CYTOTOXIC COMPOUNDS FROM THE GENUS *CENTAUREA*

MOHAMMAD SHOEB

A thesis submitted in partial fulfilment of the requirements of The Robert Gordon University for the degree of Doctor of Philosophy

This research programme was carried out in collaboration with the Department of Chemistry, University of Aberdeen

December 2005

ORIGINAL COPY TIGHTLY BOUND

TEXT BOUND CLOSE TO THE SPINE IN THE ORIGINAL THESIS

I dedicate this work to my parents

and

the rest of my family

Acknowledgements

I express my sincere and profound gratitude to my supervisor Dr. Satyajit Sarker, Reader of Pharmacy, University of Ulster at Coleraine, NI, UK for his constant guidance, encouragement, advice, invaluable suggestions and inspiration throughout the course of my study.

I also express my gratitude to my Director of Studies Dr. Stephen MacManus, School of Pharmacy, The Robert Gordon University, UK for his help, advice, guidance and inspiration during my research work

I am deeply indebted to my external supervisor Professor Marcel Jaspars, Department of Chemistry, University of Aberdeen, UK for his care, advice, help, guidance and encouragement during this work.

I would like to extend my heartfelt gratitude to my other supervisor Dr. Paul Kong, School of Life Sciences, The Robert Gordon University, UK for his help, advice and guidance during this study.

I am grateful to The Robert Gordon University, UK and School of Pharmacy for providing me the RDI funded studentship to carry out this postgraduate research program.

I would like to thank Professor Terry Healey, Head of the School of Pharmacy, RGU for financial support to attend conferences. I am also grateful to The Phytochemical Society of Europe for awarding me a travel bursary.

I am indebted to Dr. David Durham, Dr. Yash Kumarasamy, Mrs Moira Middleton, Mrs Maureen Byres, Mr Raymond Reid, Mr Brain de Jonckheere, Ms Jeanette Lowe, Ms Fiona Aitcheson and Ms Joanne Emslie, School of Pharmacy, RGU for their advice, help, cooperation and assistance during the last three years.

I extend my thanks to all other members of staff and technicians, and research students, School of Pharmacy. I am very grateful to Mr. Martin Simpson and Ms. Rosie Mearns, Research Degrees Officer, RGU for their help, support and assistance.

Special thanks to Mr Russell Gray, Department of Chemistry, University of Aberdeen for obtaining 2D NMR spectra.

I am grateful to Dr. Sezgin Celik, Department of Biology, Faculty of Science and Literature, Canakkale 18 Mart University, Turkey for sending plant materials of *Centaurea* species and Dr. Chris Wilcock, Department of Plant and Soil Science, Aberdeen University for voucher specimen of *Centaurea* species.

I would like to thank my friends Dr. Lutfun Nahar, Ben and Tim for their cooperation, help and encouragements during my work.

Thanks to the EPSRC National Mass Spectrometry Service Centre (Department of Chemistry, University of Wales, Swansea, UK for MS analyses.

I extend my warmest gratitude to my wife Farah for her inspiration, encouragement, assistance, wishes and patience during my work.

Cytotoxic compounds from the genus *Centaurea*

MOHAMMAD SHOEB

A thesis submitted in partial fulfilment of the requirements of The Robert Gordon University for the degree of Doctor of Philosophy

Abstract

This thesis, which is divided into four chapters, represents an account on the isolation, identification and the assessment of bioactivity of cytotoxic compounds from the genus *Centaurea* (Family: Asteraceae *alt.* Compositae), a large genus of about 500 species. The first three chapters deal with an introduction of natural products and *Centaurea* species, followed by the isolation and characterisation of compounds from twelve *Centaurea* species. The last chapter describes the bioactivities of extracts and isolated compounds from these species. A total of 45 compounds were isolated from twelve *Centaurea* species, and only *C. americana*, *C. cyanus*, *C. dealbata* and *C. macrocephala* had previously been studied. Four of these are novel compounds.

Four lignans arctiin, matairesinoside, matairesinol and lappaol A were isolated from the methanol extract of *C. macrocephala* seeds. Arctiin and matairesinoside were also isolated from the methanol extract of *C. americana*, *C. bornmuelleri*, *C. dealbata*, *C. huber-morathii*, *C. mucronifera*, *C. pamphylica*, *C. schischkinii* and *C. urvillei*. The methanol extract of *C. americana* also afforded 20-hydroxyecdysone, 24-hydroxyecdysone, lappaol A, arctigenin and a novel compound, 3''-*O*-caffeoyl-(9'''→3'')-arctiin. The methanol extract of *C. cyanus* produced lariciresinol 4-*O*-β-D-glucoside, berchemol, moschamine and *cis*-moschamine. Arctigenin, astragalin,

afzelin, matairesinol and a novel indole alkaloids, schischkiniin, were isolated from the methanol extract of *C. schischkinii*. Extract from *C. bornmuelleri* afforded arctigenin, astragalin, afzelin and matairesinol. The methanol extract of *C. montana* yielded berchemol, berchemol 4'-*O*-β-D-glucoside, *p*-coumaroylquinic acid, *cis*-*p*-coumaroylquinic acid, pinoresinol, pinoresinol monomethyl ether, pinoresinol dimethyl ether, pinoresinol 4-*O*-β-D-glucoside, pinoresinol 4,4'di-*O*-β-D-glucoside, pinoresinol 4-*O*-apiose-(1→2)-β-D-glucoside, centcyamine, *cis*-centcyamine, *N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine, *cis*-*N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine, moschamine, *cis*-moschamine, tryptamine and two novel compounds, flavanone-apiose-glucuronic acid and montamine. *C. gigantea* afforded arctiopicrin, 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide, chlorogenic acid, cirsiliol, isoquercetrin, orientin, isoorientin and 4''-hydroxybenzoyl-isoorientin.

General toxicity, cytotoxicity and antioxidant activity of the extracts and isolated compounds were evaluated, respectively, by the brine shrimp lethality assay, MTT assay on human colon cancer cell line (CaCo-2) and DPPH assay. Among all the species, the methanol extract of C. *bornmuelleri*, *C. gigantea*, *C. huber-morathii* and *C. montana* were the most toxic extracts in brine shrimp lethality and MTT assay. Arctigenin (IC_{50}=7.0 μM), matairesinol, montamine (IC_{50}=43.9 μM) and lappol A, schischkiniin, arctiopicrin (IC_{50}=8.5 μM) and 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (IC_{50}=26.4 μM) showed higher cytotoxicity against MTT assay. Matairesinoside (IC_{50}=2.2 × 10^{-3} mg/mL), matairesinol (IC_{50}=2.0 × 10^{-3} mg/mL) and schischkiniin (IC_{50}=3.8 × 10^{-3} mg/mL) exhibited significant free radical scavenging activities towards DPPH assay.

Contents

		Page
CHAPTER ONE		1
1.	**Introduction**	2
1.1	Background	2
1.2	Plant Derived Anticancer Agents	5
1.3	The Genus *Centaurea*	8
1.4	Phytochemistry of the Genus *Centaurea*	9
1.4.1	Sesquiterpenoids	10
1.4.2	Flavonoids	39
1.4.3	Lignans	45
1.4.3.1	Biosynthesis of Lignans	45
1.4.4	Alkaloids	48
1.4.5	Other Secondary Metabolites	50
1.5	*Centaurea* Species under Investigation	53
1.5.1	*C. americana* Nutt.	53
1.5.2	*C. bornmuelleri* Hausskn. Ex. Bornm.	53
1.5.3	*C. cyanus* L.	53
1.5.4	*C. dealbata* Willd.	54
1.5.5	*C. gigantea* Schultz Bip. ex Boiss.	54
1.5.6	*C. huber-morathii* Wagenitz	54
1.4.7	*C. macrocephala* Muss. Puschk. Ex Willd.	55
1.5.8	*C. montana* L.	55
1.5.9	*C. mucronifera* DC	55
1.5.10	*C. pamphylica* Boiss. & Heldr.	55
1.5.11	*C. schischkinii* Tzvelev.	56
1.5.12	*C. urvillei* subsp. armata Wagenitz	57
1.6	Aims and Objective of the Present Study	58

Page

CHAPTER TWO 59

2. Materials and Methods 60

2.1 Plant Materials 60

2.2 Extraction of Plant Materials 60

2.3 Isolation of Compounds 60

2.3.1 Thin Layer Chromatography (TLC) 60

2.3.2 Preparative Thin Layer Chromatography (PTLC) 61

2.3.3 Vacuum Liquid Chromatography (VLC) 62

2.3.4 Column Chromatography (CC) 63

2.3.4.1 Normal Column Chromatography 63

2.3.4.2 Sephadex LH-20 Column (Gel filtration) 63

2.3.4.3 Solid Phase Extraction (Sep-Pak) Column 64

2.3.5 High Performance (or Pressure) Liquid Chromatography (HPLC) 64

2.3.5.1 Analytical HPLC 64

2.3.5.2 Preparative HPLC 64

2.3.5.3 Semi-Preparative HPLC 65

2.4 Detection of Compounds 66

2.4.1 UV Light 66

2.4.2. Spray Reagents 66

2.4.2.1 Vanillin in Sulphuric acid 66

2.4.2.2 Anisaldehyde-Sulphuric acid 66

2.4.2.3 Modified Dragendorff's Reagent 66

2.5 Instrumentation 67

2.5.1 General Laboratory Instrument 67

		Page
2.5.1.1	Melting Point Apparatus	67
2.5.1.2	Polarimeter	67
2.5.2	Spectroscopic Instrument	67
2.5.2.1	UV Spectrophotometer	67
2.5.2.2	IR Spectrophotometer	67
2.5.2.3	Mass Spectrometry	67
2.5.2.4	NMR Spectroscopy	67
2.6	Structure Elucidation of Compounds	68
2.7	Isolation of Compounds	71
2.7.1	Isolation of Compounds from *Centaurea macrocephala*	71
2.7.2	Isolation of Compounds from *Centaurea cyanus*	72
2.7.3	Isolation of Compounds from *Centaurea dealbata*	73
2.7.4	Isolation of Compounds from *Centaurea americana*	74
2.7.5	Isolation of Compounds from *Centaurea huber-morathii*	74
2.7.6	Isolation of Compounds from *Centaurea montana*	75
2.7.7	Isolation of Compounds from *Centaurea schischkinii*	80
2.7.8	Isolation of Compounds from *Centaurea bornmuelleri*	82
2.7.9	Isolation of Compounds from *Centaurea gigantea*	82
2.7.10	Isolation of Compounds from *Centaurea pamphylica*	84
2.7.11	Isolation of compounds from *Centaurea urvillei*	85
2.7.12	Isolation of Compounds from *Centaurea mucronifera*	85

Page

CHAPTER THREE 87

3. Results and Discussion 88

3.1 Lignans 88

3.1.1 Dibenzylbutyrolactone Lignans 88

3.1.1.1 Characterisation of **SM1**, **SM2**, **SM3** and **SM4** as Arctiin, Matairesinoside, Arctigenin and Matairesinol, respectively 89

3.1.1.2 Characterisation of **SM5** as 3''-*O*-Caffeoyl-(9'''→3)-arctiin 95

3.1.1.3 Characterisation of **SM6** as Lappaol A 98

3.1.1.4 Properties of Dibenzylbutryolactone Lignans (**SM1** to **SM6**) 101

3.1.1.4.1 Properties of Arctiin (**SM1**) 101

3.1.1.4.2 Properties of Matairesinoside (**SM2**) 101

3.1.1.4.3 Properties of Arctigenin (**SM3**) 101

3.1.1.4.4 Properties of Matairesinol (**SM4**) 101

3.1.1.4.5 Properties of 3''-*O*-Caffeoyl- arctiin (**SM5**) 102

3.1.1.4.6 Properties of Lappaol A (**SM6**) 102

3.1.2 Epoxy lignans 103

3.1.2.1 Characterisation of **SM7**, **SM8** and **SM9** as Lariciresinol 4'-*O*-β-D-glucopyranoside, Berchemol and Berchemol 4'-*O*-β-D-glucopyranoside 103

3.1.2.2 Characterisation of **SM10**, **SM11** and **SM12** as pinoresinol, pinoresinol monomethyl ether and pinoresinol dimethyl ether, Respectively 110

3.1.2.3 Characterisation of **SM13**, **SM14** and **SM15** as pinoresinol 4-*O*-β-D-glucopyranoside, pinoresinol di 4-*O*-β-D-glucopyranoside and pinoresinol 4-*O*-β-D-apiofuranosyl-(1→2) glucopyranoside, respectively 113

3.1.2.4 Properties of Epoxylignans (**SM7-SM15**) 118

		Page
3.1.2.4.1	Properties of Lariciresinol *O*-4'- β-D-glucopyranose (**SM7**)	118
3.1.2.4.2	Properties of Berchemol (**SM8**)	118
3.1.2.4.3	Properties of Berchemol-O-4'- β-D-glucopyranose (**SM9**)	118
3.1.2.4.4	Properties of Pinoresinol (**SM10**)	118
3.1.2.4.5	Properties of Pinoresinol monomethyl ether (**SM11**)	119
3.1.2.4.6	Properties of Pinoresinol dimethyl ether (**SM12**)	119
3.1.2.4.7	Properties of (+) Pinoresinol-*O*-4-β-D-glucopyranoside (**SM13**)	119
3.1.2.4.8	Properties of (+) Pinoresinol 4, 4' di-β-D-glucopyranoside (**SM14**)	120
3.1.2.4.9	Properties of (+) Pinoresinol-4-*O*-β-D-apiosefuranosyl-(1→2)-β-D-glucopyranoside (**SM15**)	120
3.2	Sesquiterpene lactones	121
3.2.1	Germacranolides	121
3.2.1.1	Characterisation of **SM16** and **SM17** as Arctiopicrin and 8-*O*-(4-hydroxy-3-methylbutanoyl)-Salonitenolide, respectively	121
3.2.2	Eudesmane	126
3.2.2.1	Characterisation of **SM18** as Pterodontriol (**180**)	126
3.2.3	Properties of sesquiterpene lactones (**SM16-SM17**)	127
3.2.3.1	Properties of 8-*O*-(3-hydroxy-2-methylpropanoyl)-Salonitenolide or Arctiopicrin (**SM16**)	127
3.2.3.2	Properties of 8-*O*-(4-hydroxy-3-methylbutanoyl)-Salonitenolide (**SM17**)	127
3.2.3.3	Properties of Pterodontriol (**180**)	127
3.3	Quinic acid Derivatives	128
3.3.1	Characterisation of **SM19** and **SM20** as *trans*-3-*O*-*p*-Coumaroylquinic acid and *cis*-3-*O*-*p*-Coumaroylquinic acid, respectively	128

		Page
3.3.2	Characterisation of **SM21** as Chlorogenic acid (**168**)	131
3.3.3	Properties of Quinic acid Derivatives (**SM19-SM21**)	133
3.3.3.1	Properties of *trans*- 3-*O*-*p*-coumaroylquinic acid (**SM19**)	133
3.3.3.2	*cis*-3-*O*-*p*-coumaroylquinic acid (**SM20**)	133
3.3.3.3	Properties of *trans*-3-*O*-*p*-caffeoylquinic acid (**SM21**)	133
3.4	Ecdysteroids	134
3.4.1	Characterisation of **SM22** and SM23 as 20-hydroxy-ecdysone (**169**) and Makisterone A (**183**), respectively	134
3.4.2	Properties of Ecdysteroids (**SM22-SM23**)	136
3.4.2.1	Properties of 20-Hydroxyecdysone (**SM22**)	136
3.4.2.2	Properties of Makisterone A (**SM23**)	136
3.5	Steroids	137
3.5.1	Characterisation of **SM24** as Stigmast-4-en-3-ol (**184**)	137
3.5.2	Characterisation of **SM25** as 3-Hydroxystigmast-5,22-dien-1-one	138
3.5.3	Characterisation of Asterosterol (**SM26**)	138
3.5.4	Properties of Steroids (**SM24-SM26**)	139
3.5.4.1	Properties of Stigmast-4-en-3-ol (**SM24**)	139
3.5.4.2	Properties of 3-Hydroxystigmast-5,22-dien-1-one (**SM25**)	139
3.5.4.3	Properties of Asterosterol (**SM26**)	139
3.6	Flavonoids	140
3.6.1	Characterisation of **SM27**, **SM28** and **SM29** as Kaempferol, Astragalin and Afzelin, respectively	140
3.6.2	Characterisation of **SM30** and **SM31** as Apigenin and Isoquercetrin respectively	144
3.6.3	Characterisation of **SM32** as 6-hydroxy luteolin 6,7-dimethylether (**114**)	147

		Page
3.6.4	Characterisation of **SM33** and **SM34** as Orientin (188) and Isoorientin (129), respectively	149
3.6.5	Characterisation of **SM35** as 2'''-(4'''-Hydroxybenzoyl)-isoorientin	152
3.6.6	Characterisation of **SM36** as Flavanone 7-*O*-apiofuranosyl 1→2)-glucuronic acid (**190**)	154
3.6.7	Properties of Flavonoids (**SM27-SM36**)	157
3.6.7.1	Properties of Kaempferol (**SM27**)	157
3.6.7.2	Properties of Astragalin (**SM28**)	157
3.6.7.3	Properties of Afzelin (**SM29**)	157
3.6.7.4	Properties of Apigenin (**SM30**)	157
3.6.7.5	Properties of Isoquercetrin (**SM31**)	157
3.6.7.6	Properties of Cirisliol (**SM32**)	158
3.6.7.7	Properties of Orientin (**SM33**)	158
3.6.7.8	Properties of Isoorientin (**SM34**)	158
3.6.7.9	Properties of 4'''-Hydroxybenzoyl- isoorientin (**SM35**)	158
3.6.7.10	Properties of 7-O-Apiofuranosyl (1→4)-glucuronic acid (**SM36**)	159
3.7	Alkaloids	160
3.7.1	Isolation and Characterisation of Indole Alkaloids	160
3.7.1.1	Identification of Indole Alkaloids	160
3.7.1.2	Characterisation of **SM37** as Tryptamine (**191**)	161
3.7.1.3	Characterisation of **SM38** and **SM39** as (*E*)-*N*-4-hydroxycinnamoyl)-5-hydroxytyptamine and (*Z*)-*N*-4-hydroxycinnamoyl)-hydroxytryptamine, respectively	162
3.7.1.4	Characterisation of **SM40** and **SM41** as Centcyamine and *cis*-Centcyamine, respectively	165
3.7.1.5	Characterisation of **SM42 and SM43** as Moschamine and *cis*-Moschamine, respectively	167
3.7.1.6	Characterisation of **SM44** as Montamine (**193**)	169
3.7.1.7	Characterisation of **SM45** as Schischkiniin (**194**)	172

		Page
3.7.2	Properties of Indole Alkaloids (**SM38-SM46**)	177
3.7.2.1	Properties of Tryptamine (**SM37**)	177
3.7.2.2	Properties of (*E*) *N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**SM38**)	177
3.7.2.3	Properties of (*Z*) *N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**SM39**)	177
3.7.2.4	Properties of (*E*) *N*-(4-hydroxycinnamoyl)-5-methoxytryptamine or Centcyamine (**SM40**)	177
3.7.2.5	Properties of (*E*) *N*-(4-hydroxycinnamoyl)-5-methoxytryptamine or *cis*-Centcyamine (**SM41**)	177
3.7.2.6	Properties of (*E*) *N*-(3-methoxy-4-hydroxycinnamoyl)-5-hydroxytyptamine or Moschamine (**SM42**)	178
3.7.1.7	Properties of (*Z*) *N*-(3-methoxy-4-hydroxycinnamoyl)-5-hydroxytyptamine or *cis*-Moschamine (**SM43**)	178
3.7.2.8	Properties of Montamine (**SM44**)	178
3.7.2.9	Properties of Schischkiniin (**SM45**)	178
CHAPTER FOUR		179
Biological Activities of Secondary Metabolites from *Centaurea* Species		179
4. 1	Introduction	180
4.1.1	Background	180
4.1.2	Brine Shrimp Lethality Assay	180
4.1. 3	MTT Cytotoxicity Assay	181
4.1.4	DPPH Antioxidant Assay	182
4.2	Materials and Methods	185
4.2.1	Brine Shrimp Lethality Assay	185
4.2.1.1	Preparation of Brine Shrimp Medium (BSM)	185
4.2.1.2	Bioassay of Sample with BSM	185
4.2.2	MTT Cytotoxicity Assay	186
4.2.2.1	Materials and Chemicals	186
4.2.2.2	Cryopreservation	186

		Page
4.2.2.3	Thawing of frozen cells	187
4.2.2.4	Cell Counting	188
4.2.2.5	Cell Viability Assay	189
4.2.3	DPPH Antioxidant Assay	191
4.2.3.1	Chemical and Materials	192
4.2.3.2	Qualitative Assay	191
4.2.3.3	Quantitative Assay	191
4.3	Results and Discussion	192
4.3.1.1	Brine Shrimp Assay of *Centaurea* Extracts	192
4.3.1.2	Brine Shrimp Lethality Assay of Lignans	193
4.3.1.3	Brine Shrimp Lethality Assay of Sesquiterpene lactones	195
4.3.1.4	Brine Shrimp Lethality Assay of Quinic acid Derivatives	196
4.3.1.5	Brine Shrimp Lethality Assay of Flavonoids	197
4.3.1.6	Brine Shrimp Assay of Alkaloids	198
4.3.2.	*In vitro* Cytotoxicity Studies against Colon Cancer Cell (CaCo-2)	203
4.3.2.1	MTT Cytotoxicity Assay of *Centaurea* Extracts	201
4.3.2.2	MTT Cytotoxicity Assay of Lignans	202
4.3.2.3	MTT Cytotoxicity Assay of Sesquiterpene lactones	203
4.3.2.4	MTT Cytotoxicity Assay of Quinic acid Derivatives	204
4.3.2.5	MTT cytotoxicity assay of Flavonoids	205
4.3.2.6	MTT cytotoxicity assay of Alkaloids	206
4.3.3	DPPH Antioxidant Assay	207
4.3.3.1	DPPH Assay of *Centaurea* Extracts	207
4.3.3.2	DPPH Assay of Lignans	207
4.3.3.3	DPPH Assay of Sesquiterpene lactones	208
4.3.3.4	DPPH Assay of Quinic acid Derivatives	209
4.3.3.5	DPPH Assay of Flavonoids	209
4.3.3.6	DPPH Assay of Alkaloids	210
4.4	Conclusions	211

References **Page** 213-232

Appendix A

NMR Spectra of Isolated Compounds (Figure 78-134)

Appendix B

List of Publications

List of Schemes

		Page
Scheme 1.	Biosynthesis of carvone (mono terpenoids) in *Mentha spicata*	11
Scheme 2.	Biosynthesis of flavonoids	40
Scheme 3.	Biosynthesis of lignans	46
Scheme 4.	Biosynthesis of indole alkaloids	48
Scheme 5.	Transformation of MTT by viable cells	181
Scheme 6.	DPPH reduction	184

List of Tables

		Page
Table 1.	Plant secondary metabolites from *Centaurea* species	12
Table 2.	List of *Centaurea* species investigated	61
Table 3.	List of solvent systems used for TLC and PTLC	62
Table 4.	Dibenzylbutyrolactone-type lignan isolated from *Centaurea* species	88
Table 5.	^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz) and ^{13}C NMR data for **SM1 (137)** and **SM2 (138)**	92
Table 6.	^{1}H NMR and ^{13}C NMR data for **SM3 (139)** and **SM4 (140)**	94
Table 7.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM5 (170)**	97
Table 8.	^{1}H NMR ^{13}C NMR data and long range HMBC for **SM6 (143)**	100
Table 9.	Epoxy lignans isolated from *Centaurea* species	103
Table 10.	^{1}H NMR, ^{13}C NMR data for **SM7 (171)** and **SM8 (172)**	107
Table 11.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM9 (173)**	109
Table 12.	^{1}H NMR, ^{13}C NMR data, ^{1}H-^{1}H COSY and ^{1}H-^{13}C long range HMBC for **SM10 (144)**	112
Table 13.	^{1}H NMR for **SM13 (176)**, **SM14 (177)** and **SM15 (178)**	115
Table 14.	^{13}C NMR data for **SM13 (176)**, **SM14 (177)** and **SM15 (178)**	116
Table 15.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM16 (21)**	123
Table 16.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM17 (179)**	125
Table 17.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM19 (181)**	129
Table 18.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM21 (168)**	132
Table 19.	^{1}H NMR and ^{13}C NMR data for **SM22 (169)** and **SM23 (183)**	135
Table 20.	Flavonoids isolated from *Centaurea* species	140
Table 21.	UV absorbance of flavonoids in methanol	141
Table 22.	^{1}H NMR for **SM27 (122)**, **SM28 (123)** and **SM29 (187)**	142
Table 23.	^{13}C NMR data for **SM27 (122)**, **SM28 (123)** and **SM29 (187)**	143
Table 24.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM30 (108)**	145
Table 25.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM32 (110)**	148
Table 26.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM33 (188)**	151
Table 27.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM35 (189)**	153

		Page
Table 28.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM36 (190)**	156
Table 29.	UV absorptions of indole alkaloids in methanol	160
Table 30.	Mass data for indole alkaloids	161
Table 31.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM37 (191)**	162
Table 32.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM38 (145)** and **SM39 (192)**	165
Table 33.	^{1}H NMR, ^{13}C NMR data and long range HMBC in CD_3OD for **SM40 (146)**	167
Table 34.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM42 (148)**	168
Table 35.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM44 (193)**	171
Table 36.	Major NMR difference between **SM42 (148)** and **SM44 (193)**	172
Table 37.	^{1}H NMR, ^{13}C NMR data and long range HMBC for **SM45 (194)**	174
Table 38.	Composition of brine shrimp medium	185
Table 39.	Brine shrimp assay of *Centaurea* species	192
Table 40.	Brine shrimp assay of lignans	195
Table 41.	Brine shrimp assay of sesquiterpene lactone	195
Table 42.	Brine shrimp assay of quinic acid derivatives	197
Table 43.	Brine shrimp assay of flavonoids	197
Table 44.	Brine shrimp assay of alkaloids	200
Table 45.	Cytotoxicity activities of *Centaurea* species	201
Table 46.	Cytotoxicity activities of lignans against CaCo-2	203
Table 47.	Cytotoxic activities of sesquiterpene lactone	204
Table 48.	Cytotoxicity activities of quinic acid derivatives	204
Table 49.	Cytotoxicity activities of flavonoids	205
Table 50.	Cytotoxicity activities of alkaloids	206
Table 51.	DPPH assay of *Centaurea* species	207
Table 52.	DPPH assay of lignans	208
Table 53.	DPPH assay of sesquiterpene lactone	209
Table 54.	DPPH assay of quinic acid derivatives	209
Table 55.	DPPH assay of flavonoids	210
Table 56.	DPPH assay of alkaloids	210

List of Figures

		Page
Figure 1.	Structures of the sesquiterpene lactones from the genus *Centaurea*	29
Figure 2.	Structures of the flavonoids from the genus *Centaurea*	41
Figure 3.	Structures of the lignans from the genus *Centaurea*	47
Figure 4.	Structures of the alkaloids from the genus *Centaurea*	49
Figure 5.	Structures of the miscellaneous compounds from the genus *Centaurea*	50
Figure 6.	*Centaurea cyanus* seeds	54
Figure 7.	*Centaurea dealbata* seeds	54
Figure 8.	*Centaurea schischkinii* seeds	56
Figure 9.	*Centaurea bornmuelleri* seeds	56
Figure 10.	*Centaurea mucronifera*	57
Figure 11.	*Centaurea urvillei*	57
Figure 12.	General Scheme for structure elucidation	69
Figure 13.	HPLC chromatogram of fraction **Ma1** from *C. macrocephala*	72
Figure 14.	HPLC chromatogram of 30% Sep-Pak fraction of *C. montana*	75
Figure 15.	HPLC chromatogram of 40% Sep-Pak fraction of *C. montana*	76
Figure 16.	HPLC chromatogram of fraction **M1** from 30% Sep-Pak of *C. montana*	76
Figure 17.	HPLC chromatogram of fraction **M2** from 30% Sep-Pak of *C. montana*	77
Figure 18.	HPLC chromatogram of fraction **M4** from 30% Sep-Pak of *C. montana*	78
Figure 19.	HPLC chromatogram of fraction **M6** from 30% Sep-Pak of *C. montana*	78
Figure 20.	HPLC chromatogram of 60% Sep-Pak fraction from *C. montana*	79
Figure 21.	HPLC chromatogram of 40% Sep-Pak fraction from *C. schischkinii*	80
Figure 22.	HPLC chromatogram of 60% Sep-Pak fraction from *C. schischkinii*	81

		Page
Figure 23.	HPLC chromatogram of 30% Sep-Pak fraction from *C. gigantea*	83
Figure 24.	HPLC chromatogram of 60% Sep-Pak fraction from *C. gigantea*	83
Figure 25.	Structure of arctiin (**SM1**) and matairesinoside (**SM2**)	89
Figure 26.	Key NOE interactions of **SM1** (**137**) based on ^{1}H-^{1}H NOESY experiment	91
Figure 27.	Structure of arctigenin (**SM3**) and matairesinol (**SM4**)	93
Figure 28.	Structure of 3''-*O*-caffeoyl-(9'''→3'')-arctiin (**SM5**)	98
Figure 29.	Structure of lappaol A (**SM6**)	101
Figure 30.	Structure of lariciresinol 4'-*O*-β-D-glucopyranoside and berchemol	104
Figure 31.	Key HMBC (^{1}H→^{13}C) correlations of **SM7** (**171**)	105
Figure 32.	Key NOE interactions of **SM8** (**172**) based on ^{1}H-^{1}H NOESY experiment	106
Figure 33.	Structure of berchemol 4'-*O*-β-D-glucopyranoside (**SM9**)	108
Figure 34.	Structures of pinoresinol (**SM10**), pinoresinol monomethyl ether (**SM11**) and pinoresinol dimethyl ether (**SM12**)	111
Figure 35.	Key HMBC correlations (^{1}H→^{13}C) of **SM10** (**144**)	111
Figure 36.	Structures of pinoresinol glycosides (**SM13-SM15**)	114
Figure 37.	Structure of arctiopicrin (**SM16**)	122
Figure 38.	8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**SM17**)	124
Figure 39.	Structure of Pterodontriol (**SM18**)	126
Figure 40.	Key HMBC correlations (^{1}H→^{13}C) of substructure **A** and **B** of **SM19** (**181**)	129
Figure 41.	Structures of *trans* and *cis* quinic acid derivatives (**SM19**, **SM20**)	130
Figure 42.	Structure of chlorogenic acid (**SM21**)	131
Figure 43.	Structure of 20-hydroxyecdysone (**SM22**) and makisterone (**SM23**)	134
Figure 44.	Structure of stigmast-4-en-3-ol (**SM24**)	137

		Page
Figure 45.	Structure of 3-hydroxystigmast-5,22-dien-1-one **(SM25)**	138
Figure 46.	Structure of asterosterol **(SM26)**	139
Figure 47.	Strucutures of Flavonols **(SM27-SM29)**	141
Figure 48.	Structure of apigenin **(SM30)**	145
Figure 49.	Structure of isoquercetrin **(SM31)**	146
Figure 50.	Structure of cirsiliol **(SM32)**	147
Figure 51.	Structure of *C*-glycosides **(SM33 and SM34)**	150
Figure 52.	Structure of 2''-(4'''-hydroxybenzoyl)- isoorientin **(SM35)**	152
Figure 53.	Structure of flavanone 7-*O*-apiofuranosyl (1→4)-glucuronic acid **(SM36)**	155
Figure 54.	Structure of tryptamine **(SM37)**	162
Figure 55.	Structure of indole alkaloids **(SM38, SM40** and **SM42)**	163
Figure 56.	Key ^{1}H-^{13}C HMBC correlations of **SM38 (145)**	164
Figure 57.	Structure of indole alkaloids **(SM39, SM41** and **SM43)**	166
Figure 58.	Structure of montamine **(SM44)**	170
Figure 59.	Tentative structures of schischkiniin **(SM45)**	173
Figure 60.	Structure of schischkiniin **(SM45)**	174
Figure 61.	Proposed biogenetic pathway for the formation of schischkiniin (**194**) from a simple diketopiperazine **195**. The reduction/dehydration and dimerisation steps may be reversed	176
Figure 62.	Global energy minima for structures **194b** and **194c** (Heavy atoms and polar hydrogens only shown). The global minimum for **46b** was found 15 times (E_{tot} = 1022.68 kcal/mol) and the global minimum for **194c** was found 37 times (E_{tot} = 986.93 kcal/mol)	176
Figure 63.	Cell viability (100%) against drug concentration curve for determination of IC_{50} value by MTT assay	182
Figure 64.	Structure of resveratrol (**198**)	183
Figure 65.	Scavenging activity (100%) against drug concentration curve to determine IC_{50} value by DPPH assay	184
Figure 66.	Laminar Flow hood	187

		Page
Figure 67.	Haemocytometer	188
Figure 68.	Particle size distributor	188
Figure 69.	Untreated CaCo-2 Cell	189
Figure 70.	Plate reader	150
Figure 71.	Structures of lignans tested in the bioassays	194
Figure 72.	Structures of sesquiterpe lactones tested in the bioassays	196
Figure 73.	Structures of quinic acid derivatives tested in the bioassays	196
Figure 74.	Structures of flavonoids tested in the bioassays	198
Figure 75.	Structures of alkaloids	199
Figure 76.	Structures of podophyllotoxin and quercetin	200
Figure 77.	Comparative of biological activity studies of *Centaurea* extracts	211

Figures 78-134 (NMR Spectra of Isolated Compounds) are in Appendix A

Figure 78. ^{1}H NMR spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

Figure 79. ^{13}C NMR spectrum (CD_3OD, 150 MHz) of **SM1** (arctiin, **137**)

Figure 80. ^{1}H-^{1}H COSY spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

Figure 81. HMBC spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

Figure 82. NOESY spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

Figure 83. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM2** (matairesinoside, **138**)

Figure 84. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM3** (arctigenin, **139**)

Figure 85. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM5** (matairesinol, **140**)

Figure 86. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM4** (matairesinol, **140**)

Figure 87. HMBC spectrum (CD_3OD, 400 MHz) of **SM4** (matairesinol, **140**)

Figure 88. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

Figure 89. ^{1}H-^{1}H COSY spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

Figure 90. HMBC spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

Figure 91. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM7** (lariciresinol 4'-*O*-β-D-glucopyranoside, **171**)

Figure 92. HMBC spectrum (CD_3OD, 400 MHz) of **SM7** (lariciresinol 4'-*O*-β-D-glucopyranoside, **171**)

Figure 93. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM8** (berchemol, **172**)

Figure 94. NOESY spectrum (CD_3OD, 400 MHz) of **SM8** (berchemol, **172**)

Figure 95. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM9** (berchemol 4'-*O*-β-D-glucopyranoside, **173**)

Figure 96. HMBC spectrum (CD_3OD, 400 MHz) of **SM9** (berchemol 4'-*O*-β-D-glucopyranoside, **173**)

Figure 97. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM10** (pinoresinol, **144**)

Figure 98. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM10** (pinoresinol, **144**)

Figure 99. HMBC spectrum (CD_3OD, 400 MHz) of **SM10** (pinoresinol, **144**)

Figure 100. ROESY spectrum (CD_3OD, 400 MHz) of **SM10** (pinoresinol, **144**)

Figure 101. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM13** (pinoresinol 4-*O*-β-D-glucopyranoside, **176**)

Figure 102. HMBC spectrum of **SM15** (pinoresinol 4-*O*-β-D-apiose furanosyl-(1→2)-β-D-glucopyranoside, **178**)

Figure 103. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM16** (arctiopicrin, **21**)

Figure 104. HMBC spectrum (CD_3OD, 100 MHz) of **SM16** (arctiopicrin, **21**)

Figure 105. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

Figure 106. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

Figure 107. HMBC spectrum (CD_3OD, 400 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

Figure 108. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM19** (*trans* 3-*O*-p-coumaroyl quinic acid, **181**)

Figure 109. HMBC spectrum (CD_3OD, 400 MHz) of **SM19** (*trans* 3-*O*-p-coumaroyl quinic acid, **181**)

Figure 110. HMBC spectrum (CD_3OD, 400 MHz) of **SM21** (chlorogenic acid, **168**)

Figure 111. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

Figure 112. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM29** (afzelin, **187**)

Figure 113. ^1H-^1H COSY spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

Figure 114. HMBC spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

Figure 115. ^1H NMR spectrum (DMSO-d_6, 400 MHz) of **SM30** (apigenin, **108**)

Figure 116. ^{13}C NMR spectrum (DMSO-d_6, 100 MHz) of **SM30** (apigenin, **108**)

Figure 117. ^1H NMR spectrum (DMSO-d_6, 400 MHz) of **SM33** (orientin, **188**)

Figure 118. ^{13}C NMR spectrum (DMSO-d_6, 100 MHz) of **SM33** (orientin, **188**)

Figure 119. HMBC spectrum (DMSO-d_6, 400 MHz) of **SM33** (orientin, **188**)

Figure 120. HMBC spectrum (CD_3OD, 400 MHz) of **SM35** (4'''-hydroxybenzoyl- isoorientin, **189**)

Figure 121. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM36** (7-*O*-apiofuranosyl (1→4)-glucuronic acid, **190**)

Figure 122. HSQC spectrum (CD_3OD, 400 MHz) of **SM36** (7-*O*-apiofuranosyl (1→4)-glucuronic acid, **190**)

Figure 123. HMBC spectrum (CD_3OD, 400 MHz) of **SM36** (7-*O*-apiofuranosyl (1→4)-glucuronic acid, **190**)

Figure 124. HMBC spectrum (CD_3OD, 400 MHz) of **SM37** (tryptamine, 191)

Figure 125. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM42** (moschamine, **148**)

Figure 126. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM42** (moschamine, **148**)

Figure 127. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM42** (moschamine, **148**)

Figure 128. HMBC spectrum (CD_3OD, 400 MHz) of **SM42** (moschamine, **148**)

Figure 129. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

Figure 130. HSQC spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

Figure 131. ^1H-^1H COSY spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

Figure 132. HMBC spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

Figure 133. ^1H NMR spectrum (CD_3OD, 400 MHz) of **SM45** (schischkiniin, **194**)

Figure 134. ^1H-^1H COSY spectrum (CD_3OD, 400 MHz) of **SM45** (schischkiniin, **194**)

CHAPTER ONE

Introduction

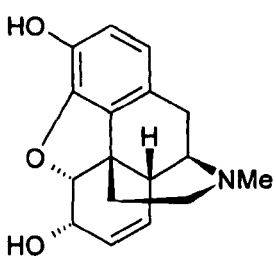
1 Introduction

1.1 Background

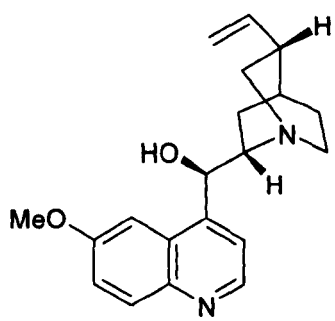
Natural Products, especially plants, have been used for the treatment of various diseases for thousands of years. Terrestrial plants have been used as medicines in Egypt, China, India and Greece from ancient time and an impressive number of modern drugs have been developed from them. The first written records on the medicinal uses of plants appeared in about 2600 BC from the Sumerians and Akkaidians.[1] The "Ebers Papyrus", the best known Egyptian pharmaceutical record, which documented over 700 drugs, represents the history of Egyptian medicine dated from 1500 BC. The Chinese *Materia Medica,* which describes more than 600 medicinal plants, has been well documented with the first record dating from about 1100 BC.[2] Documentation of the Ayurvedic system recorded in Susruta and Charaka dates from about 1000 BC.[3] The Greeks also contributed substantially to the rational development of the herbal drugs. Dioscorides, the Greek physician (100 A.D.), described in his work "De Materia Medica" more than 600 medicinal plants.[1] The World Health Organization estimates that approximately 80% of the world's inhabitants rely on traditional medicine for their primary health care.[4] Thirty-nine percent of the 520 new drugs approved between 1983 and 1994 were natural products and the proportion of antibacterials and anticancer drugs derived from natural products, was more than 60%.[2] Analysis of data on prescriptions dispensed from community pharmacies in the United States from 1959 to 1980 indicates that about 25% of them contained plant extracts or active principles derived from higher plants, and at least 119 chemical substances, derived from 90 plant species, can be considered important drugs currently in use in one or more countries.[5] Various types of new diseases have emerged recently, and microbes are becoming resistant to

current drugs. Plants can serve to combat diseases by providing new drugs in three different ways: firstly, by acting as new drugs which can be used in an unmodified state (e.g. vincristine from *Catharanthus roseus*), secondly, by providing chemical building blocks used to synthesise more complex molecules (e.g. diosgenin from *Dioscorea floribunda* for the synthesis of oral contraceptives) and thirdly, by indicating new modes of pharmacological action that allow complete synthesis of novel analogues (e.g. most of these semisynthetic analogues of penicillin from *Penicillium notatum*).[6]

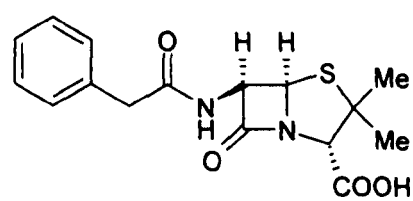
The discovery and development of new drugs from natural products requires close collaboration between chemists, pharmacologists, biochemists, taxonomists, zoologists and others. The National Cancer Institute (NCI), USA has established collaborations between organisations worldwide including Bangladesh, Brazil, China, Costa Rica, Iceland, Malaysia, Panama, Russia, South Africa and Zimbabwe for collection, extraction and isolation of natural products for their drug discovery program.[7]



1



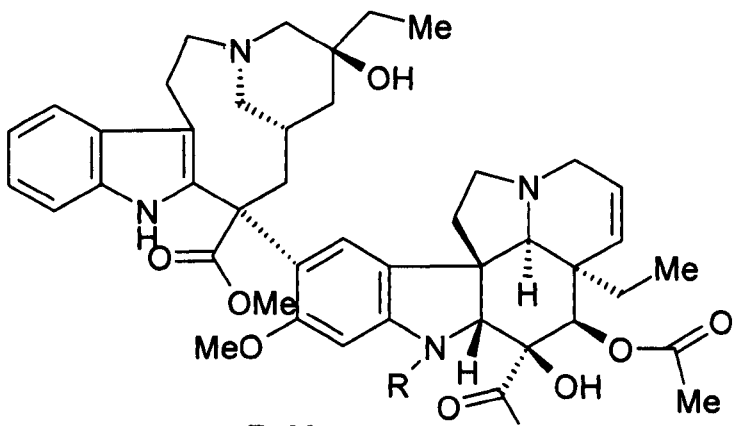
2



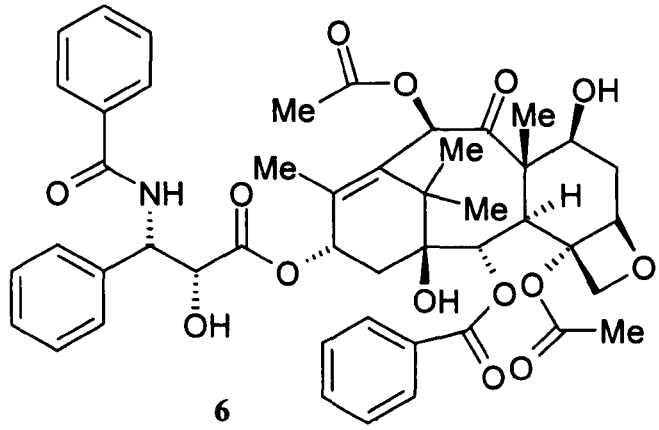
3

The isolation of morphine (**1**) from opium poppy (*Papaver somniferum*)[8] in 1816 by the German pharmacist Serturner inspired scientists to isolate active compounds

from plants for drug discovery. The French pharmacists Caventou and Pelletier reported the isolation of antimalarial drug quinine (**2**) from the bark of *Cinchona officinalis* in 1820.[9,10] The discovery of penicillin (**3**) from *Penicillium notatum*[11] by Fleming, Florey and Chain directed scientists to other natural sources in the hunt for novel bioactive compounds. The isolation of the vinca alkaloids, vinblastine (**4**) and vincristine (**5**) from *Catharanthus roseus* (formerly *Vinca rosea* Linn) introduced a new era of the use of plant material as anticancer agents. They were approved for the treatment of malignant diseases in 1974.[1] The discovery of Taxol® (**6**) from the bark of the Pacific yew, *Taxus brevifolia,* is another evidence of the success in natural product drug discovery. The structure of Taxol® was elucidated in 1971, yet was not approved for use until mid 1990's.[12] Taxol® is significantly active against ovarian cancer, advanced breast cancer, small and non-small cell lung cancer.[13]



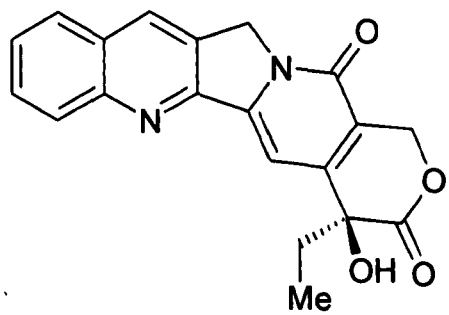
4 R=Me
5 R=CHO



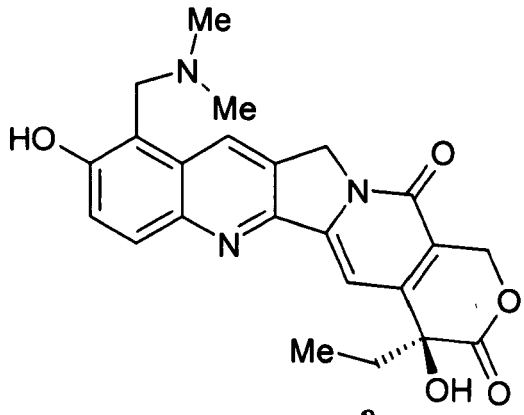
6

1.2 Plant Derived Anticancer Agents

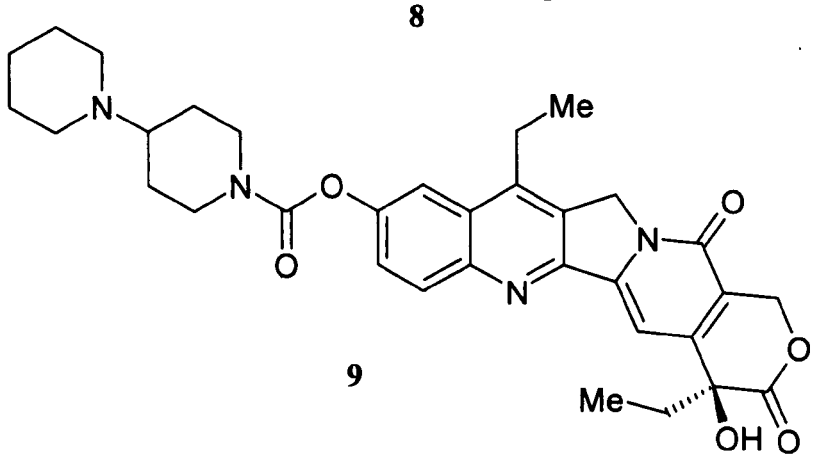
Plants have long been used in the treatment of cancer.[14] The National Cancer Institute collected about 35,000 plant samples from 20 countries and has screened around 114,000 extracts for anticancer activity. Of the 92 anticancer drugs commercially available prior to 1983 in the US and approved world wide between 1983 and 1994, 60% are of natural origin.[2] In this instance, natural origin is defined as natural products, derivatives of natural products or synthetic pharmaceuticals based on natural product models.[15]



7

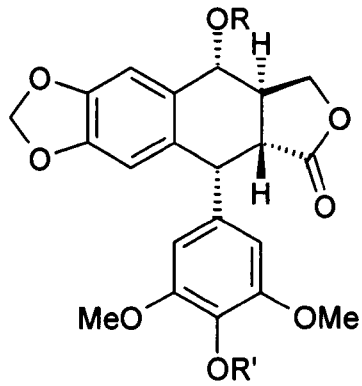


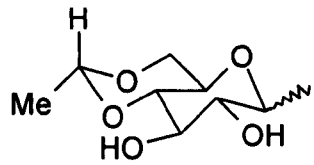
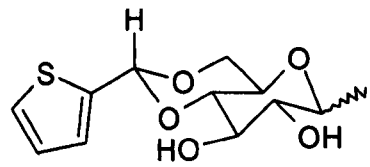
8



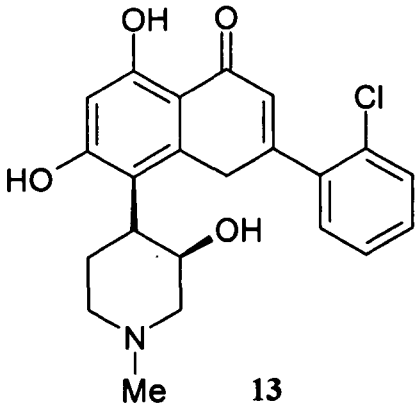
9

Camptothecin (**7**), isolated from the Chinese ornamental tree *Camptotheca acuminate,* was advanced to clinical trials by NCI in the 1970s but was dropped because of severe bladder toxicity.[16] Topotecan (**8**) and irinotecan (**9**) are semi-synthetic derivatives of 7 and have shown significant antitumour activity against ovarian and colorectal cancer, respectively.[17,18] Epipodophyllotoxin is an isomer of podophyllotoxin (**10**) which was isolated as the active antitumour agent from the roots of *Podophyllum peltatum* and *Podophyllum emodi.*[19] Etoposide (**11**) and teniposide (**12**) are two semi-synthetic derivatives of epipodophyllotoxin and have shown significant activity against small-cell lung carcinoma.[20]

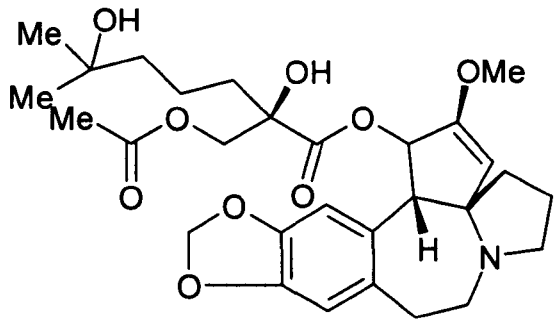


	R	R'
10	H	Me
11		H
12		H

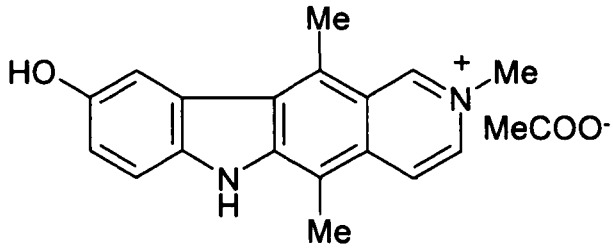
Flavopiridol (**13**) is a synthetic flavone, derived from the plant alkaloid rohitukine, which was isolated from *Dysoxylum binectariferum*.[21] It is currently in phase II clinical trials against a broad range of tumours.[22] Homoharringtonine (**14**), isolated from the Chinese tree *Cephalotaxus harringtonia* var. *drupacea,* has shown efficacy against various leukemias.[23,24] Elliptinium (**15**), a derivative of ellipticine, isolated from a Fijian medicinal plant *Bleekeria vitensis,* is marketed in France for the treatment of breast cancer.[25]



13



14



15

1.3 The Genus *Centaurea*

The genus *Centaurea* belongs to the family Asteraceae *alt.* Compositae and comprises about 500 species which are distributed mainly in Turkey, Iran, Greece, Algeria, North America and some parts of Asia.[26] *Centaurea*, which is the third largest genus of Asteraceae family in Turkey, is represented by 177 species, distributed mainly in the south-western, central, and eastern parts of the country and about 109 species are endemic (endemism ratio: 61.6%).[27] A higher endemism ratio implies that Turkey is one of the gene centres of this genus. Many species of this genus have traditionally been used to cure various ailments, e.g. diabetes, diarrhoea, rheumatism, malaria, hypertension etc.[28] In Turkey, *C. cyanus* and *C. scabiosa* are used against coughs and as liver-strengthening, itch eliminating, and ophthalmic

remedies.[29,30] *C. calcitrapa*, *C. solstitialis* and *C. melitensis* were found to have hypoglycemic effects and *C. calcitrapa*, *C. iberica* and *C. jacea* to possess an antipyretic effect.[31,32] *C. seridis* L. var. *maritime* showed antidiabetic activity in streptozotocin-induced diabetic rats and β-sitosterol 3-*β* D-glucoside was found responsible for the activity of this plant.[33] Aqueous extracts of *C. aspera* L. showed hypoglycemic activity in normal and alloxan-diabetic rats.[34] The roots of *C. ornate* has been used as traditional medicine in Spain as depuratives and cholagogues.[35] Ingestion of *C. repens* by horses was reported to cause movement disorder and neurodegenerative disorder, which is comparable to Parkinson's disease with respect to symptoms and pathology.[36] Several plant secondary metabolites including sesquiterpene lactones, lignan, alkaloids, flavonoids, etc., have been reported from various species of *Centaurea*, and many of those compounds have also been found to have different kinds of potential pharmacological properties.[28]

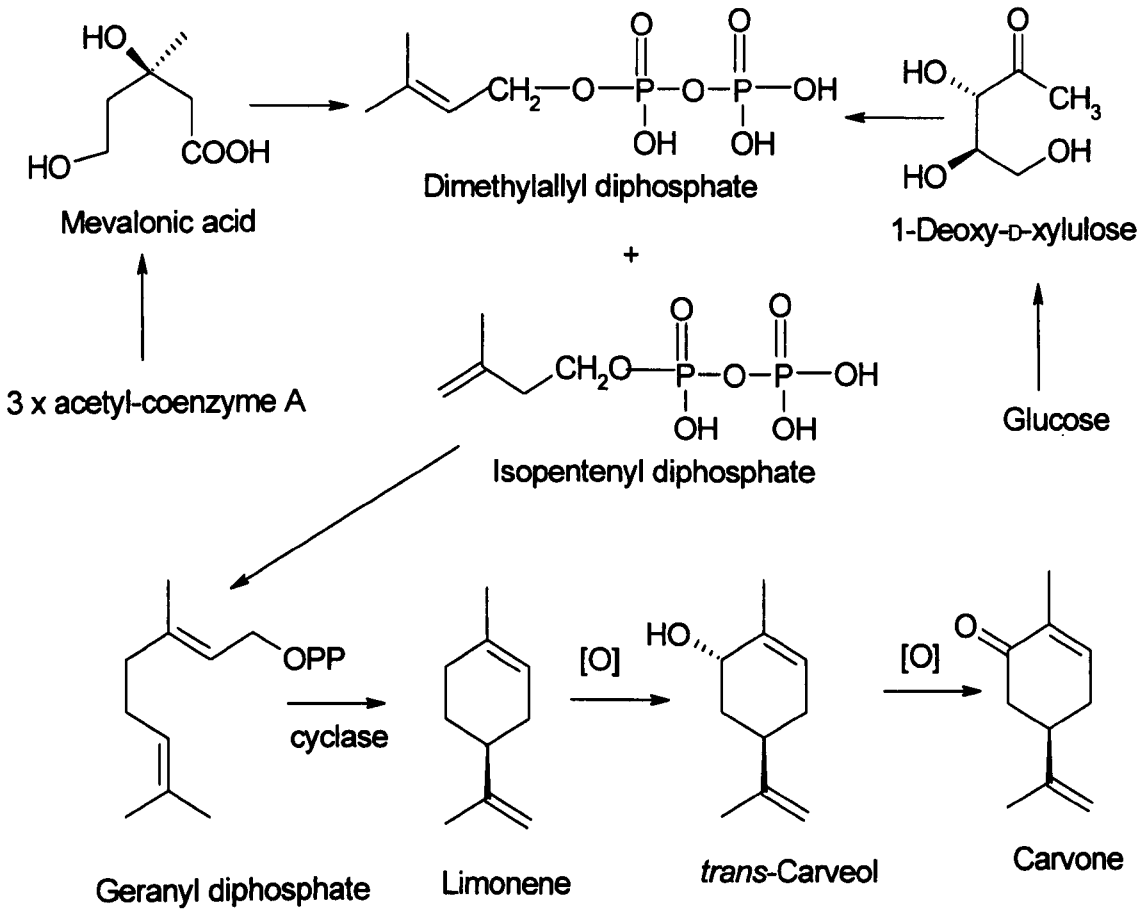
1.4 Phytochemistry of the Genus *Centaurea*

Centaurea is a big genus, but only a small portion has been investigated for secondary metabolites to date. An extensive summary of secondary metabolites isolated from *Centaurea* species is presented in **Table 1**. Sesquiterpene lactones are the most investigated secondary metabolites from this genus and the alkaloids are the least. Sesquiterpene lactones, chlorohyssopifolin A (**68**), cynaropicrin (**47**) and deacylcynaropicrin (**46**), from this genus exhibit cytotoxic and cytostatic activities.[37-39] Sesquiterpene lactones isolated from the aerial parts of *C. zuccariniana*, *C. achaia* and *C. thessala* were found to have cytotoxic effect in five human cell lines.[40]

1.4.1 Sesquiterpenoids

The terpenoids are formed by head-to-tail condensation of 5-carbon isoprene precursors, dimethylallyl diphosphate (DMAP) and isopentenyl diphosphate (IPP).[41] Two such units form monoterpenoids (**Scheme 1**) and three such units form sesquiterpenoids. Thus, terpenoids are classified according to the number of isoprene units from which they are biogenetically derived. In monoterpenoid biosynthesis, the first 10-carbon intermediate, formed from the union of DMAP and IPP, is geranyl diphosphate which may undergo further enzymatic modification to yield acyclic monoterpenes. However, monocyclic and bicyclic monoterpenoids require a cyclising enzyme to complete their biosynthesis. Lower terpenoids are volatile compounds and approximately 20000 different terpenes and terpenoids have been characterised from plant kingdom so far.[42]

Sesquiterpenoids are 15 carbon compounds formed by the assembly of three isoprenoid units and are founds in many living systems but particularly in higher plants. The biosynthesis of sesquiterpenoids initiates with the formation of farnesyl diphosphate from the condensation of geranyl diphosphate and IPP. This then undergoes a variety of enzymatic modifications to produce different sesquiterpenoids found in nature.[43] Sesquiterepenoids co-occur with monoterpenoids in plant essential oils and can usually be distinguished by their higher boiling points. The genus *Centaurea* produces a variety of compounds, the majority of them are sesquiterpene lactones. Biosynthesis of carvone, a monoterpenoid from *Mentha spicata* is shown in **Scheme 1**.[41]



Scheme 1. Biosynthesis of carvone (mono terpenoid) from *Mentha spicata*

Table 1. Plant secondary metabolites from *Centaurea* species

Name of species	Plant parts	Compounds	Type	References
C. adjarica Alb. Syn. *C. koenigii* Sosn.	Aerial	Centaurepensin (**68**) Chlorojanerin (**61**) Janerin (**78**) Repin (**77**)	Sesquiterpene lactones	44
C. aegialophila Wagenitz	Aerial	Cnicin (**17**) Dehydromelitensin (**87**) Elemanolide (**88**)	Sesquiterpene lactones	45
C. aegyptica L.	Aerial	Stigmasterol (**156**) β-sitosterol (**155**) Taraxasterol	Steroids	46
		Chlorojanerin (**61**)	Sesquiterpene lactones	
C. affinis Friv.	Aerial	Cnicin (**17**) Salonitenolide (**16**)	Sesquiterpene lactones	47
		Apigenin (**108**) Eupatorin Salvigenin (**115**)	Flavonoids	
		Arctigenin (**139**) Matairesinol (**140**)	Lignans	
C. alba L.	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Salonitenolide (**16**)	Sesquiterepen lactones	48
		Dihyrdosalonitenolide (**32**)		
		Salonitenolide 8-*O*-(4'-acetoxy-5'-hydroxy)-angelate (**24**)		
C. alexandrina Del.		Apigenin (**108**) Luteolin (**112**) Quercetin (**125**) Rutin	Flavonoids	49

C. achaia Boiss. & Heldr.	Aerial	8α-*O*-(3-Hydroxy-2-methylenepropanoyl-dehydromelitensin (**90**)	Sesquiterpene lactones	40
C. americana Nutt.	Seed	Cynaropicrin (**47**)	Sesquiterpene lactones	50-52
		Arctiin (**137**) Matairesinoside (**138**)	Lignans	
		20-Hydroxyecdysone (**169**)	Phytoecdysteroid	
		3,4-Dihydroxycinnamic acid	Phenylpropane derivatives	
		β-Sitosterol (**155**)	Steroids	
C. aspera L. var. *subinermis* DC.	Aerial	Benzoic acid *p*-Hydroxy benzoic acid	Aromatic compounds	53-54
		p-Coumaric acid	Phenyl propane derivatives	
		Scopoletin (**154**)	Coumarins	
		Eupafolin (**113**) Hispidulin (**117**) Jaceosidin (**118**) Pectolinarigenin (**116**)	Flavonoids	
		Eudesmanolide A (**103**) Eudesmanolide B (**104**) Dehydromelitensin (**87**) Isomelitensin (**86**) Melitensin (**82**) Stenophyllolide (**28**) Dihydrostenophyllolide (**29**)	Sesquiterpene lactones	
C. aspera L. var *stenophylla*	Flower	Benzoic acid *p*-Hydroxybenzoic acid	Aromatic compounds	55
		Apigenin (**108**) Eupafolin (**113**) Ethyl 7-*O*-apigenin-glucuronate	Flavonoids	
		β-Sitosterol (**155**) Stigmasterol glucoside	Steroids	

C. aspera L. var *stenophylla*	Flower	Melitensin (**82**) Dehydromelitensin (**87**) Stenophyllolide (**28**)	Sesquiterpene lactones	55
C. aspera L. subsp. *aspera* and *C. aspera* L. subsp. stenophylla (Dufour) Nyman (syn. C. stenophylla Dufour)	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Elemacarmanin (**92**) Germacranolide **25** Germacranolide **32** Germacranolide **36** Onopordopicrin (**26**) Salonitenolide (**16**) 1-Hydroxy-3-methyl-2-butenoic acid ester of salonitenolide (**22**) Salonitenolide 8-*O*-(4'-acetoxy-5'-hydroxy)-angelate (**24**) Stenophyllolide (**28**) Dihydrostenophyllolide (**29**)	Sesquiterpene lactones	56
C. attica ssp. *attica*	Aerial	Atticin (**97**) Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) 8α-[4-acetoxy-5-hydroxy-angelate]-salonitenolide (**24**) 8α-(3,4-dihydroxy-2-methylene-butanoyloxy)-dehydromelitensin (**88**) Methyl 8α-(3,4-dihydroxy-2-methylene-butanoyloxy)-6α,15-dihydroxyelema-1,3,11 (13)-trien-12-oate (**91**) Malacitanolide (**95**) 4'-acetoxy-malacitanolide (**90**) 8α-(hydroxy-4-*epi*-Sonchucarpolide (**94**)	Sesquiterpene lactones	57-58
C. bella. Trautv.	Aerial	Cebellin J	Sesquiterpene lactones	44

C. bella Trautv.	Aerial	Cebellin N (**54**) Cebellin O (**55**) Cynaropicrin (**47**) 19-Desoxychlorojanerin (**62**) 8-Deacyloxy-8α-[2-methylacryloxy]-Subluteolide (**78**)	Sesquiterpene lactones	59
C. behen L.	Leaves	Aguerin B (**50**) Cynaropicrin (**47**) Desacylcynaropicrin (**46**) Grosshemin (**58**) Mono acetyl solstitalin A (**59**)	Sesquiterpene lactones	60
		Circimaritin (**110**)	Flavonoids	
C. bombycina	Aerial	Cnicin (**17**) Salonitenolide (**16**) Salonitenolide 8-*O*- (4'-acetoxy-5'-hydroxyangelate (**24**)	Sesquiterpene lactones	61
C. bracteata Scop.	Aerial	Axillarin (**126**) Axillarin 7-glucoside Bracteoside (**135**) Centabractein (**134**) Centradixin (**133**) Hispidulin-7-sulphate Jaceidin (**120**) Jacein (**119**) Luteolin (**112**) Nepetin 7-sulphate Patuletin 7-sulphate Nepetin (**113**) Centaureidin Hispidulin (**117**) Apigenin 7-glucuronide	Flavonoids	62,63
		3,4 Dicaffeoylquinic acid 3-Caffeoylquinic acid methyl ester	Quinic acid derivatives	
		Caffeic acid Protocatechuic β-Sitosterol-3-glucoside 5- Caffeoylquinic acid	Phenylpropane derivatives	

C. calcitrapa L.	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Melitensin (**82**) Salonitenolide (**16**) 11β, 13-α Dihydroxygermacra-1(10)E,4E-diene-6β,11βH-12,6-olide-dihydro-salonitenolide (**31**) 11α, 13β-Dihydroxygermacra-1(10)E,4E-diene-6β,11βH-12,6-olide-dihydro-salonitenolide (**32**)	Sesquiterpene lactones	49,64,65
		Arctigenin (**139**) Matairesinol (**140**) Pinoresinol (**144**) 7' (*S*)-hydroxyarctigenin (**142**)	Lignans	
		Astragalin (**123**) Luteolin (**112**) Quercetin (**125**) Kaempferol (**122**) Apigenin (**108**) Naringenin	Flavonoids	
		p-Hydroxyacetophenone Chalcone (**167**) *cis*-isomer of **167** (**168**)		
C. canariensis		Aguerin A (**51**) Aguerin B (**50**) Cynaropicrin (**47**) Deacylcynaropicrin (**46**)	Sesquiterpene lactones	66
C. castellana Boiss	Aerial	Artemissifolin Dehydromelitensin (**87**) 11β,15-Dehydroxysaussurea (**85**)	Sesquiterpene lactones	67
Centaurea cineraria L. ssp. *busambarensis*	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Dehydromelitensin (**87**)	Sesquiterpene lactones	68

C. cineraria ssp. umbrosa	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Elemandienolide **88** Salvigenin (**115**) Eupatilin (**121**) Jaceosidin (**118**)	Sesquiterpene lactones Flavonoids	69
C. conifera L.	Aerial	Chlorojanerin (**61**) Chlorohyssopifolin A (**68**) Centaurepensin 17' epimer (**69**) Janerin (**78**) Repin (**77**) Subluteolide (**76**)	Sesquiterpene lactones	48,68
C. collina L.	Aerial	3β-Hydroxy-8α-epoxymethylacriloiloxy-4 (15)10 (14),11(13)-trien-(1αH),(5αH)-guaian-6,12-olide (**43**) 11β,13-Dihydro derivative of guaianolide **43** (**44**)	Sesquiterpene lactones	70
C. cyanus L.		Centcyamine (**146**) *cis*-Centcyamine (**147**) Moschamine (**148**) *cis*-Moschamine (**149**) Scopoletin (**154**) Umbelliferone (**153**) Apigenin (**108**) Apigenin 4'-*O*-β-D-glucoside Hispidulin (**117**) Isorhamnetin (**127**) Isorhamnetin 7-*O*-β-D-glucoside Kaempferol 7-*O*-β-D-glucoside Kaempferol (**122**) Luteolin (**112**) Quercetin (**125**)	Alkaloids Coumarins Flavonoids	71-73

C. cyanus L.	Aerial	Caffeic acid	Phenyl propane derivatives	72
		Chlorogenic acid (**168**) Neochlorogenic Isochlorogenic acids	Quinic acid derivatives	
C. diffusa Lam.	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**)	Sesquiterpene lactones	74
		1,2-Diangelyloxyglucose (**163**) 1-(3-Methylbutanoyloxy) 2-angelyloxyglucose (**164**)		
C. furfuracea Coss. Et Kral	Aerial	Apigenin (**108**) Apigenin 7-*O*-glucoside Apigenin 7-*O*-methylglucuronide Cirsimaritin (**110**) Hispidulin (**117**) Hispidulin 7-*O*-glucoside Patelutin 7-*O*-glucoside	Flavonoids	75-76
C. glomerata Vahl.	Aerial	Arctigenin (**139**)	Lignans	49,77
		1-Hydroxy-3-methyl-2-butenoic acid ester of salonitenolide (**22**)	Sesquiterpene lactones	
		Isomer of **22** (**23**)		
		11,13-Dihydro butenoic acid ester of salonitenolide (**33**)		
		Dihydroonopordopicrin (**34**)		
		Apigenin (**108**) Luteolin (**112**) Quercetin (**125**)	Flavonoids	
C. granata L.	Aerial	8α-Hydroxy-11β,13-dihydro onopordaldehyde (**102**)	Sesquiterpene lactones	78
		5-Hydroxy-6,7,3',4'-tetramethoxyflavone	Flavonoids	

C. granatensis	Aerial	Cnicin (**17**)	Sesquiterpene lactones	61
C. hermannii F. Hermann	Aerial	Cynaropicrin (**47**) Chlorojanerin (**61**) Janerin (**78**)	Sesquiterpene lactones	79
		β-Sitosterol (**155**)	Steroids	
C. horrida Bad.	Aerial	Horridin (**136**)	Flavonoids	80
C. incana	Aerial	Janerin (**78**) Onopordopicrin (**19**) Repin (**77**)	Sesquiterpene lactones	61,81- 82
		3,5'-Dimethyl myricetin 7-*O*-glucoside 2''-*O*-glucosyl-6-*C*-glucosylapigenin 3'Methyl myricetin 7-*O*-glucoside 6-Methoxyapigenin 6-Methoxyluteolin 6-Methoxykaempferol 6-Methoxyquercetin Apigenin-7-methylgalacturonide Hispidulin 7-*O*-glucoside 6-Methoxyquercetin 7-*O*-glucoside	Flavonoids	
C. isaurica Hub. Mor.	Aerial	Janerin (**78**)	Sesquiterpene lactones	83
		β-Sitosterol 3-*O*-glucoside	Steroids	
		Protocatechuic acid Scopoletin (**140**)	Coumarins	
		Chlorogenic acid (**168**)	Quinic acid derivatives	
		Jacein (**119**) Centaurein Kaempferol-3-*O*-β-glucopyranoside (**123**) quercetin-3-*O*-β-glucopyranoside (**126**)	Flavonoids	
		Arctiin (**137**)	Lignans	

C. inermis Velen	Aerial	Astragalin (**123**) Nepetin (**113**)	Flavonoids	84
C. kotschyi (var. kotschyi) Boiss.	Aerial	Cynaropicrin (**47**) Deacylcynaropicrin (**46**) Linichlorin B (**48**) Linichlorin B derivative (**49**)	Sesquiterpene lactones	85
C. pabotii Wagenitz	Aerial	Aguerin A (**51**) Dihydrodeacylcyaropicrin (**53**) Deacylcynaropicrin 8-*O*-(3-hydroxy-2-methylpropionate) (**52**)	Sesquiterpene lactones	64
C. pallescens Del.		Astragalin (**123**) Apigenin (**108**) Kaempferol (**122**) Luteolin (**112**) Quercetin (**125**)	Flavonoids	49
C. paniculata ssp. *castellana*	Aerial	Stigmasterol (**156**) β-sitosterol (**155**) Taraxasterol	Steroids	86
		Cirsiliol (**114**) Cirsimaritin (**110**)	Flavonoids	
		Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**)	Sesquiterpene lactones	
C. persica Boiss.	Aerial	Stigmasterol (**156**) Taraxasterol Lupeol (**157**)	Steroids	87
		Matairesinol (**140**)	Lignans	
C. monticola		Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**)	Sesquiterpene lactones	61
C. macrocephala Muss. Puschk. Ex Willd.	Seed	Arctigenin (**139**)	Lignans	88,89
		Isoorientin (**128**) Isovitexin (**129**) Isoquercitrin (**126**) Trifolin (**124**) Rutin	Flavonoids	

C. maroccana		Cnicin (**17**) Cnicin 4'-*O*- acetate (**18**) Salonitenolide (**16**) Dihydrosalonitenolide (**32**)	Sesquiterpene lactones	61
C. marschalliana Sp reng.	Aerial	Chlorojanerin (**61**) Cebellin D Acroptilin (**79**) Janerin (**78**)	Sesquiterpene lactones	44
C. malacitana		Cnicin (**17**) Cnicin 4'-*O*- acetate (**18**) Malacitanolide (**95**) 8-*O*-(-acetoxyangeloyl) salonitenolide (**24**)	Sesquiterpene lactones	90
C. melitensis L.	Aerial	Arctiopicrin (**21**) Melitensin (**82**) Melitensin-β-hydroxyisobutyrate (**83**) Dehydromelitensin-β-hydroxyisobutyrate (**84**) Onopordopicrin (**19**) Salonitenolide (**16**)	Sesquiterpene lactones	91,92
C. moschata L.	Seeds	Moschamine (**148**) *cis*-Moschamine (**149**) Moschamindole (**150**) Moschamindolol (**151**) Moschamide (**152**) Moschatine (**162**)	Alkaloid	93-96
		20-Hydroxyecdysone (**169**)	Ecdysteroids	
C. musimomum	Aerial	3 Oxo-4α-acetoxy-15-hydroxy-1αH, 5αH, 6βH, 7αH, 11βH-gui-10(14)-ene-6,12-olide (**56**) 3 Oxo-4α-hydroxy-15-hydroxy-1αH, 5αH, 6βH, 7αH, 11βH-gui-10(14)-ene-6,12-olide (**57**)	Sesquiterpene lactones	97

C. napifolia L.	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Melitensin (**82**)	Sesquiterpene lactones	98,99
		Lappaol A (**143**)	Lignans	
		Cirsimaritin (**110**) Hispidulin (**117**) Quercetin (**125**)	Flavonoids	
		1,2-Diacylated glucose (**165**)		
C. nervosa Willd.	Roots and Flower	Apigenin (**108**) Kaempferol 3-methyl ether Jaceidin (**120**) Hispidulin (**117**) Jaceosidin (**118**)	Flavonoids	100
C. nicaensis	Aerial	Amarin (**20**) Cnicin (**17**) Dihydroamarin (**35**) Elemadienolide (**93**) 11β(H),13-Dihydrocnicin Melitensin (**82**) Onopordopicrin (**19**) 11β(H),13-Dihydrosalonitenolide (**32**)	Sesquiterpene lactones	101,102
		Lappol A (**143**)	Lignans	
C. nicolai Bald.	Aerial	Kandavanolide (**39**) Salograviolide A (**38**)	Sesquiterpene lactones	103,104
		3-*O*-deacetyl-9-*O*-acetylsalograviolide A (**41**)		
		9-*O*-acetylsalograviolide A (**42**) Salograviolide B (**45**)		
		Matairesinol (**140**)	Lignans	
		Arctinal	Dithiophenes	

C. nigra L.	Seed	Arctigenin (**139**) Arctiin (**137**) Matairesinol (**140**) Matairesinoside (**138**) Thujaplicatin methyl ester (**141**)	Lignans	105
		Moschamine (**148**) *N*-(*p*-Coumaroyl)-serotonin (**145**)	Alkaloids	
C. orphanidea Heldr. & Sart. Ex Boiss.	Aerial	Salonitenolide (**16**) Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Malacitenolide (**97**) 8α-*O*-4-*epi*-Sonchucarpolide (**94**)	Sesquiterpene lactones	106
		8α-*O*-(4-Acetoxy-3-hydroxy-2-methylene-butanoyloxy)-4-*epi*-sonchucarpolide (**96**)		
		8α-*O*-(3, 4-Dihydroxy-methylene-butanoyloxy)-dehydromelitensin (**88**)		
		Apigenin (**108**) Cirsimaritin (**110**) Luteolin (**112**) 3-Methoxykaempferol	Flavonoids	
		Pinoresinol (**144**)	Lignans	
C. phaeopappoides Bordz.	Aerial	Chlorojanerin (**61**) Cynaropicrin (**47**) Janerin (**78**)	Sesquiterpene lactones	44
C. phrygia	Roots, Green parts , Flower	Apigenin (**108**) Kaempferol 3-methyl ether Kaempferol 3,6-dimethyl ether Jaceidin (**120**) Centaureidin Hispidulin (**117**) Jaceosidin (**118**)	Flavonoids	100
		Arctigenin (**139**)	Lignans	

C. pseudoscabiosa subsp. *Pseudoscabiosa* Boiss. et Buhse		Chrysin (**107**) Chrysin 7-*O*-glucuronide Chrysin 6-*C*-glucoside Chrysin 8-*C*-glucoside Chrysin 7-glucuronide, Hispidulin (**117**) Luteolin 7-glucoside, Pinocembrin 7-*O*-α-arabinopyranosyl-(1→2)-β-glucopyranoside (**130**) Chrysin 7-*O*- β – glactopyranuronoside (**131**) Baicalein 6-methyl ether Baicalein 6-methylether-7-*O*- β – galactopyranuronoside (**132**)	Flavonoids	107
C. ptosimopappoides Wagenitz	Roots	17β, 21β-Epoxy-16α-ethoxy-hopan-3β-ol (**159**) 3 β-acetoxyhop-17(21)ene (**158**) 17β,21β –Epoxyhopan-3 β-ol (**160**) 3- β-Acetoxy-17,24-dioxo-baccharane (**161**)	Triterpenes	108
	Aerial	11,13-Dihydro-desacetylcynaropicrin (**53**) Cynaropicrin (**47**) Scopoletin (**154**) Stigmasterol (**156**)	Sesquiterpene lactones Coumarins Steroids	

C. raphanina ssp. *Mixta* (DC.) Runemark	Aerial	Cnicin (**17**)	Sesquiterpene lactones	109
		Apigenin (**108**)		
		Eriodictyol 7-*O*-β-D-glucoside Kaempferol 3-*O*-β-D-glucopyranoside Apigenin 7-O-β-D-glucopyranoside Apigenin 4'-O-β-D-glucopyranoside 7,4 Dimethoxykaempferol Eriodictyol	Flavonoids	
		Matairesinol (**140**)	Lignans	
C. repens L.	Aerial	Cynaropicrin (**47**) Janerin (**78**) Repin (**71**)	Sesquiterpene lactones	110
C. salonitana	Aerial	Salonitenolide (**16**) Salonitolide (**37**) Salograviolide A (**38**) Salograviolide B (**45**) Kandavanolide (**39**) Salograviolide C (**40**) Aguerin A (**51**)	Sesquiterpene lactones	111-114
Centaurea scabiosa L.	Seed	Arctigenin (**139**) Matairesinol (**140**) Matairesinoside (**138**) 7 (*S*)-hydroxy-arctigenin (**142**)	Lignans	115
C. scoparia Sieb.		Chlorojanerin (**62**) Chloroscoparin (**63**) 8-Deacylcentaurepensin 8-*O*-(4-hydroxy)-tiglate(**64**) Chlorohyssopifoli B (**65**) Diain (**70**) Janerin (**78**) Cynaropicrin (**47**) Deacylcynaropicrin (**46**) Janerin (**78**) Cynaropicrin (**47**) Deacylcynaropicrin (**46**)	Sesquiterpene lactones	116-121

C. scoparia Sieb.	Aerial	4-β-(Chloromethyl)-3β,4α-dihydroxy-8α-(3-formyl-2-methyl-propenoyloxy)-1αH,5αH,6βH, 7αH-guai-10(14),11(13)-dien-6,12-olide (**66**) 4-β-(Chloromethyl)-3β,4α-dihydroxy-8α-(sarracenoyloxy)-1αH,5αH,6βH,7αH-guai-10(14),11(13)-dien-6,12-olide (**67**) 8α-Hydroxy-3β-(benzoyloxy)-1αH,5αH,6βH,7αH-guai-4(15),10(14)11(13)-trien-6,12-olide (**71**) 3β-Hydroxy-8α-(3,4-dimethoxybenzoyloxy)-11β,13-dihydro-1αH,5αH,6βH,7αH-guai-4(15),10(14)dien-6,12-olide (**72**) 3β,8α-*O*-Di-(4-hydroxytigloyl)-1αH,5αH,6βH,7αH-guai-4(15),10(14),11(13)triene-6,12-olide (**73**) Cebellin F 8-Angelyloxy-3-hydroxy guai-3(15),10(14),11(13)triene-6,12-olide (**75**) Desacylcynaropicrin (**46**) 8α,4'-(Hydroxytiglinate)-8-deacyloxysubliteolide (**81**) 8-Desacylrepin (**80**) Chlorohyssopifolin A (**68**) Chlorohyssopifolin B (**65**)	Sesquiterpene lactones	116-121

C. scoparia Sieb.		Apigenin (**108**) Luteolin (**112**) Salvigenin (**115**) Cirsimaritin (**110**)	Flavonoids	118
		Matairesinol (**140**) Acrtigenin (**139**)	Lignans	
C. solstitialis L. subsp. *schouwii* (DC.) Dostal	Aerail	Cynaropicrin (**47**) Aguerin B (**50**)	Sesquiterpene lactones	122
		Arctiin (**137**) Matairesinol (**140**)	Lignans	
C. sonchifolia	Aerial	Artemisiifolin Acetyl-artemisiifolin	Flavonoids	67
C. sphaerocephala subsp. Lusitanica (Boiss. et. Reute) Nyman	Aerial	Cnicin (**17**) 3α, 15-Dihydroxycostunolide (**27**)	Sesquiterpene lactones	44
C. sphaerocephala subsp *sphaerocephala*	Aerial	Cnicin (**17**) Cnicin 4'-*O*-acetate (**18**) Epoxylactone (**30**) Dehydromelitensin (**87**)	Sesquiterpene lactones	123
		4'-Acetate of elemadienolide (**89**)		
C. sulphurea Willd.		Cnicin (**17**)	Sesquiterpene lactones	61,67
		Pectolinarigenin (**116**)	Flavonoids	
C. tagananensis Svent.	Aerial	Cynaropicrin (**47**) Onopordopicrin (**19**) Dehydromelitensin 8-*O*-4(4'-Hydroxymethacrylate) (**84**) Melitensin (**82**) Deacylcynaropicrin (**46**)	Sesquiterpene lactones	67
C. thracica (Janka) Hayek	Aerial	Chlorojanerin (**61**) Cynaropicrin (**47**) Janerin (**78**)	Sesquiterpene lactones	44

C. thessala ssp. d*rakiensis*	Aerial	8α-(3,4-dihydroxy-2-methylene-butanoyloxy melitensine (**88**) 8α-(3-hydroxy-4-acetoxy -2-methylene-butanoyloxy melitensine (**89**) Acetate derivative of **99** (**100**)	Sesquiterpene lactones	40, 58
C. tweediei Hook.	Aerial	Onopordopicrin (**19**) 8-(4'-Hydroxymethacryloxy)-15-oxohelianga-1(10), 4, 11 (13)-trien-6, 12-olide (**26**) 1-Hydroxy-8-methacryloxy-15-oxoeudesm-11(13)en-6,12-olide (**98**) Methyl 1,6-dihydroxy-8-methacryloxyeudesm-11(13)-en-15-oic acid-12-oate (**101**) 15-Hydroxy-8-(4'-hydroxymethacroyloxy)-10(14), 11 (13)-guaiadien-6,12-olide (**60**) Arctigenin (**139**) Matairesinol (**140**)	Sesquiterpene lactones Lignans	124
C. uniflora subsp. nervosa	Leaves	Santamarin (**105**) Reynosin (**106**) Epoxy guianolide (**81**)	Sesquiterpene lactones	125
C. urvillei Stepposa Wagenitz	Leaves	Apigenin (**108**) Cirsimaritin (**110**) Cirsiliol (**114**) Hispidulin (**117**) Luteolin (**112**) Nepetin (**113**) Salvigenin (**115**)	Flavonoids	126

C. virgata Lam	Aerial	Apigenin (**108**) Astragalin (**123**) Hispidulin (**117**) Jaceosidin (**118**) Nepetin (**113**) Isovitexin (**128**)	Flavonoids	84
C. zuccariniana DC.	Aerial	8α-hydroxysonchucarpolide (**99**)	Sesquiterpene lactones	40

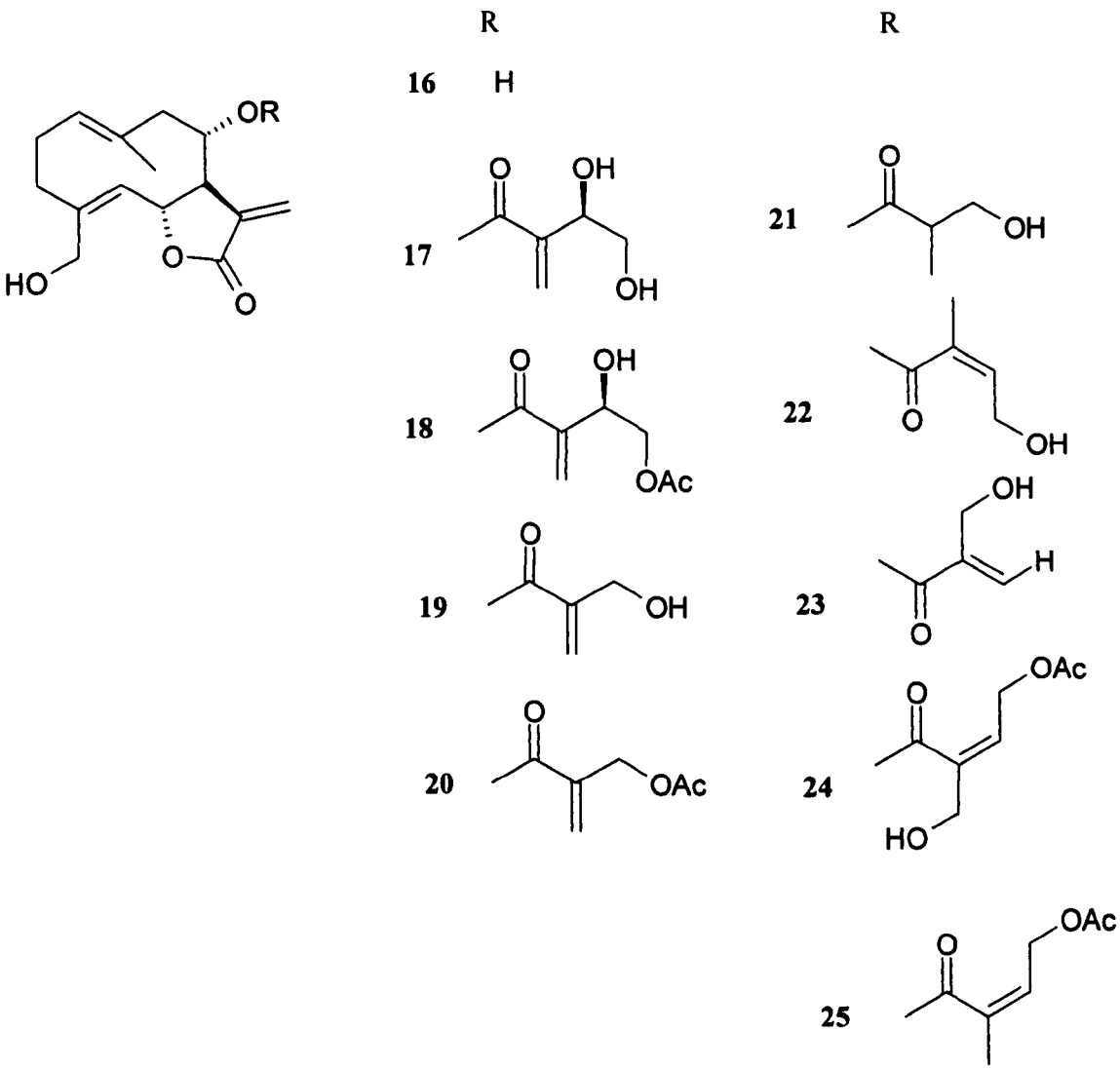
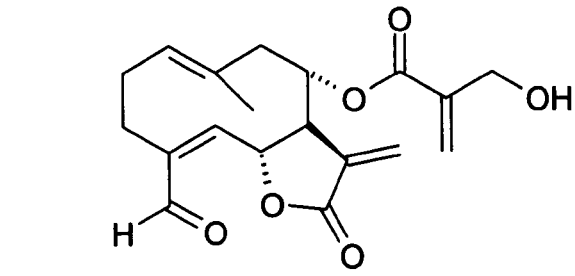
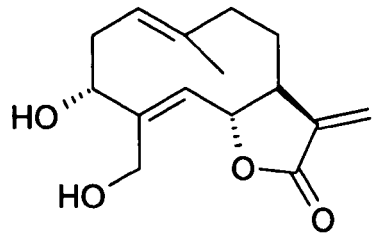


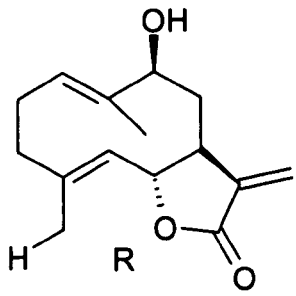
Figure 1. Structures of the sesquiterpene lactones from the genus *Centaurea*



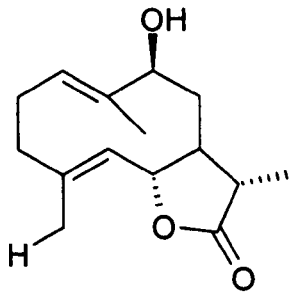
26



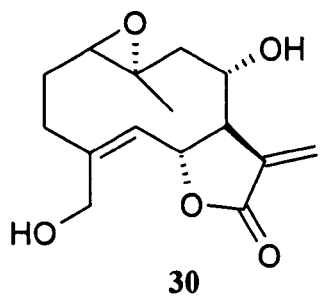
27



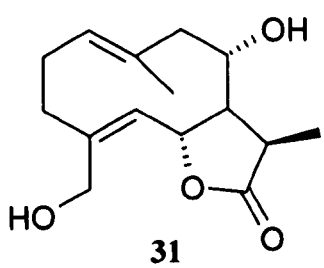
28



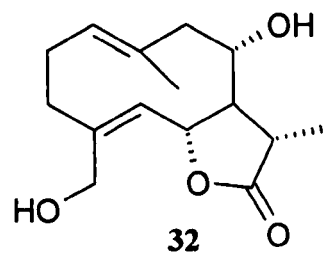
29



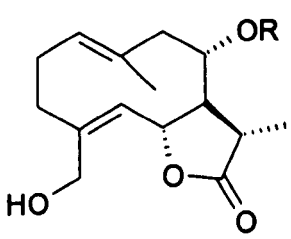
30



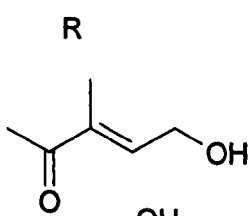
31



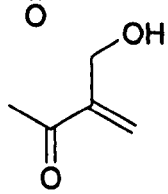
32



33



34



35

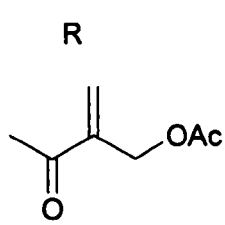
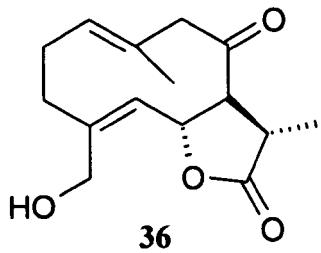
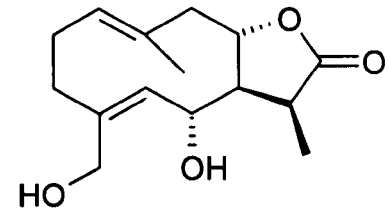


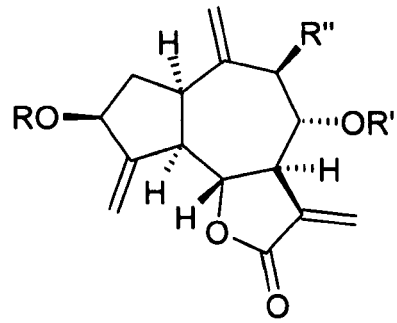
Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*



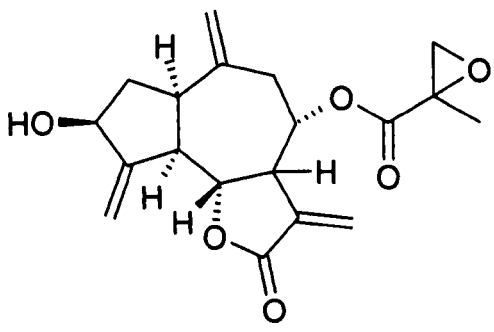
36



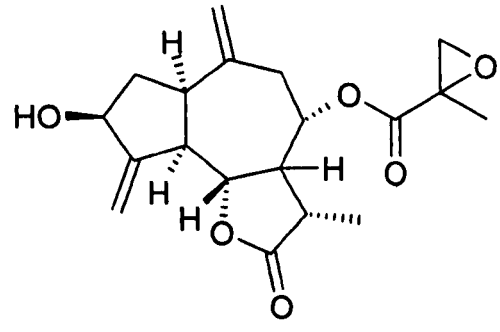
37



	R	R'	R''
38	Ac	H	OH
39	Ac	H	H
40	H	epoxy-methacryl	H
41	H	H	OAc
42	Ac	H	OAc

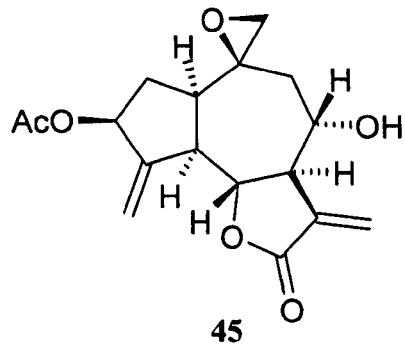


43

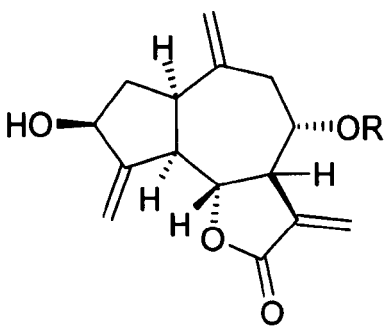


44

Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*



45



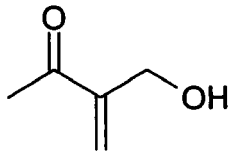
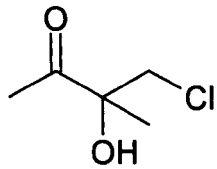
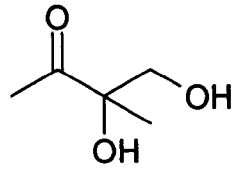
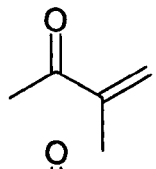
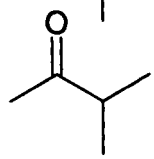
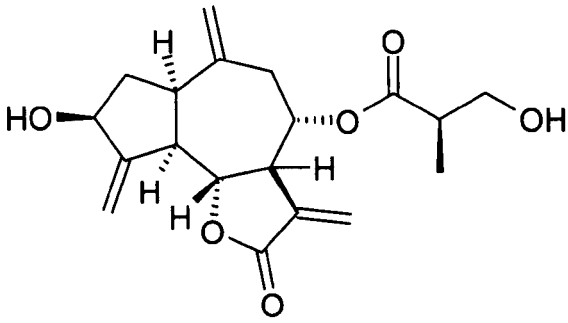
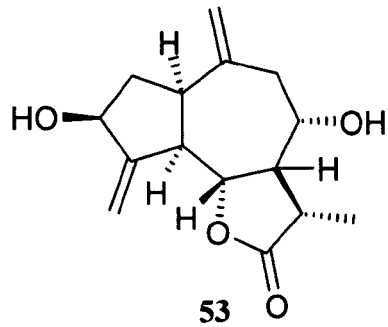
	R
46	H
47	
48	
49	
50	
51	

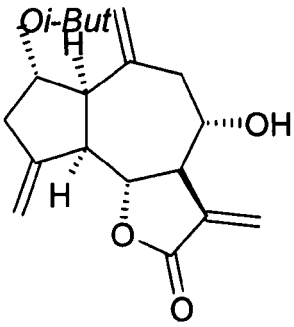
Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*



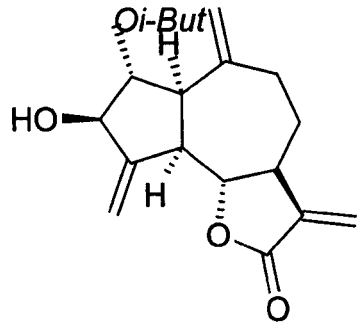
52



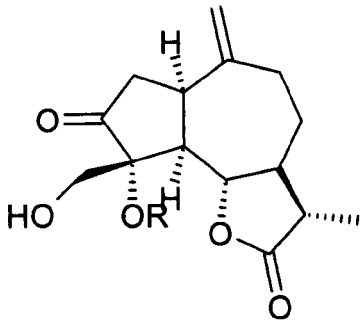
53



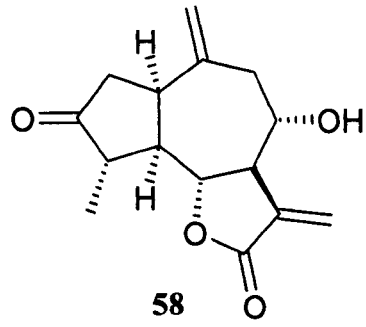
54



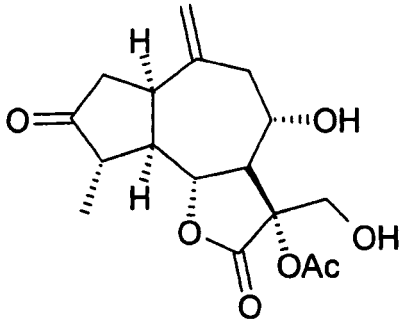
55



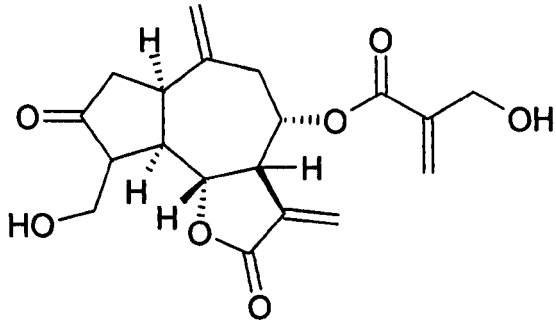
	R
56	Ac
57	H



58



59



60

Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*

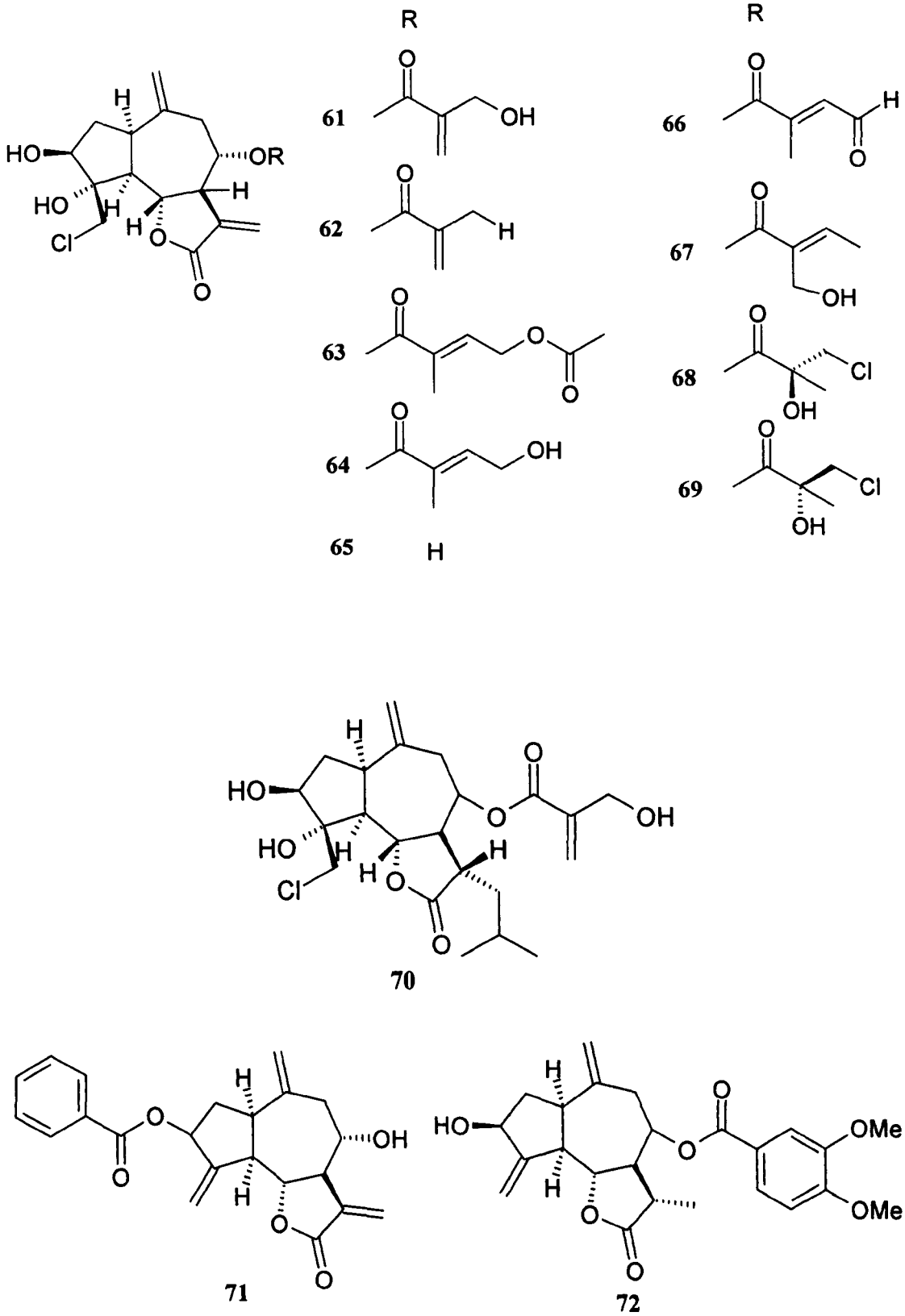


Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*

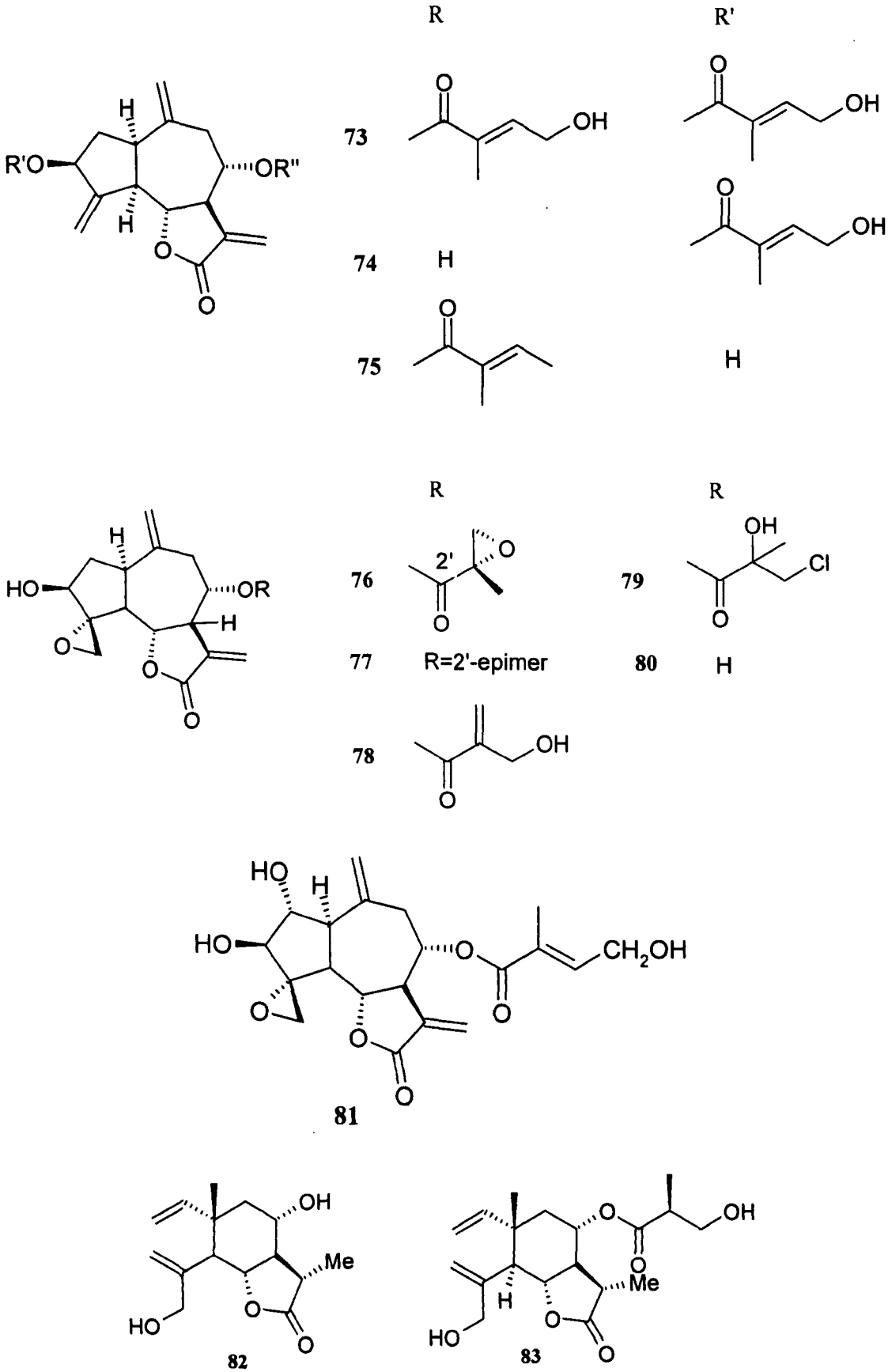
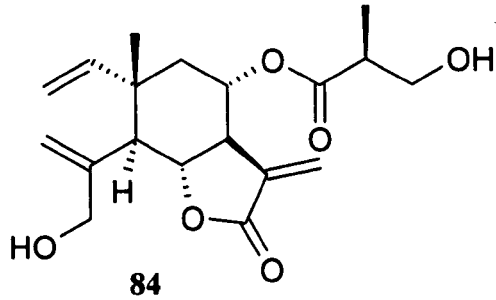
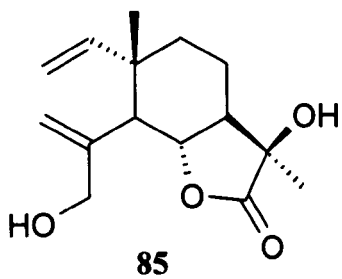


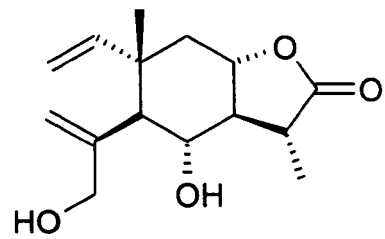
Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*



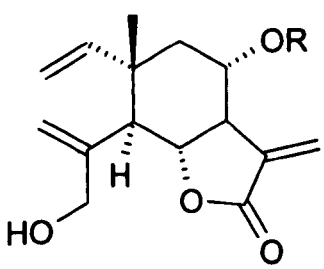
84



85



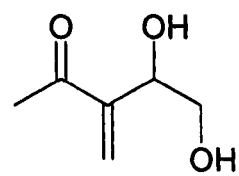
86



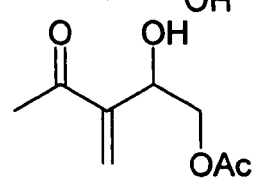
R

87 H

88



89



90

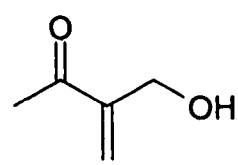
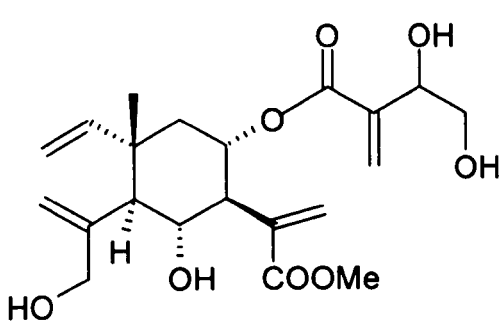
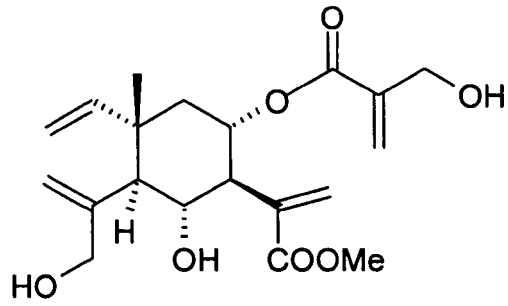


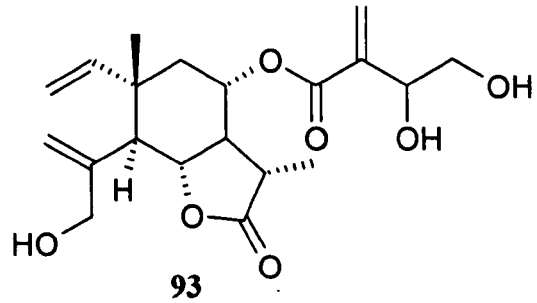
Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*



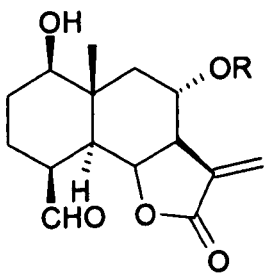
91



92

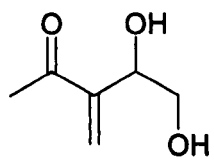


93

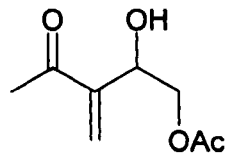


R

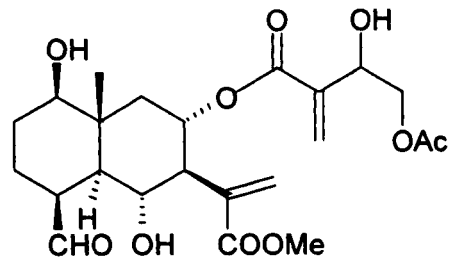
94 H



95



96



97

Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*

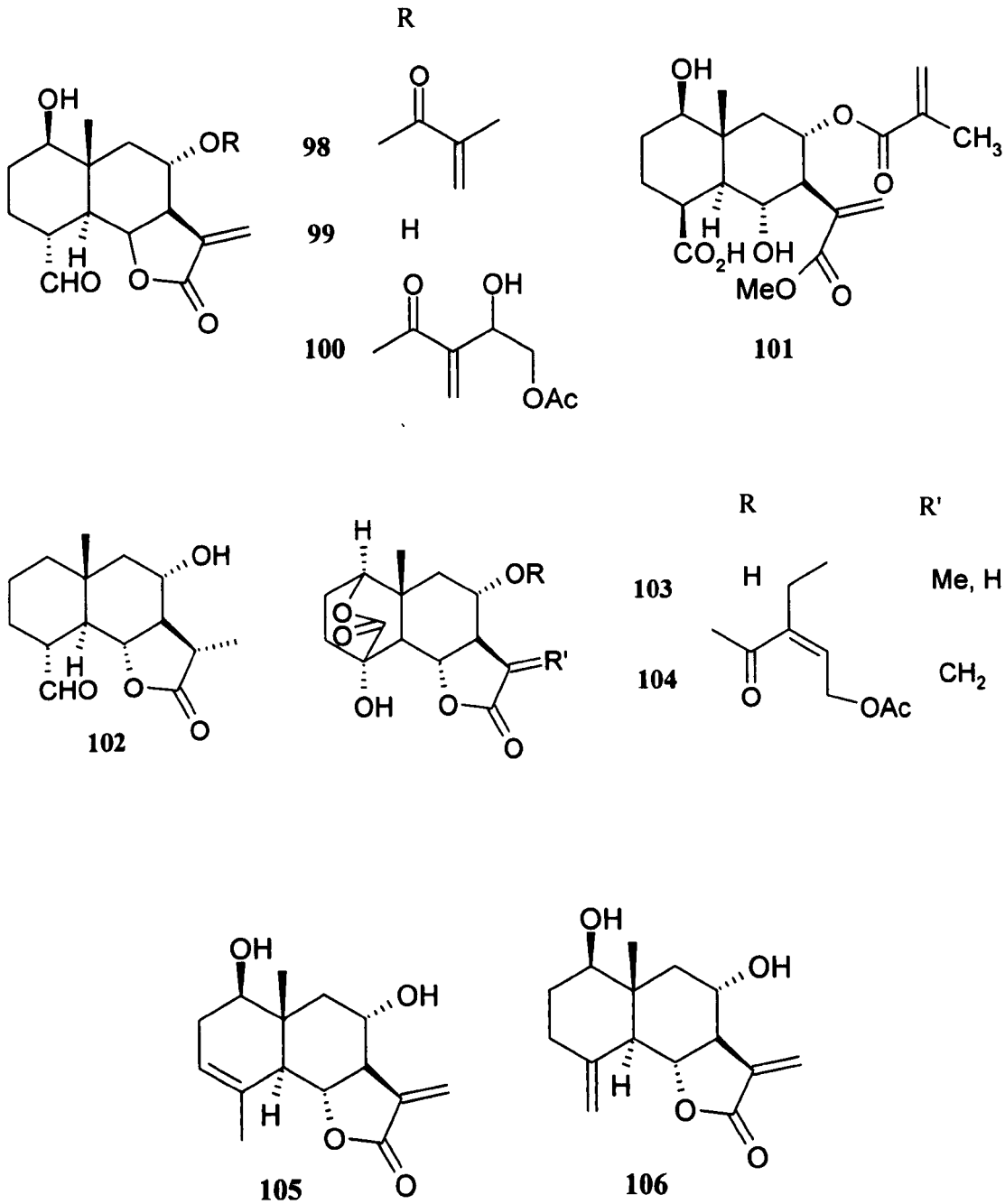
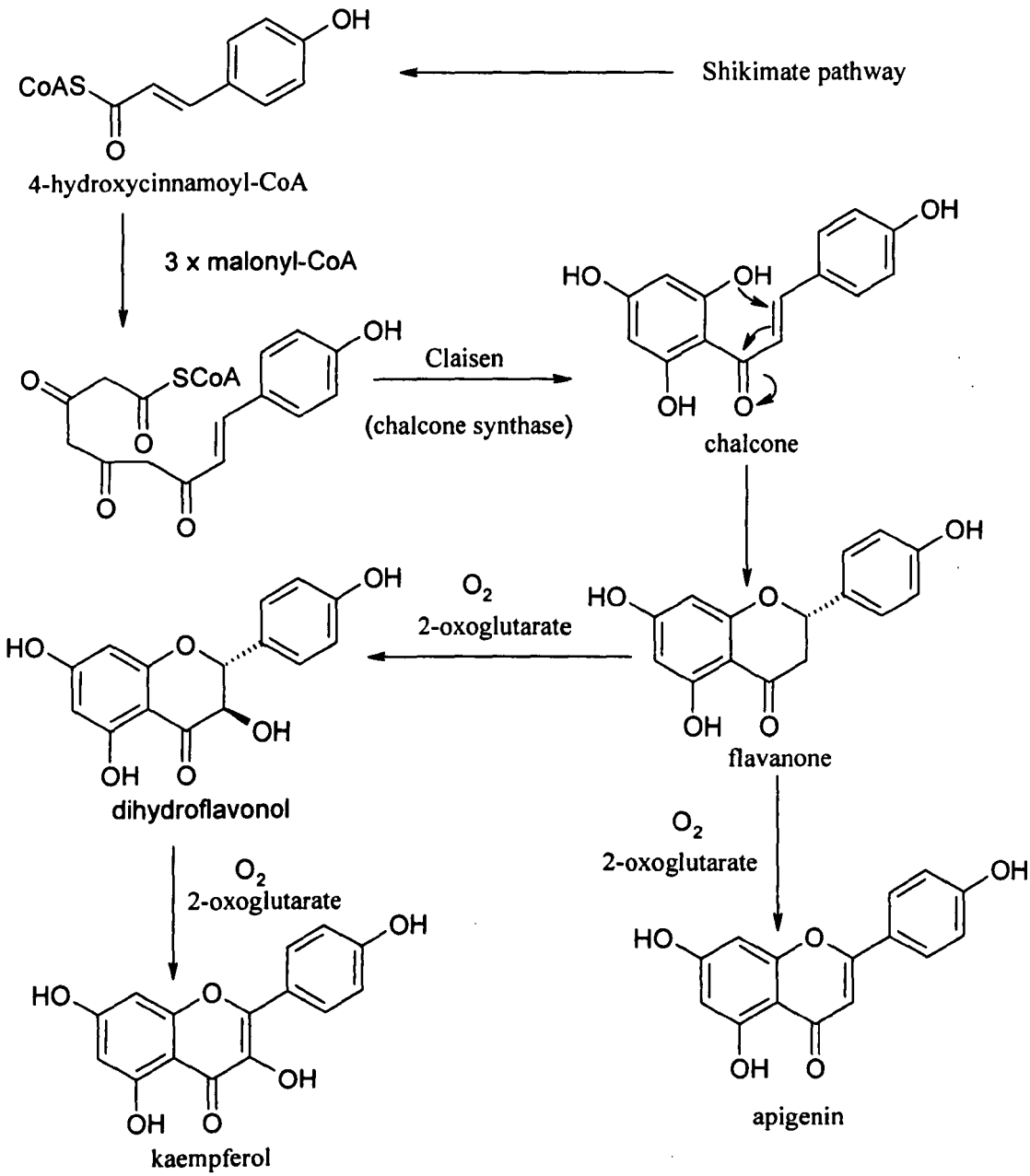


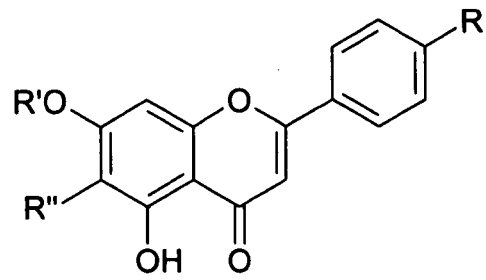
Figure 1 (Cont.). Structures of the sesquiterpene lactones from the genus *Centaurea*

1.4.2 Flavonoids

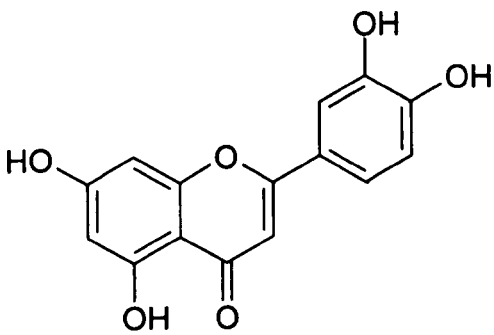
The flavonoids are a large group of natural products which are widespread in higher plants but found in some lower plants including algae.[127] They are 15-carbon compounds consisting of two aromatic rings A and B, and oxygen containing heterocyclic ring C. Depending on the central heterocyclic ring, flavonoids can be unsaturated e.g. flavones and flavonols or saturated e.g. flavanones and flavans.[128] Flavonoids are commonly associated with sugars in conjugation form and within any one class may be characterised as monoglycosidic, diglycosidic, etc. There are over 2,000 glycosides of the flavones and flavonols that have been isolated to date. Shikimate pathways, hydroxycinnamoyl-CoA and malonyl-CoA are the precursors for the biosynthesis of flavonoids. Biosynthesis of kaempferol and apigenin from these two precursors are shown in **Scheme 2**.[129] About 50 flavonoids have been reported from the genus *Centaurea* until now (**Table 1**).



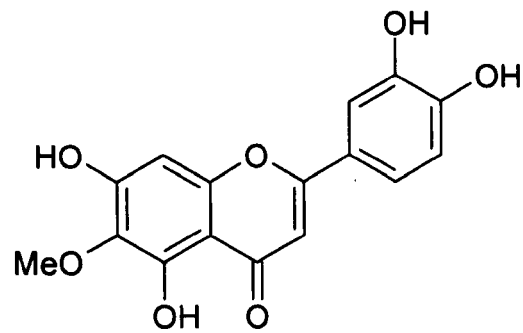
Scheme 2. Biosynthesis of flavonoids



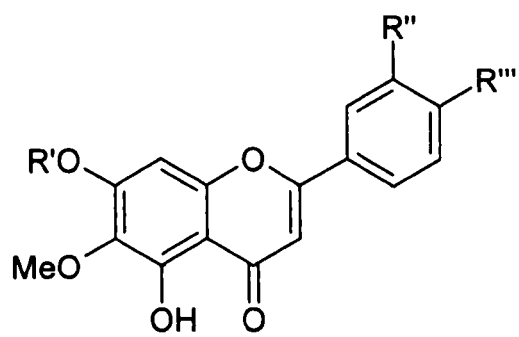
	R	R	R'
107	H	H	H
108	OH	H	H
109	OH	β-D-glucosyl	H
110	OH	Me	OMe
111	OH	methyglucuronide	OMe



112

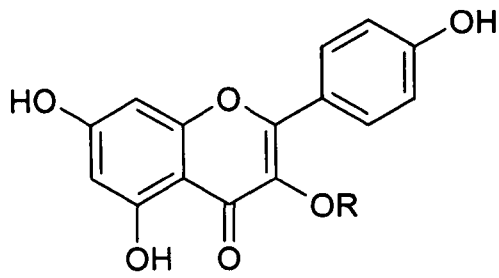


113

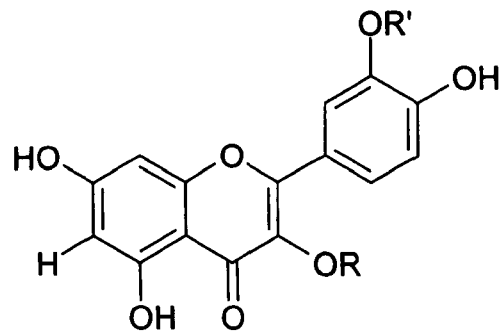


	R'	R''	R'''
114	Me	OH	OH
115	Me	H	OMe
116	H	H	OMe
117	H	H	OH
118	H	OMe	OH
119	β-D-glucosyl	OMe	OH
120	H	OMe	OH
121	H	OMe	OMe

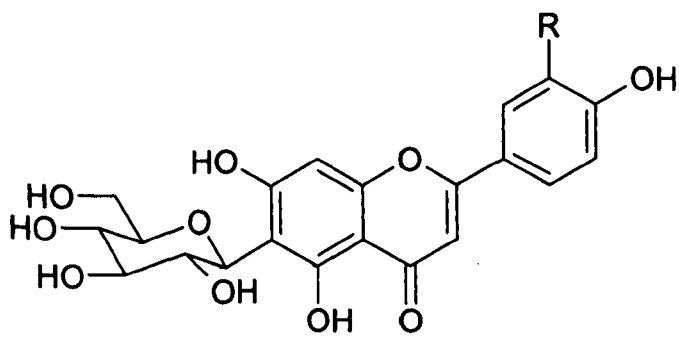
Figure 2. Structures of the flavonoids from the genus *Centaurea*



	R
122	H
123	β-D-glucosyl
124	β-D-galactosyl



	R	R'
125	H	H
126	β-D-glucosyl	H
127	H	Me



128 R=H

129 R=OH

Figure 2 (Cont.). Structures of the flavonoids from the genus *Centaurea*

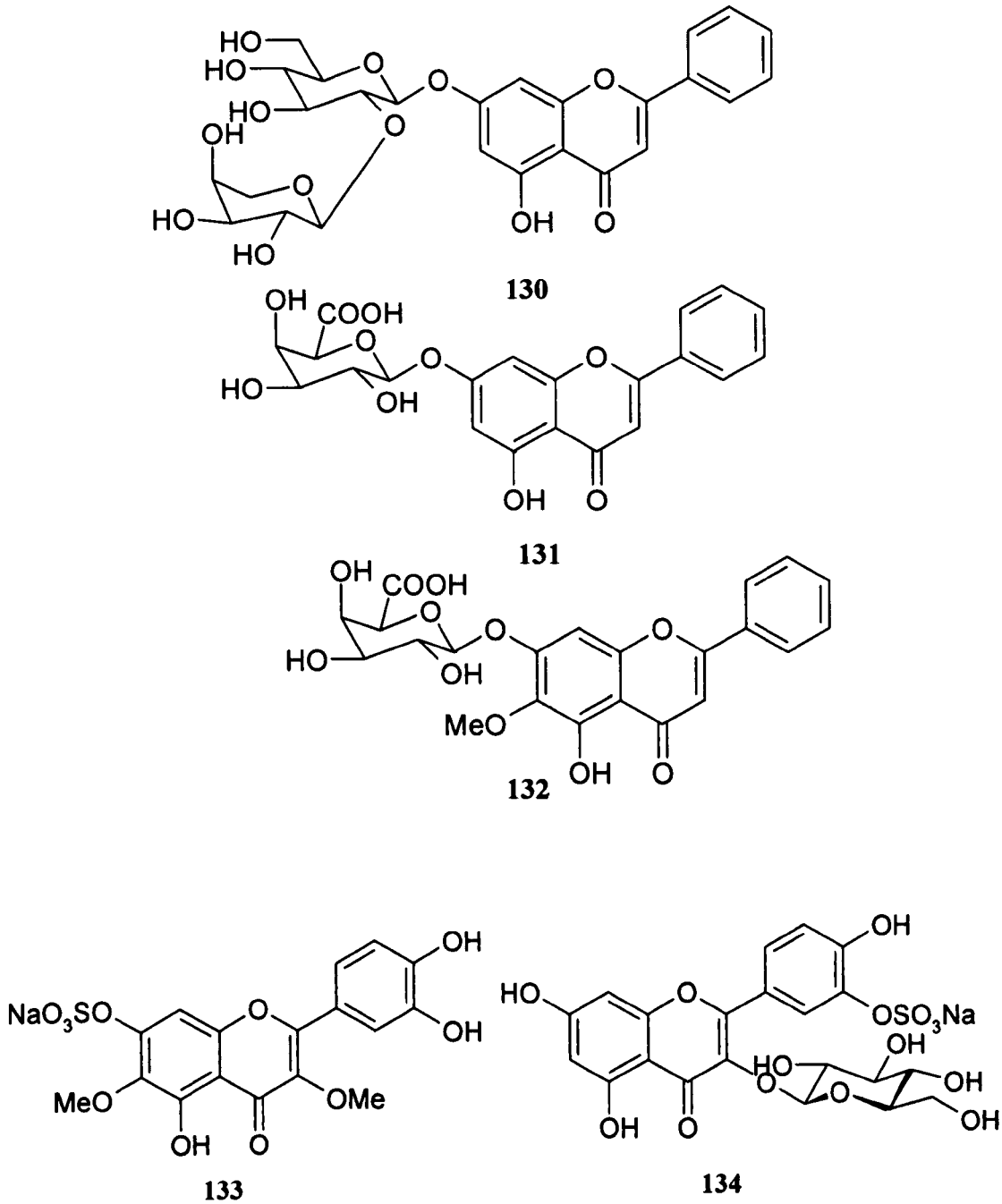
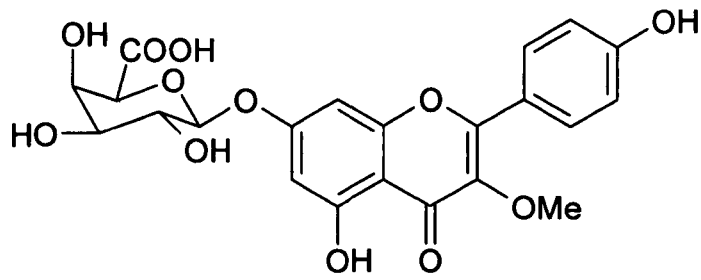
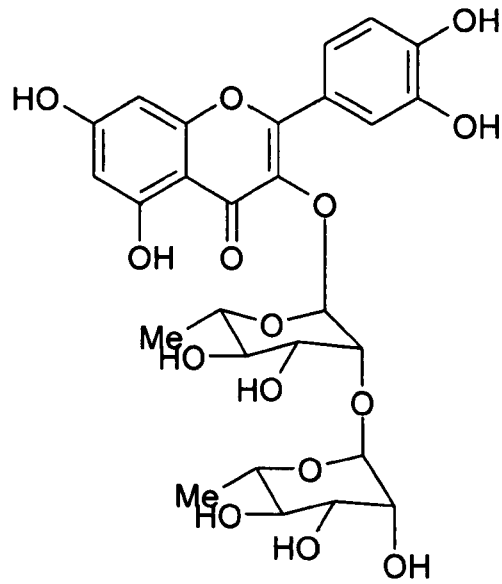


Figure 2 (Cont.). Structures of the flavonoids from the genus *Centaurea*



135



136

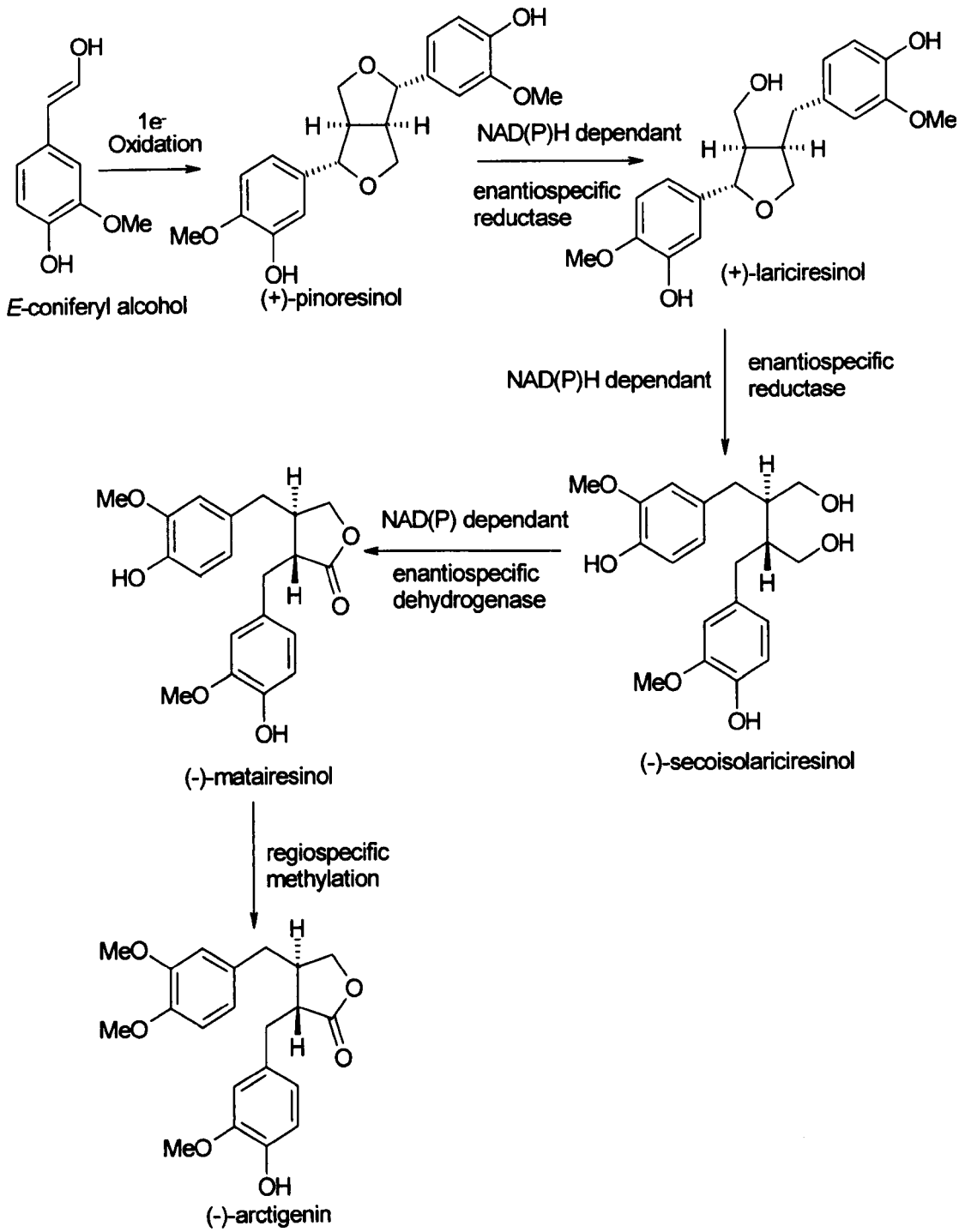
Figure 2 (Cont.). Structures of the flavonoids from the genus *Centaurea*

1.4.3 Lignans

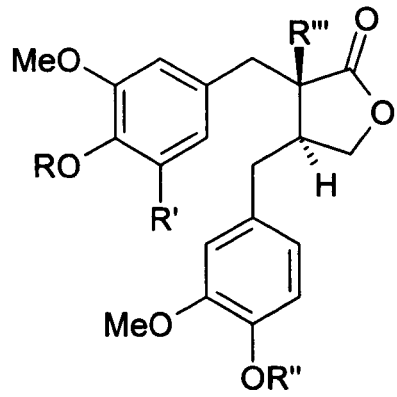
Lignans are a class of secondary metabolites widely encountered in the plant kingdom and are associated with a range of biological activities. Lignans are characterised as phenylpropanoid dimer linked by free-radical coupling through 8-8' bonds.[130] Lignans are found in roots, leaves seeds, fruits and wooden parts of plants. They are assumed to function as phytoalexins that is to provide protection for the plants against disease and pests such as wood rot fungi.[131] Secoisolariciresinol and matairesinol are the two most studied plant lignans. However, podophyllotoxin and structurally related lignans have drawn potential interest towards scientist worldwide because of their antitumor activities.[132] Usually, lignans exist in nature as glycosides and natural lignans are optically active.

1.4.3.1 Biosynthesis of Lignans

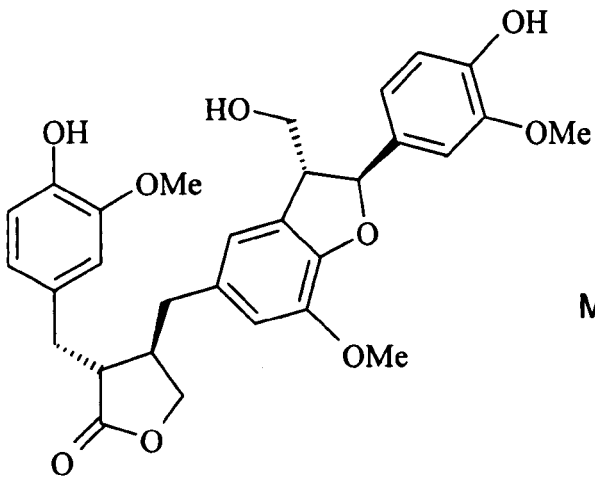
Monolignols such as coniferyl alcohol which is converted from phenylalanine by a five step enzymatic reaction-deamination, aromatic hydroxylation, *O*-methylation, CoA-dedicated ligation and NADPH mediated reduction.[133,134] The phenylpropanoid monolignols are thought to dimerise to C_{18} structures via oxidative phenolic coupling reactions followed by various oxidation, reduction or alkylation steps to give the lignans. The **Scheme 3** represents the biosynthesis of lignan in *Forsythia* species described by Lewis and coworkers.[135-138] So far about 8 lignans have been isolated from the genus *Centaurea* (**Table 1**).



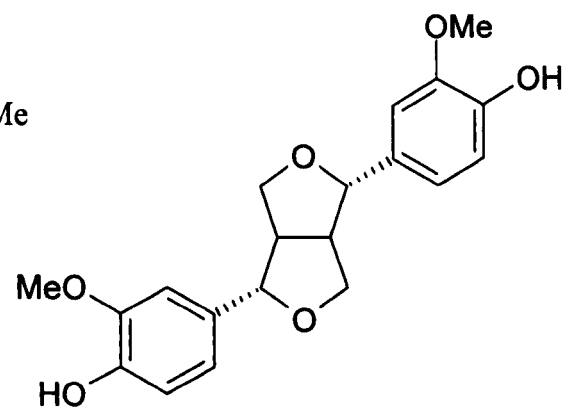
Scheme 3. Biosynthesis of lignans



	R	R'	R''	R'''
137	β-D-glucosyl	H	Me	H
138	β-D-glucosyl	H	H	H
139	H	H	Me	H
140	H	H	H	H
141	H	OMe	H	H
142	H	H	Me	OH



143

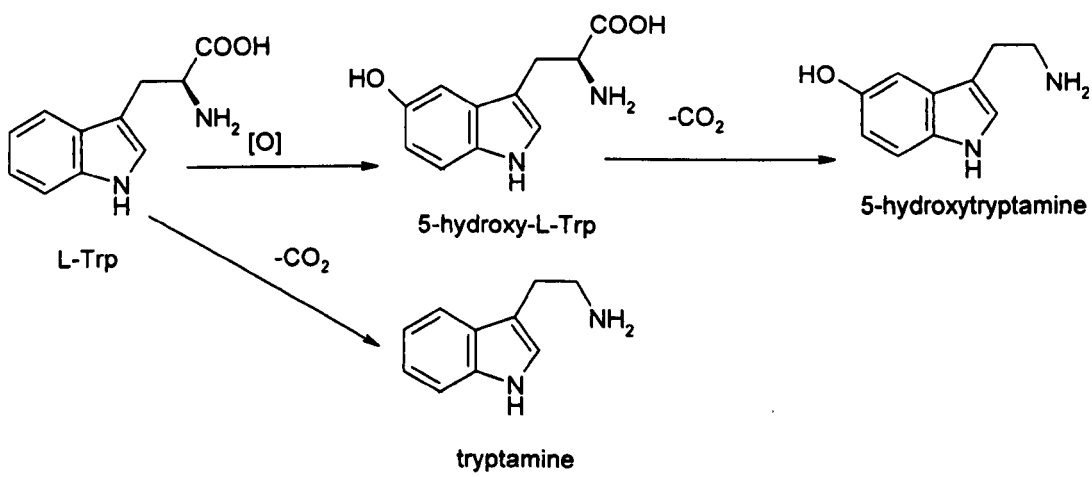


144

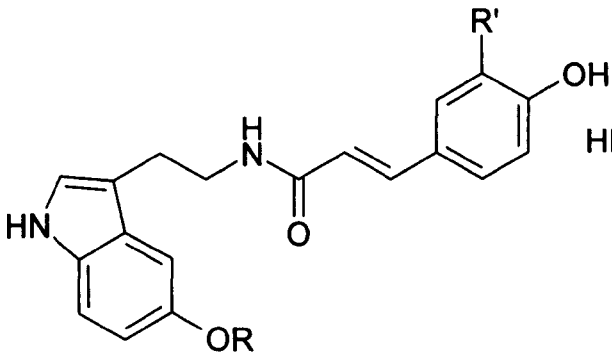
Figure 3. Structures of the lignans from the genus *Centaurea*

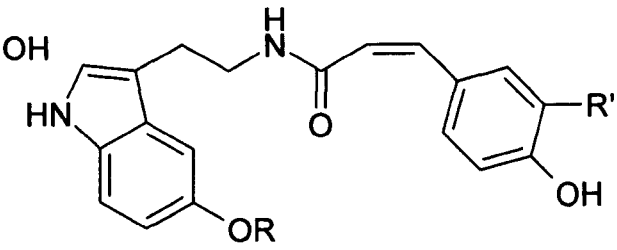
1.4.4 Alkaloids

Alkaloids are a large group of nitrogen containing secondary metabolites of plant, microbial or animal origin. More than 10000 alkaloids have been reported until now.[139] Alkaloids are usually found in the seeds, root, leaves, or bark of the plant, and generally occur as salts of various plant acids. They can be extracted from plant with acidified water or organic solvents (e.g. chloroform, methanol etc.). Usually, alkaloids are classified on the basis of the heterocyclic ring system they possess, e.g. pyrrolidine, piperidine, quinoline, isoquinoline and indole etc. Alkaloids are produced by secondary metabolism of primary metabolites, usually amino acids.[139] L-Tryptophan, originating from the shikimate pathway via anthranilic acid, acts as a precursor of a wide range of indole alkaloids. Biosynthesis of two indole alkaloids, tryptamine and 5-hydroxytryptamine is shown in **Scheme 4**.[129] About 8 alkaloids have been reported from the genus *Centaurea* until now (**Table 1**).

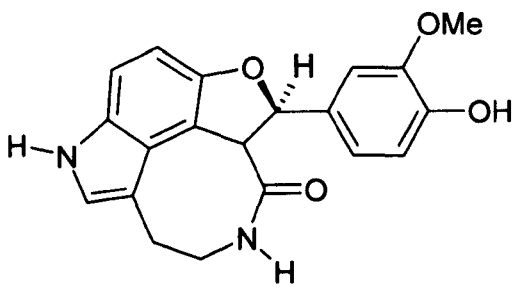


Scheme 4. Biosynthesis of indole alkaloids

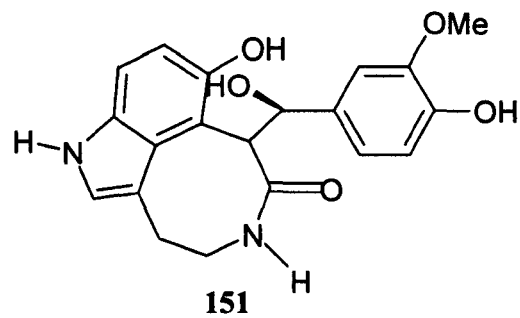




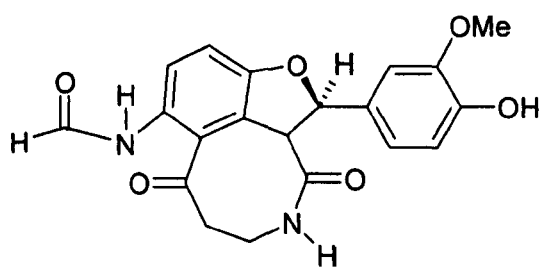
	R	R'		R	R'
145	H	H	**147**	Me	H
146	Me	H	**149**	H	OMe
148	H	OMe			



150



151

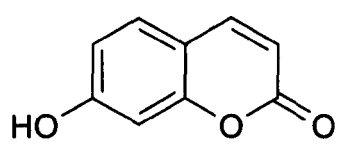


152

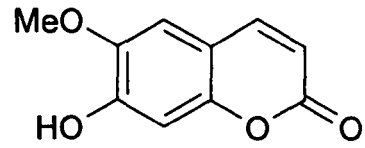
Figure 4. Structures of the alkaloids from the genus *Centaurea*

1.4.5 Other Secondary Metabolites

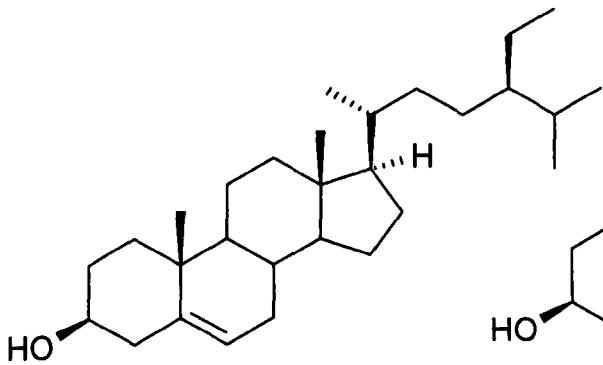
In addition to the major classes of secondary metabolites described above several coumarins, steroids, terpenoids and simple phenol derivatives have been reported from the genus *Centaurea*.



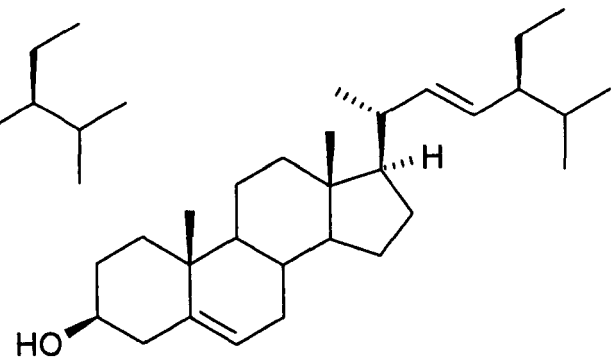
153



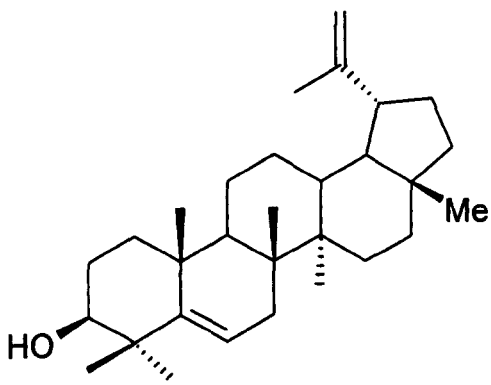
154



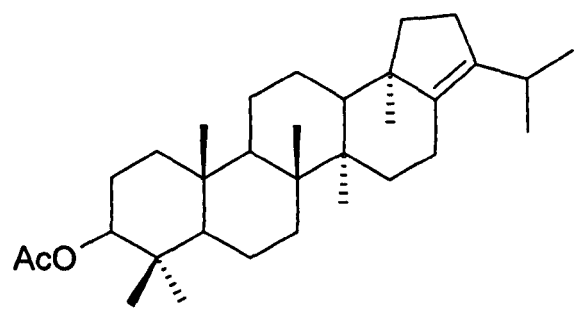
155



156

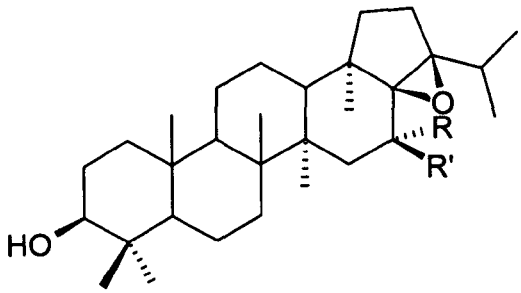


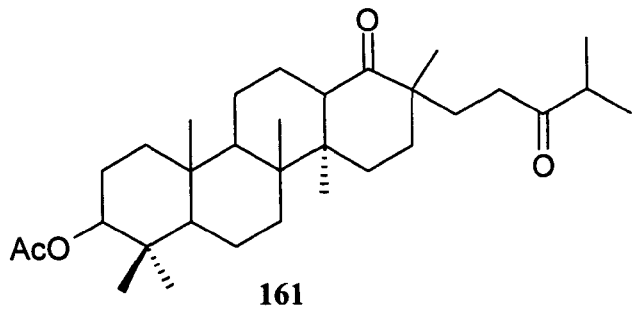
157



158

Figure 5. Structures of the miscellaneous compounds from the genus *Centaurea*

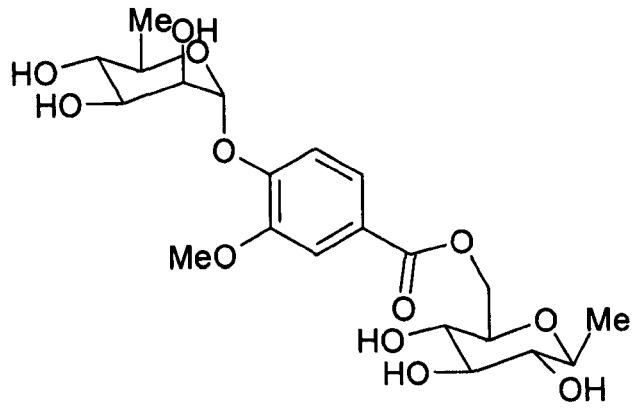




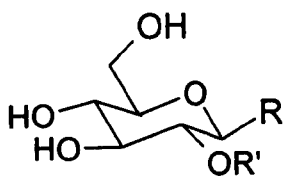
161

159 R =OCH_2CH_3, R'=H

160 R=R'=H

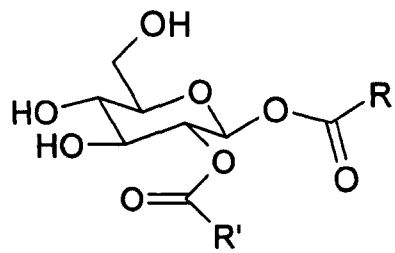


162

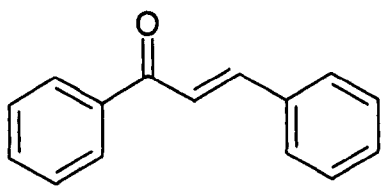


163 R, R' =Ang

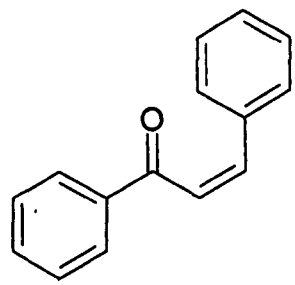
164 R=iVal, R'=Ang



165 R, R' = 2-MeBu, i-Val

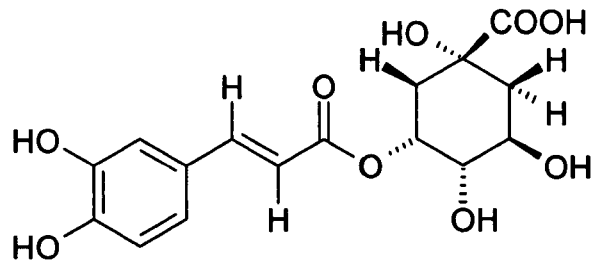


166

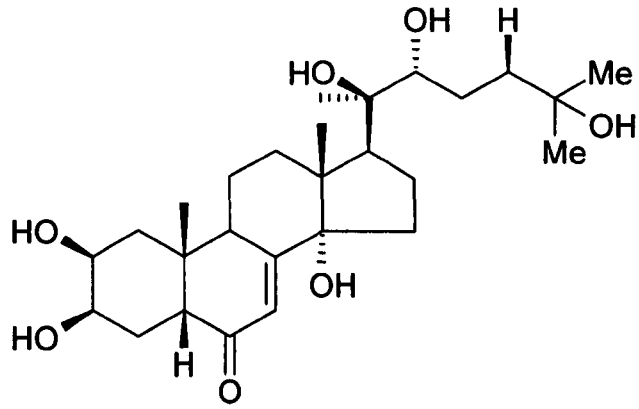


167

Figure 5 (Cont.). Structures of the miscellaneous compounds from the genus *Centaurea*



168



169

Figure 5 (Cont.). Structures of the miscellaneous compounds from the genus *Centaurea*

1.5 *Centaurea* Species under Investigation

1.5.1 *C. americana* Nutt.

C. americana Nutt, commonly known as "Jolly Joker" or "Basket flower", is indigenous to Northern Americana-Coahuila and Nuevo Leon in Mexico and Arizona, Arkansas, Kansas, Louisiana, Missouri, New Mexico, Oklahoma and Texas in the USA, and also cultivated in several other countries.[140]

1.5.2 *C. bornmuelleri* Hausskn. Ex. Bornm.

C. bornmuelleri Hausskn. Ex. Bornm (Section: *Psephelloideae*) is endemic to Turkey. This plant has natural growth in East Anatolia, Turkey (B7 Erzincan, 30.5 km from Erzincan to Gumushane, 2024 m, 39°53'29'' N 39°21'6'' E).[141] It is a perennial plant with erect stems, 35-70 cm. Flowers are violet or purple. Aches and pappus are 7 mm and 7-9 mm respectively.[142]

1.5.3 *C. cyanus* L.

C. cyanus L. (Section: *Cyanus*) well known as "corn flower" or bachelor's button", is a flowering weed endemic to Iran, Iraq, Turkey and Pakistan in Asia, and Albania, Bulgaria, Greece, Italy and Yugoslavia in Europe, and also cultivated and naturalised in many other countries of the world. It is an annual plant (15-80 cm) with blue flowers.[142]

1.5.4 *C. dealbata* Willd.

C. dealbata Willd. (Section: *Psephellus)* commonly known as "Whitewash cornflower", is indigenous to Turkey.[142]

Figure 6. *C. cyanus* seeds

Figure 7. *C. dealbata* seeds

1.5.5 *C. gigantea* Schultz Bip. ex Boiss.

C. gigantea Schultz Bip. ex Boiss., (Section: *Cynaroides)* is endemic to South East Anatolia, Turkey. The plant likes dry rocky slopes. Plant is biennial with erect stem up to 1-1.80 m. Leaves are densely adpressed-tomentose, Flowers are pale purplish but become white after drying. Usually flowers for a period of 7-8 months. Aches 5.5-6 mm; pappus 7-8 mm.[142]

1.5.6 *C. huber-morathii* Wagenitz.

C. huber-morathii Wagenitz (Section: *Psephelloideae*) is endemic to E. Anatolia, Turkey. The plant has natural growth in East Anatolia, Turkey (B7 Erzincan, 30.5 km from Erzincan to Gumushane, 2024 m, 39°53'29'' N 39°21'6'' E). Plants are perennial with erect stem (50-70 cm). Flowers are lilac-pink with marginal radiant. Aches are immature and pappus 11-13 cm. *C. huber-morathii* is closely related to *C. bornmuelleri.*[142]

1.4.7 *C. macrocephala* Muss. Puschk. Ex Willd.

C. macrocephala Muss. Puschk. Ex Willd., (Section: *Grossheimia)* commonly known as "big-head knapweed", is a splendid border plant with large, yellow, thistle-like flower heads, endemic to Armenia, Azerbaijan and Turkey, and also naturalised in many other countries of the world.[140]

1.5.8 *C. montana* L.

C. montana L., commonly known as "mountain knapweed", is an erect plant with large, reddish, blue center flower heads, native to Australia, Belgium and Italy, and also cultivated in many other countries of the world.[140]

1.5.9 *C. mucronifera* DC.

C. mucronifera DC. (Section: *Psephelloideae*) is an endemic species which is distributed in the Mediterranean, the Middle and the Eastern Anatolian regions of Turkey. It is Perennial plant with flowering stems 3-40 cm. Flowers are rose-purple, marginal radiant. Aches and pappus are 5-7 and 2-5 mm respectively. This plant prefers calcarious rocky communities.[142]

1.5.10 *C. pamphylica* Boiss. & Heldr.

C. pamphylica Boiss. & Heldr. (Section: *Calcitrapa*) is an endemic species which is distributed in the Mediterranean, Anatolian regions of Turkey. [142]

1.5.11 *C. schischkinii* Tzvelev.

C. schischkinii Tzvelev. (Section: *Psephelloideae*) is an endemic plant species to East Anatolia, Turkey and located in Erzurum city between 1800-2060 m. It is a perennial plant, the stem is erect (40-80 cm), and flowers are pink, distinctly radiant. Aches 6-7 mm and pappus 7-8 mm.[142] This plant prefers steppy areas and road sites in terms of environmental characteristics. In Turkey, the natural distributing areas of the species are Erzurum, Hinis, 1806 m, N 39°42'56'' E 41°48'01'', Erzurum, Hinis, 1830 m N 39°42'92'' E 41°48'75'', Erzurum, Tekman, 2025 m, N 39°42'23'' E 41°48'51'', Erzurum, Tekman, 2040 m, N 39°41'57'' E 41°47'69'' and Erzurum, Palandoken Mountain, 2057 m, N 39°44'32'' E 41°23'42''.[141]

Figure 8. *C. schischkinii* seeds

Figure 9. *C. bornmuelleri* seeds

1.5.12 *C. urvillei* subsp. armata Wagenitz

C. urvillei subsp. armata Wagenitz (Section: Acrocentron DC.) is widely distributed in the steppe and calcarious rocky communities in the eastern Mediterranean, the Middle and the Eastern Anatolian regions of Turkey.[142]

Figure 10. *C. mucronifera*

Figure 11. *C. urvillei*

1.6 Aims and Objective of the Present Study

There are more than 250,000 species of higher plants existing on this planet and only a small portion has been explored phytochemically. It is believed that plants can provide potential bioactive compounds for the development of new 'leads' to combat diseases. Literature review has showed that *Centaurea* species have medicinal value and are used as folk medicine in Turkey and other countries. Various bioactive compounds have been isolated from this genus and many of the species have chemical diversity. However, it is believed that many of the *Centaurea* species have not been explored yet for potential cytotoxic compounds. Therefore, the primary aim of the present work is to investigate *Centaurea* species for cytotoxic compounds.

Therefore, the objectives of the present work are to:

i) extract *Centaurea* species and assess general toxicity of different extracts;

ii) separate, purify and isolation of active compound (s) from extracts;

iii) identify and elucidate structures of isolated compounds by different spectroscopic techniques including one dimensional and two dimensional NMR;

iv) assess general toxicity (using brine shrimp lethality assay) and cytotoxicity (using MTT assay) of isolated compounds;

v) perform antioxidant activity assays of the extracts and pure compounds.

CHAPTER TWO

Materials and Methods

2 Materials and Methods

2.1 Plant Materials

Aerial parts and seeds of twelve plant species of the genus *Centaurea* were collected from Turkey, and B & T, World Seeds Sarl, Paguignan, 34210 Olonzac, France, for complete phytochemical studies. Voucher specimens for all the collections have been deposited in the herbarium of the Plant and Soil Science Department, University of Aberdeen, UK (**Table 2**).

2.2 Extraction of Plant Materials

The air dried aerial parts of each plant sample and seeds were finely ground and subjected to exhaustive Soxhlet extraction using, successively, *n*-hexane, dichloromethane (DCM) and methanol (MeOH). The individual extracts were concentrated separately, using a rotary evaporator at a maximum temperature of 40°C and under reduced pressure.

2.3 Isolation of Compounds

Modern chromatographic techniques, e.g. thin layer chromatography, preparative thin layer chromatography, vacuum liquid chromatography, column chromatography and high performance liquid chromatography were applied for separation, purification and isolation of pure and active compounds from the crude extracts.

2.3.1 Thin Layer Chromatography (TLC)

Commercially available pre-coated silica gel 60 PF_{254} plates were used for initial screening of extracts and column fractions, checking the purity of the isolated compounds and for confirming the identification of known compounds by co-TLC

with authentic samples. Different types of solvent systems, ranging from non-polar to polar were frequently used (**Table 3**).

Table 2. List of *Centaurea* species investigated

Species	Voucher number	Parts	Amount (g)	Source
Centaurea macrocephala	PHSH0001	Seeds	100	B & T
Centaurea cyanus	PHSH0002	Seeds	100	B & T
Centaurea dealbata	PHSH0003	Seeds	100	B & T
Centaurea americana	PHSH0004	Seeds	100	B & T
Centaurea huber-morathii	PHSH0005	Seeds	80	Turkey
Centaurea montana	PHSH0006	Seeds	100	B & T
Centaurea schischkinii	PHSH0007	Seeds	80	Turkey
Centaurea bornmulleri	PHSH0009	Seeds	80	Turkey
Centaurea gigantea	PHSH0010	Aerial	100	Turkey
Centaurea pamphylica	PHSH0011	Aerial	100	Turkey
Centaurea urvillei	PHSH0012	Aerial	100	Turkey
Centaurea mucronifera	PHSH0013	Aerial	100	Turkey

2.3.2 Preparative Thin Layer Chromatography (PTLC)

PTLC was applied for separation and final purification of compounds. Plates were prepared by coating glass plates (20 x 20 cm) to a thickness of 0.5 mm using slurry made from 40 g of silica gel (Kiselgel 60 PF_{254}) and 80 mL of distilled water. The plates were allowed to air dry and then activated in an oven at 60^0 C for overnight. The sample to be analysed was dissolved in a small amount of suitable solvent and applied to the plates as a uniform band 2 cm apart from the bottom edge. The plates were then developed in an appropriate solvent system (**Table 3**). Developed plates were allowed to dry and visualised under UV light (254 nm and 366 nm) or by spraying on both edges of the plate with an appropriate spray reagent. Bands of interest were scraped from plates and the compounds were eluted from silica with DCM or ethyl acetate (EtOAc) or DCM-MeOH mixture.

Table 3. List of solvent systems used for TLC and PTLC

Code number	Solvent system	Proportion
A	*n*-Hexane:EtOAc	95:5
B	*n*-Hexane:EtOAc	90:10
C	*n*-Hexane:EtOAc	85:15
D	*n*-Hexane:EtOAc	80:20
E	*n*-Hexane:EtOAc	70:30
F	*n*-Hexane:EtOAc	60:40
G	*n*-Hexane:EtOAc	50:50
H	EtOAc	100
I	*n*-Hexane:DCM	50:50
J	*n*-Hexane:DCM	30:70
K	*n*-Hexane:DCM	20:80
L	*n*-Hexane:DCM	10:90
M	DCM	100
N	DCM:MeOH	95:5
O	DCM:MeOH	90:10
P	DCM:MeOH	80:20
Q	DCM:MeOH	70:30
R	DCM:MeOH	50:50
S	MeOH	100

2.3.3 Vacuum Liquid Chromatography (VLC)

VLC involves a short column, packed with TLC grade silica (Kieselgel 60 H) and operates under reduced pressure.[143,144] The size of the column and the height of the adsorbent layer were variables throughout this study and were dependent upon the amount of extract analysed. The column was always dry-packed with silica gel under suction. The sample to be separated was adsorbed onto a small amount of silica gel

(Kieselgel 60, mesh 70-230) and dried to free-flowing particles which were then applied uniformly on the top of the packed column. Elution was carried out with solvents of increasing polarity (**Table 3**).

2.3.4 Column Chromatography (CC)

Different types of column chromatography were used during this study. They are mentioned below.

2.3.4.1 Normal Column Chromatography

The traditional column chromatography (CC) was used for fractionation and further purification of extract and fractions using normal phase column grade silica gel (mesh 70-230). The column was packed with the slurry of silica gel in an appropriate solvent (e.g. *n*-hexane or DCM). Dry sample or solution of sample (in minimum volume of column packing solvent) applied to the top of the column. The column was then eluted with different solvent with increasing polarity.

2.3.4.2 Sephadex LH-20 Column (Gel filtration)

Gel filtration or size exclusion chromatography was adopted to remove pigments (i.e., chlorophyll) and separate compounds from VLC fractions. A simple glass column was packed with the slurry of Sephadex LH-20 in appropriate solvents and the separation was achieved according to the molecular size of the compounds. The sample to be separated or purified were dissolved in a small amount of solvent and applied to the top of the adsorbent. The column was then eluted with 20% *n*-hexane in DCM followed by 100% DCM and the DCM-MeOH mixture of increasing polarity. The column was finally washed with 100% MeOH to make it suitable for further use.

2.3.4.3 Solid Phase Extraction (Sep-Pak) Column

Sep-Pak column was packed with reversed-phase silica (C_{18}) which is used for the pre-HPLC fractionation of MeOH extracts under suction. Sep-Pak columns with the following specification were used.

Column- Waters Sep-Pak Cartridge, vac 35 cc, C_{18}-10 gm

Column size-5.5 × 2.5 cm

Packing materials-silica, reversed-phase

Manufacturer-Waters, USA

2.3.5 High Performance (or Pressure) Liquid Chromatography (HPLC)

HPLC is a sophisticated technique for analysing samples qualitatively and quantitatively. Both analytical and preparative reversed-phased HPLC were used routinely for separation and purification of compounds in this study.

2.3.5.1 Analytical HPLC

Reversed-phased analytical HPLC was used for the chemical profiling of the Sep-Pak fractions, and development of a suitable mobile phase for preparative HPLC as well as analysis of the purity of isolated compound. A JASCO PU-1580 Intelligent HPLC Pump, coupled with JASCO DG-1580-53 Degasser and JASCO LG-1580-02 Ternary Gradient Unit. and Luna C_{18} column (5 μm, 250 mm x 4.6 mm) were used for this purpose.

2.3.5.2 Preparative HPLC

Preparative-HPLC separation was performed using a Dionex prep-HPLC system coupled with Gynkotek GINA50 autosampler and Dionex UVD340S Photo-Diode-Array detector. A Luna C_{18} column (10 μm, 250 mm x 21.2 mm) was used for this

purpose. This system is mentioned as Prep-HPLC-1 for describing isolation of compounds. In some cases, semi-preparative column (Luna C_{18} column; 5 µm, 250 mm x 10.0 mm) was used with preparative HPLC system instead of preparative column to purify minor fractions. This system is mentioned as Semi-Prep HPLC-1.

2.3.5.3 Semi-Preparative HPLC

Semi-Preparative HPLC separation was carried out using JASCO PU-2087 Plus intelligent Preparative Pump coupled with JASCO UV-2070 Plus Intelligent UV/VIS detector. A Luna C_{18} column (5 µm, 250 mm x 10.0 mm) was used with this system to purify minor fractions obtained from preparative HPLC. Usually, isocratic solvent system was used for this purpose. This system is mentioned as Semi-Prep HPLC-2 to describe isolation of compounds.

2.4 Detection of Compounds

The following techniques were used for the detection of compounds on TLC and PTLC plates.

2.4.1 UV Light

The developed and dried plates, either TLC or PTLC, were placed under UV light (254 and 366 nm) to observe quenching or fluorescing compounds.

2.4.2 Spray Reagents

Different types of spray reagents were used depending upon the chemical nature of compounds expecting to be present in the crude extracts or fractions.

2.4.2.1 Vanillin in Sulfuric acid[145]

This reagent was prepared by dissolving 1g vanillin in 100 mL of concentrated sulfuric acid (H_2SO_4)-ethanol (EtOH) mixture (1:1) and it was used as a general spray reagent to detect mainly non-alkaloidal compounds. TLC or PTLC plates, sprayed with this reagent were heated at 110°C for 15 minutes to develop a specific colour, indicating the type of compounds.

2.4.2.2 Anisaldehyde-Sulfuric acid

A mixture of 10 mL conc. H_2SO_4 and 20 mL glacial acetic acid were poured into 170 mL MeOH containing 1 mL anisaldehyde. The TLC plate was sprayed with this reagent and heated at 110°C for 15 min to visualize the spot.

2.4.2.3 Modified Dragendorff's Reagent[146]

This reagent was prepared by mixing equal parts (v/v) of 1.7% bismuth subnitrate dissolved in 20% acetic acid in water and a 40% aqueous solution of potassium iodide. Alkaloids gave positive colour reaction (usually orange-red colour) with this spray reagent instantly.

2.5 Instrumentation

2.5.1 General Laboratory Instrument

2.5.1.1 Melting Point Apparatus

Melting point of solid compounds was determined by using a Griffin melting point apparatus.

2.5.1.2 Polarimeter

The optical rotation was measured on ADP 220 Polarimeter, Bellingham, Stanley, Ltd.

2.5.2 Spectroscopic Instruments

2.5.2.1 UV Spectrophotometer

Ultra violet (UV) spectra were obtained in methanol using a Hewlett-Packard 8453 UV-Vis spectrometer.

2.5.2.2 IR Spectrophotometer

Infra red (IR) spectra were obtained on AVATAR 360 FT-IR spectrophotometer.

2.5.2.3 Mass Spectrometry

MS analyses were performed on a Quattro II triple quadrupole instrument.

2.5.2.4 NMR Spectroscopy

^{1}H and ^{13}C NMR spectra were obtained from Varian Unity Inova 400 MHz, Bruker AC-250, AMX 400 and DRX 600 MHz NMR spectrometer. 2D experiments (^{1}H-^{1}H COSY, HSQC, HMBC, NOESY and ROESY) were observed mainly on the Varian 400 MHz and Bruker DRX 600 MHz NMR. Different types of deuterated solvents, e.g. $CDCl_3$, CD_3OD and DMSO-d_6 were used to prepare the sample solution for these experiments.

2.6 Structure Elucidation of Compounds

Modern NMR techniques are the main tools for structure elucidation of unknown compounds. The first step is to get ^{1}H and ^{13}C NMR spectra. To distinguish between the methyl, methylene, methine and quarternary carbons, in addition to a simple decoupled ^{13}C spectrum, the technique called Distortionless Enhachment by Polarisation Transfer (DEPT), available as DEPT 45^{0}, 90^{0} and 135^{0} can be very useful.[147] UV and IR can be used to identify the presence of different chromophores and functional groups. Fast atom bombardment mass spectroscopy (FABMS) and electro-spray ionisation mass spectrometry (ESIMS) are the two commonly used techniques for mass analysis.[148] By high resolution mass spectrometry an exact mass of the compound can be obtained. From UV, IR, ^{1}H and ^{13}C NMR data and mass spectra, a tentative molecular formula can be derived. A database search is performed based on taxonomy, to yield biological activity and known compounds isolated from the plant materials studied. From the database it can be determined whether the isolated compound has previously been reported using molecular formula and spectral data. This process is known as dereplication. The databases Dictionary of Natural Products, containing publication on most natural products, and NAPRALERT are the most popular databases.[149] The process of structure determination is illustrated in **Figure 12.**

Many two dimensional NMR experiments are used to elucidate the structure. The first experiment commonly used is the heteronuclear single quantum coherence (HSQC)[147,150] experiment that shows which proton and carbons are directly attached to each other. This experiment uses an inverse detected probe designed for maximum ^{1}H sensitivity. The same data can be yielded from a ^{1}H-^{13}C COSY, J=140 Hz

correlated spectroscopy (COSY)[147,151] experiment but this is less sensitive than the HSQC. This, in conjunction with data from DEPT establishes all corresponding C-H connectivities.

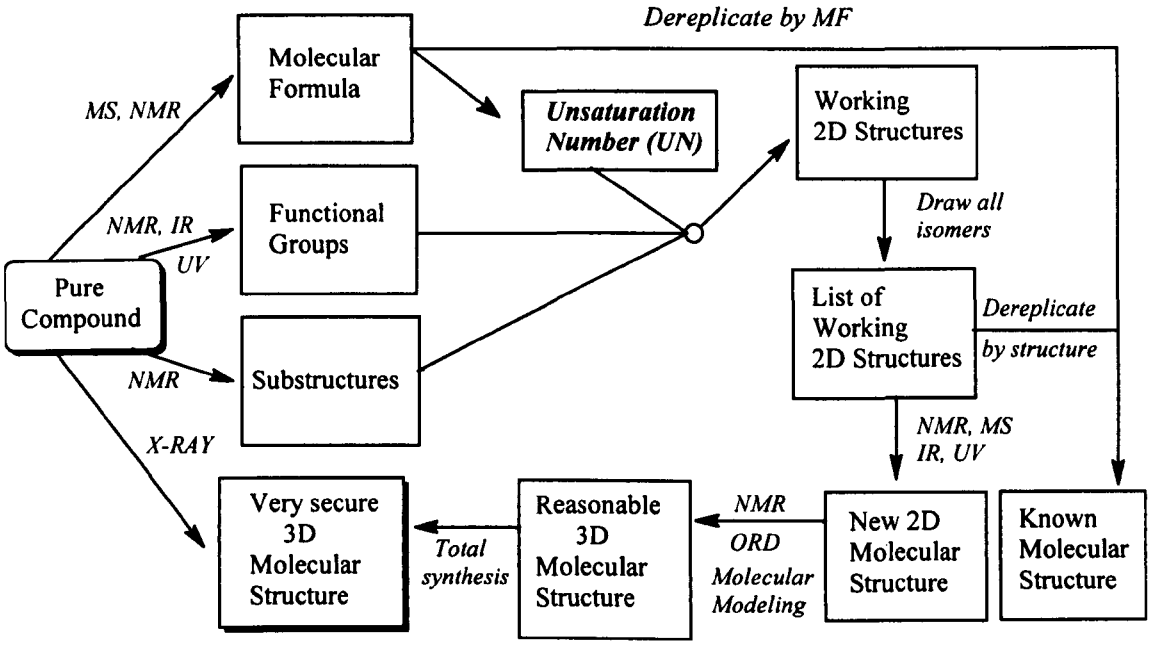


Figure 12. General scheme for structure elucidation[147]

All spin-spin coupled protons can be identified from an experiment called ^{1}H-^{1}H COSY. This experiment gives mainly geminal (two bond) and vicinal (three bond) correlation. With this information, possible substructures can be generated. These substructures can be linked together by long range correlation data. The inverse detected equivalent of the ^{1}H-^{13}C COSY (J=9 Hz) is called the heteronuclear multiple bond correlation (HMBC)[147] experiment which can detect two and three bond carbon to hydrogen connectivities. Sometimes heteronuclear single quantum coherence total correlation spectroscopy (HSQC-TOCSY) can be useful to assign crowded and mutually coupled spins. In this experiment all the protons in the same

spin system are linked together by ^{13}C chemical shifts on one axis and by ^{1}H chemical shifts on the other.

In cases where the interpretation of data is hindered by an inadequate number of protons in the spectra an alternative strategy is to use C-C correlations from a two dimensional incredible natural abundance double quantum transfer experiment (2D INADEQUATE). This allows the construction of the carbon skeleton of the molecule relatively easily, except in cases of extreme overlap in the ^{13}C spectrum. The Nuclear Overhauser Effect (NOE), which offers indirect measurement of dipolar coupling between nuclei that are in close proximity in space is a technique which helps to define the three dimensional structure. The interaction is by a through-space dipole-dipole relaxation mechanism, which is distance dependent ($1/n^{6r}$). The experiments are available in one dimensional (1D difference nOe) and two-dimensional (2D NOESY or ROESY). If the compound is crystalline, X-ray crystallography is the best way to determine the structure.

2.7 Isolation of Compounds

All *Centaurea* species listed in **Table 2** (page 61) were finely ground and Soxhlet-extracted with *n*-hexane (1.1 L), DCM (1.1 L) and MeOH (1.1 L) as described in **Section 2.2**. Thus *n*-hexane extract, DCM extract and MeOH extract were obtained separately from twelve *Centaurea* species and were stored at -20°C until further required for phytochemical or biological activity studies.

2.7.1 Isolation of Compounds from *Centaurea macrocephala*

The MeOH extract (2.0 g) of *C. macrocephala* obtained from Soxhlet-extraction process, was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 30, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient- water:MeOH= 70:30 to 20:80 over 50 min followed by 80% MeOH for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 30% MeOH, yielded fraction **Ma1** (70.5 mg, t_R = 14.1 min) and **CMa1** (314.1 mg, t_R= 16.5 min).

Fraction **Ma1** was further purified by Prep-HPLC-1 (isocratic elution with 40% MeOH in water, 20 mL/min) to obtain **CMa2** (5.8 mg, t_R = 6.8 min). Pre-HPLC-1, using the same condition as for the Sep-Pak 30% MeOH fraction, of 60% Sep-Pak fraction afforded more of **CMa1** (49.5 mg), **CMa3** (33.2 mg, t_R= 21.6 min) and fraction **Ma2** (47.2 mg, t_R= 23.6 min). Compound **CMa4** (5.2 mg, t_R= 23.4 min) was purified from fraction **Ma2** by Pre-HPLC-1, eluted with a linear gradient water:acetonitrile (MeCN)= 75:25 to 40:60 over 50 min, 20 mL/min in addition to **CMa3** (**Figure 13**).

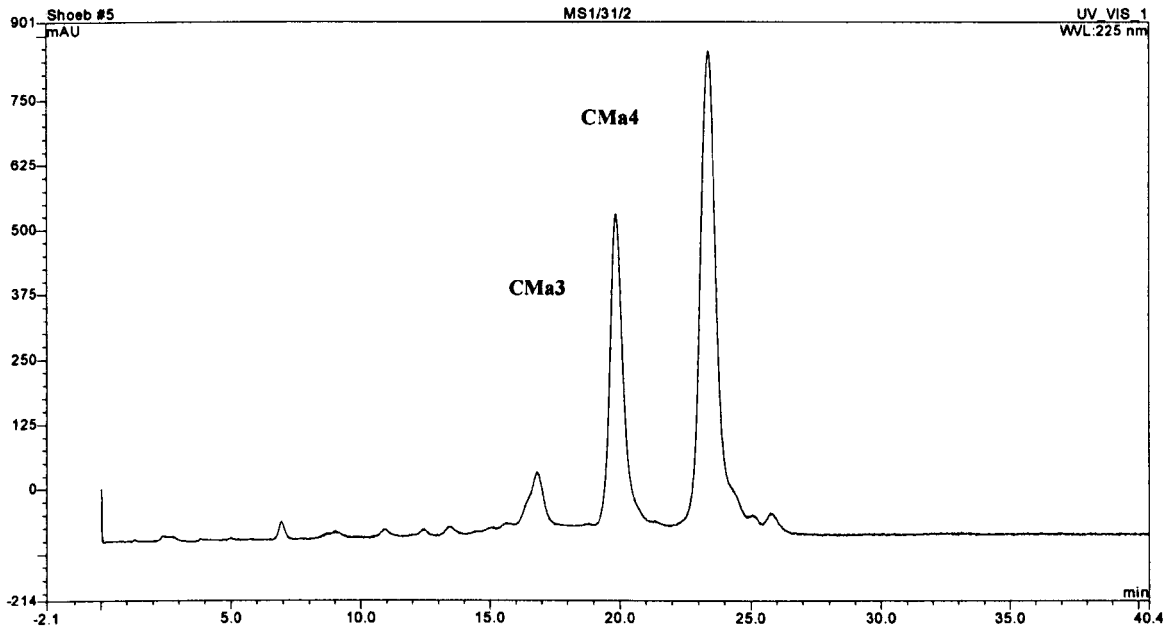


Figure 13. HPLC chromatogram of fraction **Ma1** from *C. macrocephala*

The *n*-hexane and DCM extracts were processed and their ^{1}H and ^{13}C NMR spectra showed that they contained mainly essential oil like phytol, caryophyllene oxide and long chain fatty acids and hydrocarbons. As a result, no compound of our interest was isolated from them.

2.7.2 Isolation of Compounds from *Centaurea cyanus*

The MeOH extract (2.0 g) of *C. cyanus* was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 30, 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient- water:MeCN= 80:20 to 40:60 over 50 min followed by 60% MeCN for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 40% MeOH, yielded **CC1** (4.7 mg, t_R = 17.1 min), **CC2** (10.3 mg, t_R = 22.2 min), **CC3** (12.2 mg, t_R = 26.2 min) and **CC4** (19.0 mg, t_R= 29.0 min).

The ^{1}H and ^{13}C NMR spectra together with preliminary TLC screening of *n*-hexane extract showed that it contained mainly long chain fatty alcohol, ester and acids. So this extract was not investigated further for isolation of compounds. However, the DCM extract showed interesting peaks in high and low field of the ^{1}H and ^{13}C NMR spectra. The DCM extract (100 mg) was subjected to vacuum liquid chromatography (procedure previously described on page 62). The column was packed in *n*-hexane and gradually increased in polarity from 0 to 100% DCM in *n*-hexane and finally washed with 100% MeOH and 7 fractions were obtained. Fraction obtained from 90% DCM in *n*-hexane was further purified by PTLC (page 61) using a solvent mixture of EtOAc and *n*-hexane (50/50) to afford pure compound **CC5** (5.4 mg).

2.7.3 Isolation of Compounds from *Centaurea dealbata*

The MeOH extract (2.0 g) was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 70:30 to 20:80 over 50 min followed by 80% MeOH for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 60% MeOH, yielded **CD1** (13.9 mg, t_R= 15.6 min), **CD2** (203.4 mg, t_R = 17.2 min), **CD3** (8.2 mg, t_R= 20.7 min).

The TLC, ^{1}H and ^{13}C NMR spectra of DCM extract showed the presence of similar compounds (in a negligible amount) isolated from the MeOH extract. So this extract was not processed. The ^{1}H and ^{13}C NMR spectra of *n*-hexane extract revealed that it contained mainly long chain fatty alcohol, fatty ester and acids. As a result, this extract was not processed further for isolation of compound.

2.7.4 Isolation of Compounds from *Centaurea americana*

The MeOH extract (2.0 g) was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 40:60 to 20:80 over 50 min followed by 80% MeOH for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 60% MeOH, yielded **CA1** (8.8 mg, t_R = 16.5 min), **CA2** (13.9 mg, t_R= 21.0 min), **CA3** (203.4 mg, t_R = 25.0 min), **CA4** (8.2 mg, t_R= 27.1 min), **CA5** (8.0 mg, t_R= 29.0 min), **CA6** (15.5 mg, t_R= 31.8 min), **CA7** (12.1 mg, t_R= 33.0 min) and **CA8** (9.2 mg, t_R= 34.2 min).

The TLC pattern of *n*-hexane and DCM extracts showed that these extracts contained mostly long chain alcohol and acids which were not compounds of our interest. So they were not processed for further investigation.

2.7.5 Isolation of Compounds from *Centaurea huber-morathii*

The MeOH extract (2.0 g) of *C. huber-morathii* was subjected to a Sep-Pak C_{18} (10 g) column eluting with a step gradient: 30, 40, 60, 80 and 100% MeOH in water (200 mL each). Fraction resolved from 40% aqueous MeOH was further purified by Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 70:30 to 20:80 over 50 min followed by 80% MeOH for 10 min, 20 mL/min) to afford **CH1** (9.0 mg, t_R = 15.1 min), **CH2** (180.0 mg, t_R= 17.3 min), **CH3** (11.0 mg, t_R = 26.1 min), **CH4** (5.0 mg, t_R= 21.2 min) and **CH5** (3.0 mg, t_R= 31.3 min).

The TLC pattern of the *n*-hexane extract showed that it contained mainly fatty alcohol and was not processed further. The DCM extract was fractionated by normal column chromatography (page 63) using *n*-hexane, DCM and MeOH with gradual

increase in polarity. Fractions eluted with DCM-MeOH (18:2) afforded compounds **CH1** (7.2 mg) and **CH2** (4.8 mg) which were also isolated from MeOH extract.

2.7.6 Isolation of Compounds from *Centaurea montana*

The MeOH extract (2.0 g) was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 30, 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient-water:MeCN=90:10 to 60:40 over 50 min followed by 40% MeCN for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 30% MeOH (**Figure 14**), yielded seven fractions: **M1** (30.4 mg, t_R=6.1 min), **M2** (78.2 mg, t_R=11.2 min), **M3** (50.9 mg, t_R=12.2 min), **M4** (60.0 mg, t_R=18.2 min), **M5** (925.0 mg, t_R=19.0 min), **M6** (45.4 mg, t_R=26.5 min) and **M7** (68.2 mg, t_R=27.3 min). Following the same HPLC system (**Figure 15**), 40% Sep-Pak fraction of the MeOH extract gave three pure compounds **CM14** (14.1 mg, t_R=31.5 min), **CM15** (3.2 mg, t_R=33.5 min) and **CM16** (4.1 mg, t_R=35.0 min) in addition to previous seven fractions.

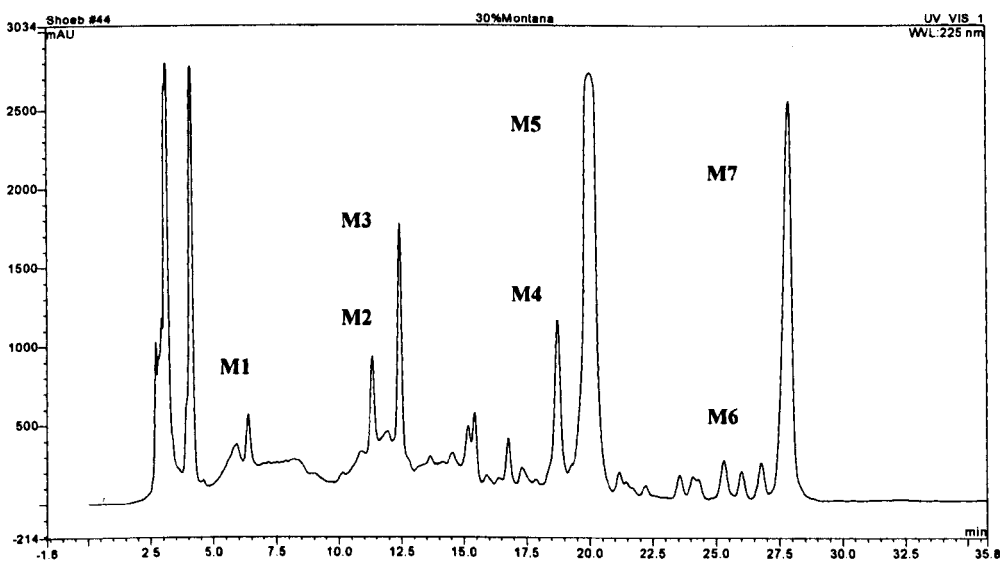


Figure 14. HPLC chromatogram of 30% Sep-Pak fraction of *C. montana*

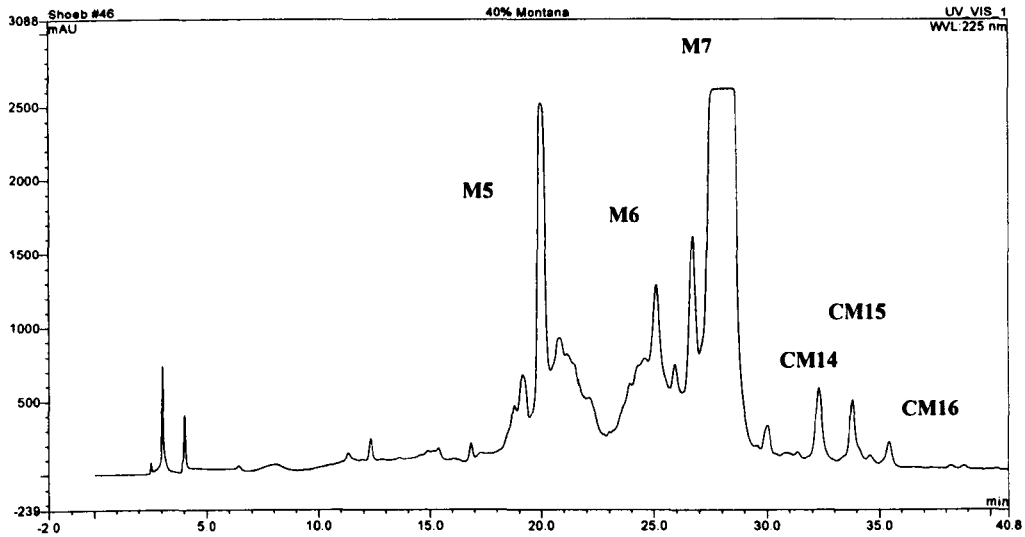


Figure 15. HPLC chromatogram of 40% Sep-Pak fraction of *C. montana*

Fraction **M1** was further purified by Semi-Prep-HPLC-2 (isocratic elution with 10% MeCN in water, 2.0 mL/min) to obtain **CM1** (7.0 mg, t_R=9.0 min), **CM2** (4.5 mg, t_R=10.0 min), **CM3** (3.0 mg, t_R=16.0 min) and **CM4** (2.0 mg, t_R=23.0 min) (**Figure 16**).

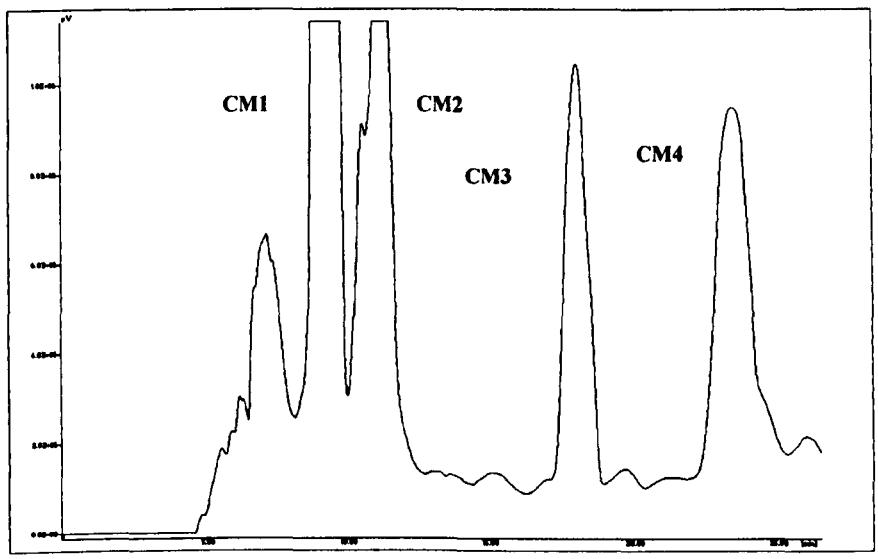


Figure 16. HPLC chromatogram of fraction **M1** from 30% Sep-Pak of *C. montana*

Fraction **M2** was further purified by Semi-Prep-HPLC-2 (isocratic elution with 12% MeCN in water, 2.0 mL/min) to obtain **CM5** (12.3 mg, t_R=42.0 min) and **CM6** (9.8 mg, t_R=70.0 min) (**Figure 17**). Pure compounds **CM7** (25.0 mg, t_R=74.0 min) and **CM8** (20 mg, t_R=84.0 min) were obtained from fraction **M5** by Prep-HPLC-1 (isocratic elution with 15% MeCN in water, 20 mL/min).

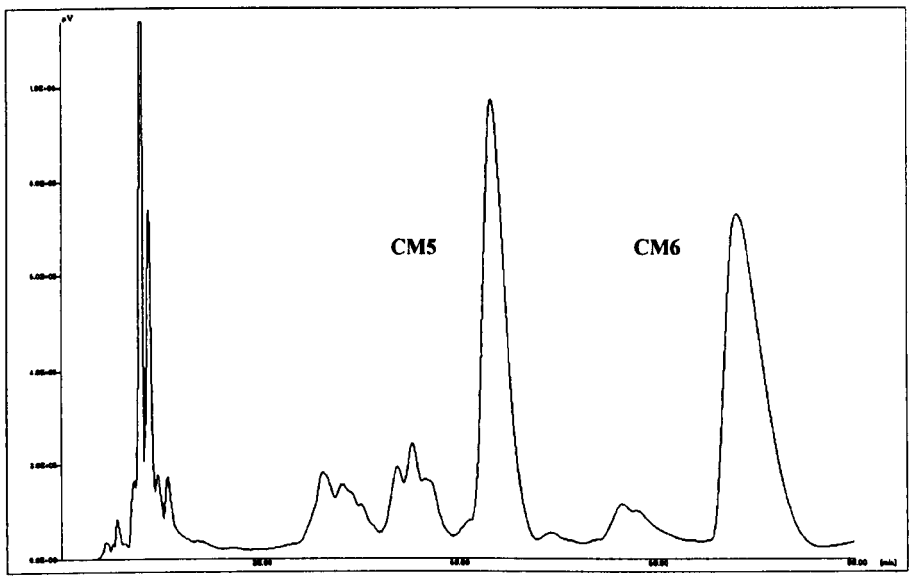


Figure 17. HPLC chromatogram of fraction **M2** from 30% Sep-Pak of *C. montana*

Fraction **M4** was purified by Semi-Prep HPLC-2 (isocratic elution with 17% MeCN in water, 2.0 mL/min) to obtain **CM9** (19.9 mg, t_R=48.0 min) in addition to **CM6** (10.1 mg, t_R=60.0 min) (**Figure 18**). **CM10** (3.5 mg, t_R=40.5 min), **CM11** (4.5 mg, t_R=49.0 min), **CM12** (6.0 mg, t_R=56.0 min) and **CM13** (22.5 mg, t_R=67.0 min) were obtained from fraction **M6** by Semi-Prep HPLC-2 (isocratic elution with 20% MeCN in water, 2.0 mL/min) (**Figure 19**).

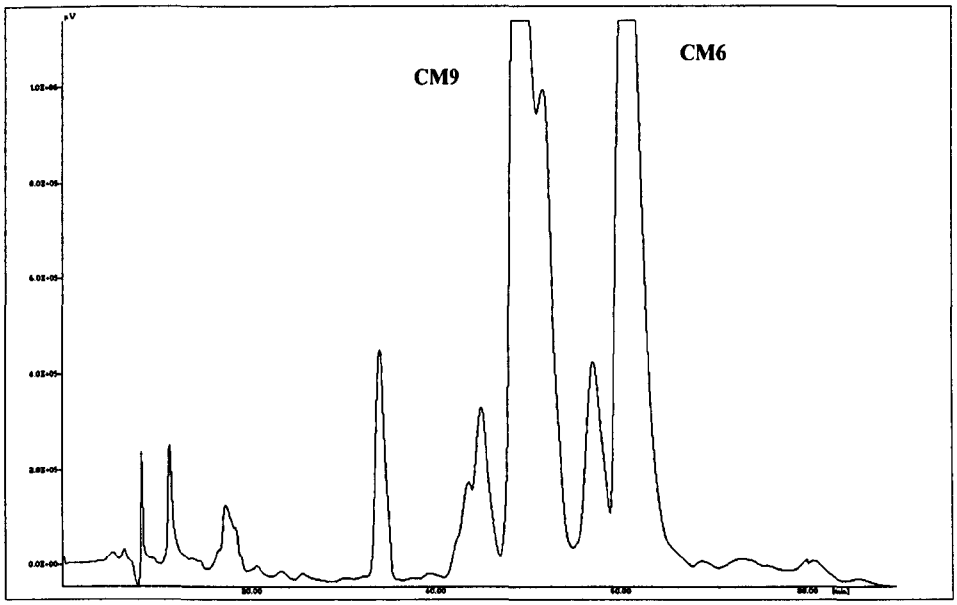


Figure 18. HPLC chromatogram of fraction **M4** from 30% Sep-Pak of *C. montana*

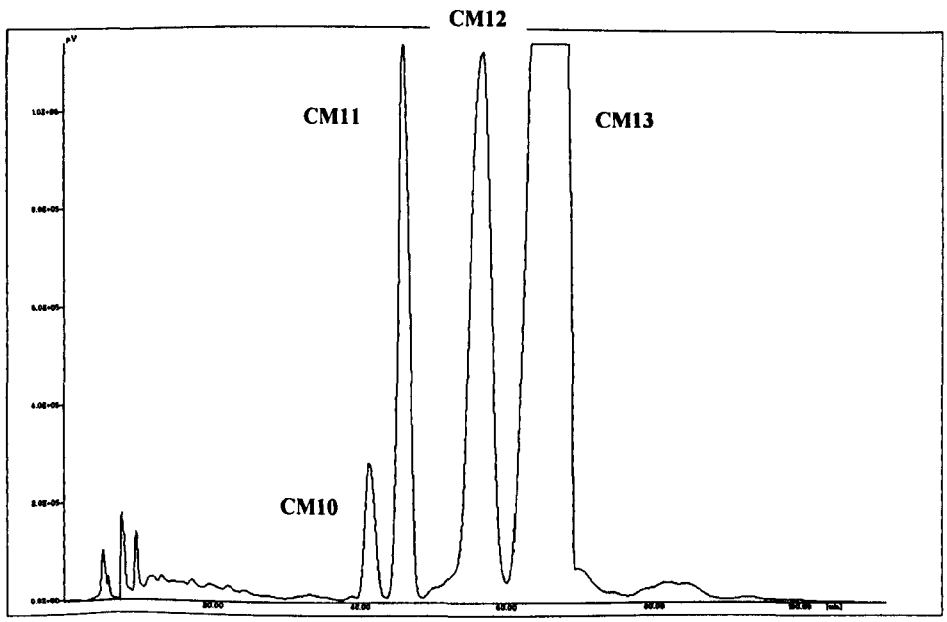


Figure 19. HPLC chromatogram of fraction **M6** from 30% Sep-Pak of *C. montana*

Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 70:30 to 0:100 over 50 min followed by 100% MeOH for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 60% MeOH, yielded two fractions **M8** (32.0 mg, t_R=17.5 min) and **M9** (33.0 mg, t_R=19.8 min) in addition to pure compound **CM13** (54.7 mg, t_R=14.3 min) (**Figure 20**). Fraction **M8** was further purified by Semi-Prep HPLC-2 (isocratic elution with 25% MeCN in water, 2.0 mL/min) to obtain **CM17** (2.0 mg, t_R=92.0 min). **CM18** (5.0 mg, t_R=31.0 min) and **CM19** (4.0 mg, t_R=34.0 min) were obtained from fraction **M9** by Prep-HPLC-1 (isocratic elution with 30% MeCN in water, 20 mL/min).

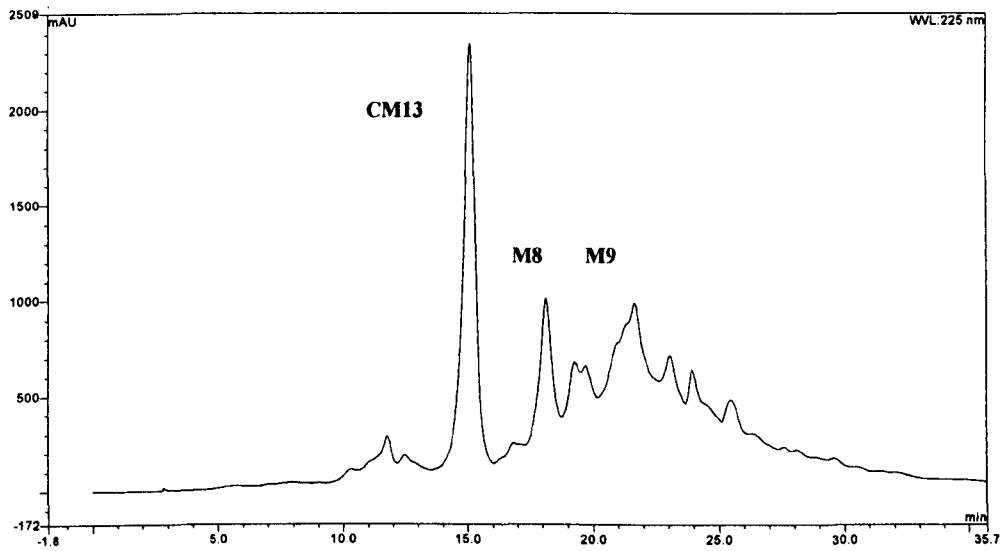


Figure 20. HPLC chromatogram of 60% Sep-Pak fraction from *C. montana*

The *n*-hexane and DCM extracts were processed and found that they contained mainly long chain fatty alcohol and acids. These extracts were not further investigated.

2.7.7 Isolation of Compounds from *Centaurea schischkinii*

The MeOH extract (2.0 g) obtained from Soxhlet-extraction method previously described on page 64 was fractionated using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 30, 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient- water:MeOH= 75:25 to 30:70 over 50 min followed by 70% MeOH for 10 min, 15 mL/min) of the Sep-Pak fraction, which was eluted with 40% MeOH, yielded fraction **S1** (14.0 mg, t_R = 8.1 min), **CS2** (59.2 mg, t_R= 24.4 min), **CS3** (518.6 mg, t_R = 28.5 min), **CS4** (11.7 mg, t_R= 32.2 min), **CS5** (29.9 mg, t_R= 33.4 min) and **CS6** (11.5 mg, t_R= 36.1 min) (**Figure 21**).

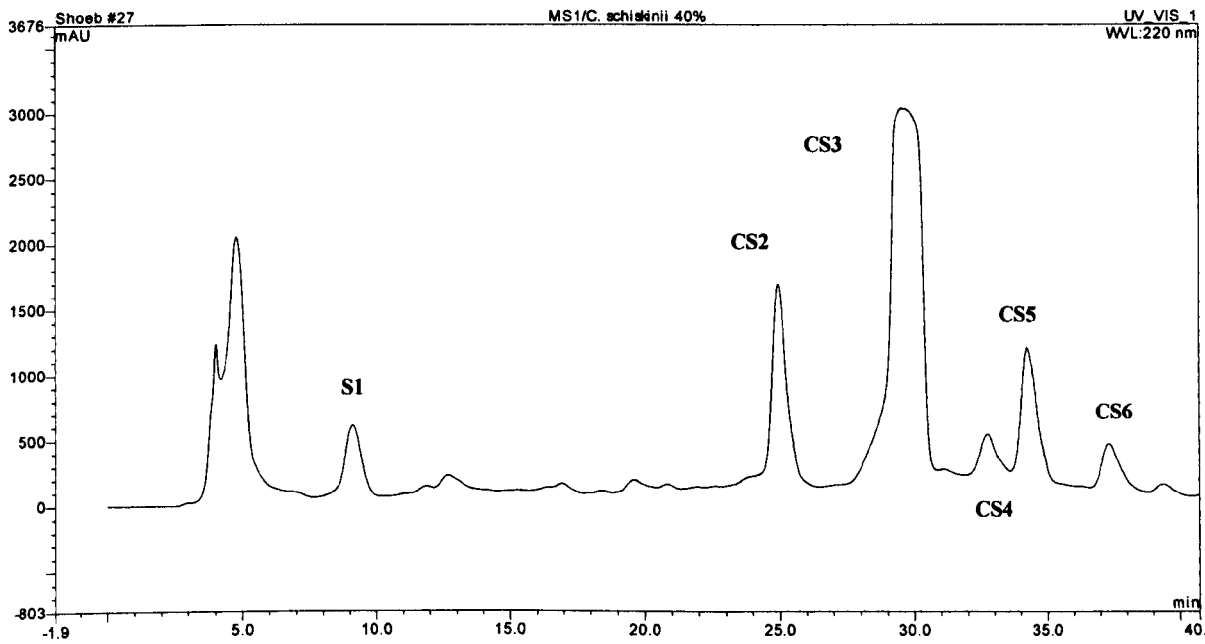


Figure 21. HPLC chromatogram of 40% Sep-Pak fraction from *C. schischkinii*

Fraction **S1** was further purified by Semi-Prep-HPLC-1 (eluted with a linear gradient-water:MeCN= 90:10 to 60:40 over 50 min followed by 40% acetronitrile for 10 min, 15 mL/min) to obtain **CS1** (7.0 mg, t_R=14.5 min). Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 60:40 to 20:80 over 50 min followed by 80% MeOH for 10 min, 15 mL/min) of the Sep-Pak fraction, which was eluted with 60% MeOH,

afforded **CS7** (13.7 mg, t_R=23.1 min) and **CS8** (27.3 mg, t_R=29.3 min) in addition to **CS3** and **CS5 (Figure 22)**.

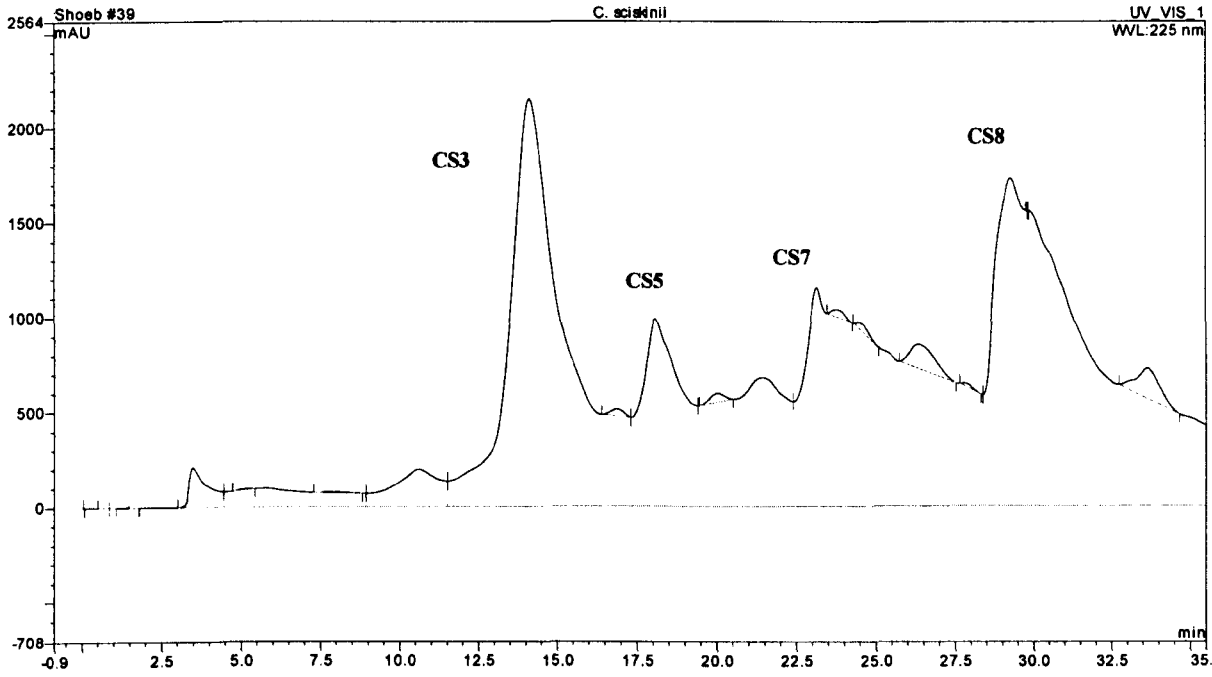


Figure 22. HPLC chromatogram of 60% Sep-Pak fraction from *C. schischkinii*

The TLC pattern of DCM extract showed the presence of similar compounds (in a negligible amount) isolated from methanol extract. So this extract was not processed further. The ^{1}H and ^{13}C NMR spectra of *n*-hexane extract suggested that it contained long chain fatty acids and alcohol which were not compounds of our interest and was not processed for further investigation.

2.7.8 Isolation of Compounds from *Centaurea bornmuelleri*

The MeOH extract (2.0 g) of *C. bornmuelleri* were applied to the reversed-phase Sep-Pak column as previously described on page 64. A step gradient of 20 to 100% aqueous MeOH in 20% increments (200 mL per each step) was used to elute different metabolites from the column and collected in five different flasks. Fraction obtained from 40% aqueous MeOH was further purified by Prep-HPLC-1 (eluted with a linear gradient- water:MeOH= 75:25 to 30:70 over 50 min followed by 70% MeOH for 10 min, 20 mL/min) to yield fraction **CB1** (13.9 mg, t_R= 21.0 min), **CB2** (535.4 mg, t_R = 25.0 min), **CB3** (5.1 mg, t_R= 27.1 min), **CB4** (10.0 mg, t_R= 29.0 min), **CB5** (25.1 mg, t_R= 31.8 min) and **CB6** (8.8 mg, t_R= 34.0 min).

The *n*-hexane extract was processed, but no compound of our interest was isolated. However, the DCM extract was fractionated by normal column chromatography (page 63) using *n*-hexane, DCM and MeOH with gradual increase in polarity. Fraction obtained from DCM-MeOH (19:1) was further purified by PTLC (page 61) using a solvent mixture of *n*-hexane-EtOAc (7:3) to afford pure compound **CB7** (4.7 mg).

2.7.9 Isolation of Compounds from *Centaurea gigantea*

The MeOH extract (2.0 g) was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 30, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient- water:MeCN= 90:10 to 60:40 over 50 min followed by 40% MeCN for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 30% MeOH (**Figure 23**), yielded compounds **CG1** (15.5 mg, t_R = 7.3 min), **CG2** (34.9 mg, t_R = 13.3 min),

CG3 (33.5 mg, t_R = 14.3 min), **CG4** (7.9 mg, t_R = 15.3 min), and **CG5** (3.5 mg, t_R = 18.0 min).

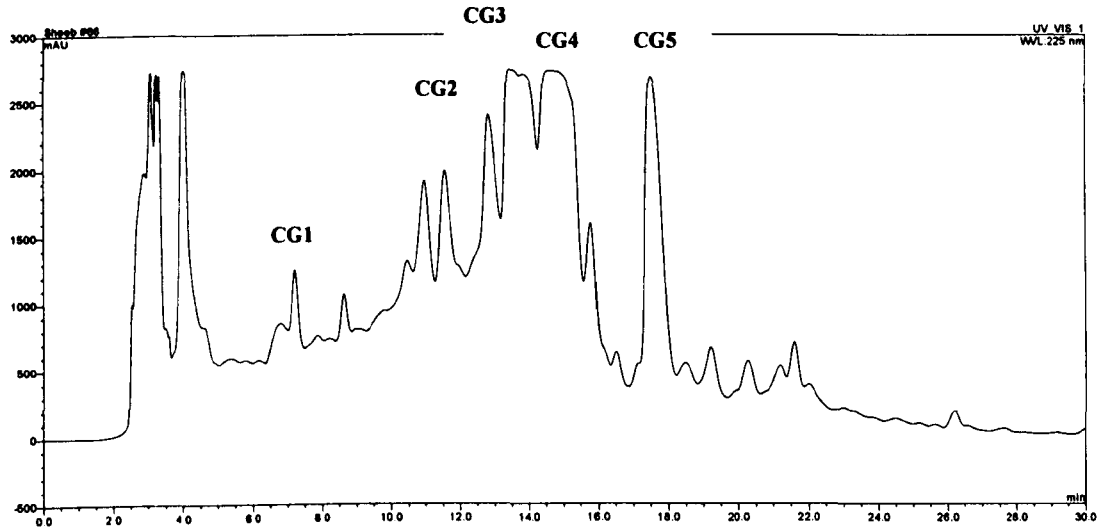


Figure 23. HPLC chromatogram of 30% Sep-Pak fraction from *C. gigantea*

Prep-HPLC-1 (eluted with a linear gradient- water:MeOH= 75:25 to 30:70 over 50 min followed by 70% MeOH for 10 min, 15 mL/min) of the Sep-Pak fraction, which was eluted with 60% MeOH (**Figure 24**), afforded **CG6** (9.0 mg, t_R=22.3 min) **CG7** (21.4 mg, t_R=24.6 min) and **CG8** (12.1 mg, t_R=25.9 min).

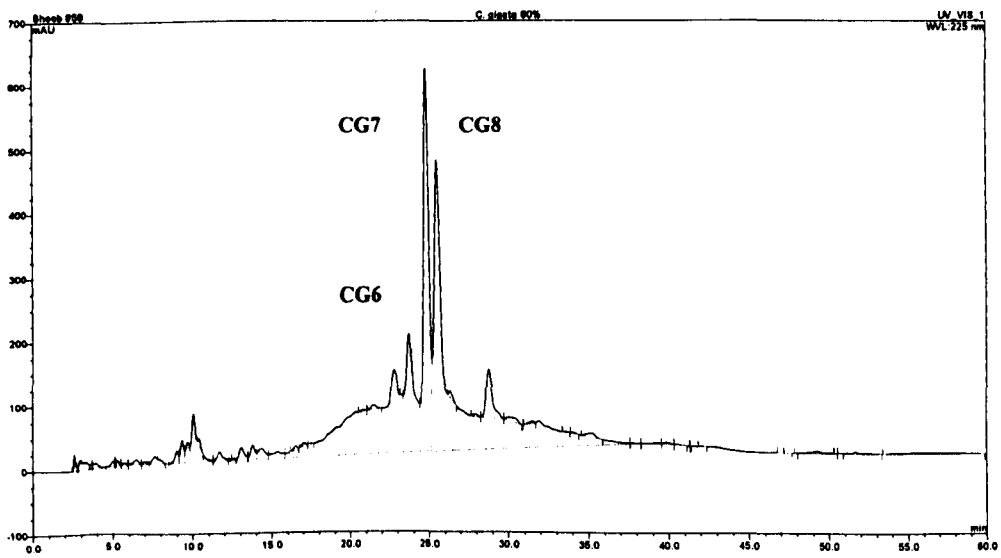


Figure 24. HPLC chromatogram of 60% Sep-Pak fraction from *C. gigantea*

The ^{1}H and ^{13}C NMR spectra of *n*-hexane extract revealed that it contained long chain fatty alcohol and hydrocarbons which were not potential candidate for our investigation. So this extract was not examined further. The DCM extract was subjected to normal phase column chromatography. The column packed in DCM and gradually increased in polarity from 0 to 50% MeOH in DCM and finally washed with 100% MeOH. After monitoring by TLC, similar fractions were pooled together and 3 pure compounds (**CG6**, **CG7** and **CG8**), which were also isolated from 60% Sep-Pak fraction of MeOH extract, were obtained.

2.7.10 Isolation of Compounds from *Centaurea pamphylica*

The methanol extract was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 40, 60, 80 and 100% MeOH in water (200 mL each). Preparative-HPLC-1 (Luna C_{18} column 10 μm, 250 mm × 21.2 mm, eluted with a linear gradient-water:MeOH= 75:25 to 30:70 over 50 min followed by 70% MeOH for 10 min, 15 mL/min) of the Sep-Pak fraction, which was eluted with 40% MeOH, yielded compound **CP1** (19.7 mg, t_R = 22.1 min), **CP2** (44.2 mg, t_R = 26.5 min) and **CP3** (12.6 mg, t_R = 30.5 min).

The *n*-hexane extract was fractionated by VLC (page 62). The column was packed in *n*-hexane and increased polarity from A to H (**Table 3**) and eight fractions were obtained. The ^{1}H and ^{13}C NMR spectra revealed that all eight fractions contained mainly long chain fatty alcohol and acids. So they were not processed further. However, DCM extract was subjected to normal phase column chromatography. The column packed in *n*-hexane-DCM (1:1) and gradually increased in polarity from I to S (**Table 3**). After monitoring by TLC, similar fractions were pooled together and

pure compound **CP4** (4.7 mg) was obtained from fraction eluted with DCM-MeOH (19:1) in addition to **CP1** and **CP2**.

2.7.11 Isolation of compounds from *Centaurea urvillei*

The MeOH extract (2.0 g) was fractionated by solid phase extraction method using a Sep-Pak C_{18} (10 g) cartridge eluting with a step gradient: 40, 60, 80 and 100% MeOH in water (200 mL each). Prep-HPLC-1 (eluted with a linear gradient-water:MeOH= 75:25 to 30:70 over 50 min followed by 70% MeOH for 10 min, 20 mL/min) of the Sep-Pak fraction, which was eluted with 40% MeOH, yielded compound **CU1** (12.4 mg, t_R = 20.6 min) and **CU2** (56.0 mg, t_R = 24.5 min).

The *n*-hexane extract was processed, and found that this extract contained mainly long chain fatty alcohol, fatty acids and esters. As a result, no compound of our interest was isolated. ^1H and ^{13}C NMR of DCM extract revealed that it contained similar compounds isolated from MeOH extract and was not investigated further.

2.7.12 Isolation of Compounds from *Centaurea mucronifera*

The *n*-hexane (100 mg) extract of *C. mucronifera* was subjected to normal phase flash chromatography. The column was packed in *n*-hexane and gradually increased in polarity from 0 to 100% DCM in *n*-hexane and finally washed with 100% MeOH. After monitoring by TLC, similar fractions were pooled together and 5 fractions were obtained. The fraction obtained from DCM-*n*-hexane (3:2) was washed several times with the same solvent mixture and collected as pure compound **CMu1** (6.2 mg) which gave single spot on TLC. The DCM extract was subjected to normal phase column chromatography. The column packed in *n*-hexane-DCM (1:1) and

gradually increased in polarity from I to S (**Table 3**). After monitoring by TLC, similar fractions were pooled together and pure compound **CMu2** (15.8 mg) and **CMu3** (8.7 mg) were obtained from fraction eluted with DCM-MeOH (18:1).

The TLC, ^{1}H and ^{13}C NMR spectra of MeOH extract showed the presence of similar compounds (**CMu2** and **CMu3**) isolated from the DCM extract. So this extract was not processed further.

CHAPTER THREE

Results and Discussion

3 Results and Discussion

3.1 Lignans

A total of fifteen lignans were isolated from *Centaurea* species in this study. Six were dibenzylbutyrolactone-type (**SM1-SM6**) and nine were epoxy-type (**SM7-SM15**) lignans. Lignan **SM5**, isolated from *C. americana*, is a new natural product.

3.1.1 Dibenzylbutyrolactone Lignans

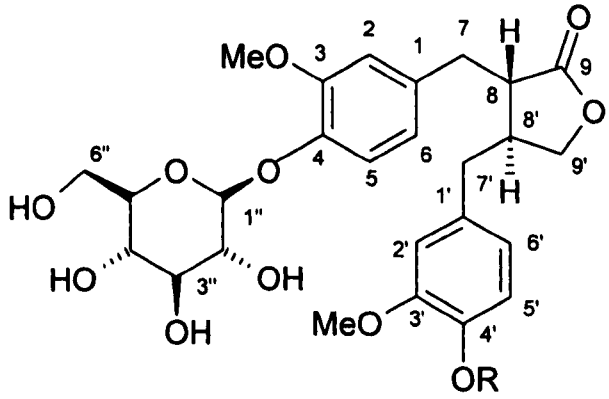
Six dibenzylbutyrolactone-type lignans isolated from different *Centaurea* species were given new code in **Table 4**.

Table 4. Dibenzylbutyrolactone-type lignan isolated from *Centaurea* species

Code (structure)	Source and isolation code	Page
SM1 (137)	*C. macrocephala* (**CMa1**), *C. schischkinii* (**CS3**),	71, 80,
	C. huber-morathii (**CH2**), *C. bornmuelleri* (**CB2**),	74, 82,
	C. mucronifera (**CMu3**), *C. urvillei* (**CU2**),	86, 85,
	C. americana (**CA3**), *C. dealbata* (**CD2**), *C. pamphylica* (**CP2**)	74, 73, 84
SM2 (138)	*C. macrocephala* (**CMa2**), *C. schischkinii* (**CS2**),	71, 80,
	C. huber-morathii (**CH1**), *C. bornmuelleri* (**CB1**),	74, 82,
	C. mucronifera (**CMu2**), *C. urvillei* (**CU1**),	86, 85,
	C. americana (**CA1**), *C. dealbata* (**CD1**), *C. pamphylica* (**CP1**)	74, 73, 84
SM3 (139)	*C. schischkinii* (**CS7**), *C. bornmuelleri* (**CB6**),	81, 82,
	C. americana (**CA6**), *C. dealbata* (**CD3**)	74, 73
SM4 (140)	*C. macrocephala* (**CMa3**), *C. schischkinii* (**CS5**),	71, 80,
	C. huber-morathii (**CH3**), *C. bornmuelleri* (**CB4**),	74, 82,
	C. pamphylica (**CA5**)	84
SM5 (170)	*C. americana* (**CA7**)	74
SM6 (143)	*C. macrocephala* (**CMa4**),*C. americana* (**CA8**)	71, 74

3.1.1.1 Characterisation of SM1, SM2, SM3 and SM4 as Arctiin (137), Matairesinoside (138), Arctigenin (139) and Matairesinol (140), respectively

All four compounds (**SM1-SM4**) displayed characteristic UV absorption maxima of dibenzylbutyrolactone type lignans. The presence of two strong UV absorbance signals at 279-282 and 225-228 nm in the UV spectrum indicated that they had aromatic benzene rings and typical of a lignan type structure.[152] A strong absorption band at 1765 cm^{-1} in the IR spectrum of each compound could be attributed to the carbonyl functionality of the lactone ring. The HRCIMS spectrum of **SM1** revealed the $[M+NH_4]^+$ ion peak at *m/z* 552.2441, suggesting *Mr*=534 and the molecular formula $C_{27}H_{34}O_{11}$ (calculated 552.2439 for $C_{27}H_{38}NO_{11}$). The ^{13}C NMR spectrum (**Figure 79**; **Table 5**) showed the presence of twentyseven carbons. The DEPT-135 spectrum indicated three methoxy groups (δ 55.7, 55.6, 55.5), four methylenes (δ 71.9, 61.6, 38.0, 34.4), thirteen methines (δ 122.0, 121.1, 117.0, 113.9, 112.7, 112.2, 102.0, 77.2, 77.0, 74.0, 70.4, 46.7, 41.5), six quarternary carbons (δ 149.8, 149.6, 148.3, 145.9, 133.3, 131.8) and one carbonyl (δ 180.4).



	R
137 (SM1)	Me
138 (SM2)	H

Figure 25. Structure of arctiin (**SM1**) and matairesinoside (**SM2**)

The ^{1}H NMR spectrum (**Figure 78; Table 5**) revealed the presence of two ABX benzene ring systems by the signals at δ_H 7.04 (d, *J*=8.2 Hz), 6.64 (dd, *J*= 8.2, 1.8 Hz), 6.74 (d, *J*=1.8 Hz), and δ 6.81 (d, *J*=8.4 Hz), 6.58 (d, *J*=2.0 Hz) and 6.59 (dd, *J*=8.4, 2.0 Hz). The ^{1}H NMR spectrum also showed additional signals for one oxymethylenes [δ_H 4.18 (dd, *J*=8.9, 7.6 Hz) and 3.93 (dd, *J*=7.8, 7.6 Hz)], two methylenes [δ_H 2.89 (dd, *J*=14.1, 7.3 Hz) and 2.83 (dd, *J*=14.1, 6.2 Hz), and 2.56 (m)], and two methines [δ_H 2.67 (m) and 2.49 (m)]. The ^{1}H and ^{13}C NMR data suggested that **SM1** was a dibenzylbutyrolactone lignan.[152] This was further confirmed by ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC experiments (**Figure 80 and 81**). Six carbon signals at δ 102.0, 77.2, 77.0, 74.0, 70.4 and 61.6 in the ^{13}C NMR could be assigned to a glucose moiety, and the doublet at δ_H 4.84 (*J*=7.6 Hz) was attributable to the anomeric proton with a β-configuration.[153] A ^{3}J ^{1}H-^{13}C long-range correlation between the anomeric proton, H-1'' (δ 4.84) and the aromatic quarternary δ_C 145.9 (C-4) in the HMBC spectrum confirmed that the glucose moiety was connected to C-4 of the dibenzylbutyrolactone ring. This was further confirmed by NOE interaction between the anomeric proton (H-1'') and H-5 (δ 7.04) of benzene ring in the ^{1}H-^{1}H NOESY experiment (**Figure 26**). The ^{1}H-^{1}H NOESY spectrum (**Figure 82**) of **SM1** also showed strong NOE interactions between H-7 (δ 2.89) and H-2 (δ 6.74), H-8 (δ 2.67) and H-7' (δ 2.56), H-8 and H-7, H-9' (δ 4.18) and H-7', and H-5 (δ 7.04) and H-6. The ^{1}H and ^{13}C NMR spectral data of **SM1** were similar to those published for arctiin[154,155] and a combination of COSY, HMBC and NOESY 2D NMR spectral analyses confirmed unequivocally its identity as arctiin (**137**). Arctiin was previously isolated from *C. americana*[51], *C. isaurica*[83] and *C. nigra*[105] (**Table 1**).

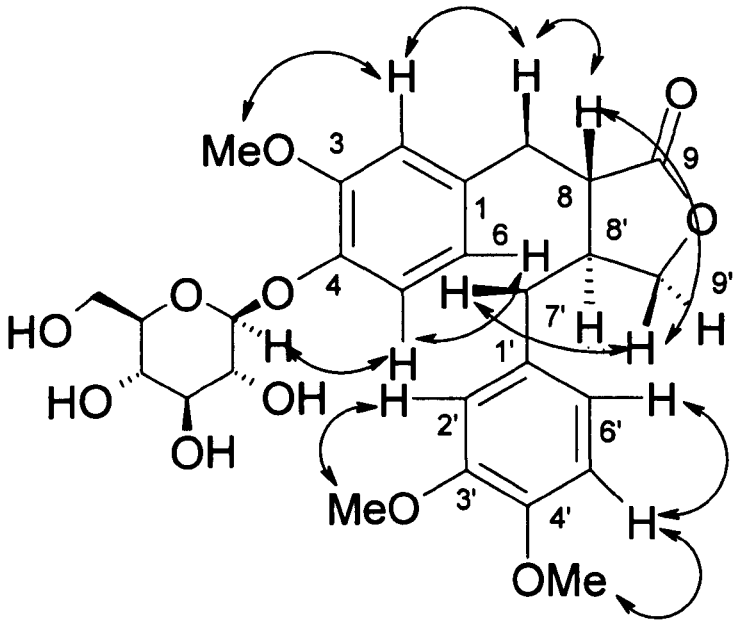


Figure 26. Key NOE interactions of **SM1** (**137**) based on ^{1}H-^{1}H NOESY experiment

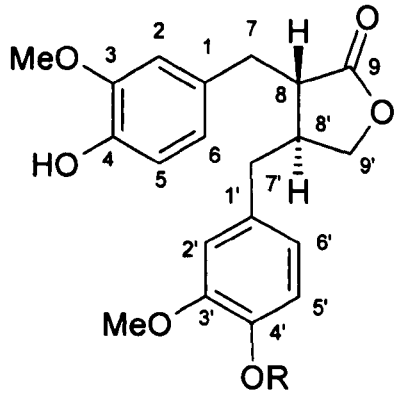
The ^{1}H and ^{13}C NMR spectra of **SM2** (**Table 5**) displayed signals similar to those of **SM1** with the exceptions that **SM2** had signals for two methoxy groups instead of three. The ESIMS spectrum of **SM2** revealed the pseudomolecular ion, $[M+Na]^{+}$ peak at *m/z* 543, suggesting *Mr*=520, and the molecular formula $C_{26}H_{32}O_{11}$. This mass data confirmed the findings from the NMR data of **SM2** that this compound contained 14 mass units less than arctiin (**SM1**). The ^{1}H and ^{13}C NMR data of **SM2** were in good agreement with the published data of matairesinoside.[154,155] Thus, the structure of **SM2** was elucidated as matairesinoside (**138**). Matairesinoside was previously isolated from *C. americana*[51], *C. nigra*[105], *C. scabiosa*[115] (**Table 1**).

Table 5. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz) and ^{13}C NMR data for **SM1 (137)** and **SM2 (138)**

Carbon number	Chemical shift δ in ppm			
	^{1}H		^{13}C	
	SM1[a]	**SM2**[b]	**SM1**[a]	**SM2**[b]
1	–	–	133.3	130.1
2	6.74, d, 1.8	6.54, d, 1.6	113.9	112.1
3	–	–	149.6	147.8
4	–	–	145.9	145.7
5	7.04, d, 8.2	6.66, d, 8.4	117.0	115.0
6	6.64, dd, 8.2, 1.8	6.64, dd, 8.4, 2.0	122.0	121.8
7	2.89, dd, 14.1, 7.3 2.83, dd, 14.1, 6.2	2.84, dd, 14.0, 5.3 2.76, dd, 14.0, 6.5	34.4	34.1
8	2.67, m	2.64, m	46.7	46.4
9	–	–	180.4	180.2
1'	–	–	131.8	133.1
2'	6.58, d, 2.0	6.72, d, 2.0	112.7	113.7
3'	–	–	149.8	149.5
4'	–	–	148.3	145.0
5'	6.81, d, 8.4	7.2, d, 8.0	112.2	116.7
6'	6.59, dd, 8.4, 2.0	6.48, dd, 8.0, 2.0	121.1	121.0
7'	2.56, m	2.58, m	38.0	37.8
8'	2.49, m	2.46, m	41.5	41.4
9'	4.18, dd, 8.9, 7.6 3.93, dd, 7.8, 7.6	4.15, dd, 8.7, 7.5 3.90, dd, 8.4, 7.5	71.9	71.7
1''	4.84, d, 7.3	4.80, d, 7.6	102.0	101.7
2''	3.45, m	3.44, m	74.0	73.7
3''	3.41, dd, 6.8, 4.8	3.44, m	77.2	76.6
4''	3.38, dd, 4.8, 3.2	3.36, m	70.4	70.1
5''	3.37, m	3.35, m	77.0	76.9
6''	3.84, m 3.67, m	3.76, m 3.66, m	61.6	61.3
O-CH_3 (3)	3.73, s	3.77, s	55.6	55.5
O-CH_3 (3')	3.78, s	3.74, s	55.5	55.2
O-CH_3 (4')	3.78, s	–	55.7	–

[a] ^{1}H NMR (600 MHz) and ^{13}C NMR (150 MHz) in CD_3OD
[b] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

The ^{1}H and ^{13}C NMR spectra (**Table 6**) of **SM3** and **SM4** showed signals similar to those of **SM1** and **SM2,** but lacked signals for the glucose moiety. This was further confirmed from their mass analyses. The ESIMS spectra of **SM3** and **SM4** showed $[M+Na]^{+}$ ions at *m/z* 395 and 381, respectively, suggesting their molecular formula as $C_{21}H_{24}O_6$ and $C_{20}H_{22}O_6$. Their NMR data were compared with published data and found to be identical with arctigenin (**139**) and matairesinol (**140**), respectively.[154,155] However, detailed 2D NMR spectral analyses were carried out to assign unambiguously all the ^{1}H and ^{13}C NMR signals of **SM3** (**139**) and **SM4** (**140**). Arctigenin (**139**) was previously isolated from *C. affinis*[47], *C. calcitrapa*[65], *C. macrocephala*[89], *C. nigra*[105] and *C. phrygia*[100] whereas *C. affinis*[47], *C. calcitrapa*[65], *C. persica*[87], *C. nicolai*[104], *C.nigra*[105] and *C. raphanina ssp*[109] afforded matairesinol (**140**) (**Table 1**).



	R
139 (SM3)	Me
140 (SM4)	H

Figure 27. Structures of arctigenin (**SM3**) and matairesinol (**SM4**)

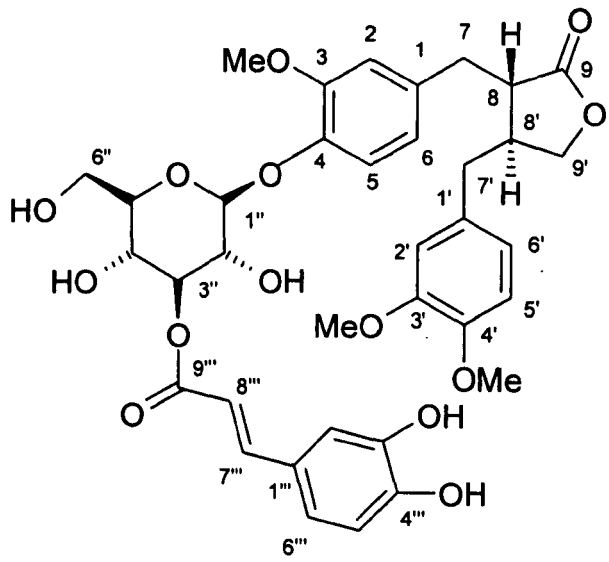
Table 6. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz) and ^{13}C NMR data for **SM3 (139)** and **SM4 (140)**

Carbon number	Chemical shift δ in ppm			
	^{1}H		^{13}C	
	SM3[a]	**SM4**[a]	**SM3**[a]	**SM4**[a]
1	–	–	129.5	130.2
2	6.63, d, 1.6	6.63, d, 2.0	112.7	112.7
3	–	–	148.0	147.8
4	–	–	145.2	145.0
5	6.66, d, 8.4	6.67, d, 8.4	114.9	114.9
6	6.55, dd, 8.4, 1.6	6.54, dd, 8.4, 2.0	121.8	121.8
7	2.85, dd, 14.0, 5.6 2.77, dd, 14.0, 7.2	2.81, dd, 14.0, 5.6 2.78, dd, 14.0, 6.8	34.2	34.1
8	2.66, m	2.61, m	46.5	46.5
9	–	–	180.3	180.4
1'	–	–	131.6	129.5
2'	6.52, d, 2.0	6.51, d, 1.6	111.9	112.1
3'	–	–	147.8	147.8
4'	–	–	149.3	145.2
5'	6.78, d, 8.0	6.64, d, 8.0	112.4	115.0
6'	6.56, dd, 8.0, 2.0	6.46, dd, 8.0, 1.6	120.8	121.0
7'	2.50, m	2.50, m	37.7	37.7
8'	2.47, m	2.48, m	41.2	41.3
9'	4.13, dd, 9.2, 7.6 3.89, dd, 7.2, 7.6	4.11, dd, 9.2, 7.6 3.82, dd, 8.8, 7.6	71.7	71.7
O-CH_3 (3)	3.75, s	3.70, s	55.3	55.1
O-CH_3 (3')	3.74, s	3.70, s	55.2	55.1
O-CH_3 (4')	3.72, s	–	55.1	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.1.1.2 Characterisation of SM5 as 3''-*O*-Caffeoyl-(9'''→3'')-arctiin (170)

An ESIMS spectrum of **SM5** displayed the pseudomolecular ion peak at *m/z* 719 $[M+Na]^+$, suggesting *Mr*=696 and the molecular formula $C_{36}H_{40}O_{14}$. The HRCIMS spectrum showed *m/z* 714.2759 $[M+NH_4]^+$ (calculated 714.2756 for $C_{36}H_{44}NO_{14}$). The ^1H and ^{13}C NMR data (**Table 7**) indicated that **SM5** was an arctiin (**SM1**) derivative. The Arctiin (**SM1**) moiety was also confirmed by ^1H-^1H COSY (**Figure 89**) and ^1H-^{13}C HMBC experiments (**Figure 90**; **Table 7**). The signals for a trisubstituted benzene ring [δ_H 7.03 (d, *J*=2.0 Hz), 6.93 (dd, *J*=8.0, 2.0 Hz) and 6.74 (d, *J*=8.0 Hz)], and two olefinic protons [δ_H 7.57 (d, *J*=16.0 Hz) and 6.32 (d, *J*=16.0 Hz)] in the ^1H NMR spectrum together with signals at δ_C 167.8, 149.3, 148.3, 145.6, 126.7, 121.8, 115.3, 114.2 and 114.0 for nine carbons in the ^{13}C NMR spectrum suggested that the other part of the molecule was a caffeoyl moiety. The ^3J ^1H-^{13}C long-range correlations from δ_H 6.74 (H-5''') to δ_C 126.7 (C-1'''), δ_H 6.93 (H-6''') to δ_C 114.2 (C-2''') and 148.3 (C-4'''), δ_H 7.57 (H-7''') to δ_C 114.2 (C-2''') and 167.8 (C-9''') in the ^1H-^{13}C HMBC spectrum also confirmed the presence of the caffeoyl moiety. A ^3J ^1H-^{13}C long-range HMBC correlation between δ_H 5.12 (H-3'') and 167.8 (C-9''') confirmed that caffeoyl group was attached to glucose molecule at C-3''. Thus, **SM5,** isolated from *C. americana,* was confirmed as 3''-*O*-caffeoyl-(9'''→3'')-arctiin (**170**). To the best of our knowledge this is a new natural product.



170

Figure 28. Structure of 3''-*O*-caffeoyl-(9'''→3'')-arctiin (**SM5**)

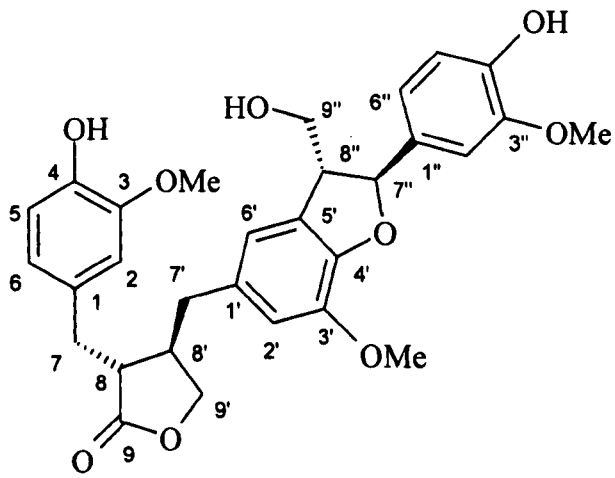
Table 7. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM5 (170)**

Carbon number	Chemical shift δ in ppm		HMBC correlations ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	–	131.5	–	–
2	6.57, d, 2.0	112.4	C-3'	C-6
3	–	149.6	–	–
4	–	148.0	–	–
5	6.79, d, 8.8	111.9	–	C-1, C-3
6	6.56, m	121.0	C-5	
7	2.81, dd, 14.1, 6.1	34.2	C-1	C-2, C-6, C-9
	2.87, dd, 14.1, 6.8			
8	2.65, m	46.4	C-9, C-8'	
9	–	180.2	–	–
1'	–	133.5	–	–
2'	6.88, d, 2.0	109.2	–	–
3'	–	149.3	–	–
4'	–	145.8	–	–
5'	7.04, d, 8.4	116.8		C-1', C-3'
6'	6.71, m	121.7	C-5'	C-4'
7'	2.50, m	37.7	–	–
8'	2.41, m	41.3	–	–
9'	4.16, 7.6, 9.2	71.7	–	C-9
	3.91, 7.6, 9.2			
1''	4.96, d, 7.2	101.1	–	C-4
2''	3.65, m	68.4	–	–
3''	5.12, m	77.4	C-4''	C-9'''
4''	3.62, m	72.1	–	C-2'''
5''	3.48, m	76.8	–	–
6''	3.84, m	61.0	–	–
	3.71, m			
1'''		126.7	–	–
2'''	7.03, d, 2.0	114.2	–	–
3'''		149.3	–	–
4'''		148.3	–	–
5'''	6.74, d, 8.0	115.3	–	C-1'''
6'''	6.93, dd, 8.0, 2.0	121.8	–	C-2''', C-4'''
7'''	7.57, d, 16.0	145.6	C-1'''	C-2''', C-9'''
8'''	6.32, d, 16.0	114.0	–	C-1'''
9'''	–	167.8	–	–
O-CH_3 (3)	3.72, s	55.4	–	C-3
O-CH_3 (3')	3.71, s	55.5	–	C-3'
O-CH_3 (4')	3.70, s	55.6	–	C-4

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.1.1.3 Characterisation of SM6 as Lappaol A (143)

The ESIMS spectrum of **SM6** revealed the pseudomolecular ion, $[M+Na]^+$ peak at *m/z* 559, indicating *Mr*=536, and the molecular formula $C_{30}H_{32}O_9$. Signals at δ_H 6.95 (d, *J*=2.0 Hz), 6.82 (dd, *J*=8.0, 2.0 Hz), 6.76 (d, *J*=8.0 Hz), 6.70 (m), 6.68 (d, *J*=8.0 Hz) and 6.57 (brs) observed in the ^{1}H NMR spectrum (**Table 8**) could be assigned to two (1,3,4)- trisubstituted benzene rings, while the signals at δ_H 6.59 (d, *J*=2.0 Hz) and 6.55 (d, *J*=2.0 Hz) were due to the protons of a (1,3,4,5)-tetra-substituted benzene ring. The ^{1}H NMR spectrum also showed additional signals for one oxymethylenes [δ_H 4.20 (dd, *J*=9.0, 7.5 Hz) and 3.92 (dd, *J*=9.0, 7.0 Hz)], two methylenes [δ_H 2.92 (dd, *J*=14.0, 5.4 Hz) and 2.85 (dd, *J*=14.0, 6.6 Hz), and 2.67 (dd, *J*=13.5, 6.5 Hz) and 2.56 (m)] and two methines [δ_H 2.54 (m) and 2.52 (m)]. The ^{1}H-^{13}C HMBC correlations (**Table 8**) from H-7 to C-9 and C-8', H-7' to C-9', C-2' and C-6', H-8' to C-9' and C-1', and H-9' to C-9, C-7' and C-8 confirmed the presence of dibenzylbutyrolactone structure, similar to that of **SM1-SM4**.



143

Figure 29. Structure of lappaol A (**SM6**)

The ^{1}H NMR signals at δ_H 5.54, 3.46 and 3.82, and the ^{13}C NMR chemical shifts at δ_C 89.2, 55.0 and 65.2 could be assigned to the protons and carbons of a dihydrofuran ring system, and a primary alcohol group. The ^{1}H-^{1}H COSY spectrum displayed correlation between H-7'' and H-8'', and H-8'' and H-9''. The ^{1}H-^{13}C HMBC correlations from H-7'' to C-2'', C-6'', C-4' and C-5', H-8'' to C-7'', C-9'', C-1'', C-4' and C-5', and H-2'' to C-7'' confirmed that the dihydrofuran system was fused with one of the two phenyl rings of the dibenzylbutyrolactone system to form the structure of lappaol A (**143**). The spectroscopic data of **SM6** were in good agreement with the published data of sesquilignan lappaol A (**143**).[156,157] The relative stereochemistry at the chiral centres in **SM6** was assigned by direct comparison of its ^{1}H and ^{13}C NMR data with published data.[156] The unambiguous assignment of all ^{1}H and ^{13}C NMR chemical shifts, based on extensive 2D NMR analysis of **SM6**, has been presented here for the first time. Lappaol A (**143**) was previously isolated from *C. napifolia*[98] and *C. nicaensis*[101] (**Table 1**).

Table 8. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR and long-range HMBC for **SM6 (143)**

Carbon number	Chemical shift δ in ppm		HMBC correlations ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	–	130.9	–	–
2	6.57, brs	113.6	C-3	C-4, C-6, C-7
3	–	149.1	–	–
4	–	146.5	–	–
5	6.68, d, 8.0	115.1	C-4, C-6	C-1, C-3
6	6.70, m	119.2	C-1	C-4, C-7
7	2.92, dd, 14.0, 5.4 2.85, dd, 14.0, 6.6	35.7	C-1, C-8	C-6, C-9, C-8'
8	2.54, m	48.0	C-9	–
9	–	181.8	–	–
1'	–	131.2	–	–
2'	6.59, d, 2.0	113.2	–	C-7'
3'	–	146.6	–	–
4'	–	147.5	–	–
5'	–	132.0	–	–
6'	6.55, d, 2.0	123.0	–	C-2', C-7'
7'	2.67, dd, 13.5, 6.5 2.56, m	39.1	C-8'	C-2', C6', C-9'
8'	2.52, m	42.8	C-9'	C-1'
9'	4.20, dd, 9.0, 7.5 3.92, dd, 9.0, 7.0	73.1	C-8'	C-8, C-9, C-7'
1''		134.9	–	–
2''	6.95, d, 2.0	110.7	C-1'', C-3''	C-4'', C-6'', C-7''
3''	–	149.2	–	–
4''	–	147.7	–	–
5''	6.76, d, 8.0	116.4	C-4'', C-6''	C-1'', C-3''
6''	6.82, dd, 8.0, 2.0	119.9		C-2'', C-4'', C-7''
7''	5.54, dd, 6.0	89.2	C-1'', C-8''	C-4', C-5', C-2'', C-6'', C-9''
8''	3.46, br. qr., 6.0	55.0	C-5', C-7'', C-9''	C-4', C-6', C-1''
9''	3.82, m	65.2	–	C-5'
O-CH_3 (3)	3.73, s	56.9	–	C-3
O-CH_3 (3')	3.90, s	56.5	–	C-3'
O-CH_3 (3'')	3.75, s	56.5	–	C-3''

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.1.1.4 Properties of Dibenzylbutryolactone Lignans (SM1 to SM6)

3.1.1.4.1 Properties of Arctiin, 137 (SM1)

Gum, $[\alpha]^{23}_{D}$ -55.3° (c 0.0033, MeOH) (Lit.[154] -52.3°, MeOH); UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3459, 1765, 1591, 1514, 1460 and 1266 cm^{-1}; CIMS *m/z* 552 $[M+NH_4]^+$; HRCIMS *m/z* 552.2441 $[M+NH_4]^+$ (calculated 552.2439 for $C_{27}H_{38}NO_{11}$); ^1H NMR (600 MHz, CD_3OD): **Table 5**; ^{13}C NMR (150 MHz, CD_3OD): **Table 5**.

3.1.1.4.2 Properties of Matairesinoside, 138 (SM2)

Gum, $[\alpha]^{23}_{D}$ -48.8° (c 0.002, MeOH) (Lit.[154] -43.2°, MeOH); UV λ_{max} (MeOH): 279, 222 nm; IR ν_{max} (neat): 3373, 1760, 1600, 1514, 1452 and 1270 cm^{-1}; ESIMS *m/z* 543 $[M+Na]^+$; Molecular formula $C_{26}H_{32}O_{11}$, ^1H NMR (400 MHz, CD_3OD): **Table 5**; ^{13}C NMR (100 MHz, CD_3OD): **Table 5**.

3.1.1.4.3 Properties of Arctigenin, 139 (SM3)

Gum, $[\alpha]^{23}_{D}$ -42.6° (c 0.0015, MeOH) (Lit.[154] -30.0°, MeOH); UV λ_{max} (MeOH): 281, 220 nm; IR ν_{max} (neat): 3367, 1757, 1595, 1512, 1449 and 1267 cm^{-1}; ESIMS *m/z* 395 $[M+Na]^+$; Molecular formula $C_{21}H_{24}O_6$; ^1H NMR (400 MHz, CD_3OD): **Table 6**; ^{13}C NMR (100 MHz, CD_3OD): **Table 6**.

3.1.1.4.4 Properties of Matairesinol, 140 (SM4)

Gum, $[\alpha]^{23}_{D}$ -47.2° (c 0.0022, MeOH) (Lit.[154] -50.0°, MeOH); UV λ_{max} (MeOH): 282, 228 nm; IR ν_{max} (neat): 3430, 1760, 1610 and 1525 cm^{-1}; ESIMS *m/z* 381 $[M+Na]^+$; Molecular formula $C_{20}H_{22}O_6$; ^1H NMR (400 MHz, CD_3OD): **Table 6**; ^{13}C NMR (100 MHz, CD_3OD): **Table 6**.

3.1.1.4.5 Properties of 3''-*O*-Caffeoyl-(9'''→3'')-arctiin, 170 (SM5)

Amorphous, $[\alpha]^{23}_{D}$ -34.5° (c 0.001, MeOH); UV λ_{max} (MeOH): 282, 225 nm; IR ν_{max} (neat): 3410, 1765, 1591, 1514 and 850 cm^{-1}; ESIMS *m/z* 719 $[M+Na]^+$; HRCIMS 714.2759 $[M+NH_4]^+$ (calculated 714.2756 for $C_{36}H_{44}NO_{14}$); ^1H NMR (400 MHz, CD_3OD): **Table 7**; ^{13}C NMR (100 MHz, CD_3OD): **Table 7**.

3.1.1.4.6 Properties of Lappaol A, 143 (SM6)

Amorphous, $[\alpha]^{23}_{D}$ -17.6° (c 0.0021, MeOH) (Lit.[157] -17.4°, MeOH); UV λ_{max} (MeOH): 281, 225 nm; IR ν_{max} (neat): 3425, 1764, 1591, 1520 and 858 cm^{-1}; ESIMS *m/z* 559 $[M+Na]^+$; Molecular formula $C_{30}H_{32}O_9$; ^1H NMR (400 MHz, CD_3OD): **Table 8**; ^{13}C NMR (100 MHz, CD_3OD): **Table 8**.

3.1.2 Epoxy lignans

Nine epoxy lignans have been isolated from *C. cyanus* and *C. montana* (**Table 9**).

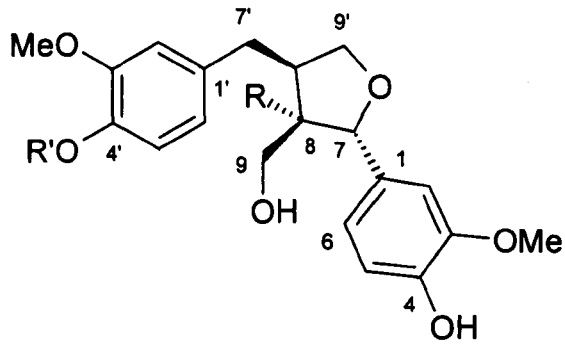
Table 9. Epoxy lignans isolated from *Centaurea* species

Code (structure)	Source and isolation code	Page
SM7 (171)	*C. cyanus* (**CC1**)	72
SM8 (172)	*C. cyanus* (**CC2**),	72,
	C. montana (**CM7**)	77
SM9 (173)	*C. montana* (**CM5**)	77
SM10 (144)	*C. montana* (**CM14**)	75
SM11 (174)	*C. montana* (**CM15**)	75
SM12 (175)	*C. montana* (**CM16**)	75
SM13 (176)	*C. montana* (**CM8**)	77
SM14 (177)	*C. montana* (**CM6**)	77
SM15 (178)	*C. montana* (**CM9**)	77

3.1.2.1 Characterisation of SM7, SM8 and SM9 as Lariciresinol 4'-*O*-β-D-glucopyranoside (171), Berchemol (172) and Berchemol 4'-*O*-β-D-glucopyranoside (173), respectively

The UV, ^{1}H and ^{13}C NMR (**Table 10**) of lignans **SM7-SM9** revealed structural similarities among them. The CIMS analysis of **SM7** showed the molecular ion peak at *m/z* 540 $[M+NH_4]^+$, and the HRCIMS showed *m/z* at 540.2437 $[M+NH_4]^+$ (calculated 540.2439 for $C_{26}H_{38}NO_{11}$). The ^{1}H NMR spectrum revealed that **SM7** had two (1,3,4)-trisubstituted benzene ring systems [δ_H 7.08 (d, *J*=8.2 Hz), 6.88 (d, *J*=1.8 Hz), 6.73 (dd, *J*=8.2, 1.8 Hz) and 6.90 (d, *J*=1.8 Hz), 6.75 (d, *J*=8.2 Hz), 6.74 (dd, *J*=8.2, 1.8 Hz)], one methylene [δ_H 2.98 (dd, *J*=13.5, 4.7 Hz) and 2.54 (dd, *J*=13.5, 11.3 Hz)], three methines [(δ_H 4.74 (d, *J*=6.9 Hz), 2.73 (m) and 2.37 (m)] and two oxygenated methylenes [δ_H 3.96 (dd, *J*=8.0, 5.2 Hz), 3.70 (dd, *J*=8.0, 6.6 Hz) and 3.82 (m), 3.64 (dd, *J*=11.6 Hz). All these data suggested that **SM7** possessed

a lariciresinol type skeleton.[158] The lariciresinol skeleton was further confirmed by ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC experiments (**Figure 31**). The ^{1}H-^{1}H COSY showed cross peaks between H-5 (δ_H 6.75) and H-6 (δ_H 6.74), H-5' (δ_H 7.08) and H-6' (δ_H 6.73), H-7 (δ_H 4.74) and H-8 (δ_H 2.37), H-8 and H-8' (δ_H 2.73), and H-8' and H-7' (δ_H 2.98) and H-9' (δ_H 3.96). The ^{3}J ^{1}H-^{13}C long range correlations were observed between H-2 (δ_H 6.90) and C-4 (δ_C 147.1), C-6 (δ_C 119.8) and C-7 (δ_C 84.1), H-7 (δ_H 4.74) and C-2 (δ_C 110.7), C-6 and C-9 (δ_C 60.5), H-9 (δ_H 3.82) and C-7, H-7' (δ_H 2.98) and C-2' (δ_C 114.4) and C-6' (δ_C 122.3), and H-9' (δ_H 3.96) and C-7' (δ_C 33.7) (**Figure 92**).



	R	R'
171 (SM7)	H	β-D-glucopyranosyl
172 (SM8)	OH	H

Figure 30. Structures of lariciresinol 4'-*O*-β-D-glucopyranoside (**SM7**) and berchemol (**SM8**)

The ^{1}H NMR signal at δ_H 4.84 and the ^{13}C NMR signal at δ_C 103.1 could be attributed to an anomeric proton and anomeric carbon, respectively, of a glucose moiety.[153] The ^{3}J ^{1}H-^{13}C long range correlation, observed in the HMBC spectrum (**Figure 92**) between δ_H 4.84 (H-1'') and δ_C 146.4 (C-4') confirmed the attachment of the glucose moiety at C-4' of the aglycone unit. Although the ^{1}H and ^{13}C NMR

data of **SM7** were identical to the published data of lariciresinol 4'-*O*-β-D-glucopyranoside[159], all ^{1}H and ^{13}C NMR signals were completely assigned on the basis of DEPT-135, COSY, HMQC and HMBC experiments. Thus, **SM7** was identified unequivocally as lariciresinol 4'-*O*-β-D-glucopyranoside (**171**).[159] This is the first report of the occurrence of this compound in *C. cyanus* and even in the genus *Centaurea*.

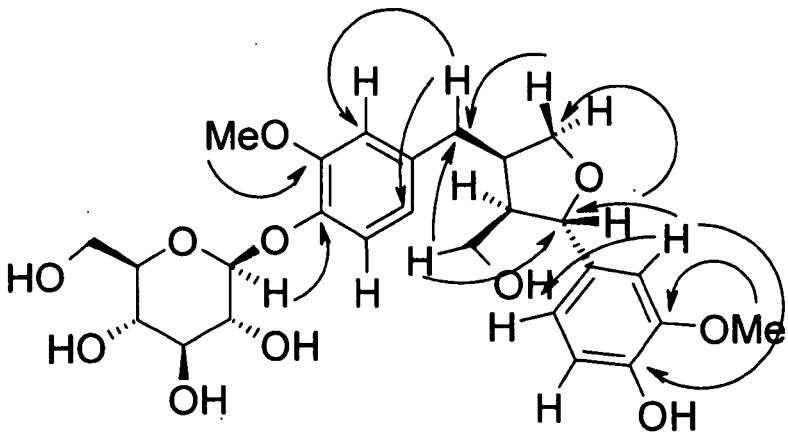


Figure 31. Key HMBC ($^{1}H \rightarrow ^{13}C$) correlations of **SM7** (**171**)

The CIMS spectrum of **SM8** showed an *m/z* at 394 $[M+NH_4]^+$, indicating the molecular formula $C_{20}H_{24}O_6$. The ^{1}H and ^{13}C NMR data (**Table 10**) of **SM8** showed structural similarities with **SM7**, but lacking a glucose moiety and one methine. The ^{13}C NMR data (**Table 10**) showed the presence of twenty carbons. The DEPT-135 indicated two methoxy (δ 56.1, 55.9), two methylenes (δ 72.1, 35.2), nine methines (δ 122.5, 121.7, 116.4, 115.7, 113.8, 112.8, 85.8, 83.4, 52.0), six quarternary carbons (δ 149.2, 148.8, 147.5, 146.0, 133.4, 130.9), and one primary alcohol group (δ 64.6). The ^{1}H NMR data (**Table 10**) showed the presence of two trisubstituted benzene ring systems [δ_H 6.92 (d, *J*=1.8 Hz), 6.72, 6.72 and 6.77 (d, *J*=1.8 Hz), 6.68 (d, *J*=8.0 Hz) and 6.61 (dd, *J*=8.0, 1.8 Hz)], two oxymethylenes [δ_H 4.03 (dd, *J*=8.0, 5.2 Hz) and

3.62 (dd, J=8.0, 6.6 Hz), and 3.76 (d, J=11.6 Hz) and 3.56 (d, J=11.6 Hz)], one oxymethine (δ_H 4.80, s), one methylenes [δ_H 3.05 (dd, J=12.8, 12.4 Hz) and 2.43 (t, J=12.8, 12.4 Hz)] and one methine [δ_H 2.54, m)] (**Figure 93**). The ^3J ^{1}H-^{13}C long-range HMBC correlations were observed between H-2 (δ_H 6.92) and C-4 (147.5), C-6 (δ_C 121.7) and C-7 (δ_C 85.8), H-6 (δ_C 6.72) and C-2 (δ_C 112.8) and C-7, H-7 (δ_H 4.80) and C-2, C-6, C-9 (δ_C 64.6) and C-9' (δ_C 72.1), H-7' (δ_H 3.05) and C-8 (δ_C 83.4) , C-2' (δ_C 113.8), C-6 and C-9', and H-9' (δ_H 4.03) and C-7, C-8 and C-7'. The relative stereochemistry at the chiral centres was determined from the ^{1}H-^{1}H NOESY experiment (**Figure 32**). The ^{1}H-^{1}H NOESY spectrum (**Figure 94**) showed strong NOE interactions between H-7 (δ_H 4.80) and H-2 (δ_H 6.92) and H-6 (δ_H 6.72), H-7' (δ_H 3.05) and H-2' (δ_H 6.77) and H-9 (δ_H 3.76), H-2 and 3-OCH_3 (δ_H 3.80), and H-2' and 3'-OCH_3 (δ_H 3.81). The relative stereochemistry of **SM8** was determined by direct comparison of its ^{1}H and ^{13}C NMR data with published data.[160] The ^{1}H and ^{13}C NMR data of **SM8** were in good agreements with published data of berchemol.[160] The UV, MS, ^{1}H NMR, ^{13}C NMR, HSQC and HMBC analyses, unequivocally confirmed **SM8** as berchemol (**172**). This is the first report on the occurrence of berchemol from any species of the genus *Centaurea*.

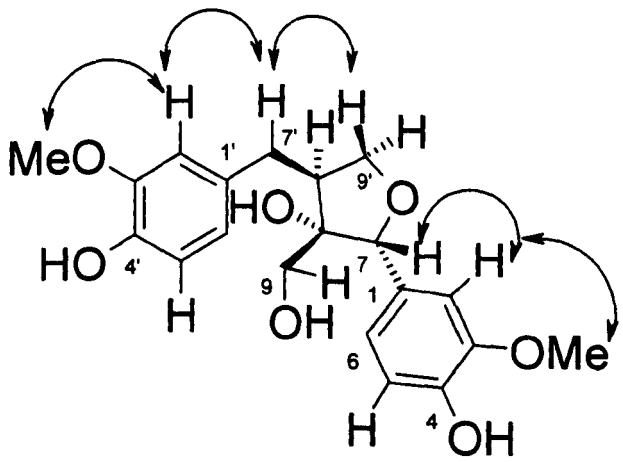


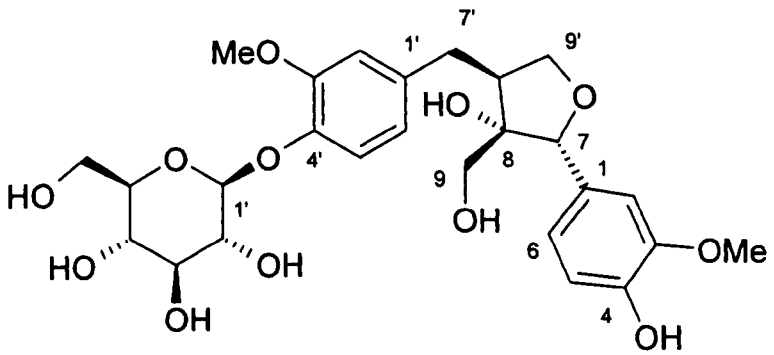
Figure 32. Key NOE interaction of **SM8** (**172**) based on ^{1}H-^{1}H NOESY experiment

Table 10. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data for **SM7 (171)** and **SM8 (172)**

Carbon number	Chemical shift δ in ppm			
	^{1}H		^{13}C	
	SM7[a]	**SM8**[b]	**SM7**[a]	**SM8**[b]
1	–	–	135.7	130.9
2	6.90, d, 1.8	6.92, d, 1.8	110.7	112.8
3	–	–	149.0	148.8
4	–	–	147.1	147.5
5	6.75, d, 8.2	6.72[c]	116.0	115.7
6	6.74, dd, 8.2, 1.8	6.72[c]	119.8	121.7
7	4.74, d, 6.9	4.80, s	84.1	85.8
8	2.37, m	–	54.0	83.4
9	3.82 3.64, dd, 11.6	3.76, d, 11.6 3.56, d, 11.6	60.5	64.6
1'	–	–	137.2	133.4
2'	6.88, d, 1.8	6.77, d, 1.8	114.4	113.8
3'	–	–	150.9	149.2
4'	–	–	146.4	146.0
5'	7.08, d, 8.2	6.68, d, 8.0	118.3	116.4
6'	6.73, dd, 8.2, 1.8	6.61, dd, 8.0, 1.8	122.3	122.5
7'	2.98, dd, 13.5, 4.7 2.54, dd, 13.5, 11.3	3.05, dd, 12.8, 12.4 2.43, t, 12.8, 12.4	33.7	35.2
8'	2.73, m	2.54, m	43.7	52.0
9'	3.96, dd, 8.0, 5.2 3.70, dd, 8.0, 6.6	4.03, dd, 8.0, 5.2 3.62, dd, 8.0, 6.6	73.5	72.1
1''	4.84[c]	–	103.1	–
2''	3.47[c]	–	75.0	–
3''	3.44[c]	–	77.9	–
4''	3.30[c]	–	71.4	–
5''	3.38[c]	–	78.2	–
6''	3.88[c] 3.68[c]	–	62.6	–
O-CH_3 (3)	3.83, s	3.80, s	56.4	55.9
O-CH_3 (3')	3.84, s	3.81, s	56.8	56.1

[a] ^{1}H NMR (600 MHz) and ^{13}C NMR (150 MHz) in CD_3OD
[b] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD
[c] Overlapped signals, assignments was confirmed from ^{1}H-^{1}H COSY and ^{1}H-^{13}C HSQC correlations

The ^{1}H and ^{13}C NMR spectrum (**Figure 95; Table 11**) of **SM9** displayed signals similar to those of berchemol (**172**) with the exception that there were additional signals which could be attributed to a glucose moiety in **SM9**. The ESIMS spectrum of **SM9** also confirmed the presence of a glucose moiety by displaying the pseudomolecular ion peak *m/z* at 561 $[M+Na]^{+}$ indicating the molecular formula $C_{26}H_{34}O_{12}$ [HRESIMS *m/z* 561.1934 $[M+Na]^{+}$]. The ^{1}H NMR signal at δ_H 4.81 (d, *J*=7.2 Hz, H-1'') could be assigned to the anomeric proton of a glucose moiety.[153] The ^{1}H-^{13}C ^{3}J long-range correlation between the anomeric proton (H-1'') and δ_C 145.2 (C-4') in the HMBC spectrum (**Figure 96; Table 11**) confirmed that the glucosyl unit was attached to C-4'. Thus, the structure of **SM9** was elucidated as berchemol 4'-*O*-β-D-glucopyranoside (**173**). The relative stereochemistry at the chiral centres in **SM9** was assigned by direct comparison of its ^{1}H and ^{13}C NMR data with published data.[161] This compound (**173**) was first isolated from *Valeriana officinalis*.[161] However, this is the first report on the occurrence of this compound in *C. montana* and even in the genus *Centaurea*.



173

Figure 33. Structure of berchemol 4'-*O*-β- D-glucopyranoside (**SM9**)

Table 11. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM9 (173)**

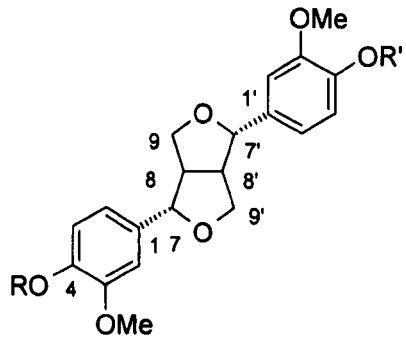
Carbon number	Chemical shift δ in ppm		HMBC correlations ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	–	129.5	–	–
2	6.92, s	111.5	C-3	C-4, C-6, C-7
3	–	147.4	–	–
4	–	146.0	–	–
5	6.72, s	114.3	–	C-1, C-3
6	6.72, s	120.3	C-1	C-2 , C-7
7	4.80, s	84.4	C-1	C-6, C-2, C-9
8	–	82.0	–	–
9	3.76, d, 11.6 3.56, d, 11.2	63.3	–	–
1'	–	135.6	–	–
2'	6.86, d, 2	113.1	C-3'	C- 4', C-6', C-7'
3'	–	149.7	–	–
4'	–	145.2	–	–
5'	7.06, d, 8.0	117.1	C-4'	C- 3', C- 1'
6'	6.74, dd, 8.0, 2.0	121.1	–	C-2', C-4'
7'	3.12, dd, 12.4, 2.4 2.52, dd, 12.4, 3.4	33.8	C-1'	C-2', C-6'
8'	2.58, m	50.5	–	
9'	4.05, dd, 8.0, 6.0 3.61, dd, 8.0, 3.2	70.6	–	C-7', C-7, C-8
1''	4.81, d, 7.2	101.8	–	C-4'
2''	3.44, m	73.7	–	–
3''	3.43, m	76.6	–	–
4''	3.34, m	70.1	–	–
5''	3.35, m	77.0	–	–
6''	3.82, m 3.64, m	61.3	–	–
O-CH_3 (3)	3.81, s	55.5	–	C-3
O-CH_3 (3')	3.82, s	55.1	–	C-3'

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.1.2.2 Characterisation of SM10, SM11 and SM12 as Pinoresinol (144), Pinoresinol monomethyl ether (174) and Pinoresinol dimethyl ether (175), respectively

The ^{13}C NMR and DEPT-135 spectra of **SM10** (**Figure 97 and 98**) displayed 10 carbon signals consisting of one methylene (δ 71.4), five methines (δ 118.8, 114.9, 109.8, 86.3, 54.1), two oxygenated quarternary carbons (δ 147.9, 146.1), one quarternary carbon (δ 132.8) and one methoxy groups (δ 55.2). All protonated carbons were assigned by ^{1}H-^{13}C HSQC experiments.[150] The ^{1}H NMR spectrum showed signals for a methine proton at δ 3.10 (m), a benzylic oxymethine proton at δ 4.66 (d, *J*=4.8 Hz), an oxymethylene at δ 4.19 (m) and 3.80 (m), and a 1,3,4 trisubstituted phenyl group [δ 6.90 (d, *J*=1.6 Hz), 6.77 (dd, *J*=8.0, 1.6 Hz) and 6.72 (d, *J*=8.0 Hz)], which could be assigned to a lignan of the 3,7-dioxobicyclo[3.3.0]octane type, containing a furofuran ring (**Table 12**).[162,163] The presence of a furofuran skeleton was also confirmed by the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC experiments (**Table 12**). The ^{1}H-^{1}H COSY spectrum showed cross peaks between H-5 (δ_H 6.72) and H-6 (δ_H 6.74), and H-6 (δ_H 6.74) and H-7 (δ_H 4.66). A ^{2}J ^{1}H-^{13}C correlation between H-7 (δ_H 4.66) to C-8 (δ_C 54.1) and a ^{3}J correlations from H-7 (δ_H 4.66) to C-9 (δ_C 71.4), and H-9 (δ_H 4.22) to C-7 (δ_C 86.3) and C-8 (δ_C 54.1) were observed in the ^{1}H-^{13}C HMBC spectrum (**Figure 99**), and thus confirmed the furofuran structure. A singlet at δ_H 3.81 in the ^{1}H NMR spectrum, together with the signal δ_C 55.2 in the ^{13}C NMR spectrum, could be assigned to a methoxy group attached to the aromatic ring. The ^{3}J ^{1}H-^{13}C long range correlation between δ_H 3.81 and δ_C 147.9 (C-3) in the HMBC spectrum confirmed the attachement of the methoxy group at C-3 position. The relative configuration at H-7 and H-8, and H-7' and H-8' were determined as *trans*-form from the coupling constant, *J*=4.48 Hz of

H-7 and H-7', and from the ^{1}H-^{1}H ROESY experiment (**Figure 35**).[164] The ^{1}H-^{1}H ROESY spectrum showed strong NOE interactions between H-2 (δ_H 6.90) and 3-OMe (δ_H 3.81), H-7 (δ_H 4.66) and H-2, H-6 (δ_H 6.74) and H-9 (δ_H 4.22). However, no NOE correlations were observed between H-7 and H-8 (δ_H 3.10). The molecular formula of **SM10** was determined as $C_{20}H_{22}O_6$ on the basis of the $[M-H]^+$ ion peak at *m/z* 357.1344 in the HRESIMS. Taking the ^{1}H and ^{13}C NMR data, and the molecular mass into account, it was clear that this molecule had two equal parts or symmetry. The ^{1}H and ^{13}C NMR data were identical to published data of pinoresinol.[165,166]



	R	R'
144 (SM10)	H	H
174 (SM11)	H	Me
175 (SM12)	Me	Me

Figure 34. Structures of pinoresinol (**SM10**), pinoresinol monomethyl ether (**SM11**) and pinoresinol dimethyl ether (**SM12**)

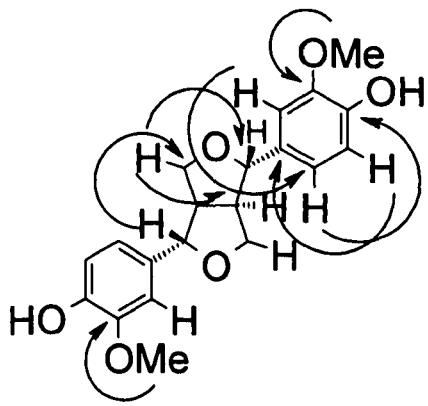


Figure 35. Key HMBC correlations ($^{1}H \rightarrow ^{13}C$) of **SM10** (**144**)

A positive optical rotation was observed for **SM10**. A combination of 1D and 2D NMR analyses confirmed the identity of **SM10** as (+)-pinoresinol (**144**). (+)-Pinoresinol was previously isolated from many other plant species[165-166], including *C. calcitrapa*[64] and *C. orphanidea*[106] (**Table 1**).

Table 12. ^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data, ^{1}H-^{1}H COSY and ^{1}H-^{13}C long range HMBC for **SM10 (144)**

Carbon number	Chemical shift δ in ppm			
	$^{1}H^{a}$	$^{13}C^{a}$	COSY (^{1}H-^{1}H)	HMBC correlation ($^{1}H \rightarrow ^{13}C$)
1/1'	–	132.6	–	–
2/2'	6.90, d, 1.6	109.8	H-2 to H-6	C-1, C-3,C-5, C-6, C-7
3/3'	–	147.9	–	–
4/4'	–	146.1	–	–
5/5'	6.72, d, 8	114.9	H-5 to H-6	C-1, C-3, C-4
6/6'	6.77, dd, 8, 1.6	118.8	H-6 to H-5	C-2, C-4, C-7
7/7'	4.66, d, 4.8	86.3	H-7 to H-8	C-2, C-6, C-8, C-9
8/8'	3.10, m	54.1	H-8 to H-7, H-9	–
9/9'	4.19, m 3.80, m	71.4	Ha-9 to Hb-9, H-8	C-7, C-8
O-CH_3 (3/3')	3.81, s	55.2	–	C-3/C-3'

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

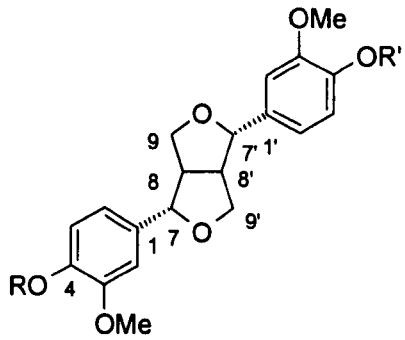
The ^{1}H and ^{13}C NMR spectra of **SM11 and SM12** indicated that they were the pinoresinol (**144**) derivatives. Their ^{1}H NMR spectra showed the presence of two 1,3,4, trisubstituted benzene rings [δ_H 7.02 (d, *J*=8.0 Hz), 6.94 (d, *J*=2.0 Hz), 6.84 (dd, *J*=8.0, 2.0 Hz), 6.91 (d, *J*=2.0 Hz), 6.77 (dd, *J*=8.0, 2.0 Hz) and 6.72 (d, *J*=8.0 Hz)] and furofuran rings [δ_H 4.71 (d, *J*=4.8 Hz), 4.65 (d, *J*=4.4 Hz), 4.21 (m), 3.83 (m) and 3.10 (m)].[162,163] The ^{1}H NMR spectrum of **SM11** and **SM12** also showed additional signals at δ_H 3.81, and 3.81 and 3.82, respectively, for the presence one, and two of methoxy groups. All spectroscopic data of **SM11** and **SM12** were compared with published data of pinoresinol monomethyl ether (**174**) and pinoresinol dimethyl ether (**175**) and found identical.[167-169] This is the first report on the occurrence of compounds **174** and **175** in the genus *Centaurea*.

3.1.2.3 Characterisation of SM13, SM14 and SM15 as Pinoresinol 4-*O*-β-D-glucopyranoside (176), Pinoresinol 4,4'-di-*O*-β-D-glucopyranoside (177) and Pinoresinol 4-*O*-β-D-apiofuranosyl-(1→2) glucopyranoside (178), respectively

The ^{1}H and ^{13}C NMR spectra (**Figure 101; Table 13-14**) of **SM13, SM14 and SM15** indicated that they were the pinoresinol (**144**) derivatives having one and two glucose units, and one apiose-glucose moiety, respectively. This fact was confirmed by mass analyses. The ESIMS spectrum of **SM13** revealed the pseudomolecular ion peak at *m/z* 543 $[M+Na]^{+}$, suggesting the molecular formula $C_{26}H_{32}O_{11}$. An additional 162 mass units compared to pinoresinol (**144**) indicated the presence of a glucose moiety in **SM13**. In the ^{1}H NMR spectrum, a doublet observed at δ 4.82 (*J*=7.6 Hz) was attributable to the anomeric proton of the glucopyranose with a β configuration.[153] A ^{3}J ^{1}H-^{13}C long range correlation observed in the HMBC spectrum (**Table 13**) from the anomeric proton at δ_H 4.82 (H-1'') to δ_C 149.8.1 (C-4) confirmed the attachment of the glucose unit at C-4 of the pinoresinol structure. A positive optical rotation was observed for the compound. Thus, the structure of **SM13** was confirmed as (+) pinoresinol 4-*O*-β-D-glucopyranoside (**176**). This compound was previously reported from *Forsythia* leaves and *Osmanthus asiaticus*.[170,171]

The ESIMS spectrum of **SM14** revealed the pseudomolecular ion peak at *m/z* 705 $[M+Na]^{+}$, suggesting the molecular formula $C_{32}H_{42}O_{16}$. It was clear that **SM14** had an extra 324 mass units compared to pinoresinol (**144**) and is due to the presence of two glucosyl moieties in **SM14**. However, due to the symmetry of the molecule both glucosyl units showed similar signals in the ^{1}H NMR spectrum (**Table 13**). A doublet at δ_H 4.84 (*J*=7.6 Hz) was attributable to the anomeric proton of a

glucopyranose with β configuration.[153] A ^{3}J ^{1}H-^{13}C long range correlation observed in the HMBC spectrum from the anomeric proton, δ_H 4.84 (H-1'') to δ_C 149.2 (C-4) confirmed that the glucosyl unit was linked C-4 of pinoresinol. Thus, **SM14** was confirmed as pinoresinol 4,4'-di-*O*-β-D-glucopyranoside (**177**). This compound was previously isolated from *Eucommia ulmoides*.[172]



	R	R'
176 (SM13)	β-D-glucopyranosyl	H
177 (SM14)	β-D-glucopyranosyl	β-D-glucopyranosyl
178 (SM15)	apiose-β-D-glucopyranosyl	H

Figure 36. Structures of pinoresinol glycosides (**SM13-SM15**)

Table 13. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz) for **SM13 (176), SM14 (177)** and **SM15 (178)**

Carbon number	Chemical shift δ in ppm		
	SM13[a]	**SM14**[a]	**SM15**[a]
2	6.99, d, 1.6	6.98, d, 2.0	6.96, d, 2.0
5	7.10, d, 8.4	7.02, d, 8.4	7.04, d, 8.4
6	6.86, dd, 8, 2	6.85, dd, 8, 2	6.84, dd, 8.0, 2.0
7	4.72, d, 4.4	4.71, d, 4.8	4.70, d, 4.8
8	3.10, m	3.10, m	3.1, m
9	4.20, m 3.88, m	4.20, m 3.84, m	4.2, m 3.82, m
2'	6.90, d, 2.0	6.98, d, 2.0	6.90, d, 2.0
5'	6.74, d, 8.0	7.02, d, 8.4	6.72, d, 8.4
6'	6.76, dd, 8.2	6.85, dd, 8, 2	6.77, dd, 8.4
7'	4.66, d, 4.4	4.71, d, 4.8	4.66, d, 4.4
8'	3.10, m	3.10, m	3.10, m
9'	4.20, m 3.78, m	4.20, m 3.84, m	4.20, m 3.82, m
1''	4.82, d, 7.6	4.84, d, 7.6	4.93, d, 7.6
2''	3.42, m	3.44, m	3.3, m
3''	3.15, m	3.20, m	3.93, m
4''	3.32, m	3.35, m	3.35, m
5''	3.32, m	3.32, m	3.61, m
6''	3.80, m 3.62, m	3.79, m 3.64, m	3.82, m 3.79, m
1'''	–	4.84, d, 7.6	5.50, d, 1.0
2'''	–	3.44, m	3.25, m
3'''	–	3.20, m	3.89, m
4'''	–	3.35, m	4.10, d, 10.0 3.70, d, 10.0
5'''	–	3.32, m	3.52, d, 8.0 3.48, d, 8.0
6'''	–	3.79, m 3.64, m	–
O-CH_3 (3)	3.82, s	3.82, s	3.81, s
O-CH_3 (3')	3.81, s	3.82, s	3.80, s

[a] ^{1}H NMR (400 MHz) in CD_3OD

Table 14. ^{13}C NMR data for **SM13 (176), SM14 (177)** and **SM15 (178)**

Carbon number	Chemical shift δ in ppm		
	SM13[a]	**SM14**[a]	**SM15**[a]
1	136.3	136.0	135.8
2	110.4	110.6	110.3
3	146.1	146.4	146.1
4	149.8	149.2	149.8
5	116.8	116.3	116.0
6	118.6	118.4	118.3
7	85.9	85.9	86.0
8	54.3	54.3	54.1
9	71.5	71.5	71.5
1'	132.6	136.0	132.6
2'	109.8	110.6	109.7
3'	146.2	146.4	146.3
4'	147.9	149.2	147.9
5'	114.9	116.3	114.9
6'	118.8	118.4	118.8
7'	86.3	85.9	86.3
8'	54.3	54.3	54.3
9'	71.5	71.5	71.5
1''	101.7	101.7	99.8
2''	73.7	73.7	76.8
3''	77.0	77.0	76.6
4''	70.1	70.1	70.2
5''	76.6	76.6	75.9
6''	61.3	61.3	61.3
1'''	–	101.7	109.0
2'''	–	73.7	77.6
3'''	–	77.0	79.6
4'''	–	70.1	74.3
5'''	–	76.6	65.0
6'''	–	61.3	
O-CH_3 (3)	55.5	55.2	55.2
O-CH_3 (3')	55.2	55.2	55.2

[a] ^{13}C NMR (100 MHz) in CD_3OD

The ESIMS spectrum of **SM15** showed the pseudomolecular ion peak at *m/z* 675 $[M+Na]^+$, and the molecular formula was determined as $C_{31}H_{40}O_{15}$ from HRESIMS (*m/z* 675.2263 $[M+Na]^+$, calculated value 675.2259) revealing that this compound had an extra 294 mass units compared to pinoresinol (**SM10**). This extra mass could be attributed to an apiose and a glucose moiety. In the ^1H NMR spectrum (**Table 14**), a doublet at δ 4.93 (*J*=7.6 Hz) was attributable to the anomeric proton of a glucopyranosyl unit with a β configuration.[153] In the ^1H-^{13}C HMBC spectrum (**Figure 102**) the glucose anomeric proton at δ_H 4.93 was correlated with the carbon signal at δ_C 149.8 (C-4) indicating that the glucose unit was attached to C-4 of pinoresinol (**SM10**). Additional signals in the ^1H NMR (δ 5.50, d, *J*=1.0 Hz and 4.10-3.52, m) and ^{13}C NMR (δ 109.0, 79.6, 77.6, 74.3, 65.0) could be assigned to an apiose unit.[173] The β orientation of apiose was considered by comparing the ^{13}C NMR data of **SM15** with those of α-D-(δ_C 104.5) and β-D-apiofuranoside (δ_C 111.5), respectively.[173,174] A ^3J ^1H-^{13}C HMBC correlation between anomeric proton of apiose (H-1''', δ_H=5.50) and δ_C 76.8 (C-2'') of glucose confirmed the linkage of apiose-(1→2)-glucose (**Figure 102**). Thus **SM15** was established as 4-*O*-β-D-apiofuranosyl-(1→2)-β-D-glucopyranosyl pinoresinol (**178**). This compound was first isolated from *Parsonsia laevigata*.[175] However, this is the first report on the occurrence of this compound and other pinoresinol glycosides in *C. montana* and even in the genus *Centaurea*.

3.1.2.4 Properties of Epoxylignans (SM7-SM15)

3.1.2.4.1 Properties of Lariciresinol 4'-*O*-β-D-glucopyranoside, 171 (SM7)

Gum, UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3439, 1595, 1510, 1460 and 1260 cm^{-1}; ESIMS *m/z* 540 $[M+NH_4]^+$; HRESIMS *m/z* 540.2437 $[M+NH_4]^+$ (calculated 540.2439 for $C_{26}H_{38}NO_{11}$); ^1H NMR (400 MHz, CD_3OD): **Table 10**; ^{13}C NMR (100 MHz, CD_3OD): **Table 10**.

3.1.2.4.2 Properties of Berchemol, 172 (SM8)

Gum, $[\alpha]^{23}_D$ + 37.8° (c 0.021, MeOH) (Lit.[160] – 7.9°, MeOH); UV λ_{max} (MeOH): 281, 228 nm; IR ν_{max} (neat): 3459, 1591, 1514, 1460 and 1266 cm^{-1}; CIMS *m/z* 394 $[M+NH_4]^+$; HRCIMS *m/z* 394.1857 $[M+NH_4]^+$ (calculated 394.1865 for $C_{20}H_{28}NO_7$); ^1H NMR (400 MHz, CD_3OD): **Table 10**; ^{13}C NMR (100 MHz, CD_3OD): **Table 10**.

3.1.2.4.3 Properties of Berchemol 4'-*O*-β-D-glucopyranoside, 173 (SM9)

Gum, $[\alpha]^{23}_D$ + 21.7° (c 0.015, MeOH) (Lit.[161] - 15.0°, DMSO); UV λ_{max} (MeOH): 279, 228 nm; IR ν_{max} (neat): 3459, 1591, 1514, 1460 and 1266 cm^{-1}; CIMS *m/z* 556 $[M+NH_4]^+$, ESIMS *m/z* 561 $[M+Na]^+$, 375 [M-glucose]$^+$, 309, 149, 121; HRESIMS *m/z* 561.1934 $[M+Na]^+$ (calculated for $C_{26}H_{34}O_{12}Na$); ^1H NMR (400 MHz, CD_3OD): **Table 11**; ^{13}C NMR (100 MHz, CD_3OD): **Table 11**.

3.1.2.4.4 Properties of Pinoresinol, 144 (SM10)

Gum, $[\alpha]^{23}_D$ + 80.4 (c 0.0033, MeOH) (Lit.[166] + 35.0°, $CHCl_3$); UV λ_{max} (MeOH): 280, 227 nm; IR ν_{max} (neat): 3455, 1596, 1518, 1464 and 1262 cm^{-1}; ESIMS *m/z* 381 $[M+Na]^+$, 341 $[M-OH]^+$, 235, 175, 137, 58; HRESIMS *m/z* 357.1339 $[M-H]^+$ (calculated 357.1344 for $C_{20}H_{21}O_6$); ^1H NMR (400 MHz, CD_3OD): **Table 12**; ^{13}C NMR (100 MHz, CD_3OD): **Table 12**.

3.1.2.4.5 Properties of Pinoresinol monomethyl ether, 174 (SM11)

Gum, UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3455, 1597, 1512, 1462 and 1267 cm^{-1}; ESIMS *m/z* 395 $[M+Na]^+$, 175, 151, 137, 58; Molecular formula $C_{21}H_{24}O_6$, ^1H NMR (250 MHz, CD_3OD):(δ_H 6.91 (d, 2.0 Hz), 6.77 (dd, 8.0, 2.0 Hz), 6.72 (d, 8.0 Hz), 4.71 (d, 4.8 Hz), 4.65 (d, 4.4 Hz), 4.20(m), 3.83 (m), 3.81 (s), 3.10 (m); ^{13}C NMR (62.5 MHz, CD_3OD): (δ_C 148.5, 147.9, 147.2, 146.1, 132.8, 118.8, 114.9, 109.9, 86.3, 71.4, 55.5, 55.5, 55.4, 54.1).

3.1.2.4.6 Properties of Pinoresinol dimethyl ether, 175 (SM12)

Gum, $[\alpha]^{23}_D$ +45.2° C (c 0.03, MeOH) (Lit.[168] + 46.0°, $CHCl_3$); UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3451, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 409 $[M+Na]^+$, 137; Molecular formula $C_{22}H_{26}O_6$, ^1H NMR (250 MHz, CD_3OD): (δ_H 7.02 (d, *J*=8.0 Hz), 6.94 (d, *J*=2.0 Hz), 6.94 (d, *J*=2.0 Hz), 6.91 (d, 2.0 Hz), 6.77 (dd, 8.0, 2.0 Hz), 6.72 (d, 8.0 Hz), 4.71 (d, 4.8 Hz), 4.65 (d, 4.4 Hz), 4.20 (m), 3.83 (m), 3.82, 3.81 (s), 3.10 (m); ^{13}C NMR (62.5 MHz, CD_3OD): (δ_C 147.9, 147.9, 146.2, 146.1, 132.8, 118.8, 114.9, 109.9, 86.3, 71.4, 55.5, 55.5, 55.5, 55.4, 55.4, 54.1).

3.1.2.4.7 Properties of Pinoresinol 4-*O*-β-D-glucopyranoside, 176 (SM13)

Gum, $[\alpha]^{23}_D$ +11.3 (c 0.044, MeOH) (Lit.[171] + 10.8°, MeOH); UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3459, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 520 $[M+Na]^+$, 367, 357 $[M\text{-glucose}]^+$, 234, 157; HRESIMS *m/z* 538.2282 $[M+NH_4]^+$ (calculated 538.2283 for $C_{26}H_{36}NO_{11}$); ^1H NMR (400 MHz, CD_3OD): **Table 13**; ^{13}C NMR (100 MHz, CD_3OD): **Table 14**.

3.1.2.4.8 Properties of Pinoresinol 4, 4' di-β-D-glucopyranoside, 177 (SM14)

Gum, UV λ_{max} (MeOH): 281, 228 nm; IR ν_{max} (neat): 3459, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 705 $[M+Na]^+$; 519 [M-monoglucose]$^+$, 357 [M-diglucose]$^+$, 151) HRESIMS *m/z* 700.2805 $[M+NH_4]^+$ (calculated 700.2811 for $C_{32}H_{46}NO_{16}$); ^1H NMR (400 MHz, CD_3OD): **Table 13**; ^{13}C NMR (100 MHz, CD_3OD): **Table 14**.

3.1.2.4.9 Properties of Pinoresinol 4-*O*-β-D-apiosefuranosyl-(1→2)-β-D-glucopyranoside, 178 (SM15)

Gum, $[\alpha]^{23}_D$ -47.1 (c 0.019, MeOH) (Lit.[175] – 47.5°, MeOH); UV λ_{max} (MeOH): 280, 222 nm; IR ν_{max} (neat): 3459, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 675 $[M+Na]^+$, 538, 312, 150; HRESIMS *m/z* 675.2263 $[M+Na]^+$ (calculated 675.2259 for $C_{31}H_{40}O_{15}Na$); ^1H NMR (400 MHz, CD_3OD): **Table 13**; ^{13}C NMR (100 MHz, CD_3OD): **Table 14**.

3.2 Sesquiterpene lactones

Three sesquiterpene lactones (**SM16-SM18**) were isolated from two *Centaurea* species in this study. Of them, two were germacranolides (**SM16** and **SM17**) and one eudesmane (**SM18**).

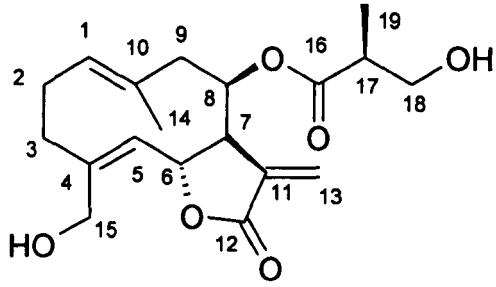
3.2.1 Germacranolides

Two germacranolides, **SM16** (isolation code **CG6**) and **SM17** (isolation code **CG7**) were isolated from *C. gigantea* (**Section 2.7.9**).

3.2.1.1 Characterisation of SM16 and SM17 as Arctiopicrin (21) and 8-*O*-(4-Hydroxy-3-methylbutanoyl)-salonitenolide (179), respectively

The ^{13}C NMR data (**Table 15**) revealed that **SM16** had 19 carbons. The DEPT-135 indicated the presence of six methylenes (δ 124.8, 63.6, 59.6, 48.4, 34.0, 25.7), six methines (δ 129.5, 128.5, 77.3, 72.8, 52.4, 42.7), two methyls (δ 15.7, 12.8), three quarternary carbons (δ 144.9, 135.9, 132.3) and two carbonyl carbons (δ 174.7, 171.0). All protonated carbons were assigned by ^{1}H-^{13}C HSQC experiments.[150] The ESIMS spectrum showed the pseudomolecular ion peak at *m/z* 373 $[M+Na]^+$, suggesting *Mr*=350 and the molecular formula $C_{19}H_{26}O_6$. The IR spectra of **SM16** exhibited bands indicative of hydroxyl (3410 cm^{-1}), lactone carbonyl (1770 cm^{-1}), cyclopentanone (1730 cm^{-1}) and olefinic carbons (1445 cm^{-1}).[97] The UV absorption at λ_{max} 218 nm indicated the presence of an αβ unsaturated carbonyl moiety.[176] All these spectroscopic information from UV, IR, and ^{13}C NMR data indicated that **SM16** was a germacranolide.[91] The ^{1}H NMR data (**Table 15**) exhibited signals at δ 6.20 (d, *J*=2.4 Hz) and 5.94 (d, *J*=2.4 Hz) for an exocyclic methylene group and at δ 5.11 (dd, *J*=10.0, 8.4 Hz) for a proton geminal to the oxygen of the lactone ring. In the ^{1}H NMR spectrum, additional doublet at δ_H 4.87 (*J*=10.0 Hz) and multiplets at δ_H

4.97, 4.93, 3.16, 2.39 and 2.34 suggested that the germacranolides are salonitenolide derivatives.[91]



21

Figure 37. Structure of arctiopicrin **(SM16)**

The salonitenolide moiety was also confirmed by the ^{1}H-^{13}C HMBC experiment which showed ^{3}J ^{1}H-^{13}C long-range correlations between H-2 (δ 2.24) and C-10 (δ 132.3), H-5 (δ 4.87) and C-3 (δ 34.0), H-6 (δ 5.11) and C-8 (δ 72.8), H-9 (δ 2.39) and C-7 (δ 52.4) and C-14 (δ 15.7), H-13 (δ 6.20) and C-7 (52.4) and C-12 (δ 171.0), and ^{2}J correlation between H-3 (δ 2.50) and C-4 (δ 144.9), and H-9 (δ 2.39) to C-8 (δ 72.7) and C-10 (δ 132.3) (**Table 15**). Additional signals at δ 3.70 (m), 3.60 (m), 2.58 (m) and 1.10 (d, J=7.2 Hz) in the ^{1}H NMR spectrum implied the presence of a 3-hydroxy-2-methylpropanoyl group which was further confirmed by the ^{1}H-^{13}C HMBC experiment. The HMBC spectrum (**Figure 104**) showed ^{2}J ^{1}H-^{13}C long-range correlations between H-17 (δ 2.58) and C-16 (δ 174.7), C-18 (δ 63.6) and C-19 (δ 12.8), and H-19 (δ 1.10) and C-17 (δ 42.7), and ^{3}J correlations between H-18 (δ 3.70) and C-16 (δ 174.7) and C-19 (δ 12.8), and H-19 (δ 1.10) and C-16 (174.7) and C-18 (δ 63.6). Thus, **SM16** was elucidated as 8-*O*-(3-hydroxy-2-methylpropanoyl)-salonitenolide or arctiopicrin (**21**). Although arctiopicrin was previously isolated from *C. melitensis*[91], this is the first report of the occurrence of this compound in *C. gigantea.*

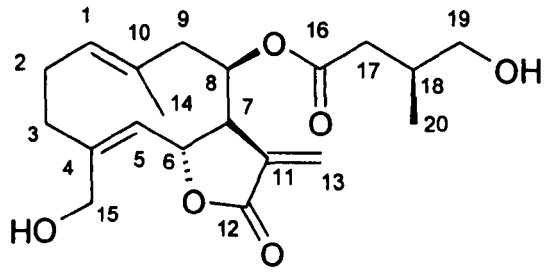
Table 15. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM16 (21)**

Carbon number	Chemical shift δ in ppm		HMBC correlation ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	4.99, m	129.5	–	–
2	2.24, d, m	25.7	–	C-10
	2.16, d, m			
3	2.50, m	34.0	C-4	–
	1.90, m			
4		144.9	–	–
5	4.87, d, 10	128.5	–	C-3
6	5.11, dd, 8.4, 10	77.3	–	C-8
7	3.16, m	52.4	–	–
8	4.97, m	72.8	–	–
9	2.39, m	48.4	C-8, C-10	C-7, C-14
	2.34, m			
10	–	132.3	–	–
11	–	135.9	–	–
12	–	171.0	–	–
13	6.20, d, 2.4	124.8	C-11	C-7, C-12
	5.94, d, 2.4			
14	1.49, s	15.7	C-10	C-9
15	4.20, d, 14	59.6	C-4	C-3, C-5
	3.97, d, 14			
16	–	174.7	–	–
17	2.58, m	42.7	C-16, C-18, C-19	
18	3.70, m	63.6	C-17	C-16, C-19
	3.60, m			
19	1.10, d, 7.2	12.8	C-17	C-16, C-18

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

The UV, IR, ^{1}H and ^{13}C NMR data (**Table 16**) revealed that **SM17** was a germacranolide like **SM16**. The ^{1}H NMR data (**Table 16**) suggested that germacranolide **SM17** was a salonitenolide derivative by displaying the signals at δ 6.06 (d, J=2.4 Hz), 5.74 (d, J=2.4 Hz), 5.13 (dd, J=10.0, 8.4 Hz), 4.97 (m), 4.93 (m), 4.77 (d, J=10.0 Hz) and 2.39 (m). The ^{13}C NMR spectrum (**Figure 106**) displayed 20 carbon signals instead of 19 carbons. The additional signal was due to the extra methylene on the side chain. The ^{1}H NMR spectrum also showed the presence of 4-hydroxy-3-methylbutanoyl group by showing the signals at δ 3.19, 2.35, 2.08, 1.90

and 0.81. This side chain was also confirmed by the ^{1}H-^{13}C HMBC experiment. The HMBC spectrum (**Figure 107**; **Table 16**) showed correlations between H-17 (δ 2.35) to C-19 (δ 65.9), C-20 (δ 17.0), H-18 (δ 1.90) to C-17 (δ 38.3), C-20 (δ 17.0), and H-20 (δ 0.81) to C-17 (δ 38.3) and C-19 (δ 65.9). The ESIMS spectrum showed the pseudomolecular ion at *m/z* 387 $[M+Na]^{+}$, suggesting *Mr*=364, and the molecular formula $C_{20}H_{28}O_6$. The mass analyses also proved that **SM17** had an extra 14 mass units compared to **SM16**. Thus, **SM17** was identified as 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**179**). This compound was previously isolated from *Onopordon laconicum*.[177] However, isolation of this compound from *C. gigantea* and even from the genus *Centaurea* is reported here for the first time.



179

Figure 38. Structure of 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**SM17**)

Table 16. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM17 (179)**

Carbon number	Chemical shift δ in ppm		HMBC correlation ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	4.93, m	129.8	–	–
2	2.11, d, 7.6	26.4	–	–
	2.06, d, 7.6			
3	2.45, m	38.3	–	C-15
	1.78, m			
4		144.9	–	–
5	4.77, d, 10.0	128.3	–	C-3, C-7, C-15
6	5.13, dd, 8.4, 10.0	76.9	–	C-8
7	3.16, m	52.0	–	–
8	4.97, m	72.7	–	–
9	2.39, m	48.7	C-8, C-10	C-7, C-14
	2.34, m			
10	–	132.7	–	–
11	–	136.9	–	–
12	–	170.2	–	–
13	6.06, d, 2.4	124.7	C-11	C-7, C-12
	5.74, d, 2.4			
14	1.36, s	17.1	C-10	C-9
15	3.9, d, 9.6	59.7	C-4	–
	3.8, d, 9.6			
16		174.1		
17	2.35, m	38.3	C-16, C-18	C-19, C-20
	2.08, m			
18	1.90, m	33.4	C-17, C-19, C-20	–
19	3.19, m	65.9	–	–
20	0.81, d, 6.4	17.0	C-18	C-17, C-19

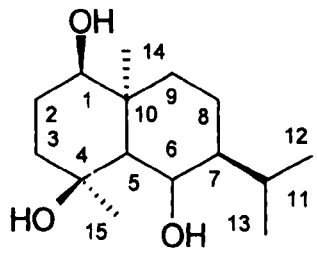
[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in DMSO-d_6

3.2.2 Eudesmane

One eudesmane-type sesquiterpene lactone, **SM18** (isolation code **CP4**), was isolated from *C. pamphylica* (**Section 2.7.10**).

3.2.2.1 Characterisation of SM18 as Pterodontriol (180)

The EIMS spectrum of **SM18** showed the molecular ion peak at *m/z* 256, and the molecular formula was calculated as $C_{15}H_{28}O_3$. The ^{13}C NMR spectrum (**Section 3.2.3.3**) of **SM18** indicated the presence of fifteen carbons. The ^{1}H NMR spectrum (**Section 3.2.3.3**) showed the presence of four methyl groups. Of them two were doublets at δ_H 1.34 and 0.96 (J=6.6 Hz) and the other two singlets at δ_H 1.22 and 1.60. A signal at δ_C 28.8 in the ^{13}C NMR spectrum was attributable to methine C-11 which implied the presence of an isopropyl group. Thus, the ^{1}H and ^{13}C NMR data of **SM18** suggested the presence of an eudesmane skeleton.[178] The ^{1}H and ^{13}C NMR data were also comparable to those published pterodontriol (**180**), previously isolated from *Laggera pterodonta* and found to be identical.[179] The occurrence of pterodontriol in *C. pamphylica* and even in the genus *Centaurea* is reported here for the first time.



180

Figure 39. Structure of pterodontriol (**SM18**)

3.2.3 Properties of Sesquiterpene lactones (SM16-SM18)

3.2.3.1 Properties of 8-*O*-(3-Hydroxy-2-methylpropanoyl)-salonitenolide or Arctiopicrin, 21 (SM16)

Gum, UV λ_{max} (MeOH): 218 nm; IR ν_{max} (neat): 3410, 1770, 1730 and 1445 cm^{-1}; ESIMS 373 $[M+Na]^+$; Molecular formula $C_{19}H_{26}O_6$; ^1H NMR (400 MHz, CD_3OD): **Table 15**; ^{13}C NMR (100 MHz, CD_3OD): **Table 15**.

3.2.3.2 Properties of 8-*O*-(4-Hydroxy-3-methylbutanoyl)-salonitenolide, 179 (SM17)

Gum, UV λ_{max} (MeOH): 218 nm; IR ν_{max} (neat): 3410, 1770, 1730 and 1445 cm^{-1}; ESIMS *m/z* 387 $[M+Na]^+$; Molecular formula $C_{20}H_{28}O_6$; ^1H NMR (400 MHz, CD_3OD): **Table 16**; ^{13}C NMR (100 MHz, CD_3OD): **Table 16**.

3.2.3.3 Properties of Pterodontriol, 180 (SM18)

MP 180° C, EIMS 256 $[M]^+$; Molecular formula $C_{15}H_{28}O_3$; ^1H NMR (250 MHz, CD_3OD): 4.22 (m), 3.67 (t, *J*=6.2 Hz), 2.25 (m), 2.14 (d, *J*=11.0 Hz), 1.98 (m), 1.96 (m), 1.88 (m), 1.57 (m), 1.34 (d), 1.04 (m), 1.60 (s), 1.22 (s), 0.96 (d) ; ^{13}C NMR (62.5 MHz, CD_3OD): 76.6 (C-1), 29.5 (C-2), 32.0 (C-3), 69.0 (C-4), 47.9 (C-5), 73.1 (C-6), 51.2 (C-7), 24.7 (C-8), 34.0 (C-9), 34.0 (C-10), 28.8 (C-11), 24.8 (C-12), 29.1 (C-13), 14.1 (C-14) and 22.7 (C-15).

3.3 Quinic acid Derivatives

A total of three quinic acid derivatives (**SM19-SM21**) were isolated in this study. Of them, **SM19** (isolation code **CM1**) and **SM20** (isolation code **CM2**) were purified from *C. montana* (**Section 2.7.6**). *C. gigantea* afforded **SM21** (isolation code **CG1**) (**Section 2.7.9**).

3.3.1 Characterisation of SM19 and SM20 as *trans*-3-*O*-*p*-Coumaroylquinic acid (181) and *cis*-3-*O*-*p*-Coumaroylquinic acid (182), respectively

The ^{13}C NMR and DEPT-135 spectra of both **SM19** and **SM20** displayed sixteen carbon signals, consisting of two methylenes (δ 39.2, 37.8), nine methines (δ 145.2, 129.8, 129.8, 115.6, 115.6, 114.6, 73.9, 71.9, 71.5), three quarternary carbons (δ 160.0, 126.1, 67.8), and two carbonyl carbons (δ 170.1, 167.9). All protonated carbons were assigned by ^{1}H-^{13}C HSQC.[150] The ESIMS spectra of **SM19** and **SM20** revealed the similar pseudomolecular ion peaks at *m/z* 361 $[M+Na]^+$, suggesting the molecular formula as $C_{16}H_{18}O_8$. Although the ESIMS and the ^{13}C NMR spectra (**Table 17**) indicated that both **SM19** and **SM20** were similar compounds, the ^{1}H NMR spectrum (**Figure 108**) revealed that they were geometrical isomers or *cis*-*trans* isomers. The ^{1}H NMR data (**Table 17**) confirmed the presence of olefinic protons by displaying the signals at δ 7.59 (d, *J*=15.5 Hz) and 6.31 (d, *J*=15.6 Hz) for **SM19** and δ 6.79 (d, *J*=12.8 Hz) and 5.77 (d, *J*=12.8 Hz) for **SM20**. The coupling constant, *J*=15.6 Hz confirmed that the olefinic protons in **SM19** were in *trans* configuration and **SM20** was its *cis* isomer.[176]

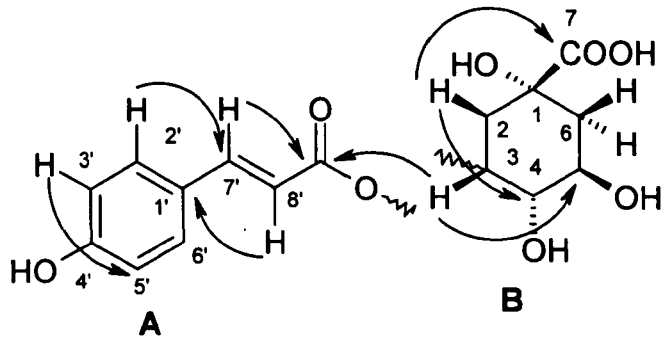


Figure 40. Key HMBC correlations ($^{1}H \rightarrow ^{13}C$) of substructure **A** and **B** of **SM19 (181)**

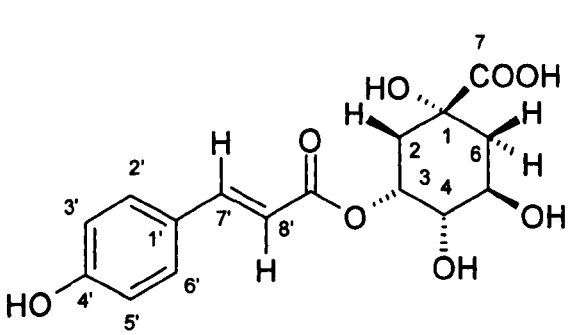
The use of 1D and 2D NMR data of **SM19** enabled the construction of two substructures, **A** and **B** (**Figure 40**). The ^{1}H NMR spectrum (**Figure 108**; **Table 17**) revealed that substructure **A** was a *p*-coumaroyl moiety, which was evident from the signals δ 7.41 (d, *J*=8.4 Hz) and 6.76 (d, *J*=8.4 Hz), and δ 7.59 (d, *J*=15.6 Hz) and 6.31 (d, *J*=15.6 Hz).

Table 17. ^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM19 (181)**

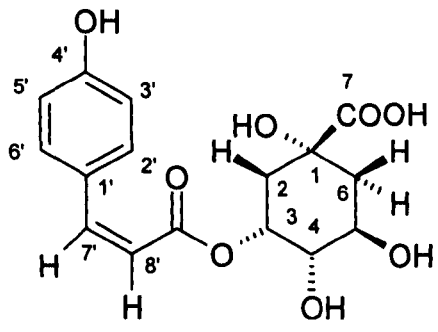
Carbon number	Chemical shift in ppm		HMBC correlation ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
1	–	67.0	–	–
2	2.13, dd, 14.8, 3.2 1.92, m	39.3	–	C-4
3	5.30, m	71.5	C-4	C-9'
4	3.63, dd, 9.6, 3.2	73.9	C-5	–
5	4.05, m	71.9	C-4	–
6	2.04, m 1.98, m	37.8	C-5	C-2, C-4
7	–	170.1	–	–
1'	–	126.1	–	–
2'	7.41, d, 8.4	129.8	–	C-4', C-6', C-7'
3'	6.76, d, 8.4	115.6	C-4'	C-1', C-5'
4'		160.0		
5'	6.76, d, 8.4	115.6	C-6'	C-1', C-3'
6'	7.41, d, 8.4	129.8	–	C-2', C-4', C-7'
7'	7.59, d, 15.6	145.2	C-8'	C-2', C-9'
8'	6.31, d, 15.6	114.6	C-9'	C-1'
9'	–	167.9	–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

The ^{1}H-^{1}H COSY spectrum also showed cross peaks between H-2'/4' to H-3'/5' and H-7' to H-8' to prove this substructure. The ^{1}H-^{13}C HMBC correlation between H-7' to C-9', H-8' to C-1', C-9', and H-2' to C-7' further confirmed the presence of a *p*-coumaroyl moiety (**Table 17**). The signals at δ 5.30 (m), 4.05 (d, m), 3.63 (dd, *J*=9.6, 3.2 Hz), 2.13 (dd, *J*=14.8, 3.2 Hz) and 2.04 (m) were observed in the ^{1}H NMR spectrum due to the presence of a quinic acid unit which was further confirmed by ^{1}H-^{1}H COSY experiment. The ^{1}H-^{1}H COSY spectrum showed cross peaks between H-2 to H-3, H-3 to H-4, H-4 to H-5 and H-5 to H-6. Finally, the ^{1}H-^{13}C HMBC correlations between H-2 to C-4, H-3 to C-2, C-4, H-4 to C-5, H-5 to C-4, and H-6 to C-2, C-4 and C-5 confirmed the presence of a quinic acid unit.[180] A ^{3}J long-range ^{1}H-^{13}C correlation between H-3 (δ_H 5.30) and δ_C 167.9 (C-9') attached *p*-coumaroyl moiety to quinic acid at C-3 (**Figure 109**). Thus, **SM19** was elucidated as *trans*-3-*O*-*p*-coumaroylquinic acid (**181**) and **SM20** was its *cis* isomer, *cis*-3-*O*-*p*-coumaroylquinic acid (**182**). This is the first report of the occurrence of these quinic acid derivatives in the *C. gigantea* and even in the genus *Centaurea*.



181

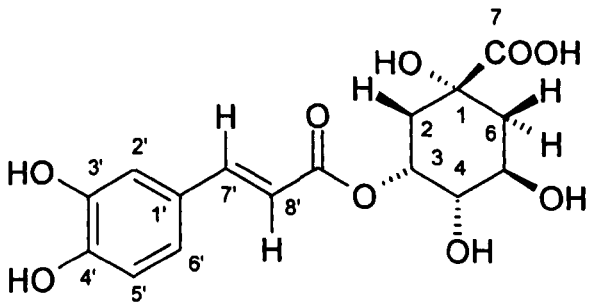


182

Figure 41. Structures of *trans* and *cis* quinic acid derivatives (**SM19** and **SM20**)

3.3.2 Characterisation of SM21 as Chlorogenic acid (168)

The ESIMS spectrum of **SM21** revealed the pseudomolecular ion peak at *m/z* 353 $[M-H]^+$, suggesting *Mr*=354 and the molecular formula $C_{16}H_{18}O_9$. The molecular mass indicated that it had an extra 16 higher mass units than compared to **SM20** which could be due to the presence of an additional hydroxyl group. The ^1H and ^{13}C NMR data (**Table 18**) also revealed that **SM21** had two substructures like **SM20** (**Figure 40**). But instead of a *p*-coumaroyl moiety (**SM20**), one of them was a *p*-caffeoyl moiety. The signals at δ 7.38 (d, *J*=15.6 Hz), 7.00 (d, *J*=2.0 Hz), 6.91 (dd, *J*=8.0, 2.0 Hz), 6.70 (d, *J*=8.0 Hz), 6.16 (d, *J*=15.6 Hz) in the ^1H NMR spectrum (**Table 18**) were due to the presence of a *p*-caffeoyl moiety which was further confirmed by the ^1H-^1H COSY and ^1H-^{13}C HMBC experiments. The ^1H NMR (δ 5.08, 4.11, 3.68, 2.43, 2.02) and ^{13}C NMR (δ 176.8, 74.0, 73.3, 71.9, 67.2, 39.4 and 38.3) signals implied that the other part of the molecule was quinic acid as in **SM20**. The quinic acid unit was also confirmed by the ^1H-^1H COSY and long range ^1H-^{13}C HMBC experiments. A ^3J long-range HMBC correlation between δ_H 5.08 (H-3) to δ_C 166.9 (C-9') indicated that the *p*-caffeoyl moiety was attached to quinic acid at C-3. Thus, the structure of **SM21** was elucidated as *trans* 3-*O*-caffeoyl quinic acid (**168**) and it is commercially available as chlorogenic acid (**168**).[181]



168

Figure 42. Structure of chlorogenic acid (**SM21**)

Chlorogenic acid, known as a defence compound against microorganism, is common in Compositae family[182] and was previously isolated from *C. cyanus*[72], *C. isaurica*[83].

Table 18. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR and long-range HMBC for **SM21 (168)**

Carbon number	Chemical shift δ in ppm		HMBC correlation ($^{1}H \rightarrow ^{13}C$)	
	^{1}H	^{13}C	^{2}J	^{3}J
1	–	69.2		–
2	2.43, dd, 14.8, 3.2	39.2	C-3	–
	1.92, m			
3	5.08, m	73.3	C-4	C-9'
4	3.68, dd, 9.6, 3.2	74.0	C-5	
5	4.11, m	71.9	C-4	
6	2.02, m	38.3	C-5	C-2, C-4
	1.98, m			
7	–	176.8	–	–
1'	–	126.2	–	–
2'	7.00, d, 2.0	115.4		C-4', C-6', C-7'
3'	–	146.3	–	–
4'	–	149.1	–	–
5'	6.70, d, 8.0	116.5	C-4'	C-1', C-3'
6'	6.91, dd, 8.0, 2.0	121.9	C-5'	C-4'
7'	7.38, d, 15.6	145.3	C-1', C-8'	C-6', C-9'
8'	6.16, d, 15.6	115.2	C-9'	C-1'
9'	–	166.9	–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.3.3 Properties of Quinic acid Derivatives (SM19-SM21)

3.3.3.1 Properties of *trans*- 3-*O*-*p*-Coumaroylquinic acid, 181 (SM19)

Gum, UV λ_{max} (MeOH): 332, 220 nm; IR ν_{max} (neat): 3459, 1765, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 361 $[M+Na]^+$, 191, 163, 104, 60; HRESIMS 339.1079 $[M+H]^+$ (calculated 339.1074 for $C_{16}H_{19}O_8$); ^1H NMR (400 MHz, CD_3OD): **Table 17**; ^{13}C NMR (100 MHz, CD_3OD): **Table 17.**

3.3.3.2 *cis*-3-*O*-*p*-Coumaroylquinic acid, 182 (SM20)

Gum, UV λ_{max} (MeOH): 279, 225 nm; IR ν_{max} (neat): 3459, 1765, 1591, 1514, 1460 and 1266 cm^{-1}; ESIMS *m/z* 361 $[M+Na]^+$, 191, 163, 104, 60; HRESIMS 339.1071 $[M+H]^+$ (calculated 339.1074 for $C_{16}H_{19}O_8$); ^1H NMR (400 MHz, CD_3OD) [δ 7.57, (d, *J*=8.4 Hz), 6.68 (d, *J*=8.4 Hz), 6.79 (d, *J*=12.8 Hz), 5.77 (d, *J*=12.8 Hz), 5.38 (m), 4.05 (d, *J*=3.2 Hz), 3.56 (dd, *J*=9.6, 3.2 Hz), 2.13 (dd, *J*=14.8, 3.2 Hz), 2.04 (m), 1.98 (m), 1.92 (m)]; ^{13}C NMR (100 MHz, CD_3OD): **Table 17.**

3.3.3.3 Properties of *trans*-3-*O*-Caffeoylquinic acid, 168 (SM21)

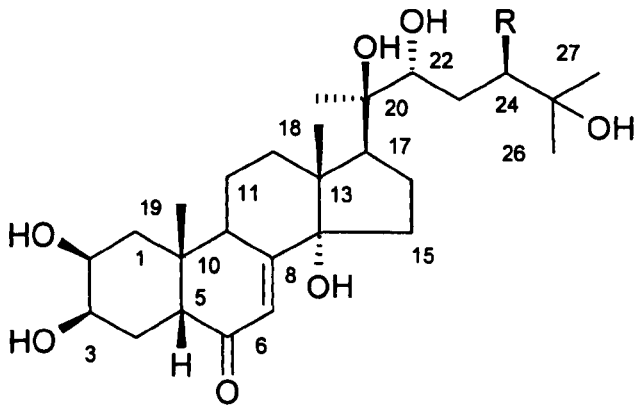
Gum, $[\alpha]^{23}_D$ -34.3° (c 0.004, MeOH) (Lit.[180] -38.0°, MeOH); UV λ_{max} (MeOH): 328, 218 nm; IR ν_{max} (neat): cm^{-1}; ESIMS *m/z* 353 $[M-H]^+$, 191, 85, 67; Molecular formula $C_{16}H_{18}O_9$; ^1H NMR (400 MHz, CD_3OD): **Table 18**; ^{13}C NMR (100 MHz, CD_3OD): **Table 18.**

3.4 Ecdysteroids

Two ecdysteroids, **SM22** (isolation code **CA2**) and **SM23** (isolation code **CA4**) were isolated from the *C. americana* seeds (**Section 2.7.4**).

3.4.1 Characterisation of SM22 and SM23 as 20-Hydroxyecdysone (169) and Makisterone A (183), respectively

The ^{13}C NMR spectrum (**Table 19**) revealed that **SM22** had twenty-seven carbons. The DEPT-135 indicated eight methylenes (δ 42.5, 38.9, 32.6, 32.6, 32.5, 27.4, 21.7, 21.3), seven methines (δ 122.2, 78.5, 68.9, 68.6, 51.9, 50.0, 33.0), six quarternary carbons (δ 166.1, 85.4, 78.1, 71.5, 48.3, 39.4), five methyl (δ 29.9, 29.0, 24.6, 21.6, 18.2) and a ketonic carbonyl (δ 203.5). The IR spectrum of the compound showed absorption bands of a hydroxyl group (3430 cm^{-1}) and a carbonyl group of an αβ-unsaturated ketone (1660 cm^{-1}). The ESIMS spectrum of **SM22** revealed the pseudomolecular ion peak at *m/z* 503 $[M+Na]^+$, and suggested the molecular formula $C_{27}H_{44}O_7$.



	R
169 (SM22)	H
183 (SM23)	Me

Figure 43. Structure of 20-hydroxyecdysone (**SM22**) and makisterone A (**SM23**)

Table 19. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz) and ^{13}C NMR data for **SM22 (169)** and **SM23 (183)**

Carbon number	Chemical shift δ in ppm			
	$^{1}H^{a}$		$^{13}C^{a}$	
	SM22	**SM23**	**SM22**	**SM23**
1	1.50, t, 12.6 1.85, m	1.50, t, 12.6, 1.85, m	38.9	36.7
2	3.79, m	3.79, m	68.9	63.4
3	3.68, d, 11.0	3.68, d, 11.0	68.6	63.4
4	1.75, m 1.90, m	1.72, m 1.86, m	32.5	33.7
5	2.33, dd, 11.0, 6.5	2.34, dd, 11.0, 6.5	51.9	55.4
6	–	–	203.5	202.6
7	5.74, d, 2.5	5.72, d, 2.0	122.2	121.0
8	–	–	166.1	175.9
9	3.12, m	3.62, m	33.0	34.9
10	–	–	39.4	39.0
11	1.75, m 1.85, m	1.75, m 1.85, m	21.3	22.1
12	1.95, m 1.75, m	1.95, m 1.75, m	32.6	33.7
13	–	–	48.3	44.5
14	–	–	85.4	83.7
15	2.00, m 1.54, m	2.00, m 1.54, m	32.6	34.2
16	1.85, m 1.80, m	1.85, m 1.80, m	21.7	26.6
17	2.33, m	2.90, m	50.0	48.0
18	0.91, s	0.91, s	18.2	18.8
19	0.94, s	0.93, s	24.6	22.0
20	–	–	78.1	78.1
21	1.19, s	1.19, s	21.6	22.0
22	3.66, m	3.72, m	78.5	77.9
23	1.33, m 1.65, m	1.33, m 1.65, m	27.4	34.8
24	1.52, m 1.80, m	2.0, m	42.5	22.0
25	–	–	71.5	72.3
26	1.19, s	1.19, s	29.9	18.8
27	1.20, s	1.19, s	29.0	17.9
24'	–	1.05, s	–	13.0

[a] ^{1}H NMR (250 MHz) and ^{13}C NMR (62.5 MHz) in CD_3OD

From IR, mass, ^{1}H and ^{13}C NMR spectra (**Table 19**), and by comparison of these data with published data, **SM22** was identified as 20-hydroxyecdysone (**169**).[183-185] This compound was previously isolated from *C. americana*.[51]

The IR, ^{1}H and ^{13}C NMR data of **SM23** were similar to those of **SM22** except the signal for an additional methyl group present in **SM23**. The ESIMS spectra supported the presence of additional methyl group by showing the $[M+Na]^{+}$ ion *m/z* at 517 corresponding to the molecular formula $C_{28}H_{46}O_7$. The ^{1}H and ^{13}C NMR spectra (**Table 19**) were compared with the published data of makisterone A (**183**) and found to be identical.[185,186] Although makisterone A was previously reported from *Penstemon venustus* and *Leuzea carthamoides*,[185,187] the occurrence of this compound in the genus *Centaurea* is reported here for the first time.

3.4.2 Properties of Ecdysteroids (SM22-SM23)

3.4.2.1 Properties of 20-Hydroxyecdysone, 169 (SM22)

Gum; UV λ_{max} (MeOH): 328, 218 nm; IR ν_{max} (neat): 3430, 1660 cm^{-1}; ESIMS *m/z* 503 $[M+Na]^{+}$, 479 $[M-H]^{+}$, 463 $[M+1-H_2O]^{+}$, 445 $[M+1-2H_2O]^{+}$, 427 $[M+1-3H_2O]^{+}$, 409 $[M+1-4H_2O]^{+}$, 383, 353, 177, 137; Molecular formula $C_{27}H_{44}O_7$; ^{1}H NMR (250 MHz, CD_3OD): **Table 19**; ^{13}C NMR (62.5 MHz, CD_3OD): **Table 19.**

3.4.2.2 Properties of Makisterone A or 20-Hydroxyecdysone-24 methyl, 183 (SM23)

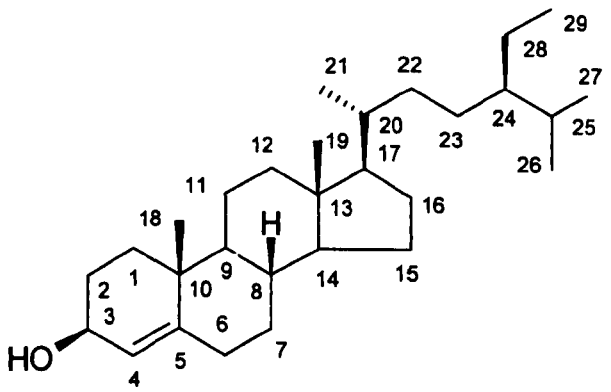
Gum; UV λ_{max} (MeOH): 328, 218 nm; IR ν_{max} (neat): 3430, 1660 cm^{-1}; ESIMS *m/z* 517 $[M+Na]^{+}$, 191, 85, 67; Molecular formula $C_{28}H_{46}O_7$; ^{1}H NMR (250 MHz, CD_3OD): **Table 19**; ^{13}C NMR (62.5 MHz, CD_3OD): **Table 19**.

3.5 Steroids

Three steroids (**SM24-SM26**) were isolated in this study. **SM24** (isolation code **CB7**) was purified from *C. bornmuelleri* (**Section 2.7.8**), **SM25** (isolation code **CC5**) from *C. cyanus* (**Section 2.7.2**) and **SM26** (isolation code **CMu1**) from *C. mucronifera* (**Section 2.7.12**).

3.5.1 Characterisation of SM24 as Stigmast-4-en-3-ol (184)

The EIMS spectrum of **SM24** revealed the molecular ion peak at *m/z* 414, corresponding to the molecular formula of $C_{29}H_{50}O$. The ^1H NMR spectrum (**Section 3.5.2.1**) showed the presence of six methyl groups (resonating between at δ_H 0.72-1.20), one olefinic proton (δ_H 5.35), several methine and methylenes (δ_H 0.78-2.34) and one methine (δ_H 3.51). The ^{13}C NMR spectrum (**Section 3.5.4.1**) displayed 29 carbon signals and the observed signal at δ_C 73.1 was due to the presence of 3-β hydroxyl group. On the basis of mass, ^1H and ^{13}C NMR data and their direct comparison with published data, **SM24** was identified as stigmast-4-en-3-ol (**184**).[188]

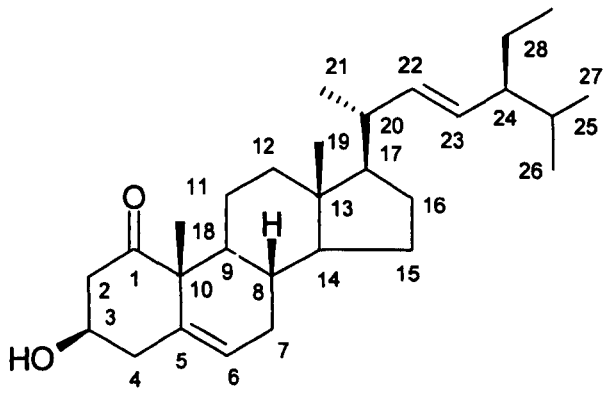


184

Figure 44. Structure of stigmast-4-en-3-ol (**SM24**)

3.5.2 Characterisation of SM25 as 3-Hydroxystigmast-5, 22-dien-1-one (185)

The EIMS spectrum of **SM25** revealed the molecular ion peak *m/z* at 426, suggesting for the molecular formula of $C_{29}H_{46}O_2$. The ^1H NMR spectrum (**Section 3.5.4.2**) showed the presence of six methyl groups (δ 0.72-1.20), one olefinic proton (δ 5.35), methine and methylene protons (δ 0.78-2.34) and one methine (δ 3.51). The ^{13}C NMR spectrum (**Section 3.5.4.2**) displayed 29 carbon signals and the signal at δ_C 201.1 was due to the presence of a carbonyl group. On the basis of mass, ^1H and ^{13}C NMR data and their direct comparison with published data, **SM25** was identified as 3-hydroxystigmast-5,22-dien-1-one (**185**).[189] This is the first report of the occurrence of this compound in the genus *Centaurea.*



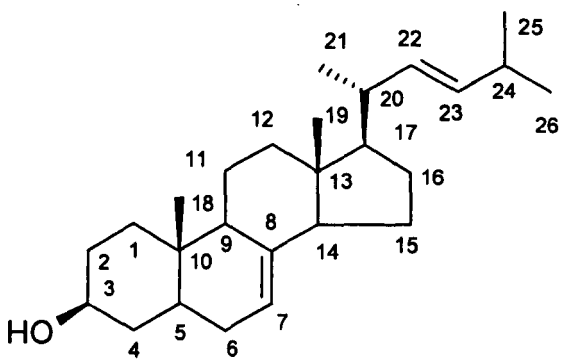
185

Figure 45. Structure of 3-hydroxystigmast-5,22-dien-1-one **(SM25)**

3.5.3 Characterisation of SM26 as Asterosterol (186)

The ^1H NMR data (**Section 3.5.4.3**) of **SM26** showed signals for a septet (δ 3.90, m), two methyl singlets (δ_H 1.21 and 0.86), two methyl doublets (δ_H 1.28 and 1.21) and two olefinic protons (δ_H 5.18 and 5.20) which implied **SM26** to be a steroid.[190] A molecular formula of $C_{26}H_{42}O$ was calculated from the EIMS spectrum which gave a

molecular ion peak at *m/z* 370. On the basis of mass, ^{1}H NMR data and their direct comparison with published data, **SM26** was identified as asterosterol (**186**).[190]



186

Figure 46. Structure of asterosterol **(SM26)**

3.5.4 Properties of Steroids (SM24-SM26)

3.5.4.1 Properties of Stigmast-4-en-3-ol, 184 (SM24)

Gum; EIMS *m/z* 414; Molecular formula $C_{29}H_{50}O$; ^{1}H NMR (400 MHz, CD_3OD): δ 5.35, 3.51, 2.34, 1.20, 0.78; ^{13}C NMR (100 MHz, CD_3OD): δ 140.2, 122.5, 78.3, 63.4, 39.0, 38.3, 38.4, 38.1, 37.0, 36.1, 35.6, 34.2, 33.5, 32.8, 32.2, 30.2, 30.0, 29.8, 29.6, 29.5, 29.4, 29.1, 28.0, 26.7, 25.7, 237, 18.2, 15.4, 14.5.

3.5.4.2 Properties of 3-Hydroxystigmast-5,22-dien-1-one, 185 (SM25)

Gum; EIMS *m/z* 426; Molecular formula $C_{29}H_{46}O_2$; ^{1}H NMR (400 MHz, CD_3OD): δ 5.35, 5.20, 5.17, 3.51, 2.34, 1.20, 0.78; ^{13}C NMR (100 MHz, CD_3OD): δ 203.0, 140.4, 122.1, 121.2, 114.5, 78.0, 63.1, 39.4, 38.9, 38.8, 38.0, 37.1, 35.6, 34.2, 32.8, 32.0, 29.7, 29.6, 29.6, 29.5, 29.4, 29.3, 28.0, 26.7, 25.7, 22.7, 18.0, 15.4, 14.1.

3.5.4.3 Properties of Asterosterol, 186 (SM26)

Gum; EIMS *m/z* 370; Molecular formula $C_{26}H_{42}O$; ^{1}H NMR (400 MHz, CD_3OD): δ_H 5.34, 5.20, 5.18, 3.9, 1.95-3.15, 1.28, 1.21, 1.21, 0.86.

3.6 Flavonoids

Ten flavonoids were isolated from five *Centaurea* species (**Table 20**). Of them, four were flavonols (**SM27**, **SM28**, **SM29** and **SM30**), two flavones (**SM31** and **SM32**), three flavonoid *C*-glycosides (**SM33**, **SM34** and **SM35**) and one flavanone (**SM36**). Flavonoid **SM36**, isolated from *C. montana*, is a new natural product.

Table 20. Flavonoids isolated from *Centaurea* species

Code (structure)	Source and isolation code	Page
SM27 (122)	*C. huber- morathi* (**CH5**)	74
SM28 (123)	*C. schischkinii* (**CS4**), *C. bornmuelleri* (**CB3**)	80, 82
SM29 (187)	*C. huber-morathii* (**CH4**), *C. schischkinii* (**CS6**), *C. bornmuelleri* (**CB5**)	74, 80, 82
SM30 (108)	*C. schischkinii* (**CS8**)	80
SM31 (126)	*C. gigantea* (**CG5**)	83
SM32 (114)	*C. gigantea* (**CG8**)	83
SM33 (188)	*C. gigantea* (**CG3**)	82
SM34 (129)	*C. gigantea* (**CG4**)	83
SM35 (189)	*C. gigantea* (**CG2**)	82
SM36 (190)	*C. montana* (**CM4**)	76

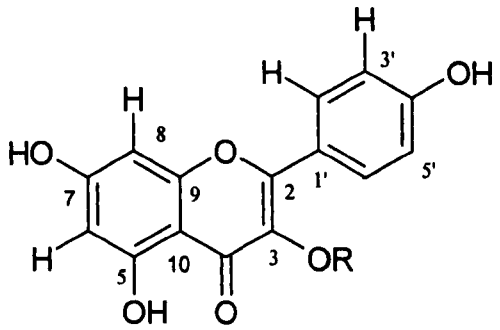
3.6.1 Characterisation of SM27, SM28 and SM29 as Kaempferol (122), Astragalin (123) and Afzelin (187), respectively

The UV (**Table 21**), IR, ^{1}H and ^{13}C NMR (**Table 22**) data of **SM27**, **SM28** and **SM29** indicated that they were flavonoids.[191] Their IR spectra showed absorption bands for conjugated carbonyl (1679 cm^{-1}) and the UV maxima were observed at 234, 252, 282 and 313 nm to support for flavonoid skeleton.[191,192] In the ^{1}H NMR spectrum (**Figure 111**; **Table 22**) two singlets at δ 6.50 (s) and 6.20 (s), and two doublets at δ 7.71 (d, *J*=8.8 Hz) and 6.92 (d, *J*=8.8 Hz) indicated that they were kaempferol derivatives.[192,193] The kaempferol type flavonol structures were also confirmed by the ^{1}H-^{1}H COSY correlations. The ^{1}H-^{1}H COSY spectrum showed

cross peaks between H-2' to H-3' and H-5' to H-6'. The ESIMS spectrum of **SM27** gave $[M-H]^+$ *m/z* ion at 285, suggesting *Mr*=286 which could be accounted for the molecular formula $C_{15}H_{10}O_6$. The spectroscopic data of **SM27** were in good agreement with literature data of kaempferol (**122**).[192,194] Kaempferol (**122**) was isolated previously from *C. calcitrapa*[49], *C. cyanus*[72] and *C. pallescens* [49] (**Table 1**).

Table 21. UV absorbance of flavonoids in methanol

Flavonoids	λ_{max} (nm)	Flavonoids	λ_{max} (nm)
SM27 (122)	215, 250, 278, 317	**SM32 (114)**	213, 252, 280, 313
SM28 (123)	213, 252, 280, 313	**SM33 (188)**	254, 272, 334
SM29 (187)	214, 250, 284, 314	**SM34 (129)**	214, 250, 280, 314
SM30 (108)	222, 254, 284, 319	**SM35 (189)**	233, 252, 295, 314
SM31 (126)	213, 248, 285, 313	**SM36 (190)**	213, 252, 343



	R
122 (SM27)	H
123 (SM28)	β-D-glucopyranosoyl
187 (SM29)	α-D-rhamnopyranosoyl

Figure 47. Structures of flavonols (**SM27-SM29**)

The ESIMS spectrum of **SM28** displayed the pseudomolecular ion peak *m/z* at 471 $[M+Na]^+$, suggesting *Mr*=448, accounted for by the molecular formula $C_{21}H_{21}O_{11}$. **SM28** is 162 mass units more than kaempferol (**SM27**). This extra mass could be assigned to a glucosyl moiety which implied that this compound was a kaempferol

glucoside. The NMR signals (**Table 22**) for the anomeric proton at δ_H 4.60 (d, *J*=7.4 Hz) and carbon at δ_C 101.6 also indicated the presence of β-glucose moiety.[153] A ^3J ^1H-^{13}C long-range correlation between the anomeric proton, δ 4.6 (H-1'') and δ_C 135.2 (C-3) in the HMBC spectrum confirmed that glucosyl unit was attached to C-3 of aglycone. All spectroscopic data were in good agreement with the published data of astragalin.[193-196] Thus **SM28** was confirmed as astragalin or kaempferol 3-*O*-β-D-glucopyranoside (**123**). This compound was previously isolated from *C. inermis*[84], *C. isaurica*[83], *C. pallescens*[49] and *C. virgata*[84] (**Table 1**).

Table 22. ^1H NMR (chemical shift, multiplicity, coupling constant *J* in Hz data for **SM27** (**122**), **SM28** (**123**) and **SM29** (**187**)

Carbon number	Chemical shift δ in ppm		
	SM27[a]	**SM28**[a]	**SM29**[a]
6	6.20, d, 2.0	6.18, d, 2.0	6.15, d, 2.0
8	6.50, d, 2.0	6.42, d, 2.0	6.32, d, 2.0
2'	7.71, d, 8.8	7.74, d, 8.8	7.72, d, 8.8
3'	6.92, d, 8.8	6.92, d, 8.8	6.90, d, 8.8
5'	6.92, d, 8.8	6.92, d, 8.8	6.90, d, 8.8
6'	7.71, d, 8.8	7.74, d, 8.8	7.72, d, 8.8
1''	–	4.60, d, 7.4	5.33, d, 1.2
2''	–	3.46[b]	4.18, dd, 3.2, 1.6
3''	–	3.46[b]	3.71[b]
4''	–	3.36[b]	3.32[b]
5''	–	3.36[b]	3.24[b]
6''	–	3.82[b] 3.62[b]	0.88, d, 5.6

[a] ^1H NMR (400 MHz) in CD_3OD
[b] Overlapped peaks

Table 23. ^{13}C NMR data for **SM27 (122)**, **SM28 (123)** and **SM29 (187)**

Carbon number	Chemical Shift δ in ppm		
	SM27[a]	**SM28**[a]	**SM29**[a]
2	158.2	158.1	158.0
3	135.4	135.2	135.0
4	182.5	178.6	178.3
5	162.0	162.0	162.0
6	99.3	98.6	98.7
7	164.3	165.1	165.3
8	94.6	93.2	93.7
9	157.4	157.4	157.4
10	104.1	104.2	104.2
1'	122.0	121.4	121.4
2'	130.7	130.7	130.7
3'	115.8	115.5	115.3
4'	161.6	160.8	160.4
5'	115.8	115.7	115.3
6'	130.7	130.7	130.7
1''	–	101.6	102.3
2''	–	73.7	70.7
3''	–	76.6	70.8
4''	–	70.1	72.0
5''	–	76.9	71.0
6''	–	61.3	16.4

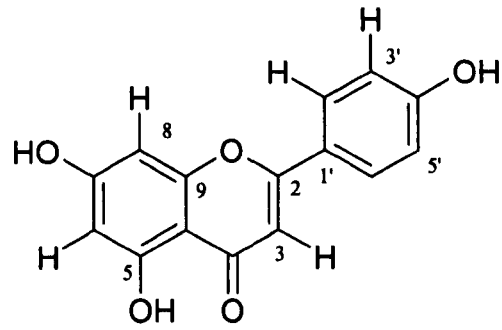
[a] ^{13}C NMR (100 MHz) in CD_3OD

The ESIMS spectrum of **SM29** gave the pseudomolecular ion at *m/z* 455 $[M+Na]^+$, calculated for the molecular formula $C_{21}H_{20}O_{10}$. **SM29** was 146 mass units larger than **SM27,** which implied that it had a rhamnose moiety. In the ^1H NMR spectrum (**Table 22**), the signals at δ 5.33 (d, *J*=1.2 Hz) and 0.88 (d, *J*=5.6 Hz), together with the ^{13}C NMR signals (**Table 23**) at δ_C 102.3 and 16.4 also suggested the presence of a rhamnose moiety. The rhamnose moiety was also confirmed by the ^1H-^1H COSY

and ^{1}H-^{13}C HMBC experiments (**Figure 114**). A long-range ^{3}J ^{1}H-^{13}C correlation between the anomeric proton, δ 5.33 (H-1'') and δ_C 135.0 (C-3) in the HMBC spectrum confirmed that rhamnosyl unit was attached to C-3 of the aglycone. All spectroscopic data were further compared with published data of afzelin and found identical.[193-197] Thus **SM29** was characterised as afzelin or kaempferol 3-*O*-α-L-rhamnopyranoside (**187**). This is a new report on the occurrence of this compound in the genus *Centaurea*.

3.6.2 Characterisation of SM30 and SM31 as Apigenin (108) and Isoquercetrin (126), respectively

The UV (**Table 21**), IR, ^{1}H and ^{13}C NMR (**Table 24**) data of **SM30** indicated that it was a flavonoid. The pseudomolecular ion peak at *m/z* 271 $[M+H]^+$ in the ESIMS (positive mode) revealed the molecular mass of 270 for this compound. The molecular formula $C_{15}H_{10}O_5$ was assigned from the HRESIMS spectrum (*m/z* at 271.0600 $[M+H]^+$, calculated value 271.0601). **SM30** had 16 lower mass units than **SM27** due to the lack of a hydroxyl group. The ^{13}C NMR spectrum (**Figure 116**) of **SM30** revealed that it had fifteen carbons. The DEPT-135 spectrum indicated the presence of seven methines (δ 129.1, 129.1, 116.6, 116.6, 103.5, 99.5, 94.6), six oxygenated quarternary carbons (δ 182.4, 164.9, 164.4, 162.1, 161.8, 158.0) and two quarternary carbons (δ 121.8, 104.3). The ^{1}H NMR spectrum (**Figure 115**) exhibited signals similar to those of **SM27** (**Table 22**) with the exception that **SM30** had an additional singlet at δ 6.72, which was attributable to H-3 of a flavone system. This was further confirmed by the ^{1}H-^{13}C HMBC experiment. The ^{3}J ^{1}H-^{13}C long range correlations were observed between H-3 and C-2 (δ_C 164.4), C-10 (δ_C 104.3) and C-1' (δ_C 121.8).



108

Figure 48. Structure of apigenin (**SM30**)

Table 24. ^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM30 (108)**

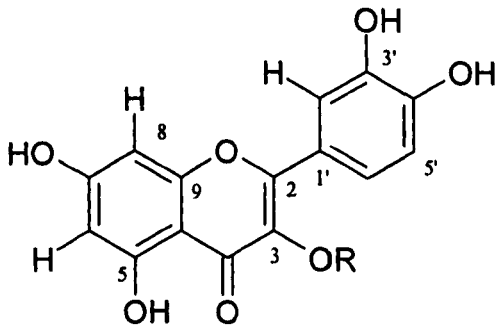
Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	–	164.4	–	–
3	6.72, s	103.5	C-2	C-1', C-10
4	–	182.4	–	–
5	–	162.1	–	–
6	6.13, d, 2.4	99.5		C-10
7		162.1		
8	6.42, d, 2.4	94.6	C-9	C-6, C-10
9	–	158.0	–	–
10	–	104.3	–	–
1'	–	121.8	–	–
2'	7.87, d, 8.8	129.1	–	C-2, C-4', C-6'
3'	6.87, d, 8.8	116.6	–	C-1', C-5'
4'		161.8	–	
5'	6.87, d, 8.8	116.6	–	C-1', C-3'
6'	7.87, d, 8.8	129.1	–	C-2, C-2', C-6'
5-OH	12.90, br s	–	–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in DMSO-d_6

The ^{1}H-^{13}C HMBC spectrum also showed key correlations between H-2' (δ_H 7.87) and C-3' (δ 129.1), H-3' (δ 6.87) and C-1' (δ 121.8), H-6 (δ 6.13) and C-10 (δ 104.3), H-8 (δ 6.42) and C-6 (δ 99.5), C-10 (δ 104.3). By analysing the ^{1}H, ^{13}C NMR, ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC data and comparing with published data, the

structure of **SM30** was elucidated as apigenin (**108**).[198] Apigenin was previously isolated from *C. pallescens*[49], *C. nervosa*[100], *C. orphanidea*[106] and *C. phrygia*[100] (**Table 1**).

The ESIMS spectrum of **SM31** displayed the pseudomolecular ion peak *m/z* at 487 $[M+Na]^+$, suggesting *Mr*=464, calculated for molecular formula $C_{21}H_{20}O_{12}$. **SM31** was 16 mass unit larger than kaempferol 3-*O*-β-D-glucopyranoside, **123** (**SM28**) due to an additional hydroxyl group. The ^1H NMR spectra data of **SM31** indicated the presence of two doublets at δ 6.36 and 6.17, and an ABX benzene spin system [δ 7.66 (d, *J*= 2.0 Hz), 7.55 (dd, *J*=8.4, 2.0 Hz), and 6.83 (d, *J*= 8.4 Hz)] on flavone ring B. The signals for the anomeric proton at δ_H 5.20 (d, *J*=7.6 Hz) and carbon at δ_C 101.6 also indicated the presence of β-glucose moiety.[153] The ^1H and ^{13}C NMR data (**Section 3.6.9.5**) were in good agreement with the published data of quercetin 3-*O*-β-D-glucopyranoside or isoquercetrin (**126**) .[196]



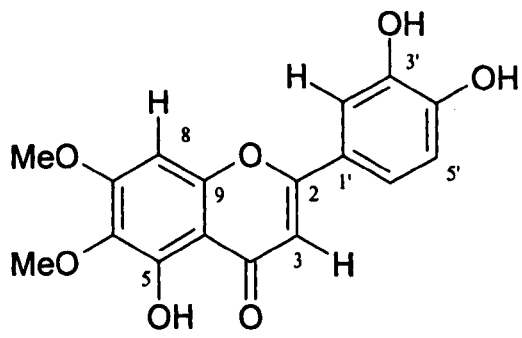
126 R=β-D-glucopyranosyl

Figure 49. Structure of isoquercetrin (**SM31**)

3.6.3 Characterisation of SM32 as 6-Hydroxy-luteolin 6,7-dimethylether (114)

The EIMS spectrum of **SM32** showed the molecular ion *m/z* at 330, accounted for by the molecular formula $C_{17}H_{14}O_7$. The ^{13}C NMR spectrum of **SM32** (**Table 25**) revealed that it had seventeen carbons. The DEPT-135 spectrum indicated the presence of five methines (δ 119.7, 116.6, 114.1, 103.3, 92.1), eight oxygenated quarternary carbons (δ 182.8, 164.9, 159.2, 153.2, 152.7, 150.6, 146.5, 132.5), two quarternary carbons (δ 122.0, 105.7) and two methoxy groups (δ 60.7, 57.1). All protonated carbons were assigned by ^1H-^{13}C HSQC experiment.[150]

The ^1H NMR spectrum (**Table 25**) indicated the presence of two singlets at δ 6.83 and 6.68, and an ABX benzene spin system [δ 7.40 (dd, *J*=8.0, 2.0 Hz), 7.39 (d, *J*= 2.0 Hz) and 6.85 (d, *J*= 8.0 Hz)] on flavone ring B. Two 3H singlets at δ_H 3.87 and 3.67 could be assigned to two methoxy groups linked directly to the aromatic rings. The signals at δ 6.83 (s) and 6.68 (s) were assigned to proton H-8 and H-3 of the flavone ring, and were confirmed by the ^1H-^{13}C HMBC experiment (**Table 25**).



114

Figure 50. Structure of cirsiliol (**SM32**)

The key HMBC correlations were observed between δ_H 6.83 (H-8) to δ_C 132.2 (H-6), 159.2 (C-7), 153.2 (C-9) and 105.7 (C-10). The ^3J ^1H-^{13}C long-range correlation

between δ_H 3.87 and δ_C 159.2 (C-7), and δ_H 3.67 and δ_C 132.5 (C-6) confirmed that methoxy groups were linked, respectively, to C-7 and C-6 on the flavone ring A. Thus the structure of **SM32** was confirmed as 6-hydroxy luteolin 6,7-dimethylether or cirsiliol (**114**). Cirsiliol was previously isolated from many plants[199] including *C. paniculata*[86].

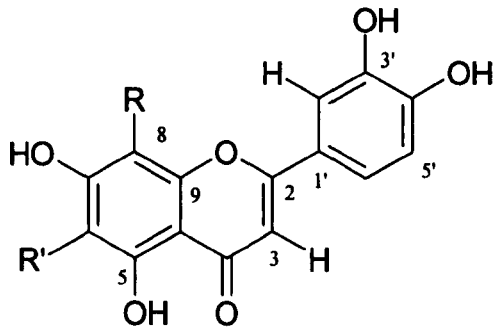
Table 25. ^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM32 (110)**

Carbon number	Chemical shift δ in ppm		HMBC (^{1}H→^{13}C)	
	^{1}H[a]	^{13}C[a]	^2J	^3J
2	–	164.9	–	–
3	6.68, s	103.3	C-2, C-4	C-1', C-10
4	–	182.8	–	–
5	–	152.7	–	–
6	–	132.5	–	–
7	–	159.2	–	–
8	6.83, s	92.1	C-7, C-9	C-6, C-10
9	–	153.2	–	–
10	–	105.7	–	–
1'	–	122.0	–	–
2'	7.39, d, 2.0	114.1	C-3'	C-4', C-6'
3'	–	146.5	–	–
4'	–	150.6	–	–
5'	6.85, d, 8.0	116.6	C-6'	C-1',C-3'
6'	7.40, dd, 8.0, 2.0	119.7	C-5'	C-2', C-4'
O-CH_3 (6)	3.67, s	60.7	–	C-6
O-CH_3 (7)	3.87, s	57.1	–	C-7
5-OH	12.92, br s	–	–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in DMSO-d_6

3.6.4 Characterisation of SM33 and SM34 as orientin (188) and isoorientin (129), respectively

The ESIMS spectrum of **SM33** showed the pseudomolecular ion peak at *m/z* 471 $[M+Na]^+$, suggesting *Mr*=448 and solving for $C_{21}H_{20}O_{11}$. The UV (**Table 21**), ^1H and ^{13}C NMR data (**Table 26**) indicated that **SM33** was a flavone. In the ^1H NMR spectrum (**Table 26**), two singlets at δ 6.58 (s) and 6.21 (s), and signals at δ 7.48 (dd, *J*=8.4, 2.0 Hz), 7.42 (d, *J*=2.0 Hz) and 6.80 (d, *J*=8.4 Hz) associated with B ring protons suggested that the compound was a luteolin derivative.[200] A singlet at δ_H 12.93 was assigned to the hydroxyl group linked to the C-5 of the flavone. The ^1H and ^{13}C NMR spectra also showed the presence of a glucose moiety. The anomeric proton signal was observed at δ_H 4.62 as a doublet (*J*=10.0 Hz) which was indicative of β configuration.[201] In the ^{13}C NMR spectrum (**Figure 118**; **Table 26**) six carbon signals at δ 82.6, 79.4, 74.0, 71.4, 71.3 and 62.3 suggested that **SM33** was a flavone *C*-glucoside.[202] The signals at δ_H 6.58 and 6.21 could be assigned to H-3 and H-6 of ring A. These assignments were confirmed by the ^1H-^{13}C HSQC and ^1H-^{13}C HMBC experiments. In the HMBC spectrum (**Figure 119**) H-3 (δ_H 6.58) showed ^3J correlations to δ_C 122.7 (C-1') and 104.7 (C-10), and ^2J correlations to δ_C 182.7 (C-4) and 164.7 (C-2). H-6 (δ_H 6.21) showed ^3J correlations to δ_C 105.2 (C-8) and 104.7 (C-10), and ^2J correlations to δ_C 163.2 (C-7) and 161.0 (C-5). The anomeric proton, δ 4.62 (H-1'') showed ^3J correlations to δ_C 163.2 (C-7) and 156.6 (C-9), and ^2J correlation to δ_C 105.2 (C-8) and confirmed *C*-glucosidation at C-8. Thus, the structure of **SM33** was elucidated as orientin (**188**). All spectroscopic data of **SM33** were in good agreement with literature data for orientin.[201-203]



	R	R'
188 (SM33)	β-D-glucopyranosoyl	H
129 (SM34)	H	β-D-glucopyranosoyl

Figure 51. Structure of *C*-glycosides **(SM33 and SM34)**

The ^{1}H and ^{13}C NMR data of **SM34** were similar to those of **SM33,** a flavonoid *C*-glycoside. The ESIMS spectrum revealed *m/z* at 471 $[M+Na]^{+}$, accounted for by the molecular formula $C_{21}H_{20}O_{11}Na$ which was the same as **SM33.** However, NMR data analyses revealed that **SM34** was in fact a positional isomer of orientin (**SM33**). The signals at δ_H 6.45 (s) in the ^{1}H NMR spectrum and δ_C 94.0 in the ^{13}C NMR spectrum were attributable to H-8 of flavone, indicating the site of *C*-glucosidation at C-6. This was also supported from the carbon chemical shift of C-6 at δ107.2. The ^{1}H and ^{13}C NMR data were compared with published data of isoorientin (**129**) and found identical.[204-205]

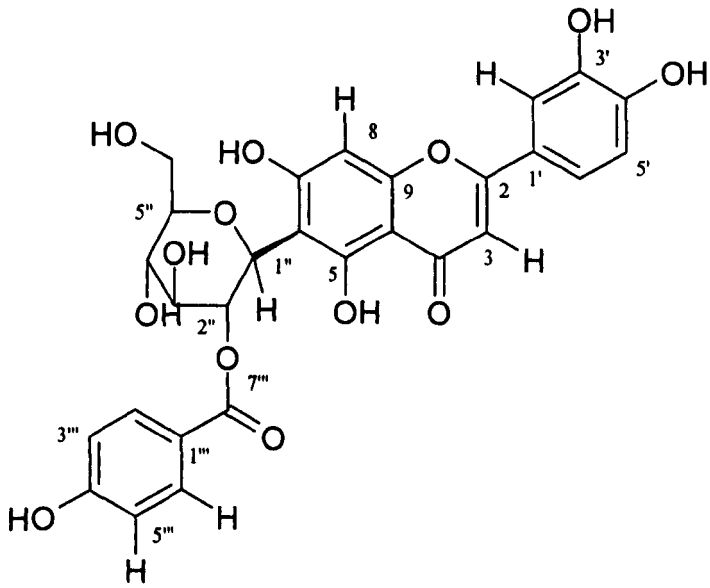
Table 26. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM33 (188)**

Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	–	164.7	–	–
3	6.58, s	103.0	C-2, C-4, C-10	C-1', C-10
4	–	182.7	–	–
5	–	161.0	–	–
6	6.21, s	98.8	C-5, C-7	C-8, C-10
7	–	163.2	–	–
8	–	105.2	–	–
9	–	156.6	–	–
10	–	104.7	–	–
1'	–	122.7	–	–
2'	7.42, d, 2.0	114.7	C-3'	C-2, C-4', C-6'
3'	–	146.4	–	–
4'	–	150.2	–	–
5'	6.80, d, 8.4	116.3	C-4'	C-1', C-3'
6'	7.48, dd, 8.4, 2.0	120.0		C-2, C-2', C-4'
1''	4.62, d, 10	74.0	C-2'', C-8	C-7, C-9, C-3'', C-5''
2''	3.79, m	71.4	C-1'', C-3''	–
3''	3.20, m	79.4	C-2'', C-4''	–
4''	3.29, m	71.3	C-2'', C-4''	–
5''	3.10, m	82.6	C-4''	C-3'
6''	3.75, m 3.50, m	62.3	–	–
5-OH	12.92, br s		–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in DMSO-d6

3.6.5 Characterisation of SM35 as 2''-(4'''-Hydroxybenzoyl)-isoorientin (189)

The ESIMS spectrum of **SM35** showed the pseudomolecular ion peak at *m/z* 591 $[M+Na]^+$, suggesting *Mr*=568 and was calculated for the molecular formula $C_{28}H_{24}O_{13}$. The UV (**Table 21**), ^1H and ^{13}C NMR data (**Table 27**) revealed that **SM35** was a flavone *C*-glycoside like **SM34**. The spectroscopic data of **SM35** were similar to those of isoorientin (**SM34**) with the exceptions that the ^1H NMR spectrum (**Table 27**) of **SM35** showed additional resonances at δ_H 7.90 (d, *J*=8.8 Hz) and 6.87 (d, *J*=8.8 Hz) which were assignable to a 4-hydroxybenzoyl moiety. In the ^{13}C NMR, seven signals at δ_C 165.4, 161.6, 128.9, 128.9, 122.1, 115.8 and 115.8 also confirmed the presence of this group.



189

Figure 52. Structure of 2''-(4'''-hydroxybenzoyl)- isoorientin (**SM35**)

The attachment of this moiety at C-2'' of the glucose unit was confirmed by a ^3J ^1H-^{13}C correlation from H-2'' (δ_H 4.15) to the carbonyl carbon C-7''' (δ_C 165.4) observed in the HMBC spectrum (**Figure 120**). Thus, the structure of **SM35** was determined as 2''-(4'''-hydroxybenzoyl)-isoorientin (**189**). This compound was

previously isolated from *Gentiana asclepiadea*.[206] However, this is a new report on the occurrence of this compound in *C. gigantea* and even in the genus *Centaurea*.

Table 27. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM35 (189)**

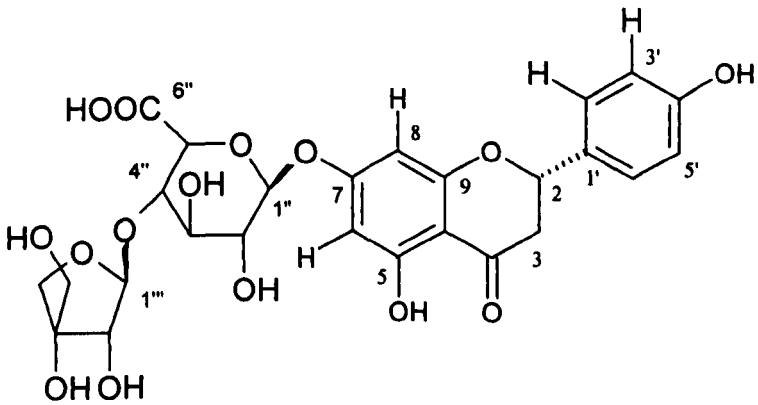
Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	–	165.0	–	–
3	6.45, s	102.6	C-2, C-4, C-10	C-1'
4	–	182.7	–	–
5	–	160.8	–	–
6	–	107.9	–	–
7	–	163.6	–	–
8	6.41, s	94.0	C-7, C-9	C-6, C-10
9	–	157.4	–	–
10	–	104.0	–	–
1'	–	122.3	–	–
2'	7.27, m	113.0	C-3'	C-6'
3'	–	145.8	–	–
4'	–	149.8	–	–
5'	6.83, m	115.7	C-1'	C-3'
6'	7.28, m	119.1	–	C-2', C-3'
1''	4.98[b]	74.1	C-6, C-2''	C-5, C-7, C-3'', C-5''
2''	4.15, m	75.3	–	C-7'''
3''	3.50, m	78.9	–	–
4''	4.05, m	71.2	–	–
5''	3.49, m	81.4	C-4''	C-3'
6''	3.90, m 3.75, m	61.7	–	–
1'''	–	122.1	–	–
2'''	7.90, d, 8.8	128.9	C-7'''	C-4''', C-6'''
3'''	6.87, d, 8.8	115.8	C-4'''	C-1''', C-5'''
4'''	–	161.6	–	–
5'''	6.87, d, 8.8	115.8	C-4'''	C-1''', C-3'''
6'''	7.90, d, 8.8	128.9	C-7'''	C-2''', C-4'''
7'''	–	165.4	–	–

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD
[b] Overlapped peak

3.6.6 Characterisation of SM36 as Flavanone 7-*O*-apiofuranosyl (1→2)-glucuronic acid (190)

The ESIMS spectrum of **SM36** showed the pseudomolecular ion peak at *m/z* 603 $[M+Na]^+$ suggesting *Mr*=580. The HRCIMS gave the pseudomolecular ion *m/z* at 598.1769 $[M+NH_4]^+$ (calculated 598.1766 for $C_{26}H_{32}NO_{15}$). The UV absorptions at 213, 252 and 343 nm and the IR absorption band for conjugated carbonyl (1679 cm^{-1}) were indicative of a flavanone skeleton.[191,207] The ^1H and ^{13}C NMR spectra (**Table 28**) of **SM36** showed the presence of one methylene [δ_H 3.12 (dd, *J*=13.0, 17.0 Hz) and 2.73 (dd, *J*=2.0, 17.0 Hz); δ_c=42.4], one oxymethine [δ_H 5.36 (dd, *J*=2.0, 13.0 Hz); δ_c=79.0] and two methines [δ_H 6.15 (d, *J*=2.0 Hz); δ_c=95.0 and 6.11 (d, *J*=2.0 Hz); δ_c=96.5] of the ring A, and 4 methines in a group of two chemically equivalent protons [δ_H 7.28 (d, *J*=8.4 Hz) and 6.77 (d, *J*=8.4 Hz)] of the ring B (**Figure 121**). These data also supported **SM36** to be a flavanone.[208] The presence of a 1,4-disubstituted benzene ring system was evident from these signals which was further confirmed by cross peaks between H-2'/6' and H-3'/5' in the ^1H-^1H COSY spectrum. The ^1H-^{13}C HMBC correlations between H-6 (δ_H 6.11) and C-8 (δ_C 95.0) and C-10 (δ_C 104.0), and H-8 (δ_H 6.15) and C-6 (δ_C 96.5), C-7 (δ_C 164.0) and C-10 (δ_C 104.0) further confirmed the flavanone skeleton. The ^1H and ^{13}C NMR spectra (**Table 28**) also revealed signals representing two sugar moieties. A doublet at δ 5.05 (d, *J*=7.2 Hz) and additional signals (δ 3.47-3.72) in the ^1H NMR implied that one of the sugars was a β-glucose derivative and the carbon signal at δ 173.0 in the ^{13}C NMR confirmed that it was a β-D-glucuronic acid moiety.[209] Five more ^{13}C NMR signals at δ 109.0, 79.9, 77.5, 74.0 and 65.0 could be assigned to the carbons of another sugar, apiose.[210] A ^3J ^1H-^{13}C long range coupling between the anomeric proton of apiose, δ_H 5.38 (H-1'') and δ_C 75.0 (C-4'') in the HMBC spectrum

(**Figure 123**) confirmed that the apiose was connected to glucuronic acid at C-4''. The presence of apiose-glucuronic acid was also confirmed by the loss of 309 mass units from the molecular mass of the compound in the ESIMS spectrum. The ^{3}J ^{1}H-^{13}C long-range HMBC correlation between δ_H 5.05 (H-1'') and δ_C 164.0 (C-7) confirmed that the glucuronic acid moiety was connected to C-7 of the flavanone skeleton. The specific optical rotation for **SM36** was found -48°. Literature review showed that all natural (-)-flavanones have been found to have *S*-chirality at C-2.[211,212] Comparing with other flavanones of established absolute configuration measured by using circular dichroism spectroscopy, **SM36** was considered as *S* stereochemistry at C-2 due to its levorotatory nature.[211,212] Thus, **SM36** was determined as flavanone 7-*O*-apiofuranosyl (1→4)-glucuronic acid (**190**). To the best of our knowledge this is a novel compound.



190

Figure 53. Structure of flavanone 7-*O*-apiofuranosyl(1→4)-glucuronic acid (**SM36**)

Table 28. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM36 (190)**

Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	5.36, dd, 2.0, 13.0	79.0	–	–
3	3.12, dd, 13.0, 17.0 2.73,dd , 2.0, 17.0	42.5	C-4	C-1'
4	–	196.0	–	–
5		162.0	–	–
6	6.11, d, 2.0	96.5	–	C-8, C-10
7		164.0	–	–
8	6.15, d, 2.0	95.0	C-7	C-6, C-10
9	–	162.0	–	–
10	–	104.0	–	–
1'	–	130.0	–	–
2'	7.28, d, 8.4	129.0	C-3'	C-2, C-4', C-6'
3'	6.77, d, 8.4	115.1	C-4'	C-1', C-5'
4'	–	158.0	–	–
5'	6.77, d, 8.4	115.1	C-4'	C-1', C-3'
6'	7.28, d, 8.4	129.0	C-5'	C-2, C-2', C-4'
1''	5.05, d, 7.2	99.0	–	C-7
2''	3.47, m	73.0,	C-3''	–
3''	3.61, m	77.8	–	–
4''	3.72, t, 9.2	75.0	–	–
5''	3.65, m	78.0	–	C-1''
6''		173.0	–	–
1'''	5.38, d, 1.0	109.0	C-2'''	C-4'', C-3'''
2'''	3.89, m	77.5		C-4''', C-5'''
3'''	–	79.9	–	–
4'''	3.91, m	74.0	–	C-5'''
5'''	3.47, m	65.0	C-3'''	C-4'''

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.6.7 Properties of Flavonoids (SM27-SM36)

3.6.7.1 Properties of Kaempferol, 122 (SM27)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 285 $[M-H]^+$; Molecular formula $C_{15}H_{10}O_6$; ^1H NMR (400 MHz, CD_3OD): **Table 22**; ^{13}C NMR (100 MHz, CD_3OD): **Table 23**.

3.6.7.2 Properties of Astragalin, 123 (SM28)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 471 $[M+Na]^+$, 285 $[M\text{-glucose-H}]^+$, 137, 70; Molecular formula $C_{21}H_{20}O_{11}$; ^1H NMR (400 MHz, CD_3OD): **Table 22**; ^{13}C NMR (100 MHz, CD_3OD): **Table 23**.

3.6.7.3 Properties of Afzelin, 187 (SM29)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 433 $[M+H]^+$, 285 $[M\text{-rhamnose-H}]^+$, 233, 137, 70; Molecular formula $C_{21}H_{20}O_{10}$; ^1H NMR (400 MHz, CD_3OD): **Table 22**; ^{13}C NMR (100 MHz, CD_3OD): **Table 23**.

3.6.7.4 Properties of Apigenin, 108 (SM30)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 271 $[M+H]^+$, HRESIMS *m/z* 271.0600 (calculated 271.0601 for $C_{15}H_{11}O_5$; ^1H NMR (400 MHz, CD_3OD): **Table 24**; ^{13}C NMR (100 MHz, CD_3OD): **Table 24**.

3.6.7.5 Properties of Isoquercetrin, 126 (SM31)

Gum; UV λ_{max} (MeOH): **Table** ; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 487 $[M+Na]^+$; Molecular formula $C_{21}H_{20}O_{12}$; ^1H NMR (400 MHz, CD_3OD): δ 7.66 (d, *J*=2.0 Hz), 7.55 (dd, *J*=8.4, 2.0 Hz), 6.83 (d, *J*=8.4 Hz), 6.36 (d, *J*=2 Hz),

6.17 (d, *J*=2 Hz), 5.2 (d, *J*=7.6 Hz), 3.82-3.20, m; ^{13}C NMR (100 MHz, CD_3OD): δ 180.1 (C-4), 165.0 (C-7), 161.9 (C-5), 157.3 (C-2), 153.0 (C-9), 150.1 (C-4'), 144.7 (C-3'), 134.4 (C-3), 122.0 (C-1'), 120.0 (C-6'), 116.3 (C-5'), 114.2 (C-2'), 104.1 (C-10), 101.6 (C-1''), 98.6 (C-6), 93.5 (C-8), 77.2 (C-5''), 77.1 (C-3''), 74.5 (C-4''), 70.0 (C-2'') and 61.3 (C-6'').

3.6.7.6 Properties of Cirsiliol, 114 (SM32)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; EIMS *m/z* 330 $[M]^{+\cdot}$ 315, 75 (100); Molecular formula $C_{17}H_{14}O_7$; ^{1}H NMR (400 MHz, CD_3OD): **Table 25**; ^{13}C NMR (100 MHz, CD_3OD): **Table 25**.

3.6.7.7 Properties of Orientin, 188 (SM33)

Gum; UV λ_{max} (MeOH): **Table**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESI-MS *m/z* 471 $[M+H]^+$; Molecular formula $C_{21}H_{20}O_{10}$; ^{1}H NMR (400 MHz, CD_3OD): **Table 26**; ^{13}C NMR (100 MHz, CD_3OD): **Table 26**.

3.6.7.8 Properties of Isoorientin, 129 (SM34)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 471 $[M+H]^+$; Molecular formula $C_{21}H_{20}O_{10}$; ^{1}H NMR (400 MHz, CD_3OD): δ 7.48 (dd, 8.4, 2.0 Hz), 7.42 (d, 2.0 Hz), 6.80 (d, 8.4 Hz), 6.58 (s), 6.45 (s), 4.60 (d, 10), 3.78-3.10 (m); ^{13}C NMR (100 MHz, CD_3OD): δ 182.7, 164.5, 161.0, 156.2, 150.2, 146.4, 122.8, 122.7, 116.3, 114.7, 114.5, 120.0, 107.2, 104.6, 103.0, 94.0, 82.5, 79.4, 71.4, 71.3, 62.4.

3.6.7.9 Properties of 4'''-Hydroxybenzoyl- isoorientin, 189 (SM35)

Gum; UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1205 cm^{-1}; ESIMS *m/z* 591 $[M+Na]^+$; Molecular formula $C_{28}H_{24}O_{13}$; ^{1}H NMR (400 MHz, CD_3OD): **Table 27**; ^{13}C NMR (100 MHz, CD_3OD): **Table 27**.

3.6.7.10 Properties of Flavanone 7-*O*-apiofuranosyl (1→4)-glucuronic acid, 190 (SM36)

Yellow amorphous; $[\alpha]^{23}_{D}$ -48° (c 0.021, MeOH); UV λ_{max} (MeOH): **Table 21**; IR ν_{max} (neat): 3459, 1679 and 1246 cm^{-1}; ESI-MS *m/z* 603 $[M+Na]^+$, 307, $[glucose+apiose+2H]^+$, 271 $[M-glucose-apiose]^+$, 104, 60; HRCIMS *m/z* 598.1769 $[M+NH_4]^+$ (calculated 598.1766 for $C_{26}H_{32}NO_{15}$); ^1H NMR (400 MHz, CD_3OD): **Table 28**; ^{13}C NMR (100 MHz, CD_3OD): **Table 28**.

3.7 Alkaloids

Nine alkaloids (**SM37-SM45**) were isolated from three *Centaurea* species in this study. All of them were indole alkaloids.

3.7.1 Isolation and Characterisation of Indole Alkaloids

Eight indole alkaloids **SM37** (isolation code **CM3**), **SM38** (isolation code **CM10**), **SM39** (isolation code **CM11**), **SM40** (isolation code **CM18**), **SM41** (isolation code **CM19**), **SM42** (isolation code **CM12**), **SM43** (isolation code **CM13**) **and SM44** (isolation code **CM17**) were isolated from *C. montana* seeds (**Section 2.7.6**). **SM42** and **SM43** were also purified from *C. cyanus* seeds (**Section 2.7.2**). *C. schischkinii* afforded alkaloid **SM45** (isolation code **CS1**, **Section 2.7.7**). **SM44** and **SM45** were novel compounds.

Table 29. UV absorptions of indole alkaloids in methanol

Alkaloids	λ_{max} (nm)	Alkaloids	λ_{max} (nm)
SM37 (191)	226, 290, 318	**SM42 (148)**	228, 260, 292, 313
SM38 (145)	226, 291, 317	**SM43 (149)**	227, 282, 291, 313
SM39 (192)	227, 280, 303, 320	**SM44 (193)**	213, 271, 274
SM40 (146)	227, 280, 303, 320	**SM45 (194)**	228, 255, 283, 301, 313
SM41 (147)	228, 257, 283, 301, 313		

3.7.1.1 Identification of Indole Alkaloids

All nine compounds (**SM37-SM45**) gave orange-red colour with Dragendorff's reagent and were visualised as blue fluorescent as well as quenching spots on TLC under UV light at 366 and 254 nm respectively indicating them to be alkaloids. The UV absorption maxima (**Table 29**) of them were typical of indole chromophore.[95] The characteristics band at 3479 cm^{-1} for N-H stretching in the IR spectrum was also

indicative of indole alkaloids. The presence of indole moieties was further confirmed by ^{1}H and ^{13}C NMR spectral data.

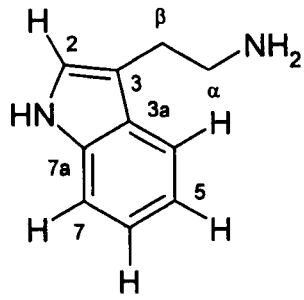
Table 30. Mass spectral data for indole alkaloids

Alkaloids	LRMS	HRMS $[M+H]^+$		Molecular
	$[M+Na]^+$	Found	Required	formula
SM37 (191)	183	–	–	$C_{10}H_{10}N_2O$
SM38 (145)	345	323.1391	323.1390	$C_{19}H_{18}N_2O_3$
SM39 (192)	345	323.1390	323.1390	$C_{19}H_{18}N_2O_3$
SM40 (146)	359	337.1551	337.1547	$C_{20}H_{20}N_2O_3$
SM41 (147)	359	337.1551	337.1547	$C_{20}H_{20}N_2O_3$
SM42 (148)	375	375.1320	3.75.1321[a]	$C_{20}H_{20}N_2O_4$
SM43 (149)	375	–	–	$C_{20}H_{20}N_2O_4$
SM44 (193)	725	725.2590[a]	725.2582[a]	$C_{40}H_{38}N_4O_8$
SM45 (194)	475	475.1857[a]	475.1858[a]	$C_{26}H_{24}N_6O_2$

[a] $[M+Na]^+$

3.1.7.2 Characterisation of SM37 as Tryptamine (191)

The ^{13}C NMR spectrum of **SM37** (**Table 31**) revealed that this compound had ten carbons. The DEPT-135 indicated the presence of two methylenes (δ 55.5, 27.5), five methines (δ 124.0, 118.0, 118.5, 121.0, 111.0) and three quarternary carbons (δ 138.0, 127.0, 109.0). All protonated carbons were assigned by ^{1}H-^{13}C HSQC experiment.[150] The ^{1}H NMR spectrum showed that **SM37** had four aromatic protons in an ABCD system. The ^{13}C NMR and ^{1}H NMR signals (7.66, d, *J*=8.0 Hz; 7.32, d, *J*=8.0 Hz; 7.07, ddd; 7.01, ddd) suggested that **SM37** was tryptamine.[213] The ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC correlations (**Figure 124**) confirmed the identity of **SM37** as tryptamine (**191**).



191

Figure 54. Structure of tryptamine (**SM37**)

Table 31. ^{1}H NMR (chemicals shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM37 (191)**

Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)
	$^{1}H^{a}$	$^{13}C^{a}$	
2	7.14, s	124.0	C-3,C-3a, C-7a
3	–	109.0	–
3a	–	127.0	–
4	7.66, d, 8.0	118.0	C-3, C-3a, C-6, C-7a
5	7.01, ddd	118.5	C-3a, C-7
6	7.07, ddd	121.0	C-4, C-7a
7	7.31, d, 8.0	111.0	C-3a, C-5
7a	–	138.0	–
α-CH_2	3.80, dd, 9.6, 4.8	27.5	C-3, C-βC
β-CH_2	3.48, d, 3.6	55.5	C-2, C-3a, C-4

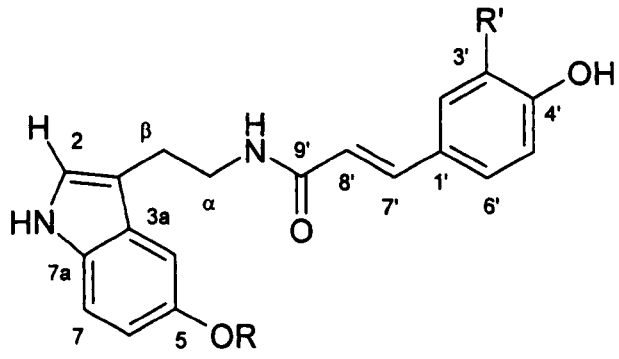
[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.7.1.3 Characterisation of SM38 and SM39 as (*E*) *N*-(4-Hydroxycinnamoyl)-5-hydroxytryptamine (145) and (*Z*) -*N*-(4-Hydroxycinnamoyl)-5-hydroxytryptamine (192), respectively

The HRESIMS data (**Table 30**) of **SM38** and **SM39** revealed the pseudomolecular ion *m/z* 323.1391 $[M+H]^{+}$ (calculated 323.1390 for $C_{19}H_{19}N_2O_3$). Analysis of the ^{1}H and ^{13}C NMR data revealed that **SM38** and **SM39 (Table 32)** had a set of two identical absorption peaks, and enabled the construction of two substructures **A** and **B (Figure 56)**. The ^{1}H NMR spectrum showed that substructure **A** had an ABX spin

system by displaying the signals at δ 7.11 (d, *J*=8.0 Hz), 6.91 (d, *J*=2.0 Hz) and 6.63 (dd, *J*=8.0, 2.0 Hz). The substructure **A** also gave rise to additional signals at δ 6.98 (s), 3.52 (t, *J*=7.2 Hz) and 2.89 (t, *J*=7.2 Hz) for **SM38** and 6.91 (s), 3.47 (t, *J*=7.2 Hz) and 2.83 (t, *J*=7.2 Hz) for **SM39** in the ^{1}H NMR spectrum. The ^{1}H and ^{13}C NMR data (**Table 32**) confirmed that substructure **A** was identical to that of a tryptamine derivative, 5-hydroxytyptamine (or serotonin).[95]

The ^{1}H NMR signals at δ 7.33 (d, *J*=8.4 Hz) and 6.75 (d, *J*=8.4 Hz) for **SM38**, and 7.35 (d *J*=8.0 Hz) and 6.67 (d, *J*=8.0 Hz) for **SM39** in the ^{1}H NMR spectrum revealed that substructure **B** had a 1,4-disubstituted benzene ring system. The presence of two olefinic protons was also evident from the signals at δ 7.41 (d, *J*=15.6 Hz) and 6.35 (d, *J*=15.6 Hz) for **SM38** and 6.57 (d, *J*=12.0 Hz) and 5.77 (d, *J*=12.0 Hz) for **SM39**. The higher coupling constant (*J*=15.6 Hz) between olefinic protons for **SM38** confirmed that it had a *trans* configuration and **SM39** possessed *cis* configuration.[214]



Alkaloids	R	R'
145 (SM38)	H	H
146 (SM40)	Me	H
148 (SM42)	H	OMe

Figure 55. Structure of indole alkaloids **(SM38, SM40 and SM42)**

The ^{13}C NMR signal at δ 168.0 indicated the presence of a carbonyl carbon. All these findings from ^{1}H and ^{13}C NMR confirmed that the substructure **B** was a *p*-coumaroyl moiety.[71] The presence of a tryptamine derived substructure and a *p*-coumaroyl moiety were further confirmed by the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC experiments (**Table 32**). The ^{1}H-^{1}H COSY revealed four different spin systems:H-7↔H-6↔H-4, $H_2\alpha \leftrightarrow H_2\beta$, H-7'↔H-8' and H-2' (H-6') ↔H-3' (H-5'). The ^{1}H-^{13}C HMBC correlations were observed between H-2 to C-3, C-3a, C-7a; H-4 to C-2, C-3a, C-4, C-7a; H-2' to C-4', C-6', and H-3' to C-1' and C-5'. The ^{1}H-^{13}C long-range correlation from $H_2\alpha$ (δ 3.52/3.47) to carbonyl carbon C-9' (δ 168.0) formed an amide linkage between the *p*-coumaroyl and serotonin moieties giving the complete structure of **SM38** as (*E*) *N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**145**) and **SM39** as (*Z*) *N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**192**).[215]

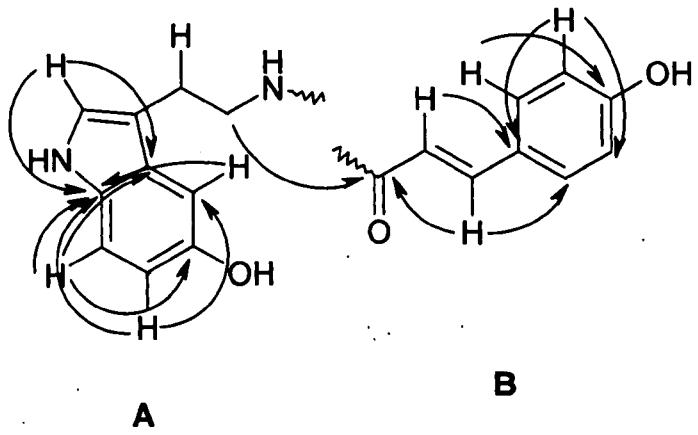


Figure 56. Key ^{1}H-^{13}C HMBC correlations of **SM38 (145)**

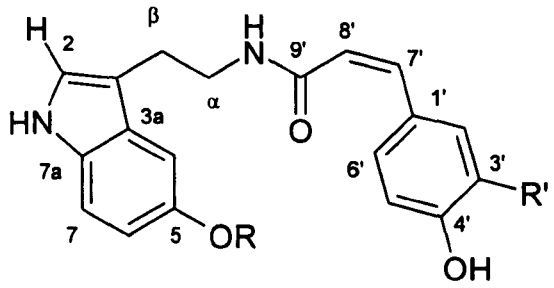
Table 32. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long-range HMBC for **SM38 (145)** and **SM39 (192)**

Carbon number	Chemical shift δ in ppm				HMBC ($^{1}H \rightarrow ^{13}C$) of **SM38**
	$^{1}H^{a}$ **(SM38)**	$^{1}H^{a}$ **(SM39)**	$^{13}C^{a}$ **(SM38)**	$^{13}C^{a}$ **(SM39)**	
2	6.98, s	6.91, s	123.0	123.0	C-3, C-3a, C-7a
3	–	–	110.5	111.0	–
3a	–	–	128.0	128.0	–
4	6.91, d, 2.0	6.91, d, 2.0	102.3	102.3	C-2, C-3a, C-6, C-7a
5	–	–	149.0	149.0	–
6	6.63, dd, 2.0, 8.0	6.63, dd, 8.0, 2.0	111.2	111.2	C-7a
7	7.11, d, 8.0	7.11, d, 8.0	111.4	111.4	C-3a, C-5, C-7a
7a	–	–	131.9	131.0	–
α-CH_2	3.52, t, 7.2	3.47, t, 7.2	40.2	39.9	C-3, C-β, C-9'
β-CH_2	2.89, t, 7.2	2.83, t, 7.2	25.2	24.9	C-2, C-3, C-3a, C- α
1'	–	–	127.0	129.3	–
2'	7.33, d, 8.4	7.35, d, 8.0	129.3	128.2	C-4', C-6'
3'	6.75, d, 8.4	6.67, d, 8.0	115.7	114.8	C-1', C-5'
4'	–	–	160.0	159.5	–
5'	6.75, d, 8.4	6.67, d, 8.0	115.7	114.8	C-1', C-3'
6'	7.33, d, 8.4	7.35, d, 8.0	129.3	128.2	C-2', C-4'
7'	7.41, d, 15.6	6.57, d, 12.0	140.6	136.7	C-2', C-9'
8'	6.35, d, 15.6	5.77, d, 12.0	117.1	120.4	C-1'
9'	–	–	168.0	168.0	–

• ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.7.1.4 Characterisation of SM40 and SM41 as Centcyamine (146) and *cis*-Centcyamine (147), respectively

The HRESIMS data (**Table 30**) of **SM40** and **SM41** revealed the pseudomolecular ion *m/z* 337.1551 $[M+H]^{+}$ (calculated 337.1547 for $C_{20}H_{21}N_2O_3$). The ^{1}H NMR and ^{13}C NMR spectra (**Table 33**) of **SM40** were similar to that of **SM38** indicating the presence of a tryptamine derived substructure and a *p*-coumaroyl moiety.



Alkaloids	R	R'
192 (SM39)	H	H
147 (SM41)	Me	H
149 (SM43)	H	OMe

Figure 57. Structure of indole alkaloids **(SM39, SM41 and SM43)**

The signals at δ 3.81 in the ^{1}H NMR spectrum and δ 55.1 in the ^{13}C NMR suggested that **SM40** had an additional methoxy group. The long range ^{1}H-^{13}C ^{3}J correlation between the methoxy signal (δ 3.81) and δ 148.7(C-5) confirmed that the methoxy group was linked to C-5 of the tryptamine derived substructure. The long range ^{1}H-^{13}C correlation between $H_2\alpha$ (δ 3.57) and the carbonyl carbon C-9' (δ 168.0) confirmed the formation of an amide linkage between the *p*-coumaroyl moiety and 5-methoxy tryptamine giving the complete structure of **SM40** as (*E*) *N*-(4-hydroxycinnamoyl)-5-methoxytryptamine or centcyamine (**146**). **SM41** had the same ^{1}H NMR and ^{13}C NMR data as **SM40** except the signals and coupling constant associated with the two olefinic protons (δ 6.62, *J*=12.0 and 5.74, *J*=12.0 Hz). A lower coupling constant (*J*=12.0 Hz) between olefinic protons revealed that **SM41** was the *cis* isomer of *N*-(4-hydroxycinnamoyl)-5-methoxytryptamine or *cis*-centcyamine (**147**). These compounds were previously isolated from *C. cyanus*.[71]

Table 33. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data, ^{1}H-^{1}H COSY and long-range HMBC in CD_3OD for **SM40 (146)**

Carbon number	Chemical shift δ in ppm			HMBC ($^{1}H \rightarrow ^{13}C$)
	$^{1}H^{a}$	$^{13}C^{a}$	COSY (^{1}H-^{1}H)	
2	7.04, s	123.0	–	C-3, C-3a, C-7a
3	–	111.2	–	–
3a	–	128.0	–	–
4	7.2, d, 2.0	111.0	–	C-3a
5	–	148.7	–	
6	6.9, dd, 8.0, 2.0	111.0	–	C-4, C-3, C-3a
7	7.2, d, 2.0	111.2	–	C-3a
7a		137.5	–	–
β-CH_2	2.98, t, 7.2	27.5	–	C-2, C-3, C-3a
α-CH_2	3.57, t, 7.2	40.4	–	C-3, C-9'
1'	–	127.0	–	–
2'	7.38, d, 8.0	129.5	H-2' to H-3'	C-3', C-4'
3'	6.75, d, 8.0	115.0	H-3' to H-2'	C-1'
4'	–	160.0	–	–
5'	6.75, d, 8.0	115.0	H-3' to H-2'	C-1'
6'	7.38, d, 8.0	129.5	H-2' to H-3'	C-2', C-3', C-4'
7'	7.42, d, 16.0	140.7	–	C-9'
8'	6.38, d, 16.0	117.5	–	C-1', C-9'
9'	–	168.0	–	–
O-CH_3 (5)	3.81, s	55.1	–	C-5

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

3.7.1.5 Characterisation of SM42 and SM43 as Moschamine (148) and *cis*-Moschamine (149), respectively

The ^{1}H and ^{13}C NMR of **SM42** and **SM43** **(Table 34)** revealed that they were geometrical isomers. The ^{1}H NMR data showed the presence of two olefinic protons at δ 7.39 (d, J=15.6 Hz) and 6.37 (d, J=15.6 Hz) for **SM42** and 6.57 (d, J=12.0 Hz) and 5.77 (d, J=12.0 Hz) for **SM43**. The higher coupling constant between olefinic protons for **SM42** confirmed that it had *trans* configuration and **SM43** was its *cis* isomer.[214] The ESIMS data **(Table 30)** of **SM42** and **SM43** revealed the

pseudomolecular ion *m/z* 353 $[M+H]^+$ which was calculated for the molecular formula $C_{20}H_{21}N_2O_4$. The ^1H NMR of **SM42** also showed the presence of a tryptamine derivative by showing the signals at δ 7.12 (d, *J*=8.4 Hz), 6.98 (s), 6.92 (d, *J*=2.0 Hz) and 6.63 (dd, *J*=8.4, 2.0 Hz), 3.52 (t, *J*=7.2 Hz) and 2.88 (t, *J*=7.2 Hz).

Table 34. ^1H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM42 (148)**

Carbon number	Chemical shift δ in ppm		HMBC ($^1H \rightarrow ^{13}C$)
	$^1H^a$	$^{13}C^a$	
2	6.98, s	122.0	C-3, C-7a
3	–	111.3	–
3a	–	128.3	–
4	6.92, d, 2.0	102.4	C-3, C-7a
5	–	150.0	–
6	6.63, dd, 8.4, 2.0	110.4	C-7a
7	7.12, d, 8.4	111.2	C-3a, C-5
7a	–	132.0	–
α-CH_2	3.53, t, 7.2	40.3	C-2, C-β, C-9'
β-CH_2	2.88, t, 7.2	25.2	C-2, C-α
1'	–	127.1	–
2'	7.05, d, 2.0	111.5	C-4'
3'	–	148.6	–
4'	–	148.1	–
5'	6.74, d, 8.4	115.3	C-1', C-3'
6'	6.94, dd, 8.4, 2.0	123.1	C-2'
7'	7.39, d, 15.6	140.8	C-9'
8'	6.37, d, 15.6	117.7	C-1', C-9'
9'	–	168.0	–
O-CH_3 (3')	3.81, s	55.2	C-3'

[a] ^1H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

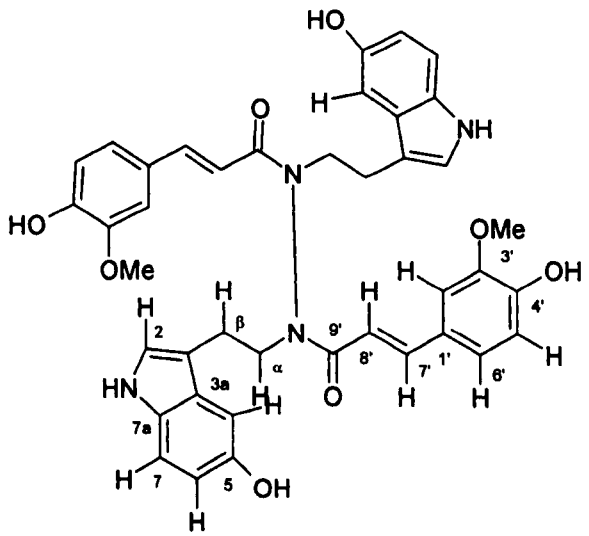
The ^{13}C NMR signal at δ 168.0 indicated the presence of a carbonyl carbon and the signals at δ 7.05 (d, *J*=2.0 Hz), 6.94 (dd, *J*=8.4, 2.0 Hz) and 6.74 (d, *J*=8.2 Hz) in the

^{1}H NMR spectrum (**Figure 125**; **Table 34**) were due to the ABX aromatic ring systems. The ^{1}H and ^{13}C NMR data confirmed that **SM42** had a 3-methoxy-4-hydroxycinnamoyl moiety. The presence of a tryptamine derivative and a 3-methoxy-4-hydroxycinnamoyl moiety were also confirmed by the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC correlations (**Figure 128**; **Table 34**). The ^{1}H-^{1}H COSY revealed four different spin systems:H-7↔H-6↔H-4, $H_2\alpha$↔$H_2\beta$, H-7'↔H-8' and H-6' ↔H-5'. The ^{1}H-^{13}C long-range correlation between $H_2\alpha$ (δ 3.52) to the carbonyl C-9' (δ 168.0) confirmed the existence of an amide linkage between the 3-methoxy-4-hydroxycinnamoyl moiety and serotonin giving the complete structure of **SM42** as (*E*) *N*-(3-methoxy-4-hydroxycinnamoyl)-5-hydroxytryptamine or moschamine (**148**). Thus, the structure **SM43** was elucidated as (*Z*) *N*-(3-methoxy-4-hydroxcinnamoyl)-5-hydroxytryptamine or *cis*-moschamine (**149**). Both of these compounds were isolated previously from *C. cyanus*, *C. moschata* and *C. nigra*.[95,71,105]

3.7.1.6 Characterisation of SM44 as Montamine (193)

The ^{13}C NMR spectrum of **SM44** (**Table 35**) displayed 20 carbons. The DEPT-135 indicated the presence of two methylenes (δ 41.0, 25.0), nine methine (δ 141.0, 121.0, 126.0, 117.5, 115.0, 111.5, 111.2, 111.1, 110.0), seven quarternary (δ 149.0, 148.0, 146.0, 133.0, 129.5, 129.0, 111.0), one carbonyl carbon (δ 172.0) and a methoxy group (δ 55.5). All protonated carbons were assigned by the ^{1}H-^{13}C HSQC experiment (**Figure 130**).[150] These carbon signals and the ^{1}H NMR data (**Figure 129**; **Table 35**) implied that **SM44** was composed of a tryptamine derived substructure and a *p*-coumaroyl moiety like moschamine (**SM42**). This fact was further confirmed by the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC experiments. The ^{1}H-^{1}H COSY spectrum

(**Figure 131**) revealed four different spin systems:H-7↔H-6↔H-4, $H_2\alpha \leftrightarrow H_2\beta$, H-7'↔H-8' and H-H-6' ↔ H-5' and the ^{1}H-^{13}C HMBC spectrum (**Figure 132**) showed key correlations between H-2 to C-3, C-3a and C-7a, H-4 and C-5 and C-7a, H-6 and C-5 and C-7a, H-7 and C-3a and C-5, H-8' and C-1' and C-9'. All spectroscopic data including UV, IR, 1D and 2D NMR suggested that **SM44** possessed a moschamine (**148**) type structure. However, the α-CH_2 and β-CH_2 protons of **SM44** gave rise to different resonance in the ^{1}H NMR spectrum (**Table 36**). For moschamine (**148**) the signals for α-CH_2 and β-CH_2 protons were observed at δ 3.53 (t, J=7.2 Hz) and 2.88 (t, J=7.2 Hz) whereas they were found at δ 2.86, m and 2.27, m respectively for **SM44**.



193

Figure 58. Structure of montamine **(SM44)**

The ESIMS spectrum showed that **SM44** had the molecular mass of 702 instead of 352. The HRESIMS gave $[M+Na]^+$ at 725.2590 (required 725.2587), counted for the molecular formula $C_{40}H_{38}N_4O_8$. The HRESIMS and ^{1}H NMR confirmed that **SM44**

was a dimer of moschamine, named montamine (**193**). Montamine (**193**), isolated from *C. montana,* is a new natural product.

Table 35. ^{1}H NMR (chemical shift, multiplicity, coupling constant *J* in Hz), ^{13}C NMR data and long-range HMBC for **SM44 (193)**

Carbon number	Chemical shift δ in ppm		HMBC ($^{1}H \rightarrow ^{13}C$)	
	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	6.93, s	126.0	C-3	C-3a, C-7a
3	–	111.0	–	–
3a	–	129.5	–	–
4	6.83, d, 2	111.2	C-5	C-7a
5	–	148.0	–	–
6	6.82, dd, 2, 8	111.1	C-5	C-7a
7	7.26, d, 8	111.5	C-7a	C-3a, C-5
7a	–	133.0	–	–
β-CH_2	2.27, m 2.19, m	25.0	C-2, C-α-CH_2	C-3
α-CH_2	2.86, m 2.77, m	41.0	C- β -CH_2	C-3, C-9'
1'		129.0	–	–
2'	6.99, d, 2	110.0	–	C-6, C-4', C-7'
3'	–	146.0	–	–
4'	–	149.0	–	–
5'	6.70, d, 8.4	115.0	–	C-1', C-3'
6'	6.91, dd, 2, 8.4	121.0	–	C-2', C-4',C-7'
7'	7.27, d, 15.6	141.0	–	C-2', C-6' C-9'
8'	6.31, d, 15.6	117.5	C-9'	C-1'
9'	–	172.0	–	–
OCH_3 (3')	3.78, s	55.5	–	C-3'

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

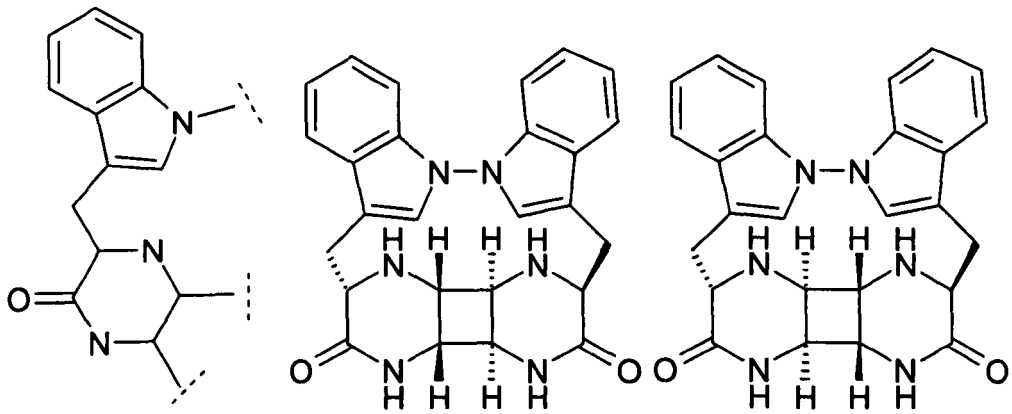
Table 36. Major differences in NMR data of **SM42 (148)** and **SM44 (193)**

Molecule	Group	^{1}H	^{13}C	HMBC ($^{1}H \rightarrow ^{13}C$)	HRESIMS $[M+Na]^+$
Moschamine (**SM42**)	α-CH_2	3.53, t, 7.2	40.3	C-2, C-9'	375.1321
	β-CH_2	2.88, t, 7.2	25.2	C-2	
Montamine(**SM44**)	α-CH_2	2.86, m	41.0	C-3, C-9', C-β-CH_2	725.2590
		2.77, m			
	β-CH_2	2.27, m	25.0	C-2, C-3, C-α-CH_2	
		2.19, m			

3.7.1.7 Characterisation of SM45 as Schischkiniin (194)

The ESIMS spectrum of **SM45** revealed the $[M+Na]^+$ ion peak at *m/z* 475 suggesting *Mr*=452 and the molecular formula was determined as $C_{26}H_{24}N_6O_2$ from its HRESIMS spectrum where the $[M+Na]^+$ ion was observed at *m/z* 475.1857 (calculated 475.1858 for $C_{26}H_{24}N_6O_2Na$). In the ^{1}H NMR spectrum of **SM45** (**Figure 133**; **Table 37**), a singlet at δ 7.15 and the signals at δ 7.65, 7.32, 7.08 and 7.01, were typical of a 3-substituted indole skeleton as described earlier in this chapter and were assigned to H-2, H-4, H-5, H-6 and H-7 by HSQC.[150] Signals at δ 137.2, 127.3, 123.9, 121.5, 118.9, 118.1, 111.2 and 108.4 in the ^{13}C NMR spectrum also fulfilled the requirement for a 3-substituted indole skeleton which was further confirmed by the ^{1}H-^{13}C HMBC. The HMBC showed correlations between H-2 to C-3a, C-7a and C-8; H-4 to C-3, C-6 and C-7a; H-5 to C-3a and C-7; H-6 to C-4 and C-7; H-7 to C-3a, C-5. In addition to the signals attributable to the 3-substituted indole skeleton, the ^{1}H and ^{13}C NMR spectra also showed signals for a methylene (δ_H 3.47 and 3.10, δ_C 27.3), three methines (δ_H 3.80-3.86, δ_C 55.5, 55.4 and 55.5) and an amide

carbonyl (δ_C 174.0). A ^1H-^{13}C long-range HMBC correlation was observed between H-8 (δ_H 3.47) to amide carbonyl (δ_C 174.0). All these signals formed the substructure **194a**.

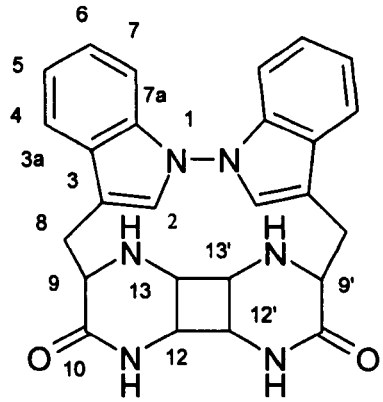


194a **194b** **194c**

Figure 59. Tentative structures of schischkiniin (**SM45**)

Taking the molecular formula and mass into account, it was clear that the ^1H and ^{13}C NMR signals actually displayed signals for just one of the two identical parts of the molecule. Therefore, the molecule must be composed of two of these part structures **194a**. When combining two **194a** structures, only structure **194** could satisfy the molecular formula and molecular mass of this compound. Further evidence to support the structure of **194** was obtained from a series of ^1H-^{13}C long-range couplings observed in its HMBC spectrum (**Table 37**). In the HMBC spectrum, H-12 showed ^2J correlation to C-13 and C-12', and ^3J to C-10 and C-13'. Similarly, H-13 displayed ^2J correlation to C-12 and C-13', and ^3J to C-9 and C-12'. Owing to overlapped ^1H NMR signals for H-9, H-12 and H-13, the ^1H-^1H NOESY was not successful in establishing the relative stereochemistry at C-9 (and C-9'), C-12 (and C-12') and C-13 (and C-13'). Although it is difficult to completely determine the relative stereochemistry of this unique molecule, the use of biogenetic speculation in

tandem with molecular mechanics may give an insight into the stereochemistry of schischkiniin (**194**).



194

Figure 60. Structure of schischkiniin (**SM45**)

Table 37. ^{1}H NMR (chemical shift, multiplicity, coupling constant J in Hz), ^{13}C NMR data and long range HMBC for **SM45 (194)**

Carbon	Chemical shift δ in ppm		HMBC correlation ($^{1}H \rightarrow ^{13}C$)	
Number	$^{1}H^{a}$	$^{13}C^{a}$	^{2}J	^{3}J
2	7.15, s	123.9	C-3	C-3a, C-7a, C-8
3	–	108.4	–	–
3a	–	127.3	–	–
4	7.65, d, 8.2	118.1	C-3a	C-3, C-6, C-7a
5	7.01, dd, 8.2, 8.2	118.9	–	C-3a, C-7
6	7.08, dd, 8.2, 8.2	121.5	C-7	C-4, C-7a
7	7.32, d, 8.2	111.2	–	C-3a, C-5
7a	–	137.2	–	–
8	3.47 dd, 4.4, 15.2	27.3	C-3, C-9	C-2, C-3a, C-10
	3.10 dd, 9.2, 15.2			
9	3.82[b]	55.5	C-8	C-12, C-3
10	–	174.0	–	–
12	3.86[b]	55.4	C-13, C-12'	C-10, C-13'
13	3.84[b]	55.5	C-12, C-13'	C-9, C-12'

[a] ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) in CD_3OD

[b] Overlapped peaks, assigned with the help of ^{1}H-^{13}C HSQC correlation.

Simple diketopiperazines such as the Trp-Gly diketopiperazine (**195**) are common natural products (**Figure 61**). On this occasion we are assuming that the naturally occurring L-Trp has been incorporated into **195**. The next steps would involve the reduction and dehydration of the Gly residue, resulting in the formation of **196**, followed or preceded by the dimerisation at the Trp indole *N* to give the dimer **197**. This dimer would then undergo a photochemically allowed 2+2 cycloaddition to give 4 possible products. Schischkiniin (**194**) is a symmetrical structure as is evident from the degeneracy of the resonances in the NMR spectra. For this reason two possible asymmetric structures can be ruled out leaving only possibilities **194b** and **194c.** Molecular mechanics calculations[216,217] to determine the global energy minima of these possible structures suggest that **194c** has the lowest total energy function 986.93 kcal/mol (**Figure 62**). Although it is possible that a natural product is enzymetically biosynthesised in a high-energy conformation, in this case a photochemically driven cycloaddition reaction will result in the lowest energy product. We therefore speculate that the relative stereochemistry of schischkiniin is as shown in **194c**. If we assume that the origin of the Trp residue is L-Trp then we may also predict the absolute stereochemistry as shown. Thus this novel indole alkaloid was identified as schischkiniin (**194**), which is a tryptophan-derived alkaloid. In addition to the indole skeleton, compound **194** also possesses a distinct macrocyclic polyamine (n = 14) structure.

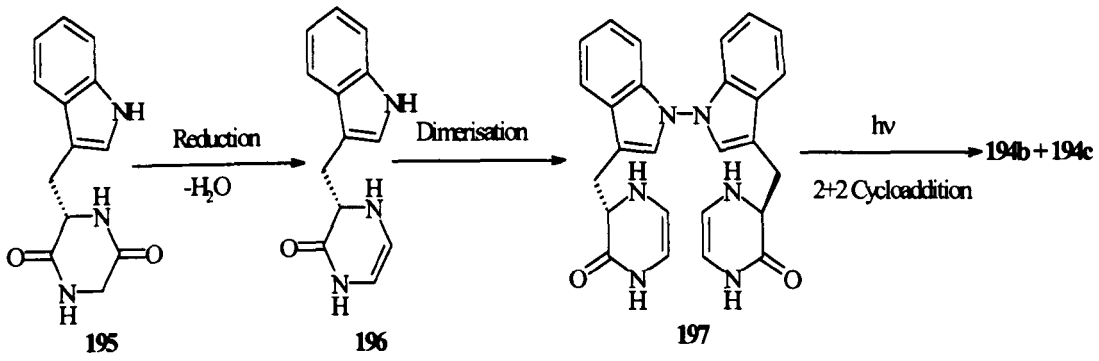


Figure 61. Proposed biogenetic pathway for the formation of schischkiniin (**194**) from a simple diketopiperazine **195**. The reduction/dehydration and dimerisation steps may be reversed

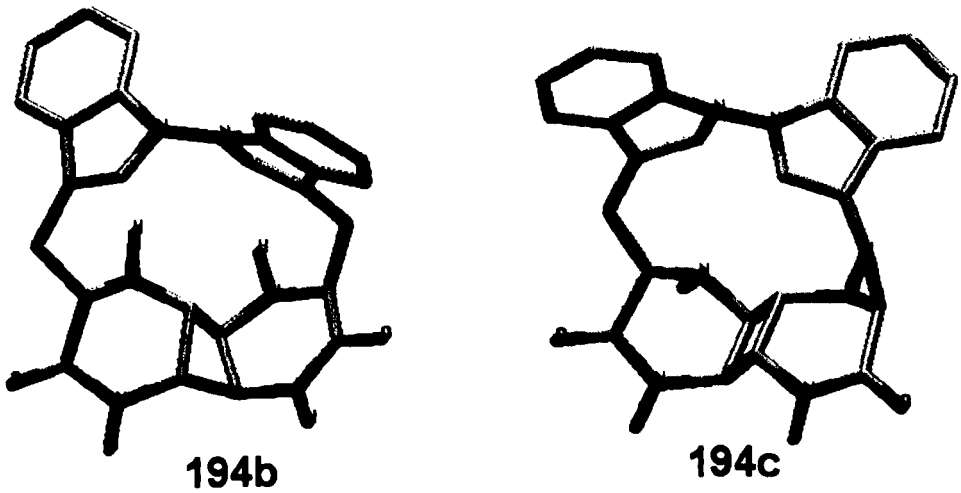


Figure 62. Global energy minima[a] for structures **194b** and **194c** (Heavy atoms and polar hydrogens only shown) The global minimum for **194b** was found 15 times (E_{tot} = 1022.68 kcal/mol) and the global minimum for **194c** was found 37 times (E_{tot} = 986.93 kcal/mol)

[a][**Modelling conditions:** Minimisations (2000) steps were carried out using MacroModel version 6.5[216] using the Merck Molecular Force Field. The generalised Born solvent accessible area continuum solvent model[217] was used to simulate H_2O solvent, due to the unavailability of parameters for MeOH. The minimisations were followed by 1000 steps of Monte Carlo conformational searching to give the global energy minima shown in the figures].

3.7.2 Properties of Indole Alkaloids (SM37-SM45)

3.7.2.1 Properties of Tryptamine, 191 (SM37)

Gum; UV: **Table 29**; LRESIMS: **Table 30**; ^{1}H NMR: **Table 31**; ^{13}C NMR: **Table 31.**

3.7.2.2 Properties of (*E*) *N*-(4-Hydroxycinnamoyl)-5-hydroxytryptamine, 145 (SM38)

Amorphous; UV: **Table 29**; IR λ_{max} (thin film): 3346, 2937, 1653, 1594, 1516, 1456, 1270, 1213, 1125, 1030 cm^{-1}; HRESIMS: **Table 30**; ^{1}H NMR: **Table 32**, ^{13}C NMR: **Table 32.**

3.7.2.3 Properties of (*Z*) *N*-(4-Hydroxycinnamoyl)-5-hydroxytryptamine, 192 (SM39)

Amorphous. UV: **Table 29**; IR λ_{max} (neat): 3346, 2937, 1653, 1594, 1516, 1456, 1270, 1213, 1125, 1030 cm^{-1}; HRESIMS **Table 30**, ^{1}H NMR: **Table 32**; ^{13}C NMR: **Table 32**.

3.7.2.4 Properties of (*E*) *N*-(4-Hydroxycinnamoyl)-5-methoxytryptamine, 146 or Centcyamine (SM40)

Amorphous. UV: **Table 29**; IR λ_{max} (neat): 3433, 3388, 2360, 1651, 1592, 1515, 1463, 1367, 1269, 1212, 1125, 1030 cm^{-1}; HRESIMS: **Table 30**; ^{1}H NMR: **Table 33**; ^{13}C NMR: **Table 33.**

3.7.2.5 Properties of (*E*) *N*-(4-Hydroxycinnamoyl)-5-methoxytryptamine or *cis*-Centcyamine, 147 (SM41)

Amorphous. UV: **Table 29**; IR λ_{max} (neat): 3433, 3388, 2360, 1651, 1592, 1515, 1463, 1367, 1269, 1212, 1125, 1030 cm^{-1}; HRESIMS: **Table 30**; ^{1}H NMR: (major signals) δ 6.62, d, J=12.0 Hz and 5.74, d, J=12.0 Hz; ^{13}C NMR: **Table 33**.

3.7.2.6 Properties of (*E*) *N*-(3-Methoxy-4-hydroxycinnamoyl)-5-hydroxytyptamine or Moschamine, 148 (SM42)

Amorphous. UV: **Table 29**; IR λ_{max} (neat): 3433, 3388, 2360, 1651, 1592, 1515, 1463, 1367, 1269, 1212, 1125, 1030 cm^{-1}; ESIMS: **Table 30**; ^{1}H NMR: **Table 34**; ^{13}C NMR: **Table 34.**

3.7.2.7 Properties of (*Z*) *N*-(3-Methoxy-4-hydroxycinnamoyl)-5-hydroxytyptamine or *cis*-Moschamine, 149 (SM43)

Amorphous. UV: **Table 29**; IR λ_{max} (neat): 3433, 3388, 2360, 1651, 1592, 1515, 1463, 1367, 1269, 1212, 1125, 1030 cm^{-1}; ESIMS: **Table 30**; ^{1}H NMR: (major signals) δ 6.57, d, J=12 Hz and 5.77, d, J=12 Hz; ^{13}C NMR: **Table 34.**

3.7.2.8 Properties of Montamine, 193 (SM44)

Gum; UV: **Table 29**; HRESIMS: **Table 30**; ^{1}H NMR: **Table 35;** ^{13}C NMR: **Table 35.**

3.7.2.9 Properties of Schischkiniin, 194 (SM45)

Gum; UV: **Table 29**; ESIMS: 475 $[M+Na]^{+}$, 430, 413, 393, 354; HRESIMS: **Table 30**; ^{1}H NMR: **Table 37**, ^{13}C NMR: **Table 37.**

CHAPTER FOUR

Biological Activities of Secondary Metabolites from *Centaurea* Species

4. 1 Introduction

4.1.1 Background

The chapters 1-3 have dealt with the description and collection of *Centaurea* species, followed by the isolation and characterisation of compounds from twelve *Centaurea* species. In this chapter the biological activities of the extracts and the characterised compounds from all these *Centaurea* species are discussed. The brine-shrimp lethality assay[218], 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide (MTT) cytotoxicity assay[219] and 2,2-diphenyl-1-picryl hydrazyl (DPPH) antioxidant assay[220] were carried out for evaluating, respectively, general toxicity, cytotoxicity and antioxidant activity of extracts and pure compounds in this study.

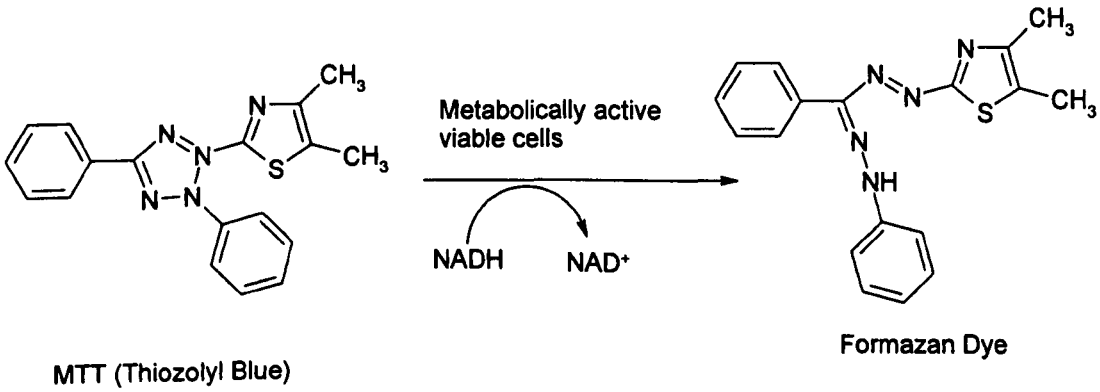
4.1.2 Brine Shrimp Lethality Assay

The brine shrimp lethality assay was performed to evaluate the general toxicity of extracts and pure compounds.[221] Ferrigni *et. al.* screened the ethanol extracts from the seeds of 41 Euphorbiaceae species for toxicity and cytotoxicity in brine shrimp assay and potato disc assay, respectively, and found a positive correlation between them.[222] It has been now established that cytotoxic compound usually shows good activity in brine shrimp assay.[223,224] This assay is recommended as a guide for the detection of antitumour and pesticidal compounds.[225] The brine shrimp assay is a simple, inexpensive and easy to perform for rapid screening of natural products.[218] The eggs of brine shrimp, *Artemia salina* Leach, are readily available at a low cost in pet shops. The eggs hatch within 48 hours, providing large numbers of larvae (nauplii) upon being placed in a brine solution under illumination, and test compounds or plant extracts of desired concentration can be added in vials containing 14 to 16 nauplii.[226] After 24 hours the number of nauplii are counted and

the percentage of deaths at each dose are recorded using the LDPLINE software program.[227] A LD_{50}, defined as the concentration of compound which causes 50% mortality of nauplii, is determined using Probit analysis described by Finney.[228]

4.1.3 MTT Cytotoxicity Assay

Cytotoxicity means toxicity to cells and is measured as functions of fundamental biochemical pathways leading to cell death. The discovery and development of cytotoxic agents for cancer therapy involves the systematic screening of large number of extracts and compounds. The screening of natural products for their cytotoxicity requires methods that are easy, reliable, rapid, sensitive and economic. The most commonly used cytotoxicity assays includes ATP measurements[229], MTT assay[219], neutral red[230], membrane integrity/LDH release assay[231], alamar blue assay[232] and sulforhodamine B (SRB) assay[233].



Scheme 5. Transformation of MTT by viable cells

The MTT assay is a well-documented cell viability assay and has been modified by several investigators since it was first developed by Mosmann.[219,234] This assay is based on the transformation of tetrazolium salt, 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide (MTT) by mitochondrial succinic dehydrogenases in viable cells yielding purple formazan crystals that are not soluble in aqueous solution (**Scheme 5**).

The amount of formazan generated by dehydrogenase enzyme is directly proportional to the number of viable cells in culture and is measured at 560 nm. In cytotoxicity screening, the most commonly determined parameter is the IC_{50}, which is defined as the drug concentrations required to reduce the absorbance by 50% of the control values. This can easily be calculated from percentage of cell viability against concentration curve as shown in **Figure 63.** In this project, the MTT assay was used to evaluate cytotoxicity of the extracts and compounds from *Centaurea* species on human colon cancer cell line (CaCo-2).

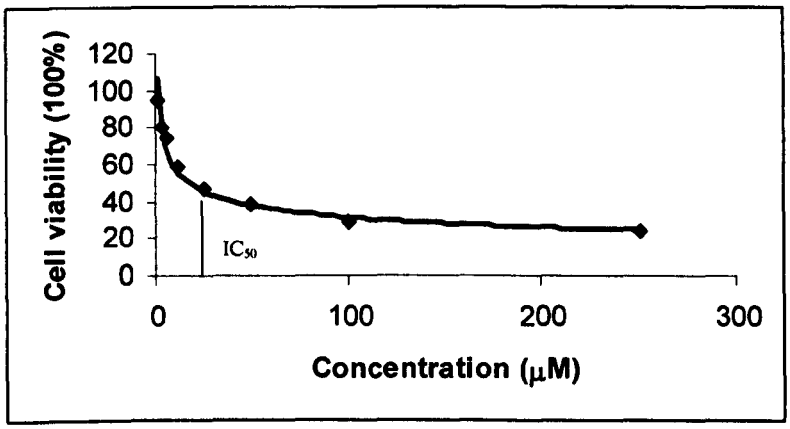
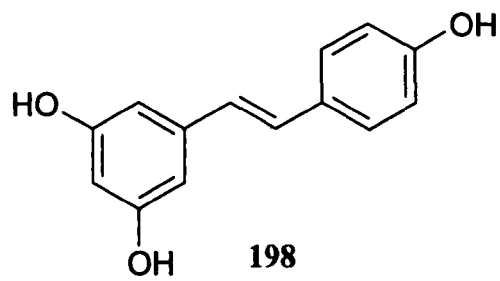


Figure 63. Cell viability (%) against drug concentration curve for determination of IC_{50} value by MTT assay

4.1.4 DPPH Antioxidant Assay

Antioxidants are important species, which have the ability to protect body from damage caused by free-radical induced oxidative stress.[235] Oxygen free radicals or reactive species (ROS) such as superoxide anion radicals ($O_2^{\cdot}$), hydrogen peroxides (H_2O_2), hydroxyl radicals (OH) and singlet oxygen ($^1O_2$) are continuously generated in cells exposed to an aerobic environment.[236] The role of reactive oxygen (ROS) has been implicated in many human degenerative diseases, including aging, cancer, and neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease and Huntington's disease.[237, 238] It is reported that hydrogen peroxide (H_2O_2) can cause

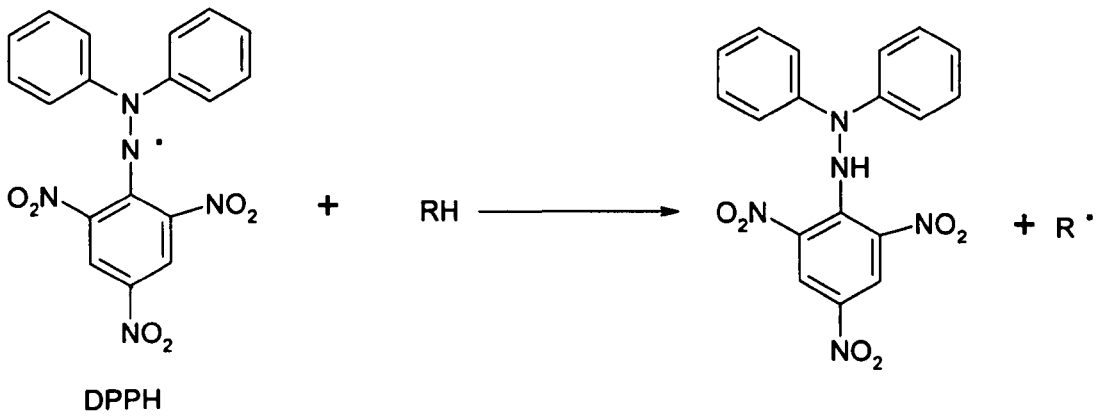
lipid peroxidation and DNA damage.[239,240] Antioxidant defence systems have co-evolved with aerobic metabolism to counteract oxidative damage from reactive oxygen species.[241] Furthermore many antioxidant compounds have shown anticancer properties.[242,243] For example, resveratrol (**198**), a phenolic natural product from grapes was reported to protect cells from oxidative damage and cell death.[244, 245] Resveratrol (**198**) was also reported to prevent carcinogenesis in a murine model.[246] One of the rapid, simple and inexpensive methods to measure antioxidant capacity of natural products involves the use of the free radical, 2,2-diphenyl-1-picryl hydrazyl (DPPH). The DPPH assay was performed to evaluate the antioxidant properties of the extracts and compounds from *Centaurea* species.[220, 247]



198

Figure 64. Structure of resveratrol (**198**)

The DPPH antioxidant assay is used to test the ability of compounds to act as free radical scavengers or hydrogen donors, and to evaluate antioxidant activity of natural products and synthetic compounds. The potential antioxidant activity of plant extracts is based on scavenging stable 2,2-diphenyl-1-picryl hydrazyl (DPPH) free radicals. Antioxidants can react with the stable free radical DPPH and produce 2,2-diphenyl-1-picryl hydrazine (**Scheme 6**).



Scheme 6. DPPH reduction

Due to its unpaired electron, 2, 2-diphenyl-2-picryl-hydrazyl radical (DPPH) gives strong absorption at 517 nm and is purple in colour. In the presence of a free radical scavenger electron becomes paired off and the absorption is quenched. The resulting decolourisation is stoichiometric with respect to the number of electrons being taken up. The change of absorbance produced in this reaction is assessed to evaluate the antioxidant potential of test samples. An IC_{50}, which is defined as the concentration of compound that causes a 50% reduction potential in absorbance compared to control can be calculated as shown in **Figure 65.**

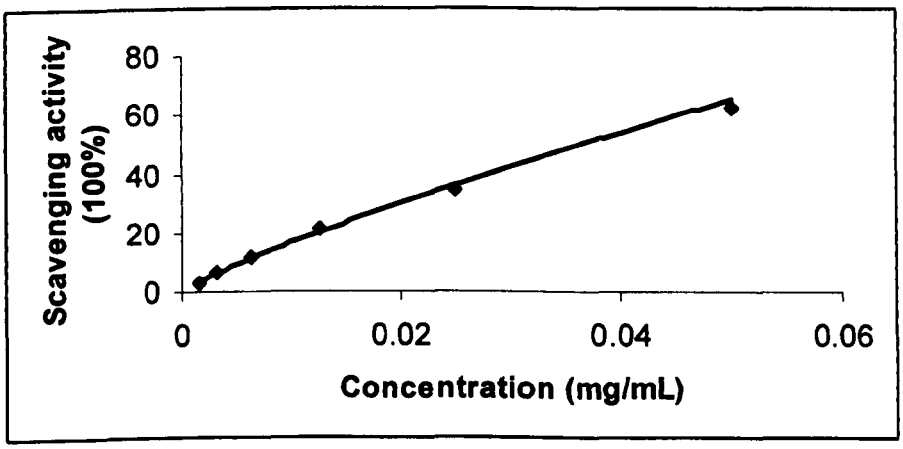


Figure 65. Scavenging activity (%) against drug concentration curve to determine the IC_{50} value by DPPH assay

4.2 Materials and Methods

4.2.1 Brine Shrimp Lethality Assay

4.2.1.1 Preparation of Brine Shrimp Medium (BSM)

Brine shrimp medium (BSM) was prepared by dissolving the material in **Table 38** in 1.25 L of distilled water. Each chemical was dissolved separately in the order shown in **Table 38** and the disodium glycerophosphate was added last to prevent the formations of any insoluble precipitate forming. The stock solution was stored in brown glass bottle in the refrigerator (-4°C).

Table 38. Composition of brine shrimp medium

Chemicals	Amount (g)
Sodium chloride	300
Calcium chloride dihydrate	3
Magnesium chloride	15
Magnesium sulphate	5
Potassium chloride	8
Glycine	60
Disodium glycerophsphate	30

4.2.1.2 Bioassay of Sample with BSM

BSM stock solution (20 mL) was mixed with distilled water (140 mL) in a 250 mL conical flask to get a working solution. Brine shrimp eggs (100 mg) were added in the solution and incubated at 25° C for 48 hours. During this time the brine shrimp eggs hatched completely. Hatched larvae were separated from unhatched eggs and were transferred into a Petri dish. A desk lamp was positioned one side so that the larvae could get light as they concentrate near the light. Dried methanol extract

(10 mg) or pure compound (1 mg) was dissolved in 1 mL of 20% aqueous DMSO solution and poured into 9.0 mL of working solution in a universal bottle. The solution (5.0 mL) was taken from the latter and added into another universal bottle containing 5.0 mL of working solution. In a similar way, serial dilutions were performed to obtain 7 different concentrations (100, 50, 25, 12.5, 6.25, 3.12, 1.56 μg/mL). About 14 to 16 nauplii were added to each bottle and all universal bottles were kept in a water bath at 25°C for 24 hours. After 24 hours the number of nauplii was counted with a magnifying glass and the percentage mortality was calculated based on 3 replicates. An LD_{50} value (i.e. the concentration of cell extract which causes 50% mortality) for each sample was determined as described in **Section 4.1.2**. Podophyllotoxin (**10**, page 200) was used as a positive control in this experiment.

4.2.2 MTT Cytotoxicity Assay

4.2.2.1 Materials and Chemicals

CaCo-2, human colon cancer cell line was obtained from the European collection of cell cultures. Earle's minimum essential medium, Earle's balanced salt solution, trypsin, L-glutamine, non essential amino acids, penicillin and streptomycin were purchased from Sigma, UK. Foetal calf serum (FCS) was purchased from Biosera, UK. 96 well plates were bought from Sero-Wel, Bibbly Sterilin Ltd., Stones, Staffs, UK.

4.2.2.2 Cryopreservation

Cells were cryopreserved to ensure the continuous supply of a cell line during this study. Healthy and viable cells 4-10 x 10^6 per ampoule were used for this purpose. Freezing medium was prepared from the growth medium containing 10% (v/v) sterile dimethyl sulphoxide. Cell pellets dissolved in 1 mL freezing medium were

poured into a sterile plastic screw-top cryotubes and immediately cooled at a rate of between 1-5°C per min. Normally the cryotubes were placed in the -20°C freezer for 3-4 hours and transferred to the -80°C for approximately 16 hours (overnight). After freezing, the ampoules were transferred to liquid nitrogen.

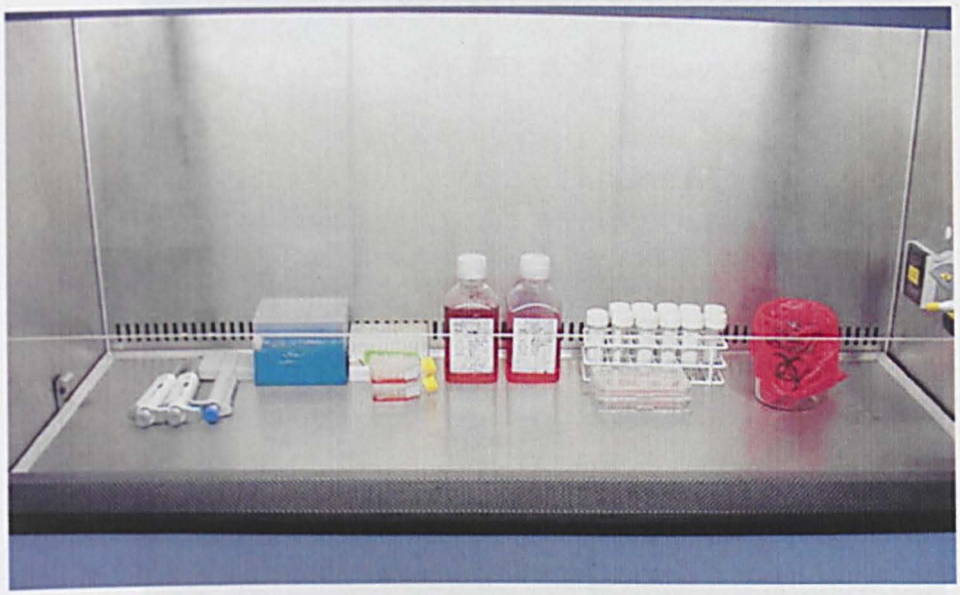


Figure 66. Laminar Flow hood

4.2.2.3 Thawing of Frozen Cells

The ampoule was removed from storage and unscrewed the cap ¼ turn to release any residual nitrogen. After 1-2 minutes (to let gas escape) it was placed in the water bath at 37°C. Special care was taken in order to prevent water entering the ampoule and contaminating the cells. When ampoules were fully thawed, they were taken to the laminar flow hood (**Figure 66**). Then the ampoule contents were transferred into a 30 mL sterile universal. The growth medium (10 mL) were added slowly to the universal, mixed and centrifuged at the lowest speed to pellet the cells at 70-100 × g. After decanting the supernatant, the pelleted cells were resuspended in fresh medium and transferred to a culture flask.

4.2.2.4 Cell Counting

Cells were counted by haemocytometer (**Figure 67**) and particle size distributor (**Figure 68**). Similar results were obtained in both cases and constant cell density was maintained for all cell culture work in this study.

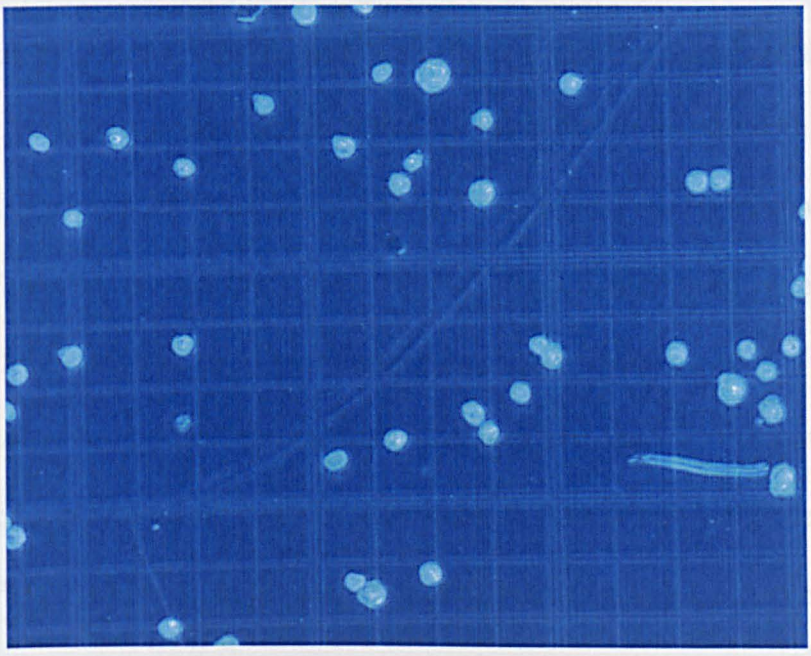


Figure 67. Haemocytometer

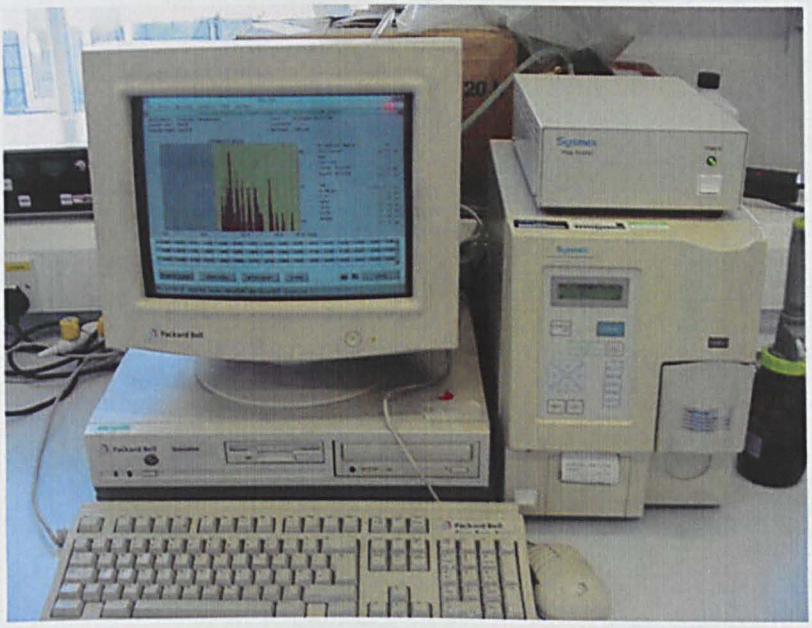


Figure 68. Particle size distributor

4.2.2.5 Cell Viability Assay

CaCo-2 cells (**Figure 69**) were maintained in Earle's minimum essential medium (Sigma), supplemented with 10% (v/v) foetal calf serum, 2 mM L-glutamine, 1% (v/v) non-essential amino acids, 100 IU/mL penicillin and 100 μg/mL streptomycin. Exponentially growing cells were seeded on 96-well plates at 1×10^4 cells per well and incubated for 24 hours before the addition of drugs. Stock solution of the test compounds was initially dissolved in DMSO or H_2O and further diluted with serum free medium. Following a 24 hours incubation at 37°C, 5% CO_2, 100 μl of various concentration of test compounds (500, 250, 100, 50, 25, 10, 5 and 1 μM) were added in each well in triplicates and cells were further incubated for 72 hours.[248,249] After 72 hours of incubation at 37°C, the medium was removed, and 100 μL of MTT reagent (1 mg/mL) in serum free medium was added to each well. The plates were incubated at 37°C for 4 hours. At the end of the incubation period, the supernatants were removed and, pure DMSO (200 μL) was added to each well and plates were

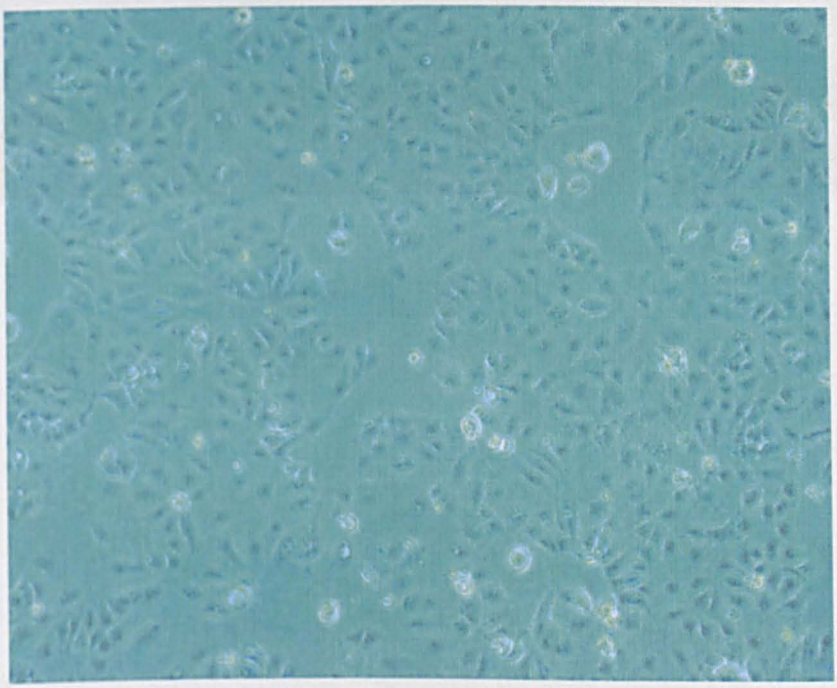


Figure 69. Untreated CaCo-2 cell

shaken gently for 15 minutes. The metabolised MTT product dissolved in DMSO was quantified by reading the absorbance at 560 nm using a micro plate reader (Dynex Technologies, USA, **Figure 70**). IC_{50} values are defined as the drug concentrations required to reduce the absorbance by 50% of the control values. They are calculated from the equation of the logarithmic line determined by fitting the best line (Microsoft Excel) to the curve formed from the data. The IC_{50} value was obtained from the equation y = 50 (50% value).

Figure 70. Plate reader

4.2.3 DPPH antioxidant assay

4.2.3.1 Chemicals and materials

2,2-Diphenyl-1-picrylhydrazyl (DPPH), molecular formula $C_{18}H_{12}N_5O_6$, was obtained from Fluka Chemie AG, Bucks. Quercetin was purchased from Avocado Research Chemicals Ltd, Shore road, Heysham, Lancs.

4.2.3.2 Qualitative Assay

Test compounds of lignans, flavonoids and alkaloids were applied on a TLC plate and sprayed with DPPH solution using an atomiser. It was allowed to develop for 30 min. The colour change (purple on white) was noted. Positive colour change indicated the presence of antioxidants.

4.2.3.3 Quantitative assay

DPPH (4 mg) was dissolved in MeOH (50 mL) to obtain a concentration of 80 μg/mL. Test compounds were also dissolved in MeOH to obtain initial concentration of 0.5 mg/mL. Serial dilutions were made up to eight different concentrations (5×10^{-2}, 2.5×10^{-2}, 1.25×10^{-2}, 6.25×10^{-3}, 3.12×10^{-3} 1.56×10^{-3}, 7.8×10^{-4} and 3.9×10^{-4} mg/mL) with a volume of 5.0 mL in each case. Each solution (5.0 mL) was mixed with DPPH solution (5.0 mL) and allowed to stand for 30 min. The UV absorbance was recorded at 517 nm. The experiment was performed in triplicate and the average absorption was determined for each concentration. The same procedure was repeated for the positive control, quercetin.

4.3 Results and Discussion

4.3.1 Brine Shrimp Lethality Assay

4.3.1.1 Brine Shrimp Lethality Assay of Centaurea Extracts

Brine shrimp lethality assay was performed to evaluate general toxicity of MeOH extracts of twelve *Centaurea* species, and the results are summarised in **Table 39**. The *C. bornmuelleri*, *C. huber-morathii*, *C. gigantea* and *C. montana* extracts were the most toxic among the extracts tested, and the LD_{50} values for them were found to be 55.2×10^{-2}, 42.4×10^{-2}, 69.2×10^{-2} and 62.5×10^{-2} mg/mL, respectively.

Table 39. Brine shrimp lethality assay of the extracts of *Centaurea* species

Methanol extracts	LD_{50}[a] (mg/mL)
C. americana	89.3×10^{-2}
C. bornmuelleri	55.2×10^{-2}
C. cyanus	71.2×10^{-2}
C. dealbata	110.0×10^{-2}
C. gigantea	69.2×10^{-2}
C. huber-morathii	42.4×10^{-2}
C. macrocephala	85.1×10^{-2}
C. montana	62.5×10^{-2}
C. mucronifera	120.0×10^{-2}
C. pamphylica	125.0×10^{-2}
C. schischkinii	76.4×10^{-2}
C. urvillei	115.5×10^{-2}
Podophyllotoxin (**10**)	2.8×10^{-3}

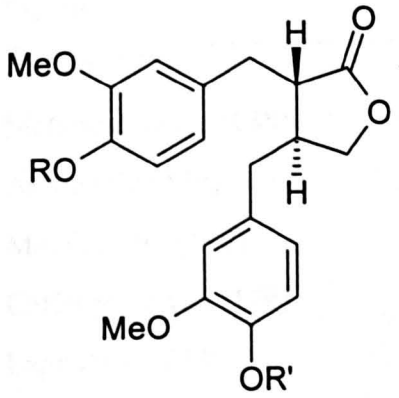
[a]most toxic:1-70 $\times 10^{-2}$ mg/mL; moderate toxic:70-100 $\times 10^{-2}$ mg/mL; non toxic: above 100 $\times 10^{-2}$ mg/mL

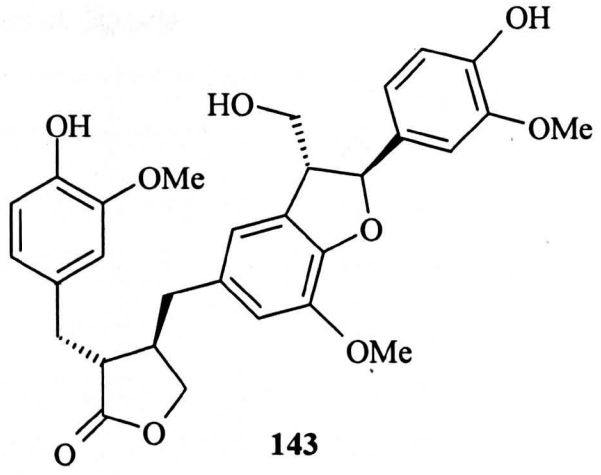
While the extracts of *C. americana*, *C. cyanus*, *C. macrocephala* and *C. schischkinii* showed moderate toxicity by displaying their LD_{50} values between 70-100 $\times 10^{-2}$ mg/mL, the extracts of *C. dealbata*, *C. mucronifera*, *C. pamphylica* and *C. urvillei*

were the least toxic having their LD_{50} values above 100 × 10^{-2} mg/mL. Podophyllotoxin (**10**, page 201), a cytotoxic lignan[250], was used as the positive control in this study and the LD_{50} value for podophyllotoxin was found to be 2.8 × 10^{-3} mg/mL.

4.3.1.2 Brine Shrimp Lethality Assay of Lignans

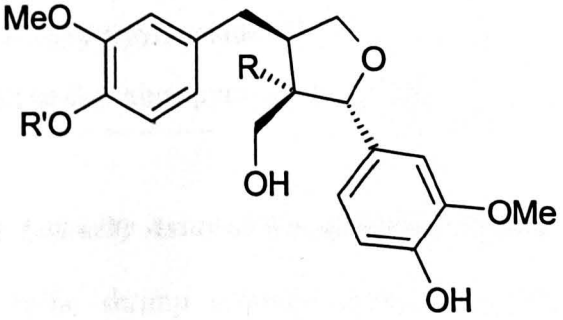
The brine shrimp lethality assay of lignans is presented in **Table 40**. Lariciresinol type lignans, lariciresinol 4'-*O*-glucoside (**171**), berchemol (**172**) and berchemol 4'-*O*-glucoside (**173**) showed moderate level of activity and their LD_{50} values were found to be, respectively, 4.5 × 10^{-2}, 3.1 × 10^{-2} and 10.3 × 10^{-2} mg/mL. Dibenzylbutyrolactone type lignans (**137-140**) showed significantly higher activity than that of lariciresinol type in this assay. For example, arctigenin (**139**), matairesinol (**140**) and lappaol A (**143**) exhibited LD_{50} values of 2.0 × 10^{-3}, 5.5 × 10^{-3} and 9.2 × 10^{-3} mg/mL, respectively. Arctiin (**137**; LD_{50}= 9.80 × 10^{-2} mg/mL) and matairesinoside (**138**; LD_{50}= 1.65 × 10^{-2} mg/mL) were moderately active in the brine shrimp lethality assay. Epoxylignans of pinoresinol-type (**144,176,178**) were found to be the least toxic among the lignans tested in this study. For example, LD_{50} values for pinoresinol (**144**), pinoresinol monoglucosde (**176**), pinoresinol di-glucoside (**177**) and pinoresinol-apiose-glucoside (**178**) were found to be 6.5 × 10^{-2}, 6.8 × 10^{-2}, 7.2 × 10^{-2} and 8.3 × 10^{-2} mg/mL, respectively.



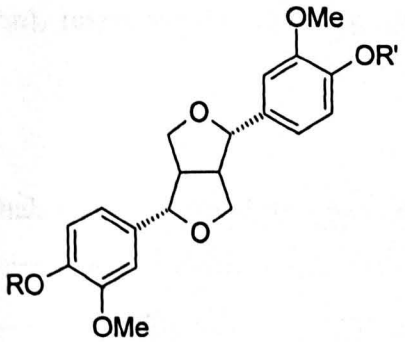


143

	R	R'
137	β-D-glucopyranosyl	Me
138	β-D-glucopyranosyl	H
139	H	Me
140	H	H
170	3-*O*-caffeoyl- β-D-glucopyranosyl	Me



	R	R'
171	H	β-D-glucopyranosyl
172	OH	H
173	OH	β-D-glucopyranosyl



	R	R'
144	H	H
176	β-D-glucopyranosyl	H
177	β-D-glucopyranosyl	β-D-glucopyranosyl
178	apiose-β-D-glucopyranosyl	H

Figure 71. Structures of lignans tested in the bioassay

Table 40. Brine shrimp lethality assay of lignans

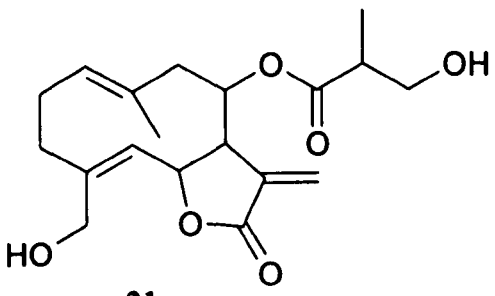
Lignans	LD_{50} (mg/mL)
Arctiin (**137**)	9.8×10^{-2}
Matairesinoside (**138**)	1.6×10^{-2}
Arctigenin (**139**)	2.0×10^{-3}
Matairesinol (**140**)	5.5×10^{-3}
Caffeoyl-arctiin (**170**)	7.6×10^{-2}
Lappaol A (**143**)	9.2×10^{-3}
Lariciresinol 4'-*O*-β-D-glucoside (**171**)	4.5×10^{-2}
Berchemol (**172**)	3.1×10^{-2}
Berchemol 4'-*O*-β-D-glucoside (**173**)	10.3×10^{-2}
Pinoresinol (**144**)	6.5×10^{-2}
Pinoresinol 4-*O*-β-D-glucopyranoside (**176**)	6.8×10^{-2}
Pinoresinol 4,4'-di -*O*-D-glucopyranoside (**177**)	7.2×10^{-2}
Pinoresinol 4-*O*-β-apiose-β-D-glucopyranoside (**178**)	8.3×10^{-2}

4.3.1.3 Brine Shrimp Lethality Assay of Sesquiterpene lactones

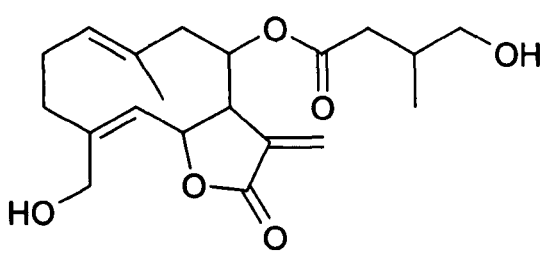
The results of the brine shrimp lethality assay of sesquiterpene lactones are summarised in **Table 41.** Both arctiopicrin (**21**) and 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**179**) showed significant toxicity (LD_{50} values of 4.7×10^{-3} and 5.8×10^{-3} mg/mL respectively), and were more toxic than the lignans tested (**137-140**).

Table 41. Brine shrimp lethality assay of sesquiterpene lactones

Sesquiterpe lactones	LD_{50} (mg/mL)
Arctiopicrin (**21**)	4.7×10^{-3}
8-*O*-(4-Hydroxy-3-methylbutanoyl)-salonitenolide (**179**)	5.8×10^{-3}



21

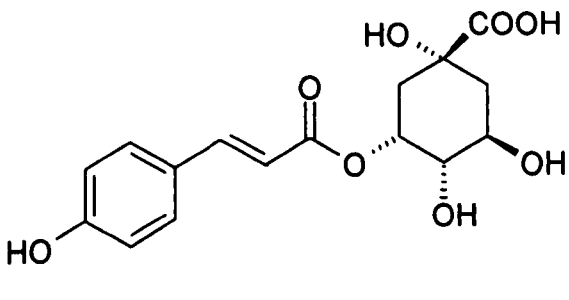


179

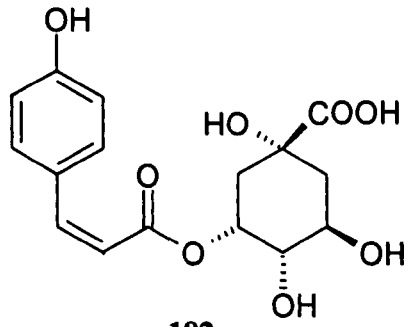
Figure 72. Structures of sesquiterpene lactones tested in the bioassay

4.3.1.4 Brine Shrimp Lethality Assay of Quinic acid Derivatives

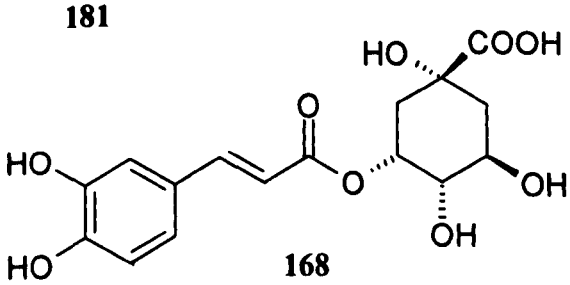
The quinic acid derivatives, isolated from the *Centaurea* species exhibited moderate levels of activity in the brine shrimp lethality assay (**Table 42**). Chlorogenic acid (**168**; LD_{50}=2.5 × 10^{-2} mg/mL) was more toxic compound than its other derivative, *p*-coumaroyl quinic acid (**181**; LD_{50}=7.8 × 10^{-2} mg/mL).



181



182



168

Figure 73. Structures of quinic acid derivatives in the bioassay

Table 42. Brine shrimp lethality assay of quinic acid derivatives

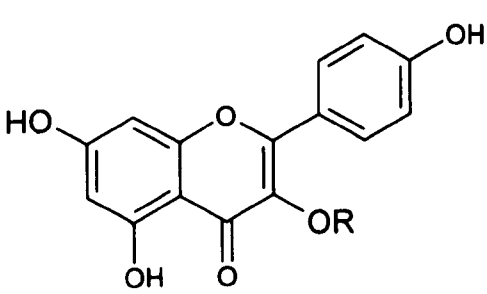
Compounds	LD_{50} (mg/mL)
p-Coumaroylquinic acid (**181**)	7.8×10^{-2}
cis-p-Coumaroylquinic acid (**182**)	10.3×10^{-2}
Chlorogenic acid (**168**)	2.5×10^{-2}

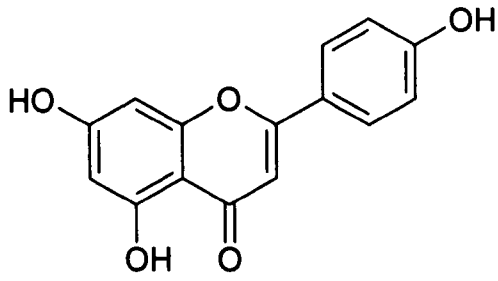
4.3.1.5 Brine Shrimp Lethality Assay of Flavonoids

Among the *Centaurea* flavonoids, tested in this study, kaempferol derivatives, astragalin (**123**) and afzelin (**187**) showed no significant activity with the LD_{50} values of 1.4×10^{-1} and 8.0×10^{-1} mg/mL respectively (**Table 43**). However, apigenin (**108**), cirsiliol (**114**) and flavanone apiose-glucuronic acid (**190**) showed significant toxicity with the LD_{50} values of 9.3×10^{-3}, 6.4×10^{-3} and 7.2×10^{-3} mg/mL, respectively.

Table 43. Brine shrimp lethality assay of flavonoids

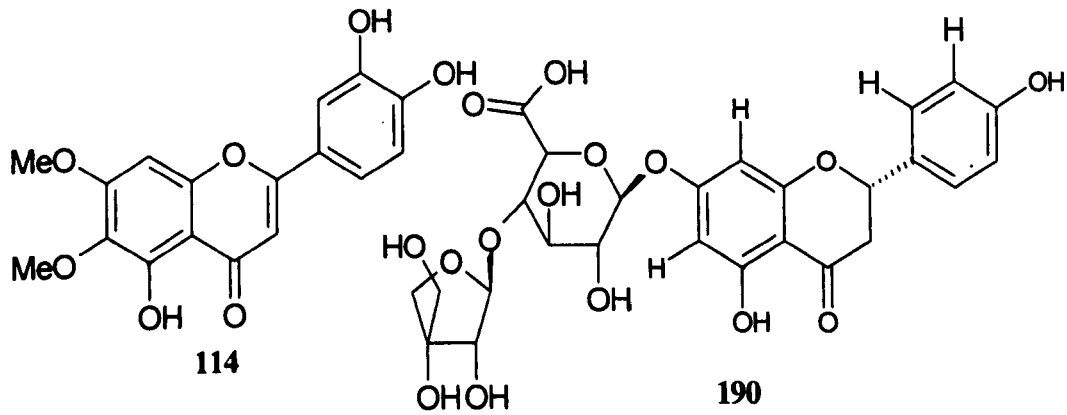
Flavonoids	LD_{50} (mg/mL)
Astragalin (**123**)	1.4×10^{-1}
Afzelin (**187**)	8.0×10^{-1}
Apigenin (**108**)	9.3×10^{-3}
Cirsiliol (**114**)	6.4×10^{-3}
Flavanone-apiose-glucuronic acid (**190**)	7.2×10^{-3}





108

	R
122	H
123	3-*O*-β-D-glucopyranosyl
187	3-*O*-α-D-rhamnosyl

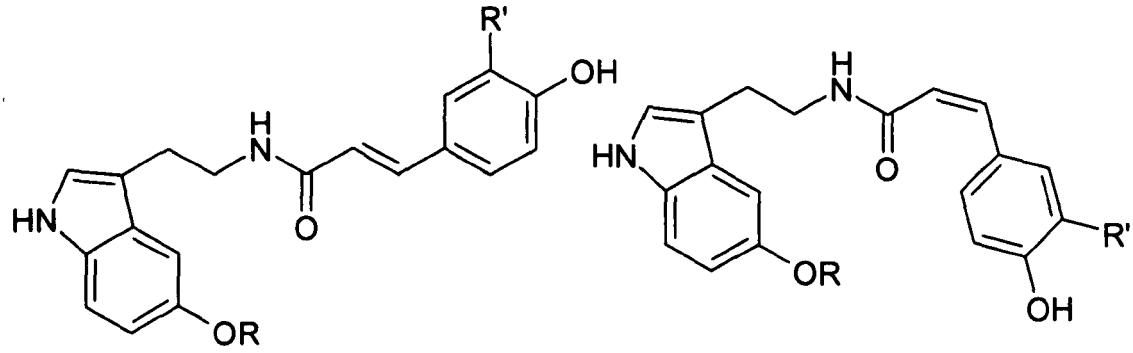


114 **190**

Figure 74. Structures of flavonoids

4.3.1.6 Brine Shrimp Lethality Assay of Alkaloids

Four alkaloids (**146, 148, 193** and **194**) were tested in the brine shrimp assay and the results are summarised in **Table 4.7**. It was observed that all tested alkaloids showed significant toxicities in this assay with LD_{50} values for centcyamine (**146**), moschamine (**148**), montamine (**193**) and schischkiniin (**194**), respectively of 15×10^{-3}, 20×10^{-3}, 3.5×10^{-3} and 7.2×10^{-3} mg/mL. Montamine (**193**), a dimer of moschamine (**148**) was the most toxic among these alkaloids.



Alkaloids	R	R'	Alkaloids	R	R'
145	H	H	**192**	H	H
146	Me	H	**147**	Me	H
148	H	OMe	**149**	H	OMe

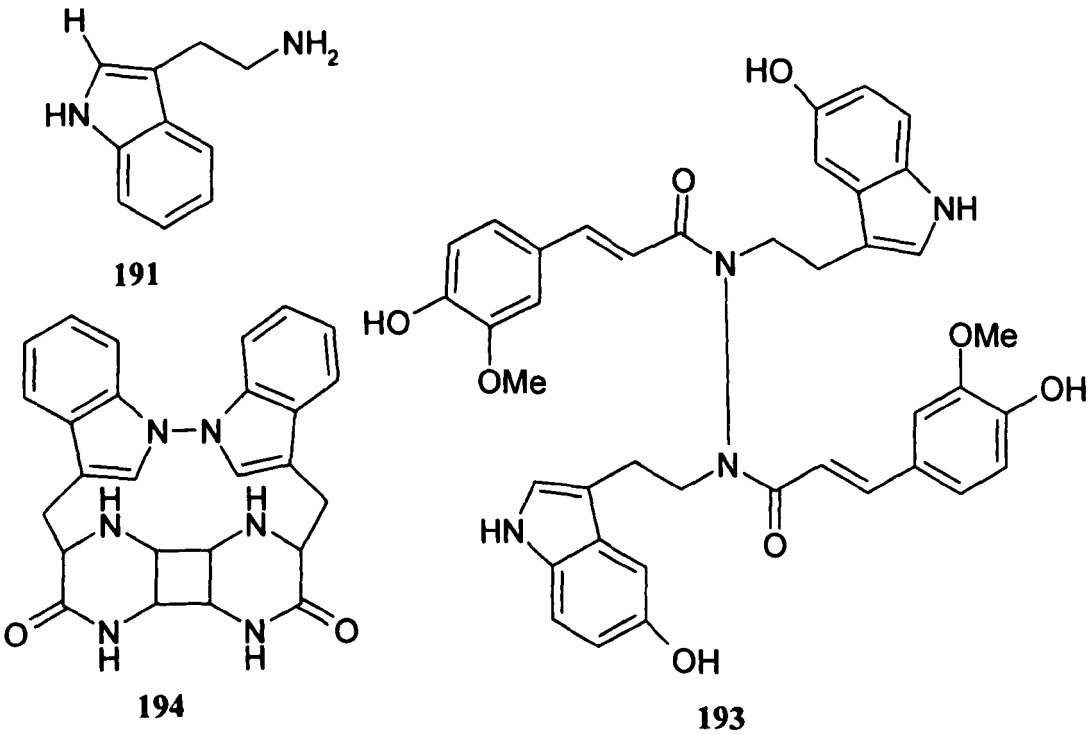
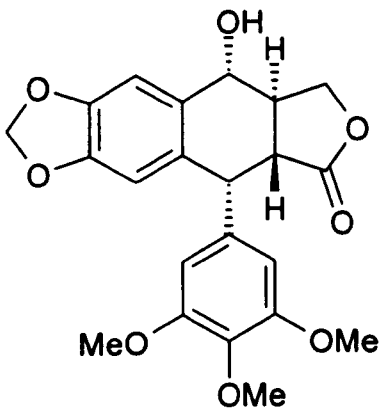


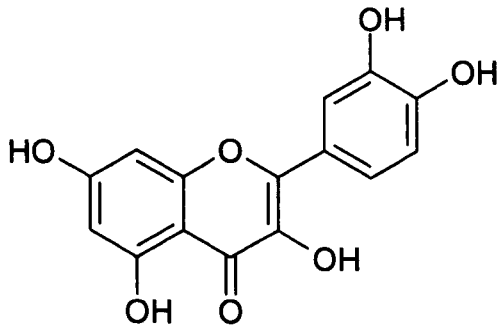
Figure 75. Structures of alkaloids tested in the bioassay

Table 44. Brine shrimp lethality assay of alkaloids

Alkaloids	LD_{50} (mg/mL)
Centcyamine (**146**)	15×10^{-3}
Moschamine (**148**)	20×10^{-3}
Montamine (**193**)	3.5×10^{-3}
Schischkiniin (**194**)	7.2×10^{-3}



10



125

Figure 76. Structures of podophyllotoxin (**10**) and quercetin (**125**)

4.3.2. *In vitro* Cytotoxicity Studies against Colon Cancer Cell (CaCo-2)

4.3.2.1 MTT Cytotoxicity Assay of Centaurea Extracts

The MeOH extracts of twelve *Centaurea* species were tested for cytotoxicity against a human colon cancer cell line (CaCo-2) using the MTT assay and the results are summarised in **Table 45**. The cytotoxic lignan podophyllotoxin[250] (**10**) was used as positive control with an IC_{50} value of 0.06 μM in this study. Among the extracts, *C. bornmuelleri*, *C. huber-morathii*, *C. gigantea* and *C. montana* showed significant cytotoxicity with the IC_{50} values of 29.8, 33.0, 43.2 and 56.4 μg/ml respectively.

Table 45. Cytotoxicity activities of *Centaurea* species

Methanol extracts	IC_{50}[a] (μg/mL)
C. americana	77.5
C. bornmuelleri	29.8
C. cyanus	69.3
C. dealbata	>100
C. gigantea	43.2
C. huber-morathii	33.0
C. macrocephala	86.2
C. montana	56.4
C. mucronifera	>100
C. pamphylica	>100
C. schischkinii	68.8
C. urvillei	>100
Podophyllotoxins (**10**)	0.06 (μM)

[a]Significantly cytotoxic:1-70 $\times 10^{-2}$ mg/mL; moderate cytotoxic:70-100 $\times 10^{-2}$ mg/mL; non cytotoxic: above 100 $\times 10^{-2}$ mg/mL

On the other hand, *C. americana*, *C. cyanus*, *C. macrocephala* and *C. schischkinii* showed moderate levels of cytotoxicity with IC_{50} values between 68-78 μg/mL. The extracts from *C. dealbata*, *C. mucronifera*, *C. pamphylica* and *C. urvillei* were the least active with IC_{50} values to be above 100 μg/mL.

4.3.2.2 MTT Cytotoxicity Assay of Lignans

The cytotoxicity results of twelve lignans against a human colon cancer cell line (CaCo-2) by MTT assay are summarised in **Table 46**. Dibenzylbutyrolignans, **137-140** were more cytotoxic than the other lignans. The IC_{50} values for arctiin (**137**), matairesinoside (**138**), arctigenin (**139**), matairesinol (**140**) and lappaol A (**143**) were 220.0, 288.0, 7.0, 124.0 and 135.0 μM respectively (**Table 46**). Both arctiin (**137**) and matairesinoside (**138**) contain glucose moiety and the cytotoxicity of glycosides decreased significantly. Arctigenin (**139**), matairesinol (**140**) and lappol A (**143**) do not have any glucose moiety and they exhibited higher cytotoxicity. Arctigenin (**139**) contained a $-OCH_3$ group in C-4' and was less polar than matairesinol (-OH group in C-4') and showed potent cytotoxicity in colon cancer cell line among all lignans tested in this study. These results suggest that the presence of appropriate hydrophobicity of lignan was important for cytotoxicity to CaCo-2.

The IC_{50} values of lariciresinol 4'-*O*-glucoside (**171**), berchemol (**172**) and berchemol 4'-*O*-glucoside (**173**) were found to be 607.0, 833.0, 1260.0 μM, respectively. They were considered to be inactive towards colon cancer cell line.

Pinoresinol (**144**) was the most cytotoxic compound (IC_{50}=233.0 μM) among all pinoresinol derivatives (**176-178**) tested in this study. The other derivatives contain

one or two sugar moieties and had higher IC_{50} values than pinoresinol (**Table 46**). Pinoresinol monoglucoside (**176**) and pinoresinol-di-glucoside (**177**) had the IC_{50} values of 705.0 and 843.2 μM, respectively. Pinoresinol-apiose-glucoside (**178**), having one glucose and one apiose moieties in the molecule, exhibited the least cytotoxic properties towards colon cancer cell line (IC_{50}=1.13 mM). The above results concluded that the presence of a sugar molecule in all the isolated compounds exhibited reduced cytotoxicity. Pinoresinol type lignans (**144, 176-178**) were less cytotoxic than dibenzylbutyrolactones (**137-140**) and more cytotoxic than the lariciresinol type lignans (**171-173**).

Table 46. Cytotoxicity activities of lignans

Lignans	IC_{50} (μM)
Arctiin (**137**)	220.0
Matairesinoside (**138**)	288.0
Arctigenin (**139**)	7.0
Matairesinol (**140**)	124.0
Lappaol A (**143**)	135.0
Lariciresinol 4'-*O*-β-D-glucoside (**171**)	607.0
Berchemol (**172**)	833.0
Berchemol 4'-*O*-β-D-glucoside (**173**)	1260.0
Pinoresinol (**144**)	233.0
Pinoresinol 4-*O*-β-D-glucopyranoside (**176**)	705.0
Pinoresinol 4, 4'di-*O*-D-glucopyranoside (**177**)	843.2
Pinoresinol 4-*O*-β-apiose-β-D-glucopyranosideside (**178**)	1130.0

4.3.2.3 MTT Cytotoxicity Assay of Sesquiterpene lactones

In vitro cytotoxicity results for two sesquiterpene lactones, arctiopicrin (**21**) and 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**179**) are summarised in **Table 47.** The respective IC_{50} were found to be 8.5 and 26.4 μM. Arctiopicrin (**21**) was more cytotoxic than 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**179**). However,

both sesquiterpene lactones showed prominent cytotoxicity than most of the lignans except arctigenin (**139**).

Table 47. Cytotoxic activities of sesquiterpene lactones

Sesquiterpene lactones	IC_{50} (μM)
Arctiopicrin (**21**)	8.5
8-*O*-(4-Hydroxy-3-methylbutanoyl)-salonitenolide (**179**)	26.4

4.3.2.4 MTT Cytotoxicity Assay of Quinic acid Derivatives

The IC_{50} values of *p*-coumaroylquinic acid (**181**), *cis-p*-coumaroylquinic acid (**182**) and chlorogenic acid (**168**) were 146.4, 325.0 and 79.0 μM, respectively (**Table 48**). The *trans* compound (**181**) showed higher cytotoxicity than the *cis*-compound (**182**) towards the colon cancer cell line. These results corresponded well to the findings from other studies, that the *trans* configuration was essential for cytotoxic activity of phenolic compounds.[251] Chlorogenic acid (**168**) showed the most cytotoxic properties among all quinic acid derivatives. They were significantly more toxic than the tested epoxy lignans (**171-173** and **176-178**). These could be due to the fact that all these compounds possess an α,β-unsaturated carbonyl moiety, which can be considered as Michael acceptor, an active moiety often employed in the design of anticancer drug.[152]

Table 48. Cytotoxicity activities of quinic acid derivatives

Compounds	IC_{50} (μM)
p-Coumaroylquinic acid (**181**)	146.4
cis-p-Coumaroylquinic acid (**182**)	325.0
Chlorogenic acid (**168**)	79.0

4.3.2.5 MTT Cytotoxicity Assay of Flavonoids

The IC_{50} values of the flavonoids tested are shown in **Table 49**. Flavonoid glycosides were found to be less cytotoxic than the corresponding aglycones. For example, the IC_{50} value for kaempferol (**122**) was 201.0 μM while that of astragalin (**123**) and afzelin (**187**) were 302.0 and 316.0 μM respectively. However, flavanone apiose-glucuronic acid (**190**) which contained a carboxylic acid (-COOH) group instead of a free hydroxy group (-OH) in its glucose moiety showed higher cytotoxicity (IC_{50}=153.4 μM) than other flavonoid glycosides. The better cytotoxicity of flavonoid **190** may be due to the presence of a carboxylic acid group which could provide more binding affinity towards receptors than an OH group. The IC_{50} value for two *C*-glycosides **188** and **189** were 290.3 and 285.7 μM respectively. They were more cytotoxic than the *O*-glycosides (**123** and **187**). Apigenin (**108**) (IC_{50}=133.0 μM) and cirsiliol (**114**) (IC_{50}=96.0 μM) showed the most potent activities among the flavonoids evaluated in this study.

Table 49. Cytotoxicity activities of flavonoids

Flavonoids	IC_{50} (μM)
Kaempferol (**122**)	201.0
Astragalin (**123**)	302.0
Afzelin (**187**)	316.0
Apigenin (**108**)	133.1
Cirsiliol (**114**)	96.0
Orientin (**188**)	290.3
Hydroxybenzoyl-isoorientin (**189**)	285.7
Flavanone-apiose-glucuronic acid (**190**)	153.4

4.3.2.6 MTT Cytotoxicity Assay of Alkaloids

The cytotoxicity of nine alkaloids are presented in **Table 50**. Among the *cis-trans* isomers (**148** and **149**), the *trans* compounds (**148**) (IC_{50}=<100 μM) showed higher cytotoxicity than the *cis*-compound (**149**; IC_{50} >200 μM). A similar phenomenon was also observed for quinic acid derivatives suggesting that the *trans* configuration is essential for cytotoxicity in both alkaloids (**148**) and quinic acid (**168**). Moschamine (**148**) (IC_{50}=81.0 μM) was found to be more cytotoxic than centcyamine (**146**) (IC_{50}=82.2 μM). Although both **146** and **148** were structurally similar, moschamine had an additional methoxy group at C-3' position. This methoxy group may be responsible for higher cytotoxicity. However, montamine (**193**; IC_{50}=43.9 μM)), a dimer of moschamine (**148**) showed the most cytotoxic properties among all alkaloids. This result suggested that dimerisation could lead to an increase in cytotoxicity. Another dimer, schischkiniin (**194**) which was a trpytophan-derived indole alkaloid exhibited IC_{50}=76.0 μM. It can thus be suggested that indole alkaloids having a *p*-coumaroyl moiety linked to them might be candidate for further study as potential anticancer agents.

Table 50. Cytotoxicity activities of alkaloids

Alkaloids	IC_{50} (μM)
N-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**145**)	125.0
cis-*N*-(4-hydroxycinnamoyl)-5-hydroxytryptamine (**192**)	411.0
Centcyamine (**146**)	82.2
cis-Centcyamine (**147**)	213.0
Moschamine (**148**)	81.0
cis-moschamine (**149**)	213.0
Tryptamine (**191**)	198.0
Montamine (**193**)	43.9
Schischkiniin (**194**)	76.0

4.3.3 DPPH Antioxidant Assay

4.3.3.1 DPPH Assay of Centaurea Extracts

The DPPH assay of the MeOH extracts of *Centaurea* species are summarised in **Table 51**. The extracts of *C. huber-morathii, C. macrocephala, C. dealbata, C. americana, C. gigantea* and *C. schischkinii* showed significant antioxidant properties having the IC_{50} values of 3.1×10^{-2}, 3.5×10^{-2}, 4.7×10^{-2}, 5.2×10^{-2}, 7.2×10^{-2} and 15×10^{-2} mg/mL respectively. Quercetin (**125**) with IC_{50} value of 2.8×10^{-3} mg/mL was used as a positive control for antioxidant activity in this study.

Table 51. DPPH assay of *Centaurea* species

Methanol extracts	IC_{50} (mg/mL)
C. americana	5.2×10^{-2}
C. bornmuelleri	63.0×10^{-2}
C. cyanus	42.0×10^{-2}
C. dealbata	4.7×10^{-2}
C. gigantea	7.2×10^{-2}
C. huber-morathii	3.1×10^{-2}
C. macrocephala	3.5×10^{-2}
C. montana	32.7×10^{-2}
C. mucronifera	53.6×10^{-2}
C. pamphylica	47.3×10^{-2}
C. schischkinii	15.0×10^{-2}
C. urvillei	51.6×10^{-2}
Quercetin (**125**)	2.8×10^{-3}

4.3.3.2 DPPH Assay of Lignans

Antioxidant activity of lignans (**137-140**) are shown in **Table 52**. Matairesinoside (**138**) and matairesinol (**140**) showed the most prominent free radical scavenging activities (IC_{50} values of 2.2×10^{-3} and 2.0×10^{-3} mg/mL, respectively) among all

the lignans tested in this study. Both matairesinoside (**138**) and matairesinol (**140**) had hydroxy groups in C-3/C-3' of benzene ring and showed higher antioxidant activities. These results suggested that the presence of C-3' hydroxy group in benzene rings increased free radical scavenging activity significantly. For arctiin (**137**) and arctigenin (**139**) methoxy groups were present instead of the hydroxy groups at C-3/C-3', and their IC_{50} values were 16.0×10^{-2} and 1.9×10^{-2} mg/mL, respectively. Lappol A (**143**) and caffeoyl arctiin (**170**) showed better scavenging activities with the IC_{50} values of 3.6×10^{-2} and 3.2×10^{-2} mg/mL, respectively, than the corresponding arctiin (**137**) and arctigenin (**138**). Lariciresinol and pinoresinol type lignans (**171 and 144**) showed moderate levels of antioxidant activities with the IC_{50} values between 1.4-3.8 $\times 10^{-2}$ mg/mL.

Table 52. DPPH assay of lignans

Lignans	IC_{50} (mg/mL)
Arctiin (**137**)	16.0×10^{-2}
Matairesinoside (**138**)	2.2×10^{-3}
Arctigenin (**139**)	1.9×10^{-2}
Matairesinol (**140**)	2.0×10^{-3}
Lappaol A (**143**)	3.6×10^{-2}
Caffeoyl-arctiin (**170**)	3.2×10^{-2}
Lariciresinol 4-*O*-β-D-glucoside (**171**)	3.8×10^{-2}
Berchemol (**172**)	3.2×10^{-2}
Berchemol 4'-*O*-β-D-glucoside (**173**)	2.1×10^{-2}
Pinoresinol (**144**)	1.4×10^{-2}
Pinoresinol 4-*O*-β-D-glucopyranoside (**176**)	3.6×10^{-2}
Pinoresinol 4, 4' di-*O*-D-glucopyranoside (**177**)	$>5 \times 10^{-2}$
Pinoresinol 4-*O*-β-apiose-β-D-glucopyranoside (**178**)	3.0×10^{-2}

4.3.3.3 DPPH Assay of Sesquiterpene lactones

The sesquiterpene lactones are not common to be used as antioxidants. However, the antioxidant activity of two sesquiterpene lactones (**21** and **179**) were evaluated and

the results (**Table 53**) indicated that they did not have any free radical scavenging properties at the test concentrations.

Table 53. DPPH assay of sesquiterpene lactones

Sesquiterpe lactones	IC_{50} (mg/mL)
Arctiopicrin (**21**)	$>5 \times 10^{-2}$
8-*O*-(4-Hydroxy-3-methylbutanoyl)-salonitenolide (**179**)	$>5 \times 10^{-2}$

4.3.3.4 DPPH Assay of Quinic acid derivatives

Of the three quinic acid derivatives (**Table 54**), chlorogenic acid (**168**) with the IC_{50} value of 2.3×10^{-2} mg/mL showed the highest antioxidant activity. *p*-Coumaroyl quinic acid (**181**; $IC_{50}=7.6 \times 10^{-2}$ mg/mL) exhibited better activity than its *cis*-isomer ($IC_{50}=10.0 \times 10^{-2}$ mg/mL).

Table 54. DPPH assay of quinic acid derivatives

Compounds	IC_{50} (mg/mL)
p-Coumaroylquinic acid (**181**)	7.6×10^{-2}
cis-p-Coumaroylquinic acid (**182**)	10.0×10^{-2}
Chlorogenic acid (**168**)	2.3×10^{-2}

4.3.3.5 DPPH Assay of Flavonoids

Flavonoids are widely used as antioxidant. The results from the DPPH assay of flavonoids are summarised in **55**. Kaempferol (**122**) had an IC_{50} value of 3.5×10^{-3} mg/mL. The IC_{50} values for astragalin (**123**) and afzelin (**187**) were 8.0×10^{-2} and 11.6×10^{-2} respectively. These results showed that glycosidation did not effect significantly the scavenging activities of these flavonoids (**108** and **123**).

Table 55. DPPH assay of flavonoids

Flavonoids	IC_{50} (mg/mL)
Kaempferol (**122**)	3.5×10^{-3}
Astragalin (**123**)	8.0×10^{-2}
Afzelin (**187**)	11.6×10^{-2}
Apigenin (**108**)	1.4×10^{-2}
Orientin (**188**)	3.5×10^{-2}
Hydroxybenzoyl-isoorientin (**189**)	4.6×10^{-2}
Flavanone-apiose-glucuronic acid (**190**)	$>5.0 \times 10^{-2}$

4.3.3.6 DPPH Assay of Alkaloids

The antioxidant activities of alkaloids are summarised in **Table 56**. Moschamine (**148**), centcyamine (**146**) and schischkiniin (**194**) showed the highest antioxidant properties with the IC_{50} values of 2.2×10^{-3}, 2.8×10^{-3} and 3.8×10^{-3} mg/mL, respectively. Although alkaloids were not well reported as antioxidant like flavonoids, the higher antioxidant properties for moschamine, centcyamine and schischkiniin were due to the presence of indole moiety in these alkaloids.

Table 56. DPPH assay of alkaloids

Alkaloids	IC_{50} (mg/mL)
N-(4-Hydroxycinnamoyl)-5-hydroxytryptamine (**145**)	1.6×10^{-2}
cis-*N*-(4-Hydroxycinnamoyl)-5-hydroxytryptamine (**192**)	4.8×10^{-2}
Centcyamine (**146**)	2.8×10^{-3}
cis-Centcyamine (**147**)	3.2×10^{-3}
Moschamine (**148**)	2.2×10^{-3}
cis-Moschamine (**149**)	4.5×10^{-3}
Tryptamine (**192**)	6.4×10^{-2}
Montamine (**193**)	3.6×10^{-2}
Schischkiniin (**194**)	3.8×10^{-3}

4.4 Conclusions

The MeOH extracts *of C. bornmuelleri*, *C. huber-morathii*, *C. gigantea* and *C. macrocephala* were the most toxic in the brine shrimp lethality assay among the tested species (**Figure 77**). They also showed significant cytotoxicity in colon cancer cell lines. However, only *C. huber-morathii* exhibited a good level of antioxidant activity. These species can be studied further for potential anticancer agents.

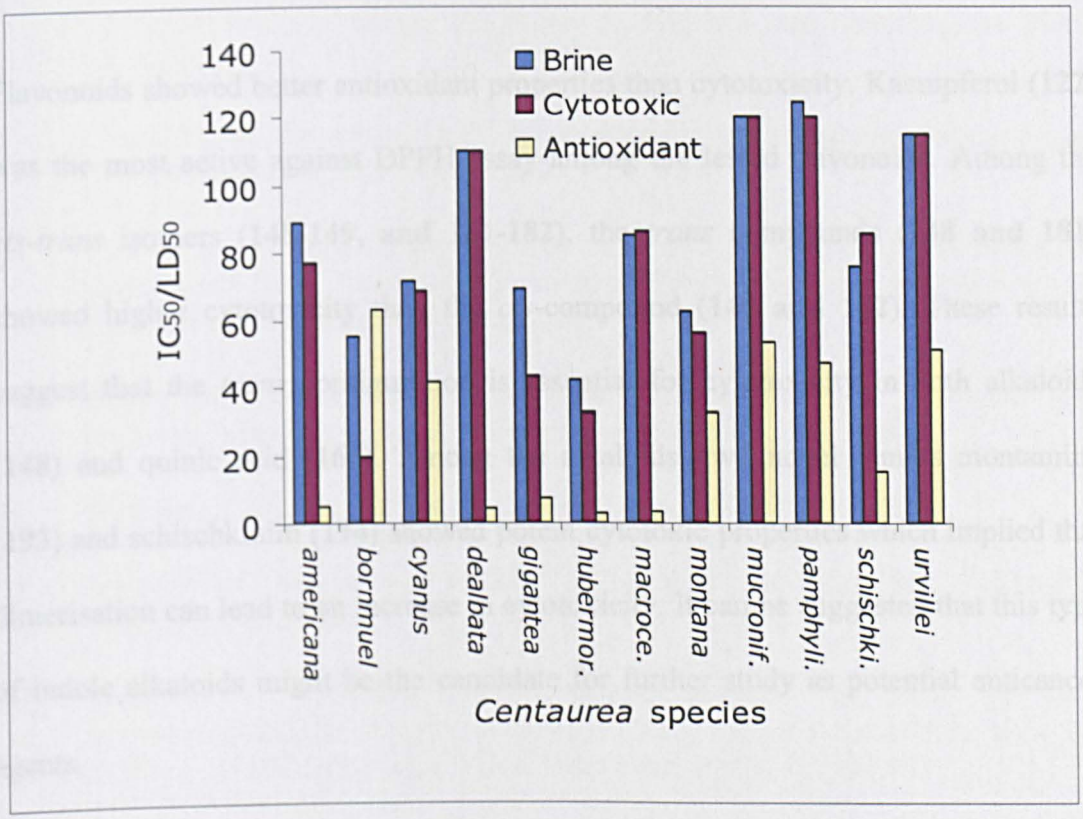


Figure 77. Comparative biological activity studies of *Centaurea* extracts against brine shrimp lethality assay (LD_{50}, mg/mL), MTT cytotoxicity assay (IC_{50}, μM) and DPPH antioxidant assay (IC_{50}, mg/mL)

Although arctiin (**137**) exhibited moderate cytotoxicity, arctigenin, **139** (aglycone) showed potent cytotoxicity. Thus, it was observed that cytotoxicity of lignan glycosides decreased significantly. Arctigenin might be the candidate for further potential anticancer agents. Matairesinol (**140**) and Matairesinoside (**138**) showed higher free radical scavenging activities. The presence of C-3/C-4' hydroxyl groups

in benzene rings played a crucial part for their antioxidant properties. Further investigation is necessary to use these compounds for antioxidant agents. Two sesquiterene lactones, arctiopicrin (**21**) and 8-*O*-(4-hydroxy-3-methylbutanoyl)-salonitenolide (**179**) showed prominent cytotoxicity but they did not show scavenging activities at the test concentrations. These two compounds can be studied further for potential cytotoxic agents.

Flavonoids showed better antioxidant properties than cytotoxicity. Kaempferol (**122**) was the most active against DPPH assay among the tested flavonoids. Among the *cis-trans* isomers (**148-149, and 181-182**), the *trans* compounds (**148 and 181**) showed higher cytotoxicity than the *cis*-compound (**149 and 182**). These results suggest that the *trans* configuration is essential for cytotoxicity in both alkaloids (**148**) and quinic acid (**168**). Among the alkaloids, two novel dimers montamine (**193**) and schischkiniin (**194**) showed potent cytotoxic properties which implied that dimerisation can lead to an increase in cytotoxicity. It can be suggested that this type of indole alkaloids might be the candidate for further study as potential anticancer agents.

References

(1) Samuelsson, G., Drugs of Natural Origin, A Textbook of Pharmacognosy. 4th revised ed., Swedish Pharmaceutical Press, Stockholm, **1999**.

(2) Cragg, G. M., Newman, D. J., Snader, K. M., Natural Products in drug discovery and development. *Journal of Natural Products,* **1997**, 60, 52-60.

(3) Kappor, L. D., CRC Handbook of Ayurvedic Medicinal Plants. CRC Press. Boca Raton, Florida, **1990**, 416-417.

(4) Farnsworth, N. R., Akerele, O., Bingel, A. S., Soejarto, D. D., Guo, Z., Medicinal Plants in therapy. *Bulletin of the World Health Organisation,* **1985**, 63, 965-981.

(5) Cragg, G. M., Newman, D. J., Medicinals for the Millennia. *Annals of the New York Academy of Sciences,* **2001**, 953, 3-25.

(6) Cox, P. A., Ethnopharmacology and the search for new drugs, In Bioactive Compounds from Plants. Ciba Foundation Symposium 154, John Wiley and Sons, Chichester, England, **1990**, 40-55.

(7) Shoeb, M., Investigations of Secondary Metabolites from Seaweeds and Cyanobacteria. MSc Thesis, University of Aberdeen, UK, **2003**.

(8) Seturner, F. W., *Gilbert's Annalen der. Physic,* **1817**, 25, 56-57.

(9) Anonymous., A Pictorial History of Herbs in Medicine and Pharmacy. *Herbalgram*, **1998**, 33-47.

(10) Buss, A. D., Waigh, R. D., Natural products as leads for new pharmaceuticals. In Burgers Medicinal Chemistry and Drug Discovery. 5th ed. M. E. Wolff, Eds., John Wiley, New York, **1995**, 1, 983-1033.

(11) Scriabine, A., Discovery and development of major drugs currently in use. In Pharmaceutical Innovation. R., Landau, B., Achilladelis & A. Scriabine, Eds., Chemical Heritage Press, **1999**, 148-270.

(12) Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., McPhail, A. T, Plant antitumor agents. VI. The isolation and structure of Taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia. Journal of American Chemical Society,* **1971**, 93, 2325-2327.

(13) Rowinsky, E. K., Onetto, N., Canetta, R. M., Arbuck, S. G., Taxol-The 1st of the Texanes, an important new class of antitumor agents. *Seminars in Oncology,* **1992**, 19, 646-662.

(14) Hartwell, J. L., Plants used against cancer. A survey, Lawrence, M. A., Quarterman Publications, **1982**, 438-439.

(15) Jaspars, M and Lawton, L. A., Cyanobacteria as a novel source of pharmaceuticals. *Current Opinion in Drug Discovery and Development,* **1998**, 1, 77-84.

(16) Potmeisel, M., Pinedo, H., Camptothecins: New Anticancer Agents. CRC Press, Boca Raton, Florida, **1995**, 149-150.

(17) Creemers, G. J., Bolis, G., Gore, M., Scarfone, G., Lacave, A. J., Guastalla, J. P., Despax, R., Favalli, G., Kreinberg, R., VanBelle, S., Hudson, I., Verweij, J., Huinink, W. W. T., Topotecan, an active drug in the second-line treatment of epithelial ovarian cancer: Results of a large European phase II study. *Journal of Clinical Oncology,* **1996**, 14, 3056-3061.

(18) Bertino, J. R, Irinotecan for colorectal cancer. *Seminars in Oncology,* **1997**, 24, S18-S23.

(19) Stahelin, H., Activity of a new glycosidic lignan derivative (VP 16-213) related to podophyllotoxin in experimental tumors. *European Journal of Cancer,* **1973**, 9, 215-221.

(20) Harvey, A. L., Medicines from nature: are natural products still relevant to drug discovery. *Trends in Pharmacological Sciences,* **1999**, 20, 196-198.

(21) Kellard, L. R., Flovopiridol, the first cyclic-dependent kinase inhibitor to enter the clinic: current status. *Expert Opinion on investigational related Drugs*, **2000**, 9, 2903-2911.

(22) Christian, M. C., Pluda, J. M., Ho, T. C, Arbuck, S, G, Murgo, A, J and Sausville, E. A., Promising new agents under development by Division of Cancer Treatment, Diagnosis, and Centers of the National Cancer Institute. *Seminars in Oncology,* **1997**, 13:2643-2655.

(23) Powell, R. G., Weisleder, D., Smith, C. R. Jr., Rohwedder, W. K., Structures of harringtonine, isoharringtonine, and homoharringtonine. *Tetrahedron Letters,* **1970**, 11, 815-818.

(24) Kantarjian, H. M., O'Brien, S., Anderlini, P., Talpaz, M., Treatment of chronic myelogenous Leukemia: Current status and investigational options. *Blood,* **1996**, 87, 3069-3081.

(25) Cragg, G. M., Boyd, M. R., Cardellina, J. H., Grever, M. R., Schepartz, S. A., Snader, K. M., Suffness, M., In Human Medicinal Agents from Plants. American Chemical Society Symposium Series 534, pp.80-95. Eds AD Kinghorn and M F Balandrin. Washington, DC: American Chemical Society, **1993**.

(26) Mabberlay, D. J., The Plant Book. 2nd eds. Cambridge University Press, Cambridge, **1997**.

(27) Davis, P. H., Flora of Turkey and the East Aegean Islands. University Press, Edinburgh, **1988**, 10.

(28) Kaij-a-Kamb, M., Amoros, M., Girre, L. Chimie et activities biologiques du genre *Centaurea*. *Pharmaceutica Acta Helvetiae,* **1992**, 67, 178-188.

(29) Poletti, A., Fiori e Piante Medicinali. Musumeci Ed., Milano, **1978**, 1, 152-153.

(30) Baser, K. H. C., Honda, G., Miki, W., Herb Drugs and Herbalist in Turkey. ILCAA, *Studia Culturae Islamicae,* Tokyo, **1986**, 27.

(31) Baytop, T., Turkiye'de Bitkilerle Tedavi, Gecmiste ve Bugun, Nobel Tip Kitabevi, Ankara, **1999**.

(32) Masso, J. L., Bertran, M. N., Adzet, T., Contribution a l'etude chimique et. pharmacologiue de quelques especes de *Centaurea* (Composees). *Plant. Med. Phytother,* **1979**, 13, 41-45.

(33) Ivorra M. D., Paya, M., Villa, A., Effect of beta-sitosterol-3-beta-D-glucoside on insulin-secretion *invivo* in diabetic rats and *in vitro* in isolated rat islets of Langerhans. *Pharmazie,* **1990**, 45, 271-3.

(34) Massol, J. L., Adzet, T., Hypoglycemic activity of *Centaurea-aspera* L., *Revista Espanola de Fisiologia*, **1976**, 32, 313-316.

(35) Vazquez, F. M., Suarez, M. A., Perez, A., Medicinal plants used in the Barros Area, Badajoz Province (Spain). *Journal of Ethnopharmacology*, **1997**, 55, 81-85.

(36) Robles, M., Wang, N., Kim, R., Choi, B. H., Cytotoxic effects of repin, a principal sesquiterpene lactone of Russian knapweed. *Journal of Neuroscience Research,* **1997**, 47, 90-97.

(37) Gonzalez, A. G., Darias, V., Alonso, G., Boada, J. N., Feria, M., Cytostatic activity of sesquiterepene lactones from Compositae of Canary islands. *Planta Medica*, **1978**, 33, 356-359.

(38) Gonzalez, A. G., Darias, V., Alonso, G., Estevez, E., The cytostatic activity of the chlorohyssopifolins, chlorinated sesquiterpene lactones from *Centaurea*. *Planta Medica,* **1980**, 40, 179-184.

(39) Nowak, G., A chemotaxonomic study of sesquiterpene lactones from subtribe Centaureinae of the Compositae. *Phytochemistry,* **1992**, 31, 2363-2368.

(40) Koukoulitsa, E., Skaltsa, H., Karioti, A., Demetzos, C., Dimas, K., Bioactive sesquiterpene lactones from *Centaurea* species and their cytotoxic/cytostatic activity against human cell lines *in vitro*. *Planta Medica,* **2002**, 68, 649-652.

(41) Harborne, J. B., "Terpenoids" in Encyclopedia of Life Science, an IT Based Learning Database, Macmillan Scientific Publishers, **2001**.

(42) Nes, W. D., Zhou, W., "Terpenoids" in Encyclopedia of Life Science, an IT Based Learning Database, Macmillan Scientific Publishers, **2001**.

(43) Charlwood, B. V., Banthorpe, D. V., Eds, Methods in Plant Biochemistry, Terpenoids, Academic Press, London, **1991**, 7.

(44) Nowak, G., Holub, M., Budesinsky. M., Sesquiterpene lactones. XXXVI. Sesquiterpene lactones in several subgenera of the genus *Centaurea* L. *Acta Societatis Botanicorum Poloniae,* **1989**. 58, 95-102.

(45) Bruno, M., Maggio, A., Paternostro, M. P., Rosselli, S., Arnold, N. A., Herz, W., Sesquiterpene lactones and other constituents of three Cardueae from Cyprus. *Biochemical Systematics and Ecology,* **2001**, 29, 433-435.

(46) El Dahmy, S., Bohlmann, F., Sarg, T. M., Ateya, A., Farrag, N., New guaianolides from *Centaurea aegyptica. Planta Medica,* **1985**, 2, 176-177.

(47) Janackovic, P., Tesevic, V., Milosavljevic, S., Vajs, V., Marin, P. D., Sesquiterpene lactones, lignans and flavones of *Centaurea affinis, Biochemical Systematics and Ecology,* **2004,** 32, 355-357.

(48) Fernandez, I., Pedro, J. R., Polo, E., Sesquiterpe lactones from *Centaurea alba* and *C. conifera. Phytochemistry,* **1995**, 38, 655-657.

(49) Ahmed, Z. F., Rimpler, H., Rizk, A. M., Hammouda, F. M., Ismail, S. I., The flavonoid constituents of certain *Centaurea* species grown in Egypt. *Phytochemistry*, **1970**, 9, 1595-1601.

(50) Ohno, N., Hirai, H., Yoshioka, H., Dominguez, X. A., Mabry, T. J., Cynaropicrin: A sesquiterpene lactone from *Centaurea americana. Phytochemistry,* **1973**, 12, 221-222.

(51) Cooper, G., Laird, A., Nahar, L., Sarker, S. D., Lignan glucosides from the seeds of *Centaurea americana* (Compositae). *Biochemical Systematics and Ecology,* **2002**, 30, 65-67.

(52) Dominguez, X. A., Gonzalez, L. A., Rojas, P., Terpenoids of *Centaurea americana. Phytochemistry,* **1972**, 11, 850-851.

(53) Cardona, M. L., Fernandez, I., Pedro, J. R., Perez, B., Sesquiterpene lactones and flavonoids from *Centaurea aspera. Phytochemistry,* **1991**, 30, 2331-2333.

(54) Cardona, L., Fernandez, I., Pedro, J. R., Vidal, R., Polyoxygenated terpenes and cyanogenic glucosides from *Centaurea aspera* var. *subinermis. Phytochemistry,* **1992**, 31, 3507-3509.

(55) Picher, M. T., Seoane, E., Tortajada, A., Flavones, sesquiterpen lactones and glycosides isolated from *Centaurea aspera* var. *stenophylla*. *Phytochemistry*, **1984**, 23, 1995-1998.

(56) Marco, J. A., Sanz-Cervera, J. F., Yuste, A., Sancenon, F., Carda, M., Sesquiterpenes from *Centaurea aspera*. *Phytochemistry*, **2005**, 66, 1644-1650.

(57) Skaltsa, H., Lazari, D., Georgiadou, E., Kakavas, S., Constantinidis, Th., Sesquiterpene lactones from *Centaurea* species: *C. thessala* subsp. drakiensis and *C. attica* subsp. *attica*. *Planta Medica*, **1999**, 65, 393-394.

(58) Skaltsa, H., Lazari, D., Panagouleas, C., Georgiadou, E., Garcia, B., Sokovic, M., Sesquiterpene lactones from *Centaurea thessala* and *Centaurea attica*. Antifungal activity. *Phytochemistry*, **2000**, 55, 903-908.

(59) Daniewski, W. M., Nowak, G., Further sesquiterpene lactones of *Centaurea bella*. *Phytochemistry*, **1993**, 32, 204-205.

(60) Oksuz, S., Ulubelen, A., Aynechi,Y., Wagner, H., A guianolide from *Centaurea behen*. *Phytochemistry*, **1982**, 31, 2747-2749.

(61) Barrero, A. F., Oltra, J. E., Alvarez, M., Raslan, D. S., Saude, D. A., Akssira, M., New sources and antifungal activity of sesquiterpene lactones. *Fitoterapia*, **2000**, 71, 60-64.

(62) Flamini, G., Pardini, M., Morelli, I., A flavonoid sulphate and other compounds from the roots of *Centaurea bracteata*. *Phytochemistry*, **2001**, 58, 1229-1233.

(63) Flamini, G., Antognoli, E., Morelli, I., Two flavonoids and other compounds from the aerial parts of *Centaurea bracteata* from Italy. *Phytochemistry*, **2001**, 57, 559-564.

(64) Marco, J. A., Sanz, J. F., Sancenon, F., Susanna, A., Rustaiyan, A., Saberi, M., Sesquiterpene lactones and lignans from *Centaurea* species. *Phytochemistry*, **1992**, 31, 3527-3530.

(65) Dawidar, A. M., Metwally, M. A., Abou-Elzahab, M., Abdel-Mogib, M., Chemical constituents of two *Centaurea* species. *Pharmazie*, **1989**, 44, 735-736.

(66) Gonzalez, A. G., Bermejo, J., Cabrera, I., Massanet, G., M., Mansilla, M., Galindo, A., Two sesquiterpene lactones from *Centaurea canariensis*. *Phytochemistry*, **1978**, 17, 955-956.

(67) Gonzalez, A. G., Barrera, B. J., Garcia, T. Z., Rosas, F. E., Sesquiterpene lactones from *Centaurea* species. *Phytochemistry*, **1984**, 23, 2071-2072.

(68) Bruno, M., Vasallo, N., Fazio, C., Gedris, T. E., Herz, W., Sesquiterpene lactones of two *Centaurea* species from Sicily. *Biochemical Systematics and Ecology,* **1998,** 26, 801-803.

(69) Bruno, M., Herz, W., Sesquiterpene lactones and flavones from *Centaurea cineraria* subsp. *umbrosa. Phytochemistry*, **1988,** 27, 1873-1875.

(70) Fernandez, I., Garcia, B., Grancha, F. J., Pedro, J. R., Two guaianolides from *Centaurea collina. Phytochemistry,* **1987**, 26, 2403-2405.

(71) Sarker, S. D., Laird, A., Nahar, L., Kumarsamy, Y., Jaspars, M., Indole alkaloids from the seeds of *Centaurea cyanus* (Asteraceae). *Phytochemistry,* **2001**, 57, 1273-1276.

(72) Litvinenko, V. I., Bubenchikova, V. N., Phytochemical study of *Centaurea cyanus. Khimiya Prirodynkh Soedinenii,* **1988,** 6, 792-795.

(73) Bubenchikova, V. N., Coumarins of plants of the genus *Centaurea. Khimiya Prirodynkh Soedinenii*, **1990,** 6, 829-830.

(74) Fortuna, A. M., Riscala, E. C. de, Catalan, C. A. N., Gedris, T. E., Herz, W., Sesquiterpene lactones and other constituents of *Centaurea diffusa. Biochemical Systematics and Ecology,* **2002**, 30, 805-808.

(75) Akkal, S., Benayache, F., Medjroubi, K., Tillequin, F., Sequin, E., Flavonoids from *Centaurea furfuracea* (Asteraceae). *Biochemical Systematics and Ecology,* **2003**, 31, 641-643.

(76) Akkal, S., Benayache, F., Benayache, S., Medjroubi, K., Jay, M., Tillequin, F., Sequin, E., A new flavone glycoside from *Centaurea furfuracea. Fitoterapia,* **1999**, 70, 368-370.

(77) El-Masry, S., Darwish, F. A., Abou-Donia, A., Abou-Karam, M. A., Grenz, M., Sesquiterpene lactones from *Centaurea glomerate. Phytochemistry,* **1985**, 24, 999-1001.

(78) Medjroubi, K., Benayache, F., Benayache, S., Akkal, S., Kaabeche, M., Tillequin, F., Seguin, E., Eudesmanolide from *Centaurea granata. Phytochemistry,* **1998**, 49, 2425-2427.

(79) Oksuz, S., Serin, S., Topcu, G., Sesquiterpene lactones from *Centaurea hermannii. Phytochemistry*, **1994**, 35, 435-438.

(80) Flamini, G., Bullerti, C., Morelli, I., Manunta, A., A new flavonoid glycoside from *Centaurea horrida. Journal of Natural Product,* **2000**, 63, 662-663.

(81) Massiot, G., Morfaux, A. M., Le Men-Olivier, L., Bouquant, J., Madaci, A., Mahmoud, A., Chopova, M., Aclinou, P. Guaianolides from the leaves of *Centaurea incana. Phytochemistry,* **1986**, 25, 258-261.

(82) Akkal, S., Benayache, F., Benayache, S., Jay, M., Flavonoids from *Centaurea incana* (Asteraceae). *Biochemical Systematics and Ecology,* **1997**, 25, 361-362.

(83) Flamini, G., Stoppelli, G., Morelli, I., Ertugral, K., Dural, H., Tugay, O., Demirelma, H., Secondary metabolites from *Centaurea isaurica* from Turkey and their chemotaxonomical significance. *Biochemical Systematics and Ecology,* **2004**, 32, 553-557.

(84) Oksuz, S., Ayyildiz, H., Johansson, C., 6-methoxylated and *C*-glycosyl flavonoids from *Centaurea* species. *Journal of Natural Products,* **1984**, 47, 902-903.

(85) Oksuz, S., Putun, E., Guaianolides from *Centaurea kotschyi. Phytochemistry,* **1983**, 22, 2615-2616.

(86) Bruno, M., Maggio, A., Rosselli, S., Gedris, T. E., Herz, W., Sesquiterpene lactones and other constituents of *Centaurea paniculata* ssp. *Castellana. Biochemical Systematics and Ecology*, **2002**, 30, 379-381.

(87) Sanz, J. F., Marco, J. A., Rustaiyan, A., Chemical constituents from *Centaurea persica* and *Senecio coronopifolius. Pharmazie*, **1990**, 45, 381-381.

(88) Reynaud, J., Couble, A., Raynaud, J., Flavonoids from *Centaurea-macrocephala* Muss Puschk Ex Willd. (Compositae). *Pharmazie,* **1992**, 47, 51-52.

(89) Ribeiro, N. L., Nahar L., Kumarasamy, Y., Mir-Babayev N., Sarker S D., Flavonoids *C*-glucosides and a lignan from *Centaurea macrocephala* (Compositae). *Biochemical Systematics and Ecology,* **2002**, 30, 1097-1100.

(90) Barrero, A. F., Oltra, J. E., Morales, V., Alvarez, M., Biomimetic Cyclization of Cnicin to Malacitanolide, a Cytotoxic Eudesmanolide from *Centaurea malacitana. Journal of Natural Products,* **1997**, 60, 1034-1035.

(91) Barrero, A. F., Sanchez, J. F., Rodriguez, I., Germacranolides from *Centaurea melitensis. Phytochemistry,* **1989**, 28, 1975-1976.

(92) Gonzalez, A. G., Arteaga, J. M., Breton, J. L., Elemanolides from *Centaurea melitensis. Phytochemistry,* **1975**, 14, 2039-2041.

(93) Sarker, S. D., Dinan L. N., Sik V., Underwood E., Waterman, P. G., Moschamide: An Unusual alkaloids from the seeds of *Centaurea moschata. Tetrahedron Letters,* **1998**, 39, 1421-1424.

(94) Sarker, S. D., Dinan L. N., Sik V., Rees, H. H., Moschatine: An unusual steroidal glycoside from *Centaurea moschata. Phytochemistry,* **1998**, 48, 1039-1043.

(95) Sarker, S. D., Savchenko, T., Whiting, P., Sik, V.; Dinan, L. N., Moschamine, *cis*-moschamine, moschamine, moschamindole and moschamindolol: four novel indole alkaloids from *Centaurea moschata*. *Natural Product Letters*, **1997**, 9, 189-199.

(96) Sarker, S. D., Savchenko T., Whiting, P., Sik V., Lafont, R., Dinan L. N., Occurance of ecdysteroids in the Genus *Centaurea* (Compositae). *Biochemical Systematics and Ecology*, **1997**, 25, 367-368.

(97) Medjroubi, K., Benayache, F., Benayache, S., Akkal, S., Khalfallah, N., Aclinou, P., Guaianolides from *Centaurea musimomum*. *Phytochemistry*, **1997**, 45, 1449-1451.

(98) Bruno, M., Fazio, C., Paternostro, M. P., Diaz, J. G., Herz, W., Sesquiterpene lactones and other constituents of *Centaurea napifolia*. *Planta Medica*, **1995**, 61, 374-375.

(99) Akkal, S., Benayache, F., Bentamene, A., Medjroubi, K., Seguin, E., Tillequin, F., Flavonoid aglycones from *Centaurea napifolia*. *Chemistry of Natural Compounds*, **2003**, 39, 219-220.

(100) Christensen, L. P., Lam, J., Flavones and other constituents from *Centaurea* species. *Phytochemistry*, **1991**, 30, 2663-2665.

(101) Bruno, M., Paternostro, M. P., Gedris, T. E., Herz, W., Sesquiterpene lactones. *Phytochemistry*, **1996**, 41, 335-336.

(102) Medjroubi, K., Bouderdara, N., Benayache, F., Akkal, S., Seguin, E., Tillequin, F., Sesquiterpene lactones of *Centaurea nicaensis*. *Chemistry of Natural Compounds*, **2003**, 39, 506-507.

(103) Tesevic, V., Milosavljevic, S., Vajs, V., Janackovic, P., Popsavin, M., Dithiophenes and other constituents of roots of *Centaurea nicolai*. *Biochemical Systematics and Ecology*, **2003**, 31, 89-90.

(104) Vajs, V., Todorovic, N., Ristic, M., Tesevic, V., Todorovic, B., Janackovic, P., Marin, P., Milosavljevic, S., Guaianolides from *Centaurea nicolai*: antifungal activity. *Phytochemistry*, **1999**, 52, 383-384.

(105) Middleton M., Cox P. J., Jaspars M., Kumarasamy Y., Nahar L., Reid R., Sarker S. D., Dibenzylbutyrolactone lignans and indole alkaloids from the seeds of *Centaurea nigra* (Asteraceae). *Biochemical Systematics and Ecology*, **2003**, 31, 653-656.

(106) Gousiadou, C., Skaltsa, H., Secondary metabolites from *Centaurea orphanidea*. *Biochemical Systematics and Ecology*, **2003**, 31, 389-396.

(107) Flamini, G., Pardini, M., Morelli, I., Ertugrul, K., Dural, H., Bagci, Y., Kargioglu, M., Flavonoid glycosides from *Centaurea pseudoscabiosa* subsp. *pseudoscabiosa* from Turkey. *Phytochemistry*, **2002**, 61, 433-437.

(108) Oksuz, S., Serin, S., Triterpenes of *Centaurea ptosimopappoides*. *Phytochemistry*, **1997**, 46, 545-548.

(109) Panagouleas, C., Skaltsa, H., Lazari, D., Skaltsounis, A., Sokovic, Antifungal activity of secondary metabolites of *Centaurea raphanina* ssp. Mixta, growing Wild in Greece. *Pharmaceutical Biology*, **2003**, 41, 266-270.

(110) Stevens, K. L., Sesquiterpene lactones from *Centaurea repens*. *Phytochemistry*, **1982**, 21, 1093-1098.

(111) Suchy, M., Herout, V., Sorm, F., Absolute configuration of cnicin and scabiolode. *Collection Czechoslov. Chemical Communication*, **1962**, 27, 2398-2403.

(112) Suchy, M., Herout, V., Sorm, F., The Structure of salonitolide, A sesquiterpenic lactone of germacrane type from *Centaurea salonitana* Vis. *Collection Czechoslov. Chemical Communication*, **1965**, 30, 2863-2864.

(113) Daniewski, W. M., Nowak, G., Pankowska, E., Georgiadis, T., Routsi, E., Rychlewska, U., Szczepanska, B., Sesquiterpene lactones of *Centaurea salonitana*. *Phytochemistry*, **1993**, 34, 445-447.

(114) Daniewski, W. M., Nowak, G., Routsi, E., Rychlewska, U., Szczepanska, B., Skibicki, P., Salograviolide A, A sesquiterpene from *Centaurea salonitana*. *Phytochemistry*, **1992**, 31, 2891-2893.

(115) Ferguson, C. A., Nahar, L., Finnie, D., Kumarasamy, Y., Raymond, R., MirBabayev, N. F., Sarker, S. D., *Centaurea scabiosa*: a source of dibenzylbutyrolactone lignans. *Biochemical Systematics and Ecology*, **2003**, 31, 303-305.

(116) Youssef, D. T. A., Sesquiterpene lactones of *Centaurea scoparia*. *Phytochemistry*, **1998**, 49, 1733-1737.

(117) Youssef, D., Frahm, A. W., Circular dichroism of C-7, C-6 *trans*-fused guainolides of *Centaurea scoparia*. *Phytochemistry*, **1997**, 45, 551-554.

(118) Youssef, D., Frahm, A. W., Constituents of the Egyptian *Centaurea scoparia*; III. Phenolic Constituents of the Aerial parts. *Planta Medica*, **1995**, 61, 570-573.

(119) Youssef, D., Frahm, A. W., Constituents of the Egyptian *Centaurea scoparia*; II. Guaianolides of the aerial parts. *Planta Medica*, **1994**, 60, 572-575.

(120) Youssef, D., Frahm, A. W., Constituents of the Egyptian *Centaurea scoparia*; Chlorinated guainolides of the aerial parts. *Planta Medica*, **1994**, 60, 267-271.

(121) Helal, A. M., Nakamura, N., Meselhy, M. R., El-Fishawy, A. M., Hattori, M., Mahran, G. H., Guaianolides from *Centaurea scoparia*. *Phytochemistry*, **1997**, 45, 551-554.

(122) Bruno, M., Diaz, J., G., Herz, W., Guaianolides and Lignans from *Centaurea solstitialis* subsp. *schouwii*. *Phytochemistry*, **1991**, 30, 4165-4166.

(123) Bruno, M., Fazio, C., Passananti, S., Paternostro, M., Diaz, J. G., Herz, W., Sesquiterpene lactones from *Centaurea sphaerocephala* ssp. *sphaerocephala*. *Phytochemistry*, **1994**, 35, 1371-1372.

(124) Fortuna, A. M., Riscala, E. C. D., Catalan, C. A. N., Gedris, T. E., Herz, W., Sesquiterpene lactones from *Centaurea tweediei*. *Biochemical Systematics and Ecology*, **2001**, 29, 967-971.

(125) Appendino, G., Gariboldi, P., Belliardo, F., Sesquiterpene lactones from *Centaurea uniflora* subsp. *nervosa*. *Phytochemistry*, **1986**, 25, 2163-2165.

(126) Ulubelen, A., Oksuz, S., Cytotoxic flavones from *Centaurea urvillei*. *Journal of Natural Products*, **1982**, 45, 373-373.

(127) Agrawal, P. K., (Eds.), Carbon-13 NMR of Flavonoids, Elsevier, Amsterdam, **1989**.

(128) Harborne, J. B. (Eds.) The Flavonoids: Advances in Research since 1980, Chapman & Hall, London, **1988**.

(129) Dewick, P. M., Medicinal Natural Products, A Biosynthetic Approach, 2nd eds., John Wiley & Sons, Ltd, England, **2001**.

(130) Seidel, V., Windhovel, J., Eaton, G., Alfermann, A. W., Arroo, R. R. J., Medarde, M., Petersen, M., Woolley, J. G., Biosynthesis of podophyllotoxin in *Linum album* cell cultures. *Planta*, **2002**, 215, 1031-1039.

(131) Raffaelli, B., Hoikkala, A., Leppala, E., Wahala, K., Enterolignans. *Journal of Chromatography B*, **2002**, 77, 29-43.

(132) Ward, R. S. Lignans, neolignans and related compounds. *Natural Product Reports*, **1997**, 14, 43-74.

(133) Davin, L. B., Lewis, N. G., in; Stafford, H. A., Ibrahim R. K., (Eds.), Phenolic Metabolism in Plants, Vol. 26, Plenum Press, New York, p. 325, Chapter. 11, **1992**.

(134) Lewis, N. G., Sarkanen Eds., Lignin and Lignan Biosynthesis, ACS Symposium Series 697, American Chemical Society, Washington, DC, **1998**.

(135) Davin, L. B., Wang. H. B., Crowell, A. L., Bedgar, D. L., Martin, D. M., Sarkanen, S., Lewis, N. G., Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science,* **1997**, 275, 362-366.

(136) Umezawa, T., Davin, L. B., Lewis, N. G., Formation of the lignans (-) secoisolariciresinol by cell free extracts of *Forsythia intermedia. Biochemical and Biophysical Research Communications*, **1990**, 171, 1008-1014.

(137) Umezawa, T., Davin, L. B., Lewis, N. G., Formation of (-) secoisolariciresinol and (-)-matairesinol with *Forsythia intermedia* cell free extracts. *Journal of Biological Chemistry,* **1991**, 266, 10210-10217.

(138) Ozawa, S., Davin, L. B., Lewis, N. G., Formation of (-)-arctigenin in *Forsythia intermedia. Phytochemistry,* **1993**, 32, 643-652.

(139) Woolley, J. G., "Alkaloids" in Encyclopedia of Life Science, an IT Based Learning Database, Macmillan Scientific Publishers, **2000**.

(140) USDA-ARS-GRIN data base, USDA, ARS, National Genetic Resources Program, Germplasm Resources Information Network-(GRIN). National Germplasm Resources Laboratory, Beltsville, Maryland, USA, **2002**.

(141) Celik, S., The Ecological Characteristics of the Species Belonging to Psephelloidea (Boiss.) Sosn. Section of the Genus Centaurea L., Anadolu University, Graduate School of Natural and Applied Sciences, **2003**.

(142) Wagentiz, G., In Flora of Turkey and the East Aegean Islands, Davis, P.H, Eds., University Press, Edinburgh, **1975**, 5.

(143) Coll, J. C., Bowden, B. F., The application of vacuum liquid-chromatography to the separation of terpene mixtures. *Journal of Natural Products,* **1986**, 49, 934-936.

(144) Pelletier, S. W., Chokshi, H. P., Desai, H. K., Separation of diterpenoid alkaloid mixtures using vacuum liquid-chromatography. *Journal of Natural Products,* **1986**, 49, 892-900.

(145) Stahl, E., Thin Layer Chromatography, 2nd ed., Springer-Verlag, Berlin, Heidelberg, New York, **1966**, 855-856.

(146) Touchtone, J. C., Dobbins, M. F., Practice of Thin-Layer Chromatography, 1st print, John Wiley & Sons Ltd, New York, **1977**, 170-171.

(147) Crews, P., Rodriguez, J, Jaspars, M., *Organic Structure Analysis*; Oxford University Press: New York, **1988**.

(148) Dass, C., In *Mass spectrometry, Clinical and biomedical applications*; Desiderio, D. M., Eds., Plenum Press, New York, **1992**, 2, 1-52.

(149) Dinan, L., Dereplication and Partial Identification of Compounds, In Method in Biotechnology, Natural Products Isolation, Sarker, S. D., (Eds.), Humana Press, London, **2005**.

(150) Bax, A., Subramanian, S., Sensitivity-enhanced two-dimensional heteronuclear shift correlations NMR-spectroscopy. *Journal of Magnetic Resonance,* **1986**, 67, 565-569.

(151) Bodenhausen, G; Freeman, R. Direct comparison of spin-lattice and spin-spin relaxation in high resolution NMR. *Journal of Magnetic Resonance,* **1977**, 28, 303-308.

(152) Ayres, D. C., Loike, J. D., Lignans, Chemical, biological and clinical properties, Cambridge University Press, **1990**.

(153) Delazar, A., Byres, M., Gibbons, S., Kumarasamy, Y., Modarresi, M., Nahar, L., Shoeb, M., Sarker, S. D., Iridoid glycosides from *Eremostachy glabra. Journal of Natural Products,* **2004**, 67, 1584-1587.

(154) Rahman, M. M. A., Dewick, P. M., Jackson, D. E., Lucas, J., Lignans of *Forsythia intermedia. Phytochemistry*, **1990**, 29, 1971-1980.

(155) Nishibe, S., Tsukamoto, H., Hisada, S., Effects of O-methylation and O-glucosylation on carbon-13 nuclear magnetic resonance chemical shifts of matairesinol, (+)-pinoresinol and (+)-epipinoresinol. *Chemical & Pharmaceutical Bulletin,* **1984**, 32, 4653-4657.

(156) Bastos, M. M. S. M., Kijjoa, A., Cardoso, J. M., Gutierrez, A. B., Herz, W., Lignans and other constituents of *Centaurea sphaerocephala ssp. Polyacantha. Planta Medica,* **1990**, 56, 403-405.

(157) Ichihara A., Oda, K., Numata, Y., Sakamura, S., Lappaol A and B, Novel lignans from *Arctium Lappa* L. *Tetrahedron Letters,* **1976**, 44, 3961-3964.

(158) Gamal, A. A. E., Takeya, K., Itokawa, H., Halim, A. F., Amer, M. M., Saad, H. E. A., Lignan Bis-glucosides from *Galium sinaicum. Phytochemistry,* **1997**, 45, 597-600.

(159) Dellagreca, M., Molinaro, A., Monaco, P., Previtera, L., 2,New lignan glucosides from *Arum-italicum. Heterocycles,* **1993**, 36, 2081-2086.

(160) Sakurai, N., Nagashima, S., Kawai, K., Inoue, T., *Chemical & Pharmaceutical Bulletin,* **1989**, 37, 3311-3315.

(161) Schumacher, B., Scholle, S., Holzl, J., Khudeir, N., Hess, S., Muller, C. E., Lignans Isolated from Valerian: Identification and Characterization of a new Olivil Derivative with Partial Agonistic Activity at A1 Adenosine Receptors. *Journal of Natural Products,* **2002**, 65, 1479-1485.

(162) Min, B. S., Na, Mi. K, Oh, S. R., A, K. S., J, G. S., Li G; Lee, S. K., Joung, H., Lee, H. K., New Furofuran and Butyrolactone Lignans with Antioxidant Activity from the Stem bark of *Styrax japonica*. *Journal of Natural Products*, **2004**, 67, 1980-1986.

(163) Duan, H., Tajkaishi, Y., Momota, H., Ohmoto, Y., Taki, T., Immunosuppresive constituents from *Saussurea medusa*. *Phytochemistry*, **2002**, 59, 85-90.

(164) Chen, I. S., Chen, T. Li., Chang, Y. L., Teng, C. M., Lin, W. Y., Chemical Constituents and Biological Activities of the Fruits of *Zanthoxylum integrifoliolum*. *Journal of Natural Products*, **1999**, 62, 8333-837.

(165) Cowan, S., Stewart, M., Abbiw, D.K., Latif, Z., Sarker, S. D., Nash, R. J., Lignans from *Strophanthus gratus*. *Fitoterapia*, **2001**, 72, 80-82.

(166) Miyazawa, M., Kasahar, H., Kameoka, H., Phenolic lignans from flower buds of *Magnolia fargessi*. *Phytochemistry*, **1992**, 31, 3666-3668.

(167) Miyazawa, M., Kasahar, H., Kameoka, H., Biotransformation of lignans: A specific microbial oxidation of (+)-Eudesmin and (+)-magnolin by *Aspergillus niger*. *Phytochemistry*, **1993**, 34, 1501-1507.

(168) Fonseca, S. F., Campello, J. D. P., Barata, L. E. S., Ruveda, E. A., ^{13}C NMR spectral analysis of Lignans from *Araucaria angustifolia*. *Phytochemistry*, **1978**, 17, 499-502.

(169) Jong, T-T., Chau, S. W., Antioxodant activities of constituents isolated from *Pandanus odoratissimus*. *Phytochemistry*, **1998**, 49, 2145-2148.

(170) Kitagawa, S., Hisada, S., Nishibe, S., Phenolic compounds from *Forsythia* leaves. *Phytochemistry*, **1984**, 23, 1635-1636.

(171) Sugiyama, S., Kikuchi, M., Studies on the Constituents of *Osmanthus* Species. Structure of Lignans glycosides from the leaves of *Osmanthus asiaticus* Nakai. *Chemical & Pharmaceutical Bulletin*, **1991**, 39, 483-495.

(172) Sih, C. J., Ravikumar, P. R., Hung, F. C., Buckner, C., Whitlock, H., Isolation and Synthesis of Pinoresinol diglucoside, a major Antihypertensive Principle of Tu-Chung (*Eucommia ulmoides*, Oliver). *Journal of American Chemical Society*, **1976**, 98, 5412-5413.

(173) Takahashi, H., Hirata, S., Minami, H., Fukuyama, Y., Triterpene and flavanone glycoside from *Rhododendron simsii*. *Phytochemistry*, **2001**, 56, 875-879.

(174) Kitagawa, I., Hori K., Sakagami, M., Hashiuchi, F., Yoshikawa, M., Ren, J., Saponin and sapogenol. XLIX. On the constituents of the roots of glycyrrhiza inflata Batalin from Xinjiang, China. Characterization of two sweet oleanane-type triterpene oligoglycosides, apioglycyrrhizin and araboglycyrrhizin. *Chemical & Pharmaceutical Bulletin,* **1993**, 41, 1350-1357.

(175) Abe, F., Yamauchi, T., Lignan glycosides from *Parsonsia laevigata. Phytochemistry,* **1989**, 28, 1737-1741.

(176) Pavia, D. L., Introduction to spectroscopy, A guide for students of Organic Chemistry, Fort Worth, TX: Harcourt College, **2001**.

(177) Lazari, D., Garcia, B., Skaltsa, H., Pedro, J. R., Harvala, C., Sesquiterpene lactones from *Onopordon laconicum* and *O. sibthorpianum. Phytochemistry,* **1998**, 47, 415-422.

(178) Ahmed, A. A., Balboul, B. A. A., Scott, Williams, H. H., Miao, B., Mabry, T. J., Eudesmane derivatives from *Iva frutescens. Phytochemistry,* **1998**, 47, 411-413.

(179) Zhao, Y., Yue, J., Lin, Z., Ding, J., Sun, H., Eudesmane sesquiterpenes from *Laggera pterodonta. Phytochemistry,* **1997**, 44, 459-465.

(180) Nishizawa, M., Izuhara, R., Kaneko, K., Koshihara, Y., Fujimoto, Y., 5-Lipoxygenase inhibitors isolated from gardeniae fructus. *Chemical & Pharmaceutical Bulletin,* **1988**, 36, 87-95.

(181) Gillet, F., Mesnard, F., Fliniaux, O., Monti, J-P., Fliniaux, M-A., Chlorogenic acid in a *Nicotiana plumbaginifolia* cell suspension. *Plant Physiology & Biochemistry,* **1999**, 37, 869-875.

(182) Ronsted, N., Strandgaard, H., Jensen, S. R., Molgaard, P., Chlorogenic acid in a *Nicotiana plumbaginifolia* cell suspension. *Biochemical Systematic and Ecology,* **2002**, 30, 1105-1108.

(183) Sarker, S. D., Girault, J. P., Lafont, R., Dinan, L. N., Ecdyesteroid xylosides from *Limnanthes douglasii. Phytochemistry*, **1997**, 44, 513-521.

(184) Coll, J., Reixach, N., Sanchez-Baeza, F., Casas, J., Camps, F., New ecdysteroids from *Polypodium vulgare. Tetrahedron,* **1994**, 50, 7247-7252.

(185) Pis, J., Budesinsky, M., Vokac, K., Laudova, V., Harmatha, J., Ecdysteroids from the roots of *Leuzea carthamoides. Phytochemistry,* **1994**, 37, 707-711.

(186) Imai, S., Hori, M., Fujioka, S., Murata, E., Goto, M., Nakanishi, K., Isolation of four new phytoecdysones makisterone A, B, C, D, and the structure of makisterone A, A C_{28} steroid. *Tetrahedron Letters,* **1968**, 9, 3883-3886.

(187) Roth, U., Konig, M., Seifert, K., Ecdysteroids from *Penstemon venustus. Phytochemistry,* **1995**, 39, 941-942.

(188) Gupta, S., Ali, M., Alam, M. S., Niwa, M., Sakai, T., 24β-Ethylcholest 4-en-3β-ol from the roots of *Lawsonia inermis*. *Phytochemistry,* **1992**, 31, 2558-2560.

(189) Ulubelen, A., Tan, N., Topou, G., Terpenoids from *Salvia cardidissima* subsp. *Candidissima*. *Phytochemistry,* **1997**, 45, 1221-1223.

(190) Kobayashi, M., Tsuru, R., Todo, K., Mitsuhashi, H., Marine Sterols-II, Asterosterol, A New C_{26} sterol from *Asterias amurensis* lutken. *Tetrahedron,* **1973**, 29, 1193-1196.

(191) Markham, K. R., Techniques of Flavonoid Identification. Academic Press, London, **1982**.

(192) Mabry, T. J., Markham, K. R., Thomas, M. B., The systematic Identification of flavonoids. Springer-Verlag, New York, **1970**.

(193) Liu, Y., Wu, Y., Yuan, Ke., Ji, C., Hou, A., Yoshida, T., Okuda, T., astrgalin 2'', 6''-di-*O*-gallate from *Loropetalum chinese*. *Phytochemistry,* **1997**, 46, 389-391.

(194) Markham, K. R., Ternai, B., Stanley, R., Geiger, H., Mabry, T. J., Carbon-13 NMR Studies of Flavonoids-III. *Tetrahedron,* **1978**, 34, 1389-1397.

(195) Markham, K. R., Chari, V. M., In 'The Flavonoids, Advances in Research,' Eds. J. B. Harbornes, T. J. Mabry, Chapman and Hall, London, **1982**.

(196) Barakat, H. H., El-Mousallamy, A. M. D., Souleman, A. M. A., Awadalla, S., Flavonoids of *Ochradenus baccatus*. *Phytochemistry,* **1991**, 30, 3777-3779.

(197) Mathes, H. W. D., Luu, B., Ourisson, G., Cytotoxic components of *Zingiber zerumbet*, *Curcuma zedoaria* and *C. domestica*. *Phytochemistry,* **1980**, 19, 2643-2646.

(198) Achari, B., Chaudhuri, C., Saha, C. R., Dutta, P. K., Pakrashi, S. C., A clerodane diterpene and other constituents of *Clerodendron inerme*. *Phytochemistry,* **1990**, 29, 3671-3673.

(199) Ulubelen, A., Topcu, G., Kaya, U., Steroidal compounds from *Teucrium chamaedrys* subsp. *chamaedrys*. *Phytochemistry*, **1994**, 36, 171-173.

(200) Mabry, T. J., Kagan, J., Rosler, H., NMR spectra of trimethylsilyl ethers of flavonoid glycosides. *Phytochemistry*, **1965**, 4, 177-183.

(201) Leitao, S. G., Monache, F. D., 2''-*O*-caffeoylorientin, from *Vitex polygama*. *Phytochemistry*, **1998**, 49, 2167-2169.

(202) Abou-Zahid, M. M., Lombardo, D. A., Kite G. C., Grayer, R, J., Veitch, N. C., Acylated flavone *C*-glycosides from *Cucumis sativus*. *Phytochemistry,* **2001**, 58, 167-172.

(203) Kato, T. and Morita, Y., *C*-glycosylflavones with acetyl substitution from Rumex acetosa L. *Chemical and Pharmaceutical Bulletin,* **1990** 38, 2277-2280.

(204) Kuo, S-H., Yen, M-H., Chung, M-I., Lin, C-N, A flavone *C*-glycosides and an aromatic glucosides from *Gentiana* species. *Phytochemistry,* **1996,** 41, 309-312.

(205) Kumazawa, T., Minatogawa, T., Matsuba, S., Sato, S., Onodera, J-I., An effective synthesis of isoorientin: the regioslective synthesis of a 6-*C*-glucosyl flavone. *Carbohydrate Research,* **2000**, 329, 705-513.

(206) Goetz, M, Jacot-Guillarmod, A., Phytochemistry of genus gentiana. 24 new *C*-glycosylflavones from leaves of *Gentiana-asclepiadeal* L. *Helvetica Chimica Acta,* **1978**, 61, 1373-1375.

(207) Sritularak, B., Likhitwitayawuid, K., Conrad, J., Kraus, W., Flavonoids from the roots of *Milletia erythrocalyx Phytochemistry,* **2002,** 61 943-947.

(208) Murphy, B. T., Cao, S., Norris, A., Miller, J. S., Ratovoson, F., Andriantsiferana, R., Rasamison, V. E., Kingston, D. G. I., Cytotoxic Flavanones of *Schizolaena hystrix* from Madagascar Rainforest. *Journal of Natural Products,* **2005**, 68, 417-419.

(209) Bergman, M., Varshavsky, L., Gottlieb, H. E., Grossman, S., The antioxidant activity of aqueous spinach extract: chemical identification of active fractions. *Phytochemistry*, **2001**, 58, 143-152.

(210) Bashir, A., Hamburger, M., Gupta, M. P., Solist, P. N., and Hostettmann, Flavonol glycosides from *Monnina sylvatica. Phytochemistry,* **1991**. 30, 3781-3784.

(211) Dewick, P. M., In; Harborne, J. B. (Eds.), The Flavonoids-Advances in Research Since 1986. Springer, New York, **1994**, 117-118.

(212) Chacha, M., Bojase-Moleta, G., Majinda, R. R. T., Antimicrobial and radical scavenging flavonoids from the stem wood of *Erythrina latissima. Phytochemistry,* **2005**, 66, 99-104.

(213) Wu, Y-Ch., Chang, F-R., Chen, Ch-Y., Tryptamine-derived amides and alkaloids from the seeds of *Annona atemoya. Journal of Natural Products,* **2005**, 68, 406-408.

(214) Jenett-Siemes, K., Weigl, R., Kaloga, M., Schula, J., Eich, E., Ipobscurines C and D: macrolactam-type indole alkaloids from the seeds of *Ipomoea obscura. Phytochemistry,* **2003**, 62, 1257-1263.

(215) Sakamura, S., Teravama, Y., Kawakatsu, S., Ichihara, A., Saito, H., Conjugated serotins related to carthrtic activity in safflower seeds (*Carthamus tinctorius* L.). *Agricultural Biological Chemistry,* **1980**, 44, 2951-2954.

(216) Mohamadi, F., Richards, N.G., Guida, W. C., Liskamp, R., Lipton, M., Caulfield, C., Chang, G., Hendrickson, T., Still, W. C., Macromodel-an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *Journal of Computer Chemistry,* **1990**, 11, 440-467.

(217) Still, W. C., Tempczyk, A., Hawley, R. C., Hendrickson, T., Seminanalytical treatment of solvation for molecular mechanics and dynamics. *Journal of American Chemical Society,* **1990**, 112, 6127-6129.

(218) Meyer, B. N., Ferrigni, R. N., Putnam, J. E., Jacobsen, L. B., Nichols, D. E. Mclaughlin, J. L., Brine Shrimp- A convenient general bioassay for active plants constituents. *Planta Medica,* **1982**, 45, 31-34.

(219) Mosmann, T., Rapid colorimetric assay for the cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of Immunological Methods,* **1983**, 65, 55-63.

(220) Takao, T., Watanabe, N., Yagi, I., Sakata, K., A simple screening method for antioxidants and isolation of several antioxidants produced by marine bacteria from fish and shellfish. *Biosci. Biotech. Biochem.* **1994**, 58, 1780-1783.

(221) Parra, A. L., Yhebra, R. S., Sardinas, I. G., Buela, L. I., Comparative study of the assay of *Artemina salina L.* and the estimate of the medium lethal dose (LD_{50} value) in mice, to determine oral acute of plant extracts. *Phytomedicine,* **2001**, 8, 395-400.

(222) Ferrigni, N. R., McLaughlin, J. L., Powell, R. G., Smith, C. R., Jr., Use of Potato Disc and Brine Shrimp Bioassay to Detect Activity and Isolate Piceatannol as the Antileukemic Principles from the Seeds of Euphorbia lagascae. *Journal of Natural Products,* **1984**, 47, 347-352.

(223) Dagne, E., Gunatilaka, A. A., Asmellash, S., Abate, D., Kingston, D. G. I., Hofmann, G. A., Johnson, R. K., Two new cytotoxic cytochalasins from *Xylaria obovata. Tetrahedron,* **1994**, 50, 5615-5620.

(224) Greca, M. D., Fiorentino, A., Mangoni, L., Molinaro, A., Monaco, P., Previtera, L., Cytotoxic 9,10-Dihydrophenanthrenes from *Juncus efusus* L. *Tetrahedron,* **1993**, 49, 3425-3432.

(225) McLaughlin, J. L., Crown gall tumours on potato discs and brine shrimp lethality: two simple bioassyas, for higher plant screening and fractionation. In: Hostettman, K., editor. Methods in plant biochemistry, Academic Press, London, **1991**.

(226) Kumarasamy, Y., Cox, P. J., Jaspars, M., Nahar, L., Sarker, S. D., Isolation, structure elucidation and biological activity of hederacine A and B, two unique alkaloids, from *Glechoma hederaceae*. *Tetrahedron*, **2003**, 59, 6403-6407.

(227) http://www.ehabsoft.com/ldpline

(228) Finney, D., J., Probit Analysis, 3rd ed., Cambridge University Press Cambridge, **1971**.

(229) Untch, M., Sevin, B.U., Perras, J. P., Angiol, R., Untch, A., Hightower, R. D., Koechli, O., Averette, H. E., Evaluation of paclitaxel (taxol), cisplatin, and the combination paclitaxel-cisplatin in ovarian cancer in vitro with ATP cell viability assay. *Gynecologic Oncology*, **1994**, 53, 44-49.

(230) Babich, H., Borenfreund, E., Cytotoxicity of T-2 toxin and its metabolites determined with neutral red cell viability assay. *Applied & Environmental Microbiology*, **1991**, 57 2101-2103.

(231) Korzeniewski, C., Callewaert, D. M., An enzyme-release assay for natural cytotoxicity. *Journal of Immunological Methods*, **1983**, 64, 313-320.

(232) O'Brien, J., Wilson, I., Orton, T., Pognan, F., Investigation of the alamar blue (resazurin) fluorescent dye for the assessement of mammalian cell toxicity. *European Journal of Biochemistry*, **2000**, 267, 5421-5426.

(233) Monks, A., Scudiero, D., Skehan, P., Shoemaker, R., Paull, K., Vistica, D., Hose, C., Langley, J., Cronise, P., VaigroWolff, A., Gray-Goodrich, M., Campbell, H., Mayo, J., Boyd, M., Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. *Journal of the National Cancer Institute*, **1991**, 83, 757-766.

(234) Hamid, R., Rotshteyn, Y., Rabadi, L., Parikh, R., Bullock, P., Comparison of alamar blue and MTT assays for high through-put screening, *Toxicology in vitro*, **2004**, 18, 703-710.

(235) Stajner, D., Millic, N., Canadanovic, B. J., An investigation into the antioxidant activity of *Allium nutans* L. *Phytotherapy Research*, **1999**, 13, 333-336.

(236) Troll, W., Wiesner, R., The roll of oxygen radicals as a possible mechanism of tumor promotion. *Annual review of Pharmacology and Toxicol*ogy, **1985**, 25, 509-528.

(237) Harman, D., Free radical theory of aging, increasing the functional life span. *Annals of the New York Academy of Sciences*, **1994**, 717, 1-15.

(238) Simonian, N. Y., Coyle, J. T., Oxidative stress in neurodegenerative disease. *Annual review of pharmacology and toxicology*, **1996**, 36, 83-106.

(239) Sies, H., Oxidative stress, third ed. Academic Press, New York, **1985**.

(240) Halliwell, R., Aruoma, O. I., DNA damage by oxygen-derived species. FEBS *Letters,* **1991**, 281, 9-19.

(241) Janseen, Y. M. W., Vanhouten, B., Borm, P. J. A., Mossman, B. T., Biology Disease-Cell and tissue responses to oxidative damage. *Laboratory Investigation,* **1993**, 69, 261-274.

(242) Johnson, I. T., Williamson, G., Musk, S. R., Anticarcinogenic factors in plants foods: a new class of nutrients. *Nutrition Reviews,* **1994**, 7, 175-2004.

(243) Dragsted. L. O., Strube, M., Larsen, J. C., Cancer-protective factors in fruits and vegetables: biochemical and biological background. *Pharmacology & Toxicology,* **1993**, 72, 116-135.

(244) Jang, M., Cai, L., Udeani, G. O., Pezzuto, J. M., Cancer chemopreventive activity of resveratrol, a natural products derived from grapes. *Science,* **1997**, 275, 218-220.

(245) Chanvitayapongs, S., Draezynska-Lusiak, B., Sun, A. Y., Amelioration of oxidative stress by antioxidants and resveratrol in PC12 cells. *Neuroreport,* **1997**, 8, 1499-1502.

(246) Clement, M. V., Hirpara, J. L., Chawdhury, S., Perviaz, S., Chemopreventive agent resveratrol, a natural product derived from grapes, triggers CD95 signaling-dependent apoptosis in human tumor cells. *Blood,* **1998**, 92, 996-1002.

(247) Kumarasamy, Y., Fergusson, M., Nahar, L., Sarker, S. D., Bioactivity of moschamindole from *Centaurea moschata. Pharmaceutical Biology,* **2002**, 40, 307-310.

(248) Chiang, Y. M., Chang, J. Y., Kuo, C. C., Chang, C. Y., Kuo, Y. H., Cytotoxic triterpenes from the aerial roots of *Ficus microcarpa. Phytochemistry,* **2005**, 66, 495-501.

(249) Nguyen, A. T., Fontaine, J., Malonne, H., Claeys, M., Luhmer, M., Duez, P., A sugar ester and an iridoid glycoside from *Scrophularia ningpoensis, Phytochemistry,* **2005**, 1186-1191.

(250) Koulman, A., Kubbinga, M. E., Batterman, S., Woerdenbag, H., J., Pras, N., Woolley, J. G., Quax, W. J., A phytochemical study of lignans in whole plants and cell suspension cultures of *Anthriscus sylvestris. Planta Medica,* **2003**, 69. 1186-1191.

(251) Nam, N-H., You, Y-J., Kim, Y., Hong, D-H., Kim, H-M., Ahn, B. Z., Syntheses of certain 3-aryl-2-propenoates and evaluation of their cytotoxicity. *Bioorganic & Medicinal Chemistry Letters*, **2001**, 11, 1173-1176.

(252) Nguyen, A. T., Sok, D. E., Michael acceptors as a tool for anticancer drug design. *Current Pharmaceutical Design,* **1996**, 2, 247-294.

Appendix A

NMR Spectra of Isolated Compounds (Figure 78-134)

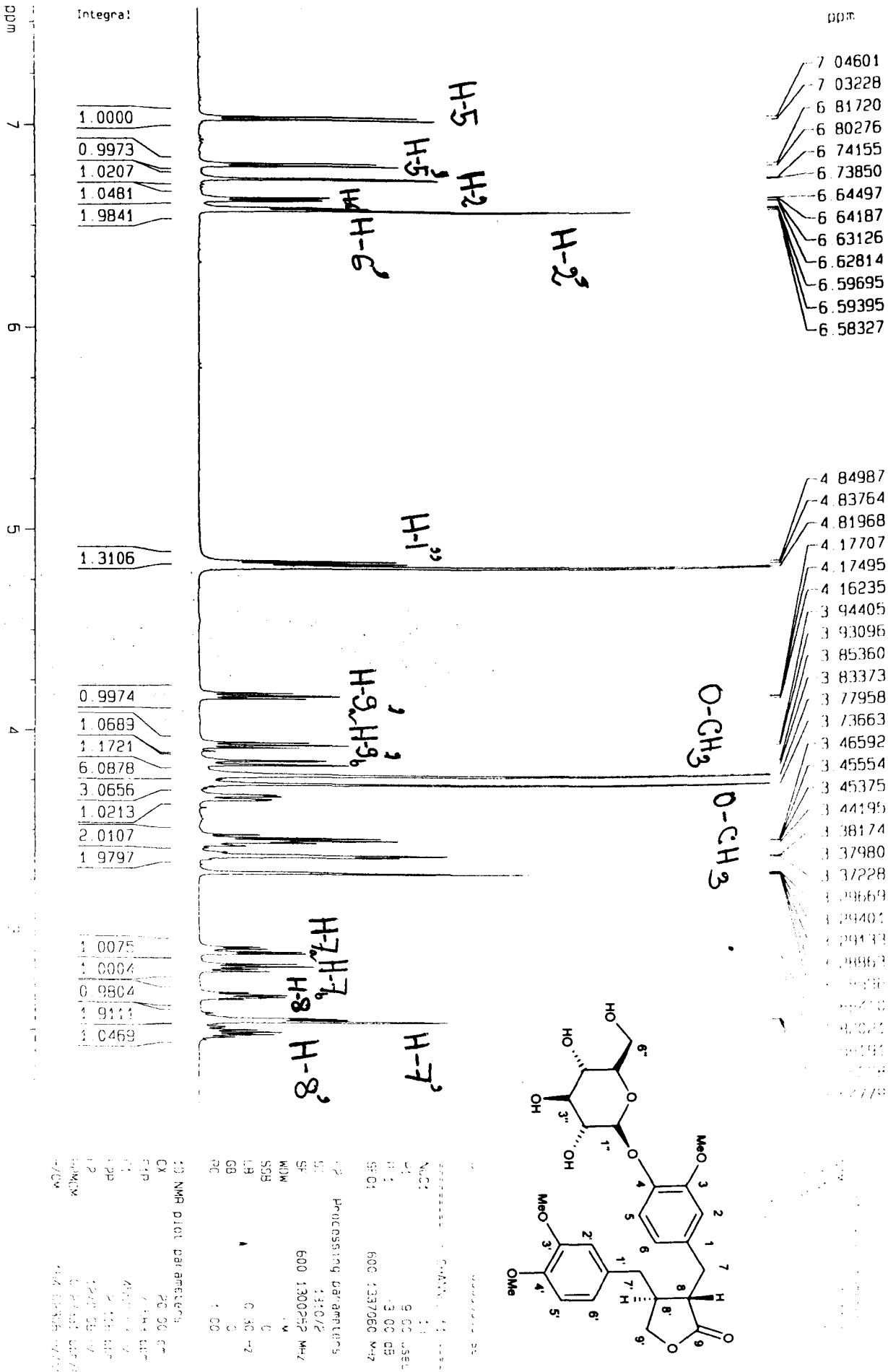


Figure 78. ^{1}H NMR spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

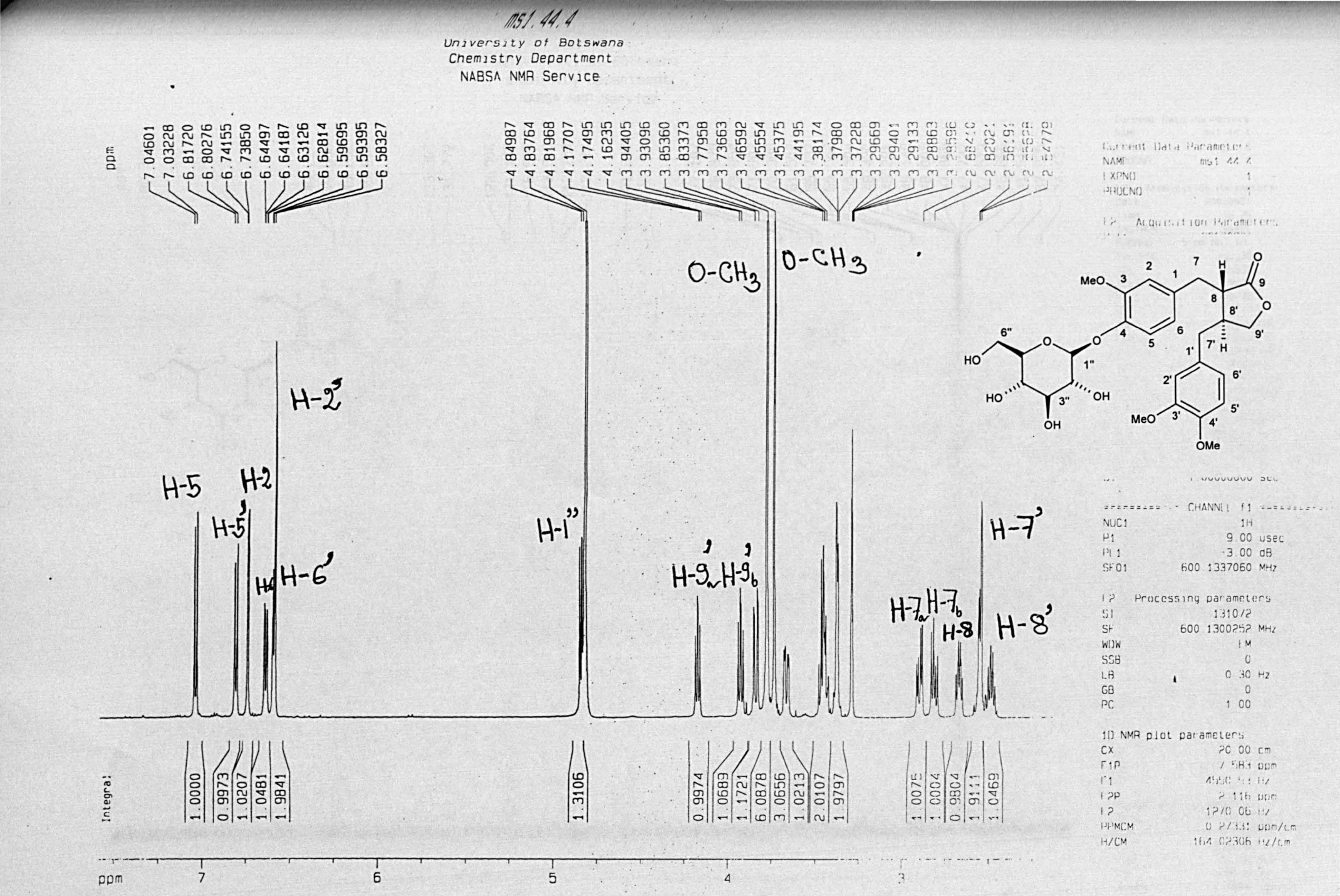
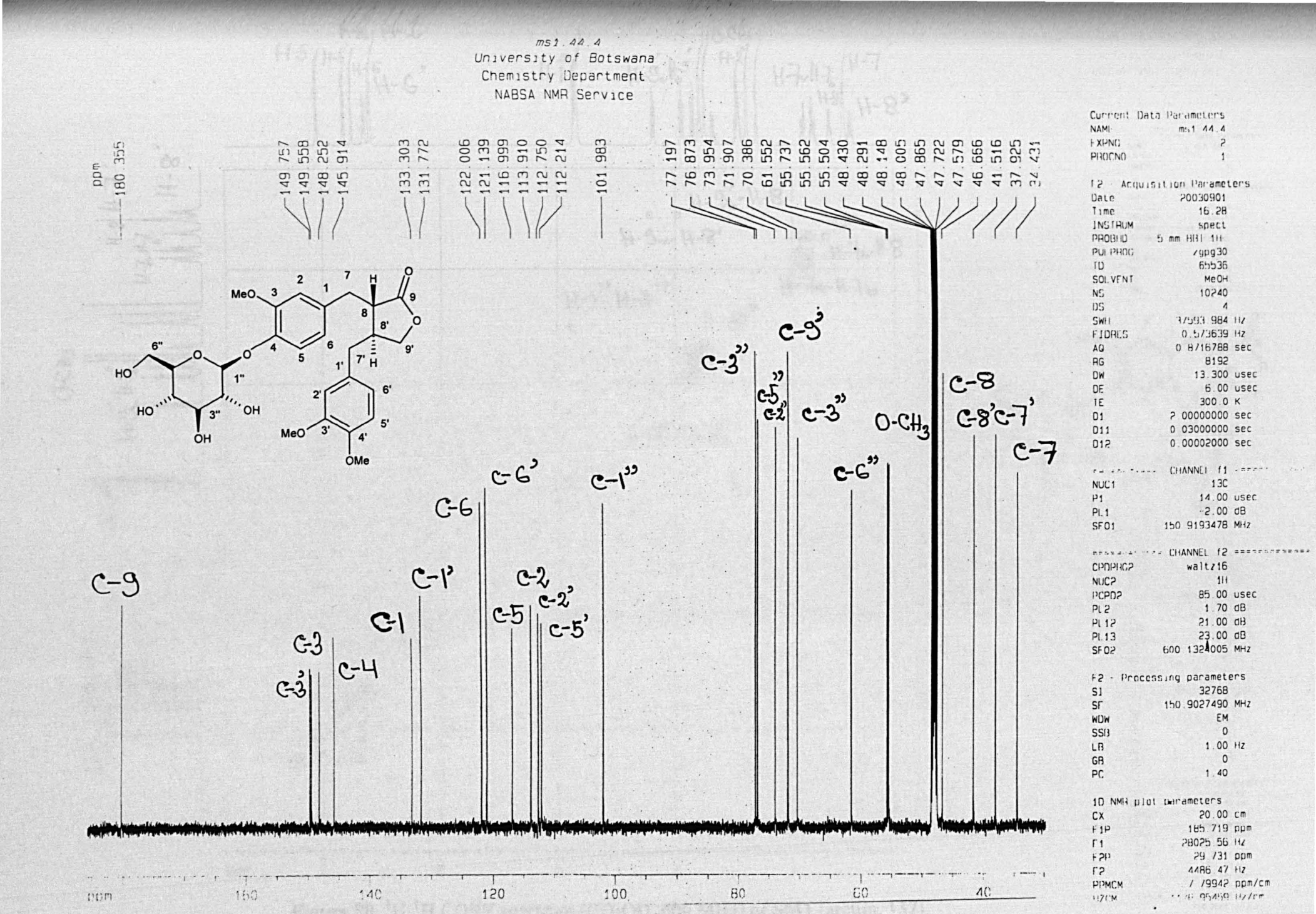


Figure 78. ^{1}H NMR spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)



Current Data Parameters

NAME	ms1.44.4
EXPNO	2
PROCNO	1

F2 - Acquisition Parameters

Date	20030901
Time	16.28
INSTRUM	spect
PROBHD	5 mm BBI 1H
PULPROG	zgpg30
TD	65536
SOLVENT	MeOH
NS	10240
DS	4
SWH	37593.984 Hz
FIDRES	0.573639 Hz
AQ	0.8716788 sec
RG	8192
DW	13.300 usec
DE	6.00 usec
TE	300.0 K
D1	2.00000000 sec
D11	0.03000000 sec
D12	0.00002000 sec

CHANNEL f1

NUC1	13C
P1	14.00 usec
PL1	-2.00 dB
SFO1	150.9193478 MHz

CHANNEL f2

CPDPRG2	waltz16
NUC2	1H
PCPD2	85.00 usec
PL2	1.70 dB
PL12	21.00 dB
PL13	23.00 dB
SFO2	600.1324005 MHz

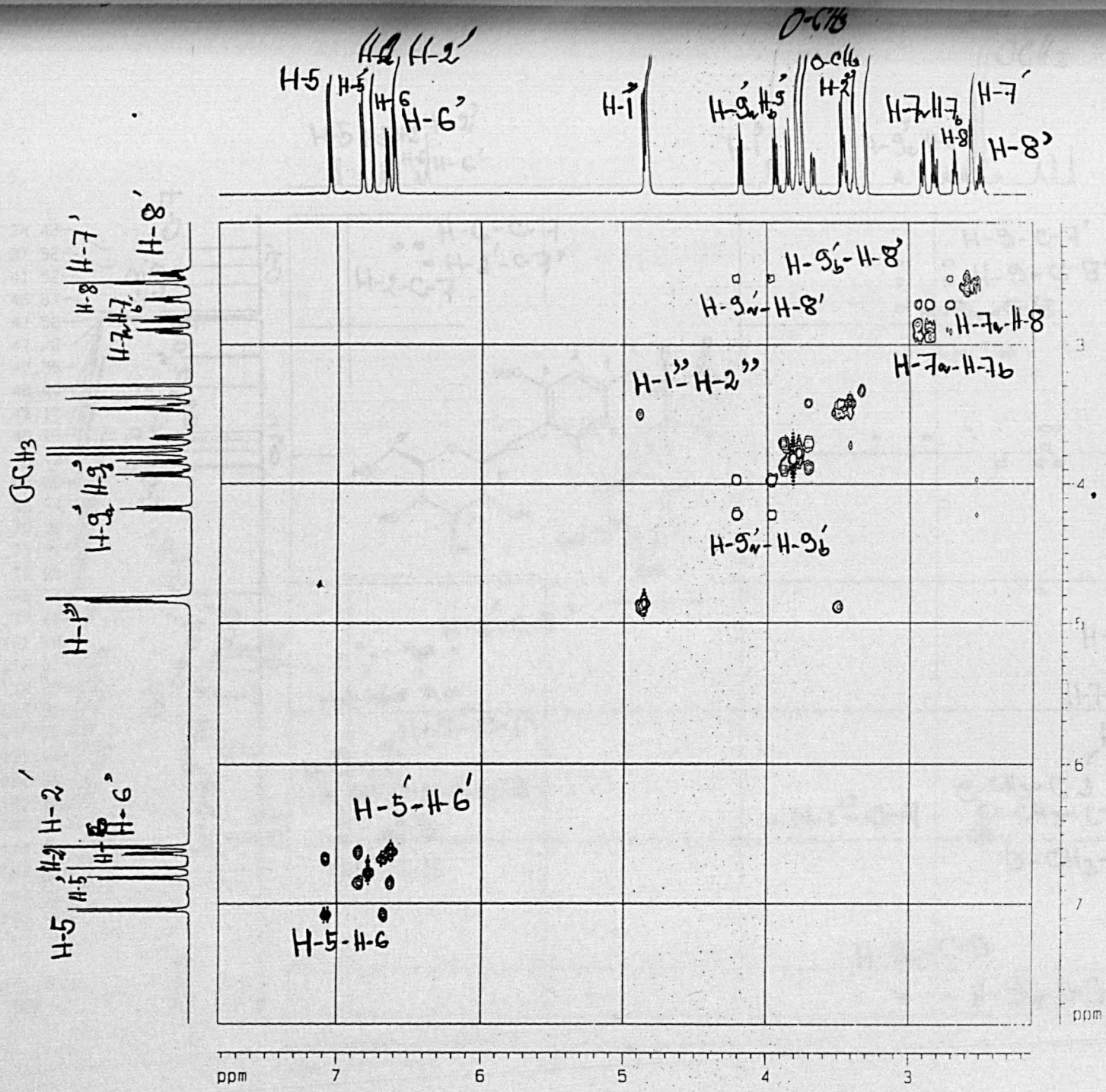
F2 - Processing parameters

SI	32768
SF	150.9027490 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

1D NMR plot parameters

CX	20.00 cm
F1P	185.719 ppm
F1	28025.56 Hz
F2P	29.731 ppm
F2	4486.47 Hz
PPMCM	7.79942 ppm/cm
HZCM	1176.95459 Hz/cm

Figure 79. ^{13}C NMR spectrum (CD_3OD, 150 MHz) of **SM1** (arctiin, **137**)



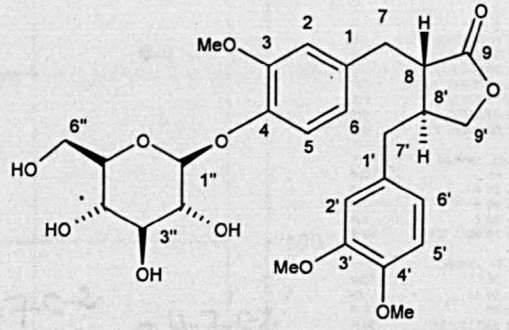


Figure 80. ^{1}H-^{1}H COSY spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

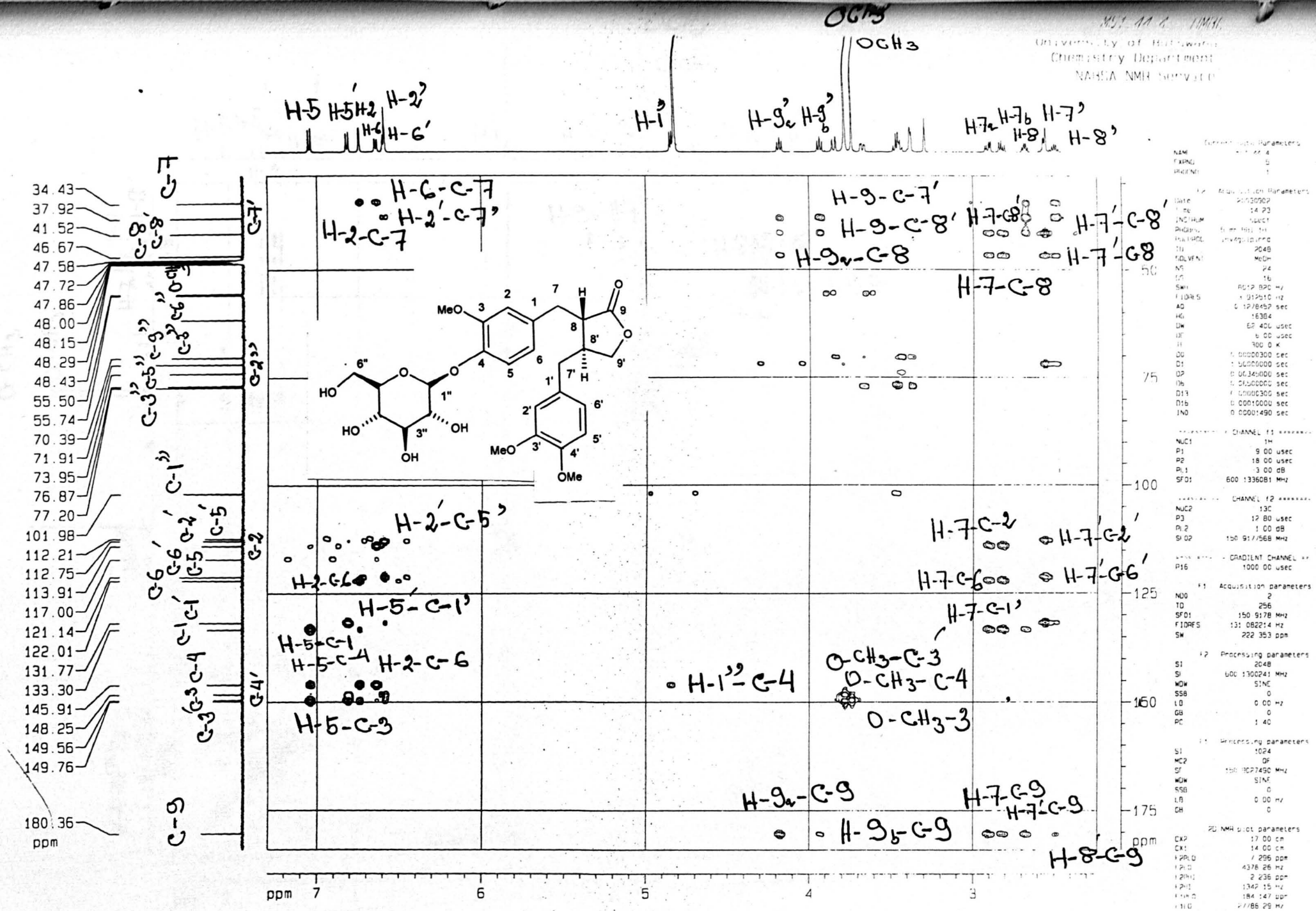


Figure 81. HMBC spectrum (CD_3OD, 600 MHz) of **SM1** (arctiin, **137**)

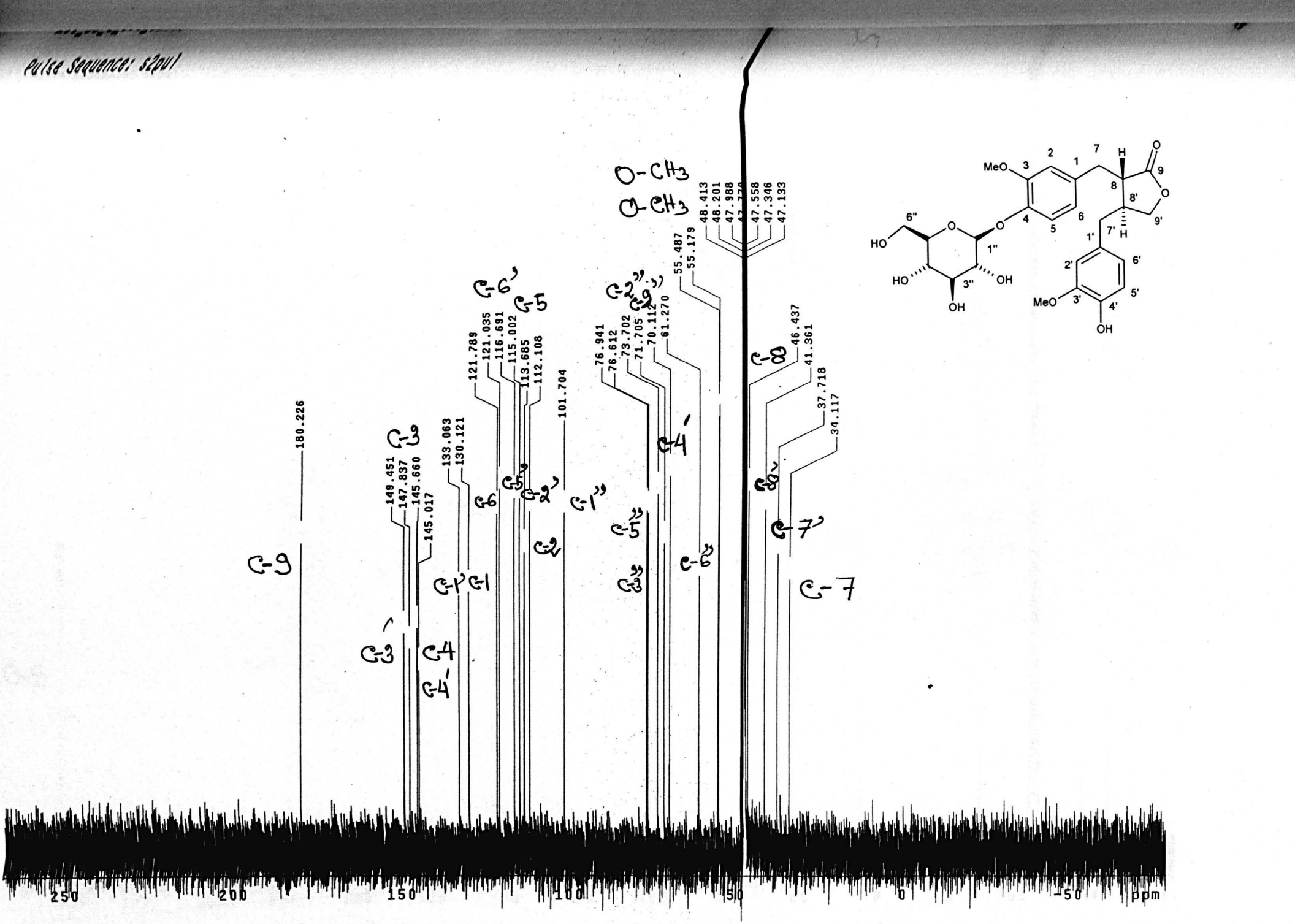


Figure 83. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM2** (matairesinoside, **138**)

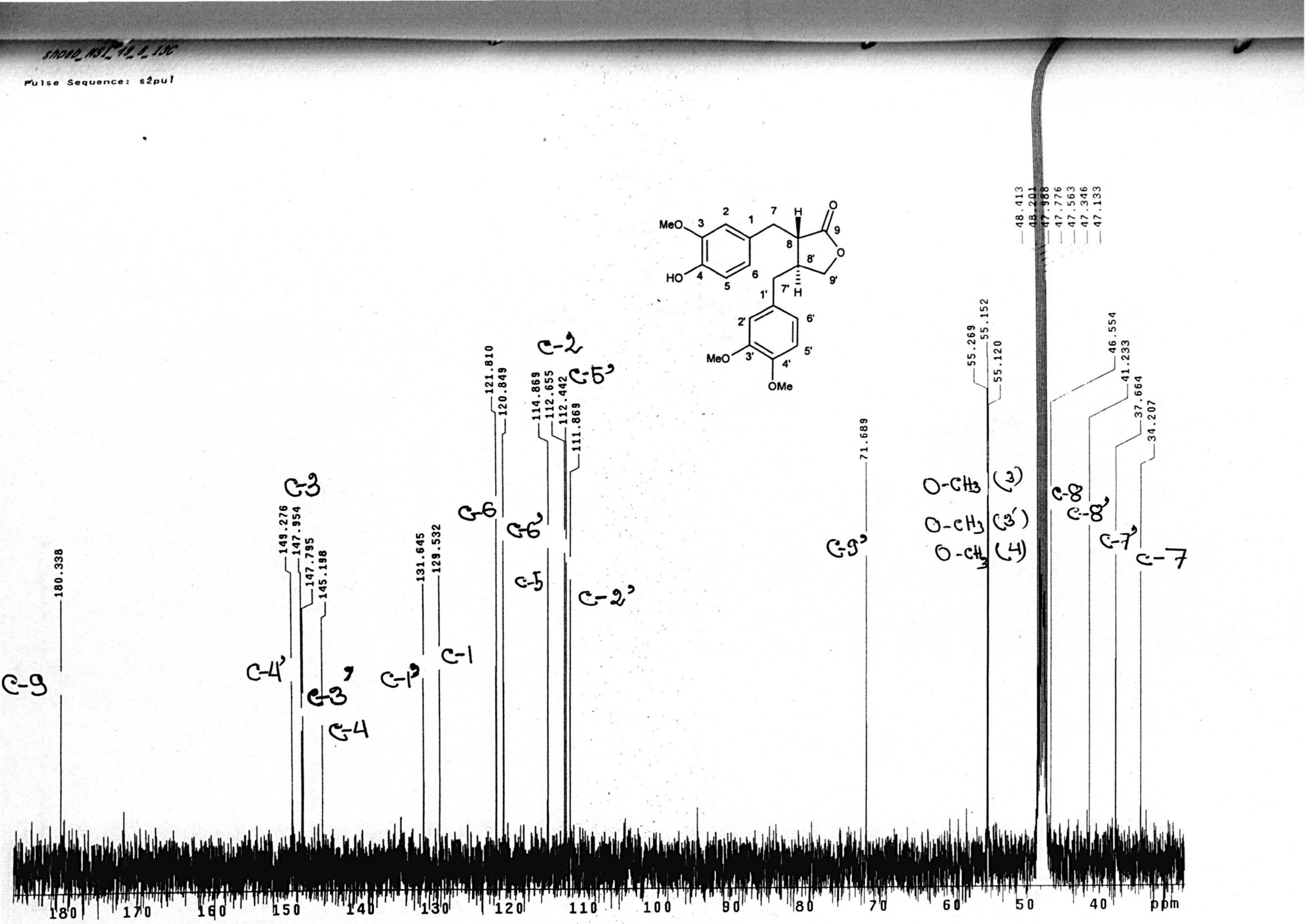


Figure 84. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM3** (arctigenin, **139**)

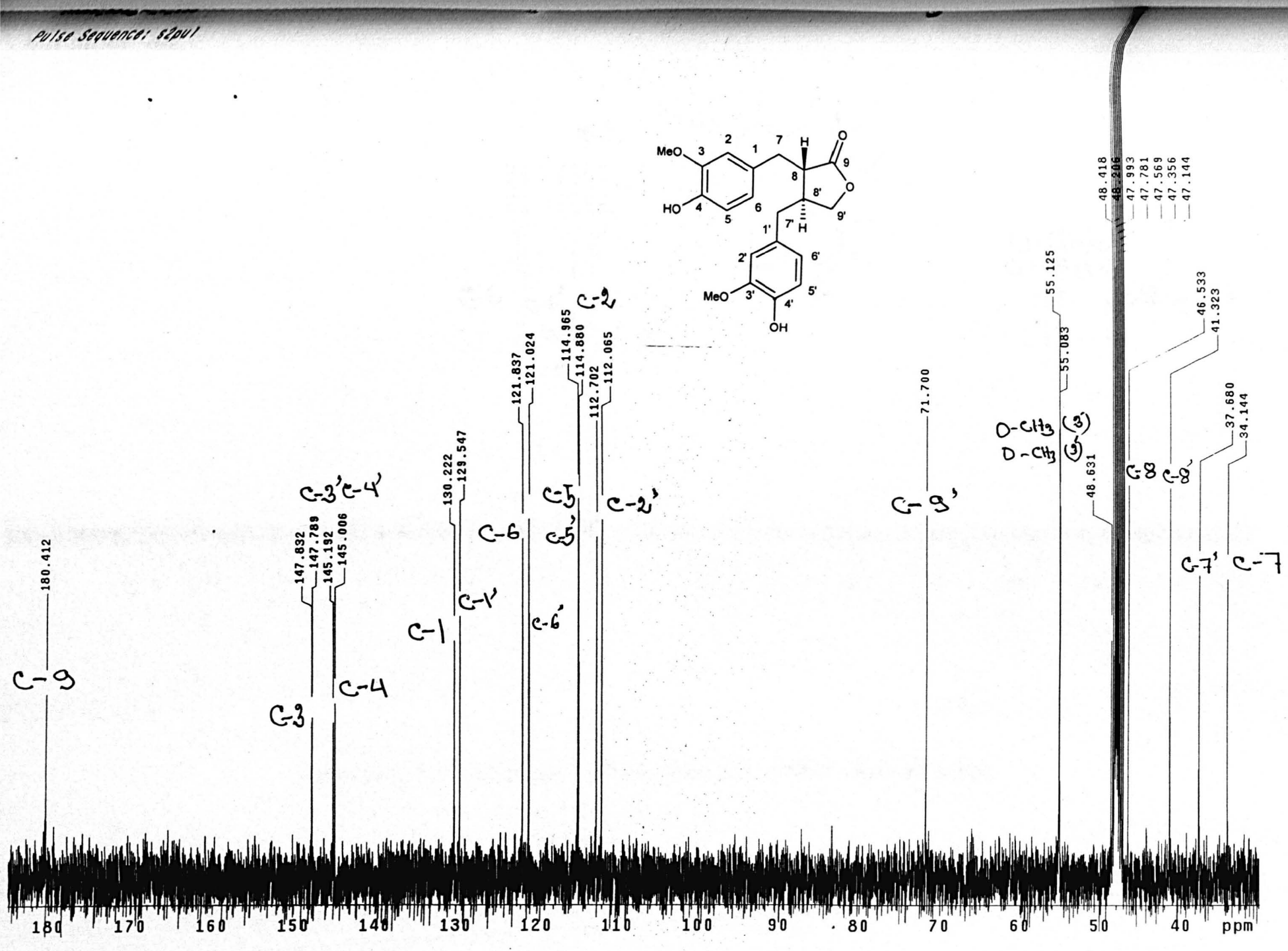


Figure 85. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM4** (matairesinol, **140**)

Pulse Sequence: dept

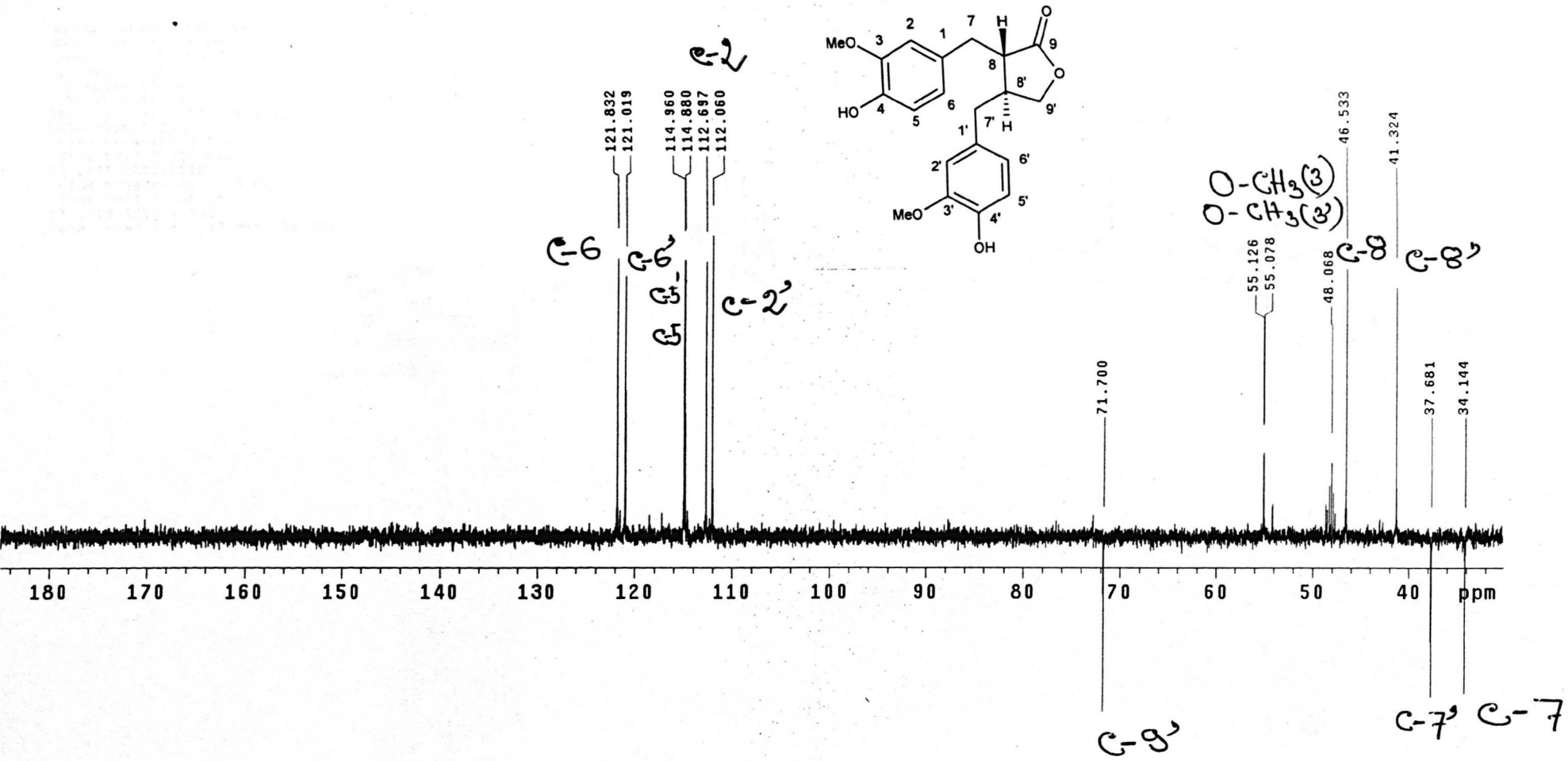


Figure 86. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM4** (matairesinol, **140**)

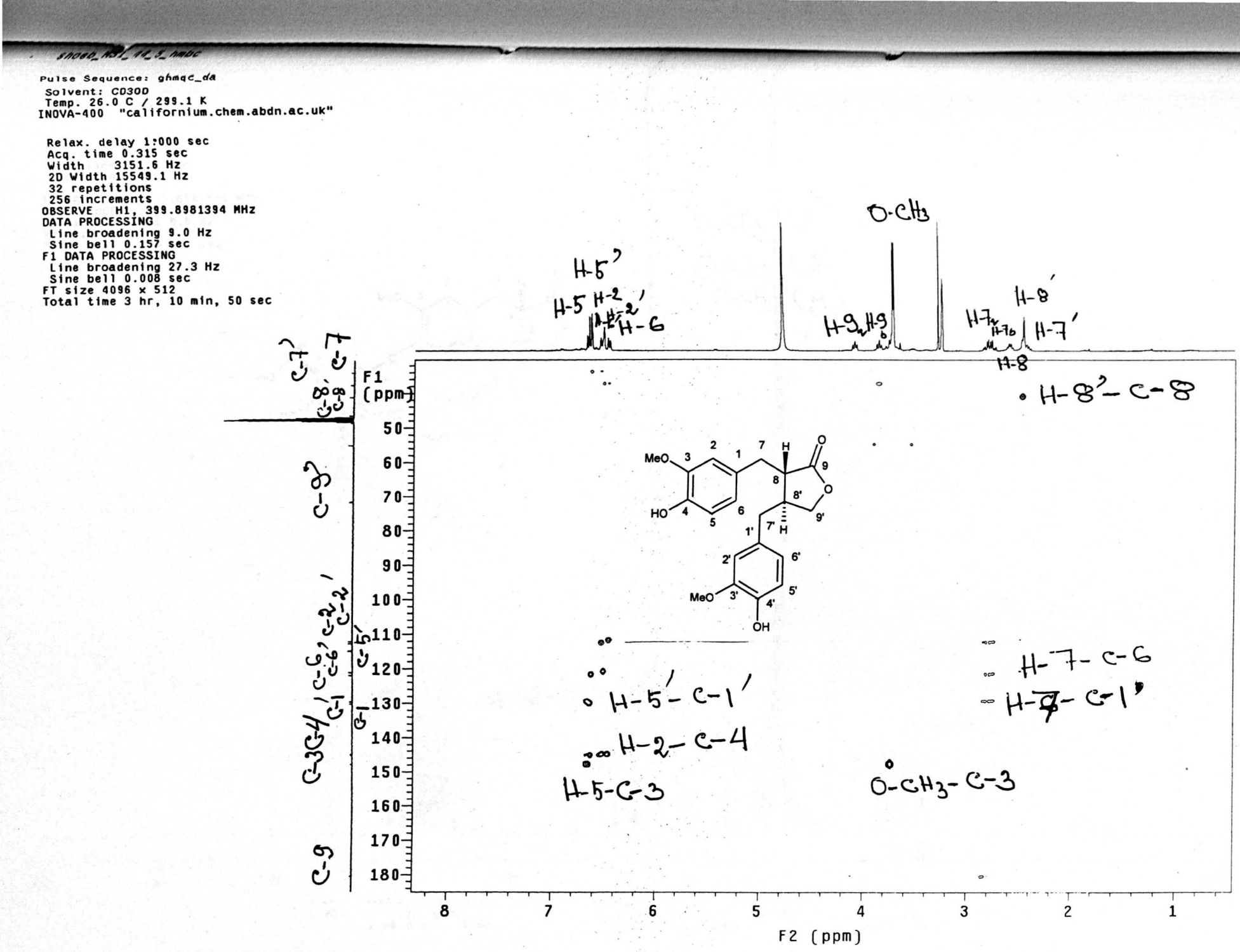


Figure 87. HMBC spectrum (CD_3OD, 400 MHz) of **SM4** (matairesinol, **140**)

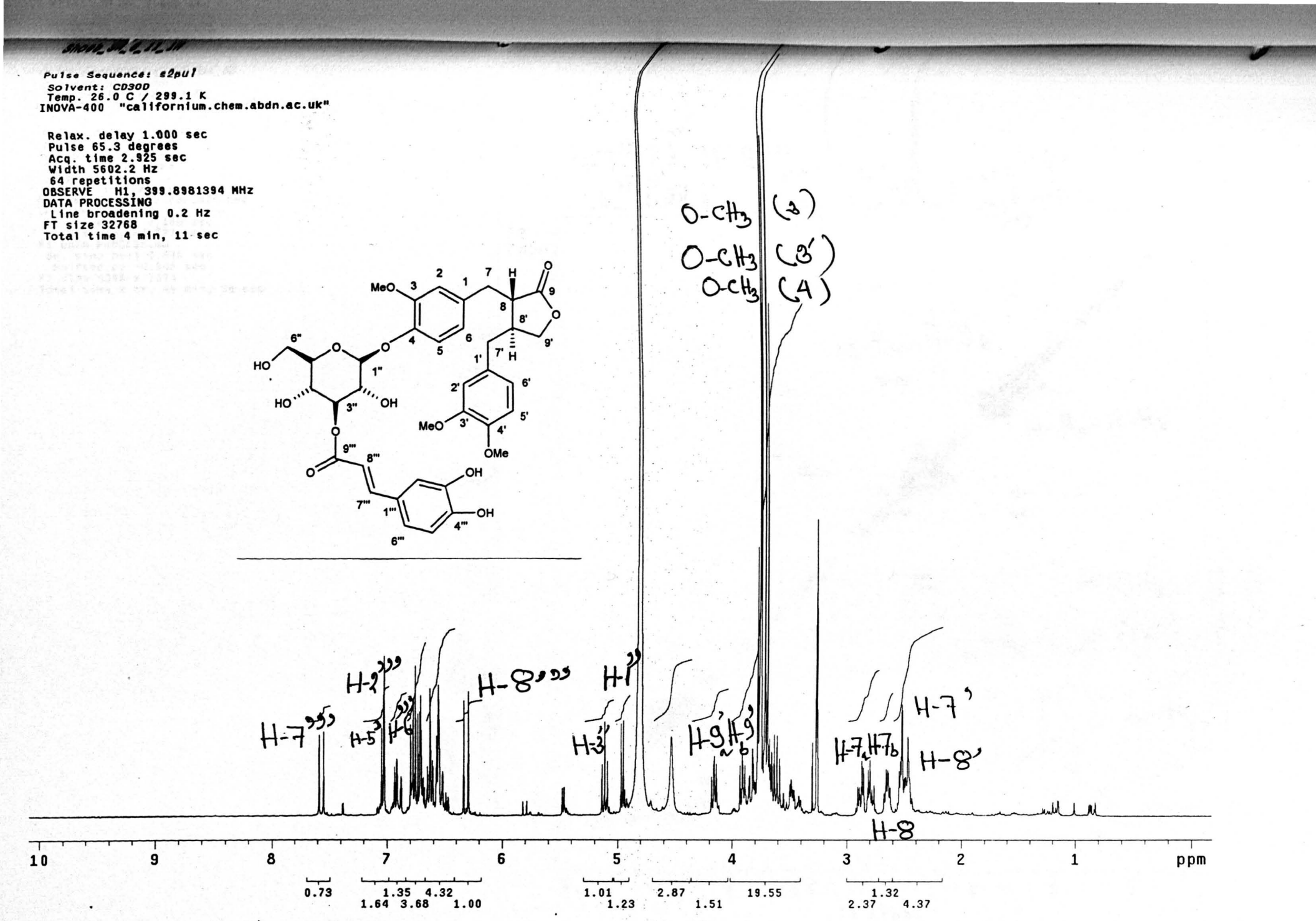


Figure 88. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

Pulse Sequence: gmqfcosy_dd
Solvent: cd3od
Temp. 26.0 C / 299.1 K
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.177 sec
Width 5602.2 Hz
2D Width 5602.2 Hz
16 repetitions
2 x 256 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Sq. sine bell 0.180 sec
Shifted by -0.177 sec
F1 DATA PROCESSING
Sq. sine bell 0.046 sec
Shifted by -0.046 sec
FT size 4096 x 1024
Total time 2 hr, 46 min, 36 sec

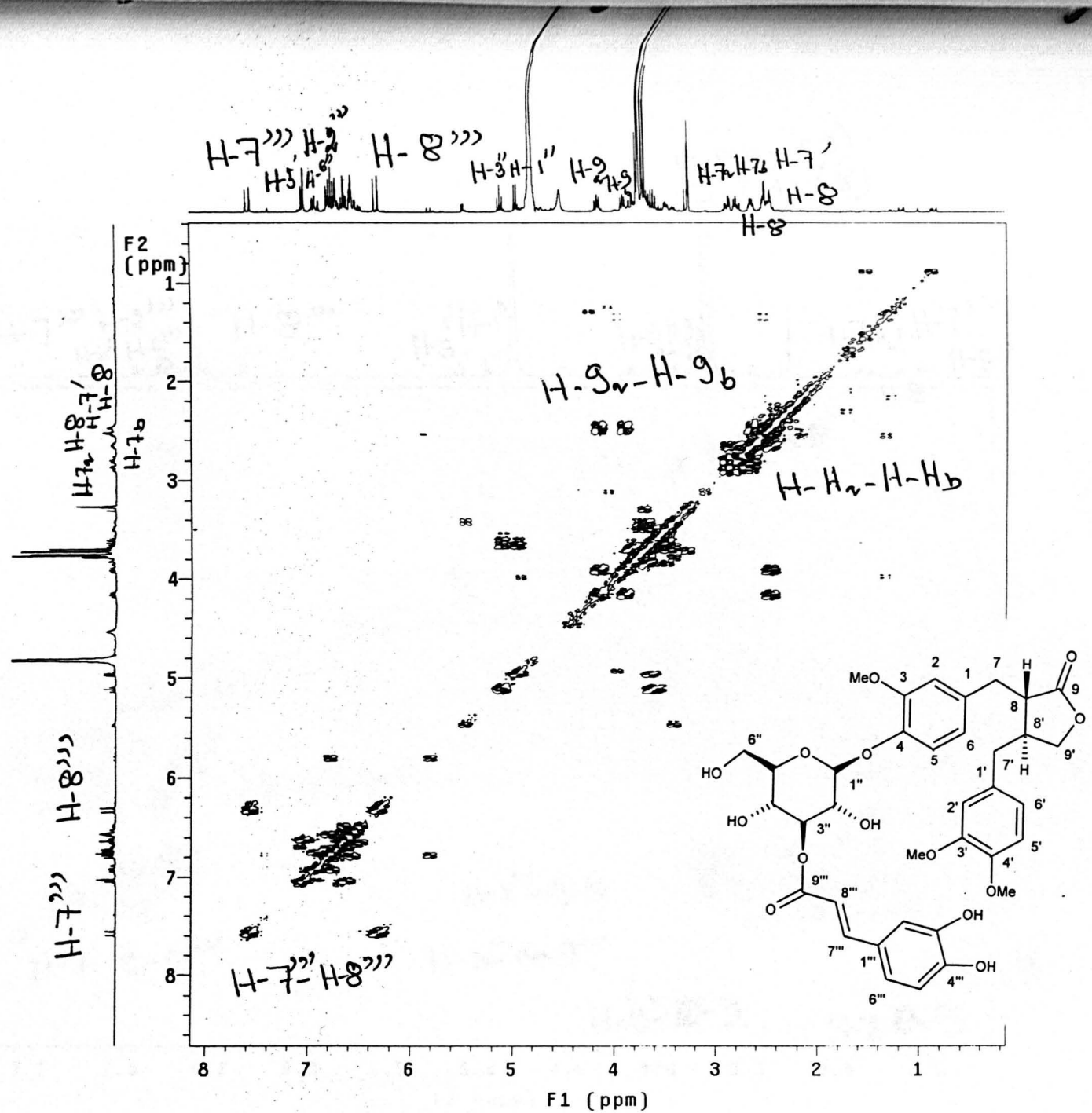


Figure 89. ^{1}H-^{1}H COSY spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

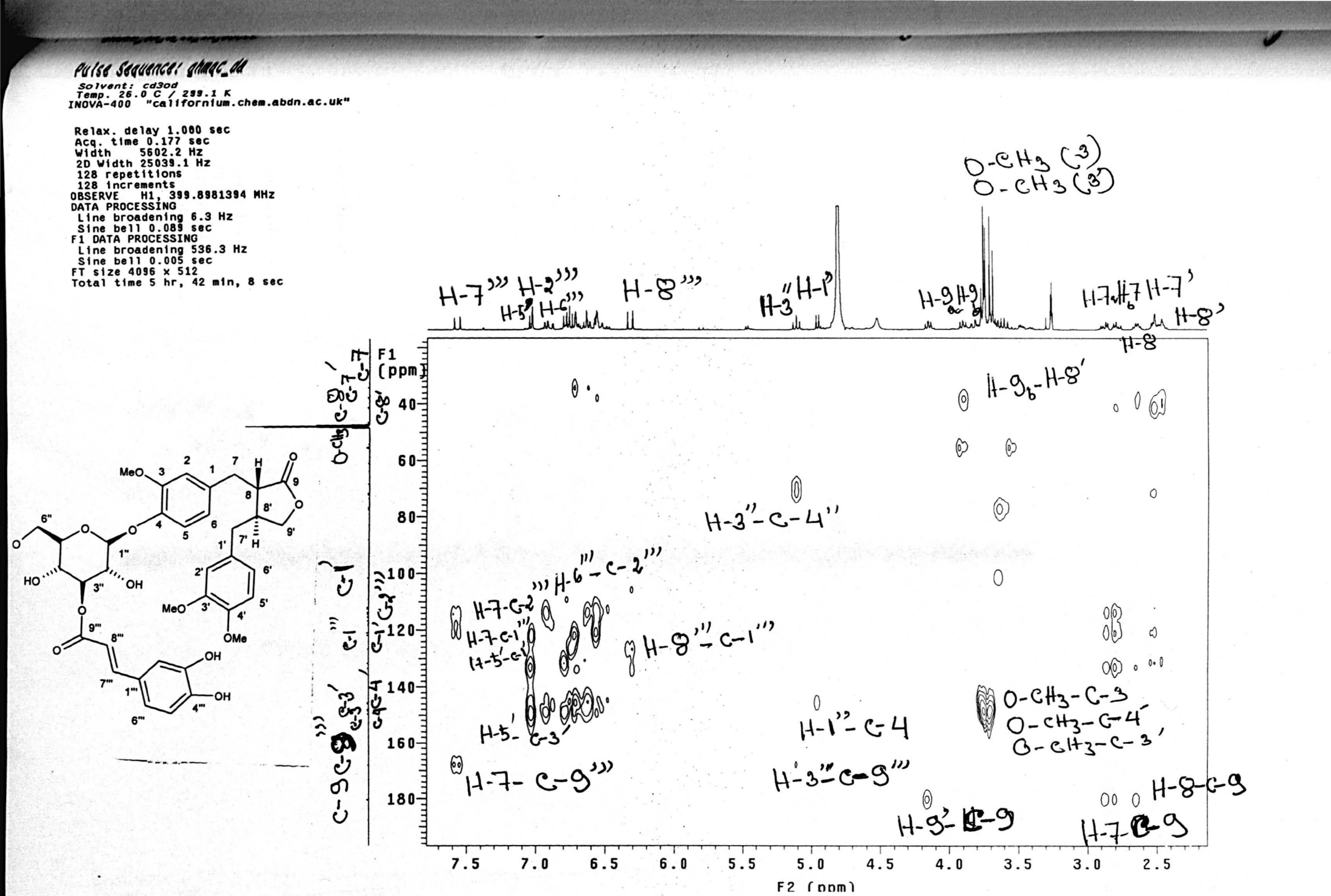


Figure 90. HMBC spectrum (CD_3OD, 400 MHz) of **SM5** (3''-*O*-caffeoyl-(9'''→3'')-arctiin, **170**)

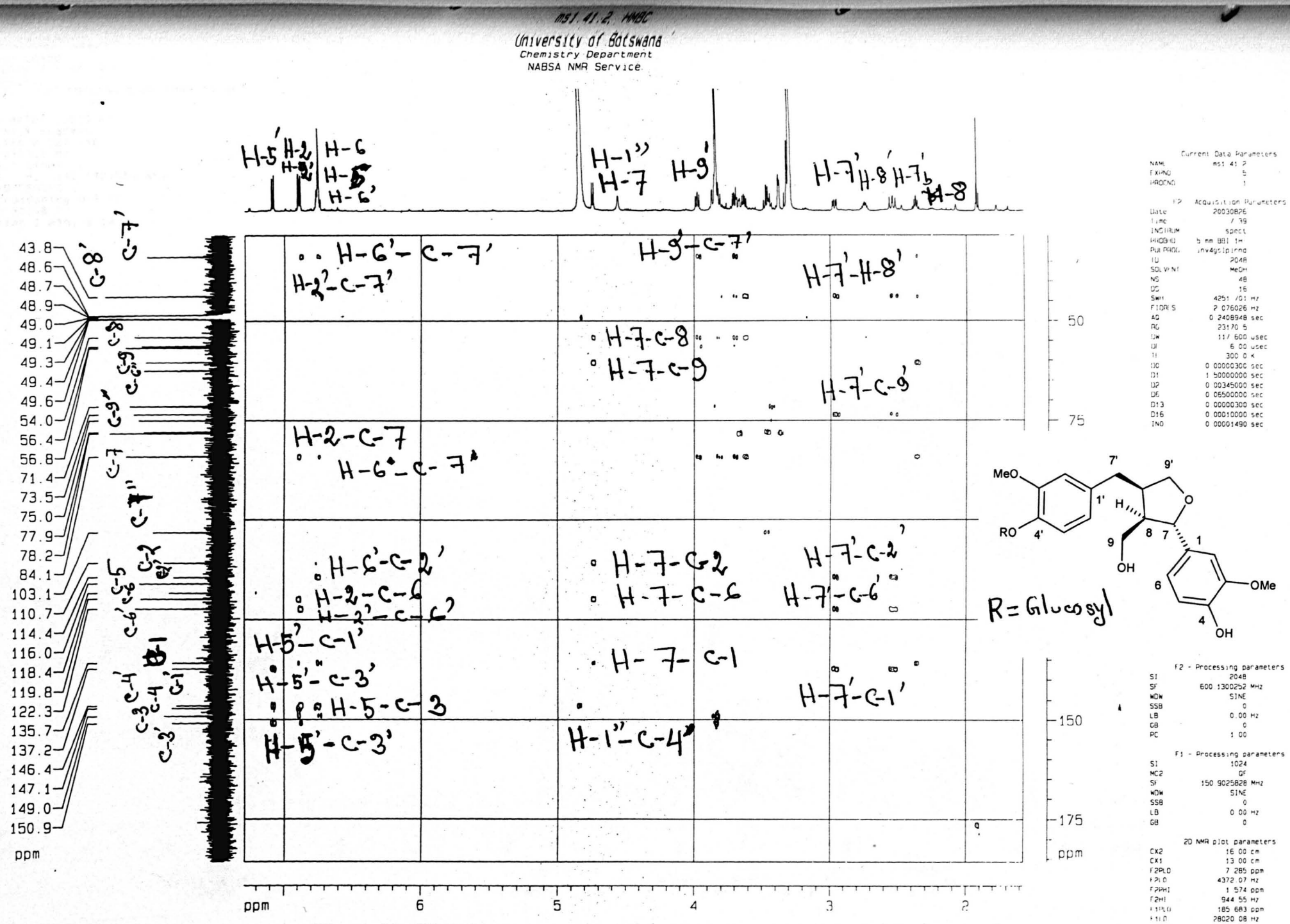


Figure 92. HMBC spectrum (CD_3OD, 400 MHz) of **SM7** (lariciresinol 4'-*O*-β-D-glucopyranoside, **171**)

MSL_41_3_CD3OD

Pulse Sequence: s2pul
Solvent: CD3OD
Temp. 26.0 C / 299.1 K
File: MS1_41_3
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Pulse 84.9 degrees
Acq. time 2.924 sec
Width 2871.0 Hz
16 repetitions
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 32768
Total time 1 min, 2 sec

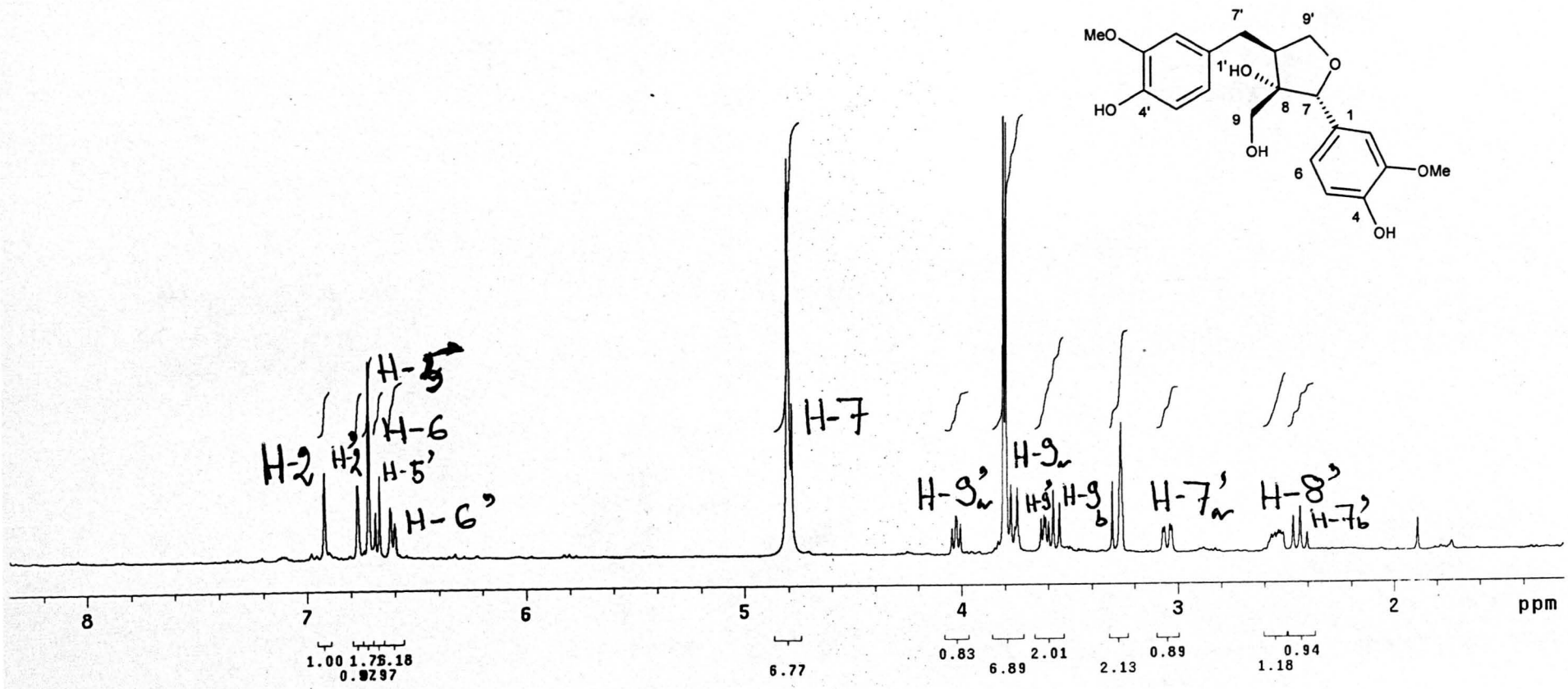


Figure 93. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM8** (berchemol, **172**)

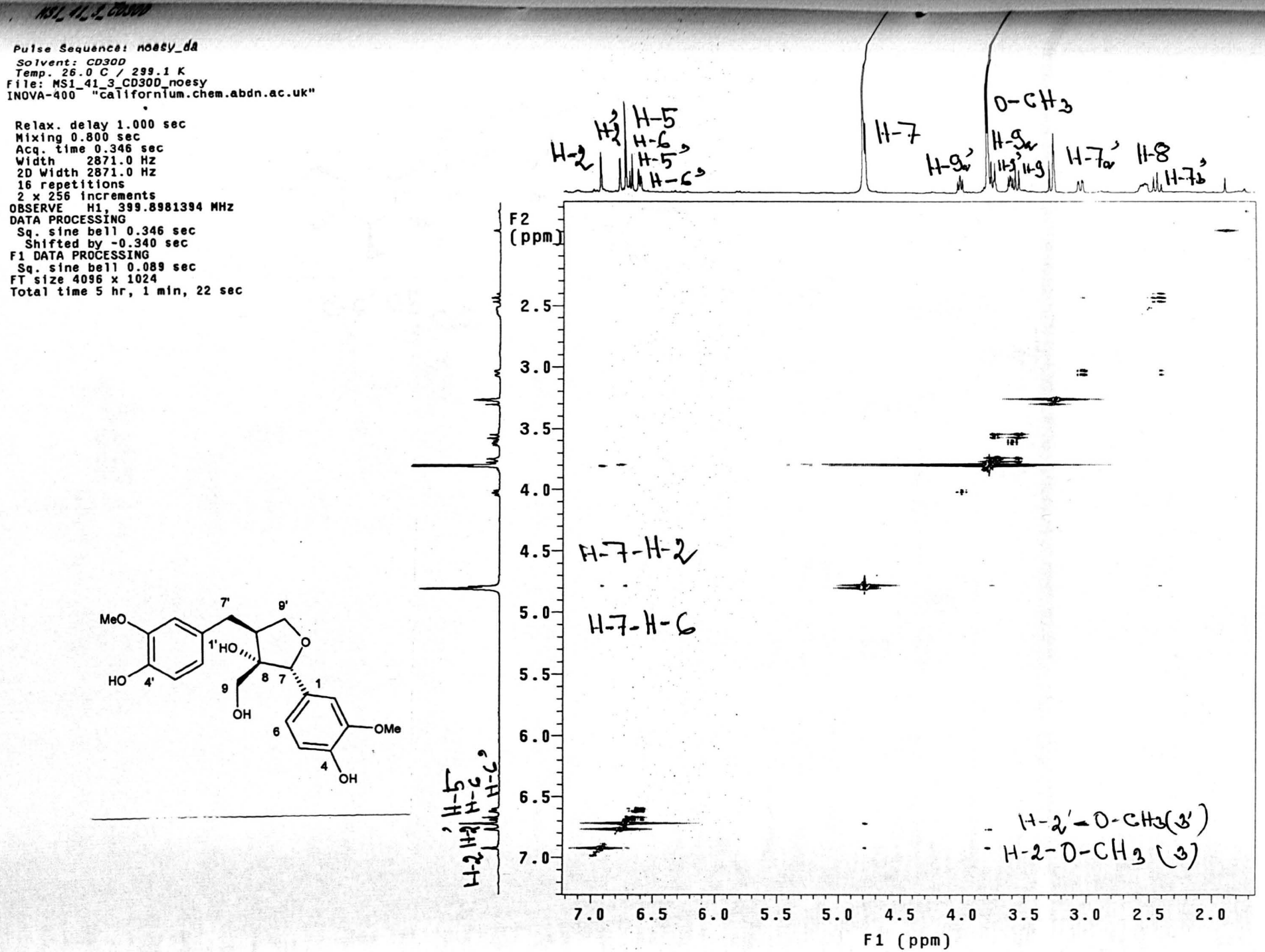


Figure 94. NOESY spectrum (CD_3OD, 400 MHz) of **SM8** (berchemol, **172**)

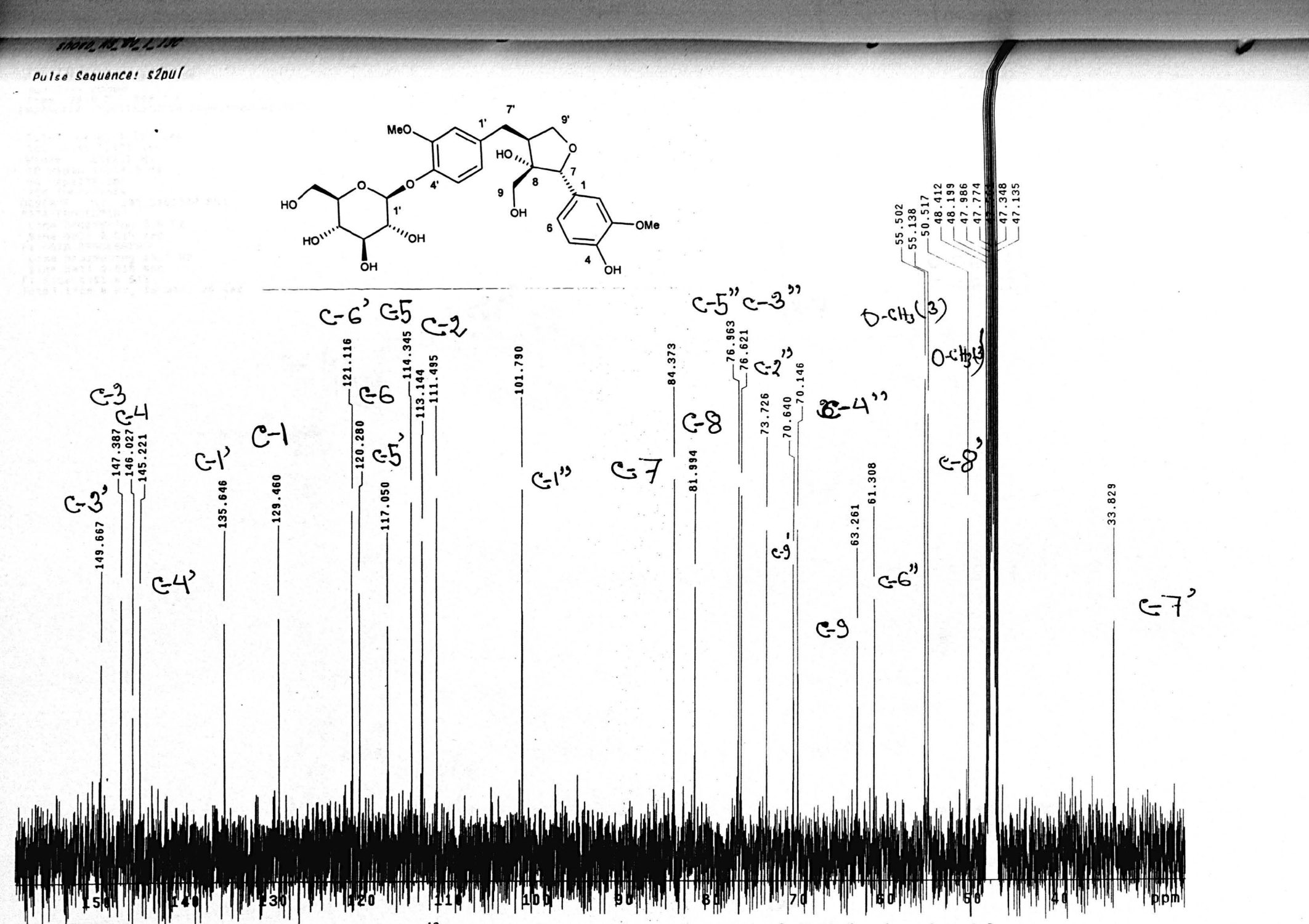


Figure 95. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM9** (berchemol 4'-*O*-β-D-glucopyranoside, **173**)

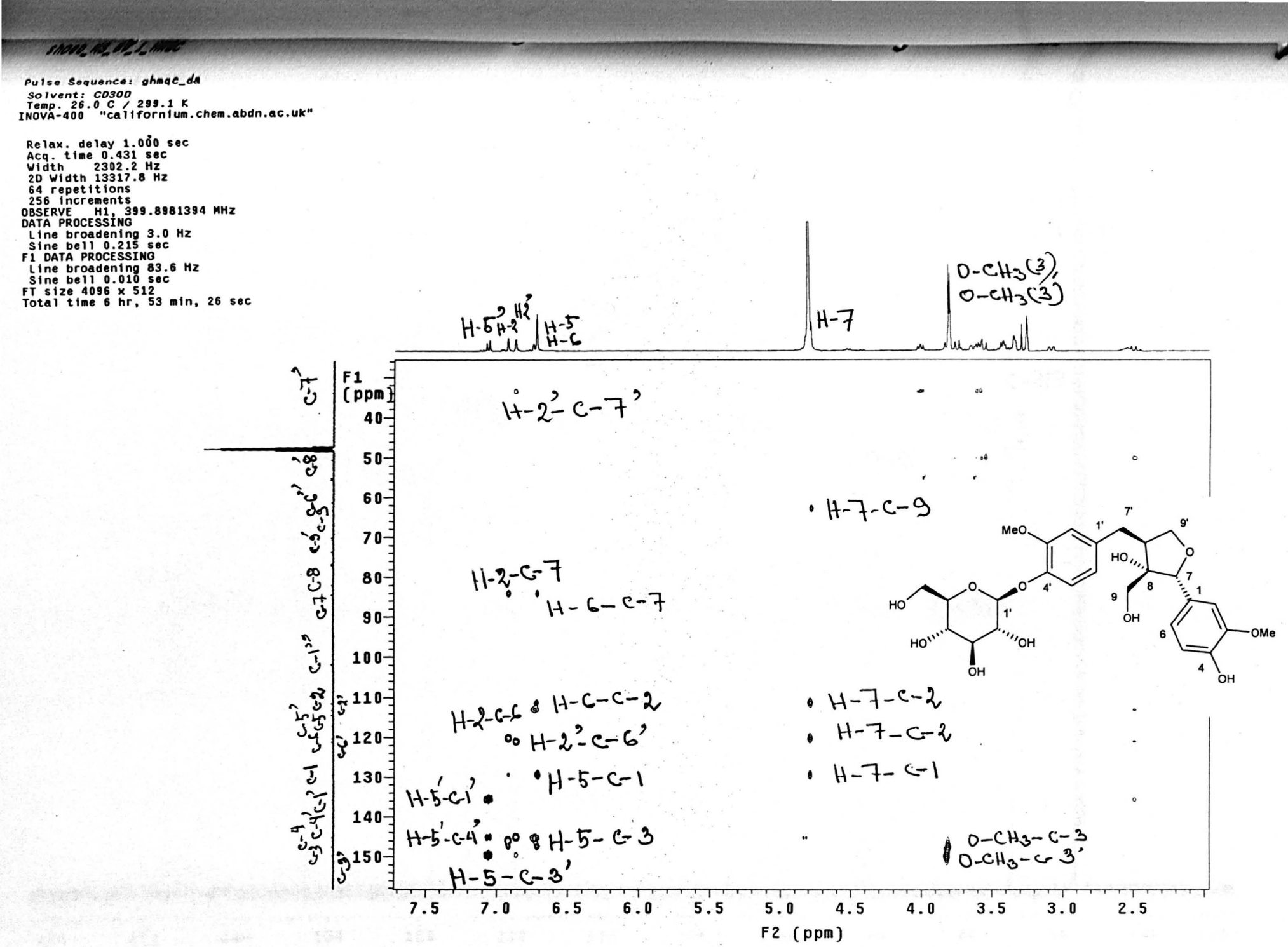


Figure 96. HMBC spectrum (CD_3OD, 400 MHz) of **SM9** (berchemol 4'-*O*-β-D-glucopyranoside, **173**)

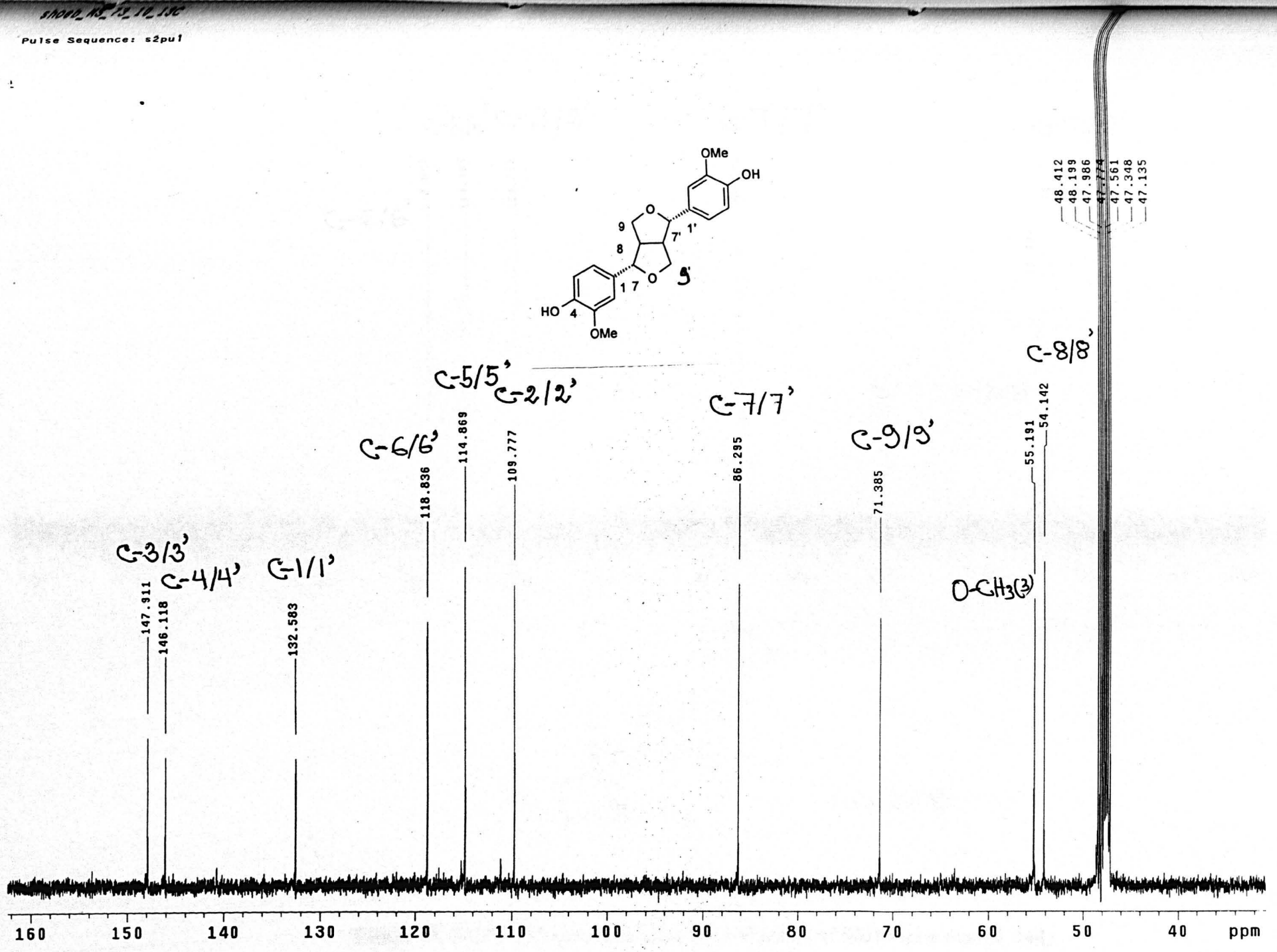


Figure 97: ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM10** (pinoresinol, **144**)

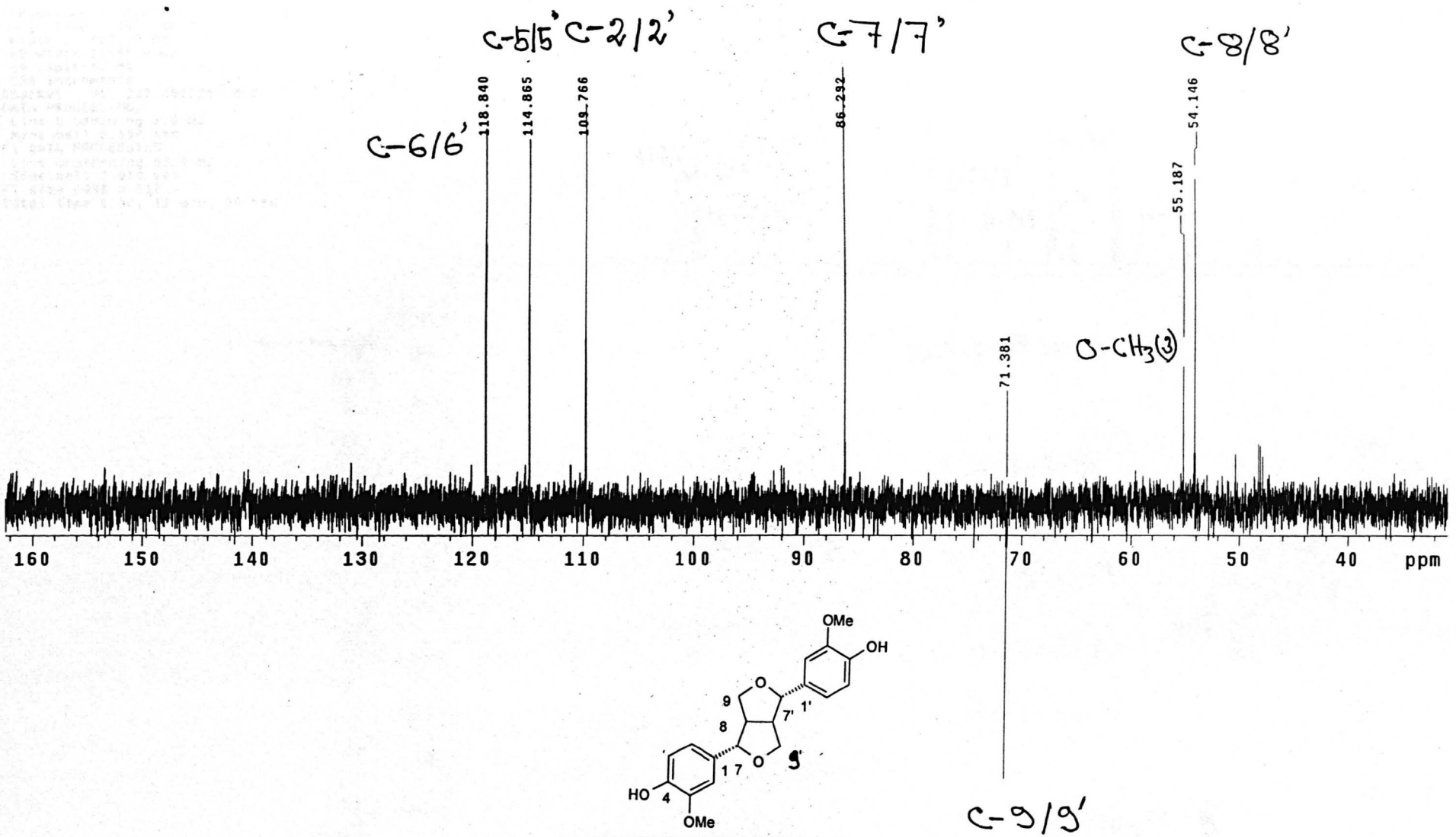


Figure 98. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM10** (pinoresinol, **144**)

Pulse Sequence: ghmqc_da
Solvent: CD3OD
Temp. 26.0 C / 299.1 K
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.307 sec
Width 3227.0 Hz
2D Width 12596.4 Hz
64 repetitions
256 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Line broadening 3.6 Hz
Sine bell 0.154 sec
F1 DATA PROCESSING
Line broadening 52.4 Hz
Sine bell 0.010 sec
FT size 4096 x 512
Total time 6 hr, 19 min, 50 sec

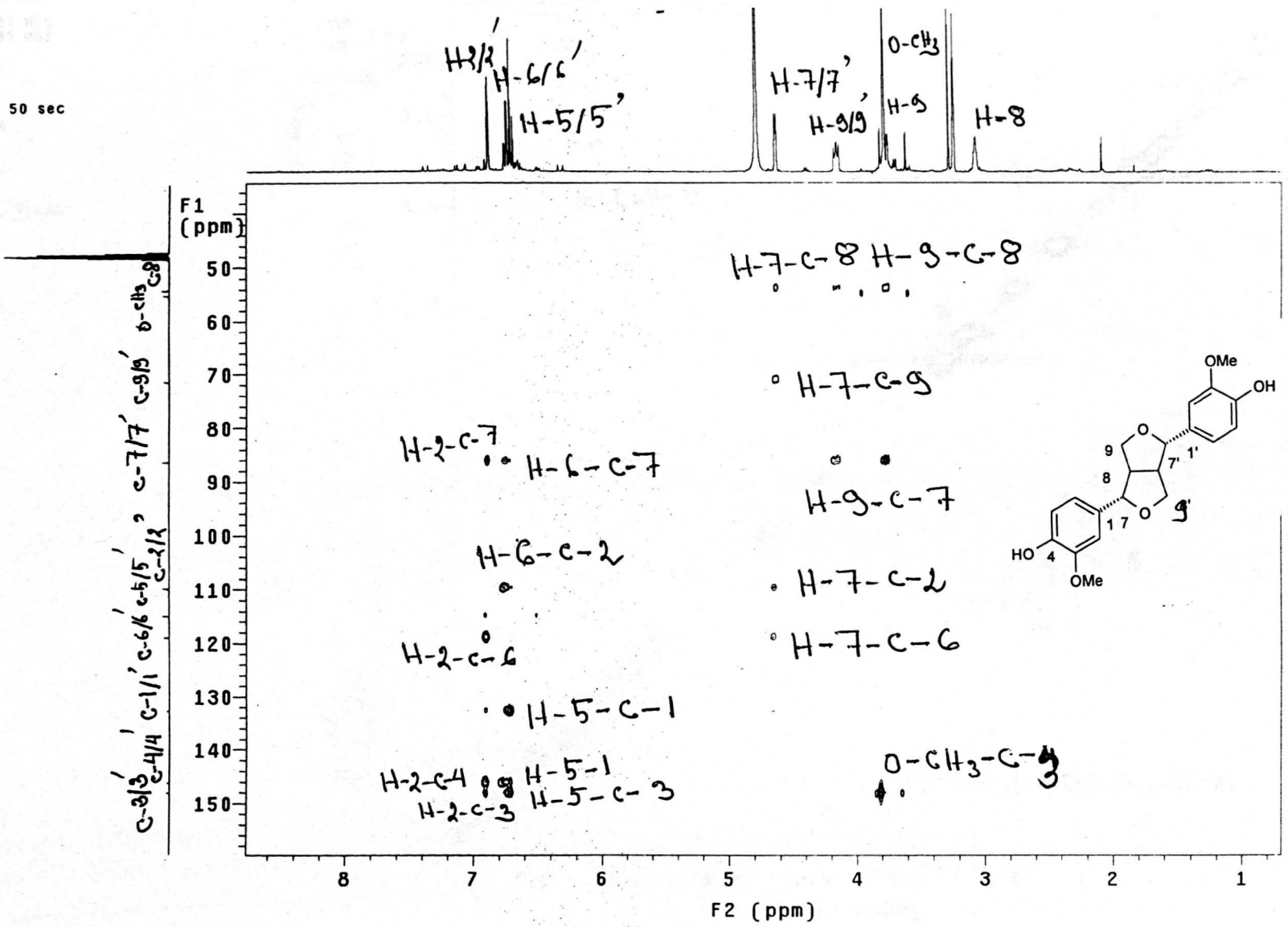


Figure 99. HMBC spectrum (CD_3OD, 400 MHz) of **SM10** (pinoresinol, **144**)

Pulse Sequence: Troesy_da
Solvent: cd3od
Temp. 26.0 C / 299.1 K
File: Shoeb_MS73_10_1H_TROESY_800ms_CD3OD
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Mixing 0.800 sec
Acq. time 0.177 sec
Width 5602.2 Hz
2D Width 5602.2 Hz
16 repetitions
2 x 256 increments
OBSERVE H1, 399.8981394 MHz
DECOUPLE H1, 399.9001392 MHz
Power 46 dB
off during acquisition
on during delay
single frequency
DATA PROCESSING
Sq. sine bell 0.177 sec
Shifted by -0.177 sec
F1 DATA PROCESSING
Sq. sine bell 0.046 sec
Shifted by -0.046 sec
FT size 4096 x 1024
Total time 4 hr, 35 min, 0 sec

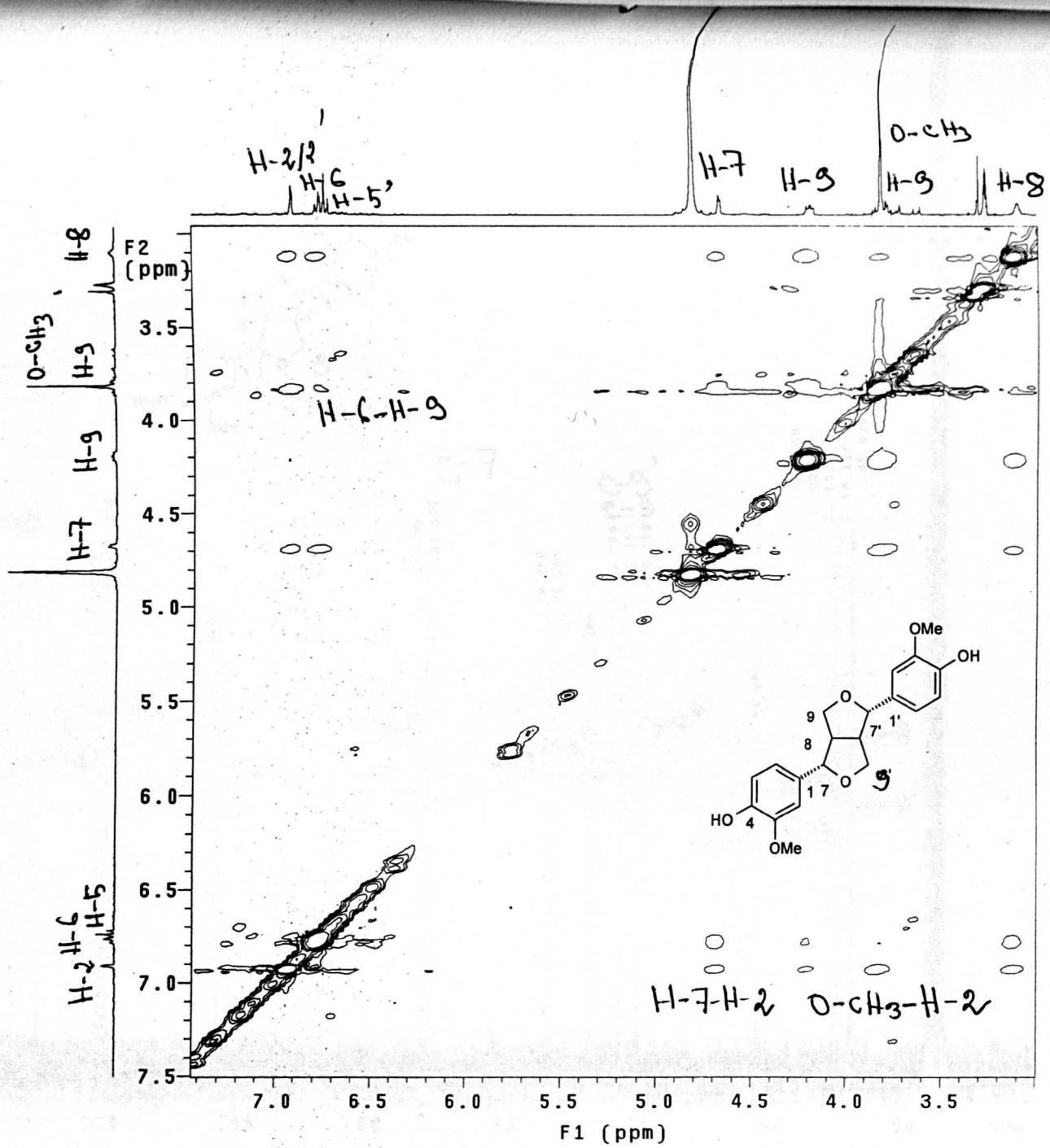


Figure 100. ROESY spectrum (CD_3OD, 400 MHz) of **SM10** (pinoresinol, **144**)

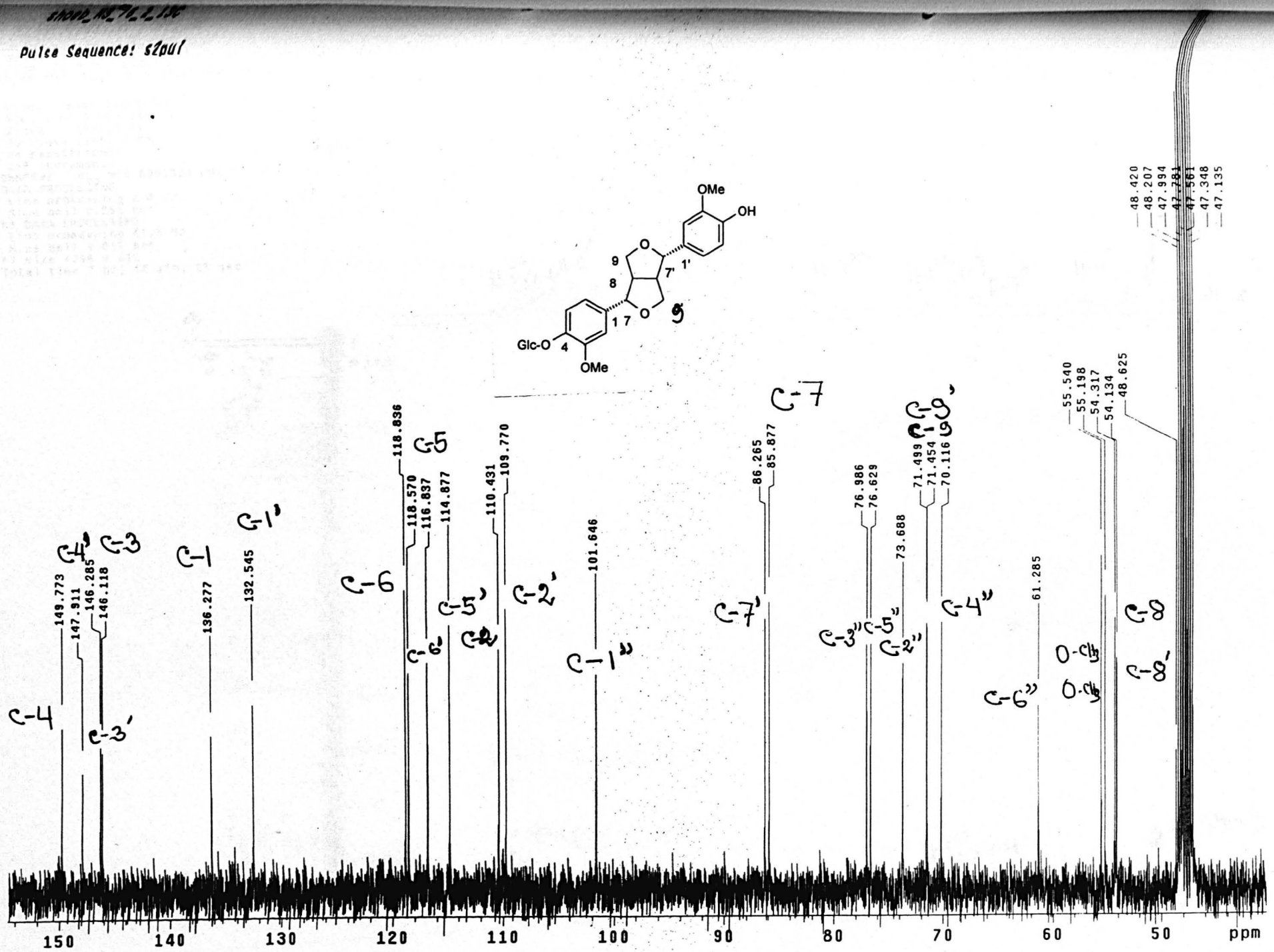


Figure 101. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM13** (pinoresinol 4-*O*-β-D-glucopyranoside, **176**)

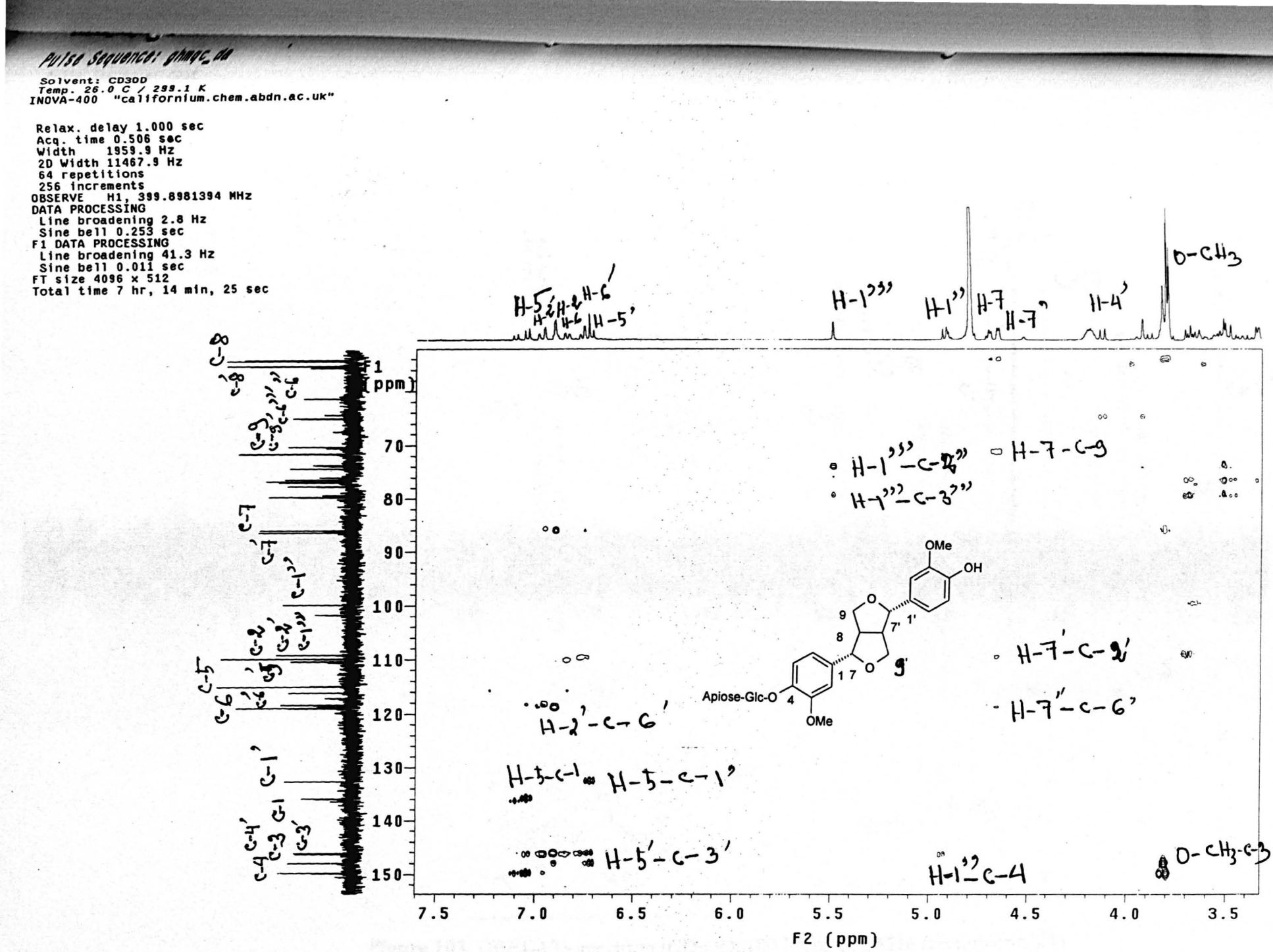


Figure 102. HMBC spectrum of **SM15** (Pinoresinol 4-*O*-β-D-apiose furanosyl-(1→2)-β-D-glucopyranoside, **178**)

Pulse Sequence: dept

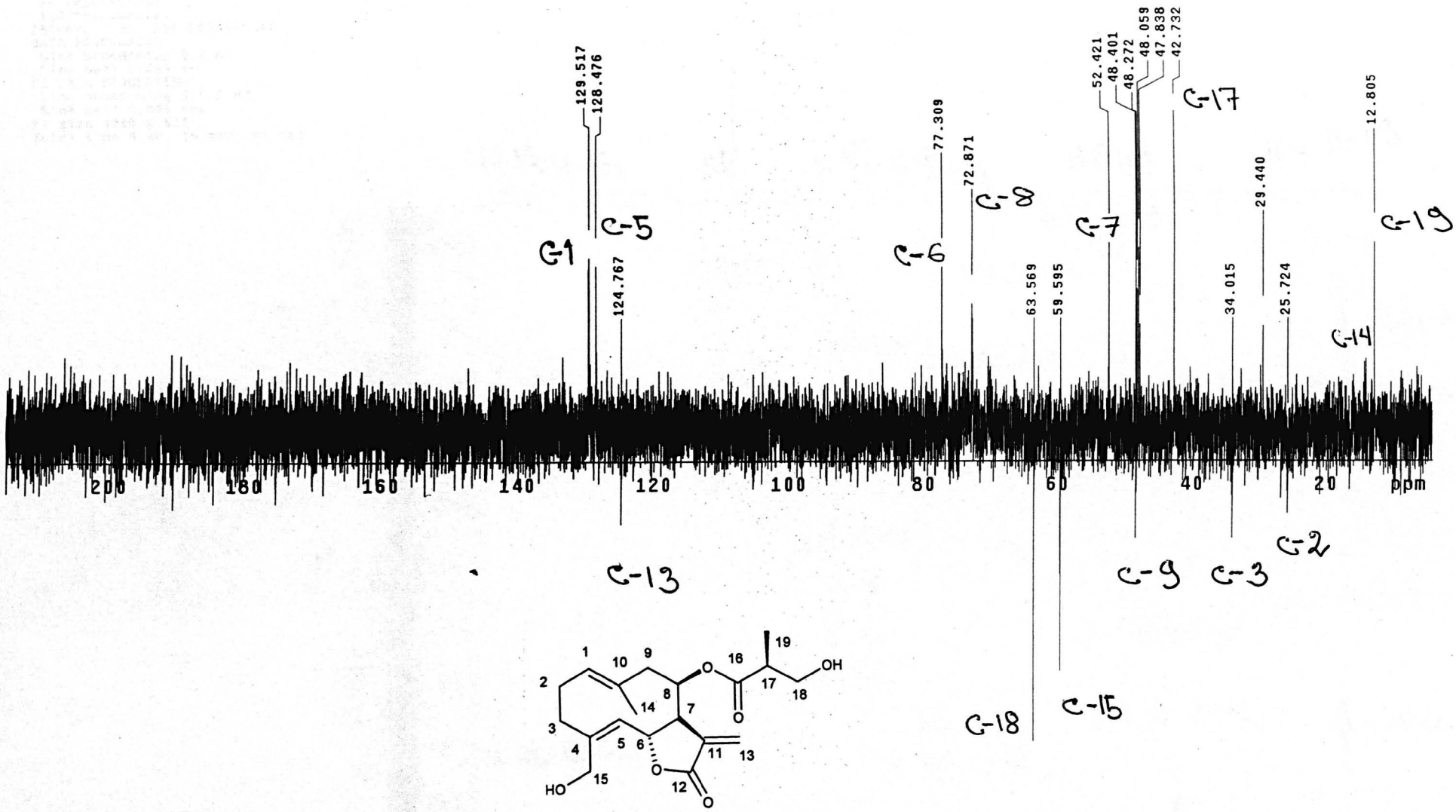


Figure 103. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM16** (arctiopicrin, **21**)

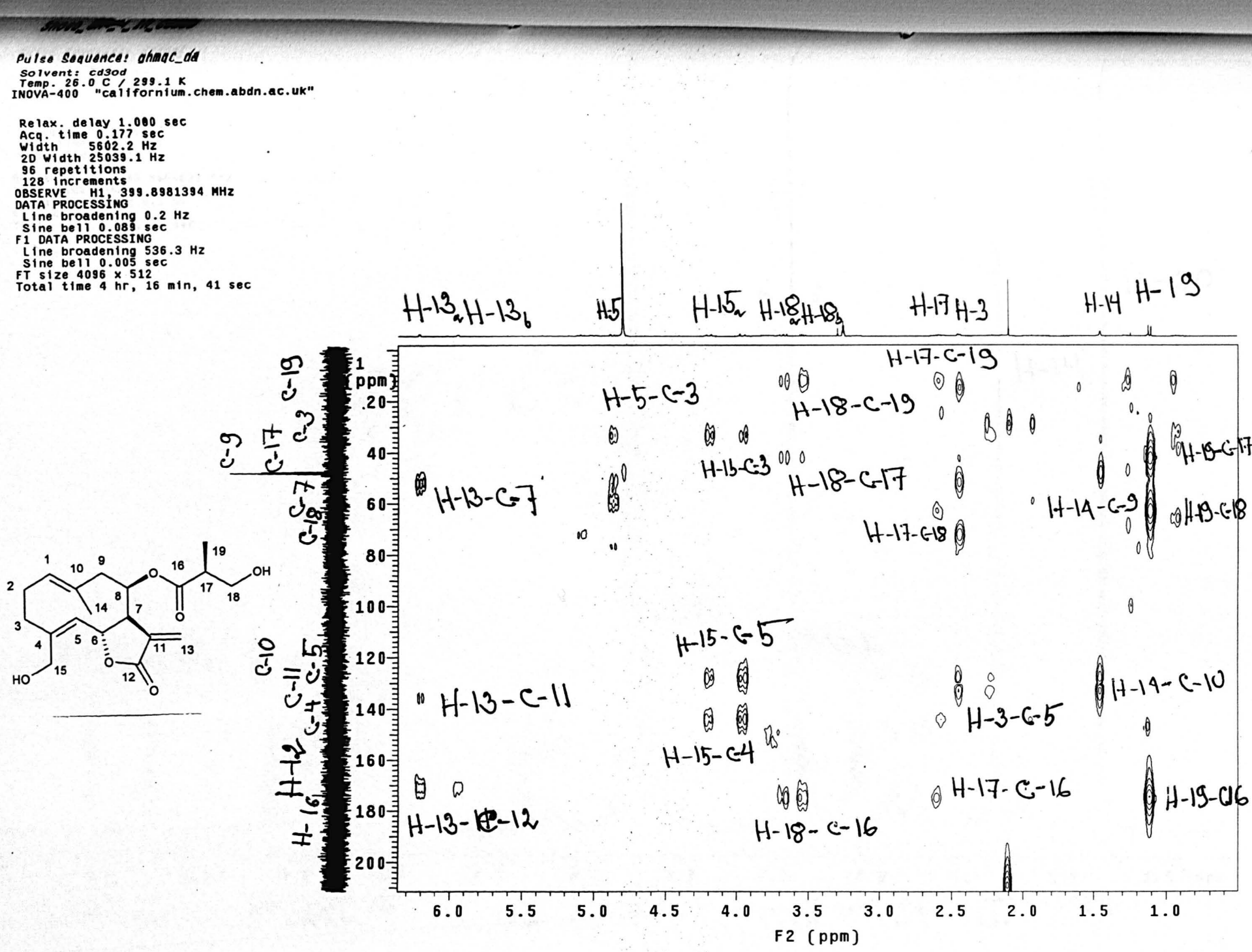


Figure 104. HMBC spectrum (CD_3OD, 100 MHz) of **SM16** (arctiopicrin, **21**)

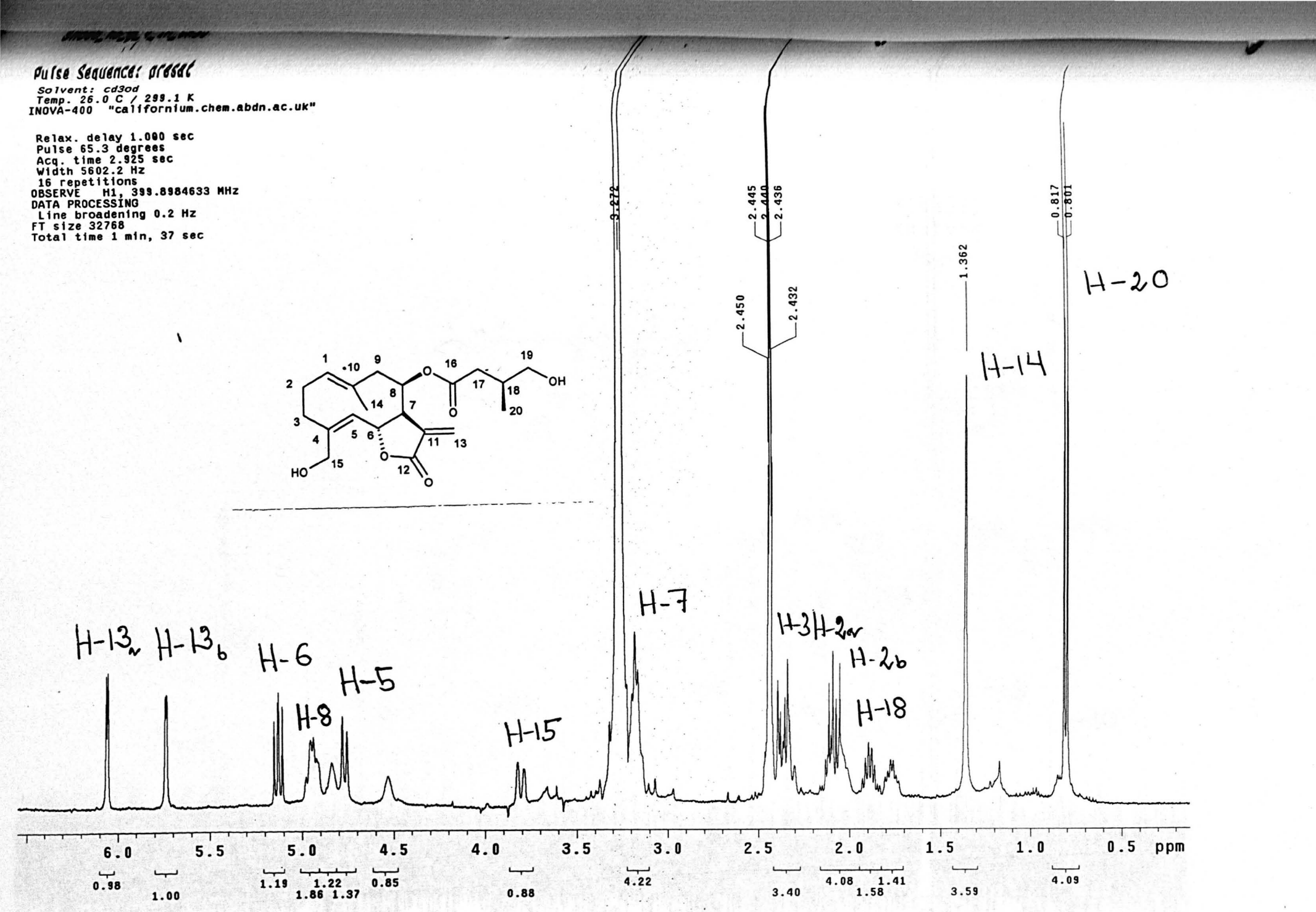


Figure 105. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

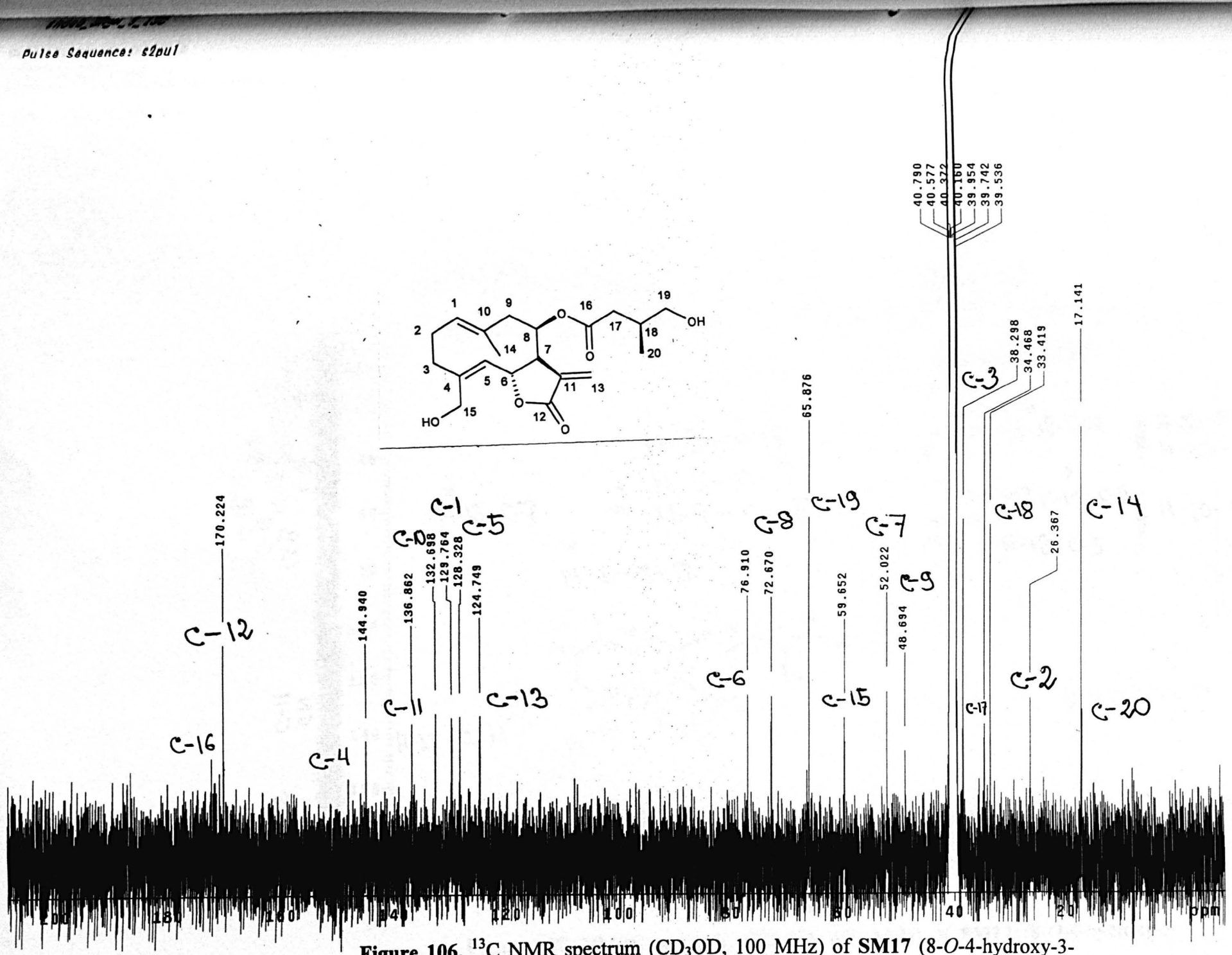


Figure 106. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

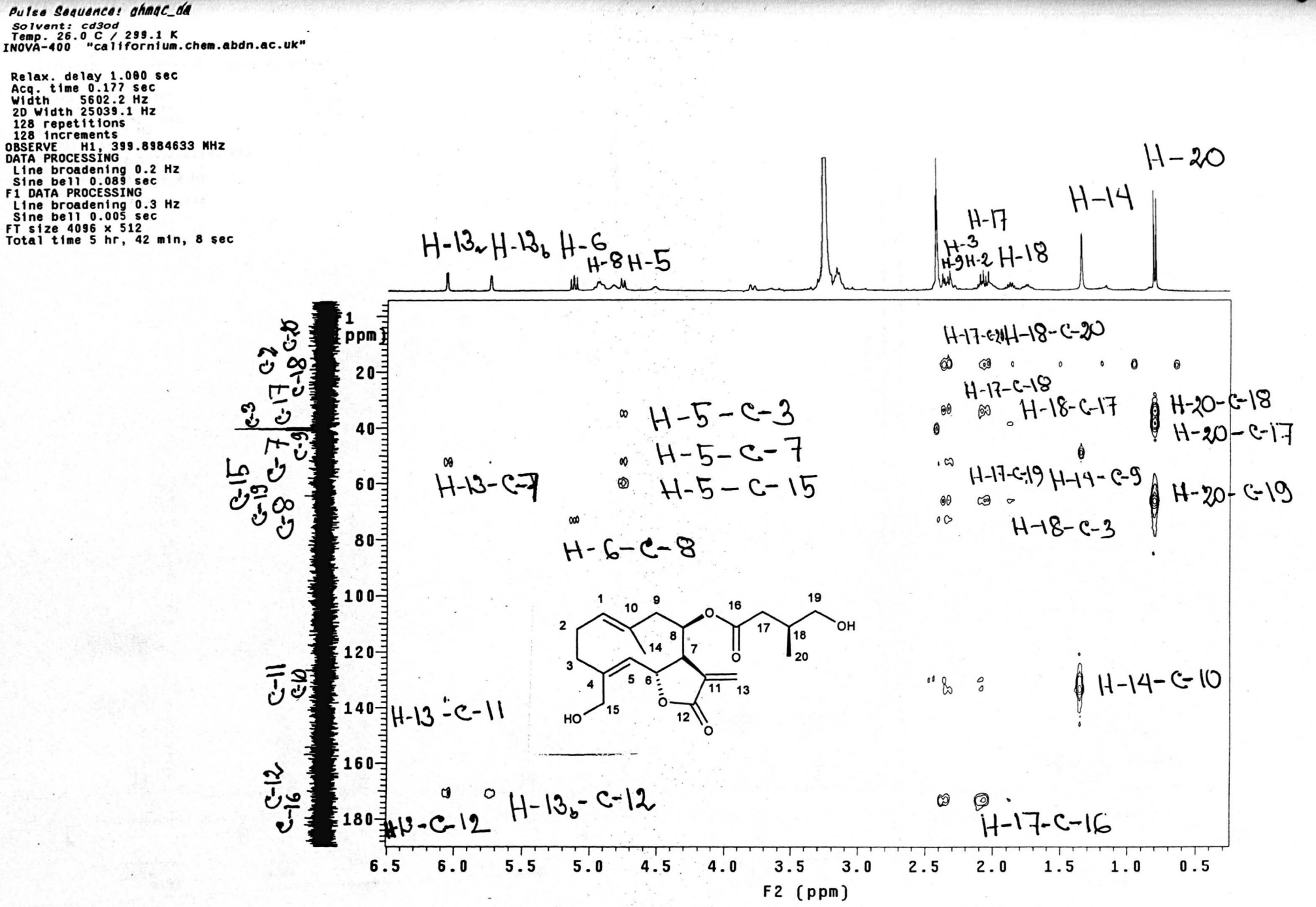


Figure 107. HMBC spectrum (CD_3OD, 400 MHz) of **SM17** (8-*O*-4-hydroxy-3-methylbutanoyl-salonitenolide, **179**)

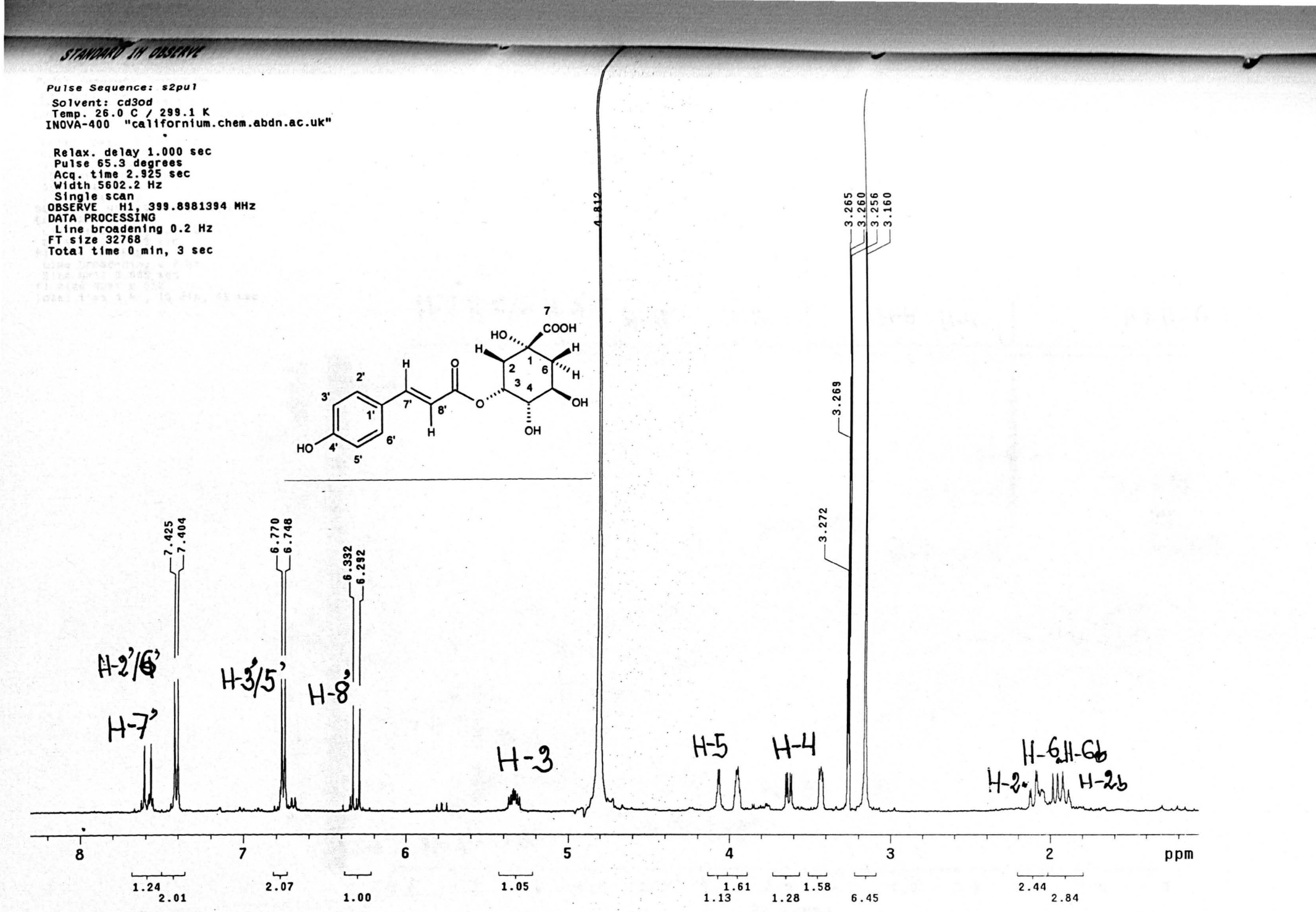


Figure 108. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM19** (*trans* 3-*O*-p-coumaroyl quinic acid, **181**)

Pulse Sequence: ghmqc_da
Solvent: cd3od
Temp. 26.0 C / 299.1 K
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.177 sec
Width 5602.2 Hz
2D Width 25039.1 Hz
96 repetitions
128 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Line broadening 0.2 Hz
Sine bell 0.089 sec
F1 DATA PROCESSING
Line broadening 0.3 Hz
Sine bell 0.005 sec
FT size 4096 x 512
Total time 4 hr, 16 min, 41 sec

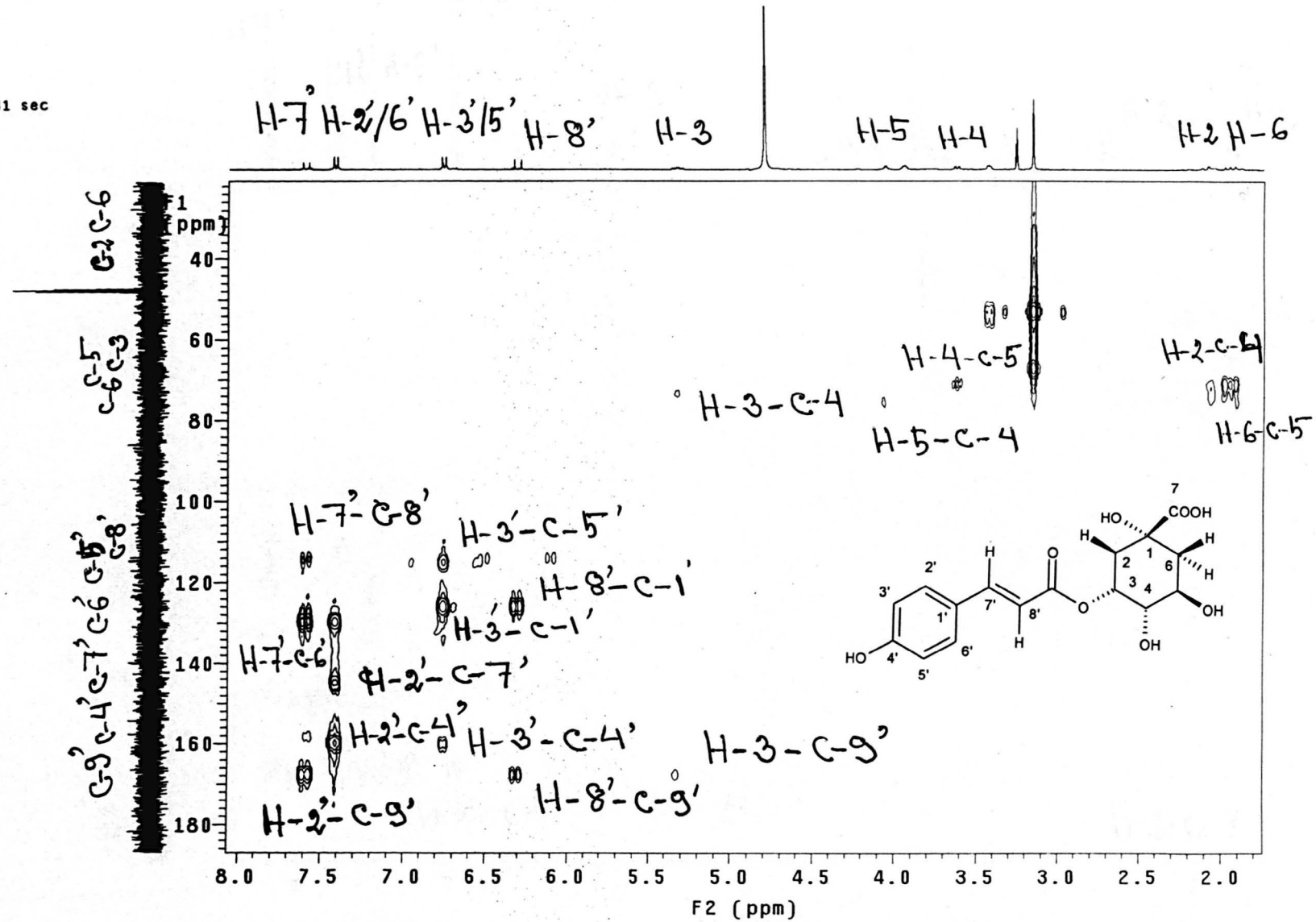


Figure 109. HMBC spectrum (CD_3OD, 400 MHz) of **SM19** (*trans* 3-*O*-p-coumaroyl quinic acid, **181**)

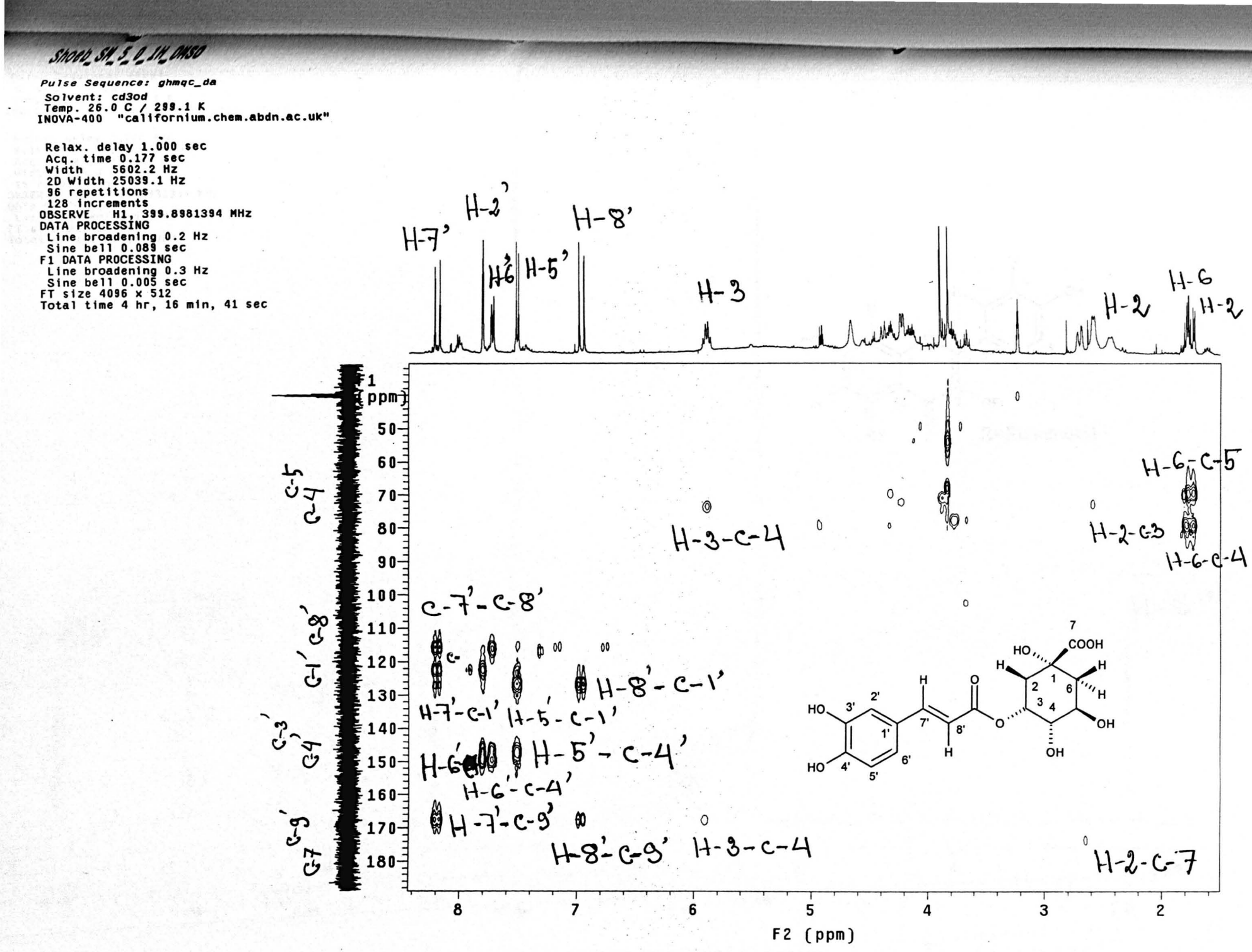


Figure 110. HMBC spectrum (CD_3OD, 400 MHz) of **SM21** (chlorogenic acid, **168**)

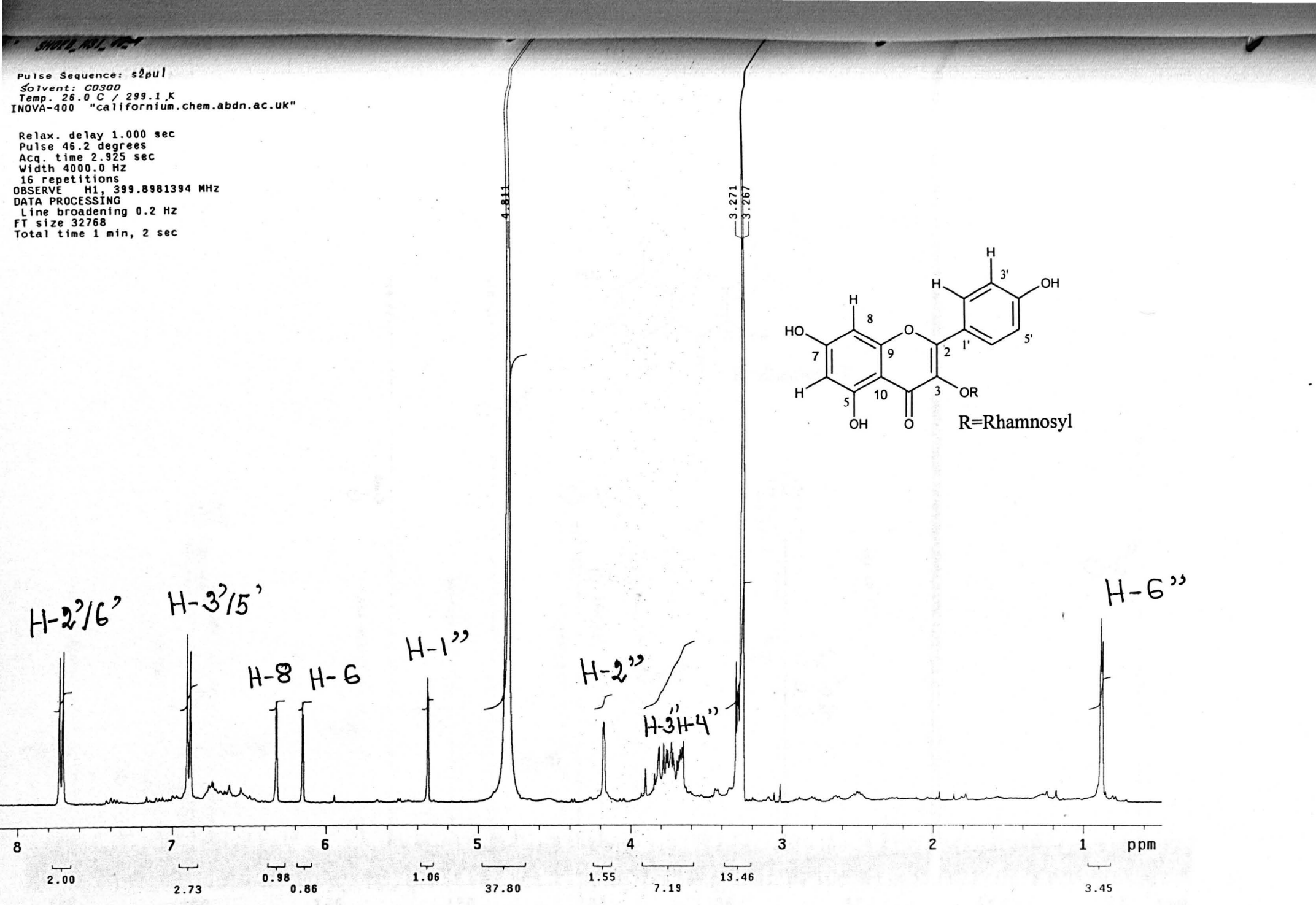


Figure 111. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

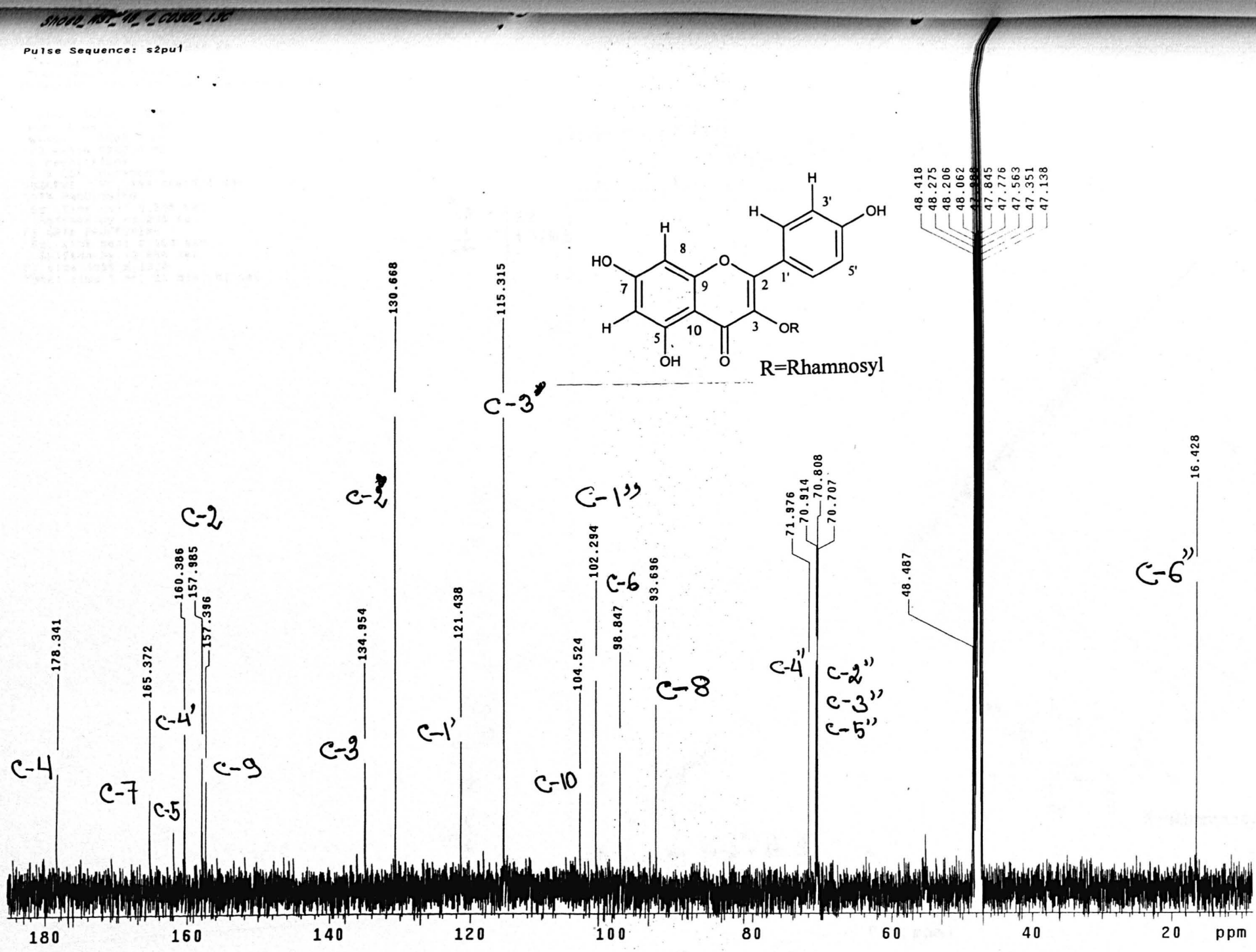


Figure 112. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM29** (afzelin, **187**)

Pulse Sequence: gmqfcops_da
Solvent: CD3OD
Temp. 26.0 C / 299.1 K
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.248 sec
Width 4000.0 Hz
2D Width 4000.0 Hz
8 repetitions
2 x 256 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Sq. sine bell 0.248 sec
Shifted by -0.248 sec
F1 DATA PROCESSING
Sq. sine bell 0.064 sec
Shifted by -0.064 sec
FT size 4096 x 1024
Total time 1 hr, 28 min, 46 sec

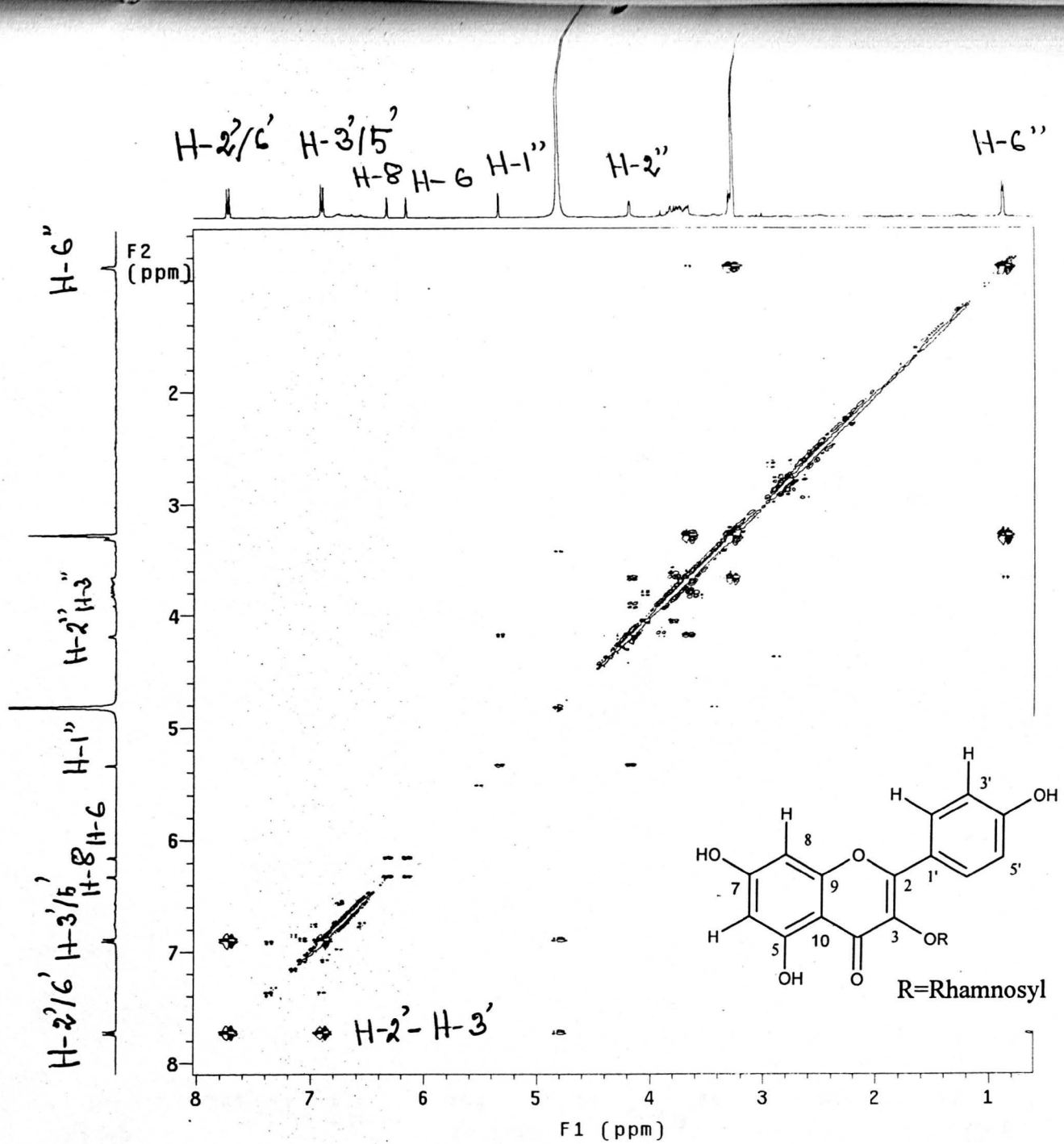


Figure 113. ^{1}H-^{1}H COSY spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

SHOEB_NSJ_10_1

```
Pulse Sequence: ghmqc_da
 Solvent: CD3OD
 Temp. 26.0 C / 299.1 K
INOVA-400  "californium.chem.abdn.ac.uk"

 Relax. delay 1.000 sec
 Acq. time 0.248 sec
 Width     4000.0 Hz
 2D Width 25039.1 Hz
 96 repetitions
 256 increments
OBSERVE   H1, 399.8981394 MHz
DATA PROCESSING
 Sine bell 0.124 sec
F1 DATA PROCESSING
 Line broadening 0.3 Hz
 Sine bell 0.005 sec
FT size 4096 x 512
Total time 9 hr, 3 min, 9 sec
```

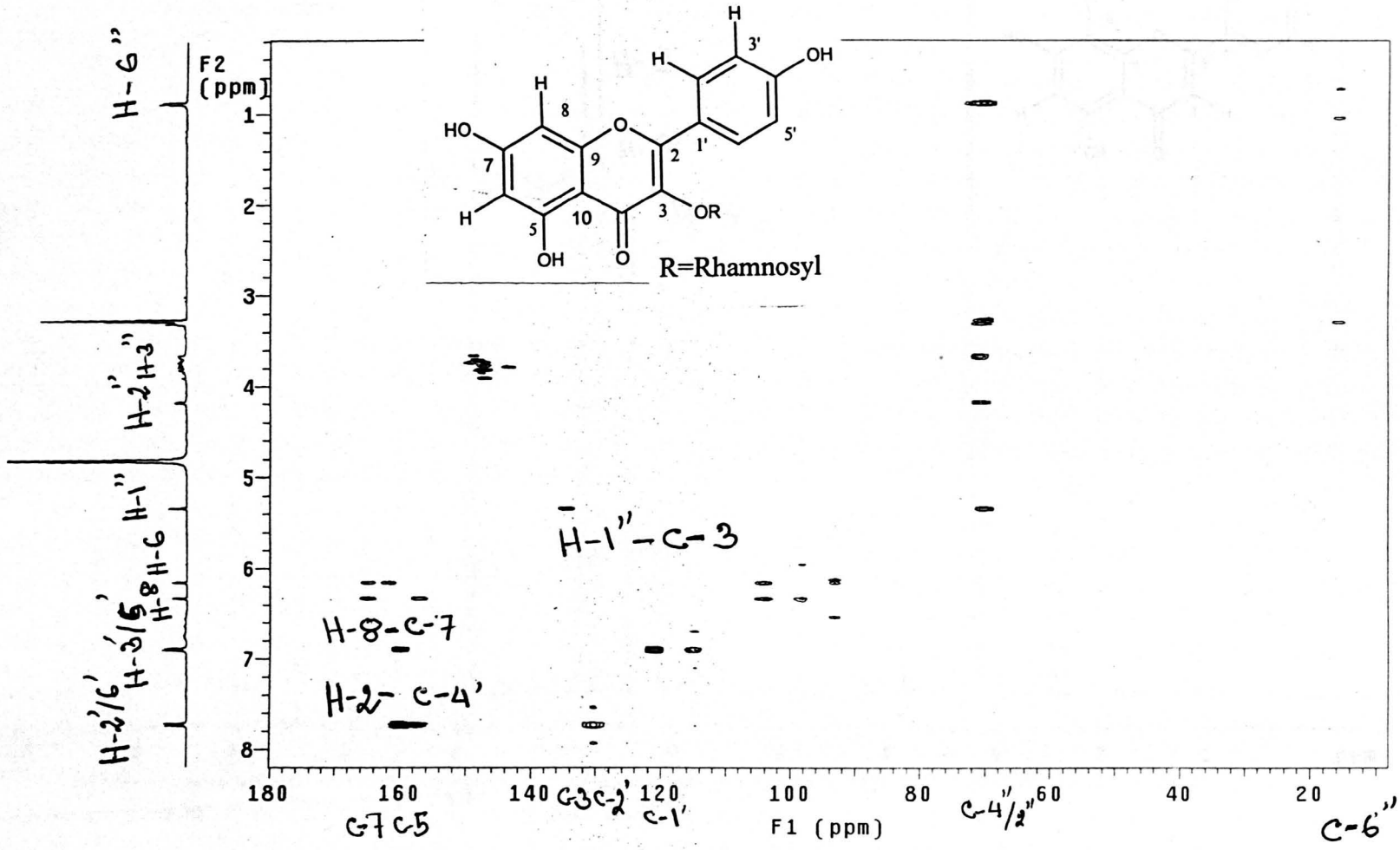


Figure 114. HMBC spectrum (CD_3OD, 400 MHz) of **SM29** (afzelin, **187**)

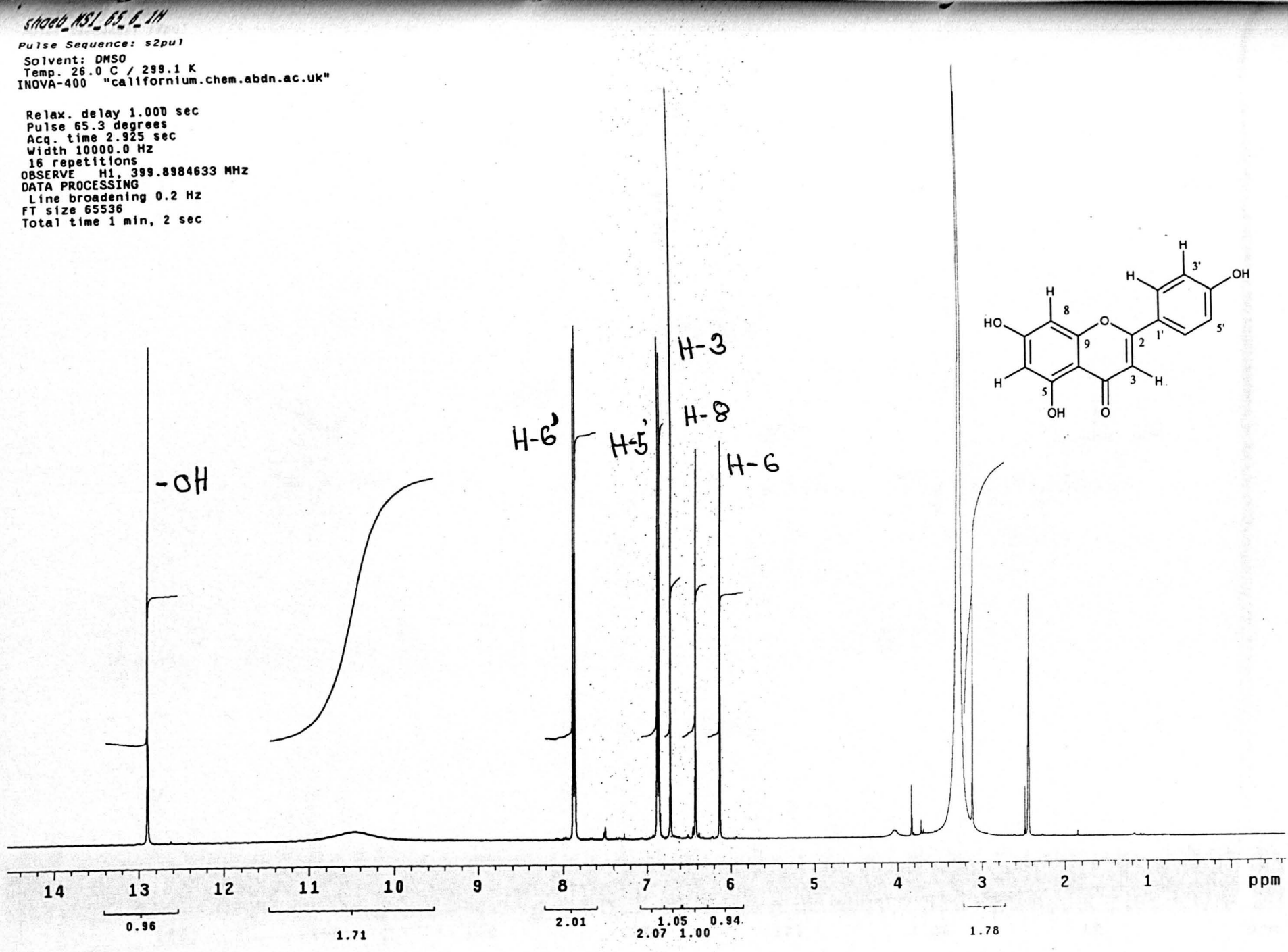


Figure 115. ^{1}H NMR spectrum (DMSO-d_6, 400 MHz) of **SM30** (apigenin, **108**)

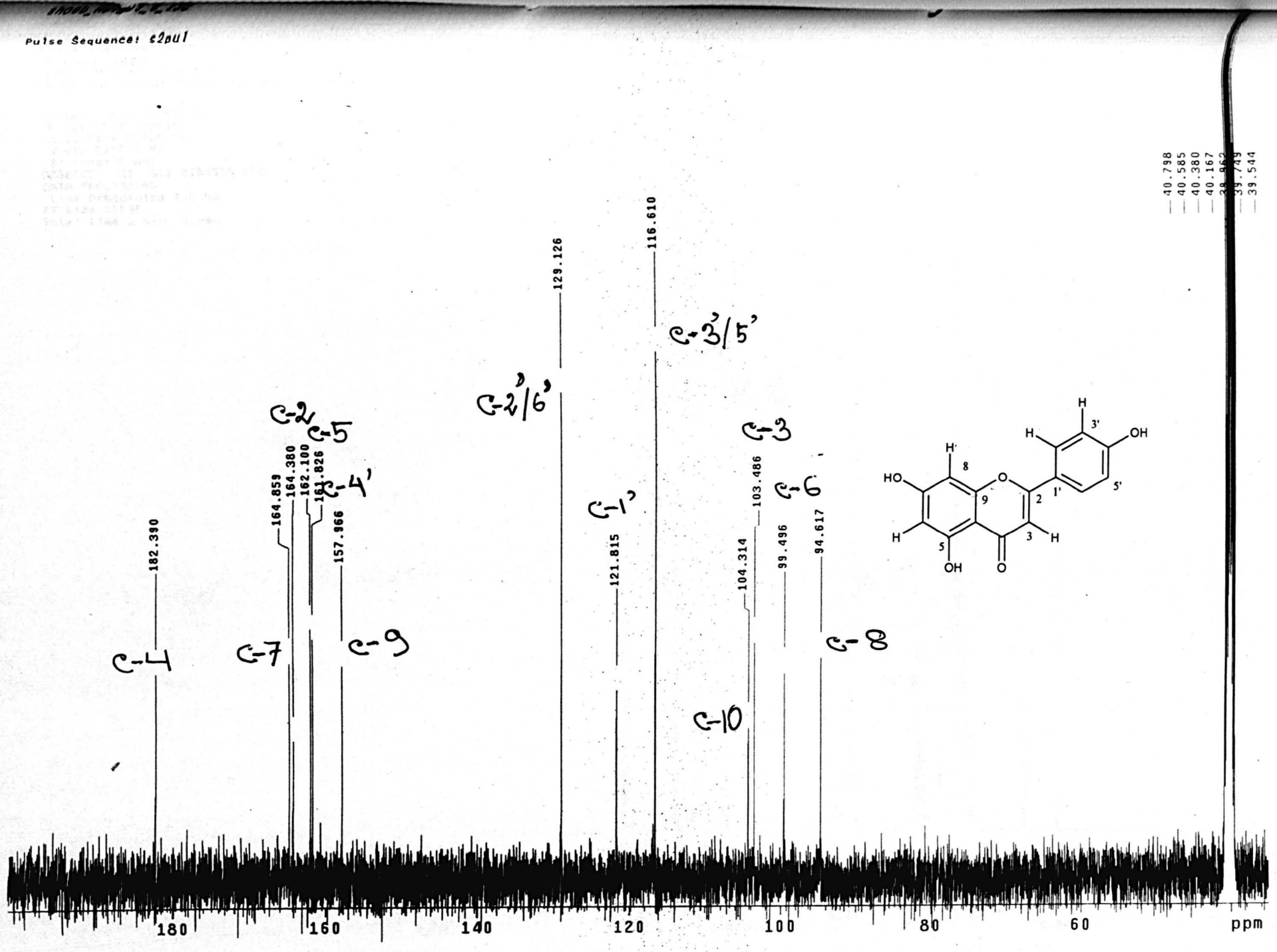


Figure 116. ^{13}C NMR spectrum (DMSO-d_6, 100 MHz) of **SM30** (apigenin, **108**)

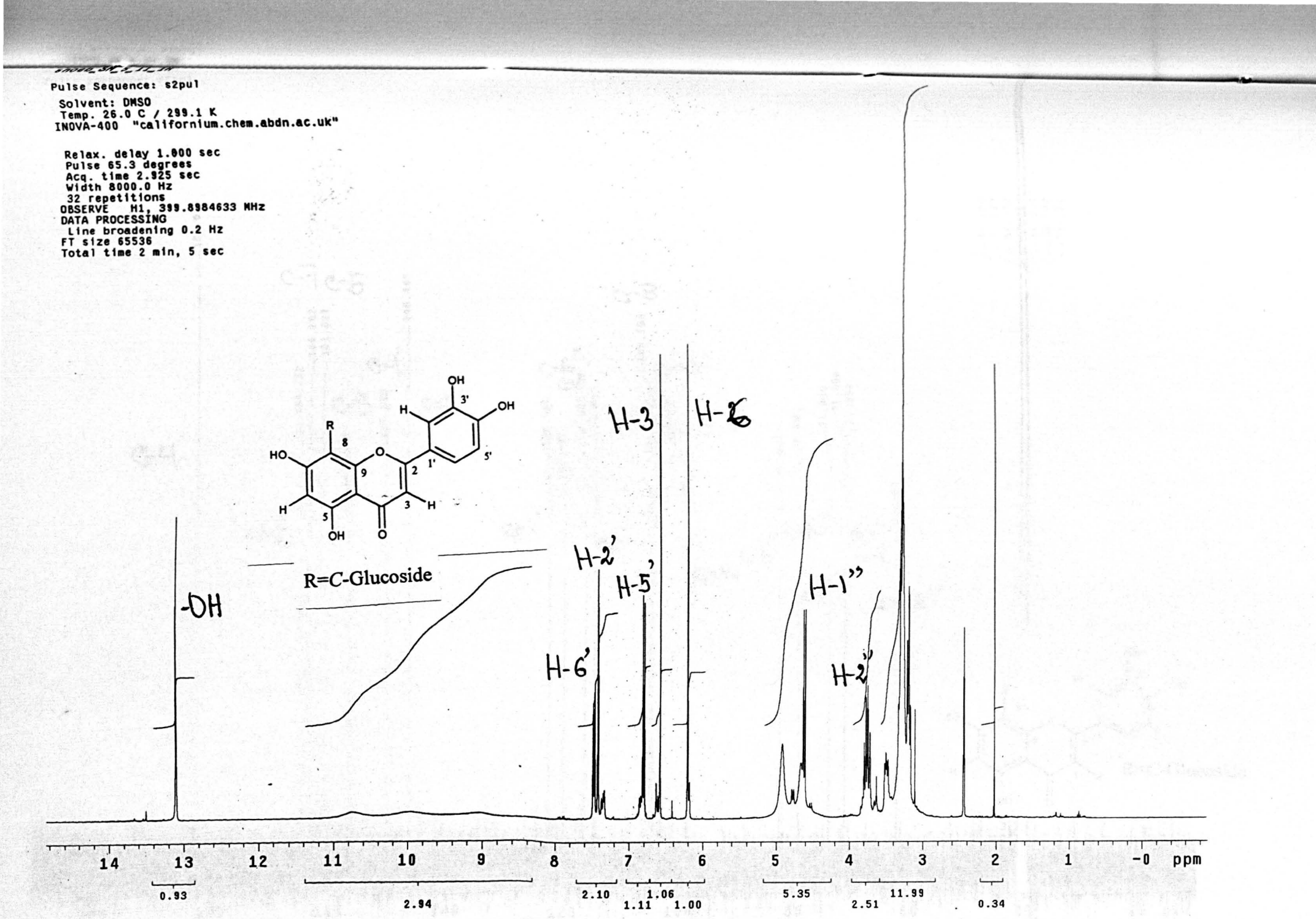


Figure 117. ^{1}H NMR spectrum (DMSO-d_6, 400 MHz) of **SM33** (orientin, **188**)

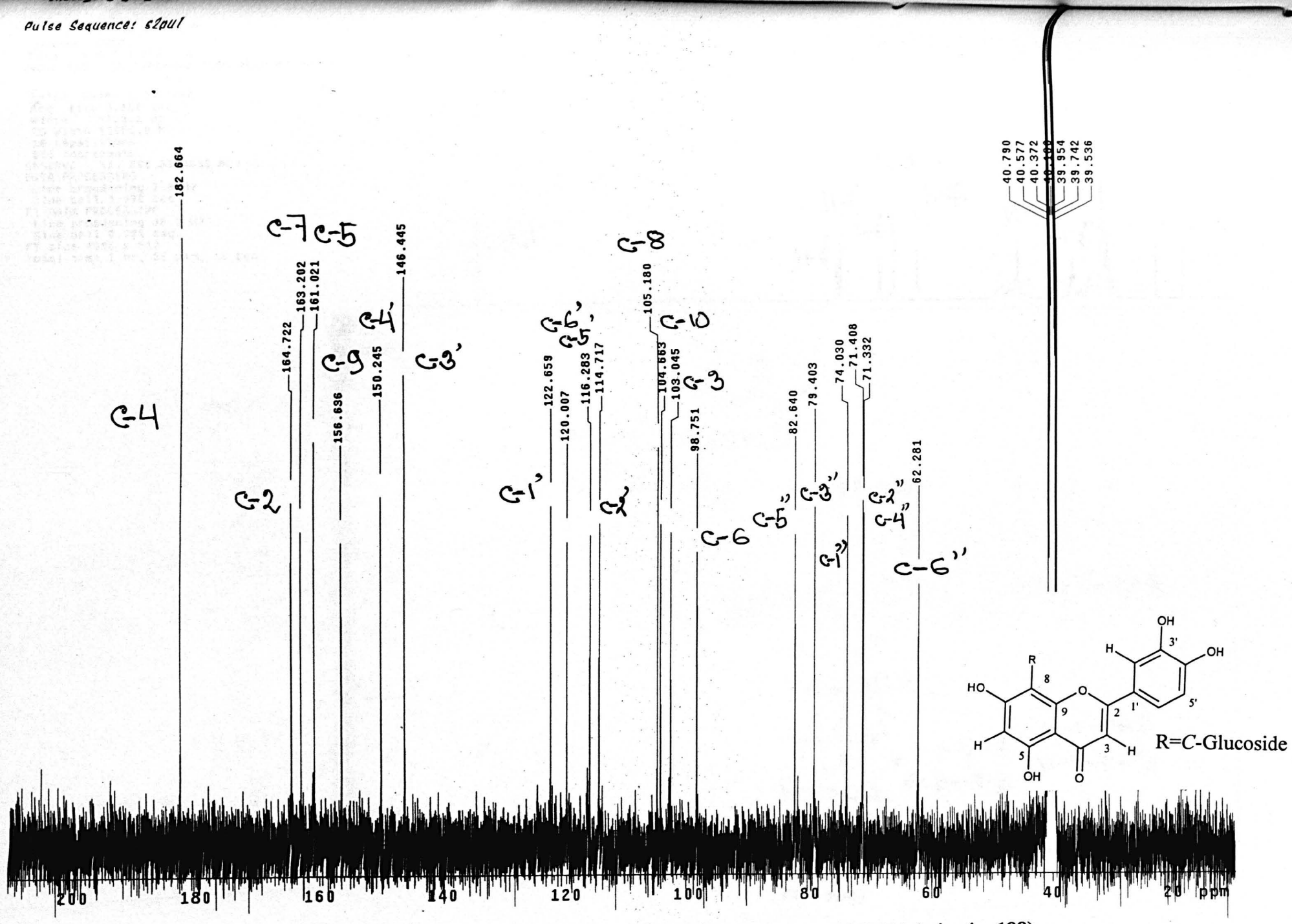


Figure 118. ^{13}C NMR spectrum (DMSO-d_6, 100 MHz) of **SM33** (orientin, **188**)

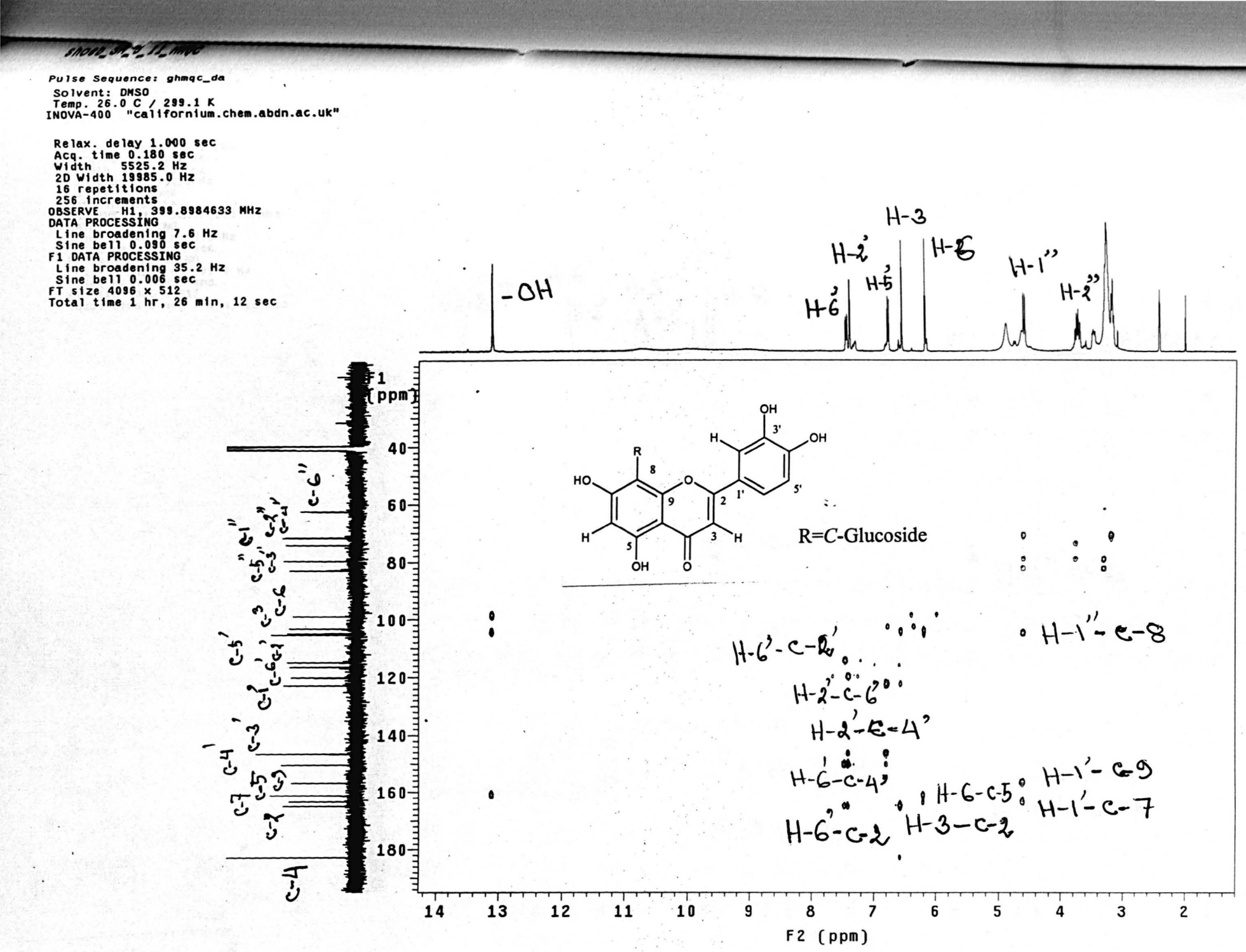


Figure 119. HMBC spectrum (DMSO-d_6, 400 MHz) of **SM33** (orientin, **188**)

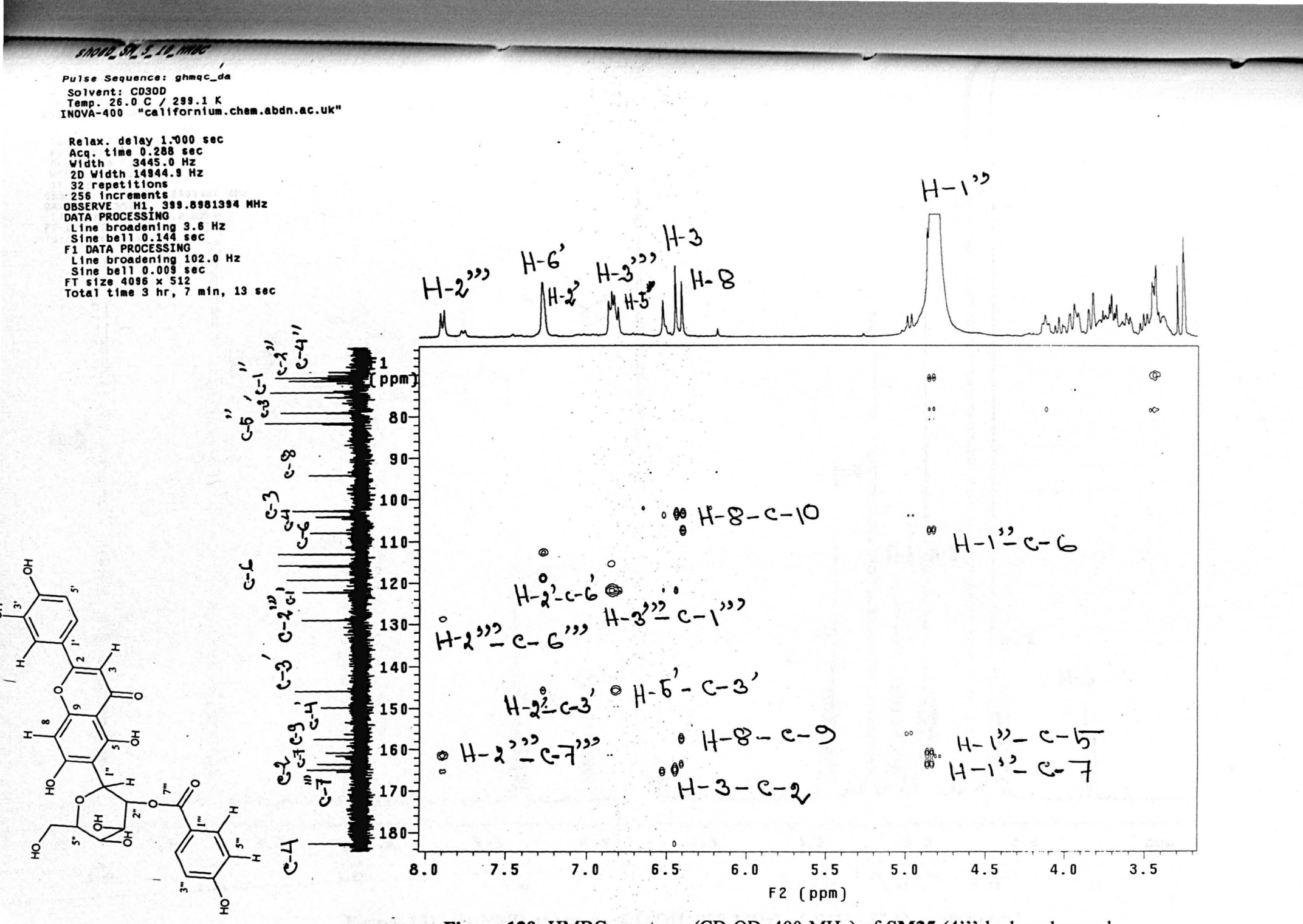


Figure 120. HMBC spectrum (CD_3OD, 400 MHz) of **SM35** (4'''-hydroxybenzoyl-isoorientin, **189**)

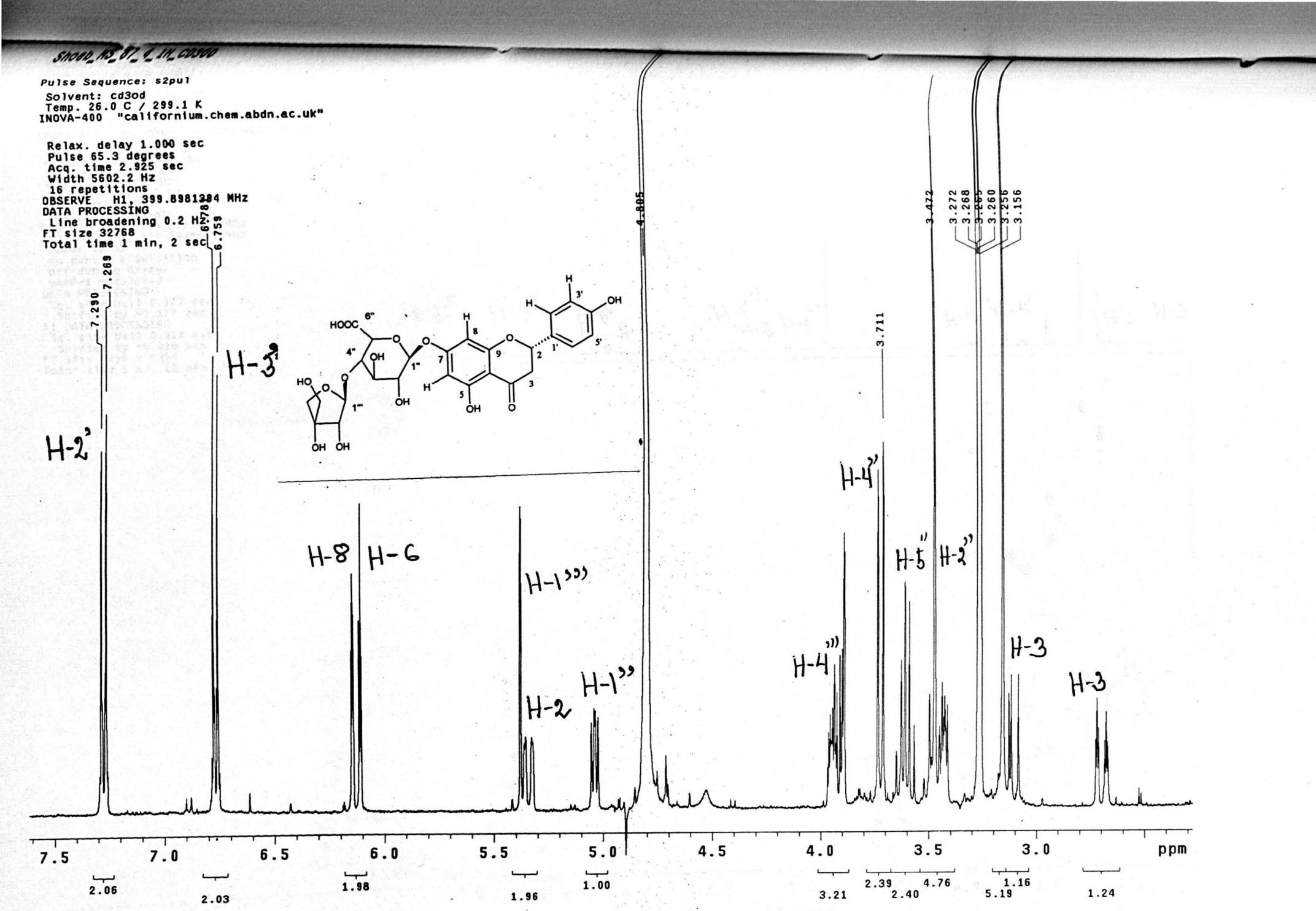


Figure 121. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM36** (7-*O*-apiofuranosyl (1→4)-glucuronic acid, **190**)

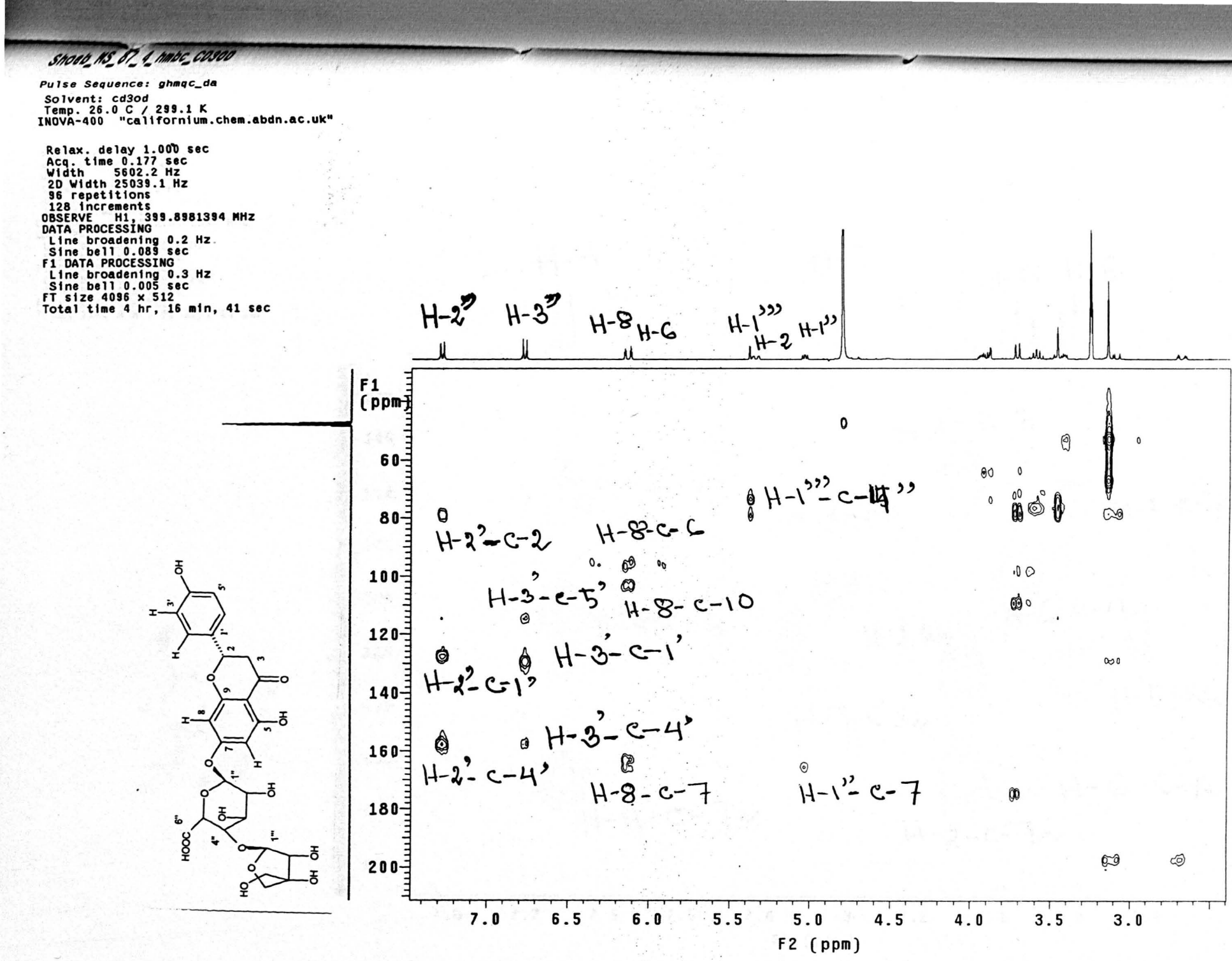


Figure 123. HMBC spectrum (CD_3OD, 400 MHz) of **SM36** (7-*O*-apiofuranosyl (1→4)-glucuronic acid, **190**)

Pulse Sequence: ghmqc_da
Solvent: cd3od
Temp. 26.0 C / 299.1 K
File: Shoeb_MS_87_3_gHMBC_CD3OD
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.177 sec
Width 5602.2 Hz
2D Width 25039.1 Hz
112 repetitions
128 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Line broadening 0.2 Hz
Sine bell 0.089 sec
F1 DATA PROCESSING
Line broadening 0.3 Hz
Sine bell 0.005 sec
FT size 4096 x 512
Total time 4 hr, 59 min, 24 sec

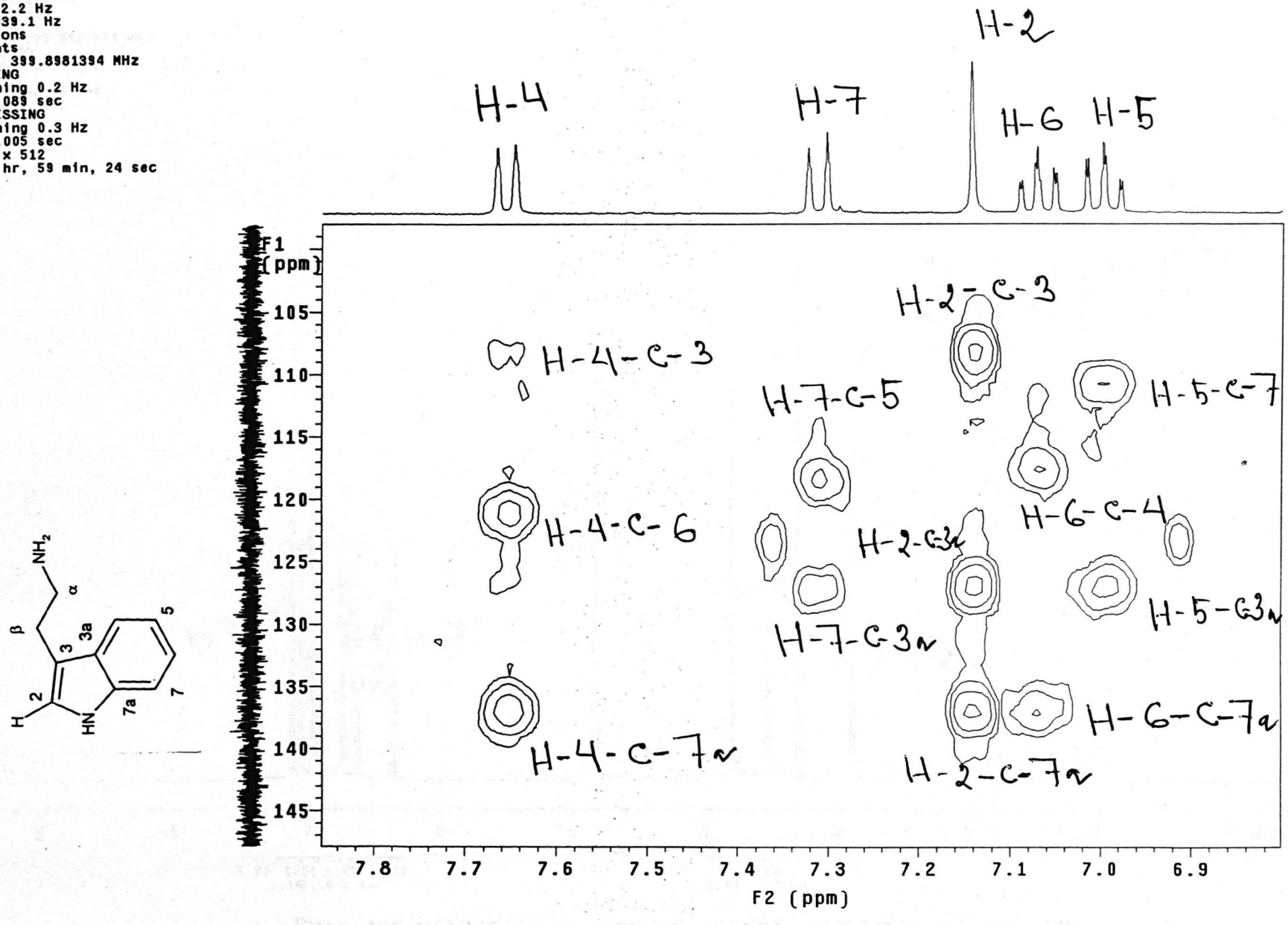


Figure 124. HMBC spectrum (CD_3OD, 400 MHz) of **SM37** (tryptamine, 191)

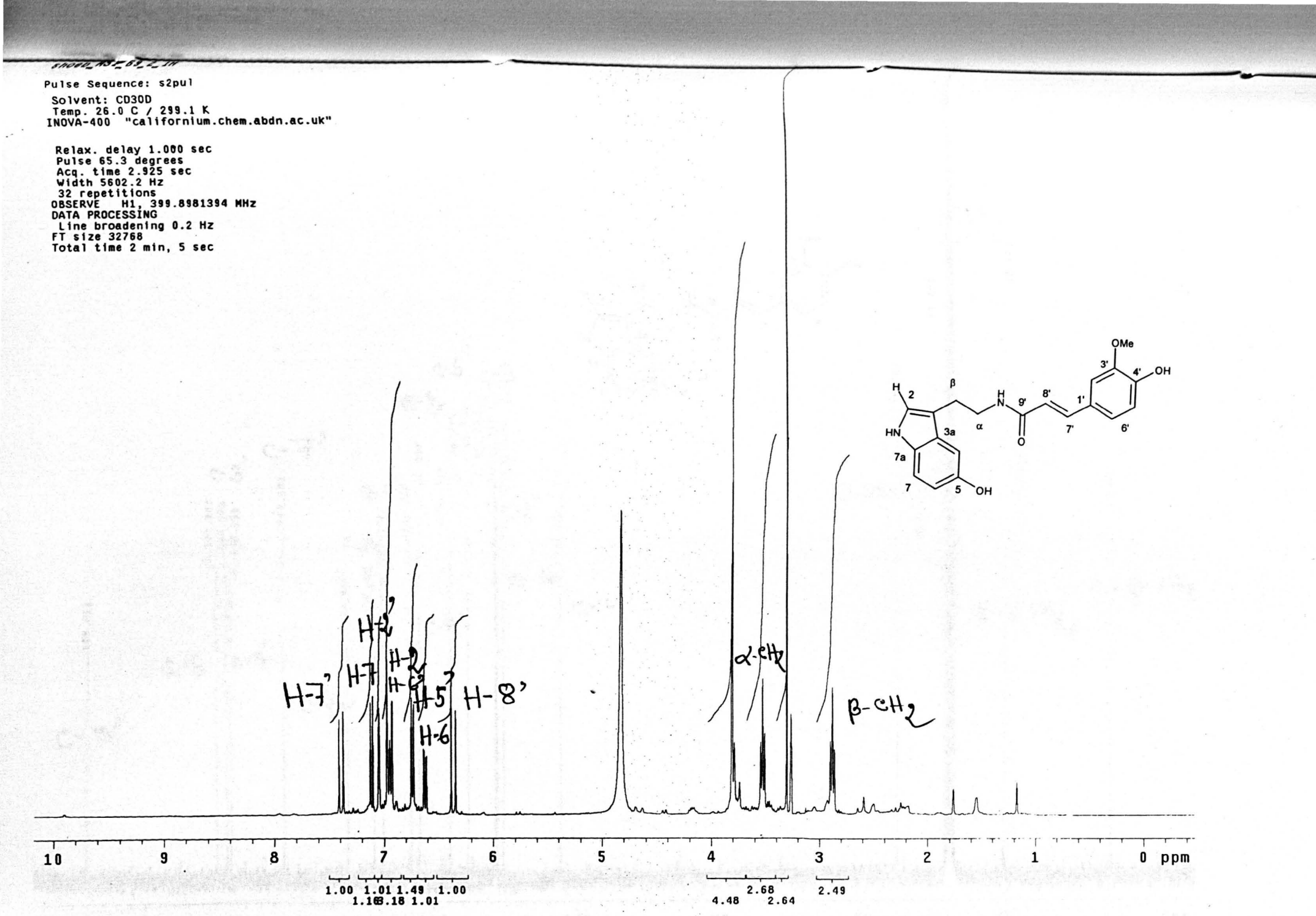


Figure 125. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM42** (moschamine, **148**)

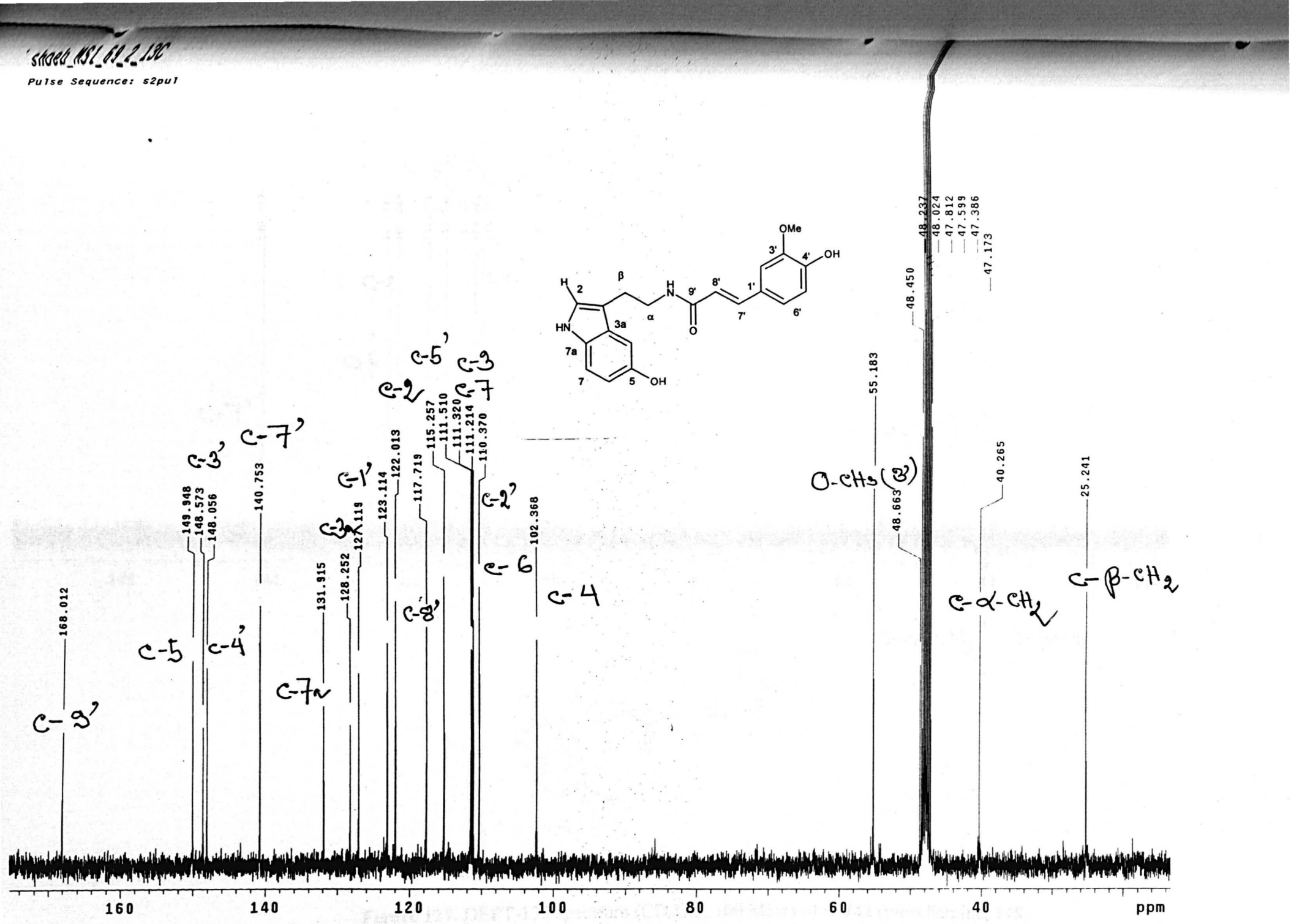


Figure 126. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of **SM42** (moschamine, **148**)

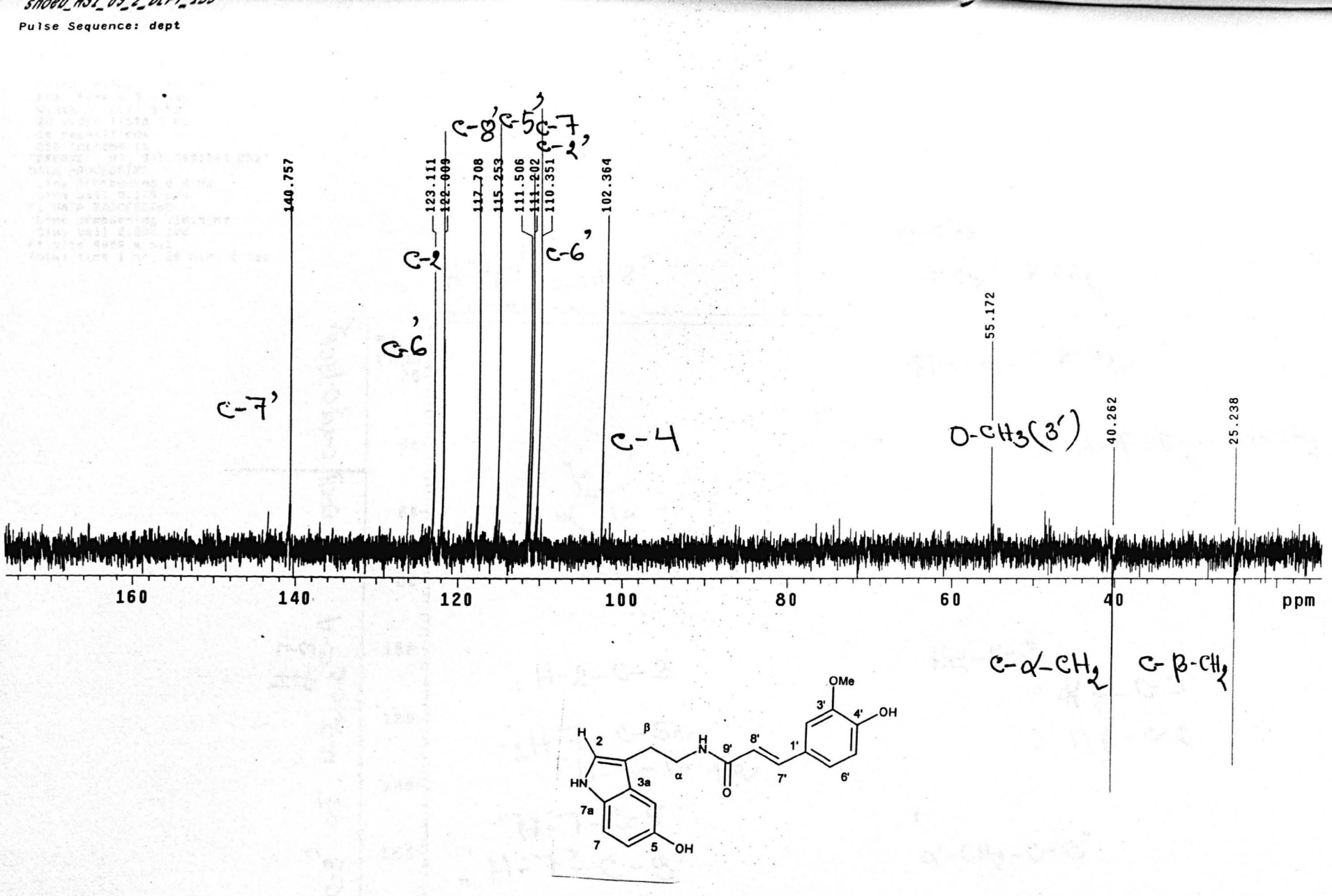


Figure 127. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM42** (moschamine, **148**)

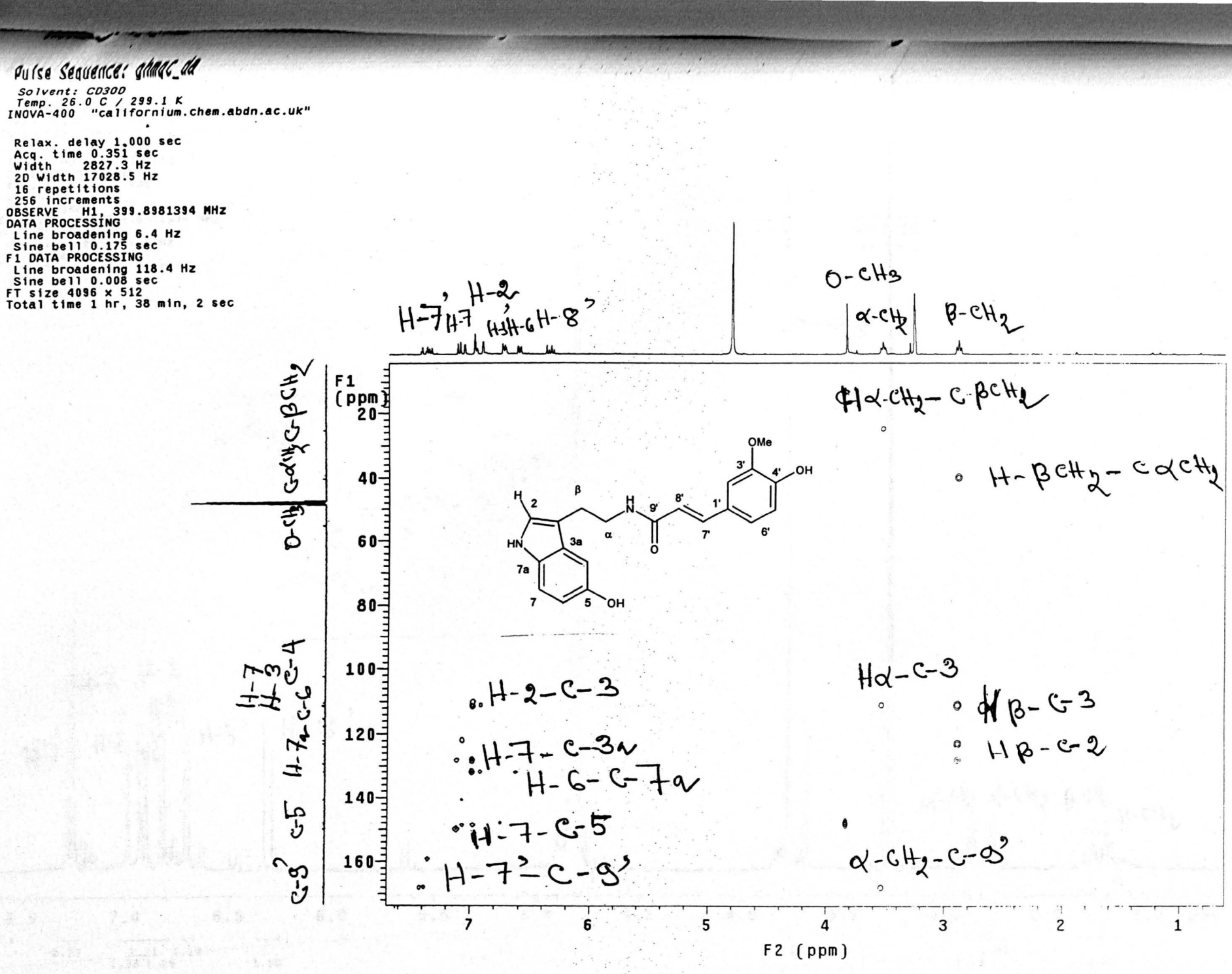


Figure 128. HMBC spectrum (CD_3OD, 400 MHz) of **SM42** (moschamine, **148**)

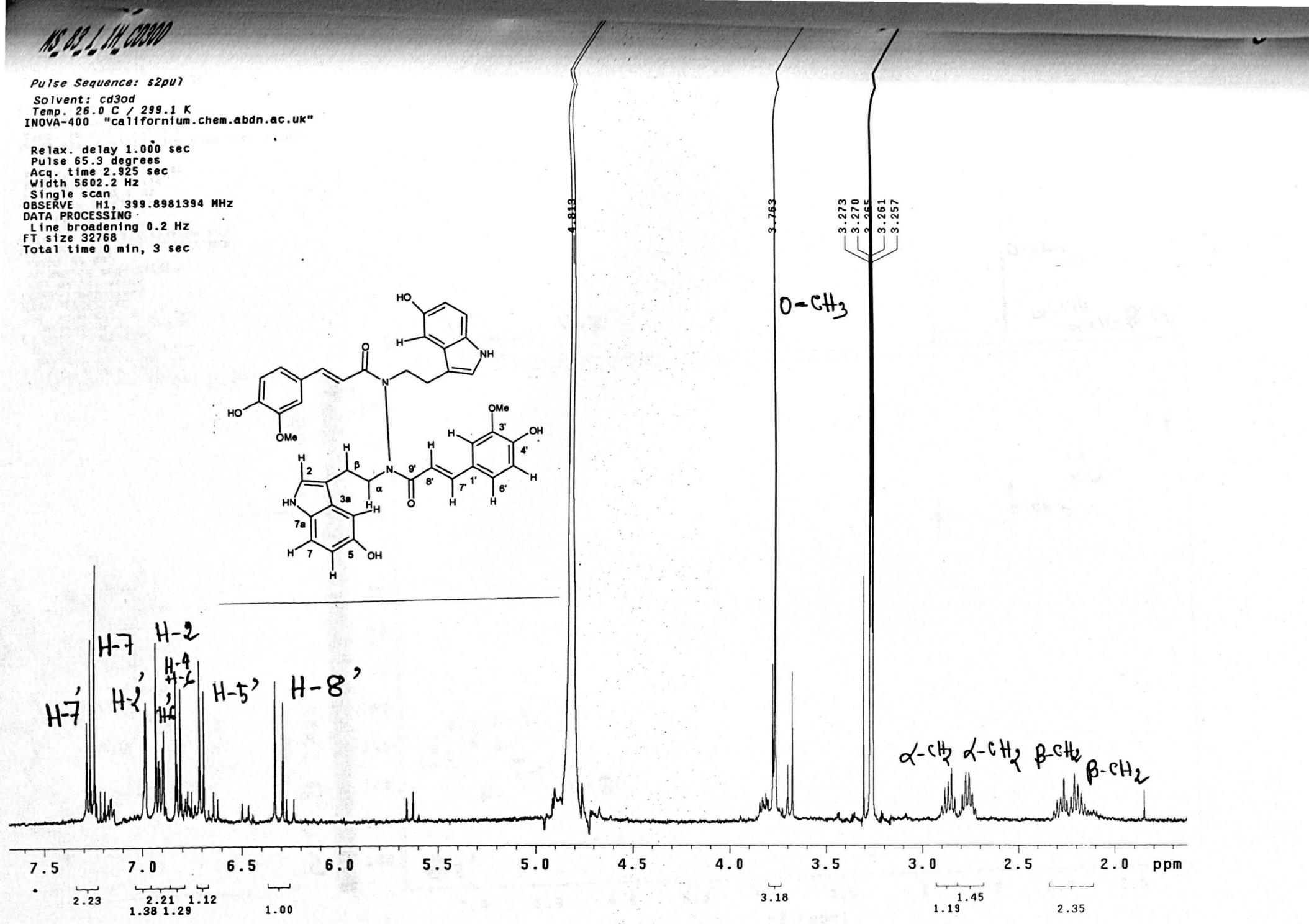


Figure 129. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

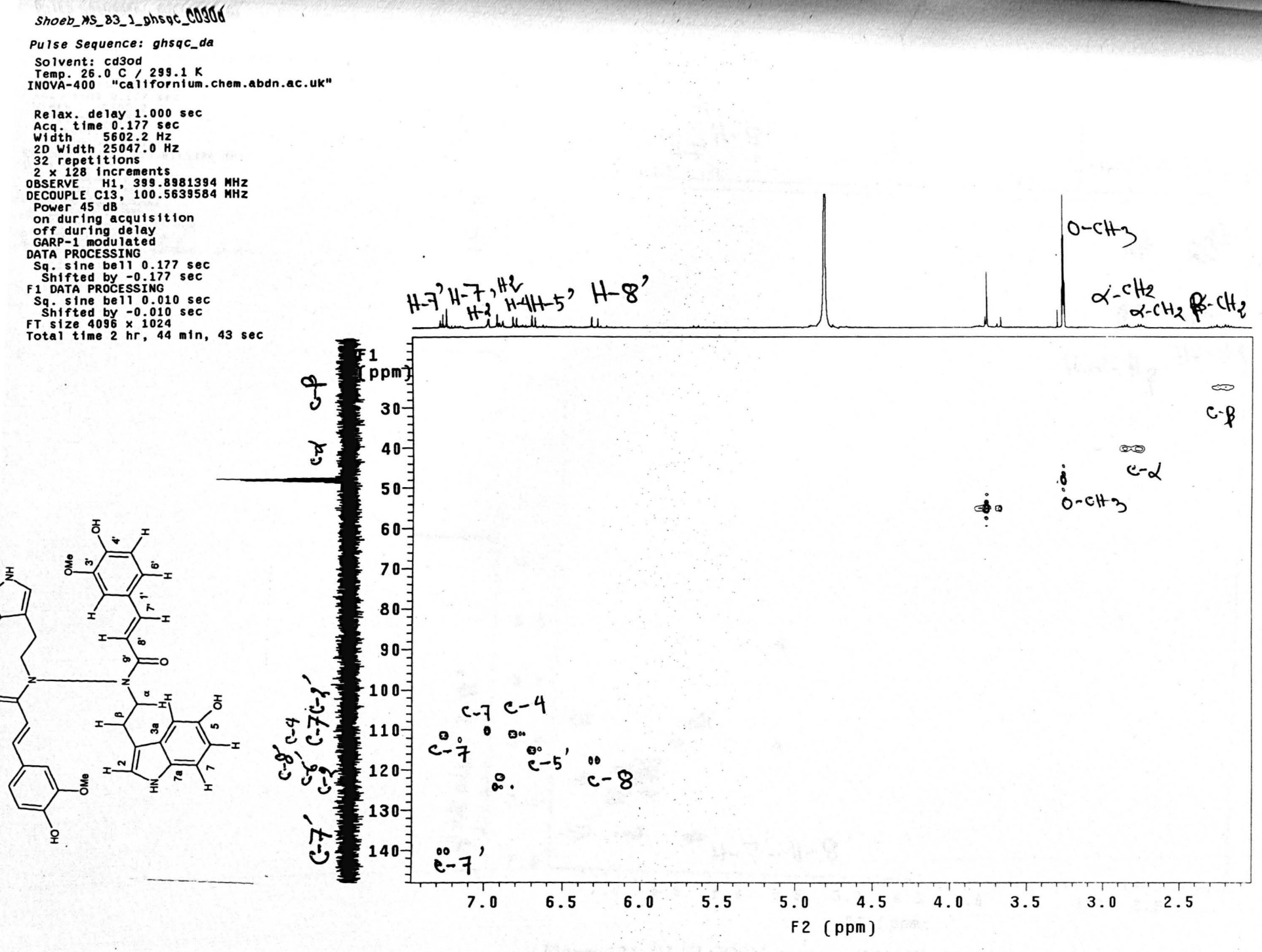


Figure 130. HSQC spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

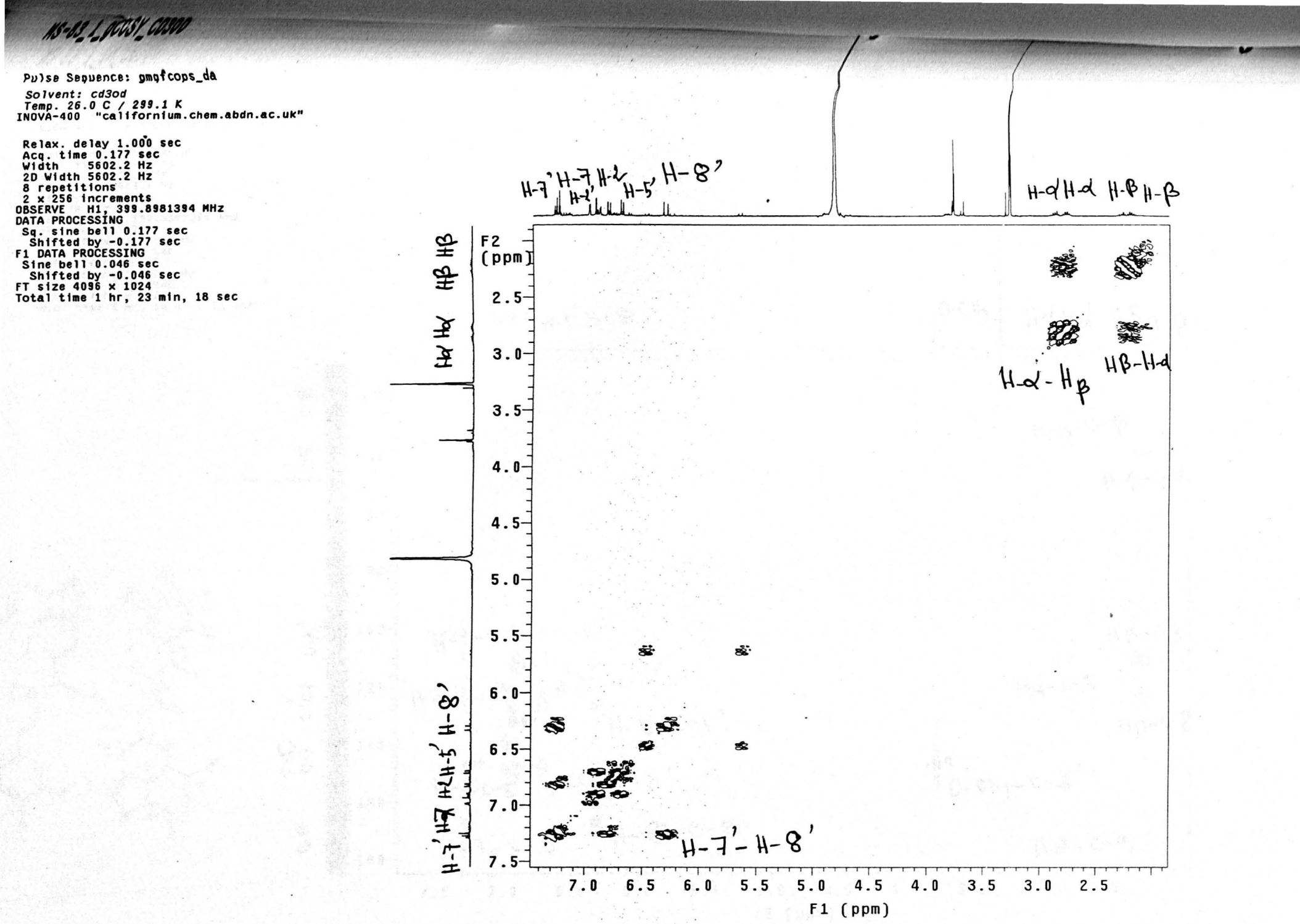


Figure 131. ^{1}H-^{1}H COSY spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

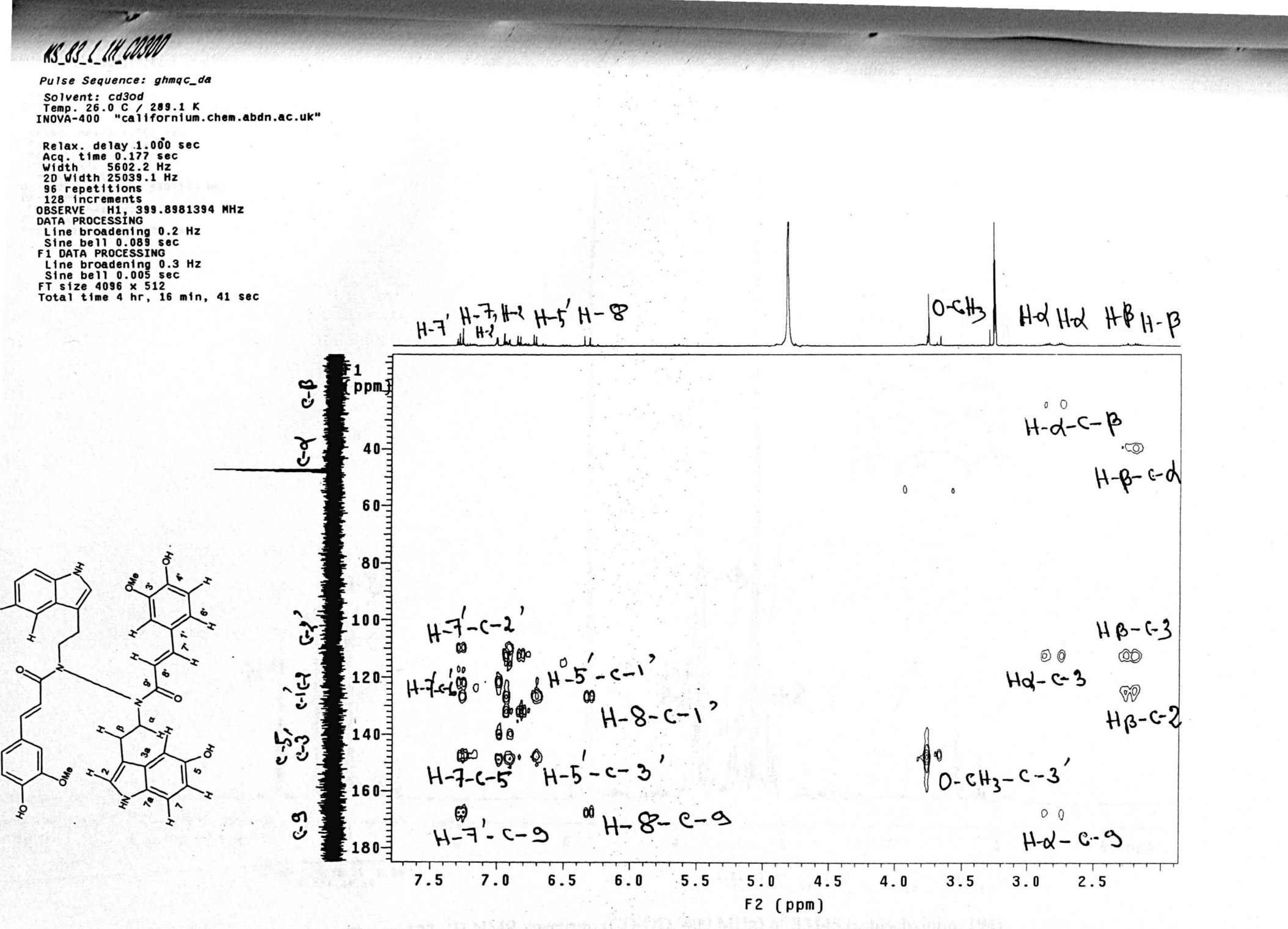


Figure 132. HMBC spectrum (CD_3OD, 400 MHz) of **SM44** (montamine, **193**)

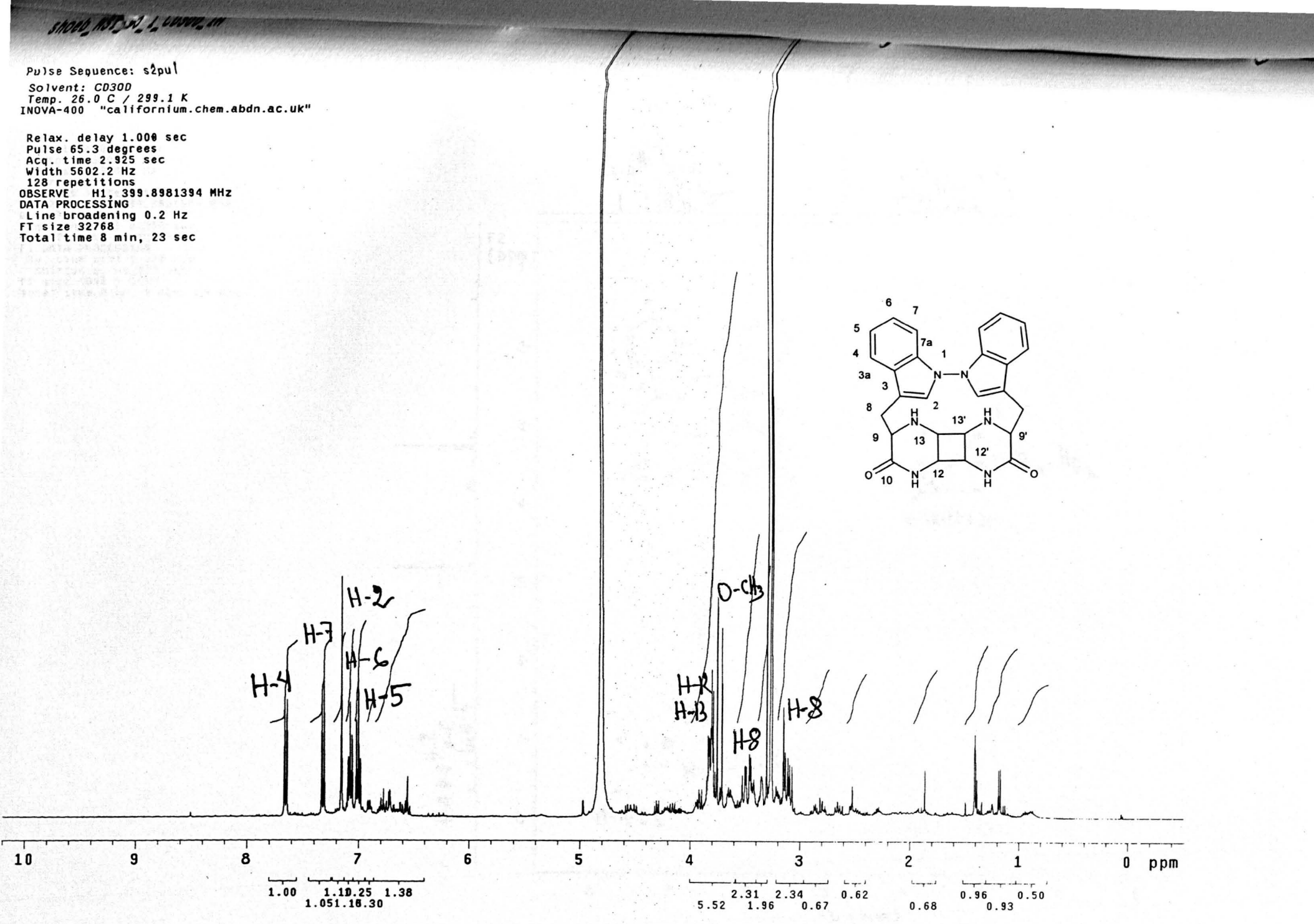


Figure 133. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of **SM45** (schischkiniin, **194**)

Pulse Sequence: gmqfcops_dd

Solvent: CD3OD
Temp. 26.0 C / 299.1 K
INOVA-400 "californium.chem.abdn.ac.uk"

Relax. delay 1.000 sec
Acq. time 0.300 sec
Width 3304.4 Hz
2D Width 3304.4 Hz
16 repetitions
2 x 256 increments
OBSERVE H1, 399.8981394 MHz
DATA PROCESSING
Sq. sine bell 0.300 sec
Shifted by -0.300 sec
F1 DATA PROCESSING
Sq. sine bell 0.078 sec
Shifted by -0.077 sec
FT size 4096 x 1024
Total time 3 hr, 5 min, 30 sec

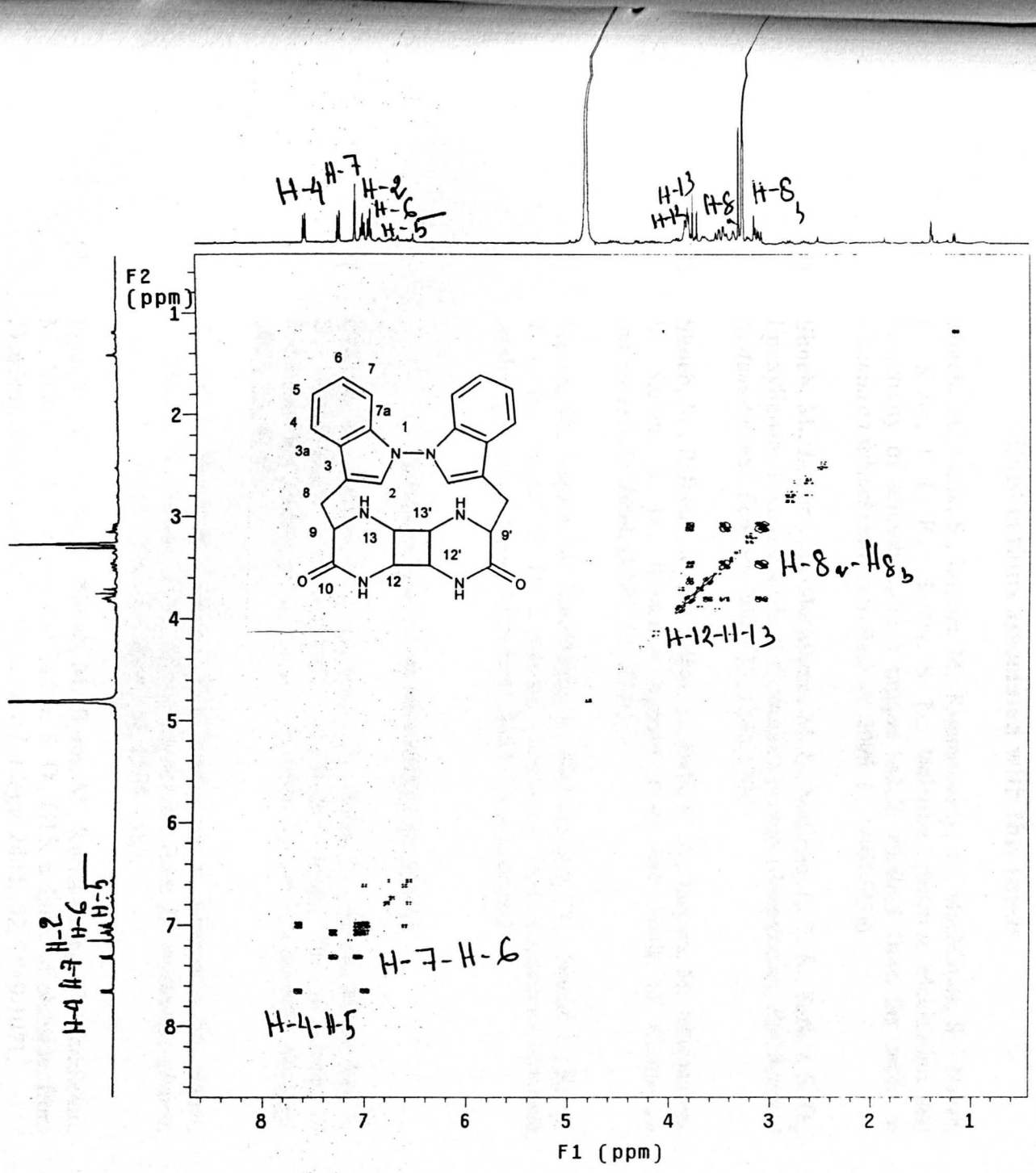


Figure 134. ^{1}H-^{1}H COSY spectrum (CD_3OD, 400 MHz) of **SM45** (schischkiniin, **194**)

Appendix B

Publications associated with this thesis

(1) **Shoeb, M.**, Celik, S., Jaspars, M., Kumarasamy, Y., MacManus, S., Nahar, L., Kong, T. L. P., Sarker, S. D., Isolation, structure elucidation and bioactivity of schischkiniin, a unique indole alkaloid from the seeds of *Centaurea schischkinii. Tetrahedron,* **2005**, 61, 9001-9006.

(2) **Shoeb, M.**, Jaspars, M., MacManus, M. S., Majinda, R. T. R., Sarker, S. D., Epoxylignans from the seeds of *Centaurea cyanus* (Asteraceae). *Biochemical Systematic and Ecology,* **2004**, 32, 1201-1204.

(3) **Shoeb, M.**, Rahman, M. M., Nahar, L., Delazar, A., Jaspars, M., MacManus, S., Sarker, S. D., Bioactive lignans from the seeds of *Centaurea macrocephala,* **2004**, *DARU* 12, 87-93.

(4) **Shoeb, M.**, Jaspars, M., MacManus, S., Kumarasamy, Y., Nahar, L., Kong, T. L. P., Sarker, S. D., Cytotoxic compounds from *Centaurea montana* seeds. *Journal of Natural Products,* **2004**, (Manuscripts)

Publications from secondary projects

(5) Delazar, A., Gibbons, S., Kumarasamy, Y., Nahar, L., **Shoeb, M.**, Sarker, S. D., Antioxidant phenylethanoid glycosides from the rhizomes of *Eremostachys glabra* (Lamiaceae). *Biochemical and Systematic Ecology,* **2005**, 33, 87-90.

(6) Delazar, A., Byres, M., Gibbons, S., Kumarasamy, Y., Modarresi, M., Nahar, L., **Shoeb, M.,** Sarker, S. D., Iridoid Glycosides from *Eremostachys glabra. Journal of Natural Products,* **2004**, 67, 1584-1587.

(7) Egan, P., Middleton, P., **Shoeb, M.**, Byres, M., Kumarasamy, Y., Middleton, M., Nahar, L., Delazar, A and Sarker, S. D., G15, a dimer of oleoside, from *Fraxinus. Biochemical and Systematic Ecology,* **2004**, 32, 1069-1071.

(8) Kumarasamy, Y., Byres, M., Cox, P. J., Delazar, A, Jaspars, M., Nahar, L., **Shoeb, M.**, and Sarker, S. D., Isolation, structure elucidation, and biological activity of flavone *C*-glycosides *Alliaria petiolata. Chemistry of Natural Compounds,* **2004**, 40, 122-128.

(9) Cox, P. J., Kumarasamy, Y., Nahar, L., Sarker, S., **Shoeb, M.**, Luteolin. *Acta Cyrstallographica,* **2003**, E 59, o975-o977.

(10) Cox, P. J., Jaspars, M., Kumarasamy, Y., Nahar, L., Sarker, S and **Shoeb, M.**, A mixed crystal of imperatorin and phellopterin, with C-H...O, C-H...π and π-π interactions. *Acta Cyrstallographica,* **2003**, C59, o520-o522.

MS1-41-2

University of Botswana
Chemistry Department
NABSA NMR Service

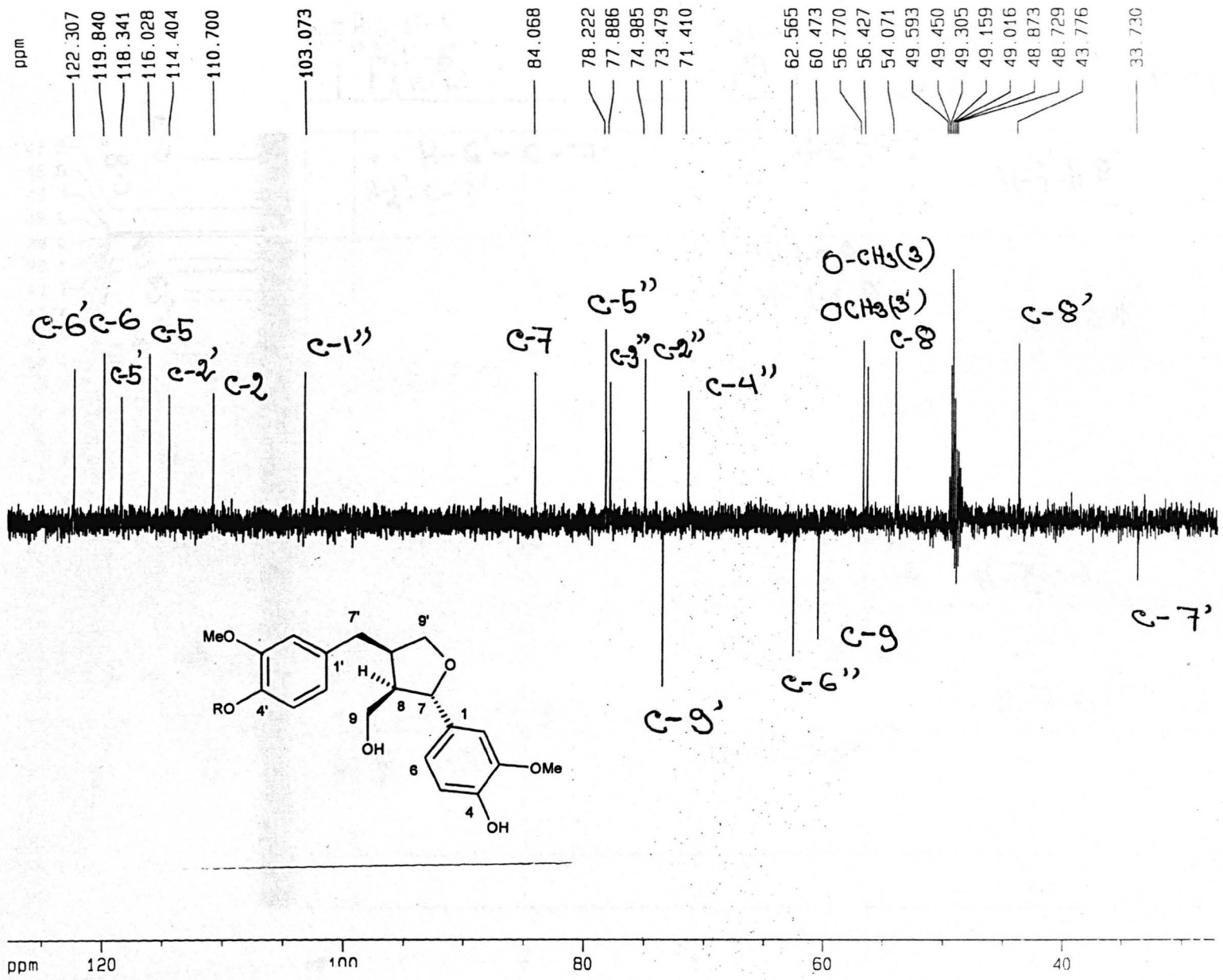


Figure 91. DEPT-135 spectrum (CD_3OD, 100 MHz) of **SM7** (lariciresinol 4'-*O*-β-D-glucopyranoside, **171**)

```
Current Data Parameters
NAME                  ms1.41.2
EXPNO                        3
PROCNO                       1

F2 - Acquisition Parameters
Date                  20030826
Time                      6.14
INSTRUM                  spect
PROBHD      5 mm BBI 1H-
PULPROG                dept135
TD                        65536
SOLVENT                   MeOH
NS                        10240
DS                            4
SWH                  35971.223 Hz
FIDRES               0.548877 Hz
AQ                   0.9110004 sec
RG                        13004
DW                   13.900 usec
DE                    6.00 usec
TE                    300.0 K
D1                   2.00000000 sec
D2                   0.00357143 sec
D12                  0.00002000 sec
DELTA                0.00001783 sec

============ CHANNEL f1 =============
NUC1                        13C
P1                    14.00 usec
P2                    28.00 usec
PL1                   -2.00 dB
SFO1                150.9178388 MHz

============ CHANNEL f2 =============
CPDPRG2                 waltz16
NUC2                        1H
P3                    10.00 usec
P4                    20.00 usec
PCPD2                 85.00 usec
PL2                    1.70 dB
PL12                  21.00 dB
SFO2                600.1324005 MHz

F2 - Processing parameters
SI                          32768
SF                  150.9025934 MHz
WDW                           EM
SSB                            0
LB                     1.00 Hz
GB                             0
PC                     1.40

1D NMR plot parameters
CX                    20.00 cm
F1P                  127.836 ppm
F1                   19290.75 Hz
F2P                   27.125 ppm
F2                   4093.19 Hz
PPMCM                5.03656 ppm/cm
HZCM                 759.8738 Hz/cm
```