ISOLATION AND IDENTIFICATION OF MOLECULAR CONSTITUENTS FROM THE PROTECTIVE PLANT POLYMER SUBERIN

by

WEIMIN WANG

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2007

UMI Number: 3288742

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Dr. Ruth E. Stark

Chair of Examining Committee

Date

Dr. Gerald Koeppl

Executive Officer

Dr. Steven G. Greenbaum

Dr. James D. Batteas

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

ISOLATION AND IDENTIFICATION OF MOLECULAR CONSTITUENTS FROM THE PROTECTIVE PLANT POLYMER SUBERIN

by

Weimin Wang

Advisor: Professor Ruth E. Stark

The study of molecular structure for biopolymer suberin can provide compositional and structural information about its biosynthesis and help to understand its biological function, informing the design of crop protection strategies. Because of the intractability of suberin, a partial degradation strategy was employed. This research develops new methods for separation and identification of degraded suberin, to elucidate new aliphatic and aromatic oligomers that preserve its essential molecular architecture.

Potato suberin was depolymerized with KOH-methanol solution. Degradation conditions were varied systematically to enhance the yield of the desired compounds. A separation protocol, in which CC was used to do preliminary separation and multiple steps of solvent extraction were combined with TLC, RP-HPLC or NP-HPLC separation, was developed to improve the isolation efficiency of the target compounds. The strategy of combining NMR analysis of through-bond connectivities and MS^n fragments allowed identification of some compounds that are challenging to isolate in pure form.

Twenty different triglycerides, three new aliphatic monomers, more than twenty aliphatic oligomers, one aromatic dimer, and two new aromatic monomers were isolated from degradation of suberin by CC, TLC, solvent extraction, and HPLC. The molecular

structures of these isolated compounds were elucidated by 1D and 2D NMR, LC/MS, and LC/MS^n. These results were used to evaluate and extend current models of suberin architecture.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor and mentor, Prof. Ruth E. Stark, who gave me the wonderful opportunity to be involved in the analytical field of molecular structural studies on natural polymers and guided me patiently in every growth step.

I would like to extend my thanks to Dr. Hsin Wang, who helped me to set up NMR experiments and Dr. Boris Arshava, who helped me with the HPLC instrument and data interpretation.

I also want to thank all my committee members, Professor Steve Greenbaum and Professor James D. Batteas, for their sound advice and strong support.

I give special thanks to my parents, Xiaoyi Wang, Jingxin Zhang, and my brother, Limin Wang. It is their love and strong support that enables me to focus on my study and research.

Table of Contents

Chapter 1 INTRODUCTION 1

1.1 Why study suberin? 1

1.2 What is known about the structure of suberin? 3

1.2.1 Monomeric units: aliphatics 4

1.2.2 Monomeric units: phenolics 7

1.2.3 Monomeric units: glycerol 8

1.2.4 A Model for Macromolecular Architecture 10

1.3 The objectives of this research 12

Chapter 2 INTRODUCTION TO METHODS 14

2.1 Introduction to Chromatography 14

2.1.1 Basic Principles of Chromatography 14

2.1.2 Column Chromatography (CC) 15

2.1.3 Thin-Layer Chromatography (TLC) 15

2.1.4 High-Performance Liquid Chromatography (HPLC) 16

2.2 Introduction to Nuclear Magnetic Resonance (NMR) 18

2.2.1 Basic Principles of NMR 18

2.2.2 Solution-State NMR 20

2.2.3 ACD database software 24

2.3 Introduction to Mass Spectrometry (MS) 25

2.3.1 Basic Principles of MS 25

2.3.2 Ionization Source 28

2.3.3 Ion Analyzer 33

2.3.4 Detector 34

2.4 Introduction to LC-MS 35

Chapter 3 SAMPLE PREPARATION AND SPECTRAL ANALYSIS 36

3.1 Preparation and isolation of potato suberin 36

3.2 Generation of potato suberin aromatics and oligomers with KOH 38

3.3 Isolation and chromatographic purification of oligomers and aromatic monomers 39

3.3.1 Preliminary separation of depolymerized suberin by CC 39

3.3.2 Separation of depolymerized suberin by TLC 40

3.3.3 Separation of depolymerized suberin by HPLC 40

3.4 Solution-state NMR characterization of soluble products of suberin 41

3.5 Mass spectral analysis of suberin 42

Chapter 4 RESULTS AND DISCUSSION: SEPARATION AND PROVISIONAL IDENTIFICATION OF DEPOLYMERIZATION PRODUCTS 44

4.1 Preliminary study of potato suberin 44

4.1.1 The effect of depolymerization conditions 45

4.1.1.1 Depolymerization time 45

4.1.1.2 Concentration of KOH-methanol solution 47

4.1.1.3 ^{1}H NMR integration results 49

4.2 Separation of depolymerized suberin #11 using CC and TLC 51

4.2.1 Fraction #11 51

4.2.2 1D and 2D NMR study of #11-TLC-4-1 52

4.2.3 ACD spectral simulation of #11-TLC-4-1 58

4.3 HPLC method development 61

4.3.1 Challenges in the separation of suberin depolymerization products 61

4.3.2 HPLC separation of aromatic dimer #11-TLC-4-1 62

4.3.3 HPLC separation of #11-HPLC-2 71

4.4 Solvent extraction methods 74

4.5 Combination of HPLC and solvent extraction methods 77

4.5.1 Solvent extraction of aromatic compound 3 77

4.5.2 HPLC separation of aromatic compound 3 82

4.5.3 HPLC separation of a triglyceride 84

4.6 Conclusions 87

Chapter 5 RESULTS AND DISCUSSION: STRUCTURAL ELUCIDATION OF ALIPHATIC AND AROMATIC COMPOUNDS 88

5.1 Aliphatic compound 1: glycerol 88

5.2 Aliphatic compound 2: triglycerides 91

5.2.1 NMR identification of triglycerides 91

5.2.1.1 Triglycerides from depolymerized suberin 91

5.2.1.2 Standard triglyceride 100

5.2.2 LC/MS and LC/MS^n of aliphatic compound **2** 102

5.2.2.1 LC/MS identification of triglycerides 103

5.2.2.1.1 LC/MS of sample 5-24-#5-HPLC-f3 105

5.2.2.1.2 LC/MS of sample 5-24-#5-HPLC-f4 106

5.2.2.1.3 LC/MS of sample 5-24-#5-HPLC-f5 106

5.2.2.1.4 LC/MS of sample 5-24-#7 108

5.2.2.2 LC/MS^n identification of triglycerides 109

5.2.2.2.1 LC/MS^n of a triglyceride, tristearin 109

5.2.2.2.2 LC/MS^n of triglyceride compounds separated from suberin 111

5.2.2.2.2.1 LC/MS^n of sample 5-24-#5-HPLC-f4 111

5.2.2.2.2.2 LC/MS^n of sample 5-24-#5-HPLC-f3 116

5.2.2.2.2.3 LC/MS^n of sample 5-24-#5-HPLC-f5 118

5.2.2.2.2.4 LC/MS^n of sample 5-24-#7 118

5.3 Aliphatic compound 3: ester oligomers 119

5.3.1 NMR identification of aliphatic compound **3** 119

5.3.2 LC/MS identification of aliphatic compound **3** 121

5.3.3 The structure of aliphatic compound **3** 123

5.3.3.1 CH_3 _____OH Trimer structures 124

5.3.3.2 CH_3 _____CH_3 Trimer structures 127

5.3.4 LC/MS^n of aliphatic compound **3** 128

5.4 Aromatic compound 1 (#11-TLC-4-1) 128

5.4.1 1D and 2D NMR of aromatic compound **1** 128

5.4.2 Mass spectrum of sample #11-TLC-4-1 128

5.5 Aromatic compound **2** 135

5.5.1 1D and 2D NMR identification of aromatic compound **2** 135

5.5.1.1 Aromatic compound **2-1** 135

5.5.1.2 Aromatic compound **2-2** 139

5.5.2 LC/MS and LC/MS^n of aromatic compound **2** 140

5.5.2.1 LC/MS and LC/MS^n of LC elution between 3.4-3.6min 141

5.5.2.2 LC/MS and LC/MS^n of LC elution between 4.3-4.8 min 144

5.5.2.3 LC/MS and LC/MS^n of LC elution between 5.2-6.1 min 147

5.6 Aromatic compound **3** 149

5.6.1 Aromatic compound **3-f6** 149

5.6.1.1 NMR of aromatic compound **3-f6** 149

5.6.1.2 LC/MS and LC/MS^n of aromatic compound **3-f6** 155

5.6.2 aromatic compound **3-f3** 158

5.6.2.1 NMR of aromatic compound **3-f3** 158

5.6.2.2 LC/MS and LC/MS^n of aromatic compound **3-f6** 158

5.7 Aromatic compounds **4** 158

5.8 Aliphatic compound **4** 162

5.9 Summary 163

Chapter 6. CONCLUSIONS AND PROSPECTUS 169

Reference List 173

List of Tables

Table	4.1	The effect of time on percent yield of depolymerized products	45
Table	4.2	The effect of KOH concentration on overall percent yield of depolymerized products	47
Table	4.3	^{1}H NMR spectral integration results	50
Table	4.4	Calculated and Observed Chemical Shifts of #11-TLC-4-1	58
Table	4.5	Mobile phase gradient HCG-NP-6 for separation of the soluble products of suberin	62
Table	4.6	HPLC HCG-NP separation of soluble products from suberin Degradation	67
Table	4.7	HPLC method NP-1 for separation of soluble products of suberin degradation	68
Table	4.8	HPLC method HCG-NP-9 for separation of soluble products of suberin	69
Table	4.9	HPLC method NP-1	73
Table	4.10	Aromatic proton and carbon chemical shifts of compound **3**	80
Table	4.11a	HMBC correlations of aromatic compound **3**	80
Table	4.11b	HMBC correlations to aliphatic chains in aromatic compound **3**	80
Table	4.12	RP-HPLC gradient Eluents for separation of aromatic mixtures	82
Table	4.13	RP-HPLC gradient eluents for separation of MeOH soluble Triglycerides	85

Table	5.1	Theoretical and experimental NMR data for compound **1**, glycerol	90
Table	5.2	gHMBC correlations of aliphatic compound **1**, glycerol	90
Table	5.3	gHMBC correlation of aliphatic compound **2**, triglyceride	95
Table	5.4	Structure calculation for the major peaks in sample 5-24-#5-HPLC-f3	105
Table	5.5	Structure calculation for the major peaks in sample 5-24-#5-HPLC-f4	106
Table	5.6	Structure calculation for the major peaks in sample 5-24-#5-HPLC-f5	107
Table	5.7	Structure calculation for the major peaks in sample 5-24-#7	108
Table	5.8	Chemical shifts and Integration of sample aliphatic compound **3**	121
Table	5.9	MS data for aliphatic compound **3**	123
Table	5.10	Proton and carbon chemical shifts of aromatic compound **2-1**	137
Table	5.11	gHMBC correlation of aromatic compound **2-1**	137
Table	5.12	proton and carbon chemical shifts of aromatic compound **2-2**	139
Table	5.13	gHMBC correlation of aromatic compound **2-2**	139
Table	5.14	gHMBC correlations of aromatic compound **3-f6**	152
Table	5.15	gHMBC correlations of aromatic compounds in group **4-1**	161
Table	5.16	HMBC correlations of aromatic compounds in group **4-2**	161

List of Figures

Figure 1.1	Schematic picture of suberin	2
Figure 1.2	Aliphatic monomeric constituents of suberized tissues	6
Figure 1.3	Phenolic constituents of suberized tissues	7
Figure 2.1	Typical HPLC system	17
Figure 2.2	The basic COSY pulse sequence	22
Figure 2.3	Pulse sequence for HMQC	23
Figure 2.4	Pulse sequence for HMBC	24
Figure 2.5	Basic schematic of mass spectrometry	26
Figure 2.6	Main parts of a Mass Spectrometer	27
Figure 2.7	Electron Ionization Spectrometry	28
Figure 2.8	Simple schematic of Electron Impact Ionization	29
Figure 2.9	Simple schematic of Atmosphere Pressure Ionization	30
Figure 2.10	Electrospray Ionization process	31
Figure 2.11	Atmosphere Pressure Chemical Ionization process	32
Figure 2.12	Atmospheric Pressure Photo Ionization Process	33
Figure 2.13	Typical 1100 LC/MSD system	35
Figure 4.1	Flow chart for analysis of depolymerized suberin	44
Figure 4.2	Percent yield vs. depolymerization time with 1.5 M KOH treatment	46
Figure 4.3	Percent yield vs. fraction number with 0.5 M KOH-methanol	48
Figure 4.4	Percent yield vs. fraction # with 1.0M KOH-methanol	49

Figure 4.5	600 MHz ^{1}H NMR of soluble products treated with 1.0 M KOH for varying times	50
Figure 4.6	TLC spectrum of fraction #11 from treatments for different depolymerization times	51
Figure 4.7	Comparison of 600 MHz ^{1}H NMR of seven fractions #11	52
Figure 4.8	Upfield region of the 600 MHz ^{1}H NMR spectrum of #11-TLC-4	53
Figure 4.9	Two-dimensional COSY NMR spectrum of #11-TLC-4	54
Figure 4.10	150 MHz ^{13}C NMR spectrum of #11-TLC-4	54
Figure 4.11	Two-dimensional gHSQC NMR spectrum of #11-TLC-4	55
Figure 4.12	Two-dimensional gHMBC NMR spectrum of #11-TLC-4	56
Figure 4.13	Tentative structure of #11-TLC-4-1	57
Figure 4.14	ACD calculation of carbon NMR spectrum for #11-TLC-4-1	59
Figure 4.15	ACD calculation of proton NMR spectrum for #11-TLC-4-1	60
Figure 4.16	Elastic Light Scattering Detection (ELSD) of HPLC for fraction #11	61
Figure 4.17	Flow chart for HPLC separation of fraction #11	63
Figure 4.18	HPLC comparison of different fractions from fraction #11	64
Figure 4.19	600 MHz ^{1}H NMR of different fractions from HPLC separation of sample #11	65
Figure 4.20	ELSD from separation of #11-HPLC-4	66
Figure 4.21	600 MHz ^{1}H NMR of #11-HPLC-4	66
Figure 4.22	ELSD of #11-HPLC-4-2 before and after the separation	67
Figure 4.23	HPLC separation of #11-HPLC-4-2-3	68

Figure 4.24	600 MHz ^{1}H NMR spectrum of four fractions from #11-HPLC-4-2-3	69
Figure 4.25	ELSD for HPLC separation of #11-HPLC-4-2-3-2	70
Figure 4.26	ELSD and UV of #11-HPLC-4-2-3-2-f5	71
Figure 4.27	600 MHz ^{1}H NMR of #11-HPLC-4-2-3-2-f5	71
Figure 4.28	600 MHz ^{1}H NMR of #11-HPLC-4 and its fractions	72
Figure 4.29	600 MHz ^{1}H-NMR spectrum of #11-HPLC-2 before HPLC separation	72
Figure 4.30	ELSD of #11-HPLC-2 separation	73
Figure 4.31	600 MHz ^{1}H NMR of different fractions from #11-HPLC-2	74
Figure 4.32	Partial flow chart of separation of 10-5-#4, illustrating extraction with hexane	75
Figure 4.33	^{1}H NMR corresponding to partial flow chart for separation of 10-5-#4	76
Figure 4.34	600 MHz ^{1}H NMR spectrum of aromatic compound **3**	78
Figure 4.35	gHSQC spectrum of aromatic compound **3**	79
Figure 4.36	gHMBC spectrum of aromatic compound **3**	79
Figure 4.37	Provisional structure of aromatic compound **3**	81
Figure 4.38	RP-HPLC separation of aromatic mixture	83
Figure 4.39	600 MHz ^{1}H NMR spectrum of aromatic compound 3 before and after HPLC separation of fractions f3, f6, and f7	83
Figure 4.40	600 MHz ^{1}H NMR of sample 5-24-#7	85
Figure 4.41	RP-HPLC separation of 5-24-#5	85

Figure 4.42	600 MHz ^{1}H NMR of 5-24-#5-HPLC-f3	86
Figure 4.43	600 MHz ^{1}H NMR of 5-24-#5-HPLC-f4	86
Figure 4.44	600 MHz ^{1}H NMR of 5-24-#5-HPLC-f5	86
Figure 5.1	600 MHz ^{1}H NMR spectrum of aliphatic compound **1**, glycerol	88
Figure 5.2	gHSQC spectrum of aliphatic compound **1**, glycerol	89
Figure 5.3	gHMBC spectrum of aliphatic compound **1**, glycerol	89
Figure 5.4	Tentative structure of aliphatic compound **2**, a triglyceride	91
Figure 5.5	600 MHz ^{1}H NMR spectrum of aliphatic compound **2**	92
Figure 5.6	gHSQC spectrum of aliphatic compound **2**	92
Figure 5.7	COSY spectrum of aliphatic compound **2**	94
Figure 5.8	gHMBC spectrum of aliphatic compound **2**	94
Figure 5.9	gHMBC spectrum of the triglyceride with jnxh=5	97
Figure 5.10	gHMBC spectrum of the triglyceride with jnxh=8	97
Figure 5.11	ACD calculation of proton spectrum for the triglyceride	98
Figure 5.12	ACD calculation of carbon spectrum for the triglyceride	99
Figure 5.13	600 MHz ^{1}H NMR spectrum of a standard triglyceride, tristearin	100
Figure 5.14	gHMBC spectrum of a standard triglyceride, tristearin	101
Figure 5.15	gHSQC spectrum of a standard triglyceride, tristearin	101
Figure 5.16	COSY spectrum of a standard triglyceride, tristearin	102
Figure 5.17	Backbone of a triglyceride	104
Figure 5.18	APCI-MS spectrum of sample 5-24-#5-HPLC-f3	105
Figure 5.19	APCI-MS spectrum of sample 5-24-#5-HPLC-f4	106
Figure 5.20	APCI-MS spectrum of sample 5-24-#5-HPLC-f5	107

Figure 5.21	APCI-MS spectrum of sample 5-24-#7	109
Figure 5.22	APCI-LC/MSD ion trap MS^n spectrum of tristearin	110
Figure 5.23	Fragmentation pathways for tristearin using the LC/MSD ion trap	111
Figure 5.24	APCI-LC/MSD ion trap MS^n spectrum of 5-24-#5-HPLC-f4	112
Figure 5.25	Fragmentation pathways for 5-24-#5-HPLC-f4 using LC/MSD ion trap	113
Figure 5.26	Structure and fragmentation pathways for 5-24-#5-HPLC-f4 using LC/MSD ion trap	114
Figure 5.27	Fragmentation pathways for 5-24-#5-HPLC-f4 using LC/MSD ion trap	115
Figure 5.28	APCI-LC/MSD ion trap MS^n spectrum of 5-24-#5-HPLC-f3	116
Figure 5.29	Structure and fragmentation pathways for 5-24-#5-HPLC-f3 using LC/MSD ion trap	117
Figure 5.30	600 MHz ^1H spectrum of aliphatic compound **3**	119
Figure 5.31	gHMQC spectrum of aliphatic compound **3**	120
Figure 5.32	gHMBC spectrum of aliphatic compound **3**	120
Figure 5.33	Positive MS spectra of aliphatic compound **3**	122
Figure 5.34	Negative MS spectrum of sample aliphatic compound **3**	122
Figure 5.35	LC-APCI-MS at time 4.7 min for #11-TLC-4	130
Figure 5.36	LC -APPI-MS at time 5.2 min for #11-TLC-4	131
Figure 5.37	LC-ESI-MS at time 5.2 min for #11-TLC-4	132
Figure 5.38	The structure of aromatic compound **1**	134
Figure 5.39	600 MHz ^1H NMR spectrum of aromatic compound **2**	

	before and after solvent extraction	135
Figure 5.40	gHMQC spectrum of aromatic compound **2**	136
Figure 5.41	COSY spectrum of aromatic compound **2**	136
Figure 5.42	gHMBC spectrum of aromatic compound **2**	138
Figure 5.43	Provisional structure of aromatic compound **2-1**	138
Figure 5.44	Provisional structure of aromatic compound **2-2**	140
Figure 5.45	APCI-LC-MSD ion trap MS spectrum of aromatic compound **2**	142
Figure 5.46	APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound **2**	143
Figure 5.47	APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound **2**	143
Figure 5.48	The possible fragmentation pathway of aromatic compound **2**	144
Figure 5.49	APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound **2**	145
Figure 5.50	APCI-LC/MSD ion trap Fragmentation pathways for peak m/z 186.9	146
Figure 5.51	Fragmentation pathways for peak186.9	146
Figure 5.52	APCI-LC-MSD ion trap MS-MS^2-MS^3 spectrum of standard compound 10-Hydroxydecanoic acid	147
Figure 5.53	600 MHz ^{1}H NMR spectrum of aromatic compound **3-f6**	149
Figure 5.54	gHMBC spectrum of aromatic compound **3-f6**	150
Figure 5.55	A portion of the gHMBC spectrum of aromatic compound **3-f6**	151
Figure 5.56	Provisional structures of aromatic compound **3-f6**	152
Figure 5.57	ACD calculation of proton spectrum for the compound shown in Figure 5.56 (b)	153
Figure 5.58	ACD calculation of carbon spectrum for the compound shown in Figure 5.56 (b)	154

Figure 5.59 APCI-LC-MSD ion trap UV and MS spectrum of aromatic compound 3-f6 155

Figure 5.60 APCI-LC-MSD ion trap MS^n spectrum of m/z at 391.2 156

Figure 5.61 APCI-LC-MSD ion trap MS^n fragmentation of the molecular ion at 391.2 m/z 157

Figure 5.62 600 MHz ^1H NMR spectrum of aromatic compound **3-f3** 158

Figure 5.63 600 MHz ^1H NMR spectrum of aromatic compounds **4** 159

Figure 5.64 gHSQC spectrum of aromatic compounds **4** 159

Figure 5.65 gHMBC spectrum of aromatic compounds **4** 160

Figure 5.66 Partial structure of aromatic compound **4-2** 162

Figure 5.67 600 MHz ^1H spectrum of aliphatic compound **4** 163

Chapter 1. INTRODUCTION

1.1 Why study suberin?

Suberin is a plant cell wall biopolymer. The occurrence of suberin in plants is variable: it can be present in epidermal tissues of underground plant parts such as the cell walls of the periderm, the outer tissues of stems in a secondary growth stage, or in the endodermis, exodermis, root and tuber phellem, and bundle sheath cells.[1] Suberin is also deposited in response to wounding and pathogenic attack, as a barrier against water and solute loss and as protection against opportunistic invasion. Many plant cell walls become suberized to avoid desiccation and control diffusion of water between the atmosphere and tissue. Waxes are associated with this polymeric material to provide an effective diffusion barrier. [2,3] The specificity of suberin location relates to its function as a protective polymer involved in thermal insulation, in water-loss prevention, or as a barrier to pathogenic attack.[4]

The definition of suberin has remained elusive despite many years of study. This uncertainty persists due to its typical deposition between the primary wall and the plasma membrane prior to secondary wall formation (Fig. 1.1), making it impossible to obtain pure structurally intact suberin in its native state without other associated materials, or to determine its specific chemical composition.[5] For example, some authors have reported that cells from tree cork layers have three main structural components: suberin, lignin, and carbohydrates.[6-9] This description implies that the poly(phenolic) domain in these tissues contains another separate structural component, namely lignin instead of suberin. Other literature takes the position that the term suberin should only be used to describe the aliphatic domain and its associated phenolics. [6,7,9-16]

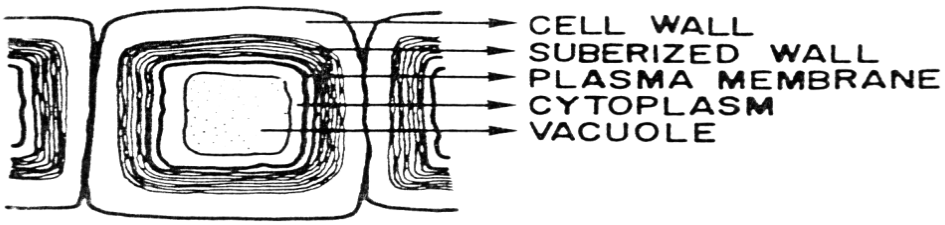


Figure 1.1. Schematic picture of suberin.[1]

Now it is generally recognized that suberin is a biomacromolecule that contains two distinct domains, a polyphenolic domain (SPPD) and a polyaliphatic domain (SPAD). The polyaromatics are located predominantly within the primary cell wall, and the polyaliphatics are located between the primary cell wall and the plasmalemma. The solid-state NMR studies of suberin by our group and others have supported the contention that the two domains are covalently cross-linked,[8,16-18] but direct evidence is lacking.

The exact qualitative and quantitative composition of suberin monomers varies among different plant species, moreover some variability occurs within the same plant species.[19] The degree of suberization in tissues is also different, ranging from highly suberized cork cells in the outer bark of woody plants to the radial-wall restricted suberization of the Casparian band in the endodermis.

The study of molecular structure in suberin should provide compositional and structural information about how the monomer units are linked together and to waxes and cell-wall matrices, and how the suberin polymer is synthesized during wound healing. Ultimately, the structure of suberin can inform the design of crop protection strategies and may lead to the development of better synthetic waterproofing materials.

1.2 What is known about the structure of suberin?

Many studies of the composition and molecular structure of suberin have been made on the natural potato or wound-induced periderm of potato.[17,20] The potato is an excellent model system to study suberization in plants because of the importance of the potato as a food stuff, the relatively high proportion of suberin in the potato peel, and the ease of obtaining suberized tissue by wound-healing, in which the tissues are suberized quickly and uniformly in response to wounding.[21-23] Other relevant structural studies of suberin have been done on commercial cork.[11,24]

Beginning 25 years ago, suberin was described as a heteropolymer of aliphatic and phenolic compounds that are linked by ester bonds.[3] Two general approaches have been used customarily to study the molecular structure of suberin. **First**, depolymerization or extraction methods have been used along with nuclear magnetic resonance (NMR) and mass spectrometry (MS) to identify monomeric and oligomeric fragments and then conceptually reconstruct the original polymer.[20,21,25] The standard procedure for the determination of suberin composition is: isolation by enzymatic treatment of the tissue with cellulases and pectinases; then dewaxing by exhaustive solvent extraction to get rid of soluble materials (waxes); finally isolation of suberin as the solvent-insoluble material. Then, the suberin is depolymerized by procedures that cleave ester bonds, e.g. alkaline hydrolysis, transesterification with methanol containing boron trifluoride or potassium hydroxide, as well as reductive cleavage with lithium aluminium hydride. The resulting monomers are separated chromatographically and identified by MS.

Second, in our group and that of Lewis, solid-state NMR has been used to identify structural features directly within intact suberin.[17,22,26,27] The information obtained from depolymerization approaches may be compromised by incomplete breakdown, because each technique can only break certain linkages; or by complete breakdown of connectivities between the monomeric units; or because time consuming multistep chemical procedures are usually required. The solid-state NMR strategy provides information that is more directly applicable to the native plant material, and the structural complexity of suberized tissue that makes it difficult to deduce detailed molecular information can be reduced by labeling with spectroscopically sensitive isotopes.[22,23] However, chemical heterogeneity reduces the spectral resolution and the particular NMR chemical shifts cannot identify interunit crosslinks definitively. Based on solid-state NMR studies of suberin, the two domains are viewed as distinct from each other and chemically linked when found within the same cells. [8,16,17,28,29]

1.2.1 Monomeric units: aliphatics

The aliphatic portion of suberin (SPAD) has been described as a three-dimensional polyester network. When examined using histochemical staining coupled with electron microscopy, the aliphatic domain of suberized tissues typically displays a characteristic multilamellar structure.[1,30,31] It has a different morphology from cutin (the aliphatics in suberized tissues are located between the primary cell wall and the plasmalemma, whereas cutin is formed as a continuous layer on the surface of external tissues). Its unique structure and chemical composition also differ from cutin, even though both are polyesters of modified fatty acids, and both have water retention and protective roles in plants. The characteristic aliphatic components of suberin are ω-

hydroxyfatty acids (C16-C28) and fatty α,ω-diacids (C16-C26). Small amounts of very long chain fatty acids and alcohols (C18-C30) may also be present. By contrast, the chemical composition of cutin is characterized by C16 and C18 families of mono-, di- and trihydroxy fatty acids as well as midchain epoxides.[1,30,32] Fig. 1.2 shows the main aliphatic structural components of the SPAD, including 1-alkanols, ω-hydroxyalkanoic acids, α,ω-dioic acids, mid-chain epoxide- and di- and tri-hydroxy-substituted octadecanoates.[11,19]

Early studies of suberin proposed that the aliphatic constituents form polyesters to which the phenolic compounds are cross-linked in a lignin-like manner (Fig.1.3 (**1,2,3)**). This first provisional model of suberin depicted α,ω-diacids and ω-hydroxyacids interesterified in a linear form and further interesterified to hydroxycinnamic acids and alcohols, with these latter constituents interconnected in a lignin-like manner. In this view, suberin would be an aliphatic-aromatic heteropolymer.[4] With depolymerization techniques, the composition of aliphatics and waxes from suberized tissues has been clarified. In considering the structures shown in Fig 1.2, however, it is important to remember that the proportion of the individual aliphatic components differs widely from species to species.

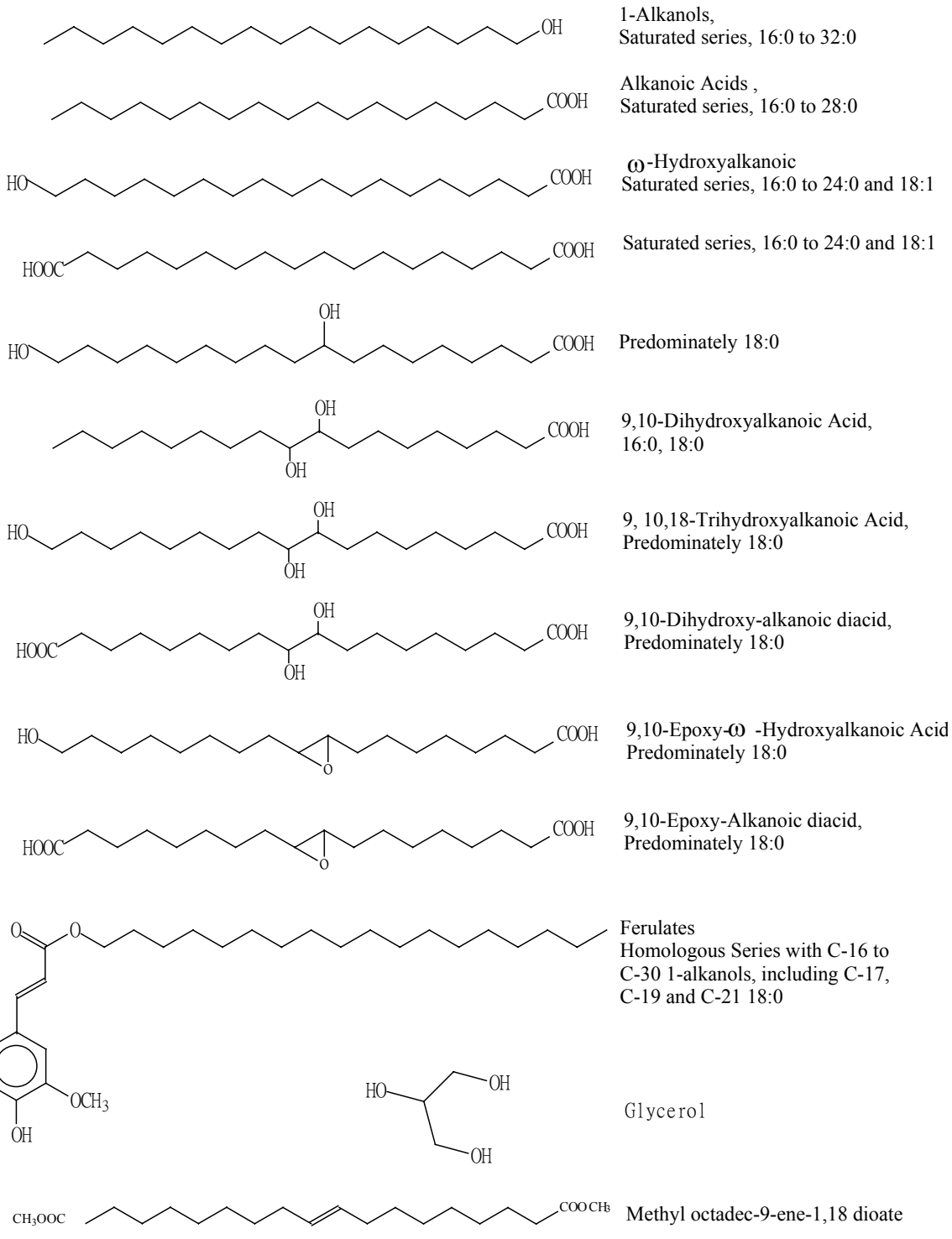
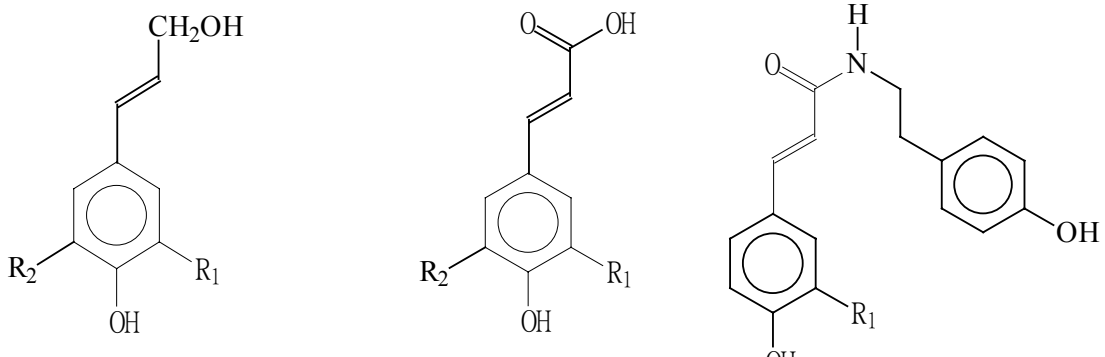


Figure 1.2. Aliphatic monomeric constituents of suberized tissues. [33]

1.2.2. Monomeric units: phenolics

Historically, the poly(phenolic) domain of suberin has been likened to lignin: a three-dimensional network of oxidatively-coupled monolignols (Fig. 1.3, e.g.. *p*-coumaryl alcohol **a**, coniferyl alcohol **b** and sinapyl alcohol **c**). However, with the help of the same techniques used for the analysis of lignin (alkaline nitrobenzene oxidation, non-destructive solid-state NMR, and thioacidolysis), the aromatic part of suberin was also shown to include structural features of a hydroxycinnamic acid-derived polymer primarily comprised of ferulic acid (Fig.1.3, **f**) and N-feruoyltyramine (Fig.1.3, **i**). Thus, the poly(phenolic) domain of suberin is not simply a lignin-like polymer composed of monolignols.[34,35] This hydroxycinnamic acid-derived polymer composition may explain the role of suberized tissues as anti-microbial barriers in plants since hydroxycinnamates are known to have anti-microbial properties.[36,37]



a R_1=R_2=H, *p*-Coumaryl Alcohol
b R_1=OH, R_2=H, Coniferyl Alcohol
c R_1=R_2=OCH_3, Sinapyl Alcohol
d R_1=R_2=H, *p*-Coumaric Acid
e R_1=OH, R_2=H, Caffeic Acid
f R_1=OCH_3, R_2=H, Ferulic Acid
g R_1=R_2=OCH_3, Sinapic Acid
h R_1=H, *p*-Coumaroyltyramine
i R_1=OCH_3, Feruloyltyramine

Figure 1.3. Phenolic constituents of suberized tissues. [33]

Therefore, the picture of phenolics in potato suberin has been revised: in addition to some monolignols,[38,39] the poly(phenolic) domain also contains a significant proportion of covalently cross-linked hydroxycinnamic acids (e.g. Fig. 1.3 , **d–g**)

[14,22,40,41,41,42] and their derivatives (e.g., feruloyltyramine **i**).[34] These units are presumably linked to the remainder of the matrix through either ester or amide linkages involving their aromatic ring carbons or phenolic hydroxyl moieties (see Section 1.3). Moreover, our group found a significant amount of monolignols in potato periderm by solid-state NMR.[23] In maize, feruoyltyramine is not detectable in suberized endodermal or hypodermal cell walls of roots, which suggests that the aromatic part of the suberin may show compositional variations in different plants or/and different tissues.[40] The new view of suberin implies other biosynthetic steps different from those involved in lignification. Furthermore, the observation that the phenolic and aliphatic parts occupy different domains in the suberin polymer has led to the hypothesis that suberin is composed of different polymers that are cross-linked to each other and whose aromatic part is also cross-linked to the primary cell wall.[5,23]

1.2.3 Monomeric units: glycerol

Though glycerol was found in early studies of suberized tissues,[43] it was ignored for a long time. The newly accumulated evidence that glycerol plays a major role in its structure has further expanded the knowledge of suberin. These studies suggest that glycerol acts as a linkage not only between acyl units, but also between the SPPD and the SPAD, because of its several possible esterifying positions.[44,45] It is known that glycerol is a major monomer released after suberin depolymerization. The analysis of water-soluble compounds of the depolymerization reaction confirmed earlier reports that glycerol is a principal monomer of suberin in oak, cotton, and potato.[13,46,47] Partial methanolysis with calcium oxide as catalyst has shown that glycerol may be present as monoacylglycerol esters of alkanoic acids, α, ω-diacids, and ferulic acid[13] as well as a

diglycerol ester linked to α, ω-diacids at both ends. Other suberin analyses have shown the presence of glycerol,[48-50] and its quantity has been positively correlated with the suberization process.[50]

The role of glycerol in the macromolecular assembly of suberin is now thought to involve the interlinking of the long-chain acid monomers. For example, the esterification of hydroxyl groups of glycerol with ω-hydroxyfatty acid or the esterification of α, ω-diacids (one of the main monomer families) with glycerol at both ends could form the backbone for growth of the macromolecular polyester structure. These latter structural units, in turn, could continue the linear extension of the polymer through esterification to another ω-hydroxyacid, or terminate the aliphatic structure through esterification to ferulic acid. Thus, it has been proposed that suberin is a poly-(acylglycerol)-polyester.[45] Glycerol may also cross-link with the aromatic and aliphatic suberin components, corresponding to the electron-opaque and electron-translucent suberin lamellae in a three-dimensional network, whereas aliphatic and aromatic suberin monomers may only form a linear polymer on their own.[46] Glycerol and glycerol-fatty acid esters are not only possible components of the suberin polymer but have also been reported in suberin-associated waxes in suberized cotton fibers.[44,48] Glycerol is a common metabolic intermediate. It is still not clear how it is incorporated into the suberin macromolecule as an ester of fatty acids and hydroxycinnamates.

Triacylglycerols in animal tissues are well known to have hydrophobic and insulating properties. The fact that suberin may be an acylglycerol-based polyester can help to rationalize the protective properties of cells and tissues where they are important constituents. We also know that acylglycerols and their derivatives, like

glycosylacylglycerols and phosphoglycerides associated with proteins, are the structural elements of both plant and animal biomembranes.[51] The existence of acylglycerols in the form of a poly(acylglycerol) polyester in the outermost cells of plants would further solidify the important role of this group of lipids as barrier structures in living bodies.

1.2.4 A Model for Macromolecular Architecture

A new structural model of suberin was suggested [33] by combining the various evidence obtained from potato tubers, cotton, *Quercus suber*, and *Pseduotsuga Meniesii,* with partial models proposed by Gil and Lopes.[28,29,52] In this conceptual model of suberin, the SPPD is described as a monolignol-hydroxycinnamoyl copolymer that is covalently linked to primary cell-wall carbohydrates. Some of the phenolic components (Fig 1.3) have been isolated, for example a small amount of monolignols (Fig 1.3, **a-c**),[38,39] a significant amount of hydroxycinnamic acids (e.g. Fig.1.3 , **d–g**) [14,22,40-42] and their derivatives (e.g., feruloyltyramine, **i**).[34] Information regarding the covalent connectivity between those phenolic components and the linkage to the cell wall is relatively sparse. Hydroxycinnamic acids are reported to be covalently linked through non-ester linkages,[22] and also attached to carbohydrate moieties.[17,23] Thus based on a presumed similarity to lignin, and supported by dimers isolated from *Q. suber* cork after permanganate oxidation,[53] these units are presumably linked to each other through either ester or amide linkages involving their aromatic ring carbons, or phenolic hydroxyl moieties.

The SPPD and SPAD are depicted to be either covalently linked through glycerol[45] (though no direct evidence has been reported), or they are viewed as directly linked between acyl groups and both the SPPD and carbohydrates of the primary cell

wall.[28,29,52] Other esters within the suberin polymer provide possible linkages between the SPPD and SPAD, such as the linear interesterication of ω-hydroxyacids. The ω-hydroxyacid-ω-hydroxyacid inter-ester structure was proposed to play a major role in cutins,[4] and a linear ester dimer of the 16-hydroxyhexadecanoic acid was found in the tomato peel cutin after a partial hydrolysis reaction.[54] Nevertheless, covalent linkages between the aliphatic polyester part and the polyaromatics in suberin could be abundant. Ferulic acid and other related small phenolics are always found in the suberin mixtures. Ferulic acid has also been found esterified to the ω-hydroxyacids.[12,55] A preliminary report of such a structure in potato suberin has been made.[56] Based on the structure of ω-hydroxyacids, it is possible that they are esterified through their acid end-group to glycerol and through their primary hydroxyl to ferulic acid. They may be the units in the suberin glycerol-aliphatic polyester that serve to link it to the surrounding polyaromatics. Such a trimeric diester, glycerol-ω-hydroxyacid-hydroxycinnamic acid, has already been identified in the extractives of green cotton fiber.[48] The linking between the glyceryl-aliphatic suberin polymer and the neighboring polyaromatics could occur via glycerol by esterification to ferulic acid.

In summary, suberized tissues possess a specialized cell-wall modification that consists of a cell wall-associated poly(phenolic) domain with a covalently attached poly(aliphatic) domain located between the cell wall and plasma membrane (Fig. 1.1). Whereas these two domains are spatially distinct, they coexist in suberized cells and are presumably linked together covalently to form the macromolecule suberin. Their chemical composition is unique and distinct from the related polymers cutin and lignin.

1.3 The objectives of this research

Knowledge of the chemical structure, macromolecular assembly and biosynthesis of suberin is accumulating because much research has been done on the subject of this protective plant polymer. But from the biosynthetic and macromolecular assembly points of view, the polymerization steps of suberin biosynthesis are still unknown, and the nature of the covalent connectivities within suberin and to the cell wall is quite speculative. For instance, the two domains of suberin SPPD and SPAD are only partially characterized; information on different inter-unit linkages (covalent bonding) between the aliphatic monomers (oligomers), between the SPPD and SPAD (which has never been verified), and new aromatic monomers and oligomers is critical to map out the essential crosslinking structure of the potato suberin and the polyester-cell wall linkage of suberin. This limitation hampers our understanding of the ability of the biopolymer to protect cell wall tissues. These gaps leave areas that need further study.

Our objective is to obtain oligomers of both aliphatic and aromatic components, exploring depolymerization conditions that produce bigger oligomers in higher proportions and developing new analytical separation and characterization strategies to identify them, in order to understand the macromolecular structure of suberin from potato wound periderm.

Specific goals include the following:

1. Separation method development: explore depolymerization conditions to find the best way to obtain the desired oligomers and aromatics. Develop appropriate analytical procedures to separate the desired products, e.g. solvent extraction, HPLC.

2. Identification method development: for oligomers and aromatics of potato suberin obtained from 1., combine different analytical tools (HPLC, NMR, LC/MS/MS) to identify them, including HPLC and MS^n methods.

3. Solid-state NMR spectroscopy has predicted a covalent cross-link between the aromatic and aliphatic domains found in suberized tissue.[17,23] If we would separate and identify aromatic-aliphatic oligomers, that would provide confirmatory evidence for the solid-state NMR conclusions. Similarly, isolation of aromatic-sugar structures would support prior proposals.

4. In the new structural model of suberin,[33] the SPPD and SPAD are depicted to be either covalently linked through glycerol [45] or directly linked via acyl or other functional groups, but those specific types of linkages have not been confirmed. We hope to find some supporting evidence for the above new structural model or to define how it must be modified.

5. The molecular structures determined for potato suberin may be compared with plant cutins, shedding light on their biosynthetic relationships, giving clues to their related functions, and establishing a database to aid future studies of these biopolymers.

Chapter 2. INTRODUCTION TO METHODS

2.1 Introduction to Chromatography

2.1.1 Basic Principles of Chromatography

Russian botanist Mikhail Tsvet invented the first chromatographic technique in 1901. He first used the term *chromatography* in print in 1906 in his papers because they dealt with colored pigments. His contribution turned the phenomenon into a method of scientific analysis.[57]

Chromatography describes a family of analytical chemistry techniques for separation and analysis of mixtures. It involves two phases, a stationary phase and a mobile phase. The mobile phase can be a gas or a liquid, whereas the stationary phase must be a liquid or solid. A mixture of analytes is introduced into the mobile phase and is carried through the system by it. As the mobile phase passes over and through the stationary phase, the components of the mixture equilibrate or partition between the two phases, resulting in differential migration rates through the system. That is, different components are retarded in their passage through the system in proportion to their interaction with the sorbent. The least retarded component emerges first; the most strongly retained component elutes last. When components pass through the system at different rates they become separated in time. Each component has a characteristic time to pass through the system, called a retention time R_t. Chromatographic separation is achieved when the retention time of a given analyte differs from that of other components in the sample of interest.

A chromatogram is the visual output of the chromatograph (plot of detector signal vs. time/volume). Different peaks on the chromatogram correspond to different components of the separated mixture.

2.1.2 Column Chromatography (CC)

Column chromatography is used in both analytical and preparative applications. In column chromatography, the solid adsorbent (stationary phase, traditionally silica gel or alumina) is placed in a vertical glass column. The mixture to be analyzed is applied to the top of it, and the liquid mobile phase (eluent) is added to the top and allowed to flow down through the column by either gravity or external pressure. Equilibrium is established for the solute between the adsorbent and the eluting solvent. Because the components in the mixture have different interactions with the stationary and mobile phases, they will be carried along with the mobile phase at varying speeds so that separation may be achieved.[58]

2.1.3 Thin-Layer Chromatography (TLC)

Thin-layer chromatography (TLC) is also used in both analytical and preparative applications. Compared with column chromatography, it is a microscale, sensitive, fast and simple technique, often used to monitor the progress of organic reactions and to check the purity of synthetic or degradative products.

TLC involves spotting the sample near one end of a sheet of glass or plate that is coated with a thin layer of adsorbent (usually silica gel or alumina). The TLC plate is placed in a covered container with a shallow layer of liquid mobile phase, which slowly rises up the TLC plate by capillary action. The different components in the mixture move up the plate at different rates due to differences in their partitioning behavior between the

mobile liquid phase and the stationary phase. It is easy to detect colored components visually. The colorless spots are visualized with ultraviolet light or by placing the plate in either iodine vapor or phosphomolybdic acid in ethanol.[58]

2.1.4 High-Performance Liquid Chromatography (HPLC)

HPLC is one of the most powerful chromatographic techniques. It is not limited by sample volatility or thermal stability. The result of separation is easily visualized; sample recovery is good; detection is highly sensitive (down to nanograms in favorable cases). It can separate a wide variety of compounds ranging from macromolecules and ionic species, labile natural products, polymeric materials, to high molecular weight polyfunctional groups. It also involves specific interactions between sample molecules and both the stationary and mobile phases. It offers a greater variety of stationary phases than GC, TLC, or CC, allowing more diverse selective interactions and possibilities for separation.[57,59]

In HPLC, solid and liquid samples are dissolved in an appropriate solvent and injected into the instrument. The analyte is forced through a column (stationary phase) by a liquid (mobile phase) at high pressure, which decreases the time the separated components remain on the stationary phase and thus the time they have to diffuse within the column. The components of the mixture are separated by selective retention within the stationary phase. As the analytes flow through, the detector deflects on a chart. Available detectors include ultraviolet (UV), refractive index (RI), mass spectrometry (MS), fluorescence, electrochemical, infrared (IR), and evaporative light scattering detector (ELSD).

The mechanisms of separation include adsorption, partition, ion pairing, ion exchange, and size exclusion. Two types of HPLC methods are distinguished by the relative polarity of the stationary phase and mobile phase. Normal phase (NP-HPLC) means that the polarity of the stationary phase (e.g. silica) is higher than that of the mobile phase (e.g. hexane or tetrahydrofuran). Reverse phase (RP-HPLC) means that the polarity of the stationary phase (e.g. chemically bonded C_{18} or C_8 alkyl chains) is less than that of the mobile phase (e.g. mixture of water with either methanol or acetonitrile). For both NP-HPLC and RP-HPLC, mixtures of solvents may be used in two elution modes: isocratic (the composition of mobile phase is constant) or gradient (composition of the mobile phase can be made to change in a predetermined way during the elution). With both techniques compounds are eluted in order of polarity. For RP-HPLC, the most polar compounds come out first, whereas for NP-HPLC, the least polar come out first. RP-HPLC dominates most current applications because it has a broad scope that allows sample types with a wide range of polarities and molecular weights to be separated, features rapid mobile phase column equilibration, and is easier, faster and more reproducible experimentally.

Figure 2.1. Typical HPLC system (www.agilent.com).

The HPLC instrumentation system consists of an injector or injection unit, degasser, pump, column, and detector as shown in Figure 2.1. The functions of each part are as follows: the solvent degasser is used to prevent bubbles in the mobile phase; the programmable high-pressure pump mixes the solvents in the prescribed ratios and pumps them through the column and past the detector; the column compartment houses and thermostats the HPLC column; the injection unit (autosampler or injector) draws prescribed volumes from sample vials and injects them onto the column; the in-line detector (UV or other) monitors the absorbance of the column effluent at a selected wavelength at regular intervals.

The very important stationary phase of columns used today contains uniform, porous particles with nominal diameters of 10, 5, 3, or 1.8 μm. Because of this property, the efficiency of HPLC separation surpasses CC and TLC. That is also the reason that a high-pressure pump is necessary to pass the mobile phase through the column.

2.2 Introduction to Nuclear Magnetic Resonance (NMR)

2.2.1 Basic Principles of NMR

Nuclear Magnetic Resonance Spectroscopy (NMR) is a physical phenomenon based upon the magnetic property of an atom's nucleus. It is one of the principal techniques used to obtain information about molecular structure.

NMR is a phenomenon that occurs when the nuclei of certain atoms (such as ^{1}H or ^{13}C) are immersed in a static magnetic field and exposed to a second oscillating magnetic field. Some nuclei experience this phenomenon, and others do not, dependent upon whether they possess a property called spin (I). The nuclei with non-zero spin quantum number (e.g. I=1/2) may be thought to spin around an axis but with only two

possible orientations, *up* and *down*. When there is no external magnetic field present, there is no energy difference between up and down spin states. If an external magnetic field $\mathbf{B_0}$ is applied, then the spin state that gives rise to a nuclear magnetic bipole aligned with the external magnetic field has the lower energy. The energy difference ΔE (= $h\nu$) lies in the range of radio frequencies. The populations also differ for the two states, with their ratio given by the Boltzmann distribution [N_{up} /N_{down} = $\exp(-\Delta E/kT)$]. The lower-energy state that has an orientation of nuclear spins parallel to the magnetic field is slightly more populated. The nuclear magnetic resonance phenomenon occurs when nuclei aligned with (or against) an applied field are induced to absorb (or release) energy and change their spin orientation. The energy absorption is a quantized process, and the energy absorbed must equal the energy difference between the two states involved.

$$E_{absorbed} = h\nu$$

The energy separation between two different states *(Larmor frequency)* is given by:

$$\Delta E = h \gamma \mathbf{B_0} / 2\pi$$

where γ is the gyromagnetic ratio of the spin **I**, and $\mathbf{B_0}$ is the magnitude of the applied static magnetic field. The observed NMR frequency ν may be expressed in terms of the gyromagnetic ratio and the applied field.

$$\nu = \gamma \mathbf{B}_0 /2\pi$$

The power of NMR spectroscopy relies on the fact that not all protons in a molecule resonate at the same frequency. This is because the protons in a molecule are surrounded by electrons and exist in slightly different electronic environments from one another. So different protons in a molecule each resonate at slightly different frequencies. This means that each different proton has a characteristic chemical shift (δ) from a reference proton,

typically expressed in dimensionless units, parts per million (ppm). Because the chemical shift is obtained by measuring Hz (from the reference) divided by ν_0, the operating frequency of the spectrometer, it is independent of spectrometer conditions. Therefore it is viewed as a characteristic property of the proton. For the spectrum of a given compound, the intensity of ^{1}H NMR absorption is strictly proportional to the number of nuclei giving rise to the absorption. This property is very important for structure determination.[57,60,61]

Spin-spin coupling refers to scalar coupling between spins, which is characterized by the coupling constant J (a measurement of the interaction between a pair of protons). The value of J is field independent and is characteristic of each bond type. It depends on chemical environment (e.g. the number and types of bond through which the nuclei are coupled). Usually it is smaller than other magnetic interactions, e.g., for a ^{13}C-^{1}H pair, J can be as much as 150 Hz; for protons separated by three bonds, J is in the range 2-9 Hz, but for geminal protons, J is ~12 Hz.[61]

In contrast to the spin-spin coupling, the dipole-dipole coupling refers to magnetic interactions through space instead of through chemical bonds. Its magnitude is proportional to the internuclear distance. Theoretically, the dipolar-dipolar coupling also depends on $(1-3\cos^2\theta)$, where θ is the angle between the internuclear vector and the applied field $\mathbf{B}_0$.[61] In liquids this coupling is averaged to zero.

2.2.2 Solution-State NMR

In solution-state NMR, the direct dipole-dipole interactions are averaged to zero by rapid intra- and intermolecular motions so that narrow ^{1}H NMR spectral lines are observed. The most reliable information for structure determination in a one-dimensional

NMR spectrum includes the spin-spin couplings between NMR-active nuclei, chemical shifts, and integrals of the proton resonances for tentative identification of proton types and their relative numbers. The most commonly used solvent-suppression technique is presaturation of solvent signal during the recycle delay using a weak radio-frequency (rf) field. Presaturation is very simple to implement and is very effective for getting a general picture of the 1D ^{1}H spectrum.

Two-dimensional (2D) NMR spectroscopy also plays an increasingly important role in the assignment of complex NMR spectra and the determination of unknown molecular structures.

Homonuclear correlation: The ^{1}H-^{1}H **COSY** (**CO**rrelated **S**pectroscop**Y**) experiment consists simply of two non-selective 90^{o} rf pulses with a variable delay t_1 between them. Data are acquired after the second pulse during another period t_2 (Figure 2.2).[61,62] In the COSY experiment, the first 90^{o} pulse excites the coupled spins. Both chemical shift and J coupling evolve during the t_1 period. It is the precession under the influence of scalar coupling during t_1 that produces the cross (off-diagonal) peaks. Using the second 90^{o} pulse, magnetization is transferred by scalar coupling. Transfer may occur between transitions of the same nucleus, corresponding to 2D crosspeaks on the diagonal line. Additionally, off-diagonal peaks show the chemical shifts of the resonances that are coupled through chemical bonds. Protons that are more than three chemical bonds apart give no cross (off-diagonal) signals because the ^{4}J coupling constants are close to 0. Therefore, only signals from protons that two or three bonds apart are visible in a COSY spectrum.[60,61] The data are usually presented as a contour plot with ^{1}H chemical shifts on each axis.

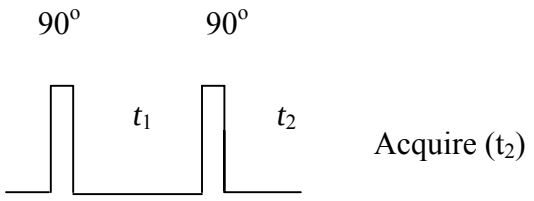


Figure 2.2. The basic COSY pulse sequence.

Heteronuclear correlation: ^{1}H-^{13}C **HMQC** (**H**eteronuclear **M**ultiple **Q**uantum **C**oherence) and **HSQC** (**H**eteronuclear **S**ingle **Q**uantum **C**oherence) experiments correlate the chemical shift of a proton with the chemical shift of directly bonded nuclei (X: carbon or nitrogen). These experiments utilize one-bond couplings and feature high-sensitivity ^{1}H detection. The idea behind HMQC is 'reverse correlation', i.e. heteronuclear correlation with detection of protons. One simple version of the HMQC sequences is shown in Figure 2.3.[62] It starts with a proton pulse followed by a delay Δ_1 of 1/(2J). During that delay, the proton magnetization (coupled to the heteronuclei) acquires anti-phase coherence. A 90° pulse applied to the X nuclei generates proton-carbon multiple quantum coherence which evolves under the influence of the X chemical shift during the t_1 evolution time (the proton chemical shift and heteronuclear coupling are refocused by using a 180° pulse at mid-evolution time). At the end of the evolution period, another 90° pulse on X transfers the magnetization back into detectable (antiphase) single quantum proton magnetization, which is then left to evolve during another delay Δ_2 of 1/(2J) to refocus the antiphase magnetization. During the proton detection period, the X-nuclei coupling interactions can be decoupled.[61,62]

Both HSQC and HMQC provide the same information. The differences are technical and involve the signal-to-noise ratio. HSQC has a sensitivity-enhanced version

that provides a factor of two improvement for methine protons only. The data are usually presented as a contour plot with ^{1}H and ^{13}C chemical shifts on each axis.

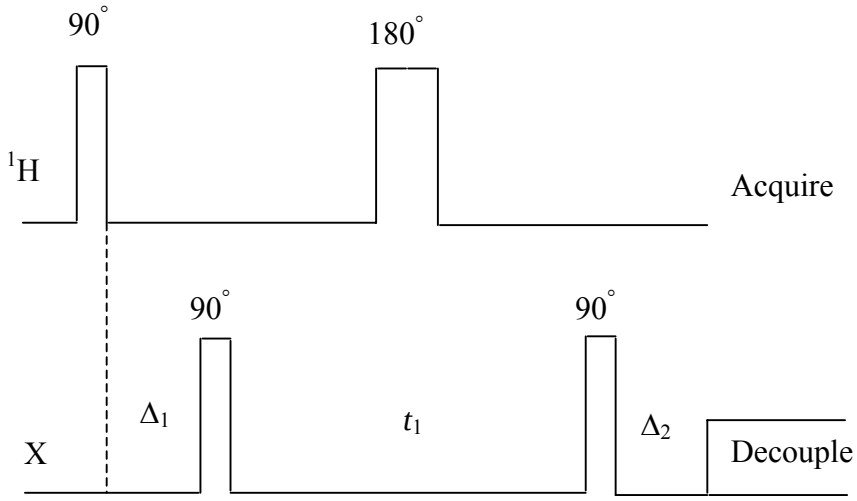


Figure 2.3. Pulse sequence for HMQC.

^{1}H-^{13}C **HMBC** (**H**eteronuclear **M**ultiple **B**ond **C**orrelation) is an experiment that correlates proton nuclei with carbon nuclei (or other X) that are separated by more than one bond, again using ^{1}H detection. The pulse sequence is shown in Figure 2.4.[62] The first 90° pulse applied to protons excites the ^{1}H spins, whereas the first 90° pulse applied to carbon serves as a low-pass J filter when the duration of the delay Δ_1 is adjusted to 1 / 2 ($^{1}J_{CH}$), i.e. it suppresses one-bond correlations and passes the long-range correlations that have smaller coupling. The delay Δ_2 is chosen so that the second 90° pulse on ^{13}C creates multiple (zero- and double-) quantum coherence for the long-range couplings. After the evolution time t_1, the magnetization is converted back into detectable single-quantum proton magnetization. The 180° proton pulse effectively removes the effect of the proton shifts from the t_1 modulation frequency by interchanging the zero- and double-

quantum components. Thus after the final 90° (^{13}C) pulse, the ^{1}H signals that originate from ^{1}H- ^{13}C multiple-quantum coherence are modulated by ^{13}C chemical shift and homonuclear proton couplings. Signals from protons that do not have a long-range coupling to ^{13}C are removed by phase cycling of the second final 90° (^{13}C) pulse [62]. The experiment is closely related to HMQC and uses the same principles to convert transverse magnetization into zero and double quantum coherence, with the exception that the delay (Δ_2) is matched to the inverse of the long-range coupling constant $^{n}J_{CH}$ and there is a J filter (Δ_1) to suppress one-bond correlations. The data are usually presented as contour plot similar to HMQC.

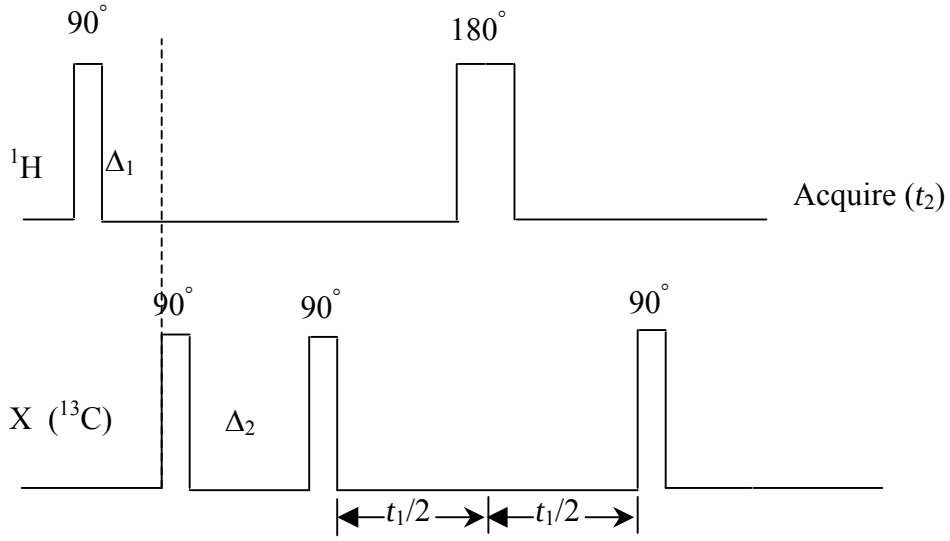


Figure 2.4. Pulse sequence for HMBC.

2.2.3 ACD database software

ACD/ChemSketch (Advanced Chemistry Development, Inc., Toronto, Canada) is a chemical drawing and spectral prediction software package from Advanced Chemistry Development, Inc. It can be used alone or integrated with other applications. ChemSketch

is used to draw chemical structures, reactions and schematic diagrams. It can also be used to design chemistry-related reports and presentations. This software is used to predict NMR spectra, chemical shifts, and coupling constants for all types of 1D and 2D NMR experiments. The predicted results can confirm the inferred structures obtained from NMR spectra of unknown samples.

The ACD software designers have compiled and checked data from thousands of literature articles to build databases of over 165,000 assigned structures for ^{1}H and ^{13}C, 8000 for ^{15}N, 13,800 for ^{19}F, and 22,600 for ^{31}P. With these assignments and structures as a basis and using proprietary correlation algorithms, the fragment-based predictions are claimed to offer superior performance to any rules-based systems (www.acdlabs.com). Typical confidence intervals for the predicted chemical shifts are ±0.1 ppm and ±1 ppm for ^{1}H and ^{13}C nuclei, respectively.

2.3 Introduction to Mass Spectrometry (MS)

2.3.1 Basic Principles of MS

Mass spectroscopy is an analytical technique for measuring the mass of molecules. Mass spectrometers separate and detect ionized atoms or molecules according to their mass-to-charge ratio (m/z). Different molecules also have distinctive fragmentation patterns that provide additional structural information. MS experiments have been used in industrial and academic laboratories for the analysis of inorganic, organic, polymer and bioorganic chemicals.

Mass spectrometers have three basic parts, namely the **ionization source**, the mass selective **analyzer**, and the **ion detector**. Figure 2.5 shows a basic schematic of mass spectrometry. The sample is introduced into the ionization source of the instrument,

in which the ions are generated by inducing either the loss or the gain of a charge (e.g. electron ejection, protonation, or deprotonation). The ions are directed electrostatically into the mass analyzer, where they are separated according to their **mass-to-charge ratios (m/z)**. The separated ions are detected and the resulting signal is sent to a data system, in which the output shows a plot of relative intensity vs. m/z. The analyzer and detector are operated under high vacuum to give the ions a reasonable chance of traveling from one end of the instrument to the other without any hindrance from air molecules [57] (www.masspec.scripps.edu).

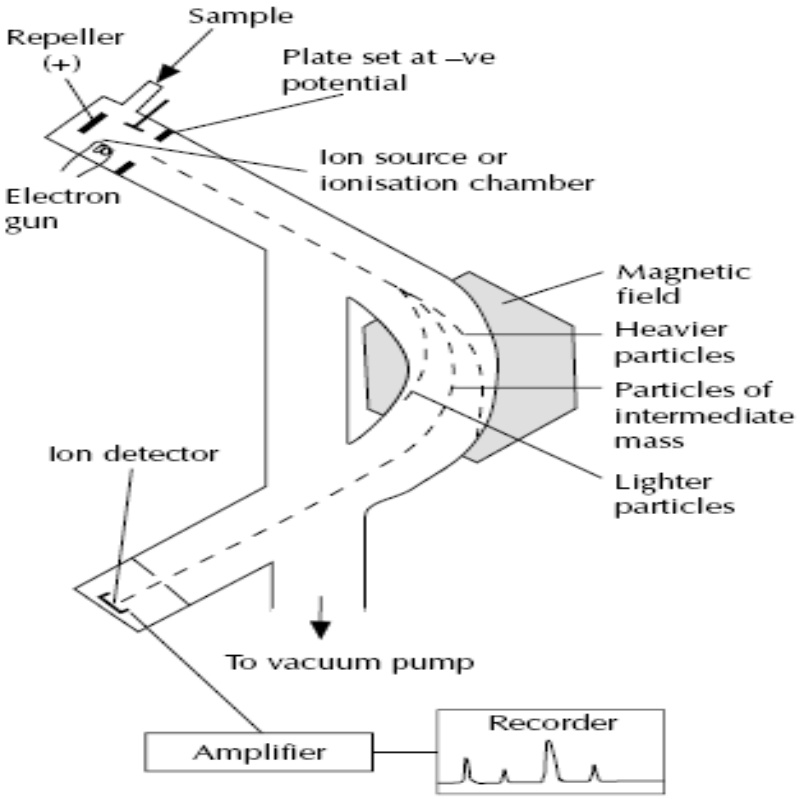
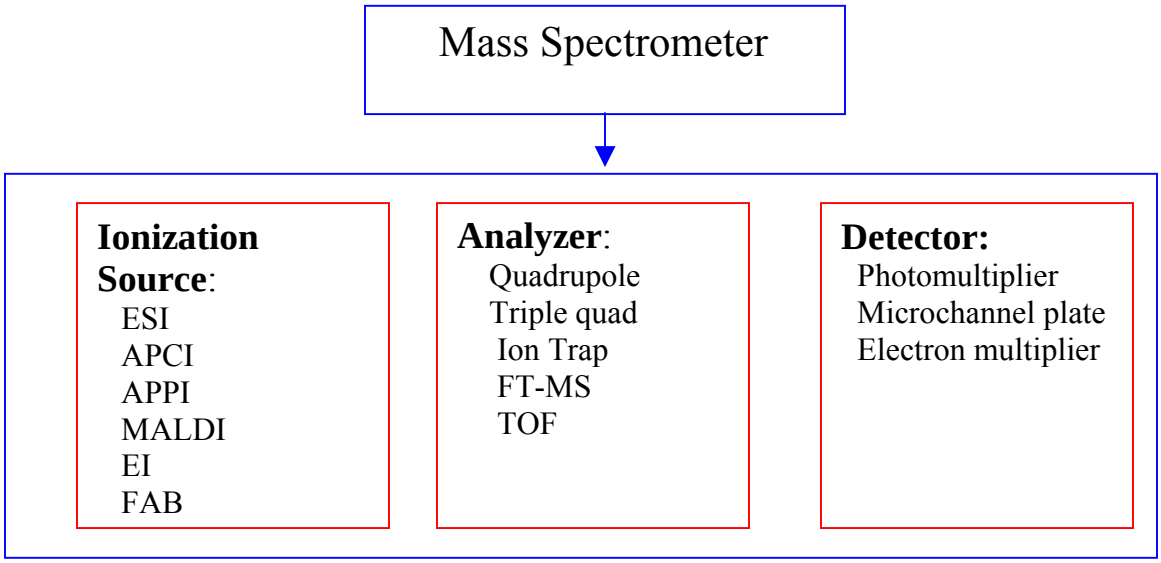


Figure 2.5. Basic schematic of mass spectrometry (www.chemsoc.org).

Figure 2.6 depicts different ionization sources, analyzers and detectors. Depending on the different measurement requirements and available equipment, we can choose

different ionization source and analyzer combinations. The sample can be injected directly into the ionization source, or it can undergo some type of chromatography before ionization, e.g. GC-MS, LC-MS, CE (capillary electrophoresis)-MS.



ESI: Electrospray Ionization

APCI: Atmospheric Pressure Chemical Ionization

APPI: Atmospheric Pressure Photoionization Ionization

MALDI: Matrix Assisted Laser Desorption Ionization

EI: Electron Impact

FAB: Fast Atom Bombardment

TOF: Time-Of-Flight

Figure 2.6. Main parts of a Mass Spectrometer (www.masspec.scripps.edu).

Tandem (MS^n) mass spectrometers are instruments that have more than one analyzer and so can be used for structure determination and sequencing studies by collisionally generating and identifying fragmentation ions inside the mass spectrometer. Two, three or more analyzers can be incorporated into commercially available tandem instruments, and the analyzers do not necessarily have to be of the same type, in which case the instrument is a hybrid one. More popular tandem mass spectrometers include those of the quadrupole-quadrupole, magnetic sector-quadrupole , and more recently, the quadrupole-

time-of-flight geometries. In Tandem MS, the ion of interest is selected with the first analyzer (MS-1) and then collided with inert gas atoms; the fragments generated by the collision are separated by a second analyzer (MS-2), and so on. In Ion Trap and Fourier transform experiments, the experiments are carried out in one analyzer, and the various events are separated in time, not in space (www.masspec.scripps.edu). The information can be used to sequence oligomers, to determine structure and connectivity of polysaccharides, and to determine the position and structure of fatty acids in complex lipids.

2.3.2 Ionization Source

Electron Impact Ionization (EI): EI introduces a high energy into molecules. It is known as "hard" ionization. In EI, one bombards the molecule with a high-energy electron beam (Figure 2.7).

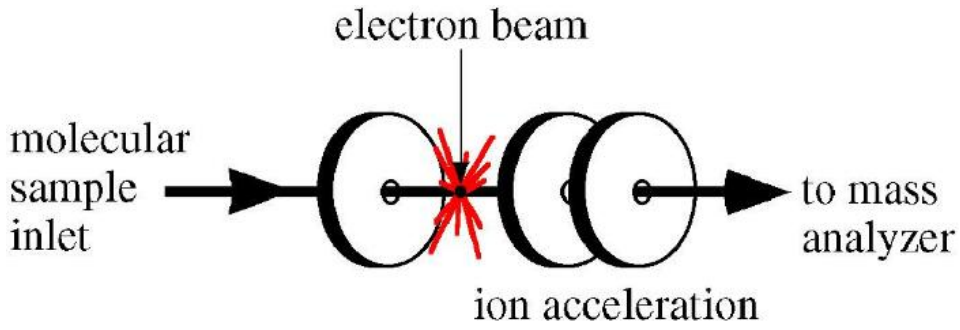


Figure 2.7. Electron Ionization Spectrometry (www.masspec.scripps.edu).

An electron that strikes a molecule may impart enough energy to remove another electron to produce singly charged ions. The molecules fragment and the positive ion paths are bent by an applied magnetic field. Ions with low mass for given charge will be deflected most by this field and will collide with the walls of the analyzer. High momentum ions will be deflected the least and will thus enter the analyzer. Ions having the proper mass-

to-charge ratio will follow the path of the analyzer, exit through the slit and collide with the collector. This event generates an electric current, which is then amplified and detected. A charged molecule that remains intact is called the **molecular ion,** which reveals the molecular weight of the compound of interest. Electrons that impart enough energy to the molecular ion can cause that ion to break into fragments, which also yields structural information. EI is most useful for compounds with molecular weight below 400, since larger molecules tend to degrade thermally during vaporization.

Before the advent of techniques that could be carried out at atmospheric pressure (API, e.g. APCI, APPI, ESI), electron impact (EI) was the main choice for ionization; the solvent elimination and ionization steps are separate. Most instruments nowadays feature API techniques, in which solvent elimination and ionization steps are combined in the source and occur at atmospheric pressure. Figures. 2.8 and 2.9 show the difference between the two ionization sources. API is known as a "soft" ionization method, as the sample is ionized by the addition or removal of a proton, with very little extra energy remaining to cause fragmentation. By raising the voltage applied to the sampling cone, extra energy can induce fragmentation if desired.

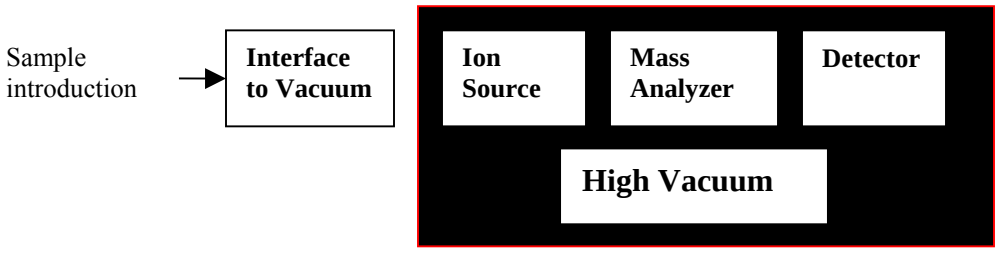


Figure 2.8. Simple schematic of Electron Impact Ionization (www.waters.com).

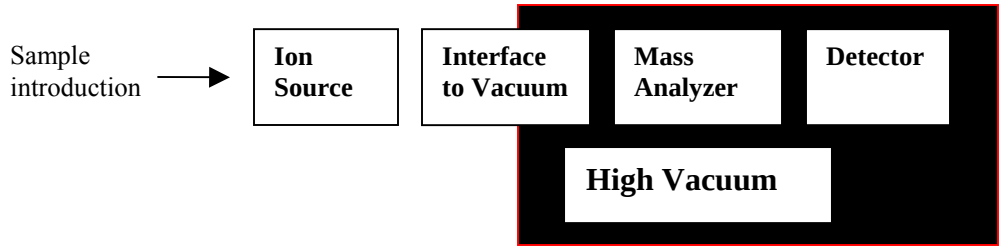


Figure 2.9. Simple schematic of Atmosphere Pressure Ionization (www.waters.com).

Electrospray Ionization (ESI) is one of the Atmospheric Pressure Ionization (API) techniques and is suitable for the sensitive analysis of polar molecules such as peptides, proteins, metal complexes and polymers. The sample solution is pumped through a capillary, where a high voltage is applied to the tip. The unique feature of ESI is that it generates ions directly from solution by creating a fine spray of highly charged droplets in the presence of a strong electric field, making it a convenient mass-based detector for HPLC and leading to routine liquid chromatography-mass spectrometry techniques. As a charged droplet diminishes in size by solvent evaporation assisted by a drying gas (nitrogen), the electric charge density on its surface increases. The mutual repulsion between like charges on this surface becomes so great that it exceeds the forces of surface tension, and ions begin to leave the droplet. Eventually only charged sample ions (either positive or negative) are released from the droplets, some of which pass through the sampling cone into an intermediate vacuum region, and through a small aperture into the analyzer (Figure 2.10). ESI can create singly or multiply-charged gaseous ions, so its applicability extends to higher molecular weight. The charge number depends on the composition and pH of the solvent and the chemical nature of the sample. ESI is preferred for compounds that are ionic or very polar or thermolabile.

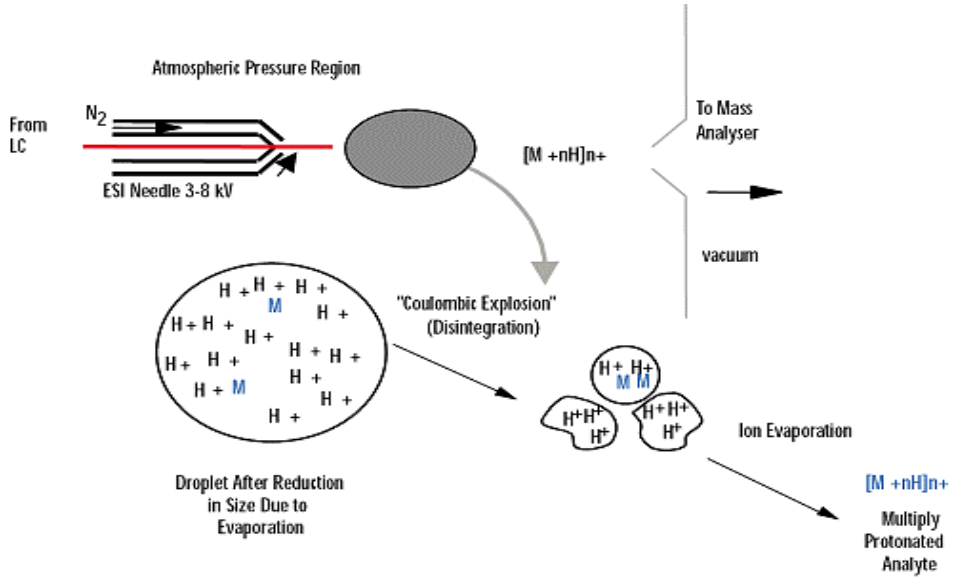


Figure 2.10. Electrospray Ionization process (www.waters.com).

Atmospheric pressure chemical ionization (APCI):

APCI is a similar interface to that used for ESI. The APCI source (Figure 2.11) contains a heated vaporizer that facilitates rapid desolvation/vaporization of the droplet. A corona discharge is used to ionize the analyte in the atmospheric pressure region. Chemical ionization of molecules is very efficient at atmospheric pressure due to the high collision frequency. APCI is widely used for analyzing relatively non-polar, semivolatile low molecular weight molecules.

The term "corona effect" describes the partial discharge around a conductor placed at a high potential. This leads to ionization and electrical breakdown of the atmosphere immediately surrounding the conductor. St. Elmo's fire is an example of a naturally occurring corona. In the case of an APCI source, the atmosphere surrounding the corona electrode consists mainly of the vapor generated from the HPLC eluent, nitrogen, and the analyte molecules. The eluent vapors are ionized by the corona effect, and they react chemically with the analyte molecules in the gas phase. Depending on the

solvents, singly charged molecular ions like $[M+H]^+$, $[M+Na]^+$ and $M^{+\cdot}$ (in the case of aromatics) and/or fragments can be produced. Sodium ion comes from glassware used to process the sample (www.masspec.scripps.edu).

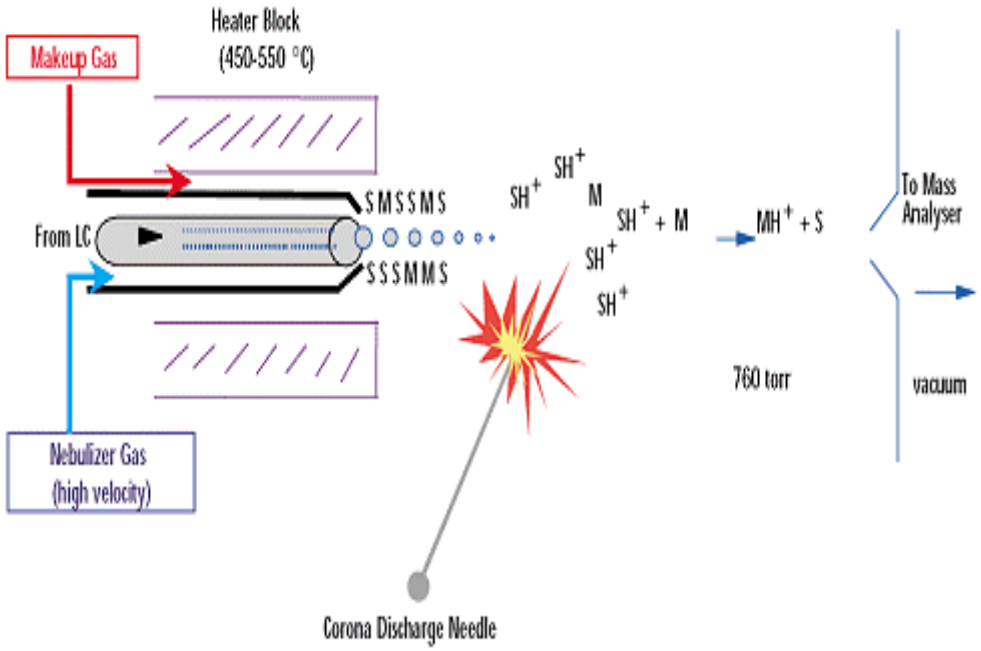


Figure 2.11. Atmosphere Pressure Chemical Ionization process (www.waters.com).

Atmospheric Pressure Photo Ionization (APPI)

APPI is a newly introduced technique. The principle is to use photons to ionize gas-phase molecules. The source (Figure 2.12) is a modified APCI source, with a UV lamp instead of a corona. Ionization can be obtained directly, or with a dopant (e.g. acetone or toluene).

Matrix-Assisted Laser Desorption Ionization (MALDI)

MALDI is a laser ionization method of vaporizing and ionizing large biological molecules such as proteins or DNA fragments. In MALDI analysis, the analyte is first co-crystallized with a large molar excess of a matrix compound (usually a UV-absorbing

weak organic acid), after which pulsed UV laser radiation of this analyte-matrix mixture results in the vaporization of the matrix that carries the analyte with it. The matrix plays a key role by strongly absorbing the laser light energy and causing the analyte to vaporize indirectly. The matrix also serves as a proton donor and receptor to ionize the analyte in positive and negative modes, respectively.

MALDI allows for determination of the molecular weight of molecules up to 500 kDa, and routinely 5 to 100 kDa (polymers, biomolecules, complexes, enzymes), depending on the analyzer. The MALDI technique can be coupled with a time-of-flight analyzer (see below) or a Fourier-transform mass spectrometer.

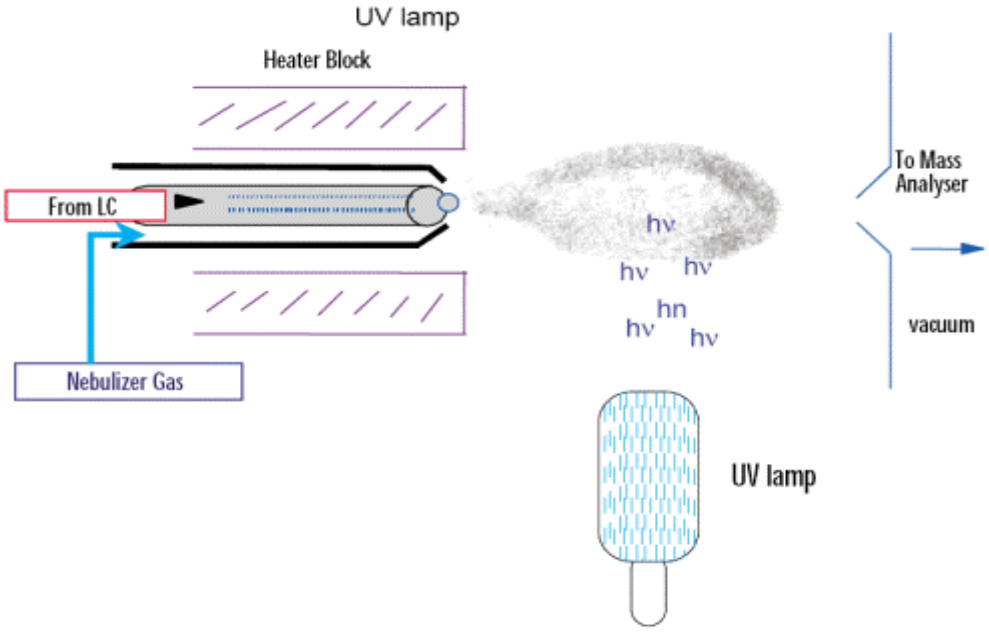


Figure 2.12 Atmospheric Pressure Photo Ionization Process (www.waters.com).

2.3.3 Ion Analyzer

The main function of the mass analyzer is to resolve the ions formed in the ionization source of the mass spectrometer according to their **mass-to-charge (m/z)** ratios. There are a number of mass analyzers currently available, as shown in Figure 2.6.

These mass analyzers have different features, including different mass range, the mass accuracy, and the achievable resolution. The compatibility of different analyzers with various ionization methods varies. For example, all of the analyzers listed above can be used in conjunction with electrospray ionization, whereas MALDI is not usually coupled to a quadrupole analyzer.

Ion Trap: The ions are trapped in a radio frequency quadrupole field. By ejecting all others ions from the trap, one ion species is isolated. The isolated ion can be fragmented subsequently by collisional activation and the fragments detected to generate a fragmentation spectrum.

Time-Of-Flight (TOF): This detection method is commonly used with MALDI. TOF analysis is based on accelerating a set of ions to a detector with the same amount of energy. Because the ions have the same energy and different mass, they reach the detector at different times. The smaller ions reach the detector earlier than the larger ones. Thus the analyzer is called time-of-flight.

2.3.4 Detector

The detector monitors the ion current, amplifies it, and then transmits the signal to the data system where it is recorded in the form of mass spectra. The **m/z** values of the ions (x-axis) are plotted against their intensities (y-axis) to show the number of components in the sample, the molecular mass of each component, and the relative abundance of the various components.

2.4 Introduction to LC-MS

Liquid Chromatography - Mass Spectrometry is a hyphenated technique, which combines the separation power of HPLC with the detection power of MS (Figure 2.13). It allows separation of complex mixtures into a series of components before introducing them into the mass spectrometer. It is used extensively for compounds that have a substantial molecular weight (higher than 400) or are too sensitive to heat to be analyzed by GC.

Figure 2.13. Typical 1100 LC/MSD system (www.agilent.com).

The most common ionization methods are ESI and APCI in positive- and negative-ion modes. The LC is done in most cases by RP-HPLC, and the buffer system should not contain involatile salts (e.g., phosphates). LC-MS can be used to separate and identify a wide variety of compounds.

Chapter 3. SAMPLE PREPARATION AND SPECTRAL ANALYSIS

3.1 Preparation and isolation of potato suberin

Biosynthesis of the plant polyester suberin occurs following the mechanical wounding of potato tubers. The wounds can be induced by slicing potatoes into disks, after which the injured tissue is allowed to synthesize the protective polymer for a period of 0-14 days. Since suberin is deposited within the cell wall, isolation of the pure material is a somewhat elusive goal. Removal of as much of the other polymeric cellular components as possible leaves behind suberized cell walls.[20,63,64] Enzymes such as cellulase and pectinase have been used to remove non-suberized cell walls and other carbohydrates. The soluble tissue components are also removed by this treatment, leaving behind an insoluble material that is highly enriched in suberized walls. The waxes associated with suberized walls and other lipids that may be present in such preparations are removed by thorough extraction with organic solvents including multi-day Soxhlet extraction. Suberization of wounded potatoes and suberin isolation followed published procedures,[18,22,65] in which about 95% of the dry mass is removed. Recently, it has been shown that trifluoroacetic acid can also be used to achieve more exhaustive cell-wall removal.[56]

Russet Burbank (Idaho) potato tubers (*Solanum Tuberosum* L.) were purchased from a local supermarket. The potato tubers were washed and scrubbed with a brush to remove surface dirt and then soaked in 25% (v/v) Chlorox bleach solution for 5-10 minutes to sterilize them. The plastic growth chambers were also briefly soaked with

25% (v/v) Chlorox bleach. All peelers, knives, forceps, gloves, and filter paper were autoclaved for about 30 minutes.

The potatoes were peeled and sliced into disks 4-5 mm thick, about 20 by 30 mm in size. Then the potato disks were rinsed briefly with double-processed tissue culture water (Sigma W-3500, St. Louis, MO). The potato disks were placed on autoclaved filter papers, which were placed in turn on plastic-covered wire mesh in a closed growth chamber to allow for air circulation on all sides. The atmosphere was kept humid by placing sterile water in weighing boats underneath the mesh. The potato disks were left to suberize at 25 $\pm$ 0.1 °C in a large refrigerated incubator for 7 days. Some water droplets were observed on the sides of the growth chambers, and the surface of the potatoes darkened noticeably after several days, as expected. Any preparations that developed mold were discarded.

When the suberization period was complete, the potato disks were collected from the growth chamber. The disks were covered by a layer of brown suberized cell walls, which was removed manually from the internal tissue. The suberized cell walls were washed and cleaned thoroughly with distilled water and a rough (wet) weight of the product was recorded.

The residual unsuberized cell walls (cellulose and pectin) were removed through a series of enzyme treatments (cellulase and pectinase). To remove extraneous cellulose, the suberized potatoes were treated with a 2.5% (w/v) solution of *Aspergillus niger* cellulase (ICN 150583, activity 110,000 units/g, Aurora, Ohio) at 37 ^{o}C for two days and at 44 ^{o}C for another two days in a shaking incubator. Upon completion, the solid suberized potato was collected using vacuum filtration and washed with water (100 mL x

3 times). To remove extraneous pectin, the suberized potato tissues were treated with 1.0 mg/mL *Aspergillus niger* pectinase (Sigma P-9179 @ 31 mg/mL and 11.8 units/mg) at 28 ^{o}C for 8 hours and at 31 ^{o}C for another 16 hours in a shaking incubator.

After the enzymatic treatments, the suberized potato tissues were dried in air. The dry suberized potato tissues were extracted with methylene chloride and methanol-methylene chloride (1:1 v/v) in a Soxhlet extractor for two days each. It was important to break up the sample into flakes rather than clumps, so that as much surface area as possible was exposed to the solvents. The suberized potato tissues were air dried in a thimble and then completely dried in a SpeedVac until the mass was constant. Depending on the geometry of the potato disks, approximately 200 g of peeled potatoes yielded about 2 g of suberized potato tissue. A CPMAS ^{13}C NMR experiment was performed to confirm that suberin peaks were clearly visible in addition to the cell-wall resonances.[65]

3.2 Generation of potato suberin aromatics and oligomers with KOH

Suberin is an intractable polyester in nature; it is frequently subjected to harsh chemical treatments for studies of molecular structure. For example, alkaline hydrolysis, alkaline nitrobenzene oxidation, cupric oxide oxidation, pyrolysis-MS and thioacidolysis are routinely used to characterize suberized tissues. Of these approaches, only alkaline hydrolysis and thioacidolysis release phenylpropanoid moieties (from primary esters and alkyl-aryl ethers, respectively) with their side chains intact, thereby allowing for a more reliable assessment of the original molecular structure.

Previous studies [66,67] established that alkaline methanolysis (KOH/CH_3OH) is the most gentle depolymerization method available to determine suberin monomeric composition, because epoxy acids can be detected indirectly in the form of

methoxyhydrins. The high temperatures and aqueous alcoholic alkaline medium used in alcoholic saponification (KOH/ROH, with water associated with the alcohol) promote the degradation of the unstable epoxy groups to *vic*-diol groups, whereas aqueous alkaline hydrolysis (KOH/H_2O) at room temperature liberates large amounts of polyphenols that interfere with the subsequent extraction of the acids into ether or chloroform.

Thus, suberin was degraded by using different concentrations of methanolic KOH solution. By adding concentrated HCl, the depolymerization was stopped at different times, and acidified to pH 4 or 5. The hydrolysis time and temperature were varied to determine optimum conditions for obtaining suberin monomers or oligomers. The mixture was filtered to remove the salt (KCl) formed; then the solvent (methanol) was evaporated to obtain the depolymerized monomers and oligomers. The residues were dissolved in $CHCl_3/CH_3OH$; the solvent soluble material was subjected to column chromatography (CC), thin-layer chromatography (TLC), and/or high-performance liquid chromatography (HPLC).

3.3 Isolation and chromatographic purification of oligomers and aromatic monomers

Chromatographic methods were used to separate the depolymerized soluble mixtures from potato subeirn. The particular methods used were CC, TLC and HPLC.

3.3.1 Preliminary separation of depolymerized suberin by CC

Depolymerized suberin soluble products were pre-purified with column chromatography using Silica gel 60 (70-230 mesh) (Mallinckrodt Baker, Inc., Paris, Kentucky). Hexane, acetone, and methanol were used as eluents of increasing polarity. All solvents were obtained from Fisher Scientific (Fair Lawn, New Jersey). Typically,

nine fractions of about 50 ml each were collected. ^{1}H NMR was used to choose fractions for further purification.

3.3.2 Separation of depolymerized suberin by TLC

The dried extract from preliminary separation by column chromatography was dissolved in $CHCl_3$ and applied to pre-coated silica gel 60 TLC plates with 10 µl micropipettes. The plates were developed using an acetone: *n*-hexane solvent system. After visualization of the plates by spraying with phosphomolybdic acid in ethanol and heating to 100-120 ^{o}C, a minimum of 1-2 µg of the soluble components (monomers or oligomers) could be detected as dark blue bands on a yellow background. The separated zones were scraped off and the compounds recovered by elution of the silica gel with solvents (e.g. $CHCl_3$ and CH_3OH).

3.3.3 Separation of depolymerized suberin by HPLC

The purification of different solvent extracts from alkaline hydrolysis of suberin was achieved using a Hewlett-Packard Model 1100 HPLC instrument equipped with a quaternary solvent delivery system. A UV detector, diode array detector (DAD), and evaporative light scattering detector (ELSD, Alltech, Burtonsville, Maryland) were used in series. The reason for using ELSD is that some of our samples lack a suitable chromophore for UV detection.

Reverse phase (RP) and normal phase (NP) HPLC separations were performed with different columns, flow rates, solvent systems, buffers, and gradient elution conditions as noted in Chapter 2. Isolated compounds were checked for purity by re-injection using the same solvent gradient and/or a different gradient. The solvents used for NP-HPLC were hexane, acetone, and isopropanol. The solvents used for RP-HPLC

were acetonitrile, methanol, and water with or without buffers. All HPLC grade solvents were purchased from Fisher Scientific.

Typical ELSD conditions for NP-HPLC were: drift tube temperature 50 ^{o}C, nitrogen gas flow rate 2.0 SLPM (standard liters per minute), nitrogen gas pressure 11.0 psi. A representative NP-HPLC solvent system used a binary gradient (hexane-acetic acid-isopropanol, 7:2:991-200:2:778).[68] A representative RP-HPLC solvent system used a gradient system (methanol-water) at different ratios (0:100 to 100:0).

Separation of depolymerized suberin by solvent extraction methods was combined with ^{1}H NMR to guide the process. The choice of solvent is crucial. At the beginning, different solvents were tried. The main solvents used were hexane, acetonitrile, methanol and ethyl acetate.

3.4 Solution-state NMR characterization of soluble products of suberin

1D H and 2D homonuclear COSY, heteronuclear HMQC and HMBC NMR spectra were acquired on a Varian (Palo Alto, CA) ^{Unity}INOVA 600 spectrometer operating at ^{1}H and ^{13}C frequencies of 599.95 MHz and 150.87 MHz, respectively. Samples were dissolved in either $CDCl_3$ or CD_3OD to provide a field-frequency lock and $CDCl_3$ containing 1% tetramethylsilane (TMS) to provide an internal chemical shift standard. The solution-state NMR spectra were acquired using a Varian HCN probe optimized for ^{1}H detection.

2D NMR experiments were used to establish through-bond connectivities within and between monomer units of the oligomer structure. A two-dimensional ^{1}H-^{1}H experiment (COSY) was used to delineate chemically bonded proton networks within each monomer unit, which generally requires experiment times of 2-5 hours. In addition,

bonded proton-carbon pairs and long-range proton-carbon bonded interactions were identified with the two-dimensional ^{1}H-^{13}C experiments HSQC and HMBC, respectively, corresponding to typical experiment times of 5 and 22 hours, respectively.

3.5 Mass spectral analysis of suberin

Despite extensive studies on the monomeric composition of suberin by GC-MS, most of the published papers present semiquantitative results. Moreover, the identification of suberin monomers by mass spectrometry, usually after derivatization to the corresponding methyl ester trimethylsilyl ethers, has been limited to a particular fraction of suberin that is volatile enough to be analyzed by GC. Therefore, for most of oligomers and aromatics obtained during our analysis of suberin, LC/MS plays an important role.

Mass spectral data were acquired on two instruments: an Agilent Technology 1100 Series LC/MSD Model G1946D (Agilent, Santa Clara, CA; CUNY Mass Spec Facility at Hunter College) and an Agilent Technology 1100 Series LC/MSD SL system(Agilent, Santa Clara, CA) (Mass Spec Facility at College of Staten Island). Electrospray ionization (ESI), atmosphere pressure chemical ionization (APCI) and atmosphere pressure photoionization (APPI) using either positive or negative mode were used to generate the ions.

For the Hunter College spectrometer, ESI ionization was carried out with a drying gas temperature of 200 ^{o}C, nebulizer pressure of 40 psi and drying gas flow rate of 13L/min, while the HPLC flow rate was 400 μl/min. APCI and APPI ionization were carried out with a drying gas temperature of 300 ^{o}C, nebulizer pressure of 60 psi and a drying gas flow rate of 5L/min, while the HPLC flow rate was 800 μl/min. The capillary

voltage was set to 4000 volts, and the scanned mass range was between 200 and 1500 m/z. Samples were introduced into the mass spectrometer in different solvent systems (as noted), depending on which ion source was used.

For the CSI spectrometer, ESI ionization was carried out with a drying gas temperature of 325 ^{o}C, nebulizer pressure of 10-70 psi and a drying gas flow rate of 4-12 L/min depending on the choice of HPLC flow rate (1-1000 μl/min). APCI and APPI ionization were carried out with a drying gas temperature of 350 ^{o}C, nebulizer pressure of 60 psi and a drying gas flow rate of 5L/min, while the HPLC flow rate was 200-1500 μl/min. For the above three ionization sources, the capillary voltage was set to 3500 volts. The mass range scanned was between 200 and 2200 m/z. Samples were introduced into the mass spectrometer in different solvent systems (as noted), depending on which ion source was used.

Chapter 4. RESULTS AND DISCUSSION: SEPARATION AND PROVISIONAL IDENTIFICATION OF DEPOLYMERIZATION PRODUCTS

4.1 Preliminary study of potato suberin

The well-established KOH-methanol method[66,67] was used to depolymerize suberin and followed by Column Chromatography (CC) and Thin-Layer Chromatography (TLC) to separate the soluble products. After separation and purification, NMR and Mass Spectrometry (MS) were used to characterize the compounds.

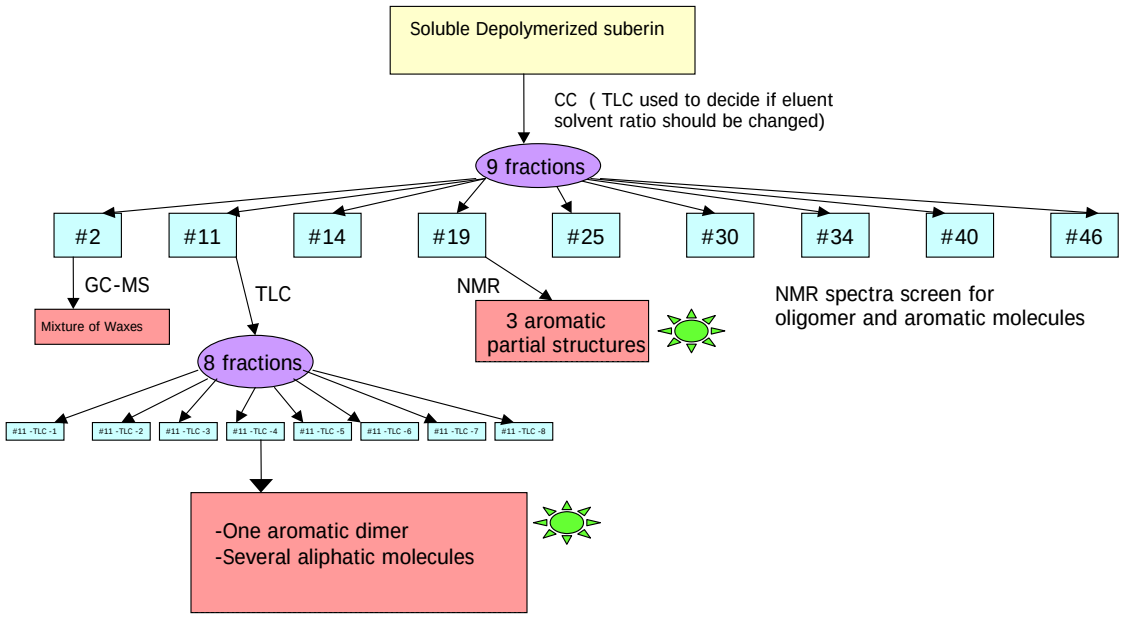


Figure 4.1. Flow chart for analysis of depolymerized suberin.

For soluble depolymerized suberin, CC was used to do further preliminary separation, in which hexane and acetone were used as a mixed solvent system. First pure hexane was used, then the relative amount of acetone was increased progressively. After pure acetone, methanol was used as the final effluent. TLC was used as a detector for CC, and nine fractions were collected (Figure 4.1).

4.1.1 The effect of depolymerization conditions

The most difficult aspect of this research is that the mixtures obtained have hundreds of components. To find the optimal conditions for production of our target products (aliphatic oligomers and aromatic compounds), different depolymerization conditions were tested.

4.1.1.1 Depolymerization time

At room temperature, 200 mg of potato periderm suberin was depolymerized with 1.5 M KOH-methanol solutions for each of nine different times (0.25 hr, 0.5 hr, 1 hr, 2 hr, 3 hr, 6 hr, 12 hr, 24 hr, 3 days). Since the depolymerized suberin was still a mixture, it was further separated by CC and TLC. Nine fractions were obtained as described above and weighed (Table 4.2). ^{1}H NMR was conducted on all nine fractions. Since fractions #11 and #19 showed NMR signals indicative of oligomers and aromatics in relatively large quantity (see Figure 4.2 and description below), they were studied further.

Table 4.1 The effect of time on percent yield of depolymerized products

Depolymerization time	Depolymerized soluble part (mg)	Unreacted suberin residue (mg)	Depolymerized portion (%)
0.25 hr	14.3	180.3	**7.15**
0.5 hr	19.1	180.6	**9.55**
1 hr	17.6	170.7	**8.80**
2 hr	16.3	179.7	**8.15**
3 hr	24.8	170.0	**12.4**
6 hr	20.8	164.8	**10.4**
12 hr	19.9	159.5	**9.95**
24 hr	18.7	160.0	**9.35**
3 days	27.6	147.3	**13.8**

Based on Table 4.1, the general trend is that as depolymerization time increases, depolymerized percent yields also increase, but the highest overall yield is less than 15% even after 3 days (filtration and handling procedures may make the numbers uncertain).

This result nevertheless exceeds the yield reported by Graca et al. [13] The longest times (12 hr, 24 hr, 3 days) do not yield more of the target products (oligomers and aromatic components). At shorter times, the dependence on depolymerization time is not obvious, which implies that only modest portions of the intact suberin structure can be broken down with methanolic KOH depolymerization methods.

Figure 4.2 shows the masses of each fraction for different treatment times (masses were used to assess the yields, since the number of moles could not be determined in the absence of molecular weight information). These trends indicate that shorter treatment times produce more of fractions #11 and #19. As these fractions appear to be rich in

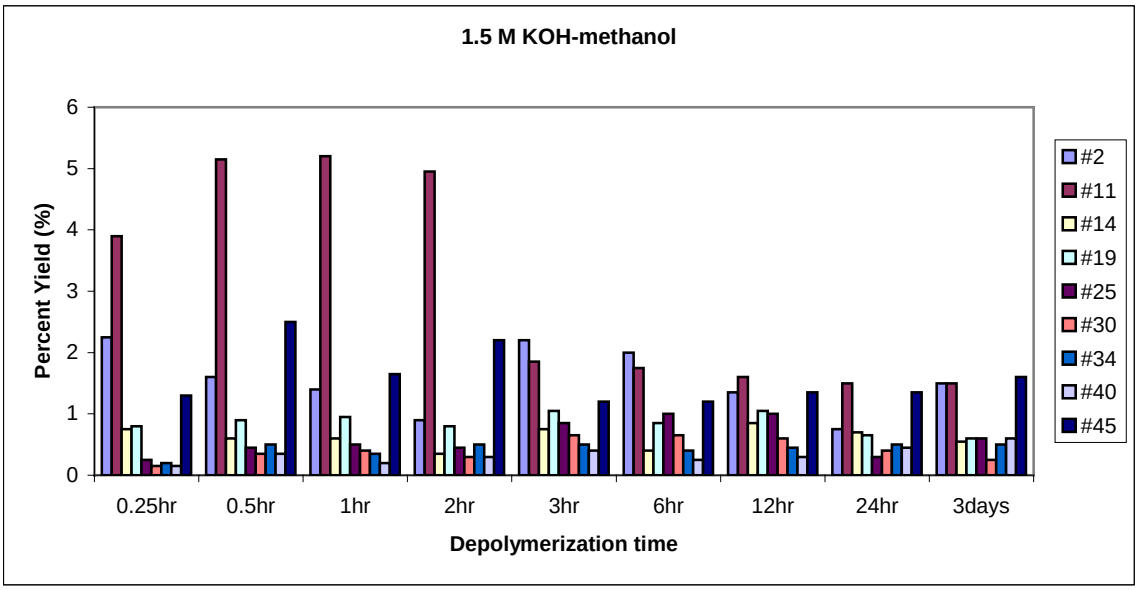


Figure 4.2. Percent yield vs. depolymerization time with 1.5 M KOH treatment (a yield of 5 % corresponds to 10 mg).

oligomeric target compounds, the observed time dependence suggests that reaction times of 0.5-2.0 hr are effective in achieving suberin breakdown to structurally informative

fragments, whereas the slightly greater overall yields found at longer times are accompanied by further breakdown to less interesting monomeric products.

4.1.1.2 Concentration of KOH-methanol solution

At room temperature, 200 mg of suberin was depolymerized with 0.5, 1.0, and 1.5 M KOH-methanol solutions and different times (0.5, 1, 2, 4 hr), respectively. Elevated temperature was avoided because doing so could lead to generation of artifacts.[1] Since the depolymerized suberin was a mixture, preliminary separation was achieved by CC and TLC. Nine fractions were obtained and weighed (Table 4.2).

Table 4.2 The effect of KOH concentration on overall yield of depolymerized products

KOH concentration (M)	Depolymerization time (hr)	Depolymerized soluble part (mg)	Unreacted suberin residue (mg)	Depolymerized percent (%)
0.5	0.5	18.9	174.4	**9.45**
	1	19.0	176.2	**9.50**
	2	19.2	178.7	**9.6**
	4	20.3	167.3	**10.15**
1.0	0.5	18.6	178.8	**9.30**
	1	18.8	178.5	**9.40**
	2	18.9	171.6	**9.45**
	4	20.4	168.9	**10.2**
1.5	0.5	19.1	180.6	**9.55**
	1	17.6	170.7	**8.80**
	2	16.3	179.7	**8.15**
	6	20.8	164.8	**10.4**

Table 4.2 shows the percent yield of depolymerized suberin with different KOH concentrations and depolymerization times. Considering uncertainties from filtration treatment and handling operations, these results show that the concentration of KOH (0.5, 1.0, 1.5 M) has little influence on the depolymerization of suberin.

Figures 4.3 and 4.4 show the percent yield vs. fraction number at 0.5 and 1.0 M KOH-methanol, respectively. ^{1}H NMR shows that fractions #11, #19, and #45 contain the target products. For corresponding fractions, different treatment conditions yield similar mixtures.

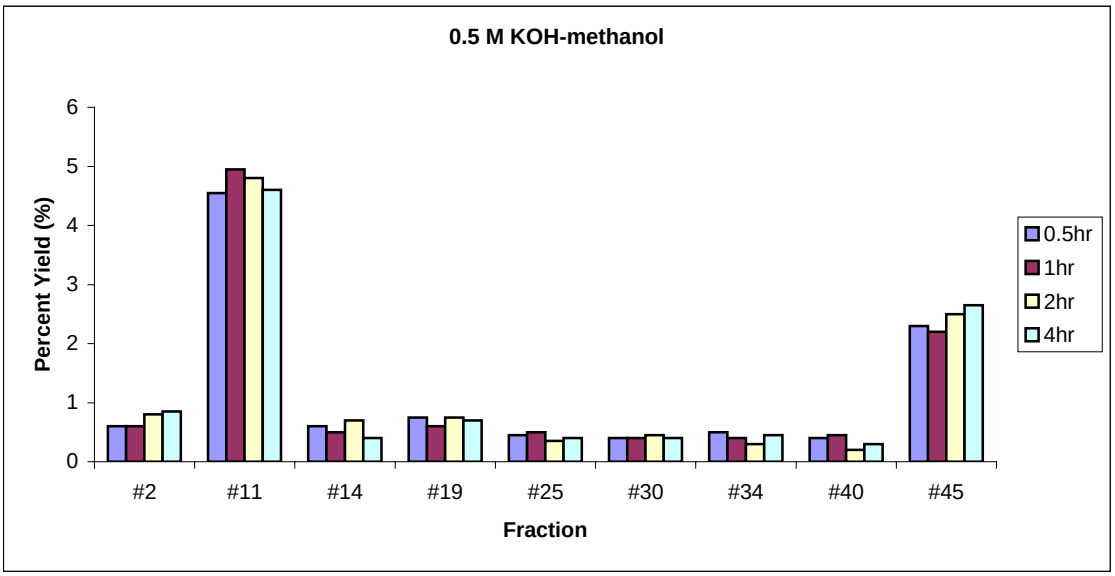


Figure 4.3. Percent yield vs. fraction number with 0.5 M KOH-methanol (a yield of 5 % corresponds to 10 mg).

By comparing the weight of each fraction, it is clear that longer treatment time does not improve the yield of target product (Figure 4.2). Therefore, for 0.5 M and 1.0 M KOH, longer depolymerization time experiments were omitted. For various molarities (0.5, 1.0 and 1.5 M), Figures 4.3 and 4.4 show that no big differences exist at short times (0.5, 1,

2, and 4 hr). Fraction #11 accounts for the largest amount of targets, followed by fraction #45, and finally the other fractions that contain very small amounts.

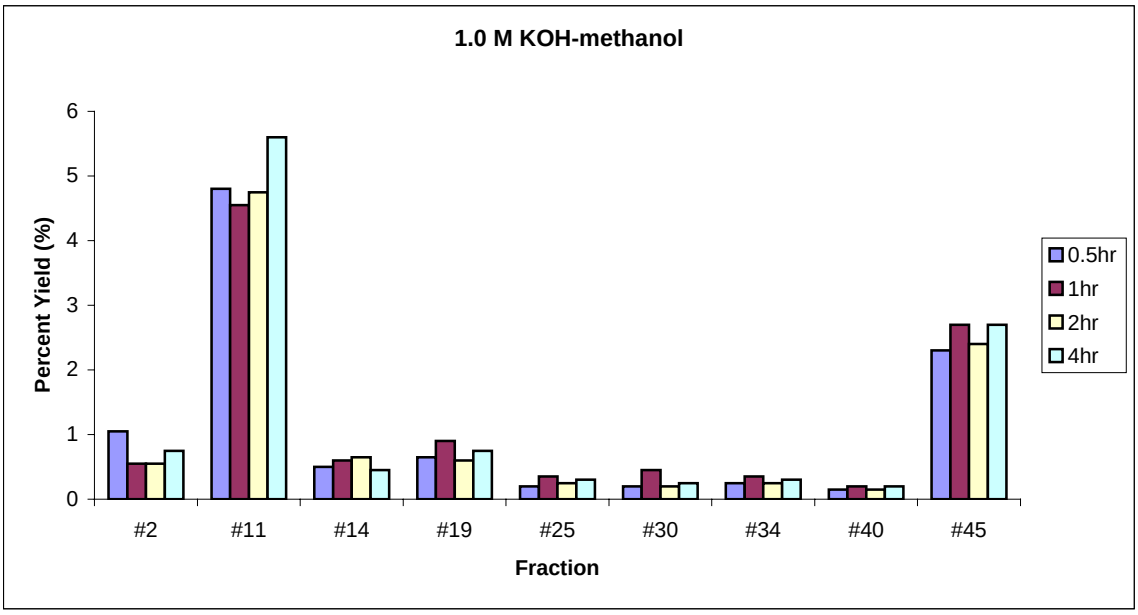


Figure 4.4. Percent yield vs. fraction number with 1.0M KOH-methanol (a yield of 5 % corresponds to 10 mg).

4.1.1.3 ^{1}H NMR integration results

Because lengthening the treatment times and varying the KOH concentration did not enhance the overall yield or proportion of target compounds, subsequent experiments focused on short reaction times and used ^{1}H NMR integration to determine the optimal conditions (Figure 4.5 and Table 4.3). The peak at 4.03 ppm was judged to be the signature for oligomers;[69,70] the peaks between 6 and 8 ppm were indicative of aromatic compounds. Each peak was compared with the resonance at 2.28 ppm, which corresponds to CH_2's adjacent to a carbonyl carbon.

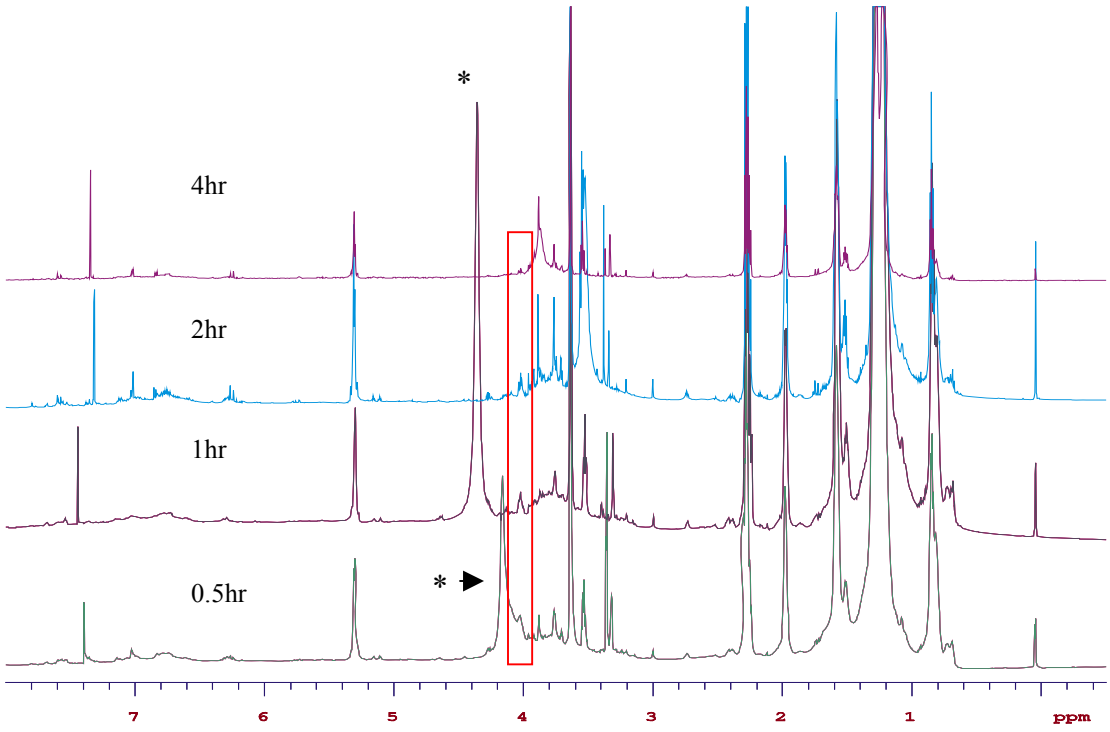


Figure 4.5. 600 MHz ^{1}H NMR of soluble products of 1.0 M KOH treatment for varying times. (The bar highlights the resonance at 4.03 ppm that is diagnostic for oligomers. The large peaks donated by *'s may be water present in the $CDCl_3$ solvent).

Table 4.3 ^{1}H NMR spectral integration results

	ratio of 4.03ppm/2.28ppm	ratio of aromatic (6-8ppm) /2.28ppm
0.5 hr	0.26	0.36
1.0 hr	0.19	0.51
2.0 hr	0.11	0.54
4.0 hr	0.09	0.68

Table 4.3 indicates that with increases in reaction time in the range 0.5-4.0 hr, the relative amount of aliphatic oligomers decreases, and the relative amount of aromatic molecules increases. That means shorter depolymerization time favors aliphatic oligomers, whereas longer depolymerization time favors aromatics.

4.2 Separation of depolymerized suberin #11 using CC and TLC

4.2.1 Fraction #11

Figure 4.2 shows that fraction #11 accounts for about 40% by weight of the depolymerization products obtained with short depolymerization time. Moreover, its ^{1}H NMR spectrum shows that it contains the desired oligomeric and aromatic compounds. Figure 4.6 shows TLC results for each fraction #11 obtained as a function of depolymerization time, revealing a mixture with at least seven spots in each case. Because different depolymerization times yielded similar compositions for fraction #11, the corresponding fractions were combined and separated further. Eight sub-fractions of decreasing polarity were obtained (from bottom to top, TLC-8, TLC-1, TLC-2, TLC-3, TLC-4, TLC-5, TLC-6, and TLC-7 (Figure 4.1).

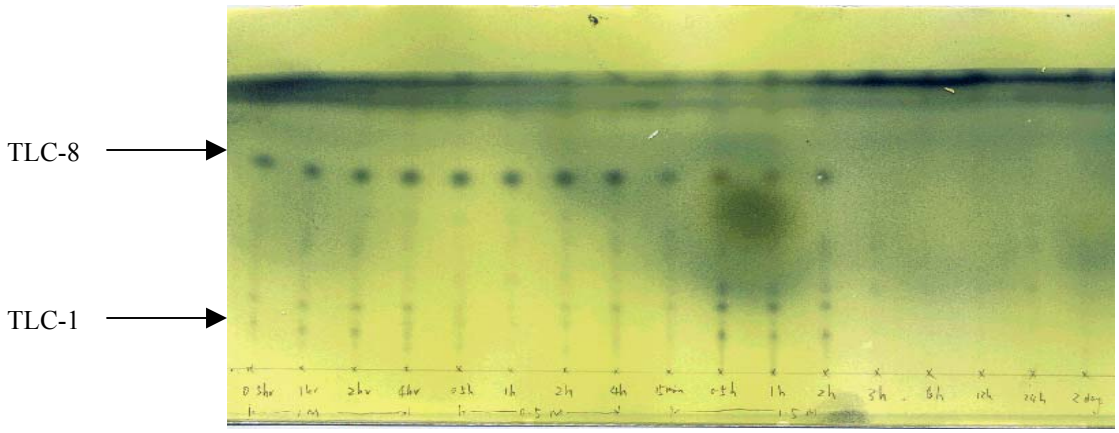


Figure 4.6. TLC spectrum of fraction #11 from treatments for different depolymerization times.

Figure 4.7 displays the ^{1}H NMR of seven fractions from sample #11, showing aromatic signals (6-8 ppm) in each fraction. At chemical shifts near 4 ppm, characteristic peaks are also found for sugars and oligomers. By comparing the NMR spectra of the

seven fractions, the two fractions TLC-4 and TLC-7 are found to have the strongest characteristic peaks of aromatic and oligomer (ester bond), respectively.

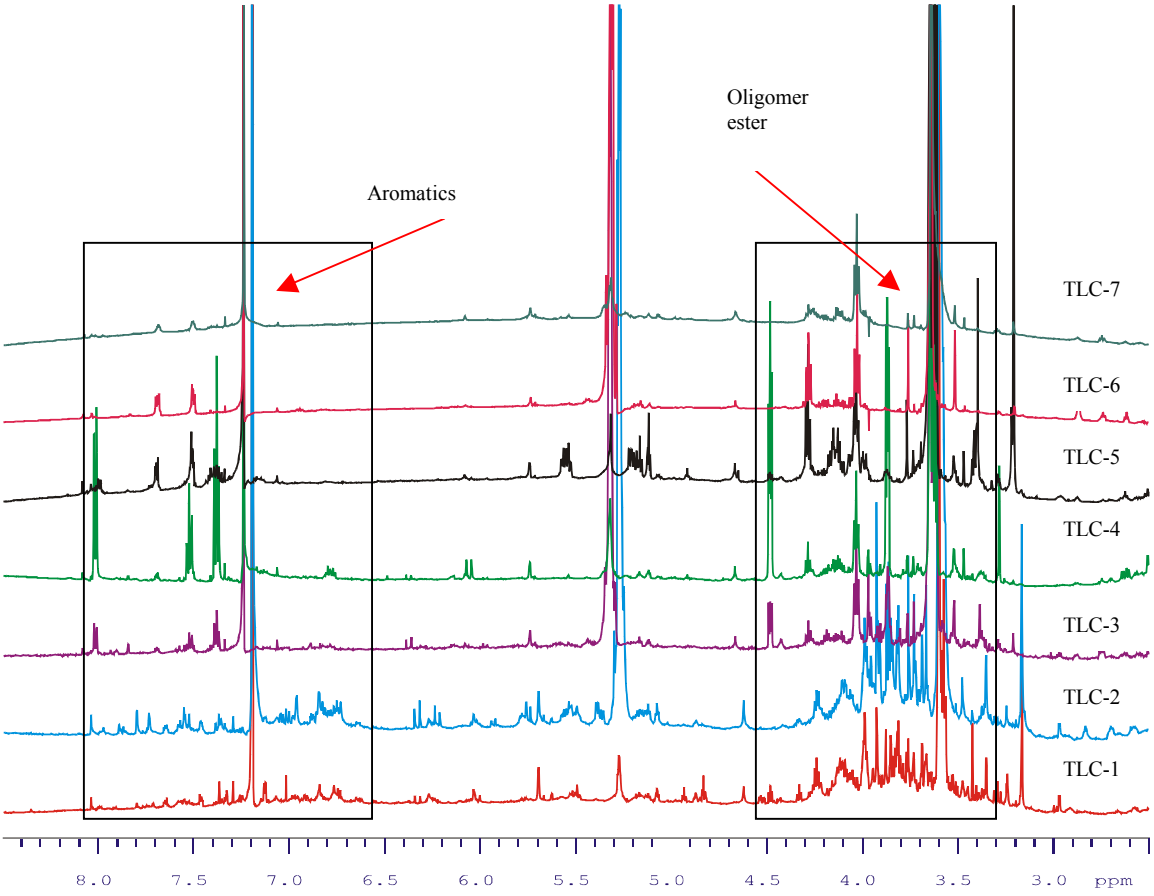


Figure 4.7. Comparison of 600 MHz ^{1}H NMR of seven fractions #11. (Aromatic and ester resonances are highlighted.)

4.2.2 1D and 2D NMR study of #11-TLC-4-1

The aromatic portion of #11-TLC-4 is designated as #11-TLC-4-1. It should be noted that other aliphatic monomers and oligomers also co-exist with it. The **^{1}H NMR** spectrum (Figure 4.8) shows the following resonances in the aromatic spectral region: **H_1** at 7.97 ppm (doublet), **H_2** at 7.48 ppm (triplet), **H_3** at 7.33 ppm (triplet). For **H_4** at 4.48 ppm (triplet) and **H_5** at 3.87 ppm (triplet), the chemical shifts at relatively low field imply aliphatic protons with connections to oxygen. Some other characteristic aliphatic peaks

also appear, e.g. –C**H**$_2$OC(O)- (4.01 ppm), -C**H**=C**H**- (5.28 ppm), -C**H**$_2$-COO- (2.28 ppm), -COO-C**H**$_3$ (3.60 ppm), HO-C**H**$_2$- (3.57 ppm), and R-C**H**$_3$ (0.81 ppm).

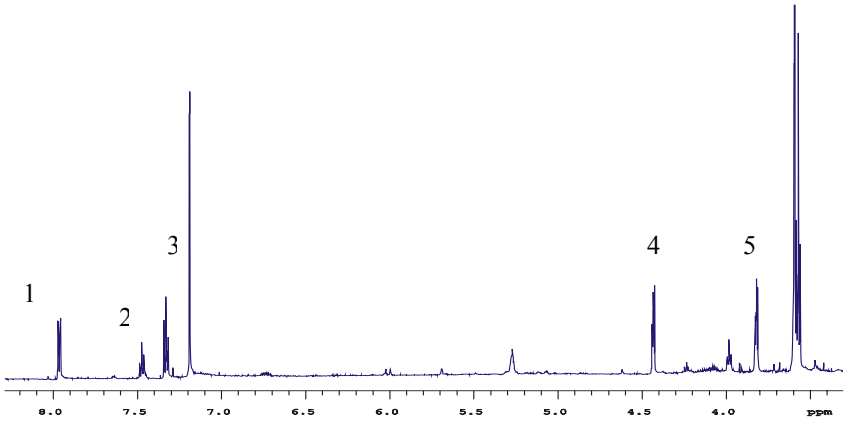


Figure 4.8. Upfield region of the 600 MHz ^{1}H NMR spectrum of #11-TLC-4.

A **COSY** experiment (usually implemented as gmqcosy) (Figure 4.9) gives information about homonuclear correlations via through-bond J coupling. The correlation between protons connected by two to three bonds can be detected. Off-diagonal peaks show that the resonances at the relevant shift positions on the two axes are coupled. Figure 4.9 shows that **H$_{1-3}$**, **H$_{2-3}$** and **H$_{4-5}$** correlations are present, which means that these pairs of protons are located on neighboring carbons. There are also correlations belonging to the aliphatic long chain.

The **^{13}C NMR** spectrum (Figure 4.10) shows 4 aromatic carbons (>130 ppm), doubly bonded carbons (~130 ppm), and 3 different carbonyl carbons, two from long-chain aliphatic molecules (~174 ppm) and another special one that resonates at ~167 ppm and is ultimately assigned to a carbonyl carbon connected to an aromatic ring (see below).

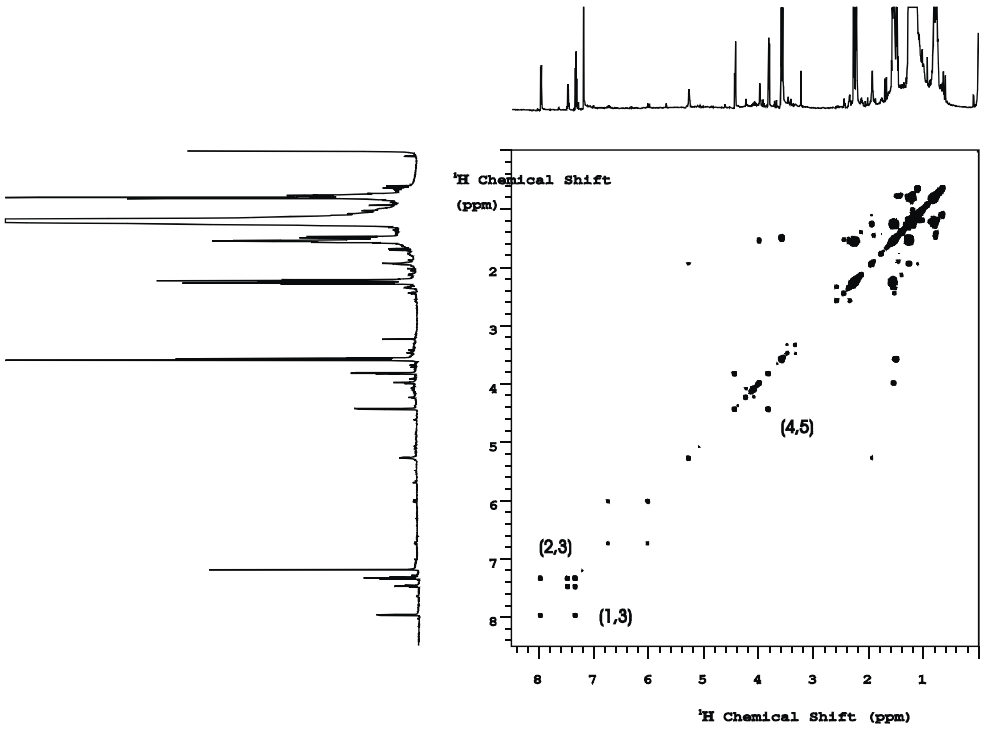


Figure 4.9. Two-dimensional COSY NMR spectrum of #11-TLC-4.

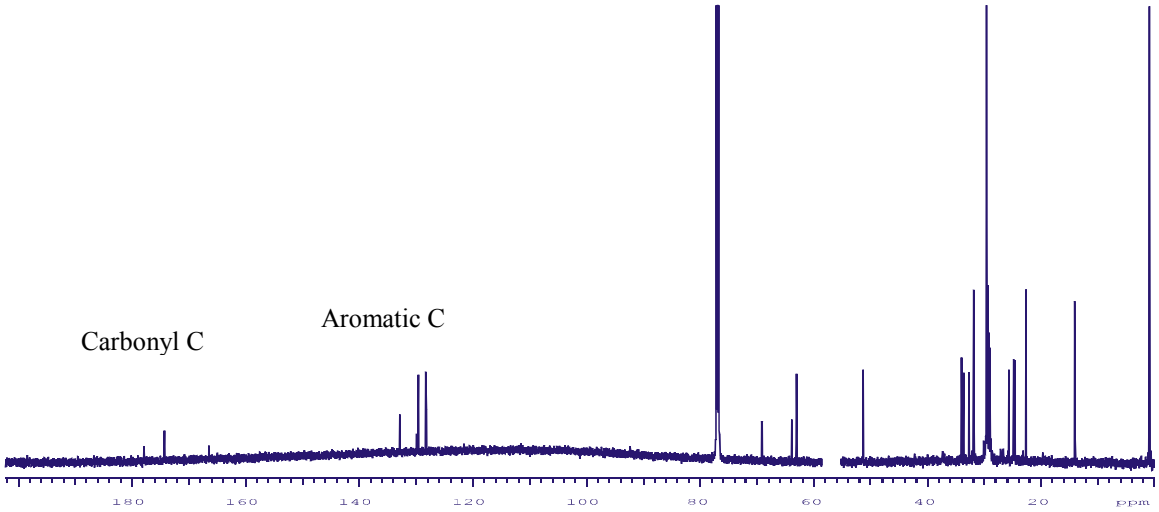


Figure 4.10. 150 MHz ^{13}C NMR spectrum of #11-TLC-4.

The **gHSQC** experiment (Figure 4.11) identifies directly bonded protons and carbons. #11-TLC-4 is still a mixture, for which Fig. 4.11 shows the following diagnostic

correlations belonging to long-chain aliphatics: -CH=CH- (129.7, 5.28 ppm), -CH_2-COO-CH_2- (64.5, 4.01 ppm). –CH_2-COO- (34.0, 2.28 ppm), R-CH_3 (14.3, 0.8 ppm), -COO-CH_3 (63.2, 3.60 ppm), and -CH_2-OH (63.2, 3.57 ppm). The other 5 characteristic peaks are **H_1-C_1** (129.7, 7.97 ppm), **H_2-C_2** (132.9, 7.48 ppm), **H_3-C_3** (128.4, 7.33 ppm), **H_4-C_4** (63.7, 4.48 ppm), and **H_5-C_5** (68.9, 3.87 ppm).

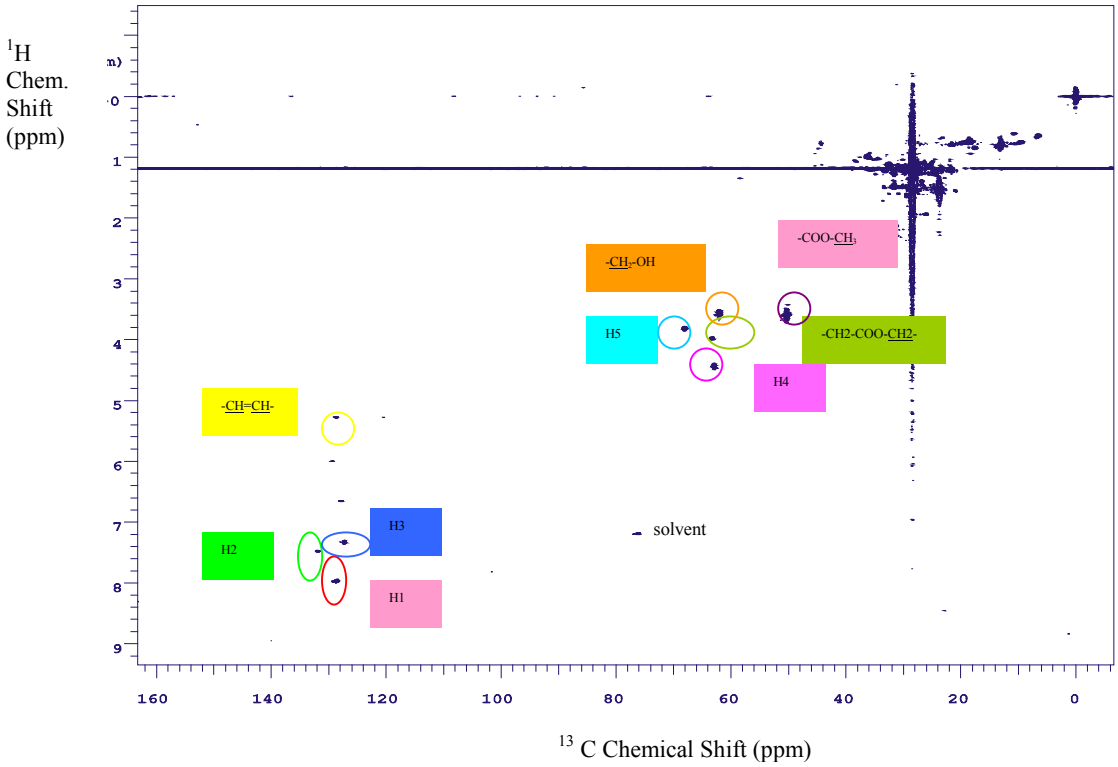


Figure 4.11. Two-dimensional gHSQC NMR spectrum of #11-TLC-4.

Finally, **gHMBC** (Figure 4.12) gives information on ^{1}H-^{13}C pairs that are typically 3 bonds apart. The cross-peak at (166.7, 7.97 ppm) indicates a correlation between aromatic proton **H_1** with the carboxylic side of a carbonyl carbon. The cross-peak at (166.7, 4.48 ppm) indicates a correlation between proton **H_4** and the carbonyl carbon from the hydroxyl side. Because **H_4** and **H_5** correlate with each other in the COSY spectrum, **H_5** must be connected to **H_4** via 3 intervening bonds. The notable downfield

chemical shift of **H_4** and **H_5** also suggests that they are bound to an oxygen. Thus, the connection should be: aromatic ring-COO-CH_2 (#4) –CH_2 (#5)-O-.

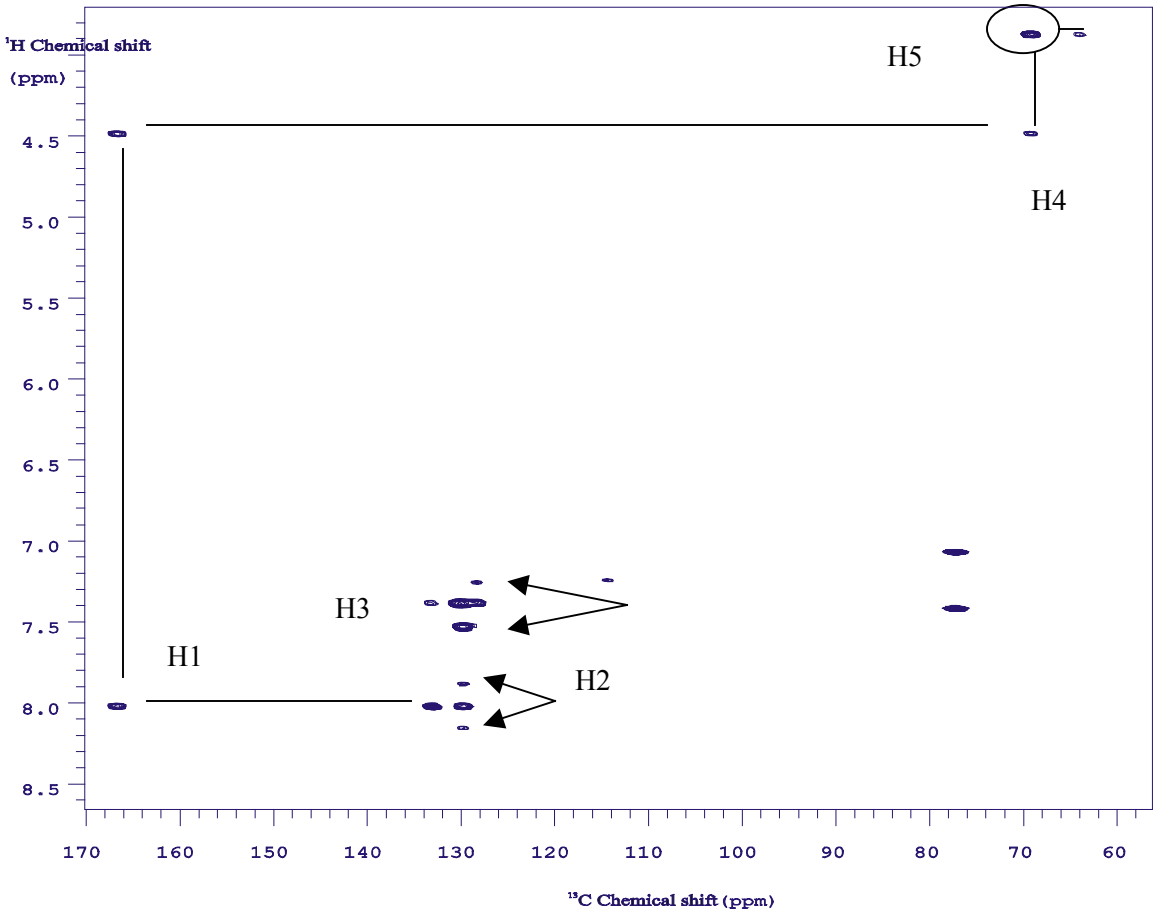
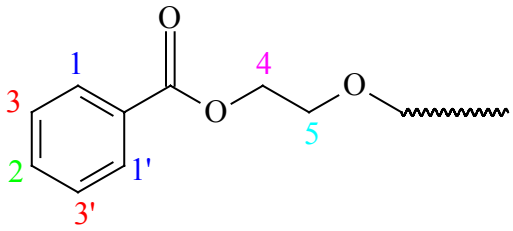


Figure 4.12. Two-dimensional gHMBC NMR spectrum of #11-TLC-4.

For the aromatic ring, there are six carbons, one of which connects to the carbonyl. For correlations to **H_1** and **H_3** in gHMBC, the symmetrical sets of signals (arrows) are attributable to a symmetrical structure because of multiple-quantum correlation (H. Wang, personal communication). Thus the aromatic ring must contain 2 pairs of protons at 6-8 ppm, each of which has the same chemical environment because of a symmetrical structure. In all, three proton signals can be seen, but there are 5 protons present: one at

the para-position (#2), two (#1 and #1') at the ortho-position, and two (#3 and #3') at the meta-position.

Thus part of the preliminary structure is:



A very special anomalous peak at (68.9, 3.87 ppm) is circled in Figure 4.12. Its chemical shifts match the $\mathbf{H_5}$-$\mathbf{C_5}$ direct correlation observed in the gHSQC spectrum (Figure 4.11), but in that case it should not be present in HMBC. Surprisingly, no other correlations are detected beyond the oxygen atom.

Although direct H-C correlations show up occasionally in HMBC spectra, they are usually split into two peaks separated by the $^2J_{CH}$ of about 145 Hz. So our peak should be a real 2- or 3- bond correlation. The most reasonable structure consistent with these observations is a dimer of the above preliminary structure (Figure 4.13).

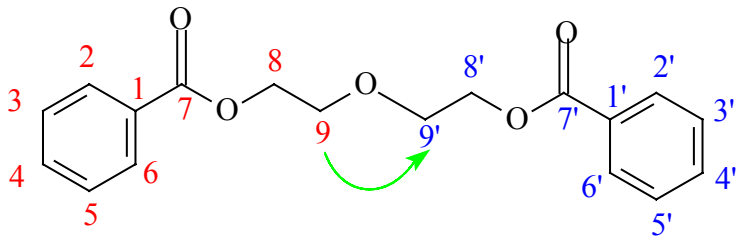


Figure 4.13. Tentative structure of #11-TLC-4-1. The positions have been renumbered compared with the partial structure above.

If this structure is correct, the peak at (68.9, 3.87 ppm) in the HMBC spectrum can be accounted for. Rather than a $\mathbf{H_9}$-$\mathbf{C_9}$ direct-bond correlation, it is an $\mathbf{H_9}$,-$\mathbf{C_9}$ 3-bond

correlation. This symmetrical structure also explains why no other correlations are detected beyond the oxygen atom.

4.2.3 ACD spectral simulation of #11-TLC-4-1

ACD software was used to predict ^{13}C and ^{1}H NMR spectra. Table 4.4 compares chemical shifts from NMR experiments and ACD calculations (numbering is taken from ACD). Both ^{13}C and ^{1}H shifts show excellent agreement between the observed and predicted spectra. Examples of ACD results appear in Figures 4.14 and 4.15.

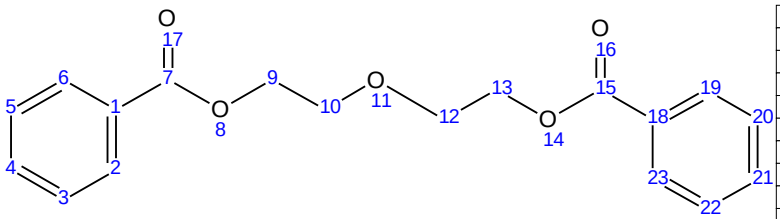
Table 4.4 Calculated and Observed Chemical Shifts of #11-TLC-4-1

Proton #	# of protons	Calc. Chemical shift (ppm)[a]	Exp. Chemical shift (ppm)
2,6,19,23	1/position	7.88	7.97
3,5,20,22	1/position	7.46	7.33
4,21	1/position	7.58	7.48
9,13	2/position	4.34	4.48
10,12	2/position	3.7	3.87

Carbon #	CH_n	Calc. Chemical shift (ppm)[a]	Exp. Chemical shift (ppm)
2,6,19,23	CH	129.3	129.8
3,5,20,22	CH	128.7	128.4
4,21	CH	133.4	133.0
9,13	CH_2	64.7	63.7
10,12	CH_2	70.1	68.9
1,18	C	129.6	130.0
7,15	C	165.9	166.7

[a] Typical error limits provided by ACD are ±1 ppm and ± 0.1 ppm for ^{13}C and ^{1}H spectra, respectively. Atom numbers correspond to the structures in Figures 4.14 and 4.15.

[b] Numbering comes from the ACD structures (Figures 4.14 and 4.15).



Carbon No.	CHn	Chem. Shifts
1	C	129.56
2	CH	129.25
3	CH	128.68
4	CH	133.36
5	CH	128.68
6	CH	129.25
7	C	165.93
9	CH_2	64.86
10	CH_2	70.12
12	CH_2	70.12
13	CH_2	64.86
15	C	165.93
18	C	129.56
19	CH	129.25
20	CH	128.68
21	CH	133.36
22	CH	128.68
23	CH	129.25

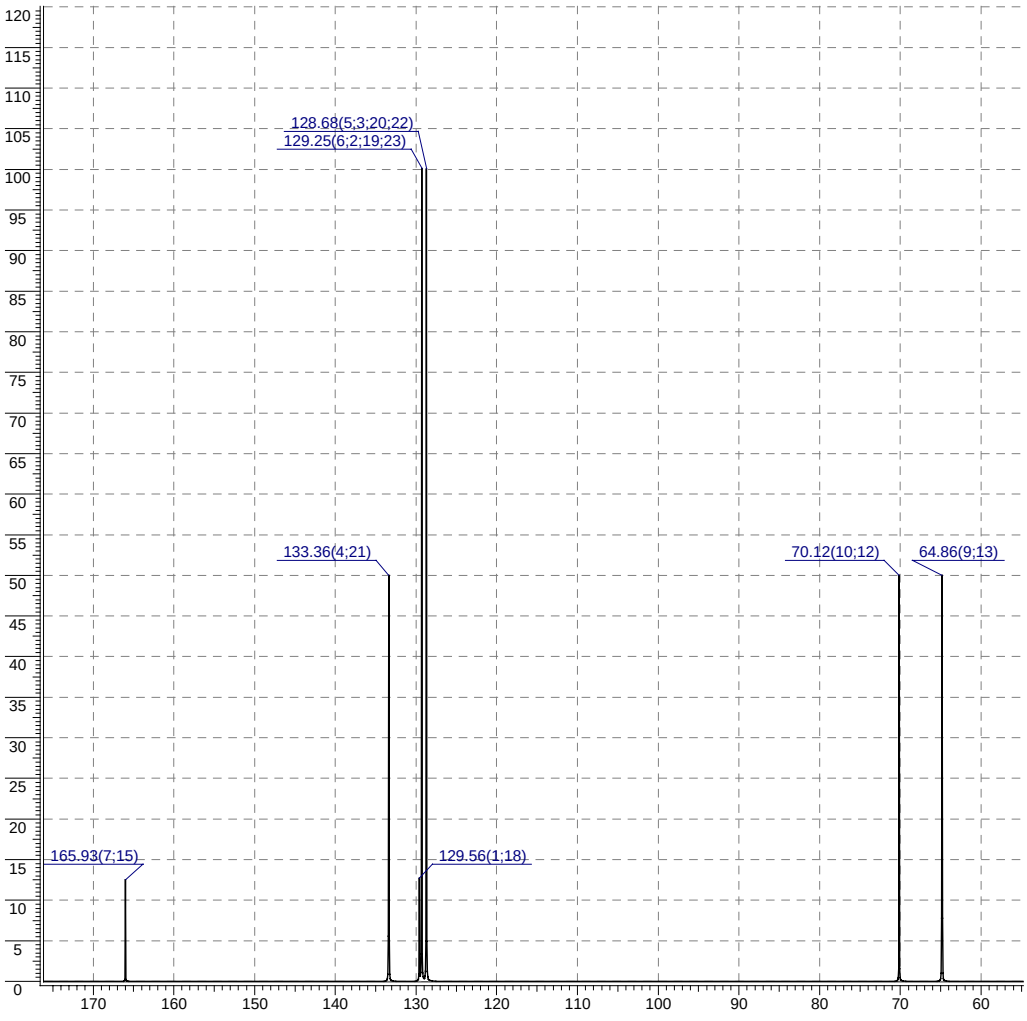
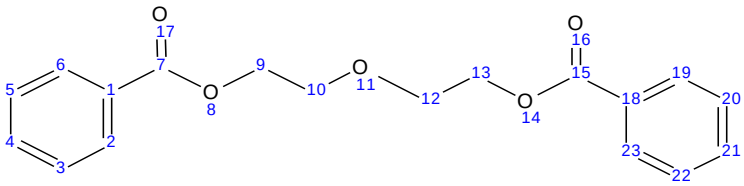


Figure 4.14. ACD calculation of carbon NMR spectrum for #11-TLC-4-1.



J	Grp.1	Grp.2	Value
^{3}J	2	3	7.89
^{4}J	2	4	1.3
^{3}J	3	4	7.12
^{5}J	5	2	0.62
^{4}J	5	3	1.38
^{3}J	5	4	7.12
^{4}J	6	2	1.85
^{5}J	6	3	0.62
^{4}J	6	4	1.3
^{3}J	6	5	7.89
^{2}J	9	9	10
^{3}J	9	10	4.43
^{2}J	10	10	10
^{4}J	10	12	0.73
^{2}J	12	12	10
^{3}J	12	13	4.43
^{2}J	13	13	10
^{3}J	19	20	7.89
^{4}J	19	21	1.3
^{5}J	19	22	0.62
^{4}J	19	23	1.85
^{3}J	20	21	7.12
^{4}J	20	22	1.38
^{5}J	20	23	0.62
^{3}J	21	22	7.12
^{4}J	21	23	1.3
^{3}J	22	23	7.89

Group	nH	Shift
2	1	7.882
3	1	7.461
4	1	7.584
5	1	7.461
6	1	7.882
9	2	4.34
10	2	3.7
12	2	3.7
13	2	4.34
19	1	7.882
20	1	7.461
21	1	7.584
22	1	7.461
23	1	7.882

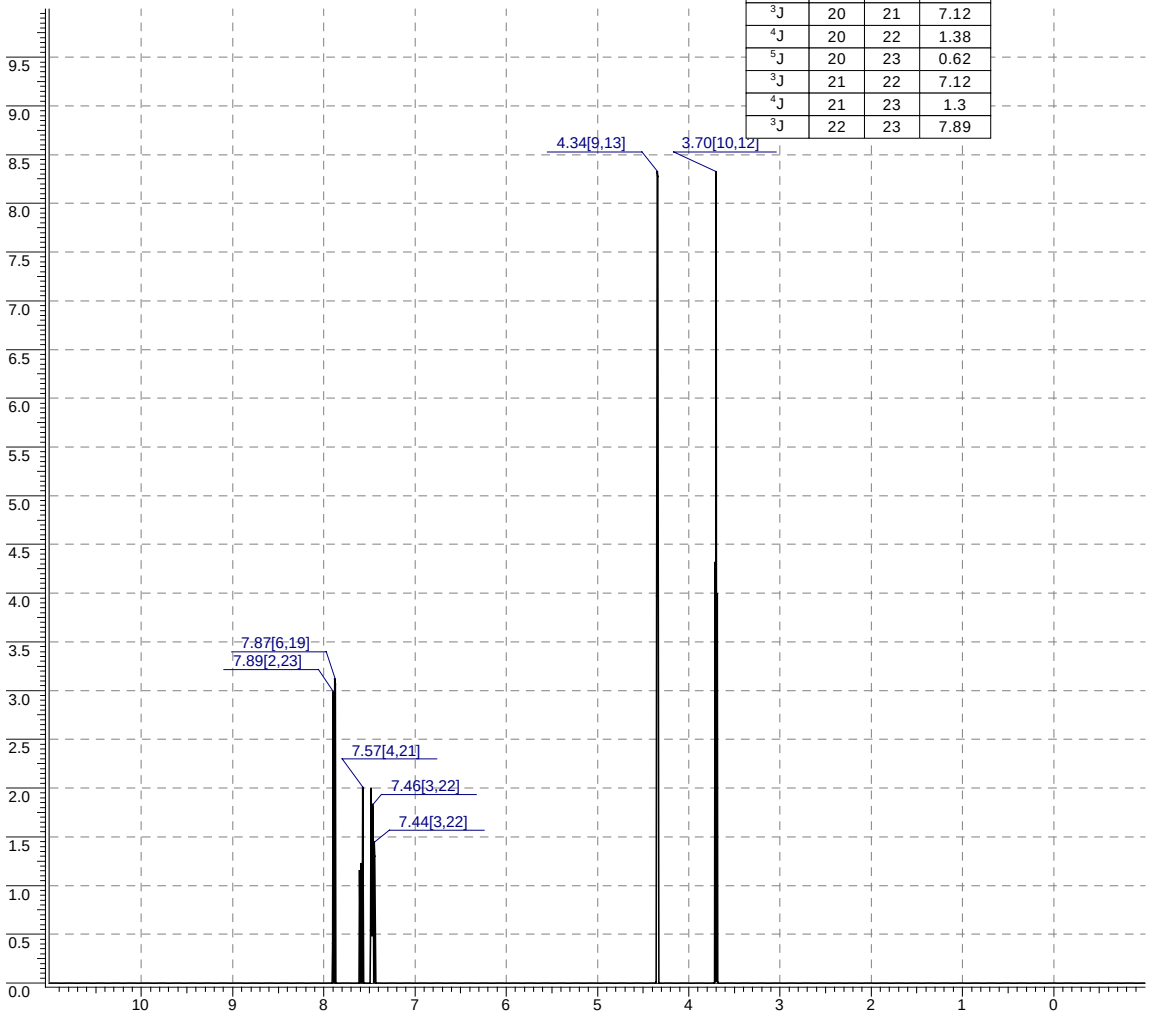


Figure 4.15. ACD calculation of proton NMR spectrum for #11-TLC-4-1.

4.3 HPLC method development

4.3.1 Challenges in the separation of suberin depolymerization products

a. Depolymerization products are mixtures of numerous compounds (maybe hundreds, see Figure 4.16). The major products of suberin depolymerization are aliphatic monomers, which have been identified previously (Figure 1.2).[33] The aromatics and oligomers account for only a small portion of the total depolymerized suberin. Usually less than 15% of the suberin can be broken down, so the absolute amount of these latter products is very modest.

b. Some of the aliphatic compounds are homologs, for which the only difference is the number of CH_2 groups. Because their polarities are very similar, it is extremely hard to separate and identify them solely using CC, TLC, and NMR, and HPLC. MS^n is needed, but in this case, the compound must be at least relatively pure.

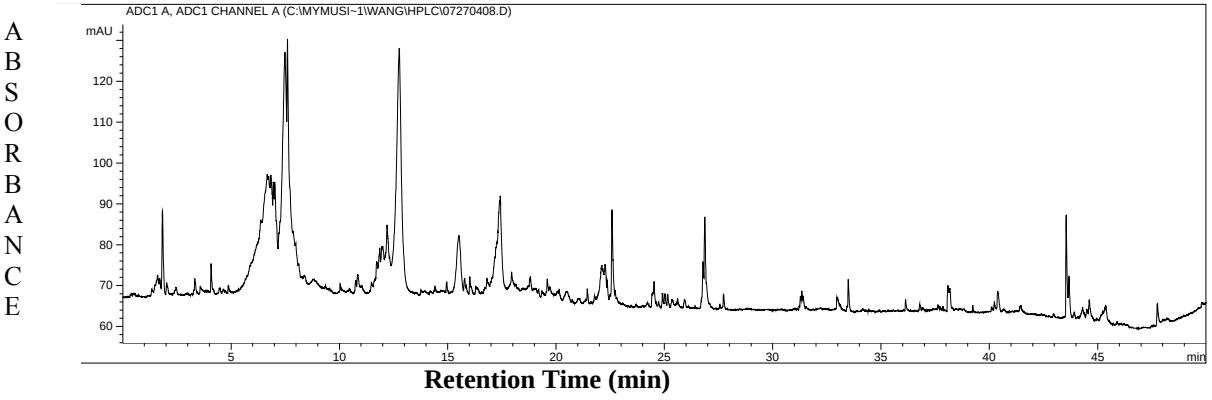


Figure 4.16. Elastic Light Scattering Detection (ELSD) of HPLC for fraction #11.

c. Few of the long-chain aliphatic compounds have UV absorption, so the compounds separated by HPLC must be detected using ELSD. However, ELSD is less sensitive than UV absorbance.

Considering all of the above issues, new separation methods were developed to obtain the desired compounds in pure form from suberin depolymerization products.

4.3.2 HPLC separation of aromatic dimer #11-TLC-4-1

In Section 4.2, the structure of the #11-TLC-4-1 (Figure 4.13) was determined provisionally by NMR and MS (see Chapter 5), but it was noted that this fraction is still a mixture. In order to obtain other pure aromatic and oligomeric compounds, an HPLC method was developed.

Figure 4.17 is a flow chart describing HPLC separation of sample #11. The 'HCG_NP-6' method (Table 4.5), a modified version of that used for fruit cutin monomers,[71] was used for the first round of separation; six fractions were obtained. Figure 4.18 displays the ELSD spectra from this separation. Fractions #11-HPLC-3, #11-HPLC 4, and #11-HPLC-5 all contain some components that were expected to appear only in fraction #11-HPLC-2, thus the purification is not satisfactory.

Table 4.5 Mobile phase gradient HCG-NP-6 for separation of the soluble products of suberin

Time (min) Flow rate=0.5ml/min	Composition of mobile phase, % (v/v)	
	Eluent A [a]	Eluent B [b]
0	99	1
1	97.5	2.5
8	80	20
33	60	40

[a] Eluent A: hexane/isopropanol/acetic acid (99.3%: 0.5%: 0.2%)
[b] Eluent B: hexane/isopropanol/acetic acid (79.8%: 20.0%: 0.2%)

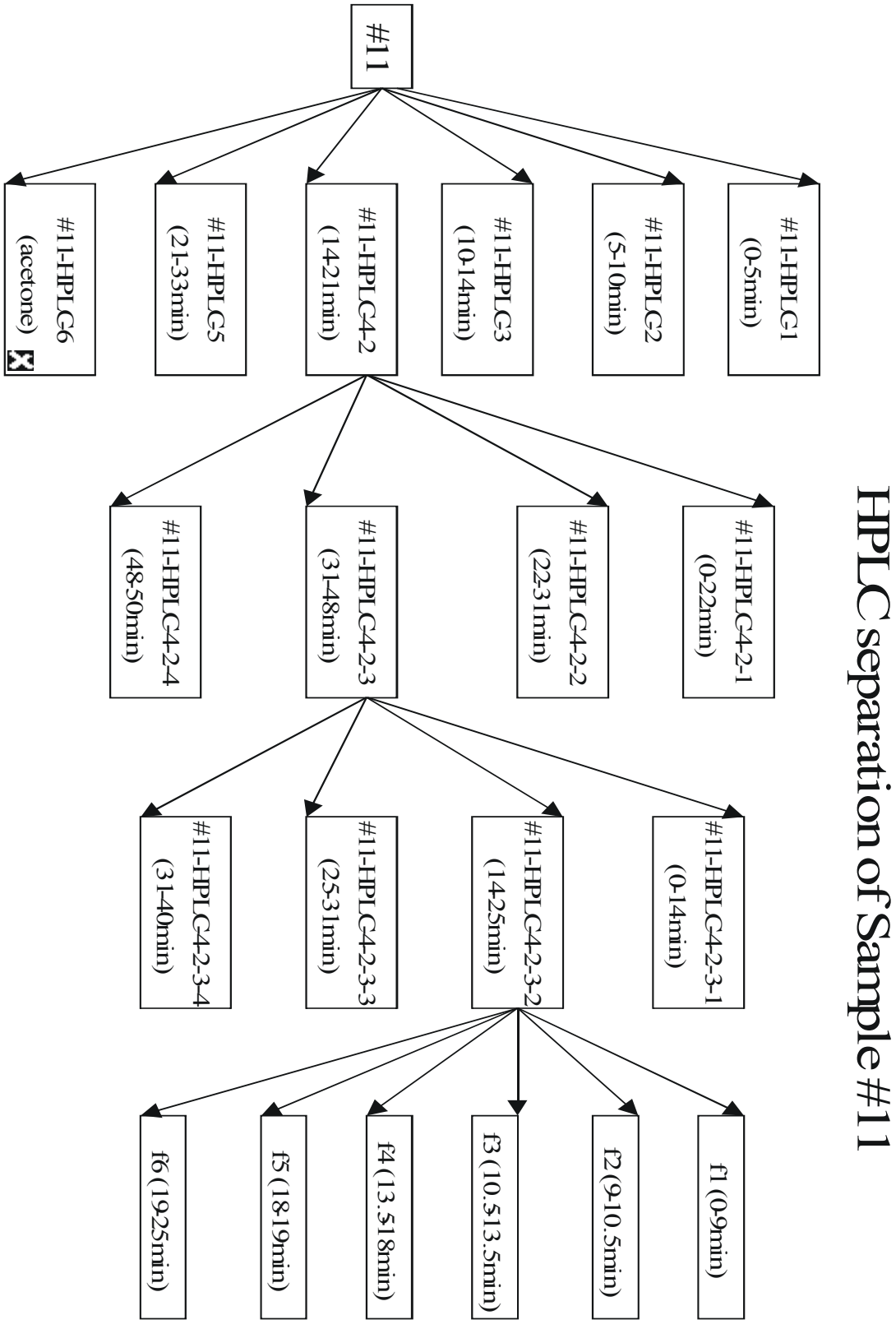


Figure 4.17. Flow chart for HPLC separation of fraction #11.

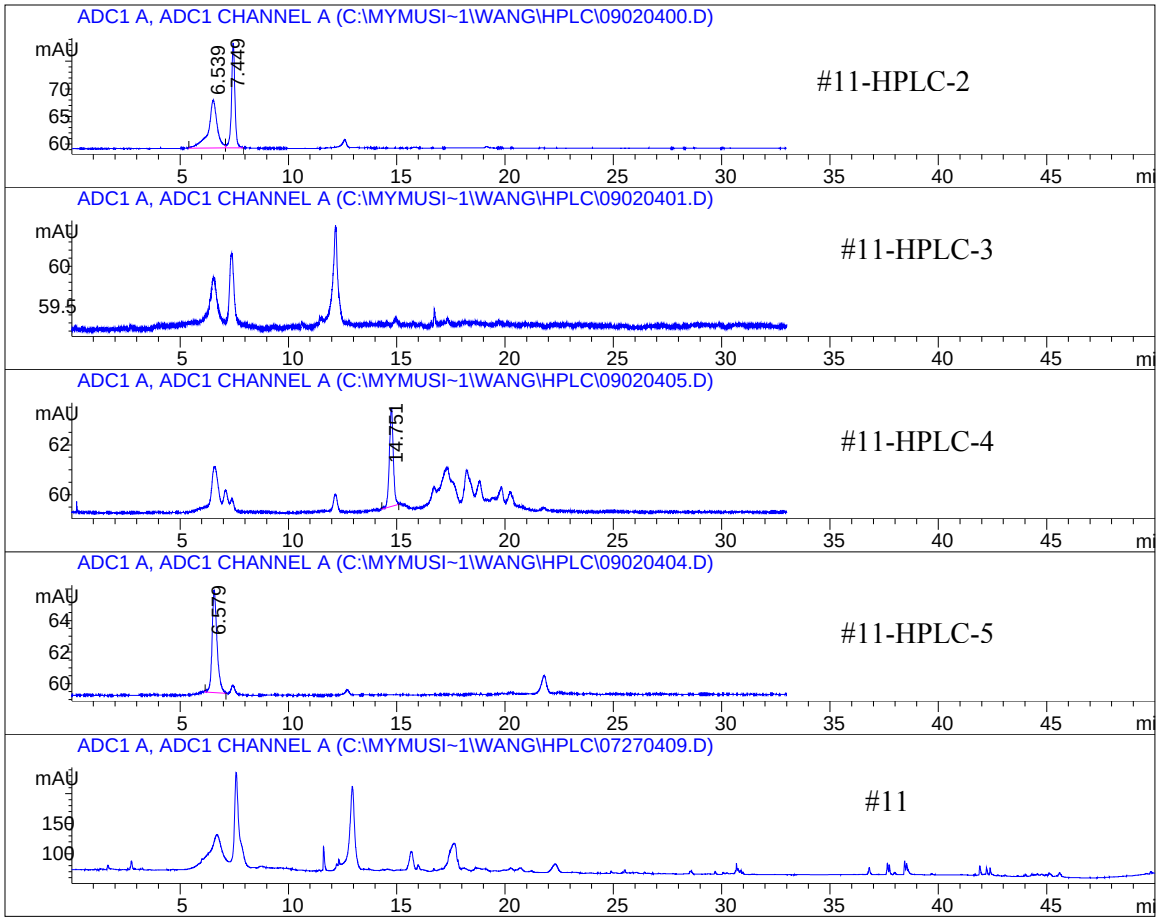


Figure 4.18. HPLC comparison of different fractions from fraction #11, using ELSD detection.

Figure 4.19 shows the ^{1}H NMR spectra of fractions #11-HPLC-1 through #11-HPLC-5. Surprisingly, fraction #11-HPLC-5, collected at elution times between 21 and 33 min, shows only a peak between 5 and 10 min after re-injection. Even though fraction #11-HPLC-4 contains mainly the aromatic dimer of interest, it is clear that all fractions contain both oligomers and aromatic compounds. The gHMBC spectrum of #11-HPLC-4 (not shown) also evidences a mixture: some of the ester bond correlations are not connected through bonds to the aromatic compound. In order to get pure product, HPLC was used to separate #11-HPLC-4 further.

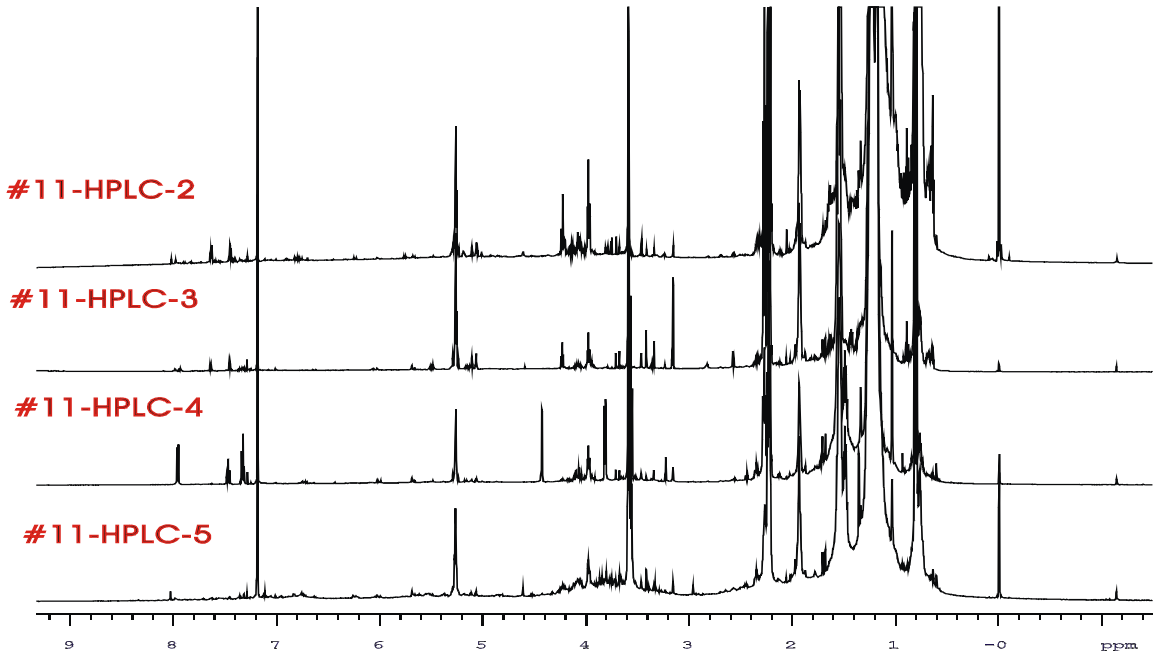


Figure 4.19. 600 MHz ^{1}H NMR of different fractions from HPLC separation of sample #11.

The HPLC method HCG_NP_6 (Table 4.5) was used to separate #11-HPLC-4 again, collecting the following fractions: 0-5 min, 5-10 min, 10-13.5 min, 13.5-16 min (#11-HPLC-4-1), 16-21 min (#11-HPLC-4-2), and 21-33 min. Fig. 4.20 shows the ELSD spectra of #11-HPLC-4, #11-HPLC-4-1, and #11-HPLC-4-2, which together with NMR (Figure 4.21) indicate a mixture of the aromatic dimer with other aliphatic oligomers and monomers.

The HPLC method HCG_NP (Table 4.6) separated #11-HPLC-4-2 into four fractions (Fig. 4.17): #11-HPLC-4-2-1: (0-22 min); #11-HPLC-4-2-2: (22-31 min); #11-HPLC-4-2-3: (31-48 min); #11-HPLC-4-2-4: (48-50 min).

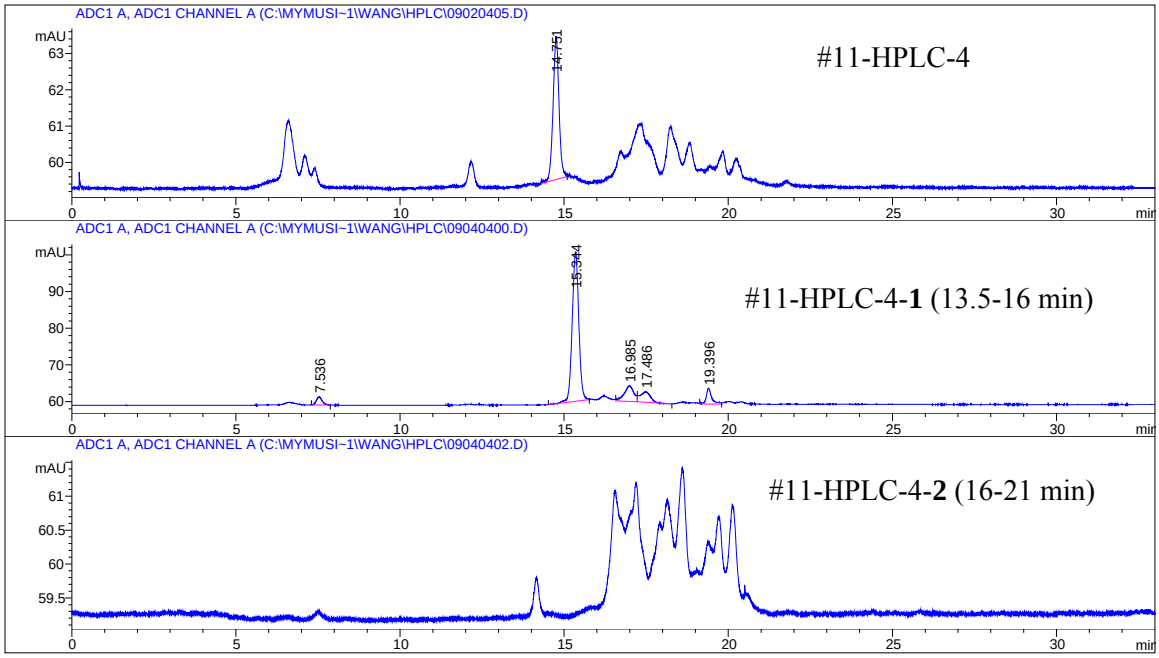


Figure 4.20. ELSD from separation of #11-HPLC-4.

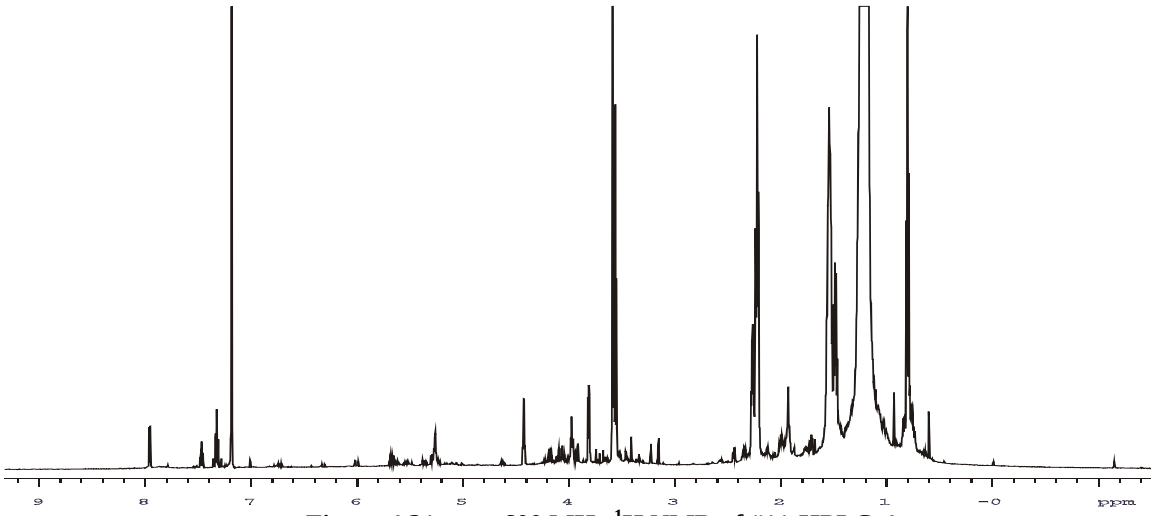


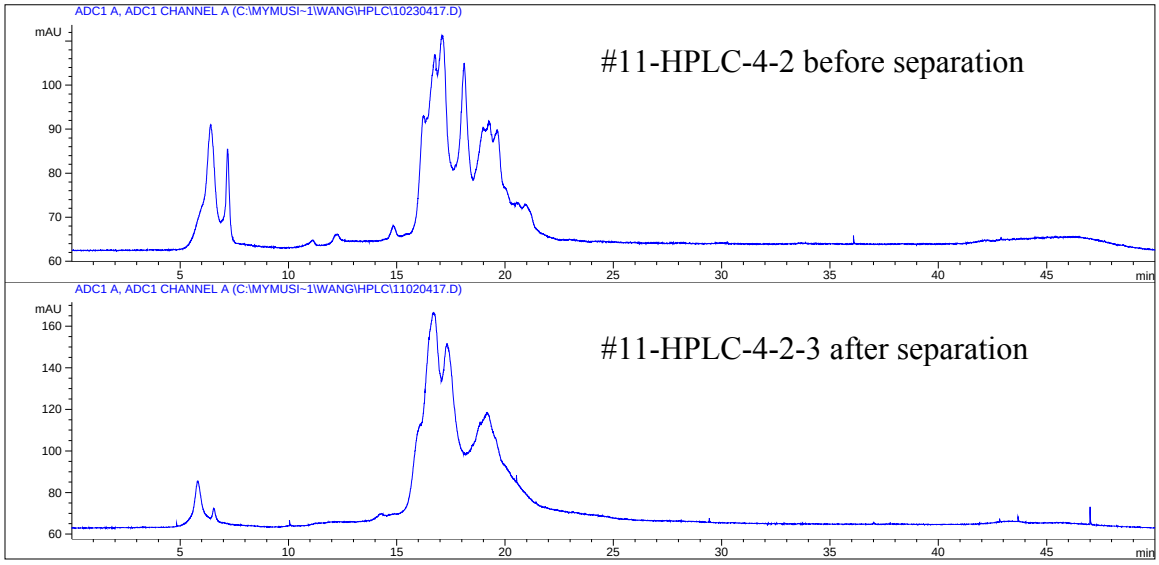
Figure 4.21. 600 MHz ^{1}H NMR of #11-HPLC-4.

Figure 4.22 shows that after this round of separation, #11-HPLC-4-2-3 contains the aromatic dimer as its principal constituent, but it still contains compounds from other fractions.

Table 4.6 HPLC HCG-NP separation of soluble products from suberin degradation

Time (min) Flow rate=1ml/min	Composition of mobile phase % (v/v)	
	Eluent A [a]	Eluent B [b]
0	99	1
1	97.5	2.5
8	80	20
33	60	40
39	0	100
42	99	1
50	99	1

[a] Eluent A: hexane/isopropanol/acetic acid (99.3%: 0.5%: 0.2%)
[b] Eluent B: hexane/isopropanol/acetic acid (79.8%: 20.0%: 0.2%)



Elution time (min)

Figure 4.22. ELSD of #11-HPLC-4-2 before and after the separation.

The HPLC method NP-1 (Table 4.7) was used to separate #11-HPLC-4-2-3 to get another four fractions (Figure 4.17): #11-HPLC-4-2-3-1: (0-14 min); #11-HPLC-4-2-3-2: (14-25 min); #11-HPLC-4-2-3-3: (25-31 min); #11-HPLC-4-2-3-4: (31-40 min). Both Figure 4.23 (ELSD spectra) and Figure 4.24 (^{1}H NMR spectra) show that #11-HPLC-4-2-3-2 contains the aromatic dimer, but it still contains other aliphatic compounds. None of the fractions is pure because they all contain fractions that appear around 10 min., e.g.

the last fraction (#11-HPLC-4-2-3-4) only has peaks eluting around 10 min after injection. The presence of both aromatic and aliphatic compounds is an unexpected result, because their structures and polarities are expected to be dissimilar.

Table 4.7 HPLC method NP-1 for separation of soluble products of suberin degradation

Time (min) Flow rate=1ml/min	Composition of mobile phase % (v/v)	
	Eluent A [a]	Eluent B [b]
0	90	10
40	85	15

[a] Eluent A: hexane/isopropanol/acetic acid (99.3%: 0.5%: 0.2%)
[b] Eluent B: hexane/isopropanol/acetic acid (79.8%: 20.0%: 0.2%)

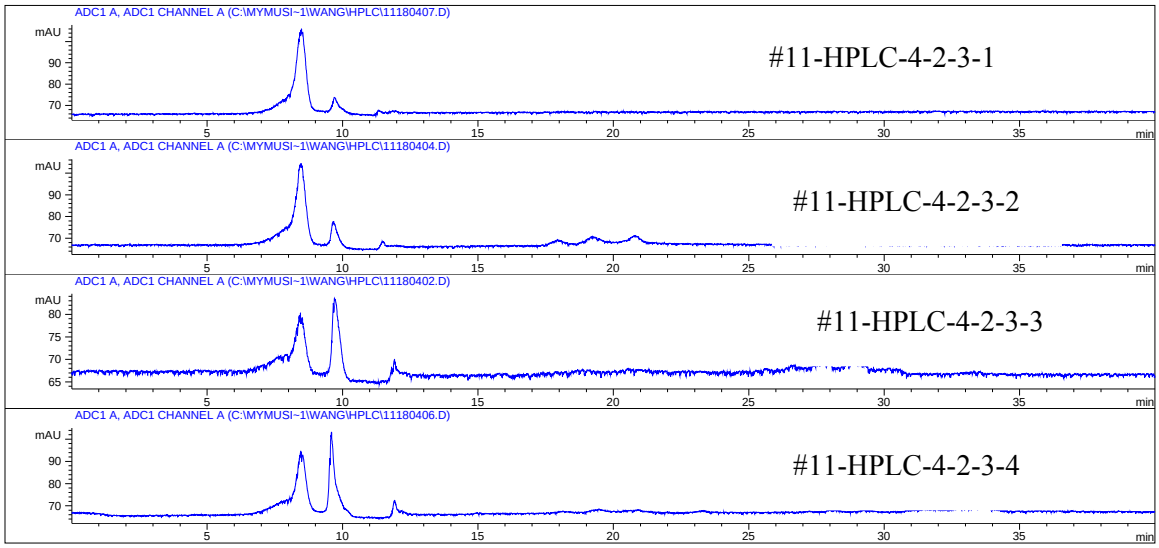


Figure 4.23. HPLC separation of #11-HPLC-4-2-3.

The HPLC method HCG-NP-9 (Table 4.8) was used to separate #11-HPLC-4-2-3-3 into 6 fractions (Figure 4.20): #11-HPLC-4-2-3-2-f1 (0-9 min); #11-HPLC-4-2-3-2-f2 (9-10.5 min); #11-HPLC-4-2-3-2-f3 (10.5-13.5 min); #11-HPLC-4-2-3-2-f4 (13.5-18 min); #11-HPLC-4-2-3-2-f5 (18-19 min); #11-HPLC-4-2-3-2-f6 (19-25 min).

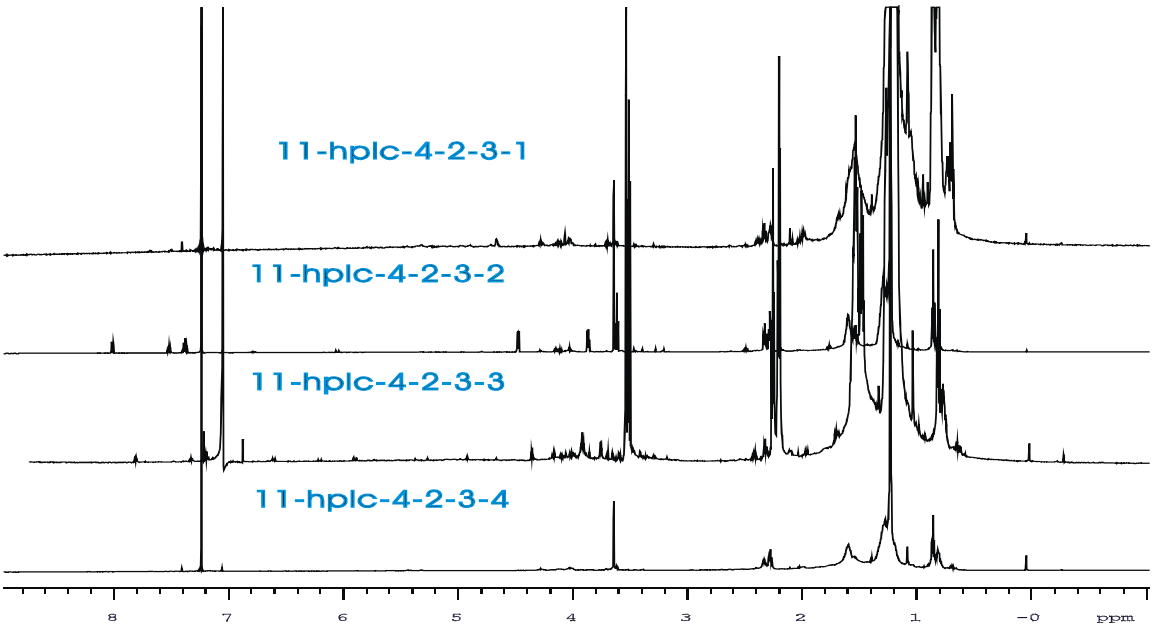


Figure 4.24. 600 MHz ^{1}H NMR spectrum of four fractions from #11-HPLC-4-2-3.

Table 4.8 HPLC method HCG-NP-9 for separation of soluble products of suberin

Time (min) Flow rate=1ml/min	Composition of mobile phase % (v/v)	
	Eluent A [a]	Eluent B [b]
0	88	12
25	85	15

[a] Eluent A: hexane/isopropanol/acetic acid (99.3%: 0.5%: 0.2%)
[b] Eluent B: hexane/isopropanol/acetic acid (79.8%: 20.0%: 0.2%)

Figure 4.25 shows ELSD traces from HPLC separation of #11-HPLC-4-2-3-2. Fraction 5 contains the aromatic dimer, but it is still a mixture and all fractions show peaks that correspond to aliphatics around 10 min. Figure 4.26 shows the ELSD and UV signals of #11-HPLC-4-2-3-2-f5; UV at 254 nm and 280 nm both show absorption. Figure 4.27 shows the ^{1}H NMR spectrum of #11-HPLC-4-2-3-2-f5. It is purer than #11-HPLC-4-2-3-2 (Figure 4.24), but it is still a mixture of the aromatic dimer and aliphatic compounds. This conclusion is evidenced by the gHMBC spectrum of #11- HPLC-4-2-3-

2-f5 (not shown), which shows no covalent connections between the aromatic and aliphatic moieties.

As noted above, ^{1}H NMR (Figures 4.27 and 4.28) shows that the compound looks relatively pure. The problem is that even if after many cycles of HPLC separation, the aromatic dimer is always tangled together with the aliphatic monomers. The aliphatic part can only be separated from the aromatic part by successive separations. This is an unexplained phenomenon, because aromatic and aliphatic compounds with different polarities should be separated easily by HPLC.

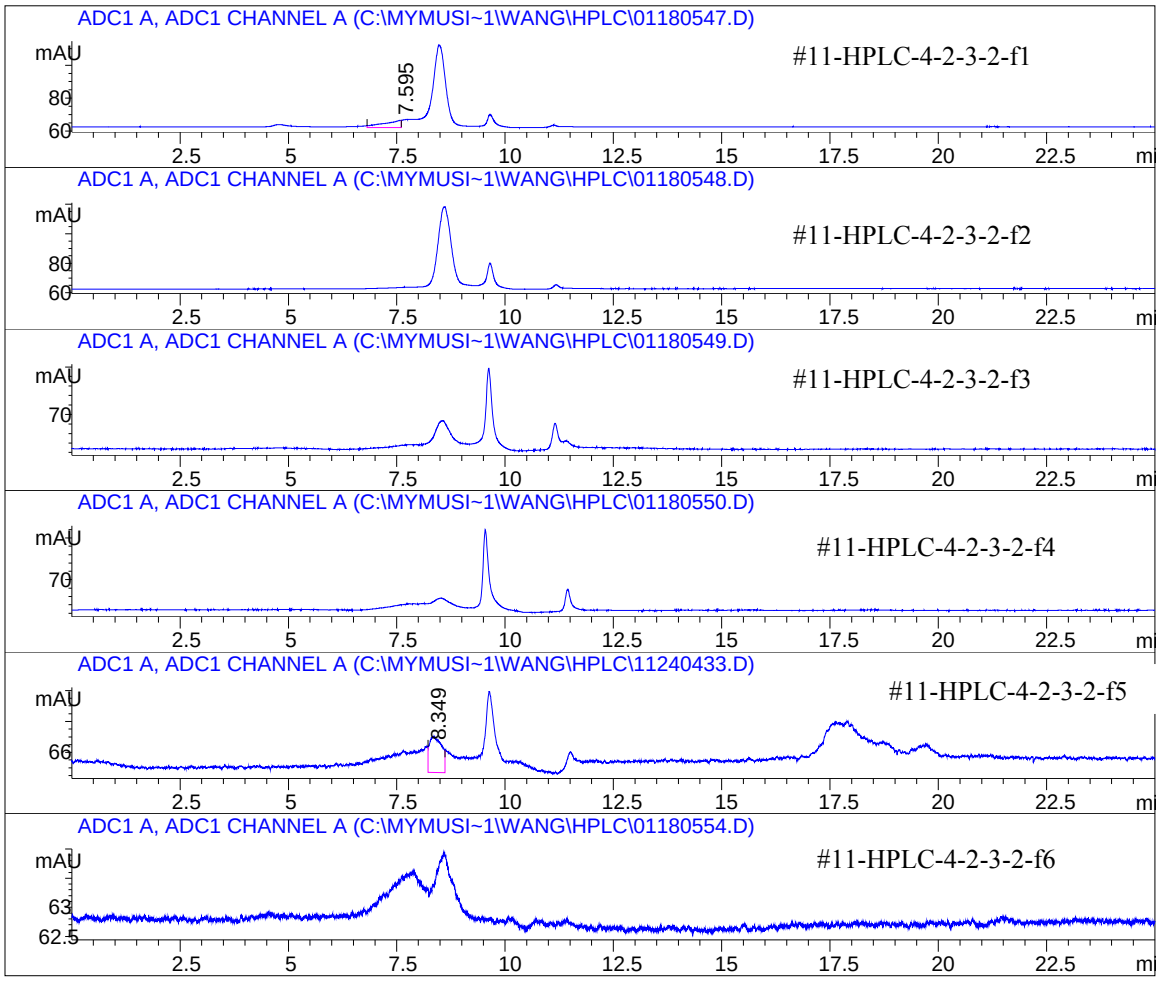
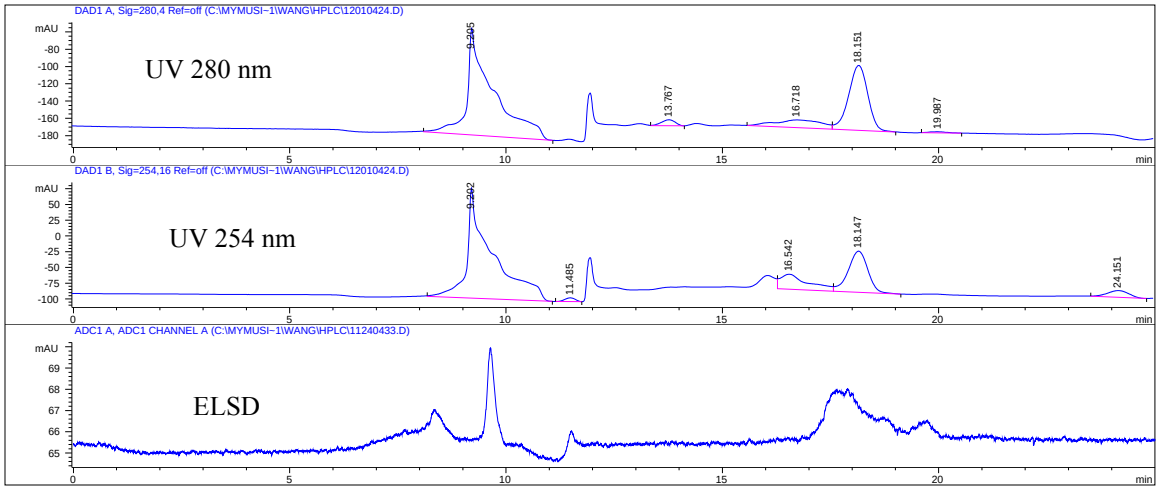


Figure 4.25. ELSD for HPLC separation of #11-HPLC-4-2-3-2.



Elution time (min)

Figure 4.26. ELSD and UV of #11-HPLC-4-2-3-2-f5.

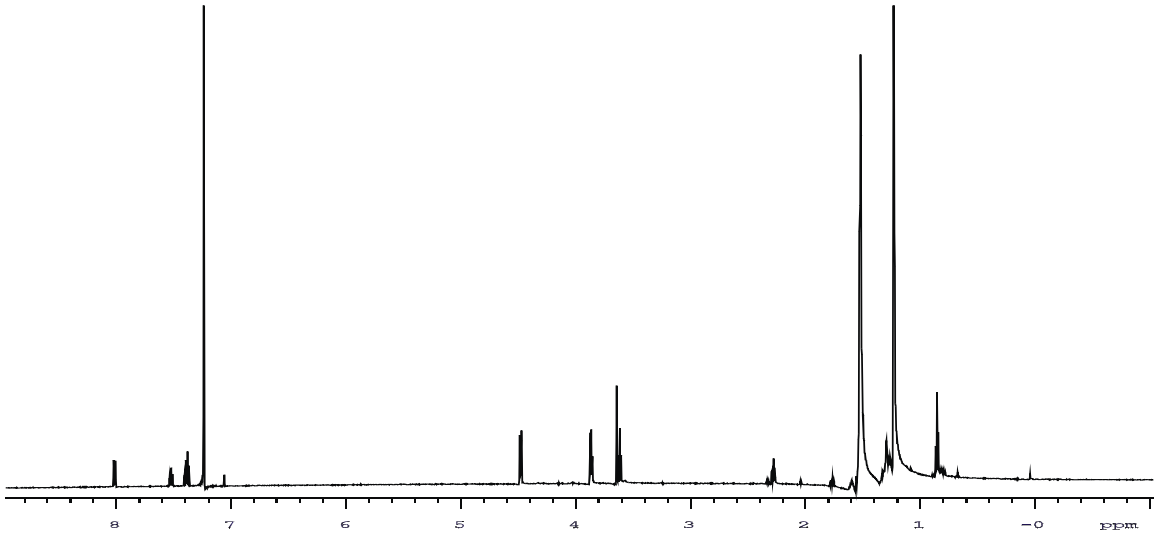


Figure 4.27. 600 MHz ^{1}H NMR of #11-HPLC-4-2-3-2-f5.

4.3.3 HPLC separation of #11-HPLC-2

This fraction isolated according to the flow chart of Figure 4.17 has the ^{1}H NMR spectrum shown in Figure 4.29, showing that it is a mixture of both aromatic and aliphatic compounds. The interesting thing about this fraction is that it contains both aromatic compounds and oligomers.

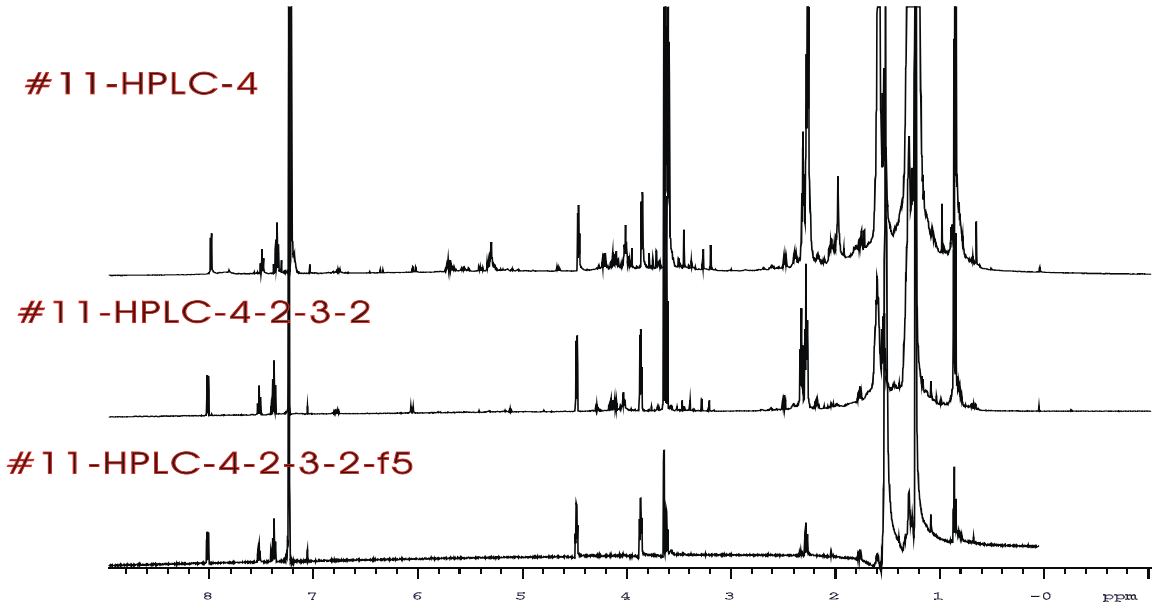


Figure 4.28. 600 MHz ^{1}H NMR of #11-HPLC-4 and its fractions.

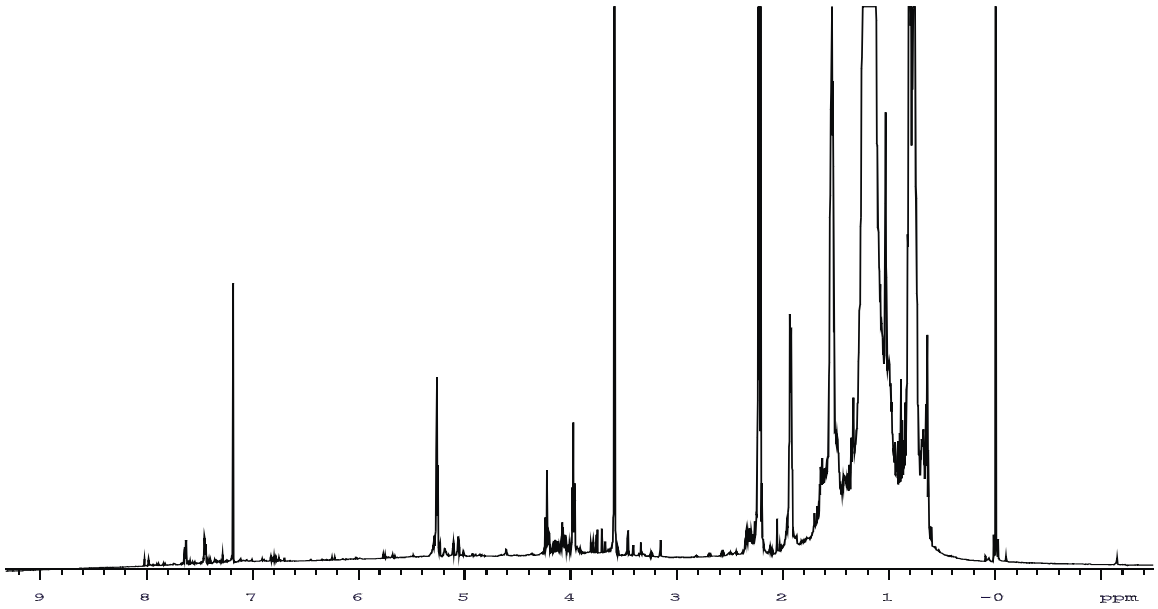


Figure 4.29. 600 MHz ^{1}H-NMR spectrum of #11-HPLC-2 before HPLC separation.

An isocratic HPLC method NP-1 (Table 4.9) was used to separate #11-HPLC-2. Figure 4.30 shows the ELSD of this #11-HPLC-2 separation, from which five fractions were collected: #11-HPLC-2-1 (0-5 min); #11-HPLC-2-2 (5-7.6 min); #11-HPLC-2-3 (7.6-9.5 min); #11-HPLC-2-4 (9.5-15 min); and #11-HPLC-2-5 (15-20 min).

Table 4.9 HPLC method NP-1

Time (min) Flow rate=0.5ml/min	Composition of mobile phase % (v/v)	
	Eluent C [c]	Hexane
0	0.2	99.8
20	0.2	99.8

[c] Eluent C : isopropanol/acetic acid (100:40)

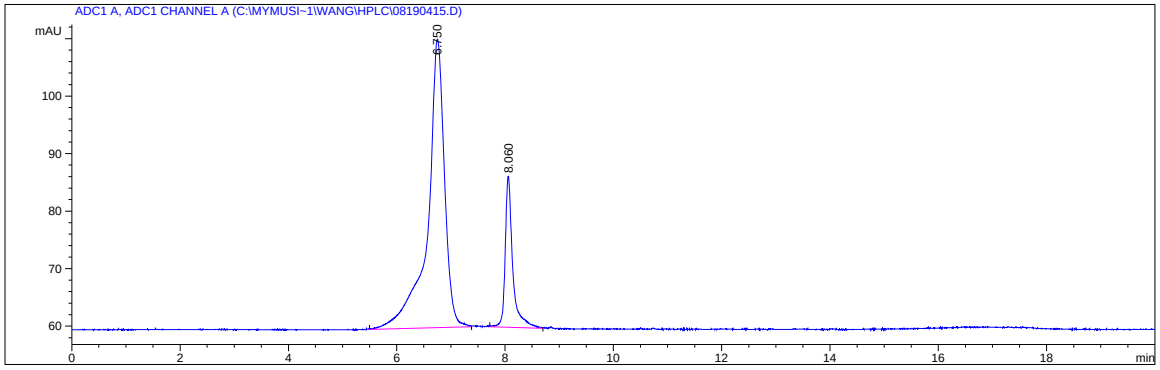


Figure 4.30. ELSD of #11-HPLC-2 separation.

Figure 4.31 shows the ^{1}H NMR spectra of these five fractions. Among them, #11-HPLC-2-2 (5-7.6 min) and #11-HPLC-2-3 (7.6-9.5 min) are mainly aliphatic monomers, which correspond to the two major peaks in the ELSD spectrum (Figure 4.30). #11-HPLC-2-1 (0-5 min), #11-HPLC-2-4 (9.5-15 min) and #11-HPLC-2-5 (15-20 min) each have more aromatic compounds and aliphatic oligomers, which are evidently present in amounts too modest to show up in the ELSD spectrum. This causes difficulty in isolating and identifying the target aromatic and aliphatic oligomers.

HPLC methods were also used to separate #11-HPLC-3 and #11-HPLC-5, both of which gave similar results to the aromatic dimer and #11-HPLC-2. Taken together, these difficulties led us to combine the five fractions of #11-HPLC-2 and proceed with alternate separation methods.

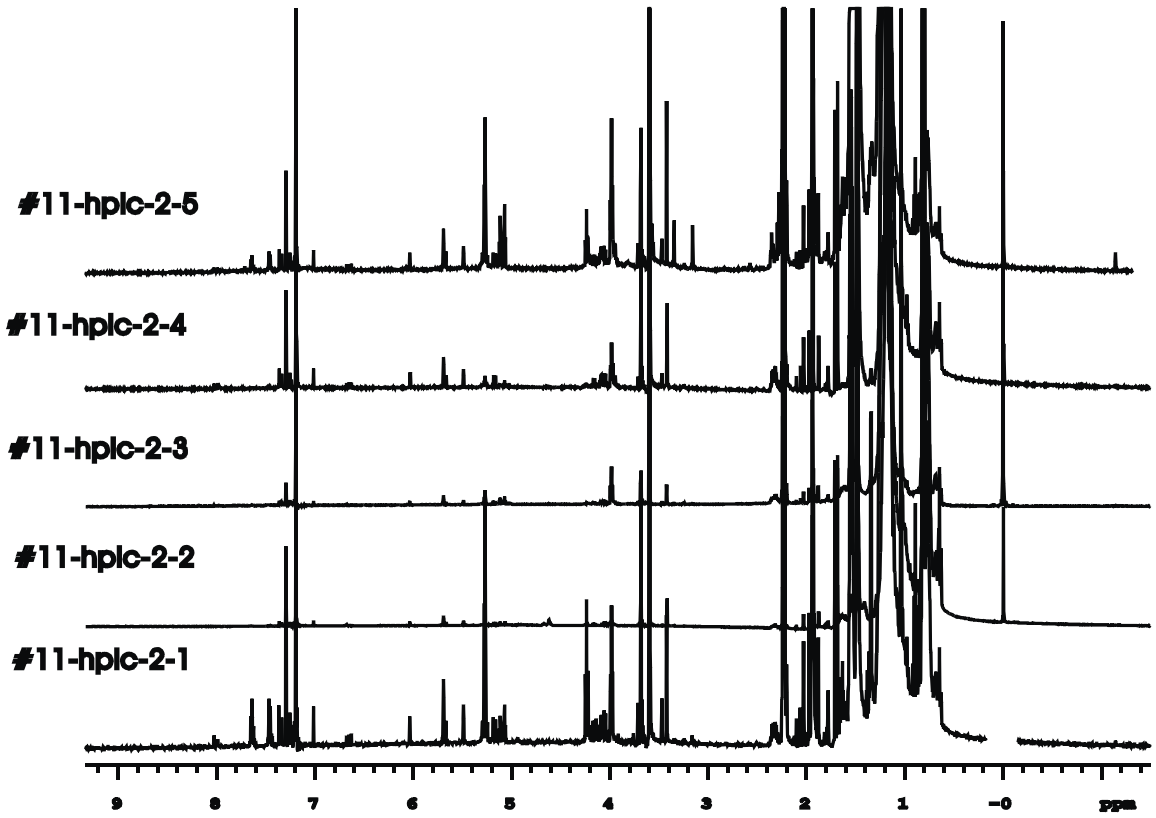


Figure 4.31. 600 MHz ^{1}H NMR of different fractions from #11-HPLC-2.

4.4. Solvent extraction methods

Fraction #11 is derived from preliminary separation by Column Chromatography (Figure 4.1). As noted above, the aromatic and aliphatic oligomers account for a very small portion of depolymerized suberin (less than 5%), and because they are distributed in different fractions after HPLC separation their relative amounts are very small. Moreover, repeated HPLC separation is required to achieve adequate purification, a process that is time-consuming, uses a lot of solvent and incurs sample losses. The separation of aromatics and aliphatic oligomers is not as complete as expected. For those reasons an attempt was made to use solvent extraction methods combined with ^{1}H NMR to guide the process. This method is simple, fast, not prone to lose precious sample, and

solvent conserving. However, the choice of solvent is crucial. At the beginning, different solvents were tried. The main solvents used were hexane, acetonitrile, methanol and ethyl acetate. Fig. 4.32 illustrates the solvent extraction process, showing part of sample #11-HPLC-2.

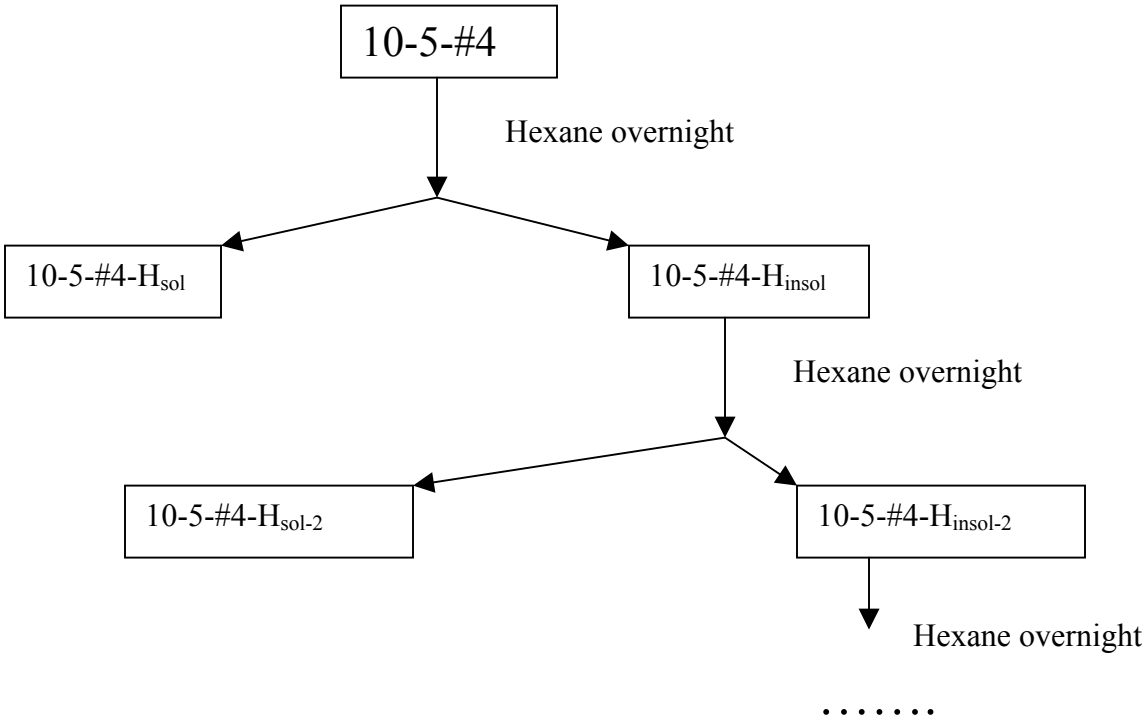


Figure 4.32. Partial flow chart of separation of 10-5-#4, illustrating extraction with hexane.

Figure 4.33 shows ^{1}H NMR spectra corresponding to the solvent extraction in Figure 4.32. As an example, Figure 4.33 (a) is the ^{1}H NMR spectrum of sample 10-5-#4. The characteristic peak at 4.05 ppm shows it contains an oligomer, is relatively pure, but still contains some monomers (revealed by HMBC, not shown). An overnight extraction with hexane was required to extract out the monomers: the hexane-soluble part didn't show the 4.05 ppm diagnostic oligomer peak, whereas the hexane-insoluble part did show this peak (Figure 4.33 b-c). Continued extraction of the hexane-insoluble part with hexane (Figure 4.32) yielded hexane-soluble and hexane-insoluble fractions with characteristic

spectral peaks at 4.03 ppm (Figure 4.33 d-e). Solvent extraction was continued for several rounds.

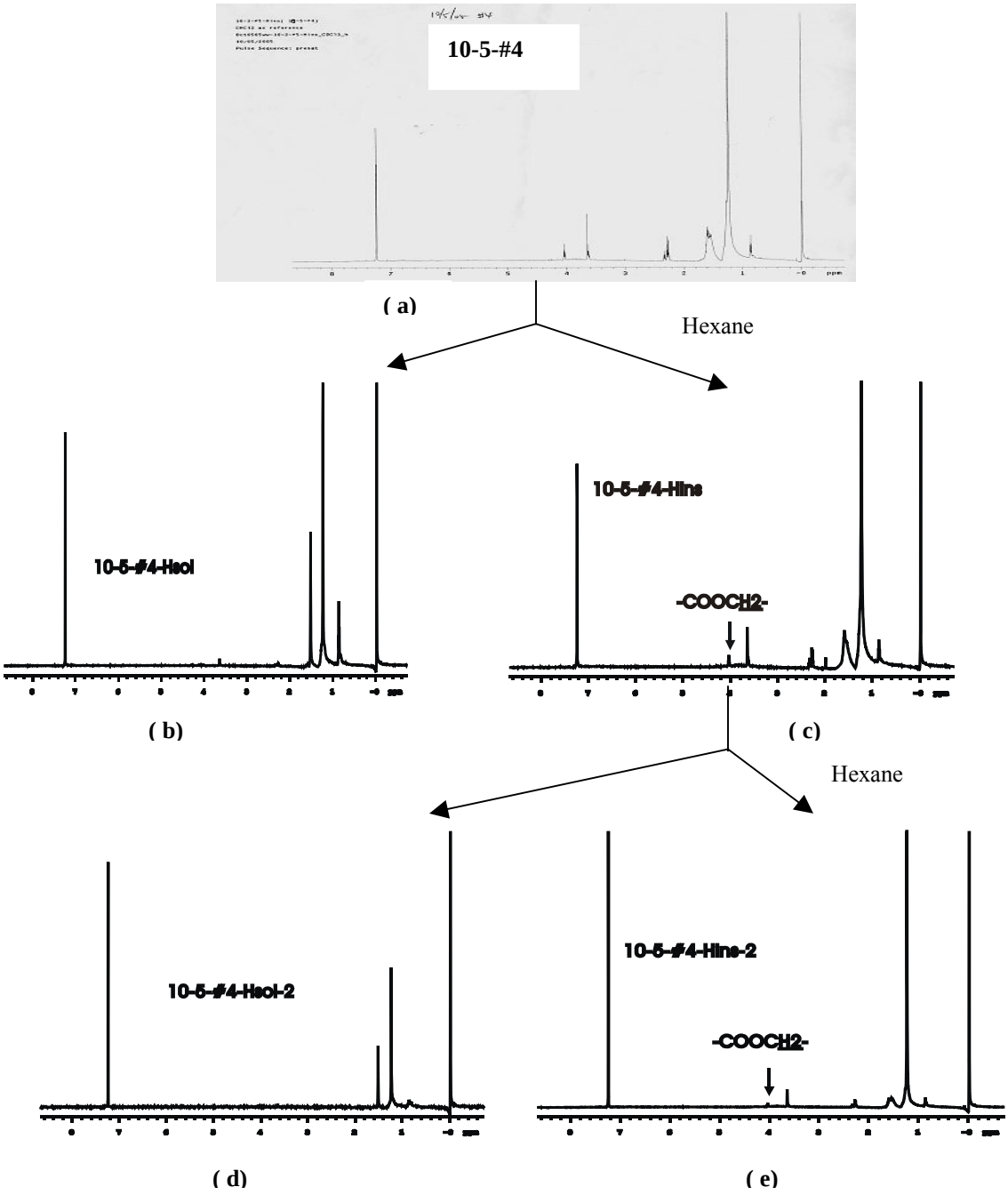


Figure 4.33. ^{1}H NMR corrsponding to partial flow chart for separation of 10-5-#4 (Fig 4.32).

Although the solvent extraction method is fast, simple, and solvent-sparing, it nevertheless shows similar shortcomings to the HPLC separation. The same solvent must

be used repeatedly to extract a given sample completely. The aromatic and aliphatic compounds seem to exist like layers of an onion that interweave with each other: only after the outside layer is peeled do we expose the inside layer. No precedent for this phenomenon has been reported.

The extraction strategy was used with solvents of differing polarity, (e.g. hexane, cyclohexane, *n*-pentane, acetonitrile, methanol, ethyl acetate); then ^{1}H NMR was used to decide which solvent should be used in the next step. To extract aromatic compounds, usually acetonitrile or methanol were used. To extract aliphatic oligomers, hexane, *n*-pentane or cyclohexane were most effective. Because a single extraction did not yield a complete separation, the above solvents were used repeatedly.

4.5 Combination of HPLC and solvent extraction methods

Several fractions that had not been purified by HPLC were subsequently separated by solvent extraction and monitored by 1D and 2D NMR. For some components that contain aromatic signals, the same solvent was used to extract the sample repeatedly to get rid of long-chain aliphatic compounds. In this case, acetonitrile, methanol and hexane were typically used in turn.

4.5.1 Solvent extraction of aromatic compound 3

The aromatic compound 3 was separated by CC and solvent extraction. As will be demonstrated, it is a compound in which the aromatic part is connected to aliphatic parts. Figure 4.34 shows the proton spectrum of this compound, which still has minor contaminants. It shows two aromatic peaks at 7.69 (m) and 7.51 ppm (m), corresponding to carbons at 128.5 and 130.5 ppm in the HSQC spectrum (highlighted in Figure 4.35),

respectively. The COSY spectrum (not shown) shows correlations between peaks at 7.69 and 7.51 ppm (the resonances are both multiplets). All of the above information indicates that this aromatic compound has a 1, 2-disubstituted benzene ring.

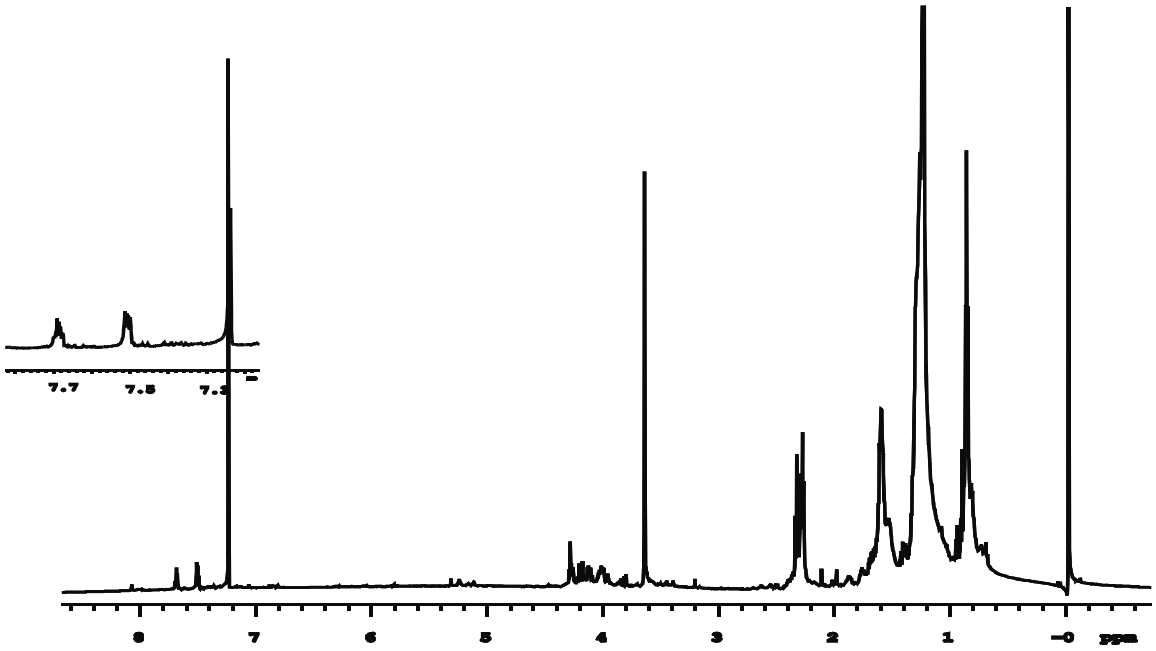
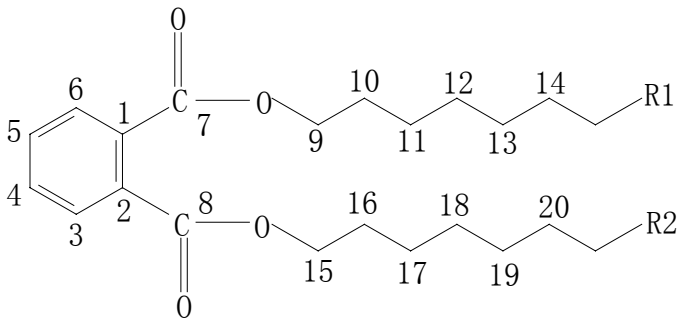


Figure 4.34. 600 MHz ^{1}H NMR spectrum of aromatic compound 3.



The gHMBC spectrum (Figure 4.36) shows that protons at 7.67 and 7.70 ppm correlate to a carbon with chemical shift 167.7 ppm. This means two carbonyl carbons (ester bond) connect to positions 1 and 2: in the gHMBC spectrum, the chemical shifts at positions 3 (7.67 ppm) and 6 (7.70 ppm) are slightly different, suggesting that the substituents on the two carbonyl carbons differ from one another.

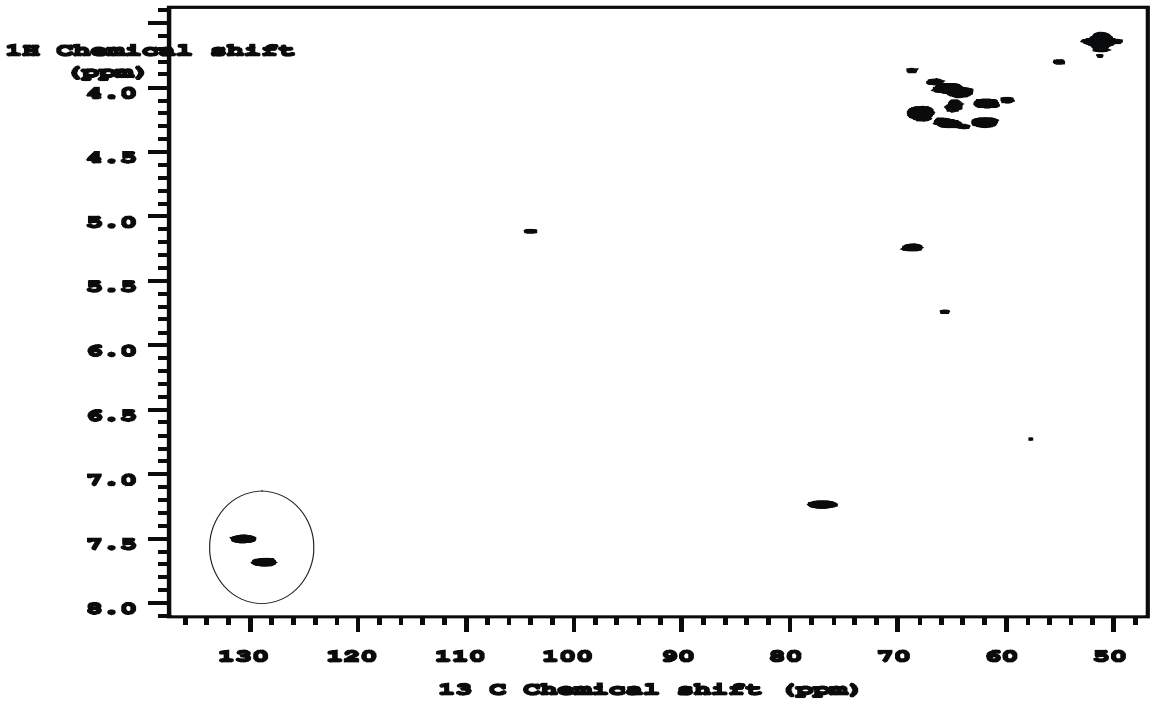


Figure 4.35. gHSQC spectrum of aromatic compound 3 (the highlighted peaks show correlations between aromatic protons and carbons as described in the text).

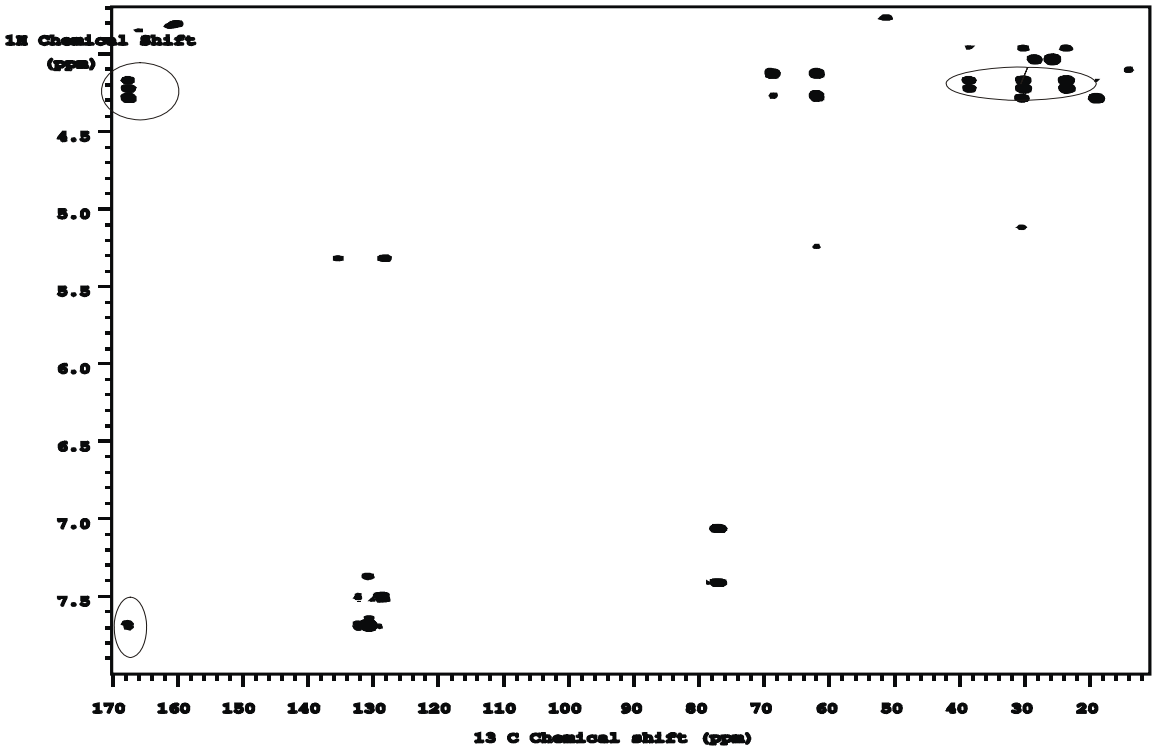


Figure 4.36. gHMBC spectrum of aromatic compound 3.

Table 4.10 Aromatic proton and carbon chemical shifts of compound 3

Position #	CH_n	Chemical shifts (ppm)	# of H's (predicted from Figure 4.37)	Proton Chemical shifts (ppm)
1	C	132.4	0	
2	C	132.4	0	
3	CH	128.5	1	7.69 (7.70-7.67)
4	CH	130.5	1	7.51 (7.51-7.50)
5	CH	130.5	1	7.51(7.51-7.50)
6	CH	128.5	1	7.69(7.70-7.67)

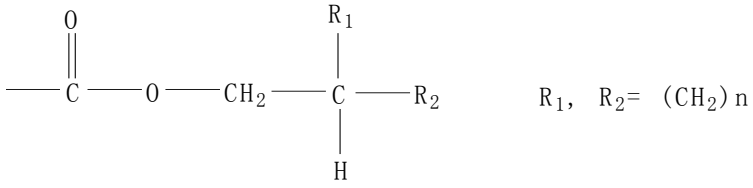
Table 4.11a HMBC (Figure 4.49) correlations of aromatic compound 3

H\C Position (ppm)	1 (132.4)	2 (132.4)	3 (128.5)	4 (130.5)	5 (130.5)	6 (128.5)	7 (167.7)	8 (167.5)
3 (7.70)	*				*			*
4 (7.51)						*		
5 (7.51)			*					
6 (7.67)		*		*			*	

Table 4.11b HMBC (Figure 4.38) correlations to aliphatic chains in aromatic compound 3

H\C Position (ppm)	7 (167.7)	8 (167.5)	? (38.6)	? (30.3)	? (23.5)	? (30.3)	? (18.8)
9 (4.17)	*		*	*	*		
9 (4.21)	*		*	*	*		
15 (4.28)		*				*	*

In the gHMBC spectrum, it is also clear that corresponding to the carbon at position 7 (structure above, 167.7 ppm), there are two cross-peaks at 4.17 and 4.21 ppm. Both of these protons display cross-peaks at 38.6, 30.3, and 23.5 ppm, implicating the following structure on the alcohol side of the methine carbon.



For the carbon at position 8 (167.5 ppm), there is only one cross-peak at 4.28 ppm, which in turn has cross-peaks at 30.3 ppm and 18.8 ppm. All of these features indicate that the other side of the ester bond for the carbon at position 8 is an unbranched chain.

After combining information from ^{1}H NMR (chemical shift, peak number, integration and J coupling constant) and 2D NMR (HMBC, HMQC, COSY correlation), a provisional structure was determined (Figure 4.37). ACD calculations (not shown) match the observations very well. This is a novel finding compared with previously reported suberin structures in which ferulic acid is esterified to a long aliphatic chain.[33] In order to confirm the final structure of the aromatic compound 3, MS information is needed (Chapter 5).

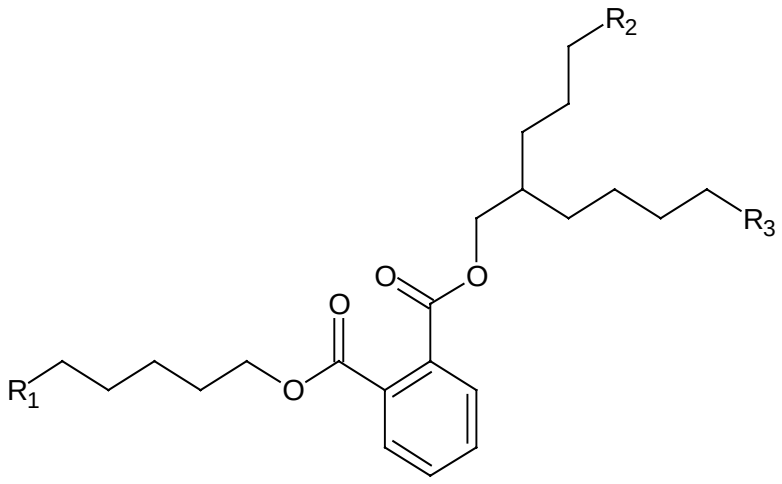


Figure 4.37. Provisional structure of aromatic compound 3.

4.5.2 HPLC separation of aromatic compound 3

After many rounds of solvent extraction using acetonitrile, methanol and hexane, with fractions sorted according to their respective ^{1}H NMR spectra, a solvent-extracted pre-purified aromatic mixture was obtained. 2D NMR shows that it is still a mixture containing one aromatic-aliphatic compound (aromatic compound **3**, Section 4.5.1), one aliphatic oligomer (later proven to be a triglyceride), and some aliphatic oligomers and monomers. Here the limitations of solvent extraction become evident: because different solvents were used to separate this relatively pure mixture, no further separation improvement could be achieved. HPLC trials were carried out with different columns, solvents, buffers, gradients, and flow rates. The reverse phase HPLC method shown in Table 4.12 was the best for separation of this mixture; the corresponding chromatogram is shown in Figure 4.38.

Table 4.12 RP-HPLC gradient Eluents for separation of aromatic mixtures

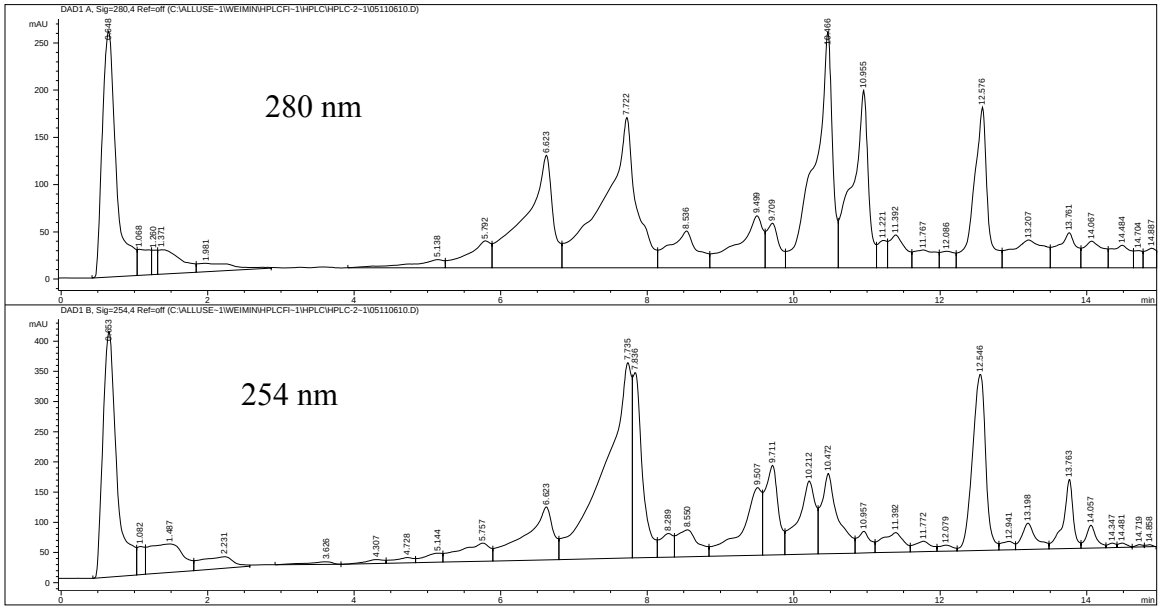
Time (min)	MeOH +0.04% TFA	Water + 0.05%TFA	Flow rate (ml/min)
0	50	50	0.9
15	100	0	0.9
Reverse phase column C: Symmetry Shield RP8, 4.6x50mm, 3.5 μm, Waters Corporation DAD=254 nm, 280 nm as detector			

Seven fractions were collected based on UV absorption at 254 and 280 nm. The fractions are: f1 (1-2 min), f2 (6.3-7.3 min), f3 (7.3-8.4 min), f4 (9-10.3 min), f5 (10.3-11 min), f6 (12.3-13.4 min), f7 (13.4-15 min). ^{1}H NMR spectra (not shown) indicate that f1, f2, f4, f5 are various aliphatic monomers, which are not our target compounds. The interesting fractions are: f3, f6, and f7; their ^{1}H NMR are shown in Figure 4.39.

f3 (7.3-8.4min), which has strong UV absorption at both 254 and 280 nm

f6 (12.3-13.4min), which has strong UV absorption at both 254 and 280 nm

f7(13.4-15min), which has strong UV absorption only at 254, not at 280 nm



Elution time (min)

Figure 4.38. RP-HPLC separation of aromatic mixture.

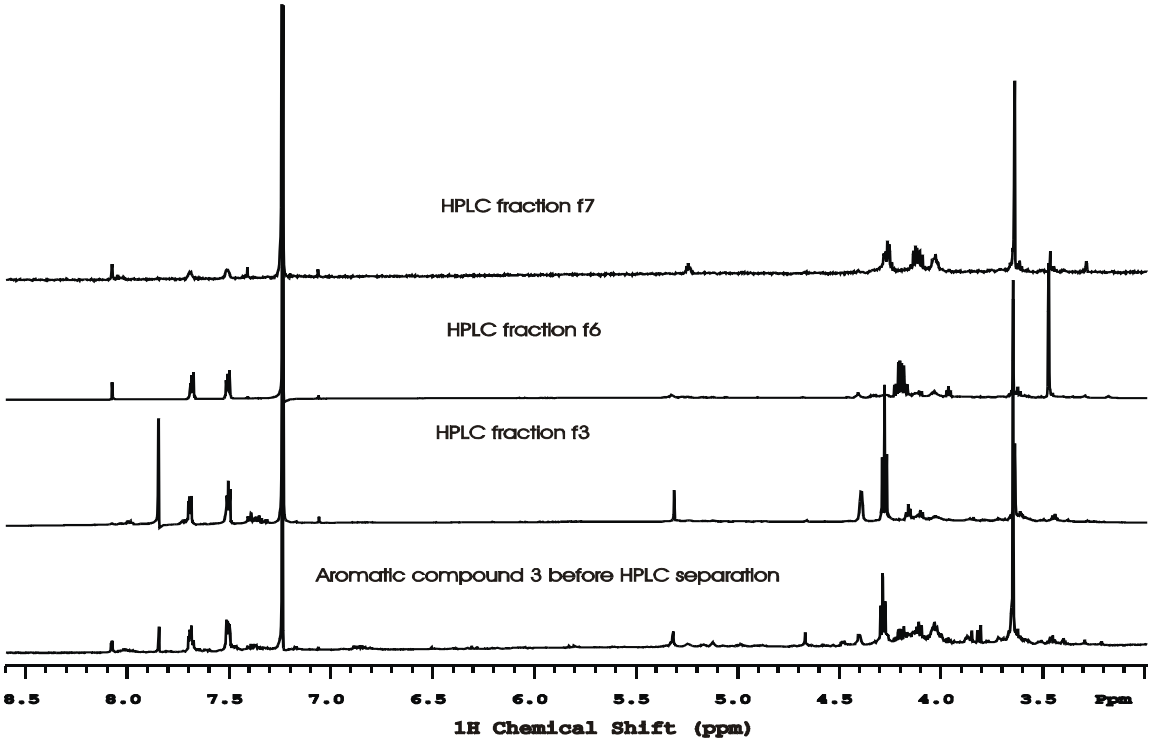


Figure 4.39. 600 MHz ^{1}H NMR spectra of aromatic compound 3 before and after HPLC separation of fractions f3, f6, and f7.

Before this HPLC separation, based on the NMR of the aromatic mixture, only one provisional structure was determined (Figure 4.37). After the HPLC separation, fractions f3 and f6 eluted at different times; their NMR spectra (Figure 4.39) have similar aromatic resonances and partial structures, but they are two different compounds. For the triglyceride compound f7 (Figure 4.39, top; discussed below), solvent extraction could not separate it from the aromatic-aliphatic compounds f6 and f3 (Figure 4.39, middle), but after reverse-phase HPLC, it was separated in pure form.

4.5.3. HPLC separation of a triglyceride

Fraction f7 from the above HPLC protocol was combined with another solvent extracted triglycerides 5-24-#5 (based on ^{1}H NMR) and dissolved in methanol. Two parts were obtained based on solubility. ^{1}H NMR shows that both of them contain characteristic peaks of triglycerides (see Section 5.2.1, NMR spectrum of standard triglyceride), but they are still mixtures with some other aliphatic oligomers and monomers.

MeOH soluble triglyceride: sample 5-24-#5, subjected to further HPLC separation.

MeOH-insoluble but hexane-soluble triglyceride: sample 5-24-#7 (^{1}H NMR, Figure 4.40), used directly in MS analysis.

Reverse phase HPLC (Table 4.13) was conducted for the methanol-soluble triglyceride sample 5-24-#5. Five fractions (Figure 4.41) were collected and subjected to ^{1}H NMR analysis: f3 (8-8.6 and 9.2-11 min), f4 (8.6-9.2 min), f5 (11-12 min) all show the characteristic peaks of a triglyceride (5.24, 4.28, and 4.13 ppm); fraction f4 is the purest one. The spectra are displayed in Figures 4.42, 4.43, and 4.44, respectively.

Table 4.13 RP-HPLC gradient eluents for separation of MeOH soluble triglycerides

Time (min)	MeOH +0.04% TFA	Water + 0.05%TFA	Flow rate (ml/min)
0	70	30	0.9
10	100	0	0.9
12	100	0	0.9
Reverse phase column C8: Symmetry Shield RP8, 4.6x50mm, 3.5 μm, Waters Corporation, DAD=254, 280 nm as detector			

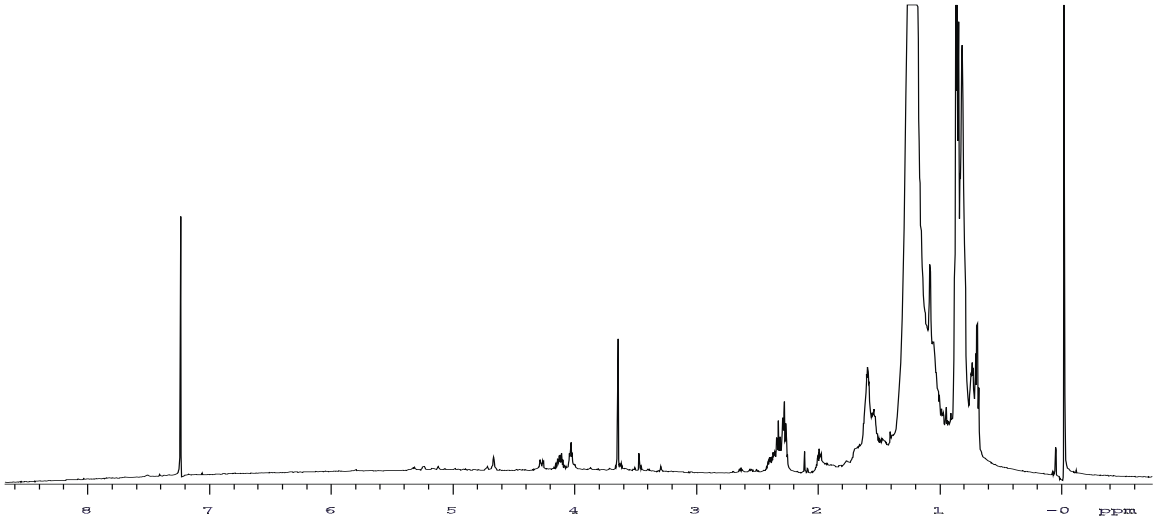
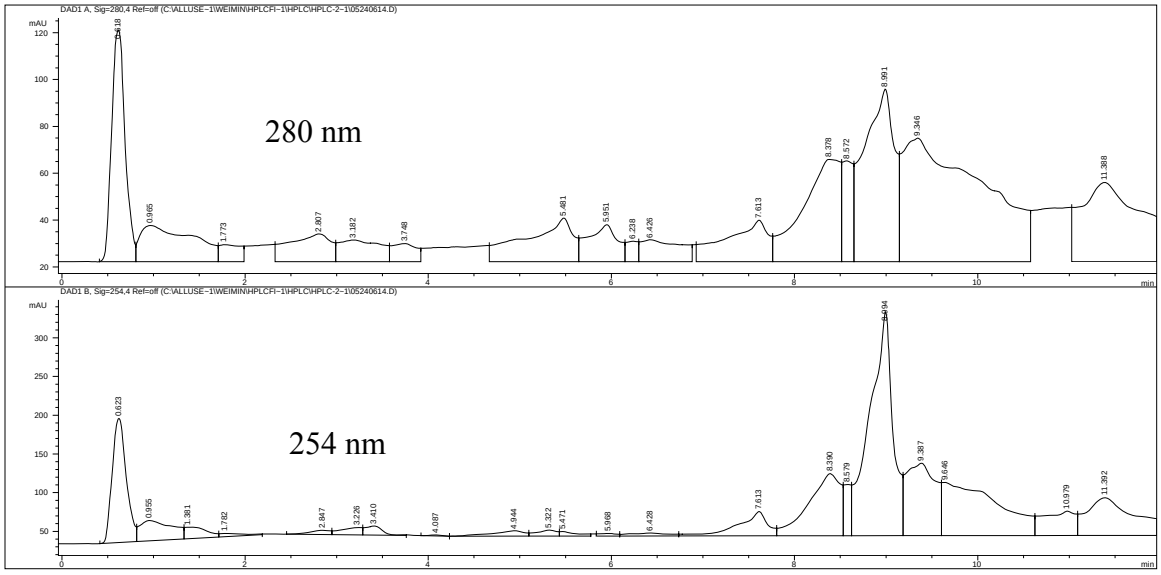


Figure 4.40. 600 MHz ^{1}H NMR of sample 5-24-#7.



Elution time (min)

Figure 4.41. RP-HPLC separation of 5-24-#5.

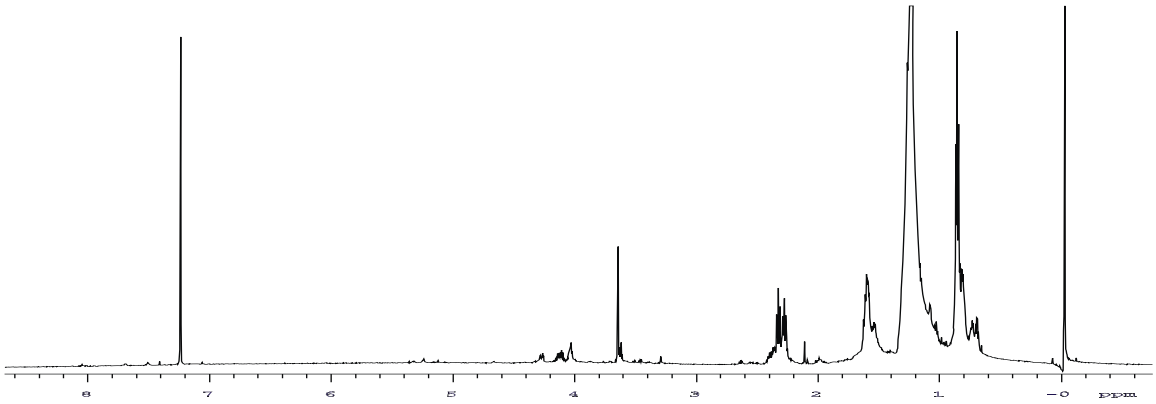


Figure 4.42. 600 MHz ^{1}H NMR of 5-24-#5-HPLC-f3.

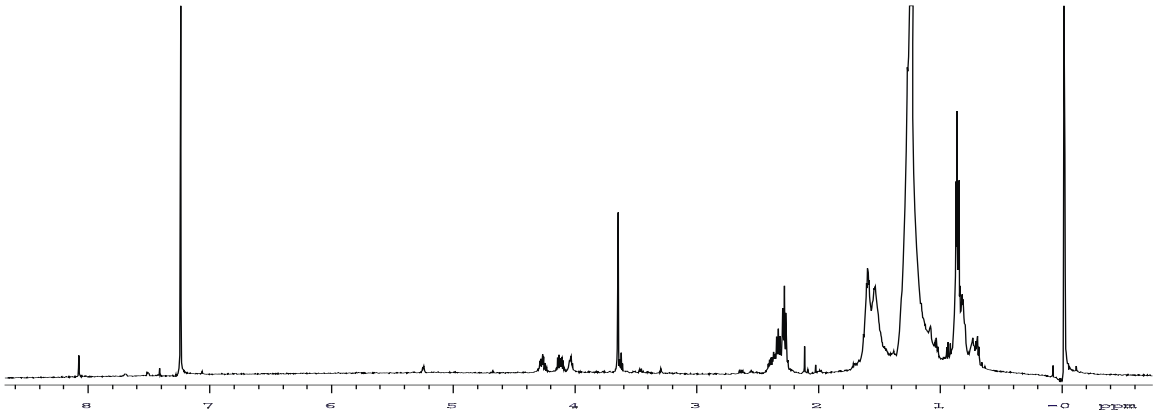


Figure 4.43. 600 MHz ^{1}H NMR of 5-24-#5-HPLC-f4.

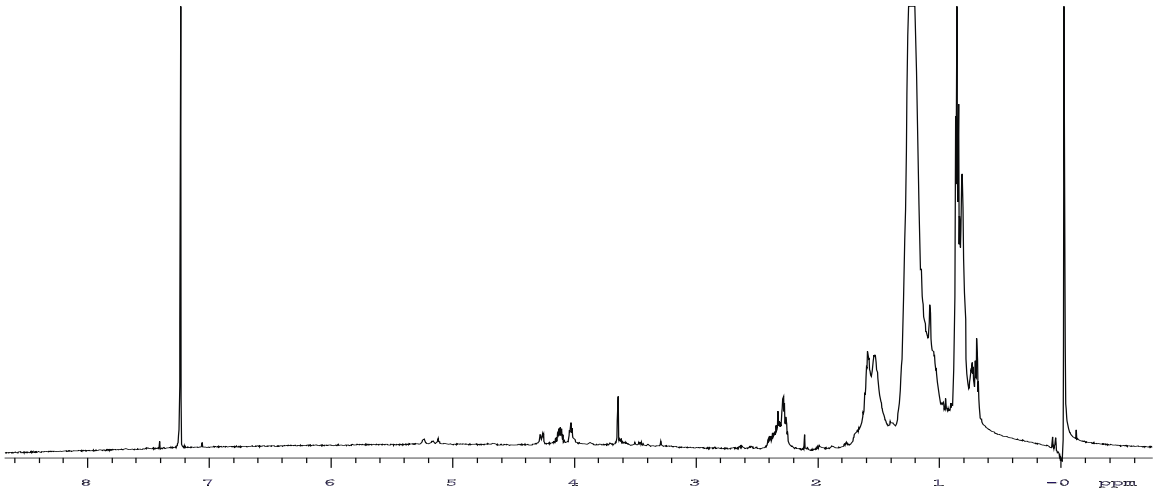


Figure 4.44. 600 MHz ^{1}H NMR of 5-24-#5-HPLC-f5.

In order to identify the structures of the above compounds with confidence, MS is needed to get the molecular weight. In case several possible structures exist, $MS^{(n)}$ is required (see Chapter 5).

4.6 Conclusions

Based on all of the above experimental experience, it is clear that the separation of suberin depolymerized products cannot be done using traditional methods of CC, TLC or HPLC. First organic solvent extraction should be used, but the resulting extracts may still contain complex mixtures differing in molecular mass, degree of unsaturation, and the number and kind of functional groups. No single analytical method is capable of separating and identifying them. Thus the analysis of suberin fragments requires multistep solvent extraction, stepwise chromatographic (CC, TLC, HPLC) resolution of relatively purer composition combined with NMR-based structural elucidation. Finally, ACD spectral simulation and MS experiments should be used to confirm the final proposals.

Chapter 5. RESULTS AND DISCUSSION: STRUCTURAL ELUCIDATION OF ALIPHATIC AND AROMATIC COMPOUNDS

Whereas many previous reports have appeared on suberin monomers,[1,33] our research objective is to obtain information about oligomers, aromatics and connections between the aliphatic and aromatic groups. In this chapter, only new aliphatic and aromatic monomers and diverse oligomers are reported.

5.1 Aliphatic compound 1: glycerol

Aliphatic compound **1** was separated by solvent extraction of the soluble suberin depolymerization products. It was elucidated as glycerol. Even though this molecule has been found earlier in potato suberin,[43,46,47] it is still reported here because it provides direct evidence related to the discovery of triglycerides (see Section 5.2). It is present in a mixture with aromatic compound **4** (see Section 5.7). 1D and 2D NMR confirm its identity.

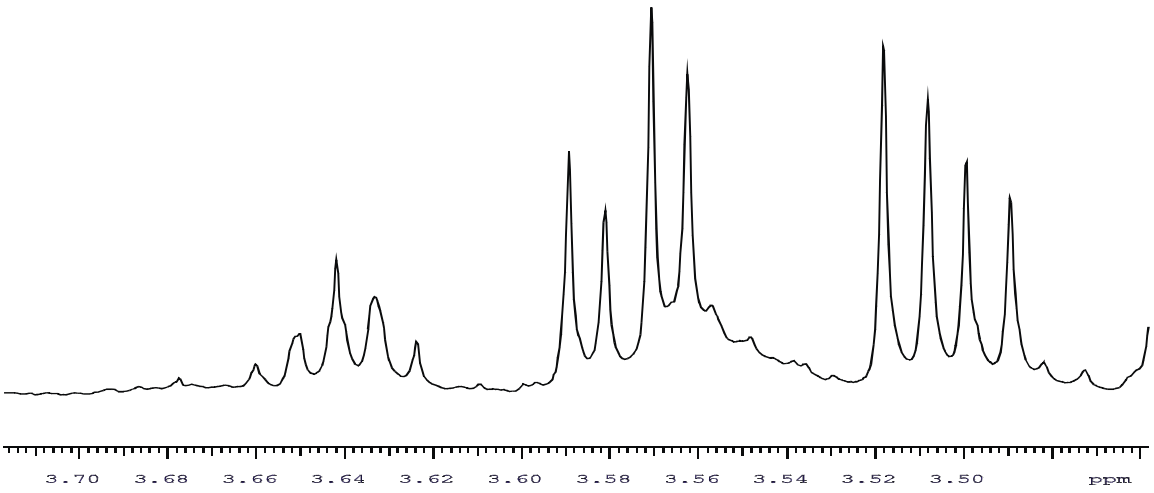


Figure 5.1. 600 MHz ^{1}H NMR spectrum of aliphatic compound 1, glycerol.

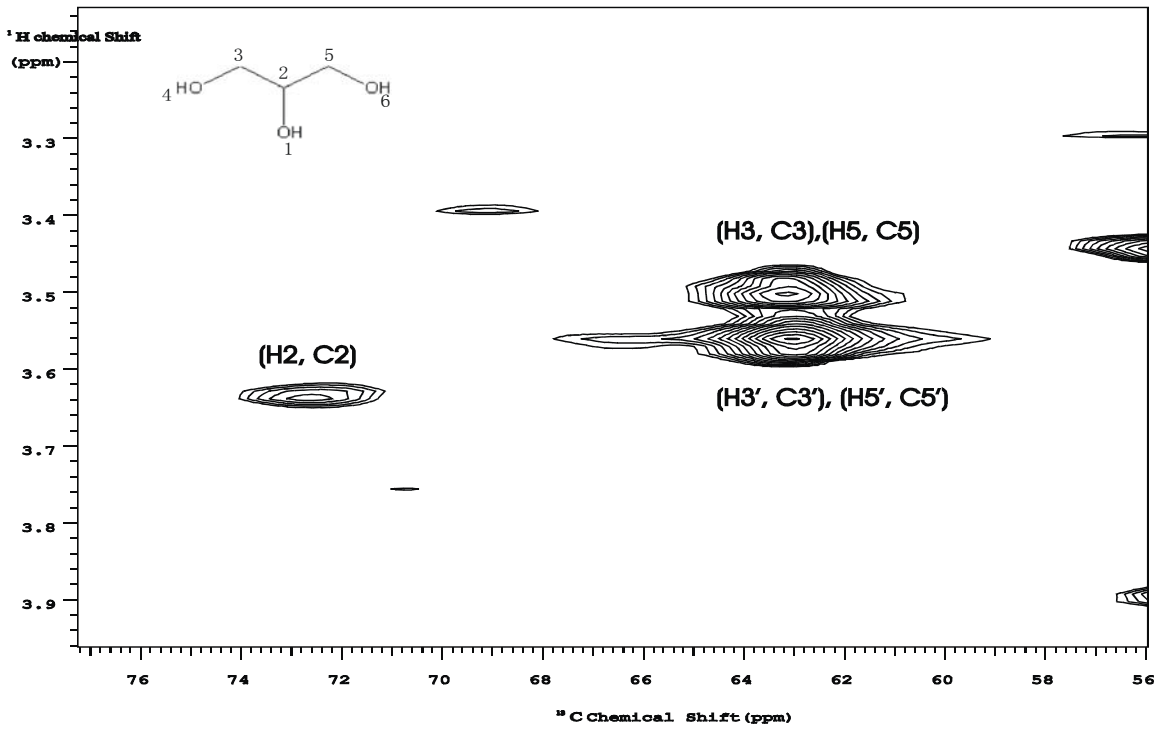


Figure 5.2. **gHSQC spectrum of aliphatic compound 1, glycerol.**

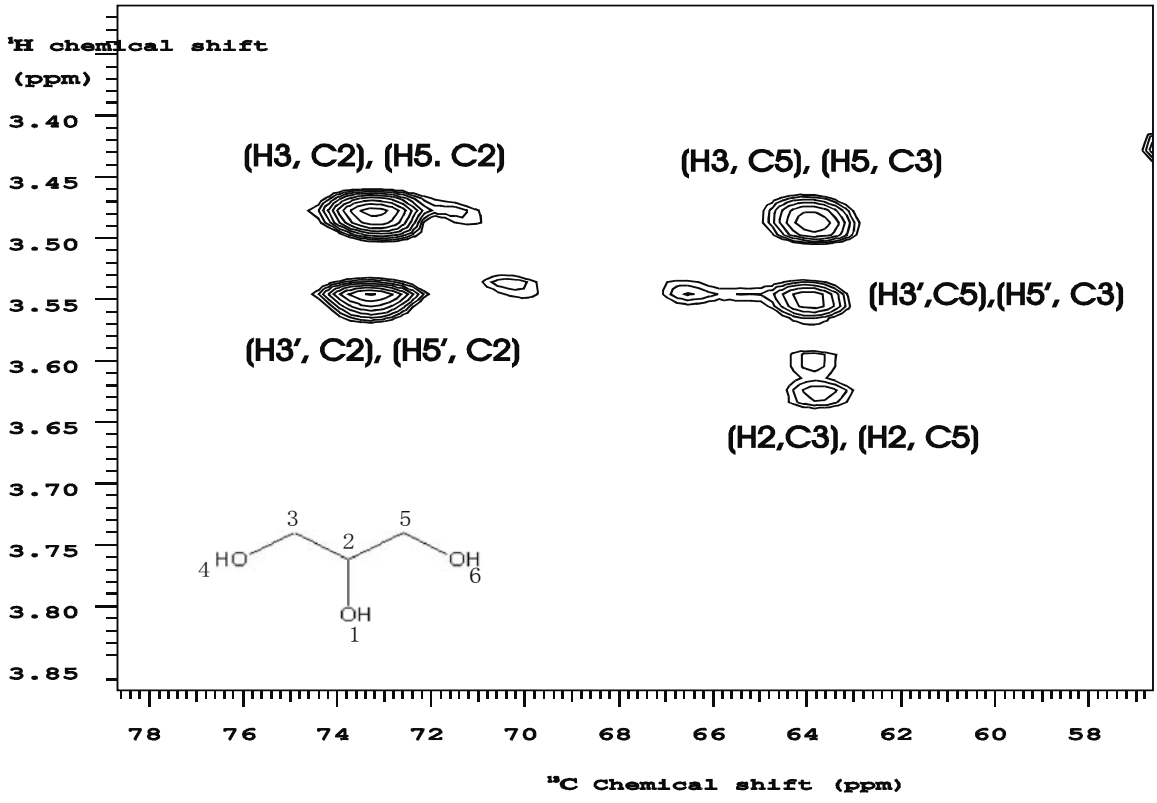


Figure 5.3. **gHMBC spectrum of aliphatic compound 1, glycerol.**

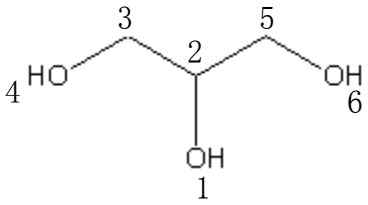


Figure 5.1 shows the ^{1}H spectrum of glycerol, revealing three proton peaks at 3.65, 3.58, 3.53 ppm and corresponding to ^{13}C's at 73.8, 64.3, 64.3 ppm, respectively (Figure 5.2, gHSQC). Table 5.1 shows the excellent match between chemical shifts obtained from experiment compared with the result calculated from ACD (not shown).

Table 5.1 Theoretical and experimental NMR data for compound 1, glycerol

Proton number	Predicted number of protons	Calc. Chemical shift (ppm)	Obs. Chemical shift (ppm)
2	1	3.77	3.65
3	2	3.63/3.55	3.58/3.53
5	2	3.63/3.55	3.58/3.53

Carbon number	CH_n	Calc. Chemical shift (ppm)	Obs. Chemical shift (ppm)
2	$\underline{C}H$	74.3	73.8
3	$\underline{C}H_2$	65.1	64.3
5	$\underline{C}H_2$	65.1	64.3

Table 5.2 gHMBC correlations (*) of aliphatic compound 1, glycerol

H\ C positions (ppm)	C-3 (64.3)	C-2 (73.8)	C-5 (64.3)
H-3 (3.53)		*	*
H-3 (3.58)		*	*
H-2 (3.65)	*		*
H-5 (3.53)	*	*	
H-5 (3.38)	*	*	

Figure 5.3 displays the gHMBC spectrum and Table 5.2 summarizes the correlations that confirm the structure of aliphatic compound **1** as glycerol. Because the molecular mass of glycerol is 92.047 Da, which is below the threshold for accurate LC/MS determinations, no MS data were obtained.

5.2 Aliphatic compound 2: triglycerides

5.2.1 NMR identification of triglycerides

5.2.1.1 Triglycerides from depolymerized suberin

From solvent extraction of the soluble depolymerization products of suberin, a partially purified material was elucidated as triglycerides with aliphatic chains R_1, R_2 and R_3 that may be C_8 to C_{18} chains with double bonds in the middle. The tentative structure is shown in Figure 5.4; confirmation is obtained by MS or MS^n as shown below.

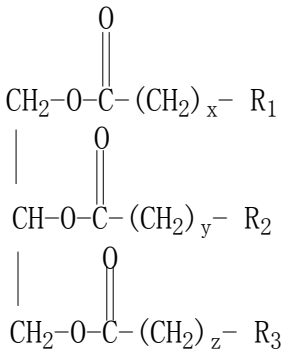


Figure 5.4. Tentative structure of aliphatic compound 2, a triglyceride.

The ^1H spectrum of this material (Figure 5.5) shows a diagnostic peak for ester oligomers at 4.03 ppm (peak 1). The most informative ^1H resonances are at 5.24 ppm (peak 3), 4.28 ppm (peak 4) and 4.13 ppm (peak 5), corresponding to carbons at 68.8, 61.9, and 61.9 ppm, respectively, in the gHSQC spectrum (Figure 5.6). The integration of ^1H resonances in the 600 MHz NMR spectrum basically supports the glycerol structure (Figure 5.4). The integral of peak 1 to peaks 2 and 3 is 1:5.5 (theoretical integration is 1:4), but from the gHMBC spectrum, it is clear there are some additional compounds with ^1H signal intensity at 4.13 ppm, which can explain the higher-than-expected integration ratio.

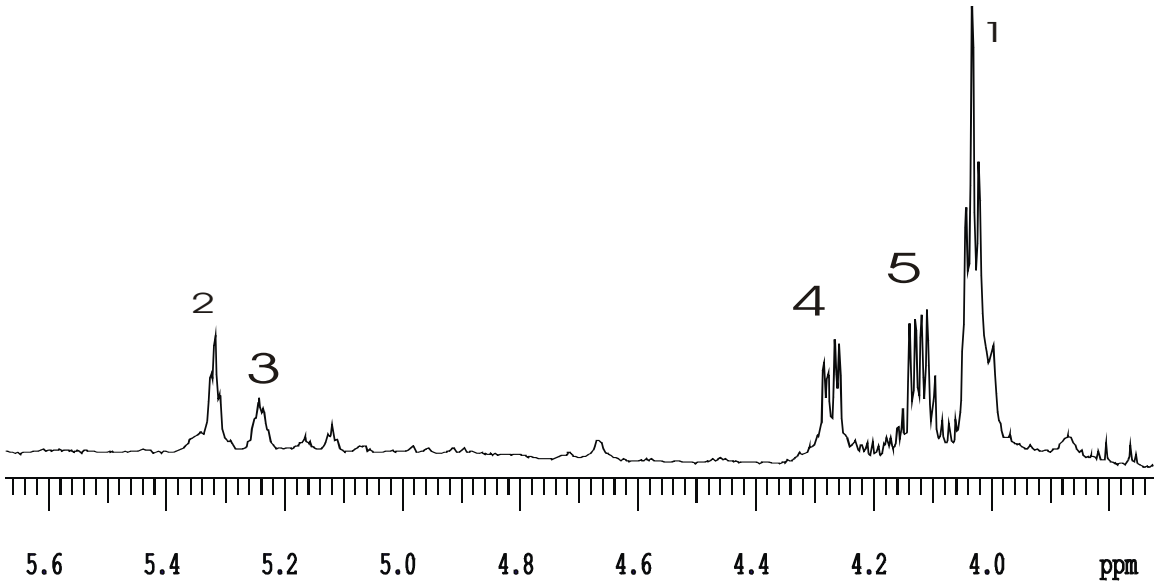


Figure 5.5. 600 MHz ^{1}H NMR spectrum of aliphatic compound 2.

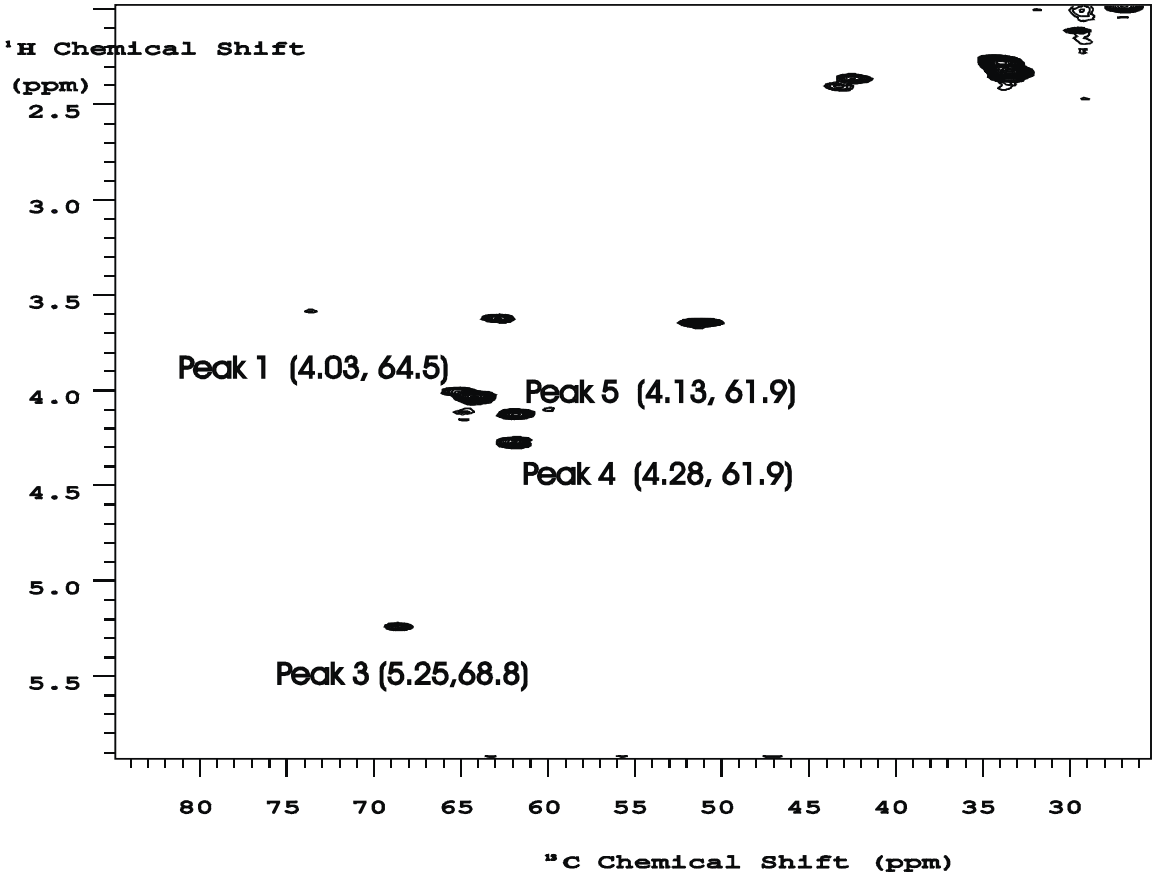
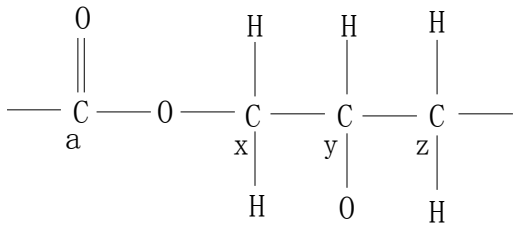


Figure 5.6. gHSQC spectrum of aliphatic compound 2.

The COSY spectrum (Figure 5.7) shows correlations between peaks 4 and 5; 3 and 4; 3 and 5. Besides these, peaks 3, 4 and 5 have no other correlations. The gHMBC spectrum (Figure 5.8) shows two correlations at (4.13, 172.9 ppm) and (4.28, 172.9 ppm). These two cross-peaks are diagnostic for ester bonds.[69] They show correlations between protons from the hydroxyl side with the carbonyl carbon. The gHSQC spectrum (Figure 5.6) shows that these two protons are correlated to the same ^{13}C, which indicates they are connected to the same carbon. When two protons from the same carbon show different chemical shifts, it might be deduced that the next carbon is a methine (tertiary carbon). That is, this methine carbon makes the two protons that are connected to its neighboring carbon inequivalent, with chemical shifts of 4.13 and 4.28 ppm. Because the chemical shift of the proton at position y (corresponding to peak 3 in Figure 5.5) is 5.24 ppm, there must be an oxygen connected to C_y (y refers to symbol in the partial structure below).

Surprisingly, the gHMBC spectrum exhibits no other correlations to the proton at position y. The protons at 4.13 and 4.28 ppm both show correlations to the carbon at 61.9 ppm, which appear at first to be correlations of the protons at position x to the carbon at position x, as observed in the gHSQC spectrum (Figure 5.6), but are unanticipated in the gHMBC spectrum. The only possible explanation is a symmetric structure, in which the chemical shifts of the two protons at position z are also 4.13 and 4.28 ppm. Thus the correlations at (4.13, 61.9) and (4.28, 61.9) actually correspond to H_z-C_x and H_x-C_z. The structural fragment consistent with these data is:



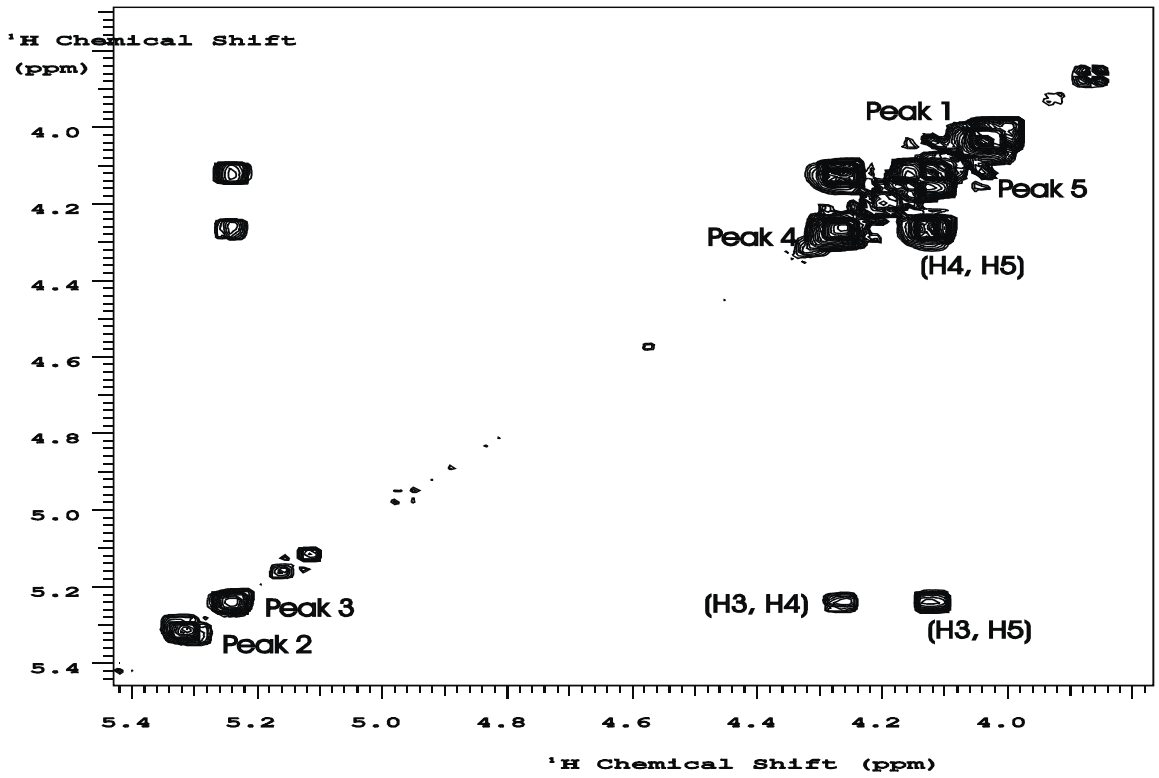


Figure 5.7. COSY spectrum of aliphatic compound 2.

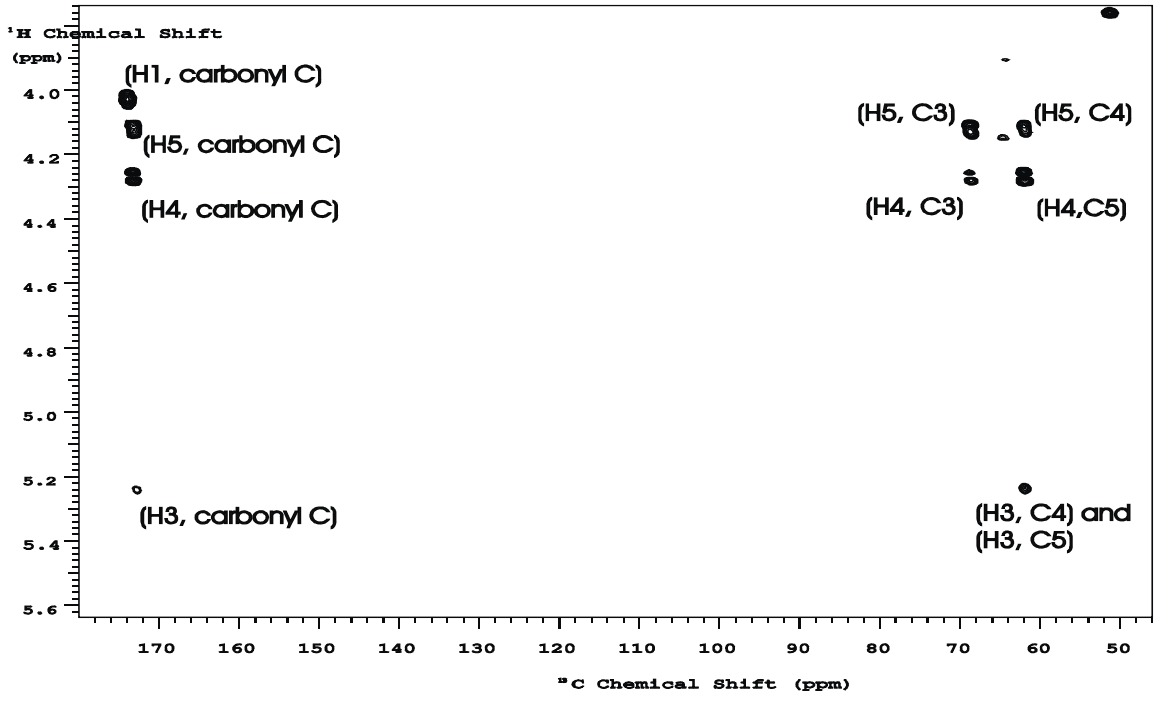


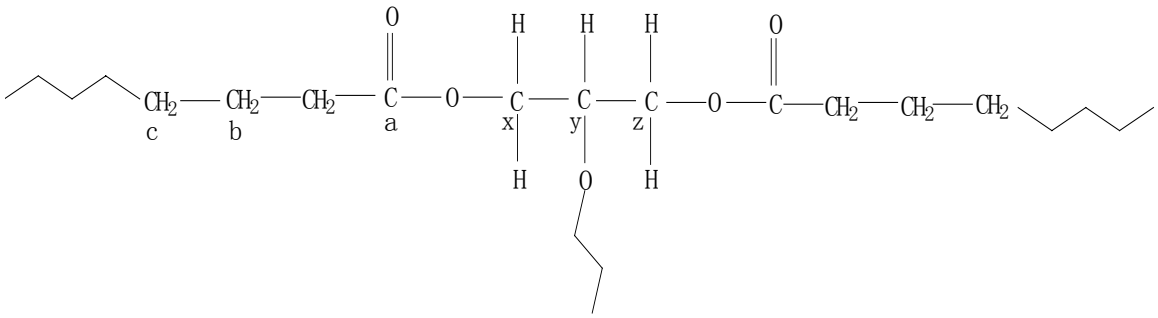
Figure 5.8. gHMBC spectrum of aliphatic compound 2.

Table 5.3 also shows another gHMBC (Figure 5.8) correlation of modest intensity for H_y at 5.24 ppm with a carbonyl carbon at 172.7 ppm. If this peak is real, this compound could be a triglyceride, i.e. the branch that comes from C_y may connect to another ester bond. The small size of this cross-peak could be explained by the fact that position y has only one proton, or because of an anomalously small (or large) vicinal coupling constant $^3J_{CH}$ that is not well matched by the delay time ½ J in the standard HMBC pulse sequence.

Table 5.3 gHMBC (Figure 5.8) correlation of aliphatic compound 2, a triglyceride

H \ C positions (ppm)	C_a (173.0)	C_x (61.9)	C_y (68.8)	C_b (28.9)	C_c (24.5)
H_x (4.13)	*	*	*		
H_x (4.28)	*	*	*		
H_y (5.24)	Another Carbonyl at (172.7)	*			
Other H (2.28)	*			*	*

- Letters refer to the symbol in the partial structure below.



Karplus and Conroy[60] predicted the theoretical dependence of vicinal coupling constants on the dihedral angle, ϕ:

$$^3J = A + B \cos \phi + C \cos 2\phi$$

This equation, typically represented as the Karplus curve, shows that if the dihedral angle is 0^o or 180^o, ^3J reaches a maximum, but if the dihedral angle is small, ^3J is small. For the

hypothesized structure, the dihedral angle of the carbonyl carbon connected to C_y is 60° from the carbonyl connected to C_a (see figure above), so that ^3J of H_y with the connected carbonyl should be about 3 Hz according to the Karplus curve (not shown). Thus the parameter jnxh was lowered to 5 (rather than the default value of 8); with a run time comparable to previous experiments, the cross peak (5.24, 172.7) appeared much stronger than before. This result confirms that another ester bond is connected to C_y. Moreover, the gHMBC spectrum with jnxh=5 (Figure 5.9) yields more detailed information than the one with jnxh=8 (Figure 5.10).

The proton spectrum (Figure 5.5) suggests that the resonance at 4.28 ppm actually corresponds to two groups of peaks, though the situation for the signal at 4.13 ppm is not so obvious. By contrast, Figure 5.9 (jnxh=5) shows clearly that the signals at 4.13 and 4.26 ppm each correspond to two different peaks. Figure 5.10 (jnxh=8) leaves this result ambiguous. Thus, protons H_x and H_z are actually inequivalent, because on the carboxylic side, the chains may be different. Based on the above information, the tentative structure was elucidated (Figure 5.4). In order to get complete structural information, MS data are needed (see below). ACD simulations of the ^1H and ^{13}C NMR spectra (Figures 5.11, and 5.12) match the tentative structure very well.

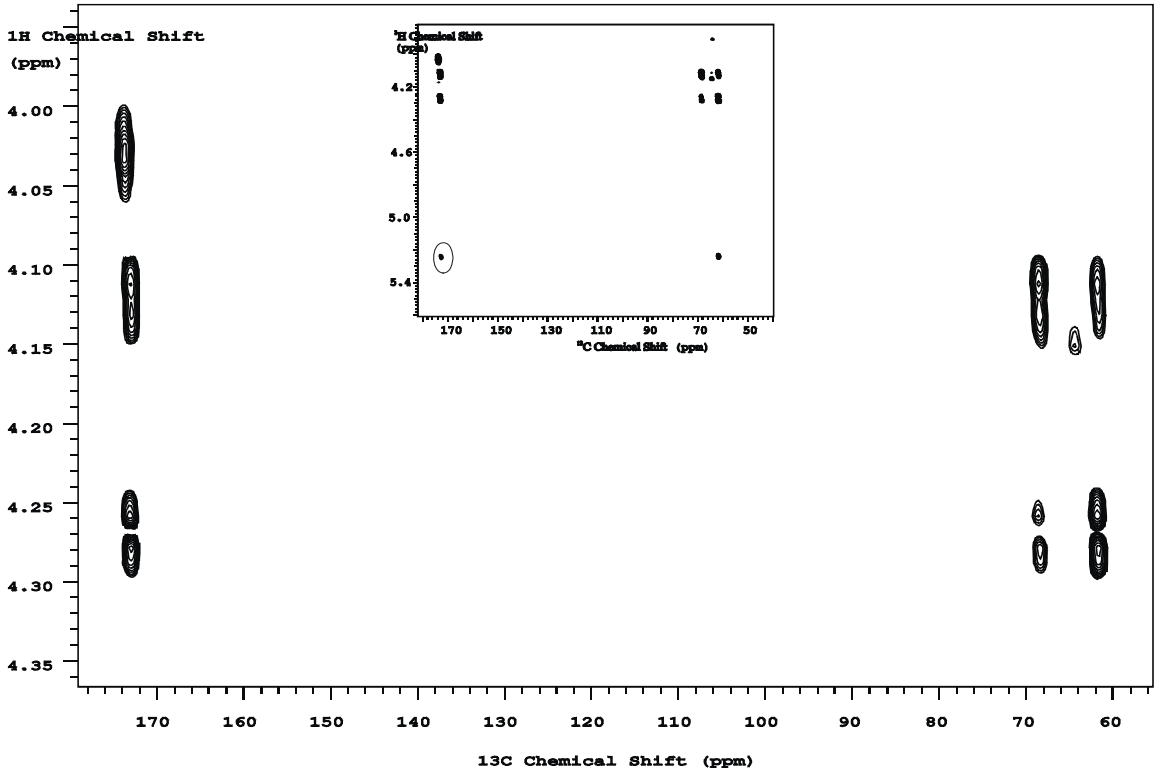


Figure 5.9. gHMBC spectrum of the triglyceride with jnxh=5. (The insert shows correlation of 5.24ppm with 172.7 ppm is stronger for jnxh=5 than that of jnxh=8)

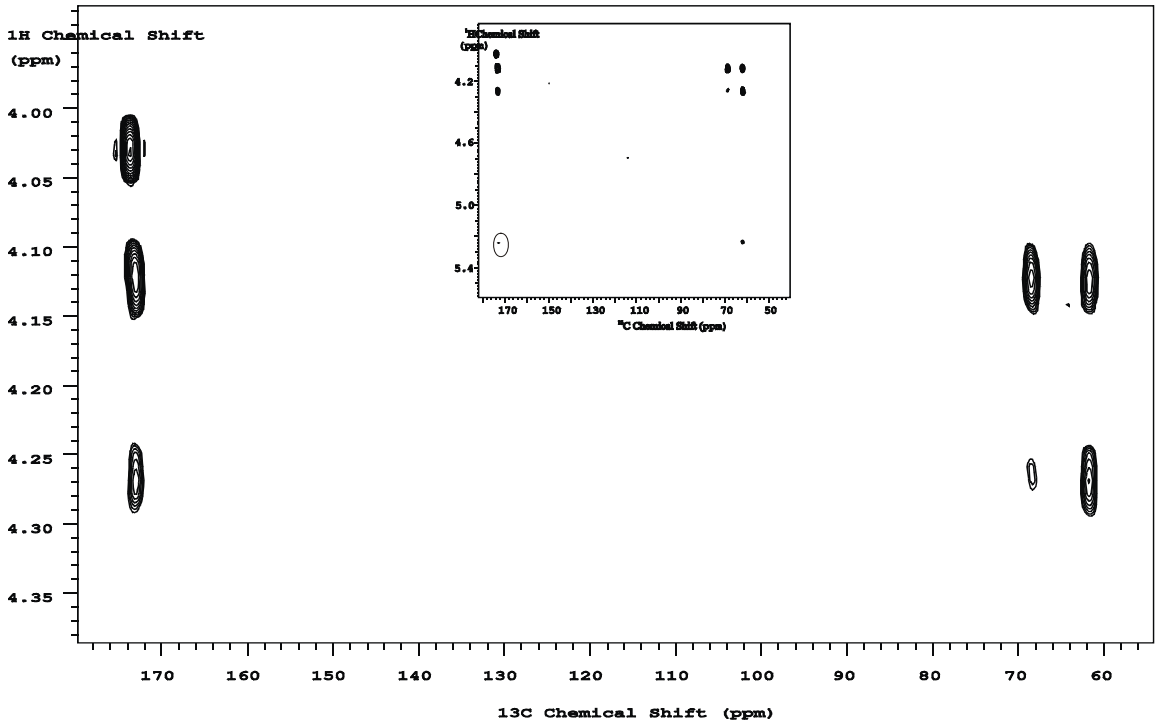
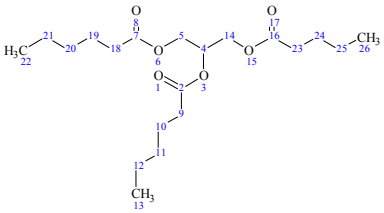


Figure 5.10. gHMBC spectrum of the triglyceride with jnxh=8.



04/Jan/2006 19:44:20 ACD/HNMR Predictor (v.9.02)

Group	nH	Shift	Error	Av.Exp
4	1	**5.26**	0.30	5.05..5.35
5<''>	1	**4.33**	0.41	4.33..4.43
5<'>	1	**4.09**	0.41	4.11..4.17
9	2	**2.28**	0.16	2.20..2.37
10	2	**1.64**	0.19	1.63..1.71
11	2	**1.29**	0.10	1.26..1.36
12<''>	1	**1.27**	0.18	1.23..1.31
12<'>	1	**1.26**	0.18	1.24..1.30
13	3	**0.88**	0.08	0.87..0.92
14<''>	1	**4.33**	0.41	4.33..4.43
14<'>	1	**4.09**	0.41	4.11..4.17
18	2	**2.34**	0.13	2.34..2.37
19	2	**1.62**	0.19	1.63..1.65
20	2	**1.30**	0.10	1.26..1.36
21<''>	1	**1.27**	0.18	1.23..1.31
21<'>	1	**1.26**	0.18	1.24..1.30
22	3	**0.89**	0.08	0.88..0.91
23	2	**2.14**	0.21	2.03..2.23
24	2	**1.55**	0.23	1.55
25	2	**1.32**	0.20	1.32
26	3	**0.88**	0.12	0.85..0.90

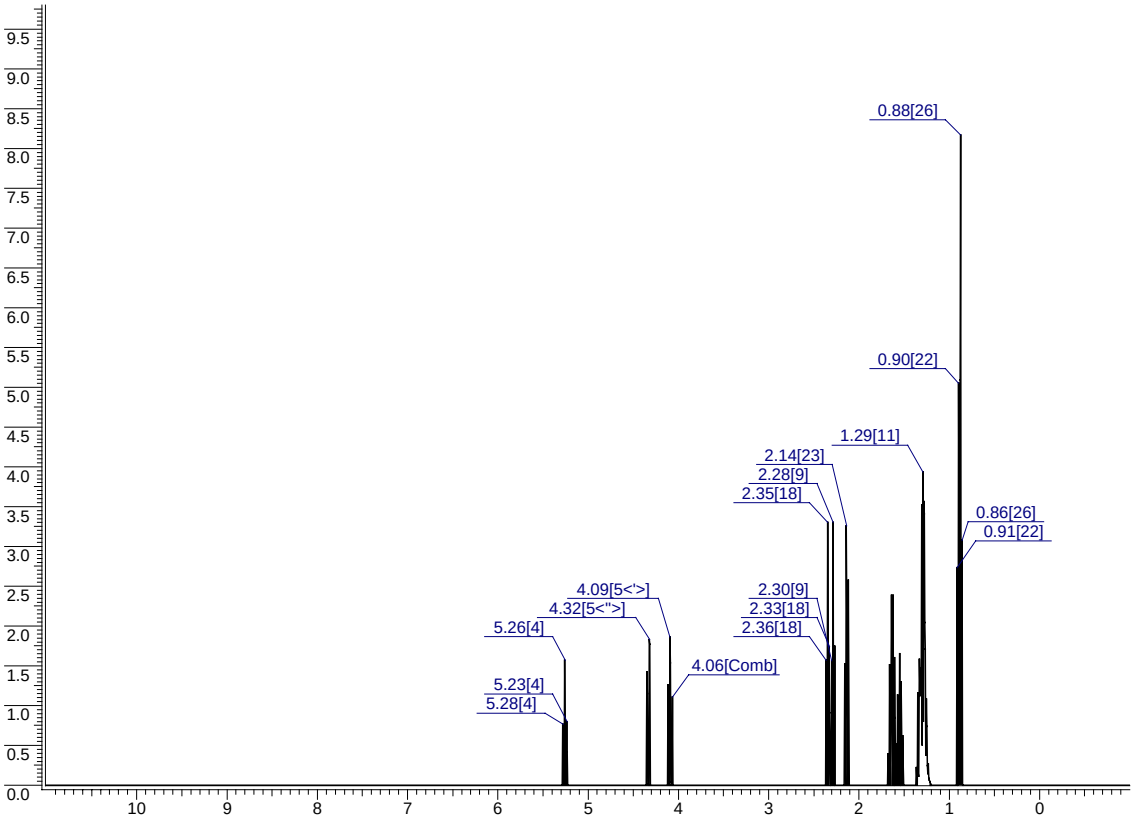
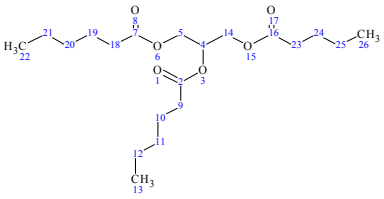


Figure 5.11. ACD calculation of proton spectrum for the triglyceride.



21/Dec/2005 11:23:58 ACD/CNMR (v.9.02)

Carbon No.	CHn	Chem. Shifts	Conf. Limits	Av.Exp
2	C	**174.73**	1.83	172.77..174.96
4	CH	**70.04**	2.78	69.09..69.11
5	CH_2	**62.03**	4.76	62.09..62.11
7	C	**175.55**	1.46	172.44..174.71
9	CH_2	**33.84**	1.15	34.13..35.00
10	CH_2	**24.77**	1.50	24.43..24.77
11	CH_2	**31.40**	1.59	31.20..31.43
12	CH_2	**22.35**	1.20	22.22..22.41
13	CH_3	**14.00**	0.74	13.82..13.98
14	CH_2	**62.03**	4.76	62.09..62.11
16	C	**173.30**	1.46	172.91..175.13
18	CH_2	**33.66**	1.21	33.93..34.02
19	CH_2	**24.48**	1.50	24.51
20	CH_2	**31.19**	1.59	31.15..31.22
21	CH_2	**22.29**	1.20	22.26..22.30
22	CH_3	**13.89**	0.74	13.85..13.90
23	CH_2	**32.88**	1.15	32.88..32.92
24	CH_2	**26.43**	4.74	26.43
25	CH_2	**21.58**	0.96	21.58
26	CH_3	**13.53**	1.15	13.51..13.55

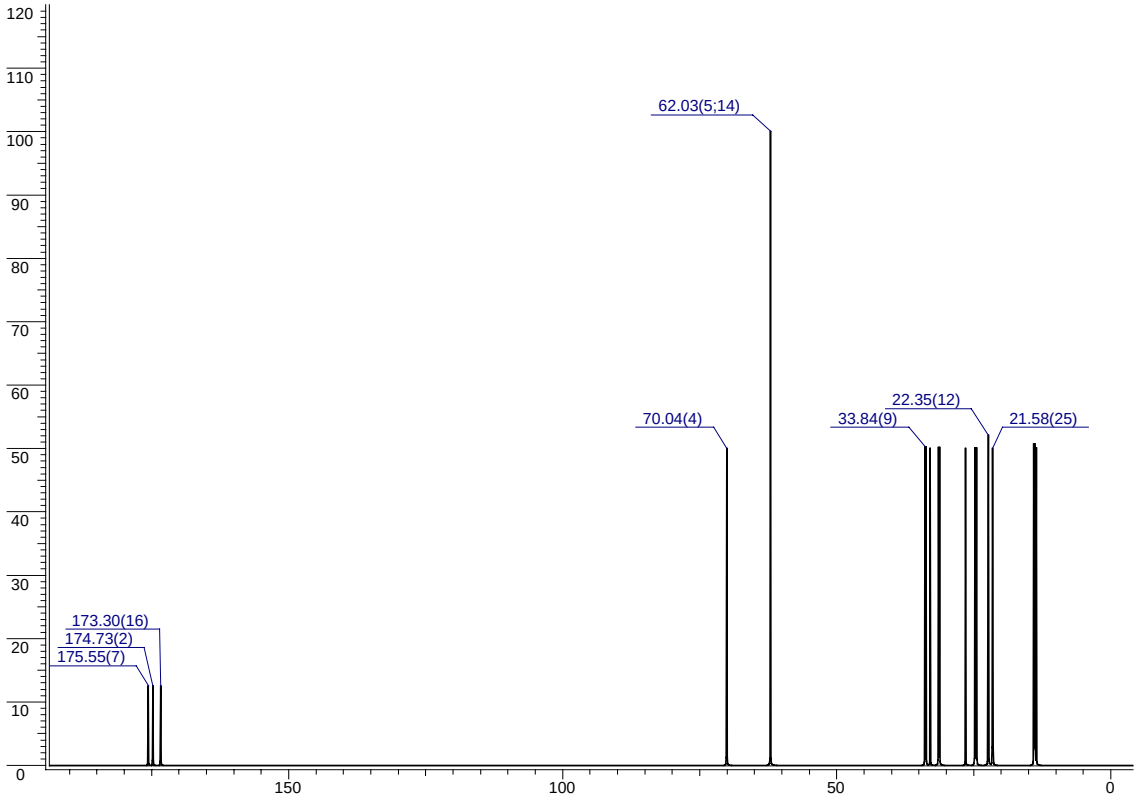
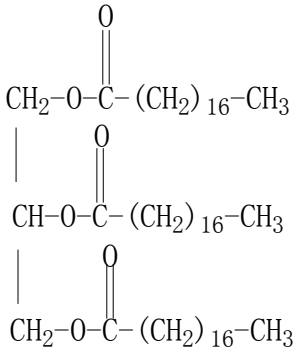


Figure 5.12. ACD calculation of carbon spectrum for the triglyceride.

5.2.1.2 Standard triglyceride: Tristearin (1,2,3-Trioctadecanoylglycerol, or Glyceryl tristearate)



1D and 2D NMR (Figures 5.15, 5.16, 5.17, and 5.18) were obtained for authentic tristearin. The spectra showed the same chemical and correlation patterns as aliphatic compound **2**, supporting the tentative structural identification.

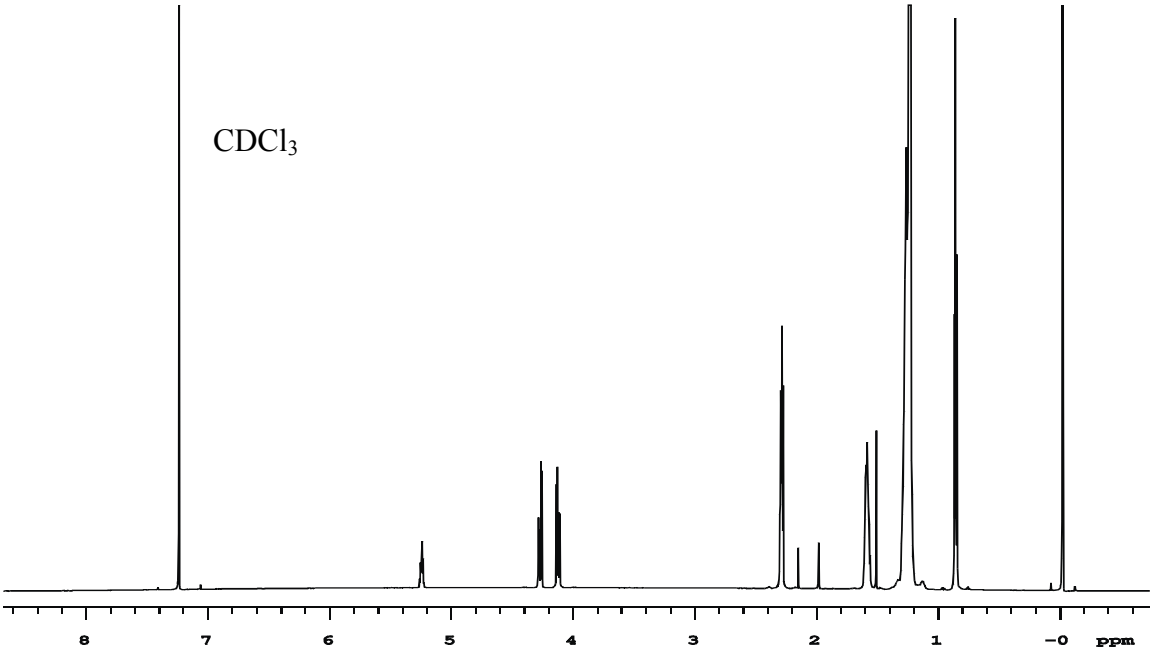


Figure 5.13. 600 MHz ^{1}H NMR spectrum of a standard triglyceride, tristearin.

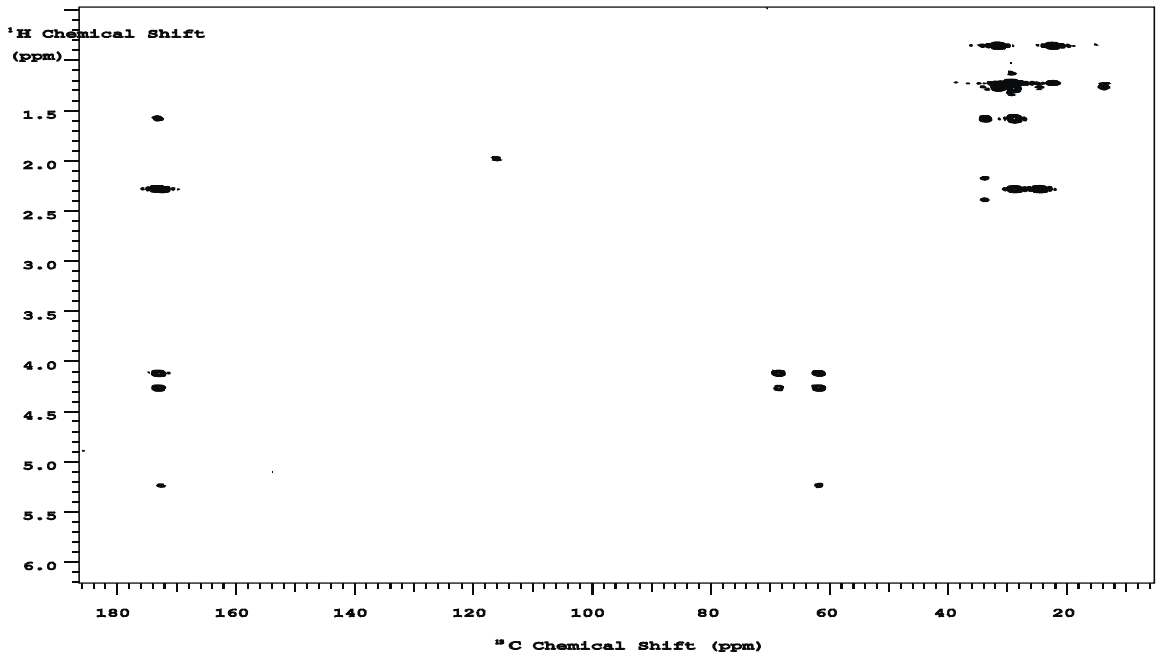


Figure 5.14. gHMBC spectrum of a standard triglyceride, tristearin.

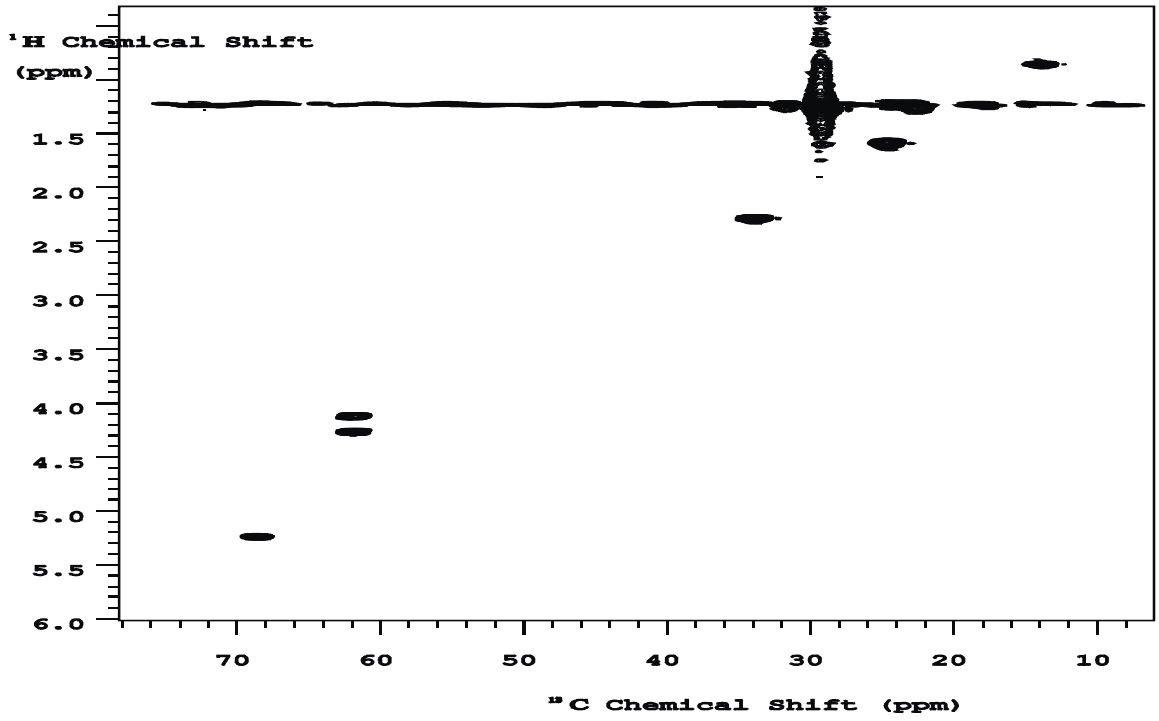


Figure 5.15. gHSQC spectrum of a standard triglyceride, tristearin.

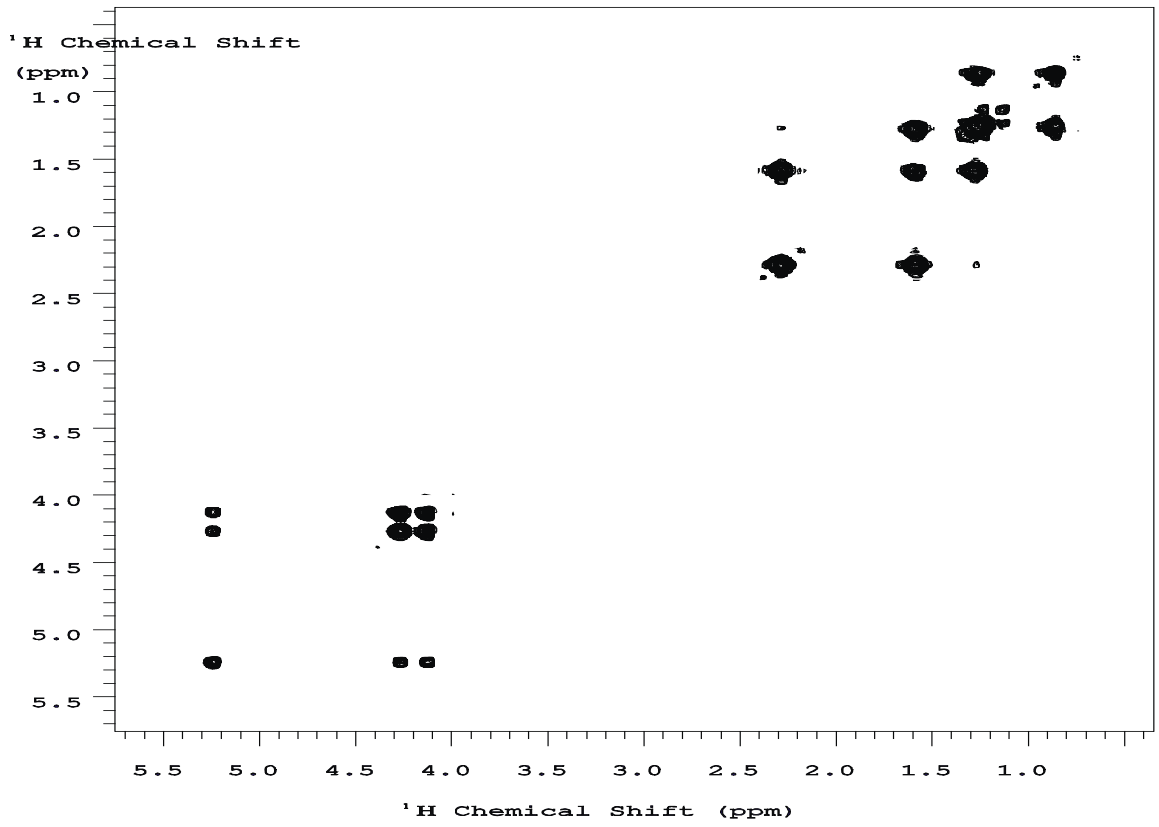


Figure 5.16. COSY spectrum of a standard triglyceride, tristearin.

5.2.2 LC/MS and LC/MS^n of aliphatic compound **2**

Section **4.5.3** above describes the isolation of triglycerides after solvent extraction and HPLC separation. Four compounds were obtained, as follows: MeOH-insoluble but hexane-soluble triglyceride (sample 5-24-#7); MeOH-soluble triglycerides (samples 5-24-#5-HPLC-f3, 5-24-#5-HPLC-f4, and 5-24-#5-HPLC-f5).

^{1}H NMR of the above four compounds (Figure 4.40, 4.42, 4.43, 4.44) are all consistent with triglycerides. Although some of them are relatively pure (e.g. sample 5-24-#5-HPLC-f4), others contain additional compounds such as ester oligomers or monomers. Since they have different solubilities and HPLC retention times, their structures must differ in some way. In order to finally determine their structures, MS or even MS^n are needed. In ordinary LC/MS, it may not be possible to determine with

certainty the origin of specific ion fragments when they may derive from more than one parent ion. Application of MS^n can resolve some of these ambiguities, though without standard compounds it cannot provide information such as the position of the fatty acids, nor is it capable of differentiating in all instances between alkylacyl and alkenylacyl species without prior separation of these subclasses.

5.2.2.1 LC/MS identification of triglycerides

Traditionally, triglycerides have been analyzed by hydrolyzing them to yield the fatty acids, which are then derivatized and analyzed as methyl esters by GC/MS.[72] This approach offers only indirect analysis of the original triglyceride, and preparation of the derivative is time consuming. Intact triglycerides can be chromatographed by HPLC, but their lack of chromophore (unless conjugated) makes detection difficult.[73] To identify each triglyceride, a large number of standards and lengthy analysis time are required. LC/MS analysis overcomes these problems, offering LC with direct mass detection but without the requirement of a chromophore, and high sensitivity, provided that the compound can be ionized.

Due to the nonpolar nature of the molecules, atmospheric pressure chemical ionization (APCI) was employed. When the most commonly used solvent system for APCI (methanol/water with acid) was used, no molecular ion was obtained because the protonated $[M+1]^+$ species is evidently not stable. Another solvent system [73] was used to ionize the triglyceride samples: 50% A + 50% B

Eluent A: water: isopropanol (60:40) + 25 mM ammonium formate

Eluent B: water: isopropanol: *n*-butanol (10:10:80) + 25 mM ammonium formate

The addition of ammonium formate to the mobile phase allowed the formation of an ammonium adduct $[M+18]^+$, which was more stable than the protonated $[M+1]^+$ species.

In nature triglycerides are made up predominantly of fatty acids with even numbers of carbons (C8, C10, C12… C18, etc.) and either 1 or 2 double bonds per acid, but in the above triglyceride, ^{1}H NMR showed no double bond signals. A generic formula was used to determine the possible combinations of double bonds and carbon chain lengths:

$$\mathbf{A}\text{ -18 } (NH_4^+) = 218.03 + 28.03 * m + 26.02 * n \ ,$$

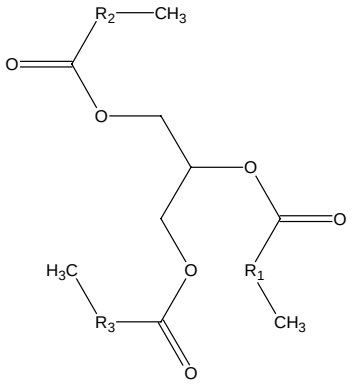


Figure 5.17. Backbone structure of a triglyceride.

where **A** is the mass of the observed adduct ion, which is M + 18 (NH_4^+). The 218.03 accounts for the backbone of the triglyceride without adding the mass of R_1, R_2, and R_3 in Figure 5.17, m is the number of ethylene ($-CH_2-CH_2-$) groups and n is the number of ethenyl (-CH=CH-) groups. Only integer solutions are reasonable, and in most cases fatty acids of the same or similar length and unsaturation are most likely.

Under the above analytical conditions, MS data were obtained for the four different triglycerides. For samples that were relatively pure, the MS spectrum was clean; for other triglyceride samples that contained some aliphatic oligomers and monomers, the corresponding MS spectra were more complex.

5.2.2.1.1 LC/MS of sample 5-24-#5-HPLC-f3

Three $[M+18]^+$ ionization peaks appeared in the MS spectrum (Figure 5.18) at m/z 419.3**,** 488.4, and 544.5, respectively. The molecular ion at 419.3 did not match the triglyceride formula, so MS^n may be required to deduce its structure.

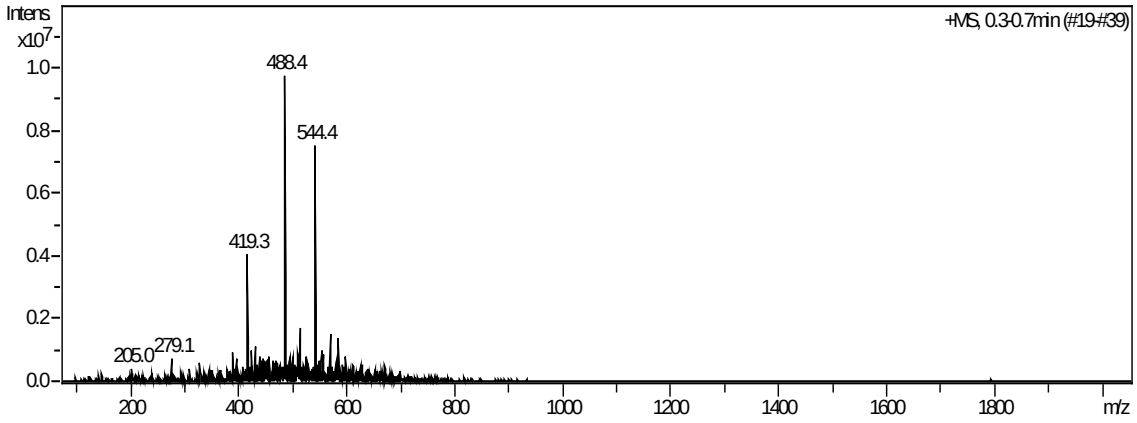


Figure 5.18. APCI-MS spectrum of sample 5-24-#5-HPLC-f3.

Table 5.4 Structure calculation for the major peaks in sample 5-24-#5-HPLC-f3

Observed Ion (m/z), $[M+18]^+$	Calc. MW	Backbone Removed	m	n	Possible Side Chains
488.4	470.4	252.37	9	0	**C8, C8, C8*** C6, C8, C10 …….
544.4	526.4	308.37	11	0	**C8, C10, C10*** C8, C8, C12 …….
419.3					Not triglyceride

***Bold represents the most likely combination.**

The results for sample 5-24-#5-HPLC-f3 are summarized in Table 5.4, showing the possible combinations of double bonds and carbon chain lengths. If the side chains have the same length, there is no problem in determining the structure; if the side chains have different chain lengths, MS cannot determine their respective positions in the triglyceride structure.

5.2.2.1.2 LC/MS of sample 5-24-#5-HPLC-f4

According to the MS spectrum (Figure 5.19), one $[M+18]^+$ ionization peak was obtained at m/z 516.4.

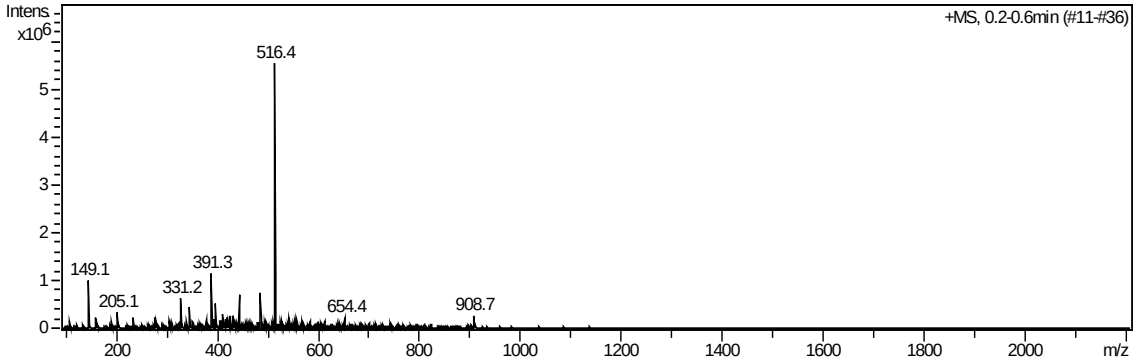


Figure 5.19. APCI-MS spectrum of sample 5-24-#5-HPLC-f4.

Table 5.5 Structure calculation for the major peaks in sample 5-24-#5-HPLC-f4

Observed Ion (m/z), $[M+18]^+$	Calc. MW	Backbone Removed	m	n	Possible Side chain
516.4	498.4	280.37	10	0	**C8, C8, C10*** C6, C8, C12 ……

***Bold represents the most likely combination for the corresponding molecular ion.**

As before, Table 5.5 used the generic formula to determine the possible combinations of double bonds and carbon chain lengths.

5.2.2.1.3 LC/MS of sample 5-24-#5-HPLC-f5

The MS spectrum (Figure 5.20) shows that multiple $[M+18]^+$ ionization peaks are obtained at m/z 488.4 (not displayed), 656.7, 684.7, 740.7, 768.8, 796.8 and 784.8. Thus as summarized in Table 5.6, sample 5-24-#5-HPLC-f5 is a mixture of many similar compounds, only some of which are triglycerides. The reason for poor signal-to-noise ratio is that sample 5-24-#5-HPLC-f5 is a mixture of many triglycerides together with some other compounds. For the molecular ion $[M+18]^+$ at 784.8, the fragmentation

pattern is different from a triglyceride, suggesting that it might be a long-chain ester oligomer.

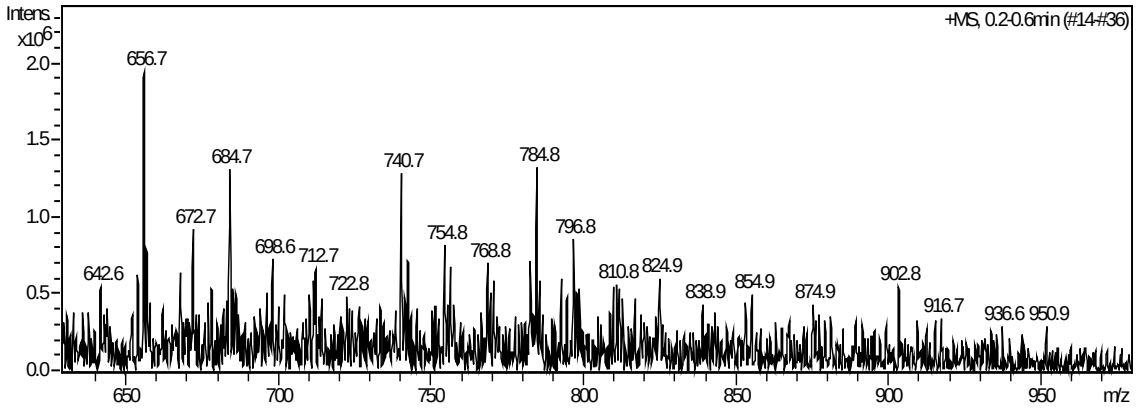


Figure 5.20. APCI-MS spectrum of sample 5-24-#5-HPLC-f5.

Table 5.6 Structure calculation for the major peaks in sample 5-24-#5-HPLC-f5

Observed Ion (m/z), $[M+18]^+$	Calc. MW	Backbone Removed	m	n	Possible Side chains
488.4					Same as 5-24-#5-HPLC-f3
656.7	638.7	420.67	15	0	**C12, C12, C12*** **C10, C12, C14*** **C10, C10, C16*** …….
684.7	666.7	448.67	16	0	**C12, C12, C14*** **C10, C12, C16*** …….
740.7	722.7	504.67	18	0	C14, C14, C14 **C10, C16, C16*** …….
768.8	750.8	532.77	19	0	**C14, C14, C16*** **C12, C16, C16*** …….
796.8	778.8	560.77	20	0	**C14, C16, C16*** C12, C16, C18 …….
784.8					Ester oligomer**

***Bold represents the most likely combination for the corresponding molecular ion.**
****NMR suggests ester oligomer but fragments cannot be identified with confidence.**

5.2.2.1.4 LC/MS of sample 5-24-#7

The MS spectrum (Fig. 5.21) shows multiple $[M+18]^+$ ionization peaks at m/z 488.5, 600.5, 656.5, 684.8, 740.8, 796.8, 824.8, 852.6, and 908.4. The structural possibilities are summarized in Table 5.7. Again, the poor signal-to-noise ratio is attributed to the fact that sample 5-24-#5-HPLC-f5 is a mixture of many triglycerides together with some other compounds.

Table 5.7 Structure calculation for the major peaks in sample 5-24-#7

Observed Ion (m/z), $[M+18]^+$	Calc. MW	Backbone Removed	m	n	Possible Side chains
488.4					Same as 5-24-#5-HPLC-f3
600.5	582.6	364.57	13	0	C10, C10, C12 **C8, C12, C12***
656.6					Same as 5-24-#5-HPLC-f5
684.7					Same as 5-24-#5-HPLC-f5
740.8					Same as 5-24-#5-HPLC-f5
768.7					Same as 5-24-#5-HPLC-f5
782.8	764.8	564.77	19.5	0	C14, C14, C17 **C14, C15, C16*** **C12, C16, C19***
796.8					Same as 5-24-#5-HPLC-f5
810.8	792.8	574.77	20.5	0	**C15, C16, C16*** C12, C16, C18
824.8	806.8	588.77	21	0	**C16, C16, C16*** C14, C16, C18
852.6	834.6	616.57	22	0	**C16, C16, C18*** C14, C16, C20
908.7	890.7	672.67	24	0	**C18, C18, C18*** C16, C18, C20……

*** Bold represents the most likely combination for the corresponding molecular ion.**

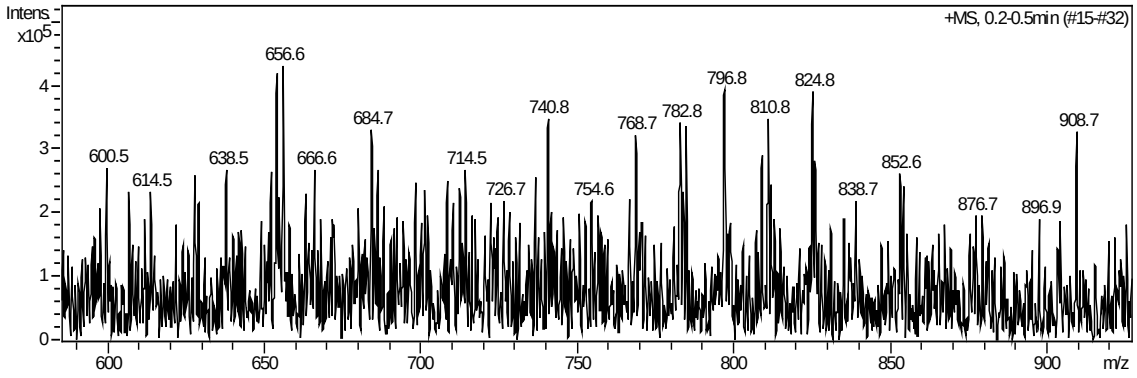


Figure 5.21. APCI-MS spectrum of sample 5-24-#7.

5.2.2.2 LC/MS^n identification of triglycerides

LC/MS can only provide molecular weight information. To obtain structural information, MS^n was employed to yield fragmentation information by adjusting some parameters in the MS experiment.

5.2.2.2.1 LC/MS^n of a standard triglyceride, tristearin

Molecular Formula: $[CH_3(CH_2)_{16}COOCH_2]_2CHOCO(CH_2)_{16}CH_3$

Molecular Weight: 891.48

Each member of a homologous series should have a similar MS^n fragmentation pattern, so experiments on this standard compound were designed to establish the characteristic fragmentation pattern of triglycerides. Figure 5.22 shows the fragmentation of tristearin using the APCI-LC/MSD ion trap. The MS spectrum shows the ammonium adduct ion $[M+18]^+$ at m/z 909.0; fragmentation of this ion (MS^2) yields a fragment ion at m/z 607.6, with a difference of 302.4 attributable to loss of $-O-\overset{\overset{O}{\|}}{C}-(CH_2)_{16}-CH_3$ NH_4^+ (C18 fatty

acid), from any of the three triglyceride positions (1,2,3). In MS^3, further fragmentation of m/z 607.6 yields an ion at 267.2, corresponding to $-C(=O)^+-(CH_2)_{16}-CH_3$ (C18 fatty acid).

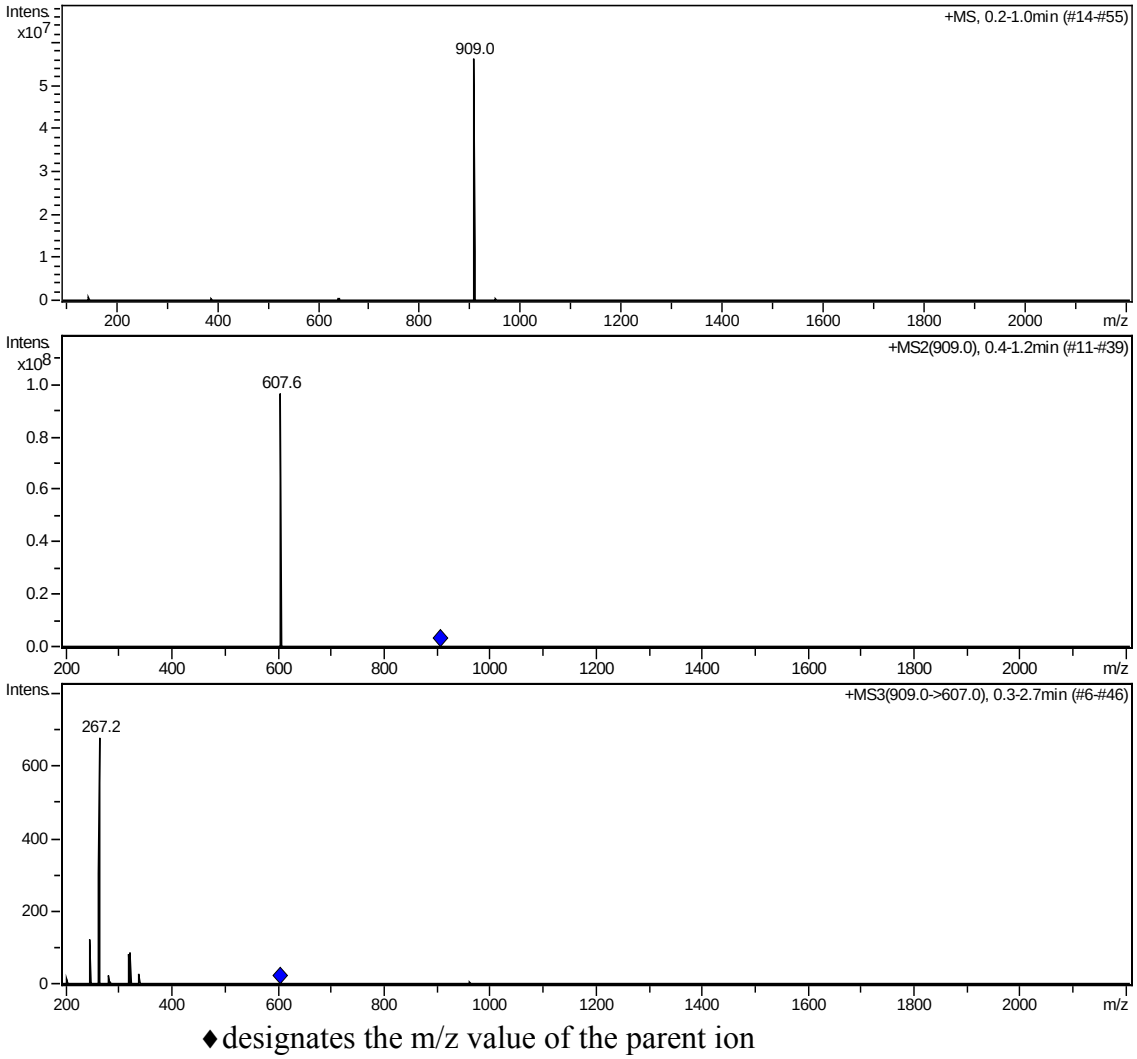


Figure 5.22. APCI-LC/MSD ion trap MS^n spectrum of tristearin (MS->MS^2->MS^3).

The deduced fragmentation pathway of tristearin is shown in Figure 5.23. As noted above, MS^n cannot reveal the position of the fatty acids, e.g. in Figure 5.23, the first loss of 302.3 could occur at position 2 of the triglyceride. The scheme shown here assumes

that for MS^2, the first ion is lost from position 1 and for MS^3, the second ion is removed from position 3 and the retained ion is at position 2.

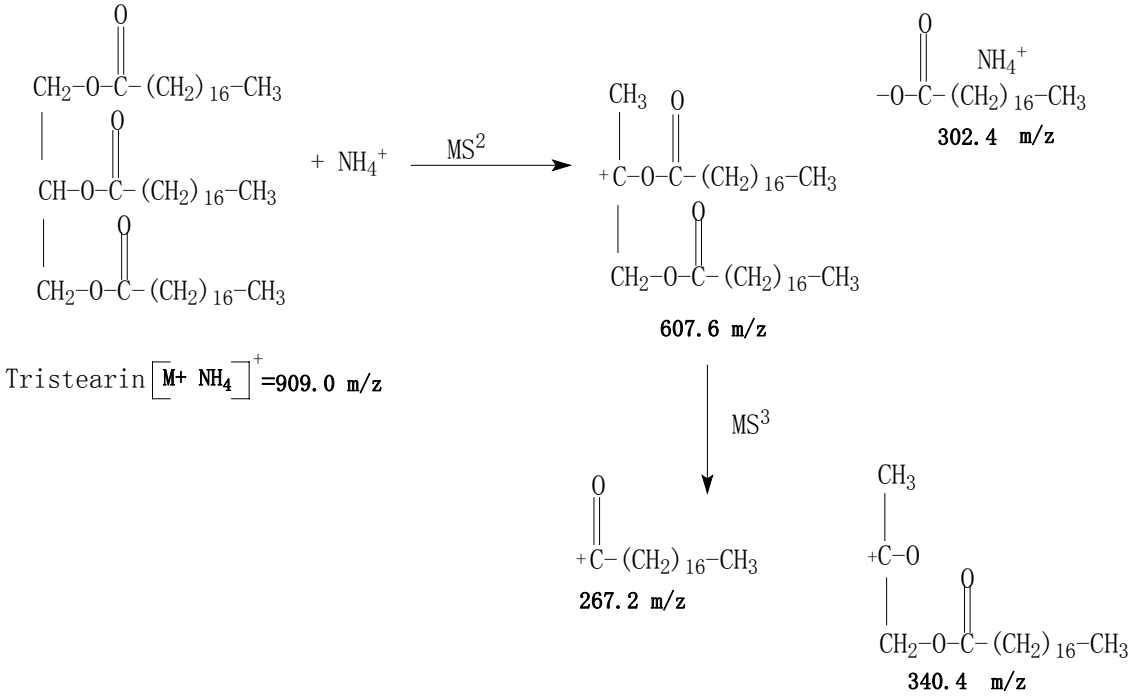


Figure 5.23. Fragmentation pathways for tristearin using the LC/MSD ion trap.

5.2.2.2.2 LC/MS^n of triglyceride compounds derived from potato suberin

For all four triglycerides obtained from suberin (sample 5-24-#7, sample 5-24-#5-HPLC-f3, sample 5-24-#5-HPLC-f4, sample 5-24-#5-HPLC-f5), $MS \rightarrow MS^2 \rightarrow MS^3$ was performed using the APCI-LC/MSD ion trap under the same experimental conditions as the standard triglyceride, tristearin.

5.2.2.2.2.1 LC/MS^n of sample 5-24-#5-HPLC-f4

Sample 5-24-#5-HPLC-f4 contains the purest triglyceride separated from suberin: its LC/MS spectrum (Figure 5.19) makes it clear that the MS spectrum has only one major $[M+18]^+$ peak at m/z 516.4. Table 5.5 shows that the most probable combination of carbon chain lengths is C8, C8, C10 (including the carbonyl carbon and end methyl

carbon). In this case, because the three chain lengths are different, there are two possible structures. Of course, other combinations like C6, C8, C12 (3 possible structures) are also theoretically possible. This kind of ambiguity cannot be resolved by LC/MS, whereas MS^n with the help of standard fragmentation patterns can be informative.

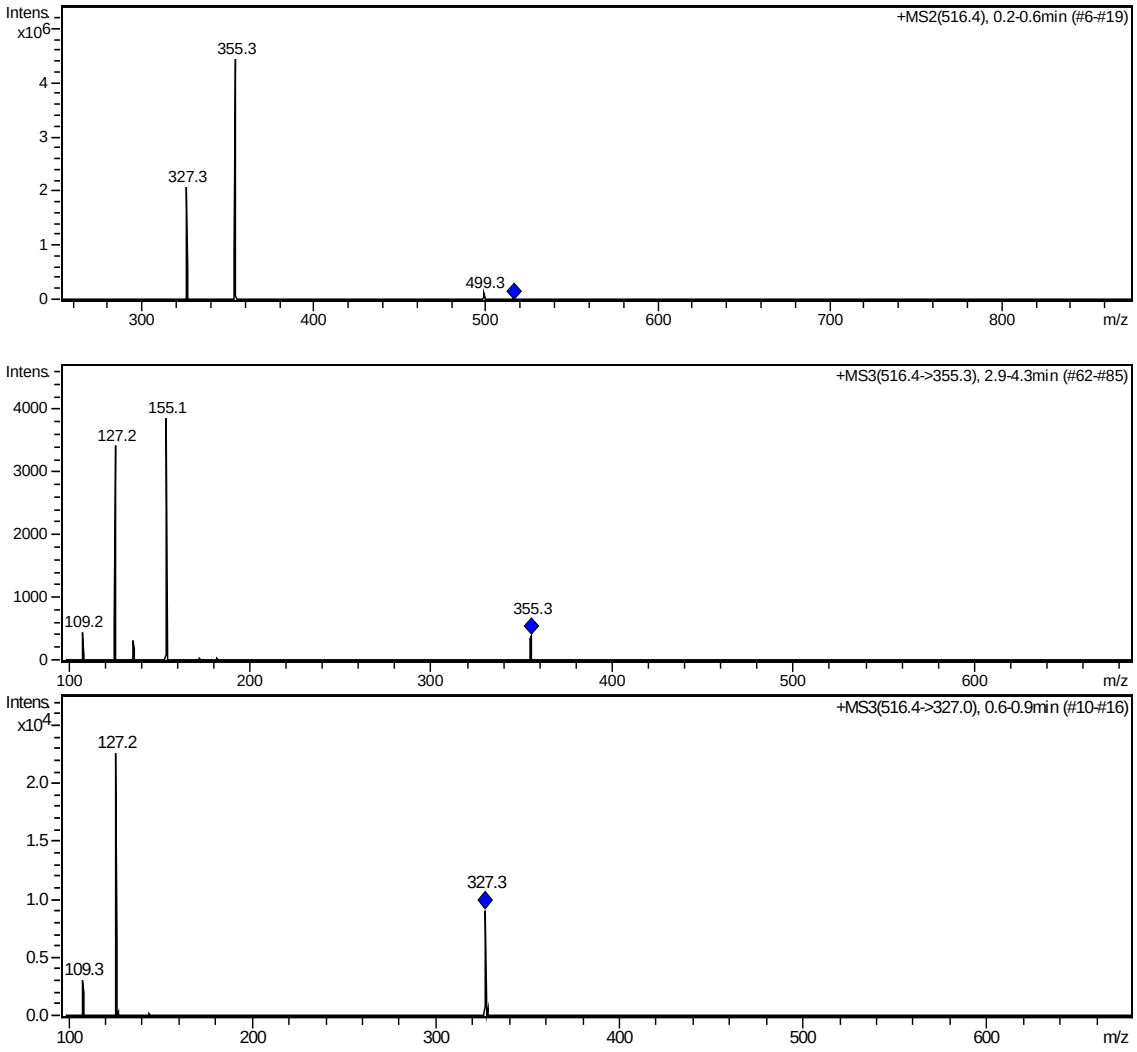


Figure 5.24. APCI-LC/MSD ion trap MS^n spectrum of 5-24-#5-HPLC-f4 (MS 516.4->MS^2->MS^3).

Figure 5.24 shows the MS^n spectrum of 5-24-#5-HPLC-f4: $[M+18]^+$ at m/z 516.4, its MS^2 yielding two fragmentation ions at m/z 327.3 and 355.3 (Figure 5.24, top). Further fragmentation (Figure 5.24, middle, MS^3) of m/z 355.3 gives two fragment ions at m/z

127.2 and 155.1, respectively, whereas further fragmentation (Figure 5.24, bottom, MS^3) of m/z 327.3 provides one fragment ion at 127.2. These MS^n data may be explained with three schemes as follows:

A. 516.4 m/z $\xrightarrow[\text{lose 161.1}]{MS^2}$ 355.3 m/z $\xrightarrow[\text{lose 200.2}]{MS^3}$ 155.1 m/z

The MS^2 difference of m/z 161.1 represents the loss of a C8 fatty acid first; MS^3 yields m/z 155.1, corresponding to the remaining C10 fatty acid. According to this scheme and following the pattern for tristearin, molecule **A** should contain C8 (position 1), C10 (position 2), and C8 (position 3). Its fragmentation pathway is shown in Figure 5.25.

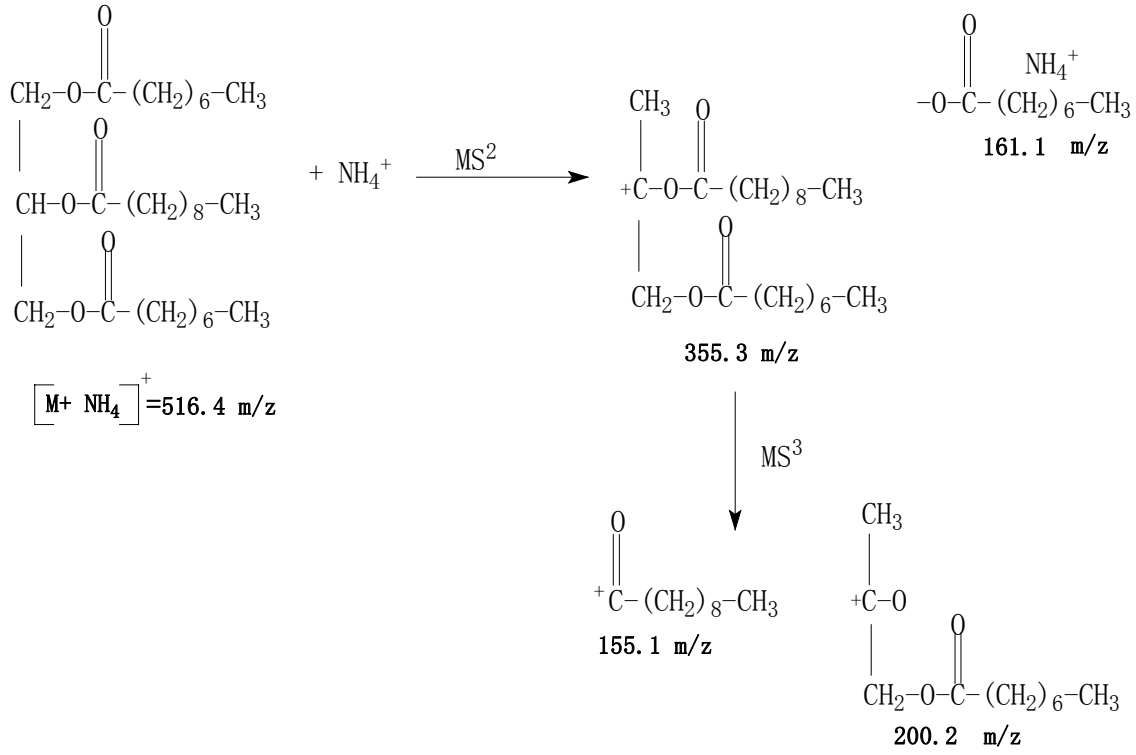


Figure 5.25. Fragmentation pathways for 5-24-#5-HPLC-f4 using the LC/MSD ion trap.

B. 516.4 m/z $\xrightarrow[\text{lose 161.1}]{MS^2}$ 355.3 m/z $\xrightarrow[\text{lose 228.1}]{MS^3}$ 127.2 m/z

The MS^2 difference of m/z 161.1 represents the loss of a C8 fatty acid first; MS^3 yields m/z 127.2, corresponding to a C8 fatty acid. Molecule **B** should then contain C8, C8, C10. Its fragmentation pathway is shown in Figure 5.26.

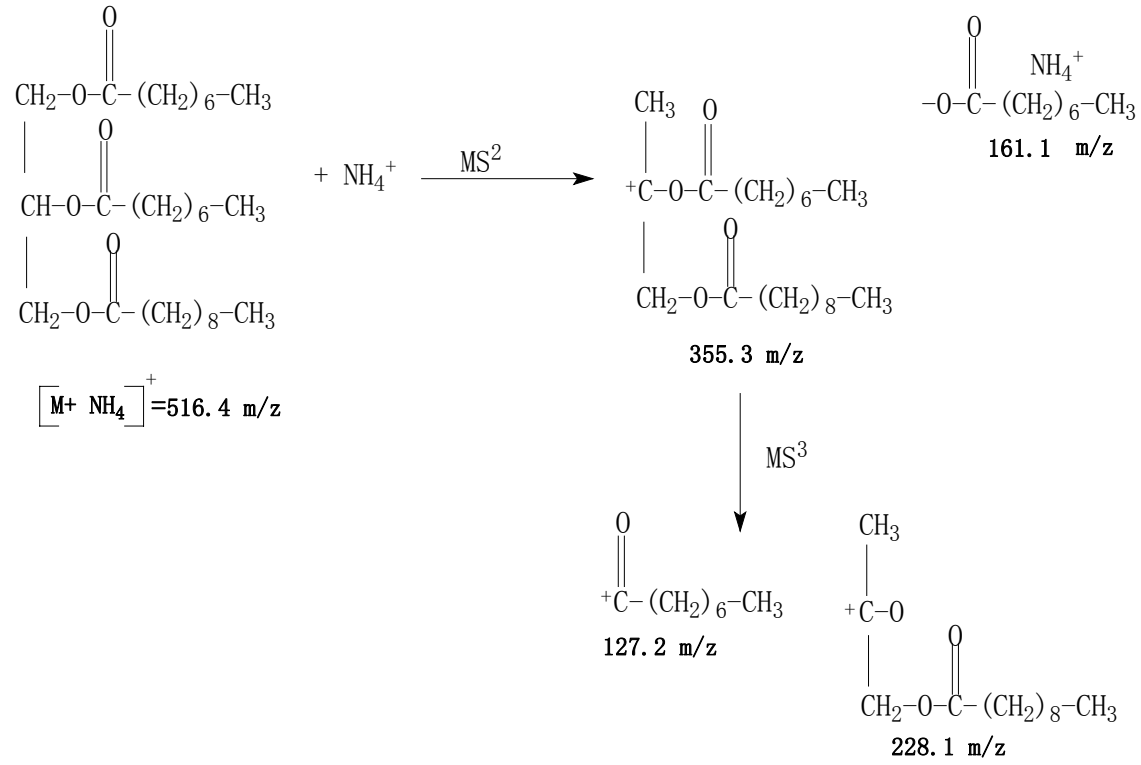
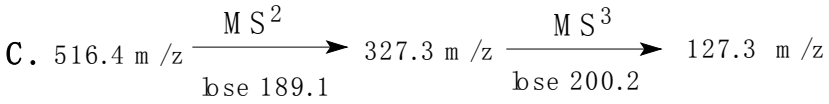


Figure 5.26. Structure and fragmentation pathways for 5-24-#5-HPLC-f4 using the LC/MSD ion trap.



The MS^2 difference of m/z 189.1 represents the loss of a C10 fatty acid first; MS^3 yields m/z 127.3, corresponding to a C8 fatty acid. The molecule **C** should then include C10, C8, C8. Its fragmentation pathway is shown in Figure 5.27.

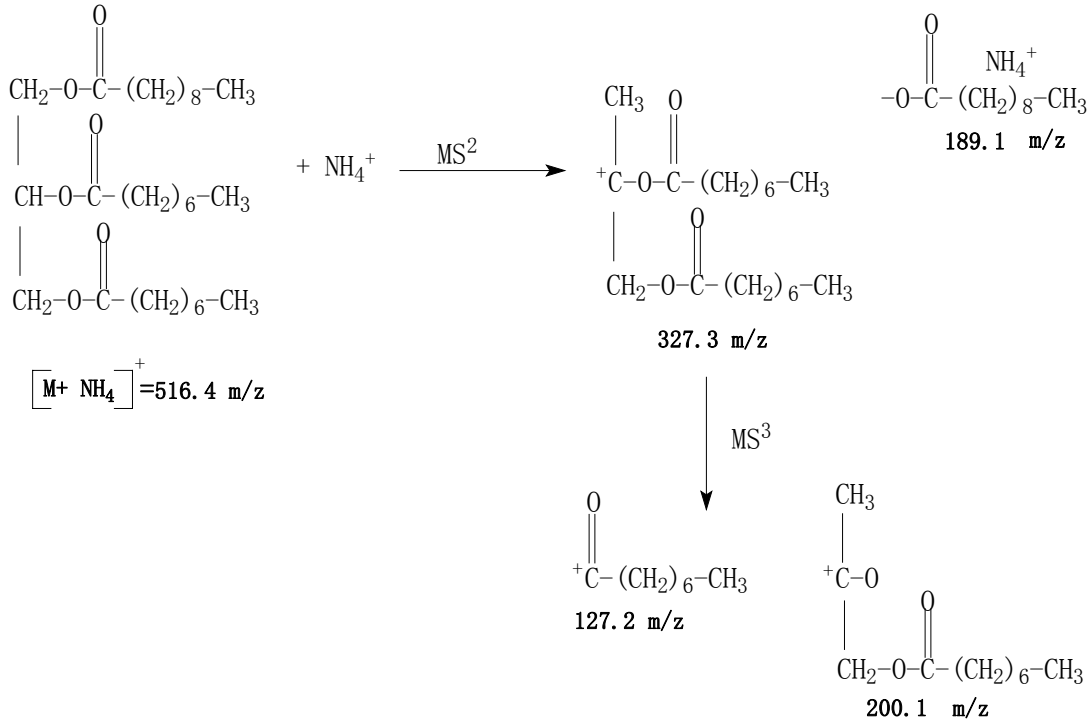
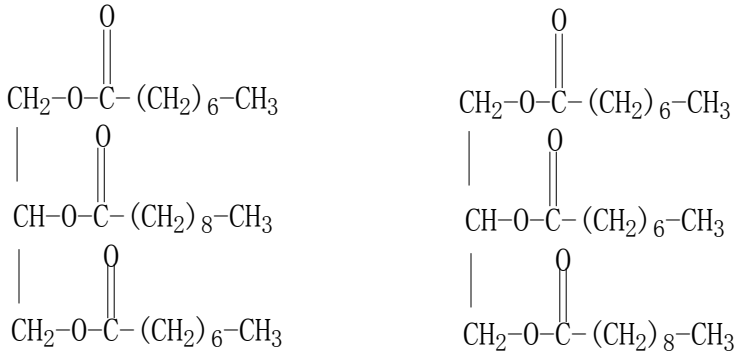


Figure 5.27. Fragmentation pathways for 5-24-#5-HPLC-f4 using the LC/MSD ion trap.

Thus, **B** and **C** are the same molecule; MS^2 simply gives fragmentation at different positions. MS^n data confirm that two structures (below) are consistent with the $[M+18]^+$ peak at m/z 516.4.



5.2.2.2.2.2 LC/MS^n of sample 5-24-#5-HPLC-f3

For $[M+18]^+$ of m/z 488.4, Table 5.4, the LC/MS spectrum, and calculations show that the most probable combination of carbon chain lengths is C8, C8, C8 in the triglyceride molecule. Of course, other combinations are theoretically possible, including C6, C8, C10; C4, C10, C10, etc.

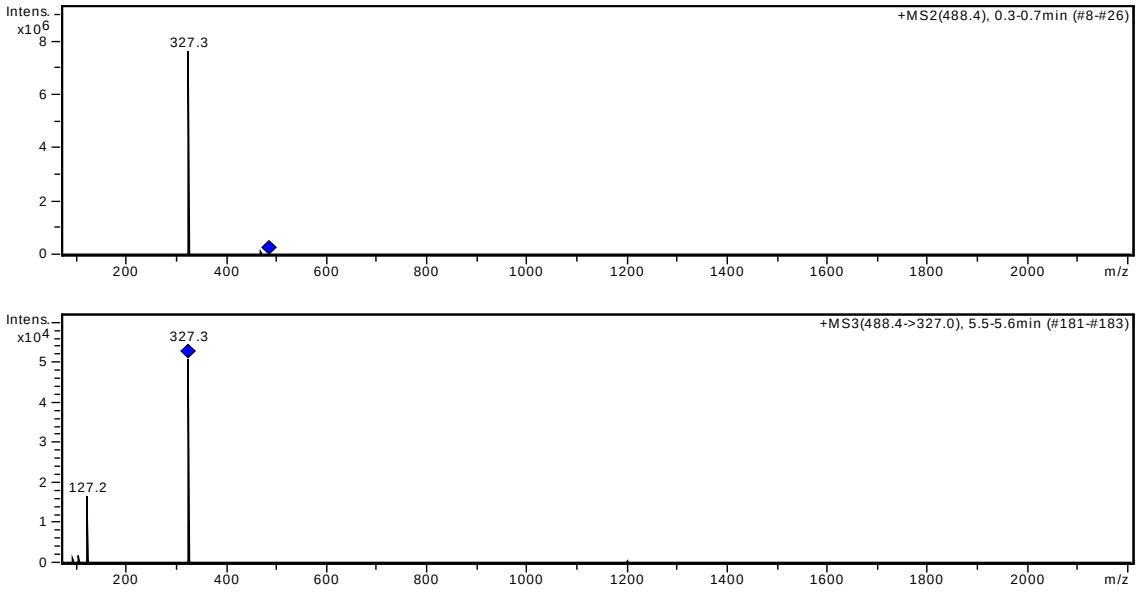


Figure 5.28. APCI-LC/MSD ion trap MS^n spectrum of 5-24-#5-HPLC-f3 (MS488.4->MS^2->MS^3).

Figure 5.28 shows the MS^n spectrum of 5-24-#5-HPLC-f3 at $[M+18]^+$ m/z 488.4, with MS^2 of this ion yielding a fragment ion at m/z 327.3 and a difference of m/z 161.1. It represents loss of a C8 fatty acid, but MS^n cannot tell which position was cleaved. Further fragmentation from MS^3 of m/z 327.2 produces a fragment ion at m/z 127.2, corresponding to removal of another C8 fatty acid. This means the third branch must also be C8. Thus the only possible structure of this molecule ($[M+18]^+$ m/z 488.4) is C8, C8, C8, shown in **bold** in Table 5.4. The MS^n results rule out other possible combinations such as C6, C8, C10, or C4, C10, C10. Figure 5.29 shows the structure and fragmentation pathways for this molecule using the LC/MSD ion trap. It should be emphasized that

Figure 5.25 is just one possible fragmentation pathway, as the first loss of m/z 161 could also occur from position 2 of the triglyceride.

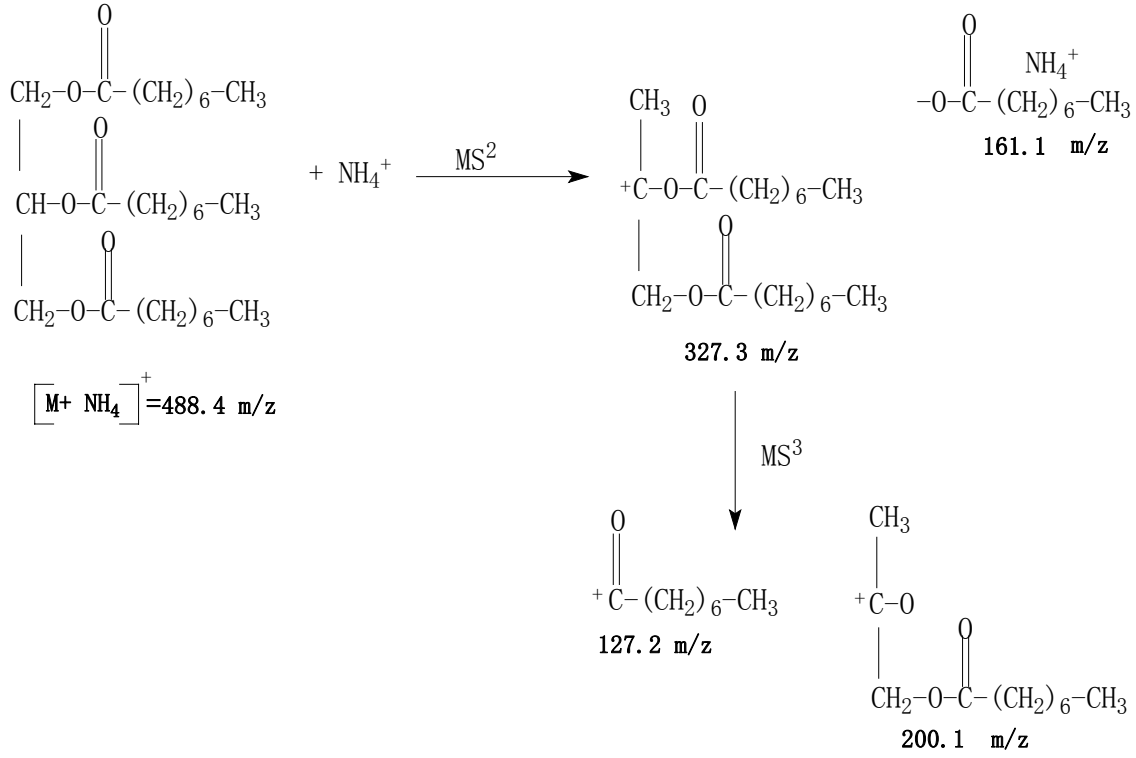
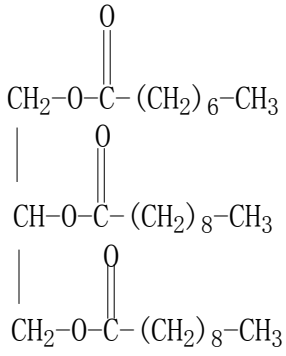
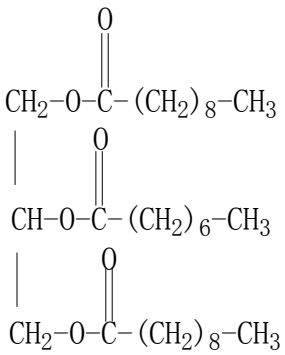


Figure 5.29. Structure and fragmentation pathways for 5-24-#5-HPLC-f3 (MS488.4->MS^2->MS^3) using the LC/MSD ion trap.

For $[M+18]^+$ 544.4, Table 5.4, the LC/MS spectrum and calculations implicate the most probable combination of carbon chain lengths as C8, C10, C10 in the triglyceride molecule. As before, other combinations are theoretically possible: C6, C8, C10, or C4, C10, C10, etc. MS^n (MS^2 and MS^3, data not shown) of the molecular ion $[M+18]^+$ with m/z 544.4 rules out all possible combinations except one C8 and two C10 units (**bold** in Table 5.4). This molecule is expected to follow the same fragmentation pattern as the triglycerides described above. Like sample 5-24-#5-HPLC-f4 at $[M+18]^+$ 516.4 (Section

5.2.2.2.2.1), MS^2 and MS^3 confirm that two isomers (C8, C10, C10 and C10, C8, C10) co-exist (see structures below).





5.2.2.2.2.3 LC/MS^n of sample 5-24-#5-HPLC-f5

Sample 5-24-#5-HPLC-f5 is a mixture of triglyceride and some other aliphatic oligomers; its MS spectrum (Figure 5.20) shows that it contains numerous compounds. Only the peaks fitting a triglyceride formula were chosen for MS^2 and MS^3 experiments: $[M+18]^+$ with m/z 488.4, 656.7, 684.7, 740.7, 768.8, 796.8. (data not shown). Each peak shows the same fragmentation pattern, confirming the combinations of three side chains shown in **bold** in Table 5.6. The hydrophobic side chain length is longer than that of Samples 5-24-#5-HPLC-f3 and 5-24-#5-HPLC-f4, as expected because 5-24-#5-HPLC-f5 elutes later than f3 and f4 in RP-HPLC.

5.2.2.2.2.4 LC/MS^n of sample 5-24-#7

Like sample 5-24-#5-HPLC-f5, sample 5-24-#7 is a mixture of triglyceride and some other aliphatic oligomers and exhibits a complex MS spectrum (Figure 5.21). As before, peaks fitting the triglyceride formula were chosen for MS^2 and MS^3: $[M+18]^+$ with m/z 488.4, 600.5, 656.6, 684.7, 740.7, 768.8, 782.8, 796.8, 810.8, 824.8, 852.6, 908.7 (data not shown). Each peak shows the same fragmentation pattern and resulting combinations

of three side chains depicted in **bold** in Table 5.7. The $[M+18]^+$ ion at m/z 908.7 also shows the same fragmentation pattern as the standard compound tristearin, thus its structure is confirmed as a C18, C18, C18 triglyceride.

5.3 Aliphatic compound 3: ester oligomers

5.3.1 NMR identification of aliphatic compound 3

Aliphatic compound **3** was separated by solvent extraction from the soluble depolymerization products of suberin (see solvent extraction strategy in Chapter 4).

Its ^{1}H NMR spectrum (Figure 5.30) shows the characteristic triplet peak at 4.05 ppm for an ester bond, with the presence of oligomers confirmed by 2D NMR in Figures 5.31 (gHMQC), and 5.32 (gHMBC). Although this compound is relatively pure, it is still a mixture of isomers. According to NMR, the end functional groups should be $–CH_3$, -OH, or $-COOCH_3$, so there are six possible combinations for end functional groups.

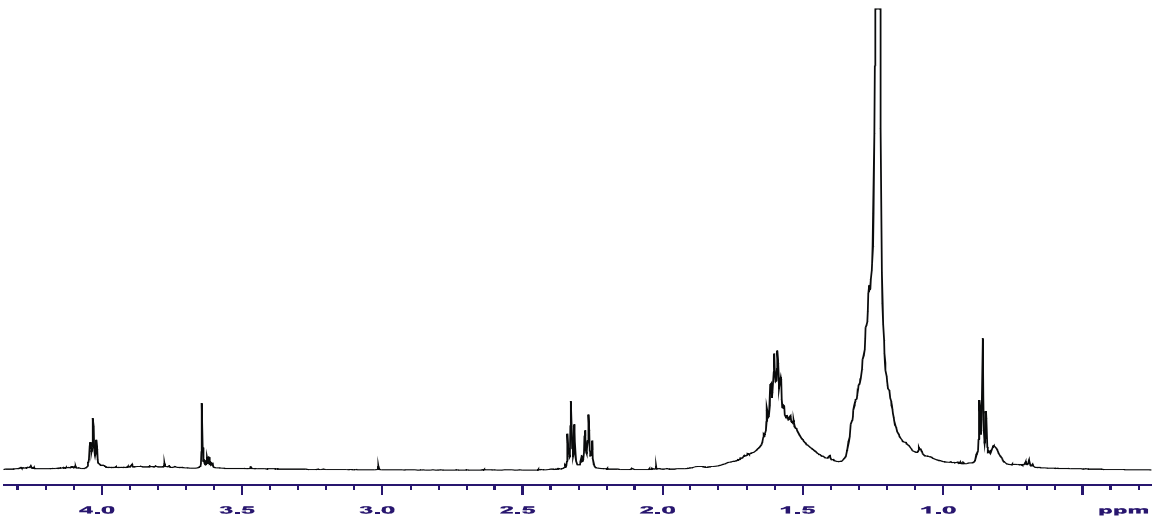


Figure 5.30. 600 MHz ^{1}H spectrum of aliphatic compound 3.

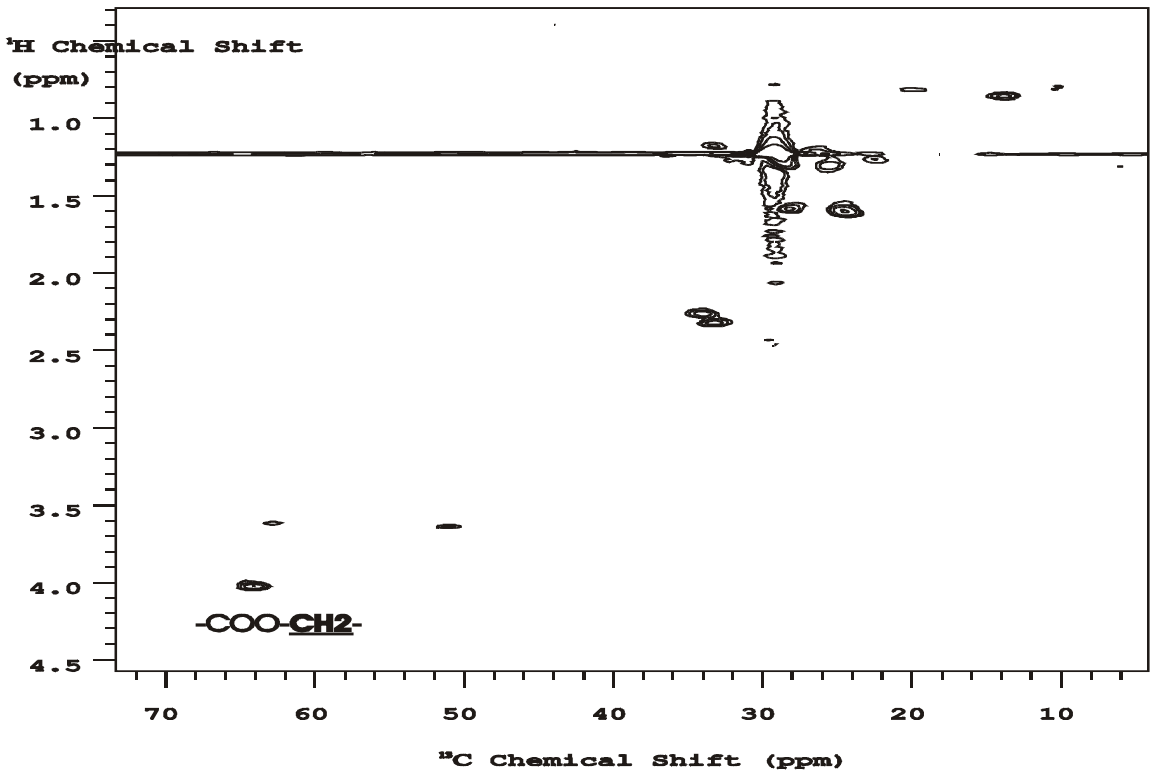


Figure 5.31. gHMQC spectrum of aliphatic compound 3.

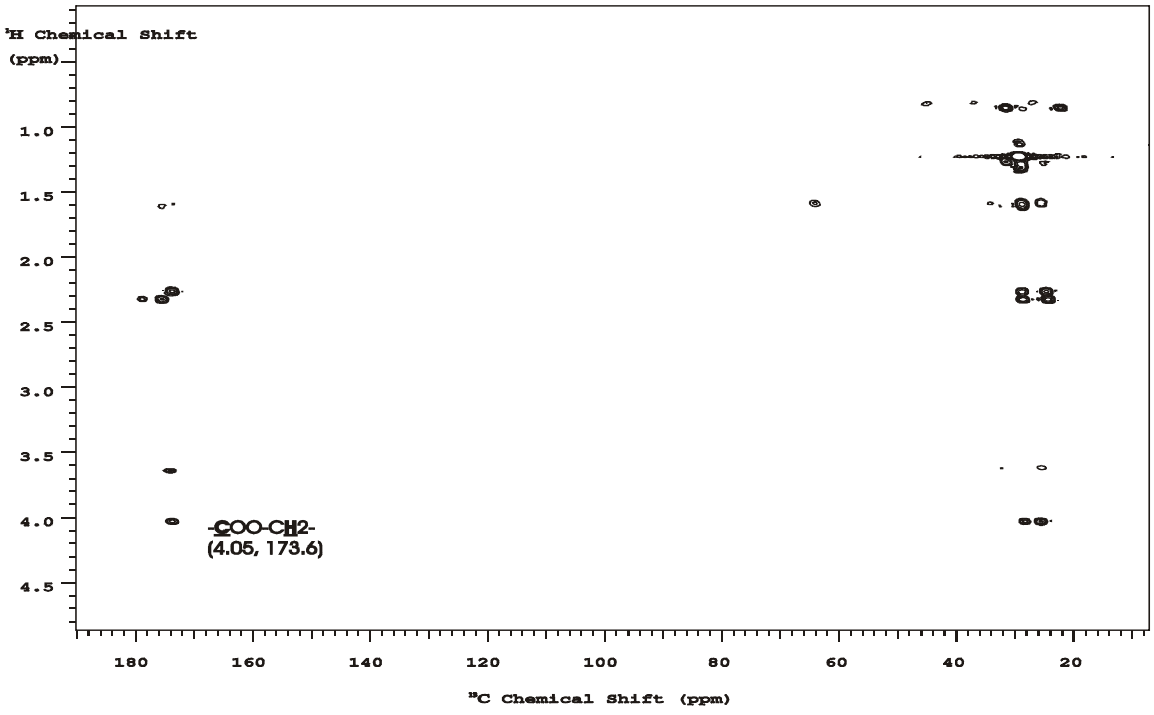


Figure 5.32. gHMBC spectrum of aliphatic compound 3.

They are either trimers or dimers (generally C_{18}, C_{18}, or C_{14}).

CH_3_____CH_3

CH_3_____OH

CH_3_____ $COOCH_3$

OH_____$COOCH_3$

OH_____OH

CH_3OOC_____$COOCH_3$

MS data ruled out all but the first two possibilities (see below).

Table 5.8 Chemical Shifts and Integration of aliphatic compound 3

Chemical shift (ppm)	4.05(t)	3.65(t)	2.27(t)	0.8(t)
Integral (Exp.)	2	0.43	2.24	4.7
Integral (Theor.) for CH_3------OH	2	1	2	1.5
Integral (Theor.) for CH_3------CH_3	2	0	2	3

The discrepancy between theoretical and experimental ^{1}H NMR integrals (Table 5.8) implies that several kinds of structures are present. ACD calculations (not shown) match the experimental chemical shifts results very well.

5.3.2 LC/MS identification of aliphatic compound **3**

Due to the non-polar nature of aliphatic compound **3**, APCI (positive, negative)-LC/MS was conducted (ESI did not produce a molecular ion). The APCI general solvent system of methanol/water with acid was used. In both negative and positive modes, a group of $[M+H]^+$, $[M-H]^-$ peaks were obtained. Additionally, the solvent system used for triglycerides (**A**: water: isopropanol (60:40) + 25mM ammonium formate; **B**: water: isopropanol: *n*-butanol (10:10:80) + 25mM ammonium formate) was used, yielding a group of $[M+NH_4]^+$ peaks that was consistent with the molecular ion obtained from $[M+H]^+$. Figures 5.33 and 5.34 show the positive and negative mode spectra,

respectively. The isotope peak for carbon is about 60%, indicating trimer structures. The MS data are summarized in Table 5.9.

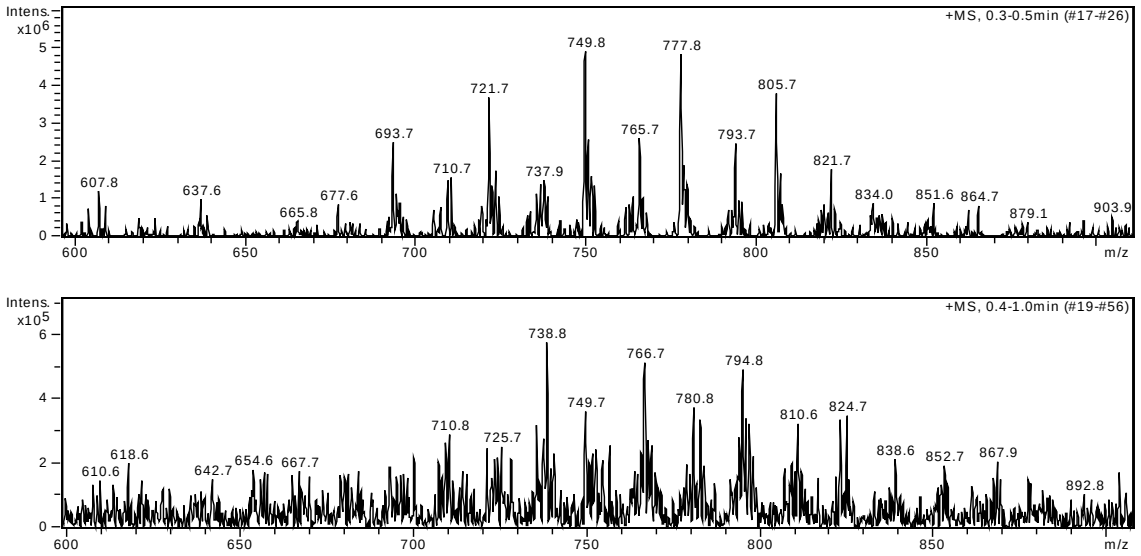


Figure 5.33. Positive MS spectra of aliphatic compound 3 Top: $[M+H]^+$; bottom: $[M+NH_4]^+$.

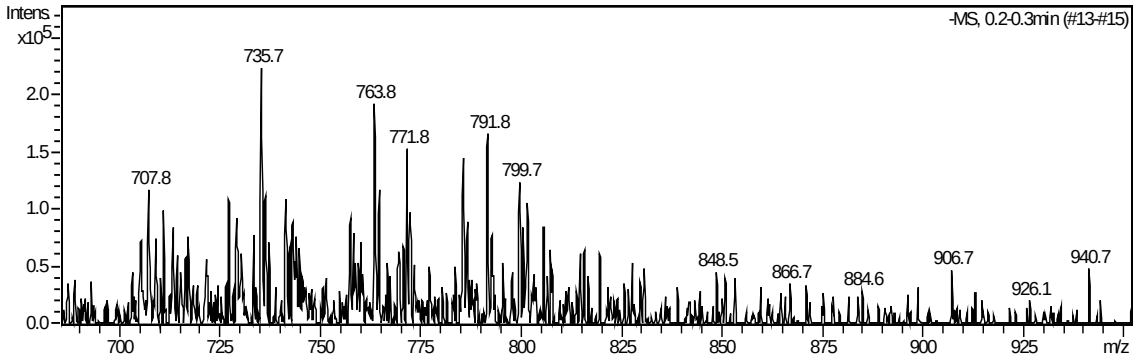


Figure 5.34. Negative MS spectrum of sample aliphatic compound 3.

Figure 5.33 shows the excellent agreement of $[M+H]^+$ and $[M+NH_4]^+$ ions, e.g. $[M+H]^+$ of m/z 693.7, 721.7, 749.8, 777.8, etc. corresponding to $[M+NH_4]^+$ of m/z 709.8, 738.8, 766.8, 794.7, etc. Together these results confirm their molecular weights as 692.7, 720.7, 748.8, 776.8, etc. (see Table 5.9).

These masses are only consistent with CH_3_____OH and CH_3_____CH_3 end groups. Moreover, the MS results show that some molecules have both positive and negative ions, but others exhibit only positive ions. For a compound with a CH_3_____OH structure, it is easy to either add or lose protons, so both negative and positive ions are possible. However, for compounds with CH_3_____CH_3 structures, it is easy to add a proton but hard to lose one, so they only show positive ions.

Table 5.9 MS data for aliphatic compound 3

Positive ion $[M+H]^+$	Negative ion $[M-H]^-$	Molecular Weight
849	847	848 (3: 18 C)
821	819	820 (2:18 C, 1:16C)
793	791	792:18C, 2:16C) or (2:18C, 1:14C)
765	763	764 (3: 16 C)
737	735	736 (2:16C, 1:14C)
709	707	708 (1:16C, 2:14C)
681	679	680 (3: 14 C)

Positive ion $[M+H]^+$	Positive ion $[M+NH_4]^+$	Negative ion	Molecular Weight
833		No	832 (3: 18 C)
805	822.7	No	804 (2:18 C, 1:16C)
777	794	No	776 (1:18 C, 2:16C)
749	766	No	748 (3: 16 C)
721	738	No	720 (1:14 C, 2:16C)
693	710	No	692 (2:14 C, 1:16C)
665	682	No	664 (3: 14 C)
637	654	No	636

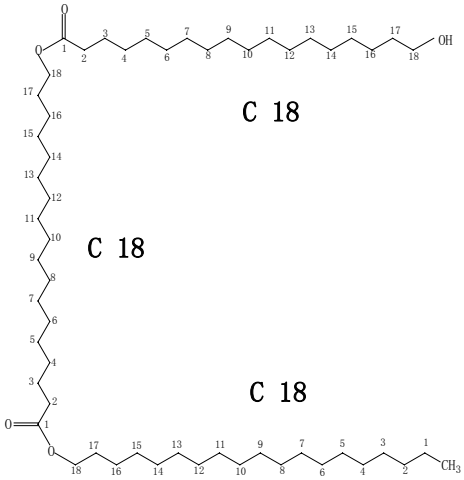
5.3.3 The structure of aliphatic compound **3**

Because the MS data display isotope peaks for carbon of about 60%, trimer ester structures are indicated. To determine the final structures, some assumptions were made: (1) fatty acids of the same or similar chain length are present, as in neutral fats; (2) because the main aliphatic monomers in potato have all even and primarily C18, C16,

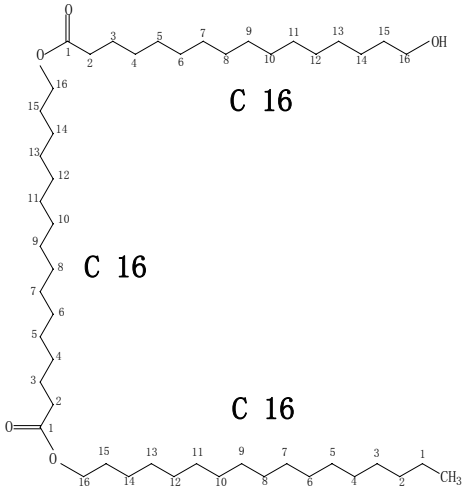
C14 chain lengths, the most common combinations in ester trimers will include C18, C16, and C14 units. Of course, other combinations are also possible, so MS^n data are needed to confirm these suppositions.

5.3.3.1 CH_3 _____OH Trimer structures

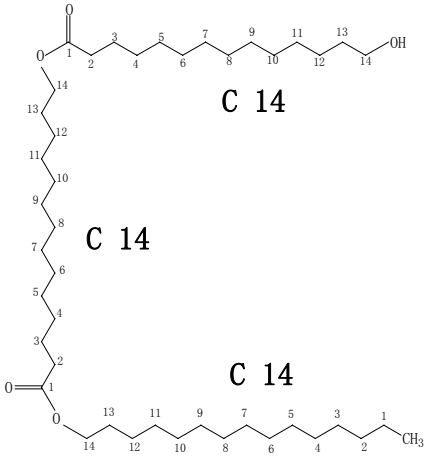
a. MW=848: the most probable structure is C18, C18, C18



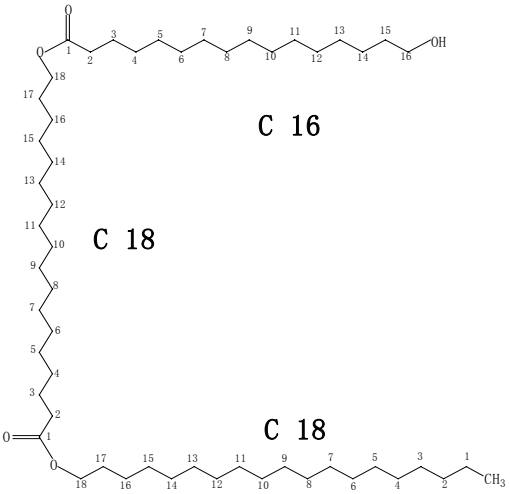
b. MW=764: the most probable structure is C16, C16, C16

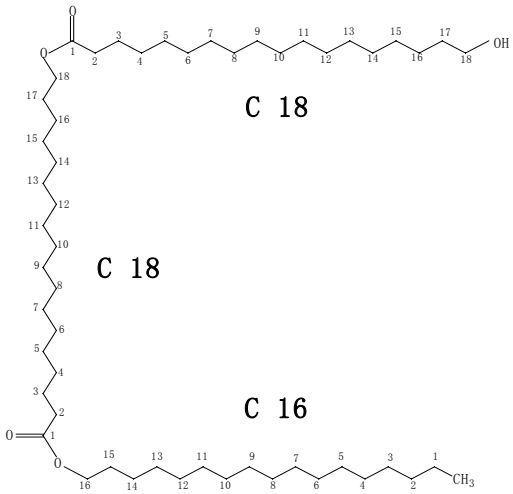


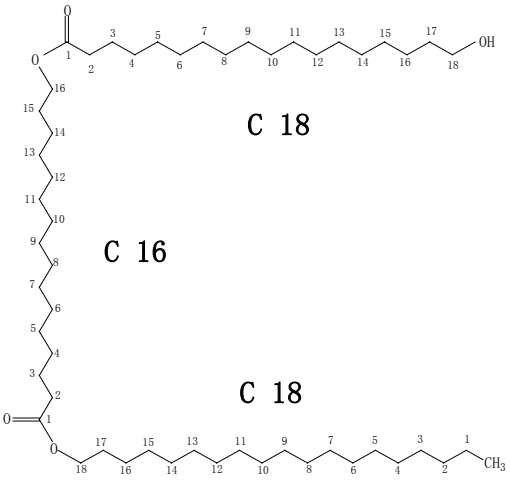
c. MW=680: the most probable structure is C14, C14, C14



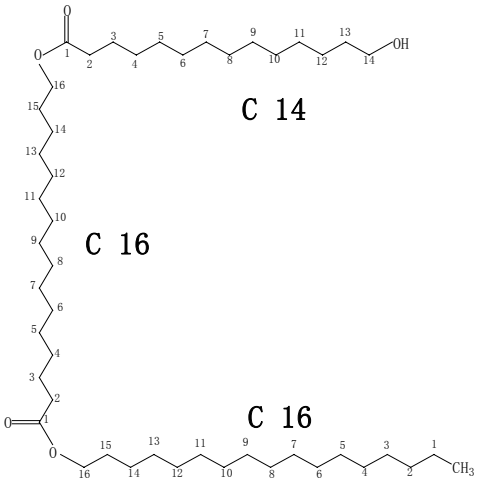
d. MW= 820: the most probable structure is 2 18 C, 1 16C (3 combinations)

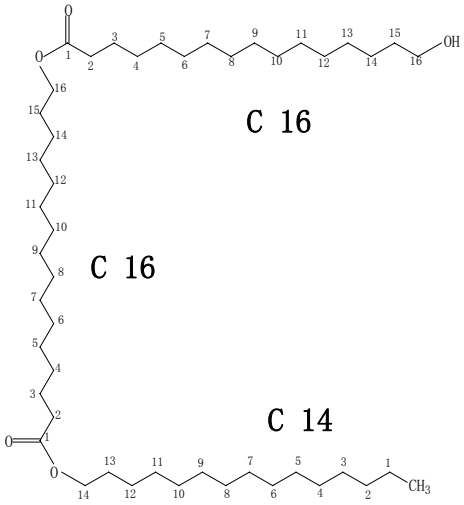


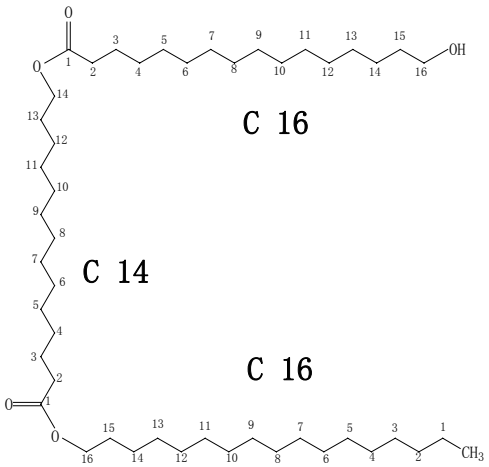




e. MW=736, the most probable structure is 2 16 C, 1 14C (3 combinations)







f. MW=792,

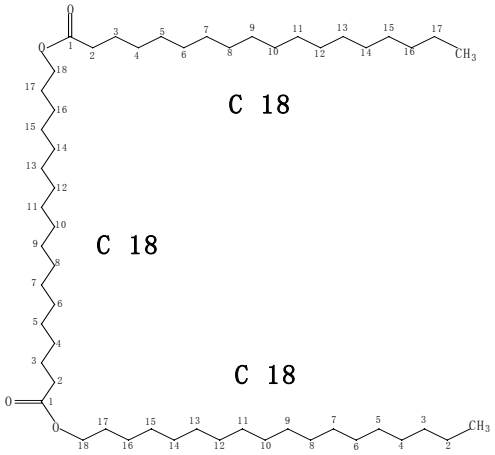
(1 18C, 2 16C): 3 possible combinations (structures not shown)

(2 18C, 1 14C): 3 possible combinations (structures not shown)

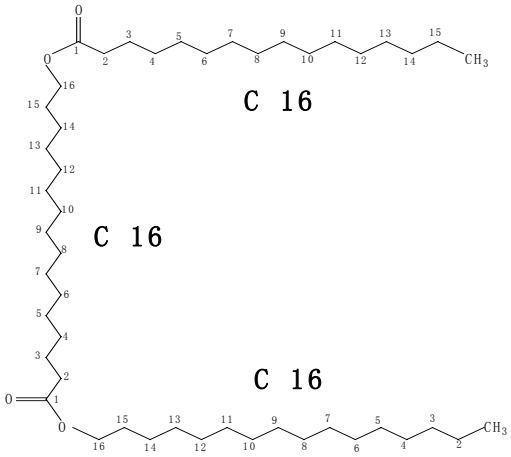
g. MW=708, (1 16C, 2 14C): 3 possible combinations (structures not shown)

5.3.3.2 CH_3 _____ CH_3 Trimer structures

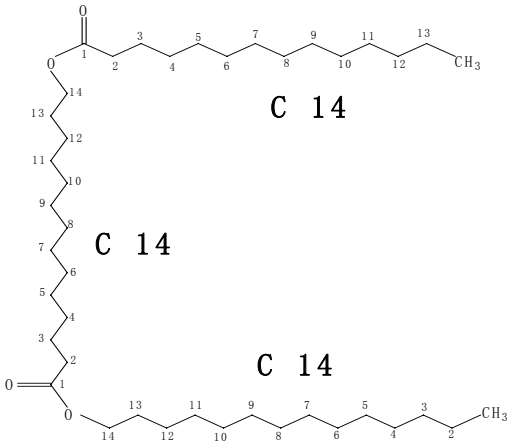
a. MW=832: the most probable structure is C18, C18, C18



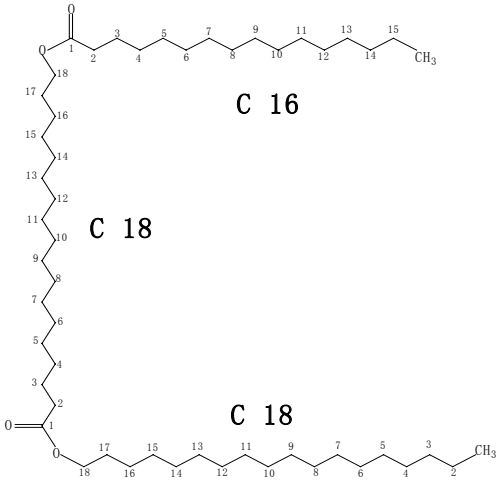
b. MW=748: the most probable structure is C16, C16, C16

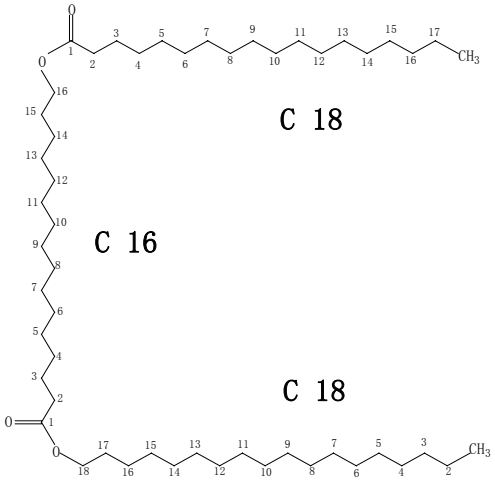


c. MW=664, the most probable structure is C14, C14, C14



d. MW=804: the most probable structure is (2 C18, 1 C16, 2 combinations)





e. MW=776: the most probable structure has 1 C18 and 2 C16 (not shown)

f. MW=720: the most probable structure has 1 C14 and 2 C16 (not shown)

g. MW=692: the most probable structure has 2 C14 and 1 C16 (not shown)

5.3.4 LC/MS^n of aliphatic compound **3**

LC/MS^n for aliphatic compound **3** did not yield fragments, suggesting that a harsher fragmentation method should be used. GC/MS is not suitable for this purpose because the molecular weight range is beyond its detection limit.

5.4 Aromatic compound 1 (#11-TLC-4-1)

5.4.1 1D and 2D NMR of aromatic compound **1** (see Section 4.2.2)

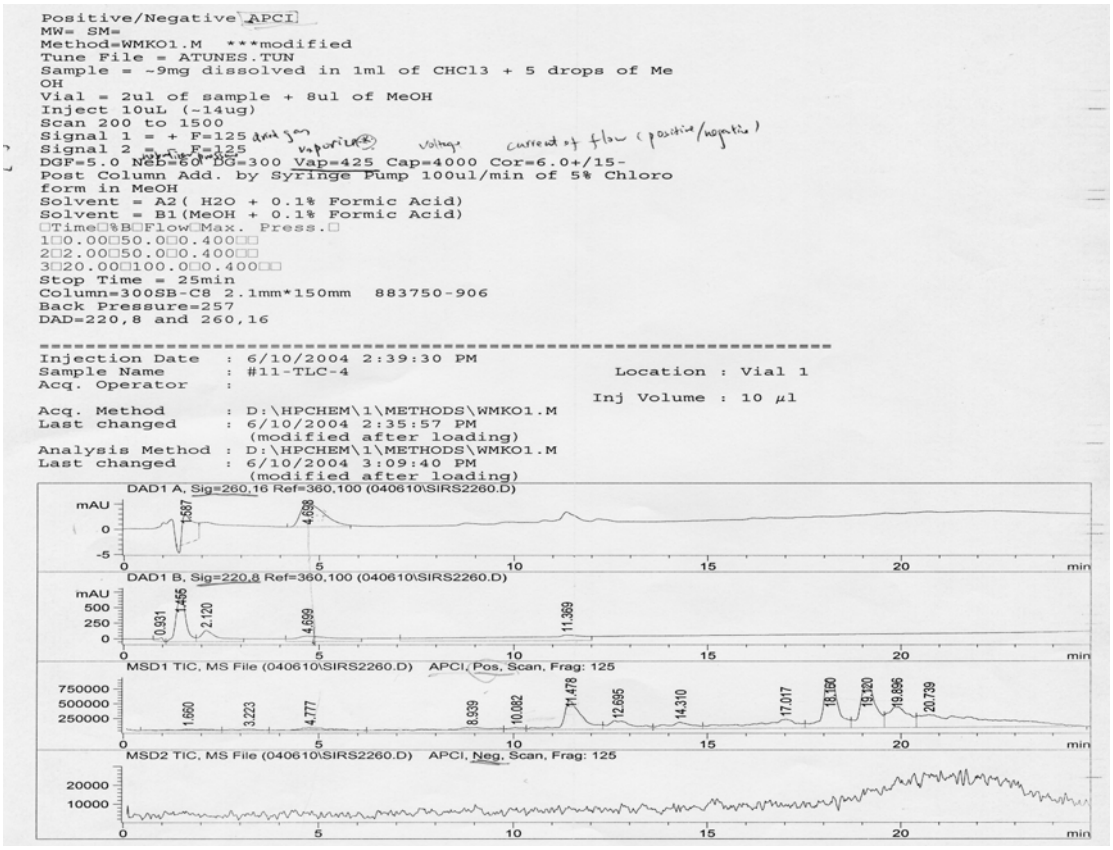
5.4.2 Mass spectrum of sample **#11-TLC-4-1**

Although the NMR spectrum reveals the functional groups and sometimes relative numbers present in the molecule, it cannot provide information on absolute molecular weight without additional assumptions. MS can provide information about molecular weight and sometimes chemical structure, so combining these two methodologies can

yield the molecular structure. Usually a mixture of compounds is initially injected into the LC for separation and then immediate entry to the mass spectrometer.

LC/MS was used to analyze sample #11-TLC-4. Figures 5.35, 5.36, and 5.37 show results obtained with APCI, APPI, and ESI ionization methods, respectively. For each method, there are two parts. The top part shows LC separation with UV detection (260 nm for aromatic molecules, 220 nm for aliphatic molecules) and the total ion current (TIC) in MS (positively and negatively charged particles, respectively). The bottom part shows the MS spectrum corresponding to an elution time (4.6 min) with measurable UV absorbance and TIC. APCI was tried first because it usually works well for aliphatic compounds. APPI is similar to APCI, whereas ESI is well-suited to the analysis of polar molecules. The expectation for this aromatic compound is the observation of UV absorption and a molecular formula of 314, corresponding to $\mathbf{C_{18}H_{18}O_5}$.

The APCI-MS spectrum should display a peak for $M+H^+$ in positive mode (H^+ comes from H_2O, MeOH and acid if added); in negative mode, $M-H^+$ or $M-Cl^-$ (M-35, and M-37, Cl^- from $CHCl_3$ in toluene) can be detected. If the molecular weights obtained from positive and negative mode match each other, the molecular ion is confirmed. However, it is also possible for some compounds to become positively charged but not negatively charged because of their chemical nature.



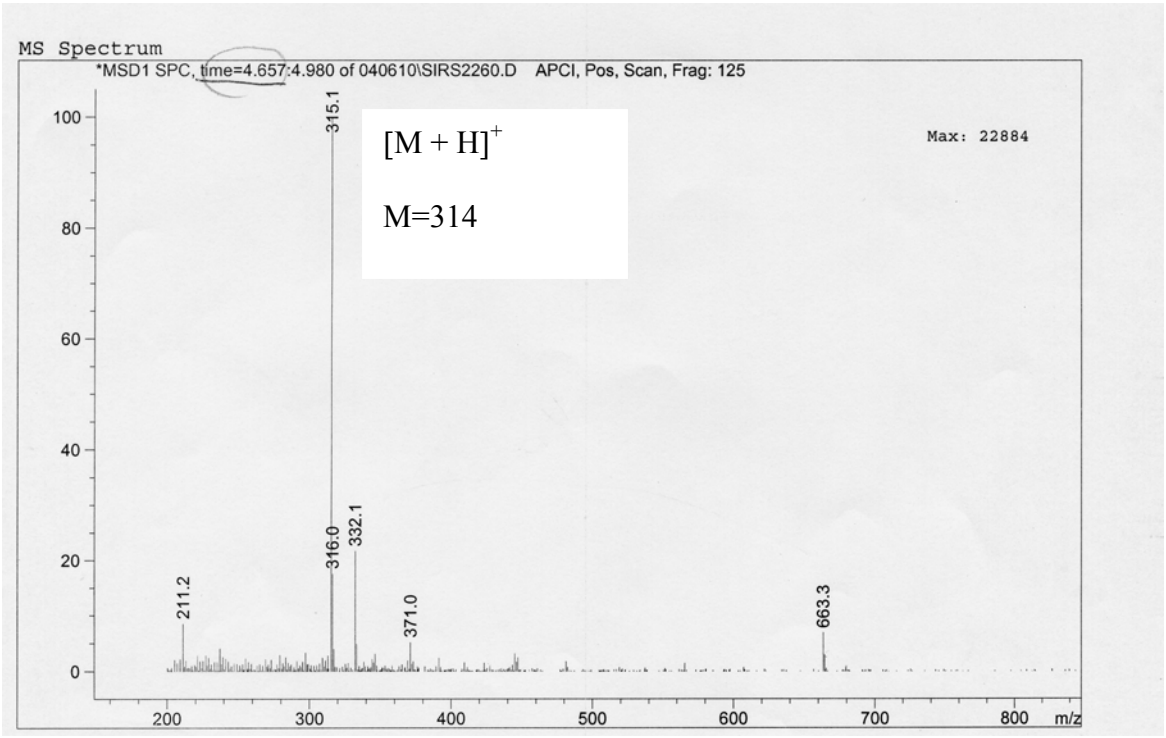
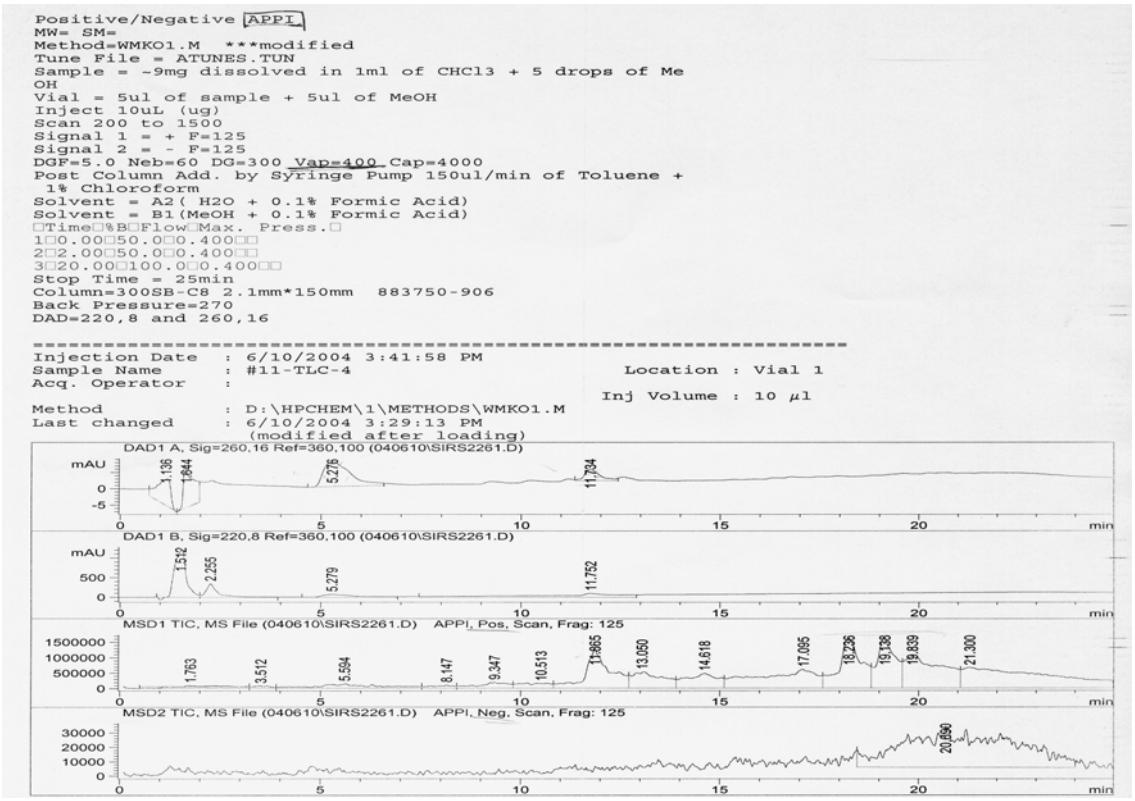


Figure 5.35. LC-APCI-MS at time 4.7 min for #11-TLC-4.



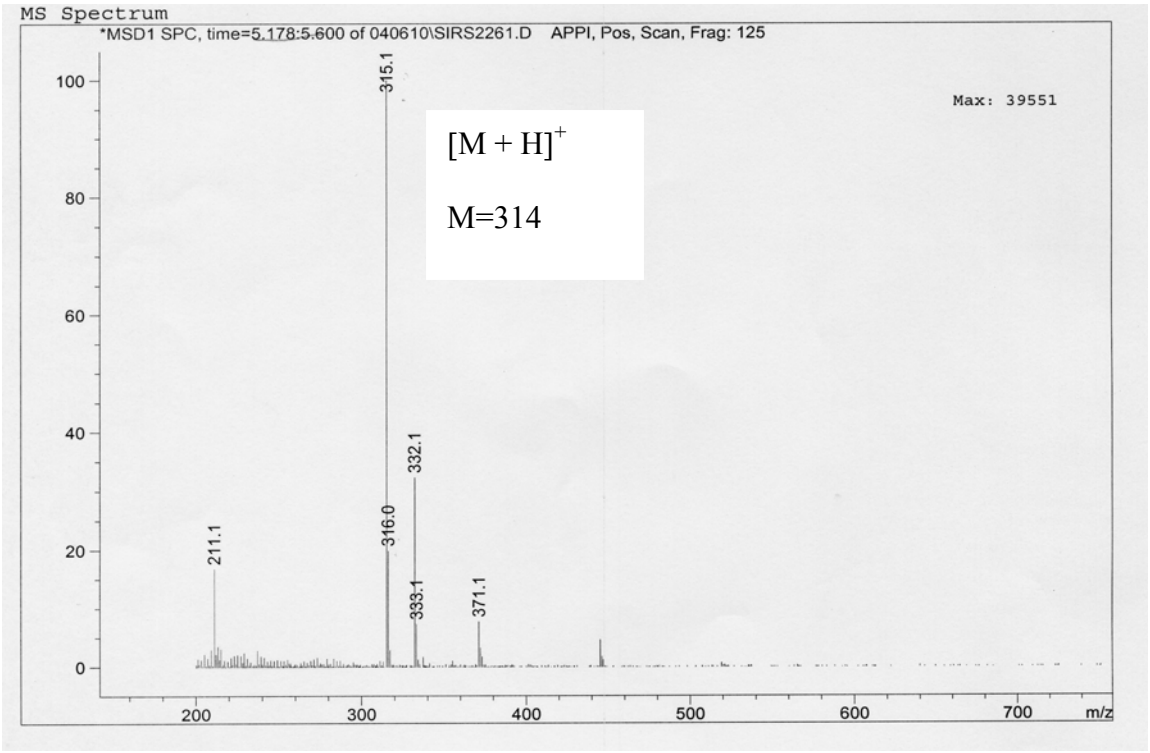
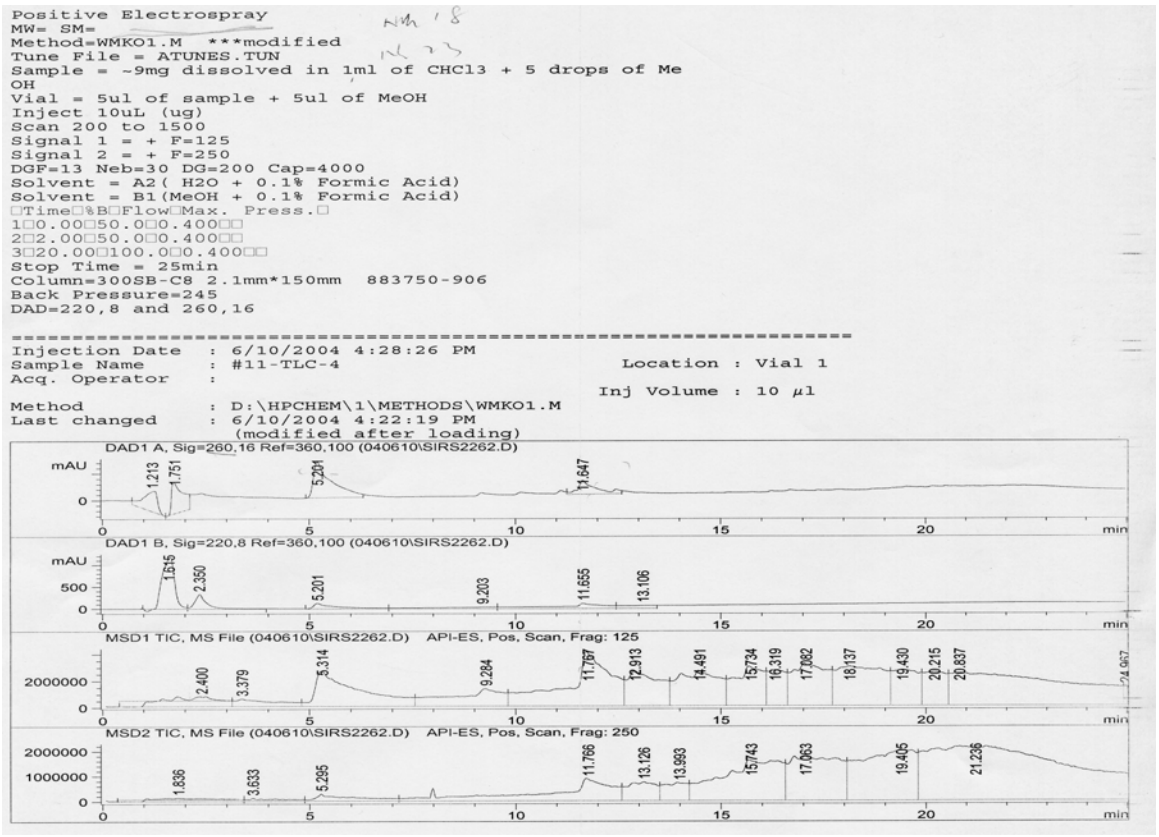


Figure 5.36. LC- APPI-MS at time 5.2 min for #11-TLC-4.



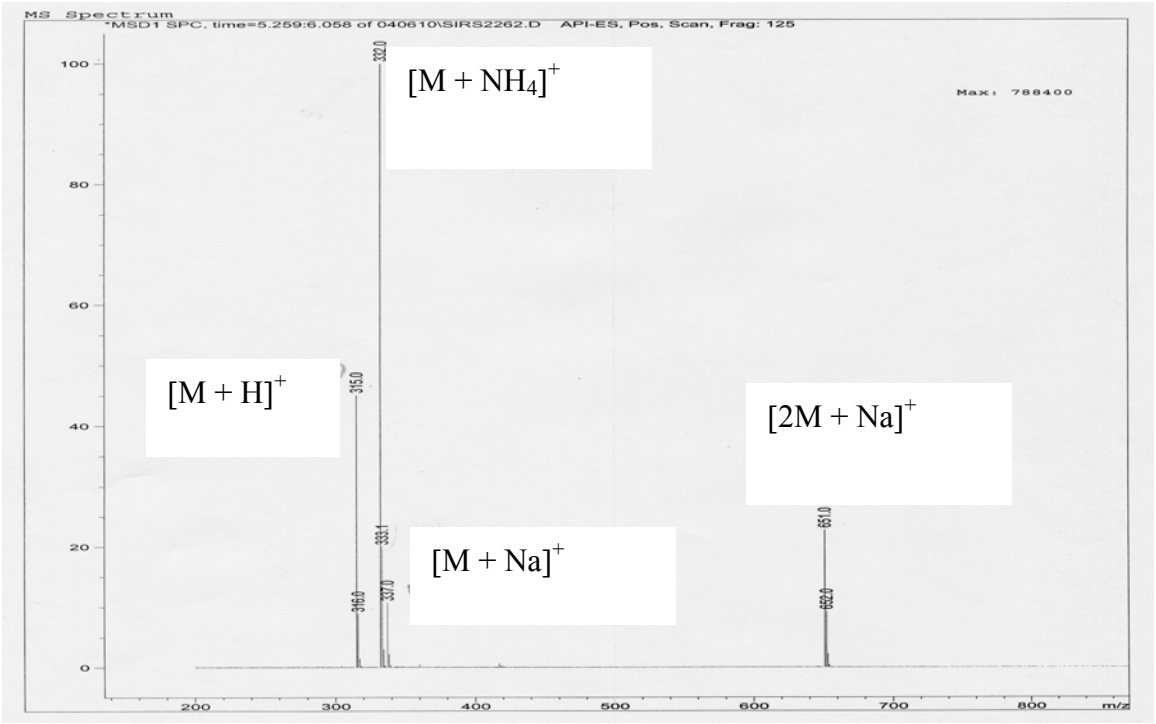


Figure 5.37. LC-ESI-MS at time 5.2 min for #11-TLC-4.

For sample #11-TLC-4, UV absorption shows one peak eluting at 4.7 min, producing an MS spectrum albeit with low TIC. Figure 5.35 (bottom) shows a molecular ion in LC-APCI-MS at 315 $[M+H]^+$ and 316 (isotopic peak), but for negative mode there is no sign of ionization. Thus based on these results, the molecular weight of #11-TLC-4-1 cannot be determined with certainty. In order to get further confirmation of the suggested structure in Figure 4.13, APPI-MS was also done (Figure 5.36), yielding similar results to APCI: 315 $[M+H]^+$, 316 (isotopic peak), but no negative ion.

Finally, ESI ionization was tried for sample #11-TLC-4, despite the fact that it usually doesn't work very well for non-polar compounds. As described in Chapter 2, the ESI-MS spectrum may have M+18 (NH_4^+), M+23 (Na^+), and M+1 (H^+) peaks in positive mode. Figure 5.37 (bottom) shows peaks at 332 ($M+NH_4^+$), 337 ($M+Na^+$), 315 ($M+H^+$); each shows an isotopic peak. This result confirms the molecular weight as 314. ESI-MS also shows a peak at 651, corresponding to a sodium ion associated with two molecules (2*313+23). This latter result also indicates that sample #11-TLC-4-1 associates easily to a positive ion: the ether bond in the central portion of the molecule has a partial negative charge on oxygen, so positive ions can be easily associated with it, whereas negative ions are rejected. This argument also accounts for the absence of negative-mode APCI and APPI spectra. APCI, APPI, and ESI-MS data together confirm the structure of #11-TLC-4-1 as shown in Figure 4.13.

It should be noted that UV detection at 220 and 260 nm shows only two peaks, but APCI-MS in positive mode shows several peaks corresponding to the aliphatic molecules in the mixture #11-TLC-4. The NMR spectrum of #11-TLC-4 also has peaks suggestive

of aliphatic monomers and oligomers. Further separation is needed to identify these latter materials (see Section 4.3).

Molecular Weight of #11-TLC-4 from APPI, APCI, and ESI Mass Spectrometry

APPI $[M+H]^+$ Isotopic peak		**APCI $[M+H]^+$ Isotopic peak**		**ESI ($[M+H]^+$, $[M+NH_4]^+$, $[M+N_a]^+$), Isotopic peak**	
LC Elution time (min)	**MW**	LC Elution time (min)	**MW**	LC Elution time (min)	**MW**
5.2	**314**	4.6	**314**	5.2	**314, 2M+Na**

Thus MS data finally confirms the structure of aromatic compound **1** (#11-TLC-4-1) based on all data from 1D NMR (^1H, ^{13}C), 2D (COSY, gHSQC, gHMBC), MS (ESI, APCI, APPI), and ACD simulation. The structure of aromatic compound **1** (#11-TLC-4-1) is shown below (Figure 5.38):

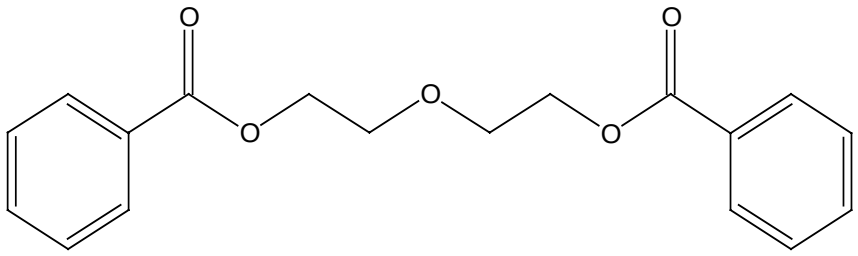


Figure 5.38. The structure of aromatic compound 1 (#11-TLC-4-1).

5.5 Aromatic compound 2

5.5.1 1D and 2D NMR identification of aromatic compound **2**

5.5.1.1 Aromatic compound **2-1**

Aromatic compound **2** was separated by solvent extraction of the soluble suberin depolymerization products. The ^{1}H NMR spectrum (Figure 5.39) shows five peaks between 6 and 8 ppm, two of which are doublets at 7.98 (A) and 6.86 (B) ppm corresponding to carbons at 132.3 and 115.1 ppm in the HMQC spectrum (Figure 5.40), respectively. The two doublet peaks have an integration ratio of 1:1 and the same J coupling constant: 8.01. The COSY spectrum (Figure 5.41) shows correlation between H_A and H_B. All of above information shows that this aromatic compound is a para-substituted benzene ring. The provisional structure of aromatic compound **2-1** is shown in Figure 5.43 and discussed further below.

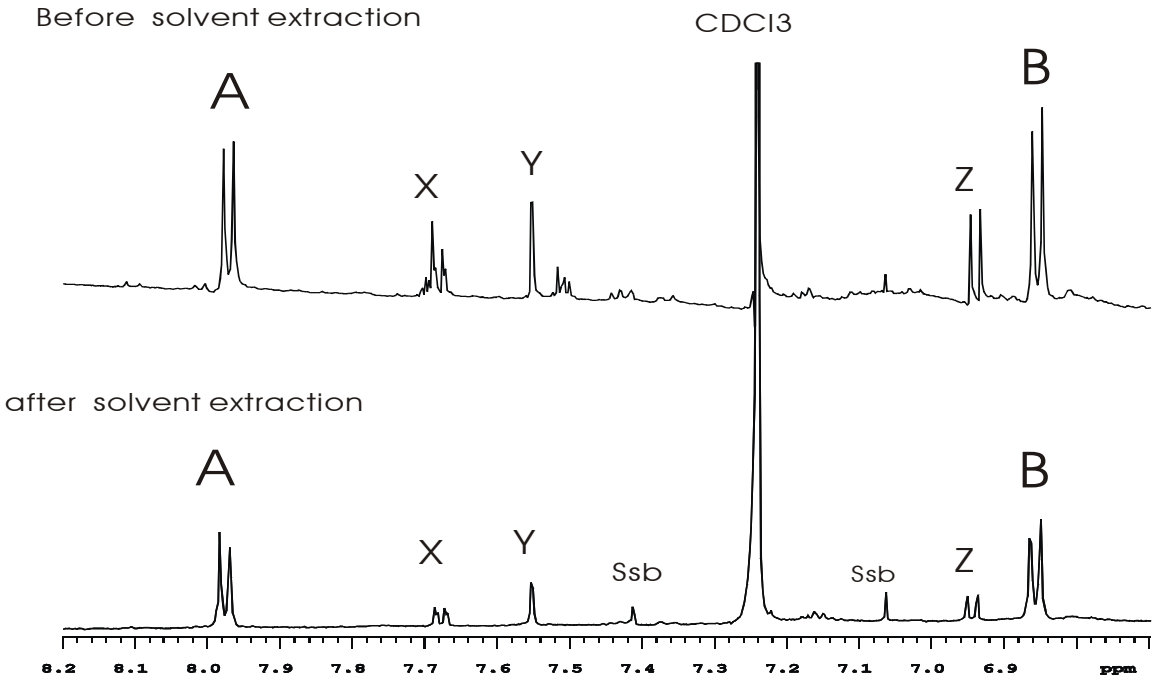


Figure 5.39. 600 MHz ^{1}H NMR spectrum of aromatic compound 2 before and after solvent extraction.

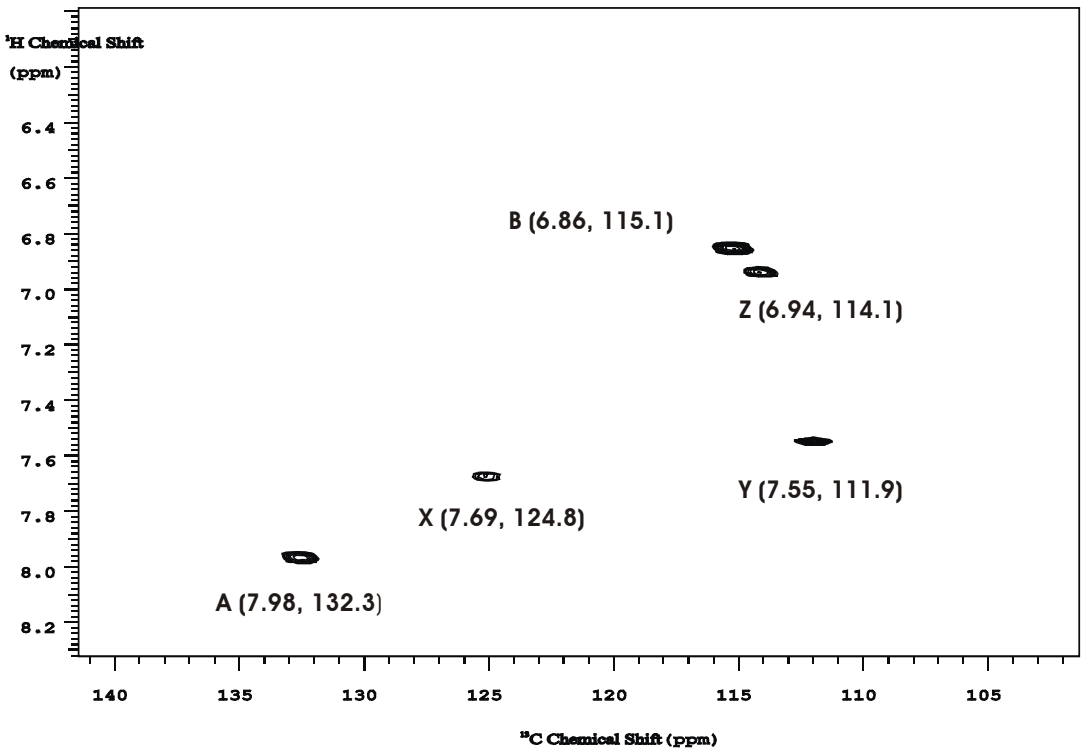


Figure 5.40. gHMQC spectrum of aromatic compound 2.

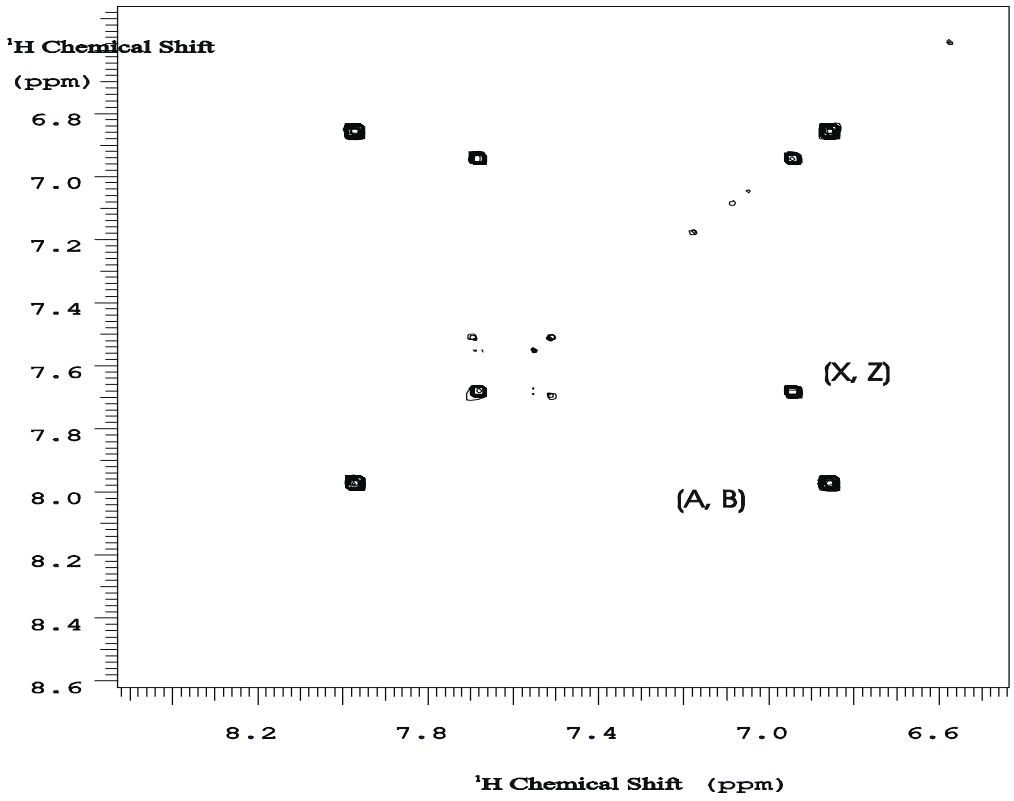


Figure 5.43. COSY spectrum of aromatic compound 2.

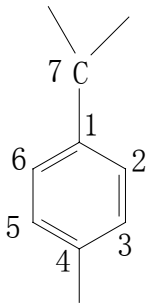


Table 5.10 Proton and carbon chemical shifts of aromatic compound 2-1

Position number	CH_n	Chemical shifts (ppm)	nH	Chemical shifts (ppm)
1	C	121.3	0	
2	CH	132.3	1	7.98
3	CH	115.1	1	6.86
4	C	160.2	0	
5	CH	115.1	1	6.86
6	CH	132.3	1	7.98

Table 5.11 gHMBC (Figure 5.43) correlation of aromatic compound 2-1

H \ C (ppm)	1 (121.3)	2 (132.3)	3 (115.1)	4 (160.2)	5 (115.1)	6 (132.3)	7 (170.1)
2 (7.98)				*		*	*
3 (6.86)	*				*		
5 (6.86)	*		*				
6 (7.98)		*		*			*

- represent cross-peaks

The gHMBC spectrum (Figure 5.42) shows that protons in positions 2 and 6 correlate to a carbon at 170.1 ppm, so the carbon at position 7 is a carbonyl carbon. For this carbon, there are no correlations with other protons, therefore aromatic compound **2-1** is most probably an acid. The chemical shift of the carbon at position 4 is 160.2 ppm; moreover it has no other correlation to protons, which indicates a hydroxyl group is probably connected to it.

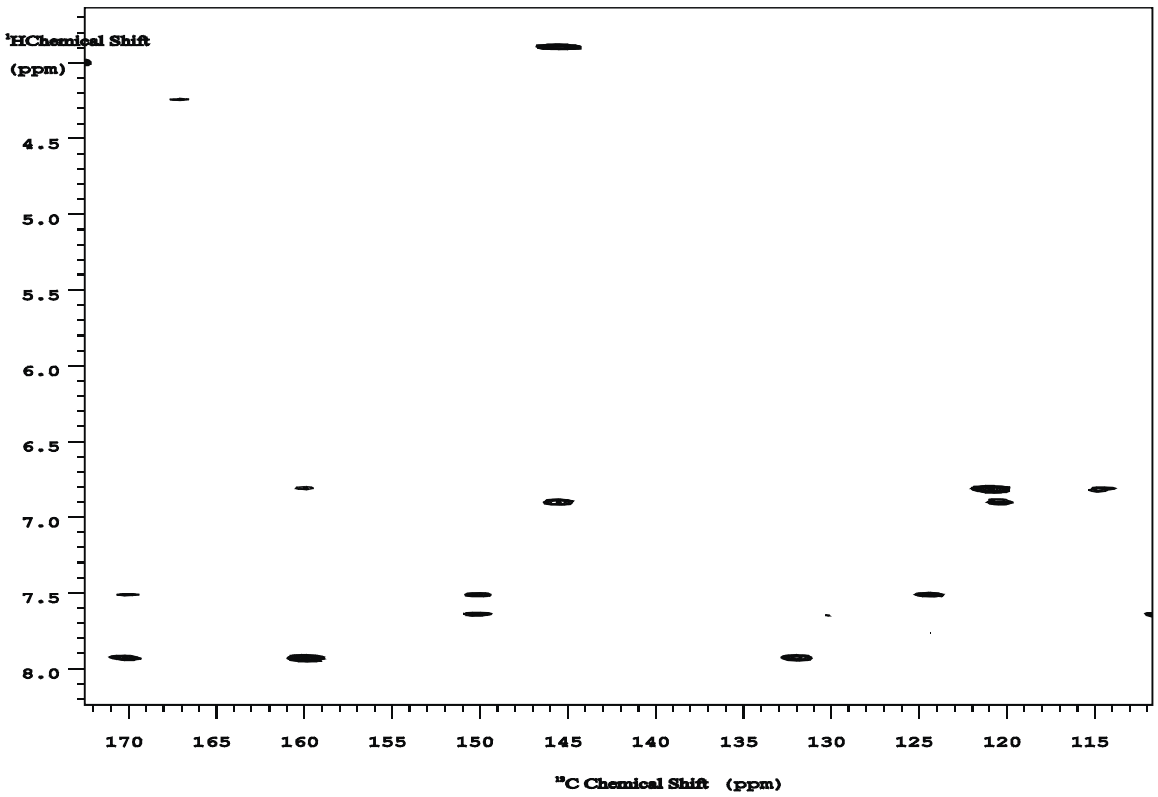
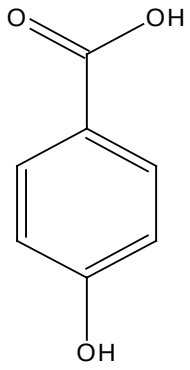


Figure 5.42. gHMBC spectrum of aromatic compound 2.

The ACD calculations (not shown) match the above NMR results very well. ^{1}H NMR (chemical shift, peak number, integration and J coupling constant) and 2D NMR (HMBC, HMQC, COSY) confirm the structure shown in Figure 5.43, which is different from the commonly known aromatics that have been found so far in suberin. MS data is desirable for confirmation of the proposed structure (see Section 5.5.2).



Monoisotopic Mass: 138.031694 Da

Figure 5.43. Provisional structure of aromatic compound 2-1.

5.5.1.2 Aromatic compound **2-2**

The ^{1}H NMR spectrum (Figure 5.39, after solvent extraction) shows five peaks between 6 and 8 ppm, three of them at 7.69 (X,d), 7.55 (Y, s), and 6.94 (Z, d) ppm corresponding to carbons at 124.8, 111.9 and 114.1 ppm in the HMQC spectrum (Figure 5.40), respectively. The COSY spectrum (Figure 5.41) shows correlation between peaks H_X and H_Z, while peak Y is a single peak that has no correlation with the above peaks. All of the information shows that this aromatic compound has a 3, 4-disubstituted benzene ring.

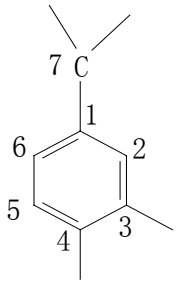


Table 5.12 Proton and carbon chemical shifts of aromatic compound 2-2

Position number	CH_n	Chemical shifts (ppm)	nH	Chemical shifts (ppm)
1	C	120.9	0	
2	CH	111.9	1	7.55
3	C	145.8	0	
4	C	150.5	0	
5	CH	114.1	1	6.94
6	CH	124.8	1	7.69

Table 5.13 gHMBC (Figure 5.42) correlation of aromatic compound 2-2

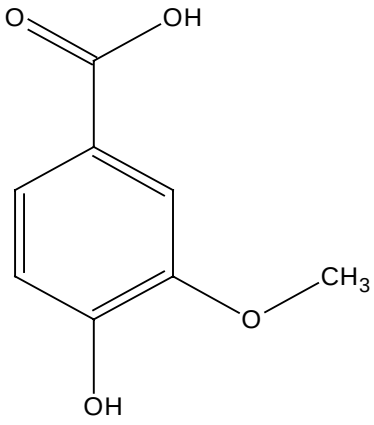
H\C (ppm)	1 (120.9)	2 (111.9)	3 (145.8)	4 (150.5)	5 (114.1)	6 (124.8)	7 (170.3)
2 (7.55)				*	*	*	*
5 (6.94)	*		*	*	*		
6 (7.69)		*		*			*

***cross-peaks**

The gHMBC spectrum (Figure 5.42) shows that protons in positions 2 and 6 correlate to a carbon at 170.3 ppm, so the carbon at position 7 is a carbonyl carbon; for this carbon,

there are no correlations with other protons, so it is most probably an acid. The chemical shift of the carbon at position 4 is 150.5 ppm, with the absence of other correlations of protons with this carbon indicating an attached hydroxyl group. Also the chemical shift of carbon at position 3 is 145.8 ppm, and there is a cross-peak at (3.93, 145.8 ppm), which indicates a methoxy group at this latter position.

The ACD calculations (not shown) match the above NMR results very well. ^{1}H NMR (chemical shift, peak number, integration and J coupling constant) and 2D NMR (HMBC, HMQC, COSY) confirm the structure in Figure 5.44, which is again different from the commonly known aromatics that have been reported so far in suberin.



Monoisotopic Mass: 168.042259 Da

Figure 5.44. Provisional structure of aromatic compound 2-2.

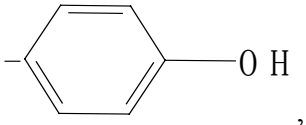
5.5.2 LC/MS and LC/MS^{n} of aromatic compound **2**

NMR spectra were used to determine the provisional structures of two aromatic compounds (Figures 5.43, 5.44), but MS data are desirable to confirm them. LC/MS was used for this analysis. Aromatic compound **2** is a mixture containing aromatic compounds **2-1** and **2-2** and other aliphatics. Both positive- and negative-mode ESI and positive-mode APCI were tried with commonly used solvent systems (acetonitrile/water with acid

and methanol/water with acid, respectively), but no ions formed. Negative-mode APCI ionization successfully produced a strong deprotonated $[M-H]^-$ ion, which further confirms the possible acid structures. Because **2** is a mixture, an LC/MS column and gradient elution were used to separate them. The best gradient was 50-100% methanol (0.1% formic acid) in 10 min with a flow rate of 0.4 ml/min. Figure 5.45 shows LC-APCI-MS negative mode results. The top graph shows LC separation with UV detection at 280 nm; the second graph gives the total ion current (TIC) in negative mode MS. The third, fourth and fifth graphs show MS spectra corresponding to LC elution (or TIC chromatogram) at 3.4-3.6 min, 4.3-4.8 min, and 5.2-6.1 min, respectively.

5.5.2.1 LC/MS and LC/MS^n of LC elution between 3.4 and 3.6 min

The MS spectrum corresponding to 3.4-3.6 min yielded $[M-H]^-$ at m/z 136.8, 166.8, 172.8, and 260.9, so their molecular weights are 138, 168, 174, and 262, respectively. Its corresponding HPLC chromatogram (Figure 5.45, top) shows UV absorption at 280 nm. It probably contains the aromatic compounds **2-1** and **2-2**. MS^n is needed to confirm their structure.

Figure 5.46 shows the MS^2 spectrum of $[M-H]^-$ at m/z 136.8, which yields one fragment at m/z 92.9. The MS^2 difference of m/z 44 represents the loss of CO_2, whereas the retained peak at m/z 92.9 represents 

, confirming the structure of aromatic compound **2-1** shown in Figure 5.43.

$$136.8\ m/z \xrightarrow[\text{lose } 44]{MS^2} 92.9\ m/z$$

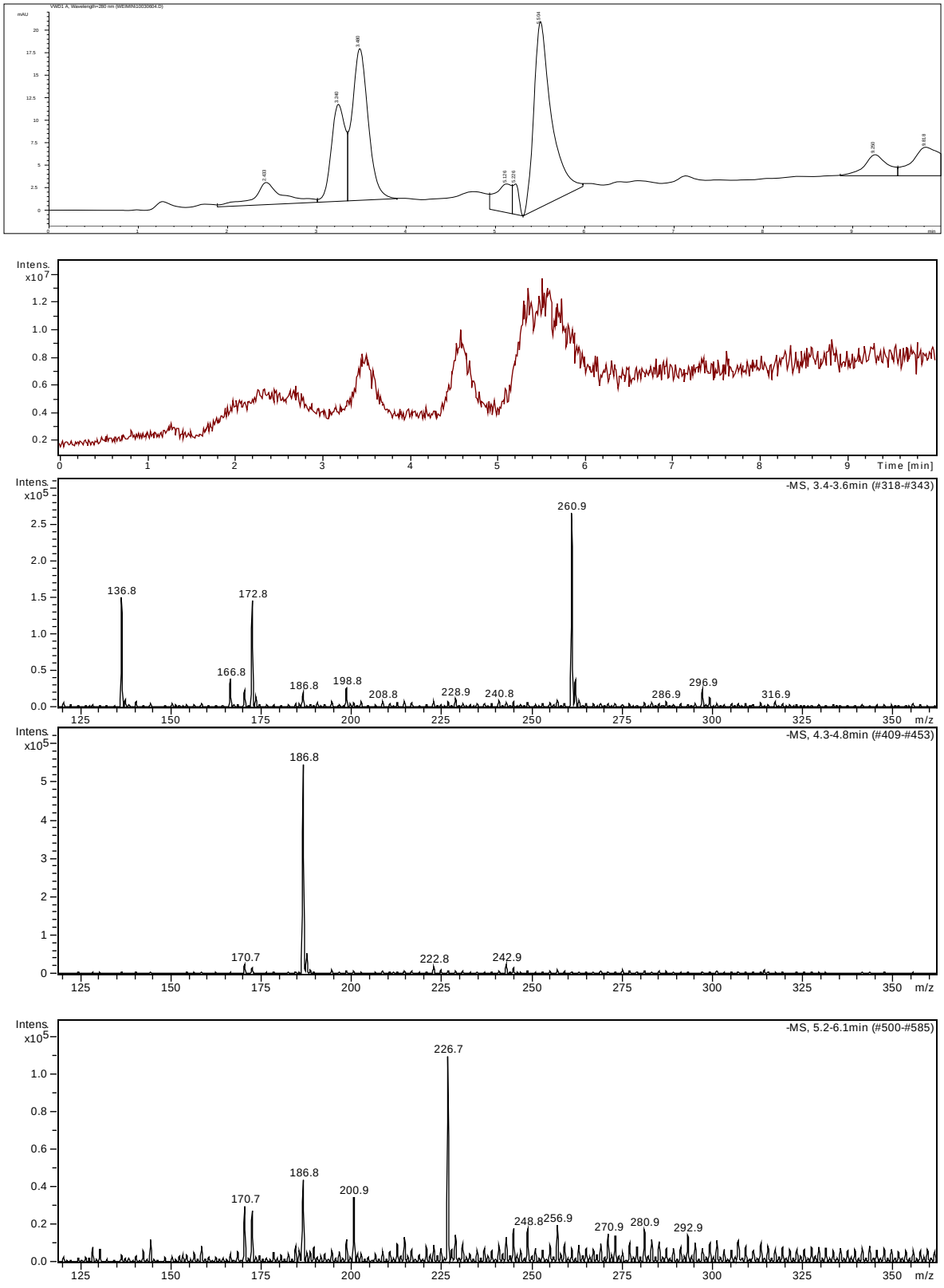


Figure 5.45. APCI-LC-MSD ion trap MS spectrum of aromatic compound 2.

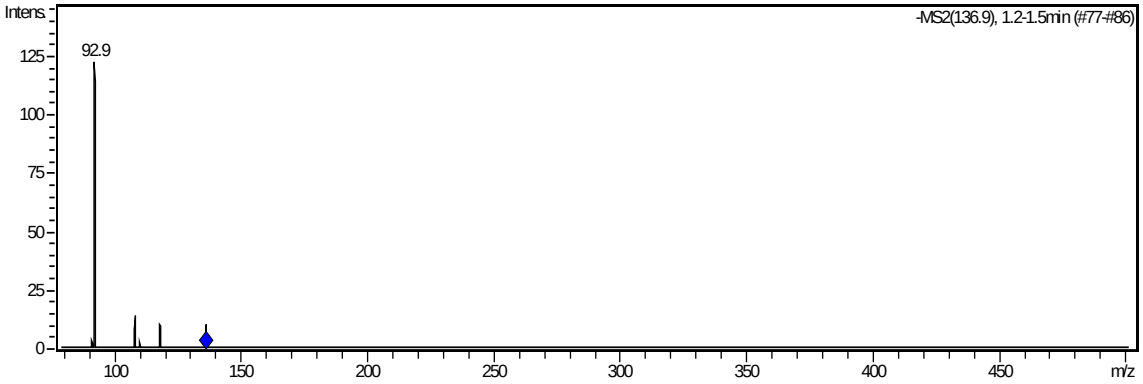
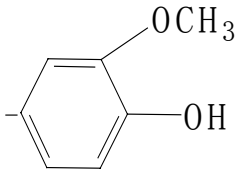


Figure 5.46. APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound 2 (MS 136.8->MS^2).

Figure 5.47 shows the MS^2 spectrum of $[M-H]^-$ at m/z 166.8, and MS^2 yields fragments at m/z 151.7 and 122.8. The difference of 15 between 166.8 and 151.7 represents loss of $-CH_3$. The MS^2 difference of m/z 44 represents the loss of CO_2, and the retained peak at m/z 122.8 represents 

, confirming the structure of aromatic compound **2-2** as shown in Figure 5.44. The fragmentation pathway is shown in Figure 5.48.

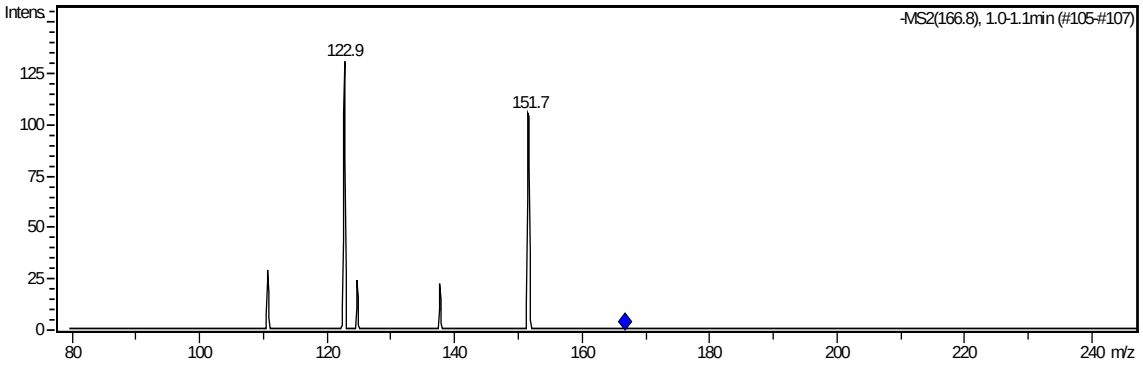


Figure 5.47. APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound 2 (MS 166.8->MS^2).

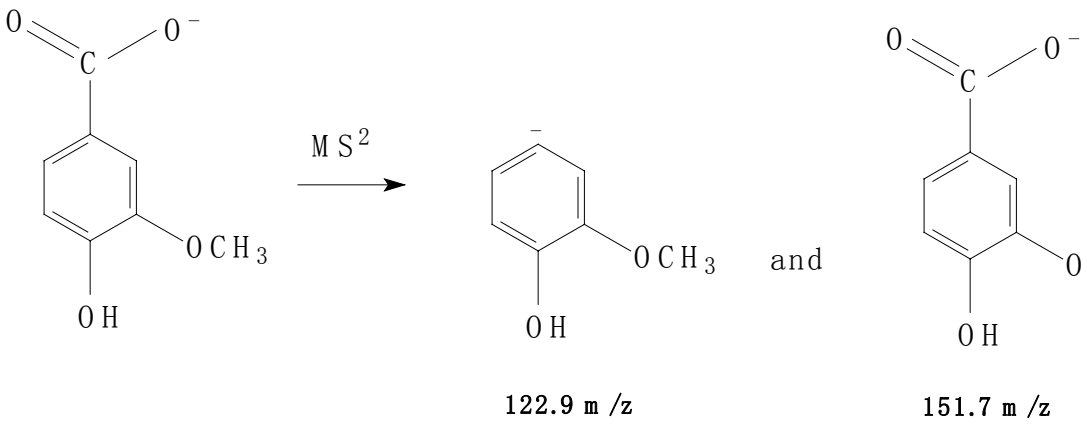


Figure 5.48. The possible fragmentation pathway of aromatic compound 2 (MS 166.8->MS^2).

5.5.2.2 LC/MS and LC/MS^n of LC elution between 4.3 and 4.8 min

The MS spectrum corresponding to 4.3-4.8 min produced $[M-H]^-$ at m/z 186.8, corresponding to a molecular weight of 188. It is notable that the $[M-H]^-$ peak at m/z 186.8 (Figure 5.45, second MS trace) has no UV absorption at 280 nm, indicating it is an aliphatic rather than aromatic compound. Figure 5.49 shows the MS and MS^n spectra of the peak with m/z 186.8.

MS^n of the $[M-H]^-$ peak at 186.9 gave fragment peaks at m/z 168.7, 140.8 and 124.8. The further fragmentation of m/z 124.9 gave fragment peaks at m/z 97. Figure 5.50 shows the fragment ions obtained by MS^2 and MS^3. MS^2 of m/z 186.8 gives 168.8, a loss of 18 usually corresponding to H_2O. Further fragmentation (MS^3) of this ion gives a fragment ion at m/z 124.8. The loss is 44, which implies loss of CO_2. MS^2 of 186.8 also gives peak at 124.9; a possible explanation for this fragment is $CH=CH-(CH_2)_7^-$, while the difference of 62 is consistent with the simultaneous loss of H_2O and CO_2. Further fragmentation of m/z 124.9 gives a fragment at m/z 97, with the loss of 28 representing two CH_2's. Therefore the possible structure of this aliphatic compound may be 10-

hydroxydecanoic acid. A possible fragmentation pathway is shown in Figure 5.50. For this monomer with a typical chain length, MS and MS^n should be done with a standard compound to confirm the structural identification.

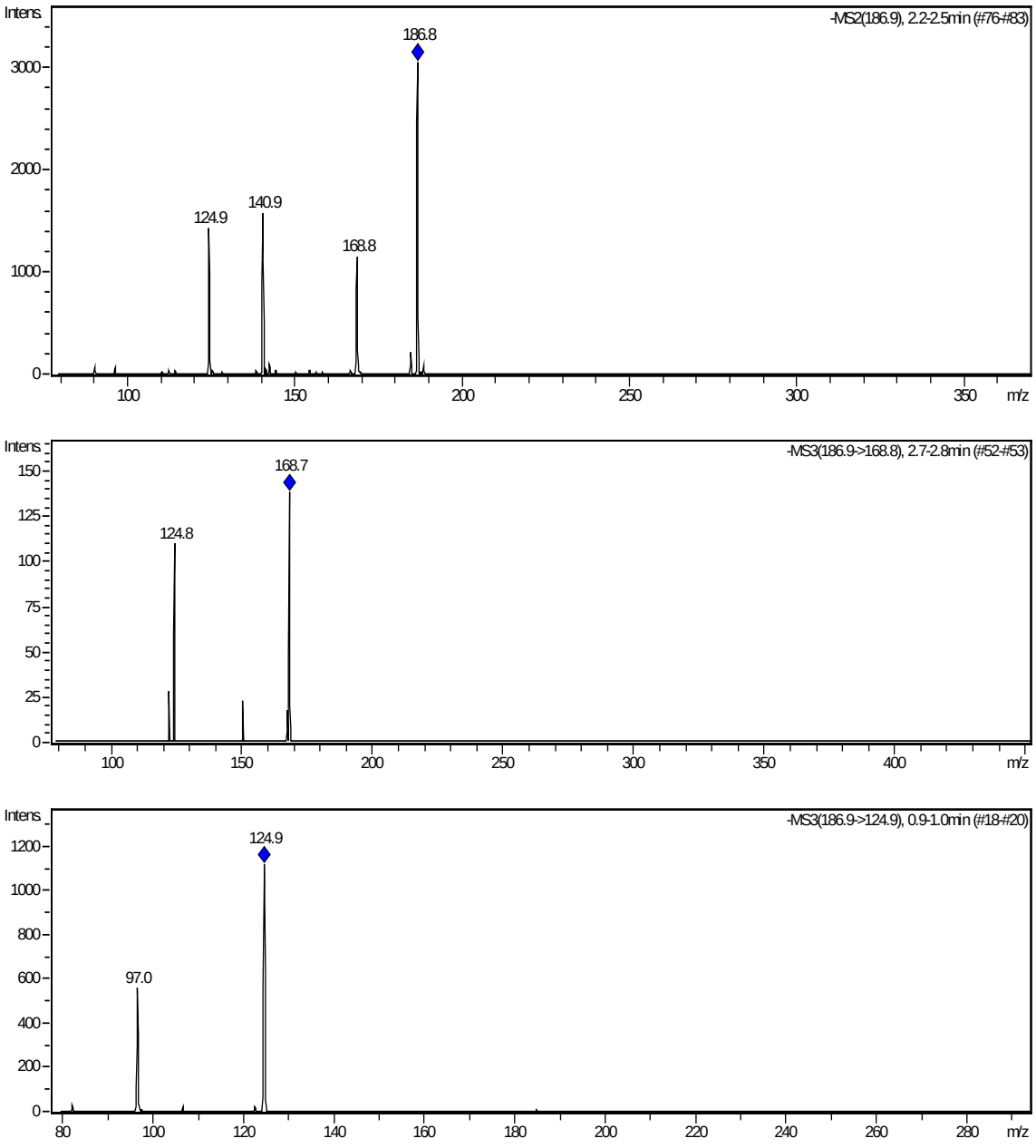


Figure 5.49. APCI-LC-MSD ion trap MS^2 spectrum of aromatic compound 2 (MS 186.9->MS^2).

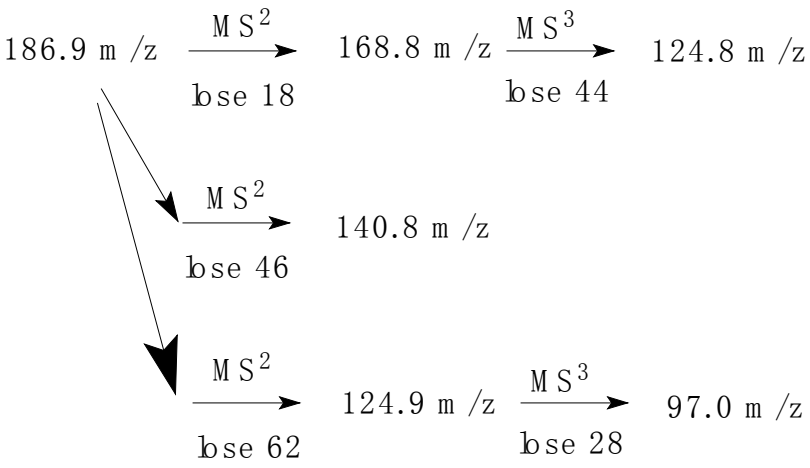


Figure 5.50. APCI-LC/MSD ion trap fragmentation pathways for the peak at m/z 186.9 (MS->MS^2->MS^3).

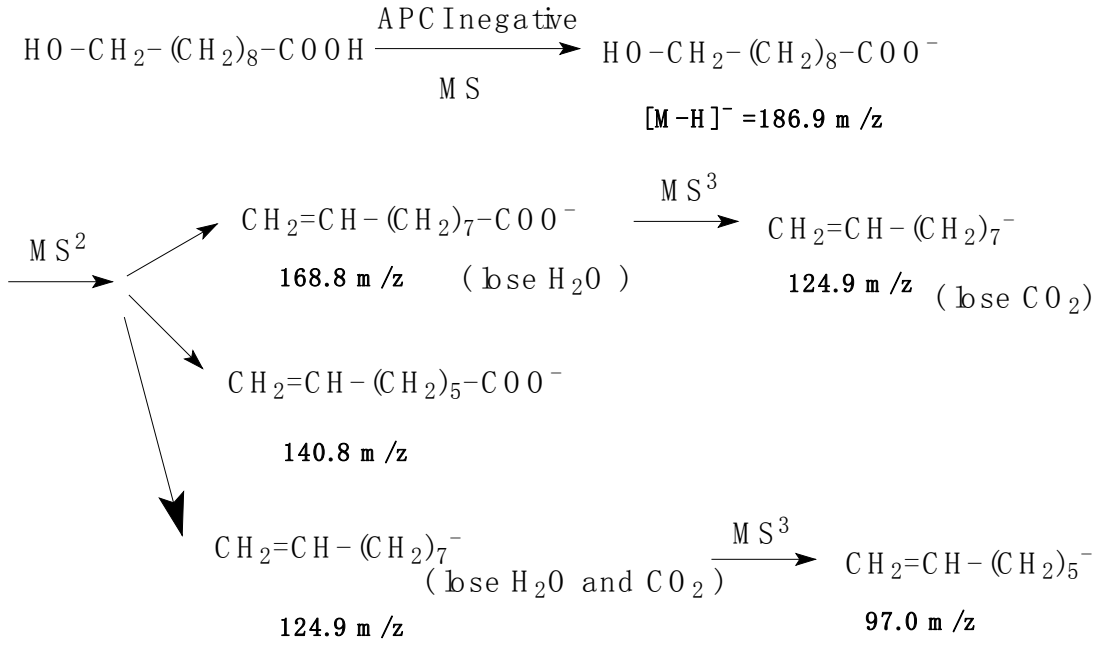


Figure 5.51. Fragmentation pathways for the peak at 186.9 (MS->MS^2->MS^3)

MS and MS^n experiments were performed on the model compound 10-hydroxydecanoic acid. Figure 5.52 shows the MS and MS^n spectra of this standard compound. Compared with the suberin product described above, the MS and MS^n patterns are almost the same. Both show a molecular ion [M-H]$^-$ at m/z 186.9, MS^2

fragments at m/z 140.8, and 168.8. The only difference is that our compound has an MS^2 fragment at m/z 124.9, while the standard does not. Given that the retention time of the standard compound is the same as the unknown, it is possible to conclude they are the same compounds (Dr. Mike Lee, personal communication).

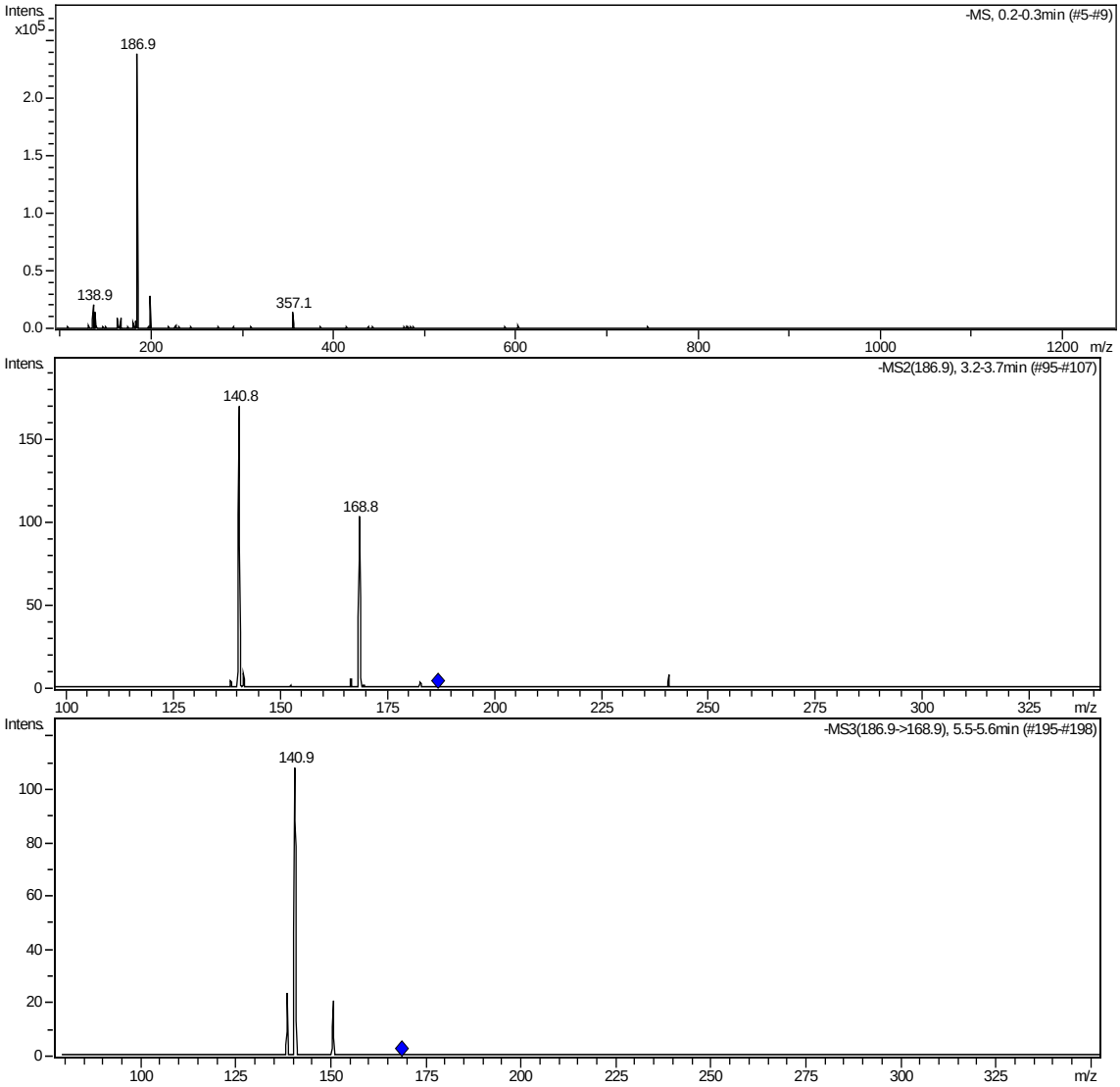
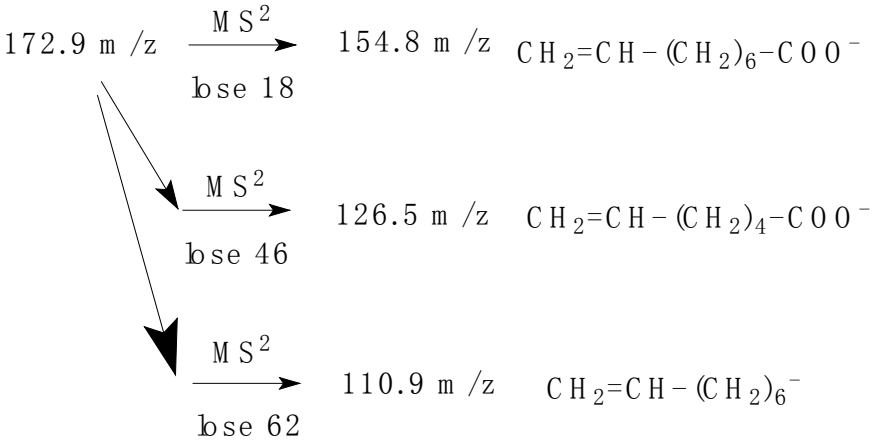


Figure 5.52. APCI-LC-MSD ion trap MS-MS^2-MS^3 spectrum of standard compound 10-hydroxydecanoic acid (MS 186.9->MS^2, 168.8->MS^3).

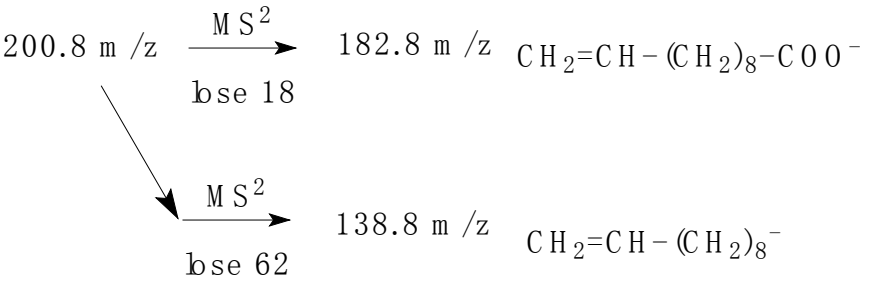
5.5.2.3 LC/MS and LC/MS^n of LC elution between 5.2 and 6.1 min

The MS spectrum corresponding to 5.2-6.1 min yielded $[M-H]^-$ at m/z 226.7, corresponding to a molecular weight of 228. Its structure has not yet been identified by MS^n. Another two peaks yielded $[M-H]^-$ at m/z 172.9 and 200.9, corresponding to molecular weights of 172 and 200. The fragmentation of m/z 172.9 and 200.9 both show similar patterns to m/z 186.9 (MS^n data not shown), therefore their structures are determined to be as follows:

For the **172.9** peak: $HO-CH_2-(CH_2)_7-COOH$



For the **200.9** peak: $HO-CH_2-(CH_2)_9-COOH$



Because the NMR spectrum (Figures 5.39-41, aliphatic area not shown) did show that aromatic compound **2** contains some aliphatic chains, the existence of those monomers is reasonable.

5.6 Aromatic compound 3

Section 4.5.1 explains the separation of aromatic compound **3** and shows its provisional structure in Figure 4.37. It also shows that after HPLC separation two fractions were obtained that incorporate this provisional structure. They are aromatic compounds **3-f3** and **3-f6** (Figure 4.39), both esters with the same 1,2 substituted benzene ring from the acid moiety, but with different aliphatics on the alcohol portion.

5.6.1 Aromatic compound **3-f6**

5.6.1.1 NMR of aromatic compound **3-f6**

The NMR spectra (^{1}H, Figure 5.53; HMBC, Figure 5.54) show the same peaks and correlations in the aromatic region, which confirms the 1,2-disubstituted benzene ring. The gHMBC spectrum shows that protons at 7.67 and 7.70 ppm correlate to a carbon with chemical shift of 167.7 ppm, which confirms that two carbonyl carbons (ester bond) connect to positions 1 and 2.

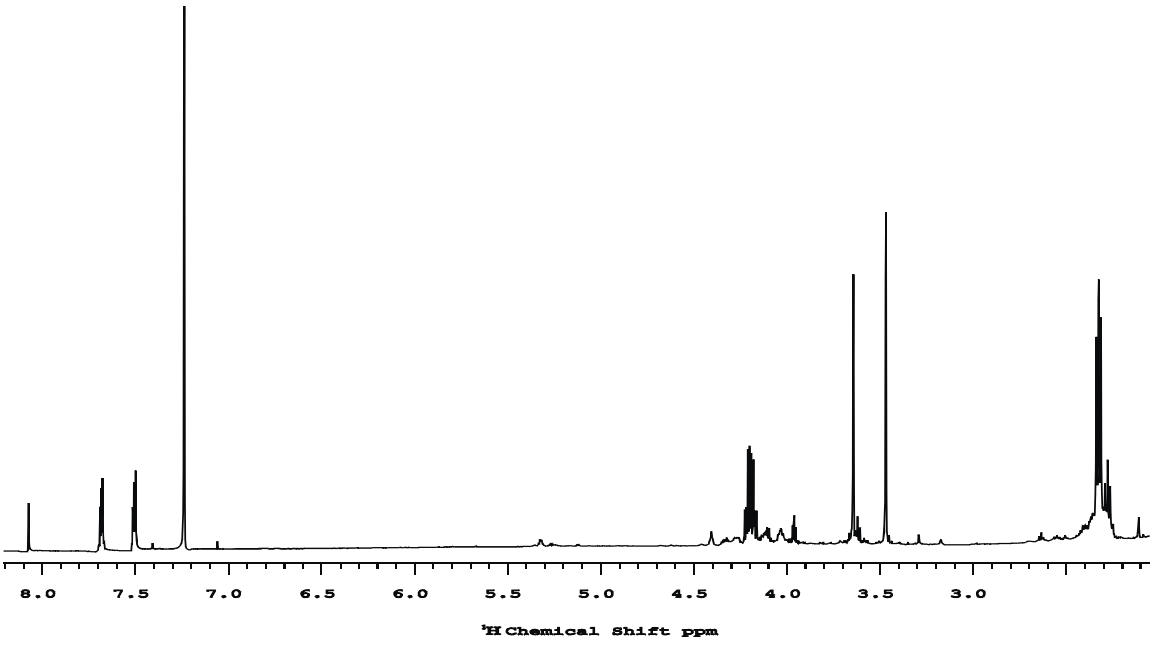


Figure 5.53. 600 MHz ^{1}H NMR spectrum of aromatic compound 3-f6.

The gHSQC spectrum (not shown) shows that the two protons at 4.17 and 4.22 ppm correspond to the same carbon at 69.9 ppm. The COSY spectrum (not shown) shows that both of the protons at 4.17 and 4.22 ppm correlate with the proton at 1.67 ppm. The gHMBC spectrum (Figure 5.54) shows that corresponding to carbon 167.7 ppm, there are 2 cross-peaks at (4.17, 167.7 ppm) and (4.21, 167.7 ppm); the proton at 4.17 ppm corresponds to cross-peaks at (4.17, 38.4 ppm), (4.17, 30.0 ppm), (4.17, 23.7 ppm); the proton at 4.21 ppm corresponds to cross-peaks at (4.21, 38.4 ppm), (4.21, 30.0 ppm), and (4.21 ppm, 23.7 ppm) (see insert to Figure 5.54). Together, this evidence indicates on the alcohol side of the ester bond, the second carbon is a methine carbon (see Section 4.5.1).

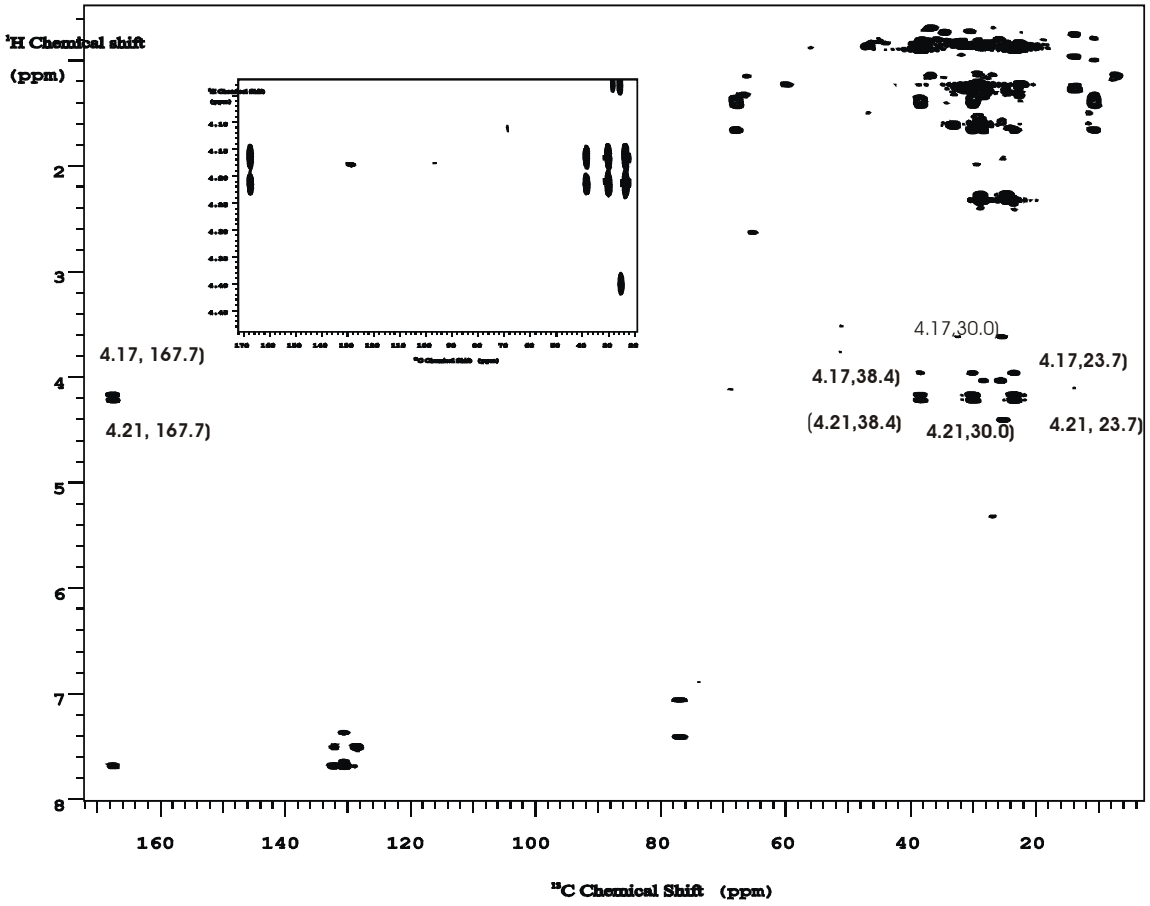


Figure 5.54. gHMBC spectrum of aromatic compound 3-f6. The insert highlights connectivities from the alcohol side of the ester bond to a methine carbon.

Table 5.14 summarizes correlations from the gHMBC spectrum (Figures 5.54, 5.55), confirming the partial structure of the alcohol moiety of the compound **3-f6**. Figure 5.56 (a) shows the provisional structure of compound **3-f6**. ACD simulations of the ^{1}H and ^{13}C NMR spectra (Figures 5.57, 5.58) match the tentative structure very well. LC/MS^{n} is needed for final confirmation of its structure.

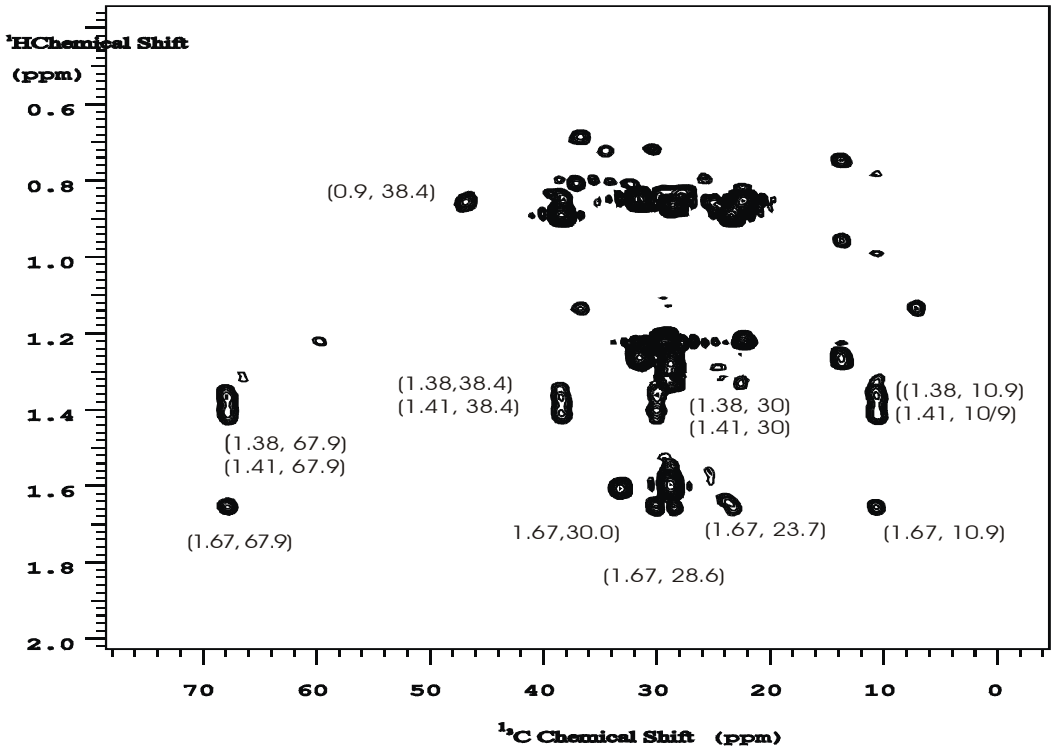


Figure 5.55. A portion of the gHMBC spectrum of aromatic compound 3-f6. Unidentified cross peaks are derived from compounds that co-elute with 3-f6.

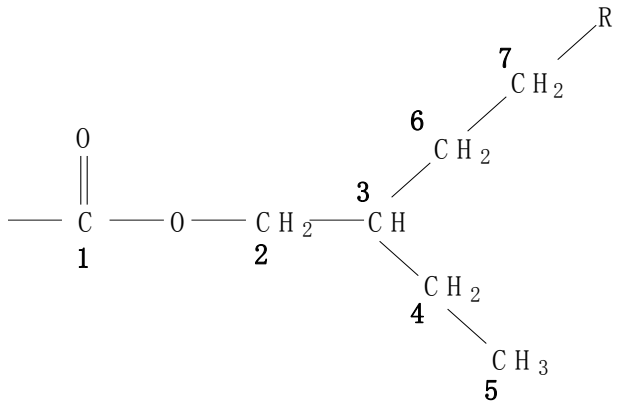
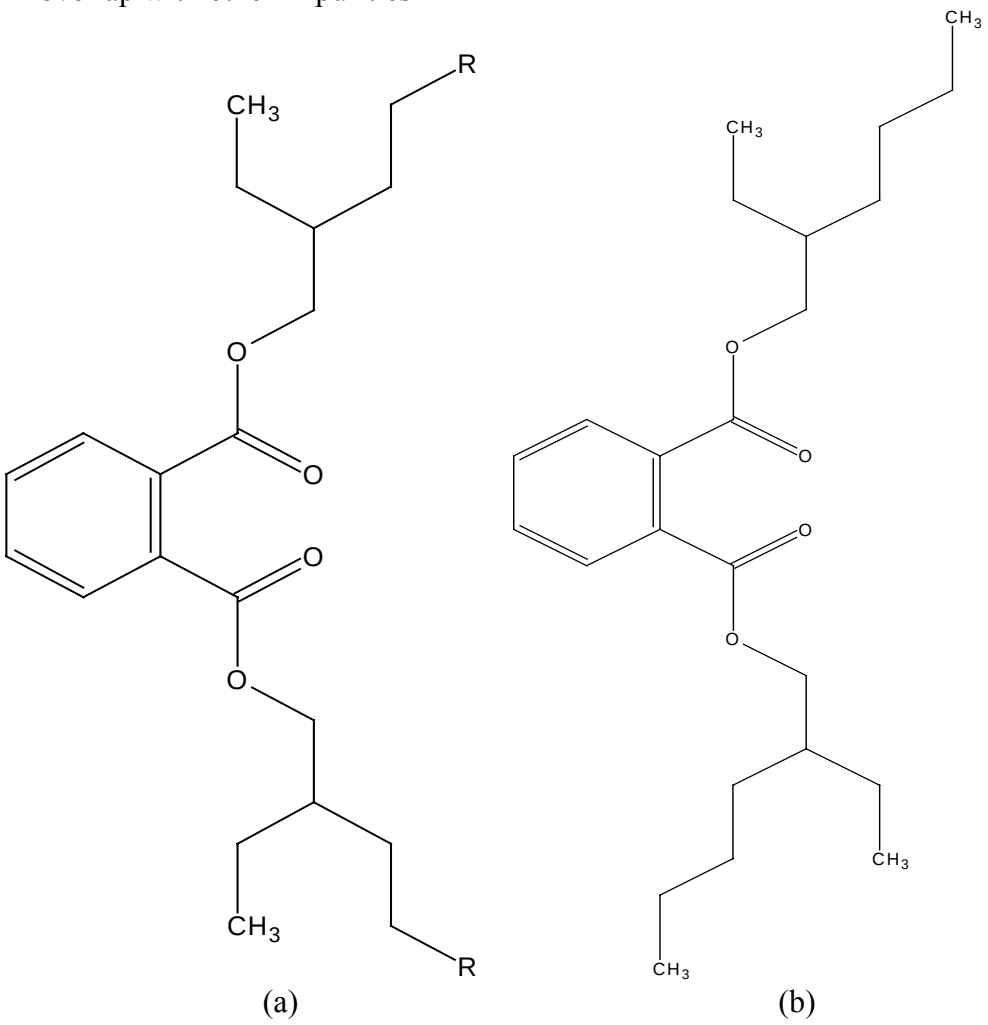


Table 5.14 gHMBC correlations for aromatic compound 3-f6 (Figures 5.54 and 5.55)

H \ C (ppm)	**1** (167.4)	**2** (67.9)	**3** (38.4)	**4** (23.7)	**5** (10.9)	**6** (30.0)	**7** (28.6)
2 (4.22, 4.17)	* *		* *	* *		* *	
3 (1.67)		*		*	*	*	*
4 (1.41, 1.38)		* *	* *		* *	* *	
5 (0.90)			*	*			
6 (uncertain**)							
7(uncertain**))							

* cross-peaks

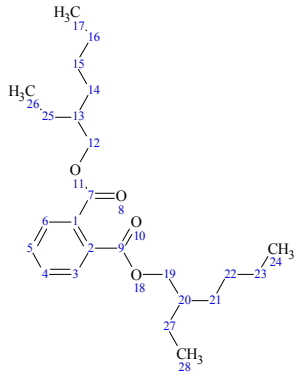
** overlap with other impurities



(a) (b)

Figure 5.56. Provisional structures of aromatic compound 3-f6. a: structure deduced from NMR; b: structure deduced from MS, including a monoisotopic mass of 390.277 D.

25/Oct/2006 13:52:41 ACD/HNMR Predictor (v.9.02)



Group	nH	Shift	Error	Av.Exp
3	1	**7.72**	0.50	7.75..8.20
4	1	**7.33**	0.38	7.42..7.82
5	1	**7.33**	0.38	7.42..7.82
6	1	**7.72**	0.50	7.75..8.20
12<''>	1	**4.19**	0.39	4.04..4.12
12<'>	1	**4.01**	0.39	4.04..4.12
13	1	**1.50**	0.12	1.62
14<''>	1	**1.43**	0.34	1.27..1.48
14<'>	1	**1.35**	0.34	1.14..1.46
15	2	**1.26**	0.28	1.12..1.37
16<''>	1	**1.33**	0.19	1.16..1.38
16<'>	1	**1.25**	0.19	1.14..1.30
17	3	**0.89**	0.03	0.89
19<''>	1	**4.19**	0.39	4.04..4.12
19<'>	1	**4.01**	0.39	4.04..4.12
20	1	**1.50**	0.12	1.62
21<''>	1	**1.43**	0.34	1.27..1.48
21<'>	1	**1.35**	0.34	1.14..1.46
22	2	**1.26**	0.28	1.12..1.37
23<''>	1	**1.33**	0.19	1.16..1.38
23<'>	1	**1.25**	0.19	1.14..1.30
24	3	**0.89**	0.03	0.89
25<''>	1	**1.50**	0.46	1.35
25<'>	1	**1.30**	0.46	1.35
26	3	**0.93**	0.06	0.82..0.90
27<''>	1	**1.50**	0.46	1.35
27<'>	1	**1.30**	0.46	1.35
28	3	**0.93**	0.06	0.82..0.90

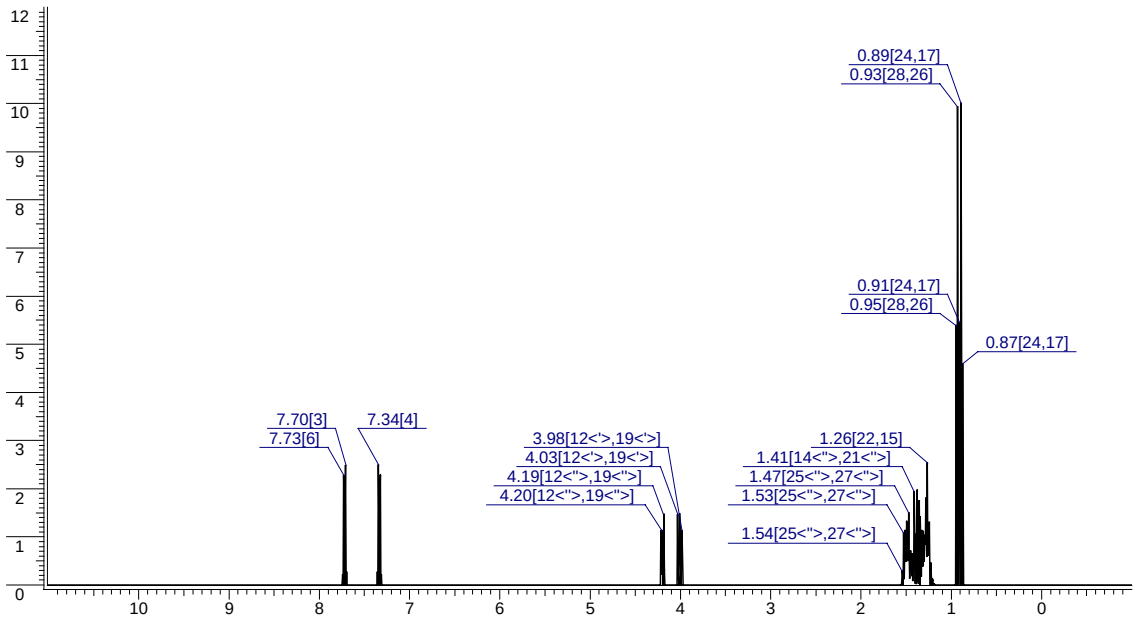
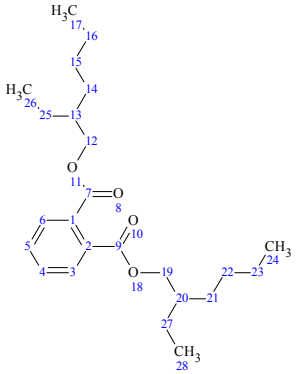


Figure 5.57. ACD calculation of proton spectrum for the compound shown in Figure 5.56 (b).



25/Oct/2006 13:55:36 ACD/CNMR (v.9.02)

Carbon No.	CHn	Chem. Shifts	Conf. Limits	Av.Exp
1	C	**132.10**	1.90	132.00..132.70
2	C	**132.10**	1.90	132.00..132.70
3	CH	**129.13**	0.79	129.79..131.21
4	CH	**131.04**	1.71	128.60..130.67
5	CH	**131.04**	1.71	128.60..130.67
6	CH	**129.13**	0.79	129.79..131.21
7	C	**168.77**	2.64	168.28..168.32
9	C	**168.77**	2.64	168.28..168.32
12	CH_2	**68.81**	2.71	68.39..68.41
13	CH	**39.37**	2.94	38.70
14	CH_2	**30.22**	2.16	29.36..30.40
15	CH_2	**28.29**	1.24	27.99..28.93
16	CH_2	**22.32**	1.13	22.15..22.98
17	CH_3	**13.82**	0.40	13.60..14.04
19	CH_2	**68.81**	2.71	68.39..68.41
20	CH	**39.37**	2.94	38.70
21	CH_2	**30.22**	2.16	29.36..30.40
22	CH_2	**28.29**	1.24	27.99..28.93
23	CH_2	**22.32**	1.13	22.15..22.98
24	CH_3	**13.82**	0.40	13.60..14.04
25	CH_2	**23.60**	1.68	22.75..23.76
26	CH_3	**11.25**	0.63	10.56..10.96
27	CH_2	**23.60**	1.68	22.75..23.76
28	CH_3	**11.25**	0.63	10.56..10.96

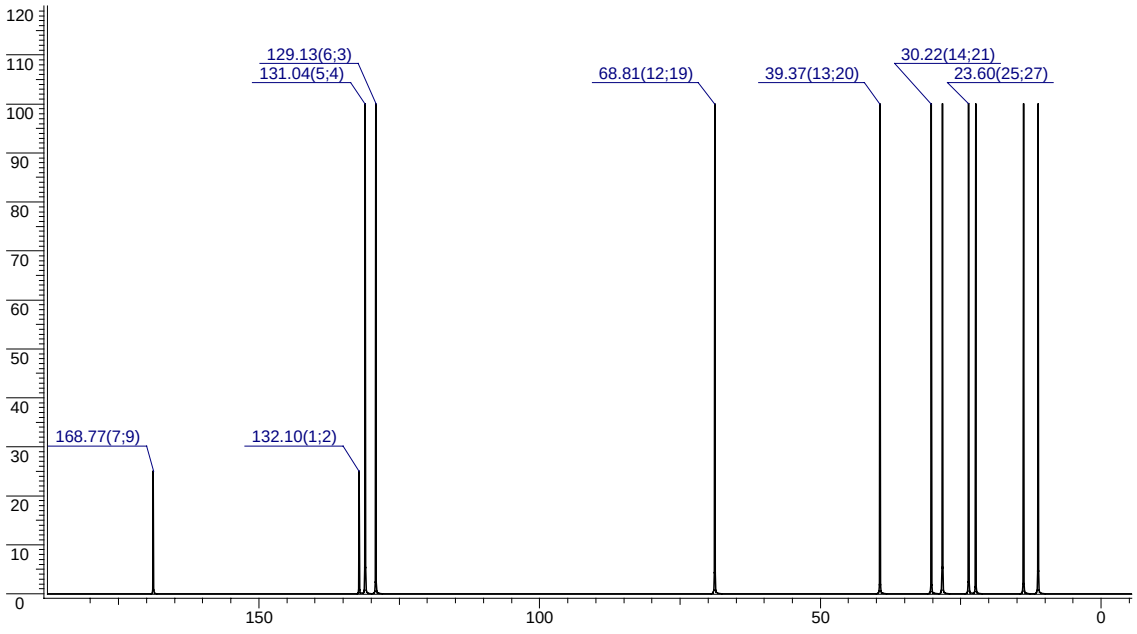


Figure 5.58. ACD calculation of carbon spectrum for the compound shown in Figure 5.56 (b).

5.6.1.2 LC/MS and LC/MS^n of aromatic compound **3-f6**

Because other compounds coexist with aromatic compound **3-f6**, some of the correlations in 2D NMR cannot be identified completely. Therefore, MS^n data are needed to work out the structure. A reverse phase C8 column and gradient elution were used to get better separation. The gradient used was from 50 to 100% methanol (0.1% formic acid) in 10 min, then holding for 5 min. Figure 5.59 shows the LC-APCI-MS positive mode results. The top graph shows LC separation with UV detection at 280 nm, and the second graph presents the total ion current (TIC) in positive mode MS. The third graph shows MS spectra corresponding to LC elution (or TIC chromatogram) at 7.0-7.4 min.

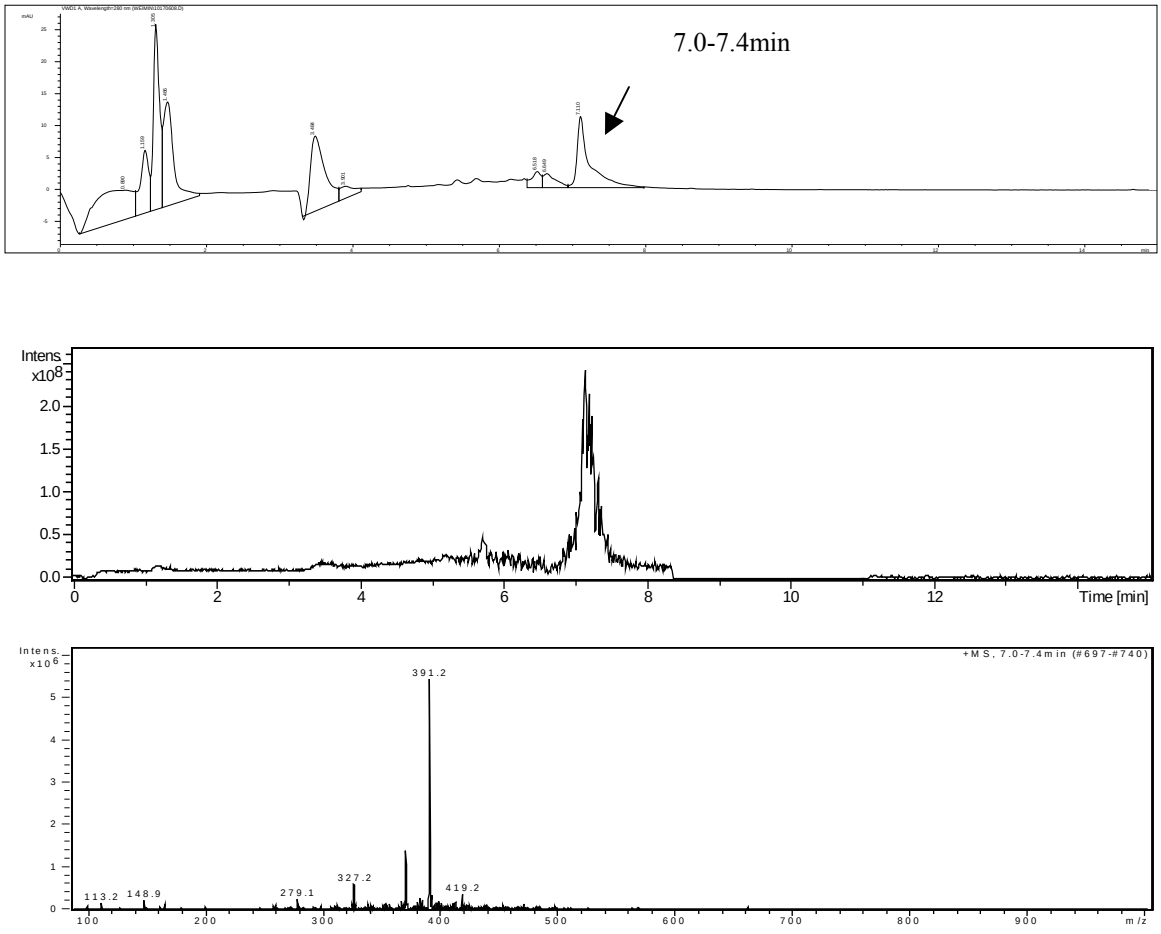


Figure 5.59. APCI-LC-MSD ion trap UV and MS spectrum of aromatic compound 3-f6. The retention times in the top trace span the range 0-15 min.

For the 7.0-7.4 min TIC peak (Figure 5.59, second trace) that yielded $[M+H]^+$ at m/z 391.2 (corresponding to molecular weight 390), the HPLC chromatogram (Figure 5.59, top) shows UV absorption at 280 nm, indicating it might be aromatic compound **3-f6**. The isotopic peak (molecular ion +1) had intensity about 26% of the original peak, implying about 24 carbons. Thus MS^n data are needed to confirm its final structure.

Figure 5.60 shows the MS^2 spectrum of $[M+H]^+$ at m/z 391.2, which produced fragment ions at m/z 279.0, 260.9, 166.9, 148.9 and 113.1. Further fragmentation MS^3 of 279.0 gave fragment ions at m/z 166.8 and 148.9. Finally, further fragmentation MS^4 of 166.9 gave fragment ions at m/z 148.8.

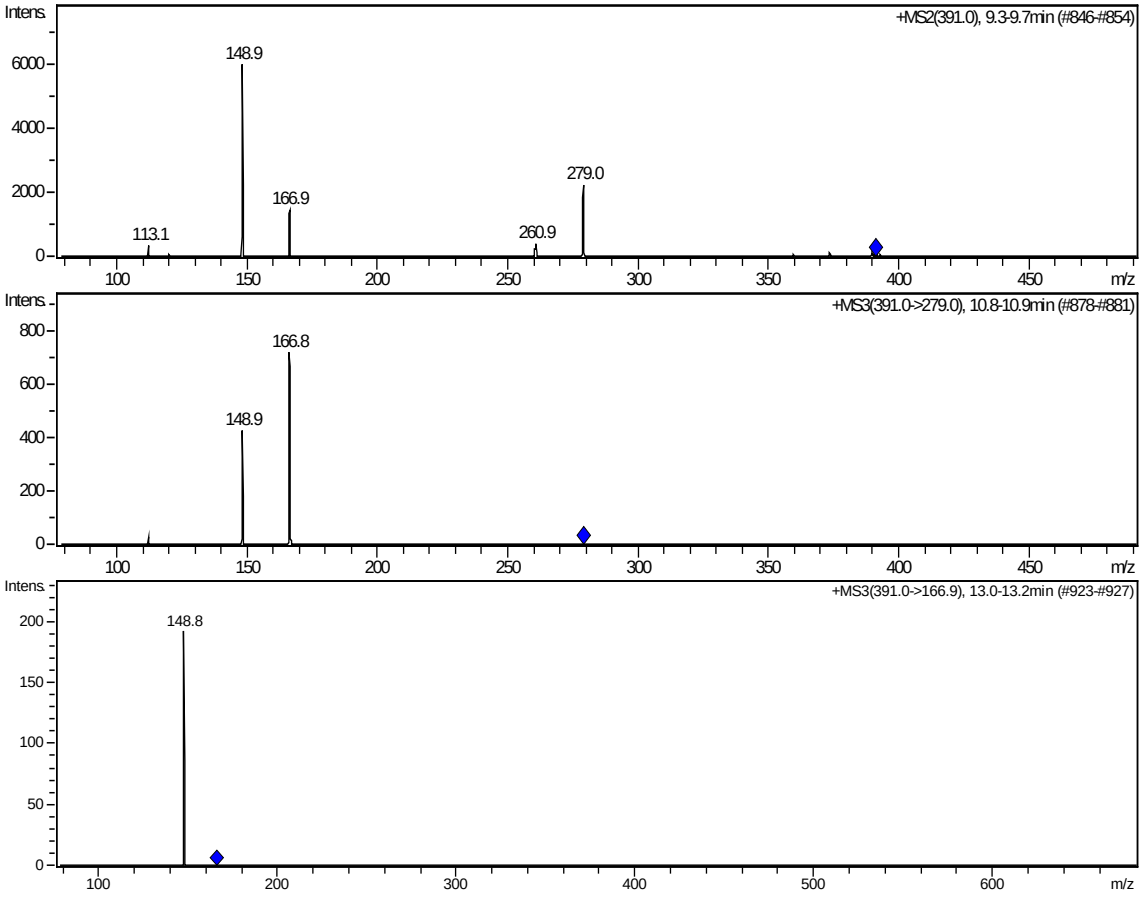


Figure 5.60. APCI-LC-MSD ion trap MS^n spectrum of m/z at 391.2.

Thus for the molecular weight of 390 determined by MS, the possible structure of aromatic compound **3-f6** that corresponds to the provisional structure in Figure 5.56 (a) is actually the structure shown in Figure 5.56 (b). Figure 5.61 shows the fragmentation pathway of aromatic compound **3-f6**.

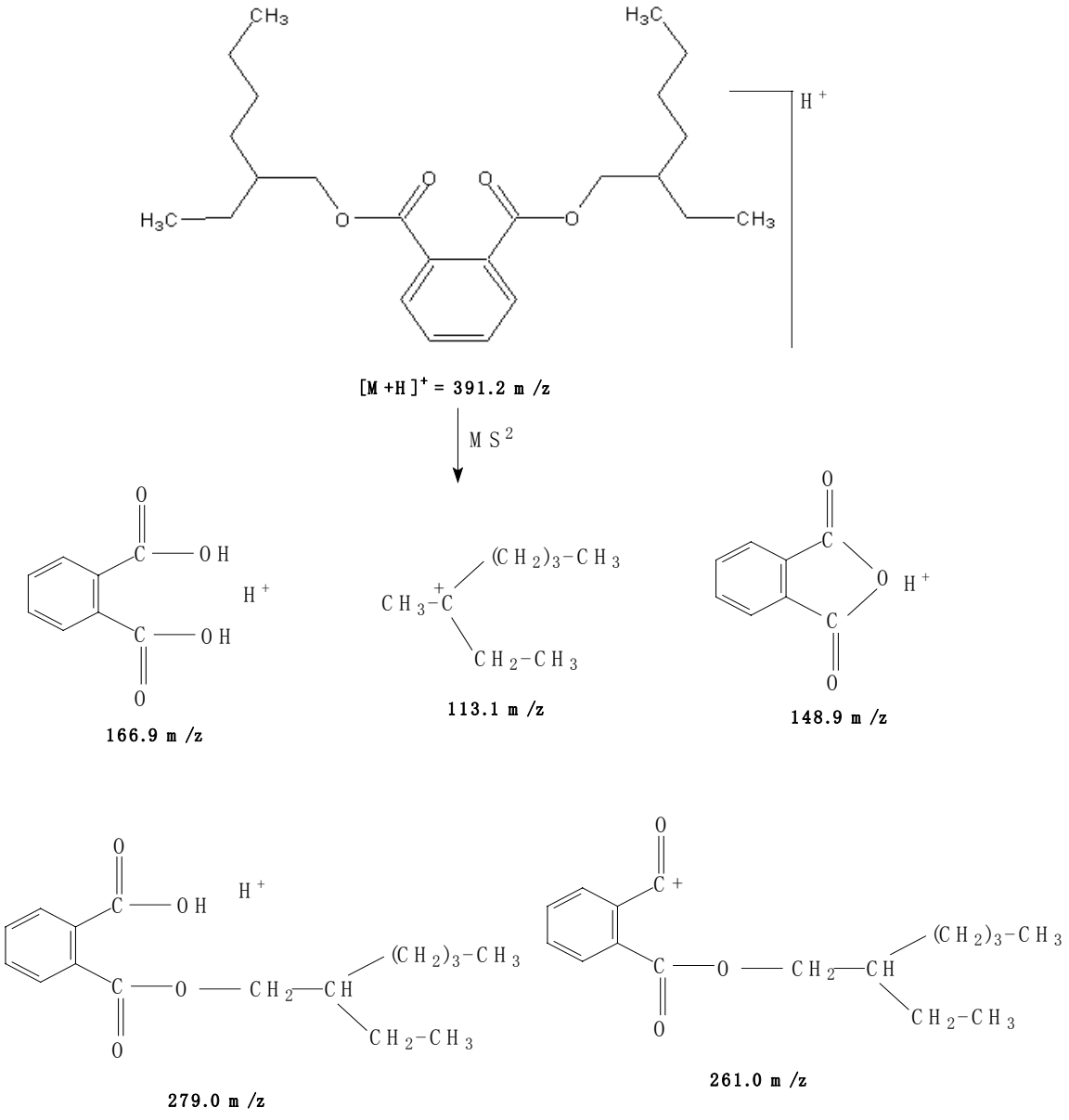


Figure 5.61. APCI-LC-MSD ion trap MS^n fragmentation of the molecular ion at m/z 391.2.

This compound was found might coming from plastic containers or solvents instead of plant material

5.6.2 Aromatic compound **3-f3**

5.6.2.1 NMR of aromatic compound **3-f3**

Figure 5.62 shows ^{1}H NMR spectrum of aromatic compound 3-f3. NMR data also confirm the 1,2-disubstituted benzene ring.

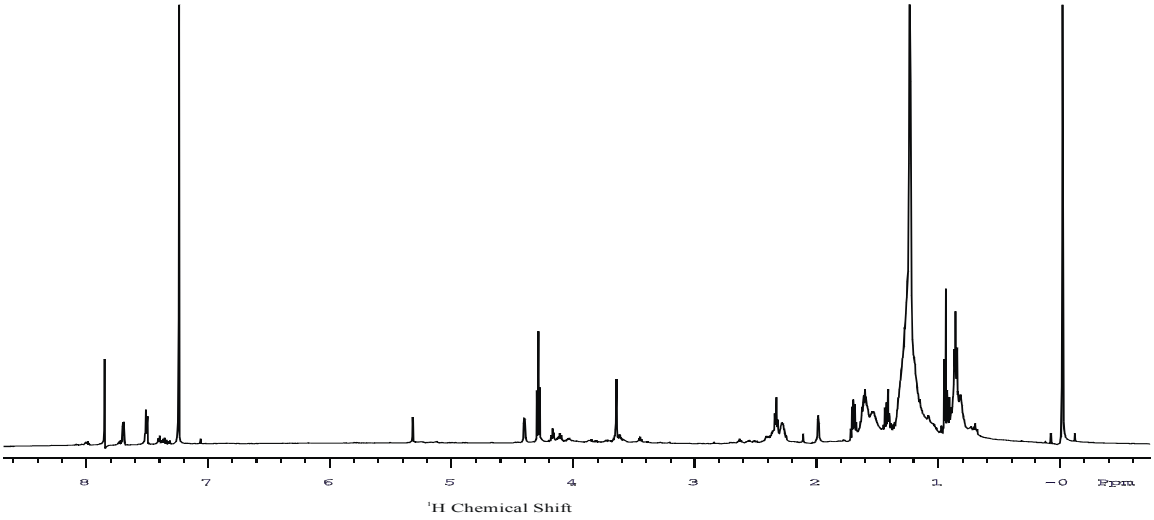


Figure 5.62. 600 MHz ^{1}H NMR spectrum of aromatic compound 3-f3.

5.6.2.2 LC/MS and LC/MS^{n} of aromatic compound **3-f6**

Satisfactory MS data of aromatic compound **3-f6** could not be obtained.

5.7 Aromatic compounds 4

Aromatic compounds **4** were separated by solvent extraction from the soluble products of suberin depolymerization. Figure 5.63 shows the proton spectrum of aromatic compounds **4**, including three groups of doublet peaks between 6 and 8 ppm. They are sorted as three groups because they are correlated pairwise in the COSY spectrum (not shown). The chemical shifts of each proton and corresponding carbon (Figure 5.64, gHSQC) are:

group 1 (aromatic compound **4-1**): peak A (7.21, 129.3 ppm), peak B (6.82,116.5 ppm)

group 2 (aromatic compound **4-2**): peak C (7.09, 130.5 ppm), peak D (6.77, 116.5 ppm)

group 3 (aromatic compound **4-3**): peak M (7.24, 129.3 ppm), peak N (6.80, 116.5 ppm)

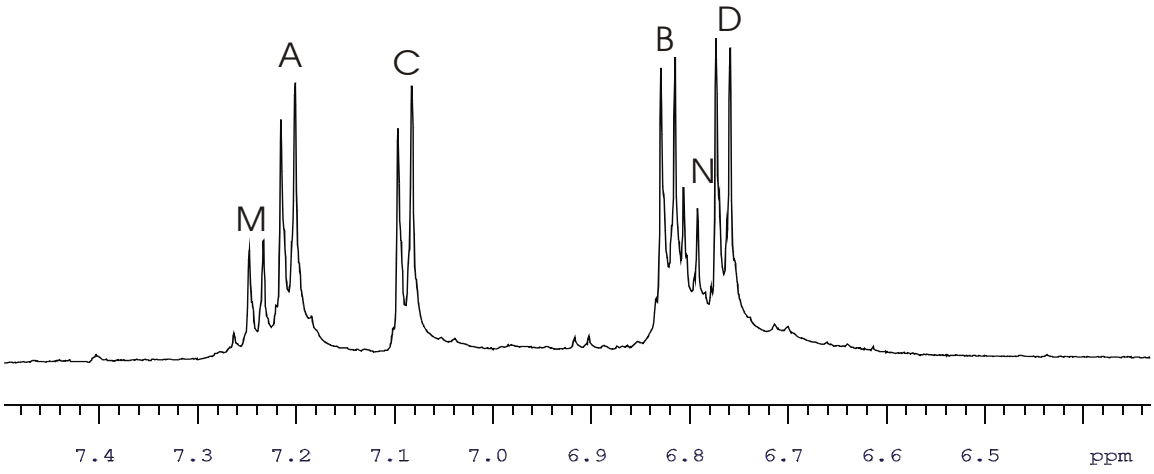


Figure 5.63. 600 MHz ^{1}H NMR spectrum of aromatic compounds 4.

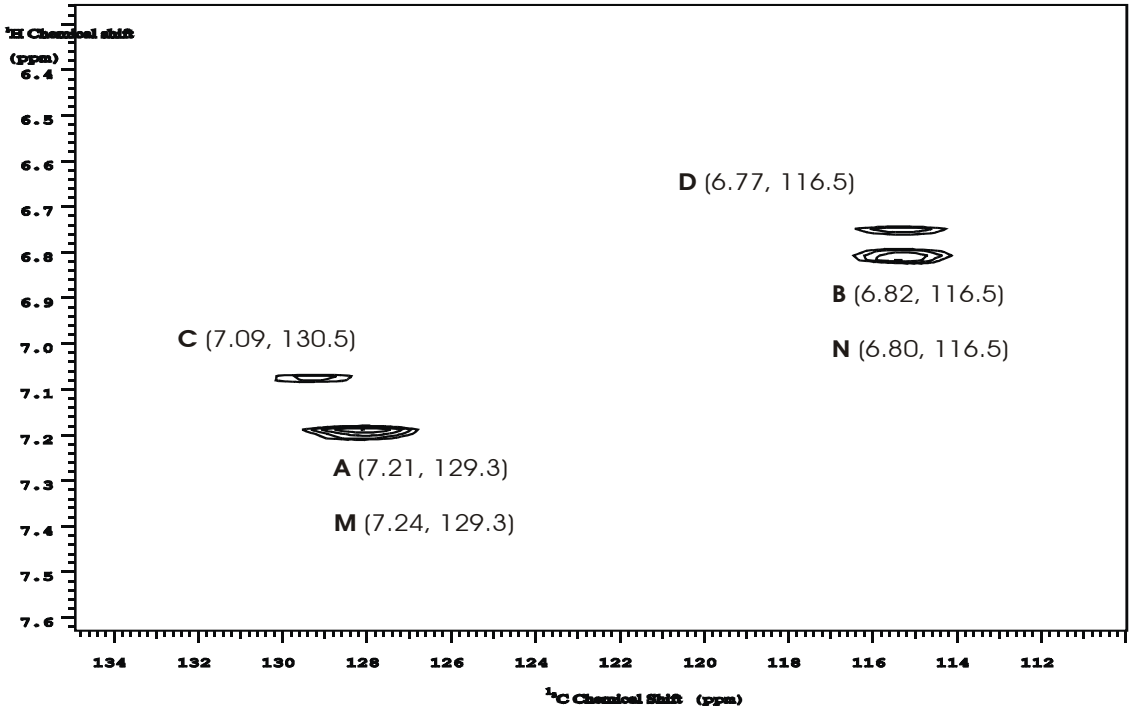


Figure 5.64. gHSQC spectrum of aromatic compounds 4.

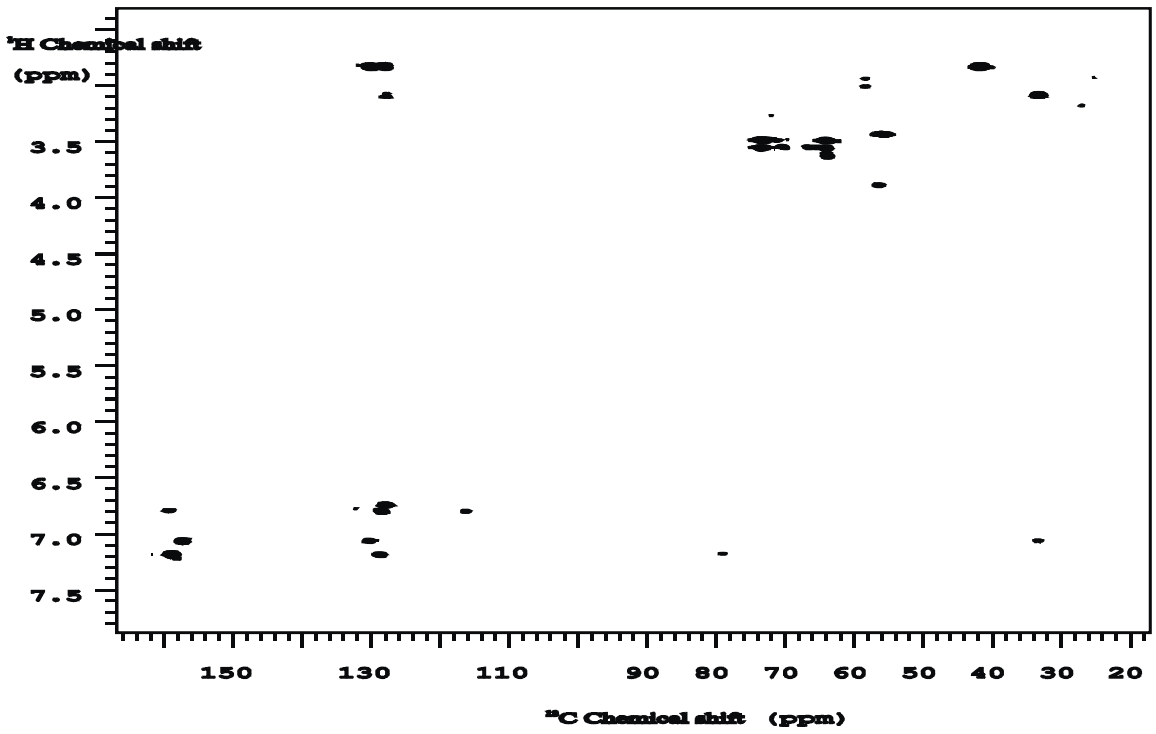


Figure 5.65. gHMBC spectrum of aromatic compounds 4.

In the COSY spectrum, peak A only correlated with B; C only correlated with C; M only correlated with N. The proton spectrum also shows they are all doublet peaks with the same J-coupling constant of 9.09 Hz. By combining all above information, these three groups of peaks represent three aromatic compounds that have a common partial structure:

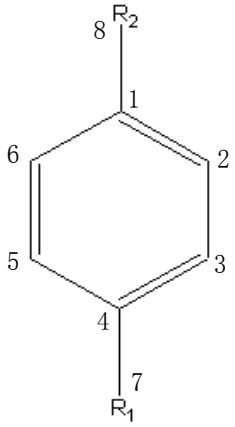


Table 5.15 gHMBC (Figure 5.65) correlations of aromatic compounds in group 4-1

H\C (ppm)	1??? (159.9 or 129.5)	2 (129.5)	3 (116.7)	4 (159.4)	5 (116.7)	6 (129.5)
2 (7.21)				*		*
3 (6.82)	*?	*?			*	
5(6.82)	*?		*			*
6 (7.21)		*		*		

(*cross-peak)

Because the carbon chemical shift for position 4 is 159.4 ppm, the atom at position 7 is likely to be oxygen. Moreover this compound is soluble in water, so R_1 should be a hydroxy (-OH) group. The structure needs further confirmation (see below). For compound **4-3**, its NMR spectrum is almost overlapping with compound **4-1**; they should have a very similar structure.

Table 5.16 HMBC (Figure 5.65) correlations of aromatic compounds in group 4-2

H\C (ppm)	1 (128.4)	2 (130.5)	3 (116.5)	4 (157.9)	5 (116.5)	6 (130.5)
2 (7.09)				*		*
3 (6.77)	*				*	
5 (6.77)	*		*			*
6 (7.09)		*		*		

(* cross-peak)

Compound **4-2** (Table 5.16) has a similar pattern of correlations as group 1. Again the carbon chemical shift at position 4 is 157.9 ppm so the atom at position 7 is oxygen. This compound is soluble in water, so R_1 should again be a hydroxy (-OH) group. However, there are some different correlations of this compound as compared with group 1.

The carbon at position 1 has correlations with two protons shown by COSY cross-peaks at (3.12, 128.4 ppm) and (2.85, 128.4 ppm), while the protons at position 2 and 6 have a correlation at (2.85, 130.9 ppm). The proton at 2.85 ppm also has a correlation

with the carbon at 42.3 ppm, and the proton at 3.12 ppm has correlation at 34.1 ppm. All of the above information is consistent with the partial structure of compound **4-2** shown in Figure 5.66. ACD calculations (not shown) match the structure very well.

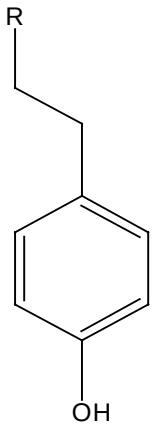


Figure 5.66. Partial structure of aromatic compound 4-2.

The final structures should be confirmed by MS and MS^n data, which are challenging to obtain and interpret for this mixture.

5.8 Aliphatic compound 4: sample 12-27-#6

This material was separated by solvent extraction of the soluble part from depolymerization products of potato suberin (see solvent extraction strategy in Chapter 4). Figure 5.67 shows the ^1H spectrum of aliphatic compound 4, in which 4.05 ppm is the characteristic peak for an ester bond; 5.32 ppm is the characteristic peak for a double bond (-CH=CH-) in the middle of the chain. 2D gHMBC and gHSQC (not shown) confirm the presence of oligomers. So this aliphatic oligomer may have a double bond in the middle of the aliphatic chain, distinguishing it from aliphatic compound **3** (Section 5.3).

The final structure should be confirmed by MS or MS^n data, which are forthcoming.

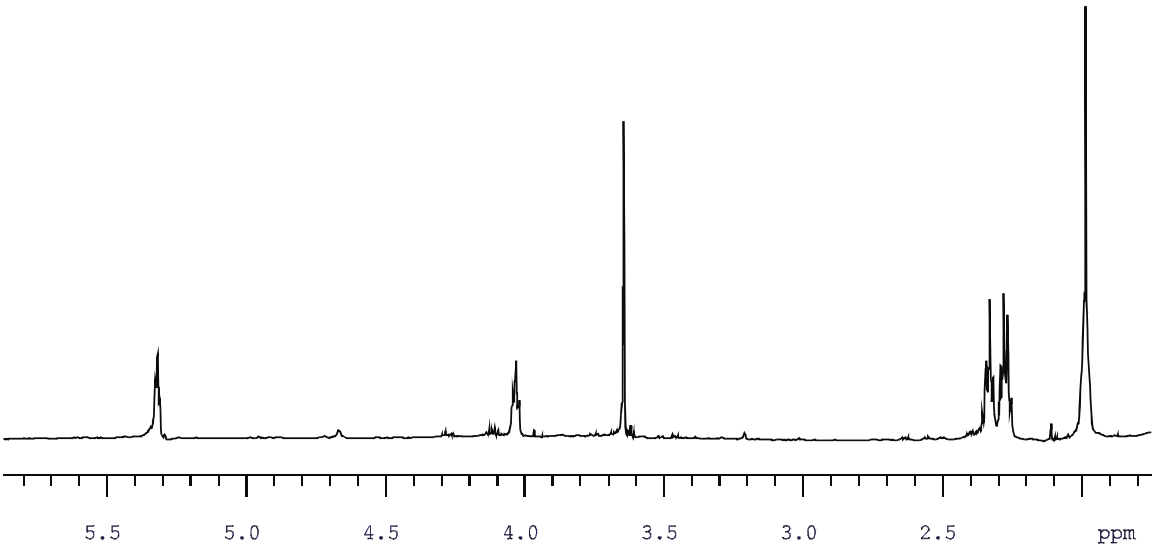


Figure 5.67. 600 MHz ^{1}H spectrum of aliphatic compound 4.

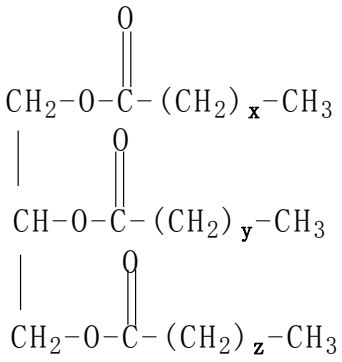
5.9 Summary

Isolation and identification of molecular constituents from the protective plant polymer suberin was accomplished by preparing a potato suberin sample, depolymerizing the potato suberin, and combining different analytical separation and identification techniques: CC, TLC, HPLC, NMR, LC/MS^n. It should be noted that the strategy of combining 2D NMR analysis of through bond connectivities and MS^n fragments allows identification of some compounds that are challenging to isolate in pure form. Several new aliphatic and aromatic compounds were identified from potato suberin depolymerization products, enriching our knowledge of the molecular structure and assembly of the biopolymer suberin.

These findings, along with others published previously, comprise a database of compounds separated from suberin along with their 1D and 2D NMR spectra, HPLC

separation conditions, molecular weight, and MS/MS fragmentation data. In addition to building our knowledge of this protective plant polymer, this database can guide the separation and elucidation efforts of future researchers.

a. Aliphatic compound **2**: family of triglycerides

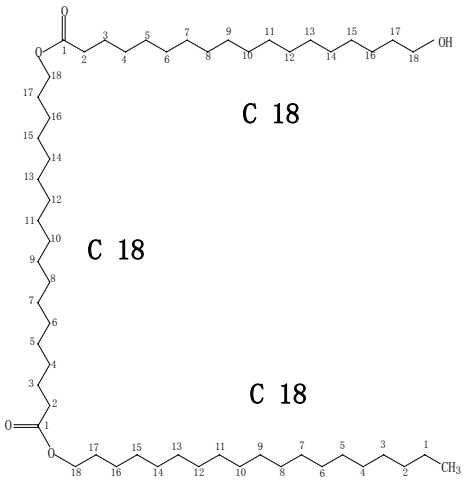


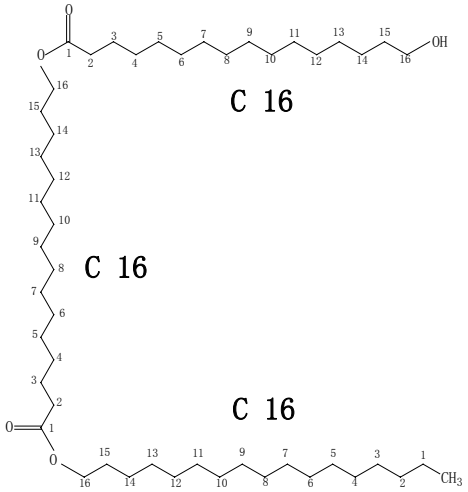
The triglycerides separated from potato suberin have a common skeleton with x, y, z determined as follows:

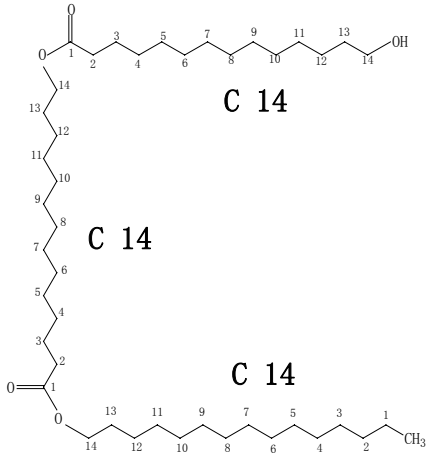
Molecular weight	**x**	**y**	**z**
470.4 *	6	6	6
498.4 *	6	6	8
544.4 *	6	8	8
582.6	6	10	10
638.7	10 8 8	10 10 8	10 12 14
666.7	10 8	10 10	12 14
722.7	8	14	14
750.8	12 10	12 14	14 14
764.8	12	13	14
778.8	12	14	14
792.8	13	14	14
806.8	14	14	14
834.6	14	14	16
890.7	16	16	16

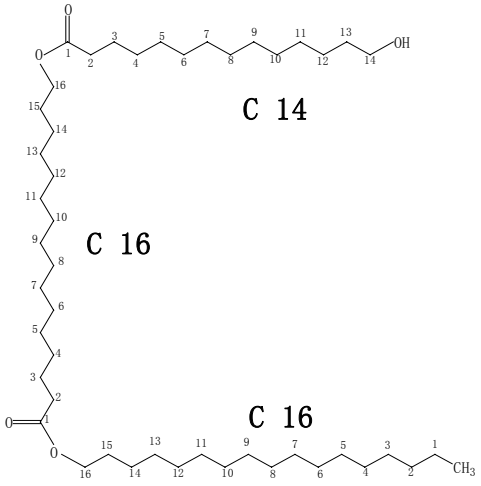
* main compound in a mixture

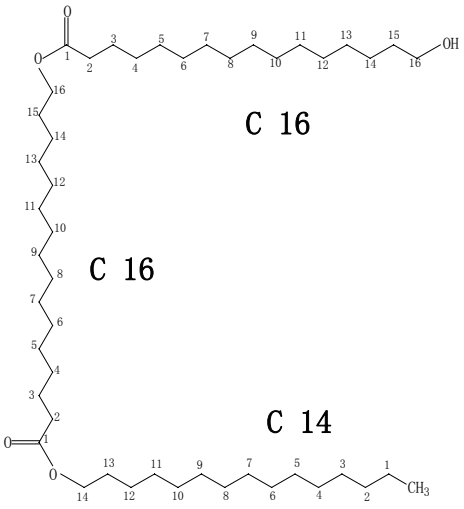
b. aliphatic compounds **3:** family of linear ester trimers (partial listing)

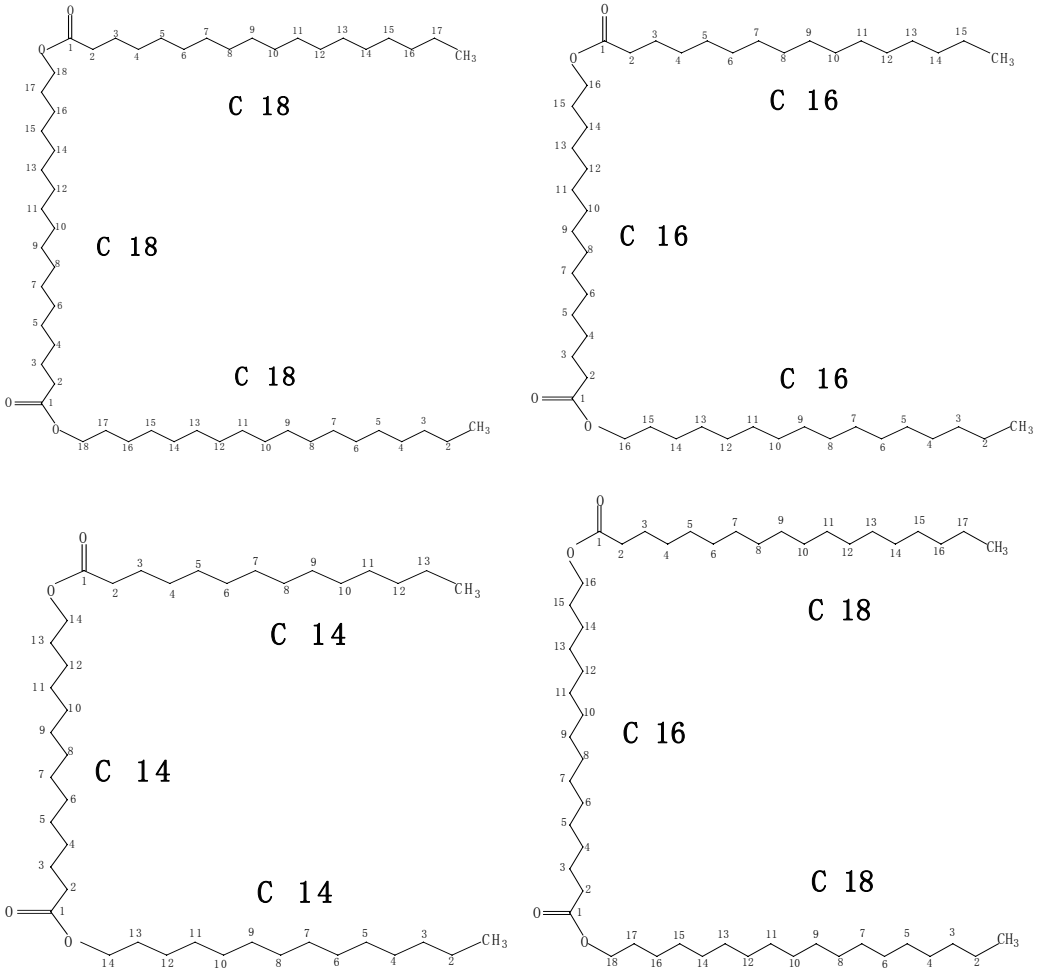








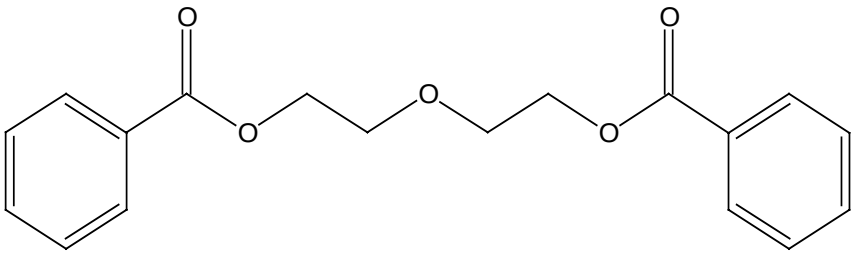




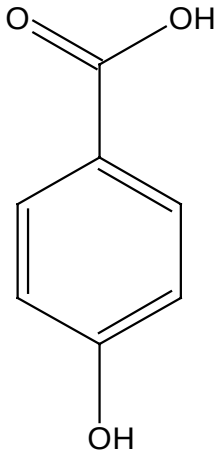
c. aliphatic compounds **4:**

They are similar oligomers to **aliphatic compounds 3,** the only difference is **aliphatic compounds 4** have double bonds in the middle of the chain. MS/MS data are needed to finally confirm their structures.

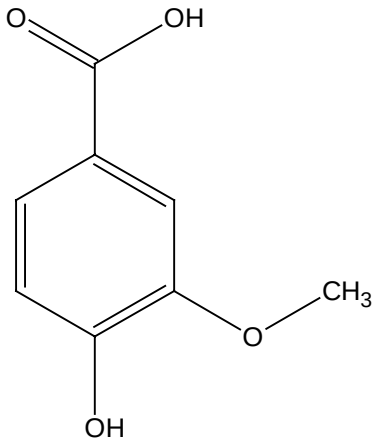
d. aromatic compound **1** (#11-TLC-4-1)



e. aromatic compound **2-1**



aromatic compound **2-2**



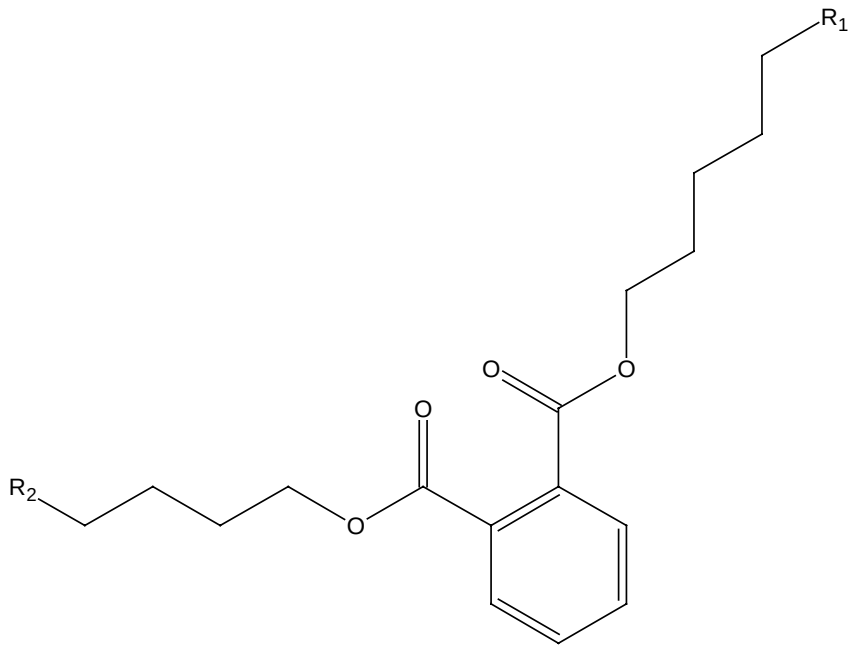
e'.co-existing aliphatic monomers with aromatic compounds **2-1** and **2-2**

HO-CH_2-$(CH_2)_7$-COOH

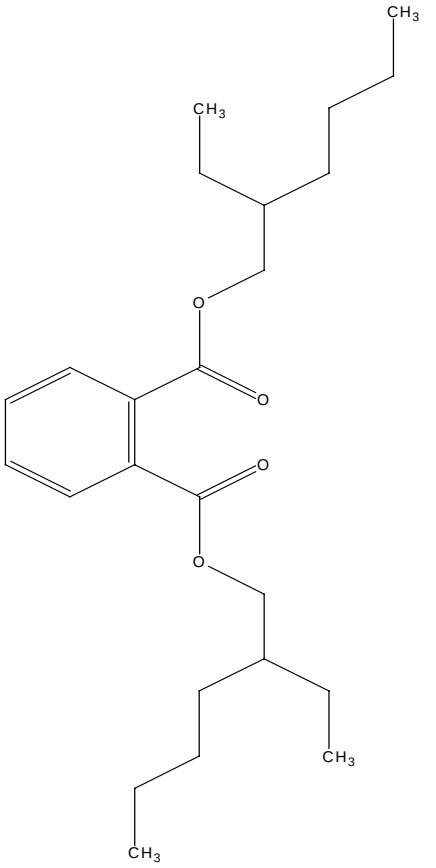
HO-CH_2-$(CH_2)_8$-COOH

HO-CH_2-$(CH_2)_9$-COOH

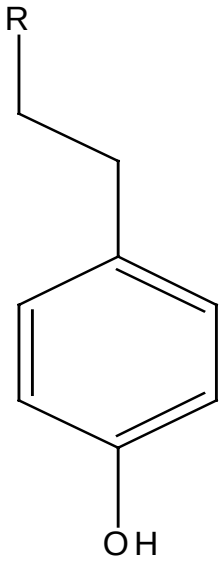
f. Aromatic compound 3-f3 (MS/MS data needed to confirm final structure**)**



g. aromatic compound 3-f6 (it is from plastic containers or solvents)



h. aromatic compound **4-2** partial structure



Chapter 6. CONCLUSIONS AND PROSPECTUS

Information on the molecular structure of the protective plant polymer suberin has historically been obtained by either solid-state NMR of intact suberin, or from elucidation of its depolymerization products. Information on aliphatic and aromatic monomers obtained by depolymerization has provided us with basic knowledge of the unique molecular structure of suberin. However, such information is insufficient for understanding its biosynthesis and macromolecular assembly. From this point of view, the polymerization steps of suberin biosynthesis are still unknown, and the nature of the covalent connectivities within suberin and to the cell wall is quite speculative. To better understand this ambiguous area and to extend the current SPPD (suberin polyphenolic domain) and SPAD (suberin polyaliphatic domain) model of suberin architecture, the isolation of aliphatic and aromatic oligomers and new monomers is critical to map out the essential crosslinking structure and polyester-cell wall linkages.

For such structural investigations, control of the depolymerization process is very important to generate soluble compounds that retain as much as the original architectural information as possible. By carefully adjusting the KOH-methanol depolymerization method, a product mixture of desired oligomers and new monomers was achieved. An analytical separation and identification protocol combined multi-step extraction with different solvents using ^{1}H NMR to guide the separation process, column chromatography, reverse phase and normal phase HPLC, and finally 1D (^{1}H, ^{13}C) and 2D NMR (COSY, HMQC, HMBC), LC/MS and LC/MS^{n}. It should be noted that combining 2D NMR analysis of through-bond connectivities and MS^{n} fragments allows

identification of some compounds that are challenging to isolate in pure form. Using this separation and structural elucidation protocol, twenty different triglycerides, three new aliphatic monomers, more than twenty aliphatic trimers, one aromatic dimer, and two new aromatic monomers were isolated from chemical degradation of suberin.

The depolymerization results show that the overall yield of aliphatic oligomers and aromatics is insensitive to KOH concentration. Whereas the relative amount of aliphatic oligomers decreases with depolymerization time, the relative number of aromatic molecules increases with depolymerization time. This result is reasonable because longer treatment time may further break down aliphatic oligomer products, and longer treatment time may also help to make the internal SPPD moieties accessible and break down more aromatics.

The identification of glycerol (**aliphatic compound 1**) and twenty different triglycerides (**aliphatic compounds 2**) provides strong supportive evidence for the current suberin conceptual model [33](Chapter 1), in which the SPPD and SPAD are depicted to be either covalently linked through glycerol [45] or directly linked between acyl groups. No direct connectivity between glycerol and the SPPD was detected, but the existence of different chain length triglycerides confirmed connectivities between glycerol and long-chain aliphatic monomers (SPAD), showing that suberin is a glycerol-derived lipid. From the molecular assembly point of view, the triglycerides play an important role in the assembly of the suberin polymer.

In the suberin conceptual model, the triglycerides (glycerol-aliphatic polyester structure of suberin) are assumed to be associated with polymeric aromatic components within suberized cell walls. Our knowledge of the aromatic domain (SPPD) now includes

the newly separated aromatic monomers (**aromatic compounds 2-1 and 2-2**), in which both hydroxyl and carboxylic functional group are all possible covalent linkage point with aliphatic ω-hydroxyacids. In addition, their architectures are revealed by the partially identified aromatic monomers (**aromatic compounds 4**), and one new aromatic dimer (**aromatic compound 1**). No direct evidence was obtained for covalent linkages between aliphatic and aromatic components or to cell-wall polysaccharides; such structural elements could go undetected because they are resistant to degradation with strong base.

Twenty provisionally identified aliphatic trimers (**aliphatic compounds 3,** saturated aliphatic chain) and oligomers (**aliphatic compound 4,** double bond in aliphatic chain) were found. The monomer sequences and midchain substituent positions within these oligomers are uncertain in the absence of MS fragmentation information, but they may be assumed from prior monomer studies.[33] The separation and identification of those new oligomers from potato suberin are essential to understand the whole molecular structure and function of suberin. The special significance of these oligomers is that they demonstrate covalent ester linkages between monomers, which are strong evidence supporting the existence of an SPAD network.

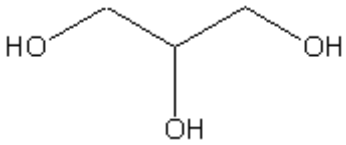
The structure of aliphatic ester oligomers obtained from potato suberin (**aliphatic compounds 3 aliphatic compounds 4**) is different from that of aliphatic ester oligomers obtained from fruit cutins. Cutin oligomers include midchain cross-link site, such as hydroxyl or ketone, but no suberin fragments exhibited these structural features. Moreover, cutins usually have very modest amount of aromatic constituents.

Further purification is required to elucidate **aromatic compound 3-f3**, **aromatic compound 4**, **aliphatic oligomers 4**, and others of significant interest. For instance, **aromatic compound 3-f3** may implicate covalent linkages between aliphatics and aromatics; **aromatic compound 4** may be a new type of aromatic compound that contains a nitrogen atom; **aliphatic oligomers 4** are aliphatic oligomers with double bonds in the middle of the aliphatic chain. Future improvement to the above separation and identification protocol might draw on experience with the successfully isolated compounds, using newly introduced HPLC equipment to achieve better separations and extensive MS^n to characterize a homologous series of new compounds.

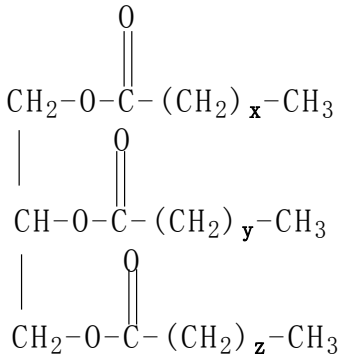
The **aromatic compound 3-f6 (** Bis(2-ethylhexyl) phthalate **)** was identified fully, but it is likely to be a contaminant from plastic containers or solvents instead of the plant material of interest. Nonetheless, it is still meaningful to identify this kind of compound as a cautionary measure for other researchers on related projects.

These findings, along with others published previously, comprise a database of compounds separated from suberin along with their 1D and 2D NMR spectra, HPLC separation conditions, molecular weight, and MS/MS fragmentation data. In addition to building our knowledge of this protective plant polymer, the database can guide the separation and elucidation efforts of future researchers.

a). Aliphatic compound 1: glycerol



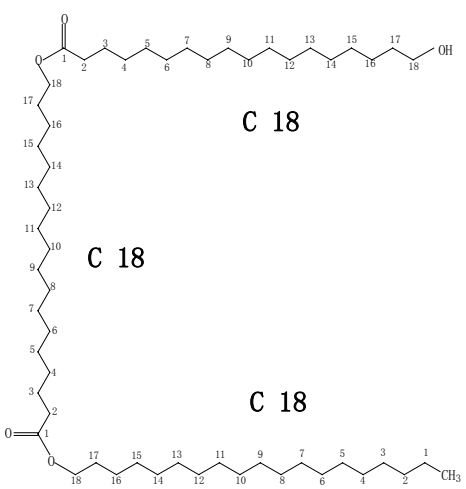
b). Aliphatic compound 2: family of triglycerides

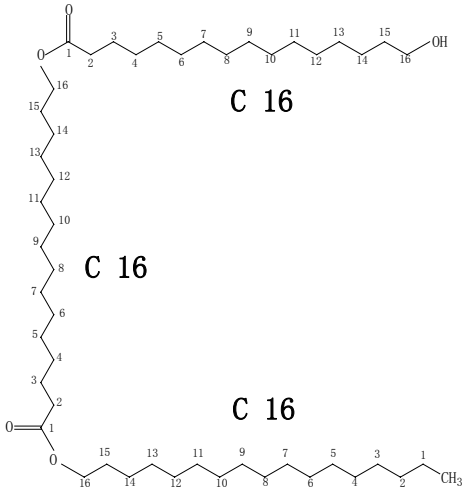


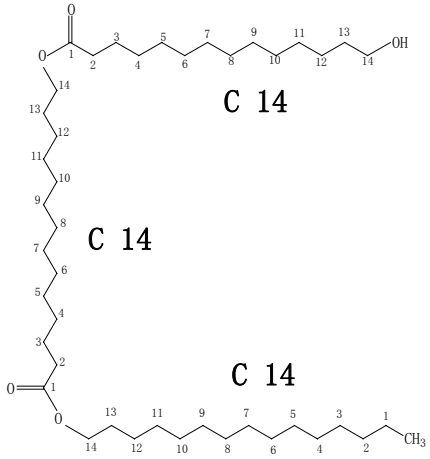
Molecular weight	**x**	**y**	**z**
470.4 *	6	6	6
498.4 *	6	6	8
544.4 *	6	8	8
582.6	6	10	10
638.7	10 8 8	10 10 8	10 12 14
666.7	10 8	10 10	12 14
722.7	8	14	14
750.8	12 10	12 14	14 14
764.8	12	13	14
778.8	12	14	14
792.8	13	14	14
806.8	14	14	14
834.6	14	14	16
890.7	16	16	16

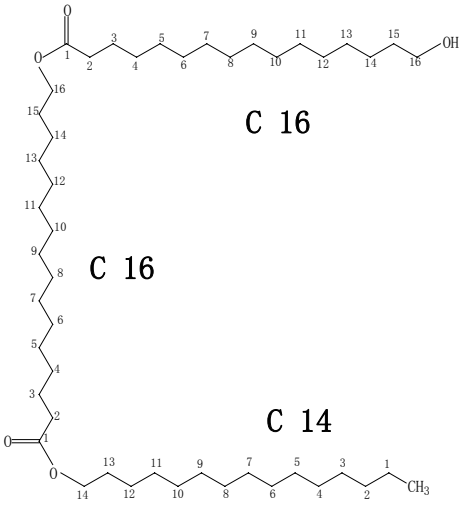
* main compound in a mixture

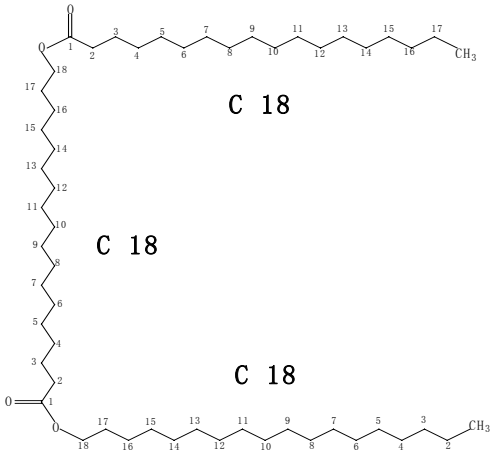
c). Aliphatic compounds 3: family of linear ester trimers (partial listing)

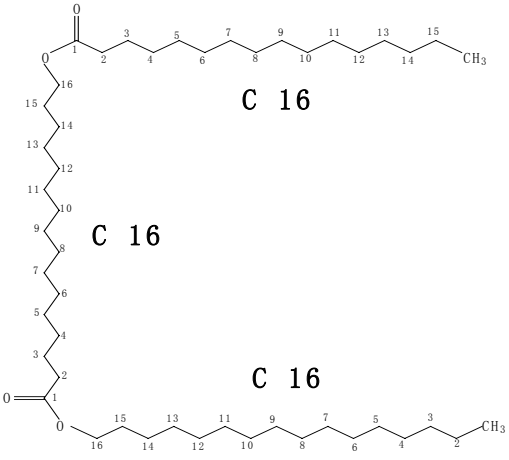


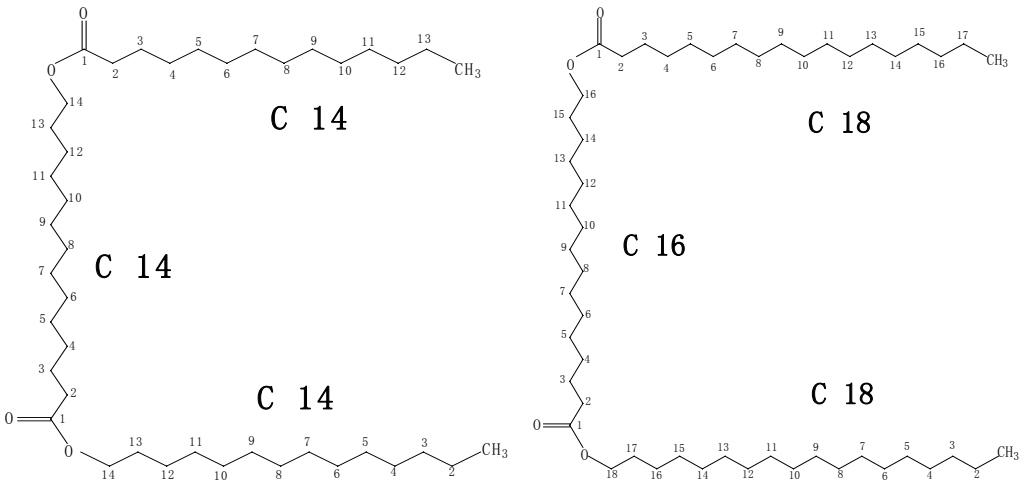








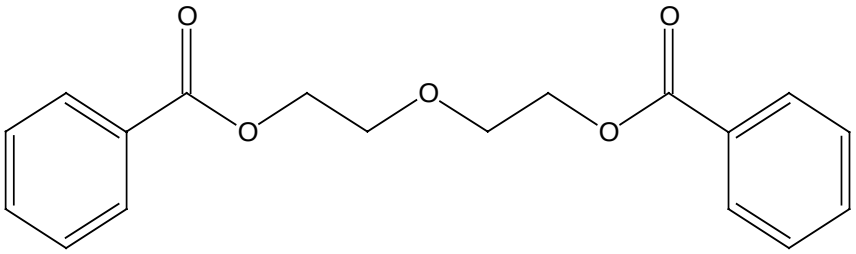




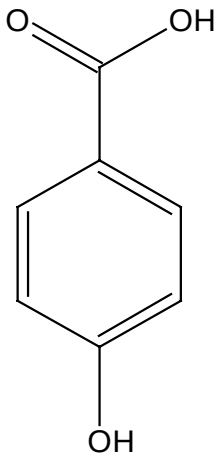
d). Aliphatic compounds 4:

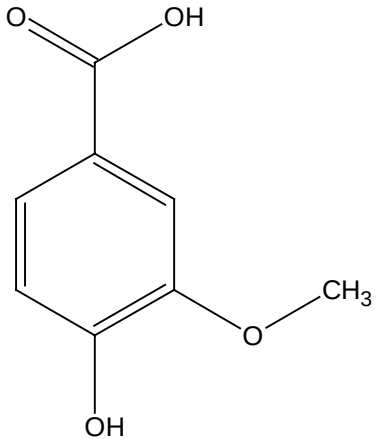
They are similar oligomers to **aliphatic compounds 3,** the only difference is **aliphatic compounds 4** have double bonds in the middle of the chain. MS/MS data are needed to finally confirm their structures.

e). Aromatic compound 1 (#11-TLC-4-1)



f). Aromatic compound 2-1 **aromatic compound 2-2**





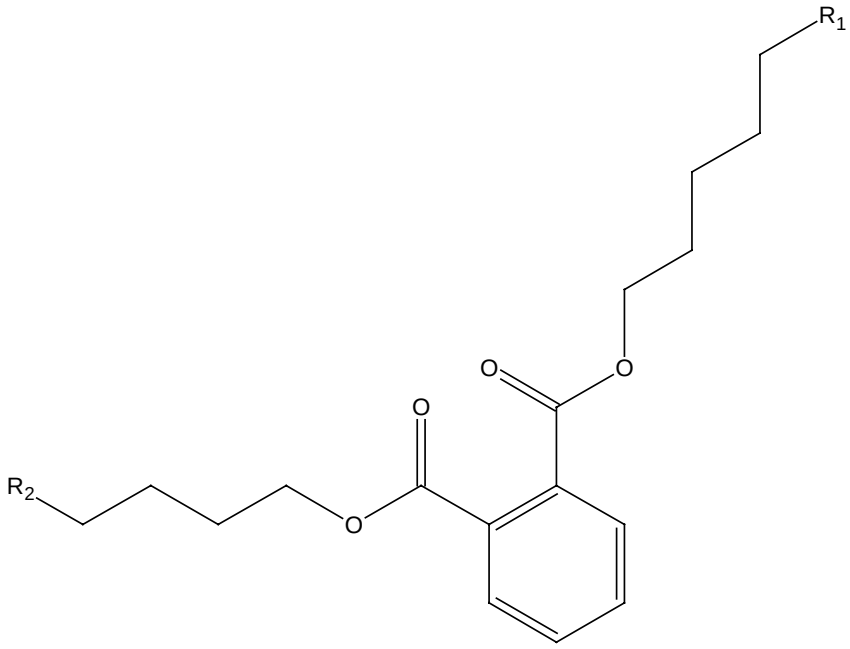
f'). co-existing aliphatic monomers with **aromatic compounds 2-1** and **2-2**

HO-CH$_2$-(CH$_2$)$_7$-COOH

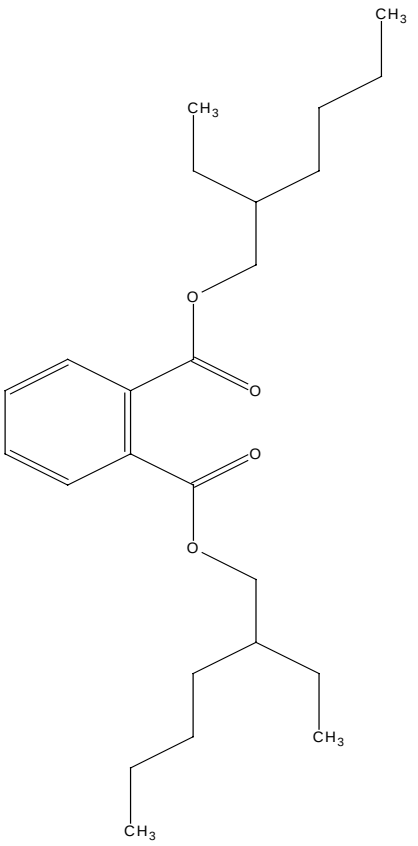
HO-CH$_2$-(CH$_2$)$_8$-COOH

HO-CH$_2$-(CH$_2$)$_9$-COOH

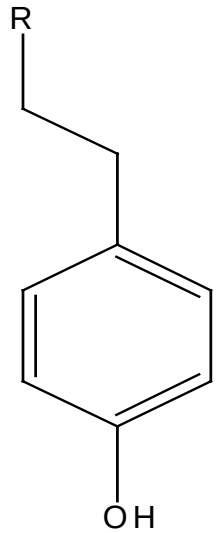
g). Aromatic compound 3-f3 (MS/MS data are needed to confirm the final structure**)**



h). Aromatic compound 3-f6 (Bis(2-ethylhexyl) phthalate: contaminant from solvent and /or tubing)



i). Aromatic compound 4-2 partial structure



Reference List

1. P. E. Kolattukudy, *Can.J.Bot.*, 1984, **62**, 2918-2933.

2. Kolattukudy, P. E. Cutin, Suberin and Waxes. Stumpf PK and Conn EE. The biochemistry of plants. [4], 571-645. 1980. New York, Academic Press.

3. P. E. Kolattukudy, *Ann.Rev.Plant Physiol.*, 1981, **32**, 539-567.

4. Kolattukudy, P. E. Lipid Polymers and Associated Phenols, Their Chemistry, Biosynthesis, and Role in Pathogenesis. Loewus FA and Runeckles VC. The structure, biosynthesis, and degradation of wood. 185-246. 1977. New York, Plenum Press.

5. M. A. Bernards and N. G. Lewis, *Phytochemistry*, 1998, **47**, 915-933.

6. A. V. Marques, H. Pereira, Meier.D., and O and Faix, *Holzforschung*, 1994, **48**, 43-50.

7. A. V. Marques, H. Pereira, D. Meier, and O and Faix, *Holzforschung*, 1999, **53**, 167-174.

8. C. P. Neto, N. Cordiero, A. Seca, F. Doingues, A. Gandini, and D and RObert, *Holzforschung*, 1996, **50**, 563-568.

9. E. Conde, E. Cadahia, M. C. Garcia-Vallejo, and J. R. and Gonzalez-Adrados, *J.Wood Chem.Tech.*, 1998, **18**, 447-469.

10. P. J Holloway, *Phytochemistry*, 1982, **21**, 2517-2522.

11. J. Graça and H Pereira, *Holzforschung*, 1997, **51**, 225-234.

12. J. Graça and H Pereira, *J.Wood Chem.Technol.*, 1998, **18**, 207-217.

13. J. Graça and H. Pereira, *J.Agric.Food Chem.*, 2000, **48**, 5476-5483.

14. J. Zeier and L. Schreiber, *Plant Physiol*, 1997, **113**, 1223-1231.

15. J. Zeier and L. Schreiber, *Planta*, 1998, **206**, 349-361.

16. M. H. Lopes, A. Sarychev, P. Neto, and A. M. Gil, *Solid State NMR*, 2000, **16**, 109-121.

17. R. E. Stark and J. R. Garbow, *Macromolecules*, 1992, **25**, 149-154.

18. B. Yan and R. E. Stark, *Macromolecules*, 1998, **31**, 2600-2605.

19. P. J. Holloway, *Phytochemistry*, 1983, **22**, 495-502.

20. P. E. Kolattukudy and V. P. Agrawal, *Lipids*, 1974, **9**, 682-691.

21. M. A. Bernards and N. G. Lewis, *Phytochemistry*, 1992, **31**, 3409-3412.

22. M. A. Bernards, M. L. Lopez, J. Zajicek, and N. G. Lewis, *J.Biol.Chem.*, 1995, **270**, 7382-7386.

23. B. Yan and R. E. Stark, *J.Agric.Food Chem.*, 2000, **48**, 3298-3304.

24. A. Marques, H. Pereira, D. Meier, and O Faix, *Holzforschung*, 1996, **50**, 393-400.

25. W. Cottle and P. E. Kolattukudy, *Plant.Physiol.*, 1982, **69**, 393-399.

26. R. E. Stark, T. Zlotnik-Mazori, L. M. Ferrantello, and J. R. Garbow, *ACS Symp.Ser.*, 1989, **399**, 214-229.

27. R. E. Stark, W. Sohn, Jr. R. A. Pacchiano, M. Al-Bashir, and J. R. Garbow, *Plant Physiol*, 1994, **104**, 527-533.

28. Gil A.M., M. Lopes, J. Rocha, and C. and Pascoal Neto, *J.Biol.Macromol.*, 1997, **20**, 293-305.

29. M. H. Lopes, A. M. Gil, A. D. Silverstre, and C. P. Neto, *J.Agric.Food Chem*, 2000, **48**, 383-391.

30. P. E. Kolattukudy, *Science*, 1980, **208**, 990-1000.

31. K. Esau, *Anatomy of Seed Plants*, John Wiley and Sons, New York, 1977.

32. K. Matzke and M. Riederer, *Planta*, 1991, **185**, 233-245.

33. M. A Bernards, *Can.J.Bot.*, 2002, **80**, 227-240.

34. J. Negrel, B. Pollet, and C. Lapierre, *Phytochemistry*, 1996, **43**, 1195-1199.

35. M. A Bernards and F. A. Razem, *Phytochemistry*, 2001, **57**, 1115-1122.

36. M. Friedman, *Journal of Agricultural and Food Chemistry*, 1997, **45**, 1523-1540.

37. M. M. Cowan, *Clinical Microbiology Reviews*, 1999, **12**, 564-582.

38. O. Borg-Olivier and B. Monties, *C.R.Acad.Sci.Paris.*, 1989, 141-147.

39. O. Borg-Olivier and B. Monties, *Phytochemistry*, 1993, **32**, 601-606.

40. L. Schreiber, K. Hartmann, M. Skrabs, and J. Zeier, *J.Exp.Botany*, 1999, **50**, 1267-1280.

41. J. Zeier, A. Goll, M. Yokoyama, I. Karahara, and L. Schreiber, *Plant Cell Environ.*, 1999, **22**, 271-279.

42. J. Zeier and L. Schreiber, *Planta*, 1999, **209**, 537-542.

43. I. Ribas and E. Blasco, *N.R.Soc.Esp.Fis.Quim.*, 1940, **36B**, 141-147.

44. A. Schmutz, T. Jenny, and U. Ryser, *Phytochemistry*, 1994, **36**, 1343-1346.

45. J. Graça and H. Pereira, *Biomacromolecules.*, 2000, **1**, 519-522.

46. L. Moire, A. Schmutz, A. Buchala, B. Yan, R. E. Stark, and U. Ryser, *Plant Physiol.*, 1999, **119**, 1137-1146.

47. J Graça and H. Pereira, *Phytochem.Anal.*, 2000, **11**, 45-51.

48. A. Schmutz, T. Jenny, N. Amrhein, and U. Ryser, *Planta*, 1993, **189**, 453-460.

49. E. Rosa and H. Pereira, *Holzforschung*, 1994, **48**, 226-232.

50. M. Bento, H. Pereira, M. Cunha, A. Moutinho, K. van den Berg, and J. Boon, *Phytochem.Anal.*, 1998, **9**, 1-13.

51. Murphy, D. Plant Lipids.Their Metabolism,Function,and Utilization . Lea, P. and Leegood, R. Plant biochemistry and molecular biology. [5]. 1993. Chichester,England, John Wiley and Sons .

52. M. H. Lopes, C. P. Neto, A. S. Barros, D. Rutledge, I. Delgadillo, and A. M Gil, *Biopolymers*, 2000, **57**, 344-351.

53. M. H. Lopes, C. P. Neto, D. Evtuguin, and A. M. Gil, *Holzforschung*, 1998, **52**, 146-148.

54. S. F. Osman, H. C. Gerard, W. F. Fett, R. A. Moreau, and R. L. Dudley, *J Agric Food Chem*, 1995, **43**, 2134-2137.

55. J. Graça and H Pereira, *Holzforschung*, 1999, **53**, 397-402.

56. D. Arrieta-Baez and R. E. Stark, *J.Agri.Food Chem.*, 2006, in press.

57. Hobart.H.W, Lynee L.M., John A.D., and Frank.A.S., *Instrumental Methods of Analysis*, Wadsworth,Inc, Belmont, California, Seventh edn., 1988.

58. Miller, J. M, *Chromatography: Concepts and Contrasts*, John Wiley & Sons, Inc., New York, 1988.

59. Sandie L., *High Performance Liquid Chromatography*, John Wiley & Sons Ltd, New York, Second edn., 1992.

60. Bovey F.A., Jelinski L., and Mirau P.A., *Nuclear Magnetic Resonance Spectroscopy*, Academic Press, Inc., San Diego, 1988.

61. Homans S.W., *A Dictionary of Concepts in NMR*, Oxford Science Publications, New York, 1997.

62. Claridge, T, *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier Inc., Oxford, 1999.

63. P. E. Kolattukudy and B. B. Dean, *Plant.Physiol.*, 1974, **54**, 116-121.

64. P. E. Kolattukudy, K. Kronman, and A. J. Poulose, *Plant.Physiol.*, 1975, **55**, 567-573.

65. R. A. Pacchiano, W. Sohn, V. L. Chlanda, J. R. Garbow, and R. E. Stark, *J.Agric.Food Chem.*, 1993, **41**, 78-83.

66. P. J. Holloway and A. H. B. Deas, *Phytochemistry*, 1973, **12**, 1721-1735.

67. R. Ekman, *Holzforschung*, 1983, **37**, 205-211.

68. H. C. Gerard, R. A. Moreau, W. F. Fett, and S. F. Osman, *J.Assoc.Oil Chemists Soc.*, 1992, **69**, 301-304.

69. X Fang, F Qiu, B Yan, H Wang, A Mort, and R. E. Stark, *Phytochemistry*, 2001, **57**, 1035-1042.

70. Tian, S., *Molecular Structures of Natural Polymers: cutin, suberin, and melanin*, 2005.

71. H. C. Gerard, S. F. Osman, W. F. Fett, and R. Moreau, *Phytochem.Anal.*, 1992, **3**, 139-144.

72. M Mossoba and M Firestone, *Food Testing and Analysis*, 1996, **2**, 24-32.

73. Mcintyre, D., *The Analysis of Triglycerides in Eidble Oils by APCI LC/MS*, Agilent Applcation Note, Publication Number: 5968-0878E, 2000.