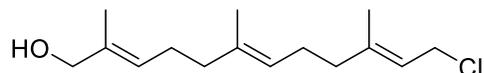
This compound was prepared and purified from **46** (0.42 g, 1.4 mmol) in the same manner as **175**, to give the title compound as a colourless oil in 88% yield over two steps (0.31 g, 1.4 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.36-5.41 (1H, t, CHCH₂OH, 7.0 Hz), 5.08-5.11 (2x1H, m, CH₃C=CH and CH₂C=CH), 4.15-4.17 (2H, d, CHCH₂OH, 7.0 Hz), 1.99-2.13 (6x2H, m, CH₂CH₂CCH₂ x2, CH₂CH₂CCH₂ x2 and CH₂CH₂CCH₂ x2), 1.68 (3H, s CH₃C=CH), 1.60 (3H, s CH₃C=CH), 0.93-1.01 (2x3H, m, CH₃CH₂C)

¹³C NMR (126 MHz, CDCl₃) δ 145.7 (CCH₂OH), 141.3 (CH₂C=CHCH₂), 131.3 (CH₃C=CH), 124.5 (CH₂C=CHCH₂), 123.5 (CH₂C=CHCH₂), 122.9 (CCHCH₂OH), 59.1 (OCH₂CH₃), 36.7 and 36.5 (CH₂CH₂C=CH x2), 26.9 and 26.2 (CH₂CH₂C=CH x2), 25.7 (CH₃C=CH), 23.5 and 23.2 (CH₃CH₂C=CH), 17.7 (CH₃C=CH), 13.7 (CH₃CH₂C), 13.2 (CH₃CH₂C)

HRMS (ES⁺): calculated for (C₁₇H₃₀O - [H₂O])⁺: 232.2191 found: 232.2189.

(2*E*,6*E*,10*E*)-12-Chloro-2,6,10-trimethyldodeca-2,6,10-trien-1-ol **179**



A stirring solution of farnesol (2 mL, 8 mmol) in anhydrous DMF (80mL) was cooled to 0 °C and collidine (6.3 mL, 48 mmol) was added followed by MsCl (1.3 mL, 16 mmol). The solution was stirred at 0 °C for 15 min and anhydrous LiCl (1.35 g, 32 mmol) was added. The reaction mixture was stirred for a further 3 h at 0 °C before it was diluted with water (80 mL). The reaction mixture was extracted with hexane (3 x 50 mL) and the combined organic fractions were washed with sat. copper sulphate (100 mL), sat. sodium bicarbonate (100 mL), brine

(100 mL), dried over sodium sulphate, and the solvent removed under reduced pressure, and dried under high vacuum for 1 h to yield the chloride, which was used without further purification.

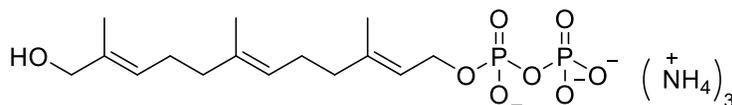
To a stirring suspension of selenium dioxide (226 mg, 2.4 mmol) and salicylic acid (331 mg, 2.4 mmol) in dichloromethane (40 mL) was added tBuOOH (5.5 mL, 70% in water, 40 mmol) and the mixture was stirred at rt for 10 min. The mixture was cooled to 0 °C, and the chloride intermediate (8 mmol) in dichloromethane (20 mL) was added. The reaction mixture was stirred overnight at 0 °C. The suspension was diluted with ether (100 mL) and successively washed with sat. aq. sodium hydrogen carbonate (200 mL), sat. aq. copper sulphate (200 mL), sat. aq. sodium sulphite (200 mL), water (200 mL), brine (200 mL), dried over sodium sulphate and solvent was removed under reduced pressure. Purification by flash chromatography on silica (dichloromethane with 1% NEt₃) yielded the title compound in 50% yield over 2 steps as a clear oil (1.04 g, 4.0 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.48-5.36 (2 H, m, 2 x C=CH), 5.10 (1 H, t, *J* = 6.0, C=CH), 4.10 (2 H, d, *J* = 6.0, CH₂Cl), 3.99 (2 H, s, CH₂OH), 2.17-1.98 (8 H, m, CH₂CH₂), 1.73, 1.66 and 1.60 (3 x 3 H, 3 x s, 3 x CH₃).

¹³C NMR (75 MHz, CDCl₃) δ 142.6, 136.2, 134.35 (3 x CH₃CCH), 126.05, 124.02, 120.32, (3 x CH₃CCH), 69.25 (CH₂OH), 42.51 (CH₂Cl), 39.22, 39.06 (2 x CCH₂CH₂), 26.22, 26.01 (2 x CCH₂CH₂), 16.24, 16.15, 13.70 (3 x CH₃).

HRMS (ES⁺): calculated for (C₁₅H₂₅O³⁵Cl + [Na])⁺: 279.1483, found: 279.1482.

(2*E*,6*E*,10*E*)-12-Chloro-2,6,10-trimethyldodeca-2,6,10-trien-1-yl tris-ammonium diphosphate
111

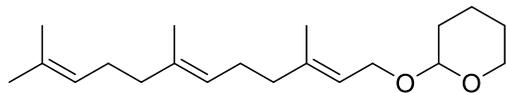


This compound was prepared and purified from **179** (1.0 g, 4 mmol) in the same manner as FDP **14**, using protocol 1 to give the title compound as a white solid in 46% yield (832 mg, 1.8 mmol).

¹H NMR (500 MHz, D₂O) δ 5.46 – 5.25 (m, 2H, 2 x C=CH), 5.15 (t, *J* = 6.3 Hz, 1H, C=CH), 4.41 (d, *J* = 5.6 Hz, 2H, CHCH₂O), 3.88 (s, 2H, CH₂OH), 2.15 – 1.91 (m, 8H, 4 x allylic CH₂), 1.64 (s, 3H, CH₃), 1.58 – 1.53 (m, 3H, 2x CH₃).

³¹P NMR (162 MHz, D₂O) δ -10.69 – -10.86 (m), -10.88 – -11.04 (m).

HRMS (ES⁻): calculated for (C₁₅H₂₈O₈P₂- 2[H] + [Na])⁻: 419.1001, found: 419.0996.

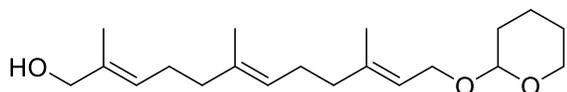
2-(((2*E*,6*E*)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran **181**

To a stirring solution of farnesol (5 g, 22.7 mmol) and 3,4-dihydropyran (8.1 mL, 88.8 mmol) in dichloromethane (50 mL) at rt was added pyridinium *p*-toluenesulfonate (PPTS, 150 mg, 0.6 mmol) and left stirring at room temperature for 3 hours before water (50 mL) was added. The aqueous layer was separated and extracted with dichloromethane (3 x 30 mL). The combined organic fractions were washed with brine (100 mL), dried over magnesium sulphate and solvent removed under reduced pressure. Purification by flash column chromatography on silica (10% EtOAc in hexane) gave the title compound as a clear oil in quantitative yield (7.0 g, 22.7 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.36 (1 H, t, J = 6.5, CHCH₂O), 5.10-5.13 (2 H, m, 2 x CCHCH₂CH₂), 4.63 (1 H, t, J = 4.0, OCHO), 4.24 (1 H, dd, J = 12.0 and 6.5, CCHCH_AO), 4.06 (1 H, dd, J = 12.0 and 7.5, CCHCH_BO), 3.97-3.39 (2 H, m, OCH₂CH₂), 2.33-1.91 (8 H, m, 2 x CCHCH₂CH₂), 1.62-1.47 (6 H, m, OCHCH₂CH₂CH₂), 1.68 and 1.60 (2 x 6 H, 2 x s, 4 x CH₃).

¹³C NMR (75 MHz, CDCl₃) δ 140.34, 135.26 and 131.34 (3 x CH₃CCH), 124.34, 123.90 and 120.55 (3 x CH₃CCH), 97.80 (OCHO), 63.67 and 62.31 (2 x CHOCH₂), 39.72 and 39.65 (2 x CH₃CCH₂), 30.73 (OCHCH₂), 26.74 and 26.30 (2 x CCHCH₂CH₂), 25.73 (OCH₂CH₂), 25.52 (OCH₂CH₂CH₂), 19.64, 17.71, 16.45 and 16.03 (4 x CH₃).

HRMS (ES⁺): calculated for (C₂₀H₃₄O₂ + [Na])⁺: 329.2457, found: 329.2463.

(2*E*,6*E*,10*E*)-2,6,10-Trimethyl-12-((tetrahydro-2H-pyran-2-yl)oxy)dodeca-2,6,10-trien-1-ol **182**

To a stirring suspension of selenium dioxide (777 mg, 7 mmol) and salicylic acid (966 mg, 7 mmol) in dichloromethane (100 mL) was added *t*BuOOH (11.2 mL, 70% in water, 114 mmol) and the mixture was stirred at rt for 10 min. The mixture was cooled to 0 °C, and 2-(((2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran **181** (7.0 g, 22.7 mmol) in dichloromethane (20 mL) was added. The reaction mixture was stirred overnight at 0 °C. The suspension was diluted with ether (100 mL) and successively washed with sat. aq. sodium hydrogen carbonate (200 mL), sat. aq. copper sulphate (200 mL), sat. aq. sodium sulphite

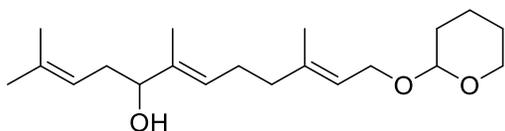
(200 mL), water (200 mL), brine (200 mL), dried over sodium sulphate and solvent was removed under reduced pressure. Purification by flash chromatography on silica (0 – 100% EtOAc in hexane) yielded the title compound in 46% yield as a clear oil (1.04 g, 3.2 mmol).

^1H NMR (300 MHz, CDCl_3) δ 5.45 – 5.25 (m, 2H, CCHCH_2), 5.20 – 4.93 (m, 1H, CCHCH_2), 4.63 (t, $J = 3.5$ Hz, 1H, OCHO), 4.36 – 4.01 (m, 2H, CCHCH_2O), 3.99 (s, 2H, OHCH_2), 3.95 – 3.40 (m, 2H, OCH_2CH_2), 2.15 – 2.00 (m, 8H, 4 x allylic CH_2), 1.70 – 1.64 (m, 6H, $\text{CHCH}_2\text{CH}_2\text{CH}_2$), 1.64 – 1.47 (m, 9H, 3 x CH_3).

^{13}C NMR (75 MHz, CDCl_3) δ 143.36, 134.72 (3 x $\text{C}=\text{CH}$), 125.90, 124.28, 122.73 (3 x $\text{C}=\text{CH}$), 97.92 (OCO), 67.50 (CH_2OH), 63.25, 63.21 (2 x OCH_2), 41.90, 41.25 (2 x CCH_2CH_2), 31.05 (OCCH_2), 26.24, 25.96 (2 x CCH_2CH_2), 23.40 ($\text{OCH}_2\text{CH}_2\text{CH}_2$), 19.21 ($\text{OCH}_2\text{CH}_2\text{CH}_2$), 15.85, 15.41, 13.69 (CCH_3).

HRMS (ES^+): calculated for $(\text{C}_{20}\text{H}_{34}\text{O}_3 + [\text{Na}]^+)$: 345.2406, found: 345.2409.

(6*E*,10*E*)-2,6,10-Trimethyl-12-((tetrahydro-2H-pyran-2-yl)oxy)dodeca-2,6,10-trien-5-ol **183**



To a stirring suspension of selenium dioxide (777 mg, 7 mmol) and salicylic acid (966 mg, 7 mmol) in dichloromethane (100 mL) was added *t*BuOOH (11.2 mL, 70% in water, 114 mmol) and the mixture was stirred at rt for 10 min. The mixture was cooled to 0 °C, and 2-(((2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran **181** (7.0 g, 22.7 mmol) in dichloromethane (20 mL) was added. The reaction mixture was stirred overnight at 0 °C. The suspension was diluted with ether (100 mL) and successively washed with sat. aq. sodium hydrogen carbonate (200 mL), sat. aq. copper sulphate (200 mL), sat. aq. sodium sulphite (200 mL), water (200 mL), brine (200 mL), dried over sodium sulphate and solvent was removed under reduced pressure. Purification by flash chromatography on silica (0 – 100% EtOAc in hexane) yielded the title compound in 14% yield as a clear oil (313 mg, 1.0 mmol).

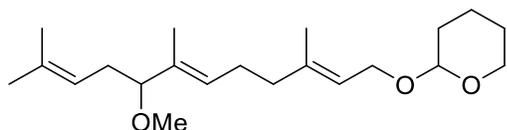
^1H NMR (300 MHz, CDCl_3) δ 5.39-5.35 (2 H, m, 2 x CCHCH_2), 5.08 (1 H, t, $J = 7.0$, CCHCH_2), 4.63 (1 H, m, OCHO), 4.33 – 4.16 (1 H, m, CCHCH_AO), 4.05 – 3.95 (2 H, m, CCHCH_BO and CHOH), 3.88 (1 H, t, $J = 7.5$, CHOH), 3.64 – 3.42 (2 H, m, OCH_2CH_2), 2.37 – 2.01 (6 H, m, CHCH_2CH and CHCH_2CH_2), 1.88-1.50 (6 H, m, $\text{CHCH}_2\text{CH}_2\text{CH}_2$), 1.72, 1.68, 1.64 and 1.58 (4 x 3 H, 4 x s, 4 x CH_3).

^{13}C NMR (125 MHz, CDCl_3) δ 142.23, 136.53 and 134.65 (3 x CH_3CCH), 124.92, 122.02 and 120.54 (3 x CH_3CCH), 97.22 (OCHO), 63.94 (CHOH), 63.74 and 62.57 (2 x CH_2O), 39.65

(CH₃CCH₂), 33.84 (CH₂CHOH), 30.92 (CH₂OCHCH₂), 26.45 (OCH₂CH₂CH₂), 25.26 (OCH₂CH₂CH₂), 19.95, 18.32, 15.27 and 11.56 (CH₃).

HRMS (ES⁺): calculated for (C₂₀H₃₄O₃ + [Na])⁺: 345.2406, found: 345.2419.

2-(((2*E*,6*E*)-8-Methoxy-3,7,11-trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran **185**



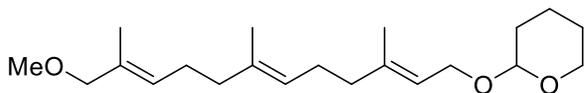
To a stirred suspension of sodium hydride (0.24 g, 60% in mineral oil, 6 mmol) in anhydrous THF (50 mL), was added a solution of the alcohol **183** (313 mg, 1.0 mmol) in anhydrous THF (20 mL) and the reaction mixture was stirred for 15 min at rt. Methyl iodide (0.933 mL, 15 mmol) was then added. Reaction was quenched with HCl (20 mL, 1M) at rt when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 20 hour). The aqueous layer was separated and extracted with ether (3 x 30 mL). The combined organic fractions were washed with NaOH (30 mL, 1 M), brine (2 x 50 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (20% EtOAc in hexane) yielded the title compound in 69% yield as a clear oil (231 mg, 0.7 mmol).

¹H NMR (300 MHz, CDCl₃) δ 5.37 (1 H, t, *J* = 6.4, CCHCH₂), 5.30 (1 H, t, *J* = 6.4, CCHCH₂), 5.09 – 4.95 (1 H, m, CCHCH₂), 4.65 – 4.58 (2 H, m, OCHO), 4.33 – 3.96 (2 H, m, CCHCH₂O), 3.93 – 3.81 (2 H, m, OCH₂CH₂), 3.50 (1 H, t, *J* = 7.0, CHOCH₃), 3.15 (3 H, s, OCH₃), 2.34–2.04 (6 H, m, CCHCH₂CH and CCHCH₂CH₂), 1.75–1.52 (6 H, m, OCH₂CH₂CH₂), 1.68, 1.59, and 1.52 (2 x 3 H and 1 x 6 H, 3 x s, 4 x CHCCH₃).

¹³C NMR (125 MHz, CDCl₃) δ 139.86, 134.19 and 132.82 (3 x CH₃CCH), 128.44, 120.92 and 120.62 (3 x CH₃CCH), 102.69 (OCHO), 87.43 (CHOCH₃), 63.87 and 62.46 (OCH₂), 55.95 (OCH₃), 39.33 (CH₃CCH₂), 32.50 (CCHCH₂CH), 30.92 (CH₂OCHCH₂CH₂), 26.05 (CH₃CCH₂CH₂), 25.81 (CHCCH₃), 25.50 (OCH₂CH₂CH₂), 19.82 (OCH₂CH₂CH₂), 17.97, 16.39 and 10.46 (3 x CHCCH₃).

HRMS (ES⁺): calculated for (C₂₁H₃₆O₃ + [Na])⁺: 359.2562, found: 359.2557.

2-(((2*E*,6*E*,10*E*)-12-Methoxy-3,7,11-trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran **186**



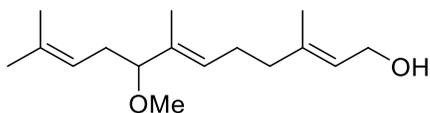
To a stirred suspension of sodium hydride (0.8 g, 60% in mineral oil, 20 mmol) in anhydrous THF (50 mL), was added a solution of the alcohol **182** (1.05 g, 3.3 mmol) in anhydrous THF (20 mL) and the reaction mixture was stirred for 15 min at rt. Methyl iodide (3.11 mL, 50 mmol) was then added. Reaction was quenched with HCl (50 mL, 1M) at rt when complete consumption of starting material was observed via TLC (20% ethyl acetate in hexane) (usually after 20 hour). The aqueous layer was separated and extracted with ether (3 x 30 mL). The combined organic fractions were washed with NaOH (60 mL, 1 M), brine (2 x 50 mL), dried over sodium sulphate, filtered, and solvent was removed under reduced pressure. Purification by flash chromatography on silica (20% EtOAc in hexane) yielded the title compound in 67% yield as a clear oil (750 mg, 2.2 mmol).

$^1\text{H NMR}$ (300 MHz, CDCl_3) δ 5.49 – 5.28 (2 H, m, 2 x CCHCH_2), 5.11 (1 H, t, $J = 6.1$, CCHCH_2), 4.70 – 4.57 (2 H, m, OCHO), 4.26 (1 H, dd, $J = 12.0$ and 6.5 , CCHCH_AO), 4.02 (1 H, dd, $J = 12.0$ and 7.5 , CCHCH_BO), 3.95-3.44 (2 H, m, OCH_2CH_2), 3.78 (2 H, s, CH_2OCH_3), 3.26 (3 H, s, OCH_3), 2.29 – 1.96 (8 H, m, 2 x CHCH_2CH_2), 1.70-1.54 (6 H, m, $\text{OCH}_2\text{CH}_2\text{CH}_2$), 1.68, 1.63, and 1.58 (3 x 3 H, 3 x s, 3 x CCH_3).

$^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 140.27, 134.94 and 132.45 (3 x CCHCH_2), 128.67, 124.16 and 121.18 (3 x CCHCH_2), 98.29 (OCHO), 79.92 (CH_2OCH_3), 64.05 and 63.67 (2 x CHOCH_2), 51.73 (CH_2OCH_3), 42.28 and 39.30 (2 x CH_3CCH_2), 31.39 (OCHCH_2), 27.47 and 26.30 (2 x $\text{CCHCH}_2\text{CH}_2$), 25.51 ($\text{OCH}_2\text{CH}_2\text{CH}_2$), 19.65 ($\text{OCH}_2\text{CH}_2\text{CH}_2$), 16.45, 16.12 and 13.87 (3 x CCH_3).

HRMS (AP^+): calculated for ($\text{C}_{21}\text{H}_{36}\text{O}_3 + [\text{Na}]^+$): 359.2549, found: 359.2562.

(2*E*,6*E*)-8-Methoxy-3,7,11-trimethyldodeca-2,6,10-trien-1-ol **143**



To a stirred solution of **185** (231 mg, 0.7 mmol) in THF (10 mL) was added a solution a aqueous solution of hydrochloric acid (3 mL, 10%) and was stirred overnight. The reaction was quenched with sat. sodium bicarbonate solution (10 mL) and the aqueous layer was separated. The aqueous layer was extracted with diethyl ether (3 x 10 mL) and the combined

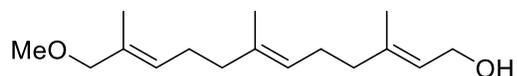
organic fractions were washed with brine (30 mL), dried over magnesium sulphate, filtered, and solvent removed under reduced pressure. Purification by flash chromatography on silica (20% EtOAc in hexane) yielded the title compound in 86% yield as a clear oil (142 mg, 0.6 mmol).

^1H NMR (300 MHz, CDCl_3) δ 5.51 – 5.37 (1 H, m, CCHCH_2), 5.32 (1 H, t, $J = 6.6$, CCHCH_2), 5.09 – 4.99 (1 H, m, CCHCH_2), 4.17 (2 H, d, $J = 6.8$, CH_2OH), 3.43 (1 H, t, $J = 7.1$, CHOCH_3), 3.18 (3 H, s, OCH_3), 2.43 – 2.02 (6 H, m, CHCH_2CHO and CHCH_2CH_2), 1.68 (6 H, s, 2 x CCH_3), 1.61 and 1.54 (2 x 3 H, 2 x s, 2 x CCH_3).

^{13}C NMR (125 MHz, CDCl_3) δ 141.6, 134.92 and 134.34 (3 x CCHCH_2), 127.50, 123.58 and 121.62 (3 x CCHCH_2), 87.69 (CHOCH_3), 59.34 (CH_2OH), 56.73 (OCH_3), 39.27 (CH_3CCH_2), 32.56 (CCHCH_2CH), 27.06 ($\text{CH}_3\text{CCH}_2\text{CH}_2$), 25.81, 17.97, 16.29 and 10.74 (4 x CHCCH_3).

HRMS (ES^+): calculated for ($\text{C}_{16}\text{H}_{28}\text{O}_2 + [\text{Na}]^+$): 275.1987, found: 275.1978.

(2*E*,6*E*,10*E*)-12-Methoxy-3,7,11-trimethyldodeca-2,6,10-trien-1-ol **142**



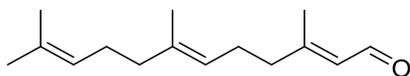
To a stirred solution of **186** (750 mg, 2.2 mmol) in methanol (20 mL) was added *p*-toluenesulfonic acid (38 mg, 0.2 mmol) and stirred at room temperature for 1 h before it was quenched with sat. sodium bicarbonate solution (20 mL). The reaction mixture was extracted with ether (3 x 10 mL) and combined organic fractions were washed with brine (30 mL), dried over anhydrous magnesium sulphate, filtered, and solvent was removed under reduced pressure. to give the title compound as a colourless oil (91 mg, 80%). Purification by flash chromatography on silica (20% EtOAc in hexane) yielded the title compound in 73% yield as a clear oil (408 mg, 1.6 mmol).

^1H NMR (400 MHz, CDCl_3) δ 5.50 – 5.32 (2 H, m, 2 x CCHCH_2), 5.18 – 5.06 (1 H, m, CCHCH_2), 4.15 (2 H, d, $J = 6.9$, CH_2OH), 3.78 (2 H, s, CH_2OCH_3), 3.27 (3 H, s, CH_2OCH_3), 2.20 – 1.98 (8 H, m, 2 x CHCH_2CH_2), 1.68, 1.63, and 1.60 (3 x 3 H, 3 x s, 3 x CCH_3).

^{13}C NMR (125 MHz, CDCl_3) δ 139.67, 135.02 and 131.94 (3 x CCHCH_2), 128.07, 124.05 and 123.41 (3 x CCHCH_2), 78.68 (CH_2OCH_3), 59.42 (CH_2OH), 57.36 (CH_2OCH_3), 39.50 and 39.26 (2 x CH_3CCH_2), 26.43 and 26.24 ($\text{CCHCH}_2\text{CH}_2$), 16.30, 15.99 and 13.82 (CCH_3).

HRMS (ES^+): calculated for ($\text{C}_{16}\text{H}_{28}\text{O}_2 + [\text{Na}]^+$): 275.1987, found: 275.1983.

(2*E*,6*E*)-3,7,11-Trimethyldodeca-2,6,10-trienal



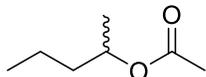
To a stirred solution of farnesol (2.5 mL, 10 mmol) in THF (60 mL) was added manganese dioxide (9 g, 104 mmol) and stirred for 48 h at rt. The reaction mixture was filtered through celite and washed with ether. The solvent was removed under reduced pressure. Purification by flash chromatography on silica (20% EtOAc in hexane) yielded the title compound in 21% yield as a clear oil (460 mg, 2.1 mmol).

^1H NMR (400 MHz, CDCl_3) δ 9.99 (d, $J = 8.1$ Hz, 1H, CHCHO), 5.87 (d, $J = 8.0$ Hz, 1H, CHCHO), 5.27 – 4.97 (m, 2H, 2 x $\text{C}=\text{CH}$), 2.33 – 1.96 (m, 8H, 2 x CH_2CH_2), 2.17 (d, $J = 0.8$ Hz, 3H, CH_3), 1.69, 1.61, 1.60 (3 x 3 H, 3 x s, 3 x CH_3).

^{13}C NMR (101 MHz, CDCl_3) δ 191.38 (CHO), 163.99 (CCHCHO), 136.57, 131.39 (2 x CCHCH_2), 127.40 (CCHCHO), 124.19, 122.42 (2 x CCHCH_2), 40.61, 39.64 (2 x CHCH_2CH_2), 26.61, 25.60 (2 x CHCH_2CH_2), 17.76, 17.55, 16.05, 14.11 (4 x CH_3).

HRMS (ES^+): calculated for ($\text{C}_{15}\text{H}_{24}\text{O} + [\text{H}]^+$): 221.1905, found: 221.1913.

rac-Pentan-2-yl acetate *rac*-188



Pentanol (1 mL, 9.2 mmol) was added to a solution of 4-dimethylaminopyridine (2.2 g, 18.4 mmol) and glacial acetic acid (1 mL, 18.4 mmol) in dichloromethane (50 mL). Dicyclohexyl carbodiimide (3.8 g, 18.4 mmol) was subsequently added and the solution was stirred overnight at room temperature. The precipitated solids were filtered off, the filtrate was concentrated in vacuo and purified by silica flash chromatography (ethyl acetate/pentane 1:4) to yield a colorless liquid (0.85 g, 6.5 mmol, 70%).

^1H NMR (500 MHz, CDCl_3): δ 4.96 – 4.85 (m, 1H, CHOCO), 2.02 (s, 3H, CH_3CO), 1.63 – 1.27 (m, 4H, 2 x CH_2), 1.20 (d, $J = 6.3$ Hz, 3H, CH_3CH), 0.91 (t, $J = 7.3$ Hz, 3H, CH_3CH_2) ppm.

^{13}C NMR (126 MHz, CDCl_3) δ 171.2 (CO), 71.2 (CHOCO), 38.5 (CH_2CHO), 21.8 (COCH_3), 20.4 (CHCH_3), 19.0 (CH_3CH_2), 14.3 ($\text{CH}_3\text{CH}_2\text{CH}_2$).

6.3. FLOW METHODS

6.3.1. General experimental

FDP, 12-OH FDP, farnesal were synthesized from commercial (2*E*,6*E*)-farnesol (see section 6.2). 2-Pentylacetate was synthesised from 2-pentanol (see section 6.2). All other chemicals were purchased from Sigma-Aldrich, Acros Chemicals or Alpha Aesar unless otherwise stated, solvents were obtained HPLC grade from Sigma-Aldrich.

Aqueous solution of lipase B from *C. antarctica* (CaLB) with a declared lipase activity of 5000 LU/g of liquid was purchased from Strem chemicals.

All ¹H and ¹³C NMR spectra were obtained on either a Bruker Avance 500 NMR (¹H: 500 MHz, ¹³C: 125 MHz) spectrometer at 25 °C in the solvent stated and are reported as chemical shifts in parts per million downfield from tetramethylsilane and all coupling constants, J, (to the nearest 0.5 Hz) in Hz. Multiplicities are reported using the following notation: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet.

Design-Expert® version 11 software was used to design design of experiment (DoE) study and analyse the data.

HPCCC

All high performance counter current chromatography (HPCCC) experiments were performed using a were performed on a Dynamic Extractions Spectrum instrument (Slough, UK) which was fitted with an analytical scale column with a column volume of 22 mL (0.8 mm I.D. PTFE tubing) and a revolution radius of 85 mm and a semi-preparative scale column with a column volume of 135 mL (1.6 mm I.D. of 1.6 mm PTFE tubing) and a revolution radius of 85 mm. β -value range of 0.52–0.86. The Spectrum HPCCC was connected to two HPLC pumps (ECOM ECP 2010), manually controlled fluidic and sample injection valves.

Segmented flow chemistry

The flow reactor (2 mL, 0.5 mm I.D. tubing) was constructed from PTFE tubing (Diba, Kinesis Ltd). The two liquid solutions (aqueous and organic) were loaded in glass syringe and injected to the flow reactor through a T-piece by using two syringe pumps (Fusion 100 Touch infusion syringe pump, KR Analytical Ltd), and the reaction mixture was collected in a glass vial at exit.

Oscillatory flow segmented chemistry

Modules of Modular-Lab standard (Eckert & Ziegler, Berlin, Germany) were used to set up the oscillatory segmented flow system with a flow reactor (4 mL, 0.5 mm I.D. tubing) was constructed from PTFE tubing (Diba, Kinesis Ltd).

6.3.2. GC-MS analysis

Gas chromatography coupled with mass spectrum (GC-MS) was performed on a Perkin Elmer Clarus 680 GC fitted with a Perkin Elmer Elite-1 column (30 m x 0.25 mm internal diameter) and a Perkin Elmer Clarus SQ 8 C mass spectrometer. The program uses an injection port temperature of 100 °C; split ratio 19:1; initial temperature 80 °C hold 2 min, ramp of 8 °C/min to 280 °C hold 3 min.

6.3.3. GC FID analysis

GC FID method **A** uses an injection port temperature of 200 °C, 5 µL was injected with a 19:1 split. The oven temperature was held at 80 °C for 1 min and then rose at 8 °C/min to 200 °C and then held for 2 min.

GC FID method **B** uses an injection port temperature of 200°C; 5 µL, split ratio 10:1; initial temperature 80 °C hold 1 min, ramp of 8 °C/min to 150 °C hold 2 min.

GC FID method **C** uses an injection port temperature of 200 °C; 5 µL, split ratio 10:1; initial temperature 80 °C hold 10 min, ramp of 8 °C/min to 100 °C hold 2 min.

6.3.4. Sesquiterpene synthases, HPLCC, flow, batch protocols

Incubation buffers and enzyme concentrations

Throughout chapter 4, AS, ADS and GDS were incubated with FDP or 12-OH FDP under various incubation systems, namely HPCCC, flow segmented system or batch. Incubation buffers and enzyme concentrations were kept identical if possible (see Table 6-19).

Table 6-19 List of incubation buffer and enzyme concentration used in chapter 3.

	Enzyme	Buffer	Glycerol (%)	Reducing agent	Enzyme (μM)
Batch			15	β -Me (5 mM)	
Flow	AS with FDP	Tris base (20 mM), MgCl_2 (3 mM), pH = 7.5.	15	β -Me (5 mM)	6
HPCCC			-	-	
Batch			10	DTT (1 mM)	
Flow	ADS with FDP	HEPES (25 mM), MgCl_2 (5 mM), pH = 7.5.	10	DTT (1 mM)	10
HPCCC			-	-	
Batch			10	β -Me (5 mM)	
Flow	GDS with FDP	Tris base (20 mM), MgCl_2 (10 mM), pH = 7.5	10	β -Me (5 mM)	12
HPCCC			10	β -Me (5 mM)	
Batch			-	DTT (1 mM)	
Flow	ADS with 12-OH FDP	Glycine (25 mM), NaOH (25 mM), MgCl_2 (5 mM), pH = 9.4	-	DTT (1 mM)	10
HPCCC			-	-	

HPCCC protocols

Method 1 (analytical column)

Aqueous and organic were pumped using the two HPLC pumps, temperature was regulated by a chiller to 28 °C. The column was initially filled with pentane (HPLC grade, Sigma-Aldrich) at 6 mL \cdot min⁻¹ from tail periphery to head-centre. The analytical column was then loaded through a loop with 10 mL AS incubation buffer (20 mM Tris base, 5 mM 2-mercaptoethanol, 10% glycerol, 3 mM MgCl_2 , pH = 7.5) containing 6 μM AS and 0.35 mM FDP at a flow rate of 0.5 mL \cdot min⁻¹ resulting in 10 mL of pentane being displaced out of the column. The instrument was rotated at 1600 rpm and pentane was pumped at the same time from tail periphery to head-centre at a flow rate of 0.5 mL \cdot min⁻¹. The pentane was collected in 5 mL fractions and analysed by GC FID (method A). The yield was calculated by comparing peak areas of the product to a calibration curve using α -humulene.

Method 2 (analytical column)

Aqueous and organic were pumped using the two HPLC pumps, temperature was regulated by a chiller to 28 °C. The column was initially filled with pentane (HPLC grade, Sigma-Aldrich) at 6 mL•min⁻¹ from tail periphery to head-centre. The analytical column was then loaded through a loop with 10 mL of incubation buffer containing a set concentration of protein and 0.35 mM FDP at a set flow rate resulting in 10 mL of pentane being displaced out of the column. The instrument was rotated at 1600 rpm and pentane was pumped at the same time from tail periphery to head-centre at a set flow rate. The first 10 mL of pentane were discarded and then two fractions (respectively 20 mL and 10 mL) were collected and analysed by GC FID (method **A**). The yield was calculated by comparing peak areas of the product to a calibration curve using α -humulene.

Method 3 (preparative column)

Aqueous and organic were pumped using the two HPLC pumps; temperature was regulated by a chiller to 28 °C. The preparative column was initially filled with pentane (HPLC grade, Sigma-Aldrich) at 10 mL•min⁻¹ from tail periphery to head-centre. The column was then loaded through a HPLC pump with 65 mL of incubation buffer containing a set concentration of protein (6 μ M for AS) and 0.7 mM FDP at 5 mL•min⁻¹ resulting in 65 mL of pentane being displaced out of the column. The instrument was rotated at 1600 rpm and pentane was pumped at the same time from tail periphery to head-centre at a set flow rate. The first 70 mL of pentane were discarded and then 70 mL of pentane (approximately 0.5 CV) were collected and analysed by GC FID (method **A**). The yield was calculated by comparing peak areas of the product to a calibration curve using α -humulene.

Preparative scale

Aqueous and organic were pumped using the two HPLC pumps; temperature was regulated by a chiller to 28 °C. The preparative column was initially filled with pentane (HPLC grade, Sigma-Aldrich) at 10 mL•min⁻¹ from tail periphery to head-centre. The column was then loaded through a HPLC pump with 65 mL of incubation buffer containing a set concentration of protein (6 μ M for AS, 12 μ M for GDS, 10 μ M for ADS) and FDP (0.7 mM, 0.045 mmol) at 5 mL•min⁻¹ resulting in 65 mL of pentane being displaced out of the column. The instrument was rotated at 1600 rpm and pentane was pumped at the same time from tail periphery to head-centre at a set flow rate. The first 70 mL of pentane were discarded and then 100 mL (20 min) of pentane were collected. Experiment was repeated four times and organic phase was pooled

together, dried (Na_2SO_4), filtered, 80% of the solvent was removed by distillation at atmospheric pressure then under reduced pressure (800 mbar minimum pressure at 30 °C water bath temperature to avoid loss of volatile product) to give a colourless oil.

(1R,7R,8aS)-Aristolochene (72)

FDP (**14**) was incubated with AS described above to give (1R,7R,8aS)-aristolochene as a colourless oil (35 mg, 94%).

^1H NMR (500 MHz, CDCl_3): δ = 5.32 (dt, J = 5.5, 1.9 Hz, 1H), 4.74 – 4.69 (m, 2H), 2.27 – 2.09 (m, 2H), 2.06 – 1.96 (m, 2H), 1.93 – 1.82 (m, 1H), 1.77 (ddd, J = 8.5, 5.4, 3.1 Hz, 1H), 1.75 – 1.74 (m, 3H), 1.74 – 1.67 (m, 1H), 1.48 – 1.21 (m, 5H), 1.19 (dt, J = 18.6, 6.9 Hz, 1H), 0.97 (d, J = 0.6 Hz, 3H), 0.85 (d, J = 6.7 Hz, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ = 150.6 (C=CH₂), 144.5 (C=CH), 118.8 (C=CH), 108.3 (C=CH₂), 44.2 (CH₂CH₂CH₂CHC), 43.3 (CHCCH₂), 38.8 (CHCCH₂), 37.8 (CH₂=CCH), 32.6 (CH₂CH₂C=CH), 31.3 (CH₂CH=C), 31.1 (CH₃CHCH₂), 27.8 (CH₃CHCH₂CH₂), 20.9 (CH₂=CCH₃), 18.1 (CH₃C), 15.7 (CH₃CH).

(S)-Germacrene D (55)

FDP (**14**) was incubated with GDS as described above to give (S)-germacrene D as a colourless oil (34 mg, 92%).

^1H NMR (500 MHz, CDCl_3): δ = 5.71 (d, J = 15.9 Hz, 1H, H₂C=C-CH=CH), 5.18 (dd, J = 15.8, 9.9 Hz, 1H, H₂C=C-CH=CH), 5.06 (dd, J = 11.0, 4.8 Hz, 1H, CH₃C=CH), 4.72 (dd, J = 2.3, 0.8 Hz, 1H, C=CH₂), 4.67 (d, J = 2.3 Hz, 1H, C=CH₂), 2.30 – 2.21 (m, 10H, CH₃C=CH H₂C=CCH₂CH₂, (CH₃)₂CHCHCH₂), 0.86 – 0.78 (m, 6H, (CH₃)₂CH).

^{13}C NMR (126 MHz, CDCl_3): δ = 149.0 (C=CH₂), 135.6 (CH=CHC=CH₂), 134.1 (CH₃C=CH), 133.7 (CH=CHC=CH₂), 129.8 (CH₃C=CH), 109.2 (CH₂=C), 53.1 (CHCHCH=CH), 40.9 (CH₂CH₂C=CH), 34.5 (CH₂CH₂C=CH₂), 32.1 (CHCHCH=CH), 29.6 (C=CHCH₂CH₂), 25.0 (CH₂CH₂C=CHCH₂), 22.9 (CH₃CHCH), 19.5 (CH₃CHCH), 16.1 (CH₃C).

(1R,4R,4aS,8aR)-Amorpha-4,11-diene (49)

FDP (**14**) was incubated with ADS as described above and purified on silver (5%) impregnated silica (hexane:EtOAc 95:5) to give (1R,4R,4aS,8aR)-amorpha-4,11-diene as a colourless oil (31 mg, 84%).

^1H NMR (500 MHz, CDCl_3): δ = 5.07 (dd, J = 8.8, 1.4 Hz, 1H, CH=CCH₃), 4.90 – 4.85 (m, 1H, C=CH_A), 4.64 (s, 1H, C=CH_B), 2.64 – 2.51 (m, 1H, CH₂CCHCH), 2.01 – 0.92 (m, 11H,

$(CH_2)_2CHCH(CH_2)_2CH$, 1.74 (dd, $J = 3.6, 3.0$ Hz, 3H, $CH_2=CCH_3$), 1.60 (s, 3H, $CH=CCH_3$), 0.88 (t, $J = 5.6$ Hz, 3H, CH_2CHCH_3).

^{13}C NMR (126 MHz, $CDCl_3$): $\delta = 148.2$ ($CH_3C=CH_2$), 134.8 ($CH_3C=CH$), 120.7 ($CH_3C=CH$), 109.9 ($CH_3C=CH_2$), 47.8 ($CH_2=CCH$), 42.0 ($CH_3CCH_2CH_2CH$), 37.8 ($C=CHCH$), 35.6 (CH_3CHCH_2), 28.0 ($CH=CCH_2CH_2$), 26.6 (CH_3CCH_2), 26.2 (CH_3CH), 26.0 ($CH_3CHCH_2CH_2$), 23.8 ($CH_3C=CH$), 22.7 ($CH_3C=CH_2$), 20.0 (CH_3CHCH_2).

Dihydroartemisinic aldehyde (**112**, **113**) (mixture of 11-*R* and 11-*S* epimers)

12-OH FDP was incubated with ADS as described above. The crude oil was purified by flash chromatography on silica gel (Gradient 0 to 25% dichloromethane in pentane) to obtain compounds **112** and **113** as colourless oil (30 mg, 70%).

1H NMR (500 MHz, $CDCl_3$): $\delta = 9.63$ (d, $J = 4.0$ Hz, 1H, CHO , (11*S*)), 9.58 (d, $J = 3.9$ Hz, 1H, CHO , (11*R*)), 5.27 (d, $J = 1.1$ Hz, 1H, $CH_3C=CH$, (11*S*)), 5.13 (s, 1H, $CH_3C=CH$, (11*R*)), 2.52 – 2.44 (m, 1H, 2 x $C=CHCH$, (11*R*) and (11*S*)), 2.42 – 2.37 (m, 1H, 2 x $OCHCH$, (11*S*) and (11*R*)), 1.96 – 1.19 (m, 20H, 2 x $(CH_2)_2CHCH(CH_2)_2$, (11*R*) and (11*S*)), 1.63 (dd, $J = 8.4, 3.6$ Hz, 6H, 2 x CCH_3), 1.08 (d, $J = 6.8$ Hz, 3H, $OCHCHCH_3$, (11*R*)), 1.05 (d, $J = 2.2$ Hz, 3H, $OCHCHCH_3$, (11*S*)), 0.99 – 0.95 (m, 1H, CH_3CHCH_2), 0.87 (d, $J = 6.5$ Hz, 6H, 2 x $CHCH_3$, (11*R*) and (11*S*)).

^{13}C NMR (126 MHz, $CDCl_3$): $\delta = 206.2$ (CHO , (11*R*)), 206.0 (CHO , (11*S*)), 136.2 ($CH=CCH_3$, (11*R*)), 136.1 ($CH=CCH_3$, (11*S*)), 119.7 ($CH=CCH_3$, (11*S*)), 119.7 ($CH=CCH_3$, (11*R*)), 48.6 ($OCHCH$, (11*R*)), 48.0 ($OCHCH$, (11*S*)), 43.6 ($CH_3CCH_2CH_2CH$, (11*S*)), 42.0 ($CH_3CCH_2CH_2CH$, (11*R*)), 41.95 ($OCHCHCH$, (11*S*)), 41.6 ($OCHCHCH$, (11*R*)), 39.3 ($C=CHCH$, (11*S*)), 36.7 ($C=CHCH$, (11*R*)), 35.6 ($CH_3CHCH_2CH_2$, (11*S*)), 35.4 ($CH_3CHCH_2CH_2$, (11*R*)), 27.9 ($CH_3CHCH_2CH_2$, (11*S*)), 27.8 ($CH_3CHCH_2CH_2$, (11*R*)), 27.5 ($CH_3CHCH_2CH_2$, (11*R*)), 26.7 (CH_3CCH_2 , (11*R*)), 26.6 (CH_3CCH_2 , (11*S*)), 25.9 ($CH_2CH_2C=CH$, (11*R*)), 25.7 ($CH_3CHCH_2CH_2$, (11*S*)), 25.3 ($CH_2CH_2C=CH$, (11*S*)), 24.0 ($CH_3C=CH$, (11*S*)), 24.0 ($CH_3C=CH$, (11*R*)), 20.0 (CH_2CHCH_3 , (11*S*)), 19.8 (CH_2CHCH_3 , (11*R*)), 12.7 (CH_3CHCH_2 , (11*S*)), 11.9 (CH_3CHCH_2 , (11*R*)).

Artemisinic-11*S*,12-epoxide **186**

12-OH FDP (**180**) was incubated with ADS as described above. The crude oil was purified by flash chromatography on silica gel (Gradient 0 to 25% dichloromethane in pentane) to obtain Artemisinic-11*S*,12-epoxide as colourless oil (2 mg)

^1H NMR (500 MHz, CDCl_3) δ 5.17 (s, 1H, C=CH), 2.82 (d, $J = 4.7$ Hz, 1H, CH_AO), 2.62 – 2.58 (m, 1H, C=CHCH), 2.56 (dd, $J = 4.7, 0.5$ Hz, 1H, CH_BO), 1.98 – 1.76 (m, 4H, OCCH and CH=CCH₂), 1.65 – 1.61 (bs, 4H, CH₃C=CH), 1.55 (s, 4H, 2 x CH₂), 1.36 (s, 4H, CH₃CO and CH), 1.28 – 1.24 (m, 3H, CH₂ and CH), 0.88 (d, $J = 6.5$ Hz, 3H, CH₃CH).

^{13}C NMR (126 MHz, CDCl_3) δ 135.72 (C=CH), 120.42 (C=CH), 58.08 (COCH₂), 51.12 (CH₂O), 45.12 (CHCO), 41.72 (CH), 38.46 (C=CHCH), 34.86 (CH₃CHCH₂), 29.71 (CH=CCH₂CH₂CHCH), 27.97 (CH₂), 26.37 (CH=CCH₂), 23.83 (CH₃C=CH), 21.64 (CH₃CO), 21.63 (CH₂CHCO), 19.74 (CH₃CH).

HRMS (EI+): calculated for ($\text{C}_{15}\text{H}_{24}\text{O}$)⁺: 220.1827, found: 220.1818.

Segmented flow protocol

For all experiments, the aqueous solution (5 mL) was prepared by dissolving FDP (0.35 mM, 0.175 μmol) and enzyme in the incubation buffer. The organic solvent is pentane with α -humulene (35 μM) or farnesal (70 μM) as internal standard (IS). The reaction mixture was collected after reaching steady state (third reactor volume), the organic layer was directly analysed by GC FID (method A). The yield was calculated by comparing peak areas of product and internal standard.

Batch protocol

The enzyme was added to the incubation buffer to a set concentration at room temperature with gentle stirring then FDP or 12-OH FDP (0.35 mM, 0.35 μmol) was added and the incubation ($V_{\text{tot}} = 10$ mL) was overlaid with pentane (5 mL) and stirred at room temperature for 36 h. The pentane overlay was removed, and the aqueous layer was further extracted with pentane (2 x 5 mL) by gentle swirling and slow separation. The combined organic extracts were pooled together in a 20 mL volumetric flask (grade A+), and volume was adjusted with fresh pentane to a total volume of 20 mL. The solution was analysed by GC FID (method A). The yield was calculated by comparing peak areas of the product to a calibration curve using α -humulene or farnesal.

6.3.5. Lipase activity

The lipase activity was spectrophotometrically determined by following the hydrolysis of *p*-nitrophenyl butyrate (*p*NPB) at 400 nm. An aliquot of enzyme (10 μ L) was added to buffer (The initial lipase activity was 10514 U/mL (45 U/mg). One unit of lipase activity (U) is defined as the amount of lipase required to release 1 nmol of *p*NPB per minute at pH 7.2 at 37 °C using *p*-nitrophenylbutyrate as substrate.

Incubation buffer

Sodium dihydrogen phosphate monohydrate (0.46 g, 33 mM), disodium hydrogen phosphate heptahydrate (1.79 g, 67 mM) and sodium chloride (0.88 g, 150 mM) were dissolved in deionised water. Triton1 X-100 (0.5 mL, 0.5% (v/v)) was added the final volume was made up to 100 mL and stored at 4 °C.

*p*NPB solution

*p*NPB (10 mg, 50 mM) was added to acetonitrile (1 mL).

Lipase solution

Aqueous solution of lipase B from *C. antarctica* (CaLB) (Strem Chemical, Cambridge, United Kingdom) (10 μ L) was added to the incubation buffer (990 μ L) and stored at 4 °C.

Protocol

Lipase solution (10 μ L) was added to incubation buffer (900 μ L) in a cuvette and solution was gently mixed and put in a Shimadzu spectrophotometer at 37 °C. When absorbance at 400 nm was constant, *p*NPB solution (10 μ L) was added and solution was immediately mixed by inversion. Absorbance at 400 nm was recorded for 5 min and $\Delta A_{400}/\text{min}$ was calculated. Reaction was performed in triplicate and a blank recorded using the same protocol using deionized water instead of *p*NPB solution.

$$\frac{\text{Units}}{\text{mL}} \text{ enzyme} = \frac{(\Delta A_{400\text{nm}}^{\text{Lipase}}/\text{min} - \Delta A_{400\text{nm}}^{\text{Blank}}) * 1.01 * 100}{0.0148 * 0.1}$$

One unit will release 1.0 nanomole (10^{-9} mole) of *p*-nitrophenol per minute at pH 7.2 at 37 °C using *p*-nitrophenol butyrate as substrate.

6.3.6. Lipase HPCCC and batch protocols

Batch protocol

A solution of octanol (30 mM) and vinyl acetate (x mM depending on the ratio alcohol: vinyl acetate) in heptane was added to a solution of Cal B lipase ($1 \text{ mg}\cdot\text{mL}^{-1}$) in aqueous buffer (20 mM phosphate buffer, pH 7.2). The solution was stirred at 400 rpm at 25°C and organic layer was sampled at regular interval of time to determine the conversion of octanol to octyl acetate. The solution was analysed by GC FID (method **B**). The conversion was calculated by comparing peak areas of the product to a calibration curve using octyl acetate.

General HPCCC protocol

Aqueous and organic were pumped using the two HPLC pumps, temperature was regulated by a chiller at the desired temperature. The analytical column was initially filled with the incubation buffer (20 mM phosphate buffer, pH 7.2) containing CalB lipase at a set concentration at $6 \text{ mL}\cdot\text{min}^{-1}$ from tail periphery to head-centre. The bobbin was then rotated at 1200 rpm and a solution of known concentration of octanol and vinyl acetate in heptane (HPLC grade, Sigma-Aldrich) was pumped through the column at $2 \text{ mL}\cdot\text{min}^{-1}$. After 1 CV, the rotation speed and the flowrate were readjusted to the desired reaction condition. The solution at the exit of the HPCCC was then collected after 3 CV and analysed by GC-FID (method **B**). The conversion was calculated by comparing peak areas of the product to a calibration curve using octanol.

HPCCC recycling protocol

Aqueous and organic were pumped using the two HPLC pumps, temperature was regulated by a chiller (20 °C). The analytical column was initially filled with the incubation buffer (20 mM phosphate buffer, pH 7.2) containing Cal B lipase ($1 \text{ mg}\cdot\text{mL}^{-1}$) at $6 \text{ mL}\cdot\text{min}^{-1}$ from tail periphery to head-centre. The bobbin was then rotated at 1600 rpm and a solution of octanol (100 mM) and vinyl acetate (300 mM) in heptane was pumped through the column at $2 \text{ mL}\cdot\text{min}^{-1}$. After 1 CV, the rotation speed was readjusted to 1600 rpm. The solution at the exit of the HPCCC

was then collected at regular intervals and analysed by GC-FID (method **B**). The yield was calculated by comparing peak areas of the product to a calibration curve using octyl acetate. Isolated yield was obtained from collecting the equivalent to 3 residence time. Solvent was removed under reduced pressure and the crude oil was purified by flash chromatography on silica gel (20% EtOAc in hexanes) to obtain octyl acetate as colourless oil (597 mg, 97%).

^1H NMR (500 MHz, CDCl_3): δ = 4.05 (t, J = 6.8 Hz, 2H, CH_2O), 2.04 (s, 3H, OCOCH_3), 1.70 – 1.53 (m, 2H, CH_2), 1.40 – 1.20 (m, 10H, 5 x CH_2), 0.88 (t, J = 6.8 Hz, 3H, CH_3CH_2) ppm.

^{13}C NMR (126 MHz, CDCl_3): δ = 171.4 (CO), 64.8 (CHOCO), 31.9 (CH_2CHO), 29.4 (CH_2), 29.3 (CH_2), 28.8 (CH_2), 26.1 (CH_2), 22.8 (CH_2), 21.2 (CH_3CO), 14.2 (CH_3CH_2) ppm.

HRMS (EI): calculated for ($\text{C}_{10}\text{H}_{20}\text{O}_2$ - [H]) $^-$: 171.1385 found: 171.1380

HPCCC protocol for the stereoselective transesterification of 2-pentanol

Aqueous and organic were pumped using the two HPLC pumps, temperature was regulated by a chiller. The analytical column was initially filled with the incubation buffer (20 mM phosphate buffer, pH 7.2) containing CalB lipase ($1 \text{ mg}\cdot\text{mL}^{-1}$) at a set concentration at $6 \text{ mL}\cdot\text{min}^{-1}$ from tail periphery to head-centre. The bobbin was then rotated at 1600 rpm and a solution of rac-2-pentanol (100 mM) and vinyl acetate (300 mM) in pentane was pumped through the column at $2 \text{ mL}\cdot\text{min}^{-1}$. The solution at the exit of the HPCCC was then collected after 3 CV and solvent was removed under reduced pressure. The crude product was purified using flash chromatography (5% EtOAc in hexanes) to give (*R*)-2-pentyl acetate as a colourless liquid (49%, >99% e.e.). The enantiomeric excess was determined by GC-FID (method **C**) using synthesized rac-2-pentyl acetate from the following procedure: Pentanol (1 mL, 9.2 mmol) was added to a solution of 4-dimethylaminopyridine (2.2 g, 18.4 mmol) and glacial acetic acid (1 mL, 18.4 mmol) in dichloromethane (50 mL). Dicyclohexyl carbodiimide (3.8 g, 18.4 mmol) was subsequently added and the solution was stirred overnight at room temperature. The precipitated solids were filtered off, the filtrate was concentrated in vacuo and purified by silica flash chromatography (ethyl acetate/pentane 1:4) to yield a racemic mixture of 2-pentyl acetate as a colourless liquid (0.85 g, 6.5 mmol, 70%).

^1H NMR (500 MHz, CDCl_3): δ 4.96 – 4.85 (m, 1H, CHOCO), 2.02 (s, 3H, CH_3CO), 1.63 – 1.27 (m, 4H, 2 x CH_2), 1.20 (d, J = 6.3 Hz, 3H, CH_3CH), 0.91 (t, J = 7.3 Hz, 3H, CH_3CH_2) ppm.

^{13}C NMR (126 MHz, CDCl_3): δ 171.2 (CO), 71.2 (CHOCO), 38.5 (CH_2CHO), 21.8 (COCH_3), 20.4 (CHCH_3), 19.0 (CH_3CH_2), 14.3 ($\text{CH}_3\text{CH}_2\text{CH}_2$).

6.3.7. Data analysis

Data analysis was performed in Microsoft Excel. Every measurement was performed in triplicate unless otherwise specified. Average values were calculated using the equation below:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

The error bars corresponding to the standard deviation were calculated according to the equation below:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

REFERENCES

- [1] J. Buckingham, *Dictionary of Natural Products*, London, **1994**.
- [2] D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.
- [3] T. Toyomasu, T. Sassa, in *Compr. Nat. Prod. II*, Elsevier, **2010**, pp. 643–672.
- [4] E. M. Davis, in *Compr. Nat. Prod. II Chem. Biol.*, Elsevier Ltd, **2010**, pp. 585–608.
- [5] T. Kushiro, Y. Ebizuka, in *Compr. Nat. Prod. II Chem. Biol.*, Elsevier Ltd, **2010**, pp. 673–708.
- [6] J. Chappell, R. M. Coates, in *Compr. Nat. Prod. II Chem. Biol.*, Elsevier Ltd, **2010**, pp. 609–641.
- [7] B. L. Trumpower, *J. Bioenerg. Biomembr.* **1981**, *13*, 1–24.
- [8] B. Nowicka, J. Kruk, *Biochim. Biophys. Acta - Bioenerg.* **2010**, *1797*, 1587–1605.
- [9] G. E. Bartley, P. A. Scolnik, *Plant Cell* **1995**, *7*, 1027–1038.
- [10] J. Gershenzon, N. Dudareva, *Nat. Chem. Biol.* **2007**, *3*, 408–414.
- [11] J. A. Pickett, R. K. Allemann, M. A. Birkett, *Nat. Prod. Rep.* **2013**, *30*, 1277–1283.
- [12] R. J. Maude, C. J. Woodrow, L. J. White, *Drug Dev. Res.* **2010**, *71*, 12–19.
- [13] A. M. Barbuti, Z. S. Chen, *Cancers (Basel)*. **2015**, *7*, 2360–2371.
- [14] M. D. Leavell, D. J. McPhee, C. J. Paddon, *Curr. Opin. Biotechnol.* **2016**, *37*, 114–119.
- [15] N. K. B. K. Ikram, H. T. Simonsen, *Front. Plant Sci.* **2017**, *8*, 1966.
- [16] L. Ruzicka, *Experientia* **1953**, *9*, 357–367.
- [17] L. K. Henry, M. Gutensohn, S. T. Thomas, J. P. Noel, N. Dudareva, R. B. Croteau, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10050–10055.
- [18] B. Singh, R. A. Sharma, *3 Biotech* **2015**, *5*, 129–151.
- [19] Y. Hoshino, E. A. Gaucher, *Mol. Biol. Evol* **2018**, *35*, 2185–2197.
- [20] L. L. Grochowski, H. Xu, R. H. White, *J. Bacteriol.* **2006**, *188*, 3192–8.
- [21] H. M. Miziorko, *Arch. Biochem. Biophys.* **2011**, *505*, 131–143.
- [22] W. Eisenreich, M. Schwarz, A. Cartayrade, D. Arigoni, M. H. Zenk, A. Bacher, *Chem. Biol.* **1998**, *5*, R221–R233.
- [23] B. M. Lange, T. Rujan, W. Martin, R. Croteau, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13172–13177.
- [24] J. M. Vinokur, T. P. Korman, Z. Cao, J. U. Bowie, *Biochemistry* **2014**, *53*, 4161–4168.
- [25] N. Dellas, S. T. Thomas, G. Manning, J. P. Noel, *Elife* **2013**, *2*, e00672.
- [26] M. Chen, C. D. Poulter, *Biochemistry* **2010**, *49*, 207–217.
- [27] J. C. VanNice, D. A. Skaff, A. Keightley, J. K. Addo, G. J. Wyckoff, H. M. Miziorko, *J. Bacteriol.* **2014**, *196*, 1055–1063.
- [28] H. Hayakawa, K. Motoyama, F. Sobue, T. Ito, H. Kawaide, T. Yoshimura, H. Hemmi,

- Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 10034–10039.
- [29] J. S. Dickschat, *Nat. Prod. Rep.* **2011**, *28*, 1917–1936.
- [30] J. A. Aaron, D. W. Christianson, *Pure Appl. Chem.* **2010**, *82*, 1585–1597.
- [31] 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.
- [32] Y. Yoshikuni, T. E. Ferrin, J. D. Keasling, *Nature* **2006**, *440*, 1078–1082.
- [33] E. Y. Shishova, L. Di Costanzo, D. E. Cane, D. W. Christianson, *Biochemistry* **2007**, *46*, 1941–1951.
- [34] J. A. Aaron, X. Lin, D. E. Cane, D. W. Christianson, *Biochemistry* **2010**, *49*, 1787–1797.
- [35] J. Durairaj, A. Di Girolamo, H. J. Bouwmeester, D. de Ridder, J. Beekwilder, A. D. van Dijk, *Phytochemistry* **2019**, *158*, 157–165.
- [36] J. Degenhardt, T. G. Köllner, J. Gershenzon, *Phytochemistry* **2009**, *70*, 1621–1637.
- [37] A. Vattekkatte, S. Garms, W. Brandt, W. Boland, *Org. Biomol. Chem.* **2018**, *16*, 348–362.
- [38] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science (80-.)*. **1997**, *277*, 1815–1820.
- [39] 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.
- [40] M. W. van der Kamp, J. Sirirak, J. Žurek, R. K. Allemann, A. J. Mulholland, *Biochemistry* **2013**, *52*, 8094–8105.
- [41] J. M. Berg, Tymoczko; J. L., L. Stryer, in *Biochemistry* (Ed.: W.H. Freeman), New York, **2002**.
- [42] J. X. Li, X. Fang, Q. Zhao, J. X. Ruan, C. Q. Yang, L. J. Wang, D. J. Miller, J. A. Faraldos, R. K. Allemann, X. Y. Chen, P. Zhang, *Biochem. J.* **2013**, *451*, 417–426.
- [43] J. R. Mathis, K. Back, C. Starks, J. Noel, C. D. Poulter, J. Chappell, *Biochemistry* **1997**, *36*, 8340–8348.
- [44] D. E. Cane, H.-T. Chiu, P.-H. Liang, K. S. Anderson, *Biochemistry* **1997**, *36*, 8332–8339.
- [45] D. S. Reddy, E. J. Corey, *J. Am. Chem. Soc.* **2018**, *140*, 16909–16913.
- [46] D. Urabe, T. Asaba, M. Inoue, *Chem. Rev.* **2015**, *115*, 9207–9231.
- [47] T. J. Maimone, P. S. Baran, *Nat. Chem. Biol.* **2007**, *3*, 396–407.
- [48] S. Moser, H. Pichler, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 5501–5516.
- [49] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, D. R. Polichuk, K. H. Teoh, D. W. Reed, T. Treynor, J. Lenihan, H. Jiang, M. Fleck, S. Bajad, G. Dang, D. Dengrove, D. Diola, G. Dorin, K. W. Ellens, S. Fickes, J. Galazzo, S. P. Gaucher, T. Geistlinger, R. Henry, M. Hepp, T. Horning, T. Iqbal, L. Kizer, B. Lieu, D. Melis, N. Moss, R. Regentin, S. Secrest, H. Tsuruta, R. Vazquez, L. F. Westblade, L. Xu, M. Yu, Y. Zhang, L. Zhao, J. Lievens, P.

- S. Covello, J. D. Keasling, K. K. Reiling, N. S. Renninger, J. D. Newman, *Nature* **2013**, *496*, 528–532.
- [50] M. Dirkmann, J. Nowack, F. Schulz, *ChemBioChem* **2018**, *19*, 2146–2151.
- [51] V. C. A. Ward, A. O. Chatzivasileiou, G. Stephanopoulos, *Biotechnol. Bioeng.* **2019**, *116*, 3269–3281.
- [52] X. Chen, C. Zhang, R. Zou, G. Stephanopoulos, H. P. Too, *ACS Synth. Biol.* **2017**, *6*, 1691–1700.
- [53] Q. Zhang, K. Tiefenbacher, *Nat. Chem.* **2015**, *7*, 197–202.
- [54] Q. Zhang, L. Catti, L.-D. Syntrivanis, K. Tiefenbacher, *Nat. Prod. Rep.* **2019**, DOI 10.1039/c9np00003h.
- [55] Q. Zhang, J. Rinkel, B. Goldfuss, J. S. Dickschat, K. Tiefenbacher, *Nat. Catal.* **2018**, *1*, 609–615.
- [56] Q. Zhang, K. Tiefenbacher, *Angew. Chemie Int. Ed.* **2019**, *58*, 12688–12695.
- [57] L. R. MacGillivray, J. L. Atwood, *Nature* **1997**, *389*, 469–472.
- [58] L.-D. Syntrivanis, S. Levi, A. Prescimone, D. T. Major, K. Tiefenbacher, *ChemRxiv* **2019**, *Preprint*, DOI 10.26434/chemrxiv.9891341.v1.
- [59] S. Picaud, P. Mercke, X. He, O. Sterner, M. Brodelius, D. E. Cane, P. E. Brodelius, *Arch. Biochem. Biophys.* **2006**, *448*, 150–155.
- [60] S. Picaud, L. Olofsson, M. Brodelius, P. E. Brodelius, *Arch. Biochem. Biophys.* **2005**, *436*, 215–226.
- [61] P. Mercke, M. Bengtsson, H. J. Bouwmeester, M. A. Posthumus, P. E. Brodelius, *Arch. Biochem. Biophys.* **2000**, *381*, 173–180.
- [62] S. H. Kim, K. Heo, Y. J. Chang, S. H. Park, S. K. Rhee, S. U. Kim, *J. Nat. Prod.* **2006**, *69*, 758–762.
- [63] J. A. Faraldos, A. K. Antonczak, V. González, R. Fullerton, E. M. Tippmann, R. K. Allemann, *J. Am. Chem. Soc.* **2011**, *133*, 13906–13909.
- [64] 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.
- [65] J. A. Faraldos, R. K. Allemann, *Org. Lett.* **2011**, *13*, 1202–1205.
- [66] J. A. Faraldos, B. Kariuki, R. K. Allemann, *J. Org. Chem.* **2010**, *75*, 1119–1125.
- [67] M. J. Calvert, S. E. Taylor, R. K. Allemann, *Chem. Commun.* **2002**, *2*, 2384–2385.
- [68] D. J. Miller, F. Yu, D. W. Knight, R. K. Allemann, *Org. Biomol. Chem.* **2009**, *7*, 962–975.
- [69] D. J. Miller, F. Yu, R. K. Allemann, *ChemBioChem* **2007**, *8*, 1819–1825.
- [70] D. J. Miller, F. Yu, R. K. Allemann, *ChemBioChem* **2007**, *8*, 1819–1825.
- [71] F. Yu, D. J. Miller, R. K. Allemann, *Chem. Commun.* **2007**, 4155–4157.

- [72] D. J. Grundy, M. Chen, V. González, S. Leoni, D. J. Miller, D. W. Christianson, R. K. Allemann, *Biochemistry* **2016**, *55*, 2112–2121.
- [73] I. Prosser, I. G. Altug, A. L. Phillips, W. A. König, H. J. Bouwmeester, M. H. Beale, *Arch. Biochem. Biophys.* **2004**, *432*, 136–144.
- [74] C. O. Schmidt, H. J. Bouwmeester, S. Franke, W. A. König, *Chirality* **1999**, *11*, 353–362.
- [75] M. T. Ansari, Z. S. Saify, N. Sultana, I. Ahmad, S. Saeed-Ul-Hassan, I. Tariq, M. Khanum, *Mini Rev. Med. Chem.* **2013**, *13*, 1879–1902.
- [76] V. Harms, A. Kirschning, J. S. Dickschat, *Nat. Prod. Rep.* **2020**, DOI 10.1039/c9np00055k.
- [77] 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.
- [78] M. Demiray, D. J. Miller, R. K. Allemann, *Beilstein J. Org. Chem.* **2019**, *15*, 2184–2190.
- [79] A. Hou, L. Lauterbach, J. S. Dickschat, *Chem. – A Eur. J.* **2020**, *26*, 2178–2182.
- [80] V. Harms, B. Schröder, C. Oberhauser, C. D. Tran, S. Winkler, G. Dräger, A. Kirschning, *Org. Lett.* **2020**, *22*, 4360–4365.
- [81] C. Oberhauser, V. Harms, K. Seidel, B. Schröder, K. Ekramzadeh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat, A. Kirschning, *Angew. Chemie Int. Ed.* **2018**, *57*, 11802–11806.
- [82] M. Demiray, X. Tang, T. Wirth, J. A. Faraldos, R. K. Allemann, *Angew. Chemie Int. Ed.* **2017**, *56*, 4347–4350.
- [83] X. Tang, R. K. Allemann, T. Wirth, *European J. Org. Chem.* **2017**, *2017*, 414–418.
- [84] R. Lauchli, K. S. Rabe, K. Z. Kalbarczyk, A. Tata, T. Heel, R. Z. Kitto, F. H. Arnold, *Angew. Chemie Int. Ed.* **2013**, *52*, 5571–5574.
- [85] K. A. Rising, C. M. Crenshaw, H. J. Koo, T. Subramanian, K. A. H. Chehade, C. Starks, K. D. Allen, D. A. Andres, H. P. Spielmann, J. P. Noel, J. Chappell, *ACS Chem. Biol.* **2015**, *10*, 1729–1736.
- [86] Y. Yoshikuni, V. J. J. Martin, T. E. Ferrin, J. D. Keasling, *Chem. Biol.* **2006**, *13*, 91–98.
- [87] M. Loizzi, V. González, D. J. Miller, R. K. Allemann, *ChemBioChem* **2018**, *19*, 100–105.
- [88] R. Li, W. K. W. Chou, J. A. Himmelberger, K. M. Litwin, G. G. Harris, D. E. Cane, D. W. Christianson, *Biochemistry* **2014**, *53*, 1155–1168.
- [89] V. J. Davison, A. B. Woodside, T. R. Neal, K. E. Stremmer, M. Muehlbacher, C. D. Poulter, *J. Org. Chem.* **1986**, *51*, 4768–4779.
- [90] O. Cascón, S. Touchet, D. J. Miller, V. Gonzalez, J. A. Faraldos, R. K. Allemann, *Chem. Commun.* **2012**, *48*, 9702–9704.
- [91] M. Brodelius, A. Lundgren, P. Mercke, P. E. Brodelius, *Eur. J. Biochem.* **2002**, *269*, 3570–3577.

- [92] M. B. Plutschack, B. Pieber, K. Gilmore, P. H. Seeberger, *Chem. Rev.* **2017**, *117*, 11796–11893.
- [93] N. Weeranoppanant, *React. Chem. Eng.* **2019**, *4*, 235–243.
- [94] O. Cascón, G. Richter, R. K. Allemann, T. Wirth, *Chempluschem* **2013**, *78*, 1334–1337.
- [95] M. N. Kashid, D. W. Agar, S. Turek, *Chem. Eng. Sci.* **2007**, *62*, 5102–5109.
- [96] L. Nord, K. Bäckström, L. G. Danielsson, F. Ingman, B. Karlberg, *Anal. Chim. Acta* **1987**, *194*, 221–233.
- [97] L. Tamborini, P. Fernandes, F. Paradisi, F. Molinari, *Trends Biotechnol.* **2018**, *36*, 73–88.
- [98] R. C. Rodrigues, C. Ortiz, Á. Berenguer-Murcia, R. Torres, R. Fernández-Lafuente, *Chem. Soc. Rev.* **2013**, *42*, 6290–6307.
- [99] J. Britton, S. Majumdar, G. A. Weiss, *Chem. Soc. Rev.* **2018**, *47*, 5891–5918.
- [100] T. Peschke, P. Bitterwolf, S. Gallus, Y. Hu, C. Oelschlaeger, N. Willenbacher, K. S. Rabe, C. M. Niemeyer, *Angew. Chemie Int. Ed.* **2018**, *57*, 17028–17032.
- [101] C. J. Hartley, C. C. Williams, J. A. Scoble, Q. I. Churches, A. North, N. G. French, T. Nebl, G. Coia, A. C. Warden, G. Simpson, A. R. Frazer, C. N. Jensen, N. J. Turner, C. Scott, *Nat. Catal.* **2019**, *2*, 1006–1015.
- [102] M. P. Thompson, I. Peñafiel, S. C. Cosgrove, N. J. Turner, *Org. Process Res. Dev.* **2019**, *23*, 9–18.
- [103] J. H. Miyazaki, R. Croteau, *Enzyme Microb. Technol.* **1990**, *12*, 841–845.
- [104] D. Valikhani, P. L. Srivastava, R. K. Allemann, T. Wirth, *ChemCatChem* **2020**, *12*, 2194–2197.
- [105] I. I. Abdallah, R. Van Merkerk, E. Klumpenaar, W. J. Quax, *Sci. Rep.* **2018**, *8*, DOI 10.1038/s41598-018-28177-4.
- [106] F. Cramer, W. Rittersdorf, W. Böhm, *Justus Liebigs Ann. Chem.* **1962**, *654*, 180–188.
- [107] R. H. Cornforth, G. Popják, *Methods Enzymol.* **1969**, *15*, 359–390.
- [108] R. K. Keller, R. Thompson, *J. Chromatogr.* **1993**, *645*, 161–7.
- [109] L. L. Danilov, T. N. Druzhinina, N. A. Kalinchuk, S. D. Maltsev, V. N. Shibaev, *Chem. Phys. Lipids* **1989**, *51*, 191–203.
- [110] V. J. Davisson, A. B. Woodside, T. R. Neal, K. E. Stremler, M. Muehlbacher, C. D. Poulter, *J. Org. Chem.* **1986**, *51*, 4768–4779.
- [111] V. Jo Davisson, A. B. Woodside, C. Dale Poulter, *Methods Enzymol.* **1985**, *110*, 130–144.
- [112] M.-A. Dessoy, L. A. Wessjohann, *Silylierte Oligophosphate, Phosphate Und Phosphite Und Verfahren Zu Deren Herstellung Und Alkylierung*, **2004**, DE102004033306 A1.
- [113] L. A. Wessjohann, M. A. Dessoy, *Polyhedron* **2014**, *70*, 133–137.
- [114] M. D. Hartley, A. Larkin, B. Imperiali, *Bioorganic Med. Chem.* **2008**, *16*, 5149–5156.

- [115] W. D. Van Horn, C. R. Sanders, *Annu. Rev. Biophys.* **2012**, *41*, 81–101.
- [116] M. Lis, H. K. Kuramitsu, *Infect. Immun.* **2003**, *71*, 1938–1943.
- [117] L.-Y. Huang, S.-H. Huang, Y.-C. Chang, W.-C. Cheng, T.-J. R. Cheng, C.-H. Wong, *Angew. Chemie Int. Ed.* **2014**, *53*, 8060–8065.
- [118] L. L. Grochowski, H. Xu, R. H. White, *J. Bacteriol.* **2006**, *188*, 3192–8.
- [119] M. F. Mabanglo, H. L. Schubert, M. Chen, C. P. Hill, C. D. Poulter, *ACS Chem. Biol.* **2010**, *5*, 517–527.
- [120] N. Dellas, J. P. Noel, *ACS Chem. Biol.* **2010**, *5*, 589–601.
- [121] B. Miroux, J. E. Walker, *Over-Production of Proteins in Escherichia Coli: Mutant Hosts That Allow Synthesis of Some Membrane Proteins and Globular Proteins at High Levels*, **1996**.
- [122] X. Zhang, F. W. Studier, *J. Mol. Biol.* **1997**, *269*, 10–27.
- [123] A. I. Derman, W. A. Prinz, D. Belin, J. Beckwith, *Science (80-.)*. **1993**, *262*, 1744–1747.
- [124] V. Vagenende, M. G. S. Yap, B. L. Trout, *Biochemistry* **2009**, *48*, 11084–11096.
- [125] C. Nakano, F. Kudo, T. Eguchi, Y. Ohnishi, *ChemBioChem* **2011**, *12*, 2271–2275.
- [126] D. C. Crans, R. J. Kazlauskas, B. L. Hirschbein, C. H. Wong, O. Abril, G. M. Whitesides, *Methods Enzymol.* **1987**, *136*, 263–280.
- [127] T. J. Bruce, M. A. Birkett, J. Blande, A. M. Hooper, J. L. Martin, B. Khambay, I. Prosser, L. E. Smart, L. J. Wadhams, *Pest Manag. Sci.* **2005**, *61*, 1115–1121.
- [128] N. Bülow, W. A. König, *Phytochemistry* **2000**, *55*, 141–168.
- [129] 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.
- [130] W. S. Wadsworth, W. D. Emmons, *J. Am. Chem. Soc.* **1961**, *83*, 1733–1738.
- [131] B. E. Maryanoff, A. B. Reitz, *Chem. Rev.* **1989**, *89*, 863–927.
- [132] W. C. Still, C. Gennari, *Tetrahedron Lett.* **1983**, *24*, 4405–4408.
- [133] S. Fortin, F. Dupont, P. Deslongchamps, *J. Org. Chem.* **2002**, *67*, 5437–5439.
- [134] F. Messik, M. Oberthür, *Synth.* **2013**, *45*, 167–170.
- [135] J. D. Moore, P. R. Hanson, *Tetrahedron Asymmetry* **2003**, *14*, 873–880.
- [136] F.-W. Sum, L. Weiler, *Can. J. Chem.* **1979**, *57*, 1431–1441.
- [137] J. E. McMurry, W. J. Scott, *Tetrahedron Lett.* **1980**, *21*, 4313–4316.
- [138] M. A. Umbreit, K. B. Sharpless, *J. Am. Chem. Soc.* **1977**, *99*, 5526–5528.
- [139] A. Roy, A. Kucukural, Y. Zhang, *Nat. Protoc.* **2010**, *5*, 725–738.
- [140] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science (80-.)*. **1997**, *277*, 1815–1820.
- [141] D. J. Grundy, M. Chen, V. González, S. Leoni, D. J. Miller, D. W. Christianson, R. K.

- Allemann, *Biochemistry* **2016**, *55*, 2112–2121.
- [142] D. J. Grundy, R. K. Allemann, Expanding the Terpenome: Complementary Approaches to Novel Terpenoids, Cardiff University, **2015**.
- [143] K. Surendra, E. J. Corey, *J. Am. Chem. Soc.* **2008**, *130*, 8865–8869.
- [144] K. Surendra, E. J. Corey, *J. Am. Chem. Soc.* **2008**, *130*, 8865–8869.
- [145] X. Tang, R. K. Allemann, T. Wirth, *European J. Org. Chem.* **2017**, *2*, 414–418.
- [146] X. Tang, M. Demiray, T. Wirth, R. K. Allemann, *Bioorg. Med. Chem.* **2017**, *in press*.
- [147] T. Zor, Z. Selinger, *Anal. Biochem.* **1996**, *236*, 302–308.
- [148] S. Picaud, L. Olofsson, M. Brodelius, P. E. Brodelius, *Arch. Biochem. Biophys.* **2005**, *436*, 215–226.
- [149] C. O. Schmidt, H. J. Bouwmeester, S. Franke, W. A. König, in *Chirality*, **1999**.
- [150] M. J. Calvert, P. R. Ashton, R. K. Allemann, *J. Am. Chem. Soc.* **2002**, *124*, 11636–11641.
- [151] M. Demiray, X. Tang, T. Wirth, J. A. Faraldos, R. K. Allemann, *Angew. Chem. Int. Ed.* **2017**, *56*, 4347–4350.
- [152] “NIST/SEMATECH e-Handbook of Statistical Methods,” can be found under <http://www.itl.nist.gov/div898/handbook/>, **n.d.**
- [153] R. Leardi, *Anal. Chim. Acta* **2009**, *652*, 161–172.
- [154] W. Taam, in *Wiley StatsRef Stat. Ref. Online*, John Wiley & Sons, Ltd, Chichester, UK, **2014**.
- [155] J. Antony, in *Des. Exp. Eng. Sci.*, Elsevier, **2014**, pp. 87–112.
- [156] Stat-Ease Inc, **n.d.**
- [157] M. Abolhasani, K. F. Jensen, *Lab Chip* **2016**, *16*, 57–59.
- [158] H. Ye, S. Ignatova, H. Luo, Y. Li, A. Peng, L. Chen, I. Sutherland, *J. Chromatogr. A* **2008**, *1213*, 145–153.
- [159] Y. H. Guan, R. N. A. M. van den Heuvel, Y.-P. Zhuang, *J. Chromatogr. A* **2012**, *1239*, 10–21.
- [160] I. A. Sutherland, D. Fisher, *J. Chromatogr. A* **2009**, *1216*, 740–753.
- [161] Y. Ito, M. Weinstein, I. Aoki, R. Harada, E. Kimura, K. Nunogaki, *Nature* **1966**, *212*, 985–987.
- [162] A. Berthod, T. Maryutina, B. Spivakov, O. Shpigun, I. A. Sutherland, *Pure Appl. Chem* **2009**, *81*, 355–387.
- [163] C. DeAmicis, Q. Yang, C. Bright, N. A. Edwards, G. H. Harris, S. Kaur, P. L. Wood, P. Hewitson, S. Ignatova, *Org. Process Res. Dev.* **2017**, *21*, 1638–1643.
- [164] I. Sutherland, C. Thickitt, N. Douillet, K. Freebairn, D. Johns, C. Mountain, P. Wood, N. Edwards, D. Rooke, G. Harris, D. Keay, B. Mathews, R. Brown, I. Garrard, P. Hewitson,

- S. Ignatova, *J. Chromatogr. A* **2013**, *1282*, 84–94.
- [165] K. Skalicka-Woniak, I. Garrard, *Phytochem. Rev.* **2014**, *13*, 547–572.
- [166] O. R. Bousquet, J. Braun, F. Le Goffie, *Tetrahedron Lett.* **1995**, *36*, 8195–8196.
- [167] C. Nioi, D. Riboul, P. Destrac, A. Marty, L. Marchal, J. S. Condoret, *Biochem. Eng. J.* **2015**, *103*, 227–233.
- [168] J. A. Dietrich, Y. Yoshikuni, K. J. Fisher, F. X. Woolard, D. Ockey, D. J. McPhee, N. S. Renninger, M. C. Y. Chang, D. Baker, J. D. Keasling, *ACS Chem. Biol.* **2009**, *4*, 261–267.
- [169] C. Ortiz, M. L. Ferreira, O. Barbosa, J. C. S. Dos Santos, R. C. Rodrigues, Á. Berenguer-Murcia, L. E. Briand, R. Fernandez-Lafuente, *Catal. Sci. Technol.* **2019**, *9*, 2380–2420.
- [170] A. Basso, S. Serban, *Mol. Catal.* **2019**, *479*, 110607.
- [171] M. Rizzi, P. Stylos, A. Riek, M. Reuss, *Enzyme Microb. Technol.* **1992**, *14*, 709–714.
- [172] G. D. Yadav, A. H. Trivedi, *Enzyme Microb. Technol.* **2003**, *32*, 783–789.
- [173] V. Lange, L. Brown, P. Angeli, *Chem. Eng. Sci.* **2018**, *192*, 551–564.
- [174] L. Wang, X. Liu, Y. Jiang, L. Zhou, L. Ma, Y. He, J. Gao, *Catalysts* **2018**, *8*, 587.
- [175] A. Bukhari, A. Idris, M. Atta, T. C. Loong, *Cuihua Xuebao/Chinese J. Catal.* **2014**, *35*, 1555–1564.
- [176] G. D. Yadav, M. P. Kamble, *Int. J. Chem. React. Eng.* **2018**, *16*, DOI 10.1515/ijcre-2017-0179.
- [177] V. Leonard, S. Lamare, M. D. Legoy, M. Graber, *J. Mol. Catal. B Enzym.* **2004**, *32*, 53–59.
- [178] R. N. Patel, A. Banerjee, V. Nanduri, A. Goswami, F. T. Comezoglu, *JAOCS, J. Am. Oil Chem. Soc.* **2000**, *77*, 1015–1019.

APPENDIX

I. GAS CHROMATOGRAMS AND CALIBRATIONS CURVES

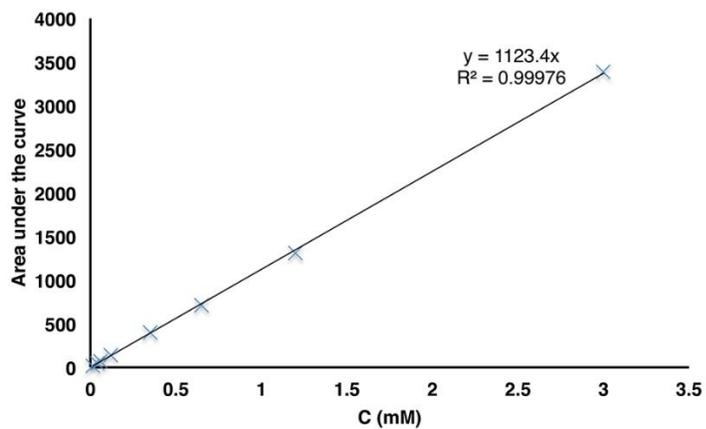


Figure A-1 Calibration curve for flame ionization detection response of α -humulene.

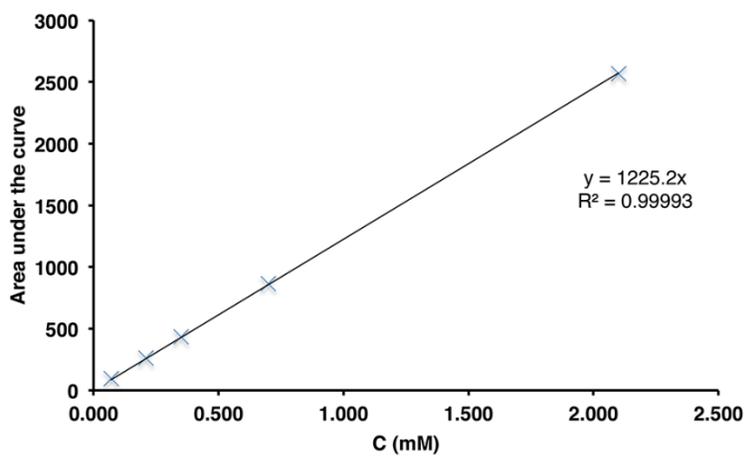


Figure A-2 Calibration curve for flame ionization detection response of farnesal.

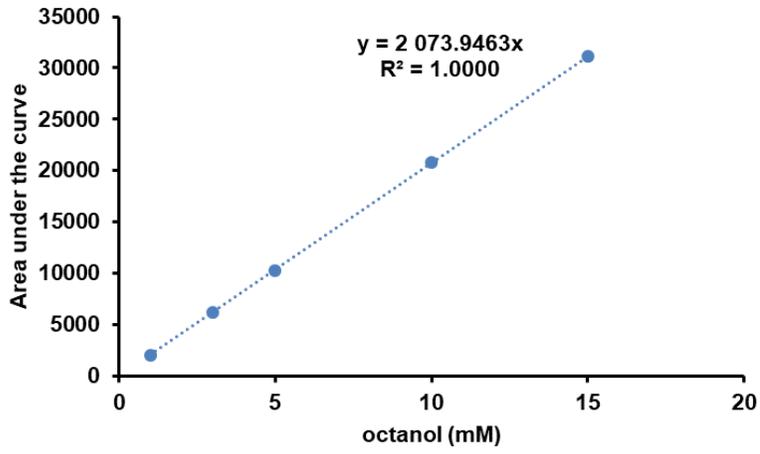


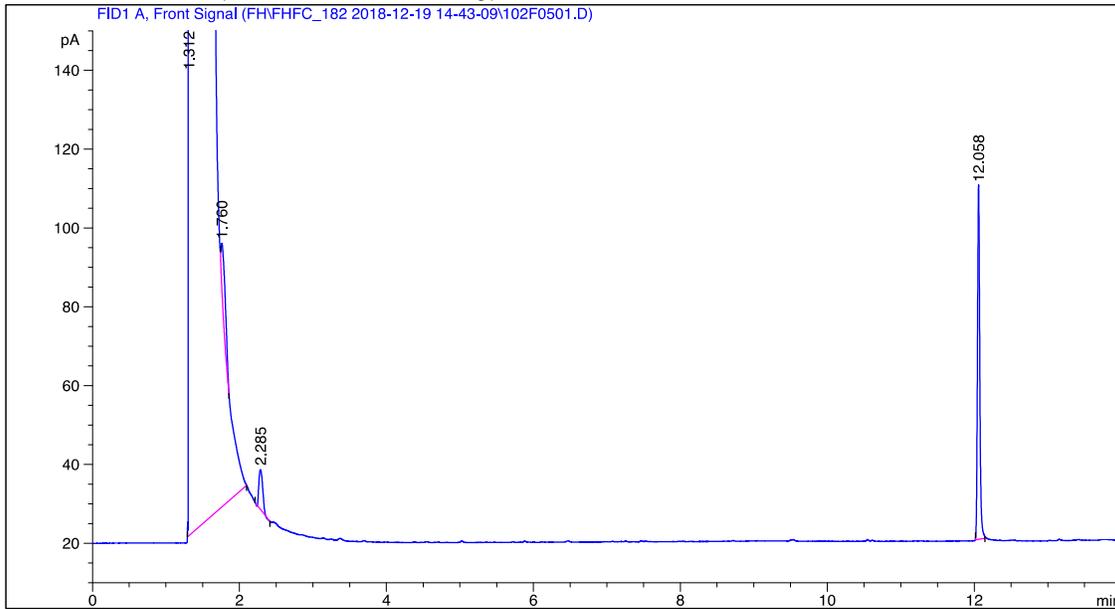
Figure A-3 Example of a calibration curve for flame ionization detection response of octanol.

Data File C:\CHEM32\1\DATA\FH\FHFC_182 2018-12-19 14-43-09\102F0501.D
 Sample Name: FH_182_run2

```

=====
Acq. Operator   : FH                               Seq. Line :    5
Acq. Instrument : Instrument 1                     Location  : Vial 102
Injection Date  : 12/19/2018 4:29:48 PM          Inj       :    1
                                                    Inj Volume: 5 µl

Acq. Method    : C:\Chem32\1\DATA\FH\FHFC_182 2018-12-19 14-43-09\FH_CALIBRATIONCURVESHORT.M
Last changed   : 8/15/2018 5:42:45 PM by RJM
Analysis Method: C:\CHEM32\1\DATA\FH\FHFC_182 2018-12-19 14-43-09\FH_CALIBRATIONCURVESHORT.M
Last changed   : 1/16/2019 10:41:10 AM by FH
                (modified after loading)
  
```



=====
 Area Percent Report
 =====

```

Sorted By      :      Signal
Multiplier     :      1.0000
Dilution       :      1.0000
Use Multiplier & Dilution Factor with ISTDs
  
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.312	BB S	0.0532	6.39093e6	1.45566e6	99.99543
2	1.760	BB X	0.0936	63.88154	11.37526	0.00100
3	2.285	BB	0.0599	35.20429	10.03268	0.00055
4	12.058	BB	0.0332	193.11363	89.56389	0.00302

Totals : 6.39122e6 1.45577e6

=====
 *** End of Report ***

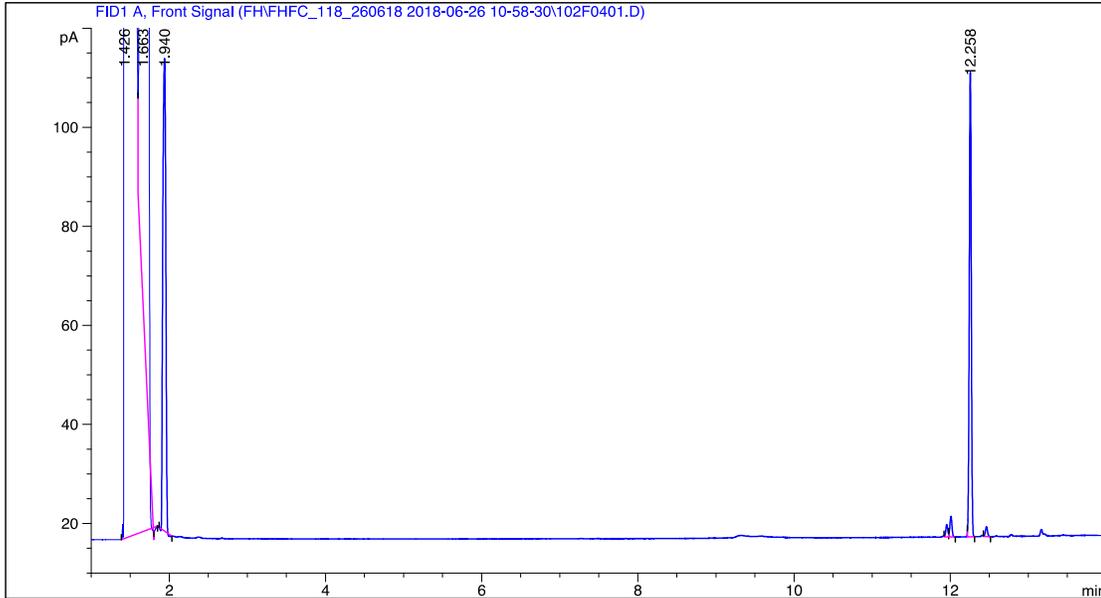
Figure A-4 Flame ionization detection response of AS with FDP by HPCCC using method 2.

Data File C:\CHEM32\1\DATA\FH\FHFC_118_260618 2018-06-26 10-58-30\102F0401.D
 Sample Name: FHfc_118_F2

```

=====
Acq. Operator   : RKA                               Seq. Line :    4
Acq. Instrument : Instrument 1                       Location  : Vial 102
Injection Date  : 6/26/2018 12:21:42 PM             Inj       :    1
                                                    Inj Volume: 5 µl

Acq. Method     : C:\Chem32\1\DATA\FH\FHFC_118_260618 2018-06-26 10-58-30\FH_
                  CALIBRATIONCURVESHORT.M
Last changed    : 6/21/2018 10:31:42 AM by RKA
Analysis Method : C:\CHEM32\1\DATA\FH\FHFC_118_260618 2018-06-26 10-58-30\FH_
                  CALIBRATIONCURVESHORT.M
Last changed    : 1/16/2019 10:43:05 AM by FH
                  (modified after loading)
  
```



Area Percent Report

```

Sorted By      : Signal
Multiplier     : 1.0000
Dilution      : 1.0000
Use Multiplier & Dilution Factor with ISTDs
  
```

Signal 1: FID1 A, Front Signal

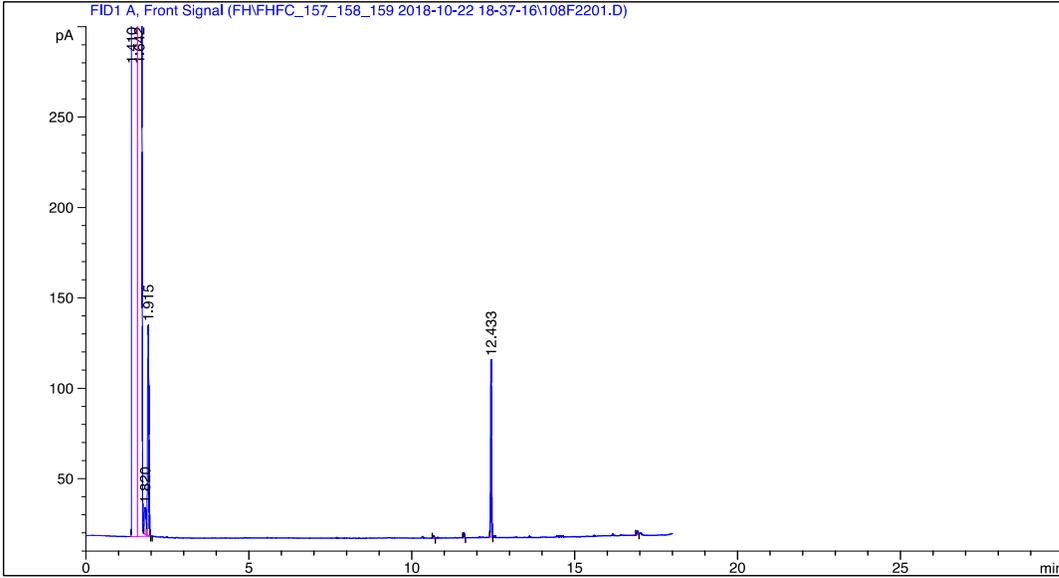
Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.426	BB S	0.0418	3.25527e6	9.79577e5	99.82280
2	1.663	BB T	0.0969	5334.65283	944.15436	0.16359
3	1.940	BB	0.0449	252.60806	95.30789	0.00775
4	11.957	BV	0.0282	4.53192	2.49670	0.00014
5	12.010	VB	0.0299	7.83406	4.17606	0.00024
6	12.258	BB	0.0280	168.08905	93.56811	0.00515
7	12.465	BB	0.0300	3.75390	1.98651	0.00012
8	15.426	BB	0.0289	4.65730	2.60378	0.00014
9	16.430	BB	0.0307	2.48672	1.27589	7.626e-5

Figure A-5 Flame ionization detection response of ADS with FDP by HPCCC using method 2.

Data File C:\CHEM32\1\DATA\FH\FHFC_157_158_159 2018-10-22 18-37-16\108F2201.D
 Sample Name: FHfc_159_run2

```

=====
Acq. Operator   : FH                               Seq. Line : 22
Acq. Instrument : Instrument 1                     Location  : Vial 108
Injection Date  : 10/23/2018 3:50:57 AM          Inj       : 1
                                                    Inj Volume: 5 µl
Acq. Method     : C:\Chem32\1\DATA\FH\FHFC_157_158_159 2018-10-22 18-37-16\FH_
                  CALIBRATIONCURVESHORT.M
Last changed    : 8/15/2018 5:42:45 PM by RJM
Analysis Method : C:\CHEM32\1\DATA\FH\FHFC_157_158_159 2018-10-22 18-37-16\FH_
                  CALIBRATIONCURVESHORT.M
Last changed    : 8/15/2018 5:42:45 PM by RJM
  
```



=====
 Area Percent Report
 =====

```

Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
  
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.410	BV S	0.0452	3.98205e6	1.10011e6	99.80553
2	1.642	VB S	0.1069	7185.88672	1120.45410	0.18011
3	1.820	BV T	0.0657	56.78963	14.82122	0.00142
4	1.915	VB X	0.0473	322.82309	116.48050	0.00809
5	10.671	BB	0.0300	2.53250	1.34253	6.347e-5
6	11.599	BB	0.0284	4.75804	2.58815	0.00012
7	12.433	BB	0.0297	182.13286	97.81767	0.00456
8	16.915	BB	0.0299	4.02694	2.15011	0.00010

Totals : 3.98981e6 1.10147e6

Figure A-6 Flame ionization detection response of GDS with FDP by HPCCC using method 2.

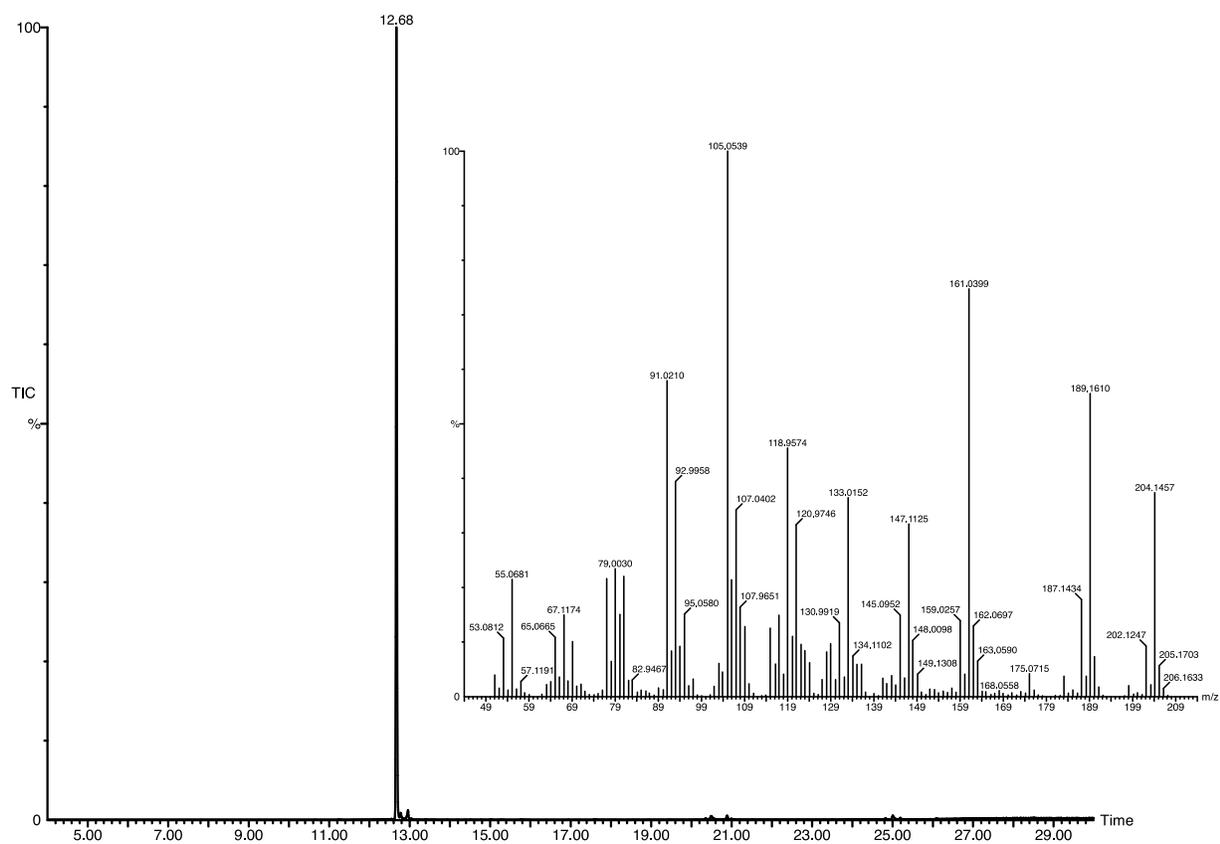


Figure A-8 Total ion chromatogram of authentic sample of FDP with aristolochene synthase by HPLC on the preparative scale method. Inset: mass spectrum of peak at 12.68 min.

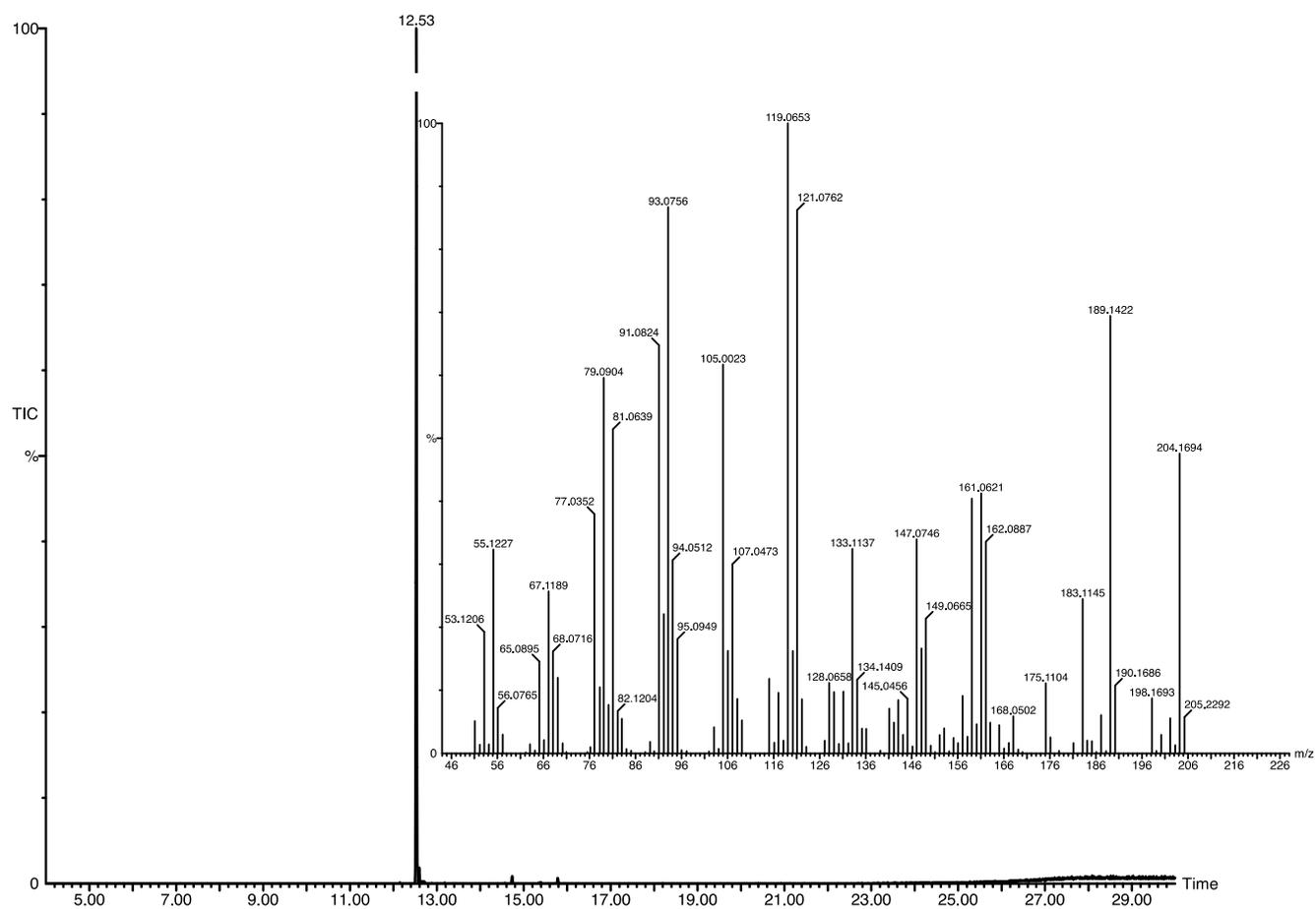


Figure A-9 Total ion chromatogram of authentic sample of FDP with amorphaadiene synthase by HPLC on the preparative scale method. Inset: mass spectrum of peak at 12.53 min.

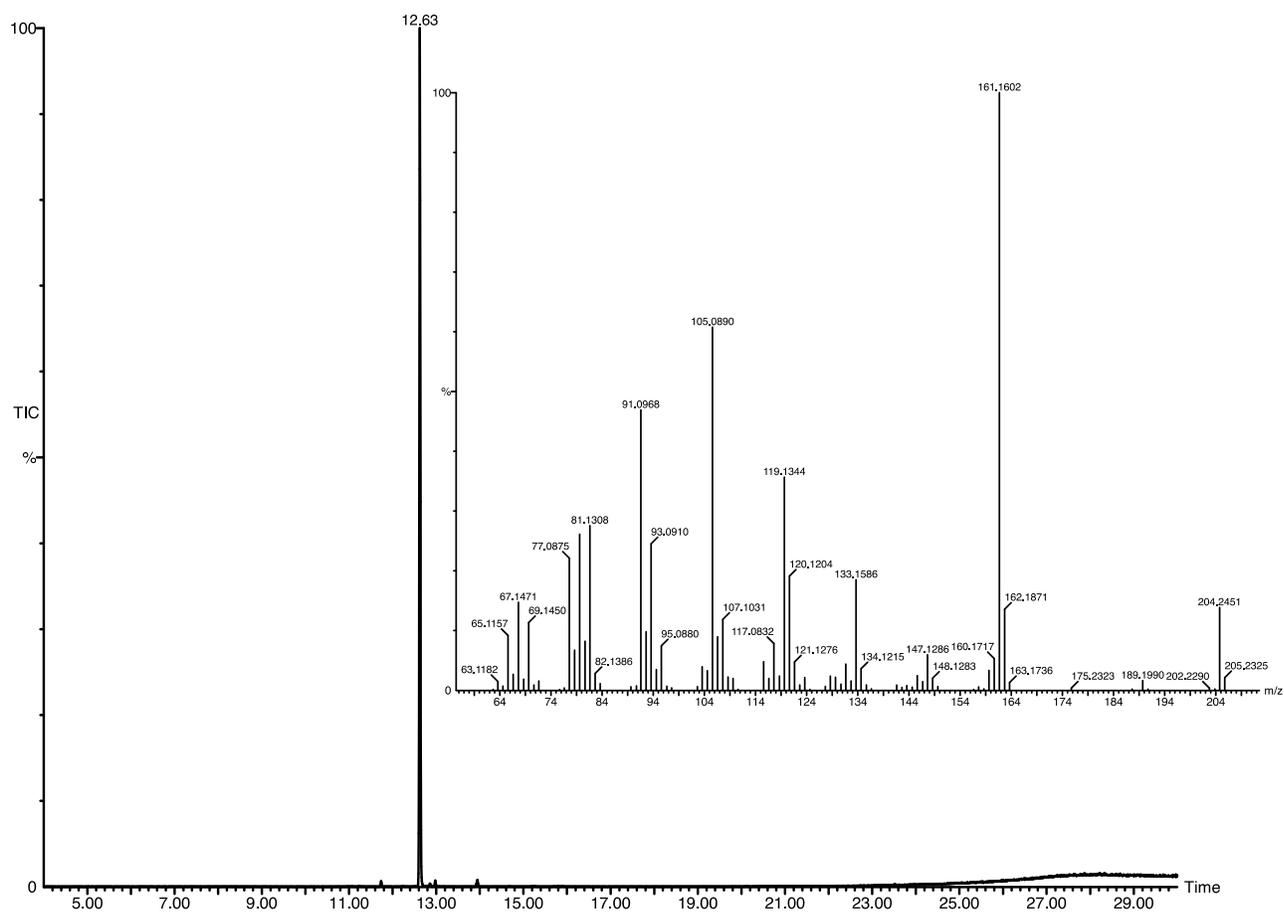


Figure A-10 Total ion chromatogram of authentic sample of FDP with germacrene D synthase by HPLC on the preparative scale method. Inset: mass spectrum of peak at 12.63 min.

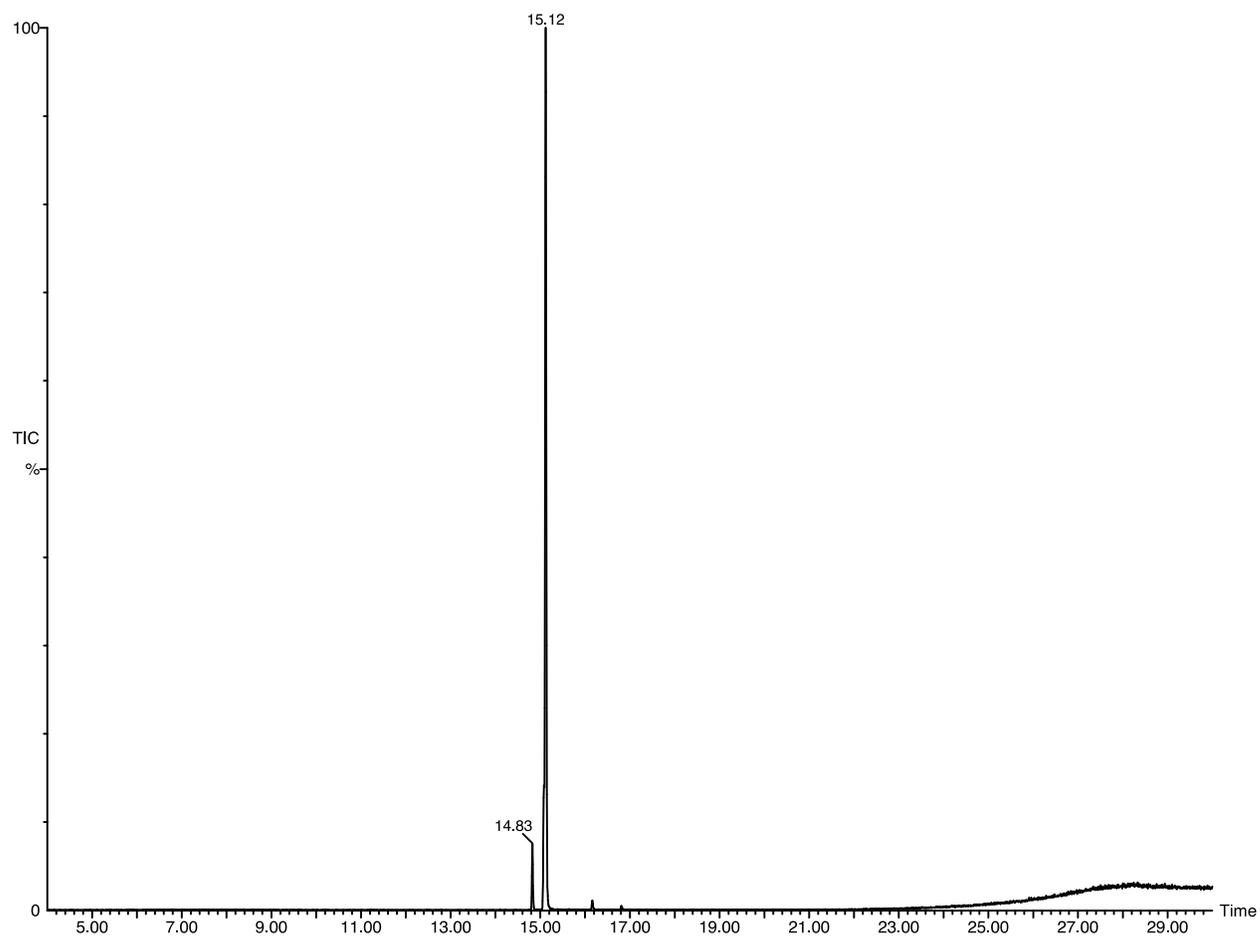


Figure A-11 Total ion chromatogram of authentic sample of 12-OH FDP with amorphaadiene synthase by HPLC on the preparative scale method.

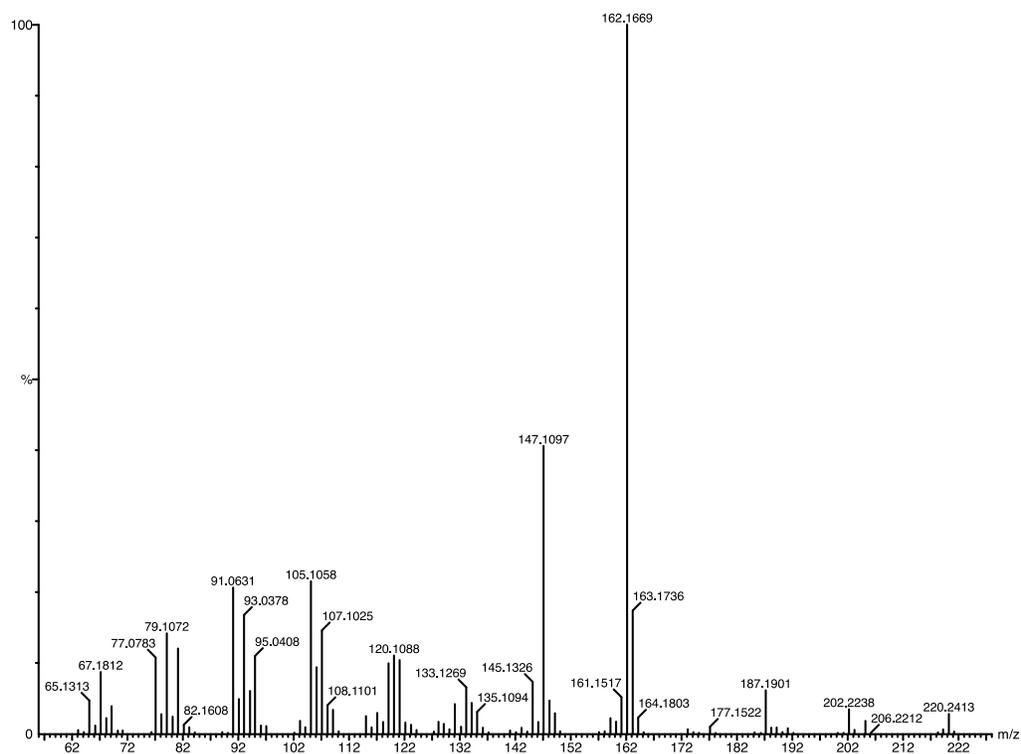


Figure A-12 Mass spectrum of peak at 15.12 min of total ion chromatogram of authentic sample of 12-OH FDP with amorphaadiene synthase by HPLC on the preparative scale method.

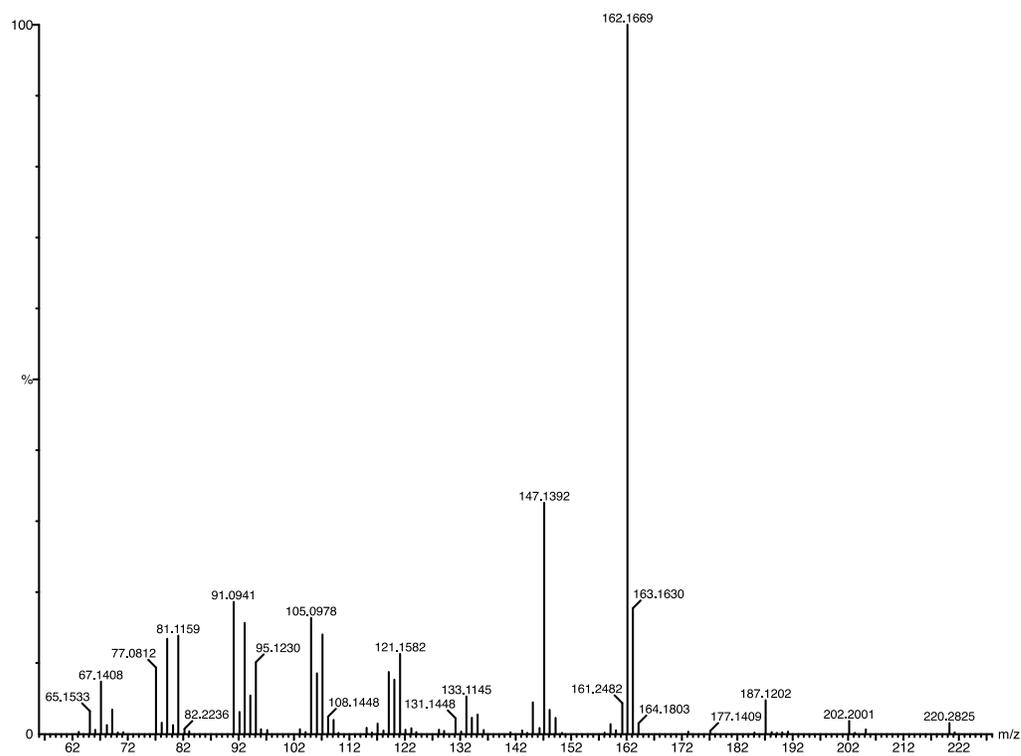
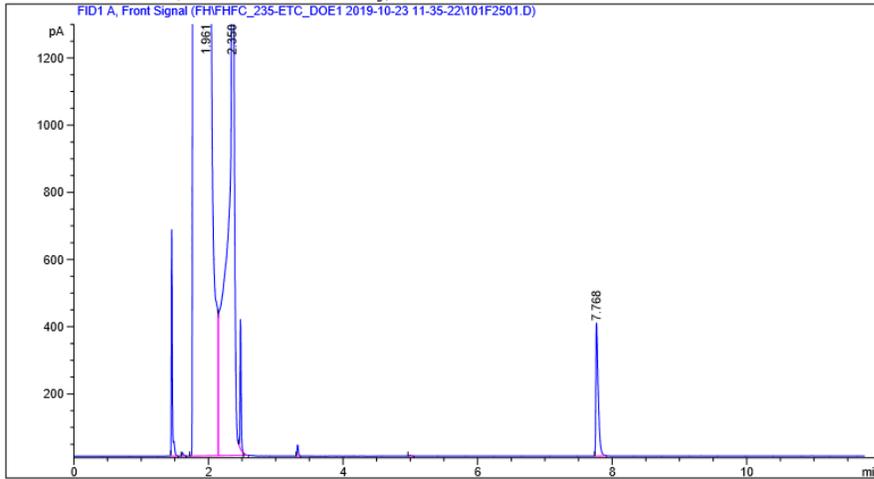


Figure A-13 Mass spectrum of peak at 15.10 min of total ion chromatogram of authentic sample of 12-OH FDP with amorphaadiene synthase by HPLCC on the preparative scale method.

Data File C:\CHEM32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\101F2501.D
 Sample Name: octanol_lmM

```
=====
Acq. Operator   : FH                               Seq. Line : 25
Acq. Instrument : Instrument 1                     Location  : Vial 101
Injection Date  : 10/23/2019 7:43:43 PM          Inj       : 1
                                                    Inj Volume: 5 µl

Acq. Method    : C:\Chem32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\FH_LIPASE1.M
Last changed   : 6/25/2019 10:50:57 AM by FH
Analysis Method: C:\CHEM32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\FH_LIPASE1.M
Last changed   : 10/25/2019 6:21:25 PM by FH
                (modified after loading)
=====
```



```
=====
                          Area Percent Report
=====
```

```
Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.455	BB	0.0151	688.84332	664.42145	0.00864
2	1.613	BB	0.0204	14.31249	10.20785	0.00018
3	1.961	BV S	0.1308	7.95450e6	7.43534e5	99.81198
4	2.350	VB S	0.0910	1.28435e4	2353.13403	0.16116
5	2.477	BB X	0.0171	430.32077	383.12177	0.00540
6	3.326	BB	0.0208	43.67822	32.25002	0.00055
7	5.002	BB	0.0299	4.20138	2.14009	5.272e-5
8	7.768	BB	0.0355	959.33185	392.75421	0.01204

```
Totals :                7.96948e6  7.47372e5
```

Instrument 1 10/25/2019 6:26:24 PM FH

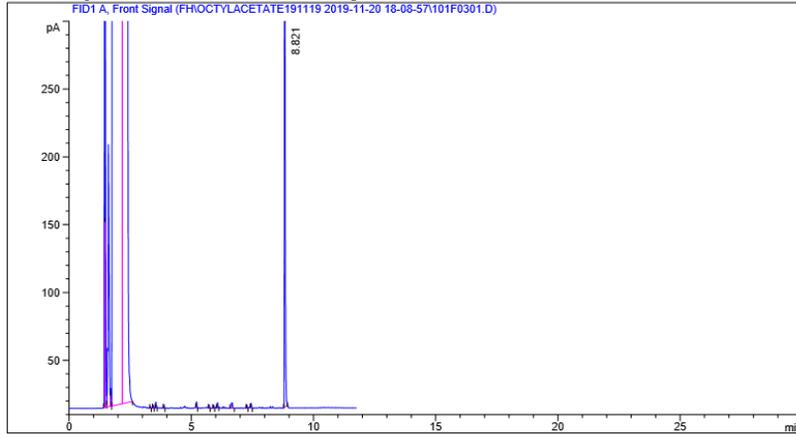
Page 1 of 2

Figure A-14 Flame ionization detection response of octanol.

Data File C:\CHEM32\1\DATA\FH\OCTYLACETATE191119 2019-11-20 18-08-57\101F0301.D
 Sample Name: octylacetate_1mM

```

=====
Acq. Operator   : FH                      Seq. Line :    3
Acq. Instrument : Instrument 1             Location  : Vial 101
Injection Date  : 11/20/2019 6:50:09 PM   Inj       :    1
                                           Inj Volume: 5 µl
Acq. Method     : C:\Chem32\1\DATA\FH\OCTYLACETATE191119 2019-11-20 18-08-57\FH_LIPASE1.M
Last changed    : 11/18/2019 2:18:45 PM by FH
Analysis Method : C:\CHEM32\1\DATA\FH\OCTYLACETATE191119 2019-11-20 18-08-57\FH_LIPASE1.M
Last changed    : 11/18/2019 2:18:45 PM by FH
=====
  
```



Area Percent Report

```

=====
Sorted By      : Signal
Multiplier     : 1.0000
Dilution      : 1.0000
Use Multiplier & Dilution Factor with ISTDs
=====
  
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.453	BV	0.0151	1146.19702	1210.10217	0.01332
2	1.485	VV	0.0238	476.31363	295.10434	0.00553
3	1.610	VV	0.0431	600.34741	193.84511	0.00697
4	1.702	VV	0.0370	29.36436	13.15514	0.00034
5	1.971	VV S	0.1382	8.58335e6	7.57702e5	99.71714
6	2.326	VB S	0.0868	2.10652e4	4044.09570	0.24473
7	3.327	BB	0.0213	1.75299	1.33397	2.037e-5
8	3.442	BB	0.0226	3.03031	2.01072	3.520e-5
9	3.554	BB	0.0276	8.51069	4.39659	9.887e-5
10	3.886	BB	0.0245	4.26820	2.68727	4.959e-5
11	5.220	BB	0.0268	8.23393	4.61553	9.566e-5
12	5.728	BB	0.0262	3.88248	2.36418	4.510e-5
13	5.912	BB	0.0270	3.94220	2.30109	4.580e-5

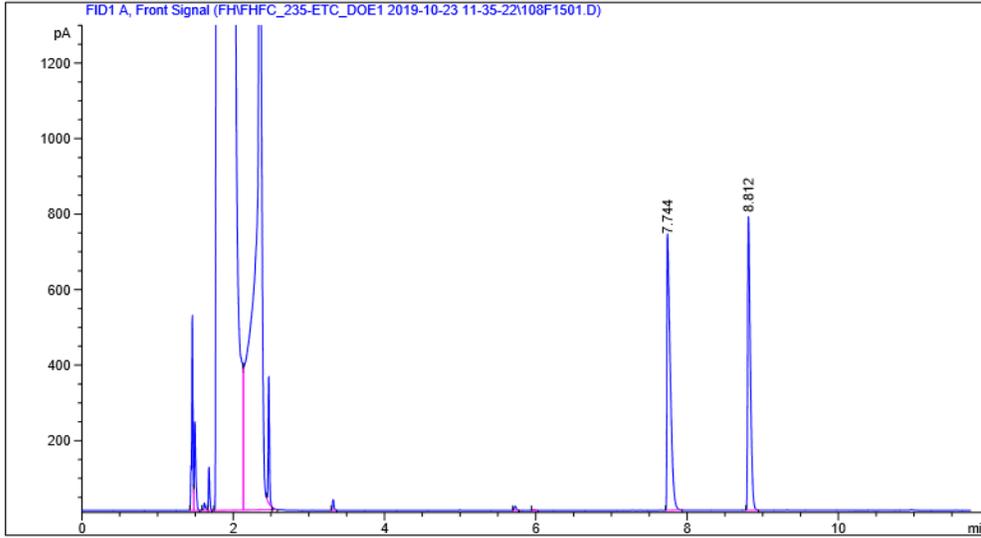
Instrument 1 12/20/2019 2:47:43 PM FH

Page 1 of 2

Figure A-15 Flame ionization detection response of octyl acetate.

Data File C:\CHEM32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\108F1501.D
 Sample Name: FHfc_239_2

```
=====
Acq. Operator   : FH                      Seq. Line : 15
Acq. Instrument : Instrument 1             Location  : Vial 108
Injection Date  : 10/23/2019 4:23:14 FM   Inj       : 1
                                           Inj Volume: 5 µl
Acq. Method     : C:\Chem32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\FH_LIPASE1.M
Last changed    : 6/25/2019 10:50:57 AM by FH
Analysis Method : C:\CHEM32\1\DATA\FH\FHFC_235-ETC_D0E1 2019-10-23 11-35-22\FH_LIPASE1.M
Last changed    : 10/25/2019 6:21:25 PM by FH
                 (modified after loading)
=====
```



=====
 Area Percent Report
 =====

```
Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.461	BV	0.0162	574.86407	509.18927	0.00828
2	1.494	VB	0.0216	328.20819	231.07214	0.00473
3	1.620	BV	0.0228	31.06871	19.28967	0.00045
4	1.681	VB	0.0168	124.66168	113.54231	0.00180
5	1.949	BV S	0.1188	6.92719e6	7.16440e5	99.74579
6	2.346	VB S	0.0932	1.20216e4	2149.60034	0.17310
7	2.471	BB X	0.0183	379.07336	333.51660	0.00546
8	3.322	BB	0.0208	38.13123	28.22597	0.00055
9	5.728	BB	0.0271	20.00497	11.59776	0.00029
10	5.979	BB	0.0276	2.86343	1.62402	4.123e-5
11	7.744	BB	0.0415	2199.84180	723.72308	0.03168
12	8.812	BB	0.0360	1934.38770	778.53345	0.02785

Instrument 1 10/25/2019 6:23:38 PM FH

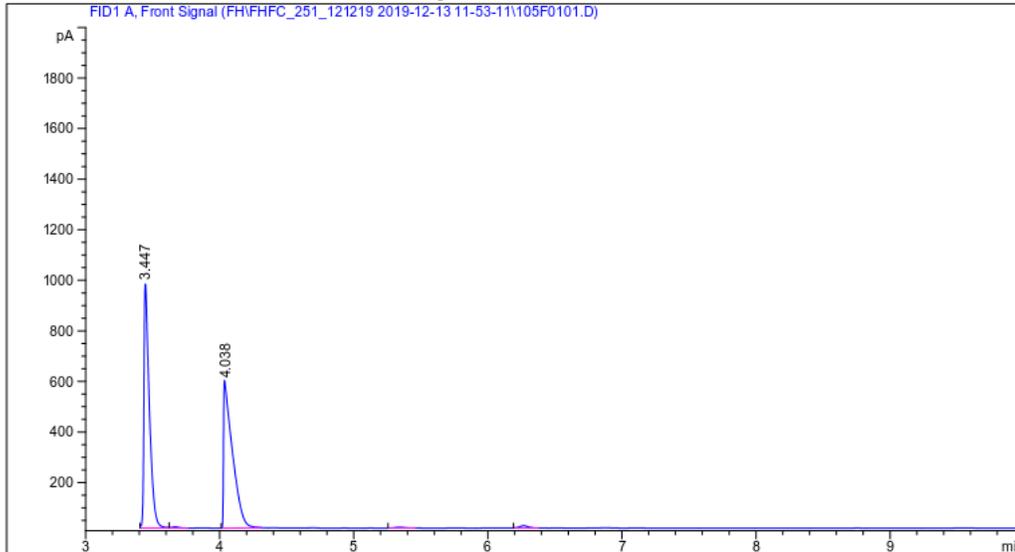
Page 1 of 2

Figure A-16 Example of a flame ionization detection response of the enzymatic transesterification of octanol with vinyl acetate by Cal B lipase (octanol at 7.744 min and octyl acetate at 8.812 min).

Data File C:\CHEM32\1\DATA\FH\FHFC_251_121219_2019-12-13_11-53-11\105F0101.D
 Sample Name: racemate amyliacetate

```

=====
Acq. Operator   : FH                               Seq. Line :    1
Acq. Instrument : Instrument 1                       Location  : Vial 105
Injection Date  : 12/13/2019 11:54:23 AM           Inj       :    1
                                                    Inj Volume: 5 µl
Different Inj Volume from Sequence !   Actual Inj Volume : 1 µl
Acq. Method    : C:\Chem32\1\DATA\FH\FHFC_251_121219_2019-12-13_11-53-11\FH_
PENTANOLISOCRATIC5MIN.M
Last changed   : 12/5/2019 5:11:50 PM by FH
Analysis Method : C:\CHEM32\1\DATA\FH\FHFC_251_121219_2019-12-13_11-53-11\FH_
PENTANOLISOCRATIC5MIN.M
Last changed   : 12/13/2019 12:43:20 PM by FH
                (modified after loading)
  
```



=====
 Area Percent Report
 =====

```

Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
  
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.422	BV	0.0317	147.87105	69.97159	0.00543
2	1.536	VB	0.0634	16.36772	3.57471	0.00060
3	1.821	BB S	0.0617	2.71049e6	5.47576e5	99.53391
4	2.282	BV T	0.0556	7066.15430	1596.40552	0.25948
5	2.444	VB T	0.0314	18.71819	9.93177	0.00069
6	3.447	BB	0.0405	2688.83496	966.20349	0.09874
7	3.668	BB	0.0554	17.04218	4.48305	0.00063
8	4.038	BB	0.0622	2683.58081	577.64923	0.09855
9	5.340	BB	0.0728	15.69835	3.28019	0.00058

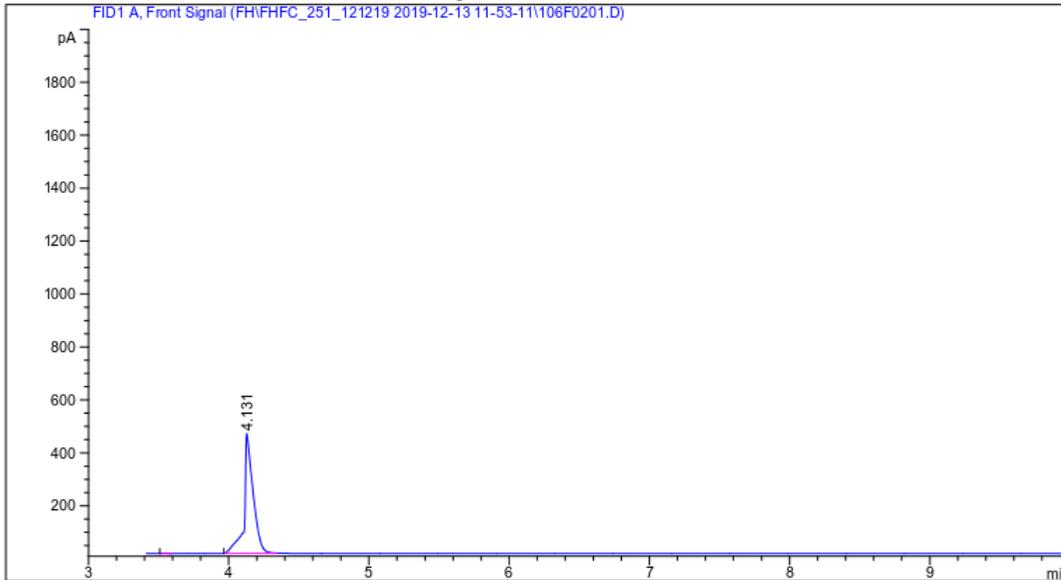
Figure A-17 Flame ionization detection response of chemically synthesised racemic mixture of 2-pentyl acetate.

Data File C:\CHEM32\1\DATA\FH\FHFC_251_121219 2019-12-13 11-53-11\106F0201.D
 Sample Name: FHfc_251p

```

=====
Acq. Operator   : FH                      Seq. Line :    2
Acq. Instrument : Instrument 1             Location  : Vial 106
Injection Date  : 12/13/2019 12:12:38 PM Inj       :    1
                                           Inj Volume: 5 µl

Acq. Method    : C:\Chem32\1\DATA\FH\FHFC_251_121219 2019-12-13 11-53-11\FH_
                PENTANOLISOCRATIC5MIN.M
Last changed   : 12/5/2019 5:11:50 PM by FH
Analysis Method: C:\CHEM32\1\DATA\FH\FHFC_251_121219 2019-12-13 11-53-11\FH_
                PENTANOLISOCRATIC5MIN.M
Last changed   : 12/13/2019 12:44:53 PM by FH
                (modified after loading)
  
```



=====
 Area Percent Report
 =====

```

Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
  
```

Signal 1: FID1 A, Front Signal

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	1.347	BV	0.0252	729.82642	402.14487	0.00430
2	1.496	VV	0.0354	23.60478	9.38223	0.00014
3	1.550	VV	0.0239	2401.91309	1410.28088	0.01416
4	2.005	VV S	0.2394	1.69266e7	8.46227e5	99.80778
5	2.349	VB S	0.1018	2.73719e4	4482.88330	0.16140
6	2.540	BB X	0.0173	54.47515	47.69345	0.00032
7	3.545	BB	0.0301	3.25974	1.64469	1.922e-5
8	4.131	BB	0.0595	2013.90759	447.74252	0.01188

Totals : 1.69592e7 8.53029e5

Instrument 1 12/13/2019 12:45:12 PM FH

Page 1 of 2

Figure A-18 Flame ionization detection response of the enzymatic transesterification of rac-2-pentanol with vinyl acetate by Cal B lipase [(R)-2-pentylacetate at 4.131 min].

II. NMR SPECTROSCOPIC ANALYSIS OF CYCLIC ETHER

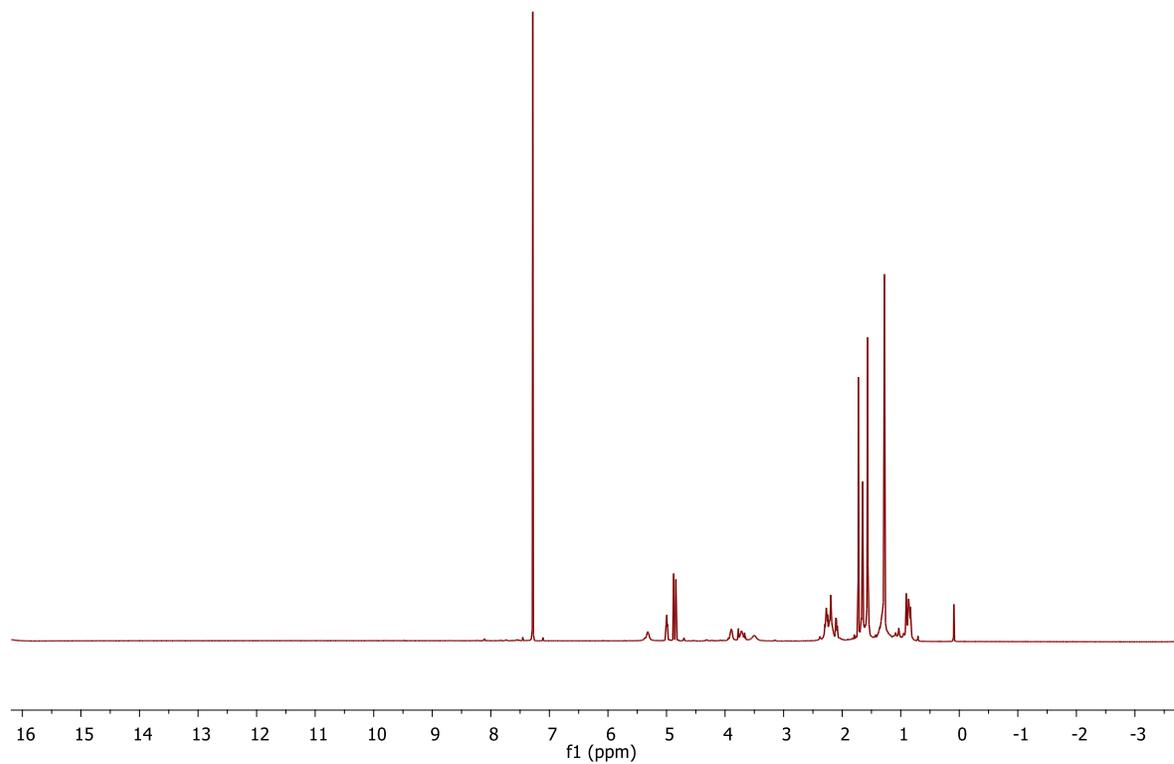


Figure A-19 Full ¹H NMR spectrum (400 MHz, CDCl₃, 298 K) of macrocyclic ether.

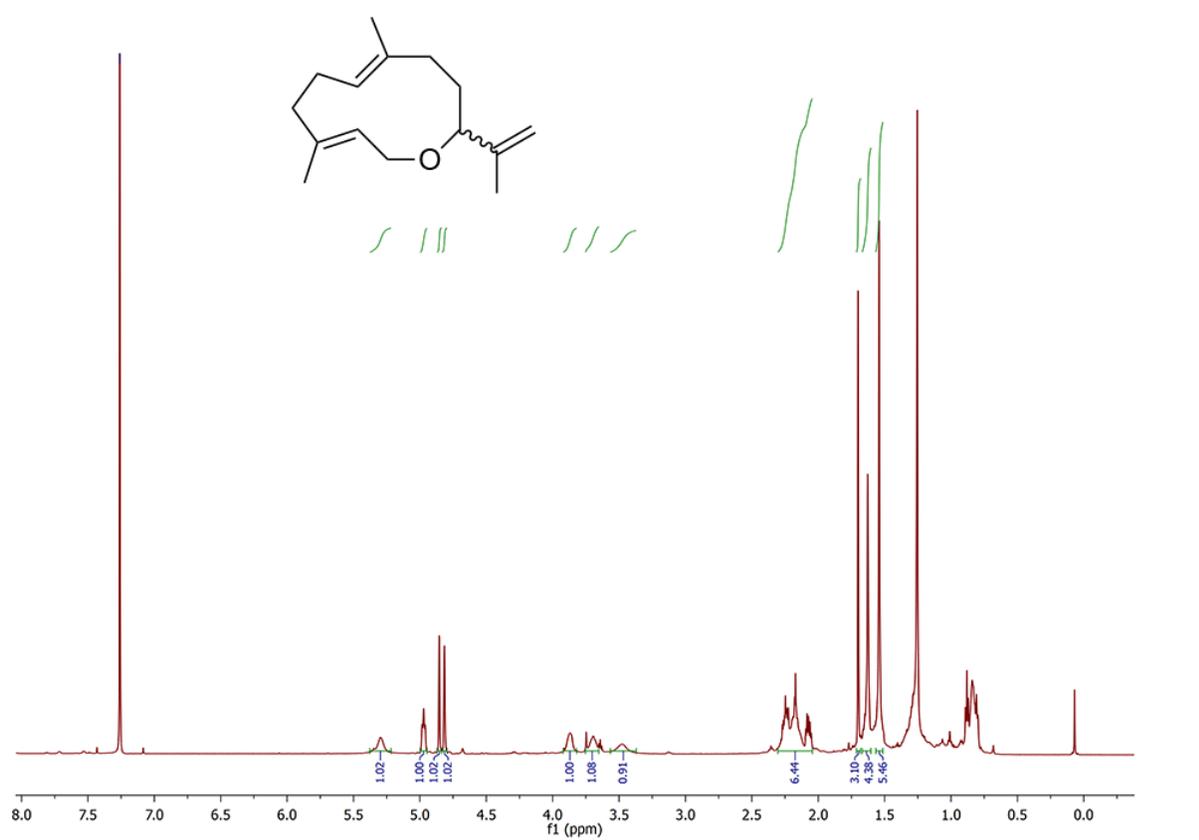


Figure A-20 ^1H NMR spectrum (400 MHz, CDCl_3 , 298 K) of macrocyclic ether.

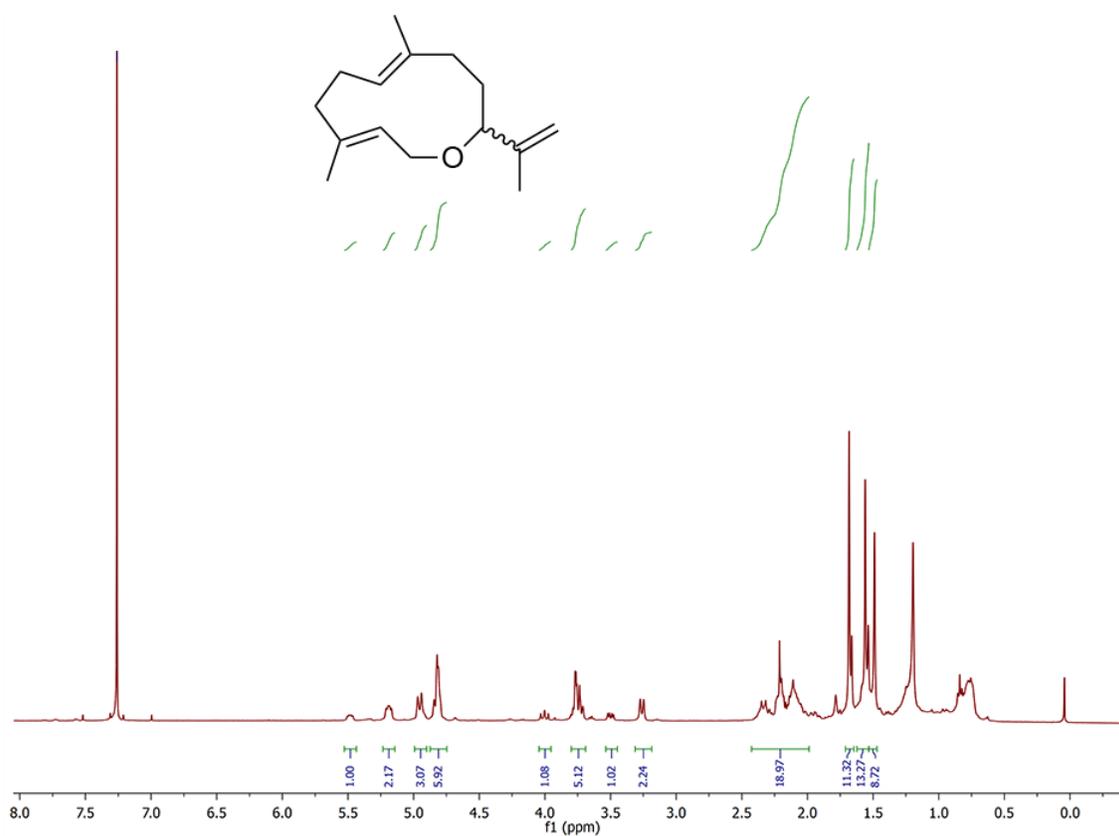


Figure A-21 ^1H NMR spectrum (400 MHz, CDCl_3 , 223 K) of macrocyclic ether.

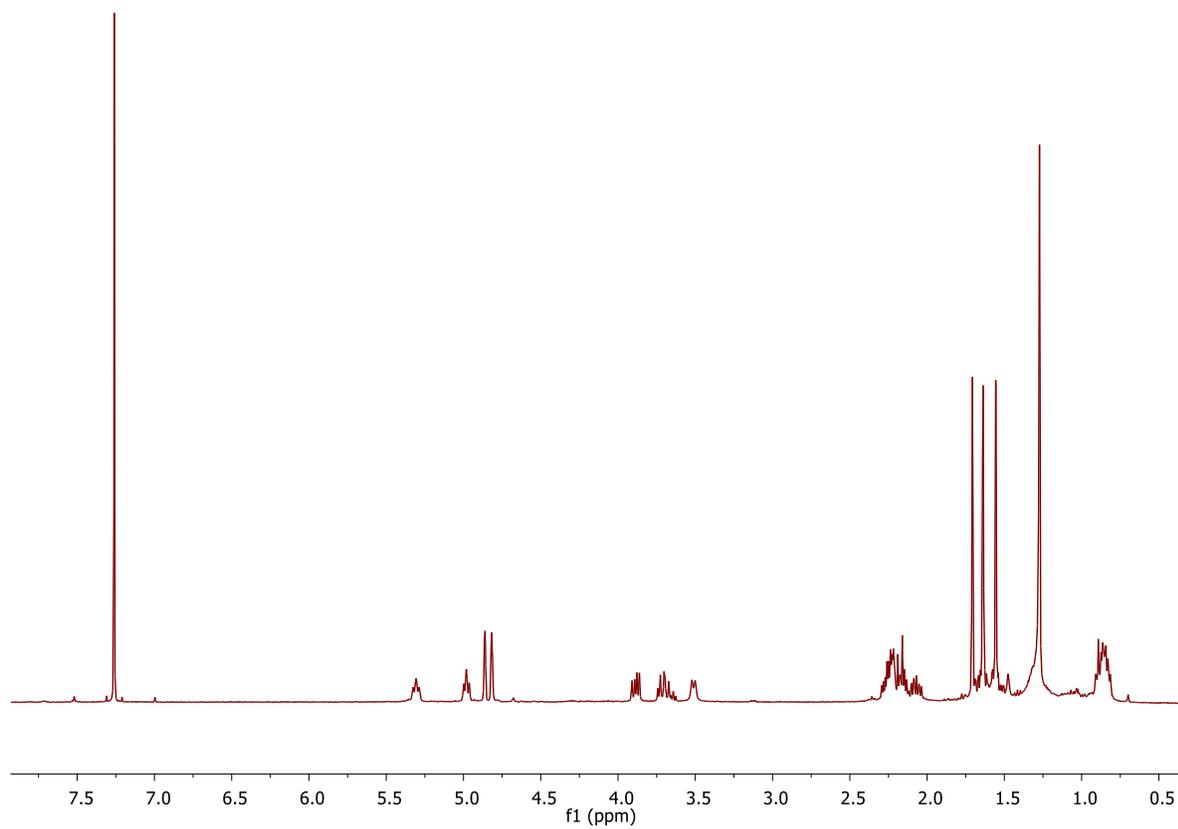


Figure A-22 ^1H NMR spectrum (400 MHz, CDCl_3 , 323 K) of macrocyclic ether.

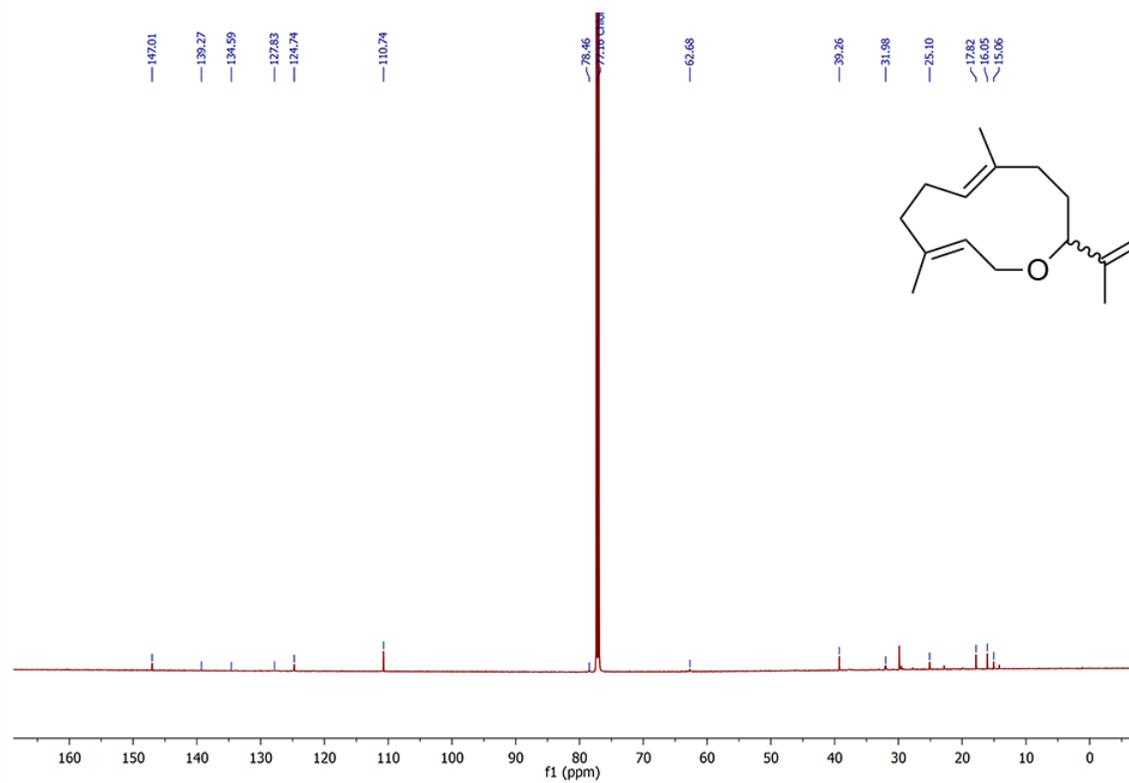


Figure A-23 ^{13}C NMR spectrum (100 MHz, CDCl_3 , 323 K) of macrocyclic ether.

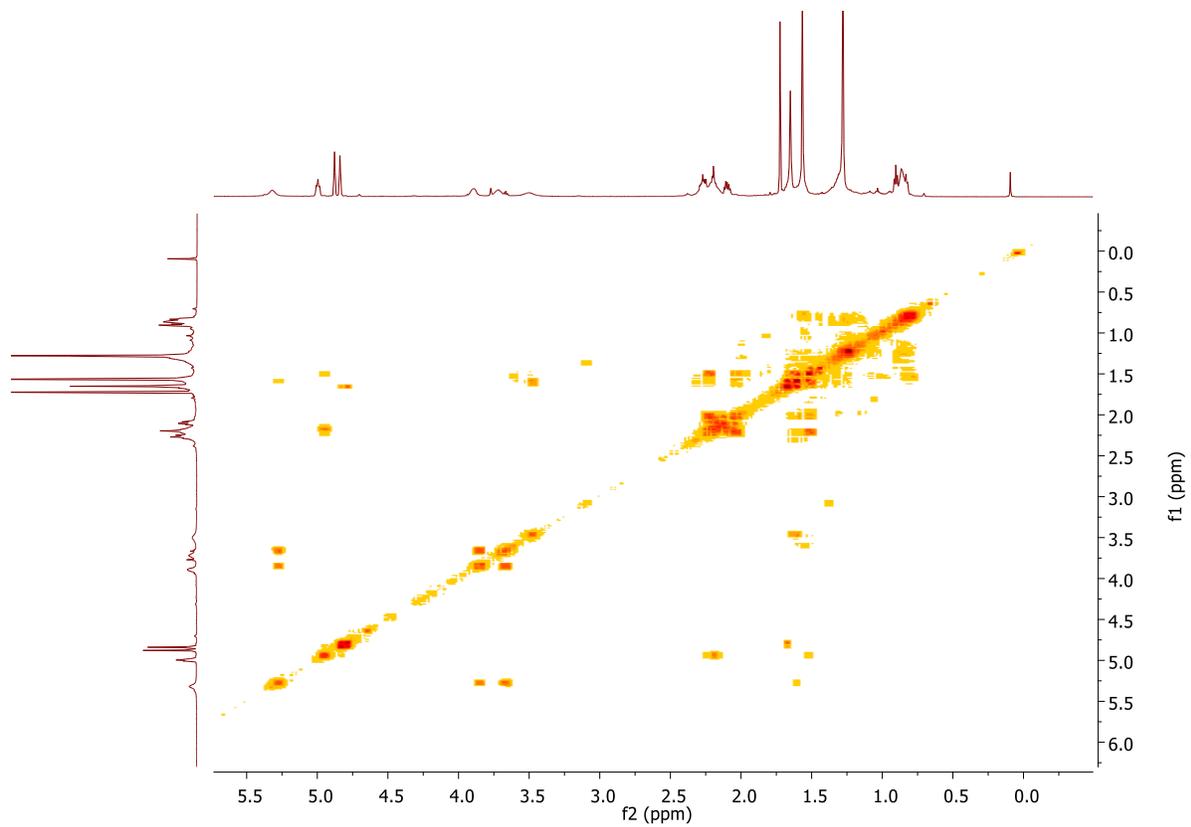


Figure A-24 COSY NMR spectrum (400 MHz, CDCl₃, 323K) of macrocyclic ether.

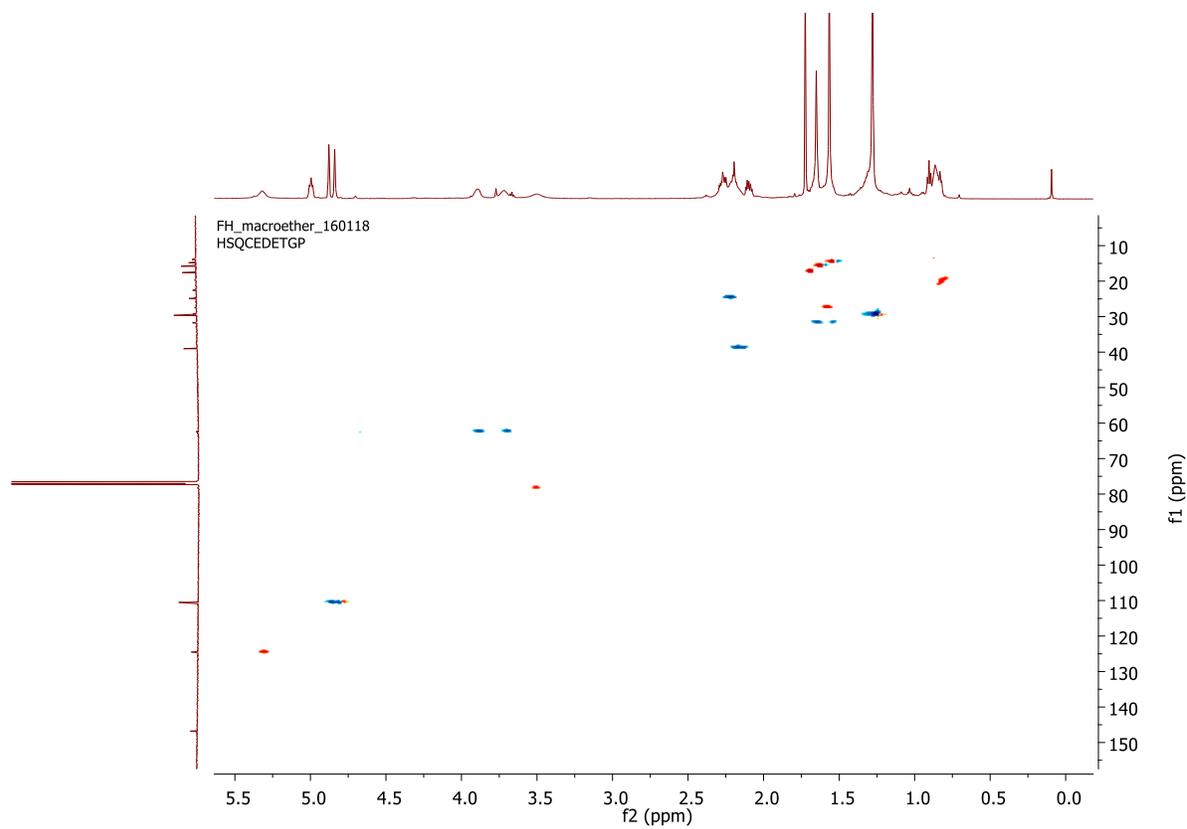


Figure A-25 Edited HSQC NMR spectrum (400 MHz, CDCl_3 , 323 K) of macrocyclic ether.

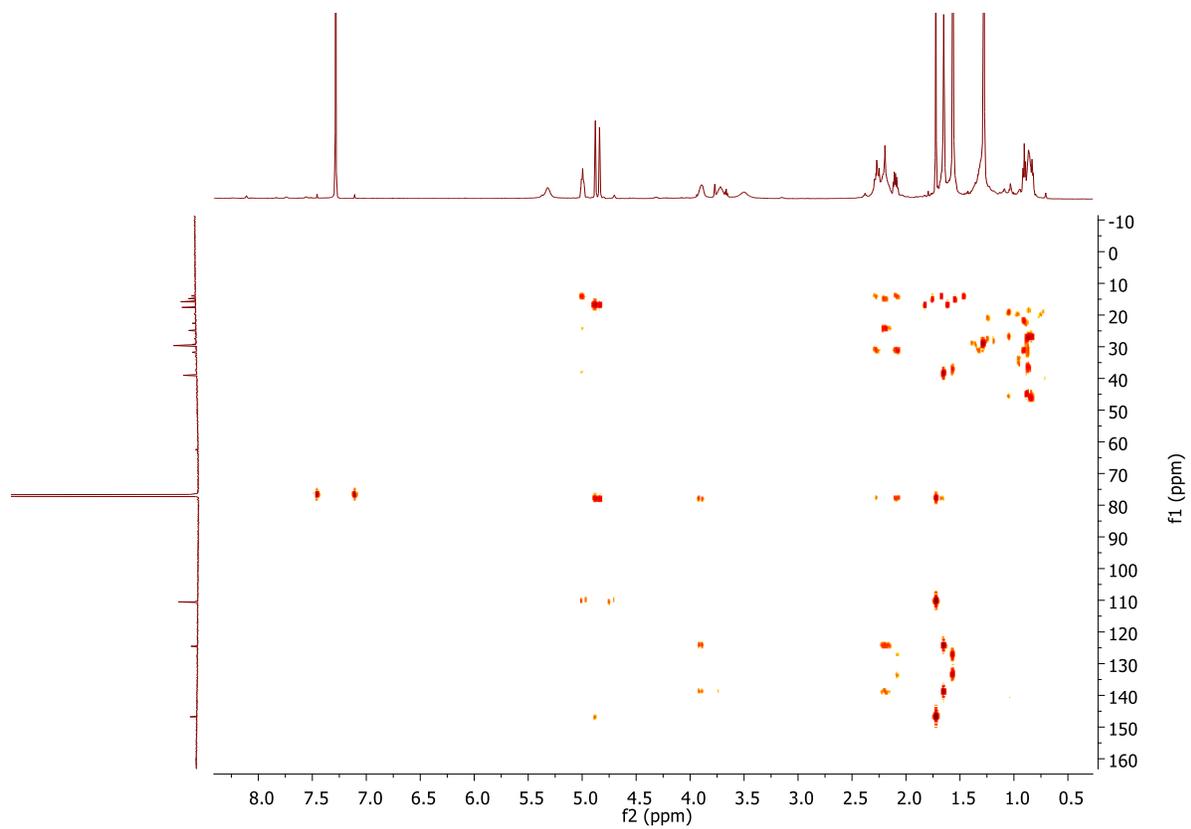


Figure A-26 HMBC NMR spectrum (400 MHz, CDCl₃, 323 K) of macrocyclic ether.

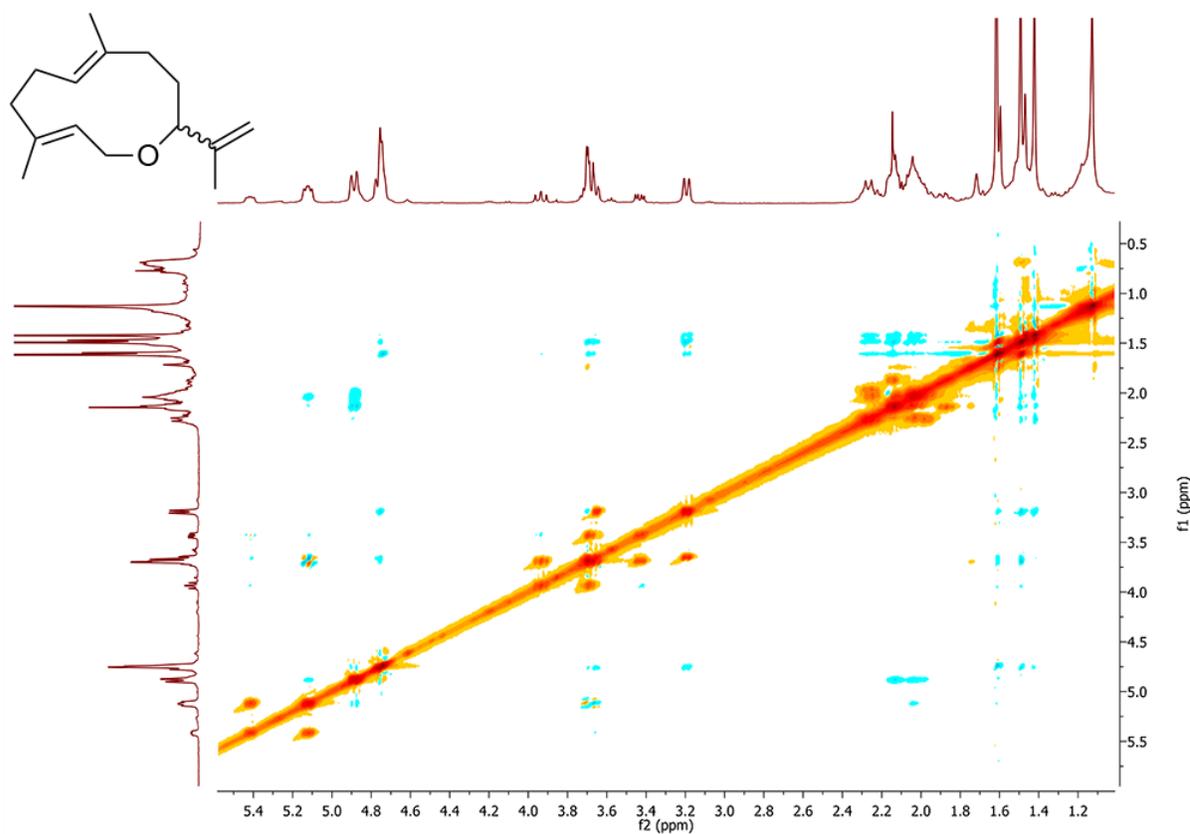


Figure A-27 NOESY NMR spectrum (400 MHz, CDCl_3 , 223 K) of macrocyclic ether.

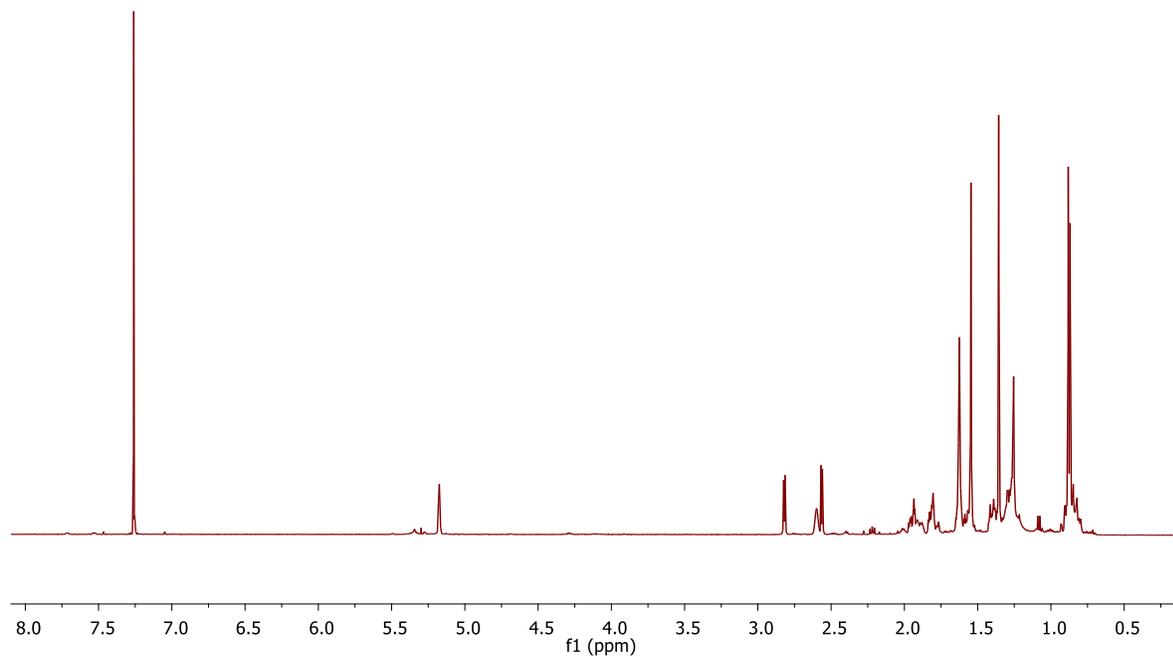
III. NMR SPECTROSCOPIC ANALYSIS OF ARTEMISINIC-11S,12-EPOXIDE

Figure A-28 ¹H NMR spectrum (500 MHz, CDCl₃, 298 K) of artemisinic-11S,12-epoxide.

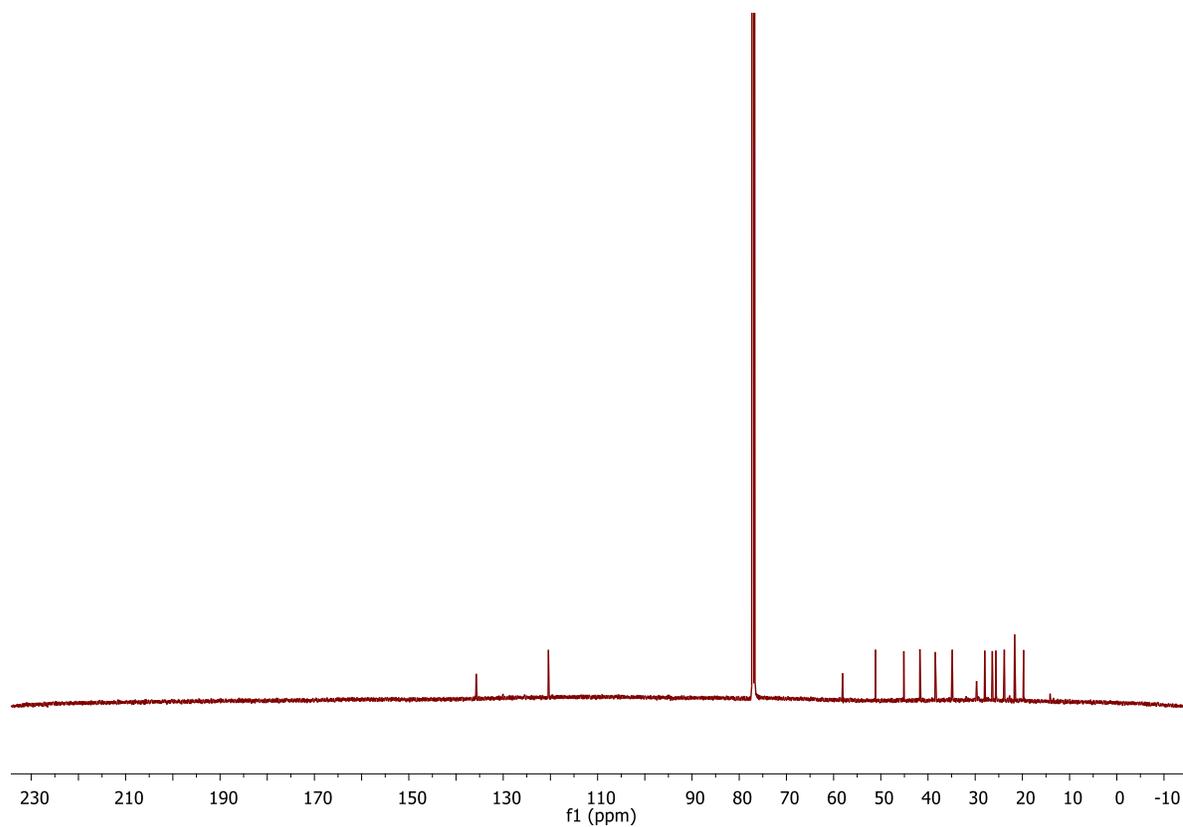


Figure A-29 ^{13}C NMR spectrum (125 MHz, CDCl_3 , 298 K) of artemisinic-11S,12-epoxide.

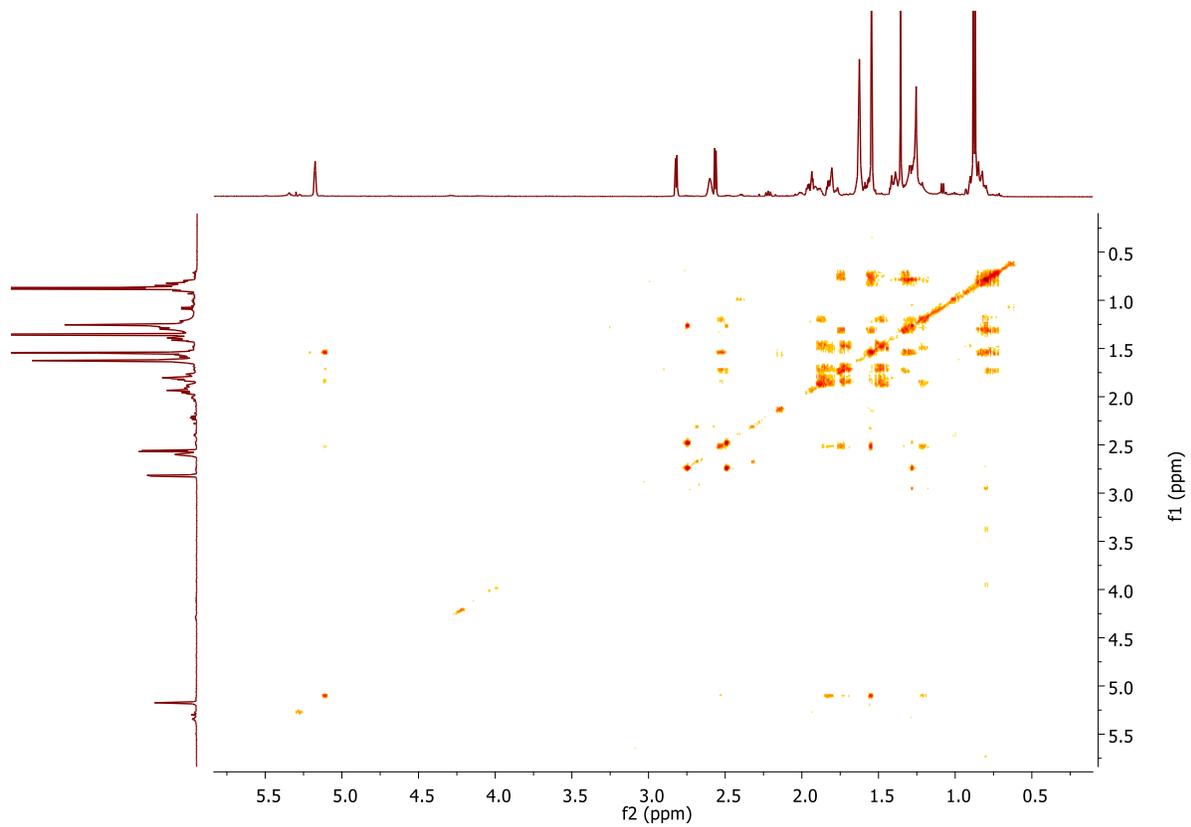


Figure A-30 COSY NMR spectrum (500 MHz, CDCl₃, 298K) of artemisinic-11S,12-epoxide.

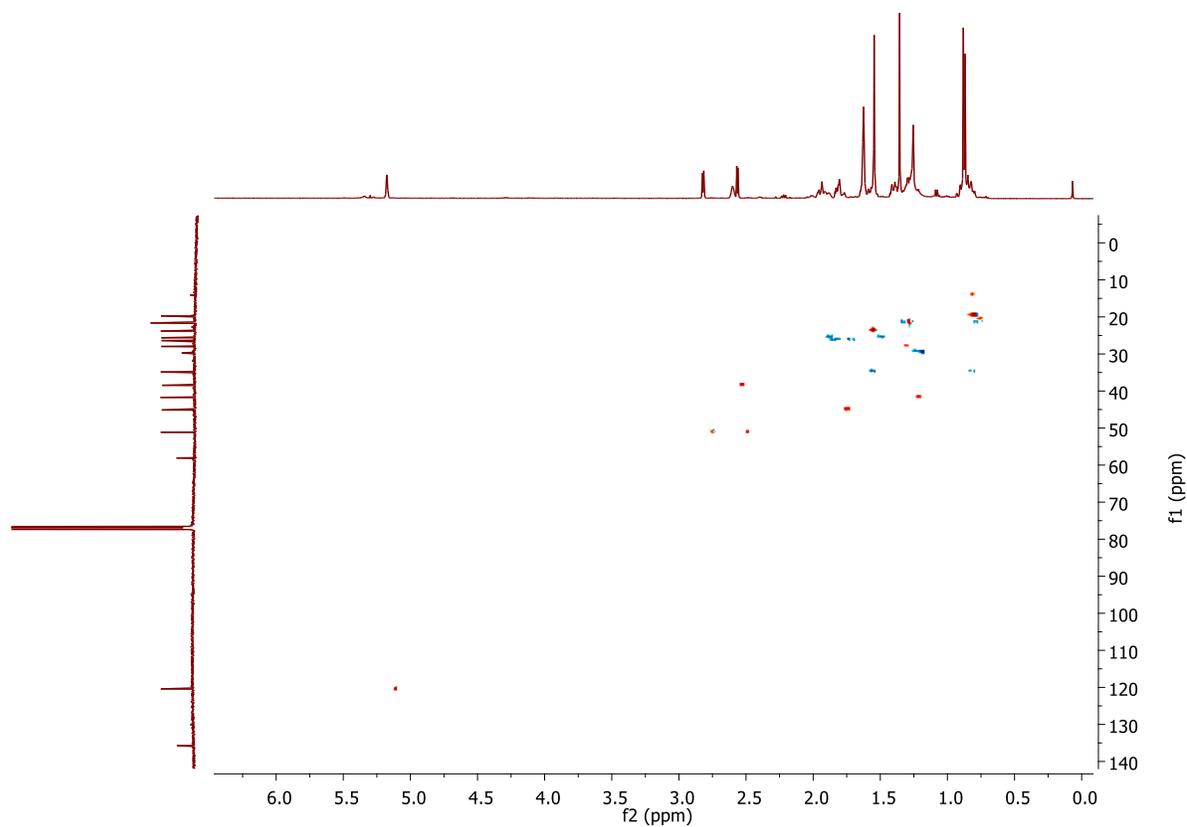


Figure A-31 Edited HSQC NMR spectrum (500 MHz, CDCl_3 , 298 K) of artemisinin-11S,12-epoxide.

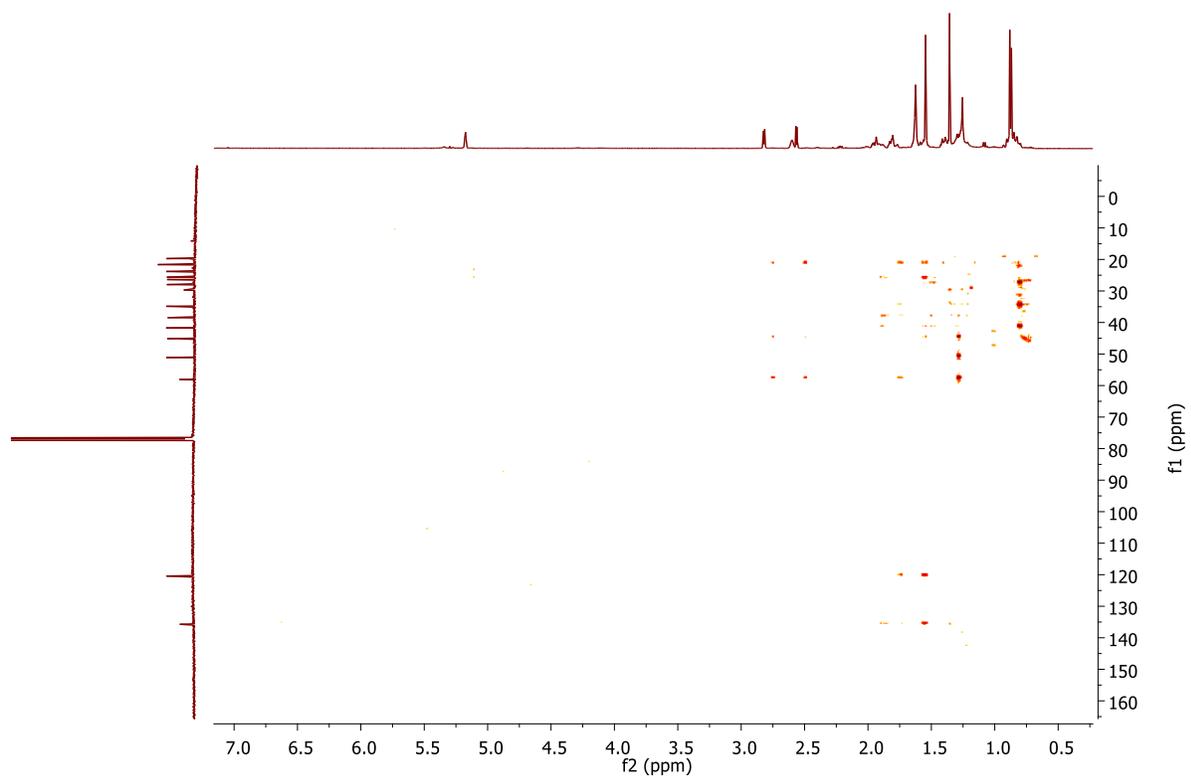


Figure A-32 HMBC NMR spectrum (500 MHz, CDCl₃, 298 K) of artemisinin-11S,12-epoxide.

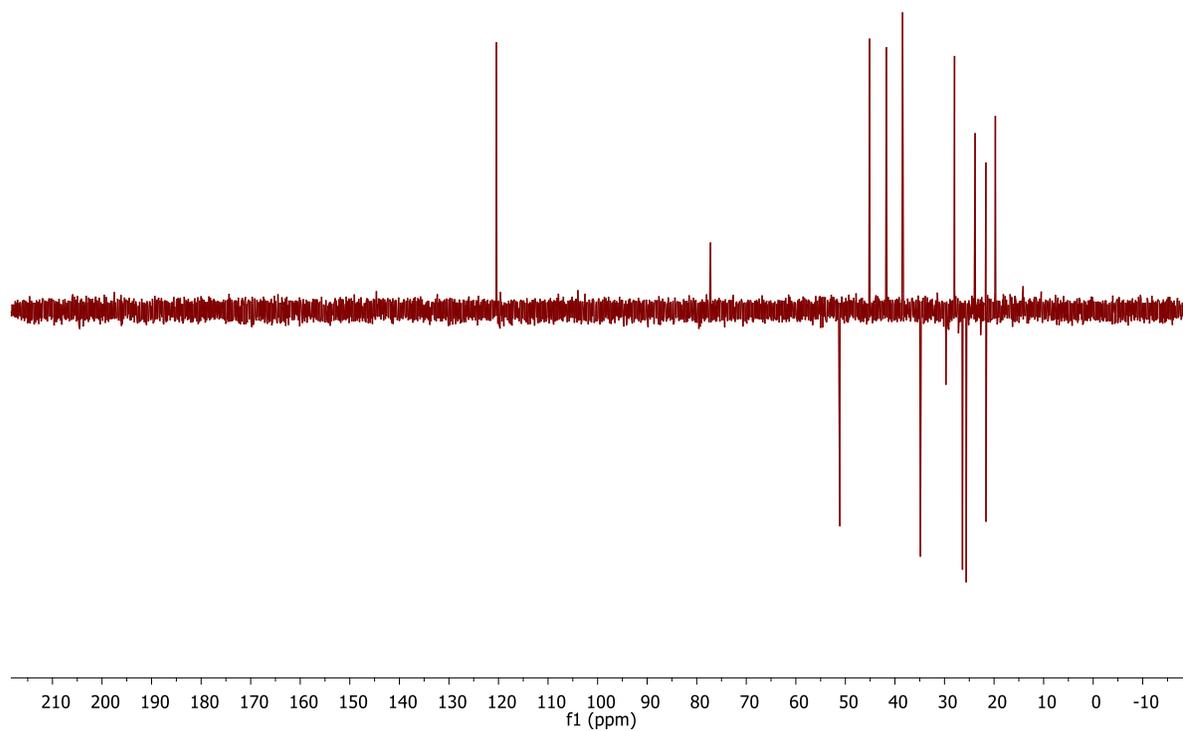


Figure A-33 DEPT 135 NMR spectrum (500 MHz, CDCl_3 , 298 K) of artemisinin-11S,12-epoxide.