

**STRUCTURE ELUCIDATION AND ISOTOPIC LABELING  
EXPERIMENTS PROVIDE INSIGHTS INTO THE BIOSYNTHESIS  
OF THREE PHYTOCHEMICALS: ALKAMIDES, AVENACIN, AND  
HYDROXYLATED FATTY ACIDS**

by

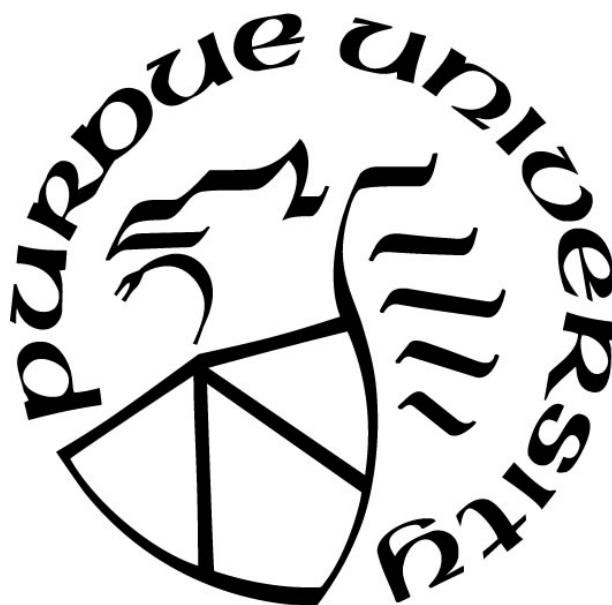
**Alicen M. Teitgen**

**A Thesis**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Doctor of Philosophy**



Department of Chemistry and Chemical Biology

Indianapolis, Indiana

December 2017

**THE PURDUE UNIVERSITY GRADUATE SCHOOL**  
**STATEMENT OF COMMITTEE APPROVAL**

Dr. Robert E. Minto, Chair

Department of Chemistry and Chemical Biology

Dr. Eric Long

Department of Chemistry and Chemical Biology

Dr. Sébastien Laulhé

Department of Chemistry and Chemical Biology

Dr. John Goodpaster

Department of Chemistry and Chemical Biology

**Approved by:**

Dr. Eric Long

Head of the Graduate Program

*To Jake – Thank you for always believing in me.*

## ACKNOWLEDGMENTS

I am forever grateful to so many people who helped me along my journey. First, I want to thank Dr. Robert E. Minto for mentoring me over the past 5 years. Dr. Minto's knowledge still intimidates me, but his wisdom and support has challenged me to become a better chemist, writer, and learner. I also want to thank Dr. Brenda Blacklock for her guidance and encouragement along the way. I've enjoyed sharing my love for running with both Dr. Minto and Dr. Blacklock, and they have been extremely compassionate towards my family and me. Thank you to Dr. Michael R. Shepard for helping me to get started in lab and for your friendship. I want to thank all the undergraduates I have had the privilege to work with over the years: Paola, Ngun, Jermell, and Bailey. A special thanks to Bailey, whose fun-loving spirit and contagious attitude helped keep my spirits up during the last year of graduate school. I also owe a special thanks to all our collaborators who made our projects possible. It truly takes a village to raise a family, especially when you start a family during graduate school. For that, I owe enormous thanks to all the family and friends who were there while I pursued my dreams. To my mom and step-dad, Nancy, and Tim, and to my parents-in-law, Barb and Tom, thank you for all the support especially in helping to take care of Jules and Eve. Thank you to my dad, Mike, for always being my #1 fan and for Sunday pizza nights. I treasure that Jules and Eve have such a special relationship with their grandparents. Thank you to the best bro ever, Kyle, for your witty personality and for always making me smile. Thank you to our community of friends, especially those from Our Lady of Grace, for unrelenting support and encouragement. Each of you has played an important role in our lives. Finally, thank you to Jules and Eve for your endless love. There is nothing better than coming home to your hugs, kisses, and the excitement you have when I walk through the door. When school was the most difficult, you were my biggest inspiration to keep working hard. Though you might not remember this time in your life, I hope you'll always remember that no dream is too big. To Jake, thank you for never giving up on me and for always believing in me. It wasn't easy and, without you, I'm sure I wouldn't have made it. Thank you for being my best friend and for helping me to achieve my dreams.



## TABLE OF CONTENTS

LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
LIST OF ABBREVIATIONS.....	xviii
ABSTRACT.....	xx
CHAPTER 1. INTRODUCTION .....	21
1.1 Evolution of Plant Natural Product Biosynthesis.....	21
1.2 Research Objectives .....	26
CHAPTER 2. IDENTIFICATION OF HYDROXYLATED FATTY ACIDS FOUND IN <i>ORYCHOPHRAGMUS VIOLACEUS</i> .....	27
2.1 Introduction .....	27
2.1.1 Lipids from specialized metabolism .....	27
2.1.2 Specialized fatty acids in the seed oils of the Brassicaceae .....	35
2.2 Materials and Methods .....	36
2.2.1 GC/MS protocols for the structure elucidation of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> , 21      37	37
2.2.1.1 FAME and pyrrolidide GC/MS methods .....	37
2.2.1.2 Metathesis and racemic 3-cyclohexen-1-ol GC/MS methods.....	38
2.2.1.3 GC/MS method for chiral separations.....	40
2.2.2 Isolation of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> from seed oil.....	41
2.2.3 Derivatives of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	41
2.2.4 Configuration of C-18 in 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	42
2.2.5 Route 1 for the attempted synthesis of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	44
2.2.6 Route 2 for the synthesis of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	47
2.2.7 Configuration of C-7 in 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	53
2.3 Results and Discussion.....	55
2.3.1 Serendipitous discovery of novel oils in <i>O. violaceus</i> seeds.....	55
2.3.2 Structural identification of novel fatty acids from <i>O. violaceus</i> seeds.....	56
2.3.3 Configuration at the C18 stereocenter of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	68
2.3.4 Configuration at the C7 stereocenter of 7,18-(OH) <sub>2</sub> -24:2Δ <sup>15,21</sup> .....	70

2.3.5	Oil composition .....	78
2.4	Conclusion.....	82
CHAPTER 3. THE CONVERGENCE OF TWO PATHWAYS IN THE		
BIOSYNTHESIS OF ALKAMIDES FROM <i>ECHINACEA</i> .....		
3.1	Introduction .....	83
3.1.1	Convergence of two metabolic pathways.....	83
3.1.2	Alkamide biosynthesis in <i>Echinacea</i> .....	87
3.2	Materials and Methods .....	96
3.2.1	Germination and growth of <i>E. purpurea</i> seedlings .....	98
3.2.2	Alkamide analysis .....	98
3.2.3	Isotopic labeling experiments.....	99
3.2.4	Equations used to calculate incorporation levels .....	100
3.2.5	Fatty acid methyl ester preparation for MS analysis.....	100
3.2.5.1	Protocols for alkamide and FAME analysis by GC/MS and LC/MS .....	101
3.2.6	Synthesis of aldehydes and [d <sub>3</sub> ]-aldehydes .....	104
3.2.7	Synthesis of alkamides and [d <sub>3</sub> ]-alkamides.....	117
3.2.8	Synthesis of fatty acids and [d <sub>3</sub> ]-fatty acids .....	121
3.2.9	Synthesis of ( <i>E,E</i> )-2,4-dodecadienoic acid .....	127
3.2.10	Synthesis of ( <i>2E,4E,8Z,10E</i> )-dodecadienoic acid.....	130
3.2.11	Synthesizing acyl-CoAs .....	133
3.2.12	Calculating concentrations of acyl-CoAs.....	133
3.3	Results and Discussion.....	135
3.3.1	Isotope labeling experiments of amino acids, Val and Ile .....	136
3.3.2	Synthesis of deuterated aldehydes.....	141
3.3.3	Synthesis of $\alpha,\beta$ -unsaturated alkamides .....	143
3.3.4	Synthesis of $\alpha,\beta$ -unsaturated fatty acids.....	145
3.3.5	Determination of isotope distribution patterns in alkamides isolated from plants provided deuterated $\alpha,\beta$ -unsaturated fatty acids and $\alpha,\beta$ -unsaturated alkamides .....	146
3.3.6	Synthesis of 12:1, 12:2, and 12:4 used for acyl-CoAs .....	147
3.3.7	Concentrations of acyl-CoAs .....	151

3.3.8	Activity assays of acyl-CoAs with EpLigase .....	152	
3.4	Conclusion.....	154	
CHAPTER 4. STRUCTURE ELUCIDATION OF PROPOSED			
INTERMEDIATES IN THE SAD2 PATHWAY.....			156
4.1	Introduction.....	156	
4.1.1	Evolution of enzymes in secondary metabolism.....	157	
4.1.2	Triterpenes.....	160	
4.1.3	Biosynthesis of the triterpene avenacin.....	163	
4.2	Method and Materials.....	174	
4.2.1	GC/MS and Q-TOF protocols for triterpene analysis .....	175	
4.2.2	Triterpene isolation from SAD2 mutant expression in <i>N. benthamiana</i> ....	177	
4.3	Results and Discussion.....	181	
4.3.1	Mutants.....	181	
4.3.2	Structure elucidation of the products from SAD2 mutants.....	186	
4.3.3	Structure elucidation of the product from wild-type SAD2 .....	186	
4.3.4	Structure elucidation from mutant A354L .....	191	
4.3.5	Structure elucidation from mutant I471M.....	195	
4.3.6	Structure elucidation of product 1 from mutant I471M .....	195	
4.3.7	Structure elucidation of product 2 from I471M .....	200	
4.3.8	Structure elucidation of product 3 from I471M .....	203	
4.3.9	Deciphering the products produced by mutant I471M.....	207	
4.3.10	Analysis of mass spectra of SAD2 mutants .....	211	
4.3.11	Effects of NMR solvent on <sup>13</sup> C chemical shifts .....	215	
4.3.12	Comparison of <sup>13</sup> C shifts of oleanene triterpene in pyridine-d <sub>5</sub> .....	216	
4.3.13	Correlation of methyl groups .....	222	
4.3.14	Future Work .....	225	
REFERENCES .....			226
APPENDIX.....			242
Assignment of β-amyrin in pyridine d <sub>5</sub> .....			242

## LIST OF TABLES

Table 1: NMR data for methyl wuhanate <b>2a</b> .....	59
Table 2: Mosher ester analysis of <b>10d</b> and <b>10e</b> .....	75
Table 3: <sup>1</sup> H chemical shifts from Mosher analysis of <b>27a</b> and <b>27b</b> .....	77
Table 4: Calculated concentrations of acyl-CoAs (12:1, 12:2, and 12:4) determined by UV-Vis and GC-MS. ....	151
Table 5: Enzyme activity of EpLigase with unsaturated CoAs. ....	153
Table 6: The fraction of terpenoids that are oxygenated per class. ....	160
Table 7: The 28 mutants that were analyzed. ....	184
Table 8: Tabulated NMR data for 12,13β-epoxy-16β-hydroxy-β-amyrin. Data was collected in pyridine-d <sub>5</sub> . ....	190
Table 9: Tabulated NMR data for 16β-hydroxy-β-amyrin. Data was collected in pyridine-d <sub>5</sub> . ....	194
Table 10: Tabulated NMR data for 15β-hydroxy-13β-β-amyrin. Data was collected in pyridine-d <sub>5</sub> . ....	199
Table 11: Tabulated NMR data for 12-keto-13β-β-amyrin. Data was collected in pyridine-d <sub>5</sub> . ....	202
Table 12: Tabulated NMR data for 12-keto-13β-β-amyrin. Data was collected in pyridine-d <sub>5</sub> . *C-7 and C-14 could not be distinguished from one another. ....	206
Table 13: Comparing the chemical shifts from pyridine-d <sub>5</sub> and CDCl <sub>3</sub> . ....	216
Table 14: Comparative table of chemical shifts of β-amyrin derivatives relative to β-amyrin. All data collected in pyridine-d <sub>5</sub> . Downfield shifts greater than 10 ppm are denoted in dark red and downfield shifts of 0.1-5 are printed in light red. Upfield shifts greater than 10 ppm are shown in dark green and upfield shifts of 0.1-5 ppm are printed in light green. ....	220
Table 15: Methyl shifts of five oleanane triterpene derivatives compared to β-amyrin in pyridine-d <sub>5</sub> . Downfield shifts > 0.1 ppm are in highlighted in yellow and upfield shifts > 0.1 ppm are highlighted in purple. ....	224
Table 16: Tabulated NMR data for β-amyrin. Data was collected in pyridine-d <sub>5</sub> . ....	245

## LIST OF FIGURES

Figure 1: An example of secondary metabolites synthesized from the primary metabolic building blocks from photosynthesis. The orange box represents a collection of primary metabolic pathways including glycolysis, Calvin cycle, Krebs cycle, oxidative phosphorylation, pentose phosphate, and citric acid cycle and green boxes are primary metabolites. Red boxes denote secondary metabolic pathways and purple boxes indicate secondary metabolites. ....	22
Figure 2: Timeline of the advancements of phytochemistry from the isolation of morphine in 1806 [3]. ....	24
Figure 3: Examples of fatty acids and their abbreviated nomenclature. ....	28
Figure 4: Important steps of fatty acid biosynthesis. The two essential enzymes include acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS) [40]. ....	29
Figure 5: Classes of FAS and PKS. <i>A.</i> FAS can be divided into type I or II enzymes. <i>B.</i> PKS can be classified as type I, II, or III. Type I can be subdivided to iterative and modular, and iterative type I enzymes can be subdivided to highly-reducing (HR), partially-reducing (PR), non-reducing (NR). ....	31
Figure 6: Modifications of saturated fatty acids such as double bond(s), triple bond(s), conjugated polyen(yn)es, epoxides, hydroxyl groups, heterocyclic groups, and cyclopropan(e)nyl rings [64]. ....	33
Figure 7: Monohydroxylated fatty acids that accumulate in plants. Ricinoleic acid accumulates at the largest concentration in seed oil and is well known for its commercial use. ....	34
Figure 8: GC oven programs for non-silylated and silylated FAME (GC1) and pyrrolidide analyses (GC2). ....	37
Figure 9: GC oven programs for silylated metathesis products (GC3) and silylated racemic 3-cyclohexen-1-ol (GC4). ....	39
Figure 10: Oven programs for silylated racemic and enantioenriched 3-cyclohexen-1-ol (GC5). ....	40

- Figure 11: Analysis of *O. violaceus* seed oil. *A, B*. TLC analyses of *O. violaceus* seed oil in solvent systems of different polarities. TLC plates were developed in heptane:ethyl ether:acetic acid at ratios of 60:40:1 (v/v) in *A* and 40:60:1 (v/v) in *B*. Shown for comparison are TLC analyses of Arabidopsis seed triacylglycerol (TAG), which lacks hydroxy fatty acids, and castor bean TAG, which is enriched in ricinoleic acid (OH-18:1), a monohydroxy fatty acid, and TAG species containing two (di-OH TAG) and three (tri-OH-TAG) monohydroxy fatty acids. *C*. GC analyses using an extended run time of fatty acid methyl esters prepared from *O. violaceus* seeds without (-trimethylsilylation) or with trimethylsilyl derivatization (+trimethylsilylation). The novel silylated FAMES were initially named Peak 1 and Peak 2 consistent with their elution order. ....56
- Figure 12: GC-MS analysis of unusual fatty acids in *O. violaceus* seeds. Mass spectra of trimethylsilylated (TMS) fatty acid methyl esters for Peaks 1 (TMS-1a) (*A*) and 2 (TMS-2a) (*C*) in the gas chromatogram in Figure 15C (lower panel). For comparison, mass spectra are shown for TMS-methyl esters of ricinoleic acid (*B*) and densipolic acid (*D*). As shown, TMS-1a and TMS-2a contain ions consistent with fragmentation products containing  $\omega$ -6 hydroxyl groups in mass spectra of TMS-methyl ricinoleic ( $m/z$  187) and densipolic ( $m/z$  185) acids, respectively. ....57
- Figure 13: Substructures based on NMR correlations and coupling data. *A-E* Assigned nuclei in the five working structures were constructed using the  $^1\text{H}$ - $^1\text{H}$ -COSY (solid-line arrows) and  $^1\text{H}$ - $^{13}\text{C}$  HMBC heteronuclear correlations (dashed-line arrows) showing  $^1\text{H}$  correlations to  $^{13}\text{C}$  in Table 2 resulting in structure *F*. Signals for C9-13 were not sufficiently dispersed for their assignment. The dotted box indicates possible locants for the hydroxyl group. ....61
- Figure 14: EI-MS spectra of FAMES. *A*. Mass spectrum of 2a with hypothetical fragmentation pathways shown in the inset. The dihydroxy ester was observed to fragment through two major channels: dehydration and  $\beta$ -scission producing M-1, M-129, and M-233 fragments or the ene fragmentation followed by the losses of H<sub>2</sub>O and MeOH. *B*. Mass spectrum of (TMS)<sub>2</sub>-2a with the origins of major ions annotated on the inset structure. The trimethylsilyl groups block the retro-ene reaction resulting in cleavage  $\beta$  to the ether oxygen, neutral loss reactions, and “deformylation” ions. An  $m/z$  74 McLafferty cleavage ion was readily apparent. *C*. A  $\gamma$ -hydrogen atom transfer and H<sup>+</sup> or TMS<sup>+</sup> migration to the carboxyl group of pyrrolidide and ester derivatives. *D*. Mass spectrum for the methyl ester of 7,18-(TMSO)<sub>2</sub>-24:1 $\Delta$ <sup>15</sup>. *E*. Mass spectrum for the methyl ester of 7,18-(TMSO)<sub>2</sub>-22:1 $\Delta$ <sup>15</sup> .....62
- Figure 15: Mass spectra of diketone from methyl wuhanate (5). ....65

- Figure 16: Electron-impact (EI) MS spectra of fatty acid pyrrolidides. *A.* Mass spectrum of the pyrrolidide of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup> (**6**). Correspondence between the proposed fragments and observed ions is shown in the inset. The even-mass ions below the structure are proposed to result from the sequential chain scission of **6**. The *m/z* 113 ion is consistent with a McLafferty cleavage. Ions marked in bold font result from the dehydration from one or, as denoted by the bold underline, two hydroxy groups. Ions above the structure are derived from cleavage processes beta to the oxygen and subsequent chemistry. *B.* The scission of the hydroxylated amide is consistent with a process reminiscent of an alkene-aldehyde ene reaction leads to the bold line fragmentation in *A.* *C.* Contrastingly, the dominant ions from (TMS)<sub>2</sub>-**6** are formed disproportionately by beta-cleavage chemistry. Ions 510 and 270 are distinguished by undergoing the formal loss of CHO. ....66
- Figure 17: Results from iodate-permanganate “von Rudloff” oxidation of **1a** and **2a**. *A.* The Δ<sup>15</sup> bonds in **1a** and **2a** were cleaved. An *m/z* 231 fragment diagnostic of a 7-hydroxy fatty acid was observed during GC/MS analysis together with an *m/z* 317 fragment, following direct silylation to the trimethylsilyl ester, or an *m/z* 259 ion when methylation preceded silylation. *B.* Additional fragmentation of **2a** due to the additional alkene present. ....68
- Figure 18: To assign the 18-hydroxy stereocenter, alkene metathesis of **2a** was performed in 2-methyl-2-butene. Ring-closing metathesis (RCM) produced in 3-cyclohexenol (**9**) together with the C1-15 metathesis product (**10**) which were then trimethylsilylated with BSTFA-TMSCl. ....69
- Figure 19: Comparison of TMS-modified enantioenriched (*R*)-(dotted line) and racemic (black dashes) cyclohexenol with the RCM product from natural **2a** (red solid line) defines the stereochemistry as *R*. ....70
- Figure 20: Attempted pathway to synthesize methyl 7-hydroxy-17-methyloctadeca-16-enoate. ....72
- Figure 21: Synthesis of methyl 7-hydroxy-16-methyheptadeca-15-enoate used for chiral GCMS analysis to determine the C-7 configuration. ....73
- Figure 22: Synthesis and functionalization of methyl 7-hydroxy-16-methyheptadeca-15-enoate. Both **10d** and **10e** were used for Mosher analysis to determine the configuration of C-7. ....74
- Figure 23: Synthesis and functionalization of methyl (*E*)-7-hydroxy-16-methyheptadeca-2,15-dienoate. Both **27a** and **27b** were used for Mosher analysis to determine the configuration of C-7. ....76
- Figure 24: Final structures of the very-long chain dihydroxy fatty acids from *O. violaceus* seed oil. The structures of mono and diunsaturated nebraskanic (**1**) and wuhanic (**2**) fatty acids, respectively, supported by the MS, NMR, IR, and degradation analyses. ....78

Figure 25: Hypothetical structure of TAG-estolides found in <i>O. violaceus</i> seed oil [107].	79
Figure 26: Mass spectra of fatty acid methyl esters (FAMES) from <i>O. violaceus</i> seeds and transgenic Arabidopsis seeds. Results shown in <i>A</i> and <i>B</i> are from GC-MS analysis of trimethylsilyl (TMS) derivatives of minor dihydroxy FAMES in <i>O. violaceus</i> seeds. Shown in <i>A</i> is a mass spectrum consistent with the methyl ester of TMS <sub>2</sub> -5,16-OH-22:1 and in <i>B</i> is a mass spectrum consistent with that of the methyl ester of TMS-3,14-OH-20:1.	80
Figure 27: Hypothetical biosynthetic pathway for nebraskanic acid. The pathways shown are consistent with the detection of C <sub>20</sub> and C <sub>22</sub> dihydroxy precursors of nebraskanic acid in seed fatty acids and acyl-CoA pools of <i>O. violaceus</i> seeds through discontinuous elongation of the 3-hydroxy intermediate arising from ricinoleoyl-CoA (Pathway 1) or by 3-hydroxylation of the full C <sub>20</sub> -elongation product of ricinoleoyl-CoA (Pathway 2). The characterization of OvFAD2-2 as an oleic acid 12-hydroxylase accounts for the ω-6 hydroxyl group of nebraskanic acid. Though not demonstrated, a second FAE1 (e.g., OvFAE1-2) may be associated with subsequent elongation steps.	81
Figure 28: Examples of natural products synthesized from the integration of fatty acid synthesis with another pathway. Sorgoleone is derived from fatty acid and polyketide synthesis. Capsaicin is synthesized from phenylpropanoid and branched-chain fatty-acid synthesis. Alkamides look to be another example of the convergence of fatty acid with another pathway, but the biosynthetic pathway is still unknown.	84
Figure 29: Biosynthetic pathway of capsaicin. Multiple arrows indicate that several steps occur, but not all of the steps are known to generate the next intermediate. An acyltransferase is thought to be responsible for the last step; a recent study has shown <i>CA01g32970</i> as a possible gene encoding the AT.	86
Figure 30: Abundant alkamides from <i>Echinacea purpurea</i> that are known as Bauer alkamides	88
Figure 31: Endogenous fatty acid amides with important biological activities	89
Figure 32: Sphingosine backbone of ceramide where R is the fatty acid residue.	90
Figure 33: <i>A</i> . Fatty acid synthesis in the chloroplast, and <i>B</i> . in the mitochondria [154].	92
Figure 34: Proposed modular pathway for alkamides biosynthesis in <i>E. purpurea</i> . Reproduced with permission from [132].	96
Figure 35: Stage 5 of <i>E. purpurea</i> seedlings used in feeding experiments	98
Figure 36: GC oven programs for alkamide analysis (GC1).	101



Figure 37: GC oven programs for FAMES analysis from fatty acids (GC2) and FAMES analysis from acyl-CoA (GC3). .....	102
Figure 38: LC method for extracted alkamides. ....	103
Figure 39: The dominant fragment $m/z$ 167 comes from the pericyclic fragmentation and elimination of hexatriene [132]. .....	135
Figure 40: Example of a stage 5 of <i>E. purpurea</i> seedling used in feeding experiments .	136
Figure 41: Enhanced isotopologue ratios in $^N\text{iBu}/12:4\Delta^{2E,4E,8Z,10E}$ resulting from feeding of <i>E. purpurea</i> seedlings with isotopically labeled precursors. (A, B) A. Feeding of $[2-^{13}\text{C}/^{15}\text{N}]$ valine significantly increased the abundance of $M + 2$ versus $M + 1$ species, indicating the retention of the C-N bond from the precursor, corrected for background Val transamination. MeJA and the presence of additional carbon sources, such as glucose (Glc), isoleucine (Ile), and chlorsulfuron (CS) had no significant effect on isotope incorporation. Isotopic envelope revealed by full-scan QqQ mass spectroscopy of $^N\text{iBu}/12:4\Delta^{2E,4E,8Z,10E}$ isolated from seedlings incubated with A. natural abundance valine and isoleucine and B. media supplemented with $[^2\text{H}_8]$ valine, unlabeled isoleucine, and in the presence of chlorsulfuron. Ions with $m/z$ 256 and 255 correspond to the $M + 8$ and $M + 7$ ions resulting from incorporation of eight deuterium atoms from the decarboxylated form of $[^2\text{H}_8]$ valine or the incorporation of seven deuterium atoms following the metabolic exchange of $[^2\text{H}_8]$ valine with 2-ketoisopentanoate via transamination that results in the loss of one of the deuterium atoms. C. Isotopic envelope from plants cultured in standard media (dashed line) or media supplemented with $[^2\text{H}_9]$ isobutylamine hydrochloride (solid line); the $M + 9$ ion at $m/z$ 257 indicates that this precursor is incorporated into the alkamide without loss of any deuterium atoms. ....	139
Figure 42: Retrosynthetic approach to synthesizing $[\text{d}_3]$ -labeled $\text{C}_{12}$ , $\text{C}_{13}$ , and $\text{C}_{14}$ $\alpha,\beta$ -unsaturated alkamides and fatty acids.....	141
Figure 43: Synthesis of aldehydes with and without deuterium labeling at the $\omega$ -end of the acyl chain. ....	143
Figure 44: Alkamide synthesis of $^N\text{iBu}/12:1\Delta^{2E}$ , $^N\text{iBu}/13:1\Delta^{2E}$ , $^N\text{iBu}/14:1\Delta^{2E}$ .....	144
Figure 45: Synthesis of fatty acids 12:1, 13:1, 14:1, $[\text{d}_3]$ -12:1, $[\text{d}_3]$ -13:1, and $[\text{d}_3]$ -14:1. ....	145
Figure 46: Synthesis of 2 <i>E</i> ,4 <i>E</i> -dodeca-2,4-dienoic acid.....	148
Figure 47: Synthesis of (2 <i>E</i> ,4 <i>E</i> ,8 <i>Z</i> ,10 <i>E</i> )-dodeca-2,4,8,10-tetraenoic acid. ....	149
Figure 48: The possible 14-desmethylation product from a SAD2 mutant that would show similar the same mechanistic function as other CYP51 enzymes. ....	156

- Figure 49: Graphical representation of the occurrence of plant CYP clans across plant phylogeny. CYP746 is not represented because it makes up less than 0.2% of the chart. Reproduced with permission from [10].....159
- Figure 50: *A.* Structure of an isoprene unit. *B.* Acyclic assembly of head-to-tail linked isoprene units that form the backbone of terpenes.....161
- Figure 51: Structures of squalene, cholesterol, and  $\beta$ -amyrin. ....161
- Figure 52: Sterol biosynthesis and triterpene biosynthesis share the squalene as a common intermediate in their biosynthetic pathways. The two pathways diverge with the cyclization of (3*S*)-2,3-epoxy-2,3-hydrosqualene. Multiple arrows indicate multiple steps previously occurred in order to synthesize squalene. ....163
- Figure 53: Two examples of triterpene saponins that have biological functions, lyonifolic acid A (antiviral) and pedunculoside (antifungal).....164
- Figure 54: Avenacin A-1, A-2, B-1, and B-2 are structurally similar avenacins. All four triterpenes have a decorated  $\beta$ -amyrin backbone and C3 is a branched sugar chain, D-glucose-D-glucose-D-arabinose [225]. The “A” is indicative of a hydroxyl group at R<sub>1</sub> and “B” means R<sub>1</sub> is a hydrogen [226]. The “1” indicates a *N*-methylantranilate group at C21 and “2” indicates a benzoate group [226]. ....165
- Figure 55: Proposed pathway for avenacin A-1 production. The dotted line indicates that several enzymatic transformations are anticipated. The sequence and chemistry from 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin to des-acyl avenacin A-1 has not been confirmed [217]. Recent work has provide insight into SAD7, SAD9, and SAD10 for the production of avenacin A-1 [232]. ....167
- Figure 56: The prototypical CYP51 reaction, where lanosterol is converted to 4,4-dimethylcholesta-8(9),14,24-trien-3- $\beta$ -ol [211].....169
- Figure 57: Disk assays testing the antifungal activity of avenacin A-1 (A1) and 12-oxo-avenacin A-1 (OA) against *Gaeumannomyces graminis* var. *tritici*, where the loading of the applied triterpene increases from left to right. Control discs were loaded with equal volumes of 75% methanol. Reproduced with permission from [203]. ....171
- Figure 58: Multifunctional CYP enzymes that perform epoxidization reactions. *A.* Two hypothetical pathways for the formation of 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin from  $\beta$ -amyrin. It is not known whether *Sad2* oxidizes intermediates 12,13 $\beta$ -epoxy- $\beta$ -amyrin or 16 $\beta$ -hydroxy- $\beta$ -amyrin. Both pathways could be active. *B.* MycG is a multifunctional CYP enzyme that hydroxylates before it epoxidizes the dienone *C.* GfsF is an example of a multifunctional CYP enzyme that epoxidizes a *trans*-alkene before hydroxylation [203]. ....172

- Figure 59: *A.* Representative members of the CYP51 subfamilies used for sequence alignment include human lanosterol (HsCYP51A1), *Mycobacterium tuberculosis* (MtCYP51B1), *Trypanosoma brucei* lanosterol 14 $\alpha$ -demethylase (TbCYP51E1), *Saccharomyces cerevisiae* lanosterol 14 $\alpha$ -demethylase (ScCYP51F1), *Avena strigosa* (AsCYP51G1), and *Avena strigosa* (AsCYP51H10). Substrate recognition sites are indicated in the black boxes. Black triangles represent residues modeled in the active site and that may interact with either or both  $\beta$ -amyirin or 16 $\beta$ -hydroxy- $\beta$ -amyirin. Red triangles represent sites identified in sad2 mutants where non-synonymous substitutions prevent the synthesis of avenacin; #1027 was a A124V mutation and #791 was the mutation P463S. *B.* Modelled precatalytic binding modes of  $\beta$ -amyirin and *C.* 16 $\beta$ -hydroxy- $\beta$ -amyirin. Reproduced with permission from [203]. .....174
- Figure 61: The structure of 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyirin collected in CDCl<sub>3</sub> with IUPAC numbering. Tabulated are <sup>1</sup>H and <sup>13</sup>C chemical shifts, carbon type, proton coupling constants and multiplicity. The key chemical shift changes versus  $\beta$ -amyirin are indicated in red boxes [203]. .....182
- Figure 62: The oxidized triterpene, 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyirin (**88**), was identified as the main product from *N. benthamiana* co-expressing wild-type CYP51H with SAD1. Numbering is based on IUPAC. Substructures based on NMR correlations and coupling data. *A-E* Nuclei on the five rings were assigned using <sup>1</sup>H-<sup>1</sup>H COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows) in Table 8 resulting in the final structure *E*. .....189
- Figure 63: Structure elucidation of compound **M2** isolated from *N. benthamiana* plants coexpressing the mutant SAD2, A354L, and the triterpene SAD1. IUPAC numbered. Substructures are based on NMR correlations and coupling data. *A-E* Nuclei were assigned and located on the five rings through the use of <sup>1</sup>H-<sup>1</sup>H COSY (dashed arrows), HMBC heteronuclear correlations (solid arrows), and <sup>1</sup>H-<sup>1</sup>H NOESY (red arrows) in Table 9 resulting in the final structure *E*. .....193
- Figure 64: Deduction of the structure of compound **90** produced by the coexpression of SAD2 mutant M471I with SAD1 in *N. benthamiana*. The elucidation molecule is the oxidized triterpene derivative, 15 $\beta$ -hydroxy- $\beta$ -amyirin. IUPAC numbered. Substructures are based on NMR correlations and coupling data *A-E* Nuclei were assigned and located on the five rings through the use of <sup>1</sup>H-<sup>1</sup>H COSY COSY (dashed arrows), HMBC heteronuclear correlations (solid arrows), and <sup>1</sup>H-<sup>1</sup>H NOESY (red arrows) in Table 10 resulting in the final structure *E*. .....198

- Figure 65: Deduction of the structure of compound **91**, a second compound produced by the coexpression of SAD2 mutant M471I with SAD1 in *N. benthamiana*. The elucidated molecule is the oxidized triterpene derivative, 12-keto-13 $\beta$ - $\beta$ -amyrin. IUPAC numbered. Substructures are based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 11 resulting in the final structure *E*.....201
- Figure 66: Structure elucidation of 12,13 $\beta$ -epoxy- $\beta$ -amyrin, **92**. This compound was isolated as M471I-3 from SAD2 mutant M471I. Substructures numbered according to IUPAC. Substructures are based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 12 resulting in the final structure *E*.....205
- Figure 67: Acid-catalyzed epoxide opening to 12 $\alpha$ -keto-13 $\beta$ - $\beta$ -amyrin and tautomerization to 12 $\beta$ -keto-13 $\beta$ - $\beta$ -amyrin.....207
- Figure 68: Crude mass spectrum with epoxide and 15 $\beta$ -hydroxy- $\beta$ -amyrin. *A*. The front half of the peak at 24 min looked to be the epoxide *B*. The tail of the peak at 24 min looked to be 15 $\beta$ -hydroxy- $\beta$ -amyrin with only one hydroxyl group silylated. ....209
- Figure 68 continued .....210
- Figure 69: Mass spectra collected for I471M-1 and A354L. *A*. Mass spectrum of 16 $\beta$ -hydroxy- $\beta$ -amyrin, the product from plants expressing SAD1 and the A354L mutant of SAD2 *B*. Mass spectrum of 15 $\beta$ -hydroxy- $\beta$ -amyrin. ....212
- Figure 69 continued. ....213
- Figure 70: Key fragmentations from the mass spectra of the isolated products. *A*. The loss of  $m/z$  279 is visible in the mass spectra for all products except 15 $\beta$ -hydroxy- $\beta$ -amyrin. *B*. The C/D/E ring fragment released from the retro Diels-Alder reaction fragmentation of 16 $\beta$ -hydroxy- $\beta$ -amyrin. *C*. The other fragment from 12-keto and 12,13 $\beta$ -epoxy seen from the loss of  $m/z$  279.....214
- Figure 71: Structure elucidation of  $\beta$ -amyrin in pyridine- $d_5$ . IUPAC numbered. Substructures based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 16 resulting in the final structure *E*...244

Figure 72: TMS-16 $\beta$ -hydroxy-12, 13-epoxy- $\beta$ -amyrin.....	246
Figure 73: TMS-12-keto-13 $\beta$ - $\beta$ -amyrin .....	247
Figure 74: TMS-12, 13 $\beta$ -epoxy- $\beta$ -amyrin.....	248

**LIST OF ABBREVIATIONS**

<i>ACN</i>	<i>acetonitrile</i>
<i>AsbASI</i>	<i>Avena strigosa</i> $\beta$ - <i>amyrin synthase</i>
<i>BSTFA</i>	<i>N,O-bis(trimethylsilyl)trifluoroacetamide</i>
<i>CDI</i>	<i>N,N-carbonyldiimidazole</i>
<i>CoA</i>	<i>coenzyme A</i>
<i>DCC</i>	<i>N,N'-dicyclohexylcarbodiimide</i>
<i>DHP</i>	<i>dihydropyran</i>
<i>DIBAL-H</i>	<i>diisobutylaluminium hydride</i>
<i>DMAP</i>	<i>4-dimethylaminopyridine</i>
<i>ESI</i>	<i>electrospray ionization</i>
<i>FA</i>	<i>fatty acid</i>
<i>FAMEs</i>	<i>fatty acid methyl esters</i>
<i>FAS</i>	<i>fatty acid synthase</i>
<i>GC/MS</i>	<i>gas chromatography-mass spectrometry</i>
<i>HPLC</i>	<i>high-pressure liquid chromatography</i>
<i>HRMS</i>	<i>high-resolution mass spectrometry</i>
<i>LC</i>	<i>liquid chromatography</i>
<i>LC/MS</i>	<i>liquid chromatography-mass spectrometry</i>
<i>MES</i>	<i>2-(N-morpholino)ethanesulfonic acid</i>
<i>MS</i>	<i>mass spectrometry</i>
<i>n-BuLi</i>	<i>n-butyllithium</i>
<i>NMR</i>	<i>nuclear magnetic resonance</i>
<i>PPh<sub>3</sub></i>	<i>triphenylphosphine</i>
<i>O=PPh<sub>3</sub></i>	<i>triphenylphosphine oxide</i>
<i>QQQ</i>	<i>triple-quadrupole mass spectrometer</i>
<i>QToF</i>	<i>quadrupole time-of-flight mass spectrometer</i>
<i>SAM</i>	<i>S-adenosyl methionine</i>
<i>TEMPO</i>	<i>2,2,6,6-tetramethyl-1-piperidinyloxy</i>

*tHMGR*

*truncated 3-hydroxy-3-methyl-glutaryl-CoA reductase*

*TMS*

*trimethylsilyl*

## ABSTRACT

Author: Teitgen, Alicen, M. PhD

Institution: Purdue University

Degree Received: December 2017

Title: Insights into Plant Biosynthesis of Three Natural Products; Alkamides, Avenacin, and Hydroxylated Fatty Acids

Major Professor: Robert Minto

Plants are an essential resource providing the world with food, oxygen to breath, and animal habitats. Within each plant, there is an integrated network of pathways that produces phytochemicals either metabolism [1, 2]. Following the discovery of the first specialized metabolite in 1806, the common view has shifted from describing a collection of “by-products” to primary metabolism [3, 4] to the products of a value-added expansion of primary metabolic processes that provides an adaptive advantage to the producing organism. Herein, we report the identification of metabolites for three lipid-related pathways and insights into their origins. The first study was initiated by the serendipitous discovery of unusual fatty acids in the seed oil of *Orychophragmus violaceus*. Structural analyses together with targeted syntheses of derivatives and molecular fragments revealed two C<sub>24</sub> dihydroxy fatty acids identified as 7,18-(OH)<sub>2</sub>-24:1Δ<sup>15</sup> and 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup> that account for ~50% of the seed oil. Evidence indicates that these hydroxylated fatty acids are made via discontinuous elongation, a novel process intermediate between fatty acid and polyketide syntheses. Second, <sup>13</sup>C and <sup>2</sup>H labeled precursors were used to gain insight into the biosynthetic pathway of alkamides produced by *Echinacea*. It was determined that the alkamide’s amide moiety is derived from valine and isoleucine, which led to the discovery of the first plant branched-chained amino acid decarboxylase. Separately, the involvement of these atypical acyl chains in alkamide biosynthesis was tested with [<sup>2</sup>H<sub>3</sub>] 2*E*-dodec-, tridec-, and tetradecenoic acids, and their isobutylamides. Analogous CoA derivatives were synthesized to characterize a ligase potentially responsible for the amide bond formation. Finally, to gain insight into the mechanism of a divergent CYP51 enzyme involved in avenacin biosynthesis, epoxy and hydroxylated triterpenes formed by mutant enzymes were identified. Collectively, these studies further the knowledge of the biosynthesis of specialized plant natural products.



## CHAPTER 1. INTRODUCTION

### 1.1 Evolution of Plant Natural Product Biosynthesis

It is estimated that there is almost a half million different plant species around the world [5]. Within each plant, there is an integrated network of enzymes that is carefully regulated, performing an array of chemical reactions to synthesize specific natural products [6]. Many of these metabolic pathways synthesize compounds that are essential to sustaining a living organism, while other pathways synthesize beneficial phytochemicals specific to the plant for genetic, physiological, or biochemical purposes [6].

Plant metabolism can be divided into two categories, primary and secondary metabolism. Recently, there has been a push to rename “secondary metabolism” to “specialized metabolism,” more positively reflecting the importance of these biochemical pathways to their hosts [1, 2]. Common in all plants, primary metabolism synthesizes metabolites that are generally necessary in cells through most developmental stages [7]. These biosynthetic pathways produce carbohydrates, amino acids, membrane and common storage lipids, and nucleic acids that are essentially the same across all organisms [7, 8]. Pathways such as glycolysis, the pentose phosphate, and the Krebs cycle are all examples of primary metabolism [8].

In contrast, secondary metabolism uses specific enzymes to produce the subset of sugars, waxes, lignin, starch, pigments, or alkaloids unique to each plant [9]. Specialized metabolism in the plant kingdom produces natural products with very diverse functions, including providing the means to communicate within the plant and its biome, defense against insects, animals, and microorganisms, and mechanisms to prevent encroachment by other plants [1]. Though there remains chemical variation within the categories used, most specialized metabolites can be classified as polyketides and fatty acids, terpenoids including steroids, phenylpropanoids, alkaloids or other nitrogen-containing natural products [10]. Nevertheless, specialized metabolism is composed of unique biosynthetic pathways derived from the building blocks and, frequently, variations upon the biosynthetic enzymes of primary metabolism (Figure 1) [3, 11]. Some natural products

can be classified as being derived from more than one metabolic category, such as capsaicin. It is derived from phenylalanine, which classifies it as a phenylpropanoid, but it is also derived from fatty acid biosynthesis.

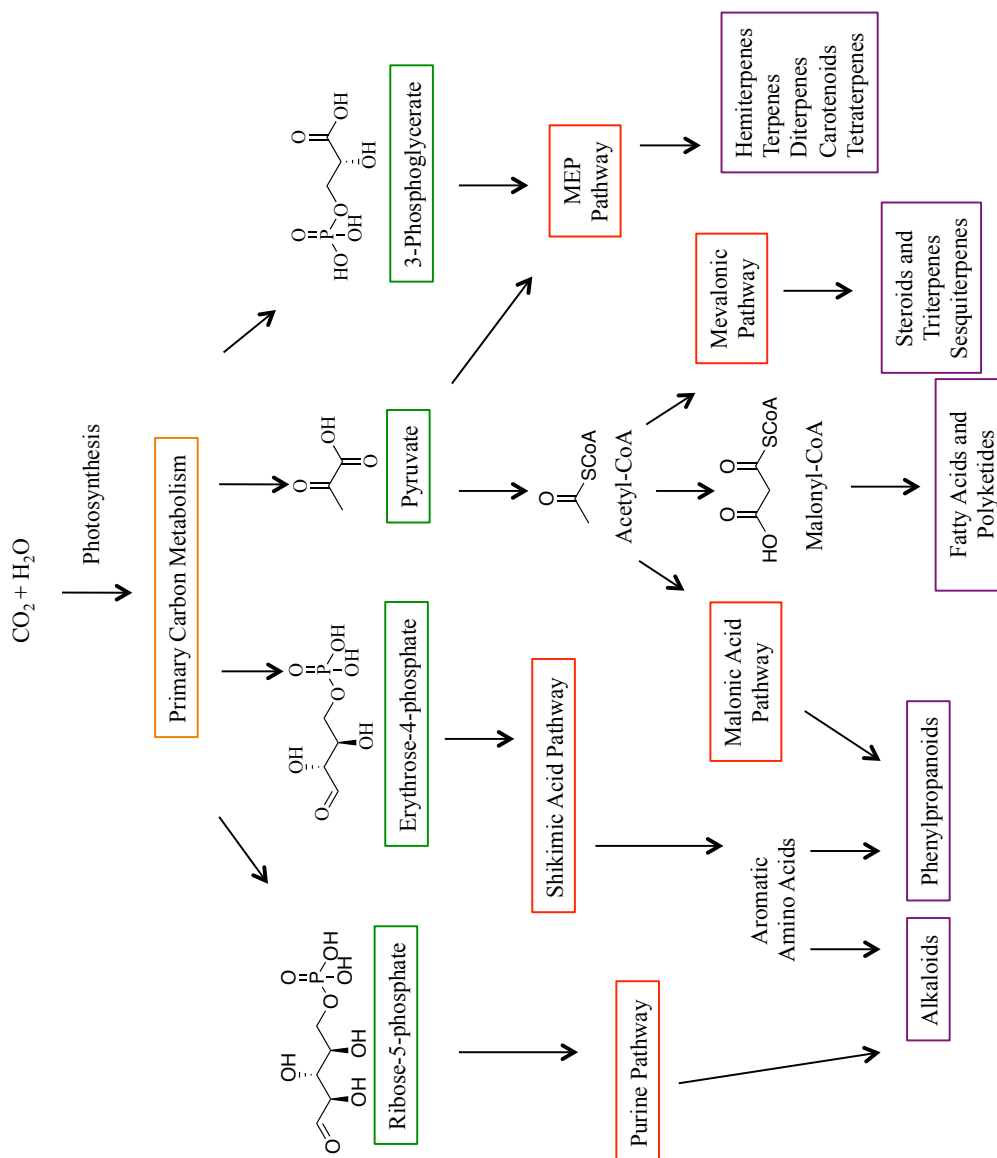


Figure 1: An example of secondary metabolites synthesized from the primary metabolic building blocks from photosynthesis. The orange box represents a collection of primary metabolic pathways including glycolysis, Calvin cycle, Krebs cycle, oxidative phosphorylation, pentose phosphate, and citric acid cycle and green boxes are primary metabolites. Red boxes denote secondary metabolic pathways and purple boxes indicate secondary metabolites.

The foundation for specialized metabolism started in 1806 with the isolation of morphine from opium poppy by Friedrich Wilhelm Sertürner (Figure 2) [3]. After morphine was isolated, there was a flurry of other secondary metabolites isolated. This increase greatly influenced trajectory of organic chemistry especially in the synthetic, pharmaceutical, and analytical areas [3]. In 1882, the first total synthesis of a plant secondary metabolite, indigo, was completed by Johann Friedrich Wilhelm Adolf von Baeyer [3]. The structure of morphine was determined in 1923, and in 1950, a total synthesis was completed [3]. Early on, little was known about the botanical relevance of these products; they were categorized into different classes based on their distribution within plants, as nothing was known about their function [3]. The dogma was that they were thought of as “by-products” of primary metabolism because they were not thought to be crucial for the plants survival [3]. In 1888, Ernst Stahl first stated that the function of secondary metabolites evolves in relation to the environment, but it was mostly disregarded for over 60 years [3, 4]. With advancements in chemical ecology, this idea slowly became more accepted [3, 12, 13]. It is now known that specialized metabolism is unique to specific phylogenetic clades of the plant kingdom and it provides the physiochemical phenotypic differences among the taxonomic groups [1]. Most plants synthesize secondary metabolites due to an evolutionary need. For example, floral scent volatiles have evolved to attract pollinators and thus enhanced fertilization [14]. Some plants produce toxic chemicals such as terpenoids, alkaloids, and phenols to ward off pathogens, competitors, or herbivores [15].

In the early 1900s, scientists used information from analogous organic reactions and structures to create postulates about the formation of natural products rather than information from biochemical evidence [3, 16]. It wasn't until the 1950s when tracer techniques using radiolabeled isotopes, particularly  $^{14}\text{C}$  and  $^3\text{H}$ , expanded, which considerably advanced mechanistic studies [3]. Further advancements occurred using organic chemistry to synthesize desired stable-isotope labeled precursors that benefited from advancements in NMR and MS techniques [3]. NMR provided the locations on labeled atoms and could be used under appropriate conditions for quantitative  $^{13}\text{C}$  NMR experiments, while MS techniques had increased in sensitivity and safety relative to NMR and radioisotopes, respectively [17, 18]. Through tracer experiments, new

biosynthetic pathways were discovered unexpectedly. Prior to radiolabeled isotopes, Ruzicka formulated the “biogenetic isoprene rule,” which stated active isoprenes such as isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP) are synthesized through the mevalonate pathway [3, 19]. Labeled acetate or mevalonate experiments confirmed this rule with the exception of the synthesis of monoterpenoids, diterpenoids, and carotenoids, which showed low incorporation [3, 20]. Twenty-five years later, Schwartz and Arigoni reported the methylerythritolphosphate (MEP) pathway as an alternative pathway to synthesize mono-, di-, and tetraterpenoids [21, 22].

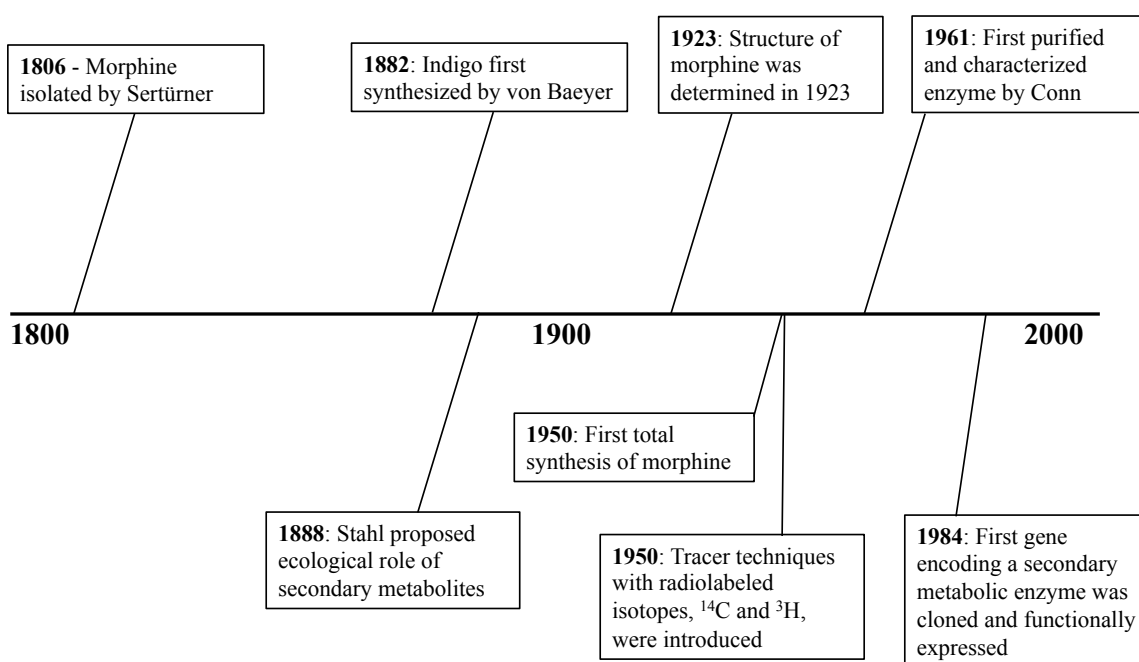


Figure 2: Timeline of the advancements of phytochemistry from the isolation of morphine in 1806 [3].

Until the 1960s, little was known about the enzymes involved in secondary metabolism. Conn’s group was the first to purify and characterize such an enzyme, phenylalanine ammonia-lyase (PAL) [23]. Two developments caused a major breakthrough in enzymology: (1) column chromatography was introduced, which allowed for a rapid and selective separation of proteins [3]; and (2) plant organ and cell cultures were developed, which allowed for a simple way to handle plants as an *in vitro* system

[3]. Through the 1970s and 1980s, these two advancements lead to numerous biosynthetic pathways being characterized at an enzymatic level [3]. For example, Grisebach and Hahlbrock began characterizing enzymes of the phenylpropanoid pathway leading to flavonoids and lignins [24, 25]. Additionally, many aspects of the monoterpenoid indole alkaloid pathway, with over 2000 structures, have been characterized at the enzymatic level from tissue and root cultures [26]. Though cell cultures have been very successful in understanding secondary metabolism, there are still countless natural products that have little-to-no *in vitro* production [3]. For some of these cases, either organ cultures have been shown to help understand the biosynthetic pathways or, in the case of terpenoid biosynthesis, the differentiated whole plant is needed for effective biosynthesis [3].

In the 1980s, genes encoding enzymes were first cloned and functionally characterized, and the molecular toolbox of recombinant DNA technology has become easier and faster ever since [27-29]. This has allowed the expression of recombinant enzymes and the recovery of the protein, which can be used for kinetic and/or structural studies [3]. Genomic tools are being used to transfer genes or pathways between organisms. Specialized metabolism can naturally adapt and be modified, partially since it is dispensable for development and growth [3]. This gives scientists the flexibility to modify organisms in hopes of improving medicinal and crop plants [3]. Of significance to humans, today, about 25% of pharmaceuticals are plant-derived, specialized metabolites or their derivatives [1, 30]. For example, acetylsalicylic acid, commonly known as aspirin, was derived from salicin, an anti-inflammatory ingredient in willow bark [1]. Of the known 400,000 plant species in the world, only a small percentage have been studied for their specialized metabolites and, of those compounds, an even smaller percentage has been studied for biological or pharmacological purposes. Learning more about the origin of the pathways of specialized metabolites along with the discovery of new specialized products will help to understand more about the evolution of secondary pathways and the indispensable functions they have for each plant. Having a better understanding of plant natural products have the potential to increase the contribution of specialized metabolites in the pharmaceutical industry.

## 1.2 Research Objectives

While specialized metabolites are distinguished by their lack of uniform distribution among taxa, they are typically categorized in broad classes based on the relatively uniform chemistry involved in their biosynthesis. Still, out of these well established metabolic themes variations unique to specific plants or pathways continue to be discovered that expand our knowledge of biological chemistry. Herein, through NMR experiments and the synthesis of organic probes, a novel dihydroxylated fatty acid, making up 50% of the seed oil in the *Brassicaceae* family, was identified. This is one of the first examples of a plant accumulating a high concentration of dihydroxy fatty acids, which apparently result from a new variation on fatty acid synthesis. We also describe a group of specialized metabolites, the alkamides, believed to derive from two biosynthetic pathways, fatty acid synthesis in the mitochondria and branched-chain amino acid synthesis. We will look at the roles of branched-chain amino acids, the  $\alpha,\beta$ -unsaturation, and acyl chain length in alkamide biosynthesis. Finally, we studied a CYP51 enzyme that mechanistically behaves differently from other members of the functionally stable CYP51 family. After protein expressions of site-directed mutants of CYP51, products were isolated and structures identified in hope of understanding the mechanism of this novel bifunctional enzyme. All of these projects added breadth to key biochemical processes, underscoring the undiscovered wealth of new chemical processes of plant secondary metabolism.

## CHAPTER 2. IDENTIFICATION OF HYDROXYLATED FATTY ACIDS FOUND IN *ORYCHOPHRAGMUS VIOLACEUS*

### 2.1 Introduction

Fatty acid biosynthesis is common among all plants; it iteratively assembles acetate units onto an acyl chain and fully reduces the carbonyl groups prior to the hydrolytic termination of the chain extension.

In this report, we describe the serendipitous discovery of novel components of the seed oil in the Brassicaceae *Orychophragmus violaceus*. This plant is native to China and is grown extensively as an ornamental for its late winter flower that resembles that of violets [31, 32]. *O. violaceus* has also been used in a number of interspecific crossing studies with *Brassica napus* to generate hybrids for characterization of chromosome stability [33, 34]. Prior reports based on GC analyses have indicated that *O. violaceus* seed oil is rich in linoleic acid (9,12-octadecadienoic acid, 18:2), but unlike most Brassicaceae, has relatively low VLCFA content [35, 36]. In contrast to these reports, we have identified two dihydroxy C<sub>24</sub> fatty acids, not detectable by typical GC protocols that together account for ~ 50% of the seed oil of *O. violaceus* seed oil, using initially thin-layer chromatographic, rather than GC analysis. Elucidation of the unprecedented chemical structures of these fatty acids was achieved using a number of analytical and chemical synthetic approaches.

#### 2.1.1 Lipids from specialized metabolism

Plants produce the majority of the lipids found on the planet, and within lipids, acyl lipids form the largest group [37]. Plant fatty acids have been shown to have several important roles. They are the basic components of cellular membranes and of surface layers, such as epicuticular wax, suberin, and cutin [37]. They are an important source of stored energy and several metabolic derivatives have biological activity [37]. More recently they have been shown to have numerous industrial uses [37]. Additionally, plant-produced  $\omega$ -6, 9 polyunsaturated fatty acids are essential nutrients for humans [37].

Within the lipid biochemistry literature, compounds typically have an abbreviated nomenclature. For example, the monosaturated fatty acid palmitoleic acid, shown in

Figure 3, is abbreviated to  $16:1\Delta^{9Z}$ . The first number represents the length of the acyl chain. The number after the colon reveals the number of double bonds present. The delta indicates the position of the double bond, referenced to the carboxyl terminus, and configuration. So, in the case of  $16:1\Delta^{9Z}$ , the acyl chain is 16 carbons in length with one double bond at position 9 that is in the *cis* (*Z*) configuration. If the fatty acid has an acetylenic functional group like crepenynic acid, the triple bond is represented with the letter 'a' (Figure 3). Crepenynic acid is abbreviated to  $18:1\Delta^{9Z, 12a}$  with a chain length of 18 carbons, a *cis* double bond starting at the 9<sup>th</sup> carbon from the carboxyl end, and an acetylenic bond at the 12<sup>th</sup> position.

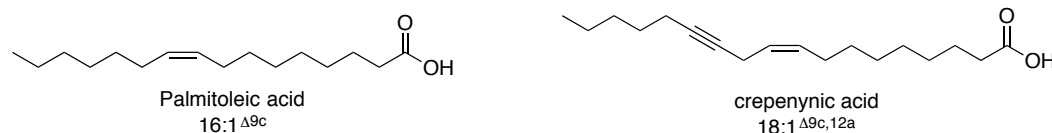


Figure 3: Examples of fatty acids and their abbreviated nomenclature.

Plant fatty acid synthesis has been heavily researched over the past 50 years. It is a key example of secondary metabolism evolving from primary metabolism. All plants produce common fatty acids such as stearic, palmitic, oleic, and linoleic acid, which are components of biological membranes and are used for energy [38]. Further modification of these basic fatty acids produces fatty acids such as crepenynic acid and ricinoleic acid, which are thought of as secondary metabolites unique to certain plant species [38].

*De novo* saturated fatty acid biosynthesis involves two enzyme complexes, fatty acid synthase (FAS) and acetyl-CoA carboxylase (ACCase) (Figure 4) [39, 40]. First though,  $\text{CO}_2$  and  $\text{H}_2\text{O}$  must be reduced to glucose through photosynthesis [6, 40]. Glucose has several roles in plants; one is as a precursor to pyruvate [6]. Pyruvate is also central to several metabolic pathways; for fatty acid synthesis to begin, it must be converted to acetyl-CoA [6]. From there, ACCase catalyzes the carboxylative activation of acetate through the formation of malonyl-CoA, an essential substrate for FAS [6, 39]. FAS in plants is a multienzyme complex which generates both palmitate and stearate [6, 37].  $\text{C}_{16}$  and  $\text{C}_{18}$  saturated fatty acids can be further modified by elongation, desaturation,



epoxidation, hydroxylation, and other functionalizations producing secondary products [6, 40].

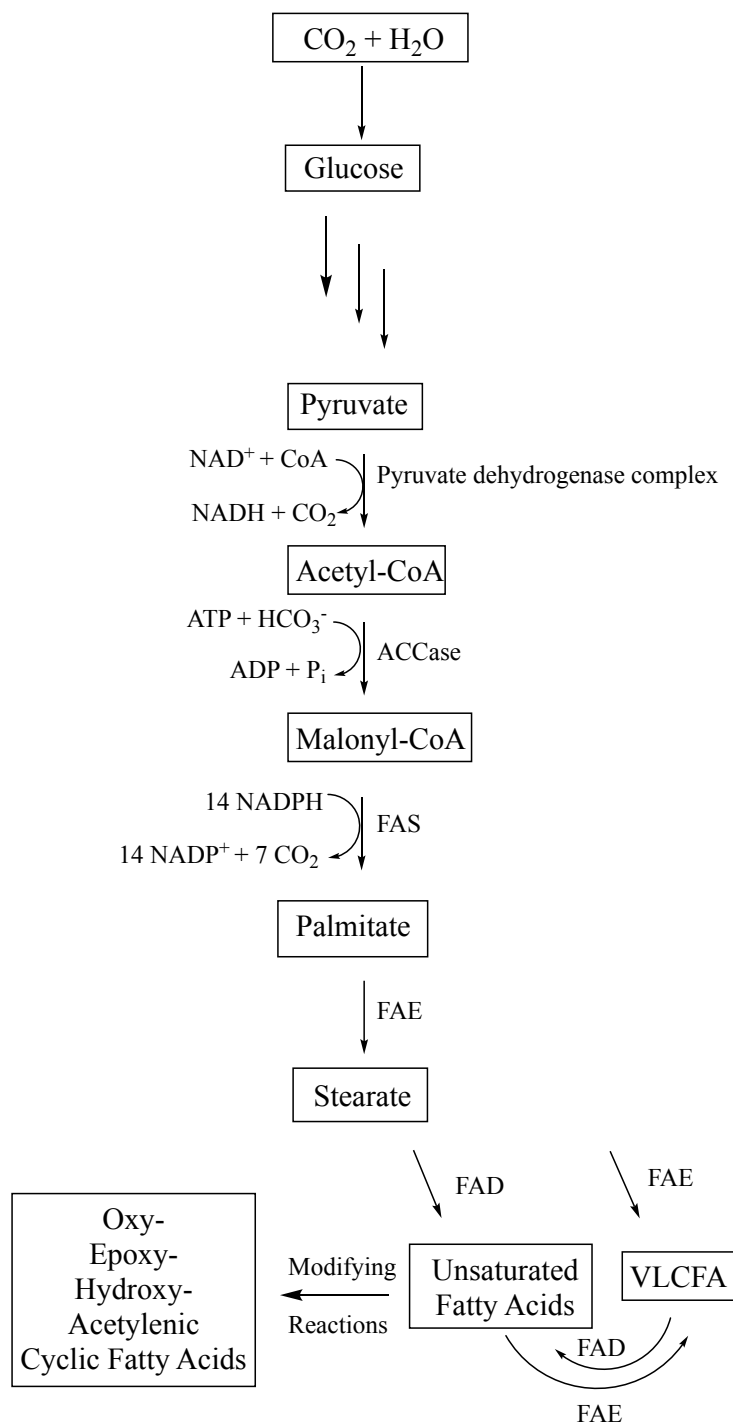


Figure 4: Important steps of fatty acid biosynthesis. The two essential enzymes include acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS) [40].

FASs are structurally and functionally related to certain of the large multienzyme protein complexes, the polyketide synthase (PKS), a highly diversified secondary metabolic enzyme. These enzymatic families share a common evolutionary history [41, 42]. They can have analogous precursors, similar chemistries, comparable structures, and common overall architectural designs [43]. Generally speaking, FASs produce a small group of molecules, whereas PKSs synthesize a large, more oxygenated class of natural products, the polyketides [43]. To date, there is estimated to be over 10,000 known polyketides [43]. Through phylogenetic analysis, it was determined that acyl transferase (AT), ketosynthase (KS), and an acyl carrier protein (ACP) are the domains for condensation reactions performed by PKS and FAS [41, 42]. The three domains are essential for both enzymes [42]. For FAS, the ketoreductase (KR), enoyl reductase (ER), and dehydratase (DH) are essential for the full reduction of the acyl chain after condensation, whereas these three domains and a host of other modifying enzymes are selectively present or absent in PKSs [41, 42].

FASs can be further subdivided into type I or type II based on the organization of the enzyme [44]. Type I uses a single large, multifunctional polypeptide and the type II enzyme complex utilizes discrete monofunctional enzymes [44]. The details of the architecture of PKSs are more complicated than FASs (Figure 5). PKSs can be divided into type I, II, or III enzymes [45]. Type I and type II are similar to FASs, but type III are enzymes that do not use ACP domains [45]. Further analysis can divide type I into iterative and modular varieties [45]. Finally iterative enzymes can be divided into three categories: non-reducing (NR), partially reducing (PR), and highly reducing (HR) [45]. The domains of type I and type II FASs are essentially the same even though the organization is different. However, NR, PR and HR PKSs all have different domains [46]. NR PKSs have the core domains, AT, KS, and ACP, along with a malonyl-CoA domain (MAT) and product template domain (PT). PR PKSs have the core domains plus KR and possibly a DH domain, but not an ER domain. Finally, HR PKSs have the core domains plus the KR, ER, and DH similar to FAS. Environmental pressures are thought to drive the evolution of secondary metabolites and could be the reason for the variety of PKSs [43]. In a recent study of eukaryotes, genes encoding FASs and PKS were the subject of phylogenetic analysis [42]. The functionality was conserved across the genes of FAS,

reflecting the evolutionary history of the organism [42]. The PKS genes function were expanded in the families and some of the genes were completely absent in some of the protest lineages suggesting some PKSs have undergone multiple duplication or deletion events through their evolutionary history allowing them to survive in nature [42].

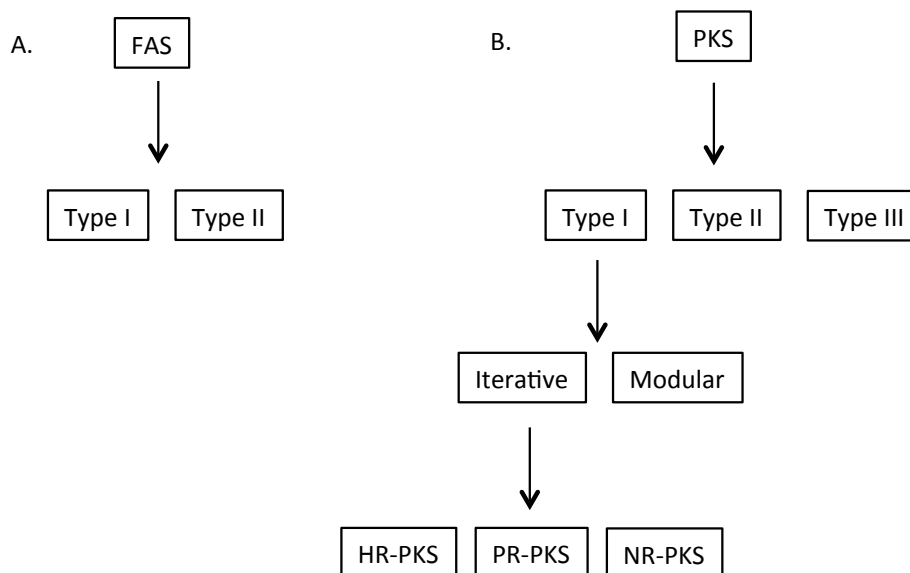


Figure 5: Classes of FAS and PKS. *A.* FAS can be divided into type I or II enzymes. *B.* PKS can be classified as type I, II, or III. Type I can be subdivided to iterative and modular, and iterative type I enzymes can be subdivided to highly-reducing (HR), partially-reducing (PR), non-reducing (NR).

In plants, type II FASs exist because there are a number of condensing enzymes involved [40]. Comparing the two, FASs seem simple compared to PKSs. However, FASs are more complicated than first understood, especially when it comes to their condensing enzyme,  $\beta$ -ketoacyl-ACP synthase (KAS). The condensation reaction occurs by three different condensing enzymes, KAS I, II, and III [40]. In the first step of FAS, the condensation reaction is catalyzed by KAS III using malonyl-ACP and acetyl-CoA as substrates [40]. After the initial condensation, a reduction by  $\beta$ -ketoacyl-ACP reductase, a dehydration by  $\beta$ -hydroxylacyl-ACP dehydrase, and a second conjugate reduction by enoyl reductase occurs to form butyrate [40]. This initial step is followed by cycles of reactions, condensation, reduction, dehydration, and a second reduction, each adding two

carbons and finally a termination step [40]. For the elongation of butyrate to palmitate, the condensation reaction from butyrate to palmitate is catalyzed by KAS I [40]. The final condensation for C<sub>16</sub> to C<sub>18</sub> is catalyzed by KAS II [40]. The termination step can occur in various ways including by acyltransferases or thioesterases [40].

Though fatty acid biosynthesis is well understood there are still several things that remain unclear. The *de novo* fatty acid synthesis occurs in the chloroplast, but fatty acids can be transferred to other organelles such as the endoplasmic reticulum for further modification. The movement of fatty acids within a cell is biochemically less well defined than the initial synthesis [39]. It is known that acyl-CoA synthetases and acyl-ACP desaturase act as acyl-CoA binding proteins to increase the solubility of acyl-CoAs, which can aid in the transfer between membranes [39, 47]. However, the exact mechanism for the targeted transfer of fatty acids from different organelles is still unclear [39].

Plants synthesize a variety of fatty acids, though only a few accumulate throughout the plant [40, 48]. The most common fatty acids produced by fatty acid synthesis are the saturated fatty acids, palmitate (16:0) and stearate (18:0). In plants, stearate is twice as abundant as palmitate [40]. Fatty acid biosynthesis primarily occurs in the chloroplast; limited fatty acid synthesis occurs in the mitochondria. The main purpose of mitochondria fatty acid synthesis is to make lipoic acid. In either case, once the acyl chain is produced further modifications can occur such as double bond(s), triple bond(s), conjugated polyen(yn)es, epoxides, hydroxyl groups, heterocyclic groups, and cycloprop(a/e)nyl rings (Figure 6) [49-55]. The examples shown are all fatty acids synthesized in the chloroplast and then modified at the ER.

Fatty acids can also be modified to compounds that have biological functions such as glycolipids, phospholipids, and sphingolipids. Glycolipids are found in the outer leaflet of cell membranes to maintain stability of the membranes and to facilitate cellular recognition [56]. Phospholipids form lipid bilayers and are the major component of cell membranes; specifically for the cell membranes of the brain and nervous tissue, the major component is sphingolipids [57, 58]. Fatty acids can exist in the energetically dense form of triacylglycerols as a metabolic fuel for energy [48]. Several fatty acids are linked to biological processes; oleic acid can reduce blood pressure, docosahexaenoic acid (DHA)

has shown prevention of cancer and cardiovascular diseases, and  $\gamma$ -linoleic acid displays anti-inflammatory properties [59]. Free fatty acids and conjugated fatty acids in the form of diacylglycerol (DAG) are important in cell signaling [48]. Jasmonic acid is an oxylipin derived from linoleic acid, which can be synthesized to methyl jasmonate. It is known the methyl jasmonate plays a key role in signaling between plants [60]. Some specialized fatty acids have industrial uses. The epoxy fatty acid, lactobacillic acid, is used as a precursor to coatings, adhesives, and biopolymers [61, 62]. Conjugated fatty acids such as linolenic acid, are used in oxygen-initiated “quick-dry” oil paints and varnishes because it can easily be oxidized [63].

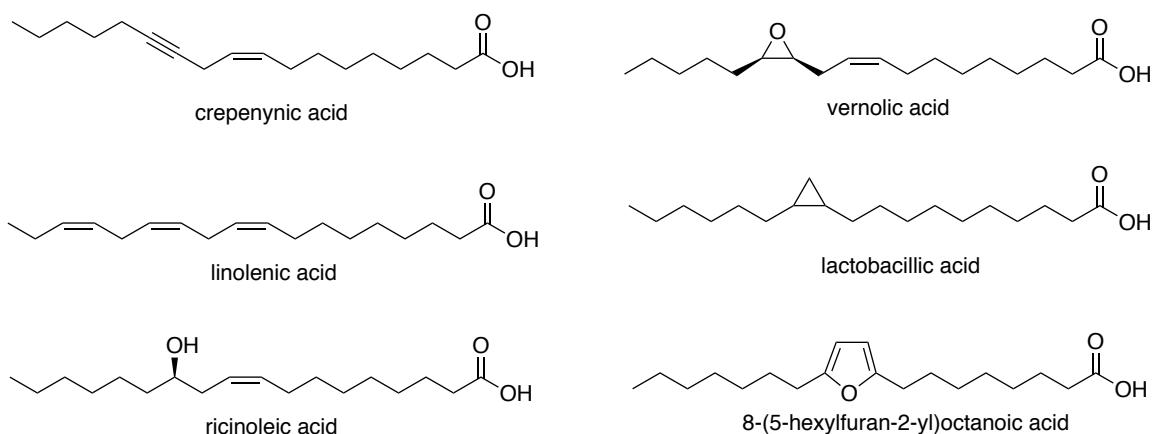


Figure 6: Modifications of saturated fatty acids such as double bond(s), triple bond(s), conjugated polyen(yn)es, epoxides, hydroxyl groups, heterocyclic groups, and cyclopropan(e)nyl rings [64].

Hydroxylated fatty acids such as ricinoleic acid are continuing to be studied for their industrial application in chemicals, foods, and cosmetics [65]. Hydroxylated fatty acids are used as starting material for the synthesis of polymers [66]. They can also be used as additives for the manufacturing of lubricants, stabilizers, and emulsifiers [67]. For example, ricinoleic acid (Figure 6) is used in lubricants, hydraulic fluids, surfactants, and cosmetics [61, 68]. As a practical consideration, the isolation of ricinoleic acid from actual sources is less than ideal for commercial use due to potent toxins and allergens present in castor seeds [69]. Therefore, new hydroxylated fatty acids, whether from an

alternative plant source or by genetic engineering, could unlock this valuable product for a wide range of industrial applications.

Very few plant species accumulate hydroxylated fatty acids. The most common is ricinoleic acid (12OH-18:1 $\Delta^{9Z}$ ), which is produced by the castor bean (*Ricinus communis*) and makes up 80-90% of the seed oil (Figure 7) [69]. Hydroxylated fatty acids can also be found in several species in the *Lesquerella* genus [70]. Some species contain the C<sub>20</sub> homologue of ricinoleic acid, lesquerolic acid (14OH-20:1 $\Delta^{11Z}$ ), while other species produce densipolic acid (12OH-18:2 $\Delta^{9Z,15Z}$ ) [70]. Coriolic acid (13OH-18:2 $\Delta^{9Z,11E}$ ) makes up 70% of the seed oil in *Coriaria* and kamlolenic acid (18OH-18:3 $\Delta^{9Z,11E,13E}$ ) is produced by *Mallotus philippinensis* [71, 72]. Polyhydroxylated fatty acids are rare in plants. Cutin, a component of plant cuticle, contains C<sub>16</sub> and C<sub>18</sub> mono, di, and trihydroxy fatty acids [48]. Though other plant species make hydroxylated fatty acids, currently ricinoleic acid it is the only commercially used oil with a hydroxyl functional group [48].

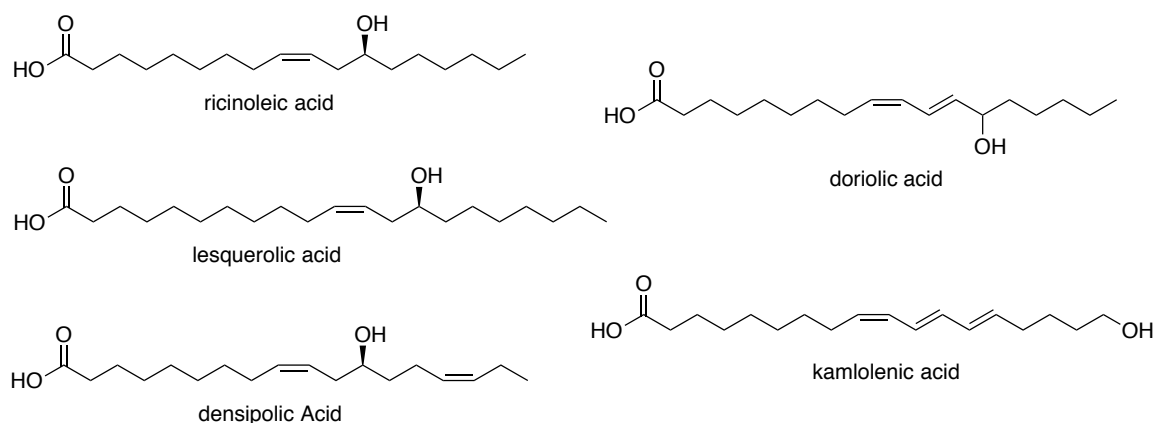


Figure 7: Monohydroxylated fatty acids that accumulate in plants. Ricinoleic acid accumulates at the largest concentration in seed oil and is well known for its commercial use.

Ricinoleic acid is synthesized from oleic acid (18:1) [73, 74]. The reaction is catalyzed by the diverged  $\Delta$ 12 desaturase,  $\Delta$ 12-hydroxylase (FAH12); molecular oxygen and NADPH are both required substrates [73, 74]. FAH12 hydroxylates 18:1 esterified to the *sn*-2 position of phosphatidylcholine, which is then hydrolyzed to form free ricinoleic acid [74]. Through a series of steps, free ricinoleic acid is stored as a triacylglycerol

(TAG) [74]. The Van de Loo group determined that FAH12 had a 67% sequence homology to an oleate desaturase (FAD2) from *Arabidopsis* [75]. This indicated that hydroxylases are closely related to desaturases and provide an excellent example of enzyme evolution [75]. Furthermore, another FAH12 was isolated from *Lesquerella fendleri* [76]. When this *Lesquerella* FAH12 from was expressed with mutant FAD2 in *Arabidopsis*, desaturation activity of the mutant phenotype was only partial suppressed [76]. Both desaturase and hydroxylase activities were still present exhibiting bifunctional properties of this FAH12; this also proves the close evolutionary relationship of the two enzymes [76].

Specialized fatty acid biosynthesis such as hydroxylated fatty acids is an excellent example of secondary metabolism evolving from primary metabolism. The fatty acids 16:0 and 18:0 are synthesized in all plant species, but each specialized fatty acid is unique to each plant species. Each specialized fatty acid identified adds another level of complexity to a once thought of simple pathway.

### 2.1.2 Specialized fatty acids in the seed oils of the Brassicaceae

Seed oils of the Brassicaceae are distinct from those of most plant families because of their high content of very long-chain fatty acids (VLCFAs) with carbon chain-lengths  $\geq C_{20}$ . Seed oils from Brassicaceae such as *Arabidopsis thaliana* and *Camelina sativa*, for example, contain >15% of the  $C_{20}$  fatty acid 11-eicosenoic acid (20:1) [77, 78]. Seed oils of other Brassicaceae such as *Crambe abyssinica* and *Brassica napus* are enriched in the  $C_{22}$  fatty acid 13-docosenoic acid (erucic acid, 22:1) [77, 79, 80], and *Lunaria annua* and *Cardamine graeca* are enriched in the  $C_{24}$  fatty acid 15-tetracosanoic acid (nervonic acid, 24:1) [80, 81]. The high content of VLCFAs make Brassicaceae seed oils valued for industrial applications, such as lubricants and chemical feedstocks [82]. Apart from their high VLCFA content, Brassicaceae seed oils are composed of typical  $C_{16}$ - $C_{18}$  saturated and unsaturated fatty acids. The only known exception to this among Brassicaceae are the seed oils of *Physaria* and *Paysonia* (formerly *Lesquerella*) species that contain 40% to 60% of the  $C_{20}$   $\omega$ -6 hydroxy fatty acid 14-OH-11-eicosenoic acid (lesquerolic acid) [83, 84].

More elaborately modified fatty acids, such as lesquerolic acid, often termed “unusual fatty acids”, occur in seed oils of diverse plant species [85, 86]. These include monohydroxy fatty acids such as ricinoleic acid (12-OH-9-octadecenoic acid) prevalent in castor bean (*Ricinus communis*) seeds, epoxy fatty acids such as vernolic acid (12-epoxy-9-octadecenoic acid) enriched in Stoke’s aster (*Stokesia laevis*) seeds, and conjugated fatty acids such as eleostearic acid (9,11,13-octadecatrienoic acid) found in high amounts in tung (*Vernicia fordii*) seeds [85, 86]. The discovery of these and other unusual fatty acids has been facilitated by gas chromatography (GC), which has enabled rapid screening of fatty acid methyl esters prepared from seeds. These analyses were largely conducted during the 1960s and 1970s. Since this time, very few “new” unusual fatty acids have been identified as major components of seed oils. Many unusual fatty acids have considerable economic potential for a range of industrial uses [82]. However, most plants that make unusual fatty acids often have limited agronomic potential without extensive breeding efforts. As a result, these plants have been used as sources of genes that can be transferred to established oilseed crops to generate novel oil functionalities [87]. In addition, the enzymes that make unusual fatty acids are divergent forms of widely-occurring fatty acid modification enzymes, such as fatty acid desaturases, and have proven to be valuable biochemical tools for understanding variant enzyme substrate specificities and reaction chemistries through heterologous expression (e.g., [88-90]).

## 2.2 Materials and Methods

$^1\text{H}$ ,  $^{13}\text{C}$ , DEPT, COSY, HSQC, HMBC, and NOESY NMR spectra were recorded on a Bruker Avance-II NMR spectrometer using  $\text{CDCl}_3$  as solvent; residual  $\text{CHCl}_3$  ( $^1\text{H}$ ,  $\delta$  7.26; 500 MHz) and  $\text{CDCl}_3$  ( $^{13}\text{C}$ ,  $\delta$  77.0; 125 MHz) were used as the internal chemical shift references.  $^{19}\text{F}$  NMR Conditions:  $\text{CDCl}_3$  was used as the solvent and trifluorotoluene (-63.72 ppm) was used as an internal standard. Peak multiplicities and shapes were annotated as follows: s, singlet; d, doublet; t, triplet; q, quartet; quint, quintet; m, multiplet; br, broad. Infrared spectra were recorded as neat films on NaCl plates with a Fourier-transform infrared spectrophotometer (Nicolet Avatar 330). Mass spectra were collected with an Agilent 6520 Q-ToF mass spectrometer or GC/MS.



Dichloromethane (DCM) was dried by distilling from  $\text{CaH}_2$ . Tetrahydrofuran (THF) and diethyl ether were distilled from sodium and benzophenone. Benzene and hexanes were used as received. Most reagents were procured from Aldrich. Acetic acid, hydrogen peroxide and borane- $\text{Me}_2\text{S}$  complex were purchased from Fisher Scientific. Basic  $\text{Al}_2\text{O}_3$  (W200) was purchased from M. Woelm Eschwege, Germany. Dess-Martin periodinane (DMP) was synthesized by the method of Boeckman [91]. Silica gel for chromatography from Dynamic Absorbents (60-200 micron, #020311) was slurry packed for flash chromatography using the indicated solvent. All glassware was oven-dried at  $120\text{ }^\circ\text{C}$  and reactions were run under nitrogen, unless otherwise noted. Quenched reaction mixtures were concentrated using a rotary evaporator with a water bath at  $25\text{ }^\circ\text{C}$  –  $45\text{ }^\circ\text{C}$ , unless otherwise specified.

### 2.2.1 GC/MS protocols for the structure elucidation of 7,18-(OH)<sub>2</sub>-24:2 $\Delta$ 15, 21

GC/MS analyses were performed using five methods with an Agilent 7890A GC/5975C MS using helium (99.99% UHP) as the carrier gas. The column and oven programs for each method are detailed below.

#### 2.2.1.1 FAME and pyrrolidide GC/MS methods

FAME analysis and pyrrolidide analysis were suspended in ACN and analyzed with a HP-5 (30 m x 250  $\mu\text{m}$  x 0.25  $\mu\text{m}$ ) column, using oven programs GC1 and GC2, respectively (Figure 8). Samples were also silylated and run under the same conditions.

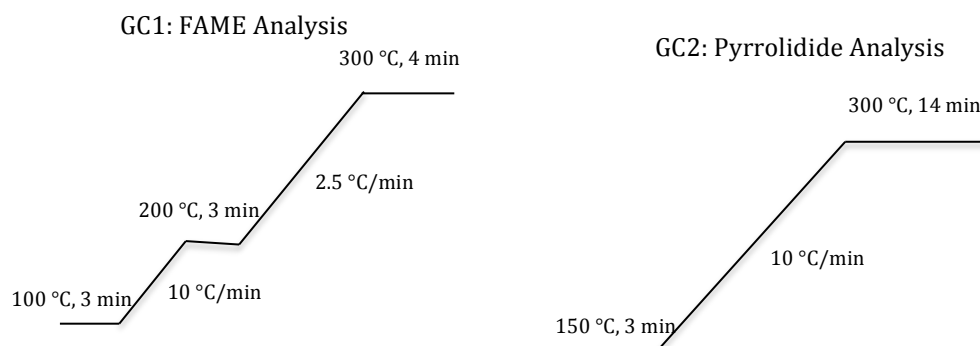


Figure 8: GC oven programs for non-silylated and silylated FAME (GC1) and pyrrolidide analyses (GC2).

GC/MS parameters were as follows:

- Split-Splitless Inlet: Splitless
- Inlet temperature: 250 °C
- Pressure: 16.834 psi
- Septum purge flow: 3 mL/min
- He flow rate: 1.56 mL/min for 32 min

The scan parameters were as follows:

- Solvent Delay: 6 min
- Scan Speed: Normal
- Start Mass: 50.00 amu; Stop Mass: 800.00 amu
- Threshold: 150; Scans/sec: 1.99

#### 2.2.1.2 Metathesis and racemic 3-cyclohexen-1-ol GC/MS methods

Silylated metathesis products and silylated achiral 3-cyclohexen-1-ol were analyzed using a VF23-MS column (Varian; 30 m x 250  $\mu$ m x 0.25  $\mu$ m). Oven programs shown in Figure 9.

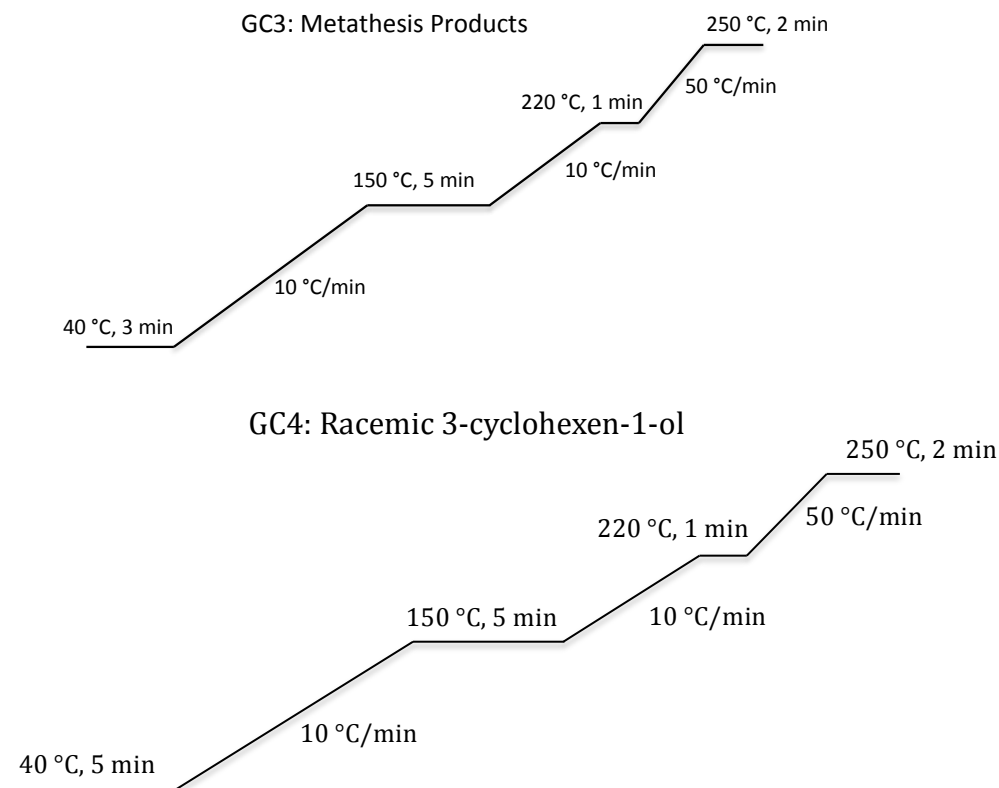


Figure 9: GC oven programs for silylated metathesis products (GC3) and silylated racemic 3-cyclohexen-1-ol (GC4).

GC/MS parameters were as follows:

- Split at a ration of 7:1
- Inlet Temperature: 250 °C
- Pressure: 15.236 psi
- Septum Purge Flow: 3mL/min
- He Flow Rate: 1.9 mL/min; 29.6 min

Scan Parameters were as follows

- Solvent Delay: 3 min
- Scan Speed: Normal
- Start Mass: 50.00 amu; Stop Mass: 550.00 amu
- Threshold: 150; Scans/sec: 1.99

### 2.2.1.3 GC/MS method for chiral separations

The separation of the enantiomers of 3-cyclohexen-1-ol was achieved using a Beta DEX 120 column (Supelco; 30 m x 250  $\mu$ m x 0.25  $\mu$ m column). Oven programs depicted in Figure 10.

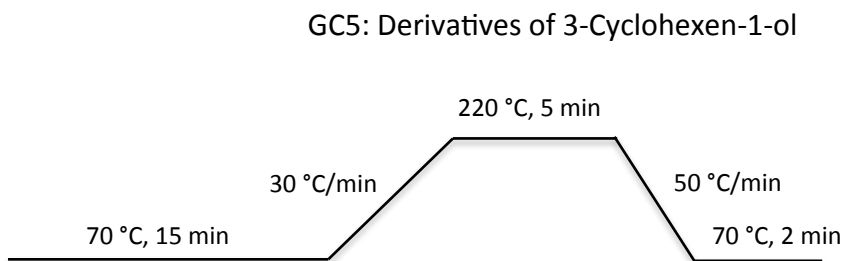


Figure 10: Oven programs for silylated racemic and enantioenriched 3-cyclohexen-1-ol (GC5).

GC/MS parameters were as follows:

- Split-Splitless Inlet: Splitless
- Inlet Temperature: 200 °C
- Pressure: 4.3792 psi
- Septum Purge Flow: 3 mL/min
- He Flow Rate: 0.659 mL/min for 30 min

Scan parameters were as follows:

- Solvent Delay: 6 min
- Scan Speed: Normal

- Start Mass: 50.00 amu; Stop Mass: 800.00 amu
- Threshold: 150; Scans/sec: 1.99

### 2.2.2 Isolation of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup> from seed oil

Seed oil (0.8918 g) was dissolved in 2.5% H<sub>2</sub>SO<sub>4</sub> in methanol. The solution was homogenized in the vial and heated at 80 °C for 1 h. The dark green solution was cooled to room temperature, and the reaction was quenched with H<sub>2</sub>O (2 mL). Hexanes (2 x 1 mL) was added, the mixture was vortexed, centrifuged, and organic layer removed. Crude FAMES were concentrated under a stream of N<sub>2</sub>. The oil was purified by SiO<sub>2</sub> flash chromatography with hexane: Et<sub>2</sub>O: acetic acid (40:60:1) (R<sub>f</sub> = 0.25) to yield 36% (0.3207 g). IR (neat)  $\bar{\nu}$  3385, 3005, 2929, 2855, 1740, 1653, 1456, 1437, 1363, 1259, 1200, 1173, 1065, 1016, 870; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  5.53 (m, 2H), 5.38 (m, 1H), 5.26 (m, 1H), 3.66 (s, 3H), 3.14 (d, *J* = 6.6 Hz, 2H), 2.47 (t, *J* = 7.4 Hz, 2H), 2.39 (m, 4H), 2.29 (t, *J* = 7.3 Hz, 2H), 2.03 (m, 3H), 1.59 (m, 8H), 1.25-1.34 (m, 13H), 0.95 (t, *J* = 7.5 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  174.2, 133.5, 132.2, 128.5, 125.0, 71.8, 71.1, 51.4, 37.5, 37.2, 36.7, 35.4, 34.0, 29.6, 29.4, 29.3, 29.19, 29.15, 27.4, 25.6, 25.3, 24.9, 23.5, 20.5, 14.3 ; HRMS (ESI) *m/z* calcd for [M + H<sup>+</sup>] C<sub>25</sub>H<sub>47</sub>O<sub>4</sub> 411.3474, found 411.3472.

### 2.2.3 Derivatives of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup>

**Fatty Acid Pyrrolidides (6):** In a conical vial, **1a/2a** (*ca.* 1 mg) was azeotropically dried with benzene (3 x 1 ml) under a stream of N<sub>2</sub>. The residue was then incubated with pyrrolidine (4.2  $\mu$ L) and glacial acetic acid (2  $\mu$ L) in a heating block at 80 °C for 1 h. After cooling, H<sub>2</sub>O (1 mL) and diethyl ether/hexane (1:1, 3 mL) were added. The mixture was thoroughly extracted by vortex mixing and centrifugation. The resulting organic layer was washed twice with H<sub>2</sub>O (1 mL), the solvent evaporated was under a stream of N<sub>2</sub>, and the pyrrolidides were redissolved in 1 mL of hexane for GC/MS analysis (Method GC2).

**Silylation of Intact Fatty Acid Pyrrolidides (6a):** For the conversion of **6** to a TMS derivative, the hexane was removed from the alcohol and then BSTFA + 2% TMSCl (100  $\mu$ L) was added to the vial. After overnight stirring at room temperature, the

silylating reagent was eliminated with a stream of N<sub>2</sub>. The bis(silyloxy) derivative was analyzed as a solution in 200  $\mu$ L of hexanes solution by GC-MS (Method GC2).

**Diketone (5):** In a 5-mL vial, a mixture of **1a/2a** (1.8 mg, 4.5 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> and Dess-Martin reagent (4.0 mg, 9.4 mmol, 2.1 equiv) was added. After stirring for 2 h at room temperature, the solvent was removed by rotary evaporation. The crude product was dissolved in 5 mL hexanes and the solution was run through a SiO<sub>2</sub> column constructed from a Pasteur pipette. The column was washed with hexane (2 mL) and the hexane was removed under reduced pressure. IR (neat)  $\bar{\nu}$  3008, 2930, 2855, 1738, 1715, 1653, 1558, 1457, 1436, 1411, 1362, 1199, 1171, 1101, 665; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  5.56 (dt,  $J$  = 1.63, 7.5, 11 Hz, 1H), 5.39 (m, 1H), 5.37 (m, 1H), 5.34 (m, 1H), 3.66 (s, 3H), 3.63 (m, 1H), 3.57 (m, 1H), 2.31 (t,  $J$  = 7.5 Hz, 2H), 2.22 (t,  $J$  = 6.5 Hz, 2H), 2.13-2.16 (m, 2H), 2.05 (m, 4H), 1.63 (quint,  $J$  = 7.5 Hz, 2H), 1.53 (m, 2H), 1.37-1.47 (m, 6H), 1.26-1.36 (m, 12H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  211.2, 208.4, 174.0, 133.7, 132.9, 127.1, 120.9, 51.5, 42.8, 42.4, 42.2, 41.8, 33.8, 31.6, 29.3, 29.2, 29.1, 28.7, 27.5, 24.7, 23.8, 23.4, 21.6, 20.5, 14.3; HRMS (ESI)  $m/z$  calcd for [M + H]<sup>+</sup> C<sub>25</sub>H<sub>42</sub>O<sub>4</sub> 407.3156, found 407.3154.

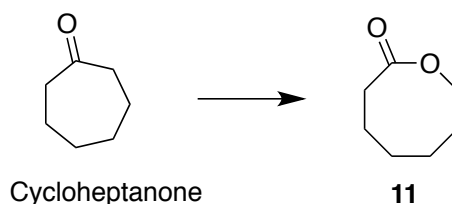
#### 2.2.4 Configuration of C-18 in 7,18-(OH)<sub>2</sub>-24:2 $\Delta$ <sup>15,21</sup>

**Fragmentation of 2b by Alkene Metathesis:** Ester **2b** (*ca.* 1 mg) was dissolved in a dry CH<sub>2</sub>Cl<sub>2</sub>:2-methyl-2-butene mixture (1:1, 1 mL) before a CH<sub>2</sub>Cl<sub>2</sub> solution of second-generation Grubbs catalyst (Aldrich 569747, 3.2 mg/ml, 60  $\mu$ L) was added. The solution was stirred at room temperature overnight. Activated carbon (100 mg) was added to absorb the catalyst; the suspension was stirred for 10 min before the mixture was filtered through silica gel, rinsing the absorbent with CH<sub>2</sub>Cl<sub>2</sub>:Et<sub>2</sub>O (1:1). A solution of BSTFA + 2% TMSCl agent (200  $\mu$ L) was added to the solution, which was stirred at room temperature for 2 h to silylate the metathesis products. GC/MS analysis was carried out using method GC3 for **9** and methods GC4 and GC5 for **10**.

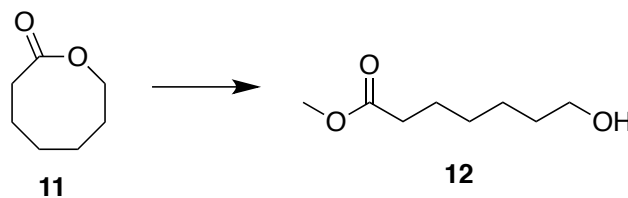
**Racemic 3-cyclohexen-1-ol (9b):** This procedure is based on the work of Donohoe [92]. All glassware was washed with 1 M NaOH, rinsed with H<sub>2</sub>O, and oven-dried at 120 °C. To a round-bottomed flask containing 1,4-cyclohexadiene (0.47 mL, 5 mmol, 1.0 equiv) and dry THF (5 mL), 9-BBN (0.5 M in THF, 10 mL, 5 mmol, 1.0

equiv) was added dropwise. The solution was stirred at room temperature overnight. The reaction was quenched by the sequential dropwise addition of NaOH solution (3 M, 1.5 mL) and H<sub>2</sub>O<sub>2</sub> (30% (v/v), 1.5 mL). The resulting solution was heated to reflux for 1 h, cooled, and then poured into brine (25 mL). The product was extracted with ether (2 x 25 mL), dried with MgSO<sub>4</sub>, and the solvent was removed with a rotary evaporator until ~1 mL remained. Alcohol **9b** was purified by SiO<sub>2</sub> flash chromatography (pentane:Et<sub>2</sub>O, 2:1). A solution of BSTFA + 2% TMSCl agent was added and the resulting mixture was stirred at room temperature for 2 h to provide racemic **9b**, which was analyzed by GC/MS (Methods GC4 and GC5) without further purification.

**(R)-3-Cyclohexen-1-ol ((R)-9):** Using a method modified from Brown [93], diisopinocampheylborane was prepared by the dropwise addition (*S*)-(-)- $\alpha$ -pinene (7.8 mL, 49.1 mmol, 2.0 equiv) to a stirred cooled solution of BMS•(CH<sub>3</sub>)<sub>2</sub>S (2.4 mL, 25.0 mmol) and 9 ml of THF at 0 °C, and the mixture was allowed to react for 3 h. In a separate flask, 1,4-cyclohexadiene (0.47 mL, 5 mmol, 1.0 equiv) was cooled to 0 °C and then a portion of the diisopinocampheylborane solution (1.8 mL, 5 mmol) was added dropwise. The solution was stirred at 0 °C for 2 h and then maintained at 4 °C in a refrigerator for approximately 18 h. Sodium hydroxide (3 M, 1.5 mL) was added followed by the dropwise addition of H<sub>2</sub>O<sub>2</sub> (30% w/v, 1.5 mL). The solution was heated to reflux for 1 h and then cooled to room temperature before pouring into 25 ml of brine. The reaction mixture was extracted with ether (2 x 25 mL) and the separated organic phase dried with MgSO<sub>4</sub> before the solvent was carefully evaporated with a rotary evaporator until ~1 mL remained. The crude solution was purified by silica gel flash chromatography (pentane:Et<sub>2</sub>O, 2:1). An aliquot of the alcohol was derivatized to enantioenriched **9a** with BSTFA + 2% TMSCl agent at room temperature for 1 h and the product was analyzed without further purification.

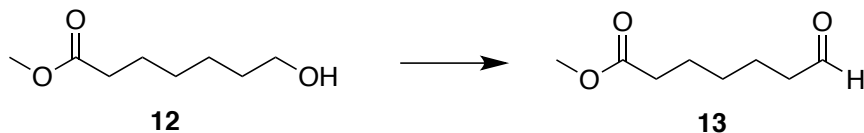
2.2.5 Route 1 for the attempted synthesis of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup>

**2-Oxocanone (11):** Using the method from Kai [94], cycloheptanone (2.6454 g, 23.58 mmol) was stirred in DCM (as-received, 75 mL). Solid 77% m-CPBA (5.5825 g, 24.91 mmol, 1.05 equiv) was added to the ketone solution and the reaction was allowed to proceed at room temperature for 5 days. The reaction was quenched and the organic layer was washed with saturated NaHCO<sub>3</sub> (4 X 100 mL), followed by H<sub>2</sub>O (200 mL). The DCM solution was dried with Na<sub>2</sub>SO<sub>4</sub> prior to being concentrated under reduced pressure, and being carried on to the next step.

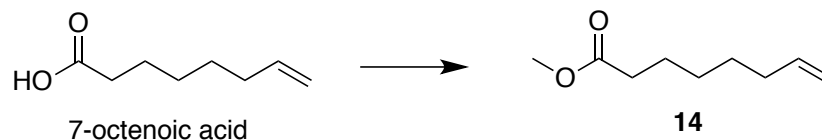


**Methyl 7-hydroxyheptanoate (12):** Lactone **11** (~23 mmol) was dissolved in 33 mL methanol. H<sub>2</sub>SO<sub>4</sub> (220 μL) was added and the reaction was stirred for 15 h at room temperature. The solvent was removed and the residue dissolved in diethyl ether (35 mL). The organic layer was washed with H<sub>2</sub>O (2 x 35 mL) and dried with Na<sub>2</sub>SO<sub>4</sub>. Ether was removed to yield methyl 7-hydroxyheptanoate (0.8277 g, 22% yield). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 3.67 (s, 3H), 3.64 (t, *J* = 6.5 Hz, 2H), 2.31 (t, *J* = 7.5 Hz, 2H), 1.64 (quint, *J* = 7.4 Hz, 2H), 1.57 (quint, *J* = 6.9 Hz, 2H), 1.34-1.37 (m, 4H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 174.3, 63.0, 51.6, 34.1, 32.7, 29.0, 25.5, 25.0. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [94].



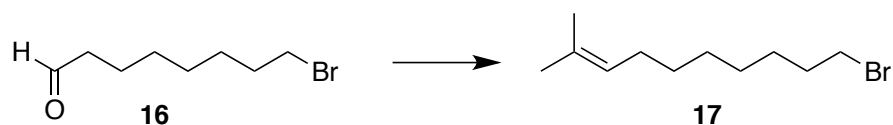


**Methyl 7-oxoheptanoate (13):** Pyridinium dichromate (18.5051 g, 49.2 mmol, 1.5 equiv) and Celite (6.6322 g) were suspended with the aid of a magnetic stirrer in 53 mL of DCM. Methyl 7-hydroxyheptanoate, dissolved in DCM (10 mL), was added dropwise via an addition funnel and reaction was allowed to stir for 4 h at room temperature. The solvent was removed and the crude product was purified by column chromatography with hexane: EtOAc (4:1) to yield methyl 7-oxoheptanoate (2.2486 g, 43%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (t,  $J = 1.5$  Hz, 1H), 3.67 (s, 3H), 2.44 (td,  $J = 7.3, 1.5$  Hz, 2H), 2.32 (t,  $J = 7.5$  Hz, 2H), 1.65 (quint,  $J = 7.5$  Hz, 4H), 1.36 (m, 2H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  202.5, 174.0, 51.6, 43.8, 33.9, 28.7, 24.8, 21.8.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [94].



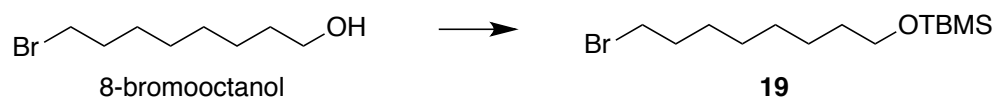
**Methyl 7-octenoate (14):** 7-Octenoic acid (0.1842 g, 1.30 mmol) was dissolved in methanol (0.284 mL, 5.4 equiv).  $\text{H}_2\text{SO}_4$  (29  $\mu\text{L}$ , 0.54 mmol, 0.42 equiv) was added to initiate the reaction and the solution was stirred overnight. The reaction was quenched with  $\text{H}_2\text{O}$  and allowed to stir for 5 minutes. The solution was extracted twice with diethyl ether, dried with  $\text{MgSO}_4$ , and the solvent was removed to yield methyl 7-octenoate (0.1990 g, 98%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  5.80 (ddt,  $J = 14, 10, 6.5$  Hz, 1H), 4.99 (dq,  $J = 17, 1.5$  Hz, 1H), 4.94 (br d,  $J = 10$  Hz, 1H), 3.67 (s, 3H), 2.31 (t,  $J = 7.5$  Hz, 2H), 2.05 (q,  $J = 7$  Hz, 2H), 1.63 (quint,  $J = 7.5$  Hz, 2H), 1.36 (m, 4H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  174.4, 139.0, 114.5, 51.6, 34.2, 33.7, 28.7, 28.6, 25.0.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [95].





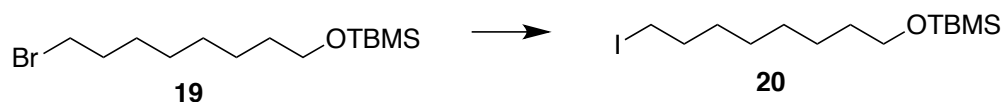
**10-Bromo-2-methyldec-2-ene (17):** Isopropyltriphenylphosphonium iodide (1.1289 g, 2.63 mmol, 1.05 equiv) was suspended in dry diethyl ether (13 mL). Potassium *tert*-butoxide (0.2820 g, 2.50 mmol, 1 equiv) was added and the reaction mixture immediately turned dark orange. The reaction was allowed to stir at room temperature for 30 min. The aldehyde **16** (0.5215 g, 2.51 mmol, 1 equiv) was dissolved in dry ether (0.5 mL) and added to the dark orange solution leading to the formation of a white precipitate. The suspension was stirred at room temperature for 2.5 h. The reaction was quenched with H<sub>2</sub>O (10 mL) and extracted with diethyl ether (20 mL). The solvent was removed from the organic layer and the crude product was purified by column chromatography with pentane: Et<sub>2</sub>O (5:1) to yield give a clear oil, 10-bromo-2-methyldec-2-ene (0.2465 g, 42%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 5.10 (t, *J* = 7.2 Hz, 1H), 3.40 (t, *J* = 6.9 Hz, 2H), 1.96 (m, 2H), 1.85 (quint, *J* = 7.2 Hz, 2H), 1.68 (s, 3H), 1.59 (s, 3H), 1.43 (m, 2H), 1.29-1.35 (m, 6H).

#### 2.2.6 Route 2 for the synthesis of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup>

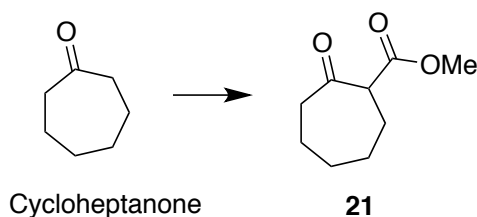


**(8-Bromooctoxy)(*tert*-butyl)dimethylsilane (19):** 8-Bromooctanol (8.3876 g, 40.12 mmol) was dissolved in DCM (200 mL) and the cooled to 0 °C with an ice-water bath. Imidazole (13.6546 g, 200.57 mmol, 5 equiv) and TBDMSCl (12.3346 g, 81.84 mmol, 2 equiv) were added followed by a catalytic amount of DMAP (0.4878 g, 3.99 mmol, 0.1 equiv). The solution was stirred at room temperature for 24 h. The reaction was quenched with saturated NH<sub>4</sub>Cl, extracted with dichloromethane, and the solvent was removed. The product was purified via column chromatography with hexane: EtOAc (50:1) to yield (8-bromooctoxy)(*tert*-butyl)dimethylsilane (10.4518 g, 81%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 3.60 (t, *J* = 6.5 Hz, 2H), 3.40 (t, *J* = 6.7 Hz, 2H), 1.85 (quint, *J* = 7

Hz, 2H), 1.51 (m, 2H), 1.42 (m, 2H), 1.31 (m, 6H), 0.89 (s, 9H), 0.05 (s, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  63.4, 34.1, 33.0, 29.4, 29.0, 28.9, 28.3, 26.1, 25.9, 18.5, -5.1.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [98].

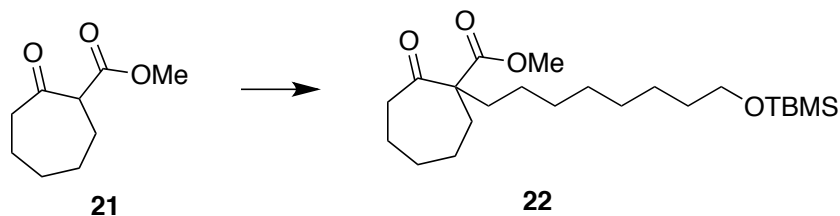


**(8-Iodooctoxy)(*tert*-butyl)dimethylsilane (20):** The protected bromooctanol **19** (10.4518 g, 32.34 mmol) was dissolved in acetone (86 mL). Potassium iodide (26.8401 g, 161.69 mmol, 5 equiv) was added and reaction was stirred at reflux for 18 h. After cooling to room temperature, the reaction mixture was filtered over Celite and the solvent was removed to yield (8-Iodooctoxy)(*tert*-butyl)dimethylsilane (10.7604 g, 89%), that was used as without further purification.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  3.60 (t,  $J = 6.5$  Hz, 2H), 3.19 (t,  $J = 7.3$  Hz, 2H), 1.82 (quint,  $J = 7$  Hz, 2H), 1.51 (m, 2H), 1.39 (m, 2H), 1.31 (m, 6H), 0.89 (s, 9H), 0.05 (s, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  63.2, 33.6, 32.8, 30.5, 29.2, 28.5, 26.0, 25.7, 18.4, 7.2, -5.3.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [98].

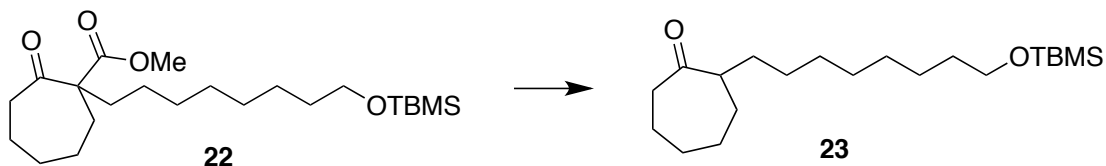


**Methyl 2-oxocycloheptanecarboxylate (21):** NaH (60% w/w suspension in mineral oil, 3.4763 g, 86.93 mmol, 1.9 equiv) was suspended in an ether (12 mL) solution of dimethyl carbonate (7.5 mL, 89.0 mol, 2 equiv). Cycloheptanone (5.3 mL, 45.0 mmol) was added dropwise to the suspension and allowed to stir at room temperature for 6 h. The solution was cooled in 0 °C and glacial acetic acid was added dropwise, followed by the slow addition of cold  $\text{H}_2\text{O}$ . The solution was extracted with diethyl ether and washed with brine (2 x 10 mL). The extract was dried with  $\text{MgSO}_4$ , filtered, and the solvent was

removed. The product was purified by Kugelrohr distillation to yield methyl 2-oxocycloheptanecarboxylate (2.1252 g, 28%).  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [99].

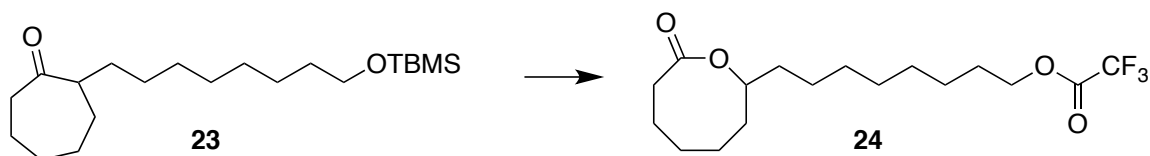


**Methyl 1-(8-(*tert*-butyldimethylsiloxy)octyl)-2-oxocycloheptane-1-carboxylate (22):** Methyl 2-oxocycloheptanecarboxylate (0.3737 g, 2.20 mmol, 1 equiv) and (8-iodooctoxy)(*tert*-butyl)dimethylsilane (20) (1.3176 g, 3.56 mmol 1.6 equiv) were dissolved in acetone (5.2 ml).  $\text{Ce}_2\text{CO}_3$  (2.1487 g, 6.59 mmol, 3 equiv) was added and the reaction was heated at reflux (66 °C) for 16 h. The reaction was cooled to room temperature, filtered through Celite, and the solvent was removed. The crude product was purified via column chromatography with hexane: EtOAc (5:1) to yield Methyl 1-(8-(*tert*-butyldimethylsiloxy)octyl)-2-oxocycloheptane-1-carboxylate (0.8914 g, 98%). IR (neat)  $\bar{\nu}$  2929, 2856, 1741, 1714, 1457, 1255, 1222, 1099, 836, 776;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  3.71 (s, 3H), 3.59 (t,  $J = 6.5$  Hz, 2H), 2.60 (m, 1H), 2.48 (m, 1H), 2.14 (m, 1H), 1.95 (m, 1H), 1.43-1.72 (m, 12H), 1.27 (m, 6H), 1.19 (m, 2H) 0.89 (s, 9H), 0.04 (s, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  209.8, 173.2, 63.4, 52.2, 42.2, 35.6, 30.5 33.0, 32.9, 30.14, 30.12, 29.5, 29.4, 26.1, 25.9, 25.8, 25.1, 24.7, 18.5, -5.1. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+$   $\text{C}_{23}\text{H}_{45}\text{O}_4\text{Si}$  413.3087, found 413.3097.

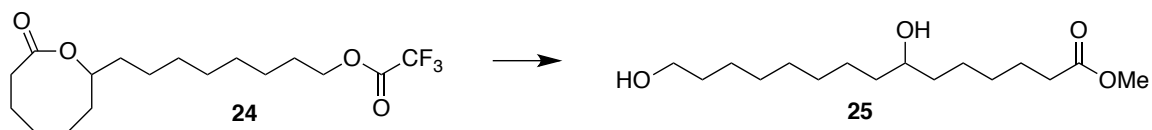


**2-(8-(*tert*-Butyldimethylsiloxy)octyl)cycloheptan-1-one (23):** Using a method modified from Carling [99], the cycloheptanone derivative 22 (0.8914 g, 2.16 mmol) was

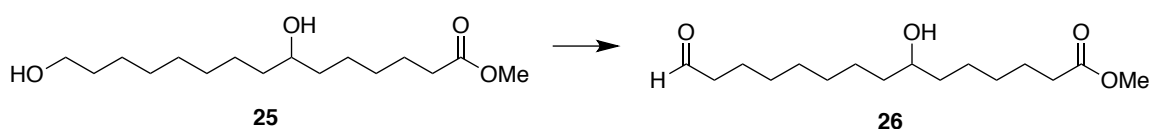
dissolved in DMSO (2.5 mL). Lithium chloride (0.2789 g, 6.58 mmol, 3 equiv) and water (58  $\mu$ L, 3.22 mmol, 1.5 equiv) were added. The reaction was heated at 120  $^{\circ}$ C for 2 h. The solution was cooled to room temperature, extracted with pentane, and washed with brine. The organic layer was dried with  $MgSO_4$ . The solvent was removed to yield 2-(8-(*tert*-butyldimethylsiloxy)octyl)cycloheptan-1-one (0.5638 g, 74%). IR (neat)  $\bar{\nu}$  2928, 2855, 1704, 1462, 1255, 1099, 835, 775;  $^1H$  NMR  $\delta$  (500 MHz,  $CDCl_3$ ) 3.59 (t,  $J$  = 6.8 Hz, 2H), 2.38-2.53 (m, 3H), 1.85 (m, 3H), 1.53-1.67 (m, 3H), 1.49 (quint,  $J$  = 6.8 Hz, 2H), 1.25-1.38 (m, 14H), 0.89 (s, 9H), 0.04 (s, 6H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$  216.7, 63.3, 52.4, 42.6, 32.9, 32.4, 31.3, 29.7, 29.6, 29.5, 29.4, 28.4, 27.2, 25.8, 25.6, 24.7, 18.4, -5.3.  $m/z$  75, 131, 201



**8-Trifluoroacetoxyoctyl-2-oxocanone (24):** The TBDMS-protected 2-(8-*tert*-butyldimethylsiloxyoctanyl)cycloheptanone **23** (0.0403 g, 0.11 mmol) was dissolved in DCM (1.5 mL). Urea hydrogen peroxide complex (0.1069 g, 1.14 mmol, 10 equiv) was added followed in dropwise fashion by trifluoroacetic anhydride (40  $\mu$ L, 0.28 mmol, 2.4 equiv). The reaction was stirred at room temperature overnight before being quenched with saturated  $NaHCO_3$ . DCM (5 mL) was added and the organic layer was extracted. The solvent was removed. The product was purified via column chromatography with hexane: EtOAc (20:1) to yield 8-trifluoroacetoxyoctyl-2-oxocanone (0.0215 g, 54%). IR (neat)  $\bar{\nu}$  2932, 2859, 1786, 1729, 1461, 1403, 1352, 1221, 1166, 1006, 777, 731 ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$  4.53 (m, 1H), 4.34 (t,  $J$  = 6.7 Hz, 2H), 2.43-2.56 (m, 2H), 1.90 (m, 1H), 1.82 (m, 1H), 1.55-1.78 (m, 6H), 1.40-1.54 (m, 4H), 1.22-1.40 (m, 10H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$  176.9, 157.6 (q,  $J$  = 41.8 Hz), 114.5 (q,  $J$  = 283.7 Hz), 78.8, 68.2, 37.3, 35.6, 32.5, 31.6, 29.2, 28.9, 28.8, 28.1, 26.4, 25.7, 25.4, 23.9.  $m/z$  98, 127

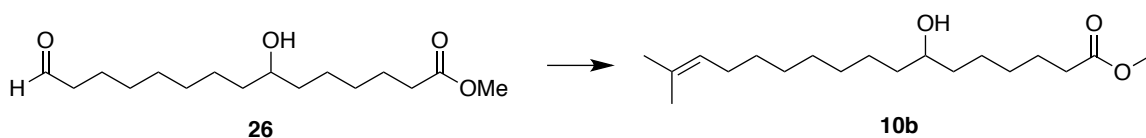


**Methyl 7,15-dihydroxypentadecanoate (25):** A solution of trifluoroacetate **24** (0.1778 g, 0.50 mmol) in MeOH (0.5 mL) was prepared to which a solution of 24% tetramethylammonium hydroxide (0.5 mL) in methanol was added; the reaction was stirred at room temperature for 18 h. To this crude solution of the intermediate carboxylate salt, dimethylformamide (DMF) (5 mL) and methyl iodide (0.5 mL) were added and the mixture was then stirred until a precipitate formed (~20 min). The product was extracted from the reaction milieu with ethyl acetate and the collected organic layers were washed with water:brine (1:1) several times until all the DMF was removed. The product was purified via column chromatography with hexane: EtOAc (3:1) to yield methyl 7,15-dihydroxypentadecanoate (0.0871 g, 60%). IR (neat)  $\bar{\nu}$  3358, 2929, 2855, 1739, 1462, 1439, 1353, 1260, 1199, 1172, 1109, 1056, 883, 801, 723, 672;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  3.67 (s, 3H), 3.64 (t,  $J = 6.7$  Hz, 2H), 3.57 (m, 1H), 2.31 (t,  $J = 7.5$ , 2H), 1.64 (quint,  $J = 7.4$  Hz, 2H), 1.56 (quint,  $J = 6.9$ , 2H), 1.27-1.45 (m, 12H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  174.3, 71.9, 63.1, 51.5, 37.5, 37.2, 34.0, 32.8, 29.6, 29.5, 29.3, 29.2, 25.7, 25.6, 25.3, 24.9.



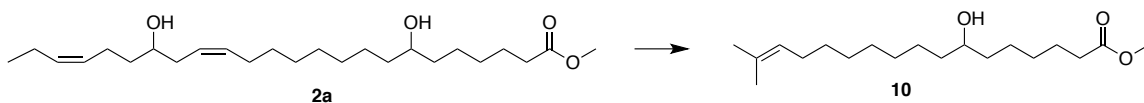
**Methyl 7-hydroxy-15-oxopentadecanoate (26):** Methyl 7,15-dihydroxypentadecanoate (0.0404 g, 0.14 mmol) was dissolved in DCM (1.2 mL). TEMPO (0.0022 g, 0.014 mmol, 0.1 equiv) and (diacetoxyiodo)benzene (0.0519 g, 0.16 mmol, 1.15 equiv) were added to the mixture. The reaction stirred at room temperature for 2.5 h. The mixture was diluted with DCM (5 mL) and the reaction was then quenched with saturated  $\text{Na}_2\text{S}_2\text{O}_3$ . The solution was extracted with DCM and washed with brine. The organic layer was dried with  $\text{Na}_2\text{SO}_4$  and the solvent was removed. The product was purified via column chromatography with hexane: EtOAc (3:1) to yield methyl 7-

hydroxy-15-oxopentadecanoate (0.0313 g, 73%). IR (neat)  $\bar{\nu}$  3446, 2930, 2855, 1734, 1700, 1684, 1653, 1559, 1540, 1521, 1456, 1436, 1418, 1172, 1012, 668;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (t,  $J = 2$  Hz, 1H), 3.67 (s, 3H), 3.57 (m, 1H), 2.42 (td,  $J = 7.4, 1.8$  Hz, 2H), 2.31 (t,  $J = 7.5$  Hz, 2H), 1.64 (m, 4H), 1.37-1.48 (m, 6H), 1.29-1.36 (m, 10H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  202.8, 176.0, 71.8, 51.5, 43.9, 37.5, 37.3, 34.0, 29.4, 29.3, 29.2, 29.1, 25.5, 25.3, 24.9, 22.0.

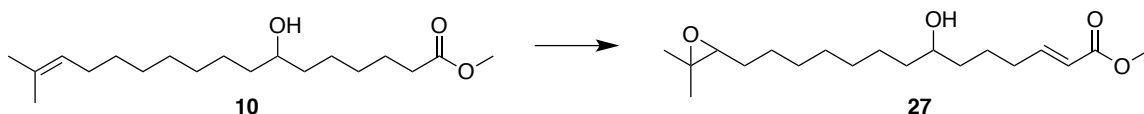


**Methyl (*E*)-7-hydroxy-16-methylheptadeca-2,15-dienoate (10b):** Isopropyl triphenylphosphonium iodide (0.0996 g, 0.230 mmol, 2.1 equiv) was suspended in THF (0.95 mL). The solution was cooled to 0 °C and *n*-butyllithium (0.12 mL, 0.234 mmol, 2.15 equiv) was added dropwise. The dark red solution was cooled to -78 °C with a dry ice/acetone bath. Aldehyde **26** (0.0313, 0.11 mmol) dissolved in THF (0.3 mL) was added slowly. The solution was stirred in the acetone bath slowly warming for 1 h to 0 °C prior to being quenched with saturated  $\text{NaHCO}_3$ . The THF layer was isolated. The aqueous layer was then extracted with diethyl ether, the organic layers were combined, and the solvent was removed. The product was purified via column chromatography with hexane: EtOAc (10:1 to 5:1 to 3:1) to yield methyl (*E*)-7-hydroxy-16-methylheptadeca-2,15-dienoate (0.0056 g, 16%). IR (neat)  $\bar{\nu}$  420, 2927, 2854, 1741, 1456, 1436, 1375, 1258, 1199, 1172, 1105, 1014;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  5.12 (t,  $J = 6.7$  Hz, 1H), 3.66 (s, 3H), 3.57 (m, 1H), 2.31 (t,  $J = 7.5$  Hz, 2H), 1.95 (m, 2H), 1.68 (s, 3H), 1.63 (m, 2H), 1.59 (s, 3H), 1.28-1.46 (m, 18H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  174.2, 131.1, 124.9, 71.9, 51.4, 37.5, 37.2, 34.0, 29.9, 29.7, 29.5, 29.3, 29.2, 28.0, 25.7, 25.6, 25.3, 24.9, 17.6. *m/z* 312, 294, 159, 69



2.2.7 Configuration of C-7 in 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup>

**Methyl 7-hydroxy-16-methylheptadec-15-enoate (10):** Ester **2a** (0.2349 g, 0.572 mmol) was dissolved in a dry CH<sub>2</sub>Cl<sub>2</sub>:2-methyl-2-butene mixture (1:1, 10 mL). The solution was sparged with nitrogen before second-generation Grubbs catalyst (0.0195 g, 0.022 mmol, 0.04 equiv) was added. The solution was stirred at room temperature for 16 h under nitrogen. Activated carbon was added to absorb the catalyst; the suspension was stirred for 10 min before the mixture was filtered through silica gel, rinsing the packing with CH<sub>2</sub>Cl<sub>2</sub>: Et<sub>2</sub>O (1:1). The solvent was removed and the crude product was purified by SiO<sub>2</sub> flash chromatography with hexane: EtOAc (3:1) to yield **10** (0.1251 g, 70%). Spectral data matches previous reported from earlier in the materials and methods above.



**Methyl (*E*)-14-(3,3-dimethyloxiran-2-yl)-7-hydroxytetradec-2-enoate (27):** Sodium bis(trimethylsilyl)amide (0.6 M in THF, 2.6 mL, 1.535 mmol, 6.2 equiv) was cooled to -78 °C with an acetone/dry ice bath. Ester **10** (0.0774 g, 0.2476 mmol) was azeotropically dried twice with benzene before being dissolved in THF (1 mL) and added to the cooled sodium bis(trimethylsilyl)amide solution. The reaction mixture was stirred for 30 minutes at -78 °C and then a solution of phenylselenenyl chloride (0.0584 g, 0.3049 mmol, 1.2 equiv) in THF (1 mL) was added. After stirring for 1 h at -78 °C and then slowly being warmed to room temperature, the reaction was quenched with saturated NH<sub>4</sub>Cl and extracted with diethyl ether. The organic extract was dried with MgSO<sub>4</sub>, filtered, and the solvent was removed by rotary evaporation. The crude α-phenylseleno intermediate was used in the next step without further purification.

The  $\alpha$ -phenylseleno intermediate was dissolved in DCM (5 mL) that had been passed through pipet column of  $\text{Al}_2\text{O}_3$  prior to use. The solution was cooled in an ice bath and 10%  $\text{H}_2\text{O}_2$  (2.2 mL) added dropwise by pipet. The ice bath was removed and solution stirred for 1 h before being quenched with saturated  $\text{NaHCO}_3$ . The solution was extracted with DCM and the organic layer was washed with brine, dried with  $\text{MgSO}_4$ , filtered, and the solvent was removed with a rotary evaporator. The residual oil was purified by  $\text{SiO}_2$  flash chromatography with hexane: EtOAc, 3:1 to yield **27** (0.0185 g, 25%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.7, 7$  Hz, 1H), 5.82 (dt,  $J = 15.6, 1.5$  Hz, 1H), 3.71 (s, 3H), 3.57 (m, 1H), 2.70 (t,  $J = 6.0$  Hz, 1H), 2.22 (m, 2H), 1.24-1.53 (m, 18H), 1.30 (s, 3H), 1.25 (s, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  167.3, 149.4, 121.3, 71.8, 64.7, 58.3, 51.5, 37.7, 37.0, 32.3, 29.69, 29.65, 29.5, 29.0, 26.7, 25.7, 25.0, 24.3, 18.9.

**Mosher Ester Analysis:** This procedure was based on the work reported by Hoye [100]. Chloroform used in the reaction was purified by passage through a pipet column with basic  $\text{Al}_2\text{O}_3$ . Stock solutions were made for all reagents to facilitate reagent handling; their concentrations were approximately 0.2 M starting material, 0.9 M 4-dimethylaminopyridine (DMAP), and 0.4 M for *N, N'*-dicyclohexylcarbodiimide and either (*R*)- or (*S*)- $\alpha$ -methoxy- $\alpha$ -trifluoromethylphenylacetic acid. Two reactions were run in parallel: the starting alcohol (*ca.* 5 mg) and one of the (*R*)- or (*S*)- $\alpha$ -methoxy- $\alpha$ -trifluoromethylphenylacetic acids (3.1 equiv) were stirred together in a small oven-dried vial. *N, N'*-Dicyclohexylcarbodiimide (3.1 equiv) was added followed by 4-dimethylaminopyridine (1 equiv). The mixture was allowed to stir at room temperature overnight. As an initial cleanup, the esterification mixture was filtered through a paper-plugged pipet and the solvent was removed. The crude product was purified by  $\text{SiO}_2$  flash chromatography with hexane: EtOAc, 10:1 ( $R_f = 0.38$ ).  $^1\text{H}$  and  $^{19}\text{F}$  NMR spectra were collected of the esters; relevant data is described in the Results and Discussion section.

## 2.3 Results and Discussion

### 2.3.1 Serendipitous discovery of novel oils in *O. violaceus* seeds

Thin-layer chromatography (TLC) separation of neutral lipids from the seeds of several Brassicaceae species was conducted as part of a study to assess correlations between triacylglycerol fatty acid composition and vitamin E content of seeds. *O. violaceus* was one species chosen because previous reports indicated that had an unusual seed oil, enriched in linoleic acid (18:2) and low in VLCFAs [35, 36]. Unexpectedly, *O. violaceus* seed extracts lacked a typical triacylglycerol (TAG) band on a silica TLC matrix using a neutral lipid solvent system (Figure 11A). Abundant bands were instead detected with I<sub>2</sub> near the origin of the TLC plate, suggesting that the oil contains fatty acids with greater polarity than those of fatty acids normally found in seeds of most plants. Subsequent TLC separations using solvent systems with increased polarity revealed that *O. violaceus* seed oil had even less mobility than castor bean oil, which is enriched in TAG with three monohydroxy fatty acids (ricinoleic acid) (Figure 11B).

Initial attempts to analyze the composition of *O. violaceus* seeds by GC indicated a fatty acid composition with a high 18:2 content, similar to published results [35, 36], but no apparent novel fatty acids. Fatty acid profiles nearly identical to those of published results [35, 36] were observed even after silylation to enhance the mobility of potentially hydroxylated fatty acid methyl esters (FAMES) (Figure 11C). It was not until GC run times were extended beyond those required for resolution of C<sub>24</sub> fatty acids that two abundant peaks (designated Peaks 1 and 2) eluted (Figure 11C). These apparent FAMES **1a** and **2a**, respectively, in a ratio of ~1:3 accounted for ~50 wt% of the total fatty acids. The TLC-based oil profile and novel fatty acids were observed in multiple independent accessions of *O. violaceus* seeds.

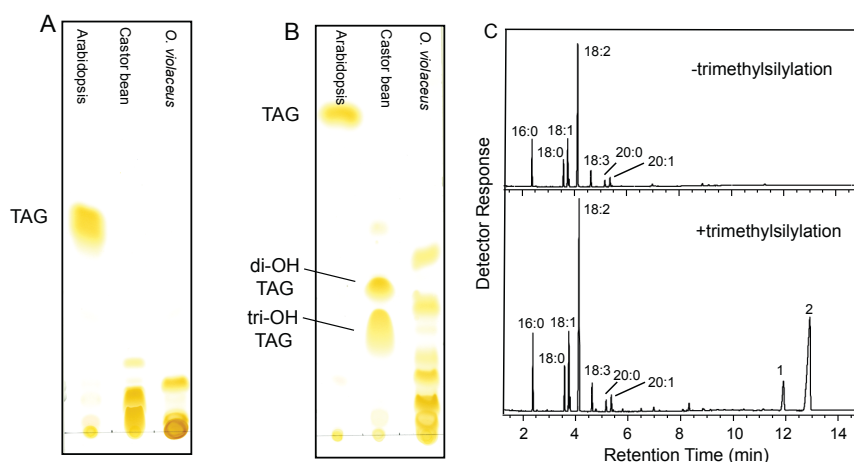


Figure 11: Analysis of *O. violaceus* seed oil. *A, B*. TLC analyses of *O. violaceus* seed oil in solvent systems of different polarities. TLC plates were developed in heptane:ethyl ether:acetic acid at ratios of 60:40:1 (v/v) in *A* and 40:60:1 (v/v) in *B*. Shown for comparison are TLC analyses of Arabidopsis seed triacylglycerol (TAG), which lacks hydroxy fatty acids, and castor bean TAG, which is enriched in ricinoleic acid (OH-18:1), a monohydroxy fatty acid, and TAG species containing two (di-OH TAG) and three (tri-OH-TAG) monohydroxy fatty acids. *C*. GC analyses using an extended run time of fatty acid methyl esters prepared from *O. violaceus* seeds without (-trimethylsilylation) or with trimethylsilyl derivatization (+trimethylsilylation). The novel silylated FAMES were initially named Peak 1 and Peak 2 consistent with their elution order.

### 2.3.2 Structural identification of novel fatty acids from *O. violaceus* seeds

GC-MS and NMR studies were undertaken to determine the chemical structures of the two abundant novel fatty acids. Using a simple TLC method, FAMES prepared from *O. violaceus* seeds were separated, and the novel FAMES that displayed much higher polarities than typical FAMES were purified from the TLC matrix. GC-MS of the trimethylsilylated (TMS) novel FAMES indicated that fragmentation of the Peak 1 analyte ((TMS)<sub>2</sub>-**1a**) yielded ions identical to those corresponding to the ω-6 portion of the silylated methyl ester of ricinoleic acid, and fragmentation of the derivative in Peak 2 ((TMS)<sub>2</sub>-**2a**) generated a subset of ions differing by only two fewer mass units from the Peak 1 component as well as ions identical to those corresponding to the ω-6 portion of the methyl ester of silylated densipolic acid (Figure 12). The difference in two mass units indicated that the Peak 2 FAME has an additional double bond, most likely in the ω-3

position, relative to the Peak 1 FAME. Apart from methyl-proximate fragments of the acyl chain that were identical to those formed from  $\omega$ -6 hydroxylated fatty acids, the overall fragmentation patterns of silylated **1a** and **2a** suggested that the remainder of their structures is considerably larger than those of the known C<sub>18</sub> and C<sub>20</sub> monohydroxy fatty acids.

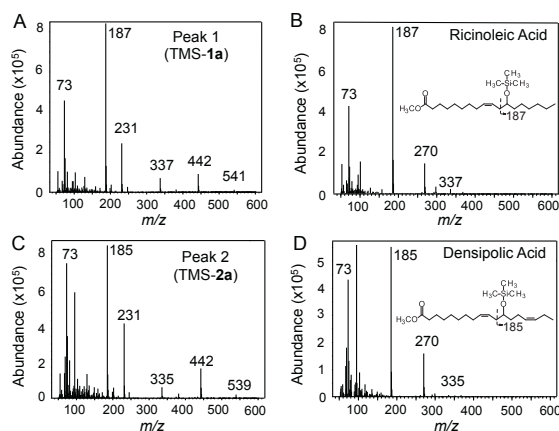


Figure 12: GC-MS analysis of unusual fatty acids in *O. violaceus* seeds. Mass spectra of trimethylsilylated (TMS) fatty acid methyl esters for Peaks 1 (TMS-**1a**) (A) and 2 (TMS-**2a**) (C) in the gas chromatogram in Figure 15C (lower panel). For comparison, mass spectra are shown for TMS-methyl esters of ricinoleic acid (B) and densipolic acid (D).

As shown, TMS-**1a** and TMS-**2a** contain ions consistent with fragmentation products containing  $\omega$ -6 hydroxyl groups in mass spectra of TMS-methyl ricinoleic ( $m/z$  187) and densipolic ( $m/z$  185) acids, respectively.

A similar approach to the initial purification was used to produce the methyl ester of **2a** on a larger scale. FAMES were prepared from *O. violaceus* seed oil. The crude material was purified by column chromatography with a solvent system of hexane: diethyl ether: acetic acid (40:60:1). The purified compound was found at an  $R_f$  of 0.25; removal of solvent allowed for further analysis. The methyl ester **2a** provided a quasimolecular ion  $[M + H]^+$  with an  $m/z$  of 411.3472 consistent with the molecular formula C<sub>25</sub>H<sub>46</sub>O<sub>4</sub> (calculated  $m/z$  411.3469). M-18 and M-36 water loss peaks in the ESI spectrum suggested the presence of two alcohols. A functional group survey by IR spectroscopy indicated an OH stretching band at 3386 cm<sup>-1</sup>, alkene C-H, C-C and out-of-plane bending stretches at 3005, 1653 cm<sup>-1</sup> and 724 cm<sup>-1</sup>, respectively, a carbonyl stretch consistent with an ester at 1740 cm<sup>-1</sup>, and C-O stretches at 1199, 1171, and 1066 cm<sup>-1</sup> for

methyl esters and secondary alcohols. The  $^{13}\text{C}$  NMR and DEPT-135 spectra revealed 25 different carbons (Table 1) including one quaternary, six methine, sixteen methylene, and two methyl groups. The existence of two alkenes was established by the four  $^{13}\text{C}$  vinyl C-H resonances at 133.5, 132.2, 128.5 and 125.0 ppm. With three oxygen-bearing carbons at  $\delta$  51.4, 71.1, and 71.8, and a carbonyl peak at 174.2 ppm, a formula of  $\text{C}_{25}\text{H}_{46}\text{O}_4$  was confirmed.

Table 1: NMR data for methyl wuhanate **2a**

Carbon #	$\delta^{13}\text{C}$	$\delta^1\text{H}$ (m)	$J$ (Hz)	DEPT	$^1\text{H}$ - $^1\text{H}$ COSY Correlations	$^{13}\text{C}$ - $^1\text{H}$ HMBC Correlations
1'	51.4	3.66 s	-	CH <sub>3</sub>	H2w	
1	174.2	-	-	C		H1', H2, H3
2	34.0	2.31 t	7.5	CH <sub>2</sub>	H3	H3
3	24.9	1.63 qn	7.5	CH <sub>2</sub>		H4
4	25.6	1.40 m; 1.30 m	-	CH <sub>2</sub>		H3, H5
5	25.3	1.42 m; 1.32 m	-	CH <sub>2</sub>		
6	37.2	1.37-1.47 m	-	CH <sub>2</sub>	H5	H5
7	71.8	3.57 m	-	CH	H8, H6	
8	37.5	1.37-1.47 m	-	CH <sub>2</sub>		
9	29.6	1.30 m	-	CH <sub>2</sub>		
10	29.4	1.30 m	-	CH <sub>2</sub>		
11	29.3	1.30 m	-	CH <sub>2</sub>		
12	29.19	1.30 m	-	CH <sub>2</sub>		
13	29.15	1.30 m	-	CH <sub>2</sub>		
14	27.4	2.04 m	-	CH <sub>2</sub>	H17, H15	H15, H16
15	133.5	5.56 dtt	11, 7.5, 1,63	CH	H17, H14	H14, H16, H17
16	125	5.39 m	-	CH	H17, H14	H14, H15, H17, H18
17	35.4	2.22 t	6.5	CH <sub>2</sub>	H14	H15, H16, H19
18	71.1	3.63 m	-	CH	H19, H17	H16, H19, H17, H20
19	36.7	1.53 m	-	CH <sub>2</sub>	H18, H20	H21, H17, H20
20	23.5	2.16 m; 2.13 m	-	CH <sub>2</sub>	H19	H22, H21, H18, H19
21	128.5	5.34 m	-	CH	H23, H20	H19, H20, H23
22	132.2	5.37 m	-	CH	H23, H20	H20, H23, H24
23	20.5	2.05 m	-	CH <sub>2</sub>	H24	H22, H21, H24
24	14.3	0.96 t	7.5	CH <sub>3</sub>	H23	H22, H23

Initial substructures for the predicted linear dihydroxylated unsaturated ester were assembled from the COSY and HMBC data. C1 and C1' were assigned based on their chemical shifts, where C1 corresponded to the carbonyl and C1' represents the CH<sub>3</sub> carbon of the methyl ester. Long-range coupling of H1' to H2 and <sup>3</sup>J<sub>HH</sub> coupling between H2 to H3 was visible in the <sup>1</sup>H-<sup>1</sup>H COSY. An H4 to C3 correlation in the HMBC extended the spin system (Figure 13A). The COSY furnished the pairs of methylenes surrounding each alcohol generating two additional fragments (Figure 13B,C). The correlations from the COSY and HMBC spectra identified the sets of vinylic carbons and protons; C15 and C16 are the carbons of one alkene and C21 and C22 are the carbons of the other. The connectivity in the C21-C22 alkene with could be expanded using coupling data in both directions along the alkyl chain (Figure 13D). The COSY spectrum showed H22 correlating to H23 and H23 coupled to H24. The HMBC further confirmed the ω end of the alkyl chain with a crosspeak of C22 to H24. The other vinylic hydrogen H21 correlated to H20, and C21 displayed long-range coupling to H19. For the other alkene functionality, both C15 and C16 correlate to H14 and H17 with crosspeak intensity favoring H17 being allylic to C16 (Figure 13E). <sup>n</sup>J<sub>CH</sub> correlations to the carboxyl-proximate side of H14 from C15 could not be unambiguously assigned. Fragment C connected fragments D and E together; C19 from fragment C appeared in the correlations observed for fragment D and C17 in fragment C is also present in the set of crosspeaks observed with fragment E. Although the methylene resonances for C9-C13 could not be fully assigned due to the low chemical shift dispersion in the range of 1-2 ppm, only one supported connectivity appeared to be possible, which is depicted in Figure 13F. Nevertheless, the exact position of the second hydroxyl position could only be hypothesized due to the similarity in chemical shifts of the alkyl chain.



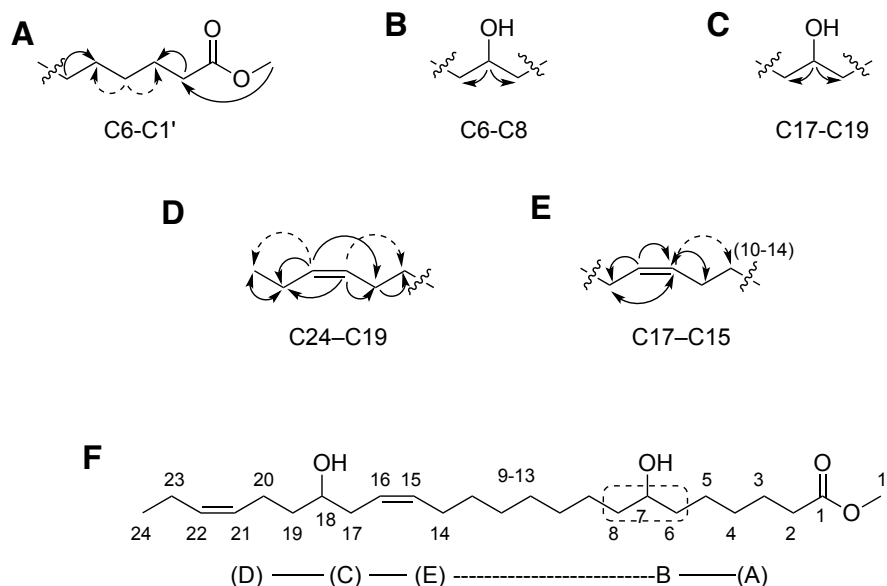


Figure 13: Substructures based on NMR correlations and coupling data. *A-E* Assigned nuclei in the five working structures were constructed using the  $^1\text{H}$ - $^1\text{H}$ -COSY (solid-line arrows) and  $^1\text{H}$ - $^{13}\text{C}$  HMBC heteronuclear correlations (dashed-line arrows) showing  $^1\text{H}$  correlations to  $^{13}\text{C}$  in Table 2 resulting in structure *F*. Signals for C9-13 were not sufficiently dispersed for their assignment. The dotted box indicates possible locants for the hydroxyl group.

The low NMR chemical shift dispersion for eight of the  $\text{CH}_2$  groups between the alkene and the carboxyl group impeded our development of an incontrovertible case for the structure. Both the unsilylated electron impact (EI) mass spectrum for **2a** (Figure 14A) and disilylated (Figure 14B) FAMES for **2a**, as well as **1a** and 5,16-OH-22:1 $\Delta^{13}$  (Figure 14D, E) were found to have anticipated fragments, as well as distinctive ions attributed to novel ene and rearrangement processes that aided in determining the correct connectivity shown in Figure 13F. Figure 14C shows  $\gamma$ -hydrogen atom transfer and  $\text{H}^+$  or  $\text{TMS}^+$  migration to the carboxyl group of pyrrolidide and ester derivatives may result in a molecular ion **3**, which subsequently loses formaldehyde. This is postulated to give rise to the 29 mass unit formal “CHO” loss from the  $\beta$ -cleavage ions of hydroxy fatty acids, such as **4**.

Figure 14: EI-MS spectra of FAMES. *A.* Mass spectrum of **2a** with hypothetical fragmentation pathways shown in the inset. The dihydroxy ester was observed to fragment through two major channels: dehydration and  $\beta$ -scission producing M-1, M-129, and M-233 fragments or the ene fragmentation followed by the losses of H<sub>2</sub>O and MeOH. *B.* Mass spectrum of (TMS)<sub>2</sub>-**2a** with the origins of major ions annotated on the inset structure. The trimethylsilyl groups block the retro-ene reaction resulting in cleavage  $\beta$  to the ether oxygen, neutral loss reactions, and “deformylation” ions. An *m/z* 74 McLafferty cleavage ion was readily apparent. *C.* A  $\gamma$ -hydrogen atom transfer and H<sup>+</sup> or TMS<sup>+</sup> migration to the carboxyl group of pyrrolidide and ester derivatives. *D.* Mass spectrum for the methyl ester of 7,18-(TMSO)<sub>2</sub>-24:1 $\Delta$ <sup>15</sup>. *E.* Mass spectrum for the methyl ester of 7,18-(TMSO)<sub>2</sub>-22:1 $\Delta$ <sup>15</sup>



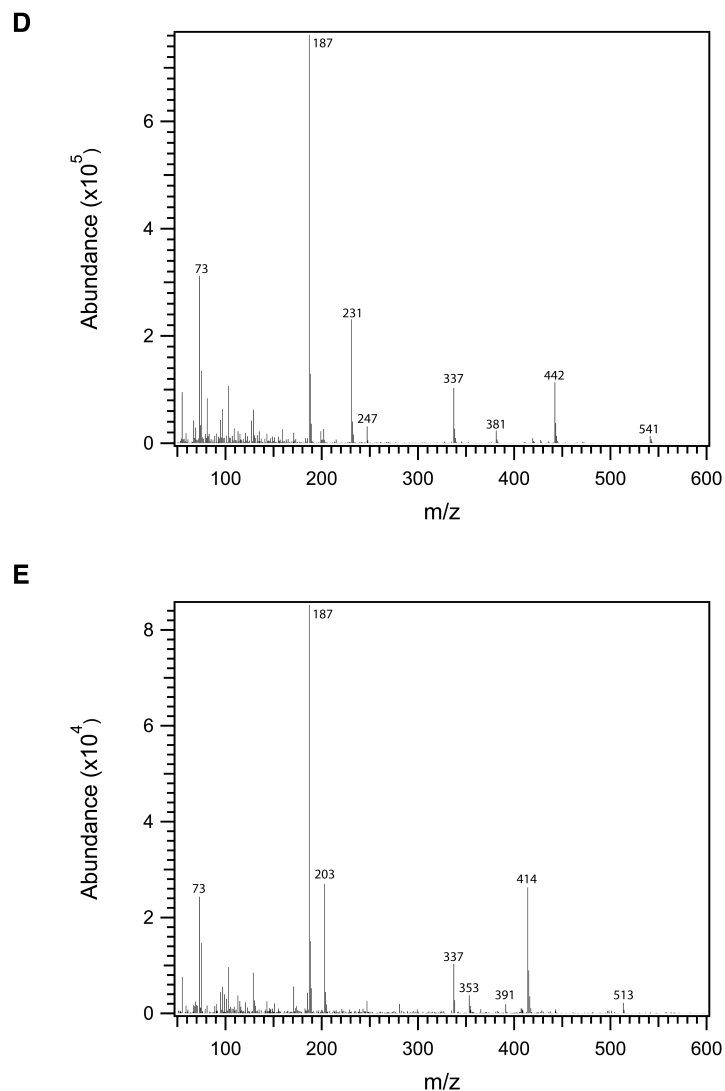


Figure 14 continued

In an attempt to rigorously situate the hydroxy groups on the backbone, the two alcohol groups in **2a** were oxidized with Dess-Martin periodinane to produce **5**. Upon oxidation, the expected formula for the ketone was confirmed (+ESI mode:  $C_{25}H_{43}O_4$  observed 407.3154, calculated  $m/z$  407.3156; -ESI mode:  $C_{25}H_{41}O_4$  observed 405.2979, calculated  $m/z$  405.3005). Both alcohol multiplets at 3.57 and 3.63 ppm disappeared in the  $^1H$  NMR leading to a new resonance at  $\delta$  3.14 for protons alpha to the  $\beta,\gamma$  – unsaturated ketone and a doubled IR stretch at 1738 and 1715  $cm^{-1}$ , the latter due to the appearance of an unconjugated ketone. Finally, weak McLafferty fragmentation ions were observed at  $m/z$  172 and 338 that first suggested the hydroxyl groups were at C-7

and C-18 (Figure 15). Other NMR data supported the structure of the ketone but the data was still deemed insufficient to finalize the structure.

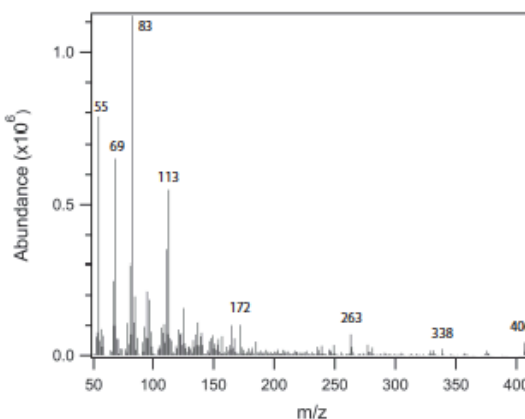


Figure 15: Mass spectra of diketone from methyl wuhanate (**5**).

Three chemical methods finally ensured the unambiguous assignment of the connectivity: pyrrolidide derivatization, von Rudloff degradation, and alkene metathesis. To determine the position of the hydroxyl groups, **2a** was incubated with pyrrolidine and glacial acetic acid to form **6**. The MS for **6** exhibited two homologous series incremented by  $\text{CH}_2$  groups (Figure 16A). The first series for the diol covered the range from  $m/z$  126 to 168. Following a major  $m/z$  198 fragment resulting from a  $\beta$ -cleavage relative to the 7-hydroxyl group, each ion in the  $m/z$  180 to 264 (C8-13) series included a water loss. Beyond  $m/z$  264, gaps of 26 amu localized the double bonds. Bis-trimethylsilylation of **2b** activated the ethers toward  $\beta$ -scission and provided intense ions neighboring each oxygenated carbon (Figure 16C). Collectively, the MS fragmentation indicated that the alcohols were at positions 7 and 18.

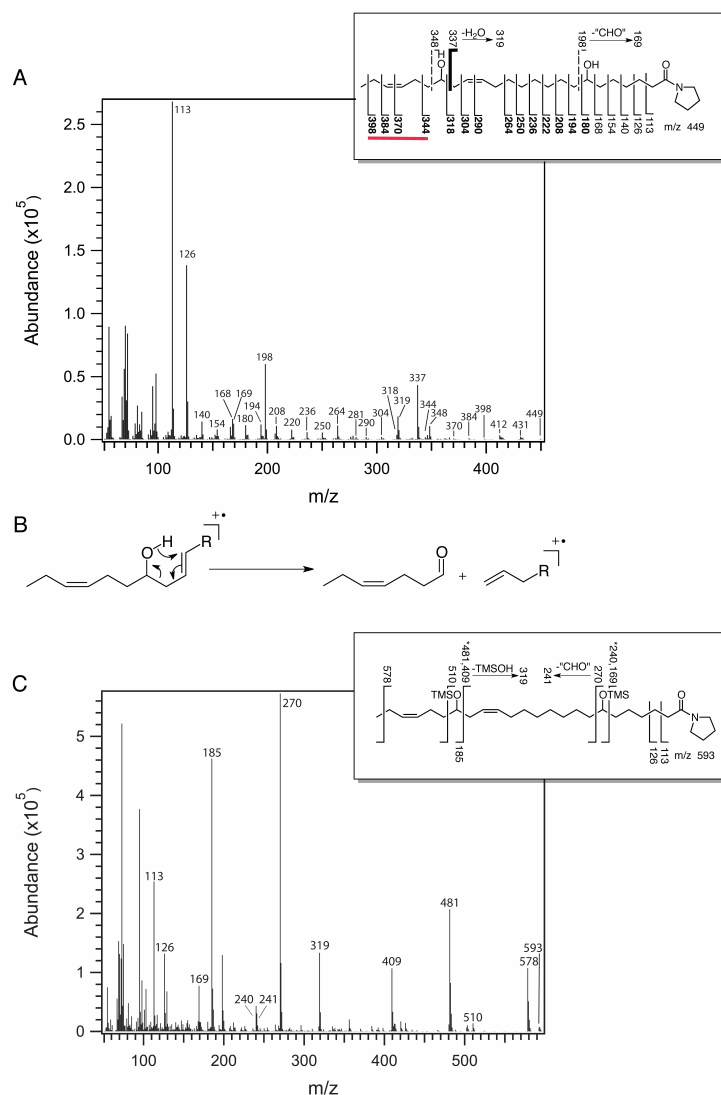


Figure 16: Electron-impact (EI) MS spectra of fatty acid pyrrolidides. *A*. Mass spectrum of the pyrrolidide of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup> (**6**). Correspondence between the proposed fragments and observed ions is shown in the inset. The even-mass ions below the structure are proposed to result from the sequential chain scission of **6**. The *m/z* 113 ion is consistent with a McLafferty cleavage. Ions marked in bold font result from the dehydration from one or, as denoted by the bold underline, two hydroxy groups. Ions above the structure are derived from cleavage processes beta to the oxygen and subsequent chemistry. *B*. The scission of the hydroxylated amide is consistent with a process reminiscent of an alkene-aldehyde ene reaction leads to the bold line fragmentation in *A*. *C*. Contrastingly, the dominant ions from (TMS)<sub>2</sub>-**6** are formed disproportionately by beta-cleavage chemistry. Ions 510 and 270 are distinguished by undergoing the formal loss of CHO.

To clarify the location of the carboxyl-proximate hydroxyl group in the acyl chains of **1a** and **2a**, a von Rudloff oxidation was conducted to cleave the FAMES at their double bonds. The oxidation products for Peak 1 (**1a**) and Peak 2 (**2a**) were consistent with cleavage at the  $\omega$ -9 (or  $\Delta^{15}$ ) double bond of the C<sub>24</sub> FAME. The mass spectra of one of the products following disilylation **7** or esterification + silylation **8** fragmented in a manner expected for a 7-hydroxyl group (Figure 17A). Both spectra contained the  $m/z$  231 ion expected for fragmentation at the site of hydroxylation for the “front-end” 7-hydroxylated product. The “back-end” fragment differed by 58 mass units depending on whether the terminal carboxyl group is methylated ( $m/z$  259) or trimethylsilylated ( $m/z$  317); both ions from the von Rudloff cleavage placed the first alkene or double bond at C15. Peak 2 (**2a**) had additional oxidation products due to the second alkene present. In the mass spectra of one of the products for **2a** had a parent  $m/z$  of 528, indicative of  $\Delta^{21}$  or  $\omega$ -3 double bond (Figure 17B).

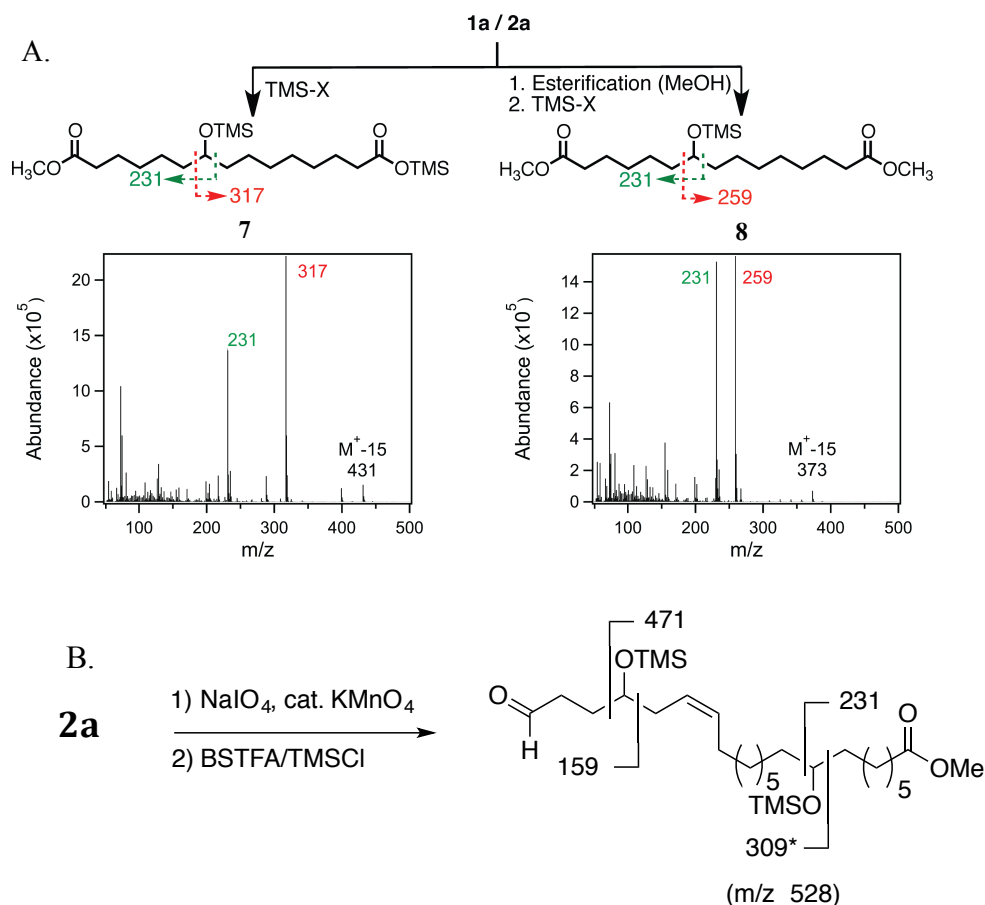


Figure 17: Results from iodate-permanganate “von Rudloff” oxidation of **1a** and **2a**. *A.* The  $\Delta^{15}$  bonds in **1a** and **2a** were cleaved. An  $m/z$  231 fragment diagnostic of a 7-hydroxy fatty acid was observed during GC/MS analysis together with an  $m/z$  317 fragment, following direct silylation to the trimethylsilyl ester, or an  $m/z$  259 ion when methylation preceded silylation. *B.* Additional fragmentation of **2a** due to the additional alkene present.

### 2.3.3 Configuration at the C18 stereocenter of 7,18-(OH)<sub>2</sub>-24:2 $\Delta^{15,21}$

All attempts to produce Mosher and *O*-methylmandelate derivatizations of the *O. violaceus*-derived FAMES in an effort to assign the configurations of the hydroxylated carbons led to decomposition. As an alternative approach, FAME **2a** was converted by alkene cross-metathesis with 2-methyl-2-butene [101] into two products fully consistent with the sites of hydroxylation and desaturation above, 3-cyclohexen-1-ol (**6a**) and 7-hydroxy-16-methyl-15-heptadecenoate (**7a**) (Figure 18).



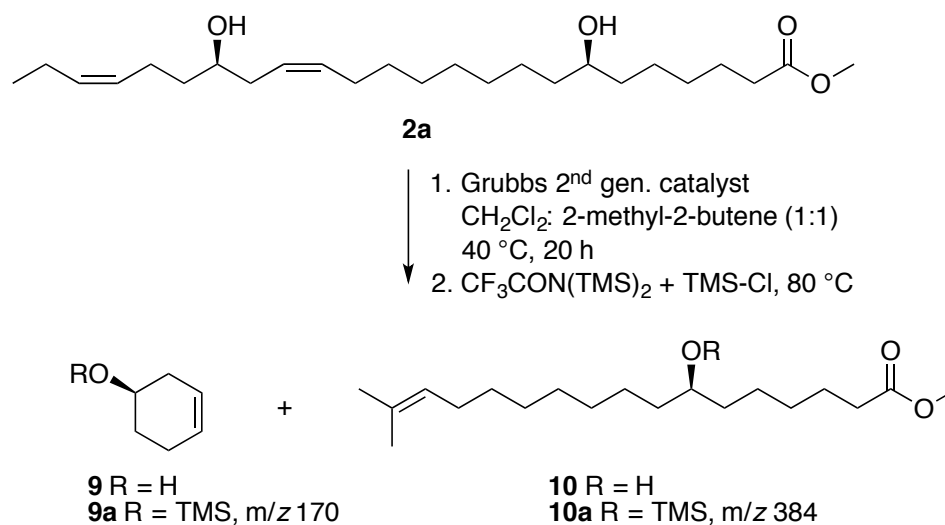


Figure 18: To assign the 18-hydroxy stereocenter, alkene metathesis of **2a** was performed in 2-methyl-2-butene. Ring-closing metathesis (RCM) produced in 3-cyclohexenol (**9**) together with the C1-15 metathesis product (**10**) which were then trimethylsilylated with BSTFA-TMSCl.

Following silylation of the mixture, cyclohexenol-derived **9a** was identifiable from its dominant retro-Diels Alder fragment (*m/z* 116). This product was found to be identical to a synthetic standard based on its retention time and ion profile. Resolutions of the racemic and a sample of enantioenriched **9a**, prepared by an enantioselective reduction, were achieved by chiral GC/MS (Figure 19). The ring-closing metathesis product coeluted with the (*R*)-isomer of **9a** at 16.85 min, 0.14 min after the (*S*)-silyl ether, and permitted the assignment of the (18*R*)-stereocenter.

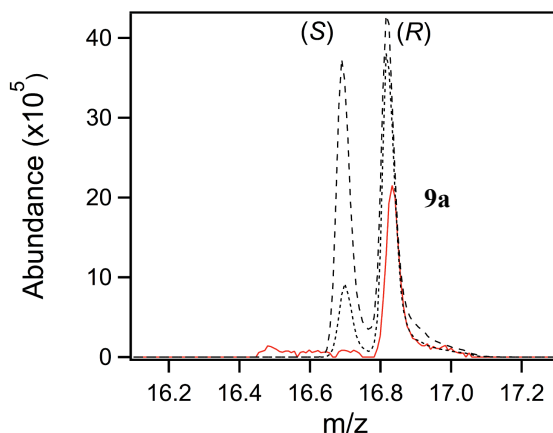


Figure 19: Comparison of TMS-modified enantioenriched (*R*)-(dotted line) and racemic (black dashes) cyclohexenol with the RCM product from natural **2a** (red solid line) defines the stereochemistry as *R*.

#### 2.3.4 Configuration at the C7 stereocenter of 7,18-(OH)<sub>2</sub>-24:2Δ<sup>15,21</sup>

Trying to take a similar approach for C7 proved more difficult than C18. A couple of synthetic pathways were attempted to synthesize **10**. Cycloheptanone was oxidized with *m*-CPBA to form the lactone **11**, which was immediately opened with methanol and H<sub>2</sub>SO<sub>4</sub> to form **12** (Figure 20A). The alcohol **12** was oxidized to the aldehyde to use in a Wittig reaction. However, after the addition of the aldehyde to the ylide the reaction, decomposition ensued. Therefore, an alternative method was used to reach **14**.

Fisher esterification of 7-octenoic acid led to the desired methyl ester, **14**, which was then oxidized with *m*-CPBA to produce the epoxide **15**. The epoxide was stored for later use with **17** from Figure 20B. To synthesize **17**, 8-bromooctanol was oxidized to the aldehyde **16** using PDC. The aldehyde was used in a Wittig reaction with isopropyltriphenylphosphonium iodide to form **17**. In order to synthesize **10** from the alkene **17** and the epoxide **15**, first a lithium-bromine exchange needed to occur. Lithium with 4, 4'-*bi-tert*-butylbiphenyl was tried for the metal-bromine substitution, but the lithiated species never seemed to form. Instead of lithium, magnesium was used for the metal-halogen exchange. Ethylene bromide was a good activating reagent to use because it would not interfere in the next step. It was determined that the formation of the magnesium complex was best done by waiting for the reaction to become cloudy through

the action of the activator on the metal before adding the alkyl bromide. After the magnesium complex formed, it was converted to an organocuprate with the addition of  $\text{CuBr}\cdot\text{SMe}_2$ . A catalytic amount of TMSOTf was added to selectively silylate the epoxide to the secondary alcohol [102]. Under these conditions, the procedure was tested using *tert*-butyl glycidyl ether and bromooctane as a model system. Selective acylation of to 1-(*tert*-butyloxy)undecan-2-ol was successful. However, when using the alkyl bromide **17** and epoxide **15**, the product polymerized. Although it was realized that this synthetic route would have led to a final product one carbon longer than the desired alkene metathesis product, the synthetic pathway was still worth exploring. If all the steps had been successful, it would have been possible to complete the synthesis starting with 7-bromoheptanol to reach the desired product. The problem with the polymerization of the final product was unlikely to be related to the chain length difference, but rather occurred with production of an alkoxide in the presence of the ester moiety.

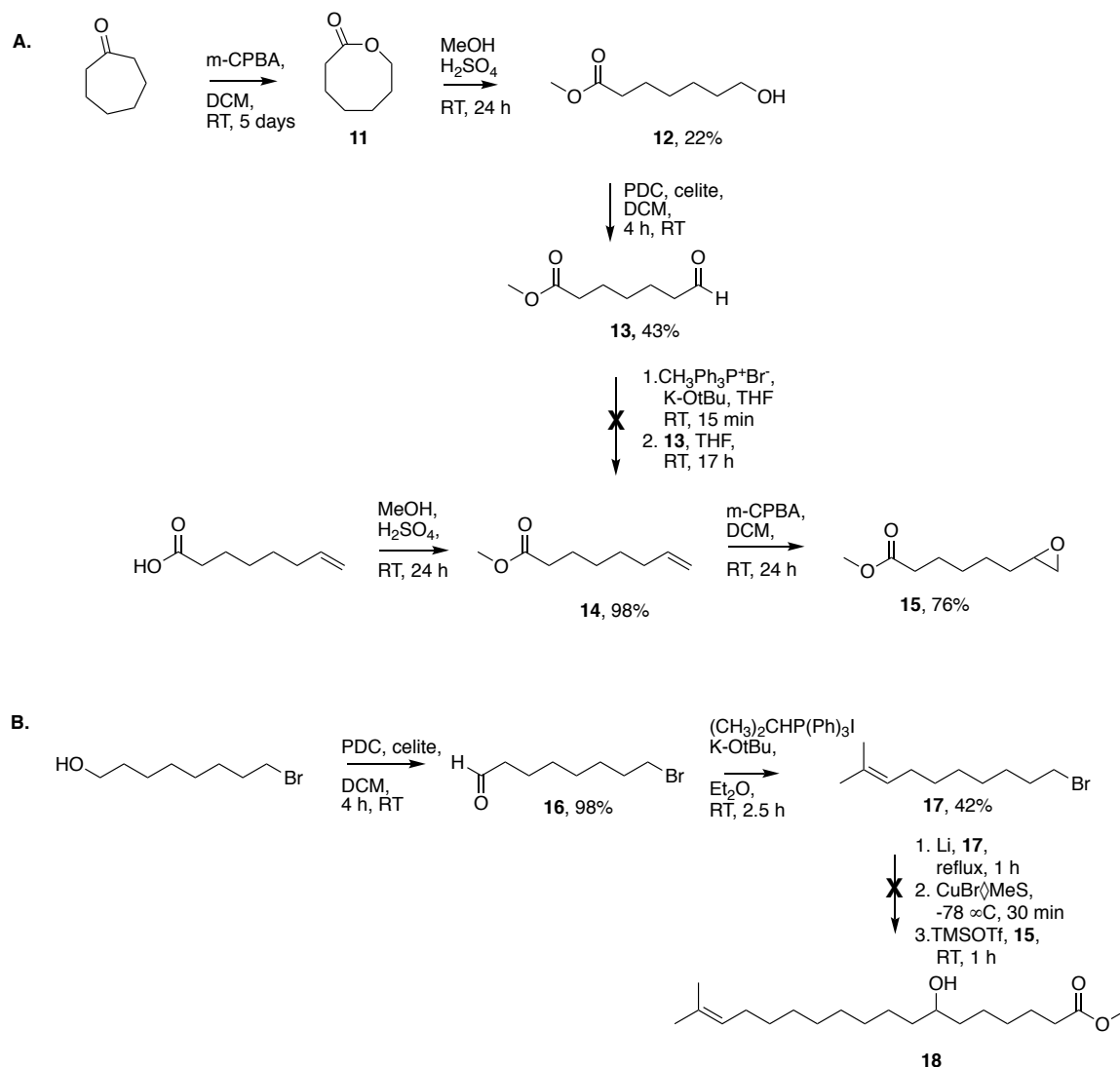


Figure 20: Attempted pathway to synthesize methyl 7-hydroxy-17-methyloctadeca-16-enoate.

A second pathway was devised that employs a convergent pathway (Figure 21A/B). The first intermediate was generated by protecting 8-bromooctanol as an TBDMS silyl ether (**19**) and then substituting the bromo group with a better iodo leaving group (**20**).

Separately, cycloheptanone underwent acylation with dimethyl carbonate to form the  $\beta$ -ketoester **21**. Activated **21** was alkylated by the previously synthesized **20** to form **22** and decarboxylated using LiCl at 180 °C to produce **23**. The cycloheptanone ring was regioselectively oxidized using UHP and TFAA to form the lactone **24**. It was observed

that, in the process, the TBDMS group was removed replaced by a trifluoroacetate group. The lactone was opened and the trifluoroacetate group was concomitantly removed with 24% tetramethylammonium hydroxide (TMAH) in methanol producing the diol **25**. The terminal alcohol needed to be selectively oxidized to the aldehyde. Using TEMPO and (bisacetoxyiodo)benzene (BAIB), with a short duration for the reaction, the aldehyde **26** was formed. This aldehyde was then immediately used in a Wittig reaction with isopropylidene triphenylphosphonium ylide completing the successful synthesis of **10b**. Unfortunately, neither racemic **10b** nor its derivatives could be separated by chiral GC/MS.

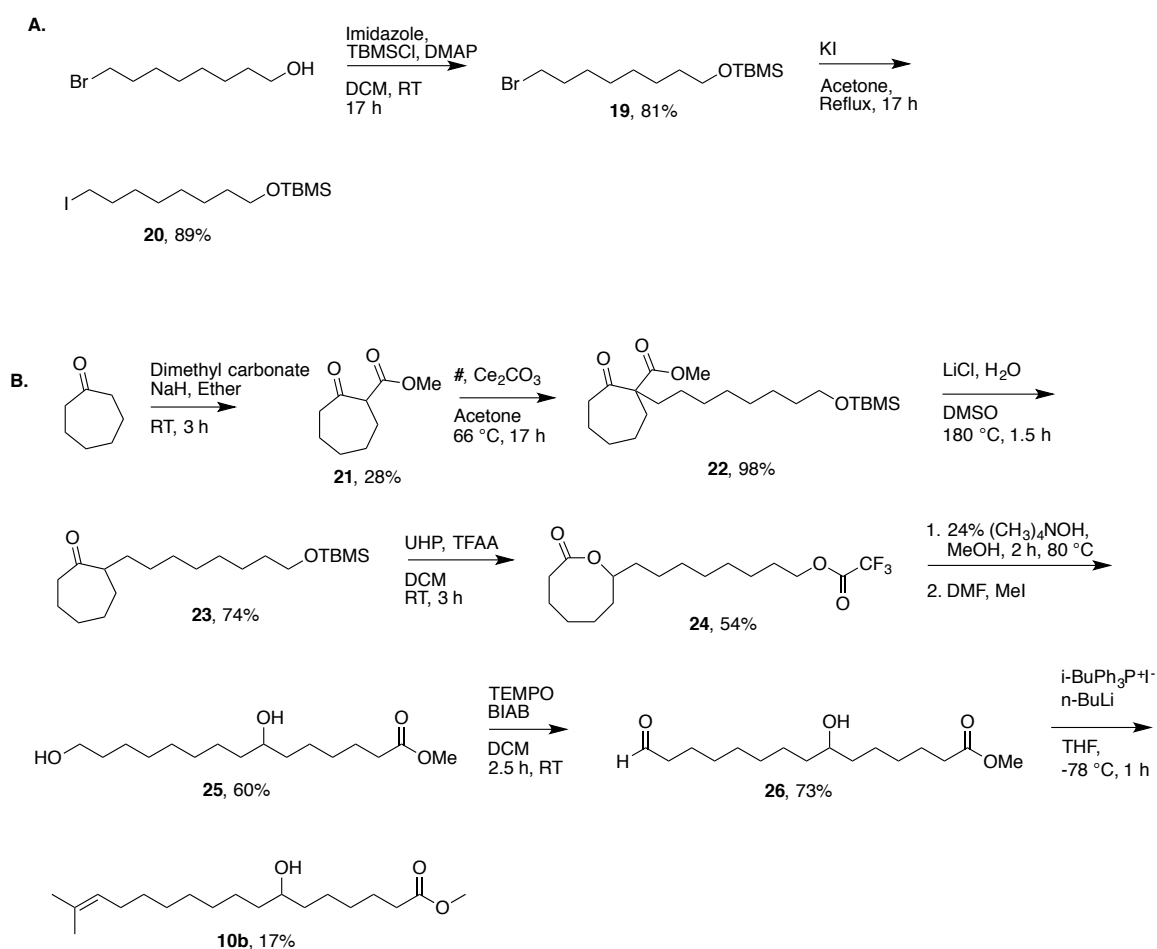


Figure 21: Synthesis of methyl 7-hydroxy-16-methyheptadeca-15-enoate used for chiral GC/MS analysis to determine the C-7 configuration.

The similarity of the alkyl chains of **10** around C7 precluded the stereochemical assignment. The previous metathesis reaction was scaled-up to convert **2a** into two products, 3-cyclohexen-1-ol (**9**) and 7-hydroxy-16-methyl-heptadec-15-enoate (**10**) was again used (Figure 22). Purifying by column chromatography allowed **7a** to be isolated for further analysis. Using a method from Hoye, diastereomeric (*S*)- and (*R*)- $\alpha$ -methoxy- $\alpha$ -trifluoromethylphenylacetic acid (MTPA) esters were prepared from **7a** in two mirrored experiments, **7b** and **7c** [100]. The corresponding (*S*)- or (*R*)-MTPA is dissolved in dry DCM. A carboxylic acid-activating agent, in this case *N,N'*-dicyclohexylcarbodiimide (DCC), was used to convert the carboxylic acid to an acid chloride. The alcohol was added to reaction mixture and was acylated by the acid chloride to form the diastereomeric Mosher esters, designated *R* or *S* specifying the stereochemistry of the acid that was used.

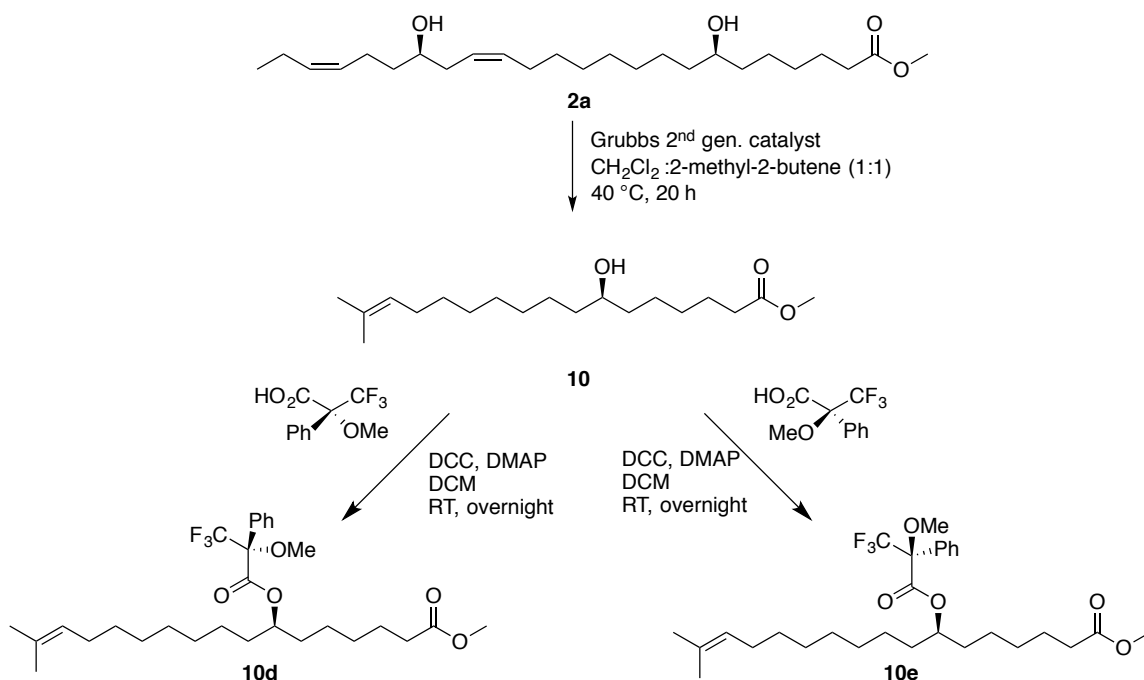
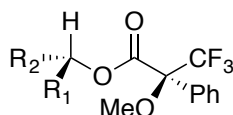


Figure 22: Synthesis and functionalization of methyl 7-hydroxy-16-methylheptadeca-15-enoate. Both **10d** and **10e** were used for Mosher analysis to determine the configuration of C-7.

NMR analysis was completed on compounds **10d** and **10e**. From  $^1\text{H}$  NMR analysis of the Mosher esters of **10d** and **10e**, the proton shift differences were calculated ( $\Delta\delta^{\text{SR}}$ ) between the diastereomeric esters. Again the similarity in the acyl chains around C7 provided some difficulty. Any variations at C6 and C8 were useless because the protons on these carbons were indistinguishable. Instead, we considered only assigned protons further down the acyl chain towards the alkene and the  $\alpha$ -methyl ester for assignable peaks. For **10d** and **10e**, the chemical shift differences were detected at C2, C14, and C15 in the  $^1\text{H}$  NMR (Table 2). Based on the methods described in the literature, protons with positive  $\Delta\delta^{\text{SR}}$  (C14 and C15) are found in R<sub>1</sub> and those that are negative (C2) are in R<sub>2</sub>, thus the configuration of C7 must be *R*.

Table 2: Mosher ester analysis of **10d** and **10e**



7-OH fragment $^1\text{H}$ NMR			
Carbon #	$\delta$ S-ester ppm	$\delta$ R-ester ppm	$\Delta\delta^{\text{SR}} (= \delta_{\text{S}} - \delta_{\text{R}})$
C2	2.245	2.286	-0.041
C14	1.9513	1.9416	0.0097
C15	5.1067	5.1046	0.0021

A dehydrogenative strategy was pursued in order to increase the dissimilarity between the alkyl chains surrounding C7. The method used needed to alter the alkyl chain without changing the stereochemistry of C7. Starting from the alkene metathesis product **10**, a selenoxide elimination reaction was used to insert an alpha-beta double bond into compound **27** (Figure 23). During the elimination, the preexisting alkene was oxidized to the epoxide. The new alkene added an additional functional group to the long alkyl chain.

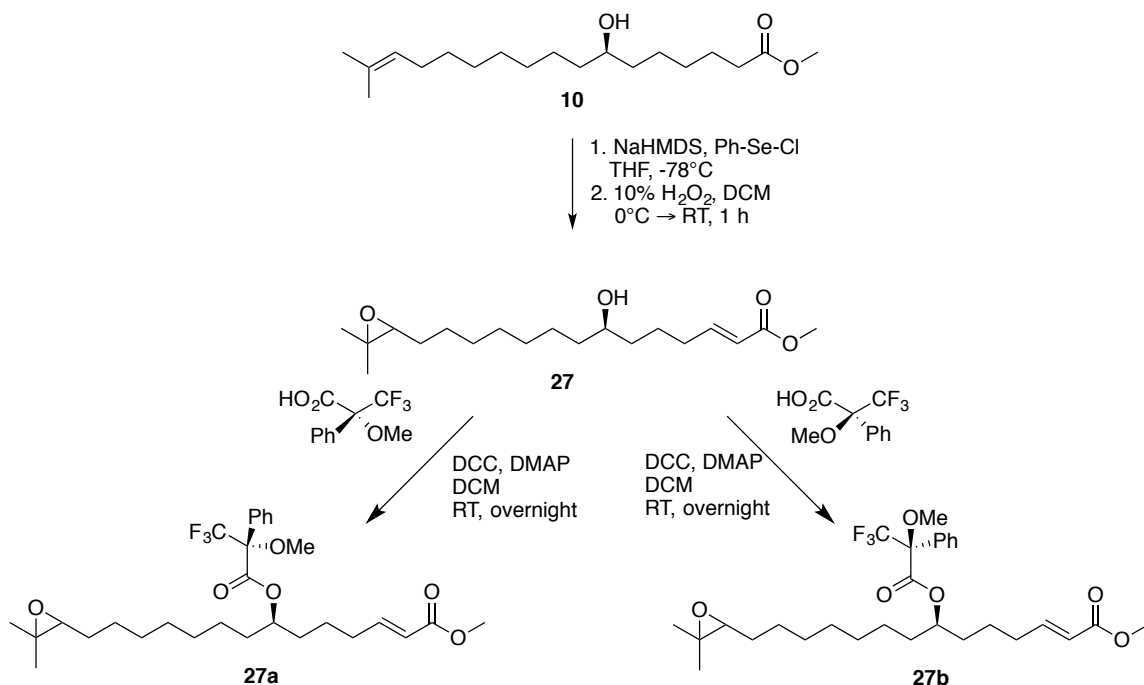


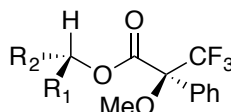
Figure 23: Synthesis and functionalization of methyl (*E*)-7-hydroxy-16-methyheptadeca-2,15-dienoate. Both **27a** and **27b** were used for Mosher analysis to determine the configuration of C-7.

For **27a** and **27b**, the chemical shift differences were detected at C2, C3, and C15 (Table 3). The addition of the alpha-beta double bond caused a larger shift difference at C2 and C3. However the epoxide formation diminished the chemical shift differences on the omega end of the chain. C15 and C14 for the *R*-acid derived ester became more shielded and the protons on C14 could no longer be unambiguously assigned. The difference in chemical shift had approximately the same magnitude at C15 whether it was the epoxide or an alkene. As with **10**, the only protons used for to calculate the different in chemical shifts were those that could unambiguously be assigned, and C6 and C8 were still indistinguishable. The protons that are positive (C15) are found in the  $\omega$ -end and those that are negative (C2 and C3) neighbor the methyl ester, thus the configuration of C7 must be *R*. In the paper by Hoye, they tested their experiment method on (-)-menthol [100]. Their  $\Delta\delta^{\text{SR}}$  values ranged from |0.32| to |0.01| [100]. Other studies have similarly found the effects observed from Mosher analysis is typically observed to the 2nd decimal place due to the field strength of the instrument [103-105]. This indicates our most



reliable data from **10d** and **10e** is at C2 and for **27a** and **27b**, C2 and C3. Though more recently, Guido *et al.* published an article showing new precision measurements could be used for  $^1\text{H}$  NMR experiment [106]. They reported that the standard two decimal rule adequately reflect  $^1\text{H}$  chemical shifts at a magnetic field strength of  $\leq 100$  MHz, but with an increase in proton field strength to  $\geq 300$  MHz, data can be meaningfully analyzed in the 4th decimal place [106]. In our case, our 500 MHz instrument allows us to carry out the data to 4 decimal places, thus using all the chemicals shifts from the unambiguously assigned protons providing further evidence C7 is *R*.

Table 3:  $^1\text{H}$  chemical shifts from Mosher analysis of **27a** and **27b**.



alpha,beta 7-OH $^1\text{H}$ NMR			
Carbon #	$\delta$ S-ester ppm	$\delta$ R-ester ppm	$\Delta\delta^{\text{SR}} (= \delta_{\text{S}} - \delta_{\text{R}})$
C2	6.843	6.901	-0.058
C3	5.738	5.79	-0.052
C15	2.701	2.699	0.002

$^{19}\text{F}$  NMR spectra of each set of Mosher esters were also analyzed as a further check on the C7 configuration. However, it is important to keep in mind that work done by Kakisawa and Kashman has shown  $^{19}\text{F}$  NMR lacks reliability and is dependent on the bulkiness of the  $\text{R}_1$  and  $\text{R}_2$  groups [103]. In the preferred conformation, the trifluoromethyl group is eclipsed with the carbonyl group, and the chemical shift difference of the  $\text{R}_1$  and  $\text{R}_2$  depends on the phenyl group [103]. The aryl group produces an anisotropic, magnetic shielding effect on the protons on  $\text{R}_1$  for the *R* acid or  $\text{R}_2$  for the *S* acid. The shielding effect by the phenyl group depends on the steric bulkiness of the  $\text{R}_1$  and  $\text{R}_2$  substituents [103]. Due to the high similarity of the alkyl chains on either side of **10** and **27**, there was no chemical shift difference in  $^{19}\text{F}$  NMR spectra. Nebraskanic acid (**1**) was thought to be the monounsaturated version of wuhanic acid (**2**) and, therefore, the stereochemistry at the hydroxyl groups is presumed to be the same.

Finally, by a combination of  $^1\text{H}$ - $^1\text{H}$  NOESY,  $^3J_{\text{HH}}$ , and IR analysis, the double bond was determined to be *Z*. NOESY signals were observed between H15-H16 and H15-H14, and no correlation was found for H15-H17. Similarly, a NOESY crosspeak was found for H21-H20 but not H21-H23. C=C stretching and out-of-plane bending vibrations at 1653 and 724  $\text{cm}^{-1}$ , respectively, further support the *cis* stereochemistry. The *cis*  $^3J_{\text{HH}}$  coupling constant at C15 was clearly visible at 11 Hz. This establishes the final structure of wuhanic acid as (7*R*,15*Z*,18*R*,21*Z*)-7,18-dihydroxytetracos-15,21-dienoic acid **2** and monounsaturated nebraskanic acid **1** shown in Figure 24.

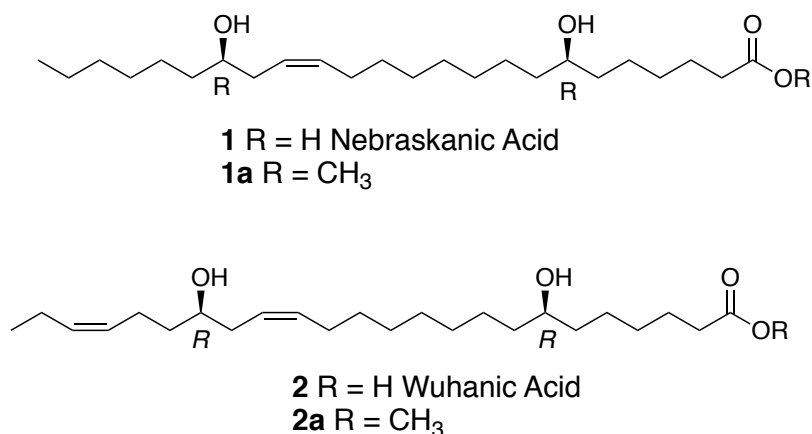


Figure 24: Final structures of the very-long chain dihydroxy fatty acids from *O. violaceus* seed oil. The structures of mono and diunsaturated nebraskanic (**1**) and wuhanic (**2**) fatty acids, respectively, supported by the MS, NMR, IR, and degradation analyses.

### 2.3.5 Oil composition

Gas chromatographic analysis of the total fatty acids of *O. violaceus* seeds revealed not only the accumulation of nebraskanic and wuhanic acids to 50 wt% of the total fatty acids, but also the presence of C<sub>18</sub>-C<sub>24</sub> monohydroxy fatty acids that accounted for <2% of the total fatty acids. These included C<sub>18</sub>-C<sub>24</sub> fatty acids with  $\omega$ -6 hydroxylation based on GC-MS analyses. TLC separation of the *O. violaceus* seed oil indicated no detectable TAG containing exclusively non-hydroxy fatty acids. However, at least two of the TLC-separated bands from *O. violaceus* oil contained >60 wt% of **1**

and **2**, suggesting that the oil from these seeds is more complex than that typically found in hydroxyl fatty acid-containing seeds such as those of castor bean. This may reflect the occurrence of TAG estolides in *O. violaceus* seed oil as shown previously for seed oils containing hydroxy fatty acids (Figure 25) [107-111]. Given that VLCFAs are typically excluded from the *sn*-2 position of Brassicaceae seed oils [112], it is likely that the C<sub>24</sub> dihydroxy fatty acids are at the *sn*-1 and *sn*-3 positions of *O. violaceus* TAG.

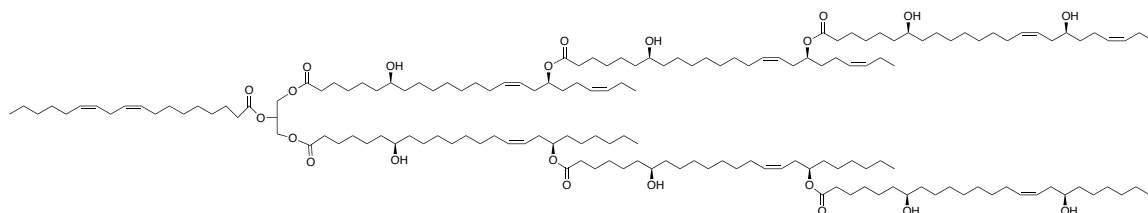


Figure 25: Hypothetical structure of TAG-estolides found in *O. violaceus* seed oil [107].

Based on the structures of nebraskanic and wuhanic acids, biosynthetic pathways for these fatty acids likely involve two hydroxylations to form the 7- and 18-hydroxyl groups and elongation reactions to extend a de novo C<sub>18</sub> fatty acid to a C<sub>24</sub> fatty acid. Given the position of the 18-hydroxyl group in the  $\omega$ -6 position, it is likely that it is formed by a FAD2-type hydroxylase, similar to biosynthesis of the  $\omega$ -6 hydroxyl group in ricinoleic acid in castor bean seeds and lesquerolic acid in *P. fendleri* seeds [113, 114]. Presumably, this reaction uses oleic acid bound to PC, as previously shown for ricinoleic acid synthesis in castor bean seeds [115]. Fatty acid elongation likely requires one or more *FAEI*-encoded  $\beta$ -ketoacyl-CoA synthase for three cycles of chain elongation from C<sub>18</sub> to C<sub>24</sub>.

Unclear is the biosynthetic origin of the 7-hydroxyl group. Clues were generated from detailed mass spectrometric analyses of fatty acids and acyl-CoAs of *O. violaceus* seeds. The Cahoon group, used selected ion searches of silylated fatty acid methyl esters prepared from *O. violaceus* seeds, compounds with prominent *m/z* 73, 185, and 187 ions arising from the  $\omega$ -6 hydroxyl group were detected. This included trace amounts (<0.1% of total fatty acids) of the C<sub>22</sub> precursors (5,16-(OH)<sub>2</sub>-22:1 $\Delta$ <sup>13</sup>, designated “indianic acid”, and 5,16-(OH)<sub>2</sub>-22:2 $\Delta$ <sup>13,19</sup>) of nebraskanic and wuhanic acids and a compound with a mass spectrum containing fragments partially consistent with those expected for the C<sub>20</sub>

precursor of nebraskanic acid (3,14-(OH)<sub>2</sub>-20:1 $\Delta^{11}$ ) (Figure 26 *A,B*). In addition, C<sub>24</sub> dihydroxy fatty acyl-CoAs and C<sub>22</sub> and C<sub>20</sub> dihydroxy acyl-CoAs were detected in developing *O. violaceus* seeds by ESI-MS/MS analyses of acyl-CoA pools.

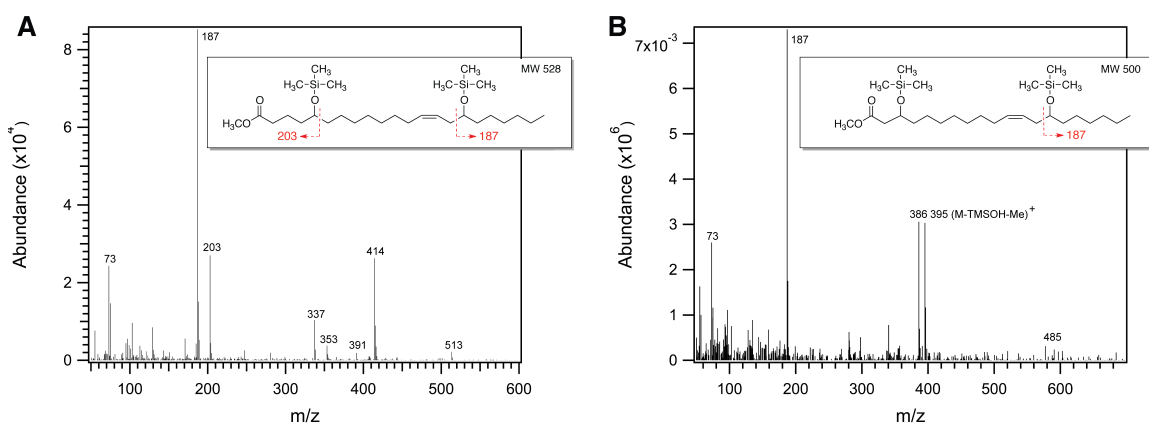


Figure 26: Mass spectra of fatty acid methyl esters (FAMES) from *O. violaceus* seeds and transgenic *Arabidopsis* seeds. Results shown in *A* and *B* are from GC-MS analysis of trimethylsilyl (TMS) derivatives of minor dihydroxy FAMES in *O. violaceus* seeds. Shown in *A* is a mass spectrum consistent with the methyl ester of TMS<sub>2</sub>-5,16-OH-22:1 and in *B* is a mass spectrum consistent with that of the methyl ester of TMS-3,14-OH-20:1.

Collectively, these findings are most consistent with a biosynthetic pathway involving  $\omega$ -6 hydroxylation of oleoyl-PC to form ricinoleoyl (12-OH-18:1 $\Delta^9$ )-PC and subsequent release of ricinoleic acid to the acyl-CoA pool (Figure 27). This is followed by elongation of the C<sub>18</sub> ricinoleoyl-CoA with introduction of a 3-hydroxyl group at the C<sub>20</sub> stage to form 3,14-OH-20:1 $\Delta^{11}$ -CoA and two further elongation cycles to sequentially generate 5,16-(OH)<sub>2</sub>-22:1 $\Delta^{13}$ -CoA and the CoA ester of nebraskanic acid (7,18-(OH)<sub>2</sub>-24:1 $\Delta^{13}$ -CoA). The  $\omega$ -3 unsaturation of wuhanic acid likely is introduced by FAD3, presumably at the C<sub>18</sub> stage using ricinoleoyl-CoA as a substrate [116] followed by the same elongation-hydroxylation pathway used for nebraskanic acid synthesis. Alternatively, the 3-hydroxy group could arise from a  $\Delta^3$  desaturase- or cytochrome P450-related hydroxylase (Figure 27, Pathway 2) or be introduced during the C<sub>18</sub> to C<sub>20</sub> elongation by “missed” dehydration of the  $\beta$ -hydroxyl-acyl-CoA intermediate

such as that found in *Arabidopsis pas2* 3-hydroxy-acyl-CoA dehydratase mutants [117] (Figure 27, Pathway 1). In the case of the *pas2-1* mutant, 3-hydroxy-fatty acyl-CoAs are detectable in acyl-CoA pools but are not further elongated [117]. 3-Hydroxy-fatty acids also occur in bacteria, mammals, and fungi but are also not further elongated [118-121]. As such, discontinuous elongation of a 3-hydroxy-acyl-CoA in *O. violaceus* would represent an unprecedented pathway in hydroxy fatty acid metabolism. This pathway is supported by the ability to partially reconstruct the nebraskanic acid biosynthesis in *Arabidopsis* seeds by co-expression of the *O. violaceus* FAE1 and FAD2-type 12-oleic acid hydroxylase.

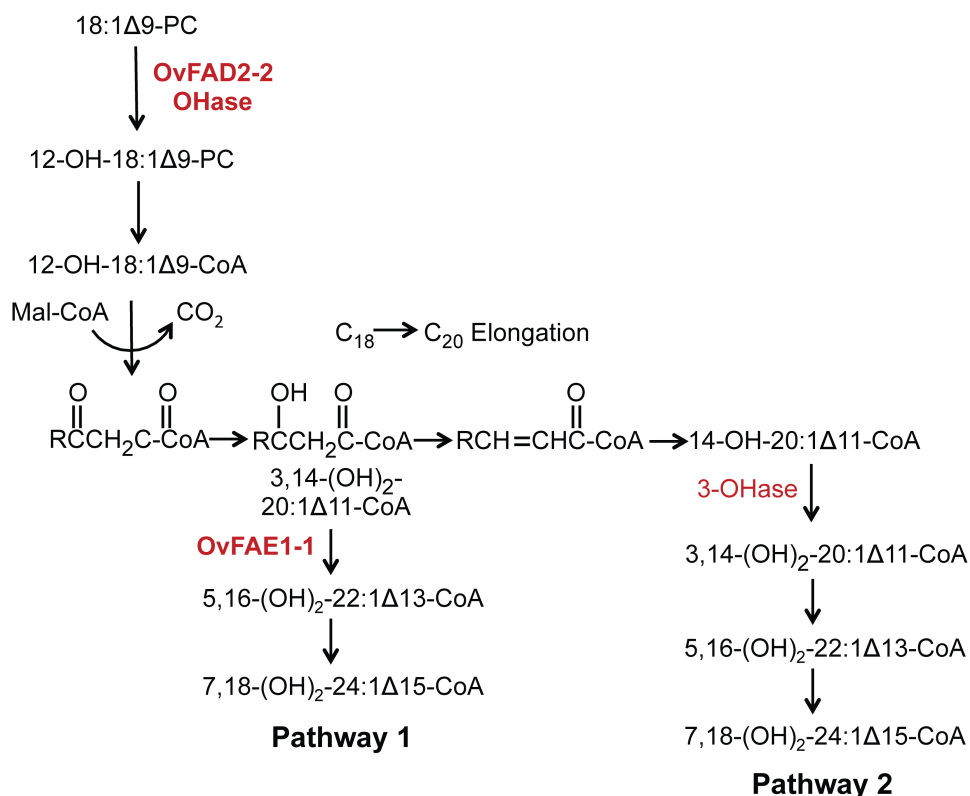


Figure 27: Hypothetical biosynthetic pathway for nebraskanic acid. The pathways shown are consistent with the detection of  $\text{C}_{20}$  and  $\text{C}_{22}$  dihydroxy precursors of nebraskanic acid in seed fatty acids and acyl-CoA pools of *O. violaceus* seeds through discontinuous elongation of the 3-hydroxy intermediate arising from ricinoleoyl-CoA (Pathway 1) or by 3-hydroxylation of the full  $\text{C}_{20}$ -elongation product of ricinoleoyl-CoA (Pathway 2). The characterization of OvFAD2-2 as an oleic acid 12-hydroxylase accounts for the  $\omega$ -6 hydroxyl group of nebraskanic acid. Though not demonstrated, a second FAE1 (e.g., OvFAE1-2) may be associated with subsequent elongation steps.

## 2.4 Conclusion

In this chapter, a seed oil rich in two very-long chain dihydroxy fatty acids was identified in the Brassicaceae *O. violaceus*. The structures of these fatty acids were elucidated using multiple and independent approaches as 7,18-dihydroxy-24:1 $\Delta^{15}$  (nebraskanic acid) and 7,18-dihydroxy-24:2 $\Delta^{15,21}$  (wuhanic acid). The detection of C<sub>20</sub>-C<sub>22</sub> dihydroxy intermediates in the fatty acid and acyl-CoA pools of *O. violaceus* seeds is consistent with the introduction of the 7-OH group as a 3-OH precursor at the C<sub>20</sub> stage of fatty acid elongation (Fig. 31, Pathway 1). The synthesis of nebraskanic acid in Arabidopsis seeds without the introduction of a second hydroxylase is consistent with the formation of the 3-OH as a fatty acid elongation intermediate during the conversion ricinoleoyl-CoA to the C<sub>20</sub> product (Figure 31, Pathway 1). Because of the low accumulation of nebraskanic acid in the engineered Arabidopsis seeds, it is likely that additional components of the very long-chain fatty acid biosynthesis are also specialized for nebraskanic and wuhanic acid TAG biosynthesis. For example, a second *O. violaceus* FAE1 may be required for more efficient elongation of the C<sub>20</sub> and C<sub>22</sub> dihydroxy fatty acyl-CoAs and/or the 3-hydroxy-acyl-CoA dehydratase in *O. violaceus* seeds may not effectively use ricinoleoyl-CoA as a substrate. In addition, the  $\omega$ -3 double bond of wuhanic acid likely arises from FAD3 activity, but it is currently unclear in which biosynthetic step that this double bond is introduced. Elucidation of these details of the biosynthesis of nebraskanic and wuhanic acids will provide the basis for an alternative pathway for hydroxy fatty acid biosynthesis of biochemical and biotechnological significance. In addition, our findings also unveil *O. violaceus* as a potential industrial crop with oil properties for applications such as lubricants and chemical feedstock synthesis distinct from those of castor bean oil. Although *O. violaceus* is currently used for ornamental purposes, extensive germplasm exists in China for selection and breeding of cultivars with suitable agronomic performance [31, 32].

## CHAPTER 3. THE CONVERGENCE OF TWO PATHWAYS IN THE BIOSYNTHESIS OF ALKAMIDES FROM *ECHINACEA*

### 3.1 Introduction

A goal of the research described in this chapter was to identify means to probe both the acyl chain and the amine moieties of alkamides in *E. purpurea*, to gain insight into poorly understood desaturation, chain-length determining, and amide forming aspects of this pathway. First, isotopic labeling experiments were completed in order to determine if branched chain amino acids (BCAA), specifically Val and Ile, are verifiably precursors to the amine moiety [1]. Further desaturation of the acyl chain must be occurring in order to provide the more complex alkene and alkyne functionality of the accumulating alkamides. If the fatty acid synthesis and branched-chain amino acids pathways converge to make alkamides, fatty acids may be desaturated prior to merging with the BCAA; alternatively, saturated fatty acids might be utilized for alkamide synthesis and then the alkamides are further desaturated. Given that the natural products universally bear the conjugated *2E* alkene, short-chain, unsaturated fatty acids (12:1 $\Delta^{2E}$ , 13:1 $\Delta^{2E}$ , 14:1 $\Delta^{2E}$ ) and alkamides ( $^{N_1}$ iBu/12:1 $\Delta^{2E}$ ,  $^{N_1}$ iBu/13:1 $\Delta^{2E}$ ,  $^{N_1}$ iBu/14:1 $\Delta^{2E}$ ) were synthesized. These compounds were used in plant experiments attempting to intersect the biosynthetic pathway. Finally, acyl-CoAs (12:1, 12:2, and 12:4) were synthesized and used in an activity assay with a cloned enzyme, EpLigase, isolated by the Nikolau group [122].  $K_m$  and  $V_{max}$  were calculated and compared to 12:0. Confirmation of the product formed from the assay was completed.

#### 3.1.1 Convergence of two metabolic pathways

Fatty acids are most often modified with additional functional groups on the alkyl chain. Other chemistry can occur naturally at the carboxylic acid moiety within plants. An example of fatty acid modification is the biosynthesis of sorgoleone, an allelochemical that is secreted from the roots of sorghum (Figure 28) [123]. The convergence of two metabolic pathways, fatty acid synthesis and polyketide synthesis, leads to the biosynthesis of sorgoleone [124]. *De novo* fatty acid synthase and desaturase produce 16:3 $\Delta^{9Z,12Z,15Z}$ -CoA [123]. This fatty acid is the starter unit for a polyketide

synthase that produces a resorcinolic lipid intermediate that is further oxidized to yield sorgoleone [123]. Through  $^{13}\text{C}$ -labeled substrate experiments, a better understanding of the biosynthesis of sorgoleone was determined [124].  $[2-^{13}\text{C}]$ Glucose incorporated into acyl chain and into the ring of sorgoleone, whereas  $[1-^{13}\text{C}]$ acetate incorporated into the ring only suggesting PKS and FAS are taking place in different cellular compartments.

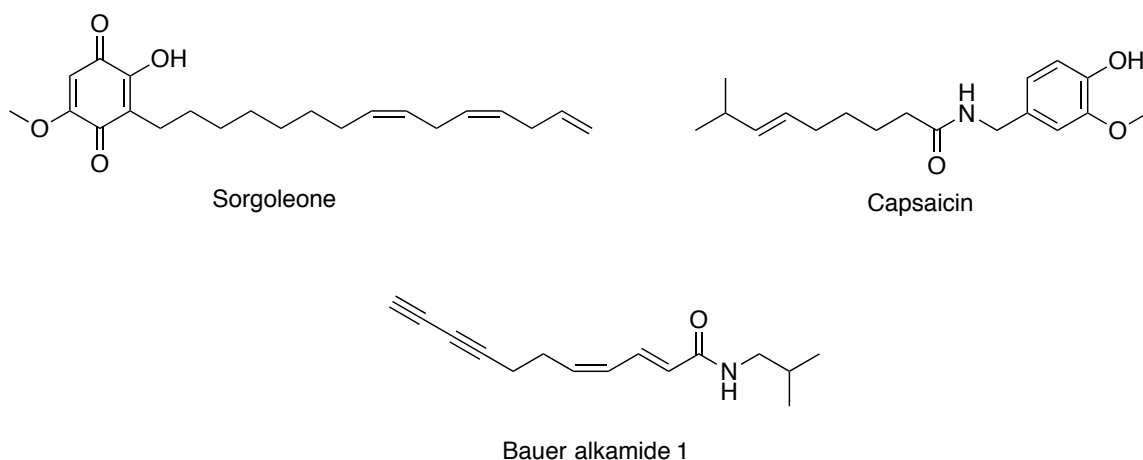


Figure 28: Examples of natural products synthesized from the integration of fatty acid synthesis with another pathway. Sorgoleone is derived from fatty acid and polyketide synthesis. Capsaicin is synthesized from phenylpropanoid and branched-chain fatty-acid synthesis. Alkamides look to be another example of the convergence of fatty acid with another pathway, but the biosynthetic pathway is still unknown.

Capsaicinoids are secondary metabolites responsible for the hot taste in peppers, where capsaicin and dihydrocapsaicin are the main determinants of capsaicinoid content [125]. Capsaicinoids bind to the vanilloid receptor subtype I (TRPV1), which can also be stimulated with heat, explaining the similar sensation given by capsaicinoids. Capsaicin and other capsaicinoids are synthesized by convergence of two metabolic pathways, phenylpropanoid synthesis and branched-chain fatty-acid synthesis (Figure 29) [126]. Through a series of steps, phenylalanine is converted to vanillylamine through phenylpropanoid synthesis [126]. Separately, pyruvate is converted to valine, which is synthesized to isobutyryl-CoA [126]. Then through branched chain fatty acid synthesis, 8-methyl-6-nonenoyl-CoA is produced [126]. Finally, the two pathways join through the condensation of vanillylamine and 8-methyl-6-nonenoyl-CoA to yield capsaicin [125].



Many of the enzymes in this pathway are known, but they are not well characterized and the regulation not understood. In particular, the key acyl transferase is unknown, but in recent work a gene, *CA01g32970*, has been identified as a possibility to encode the AT [126]. Further studies still need to be completed to characterize the enzyme and better understand regulation of the pathway.

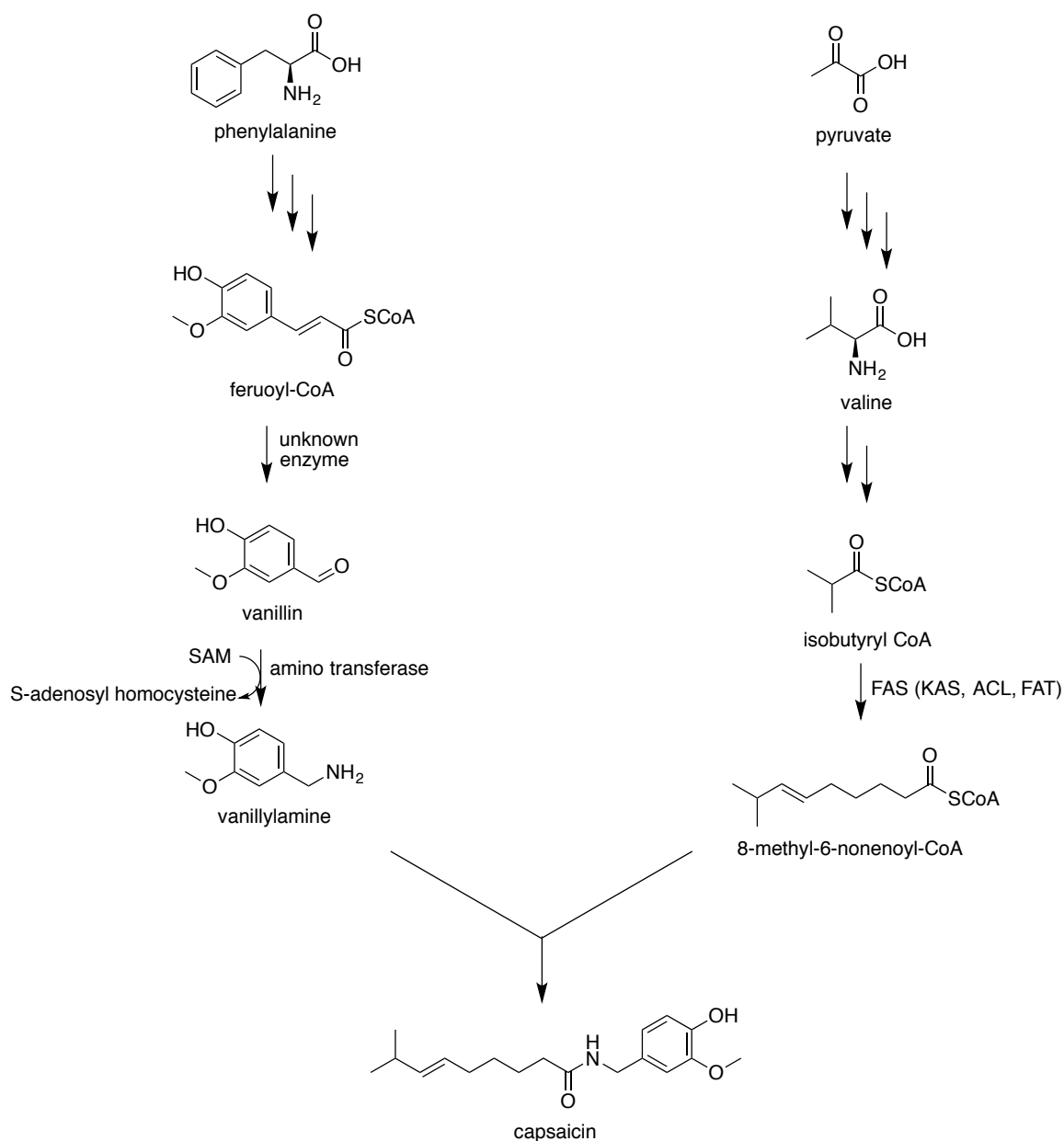


Figure 29: Biosynthetic pathway of capsaicin. Multiple arrows indicate that several steps occur, but not all of the steps are known to generate the next intermediate. An acyltransferase is thought to be responsible for the last step; a recent study has shown *CA01g32970* as a possible gene encoding the AT.

The *Echinacea* genus from the Asteraceae family produces alkamides, secondary metabolites similar in structure to capsaicin. It has been proposed that the alkamides produced by *Echinacea* can act as immunomodulatory agents in humans, but continued

research is needed to verify a biological purpose [127]. The biosynthetic pathway is also unknown; however, similar to sorgoleone and capsaicin, alkamides also seem to be derived from two pathways. The structure resembles capsaicin in that all alkamides have an amide moiety. Though in alkamides amino acids, such as valine or isoleucine, look to derive the amine instead of the acyl chain like in capsaicin. Like sorgoleone and capsaicin, the alkyl chain looks to be derived from fatty acid biosynthesis. Alkamides are unique for their several desaturations along the alkyl chains. It has been proposed that the branched-chain amino acid could react with the carboxylic acid end of a fatty acid to produce the amide moiety similar to the biosynthesis of alkamides from *Acmella radicans* (Asteraceae) [128]. If this is true, it would be another example of fatty acids being further modified in plants, the convergence of two biosynthetic pathways, and further evolution of secondary metabolism.

### 3.1.2 Alkamide biosynthesis in *Echinacea*

The purple coneflowers *Echinacea* (Asteraceae) are native to North America and Europe [129]. They are well known for their medicinal uses including the prevention of colds, influenza, and upper respiratory tract infections [129]. Of the thirteen known species, *Echinacea purpurea*, *E. pallida*, and *E. angustifolia*, are the most used for alternative medical purposes [129]. *Echinacea* contains a family of metabolites, the alkamides, in relatively high concentrations that could be one of the origins of the medicinal properties of *Echinacea* [129].

The convergence commercial holistic *Echinacea* therapeutics contain one or more of the *Echinacea* species that produces alkamides [130]. Alkamides also known as alkyl-amides are found in 33 plant families and comprise about 200 chemically related compounds [1, 131]. They contain an acyl chain and amine functional groups connected by an amide [132]. In *Echinacea*, the acyl moiety is 11 to 16 carbons in length, typically with a *trans*-configured double bond located at the alpha position [131]. The degree to which the acyl chain is desaturated varies between alkamides. Most of the acyl chains are further desaturated and include a double bond at the 4, 8, and/or 10 positions; the acetylenic bonds are found at the 8 or 10 position [132, 133]. In *Echinacea*, their amide

end is generally an isobutyl or 2-methylbutylamide [134]. In Figure 30, eight Bauer alkamides are shown that correspond to the typical structural variations found in *Echinacea*. Originally Bauer renamed the alkamides from the systematic IUPAC naming system to a lab-specific BA nomenclature (Bauer alkamides, BA) 1, 2, 3...etc. However this naming system provides no insight to the chemical structure of each alkamide. Therefore, during the course of our research, a new naming system was developed and published based on that is the more widely accepted and descriptive fatty acid nomenclature. For example, BA2 can be named  $^N\text{iBu}/11:2\Delta^{2Z,4E,8a,10a}$ . The “ $^N\text{iBu}$ ” indicates an isobutyl group is bonded to the amide nitrogen to a fatty acid 11-carbon chain length that has two double bonds at the 2 and 4 position, which are *Z* and *E* configurations, respectively. The two acetylenic bonds are at the 8 and 10 positions. Similar to shorthand fatty acid nomenclature, the number after the colon only indicates carbon-carbon double bonds [1].

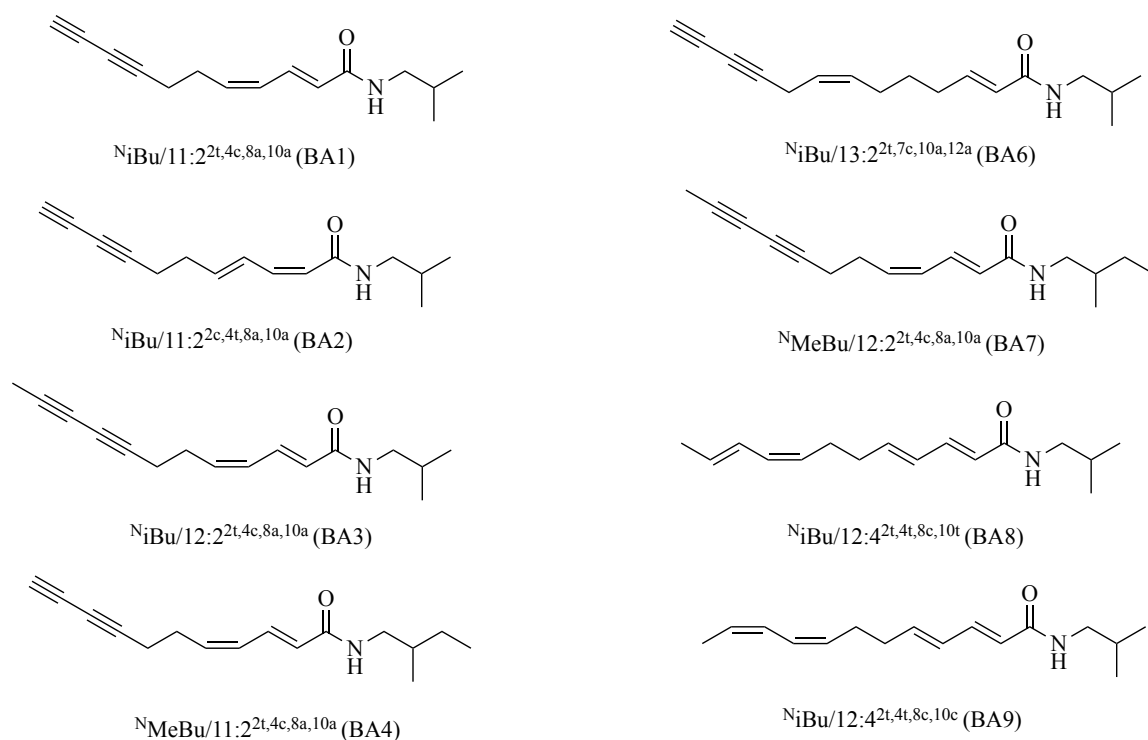


Figure 30: Abundant alkamides from *Echinacea purpurea* that are known as Bauer alkamides

The structure of alkamides most resembles that of common endocannabinoids, such as anandamide; alkamides have been found to have analogous pharmacological activities (Figure 31) [1, 132, 135]. Endocannabinoids bind to CB<sub>2</sub> receptors of T cells, macrophages, and hematopoietic cells. The *N*-alkyl branch stimulates the CB<sub>2</sub> receptors in the immune system. Another endogenous fatty acid alkamide is *N*-palmitoylethanolamine (Figure 31). It is not considered an endocannabinoid because it lacks the ability to bind to CB<sub>1</sub> and CB<sub>2</sub>, but it does enhance anandamide activity [136-138]. Due to their structural similarities, it has been proposed that alkamides could be functional analogs of anandamide or *N*-palmitoylethanolamine, and as a result could be beneficial in the fight against colds and the influenza virus [130, 135]. However, in 2011, a study reported patients taking *Echinacea* extracts showed no improvement in cold symptoms compared to those not taking the extracts [132, 139]. Another study showed a potential role for alkamides in plant signaling [140, 141]. It was found that alkamides modified root growth and differentiation in *Arabidopsis* by interacting with a cytokinin-signaling pathway [140, 141]. Plants synthesize *N*-acylethanolamines (NAE), NAE-12:0 and NAE 14:0, which are structurally similar to alkamides. NAEs are used for plant defense by signal transduction and in innate immunity [142-144]. Other groups have proposed that alkamides specifically could stimulate jasmonic acid biosynthesis [145].

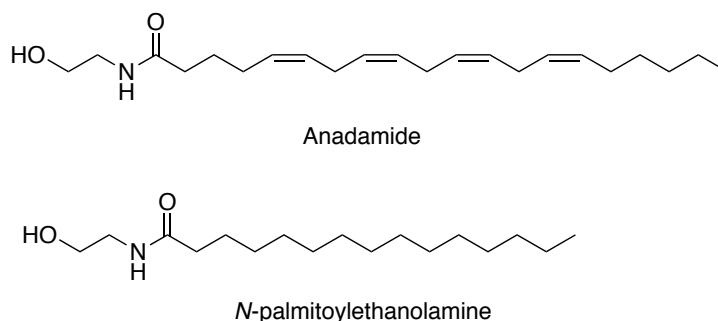


Figure 31: Endogenous fatty acid amides with important biological activities

Alkamides also resemble the structure of sphingolipids, critical components of eukaryotic cell membranes [146]. The basic structure of ceramide, a sphingolipid from animal systems, shown in Figure 32 can be further modified by differences in chain length, degree of unsaturation, insertion of additional hydroxyl groups, and methyl

branching [146]. In animal systems, ceramide contains a  $\Delta^4$ -*trans* double bond that is generated by a desaturase and not by a dehydrogenase [146]. Sphingolipids also have an additional C8 unsaturation that can be either in the *cis* or *trans* configuration [146]. A  $\Delta^8$  fatty acid desaturase discovered by Chen produced 8*E* and 8*Z* in a 4:1 ratio in sphingolipids [147]. Tian *et. al.* discovered a  $\Delta^{10E}$ -sphingolipid desaturase that catalyzes desaturation in the sphingolipid fusaruside. This example may serve as a window into how desaturation occurs in alkamide biosynthesis [148].

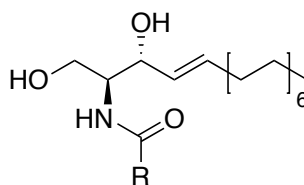


Figure 32: Sphingosine backbone of ceramide where R is the fatty acid residue.

The location where alkamides are abundant fluctuates across *Echinacea* species. Both *E. angustifolia* and *E. purpurea* have abundant alkamides in the roots and certain aerial parts, whereas *E. pallida* have about the same abundance in their aerial parts, but lower concentration in the roots [132, 149]. Beyond the knowledge of the structures of alkamides and their taxonomical distribution, only a few aspects of the biosynthetic pathway has been explored.

The enzymatic conversion of fatty acids to generate complex, specialized metabolites such as anadamide, *N*-palmitoylethanolamine, and sorgoleone is well known [132, 135, 150-152]. Endocannabinoid biosynthesis may give insight to the biosynthesis of alkamides. Endocannabinoids are synthesized by transferring arachidonic acid ( $C_{20}$ ) from a phospholipid to the primary amine of a phosphatidylethanolamine [128, 132]. Previous work completed by Bohlmann and Dallwitz in 1974 helps support this hypothesis for alkamides [153]. Using metabolites from the crepeneynate pathway and a cell-free extract of *E. purpurea*, Bohlmann and Dallwitz showed a low level of (0.1%) incorporation from  $[10-^{14}C_1]$ methyl oleate ( $18:1\Delta^{9Z}$ ),  $[16-^{14}C_1]$ methyl enediynoate ( $18:2\Delta^{9Z,12a,14a}$ ),  $[12-^{14}C_1]$ anacyclin ( $^{N}iBu/14:2\Delta^{2E,4E,8a,10a}$ ) into  $^{N}iBu/11:2\Delta^{2E,4E,8a,10a}$  [153].

Though the incorporation was low, it was hypothesized that alkamides are derived through oleic acid biosynthesis [153]. Intermediates of the pathway would have to undergo stepwise desaturation followed by chain shortening to achieve the desired chain length [153]. In addition, a study of alkamides within *A. radicans* cell cultures showed the amide moiety was derived from decarboxylated L-[<sup>2</sup>H<sub>8</sub>]valine and L-[<sup>2</sup>H<sub>8</sub>]phenylalanine [128]. These observations along with literature on acetylenic fatty acid biosynthesis were the basis for the Minto group to propose that the acyl chain of alkamides is derived from *de novo* fatty acid biosynthesis and the amine from decarboxylated amino acids [132].

While not specifically addressed by the study completed by Bohlmann and Dallwitz, fatty acids can be synthesized by two different pathways either via the chloroplast or mitochondria [154]. For most eukaryotes, *de novo* fatty acid synthesis occurs in the cytosol, but in plants it occurs in the chloroplast [154]. The chloroplastic pathway uses pyruvate to create the basic building block acetyl-CoA which is then carboxylated to malonyl-CoA [154]. This pathway synthesizes 16:0-ACP which can be hydrolyzed to produce palmitic acid (16:0) or portioned to palmitoyl CoA (Figure 33A) [154]. Furthermore, 16:0-ACP can be elongated to stearyl-ACP (18:0-ACP), which can either be released to the CoA pool, incorporated into phospholipids, transferred to complex lipids or desaturated to 18:1-ACP. 18:1-ACP can be hydrolyzed to produce 18:1 [154]. Again, plant fatty acid synthesis slightly differs from other eukaryotes. In plants, except for the *Poaceae* family, mitochondrial fatty acid uses cytosolic malonate because it lacks the ability to produce malonate directly [154-158]. Mitochondrial fatty acid synthase synthesizes three main fatty acids C<sub>8</sub>, C<sub>16</sub> and C<sub>18</sub> [158]. C<sub>8</sub>-ACP is an intermediate in the production of lipoic acid, an essential sulfur-containing cofactor (Figure 33B) [154, 157, 158]. Mitochondrial fatty acid synthase also produces minor fatty acids, such as C<sub>10</sub>, C<sub>12</sub>, and C<sub>14</sub>, and 3-hydroxylated forms of each of the acyl intermediates [158].

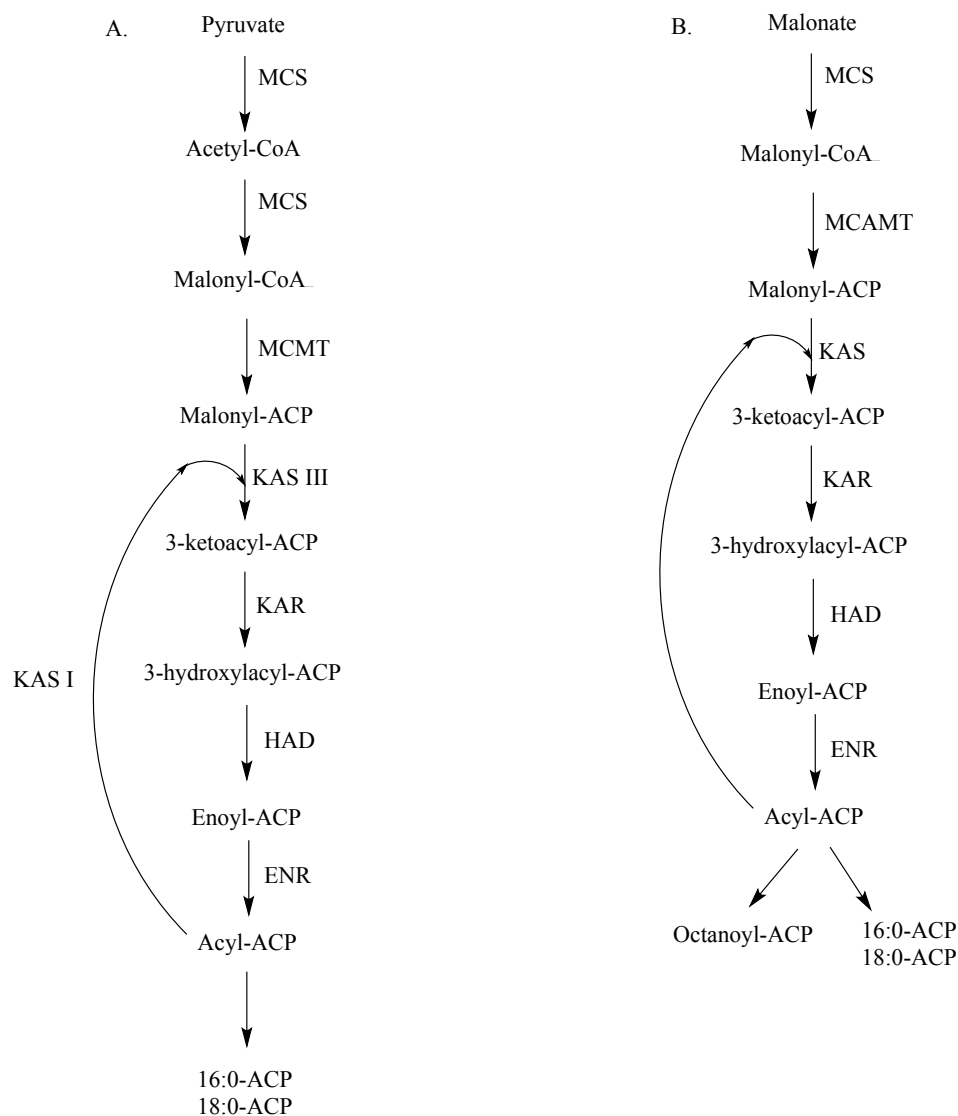


Figure 33: *A.* Fatty acid synthesis in the chloroplast, and *B.* in the mitochondria [154].

Our group has worked on gaining insight to the biosynthetic pathway of alkamides by using labeled compounds to probe both the acyl chain and the amine moiety. Initial experiments were completed with  $[1-^{13}\text{C}_1]$ glucose. Feeding experiments completed with  $[1-^{13}\text{C}_1]$ glucose showed an increased labeling ( $M + 2$  10.9%,  $M + 3$  4.6%,  $M + 4$  1.4%) in comparison to  $[1-^{13}\text{C}_1]$ acetate ( $M + 2$  2.28%) [132]. The impact on labeling of  $[6-^{13}\text{C}_1]$ glucose was similar to that of  $[1-^{13}\text{C}_1]$ glucose [132]. These observations appear to rule out loading of the acetyl-CoA pool from the oxidative pentose phosphate pathway



(OPPP) as a source of intermediary metabolites for alkamide biosynthesis. During the oxidative phase, C-1 is lost as CO<sub>2</sub>, often leading to the formation of fructose 6-phosphate; therefore, comparable labeling of [1-<sup>13</sup>C<sub>1</sub>]glucose and [6-<sup>13</sup>C<sub>1</sub>]glucose would not be observed if OPPP was providing the acetyl-CoA for alkamides biosynthesis [132].

[1,2-<sup>13</sup>C<sub>2</sub>]Glucose showed an envelope of M + 1 to M + 8 which was shifted to M + 4 to M + 10 when [U-<sup>13</sup>C<sub>6</sub>]glucose was used [132]. The isotopomer peaks of even integers were present at a higher abundance than the odd-increment peaks consistent with the metabolism of the carbohydrate to C<sub>2</sub> units [132]. Alkamides isolated from seedlings fed with sodium[1,2-<sup>13</sup>C<sub>2</sub>]acetate and [U-<sup>13</sup>C<sub>6</sub>]glucose gave HSQC data [132]. Cross-peaks showed bonded <sup>13</sup>C-<sup>13</sup>C unit interspersed between <sup>12</sup>C atoms (<sup>12</sup>CH-<sup>3</sup>CH-<sup>13</sup>CH-<sup>12</sup>CH) [132]. This reinforced that the acetate units were being incorporated.

Deuterium and carbon-13 labeled fatty acids [15,15,16,16,16-<sup>2</sup>H<sub>5</sub>]palmitate ([<sup>2</sup>H<sub>5</sub>]16:0), [17,17,18,18,18]-<sup>2</sup>H<sub>5</sub>]stearate ([<sup>2</sup>H<sub>5</sub>]18:0), [12,12,12-<sup>2</sup>H<sub>3</sub>]laurate ([<sup>2</sup>H<sub>3</sub>]12:0), [U-<sup>13</sup>C<sub>8</sub>]octanoate ([<sup>13</sup>C<sub>8</sub>]8:0), and [U-<sup>13</sup>C<sub>18</sub>]oleate ([<sup>13</sup>C<sub>18</sub>]18:1Δ<sup>9Z</sup>) were provided to seedlings in an attempt to intersect the acyl portion of alkamide biosynthesis [132]. [<sup>2</sup>H<sub>5</sub>]16:0 and [<sup>2</sup>H<sub>5</sub>]18:0 showed no intact incorporation into BA 8/9, but there was intact incorporation in downstream fatty acids [132]. Therefore, as hoped, they were taken up by the plants and converted to a form capable of being used for fatty acid synthesis [159]. The fatty acids were also broken down into acetate units. These labeled acetate units were observed in <sup>N</sup>iBu/12:4Δ<sup>2E,4E,8Z,10E/Z</sup> as the only route to isotopic incorporation by an enhancement of M + 2 peak (1.12% for [<sup>2</sup>H<sub>5</sub>]16:0 and 1.77% for [<sup>2</sup>H<sub>5</sub>]18:0) [132]. Short chain fatty acids, [<sup>2</sup>H<sub>3</sub>]12:0, and [<sup>13</sup>C<sub>8</sub>]8:0, also showed no intact incorporation into <sup>N</sup>iBu/12:4Δ<sup>2E,4E,8Z,10E/Z</sup>; a small incorporation of <sup>13</sup>C from the octanoate was detected at M + 2 (2.70%) [132].

Oleate is a precursor to polyunsaturated fatty acid only produced by the plastids which can further be desaturated to 18:2 [154]. While desaturation can occur in the chloroplast, most oleate is transported to the endoplasmic reticulum to produce polyunsaturated fatty acids [40]. Consistent with the above observations, [<sup>13</sup>C<sub>18</sub>]18:0Δ<sup>9Z</sup> was a precursor to 18:2 and 18:3 but was not a precursor to alkamide biosynthesis. [<sup>13</sup>C<sub>18</sub>]18:0Δ<sup>9Z</sup> showed no intact incorporation in BA 8/9, but led to indirect incorporation as M + 2 (2.23%) [132]. [<sup>13</sup>C<sub>18</sub>]18:0Δ<sup>9Z</sup> provided an *in planta* control for lipid uptake by

showing intact labeling of  $18:2\Delta^{9Z,12Z}$  (2.38%) and  $18:3\Delta^{9Z,12Z,15Z}$  (0.142%) [132]. Collectively, this data does not appear to support the expectation that the acyl chains of alkamides biosynthesis came from plastidal FAS [132].

To examine the potential origins of the acyl chain in alkamides biosynthesis, an inhibition study of fatty acid synthesis was undertaken. Metronidazole inhibits several enzymes in plant fatty acid biosynthesis including stearyl and linoleoyl desaturases and the elongase acting on palmitate [160, 161]. Metronidazole strongly perturbed the labeling patterns in BA 8/9. Plants fed  $[U-^{13}C_6]$ glucose and varying concentrations of metronidazole showed a significant reduction in the  $^{13}C$  content within  $N_iBu/12:4\Delta^{2E,4E,8Z,10E/Z}$  from 33% to 2.7% [132]. There was also a significant decrease in polyunsaturated fatty acids, which is consistent with reports that show metronidazole inhibits ferredoxin desaturation pathways for plant lipid biosynthesis [161]. For saturated fatty acids, which mostly come from the chloroplast, there seemed to be no simple concentration-dependent changes [132]. The significant reduction in labeling of  $N_iBu/12:4\Delta^{2E,4E,8Z,10E/Z}$  within fatty acids showed that metronidazole obstructed alkamides biosynthesis significantly more than plastidal fatty acids [132]. The metronidazole inhibition study along with the patterns from stable isotope labeling fatty acids and lack of oleate incorporation led our group to look for a *de novo* non-chloroplastic fatty acid pathway for alkamides biosynthesis [132].

Mitochondria produce short-, medium-, and long-chain fatty acids ranging from  $C_8$ - $C_{18}$  with an even number of carbons [158]. This is achieved through the action of a type II fatty acid synthase (FAS). FAS are “type II” if they require each of the conserved enzymes  $\beta$ -ketoacyl-ACP synthase,  $\beta$ -ketoacyl-ACP reductase,  $\beta$ -hydroxyl-ACP dehydrase, and enoyl-ACP reductase as individual subunits. Different from plastidal FAS, the KAS in mitochondria catalyzes all of the condensation steps; it does not have chain length specificity [162]. Mitochondria FAS uses malonate formed by carboxylation of acyl-CoA units by a multifunctional ACCase in the cytosol [163]. The mfACCcase within the cytosol is different than the ACCcase present within the plastid as it can be inhibited by several classes of herbicides [132]. Select herbicides, fenoxaprop-p-ethyl and tepraloxydim, were used for probing mitochondria FAS. When 50 mM  $[U-^{13}C_6]$ glucose was fed with 100 nM concentrations of both inhibitors, labeling of BA 8/9 decreased by

88-89% [132]. There was a reduction in VLCFA (24:0) from 51.4% (control) to 30.3% (fenoxaprop-p-ethyl) and 36.8% (tepraloxymid) [132]. There was also a reduction observed in oleate from 19.2% to about 13.0% for both inhibitors [132]. It is worth noting that there was no reduction in labeling of plastidial fatty acids 18:2 $\Delta^{9Z,12Z}$  and 18:3 $\Delta^{9Z,12Z,15Z}$  [132]. This was interpreted to show there was a significant change of the cytosolic malonate pool by the mfACCase inhibitors that blocked alkamide labeling, therefore fatty acid synthesis coming from the mitochondria must contribute to the acyl chain formation of alkamide biosynthesis [132].

As stated previously, the mitochondrion produces three major acyl-ACPs, C<sub>8</sub>, C<sub>16</sub>, and C<sub>18</sub>, along with minor acyl-ACPs, C<sub>10</sub>, C<sub>12</sub>, and C<sub>14</sub>. Since the fatty acid synthesis from the mitochondria must be contributing to the acyl chain formation of alkamide biosynthesis, this describes the formation of even-chain length alkamides, but there are alkamides, such as <sup>N</sup>iBu/13:2 $\Delta^{2E,7Z,10a,12a}$ , that have an odd acyl chain length. It has been proposed that odd-chain alkamides could be formed through alpha-ketoester synthase ( $\alpha$ KES),  $\alpha$ -oxidation of fatty acids, or synthesis initiated with a propionyl-CoA starter unit [132]. The main source of fatty acid synthase is acetyl-CoA, but for uneven chain lengths, propionyl-CoA can be substituted [164, 165]. Propionyl-CoA is an uncommon precursor in eukaryotes except plants, it can be incorporated by *Corynebacterium glutamicum*, which uses *fasA* and *fasB* to synthesize uneven fatty acids 15:0, 17:0, 17:1 [166]. Higher plants and bacteria also use propionyl-CoA as an intermediate in polyketide synthesis [132, 167]. When 1 mM [U-<sup>13</sup>C<sub>3</sub>]propionate was fed to *E. purpurea* seedlings, but no M + 3 labeling occurred in either odd or even-chain alkamides and thus suggesting that propionyl-CoA is not a starter unit for odd-chain FAS in this circumstance [132].

Heteronuclear single-quantum coherence (HSQC) analysis was used to look at the pattern of [1,2-<sup>13</sup>C<sub>2</sub>]acetate incorporation of BA6 [132]. The most noticeable observation occurred at C13 with a lack of coupling partner [132]. If  $\alpha$ -oxidation was taking place C13 should couple with C12 in the acyl chain and the carboxyl carbon would be uncoupled to other incorporated carbon-13 atoms. On the contrary, this result supports the loss of coupling between C12 and C13 could occur via  $\omega$ -oxidation of the acyl chain [132].

In summary, studies by the Minto lab support a model that the acyl chain of the alkamides is formed through mitochondrial fatty acid biosynthesis. Diversification to produce odd-chain alkamides, such as BA6, could occur from  $\omega$ -oxidation instead of an initially hypothesized propionyl-CoA starter unit or  $\alpha$ -oxidation. The current evolution of our modular pathway for alkamides is shown in Figure 34 [132].

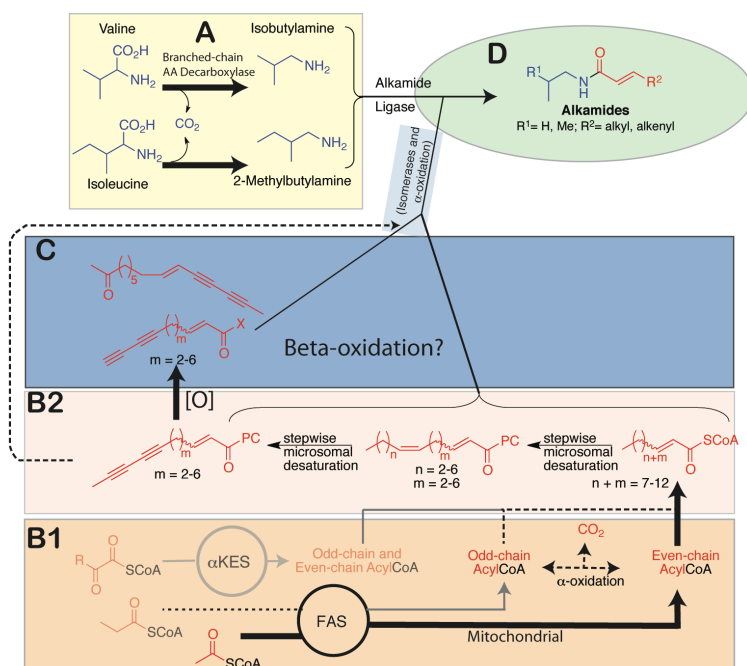


Figure 34: Proposed modular pathway for alkamides biosynthesis in *E. purpurea*.  
Reproduced with permission from [132].

### 3.2 Materials and Methods

Dichloromethane (DCM) was dried by distillation from  $\text{CaH}_2$ . THF and  $\text{Et}_2\text{O}$  were distilled from sodium and benzophenone. Benzene and hexanes were used as received. 1,3-Diaminopropane was dried by distillation from barium oxide. Unless otherwise stated, reagents were procured from Aldrich. Acetic acid, hydrogen peroxide and borane- $\text{Me}_2\text{S}$  complex were purchased from Fisher Scientific. Dess-Martin periodinane (DMP) was synthesized by the method of Boeckman [91]. Silica gel for chromatography from Dynamic Absorbents (60-200 micron, #020311) was slurry packed

for silica gel flash chromatography using the indicated solvent. All glassware was oven-dried at 120 °C and reactions run under nitrogen unless otherwise noted.

Stable isotope labeled small-molecule precursors [deuterated and  $^{13}\text{C}$  isotopologs of glucose (99%), valine (98%), isoleucine (98%), 2-methylpropylamine HCl (99.5%), and butyric acid (99%)] were obtained from Cambridge Isotope Laboratories (Tewksbury, MA). Chlorsulfuron was purchased from Fluka. *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) was purchased by Sigma Aldrich. The components used to make liquid and solid growth media were Plant Preservation Mixture (PPM) (Plant Cell Technology; Washington, DC), MS Basal medium with Gamborg vitamins (MP Biomedicals; Solon, OH), and sucrose and MES (MP Biomedicals; Waltham, MA). Murashige and Skoog basal salt mixture (MS) contains ammonium nitrate, boric acid, calcium chloride, cobalt chloride, cupric sulfate, disodium EDTA, ferrous sulfate, magnesium sulfate, manganese sulfate, molybdic acid, potassium sulfate, potassium nitrate, monobasic potassium phosphate, and zinc sulfate. The pH of the solid growth media was adjusted to 5.4 by 1M KOH. *E. purpurea* seeds were either purchased from Prairie Nursery (Westfield, WI) or provided by the North Central Regional Plant Introduction Station, USDA (Ames, IA).

$^1\text{H}$ ,  $^{13}\text{C}$ , DEPT, COSY, HSQC, HMBC, and NOESY NMR spectra were recorded on a Bruker Avance-II NMR spectrometer using  $\text{CDCl}_3$  as solvent; residual  $\text{CHCl}_3$  ( $^1\text{H}$ ,  $\delta$  7.26; 500 MHz) and  $\text{CDCl}_3$  ( $^{13}\text{C}$ ,  $\delta$  77.1; 125 MHz) were used as internal chemical shift reference. Peak multiplicities and shapes were annotated as follows: s, singlet; d, doublet; t, triplet; q, quartet; quint, quintet; m, multiplet; br, broad. Infrared spectra were recorded as neat films on NaCl plates with a Fourier-transform infrared spectrophotometer (Nicolet Avatar 330). High-resolution mass spectra were collected with an Agilent 6520 Q-ToF mass spectrometer. Mass spectra for fatty acid methyl ester analysis was collected via an Agilent 7890A GC/5975C MS, and mass spectra from labeled isotope feeding experiments were obtained with an Agilent 6410 Triple Quadrupole MS. Sample preparation for FAME and alkamide analyses of *E. purpurea* was completed in glass conical tubes. An IEC HN-SII centrifuge was used at 2000 rpm during FAME preparation and other analyses unless otherwise stated.

### 3.2.1 Germination and growth of *E. purpurea* seedlings

*E. purpurea* seeds (Prairie Nursery, Westfield, WI) were cold-stratified in a 10% aqueous solution of Plant Preservation Mixture (PPM; Plant Cell Technologies, Washington, DC) for 16 h at 4 °C. The solution was decanted and seeds were washed sequentially with 70% ethanol (aq.), 0.5% NaOCl bleach solution containing 0.02% Tween-20, and repeatedly with autoclaved nanopure H<sub>2</sub>O until the odor of bleach was eliminated. Typically, about 2 g of seed were treated with 10 mL of each solution. The seeds were then placed on ½-strength MS agar plates containing sucrose (11 mM), MES (2.56 mM), pH 5.4, PPM (0.2%) and placed in an incubator at 22 °C with a 18 h light/6 h dark cycle. The light intensity was 100- $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  at the level of the plants. Seedlings for labeling experiments were staged for use when the first leaf unfurled (“Stage 5”).



Figure 35: Stage 5 of *E. purpurea* seedlings used in feeding experiments

### 3.2.2 Alkamide analysis

Alkamides were extracted using a method modified by Hudaib et al. [168]. The seedlings were homogenized and ground with two aliquots of 70% v/v methanol. Samples were extracted with 2 mL n-hexane:ethyl acetate (1:1 v/v), vortexed, and centrifugation at 3500 rpm for 5 min. The organic phases were collected, pooled, and evaporated in a heated block (50-60 °C) under a stream of N<sub>2</sub> gas. For LC-MS, analysis samples were dissolved in 200  $\mu\text{L}$  of hexanes. For GC-MS analysis, the dried samples were dissolved in 1 ml of acetonitrile, and silylated by incubating at 70 °C for 20 min with 70  $\mu\text{l}$  of *N,O*-bis(trimethylsilyl)trifluoroacetamide containing 1%

chlorotrimethylsilane. After evaporating the excess reagents under a stream of N<sub>2</sub> gas, the samples were dissolved in acetonitrile.

### 3.2.3 Isotopic labeling experiments

Stable isotope incorporation experiments were performed in sterile 90-mm polystyrene Petri dishes each containing a 70-mm Whatman No. 1 filter disc and flooded with 6 mL of ½-strength MS containing MES (2.56 mM), PPM (0.2%). Chlorsulfuron (30 nM), methyl jasmonate (10 µM), amino acids (15 mM) and other precursors were added to the MS medium as required. Seedlings were placed on top of the paper and incubated at the germination conditions for 3 days. To synthesize deuterium-labeled alkamides and fatty acids, experiments were performed in sterile 90-mm glass Petri dish without a filter disc. Six mL of ½-strength MS containing MES (2.56 mM) and PPM (0.2%) liquid media with methyl jasmonate (10 µM) was placed in the dish. The d<sub>3</sub>-labeled lipids were dissolved in DMSO (50 mM) and 60 µL was added to the liquid media. L-valine-d<sub>8</sub> (3.6 mg), used as an internal standard, was added to the liquid media.

For LC-MS analysis, three seedlings were homogenized and extracted using a mortar and pestle with two aliquots of 70% (v/v) aqueous methanol. After centrifugation for 5 min, the pooled supernatants were extracted with two 2-mL aliquots of hexane-EtOAc (1:1 v/v). The organic phases were collected and pooled for analysis. The MS parameters for triple quadrupole analysis (Agilent 6410) were as follows: MS2 full scan method, ESI positive mode, Fragmentor voltage: 135 V, Gas temperature: 300 °C, Gas flow: 10 L/min, Nebulizer pressure: 45 psi, Capillary: +/- 4000 V. The chromatographic parameters were: Waters XBridge-C<sub>18</sub> column (4.6 x 150 mm, 5 µm particle size); Solvent A: H<sub>2</sub>O + 0.1% formic acid. Solvent B: CH<sub>3</sub>CN + 0.1% formic acid; Flow rate: 1 mL/min; 70 → 100% B over 10 min with a 1-min hold and then ramped back to 70% B over 3 min; post time: 2 min.

For GC/MS, the organic extracts were evaporated under a stream of N<sub>2</sub> gas, dissolved in 1 mL of acetonitrile and silylated by incubating at 70 °C for 20 min with 70 µL of *N,O*-bis(trimethylsilyl)trifluoroacetamide containing 1% chlorotrimethylsilane.

After evaporating the excess reagents under a stream of N<sub>2</sub> gas, the samples were dissolved in 200 µL of acetonitrile and subjected to GC/MS analysis.

For toxicity tests stock solutions of the synthesized compounds, 14:0 and <sup>N</sup>iBu/14:1Δ<sup>2E</sup> were made with 2-methoxyethanol. Dilutions from the stock solutions were made and 100, 50, 30, 20, and 5 mM concentrations were tested with *E. purpurea*. Two seedlings per concentration were used and 2 µL of the respective concentration were placed on both leaves of a stage 5 seedling; they were then grown for 3 days.

### 3.2.4 Equations used to calculate incorporation levels

Two different equations were used to determine the isotopic incorporation levels in the analyzed natural products; the first determined the percentage of labeling across a complete isotopic envelope while the second was used to calculate the relative intensity of a specific isotopic peak by percentage. Equation 1 uses the sum of all abundances from mass isotopic peaks (M + 1, M + 2, M + 3...etc.) divided by the whole envelope. For the second equation, the abundance of the specific ion, for example M + 2, was divided by the parent ion.

$$\% \text{ envelope labeled} = \frac{[(M+1) + (M+2) + (M+3)...]}{[(M) + (M+1) + (M+2) + (M+3)...]} \times 100 \quad (1)$$

$$\% \text{ isotopic peak} = \frac{[(M+n)]}{[(M)]} \times 100 \quad (2)$$

### 3.2.5 Fatty acid methyl ester preparation for MS analysis

Seedlings were also analyzed for their fatty acid content. Fatty acids from the plant were converted to their respective methyl esters for GC/MS analysis. Fatty acid methyl esters (FAMES) were prepared by the addition of 1 mL of 2% H<sub>2</sub>SO<sub>4</sub> in methanol to three seedlings in a glass conical vial. The mixture was heated at 80 °C for 2 h. The reaction was quenched with 1 mL of H<sub>2</sub>O and extracted with two 1-mL hexanes. The



organic layers were combined and evaporated at 60 °C under a stream of N<sub>2</sub> gas before dissolving in 200 µL of hexane for GC/MS analysis.

### 3.2.5.1 Protocols for alkamide and FAME analysis by GC/MS and LC/MS

Chromatography by either gas or liquid chromatography was used for analyte separation in most experiments. Samples were injected via the system autosampler. The column and oven programs for GC/MS analysis were as follows.

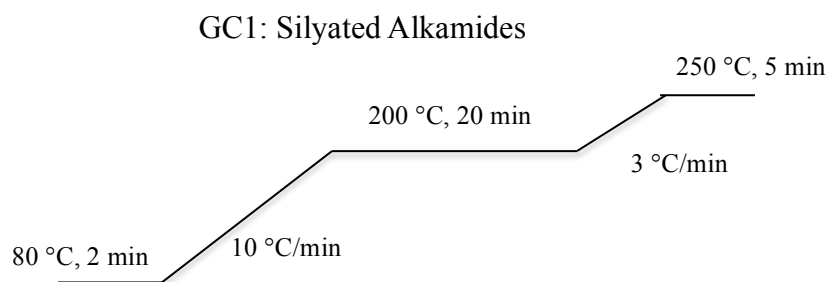


Figure 36: GC oven programs for alkamide analysis (GC1).

Alkamides analysis were suspended in ACN, silylated with TMS, and analyzed with a HP-5 (30 m x 250 µm x 0.25 µm) column.

#### **GC/MS parameters were as follows:**

- Split-Splitless Inlet: Splitless
- Inlet Temperature: 250 °C
- Pressure: 15.408 psi
- Septum Purge Flow: 3 mL/min
- He Flow Rate: 1.56 mL/min for 56 min

The scan parameters were as follows:

- Solvent Delay: 6 min
- Scan Speed: Normal
- Start Mass: 50.00 amu; Stop Mass: 800.00 amu
- Threshold: 150; Scans/sec: 1.99

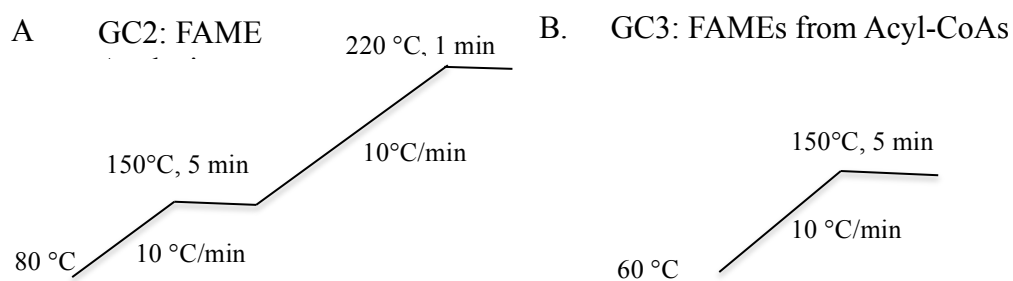


Figure 37: GC oven programs for FAMES analysis from fatty acids (GC2) and FAMES analysis from acyl-CoA (GC3).

FAMES, including those determined during the analysis of acyl-CoAs, were suspended in hexanes and analyzed with a VF23 (30 m x 250  $\mu$ m x 0.25  $\mu$ m) column.

GC/MS parameters were as follows:

- Split-Splitless Inlet: Splitless
- Inlet Temperature: 250 °C
- Pressure: 16.834 psi
- Septum Purge Flow: 3 mL/min;
- He Flow Rate: A. 1.89 mL/min for 14 min; B. 1.89 mL/min for 24.6 min

The scan parameters were as follows:

- Solvent Delay: 6 min
- Scan Speed: Normal
- Start Mass: 50.00 amu; Stop Mass: 800.00 amu
- Threshold: 150; Scans/sec: 1.99

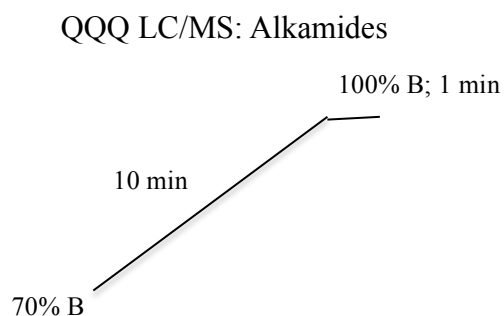


Figure 38: LC method for extracted alkamides.

Alkamide analysis was completed on a triple quadrupole MS. Samples were injected via the autosampler with the use of a Waters XBridge-C<sub>18</sub> column (4.6 x 150 mm, 5 µm particle size). The mobile phase parameters are listed below.

The source parameters were set as follows:

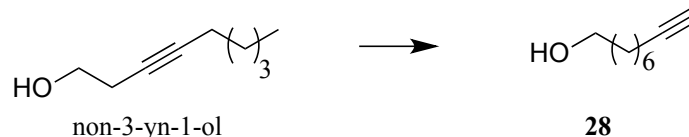
- MS2 full scan method
- ESI positive mode
- Fragmentation: 135 V
- Gas temp: 300 °C
- Gas flow: 10 L/min
- Nebulizer: 45 psi
- Capillary: +/- 4000 V

The parameters for the chromatograph were:

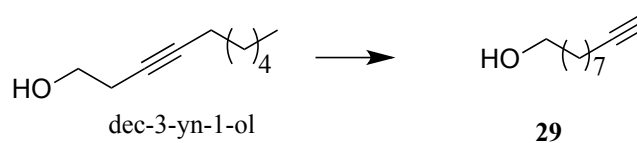
- Solvent A: H<sub>2</sub>O + 0.1% formic acid. Solvent B: CAN + 0.1% formic acid
- Flow rate: 1 mL/min; 70 → 100% B over 10 min with a 1-min hold and then ramped back to 70% B during the completion of the run
- Run time: 14 min
- Post time: 2 min

### 3.2.6 Synthesis of aldehydes and [d<sub>3</sub>]-aldehydes

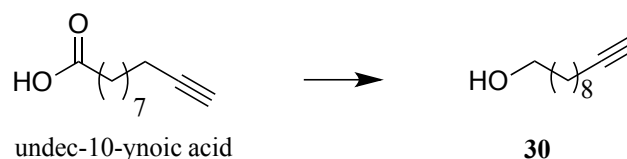
**General synthesis of terminal alkynes:** A three-necked flask was equipped with a stopper, an addition funnel, and condenser. Dry 1,3-diaminopropane (30 mL, 360 mmol, 36 equiv) was added to the round-bottom flask under N<sub>2</sub>. Lithium (0.4200 g, 60.52 mmol, 6 equiv) was rinsed with hexane and cut into small pieces before being added to the 1,3-diaminopropane. The mixture was stirred at RT for 30 min and then heated at 70 °C for 4 h. The lithium reacted giving an initial blue color and then a white solid formed giving the suspension a gray-blue color. To this mixture, *t*-BuOK (4.4200 g, 39.4 mmol, 4 equiv) was added. The reaction was stirred for 20 min at room leading to a pale green solution and then the alcohol (9.98 mmol) was added dropwise over a 10-minute period. The addition funnel was washed with 1,3-diaminopropane (2 mL). After stirring for 30 min at room temperature, the reddish-orange solution was slowly poured over 100 mL of an ice/water mixture. The quenched reaction was extracted with hexane (4 x 50 mL). The organic phases were washed once with H<sub>2</sub>O (100 mL), 10% HCl (100 mL), and brine (100 mL). The organic layer was dried over sodium sulfate and solvent was removed by reduced pressure.



**8-Nonyn-1-ol (28):** Clear oil (72%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) 3.63 (t,  $J = 6.5$  Hz, 2H), 2.18 (td,  $J = 7.1, 2.5$  Hz, 2H), 1.93 (t,  $J = 2.5$  Hz, 1H), 1.47-1.63 (m, 4H), 1.23-1.44 (m, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  84.7, 68.1, 63.0, 32.7, 28.9, 28.7, 28.4, 25.6, 18.4.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [169].



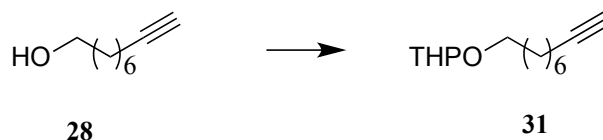
**9-Decyn-1-ol (29):** Clear oil (80%). NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  3.64 (t,  $J = 6.6$  Hz, 2H), 2.18 (td,  $J = 7.1, 2.6$  Hz, 2H), 1.93 (t,  $J = 2.6$  Hz, 1H), 1.49-1.60 (m, 4H), 1.31-1.43 (m, 8H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  84.7, 68.1, 63.0, 32.8, 29.3, 29.0, 28.7, 28.4, 25.7, 18.4.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [170].



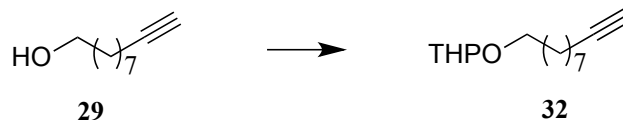
**10-Undecynol (30):** Lithium aluminum hydride (0.5588 g, 14.7 mmol, 2.1 equiv) was dissolved in dry THF (40 mL) in a 100-mL round-bottom flask. Undec-10-ynoic acid (1.2687 g, 6.96 mmol) dissolved in THF (10 mL) was added slowly and  $\text{H}_2$  gas was given off. The solution stirred at room temperature for 1 h. The reaction was quenched with  $\text{H}_2\text{O}$  (10 mL). The organic layer was washed with 15% NaOH (7 mL) and  $\text{H}_2\text{O}$  (5 mL). The salt was collected and washed with EtOAc. The organic layer was dried with  $\text{MgSO}_4$  and solvent was removed under reduced pressure to yield 10-undecynol (1.0797 g, 92%). IR (neat)  $\bar{\nu}$  3309, 2929, 2855, 2116, 1462, 1056;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  3.64 (t,  $J = 6.8$  Hz, 2H), 2.18 (td,  $J = 7.1, 2.7$  Hz, 2H), 1.94 (t,  $J = 2.5$  Hz, 1H), 1.49-1.59

(m, 4H), 1.30-1.42 (m, 10H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  84.8, 68.1, 63.1, 32.8, 29.4, 29.3, 29.0, 28.7, 28.5, 25.7, 18.4. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{11}\text{H}_{21}\text{O}$  169.1587, found 169.1612.

**General procedure to synthesize THP-protected alcohols:** Alcohol (5.50 mmol) was dissolved in dry DCM (3.5 mL). DHP (6.60 mmol, 1.2 equiv) was added, and the solution stirred for 30 min. TsOH (0.55 mmol, 0.1 equiv) was added and reaction was stirred at room temperature for 20 h. The reaction was quenched by the addition of 1 M NaOH (3 mL) and mixture was allowed to stir for 30 min. The mixture was extracted with DCM, washed with brine, dried with  $\text{MgSO}_4$ , and solvent was removed by rotary evaporation to give a dark red solution. The crude product was purified by column chromatography with hexane: EtOAc (5:1).

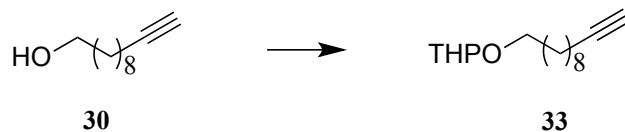


**2-(8-Nonyn-1-oxy)tetrahydro-2H-pyran (31):** Clear oil (96%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (dd,  $J = 2.8, 4.3$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 2.18 (td,  $J = 7.1, 2.7$  Hz, 2H), 1.93 (t,  $J = 2.7$  Hz, 1H), 1.82 (m, 1H), 1.71 (m, 1H), 1.50-1.63 (m, 8H), 1.31-1.44 (m, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 84.7, 68.1, 67.6, 62.4, 30.8, 29.7, 28.9, 28.7, 28.4, 26.1, 25.5, 19.7, 18.4.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [171].



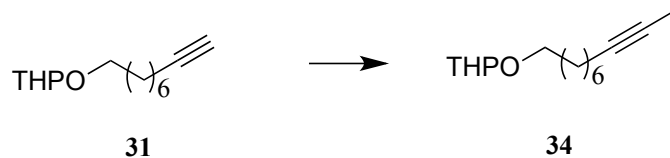
**2-(9-Decyn-1-oxy)tetrahydro-2H-pyran (32):** Clear oil (93%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (t,  $J = 3.6$  Hz, 1H), 3.87 (m, 1H), 3.72 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 2.17 (td,  $J = 7.1, 2.6$  Hz, 2H), 1.93 (t,  $J = 2.6$  Hz,

1H), 1.82 (m, 1H), 1.71 (m, 1H), 1.49-1.61 (m, 8H), 1.32-1.42 (m, 8H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 98.9, 84.8, 68.0, 67.6, 62.4, 30.8, 29.7, 29.3, 29.0, 28.7, 28.5, 26.2, 25.5, 19.7, 18.4. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [170].



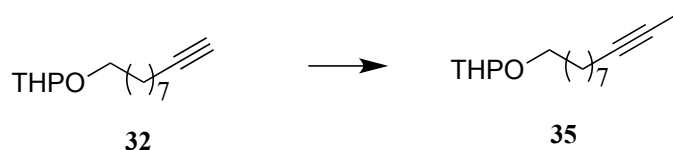
**2-(10-Undecyn-1-oxy)tetrahydro-2H-pyran (33):** Clear oil (88%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 4.57 (dd, *J* = 3, 4.5 Hz, 1H), 3.85-3.89 (m, 1H), 3.73 (dt, *J* = 9.5, 7.0 Hz, 1H), 3.48-3.52 (m, 1H), 3.38 (dt, *J* = 9.5, 6.8 Hz, 1H), 2.18 (td, *J* = 7.3, 2.8 Hz, 2H), 1.93 (t, *J* = 2.8 Hz, 1H), 1.80-1.84 (m, 1H), 1.74-1.69 (m, 1H), 1.49-1.62 (m, 8H), 1.28-1.41 (m, 10H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 98.9, 84.8, 68.1, 67.7, 62.3, 30.8, 29.7, 29.42, 29.39, 29.0, 28.7, 28.5, 26.2, 25.5, 19.7, 18.4. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [172].

**General procedure of methylation and d<sub>3</sub>-methylation:** THP-protected alcohol (3.0 mmol) was dissolved in THF (5 mL). The solution was cooled to 0 °C and *n*-BuLi (2.4 M, 4.5 mmol, 1.5 equiv) was added slowly followed by the addition of MeI (4.5 mmol, 1.5 equiv). The solution was stirred at 0 °C for 3 h after which H<sub>2</sub>O (15 mL) was added. The mixture was extracted with hexane, washed with H<sub>2</sub>O, and brine. The organic layer was dried with MgSO<sub>4</sub> and the solvent was removed.

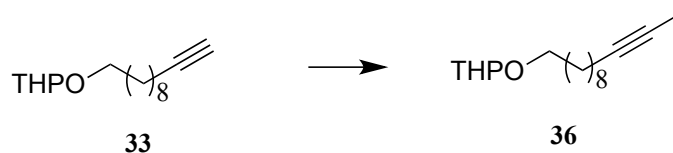


**2-(8-Decyn-1-oxy)tetrahydro-2H-pyran (34):** Clear oil (58%). IR (neat)  $\bar{\nu}$  2934, 2858, 1136, 1120, 1078, 1033, 986; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 4.57 (dd, *J* = 2.9, 4.3 Hz, 1H), 3.87 (m, 1H), 3.73 (dt, *J* = 9.6, 6.9, 1H), 3.50 (m, 1H), 3.38 (dt, *J* = 9.6, 6.7, 1H), 2.11 (m, 2H), 1.84 (m, 1H), 1.77 (t, *J* = 2.6 Hz, 3H), 1.71 (m, 1H), 1.48- 1.58 (m, 6H),

1.47 (quint,  $J = 7.3$  Hz, 2H), 1.28-1.40 (m, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 79.4, 75.3, 67.6, 62.3, 30.8, 29.7, 29.03, 29.02, 28.8, 26.2, 25.5, 19.7, 18.7, 3.5.

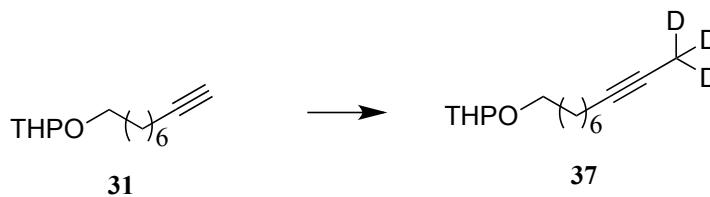


**2-(9-Undecyn-1-oxo)tetrahydro-2H-pyran (35):** Yellow oil (95%). IR (neat):  $\bar{\nu}$  2932, 2856, 1454, 1351, 1200, 1136, 1120, 1078, 1032, 988;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (t,  $J = 3.6$  Hz, 1H), 3.87 (m, 1H), 3.72 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 2.10 (m, 2H) 1.83 (m, 1H), 1.78 (t,  $J = 2.5$  Hz, 3H), 1.71 (m, 1H), 1.50-1.62 (m, 6H), 1.46 (quint,  $J = 7.3$  Hz, 2H), 1.30-1.39 (m, 8H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 79.4, 75.3, 67.7, 62.3, 30.8, 29.7, 29.4, 29.1, 29.0, 28.8, 26.2, 25.5, 19.7, 18.7, 3.5. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{16}\text{H}_{29}\text{O}_2$  253.2162 found 253.2157.

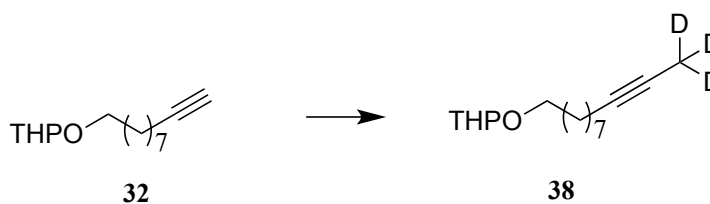


**2-(10-Dodecyn-1-oxo)tetrahydro-2H-pyran (36):** Yellow oil (46%). IR (neat)  $\bar{\nu}$  2929, 2855, 1454, 1351, 1200, 1136, 1120, 1078, 1032, 988;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (t,  $J = 3.5$  Hz, 1H), 3.84-3.89 (m, 1H), 3.72 (dt,  $J = 10.0, 6.8$  Hz, 1H), 3.47-3.51 (m, 1H), 3.37 (dt,  $J = 9.5, 6.8$  Hz, 1H), 2.10 (m, 2H), 1.79-1.86 (m, 1H), 1.77 (t,  $J = 2.5$  Hz, 3H), 1.68-1.74 (m, 1H), 1.49-1.58 (m, 6H), 1.46 (quint,  $J = 7.3$  Hz, 2H), 1.25-1.38 (m, 10H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 79.4, 75.3, 67.7, 62.3, 30.8, 29.8, 29.5, 29.4, 29.1, 29.0, 28.9, 26.2, 25.5, 19.7, 18.7, 3.5.

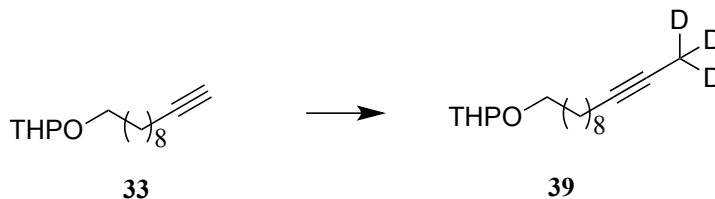




**[10,10,10-<sup>2</sup>H<sub>3</sub>]-2-(8-Decyn-1-oxy)tetrahydro-2*H*-pyran (37):** Yellow oil (79%). IR (neat)  $\bar{\nu}$  2934, 2857, 2216, 2213, 1163, 1120, 1078, 1033; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  4.57 (dd,  $J = 3.0, 4.3$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 2.11 (t,  $J = 7.1$  Hz, 2H), 1.83 (m, 1H), 1.71 (m, 1H), 1.50-1.60 (m, 6H), 1.47 (quint,  $J = 7.4$  Hz, 2H), 1.28-1.41 (m, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  98.9, 79.4, 75.3, 67.7, 62.3, 30.8, 29.7, 29.03, 29.02, 28.8, 26.2, 25.5, 19.7, 18.7.

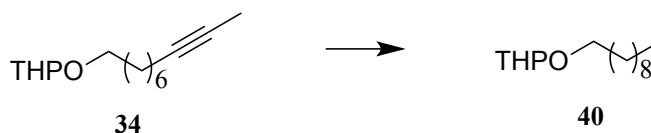


**[11,11,11-<sup>2</sup>H<sub>3</sub>]-2-(9-Undecyn-1-oxy)tetrahydro-2*H*-pyran (38):** Yellow oil (73%). IR (neat)  $\bar{\nu}$  2932, 2856, 2213, 2113, 1200, 1136, 1120, 1078, 1032, 988; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  4.57 (t,  $J = 3.6$  Hz, 1H), 3.84-3.89 (m, 1H), 3.72 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.47-3.52 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 2.11 (t,  $J = 7.1$  Hz, 2H), 1.78-1.86 (m, 1H), 1.68-1.74 (m, 1H), 1.50-1.62 (m, 6H), 1.46 (quint,  $J = 7.3$  Hz, 2H), 1.25-1.39 (m, 8H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  98.9, 79.4, 75.3, 67.7, 62.3, 30.8, 29.7, 29.4, 29.1, 29.0, 28.8, 26.2, 25.5, 19.7, 18.7.

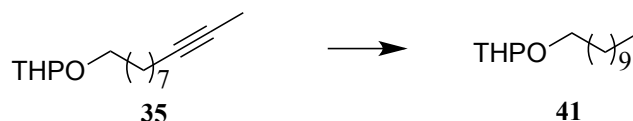


**[12,12,12-<sup>2</sup>H<sub>3</sub>]-2-(10-Dodecyn-1-oxy)tetrahydro-2*H*-pyran (39):** Yellow oil (88%). IR (neat)  $\bar{\nu}$  2930, 2855, 2213, 2113, 1464, 1351, 1200, 1136, 1120, 1078, 1033, 988; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  4.57 (dd,  $J = 3.0, 4.0$  Hz, 1H), 3.84-3.89 (m, 1H), 3.72 (dt,  $J = 9.5, 6.9$  Hz, 1H), 3.47-3.51 (m, 1H), 3.37 (dt,  $J = 9.5, 6.8$  Hz, 1H), 2.10 (t,  $J = 7.3$  Hz, 2H) 1.79-1.86 (m, 1H), 1.68-1.74 (m, 1H), 1.49-1.61 (m, 6H), 1.46 (quint,  $J = 7.8$  Hz, 2H), 1.23-1.38 (m, 10H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  98.9, 79.4, 75.3, 67.7, 62.3, 30.8, 29.8, 29.5, 29.4, 29.1, 29.0, 28.9, 26.2, 25.5, 19.7, 18.7.

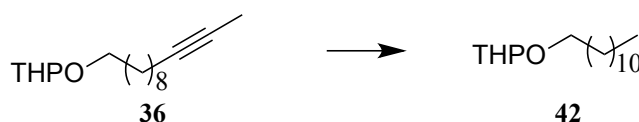
**General procedure for the reduction of alkynes:** NaI (5.76 mmol, 3 equiv) was dissolved in THF (3 mL) and cooled to -20 °C in an acetone/dry ice bath. The Wilkinson's catalyst (0.194 mmol, 0.1 equiv) was added. The acetylenic compound (0.4964 g, 1.94 mmol) was added slowly. After purging with argon, the mixture was flushed with hydrogen and heated at 45 °C for 2.5 h. The mixture was cooled to room temperature and was stirred for 17 h. Excess catalyst was filtered and the THF removed. The crude product was purified by column chromatography with hexane: EtOAc (5:1).



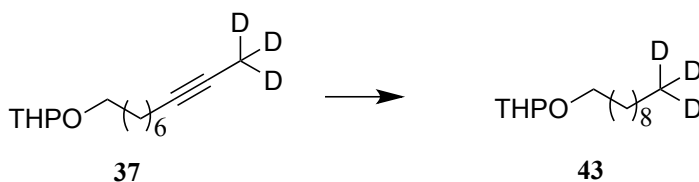
**2-Decanoxytetrahydro-2*H*-pyran (40):** Clear oil (63%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  4.58 (t,  $J = 3.5$  Hz, 1H), 3.88 (m, 1H), 3.73 (dt,  $J = 9.5, 7.0$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.5, 6.8$  Hz, 1H), 1.83 (m, 1H), 1.72 (m, 1H), 1.49-1.62 (m, 6H), 1.23-1.39 (m, 14H), 0.88 (t,  $J = 6.9$  Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  98.9, 67.7, 62.3, 31.9, 30.8, 29.8, 29.6, 29.5, 29.4, 29.3, 26.2, 25.5, 22.7, 19.7, 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [173].



**2-Undecanoxytetrahydro-2H-pyran (41):** Clear oil (65%). IR (neat):  $\bar{\nu}$  2925, 2854, 1464, 1352, 1260, 1201, 1127, 1079, 1034, 988, 870, 814;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (t,  $J = 3.6$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.5, 6.9$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 1.82 (m, 1H), 1.71 (m, 1H), 1.49-1.62 (m, 6H), 1.24-1.35 (m, 16H), 0.88 (t,  $J = 6.9$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.8, 67.7, 62.3, 31.9, 30.8, 29.8, 29.62, 29.60, 29.5, 29.4, 29.3, 26.2, 25.5, 22.7, 19.7, 14.11.

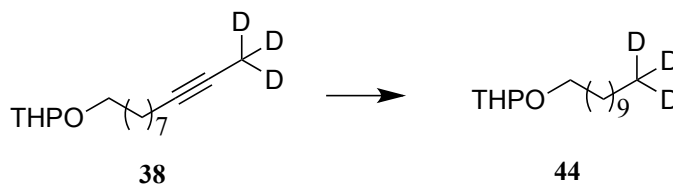


**Dodecanoxytetrahydro-2H-pyran (42):** Clear oil (90%). IR (neat)  $\bar{\nu}$  2924, 2853, 1119, 1078, 1033;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (dd,  $J = 3.0, 4.0$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.5, 7.0$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.5, 6.8$  Hz, 1H), 1.80-1.86 (m, 1H), 1.67-1.74 (m, 1H), 1.50-1.62 (m, 6H), 1.24-1.35 (m, 18H), 0.88 (t,  $J = 6.5$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 67.7, 62.3, 31.9, 30.8, 30.3, 29.8, 29.7, 29.6, 29.6, 29.5, 29.4, 26.3, 25.5, 22.7, 19.7, 14.11. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+$   $\text{C}_{17}\text{H}_{35}\text{O}_2$  271.2632, found 271.2632.

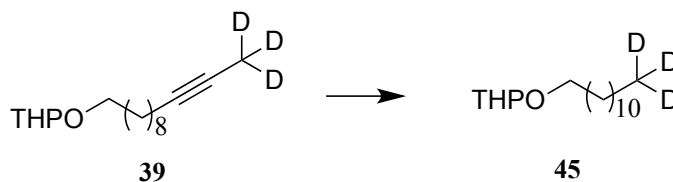


**[10,10,10- $^2\text{H}_3$ ]-2-Decanoxytetrahydro-2H-pyran (43):** Clear oil (78%). IR (neat)  $\bar{\nu}$  2924, 2854, 2212, 1122, 1078, 1033;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.58 (dd,  $J = 2.9, 4.3$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.6, 6.9$  Hz, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H),

1.83 (m, 1H), 1.71 (m, 1H), 1.49-1.62 (m, 6H), 1.23-1.38 (m, 14H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 67.7, 62.3, 31.8, 30.8, 29.8, 29.62, 29.60, 29.5, 29.4, 26.2, 25.5, 22.4, 19.7.



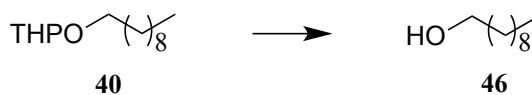
**[11,11,11- $^2\text{H}_3$ ]-2-Undecanoxytetrahydro-2H-pyran (44):** Clear oil (71%). IR (neat)  $\bar{\nu}$  2924, 2854, 2213, 1464, 1351, 1200, 1183, 1122, 1078, 1033, 988, 905, 869;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.58 (t,  $J = 3.6$  Hz, 1H), 3.87 (m, 1H), 3.73 (dt,  $J = 9.6, 7.0$  Hz, 1H), 3.50 (m, 1H), 3.38 (dt,  $J = 9.6, 6.7$  Hz, 1H), 1.82 (m, 1H), 1.71 (m, 1H), 1.49-1.62 (m, 6H), 1.25-1.35 (m, 16H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.9, 67.7, 62.3, 31.8, 30.8, 29.8, 29.62, 29.60, 29.5, 29.4, 26.2, 25.5, 22.4, 19.7. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{16}\text{H}_{30}\text{D}_3\text{O}_2$  274.2820, found 274.2819.



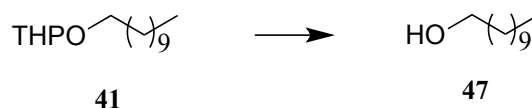
**[12,12,12- $^2\text{H}_3$ ]-2-Dodecanoxytetrahydro-2H-pyran (45):** Light yellow oil (49%). IR (neat)  $\bar{\nu}$  2924, 2853, 2213, 1122, 1078, 1032;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  4.57 (dd,  $J = 3.8, 4.3$  Hz, 1H), 3.85-3.90 (m, 1H), 3.73 (dt,  $J = 9.5, 7.0$  Hz, 1H), 3.48-3.52 (m, 1H), 3.38 (dt,  $J = 9.5, 6.6$  Hz, 1H), 1.80-1.86 (m, 1H), 1.68-1.74 (m, 1H), 1.49-1.62 (m, 6H), 1.23-1.38 (m, 18H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  98.8, 67.7, 62.3, 31.8, 30.8, 29.8, 29.7, 29.64, 29.62, 29.61, 29.5, 29.4, 26.3, 25.5, 22.4, 19.7.

**General procedure for removing a THP-protecting group:** The protected alcohol (0.4938 g, 2.01 mmol) was dissolved in methanol (12 mL) in a 25-mL round-bottom flask. Solid *p*-toluenesulfonic acid monohydrate (0.0760 g, 0.400 mmol, 0.2 equiv)

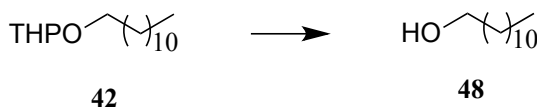
was added. The reaction was stirred at RT for 2 h before NaOH (1 M, 10 mL) was added followed by H<sub>2</sub>O (26 mL). The mixture was extracted with diethyl ether (10 mL). The combined organic layers were washed with brine (10 mL), dried with MgSO<sub>4</sub>, and ether was removed by rotary evaporation.



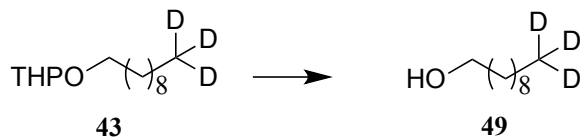
**Decanol (46):** Clear oil (91%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 3.63 (t, *J* = 6.6 Hz, 2H), 1.56 (quint, *J* = 6.9 Hz, 2H), 1.24-1.37 (m, 14H), 0.88 (t, *J* = 6.8 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 63.1, 32.8, 31.9, 29.6, 29.5, 29.4, 29.3, 25.8, 22.7, 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [174].



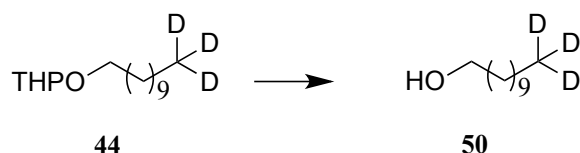
**Undecanol (47):** Clear oil (61%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 3.64 (t, *J* = 6.6 Hz, 2H), 1.56 (quint, *J* = 6.7 Hz, 2H), 1.24-1.37 (m, 16H), 0.87 (t, *J* = 6.6 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 63.1, 32.8, 31.9, 29.61, 29.60, 29.4, 29.3, 25.8, 22.7, 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [175].



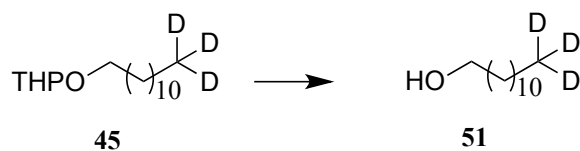
**Dodecanol (48):** Clear oil (54%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 3.64 (t, *J* = 6.8 Hz, 2H), 1.56 (quint, *J* = 6.8 Hz, 2H), 1.22-1.35 (m, 18H), 0.88 (t, *J* = 6.8 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 63.1, 32.8, 31.9, 29.7, 29.64, 29.62, 29.60, 29.4, 29.3, 25.7, 22.6, 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [176].



**[10,10,10-<sup>2</sup>H<sub>3</sub>]-Decanol (49):** Clear oil (76%). IR (neat)  $\bar{\nu}$  3336 (br), 2924, 2854, 2212, 1463, 1056; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  3.64 (t,  $J$  = 6.7 Hz, 2H), 1.57 (quint,  $J$  = 7.0 Hz, 2H), 1.21-1.36 (m, 14 H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  63.1, 32.8, 31.8, 29.6, 29.5, 29.4, 29.3, 25.7, 22.4.



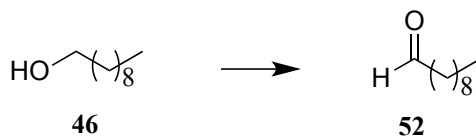
**[11,11,11-<sup>2</sup>H<sub>3</sub>]-Undecanol (50):** Light yellow oil (93%). IR (neat):  $\bar{\nu}$  3330 (br), 2924, 2854, 2213, 1464, 1056; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  3.63 (t,  $J$  = 6.7 Hz, 2H), 1.56 (quint,  $J$  = 7.0 Hz, 2H), 1.26-1.38 (m, 16H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  63.1, 32.8, 31.8, 29.61, 29.60, 29.59, 29.4, 29.3, 25.7, 22.4.



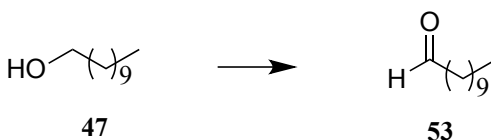
**[12,12,12-<sup>2</sup>H<sub>3</sub>]-Dodecanol (51):** Light yellow oil (72%) IR (neat)  $\bar{\nu}$  3329, 2924, 2854, 2213, 1463, 1056; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  3.64 (t,  $J$  = 6.6 Hz, 2H), 1.56 (quint,  $J$  = 6.8 Hz, 2H), 1.25 (m, 18H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  63.1, 32.8, 31.8, 29.7, 29.64, 29.62, 29.60, 29.4, 29.3, 29.2, 25.8, 22.4.

**General procedure for the oxidation of alcohols to aldehydes:** DMP (0.7144 g, 1.68 mmol, 1.1 equiv) was dissolved in DCM (10 mL) in a 50-mL round-bottom flask. In a separate vial, the respective alcohol (0.2462 g, 1.53 mmol) was dissolved in DCM (10 mL) and added to the stirred DMP solution. The reaction stirred at room temperature for

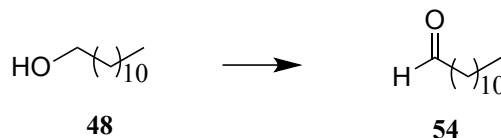
2 h. Et<sub>2</sub>O (25 mL) was added followed by NaOH (1 M, 10 mL), and then the mixture stirred for 10 min at room temperature. The Et<sub>2</sub>O layer was extracted, washed with H<sub>2</sub>O, dried with MgSO<sub>4</sub>, and solvent was removed. The crude product was purified by column chromatography with hexane: EtOAc (1:1).



**Decanal (52):** Clear oil (87%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) 9.76 (t, *J* = 1.8 Hz, 1H), 2.41 (dt, *J* = 1.8, 7.4 Hz, 2H), 1.62 (quint, *J* = 7.0 Hz, 2H), 1.25-1.30 (m, 12H), 0.87 (t, *J* = 6.7 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 203.0, 43.9, 31.9, 29.4, 29.3, 29.2, 29.1, 22.7, 22.1, 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [177].

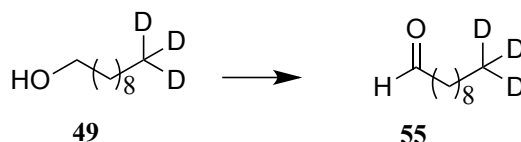


**Undecanal (53):** Clear oil (86%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 9.76 (t, *J* = 1.9 Hz, 1H), 2.41 (t, *J* = 1.9, 7.3 Hz, 2H), 1.62 (quint, *J* = 7.0 Hz, 2H), 1.25-1.30 (m, 14H), 0.88 (t, *J* = 6.7 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 203.0, 43.9, 31.9, 29.6, 29.5, 29.4, 29.2, 29.1, 22.7, 22.1, 14.1. <sup>1</sup>H data were similar to that previously reported [178].

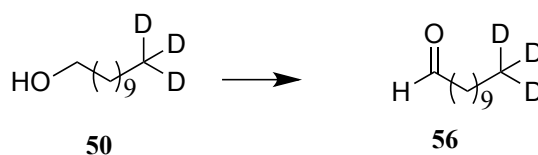


**Dodecanal (54):** Clear oil (94%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 9.96 (t, *J* = 1.8 Hz, 1H), 2.41 (dt, *J* = 1.8, 7.4 Hz, 2H), 1.62 (quint, *J* = 7.1 Hz, 2H), 1.22-1.38 (m, 16H), 0.87 (t, *J* = 6.7 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 203.0, 43.9, 31.9, 29.9, 29.6,

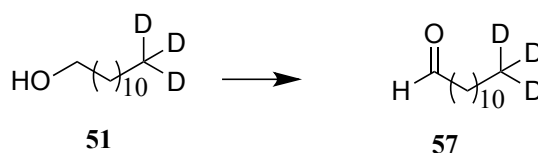
29.4, 29.3, 29.2, 24.7, 22.7, 22.1, 14.1.  $^1\text{H}$  and  $^{13}\text{C}$  data were similar to that previously reported [179]



**[10,10,10- $^2\text{H}_3$ ]-Decanal (55):** Light yellow (58%). IR (neat)  $\bar{\nu}$  2926, 2855, 2713, 2212, 2085, 1728, 1466, 948;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (t,  $J = 1.9$  Hz, 1H), 2.42 (td,  $J = 7.4, 1.9$  Hz, 2H), 1.63 (quint,  $J = 7.3$  Hz, 2H), 1.21-1.33 (m, 12H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  202.9, 43.9, 31.8, 29.4, 29.35, 29.25, 29.20, 22.4, 22.1.

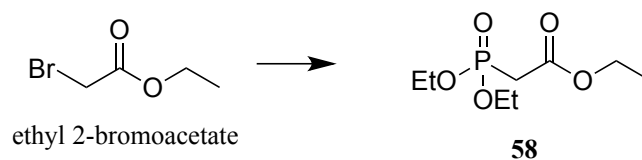


**[11,11,11- $^2\text{H}_3$ ]-Undecanal (56):** Clear oil (80%). IR (neat)  $\bar{\nu}$  2924, 2855, 2212, 1728, 1463, 1256;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (t,  $J = 1.8$  Hz, 1H), 2.41 (td,  $J = 1.8, 7.4$  Hz, 2H), 1.63 (quint,  $J = 7.3$  Hz, 2H), 1.25-1.30 (m, 14H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  203.0, 43.9, 31.8, 29.5, 29.4, 29.34, 29.30, 29.2, 22.4, 22.1.

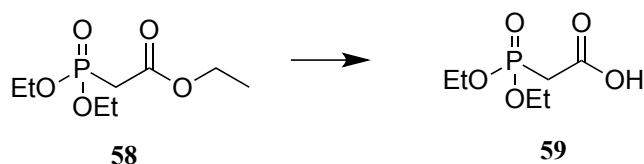


**[12,12,12- $^2\text{H}_3$ ]-Dodecanal (57):** Clear oil (96%). IR (neat)  $\bar{\nu}$  2925, 2854, 2713, 2213, 1727, 1463;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (s, 1H); 2.41 (t,  $J = 7.3$  Hz, 2H), 1.62 (quint,  $J = 7.0$  Hz, 2H), 1.25-1.30 (m, 16H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  203.0, 43.9, 31.8, 29.6, 29.5, 29.4, 29.36, 29.34, 29.2, 22.4, 22.1.

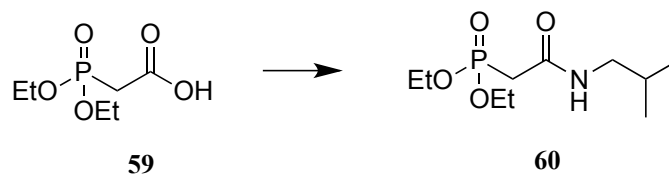


3.2.7 Synthesis of alkamides and [d<sub>3</sub>]-alkamides

**Ethyl 2-(diethoxyphosphoryl)acetate (58):** Ethyl bromoacetate (6.2094 g, 31.20 mmol) and triethyl phosphite (5.4481 g, 32.76 mmol, 1.05 equiv) were added to a round-bottom flask. The mixture was heated at 100 °C for 6 h. The reaction was cooled to room temperature to give clear liquid, ethyl 2-(diethoxyphosphoryl)acetate (6.9929 g, 99%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 4.14 (m, 6H), 2.93 (d, *J* = 21.5 Hz, 2H), 1.32 (t, *J* = 7 Hz, 6H), 1.26 (t, *J* = 7.3 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 167.8 (d, *J* = 6.3 Hz), 62.7 (d, *J* = 6.3 Hz), 61.5, 34.4 (d, *J* = 133.8 Hz), 16.3 (d, *J* = 6.3 Hz), 14.1. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [180].

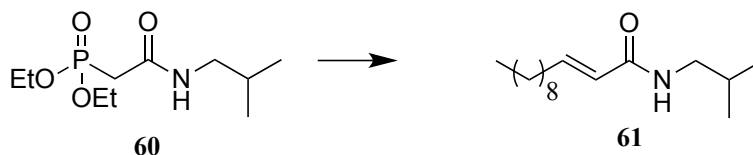


**2-(Diethoxyphosphoryl)acetic acid (59):** Triethyl phosphonoacetate **58** (5.0046 g, 22.32 mmol) was dissolved in H<sub>2</sub>O (3 mL) and cooled to 0 °C. NaOH (10 N, 2.7 mL, 1.2 equiv) was added dropwise to the stirred reaction mixture. When the base was added, the reaction was stirred for 1 h at 0 °C before being warmed to room temperature and stirred for an additional 2 h. The solution was cooled back down to 0 °C and 12 N HCl was added until the solution was at a pH of 1. The solution was extracted with DCM and washed with saturated NaCl. The organic layer was dried down to give a pale yellow liquid, 2-(diethoxyphosphoryl)acetic acid (4.0198 g, 92%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 4.20 (m, 4H), 3.00 (d, *J* = 22 Hz, 2H), 1.34 (t, *J* = 7.3 Hz, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 167.9 (d, *J* = 5.0 Hz), 63.5 (d, *J* = 6.3 Hz), 34.3 (d, *J* = 133.8 Hz), 16.4 (d, *J* = 6.3 Hz). <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [181].

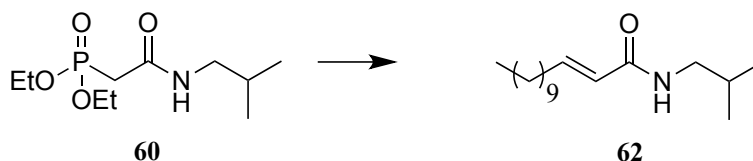


**Diethyl (2-isobutylamino-2-oxoethyl)phosphonate (60):** 2-(Diethoxyphosphoryl)acetic acid **59** (0.3943 g, 2.01 mmol) was dissolved in dry DCM (5 mL). CDI (0.3244 g, 2.01 mmol, 1 equiv) was dissolved in DCM (5 mL) and added to the stirred solution. After stirring at room temperature for 30 min, isobutylamine (0.146 g, 2.01 mmol, 1 equiv) was added and the reaction was stirred at room temperature overnight. The reaction was quenched with saturated NaHCO<sub>3</sub>, extracted with DCM, and dried with MgSO<sub>4</sub>. The solvent was removed to give a light yellow oil, diethyl (2-isobutylamino-2-oxoethyl)phosphonate (0.2601 g, 51 %). IR (neat)  $\bar{\nu}$  3295, 3084, 2961, 2931, 2871, 1659, 1556, 1443, 1319, 1292, 1246, 1098, 969, 731; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  4.09-4.15 (m, 4H), 3.08 (t, *J* = 6.6 Hz, 2H), 2.82 (d, *J* = 20.5, 2H), 1.77 (sept, *J* = 6.8 Hz, 1H), 1.32 (t, *J* = 7 Hz, 6H), 0.91 (d, *J* = 7 Hz, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  163.9 (d, *J* = 13.6 Hz), 62.7 (d, *J* = 25.7), 47.2, 35.5, 34.4, 28.4, 20.0, 16.3 (d, *J* = 24.3); HRMS (ESI) *m/z* calcd for [M + H]<sup>+</sup> C<sub>10</sub>H<sub>22</sub>NO<sub>4</sub>P 252.1359, found 252.1347.

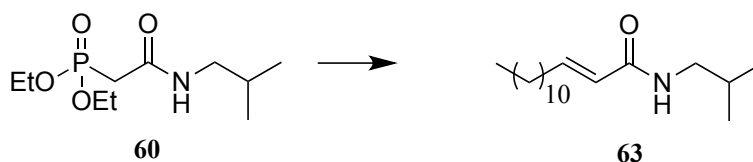
**General procedure for the synthesis of <sup>N</sup>iBu/12:1<sup>2E</sup>, <sup>N</sup>iBu/13:1<sup>2E</sup>, <sup>N</sup>iBu/14:1<sup>2E</sup>, d<sub>3</sub>-<sup>N</sup>iBu/12:1<sup>2E</sup>, d<sub>3</sub>-<sup>N</sup>iBu/13:1<sup>2E</sup>, d<sub>3</sub>-<sup>N</sup>iBu/14:1<sup>2E</sup>:** A procedure similar to the method of Matovic was used [182]. NaH (60%, 0.87 mmol, 1.1 equiv) is suspended in dry ether (4 mL) and cooled to 0 °C, then the phosphonate **60** (0.2000 g, 0.79 mmol) was added slowly and H<sub>2</sub> given off. After stirring for 15 min, the aldehyde (0.87 mmol, 1.1 equiv) in diethyl ether (0.3 mL) was added and reaction stirred for 10 min. The ice bath was removed and the reaction was quenched with H<sub>2</sub>O (10 mL). The reaction was extracted with hexanes (10 mL), washed with H<sub>2</sub>O (2 x 5 mL), brine (5 mL), and dried with MgSO<sub>4</sub>. The solvent was removed and the crude product was purified by column chromatography with hexane: EtOAc (3:1).



**(E)-N-Isobutyl-2-dodecenamide (61):** White crystals (51%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.82 (dt,  $J = 7, 15$  Hz, 1H), 5.75 (d,  $J = 15$  Hz, 1H), 3.15 (t,  $J = 6.8$  Hz, 2H), 5.45 (br, 1H), 2.16 (q,  $J = 7.3$  Hz, 2H), 1.79 (sept,  $J = 6.8$  Hz, 1H), 1.43 (m, 2H), 1.25-1.30 (m, 12H), 0.92 (d,  $J = 7$  Hz, 6H), 0.88 (t,  $J = 7$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.1, 144.8, 123.6, 46.8, 32.0, 31.9, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.7, 20.1, 14.1;  $^1\text{H}$  data was similar to that previously reported [183].

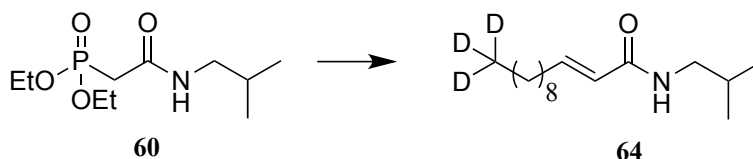


**(E)-N-Isobutyl-2-tridecenamide (62):** White crystals (49%). IR (KBr pellet)  $\bar{\nu}$  3304, 2951, 2924, 2854, 1668, 1631, 1550, 1466, 1371, 1265, 1160, 982, 721;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.82 (dt,  $J = 15, 7$  Hz, 1H), 5.75 (d,  $J = 15.5$  Hz, 1H), 5.49 (br, 1H), 3.14 (t,  $J = 6.5$  Hz, 2H), 2.16 (q,  $J = 7.3$  Hz, 2H), 1.79 (sept,  $J = 6.8$  Hz, 1H), 1.43 (m, 2H), 1.23-1.29 (m, 14H), 0.92 (d,  $J = 6.5$  Hz, 6H), 0.88 (t,  $J = 7$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.1, 144.8, 123.6, 46.8, 32.0, 31.9, 29.6, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.7, 20.1, 14.1; HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{17}\text{H}_{34}\text{NO}$  268.2635, found 268.2623.

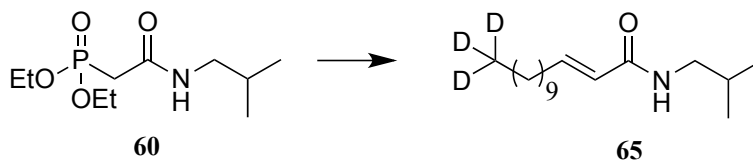


**(E)-N-Isobutyl-2-tetradecenamide (63):** White crystals (52%). IR (KBr pellet)  $\bar{\nu}$  3309, 2955, 2921, 2850, 1665, 1619, 1543, 1468, 1370, 986, 972;  $^1\text{H}$  NMR (500 MHz,

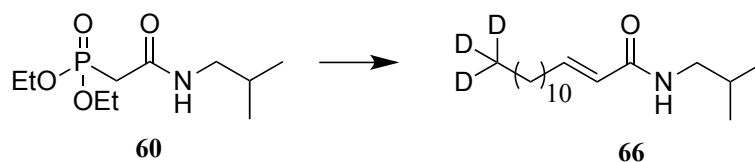
CDCl<sub>3</sub>) δ 6.82 (dt,  $J = 15, 7$  Hz, 1H), 5.75 (d,  $J = 15.5$  Hz, 1H), 5.42 (br, 1H), 3.15 (t,  $J = 6.5$  Hz, 2H), 2.16 (q,  $J = 7.3$  Hz, 2H), 1.80 (sept,  $J = 6.7$  Hz, 1H), 1.43 (m, 2H), 1.25-1.31 (m, 18H), 0.92 (d,  $J = 6.5$  Hz, 6H), 0.88 (t,  $J = 6.8$  Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 166.1, 144.8, 123.6, 46.8, 32.0, 31.9, 29.64, 29.62, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.7, 20.1, 14.1; HRMS (ESI)  $m/z$  calcd for  $[M + H]^+$  C<sub>18</sub>H<sub>36</sub>NO 282.2791, found 282.2799.



**[12,12,12-<sup>2</sup>H<sub>3</sub>]-(*E*)-*N*-Isobutyl-2-dodecenamide (64):** White crystals, (30%). IR (neat)  $\bar{\nu}$  3301, 3084, 2957, 2921, 2851, 2212, 1667, 1553, 1466, 1369, 1339, 1271, 1224, 1160, 979, 665 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 6.82 (dt,  $J = 15.2, 7.3$  Hz, 1H), 5.75 (d,  $J = 15.3$  Hz, 1H), 5.49 (br, 1H), 3.14 (t,  $J = 7.1$  Hz, 2H), 2.15 (q,  $J = 7.2$  Hz, 2H), 1.79 (nonet,  $J = 6.7$  Hz, 1H), 1.43 (quint,  $J = 7.1$  Hz, 2H), 1.25-1.33 (m, 12H), 0.92 (d,  $J = 6.7$  Hz, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 166.1, 144.7, 123.6, 46.8, 32.0, 31.8, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.4, 20.1.



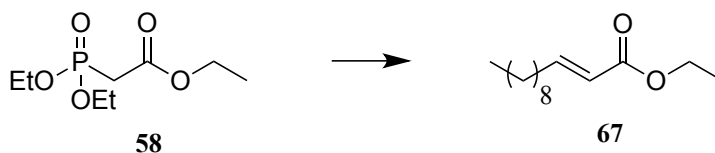
**[13,13,13-<sup>2</sup>H<sub>3</sub>]-(*E*)-*N*-Isobutyl-2-tridecenamide (65):** White crystals (24%). IR (neat)  $\bar{\nu}$  3304, 2921, 2211, 2851, 1666, 1625, 1552; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 6.82 (dt,  $J = 7.1, 15.3$  Hz, 1H), 5.75 (d,  $J = 15.3$  Hz, 1H), 5.51 (br, 1H), 3.14 (t,  $J = 6.5$  Hz, 2H), 2.15 (q,  $J = 7.2$  Hz, 2H), 1.79 (nonet,  $J = 6.7$  Hz, 1H), 1.43 (quint,  $J = 7.2$  Hz, 2H), 1.23-1.31 (m, 14H), 0.92 (d,  $J = 6.7$  Hz, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 166.1, 144.7, 123.6, 46.8, 32.0, 31.8, 29.6, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.4, 20.1; HRMS (ESI)  $m/z$  calcd for  $[M + H]^+$  C<sub>17</sub>H<sub>31</sub>D<sub>3</sub>NO 271.2823, found 271.2818.



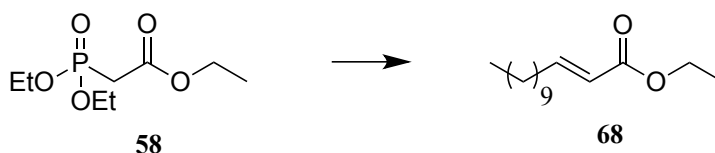
**[14,14,14-<sup>2</sup>H<sub>3</sub>]-(*E*)-*N*-Isobutyl-2-tetradecenamide (66):** White crystals (44%). IR (neat)  $\bar{\nu}$  3304, 3083, 2956, 2920, 2850, 2212, 1666, 1624, 1552, 1466; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  6.83 (dt,  $J = 7.3, 15.3$  Hz, 1H), 5.75 (d,  $J = 15.2$  Hz, 1H), 5.42 (br, 1H), 3.15 (t,  $J = 6.5$  Hz, 2H), 2.16 (q,  $J = 7.1$  Hz, 2H), 1.80 (nonet,  $J = 6.7$  Hz, 1H), 1.43 (quint,  $J = 7.0$  Hz, 2H), 1.25-1.37 (m, 16H), 0.92 (d,  $J = 6.7$  Hz, 6H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  166.1, 144.8, 123.6, 46.8, 32.0, 31.8, 29.6, 29.5, 29.4, 29.3, 29.2, 28.6, 28.3, 22.4, 20.1; HRMS (ESI)  $m/z$  calcd for  $[M + H]^+$  C<sub>18</sub>H<sub>33</sub>D<sub>3</sub>NO 285.2980, found 285.2999.

### 3.2.8 Synthesis of fatty acids and [d<sub>3</sub>]-fatty acids

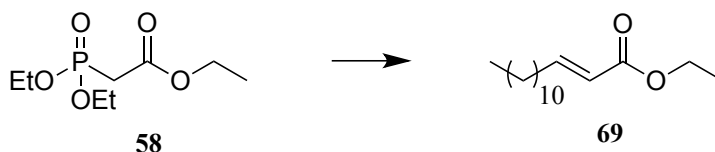
**General procedure for the synthesis of  $\alpha,\beta$ -unsaturated ethyl esters:** NaH (60%, 1.1 mmol, 1.1 equiv) was suspended in dry ether (4 mL) in a 15-mL round-bottom flask and cooled to 0 °C. The phosphonate **58** (0.22 g, 0.98 mmol) was added slowly; H<sub>2</sub> was given off. After stirring for 15 min, the aldehyde (1.08 mmol, 1.1 equiv) in diethyl ether (0.3 mL) was added dropwise and reaction was stirred for 10 min. The ice bath was removed and the reaction was quenched with H<sub>2</sub>O (10 mL). The mixture was extracted with hexane (10 mL). The organic layer washed with H<sub>2</sub>O (2 x 5 mL), saturated brine (5 mL), and dried with MgSO<sub>4</sub>. The solvent was removed and the crude product was purified by column chromatography with hexane: EtOAc (20:1).



**Ethyl (*E*)-2-dodecenoate (67):** Clear oil (51%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 16.0, 6.9$  Hz, 1H), 5.80 (d,  $J = 15.5$  Hz, 1H), 4.18 (q,  $J = 7.2$  Hz, 2H), 2.19 (q,  $J = 7.3$  Hz, 2H), 1.42-1.46 (m, 2H), 1.26-1.30 (m, 15H), 0.88 (t,  $J = 6.8$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 31.9, 29.5, 29.4, 29.3, 29.1, 28.0, 22.7, 14.2, 14.1.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to previously reported data [184]

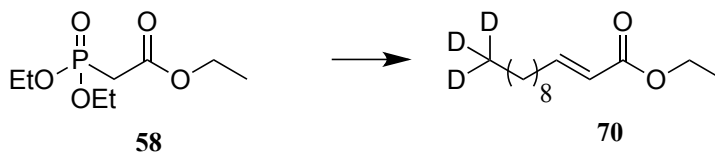


**Ethyl (*E*)-2-tridecenoate (68):** Clear oil (59%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.5, 7$  Hz, 1H), 5.81 (d,  $J = 15.5$  Hz, 1H), 4.18 (q,  $J = 7.2$  Hz, 2H), 2.19 (q,  $J = 7.2$  Hz, 2H), 1.42-1.46 (m, 2H), 1.26-1.30 (m, 17H), 0.88 (t,  $J = 7$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 31.9, 29.6, 29.5, 29.4, 29.3, 29.1, 28.0, 22.7, 14.3, 14.1.  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were similar to that previously reported [185].

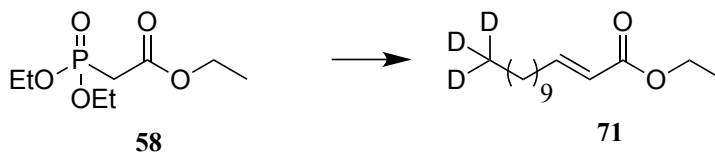


**Ethyl (*E*)-2-tetradecenoate (69):** Clear oil (48%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.5, 7$  Hz, 1H), 5.80 (d,  $J = 15.5$  Hz, 1H), 4.18 (q,  $J = 7$  Hz, 2H), 2.18 (q,  $J = 7.2$  Hz, 2H), 1.41-1.47 (m, 2H), 1.26-1.30 (m, 19H), 0.88 (t,  $J = 6.8$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 31.9, 29.62, 29.61, 29.5, 29.4, 29.3,

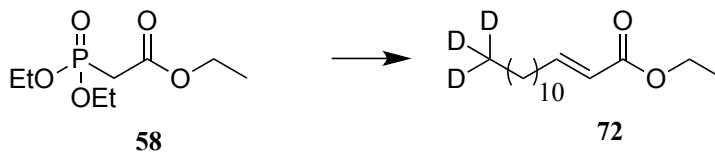
29.1, 28.0, 22.7, 14.3, 14.1.  $^1\text{H}$  spectral data was similar to that previously reported in reference [186].



**[12,12,12- $^2\text{H}_3$ ]-Ethyl (*E*)-2-dodecenoate (70):** Clear oil (39%). IR (neat)  $\bar{\nu}$  2925, 2855, 1724, 2212, 1723, 1654, 1462, 1264, 1180, 1044;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.7, 7.0$  Hz, 1H), 5.81 (dt,  $J = 15.7, 1.5$  Hz, 1H), 4.18 (q,  $J = 7.1$  Hz, 2H), 2.19 (q,  $J = 7.0$  Hz, 2H), 1.45 (quint,  $J = 7.2$  Hz, 2H), 1.25-1.36 (m, 15H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 29.7, 29.5, 29.4, 29.3, 29.1, 28.0, 22.4, 14.3.



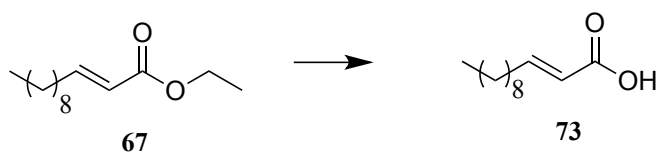
**[13,13,13- $^2\text{H}_3$ ]-Ethyl (*E*)-2-tridecenoate (71):** Clear oil (96%). IR (neat): 2925, 2855, 2212, 1724, 1264, 1180  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.7, 7.0$  Hz, 1H), 5.81 (d,  $J = 15.7$  Hz, 1H), 4.18 (q,  $J = 7.1$  Hz, 2H), 2.18 (q,  $J = 7.1$  Hz, 2H), 1.45 (quint,  $J = 7.2$  Hz, 2H), 1.25-1.36 (m, 17H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 29.7, 29.6, 29.5, 29.4, 29.3, 29.1, 28.0, 22.4, 14.3. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{15}\text{H}_{26}\text{D}_3\text{O}_2$  244.2350, found 244.2357.



**[14,14,14- $^2\text{H}_3$ ]-Ethyl (*E*)-2-tetradecenoate (72):** Clear oil (79%). IR (neat)

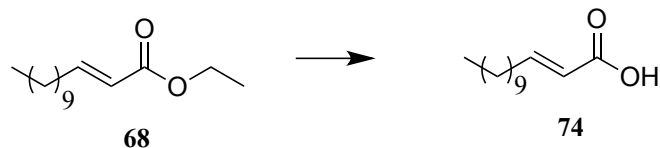
$\bar{\nu}$  2924, 2854, 2213 1723, 1264, 1179;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.96 (dt,  $J = 15.7$ , 7.4 Hz, 1H), 5.81 (d,  $J = 15.6$  Hz, 1H), 4.18 (q,  $J = 7.1$  Hz, 2H), 2.19 (q,  $J = 7.1$  Hz, 2H), 1.44 (m, 2H), 1.23-1.30 (m, 19H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8, 149.5, 121.2, 60.1, 32.2, 29.63, 29.61, 29.5, 29.4, 29.3, 29.15, 29.13, 28.0, 22.4, 14.3. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{16}\text{H}_{28}\text{D}_3\text{O}_2$  258.2507, found 258.2521.

**General procedure for the synthesis of  $\alpha,\beta$ -unsaturated fatty acids:** Ester **67** (10.0 mmol) was dissolved in ethanol (15 mL) in a 100-mL flask and then 1 M KOH (20 mL, 20 equiv) was added. The reaction was heated at reflux for 18 h. It was cooled to room temperature and the ethanol was removed by rotary evaporation. Hexanes were added and 1 M HCl was added by pipet until the aqueous layer was acidic as determined by pH paper. The hexane layer was collected, dried with  $\text{NaSO}_4$ , and the solvent removed under reduced pressure. The crude product was purified by column chromatography with hexane: EtOAc (10:1) containing 1% acetic acid.

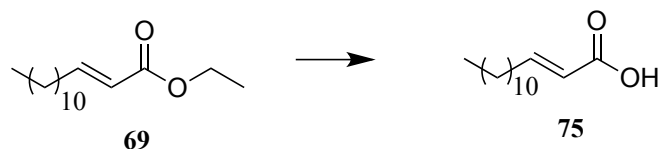


**(E)-2-Dodecenoic acid (73):** Colorless oil (76%). IR (neat)  $\bar{\nu}$  3120 (br), 2955, 2926, 2855, 1697, 1651, 1464, 1420, 1307, 1464, 1420, 1307, 1286, 982;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.08 (dt,  $J = 15.5$ , 7 Hz, 1H), 5.82 (dt,  $J = 15.5$ , 1.5 Hz, 1H), 2.23 (q,  $J = 7.3$  Hz, 2H), 1.47 (quint,  $J = 7.13$  Hz, 2H), 1.29-1.27 (m, 12H), 0.88 (t,  $J = 7$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  171.7, 152.5, 120.5, 32.3, 31.9, 29.5, 29.4, 29.3, 29.1, 27.9, 22.7, 14.1. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{12}\text{H}_{23}\text{O}_2$  199.1693, found 199.1690.



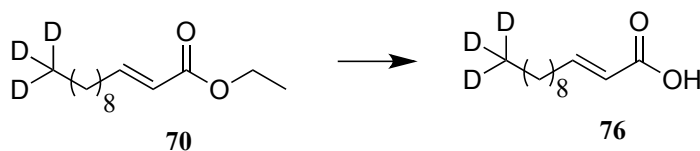


**(E)-2-Tridecenoic acid (74):** Colorless solid (83%). IR (KBr pellet)  $\bar{\nu}$  3119 (br), 2955, 2925, 2854, 1699, 1652, 1466, 1420, 1284, 1225, 983, 931, 862, 721, 543;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.09 (dt,  $J = 15.5, 7.0$  Hz, 1H), 5.82 (dt,  $J = 15.5, 1.5$  Hz, 1H), 2.23 (q,  $J = 7.3$  Hz, 2H), 1.46 (quint,  $J = 7.3$  Hz, 2H), 1.29-1.26 (m, 14H), 0.88 (t,  $J = 7.3$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  171.7, 152.5, 120.5, 32.3, 31.9, 29.6, 29.5, 29.4, 29.3, 29.1, 27.9, 22.7, 14.1. HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{13}\text{H}_{25}\text{O}_2$  213.1849, found 213.1848.

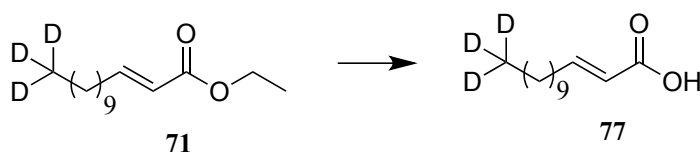


**(E)-2-Tetradecenoic acid (75):** Colorless solid (56%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.08 (dt,  $J = 15.5, 7.0$  Hz, 1H), 5.82 (d,  $J = 15.5$  Hz, 1H), 2.23 (q,  $J = 7$  Hz, 2H), 1.46 (quint,  $J = 7.8$  Hz, 2H), 1.26-1.32 (m, 18H), 0.88 (t,  $J = 6.8$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  171.8, 152.5, 120.5, 32.3, 31.9, 29.62, 29.61, 29.5, 29.4, 29.3, 29.2, 27.9, 22.7, 14.1.  $^1\text{H}$  spectral data was similar to that previously reported by [187].

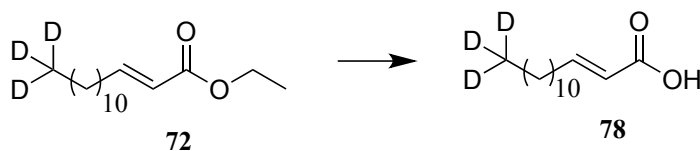
**General procedure for the synthesis of the deuterated  $\alpha,\beta$ -unsaturated fatty acids:** The ester (0.30 mmol) was dissolved in ethanol (2 mL) and 1 M KOH (1.5 mL, 5 equiv) was added. The reaction mixture was heated at reflux for 18 h. It was cooled to room temperature and the ethanol was removed by rotary evaporation. Hexanes were added and 1 M HCl was added by pipet until the aqueous layer was acidic as determined by pH paper. The hexane layer was collected, dried with  $\text{NaSO}_4$ , and the solvent removed under reduced pressure. The crude product was purified by column chromatography with hexane: EtOAc (10:1) containing 1% acetic acid.



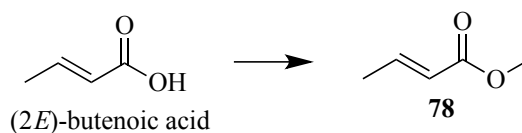
**[12,12,12-<sup>2</sup>H<sub>3</sub>]-(*E*)-2-Dodecenoic acid (76):** Colorless oil (46%). IR (neat)  $\bar{\nu}$  3059 (br), 2925, 2855, 2212, 1689, 1462, 1420, 1285  $\text{cm}^{-1}$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.08 (dt,  $J = 15.4, 6.8$  Hz), 5.82 (d,  $J = 15.5$  Hz, 1H), 2.23 (q,  $J = 7.0$  Hz, 2H), 1.47 (quint,  $J = 7.1$  Hz, 2H), 1.20-1.32 (m, 12H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  172.0, 152.5, 120.5, 32.3, 31.8, 29.7, 29.5, 29.4, 29.3, 29.1, 27.9, 22.4.



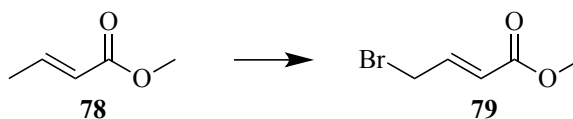
**[13,13,13-<sup>2</sup>H<sub>3</sub>]-(*E*)-2-Tridecenoic acid (77):** Colorless solid (33%). IR (neat)  $\bar{\nu}$  3018 (br), 2921, 2211, 2849, 1692, 1654, 1466, 1419, 1286; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.08 (dt,  $J = 15.6, 7.4$  Hz, 1H), 5.82 (d,  $J = 15.6$  Hz, 1H), 2.23 (q,  $J = 7.2$  Hz, 2H), 1.46 (quint,  $J = 7.1$  Hz, 2H), 1.22-1.30 (m, 14H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  172.0, 152.5, 120.6, 32.3, 31.8, 29.6, 29.5, 29.4, 29.3, 29.1, 27.9, 22.4.



**[14,14,14-<sup>2</sup>H<sub>3</sub>]-(*E*)-2-Tetradecenoic acid (78):** Colorless solid (56%). IR (neat)  $\bar{\nu}$  3031 (br), 2920, 2849, 2215, 1692, 1654, 1466, 1286; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.08 (dt,  $J = 15.6, 7.4$  Hz, 1H), 5.82 (d,  $J = 15.7$  Hz, 1H), 2.23 (q,  $J = 7.2$  Hz, 2H), 1.46 (quint,  $J = 7.2$  Hz, 2H), 1.25-1.30 (m, 16H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  171.9, 152.5, 120.5, 32.3, 31.8, 29.6, 29.5, 29.4, 29.37, 29.35, 29.14, 29.13, 27.9, 22.4; HRMS (ESI)  $m/z$  calcd for  $[\text{M} + \text{H}]^+ \text{C}_{14}\text{H}_{24}\text{D}_3\text{O}_2$  230.2194, found 230.2199.

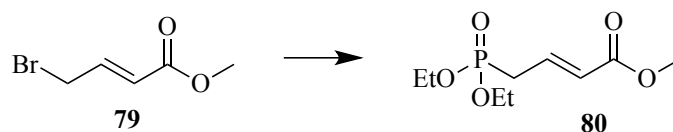
3.2.9 Synthesis of (*E,E*)-2,4-dodecadienoic acid

**Methyl (*E*)-2-butenoate (78):** Crotonic acid (12.2100 g, 0.1419 mol) was dissolved in MeOH (35 mL, 0.86 mol, 6 equiv) in a 100-mL flask. Sulfuric acid (1.362 mL, 0.0256 mol, 0.18 equiv) was added and reaction was heated to reflux for 20 h. The solution was cooled to room temperature and H<sub>2</sub>O (35 mL) was added. The mixture was extracted with diethyl ether (50 mL). The organic layer was washed with saturated NaHCO<sub>3</sub>, and dried with MgSO<sub>4</sub>. The solvent was removed to give the pale yellow oil methyl (*E*)-2-butenoate (7.3729 g, 54%). IR (neat)  $\bar{\nu}$  3050, 2976, 1701, 1656, 1444, 1316, 1224, 1103, 967, 908; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  6.98 (dq,  $J = 15.5, 7$  Hz, 1H), 5.85 (dq,  $J = 15, 2$  Hz, 1H), 3.72 (s, 3H), 1.88 (dd,  $J = 7, 2$  Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  167.2, 144.9, 122.6, 51.5, 18.1. HRMS (ESI)  $m/z$  calcd for [M + H]<sup>+</sup> C<sub>5</sub>H<sub>9</sub>O<sub>2</sub> 101.0597, found 101.0576. <sup>1</sup>H data were similar to that previously reported [188]. <sup>13</sup>C data were similar to that previously reported [189].

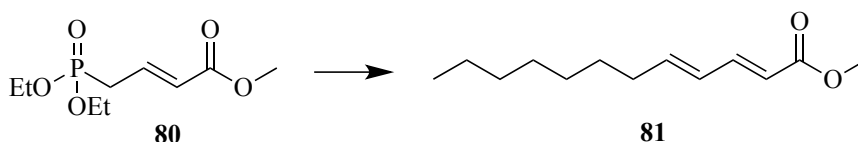


**Methyl (*E*)-4-bromobut-2-enoate (79):** Methyl (*E*)-2-butenoate (1.9438 g, 19.37 mmol) was dissolved in benzene (25 mL) in a 50-mL flask and *N*-bromosuccinimide (3.7923 g, 21.31 mmol, 1.1 equiv) was added. The solution was heated to reflux and benzoyl peroxide (0.0567 g, 0.2341 mmol, 0.012 equiv) was added through the top of the condenser. The reaction was heated at reflux for 3 h. Once the solution cooled the succinimide crystals were filtered. The filtrate was concentrated to yield methyl *E*-4-bromobut-2-enoate (3.4322 g, 99%). IR (neat)  $\bar{\nu}$  3062, 2999, 2952, 1724, 1657, 1436, 1284, 1136, 924, 837, 730; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.01 (dt,  $J = 15, 7$  Hz, 1H),

6.03 (dt,  $J = 15.5, 1.3$  Hz, 1H), 4.01 (dd,  $J = 7.3, 1.3$  Hz, 2H), 3.76 (s, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.1, 142.1, 124.4, 52.0, 29.2.  $^1\text{H}$  data were similar to that previously reported [190].

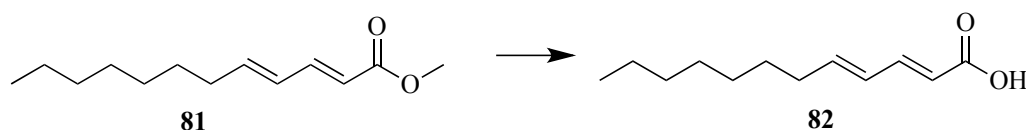


**Methyl (*E*)-4-(diethoxyphosphoryl)-2-butenoate (80):** Methyl *E*-4-bromobut-2-enoate (3.4322 g, 19.2 mmol) and triethyl phosphite (3.9054 g, 23.4 mmol, 1.2 equiv) were mixed together and heated at 100 °C for 18 h. The reaction was cooled to room temperature and dissolved in hexane:EtOAc (3:1). The solution was loaded onto a silica gel column to removed excess triethyl phosphite. Solvent was removed to yield methyl (*E*)-4-(diethoxyphosphoryl)-2-butenoate (3.2762 g, 13.8 mmol, 72%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.88 (sextet,  $J = 7.7$  Hz, 1H), 5.97 (ddt,  $J = 15.6, 5, 1.4$  Hz, 1H), 4.10 (m, 4H), 3.74 (s, 3H), 2.75 (ddd,  $J = 9.5, 8, 1.5$  Hz, 2H), 1.32 (t,  $J = 7$  Hz, 6H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.2, 137.8, 125.6, 62.5, 51.8, 31.3, 30.2, 16.5.  $^1\text{H}$  spectral data was similar to that previously reported [191].

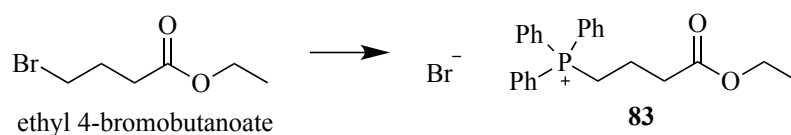


**Methyl (*E,E*)-dodeca-2,4-dienoate (81):** To a mixture of NaH (60%, 0.2072 g, 5.18 mmol, 1.2 equiv) suspended in diethyl ether (16 mL), methyl (*E*)-4-(diethoxyphosphoryl)-2-butenoate (1.0051 g, 4.25 mmol) was added dropwise at 0 °C;  $\text{H}_2$  gas given off during the addition. After the reaction was stirred for 30 minutes, octanal (1.40 mL, 8.95 mmol, 2.1 equiv) was added. The solution was heated at reflux for 2 h before cooling to room temperature and quenched with  $\text{H}_2\text{O}$  (15 mL). The mixture was extracted with hexanes (2 x 10 mL). The organic layer was washed with  $\text{H}_2\text{O}$  (10 mL), washed with brine (10 mL), dried with  $\text{MgSO}_4$ , and the solvent was removed by reduced

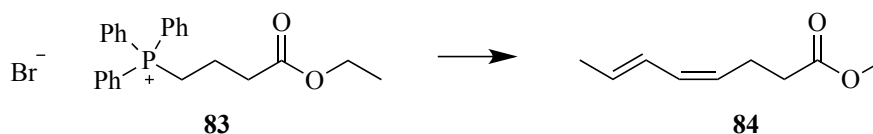
pressure. Crude product was purified with silica gel column hexane:ethyl acetate (20:1) to yield methyl (*E,E*)-dodeca-2,4-dienoate (0.1442 g, 16%).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.26 (dd,  $J = 9.9, 15.3$  Hz, 1H), 6.14 (m, 2H), 5.78 (d,  $J = 15.4$  Hz, 1H), 3.73 (s, 3H), 2.16 (q,  $J = 6.9$  Hz, 2H), 1.42 (quint,  $J = 7.0$  Hz, 2H), 1.26-1.28 (m, 8H), 0.88 (t,  $J = 6.9$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  167.7, 145.4, 144.9, 128.3, 118.6, 51.4, 32.9, 31.7, 29.1, 29.0, 28.7, 22.6, 14.1.  $^1\text{H}$  and  $^{13}\text{C}$  data were similar to that previously reported [192].



**(2*E*,4*E*)-Dodecadienoic acid (82):** Methyl (2*E*,4*E*)-dodecadienoate (0.1442 g, 0.6856 mmol) was dissolved in ethanol (1.688 mL) and 1M KOH (5.070 mL, 5.070 mmol, 7.4 equiv) was added. The reaction was heated at reflux overnight. After cooling to room temperature, hexanes were added followed by 1M HCl until the aqueous layer was acidic. The organic layer was extracted and solvent removed to yield (2*E*, 4*E*)-dodecadienoic acid (0.1023 g, 76%). IR (neat)  $\bar{\nu}$  3119 (br), 2955, 2925, 2854, 1699, 1652, 1466, 1420, 1284, 1225, 983, 931, 862, 721, 543;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.34 (m, 1H), 6.19 (m, 2H), 5.78 (d,  $J = 15.3$ , 1H), 2.18 (q,  $J = 7.1$  Hz, 2H), 1.65 (quint,  $J = 7.5$  Hz, 2H), 1.43 (m, 2H), 1.26-1.69 (m, 6H), 0.88 (t,  $J = 7.1$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  172.3, 147.5, 146.3, 128.2, 118.1, 33.1, 31.8, 29.1, 29.0, 28.6, 22.6, 14.1.  $^1\text{H}$  spectral data was similar to that previously reported by [193].

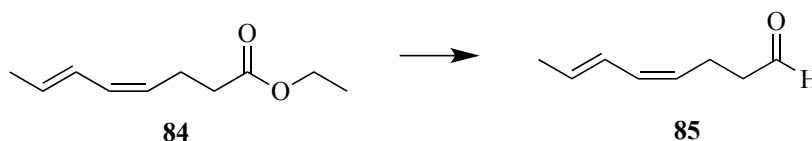
3.2.10 Synthesis of (2*E*,4*E*,8*Z*,10*E*)-dodecadienoic acid

**(4-Ethoxy-4-oxobutyl)triphenylphosphonium bromide (83):** Ethyl 4-bromobutanoate (10.0128 g, 51.33 mmol) was dissolved in acetonitrile (12 mL). Triphenylphosphine (13.7129 g, 51.33 mmol, 1 equiv) was added and the reaction was heated at 90 °C for 3 days. To the solution, toluene (50 mL) was added dropwise; crystals formed during the addition. The mixture was cooled to room temperature and toluene was decanted. The wet crystals were dried under high vacuum to yield (4-ethoxy-4-oxobutyl)triphenylphosphonium bromide (20.3099 g, 87%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.87 (m, 6H), 7.79 (m, 3H), 7.70 (m, 6H), 4.09 (q, *J* = 7.2 Hz, 2H), 4.00 (m, 2H), 2.87 (t, *J* = 5.5 Hz, 2H), 1.92 (m, 2H), 1.23 (t, *J* = 7.0 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 173.4, 135.1 (d, *J* = 2.85 Hz), 133.9 (d, *J* = 10.0 Hz), 130.6 (d, *J* = 12.6 Hz), 60.8, 33.4 (d, *J* = 18.75 Hz), 22.1, 21.7, 18.2, 14.3. <sup>1</sup>H and <sup>13</sup>C spectral data were similar to that previously reported [194].

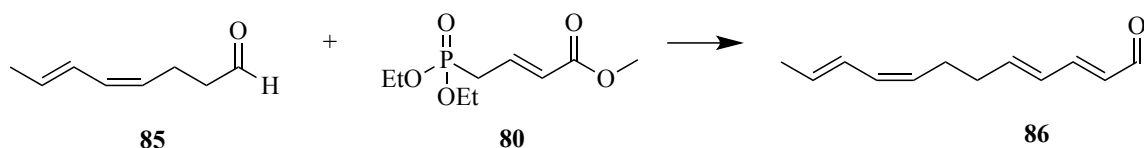


**Ethyl (Z,E)-octa-4,6-dienoate (84):** In THF (30 mL), (4-ethoxy-4-oxobutyl)triphenylphosphonium bromide (10.1752 g, 22.3 mmol) was suspended and potassium *tert*-butoxide (2.5750 g, 22.9 mmol, 1 equiv) was added; the solution turned bright orange. The reaction was stirred for 30 min at room temperature. Crotonaldehyde (1.85 mL, 22.3 mmol, 1 equiv) was added and reaction was stirred at room temperature for 18 h giving a pale orange solution. The reaction was quenched with H<sub>2</sub>O (100 mL). The solution was extracted with ethyl acetate. The ethyl acetate layer was dried with MgSO<sub>4</sub>, and the solvent was removed by rotary evaporation. The crude product was

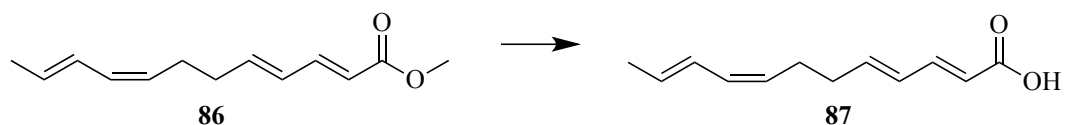
purified by silica gel column hexane:ethyl acetate (20:1) to yield ethyl (4*Z*,6*E*)-octadienoate (1.5789 g, 42%). IR (neat)  $\bar{\nu}$  3020, 2981, 2915, 1737, 1448, 1372, 1349, 1254, 1172, 1099, 1041, 985, 947, 926, 821  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  6.33 (dd,  $J = 12, 13$  Hz, 1H), 5.97 (t,  $J = 10.7$  Hz, 1H), 5.70 (dq,  $J = 14.5, 7.1$  Hz, 1H), 5.24 (dd,  $J = 7.5, 18$  Hz, 1H), 4.13 (t,  $J = 7$  Hz, 2H), 2.48 (qt,  $J = 7.3$  Hz, 2H), 2.37 (t,  $J = 7.5$  Hz, 2H), 1.77 (d,  $J = 6.5$  Hz, 3H), 1.25 (t,  $J = 7.3$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  173.2, 130.2, 130.0, 127.0, 126.8, 60.5, 34.5, 23.4, 18.4, 14.4.



**(*Z,E*)-Octa-4,6-dienal (85):** Ethyl (*Z,E*)-octa-4,6-dienoate (0.7715 g, 4.59 mmol) was dissolved in diethyl ether (5.4 mL) and cooled to  $-78$  °C. with an acetone/dry ice bath. DIBAL (5.3 mL, 5.0 mmol, 1.1 equiv) was added dropwise. After the addition was complete, the reaction was stirred for 30 min at  $-78$  °C. The solution was warmed to  $0$  °C with an ice-water bath and 10 g of 4 N HCl was cooled and added dropwise. The mixture was then warmed to room temperature. The mixture was extracted with diethyl ether. The organic layer was washed with brine, dried with  $\text{MgSO}_4$ , and the solvent was removed under reduced pressure. The crude product was purified by silica gel column hexane:ethyl acetate (5:1) to yield (*Z,E*)-octa-4,6-dienal (0.2473 g, 1.99 mmol, 43%). IR (neat)  $\bar{\nu}$  3020, 2959, 2914, 2852, 2724, 1726, 1653, 1449, 1409, 1388, 1376, 1055, 985, 948, 820;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.80 (t,  $J = 1.5$  Hz, 1H), 6.34 (t,  $J = 13$  Hz, 1H), 6.01 (t,  $J = 11$  Hz, 1H), 5.74 (dq,  $J = 14.5, 6.5$  Hz, 1H), 5.26 (dd,  $J = 7.5, 17.5$  Hz, 1H), 2.53 (m, 4H), 1.80 (d,  $J = 6.5$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  202.1, 130.6, 130.1, 126.6, 124.1, 43.9, 20.6, 18.4.



**Methyl (*E,E,Z,E*)-dodeca-2, 4, 8, 10-tetraenoate (86):** NaH (0.09382 g, 2.33 mmol, 1.1 equiv, 60%) was suspended in diethyl ether (6 mL) in a 25-mL round-bottom flask. Phosphonate ester (0.4553 g, 1.93 mmol) was added dropwise; H<sub>2</sub> given off and the solution turned yellow. After the reaction was stirred for 30 min at room temperature, *Z,E*-octa-4,6-dienal (0.2632 g, 2.12 mmol, 1.1 equiv) in diethyl ether (0.5 mL) was added. The solution was stirred for a further 22 h at room temperature. The reaction was quenched with H<sub>2</sub>O (10 mL) and extracted with hexanes (10 mL). The organic layer was washed with H<sub>2</sub>O (5 mL), washed with brine (5 mL), dried with MgSO<sub>4</sub>, and the solvent was removed by rotary evaporation. The crude product was purified by methyl (*E,E,Z,E*)-dodeca-2,4,8,10-tetraenoate (0.0225 g, 6 %). IR (neat)  $\bar{\nu}$  2925, 2855, 1724, 2212, 1723, 1654, 1462, 1264, 1180, 1044 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.50 (m, 1H), 6.27 (m, 1H), 6.03 (m, 2H), 5.72 (m, 3H), 5.61 (m, 1H), 3.80 (s, 3H), 2.01 (m, 4H), 1.58 (d,  $J$  = 6.4 Hz, 3H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  167.6, 142.9, 141.7, 130.3, 131.3, 129.52, 129.50, 126.9, 119.9, 52.0, 33.01, 33.03, 19.3.



**(*E,E,Z,E*)-Dodeca-2, 4, 8, 10-tetraenoic acid (87):** Methyl (*E,E,Z,E*)-dodeca-2,4,8,10-tetraenoate (0.0318 g, 0.162 mmol) was dissolved in ethanol (0.5 mL) and 1 M KOH (0.551 mL, 0.551 mmol, 3.4 equiv) was added. The reaction was heated at reflux for 2 h. After cooling to room temperature, hexanes were added. 1 M HCl was added until the aqueous layer was acidic. The organic layer was extracted and solvent removed to yield (*2E,4E,8Z,10E*)-dodecatetraenoic acid (0.0249 g, 0.1295 mmol, 80%). IR (neat)  $\bar{\nu}$  3119 (br), 2955, 2925, 2854, 1699, 1652, 1466, 1420, 1284, 1225, 983, 931, 862, 721, 543 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.50 (m, 1H), 6.27 (m, 1H), 6.03 (m, 2H), 5.72



(m, 3H), 5.61 (m, 1H), 2.01 (m, 4H), 1.58 (d,  $J = 6.4$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  167.6, 142.9, 141.7, 130.3, 131.3, 129.52, 129.50, 126.9, 119.9, 33.01, 33.03, 19.3.

### 3.2.11 Synthesizing acyl-CoAs

Fatty acid (~0.25 mmol) was dissolved in 2 mL of THF.  $\text{Et}_3\text{N}$  (60  $\mu\text{L}$ ) and ethyl chloroformate (40  $\mu\text{L}$ ) were added resulting in the formation of a white precipitate. The mixture was stirred for 1 h. The mixture was filtered through a pipet with a glass wool plug. The solvent was removed by rotary evaporation and the crude mixed anhydride was checked by  $^1\text{H}$  NMR.

The trilithium salt of CoA was dissolved in a solution of  $\text{NaHCO}_3$  (50 mg) in  $\text{H}_2\text{O}:\text{THF}$  (3:2) and the reaction flask was flushed with  $\text{N}_2$ . The mixed anhydride was dissolved in THF (2 mL) in a separate flask and was then added to the CoA solution. Water was added until the organic and aqueous phases were miscible and the reaction mixture was stirred at room temperature for 2 h.

The reaction was concentrated by rotary evaporation to remove most of the THF prior to transferring the solution to a conical tube by a Pasteur pipet. The aqueous layer was acidified with 5%  $\text{HClO}_4$  until a precipitate stopped forming. The mixture was centrifuged and the supernatant was removed with a pipet. The solid was purified by washing with 0.8%  $\text{HClO}_4$  (1 mL). During this step, the mixture was vortexed to ensure mixing, centrifuged, and decanted. Washing steps using the same operations were carried out with dry acetone (2 x 1 mL) and diethyl ether (1 mL). The solid was dried in 60  $^\circ\text{C}$  heat block under  $\text{N}_2$  to remove excess diethyl ether. The acyl-CoA sample was dissolved in 80 mM  $\text{NaHCO}_3$  (250  $\mu\text{L}$ ) and stored at -80  $^\circ\text{C}$ .

### 3.2.12 Calculating concentrations of acyl-CoAs

To calculate the concentration calculated by UV-vis based on the amount of CoA present, following a similar procedure completed by Merrill [195]. The CoA sample (5  $\mu\text{L}$ ) was added to a separate vial and dissolved in 0.1 M potassium phosphate (pH 7.0).

The sample was placed in a cuvette and UV-VIS readings were taken at 260 nm [195]. The absorbance at 260 nm was measured and the concentration of the diluted sample was calculated using equation 2, where  $A$  is the absorbance, the path length is 1, and the extinction coefficient was  $15,400 \text{ M}^{-1}\text{cm}^{-1}$  [195].

$$C_1 = \frac{A}{1 \cdot 15,400} \quad (3)$$

To determine the original concentration of the CoA sample,  $C_1V_1=C_2V_2$  was used to solve for  $C_2$ .  $C_1$  was the concentration calculated from equation 1.  $V_1$  was 2 mL and  $V_2$  was 0.005 mL.

The concentration was independently calculated by GC-MS based on the fatty acid content. Concentration curves were made from analysis of methyl esters of 12:1, 12:2, 12:4, and 15:0. The slopes of each methyl ester were calculated by graphing the corrected area of each peak of the total ion chromatogram versus the concentrations of the FAME solutions (0.4, 0.2, 0.08, 0.04, 0.026, 0.008 mM). Sensitivity factors for the methyl esters of 12:1, 12:2, and 12:4 were calculated from the ratio of the slope for the standard curve of each analyte to that of the methyl ester of 15:0.

The acyl-CoA concentrations were determined by the conversion of the CoA products to FAMES prior to GC/MS analysis. A 40 mM solution of 15:0 in hexanes was prepared as an internal standard. FAMES analysis was prepared by pipetting the CoA sample (10  $\mu\text{L}$ ) and 15:0 solution (37  $\mu\text{L}$ ) into a conical tube. Sulfuric acid (2%) in methanol (2 mL) was added and the solution was heated for 2 h at 80  $^\circ\text{C}$ . The solution was cooled and  $\text{H}_2\text{O}$  (1 mL) and extracted with hexanes (2 x 1 mL). The hexanes were dried down and the sample was re-dissolved in hexanes (200  $\mu\text{L}$ ) and run on GC-MS.

The following calculations were used to calculate the concentration of each CoA ( $X = 1, 2, \text{ and } 4$ ) and the percent yield from the acyl CoA synthesis.

$$\text{Corrected area} = (\text{experimental area of } 12:X) \cdot (\text{sensitivity factor}) \quad (4)$$

$$\mu\text{mol of } 12:X - \text{CoA} = \frac{\text{Corrected Area of } 12:X}{\text{Explicit Area of } 15:0} \cdot \mu\text{mol of } 15:0 \quad (5)$$

$$[12:X - \text{CoA}] = \frac{\mu\text{mol of } 12:X - \text{CoA}}{10 \mu\text{L} \cdot 10^3} \quad (6)$$

### 3.3 Results and Discussion

Previously, to study biosynthetic pathways radiolabeled [ $^{14}\text{C}$ ] compounds were used due to their enhanced sensitivity. However, stable isotopic precursors like deuterium can be utilized to determine isotopomer patterns in GC/MS and LC/MS, potentially providing additional data. The most prevalent alkamide present in *E. purpurea* is BA8 ( $^{\text{N}}\text{iBu}/12:4\Delta^{2\text{E},4\text{E},8\text{Z},10\text{E}}$ ) followed by BA9 ( $^{\text{N}}\text{iBu}/12:4\Delta^{2\text{E},4\text{E},8\text{Z},10\text{Z}}$ ) [1]. BA9 accumulates at 1/5<sup>th</sup> the level of BA8 [1]. Both BA8 and BA9 were used for correlating incorporation of labeled compounds because they are the most abundant within *E. purpurea* seedlings. They both have similar retention times in GC/MS and LC/MS, and they also have a distinctive GC/MS fragmentation pattern. The dominant fragment produced has an  $m/z$  of 167 that is thought to arise from a pericyclic fragmentation eliminating hexatriene (Figure 39) [132]. This fragmentation contains the six carbonyl-proximate carbons within the acyl chain including the *N*-isobutylamide [132].

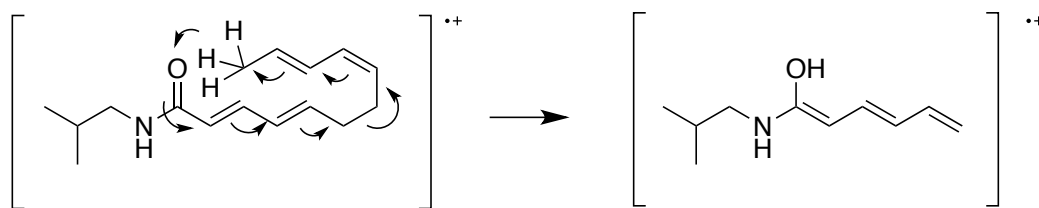


Figure 39: The dominant fragment  $m/z$  167 comes from the pericyclic fragmentation and elimination of hexatriene [132].

*E. purpurea*, grown as a field plant, collects alkamides in the roots, flowers, and seed coats [1, 132]. It had been determined from an earlier study that alkamides were actively produced late during seedling development [132]. The seedlings were characterized in five phenotypically distinct stages [132]. Stage 1 and stage 2 included the emergence from the seed coat and the appearance of the shoot apical meristem. In stage 3, the cotyledons were formed and emerged from the seed coat. Stage 4 corresponded to the first true leaf being formed. Then, in stage 5, the first true leaf fully formed and begin to unfurl (Figure 40). Alkamide production increases in germination stages 3-5 [1, 132]. The use of stage 5 seedlings provides a faster way for alkamide analysis with less variables such as soil, humidity, and microorganisms [1, 132]. Production of alkamides could be amplified by the application of 10  $\mu$ M of methyl jasmonate (MeJA) in the media. Over a 72-h period, *E. purpurea* treated with MeJA showed a 2-fold increase in accumulation of alkamides [1]. All seedlings in the current studies were stage 5 and were supplemented with MeJA via the liquid media and grown for three consecutive days for all feeding experiments.



Figure 40: Example of a stage 5 of *E. purpurea* seedling used in feeding experiments

### 3.3.1 Isotope labeling experiments of amino acids, Val and Ile

Our group has hypothesized that amino acids and fatty acids are associated with the biosynthesis of alkamides due to the two moieties that are present in alkamides, i.e. the amine moiety that seems to be acylated by a fatty-acid derivative. The Nikolau group has explored the accumulation pattern of fatty acids and amino acids in 13 plant tissues and organs [1]. After GC/MS and LC/MS analysis, 371 analytes were identified, Pearson

correlation coefficients were calculated, and the coefficients were clustered together if they had a correlation analysis value of greater than 0.7. This led the 371 compounds to be segregated into 21 separate clades labeled A-U [1].

All eight alkamides measured in this study previously shown clustered in clade U. The branched-chain amino acids, valine and isoleucine, were clustered in clad N; this clade is statistically distant from clade U [1]. The fatty acids identified with 14-carbons atoms or longer were allocated among eight different clades (C, F, H, I, J, N, O, and P). There were no fatty acids with a 12-carbon acyl chain found in the alkamides-enriched clade U. Thus, metabolite-profiling studies did not support the starting hypothesis that there is a tissue-dependent metabolic correlation between the accumulated levels of the potential precursors, amino acids and fatty acids, and the alkamides produced [1].

An experiment was designed to use isotopic tracers to test whether Val and Ile are precursors of the amine moiety and to identify the probable chemical transformations leading to that amine segment [1]. Initial studies were completed to identify the labeling patterns of Val and the alkamides. In order to complete this, seedlings were cultured in media containing uniformly labeled  $^{13}\text{C}_6$  glucose ( $[\text{U-}^{13}\text{C}_6]\text{Glc}$ ) and the isotopic composition of the alkamides was then analyzed using LC-MS and GC-MS. Valine was shown to have extensive labeling by  $[\text{U-}^{13}\text{C}_6]\text{Glc}$ , extending to an  $M + 5$  peak with an abundance of 7.4%. Alkamides were analyzed using the GC-MS. The  $m/z$  peak at 167 for alkamides  $^{\text{N}}\text{iBu}/12:4\Delta^{2\text{E},4\text{E},8\text{Z},10\text{E}}$  revealed enhancement through a signal at  $M + 10$ . The  $M + 10$  peak arises from fully labeling the isobutylamine moiety plus the six carbon atoms for the acyl moiety of the alkamides. The incorporation of the labeled isobutylamine moiety provided initial evidence that valine is a precursor to alkamide biosynthesis.

To test the involvement of *de novo* branched chain amino acid (BCAA) biosynthesis, the same experiment was completed except chlorsulfuron was used. Chlorsulfuron is known as an inhibitor of acetolactate synthase, which catalyzes the first step of branched-chain amino acid (BCAA) synthesis. This resulted in a decrease in seedling growth by 50% over a 72 h period. Alkamide production and signals above the  $M + 6$  ion of the  $m/z$  167 ion were reduced. Also,  $[\text{U-}^{13}\text{C}_6]\text{Glc}$  showed minor incorporation into Val ( $\leq 1\%$ ) and alkamides. However seedling growth and alkamides production could be restored to wild type level if supplemented with 15 mM Val and 15

mM Ile. Both these experiments further suggested the amine moiety of this alkamides originates from Val biosynthesis. Furthermore, these experiments established an experimental system in which alkamides biosynthesis was dependent on externally proved BCAAs precursors.

In order to determine the origin of the C1'-N bond of alkamides, [2-<sup>13</sup>C, <sup>15</sup>N]Val was fed to seedlings. LC-MS analysis of <sup>N</sup>iBu/12:4Δ<sup>2E,4E,8Z,10E</sup> revealed an M + 2 peak with an accumulated level of 4.2% enhancement above the control. This verified that the C-N bond from valine remains intact in alkamides (Figure 41A). The mechanism by which the isobutylamine moiety is generated can be probed by examining the fate of the hydrogen at C-2 of Val. Fully deuterated-Val, [<sup>2</sup>H<sub>8</sub>]Val, was feed to seedlings and alkamide <sup>N</sup>iBu/12:4Δ<sup>2E,4E,8Z,10E</sup> was examined by electrospray MS. The MS exhibited two new M + 7 and M + 8 peaks at a ratio of 1:1.1 (Figure 41C). The lower mass peak of M + 7 can be attributed to background transamination of valine [1]. It may be noted that this is consistent with an increase in the M + 1 / M ion ratio (0.8%-1.7%) from the <sup>13</sup>C-<sup>15</sup>N-Val labeling experiment, which had been ascribed to minor <sup>15</sup>N scrambling due to amino acid transamination (Figure 41D). Because of the complete incorporation of the deuterium atoms from [<sup>2</sup>H<sub>8</sub>]Val, the most likely mechanism for generating the amine moiety was anticipated to be a pyridoxal phosphate (PLP)-dependent decarboxylation [1]. If decarboxylation requiring oxidation at C-2 had occurred, the incorporation of the deuterium at C-2 would not have been observed. The PLP-mediated mechanism would lead to isobutylamine as an intermediate, but isobutylamine is not observed to accumulate in *Echinacea* seedlings. However, when 0.25 mM [<sup>2</sup>H<sub>8</sub>]isobutylammonium chloride is used in the growth medium, there was the appearance of a peak 9 amu larger than the parent ion in the LC-MS spectrum of 12:4Δ<sup>2E,4E,8Z,10E</sup>. Of the accumulated alkamide, <sup>N</sup>iBu/12:4Δ<sup>2E,4E,8Z,10E</sup>, the labeling of M + 9 was 3.7% (Figure 43C) [1]. This established isobutylamine can function as an intermediate in the alkamide biosynthetic pathway. Finally, [<sup>2</sup>H<sub>10</sub>]Ile was fed to *E. purpurea* seedlings that provided parallel evidence to the [<sup>2</sup>H<sub>8</sub>]Val observations, thus confirming Ile as an intermediate in the synthesis of the 2-methylbutylamide moiety of alkamides.

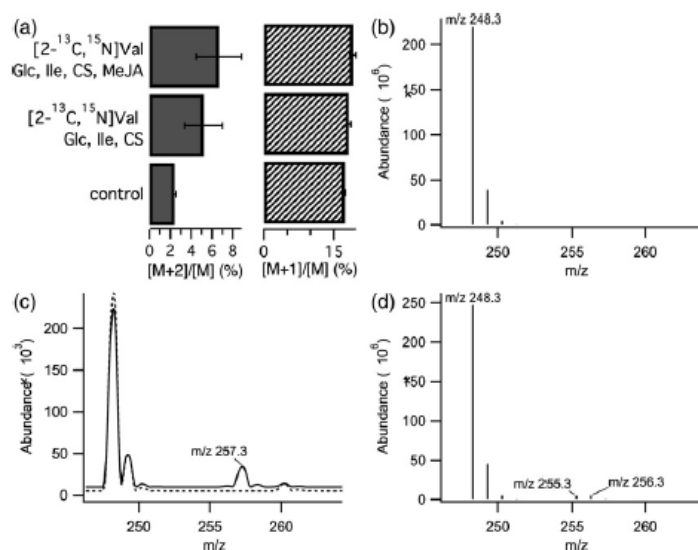


Figure 41: Enhanced isotopologue ratios in  $N_{15}Bu/12:4\Delta^{2E,4E,8Z,10E}$  resulting from feeding of *E. purpurea* seedlings with isotopically labeled precursors. (A, B) A. Feeding of  $[2-^{13}C/^{15}N]$  valine significantly increased the abundance of M + 2 versus M + 1 species, indicating the retention of the C-N bond from the precursor, corrected for background Val transamination. MeJA and the presence of additional carbon sources, such as glucose (Glc), isoleucine (Ile), and chlorsulfuron (CS) had no significant effect on isotope incorporation. Isotopic envelope revealed by full-scan QqQ mass spectroscopy of  $N_{15}Bu/12:4\Delta^{2E,4E,8Z,10E}$  isolated from seedlings incubated with A. natural abundance valine and isoleucine and B. media supplemented with  $[^2H_8]$ valine, unlabeled isoleucine, and in the presence of chlorsulfuron. Ions with  $m/z$  256 and 255 correspond to the M + 8 and M + 7 ions resulting from incorporation of eight deuterium atoms from the decarboxylated form of  $[^2H_8]$ valine or the incorporation of seven deuterium atoms following the metabolic exchange of  $[^2H_8]$ valine with 2-ketoisopentanoate via transamination that results in the loss of one of the deuterium atoms. C. Isotopic envelope from plants cultured in standard media (dashed line) or media supplemented with  $[^2H_9]$ isobutylamine hydrochloride (solid line); the M + 9 ion at  $m/z$  257 indicates that this precursor is incorporated into the alkamide without loss of any deuterium atoms.

Due to the Val and Ile-labeling experiments and the amino acid decarboxylases that catalyze PLP-dependent reactions, the Nikolau group looked to identify an enzyme that generating the amine moiety of alkamides. They used two strategies to help identify this enzyme; one based on sequence homology among PLP-dependent amino acid decarboxylases and the other based on the expectation that the expression of the genetic elements responsible for alkamides biosynthesis should positively correlate with alkamide accumulation [1].

First, the NCBI conserved-domain database for *E. purpurea* transcriptome data was searched for genes that encoded amino sequences related to known PLP-dependent amino acid decarboxylases, and found 21 *E. purpurea* transcripts to encode such enzyme [1]. A homology-based phylogenetic tree was constructed that contained 595 PLP-dependent enzymes, which were divided into 4 clades (Groups A-D) [1]. Group A contained seven of the *E. purpurea* sequences. Group B contained the biochemically characterized valine decarboxylase from *S. viridifaciens* (VImD) and a single *E. purpurea* sequence (contig epa\_952). Group C contained a single *E. purpurea* sequence (epa\_3096). Group D did not contain PLP-dependent enzyme sequences from plants.

Second, the accumulation of the alkamide  $N_iBu/12:4^{2E,4E,8Z,10E}$  among 20 tissues and organs was correlated with the expression patterns of the 21 *E. purpurea* transcripts [1]. This led to the identity of a single putative PLP-dependent amino acid decarboxylase (contig Epa\_11279) with a correlation coefficient of 0.8 to the accumulation of alkamide  $N_iBu/12:4\Delta^{2E,4E,8Z,10E}$ , whereas all other correlation coefficients were below 0.3.

From the homology and correlating expression patterns, Epa\_952 and, Epa\_11279 were identified as a potential PLP-enzymes [1]. This led to the heterologous expression of Epa\_952 and Epa\_11279 ORFs in *Escherichia coli* and assaying for decarboxylation enzymatic activity in hopes to identify the BCAA decarboxylase. Epa\_952 did not support enzymatic decarboxylation of either valine or isoleucine. Epa\_11279 did catalyze the decarboxylation of both valine and isoleucine, and thus must encode the BCAA decarboxylase that can generate the amine moieties of alkamides. Further biochemical characterization identified that Epa\_11279 encodes a type II BCAA decarboxylase because it is dependent on the PLP cofactor. The spatial expression pattern of the *Echinacea* BCAA decarboxylase was investigated at the mRNA and protein level. The roots were found to accumulate the largest concentration of alkamide  $N_iBu/12:4\Delta^{2E,4E,8Z,10E}$  and have the highest expression of the BCAA decarboxylase at both the mRNA and protein levels [1]. The labeling experiments established that Val and Ile are direct precursors to the amine moiety of alkamides in *Echinacea*. This led to the discovery and characterization of a new BCAA decarboxylase. The BCAA converts the amino acids Val and Ile to isobutylamine and 2-methylbutylamine, respectively, which



we propose is then subsequently acylated with fatty acid derivatives to generate the alkamides found in *Echinacea*.

### 3.3.2 Synthesis of deuterated aldehydes

It has yet to be definitively determined whether the alkamides themselves are being unsaturated or if the fatty acids undergo desaturation before amide bond formation. All eight alkamides previously shown in Figure 30 have  $\alpha,\beta$ -unsaturation, and the majority of them have a *trans* configuration. To probe the biosynthetic pathway,  $\alpha,\beta$ -unsaturated,  $d_3$ -labeled fatty acids and alkamides needed to be synthesized. Alkamides and fatty acids with varying chain lengths of  $C_{12}$ ,  $C_{13}$ , and  $C_{14}$  were prepared and tested. Routes were developed in order to incorporate a *trans* double bond at the C2 position and  $d_3$  label at the end of the acyl chain in the biosynthetic probes. Retrosynthetically, it was expected that the two types of compounds could be synthesized in a similar manner (Figure 42). A Horner-Wadsworth-Emmons (HWE) reaction could be utilized for the *trans* double bond with a phosphonate and an aldehyde with  $d_3$ -labeling at the end of the acyl chain. After the formation of the  $\alpha, \beta$ -*trans* double bond, the ester would be hydrolyzed to the fatty acid, respectively.

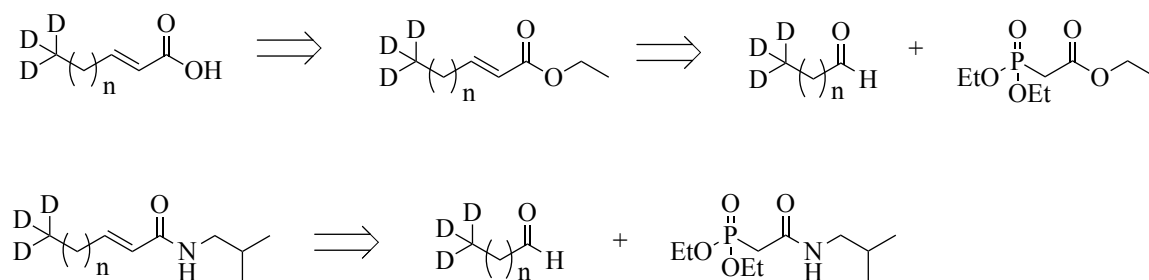


Figure 42: Retrosynthetic approach to synthesizing  $[d_3]$ -labeled  $C_{12}$ ,  $C_{13}$ , and  $C_{14}$   $\alpha,\beta$ -unsaturated alkamides and fatty acids.

First the aldehydes were synthesized to ensure that the labeling could be completed at the end of the acyl chain. To synthesize the aldehydes, the  $CD_3$  was installed at the  $\omega$ -end of the acyl chain (Figure 43). In the first step to produce decanal

and undecanal, starting materials underwent a zipper reaction to produce terminal alkynes **28** and **29**. For dodecenal, the starting carboxylic acid had a terminal alkyne and therefore the carboxylic acid was reduced to form alcohol **30**. With this reduction, the general procedures leading to all three aldehydes were then the same. The alcohol was protected with dihydropyran to form the tetrahydropyranyl ether. In testing the synthetic route, the protected alcohol was methylated with  $\text{CH}_3$  and carried through to the aldehyde. Once the unlabeled aldehyde was successfully made, the protected alcohols (**31**, **32**, **33**) were methylated with  $\text{ICD}_3$  to put the deuterium at the end of the acyl chain. Reduction of the alkyne was attempted under a hydrogen atmosphere with Wilkinson's catalyst, but this catalyst caused deuterium scrambling. Therefore, the alkyne was reduced with  $\text{H}_2$  and Pd/C. The THP-protecting group was removed by alcoholysis with *p*-toluenesulfonic acid and methanol. Finally the alcohol was oxidized to the aldehyde by using Dess-Martin periodinane.

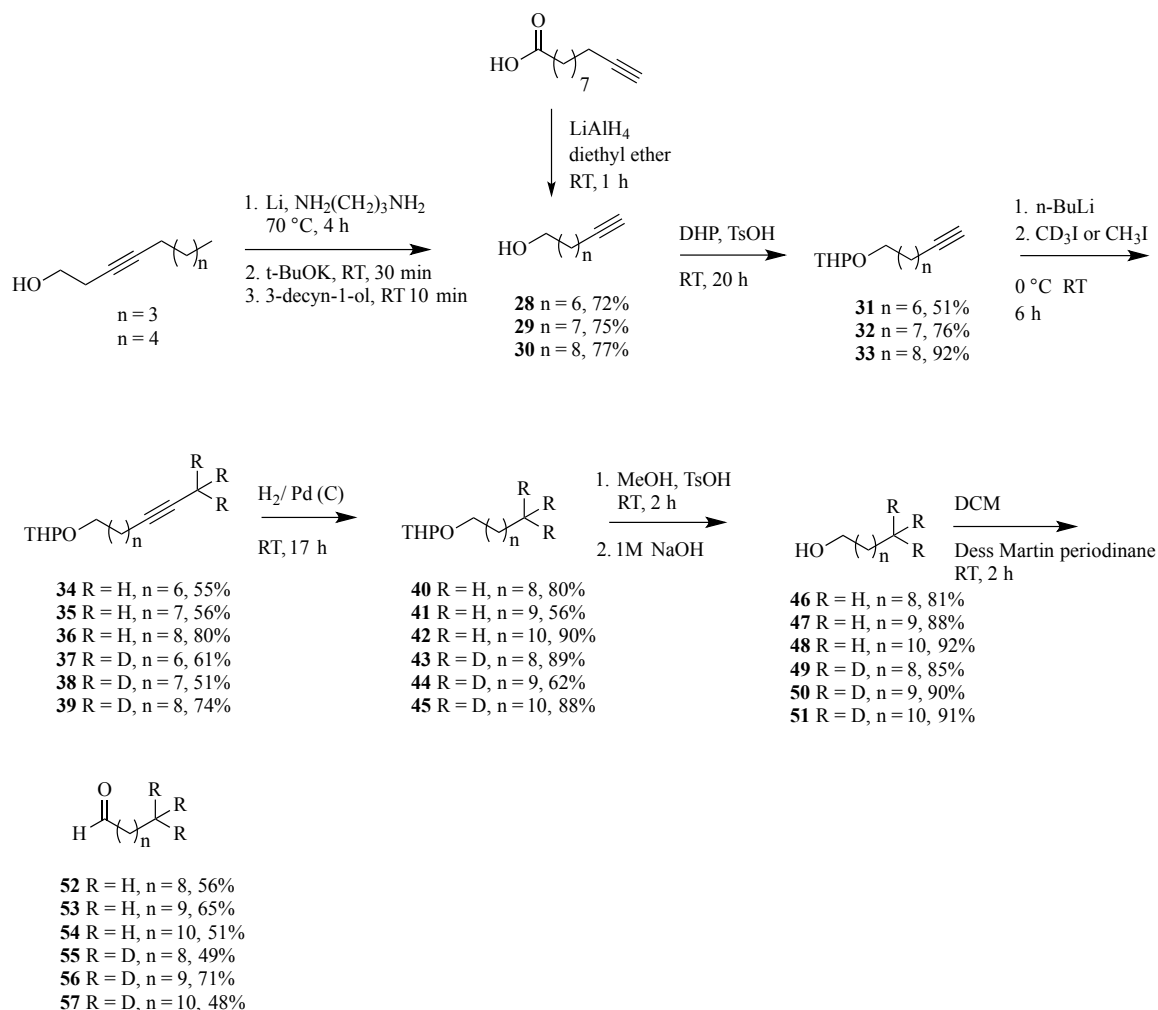


Figure 43: Synthesis of aldehydes with and without deuterium labeling at the  $\omega$ -end of the acyl chain.

### 3.3.3 Synthesis of $\alpha,\beta$ -unsaturated alkamides

Alkamides  $^{\text{N}}\text{iBu}/12:1\Delta^{2\text{E}}$ ,  $^{\text{N}}\text{iBu}/13:1\Delta^{2\text{E}}$ ,  $^{\text{N}}\text{iBu}/14:1\Delta^{2\text{E}}$  were synthesized initially without the labeled aldehyde (Figure 44). Once the synthetic pathway was optimized without any labeling, the  $[\text{}^2\text{H}_3]$ -aldehydes were used. To synthesize each alkamide, ethyl 2-bromoacetate was converted by an Arbuzov reaction to the alkyl phosphonate, **58**, with triethylphosphite. The alkyl phosphonate was hydrolyzed to yield 2-(diethoxyphosphoryl)acetic acid **59**. A carboxylic acid cannot be used directly to make an amide because amines are basic; activation of the acid to the acyl chloride normally

circumvents this problem. Initially, we used  $\text{SOCl}_2$  to make the acid chloride before using isobutylamine to form the amide. However, this resulted in a low yield and side products; therefore, CDI was used to make an anhydride before being treated with isobutylamine to complete the isobutylamide moiety of the alkamides. Finally, **60** was deprotonated and either commercial aldehydes or synthetically derived deuterated aldehydes from Figure 45 was used to complete HWE reaction and insert form the  $\alpha,\beta$ -trans double bond with deuterium at the end of the acyl chain.

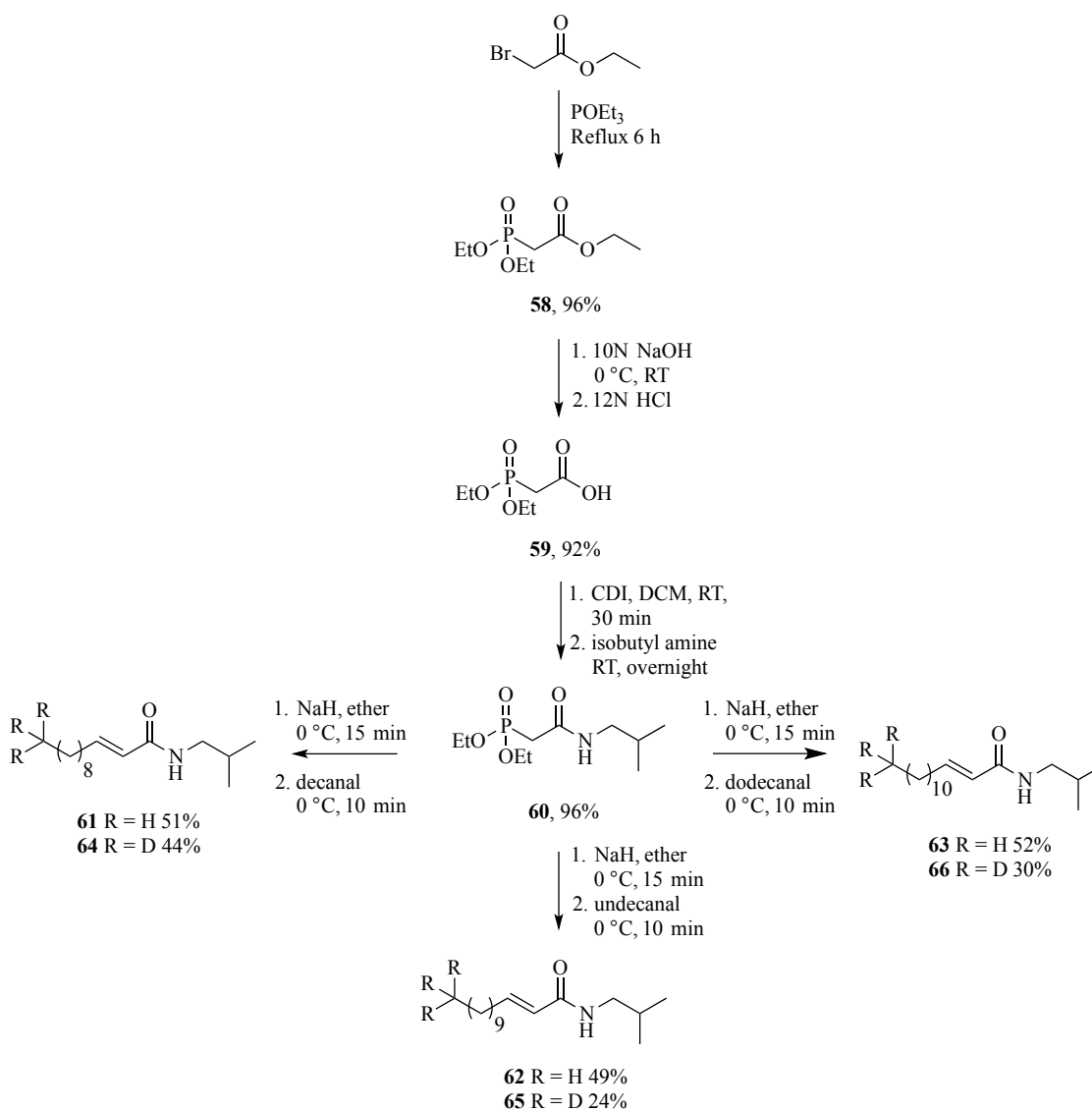


Figure 44: Alkamide synthesis of  $^{\text{N}}\text{iBu}/12:1\Delta^{2\text{E}}$ ,  $^{\text{N}}\text{iBu}/13:1\Delta^{2\text{E}}$ ,  $^{\text{N}}\text{iBu}/14:1\Delta^{2\text{E}}$

3.3.4 Synthesis of  $\alpha,\beta$ -unsaturated fatty acids

A synthetic pathway was designed to incorporate  $d_3$ -labeling at the end of the acyl chain of 12:1, 13:1, and 14:1 (Figure 45). Phosphate **58** was deprotonated with NaH to form the carbanion and this underwent a HWE reaction with the respective aldehydes from Figure 45 to form the  $\alpha,\beta$ -unsaturated ethyl esters **67** – **72**. Each ester was hydrolyzed with 10% KOH and after workup with concentrated HCl afforded the  $\alpha,\beta$ -unsaturated fatty acids **73**-**78**.

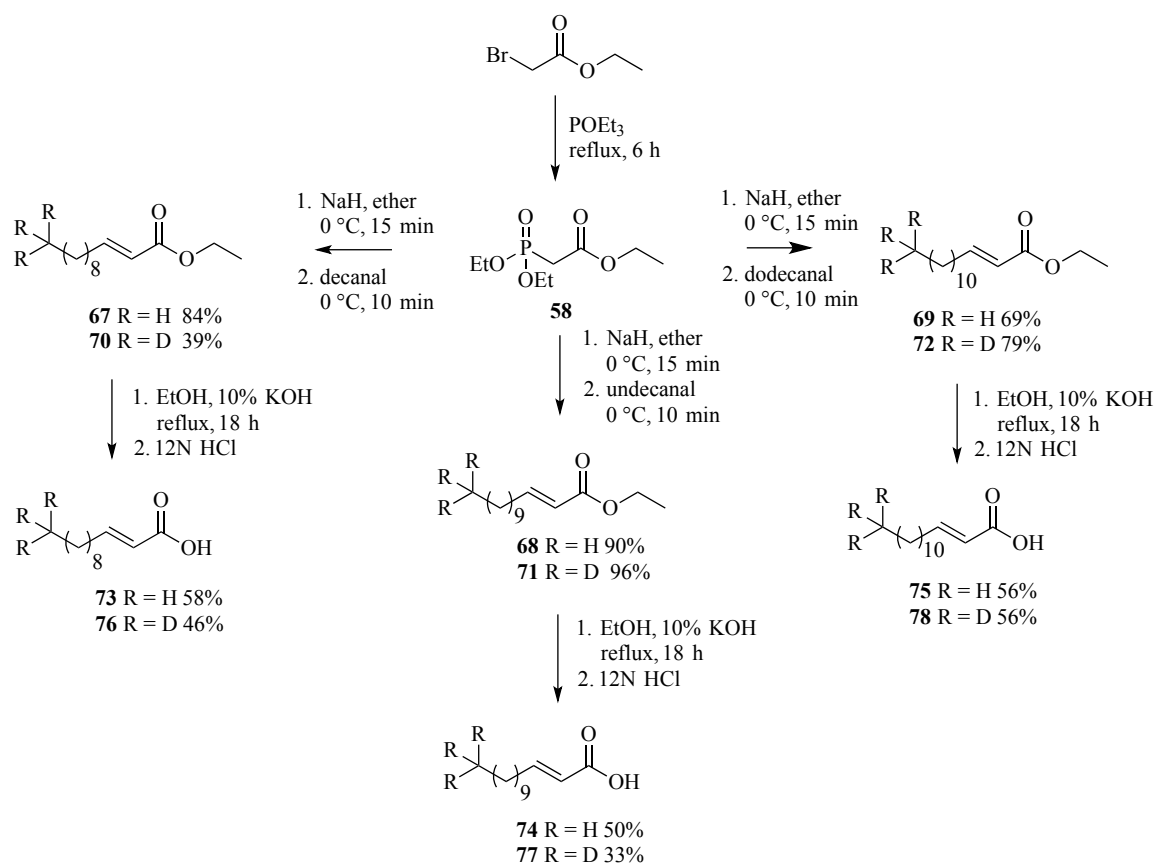


Figure 45: Synthesis of fatty acids 12:1, 13:1, 14:1,  $[d_3]$ -12:1,  $[d_3]$ -13:1, and  $[d_3]$ -14:1.

### 3.3.5 Determination of isotope distribution patterns in alkamides isolated from plants provided deuterated $\alpha,\beta$ -unsaturated fatty acids and $\alpha,\beta$ -unsaturated alkamides

Toxicity tests were completed with unlabeled synthesized compounds with the expectation that the  $d_3$ -labeled compounds would behave similarly. Conditions for the media were prepared in a manner identical to that of the labeled amino acid feeding experiments. Initially, DMSO was used to make stock solutions, but the solvent was switched to 2-methoxyethanol. When DMSO was placed on the leaf as an control it was found to be toxic to the leaves as evidenced by the leaf turning brown and wilting.

In the initial toxicity results, the 14:0 fatty acid was toxic to the plant at all concentrations tested (5 – 100 mM, 2  $\mu$ L application). Further serial dilutions of the fatty acid were investigated and the fatty acid was found to be no longer toxic at concentrations of 0.5 mM or lower. Contrastingly, the alkamide was not toxic to the plant at any of the concentrations tested ( $\leq$ 100 mM). However, at higher concentrations a white crystal was observed to accumulate on the leaves indicating that a significant amount of the alkamide was not absorbed into the leaf. The alkamide experiment presents a problem because if the synthesized material failed to enter the plant, labeling would not occur and a false negative could easily result.

The Chapman group previously published a study in which they fed *N*-acylethanolamines (NAEs) to plants at concentrations of 40  $\mu$ M and 400  $\mu$ M by adding the NAEs as DMSO stock solutions to the media [196]. In order to feed the plants the synthesized compounds through the media, glass plates were used without filter paper. Since the toxicity levels had already been tested on the leaves above the concentrations of the Chapman paper, the  $d_3$ -labeled compounds were used to at a final concentration of 500  $\mu$ M in the liquid media.

After growing for 3 days, the alkamides were extracted and analyzed by LC/MS. The media was also analyzed to determine if any of the synthetic probe was still present. A positive outcome would lead to an  $M + 3$  peak in the isotopic envelope for the parent ions of BA 6 ( $^{N_1}\text{iBu}/13:2\Delta^{2E,7Z,10a,12a}$ ) and BA8 ( $^{N_1}\text{iBu}/12:4\Delta^{2E,4E,8Z,10E}$ ) / BA9 ( $^{N_1}\text{iBu}/12:4\Delta^{2E,4E,8Z,10Z}$ ). However, for [ $^2\text{H}_3$ ]- $^{N_1}\text{iBu}/12:1\Delta^{2E}$ ,  $^{N_1}\text{iBu}/13:1\Delta^{2E}$ , and  $^{N_1}\text{iBu}/14:1\Delta^{2E}$  the envelope of only extended to  $M + 2$ . When the media was tested

deuterated alkamide was found in the media. Similar to the applications to leaves, no evidence was gained that alkamides entered the plant, thwarting any examination of whether less-unsaturated alkamides might be intermediates on the pathway. Initial results showed no labeling, but future experiments to try and certify the plants take up the alkamides needs to be explored to confirm these results. For [ $^2\text{H}_3$ ]-12:1, 13:1, 14:1, labeling of BA6 and BA 8/9 should have incorporated the  $\text{d}_8$ -valine used as a control and labeling with the fatty acids would have resulted in a  $M + 11$  envelope. However,  $M + 7$  and  $M + 8$  peaks were visible as expected, but not  $M + 11$ . When the media was analyzed, there was no labeled fatty acid present. Unlike the alkamides, the roots do seem to be taking in the fatty acids, but they do not look to have intact incorporation into alkamide biosynthesis.

Though the synthesized [ $^2\text{H}_3$ ] compounds did not show any labeling, additional work needs to be completed in order to insure the alkamides are entering the plant cell. Also, based on the media, the fatty acids do look to be taken up by the plant, but not incorporated into the alkamide pathway. However, if the plant is using the synthesized fatty acid it should be found in other pathways. For example, running FAMES analysis on the plant extract would allow us to see incorporation of the  $\text{d}_3$ -fatty acids into the fatty acid biosynthesis pathway. Further studies need to be completed in order to uncover the origin of the  $\alpha,\beta$ -double bond found in alkamides of *E. purpurea*.

### 3.3.6 Synthesis of 12:1, 12:2, and 12:4 used for acyl-CoAs

The same synthesis shown in Figure 45 was used to make (*E*)-dodeca-2-enoic acid (**73**). (*2E,4E*)-Dodeca-2,4-dienoic acid, (*2E*)-butenoic acid was esterified to methyl ester **78** (Figure 46). Allylic bromination of **78** was completed in excellent yield using *N*-bromosuccinimide (NBS). Heating of methyl (*E*)-4-bromobut-2-enoate with triethylphosphite provided phosphonate, **80**. Horner-Wadsworth-Emmons (HWE) reaction was completed by deprotonating **80** with NaH followed by the addition of octanal in order to synthesize the second double bond in the acyl chain. Finally, the methyl ester (**81**) was hydrolyzed to yield *2E,4E*-dodeca-2,4-dienoic acid, **82**.

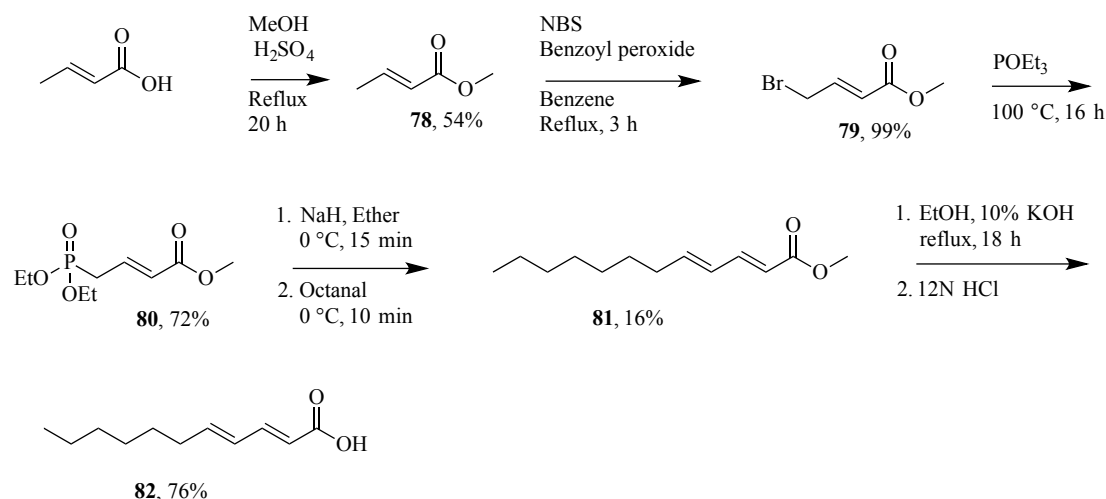


Figure 46: Synthesis of 2E,4E-dodeca-2,4-dienoic acid.

For the synthesis of (2E,4E,8Z,10E)-dodeca-2,4,8,10-tetraenoic acid, ethyl 4-bromobutanoate was first treated with triphenylphosphine to form phosphonium salt **83** (Figure 47). Potassium *tert*-butoxide was used to deprotonate **83** and crotonaldehyde were coupled by a Wittig reaction to form **84**. Ethyl (4Z,6E)-octa-4,6-dienoate was carefully reduced at -80 °C with a dropwise addition of DIBAL to yield the aldehyde **85**. The phosphonate, **80**, from the synthesis of 12:2 was deprotonated with sodium hydride and (4Z,6E)-octa-4,6-dienal was added. Though different conditions were tested, the yield for this HWE reaction was low to yield methyl (2E,4E,8Z,10E)-dodeca-2,4,8,10-tetraenoate (**86**). Finally the methyl ester was hydrolyzed to (2E,4E,8Z,10E)-dodeca-2,4,8,10-tetraenoic acid.



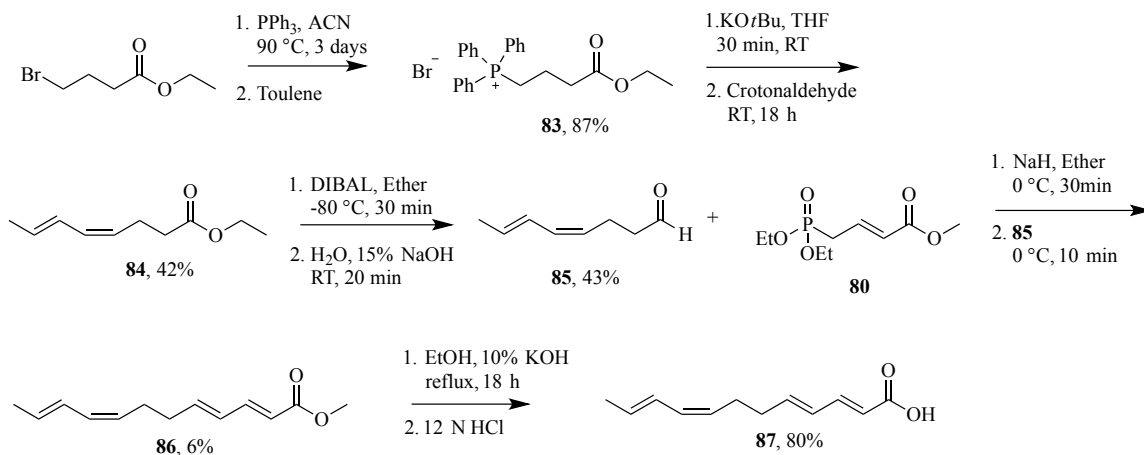


Figure 47: Synthesis of (2E,4E,8Z,10E)-dodeca-2,4,8,10-tetraenoic acid.

Acyl-CoAs are central metabolic intermediates in plants and are formed by several reactions including the condensation reaction of free fatty acids and CoA catalyzed by acyl CoA synthetase (ACS) and importantly in plants acyl CoA:lysophosphatidylcholine acyl transferase (LPCAT) converts other glycerolipids to acyl CoAs [197, 198]. However, the non-enzymatic preparation of these acyl-CoAs was not trivial. Commercial 12:0 was used while the procedure was being optimized. Initial attempts to synthesize 12:0-CoA were completed with a procedure modified by Sanchez [199]. Extreme caution had to be taken during the neutralization and precipitation of the acyl-CoA. If an excess amount of 2 M HCl was added, the acyl-CoA was completely hydrolyzed. The highest percent yield achieved using this unmodified procedure was around 13%. To optimize the yield, three key aspects of the procedure were examined; the formation of the mixed anhydride, formation of the acyl-CoA, and purification of the acyl-CoA.

To start, the extent of formation of the mixed anhydride was checked. After evaporation of the methylene chloride solvent, NMR analysis showed that free fatty acid was still present. In the procedure of Levoine, 2,6-lutidine is used instead of triethylamine [200]. When using this procedure, no mixed anhydride was seen via NMR. Therefore different conditions were tested using the Sanchez procedure as a starting point. The solvent was changed from DCM to THF. Triethylamine and ethylchloroformate were

added in direct succession instead of having a 30-minute stir in between the two. Finally, the amount of triethylamine and ethyl chloroformate was doubled to help drive the reaction to completion. All of these conditions slightly increased the percent yield, but further analysis was done on the formation of the acyl-CoA and purification.

The CoA was the limiting reagent to confirm all the CoA was being consumed, after neutralization, the supernatant was tested via UV-Vis for CoA absorption. Free CoA was found to be present in the supernatant. Therefore, the mixture of THF and CoA solution was mixed longer to ensure the mixed anhydride and CoA were interacting in the heterogeneous mixture. The reaction seemed to be optimized after 2 hours of vigorous stirring. After a few attempts where no product was produced, the conditions of the mixing the anhydride with the CoA solution were adjusted to that reported in the *CRC Handbook of Chromatography* [201]. The base was changed from  $\text{KHCO}_3$  to  $\text{NaHCO}_3$ . It was noted that a homogenous solution of water:THF must be achieved in order for the reaction to mix properly. Collectively, changing the solvent to THF, doubling the amount of triethylamine and ethylchloroformate, changing the base to  $\text{NaHCO}_3$  for the CoA solution, and 2 h of vigorous stirring of the anhydride/CoA mixture, the yield improved to 32%.

After purification, the acyl-CoA was dissolved in 80 mM  $\text{NaHCO}_3$ . Some of the solutions were cloudy meaning additional purification was needed. Solid-phase extraction with a  $\text{C}_{18}$  column was completed to eliminate unreacted starting materials. After loading the sample onto the column, the column was washed with  $\text{H}_2\text{O}$ . The product was eluted with 50% MeOH. When free CoA was tested on the column, it eluted with the  $\text{H}_2\text{O}$  washes. When free fatty acid was tested, it was not seen in any of the washes until the CoA concentration was tested. By UV-Vis the percent yield was around 50%, but when the concentration was determined by GC-MS the percent yield was over 200%. Therefore unreacted fatty acid was not being purified from the acyl-CoA sample via the  $\text{C}_{18}$  column. Finally, a new purification procedure from Al-Arif was followed using 5%  $\text{HClO}_4$  to precipitate acyl-CoA from solution followed by washes with 0.8%  $\text{HClO}_4$ , acetone, and diethyl ether [202]. The yield was optimized at 39%.

### 3.3.7 Concentrations of acyl-CoAs

Concentrations of each acyl-CoA were calculated using absorbance from the UV-VIS and the correlations areas of the FAMES from GC-MS analysis. Previous studies have shown concentration analyses of acyl-CoAs by UV-Vis [195, 197]. Following the procedure by Merrill *et. al.*, the stock solutions were diluted in a phosphate buffer and the absorbance was recorded at 260 nm. To test the procedure, commercial 16:0-CoA was used. A stock solution was made and absorbance was recorded in the phosphate buffer. The UV-Vis technique showed less than a 10% error with 16:0-CoA, which verified the use for our purposes. However, if the free-CoA was not entirely removed during the extraction process, the quantity of the acyl-CoA would be inflated. Therefore, gas chromatography was used as an independent technique to verify the acyl-CoA concentrations. For GC-MS analysis, the acyl-CoAs were derived to their methyl esters, using 15:0 as an internal standard. The relative response for the GC/MS in the FAMES analysis was determined for the esters of 12:1 and 12:2, with respect to 15:0. The slopes from each concentration curve were used to determine the sensitivity of 12:X compared to the internal standard, 15:0 (Table 4). There was a limited amount of 12:4 free fatty acid available from our synthesis; therefore, the sensitivity of methyl 12:2 was used to calculate the concentrations of methyl 12:4. FAMES analysis was completed with each acyl-CoA and 15:0 as a standard. The correlation areas were determined and Equations 2-4 were used to calculate the concentration of the acyl-CoAs (Table 5).

Table 4: Calculated concentrations of acyl-CoAs (12:1, 12:2, and 12:4) determined by UV-Vis and GC-MS.

Acyl-CoA	UV-Vis	GC-MS	
	Concentration of Acyl-CoA (mM)	Sensitivity	Concentration of Acyl-CoA (mM)
12:1CoA	2.7 mM	0.402599	3.2 mM
12:2CoA	35.4 mM	0.1487589	11.2 mM
12:4CoA	14.3 mM	0.1487589	10.5 mM

The concentrations between 12:1CoA and 12:4CoA were reasonable between the two analyses. The concentration for 12:2CoA was much higher from the UV-Vis than the GC-MS. Therefore, there must be excess CoA in the sample that was not washed away during purification. Because CoA exists in cells, the excess CoA was not expected to have deleterious effect on the assays and therefore the sample was still used for further analysis. The concentrations from the GC-MS were used for the enzymatic feedings.

### 3.3.8 Activity assays of acyl-CoAs with EpLigase

The Nikolau group isolated an enzyme, EpLigase, in the hope that it plays a key role in alkamide biosynthesis by catalyzing the reaction to form the amide bond. As previously discussed, the isolation of the BCAA decarboxylase Epa\_11279 was shown to decarboxylate valine and isoleucine to form isobutylamine and 2-methylbutylamine providing the alkyl amide nitrogen in alkamides. EpLigase was tested using activity assay with fatty acid acyl-CoAs and isobutylamine for the formation of the amide bond. Initial activity of the protein was checked using each fatty acid CoAs chain-length from C<sub>3</sub>-C<sub>18</sub>; only C<sub>12</sub>-CoA showed any activity. The synthesized acyl-CoAs, 12:1-CoA, 12:2-CoA, and 12:4-CoA were also tested with the protein and 20 mM isobutylamine. The  $k_{\text{cat}}$  and  $V_{\text{max}}$  were compared to 12:0-CoA (Table 5).

Table 5: Enzyme activity of EpLigase with unsaturated CoAs.

	C12:0 CoA		C12:1 CoA		C12:2 CoA		C12:4 CoA	
	Isobutylamine	2-Methylbutylamine	Isobutylamine	2-Methylbutylamine	Isobutylamine	2-Methylbutylamine	Isobutylamine	2-Methylbutylamine
Conc of protein (M)	1.39E-08	1.39E-08	1.39E-08	1.39E-08	9.96E-07	9.96E-07	9.96E-07	9.96E-07
V max (mol/s* L)	1.097E-07 +/- 3.45E-08	6.15E-08 +/- 8.12E-09	3.17E-08 +/- 4.78E-09	1.42E-08 +/- 1.55E-09	1.21E-07 +/- 1.90E-08	3.22E-10 +/- 5.14E-11	0.00	1.80
kcat	7.87	4.41	2.27	1.02	0.12			
kcat/Km	60269.80	39844.86	33078.28	15859.70	607.17			

Initially 12:2-CoA and 12:4-CoA showed almost no activity with the protein. When the higher concentrations were used, the enzymatic activity was detected. However, looking at the catalytic efficiency values, the  $k_{\text{cat}}/K_m$  values for 12:2-CoA and 12:4-CoA were significantly lower than 12:1-CoA. 12:1-CoA has a much higher  $k_{\text{cat}}/K_m$ , and the activity is comparable to that of 12:0-CoA.

Finally, the products from the *in vitro* assays were confirmed by GC-MS. All three synthesized acyl-CoAs produced the respective alkamide. It has been determined that EpLigase shows activity with acyl-CoAs of 12 carbons. However, after one unit of unsaturation, the  $K_m$  drastically increases showing the enzyme has low affinity for those substrates. It is still unknown whether the initial unsaturation in the acyl chain comes before or after the amide bond is formed, but these results may indicate that the majority of desaturation comes after the amide bond. As  $\alpha,\beta$ -unsaturation occurs in all alkamides, there is a possibility the  $\alpha,\beta$ -unsaturation may come before the amide bond is formed.

### 3.4 Conclusion

Through experiments described in this chapter, further insight was gained into several aspects of alkamide biosynthesis in *E. purpurea*. It was determined that the amide moiety is synthesized from the decarboxylation of isoleucine and valine. Both [ $^2\text{H}_8$ ]Val and [ $^2\text{H}_{10}$ ]Ile were both fully labeled into  $^{\text{N}}\text{iBu}/12:4\Delta^{2\text{E},4\text{E},8\text{Z},10\text{E}}$ . The decarboxylation of valine leads to isobutylamine. Though isobutylamine does not seem to accumulate in *Echinacea* seedlings, [ $^2\text{H}_8$ ]isobutylamine chloride does incorporate into  $^{\text{N}}\text{iBu}/12:4\Delta^{2\text{E},4\text{E},8\text{Z},10\text{E}}$ . Further proving the amide moiety is derived from valine and isoleucine. This contributed to our collaboration with the Nikolau group that discovered the branched-chain decarboxylase that catalyzes this reaction. Additionally, the Nikolau group isolated a ligase that potentially connects an acyl-CoA with isobutylamine to form the amide moiety. We synthesized three  $\text{C}_{12}$ -CoAs with varying degrees of unsaturation. Along with standard 12:0CoA, 12:1CoA showed promise as the acyl intermediate in the alkamide pathway. Finally, in an attempt to better understand the modification of the unique acyl chain of alkamides, three fatty acids (12:1 $\Delta^{2\text{E}}$ , 13:1 $\Delta^{2\text{E}}$ , and 14:1 $\Delta^{2\text{E}}$ ) and

three alkamides ( $^N\text{iBu}/12:1\Delta^{2E}$ ,  $^N\text{iBu}/13:1\Delta^{2E}$ ,  $^N\text{iBu}/14:1\Delta^{2E}$ ) were synthesized with  $d_3$ -labeling at the end of the acyl chain. Though initial studies showed no labeling in  $^N\text{iBu}/12:4\Delta^{2E,4E,8Z,10E}$  from these six synthesized compounds, further studies are needed to ensure and optimize the internalization of the proposed intermediates by the seedlings.

## CHAPTER 4. STRUCTURE ELUCIDATION OF PROPOSED INTERMEDIATES IN THE SAD2 PATHWAY

### 4.1 Introduction

Our goal was to identify the products of SAD2 mutants in order to give insight into the bifunctional character of AsCYP51H10 and thus providing further understanding of the biosynthetic pathway of avenacin A-1. Subcloning of the SAD2 gene from *A. strigora* allowed for its transient co-expression with SAD1 in *Nicotiana benthamiana* leaves [203]. The design of site-directed mutants from SAD2 that were proposed to disrupt the binding of either beta-amyrin or 16 $\beta$ -hydroxy-beta-amyrin was undertaken. Co-expression of each mutants with SAD1, similar to that which had been carried out with the wild-type enzyme, made several products including monooxygenated and deoxygenated triterpenes and possibly a desmethylated product. The goal of this portion of the project was to fully identify the mutant SAD2-derived metabolites and to further evaluate whether the 14-desmethyl product, suggested by GC-MS data, was present. The 14-desmethyl product would be interesting because it would show the conserved mechanistic property of the CYP51 family (Figure 48).

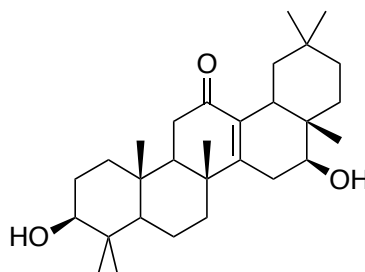


Figure 48: The possible 14-desmethylation product from a SAD2 mutant that would show similar the same mechanistic function as other CYP51 enzymes.

Additionally, it was felt that if 12,13 $\beta$ -epoxy- $\beta$ -amryin and/or 16 $\beta$ -hydroxy- $\beta$ -amryin were identified and could be isolated, it would provide an opportunity to feed single oxidized products back to the wild-type enzyme to see whether 12,13 $\beta$ -epoxy- $\beta$ -amryin can be converted to 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amryin or if 16 $\beta$ -hydroxy- $\beta$ -amryin is converted to 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amryin. This would aid in



determining the sequence of events to form 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin. The structures of intermediates from the wild-type protein and two different mutant proteins were elucidated by several NMR experiments including  $^1\text{H}$ ,  $^{13}\text{C}$ , DEPT, COSY, HSQC, HMBC, and NOESY.

#### 4.1.1 Evolution of enzymes in secondary metabolism

Since the 1970s, knowledge of the enzymes involved in specialized metabolism has expanded. Like primary metabolic enzymes, most plant secondary metabolic enzymes display high substrate specificity for both core pathways and modifying reactions for structural diversity, and there is no evidence of unspecific side reactions [3]. Unlike enzymes from primary metabolism, generally, enzymes of secondary metabolism tend to lack feedback regulation [3]. Many enzymes in secondary pathways are multifunctional such as cytochrome P450s (CYPs) [3]. It is also recognized that enzymes of specialized metabolism are found in 3-5 fold lower concentrations than enzymes of primary metabolism [3], which can provide difficulty in identifying them and understanding their functions.

Oxygenation is a common feature amongst a majority of secondary metabolites [10]. In plants, the cytochrome P450 (CYP) superfamily is the largest class of plant metabolic enzymes, and they most commonly catalyze a variety of monooxygenations, usually by inserting one oxygen molecule into an aliphatic position [203]. There are over 5,100 sequences of plant CYPs that have been identified [204]. They have several important roles in plants such as signaling, defense, polymerization, and the development of complex anatomical and chemical structures [204]. They are excellent examples of the evolution of plant metabolic enzymes.

CYPs share a catalytic center in which iron is coordinated to a thiolate of a conserved cysteine [205]. They are similar in topology and three-dimensional fold, but they vary at the amino acid level [206, 207]. Though CYPs play an important role in secondary metabolism, many of the founding members of the CYP families catalyze primary metabolic reactions [10]. Through gene duplication, members from primary or existing secondary pathways have led to multi-member gene families with a widened array of substrates and greater functional diversity to synthesize new secondary

metabolites [10]. Advances in sequencing technologies have led to an increase in sequencing of a diverse number of aquatic and land species resulting in new plant CYP families being identified [10]. By 1996, plants had already broken the nomenclature system established ten year earlier [208]. To date, there are 127 plant CYP450 families [10].

Plant CYPs were originally divided into two different clades: A-type and non-A-type [209]. It was thought that A-type enzymes participate in the biosynthesis of specialized metabolites; whereas, non-A-type function in lipid and hormone metabolism [209]. Non-A-type CYPs were thought to be more divergent than A-type [210]. Now it is known that non-A-type enzymes are involved in specialized metabolites and this simplified view is inaccurate [10]. Instead, CYP enzymes from land plants can be divided into 11 phylogenetically distinct clans, CYP51, 71, 72, 74, 85, 86, 97, 710, 711, 727 and 746 [10]. The original A-type CYPs are now contained in the CYP71 clan, and non-A-type are spread across the other 10 clans [10]. Either through gene duplication of general metabolic CYPs or divergence from pre-existing specialized pathways, clans have expanded over time [10]. The membership of some clans such as CYP71, CYP85, and CYP86 has expanded more than others like CYP51, CYP710, and CYP74 (Figure 49) [10]. One of the most highly conserved clans is CYP51. In all CYPs across animals, microbes, and plants sequences are less than 30% identical except within the CYP51 family [205]. The CYP51 family is conserved across all phyla with a sequence identity in the range of 30-40% [205]. The CYP51 is the only family in the CYP51 clan and it catalyzes 14 $\alpha$ -demethylation reaction for sterol biosynthesis. Through evolution within the family, changes in the CYP51 sequences have been seen, though less frequent than other families. However, even after changes to sequence, the catalytic function across all phyla is the same. The CYP51 continually remains an interesting family to study due to the preserved metabolic role across phyla and the possibility of diversification in new species [211].

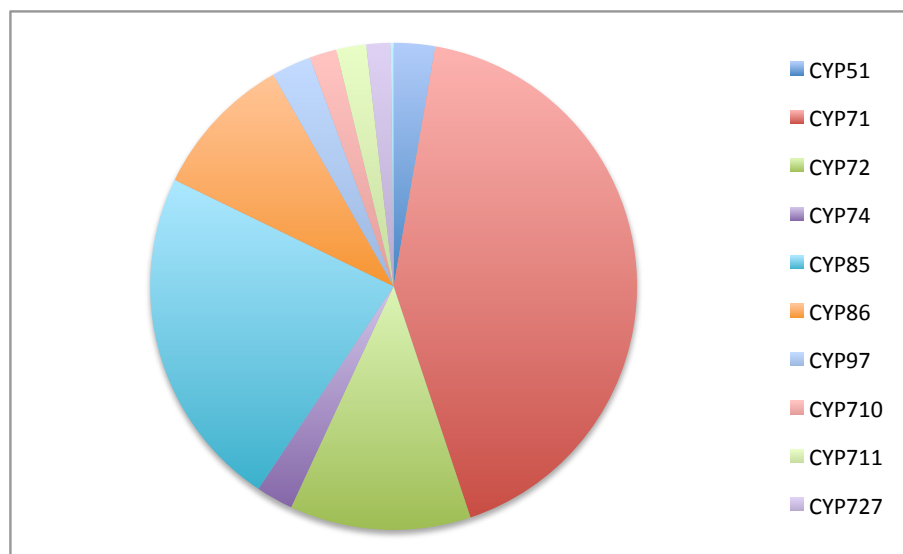


Figure 49: Graphical representation of the occurrence of plant CYP clans across plant phylogeny. CYP746 is not represented because it makes up less than 0.2% of the chart. Reproduced with permission from [10]

One of the oldest and most structurally diverse classes of plant natural products is the terpenoids [10]. In fact, there are more than 43,000 known terpenoids, and most of them are oxygenated at one or more carbon positions (Table 6) [10]. CYPs are most often the enzyme involved in the oxygenation of terpenoids [10]. Oxygenation diversifies the products produced, increases their polarity, and decreases the metabolites volatility, which all can benefit a plant [10]. Currently, it is known that the CYPs used for terpenoid biosynthesis are distributed over 8 of the 11 clans [10]. In four of the clans, the CYPs from primary metabolism have started to proliferate to form specialized terpenoids [10]. This shows that a key feature in the evolution of CYPs is the recruitment from primary pathways [10].

Table 6: The fraction of terpenoids that are oxygenated per class.

Class of Terpenoid	Number of Terpenoids	Percent Oxygenated
C <sub>10</sub> monoterpene	3,816	99.8%
C <sub>15</sub> sesquiterpene	13,211	96.7%
C <sub>20</sub> diterpene	11,609	98.8%
C <sub>30</sub> triterpene	14,505	95.2%

Not only are CYPs involved in the synthesis of terpenoids, but also in the biosynthesis of other specialized metabolites such as phenylpropanoids and alkaloids [10]. Therefore they are a key component of specialized metabolism and essential in the evolution of natural product diversity. As herbivores and pathogens have evolved, plants have developed new defense-related molecules [10]. Oxygenation reactions are a driver of structurally diverse natural products, significantly modulating their physiochemical properties. CYPs have also shown to have catalytic flexibility, which makes them ideal for the evolution of new function. Because they are already a large fraction of plant genes, it is hypothesized these enzymes will continue to be important for plants to adapt to their environment.

#### 4.1.2 Triterpenes

Isoprenoids, also known as terpenes, are compounds composed of isoprene (2-methyl-1,3-butadiene) units (Figure 50). They are produced by plants, animals, and microorganisms and have recognized commercial pharmacological and agricultural value [212, 213]. In plants, isoprenoids are usually found in the essential oils and the C<sub>30</sub> carotenoids provide the orange, yellow, and red pigments found in plants [214]. They can range from small volatile oils with one isoprene unit to natural rubber containing about 4,000 isoprene units [214]. Isoprenoids are broadly classified based on the number of isoprene units in the structure; monoterpenes (C<sub>10</sub>H<sub>16</sub>), sesquiterpenes (C<sub>15</sub>H<sub>24</sub>), diterpenes (C<sub>20</sub>H<sub>32</sub>), triterpenes (C<sub>30</sub>H<sub>48</sub>), tetraterpenes (C<sub>40</sub>H<sub>56</sub>), carotenoids (C<sub>40</sub>H<sub>56</sub>) and polyterpenes (C<sub>5</sub>H<sub>8</sub>)<sub>n</sub> designate some of the most prolific groups [214]. Furthermore within each class compounds can be classified into subclasses based on whether they are acyclic, monocyclic, or bicyclic, as examples [214].

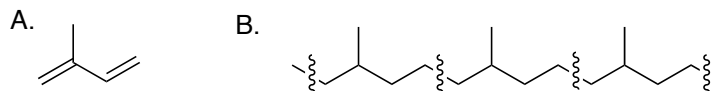


Figure 50: *A.* Structure of an isoprene unit. *B.* Acyclic assembly of head-to-tail linked isoprene units that form the backbone of terpenes.

Triterpenes,  $C_{30}H_{48}$ , are formed from the linkage of 6 isoprene units and are produced by animals, plants and fungi. This large group is structurally diverse with over 200 different skeleton structures known from natural sources or engineered enzymatic reactions [215]. Over 100 distinct skeleton structures are known in the realm of natural products. [215]. The most common structures are 6-6-6-5 tetracyclic, 6-6-6-6-5 pentacyclic, or 6-6-6-6-6 pentacyclic fused rings [215]. However, a wide range of acyclic, monocyclic, bicyclic, tricyclic, and hexacyclic triterpenoids have been also been isolated [215].

The most well known acyclic triterpene is squalene (Figure 51). Squalene is widely distributed in nature. For example, it is found in vegetable oils, in fungi, in human earwax, and it makes up more than half of liver oil in several shark species [214]. Squalene is also an important intermediate in the biosynthesis of cholesterol, an essential component to animal cells, and  $\beta$ -amyrin, a natural product isolated in a variety of plants (Figure 51) [216, 217].

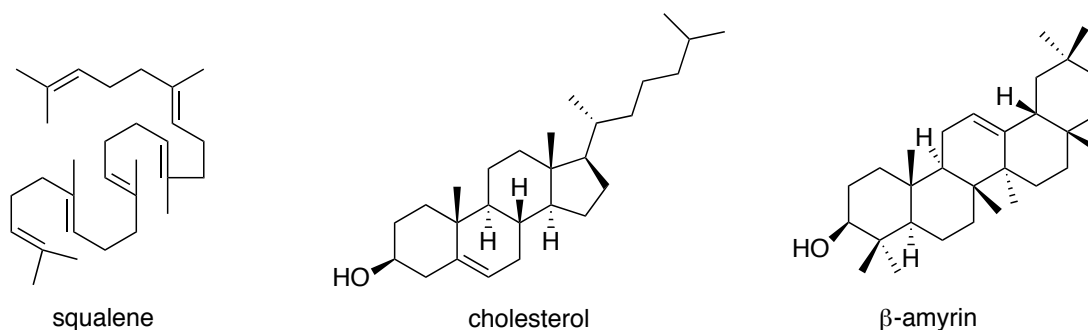


Figure 51: Structures of squalene, cholesterol, and  $\beta$ -amyrin.

Although squalene is an acyclic triterpene,  $\beta$ -amyrin is a pentacyclic triterpene, and cholesterol is categorized as a sterol, they still share a biochemical relationship in their biosynthesis (Figure 52) [214]. The biosynthesis of  $\beta$ -amyrin and cholesterol share a

biosynthetic pathway until the cyclization of 3*R*-2,3-epoxy-2,3-dihydrosqualene where the pathway diverges into sterol biosynthesis and triterpene biosynthesis. Therefore, squalene is an important triterpene that provides a biochemical connection between structurally different compounds, cholesterol and  $\beta$ -amyrin, across kingdoms.

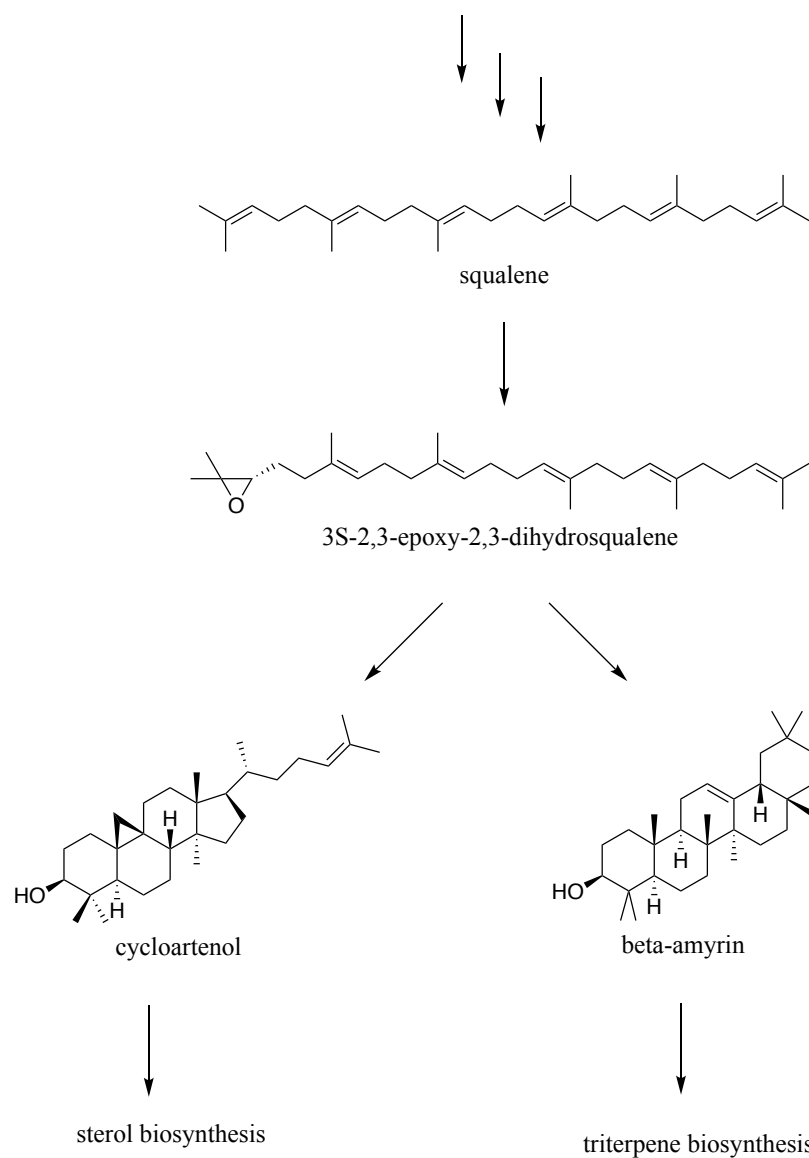


Figure 52: Sterol biosynthesis and triterpene biosynthesis share the squalene as a common intermediate in their biosynthetic pathways. The two pathways diverge with the cyclization of (3*S*)-2,3-epoxy-2,3-dihydrosqualene. Multiple arrows indicate multiple steps previously occurred in order to synthesize squalene.

#### 4.1.3 Biosynthesis of the triterpene avenacin

Triterpenes can be further subdivided based on the functional groups attached to the triterpene backbone. Triterpene saponins are glycosylated plant secondary metabolites, and they are the largest class of plant-derived natural products [203]. Triterpene saponins

are components of the cell membrane and are most well known for their antimicrobial properties [218]. They have uses in pharmaceuticals as foaming agents and in industry as agrochemicals and food additives [203, 219]. Lyonifolic acid A has shown to have antiviral activity against influenza A and pedunculoside has shown antifungal activity (Figure 53) [220, 221]. Triterpene saponins play important roles in plant defense and the fluidity of the cell membranes [218, 219].

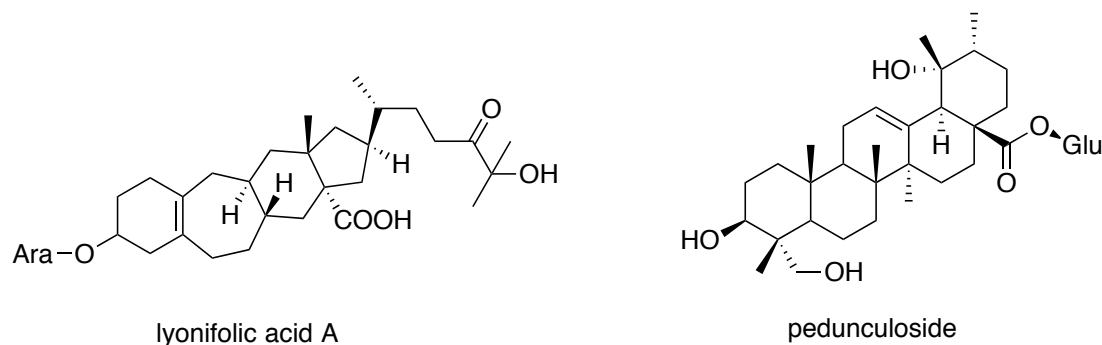
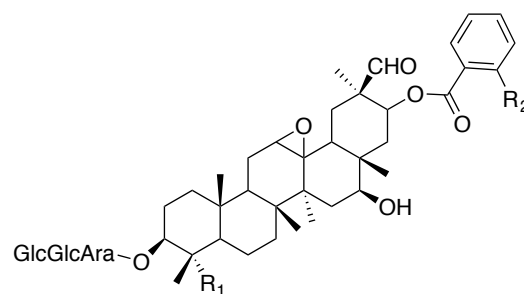


Figure 53: Two examples of triterpene saponins that have biological functions, lyonifolic acid A (antiviral) and pedunculoside (antifungal).

Cultivated oats (*Avena* sp.) are unique because they produce a family of defense-related triterpenes known as the avenacins; no other cereal is known to contain these compounds [219, 222-224]. Four structurally similar compounds, avenacin A-1, A-2, B-1, and B-2, are found in the roots (Figure 54) [225, 226]. Avenacin A-1 is the most abundant metabolite though the relative proportions of the avenacins varies during development [225]. Avenacin A-1 is synthesized and stored in the epidermal cells of the roots providing protection against the soil-borne fungal pathogen *Gaeumannomyces graminis* var *tritici*, a take-all disease of wheat [219]. All four of these triterpenes have a  $\beta$ -amyrin core and a branched sugar chain at C3 [225]. Nomenclature of the avenacins emphasizes structural variation at the C4 and C21 positions. “A” indicates a hydroxyl group at R<sub>1</sub>, whereas “B” denotes that R<sub>1</sub> is a hydrogen [226]. The “1” indicates a *N*-methylanthranilate group at C21 and “2” indicates a benzoate group [226].





Avenacin A-1:  $R_1=OH$   $R_2=NHCH_3$       Avenacin A-2:  $R_1=OH$   $R_2=H$   
 Avenacin B-1:  $R_1=H$   $R_2=NHCH_3$       Avenacin B-2:  $R_1=H$   $R_2=H$

Figure 54: Avenacin A-1, A-2, B-1, and B-2 are structurally similar avenacins. All four triterpenes have a decorated  $\beta$ -amyrin backbone and C3 is a branched sugar chain, D-glucose-D-glucose-D-arabinose [225]. The “A” is indicative of a hydroxyl group at  $R_1$  and “B” means  $R_1$  is a hydrogen [226]. The “1” indicates a *N*-methylantranilate group at C21 and “2” indicates a benzoate group [226].

Due to the highly fluorescent nature of avenacin A-1, saponin-deficient (SAD) mutants prepared by chemical mutagenesis could be identified by their lack of root fluorescence. This provided a ready entry to the study of the avenacin biosynthetic pathway. Currently, there are ten loci known for avenacin biosynthesis (SAD1-10) [219, 227-229]. After a more thorough genetic analysis, a gene cluster was revealed; SAD1, SAD2, SAD7, SAD9 and SAD10 loci are linked to the SAD1 locus [229, 230]. It has been shown that SAD1 mutants result from non-synonymous, premature termination, and splicing error mutations within the  $\beta$ -amyrin synthase gene (*AsbAS1*) causing an accumulation of 2,3-oxidosqualene and absence of  $\beta$ -amyrin [217, 231]. Other SAD mutants that have been identified through biochemical analysis have shown to accumulate putative intermediates. SAD2 mutants blocked hydroxylation, SAD7 and SAD10 mutants impaired acylation at the C21 position, SAD9 mutants inhibit *N*-methylation, and SAD3 and SAD4 mutants prevented glucosylation [229-231]. Recent work done by the Osbourn group has provided insight into the triterpene cyclization of SAD1 [232]. In collaboration with our group, we hope to provide further insight into the function of SAD2.

To date, triterpenoid saponins are synthesized from mevalonic acid (MVA) via the cytosolic isoprenoid pathway [217, 223, 230]. The chloroplastic methylerythritol

phosphate (MEP) gives rise to mono-, di- and tetraterpenes are synthesized by plants and cyanobacteria through a chemically independent route [223, 230]. The beginning of the avenacin pathway is shared with the sterol biosynthetic pathway. Subsequent to the production of 2,3-oxidosqualene, triterpenoid synthesis splits from the sterol pathway through enzymatically guided conformationally disparate transition states [223, 230]. In a sterol and cycloartenol cyclases, 2,3-oxidosqualene is cyclized via a chair-boat-chair transition state into a tetracyclic products in animals and plants, respectively. For avenacin, 2,3-oxidosqualene is cyclized by the triterpene cyclase enzyme  $\beta$ -amyirin synthase, SAD1, through a chair-chair-chair-boat conformation to form pentacyclic  $\beta$ -amyirin [217, 223]. The Osbourn group recently reported the characterization of the *saponin-deficient 2*, or SAD2, gene that is required for avenacin biosynthesis [233]. Plants with defective SAD2 alleles accumulated  $\beta$ -amyirin. Osbourn's group was also able to show that SAD2, also referred to as AsCYP51H10, catalyzed the formation of 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyirin [203]. After the formation of the 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyirin, details concerning the progression of the pathway can only be speculated (Figure 55).

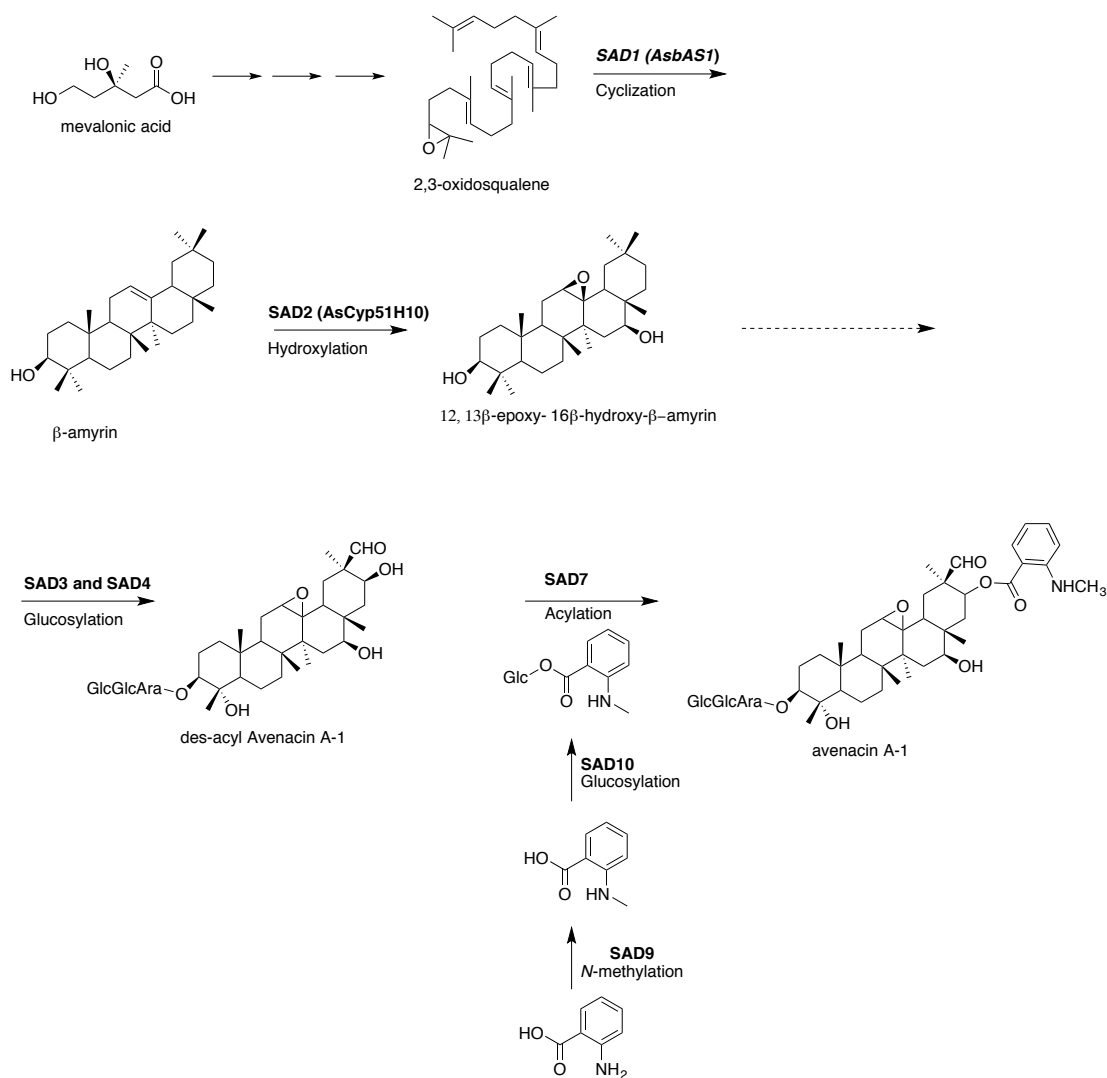


Figure 55: Proposed pathway for avenacin A-1 production. The dotted line indicates that several enzymatic transformations are anticipated. The sequence and chemistry from 12,13β-epoxy-16β-hydroxy-β-amyirin to des-acyl avenacin A-1 has not been confirmed [217]. Recent work has provide insight into SAD7, SAD9, and SAD10 for the production of avenacin A-1 [232].

It has been proposed that a series of cytochrome P450 (CYP) monooxygenases hydroxylate several intermediates of the pathway after 12,13β-epoxy-16β-hydroxy-β-amyirin [219]. Other steps that appear to occur include the formation of an *N*-methylanthranilate group through a coordinated system of methylating anthranilate (SAD9), which is subsequently glucosylated (SAD10) and then transferred to the

triterpene backbone by an acyltransferase (SAD7) [229]. Also, the linking of a branched trisaccharide is most likely effected by three different glycosyltransferases [219]. It has been proposed that a dehydrogenase, has been seen in sesquiterpene synthesis, may add the aldehyde by a pair of successive oxidations [234]. Due to the complexity and low abundance of the intermediates and the lack of availability of commercial standards for intermediates after  $\beta$ -amyrin has provided a challenge for saponin research [219]. Advances in chromatographic and spectroscopic techniques have expanded metabolic profiling of complex structures and thus have overcome a majority of these difficulties [219, 235, 236]. An alternative way to understand the biosynthetic pathway of avenacin A-1 is to study the intermediates produced by different enzymes involved in the pathway. As previously stated, it has been proposed that hydroxylations to avenacin A-1 occur via a series of CYP monooxygenases [217, 219]. For our project, a method to access  $\beta$ -amyrin derivatives could be beneficial in aiding to the understanding of the function of SAD2.

The cytochrome P450 (CYP) superfamily has been identified in all domains of life: animals, plants, fungi, bacteria, and viruses; it is the largest class of plant metabolic enzymes [203]. They catalyze a vast array of oxidation reactions, referred to by Coon et al. as a “diversosome”, most commonly inserting one oxygen atom into an aliphatic position in their substructures [203, 237]. There are thousands of CYP sequences known [208]. Due to large number of CYPs, and to help unify the enzymes across kingdoms a system of nomenclature has been developed [208]. For example, CYP51H10, CYP stands for cytochrome P450, 51 represents the family, H is the subfamily, and it was the 10<sup>th</sup> gene discovered in this subfamily. For genes to be categorized in the same family, the amino acid sequence identity must be at least 40% a cut off of 55% identity is applied for each subfamily [208]. During the modern practice of classifying a new sequence into an existing group, clustering takes priority. Use an agglomerative clustering algorithm rather than emphasizing the absolute % identity given [208]. Due the number of enzymes and that the names are assigned in historic order, the original capacity of the nomenclature system has been periodically extended and quite possibly will have to be further expanded. The CYP51 family is one of the most conserved enzyme families and is the only family in the CYP51 clan. The subfamily CYP51G catalyzes 14 $\alpha$ -demethylation

reaction for sterol biosynthesis. The discovery of CYP51H10 was unique because it was a CYP51 enzyme that mechanistically behaved differently than most members in this family.

The exceptional diversity among CYP enzymes is believed to have originated from a small number of ancestral CYP enzymes that duplicated and diverged to accept a broad range of substrates, but the cytochrome P450 14 $\alpha$ -sterol demethylase (CYP51) is unique because the structure of its active site has been maintained and the primary function of the enzyme has remained conserved across the living world [203, 238, 239]. CYP51 is one of the most ancient of the P450 families and can be found in all biological kingdoms including plants and fungi [186, 211]. The primary function of CYP51 resides in the critical conversion of lanosterol, eburicol, and obtusifoliol into membrane sterols such as cholesterol in humans, ergosterol in fungi, and sitosterol/brassinosteroids in plants, respectively [211]. The sterol substrate can vary but the sequence of oxidative conversions remains the same. Initially, the methyl at C14 is oxidized to an alcohol that is subsequently oxidized to the aldehyde. Finally, the aldehyde is oxidatively cleaved as formic acid with the concomitant of a double bond at C14-C15 (Figure 56) [211].

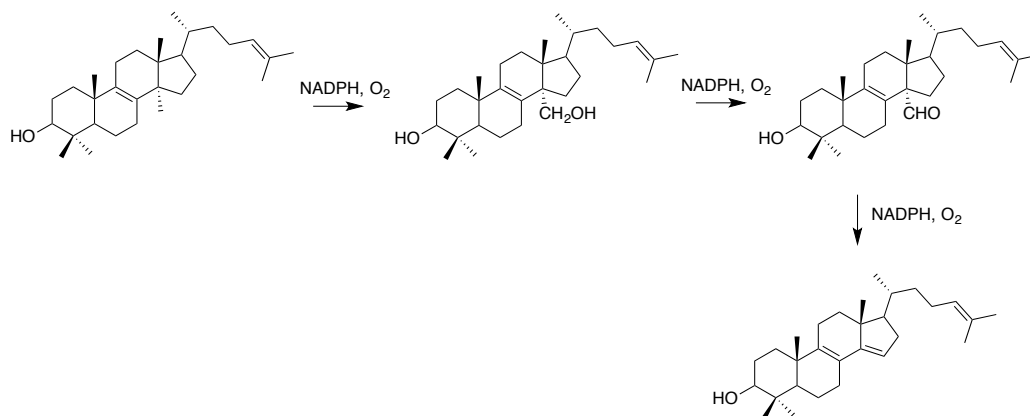


Figure 56: The prototypical CYP51 reaction, where lanosterol is converted to 4,4-dimethylcholesta-8(9),14,24-trien-3- $\beta$ -ol [211].

For plants, the obtusifoliol 14-demethylase falls into the clan CYP51G [233]. The Osbourn group cloned and characterized a new CYP enzyme belonging to the CYP51

sterol demethylase family, SAD2 [203, 233]. However, this enzyme was unable to be expressed in yeast using the standard system for CYP51G enzymes [233]. SAD2 had a 46% amino acid sequence identity with CYP51G, but only 32% identity in the active site residues [203, 233]. Phylogenetic analysis showed SAD2 belonged to a newly defined, but functionally uncharacterized subfamily, CYP51H, thus designated AsCYP51H10 [233]. When comparing the amino acid sequences of CYP51 enzymes from bacteria, fungi, protozoa, plants, and animals, 34 amino acids were fully conserved [233]. However, in AsCYP51H10, six of these residues were not conserved. Using 3D modeling, the structure of AsCYP51H10 and an oat 14 $\alpha$ -demethylase, AsCYP51G1 were compared to the crystal structure of MtCYP51B1 (*Mycobacterium tuberculosis*) [233]. A major difference in AsCYP51H10 compared to AsCYP51G1 was in the active site size. The volume of AsCYP51H10 was  $568 \pm 96 \text{ \AA}^3$  compared to MtCYP51B1 was  $343 \pm 62 \text{ \AA}^3$  and AsCYP51G1 was  $346 \pm 108 \text{ \AA}^3$  [233]. The ballooning of the active site correlates with the different functionality of AsCYP51H10 and thus makes it an inaugural member of the CYP51H subfamily [233]. Previous work done by Aoyama's group suggested that AsCYP51H10 expressed in yeast was able to oxygenate  $\beta$ -amyrin at two positions, one of which was the epoxidation of the 12-alkene [240]. Though the full structure was not determined, the MS spectrum from Osbourn's group and yeast expression were in agreement [203, 240]. AsCYP51H10 was coexpressed in *Nicotiana* plants with AsbAS1 to investigate its biochemical function [203]. The complete identification of the product from the coexpression of AsbAS1 and CYP51H10 was determined by NMR analysis to yield 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin [203].

The compound  $\beta$ -amyrin does not seem to possess antifungal properties [227]. Therefore, the other functional groups of avenacin A-1 must play an important role in its antimicrobial activity. The function groups instilled into the product from SAD2 both appear in avenacin A-1. Currently, there are no standards available of avenacin A-1 lacking the epoxy or 16-hydroxyl function group [203]. However, the Osbourn group was able to generate and purify 12-oxo-avenacin A-1 (OA) [203]. Disk assays were used to study the antifungal activity of the epoxide moiety [203]. *Gaeumannomyces graminis* var. *tritici* is a fungus known to be sensitive to avenacin (45). As the concentration increased, avenacin A-1 inhibited growth of the fungus, but 12-oxo-avenacin did not (Figure 57).

This indicates that the C12-C13 epoxide is critical for the antifungal activity, and thus the enzyme AsCYP51H10 plays an important role in the antimicrobial properties of avenacin A-1.

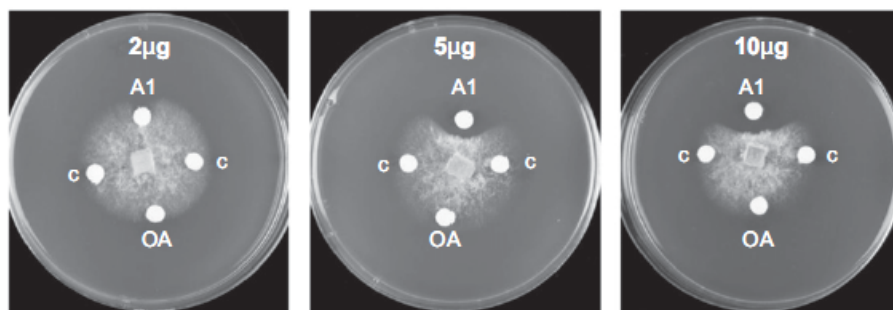


Figure 57: Disk assays testing the antifungal activity of avenacin A-1 (A1) and 12-oxo-avenacin A-1 (OA) against *Gaeumannomyces graminis* var. *tritici*, where the loading of the applied triterpene increases from left to right. Control discs were loaded with equal volumes of 75% methanol. Reproduced with permission from [203].

It was recognized from the identified product of Sad2 oxidation that the pathway to the epoxidized triterpene diol likely proceeds through two different intermediates, 12,13 $\beta$ -epoxy- $\beta$ -amryin or 16 $\beta$ -hydroxy- $\beta$ -amyryin (Figure 58A) [203]. Other microbial multifunctional P450s have been shown to both epoxidize and hydroxylate substrates. There are examples of both possible reaction mechanisms. GfsF found in *Streptomyces* synthesizes the antibiotic FD-891. It first catalyzes the epoxidation of 25-O-methyl FD-892 and then hydroxylates (Figure 58B), whereas MycG from *Micromonospora griseorubida* hydroxylates mycinamycin IV (M-IV) before epoxidation of M-V (Figure 58B) [203, 241, 242].

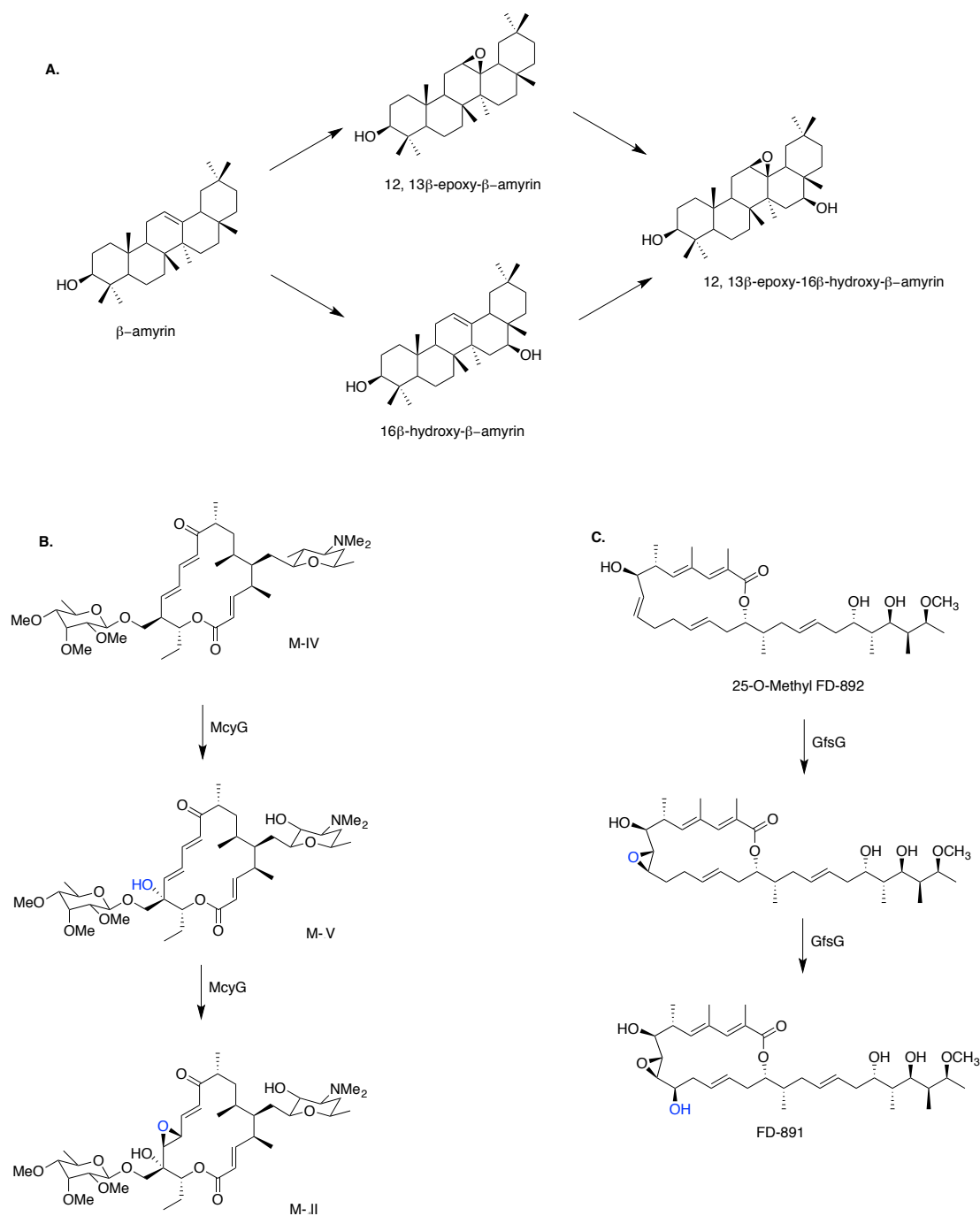


Figure 58: Multifunctional CYP enzymes that perform epoxidation reactions. *A.* Two hypothetical pathways for the formation of 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin from  $\beta$ -amyrin. It is not known whether *Sad2* oxidizes intermediates 12,13 $\beta$ -epoxy- $\beta$ -amyrin or 16 $\beta$ -hydroxy- $\beta$ -amyrin. Both pathways could be active. *B.* *MycG* is a multifunctional CYP enzyme that hydroxylates before it epoxidizes the dienone *C.* *GfsF* is an example of a multifunctional CYP enzyme that epoxidizes a *trans*-alkene before hydroxylation [203].



Geisler *et. al.* used a sequence alignment from other CYP51 enzymes and molecular modeling experiments to predict the order of the reactions carried out by AsCYP51H10 [203]. For molecular modeling the crystal structure of human CYP51 complexed with the fungal CYP inhibitor ketoconazole was used as a template. Representative members of the CYP51 subfamilies were used for sequence alignment substrate recognition sites (SRSs) were highlighted as shown in Figure 59A [203]. Of the five-substrate recognition sites, SRS5 looks to be key determinant for the size of the active site. At residues 353 and 354 all the sterol demethylase subfamilies have either proline or leucine residue, whereas AsCYP51H10 contains an alanine [203]. In addition, previous works has shown that the regiospecificity and selectivity can be altered by substitutions in SRS5 in a range of families including CYP51 [212, 243-246].

Molecular modeling was used to gain insight into the shape and binding orientation of the triterpene at the active site. Previous work investigated the molecular modeling of a CYP to predict the insertion sites discovering the average distance of the oxidation of the carbon atoms ad the heme iron was 4.14 Å and all were within 6 Å [247]. Again, the crystal structure of human CYP51 complexed with the fungal CYP inhibitor ketoconazole was used as a template. The precatalytic binding complexes were identified by minimizing the binding energy of beta-amyirin or its epoxidized or 16β-hydroxylated form and examining those enzyme-substrate complexes that had an average distance of the heme iron distal oxygen ligand (atom O1) to the hydroxylation site (C16) or to the epoxidation sites (C12, C13) that was 5 Å or less [203, 247]. This resulted in identifying precatalytic binding modes for 16β-hydroxy-β-amyirin and β-amyirin, but not for 12,13β-epoxy-β-amyirin suggesting that 16β-hydroxy-β-amyirin is formed first [203]. The SRSs that appear to have the most significant interaction in the precatalytic complex with β-amyirin include SRS1, SRS2, SRS4, SRS5, and SRS6, and those that might participate in the binding of 16β-hydroxy-β-amyirin include SRS1, SRS4, and SRS5 (Figure 59B/C) [203]. This further suggests that AsCYP51H could act as a multifunctional enzyme that is capable of two different oxidative reactions [203].

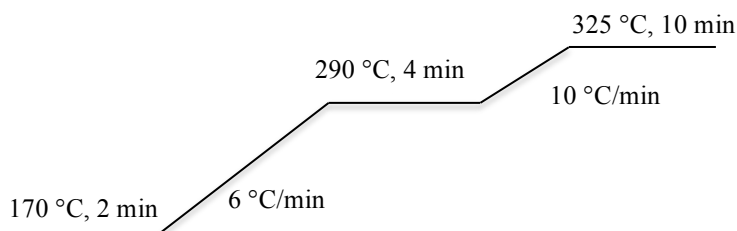


were used as internal chemical shift references. Temperature was adjusted to 298 K for all spectra collected. Pulse programs were adjusted as follows: DEPT NS = 256; COSY TD = 512, SW = 6680 Hz; HSQC NS = 8, TD = 512, SW = F1 = 20.7 MHz F2 = 2940 Hz; HMBC NS = 16, TD = 512, SW = F1 = 27.9 MHz F2 = 6493 Hz, *J* filter = 10 Hz; NOESY TD = 512, SW = 4995 Hz, TIME = 19.40. Peak multiplicities and shapes were annotated as follows: s, singlet; d, doublet; t, triplet; q, quartet; qn, quintet; m, multiplet; br, broad. Initially carbons were numbered in an increasing fashion from downfield to upfield chemical shifts. All correlations were originally made using this numbering system. Once the product was fully assigned, the carbon-shift based numbers were re-assigned IUPAC numbers to compare chemical shifts between identified products and other triterpenes. The labeling of diastereotopic hydrogens denotes the most downfield number as “a” and upfield as “b.” Infrared spectra were recorded in KBr pellets with a Fourier-transform infrared spectrophotometer (Nicolet Avatar 33). Mass spectra were collected with an Agilent 6520 Q-TOF mass spectrometer or GC/MS.

#### 4.2.1 GC/MS and Q-TOF protocols for triterpene analysis

GC/MS analyses were performed using an Agilent 7890A GC/5975C MS. Samples were dissolved in ACN, silylated with BSTFA + 2% TMSCl, and the triterpenes were separated with a HP-5 (30 m x 250  $\mu$ m x 0.25  $\mu$ m) column.

##### GC 1: Triterpene



GC/MS parameters were as follows:

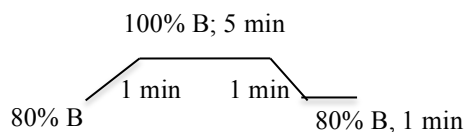
- Split-Splitless Inlet: Splitless; Inlet temperature: 280 °C; Pressure: 17.222 psi; Septum purge flow: 3 mL/min
- He flow rate: 15 mL/min for 2 min

The scan parameters were as follows:

- Solvent delay: 6 min
- Scan speed: Normal
- Start mass: 58.00 amu; Stop mass: 800.00 amu
- Threshold: 150; Scans/sec: 1.99

Triterpene high-resolution mass spectrum was obtained for each compound using Q-ToF LC/MS analysis. Q-ToF analyses were performed using an Agilent 6520 Quadruple Time-of-Flight MS.

#### Q-ToF LC/MS: Triterpene



The source parameters were set as follows:

- MS2 full scan method
- ESI positive mode
- Fragmentation: 135 V
- Gas temp: 315 °C
- Gas flow: 10 L/min
- Nebulizer: 50 psi

- Capillary: +/- 3500 V

For the product from mutants I471M-2 and I471M-3, and the triterpene standard  $\beta$ -amyryn, the parameters for the chromatograph were:

- Solvent A: H<sub>2</sub>O + 0.1% formic acid. Solvent B: MeOH + 0.1% formic acid
- Flow rate: 1 mL/min; 80% to 100% B over 1 min with a 5-min hold and then ramped back to 80% B over 1 min
- Run time: 6 min
- Post time: 1 min

For the product from wild-type CYP51H10 and mutant CYPs A345L and I471M-1, the parameters for the chromatograph were:

- Solvent A: H<sub>2</sub>O + 0.1% NH<sub>4</sub>OAc. Solvent B: MeOH + 0.1% NH<sub>4</sub>OAc
- Flow rate: 1 mL/min; 50% to 100% B over 1 min with a 5-min hold and then ramped back to 50% B over 1 min
- Run time: 6 min
- Post time: 1 min

#### 4.2.2 Triterpene isolation from SAD2 mutant expression in *N. benthamiana*

James Reed and Michael Stephenson at the John Innes Centre isolated the triterpene products using the procedures that follow. We collected IR, GC/MS and HR-MS after receiving the samples to ensure that they were not degraded during transport prior to the NMR analysis.

**CYP51H10 wild-type product:** For the isolation of the major product of wild-type Sad2, about 70 *N. benthamiana* plants were vacuum infiltrated with constructs that led to the coexpression of three enzymes, tHMGR, AsbAS1, and CYP51H10. Leaves were harvested from the plants after 5 days and lyophilized. Using a mortar and pestle, the leaf material was pulverized in liquid N<sub>2</sub> and then a mixture of ethanol/H<sub>2</sub>O/KOH 9:1:1 (v:v:w) was added (6.5 mL/g of dry tissue) to disrupt the tissue and release acyl lipids as

carboxylate salts. The leaves were saponified for 2 hours at 65 °C. The ethanol was removed by rotary evaporation and the remaining mixture was extracted several times with ethyl acetate. The combined organic fractions were pooled together and concentrated to an oil by rotary evaporation. The crude product was suspended in toluene and loaded on a SNAP KP-Sil 25-g silica gel column (Biotage). Using an Isolera One flash chromatograph and a step gradient of hexane:ethyl acetate ranging from 6% ethyl acetate to 100% ethyl acetate, 9-mL fractions were collected. After the analysis by TLC and GC-MS, fractions containing product **88** were combined and the solvent was removed. A second purification was performed using a SNAP Ultra 10-g column (Biotage) again with a gradient of hexane:ethyl acetate from 0 to 50% ethyl acetate. After TLC and GC-MS analysis, the 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy-containing fractions deemed pure were concentrated to yield a yellow solid. The solid was further purified by decolorization with charcoal; and recrystallized from an ethanol:water mixture to give the final product. IR (KBr pellet): 3493, 2996, 2985, 2957, 2861, 1458, 1389, 1368, 1357, 1106, 1061, 1052, 1035, 1024, 1001  $\text{cm}^{-1}$ ; HRMS (ESI)  $m/z$  calc for  $[\text{M} + \text{H}]^+$   $\text{C}_{30}\text{H}_{51}\text{O}_3$  459.3834, found 459.3833.

**CYP51H10-A354L product:** Approximately 45 *N. benthamiana* plants were vacuum infiltrated with constructs that led to the coexpression of three enzymes, tHMGR, AsbAS1, and CYP51H10-A354L. Leaves were harvested from the plants after 5 days and lyophilized. Using a mortar and pestle, the leaf material was pulverized in liquid  $\text{N}_2$  and then a mixture of ethanol/ $\text{H}_2\text{O}$ /KOH 9:1:1 (v:v:w) was added (6.5 mL/g of dry tissue) to disrupt the tissue and hydrolyze acyl lipids to their carboxylate salt constituents. The leaves were saponified for 2 hours at 65 °C. The ethanol was removed by rotary evaporation and the remaining mixture was extracted several times with ethyl acetate. The combined organic fractions were pooled together and concentrated to an oil by rotary evaporation. The crude product was suspended in toluene and loaded on a SNAP Ultra 50-g silica gel column (Biotage). A step-wise gradient of hexane:ethyl acetate ranging from 6% ethyl acetate to 100% ethyl acetate was used; 9-mL fractions were collected. After the analysis by TLC and GC-MS, fractions containing **M2** product were combined and solvent was removed to yield an orange solid. The solid was dissolved in toluene and

a second purification was performed using a SNAP Ultra 10-g column (Biotage) with a gradient of hexane:ethyl acetate from 7% to 70% ethyl acetate. After TLC and GC-MS analysis, the **M2**-containing fractions deemed pure were concentrated. The orange material was dissolved in methanol purified further by reverse-phase semi-preparative HPLC was used with a Phenomenex Luma column (C<sub>18</sub>, 5 μm, 100 Å, 250x10 mm). A methanol:H<sub>2</sub>O gradient was run from 85% methanol for 3 minutes to 98% over 20 minutes at 3 mL/min and held for 9 additional minutes. Dionex charged-aerosol detector monitored fractions and **M2** were collected at around 24 minutes. IR (KBr pellet): 3340, 2950, 2876, 1458, 1385, 1360, 1060, 1030, 998, 907, 737 cm<sup>-1</sup>; HRMS (ESI) *m/z* calc for [M + H]<sup>+</sup> C<sub>30</sub>H<sub>51</sub>O<sub>2</sub> 443.3889, found 443.3850.

**CYP51H10-I471M product 1:** For I471M mutant, 144 *N. benthamiana* plants were vacuum infiltrated with constructs that led to the coexpression of three enzymes, tHMGR, AsbAS1, and CYP51H10-I471M. Leaves were harvested from the plants after 5 days and lyophilized. Crude material was extracted using Büchi SpeedExtractor in ethanol at 100 °C and 150 bar. The ethanol was dried down by rotary evaporation to give a green material for saponification. A mixture of ethanol/H<sub>2</sub>O/KOH 9:1:1 (v:v:w) was added (6.5 mL/g of dry tissue) to disrupt the tissue and hydrolyze acyl lipids to their carboxylate salts. The leaves were saponified for 2 hours at 65 °C. The ethanol was removed by rotary evaporation. The remaining mixture was dissolved in ethyl acetate and washed several times with H<sub>2</sub>O. The ethyl acetate was removed to yield an orange-brown solid. The material was dissolved in dichloromethane and absorbed onto silica gel 60 (Material Harvest UK), which was loaded onto a KP Sil 25-g column (Biotage). Flash column chromatography was completed using a step gradient of hexane:ethyl acetate from 0% to 100% ethyl acetate collecting 9-mL fractions. After the analysis by TLC and GC-MS, fractions 40-50 were combined and the solvent was removed. The residue was dissolved in dichloromethane and loaded onto a 10-g SNAP Ultra column (Biotage) prior to eluting with a step gradient of hexane:ethyl acetate ranging from 0% to 100% ethyl acetate was used. After GC-MS analysis, the **90**-containing fractions of highest purity were concentrated to yield a red solid. This material was further purified by decolorization with charcoal to yield a yellow solid before trituration in methanol gave the final product

as a white solid. IR (KBr pellet): 3452, 2933, 2863, 1465, 1375, 1101, 1065, 974, 719  $\text{cm}^{-1}$ ; HRMS (ESI)  $m/z$  calc for  $[\text{M} + \text{H}]^+$   $\text{C}_{30}\text{H}_{51}\text{O}_2$  443.3889, found 443.3797.

**CYP51H10-I471M product 2:** *N. benthamiana* plants were vacuum infiltrated with constructs that led to the coexpression of three enzymes, tHMGR, AsbAS1, and CYP51H10-I471M. Leaves were harvested from the plants after 5 days and lyophilized. Crude material was extracted using Büchi SpeedExtractor in ethanol at 100 °C and 150 bar. The ethanol was dried down by rotary evaporation to give a green material. A mixture of ethanol/ $\text{H}_2\text{O}$ /KOH 9:1:1 (v:v:w) was added (6.5 mL/g of dry tissue) to disrupt the tissue and hydrolyze acyl lipids to carboxylate salts. The leaves were saponified for 2 hours at 65 °C. The ethanol was removed by rotary evaporation. The remaining mixture was dissolved in ethyl acetate and washed several times with  $\text{H}_2\text{O}$ . The ethyl acetate was removed to yield an orange-brown solid. The material was dissolved in dichloromethane and absorbed onto silica gel 60 (Material Harvest UK), which was loaded onto a KP Sil 25-g column (Biotage). Flash column chromatography was completed using a step gradient of hexane:ethyl acetate from 0% to 100% ethyl acetate collecting 9-mL fractions. After the analysis by TLC and GC-MS, fractions 51-60 were combined and the solvent was removed. The residue was dissolved in dichloromethane and loaded onto a 10-g SNAP Ultra column (Biotage) prior to eluting with a step gradient of hexane:ethyl acetate ranging from 0% to 5% before ramping to 100% ethyl acetate. Fractions containing the product were purified again by SNAP 10-g Ultra column (Biotage) using a step gradient of hexane:ethyl acetate from 0% to 30% and 30% to 100% ethyl acetate. Fractions containing **91** combine and absorbed onto silica gel 60 (Material Harvest UK) and then loaded onto a SNAP 10-g Ultrage (Biotage). Flash column chromatography was performed using hexane:ethyl acetate ranging from 5% to 25% before ramping to 100% ethyl acetate. Fractions of higher purity were combined, concentrated, and decolorized with charcoal to yield a yellow solid. The yellow solid was triturated using ethanol to give the final product as a crystalline solid. IR (KBr pellet): 3453, 2967, 2945, 2852, 1698, 1468, 1385, 1368, 1083, 1045, 1035, 992  $\text{cm}^{-1}$ ; HRMS (ESI)  $m/z$  calc for  $[\text{M} + \text{H}]^+$   $\text{C}_{30}\text{H}_{51}\text{O}_2$  443.3889, found 443.3899.



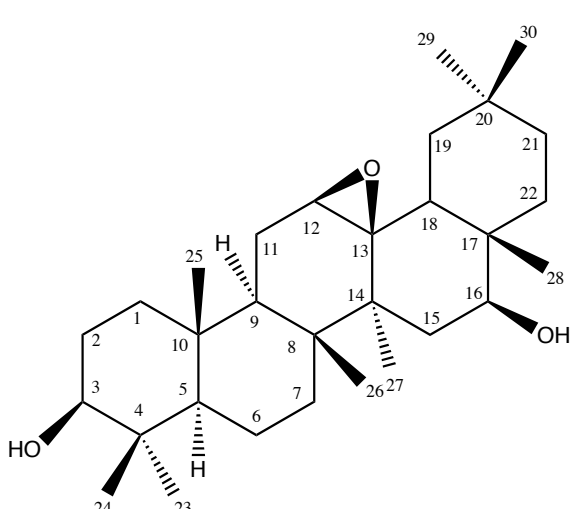
**CYP51H10-I471M product 3:** *N. benthamiana* plants were vacuum infiltrated with constructs that led to the coexpression of three enzymes, tHMGR, AsbAS1, and CYP51H10-I471M. Leaves were harvested from the plants after 5 days and lyophilized. The leaf material was pulverized in liquid N<sub>2</sub> using a mortar and pestle, and then a mixture of ethanol/H<sub>2</sub>O/KOH 9:1:1 (v:v:w) was added (6.5 mL/g of dry tissue) to disrupt the tissue and saponify acyl lipids to their constituent carboxylate salts. The leaves were saponified for 3 hours at 65 °C. The mixture was filtered through an 8 cm x 10 cm column of diatomaceous earth (celite – Sigma) and the solvent was removed. The leaf material was reextracted with ethyl acetate (250 mL) overnight. The ethanol and ethyl acetate fractions were combined and dried. The crude material was dissolved in H<sub>2</sub>O (500 mL) and extracted with hexane (4 x 500 mL). The hexane fractions were combined, washed with brine (500 mL), and the solvent was then removed by rotary evaporation to give an orange solid. The crude solid was loaded on a SNAP Ultra 50-g column (Biotage). The material was purified using flash column chromatography with 15% ethyl acetate:hexane. After analysis by GC-MS, the fractions containing **93** were combined and the solvent was removed. The solid was further purified by dissolving the triterpene in dichloromethane and decolorizing with charcoal; the final sample was recrystallized from methanol to give the final product. IR (KBr pellet): 3312, 2928, 2850, 1471, 1389, 1368, 1047, 1033, 1000, 964 cm<sup>-1</sup>; HRMS (ESI) *m/z* calc for [M + H]<sup>+</sup> C<sub>30</sub>H<sub>51</sub>O<sub>2</sub> 443.3889, found 443.3907.

## 4.3 Results and Discussion

### 4.3.1 Mutants

Previous work showed that mutants of AsCYP51H10 were defective in avenacin synthesis but accumulated β-amyryn [227, 233]. AsCYP51H10 was coexpressed in *Nicotiana* plants with AsbAS1 and the product was identified by NMR analysis in CDCl<sub>3</sub> to be 12,13β-epoxy-16β-hydroxy-β-amyryn (Figure 61) [203]. When β-amyryn is oxidized, key chemical shift changes in the <sup>13</sup>C spectra are observed at C12 and C13, where resonances at δ 121.7 and 145.2 ppm are replaced by peaks at δ 54.1 ppm and 66.9 ppm indicative of an epoxy moiety. Also, in the <sup>1</sup>H spectrum, a new signal is observed

downfield at  $\delta$  4.16 ppm reflecting the presence of a new carbon-oxygen bond. Through heteronuclear multiple bond correlation and NOE experiments, the new hydroxyl group was determined to be on the  $\beta$ -face at C16, thus defining the intermediate as 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin (**88**). Both modifications are maintained in the final product of avenacin A-1 [203].



IUPAC #	$\delta^{13}\text{C}$	Carbon Type	$\delta^1\text{H}$	(m), $J_{HH}$ (Hz)
1	38.03	CH <sub>2</sub>	1.52, 0.88	
2	27.12	CH <sub>2</sub>	1.59, 1.54	
3	78.89	CH	3.17	dd, $J = 11.2, 4.4$ Hz
4	38.81	C	-	
5	55.42	CH	0.66	dd, $J = 11.2, 4.4$ Hz
6	17.76	CH <sub>2</sub>	1.51	
7	34.17	CH <sub>2</sub>	1.39, 1.19	
8	41.38	C	-	
9	45.72	CH	1.33	
10	37.19	C	-	
11	21.59	CH <sub>2</sub>	1.86, 1.70	
12	54.13	CH	2.76	d, $J = 4.6$ Hz
13	66.90	C	-	
14	42.35	C	-	
15	34.17	CH <sub>2</sub>	1.82, 1.44	
16	67.15	CH	4.16	dd, $J = 11.1, 4.8$ Hz
17	38.47	C	-	
18	49.55	CH	1.16	
19	41.25	CH <sub>2</sub>	1.69, 1.14	
20	30.73	C	-	
21	34.17	CH <sub>2</sub>	1.34, 1.15	
22	30.70	CH <sub>2</sub>	1.81, 1.24	
23	28.00	CH <sub>3</sub>	0.77	s
24	15.35	CH <sub>3</sub>	0.97	s
25	15.04	CH <sub>3</sub>	0.86	s
26	20.36	CH <sub>3</sub>	1.11	s
27	23.40	CH <sub>3</sub>	1.22	s
28	21.34	CH <sub>3</sub>	1.03	s
29	33.17	CH <sub>3</sub>	0.81	s
30	23.69	CH <sub>3</sub>	0.91	s

Figure 61: The structure of 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin collected in CDCl<sub>3</sub> with IUPAC numbering. Tabulated are <sup>1</sup>H and <sup>13</sup>C chemical shifts, carbon type, proton coupling constants and multiplicity. The key chemical shift changes versus  $\beta$ -amyrin are indicated in red boxes [203].

It is known that AsCYP51H10 synthesizes 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin, but the order of the oxidation events in which the enzyme functions is unknown. Mutants of SAD2 have been studied by computational modeling help understand the function of AsCYP51H10 [203]. The Osbourn group identified the location of seven SAD2 mutants. Of those seven, four have single mutations on residues that are critical to the heme binding or play roles in maintaining the three-dimensional structure of the CYP enzyme,

and therefore they were discarded [203]. The three remaining, A124V, P463S, T113I, were tested individually by computational modeling in hope one of these mutants would accumulate the 12,13 $\beta$ -epoxy- or 16 $\beta$ -hydroxy intermediates [203]. Unfortunately not one of the mutants were predicted from computational modeling to be able to form pre-catalytic complexes [203]. Consistent with this, the seven mutants accumulated  $\beta$ -amyrin, but did not accumulate either potential intermediate [203, 233]. New mutants needed to be made in order to try and isolate 16 $\beta$ -hydroxy- $\beta$ -amyrin and/or 12,13 $\beta$ -epoxy- $\beta$ -amyrin.

As previously stated, at positions A353 and A354, the presence of small residues appear to be relevant for the alcohol binding. Therefore a series of mutants were made for the residues at 353 and 354. The other amino acids found in CYP51 enzymes at 353 and 354 were threonine (T), leucine (L), valine (V), isoleucine (I), alanine (A), proline (P), and serine (S). Seven different amino acids at two different residues resulted in a possible 49 mutants. Due to the strategy used to produce the variant proteins, only 28 of the 49 mutants were analyzed (Table 7).

Table 7: The 28 mutants that were analyzed.

<b>Indels</b>	<b>Amino Acids at 353 and 354</b>
ACTCTT	TL
ACTGTT	TV
ATTACT	IT
ATTATT	II
ATTGCT	IA
CCTATT	PI
CCTCCT	PP
CCTCTT	PL
CCTGCT	PA
CCTTCA	PS
CTTCCT	LP
CTTCTT	LL
CTTGCT	LA
CTTGTT	LV
CTTCA	LS
GCAGCG	AA
GCTGCT	AA
GCTCTT	AL
GCTGTT	AV
GCTTCA	AS
GTTATT	VI
G TTCCT	VP
GTTCTT	VL
GTTGCT	VA
GTTTCA	VS
TCACCT	SP
TCACTT	SL
TCAGTT	SV

The mutants were infiltrated in *Nicotiana benthamiana* with SAD1 and the products were extracted and analyzed by GC/MS. The series of mutants of sad2 from

Table 7 were observed to make a range of monooxidized  $\beta$ -amryin products (initially designated M1-4), several dioxidized  $\beta$ -amryin products (D1-3), a 14-demethylation product, and a minor compound with  $m/z$  440. Through agglomerative hierarchical clustering and principal component analysis, the mutants that make these products were divided into five different groups. Group 1 was best described as wild-type group producing D1 as the main triterpene. Group 2 was comprised of inactive mutants. Group 3 predominately made M2. Group 4 mutants made a mixture of products that included M1, M3, the demethylation product, and  $m/z$  440. Finally group 5, produced a variety of oxidation products in addition to the WT products was M4, D2, and D3.

The mutants that were most interesting were I471M and A353X-A354X' mutants X' = I, L, or V. I471M appears to produce a distinctive group of monooxygenated products and, through the apparent product of a demthylation triterpene, shifts the enzyme substantially back to the CYP51G. Mutant I471M was present in group 4 producing M1, M3, the demethylation product, and  $m/z$  440. Mass vacuum infiltration of *N. benthamiana* leaves was used to scale-up this screening hit, where products of the coexpression were separated, purified, and identified. A354X' with either I or L made exclusively the M2 product and very little wild type products. A354V mutants are intermediates between WT and A3354L/I mutants. With the goal of requiring different binding of  $\beta$ -amryin in order for the enzymatic reactivity to produce the epoxidation product, mutant A353A-A354L was used in large-scale vacuum infiltrations of *Nicotiana benthamiana* leaves and the product, M2 was isolated. The rigorous elucidation of the wild-type, A354L, and I471M products was the prime goal of our effort. Additionally, mutants of *sad2* could provide alternative oxidation patterns for making avenacin derivatives. Semi-synthetic derivatives of triterpene saponins and their biosynthetic intermediates are currently being explored for possible applications in the pharmaceutical, cosmetic, and/or food industries [248-251].

#### 4.3.2 Structure elucidation of the products from SAD2 mutants

$^1\text{H}$ ,  $^{13}\text{C}$ , DEPT, COSY, HSQC, HMBC, and NOESY NMR spectra were recorded. The following structure elucidation is described using IUPAC numbering. All of the  $^1J_{\text{CH}}$  correlations that assisted in assigning the structure were assigned from the HSQC spectrum. For all products isolated, each of the eight methyl signals was a singlet consistent with the  $\beta$ -amyirin skeleton. For the HMBC, the majority of the assignable cross-peaks were attributable to carbons that coupled to hydrogens on the eight well-resolved methyl groups. As a result, most of the connectivity of the structure could be characterized from these correlations.

#### 4.3.3 Structure elucidation of the product from wild-type SAD2

The structure of the epoxydiol from wild-type Sad2 had been identified earlier by the Osbourn lab, partially through NMR data collected in deuteriochloroform [203]. After many unsuccessful efforts during the current project were made to purify the oxygenated triterpenes, it was found that the extreme acid sensitivity of the epoxidized triterpenes required that all spectra were collected in pyridine- $d_5$ . Employing a reliable consistent NMR solvent allowed for the comparison of the chemical shifts from the product of wild-type Sad2 with the products produced by the mutants.

The structural elucidation of compounds from the mutant SAD2 enzymes relied, in part, on the detailed evaluation of a data set believed to be attributable to the 12,13-epoxy-16- $\beta$ -hydroxy- $\beta$ -amyirin, initially dubbed compound **88**.

After isolation, compound **88** was derivatized with TMS and analyzed by GC-MS; under these conditions, it had a distinctive ion with to give an  $m/z$  of 602. This indicated the presence of two derivatizable hydroxyl groups and suggested the further addition of a non-hydroxylic oxygen to  $\beta$ -amyirin, an observation that was supported by  $^1\text{H}$  NMR,  $^{13}\text{C}$  NMR, and IR data. The full set of NMR experiments used to determine the structure of **88** included  $^1\text{H}$ ,  $^{13}\text{C}$ , DEPT-135, DEPT-90, HSQC, HMBC, COSY, and NOESY; all spectra were collected in pyridine- $d_5$  solvent. Based on the overall similarity of mass spectrum of **88** with known derivatives of  $\beta$ -amyirin, compound **88** was expected to have a  $\beta$ -amyirin skeleton. According to data from the  $^{13}\text{C}$  spectrum, DEPT-135, and DEPT-90

experiments, there were thirty unique carbons including eight methyls, nine methylenes, six methines and seven quaternary carbons. The double bond of  $\beta$ -amyrin was conspicuously no longer present. Based on  $^{13}\text{C}$  peaks observed at 67.4 ppm and 54.5 ppm and HSQC data correlating the more upfield  $^{13}\text{C}$  to a proton at 2.89 ppm, the alkene appeared to have been oxidized to an epoxide. Tabulated spectral data is found in Table 8.

The detailed assignment of the carbon-13 and proton NMR data is as follows. The secondary carbons attached to the hydroxyl groups provided the most easily identifiable entry point. They are both attached to the most electronegative atoms in the structure leading to their downfield resonance in the range of 60-85 ppm in the  $^{13}\text{C}$  NMR and will have a scalar coupling to an exchangeable hydrogen that is downfield in the  $^1\text{H}$  NMR spectrum. Two carbons met this criteria; the carbon at 78.3 ppm (C3) correlating to the hydrogen at 3.44 ppm and the carbon at 66.0 ppm (C16) correlating to the hydrogen at 4.57 ppm. Only two pairs of methyl groups were observed to correlate to each other in the HMBC spectrum. There are two geminal dimethyl groups on the  $\beta$ -amyrin framework; one pair was anticipated to be C23 and C23 and the other set was C29 and C30. C3 had cross peaks to the geminal dimethyl groups H23 and H24 revealing a segment of ring A (Figure 62A); collectively, this defined several strategic starting points for the remaining assignments. Continuing on ring A, the quaternary carbon C4 and methine carbon C5 showed coupling to H23 and H24. C5 correlated to the methyl group H25 at the ring fusion; C10 and C1 were also located with HMBC correlations to Me25. Finally, the COSY showed a cross peak between protons H3 and H2, completing ring A.

Proceeding from the bridgehead atoms at the A/B fusion, C5 and C10, an HMBC crosspeak was observed that linked C5 to H6 (Figure 62B). C9 additionally correlated back to H25, the methyl group attached to C10. Also, C9 showed long-range coupling to H26, tentatively assigning C26 as the methyl group at the bridgehead from ring B to ring C. There were two nearly chemically equivalent signals at 35.0 ppm in the  $^{13}\text{C}$  spectrum that displayed a correlation to H26; this assigned the chemical shift of C7 (the overlapping carbon resonance is assigned to C21 *vide infra*). Finally the substructure containing quaternary carbons C8 and C14 was identifiable from C4 cross-peaks to H26 as well as to H27. C8 and C14 could be distinguished because the HMBC showed weak coupling from C8 to H7.

Going from ring B to ring C, the quaternary carbon C13 at 67.4 ppm correlated to H27 allowing it to be the other bridgehead carbon between ring C and ring D (Figure 62C). C13 was most likely attached to the non-hydroxylic oxygen based on its chemical shift. This was confirmed by C12 at 54.2 ppm in  $^{13}\text{C}$  spectrum correlating to a proton at 2.89 ppm in the HSQC indicative of the presence of an epoxide. The proton resonance is a doublet with a small 4 Hz coupling constant that supports the assignment of an epoxide. The doublet represents H12 $_{\alpha}$  coupling to the H11 $_{\alpha}$ , the coupling constant to H11 $_{\beta}$  is expected to be zero. Finally, the COSY had a cross peak from H12 to H11 and thus completed ring C.

C15 was correlated by the HMBC spectrum to the methyl at the ring fusion of C/D (H27) (Figure 62D). For the product from wild-type SAD2 the second hydroxyl group was found on ring D. The oxymethine hydrogen at 4.57 ppm showed a  $^1J$  correlation to the carbon at 66.0 ppm (C16) in the HSQC spectrum. C16 has an HMBC cross-peak to H28 and could be extended via the COSY data to both H15a and H15b. The remaining correlations from methyl H<sub>3</sub>28 to carbon C17 and methine C18 were located, thus completing ring D.

Starting at C18 between ring D and ring E, the COSY showed a crosspeak from H18 to H19a (Figure 62E). The HMBC connected C19 to the geminal dimethyl groups, H29 and H30. The quaternary carbon, C20, and two methylenes, C22 and C21, all correlated to H29 and H30. The methyl groups were distinguishable because C22 possessed a defining crosspeak to H28.

The NOESY confirmed the relative stereochemical arrangements. The hydroxyl group at C3 is in the beta-position (equatorial) making the hydrogen axial, which is confirmed by the axial-axial coupling constant to C2 $_{\beta}$  of 11 Hz. It can only have a  $^3J_{aa} = 7-12$  if the hydroxyl is beta [252]. The axial hydrogen at H3 correlated to H23 indicating it is on the same side of the ring as H3. Since the chair conformation is the most stable arrangement for the ring, H23 must be in the equatorial position. The 1,3-diaxial methyl groups on ring A, B, and C correlate to across the trans-fused rings, H24 to H25 and H25 to H26. H16 had a cross-peak to axial H27, thus making hydroxyl group equatorial/beta. Again this can be further confirmed by the vicinal  $^3J_{aa} = 8.3$  Hz. The NOE from H16 and H27 supports ring D occupying predominantly a chair or near-chair conformation. H12



correlated to H18, but also to H19b at 1.25 ppm. Axial protons are upfield to their equatorial proton. Therefore H12 correlating to the axial proton, H19b, places the epoxy group on the beta face of the triterpene. Because of the cis-ring fusion of rings D/E, there is some flexibility in the conformation of the two rings. Rings A, B, and C are locked in a chair conformation because of the trans-ring fusion. As stated previously ring D, must be in a chair conformation based on the NOEs present. On ring E, the NOE from H30 to H22b and H30 to H18 means it is also in a chair conformation. This completed the assignment of the wild-type product as 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin, **88**.

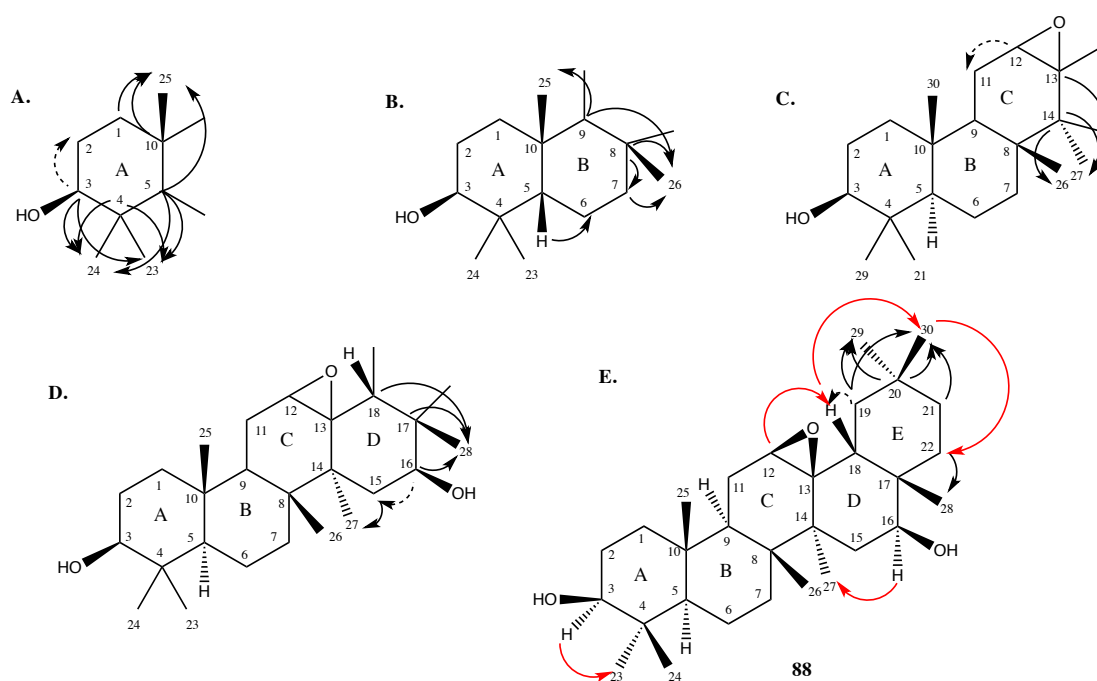


Figure 62: The oxidized triterpene, 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin (**88**), was identified as the main product from *N. benthamiana* co-expressing wild-type CYP51H with SAD1. Numbering is based on IUPAC. Substructures based on NMR correlations and coupling data. *A-E* Nuclei on the five rings were assigned using <sup>1</sup>H-<sup>1</sup>H COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows) in Table 8 resulting in the final structure *E*.

Table 8: Tabulated NMR data for 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin. Data was collected in pyridine- $d_5$ .

IUPAC	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{\text{HH}}$ (Hz)	DEPT	$^1\text{H}-^1\text{H}$ COSY Correlations	$^{13}\text{C}-^1\text{H}$ HMBG Correlations	$^1\text{H}-^1\text{H}$ NOESY Correlations
1	38.8	1.53 m, 0.99 m		CH <sub>2</sub>	2	25	
2	28.4	1.86 m		CH <sub>2</sub>			
3	78.3	3.44 dt	11.0, 5.0	CH	2	23, 24	5, 1b, 23, 2
4	39.8	-		C		23, 24	
5	56.2	0.81 m		CH		6, 23, 24, 25	
6	18.6	1.56 m		CH <sub>2</sub>			
7	35.0	*		CH <sub>2</sub>		26	
8	42.1	-		C		7, 26, 27	
9	46.5	1.48 m		CH		25, 26	
10	37.8	-		C		25	
11	22.5	1.93 m, 1.84 m		CH <sub>2</sub>			
12	54.5	2.89 d	4	CH	11		18, 19b
13	67.4	-		C		27	
14	42.9	-		C		26, 27	
15	35.2	2.32 m, 1.78 m		CH <sub>2</sub>		27	
16	66.0	4.57 dt	8.3, 5.3	CH	15a, 15b	28	19a, 15b, 27
17	39.6	-		C		28	
18	50.3	1.40 m		CH	19a	28	
19	41.9	1.90 m, 1.25 m		CH <sub>2</sub>	18	29, 30	
20	31.2	-		C		29, 30	
21	35.0	*		CH <sub>2</sub>		29, 30	
22	31.8	2.41 m, 1.39 m		CH <sub>2</sub>	21	28, 29, 30	30
23	29.0	1.23 s		CH <sub>3</sub>		24	
24	16.6	1.05 s		CH <sub>3</sub>		23	25
25	15.7	0.91 s		CH <sub>3</sub>			24, 26
26	21.2	1.35 s		CH <sub>3</sub>			25
27	23.9	1.36 s		CH <sub>3</sub>			16
28	23.0	1.50 s		CH <sub>3</sub>			
29	33.7	0.95 s		CH <sub>3</sub>		30	
30	24.1	0.86 s		CH <sub>3</sub>		29	22b
OH		5.23					
OH		5.23					

\* denotes the overlapping resonance of C7 and C21 and therefore the protons are indistinguishable.

#### 4.3.4 Structure elucidation from mutant A354L

Mutant A354L was identified through the DNA sequencing of constructs for the AHC-derived cluster “Group 3,” which produced only M2. Compound **M2** was derivatized with TMS and analyzed by GC-MS displaying an  $m/z$  of 586. This parent ion indicated the mass of compound **M2** was equal to  $\beta$ -amryin plus another hydroxylic oxygen derivatized with two TMS group. The presence of two hydroxyl groups was confirmed in the  $^1\text{H}$  NMR spectrum by the presence of two signals at 3.47 ppm and 4.57 ppm that each integrated for one hydrogen and a broad IR stretch at  $3340\text{ cm}^{-1}$ . As was determined by the  $^{13}\text{C}$ , DEPT-135, and DEPT-90 spectra, there were thirty unique carbons, eight methyls, nine methylenes, six methines, and seven quaternary carbons. The unmodified trisubstituted alkene was confirmed by the downfield chemical shifts in  $^1\text{H}$  and  $^{13}\text{C}$  NMRs. Tabulated spectral data can be found in Table 9. Original data was collected in  $\text{CDCl}_3$  before needing to switch NMR solvents to pyridine- $d_5$  for consistency. The data described below is from pyridine- $d_5$ .

The detailed assignment of the carbon-13 and proton NMR data is as follows. Rings A/B were assembled via an analogous set of HMBC and COSY correlations as were found for compound **88** (Figure 63A/B).

Similar to **88**, the quaternary carbons C14 and C8 both correlated to the methyl groups H26 and H27. However, additional long-range connectivity between C8 and H9 as well as C14 to H12 bolstered ring C.

Ring C retained the C12-C13 alkene of  $\beta$ -amryin with  $^{13}\text{C}$  peaks at 145.0 ppm and 122.9 ppm with the methine C12 (122.9 ppm) coupling to the triplet found in the  $^1\text{H}$  NMR spectrum at 5.37 ppm (Figure 63C). In the HMBC, C12 coupled with H11. Furthermore, C11 showed weak correlation to H9 and H12 confirming C11 was the allylic methylene. The quaternary carbon at 145.0 ppm, C13, coupled to the H27 methyl hydrogens at the C/D bridgehead, and thus completing ring C.

Progressing from ring C, C12 and C13 both correlated to the methine H18 on ring D (Figure 63D). The quaternary carbon, C17 coupled to H18 as well as the methyl group H28. Critically, the hydroxylated carbon, C16, showed a cross-peak to H28. This pinpointed the second hydroxyl group on compound **M2**. Finally, C16 correlated to H15 and H15 coupled back H27, the methyl group at the ring junction C/D closing ring D.

The correlations on ring E are very similar to **88**, where the HMBC data supported the coupling of C19 to the methine H18 and methyls H29 and H30. Due to correlations to each other on the HMBC, the geminal dimethyl group on ring E was H29 and H30. The quaternary carbon C20 and methylene C21 also couple to H29 and H30. Finally, the last methylene C22 correlated to H28 completing ring E.

The NOESY provided sufficient relaxation data to assign the stereochemistry. H3, assigned to the axial, alpha position consistent with the substrate provided by SAD1, correlated to H23 and H5, placing H23 in the equatorial position and H5 in the axial positions both on the alpha face. The absence of a crosspeak linking H3 to Me25 aligns with the *trans* A/B ring fusion provided by  $\beta$ -amyirin. H16 did not correlate to H28, but did correlate to H27 therefore the stereochemistry of hydroxyl group at C16 was determined to be beta. Lastly, for the geminal dimethyl group with H29 and H30, H18 correlated to H30 revealing their diaxial *syn*-facial relationship. The conformations of rings D and E were both confirmed to be in chair conformations from the NOEs similar to those found in **88**. Therefore, the identity of **M2** was 16 $\beta$ -hydroxy- $\beta$ -amyirin, **89** (Figure 63E). 16 $\beta$ -Hydroxy- $\beta$ -amyirin, also known as maniladiol, has been previously reported in literature [253, 254].  $^1\text{H}$  and  $^{13}\text{C}$  NMR data was similar to that reported by [254] in  $\text{CDCl}_3$  and by [253] pyridine- $d_5$ .

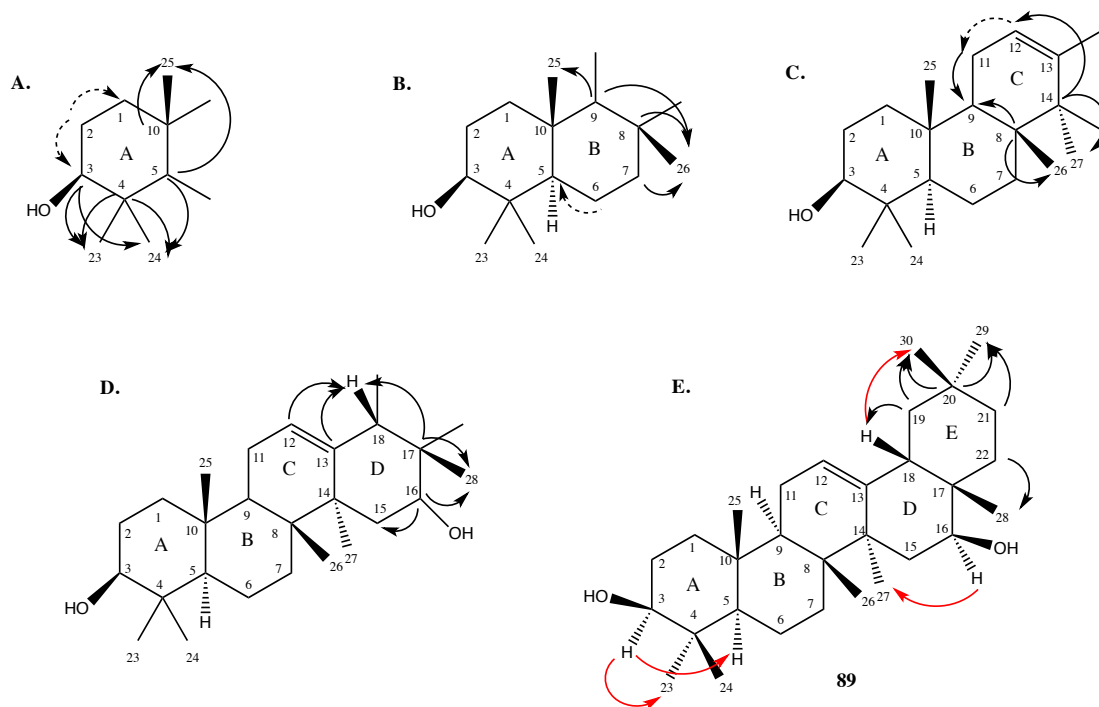


Figure 63: Structure elucidation of compound **M2** isolated from *N. benthamiana* plants coexpressing the mutant SAD2, A354L, and the triterpene SAD1. IUPAC numbered.

Substructures are based on NMR correlations and coupling data. *A-E* Nuclei were assigned and located on the five rings through the use of  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows), HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 9 resulting in the final structure *E*.

Table 9: Tabulated NMR data for 16 $\beta$ -hydroxy- $\beta$ -amyrin. Data was collected in pyridine- $d_5$ .

IUPAC	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{HH}$ (Hz)	DEPT	$^1\text{H}-^1\text{H}$ COSY Correlations	$^{13}\text{C}-^1\text{H}$ HMBC Correlations	$^1\text{H}-^1\text{H}$ NOESY Correlations
1	39.6	1.61 m; 1.04 m		CH <sub>2</sub>		25	
2	28.6	1.83 m		CH <sub>2</sub>		3w, 1a	
3	78.5	3.47 dd	11.4, 4.9	CH	2, 1bw	23, 24	5, 23
4	39.9	-		C		5w, 23, 24	
5	56.2	0.89 m		CH	6b	23, 24, 25	23
6	19.3	1.58 m; 1.42 m		CH <sub>2</sub>		5w	
7	33.5	1.58 m; 1.39 m		CH <sub>2</sub>		26	
8	40.7	-		C		18w, 26, 27	
9	47.8	1.67 m		CH		25, 26	
10	37.7	-		C		25	
11	24.4	1.95 m		CH <sub>2</sub>		12w, 9w	
12	122.9	5.37 t	3.6	CH	18w, 11	18, 11	18
13	145.0	-		C		18, 27	
14	44.4	-		C		12, 18, 15a/b, 26, 27	
15	37.0	2.11 m; 1.89 m		CH <sub>2</sub>		27	
16	64.9	4.57 dd	4.5, 11.2	CH	15a/b	18w, 15a/b, 28	21a, 27
17	38.5	-		C		18, 28	
18	50.1	2.37 dd	4.6, 14.0	CH	12w, 19	28	28, 30
19	47.6	1.91 m; 1.19 m		CH <sub>2</sub>		18, 29, 30	
20	31.7	-		C	21a/b, 22a	19aw, 29, 30	
21	35.2	1.63 m; 1.21 m		CH <sub>2</sub>	21	29, 30, 28w	
22	31.6	2.47 m		CH <sub>2</sub>		28	
23	29.2	1.27 s		CH <sub>3</sub>		3w, 24	
24	17.0	1.08 s		CH <sub>3</sub>		5, 23	
25	16.3	0.99 s		CH <sub>3</sub>		9w	
26	17.7	1.11 s		CH <sub>3</sub>		9w, 7w	
27	27.8	1.40 s		CH <sub>3</sub>		15b	
28	23.0	1.19 s		CH <sub>3</sub>		16, 18w	
29	33.9	0.95 s		CH <sub>3</sub>		19aw, 19b, 30	
30	24.6	1.01 s		CH <sub>3</sub>		19a, 29	

#### 4.3.5 Structure elucidation from mutant I471M

AsbAS1, tHMGR, and CYP51H10-I471M were coexpressed in *N. benthamiana* leaves, and after purification, characterization of the sample was attempted. I471M, a single point mutation in CYP51H had been determined through screening to accumulate M1, M3, and two additional products. Initial NMR spectra in deuteriochloroform showed that the initial I471M sample which had been expected to be the purified epoxide, was likely a mixture of triterpenes, even though the chromatography (HPLC, TLC) had only indicated one component. Over a couple of hours in  $\text{CDCl}_3$ ,  $^1\text{H}$  NMR analysis revealed four components in a ratio of approximately 3.6, 53, 39, and 4.5. Based on a cursory analysis, the compounds seemed to be the likely epoxide, epoxide isomer 1, epoxide isomer 2, and a fourth compound, respectively. Data collected over a week showed epoxide isomer 1 and the epoxide to be decaying overtime primarily to epoxide isomer 2. By 9 days the epoxide and epoxide isomer 1 were completely transformed to the more stable product. A second purification was completed using material from a second plant infiltration experiment and two compounds were isolated I471M-1 and I471M-2 avoiding any contact with acidic solvent. For NMR analysis, the solvent was switched to pyridine- $d_5$  and the components no longer decayed.

#### 4.3.6 Structure elucidation of product 1 from mutant I471M

Compound I471M-1 was believed to be a  $\beta$ -amyryn derivative. After isolation, the compound was derivitized with TMS and analyzed by GC-MS to give an  $m/z$  of 514. Initially based on the mass spectra, isolate I471M-1 was thought to be  $\beta$ -amyryn with one oxygen that could be silylated and one additional non-hydroxylic oxygen. However, the  $^1\text{H}$  NMR presented two signals at 3.48 and 3.98 ppm suggesting the presence of two hydroxyl groups. The structural elucidation of the mutant relied on complete evaluation of the data set in order to resolve whether the additional oxygen present is non-hydroxylic or non-TMS derivitized hydroxyl group.

According to the  $^{13}\text{C}$ , DEPT 135, and DEPT 90 there were twenty-nine unique carbons, eight methyls, nine methylenes, six methines, and six quaternary carbons. However, since this compound was believed to be a  $\beta$ -amyryn derivative, there was a

concern that insufficient signal dispersion or accidental by chemical shift equivalent carbons might be present. The  $^{13}\text{C}$  NMR confirmed that an alkenyl functionality similar to that of  $\beta$ -amyirin was present in the structure of I471M-1 with two  $^{13}\text{C}$  shifts at 158.6 ppm and 118.8 ppm. In addition, the HSQC showed a CH coupling between C12 at 118.8 ppm to hydrogen at 5.79 ppm indicative of an  $\text{sp}^2$ . Tabulated spectral data can be found in Table 10.

The detailed assignment of the carbon-13 and proton NMR data is as follows. Starting at C3, a network of correlations equivalent to those used in the structure elucidation of **88** uncovered the molecular framework of ring A; the HMBC and COSY correlations are shown in Figure 61A. Ring B also showed similar correlations as **88**, but the methyl group at C26 had a proton shift significantly more downfield, moving from 1.35 ppm to 1.74 ppm in the  $^1\text{H}$  NMR spectrum (Figure 64B). This was presumably due to the presence of a proximate oxygenated functionality.

The quaternary carbon at C14 also coupled to H26 and H27, but C14 had a weak cross-peak to H26 and a stronger cross-peak to H27. This allowed the quaternary carbons C8 and C14 to be distinguished from one another; therefore C8 connected ring B to ring C and C14 was at the bridgehead of rings C/D (Figure 64C). It also helped to confirm the presence of the alkene from  $\beta$ -amyirin was still intact. C13, the quaternary carbon at 158.6 ppm correlated to H26 and H27. The other double bonded carbon, C12 showed weak coupling to H11b in the HMBC and was further confirmed from the cross-peak between the allylic and vinylic hydrogens in the COSY.

The key correlation came from cross-peak between the carbon at 70.2 ppm (C15) and to the methyl at the bridgehead between ring C/D, H27 (Figure 64D). This methine carbon must be attached to an oxygen due to its downfield chemical shift in the  $^{13}\text{C}$  spectrum. HSQC data linked this carbon to the methine hydrogen at 3.98 ppm, leading us to surmise there was a hydroxyl group attached to C15 that was not silylated during analysis for GC-MS. H15 coupled to H16a in the COSY. At the D/E fusion, C17 and C18 at the D/E ring fusion coupled to the H28.

At the ring fusion of D/E, the COSY spectrum showed the correlation H18 to H19a. C19 then coupled to C29 and C30 therefore identifying the other geminal dimethyl group. They also displayed cross-peaks to each other. The last methylene group was C22



showed a cross-peak to H28. The quaternary carbon, C20, of the geminal dimethyl group was hidden underneath another carbon peak and was identified by the HMBC. Methyl carbon C28 correlated to the geminal dimethyl, but that would be a coupling of  ${}^6J_{\text{CH}}$  which too far for our parameters to have coupled. There for the quaternary carbon C20 must be directly underneath the C28 at 29.9 ppm.

The NOESY experiment was central to the stereochemical analysis (Figure 64E). Assuming this compound is a  $\beta$ -amyrin derivative means the hydroxyl group at C3 is in the beta-position (equatorial). Again, the  ${}^3J_{\text{aa}}$  coupling constant of 10.7 Hz further confirms that hydrogen occupies the axial position. The axial hydrogen at H3 correlated to H23 indicating it is on the same face of the ring as H3. Since the ring is in a chair conformation, H23 must be in the equatorial position. There was a cross-peak between H15 to H27, H27 is axial leading to the conclusion that the  $\beta$ -hydroxyl group axial. Finally H18 had discernable cross-peaks to H28 and H30, in accord with the expected *cis* ring junction at ring D/E. Therefore, the final structure elucidation for the first compound isolated from I471M is 15 $\beta$ -hydroxy- $\beta$ -amyrin, **90**.

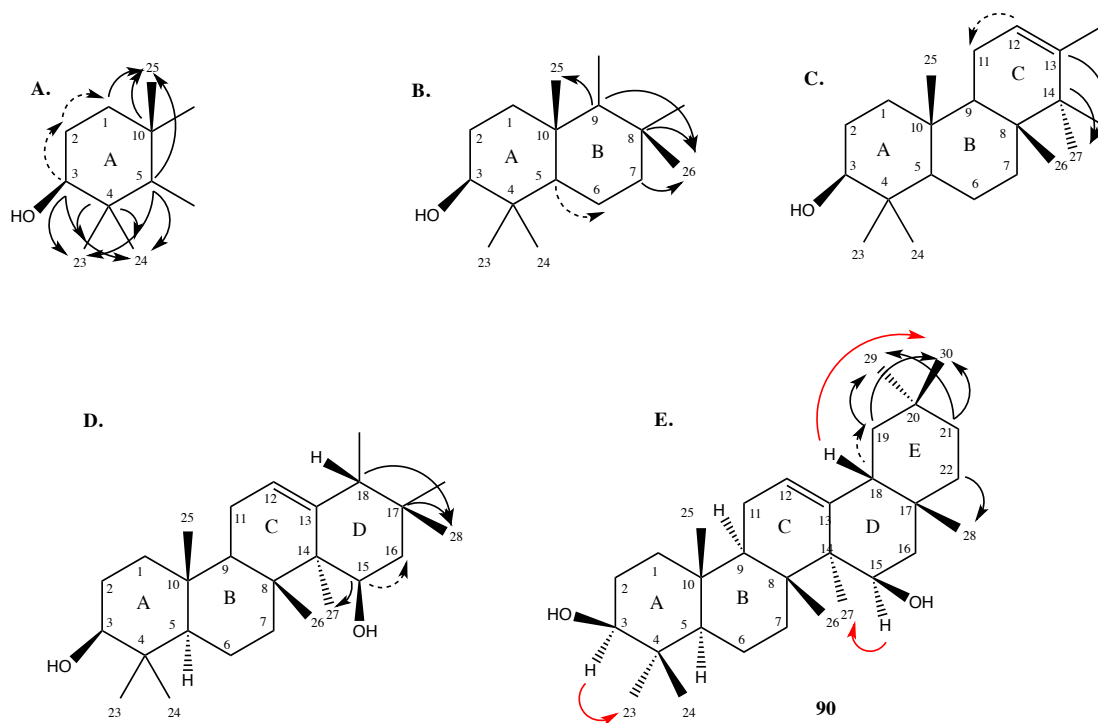


Figure 64: Deduction of the structure of compound **90** produced by the coexpression of SAD2 mutant M471I with SAD1 in *N. benthamiana*. The elucidation molecule is the oxidized triterpene derivative, 15 $\beta$ -hydroxy- $\beta$ -amyrin. IUPAC numbered. Substructures are based on NMR correlations and coupling data *A-E* Nuclei were assigned and located on the five rings through the use of  $^1\text{H}$ - $^1\text{H}$  COSY COSY (dashed arrows), HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 10 resulting in the final structure *E*.

Table 10: Tabulated NMR data for 15 $\beta$ -hydroxy-13 $\beta$ -amyrin. Data was collected in pyridine-d<sub>5</sub>.

IUPAC #	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{\text{HH}}$ (Hz)	DEPT	$^1\text{H}-^1\text{H}$ COSY Correlations	$^{13}\text{C}-^1\text{H}$ HMBC Correlations	$^1\text{H}-^1\text{H}$ NOESY Correlations
1	38.9	1.56 m, 0.94 m	-	CH <sub>2</sub>		25	
2	28.6	1.86 m	-	CH <sub>2</sub>	1b		
3	78.6	3.48 dt	10.7, 5.4	CH	2	23, 24	5, 23
4	40.0	-	-	C		23, 24	
5	56.8	0.9 m	-	CH	6b	23, 24, 25	
6	20.0	1.72 m, 1.64 m	-	CH <sub>2</sub>		26	
7	43.8	2.23 m, 1.46 m	-	CH <sub>2</sub>		26	
8	39.9	-	-	C		26	
9	50.4	1.7 m	-	CH		25, 26	
10	38.6	-	-	C		25	
11	38.4	2.11 m, 1.77 m	-	CH <sub>2</sub>			
12	118.8	5.79 d	12	CH	11a, 11b	11b	
13	158.6	-	-	C		26, 27, 11b w	
14	43.2	-	-	C		26 w, 27	
15	70.2	3.98 dd	4.5, 7.5	CH	16a	27	16a, 16bw 27
16	30.2	2.26 m, 1.98 m	-	CH <sub>2</sub>			
17	35.9	-	-	C		28	
18	39.2	2.57 dd	3, 14	CH	19a	30	
19	36.9	1.42 m, 1.14 m	-	CH <sub>2</sub>		29, 30	
20	-	-	-	C			
21	34.1	1.47 m, 1.30 m	-	CH <sub>2</sub>		29, 30	
22	36.1	1.50 m, 1.16 m	-	CH <sub>2</sub>		28	
23	29.1	1.27 s	-	CH <sub>3</sub>		24	
24	16.8	1.10 s	-	CH <sub>3</sub>		23	25
25	16.7	1.07 s	-	CH <sub>3</sub>			24, 26
26	26.4	1.74 s	-	CH <sub>3</sub>			25
27	22.2	1.08 s	-	CH <sub>3</sub>			
28	29.9	1.09 s	-	CH <sub>3</sub>			
29	33.6	1.03 s	-	CH <sub>3</sub>		30	
30	30.7	1.04 s	-	CH <sub>3</sub>		29	

#### 4.3.7 Structure elucidation of product 2 from I471M

Mutant I471M-2 was also believed to be a  $\beta$ -amyrin derivative. After isolation, the compound was derivitized with TMS and analyzed by GC-MS to give a  $m/z$  of 514. This indicated the presence of one hydroxyl group from  $\beta$ -amryin and one additional non-hydroxylic oxygen.  $^1\text{H}$  NMR and IR confirmed the hydroxyl group. Mutant I471M-2 provided a quasimolecular ion  $[\text{M} + \text{H}]^+$  with an  $m/z$  of 443.3899 consistent with the molecular formula  $\text{C}_{30}\text{H}_{51}\text{O}_2$  (calcd  $m/z$  443.3889). The structural elucidation of the mutant relied on complete evaluation of the data set in order to determine the location and form of the oxygen.

According to the  $^{13}\text{C}$ , DEPT 135, and DEPT 90 there were thirty unique carbons consistent with a triterpene, eight methyls, ten methylenes, five methines, and seven quaternary carbons. The beta-hydroxy group at C3 of ring A presented signals for an oxygenated carbon at 78.1 ppm in the  $^{13}\text{C}$  spectrum and 3.42 ppm in the  $^1\text{H}$  spectrum. The  $^{13}\text{C}$  spectrum revealed a carbonyl group at 212.2 ppm. According to the HSQC, the carbonyl carbon does not correlate to any hydrogens so therefore must be a ketone. Tabulated spectral data can be found in Table 11.

The detailed assignment of the carbon-13 and proton NMR data is as follows. Ring A and ring B displayed the same COSY and HMBC correlations as **88** (Figure 65A/B). The most noticeable change from **88** was the hydrogens at the methyl group C23 became the most downfield methyl group.

Both quaternary carbons C8 and C14 correlated to the hydrogens on methyl groups H26 and H27 (Figure 65C). However, C8 correlated back to H7 thus distinguishing the two carbons from one another. There was no longer an alkene present from  $\beta$ -amryrin as seen from the lack of peaks from 120-145 ppm in the  $^{13}\text{C}$  spectrum and no hydrogen peak around 5.25. Instead C13 had become a methine as seen in the DEPT90. This methine at 2.85 ppm in the  $^1\text{H}$  spectrum split into a doublet and in the HMBC coupled to H27. C12 was then identified as the carbonyl that correlated to H13 and H11. This pointed to the oxidation of the double bond of  $\beta$ -amyrin to the ring C ketone.

According to the HMBC, C18 correlated to H13, which was confirmed in the COSY with H18 coupling to H13. The HMBC showed C15 to H27, which gave initial

insights to ring D (Figure 65D). C18 coupled to the methyl group H28. The methyl group became a key component of ring D because it has had cross-peaks with C16, C17, and C22. To complete ring D, in the COSY H16b correlated H15a. The remaining ring E correlations are the same as the previous structures.

The NOESY completed all the stereochemistry specifications including H23 in the equatorial position because it had a correlation to H3 in the axial position (Figure 65E). The  $^3J_{aa}$  coupling constant of 10.7 Hz for H3 confirms that it is in the alpha position. H13 correlated H18, H28, and H26 indicating that H13 was axial in the thermodynamically most favorable trans C/D fusion. Last H18 had a cross-peak to H30 and H28 making it equatorial in keeping with the chair conformations of ring D/E. Thus the second compound isolated from I471M is the compound 12-keto-13 $\beta$ - $\beta$ -amyrin, **91**.

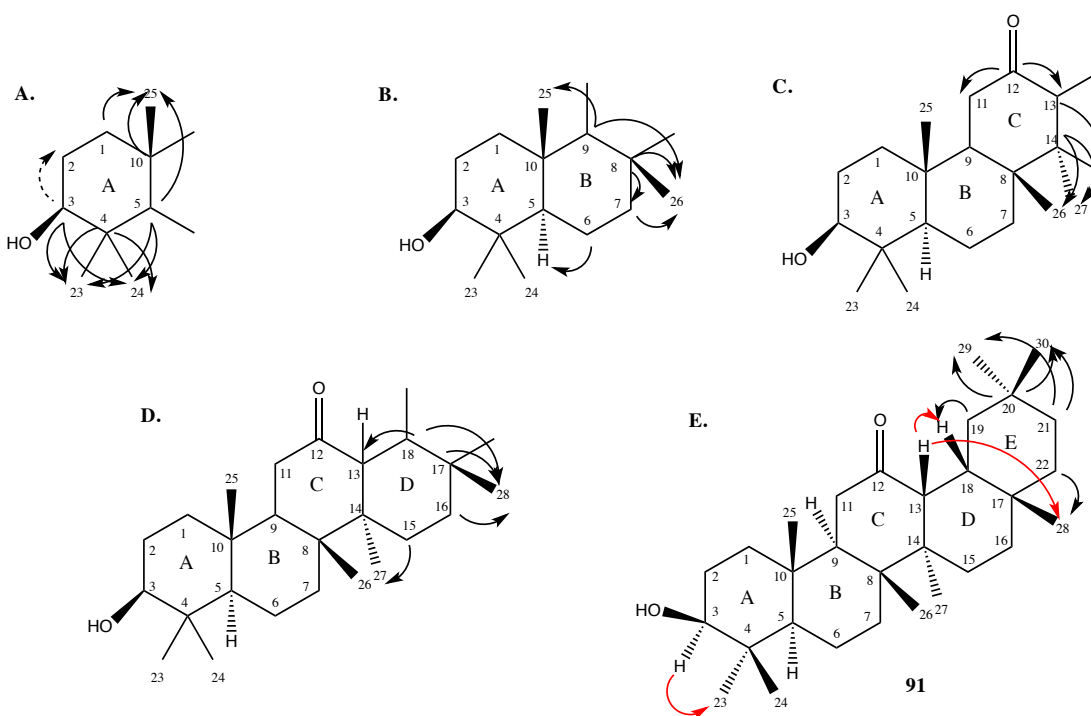


Figure 65: Deduction of the structure of compound **91**, a second compound produced by the coexpression of SAD2 mutant M471I with SAD1 in *N. benthamiana*. The elucidated molecule is the oxidized triterpene derivative, 12-keto-13 $\beta$ - $\beta$ -amyrin. IUPAC numbered. Substructures are based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 11 resulting in the final structure *E*.

Table 11: Tabulated NMR data for 12-keto-13 $\beta$ - $\beta$ -amyrin. Data was collected in pyridine-d<sub>5</sub>.

IUPAC #	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{HH}$ (Hz)	DEPT	$^1\text{H}-^1\text{H COSY}$ Correlations	$^{13}\text{C}-^1\text{H HMBC}$ Correlations	$^1\text{H}-^1\text{H NOESY}$ Correlations
1	38.7	1.44 m; 0.89 m		CH <sub>2</sub>		25	
2	28.3	1.81 m		CH <sub>2</sub>	3		
3	78.1	3.42 dt	10.4, 5.40	CH		23, 24	5, 23, 25
4	39.8	-		C		23, 24	
5	55.8	0.84 m		CH		23, 24, 25	
6	19.1	1.66 m; 1.46 m		CH <sub>2</sub>			
7	32.5	1.43 m; 1.31 m		CH <sub>2</sub>		26	
8	42.2	-		C		7, 26, 27	
9	50.4	1.70 m		CH		25, 26	
10	37.5	-		C		25	
11	39.2	2.42 dd; 2.28 dd	5.0, 16.8; 13.2, 16.7	CH <sub>2</sub>			
12	212.2	-		C		11ab, 4	
13	50.6	2.85 d	3.75	CH		27	18, 28, 26
14	42.7	-		C		26, 27	
15	26.6	1.72 m; 0.92 m		CH <sub>2</sub>		27	
16	27.0	1.95 m; 0.78 m		CH <sub>2</sub>		28	
17	33.2	-		C		28	
18	37.8	2.15 m		CH	13	13, 28, 30	13, 28, 30
19	35.5	1.41 m; 1.13 m		CH <sub>2</sub>		29, 30	
20	31.6	-		C		29, 30	
21	37.7	1.39 m		CH <sub>2</sub>			
22	38.3	1.55 m; 1.20 m		CH <sub>2</sub>		28	
23	28.9	1.25 s		CH <sub>3</sub>		24	
24	16.6	1.07 s		CH <sub>3</sub>		23	
25	15.8	0.88 s		CH <sub>3</sub>			
26	16.4	1.10 s		CH <sub>3</sub>	13		3
27	21.4	1.00 s		CH <sub>3</sub>	13		
28	27.7	0.93 s		CH <sub>3</sub>			
29	34.2	0.96 s		CH <sub>3</sub>		30	
30	23.8	1.04 s		CH <sub>3</sub>		29	

#### 4.3.8 Structure elucidation of product 3 from I471M

In the initial attempts to analyze the products of the SAD2 I471M mutant, it decomposed, apparently from small amounts of HCl present in  $\text{CDCl}_3$ . After the second purification, two products were identified I471M-1 and I471M-2, but neither were the epoxide. The epoxide seemed to be degrading through the purification process and so the conditions were simplified to avoid elevated temperatures and only one isocratic flash column with pH neutral solvents was performed. This allowed the presumptive epoxide, I471M-3, to be isolated in about 70% purity which could be further purified by crystallization with hot methanol.

Compound I471M-3 was derivitized with TMS and analyzed by GC-MS to give an  $m/z$  of 514. Similar to compound I471M-2, this MS data indicated the presence of one hydroxyl group from  $\beta$ -amryin and one additional non-hydroxylic oxygen, confirmed by  $^1\text{H}$  NMR and IR. Mutant I471M-3 provided a quasimolecular ion  $[\text{M} + \text{H}]^+$  with an  $m/z$  of 443.3907 consistent with the molecular formula  $\text{C}_{30}\text{H}_{51}\text{O}_2$  (calcd  $m/z$  443.3889). The structural elucidation of the mutant relied on complete evaluation of the data set in order to determine the location and identity of the oxygen functionality.

According to the  $^{13}\text{C}$ , DEPT 135, and DEPT 90 experiments, there were thirty unique carbons consistent with a triterpene, eight methyls, ten methylenes, five methines and seven quaternary carbons. Based on the  $^{13}\text{C}$  spectrum, the  $3\beta$ -hydroxyl group on the A ring was still present due to the  $^{13}\text{C}$  peak at 78.3 ppm bearing a proton at  $\delta$  3.45 ppm. The alkene present in  $\beta$ -amryin appeared to be the modification site due to the lack of signals at 145 and 122 ppm. Tabulated spectral data can be found in Table 12.

The detailed assignment of the carbon-13 and proton NMR data is as follows. Ring A and ring B displayed the same correlations as **88**. All of the HMBC and COSY correlations are shown in Figure 66A/B.

When assigning rings C and D, quaternary carbons C8 and C14 both correlated to the methyl groups H26 and H27, and no other cross-peaks were present to distinguish the two apart (Figure 66C). Spectroscopic data is consistent with the oxidization of the alkene to the epoxide, which has peaks at 67.4 ppm and 55.0 ppm with the relatively upfield  $^{13}\text{C}$  peak correlating to a hydrogen at  $\delta$  2.79 ppm. The COSY showed a

correlation between H12 and allylic H11b. To finalize ring C, the HMBC revealed C13 coupled to H27.

On ring D, C15 correlated to back H27 and was further confirmed by the COSY spectrum that showed a cross-peak from H27 to H15a (Figure 66D). The last methine to be assigned was C18. It is notable because of its downfield shift in the  $^{13}\text{C}$  spectrum (48.9 ppm) and upfield shift in the  $^1\text{H}$  (1.11 ppm) is characteristic of a hydrogen at the *cis*-fused D/E rings. C18 was linked through the HMBC to the methyl H28. The quaternary carbon C17 and the methylene C16 was also coupled to H28.

Finally on ring E, the COSY yielded a strong  $^3J$  correlation between H18 and the likely axial H19a, which was extended by the correlation of C19 to H29 and H30. The last quaternary carbon C20 and methylene C21 both coupled to H29 and H30. Finally C22 correlated back to the methyl group at the ring D/E bridgehead, H28.

Similar to compound **88**, the NOESY was central to the stereochemistry specifications and gave similar correlations (Figure 66E). Assuming this compound is a  $\beta$ -amyirin derivative, the hydroxyl group at C3 is in the beta-position (equatorial). The NOESY confirmed that H23 is in the equatorial position because it had a correlation to H3 in the axial position and H30 is axial due to its correlation to H18. Critically, H12 correlated to H18 and H19b. Due to the *cis*-fusion of rings D/E and both rings are in a chair conformation, H12 must be alpha in order to correlate to axial H19b. With the improvements to the purification conditions and a different NMR solvent, it could finally be shown that one of the products of the I471M mutant, I471M-3 was in fact 12,13 $\beta$ -epoxy- $\beta$ -amyirin, **92**.



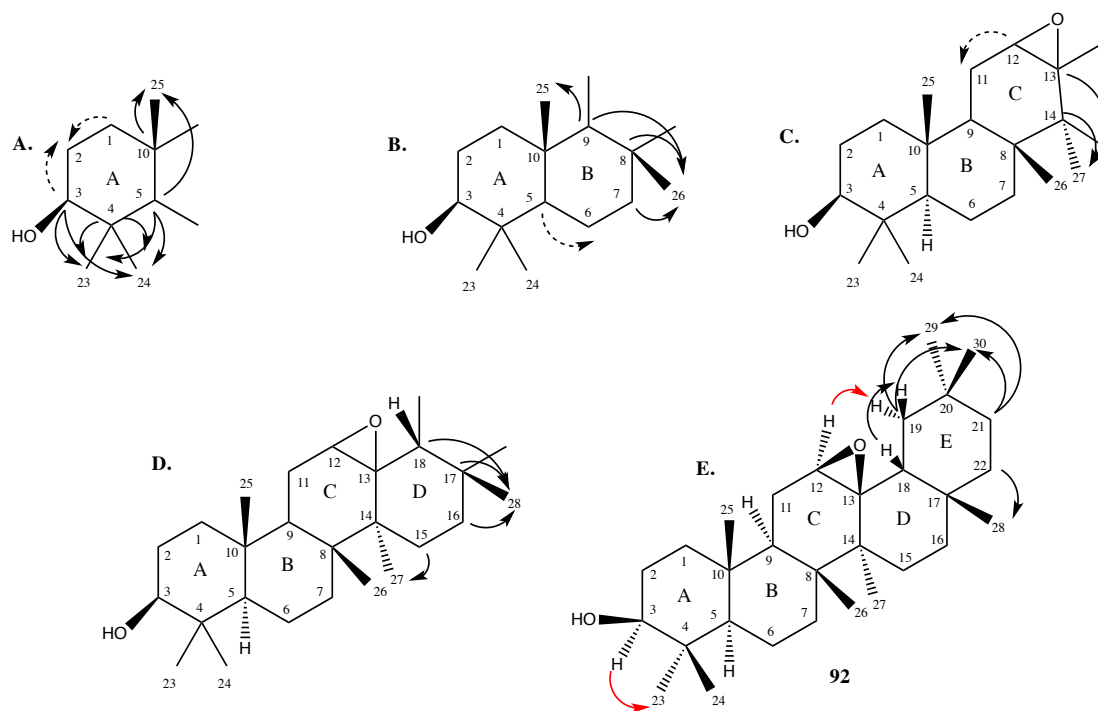


Figure 66: Structure elucidation of 12,13 $\beta$ -epoxy- $\beta$ -amyrin, **92**. This compound was isolated as M471I-3 from SAD2 mutant M471I. Substructures numbered according to IUPAC. Substructures are based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 12 resulting in the final structure *E*.

Table 12: Tabulated NMR data for 12-keto-13 $\beta$ - $\beta$ -amyrin. Data was collected in pyridine-d<sub>5</sub>. \*C-7 and C-14 could not be distinguished from one another.

IUPAC #	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{HH}$ (Hz)	DEPT	$^1\text{H}-^1\text{H}$ COSY Correlations	$^{13}\text{C}-^1\text{H}$ HMBC Correlations	$^1\text{H}-^1\text{H}$ NOESY Correlations
1	38.8	1.55 m, 1.01 m	-	CH <sub>2</sub>	2 a/b	25	
2	28.3	1.86 m	-	CH <sub>2</sub>			
3	78.3	3.45 m	-	CH	2 a/b	23, 24	5, 23, 2 a/b
4	39.7	-	-	C		23, 24	
5	56.2	0.84 m	-	CH	6	23, 24, 25	6a/b
6	18.6	1.58 m	-	CH <sub>2</sub>	5		
7	34.8	1.45 m, 1.19 m	-	CH <sub>2</sub>		26	
8	41.8*	-	-	C		26, 27	
9	47.1	1.52 m	-	CH		25, 26	
10	37.9	-	-	C		25	
11	22.3	1.93 m, 1.80 m	-	CH <sub>2</sub>			
12	55.0	2.79 d	3	CH	11b		18, 19b
13	67.4	-	-	C		27	
14	40.9*	-	-	C		26, 27	
15	25.1	2.04 m, 1.10 m	-	CH <sub>2</sub>		27	
16	28.3	2.06 m, 0.99 m	-	CH <sub>2</sub>		28	
17	34.0	-	-	C		28	
18	48.9	1.11 m	-	CH	19a	28	30
19	41.7	1.80 m, 1.19 m	-	CH <sub>2</sub>		29, 30	
20	31.2	-	-	C		29, 30	
21	35.1	1.34 m, 1.10 m	-	CH <sub>2</sub>		29, 30	
22	37.7	1.49 m, 1.25 m	-	CH <sub>2</sub>		28	
23	28.9	1.25 s	-	CH <sub>3</sub>		24	
24	16.6	1.07 s	-	CH <sub>3</sub>		23	
25	15.7	0.95 s	-	CH <sub>3</sub>			2a/b, 11a
26	21.1	1.32 s	-	CH <sub>3</sub>			
27	22.4	1.24 s	-	CH <sub>3</sub>			
28	29.1	1.25 s	-	CH <sub>3</sub>			
29	33.6	0.96 s	-	CH <sub>3</sub>		30	27
30	23.7	0.806 s	-	CH <sub>3</sub>		29	

#### 4.3.9 Deciphering the products produced by mutant I471M

From the first sample,  $^1\text{H}$  NMR analysis revealed four components: the likely epoxide, isomer 1, isomer 2, and a fourth compound corresponding to approximately 3.6, 53, 39, and 4.5% of the sample, respectively. A subsequent transient expression of I471M provided new leaf material, from which two compounds were purified collected that both gave an  $m/z$  of 514 upon silylation indicating the two compounds had  $m/z$  equivalent to  $\beta$ -amyrin plus oxygen, but also that the oxygen was not a hydroxyl group. The compound I471M-2 was easily identified as 12-keto-13 $\beta$ - $\beta$ -amyrin and was also identified as isomer 1 from the initial sample. It was concluded that the initial sample, believed to be the epoxide, was rapidly opened by acid catalysis to 12 $\alpha$ -keto-13 $\beta$ - $\beta$ -amyrin **93** (isomer 2), which then slowly isomerized to 12-keto-13 $\beta$ - $\beta$ -amyrin, **91** (I471M-2) (Figure 67) [225]. From the original sample, this explained the presence of three of the four compounds. The fourth component was never fully identified.

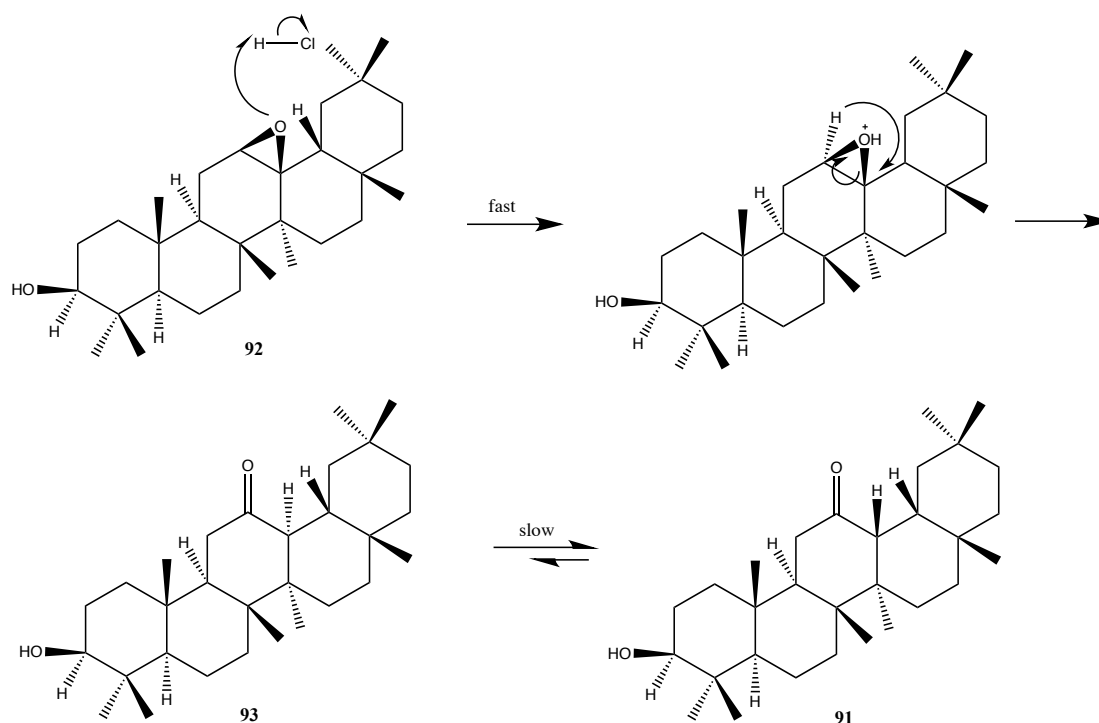


Figure 67: Acid-catalyzed epoxide opening to 12 $\alpha$ -keto-13 $\beta$ - $\beta$ -amyrin and tautomerization to 12 $\beta$ -keto-13 $\beta$ - $\beta$ -amyrin.

However, I471M-1 was not as easily identified mainly because the compound contained a hydroxyl group, which should have been indicated in the mass spectrum by the doubly silylated  $m/z$  586 ion. Initially, I471M-1 gave a parent ion with an  $m/z$  of 514 (Figure 68). Due to the hindered position of the  $15\beta$ -hydroxy group, the second hydroxyl group was not readily silylated. The non-silylated hydroxyl group is consistent with the mass spectrum. The base peak of  $m/z$  234 is consistent with a fragment containing a portion of ring C + D/E plus an oxygen and a small peak of 280 for rings A/B. It was determined that in the crude data, there were two components at 24.4 min from GC/MS, at 24.33 min was the epoxide (I471M-3) and 24.45 was the  $15\beta$ -hydroxy compound (I471M-1). The loss of the methyl to give  $m/z$  219 was only seen in 12-keto and 12,13 $\beta$ -epoxy triterpenes.

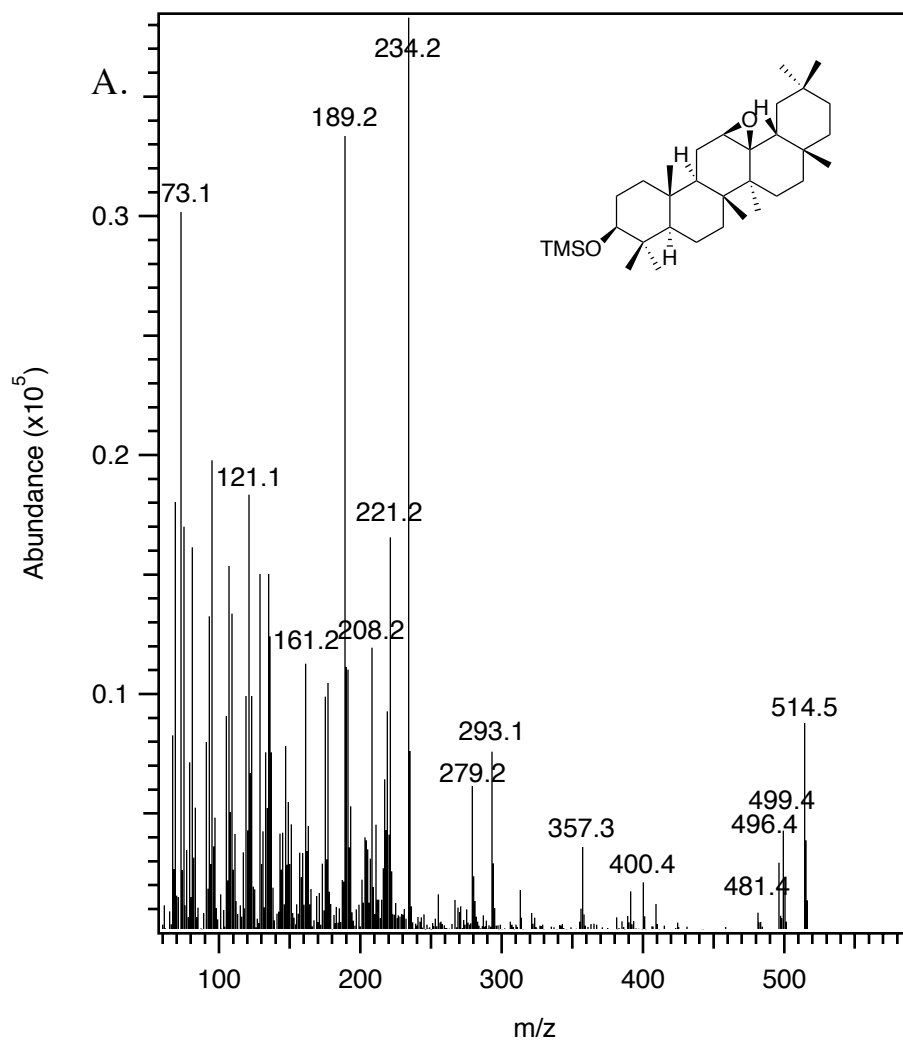


Figure 68: Crude mass spectrum with epoxide and 15 $\beta$ -hydroxy- $\beta$ -amyirin. *A.* The front half of the peak at 24 min looked to be the epoxide *B.* The tail of the peak at 24 min looked to be 15 $\beta$ -hydroxy- $\beta$ -amyirin with only one hydroxyl group silylated.

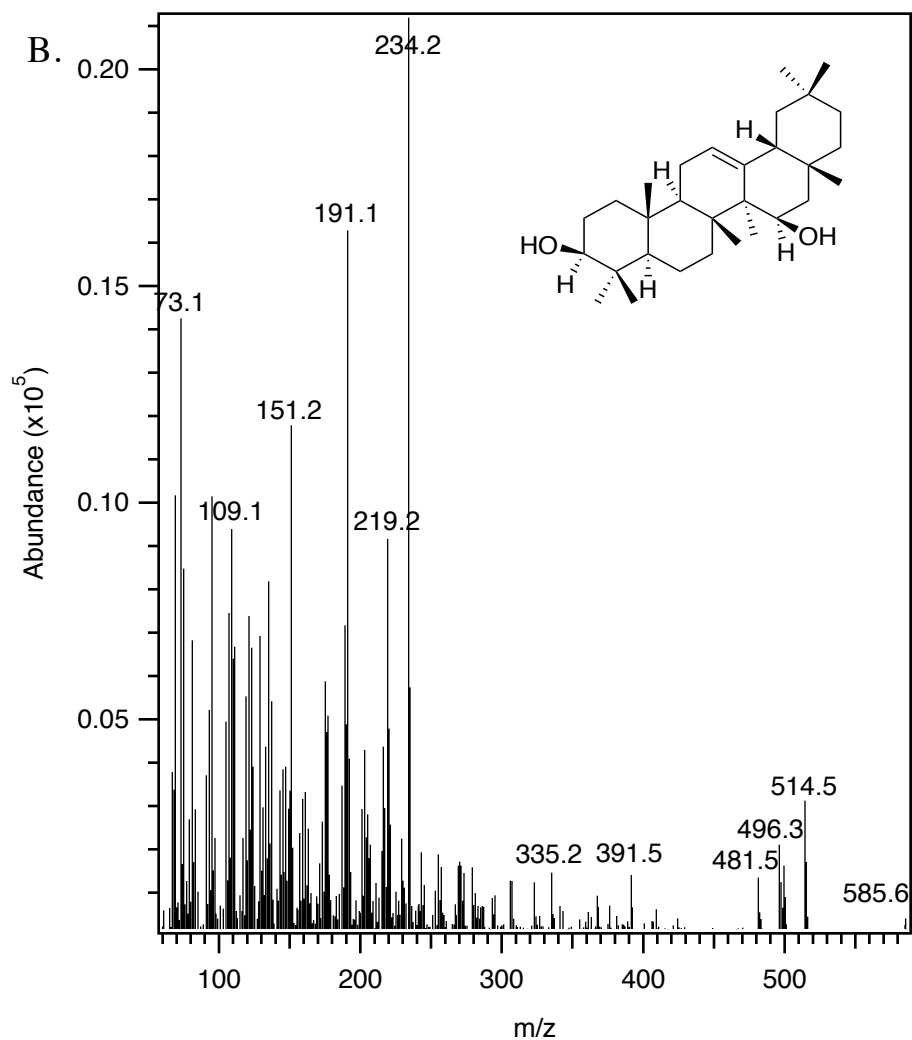


Figure 68 continued

From the second sample collected, the higher temperature during purification destroyed the epoxide to the ketone (I471M-2), revealing the 15 $\beta$ -hydroxy- $\beta$ -amyryn that could subsequently be extracted. Had the epoxide never been destroyed, 15 $\beta$ -hydroxy- $\beta$ -amyryn would probably not have been found. 15 $\beta$ -Hydroxy- $\beta$ -amyryn is an interesting product; it is the thermodynamically less stable alcohol. Positioning of the substrate in the enzyme active site seems to be relatively flexible in order to produce this compound along with 16 $\beta$ -hydroxy- $\beta$ -amyryn and the  $\beta$ -epoxide. 15 $\beta$ -Hydroxy- $\beta$ -amyryn appears to come from oxygen insertion by the CYP450 to the same face of  $\beta$ -amyryn. To check the

stability of the compounds, both I471M-1 and I47M-2 were dissolved in  $\text{CDCl}_3$ ; neither triterpene decayed over time as monitored by  $^1\text{H}$  NMR spectroscopy.

#### 4.3.10 Analysis of mass spectra of SAD2 mutants

The Osbourn group collected original mass spectra data after infiltration and isolation of the compounds. Initially, the  $15\beta$ -hydroxy position did not silylate and this precluded the assignment. We wanted to replicate the GC/MS results completed by the Osbourn group with the isolated compounds. For the  $15\beta$ -hydroxy- $\beta$ -amyrin, we were able to silylate the 15-hydroxy position as seen in Figure 69. However the mass spectrum was still quite different than the other mass spectra collected. All products were derivatized with TMS before GC/MS analysis. The mass spectra for 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin, 12,13 $\beta$ -epoxy- $\beta$ -amyrin, and 12-keto- $\beta$ -amyrin can all be found in the Appendix.

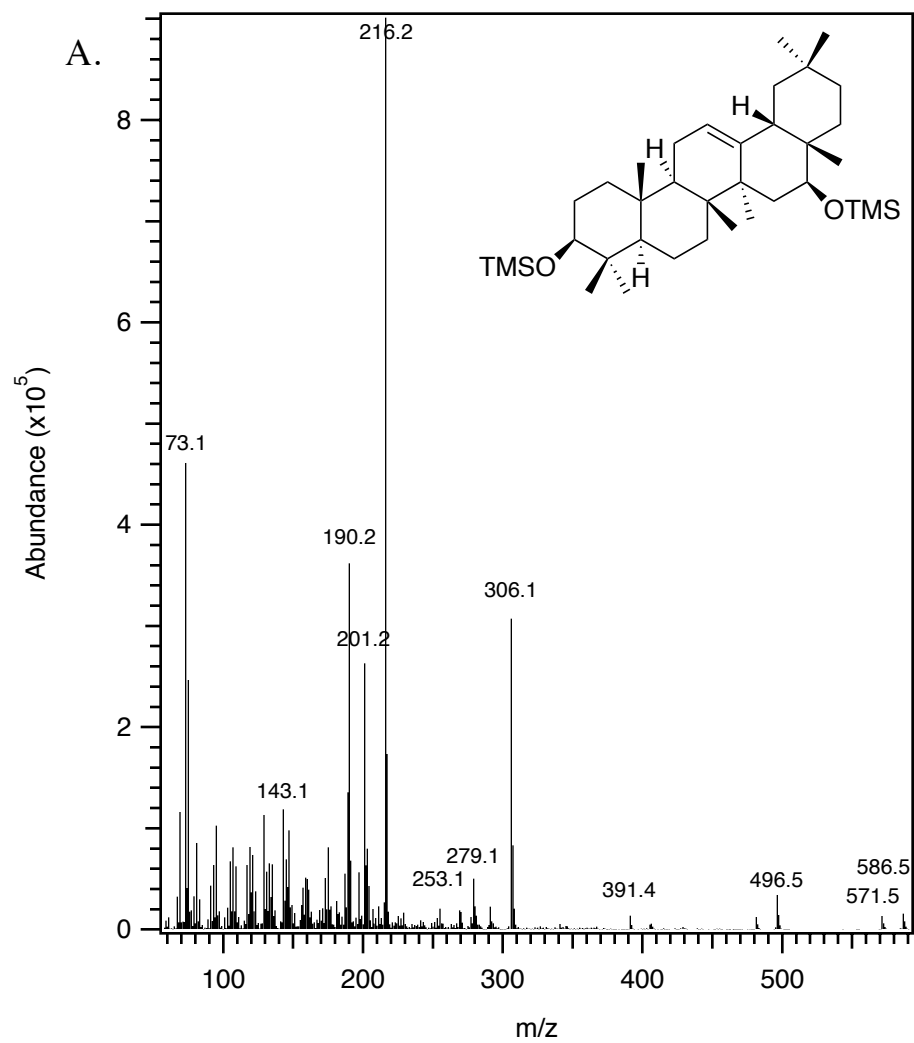


Figure 69: Mass spectra collected for I471M-1 and A354L. *A.* Mass spectrum of 16 $\beta$ -hydroxy- $\beta$ -amyrin, the product from plants expressing SAD1 and the A354L mutant of SAD2 *B.* Mass spectrum of 15 $\beta$ -hydroxy- $\beta$ -amyrin.



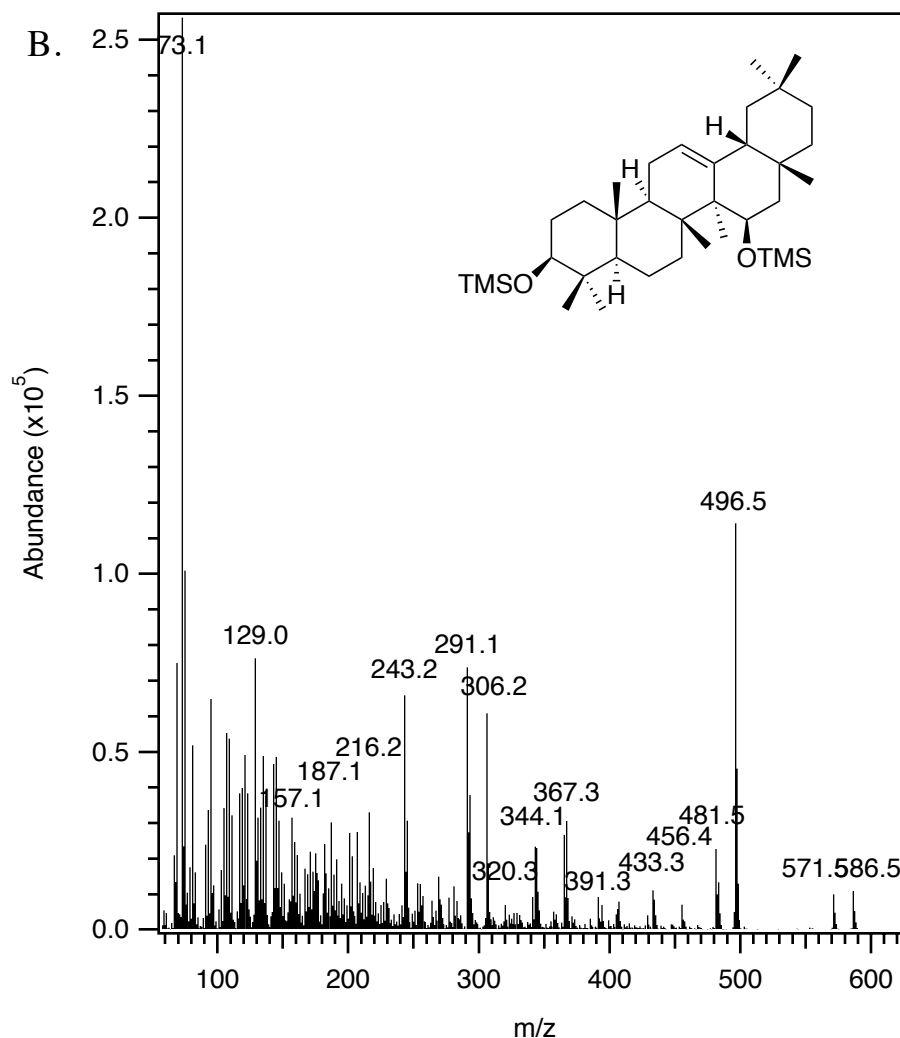


Figure 69 continued.

In all the samples, there is a loss of TMSOH from the parent ion thought to be from the 3-position of  $\beta$ -amyrin. The main product of the wild-type enzyme and 16 $\beta$ -hydroxy- $\beta$ -amyrin undergo a retro Diels-Alder cleavage to give an  $m/z$  of 279 representing rings A/B + TMS less H<sup>+</sup> (Figure 70). In a retro Diels-Alder reaction of the parent radical cation, the charge can be localized to either fragment; for 16 $\beta$ -hydroxy- $\beta$ -amyrin, we see predominantly the other fragment  $m/z$  306. A McLafferty rearrangement of 12-keto-13 $\beta$ - $\beta$ -amyrin leads to the  $m/z$  279 ion and the C/D/E fragment with an  $m/z$  of 234. Ready rearrangement of 12,13 $\beta$ -epoxy- $\beta$ -amyrin leads to the fragmentation manifold

12-keto-13 $\beta$ - $\beta$ -amyrin and the  $m/z$  279 and 234 ions. Interestingly, 15 $\beta$ -hydroxy- $\beta$ -amyrin does not undergo the retro Diels-Alder as seen by the lack of  $m/z$  of 279.

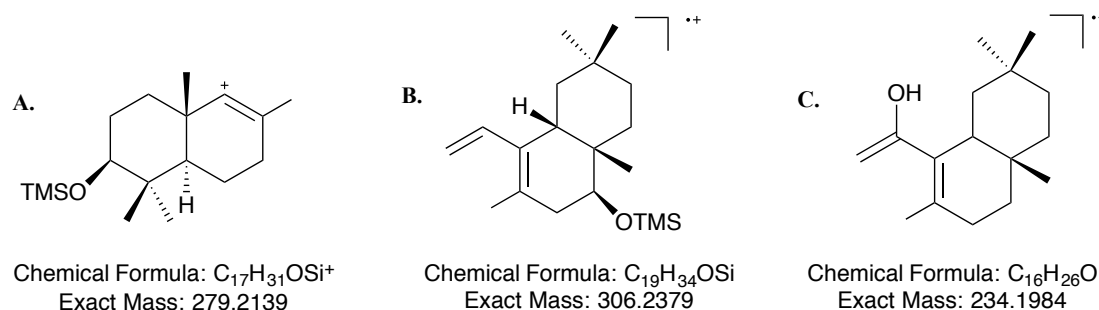


Figure 70: Key fragmentations from the mass spectra of the isolated products. *A.* The loss of  $m/z$  279 is visible in the mass spectra for all products except 15 $\beta$ -hydroxy- $\beta$ -amyrin. *B.* The C/D/E ring fragment released from the retro Diels-Alder reaction fragmentation of 16 $\beta$ -hydroxy- $\beta$ -amyrin. *C.* The other fragment from 12-keto and 12,13 $\beta$ -epoxy seen from the loss of  $m/z$  279.

Although 15 $\beta$ -hydroxy- $\beta$ -amyrin and 16 $\beta$ -hydroxy- $\beta$ -amyrin are both hydroxylated on ring D, even when both are silylated, their fragmentation patterns are very different. As stated previously, for 16 $\beta$ -hydroxy- $\beta$ -amyrin, the loss of TMSOH is seen, as in all  $\beta$ -amyrin derivatives, with a small peak at 496. The pentacyclic ring is cleaved at rings A/B + TMS – H<sub>2</sub> to give an  $m/z$  of 279. When the A/B ring is cleaved, there is a larger peak for C/D/E + TMSO at 306 and the subsequent loss of TMSOH giving the base peak of  $m/z$  216. For 15 $\beta$ -hydroxy- $\beta$ -amyrin the same fragmentation pattern is not observed. Instead, the loss of the ring D TMSOH is much faster in 15 $\beta$ -hydroxy- $\beta$ -amyrin than 16 $\beta$ -hydroxy- $\beta$ -amyrin due to the 1,3-diaxial interaction between the methyl CH<sub>3</sub> at C8 and the OTMS at C15. The retro Diels-Alder fragmentation is not observed in 15 $\beta$ -OH.

#### 4.3.11 Effects of NMR solvent on $^{13}\text{C}$ chemical shifts

We wanted to compare the chemical shifts from the isolated triterpenes to understand how new moieties effect the chemical shifts of neighboring groups and facilitate future spectral analysis of oleanane natural products. Most data sets for  $\beta$ -amyrin and its derivatives are collected in  $\text{CDCl}_3$  [255-257]. The original data set for the  $16\beta$  product from A354L, were collected in  $\text{CDCl}_3$  as well. A previous study by Agrawal reported that, general,  $^{13}\text{C}$  shifts were not very sensitive to different solvents, but there was a noticeable difference in the use of chloroform to pyridine [258]. Their study was conducted with steroids and they noticed an upfield shift (0.6-0.9 ppm) for C3 and downfield shift up to 0.9 ppm for the remaining carbons on a monohydroxylated steroid [258]. For our experiments, for  $16\beta$ -hydroxy- $\beta$ -amyrin, was assigned anew in pyridine- $d_5$  and data was acquired and analyzed for the reference compound  $\beta$ -amyrin in pyridine- $d_5$  as well (full assignment in Appendix). The chemical shifts were compared between  $\text{CDCl}_3$  and pyridine- $d_5$  (Table 13). The solvent effects that Agrawal reported for sterols were similar to the effects observed with triterpenes. Both  $16\beta$ -hydroxy- $\beta$ -amyrin and  $\beta$ -amyrin had an upfield shift of 0.6 ppm for C3 and a downfield shift for the other twenty-nine carbons. An upfield shift of 1.2 ppm was also observed for the second oxygen-bearing carbon in  $16\beta$ -hydroxy- $\beta$ -amyrin. Most shifts were less than 1 ppm difference, but there were a few exceptions in both  $16\beta$ -hydroxy- $\beta$ -amyrin and  $\beta$ -amyrin (highlighted in Table 13). The biggest effects are closest to the oxygen, where the  $\Delta\delta$  was 1 to 1.4. For example, in  $16\beta$ -hydroxy- $\beta$ -amyrin samples there was a downfield shift of 1.4 ppm observed at C13. Also, for  $\beta$ -amyrin, there was a shift difference of 1.7 ppm at C8. Since the solvent does affect the carbon chemical shifts, collecting a broader data set of triterpenes in pyridine- $d_5$  was undertaken in order to compare chemical shifts.

Table 13: Comparing the chemical shifts from pyridine-d<sub>5</sub> and CDCl<sub>3</sub>.

16 $\beta$ -Hydroxy- $\beta$ -amyrin				$\beta$ -Amyrin			
IUPAC	$\delta^{13}\text{C}$ CDCl <sub>3</sub>	$\delta^{13}\text{C}$ pyr d <sub>5</sub>	$\Delta\delta^{13}\text{C}$	IUPAC #	$\delta^{13}\text{C}$ CDCl <sub>3</sub>	$\delta^{13}\text{C}$ pyr d <sub>5</sub>	$\Delta\delta^{13}\text{C}$
1	38.7	39.6	0.9	1	38.7	39.5	0.8
2	27.3	28.6	1.3	2	27.3	28.5	1.2
3	79.1	78.5	-0.6	3	79	78.4	-0.6
4	38.9	39.9	1	4	38.8	39.7	0.9
5	55.3	56.2	0.9	5	55.3	56.1	0.8
6	18.4	19.3	0.9	6	18.5	19.1	0.6
7	32.7	33.5	0.8	7	32.8	33.4	0.6
8	40	40.7	0.7	8	38.8	40.5	1.7
9	46.9	47.8	0.9	9	47.7	48.4	0.7
10	37	37.7	0.7	10	37	37.6	0.6
11	23.6	24.4	0.8	11	23.6	24.2	0.6
12	122.5	122.9	0.4	12	121.8	122.7	0.9
13	143.6	145.0	1.4	13	145.1	145.6	0.5
14	43.9	44.4	0.5	14	41.8	42.3	0.5
15	35.7	37.0	1.3	15	26.2	26.5	0.3
16	66.1	64.9	-1.2	16	27	27.2	0.2
17	37.4	38.5	1.1	17	32.5	33.1	0.6
18	49.2	50.1	0.9	18	47.4	47.9	0.5
19	46.6	47.6	1	19	46.9	47.4	0.5
20	31	31.7	0.7	20	31.1	31.5	0.4
21	34.3	35.2	0.9	21	34.8	35.3	0.5
22	30.7	31.6	0.9	22	37.2	37.8	0.6
23	28.2	29.2	1	23	28.2	29.1	0.9
24	15.7	17.0	1.3	24	15.5	16.9	1.4
25	15.6	16.3	0.7	25	15.6	16.1	0.5
26	16.9	17.7	0.8	26	16.9	17.4	0.5
27	27.2	27.8	0.6	27	26	26.2	0.2
28	21.6	23.0	1.4	28	28.4	28.9	0.5
29	33.3	33.9	0.6	29	33.3	33.8	0.5
30	24.1	24.6	0.5	30	23.7	24.1	0.4

4.3.12 Comparison of <sup>13</sup>C shifts of oleanene triterpene in pyridine-d<sub>5</sub>

When comparing the chemical shifts, each chemical shift was compared to  $\beta$ -amyryn using the following formula.

$$\Delta\text{ppm} = \text{corresponding product (ppm)} - \beta\text{-amyryn (ppm)}$$

Shifts that had a significant downfield shift are highlighted in red and significant upfield shifts are highlighted in green with difference  $\pm 10$  shaded darker and differences that were  $\pm 5$  shaded lighter (Table 14).

Looking across the table, the largest chemical shift differences correspond to the sites directly perturbed addition of oxygen(s). The addition of the hydroxyl group occurs from the products of A354L ( $\Delta 37.7$  ppm), I471M-1 ( $\Delta 43.7$ ), and wild-type ( $\Delta 38.8$ ). The

addition of the electronegative group caused the attached carbon to be deshielded, which is seen by the positive change versus the unmodified alicyclic carbon. In olean-12-enes, like  $\beta$ -amyrin, the C12 and C13 alkene resonances are characteristically at 122 and 144 ppm, respectively [259]. There was a large increase in shielding at C12 and C13 when the double bond was oxidized to the epoxide in wild type product and I471M-3. Although oxygen is electronegative it does not deshield the carbon atoms compared to alkene so the chemical shift difference is large in the negative direction because of the hybridization change of the carbon from  $sp^2$  to  $sp^3$ . Also, the oxirane is similar to cyclopropane in that the planar three-membered ring creates an even greater shielding effect through magnetic anisotropy [260].

A large change also occurred when the alkene was oxidized to a ketone I471M-2. Though in this case, C12 has a large positive change ( $\Delta+89.5$  ppm) because the electrons are now in a polarized pi system from the  $sp^2$  hybridized carbon and now there are effectively two electron-withdrawing groups; therefore, the carbonyl group deshields C12. Magnetic anisotropy from the pi system from the carbonyl group creates shielding cones above and below the plane, which causes shielding at C18 and C19. However, C13 has a large negative change ( $\Delta-95$  ppm) from  $\beta$ -amyrin due to the change of hybridization state.

An inductive effect is seen by the addition of a  $16\beta$ -hydroxy group in the products from the A354L mutant and the wild-type enzyme. C15 and C17 alpha to C16 are shifted downfield by 5.4 to 10.5 ppm due to the deshielding by the electron-withdrawing group, -OH. When C16 is hydroxylated, there are noticeable upfield shifts at  $\gamma$ -C22 and  $\gamma$ -C28. In a previous study, Lambert studied the shielding effects the relationship between the gamma carbon and the shielding effect from the dihedral angle [261]. The  $\gamma$  effect is important because it is the simplest shielding effect that is dependent on the dihedral angle and it usually has the opposite sign from the  $\alpha$  and  $\beta$  effects [261]. We observe the opposite effect at  $\gamma$ -C22 and  $\gamma$ -C28 compared to  $\alpha$ -C16 or  $\beta$ -C15 and  $\beta$ -C17. In the report by Lambert, they reported that substituting a hydroxyl group at the equatorial position on a cyclohexane would cause the  $\gamma$ -carbons with a  $\phi = 60-90^\circ$  to be shielded by -4.10 and -6.31 ppm. Our results for C22 and C28 follow a similar pattern and are shielded by -6.2 ppm and -5.9 ppm.

15 $\beta$ -Hydroxy- $\beta$ -amryin and 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amryin had the largest number of carbons affected relative to unmodified  $\beta$ -amryin, even though 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amryin had an addition of two functional groups and 15 $\beta$ -hydroxy- $\beta$ -amryin only had one. For 15 $\beta$ -hydroxy- $\beta$ -amryin, C7, C11, C13, C15, C26, C30 had a  $>5$  shift downfield and C18 and C19 had a  $>5$  effect upfield. For the wild-type product, had a  $>5$  shift difference downfield at C15, C16, and, C17 and a  $>5$  effect upfield at C12, C13, C19, C22, and C28. One of the reasons for the number of carbons effected in 15 $\beta$ -hydroxy- $\beta$ -amryin could be due to the fact the hydroxyl group is in an axial position potentially destabilizing the chair conformation of ring D, leading to unexpected shielding effects. There are additional interactions attributable to the axial OH group. The hydroxyl group at C15 has an 1,3-diaxial interaction with methyl C26, a significant effect in conformationally restricted cyclohexanes, appears to cause steric deshielding of the  $^{13}\text{C}$  registered by a downfield shift of 9 ppm at methyl C26 [259]. Another effect, shifts C7 downfield with the addition of the axial hydroxyl group at C15 because the gauche interaction between C15 and C7 [259].

The 1,3-diaxial interaction between the OH at C15 and C28 would be expected to cause a significant downfield shift of methyl C28. However, the lack of this influence gives insight into the D/E ring fusion for 15 $\beta$ -hydroxy- $\beta$ -amryin. C28 appears not to be in an axial position due to the lack of deshielding from a 1,3-diaxial interaction. The *cis*-ring fusion of ring D/E allows for some flexibility in the molecular configuration. The flip of ring E also explains the shift differences at C18 ( $\Delta$ -8.7 ppm), C19 ( $\Delta$ -10.5 ppm), and C30 ( $\Delta$ 6.6 ppm). C18 and C19 become more shielded from the magnetic anisotropy-shielding cone of the alkene [260]. C30 flips from equatorial to an axial position where it experiences a 1,3-diaxial interaction from H22 that deshields C30 causing a downfield shifts [260]. Product **90** was the only product to undergo a ring flip at ring E.

Another interesting observation is the carbons with the largest shift difference in 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amryin were the same compared to 16 $\beta$ -hydroxy- $\beta$ -amryin and 12,13 $\beta$ -epoxy- $\beta$ -amryin. The epoxide compound has upfield shifts at C12 (-67.7 ppm), C13 (-78.2 ppm), and C19 (-5.7 ppm). In epoxy-diol compound there are upfield shifts observed for C12 (-68.2 ppm), C13 (-78.2 ppm), and C19 (-5.5). For 16 $\beta$ -hydroxy- $\beta$ -amryin, downfield shifts occur at C15 (10.5 ppm), C16 (37.7 ppm), and C17 (5.4 ppm).

There are upfield shifts at C22 (-6.2 ppm) and C28 (-5.9 ppm). For the epoxy-diol, the effects at those same carbons were downfield for C15 (8.7 ppm), C16 (38.8 ppm), and C17 (6.5 ppm) and upfield C22 (-6.0) and C28 (-5.9 ppm). The effects from 16 $\beta$ -hydroxy- $\beta$ -amyirin and 12,13 $\beta$ -epoxy- $\beta$ -amyirin are additive effects in 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amyirin.

Table 14: Comparative table of chemical shifts of  $\beta$ -amyrin derivatives relative to  $\beta$ -amyrin. All data collected in pyridine- $d_5$ . Downfield shifts greater than 10 ppm are denoted in dark red and downfield shifts of 0.1-5 are printed in light red. Upfield shifts greater than 10 ppm are shown in dark green and upfield shifts of 0.1-5 ppm are printed in light green.

IUPAC carbon	$\beta$ -amyrin		12, 13 $\beta$ -epoxy-16 $\beta$ -Hydroxy- $\beta$ -amyrin		16 $\beta$ -Hydroxy- $\beta$ -amyrin		$\Delta\Delta$ 354L
	$\delta^{13}C$	$\delta^{13}C$	$\delta^{13}C$	$\Delta\delta$	$\delta^{13}C$	$\Delta\delta$	
1	39.5	38.8	-0.7		39.6	0.1	
2	28.5	28.4	-0.1		28.6	0.1	
3	78.4	78.3	-0.1		78.5	0.1	
4	39.7	39.8	0.1		39.9	0.2	
5	56.1	56.2	0.1		56.2	0.1	
6	19.1	18.6	-0.5		19.3	0.2	
7	33.4	35.0	1.6		33.5	0.1	
8	40.5	42.1	1.6		40.7	0.2	
9	48.4	46.5	-1.9		47.8	-0.6	
10	37.6	37.8	0.2		37.7	0.1	
11	24.2	22.5	-1.7		24.4	0.2	
12	122.7	54.5	-68.2		122.9	0.2	
13	145.6	67.4	-78.2		145.0	-0.6	
14	42.3	42.9	0.6		44.4	2.1	
15	26.5	35.2	8.7		37.0	10.5	
16	27.2	66.0	38.8		64.9	37.7	
17	33.1	39.6	6.5		38.5	5.4	
18	47.9	50.3	2.4		50.1	2.2	
19	47.4	41.9	-5.5		47.6	0.2	
20	31.5	31.2	-0.3		31.7	0.2	
21	35.3	35.0	-0.3		35.2	-0.1	
22	37.8	31.8	-6		31.6	-6.2	
23	29.1	29.0	-0.1		29.2	0.1	
24	16.9	16.6	-0.3		17.0	0.1	
25	16.1	15.7	-0.4		16.3	0.2	
26	17.4	21.2	3.8		17.7	0.3	
27	26.2	23.9	-2.3		27.8	1.6	
28	28.9	23.0	-5.9		23.0	-5.9	
29	33.8	33.7	-0.1		33.9	0.1	
30	24.1	24.1	0		24.6	0.5	



Table 14 continued'

pyr-d5 IUPAC carbon	β-amyrin		15β-Hydroxy-β-amyrin		12-keto-13β-amyrin		12, 13β-epoxy-β-amyrin	
	δ <sup>13</sup> C	ΔH7IM-1	δ <sup>13</sup> C	ΔH7IM-1	δ <sup>13</sup> C	ΔH7IM-2	δ <sup>13</sup> C	ΔH7IM-3
1	39.5	-0.6	38.9	-0.6	38.7	-0.8	38.8	-0.7
2	28.5	0.1	28.6	0.1	28.3	-0.2	28.3	-0.2
3	78.4	0.2	78.6	0.2	78.1	-0.3	78.3	-0.1
4	39.7	0.3	40.0	0.3	39.8	0.1	39.7	0
5	56.1	0.7	56.8	0.7	55.8	-0.3	56.2	0.1
6	19.1	0.9	20.0	0.9	19.1	0	18.6	-0.5
7	33.4	10.4	43.8	10.4	32.5	-0.9	34.8	1.4
8	40.5	-0.6	39.9	-0.6	42.2	1.7	41.8*	1.3
9	48.4	2	50.4	2	50.4	2	47.1	-1.3
10	37.6	1	38.6	1	37.5	-0.1	37.9	0.3
11	24.2	14.2	38.4	14.2	39.2	15	22.3	-1.9
12	122.7	-3.9	118.8	-3.9	212.2	89.5	55.0	-67.7
13	145.6	13	158.6	13	50.6	-95	67.4	-78.2
14	42.3	0.9	43.2	0.9	42.7	0.4	40.9*	-1.4
15	26.5	43.7	70.2	43.7	26.6	0.1	25.1	-1.4
16	27.2	3	30.2	3	27.0	-0.2	28.3	1.1
17	33.1	2.8	35.9	2.8	33.2	0.1	34.0	0.9
18	47.9	-8.7	39.2	-8.7	37.8	-10.1	48.9	1
19	47.4	-10.5	36.9	-10.5	35.5	-11.9	41.7	-5.7
20	31.5	-	-	-	31.6	0.1	31.2	-0.3
21	35.3	-1.2	34.1	-1.2	37.7	2.4	35.1	-0.2
22	37.8	-1.7	36.1	-1.7	38.3	0.5	37.7	-0.1
23	29.1	0	29.1	0	28.9	-0.2	28.9	-0.2
24	16.9	-0.1	16.8	-0.1	16.6	-0.3	16.6	-0.3
25	16.1	0.6	16.7	0.6	15.8	-0.3	15.7	-0.4
26	17.4	9	26.4	9	16.4	-1	21.1	3.7
27	26.2	-4	22.2	-4	21.4	-4.8	22.4	-3.8
28	28.9	1	29.9	1	27.7	-1.2	29.1	0.2
29	33.8	-0.2	33.6	-0.2	34.2	0.4	33.6	-0.2
30	24.1	6.6	30.7	6.6	23.8	-0.3	23.7	-0.4

Downfield Shift Δ &gt; |10|

Downfield Shift Δ &gt; |5|

Upfield Shift Δ &gt; |10|

Upfield Shift Δ &gt; |5|

#### 4.3.13 Correlation of methyl groups

Pentacyclic triterpenes contain 8 methyl groups, which are distinctive in the proton spectrum by the intense singlet peaks. Establishing the identity of methyl groups could be beneficial and useful for studying new triterpenes. Several have tried to make this correlation, but assumptions were always made for methyls C23, C24, and C25 until a study by Ito that analyzed the regular changes that occur when 3 $\beta$ -hydroxy oleanenes are converted to 3 $\beta$ -acetoxy derivatives [262-265]. The areas they saw the most difference was from 0.78-0.99 ppm [265]. In order to unambiguously assign these three groups, 3 $\beta$ -hydroxyoleanenes, in the camelliagenin and barringtogenol families were selectively dideuterated at either C23 or C24. They observed the methyl  $^1\text{H}$  NMR signal intensity at 0.99-1.01 decreased for C23- $\text{d}_2$  and signals at 0.80 ppm became less pronounced for C24- $\text{d}_2$ , thereby deducing the signal at 0.91-0.94 to be C25.

Looking at the assigned methyl groups for  $\beta$ -amyrin and the five products in Table 15, the trend for our data is the same as Ito, but the range is bigger. For C23, the average methyl peak was around 1.26 ppm ranging from 1.23-1.28 ppm. For C24, the average methyl peak was found at 1.08 ppm ranging from 1.05-1.10. For C25, the average methyl peak was found at 0.97 ppm ranging from 0.88 – 1.07 ppm. Therefore, with Ito's results one might be able to apply their results to derivatives of camelliagenin and barringtogenol, but not across all pentacyclic triterpenes. Further, when new functionality is added to rings C and D, methyl groups at carbons 23, 24, 25, and 29 had <0.1 ppm change, except for the 12-keto-13 $\beta$ - $\beta$ -amyrin effected methyl-25 with an upfield shift of 0.12 ppm.

The other four methyls -26, -27, -28, and -30 were affected to different degrees depending on the substitution. Both hydroxylated compounds 15 $\beta$ -hydroxy- $\beta$ -amyrin and 16 $\beta$ -hydroxy- $\beta$ -amyrin remove electron density from and thus deshielded the C26, C27, and C28 methyl groups. 16 $\beta$ -Hydroxy- $\beta$ -amyrin had a <0.1 ppm effect on C26 whereas 15 $\beta$ -hydroxy- $\beta$ -amyrin had a large chemical shift difference of 0.70 ppm for C26. This can be attributed to the 1,3-diaxial interaction discussed previously. Superimposing the effects of 16 $\beta$ -hydroxy- $\beta$ -amyrin on the NMR chemical shift data for 12,13- $\beta$ -epoxy- $\beta$ -amyrin allows the prediction of the methyls -26, -27, and -28 chemical shifts observed in the wild-type product. For example, the addition of the effects from 16 $\beta$ -hydroxy- $\beta$ -

amyrin to those of 12,13 $\beta$ -epoxy- $\beta$ -amyrin, C26 = 0.36, C27 = 0.14, and C28 = 0.54 ppm gives almost the exact shift differences from  $\beta$ -amyrin for 16 $\beta$ -hydroxy-12,13 $\beta$ -epoxy- $\beta$ -amyrin C26 = 0.31, C27 = 0.11, and C28 = 0.55 ppm. This additivity effect could be useful in predicting chemical shifts with more than one substitution on  $\beta$ -amyrin derivative. Though correlations can be made from the different  $\beta$ -amyrin derivatives, the methyl groups alone can only aid in determining an unknown structure. Full assignment must be completed using several NMR experiments, IR, and GC-MS analysis to achieve a high confidence structure.

Table 15: Methyl shifts of five oleanane triterpene derivatives compared to  $\beta$ -amyrin in pyridine- $d_5$ . Downfield shifts  $> 0.1$  ppm are in highlighted in yellow and upfield shifts  $> 0.1$  ppm are highlighted in purple.

Methyls	$\beta$ -amyrin	12, 13 $\beta$ -epoxy-16 $\beta$ -Hydroxy- $\beta$ -amyrin		16 $\beta$ -Hydroxy- $\beta$ -amyrin	
	$\delta^1\text{H}$	$\delta^1\text{H}$	$\Delta\text{WT}$	$\delta^1\text{H}$	$\Delta\Delta 354\text{L}$
23	1.28	1.23	-0.05	1.27	-0.01
24	1.10	1.05	-0.05	1.08	-0.02
25	1.00	0.91	-0.09	0.99	-0.01
26	1.04	1.35	0.31	1.11	0.07
27	1.25	1.36	0.11	1.40	0.15
28	0.95	1.50	0.55	1.19	0.24
29	0.94	0.95	0.01	0.95	0.01
30	0.93	0.86	-0.07	1.01	0.08

Downfield Shift  $\Delta > 0.1$   
Upfield Shift  $\Delta > 0.1$

Methyls	$\beta$ -amyrin	15 $\beta$ -Hydroxy- $\beta$ -amyrin		12-keto-13 $\beta$ - $\beta$ -amyrin		12, 13 $\beta$ -epoxy- $\beta$ -amyrin	
	$\delta^1\text{H}$	$\delta^1\text{H}$	$\Delta 1471\text{M-1}$	$\delta^1\text{H}$	$\Delta 1471\text{M-2}$	$\delta^1\text{H}$	$\Delta 1471\text{M-3}$
23	1.28	1.27	-0.01	1.25	-0.03	1.25	-0.03
24	1.10	1.10	0.00	1.07	-0.03	1.07	-0.03
25	1.00	1.07	0.07	0.88	-0.12	0.95	-0.05
26	1.04	1.74	0.70	1.10	0.06	1.32	0.28
27	1.25	1.08	-0.17	1.00	-0.25	1.24	-0.01
28	0.95	1.09	0.14	0.93	-0.02	1.25	0.30
29	0.94	1.03	0.09	0.96	0.02	0.96	0.02
30	0.93	1.04	0.11	1.04	0.11	0.81	-0.12

#### 4.3.14 Future Work

Both proposed Sad2 intermediates, 12,13 $\beta$ -epoxy- $\beta$ -amyrin and 16 $\beta$ -hydroxy- $\beta$ -amyrin, were isolated from two different mutants. 16 $\beta$ -Hydroxy- $\beta$ -amyrin was isolated from the mutant A354L from the substrate recognition site SRS5 and 12,13 $\beta$ -epoxy- $\beta$ -amyrin was isolated from I471M from the substrate recognition site SRS6. Both substrate recognition sites were predicted from computational modeling to play a catalytic role in the conversion of  $\beta$ -amyrin to 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin. The compounds will need to be fed to the wild type enzyme in yeast in order to determine if 12,13 $\beta$ -epoxy- $\beta$ -amyrin can be converted to 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin or if 16 $\beta$ -hydroxy- $\beta$ -amyrin can be converted to 12,13 $\beta$ -epoxy-16 $\beta$ -hydroxy- $\beta$ -amyrin. This part of the project is still in progress. Lastly, the desmethylation product was never identified. At this time, we do not have further evidence to support this hypothesis. The identification of the desmethylation product would give concrete evidence that the divergent enzyme *AspCyp51H10* retained a native plant CYP51G activity.

## REFERENCES

1. Rizhsky, L., et al., *Integrating metabolomics and transcriptomics data to discover a biocatalyst that can generate the amine precursors for alkalamide biosynthesis*. Plant J., 2016. **88**(5): p. 775-793.
2. Pichersky, E. and E. Lewinsohn, *Convergent evolution in plant specialized metabolism*. Annu. Rev. Plant Biol., 2011. **62**: p. 549-566.
3. Hartmann, T., *From waste products to ecochemicals: fifty years research of plant secondary metabolism*. Phytochemistry 2007. **68**(22-24): p. 2831-2846.
4. Stahl, E., *Pflanzen und Schnecken Biologische Studien uber Schutzmittel der Pflanzen gegen Schneckenfrass*. Jenaer Zeitschrift fur Medizin und Naturwissenschaften. Vol. 22. 1888.
5. Camara-Leret, R., et al., *Fundamental species traits explain provisioning services of tropical American palms*. Nat. Plants, 2017. **3**: p. 16220.
6. Cox, M.M., Nelson, D. L., Lehninger, A.L., *Lehninger Principles of Biochemistry*. 2008, New York: W.H. Freeman.
7. Hedden, P., *Primary and Secondary Metabolism of Plants and Plant Cell Cultures III edited by J. Schripsema and R. Verpoorte*. Plant Growth Regul., 1996. **19**(3): p. 267.
8. *Primary and Secondary Metabolites*. 2016 [cited 2017; Available from: <https://www.boundless.com/microbiology/textbooks/boundless-microbiology-textbook/industrial-microbiology-17/industrial-microbiology-198/primary-and-secondary-metabolites-999-5345/>].
9. *Secondary Metabolites in Plants*. [cited 2017 2017]; Available from: <http://www.biologyreference.com/Re-Se/Secondary-Metabolites-in-Plants.html>.
10. Hamberger, B. and S. Bak, *Plant P450s as versatile drivers for evolution of species-specific chemical diversity*. Philos. Trans. R. Soc., B, 2013. **368**(1612): p. 20120426/1-20120426/16.
11. Reznik, H., *Vergleichende Biochemie der Phenylpropane*, in *Ergebnisse der Biologie*, H. Autrum, et al., Editors. 1960, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 14-46.
12. Rosenthal, G.A., D.H. Janzen, and Editors, *Herbivores: Their Interaction with Secondary Plant Metabolites*. 1979, New York, London: Academic Press. 718 pp.
13. Harborne, J.B., *Phytochemical Ecology*. 1972, Academic Press: London, New York.
14. Dudareva, N. and E. Pichersky, *Floral scent metabolic pathways: their regulation and evolution*. 2006, Boca, Raton, FL: CRC Press LLC.
15. War, A.R., et al., *Mechanisms of plant defense against insect herbivores*. Plant Signal Behav., 2012. **7**(10): p. 1306-20.
16. Robinson, T., *The Biochemistry of Alkaloids: Molecular Biology, Biochemistry, and Biophysics*. Vol. 3. 1968, New York: Springer-Verlag. 149 pp.
17. Marston, A., *Role of advances in chromatographic techniques in phytochemistry*. Phytochemistry 2007. **68**(22-24): p. 2786-2798.

18. Eisenreich, W. and A. Bacher, *Advances of high-resolution NMR techniques in the structural and metabolic analysis of plant biochemistry*. *Phytochemistry* 2007. **68**(22-24): p. 2799-2815.
19. Ruzicka, L., *Isoprene rule and the biogenesis of terpenic compounds*. *Experientia*, 1953. **9**: p. 357-67.
20. Loomis, W.D., *Biosynthesis and metabolism of monoterpenes*. *International Flavours and Food Additives*, 1975. **6**: p. 292-296.
21. Schwartz, M.K., *Terpen-Biosynthese in Ginkgo biloba: eine überraschende Geschichte*. PhD Thesis. 1994, Zürich, Switzerland.
22. Schwarz, M. and D. Arigoni, *Ginkgolide biosynthesis*. Vol. 2. 1999, Zurich, Switz: Elsevier Science B.V. 367-400.
23. Koukol, J. and E.E. Conn, *Metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of Hordeum vulgare*. *J. Biol. Chem.*, 1961. **236**: p. 2962-8.
24. Hahlbrock, K., *Flavonoids*. *Biochem. Plants*, 1981. **7**: p. 425-56.
25. Grisebach, H., *Lignins*. *Biochem. Plants*, 1981. **7**: p. 457-78.
26. Facchini, P.J., *Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications*. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 2001. **52**: p. 29-66.
27. Kuhn, D.N., et al., *Induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor*. *Proc. Natl. Acad. Sci. U. S. A.*, 1984. **81**(4): p. 1102-6.
28. Ryder, T.B., et al., *Elicitor rapidly induces chalcone synthase mRNA in Phaseolus vulgaris cells at the onset of the phytoalexin defense response*. *Proc. Natl. Acad. Sci. U. S. A.*, 1984. **81**(18): p. 5724-8.
29. Edwards, K., et al., *Rapid transient induction of phenylalanine ammonia-lyase mRNA in elicitor-treated bean cells*. *Proc. Natl. Acad. Sci. U. S. A.*, 1985. **82**(20): p. 6731-5.
30. Newman, D.J. and G.M. Cragg, *Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010*. *J. Nat. Prod.*, 2012. **75**(3): p. 311-335.
31. Luo, P., Z.Q. Lan, and Z.Y. Li, *Orychopragmus violaceus, a potential edible-oil crop*. *Plant Breeding*, 1994. **113**(1): p. 83-85.
32. Zhou, L., J. Wu, and S. Wang, *Orychopragmus*, in *Wild Crop Relatives: Genomic and Breeding Resources*, C. Kole, Editor. 2011, Springer Berlin Heidelberg. p. 199-225.
33. Li, Z.-y. and X.-h. Ge, *Unique chromosome behavior and genetic control in Brassica×Orychopragmus wide hybrids: a review*. *Plant Cell Reports*, 2007. **26**(6): p. 701-710.
34. Zhao, Z.-g., et al., *Production and characterization of intergeneric somatic hybrids between Brassica napus and Orychopragmus violaceus and their backcrossing progenies*. *Plant Cell Reports*, 2008. **27**(10): p. 1611-1621.
35. Xu, C. and Z. Li, *Seed quality and genetic analysis of F<sub>12</sub> progenies from intergeneric hybrids between Brassica napus and Orychopragmus violaceus*. *Chinese Journal of Oil Crop Sciences*, 2011. **33**: p. 20-24.

36. Wang, R., et al., *Orychopragmus violaceus L., a marginal land-based plant for biodiesel feedstock: Heterogeneous catalysis, fuel properties, and potential*. Energy Conversion and Management, 2014. **84**: p. 497-502.
37. Harwood, J.L., *Recent advances in the biosynthesis of plant fatty acids*. Biochim. Biophys. Acta, Lipids Lipid Metab., 1996. **1301**(1/2): p. 7-56.
38. Herbert, R.B., *The Biosynthesis of Secondary Metabolites. 2nd Ed.* 1989: Routledge, Chapman & Hall. 224 pp.
39. Brown, A.P., A.R. Slabas, and J.B. Rafferty, *Fatty acid biosynthesis in plant - metabolic pathways, structure and organization*. Adv. Photosynth. Respir., 2009. **30**(Lipids in Photosynthesis): p. 11-34.
40. Harwood, J.L. *AOCS Lipid Library*. 2010 [cited 2017; Available from: <http://lipidlibrary.aocs.org/Biochemistry/content.cfm?ItemNumber=40304>].
41. Jenke-Kodama, H., et al., *Evolutionary implications of bacterial polyketide synthases*. Mol. Biol. Evol., 2005. **22**(10): p. 2027-2039.
42. Kohli, G.S., et al., *Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes*. ISME J, 2016. **10**(8): p. 1877-90.
43. Smith, S. and S.-C. Tsai, *The type I fatty acid and polyketide synthases: a tale of two megasynthases*. Nat. Prod. Rep., 2007. **24**(5): p. 1041-1072.
44. Smith, S., A. Witkowski, and A.K. Joshi, *Structural and functional organization of the animal fatty acid synthase*. Prog. Lipid Res., 2003. **42**(4): p. 289-317.
45. Hertweck, C., *The Biosynthetic Logic of Polyketide Diversity*. Angew. Chem., Int. Ed., 2009. **48**(26): p. 4688-4716.
46. Gallo, A., M. Ferrara, and G. Perrone, *Phylogenetic study of polyketide synthases and nonribosomal peptide synthetases involved in the biosynthesis of mycotoxins*. Toxins, 2013. **5**: p. 717-742.
47. El Tahchy, A., et al., *Thioesterase overexpression in Nicotiana benthamiana leaf increases the fatty acid flux into triacylglycerol*. FEBS Lett., 2017. **591**(2): p. 448-456.
48. Gunstone, F.D., Harwood, J. L., and Dijkstra, A. J., *The Lipid Handbook*. 3rd ed. 2007, BocaRaton, Florida: CRC Press.
49. Kishino, S., et al., *Microbial production of conjugated fatty acids*. Lipid Technol., 2009. **21**(8/9): p. 177-181.
50. Cahoon, E.B., et al., *Biosynthetic origin of conjugated double bonds: production of fatty acid components of high-value drying oils in transgenic soybean embryos*. Proc. Natl. Acad. Sci. U. S. A., 1999. **96**(22): p. 12935-12940.
51. Cahoon, E.B., et al., *Transgenic production of epoxy fatty acids by expression of a cytochrome P450 enzyme from Euphorbia lagascae seed*. Plant Physiol., 2002. **128**(2): p. 615-624.
52. Dyk, M.S.V., J.L.F. Kock, and A. Botha, *Hydroxy long-chain fatty acids in fungi*. World J. Microbiol. Biotechnol., 1994. **10**(5): p. 499-504.
53. Hannemann, K., et al., *The common occurrence of furan fatty acids in plants*. Lipids, 1989. **24**(4): p. 296-8.
54. Dembitsky, V.M. and M. Srebnik, *Natural halogenated fatty acids: their analogues and derivatives*. Prog. Lipid Res., 2002. **41**(4): p. 315-367.
55. Greenberg, A. and J. Harris, *Cyclopropenoid fatty acids*. J. Chem. Educ., 1982. **59**(7): p. 539-43.



56. Kopitz, J. *Glycolipids: structure and function*. 1997. Chapman & Hall.
57. Ansell, G.B., et al., *B.B.A. Library, Vol. 3: Form and Function of Phospholipids. 2nd ed.* 1973, New York, NY: Elsevier. 494 pp.
58. Gault, C.R., L.M. Obeid, and Y.A. Hannun, *An overview of sphingolipid metabolism: from synthesis to breakdown*. Adv. Exp. Med. Biol., 2010. **688**(Sphingolipids as Signaling and Regulatory Molecules): p. 1-23.
59. Ibarguren, M., D.J. Lopez, and P.V. Escriba, *The effect of natural and synthetic fatty acids on membrane structure, microdomain organization, cellular functions and human health*. Biochim. Biophys. Acta, Biomembr., 2014. **1838**(6): p. 1518-1528.
60. Farmer, E.E. and C.A. Ryan, *Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves*. Proc. Natl. Acad. Sci. U. S. A., 1990. **87**(19): p. 7713-16.
61. Cahoon, E.B., et al., *Modifying Vegetable Oils for Food and Non-food Purposes*, in *Oil Crops*, J. Vollmann and I. Rajcan, Editors. 2010, Springer New York: New York, NY. p. 31-56.
62. Greenspan, F.P. and R.J. Gall, *Epoxy fatty acid ester plasticizers. Preparation and properties*. J. Am. Oil Chem. Soc., 1956. **33**: p. 391-4.
63. Kulkarni, A.S., et al., *Chemical composition of some potential paint oils of Central India*. Paintindia, 1992. **42**(1): p. 58-9.
64. Fahy, E., et al., *A comprehensive classification system for lipids. [Erratum to document cited in CA143:320906]*. J. Lipid Res., 2010. **51**(6): p. 1618.
65. Kim, K.-R. and D.-K. Oh, *Production of hydroxy fatty acids by microbial fatty acid-hydroxylation enzymes*. Biotechnol. Adv., 2013. **31**(8): p. 1473-1485.
66. Mutlu, H. and M.A.R. Meier, *Castor oil as a renewable resource for the chemical industry*. Eur. J. Lipid Sci. Technol., 2009. **112**(1): p. 10-30.
67. Ogunniyi, D.S., *Castor oil: A vital industrial raw material*. Bioresource Technology, 2006. **97**(9): p. 1086-1091.
68. Kinney, A.J. and T.E. Clemente, *Modifying soybean oil for enhanced performance in biodiesel blends*. Fuel Process. Technol., 2005. **86**(10): p. 1137-1147.
69. Smith, M., H. Moon, and L. Kunst, *Production of hydroxy fatty acids in the seeds of Arabidopsis thaliana*. Biochem. Soc. Trans., 2000. **28**(6): p. 947-950.
70. Smith, C.R., Jr., et al., *Densipolic acid: a unique hydroxydienoic acid from Lesquerella densipila seed oil*. J. Org. Chem., 1962. **27**: p. 3112-17.
71. Blondin, G.A., *Isolation, properties, and structural features of divalent cation ionophores derived from beef heart mitochondria*. Ann. N. Y. Acad. Sci., 1975. **264**: p. 98-111.
72. Chisholm, M.J. and C.Y. Hopkins, *Kamlolenic acid and other conjugated fatty acids in certain seed oils*. J. Am. Oil Chem. Soc., 1966. **43**(6): p. 390-2.
73. James, A.T., H.C. Hadaway, and J.P.W. Webb, *The biosynthesis of ricinoleic acid*. Biochem. J., 1965. **95**: p. 448-52.
74. Bafor, M., et al., *Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor-bean (Ricinus communis) endosperm*. Biochem. J., 1991. **280**(2): p. 507-14.

75. van de Loo, F.J., et al., *An oleate 12-hydroxylase from Ricinus communis L. is a fatty acyl desaturase homolog*. Proc. Natl. Acad. Sci. U. S. A., 1995. **92**(15): p. 6743-7.
76. Broun, P., S. Boddupalli, and C. Somerville, *A bifunctional oleate 12-hydroxylase: desaturase from Lesquerella fendleri*. Plant J., 1998. **13**(2): p. 201-210.
77. Appelqvist, L.-A., *Lipids in Cruciferae I. Fatty acid composition in seeds of some Svalöf varieties and strains of rape, turnip rape, white mustard and false flax*. Acta Agriculturae Scandinavica, 1968. **18**(1-2): p. 3-21.
78. Miller, R., et al., *Search for new industrial oils. XIII. Oils from 102 species of cruciferae*. Journal of the American Oil Chemists' Society, 1965. **42**(10): p. 817-821.
79. Earle, F.R., et al., *Compositional differences among crambe samples and between seed components*. Journal of the American Oil Chemists Society, 1966. **43**(5): p. 330-333.
80. Miwa, T.K. and I.A. Wolff, *Fatty acids, fatty alcohols, wax esters, and methyl esters from Crambe abyssinica and Lunaria annua seed oils*. Journal of the American Oil Chemists Society, 1963. **40**(12): p. 742-744.
81. Jart, A., *The fatty acid composition of various cruciferous seeds*. Journal of the American Oil Chemists' Society, 1978. **55**(12): p. 873-875.
82. Dyer, J.M., et al., *High-value oils from plants*. Plant J, 2008. **54**(4): p. 640-55.
83. Mikolajczak, K.L., F.R. Earle, and I.A. Wolff, *Search for new industrial oils. VI. Seed oils of the genus Lesquerella*. Journal of the American Oil Chemists Society, 1962. **39**(2): p. 78-80.
84. Salywon, A.M., et al., *Evaluation of new Lesquerella and Physaria (Brassicaceae) oilseed germplasm*. Am J Bot, 2005. **92**(1): p. 53-62.
85. Badami, R.C. and K.B. Patil, *Structure and occurrence of unusual fatty acids in minor seed oils*. Progress in Lipid Research, 1980. **19**(3-4): p. 119-153.
86. Smith, C.R., *Occurrence of unusual fatty acids in plants*. Progress in the Chemistry of Fats and other Lipids, 1971. **11**: p. 137-177.
87. Cahoon, E.B. and A.J. Kinney, *The production of vegetable oils with novel properties: Using genomic tools to probe and manipulate plant fatty acid metabolism*. European Journal of Lipid Science and Technology, 2005. **107**(4): p. 239-243.
88. Broun, P., et al., *Catalytic plasticity of fatty acid modification enzymes underlying chemical diversity of plant lipids*. Science, 1998. **282**(5392): p. 1315-7.
89. Cahoon, E.B., et al., *Redesign of soluble fatty acid desaturases from plants for altered substrate specificity and double bond position*. Proc Natl Acad Sci U S A, 1997. **94**(10): p. 4872-7.
90. Reed, D.W., et al., *Mechanistic study of an improbable reaction: alkene dehydrogenation by the delta12 acetylenase of Crepis alpina*. J Am Chem Soc, 2003. **125**(35): p. 10635-40.
91. Boeckman, R.K., P. Shao, and J.J. Mullins, *The Dess-Martin periodinane: 1,1,1-triacetoxy-1,1-dihydro-1,2-benziodoxol-3(1H)-one*. Org. Syn., 2000. **77**: p. 141.

92. Donohoe, T.J., et al., *Scope of the directed dihydroxylation: application to cyclic homoallylic alcohols and trihaloacetamides*. *Org. Biomol. Chem.*, 2003. **1**: p. 2173-2186.
93. Brown, H.C., *Hydroboration*. 76. *Hydroboration of cyclic dienes with representative hydroborating agents*. *J. Org. Chem.*, 1986. **51**: p. 445-449.
94. Kai, K., et al., *Structure-activity relationship study of flowering-inducer FN against Lemna paucicostata*. *Tetrahedron*, 2008. **64**(28): p. 6760-6769.
95. Yun, J.I., et al., *Cross-metathesis of allyl halides with olefins bearing amide and ester groups*. *Tetrahedron*, 2012. **68**(4): p. 1177-1184.
96. Kang, B. and S. Chang, *A facile synthetic route to (+)-allosedamine via hydrolytic kinetic resolution and olefin metathesis*. *Tetrahedron*, 2004. **60**(34): p. 7353-7359.
97. Nomura, M. and Y. Fujihara, *Studies on the synthesis of physiologically active substances. IV. Synthesis of trans-2- and trans-3-C10 unsaturated carboxylic acids as a sex pheromone of Apis mellifera*. *Yukagaku*, 1988. **37**(6): p. 453-7.
98. Aiguabella, N., et al., *Pauson-Khand Adducts of N-Boc-propargylamine: A New Approach to 4,5-Disubstituted Cyclopentenones*. *Org. Lett.*, 2013. **15**(11): p. 2696-2699.
99. Carling, R.W., J.S. Clark, and A.B. Holmes, *Synthesis of medium ring ethers. Part 2. Synthesis of the fully saturated carbon skeleton of Laurencia non-terpenoid ether metabolites containing seven-, eight- and nine-membered rings*. *J. Chem. Soc., Perkin Trans. 1*, 1992(1): p. 83-94.
100. Hoye, T.R., C.S. Jeffrey, and F. Shao, *Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons*. *Nat. Protoc.*, 2007. **2**(10): p. 2451-2458.
101. Kwon, Y., et al., *Simple determination of double-bond positions in long-chain olefins by cross-metathesis*. *Angew. Chem., Int. Ed. Engl.*, 2011. **50**: p. 8275-8278.
102. Molander, G.A. and K.L. Bobbitt, *Trimethylsilyl trifluoromethanesulfonate mediated dialkylcuprate addition to epoxy esters: an unusual intramolecular transesterification process*. *J. Org. Chem.*, 1992. **57**(18): p. 5031-4.
103. Ohtani, I., et al., *High-field FT NMR application of Mosher's method. The absolute configurations of marine terpenoids*. *J. Am. Chem. Soc.*, 1991. **113**(11): p. 4092-6.
104. Rieser, M.J., et al., *Determination of absolute configuration of stereogenic carbinol centers in annonaceous acetogenins by proton and fluorine 19-NMR analysis of Mosher ester derivatives*. *J. Am. Chem. Soc.*, 1992. **114**(26): p. 10203-13.
105. Dale, J.A. and H.S. Mosher, *Nuclear magnetic resonance enantiomer reagents. Configurational correlations via nuclear magnetic resonance chemical shifts of diastereomeric mandelate, O-methylmandelate, and  $\alpha$ -methoxy- $\alpha$ -trifluoromethylphenylacetate (MTPA) esters*. *J. Amer. Chem. Soc.*, 1973. **95**(2): p. 512-19.
106. Pauli, G.F., et al., *Essential parameters for structural analysis and dereplication by (1)H NMR spectroscopy*. *J Nat Prod*, 2014. **77**(6): p. 1473-87.

107. Smith, M.A., et al., *Characterization of novel triacylglycerol estolides from the seed oil of Mallotus philippensis and Trewia nudiflora*. *Lipids*, 2013. **48**(1): p. 75-85.
108. Hayes, D.G., R. Kleiman, and B.S. Phillips, *The triglyceride composition, structure, and presence of estolides in the oils of Lesquerella and related species*. *J. Am. Oil Chem. Soc.*, 1995. **72**(5): p. 559-69.
109. Lin, J.-T., et al., *Identification of (12-ricinoleoylricinoleoyl)diricinoleoylglycerol, an acylglycerol containing four acyl chains, in castor (ricinus communis L.) oil by LC-ESI-MS*. *J. Agric. Food Chem.*, 2006. **54**(10): p. 3498-3504.
110. Isbell, T.A., M.R. Edgcomb, and B.A. Lowery, *Physical properties of estolides and their ester derivatives*. *Ind. Crops Prod.*, 2001. **13**(1): p. 11-20.
111. Cermak, S.C., K.B. Brandon, and T.A. Isbell, *Synthesis and physical properties of estolides from lesquerella oil and castor oil fatty acid esters*. *Ind. Crops Prod.*, 2006. **23**(1): p. 54-64.
112. Sun, C., Y.Z. Cao, and A.H. Huang, *Acyl coenzyme a preference of the glycerol phosphate pathway in the microsomes from the maturing seeds of palm, maize, and rapeseed*. *Plant Physiol*, 1988. **88**(1): p. 56-60.
113. Broun, P., S. Boddupalli, and C. Somerville, *A bifunctional oleate 12-hydroxylase: desaturase from Lesquerella fendleri*. *Plant J*, 1998. **13**(2): p. 201-10.
114. van de Loo, F.J., et al., *An oleate 12-hydroxylase from Ricinus communis L. is a fatty acyl desaturase homolog*. *Proc Natl Acad Sci U S A*, 1995. **92**(15): p. 6743-7.
115. Bafor, M., et al., *Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor-bean (Ricinus communis) endosperm*. *Biochem J*, 1991. **280 ( Pt 2)**: p. 507-14.
116. Engeseth, N. and S. Stymne, *Desaturation of oxygenated fatty acids in Lesquerella and other oil seeds*. *Planta*, 1996. **198**(2): p. 238-245.
117. Bach, L., et al., *The very-long-chain hydroxy fatty acyl-CoA dehydratase PASTICCINO2 is essential and limiting for plant development*. *Proc Natl Acad Sci U S A*, 2008. **105**(38): p. 14727-31.
118. Gmeiner, J. and H.H. Martin, *Phospholipid and lipopolysaccharide in Proteus mirabilis and its stable protoplast L-form. Difference in content and fatty acid composition*. *Eur J Biochem*, 1976. **67**(2): p. 487-94.
119. Yano, I., et al., *Occurrence of 2- and 3-hydroxy fatty acids in high concentrations in the extractable and bound lipids of Flavobacterium meningosepticum and Flavobacterium IIb*. *Lipids*, 1976. **11**(9): p. 685-8.
120. Venter, P., et al., *Production of 3R-hydroxy-polyenoic fatty acids by the yeast Dipodascopsis uninucleata*. *Lipids*, 1997. **32**(12): p. 1277-83.
121. Jin, S.J., C.L. Hoppel, and K.Y. Tserng, *Incomplete fatty acid oxidation. The production and epimerization of 3-hydroxy fatty acids*. *J Biol Chem*, 1992. **267**(1): p. 119-25.
122. La Lone, C.A., et al., *Endogenous levels of Echinacea alkylamides and ketones are important contributors to the inhibition of prostaglandin E2 and nitric oxide production in cultured macrophages*. *J. Agric. Food Chem.*, 2009. **57**(19): p. 8820-8830.

123. Dayan, F.E., I.A. Kagan, and A.M. Rimando, *Elucidation of the biosynthetic pathway of the allelochemical sorgoleone using retrobiosynthetic NMR analysis*. J. Biol. Chem., 2003. **278**(31): p. 28607-28611.
124. Dayan, F.E., et al., *Sorgoleone*. Phytochemistry (Elsevier), 2010. **71**(10): p. 1032-1039.
125. Diaz, J., et al., *Fusarium confers protection against several mycelial pathogens of pepper plants*. Plant Pathol., 2005. **54**(6): p. 773-780.
126. Zhang, Z.-X., et al., *Discovery of putative capsaicin biosynthetic genes by RNA-Seq and digital gene expression analysis of pepper*. Sci. Rep., 2016. **6**: p. 34121.
127. Woelkart, K. and R. Bauer, *The role of alkamides as an active principle of Echinacea*. Planta Med., 2007. **73**(7): p. 615-623.
128. Cortez-Espinosa, N., et al., *Valine and phenylalanine as precursors in the biosynthesis of alkamides in Acemella radicans*. Nat. Prod. Commun., 2011. **6**(6): p. 857-861.
129. Minto, R.E. and B.J. Blacklock, *Biosynthesis and function of polyacetylenes and allied natural products*. Prog. Lipid Res., 2008. **47**(4): p. 233-306.
130. Barnes, J., et al., *Echinacea species (Echinacea angustifolia (DC.) hell., Echinacea pallida (Nutt.) Nutt., Echinacea purpurea (L.) Moench): A review of their chemistry, pharmacology and clinical properties*. J. Pharm. Pharmacol., 2005. **57**(8): p. 929-954.
131. Greger, H., *Alkamides: structural relationships, distribution and biological activity*. Planta Med., 1984. **50**(5): p. 366-75.
132. Shepard Jr, M.R., *Crepenynate-derived specialized natural products*. 2013.
133. Bauer, R., P. Remiger, and H. Wagner, *Alkamides from the roots of Echinacea purpurea*. Phytochemistry, 1988. **27**(7): p. 2339-42.
134. Molina-Torres, J., A. Garcia-Chavez, and E. Ramirez-Chavez, *Antimicrobial properties of alkamides present in flavoring plants traditionally used in Mesoamerica: affinin and capsaicin*. J. Ethnopharmacol., 1999. **64**(3): p. 241-248.
135. Muccioli, G.G., *Endocannabinoid biosynthesis and inactivation, from simple to complex*. Drug Discovery Today, 2010. **15**(11/12): p. 474-483.
136. Jonsson, K.-O., et al., *Effects of homologues and analogues of palmitoylethanolamide upon the inactivation of the endocannabinoid anandamide*. Br. J. Pharmacol., 2001. **133**(8): p. 1263-1275.
137. Ho, W.S.V., D.A. Barrett, and M.D. Randall, *'Entourage' effects of N-palmitoylethanolamide and N-oleoylethanolamide on vasorelaxation to anandamide occur through TRPV1 receptors*. Br. J. Pharmacol., 2008. **155**(6): p. 837-846.
138. O'Sullivan, S.E. and D.A. Kendall, *Cannabinoid activation of peroxisome proliferator-activated receptors: Potential for modulation of inflammatory disease*. Immunobiology, 2010. **215**(8): p. 611-616.
139. Barrett, B., et al., *Echinacea for treating the common cold: a randomized trial*. Ann. Intern. Med., 2010. **153**(12): p. 769-77.
140. Lopez-Bucio, J., et al., *Cytokinin receptors are involved in alkamide regulation of root and shoot development in Arabidopsis*. Plant Physiol., 2007. **145**(4): p. 1703-1713.

141. Ramirez-Chavez, E., et al., *Alkamides isolated from plants promote growth and alter root development in Arabidopsis*. Plant Physiol., 2004. **134**(3): p. 1058-1068.
142. Suzuki, H., et al., *Signals for local and systemic responses of plants to pathogen attack*. J. Exp. Bot., 2004. **55**(395): p. 169-179.
143. Kang, L., et al., *Overexpression of a fatty acid amide hydrolase compromises innate immunity in Arabidopsis*. Plant J., 2008. **56**(2): p. 336-349.
144. Tripathy, S., B.J. Venables, and K.D. Chapman, *N-acylethanolamines in signal transduction of elicitor perception. Attenuation of alkalization response and activation of defense gene expression*. Plant Physiol., 1999. **121**(4): p. 1299-1308.
145. Mendez-Bravo, A., et al., *Alkamides activate jasmonic acid biosynthesis and signaling pathways and confer resistance to Botrytis cinerea in Arabidopsis thaliana*. PLoS One, 2011. **6**(11): p. e27251.
146. Sperling, P., U. Zahringer, and E. Heinz, *A sphingolipid desaturase from higher plants. Identification of a new cytochrome b5 fusion protein*. J. Biol. Chem., 1998. **273**(44): p. 28590-28596.
147. Chen, M., J.E. Markham, and E.B. Cahoon, *Sphingolipid  $\Delta 8$  unsaturation is important for glucosylceramide biosynthesis and low-temperature performance in Arabidopsis*. Plant J., 2012. **69**(5): p. 769-781.
148. Tian, Y., et al.,  *$\Delta 10(E)$ -Sphingolipid desaturase involved in fusaricide mycosynthesis and stress adaptation in fusarium graminearum*. Sci Rep, 2015. **5**: p. 10486.
149. Bauer, R., I.A. Khan, and H. Wagner, *TLC and HPLC analysis of Echinacea pallida and E. angustifolia roots*. Planta Med., 1988. **54**(5): p. 426-30.
150. Baerson, S.R., et al., *A functional genomics approach for the identification of genes involved in the biosynthesis of the allelochemical sorgoleone*. ACS Symp. Ser., 2006. **927**(Natural Products for Pest Management): p. 265-276.
151. Cook, D., et al., *Molecular and biochemical characterization of novel polyketide synthases likely to be involved in the biosynthesis of sorgoleone*. ACS Symp. Ser., 2007. **955**(Polyketides): p. 141-151.
152. Sun, Y.-X., et al., *Biosynthesis of anandamide and N-palmitoylethanolamine by sequential actions of phospholipase A2 and lysophospholipase D*. Biochem. J., 2004. **380**(3): p. 749-756.
153. Bohlmann, F. and E. Dallwitz, *Polyacetylene compounds. 228. Biogenesis of polynamides*. Chem. Ber., 1974. **107**(6): p. 2120-2.
154. Li-Beisson, Y., et al., *Acyl-lipid metabolism*. Arabidopsis Book, 2013. **11**: p. e0161.
155. Focke, M., et al., *Fatty acid biosynthesis in mitochondria of grasses: Malonyl-coenzyme A is generated by a mitochondrial-localized acetyl-coenzyme A carboxylase*. Plant Physiol., 2003. **133**(2): p. 875-884.
156. Heazlewood, J.L., et al., *Towards an analysis of the rice mitochondrial proteome*. Plant Physiol., 2003. **132**(1): p. 230-242.
157. Wada, H., D. Shintani, and J. Ohlrogge, *Why do mitochondria synthesize fatty acids? Evidence for involvement in lipoic acid production*. Proc. Natl. Acad. Sci. U. S. A., 1997. **94**(4): p. 1591-1596.

158. Gueguen, V., et al., *Fatty acid and lipoic acid biosynthesis in higher plant mitochondria*. J. Biol. Chem., 2000. **275**(7): p. 5016-5025.
159. Birch, G.L., *Lipidomic profiling of Dictyostelium discoideum*. 2011.
160. Jones, A.V.M., et al., *Inhibition of plant fatty acid synthesis by nitroimidazoles*. Biochem. J., 1981. **198**(1): p. 193-8.
161. Harwood, J.L., et al., *Inhibition of fatty acid biosynthesis by metronidazole*. Biochem Soc Trans, 1980. **8**(5): p. 535-6.
162. Yasuno, R., P. von Wettstein-Knowles, and H. Wada, *Identification and molecular characterization of the  $\beta$ -ketoacyl-[acyl carrier protein] synthase component of the arabidopsis mitochondrial fatty acid synthase*. J. Biol. Chem., 2004. **279**(9): p. 8242-8251.
163. Konishi, T. and Y. Sasaki, *Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance toward herbicides*. Proc. Natl. Acad. Sci. U. S. A., 1994. **91**(9): p. 3598-601.
164. Laliotis, G.P., I. Bizelis, and E. Rogdaki, *Comparative approach of the de novo fatty acid synthesis (lipogenesis) between ruminant and non ruminant mammalian species: from biochemical level to the main regulatory lipogenic genes*. Curr. Genomics, 2010. **11**(3): p. 168-183.
165. Garton, G.A., *Fatty acid metabolism in ruminants*. Int. Rev. Biochem., 1977. **14**: p. 337-70.
166. Radmacher, E., et al., *Two functional FAS-I type fatty acid synthases in Corynebacterium glutamicum*. Microbiology (Reading, U. K.), 2005. **151**(7): p. 2421-2427.
167. Eckermann, S., et al., *New pathway to polyketides in plants*. Nature (London), 1998. **396**(6709): p. 387-390.
168. Hudaib, M., et al., *GC-MS analysis of the lipophilic principles of Echinacea purpurea and evaluation of cucumber mosaic cucumovirus infection*. J. Pharm. Biomed. Anal., 2002. **29**(6): p. 1053-1060.
169. Organ, M.G. and H. Ghasemi, *Metal-Catalyzed Coupling Reactions on an Olefin Template: The Total Synthesis of (13E,15E,18Z,20Z)-1-Hydroxypentacos-13,15,18,20-tetraen-11-yn-4-one 1-Acetate*. J. Org. Chem., 2004. **69**(3): p. 695-700.
170. Li, Z., et al., *Synthesis of photo-activatable analogs of lysophosphatidic acid and covalent labeling of plasma proteins*. J. Org. Chem., 2006. **71**(2): p. 629-635.
171. Akakabe, Y., et al., *Concise synthesis of (8Z,11Z,14Z)-8,11,14-heptadecatrienal, (7Z,10Z,13Z)-7,10,13-hexadecatrienal, and (8Z,11Z)-8,11-heptadecadienal, components of the essential oil of marine green alga Ulva pertusa*. Biosci. Biotechnol. Biochem., 2005. **69**(7): p. 1348-1352.
172. Stille, J.K. and J.H. Simpson, *Stereospecific palladium-catalyzed coupling reactions of vinyl iodides with acetylenic tin reagents*. J. Am. Chem. Soc., 1987. **109**(7): p. 2138-52.
173. Kumar, B., et al., *Allyl tetrahydropyranyl ether: a versatile alcohol/thiol protecting reagent*. Tetrahedron Lett., 2009. **50**(46): p. 6236-6240.
174. Dieskau, A.P., J.-M. Begouin, and B. Plietker, *Bu<sub>4</sub>N[Fe(CO)<sub>3</sub>(NO)]-catalyzed hydrosilylation of aldehydes and ketones*. Eur. J. Org. Chem., 2011. **2011**(27): p. 5291-5296, S5291/1-S5291/15.

175. Vechorkin, O. and X. Hu, *Nickel-catalyzed cross-coupling of non-activated and functionalized alkyl halides with alkyl Grignard reagents*. *Angew. Chem., Int. Ed.*, 2009. **48**(16): p. 2937-2940.
176. Dieskau, A.P. and B. Plietker, *A mild ligand-free iron-catalyzed liberation of alcohols from allylcarbonates*. *Org. Lett.*, 2011. **13**(20): p. 5544-5547.
177. Boeck, F., et al., *Mixed phosphane  $\eta^5$ -CpRuCl(PR3)2 complexes as ambifunctional catalysts for anti-markovnikov hydration of terminal alkynes*. *J. Am. Chem. Soc.*, 2011. **133**(21): p. 8138-8141.
178. Chen, G.-Q., et al., *Selective oxidation of terminal aryl and aliphatic alkenes to aldehydes catalyzed by iron(III) porphyrins with triflate as a counter anion*. *Chem. Commun. (Cambridge, U. K.)*, 2011. **47**(39): p. 10963-10965.
179. Hodgson, D.M., M.J. Fleming, and S.J. Stanway, *The reactivity of epoxides with lithium 2,2,6,6-tetramethylpiperidide in combination with organolithiums or grignard reagents*. *J. Org. Chem.*, 2007. **72**(13): p. 4763-4773.
180. Sharma, V., G.T. Kelly, and C.M.H. Watanabe, *Exploration of the molecular origin of the azinomycin epoxide: Timing of the biosynthesis revealed*. *Org. Lett.*, 2008. **10**(21): p. 4815-4818.
181. Mitchell, I.S., G. Pattenden, and J. Stonehouse, *A total synthesis of the antitumour macrolide rhizoxin D*. *Org. Biomol. Chem.*, 2005. **3**(24): p. 4412-4431.
182. Matovic, N.J., et al., *Polyunsaturated alkyl amides from Echinacea: synthesis of diynes, enynes, and dienes*. *J. Org. Chem.*, 2011. **76**(11): p. 4467-4481.
183. Xiao, W., et al., *One-pot synthesis of (2E)- and (2E, 4E)-unsaturated carboxylic acid amides via organotellurium reagents*. *Heteroat. Chem.*, 1990. **1**(3): p. 245-9.
184. Leung, P.S.-W., Y. Teng, and P.H. Toy, *Chromatography-free wittig reactions using a bifunctional polymeric reagent*. *Org. Lett.*, 2010. **12**(21): p. 4996-4999.
185. Raghavan, S. and S.G. Subramanian, *Toward a modular, bidirectional synthesis of (-)-mucocin*. *Tetrahedron*, 2011. **67**(39): p. 7529-7539.
186. Yoshida, T., et al., *Crucial structural factors and mode of action of polyene amides as inhibitors for mitochondrial NADH-ubiquinone oxidoreductase (Complex I)*. *Biochemistry*, 2007. **46**(36): p. 10365-10372.
187. Seto, Y., et al., *Novel cyclic peptide, epichlicin, from the endophytic fungus, Epichloe typhina*. *Biosci., Biotechnol., Biochem.*, 2007. **71**(6): p. 1470-1475.
188. T.Yamaji, T.S., K.Hayamizu, M.Yanagisawa and O.Yamamoto. *Spectral Database for Organic Compounds SDBS*. 2017]; Available from: <http://sdb.sdb.aist.go.jp>
189. Sequin, U., *Carbon-13 NMR spectral differences between corresponding methyl esters, phenyl esters and 2-substituted chromones*. *Helv. Chim. Acta*, 1981. **64**(8): p. 2654-64.
190. Satoh, S., H. Suginome, and M. Tokuda, *Electrochemical conjugate additions of the allyl groups in substituted allyl halides to  $\alpha,\beta$ -unsaturated esters*. *Bull. Chem. Soc. Jpn.*, 1981. **54**(11): p. 3456-9.
191. Baeckstroem, P., et al., *Synthesis and characterization of all four isomers of methyl 2,4-decadienoate for an investigation of the pheromone components of Pityogenes chalcographus*. *Tetrahedron*, 1988. **44**(9): p. 2541-8.



192. Ishbaeva, A.U., R.N. Shakhmaev, and V.V. Zorin, *Synthesis of (2E,4E)-dodeca-2,4-dien-1-yl isovalerate, the main component of rootstock oil of Echinacea purpurea*. Russ. J. Org. Chem., 2010. **46**(2): p. 174-176.
193. Oka, M., et al., *Glidobactins A, B and C, new antitumor antibiotics. II. Structure elucidation*. J. Antibiot., 1988. **41**(10): p. 1338-50.
194. Maulucci, N., et al., *Molecular insights into azumamide E histone deacetylases inhibitory activity*. J Am Chem Soc, 2007. **129**(10): p. 3007-3012.
195. Merrill, A.H., Jr., S. Gidwitz, and R.M. Bell, *Facile enzymatic synthesis of fatty acylcoenzyme A thioesters*. J Lipid Res, 1982. **23**(9): p. 1368-73.
196. Blancaflor, E.B., et al., *N-Acylethanolamines: lipid metabolites with functions in plant growth and development*. Plant J., 2014. **79**(4): p. 568-583.
197. Larson, T.R. and I.A. Graham, *A novel technique for the sensitive quantification of acyl CoA esters from plant tissues*. Plant J., 2001. **25**(1): p. 115-125.
198. Lager, I., et al., *Plant Acyl-CoA:Lysophosphatidylcholine Acyltransferases (LPCATs) Have Different Specificities in Their Forward and Reverse Reactions*. J. Biol. Chem., 2013. **288**(52): p. 36902-36914.
199. Sanchez, M., D.G. Nicholls, and D.N. Brindley, *The relationship between palmitoyl-coenzyme A synthetase activity and esterification of sn-glycerol 3-phosphate in rat liver mitochondria*. Biochem. J., 1973. **132**(4): p. 697-706.
200. Levoine, N., et al., *Synthesis and biological testing of acyl-CoA-ketoprofen conjugates as selective irreversible inhibitors of COX-2*. Bioorg. Med. Chem., 2002. **10**(3): p. 753-757.
201. Mukherjee, K.D.W., N., *CRC Handbook of Chromatography: Analysis of Lipids*. 1993, Boca Raton, FL: CRC Press, Inc.
202. Al-Arif, A. and M. Blecher, *Synthesis of fatty acyl CoA and other thiol esters using N-hydroxysuccinimide esters of fatty acids*. J. Lipid Res., 1969. **10**(3): p. 344-5.
203. Geisler, K., et al., *Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants*. Proc. Natl. Acad. Sci. U. S. A., 2013. **110**(35): p. E3360-E3367.
204. Nelson, D. and D. Werck-Reichhart, *A P450-centric view of plant evolution*. Plant J., 2011. **66**(1): p. 194-211.
205. Bak, S., et al., *Cytochromes P450*. Arabidopsis Book, 2011. **9**: p. e0144.
206. Graham, S.E. and J.A. Peterson, *How Similar Are P450s and What Can Their Differences Teach Us?* Arch. Biochem. Biophys., 1999. **369**(1): p. 24-29.
207. Werck-Reichhart, D. and R. Feyereisen, *Cytochromes P450: a success story*. Genome Biol, 2000. **1**(6): p. REVIEWS3003.
208. Nelson, D.R., *Cytochrome P450 nomenclature*. Methods Mol. Biol. . Vol. 107. 1998: Humana Press Inc. 15-24.
209. Durst, F. and D.R. Nelson, *Diversity and evolution of plant P450 and P450-reductases*. Drug Metab. Drug Interact., 1995. **12**(3-4): p. 189-206.
210. Paquette, S.M., S. Bak, and R. Feyereisen, *Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of Arabidopsis thaliana*. DNA Cell Biol., 2000. **19**(5): p. 307-317.

211. Lepesheva, G.I. and M.R. Waterman, *Sterol 14 $\alpha$ -demethylase cytochrome P 450 (CYP51), a P450 in all biological kingdoms*. Biochim. Biophys. Acta, Gen. Subj., 2007. **1770**(3): p. 467-477.
212. Lentz, O., et al., *Altering the regioselectivity of cytochrome P450 CYP102A3 of Bacillus subtilis by using a new versatile assay system*. ChemBioChem, 2006. **7**(2): p. 345-350.
213. Vranova, E., D. Coman, and W. Gruissem, *Structure and dynamics of the isoprenoid pathway network*. Mol. Plant, 2012. **5**(2): p. 318-333.
214. Eastman, R.H.K., Ronald H. *Isoprenoid*. August 15, 2008 [cited 2017 April 30]; Available from: <https://www.britannica.com/science/isoprenoid>.
215. Xu, R., G.C. Fazio, and S.P.T. Matsuda, *On the origins of triterpenoid skeletal diversity*. Phytochemistry (Elsevier), 2004. **65**(3): p. 261-291.
216. Clayton, R.B., *Biosynthesis of sterols, steroids, and terpenoids. II. Phytosterols, terpenes, and the physiologically active steroids*. Quart. Rev. (London), 1965. **19**(3): p. 201-230.
217. Haralampidis, K., et al., *A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots*. Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(23): p. 13431-13436.
218. Dinda, B., et al., *Naturally Occurring Triterpenoid Saponins*. Chem. Biodiv., 2010. **7**(10): p. 2327-2580.
219. Townsend, B., H. Jenner, and A. Osbourn, *Saponin glycosylation in cereals*. Phytochem. Rev., 2006. **5**(1): p. 109-114.
220. Lv, X.-J., et al., *Antiviral triterpenes from the twigs and leaves of Lyonia ovalifolia*. J. Nat. Prod., 2016. **79**(11): p. 2824-2837.
221. Haraguchi, H., et al., *Antimicrobial triterpenes from Ilex integra and the mechanism of antifungal action*. Phytother. Res., 1999. **13**(2): p. 151-156.
222. Osbourn, A.E., et al., *An oat species lacking avenacin is susceptible to infection by Gaeumannomyces graminis var. tritici*. Physiol. Mol. Plant Pathol., 1994. **45**(6): p. 457-467.
223. Hostettmann, K., A. Marston, and Editors, *Saponins*. 1995: Cambridge Univ Press. 560 pp.(approx.).
224. Trojanowska, M.R., et al., *Biosynthesis of avenacins and phytosterols in roots of Avena sativa cv. Image*. Phytochemistry, 2000. **54**(2): p. 153-164.
225. Crombie, W.M.L. and L. Crombie, *Distribution of avenacins A 1, A 2, B 1 and B 2 in oat roots: their fungicidal activity towards 'take-all' fungus*. Phytochemistry, 1986. **25**(9): p. 2069-2073.
226. Piasecka, A., N. Jedrzejczak-Rey, and P. Bednarek, *Secondary metabolites in plant innate immunity: conserved function of divergent chemicals*. New Phytol, 2015. **206**(3): p. 948-964.
227. Papadopoulou, K., et al., *Compromised disease resistance in saponin-deficient plants*. Proc. Natl. Acad. Sci. U. S. A., 1999. **96**(22): p. 12923-12928.
228. Mugford, S.T., et al., *A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats*. Plant Cell, 2009. **21**(8): p. 2473-2484.

229. Mugford, S.T., et al., *Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat*. *Plant Cell*, 2013. **25**(3): p. 1078-1092.
230. Qi, X., et al., *A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants*. *Proc. Natl. Acad. Sci. U. S. A.*, 2004. **101**(21): p. 8233-8238.
231. Trojanowska, M.R., et al., *Investigation of avenacin-deficient mutants of *Avena strigosa**. *Phytochemistry*, 2001. **56**(2): p. 121-129.
232. Salmon, M., et al., *A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases*. *Proc. Natl. Acad. Sci. U. S. A.*, 2016. **113**(30): p. E4407-E4414.
233. Qi, X., et al., *A different function for a member of an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense*. *Proc. Natl. Acad. Sci. U. S. A.*, 2006. **103**(49): p. 18848-18853.
234. De Kraker, J.-W., et al., *Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene A hydroxylase and NADP+-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis*. *Plant Physiol.*, 2001. **125**(4): p. 1930-1940.
235. Huhman, D.V. and L.W. Sumner, *Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer*. *Phytochemistry*, 2002. **59**(3): p. 347-360.
236. Marston, A., J.L. Wolfender, and K. Hostettmann, *Analysis and isolation of saponins from plant material*, in *Saponins in Food, Feedstuffs and Medicinal Plants*, W. Oleszek and A. Marston, Editors. 2000, Springer Netherlands: Dordrecht. p. 1-12.
237. Coon, M.J., A.D. Vaz, and L.L. Bestervelt, *Cytochrome P450 2: peroxidative reactions of diversozymes*. *FASEB J*, 1996. **10**(4): p. 428-34.
238. Nelson, D.R., et al., *Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot*. *Plant Physiol.*, 2004. **135**(2): p. 756-772.
239. Lepesheva, G.I., C. Virus, and M.R. Waterman, *Conservation in the CYP51 family. Role of the B' helix/BC loop and helices F and G in enzymatic function*. *Biochemistry*, 2003. **42**(30): p. 9091-9101.
240. Kunii, M., et al.,  *$\beta$ -Amyrin oxidation by oat CYP51H10 expressed heterologously in yeast cells: the first example of CYP51-dependent metabolism other than the 14-demethylation of sterol precursors*. *Biol. Pharm. Bull.*, 2012. **35**(5): p. 801-804.
241. Anzai, Y., et al., *Function of cytochrome P450 enzymes MycCI and MycG in *Micromonospora griseorubida*, a producer of the macrolide antibiotic mycinamicin*. *Antimicrob. Agents Chemother.*, 2012. **56**(7): p. 3648-3656.
242. Kudo, F., et al., *Cloning and characterization of the biosynthetic gene cluster of 16-membered macrolide antibiotic FD-891: involvement of a dual functional cytochrome P450 monooxygenase catalyzing epoxidation and hydroxylation*. *ChemBioChem*, 2010. **11**(11): p. 1574-1582.

243. Born, S.L., et al., *Characterization of the progesterone 21-hydroxylase activity of canine cytochrome P450 PBD-2/P450 2B11 through reconstitution, heterologous expression, and site-directed mutagenesis*. *Drug Metab. Dispos.*, 1995. **23**(7): p. 702-7.
244. Schalk, M. and R. Croteau, *A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase*. *Proc. Natl. Acad. Sci. U. S. A.*, 2000. **97**(22): p. 11948-11953.
245. Liu, J., et al., *The effect of reciprocal active site mutations in human cytochromes P450 1A1 and 1A2 on alkoxyresorufin metabolism*. *Arch. Biochem. Biophys.*, 2004. **424**(1): p. 33-43.
246. Cools, H.J., et al., *Heterologous expression of mutated eburicol 14 $\alpha$ -demethylase (CYP51) proteins of *Mycosphaerella graminicola* to assess effects on azole fungicide sensitivity and intrinsic protein function*. *Appl. Environ. Microbiol.*, 2010. **76**(9): p. 2866-2872.
247. Yuki, H., et al., *Prediction of sites of metabolism in a substrate molecule, instanced by carbamazepine oxidation by CYP3A4*. *Bioorg. Med. Chem.*, 2012. **20**(2): p. 775-783.
248. Augustin, J.M., et al., *Molecular activities, biosynthesis and evolution of triterpenoid saponins*. *Phytochemistry (Elsevier)*, 2011. **72**(6): p. 435-457.
249. Moses, T., et al., *Bioengineering of plant (tri)terpenoids: from metabolic engineering of plants to synthetic biology in vivo and in vitro*. *New Phytol.*, 2013. **200**(1): p. 27-43.
250. Moses, T., K.K. Papadopoulou, and A. Osbourn, *Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives*. *Crit. Rev. Biochem. Mol. Biol.*, 2014. **49**(6): p. 439-462.
251. Sparg, S.G., M.E. Light, and J. Van Staden, *Biological activities and distribution of plant saponins*. *J. Ethnopharmacol.*, 2004. **94**(2-3): p. 219-243.
252. Crews, P., J. Rodriguez, and M. Jaspars, *Organic Structure Analysis*. 1998, New York, New York: Oxford University Press. 544 pp.
253. Arriaga-Giner, F.J., et al., *2 $\beta$ -Hydroxyhauthriwaic acid, a clerodane type diterpenoid and other terpenoids from three *Baccharis* species*. *Phytochemistry*, 1986. **25**(3): p. 719-21.
254. Quijano, L., et al., *The molecular structure of maniladiol from *Baccharis salicina**. *Phytochemistry*, 1998. **49**(7): p. 2065-2068.
255. Knight, S.A., *Carbon-13 NMR spectra of tetra- and pentacyclic triterpenoids*. *Org. Magn. Reson.*, 1974. **6**(11): p. 603-11.
256. Ercil, D., et al., *Chemical constituents of *Linaria aucheri**. *Turk. J. Chem.*, 2004. **28**(1): p. 133-139.
257. Seo, S., Y. Tomita, and K. Tori, *Carbon-13 NMR spectra of urs-12-enes and application to structural assignments of components of *Isodon japonicus* tissue cultures*. *Tetrahedron Lett.*, 1975(1): p. 7-10.
258. Agrawal, P.K., et al., *Carbon-13 NMR spectroscopy of steroidal saponins and steroidal saponins*. *Phytochemistry*, 1985. **24**(11): p. 2479-96.
259. Mahato, S.B. and A.P. Kundu, *Review article number 98: 13C NMR spectra of pentacyclic triterpenoids - a compilation and some salient features*. *Phytochemistry*, 1994. **37**(6): p. 1517-75.

260. Reich, H.J. *H-NMR-2-Chemical Shift*. 2017 [cited 2017; Available from: <https://www.chem.wisc.edu/areas/reich/nmr/05-hmr-02-delta.htm>].
261. Lambert, J.B. and A.R. Vagenas, *Dependence of the  $\gamma$  carbon-13 shielding effect on the dihedral angle*. *Org. Magn. Reson.*, 1981. **17**(4): p. 265-9.
262. Lehn, J.M., *Nuclear magnetic resonance (N.M.R.) studies of natural products. II. Triterpenes of the dammarane series: the methyl groups*. *Bull. Soc. Chim. Fr.*, 1962: p. 1832-7.
263. Cohen, A.I., et al., *Nuclear magnetic resonance spectra of lanostane derivatives*. *Tetrahedron*, 1965. **21**(11): p. 3171-83.
264. Tursch, B., R. Savoir, and G. Chiurdoglu, *Triterpenes. V. Nuclear magnetic resonance spectroscopy in triterpene compounds. Methyl groups in the oleanenic series*. *Bull. Soc. Chim. Belg.*, 1966. **75**(1-2): p. 107-15.
265. Ito, S., M. Kodama, and M. Sunagawa, *Correlation of methyl signals in  $3\beta$ -hydroxyoleanenes*. *Tetrahedron Lett.*, 1967(40): p. 3989-92.

## APPENDIX

### Assignment of $\beta$ -amyrin in pyridine $d_5$

There is no published NMR data of  $\beta$ -amyrin collected in pyridine- $d_5$  solvent. For the full investigation of the chemical shift differences in the oleanane family of triterpenes, the assignments of  $\beta$ -amyrin in pyridine  $d_5$  was undertaken. This gave a valid standard with which to compare the chemical shifts of the four mutant Sad2 products and wild-type Sad2 product. Again, initial identification was numbered based on  $^{13}\text{C}$  NMR but then switched to IUPAC once the structural was fully elucidated.

The same set of NMR experiments used for the wild-type product and mutant products were also use to collected data for  $\beta$ -amyrin in pyridine- $d_5$ . The hydroxyl group was confirmed on the HSQC with  $^{13}\text{C}$  peaks at 78.7 ppm correlating to a proton at 3.47 ppm. The double bond of beta-amyrin was present at 145.6 and 122.7 ppm in the  $^{13}\text{C}$  spectrum with the correlating proton at 5.25 ppm. Tabulated spectral data can be found in Table 16.

Starting with ring A, the hydroxyl-bearing C3 correlated to H23 and H24 making H23 and H24 the geminal dimethyl group on ring A (Figure 71A). The quaternary carbon C4 along with the methine C5 were identified through their coupling to H23 and H24. Additionally, C5 correlated to methyl H25. The quaternary carbon C10 and the methine C1 also coupled to H25. Finally, in the COSY, H3 showed a cross-peak to H2a/b making C2 the methylene alpha to the hydroxylated carbon.

On ring B, methine group C9 correlated to the methyl H25 at the bridgehead between ring A/B and the methyl H26 leading to the bridgehead of rings B/C (Figure 71B). Methylene H7 was coupled to H26. Quaternary carbons C14 and C8 both had correlation peaks to the methyl groups H27 and H26; through the correlation between C8 and H9, the two quaternary carbons could be distinguished from one another. The COSY showed a cross-peak from H5 to H6a/b.

On ring C, the double bond was easily identified by the  $^{13}\text{C}$  resonances at 122.7 (C13) and 145.6 ppm (C12), similar to the standards reported for 12-oleanenes in other solvents (Figure 71C). C12 correlated to the hydrogen at 5.25 ppm in the HSQC making

it the methine. In the HMBC spectrum, C12 coupled to the allylic hydrogens H11a/b. Finally, C13 showed an important cross peak to H27 confirming assigning the methyl group C27 alpha to C14 and the methyl C26 alpha to C8.

On ring D, the methylene C15 correlated to H27 (Figure 74D). In the COSY, there was a cross peak between H16b and H15. In the HMBC C16 correlated to the methyl H28. The last methine to be assigned was C18, which coupled to H28. Additionally, the showed the quaternary carbon C17 was attached to the methyl, H28.

Finally on ring E, the COSY coupled H18 to H19a (Figure 74E). C19 correlated to H29 and H30, which made C29 and C30 the other geminal dimethyl groups. The last quaternary carbon must be C20 confirmed by correlation to H29 and H30. C21 had two cross-peaks, linking it to H29 and H30. To finish ring E, it was observed that C22 coupled to H28. The NOESY confirmed the stereospecific groups; H3 correlated the axial hydroxgen H5 and methyl H23. Since H5 is axial in  $\beta$ -amyrin, this makes H3 axial and C23 equatorial. Critically, H18 correlated to H30 making it beta and H29 alpha. This allowed beta-amyrin to be fully assigned in pyridine- $d_5$  and used to compare to the chemical shifts of other beta-amyrin derivatives. IR (KBr pellet): br 3276, 2945, 2867, 2852, 1463, 1387, 1359, 1036  $\text{cm}^{-1}$ ; HRMS (Dual ESI)  $m/z$  calc for  $[\text{M} + \text{H}]^+$   $\text{C}_{30}\text{H}_{51}\text{O}$  427.3940, found 427.3939.

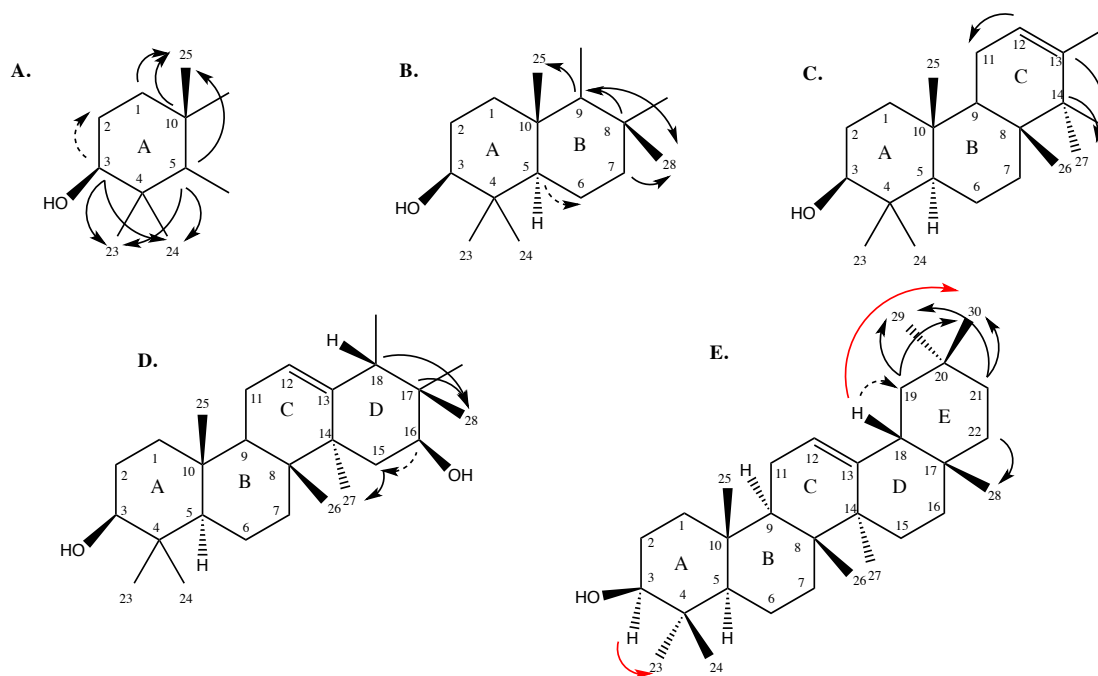
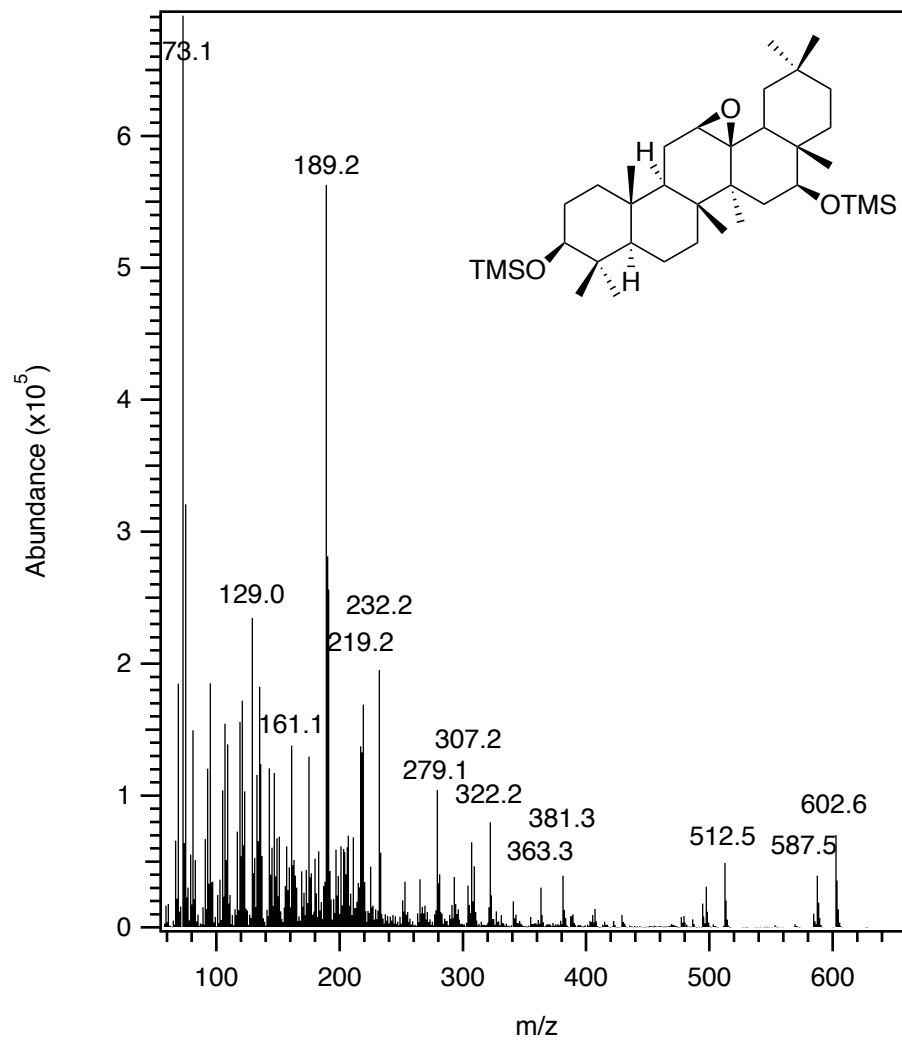


Figure 71: Structure elucidation of  $\beta$ -amyrin in pyridine- $d_5$ . IUPAC numbered. Substructures based on NMR correlations and coupling data. *A-E* Assigned nuclei on the five rings were constructed using  $^1\text{H}$ - $^1\text{H}$  COSY (dashed arrows) and HMBC heteronuclear correlations (solid arrows), and  $^1\text{H}$ - $^1\text{H}$  NOESY (red arrows) in Table 16 resulting in the final structure *E*.



Table 16: Tabulated NMR data for  $\beta$ -amyrin. Data was collected in pyridine- $d_5$ .

IUPAC #	$\delta^{13}\text{C}$	$\delta^1\text{H(m)}$	$J_{HH}$ (Hz)	DEPT	$^1\text{H}-^1\text{H}$ COSY Correlations	$^{13}\text{C}-^1\text{H}$ HMBC Correlations	$^1\text{H}-^1\text{H}$ NOESY Correlations
1	39.5	1.63 m; 1.05 m	-	CH <sub>2</sub>		25	
2	28.5	1.86 m	-	CH <sub>2</sub>			
3	78.4	3.47 m	-	CH	2a/b	23, 24	5, 1bw, 23, 2 a/b
4	39.7	-	-	C		23, 24	
5	56.1	0.90 m	-	CH	6a/b	23, 24, 25	
6	19.1	1.65 m; 1.47 m	-	CH <sub>2</sub>			
7	33.4	1.67 m; 1.46 m	-	CH <sub>2</sub>		26	
8	40.5	-	-	C		9, 26, 27	
9	48.4	1.70 m	-	CH		25, 26	
10	37.6	-	-	C		25	
11	24.2	1.95 m	-	CH <sub>2</sub>			
12	122.7	5.25 m	-	CH	11a/b		11, 27
13	145.6	-	-	C		27	
14	42.3	-	-	C		26, 27	
15	26.5	1.80 m; 1.00 m	-	CH <sub>2</sub>		27	
16	27.2	2.05 m; 0.83 m	-	CH <sub>2</sub>	15a	28	
17	33.1	-	-	C		28	
18	47.9	2.08 m	-	CH	19a	28	30
19	47.4	1.80 m; 1.15 m	-	CH <sub>2</sub>		29, 30	
20	31.5	-	-	C			
21	35.3	1.41 m; 1.15 m	-	CH <sub>2</sub>		29, 30	
22	37.8	1.49 m; 1.27 m	-	CH <sub>2</sub>		28	
23	29.1	1.28 s	-	CH <sub>3</sub>		24	
24	16.9	1.10 s	3H	CH <sub>3</sub>		23	25
25	16.1	1.00 s	-	CH <sub>3</sub>			24, 26
26	17.4	1.04 s	3H	CH <sub>3</sub>			
27	26.2	1.25 s	-	CH <sub>3</sub>			
28	28.9	0.95 s	-	CH <sub>3</sub>			30
29	33.8	0.94 s	-	CH <sub>3</sub>		30	
30	24.1	0.93 s	-	CH <sub>3</sub>		29	

Figure 72: TMS-16 $\beta$ -hydroxy-12,13-epoxy- $\beta$ -amyrin

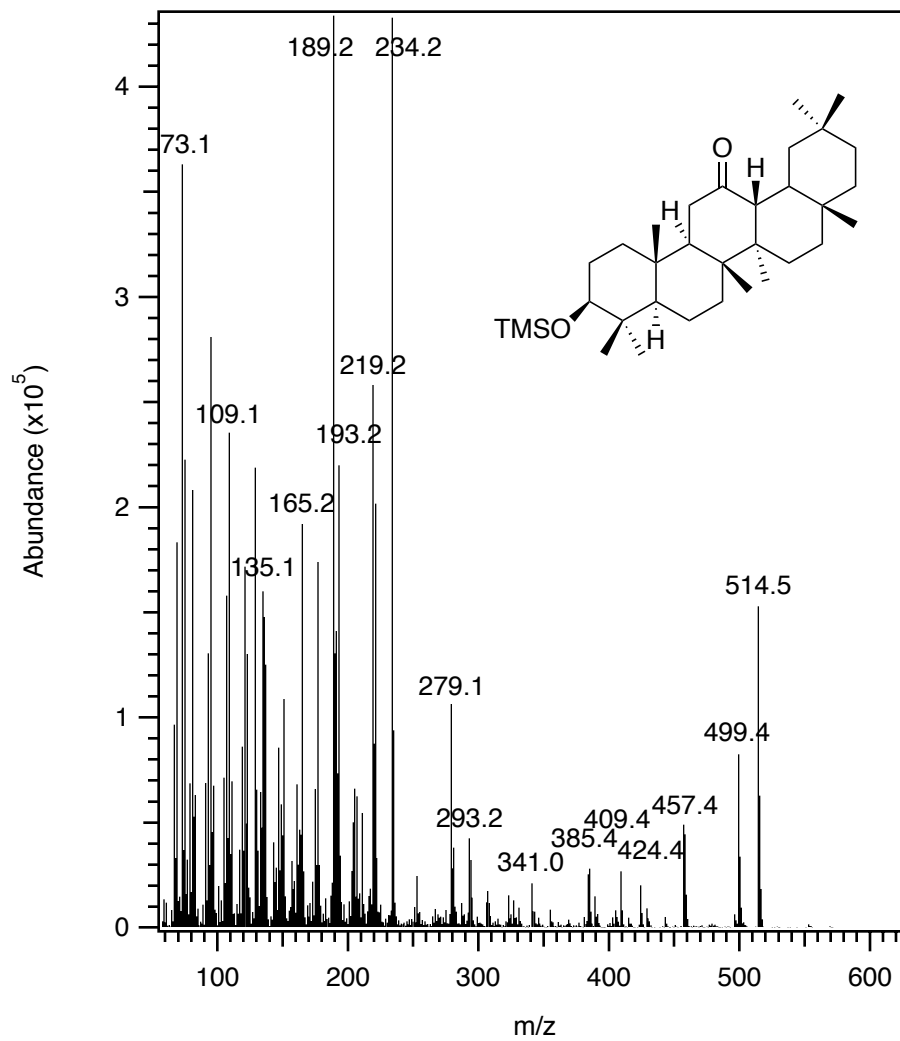
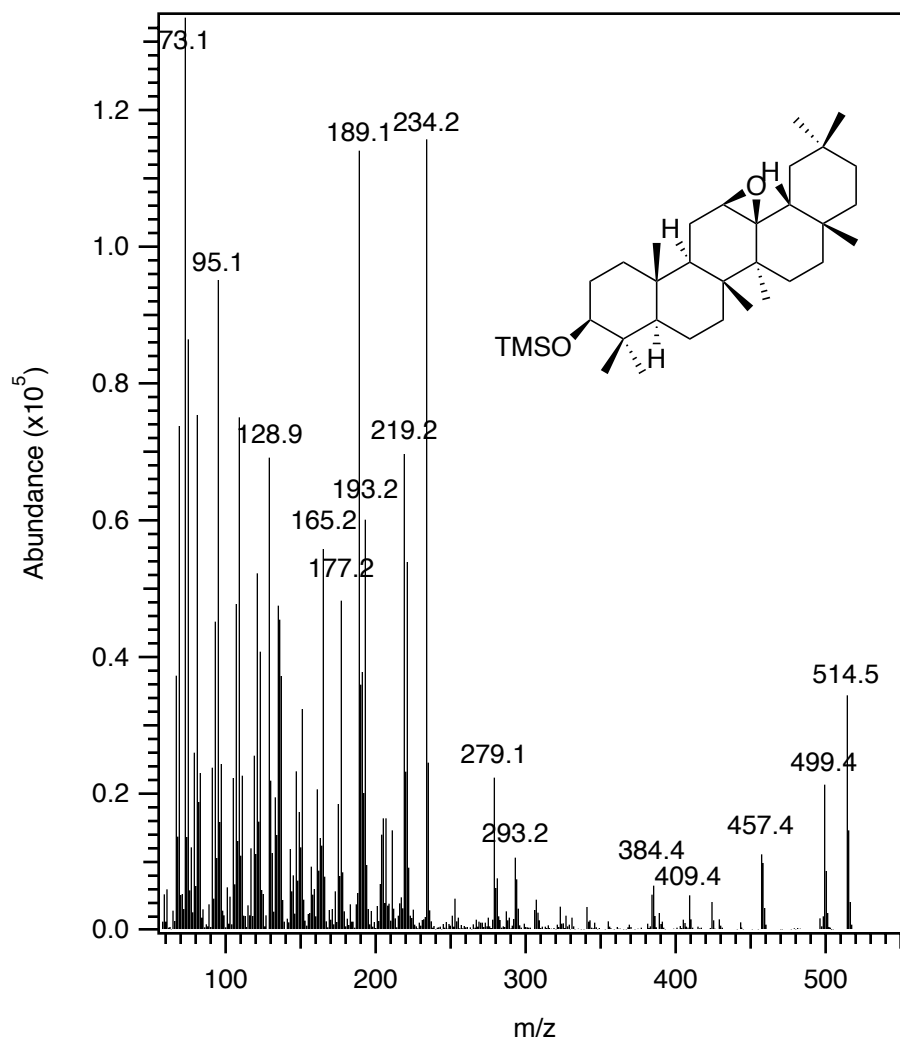


Figure 73: TMS-12-keto-13β-β-amyrin

Figure 74: TMS-12, 13 $\beta$ -epoxy- $\beta$ -amyirin