Targeted isolation of diverse alkaloids from *Alchornea rugosa*, *Persea americana*, and *Gymnema inodorum* by dereplication methods using LC-MS/MS-based molecular networking

LC-MS/MS 기반 분자네트워킹의 선택적 표적화 방법을 이용한 *Alchornea rugosa, Persea americana*와 *Gymnema inodorum*의 다양한 alkaloids 표적분리 연구

2022 년 8 월

서울대학교 대학원

약학과 생약학 전공

Doan Thi Phuong

Targeted isolation of diverse alkaloids from *Alchornea rugosa*, *Persea americana*, and *Gymnema inodorum* by dereplication methods using LC-MS/MS-based molecular networking

LC-MS/MS 기반 분자네트워킹의 선택적 표적화 방법을 이용한 *Alchornea rugosa*, *Persea americana*와 *Gymnema inodorum*의 다양한 alkaloids 표적분리 연구

지도교수 오 원 근

이 논문을 약학박사 학위논문으로 제출함
2022 년 6 월

서울대학교 대학원
약학과 생약학전공
Doan Thi Phuong

Doan Thi Phuong의 약학박사 학위논문을 인준함
2022 년 7 월

위 원 장 ＿＿＿＿＿＿＿＿＿＿＿ (인)

부위원장 ＿＿＿＿＿＿＿＿＿＿＿ (인)

위 원 ＿＿＿＿＿＿＿＿＿＿＿ (인)

위 원 ＿＿＿＿＿＿＿＿＿＿＿ (인)

위 원 ＿＿＿＿＿＿＿＿＿＿＿ (인)

Abstract

Targeted isolation of diverse alkaloids from *Alchornea rugosa*, *Persea americana*, and *Gymnema inodorum* by dereplication methods using LC-MS/MS-based molecular networking

Natural products (NPs) that are small molecules derived from the environment by secondary metabolic pathways have played significant roles in drug discovery, ecology, and biotechnology. The conventional investigation of natural products in the early years such as "grind and find" or bio-guided isolation has successfully found important medicines including morphine (painkiller), aspirin (anti-inflammation), digitoxin (cardiac glycoside), silymarin (treatment of liver disease), quinine, artemisinin (antimalarial), and pilocarpine (treating glaucoma). Since the first plant-derived medicine was founded in 1803, the number of new natural products reported every year has risen dramatically. As many easily accessed natural products have been isolated for a long history, recent studies have faced a commonplace challenge – the "rediscovery of known compounds". In this regard, analytical tools, especially mass spectrometry (MS), have recently been efficiently applied in the NP field for dereplication and rapid identification of new NPs in complex mixtures.

Remarkably, a recent technique using tandem MS/MS data sets called molecular networking (MN) has transited the traditional "grind and find" to the hypothesis-driven targeting isolation since its first introduction a decade ago. MN is the computer-based approach aiming to arrange MS/MS data, automatically compare it to available databases for the identification of specific metabolites, and give the best

visualization of related compounds in the same group (cluster). This dereplication technique has been widely and successfully applied in NPs leading to the discovery of broad types of novel secondary metabolites that could be promising for research and development of new drugs.

Herein, the MN was successfully applied to target and isolate various types of alkaloids from different plant sources.

Part 1: Guanidine alkaloids from *Alchornea rugosa* and their glucose uptake and autophagy activities

Natural guanidines that are classified as alkaloids and as unusually modified peptides with guanidine functionalities are generally found in microorganisms and marine invertebrates but are rarely found in terrestrial plants and animals. The substantial hydrophilicity of guanidines makes them suitable for the development of drugs, some of which have been used clinically. It is noteworthy that although the natural guanidines in plants account for only a small number compared to other sources, plant guanidines have been frequently reported in some species belonging to the genus *Alchornea* including *A. floribunda*, *A. hirtella*, and *A. cordifolia*. Based on MS/MS-based molecular networking analysis, three pairs of novel configurationally semistable diastereomers featuring an unprecedented 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amine scaffold named rugonidines A–F (**1–6**), and nine new guanidine-fuse catechins named rugonines A–H (**7–15**) were isolated from *Alchornea rugosa*. Compounds **1–6** possess a 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-aminefused catechin skeleton that features a naturally unique scaffold. Their structures were elucidated by NMR spectroscopy in combination with quantum-

chemical calculations. The biological activities of **1–6** were tested in glucose uptake levels in differentiated 3T3-L1 adipocytes using 2-deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-d-glucose (2-NBDG) as a fluorescent-tagged glucose probe and compounds **1–3** showed a significant increase glucose uptake levels. Structures of **7–15** were determined by NMR, HRESIMS/MS, and CD data analysis. These guanidine-fused catechins were evaluated for their bioactivities in the autophagy modulation assay. The results indicated that compounds **10–13** had potential autophagy inhibition effects compared to the chloroquine which is known as autophagy inhibitor.

Part 2: Alkaloids from *Persea americana* and their SIRT1 activities

Sirtuin 1 (SIRT1) is a nicotinamide adenosine dinucleotide (NAD)$^+$-dependent deacetylase belonging to the mammalian sirtuins family that plays important roles in cellular and organismal processes, including metabolism and aging. Through deacetylation of various substrates, including histone proteins (acetylated histones H4K16 and H3K56) and non-histone targets (p53) to generate nicotinamide and the acetyl group, SIRT1 is involved in a broad range of physiological functions, including control of gene expression, metabolism, and aging. The function of SIRT1 in senescence, which leads to extending life span, ameliorates cellular senescence, and consequently prevents aging-related diseases, is mainly achieved by catalyzing the deacetylation of various downstream transcription factors.

Avocado seeds are the by-products of the food industry with rich bioactive compounds but the secondary metabolites have not been yet studied and applied. Herein, the phytochemicals in avocado seeds were dereplicated by using an LC-

MS/MS-based molecular networking. A new finding in alkaloid compositions from avocado seeds was discovered allowing the isolation of seven new quinolone-alkaloids and two new benzoxazinone-alkaloids. Structures of isolated compounds were identified by NMR spectroscopy methods in combination with ECD calculations for absolute configuration elucidation. Later, the activities of isolated compounds were tested for their SIRT1 activities in HEK293 cells. The results showed that compound **1** has the most potent effect on SIRT1 activation with an elevated $NAD^+/NADH$ ratio and it could be a promising candidate for further investigation of anti-aging agents.

Part 3: Oleanane triterpenoids from *Gymnema inodorum* and their insulin mimetic activities

The molecular networking of *n*-hexane, EtOAc, and *n*-BuOH fractions of *G. inodorum* was generated and showed four main clusters including oleanane – triterpenoids (cluster 1), flavonoids (cluster 2), methyl anthranilate derivatives of oleanane triterpenoids (cluster 3), and benzoyl derivatives of oleanane triterpenoids (cluster 4). Seventeen compounds including 13 new oleananes (**4**–**7**), benzoyl derivatives of oleanane (**3**), and methyl anthranilate derivatives of oleanane triterpenoids (**1, 2, 10**–**18**), and four known compounds were isolated and their structures were identified by analysis of the NMR spectra. Compounds **1**–**9** were tested in an insulin-mimetic model to evaluate their biological activities. The results suggested that compounds **3**, **5**, **8**, and **9** showed potential stimulatory effects on the uptake of 2-NBDG in 3T3-L1 adipocyte cells. Compounds **1**, **8**, **9**, and **10**–**18** were evaluated for their effects on antimuscle atrophic activities. The results suggested that compound **12** showed the most potent effect on muscle atrophy at 2 μM. Later,

the SAR was discussed to suggest the crucial point for antimuscle atrophy of oleanane triterpenoids from *G. inodorum*.

Keywords: *Alchornea rugosa, Persea americana, Gymnema inodorum*, molecular networking, alkaloids, oleanane triterpenoids, glucose uptake, autophagy inhibition, SIRT1 activation, insulin mimetic.

Student Number: 2017–26762

Table of Contents

List of Scheme

List of Tables

List of Figures

List of Abbreviations

$[\alpha]_D^{20}$	Specific rotation at 20-celsius degree
2-NBDG	2-[N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino]-2-deoxyglucose
br s	broad singlet
Calcd.	calculated
CD	circular dichroism
CE	Cotton effect
CMAE	Corrected mean absolute
CMAE	Corrected mean absolute error
COSY	Correlation spectroscopy
CQ	Chloroquine
d	Doublet
dd	Doublet of doublet
DBE	Double bond equivalents
DFT	Density Functional Theory
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DMSO-d_6	Deuterated dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
ECD	Electronic circular dichroism
EtOAc	Ethyl acetate
EtOH	Ethanol
FA	Formic acid
FBMN-GNPS	Feature-Based Molecular Networking-GNPS
FBS	Fetal bovine serum
GFP-LC3	Green Fluorescent Protein - Light Chain 3
GNPS	Global Natural Product Social Molecular Networking
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High-performance liquid chromatography
HRESIMS	High-resolution electrospray ionization mass spectrometry
HSQC	Heteronuclear Single Quantum Coherence
IR	Infrared Spectroscopy
LC3-II	Light Chain 3
LC-MS/MS	Liquid Chromatography with tandem mass spectrometry
m/z	Mass-to-charge ratio
MAE	Mean absolute error
MaxErr	Maximum error
MeOH	Methanol
MPLC	Medium pressure liquid chromatography
MTT	(3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide
NAD	Nicotinamide adenosine dinucleotide
n-BuOH	*n*- butanol
NMR	Nuclear Magnetic Resonance Spectroscopy
NOESY	Nuclear Overhauser Effect Spectroscopy
OR	Optical rotation
PBS	Phosphate-buffered saline

PVDF	Polyvinylidene Fluoride
QTOF	Quadrupole Time of Flight Mass Spectrometer
R^2	Coefficient of determination
RAPA	Rapamycin
RMD	Relative mass defect
RMSE	Root-mean-square
rmse	Root-mean-square error
ROESY	Rotating Frame Overhauser Enhancement Spectroscopy
SD	Standard deviations
SIRT1	Sirtuin 1
t	triplet
TAME	targeting aging by metformin
TDDFT	time-dependent DFT
TFA	Trifluoroacetic acid
t_R	retention time
UPLC	Ultra High Performance Liquid Chromatography
UPLC-QTOF	Ultra-high performance liquid chromatography-quadrupole
MS/MS	time-of-flight mass spectrometry
WHO	World Health Organization

Part 1: Guanidine alkaloids from *Alchornea rugosa* and their glucose uptake and autophagy activities

1. Introduction

1.1. Study background

1.1.1. Natural guanidines

Natural guanidines are classified as alkaloids and as unusually modified peptides with guanidine functionalities (Berlinck, 2020; Berlinck et al., 2012). Guanidines derived from natural products are generally found in microorganisms and marine invertebrates but are rarely found in terrestrial plants and animals (Berlinck, 2020). Since the guanidine group has limited occurrence in nature, guanidine-bearing compounds are structurally unique and often exhibit potent biological activities (Berlinck, 2020; Berlinck et al., 2012). The substantial hydrophilicity of guanidines makes them suitable for the development of drugs, some of which have been used clinically, such as streptomycin (Berlinck et al., 2021). Several guanidines from plants have been reported to have antidiabetic (Perla & Jayanty, 2013), anti-inflammation (Song et al., 2015), antitumor (Mao et al., 2015), antifungal (Kohyama & Ono, 2013), and antifeedant (Rowan et al., 1990) activities. In addition to the reported activities of natural guanidines, a marine guanidine named monanchocidin has shown potential autophagy induction and lysosomal membrane permeabilization effects (Dyshlovoy et al., 2015). In addition, metformin, a well-known biguanide that is the first-line therapy in the treatment of type 2 diabetes, has been found to help attenuate hallmarks of aging by enhancing autophagy (Bharath et al., 2020; Kodali et al., 2021). Moreover, guanidine hydrochloride has been used for treating muscle weakness since the 1960s (Minot et al., 1939), for treating amyotrophic lateral sclerosis (Norris, F. H., Calanchini, P. R., Fallat, R. J., Panchari, S., & Jewett, 1974), and for treating infantile and juvenile spinal muscular atrophy (Angelini et al., 1980). It is noteworthy that although the natural guanidines in plants account for only a small number compared to other sources (Berlinck et al., 2012, 2017), plant guanidines have been frequently reported in some species belonging to the genus *Alchornea* including *A. floribunda*, *A. hirtella*, and *A. cordifolia* (Berlinck et al., 1999; Berlinck & Kossuga, 2005)
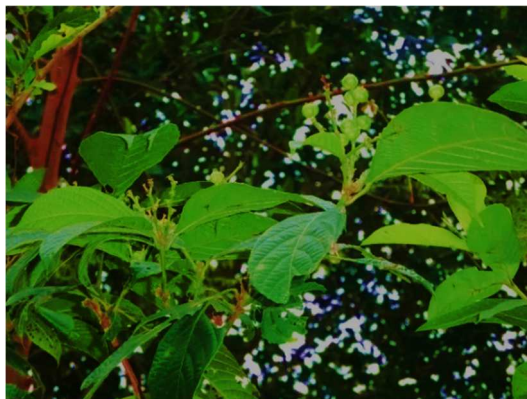
Figure 1. *Alchornea rugosa* (Lour.) Müll.Arg. Leaves of *A. rugosa* were collected in Quang Trung commune, Ngoc Lac district, Thanh Hoa Province, Vietnam (20°8′3″N, 105°24′15″E) in May 2016.

1.1.2. *Alchornea rugosa* (Lour.) Müll.Arg.

 Alchornea rugosa (Lour.) Müll.Arg., belonging to the family Euphorbiaceae, is widespread from India to Australia and grows in lowland rainforests as a dioecious shrub or tree up to 10 meters tall (Martínez et al., 2017; Wiart, 2007). Traditional uses in Vietnam, Laos, and Cambodia have indicated that the seeds of *A. rugosa* may be eaten to relieve the bowels from costiveness. Roots and leaves are used as a decoction to reduce fever and treat malaria (Wiart, 2007). Although the genus *Alchornea* comprises 55 accepted species, which occur in highly varied ecosystems across all continents with a largely pantropical distribution, most of the studies on this genus focus on *A. cordifolia,* which grows mainly in Africa (Martínez et al., 2017). Leaves and bark tissues of *A. rugosa* was first reported in 1969 to contain alkaloids, including natural guanidines named N^1,N^1-diisopentenylguanidine, N^1,N^2,N^3-triisopentenylguanidine, and two hexahydroimidazo[1,2-*α*]pyrimidines alkaloids (alchornine and alchornidine) (Hart et al., 1969; Martínez et al., 2017). These guanidine alkaloids have shown antibacterial, antiangiogenic, and anti-inflammatory activities in previous studies (Lamikanra et al., 1990; Mavar-Manga et al., 2008).
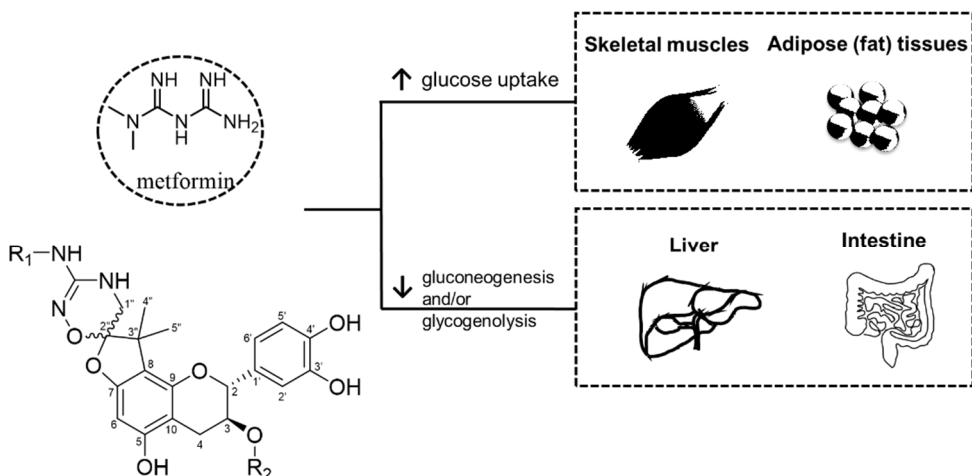
Figure 2. Structure and mechanisms of action of metformin in type 2 diabetes and structures of type 1 guanidines isolated from *A. rugosa.*

1.1.3. Diabetes

Diabetes is a metabolic disease characterized by chronic hyperglycemia that could lead to complications in the heart, blood vessels, eyes, kidneys, and nerves over time (Alberti & Zimmet, 1998; Association, 2014). It is estimated that approximately 90% of diabetes cases are type II diabetes, which is caused by *β*-cell dysfunction and insulin resistance (Association, 2014; Harding et al., 2019). According to the WHO, the number of cases and the prevalence of diabetes have been steadily increasing over the past few decades, with approximately 422 million people experiencing this disease worldwide and 1.6 million deaths directly caused by diabetes each year (World health organisation (WHO), 2020). The first-line therapy is a comprehensive lifestyle including weight management and physical activities accompanied by biguanides (metformin), followed by oral antihyperglycemic agents and insulin (Association, 2021). In addition to drug therapy, traditional medicines have been described in pharmacopeias and well-established documents in WHO monographs (Governa et al., 2018; WHO, 2007; World Health Organisation, 1999; World Health Organization, 2009). A recent review reported that 2004 plants belonging to 1112 genera and 197 families have been reported across 92 countries worldwide for diabetes management (Aumeeruddy & Mahomoodally, 2021). *A. cordifolia,* a species belonging to the genus of *Alchornea*, has antidiabetic activity associated with its *n*-butanol fraction (Mohammed et al., 2012).

3
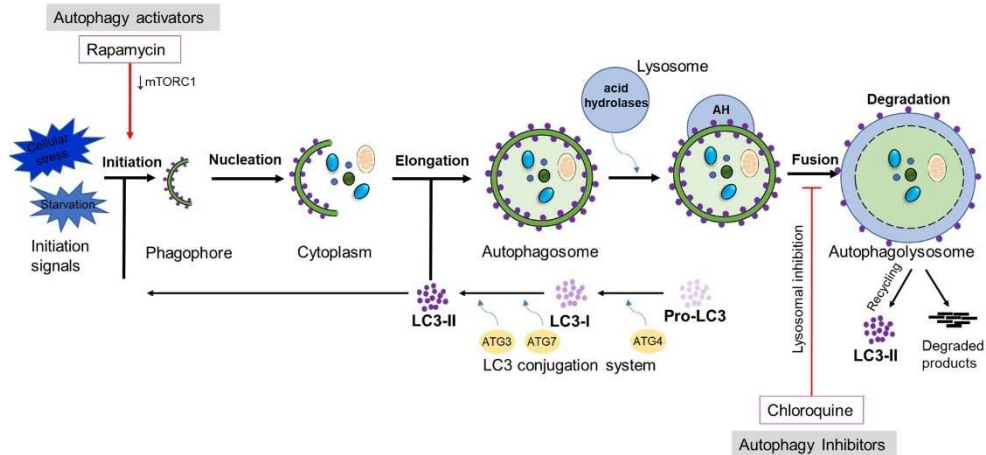
Figure 3. Overview of the autophagic machinery. Initiation of autophagy involves the formation of an isolation membrane (phagophore). This critical step is stimulated by starvation or cellular stress, or some small molecules such as rapamycin. The phagophore then sequesters cytoplasm or selected substrates, elongates (with the presence of LC3-II), and eventually matures into an autophagosome with a completed double-membrane. The autophagosome fuses with a lysosome that contains hydrolase acids to form an autophagolysosome. The following breakdown step results in degraded products and LC3-II which can be recycled for other cycles of autophagy.

1.1.4. Autophagy regulator

Autophagy is an intracellular self-degradative process that eliminates cytoplasmic materials and recycles damaged organelles, and nonfunctional proteins in the lysosomes to maintain basic energy levels for cellular regeneration and homeostasis (Mizushima & Komatsu, 2011). The initiation step of autophagy involves the formation of a double membrane (phagophore) that is regulated by multiple signaling mechanisms. The phagophore then elongates and eventually matures into an autophagosome, which is subsequently trafficked to fuse with a lysosome (Vakifahmetoglu-Norberg et al., 2015). Malfunction or overactivation of the autophagic machinery is associated with many human disorders including cardiomyopathies, infectious diseases, Crohn's disease, and neurodegenerative disorders (Alzheimer's, Huntington's, and Parkinson's diseases), and cancer (Vakifahmetoglu-Norberg et al., 2015). The hyperactivation of autophagy has also

been found to be associated with muscle loss in various conditions and to induce muscle degeneration and weakness (Sartori et al., 2021b). The importance of autophagy in pathophysiology has drawn the attention of many researchers in recent decades in the search for new potential therapeutic drugs. Therefore, autophagy modulators have been investigated as potential candidates for the treatment of autophagy-related diseases. Many small molecules have been reported as autophagy activators or inhibitors (Vakifahmetoglu-Norberg et al., 2015) including curcumin, tanshinone IIa, ursolic acid, resveratrol, and ginsenoside 20(S) Rg-3 (enhancing autophagy), and 6-gingerol (inhibiting autophagy) (Sohn & Park, 2017). Interestingly, the guanidine derivative N^2,N^4-dibenzylquinazoline-2,4-diamine (DBeQ) inhibits autophagosome maturation (Chou et al., 2011), and metformin, a biguanide which induces autophagy, are undergoing clinical trials.

1.2. Purpose of Research

Inspired by the unique structure and the activities of guanidine derivatives, molecular networking-guided isolation was applied to investigate natural guanidines from *A. rugosa*. The shared guanidine moiety in the structures of metformin and rugonidines A–F, which contains guanidine derivatives fused to the catechin skeleton via 1,6-dioxaspiro linkage, and by the antidiabetic effects of a species in the same genus, the isolated compounds were evaluated for their antidiabetic activity through the glucose uptake assay in 3T3-L1 adipocyte cells. For rugonines, the other type of guanidines isolated from *A. rugosa*, their activities have been performed in an autophagy screening to seek autophagy modulator agents due to their structural similarities with some autophagy activators or inhibitors which are under investigation for new therapeutic medicines.

2. Targeted isolation of guanidine derivatives in *A. rugosa*

2.1. Targeted isolation of guanidines from *A. rugosa* by using LC-MS/MS-based molecular networking

The chemical profile of the alkaloid fraction from the *A. rugosa* leaf extract was determined by UPLC-QTOF MS/MS analysis. High-resolution electrospray ionization mass spectrometry (HRESIMS) data (Figure 4) were then preprocessed via MZmine 2 software (Pluskal et al., 2010) and subsequently analyzed on the Global Natural Products Social platform (GNPS, https://gnps.ucsd.edu) hosted by the GNPS (M. Wang et al., 2016a). The MS/MS molecular networking displayed six clusters that were over three nodes (Figure 5). The dereplication process afforded 10 hits, four of which were previously described in the genus *Alchornea* (Figure 5 and Table 1). The largest cluster A with 20 nodes consisted mostly of small-molecule metabolites, some nodes of which could be assigned to known alkaloids reported in this genus, such as alchornin [m/z 208, (M + H)$^+$] (Hart et al., 1969), pteroginine [m/z 194, (M − H$_2$ + H)$^+$] (Barrosa et al., 2014a), pterogynidine [m/z 194, (M − H$_2$ + H)$^+$] (Barrosa et al., 2014a), isoalchorneine [m/z 240, (M + H$_2$ + H)$^+$) (Khuong-Huu et al., 1972), and alchornedine [m/z 210, (M + H)$^+$] (Barrosa et al., 2014a). The alchornealaxine with a molecular weight of 481 Da, obtained by conjugation between prenylated guanidyl and epicatechin moieties, was observed in cluster C (Figure 6). Analysis of the MS/MS fragmentation data revealed that alchornealaxine (black circle edge) and the other nodes in cluster C shared a similar pattern of product ions, suggesting the presence of derivatives with guanidyl moieties. Two typical ions at m/z 262 and m/z 151, which represented guanidine-attached fragments, were observed in all MS/MS patterns of the nodes in cluster C (Table 1, Figures 4, 7). According to molecular networking-based discovery, eight new guanidine-derived derivatives (**1**–**8**) were isolated. The different mass losses between those nodes suggested that the loss of a hydroxyl (398.1725, **14**), an additional rhamnose sugar (628.2881, **7**–**8**), the loss of an isopentenyl moiety (414.1670, **10**–**11**), a replace isopentenyl moiety by a rhamnose unit (560.2250, **13**). In addition, cluster **D** consisted of six nodes at m/z 644.2815, 576.2189, 576.2211, 498.2237, 482.2289, and 430.1604 that differed from those of the reported guanidine derivatives (Figure 5). Subsequent fractionation of the EtOH extract by column chromatography separation, including on Daion HP-20, MPLC, Sephadex LH-20, and

semipreparative HPLC, allowed the isolation of three pairs of diastereomers **1**–**6**. Compounds **1**–**6** possess a 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amine-fused catechin skeleton that features a naturally unique scaffold (Figure 8). Furthermore, eight new compounds (**7**–**14**) that share a similar skeleton to alchornealaxine in which guanidine conjugated with flavanol were also targeted and isolated, then their structures were determined to prove that the suggested nodes by molecular networking were a good match (Figure 9).
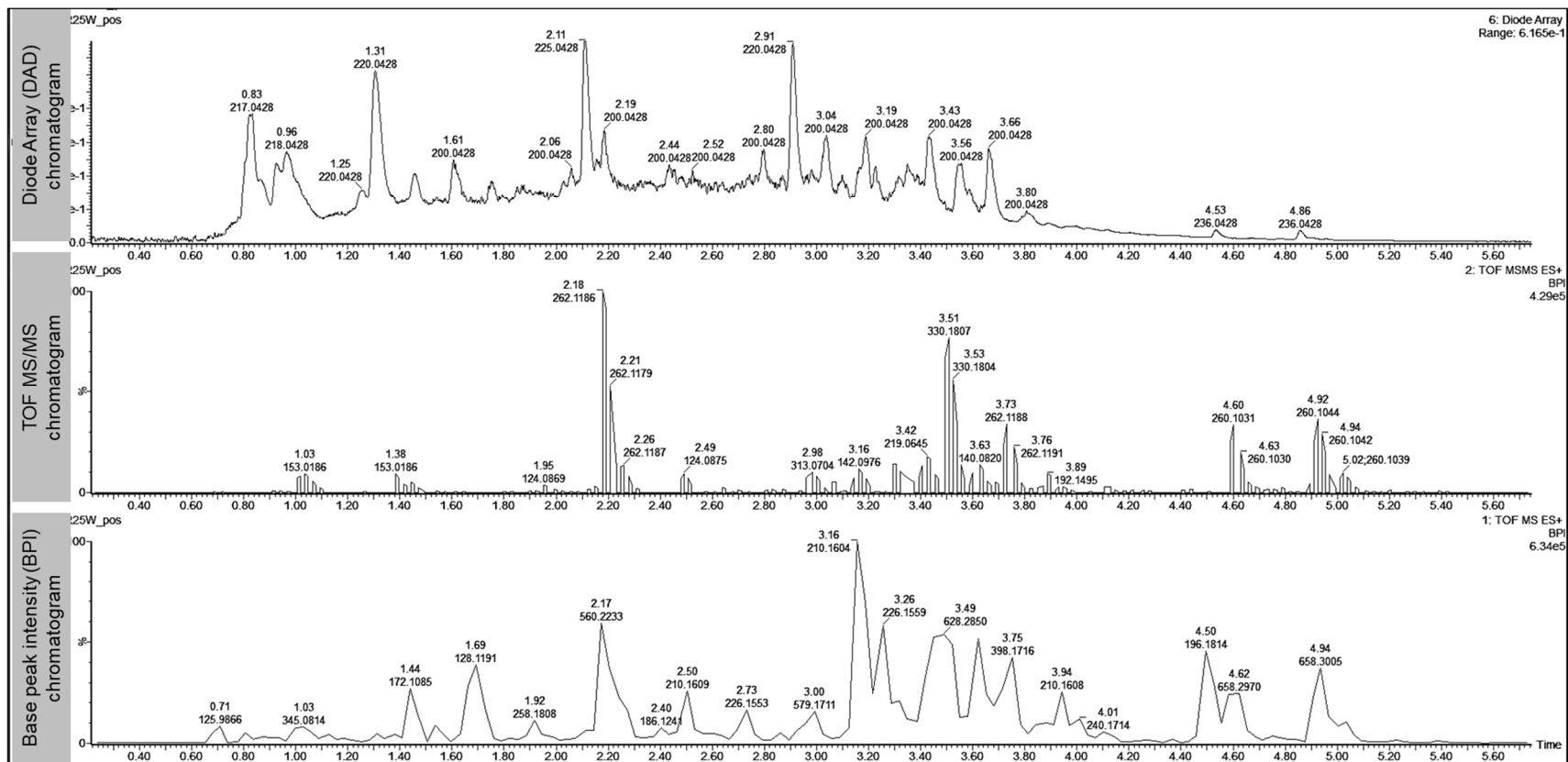
Figure 4. Base peak intensity, TOF MS/MS, and DAD chromatographies of alkaloid fraction of *A. rugosa* leaves
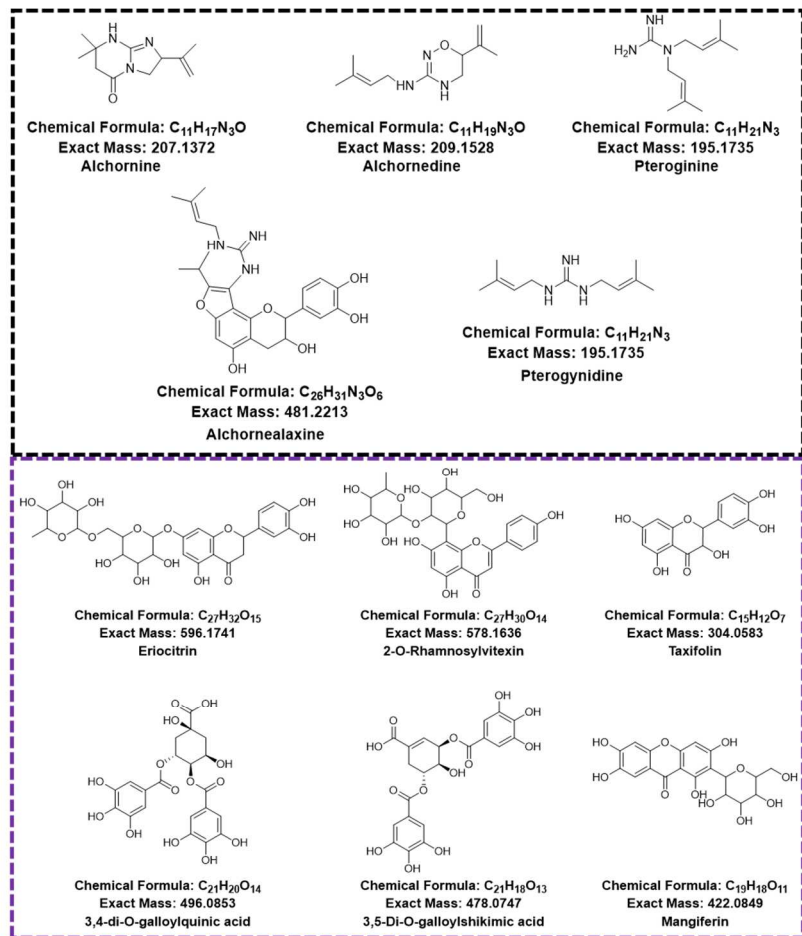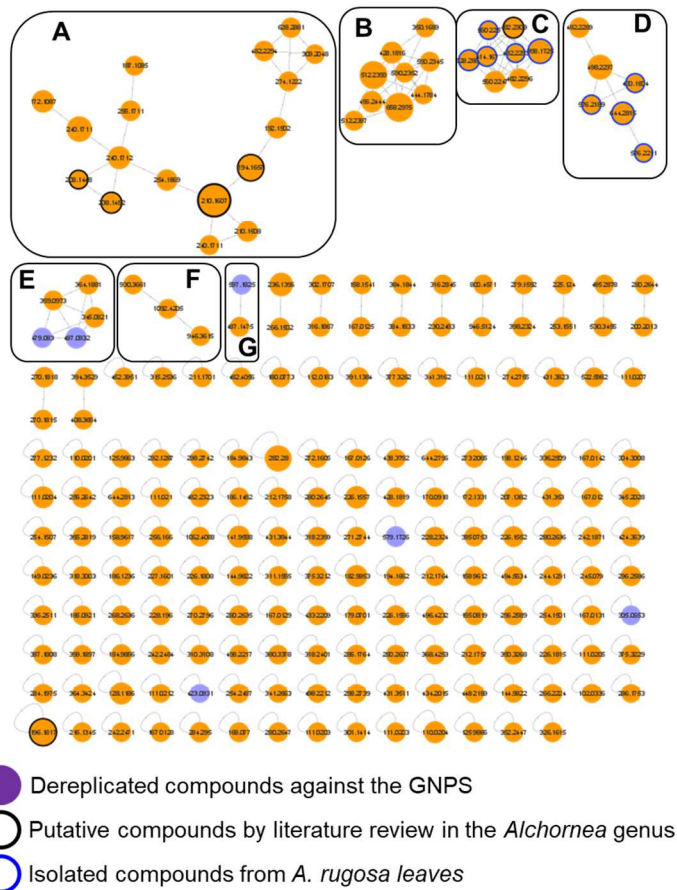
Figure 5. Dereplicated compounds against the GNPS and literature reviews in the *Alchornea* genus
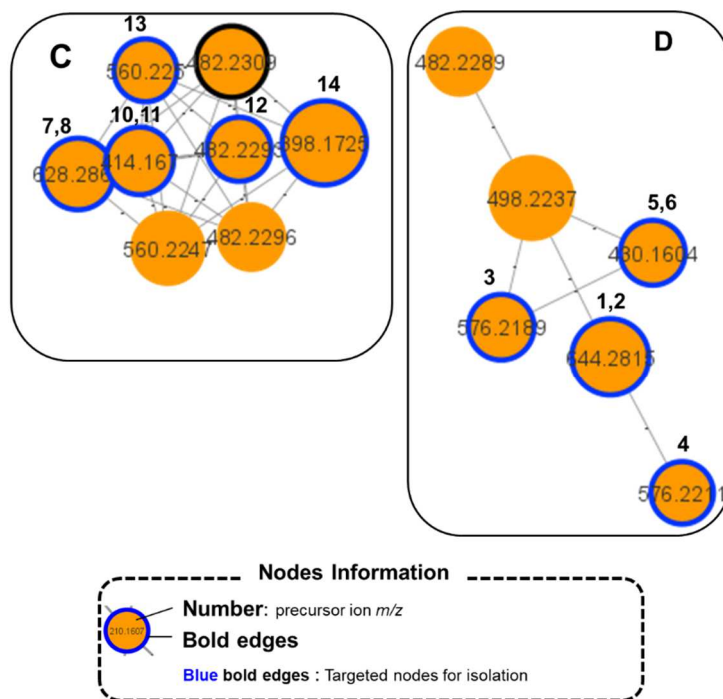
Figure 6. Molecular networking of the 25% MeOH fraction of the leave extraction of *A. rugosa*. A cluster of 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amine-fused catechin skeleton and targeted isolation nodes (this network is accessible at the following address:

http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=cb23e27317a24150abb45fad3fba59a8)
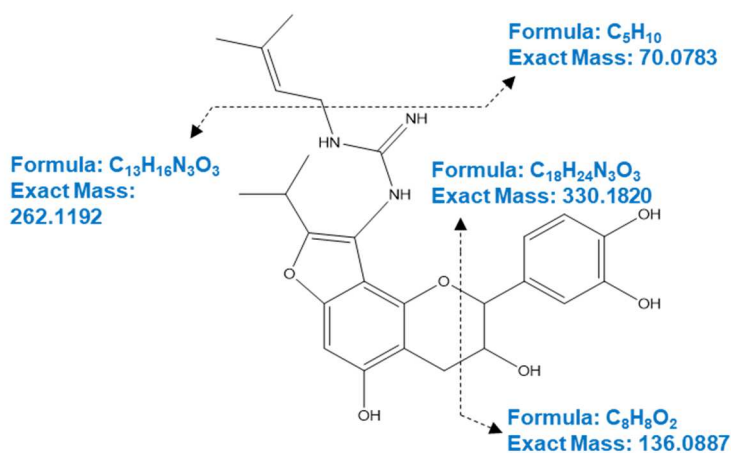
Figure 7. Mass fragmentation analysis of a known guanidine-epicatechin (alchornealaxine)

Table 1. Identified nodes in the fraction AR25 using HRESI-qTOF MS/MS against GNPS libraries or literature reviews in *Achornea* genus

Cluster group	t_R (min)	*m/z* [M+H]$^+$	Calcd.	MassDiff (ppm)	RMD	MS/MS fragments	Molecular formula	Compound name	GNPS spectral Library ID/ref.
A	2.94	208.1458	208.1450	3.80	700	140.0825	$C_{11}H_{17}N_3O$	Alchornine	(Hart et al., 1969)
A	3.14	210.1686	210.1705	-9.00	802	142.0983	$C_{11}H_{19}N_3O$	Alchornedine	(Barrosa et al., 2014a)
A	3.85	194.1657 [M-H$_2$+H]+	194.1657	-0.18	853		$C_{11}H_{21}N_3$	Pterogynine	(Hart et al., 1969)
B	4.63	658.2983	658.2978	1.10	453	154.0967, 218.0865, 260.1033, 328.1661, 462.2052,	$C_{33}H_{43}N_3O_{11}$		
B	4.82	428.1826	428.1822	0.90	426	164.0809, 218.0877, 242.0922, 260.1037	$C_{22}H_{25}N_3O_6$		
B	4.98	658.2987	658.2976	1.70	454	154.0970, 218.0873, 260.1041,328.1669, 512.2415	$C_{33}H_{43}N_3O_{11}$		
B	5.04	512.2415	512.2397	3.50	471	218.0878, 260.1037, 328.1661	$C_{27}H_{33}N_3O_7$		
B	5.21	512.2401	512.2437	0.80	469	164.0771, 191.0701, 218.0882, 242.0926, 260.1041, 328.1659,	$C_{27}H_{33}N_3O_7$		
B	5.97	496.245	496.2448	0.40	494	154.1002, 164.0776, 218.0880, 242.0925, 260.1030, 328.1658	$C_{27}H_{33}N_3O_6$		
B	6.10	496.2454	496.2448	1.20	495	164.0785, 218.0869, 242.0916, 260.1023, 328.1665	$C_{27}H_{33}N_3O_6$		
B	6.25	496.2451	496.2448	0.60	494	163.0534, 191.0736, 218.0839, 233.0818, 260.1037, 328.1644, 369.1326	$C_{27}H_{33}N_3O_6$		
B	6.48	496.2461	496.2448	2.60	496	163.0492, 191.0700, 218.0876, 233.0822, 260.1039, 328.1698, 369.1346	$C_{27}H_{33}N_3O_6$		
C	2.07	414.1674	414.1665	2.20	404	151.0374, 161.0599, 178.0588, 203.0640, 220.0885, 245.0932, **262.1190**	$C_{21}H_{23}N_3O_6$		
C	2.19	560.2274	560.2244	5.40	406	178.0599, 220.0888, 245.0925, **262.1200**, 414.1677	$C_{27}H_{33}N_3O_{10}$		
C	2.23	414.1647	414.1665	-4.30	398	161.0631, 178.0576, 203.0616, 220.0889, 245.0914, **262.1189**	$C_{21}H_{23}N_3O_6$		
C	3.49	628.2877	628.2870	1.10	458	161.0575, 220.0933, 245.0935, **262.2298**, 330.1815, 482.2284, 560.2220	$C_{32}H_{41}N_3O_{10}$		

C	3.53	628.2874	628.2870	0.60	457	178.0648, 203.0936, 245.0933, **262.1196**, 330.1818	$C_{32}H_{41}N_3O_{10}$		
C	3.74	398.1757	398.1716	10.3	441	151.0429, 161.0610, 178.0584, 203.0638, 220.0916, 245.0939, **262.1207**	$C_{21}H_{23}N_3O_5$		
D	3.64	644.2822	644.2819	0.50	438	140.0823, 156.1132, 191.0954, 203.0713, 260.1041, 278.1111, 330.1815, 430.1618, 498.2234	$C_{32}H_{41}N_3O_{11}$		
C, D	4.11	482.2316	482.2291	5.20	480	161.0679, 178.0596, 192.1441, 203.0670, 220.0907, 245.043, **262.1192**, 330.1820	$C_{26}H_{31}N_3O_6$	Alchornealaxine	(Tapondjou et al., 2016)
E	1.37	479.083	479.0826	0.91	173	125.0246, **153.0186**, 171.0323, 309.0615, 351.5561	$C_{21}H_{18}O_{13}$	3,5-di-O-galloylshikimic acid	o
E	1.38	497.0932	497.0931	0.11	187	125.0246, **153.0186**, 309.0615	$C_{21}H_{20}O_{14}$	3,4-di-O-galloylquinic acid	o
G	3.27	597.1825	597.1819	1.00	306	121.1045, 208.1438, **226.1544**, 350.0752	$C_{27}H_{32}O_{15}$	Eriocitrin	o
single node	4.50	196.1817	196.1814	1.46	926	**128.1189**, 152.0782, 175.0792, 196.1827	$C_{11}H_{21}N_3$	Pterogynidine	(Lopes et al., 2009; Mavar-Manga et al., 2008)
single node	3.40	305.0653	305.0661	-2.75	214	110.0214, 122.5465, 139.5476, 225.1362, 245.0834	$C_{15}H_{12}O_7$	Taxifolin	o
single node	2.10	423.0931	423.0927	0.97	220	102.0325, **110.0199**, 132.0642, 186.0953	$C_{19}H_{18}O_{11}$	Mangiferin	o
single node	2.98	579.1726	579.1714	2.14	298	**283.0603**, 313.0707, 337.0706, 361.0718, 379.0810	$C_{27}H_{30}O_{14}$	2-O-Rhamnosylvitexin	o
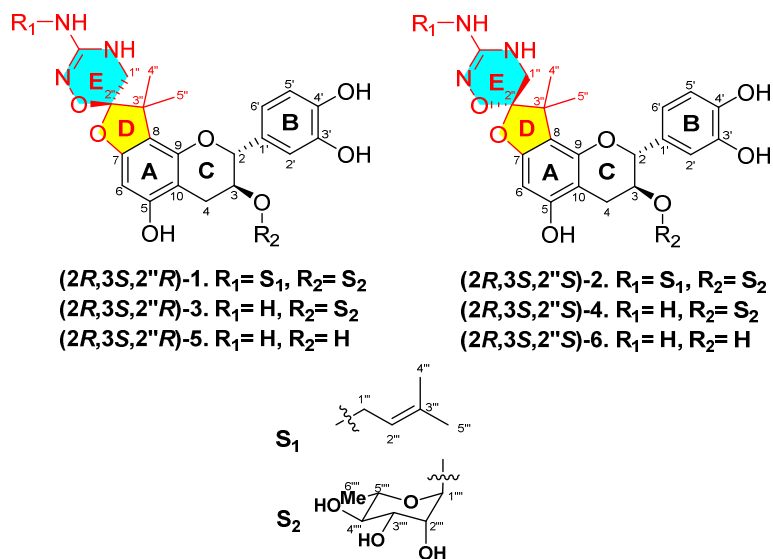
Figure 8. Chemical structures of isolated compounds 1–6 from *A. rugosa* leaves
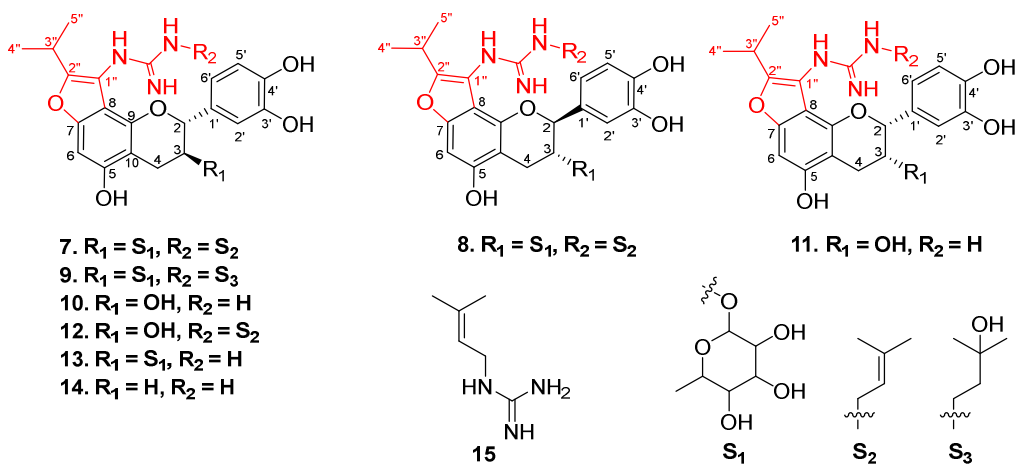
Figure 9. Chemical structures of isolated compounds 7–15 from *A. rugosa* leaves

3. Structure elucidation of 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amines from *A. rugosa*

3.1. Compound **1**

Rugonidine A (**1**) was obtained as a pale-yellow amorphous powder. Its molecular formula was established as $C_{32}H_{41}N_3O_{11}$ based on the HRESIMS ion peak at *m/z* 644.2833 [(M + H)$^+$, calcd. 644.2819], which accounted for 14 double bond equivalents (DBEs). The IR spectrum of **1** showed characteristic absorption bands for amine (3217 cm^{-1}), oxime (1689 cm^{-1}), alkenes (1523, 1647 cm^{-1}), and N–O (800 cm^{-1}) functionalities. The ^{1}H NMR spectrum of **1** showed two N–H protons (δ_H 9.59 and 9.08), four hydrogen signals at aromatic rings A and B (δ_H 6.79, 6.73, 6.66, and 5.96), an olefinic proton (δ_H 5.23, H-2'''), seven protons on oxygenated carbons (δ_H 3.12 – 4.63 ppm) including an anomeric signal (δ_H 4.05, H-1''''), an *N*-methylene group (δ_H 3.89/3.82, H-1'''), one methylene (δ_H 2.78/2.53, H-4), and five methyl groups (δ_H 1.11–1.72 ppm). The ^{13}C NMR spectrum of **1** exhibited 32 carbons, including a guanidine carbon signal (δ_C 159.0), twelve aromatic carbons (δ_C 89.3–156.0), one double bond (δ_C 136.3/118.6), a signal typical of a dioxaspiro carbon at δ_C 108.4 (Li et al., 2014), an anomeric carbon (δ_C 100.1), six oxygenated carbons (δ_C 68.9–79.1), one *N*-methylene group (δ_C 45.7), one quaternary carbon (δ_C 44.2), one methylene (δ_C 27.7), and five methyls (δ_C 17.8–25.3). The presence of a guanidine moiety was suggested by two nitrogenated protons as well as a carbon signal at δ_C 159.0, and the molecular formula of $C_{21}H_{23}N_3O_7$ for **1** resembled those of reported guanidine derivatives (Barrosa et al., 2014b; Tapondjou et al., 2016).

The HMBC correlations from H-2 (δ_H 4.63) to C-1' (δ_C 129.5), C-2' (δ_C 114.0), and C-6' (δ_C 118.0) and from H$_2$-4 (δ_H 2.78, 2.53) to C-10 (δ_C 101.0) suggested a catechin unit (C6-C3-C6). The heterocyclohexene ring of **1** (E ring), which was similar to that of 3-amino-1,2,4-oxadiazine scaffold (Tang et al., 2020), was determined by HMBC correlations from H$_2$-1'' (δ_H 3.95, 3.85) and 1''-NH (δ_H 9.59) to the guanidine carbon (δ_C 159.0) and C-2'' (δ_C 108.4) (Figure 12). The position of the isoprenyl fragment was indicated by the correlation from H$_2$-1''' (δ_H 3.89/3.82) to C=N (δ_C 159.0). The hexose sugar was recognized by an anomeric (H-1'''', δ_H 4.05/δ_C 100.1), four oxygenated (δ_H 3.12 – 3.48; δ_C 68.9 – 71.8), and one methyl (δ_H 1.11/δ_C 17.8) signals. ^{1}H–^{1}H COSY (Figure 13) and NOESY data of **1** (Figure 18), as well as the mass loss of 146 Da, revealed the presence of a rhamnose moiety. The HMBC correlation from

H-1"" to C-3 (δ_C 72.6) indicated that the hexose was linked to C-3 of the catechin moiety and that the relative configuration was an α-orientation based on the magnitude of the $^1J_{CH}$ value (169.0 Hz) (Pretsch et al., 2009). These results accounted for only 13 out of 14 DBEs of **1**, and thus suggested the presence of one additional ring. The HMBC spectrum of **1** implied the presence of an additional furan ring fused to the catechin moiety at C-7 and C-8, based on HMBC cross-peaks from H$_3$-4" and H$_3$-5" to C-8 (δ_C 109.6). The linkage of the D-E ring through C-2" was also proven by HMBC correlations from the H$_3$-4", H$_3$-5", and H$_2$-1" to the dioxospiro carbon at C-2". The configurations at C-2 and C-3 of **1** were identified as 2*R*,3*S* based on the NMR doublet for H-2 (δ_H 4.63, $J = 8.1$ Hz), and the carbon signals for C-2 at δ_C 79.1 and C-3 at δ_C 72.6 (Schmidt et al., 2010). The negative Cotton effect (CE) in the range 280–290 nm of the circular dichroism (CD) experiment of **1** (Figures 20) was also typical for (+)-catechin (Kim et al., 2016; Korver & Wilkins, 1971). The absolute configuration of the rhamnose was established by acid hydrolysis followed by conversion to its corresponding thiocarbamoyl-thiazolidine carboxylate derivative with L-cysteine methyl ester and *o*-tolyl isothiocyanate (Tanaka et al., 2007). The consistent HPLC retention times for the sugar in **1** and the authentic L-rhamnose derivative confirmed that the stereochemistry of the sugar moiety was that of L-rhamnose. To further confirm the unique 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amine-fused catechin moiety of **1**, NMR calculations for two possible isomers of the core structure of **1**, (2*R*,3*S*,2"*R*)-**1** (**I**) and (2*R*,3*S*,2"*S*)-**1** (**II**), were performed with density functional theory (DFT) at the B3LYP/6-311G (2df,2pd) level. The calculated ^{13}C NMR data matched well with the experimental data ($R^2 = 0.9972$ for **I** and 0.9967 for **II**, Figure 17A). The corrected mean absolute (CMAE) and root-mean-square (RMSE) errors for calculated ^{13}C NMR chemical shifts of 2.32 and 2.56 (**I**) and 2.61 and 2.69 (**II**), respectively, were similar to those seen with structurally well-established natural products (Hehre et al., 2019) (Figure 17B). The DP4+ probability of **I** was 100%, suggesting the 2*R*,3*S*,2"*R* relative configuration assigned for **1**.
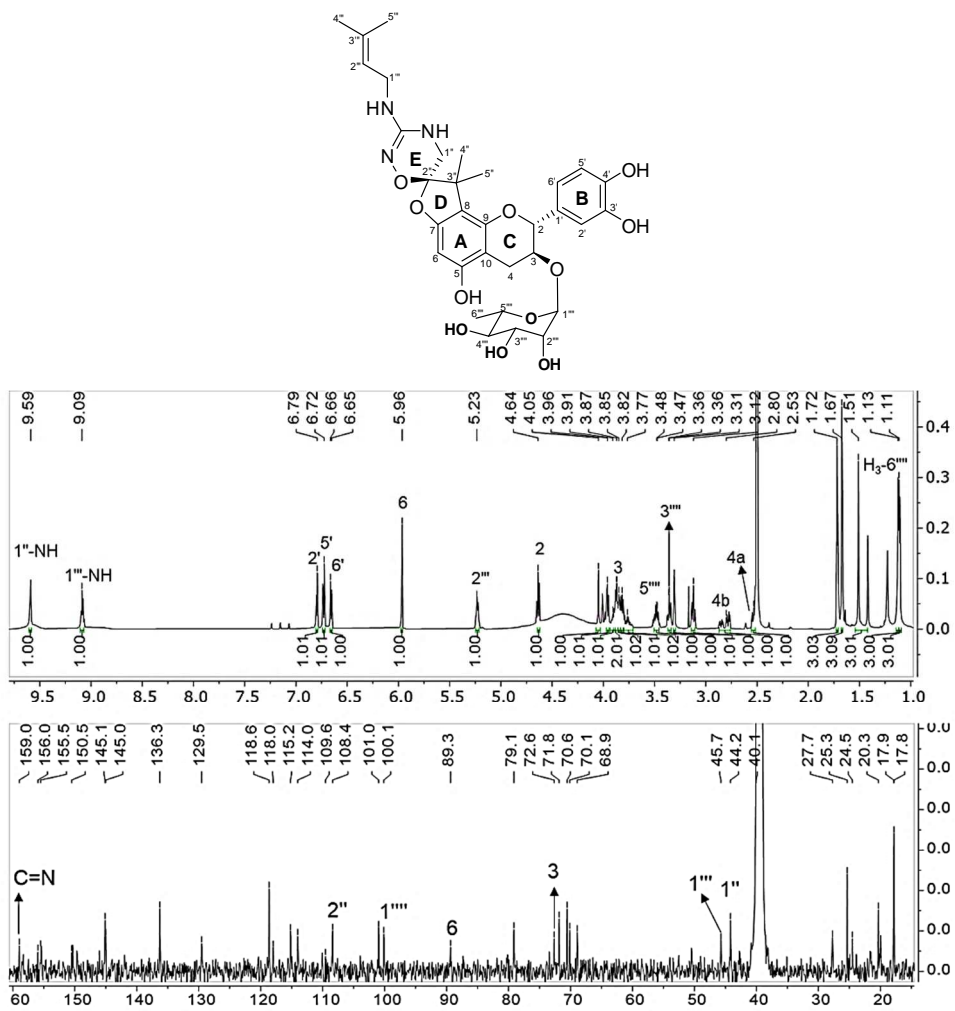
Figure 10. ^{1}H and ^{13}C NMR spectra of compound **1** (600/150 MHz, DMSO-d_6)
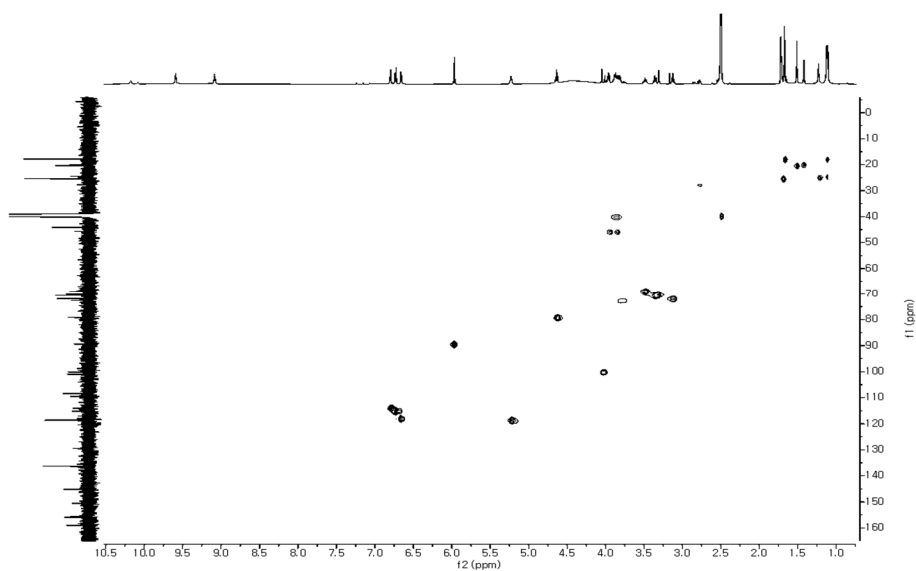
Figure 11. HSQC spectrum of compound **1**
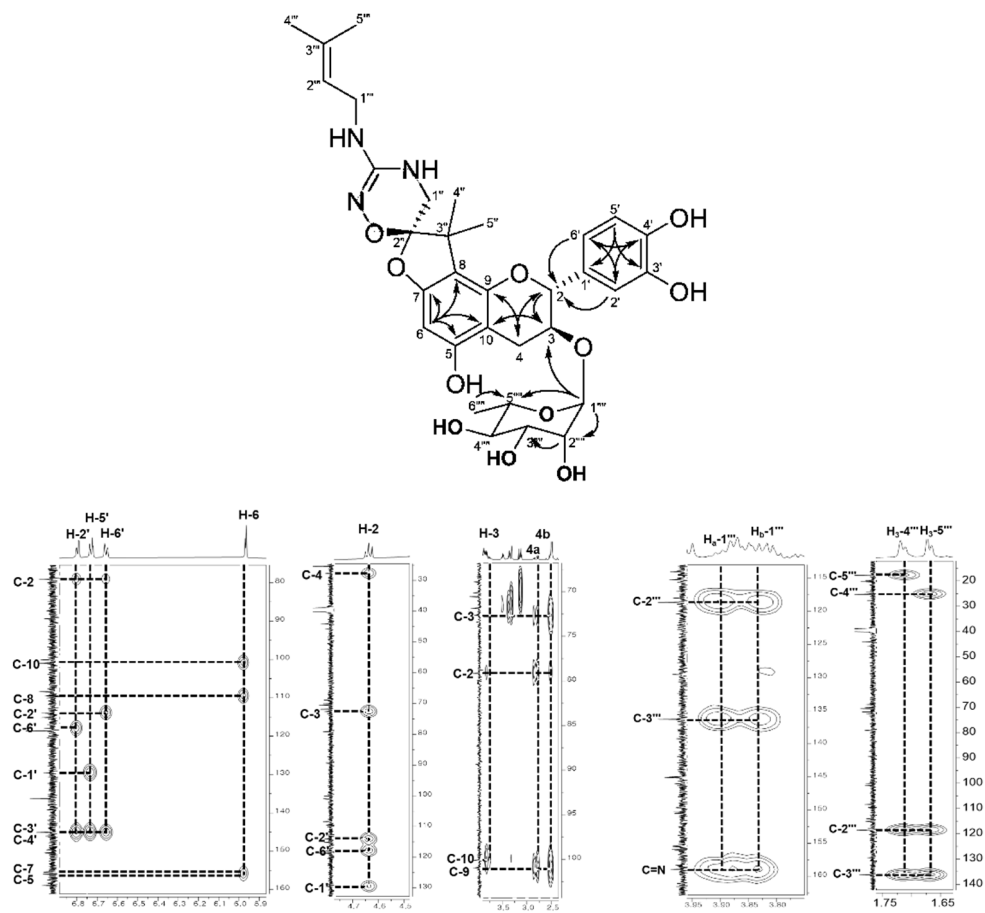
Figure 12. HMBC spectrum of compound **1**
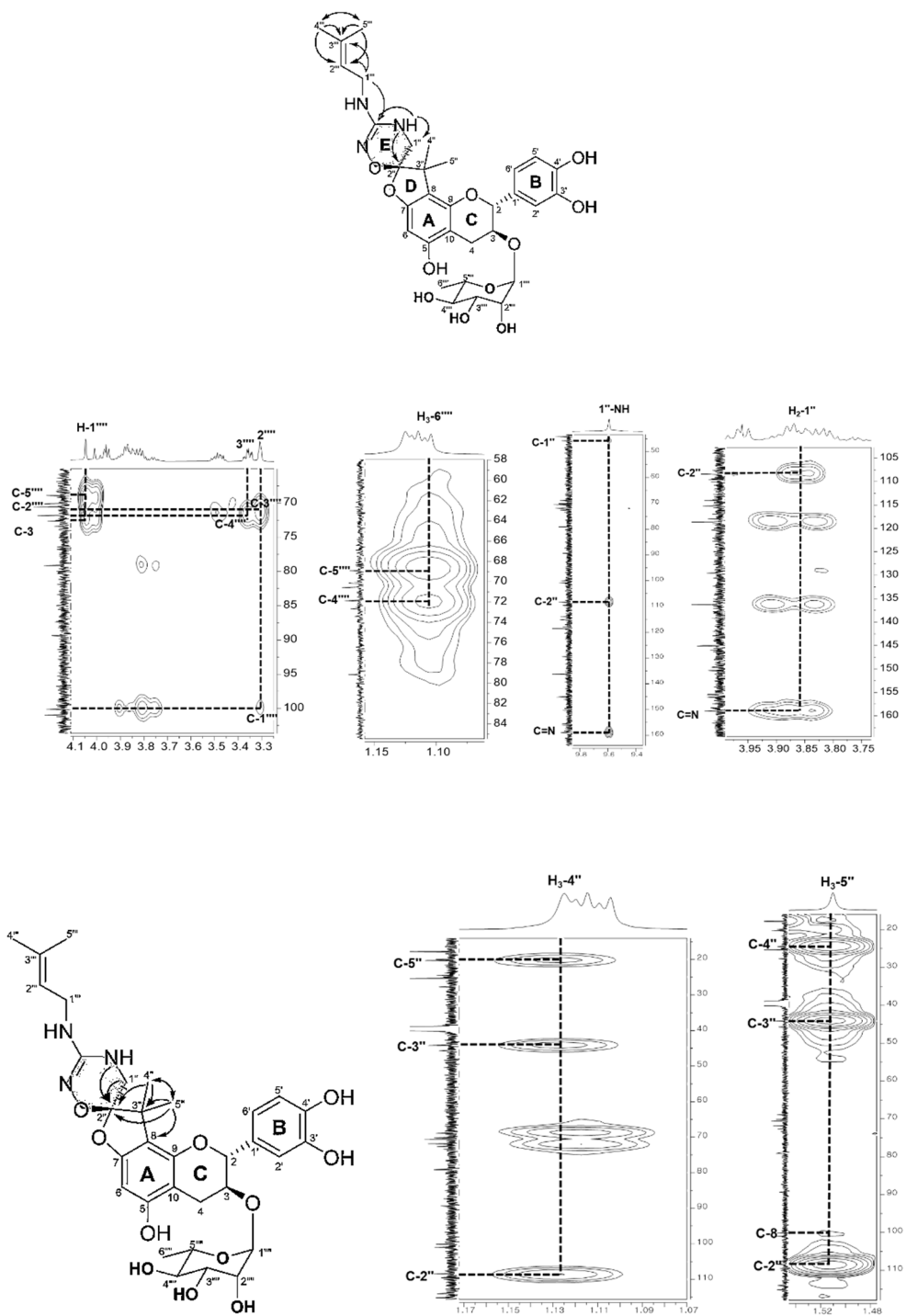
Figure 10. HMBC spectrum of compound **1** - continued
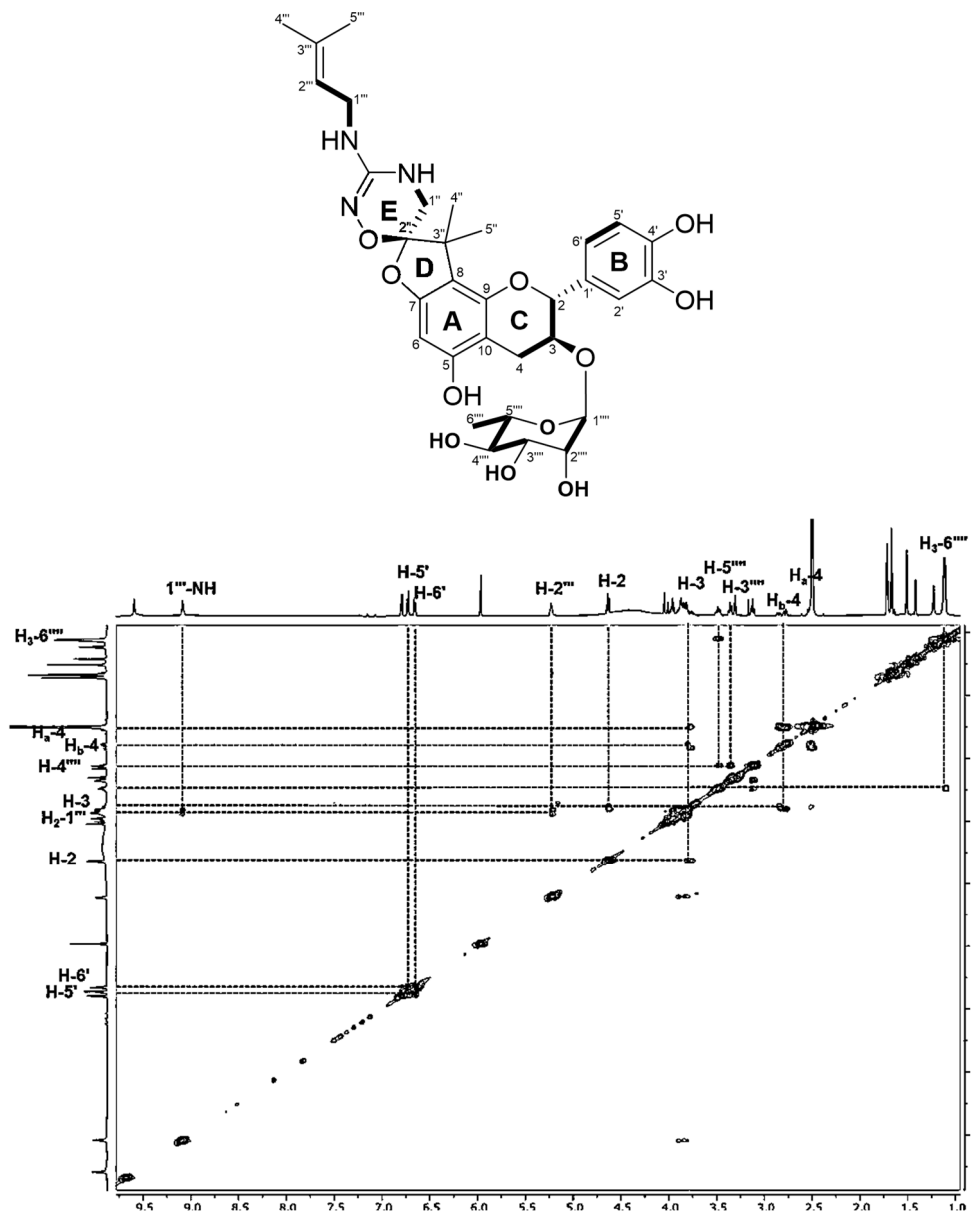
Figure 13. ^1H–^1H COSY spectrum of compound **1**

3.2. Compound **2**

The HRESIMS and NMR spectra of **2** (rugonidine B), which was isolated from a 1:1 mixture with its diastereomer **1** (Figure 14) by HPLC ($[\alpha]_D^{20}$ +30 and –17, respectively), resembled those of **1,** suggesting that it has the same molecular formula and planar structure as **1** (Tables 2 and 3). Therefore, **2** was assumed to be the 2″-epimer of **1**. Moreover, the ^{13}C NMR calculation for **2** displayed a 99.72 % DP4+ probability matching with the isomer **II** (Figure 20D), suggesting the 2*R*,3*S*,2″*S* relative configuration for **2**. The absolute configurations of the two diastereomers were determined based on electronic circular dichroism (ECD) calculations by time-dependent DFT (TDDFT). The ECD spectra were calculated for two isomers of **1** (**I** and **II**) using TDDFT at the B3LYP/def-TZVP level followed by the calculated Boltzmann weighted mean, and the CD curves were obtained from the Gaussian output using the program SpecDis ver. 1.71 (Bruhn et al., 2013a). As a result, the experimental CD data of **1** were fit to the Boltzmann-averaged computed ECD spectrum of 2*R*,3*S*,2″*R* with a negative CE at 248 nm, while **2** was confidently assigned as 2*R*,3*S*,2″*S* (Figure 24A). Interestingly, interday experiments on the CD spectra of **1** and **2** revealed that **2** slowly converted to **1** after 30 days, which was indicated by the (–) CE at approximately 240 nm for both diastereomers (Figure 24B). The opposite optical rotations, which were measured directly after repurifying $[\alpha]_D^{20}$ + 30 (**1**), and –17 (**2**), and their CD spectra obtained after one, two, and five days (Figure 25), indicated that these two diastereomers were stable for five days with (–) CE for **1** and (+) CE for **2** at approximately 240 nm. However, the CD spectrum of **2** in the 30th-day experiment exhibited a pattern similar to that of **1**, suggesting that these compounds occurred as configurationally semistable diastereomers and that the isomer of 2*R*,3*S*,2″*R* -**1** was more stable.
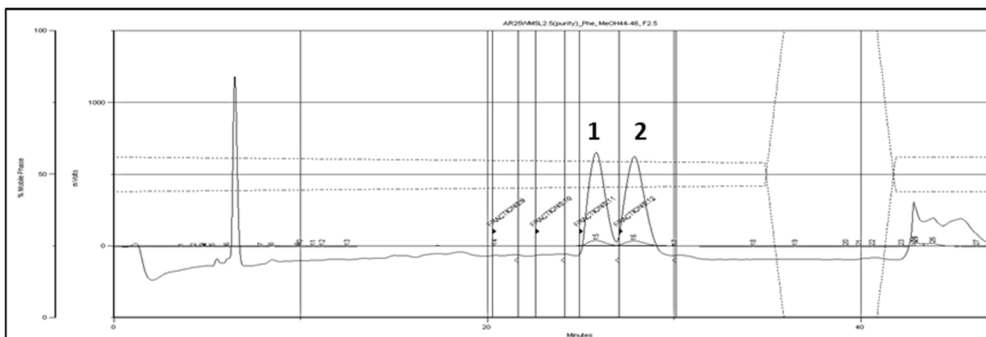
Figure 14. HPLC chromatography for isolation of compounds **1** and **2**. HPLC Gilson system, 5μm YMC Phenyl Hexyl column. HPLC isolation conditions were as follows: a gradient condition of 48–50 % MeOH/H$_2$O (0.1 % TFA) in 30 min followed by 5 min of 100 % MeOH and then 5 min of 48 % MeOH/ H$_2$O (0.1% TFA) for re-stabilizing.
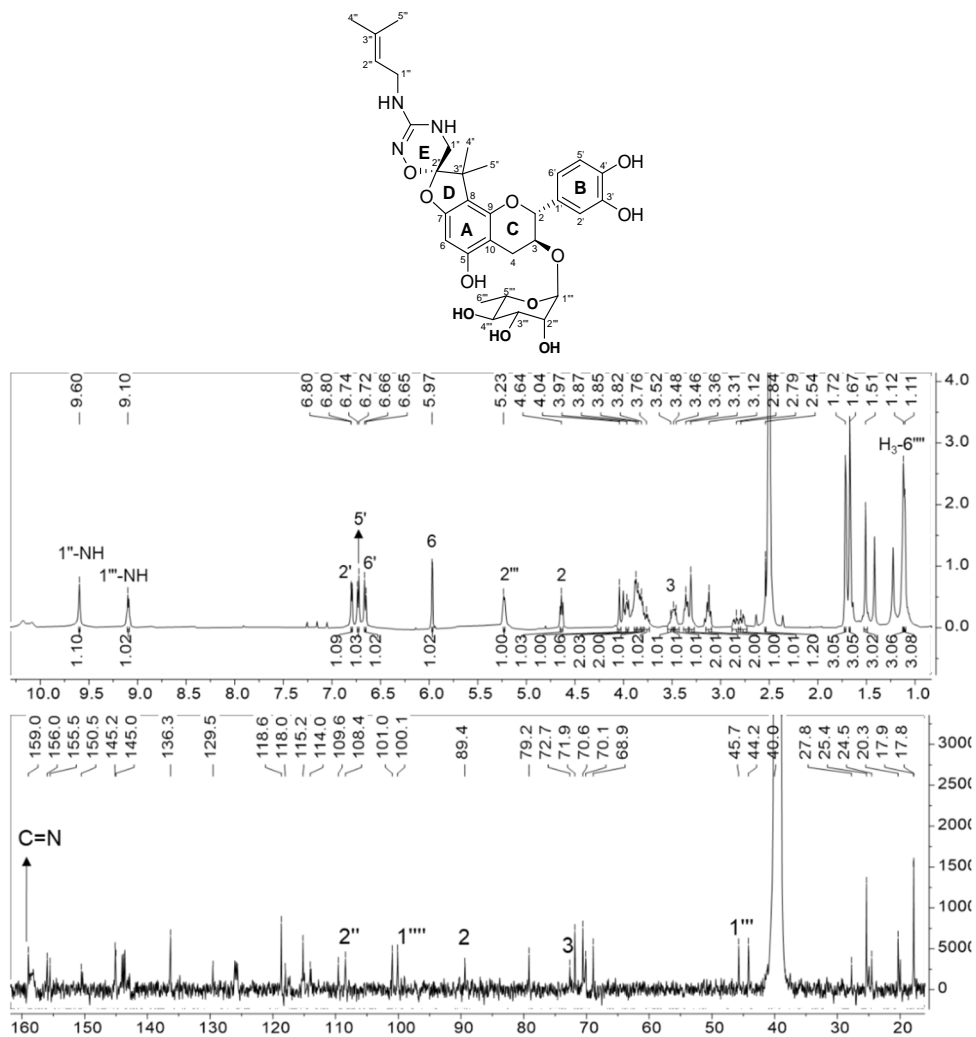
Figure 15. ^{1}H and ^{13}C NMR spectra of compound **2** (500/125 MHz, DMSO-d_6)

Figure 16. HSQC spectrum of compound **2**
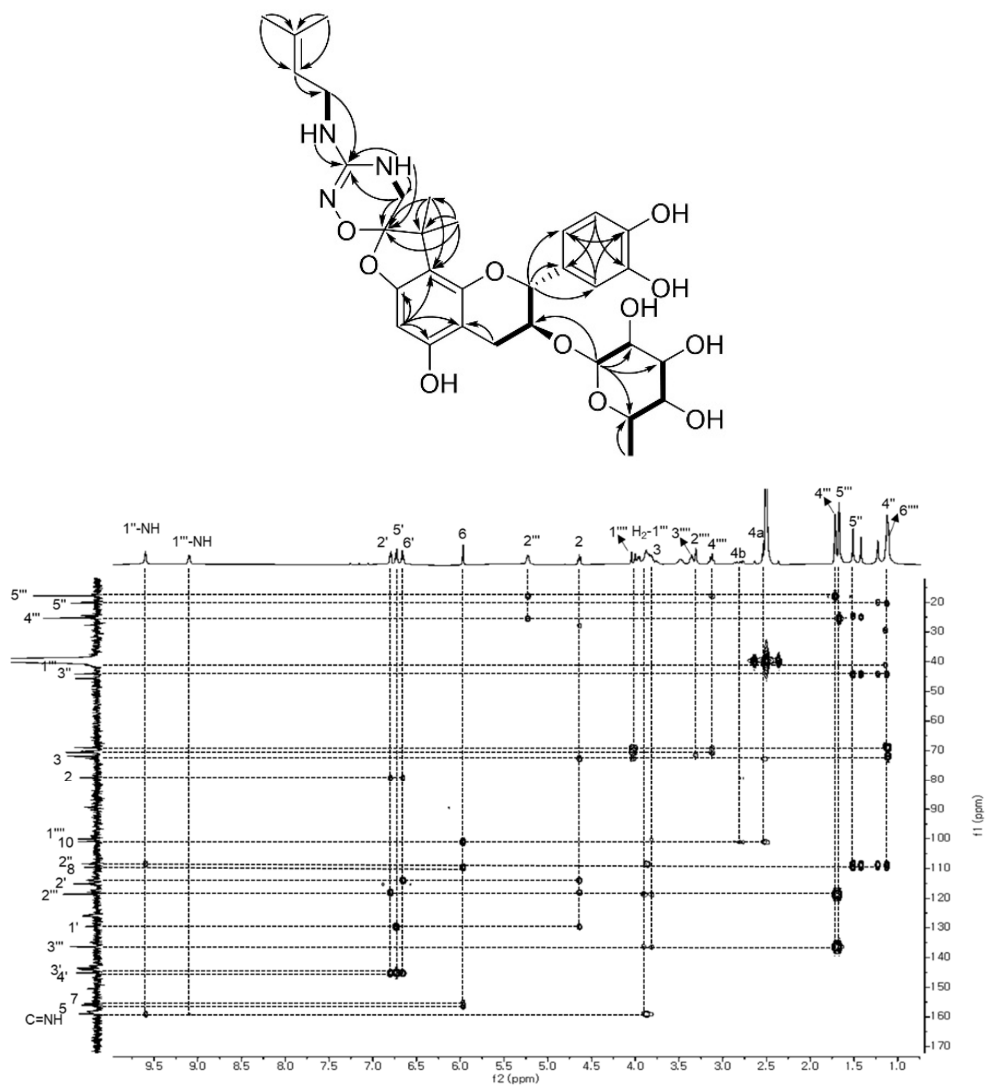
Figure 17. HMBC spectrum of compound **2**
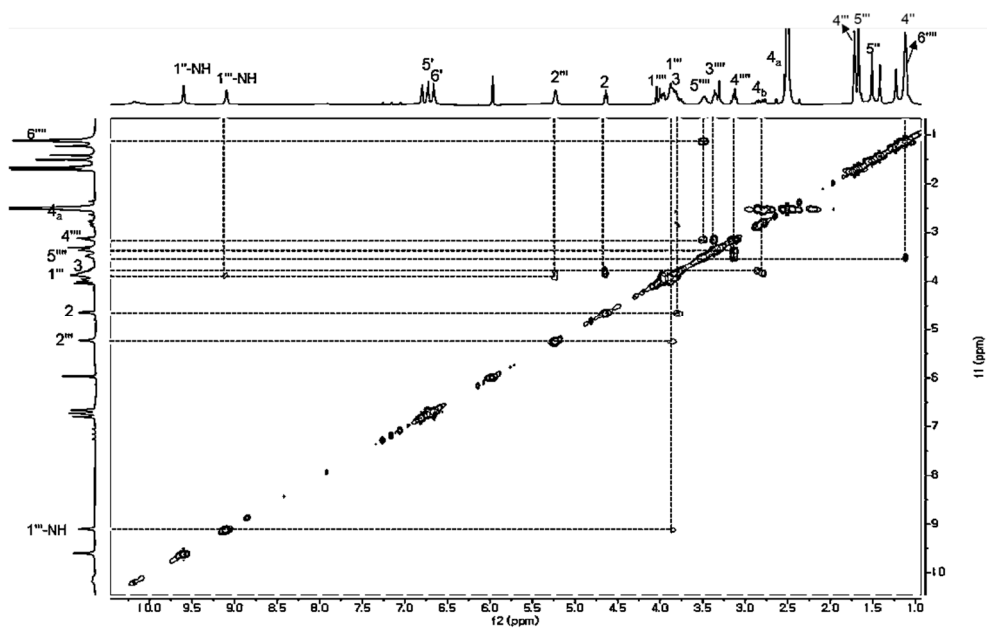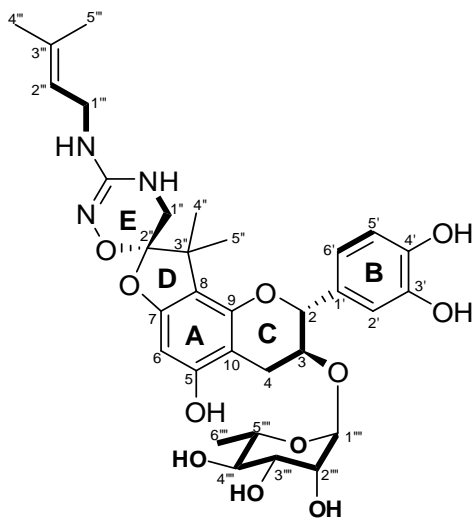
Figure 18. ^1H–^1H COSY spectrum of compound **2**
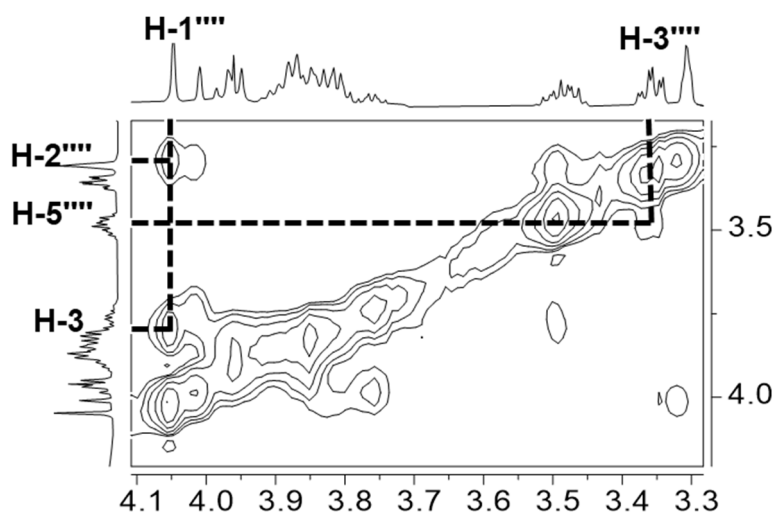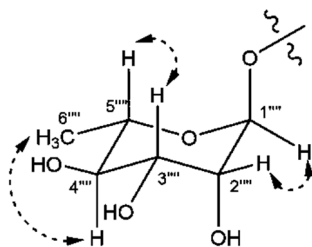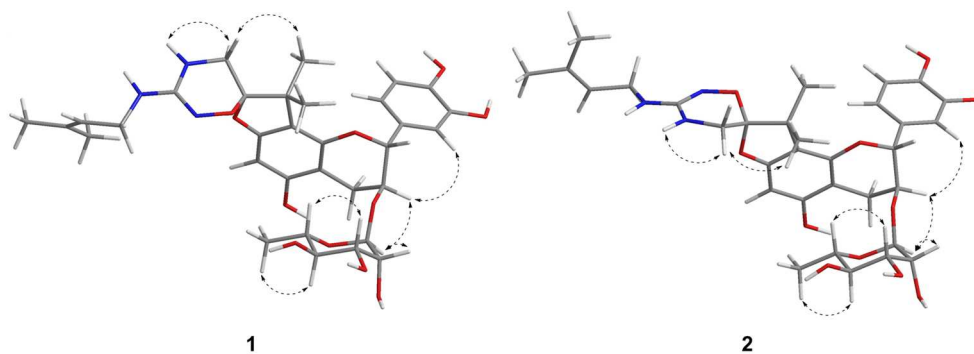
Figure 19. NOESY spectra of compounds **1** and **2**
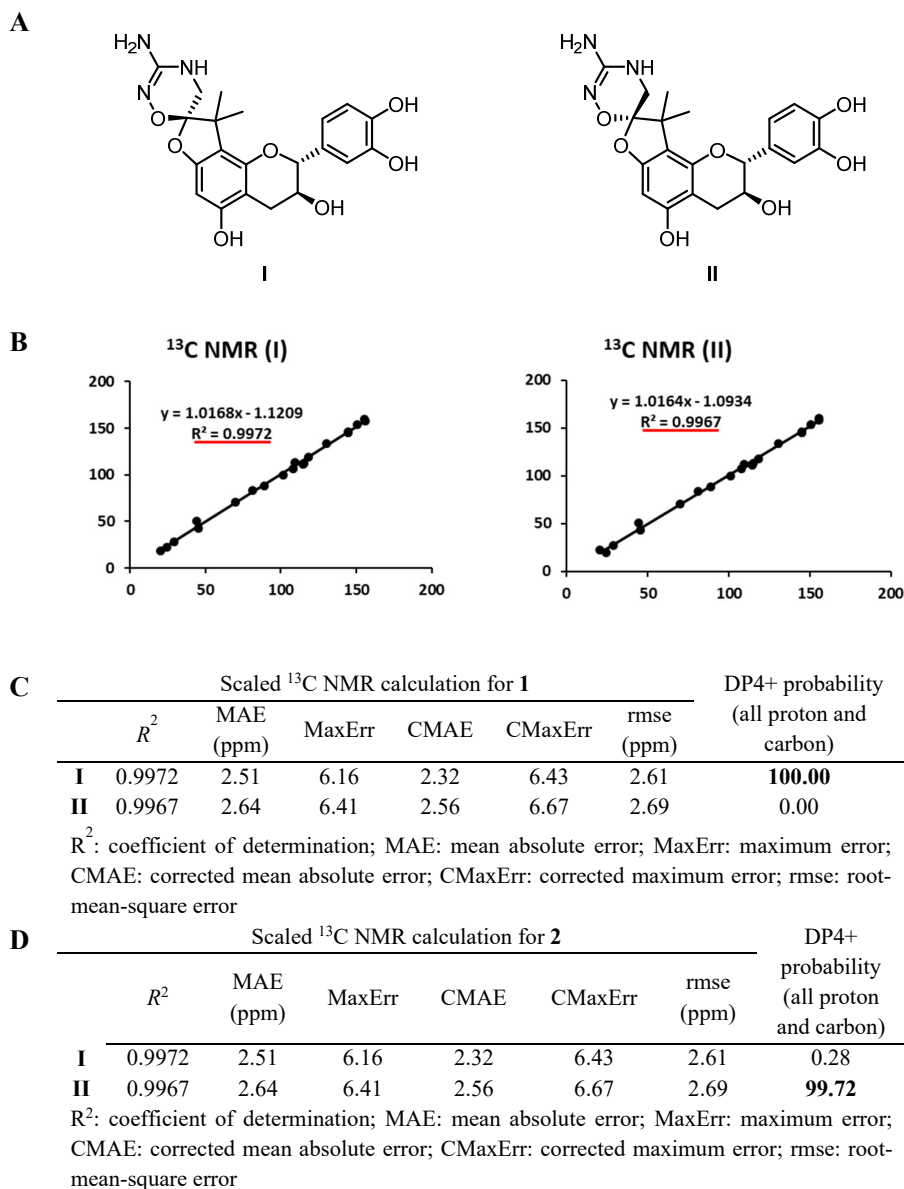
Figure 20. ^{13}C NMR calculation for two possible isomers of the core structure of **1** and **2**. (A) Linear correlation plots of predicted versus experimental ^{13}C NMR chemical shifts. (B) Statistical parameters of calculated ^{13}C NMR chemical shifts compared with experimental data of 1 and DP4+ probability for all proton and carbon calculations.
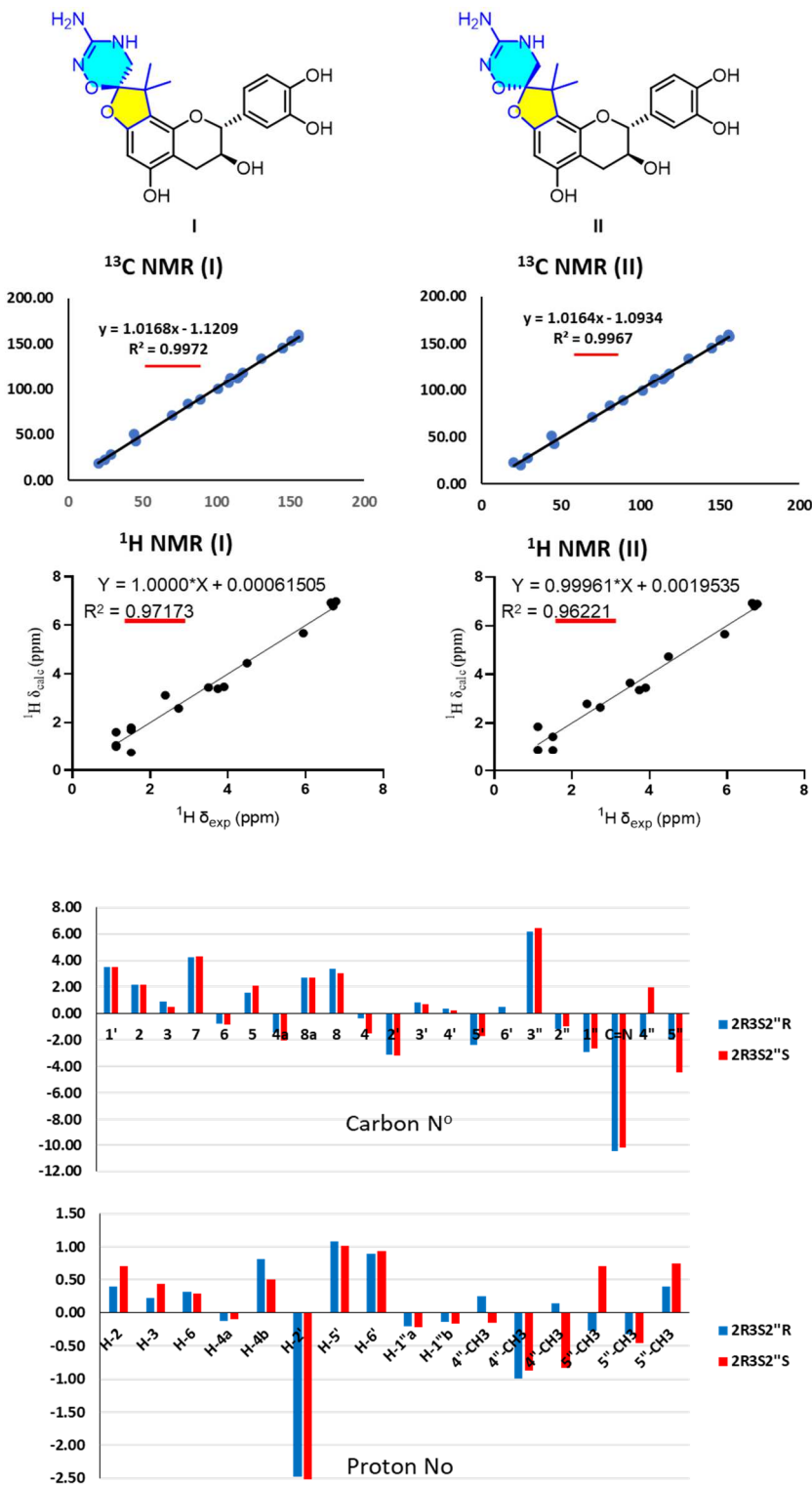
Figure 21. Linear correlation and relative errors between the calculated and the experimental ^{1}H and ^{13}C NMR chemical shifts for **I** and **II**.

Isomer Nº		1	2
DP4 (%)	H	100.00	0.00
	C	79.25	20.75
	H+C	*100.00*	*0.00*
	J	-	-
	all data	**100.00**	**0.00**

Settings	ype of data (shift		Default settings				Custom settings				Most Likely Isomers		
Custom	Shielding tensors		TMS	σ	v		TMS	σ	v		Rank	Isomer	Probability
		H	31.831	0.185	14.18	H	31.831	0.185	14.18		1*	1	100.00
		C	192.29	2.306	11.38	C	192.29	2.306	11.38		2*	2	0.00
		J	-	0.992	3.06	J	-	0.992	3.06		3*	-	-
		Slope scaling J		0.9509		Slope scaling J		0.9509			4*	-	-
		Intercept scaling J		-0.1405		Intercept scaling J		-0.1405					

	Isomer Nº	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	H	100.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	79.25	20.75	-	-	-	-	-	-	-	-	-	-	-	-	-
DP4 (%)	*H+C*	*100.00*	*0.00*	-	-	-	-	-	-	-	-	-	-	-	-	-
	J	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	all data	100.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-

Type	Exp	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
c	129.5	51.902	46.0584													
c	79.1	113.85	100.838													
c	72.6	129.21	114.843													
c	155.5	20.298	18.0208													
c	89.3	107.5	95.3236													
c	156.0	23.162	20.0717													
c	101.0	93.445	83.4421													
c	150.5	28.032	25.0084													
c	109.6	77.723	69.2772													
c	27.7	180.66	161.196													
c	114.0	79.447	70.4716													
c	145.0	37.409	33.4024													
c	145.1	37.963	33.9111													
c	115.2	77.833	68.2383													
c	118.0	70.455	63.1262													
c	44.2	154.08	136.166													
c	108.4	84.886	75.0279													
c	45.7	163.5	144.467													
c	159.0	32.165	28.341													
c	20.3	192.83	166.838													
c	24.5	188.15	169.247													
h	4.64	32.359	28.3165													
h	3.76	33.546	29.4393													
h	5.96	30.951	27.3874													
h	2.78	34.69	30.5963													
h	2.53	34.082	30.3679													
h	6.8	32.955	29.164													
h	6.73	29.392	26.0614													
h	6.66	29.642	26.1863													
h	3.95	33.573	29.6599													
h	3.86	33.664	29.7604													
h	1.51	35.547	31.7346													
h	1.51	36.828	32.3752													
h	1.51	35.657	32.3379													
h	1.13	36.478	31.3146													
h	1.13	36.53	32.3471													
h	1.13	35.793	31.2807													

Figure 22. DP4+ probability of compound **1** (carbon and proton NMR chemical shift calculation).

Isomer Nº		1	2
DP4 (%)	H	0.07	99.93
	C	80.87	19.13
	H+C	*0.28*	*99.72*
	J	-	-
	all data	0.28	99.72

Settings	ype of data (shift				Default settings				Custom settings				Most Likely Isomers		
Custom	Shielding tensors				TMS	σ	v		TMS	σ	v		Rank	Isomer	robability
				H	31.831	0.185	14.18	H	31.831	0.185	14.18		1*	2	99.72
				C	192.29	2.306	11.38	C	192.29	2.306	11.38		2*	1	0.28
				J	-	0.992	3.06	J	-	0.992	3.06		3*	–	–
				Slope scaling J		0.9509		Slope scaling J		0.9509			4*	–	–
				Intercept scaling J		-0.1405		Intercept scaling J		-0.1405					

	Isomer Nº	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	H	0.07	99.93	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	80.87	19.13	-	-	-	-	-	-	-	-	-	-	-	-	-
DP4 (%)	H+C	0.28	99.72	-	-	-	-	-	-	-	-	-	-	-	-	-
	J	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	all data	0.28	99.72	–	–	–	–	–	–	–	–	–	–	–	–	–

Type	Exp	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
c	129.5	51.902	46.0584													
c	79.2	113.85	100.838													
c	72.7	129.21	114.843													
c	155.5	20.298	18.0208													
c	89.4	107.5	95.3236													
c	155.8	23.162	20.0717													
c	101.0	93.445	83.4421													
c	150.5	28.032	25.0084													
c	109.6	77.723	69.2772													
c	27.8	180.66	161.196													
c	114.0	79.447	70.4716													
c	145.0	37.409	33.4024													
c	145.2	37.963	33.9111													
c	115.2	77.833	68.2383													
c	118.0	70.455	63.1262													
c	44.2	154.08	136.166													
c	108.4	84.886	75.0279													
c	45.7	163.5	144.467													
c	159.0	32.165	28.341													
c	20.3	192.83	166.838													
c	24.5	188.15	169.247													
h	4.64	32.359	28.3165													
h	3.82	33.546	29.4393													
h	5.97	30.951	27.3874													
h	2.82	34.69	30.5963													
h	2.54	34.082	30.3679													
h	6.8	32.955	29.164													
h	6.73	29.392	26.0614													
h	6.66	29.642	26.1863													
h	3.96	33.573	29.6599													
h	3.86	33.664	29.7604													
h	1.12	35.547	31.7346													
h	1.12	36.828	32.3752													
h	1.12	35.657	32.3379													
h	1.51	36.478	31.3146													
h	1.51	36.53	32.3471													
h	1.51	35.793	31.2807													

Figure 23. DP4+ probability of compound **2** (carbon and proton NMR chemical shift calculation).
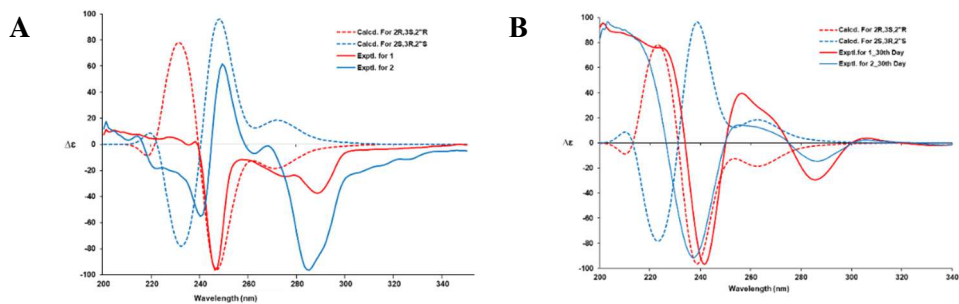
Figure 24. Experimental and calculated ECD spectra of **1** and **2**.

The calculated ECD curves were generated by SpecDis ver. 1.71 (Bruhn et al., 2013a). (**A**) Experimental data recorded directly after separating **1** and **2**. (**B**) Experimental data recorded 30 days after separation of **1** and **2**.
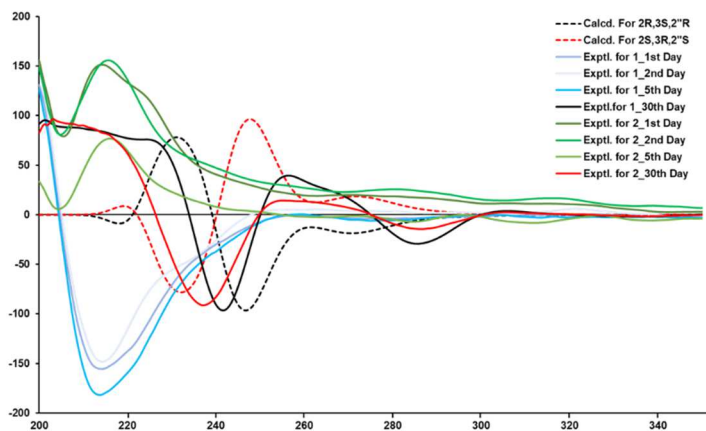
Figure 25. Inter-days experimental ECD spectra of compounds **1** and **2** by time intervals on 1st, 2nd, 5th, and 30th-day after re-separating from HPLC.

3.3. Compounds 3 and 4

Rugonidines C (**3**) and D (**4**) were successfully separated by HPLC at a 1:1 ratio ($[\alpha]_D^{20}$ +6 and −3, respectively, Figure 26). The molecular formula of both isomers, $C_{27}H_{33}N_3O_{11}$, was established by the HRESIMS ion peak at *m/z* 576.2201 ([M + H]$^+$, calcd 576.2193), indicating 13 DBEs, including six rings and seven double bonds. The loss of 68 Da in the HRESIMS data and the lack of one double bond along with an *N*-methylene and two methyl groups in the NMR data (Tables 2, 3) of **3** suggested the absence of an isoprenyl unit in **3** compared with **1**. The *α*-rhamnosyl moiety at C-3 of **3** was also determined by analyzing its COSY and HMBC spectra. Similar to **1**, the sugar unit of **3** was assigned as *α*-L-rhamnose based on its $^1J_{CH}$ value (171.0 Hz) (Pretsch et al., 2009) and the consistent retention time in HPLC compared to that of the authentic L-rhamnose derivative after hydrolysis of **3**. The absolute configuration of **3** was identified as 2*R*,3*S*,2″*R*-rugonidine C by comparative ECD and optical rotation (OR) analysis. Compounds **3** and **4** were also found as a pair of 2″-configurational diastereomers. According to the ECD and OR comparison, compound **4** was determined to be 2*R*,3*S*,2″*S*-rugonidine D.
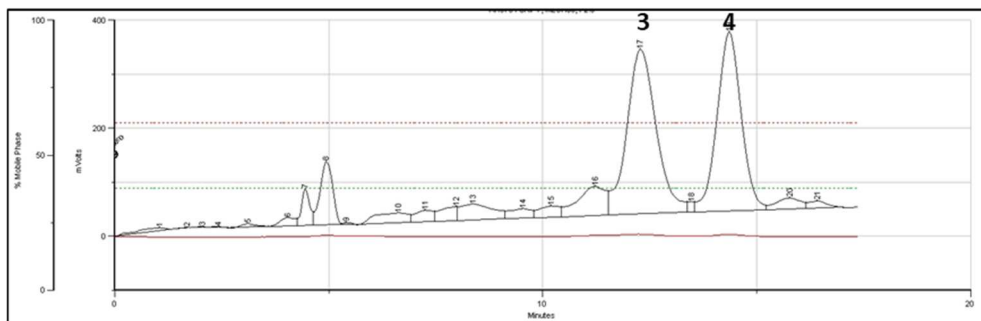
Figure 26. HPLC chromatography for isolation of compounds **3** and **4** by 5μm YMC Phenyl Hexyl column using conditions as follows: an isocratic of 38% MeOH/Water (0.1% TFA) in 30 min followed by 5 min of 100% MeOH and then 5 min of 38% MeOH/Water (0.1% TFA) for re-stabilizing
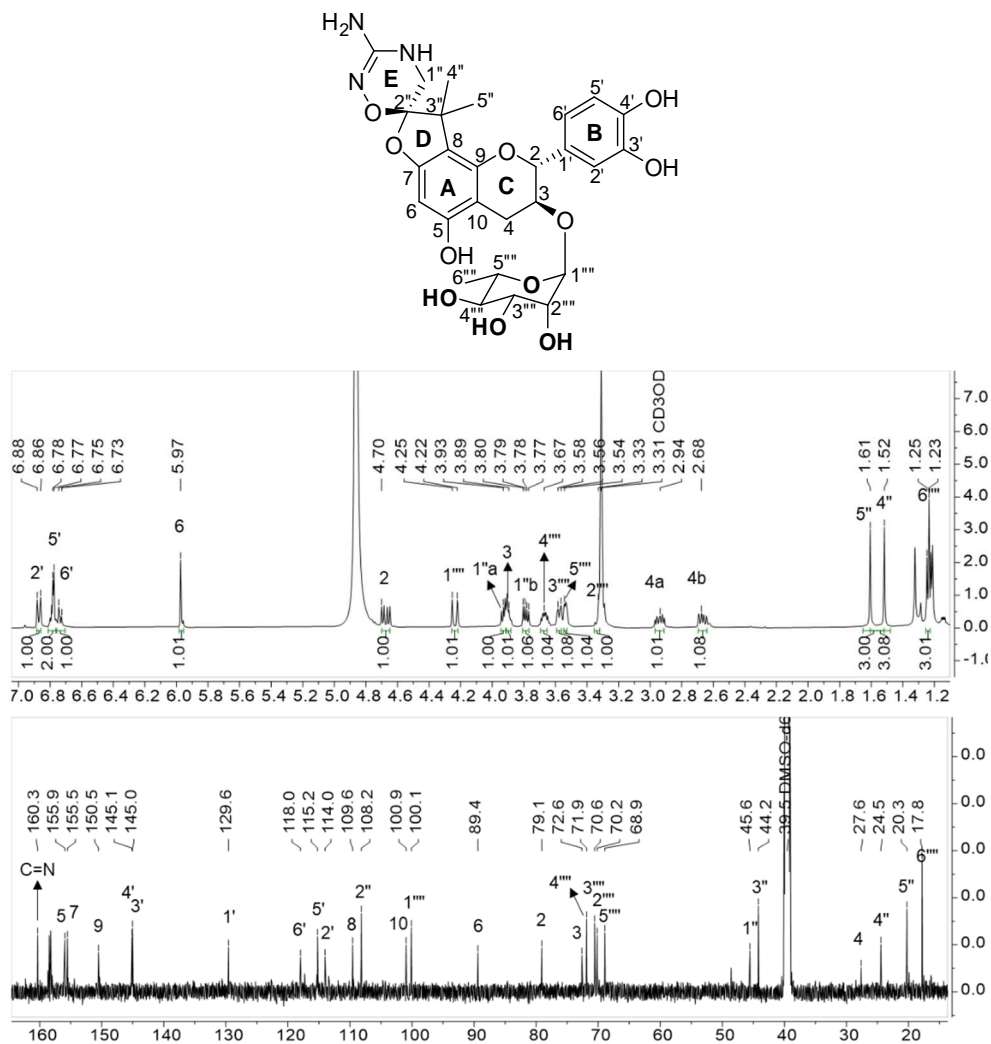
Figure 27. ^{1}H spectrum of **3** (500 MHz, methanol-d_4) after re-separating from **4** and ^{13}C NMR spectrum of compound **3** (150 MHz, DMSO-d_6)
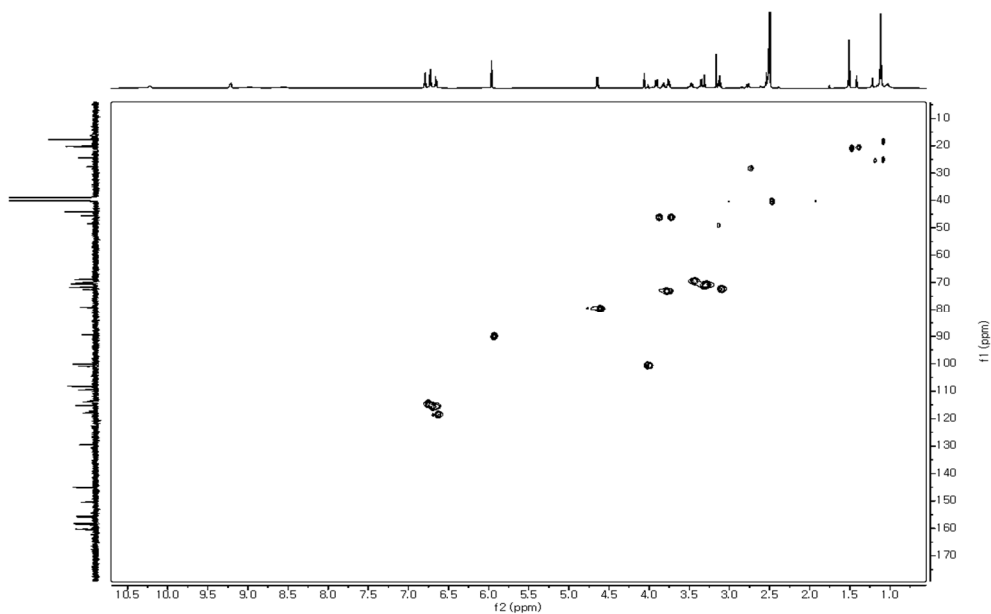
Figure 28. HSQC spectrum of compound **3** (600MHz, DMSO-d_6)
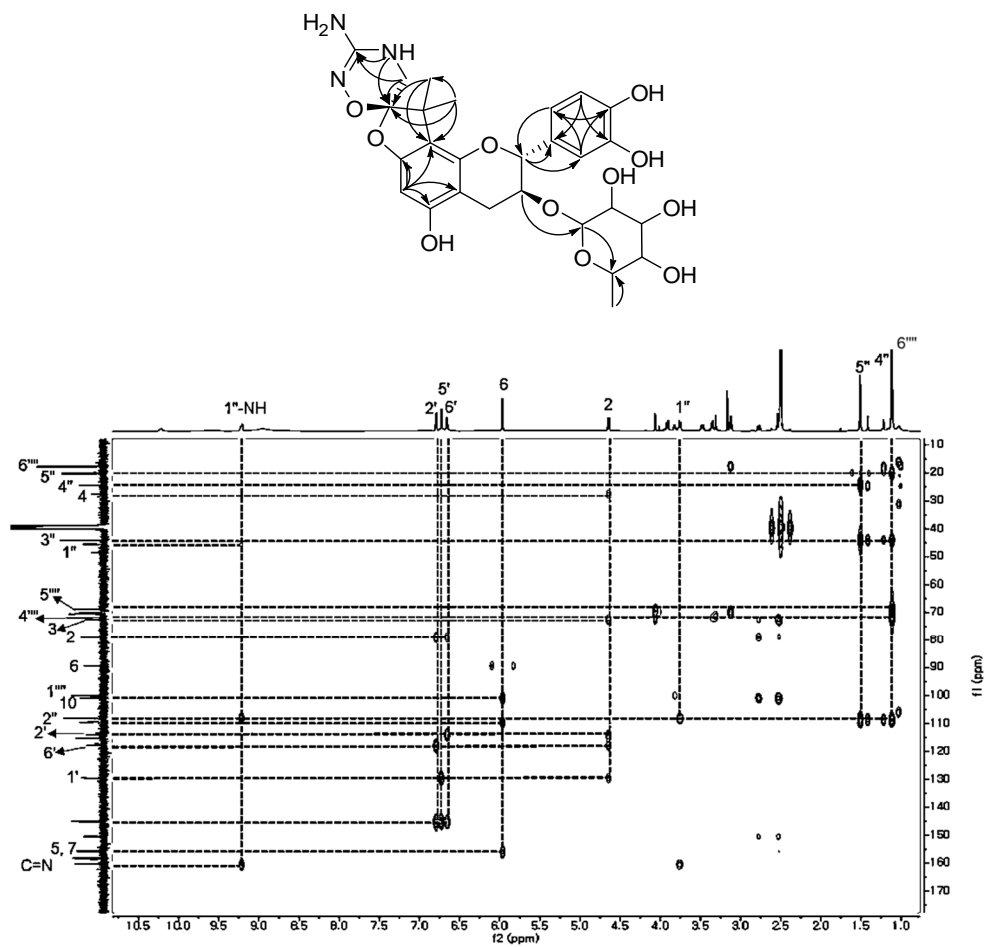
Figure 29. HMBC spectrum of compound **3** (600MHz, DMSO-d_6)
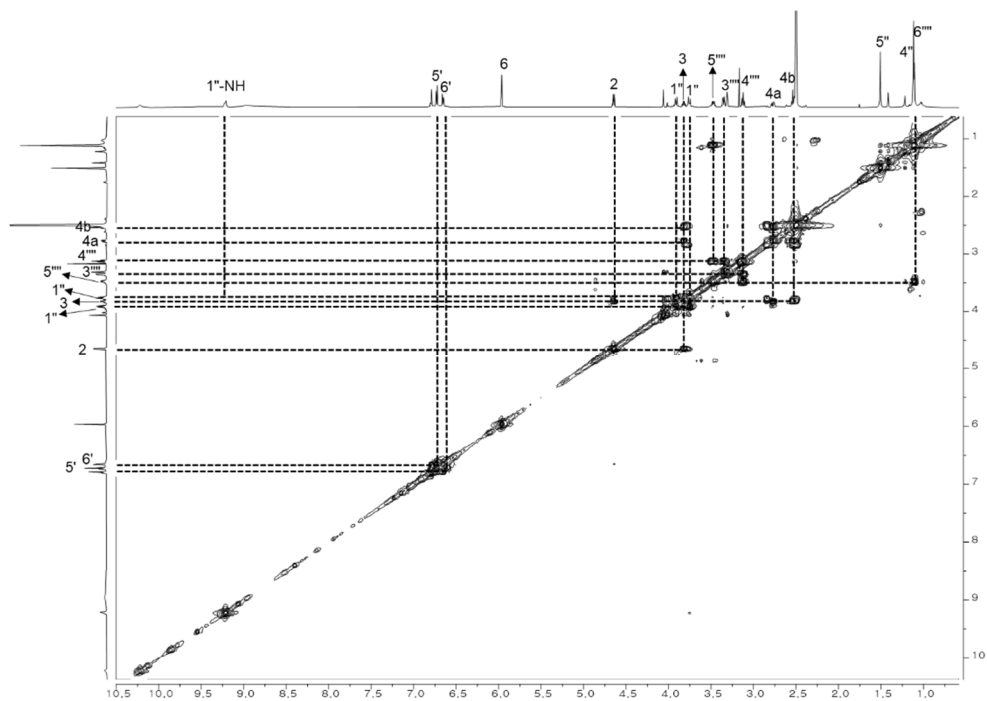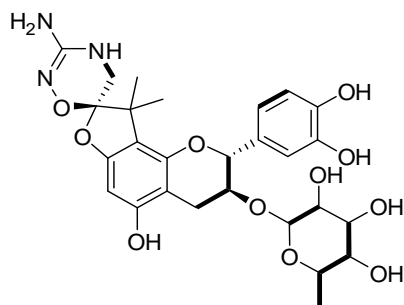
Figure 30. ¹H–¹H COSY spectrum of compound **3**
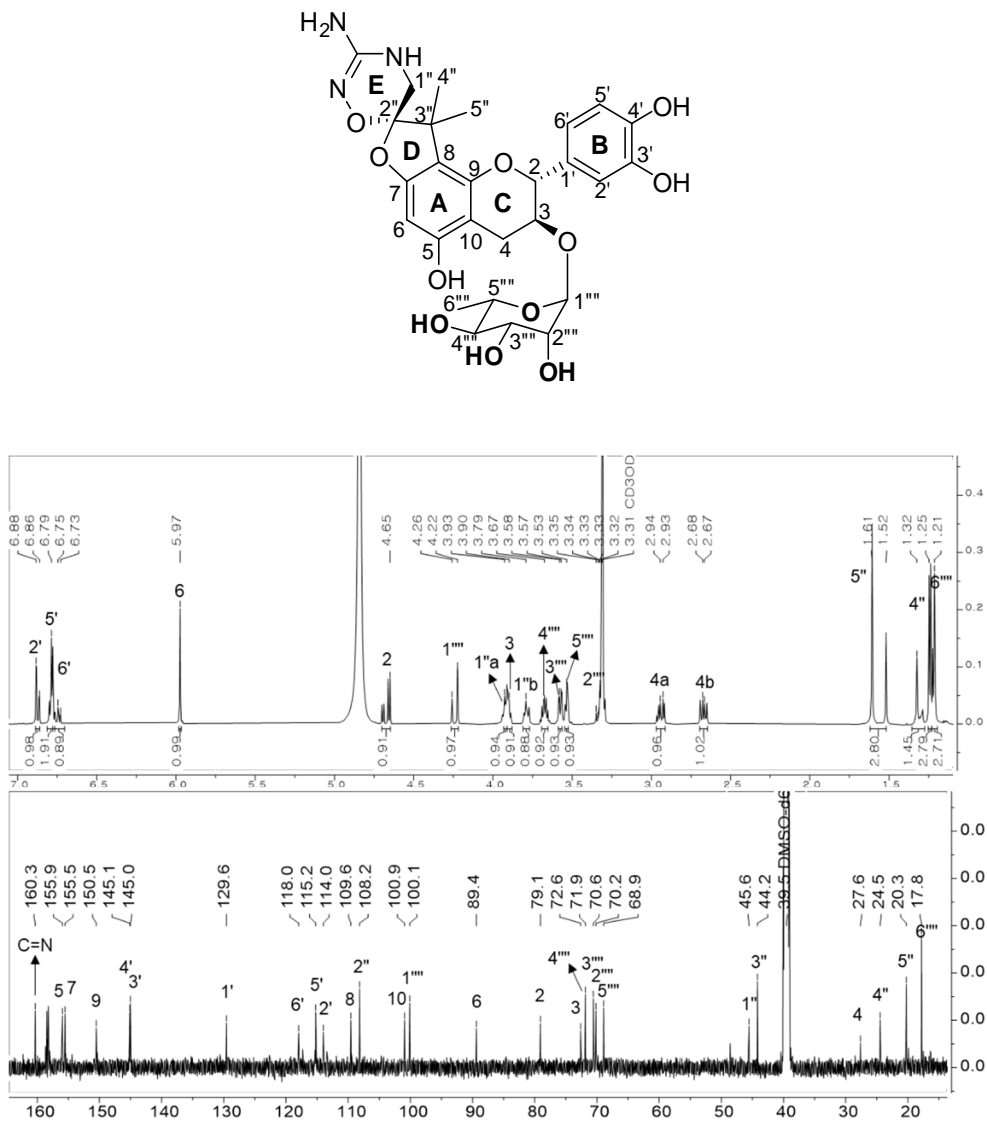
Figure 31. [1]H spectrum of **4** (500 MHz, methanol-d_4) after re-separating from **3** and ^{13}C NMR spectrum of compound **3** (150 MHz, DMSO-d_6)
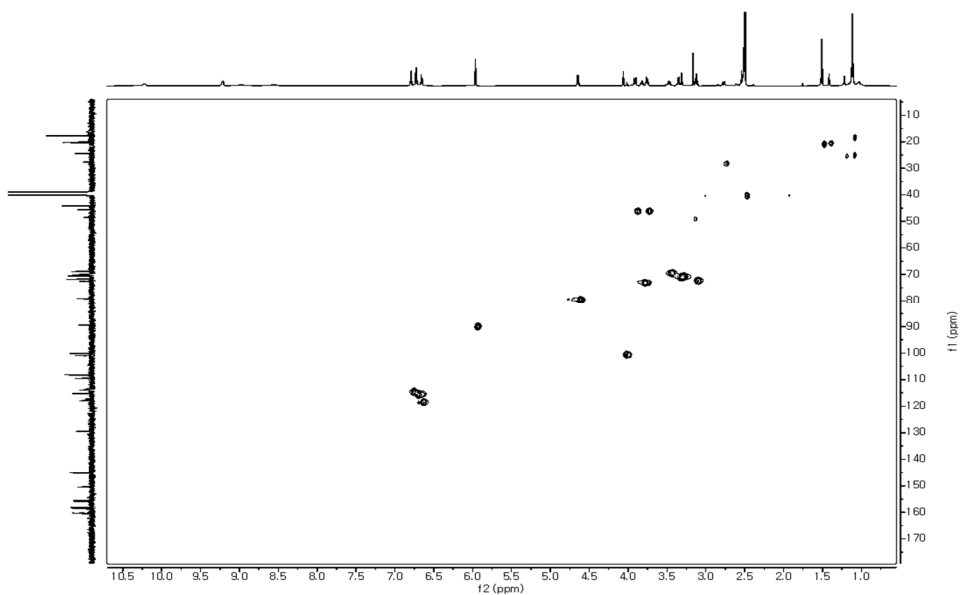
Figure 32. HSQC spectrum of compound **4**
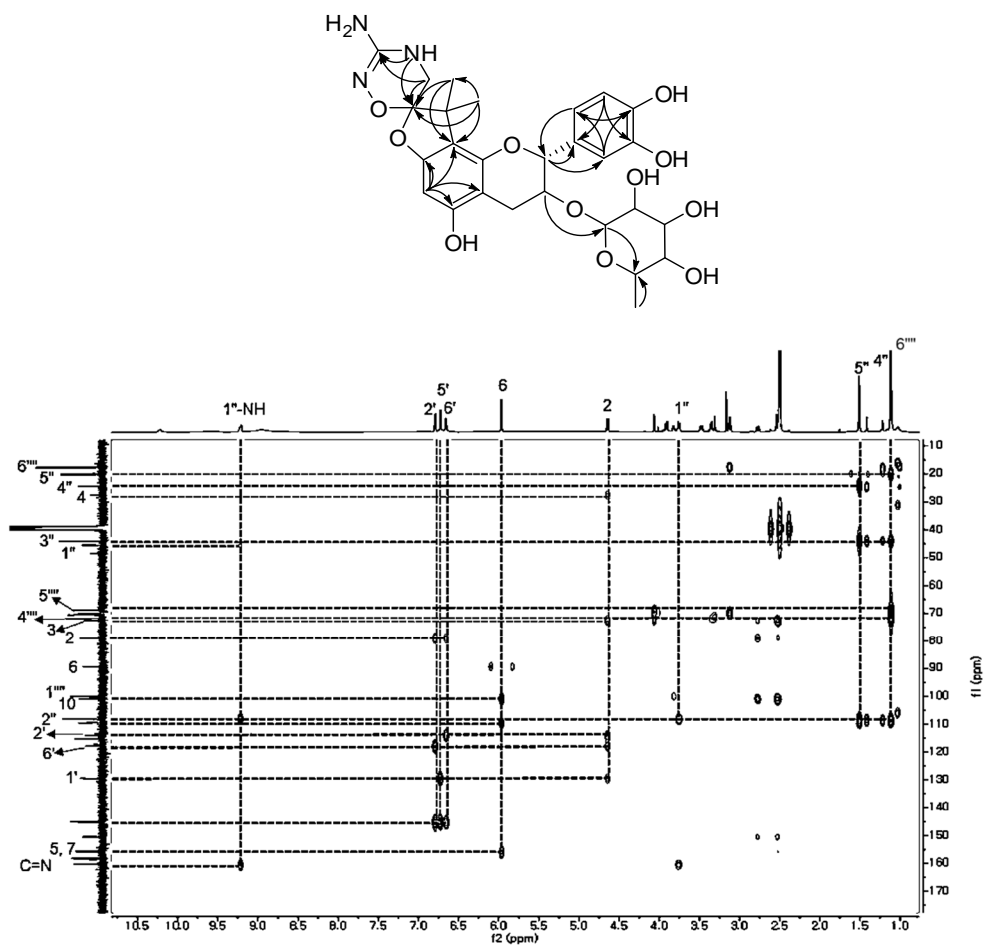
Figure 33. HMBC spectrum of compound **4**

35
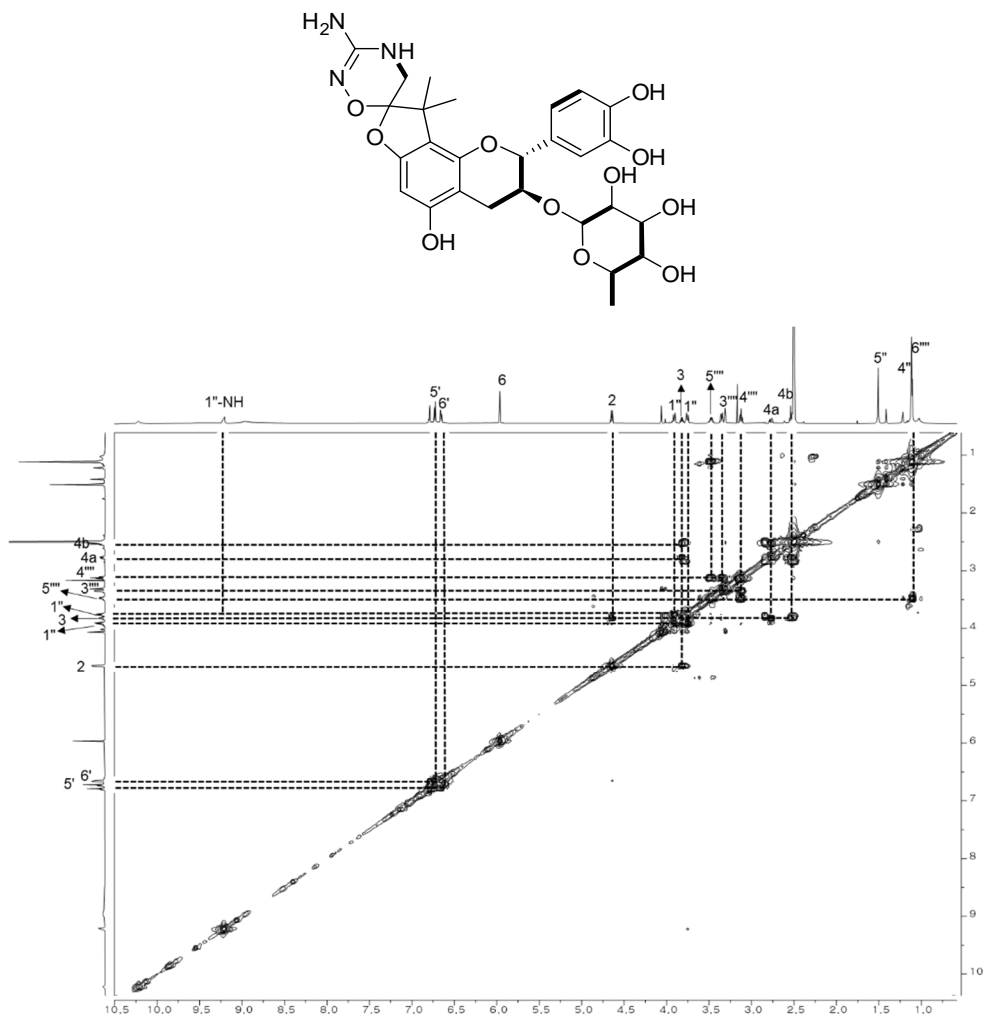
Figure 34. ^{1}H–^{1}H COSY spectrum of compound **4**

$^{3}J_{\text{H2''''-H3''''}}$ = 3.2 Hz (H2'''': equatorial)
$^{3}J_{\text{H3''''-H4''''}}$ = 9.2 Hz (H4'''': axial)
$^{3}J_{\text{H4''''-H5''''}}$ = 9.2 Hz (H5'''': axial)
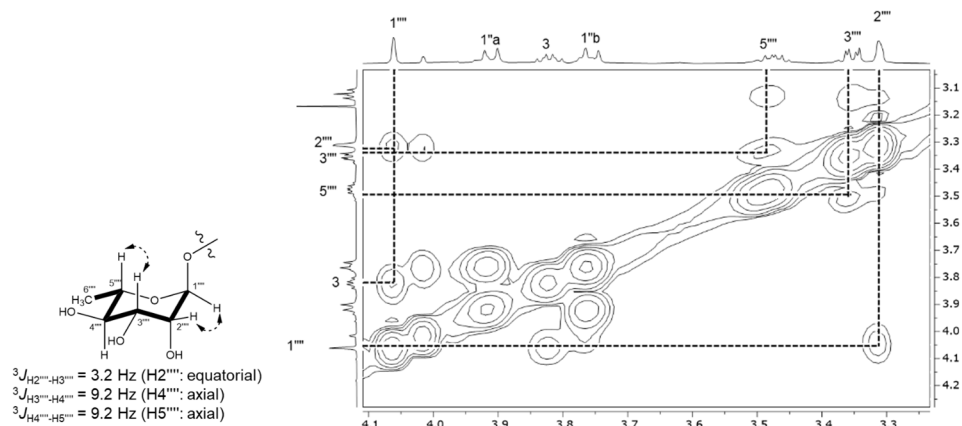
Figure 35. NOESY spectrum of compounds **3** and **4**
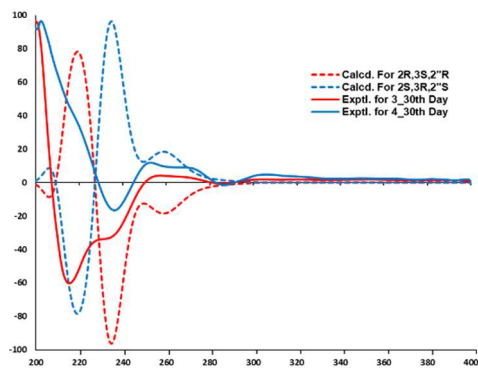
Figure 36. Experimental ECD spectra of compounds **3** and **4** after 30 days of re-separating by HPLC

3.4. Compounds **5** and **6**

Compound **5** was obtained as pale-yellow amorphous powder and was isolated from a 1:1 mixture with its diastereomer **6** by HPLC ($[\alpha]_D^{20}$ +2 and – 4, respectively) (Figure 37). Their chemical formula, $C_{21}H_{23}N_3O_7$, was deduced from HRESIMS ion peaks at m/z 430.1635 and 430.1609 $[M + H]^+$ (calcd 430.1614), implying 12 DBEs. The HRESIMS and NMR data of **5** and **6** suggested that they contained fewer rhamnose units than compounds **3** and **4** (Tables 2, 3). The planar structure of **5** was determined based on its NMR spectra, which shared a pattern similar to that of the core skeleton of **1**–**4**. The absolute configuration of **5** was assigned by the ECD and OR data compared to those of **1**. Therefore, **5** was determined to be 2*R*,3*S*,2″*R*-rugonidine E, while its diastereomer **6** was identified as 2*R*,3*S*,2″*S*-rugonidine F. The configurationally semistable phenomenon was also observed in pairs of **3** and **4** as well as **5** and **6** after thirty days (Figure 42).
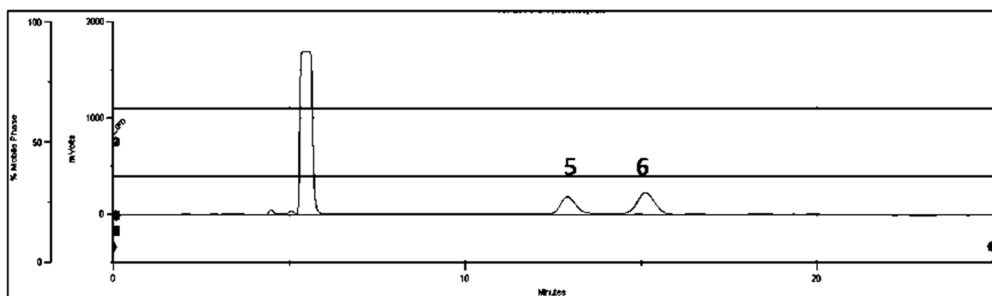
Figure 37. HPLC chromatography for isolation of compounds **5** and **6** by 5μm YMC Phenyl Hexyl column using conditions as follows: an isocratic of 36% MeOH/H_2O (0.1% TFA) in 30 min followed by 5 min of 100% MeOH and then 5 min of 36% MeOH/ H_2O (0.1% TFA) for re-stabilizing
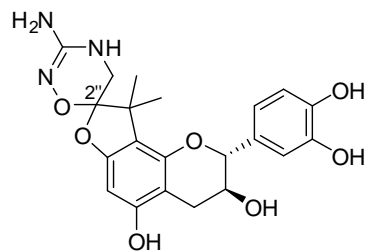
5. 2"*R*
6. 2"*S*
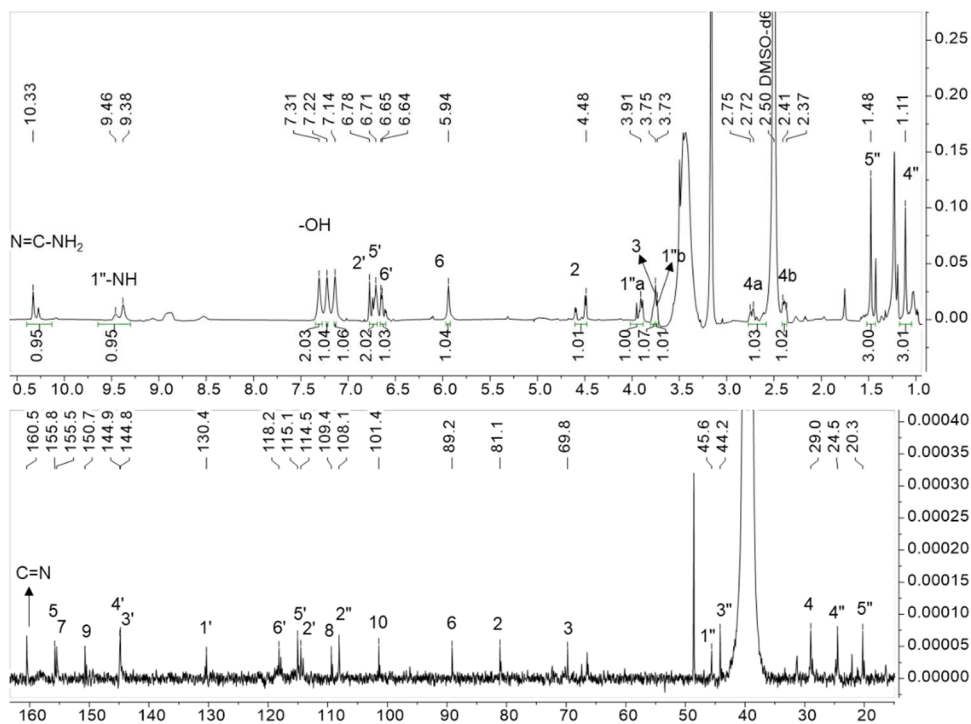
Figure 38. ^1H and ^{13}C NMR spectra of compounds **5, 6** (600/150 MHz, DMSO-d_6)
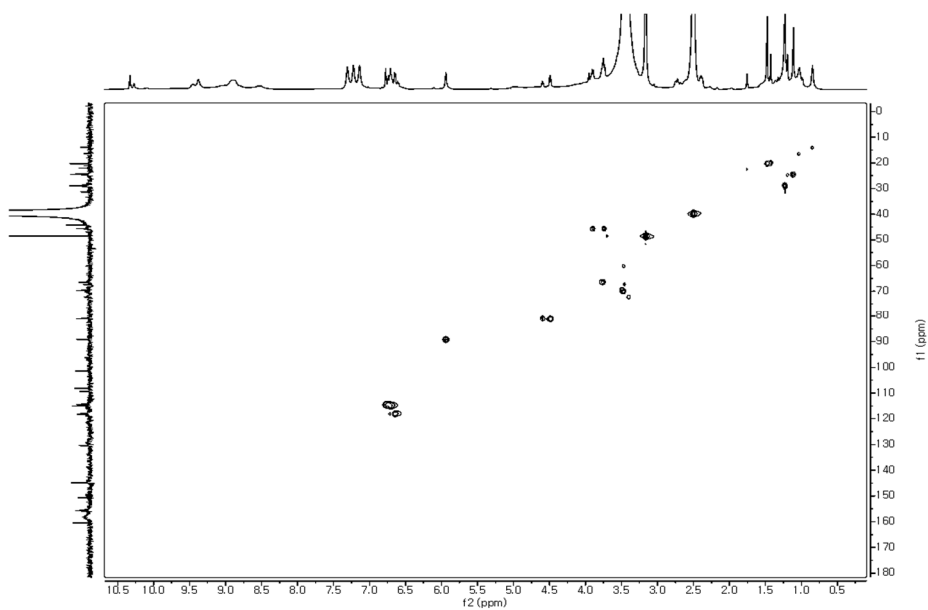
Figure 39. HSQC spectrum of compounds **5, 6**
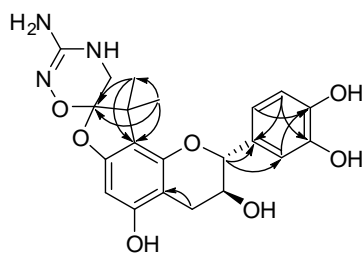
5. 2″*R*
6. 2″*S*
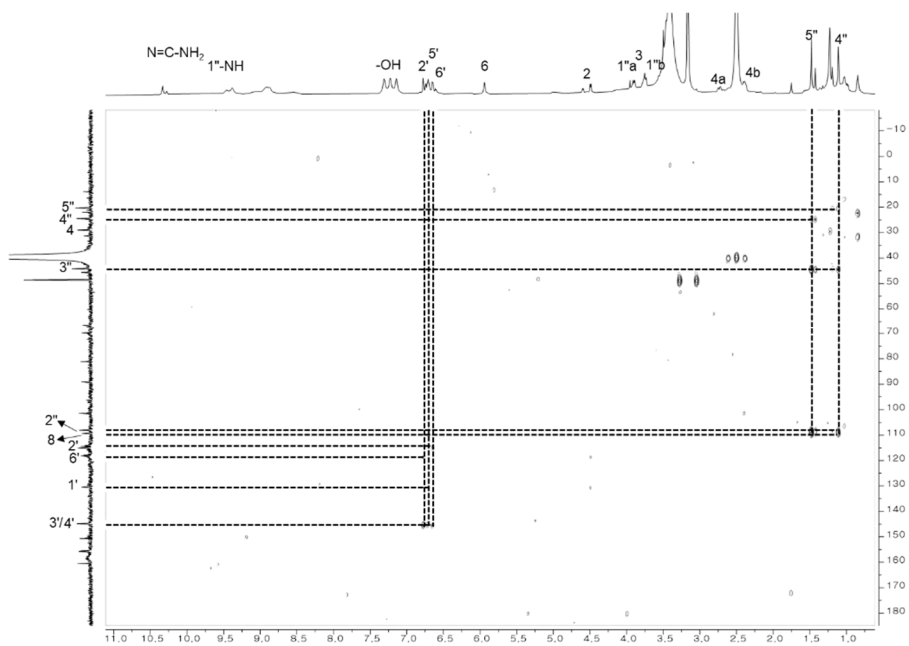
Figure 40. HMBC spectrum of compounds **5, 6**

5. 2"*R*
6. 2"*S*

Figure 41. ¹H–¹H COSY spectrum of compounds **5, 6**

Figure 42. Experimental ECD spectra of compounds **5** and **6** after 30 days of re-
separating by HPLC

4. Structure elucidation of guanidine-conjugated flavonoids from *A. rugosa*

4.1. Compound **7**

Compound **7** was isolated as an amorphous powder with $[\alpha]_D^{25}$ +54.9 (*c* 0.2, MeOH). The molecular formula, $C_{32}H_{41}N_3O_{10}$, was deduced from its HRESIMS ion peak at *m/z* 628.2859 $[M + H]^+$ (calcd for $C_{32}H_{42}N_3O_{10}$, 628.2865). The IR spectrum of **7** showed absorption bands characterized by hydroxyl or amine (3704 cm^{-1}), C-H in heteroaromatic rings (2967 cm^{-1}), C=NH or aromatic C (1682 cm^{-1}), benzofuran (1203 cm^{-1}), and C-O (1032 cm^{-1}). The ^{1}H NMR spectrum of **7** showed two N-H protons (δ_H 11.93, 11.65 ppm, Figure 43), four aromatic signals (δ_H 6.76, 6.73, 6.66, and 6.10), one olefinic proton at δ_H 5.26 (t, *J* = 6.9 Hz), an anomeric signal (δ_H 4.35 ppm), one *N*-methylene group at δ_H 3.81 (d, *J* = 6.9 Hz), six protons on oxygenated carbon (δ_H 3.33–4.71 ppm), one methylene group (δ_H 2.86/2.75 ppm), one methine group (δ_H 2.78 ppm), and five methyl groups (δ_H 1.73, 1.69, 1.25, 1.13, and 1.11 ppm). The ^{13}C NMR spectrum of **7** showed 32 carbon signals, including a guanidine carbon (δ_C 147.6), fourteen aromatic signals (δ_C 95.9–158.8), two olefinic carbons (δ_C 120.1/138.5), one anomeric carbon at δ_C 102.0 ppm, six oxygenated carbons (δ_C 70.3–81.2), an *N*-methylene group (δ_C 41.9), a methylene group (δ_C 27.3), one methine (δ_C 26.2), and five methyl group signals (δ_C 17.9–25.7). The HMBC correlations from H-2 (δ_H 4.71) to C-1' (δ_C 131.6), C-3 (δ_C 75.4), C-4 (δ_C 27.3), and C-9 (δ_C 155.5); from H-6 (δ_H 6.10) to C-5 (δ_C 158.8), C-7 (δ_C 157.3), C-10 (δ_C 100.8), and C-8 (δ_C 95.9), and from H$_2$-4 (δ_H 2.86/2.75) to C-5, C-9, and C-10 suggested the presence of a C6-C3-C6 unit. A hexose sugar moiety was revealed by the mass loss of 146 Da in HRESIMS/MS data as well as signals of an anomeric (δ_H 4.35/δ_C 102.0), four oxygenated methine groups (δ_H 3.33–3.65; δ_C 70.3–73.9), and a doublet methyl group (δ_H 1.25 (d, *J* = 6.3 Hz)/δ_C 17.9). This rhamnose unit was proven by the ^{1}H-^{1}H COSY spin system (Figure 46). The HMBC cross peak from H-1'''' (δ_H 4.35) to C-3 indicated that the rhamnose was linked to the catechin moiety of **7** at C-3. The small coupling constant of H-1'''' (d, *J* = 1.5 Hz), as well as the large $^1J_{CH}$ (169.6 Hz), indicated that the relative configuration of the sugar moiety was an *α*-oriented. In addition, the NMR data of **7** exhibited a guanidine unit characterized by the carbon signal at δ_C 147.6 with two N-H signals at δ_H 11.93 and 11.65 that shared the similarities to those of guanidine derivatives reported in the *Alchornea* genus (Barrosa et al., 2014a; Doan et al., 2021a; Tapondjou et al., 2016). An isoprenyl unit

in **7** elongated from the guanidine group was indicated by the presence of the *N*-methylene group (δ_H 3.81/δ_C 41.9), a double bond (δ_H 5.26/δ_C 120.1, 138.5), and two methyl groups (δ_H 1.69/δ_C 18.0; δ_H 1.73/δ_C 25.7), and the HMBC cross-peaks from H$_2$-1''' to guanidine carbon (δ_C 147.6). Moreover, the HMBC correlations from two doublet methyls (H$_3$-4'', δ_H 1.11, and H$_3$-5'', δ_H 1.13) to methine C-3'' (δ_C 26.2) and the other olefinic carbon (δ_C 131.3, C-2''), as well as from 1''-NH (δ_H 11.93) and C=NH (δ_H 11.65) to the double bond C-1''/C-2'', suggested that the other five-carbon chain was linked to the guanidine moiety (Figure 45). As suggested by HRESIMS, the molecular formula of **7** was C$_{32}$H$_{41}$N$_3$O$_{10}$, which consisted of 14 double bond equivalents (DBE); however, only 13 out of 14 DBEs had been assigned. Therefore, an additional ring of **7** through C-8/C-1'' and C-1''/7-OH was suggested because of the consistency with the conjugation reported in alchornealaxine (Tapondjou et al., 2016).

The relative orientations of the rhamnose moiety were determined by the coupling constants; in particular, $^2J_{\text{H-H}}$ of H-1'''' and H-2'''' indicated the equatorial orientation of H-2'''' while the large coupling constants of H-3'''', H-4'''', and H-5'''' suggested axial orientations of those protons in the sugar unit. Moreover, the NOESY correlations between H-1''''/H-2'''', H-3''''/H-5'''', and H-4''''/H$_3$-6'''' demonstrated the relative configuration of rhamnose sugar in **7** (Figure 47). The absolute configuration of the rhamnose was established by acid hydrolysis, followed by conversion to the corresponding thiocarbamoyl-thiazolidine carboxylate derivative with L-cysteine methyl ester and *o*-tolyl isothiocyanate (Tanaka et al., 2007). The consistent HPLC retention times for the sugar in **7** and the authentic L-rhamnose derivative confirmed that the stereochemistry of the sugar moiety was that of α-L-rhamnose. The absolute configurations at C-2, C-3 of **7** were determined to be 2*R*,3*S* based on the large coupling constants of H-2 (d, *J* = 7.3 Hz)/H-3 (q, *J* = 7.3 Hz), which suggested a 2,3-*trans* flavan-3-ol, and its CD data with negative CEs of approximately 290 and 240 nm (Figure 41) (Slade et al., 2005). Hence, the structure of compound **7** was identified as (2*R*,3*S*)-rugonine A.
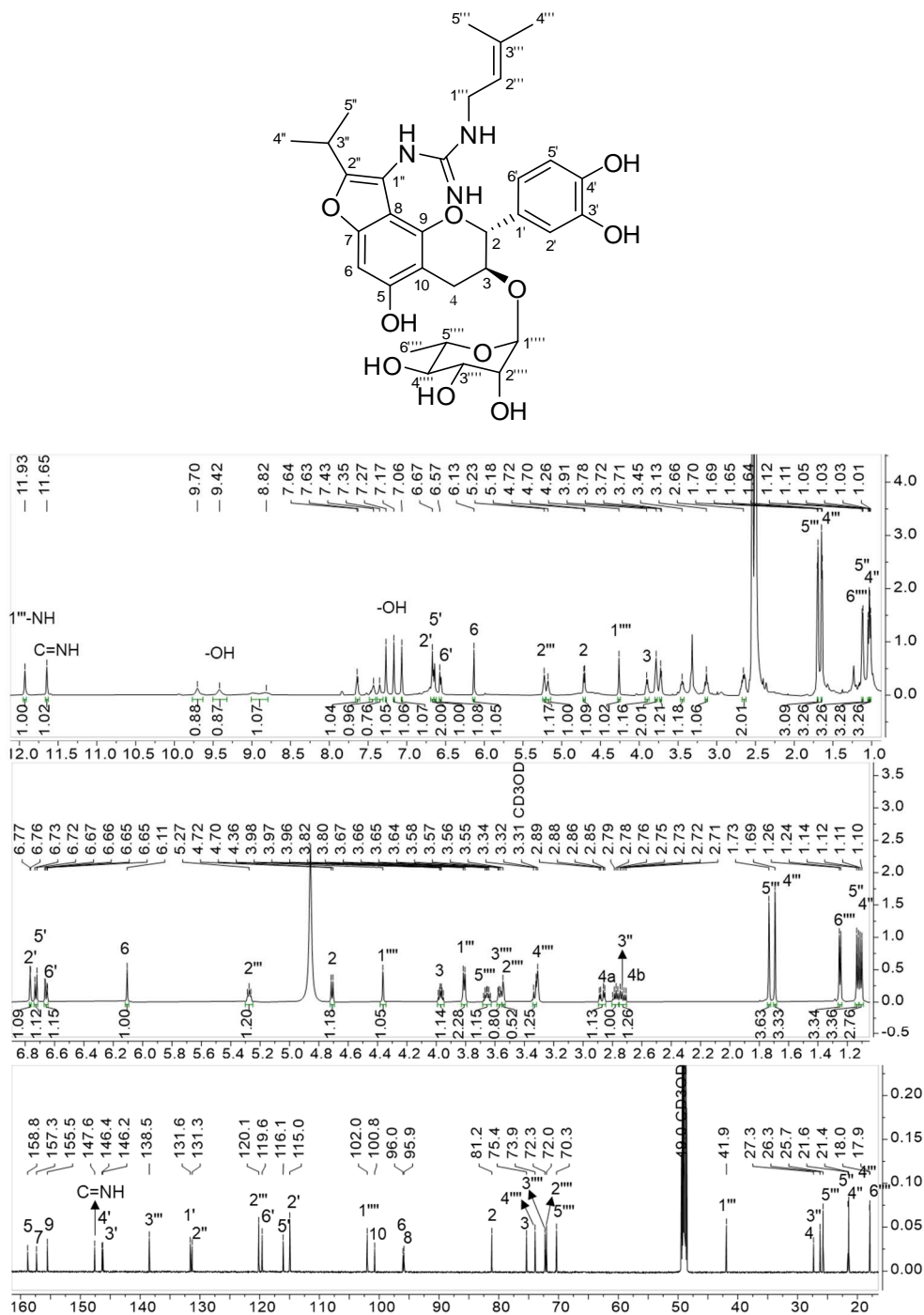
Figure 43. ^{1}H NMR (600 MHz, DMSO-d_6 and methanol-d_4) and ^{13}C NMR (150 MHz, methanol-d_4) spectra of compound **7**
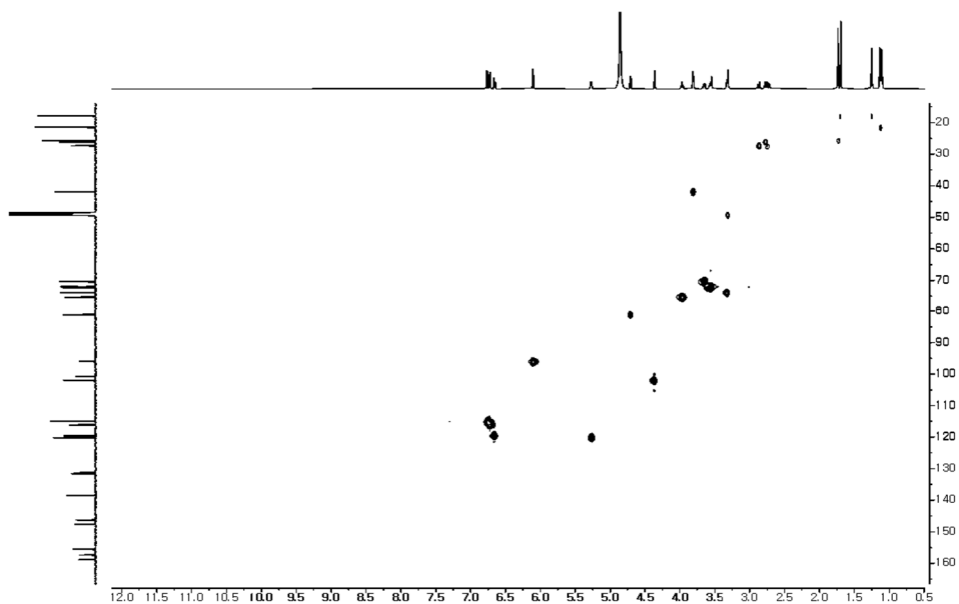
Figure 44. HSQC spectrum of compound **7**
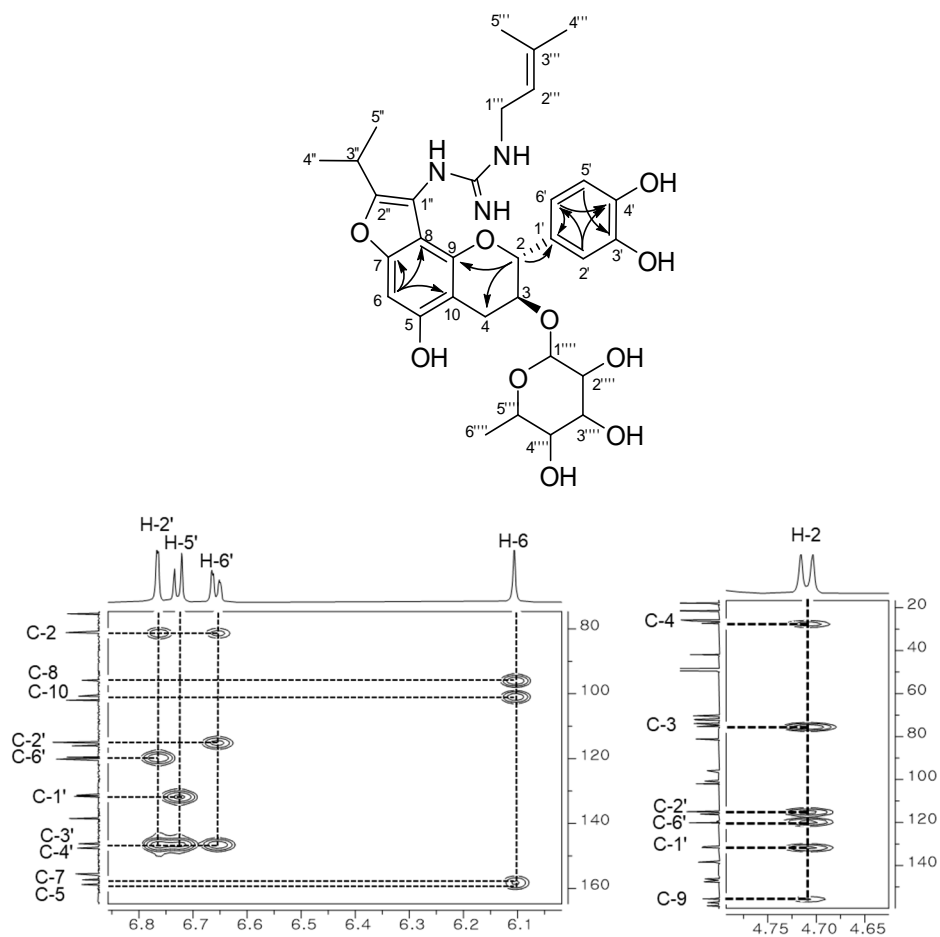
Figure 45. HMBC spectrum of compound **7**
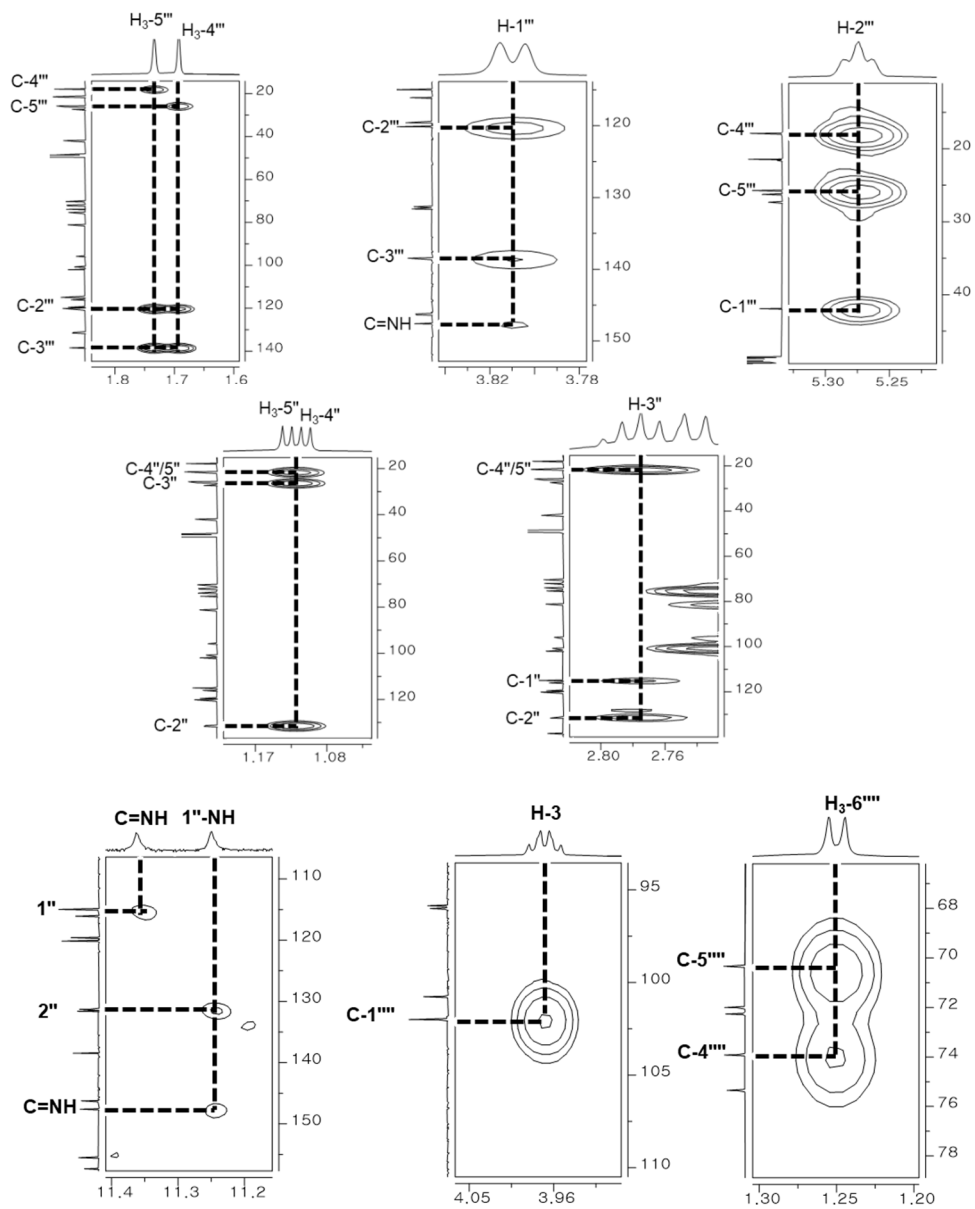
Figure 45 (continued). HMBC spectrum of compound **7**
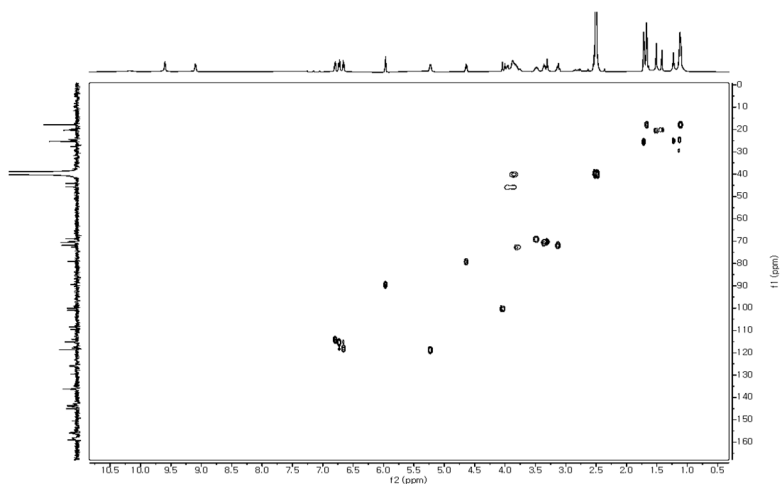
Figure 46. ^1H–^1H COSY spectrum of compound **7**
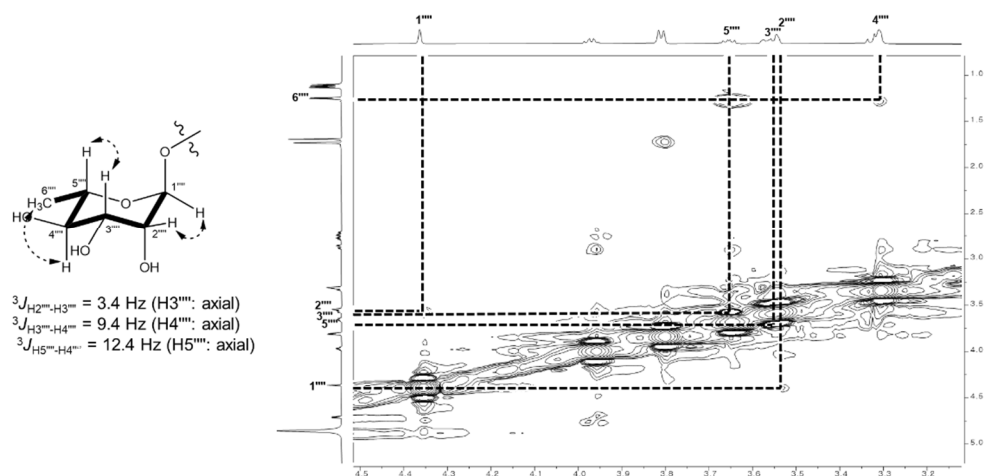
Figure 47. Key NOESY correlations of compound **7**
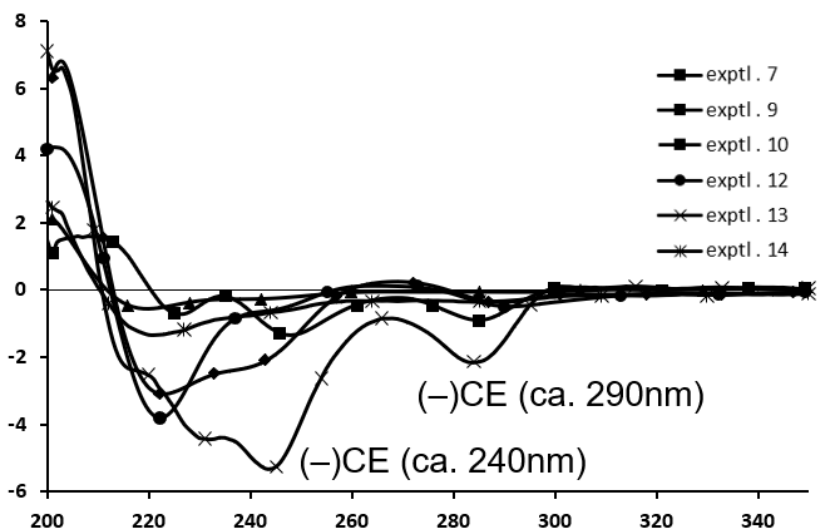
Figure 48. Experimental CD spectra of compounds **7, 9–10, 12–14**

4.2. Compound **8**

Compound **8** was isolated as an amorphous powder with $[\alpha]_D^{25}$ −78.0 (*c* 0.1, MeOH). The molecular formula, $C_{32}H_{41}N_3O_{10}$, was deduced from its HRESIMS ion peak at *m/z* 628.2870 [M + H]$^+$ (calcd for $C_{32}H_{42}N_3O_{10}$, 628.2865). The IR spectrum of **8** exhibited the absorption bands of hydroxyl or amine (3252 cm^{-1}), C-H in heteroaromatic rings (2976 cm^{-1}), C=NH or aromatic C (1668 cm^{-1}), benzofuran (1200 cm^{-1}), and C-O (1072 cm^{-1}). The NMR data of **8** were similar to those of **7**, suggesting a planar structure similar to that of **7**. The sugar moiety of **8** was also determined by NMR coupling constants and NOESY correlations analyses, and by comparing the HPLC retention time with the authentic L-rhamnose derivative, suggesting that the stereochemistry of the sugar moiety was that of *α*-L-rhamnose. Moreover, the 2,3-*trans* flavan-3-ol, which was identified in **7**, was also seen in **8** based on its NMR data for the same positions (δ_H 4.67, d, *J* = 7.3 Hz/δ_C 81.1 for C-2 and δ_H 3.93, d, 7.3 Hz/δ_C 75.2 for C-3). The different features in **7** and **8** were observed in their optical rotations (+55 and −78, respectively), and the opposite CE at 290 nm in the CD spectra. The CD data of **8** showed a positive CE at 290 nm and a negative CE at 240 nm (Figure 54), indicating that the absolute configuration of **8** was 2*S*, 3*R* (Slade et al., 2005). Therefore, the structure of compound **8** was identified as (2*S*,3*R*)-rugonine B.
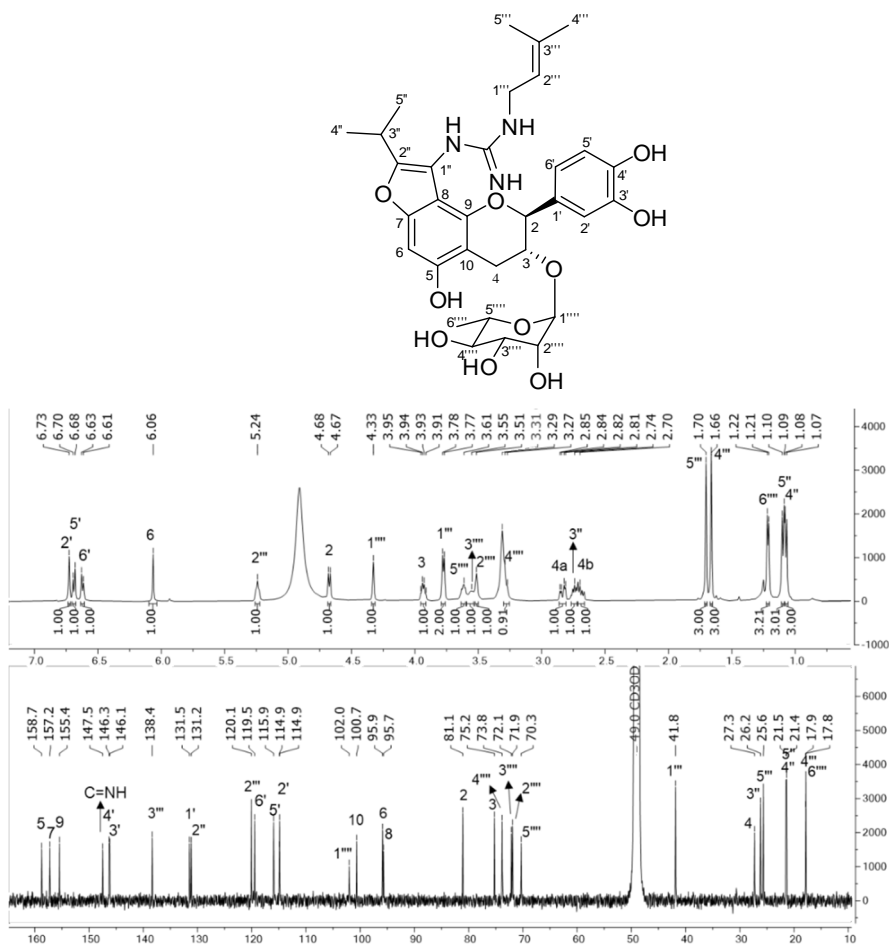
Figure 49. ^{1}H and ^{13}C NMR spectra of compound **8** (600/150 MHz, methanol-d_4)
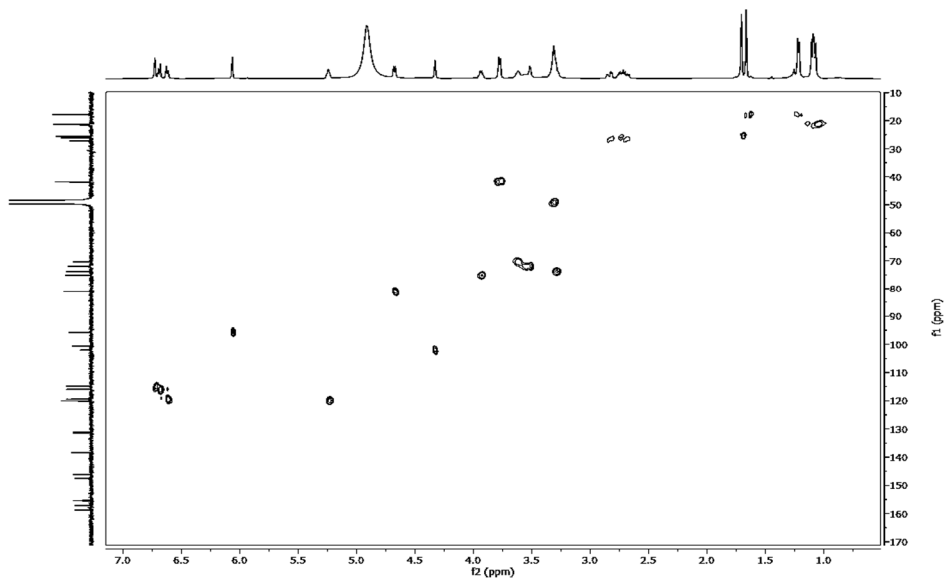
Figure 50. HSQC spectrum of compound **8**
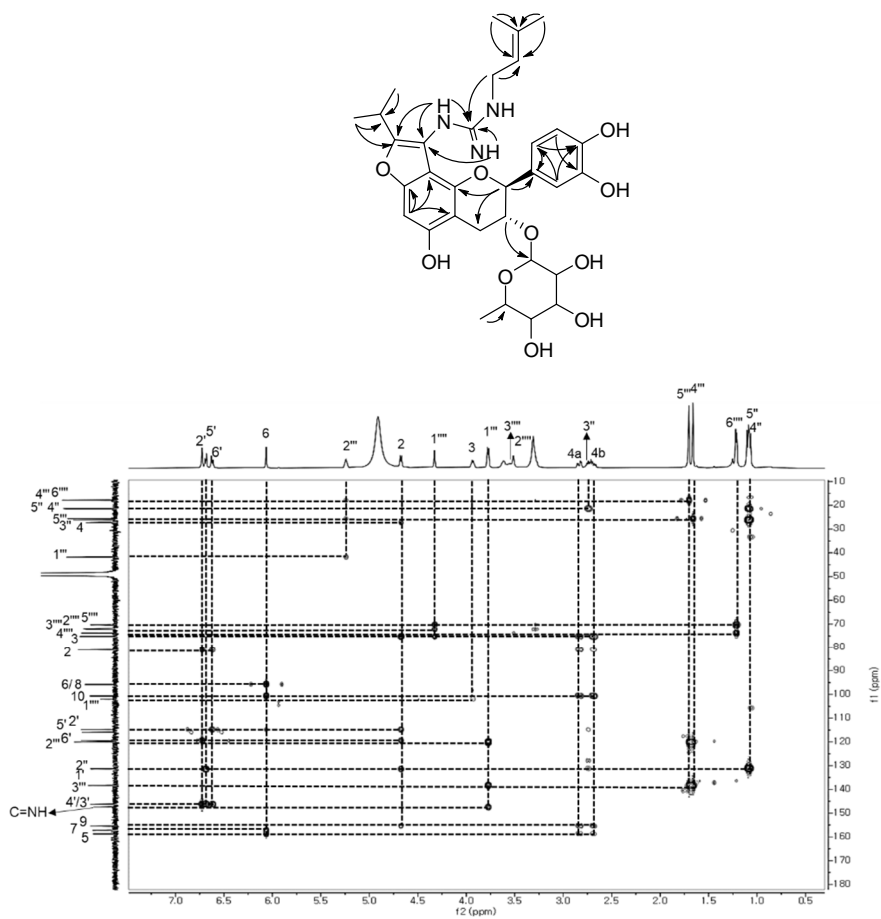
Figure 51. HMBC spectrum of compound **8**
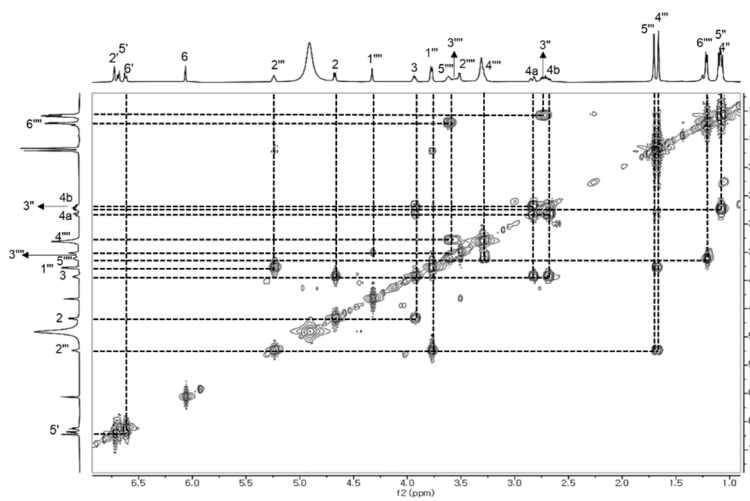
Figure 52. ^{1}H–^{1}H COSY spectrum of compound **8**
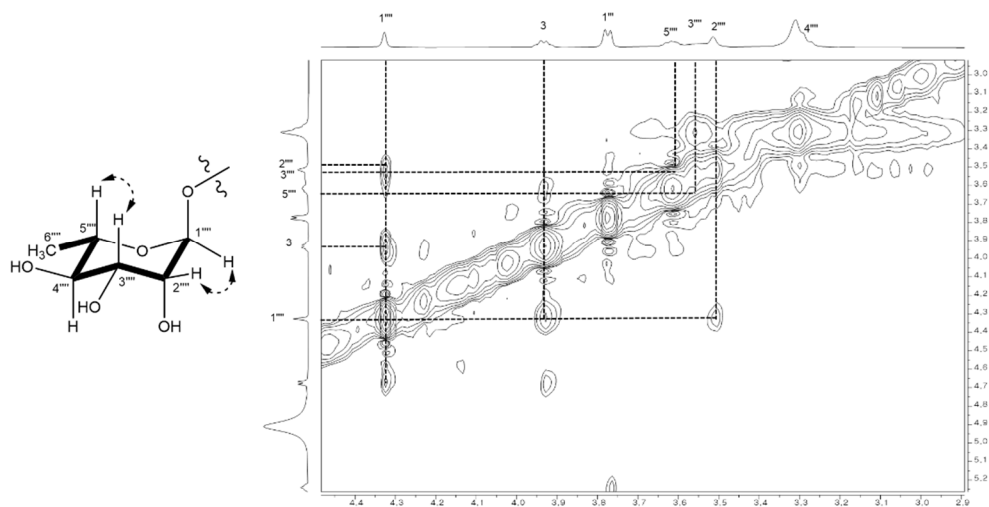
Figure 53. Key NOESY correlations of compound **8**
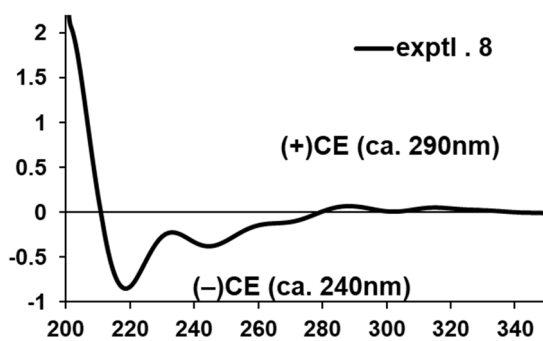
Figure 54. Experimental CD spectra of compound **8**

4.3. Compound **9**

Compound **9** was isolated as an amorphous powder with $[\alpha]_D^{25}$ +199.6 (*c* 0.1, MeOH). The molecular formula, $C_{32}H_{43}N_3O_{10}$, was deduced from its HRESIMS ion peak at *m/z* 646.2974 $[M + H]^+$ (calcd for $C_{32}H_{44}N_3O_{11}$, 646.2976). The IR spectrum of **3** displayed the absorption bands of hydroxyl or amine (3704 cm^{-1}), C-H in heteroaromatic rings (2922 cm^{-1}), C=NH or aromatic C (1682 cm^{-1}), benzofuran (1195 cm^{-1}), and C-O (1057 cm^{-1}) functional groups. The NMR data of **9** (Table 4) suggested that **9** was similar to **7** and **8**, excluding the lack of an olefinic bond at δ_H 5.26/δ_C 120.1, 138.5, and the presence of an additional methylene at δ_H 1.75/δ_C 42.4, and an oxygenated quaternary carbon at δ_C 70.9 ppm. The COSY correlation between H$_2$-1''' (δ_H 3.34) and H$_2$-2''' (δ_H 1.75), as well as HMBC cross-peaks from H$_2$-2''' to C-3''' (δ_C 70.9), C-4''', and C-5''' (δ_C 29.5) revealed the presence of a 4-hydroxyl-4-methyl pentyl moiety in **9** instead of the isoprenyl moiety in **7**. The rhamnose moiety in **9** was also determined by the same method as in **7**–**8** by NMR coupling constants and NOESY analysis, as well as comparing the HPLC retention time to the authentic L-rhamnose derivative suggesting that the stereochemistry of the sugar moiety was that of *α*-L-rhamnose. The absolute configurations at C-2, C-3 were identified based on the large coupling constants of H-2 (δ_H 4.70, *J* = 7.3 Hz)/H-3 (δ_H 3.96, *J* = 7.7, 5.4 Hz) indicating a 2,3-*trans* flavan-3-ol, and the CD data which showed a negative CEs at approximately 290, 240 nm similar to those of **7** (Figure 48). Therefore, the structure of compound **9** was identified as (2*R*,3*S*)-rugonine C.
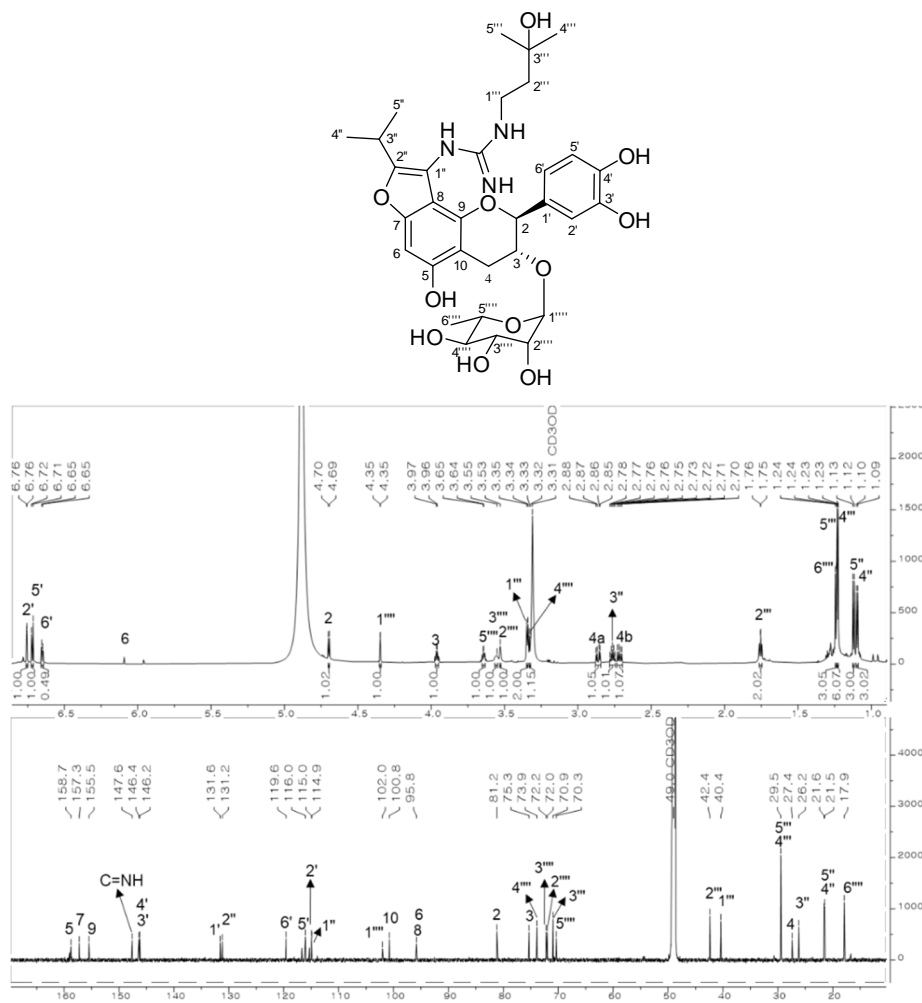
Figure 55. ^{1}H and ^{13}C NMR spectra of compound **9** (800/200 MHz, methanol-d_4)
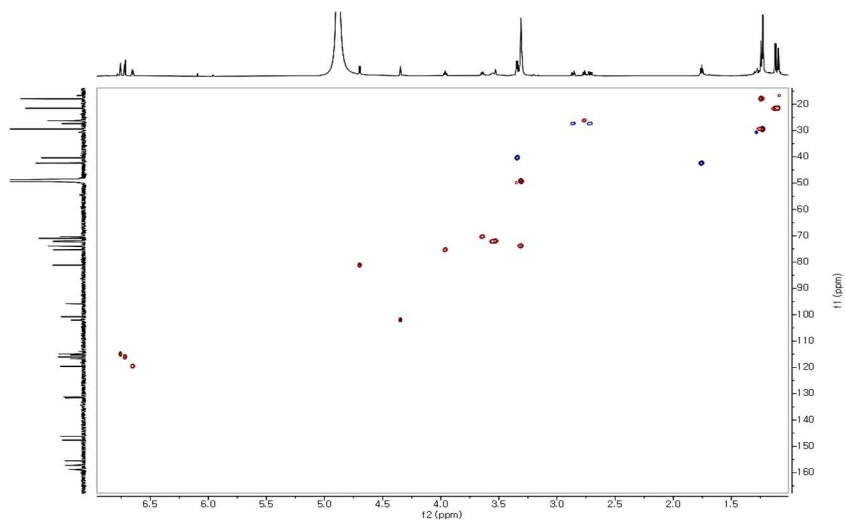
Figure 56. HSQC spectrum of compound **9**
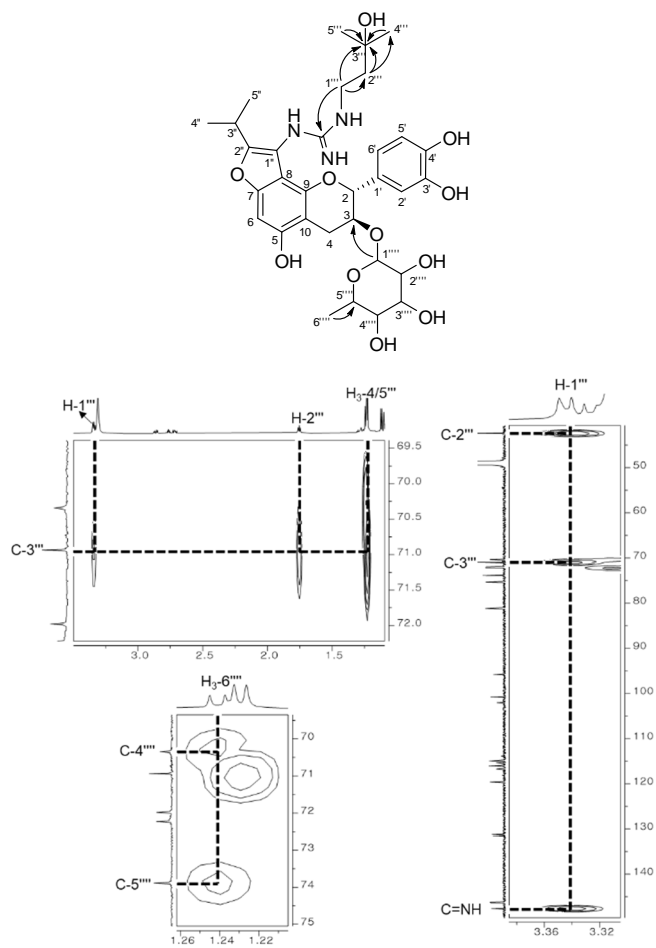
Figure 57. HMBC spectrum of compound **9**
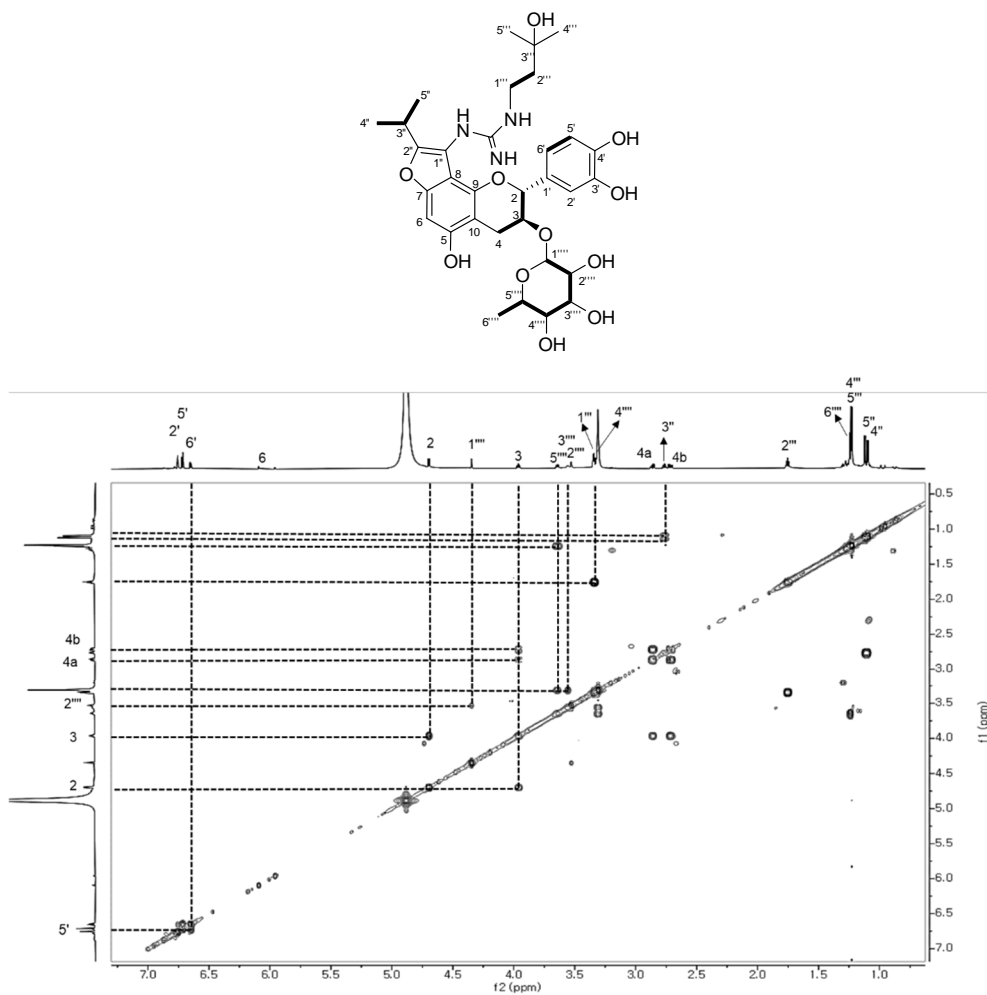
Figure 58. ^1H–^1H COSY spectrum of compound **9**
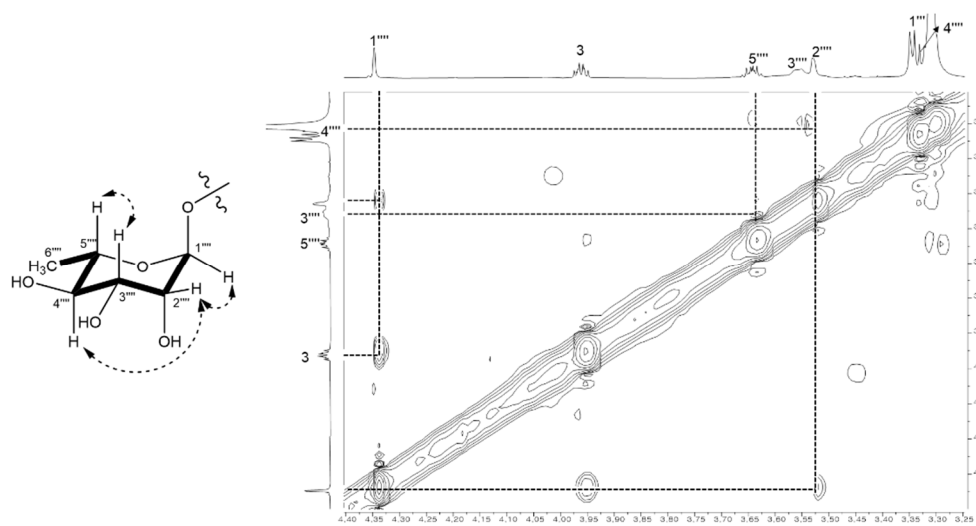
Figure 59. Key NOESY correlations of compound **9**

56

4.4. Compound 10

Compound **10** was isolated as an amorphous powder with $[\alpha]_D^{25}$ +27.5 (*c* 0.2, MeOH). The molecular formula, $C_{21}H_{23}N_3O_6$, was deduced from its HRESIMS ion peak at *m/z* 414.1649 $[M + H]^+$ (calcd for $C_{21}H_{24}N_3O_6$, 414.1665). The ^1H and ^{13}C NMR spectra of **10** showed a similar pattern to the core guanidine-fused catechin skeleton of **7** with the absence of sugar and isopentenyl moieties. The relative configuration of **10** at C-2/3 was identified based on its large coupling constants of H-2 (δ_H 4.54, *J* = 6.9 Hz), H-3 (δ_H 3.83, *J* = 12.6, 6.9 Hz) indicated a 2,3-*trans* flavan-3-ol as those of **7–9**. The absolute configuration of **10** was determined to be 2*R*,3*S* by experimental CD data that showed negative CEs at approximately 290 nm and 240 nm (Figure 48), which is typical for (+)-catechin (Kim et al., 2016; Korver & Wilkins, 1971; Slade et al., 2005). Hence, compound **10** was identified as (2*R*,3*S*)-rugonine D.
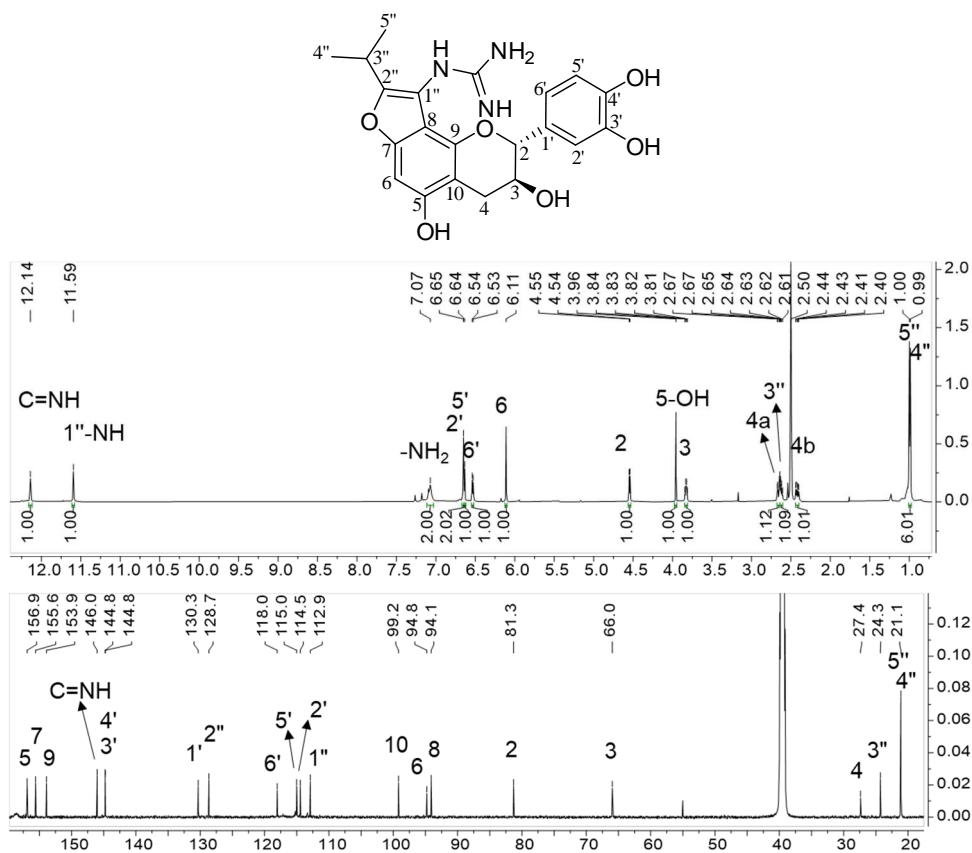
Figure 60. ^1H and ^{13}C NMR spectra of compound **10** (600/150 MHz, DMSO-d_6)
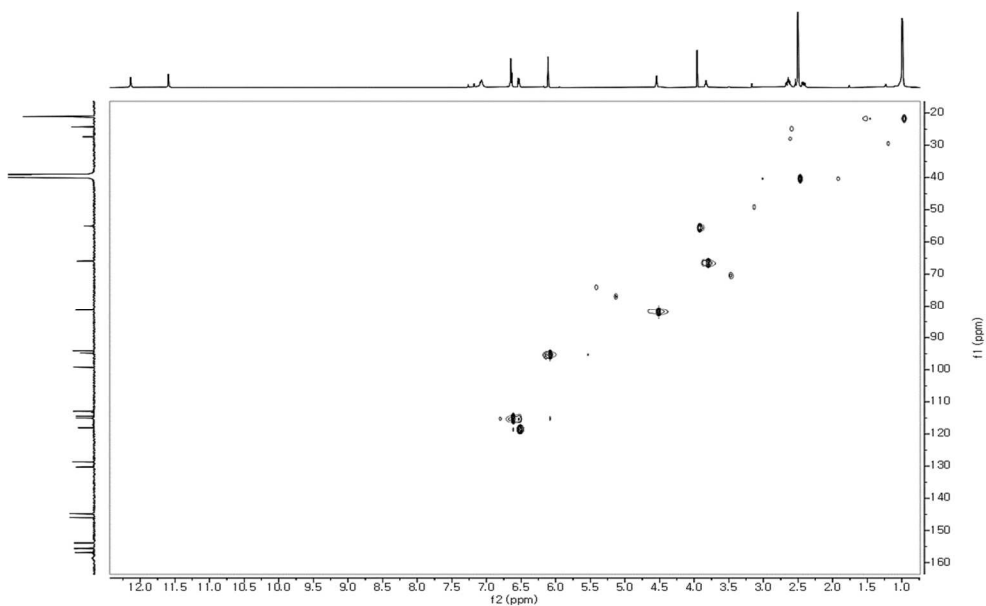
Figure 61. HSQC spectrum of compound **10**
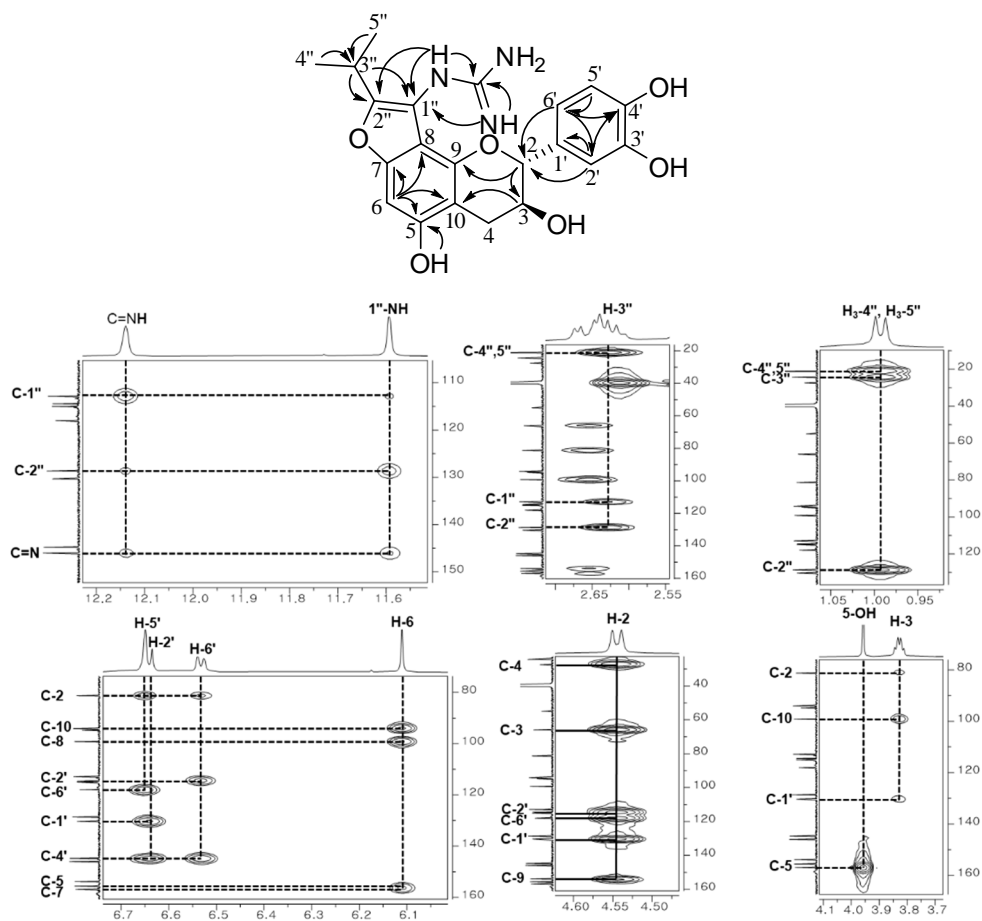
Figure 62. HMBC spectrum of compound **10**

4.5. Compound 11

Compound **11** was isolated as an amorphous powder with $[\alpha]_D^{25}$ −16.4 (*c* 0.2, MeOH). The molecular formula of **11** was the same as that of **10**, $C_{21}H_{23}N_3O_6$, which was deduced from its HRESIMS ion peak at *m/z* 414.1675 $[M + H]^+$ (calcd for $C_{21}H_{24}N_3O_6$, 414.1665). The ^1H and ^{13}C NMR spectra of **11** shared the similarity to those of **10** excluded signals at δ_H 4.79 (s) (H-2) and 4.02 (s) (H-3). These features and the opposite optical rotation between **10** and **11** suggested that they might have different orientations at C-2 and C-3. In particular, based on the small coupling constants at C-2 and C-3, the relative configuration of **11** was identified as 2,3-*cis* flavan-3-ol (Slade et al., 2005). The absolute configuration of **11** was identified as 2*S*, 3*S* by the CD spectrum, which showed negative CEs at 290 and 240 nm (Figure 67) (Kang et al., 2017; Slade et al., 2005). Hence, the structure of **11** was determined to be (2*S*,3*S*)- rugonine E.
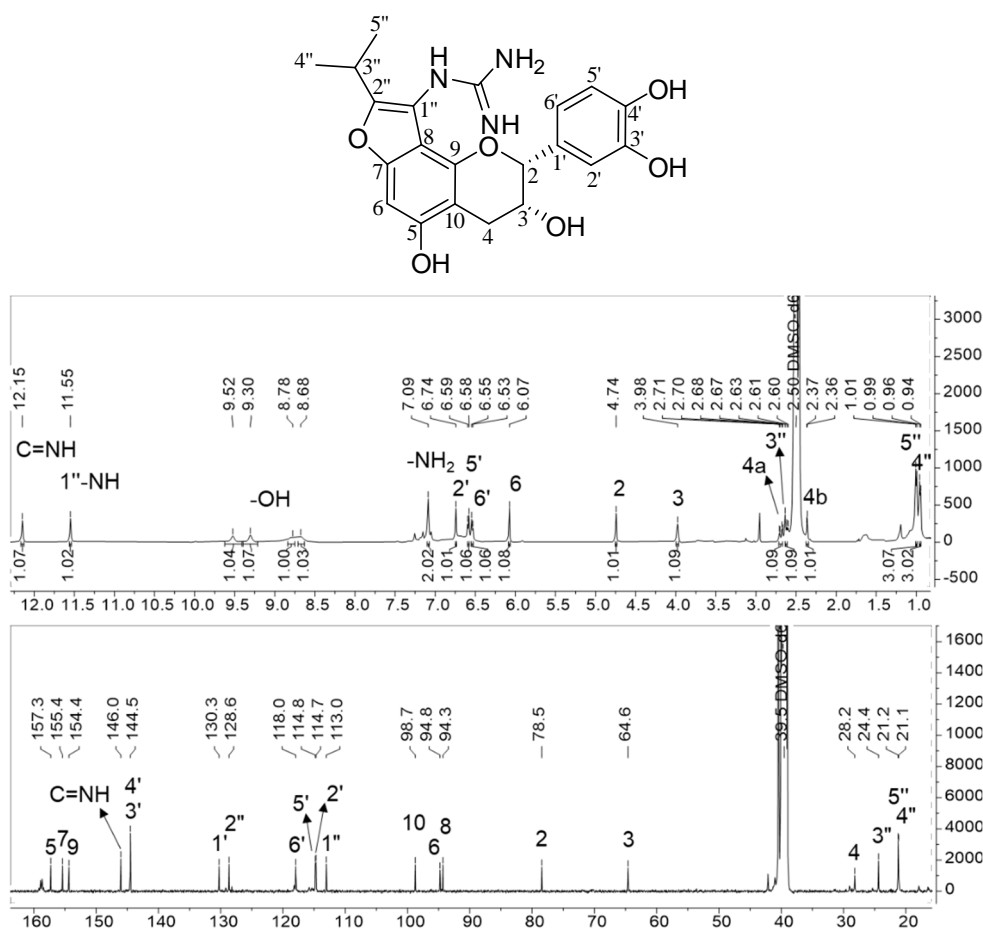
Figure 63. ^1H and ^{13}C NMR spectra of compound **11** (500/125 MHz, DMSO-d_6)
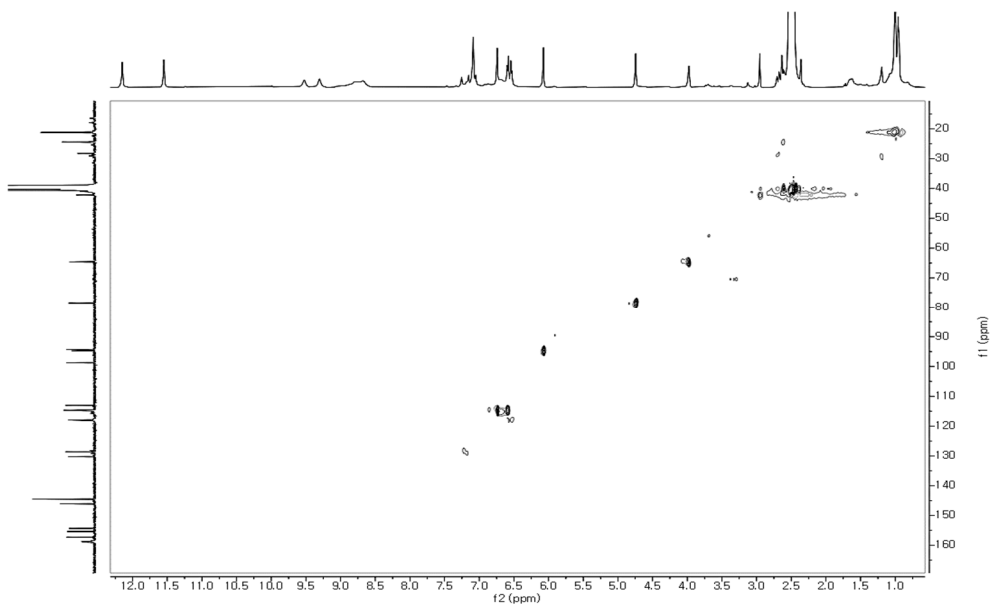
Figure 64. HSQC spectrum of compound **11**
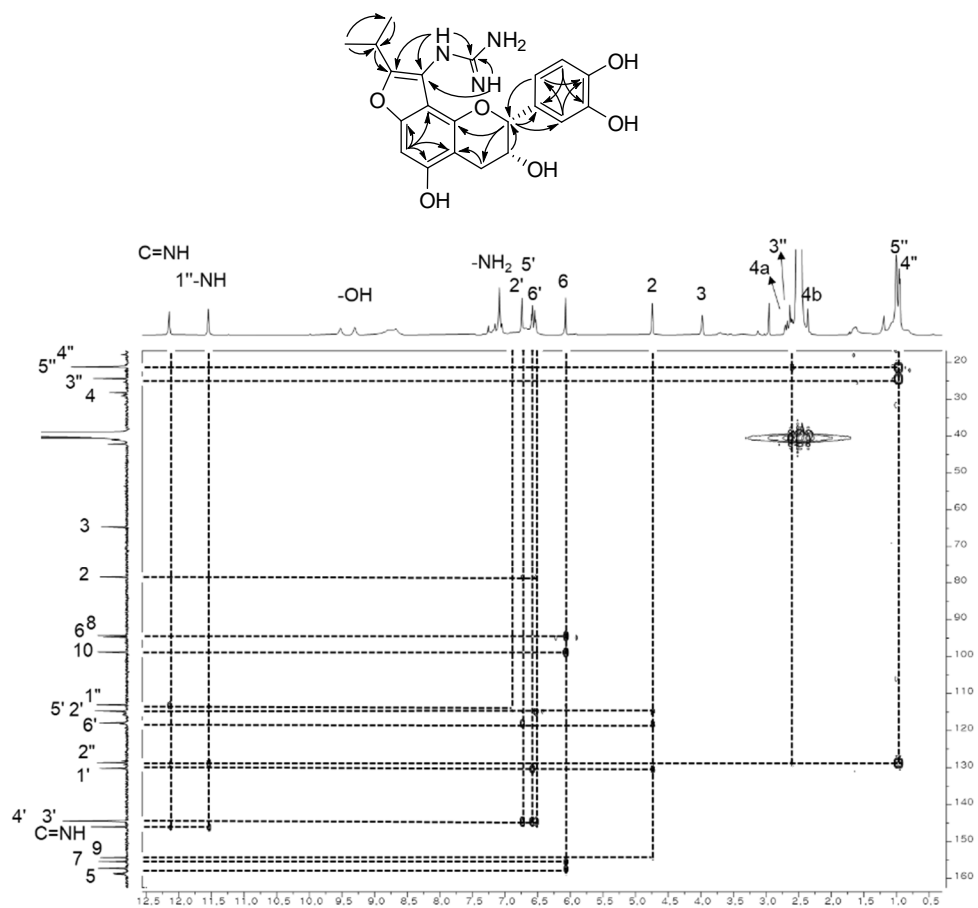
Figure 65. HMBC spectrum of compound **11**
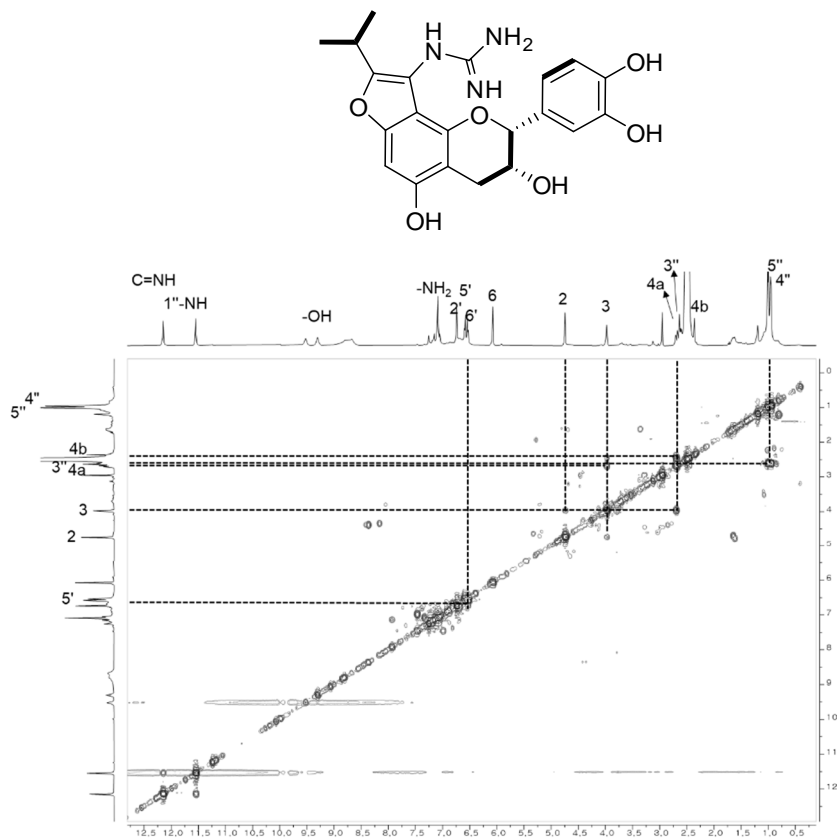
Figure 66. ^1H–^1H COSY spectrum of compound **11**
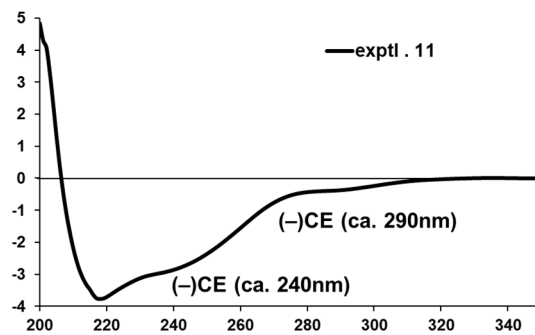
Figure 67. Experimental CD spectra of compound **11**

4.6. Compound 12

Compound **12** was isolated as an amorphous powder with $[\alpha]_D^{25}$ +34.1 (*c* 0.2, MeOH). The molecular formula, $C_{26}H_{31}N_3O_6$, was deduced from its HRESIMS ion peak at *m/z* 482.2316 $[M + H]^+$ (calcd for $C_{26}H_{32}N_3O_6$, 482.2291). The ^1H and ^{13}C NMR spectra of **12** shared a pattern similar to those of **7** except for the lacking of the sugar moiety. This led to the conclusion that **12** is another derivative of guanidine-catechin with an isoprenyl substituent that included two methyl groups at C-4''' (δ_H 1.64, δ_C 25.3) and C-5''' (δ_H 1.68, δ_C 17.8), an olefinic bond at C-3''' (δ_C 135.0), C-2''' (δ_H 5.22, δ_C 120.2), and methylene at C-1''' (δ_H 3.75, δ_C 44.1). The connection between the isopentyl group and guanidine moiety was indicated via the HMBC correlation from H-1''' (δ_H 3.75 ppm) to δ_C 146.8 ppm. Similar to **7,** the absolute configurations of **12** were elucidated as 2*R*,3*S* by the large coupling constants of H-2 (*J* = 5.8 Hz), H-3 (*J* = 6.2 Hz) indicating a 2,3-*trans* flavan-3-ol, and negative CEs at approximately 280−290 nm and 240 nm in CD spectra (Figure 48). Therefore, compound **12** was determined to be (2*R*,3*S*)- rugonine F.
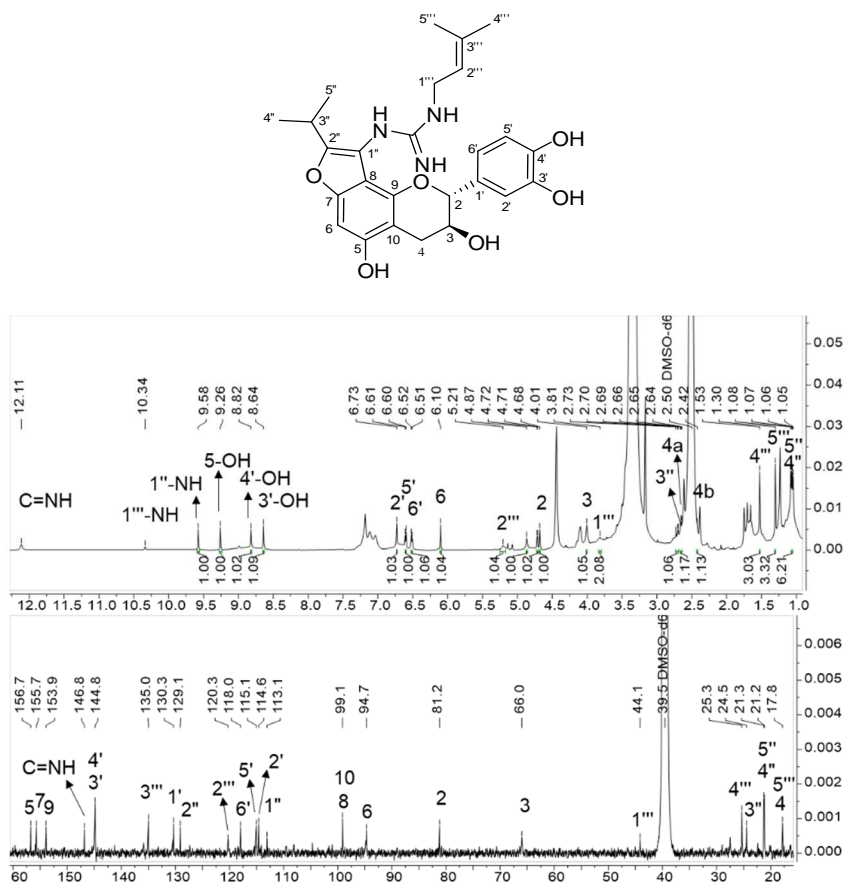
Figure 68. ^1H and ^{13}C NMR spectra of compound **12** (600/150 MHz, DMSO-d_6)

62
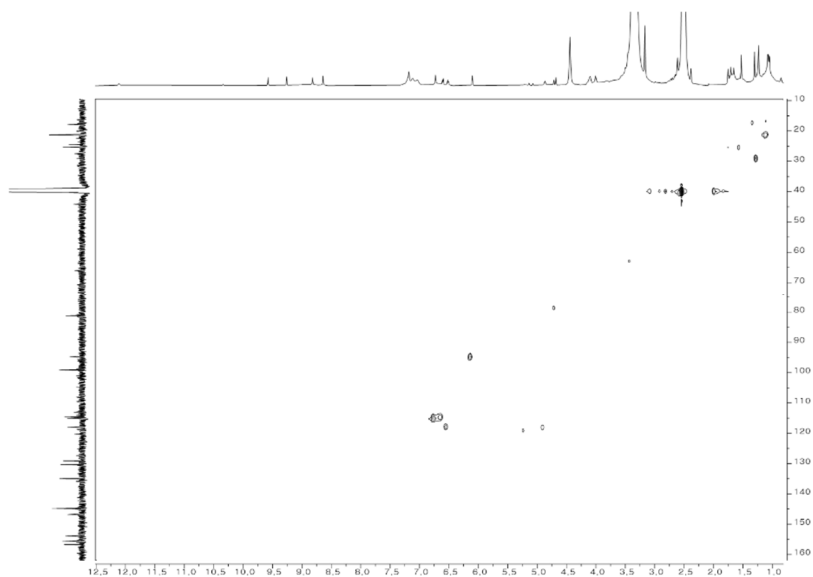
Figure 69. HSQC spectrum of compound **12**
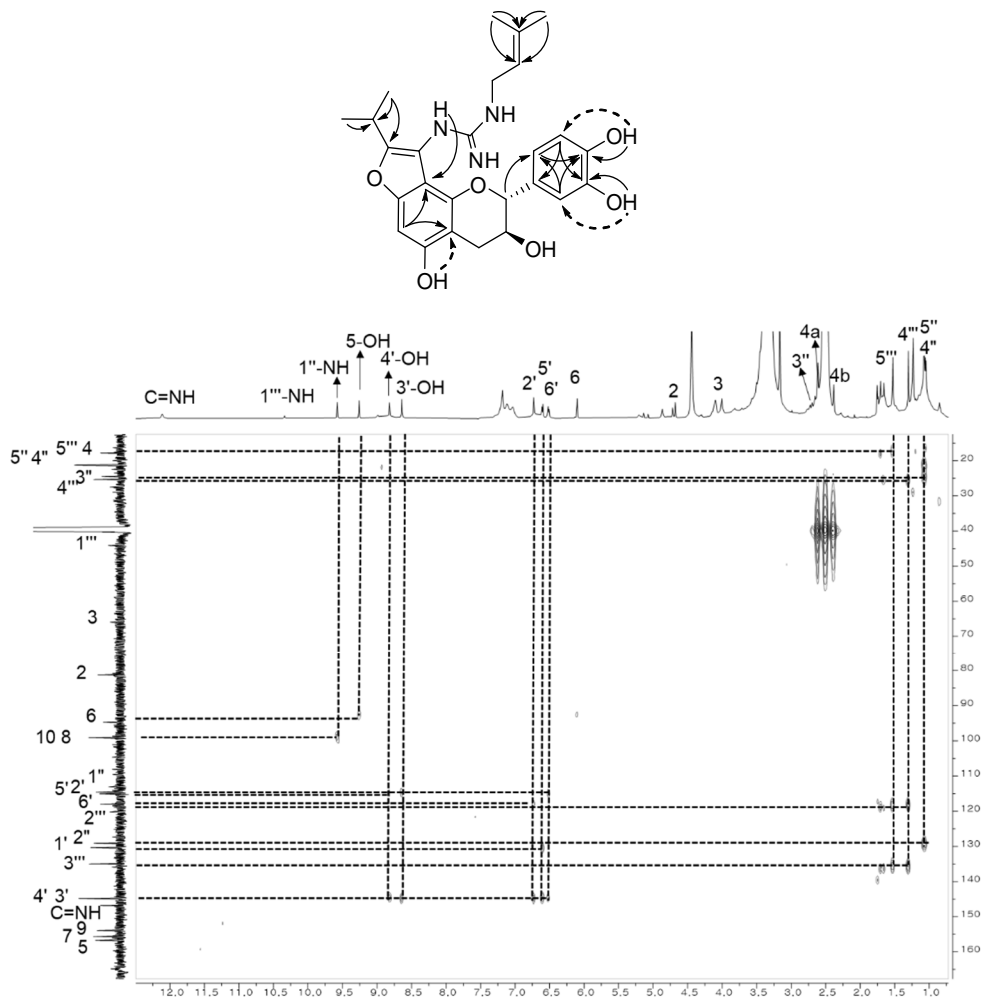
Figure 70. HMBC spectrum of compound **12**

4.7. Compound 13

Compound **13** was isolated as an amorphous powder with $[\alpha]_D^{25}$ + 47.5 (*c* 0.2, MeOH). The molecular formula, $C_{27}H_{33}N_3O_{10}$, was deduced from its HRESIMS ion peak at *m/z* 560.2272 $[M + H]^+$ (calcd for $C_{27}H_{34}N_3O_{10}$, 560.2244). The ^1H and ^{13}C NMR data of compound **13** shared the similarity with compound **7** excluding the absence of the isoprenyl substituent. The sugar moiety was determined by an anomeric signal (δ_H 4.30, δ_C 99.9 ppm), four oxygenated carbons (δ_H 3.15–3.47 ppm, and δ_C 69.0–72.0 ppm), and an additional methyl at δ_H 1.14, δ_C 17.9 ppm, which indicated the presence of a rhamnose moiety in **13**. Based on the HMBC correlations from anomeric proton H-1'''' (δ_H 4.30) to C-3 (δ_C 72.04 ppm), the sugar moiety of **13** was identified as 3-O-rhamnoside. The relative configuration of the sugar moiety was determined by coupling constants and NOESY analyses. The small coupling constant of H-1 (δ_H 4.30, s) and large $^1J_{CH}$ (170.0 Hz) suggested an α orientation of C-1'''', and the observed large coupling constants $^3J_{H-H}$ of H-4'''' (9.1 Hz) and H-5'''' (12.4, 6.1 Hz) indicated that H-3''''/H-4''''/H-5'''' were in axial orientation. Furthermore, the NOESY correlations of **13** were confirmed for the relative configuration in the sugar moiety as established by *J* values (Figure 75). The absolute configuration of the rhamnose was established by acid hydrolysis followed by conversion to the corresponding thiocarbamoyl-thiazolidine carboxylate derivative with L-cysteine methyl ester and *o*-tolyl isothiocyanate (Tanaka et al., 2007). The consistent HPLC retention times for the sugar in **13** and the authentic L-rhamnose derivative confirmed that the stereochemistry of the sugar moiety was that of α- L-rhamnose. The absolute configurations of **13** at C-2, and C-3 were determined to be 2*R*, 3*S* based on their large coupling constants at H-2 (*J* = 6.7 Hz)/H-3(*J* = 12.7, 6.5 Hz), which indicated a 2,3-*trans* flavan-3-ol, and the CD data, which showed a negative CE at approximately 290 nm and 240 nm (Figure 48). Finally, the structure of **13** was identified as (2*R*,3*S*)- rugonine G.
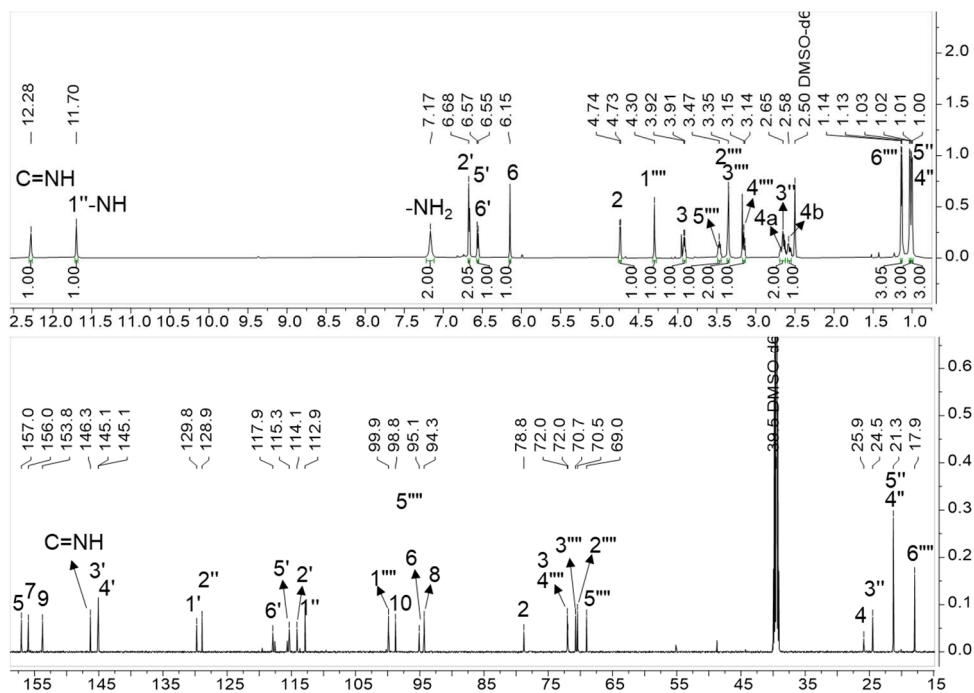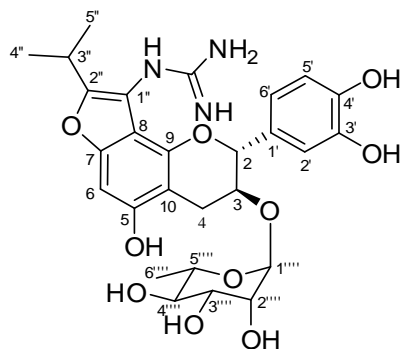
Figure 71. ^1H and ^{13}C NMR spectra of compounds 13 (600/150 MHz, DMSO-d_6)
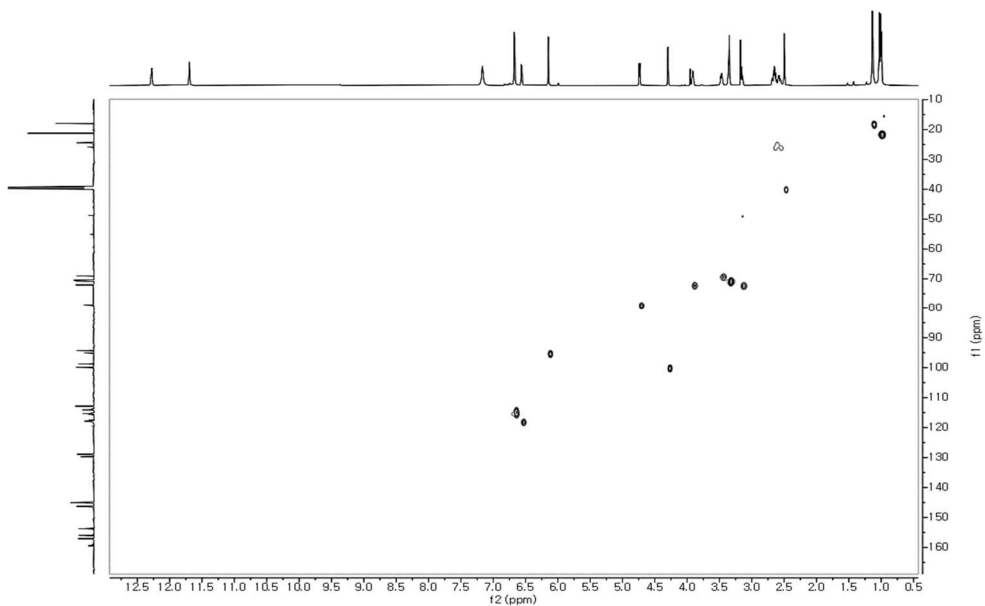
Figure 72. HSQC spectrum of compound **13**
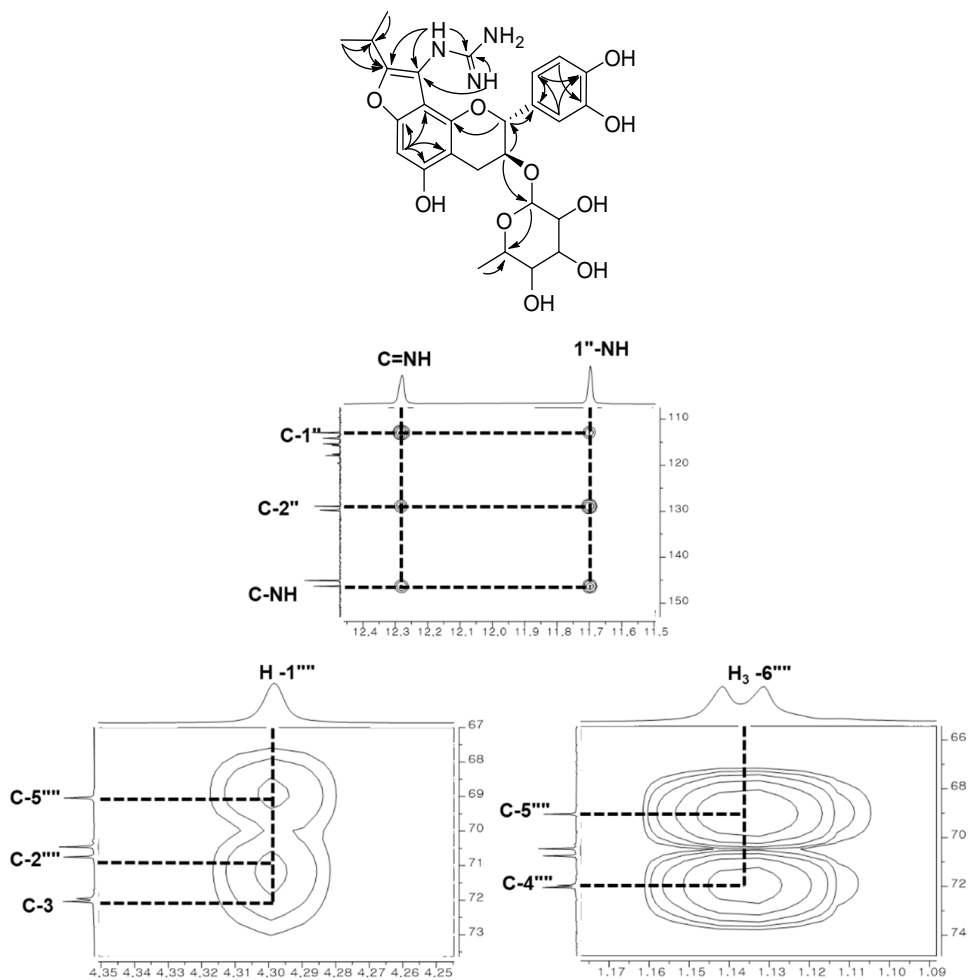
Figure 73. HMBC spectrum of compound **13**
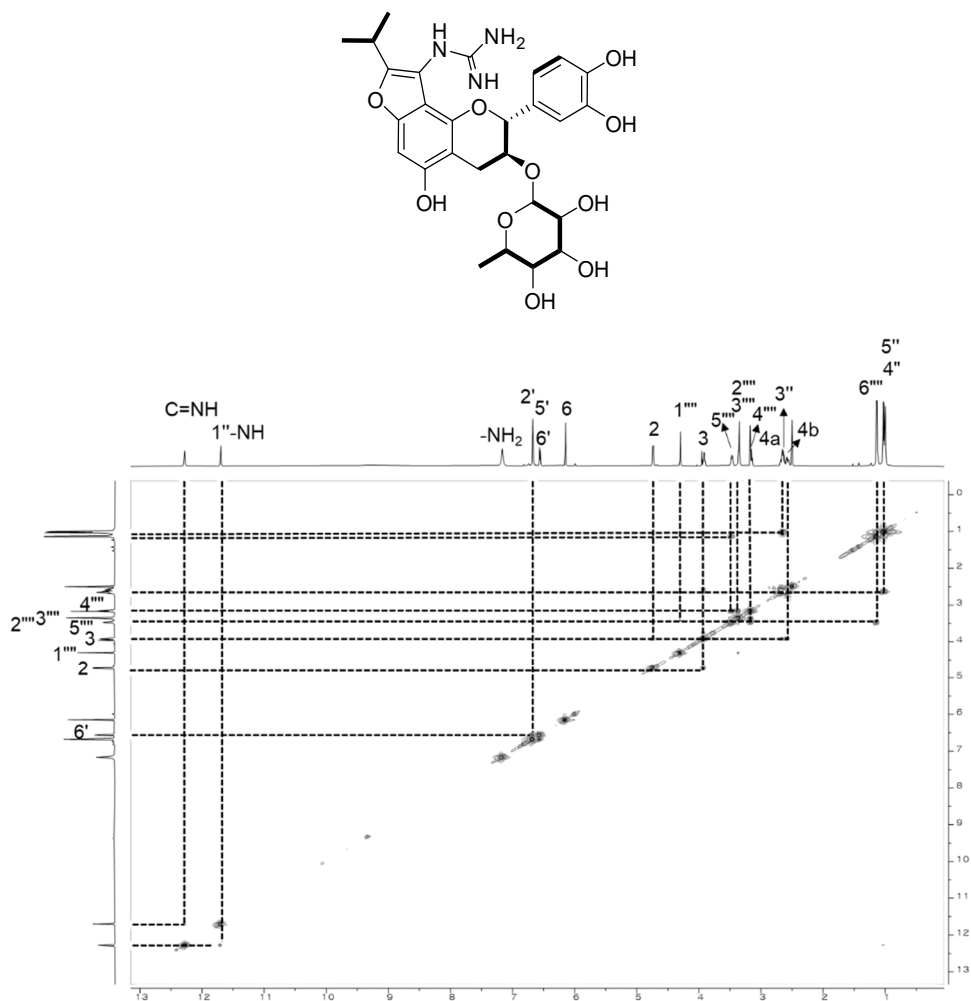
Figure 74. ^{1}H–^{1}H COSY spectrum of compound **13**
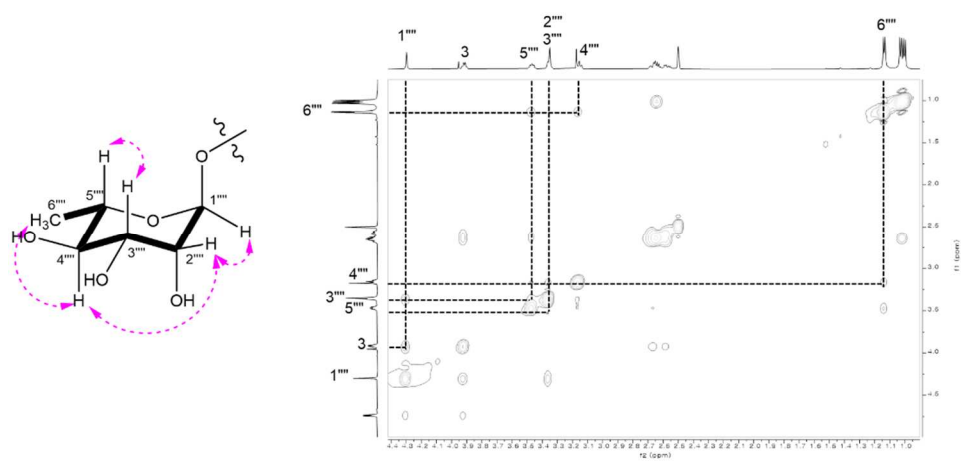
Figure 75. Key NOESY correlations of compound **13**

4.8. Compound **14**

Compound **14** was isolated as an amorphous powder with $[\alpha]_D^{25}$ +20.3 (*c* 0.2, MeOH). The molecular formula, $C_{21}H_{23}N_3O_5$, was deduced from its HRESIMS ion peak at *m/z* 396.1581 [M – H]⁻ (calcd. for $C_{21}H_{22}N_3O_5$, 396.1565). The ¹H NMR and ¹³C NMR data of **14** showed similarities to those of **10** and **11** and the mass was different by 16 Da, suggesting that the structure of **14** differed from those of **10** and **11** by less than one hydroxy group. The differences were also indicated by the presence of methylene at δ_C 28.6/δ_H 2.00, 1.82 instead of an oxygenated group as in **10** and **11**, and the greater upfield shift of C-4 (δ_C 18.8) in **14** compared to the other compounds. The HMBC correlations from H-2 (δ_H 4.80) to C-3 (δ_C 28.6), C-4 and the ¹H-¹H COSY spin system of H-2/H_2-3/H_2-4 suggested that **14** contained a C6-C3-C6 ring similar to that of luteoliflavan (Roemmelt et al., 2003). Thus, the planar structure of **14** was elucidated, as shown in Figure 76. The configuration of C-2 was determined by comparing its ECD and NMR data with references. To date, there only a few flavans have been reported to occur naturally and all of them have the 2*R* absolute configuration that would be expected from the flavanone origin (Slade et al., 2005). Flavans showed a low specific rotation, making firm conclusions difficult, and their configuration could be determined only by studying their CD data (Slade et al., 2005). Experimentally, the CD spectrum of **14** showed a negative cotton effect (CE) at a ¹L_b of 285 nm (Figure 48) which indicated the 2*R* absolute configuration (Slade et al., 2005). Moreover, the NMR data of **14** shared a similarity with those of luteoliflavan 5-glucoside (Roemmelt et al., 2003), suggesting a 2*R* configuration. Therefore, compound **14** was identified as (2*R*)-rugonine H.
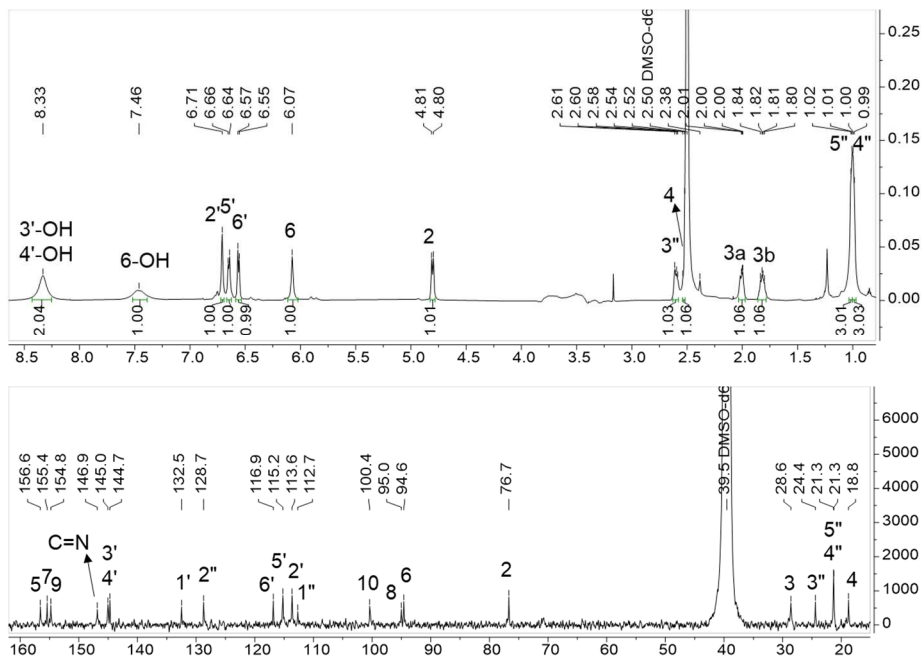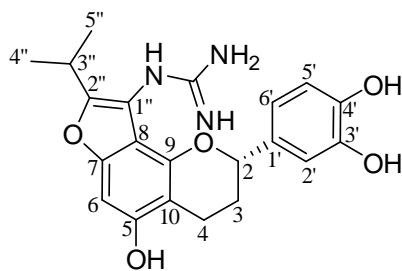
Figure 76. ^1H and ^{13}C NMR spectra of compounds **14** (500/125 MHz, DMSO-d_6)
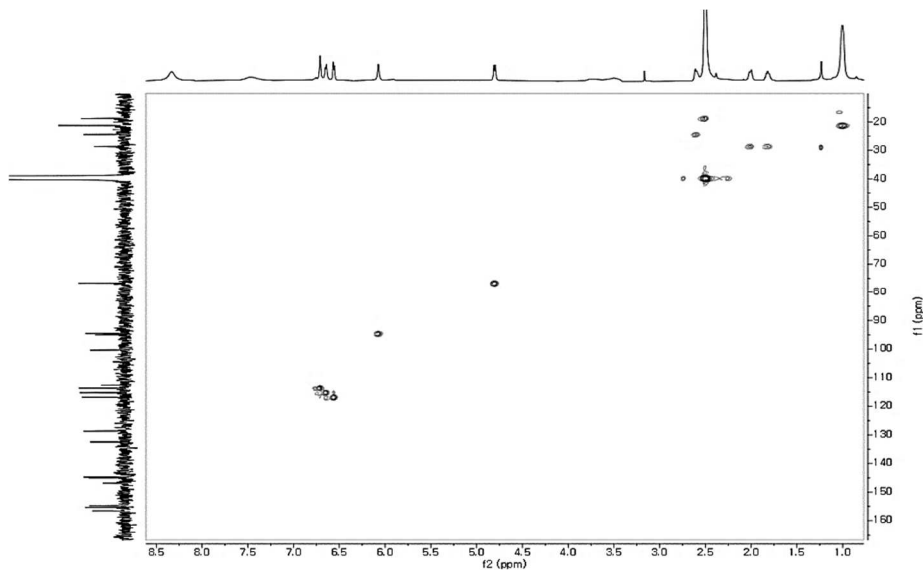
Figure 77. HSQC spectrum of compound **14**
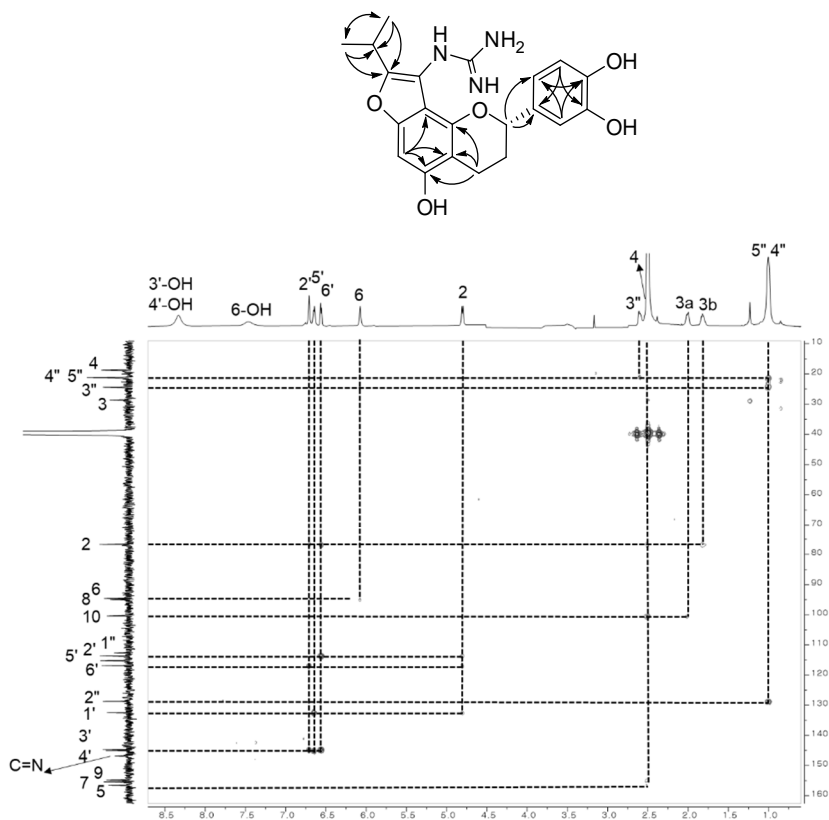
Figure 78. HMBC spectrum of compound **14**
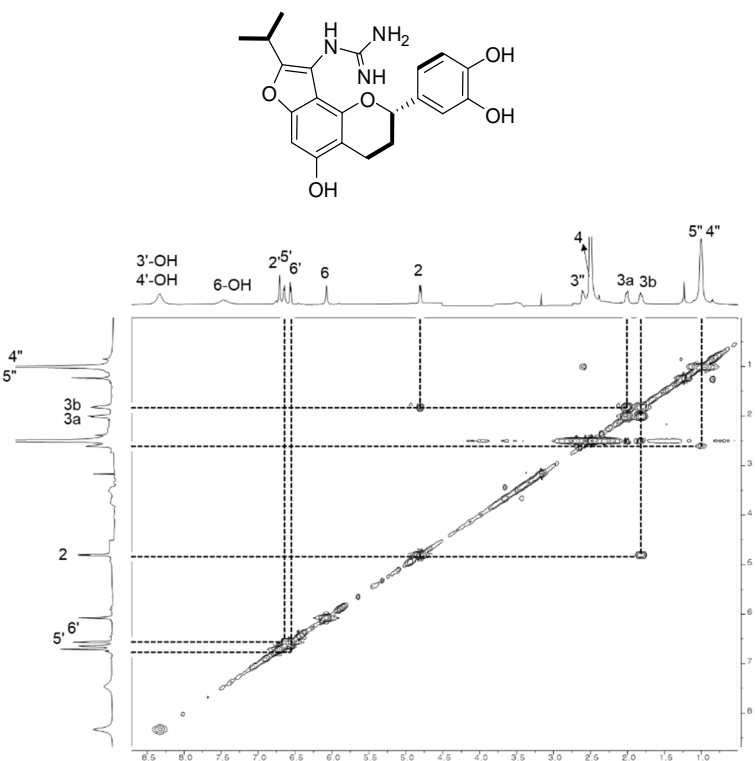
Figure 79. ¹H–¹H COSY spectrum of compound **14**

70

4.9. Compound 15

Compound **15** was isolated as an amorphous powder. The molecular formula, $C_6H_{13}N_3$. The ^1H NMR and ^{13}C NMR data of **15** were identical with a compound named galegine as compared to the literature (Monache et al., 1999).

Galegine
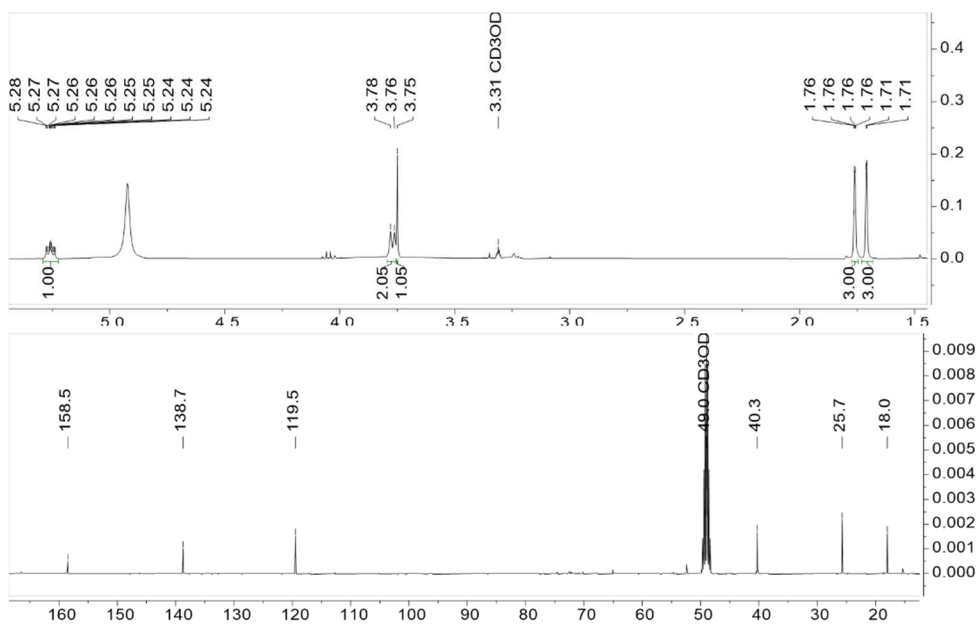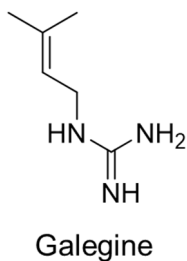
Figure 80. ^1H and ^{13}C NMR spectra of compounds **15** (600/150 MHz, methanol-d_4)

5. Biological activities of compounds 1–15

5.1. Biological activities of compounds 1–6

Recent studies of antidiabetic effects of *A. cordifolia* leaf extracts (Eliakim-Ikechukwu & Obri, 2010; Mohammed et al., 2012) suggested the evaluation of **1–6** for stimulatory effects on the uptake of 2-NBDG by 3T3-L1 adipocyte cells. Firstly, six isolated compounds had been screened for their glucose uptake stimulatory activities on 3T3-L1 adipocyte cells at a concentration of 20 μM and compared to insulin as the positive control (100 nM). The cell viability result indicated that none of those compounds showed cytotoxicity at that concentration (Figure 81).
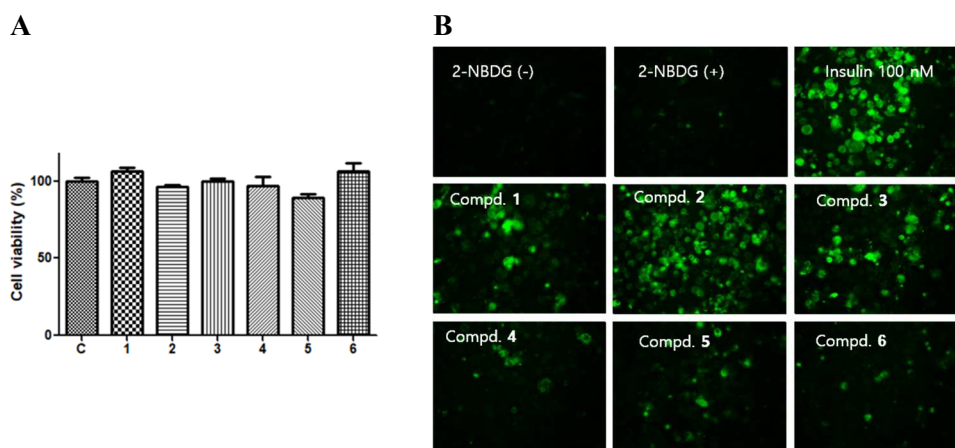
Figure 81. Cell viability and screening of glucose uptake effects of compounds **1–6**. (A) The cell viability of compounds **1–6** from *A. rugosa* using MTT assay on glucose uptake by 3T3-L1 adipocytes using a fluorescent derivative of glucose (2-NBDG). The compounds were treated to differentiated cells at 20 μM directly after being separated from the other diastereomer of each pair. Insulin (100 nM) was used as a positive control. (B) After 1h of incubation with 2-NBDG, fluorescent signals were measured at Ex/Em = 450/535 nm.

Compounds **1–6** (20 μM) were used to treat 2-NBDG uptake in differentiated 3T3-L1 adipocytes. As shown in Figure 82A, compounds **1–3** exhibited the most potent stimulatory effects compared to insulin. Measurement of fluorescent signals after treatment was performed using fluorescence microscopy to assess the efficacy of the transport of 2-NBDG into cells. Increased signal intensities after treatment with the compounds were readily observed in cells treated with **1–3** (Figure 82B). These fluorescence intensities indicated that 2-NBDG uptake increased in a dose-dependent manner for the selected compounds (Figure 82C).

72
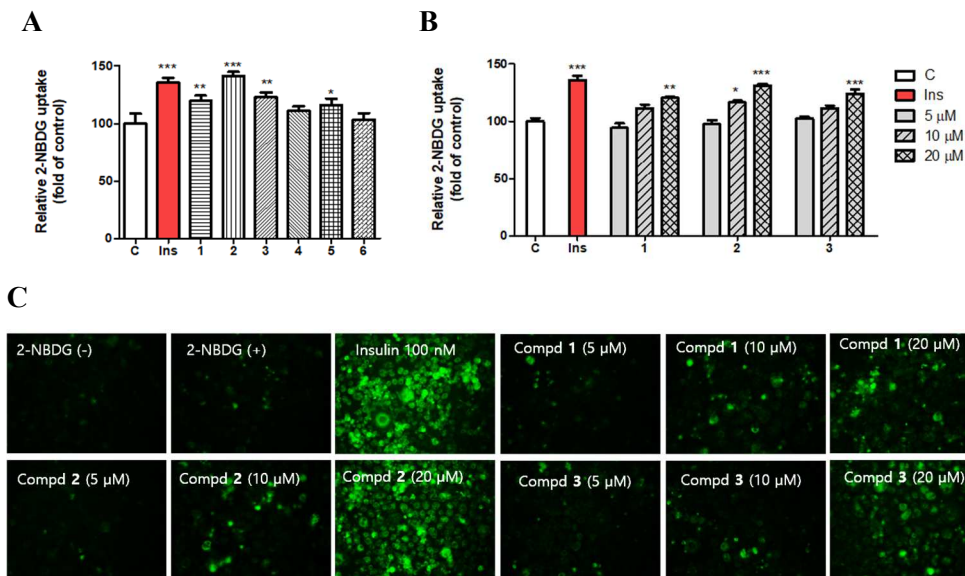
Figure 82. Glucose uptake effects of compounds **1–3** in a dose-dependent manner in differentiated 3T3-L1 adipocytes. (A) Relative glucose uptake levels with **1–6** versus glucose uptake by 3T3-L1 adipocytes using a fluorescent derivative of glucose, 2-NBDG. The compounds were added to differentiated cells at 20 μM. Insulin (100 nM) was used as a positive control. (B) Differentiated 3T3-L1 adipocytes were treated with compounds **1–3** at concentrations of 5, 10, and 20 μM or insulin at 100 nM. (C) Compounds **1–3** concentrations dependently increase glucose uptake in 3T3-L1 adipocytes. After 1 h of incubation, the fluorescent signals were measured at Ex/Em = 450/535 nm. Data are expressed as the mean ± SD (n=3), each performed in triplicate; $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$, compared to the control.

5.2. Biological activities of compounds **7–15**

5.2.1. Biological activities of compounds **7–15** in autophagy modulating

To screen autophagy regulatory activities of compounds **7–15**, the GFP-LC3 stably expressing HEK293 cells were administrated. In the HEK293 cells, the formation of punctuates could be observed by using chloroquine (CQ) and rapamycin (RAPA) which were known to inhibit and induce autophagy, respectively. In the confocal microscopic image, the tested CQ and RAPA showed the formation of punctuates, and a GFP signal was detected in the cell cytosols. The results indicated that the formation of punctuates in the GFP-LC3 stable expressing HEK293 cells reveals autophagy regulation of compounds. Nine isolated compounds from *A. rugosa* (**7–15**) were treated at a concentration of 20 μM (the concentration that does not show cytotoxicity, Figure 83A) for 24 h the cytosols were checked under a confocal microscope. Compared with the control groups, a significant increase of LC3 puncta in HEK293-GFP-LC3 cells has been observed after treatment with compounds **10–13** at a concentration of 20 μM (Figure 83B).

(A)   **(B)**
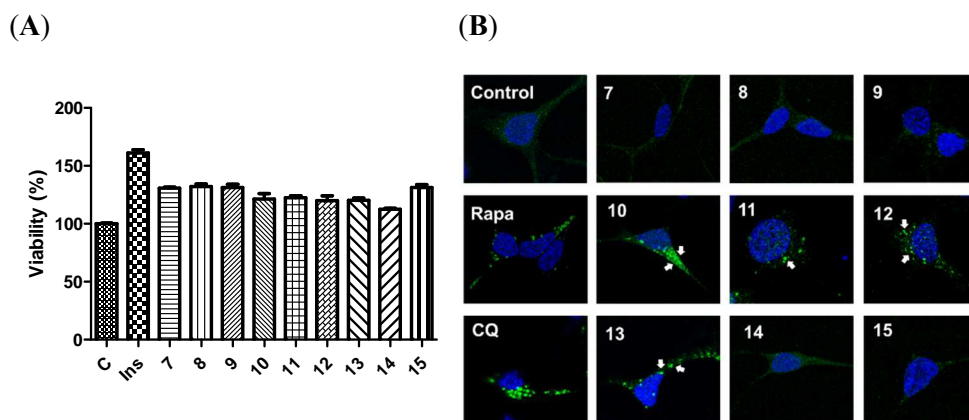
Figure 83. Cell viability in myoblast C2C12 cell line and screening of autophagy regulation effects of compounds **7–15** in HEK293-GFP-LC3 cells. (A) The cell viability of compounds **7–15** from *A. rugosa* using MTT assay on HEK293 cells. Tested compounds were treated at 20 μM compared to positive control insulin at 20 nM. (B) Screening the regulation of autophagy for isolated compounds (**7–15**) from *A. rugosa*. HEK293 cells stably expressing GFP-LC3 were treated with 20 μM of tested compounds for 24 h, and confocal imaging was used to examine the production of LC3-GFP punctuates.

5.2.2. Compounds **10**–**13** inhibit the autophagy flux

To investigate autophagy regulation by the tested compounds, the protein expression levels of LC3B and p62 were detected. The LC3B level was calculated as the LC3B II/I level, indicating the transition of LC3B I to II in autophagy. The p62 protein level was detected to monitor the degradation level in autophagosomes. CQ and RAPA were used as an autophagy inhibitor and activator, respectively. Under the CQ treatment, both the protein levels of LC3B and p62 were increased. Since CQ blocks the autophagic system, p62 could not be degraded, and LC3B II and I could not be used properly; therefore, they accumulated in the cells. In contrast, in the RAPA-treated group, autophagy was induced, leading to a strong LCB I to II transition; hence, a decrease in the LC3B I level and an increase in the LC3II level were observed in the Western blot results. Moreover, the p62 level was decreased due to the degradation of autophagolysosomes (Wang et al., 2015). As a result, compounds **10**–**13** at various concentrations of 5 and 20 μM exhibited the same effects in the CQ-treated group, which showed increases in the LC3B and p62 levels (Figure 84A). Moreover, the protein expression levels of p62 and LC3 II increased in a dose-dependent manner compared to those of the control group (Figure 84A). This finding suggested that these compounds inhibited protein degradation by the autolysosome by blocking the fusion of autophagosomes and lysosomes.

To prove this hypothesis, GFP-mRFP-LC3 was transfected into HEK293 cells, and autophagosomes were examined using confocal microscopy. GFP-mRFP-LC3 transfected cells were treated with compounds **10**–**13** at a concentration of 20 μM, and DAPI staining was performed on glass slides. GFP and mRFP were activated in the autophagosome, but when the autophagosome became an autolysosome, GFP was deactivated due to the autolysosome's pH. Therefore, in the RAPA-treated group in which autophagy was induced, GFP activation was lost, but mRFP puncta were activated and expressed. Thus, the combined image showed red fluorescence puncta that indicated the increased autophagy flux and increased numbers of autolysosomes under RAPA treatment compared to the control group (Figure 84B). In contrast, the CQ treatment group had a higher quantity of yellow puncta than the control group in the merged image, suggesting that autophagosome activation was prevented by CQ treatment, and autophagosomes accumulated in the cytosol, revealing a large number of GFP- and mRFP-activated autophagosomes (Figure 84B). Based on this result,

the effects of compounds **10–13** on autophagosomes were also analyzed, and the yellow puncta in these treatment groups appeared strongly. According to these findings, guanidine-conjugated catechins (**10–13**) isolated from *A. rugosa* leaves have an inhibitory effect on autophagic flux.

Based on the structures and activities of **7**–**15**, a gross SAR for this guanidine type can be delineated as follows: (1) the core skeleton of guanidine-conjugated flavan-3-ol is crucial for autophagy modulation activity since the activities were observed in **10–13** but not **14**, which has no hydroxy functional group at C-3 and showed no effect in the autophagy assay (Figure 83B); (2) the addition of one substituent, such as one sugar moiety or a side chain, does not affect the activities, but two or more substituents have an effect, as indicated by the activities of **12** and **13** and the lack of activities observed for **7**–**9** (Figure 9); (3) the different configurations at C-2 and C-3 did not affect the autophagy inhibition effects, as shown for **10** and **11**, which differed in the C-2 and C-3 absolute configurations but both showed activity.
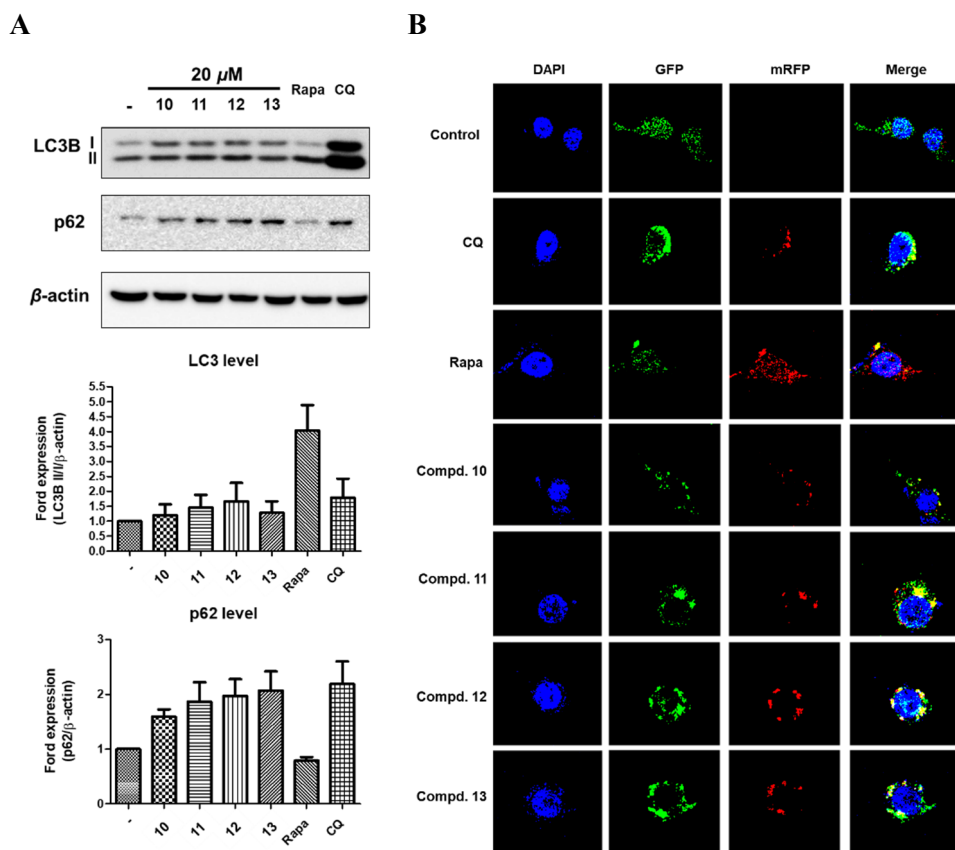
Figure 84. The autophagy marker protein expression level of compounds treated cells and the GFP-mRFP-tagged LC3 punctuate images with the treatment of compounds **10–13**. (A) The level of autophagy marker protein expression in cells treated with compounds **10–13**. The LC3B was determined using the LC3B II/I ratio, which was then standardized against the control group (non-treated group). The p62 protein level was detected, and the expression level was normalized by dividing the protein level of the control group. (CQ: chloroquine treatment group at 25 μM, Rapa: rapamycin treatment group at 0.25 μM). (B) The GFP-mRFP-tagged LC3 punctuates images with the treatment of compounds. The autophagy-regulated compounds **10–13** were treated in HEK293 cells and the LC3 puncta were detected by using confocal microscopy.

Although guanidines have not been frequently found in natural products, especially plant materials, in contrast to other types of alkaloids, their substantial hydrophilicity makes them suitable for the drug development, and some have even been used clinically, such as streptomycin (Berlinck et al., 2021). Since the guanidine

functional group has limited occurrence in nature, guanidine-bearing molecules are structurally unique and often exhibit potent biological activities (Berlinck et al., 2012). By using HRESI-MS/MS-based molecular networking, the interesting cluster of guanidines from *A. rugosa* leaves has been quickly investigated in order to isolate new natural guanidines. In addition to biological effects that have been reported previously, the new finding of autophagy inhibition activities in plant guanidines (**10**–**13**) contributes to the potential activities of natural products in general and guanidine derivatives in particular. Notably, guanidine derivatives have been reported to show autophagy regulation, both induction and inhibition, as exemplified by metformin and DBeQ, respectively. Our results suggested that natural guanidines (**10**–**13**) exhibited potential autophagy inhibition activities which are promising for targeted therapy in cancer (Xie et al., 2021) and also for the treatment of muscle atrophy (Sartori et al., 2021a).

6. Experimental Section

6.1. Materials

6.1.1. Plant material

Leaves of the plant material were collected in Quang Trung commune, Ngoc Lac district, Thanh Hoa Province, Vietnam, in May 2016. Samples were identified as *Alchornea rugosa* (Lour.) Müll.Arg. based on morphological characteristics by Dr. P. T.Thuong, Division of Herbal Products, Vietnam-Korea Institute of Science and Technology (VKIST), Hanoi, Vietnam. A voucher specimen was deposited in the Herbarium of VKIST under accession number VKIST-HP-01.

6.1.2. Chemicals and reagents

Column chromatography (CC) was performed by using Diaion HP-20 resin (250–850 μm particle size, Merck, Germany), and medium-pressure liquid chromatography (MPLC) was carried out with a Biotage Isolera One MPLC system equipped with an RP-C$_{18}$ column (120 g, Grace), Sephadex LH-20 (Sigma-Aldrich, Inc., St. Louis, MO, USA), and RP-C$_{18}$ (40–63 μm particle size, Merck, Germany). TLC silica gel 60 RP-C$_{18}$ F254 plates (Merck, Darmstadt, Germany) were used for thin-layer chromatography. Industrial grades of MeOH, EtOH, and EtOAc were used for extraction, partition, and fractionation steps. Analytical grades of MeOH, MeCN, formic acid (FA), and trifluoroacetic acid (TFA) were used for isolation. All solvents were purchased from DaeJung Pure Chemical Engineering Co Ltd. (Siheung, Republic of Korea). Deuterated NMR solvents and dimethyl sulfoxide (DMSO) were purchased from BK Instuments Inc. (Daejeon, Republic of Korea).

6.1.3. Equipment and software

Optical rotations were recorded on a JASCO P-2000 polarimeter (JASCO International Co. Ltd., Tokyo, Japan) with a 10 mm path length cell at a specified temperature. UV and ECD spectra were obtained from a ChirascanPlus CD spectrometer (Applied Photophysics, Leatherhead, UK). IR spectra were measured on a JASCO FT/IR-4200 spectrometer (Thermo Electron Corp., Waltham, MA, USA). NMR data were measured by using AVANCE-500 MHz (Bruker, Rheinstetten, Germany) and JNM-ECA-600 MHz (JEOL, Tokyo, Japan) spectrometers coupled with a 5 mm cryoprobe (Bruker, Germany). NMR spectra were processed on

MestReNova software (Mastrelab Research, USA). HRESIMS/MS data were collected by an UHPLC-ESI-qTOF-MS system with a Waters Xevo G2 QTOF mass spectrometer (Waters MS Technologies, UK), a Water BEH C_{18} (100×2.1 mm, 1.7 μm, Water, USA), and analyzed by MassLynx SCN 855 software. The preparative HPLC system used for isolation included a Gilson 321 pump and Gilson UV/VIS 151 detector (Gilson Inc., USA), OptimaPak C_{18} column (10 × 250 mm, 5 μm particle size, RS Tech, Seoul, Korea), and a YMC-Triart Phenyl column (250 × 10.0 mm i.d., 5 μm particle size, YMC Co., Ltd.). The conformational search was performed using CONFLEX 7 Rev. (CONFLEX Corporation, Japan); ECD calculations were done by TURBOMOLE 3.4 (COSMOlogic GmbH & Co. KG, Germany). Samples preparation for HRESI MS/MS analyses and separation used sonicator Power Sonic 420 (Hwashin Tech Co., Ltd., Republic of Korea), evaporator EYELA KSB-202 (Eyela Singapore Pte. Ltd., Singapore), centrifuge MICRO17TR (Hanil Scientific Inc., Korea). Biological experiments used autoclave LAC-5080S (DAIHAN LABOTECH Co. Ltd., Korea), clean bench Class II Biological Safety Cabinet (Esco Technologies, Inc., USA), microplate reader VersaMax (Molecular Devices, USA)

6.2. Extraction and isolation

Dry leaves of *A. rugosa* (10.0 kg) were extracted with EtOH at room temperature for 24 h (3 × 20 L) to obtain the EtOH extract (300 g) after freeze-drying. This extract was subsequently suspended in distilled H_2O and applied to a Diaion HP-20 column (15 × 65 cm i.d.). The extract was eluted with H_2O, 25%, 50%, 100% MeOH, and acetone. The 25% MeOH fraction (AR25, 40 g) was partitioned between EtOAc and H_2O. The organic layers were combined to obtain 15.0 g of the EtOAc fraction. The aqueous layer (18.0 g) was concentrated and loaded on an MPLC with a 10–50% MeOH gradient system as a mobile phase, and 8 fractions (M1–M8) were obtained. Fraction M2 (5.0 g) was separated by size-exclusion chromatography using a Sephadex LH-20 (Φ 5.0×80.0 cm i.d.) and eluted with 100% MeOH. Three subfractions (L1–L3) were obtained. L2 (4.0 g) was repeatedly separated by MPLC using a gradient of 20–40% MeOH, which yielded eight subfractions (M1–M8). Fraction M2 (1.0 g) was again divided into seven fractions using a Sephadex LH-20 (Φ 2.5×75.0 cm i.d.) by eluting with 100% MeOH. Fraction L2.2.2 (300.0 mg) was separated by HPLC with an isocratic mobile phase of 15% MeCN/H_2O (0.1% TFA)

to obtain a mixture of compounds **1**–**2** (12.0 mg), and compound **12** (31.6 mg). Fraction L2.2.5 (128.0 mg) was separated by HPLC using an isocratic phase of 17% MeCN/H_2O (0.1% TFA) to yield a mixture of compounds **3**–**4** (6.0 mg), and **11** (18.0 mg). Similarly, fraction L2.2.7 (90 mg) was subjected to HPLC using 15% MeCN/H_2O (0.1% TFA) to obtain compounds **5**–**6** (4.0 mg), and **8** (12.5 mg). These mixtures were further separated by an HPLC Gilson system equipped with a YMC-Triart Phenyl column using MeOH/H_2O (0.1% TFA) to yield **1** (5.0 mg) and **2** (4.0 mg) [from a gradient of 48−50% MeOH/ H_2O (0.1% TFA)], **3** (2.0 mg) and **4** (2.5 mg) [from an isocratic phase of 38% MeOH/H_2O (0.1% TFA)], and **5** (1.0 mg) and **6** (1.0 mg) [from an isocratic phase of 36% MeOH/H_2O (0.1% TFA)]. Purification of sub-fraction L2.5 using a semipreparative HPLC and a 5μm Optimapak column (15% MeCN/H_2O (0.1% TFA)) isocratic, flow rate: 3.0 mL/min, UV detector at two wavelengths of 201 and 280 nm) obtained compound **13** (25.0 mg). With the same condition on the semipreparative HPLC, fraction L2.6 yielded compounds **7** (31.8 mg) and **10** (12.0 mg). Fraction M5 (3.5 g) was chromatographed on a Sephadex LH-20 (Φ 5.0×80.0 cm i.d.) eluted with 100% MeOH to afford six sub-fractions (L1–L6). Fraction M5.2 afforded 4.5 mg of compound **14** after isolating with 35% MeOH/H_2O (0.1% TFA) and a flow rate of 3 mL/min on a semipreparative HPLC equipped with a 5μm Optimapak column. Fraction M5.5 was applied to the HPLC system using5μm Optimapak column eluted with 35% MeOH/H_2O (0.1% TFA) at a flow rate of 3 mL/min to give compound **9** (4.0 mg). Fraction M8 (2.1 g) was applied to an LH-20 column to yield compound **15** (14.1 mg).
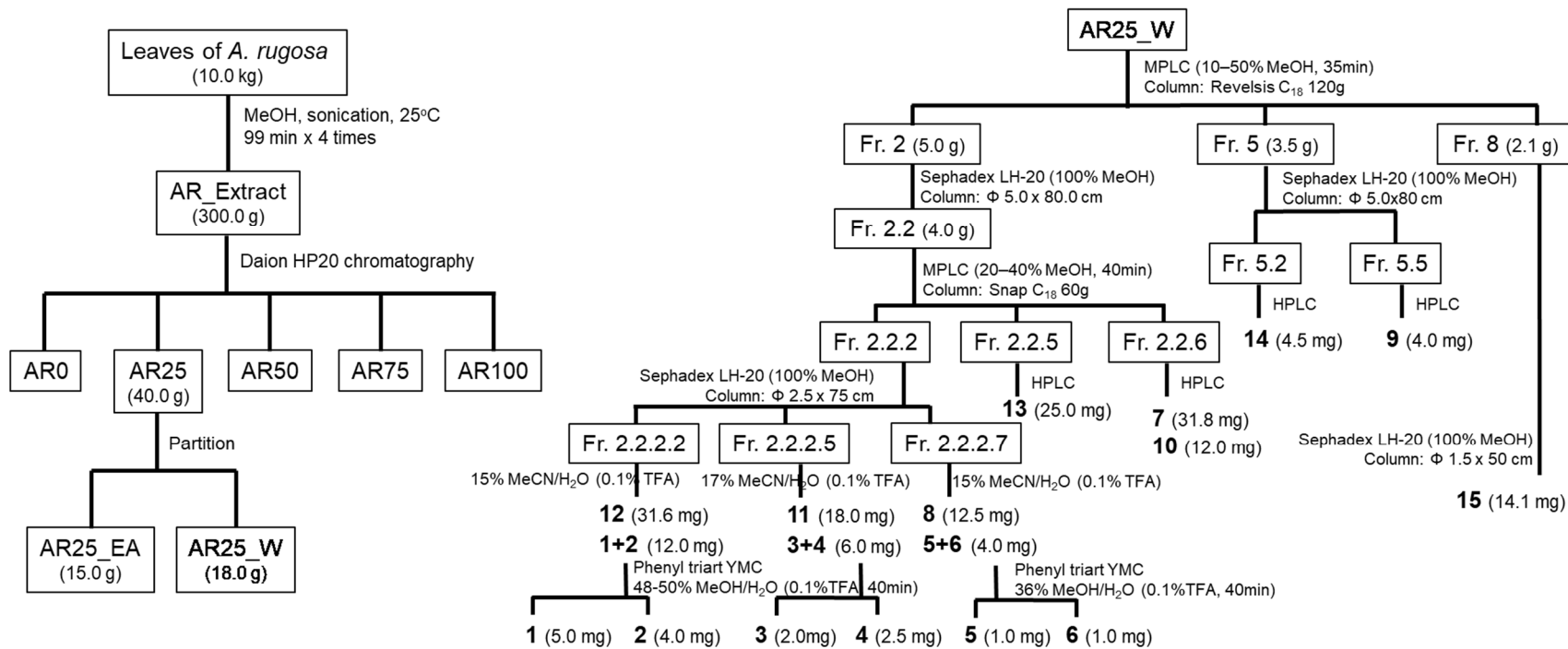
Scheme 1. Isolation scheme of compounds **1–15** from *A. rugosa* leaves

6.3. Physical and chemical characteristics of isolated compounds

6.3.1. Compound **1**

Rugonidine A (**1**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +30 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 230 (0.34), 280 (0.23);

IR (KBr) ν_{max} 3276, 3217, 2979, 2925, 2373, 2352, 2321, 2309, 1706,1697,
1617, 1648, 1637, 1523, 1454, 1314, 1201, 1135, 1046, 1017, 946, 838, 800,
719 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3;

HRESIMS *m/z* 644.2833 [M + H]$^+$, (calcd for $C_{32}H_{42}N_3O_{11}$, 644.2819).

6.3.2. Compound **2**

Rugonidine B (**2**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ −17 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 237 (0.34), 280 (0.23);

IR (KBr) ν_{max} 3276, 3222, 2979, 2925, 2854, 2373, 2352, 2322, 2341, 1747,
1706, 1680, 1627, 1523, 1454, 1316, 1203, 1136, 1046, 800, 724 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3; HRESIMS *m/z* 644.2827 [M + H]$^+$,
(calcd for $C_{32}H_{42}N_3O_{11}$, 644.2819).

6.3.3. Compound **3**

Rugonidine C (**3**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +20 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 215 (0.83), 281 (0.65);

IR (KBr) ν_{max} 3244, 2929, 1706, 1679, 1626, 1527, 1456, 1442, 1201, 1136,
1046, 1019, 801, 721 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3;

HRESIMS *m/z* 576.2201 [M + H]$^+$ (calcd for $C_{27}H_{34}N_3O_{11}$, 576.2193).

6.3.4. Compound **4**

Rugonidine D (**4**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ −5 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 215 (0.03), 281 (0.63);

IR (KBr) ν_{max} 3313, 3249, 2925, 2852, 2363, 2350, 2310, 1747, 1706, 1689, 1647, 1625, 1611, 1543, 1523, 1458, 1201, 1135, 1114, 1045, 841, 803, 721 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3;

HRESIMS *m/z* 576.2201 [M + H]$^+$ (calcd for $C_{27}H_{34}N_3O_{11}$, 576.2193).

6.3.5. Compound **5**

Rugonidine E (**5**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +2 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 215 (0.95), 281 (0.76);

IR (KBr) ν_{max} 3243, 2962, 2924, 2848, 1706, 1679, 1627, 1524, 1456, 1199, 1137, 1043, 843, 801, 723 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3;

HRESIMS *m/z* 430.1635 [M + H]$^+$ (calcd. for $C_{21}H_{24}N_3O_7$, 430.1614).

6.3.6. Compound **6**

Rugonidine F (**6**)

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ −4 (*c* 0.2, MeOH);

UV(MeOH) λ_{max} (log ε) 213 (0.74), 280 (0.63);

IR (KBr) ν_{max} 3276, 3222, 2979, 2925, 2854, 2373, 2352, 2322, 2341, 1747, 1076, 1680, 1627, 1524, 1454, 1316, 1203, 1136, 1046, 800, 724 cm^{-1};

^{13}C and ^{1}H NMR data, see Tables 2 and 3;

HRESIMS *m/z* 430.1609 [M + H]$^+$, (calcd. for $C_{21}H_{24}N_3O_7$, 430.1614).

Table 2. ^{1}H NMR data for **1–6** from *A. rugosa* (δ_H in ppm and *J* in Hz)

No.	**1**[a]	**2**[b]	**3**[b,*]	**4**[a,*]	**5**[a]	**6**[a]
2	4.63, d (8.1)	4.64, d (6.1)	4.70, d (8.1)	4.65, d (8.1)	4.49, d (7.1)	4.49, d (7.1)
3	3.76, td (8.1, 7.9, 5.6)	3.82, m	3.90, m	3.90, qd (8.1, 5.6, 2.3)	3.50, m	3.50, m
4β	2.78, dd (5.6, 16.1)	2.82, dd (5.6, 16.0)	2.94, dt (5.8, 16.0)	2.94, ddd (16.1, 7.7, 5.6)	2.73, dd (16.1, 4.2)	2.73, dd (16.1, 4.2)
4α	2.53, m	2.54, m	2.67, dd (16.0, 8.9)	2.67, dd (16.1, 9.0)	2.39, dd (16.1, 8.3)	2.39, dd (16.1, 8.3)
6	5.96, s	5.97, d (2.4)	5.97, br s	5.97, s	5.94, s	5.94, s
2'	6.79, d (1.9)	6.80, d (2.0)	6.86, d (2.0)	6.88, d (2.0)	6.78, br s	6.78, br s
5'	6.73, d (8.2)	6.73, d (8.2)	6.79, d (8.2)	6.79, d (8.2)	6.71, m	6.71, m
6'	6.66, dd (1.9, 8.2)	6.66, d (8.2)	6.74, dd (8.2, 2.0)	6.74, dd (8.2, 2.0)	6.65, d (7.0)	6.65, d (7.0)
1"a	3.95, d (6.9)	3.96, dd (8.3, 11.9)	3.93, m	3.92, m	3.90, d (11.0)	3.90, d (11.0)
1"b	3.85, dd (6.9, 11.0)	3.86, dd (6.0, 8.3)	3.79, dd (11.9, 5.9)	3.79, dd (11.9, 6.9)	3.74, d (11.7)	3.74, d (11.7)
4"	1.13, s	1.12, s	1.52, s	1.52, s	1.48, s	1.48, s
5"	1.51, s	1.51, s	1.61, s	1.61, s	1.11, s	1.11, s
1'''a	3.89, dd (6.5, 8.6)	3.88, d (6.2)				
1'''b	3.82, dd (6.0, 8.6)	3.82, m				
2'''	5.23, d (6.5)	5.23, d (6.2)				
4'''	1.72, s	1.71, s				
5'''	1.67, s	1.67, s				
1''''	4.05, s	4.06, s	4.22, s	4.22, s		
2''''	3.31, m	3.31, t (3.9)	3.33, m	3.53, m		
3''''	3.36, td (3.2, 9.2)	3.35, td (2.9, 9.3)	3.57, d (3.5)	3.57, m		
4''''	3.12, td (3.2, 9.2)	3.12, td (2.9, 9.3)	3.58, dd (9.5, 3.5)	3.58, m		
5''''	3.48, dd (5.9, 9.2)	3.49, td (6.0, 2.9)	3.54, dd (6.1, 3.5)	3.54, m		
6''''	1.11, d (6.4)	1.11, d (6.0)	1.24, d (6.1)	1.24, d (6.2)		
C=N						
1"-NH	9.59, s	9.60, s			9.38, s	9.38, s
1'''-NH	9.08, d (6.0)	9.09, q (6.0)			10.33, s	10.33, s

[a] Recorded at 600 MHz, [b] Recorded at 500 MHz, *Recorded in MeOH

Table 3. ^{13}C NMR data for **1–6** from *A. rugosa* in DMSO-d_6 (δ_C, type)

No.	1[a]	2[b]	3[a]	4[a]	5[a]	6[a]
	δ_C, type	δ_C, type	δ_C, type	δ_C, type	δ_C, type	δ_C, type
2	79.1, CH	79.2, CH	79.1, CH	79.1, CH	81.1, CH	81.1, CH
3	72.6, CH	72.7, CH	72.6, CH	72.6, CH	69.8, CH	69.8, CH
4	27.7, CH$_2$	27.8, CH$_2$	27.6, CH$_2$	27.6, CH$_2$	29.0, CH$_2$	29.0, CH$_2$
5	156.0, C	156.0, C	155.9, C	155.9, C	155.8, C	155.8, C
6	89.3, CH	89.4, CH	89.4, CH	89.4, CH	89.2, CH	89.2, CH
7	155.5, C	155.5, C	155.5, C	155.5, C	155.5, C	155.5, C
8	109.6, C	109.6, C	109.6, C	109.6, C	109.4, C	109.4, C
9	150.5, C	150.5, C	150.5, C	150.5, C	150.7, C	150.7, C
10	101.0, C	101.0, C	100.9, C	100.9, C	101.4, C	101.4, C
1'	129.5, C	129.5, C	129.6, C	129.6, C	130.4, C	130.4, C
2'	114.0, CH	114.0, CH	114.0, CH	114.0, CH	114.5, CH	114.5, CH
3'	145.0, C	145.0, C	145.0, C	145.0, C	144.9, C	144.9, C
4'	145.1, C	145.2, C	145.1, C	145.1, C	144.9, C	144.9, C
5'	115.2, CH	115.2, CH	115.2, CH	115.2, CH	115.1, CH	115.1, CH
6'	118.0, CH	118.0, CH	118.0, CH	118.0, CH	118.2, CH	118.2, CH
1"	45.7, CH$_2$	45.7, CH$_2$	45.6, CH$_2$	45.6, CH$_2$	45.6, CH$_2$	45.6, CH$_2$
2"	108.4, C	108.4, C	108.2, C	108.2, C	108.1, C	108.1, C
3"	44.2, C	44.2, C	44.1, C	44.1, C	44.2, C	44.2, C
4"	24.5, CH$_3$	24.5, CH$_3$	24.5, CH$_3$	24.5, CH$_3$	24.5, CH$_3$	24.5, CH$_3$
5"	20.3, CH$_3$	20.3, CH$_3$	20.3, CH$_3$	20.3, CH$_3$	20.3, CH$_3$	20.3, CH$_3$
1'''	40.1, CH$_2$	40.0, CH$_2$				
2'''	118.6, CH	118.6, CH				
3'''	136.3, C	136.3, C				
4'''	25.3, CH$_3$	25.4, CH$_3$				
5'''	17.9, CH$_3$	17.9, CH$_3$				
1''''	100.1, CH	100.1, CH	100.1, CH	100.1, CH		
2''''	70.1, CH	70.1, CH	70.2, CH	70.2, CH		
3''''	70.6, CH	70.6, CH	70.6, CH	70.6, CH		
4''''	71.8, CH	71.9, CH	71.9, CH	71.9, CH		
5''''	68.9, CH	68.9, CH	68.9, CH	68.9, CH		
6''''	17.8, CH$_3$	17.8, CH$_3$	17.8, CH$_3$	17.8, CH$_3$		
C=N	159.0, C	159.0, C	160.3, C	160.3, C	160.5, C	160.5, C

[a] Recorded at 150 MHz, [b] Recorded at 125 MHz

6.3.7. Compound **7**

(2*R*,3*S*)-Rugonine A (**7**)

Light brown powder;

$[\alpha]_D^{25}$ + 54.9 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 223 (3.28), 282 (2.46);

IR (KBr) ν_{max} 3704, 3237, 2967, 2942, 2869, 1683, 1629, 1461, 1203, 1136, 1054, 1033, 1012, 837, 800, 721 cm^{-1};

^{1}H and ^{13}C NMR (500 and 125 MHz in methanol-*d*$_4$), see Table 4;

HRESIMS *m/z* 628.2859 [M + H]$^+$ (calcd for $C_{32}H_{42}N_3O_{10}$, 628.2865).

6.3.8. Compound **8**

(2*S*,3*R*)-Rugonine B (**8**)

Light brown powder;

$[\alpha]_D^{25}$ –78.0 (*c* 0.1, MeOH);

UV(MeOH) λ_{max} (log ε) 228 (2.27), 280 (1.56);

IR (KBr) ν_{max} 3252, 2977, 2937, 1668, 1613, 1441, 1383, 1279, 1200, 1146, 1072, 983, 841, 718 cm^{-1};

^{1}H and ^{13}C NMR (500 and 125 MHz in methanol-*d*$_4$), see Table 4;

HRESIMS *m/z* 628.2870 [M + H]$^+$ (calcd for $C_{32}H_{42}N_3O_{10}$, 628.2865).

6.3.9. Compound **9**

(2*R*,3*S*)-Rugonine C (**9**)

Light yellow powder;

$[\alpha]_D^{25}$ +119.6 (*c* 0.1, MeOH);

UV(MeOH) λ_{max} (log ε) 228 (2.27), 280 (1.56);

IR (KBr) ν_{max} 3305, 2972, 1676, 1618, 1443, 1384, 1200, 1140, 1076, 1041, 979, 837, 798, 724 cm^{-1};

^{1}H and ^{13}C NMR (800 and 200 MHz in methanol-*d*$_4$), see Table 4;

HRESIMS *m/z* 646.2974 [M + H]$^+$ (calcd for $C_{32}H_{44}N_3O_{11}$, 646.2976).

6.3.10. Compound **10**

(2*R*,3*S*)-Rugonine D (**10**)

Light brown powder;

$[\alpha]_D^{25}$ +27.5 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 225 (2.63), 280 (1.85);

IR (KBr) v_{max} 3705, 3194, 2977, 2934, 2864, 1683, 1613, 1530, 1456, 1196, 1058, 1028, 1009, 826 cm^{-1};

^{1}H and ^{13}C NMR (600 and 150 MHz in DMSO-d_6), see Table 5;

HRESIMS *m/z* 414.1649 [M + H]$^+$ (calcd for $C_{21}H_{24}N_3O_6$, 414.1665)

6.3.11. Compound **11**

(2*R*,3*R*)-Rugonine E (**11**)

Light brown powder

$[\alpha]_D^{25}$ −16.4 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 223 (2.49), 280 (1.70);

IR (KBr) v_{max} 3179, 1682, 1611, 1527, 1446, 1376, 1281, 1200, 1137, 1105, 1047, 1024, 1001, 825, 800, 766, 721 cm^{-1};

^{1}H and ^{13}C NMR (600 and 150 MHz in DMSO-d_6), see Table 5;

HRESIMS *m/z* 414.1675 [M + H]$^+$ (calcd for $C_{21}H_{24}N_3O_6$, 414.1665).

6.3.12. Compound **12**

(2*R*,3*S*)-Rugonine F (**12**)

Light brown powder

$[\alpha]_D^{25}$ +34.1 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 225 (2.63), 280 (1.85);

IR (KBr) v_{max} 3705, 2972, 2923, 1678, 1619, 1205, 1136, 1058, 1033, 1013, 718 cm^{-1};

^{1}H and ^{13}C NMR (600 and 150 MHz in DMSO-d_6), see Table 5;

HRESIMS *m/z* 482.2316 [M + H]$^+$ (calcd for $C_{26}H_{32}N_3O_6$, 482.2291).

6.3.13. Compound **13**

(2*R*,3*S*)-Rugonine G (**13**)

Light brown powder;

$[\alpha]_D^{25}$ +47.5 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 223 (3.28), 282 (2.46);

IR (KBr) ν_{max} 3700, 3346, 2972, 2942, 2869, 2843, 1673, 1619, 1525, 1456, 1196, 1146, 1053, 1032, 1013, 807, 722 cm^{-1};

^{1}H and ^{13}C NMR (600 and 150 MHz in DMSO-d_6), see Table 5;

HRESIMS *m/z* 560.2272 [M + H]$^+$ (calcd for $C_{27}H_{34}N_3O_{10}$, 560.2244).

6.3.14. Compound **14**

(2*R*)-Rugonine H (**14**)

Light brown powder;

$[\alpha]_D^{25}$ +20.3 (*c* 0.2, MeOH);

UV (MeOH) λ_{max} (log ε) 212 (3.08), 280 (2.36);

IR (KBr) ν_{max} 3173, 2967, 1692, 1623, 1461, 1383, 1348, 1284, 1161, 1092, 1023, 998, 821, 763 cm^{-1};

^{1}H and ^{13}C NMR (500 and 125 MHz in DMSO-d_6), see Table 5;

HRESIMS *m/z* 396.1581 [M − H]$^-$ (calcd for $C_{21}H_{22}N_3O_5$, 396.1565).

Table 4. ^{1}H and ^{13}C NMR data for compounds **7–9** from *A. rugosa* in methanol-d_4

Pos.	7[b] δ_H (*J in* Hz)	δ_C	8[a] δ_H (*J in* Hz)	δ_C	9[c] δ_H (*J in* Hz)	δ_C
2	4.71, d (7.3)	81.2	4.67, d (7.3)	81.1	4.70, d (7.3)	81.2
3	3.97, q (7.3)	75.4	3.93, q (7.3)	75.2	3.96, td (5.4, 7.7)	75.3
4	2.86, dd (5.5, 16.5)	27.3	2.83, dd (5.5, 16.3)	27.3	2.86, dd (5.4, 16.2)	27.4
	2.75, dd (7.7, 16.5)		2.69, dd (7.8, 16.3)		2.72, dd (5.4, 16.2)	
5		158.8		157.2		158.7
6	6.10, s	96.0	6.06, s	95.9	6.09, s	95.8
7		157.3		155.7		157.3
8		95.9		95.7		95.8
9		155.5		155.4		157.3
10		100.8		100.7		100.8
1'		131.6		131.5		131.6
2'	6.76, d (2.0)	115.0	6.73, d (2.0)	114.9	6.76, d (2.1)	115.0
3'		146.2		146.1		146.2
4'		146.4		146.3		146.4
5'	6.73, d (7.8)	116.1	6.69, d (8.1)	115.9	6.72, d (8.1)	116.0
6'	6.66, dd (2.3, 8.3)	119.6	6.62, dd (2.0, 8.1)	119.5	6.65, dd (2.1, 8.1)	119.6
1"		115.0		114.9		114.9
2"		131.3		131.2		131.2
3"	2.78, q (7.3)	26.2	2.75, q (7.3)	26.2	2.76, p (7.0)	26.2
4"	1.11, d (7.0)	21.5	1.07, d (7.1)	21.4	1.10, d (7.0)	21.5
5"	1.13, d (7.0)	21.6	1.10, d (7.1)	21.5	1.12, d (7.0)	21.6
C=NH	11.65*	147.6		147.6		147.6
1'''	3.81, d (6.9)	41.9	3.77, d (7.0)	41.8	3.34, d (6.9)	40.4
2'''	5.26, t (6.9)	120.1	5.23, t (7.3)	120.1	1.75, dd (6.9, 8.2)	42.4
3'''		138.5		138.4		70.9
4'''	1.69, s	18.0	1.66, s	17.9	1.23, d (5.0)	29.5
5'''	1.73, s	25.7	1.70, s	25.6	1.23, d (5.0)	29.5
1''''	4.35, d (1.5)	102.0	4.33, s	102.0	4.35, d (1.7)	102.0
2''''	3.54, d (3.4)	72.0	3.51, br s	71.9	3.53, q (1.7)	72.0
3''''	3.57, dd (3.4, 9.4)	72.3	3.55, br s	72.1	3.56, d (9.5)	72.2
4''''	3.33, d (9.4)	73.9	3.28, d (10.0)	73.8	3.33, d (7.4)	73.9
5''''	3.65, dq (6.3, 12.4)	70.3	3.62, dt (3.4, 6.1)	70.3	3.64, dd (6.2, 9.5)	70.3
6''''	1.25, d (6.3)	17.9	1.21, d (6.1)	17.8	1.24, d (6.2)	17.9
1"-NH	11.93*					

[a] ^{1}H and ^{13}C NMR spectra were acquired at 500 and 125 MHz, respectively

[b] ^{1}H and ^{13}C NMR spectra were acquired at 600 and 150 MHz, respectively

[c] ^{1}H and ^{13}C NMR spectra were acquired at 800 and 200 MHz, respectively

* data recorded in DMSO-d_6

Table 5. ^{1}H and ^{13}C NMR data for compounds **10–14** from *A. rugosa* in DMSO-d_6

No.	10[b] δ_H (J in Hz)	10[b] δ_C	11[a] δ_H (J in Hz)	11[a] δ_C	12[b] δ_H (J in Hz)	12[b] δ_C	13[b] δ_H (J in Hz)	13[b] δ_C	14[a] δ_H (J in Hz)	14[a] δ_C
2	4.54, d (6.9)	81.3	4.74, s	78.5	4.71, d (5.8)	81.2	4.74, d (6.7)	78.8	4.80, d (9.9)	76.7
3	3.83, dd (12.6, 6.9)	66.0	3.98, s	64.6	4.01, t-like (6.2)	66.0	3.91, dd (12.7, 6.5)	72.0	2.00, d (7.3); 1.82, q (9.9, 7.3)	28.6
4	2.66, dd (15.8, 4.7); 2.42, dd (6.1, 7.5)	27.4	2.60, dd (4.4, 16.2); 2.46, s	28.2	2.63, dd (12.6, 6.2); 2.41, dd (16.7, 8.0)	17.8	2.67, dd (16.3, 5.5); 2.57, dd (16.1, 7.2)	25.9	2.54, d (10.6)	18.8
5		156.9		157.3		156.7		157.0		156.6
6	6.11, s	94.8	6.07, s	94.8	6.10, s	94.7	6.15, s	95.1	6.07, s	94.6
7		155.6		155.4		155.7		156.0		155.4
8		94.1		94.3		99.1		94.3		95.0
9		153.9		154.4		153.9		153.8		154.8
10		99.2		98.7		99.1		98.8		100.4
1'		130.3		130.3		130.3		129.8		132.5
2'	6.65, s	114.5	6.74, d (2.2)	114.7	6.73, s	114.6	6.66, s	114.1	6.71, s	113.6
3'		144.8		144.5		144.8		145.1		145.0
4'		144.8		144.5		144.8		145.1		144.7
5'	6.64, s	115.0	6.59, d (8.1)	114.8	6.60, d (8.1)	115.1	6.68, s	115.3	6.65, d (7.0)	115.2
6'	6.53, d (7.2)	118.0	6.54, d (7.1)	118.0	6.52, d (8.1)	118.0	6.56, d (8.1)	117.9	6.56, d (7.1)	116.9
1"		112.9		113.0		113.1		112.9		112.7
2"		128.7		128.6		129.1		128.9		128.7
3"	2.62, dd (13.7, 6.7)	24.3	2.62, dd (13.9, 7.0)	24.4	2.67, d (4.6)	24.5	2.63, m	24.5	2.60, m	24.4
4"	1.00, s	21.1	1.00, d (7.0)	21.2	1.01, s	21.2	1.00, d (6.9)	21.3	1.00, d (6.7)	21.3
5"	0.99, s	21.1	0.95, d (7.0)	21.1	1.01, s	21.3	1.03, d (7.0)	21.3	1.01, d (6.7)	21.3
C=NH	12.14, s	146.0	12.15, s	146.0	12.11, s	146.8	12.28, s	146.3		146.9
1"-NH	11.59, s		11.54, s		9.58, s		11.70, s		8.33, br s	
1"'-NH or NH$_2$	7.07, br s		7.09, s		10.34, s		7.17, s		7.46, br s	
1"'					3.81, s	44.1				
2"'					5.21, s	120.3				
3"'						135.0				
4"'					1.53, s	25.3				

Position						
5'''			1.30, s	17.8		
1''''					4.30, s	99.9
2''''					3.36, br s	70.5
3''''					3.35, s	70.7
4''''					3.15, d (9.1)	72.0
5''''					3.47, dq (12.4, 6.1)	69.0
6''''					1.14, d (6.1)	17.9
3-OH		9.52, s				
5- OH	3.96, s	9.30, s	9.26, s			
3'- OH		8.68, br s	8.64, s			
4'- OH		8.80 br s	8.82 s			

[a] ^{1}H and ^{13}C NMR spectra were acquired at 500 and 125 MHz, respectively
[b] ^{1}H and ^{13}C NMR spectra were acquired at 600 and 150 MHz, respectively

6.4. LC-MS/MS-based molecular networking

HRESI-MS/MS data were acquired from a Waters Acquity UPLC system (Waters Co.) equipped with a Waters Xevo G2 QTOF mass spectrometer (Waters Co., Manchester, UK) and electrospray ionization (ESI) interface, operating in positive-ion mode. A Waters Acquity UPLC BEH C_{18} column (150 mm × 2.1 mm, 1.7 μm) was used with a MeCN/H_2O gradient system (10:90 to 90:10) for 14 min. The flow rate was set at 300 μL/min, and the injection volume was 2.0 μL. The temperatures of the autosampler and column oven were maintained at 15 and 40 °C, respectively. Source parameters were set as follows: a capillary voltage of 2.5 kV, cone voltage of 40 V, source temperature of 120 °C, desolvation gas temperature of 350 °C, cone gas flow of 50 L/h, and desolvation gas flow of 800 L/h. The acquisition rate was 0.2 s. Data were centroided during acquisition using an independent reference lock-mass ion via the LockSpray interface to ensure accuracy and precision. Peak picking, chromatogram deconvolution, and other data processing of MS/MS data were performed by MZmine2 software v32 (Pluskal et al., 2010). Eventually, the .mgf preclustered spectral data file and its corresponding .csv metadata file (for RT, areas, and formula integration) were exported using the dedicated "Export for GNPS" and "Export to CSV file" built-in options. A molecular network was created using the online workflow at GNPS (http://gnps.ucsd.edu*)*. The spectra in the network were then searched against GNPS spectral libraries. All matches kept between network spectra and library spectra were required to have a score above 0.65 and at least 4 matched peaks. The molecular networking data were analyzed and visualized using Cytoscape (ver. 3.6.0) (Shannon, 2003). All of the results and parameters can be accessed with the GNPS job id for molecular networking feature-based analysis at http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=cb23e27317a24150abb45fad3fba59a8.

6.5. Absolute configurations for sugar in **1**–**4**, and **7**–**9, 13**

Compounds **1**–**4,** and **7**–**9, 13** (0.5 mg) were hydrolyzed with 2 N HCl (0.5 mL) at 100 °C for 18 h (Chaturvedula & San Miguel, 2012). After cooling, the mixture was evaporated *in vacuo* and diluted with 1 mL of distilled H_2O. After neutralization (pH 7), the reacted mixture was extracted with EtOAc (3 × 1.0 mL) to obtain an aqueous layer that was dried under vacuum. Each hydrolysate of **1**–**4**, and **7**–**9, 13**

93

was then dissolved in anhydrous pyridine (0.25 mL) and heated with L-cysteine methyl ester hydrochloride (2.5 mg) at 60 °C for 1.5 h, and then 0.25 mL of pyridine containing *o*-tolyl isothiocyanate (2.5 mg) was added to the mixture and reacted at 60 °C for an additional 1.5 h. The reaction mixture was analyzed by a UPLC column [YMC C$_{18}$, 4.6 × 250 mm (5 μm); 10–90% MeCN/H$_2$O (0.1% formic acid); UV detection at 250 nm; 1.0 mL/min]. The sugar in **1**–**4**, and **7**–**9, 13** was determined to be L-rhamnose (t$_R$ = 21.26 for compounds **1** and **2** and 21.27 min for compounds **3** and **4,** and **7**–**9, 13**), which was consistent with that of the derivative containing authentic L-rhamnose (t$_R$ = 21.26 min) (Shin et al., 2015; Tanaka et al., 2007). As a result, the rhamnose moiety in compounds **1**–**4**, and **7**–**9, 13** was assigned the L-configuration.

6.6. Computational Methods

Conformational analyses of compounds **I** and **II** were simulated using molecular mechanics force-field (MMFF94s) calculations with a search limit of 1.0 kcal/mol in CONFLEX 7 (Conflex Corp., Tokyo, Japan). The CONFLEX searches gave 15 and 13 stable conformers for compounds **I** and **II**, respectively. The conformers with a Boltzmann population of over 1% were geometrically optimized at the B3LYP/6-31g level in gas by TmoleX 4.3 and Turbomole (COSMOLogic GmbH, Leverkusen, Germany) at the def-SV(P) basis set.

6.6.1. NMR Calculations

^{13}C NMR chemical shifts for these conformers were calculated at the mPW1PW91/6-311+G(2df,2pd) level after geometry optimization. Boltzmann-weighted averages of the chemical shifts were calculated to scale them against the experimental values. The experimental and calculated data were analyzed with the linear correlation coefficients R^2 and the probability DP4+ method for stereoisomers (Grimblat et al., 2019). The parameters a and b of the linear regression $\delta_{calcd} = a\delta_{exptl} + b$; the correlation coefficient, R^2; the mean absolute error (MAE) defined as $\Sigma_n |\delta_{calcd} - \delta_{exptl}|/n$; and the corrected mean absolute error (CMAE), defined as $\Sigma_n |\delta_{corr} - \delta_{exptl}|/n$, where $\delta_{corr} = (\delta_{calcd} - b)/a$ and therefore corrects for systematic errors were presented.

Table 6. Comparison of ^{13}C NMR experimental values for **1** and calculated chemical shifts for **I**

Carbon	1		I		
	δ_{exptl}	δ_{calcd}	δ_{corr}	$\Delta\delta_{calcd}$[a]	$\Delta\delta_{corr}$[b]
2	81.1	83.3	83.0	2.2	1.9
3	69.8	70.7	70.6	0.9	0.8
4	29.0	28.6	29.2	0.4	0.3
5	155.8	157.4	155.9	1.6	0.1
6	89.2	88.4	88.1	0.8	1.1
7	155.5	159.7	158.2	4.2	2.7
8	109.4	112.8	112.0	3.4	2.6
9	150.7	153.4	152.0	2.7	1.3
10	101.4	99.9	99.4	1.5	2.0
1'	130.4	133.9	132.8	3.5	2.4
2'	114.5	111.4	110.6	3.1	3.9
3'	144.9	145.7	144.4	0.8	0.5
4'	144.9	145.3	144.0	0.4	0.9
5'	115.1	112.7	111.9	2.4	3.2
6'	118.2	118.7	117.9	0.5	0.3
1"	45.6	42.7	43.1	2.9	2.5
2"	108.1	106.9	106.3	1.2	1.8
3"	44.2	50.4	50.6	6.2	6.4
4"	20.3	18.7	19.5	1.6	0.8
5"	24.5	22.5	23.2	2.0	1.3
C=N	160.5	150.0	148.7	10.5	11.9

[a]$\Delta\delta_{calcd} = \left| \delta_{calcd} - \delta_{exptl} \right|$; [b]$\Delta\delta_{corr} = \left| \delta_{corr} - \delta_{exptl} \right|$

Table 7. Comparison of ^{13}C NMR experimental values for **2** and calculated chemical shifts for **I**

Carbon	**2**			**I**	
	δ_{exptl}	δ_{calcd}	δ_{corr}	$\Delta\delta_{calcd}$[a]	$\Delta\delta_{corr}$[b]
2	79.2	83.2	83.0	4.0	3.8
3	72.7	70.7	70.6	2.0	2.1
4	27.8	28.6	29.3	0.8	1.5
5	156.0	157.4	155.9	1.4	0.1
6	89.4	88.4	88.1	1.0	1.3
7	155.5	159.7	158.2	4.2	2.7
8	109.6	112.8	112.0	3.2	2.4
9	150.5	153.4	152.0	2.9	1.5
10	101.0	99.9	99.4	1.1	1.6
1'	129.5	133.9	132.8	4.4	3.3
2'	114.0	111.4	110.6	2.6	3.4
3'	145.0	145.7	144.4	0.7	0.6
4'	145.2	145.3	144.0	0.1	1.2
5'	115.2	112.7	111.9	2.5	3.3
6'	118.0	118.7	117.9	0.7	0.1
1"	45.7	42.7	43.1	3.0	2.6
2"	108.4	106.9	106.3	1.5	2.1
3"	44.2	50.4	50.6	6.2	6.4
4"	24.5	18.7	19.5	5.8	5.0
5"	20.3	22.5	23.2	2.2	2.9
C=N	159.0	150.0	148.7	9.0	10.3

[a]$\Delta\delta_{calcd} = \left| \delta_{calcd} - \delta_{exptl} \right|$; [b]$\Delta\delta_{corr} = \left| \delta_{corr} - \delta_{exptl} \right|$

6.6.2. ECD Calculations

The conformers of compound **I** were used for ECD calculations based on the one selected from the NMR calculation. The theoretical computational ECD spectra of the optimized conformers were obtanied by time-dependent density functional theory (TD-DFT) at the B3LYP/def-SV(P) functional level according to the Boltzmann distributions. The calculated ECD data were generated using Gaussian functions for each transition (σ is the width of the band at a height of $1/e$). The overall calculated ECD curve was generated using SpecDis 1.71 (Bruhn et al., 2013a) from dipole-length rotational strengths by applying Gaussian band shapes. The two empirical parameters used by SpecDis to optimize the fit between the experimental and predicted spectra were adjusted as follows: the half-width of the CD bands, σ, was set to 0.16–0.24 eV, and the UV shifts were set to +10 nm to +41 nm.

Table 8. Gibbs free energies and equilibrium populations of low-energy conformers of 2*S*,3*R*,2″*R*-**1** (**I**).

Conformers	In MeOH	
	ΔG^a	P (%)[b]
I-1	0.00	16.71
I-2	0.01	16.36
I-3	-0.14	21.07
I-4	0.10	14.17
I-5	1.08	2.69
I-6	1.18	2.29
I-7	1.22	2.11
I-8	1.21	2.18
I-9	0.99	3.12
I-10	2.00	0.57
I-11	1.14	2.45
I-12	1.22	2.13
I-13	1.19	2.25
I-14	1.08	2.68
I-15	0.35	9.22

[a]B3LYP/6-31g, in kcal/mol. [b]From ΔG values at 298.15K and 1 atm

Table 9. Gibbs free energies[a] and equilibrium populations[b] of low-energy conformers of 2*S*,3*R*,2″*S*-**1** (**II**)

Conformers	In MeOH	
	ΔG	P (%)
II-1	0.00	5.25
II -2	-1.02	29.18
II -3	-0.64	15.43
II -4	0.42	2.59
II -5	-0.60	14.44
II -6	-0.57	13.75
II -7	0.17	3.96
II -8	0.23	3.53
II -9	0.39	2.73
II -10	0.65	1.75
II -11	1.35	0.54
II -12	0.21	3.68
II -13	0.30	3.16

[a]B3LYP/6-31g, in kcal/mol. [b]From ΔG values at 298.15K and 1 atm

6.7. Cell Viability Assay

3T3-L1 cells were cultured and seeded in 96-well plates at 3000 cells/well in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS). After 24 h of incubation, the tested compound was dissolved in DMSO and treated for 24 h. Cell viability was analyzed by (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide (MTT) (Sigma, MO, USA). Then, 20 μL of the MTT solution (2 mg/mL) was added to the wells. The plates were incubated at 37 °C in the dark for the MTT reaction. After 2 h of incubation, the supernatant was removed. The formazan crystals were dissolved in 100 μL of DMSO. The absorbance was obtained at 570 nm using a microplate reader (VersaMaxTM, Randor, PA, USA).

6.8. Differentiation of 3T3-L1 Preadipocytes

3T3-L1 preadipocyte cells were cultured with DMEM media (HyClone, UT, USA) with 10% calf serum, 100 U/mL penicillin, and 100 mg/mL streptomycin (HyClone, UT, USA) in 5% CO_2 at 37 °C. At 48 h postconfluence, differentiation was initiated. The culture media was changed to DMEM with 10% fetal bovine serum (FBS) (HyClone, UT, USA) containing 1 μM dexamethasone (Sigma, MO, USA), 0.52 mM 3-isobutyl-1-methylxanthine (Sigma, MO, USA), and 1 μg/mL insulin (Roche, Germany). After 2 days, the media contents were changed to DMEM with 10% FBS, 1 μg/mL insulin, 100 U/mL penicillin, and 100 mg/mL streptomycin. The medium was exchanged after 2 days in DMEM supplemented with 10% FBS medium. The media was replaced every two days until the induction of adipogenesis.

6.9. Measurement of Glucose Uptake Level

The glucose uptake level was analyzed with a fluorescent derivative of glucose 2-[N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino]-2-deoxyglucose (2-NBDG) (Invitrogen, OR, USA) as previously described with small modifications (Pham, Ha, et al., 2018). The 3T3-L1 cells were seeded into a 96-well plate. After adipocytes were fully differentiated, the media was replaced with glucose-free media with test compounds and with or without 50 μM 2-NBDG for 1 h. After treatment, the cells were washed two times with phosphate-buffered saline (PBS). A 70 μL of 1% Triton X-100 in PBS and 0.1 M K_3PO_4 were added to each well for cell lysis. The 2-NBDG fluorescence signal was obtained by a plate reader (VICTOR X3, PerkinElmer, MA,

USA) with 450 nm excitation and 535 nm emission.

6.10. Cell culture and viability assay for the HEK293 cell line

A Green Fluorescent Protein - Light Chain 3 (GFP-LC3) transfected stable HEK293 cell line was generously contributed by Professor Junsoo Park (Yonsei University, Republic Korea), while MCF-7 and HEK293 cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA). The cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Welgne, Gyeongsan-si, Republic of Korea) supplemented with 10% fetal bovine serum (FBS; Gibco Island, NY, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin at 37 °C and 5% CO_2. The cells were maintained at 70% to 80% confluency and subcultured every 2 days. The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) method was applied to measure cell viability. The cells were seeded into 96-well plates (1 × 10^4 cells per well) and incubated for 24 h. After cells attached to the bottom of the well plate, the cells were rinsed with phosphate-buffered saline (PBS) and treated with 20 μM of the tested compounds **7–15**. After 24 h of treatment, 20 μL of MTT solution (2 mg/mL) was added to each well. The MTT solution was incubated in the incubator for 4 h at 37 °C before being discarded. A formazan crystal was dissolved in 100 μL DMSO to measure the absorbance at 570 nm of each well. The analysis was performed to assess the % of cell viability between the control and treatment groups. Autophagosome formation was confirmed in GFP-LC3 stable HEK 293 cells. The cells were seeded into a 6-well plate with coverslips and incubated for 24 h. The cells were treated with a noncytotoxic concentration of isolated compounds. After 24 h of treatment, the cells were washed three times with the PBS, then fixed with 4% paraformaldehyde, and coverslips were carefully lifted to prepare the sliding glass for microscopic analysis.

6.11. Protein expression analysis by Western blot

MCF-7 cells were seeded onto well plates and incubated for 24 h. Isolated compounds that showed enhanced formation of autophagosomes were treated for 24 h, and the cells were harvested. The cells were lysed with 150 μL of lysis buffer [120 mM NaCl, 40 mM Tris (pH 8), and 0.1% NP40 (Nonidet P-40)] and centrifuged at 10,000 rpm for 15 min. The supernatants obtained from the lysates were used as the

protein extract. Protein concentrations were calculated using the BCA protein assay method (Bio-Rad, CA, USA). The protein extracts were diluted with 5X sample buffer and boiled at 95 °C. Thirty-microgram protein samples were injected into each well and electrophoresed on 6 to 15 % gradient SDS-polyacrylamide gels. Protein lanes were transferred to polyvinylidene fluoride (PVDF) membranes, which were then incubated with primary antibodies (LC3B, p62, β-actin) at 4 °C for 24 h. The membranes were further incubated with secondary anti-mouse or anti-rabbit antibodies. Finally, the protein bands were detected using enhanced chemiluminescence Western blotting detection ECL buffer. Protein bands were quantified by ImageJ software.

6.12. Confocal microscopy images

HEK293 cells were seeded on sterilized glass coverslips and incubated for 24 h. GFP-mRFP-LC3 (ptfLC3) plasmid, which was kindly provided by Yonsei University (Wonju, Republic of Korea), was applied to cells for transfection using Lipofectamine 2000 (Thermo Fisher Scientific, Waltham, MA, USA). After transfection, the cells were treated with the compounds and incubated for 24 h. The cells were fixed with 4% paraformaldehyde. After fixation, the coverslips were washed three times with PBS and stained with DAPI solution. Images of the autophagic flux induced by the test compounds were obtained with a Confocal Scope TCS8 (LEICA, Wetzlar, Germany).

6.13. Statistical Analysis

Graphpad Prism 5 (GraphPad Sofware, La Jolla, CA, USA) was used for statistical analysis. Data are expressed as the means ± standard deviations (SDs) of three independent experiments and were calculated and plotted by one-way analysis of variance (ANOVA). Statistically significant differences are indicated as follows: [*]$p < 0.05$, [**]$p < 0.01$, and [***]$p < 0.001$

7. Conclusions

By using the HRESI-MS/MS-based molecular networking, the interesting cluster of guanidines from *A. rugosa* leaves has been quickly investigated for targeting the isolation of the new natural guanidines. Besides biological effects that have been reported previously, the new finding of glucose uptake of 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amines (**1**–**3**), as well as autophagy inhibition activities from plant guanidines (**10**–**13**), contributes to potential activities of natural products in general and guanidines derivatives in particular. It is noteworthy that guanidine derivatives have been reported to show autophagy regulation effects in both induction and inhibition manners by metformin and DBeQ, respectively. Our results suggested that natural guanidines (**10**–**13**) exhibited potential autophagy inhibition activities which are promising for targeted therapy in cancer (Xie et al., 2021) and also for the treatment of muscle atrophy (Sartori et al., 2021a). In summary, the application of molecular networking allowed us to investigate six new guanidine alkaloids of 1,6-dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amine scaffold (rugonidines A–F) and eight new guanidine-fused flavan derivatives (rugonines A–H) from the leaves of *A. rugosa*. The structures of all isolated compounds have been comprehensively elucidated by NMR, HEESIMS data analyses, and ECD and ^{13}C NMR quantum calculations. These compounds consequently were tested and compounds **1**–**3** showed potential effects in stimulating glucose uptake, while compounds **10**–**13** exhibited potential autophagy inhibition activities in HEK293 cells. The results suggested the advantage of HRESI-qTOF-MS/MS-based molecular networking for targeting the isolation of guanidine derivatives from *A. rugosa* as well as their inhibition activities on autophagy in HEK 293 cells that could be promising for future drug discovery of cancer and muscle atrophy treatment.

Part 2: Alkaloids from *Persea americana* and their SIRT1 activities

1. Introduction

1.1. Study background

1.1.1. *Persea americana* Mill. (Lauraceae)

Persea americana Mill. (Lauraceae), commonly known as avocado, is a native tree in Mesoamerica and Central America and is now cultivated worldwide in tropical and subtropical regions (Yao & Xu, 2021). *P. americana* has eight well identified varieties that related to geographical ecotypes, of which the commercial avocado is made up of three that are referred to the Mexican (*P. americana* var. drymifolia), the West Indian (*P. americana* var. americana), and the Guatemalan (*P. americana* var. guatemalensis) (Ibarra-Laclette et al., 2015). The selection of these three races and their hybrids, most frequently Mexican×Guatemalan or "Hass avocado", are the foundation of commercial avocado production (Chanderbali et al., 2008). The size and shape of an avocado's fruit and seed can vary greatly depending on the variety. The Hass variety's fruits have an ovoid to pear form, and when they are fully ripe, they have tough, leathery skin that is dark brown or black. Additionally, compared to most other common types, it has a smaller seed (Dabas et al., 2013). Due to the nutritional compositions and biological activities of avocado pulp, which are rich in grease, the fruit has gained considerable popularity in recent years (Dreher & Davenport, 2013). Its health effects have been indicated through biological studies, observation and clinical trial including reduce risks of cardiovascular disease, overweight or obese, alleviate cognitive function, and activate colonic microbiota system (Dreher et al., 2021). The high contain of monounsaturated fatty acid in avocado mesocarp has been suggested to reduce risk factors for cardiovascular disease (Pieterse et al., 2005). The pulp also includes a variety of bioactive phytochemicals, such as carotenoids (lutein, zeaxanthin, carotene, and cryptoxanthin), B vitamins, vitamins C and E, D-mannoheptulose, sitosterol, and persenone A and B, which have shown to have antifungal, anticancer, and antioxidant properties (Lu et al., 2005).

In addition to the edible portion of avocados, large amounts of byproducts, including peels and seeds, are produced, which could affect the environment (Rodríguez-Carpena et al., 2011). According to ethnopharmacology, the Aztec and

Mayan societies used avocado seed decoctions to cure parasite and mycotic ailments. (Dabas et al., 2013). Additionally, seed formulations have been suggested for use in the treatment of diabetes and digestive disorders, snakebite, as well as abortive agent and contraceptive (Duke & Martinez, 1994). It is said that placing a portion of the seed or the decoction into a tooth cavity may ease toothache. Dandruff is supposedly treated with avocado seed powder, and skin eruptions are occasionally treated with seed oil (Dabas et al., 2013). Avocado seed paste used topically has been used to treat arthritis, and the Nigerian added to soups or pudding to manage hypertension (Ozolua et al., 2010). The pulverized seeds were traditionally ingested in African medicine to alleviate diarrhea and wheezing (Dabas et al., 2013). The current bioactive effects of avocado seeds have been investigated include anticancer (Lee et al., 2008), anti-inflammatory (Rosenblat et al., 2011), antidiabetic (Edem et al., 2009), hypocholesterolemic (Pahua-Ramos et al., 2012), antimicrobial (Raymond Chia & Dykes, 2010), insecticidal (Leite et al., 2009) and dermatological uses (Rosenblat et al., 2011). Avocado seed phytochemical studies have identified a variety of natural products, including procyanidins (Wang et al., 2010), acetogenins (Rodríguez-Sánchez et al., 2013), fatty acids, amines (Arlene et al., 2015), alkaloids (Oboh et al., 2016), phenolic compounds (Saavedra et al., 2017; Segovia et al., 2016), triterpenoids (Abubakar et al., 2017) and pigments (Hatzakis et al., 2019). The presence of C_7 sugars, such as mannoheptulose and perseitol, as the primary phloem transported sugars and as respiratory substrates makes avocados distinct from other fruit (X. Liu et al., 1999, 2002). These C7 sugars have been identified as the "tree factor," which slows down the ripening of fruit while it ripens on the tree (Bertling & Bower, 2005) and may be related to the variations in the postharvest ripening speed of fruit (Landahl et al., 2009).

Although the avocado seed accounts for up to 16% of the total weight of the fruit, has a diverse phytochemical profile, and a long history of ethnobotanical use, it is largely regarded as a waste product and underutilized resource (Dabas et al., 2013; Otaigbe et al., 2016), modern scientific studies into phytochemicals and avocado seed bioactivities is still in early stages.

Figure 85. *Persea americana* Mill. (Hass avocado)

1.1.2. Molecular networking

Molecular networking was first introduced in 2012 (Watrous et al., 2012) and has become increasingly popular as a dereplication strategy in the natural products field. Molecular networking that was created on the GNPS platform (https://gnps.ucsd.edu/, Wang et al., 2016a) can analyze MS/MS data against GNPS libraries to match either putative analogs of known compounds or those that belong to the same class. Moreover, based on the similarities of fragmentation patterns, nodes that shared similar MS/MS profiles were grouped into the same cluster and connected via the edges (cosine score) (Yang et al., 2013). Therefore, molecular networking has been applied to natural products not only as a dereplication technique but also for the discovery of a variety of small molecules (Ryu et al., 2021).

Recent studies that involved avocado have applied HR-MS/MS-based techniques such as quantification or metabolomics to investigate the chemical profile of avocado pulp, peel, and seeds. Analyzing the MS/MS fragmentations allowed to dereplicate flavonoids and phenolic components from all parts of avocado including the mesocarp (Di Stefano et al., 2017; Hurtado-Fernández et al., 2011), peels (Figueroa et al., 2018), and seeds (Do et al., 2022; Younis et al., 2022). This work also could be done by a molecular networking-based dereplication strategy that matches reported compounds against GNPS libraries, as well as double checking with avocado-related studies. By providing crucial information on similarities

between the discovered peaks, this method gives effective and efficient ways to explore phytochemicals, expediting the discovery of new metabolites.

1.1.3. SIRT1

Sirtuin 1 (SIRT1) is a nicotinamide adenosine dinucleotide (NAD)-dependent deacetylase that belongs to the mammalian sirtuin family and plays important roles in cellular and organismal processes, including metabolism and aging (Hall et al., 2013; Rahman & Islam, 2011; Xu et al., 2020). Through the deacetylation of various substrates, including histone proteins (acetylated histones H4K16 and H3K56) and nonhistone targets (p53) to generate nicotinamide and the acetyl group, SIRT1 is involved in a broad range of physiological functions, including controlling gene expression, metabolism, and aging (Rahman & Islam, 2011). The function of SIRT1 in senescence, which extends life spans, ameliorates cellular senescence, and consequently prevents aging-related diseases, is mainly achieved by catalyzing the deacetylation of various downstream transcription factors (Brooks & Gu, 2009). Among these factors, p53 is a crucial factor that contributes to cellular senescence through the regulation of cell cycle arrest in response to DNA damage (Yi & Luo, 2010).
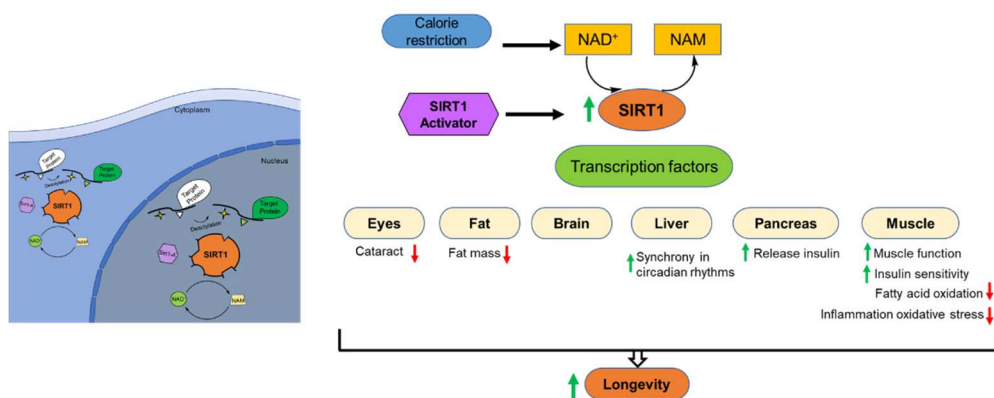
Figure 86. Role of Sirt1 in mammal aging and longevity regulation. Sirt1 may establish a connection between the nutritional condition of the cell and the control of its metabolism by decoding NAD+ changes in special tissues. Sirt1 can alter the patterns of gene expression in target organs (brain, liver, fat, pancreatic beta cells, and muscle) via controlling transcriptional co-regulators like PGC-1 or by directly interacting with transcription factors.

SIRT1 activity can be regulated through NAD^+/NADH and nicotinamide by the SIRT1 protein level and by phosphorylation (Haigis & Sinclair, 2010). An elevated NAD^+/NADH ratio may promote cellular health through mechanisms related to SIRT1 activation (Jang et al., 2012).

1.2. Purpose of Research

In this study, we aimed to apply UPLC qTOF-MS/MS-based molecular networking analysis to navigate the chemical profile of secondary metabolites from avocado seeds and to target unknown compounds for further isolation. The isolated compounds were extensively determined using NMR spectroscopy, chemical reactions, and ECD calculations. Then, all isolated compounds were tested for their effects on SIRT1 activity in a p53-mediated transcriptional luciferase reporter cell-based assay.

2. Targeted isolation of alkaloids from avocado seeds

2.1. Targeted isolation of alkaloids from *P. americana* by using LC-MS/MS-based molecular networking

The avocado seeds were extracted with *n*-hexane solvent to remove fat, and the residue was re-extracted three times with 70% EtOH. The combined extracts were dried and suspended in acetone to precipitate the perseitol sugars. Then, the supernatant was dried under a vacuum and partitioned with EtOAc, *n*-BuOH, and water. The UPLC-qTOF-MS/MS data of fractions n-hexane, EtOAc, and *n*-BuOH (Figure 87) were processed by Mzmine2 (Pluskal et al., 2010) and applied to the FBMN-GNPS platform (hosted by https://gnps.ucsd.edu, (Wang et al., 2016a)) to generate molecular networking. The result was visualized through the Cytoscape ver. 3.6.0 (Shannon, 2003). The dereplication of secondary metabolites from avocado seeds was identified against GNPS libraries and by comparison with compounds reported in the *P. americana* (Scifinder) databases (Figure 88, Table 10). As shown in Figure 88, cluster A displayed even *m/z* nodes of 492, 496, 512, 520, and 522, which are typical characteristics of alkaloids. Therefore, experiments were performed to further isolate these peaks and determine their structures. As a result, nine new alkaloids were isolated (Figure 89), and their structures were elucidated by NMR spectroscopy in combination with chemical reactions and ECD calculations.
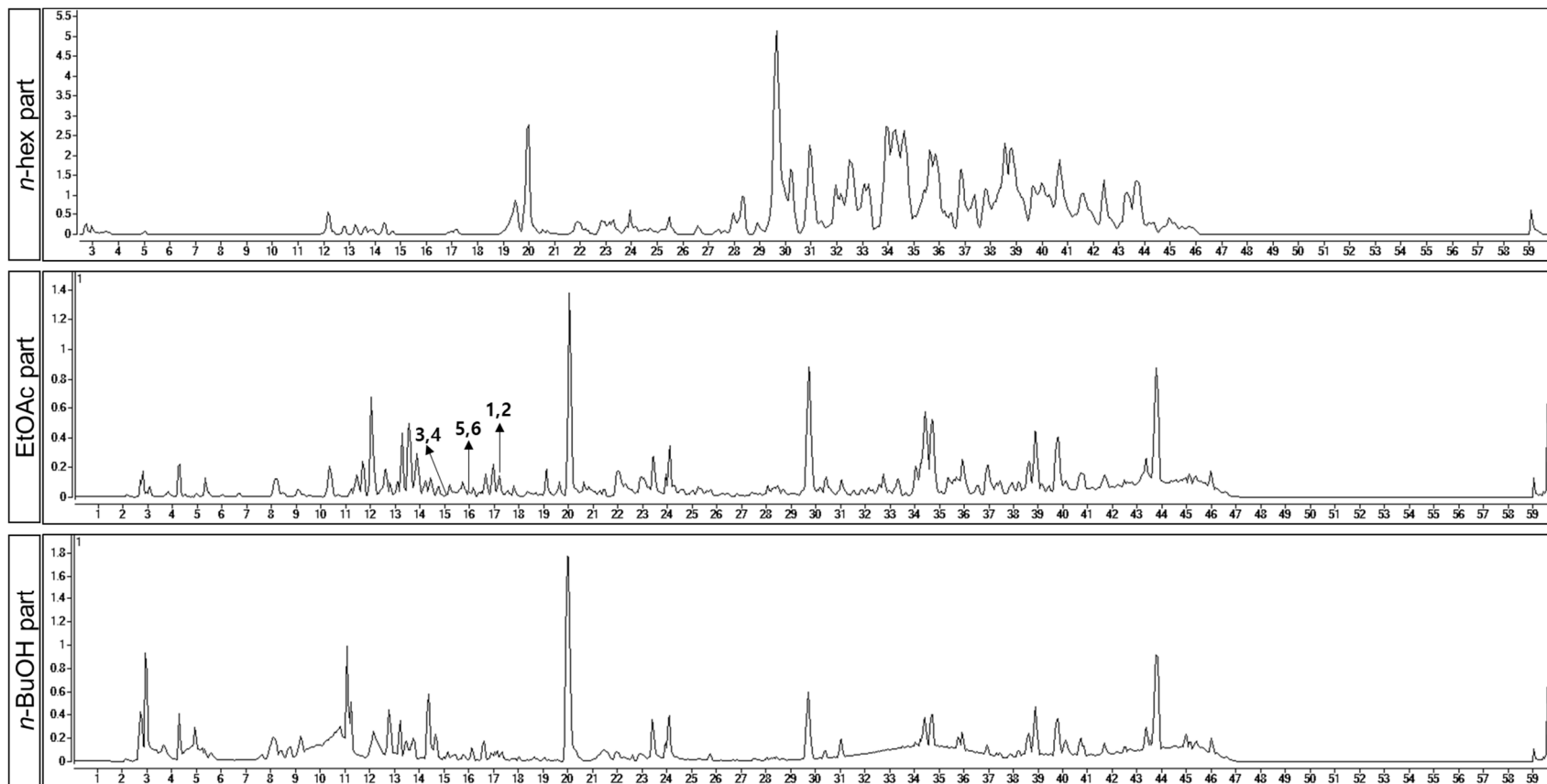
Figure 87. Total ion chromatography (TIC) of *n*-hexane, EtOAc, and *n*-BuOH fractions of avocado seeds in positive mode.
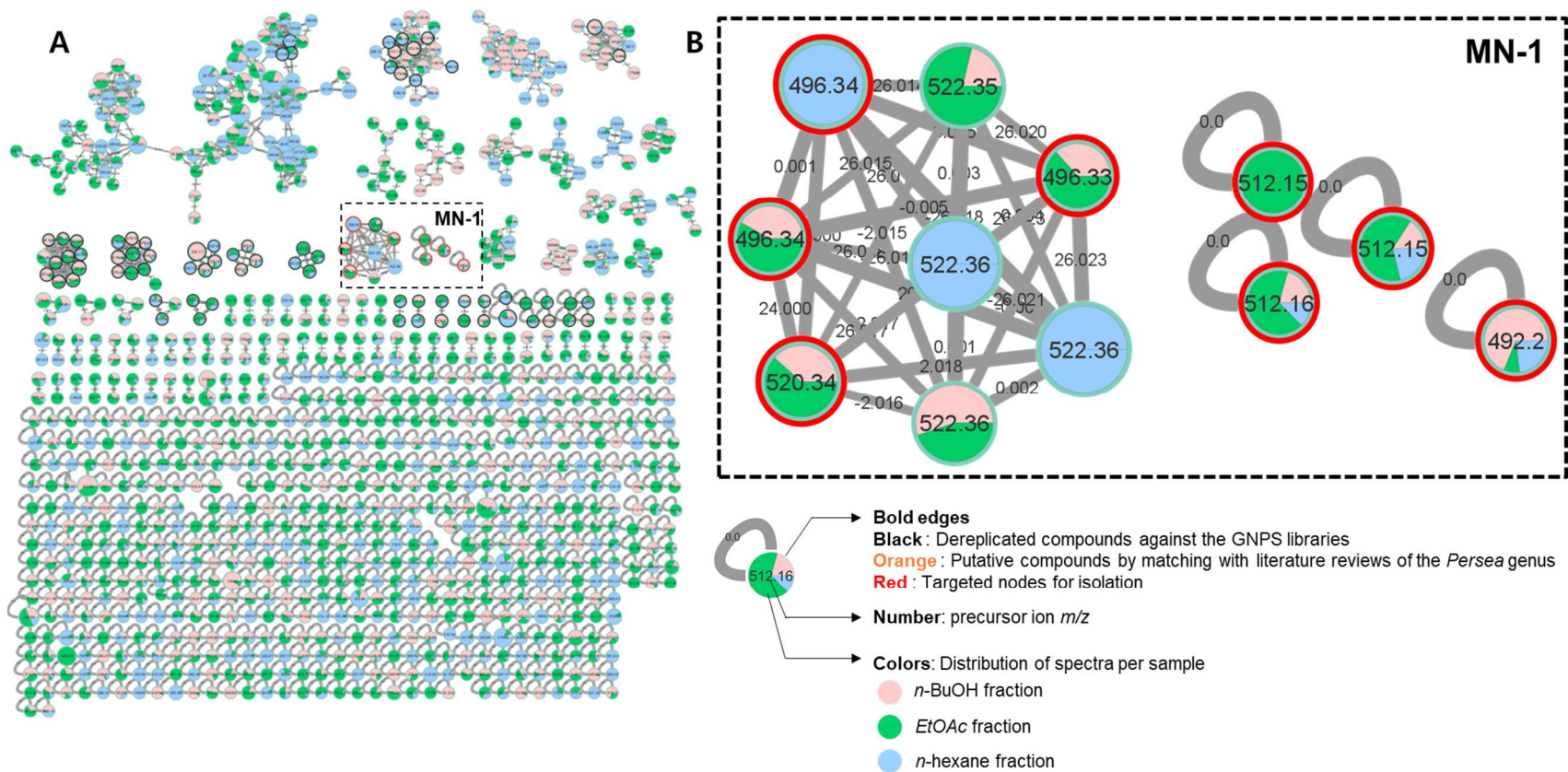
Figure 88. Molecular networking of *n*-hexane, EtOAc, and *n*-BuOH fractions of avocado seeds in pos mode.

This network is accessible at the following address: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=618fd6ab564f4d32aee9cb44d1fb8be7

Table 10. Putative compounds detected in EA fraction of Avocado seeds by UPLC-qTOFMS/MSs in negative mode

RT	*m/z*	MW (Calc.)	MF	Compound name	Ref
1.05	377.0824	378.2770	$C_{23}H_{38}O_4$	2-Hydroxy-4-oxoheneicosa-5,12,15-trien-1-yl acetate	(Kawagishi et al., 2001)
1.06	211.0820	212.0896	$C_7H_{16}O_7$	Perseitol	(Liu et al., 1999, 2002)
1.06	377.0835	378.2770	$C_{23}H_{38}O_4$	Persenone A	(Domergue et al., 2000)
1.15	191.0363	192.0634	$C_7H_{12}O_6$	Quinic acid	(Pahua-Ramos et al., 2012)
1.18	341.1090	342.1162	$C_{12}H_{22}O_{11}$	Sucrose	(Liu et al., 1999, 2002)
1.28	191.0238	192.0270	$C_6H_8O_7$	Citric acid	(Calderón-Oliver et al., 2016)
1.30	399.1463	400.1522	$C_{22}H_{24}O_7$	7-Hydroxy-6,8-di-C-methylflavanone 7-O-arabinoside	(Calderón-Oliver et al., 2016)
1.48	413.0853	414.3862	$C_{29}H_{50}O$	(-)-β-Sitosterol	(Berasategi et al., 2012)
1.51	315.1072	316.1158	$C_{14}H_{20}O_8$	hydroxytyrosol glucoside	(Araújo et al., 2018)
1.59	153.0558	154.0266	$C_7H_6O_4$	Protocatechuic acid	(Pahua-Ramos et al., 2012)
1.71	353.0856	354.0951	$C_{16}H_{18}O_9$	3-O-caffeoylquinic acid	(Kosińska et al., 2012)
1.71	353.0837	354.0951	$C_{16}H_{18}O_9$	Chlorogenic acid	(Kosińska et al., 2012)
1.72	353.0872	354.0951	$C_{16}H_{18}O_9$	4-caffeoylquinic acid	(Araújo et al., 2018)
1.97	299.1166	300.0634	$C_{16}H_8O_6$	Kaempferide	(Pahua-Ramos et al., 2012)
2.11	299.1152	300.1209	$C_{14}H_{20}O_7$	Tyrosol glucoside	(Araújo et al., 2018)
2.22	163.0402	164.0473	$C_9H_8O_3$	Coumaric acid	(Rodríguez-Carpena et al., 2011)
2.23	337.0945	338.1002	$C_{16}H_{18}O_8$	3-*p*-Coumaroylquinic acid	(Kosińska et al., 2012)
2.39	353.0860	354.0951	$C_{16}H_{18}O_9$	1-caffeoylquinic acid	(Araújo et al., 2018)
2.39	353.0856	354.0951	$C_{16}H_{18}O_9$	Neochlorogenic acid	(Kosińska et al., 2012)
2.50	577.1352	578.1424	$C_{30}H_{26}O_{12}$	Procyanidin B	(Wang et al., 2010)
2.83	289.0715	290.0790	$C_{15}H_{14}O_6$	Epicatechin	(Villa-Rodríguez et al., 2011)
3.01	337.0924	338.3185	$C_{22}H_{42}O_2$	Docosenoic acid	(Bora et al., 2001)
3.03	289.0715	290.0790	$C_{15}H_{14}O_6$	Catechin	(Villa-Rodríguez et al., 2011)
3.04	337.0918	338.1002	$C_{16}H_{18}O_8$	5-O-caffeoylquinic acid	(Kosińska et al., 2012)
3.22	577.1324	578.1424	$C_{30}H_{26}O_{12}$	Procyanidin dimer B	(Kosińska et al., 2012)
3.25	443.1542	444.1995	$C_{21}H_{32}O_{10}$	Penstemide	(Araújo et al., 2018)
3.36	329.0873	330.0951	$C_{14}H_{18}O_9$	vanillic acid glucoside	(Araújo et al., 2018)
3.46	377.1824	378.2770	$C_{19}H_{34}O_4$	16-Heptadecyne-1,2,4-triol, 1-acetate	(Hashimura et al., 2001)
3.52	567.2462	568.4280	$C_{40}H_{56}O_2$	Lutein A	(Ashton et al., 2006)
3.57	225.1131	226.1933	$C_{14}H_{26}O_2$	Tetradecenoic acid	(Bora et al., 2001)
3.58	441.1992	442.1839	$C_{21}H_{30}O_{10}$	8'-hydroxyabscisic acid β-D-glucoside	(Araújo et al., 2018)
3.63	441.1933	442.0900	$C_{22}H_{18}O_{10}$	Catechin 3-gallate	(Kosińska et al., 2012)
3.67	465.1084	464.0955	$C_{21}H_{20}O_{12}$	Hyperoside	(Kosińska et al., 2012)

3.96	595.1320	596.1377	$C_{26}H_{28}O_{16}$	Quercetin-3-*O*-arabinosylglucoside	(Kosińska et al., 2012)
4.13	575.1198	576.1268	$C_{30}H_{24}O_{12}$	Procyanidin A3	(Wang et al., 2010)
4.15	441.1769	442.0900	$C_{22}H_{18}O_{10}$	Epicatechin gallate	(Kosińska et al., 2012)
4.16	441.1780	445.0900	$C_{22}H_{18}O_{13}$	Epicatechin gallate	(Kosińska et al., 2012)
4.18	575.1196	576.1268	$C_{30}H_{24}O_{12}$	Procyanidin A2	(Wang et al., 2010)
4.24	575.1192	576.1268	$C_{30}H_{24}O_{12}$	Procyanidin A4	(Wang et al., 2010)
4.72	379.1968	380.2927	$C_{23}H_{40}O_4$	2-Hydroxy-4-oxoheneicosa-5,12-dien-1-yl acetate	(Kawagishi et al., 2001)
4.98	433.1121	434.0849	$C_{20}H_{18}O_{11}$	Quercetin 3-arabinoside	(Kosińska et al., 2012)
5.11	551.2128	552.4331	$C_{40}H_{56}O$	Cryptoxanthin	(Lu et al., 2005)
5.41	319.2117	320.2715	$C_{21}H_{36}O_2$	1-Hydroxyheneicosa-2,12,15-trien-4-one	(Kawagishi et al., 2001)
5.43	535.1568	536.4382	$C_{40}H_{56}$	*β*-Carotene	(Ashton et al., 2006)
5.77	465.2127	464.0955	$C_{21}H_{20}O_{12}$	Quercetin hexoside	(Melgar et al., 2018)
7.02	301.2018	302.0427	$C_{15}H_{10}O_7$	Quercetin	(Kosińska et al., 2012)
7.07	415.2332	416.3654	$C28H48O2$	*γ*-Tocopherol	(Lu et al., 2005)
7.33	483.2439	485.0772	$C_{21}H_{21}O_{11}.Cl$	Cyanidin 3-glucoside	(Cox et al., 2004)
7.76	415.2340	416.4018	$C_{29}H_{52}O$	Sitostanol	(Berasategi et al., 2012)
8.95	327.2180	328.2614	$C_{19}H_{36}O_4$	16-Heptadecene-1,2,4-triol, 1-acetate	(Domergue et al., 2000)
9.05	465.1193	464.0955	$C_{21}H_{20}O_{12}$	Quercetin 3-glucoside	(Kosińska et al., 2012)
9.40	465.1191	464.0955	$C_{21}H_{20}O_{12}$	Quercetin 3-O-galactoside	(Kosińska et al., 2012)
9.87	429.1589	430.3811	$C_{29}H_{50}O_2$	(+)-*α*-Tocopherol	(Lu et al., 2005)
10.64	351.2183	352.2614	$C_{21}H_{36}O_4$	1-(Acetyloxy)-2-hydroxy-5,16-nonadecadien-4-one	(Rodríguez-Sánchez et al., 2013)
10.83	285.2068	286.0477	$C_{15}H_{10}O_6$	Kaempferol	(Pahua-Ramos et al., 2012)
10.85	285.2074	286.0477	$C_{15}H_{10}O_6$	Luteolin	(Owolabi et al., 2010)
11.28	353.2338	354.2770	$C_{21}H_{38}O_4$	1-(Acetyloxy)-2-hydroxy-5-nonadecen-4-one	(Rodríguez-Sánchez et al., 2013)
11.46	351.2182	352.2614	$C_{21}H_{36}O_4$	1-(Acetyloxy)-2-hydroxy-16,18-nonadecadien-4-one	(Rodríguez-Sánchez et al., 2013)
12.01	353.2334	354.2770	$C_{21}H_{38}O_4$	1-Monolinolein	(Hashimura et al., 2001)
12.13	401.2551	402.3862	$C_{28}H_{50}O$	Campestanol	(Berasategi et al., 2012)
12.66	491.2685	492.0904	$C_{22}H_{20}O_{13}$	Isorhamnetin glucuronide	(Melgar et al., 2018)
12.90	269.2122	270.0528	$C_{15}H_{10}O_5$	Apigenin	(Owolabi et al., 2010)
13.21	227.1648	228.2089	$C_{14}H_{28}O_2$	Myristic acid	(Villa-Rodríguez et al., 2011)
13.68	315.2548	316.2977	$C_{19}H_{40}O_3$	1,2,4-Nonadecanetriol	(Oberlies et al., 1998)
13.92	335.1068	336.3028	$C_{22}H_{40}O_2$	Docosadienoic acid	(Villa-Rodríguez et al., 2011)
14.36	425.1645	426.3862	$C_{30}H_{50}O$	Lupeol	
14.44	277.1830	278.2246	$C_{18}H_{30}O_2$	Linolenic acid	(Ozdemir & Topuz, 2004)
15.73	399.2756	400.3705	$C_{28}H_{48}O$	Campesterol	(Berasategi et al., 2012)
15.93	439.2712	440.4018	$C_{31}H_{52}O$	24-Methylenecycloartanol	(Berasategi et al., 2012)
16.49	325.2372	326.2457	$C_{18}H_{36}O_4$	1,2,4-Hexadecanetriol, 1-acetate	(Hashimura et al., 2001)
16.52	325.2392	326.2457	$C_{19}H_{34}O_4$	Avocadenone acetate	(Brown, 1972)

17.08	281.2112	282.2559	$C_{18}H_{34}O_2$	Octadecenoic acid	(Villa-Rodríguez et al., 2011)
17.52	297.2444	298.2508	$C_{18}H_{34}O_3$	6-(12-Tridecen-1-yl)-1,3-dioxane-4-methanol	(Oberlies et al., 1998)
18.05	255.1966	256.2402	$C_{16}H_{32}O_2$	Palmitic acid	(Ozdemir & Topuz, 2004)
18.29	277.2180	278.2246	$C_{18}H_{30}O_2$	*cis*-Octadecatrienoic acid	(Villa-Rodríguez et al., 2011)
18.30	311.1689	312.3028	$C_{20}H_{40}O_2$	Eicosanoic acid	(Ozdemir & Topuz, 2004)
19.06	253.2175	254.2246	$C_{16}H_{30}O_2$	Hexadecenoic acid	(Villa-Rodríguez et al., 2011)
19.06	277.2531	278.2246	$C_{18}H_{30}O_2$	Gamma-linolenic acid	(Plaza et al., 2009)
19.16	269.2131	270.2559	$C_{17}H_{34}O_2$	Heptadecanoic acid	(Bora et al., 2001)
19.21	253.2167	254.2246	$C_{16}H_{30}O_2$	Palmitoleic acid	(Ozdemir & Topuz, 2004)
19.53	379.1586	380.2927	$C_{23}H_{40}O_4$	2-Hydroxy-4-oxoheneicosa-12,15-dien-1- yl acetate	(Kawagishi et al., 2001)
19.55	279.2330	280.2402	$C_{18}H_{32}O_2$	Octadecadienoic acid	(Villa-Rodríguez et al., 2011)
19.71	241.2166	242.2246	$C_{15}H_{30}O_2$	Pentadecanoic acid	(Bora et al., 2001)
19.75	279.2326	280.2402	$C_{18}H_{32}O_2$	Linoleic acid	(Ozdemir & Topuz, 2004)
20.10	267.2328	268.2402	$C_{17}H_{32}O_2$	Heptadecenoic acid	(Villa-Rodríguez et al., 2011)
20.60	281.2494	282.2559	$C_{18}H_{34}O_2$	Vaccenic acid	(Plaza et al., 2009)
20.60	381.1740	382.3083	$C_{23}H_{42}O_4$	12,15-Heneicosadiene-1,2,4-triol, 1-acetate	(Rodríguez-Sánchez et al., 2013)
20.60	281.2485	282.2559	$C_{18}H_{34}O_2$	Oleic acid	(Landahl et al., 2009)
20.75	307.2639	308.2715	$C_{20}H_{36}O_2$	Eicosadienoic acid	(Villa-Rodríguez et al., 2011)
21.20	283.2638	284.2351	$C_{17}H_{32}O_3$	16-Heptadecyne-1,2,4-triol	(Oberlies et al., 1998)
21.22	283.2647	284.2715	$C_{18}H_{36}O_2$	Stearic acid	(Ozdemir & Topuz, 2004)
21.33	367.3605	368.3654	$C_{24}H_{48}O_2$	Tetracosanoic acid	(Bora et al., 2001)
23.46	325.1846	326.2457	$C_{19}H_{34}O_4$	16-Heptadecyne-1,2,4-triol, 1-acetate	(Domergue et al., 2000)

1. (4*S*) R₁ = H, R₂ = S₁
2. (4*R*) R₁ = H, R₂ = S₁
3. (4*R*) R₁ = H, R₂ = S₂
4. (4*S*) R₁ = H, R₂ = S₂
5. (4*S*) R₁ = α– OH, R₂ = S₁
6. (4*R*) R₁ = β– OH, R₂ = S₁

7. R = S₁

8. (4*S*) R₁ = α– OMe, R₂ = S₁
9. (4*R*) R₁ = β– OMe, R₂ = S₁

S₁

S₂

10

11
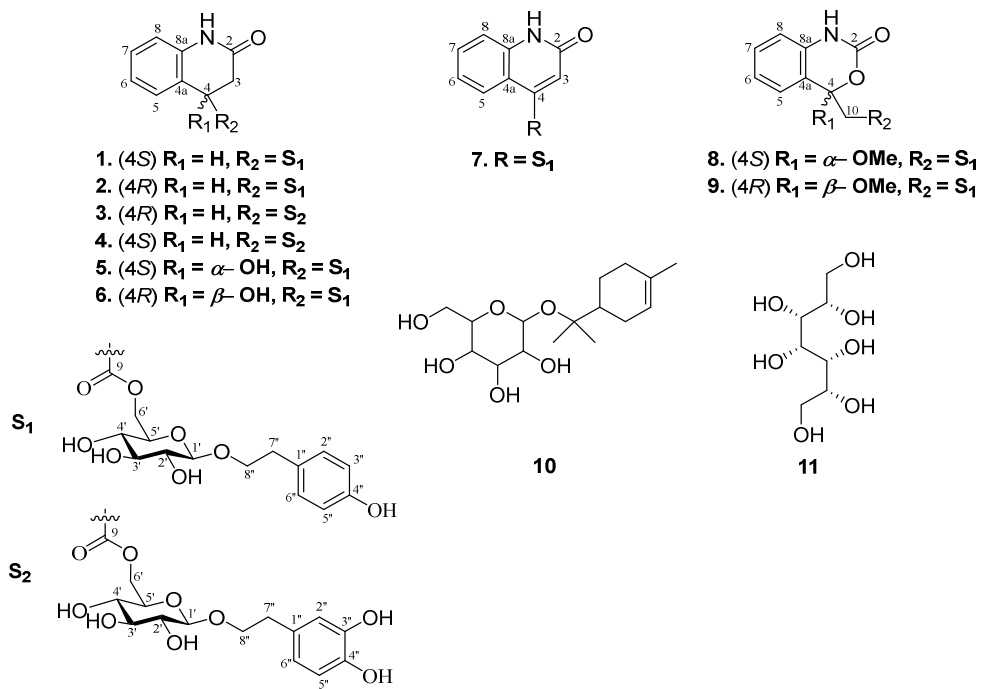
Figure 89. Chemical structures of isolated compounds **1–11** from *P. americana.*

3. Structure elucidation of alkaloids from *P. americana*

3.1. Compound **1**

Compound **1** was obtained as a pale yellow powder and has a molecular formula of $C_{24}H_{27}NO_{10}$, which was deduced from its HRESIMS ion in positive mode at *m/z* 472.1617 $[M - H]^-$, (calcd. for $C_{24}H_{26}NO_9$ 472.1608). The IR spectrum showed absorption bands at 3339 cm^{-1} for hydroxy or amine, 1720 cm^{-1} for carbonyl functional groups, and 1618 and 1517 cm^{-1} for aromatic rings. The ^1H NMR spectrum showed signals for eight aromatic protons at δ_H 7.24 (1H, dd, *J* = 10.6, 7.3 Hz, H-5), 7.18 (1H, q, *J* = 7.3 Hz, H-7), 7.02 (2H, dd, *J* = 15.9, 8.3 Hz, H-2", H-6"), 6.93 (1H, m, H-6), 6.88 (1H, dd, *J* = 7.7, 2.6 Hz, H-8), and 6.68 (2H, t, *J* = 9.1 Hz, H-3", H-5"), an anomeric proton (δ_H 4.27 ppm, H-1'), two pairs of oxygenated methylenes at δ_H 4.44/4.21 (H$_2$-6') and δ_H 2.81 (H$_2$-7"), one methine group (δ_H 3.76, t, *J* = 6.4 Hz, H-4), and two methylene groups at δ_H 3.07/2.84 (H$_2$-3) and 3.91/3.67 (H$_2$-8"). The ^{13}C NMR spectrum of **1** revealed the presence of 24 carbon signals, including two carbonyl groups (δ_C 181.1 and 172.6 ppm), 12 aromatics (δ_C 110.9–156.8 ppm), one anomeric (δ_C 104.4), six oxygenated (δ_C 65.9–77.8), one methine (δ_C 43.4 ppm), and two methylene (δ_C 36.4 and 35.5 ppm) carbons. The HMBC and ^1H-^1H COSY NMR data of **1** indicated the presence of a salidroside moiety (Candido et al., 2016; Yang et al., 2021) and an alkaloid moiety that was identified as 2-oxindole-3-acetic acid (O-3AA) or 2-oxo-1,2,3,4-tetrahydroquinoline-4-carboxylic acid (2O-CA) (Liu et al., 2016). The presence of an A$_2$B$_2$-type aromatic ring and ethylene moiety linked to the aromatic ring at C-1" was demonstrated by the cross HMBC from H$_2$-7" (δ_H 2.81 ppm) and H$_2$-8" (δ_H 3.91, 3.67 ppm) to C-1" (Figure 92). The coupling constant (d, *J* = 7.8 Hz) of the anomeric proton signal and carbon data at δ_C 104.4 ppm suggested a β-linked D-glucose (Roslund et al., 2008; Tanaka, 1985). The cross-HMBC peaks from the anomeric proton H-1' (δ_H 4.27 ppm) to C-8" (δ_C 72.3 ppm) and from H$_2$-6' (δ_H 4.44, 4.21 ppm) to the carbonyl group at δ_C 172.6 ppm (C-9) indicated that the salidroside moiety was connected to the alkaloid unit through an ester linkage (C-9). When elucidating the structure of the alkaloid unit of **1**, HMBC correlations of the alkaloid unit could not be used to distinguish the five- or six-member heteroring of **1** (Figure 93). The HMBC cross-peaks from H$_2$-3 (δ_H 3.07, 2.84 ppm) to C-2 (δ_C 181.1 ppm), C-4 (δ_C 43.4 ppm), C-4a (δ_C 130.6 ppm), C-9 (δ_C 172.6 ppm), and from H-5 (δ_H 7.24 ppm) to C-4 made it impossible to determine

structures **1a** (2O-CA) and **1b** (O-3AA) (Figure 93). Previous studies on a moiety similar to **1** indicated that quantum chemical calculations, including ^{13}C NMR, ECD, and ORD, could not be helpful for the identification of the five- or six-membered rings in **1** (Liu et al., 2016; Zhang et al., 2014). Therefore, to confirm the alkaloid unit in **1**, a de-esterification reaction was performed with **1** to obtain the alkaloid moiety. The product was compared to two authentic compounds (O-3AA and 2O-CA) under the HPLC condition of 10–90% MeCN/H_2O (0.1% formic acid) for 20 min. The retention time of the reacted product suggested that the alkaloid moiety in **1** had an identity of 2O-CA (Figure 94). The absolute configuration at C-4 of **1** was identified by comparing the experimental data to the ECD calculation (Figure 101). As shown in Figure 101, the experimental CD data of **1** exhibited (–) cotton effects (CE) at 205, 245 nm, and (+) CEs at 220, and 280 nm, and the compound had similar ECD data as that calculated for (4*S*)–**1**. Hence, the structure of **1** was identified, and the compound was named avoquinoside A.
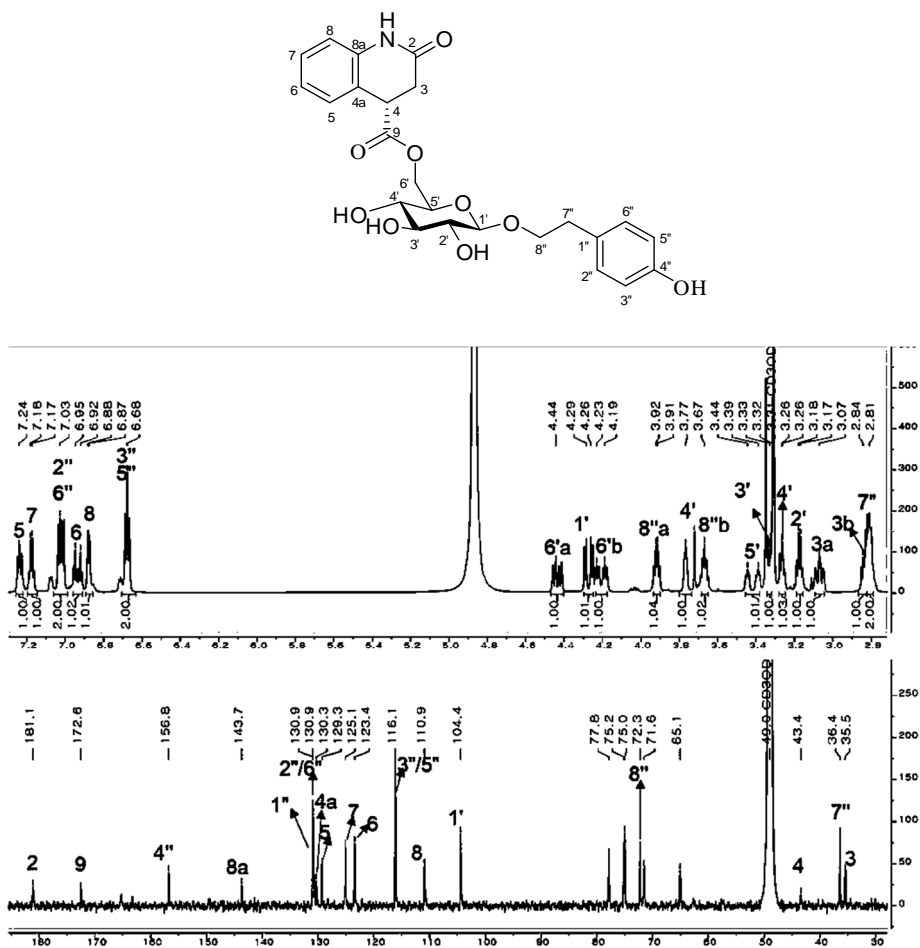
Figure 90. ^{1}H and ^{13}C NMR spectra of compound **1** (800/ 200 MHz, methanol-d_4)
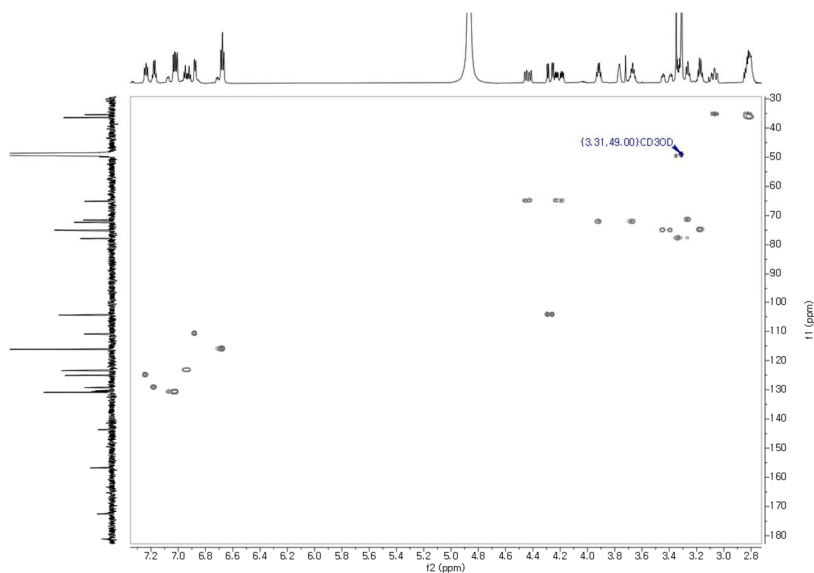
Figure 91. HSQC spectrum of compound **1**

Figure 92. HMBC spectrum of compound **1**
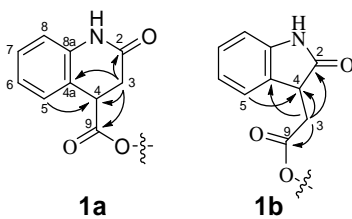
Figure 93. Possible structures of **1 (1a, 1b)** by HMBC correlations

Figure 94. The HPLC chromatography of two authentic compounds and reacted solution of compounds 1 and 2 for identification of alkaloid moiety in **1** and **2**.
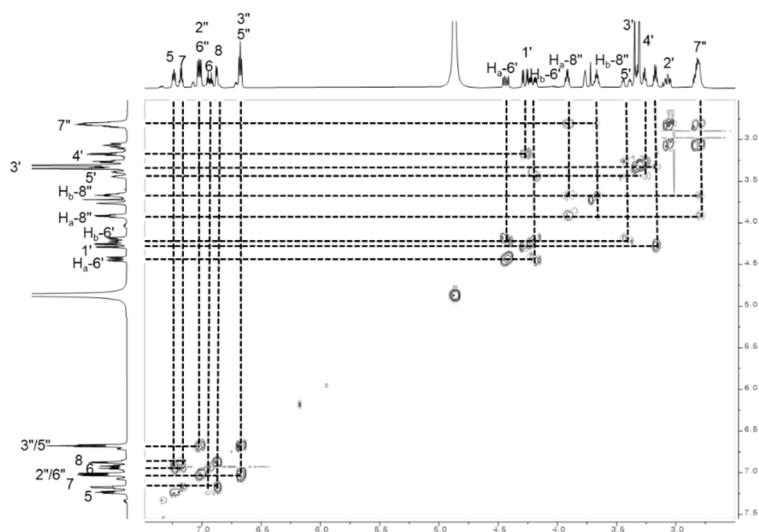
Figure 95. ^{1}H–^{1}H COSY spectrum of compound **1**

3.2. Compound 2

Compound **2** was separated from the mixture of compound **1** diastereomers by an HPLC equipped with a YMC Triart henyl hexyl column using a gradient of 45–55% MeOH/H$_2$O (0.1% formic acid) (Figure 96). The ^{1}H and ^{13}C NMR data of **2** shared similarities to those of **1** (Tables 11, 12), indicating that **2** had the same planar structure as **1**. The difference between the two compounds was illustrated in their CD data, which showed opposite orientations at C-4. Compared to the calculated ECD data of (4R)–**1** and (4S)–**1** (Figure 101), the absolute configuration of **2** was determined to be (4R)–**1,** which exhibited (+) cotton effects (CE) at 205 and 245 nm and (–) CEs at 220 and 280 nm. Therefore, the structure of **2** was identified to be avoquinoside B. Interestingly, the interday CD experiments of **1** and **2** after re-HPLC was performed on the 0th, 3rd, 10th, and 15th days showed that the CEs at the curtained wavelengths (205, 220, 245, and 280 nm) gradually reduced and approached 0 (Figure 102). This phenomenon suggested that the two compounds tended to convert back into a mixture of diastereomers.
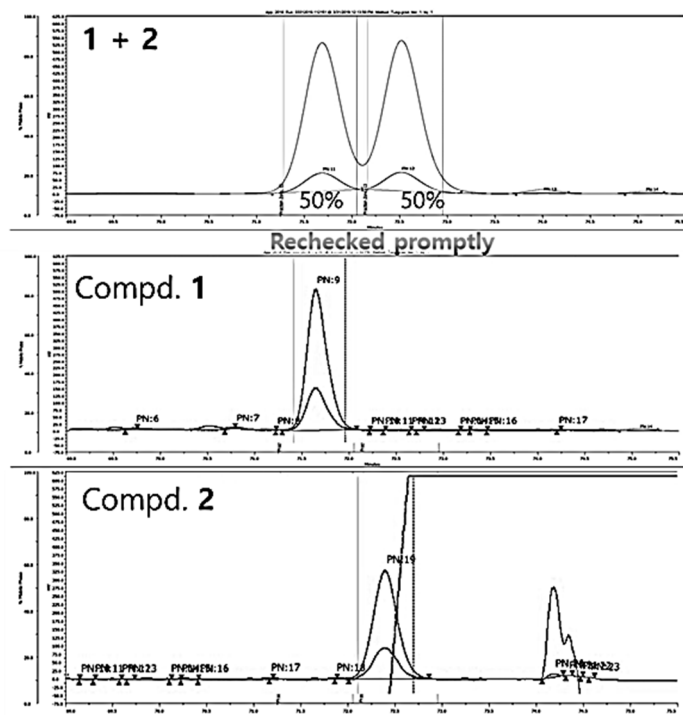
Figure 96. Re-HPLC separation of **1** and **2**. The HPLC separation used a YMC phenyl hexyl column (5μm, 250 x 10 mm) in a gradient of 45–55 % MeOH/H$_2$O (0.1% formic acid) for 75 min, the flow rate of 2.0 mL/min coupled with UV detector that was set at 210 and 280 nm.
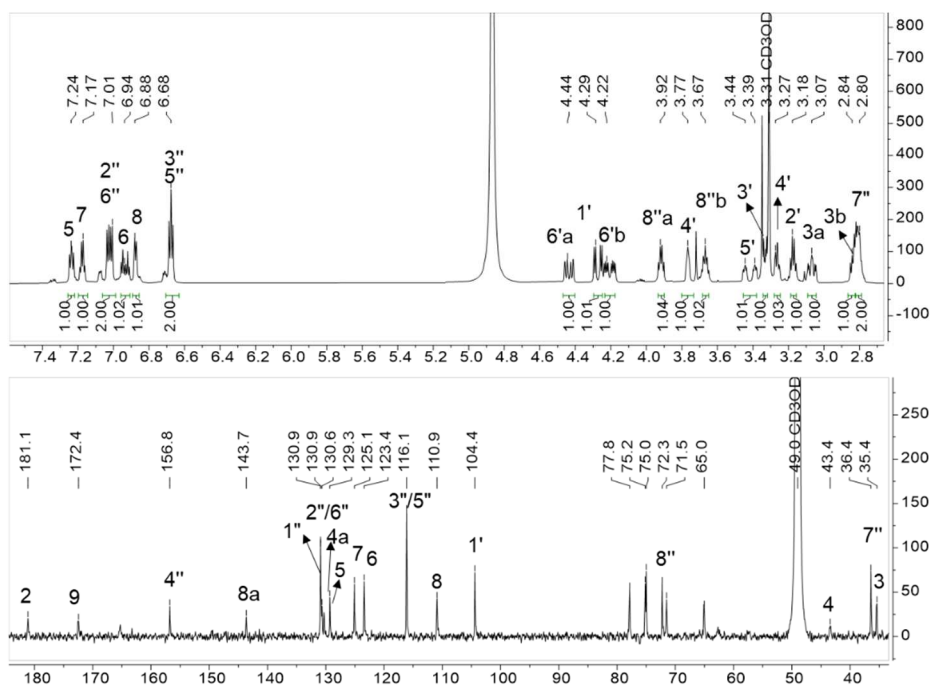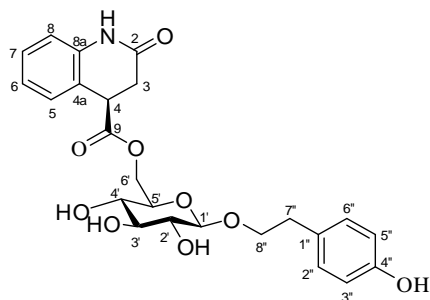
Figure 97. ^{1}H and ^{13}C NMR spectra of compound **2** (800/200 MHz, methanol-d_4)
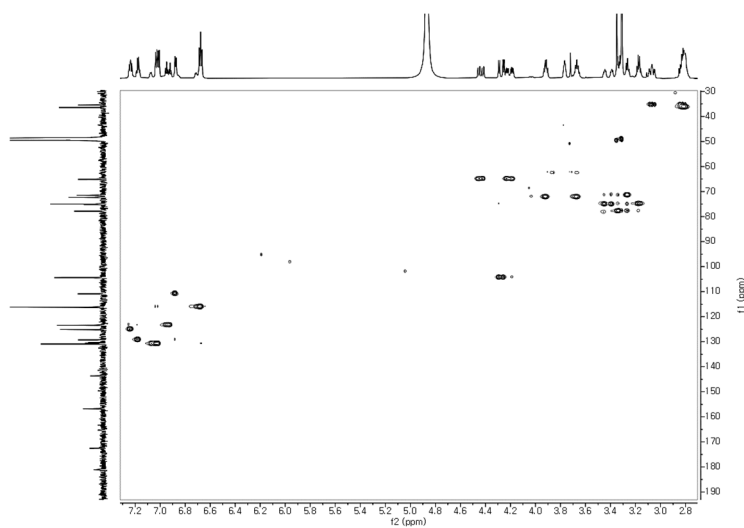
Figure 98. HSQC spectrum of compound **2**
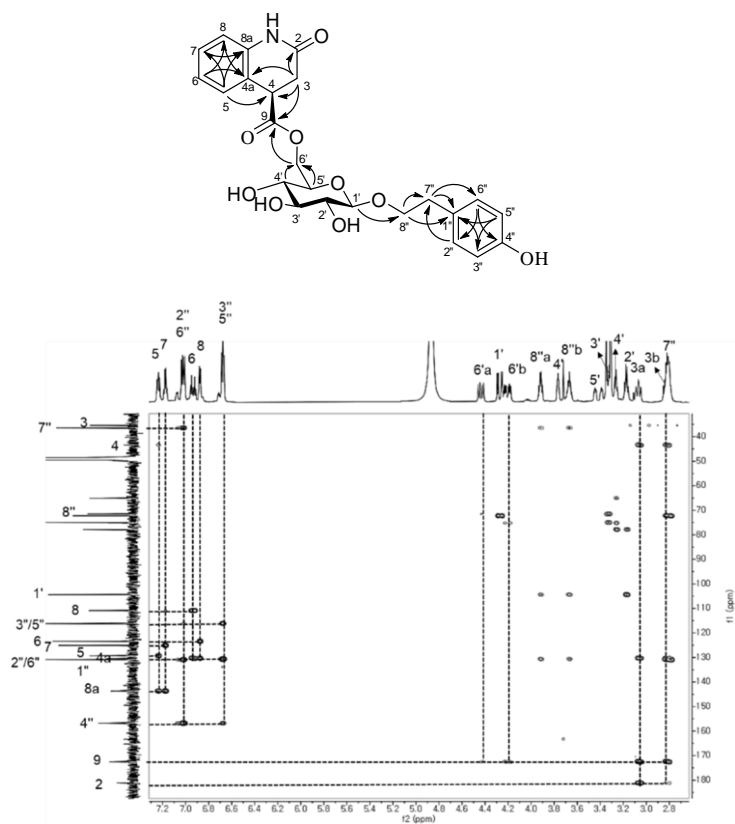
Figure 99. HMBC spectrum of compound **2**

Figure 100. ^{1}H–^{1}H COSY spectrum of compound **2**

122
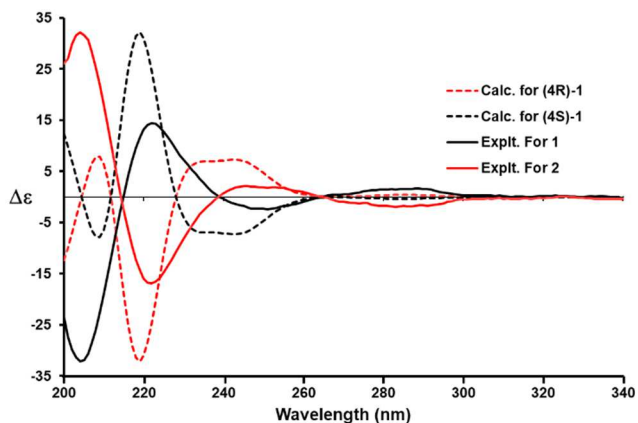
Figure 101. The experimental spectra of **1** and **2**, which were recorded directly after the compounds were separated from their diastereomer mixtures and calculated ECD data for (4*R*)–**1** and (4*S*)–**1**. The calculated ECD curves were generated by SpecDis ver. 1.71.40 (Bruhn et al., 2013b)
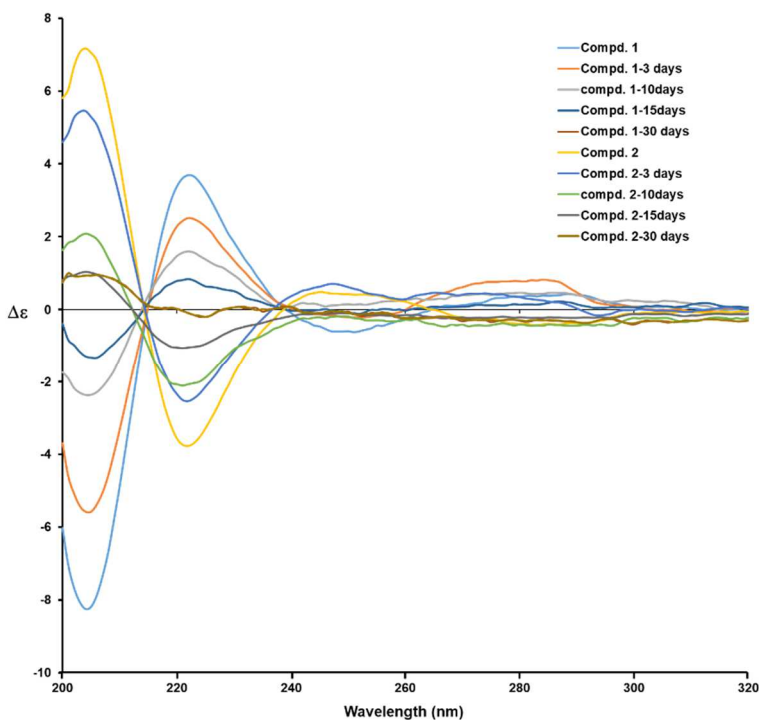
Figure 102. The inter-day CD data of compounds **1, 2** at 0th, 3rd, 10th, 15th and 30th day after separating from their diastereomer mixtures.

3.3. Compound 3

Compound **3** was obtained as a pale yellow powder with a molecular formula of $C_{24}H_{27}NO_{10}$, as deduced from its HRESIMS ion in positive mode at 488.1518 [M – H]$^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557). The IR spectrum showed absorption bands for hydroxy or amine (3345 cm^{-1}) and carbonyl (1722 cm^{-1}) functional groups and 1607 and 1513 cm^{-1} for aromatic rings. The ^{1}H and ^{13}C NMR spectra of **3** were similar those of **1** and **2** and included two aromatic rings with the presence of seven aromatic protons (δ_H 7.23, 7.17, 6.93, 6.87, 6.68, 6.66, and 6.53 ppm), as well as twelve aromatic carbons (δ_C 110.9–146.1 ppm), a sugar unit with an anomeric signal (δ_H 4.27; δ_C 104.4), and two carbonyl carbons at δ_C 181.2 and 172.6 ppm. The difference was that in **3**, one ABX aromatic system was present instead of the A_2B_2-type aromatic ring of the salidroside moiety in **1** and **2**. The lack of one aromatic proton signal and the presence of an additional hydroxyl group in **3** supported the attachment to the salidroside moiety at C-3″ (δ_C 146.1 ppm). The experimental CD data of **3** showed a similar pattern of (4*R*)–**1** with a (–) CE at approximately 220 nm and (+) CE at approximately 240 nm (Figure 111). Hence, the absolute configuration of **3** was identified as (4*R*) and named avoquinoside C.
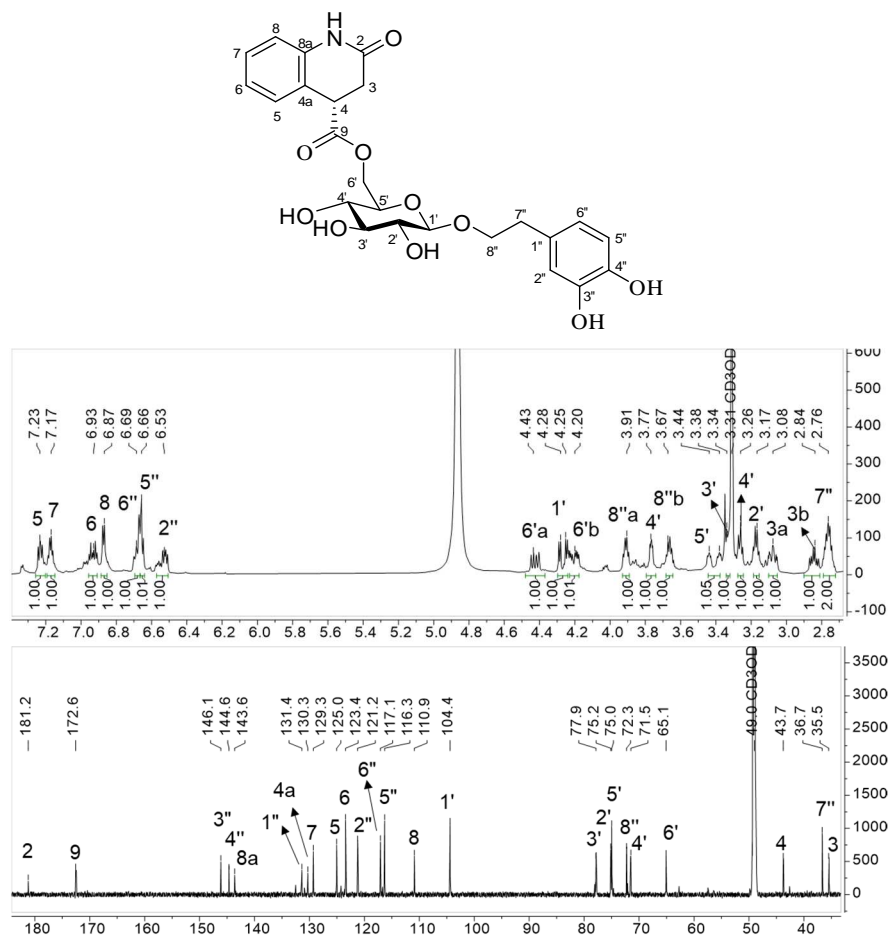
Figure 103. ^1H and ^{13}C NMR spectra of compound **3** (800/200 MHz, methanol-d_4)
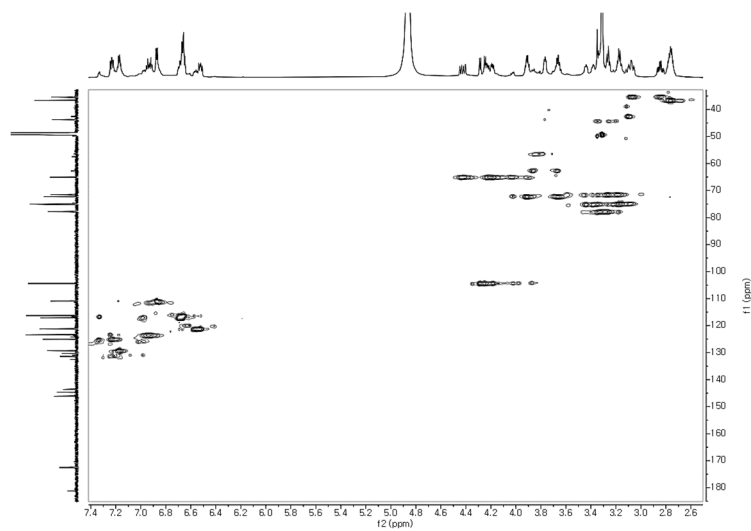
Figure 104. HSQC spectrum of compound **3**

Figure 105. HMBC spectrum of compound **3**
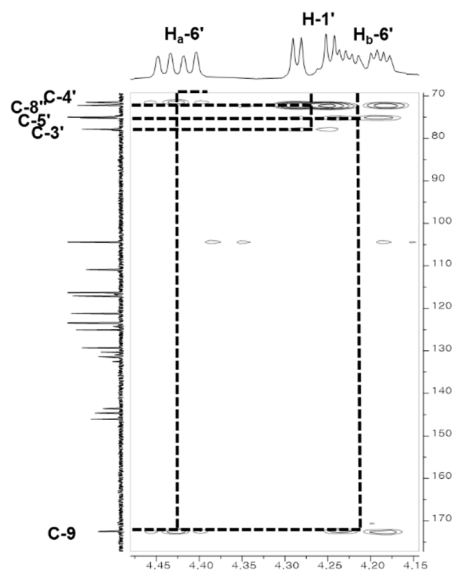
Figure 105 (continued). HMBC spectrum of compound **3**
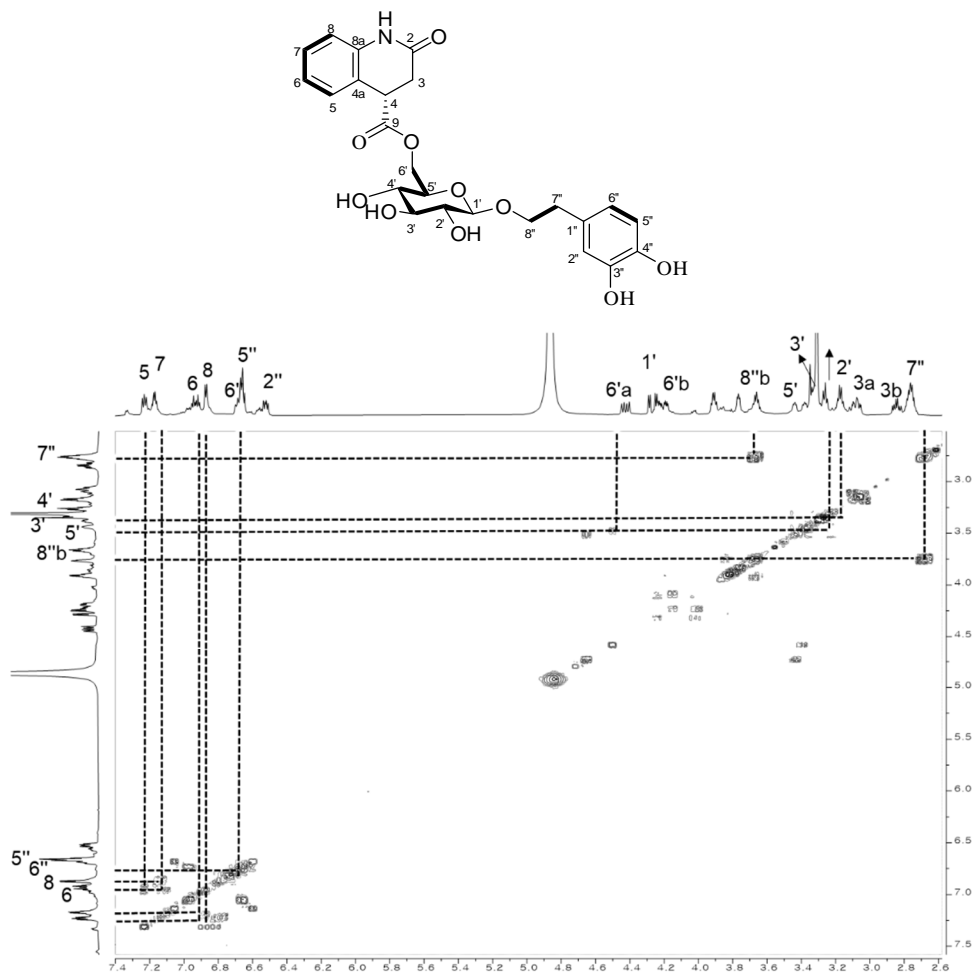
Figure 106. ^{1}H–^{1}H COSY spectrum of compound **3**

3.4. Compound 4

Compound **4** was separated from the mixture of compound **3** diastereomers by an HPLC equipped with a YMC Triart phenyl hexyl column using a gradient of 40–45% MeOH/H$_2$O (0.1% formic acid). Compound **4** was obtained as a pale yellow powder that possessed the same molecular formula, which was deduced from its HRESIMS ion in positive mode at *m/z* 488.1505 [M – H]$^-$ (calcd. for C$_{24}$H$_{26}$NO$_{10}$ 488.1557), as that of **3** (C$_{24}$H$_{27}$NO$_{10}$). The 1D and 2D NMR spectra of **4** were similar to those of **3**, suggesting that they possessed the same planar structure. The experimental CD data of **4** showed a similar feature to the CEs of the ECD data for (4*S*)–**1**, in which a (+) CE at 220 nm and a (–) CE at 240 nm were found (Figure 109). Thus, compound **4** was determined as shown in Figure 87 and was named avoquinoside D.
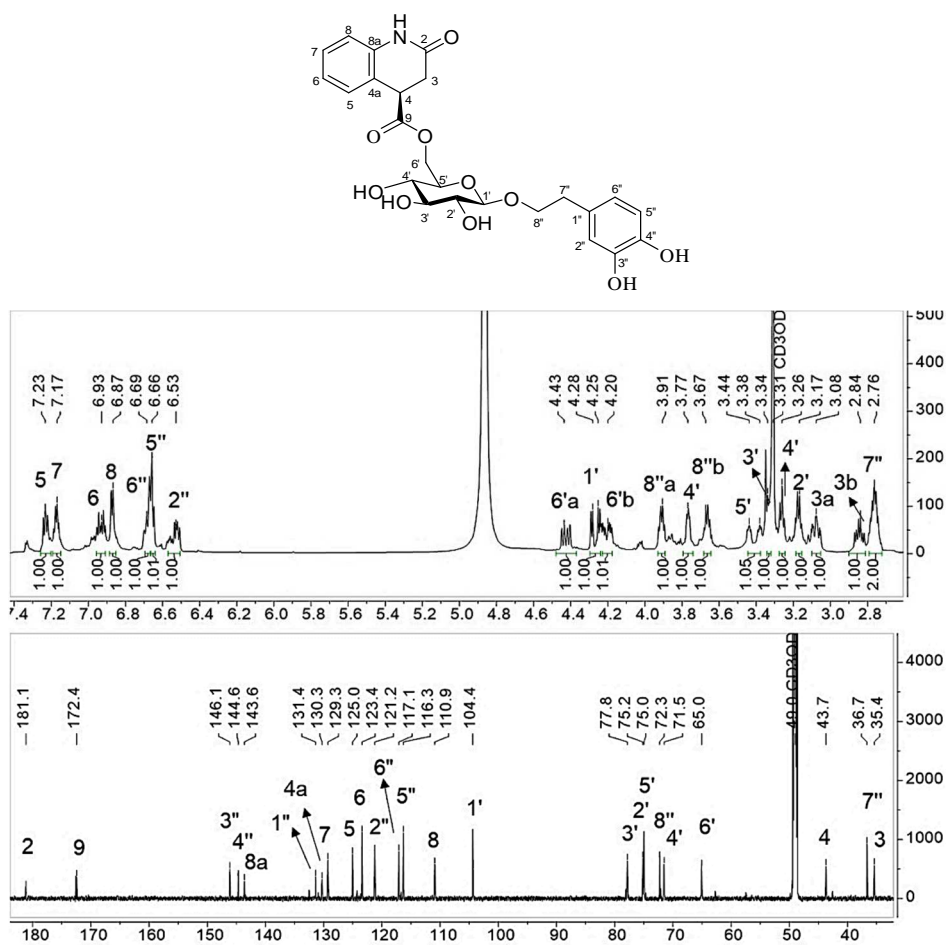
Figure 107. ^{1}H and ^{13}C NMR spectra of compound **4** (800/200 MHz, methanol-*d*$_4$)

128
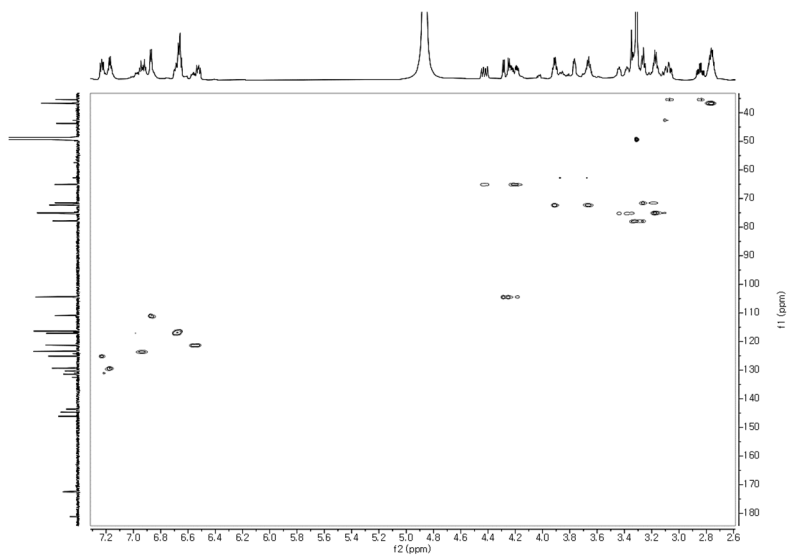
Figure 108. HSQC spectrum of compound **4**
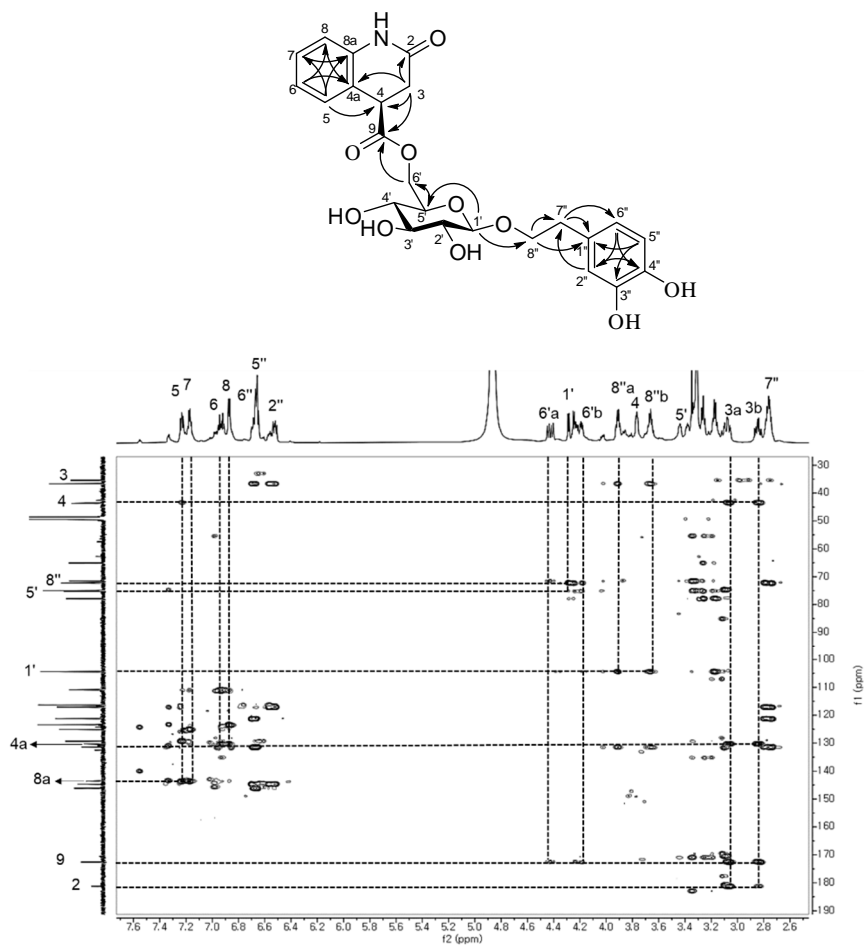
Figure 109. HMBC spectrum of compound **4**
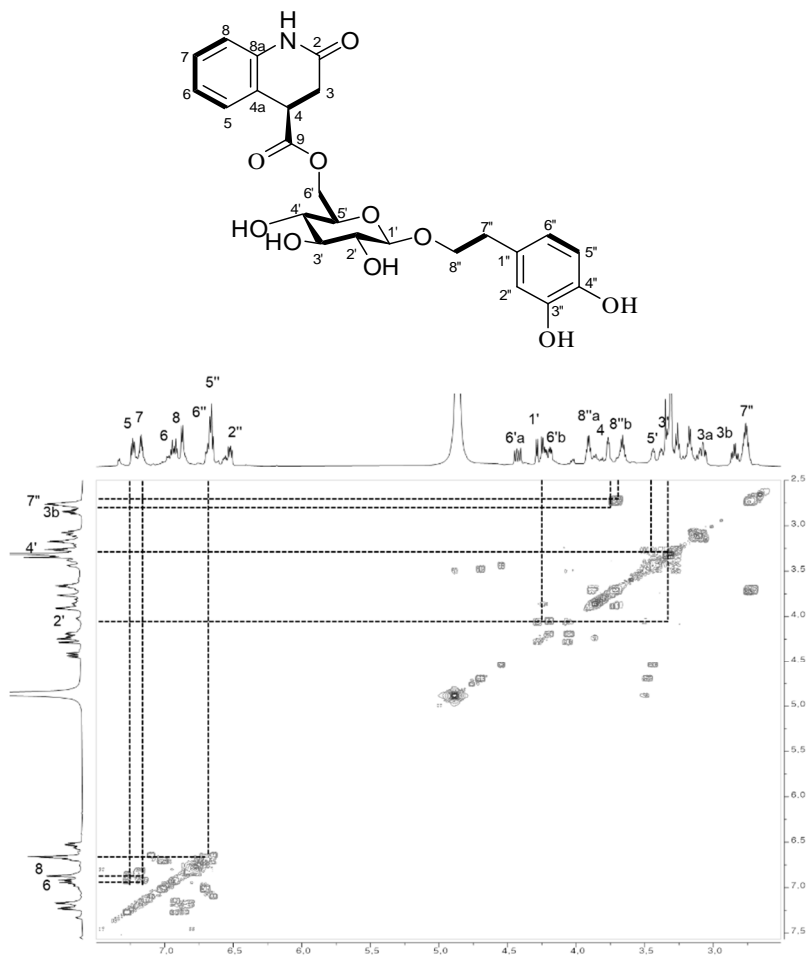
Figure 110. ^{1}H–^{1}H COSY spectrum of compound **4**
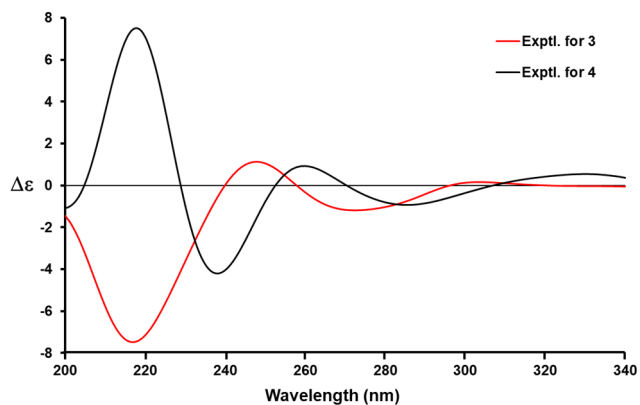
Figure 111. The experimental spectra of **3** and **4**, which were recorded directly after the compounds were separated from their diastereomer mixtures.

3.5. Compound **5**

Compound **5** was obtained as a pale yellow powder with a molecular formula of $C_{24}H_{27}NO_{10}$, as deduced from its HRESIMS ion in positive mode at *m/z* 488.1557 $[M - H]^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557). The IR spectrum showed absorption bands for hydroxy or amine (3370 cm^{-1}) and carbonyl (1722 cm^{-1}) functional groups and 1612 and 1519 cm^{-1} for aromatic rings. The ^1H and ^{13}C NMR spectra of **5** were similar to those of **1** and **2** except for the presence of an additional quaternary carbon at δ_C 74.7 ppm that was oxygenated. The HMBC cross-peaks from H_2-3 (δ_H 3.09 ppm) to C-4 (δ_C 74.7 ppm), C-2 (δ_C 180.7 ppm), C-9 (δ_C 170.4 ppm), and C-4a (δ_C 131.6 ppm), as well as from H-5 (δ_H 7.33 ppm) to C-4, suggested that the additional hydroxy group was located on C-4. Similar to **1**–**4**, the anomeric signal of the sugar unit was observed at δ_H 4.26 (d, *J* = 7.8 Hz) and δ_C 104.3 ppm, indicating the presence of a β-linked D-glucose in **5**. The linkage from the salidroside moiety to the alkaloid unit was indicated by the HMBC cross-peaks from H_2-6' (δ_H 4.20, 4.04 ppm) to C-9, which were similar to those of **1**–**4**. The experimental CD spectrum of **5** showed (+) CEs at 235–245 nm and 275–285 nm bands that were similar to the calculated data for (4*R*)–**5** (Figure 120). Thus, the absolute configuration at C-4 in **5** was assigned as 4*R*, and the compound was named asavoquinoside E.
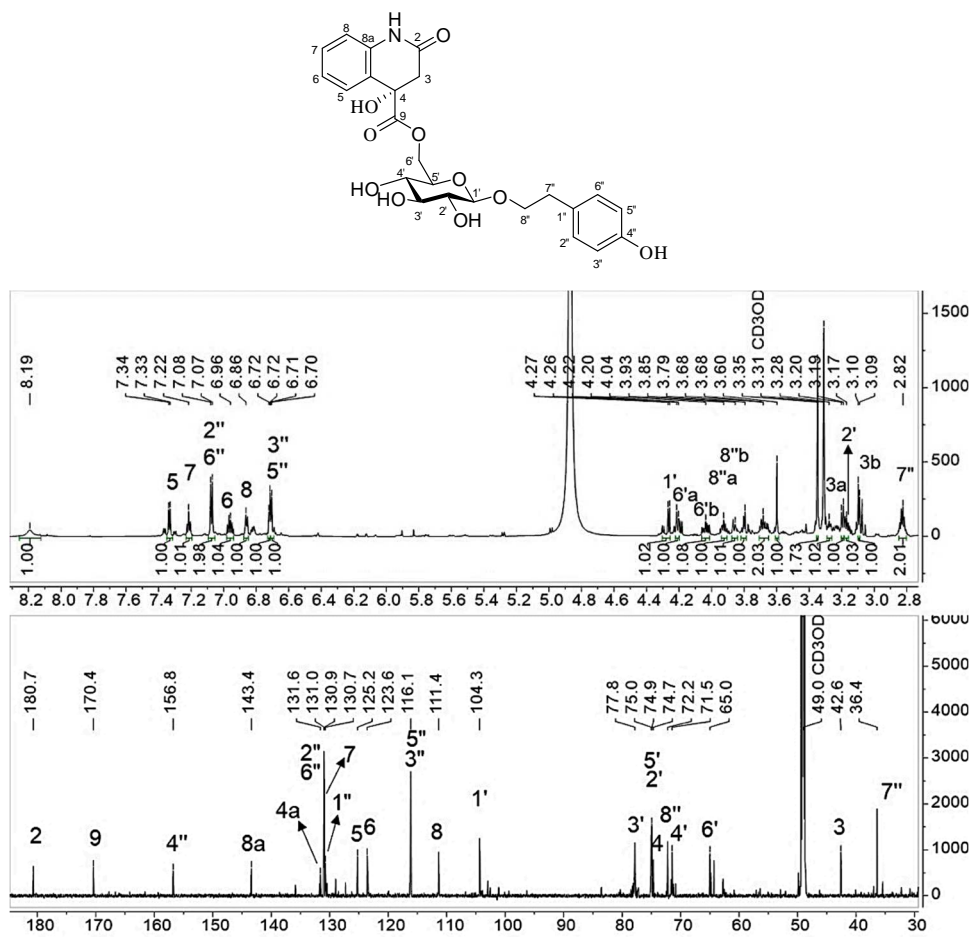
**Figure 112.** ^{1}H and ^{13}C NMR spectra of compound **5** (800/200 MHz, methanol-d_4)
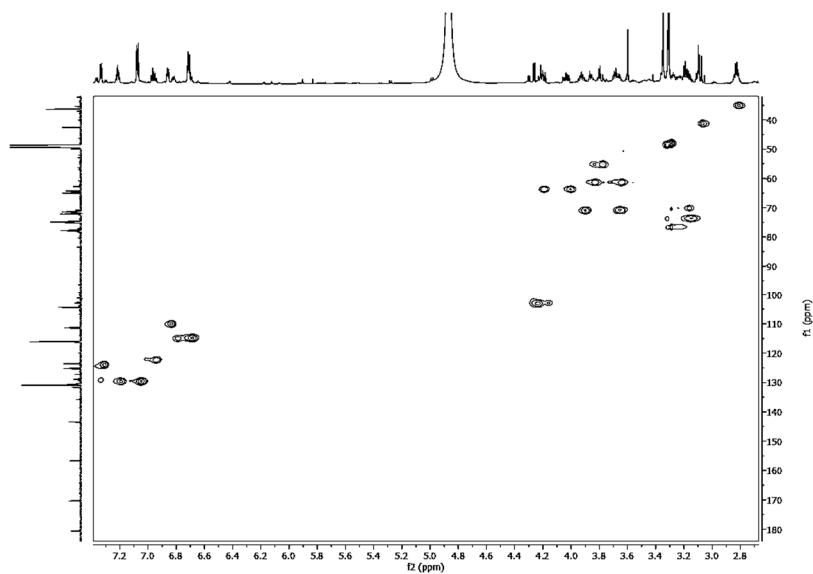
Figure 113. HSQC spectrum of compound **5**
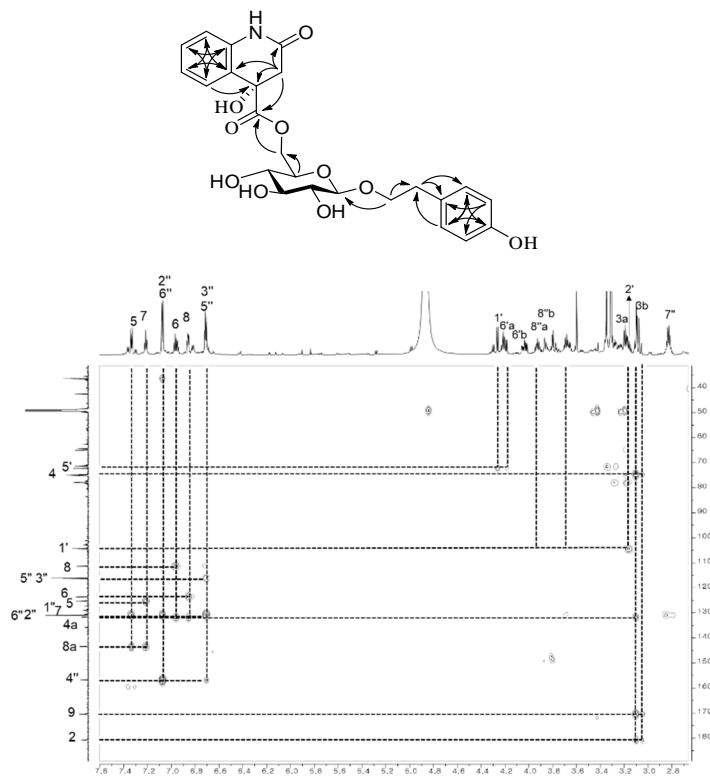
Figure 114. HMBC spectrum of compound **5**
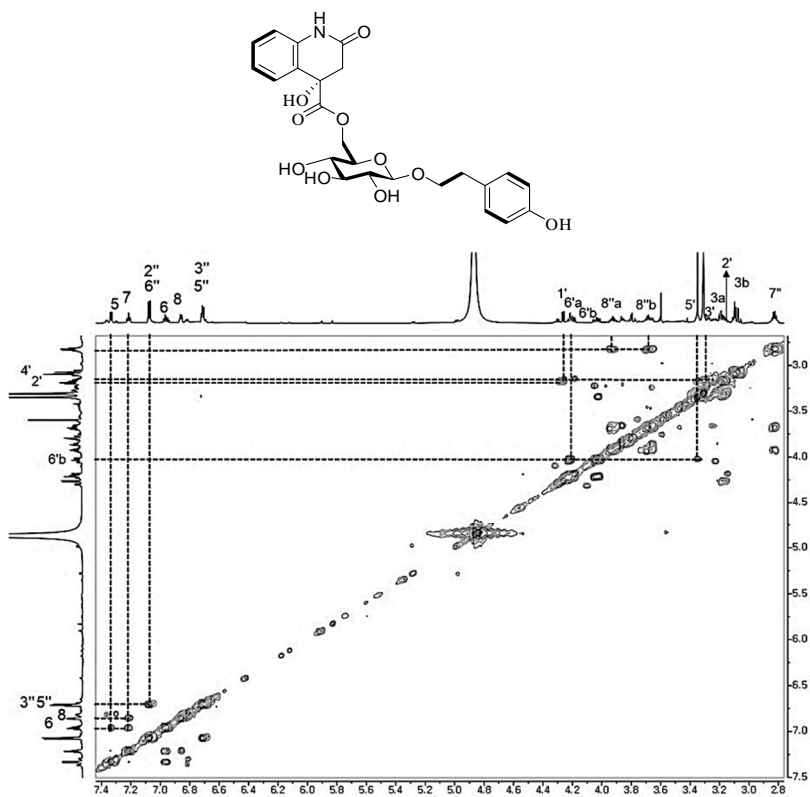
Figure 115. ^{1}H–^{1}H COSY spectrum of compound **5**

133

3.6. Compound 6

Compound **6** was obtained as a pale yellow powder and has a molecular formula of $C_{24}H_{27}NO_{10}$, as deduced from its HRESIMS ion in positive mode at *m/z* 488.1565 [M – H]⁻, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557). The difference between **5** and **6** was revealed by comparing their CD data to the calculated ECD for (4*R*)–**5** and (4*S*)–**5**, which suggested that the compounds are diastereomers at C-4. As a result, the CD spectrum of **6** was identical with the calculated ECD data of (4*S*)–**5**, which showed an (+) CE at approximately 230–240 nm and (–) CE at 275–280 nm (Figure 120). Hence, **6** was named avoquinoside F.
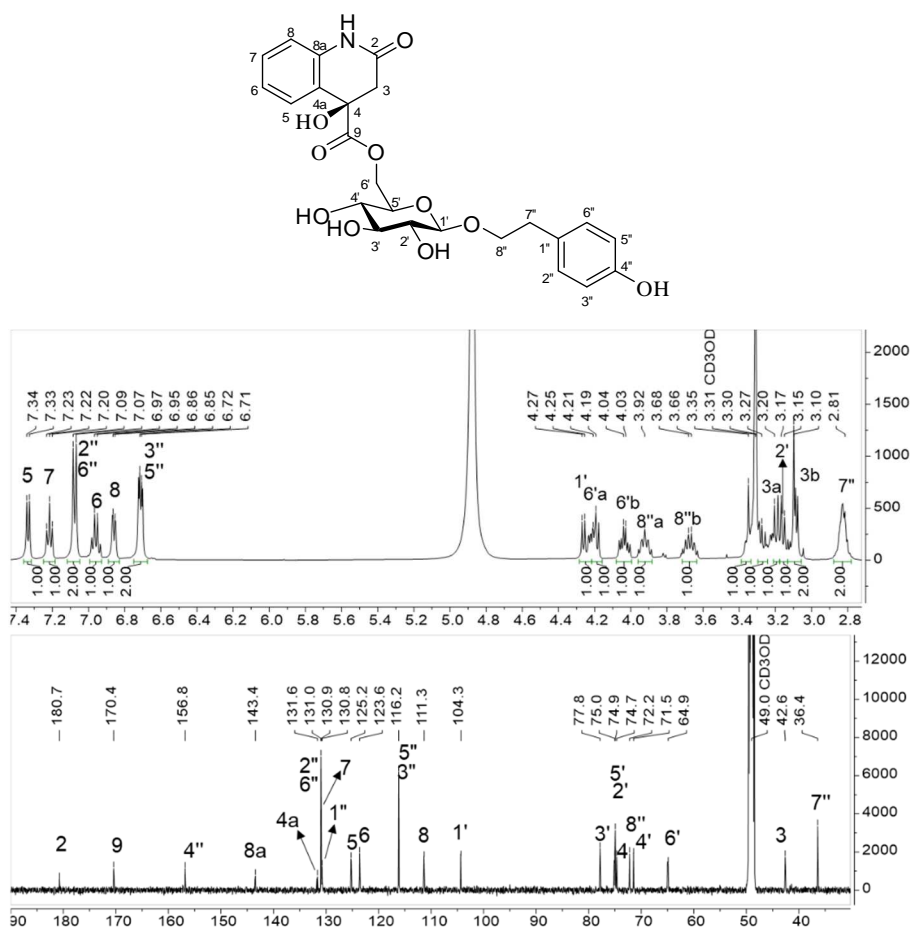
Figure 116. ¹H and ¹³C NMR spectra of compound 6 (500/125 MHz, methanol-*d₄*)
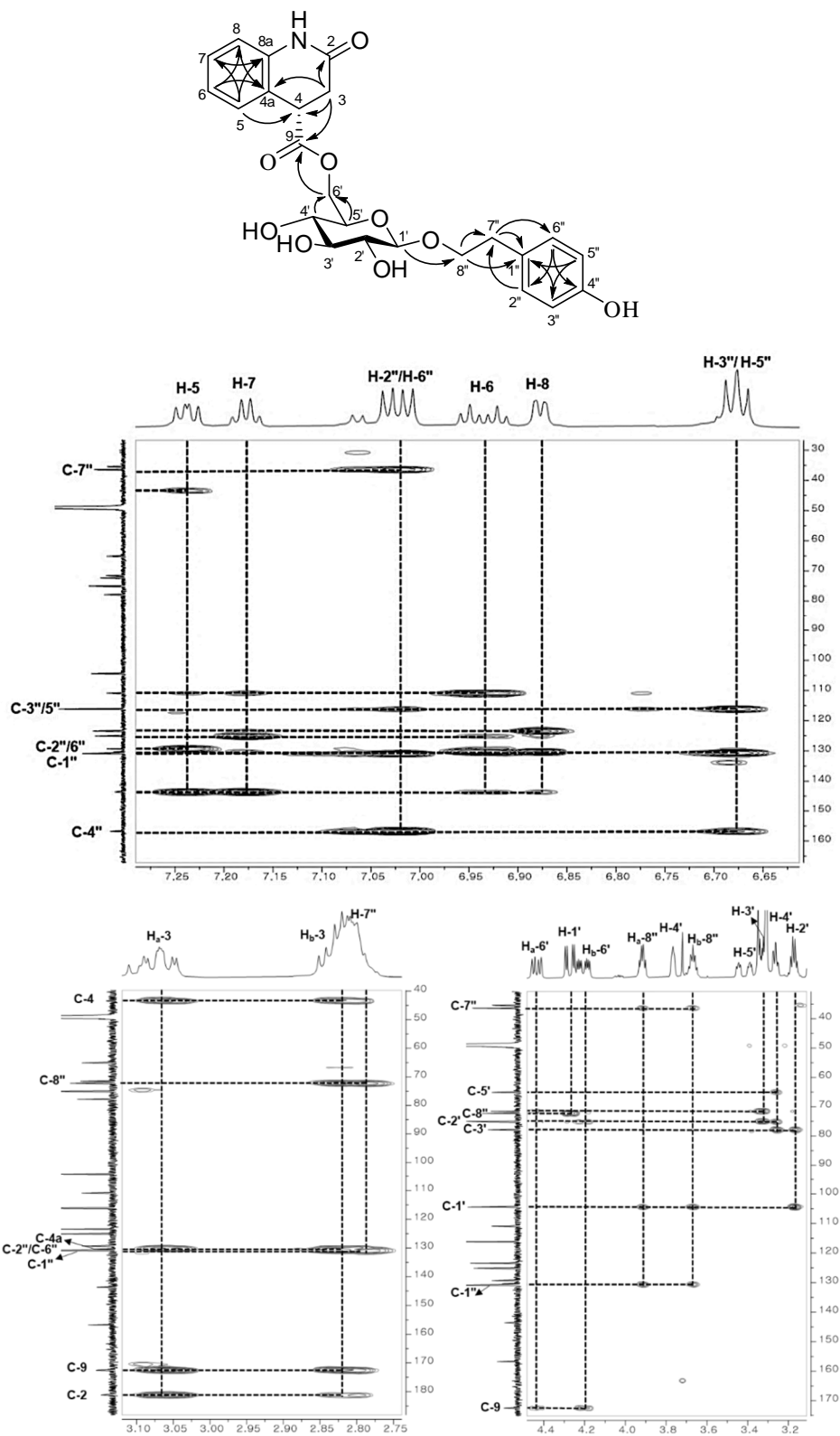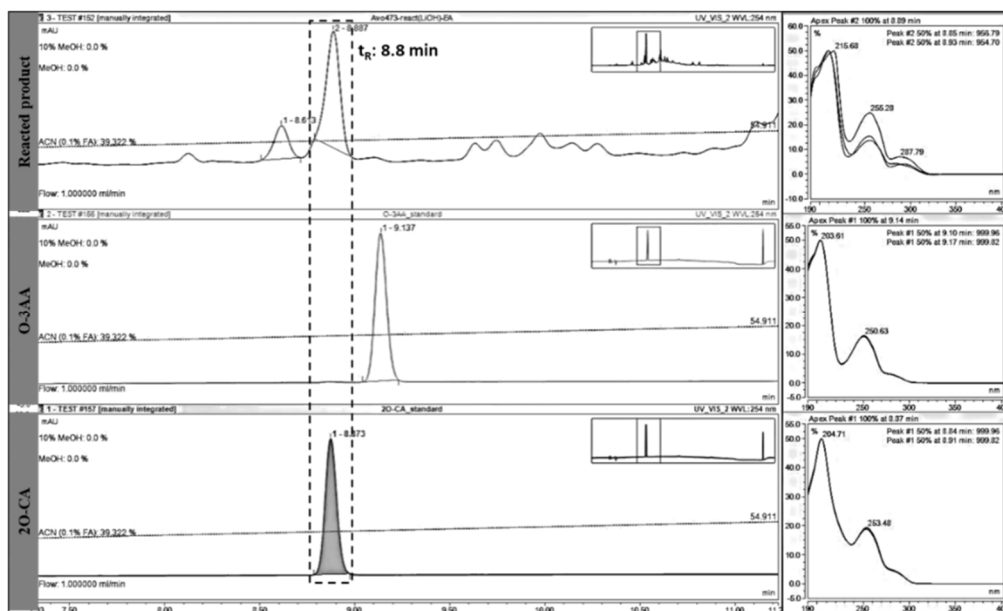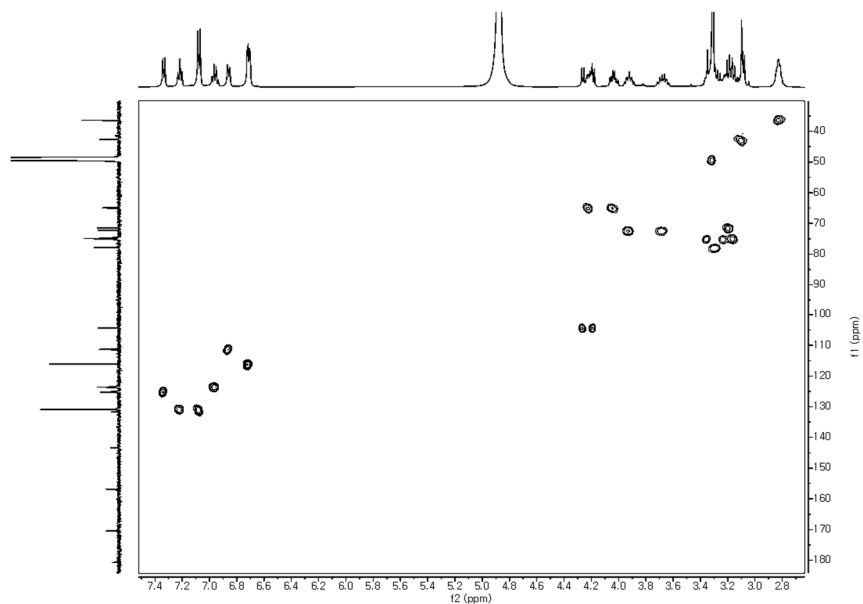
Figure 117. HSQC spectrum of compound **6**
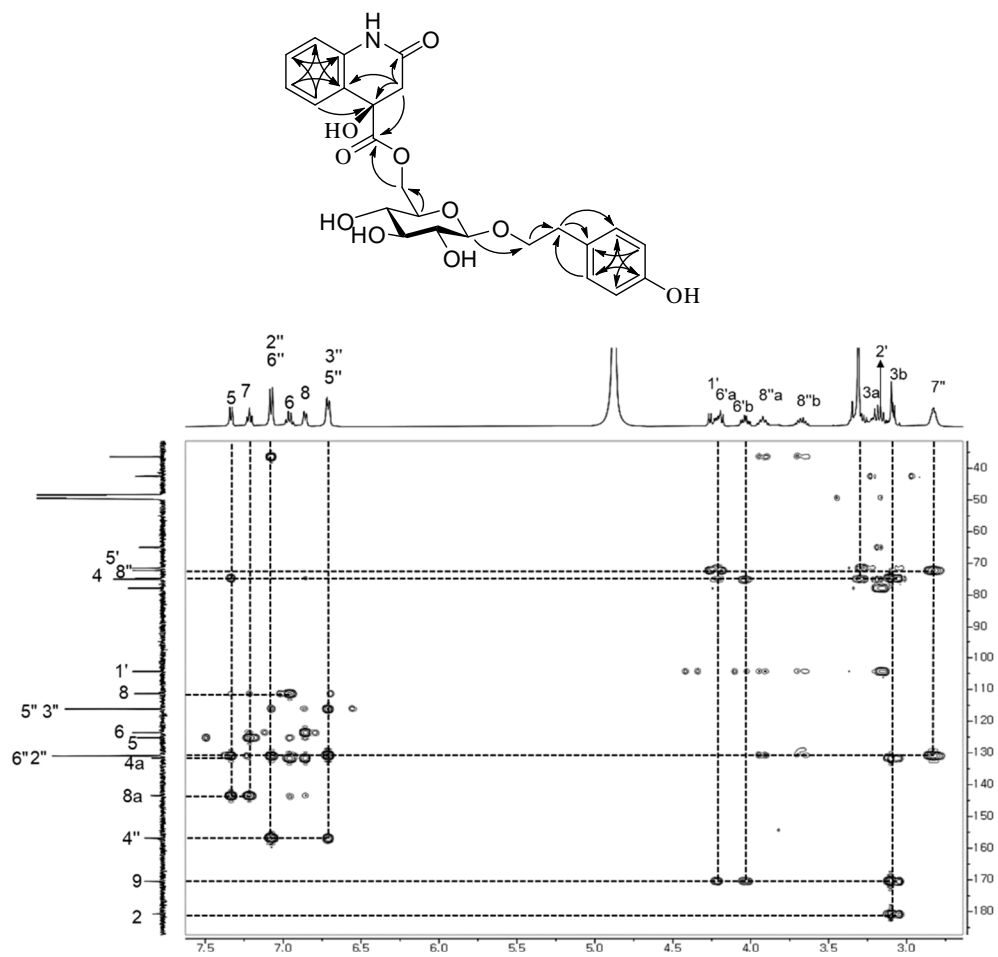
Figure 118. HMBC spectrum of compound **6**

135
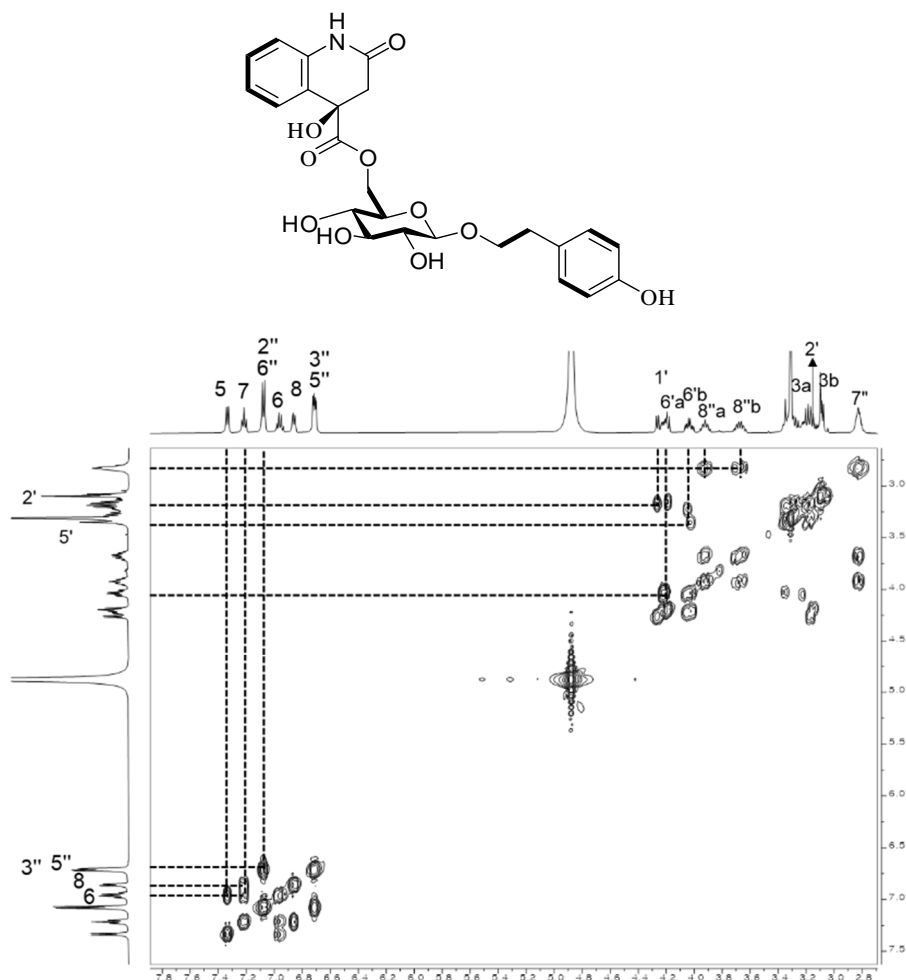
Figure 119. ^1H–^1H COSY spectrum of compound **6**
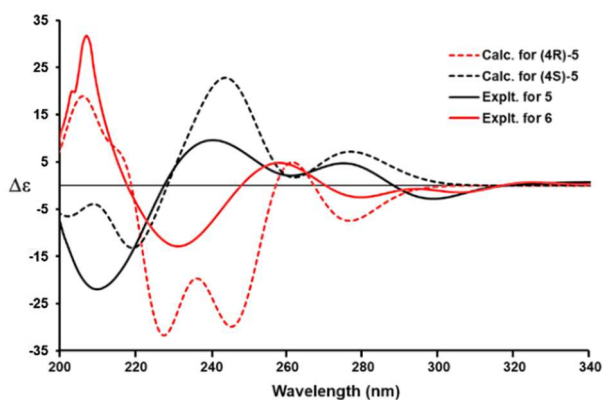
Figure 120. The experimental spectra of **5** and **6**, which were recorded directly after the compounds were separated from their diastereomer mixtures and calculated ECD data for (4*R*)–**5** and (4*S*)–**5**. The calculated ECD curves were generated by SpecDis ver. 1.71.40 (Bruhn et al., 2013b).

3.7. Compound 7

Compound **7** was obtained as a pale yellow powder. Its molecular formula was established to be $C_{24}H_{25}NO_9$ based on the HRESIMS data at m/z 470.1440 [M – H]⁻, (calcd. for $C_{24}H_{24}NO_9$ 470.1451). The IR spectrum showed absorption bands for hydroxy or amine (3375 cm⁻¹), carbonyl (1713 cm⁻¹) functional groups, and aromatic rings (1613 and 1514 cm⁻¹). The ¹H and ¹³C NMR spectra of **7** shared similar patterns at the salidroside moieties of **1–6** except for the alkaloid unit. However, there were several differences, including the disappearance of an oxygenated quaternary carbon (δ_C 74.7 ppm) in the hetero ring of **5** and **6** the presence of an olefinic bond at δ_C 140.0 (C-4) and δ_C 122.1 (C-3); δ_H 6.76 (H-3) in **7**. The HMBC cross-peaks from H-3 to C-2 (δ_C 170.5 ppm), C4, and C-4a (δ_C 121.3 ppm), as well as HMBC correlations from H-5 (δ_H 8.47 ppm) to C-4, suggested the presence of a 2-oxo-1,2-dihydroquinoline-4-carboxyl moiety (Baba et al., 2020). The linkage between salidroside and the alkaloid moiety was observed by correlations from H₂-6' at δ_H 4.60 (d, J = 11.6 Hz) and 4.51 (dd, J = 11.6, 5.7 Hz) to C-9 (δ_C 166.8 ppm) (Figure 123). Therefore, the structure of **7** was identified as shown in Figure 121, and the compound was named avoquinoside G.
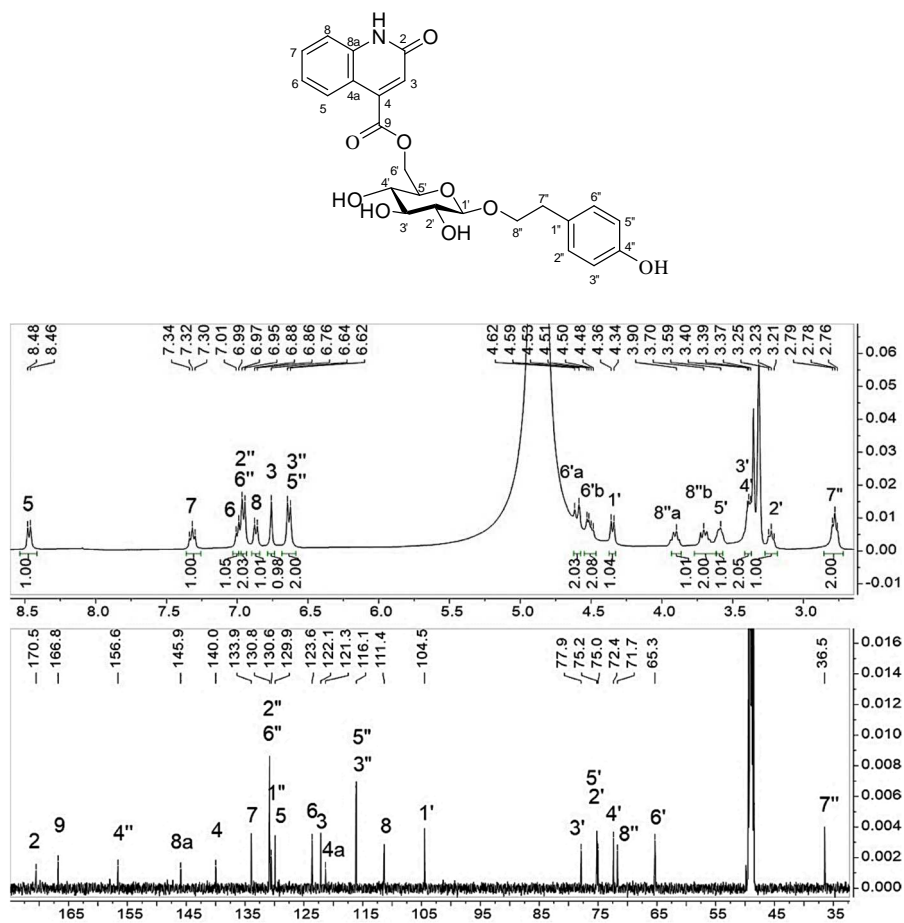
Figure 121. ^{1}H and ^{13}C NMR spectra of compound **7** (600/150 MHz, methanol-d_4)
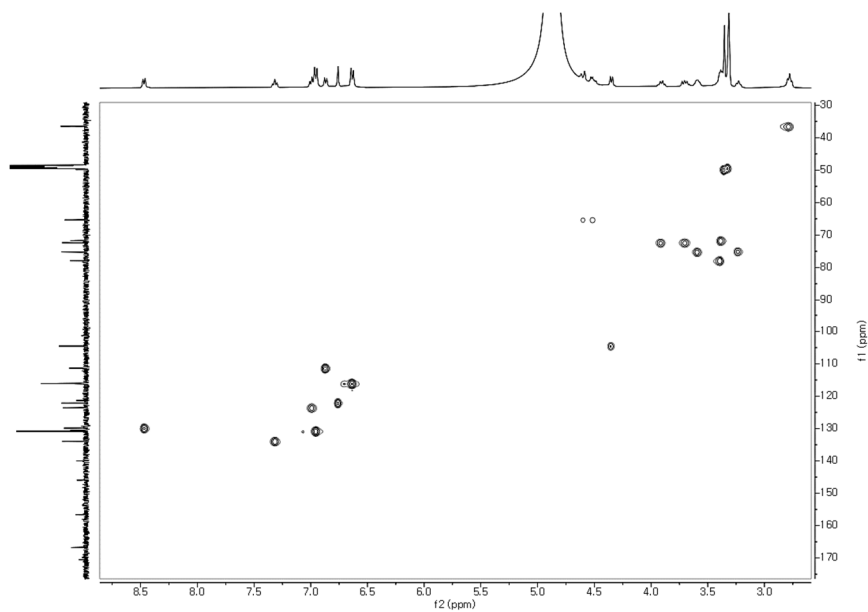
Figure 122. HSQC spectrum of compound **7**
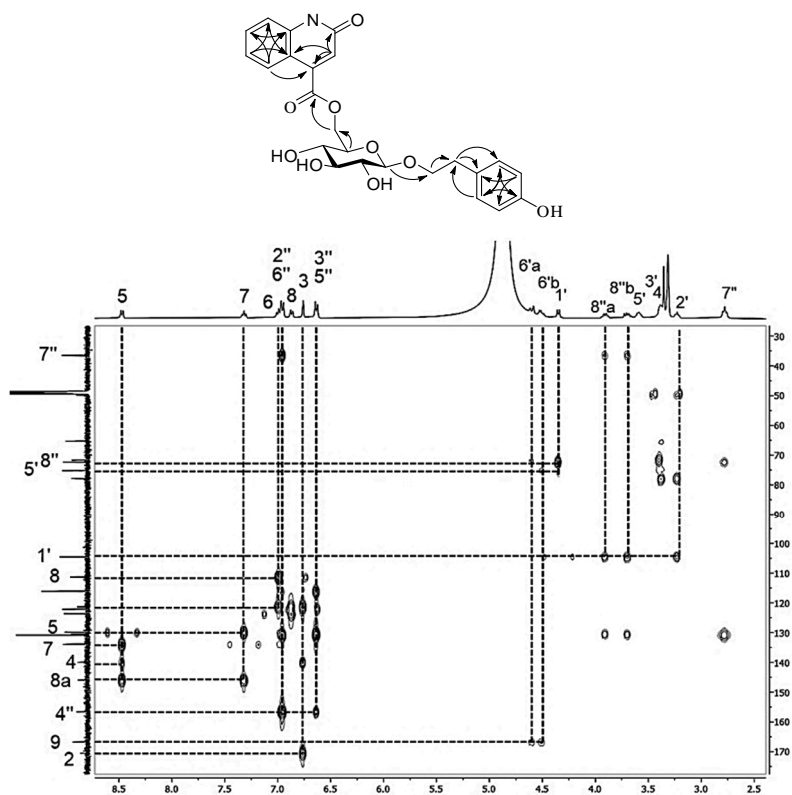
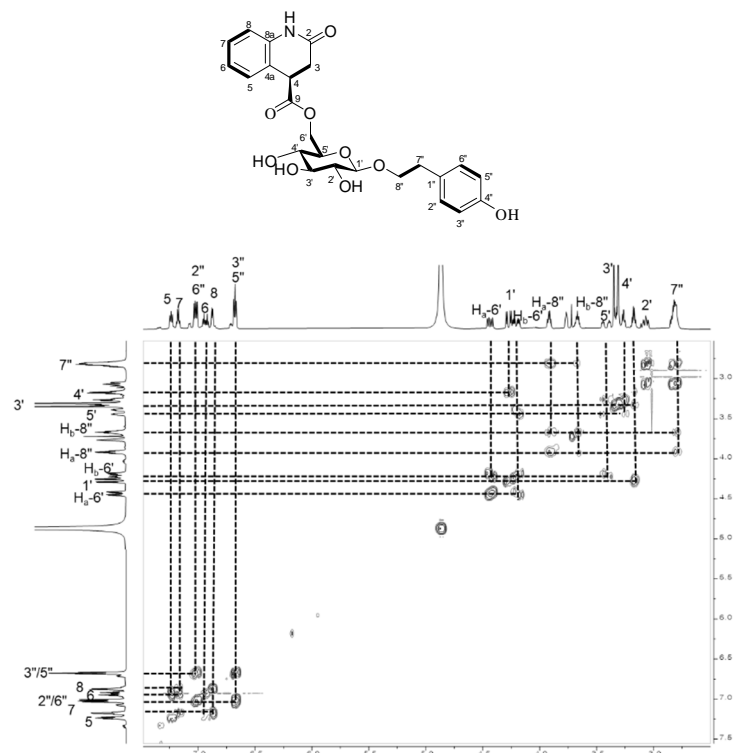**Figure 123.** HMBC spectrum of compound **7**
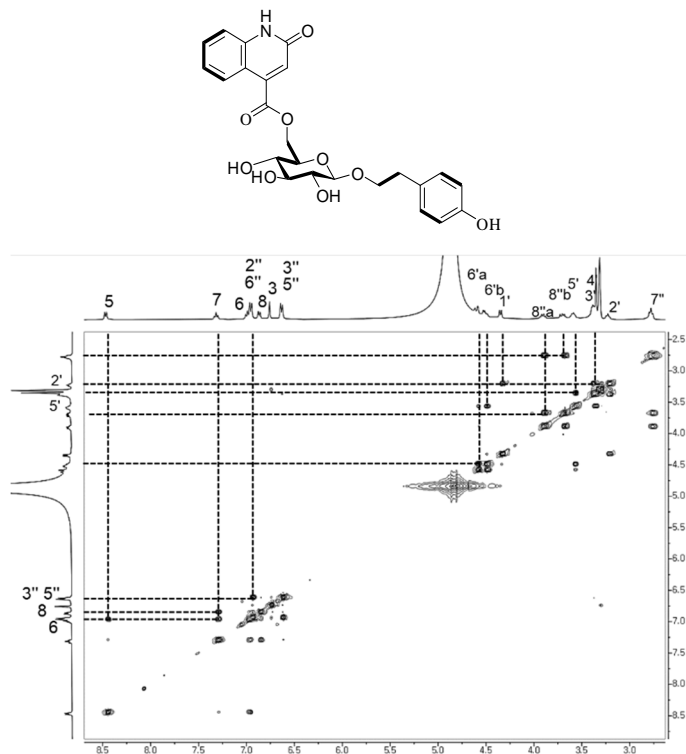
Figure 124. ¹H–¹H COSY spectrum of compound **7**

3.8. Compound **8**

Compound **8** was obtained as a pale yellow powder. Its molecular formula was established to be $C_{25}H_{29}NO_{11}$ based on the HRESIMS data at *m/z* 542.1646 [M + Na + H]$^+$, (calcd. for $C_{25}H_{29}NO_{10}Na$, 542.1638). The IR spectrum showed absorption bands for hydroxy or amine (3370 cm^{-1}), carbonyl (1716 cm^{-1}) functional groups, and aromatic rings (1602 and 1519 cm^{-1}). Interestingly, the ^{1}H and ^{13}C NMR data of **8** revealed an additional methoxy group at δ_H 3.12 (δ_C 50.7, C-11) and a carbon signal at δ_C 107.0 (C-4), suggesting the presence of a dioxaspiro skeleton. The positions of the methoxy and the dioxaspiro carbon (C-4) were supported by the HMBC correlations from H$_2$-10 (δ_H 3.27/3.20 ppm) to C-4, C-4a (δ_C 118.5), C-9 (δ_C 169.6), from MeO-11 (δ_H 3.12, s) to C-4, and from H-5 (δ_H 7.24, t, *J* = 7.8 Hz) to C-4. The upfield chemical shifts of the carbonyl group at C-9 were explained by the linkage with the methylene group (C-10). The chemical shift at δ_C 152.3 (C-2) also suggested the presence of a carbamide (NH-COO) at **8** by the disappearance of the amide group (NH-CO) in **1**–**6.** The chemical shifts of the alkaloid moiety in **8** showed similarities to those of the benzoxazinone skeleton that was previously reported in a natural product (Gala et al., 2009) and synthetic compounds (Kobayashi et al., 2011; Larin et al., 2021; Suárez-Castillo et al., 2012). The absolute configuration at C-4 in **8** was determined by comparing their experimental CD spectrum to the calculated ECD for two isomers (4*R*)–**8** and (4*S*)–**8**. As the CD data of **8** matched the Boltzmann-averaged computed ECD spectrum of (4*S*) with a (–) CE at 240 nm and (+) CE at 280 nm (Figure 133), the structure of **8** was identified and named avoxazinoside A.
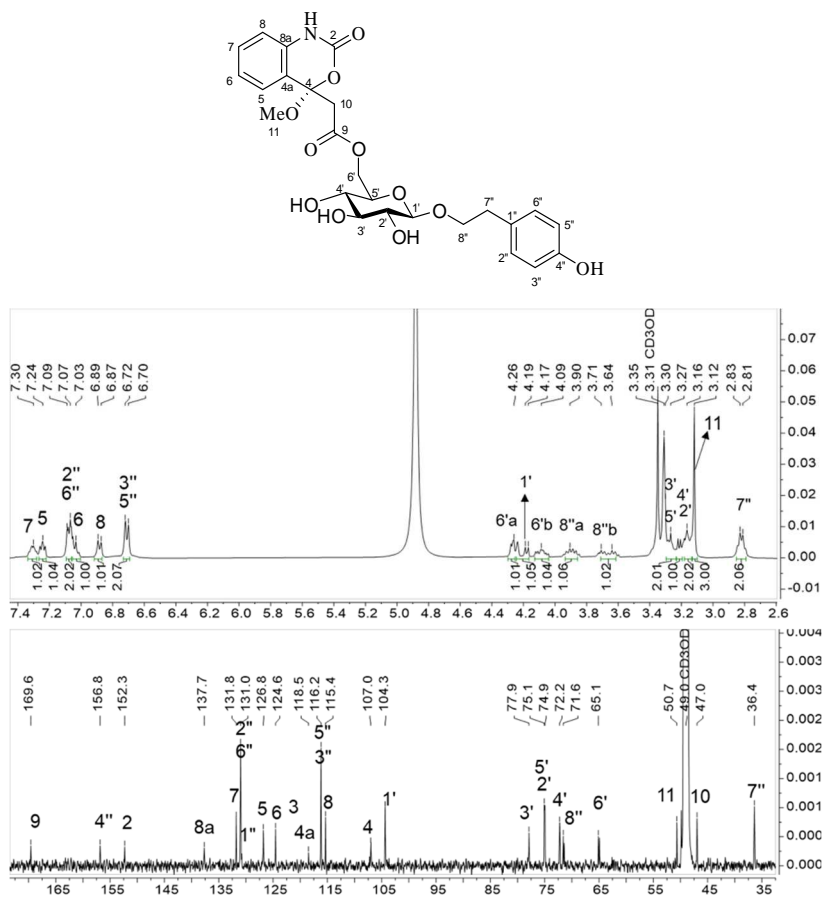
Figure 125. ^{1}H and ^{13}C NMR spectra of compound **8** (500/125 MHz, methanol-d_4)
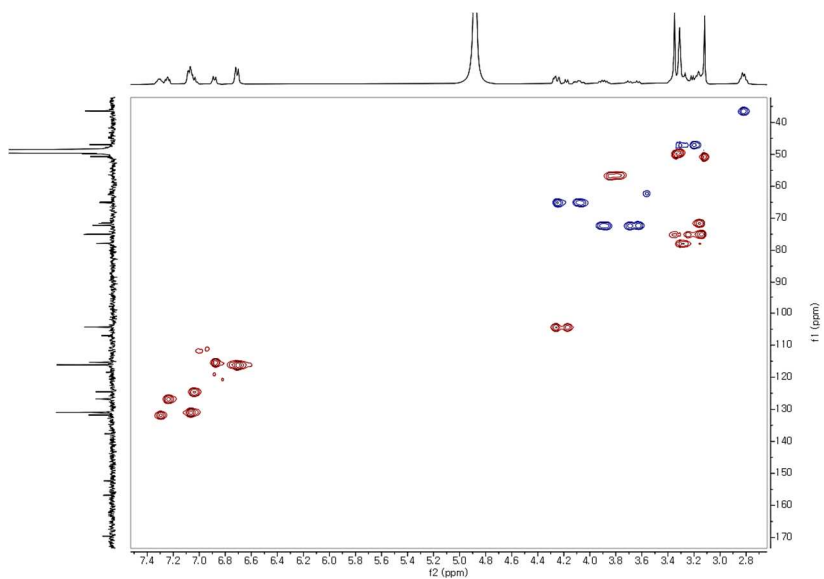
Figure 126. HSQC spectrum of compound **8**

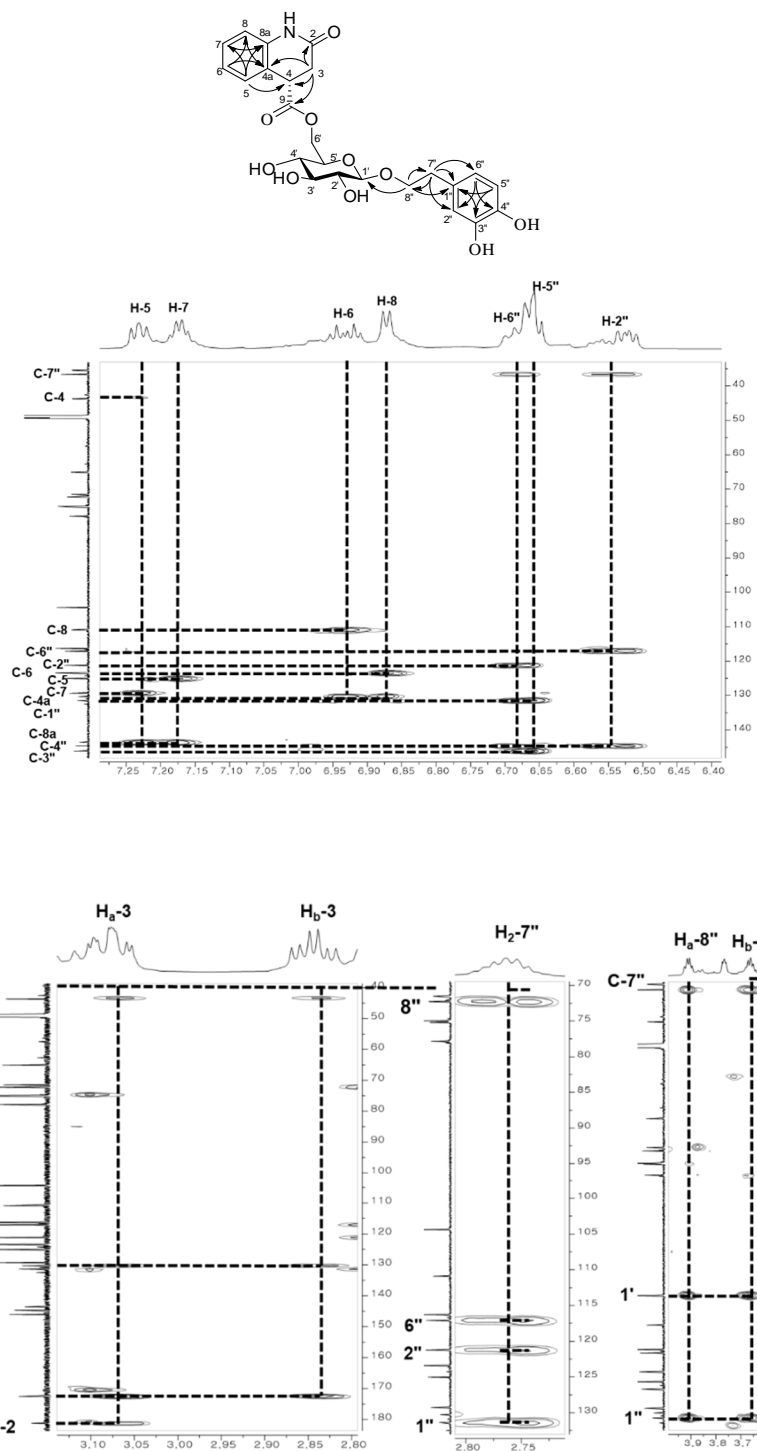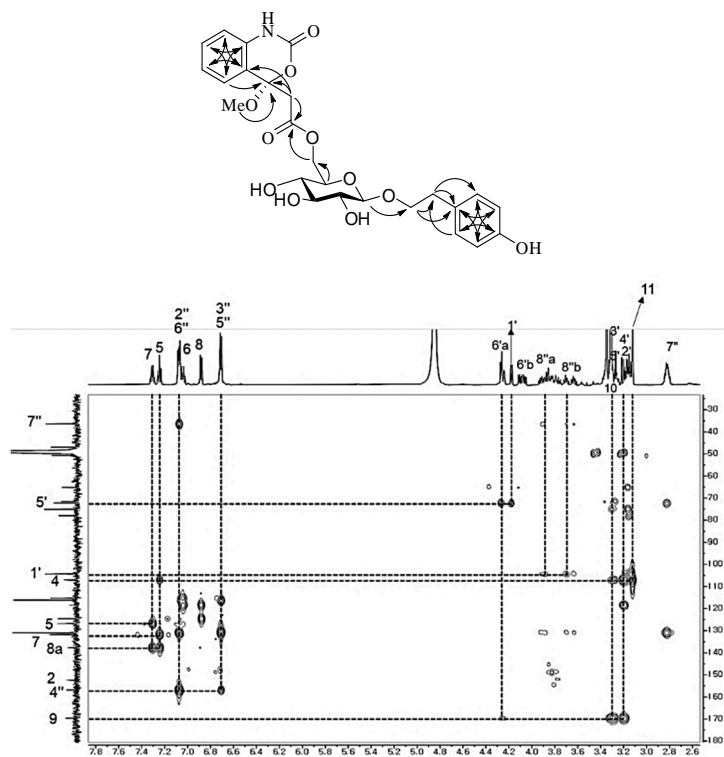**Figure 127.** HMBC spectrum of compound **8**
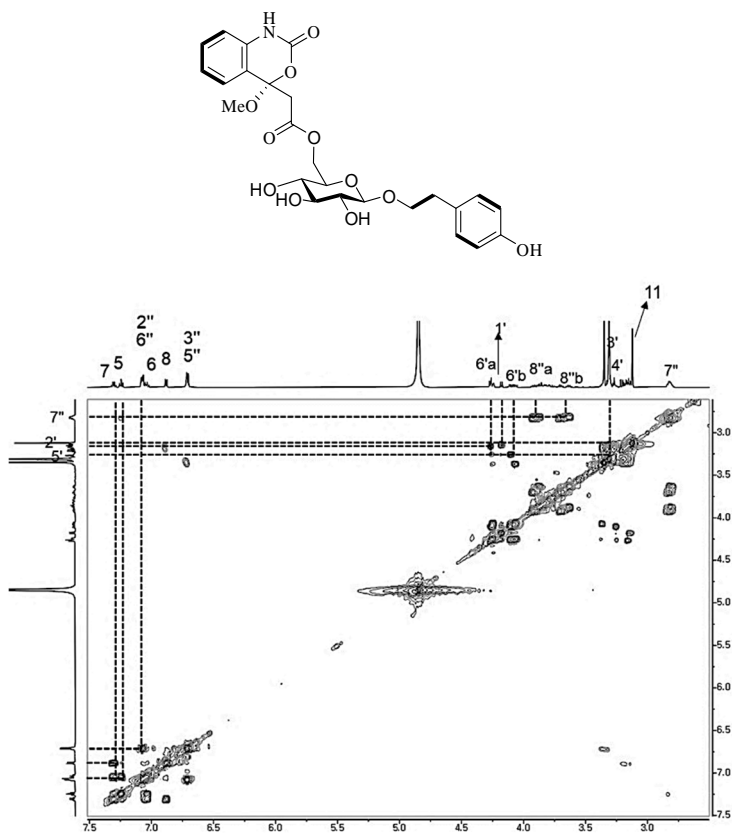
Figure 128. ¹H–¹H COSY spectrum of compound **8**

142

3.9. Compound 9

Compound **9** was isolated from the mixture of compound **8** diastereomers by an HPLC equipped with a YMC Triart phenyl hexyl column using a gradient of 45–48% MeOH/H$_2$O (0.1% formic acid). The absolute configuration of the two diastereomers was determined by comparing their experimental CD spectra to the calculated ECD for two isomers, (4*R*)–**8** and (4*S*)–**8**. The CD data of **9** were assigned as (4*R*), and the compound was named avoxazinoside B. Similar to three pairs of diastereomers (**1**–**6**), it was observed that **8** and **9** were also slowly converted to a mixture of the two compounds after 30 days (Figure 134).
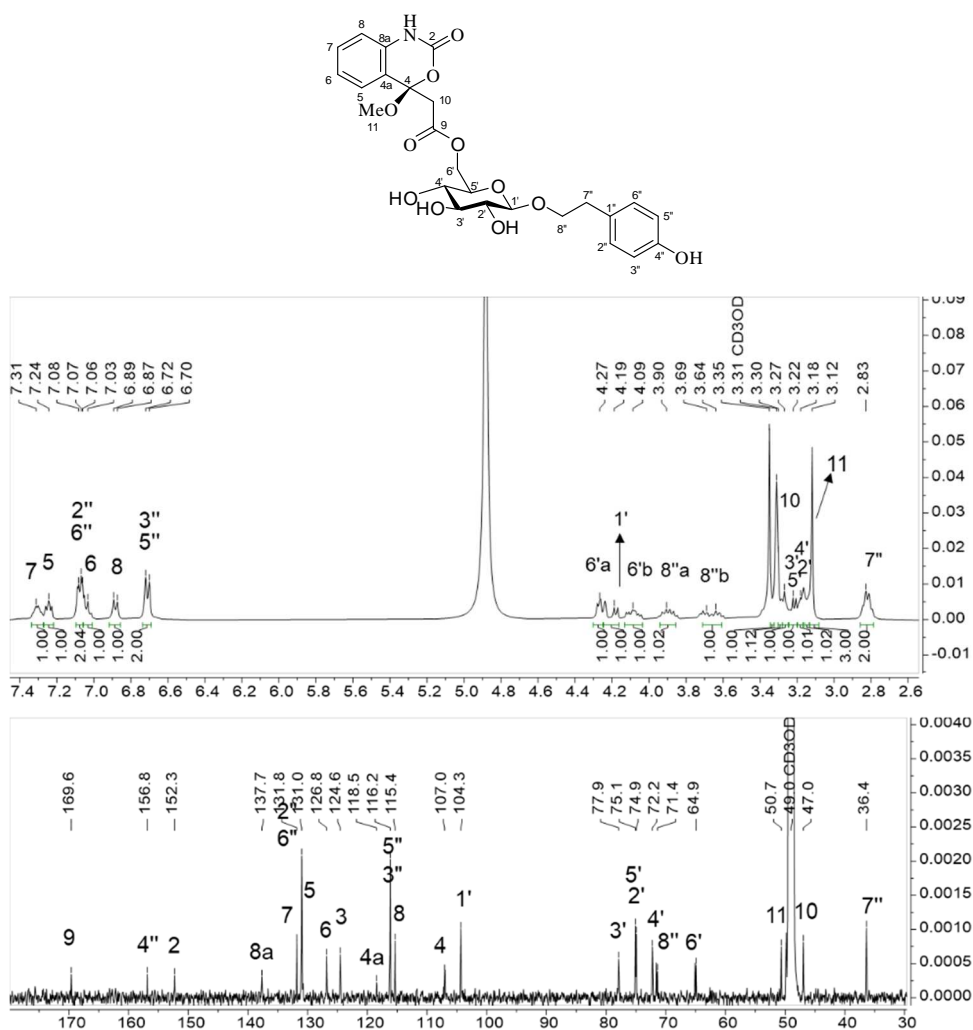
Figure 129. ¹H and ¹³C NMR spectra of compound **9** (600/150 MHz, methanol-*d*₄)

143
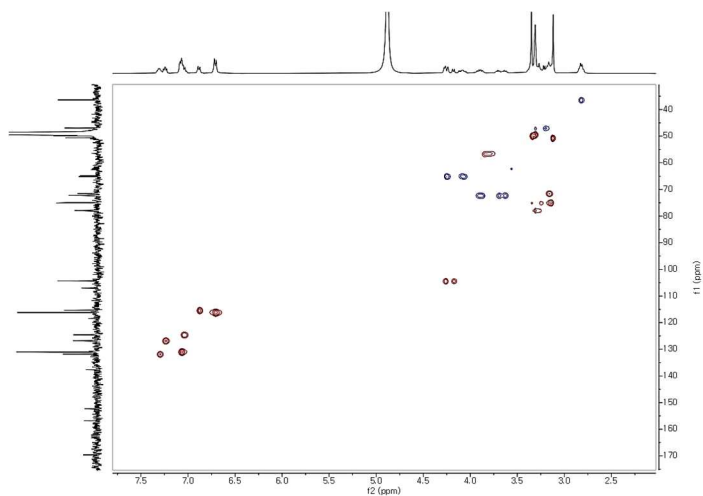
Figure 130. HSQC spectrum of compound **9**
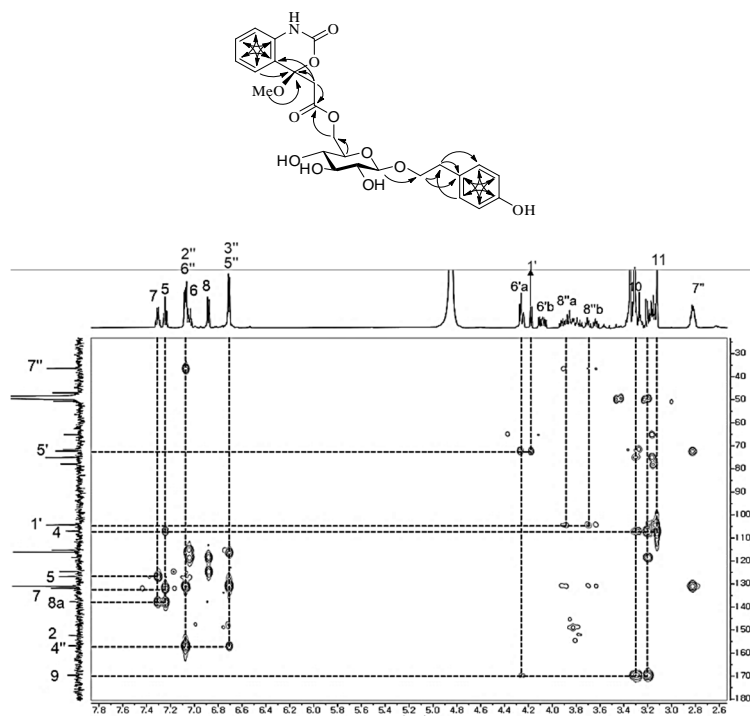
Figure 131. HMBC spectrum of compound **9**
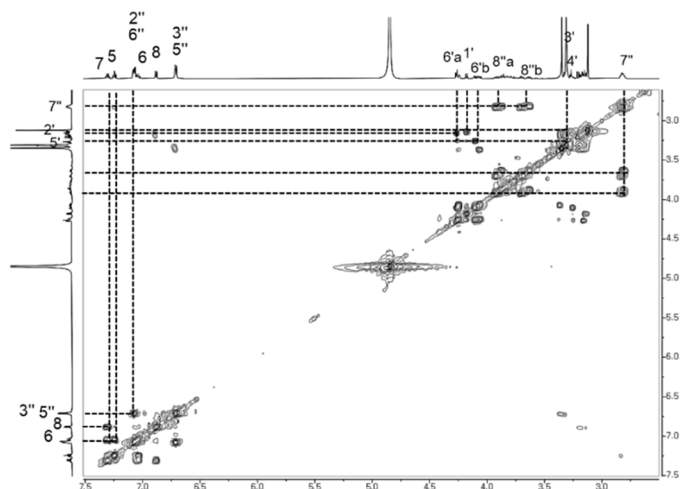
Figure 132. ^{1}H–^{1}H COSY spectrum of compound **9**
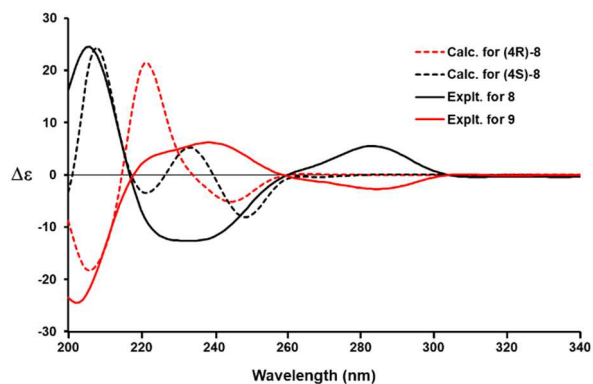
Figure 133. The experimental calculated ECD spectra of **8** and **9**. The experimental data were recorded directly after the compounds were separated from their diastereomer mixtures and calculated ECD data for (4*R*)–**8** and (4*S*)–**8**. The calculated ECD curves were generated by SpecDis ver. 1.71.40 (Bruhn et al., 2013b)
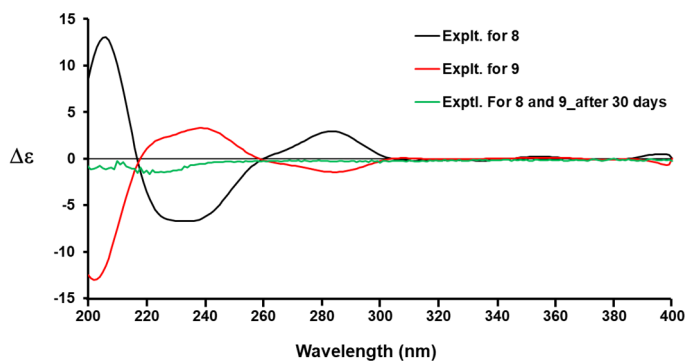
Figure 134. The CD experimental data of compounds **8** and **9** after 30 days

4. Biological activities of compounds 1–11

SIRT1 is an NAD$^+$-dependent deacetylase enzyme that removes acetyl groups from specific nuclear proteins and has been linked to the regulation of aging, cancer, and longevity (Brooks & Gu, 2009). The function of Sirt1 is mainly achieved by regulating a variety of transcription factors and transcriptional coregulators, including p53, Ku70, FoxO1, NF-κB, PPARγ, and p300 (Leibiger & Berggren, 2006). Among those factors, p53 is a crucial contribution to cellular senescence through the regulation of cell cycle arrest in response to oxidative stress and DNA damage (Yi & Luo, 2010). SIRT1-mediated p53 deacetylation prevents p53-dependent cellular senescence and has an important function in neuronal survival (Yi & Luo, 2010). Therefore, the identification of SIRT1 activators may be beneficial for the future treatment of neurodegenerative disorders.

To check the SIRT1 activities of all isolated compounds (1–11), the cytotoxicities of 1–11 at 20 μM were first confirmed by the MTT assay, and none of those compounds exhibited a cytotoxic effect at the tested concentration (Figure 135A). The effects of 1–11 on SIRT1 activity were determined using a p53-mediated transcriptional luciferase reporter cell-based assay. In particular, a luciferase reporter plasmid that contained upstream p53 binding sites (PG13-luc), a mammalian p53 expression vector (myc-tagged p53), and a SIRT1 expression vector (flag-tagged SIRT1) were cotransfected into HEK293 cells. Considering that SIRT1 negatively regulates p53, SIRT1 activators were detected to decrease the expression of the luciferase reporter in our assay, and resveratrol, which activates SIRT1 both *in vitro* and *in vivo* (Borra et al., 2005), was used as a positive control. As a result, among the eleven tested compounds, 1 significantly reduced the luciferase activity to approximately 45%, and its diastereomer, compound 2, only slightly decreased the luciferase activity to approximately 65% compared to those of the vehicle alone and the positive control, *trans*-resveratrol, at 20 μM (Figure 135B). Then, the concentration response was determined for compounds 1 and 2, and 1 activated SIRT1 deacetylation activity in a dose-dependent manner (Figure 135C). At 10 μM, downregulated luciferase activity, which was controlled by SIRT1 activation, was observed after treatment with compound 1 but not compound 2. Therefore, a greater potential for SIRT1 activation may occur with compound 1 than with compound 2 (Figure 135C). Moreover, SIRT1 activity is coupled to the hydrolysis of NAD$^+$;

therefore, the intracellular levels of NAD^+ and nicotinamide play an important role in regulating enzyme activity (Aguilar-Arnal et al., 2016). Furthermore, the NAD^+ to NADH ratio in the whole cell lysate of HEK293 cells that overexpress SIRT1 was tested to confirm the effect of **1** on SIRT1 activation. At 10 and 20μM, compound **1** significantly increased the NAD^+/NADH ratio in a dose-dependent manner compared to that of the negative control treated with DMSO (Figure 135D). Therefore, **1** could be a promising SIRT1 activator that activates SIRT1 deacetylation activity with an elevated NAD^+/NADH ratio.
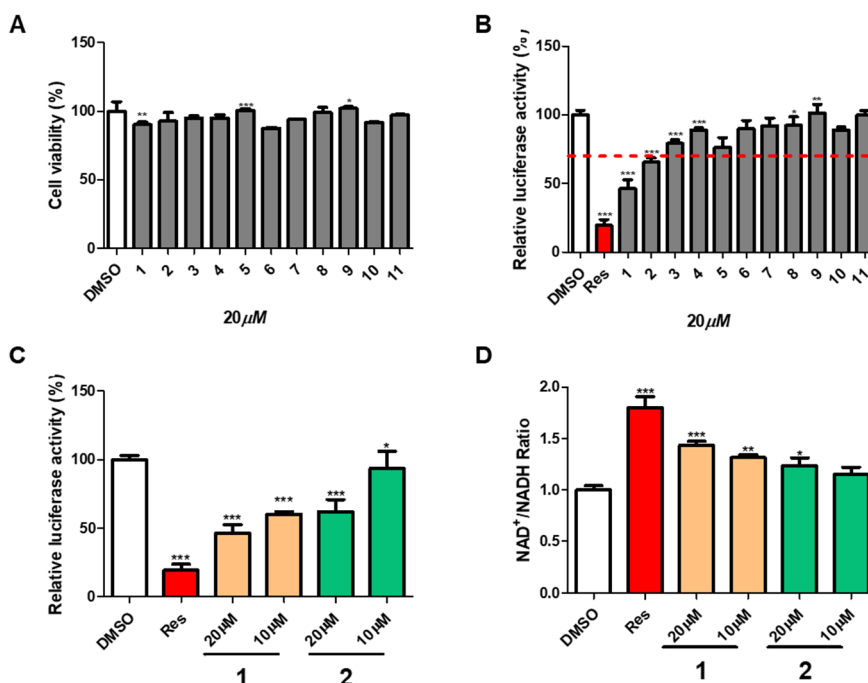
Figure 135. Effects of isolated compounds **1–11** on SIRT1 activities *in vitro*. (A) An MTT assay was performed to evaluate the cytotoxic effects of these compounds on HEK293 cells. The HEK293 cells were exposed to vehicle or 20 μM **1–11** for 24 h. The cell viability after the treatment was measured by the MTT assay. (B) HEK293 cells, after 24 h of cotransfection with four plasmids (PG13-luc, myc-p53, flag-SIRT1, and RSV-β-gal), were treated with DMSO, *trans*-resveratrol (20 μM), or the tested compounds (**1–11**). Compounds **1** and **2** reduced the luciferase activity that was controlled by SIRT1 activation. (C) Compound **1** activated SIRT1 in a dose-dependent manner. (D) Compound **1** elevated the NAD^+ to NADH ratio in HEK293 cells that overexpressed SIRT1. Data are presented as the mean ± SD (n = 3), $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ compared to the vehicle-treated samples.

For the first time, the alkaloid components from avocado seeds, which is the largest by-products of the nutrient fruit, were investigated and evaluated in their biological activities. Compound **1**, the most abundant among those alkaloids, showed the potential effect to activate SIRT1 that contributes to promoting healthy aging and regulating lifespan, as well as preventing cardiovascular disease, neurodegeneration, and cancer. Thus, this finding brings new insight into alkaloid types in avocado seeds and their potential for new drug discovery.

5. Materials and methods

5.1. Materials

5.1.1. Plant material

Ripe Hass avocados (*Persea americana* Mill.) were purchased from the local grocery store in March 2020 and were authenticated by Professor W. K. Oh. The avocado seeds were collected after removing the flesh and were stored. A voucher specimen (SNU2020-08) was deposited at the College of Pharmacy, Seoul National University, Seoul, Korea.

5.1.2. Equipment and software

Optical rotations were determined with a JASCO P-2000 polarimeter (JASCO International Co. Ltd., Tokyo, Japan). A Nicolet 6700 FT-IR spectrometer (Thermo Electron Corp., Waltham, MA, USA) was used to measure the IR spectra. The CD experiments were conducted using a Chirascan CD spectrophotometer, and the CD spectra were analyzed and visualized using Pro-Data Viewer software version 4.4.2.0 (Applied Photophysics, Leatherhead, UK). The conformational analysis was carried out with a Conflex 8 instrument (Conflex Corp., Tokyo, Japan). The molecular geometry analysis was simulated and visualized with TmoleX 4.3 and Turbomole (COSMOLogic GmbH, Leverkusen, Germany). Calculated spectra were plotted as sums of Gaussians with a 0.21 eV exponential half-width using the program Specdis (Bruhn et al., 2013b). The NMR data were analyzed using a JNM-ECA 800 MHz spectrometer (JEOL Ltd., Tokyo, Japan) coupled with a 5 mm CPTCI cryoprobe (Bruker, Germany) and an AVANCE 500 MHz spectrometer (Bruker, Billerica, MA USA) or a JNM-ECA 600 MHz spectrometer (JEOL Ltd., Tokyo, Japan). HPLC was carried out using a Gilson system with a UV detector (200 and 280 nm), an OptimaPak C_{18} column (10×250 mm, 5 μm particle size, RS Tech Korea), and a YMC Phenyl hexyl column (10×250 mm, 5 μm particle size, YMC Co., Ltd., Japan).

5.1.3. Chemicals and reagents

Silica gel (Merck, 63−200 μm particle size) and RP-18 (Merck, 75 μm particle size), Sephadex LH-20 from Sigma–Aldrich (St. Louis, MO, USA), and Daion HP 20 (Merck, 250−850 μm particle size) were used for column chromatography. TLC was developed with silica gel 60 F_{254} and RP-18 F_{254} plates from Merck. All solvents

were purchased from Daejung Chemicals & Metals Co. Ltd. (Si-heung, Korea).

5.2. Extraction and isolation

The air-dried avocado seeds (4.0 kg) were powdered and defatted by extraction with *n*-hexane (2 h × 3 times). The residue was sonicated with 70% EtOH (2 h × 3 times) at room temperature. The combined extract (270.0 g) was suspended in acetone to generate its precipitate (perseitol sugars) and solution. The solution layer was dried and suspended in water to partition with EtOAc and BuOH, respectively. The EtOAc fraction (90.0 g) was subjected to silica gel column chromatography, and the elution involved a gradient mixture of *n*-hexane/EtOAc (4:1 to 0:1) and then EtOAc/MeOH (8:1 to 0:1) to afford ten fractions (N1–N10). Fraction N8 was successfully separated with open RP C_{18} column chromatography using MeOH/H_2O (20–100%) to produce 12 subfractions (R1–R12). Fraction N9 (8.0 g) was subjected to RP-C_{18} column chromatography by a gradient system of 30–100% MeOH to obtain nine subfractions (R1–R9). Fraction R4 (1.1 g) was applied to a Sephadex LH-20 column, which was eluted with 100% MeOH to yield six subfractions (L1–L6). Fraction L3 (50.5 mg), which contained a mixture of two diastereomers, was further isolated by semipreparative HPLC (YMC Triart C_{18}, 10 × 250 mm, 45% to 50% MeOH in H_2O (0.1% formic acid) to yield compounds **7** (3.0 mg) and **8** (2.9 mg). Fraction R5 (2.5 g) was loaded onto Sephadex LH-20 using 100% MeOH to produce 10 subfractions (L1–L10). Subfraction L3 (80.5 mg) was separated by preparative HPLC equipped with an OptimaPak column (10 × 250 mm, 5 μm particle size, RS Tech, Korea) to obtain a mixture of compounds **3** and **4**. The semipreparative HPLC with a YMC Triart phenyl hexyl column (YMC Triart C_{18}, 10 × 250 mm) on two pairs of two diastereomers (compounds **3–4** and **5–6**) yielded compounds **3** (1.7 mg) and **4** (2.0 mg), as well as compounds **5** (2.7 mg) and **6** (5.0 mg). Similarly, fraction L4 (50.0 mg) was subjected to semipreparative HPLC with an Optima Pak C_{18} column (10 × 250 mm, 5 μm particle size) to obtain compounds **1** and **2** (15.4 mg) as a mixture of diastereomers. Then, this mixture was applied to MPLC (YMC Triart Phenyl, 45 to 55% MeOH in H_2O with 0.1% formic acid) to produce compounds **1** (5.4 mg) and **2** (5.5 mg). Fraction L5 (44.0 mg) was subjected to HPLC with an Optima Pak C_{18} column to yield compound **9** (3.7 mg). Fraction R8 (200.0 mg) was subjected to passage over a Sephadex LH-20 column using 100% MeOH

to produce seven subfractions. Fraction R8L3 was further purified by semipreparative HPLC (MeCN/H$_2$O, 25/75) to yield compound **10** (9.1 mg).
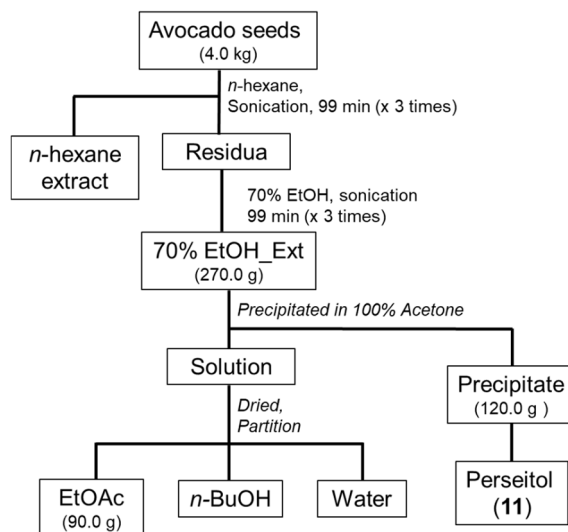
Scheme 2. Extraction scheme of avocado seeds
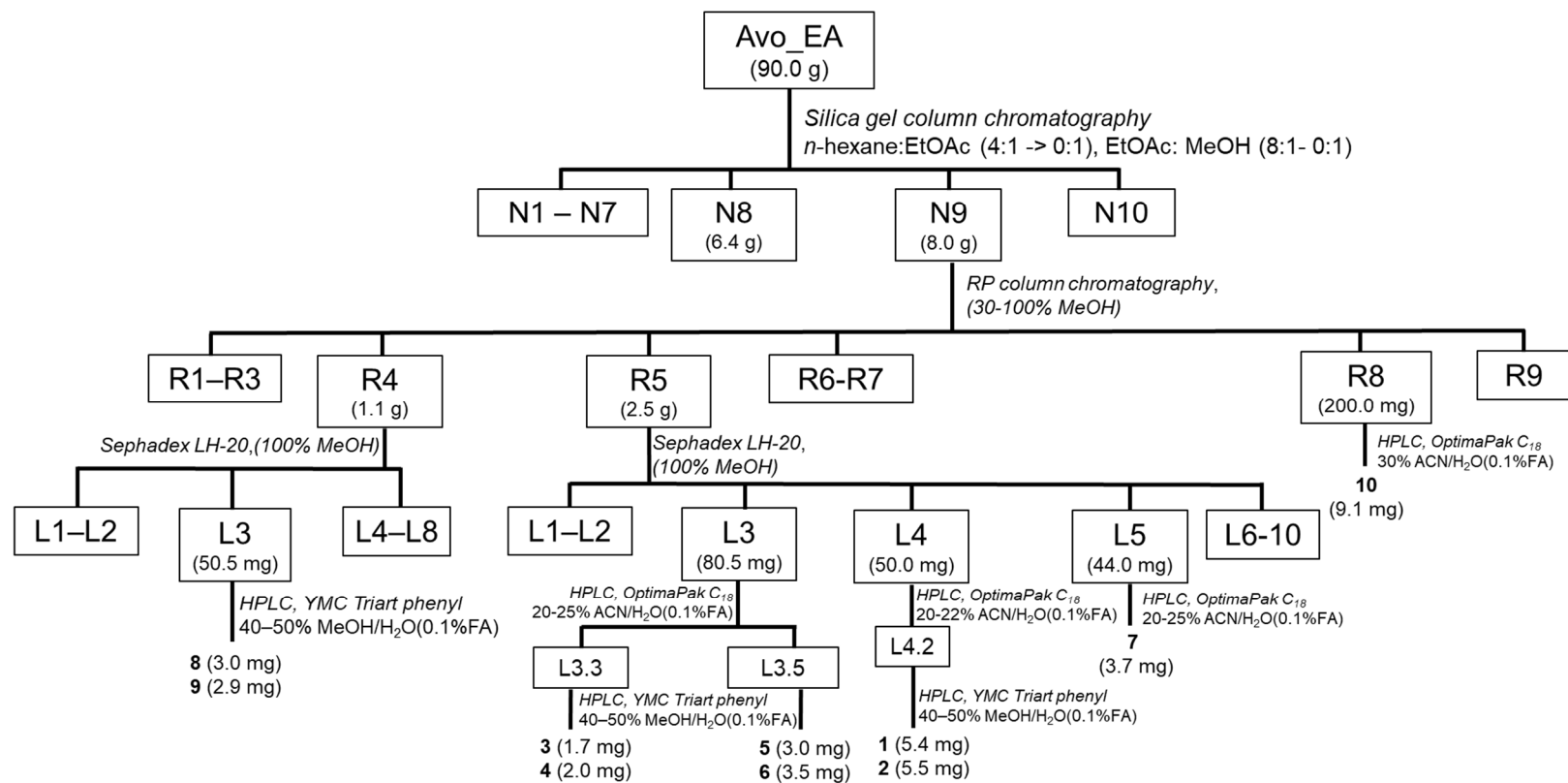
Scheme 3. Isolation scheme of compounds **1–11** from avocado seeds

5.3. Physical and chemical characteristics of isolated compounds

5.3.1. Compound **1**

Avoquinoside A **(1)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +71 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 230 (1.10), 280 (0.43);

IR (KBr) v_{max} 3339, 2925, 2884, 2855, 1720, 1618, 1517, 1472, 1442, 1361, 1235, 1205, 1079, 1030, 1024, 832, 756 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 472.1617 [M − H]$^-$, (calcd. for $C_{24}H_{26}NO_9$ 472.1608).

5.3.2. Compound **2**

Avoquinoside B **(2)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +71 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 230 (1.10), 280 (0.43);

IR (KBr) v_{max} 3339, 2925, 2884, 2855, 1720, 1618, 1517, 1472, 1442, 1361, 1235, 1205, 1079, 1030, 1024, 832, 756 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 472.1617 [M − H]$^-$, (calcd. for $C_{24}H_{26}NO_9$ 472.1608).

5.3.3. Compound **3**

Avoquinoside C **(3)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +67 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 220 (1.74), 280 (1.00);

IR (KBr) v_{max} 3344, 2887, 2357, 2312, 1722, 1607, 1512, 1473, 1269, 1235, 1205, 1076, 1036, 754 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 488.1518 [M − H]$^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557).

5.3.4. Compound **4**

Avoquinoside D **(4)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +67 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 220 (1.74), 280 (1.00);

IR (KBr) v_{max} 3344, 2887, 2357, 2312, 1722, 1607, 1512, 1473, 1269, 1235, 1205, 1076, 1036, 754 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 488.1505 [M – H]$^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557).

5.3.5. Compound **5**

Avoquinoside E **(5)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +68 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε220 (1.73), 280 (0.70);

IR (KBr) v_{max} 3370, 2933, 2887, 2853, 2381, 2357, 2307, 1722, 1612, 1518, 1473, 1443, 1364, 1269, 1081, 1036, 833, 749 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 488.1557 [M – H]$^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557).

5.3.6. Compound **6**

Avoquinoside F **(6)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +65 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 220 (1.73), 280 (0.70);

IR (KBr) v_{max} 3343, 2889, 2380, 2630, 2310, 1714, 1618, 1603, 1512, 1472, 1442, 1336, 1270, 1200, 107, 1044, 1024, 756 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS *m/z* 488.1565 [M – H]$^-$, (calcd. for $C_{24}H_{26}NO_{10}$ 488.1557).

5.3.7. Compound **7**

Avoquinoside G **(7)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +57 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 225 (1.04), 252 (1.20), 320 (0.70);

IR (KBr) v_{max} 3375, 2891, 2851, 2377, 2357, 2347, 2307, 1713, 1653, 1613, 1558, 1513, 1464, 1439, 1359, 1324, 1200, 1079, 750 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS m/z 470.1440 [M – H]$^-$, (calcd. for $C_{24}H_{24}NO_9$ 470.1451).

5.3.8. Compound **8**

Avoxazinoside A **(8)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +63 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 220 (1.79), 240 (1.48), 280 (1.00);

IR (KBr) v_{max} 3370, 2888, 2382, 2307, 1716, 1602, 1519, 1439, 1365, 1255, 1111, 1082, 1041, 1022, 828, 759 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS m/z 542.1646 [M+Na+H]$^+$, (calcd. for $C_{25}H_{29}NO_{10}Na$, 542.1638).

5.3.9. Compound **9**

Avoxazinoside B **(9)**

Pale yellowish amorphous powder;

$[\alpha]_D^{20}$ +63 (*c* 0.1, MeOH);

UV (MeOH) λ_{max} (log ε) 225 (1.40), 240 (1.34), 280 (0.95);

IR (KBr) v_{max} 3370, 2888, 2382, 2307, 1716, 1602, 1519, 1439, 1365, 1255, 1111, 1082, 1041, 1022, 828, 759 cm^{-1};

^{1}H and ^{13}C NMR data, see Tables 11, 12;

HRESIMS m/z 542.1629 [M+Na+H]$^+$, (calcd. for $C_{25}H_{29}NO_{10}Na$, 542.1638).

Table 11. ^{1}H NMR data of compounds **1–9** from *P. americana* in methanol-d_4 (mult., *J* in Hz)

Pos.	1[a]	2[a]	3[a]	4[a]	5[a]	6[c]	7[b]	8[c]	9[c]
3	3.07 (m)	3.07 (m)	3.08 (ddq, 16.5, 7.9, 4.2, 3.6)	3.08 (ddq, 16.5, 7.9, 4.2, 3.6)	3.09 (m)	3.09 (m)	6.76 (s)		
	2.84 (m)	2.84 (m)	2.84 (td, 16.5, 7.9)	2.84 (td, 16.5, 7.9)					
4	3.76 (t, 6.4)	3.77 (t, 6.4)	3.77 (dd, 7.9, 4.9)	3.77 (dd, 7.9, 4.9)					
5	7.24 (dd, 10.6, 7.3)	7.24 (dd, 10.7, 7.2)	7.23 (dd, 10.5, 7.5)	7.23 (dd, 10.5, 7.5)	7.33 (d, 7.4)	7.33 (d, 7.9)	8.47 (d, 7.6)	7.24 (t, 7.8)	7.24 (t, 7.8)
6	6.93 (m)	6.94 (m)	6.93 (m)	6.93 (m)	6.96 (q, 7.4)	6.96 (q, 7.9)	7.00 (d, 7.6)	7.04 (dd, 3.2, 7.3)	7.04 (dd, 3.2, 7.3)
7	7.18 (q, 7.3)	7.17 (q, 7.5)	7.17 (q, 7.2)	7.17 (q, 7.2)	7.22 (tt, 7.7, 1.5)	7.22 (t, 7.9)	7.32 (t, 7.5)	7.31 (q, 7.3)	7.31 (q, 7.3)
8	6.88 (dd, 7.7, 2.6)	6.88 (dd, 7.7, 2.8)	6.87 (d, 7.8)	6.87 (d, 7.8)	6.86 (dd, 7.9, 4.3)	6.86 (dd, 7.9, 2.4)	6.87 (d, 7.5)	6.88 (d, 7.9)	6.88 (d, 7.9)
10								3.27 (m) 3.20 (m)	3.27 (m) 3.20 (m)
11								3.12 (s)	3.12 (s)
1'	4.27 (d, 7.8)	4.27 (d, 7.8)	4.27 (m)	4.27 (m)	4.26 (d, 7.8)	4.26 (d, 7.9)	4.35 (d, 7.7)	4.18 (d, 7.8)	4.18 (d, 7.8)
2'	3.17 (q, 8.2)	3.17 (q, 8.3)	3.17 (q, 8.7)	3.17 (q, 8.7)	3.17 (dd, 7.5, 2.0)	3.16 (dd, 9.6, 1.6)	3.23 (t, 7.7)	3.14 (d, 3.7)	3.14 (d, 3.7)
3'	3.33 (m)	3.33 (q, 4.6)	3.34 (m)	3.34 (m)	3.28 (dt, 9.1, 4.5)	3.28 (m)	3.59 (dd, 13.6, 6.2)	3.28 (m)	3.28 (m)
4'	3.26 (td, 9.4, 9.4, 2.8)	3.27 (td, 8.1, 6.4, 2.8)	3.26 (t, 9.4)	3.26 (t, 9.4)	3.20 (d, 9.6)	3.20 (d, 9.6)	3.38 (d, 6.2)	3.16 (m)	3.16 (m)
5'	3.42 (m)	3.42 (m)	3.41 (m)	3.41 (m)	3.35 (br s)	3.35 (br s)	3.38 (d, 5.7)	3.25 (dd, 2.1, 5.9)	3.25 (dd, 2.1, 5.9)
6'	4.44 (m)	4.44 (m)	4.43 (q, 12.1)	4.43 (q, 12.1)	4.20 (m)	4.22 (ddd, 12.0, 5.8, 2.0)	4.60 (d, 11.6)	4.26 (m)	4.26 (m)
	4.21 (m)	4.21 (m)	4.20 (m)	4.20 (m)	4.04 (m)	4.03 (td, 12.0, 5.8)	4.51 (dd, 11.6, 5.7)	4.08 (m)	4.08 (m)
2"	7.02 (dd, 15.9, 8.3)	7.02 (dd, 15.9, 8.3)	6.53 (m)	6.53 (m)	7.08 (d, 8.2)	7.08 (d, 8.4)	6.96 (d, 8.3)	7.08 (dd, 4.1, 8.5)	7.08 (dd, 4.1, 8.5)
3"	6.68 (t, 9.1)	6.68 (t, 9.1)			6.71 (dd, 8.6, 5.01)	6.71 (dd, 8.4, 3.2)	6.63 (d, 8.3)	6.71 (dd, 2.3, 8.5)	6.71 (dd, 2.3, 8.5)
5"	6.68 (t, 9.1)	6.68 (t, 9.1)	6.66 (m)	6.66 (m)	6.71 (dd, 8.6, 5.01)	6.71 (dd, 8.4, 3.2)	6.63 (d, 8.3)	6.71 (dd, 2.3, 8.5)	6.71 (dd, 2.3, 8.5)

6"	7.02 (dd, 15.9, 8.3)	7.02 (dd, 15.9, 8.3)	6.68 (dd, 13.7, 2.3)	6.68 (dd, 13.7, 2.3)	7.08 (d, 8.2)	7.08 (d, 8.4)	6.96 (d, 8.3)	7.08 (dd, 4.1, 8.5)	7.08 (dd, 4.1, 8.5)
7"	2.81 (m)	2.81 (m)	2.76 (m)	2.76 (m)	2.83 (m)	2.82 (qd, 7.2, 6.2, 3.7)	2.78 (t, 8.3)	2.82 (m)	2.82 (m)
8"	3.91 (td, 9.2, 8.8, 6.4)	3.92 (td, 9.2, 8.6, 6.4)	3.91 (q, 8.4)	3.91 (q, 8.4)	3.92 (ddt, 17.4, 8.6, 1.9)	3.92 (tt, 9.8, 7.4)	3.91 (dd, 16.8, 8.3)	3.88 (m)	3.88 (m)
	3.67 (td, 9.4, 9.2, 4.6)	3.67 (m)	3.67 (q, 8.4)	3.67 (q, 8.4)	3.67 (dt, 8.4, 1.7)	3.67 (dtd, 16.5, 9.8, 7.4)	3.69 (dd, 16.8, 8.3)	3.66 (m)	3.66 (m)

[a] Recorded at 800 MHz; [b] Recorded at 600 MHz; [c] Recorded at 500 MHz

Table 12. ^{13}C NMR data of compounds **1–9** from *P. americana* in methanol-d_4

Pos.	**1**[a]	**2**[a]	**3**[a]	**4**[a]	**5**[a]	**6**[c]	**7**[b]	**8**[c]	**9**[c]
2	181.1	181.1	181.2	181.1	180.7	180.7	170.5	152.3	152.3
3	35.5	35.4	35.5	35.4	42.6	42.6	122.1		
4	43.4	43.4	43.7	43.7	74.7	74.7	140.0	107.0	107.0
4a	130.6	130.3	130.3	130.3	131.6	131.7	121.3	118.5	118.5
5	125.1	125.1	125.0	125.0	125.2	125.3	129.9	126.8	126.8
6	123.4	123.4	123.4	123.4	123.6	123.6	123.6	124.6	124.6
7	129.3	129.3	129.3	129.3	130.9	130.9	133.9	131.8	131.8
8	110.9	110.9	110.9	110.9	111.4	111.4	111.4	115.4	115.4
8a	143.7	143.7	143.6	143.6	143.4	143.4	145.9	137.7	137.7
9	172.6	172.4	172.6	172.4	170.4	170.4	166.8	169.6	169.6
10								47.0	47.0
11								50.7	50.7
1'	104.4	104.4	104.4	104.4	104.3	104.4	104.5	104.3	104.3
2'	75.0	75.0	75.0	75.0	74.9	74.9	75.0	74.9	74.9
3'	77.8	77.8	77.9	77.8	77.8	77.8	77.9	77.9	77.9
4'	71.6	71.5	71.5	71.5	71.5	71.5	72.4	71.6	71.4
5'	75.2	75.2	75.2	75.2	75.0	75.2	75.2	75.1	75.1
6'	65.1	65.0	65.1	65.0	65.0	65.0	65.3	65.1	64.4
1"	130.92	130.9	131.4	131.4	130.7	130.8	130.6	130.8	130.8
2"	130.88	130.9	121.2	121.2	131.0	131.0	130.8	131.0	131.0
3"	116.1	116.1	146.1	146.1	116.1	116.2	116.1	116.2	116.2
4"	156.8	156.8	144.6	144.6	156.8	156.8	156.6	156.8	156.8
5"	116.1	116.1	116.3	116.3	116.1	116.2	116.1	116.2	116.2
6"	130.88	130.9	117.1	117.1	131.0	131.0	130.8	131.0	131.0
7"	36.4	36.4	36.7	36.7	36.4	36.4	36.5	36.4	36.4
8"	72.3	72.3	72.3	72.3	72.2	72.3	71.7	72.2	72.2

[a] Recorded at 200 MHz; [b] Recorded at 150 MHz; [c] Recorded at 125 MHz

5.4. UPLC-qTOF-MS/MS analysis and molecular networking generation

An aliquot (5.0 g) of the avocado seeds was defatted by *n*-hexane extraction and then extracted by 70% EtOH at room temperature under sonication (1 h, 3 times). The aliquot of combined extracts was dried and diluted in MeOH (HPLC grade) to a concentration of 2 mg/ml. The other extract was suspended in water and subsequently partitioned with *n*-hexane, EtOAc, and *n*-BuOH. The organic layers of each partition were also concentrated and prepared at a concentration of 2 mg/ml in MeOH. All samples were filtered through a 0.2 μm membrane and stored at 4 °C before use in UPLC-qTOF-MS/MS analysis. The MS/MS data were acquired from an Agilent 6530 Q-TOF mass spectrometer equipped with an Agilent 1260 Infinity UPLC (Agilent Technologies, Santa Clara, CA, USA). A YMC Triart C_{18} column (150 mm × 3.0 mm i.d., 5 μm) was used for separation. The elution gradient consisted of H_2O (A) and MeCN (B) (both were buffered with 0.1% formic acid) and was increased from 10% B to 90% B in 40 min (0.3 mL/min), held for 12 min (1 mL/min), returned to 10% B in 0.1 min, and then maintained for 5 min (0.3 mL/min). Nitrogen was used as the drying gas and collision gas in the ESI source. The electrospray ion source parameters were as follows: drying gas flow rate, 10 L/min; heated capillary temperature, 350 °C; sheath gas temperature, 350 °C; flow, 12 L/min; nebulizer pressure, 30 psi; and VCap, fragmentor, skimmer, and octopole RF peak voltages set at 4000, 180, 60, and 750 V, respectively. The detection was carried out in positive and negative electrospray ionization modes, and the spectra were recorded by MS scanning in the range of *m/z* 100–1000. The MS/MS analyses were carried out by targeted fragmentation, and the collision energy was set at 50 eV. Data files were converted to mzXML format by ProteoWizard software, version 3.0.20315-7da487568 (Chambers et al., 2012). Peak picking, chromatogram deconvolution, and other data processing for the MS/MS data were performed by MZmine2 software v32 (Pluskal et al., 2010). Eventually, the.mgf clustered spectral data file and its corresponding.csv metadata file (for RT, areas, and formula integration) were exported using the dedicated "Export for GNPS" and "Export to CSV file" built-in options. A molecular network was created using the online workflow at GNPS (http://gnps.ucsd.edu). The spectra in the network were then searched against GNPS spectral libraries. All matches between the network spectra and library spectra that were retained had to have a score above 0.7 and at least 3

matched peaks. The molecular networking data were analyzed and visualized using Cytoscape (ver. 3.6.0) (Shannon et al., 2003). All of the results and parameters can be accessed with the GNPS job id for molecular networking feature-based analysis at

https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=618fd6ab564f4d32aee9cb44d1fb8be7.

5.5. Deesterification reaction of compounds **1** and **2** to determine the alkaloid moiety
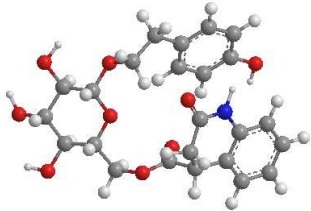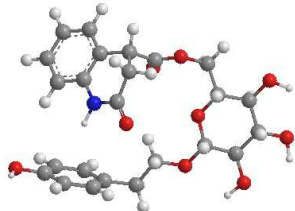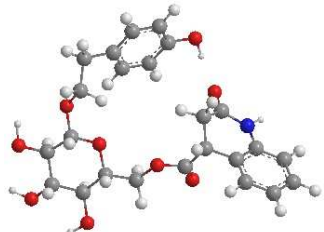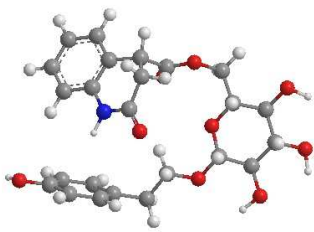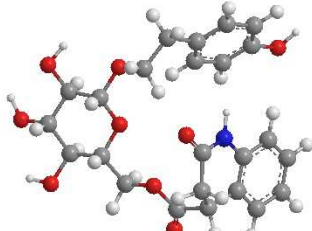
The mixture of compounds **1** and **2** (1.5 mg, 1 eq) was dissolved in MeOH (200 μL) and treated with 1 M LiOH (10 eq) in THF/H_2O (3:1) by stirring for 1 h. The products were neutralized by 1 N HCl to obtain the corresponding alkaloid moiety that contains carboxylic acid. After the reacted mixture was dried and partitioned by EtOAc/H_2O (3 times), the EtOAc layer was checked by HPLC for comparison with two standard compounds, 2-oxo-1,2,3,4-tetrahydroquinoline-4-carboxylic acid (2O-CA) and 2,3-dihydro-2-oxo-1H-indole-3-acetic acid (O-3AA). The HPLC conditions were set up as follows: a gradient of 10–90% MeCN/H_2O (0.1% formic acid) in 20 min, 100% MeCN in 5 min, and restabilization in 5 min at 10% MeCN/H_2O (0.1% formic acid). The UV detector was set up at four wavelengths of 201, 254, 280, and 330 nm. The alkaloid moiety of **1** and **2** was identified as 2O-CA as they shared the same retention time on the HPLC chromatogram at 8.8 min.

5.6. Computational ECD calculations

Conformational analyses of compounds **1**–**6**, **8**, and **9** were simulated using molecular mechanics force-field (MMFF94s) calculations with a search limit of 1.0 kcal/mol in Conflex 7 (Conflex Corp., Tokyo, Japan). The conformers with a Boltzmann population of over 1% were geometrically optimized by TmoleX 4.3 and Turbomole (COSMOLogic GmbH, Leverkusen, Germany) at the def-SV(P) basis set for all atoms and the B3LYP functional level in the gas phase. The computational ECD spectra of the optimized conformers were represented by time-dependent density functional theory (TDDFT) using 6-31G/B3LYP according to the Boltzmann distributions. The ECD data were generated using Gaussian functions for each transition (σ is the width of the band at a height of 1/e). The overall calculated ECD curve that was generated by using SpecDis 1.71 version 1.71 (Bruhn et al., 2013b)

was used for comparison to experimental data.
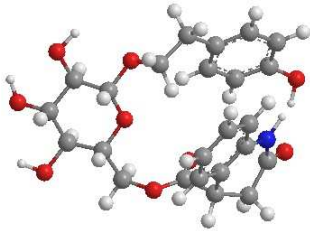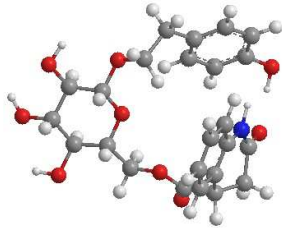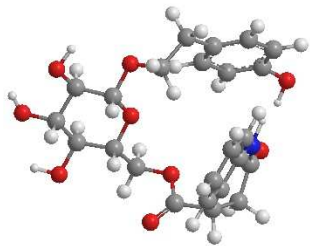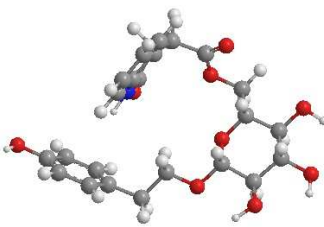
Table 13. Gibbs free energies and equilibrium populations of low-energy conformers

of 4*R*-**1**

Conformers	In MeOH	
4*R* - **1**	ΔG^{a}	$P(\%)^{b}$
I	0.00	1.16
II	0.29	0.71
III	-2.62	95.94
IV	0.25	0.76
V	-0.13	1.43

[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm
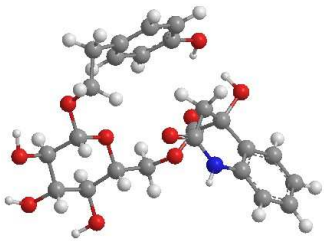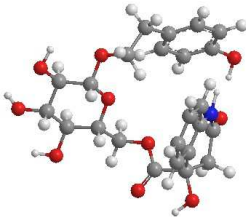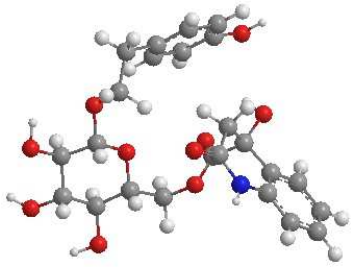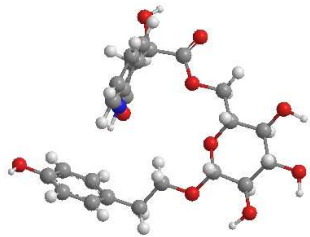
Table 14. Gibbs free energies and equilibrium populations of low-energy conformers

of 4*S*-**1**

Conformers	In MeOH	
4*S*-**1**	ΔG[a]	$P(\%)$[b]
I	0.00	78.27
II	1.06	13.01
III	1.31	8.65
IV	4.14	0.07
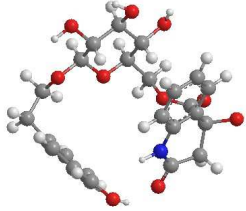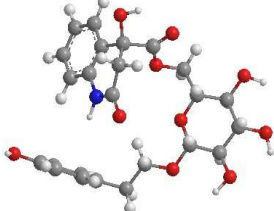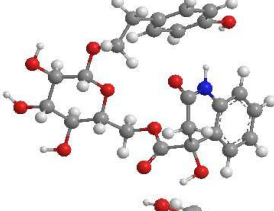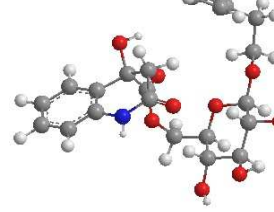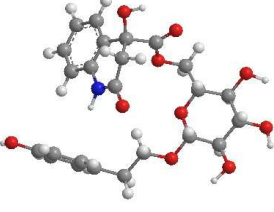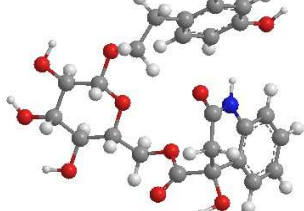
[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm

Table 15. Gibbs free energies and equilibrium populations of low-energy conformers

of 4*R*-**5**

Conformers	In MeOH	
4*R* - **5**	ΔG^a	$P(\%)^b$
I	0.00	53.45
II	0.23	35.98
III	0.98	10.26
IV	3.04	0.31
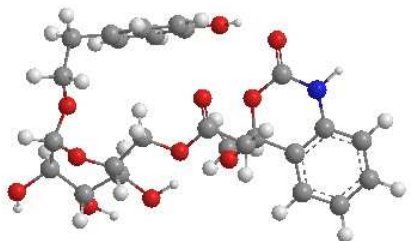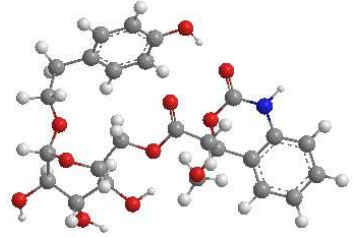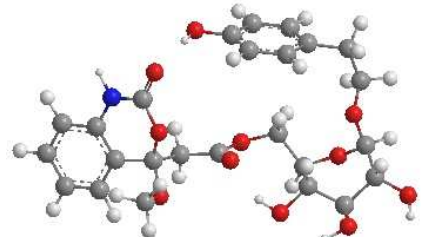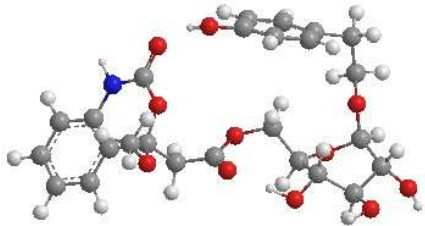
[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm

Table 16. Gibbs free energies and equilibrium populations of low-energy conformers of 4*S*-**5**

Conformers	In MeOH	
4*R* - **5**	ΔG^{a}	$P(\%)^{b}$
I	0.00	95.26
II	5.17	0.02
III	3.99	0.11
IV	1.81	4.52
V	5.31	0.01
VI	4.24	0.07
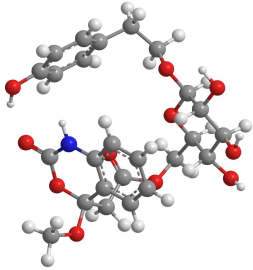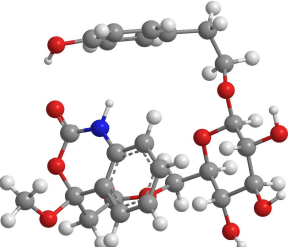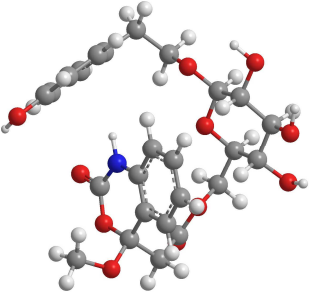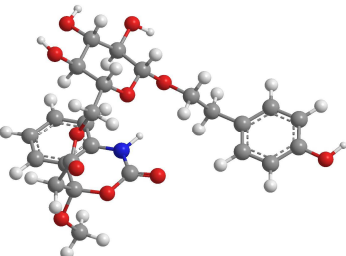
[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm

Table 17. Gibbs free energies and equilibrium populations of low-energy conformers of 4*R*-**8**

Conformers	In MeOH	
4*R*-**8**	ΔG[a]	$P(\%)$[b]
I	0.00	87.34
II	1.18	11.92
III	4.35	0.06
IV	2.88	0.68

[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm

Table 18. Gibbs free energies and equilibrium populations of low-energy conformers

of 4*S*-**8**

Conformers	In MeOH	
4*S* - **8**	ΔG^a	$P(\%)^b$
I	0.00	92.44
II	1.55	6.81
III	3.64	0.20
IV	3.04	0.55

[a] B3LYP/6-31G, in kcal/mol, [b] From ΔG values at 298.15K and 1 atm

5.7. Sirt1 deacetylation assay with a luciferase reporter cell-based assay

HEK293 cells were maintained in high glucose Dulbecco's modified Eagle's medium (DMEM; Welgene, Deajeon, Korea) supplemented with 10% fetal bovine serum (FBS; HyClone, GE Healthcare Life Sciences, UK), 100 U penicillin, and 100 μg/mL streptomycin (HyClone). The HEK293 cells (5,000 per well) were seeded in 48-well plates and maintained overnight. The cells were transfected with 4 plasmids, PG13-luc plasmid (wt p53 binding sites), myc-tagged p53 plasmid (myc-p53), flag-tagged Sirt1 plasmid (flag-Sirt1), and RSV-β-galactosidase plasmid as an internal standard, to investigate SIRT1 deacetylation activity via a luciferase reporter. PEI transfection reagent (Polyscience, Inc., PA, USA) was employed to deliver the plasmids into cells. At five hours posttransfection, the medium was replaced, and the cells were treated with vehicle, resveratrol, or the compounds of interest for 24 h before luciferase and β-galactosidase expression was analyzed by using passive lysis buffer (Promega, WI, USA) and a firefly luciferase assay kit (Promega, WI, USA). The luciferase activity was normalized to β-galactosidase activity to obtain the final value representative of SIRT1 deacetylation activity.

5.8. NAD$^+$/NADH ratio measurements

For the overexpression of SIRT1, the flag-tagged Sirt1 plasmid was transfected into HEK293 cells and then treated with vehicle, resveratrol, or the compounds of interest for 12 h. The NAD$^+$/NADH quantification kit (Red Fluorescence, AAT Bioquest, Inc., CA, USA) was used to measure the NAD$^+$ to NADH ratio from the whole cell lysate according to the manufacturer's instructions. The fluorescence signal generated from the enzymatic cycling reaction was monitored at excitation and emission wavelengths of 540 and 590 nm, respectively, using a fluorescence microplate reader (Spectra Max GEMINI XPS, Molecular Devices, CA, USA).

5.9. Statistic analyses

GraphPad PRISM (GraphPad Software, La Jolla, CA, USA) was used for graphical representation and statistical calculation. Data are presented as the means $\pm$ standard deviations (SDs) of three or four independent experiments. Multiple comparisons were conducted with one-way analysis of variance (ANOVA), followed by Dunnett's multiple comparison test ($^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$

compared to the control).

6. Conclusions

In summary, by utilizing a modern technique called UPLC-qTOF-MS/MS-based molecular networking, the phytochemicals in avocado seeds were dereplicated. Moreover, a procedure to quickly access alkaloids in avocado seed extract was successfully applied to isolate nine new compounds, including seven quinolone alkaloids and two benzoxazinone alkaloids. The structures of these compounds were reported for the first time and were fully determined by a combination of NMR spectroscopy analysis and ECD calculations. All isolated compounds were tested for their SIRT1 activities in HEK293 cells, and the results suggested that compound **1** showed the strongest SIRT1 stimulation activity with an elevated NAD^+/NADH ratio and could be a promising SIRT1 activator for the discovery of new drugs.

Part 3: Oleanane triterpenoids from *Gymnema inodorum* and their insulin-mimetic and muscle cell proliferation activities

1. Introduction

1.1. Study background

1.1.1. *Gymnema inodorum* (Lour.) Decne

Gymnema inodorum (Lour.) Decne is a climbing plant with a slender but vigorous woody stem (Wang et al., 2008) belonging to the family Apocynaceae which contains about 51 species distributed in Asia and Africa. *G. inodorum* name comes from the Hindi word "Gurmar" which means "destroyer of sugar" and it is also well-known for its antidiabetic effect as its name either in the traditional uses of their leaves (Dunkhunthod et al., 2021) or recent *in vitro* studies (Shimizu et al., 2001; Srinuanchai et al., 2021; Trang et al., 2021), and even in clinical trial (Anchalee et al., 2010). In addition, *G. inodorum* has been known to have therapeutic effects in curing other certain diseases, such as rheumatic arthritis, and gout (Dunkhunthod et al., 2021). The phytochemicals in *G. inodorum* have been reported including polyphenol and flavonoids (Dunkhunthod et al., 2021), oleanane triterpenoids (An et al., 2020), pregnane glycosides (Trang et al., 2021). Among reported phytochemicals from *G. inodorum*, methyl anthranilate derivatives that have been considered as crucial contributors to flavor odor in foods showed potential stimulation effects on glucose uptake in 3T3-L1 adipocyte cells (An et al., 2020). Methyl anthranilate is one of the most potent active floral odorants, occurring mainly in grapes and strawberries, and has been applied to the formulation of food additives or edible flavors (Thompson & Quaife, 2001). This compound has been reported to exhibit a flavor strong enough to exceed perception thresholds (Zhou et al., 2019). Therefore, it has been employed to analyze food quality, such as distinguishing wines made from *Vitis vinifera* grapes and those made from *Vitis labruscana* grapes or applied as a marker of the floral origin of citrus honey (Perry et al., 2019).

Figure 136. *Gymnema inodorum* (Lour.) Decne. The leaves of *G. inodorum* were collected in September 2018 in Hoai Duc district, Hanoi City, Vietnam (20°59′30.6″ N 105°43′49.8″ E).

1.1.2. Relative mass defect

The difference between the exact mass and nominal mass of a compound is known as a mass defect (Sleno, 2012). High-resolution mass spectrometers, which can measure exact mass, have become more and more common in analytical and bioanalytical chemistry since their introduction. Since every element's isotope produces a small amount of energy upon forming and stabilizing its nucleus, this energy causes the phenomenon known as a mass defect (nuclear binding energy) (Sleno, 2012). Each element has a distinct mass defect, therefore, the absolute value of an anion's mass defect reveals the ion's elemental makeup (Ekanayaka et al., 2015). Table 19 showed the absolute mass defects of common organic atoms in natural products.

Table 19. Natural abundances, exact mass, and mass defects of common elements in natural products.

Atom	Isotope	natural abundance (%)	Exact mass	Mass defect
H	^{1}H	99.9885	1.00783	0.00783
	^{2}H	0.0115	2.01410	0.01410
C	^{12}C	98.9300	12.00000	0.00000
	^{13}C	1.0700	13.00335	0.00335
O	^{16}O	99.7570	15.99491	-0.00509
	^{17}O	0.0380	16.99913	-0.00087
N	^{14}N	99.6320	14.00307	0.00307
	^{15}N	0.3680	15.00011	0.00011

Relative mass defect (RMD) in ppm, a normalization of the absolute mass defect

to the mass of the anion, has been introduced as an alternate and promising approach and was calculated as (mass defect/measured mono isotopic mass)$\times 10^6$ (Ekanayaka et al., 2015). By calculation of varieties of secondary metabolites from nature, Rubner and Intelmann provided a "molecular map" that showed a plot of RMD versus the *m/z* and has been useful for further analysis of natural products (Rubner & Intelmann, 2014). In particular, phenolics and flavonoids showed a relatively low RMD compared to the other type of compounds (RMD 100–300 ppm), and the higher molecular (*m/z*) indicated the level of glycosylation in those molecules. A similar trend at higher value of RMD was observed in terpenoids, triterpenoids, and saponins. The associated between RMD and *m/z* in the plot of "molecular map" enable quick and thorough overviews of secondary metabolites that can be found using HR-MS/MS (Rubner & Intelmann, 2014).

1.1.3. Natural products and type 2 diabetes

The epidemic of diabetes mellitus and its complications is a major global health problem (Chae & Shin, 2012). The global prevalence of diabetes and impaired glucose tolerance has sharply increased recently. The International Diabetes Federation reported the prevalence of adult diabetes as an estimated 425 million in 2017, with one out of 11 adults suffering from diabetes (Gomes et al., 2019). Furthermore, the number of persons with type 2 diabetes is expected to increase from 405 million in 2018 to 510 million in 2030 (Basu et al., 2019). With respect to socioeconomic burden, around 12% of overall global health-care expenditures are spent on the treatment of diabetes (Bommer et al., 2018).

Insulin is a peptide hormone from pancreatic β cells that increases glucose uptake by playing a critical role in glucose homeostasis (Gómez-Huelgas et al., 2015). Insulin binds to the insulin receptor, inducing conformational changes and successively resulting in autophosphorylation of the receptor (Sciacca et al., 2010). Insulin plays a critical role to maintain glucose homeostasis by orchestrating hepatic glucose production and peripheral glucose utilization (Alvarez & Ashraf, 2010). However, impaired insulin circulation reduces glucose transportation into skeletal muscle and adipose tissues by inducing high glucose levels in the blood (Nielsen et al., 2015). Insulin mimetics have been considered as strong candidates to treat type 2 diabetes, and they exhibit two major mechanisms by which compounds act like

insulin (Nankar & Doble, 2013). Insulin mimetics activate the tyrosine kinase domain of the insulin receptor, resulting in autophosphorylation of the receptor, activating downstream signaling pathways crucial for insulin metabolism. These agents are also known to inhibit protein tyrosine phosphatases to dephosphorylate insulin receptors and insulin receptor substrates. Much effort has been made to discover insulin mimetics from natural products, because of their potential lesser side effects and the possibility of penetrating the blood–brain barrier (Zhang et al., 1999).

1.1.4. Muscle atrophy

Skeletal muscles are the most abundant tissues accounted for 40–50 % of the total mass in healthy-weight individuals and are the protein reservoir in the human body (Sartori et al., 2021b). They not only control locomotion but also are fundamental for breathing, eating, energy expenditure, as well as for glucose, amino acids, and lipids homeostasis, and for maintaining a high quality of life (Sartori et al., 2021b). Skeletal muscles are heterogeneous concerning fiber type and metabolic properties and therefore they vary, often drastically, in their response to the same stimulus (Sartori et al., 2021b). Most of the pathways regulating muscle mass, such as insulin/IGF1 or TGFβ/activin/BMP, impinge both on protein synthesis and degradation, changes in protein turnover leading to muscle hypertrophy or atrophy, and muscle atrophy results from decreased protein synthesis and increased protein degradation (Sartori et al., 2021b). In muscle atrophy, the amino acids released by lysosome and proteasome during protein breakdown directly stimulate mTOR, and therefore, protein synthesis might increase (Sartori et al., 2021b). Recently, growing evidence has shown that natural products play a key role in the prevention and treatment of skeletal muscle atrophy such as resveratrol (Momken et al., 2011; D.-T. Wang et al., 2014a, 2014b), salidroside (X. Chen et al., 2016), matrine (Li Chen et al., 2019), imperatorin (Linlin Chen, Xu, et al., 2020), parthenolide (Quanjun et al., 2013), ursolic acid (Yu et al., 2017), and cryptotanshinone (Linlin Chen, Yang, et al., 2020). Structurally, oleanane triterpenoids isolated from *G. inodorum* shared a similar backbond to ursolic acid that showed an effect in the treatment of muscle atrophy (Yu et al., 2017).

1.2. Purpose of Research

Molecular networking that has been used as a dereplication strategy in the natural products field was applied to investigate phytochemicals in *G. inodorum*. Based on analyzing RMD and molecular network against GNPS libraries as well as a literature review of components from *Gymnema* genus, putative analogs of known compounds or those that belong to the same class that was observed in the same cluster were identified. Clusters of interesting compounds that were not assigned by comparison with available databases were further targeted and isolated for structure determination and biological activities evaluation.

2. **Targeted isolation of alkaloids from *G. inodorum***

2.1. Targeted isolation of alkaloids from *G. inodorum* by using LC-MS/MS-based molecular networking

The aerial part of *G. inodorum* was extracted with 70% EtOH. Then, the crude extract was dried under a vacuum and partitioned between *n*-hexane, EtOAc, *n*-BuOH, and water. The UPLC-qTOF-MS/MS data of fractions *n*-hexane, EtOAc, *n*-BuOH, and water were processed by Mzmine2 (Pluskal et al., 2010) and applied to the FBMN-GNPS platform (hosted by https://gnps.ucsd.edu, (M. Wang et al., 2016a)) to generate molecular networking. The result was visualized through the Cytoscape ver. 3.6.0 (Shannon, 2003). The dereplication of secondary metabolites from avocado seeds was identified against GNPS libraries and by comparison with compounds reported in the *Gymnema* genus (Scifinder) databases (Figure 138). As shown in Figure 138, four main clusters were visualized and putatively assigned to belong to oleanane triterpenoids (cluster 1), flavonoids (cluster 2), *N*-methyl anthranilate derivatives of oleanane triterpenoids (cluster 3), and benzoyl derivatives of oleanane triterpenoids. Therefore, further experiments were performed to isolate targeting peaks that were not reported before in the literature. As a result, fourteen new alkaloids were isolated, and their structures were elucidated by NMR spectroscopy.
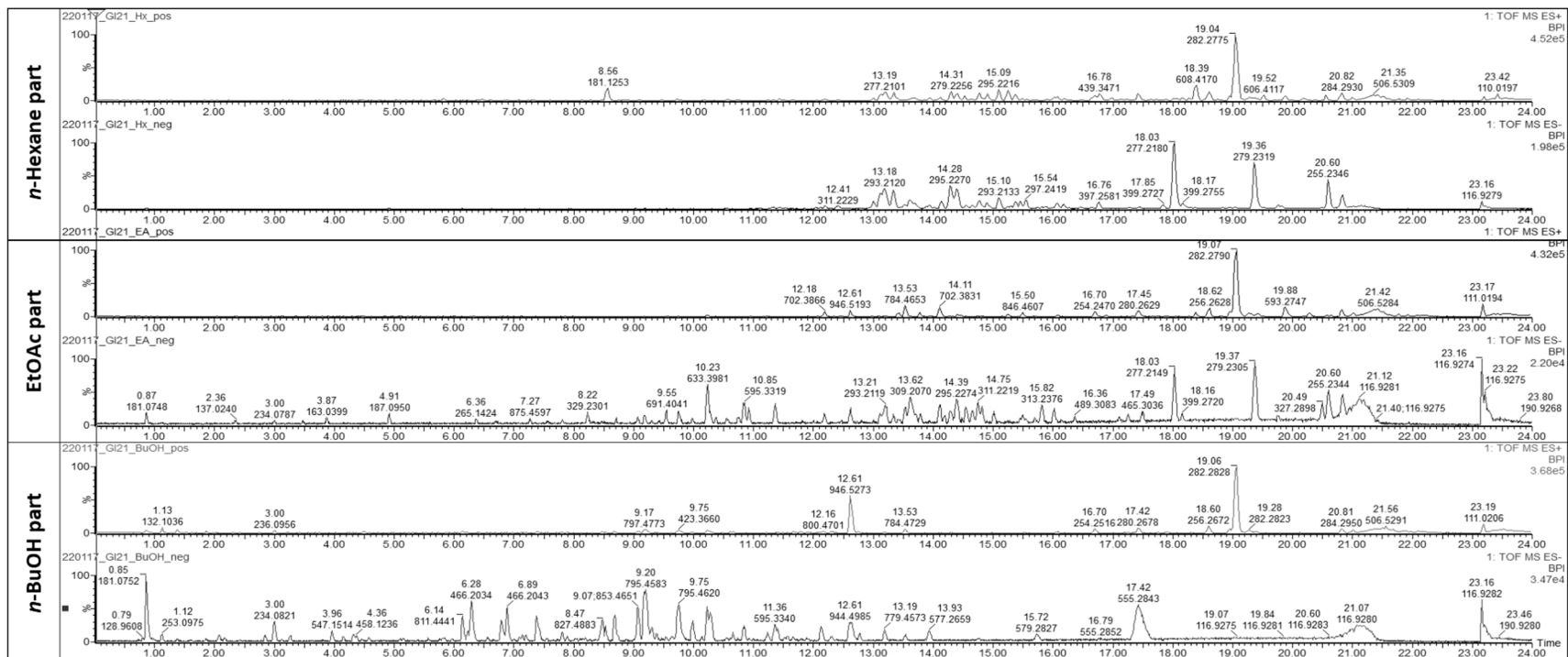
Figure 137. Total ion chromatography (TIC) of the *n*-hexane, EtOAc, and *n*-BuOH fractions of *G. inodorum* aerial extract in negative and positive modes.
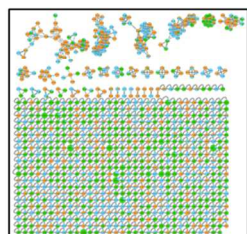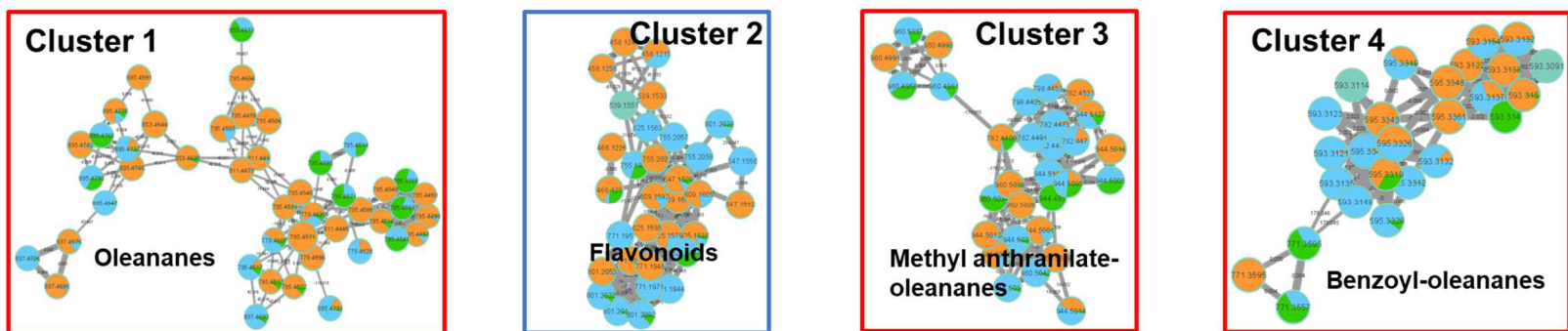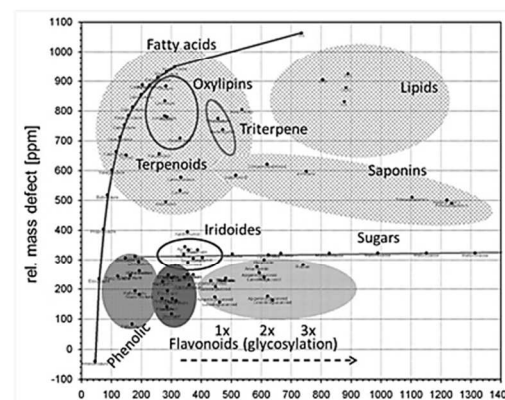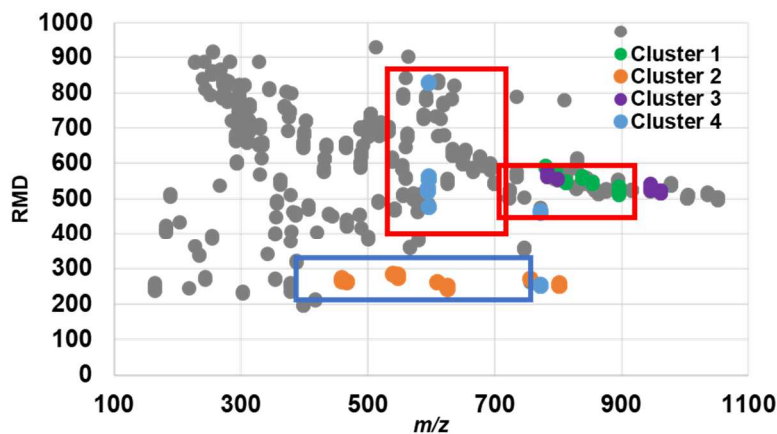
Figure 138. Molecular networking of the *n*-hexane, EtOAc, and *n*-BuOH fractions of *G. inodorum* aerial extract and the relative mass defect (RMD) analysis of these fractions.

(this network is accessible at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=924954612476435ca4a0c803a7dacd9d)
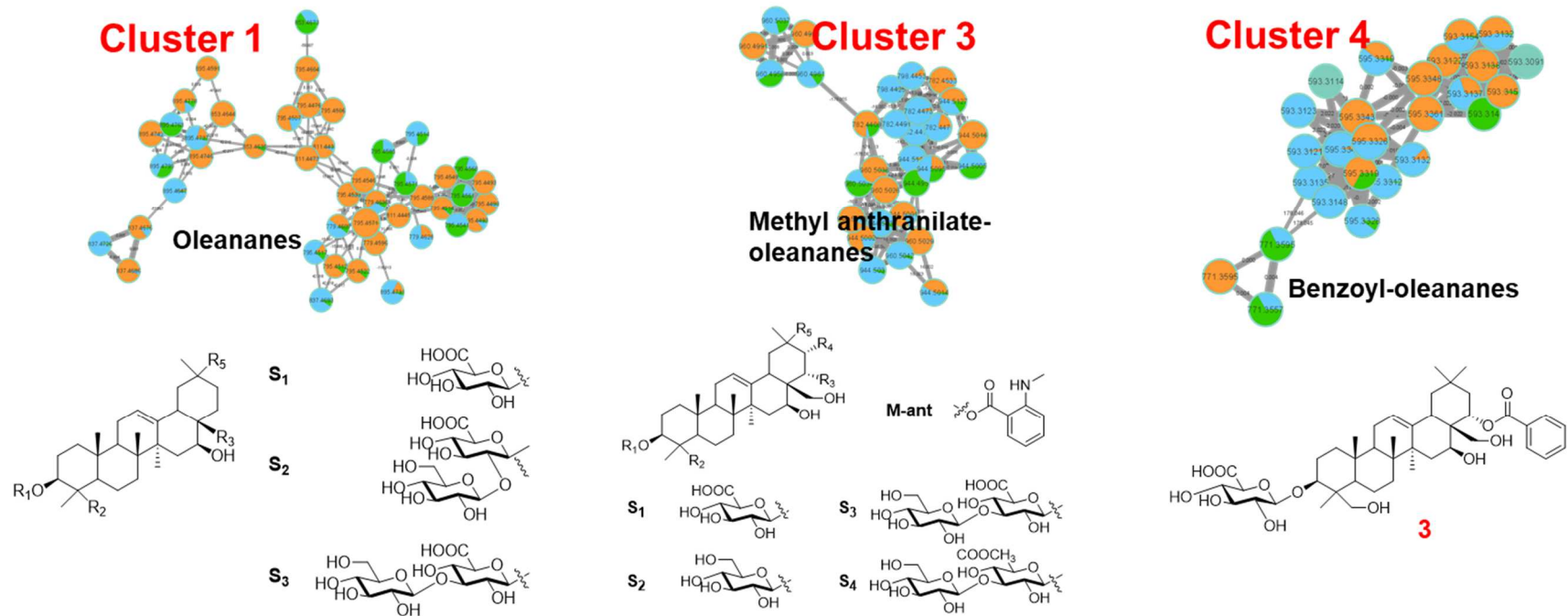
Figure 139. Targeted molecules for isolation from molecular networking of *G. inodorum*

	R_1	R_2	R_3	R_5
4	S_2	CH_2OH	CH_2OH	CH_3
5	S_2	CH_3	CH_3	CH_3
6	S_2	CH_3	CH_2OH	CH_2OH
7	S_2	CH_3	CH_3	CH_2OH
15	S_1	CH_3	CH_2OH	CH_3
16	S_3	CH_3	CH_2OH	CH_3

	R_1	R_2	R_3	R_4	R_5
1	S_3	CH_3	M-ant	H	CH_3
2	S_1	CH_2OH	M-ant	H	CH_2OH
8	S_1	CH_2OH	M-ant	H	CH_3
9	S_3	CH_2OH	M-ant	H	CH_3
10	S_2	CH_3	M-ant	H	CH_2OH
11	$S_{4\,(\alpha\text{-Glc})}$	CH_3	M-ant	H	CH_3
12	$S_{4\,(\beta\text{-Glc})}$	CH_3	M-ant	H	CH_3
13	S_2	CH_2OH	M-ant	H	CH_3
14	S_3	CH_3	H	M-ant	CH_3
17	S_1	CH_3	M-ant	H	CH_3
18	H	CH_3	M-ant	H	CH_3
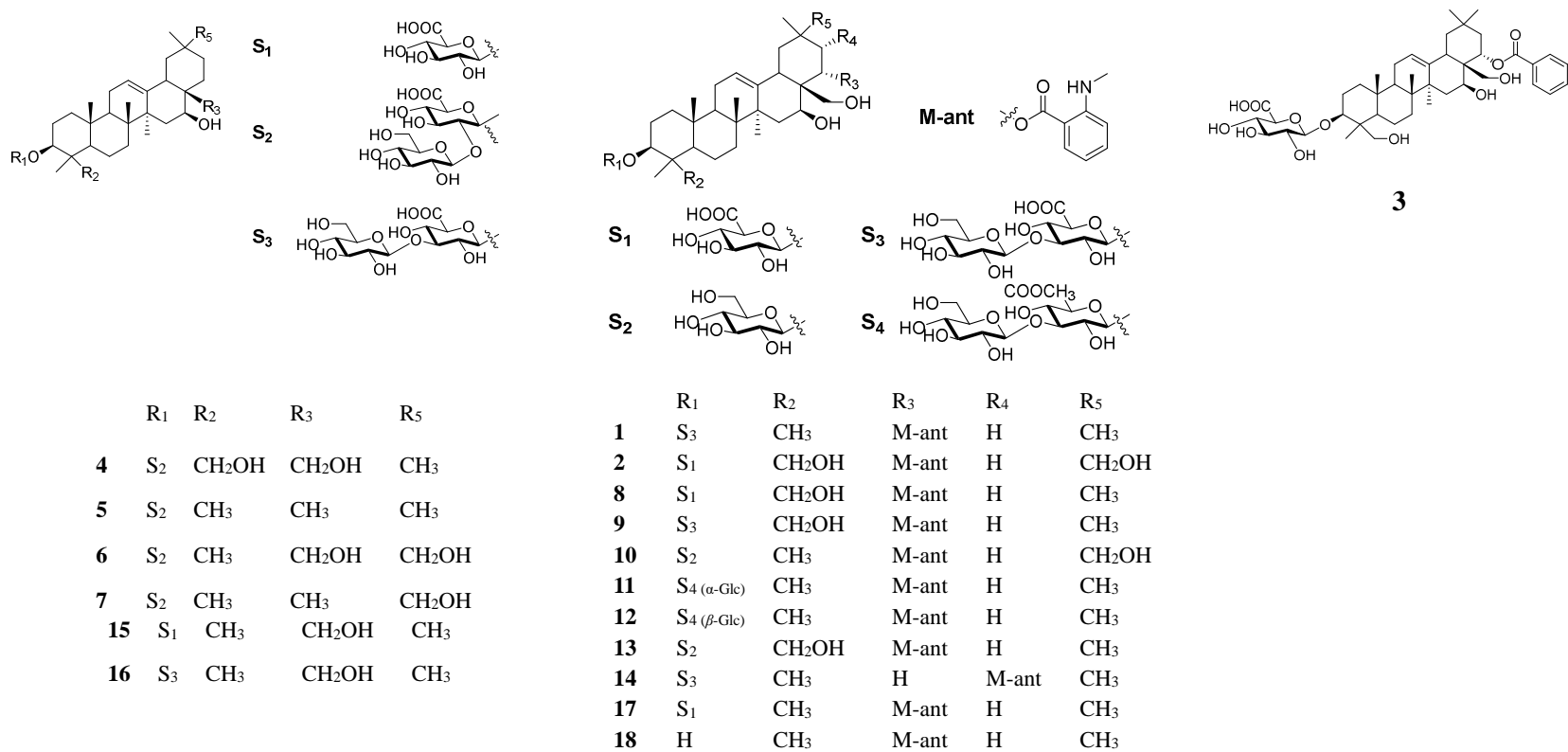
Figure 140. Chemical structures of isolated compounds **1–18** from *G. inodorum*

3. Structure elucidation of alkaloids from *G. inodorum*

3.1. Compound **1**

Compound **1** was isolated as a white, amorphous powder with an $[\alpha]_D^{25}$ – 6.2 (*c* 0.5, MeOH). Its molecular formula of $C_{50}H_{75}NO_{16}$ was suggested from a deprotonated highresolution electrospray ionization mass spectrometric (HRESIMS) ion peak at *m/z* 944.5023 [M–H]$^-$ (calcd for $C_{50}H_{74}NO_{16}$, 944.5007), indicating 14 degrees of unsaturation. The IR bands at 3394, 1605, and 1519 cm^{-1} were indicative of NH stretch, olefinic group, and NH bend absorbances. The ^{1}H NMR spectrum showed signals for four aromatic protons (δ_H 7.92 (dd, *J* = 8.0, 1.3 Hz, H-7'), 7.39 (ddd, *J* = 7.0, 6.5, 1.4 Hz, H-5'), 6.77 (dd, *J* = 8.1, 1.1 Hz, H-4'), and 6.64 (ddd, *J* = 7.0, 6.4, 1.3 Hz, H-6'), one olefinic proton (δ_H 5.36, br s, H-12), two anomeric protons (δ_H 4.58 (d, *J* = 7.8 Hz, H-1''') and 4.44 (d, *J* = 7.8 Hz, H-1''))], one nitrogenated methyl group (δ_H 2.88, s, H-8'), and seven methyl groups (δ_C 1.31, 1.11, 1.07, 1.04, 0.99, 0.98, and 0.86). The ^{13}C NMR spectrum exhibited resonances for two carbonyl carbons (δ_C 172.3 and 169.1), six aromatic carbons (δ_C 151.4, 135.6, 133.1, 117.0, 113.3, and 113.1), two olefinic carbons (δ_C 142.7 and 125.0), and two anomeric carbons (δ_C 106.6 and 105.2), suggesting the occurrence of pentacyclic triterpene aglycone with two sugar substituents. The HMBC correlations from the nitrogenated methyl signal at δ_H 2.88 to an aromatic carbon (δ_C 151.4, C-3') and from the aromatic proton signal at δ_H 7.92 to a carbonyl carbon (δ_C 169.1, C-1') suggested the presence of an anthranilate group (Figure 1A). Its position was confirmed based on the HMBC correlation from H-22 (δ_H 5.63, dd, *J* = 12.2, 3.8 Hz) to C-1'. The COSY spectrum showed correlations for H-15 (δ_H 1.79)/H-16 (δ_H 4.70) and H-21 (δ_H 1.83)/H-22 (δ_H 5.63), indicating C-16 and C-22 to be oxygenated (Figure 144). The oxygenated methylene group at C-17 was confirmed based on an HMBC correlation from H-28 (δ_H 3.59, m) to C-17 (δ_C 46.5). The presence of a glucuronic acid moiety was determined based on the HMBC correlation from H-5'' of glucuronic acid (GlcA) (δ_H 3.81, m) to C-6'' (δ_C172.3). In the HMBC spectrum, the anomeric proton of the glucuronic acid unit showed a correlation with C-3 (δ_C 91.0) of the aglycone, and the anomeric proton of glucose was correlated with C-3''. The attachment of an *N*-methyl anthranilate moiety at C-22 was deduced as α-oriented based on the NOESY correlation between H-18 (δ_H 2.59, dd, *J* = 11.0, 3.0 Hz) and H-22 (δ_H 5.63, dd, *J* = 12.2, 3.8 Hz) (Figure 145). The NMR data observed for compound **1** were similar to

those of (3β,16β,22α)-22-(N-methylanthraniloxy)- 16,23,28-trihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid, which was isolated previously from *G. inodorum*, except for the absence of a C-23 hydroxy group of the aglycone in compound **1** (Shimizu et al., 2001). Thus, compound **1** was characterized as (3β,16β,22α)-22-(N-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-Dglucopyranosiduronic acid.
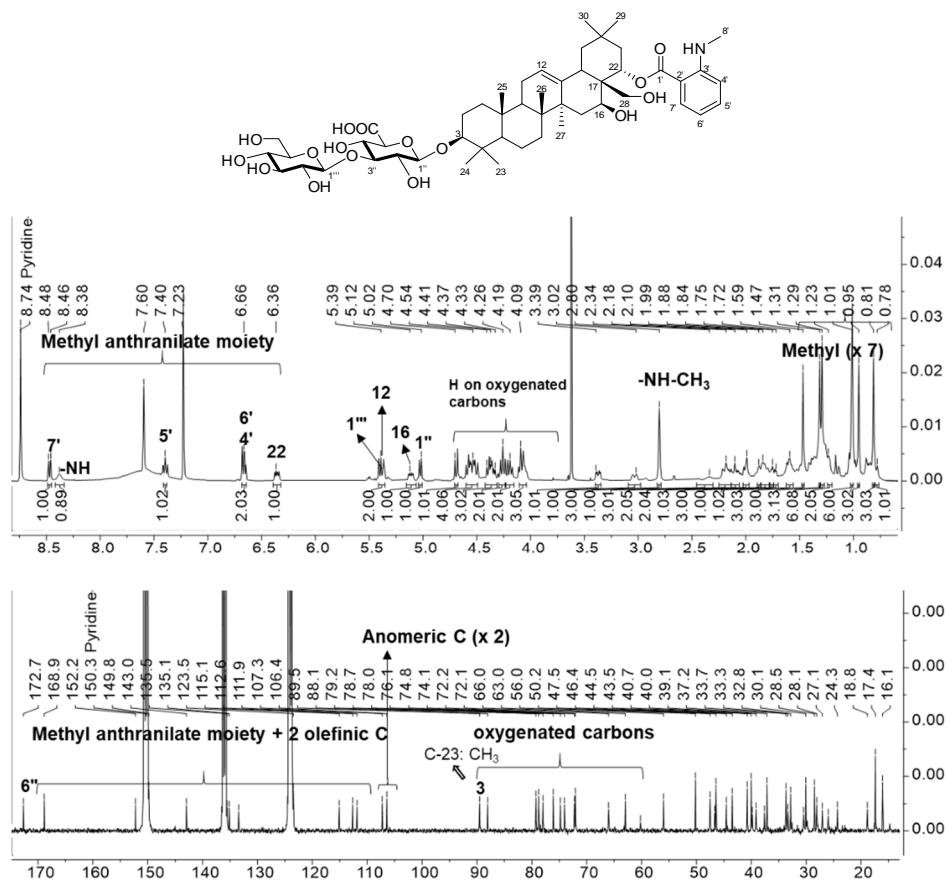
Figure 141. ^{1}H and ^{13}C NMR spectra of compound **1** (400/100 MHz, pyridine-d_5)
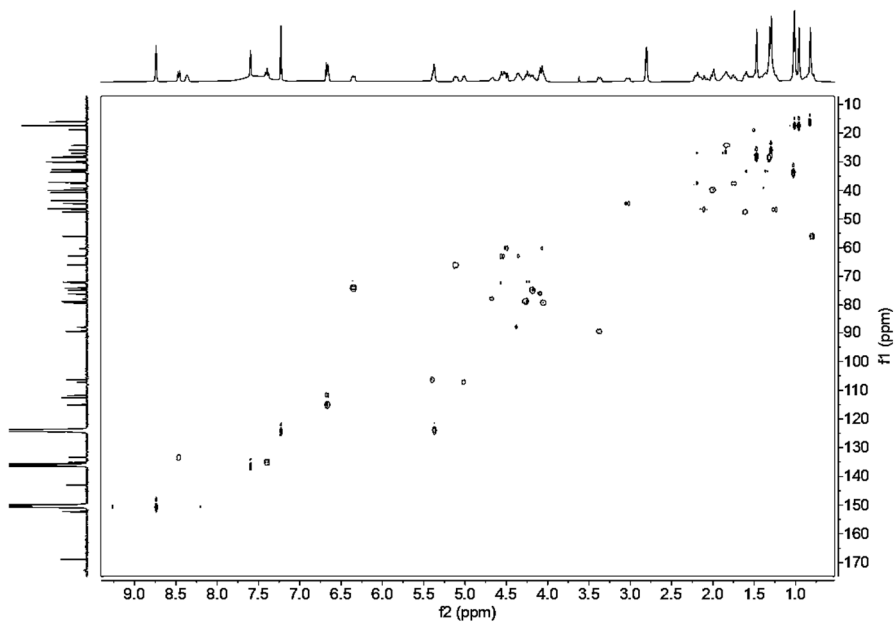
Figure 142. HSQC spectrum of compound **1**
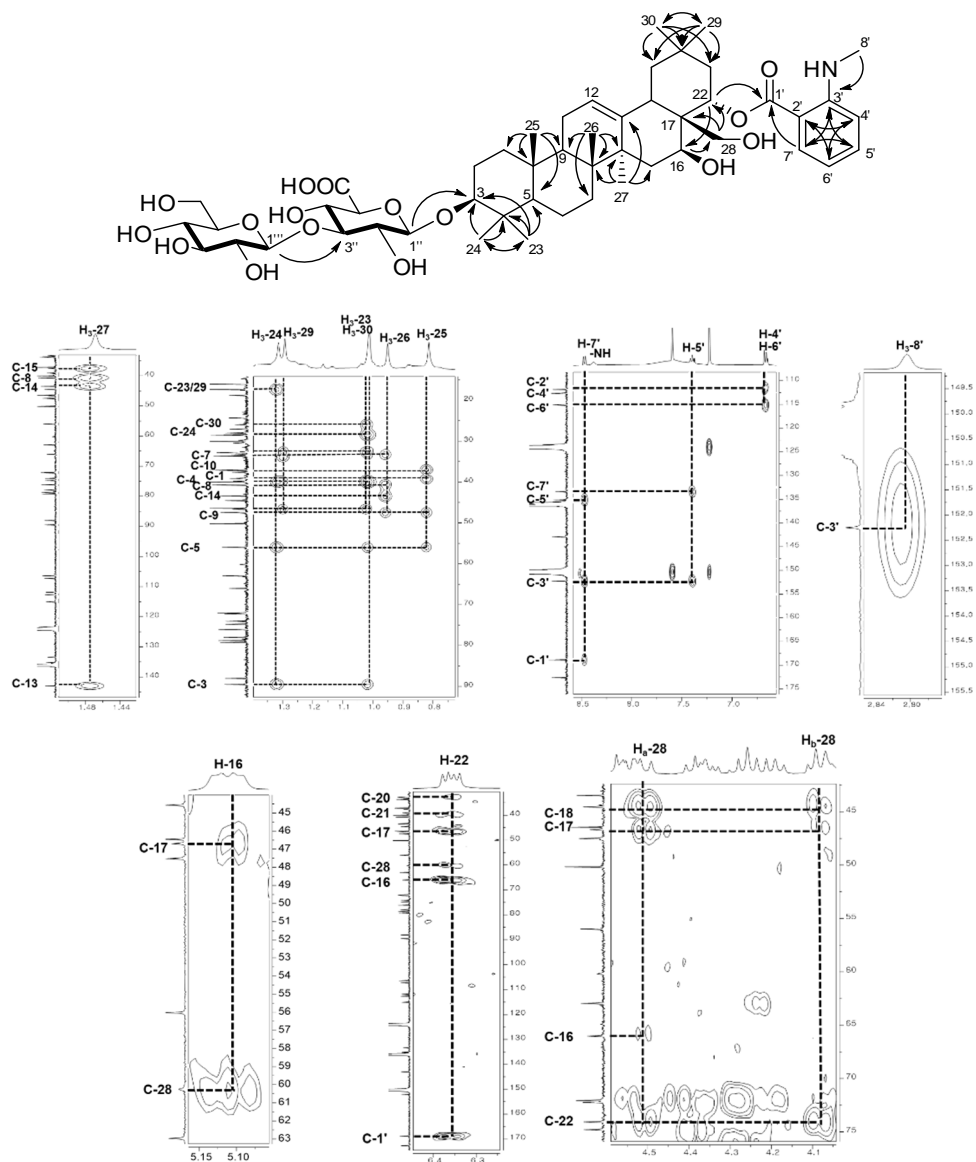
Figure 143. HMBC spectrum of compound **1**

Figure 144. $^1H–^1H$ COSY spectrum of compound **1**
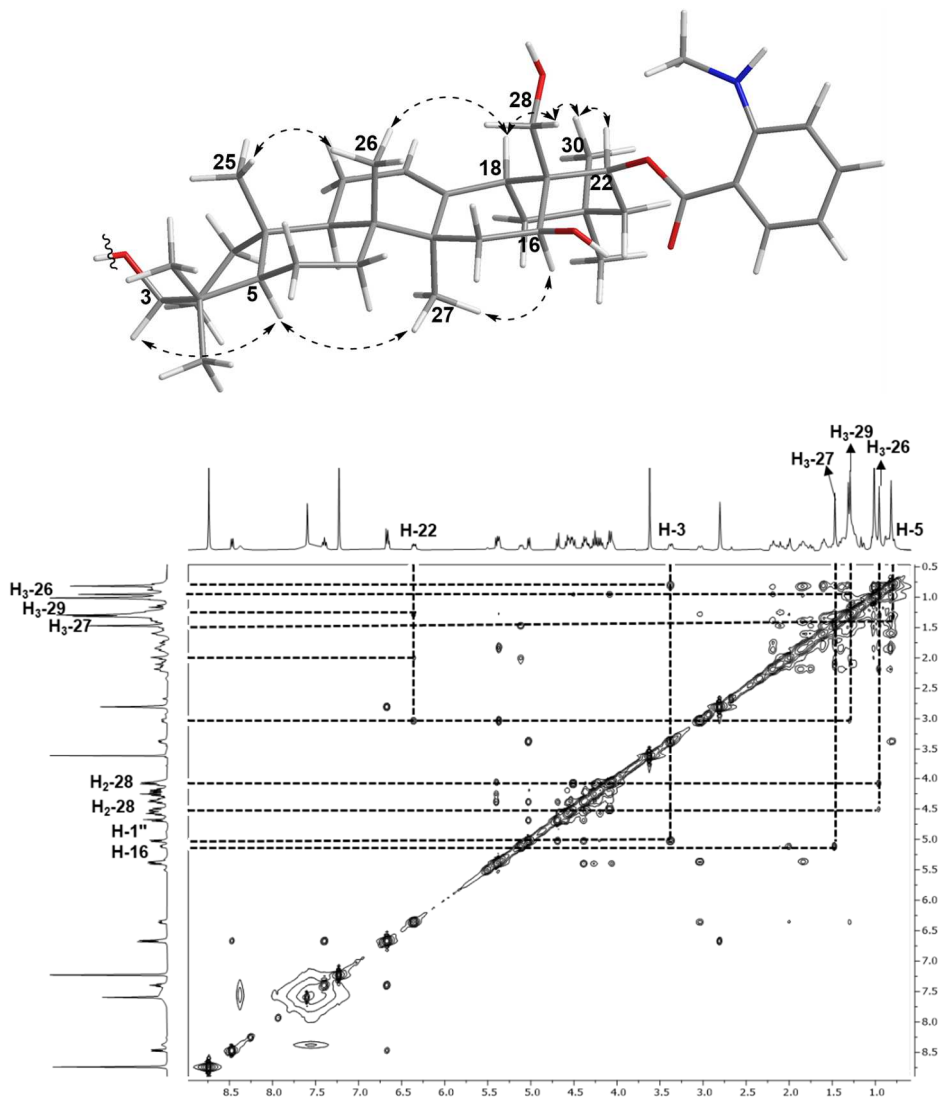
Figure 145. NOESY spectrum of compound **1**

3.2. Compound 2

Compound **2** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ – 33 (*c* 0.1, MeOH). The molecular formula $C_{44}H_{65}NO_{13}$ was determined from the HRESIMS ion peak at *m/z* 814.4402 [M–H]$^-$ (calcd for $C_{44}H_{64}NO_{13}$, 814.4377), indicating 13 degrees of unsaturation. The IR bands at 3395, 1607, and 1520 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **2** were similar to those of compound **1**. Analysis of its ^{1}H and ^{13}C NMR data with the aid of the HSQC spectrum revealed signals of four aromatic protons [δ_H 7.89 (dd, *J* = 8.0, 1.6 Hz, H-7'), 7.34 (ddd, *J* = 8.6, 7.1, 1.7 Hz, H-5'), 6.68 (d, *J* = 8.4 Hz, H-4'), and 6.59 (ddd, *J* = 8.4, 7.0, 1.6 Hz, H-6')], a nitrogenated methyl (δ_H 2.86, s, H$_3$-8'), and a carbonyl carbon (δ_C 169.6, C-1'), consistent with the presence of a methyl anthranilate substituent in compound **2**. The major difference for **2** compared to **1** was the presence of two hydroxy groups at two methyl groups (C-23 and C-29) and the absence of any glucose unit. Two pairs of oxygenated methylene groups [δ_H 3.63/3.30 (H$_2$-23) and 3.23/ 3.22 (H$_2$-29)] showed HMBC correlations with C-4 (δ_C 43.9) and C-20 (δ_C 38.3), respectively, indicating that both C-23 and C-29 are substituted with a hydroxy group. The carbon chemical shift of C-3" was present at δ_C 77.8, which was relatively upfield shifted compared to C-3" (δ_C 86.7) in compound **1**, suggesting no functional group or sugar was attached. The configuration of C-22 was suggested based on a NOESY correlation between H-18 (δ_H 2.65, d, *J* = 12.8 Hz) and H-22 (δ_H 5.63, dd, *J* = 12.3, 3.9 Hz). Therefore, compound **2** was assigned as (3β,16β,22α)-22-(*N*-methylanthraniloxy)-16,23,28,29-tetrahydroxyolean-12-en-3-yl-O-β-D-glucopyranosiduronic acid.
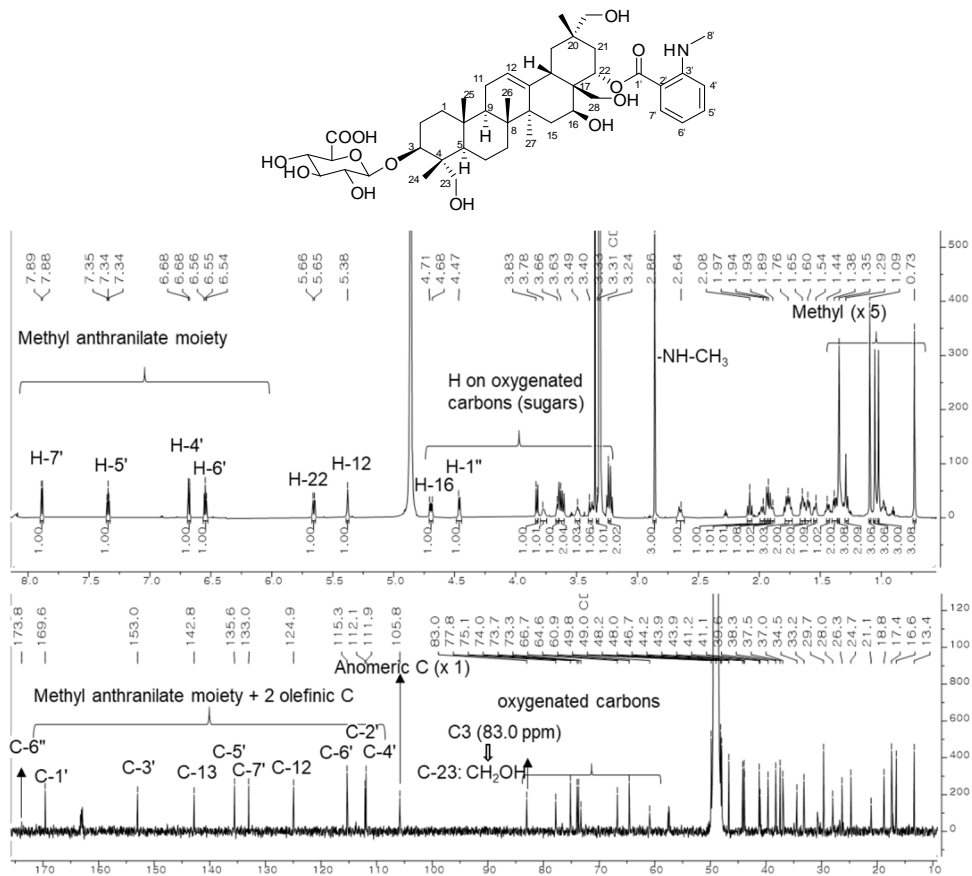
Figure 146. ^{1}H and ^{13}C NMR spectra of compound **2** (800/200 MHz, methanol-d_4)
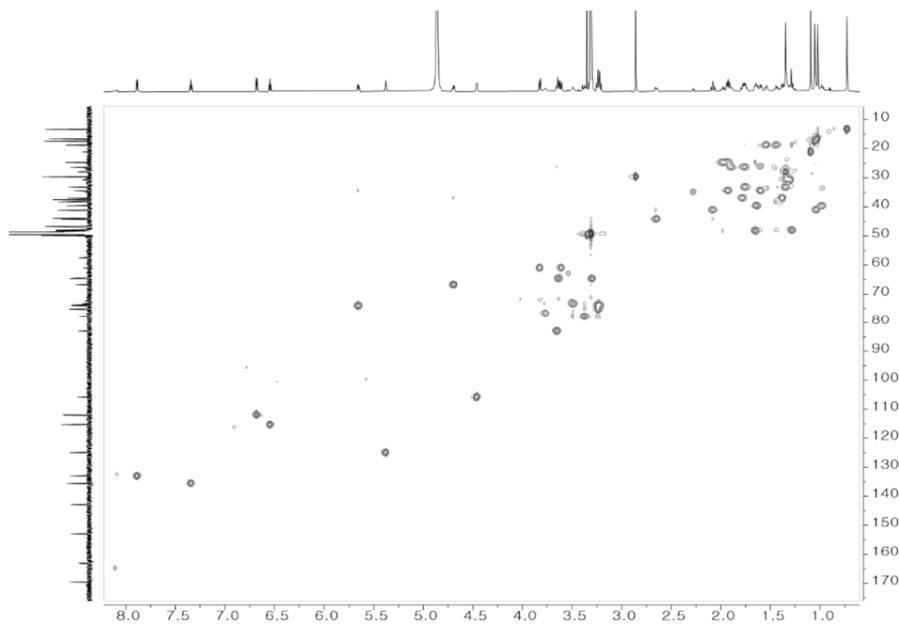
Figure 147. HSQC spectrum of compound **2**
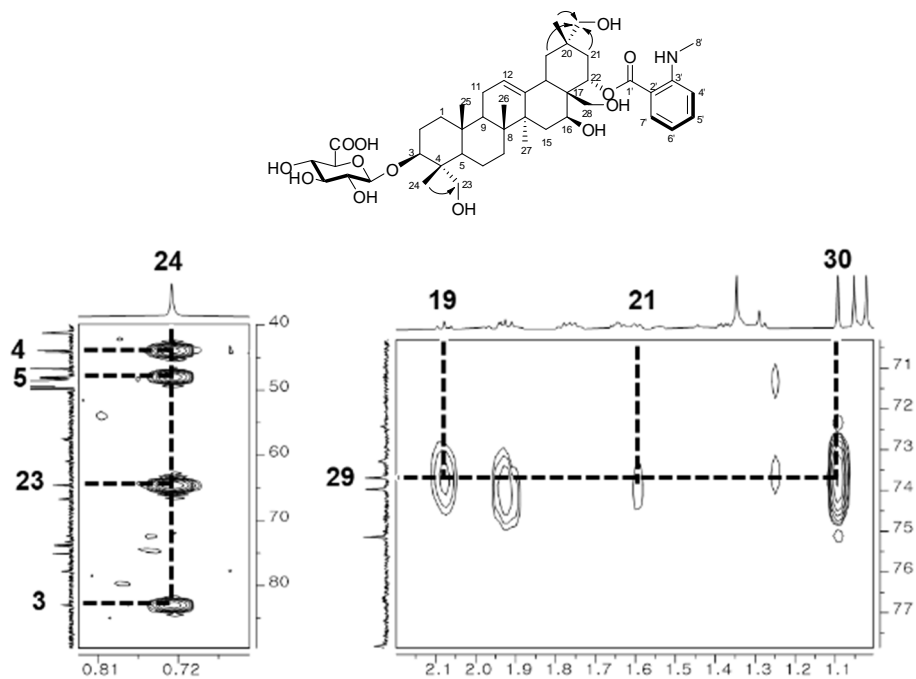
Figure 148. HMBC spectrum of compound **2**
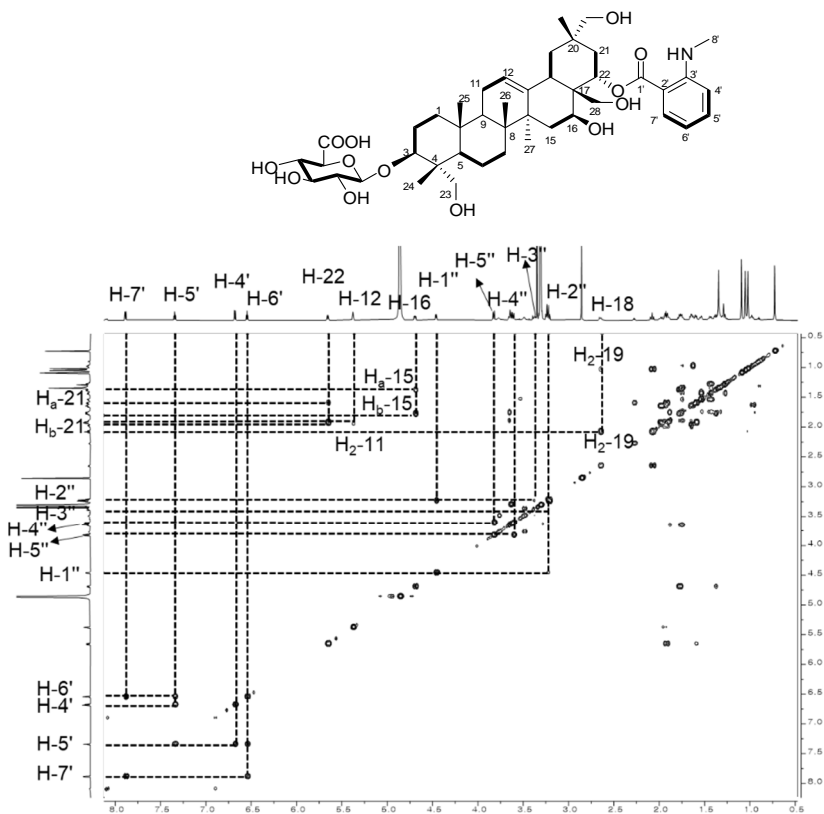
Figure 149. ^{1}H–^{1}H COSY spectrum of compound **2**
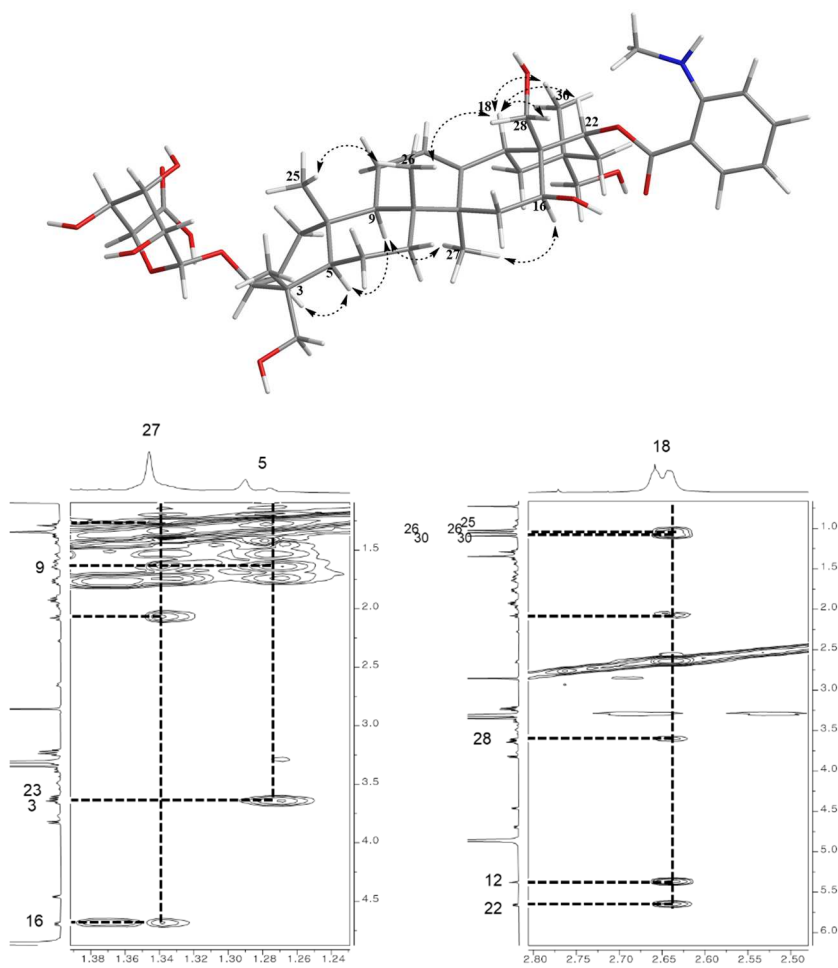
Figure 150. ROESY spectrum of compound **2**

3.3. Compound 3

Compound **3**, a white, amorphous powder with an $[\alpha]_D^{25}$ + 8.2 (*c* 0.6, MeOH), was found to possess a molecular formula of $C_{43}H_{62}O_{12}$ based on its HRESIMS peak at *m/z* 769.4150 [M–H]$^-$ (calcd for $C_{43}H_{61}O_{12}$, 769.4163), which indicated 13 degrees of unsaturation. The IR bands at 3359 and 1606 cm^{-1} were indicative of the presence of hydroxy and olefinic groups. The NMR data of compound **3** showed similar patterns to those of compound **2** except the absence of any signals on the methyl amine group at C-2' of the aromatic ring and for a hydroxy group at C-29 of the aglycone. The ^{1}H NMR data of **3** showed a typical A_2B_2X system in the aromatic ring [δ_H 8.04 (dd, *J* = 8.0, 1.0 Hz, H-2', H-6'), 7.57 (m, H-4'), and 7.46 (m, H-3', H-5')]. The HMBC correlation from H-22 at δ_H 5.64 (dd, *J* = 11.9, 3.8 Hz) to C-7' (δ_C 167.8) indicated the position of a benzoic acid substituent. The HMBC correlations between two methyl groups [δ_H 1.11 (s, H$_3$-30) and 1.00 (s, H$_3$-29)] and a quaternary carbon (δ_C 32.9, C-20) suggested that neither C-29 nor C-30 was substituted. The proton on H-22 was determined as being *β*-oriented based on a NOESY correlation between H-18 (δ_H 2.56, d, *J* = 11.0 Hz) and H-22 (δ_H 5.64, dd, *J* = 11.9, 3.8 Hz). Compound **3** was characterized as (3*β*,16*β*,22*α*)-22-benzoyloxy-16,23,28-trihydroxyolean-12-en-3-yl-O-*β*-D-glucopyranosiduronic acid.
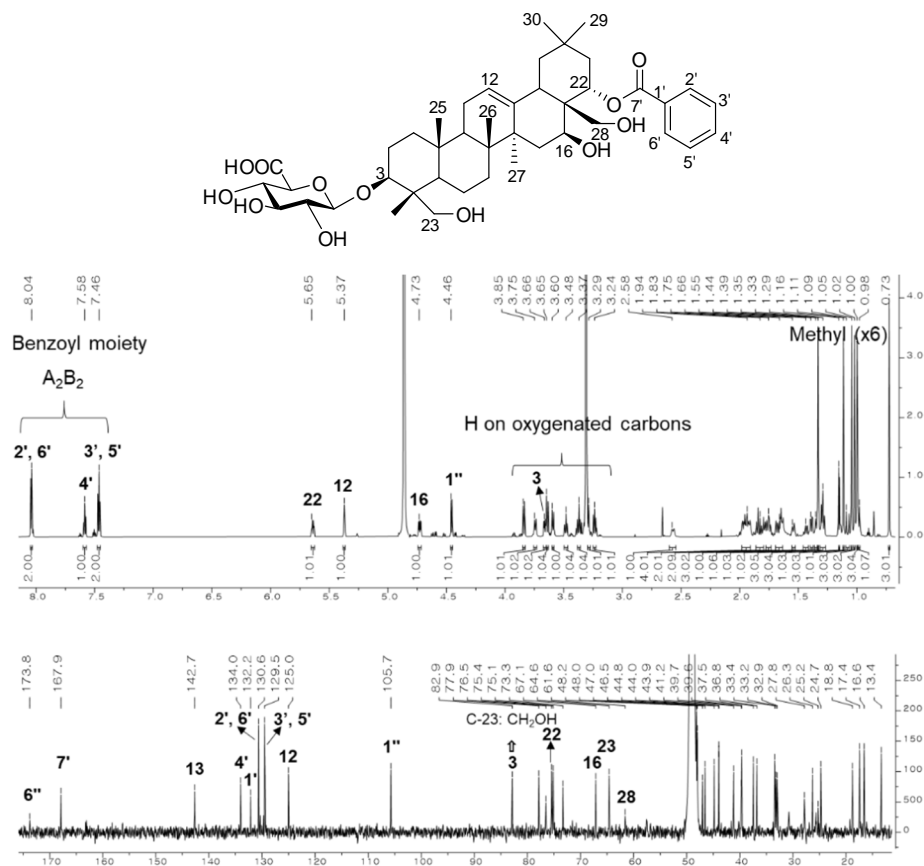
Figure 151. ^{1}H and ^{13}C NMR spectra of compound **3** (800/200 MHz, methanol-d_4)
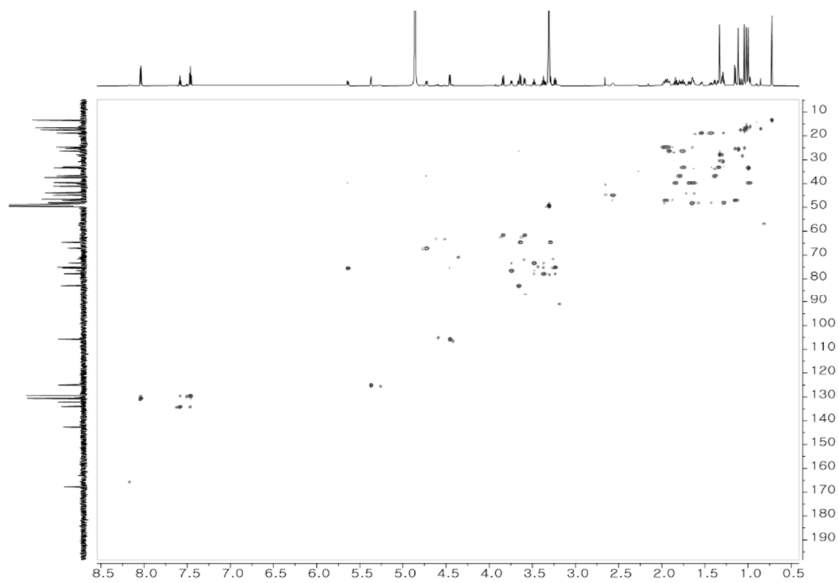
Figure 152. HSQC spectrum of compound **3**
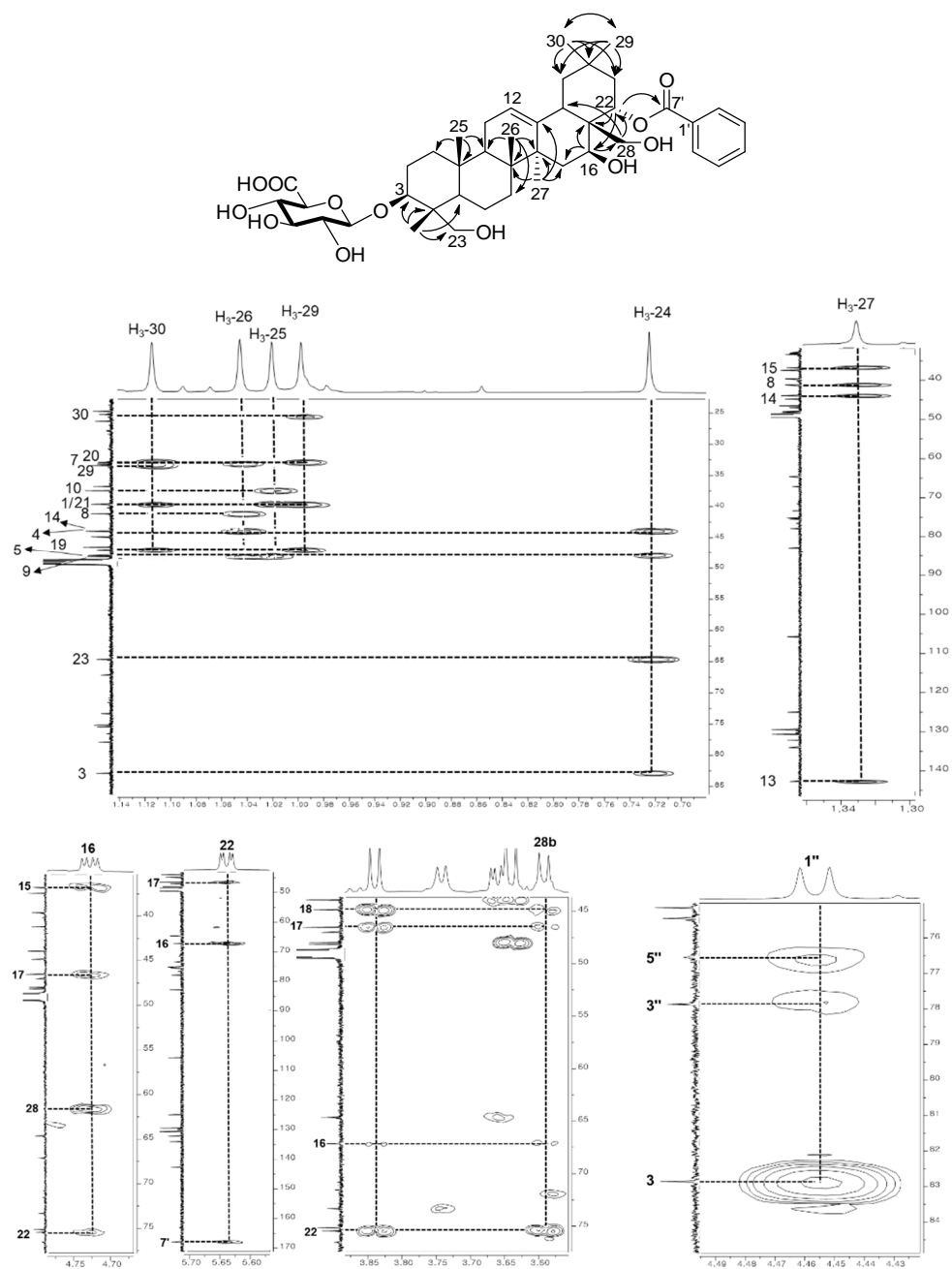
Figure 153. HMBC spectrum of compound **3**
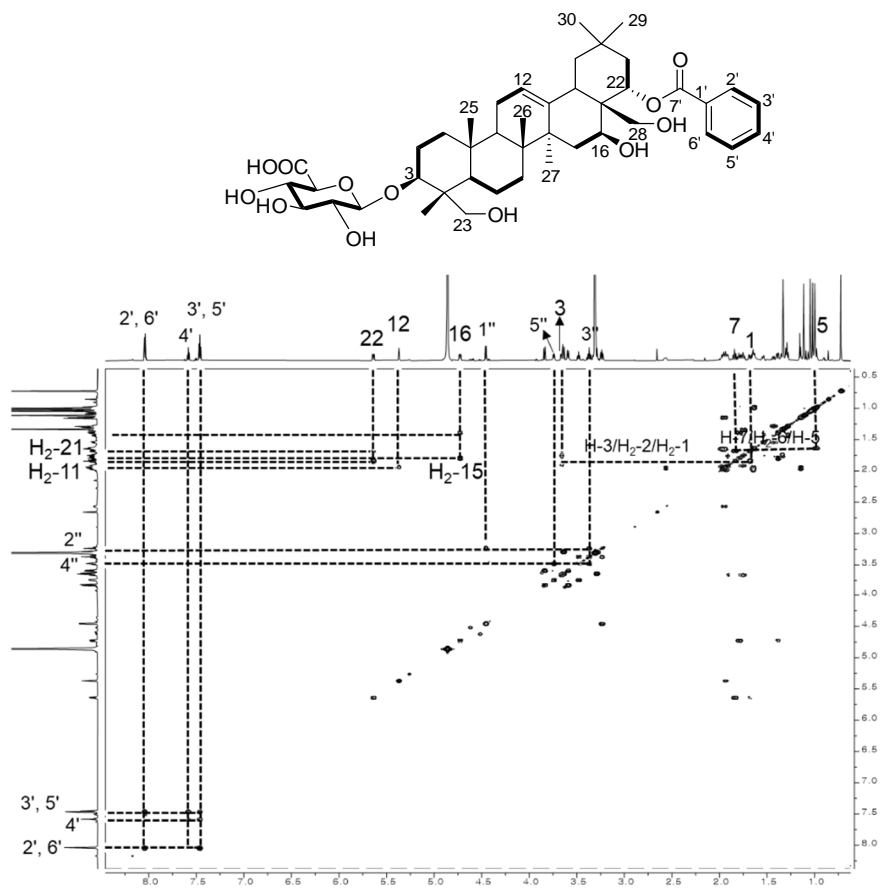
Figure 154. ^1H–^1H COSY spectrum of compound **3**

3.4. Compound 4

Compound **4** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ + 51.6 (*c* 0.6, MeOH). The molecular formula $C_{42}H_{68}NO_{15}$ was determined from the HRESIMS ion peak at *m/z* 811.4472 [M−H]⁻ (calcd for $C_{42}H_{67}O_{15}$, 811.4480), consistent with nine degrees of unsaturation. The IR bands at 3360 and 1607 cm⁻¹ were indicative of hydroxy and olefinic groups. The ¹H and ¹³C NMR data of **4** were similar to those of compound **1**. The major difference was the absence of signals for an *N*-methylanthranilate group at C-22 and the evidence for a different glycosidic linkage. The ¹H NMR data of compound **4** showed six singlet methyl groups (δ_H 1.38, 1.10, 1.04, 1.00, 0.94, and 0.91). The HMBC correlations from H-1" (δ_H 5.21, d, *J* = 7.1 Hz) to C-3 (δ_C 82.3) and from H-1‴ (δ_H 5.46, d, *J* = 7.8 Hz) to C-2" (δ_C 83.9) supported the glycosidic position. The carbon chemical shift of C-2" was shifted downfield, because of the attachment of a glucopyranosyl unit at C-2". Therefore, compound **4** was determined as (3β,16β)-16,23,28-trihydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid.
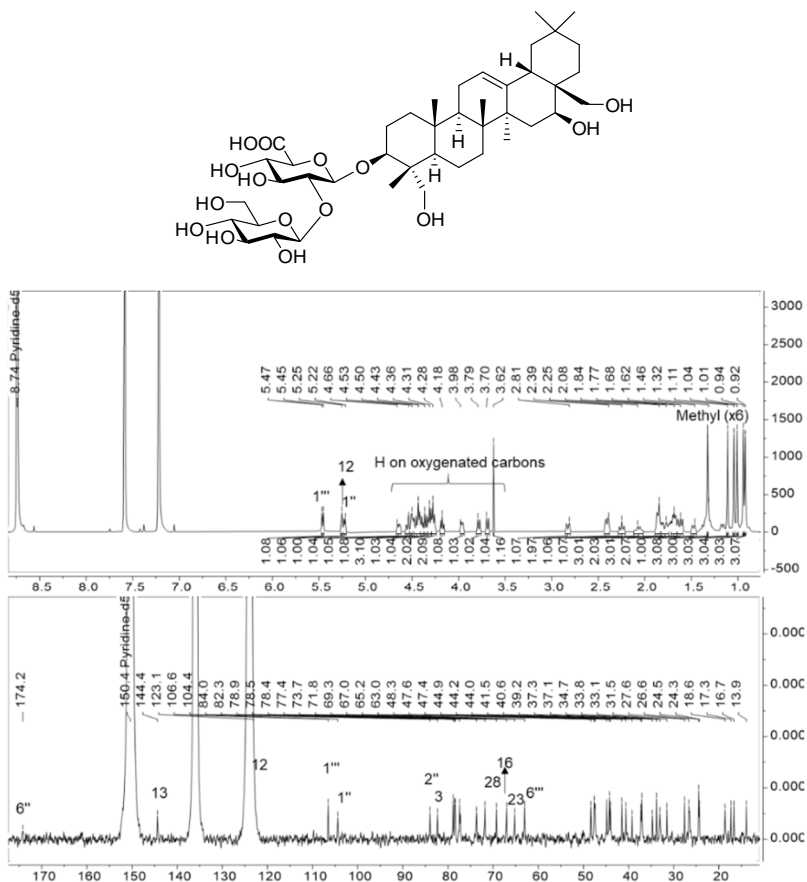
Figure 155. ¹H and ¹³C NMR spectra of compound **4** (500/125 MHz, pyridine-*d₅*)
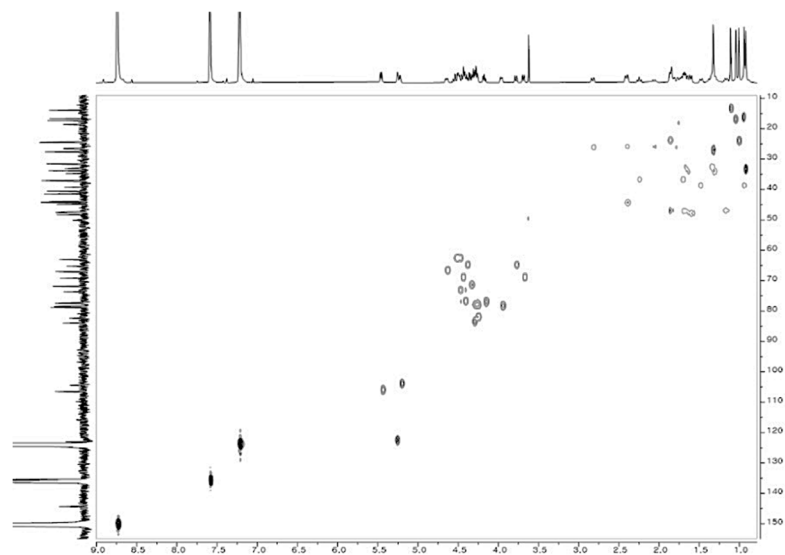
Figure 156. HSQC spectrum of compound **4**
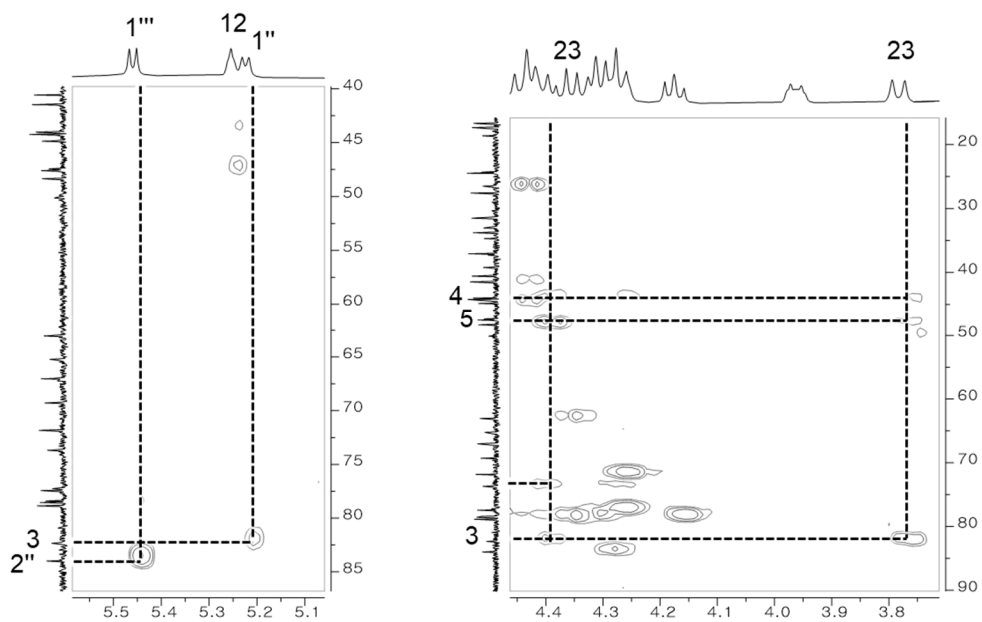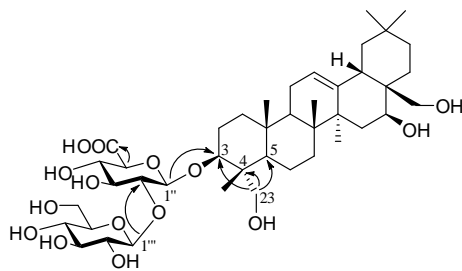
Figure 157. HMBC spectrum of compound **4**
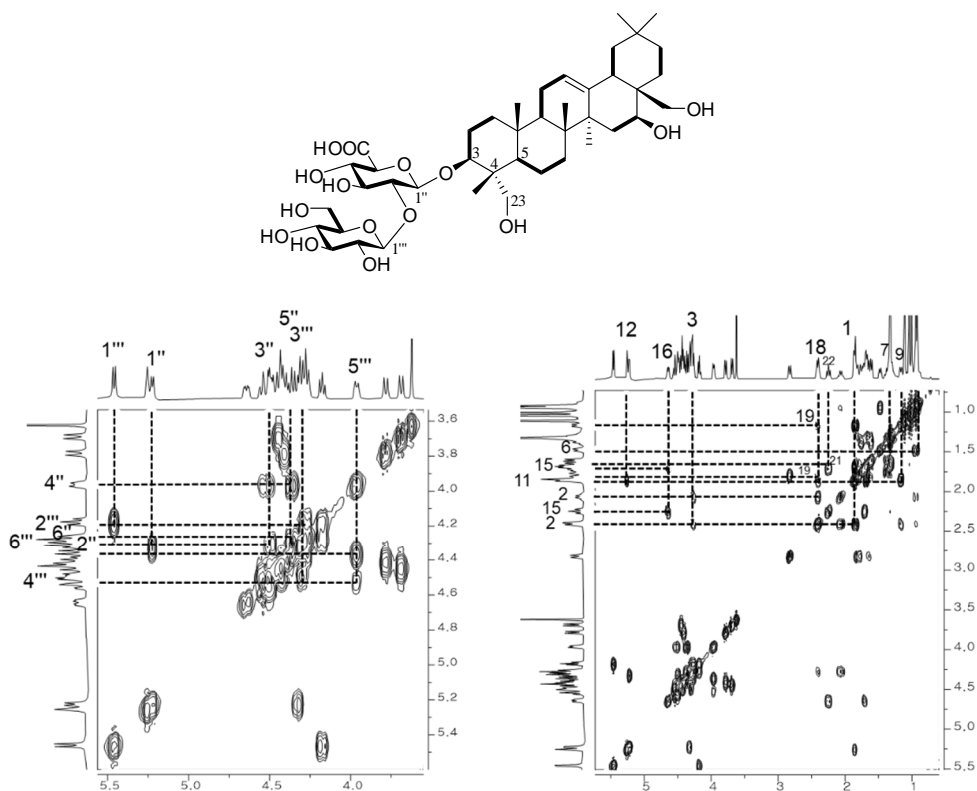
Figure 158. ^{1}H–^{1}H COSY spectrum of compound **4**
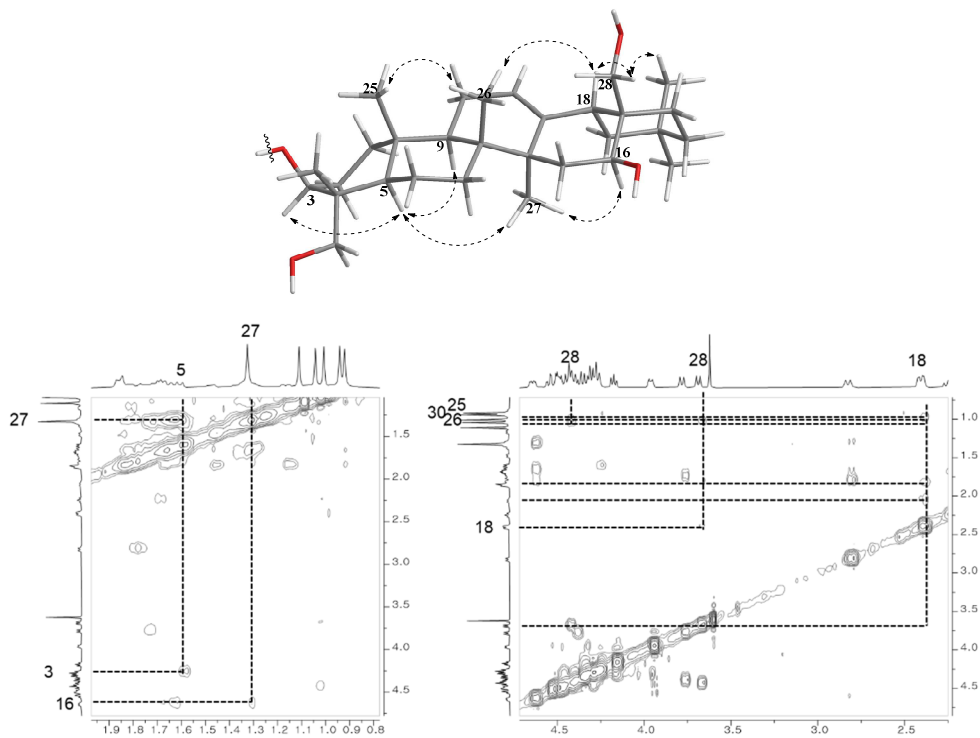
Figure 159. NOESY spectrum of compound **4**

3.5. Compound 5

Compound **5**, a white, amorphous powder with an $[\alpha]_D^{25}$ + 51.6 (*c* 0.6, MeOH), was found to possess a molecular formula of $C_{42}H_{68}O_{13}$, based on its HRESIMS peak at *m/z* 779.4568 [M–H]$^-$ (calcd for $C_{42}H_{67}O_{13}$, 779.4581), indicating nine degrees of unsaturation. The IR bands at 3361 and 1606 cm^{-1} were indicative of hydroxy and olefinic groups. Its ^{1}H and ^{13}C NMR data revealed that compound **5** possesses a similar structure to compound **4** except for the absence of hydroxy groups at C-23 and C-28. The ^{1}H NMR spectrum of compound **5** showed eight singlet methyl group resonances (δ_H 1.23, 1.09, 1.02, 0.99, 0.92, 0.89, 0.87, and 0.79). The glucoside unit position was established based on HMBC correlations from H-1'' (δ_H 4.46, d, *J* = 7.8 Hz) to C-3 (δ_C 91.6) and from H-1''' (δ_H 4.69, d, *J* = 7.8 Hz) to C-2'' (δ_C 80.8). The carbon chemical shift of C-2'' was shifted downfield, because of the attachment of a glucopyranosyl at C-2''. Compound **5** was characterized as (3*β*,16*β*)- 16-hydroxyolean-12-en-3-yl-2-O-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronic acid.
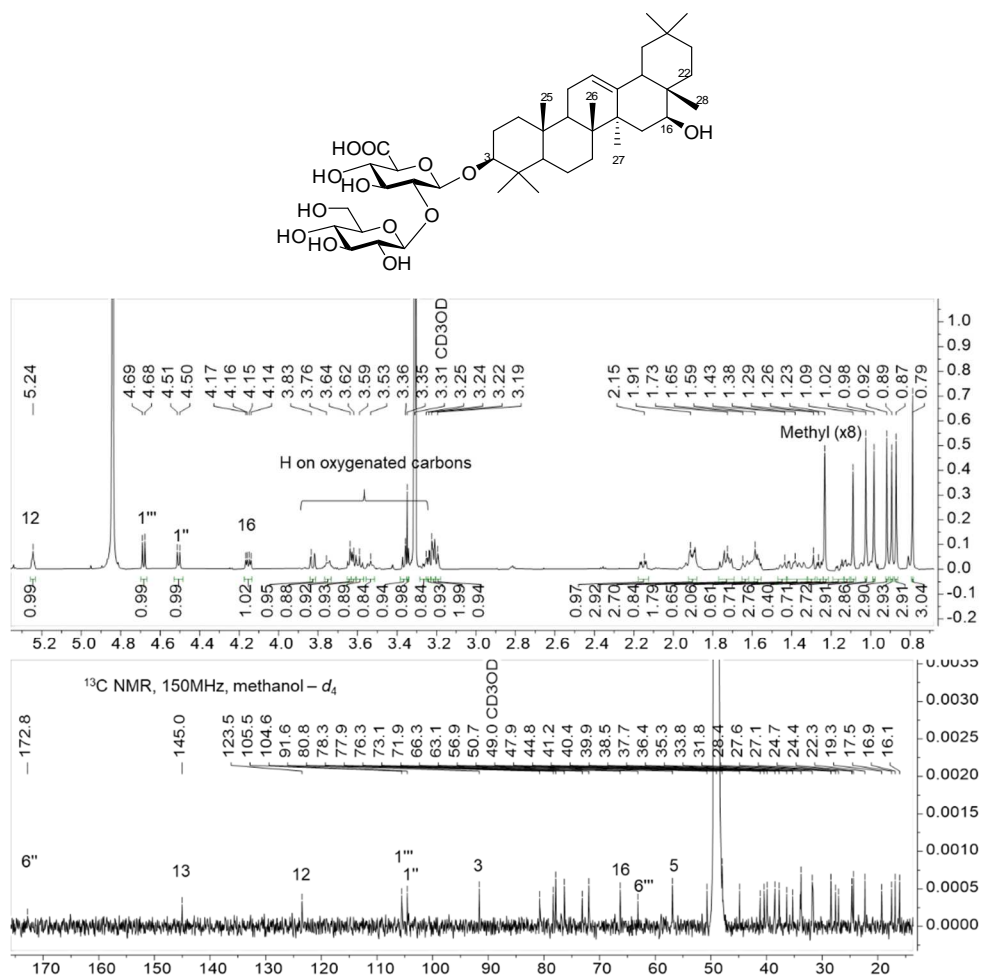
Figure 160. ^{1}H and ^{13}C NMR spectra of compound **5** (800/200 MHz, methanol-d_4)
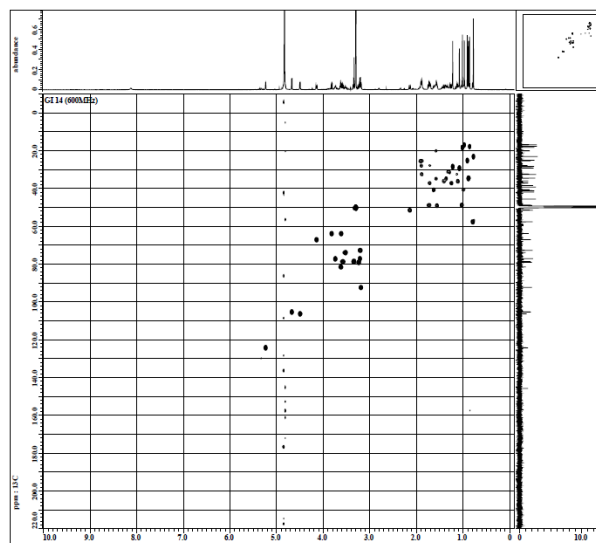
Figure 161. HSQC spectrum of compound **5**

Figure 162. ^{1}H–^{1}H COSY spectrum of compound **5**

Figure 163. NOESY spectrum of compound **5**

198

3.6. Compound 6

Compound **6** was isolated as a white, amorphous powder with $[\alpha]_D^{25}$ –34.6 (*c* 0.1, MeOH). Its molecular formula of $C_{42}H_{68}O_{15}$ was deduced from the deprotonated HRESIMS peak at *m/z* 811.4470 [M–H]⁻ (calcd for $C_{42}H_{67}O_{15}$, 811.4480) showing nine degrees of unsaturation. The IR bands at 3360 and 1605 cm⁻¹ were indicative of the occurrence of hydroxy and olefinic groups. The ¹H and ¹³C NMR data were similar to those of compound **5** except for the presence of hydroxy groups at C-28 and C-29. Six singlet methyl groups (δ_H 1.39, 1.28, 1.23, 1.12, 1.03, and 0.85) were observed, and oxygenated methylene protons were present at δ_H 3.62 (s, 2H, H₂-29). Therefore, compound **6** was proposed structurally as (3*β*,16*β*)-16,28,29-trihydroxyolean-12-en-3-yl-2-O-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronic acid.
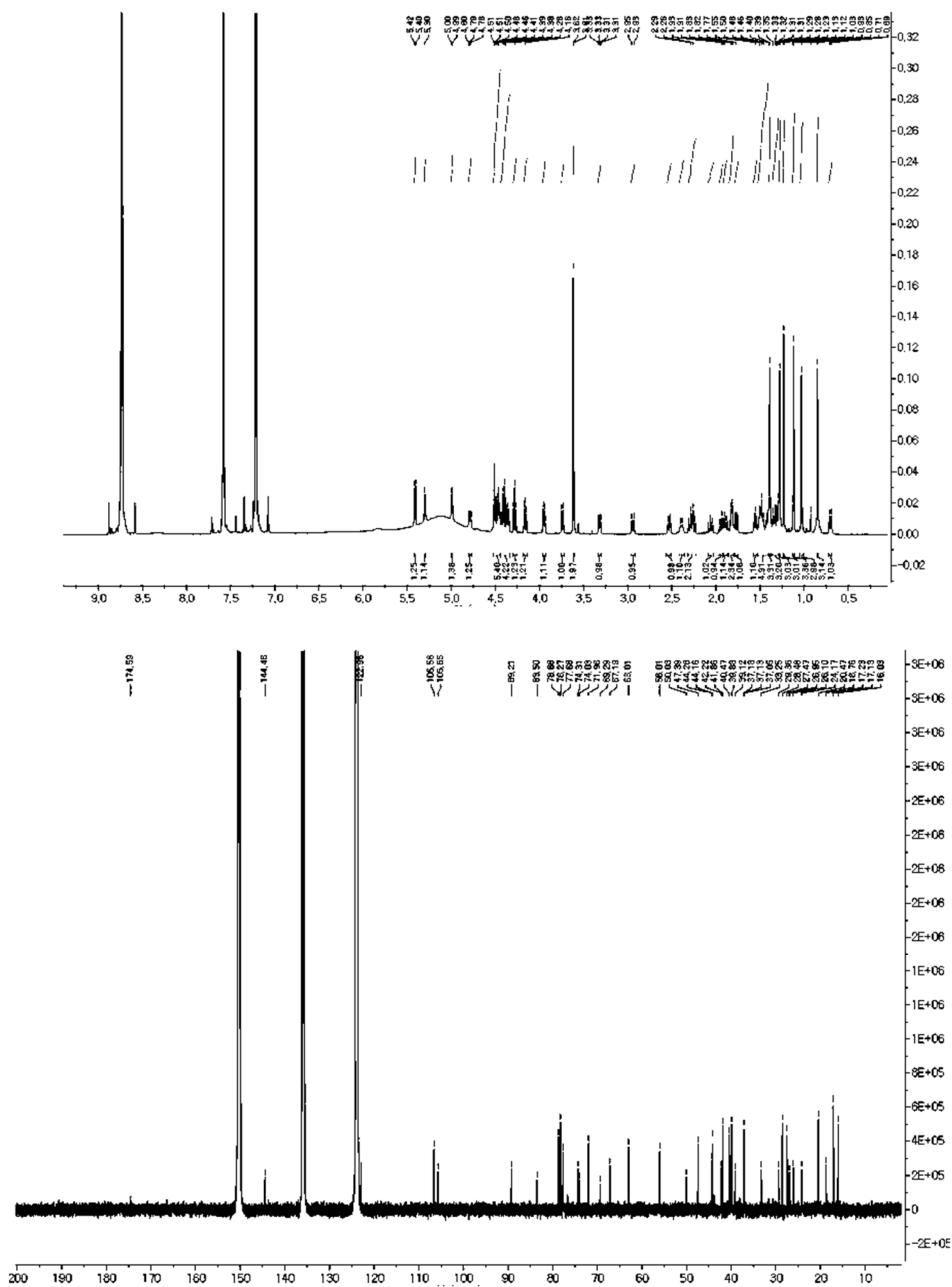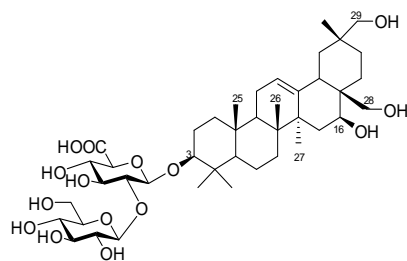
**Figure 164.** ^{1}H and ^{13}C NMR spectra of compound **6** (600/150 MHz, methanol-d_4)

3.7. Compound 7

Compound **7** was isolated as a white, amorphous powder, with an $[\alpha]_D^{25}$ +19.8 (*c* 0.1, MeOH). The chemical formula of $C_{42}H_{68}O_{14}$ was suggested from the deprotonated HRESIMS peak at *m/z* 795.4520 [M−H]$^-$ (calcd for $C_{42}H_{67}O_{14}$, 795.4531), indicating nine degrees of unsaturation. The IR bands at 3358 and 1606 cm^{-1} showed hydroxy and olefinic groups to be present. The NMR spectrum of compound **7** showed similar patterns to those of compound **5** except the presence of a signal for a hydroxy group at C-29. Seven singlet methyl groups (δ_H 1.40, 1.28, 1.21, 1.21, 1.17, 1.05, and 0.88) were shown, and oxygenated methylene protons were present at δ_H 3.60 (s, H$_2$-29). Compound **7** was characterized therefore as (3β,16β)-16,29-dihydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid.
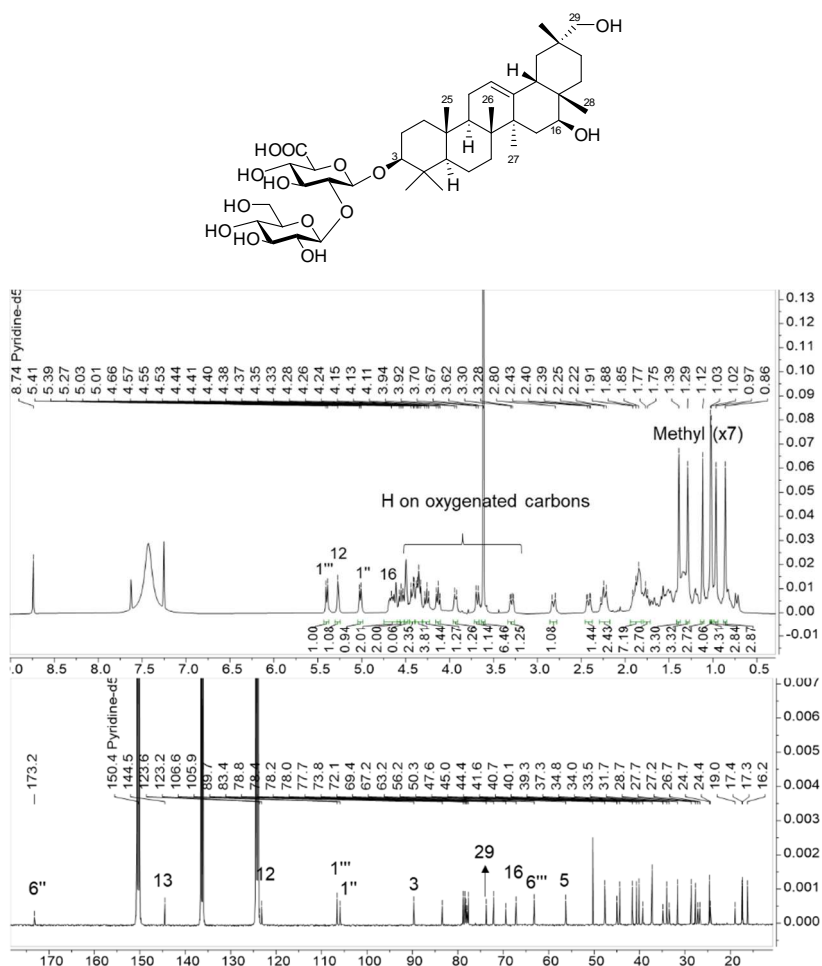
Figure 165. ^{1}H and ^{13}C NMR spectra of compound **7** (500/125 MHz, methanol-d_4)

201

Figure 166. HSQC spectrum of compound **7**
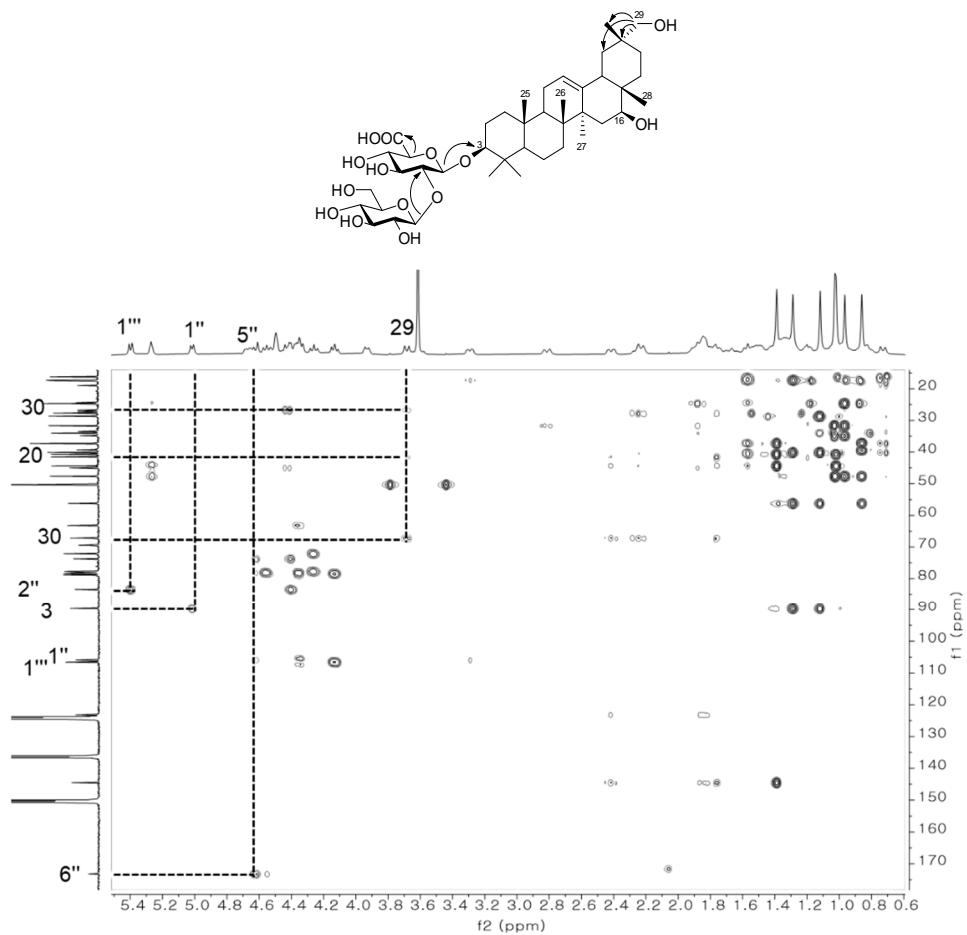
Figure 167. HMBC spectrum of compound **7**

Figure 168. ^{1}H–^{1}H COSY spectrum of compound **7**

3.8. Compound 8

Compound **8** was identified as (3β,16β,22α)-22-(N-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-β-D-glucopyranosiduronic acid based on the reported compounds in literature (Shimizu et al., 2001).
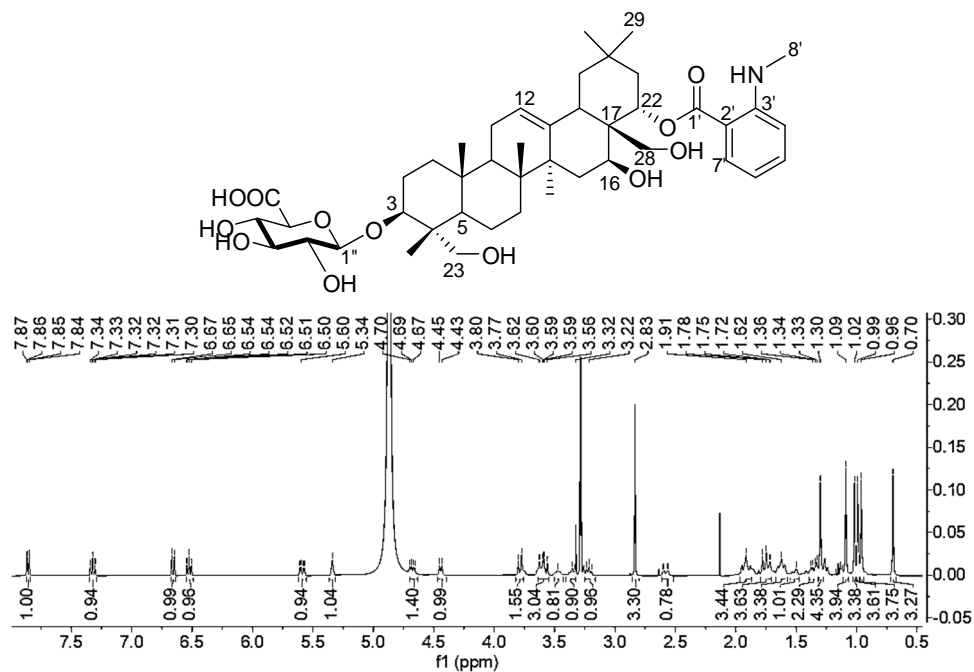
Figure 169. ^{1}H and ^{13}C NMR spectra of compound **8** (400/100 MHz, methanol-d_4)

3.9. Compound 9

The ^{1}H and ^{13}C NMR data of compound **9** suggested that **9** was (3β,16β,22α)-22-(N-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid that was reported in literature (Shimizu et al., 2001).
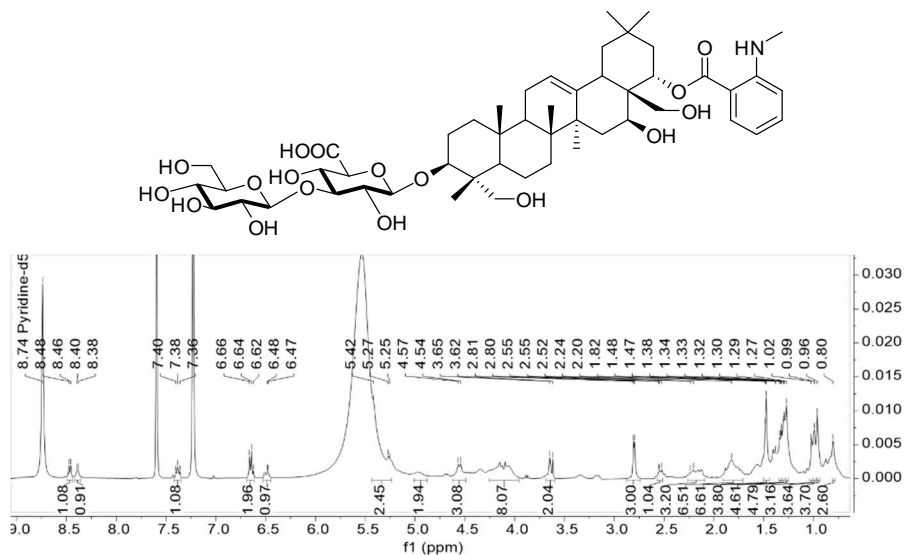
Figure 170. ^{1}H and ^{13}C NMR spectra of compound **9** (400/100 MHz, pyridine-d_5)

3.10. Compound 10

Compound **10** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +21 (*c* 0.1, MeOH). The molecular formula $C_{44}H_{67}NO_{11}$ was determined from the HRESIMS ion peak at *m/z* 960.4989 $[M - H]^-$, (calcd for $C_{44}H_{66}NO_{11}$, 960.4957), indicating 12 degrees of unsaturation. The IR bands at 3383, 1607, and 1514 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^1H and ^{13}C NMR data of **10** were similar to those of the oleanane skeleton of **8** with *N*- methyl anthranilate moiety. The differences between **10** and **8** were observed in the sugar moiety and the position of a methylene hydroxyl on the oleanane skeleton. In particular, the downfield shifted resonance of C-3 at 89.4 ppm indicated that there were two methyl groups located at C-4 (C-23, δ_H 0.98/ δ_C 17.5 and C-24, δ_H 1.29/ δ_C 28.5). HMBC correlations from H_2-29 (δ_H 4.13/4.03) to C-30 (δ_C 28.8), C-20 (δ_C 37.6), C-21 (δ_C 35.5), and C-19 (δ_C 41.7) suggested that the methylene hydroxy functional group was C-29 in **10** instead of at C-23 as in **8**. Moreover, the structure of **10** showed the presence of a glucose unit with an anomeric signal at δ_H 5.47 (*J* = 7.6 Hz)/ δ_C 106.4 ppm instead of the glucuronic acid unit as in **8**. The large coupling constant of the anomeric proton indicated that the sugar was in *β*-orientation, and the absolute configuration of sugar was identified to be *β*-D-glucose by comparing to authentic compounds after acid hydrolysis and sugar derivatizing. Therefore, **10** was identified as (3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28,29-trihydroxyolean-12-en-3-yl-*O*-*β*-D-glucopyranosiduronic acid.
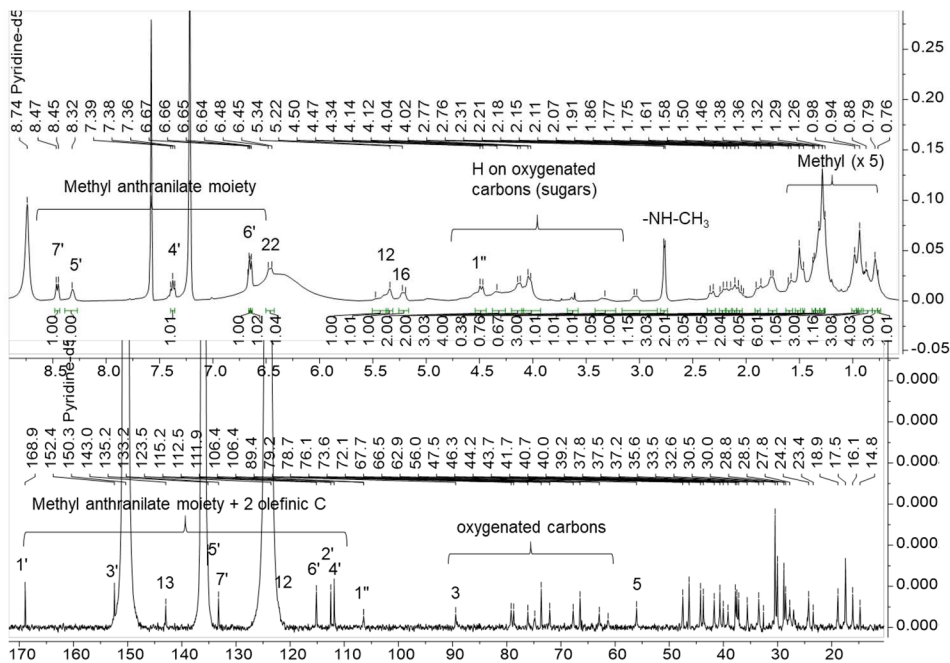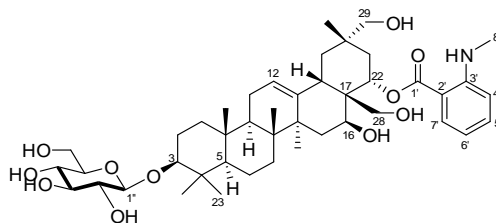
Figure 171. ^{1}H and ^{13}C NMR spectra of compound **10** (400/100 MHz, pyridine-d_5)
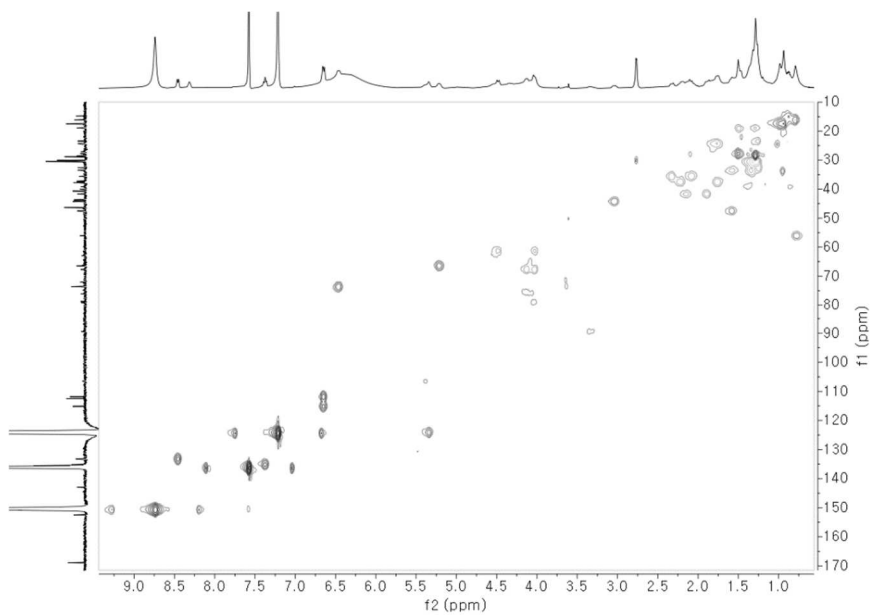
Figure 172. HSQC spectrum of compound **10**
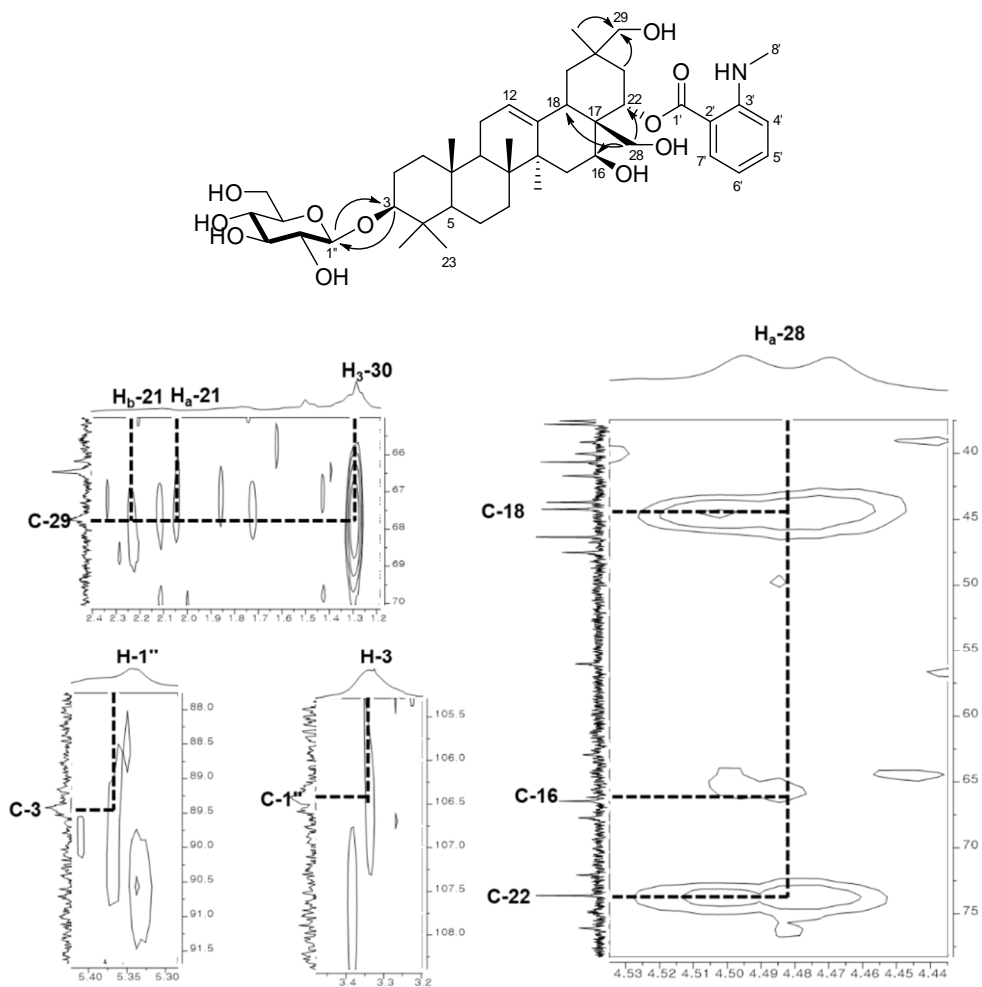
Figure 173. HMBC spectrum of compound **10**
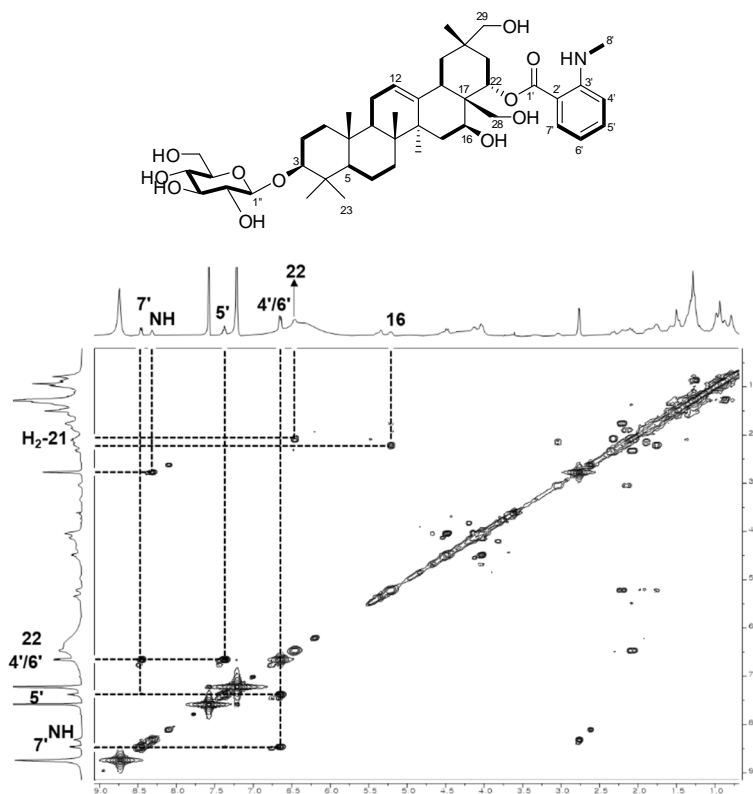
Figure 174. ^{1}H–^{1}H COSY spectrum of compound **10**
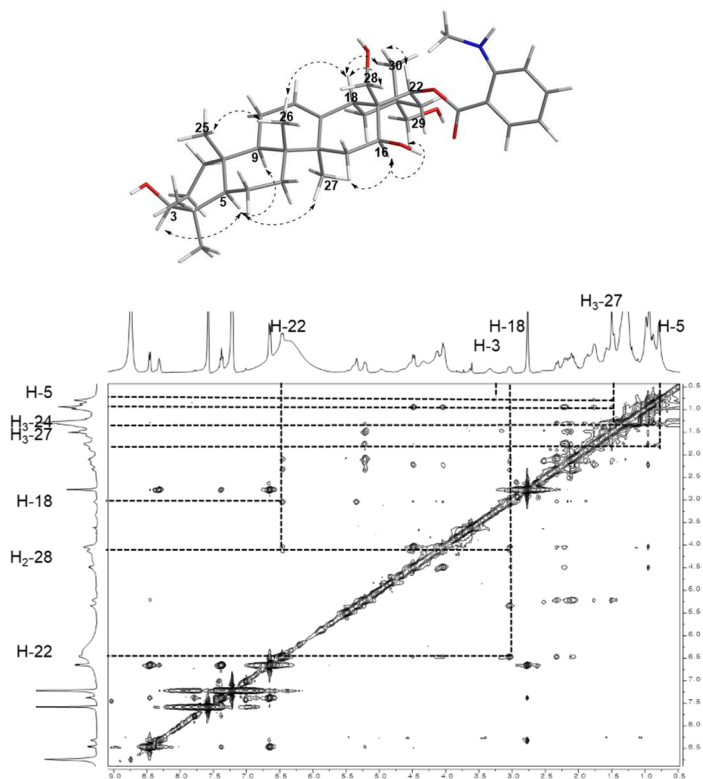
Figure 175. NOESY spectrum of compound **10**

3.11. Compound 11

Compound **11** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +27.2 (*c* 0.1, MeOH). The molecular formula $C_{51}H_{77}NO_{16}$ was determined from the HRESIMS ion peak at *m/z* 960.5447 [M + H]$^+$, (calcd for $C_{51}H_{76}NO_{16}$, 960.5433), indicating 13 degrees of unsaturation. The IR bands at 3376, 1678, and 1510 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **11** were similar to **1** except for the presence of an additional methoxy functional group at δ_H 3.74 (s)/ δ_C 52.6 ppm. The HMBC correlation from the methoxy signal (δ_H 3.74) to C-6" (δ_C 170.7) suggested that the additional 7"-OMe was attached at the carboxylic group position. Additionally, the resonance of the anomeric proton signal of glucose moiety showed a small coupling constant (δ_H 5.37, *J* = 3.4 Hz) suggesting an *α*-orientation of glucose unit in **11.** Therefore, compound **11** was identified as methyl (3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-*O*-*α*-D-glucopyranosyl-*β*-D-glucopyranosiduronate.
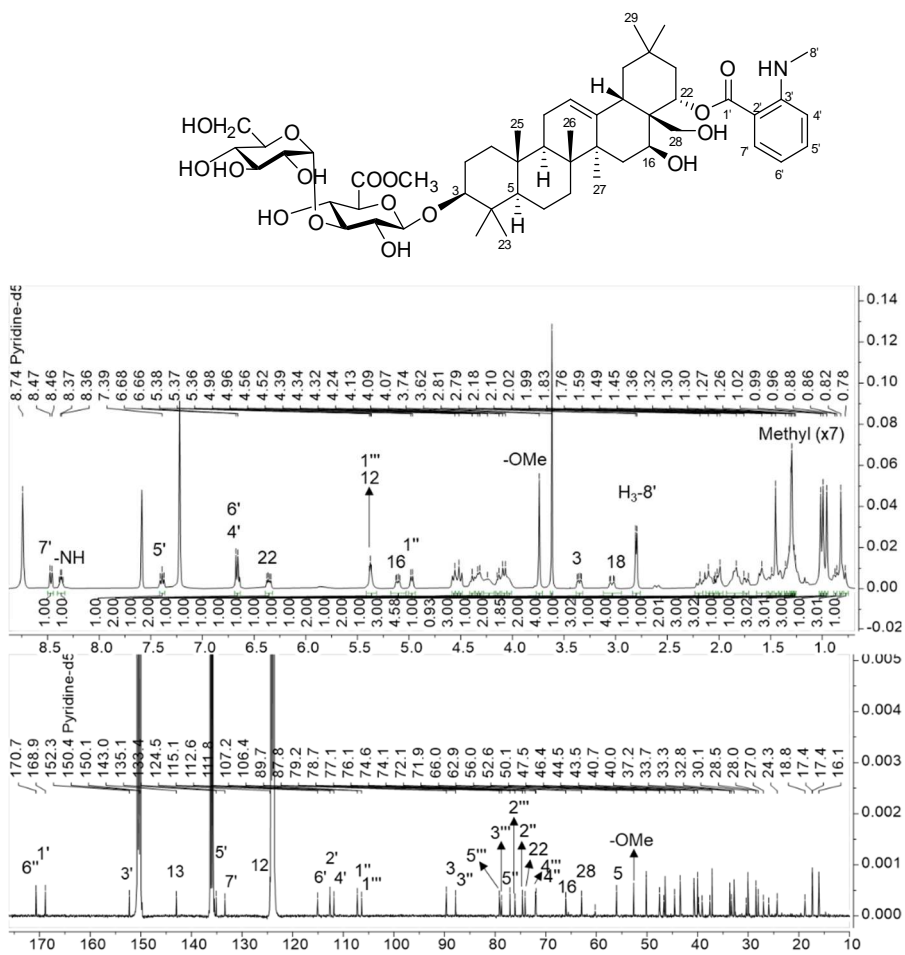
Figure 176. ^{1}H and ^{13}C NMR spectra of compound **11** (400/100 MHz, pyridine-d_5)
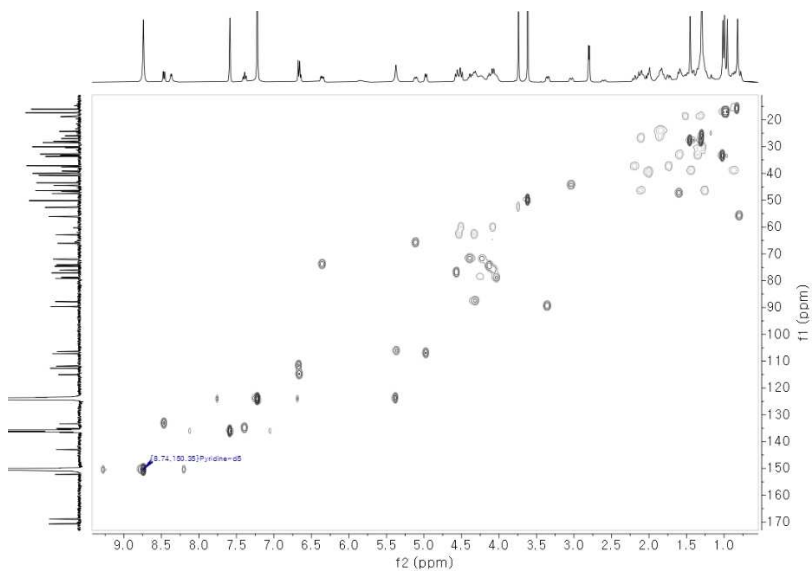
Figure 177. HSQC spectrum of compound **11**
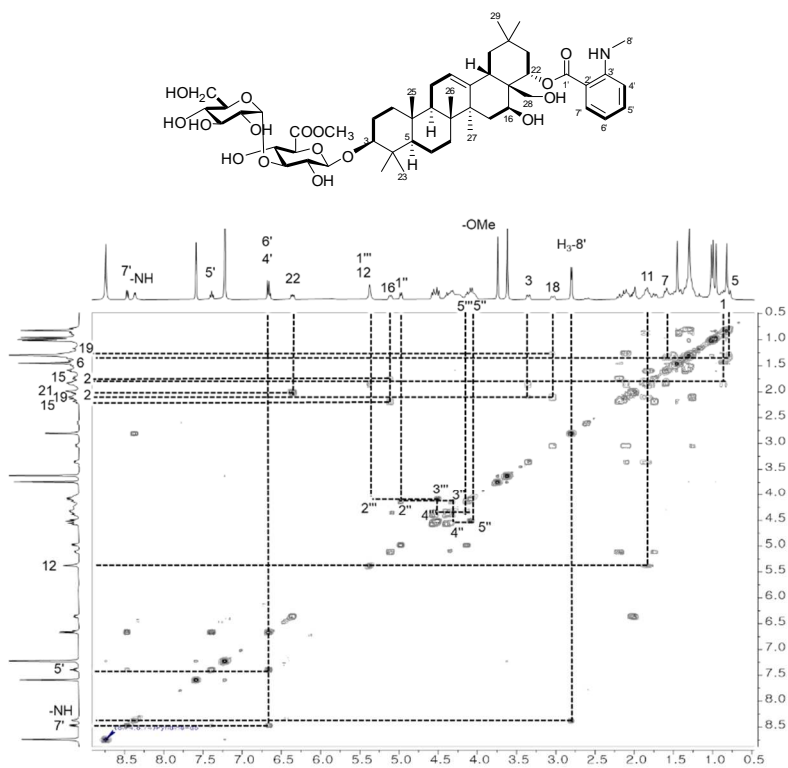
Figure 178. ^{1}H–^{1}H COSY spectrum of compound **11**
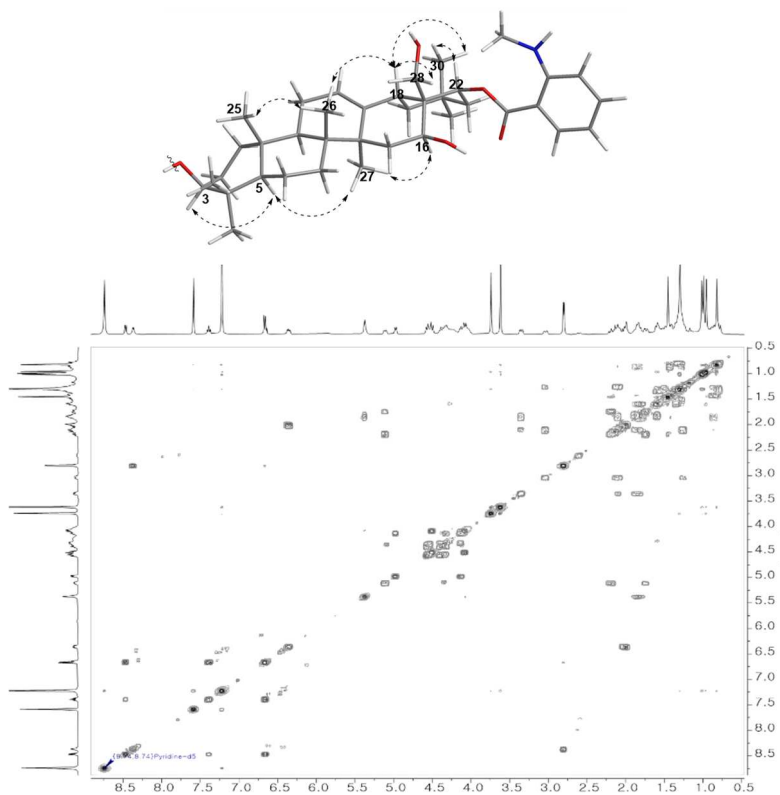
Figure 179. NOESY spectrum of compound **11**

3.12. Compound **12**

Compound **12** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +12.7 (*c* 0.1, MeOH). The molecular formula $C_{44}H_{65}NO_{11}$ was determined from the HRESIMS ion peak at *m/z* 784.4612 [M + H]$^+$, (calcd for $C_{44}H_{65}NO_{11}$, 784.4636), indicating 12 degrees of unsaturation. The IR bands at 3365, 1608, and 1514 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **12** were similar to those of the oleanane skeleton of **1** with *N*-methyl anthranilate moiety. The differences between **12** and **1** were observed in the sugar moieties that were indicated by their NMR spectroscopy. Instead of two sugar units in **1**, the structure of **12** was an absence of a glucose unit. Therefore, **12** was identified as (3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-*β*-D-glucopyranosiduronic acid.

Figure 180. ¹H and ¹³C NMR spectra of compound **12** (400/100 MHz, pyridine-d_5)

Figure 181. HSQC spectrum of compound **12**
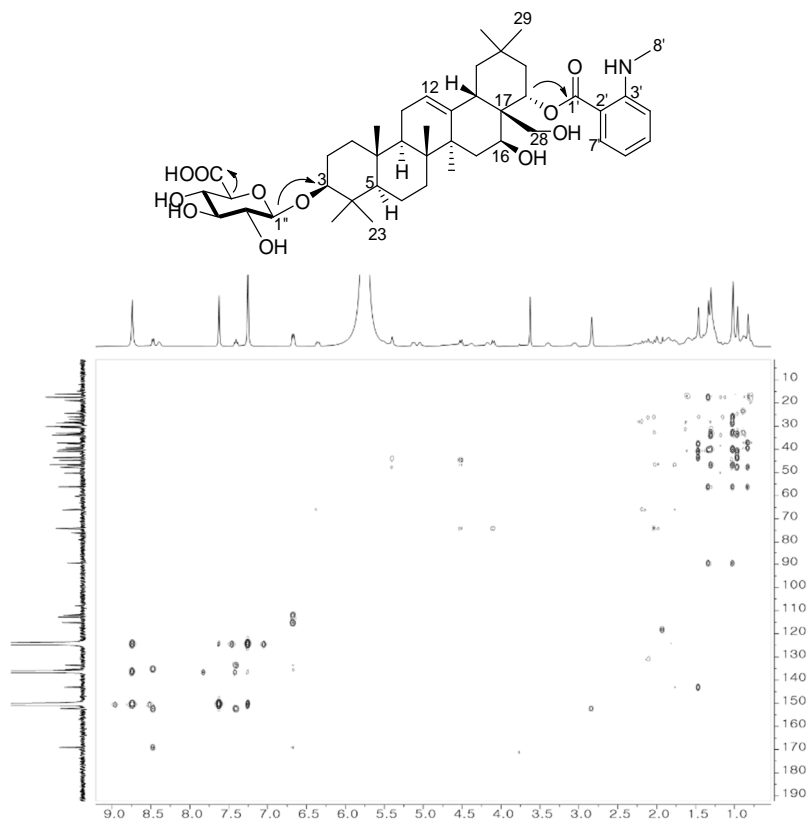
Figure 182. HMBC spectrum of compound **12**
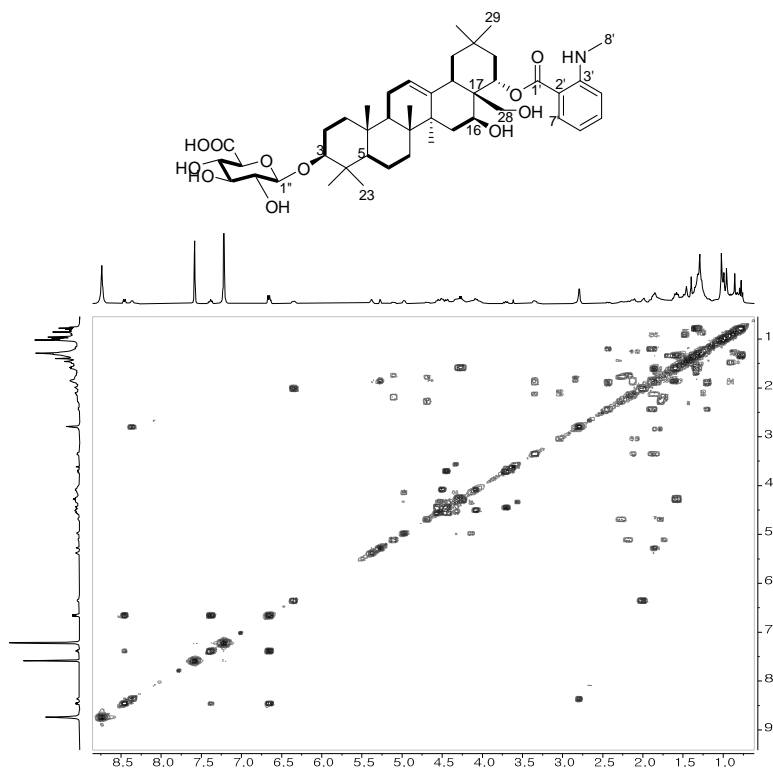
Figure 183. ^{1}H–^{1}H COSY spectrum of compound **12**

3.13. Compound **13**

Compound **13** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +10.6 (*c* 0.1, MeOH). The molecular formula $C_{51}H_{77}NO_{16}$ was determined from the HRESIMS ion peak at *m/z* 960.4907, (calcd for $C_{51}H_{76}NO_{16}$, 960.4957), indicating 12 degrees of unsaturation. The IR bands at 3379, 1608, and 1520 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **13** were similar to those of the oleanane skeleton of **8** with *N*-methyl anthranilate moiety. The differences between **13** and **8** were observed in the sugar moiety. The structure of **13** showed the presence of a glucose unit with an anomeric signal at δ_H 4.59 (*J* = 7.8 Hz)/ δ_C 150.1 ppm instead of the glucuronic acid unit as in **8**. The large coupling constant of the anomeric proton indicated that the sugar was in *β*-orientation, and the absolute configuration of sugar was identified to be *β*-D-glucose by comparing to authentic compounds after acid hydrolysis and sugar derivatizing. Therefore, **13** was identified as (3*β*,16*β*,22*α*)-22-(N-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-*β*-D-glucopyranoside.
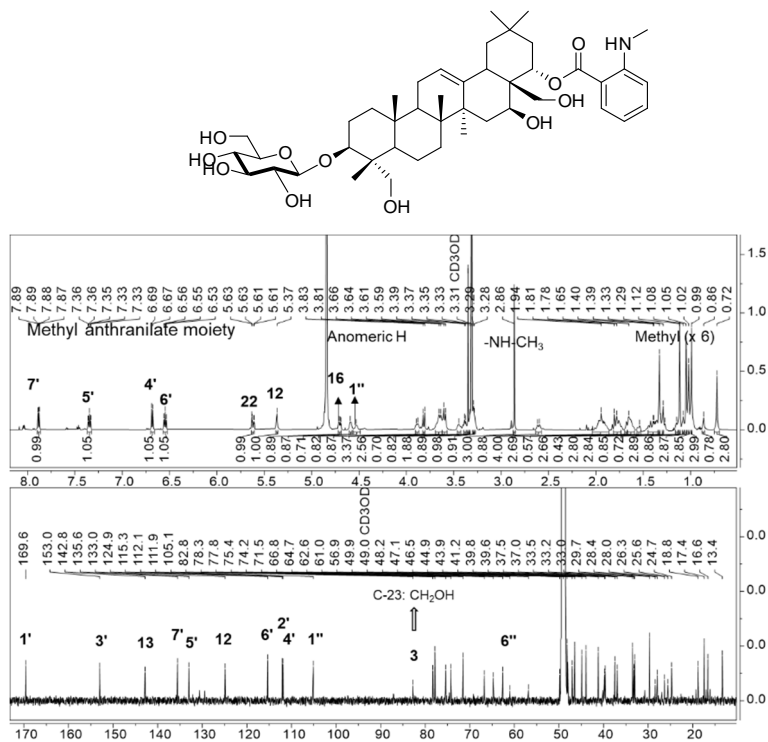
Figure 184. ^{1}H and ^{13}C NMR spectra of compound **13** (400/100 MHz, methanol-d_4)
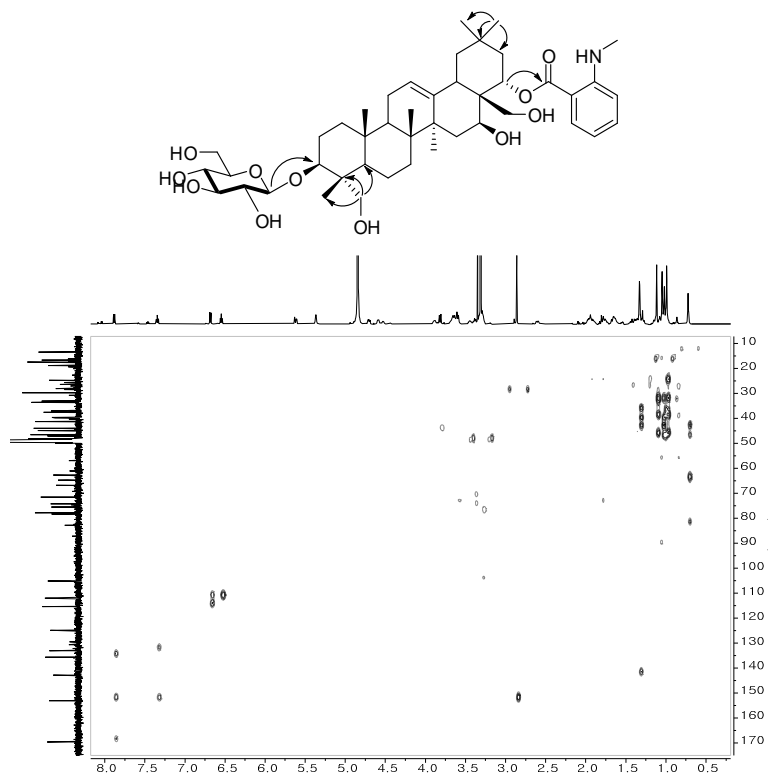
Figure 185. HMBC spectrum of compound **13**
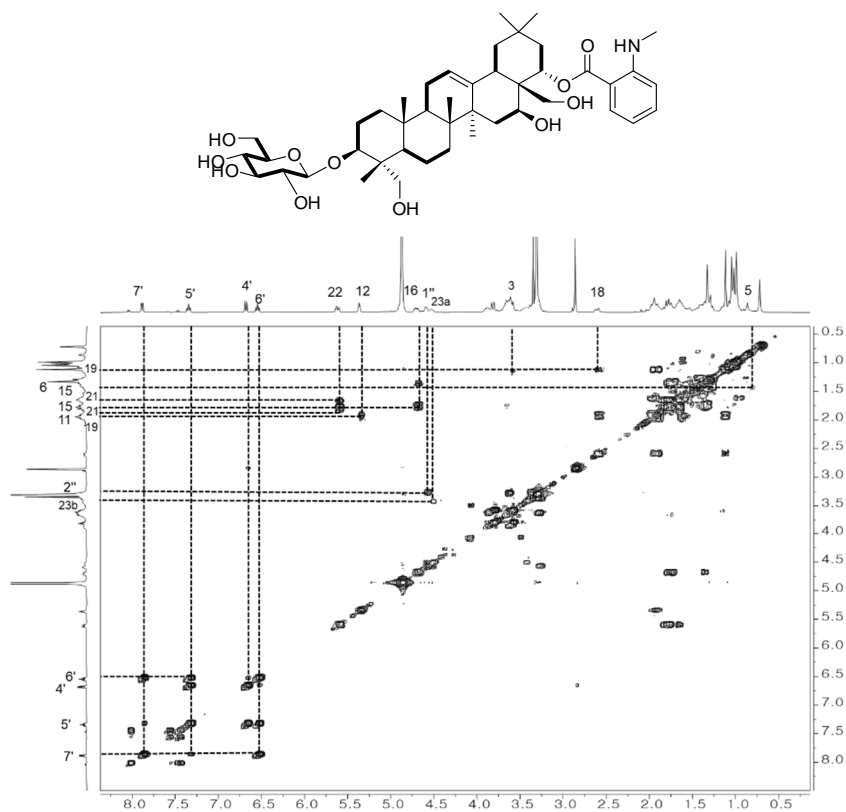
Figure 186. ^{1}H–^{1}H COSY spectrum of compound **13**
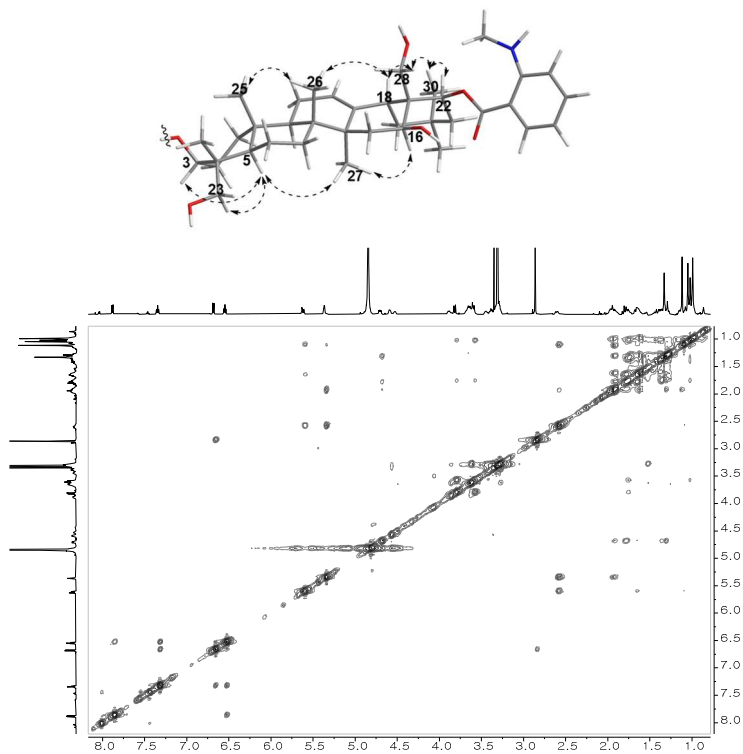
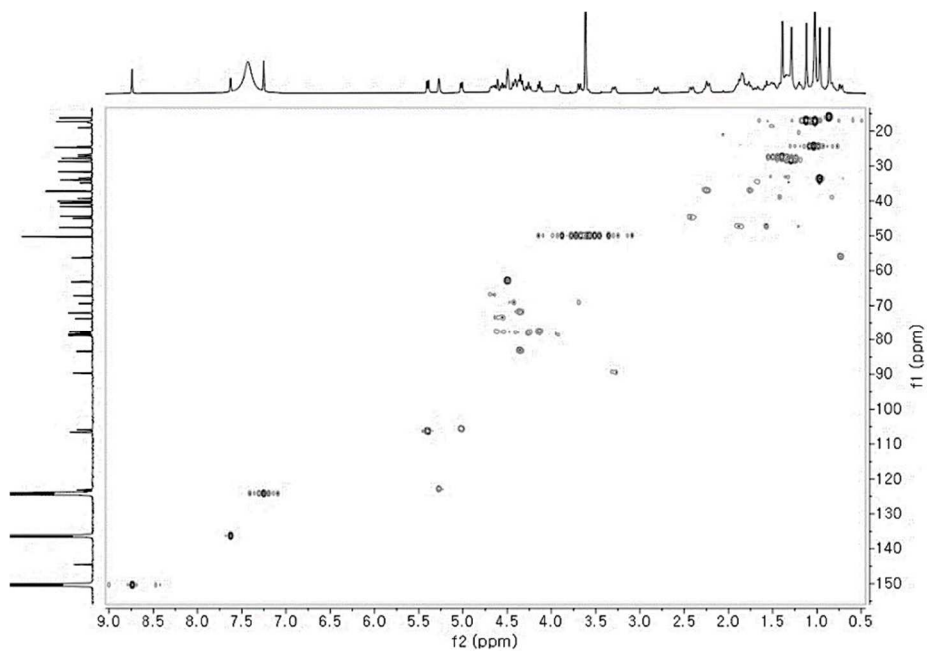**Figure 187.** NOESY spectrum of compound **13**

218

3.14. Compound 14

Compound **14** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +6.0 (*c* 0.1, MeOH). The molecular formula $C_{50}H_{75}NO_{17}$ was determined from the HRESIMS ion peak at *m/z* 962.5108 [M + H]$^+$, (calcd for $C_{50}H_{76}NO_{17}$, 962.5113), indicating 13 degrees of unsaturation. The IR bands at 3376, 1678, and 1510 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **14** were similar to those of oleanane skeleton and sugar moieties in **1** except for the position of the *N*-methyl anthranilate unit and an additional hydroxy group. The presence of the hydroxyl at C-22 was indicated by the additional oxygenated signal at C-22 δ_H 5.18 (*J* = 10.8)/ δ_C 71.7 ppm. And the positions of the 21-*N*-methyl anthranilate unit and 22-OH were revealed by the HMBC correlations from two methyl groups at δ_H 3.74 (H$_3$-29) and δ_H 3.74 (H$_3$-30) to C-21 (δ_C 80.1); and from H-22 to C-16 (δ_C 68.5) and C-28 (δ_C 58.6). The COSY correlation between H-21 and H-22 was also supported for the identification of H-21/H-22. The large $^2J_{HH}$ (10.8 Hz) between H-21, and H-22 suggested that they were oriented in different faces. Thus, **14** was identified as methyl (3*β*,16*β*,21*β*,22*α*)-21-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-*O*-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronate.
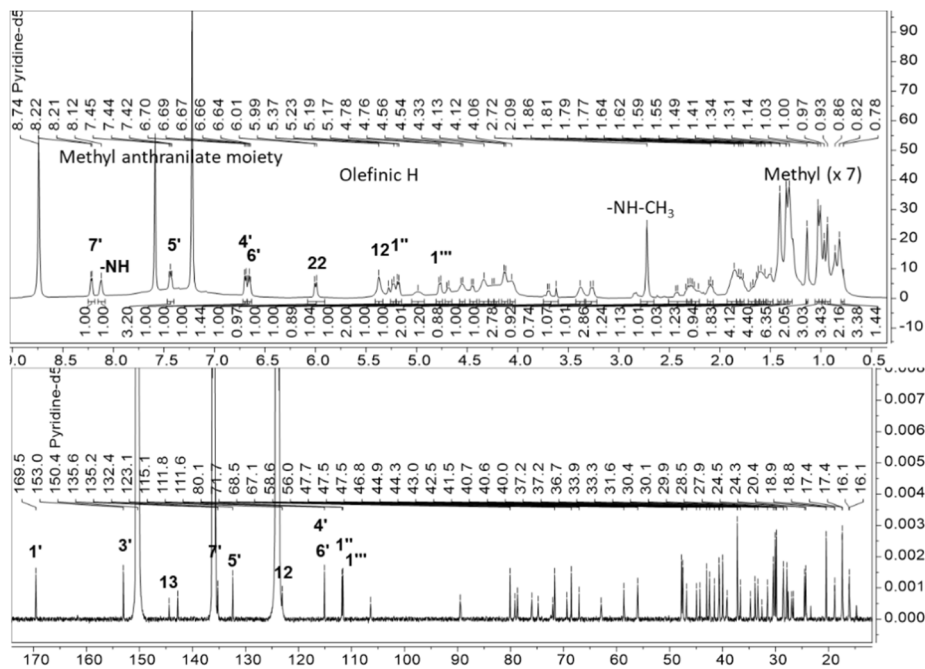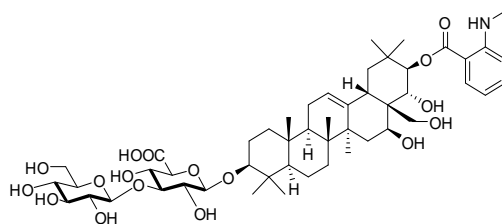
Figure 188. ^{1}H and ^{13}C NMR spectra of compound 14 (400/100 MHz, pyridine-d_5)
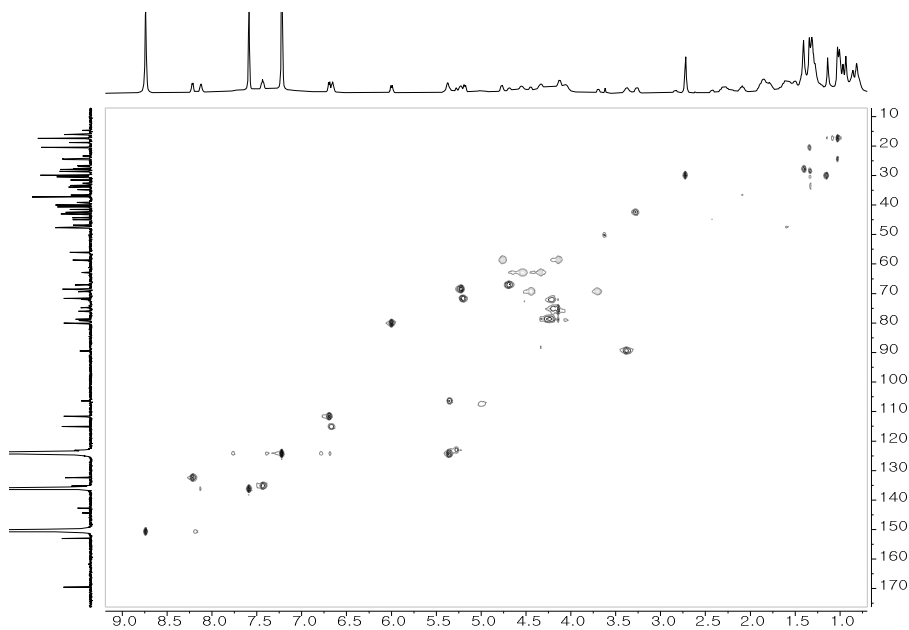
Figure 189. HSQC spectrum of compound 14
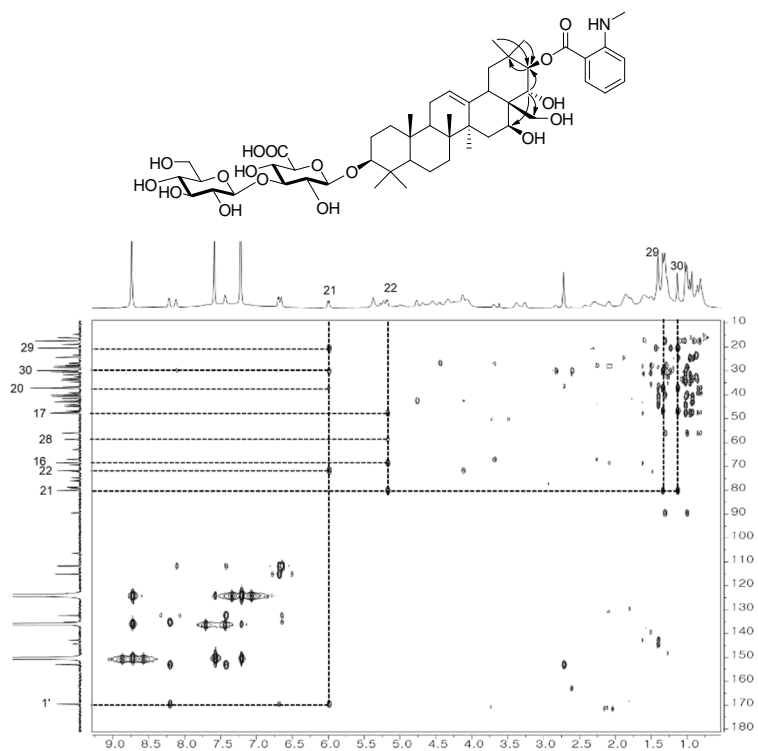
Figure 190. HMBC spectrum of compound **14**
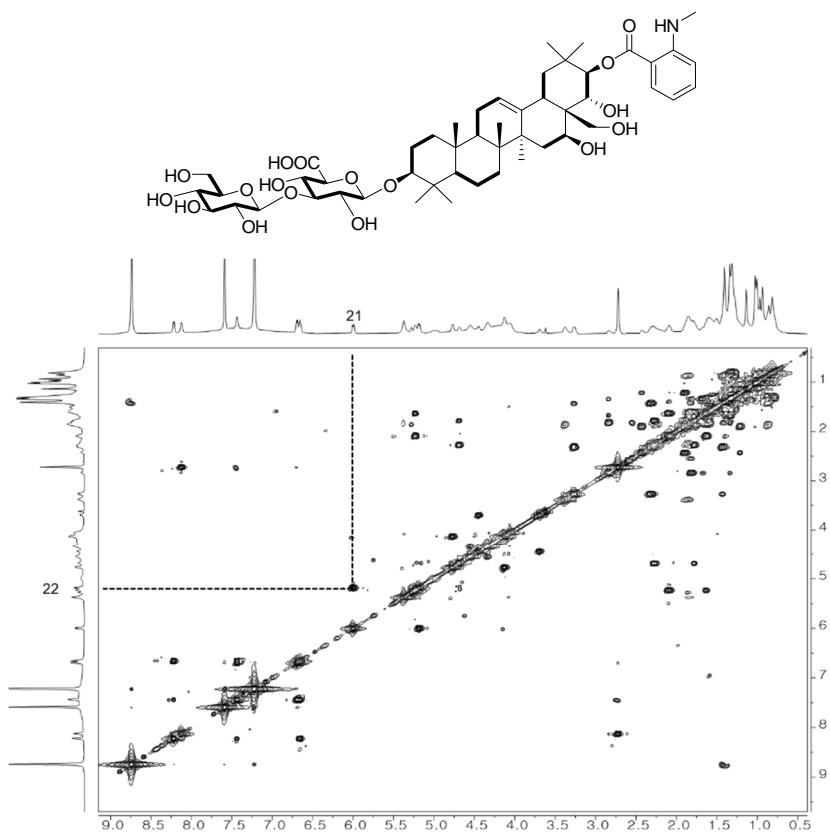
Figure 191. ^{1}H–^{1}H COSY spectrum of compound **14**
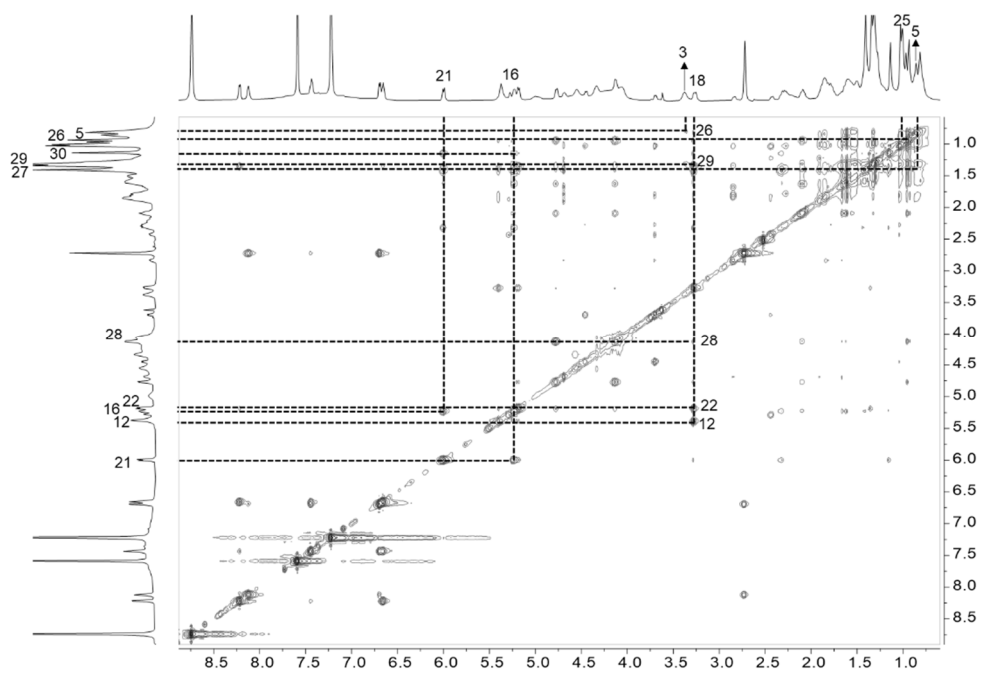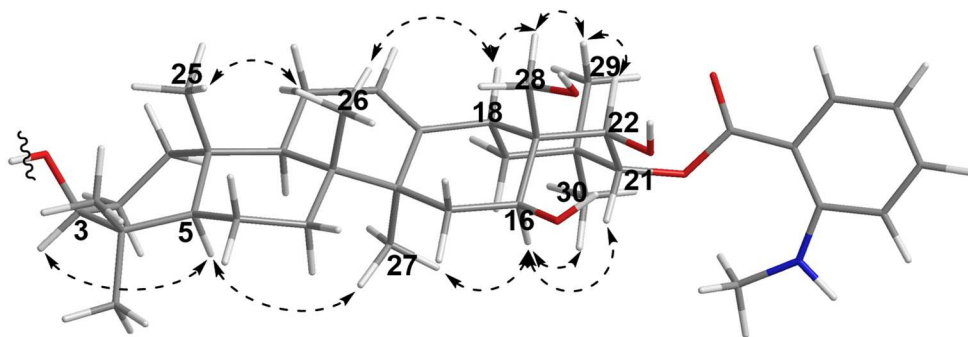
Figure 192. NOESY spectrum of compound **14**

3.15. Compound 15

Compound **15** was identified as (3β,16β)-16,28-dihydroxyolean-12-en-3-yl-β-D-glucopyranosiduronic acid based on the reported compounds in literatures (H.-T.-T. Pham et al., 2019; Ye et al., 2000).
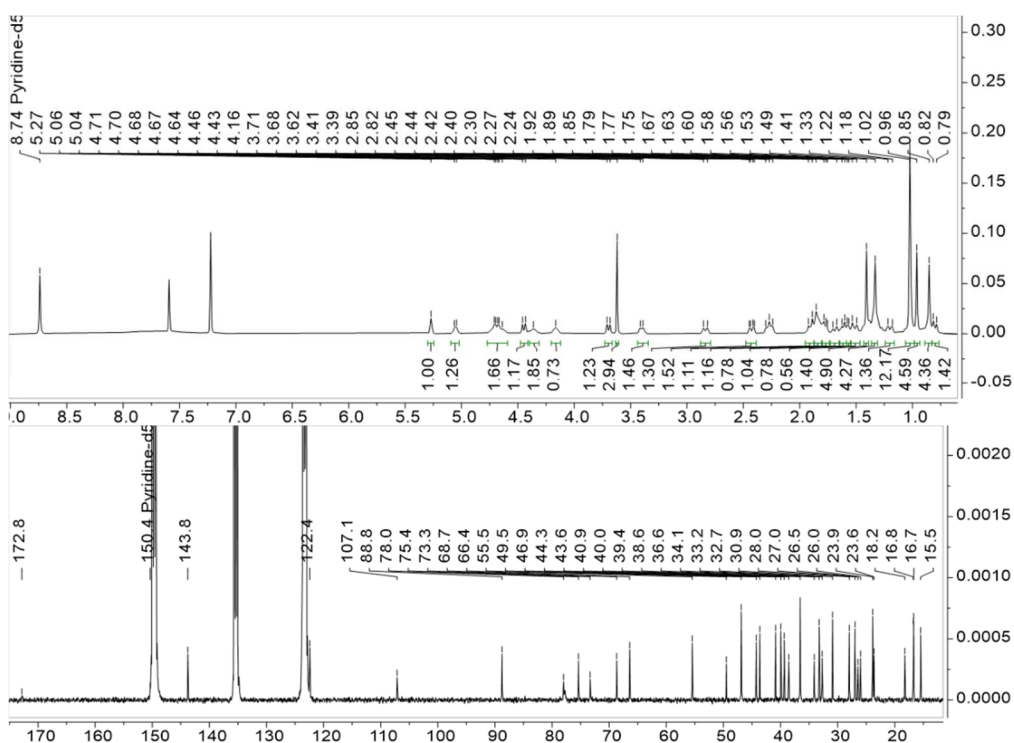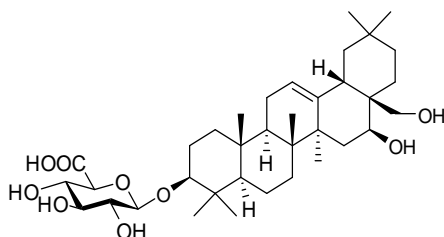
Figure 193. ^{1}H and ^{13}C NMR spectra of compound **15** (400/100 MHz, pyridine-d_5)

3.16. Compound 16

Compound **16** was identified as (3*β*,16*β*)-16,28-dihydroxyolean-12-en-3-yl-3-O-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronic acid based on the reported compounds in literature (H. T. T. Pham, Hoang, et al., 2018).
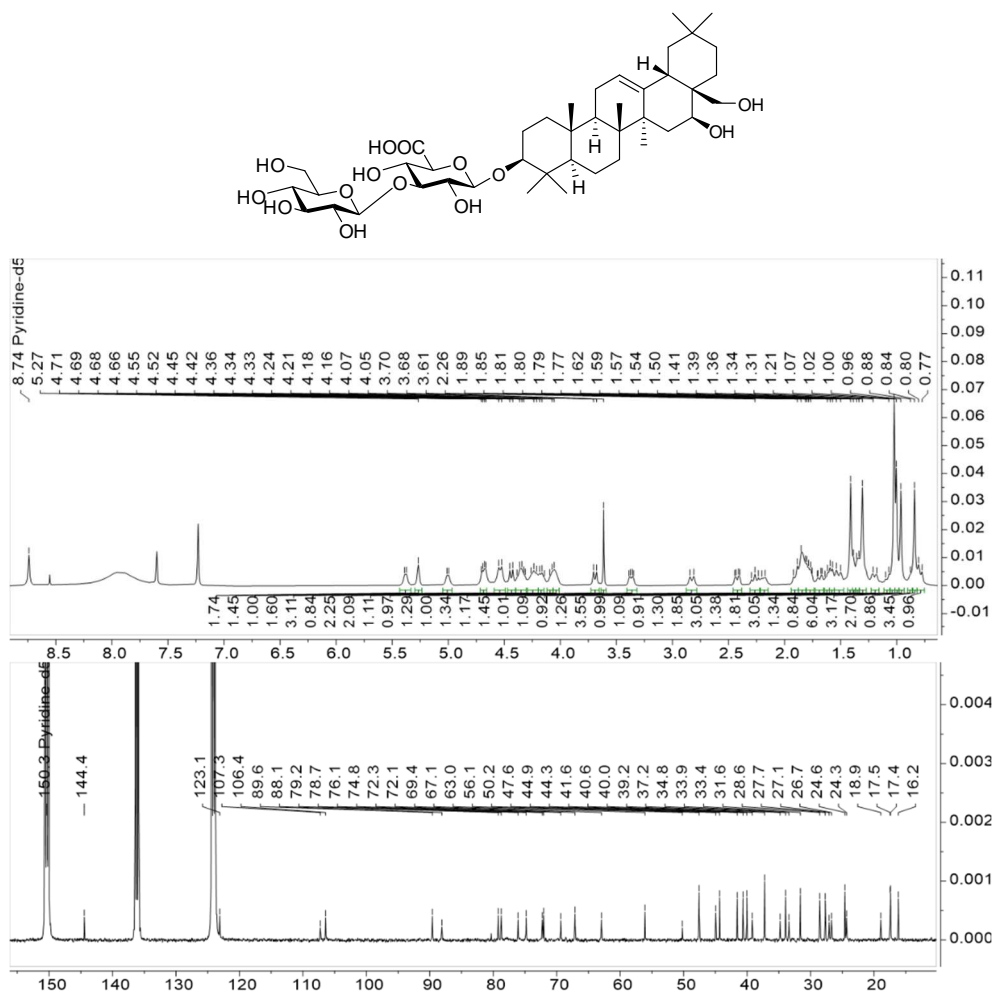
Figure 194. ^{1}H and ^{13}C NMR spectra of compound **16** (400/100 MHz, pyridine-d_5)

3.17. Compound **17**

Compound **17** was obtained as a white, amorphous powder with an $[\alpha]_D^{25}$ +30.0 (*c* 0.1, MeOH). The molecular formula $C_{51}H_{71}NO_{16}$ was determined from the HRESIMS ion peak at *m/z* 960.5349 [M + H]$^+$, (calcd for $C_{51}H_{71}NO_{16}$, 960.5321), indicating 13 degrees of unsaturation. The IR bands at 3379, 1608, and 1525 cm^{-1} were indicative of NH stretch, olefinic, and NH bend absorbances. The ^{1}H and ^{13}C NMR data of **17** share a similarity to those of **11** except for the presence of a coupling constant of the anomeric proton in the glucose unit. Compound **17** showed a large *J* couping at H-1''' (δ_H 5.29, *J* = 7.8 Hz) indicating that glucose was attached in *β*-orientation. Therefore, **17** was identified as methyl (3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-*O*-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronate.
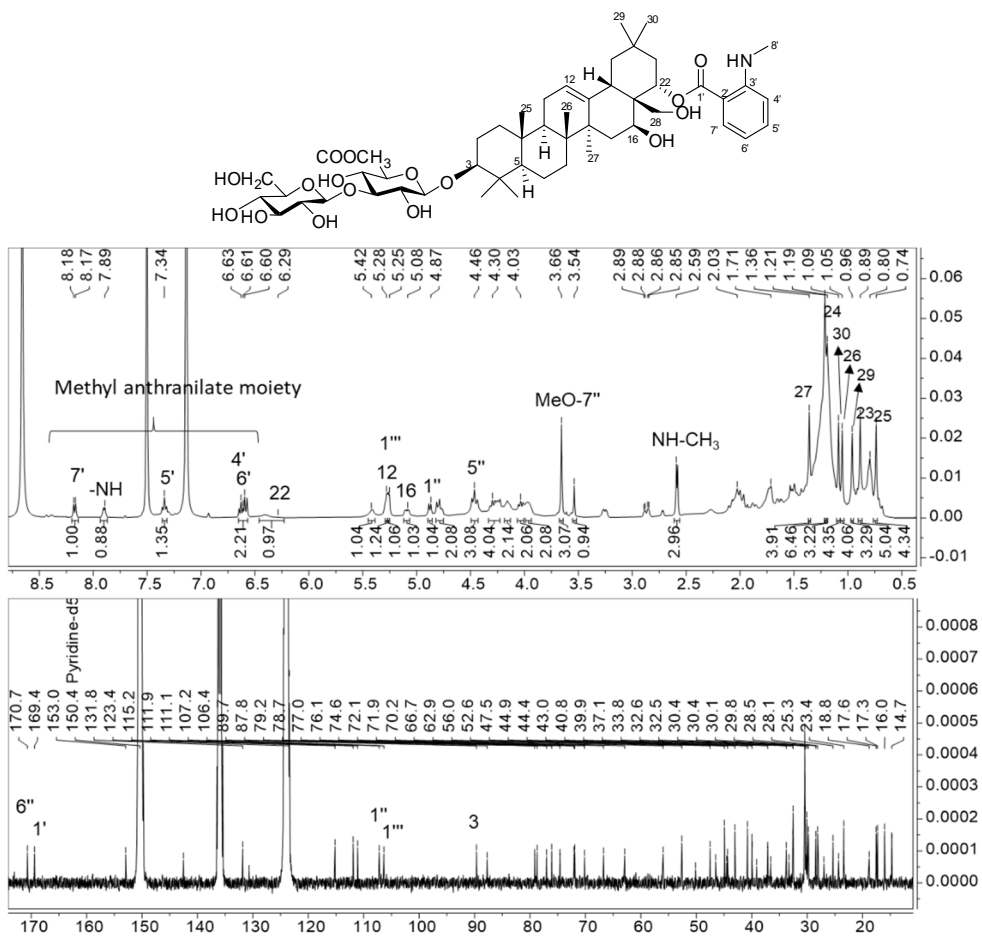
Figure 195. ^{1}H and ^{13}C NMR spectra of compound **17** (400/100 MHz, pyridine-d_5)
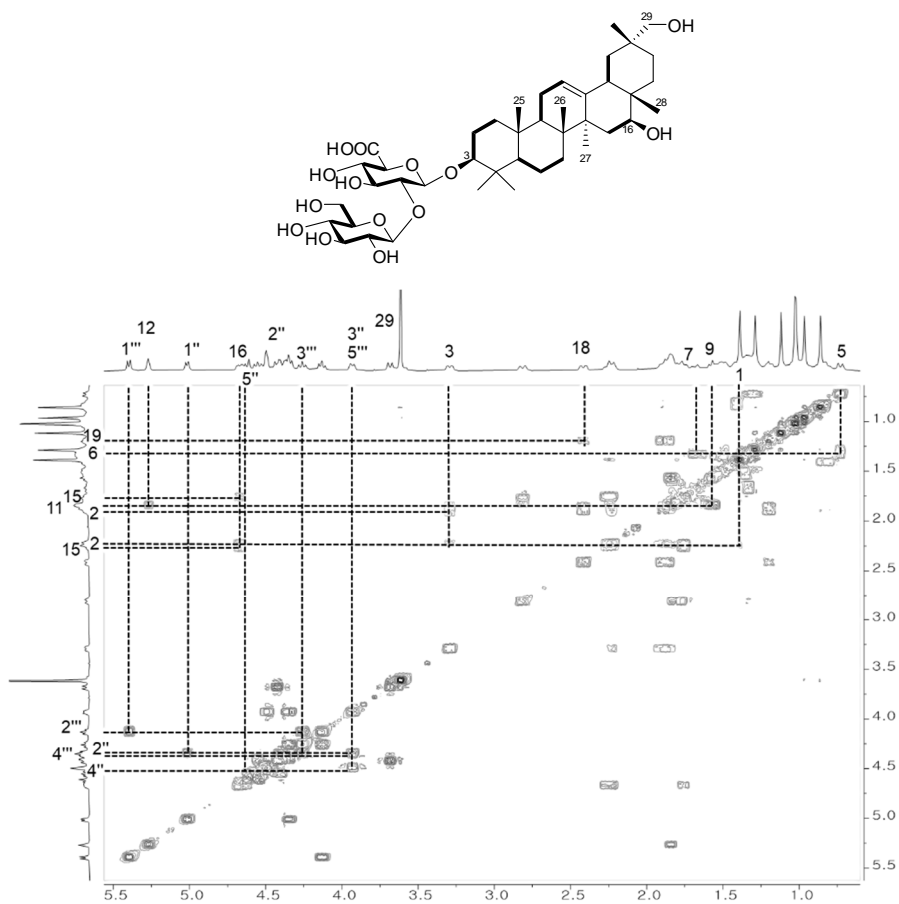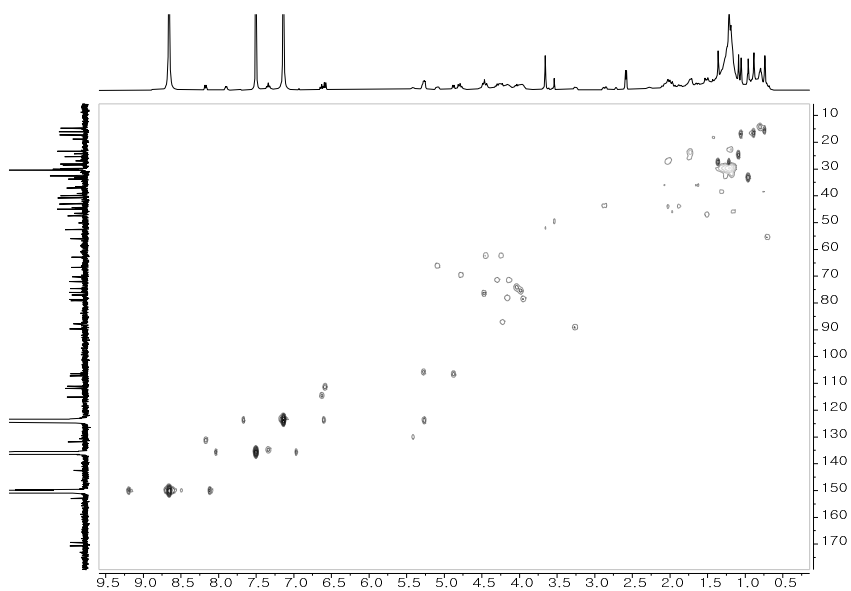

Figure 196. HSQC spectrum of compound **17**
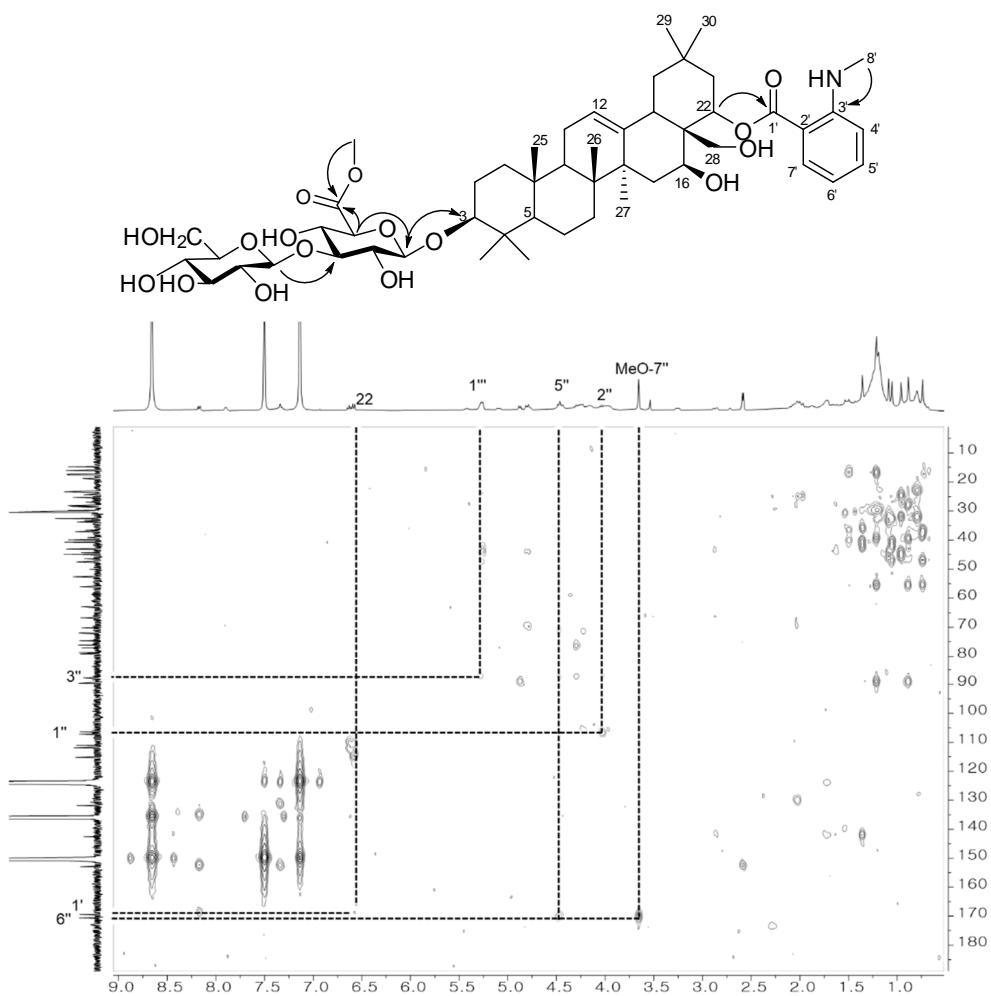
Figure 197. HMBC spectrum of compound **17**
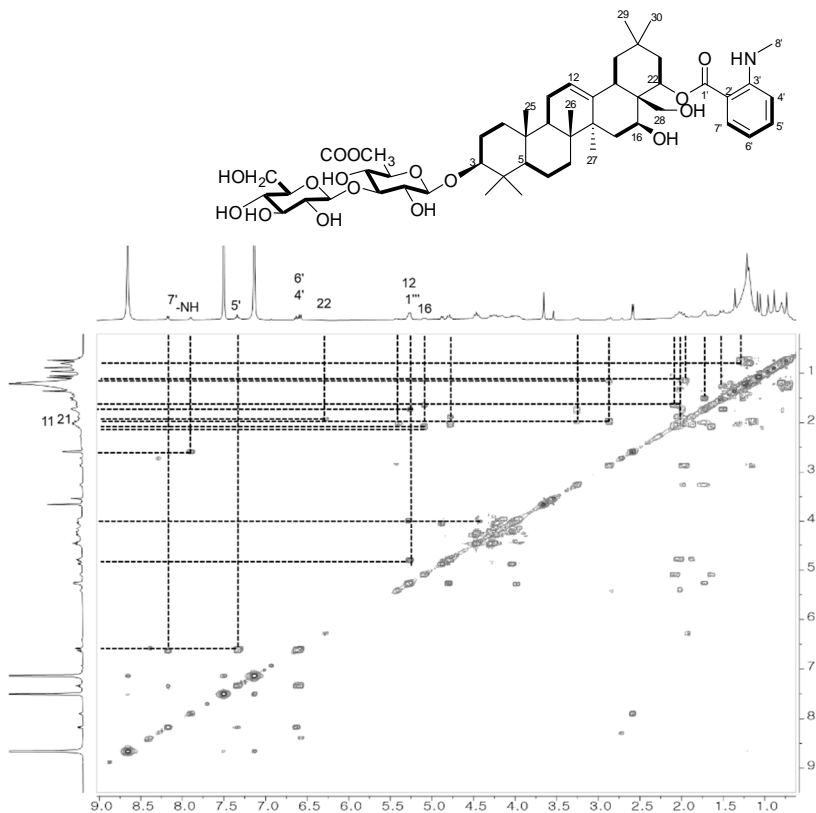
Figure 198. ^{1}H–^{1}H COSY spectrum of compound **17**
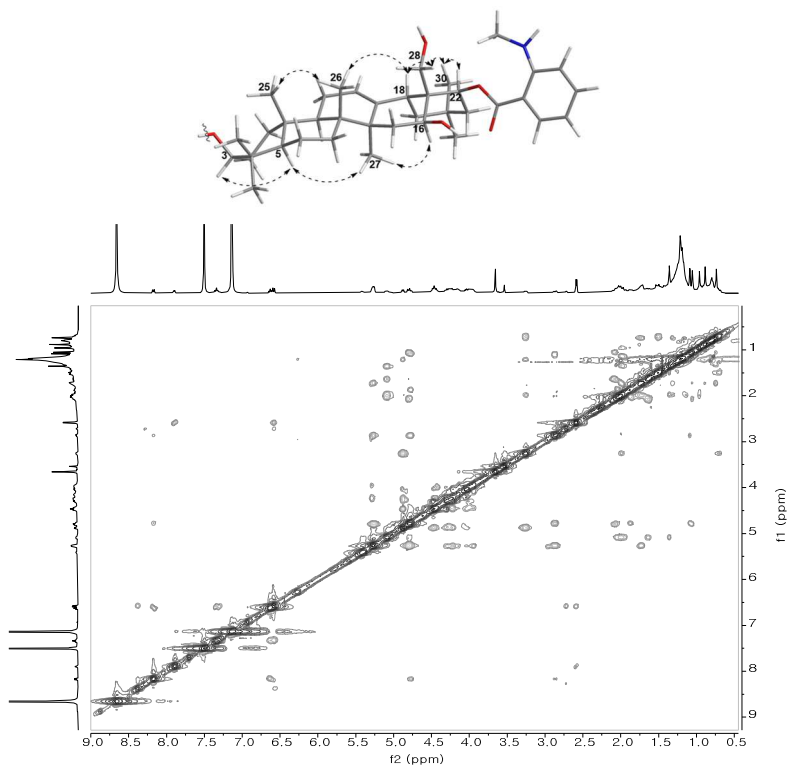
Figure 199. ROESY spectrum of compound **17**

228

3.18. Compound 18

Compound **18** was obtained as the aglycon of **1** by the acid hydrolysis reaction. The structure of **18** was confirm based on its 1H NMR data indicated the same pattern as **1** except the absence of two sugar moiety as in **1**. Hence, compound **18** was identified as (3β,16β,22α)-22-(N-methylanthraniloxy)-3,16,28-trihydroxyolean-12-ene.
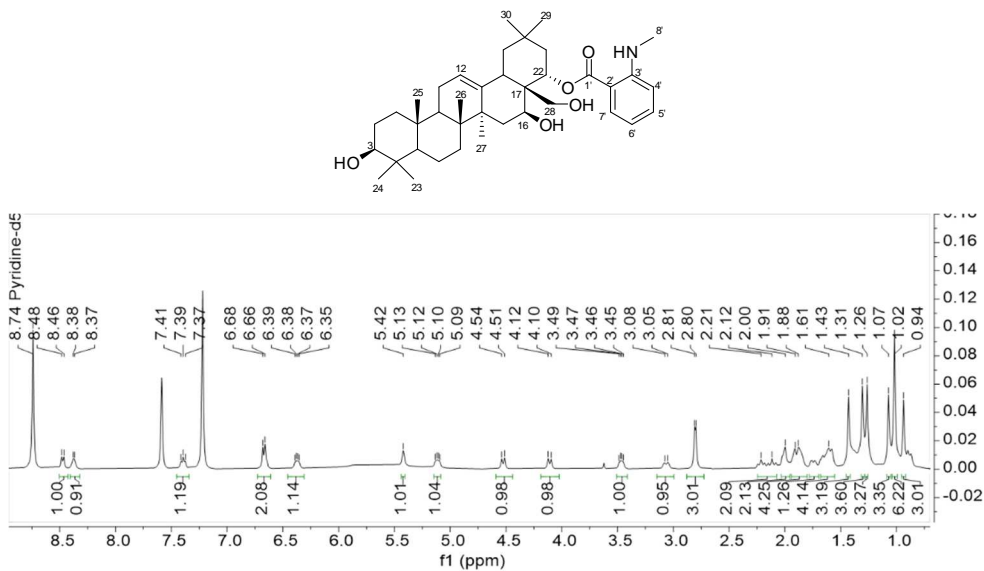
Figure 200. ^{1}H and ^{13}C NMR spectra of compound **18** (400/100 MHz, pyridine-d_5)

4. Biological activities of compounds 1–18

4.1. Biological activities of compounds **1–9**

To examine the glucose uptake of isolates **1–9**, the uptake of 2-deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-glucose (2-NBDG) in 3T3-L1 adipocytes was measured. 2-NBDG, a fluorescent glucose analogue, is used for determining the effects on glucose uptake to find insulin mimetic agents. The isolated compounds at a concentration of 20 μM were added to the differentiated 3T3-L1 adipocytes with 2-NBDG (Figure 201A). Dimethyl sulfoxide (DMSO) and insulin were used as negative and positive controls, respectively. An MTT assay was carried out at 20 μM in advance (Figure 201A). Among all isolates, compounds **3**, **5**, **8**, and **9** showed significant 2-NBDG uptake effects at 20 μM, and these four potential compounds were selected to analyze their glucose uptake effect as several different concentrations (5, 10, and 20 μM). Another MTT assay was performed with three different concentrations (5, 10, and 20 μM) to determine a safe dose of the compounds tested (Figure 201B). Four compounds (**3**, **5**, **8**, and **9**) increased glucose uptake in a dose-dependent manner (Figure 202B). Fluorescent signals were measured using fluorescent microscopy for analyzing the transport efficacy of 2-NBDG into cells. Fluorescent intensities of cells treated with compounds **3**, **5**, **8**, and **9** (20 μM) showed higher levels than a positive control Figure 202C. These results are in good agreement with previous studies showing that extracts of *G. inodorum* lowered the postprandial peak of plasma glucose level in healthy humans (Anchalee et al., 2010).

A
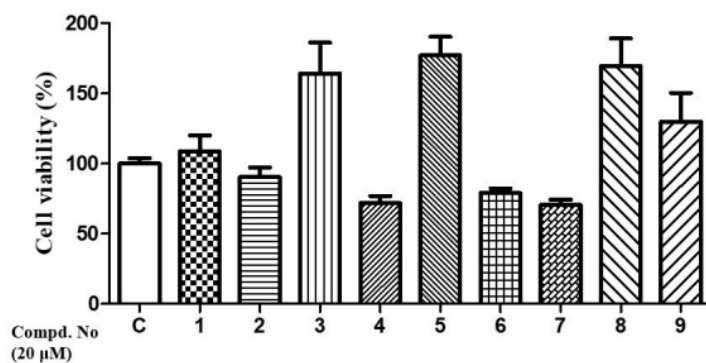
B
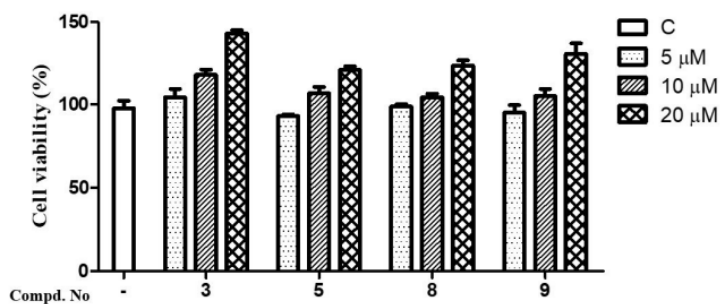
Figure 201. The cell proliferation rate of **1–9** from *G. inodorum* in 3T3-L1 adipocytes. (A) Compounds **1–9** were treated at a concentration of 20 μM in 3T3-L1 adipocytes. (B) Compounds **3, 5, 8**, and **9** were treated at different concentrations (5, 10, and 20 μM) in 3T3-L1 adipocytes.

231
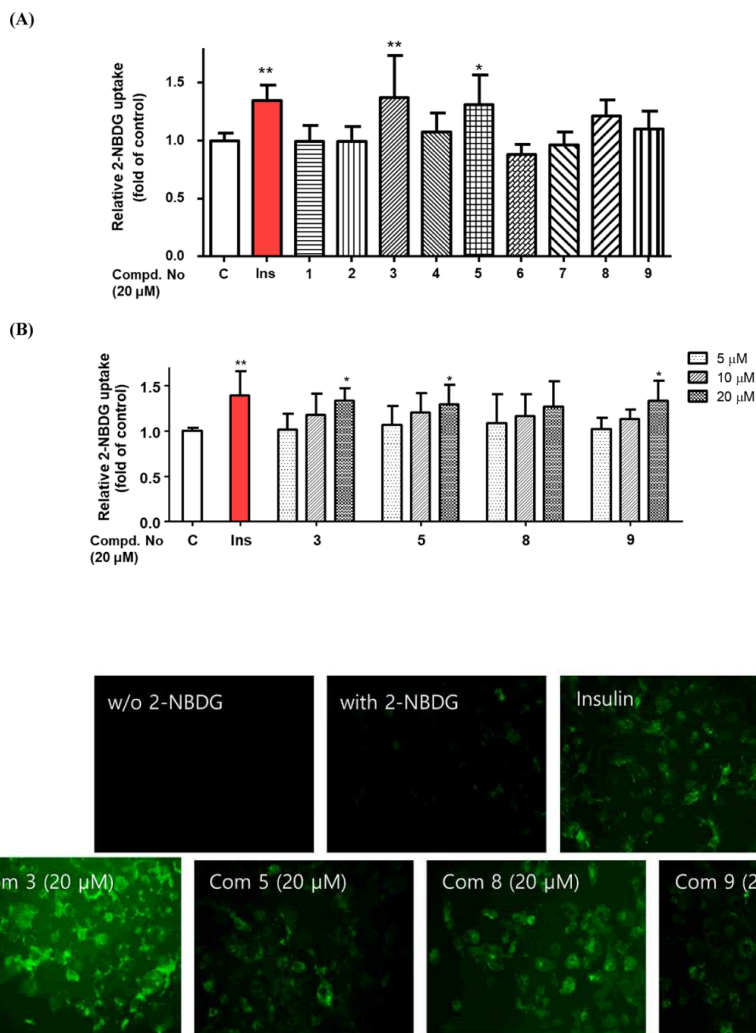
Figure 202. Glucose uptake activities of compounds **1–9** in 3T3-L1 adipocytes

(A) Stimulation effects of compounds **1–9** on glucose uptake in 3T3-L1 adipocytes using a fluorescent analogue of glucose (2-NBDG). 3T3-L1 adipocytes were exposed to 100 nM insulin and the test compounds for 1 h in the presence of 2-NBDG. Glucose uptake was measured at ex/em = 450/535 nm using a fluorescence microplate reader. The results were calculated as the means $\pm$ SD (n = 3). Green fluorescent signals were measured and expressed as the means $\pm$ SDs (n = 3); $^{*}p <$ 0.05, $^{**}p < 0.01$, and $^{***}p < 0.001$. (B) 3T3-L1 adipocytes were exposed to compounds **3, 5, 8**, and **9** at various concentrations (5, 10, and 20 μM) for 1 h. (C) The cells were examined using a fluorescence microscope (treated with 20 μM). The green fluorescent signals significantly increased, which indicated the successful transport of 2-NBDG into these cells.

4.2. Biological activities of compounds 1, 8–18

The cell proliferation ratio of compounds **1**, **8–18** was performed in dexamethasone (Dexa) plus LPS model on myotube at a concentration of 2 μM. The result shown in Figure 203 indicated that the cell proliferation ratios of all tested compounds increase compared to the negative control. Furthermore, compounds **1**, **8**, and **12** were relatively high compared to the positive control (Figure 203). This finding suggested further biological activity studies.
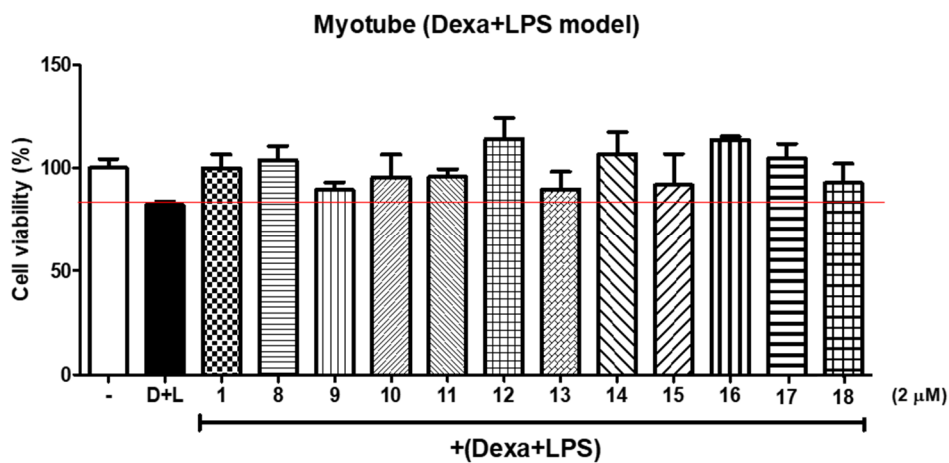
Figure 203. Cell proliferation ratio of compounds **1**, **8–18** in myoblast C2C12 cell line

In muscle atrophy two of E3 ubiquitin ligases which are MuRF-1 (muscle ring finger 1) and MAFbx (muscle atrophy F-box) have been reported to be increased (Bodine et al., 2001). Therefore, to screening effects of isolated compounds on muscle atrophy, the MuRF-1 level has been tested by PRR method in dexamethasone (Dexa) model on myotube. The results shown in Figure 204 indicated that compound **12** displayed the best activities with the lowest level of MuRF-1 detected. Compounds **8**, **9**, **15**, **16**, and **18** also exhibited the strongly decrease of MuRF-1.
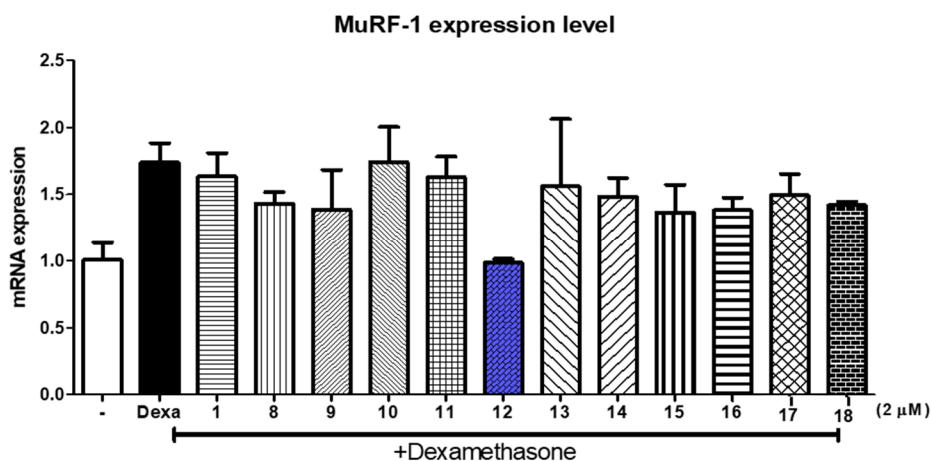
Figure 204. MuRF-1 expression levels of compounds **1**, **8–18** in myoblast C2C12 cell line

4.3. Structure-activity relationships (SARs) discussion and further studies

G. inodorum and other species in the genus of *Gymnema* have been widely used in diabetes treatment. The phytochemical and biological studies suggested that triterpenoids from this genus may contribute to the antidiabetes effect. Herein, the insulin-mimetic activities of compounds **1–9** that showed oleanane skeleton with some different substituents were tested for their activities. Based on the structures of **1–9** and the better activity of **3**, **5**, **8**, and **9** compared to the others in glucose uptake effects, a gross SAR for this triterpenoid type can be delineated as follows:

(1) compounds **1** and **9** have the same oleanane skeleton with an *N*-methyl anthranilate substituent at C-22 and 2 sugars at C-3. The only difference in both compounds was observed at C-23 (23-Me in **1** and 23-CH$_2$OH in **9**). As shown in Figure 201, **9** exhibited good activities compare to insulin as the positive control, but it was not observed the same in **1**. Therefore, the 23- CH$_2$OH may be important for activities of the oleanane skeleton in the insulin-mimetic effect; (2) Compare to **9,** compound **8** which also contains a methylene hydroxyl at C-23 in its structure has one sugar less at C-3 but this compound also showed as good activities like **9**. This evidence suggested that the number of sugar chains may not affect the activity; (3) Compound **3** contains one different substituent at C-22 compared to **8**, however, both compounds showed activities in the biological assay indicating that the 22-substituent may not affect the activities of this type of compound; (4) Compound **2**

234

shared a similar structure to **8** but has an additional methylene hydroxyl at C-29 and showed no activity. This finding may suggest the position of hydroxy methylene and the number of that substituents may contribute to the activity. For further confirmation of this hypothesis, the other structure with only 29-CH_2OH should be evaluated together; (5) compound 5 showed a methyl-substituted at C-17 without any other hydroxy functional groups but the activity was also seen. This result suggested that the 17-Me may play an important role in the activity of the oleanane skeleton.

A manner was applied to revise the SAR in muscle atrophy activities of compounds **1**, **8**–**18**. (1) the N-Mant at C-22 may contribute to the activity that reduce of MuRF-1 level when compared compounds **12** and **15**, but the position of N-Mant on C-21 or C-22 of oleanane skeleton may not affect to activity (compound **14** versus **1**); (2) the oleanane skeleton that contain one sugar showed better activity compared to aglycon and two-sugar glycoside (activity **12** > **18** > **1**); (3) the 24-CH_2OH substituent may reduce the activity that were seen in **8** compared to **12**, however, 24-CH_2OH (**13**) showed better effect than 29-CH_2OH (**10**) ; (4) the Glucuronic acid or glucose moieties at C-3 may not affect to the activity indicated by compound **8** (3-GlcA) similar to **13** (3-Glc).

Interestingly, the plant can produce a variety of derivatives by itself that could be well visualized with the aid of modern techniques such as UPLC-qTOF-MS/MS and molecular networking or other similar tools. This inspires us to quickly access their phytochemicals and continue to research SAR studies and also find out more mechanisms of activities.

5. Experimental Section

5.1. Materials

5.1.1. Plant material

The leaves of *G. inodorum* were collected in September 2018 in Hoai Duc district, Hanoi City, Vietnam (20°59′30.6″ N 105°43′49.8″ E). The sample was identified botanically by Dr. Ha Thanh Tung Pham, Department of Botany, Hanoi University. A voucher specimen (SNU2018-11) was deposited in the herbarium of the College of Pharmacy at Seoul National University.

5.1.2. Chemicals and reagents

Column chromatography (CC) was performed by using Diaion HP-20 resin (250–850 μm particle size, Merck, Germany), and medium-pressure liquid chromatography (MPLC) was carried out with a Biotage Isolera One MPLC system equipped with an RP-C$_{18}$ column (120 g, Grace), Sephadex LH-20 (Sigma-Aldrich, Inc., St. Louis, MO, USA), and RP-C$_{18}$ (40–63 μm particle size, Merck, Germany). TLC silica gel 60 RP-C$_{18}$ F254 plates (Merck, Darmstadt, Germany) were used for thin-layer chromatography. Industrial grades of MeOH, EtOH, and EtOAc were used for extraction, partition, and fractionation steps. Analytical grades of MeOH, MeCN, formic acid (FA), and trifluoroacetic acid (TFA) were used for isolation. All solvents were purchased from DaeJung Pure Chemical Engineering Co Ltd. (Siheung, Republic of Korea). Deuterated NMR solvents and dimethyl sulfoxide (DMSO) were purchased from BK Instuments Inc. (Daejeon, Republic of Korea).

5.1.3. Equipment and software

Optical rotations were recorded on a JASCO P-2000 polarimeter (JASCO International Co. Ltd., Tokyo, Japan) with a 10 mm path length cell at a specified temperature. UV and ECD spectra were obtained from a ChirascanPlus CD spectrometer (Applied Photophysics, Leatherhead, UK). IR spectra were measured on a JASCO FT/IR-4200 spectrometer (Thermo Electron Corp., Waltham, MA, USA). NMR data were measured by using AVANCE-500 MHz (Bruker, Rheinstetten, Germany) and JNM-ECA-600 MHz (JEOL, Tokyo, Japan) spectrometers coupled with a 5 mm cryoprobe (Bruker, Germany). NMR spectra were processed on MestReNova software (Mastrelab Research, USA). HRESIMS/MS data were

collected by an UHPLC-ESI-qTOF-MS system with a Waters Xevo G2 QTOF mass spectrometer (Waters MS Technologies, UK), a Water BEH C18 (100×2.1 mm, 1.7 μm, Water, USA), and analyzed by MassLynx SCN 855 software. The preparative HPLC system used for isolation included a Gilson 321 pump and Gilson UV/VIS 151 detector (Gilson Inc., USA), Optima Pak C_{18} column (10 × 250 mm, 5 μm particle size, RS Tech, Seoul, Korea), and a YMC-Triart Phenyl column (250 × 10.0 mm i.d., 5 μm particle size, YMC Co., Ltd.). The conformational search was performed using CONFLEX 7 Rev. (CONFLEX Corporation, Japan); ECD calculations were done by TURBOMOLE 3.4 (COSMOlogic GmbH & Co. KG, Germany). Samples preparation for HRESI MS/MS analyses and separation used sonicator Power Sonic 420 (Hwashin Tech Co., Ltd., Republic of Korea), evaporator EYELA KSB-202 (Eyela Singapore Pte. Ltd., Singapore), centrifuge MICRO17TR (Hanil Scientific Inc., Korea). Biological experiments used autoclave LAC-5080S (DAIHAN LABOTECH Co. Ltd., Korea), clean bench Class II Biological Safety Cabinet (Esco Technologies, Inc., USA), microplate reader VersaMax (Molecular Devices, USA)

5.2. Extraction and isolation

The leaves of *G. inodorum* (5.0 kg) were extracted with 70% EtOH (15 L) three times (1 day each) at room temperature. The crude extract (320 g) was partitioned with *n*-hexane (3 × 10 L), EtOAc (3 × 10 L), *n*-BuOH (3 × 10 L), and H_2O. The *n*-BuOH-soluble partition (100 g) was loaded onto a Daion HP-20 column to obtain seven fractions. Fraction H90-1 (50 g) was chromatographed via an MPLC equipped with an RP-C_{18} column (Revelevis ® Buchi, C-18 40μm, 4 x 20 cm i.d.) eluting with MeOH/H_2O (10–100%, 120 min) as solvents to give nine subfractions. Fraction M7 (5.4 g) was loaded to a Sephadex LH-20 using 100% MeOH for eluting to obtain eight subfractions. Fraction L6 (100 mg) was purified by semipreparative HPLC to isolate compounds **9** (9.0 mg) and **10** (11.0 mg). Fraction L7 (85.0 mg) was further purified by semipreparative HPLC (44% aqueous MeCN, flow rate 3 mL/min) to afford compounds **2** (4.0 mg) and **3** (5.0 mg). Fraction M8 (10.0 g) was also applied to a Sephadex LH-20 with 100% MeOH to yield 7 fractions. Fraction L4 (120.0 mg) was further isolated by a semipreparative HPLC to give compounds **4** (3.5 mg) and **6** (4.0 mg). Fraction L5 (2.5 g) was loaded on MPLC equipped with an RP-C_{18}

column (cm) with 40–100% MeOH to obtain thirteen subfractions. Further purification of fraction M2 (250.0 mg) using semipreparative HPLC (33% aqueous MeCN, flow rate 3 mL/min) yielded compounds **7** (20.0 mg). Fraction M3 (200.0 mg) was subjected to semipreparative HPLC to isolate compounds **16** (62.7 mg). Similarly, the separation by semipreparative HPLC of fraction M4 (80.0 mg) yielded compounds **14** (5.0 mg) and **15** (20.2 mg); fraction M6 (70.2 mg) gave compound **13** (5.0 mg); fraction M7 (60.0 mg) produced compounds **5** (4.0 mg), and **8** (4.5 mg); and fraction M10 (98.0 mg) yielded compound **17** (9.5 mg). Fraction M9 (2.5 g) was the major of compound **1** that was purified via HPLC using 40% MeCN/H_2O. Further mild acid hydrolysis of **1** gave the aglycon (compound **18**, 12.0 mg). Fraction H90-2 (2.0 g) was consequently loaded to an MPLC (30 – 100 % MeOH/H_2O with a column of Watcher ® Flash Cartridge C18 40-63μm, 3 x 14 cm i.d.)), and a Sephadex LH-20 using 100 % MeOH to give subfraction L6 (230.0 mg) that was subsequently separated to give compounds **11** (19.0 mg) and **12** (15.0 mg).
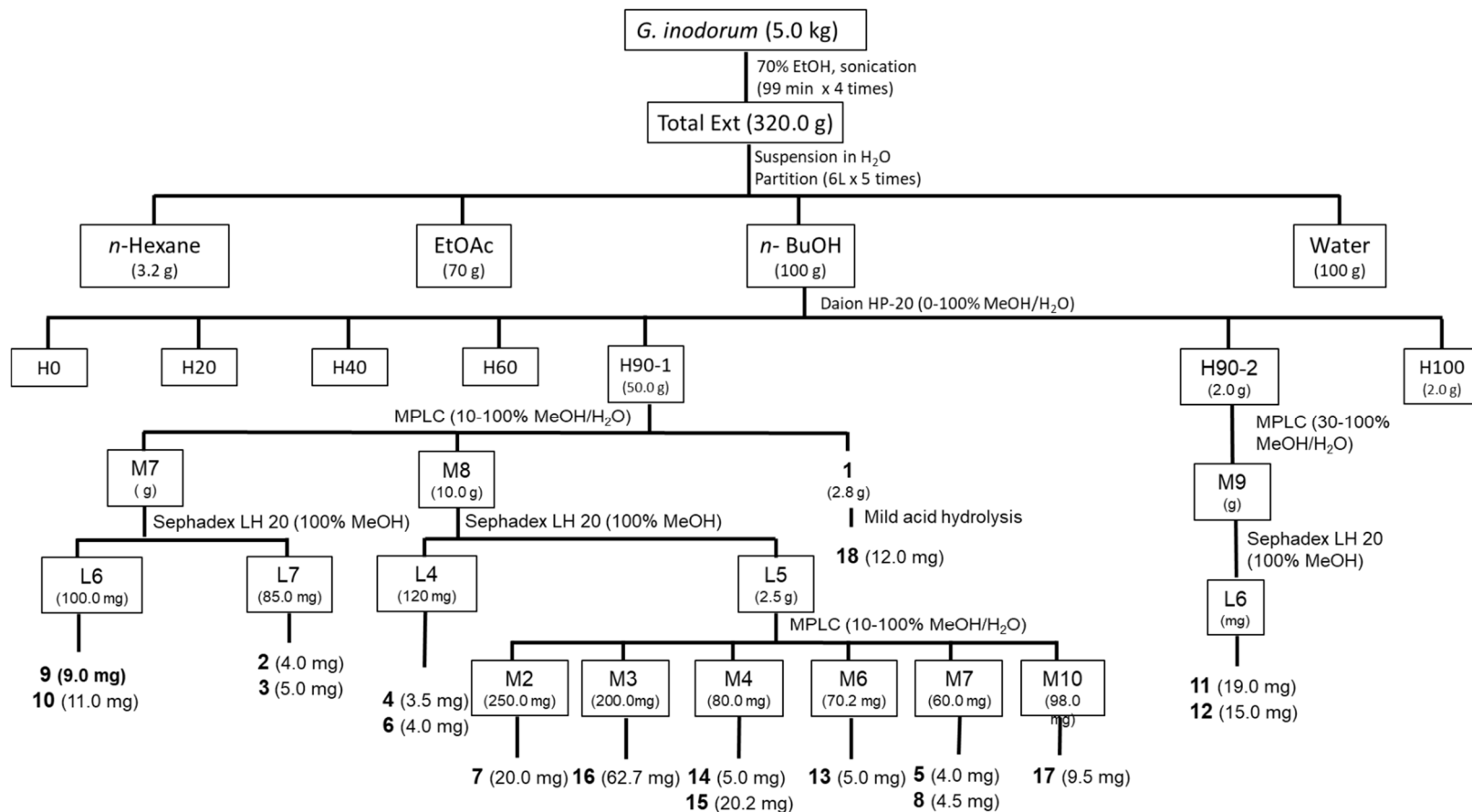
Scheme 4. Isolation scheme of compounds **1–19** from *G. inodorum*

5.3. Physical and chemical characteristics of isolated compounds

5.3.1. Compound **1**

(3β,16β,22α)-22-(N-Methylanthraniloxy)-16,28-dihydroxyolean12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid (**1**)

White amorphous powder

Molecular formula: $C_{50}H_{75}NO_{16}$

HRESIMS/MS *m/z* 944.5023 [M – H]$^-$, (calcd for $C_{50}H_{74}NO_{16}$, 944.5007)

^{1}H and ^{13}C NMR data, see Tables 20, 21

$[\alpha]_D^{25}$ – 6.2 (*c* 0.5, MeOH)

UV (MeOH) λ_{max} (log ε) 223 (1.70), 257 (1.24), 357 (1.10)

IR (KBr) v_{max} 3395, 2925, 1725, 1625 cm^{-1}

5.3.2. Compound **2**

(3β,16β,22α)-22-(N-Methylanthraniloxy)-16,23,28,29-tetrahydroxyolean-12-en-3-yl-O-β-D-glucopyranosiduronic acid (**2**)

White amorphous powder

Molecular formula: $C_{44}H_{65}NO_{13}$

HRESIMS/MS *m/z* 814.4402 [M – H]ˉ , (calcd for $C_{44}H_{64}NO_{13}$, 814.4377)

^{1}H and ^{13}C NMR data, see Tables 20, 21

$[\alpha]_D^{25}$ – 33 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log ε) 223 (1.69), 257 (1.25), 357 (1.10)

IR (KBr) v_{max} 3399, 2926, 1720, 1455, 1083 cm^{-1}

5.3.3. Compound **3**

*(3β,16β,22α)-22-Benzoyloxy-16,23,28-trihydroxyolean-12-en-3-yl-O-β-D-glucopyranosiduronic acid (***3***)*

White amorphous powder

Molecular formula: $C_{43}H_{62}O_{12}$

HRESIMS/MS *m/z* 769.4150 [M – H]$^-$, (calcd for $C_{43}H_{61}O_{12}$, 769.4163)

^{1}H and ^{13}C NMR data, see Tables 20, 21

$[\alpha]_D^{25}$ + 8.2 (c 0.6, MeOH)

UV (MeOH) λ_{max} (log ε) 216 (3.17)

IR (KBr) v_{max} 3403, 2920, 1720, 1620 cm^{-1}

5.3.4. Compound **4**

(3β,16β)-16,23,28-Trihydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid (**4**)

      White amorphous powder

      Molecular formula: $C_{42}H_{68}O_{15}$

      HRESIMS/MS *m/z* 811.4472 [M – H]⁻, (calcd for $C_{42}H_{67}O_{15}$, 811.4480)

      ^{1}H and ^{13}C NMR data, see Tables 20, 21

      $[\alpha]_D^{25}$ + 51.6 (*c* 0.6, MeOH)

      UV (MeOH) λ_{max} (log ε) 217 (3.20)

      IR (KBr) ν_{max} 3389, 2913, 1731, 1615, 1435, 1085 cm⁻¹

5.3.5. Compound **5**

(3β,16β)-16-Hydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosylβ-D-glucopyranosiduronic acid (**5**)

      White amorphous powder

      Molecular formula: $C_{42}H_{68}O_{13}$

      HRESIMS/MS *m/z* 779.4568 [M – H]⁻, (calcd for $C_{42}H_{67}O_{13}$, 779.4581)

      ^{1}H and ^{13}C NMR data, see Tables 20, 21

      $[\alpha]_D^{25}$ + 51.6 (*c* 0.6, MeOH)

      UV (MeOH) λ_{max} (log ε) 217 (2.72)

      IR (KBr) ν_{max} 3410, 2910, 1620 cm⁻¹

5.3.6. Compound **6**

(3β,16β)-16,28,29-Trihydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid (**6**)

      White amorphous powder

      Molecular formula: $C_{42}H_{68}O_{15}$

      HRESIMS/MS *m/z* 811.4470 [M – H]⁻, (calcd for $C_{42}H_{67}O_{15}$, 811.4480)

      ^{1}H and ^{13}C NMR data, see Tables 20, 21

      $[\alpha]_D^{25}$ –34.6 (*c* 0.1, MeOH)

      UV (MeOH) λ_{max} (log ε) 217 (2.78)

      IR (KBr) ν_{max} 3413, 2909, 1615 cm⁻¹

5.3.7. Compound **7**

(3β,16β)-16,29-Dihydroxyolean-12-en-3-yl-2-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid (**7**)

White amorphous powder

Molecular formula: $C_{42}H_{68}O_{14}$

HRESIMS/MS *m/z* 795.4520 [M – H]$^-$, (calcd for $C_{42}H_{67}O_{14}$, 795.4531)

^{1}H and ^{13}C NMR data, see Tables 20, 21

$[\alpha]_D^{25}$ +19.8 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log ε) 217 (2.68)

IR (KBr) ν_{max} 3403, 2911, 1630 cm^{-1}

5.3.8. Compound **8**

(3β,16β,22α)-22-(*N*-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-*β*-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{44}H_{65}NO_{12}$

HRESIMS/MS *m/z* 798.4431 [M–H]$^-$, (calcd for $C_{36}H_{56}O_8$, 798.4429)

$[\alpha]_D^{25}$ –4.1 (*c* 0.1, MeOH) (Shimizu et al., 2001)

5.3.9. Compound **9**

(3β,16β,22α)-22-(*N*-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-3-O-*β*-D-glucopyranosyl-*β*-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{50}H_{75}NO_{17}$

HRESIMS/MS *m/z* 960.4970 [M–H]$^-$, (calcd for $C_{50}H_{75}NO_{17}$, 960.4975)

$[\alpha]_D^{25}$ –7.5 (*c* 0.1, MeOH) (Shimizu et al., 2001)

5.3.10. Compound **10**

(3β,16β,22α)-22-(*N*-methylanthraniloxy)-16,28,29-trihydroxyolean-12-en-3-yl-O-*β*-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{44}H_{67}NO_{11}$

HRESIMS/MS *m/z* 960.4989 [M – H]⁻, (calcd for $C_{44}H_{66}NO_{11}$, 960.4957)

¹H and ¹³C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 21.2 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log *ε*) 223 (1.71), 257 (1.25), 357 (1.11)

IR (KBr) v_{max} 3383, 2927, 2371, 1607, 1514, 1427, 1372, 1244, 1160, 1081, 1043, 1026, 954, 752 cm⁻¹

5.3.11. Compound **11**

Methyl   (3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-O-*α*-D-glucopyranosyl-*β*-D-glucopyranosiduronate

White amorphous powder

Molecular formula: $C_{51}H_{77}NO_{16}$

HRESIMS/MS *m/z* 960.5447 [M + H]⁺, (calcd for $C_{51}H_{76}NO_{16}$, 960.5433)

¹H and ¹³C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 27.2 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log *ε*) 223 (1.70), 257 (1.24), 357 (1.11)

IR (KBr) v_{max} 3376, 2945, 2906, 2376, 2356, 2307, 1743, 1678, 1510, 1426, 1337, 1243, 1163, 107, 1040, 1025, 956, 753 cm⁻¹

5.3.12. Compound **12**

(3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-*β*-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{44}H_{65}NO_{11}$

HRESIMS/MS *m/z* 784.4612 [M + H]⁺, (calcd for $C_{44}H_{65}NO_{11}$, 784.4636)

¹H and ¹³C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 12.7 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log *ε*) 223 (1.72), 257 (1.24), 357 (1.10)

IR (KBr) v_{max} 3365, 2941, 2873, 2356, 2312, 1751, 1672, 1608, 1514, 1427, 1362, 1240, 1160, 1077, 1027, 954, 752 cm⁻¹

5.3.13. Compound **13**

(3*β*,16*β*,22*α*)-22-(*N*-methylanthraniloxy)-16,23,28-trihydroxyolean-12-en-3-yl-*β*-

D-glucopyranoside

White amorphous powder

Molecular formula: $C_{51}H_{77}NO_{16}$

HRESIMS/MS *m/z* 960.4907, (calcd for $C_{51}H_{76}NO_{16}$, 960.4957)

^{1}H and ^{13}C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 10.6 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log ε) 223 (1.71), 257 (1.23), 357 (1.11)

IR (KBr) v_{max} 3379, 2951, 2381, 2356, 2317, 1682, 1608, 1520, 1427, 1240, 1081, 1023, 954, 752 cm^{-1}

5.3.14. Compound **14**

Methyl (3β,16β,21β,22α)-21-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronate

White amorphous powder

Molecular formula: $C_{50}H_{75}NO_{17}$

HRESIMS/MS *m/z* 962.5108 [M + H]$^+$, (calcd for $C_{50}H_{76}NO_{17}$, 962.5113)

^{1}H and ^{13}C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 6.0 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log ε) 223 (1.71), 256 (1.23), 357 (1.10)

IR (KBr) v_{max} 3379, 2946, 2381, 2351, 2312, 1751, 1682, 1608, 1520, 1427, 1240, 1081, 1023, 954 cm^{-1}

5.3.15. Compound **15**

(3β,16β)-16,28-dihydroxyolean-12-en-3-yl-β-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{36}H_{58}O_9$

HRESIMS/MS *m/z* 617.4056 [M – H_2O+ H]$^+$, (calcd for $C_{36}H_{56}O_8$, 617.4053)

5.3.16. Compound **16**

(3β,16β)-16,28-dihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronic acid

White amorphous powder

Molecular formula: $C_{42}H_{68}O_{14}$

HRESIMS/MS *m/z* 795.4514 [M – H]$^-$, (calcd for $C_{42}H_{67}O_{14}$, 795.4531)

5.3.17. Compound **17**

Methyl   (3β,16β,22α)-22-(*N*-methylanthraniloxy)-16,28-dihydroxyolean-12-en-3-yl-3-O-β-D-glucopyranosyl-β-D-glucopyranosiduronate

White amorphous powder

Molecular formula: $C_{51}H_{71}NO_{16}$

HRESIMS/MS *m/z* 960.5349 [M + H]$^+$, (calcd for $C_{51}H_{71}NO_{16}$, 960.5321)

^{1}H and ^{13}C NMR data, see Tables 22, 23

$[\alpha]_D^{25}$ + 30.0 (*c* 0.1, MeOH)

UV (MeOH) λ_{max} (log ε) 223 (1.70), 257 (1.24), 356 (1.11)

IR (KBr) ν_{max} 3379, 2922, 2853, 2366, 2312, 1746, 1682, 1608, 1525, 1460, 1377, 1244, 1165, 1077, 1032, 747 cm^{-1}

Table 20. ^{1}H NMR data of compounds **1–7** from *G. inodorum* in methanol-d_4 (mult., *J* in Hz)

Pos.	1^a	2^c	3^c	4^a	5^c	6^b	7^a
1	1.90, m; 1.88, m	1.66, m; 0.97, m	1.64, m; 0.98, m	1.47, m; 0.94, m	1.65, m; 1.00, m	1.81, m; 1.38, m	1.43, m; 0.85, m
2	1.70, m; 1.84, m	1.89, m; 1.78, m	1.92, m; 1.76, m	2.82, m; 1.78, m	1.90, m; 1.71, m	2.41, m; 1.90, m	2.36, m; 1.92, m
3	3.19, dd (11.6, 4.4)	3.64, m	3.66, dd (11.7, 4.5)	4.26, m	3.18, m	3.33, dd (11.7, 4.3)	3.31, br d (9.2)
4							
5	0.84, m	1.29, m	1.29, m	1.60, m	1.55, m	1.55, m	0.71, m
6	1.46, m; 1.60, m	1.54, m; 1.45, m	1.54, m; 1.43, m	1.76, m; 1.40, m	1.58, m; 1.44, m	1.58, m; 1.44, m	1.49, m; 1.32, m
7	1.62, m; 1.38, m	1.74, m; 1.32, m	1.75, m; 1.34, m	1.67, m; 1.35, m	1.74, m; 1.32, m	1.74, m; 1.32, m	1.51, m; 1.34, m
8							
9	1.59, m	1.64, m	1.65, m	1.69, m	1.55, m	1.55, m	1.57, m
10							
11	1.92, m; 1.90, m	1.96, m; 1.94, m	1.97, m; 1.94, m	1.89, m; 1.88, m	1.91, m; 1.90, m	1.81, m; 1.80, m	1.86, m; 1.85, m
12	5.36, br s	5.38, t-like s	5.37, br s	5.25, t-like s	5.24, t-like s	5.25, s	
13							
14							
15	1.79, d (13.0); 1.38, m	1.78, m; 1.38, dd (11.0, 4.0)	1.79, d (13.0); 1.38, m	2.24, m; 1.70, m	1.72, m; 1.24, m	2.27, m; 1.79, m	2.11, m; 1.63, m
16	4.70, dd (11.6, 5.1)	4.69, dd (11.7, 5.2)	4.73, dd (11.7, 5.0)	4.64, dd (11.0, 4.5)	4.16, dd (11.5, 4.8)	4.78, dd (11.0, 4.0)	4.69, dd (11.0, 4.7)
17							
18	2.59, dd (11.0, 3.0)	2.65, d (12.8)	2.56, d (11.0)	2.40, dd (8.0, 2.0)	2.14, dd (14.0, 4.4)	2.53, dd (13.9, 4.4)	2.46, dd (13.4, 4.1)
19	1.19, m; 1.94, m	2.08, t (14.0); 1.07, m	1.95, m, 1.15, m	1.84, m; 1.15, m	1.73, m; 1.03, m	1.73, m; 1.03, m	2.29, m; 1.40, m
20							
21	1.83, m; 0.9, m	1.92, m; 1.60, m	1.84, m; 1.68, m	1.63, m; 1.30, m	1.61, m; 1.41, m	1.61, m; 1.41, m	2.03, m; 1.39, m
22	5.63, dd (12.2, 3.8)	5.63, dd (12.3, 3.9)	5.64, dd (11.9, 3.8)	2.39, m; 2.04, m	1.88, m; 1.14, m	1.88, m; 1.14, m	2.56, m; 1.46, m
23	1.07, s	3.63, m; 3.30, m	3.63, d (11.7);	4.39, m; 3.78, d	1.09, s	1.28, s	1.28, s
24	0.86, s	0.73, s	0.73, s	1.10, s	0.87, s	1.12, s	1.05, s
25	0.99, s	1.02, s	1.02, s	0.94, s	0.99, s	0.85, s	0.88, s
26	1.04, s	1.05, s	1.05, s	1.04, s	1.02, s	1.03, s	1.17, s
27	1.31, s	1.35, s	1.33, s	1.38, s	1.23, s	1.39, s	1.40, s
28	3.79, m; 3.59, m	3.83, d (11.0); 3.61, m	3.84, d (11.8); 3.58, d (11.7)	4.44, m; 3.68, d (10.6)	0.79, s	4.47, m; 3.72, d (11.0)	1.21, s
29	0.98, s	3.23, m; 3.22, m	1.00, s	0.91, s	0.92, s	3.62, s	3.60, s
30	1.11, s	1.09, s	1.11, s	1.00, s	0.89, s	1.23, s	1.21,s
1'							
2'			8.04, dd (8.0, 1.0)				
3'			7.46, m				
4'	6.77, dd (8.1, 1.1)	6.68, d (8.4)	7.57, m				

Position							
5'	7.39, ddd (7.0, 6.5, 1.4)	7.34, ddd (8.6, 7.1, 1.7)	7.46, m				
6'	6.64, ddd (7.0, 6.4, 1.3)	6.59 ddd (8.4, 7.0, 1.6)					
7'	7.92, dd (8.0, 1.3)	7.89, dd (8.0, 1.6)	8.04, dd (8.0, 1.0)				
8'	2.88, s	2.86, s					
1"	4.44, d (7.8)	4.46, d (7.8)	4.46, d (7.8)	5.21, d (7.1)	4.46, d (7.8)	4.99, d (7.5)	5.01, d (7.5)
2"	3.28, m	3.24, m	3.24, m	3.64, m	3.64, m	4.34, m	4.36, m
3"	3.57, m	3.37, m	3.37, m	3.52, m	3.52, m	4.22, m	4.51, m
4"	3.62, m	3.48, m	3.48, m	3.20, m	3.20, m	4.10, m	4.29, m
5"	3.81, m	3.73, m	3.73, m	3.72, m	3.72, m	4.40, m	4.41, m
6"							
1'''	4.58, d (7.8)			5.46, d (7.8)	4.69, d (7.8)	5.44, d (7.2)	5.43, d (7.4)
2'''	3.44, m			3.52, m	3.52, t (8.6)	4.15, t (8.2)	4.49, m
3'''	3.37, t (7.9)			3.34, m	3.34, m	4.20, m	4.17, m
4'''	3.26, m			3.21, m	3.21, m	4.01, m	4.40, m
5'''	3.31, m			3.25, m	3.25, m	4.05, m	3.96, m
6'''				4.54, m; 4.48, m	3.84, dd (11.9, 2.2), 3.63, m	4.50, m; 4.49, m	4.51, m; 4.50, m

[a] NMR 500 MHz. [b] NMR 600 MHz. [c] NMR 800 MHz

Table 21. ^{13}C NMR data of compounds **1–7** from *G. inodorum*

Pos.	1[a]	2[c]	3[c]	4[a]	5[c]	6[b]	7[a]
1	39.8, CH$_2$	39.6, CH$_2$	39.6, CH$_2$	39.1, CH$_2$	39.9, CH$_2$	39.8, CH$_2$	39.1, CH$_2$
2	26.9, CH$_2$	26.3, CH$_2$	26.3, CH$_2$	26.5, CH$_2$	27.1, CH$_2$	26.9, CH$_2$	26.9, CH$_2$
3	91.0, CH	83.0, CH	82.8, CH	82.3, CH	91.6, CH	89.2, CH	89.3, CH
4	40.2, C	43.9, C	43.9, C	43.9, C	40.4, C	40.3, C	39.9, C
5	56.8, CH	47.9, CH	47.9, CH	48.3, CH	56.1, CH	56.2, CH	56.0, CH
6	19.3, CH$_2$	18.8, CH$_2$	18.8, CH$_2$	18.5, CH$_2$	19.3, CH$_2$	19.3, CH$_2$	18.8, CH
7	33.7, CH$_2$	33.2, CH$_2$	33.2, CH$_2$	32.9, CH$_2$	33.2, CH$_2$	33.2, CH$_2$	33.3, CH$_2$
8	41.2, C	41.2, C	41.2, C	40.5, C	41.2, C	41.2, C	40.6, C
9	48.1, CH	48.2, CH	48.2, CH	47.5, CH	48.2, CH	48.2, CH	47.5, CH
10	37.7, C	37.5, C	37.5, C	36.9, C	37.7, C	37.7, C	37.2, C
11	24.7, CH$_2$	24.7, CH$_2$	24.7, CH$_2$	24.2, CH$_2$	24.7, CH$_2$	24.2, CH$_2$	24.2, CH$_2$
12	125.0, CH	124.9, CH	125.1, CH	122.9, CH	123.5, CH	122.9, CH	122.8, CH
13	142.7, C	142.8, C	142.7, C	144.3, C	145.0, C	144.5, C	145.1, C
14	43.8, C	43.9, C	44.0, C	44.1, C	44.8, C	44.8, C	44.2, C
15	36.9, CH$_2$	36.9, CH$_2$	36.8, CH$_2$	37.2, CH$_2$	36.4, CH$_2$	36.4, CH$_2$	36.9, CH$_2$
16	66.7, CH	66.7, CH	67.2, CH	66.9, CH	66.3, CH	67.2, CH	64.9, CH
17	46.5, C	46.7, C	46.5, C	41.4, C	38.5, C	38.5, C	38.8, C
18	44.8, CH	44.2, CH	44.9, CH	44.8, CH	50.7, CH	44.2, CH	49.4, CH
19	47.0, CH$_2$	41.1, CH$_2$	47.0, CH$_2$	47.3, CH$_2$	47.9, CH$_2$	47.9, CH$_2$	42.2, CH
20	33.0, C	38.3, C	32.9, C	31.4, C	37.7, C	37.7, C	37.1, C
21	39.7, CH$_2$	34.5, CH$_2$	39.7, CH$_2$	34.6, CH$_2$	35.3, CH$_2$	35.3, CH$_2$	29.8, CH$_2$
22	74.4, CH	73.9, CH	75.5, CH	26.4, CH$_2$	31.6, CH	31.6, CH	31.1, CH
23	28.4, CH$_3$	64.6, CH$_2$	64.6, CH$_2$	64.6, CH$_2$	28.4, CH$_3$	28.5, CH$_3$	28.5, CH$_3$
24	16.9, CH$_3$	13.4, CH$_3$	13.4, CH$_3$	13.8, CH$_3$	16.9, CH$_3$	17.2, CH$_3$	17.5, CH$_3$
25	16.1, CH$_3$	16.6, CH$_3$	16.6, CH$_3$	16.6, CH$_3$	16.1, CH$_3$	16.0, CH$_3$	16.0, CH$_3$
26	17.4, CH$_3$	17.4, CH$_3$	17.4, CH$_3$	17.3, CH$_3$	17.5, CH$_3$	17.1, CH$_3$	17.1, CH$_3$
27	27.9, CH$_3$	28.0, CH$_3$	27.8, CH$_3$	27.5, CH$_3$	27.6, CH$_3$	27.5, CH$_3$	27.8, CH$_3$
28	60.9, CH$_2$	60.9, CH$_2$	61.5, CH$_2$	69.2, CH$_3$	22.3, CH$_3$	69.3, CH$_2$	22.9, CH$_2$
29	33.5, CH$_3$	73.7, CH$_2$	33.4, CH$_3$	33.7, CH$_3$	24.4, CH$_3$	74.3, CH$_2$	74.4, CH$_2$
30	25.5, CH$_3$	21.1, CH$_3$	25.3, CH$_3$	24.4, CH$_3$	33.8, CH$_3$	20.5, CH$_3$	20.5, CH$_3$
1'	169.1, C	169.6, C	132.2, CH				
2'	113.3, C	112.1, C	130.6, CH				
3'	151.4, CH	151.4, CH	129.5, CH				
4'	113.1, CH	111.9, CH	134.0, CH				
5'	135.6, CH	135.6, CH	129.5, CH				
6'	117.0, CH	115.3, CH	130.6, CH				
7'	133.1, CH	133.0, CH	167.8, C				
8'	30.7, CH$_3$	29.7, CH$_3$					
1"	106.6, CH	105.8, CH	105.7, CH	104.3, CH	105.8, CH	105.6, CH	105.7, CH
2"	75.5, CH	75.1, CH	75.2, CH	83.9, CH	80.8, CH	83.5, CH	83.4, CH
3"	86.7, CH	77.8, CH	77.8, CH	73.6, CH	77.9, CH	74.3, CH	76.9, CH
4"	71.8, CH	73.3, CH	73.3, CH	78.5, CH	76.3, CH	78.3, CH	78.3, CH
5"	76.3, CH	76.6, CH	76.6, CH	77.2, CH	76.9, CH	77.7, CH	78.2, CH
6"	172.3, C	173.8, C	173.8, C	174.0, C	173.8, C	174.6, C	174.5, C
1'''	105.2, CH			106.5, CH	104.6, CH	106.6, CH	106.5, CH
2'''	74.8, CH			77.3, CH	73.1, CH	73.1, CH	73.9, CH
3'''	77.8, CH			78.3, CH	77.8, CH	77.7, CH	77.6, CH
4'''	71.6, CH			71.7, CH	71.9, CH	71.9, CH	71.9, CH
5'''	78.2, CH			78.8, CH	78.3, CH	78.7, CH	78.7, CH
6'''	62.7, CH$_2$			62.9, CH$_2$	63.1, C	63.0, CH$_2$	63.0, CH$_2$

[a] NMR 125 MHz. [b] NMR 150 MHz. [c] NMR 200 MHz

248

Table 22. ^{1}H NMR data of compounds **10–14, 17** from *G. inodorum* (mult., *J* in Hz)

Pos.	10 [a]	11 [a]	12 [a]	13 [b]	14 [a]	17 [a]
1	1.37, d (5.0) 0.86, m	1.45, m 0.87, m	1.39, m 0.88, m	1.31, m	1.41, s	1.31, m 0.76, m
2	1.58, d (10.5) 1.37, d (5.0)	2.10, m 1.84, m	2.18, m 1.88, d (2.3)	1.78, m 1.52, m	2.21, m 1.80, m	2.01, m 1.74, m
3	3.32, s	3.35, dd (11.4, 3.9)	3.37, dd (11.6, 4.2)	3.63, m	3.38, br s	3.26, dd (11.4, 4.6)
5	0.76, s	0.77, d (7.3)	0.79, d (7.3)	0.87, d (4.6)	0.78, br s	0.70, d (11.9)
6	1.46, s 1.29, s	1.50, m 1.32, m	1.51, d (13.0) 1.33 (br s)	1.45, m 1.30, m	1.50, m 1.31, m	1.42, m 1.24, m
7	2.23, d (12.4) 1.76, d (8.7)	1.60, d (9.8) 1.36, s	1.34, m 1.59, d (2.8)	1.80, m 1.58, m	1.55, m 1.39, m	1.52, m 1.26, m
9	1.58, d (10.5)	1.60, d (9.8)	1.60, m	1.63, m	1.59, m	1.51, m
11	1.76, d (8.7)	1.84, m	1.84, m	1.96, m	1.85, m 1.78, m	1.73, m
12	5.34, s	5.37, t (3.8)	5.36, d (4.0)	5.37, t (3.9)	5.28, br s	5.27, d (6.7)
15	2.23, d (12.4) 1.76, d (8.7)	2.19, d (12.8) 1.74, dd (13.1, 4.5)	2.18 , m 1.74, dd (11.3, 4.8)	1.79, d (12.0) 1.39, d (5.1)	2.08, m 1.62, m	2.07, d (12.4) 1.64, dd (13.0, 5.2)
16	5.21, d (11.9)	5.11, dd (11.2, 4.9)	5.11, dd (11.3, 5.0)	4.70, dd (11.7, 5.1)	5.23, d (10.7)	5.09, dd (11.5, 5.4)
18	3.03, d (9.6)	3.04, dd (14.7, 3.9)	3.03, dd (13.9, 5.8)	2.61, dd (14.9, 3.9)	3.27, d (11.2)	2.87, dd (14.7, 4.5 Hz)
19	2.16, d (14.2) 1.87, d (14.2)	2.10, m 1.26, m	2.10 (1H, br s) 1.23, m	1.16, m 1.96, m	2.28, m 1.41, s	2.01, m 1.88, m
21	2.31, d (11.9) 2.06, d (11.0)	1.99, m	2.00, d (10.7)	1.82, m 1.69, m	6.00, d (10.8)	1.97, m 1.16, m
22	6.46, d (8.8)	6.36, dd (10.5, 5.6)	6.36, dd (10.5, 5.4)	5.62, dd (12.2, 3.8)	5.18, d (10.7)	6.29, br s
23	0.98, s	0.99, s	1.01, s	4.52, d (7.8) 3.46, d (6.4)	1.00, s	0.89, s
24	1.29, s	1.30, s	1.31, s	1.02, s	1.34, s	1.21, s
25	0.79, s	0.82, s	0.81, s	0.72, s	0.82, s	0.75, s
26	0.94, s	0.96, s	0.95, s	1.04, s	0.93, s	1.06, s
27	1.50, s	1.45, s	1.47, s	1.33, s	1.41, s	1.36, s
28	4.48, d (10.5) 4.03, d (10.5)	4.50, d (10.4) 4.08, d (10.4)	4.07, s 4.51, d (10.6)	3.82, d (10.9) 3.60, d (10.9)	4.77, d (10.6) 4.13, d (10.6)	5.26, d (10.5) 4.80, d (10.5)
29	4.13, d (10.1) 4.03, d (10.5)	1.30, s	1.28, s	0.98, s	1.34, s	1.09, s
30	1.29, s	1.01, s	1.01, s	1.11, s	1.14, s	0.96, s
4'	6.66, dd (8.2, 4.6)	6.67, d (8.0)	6.66, dd (8.0, 5.7)	6.68, d (8.5)	6.69, d (8.6)	6.59, d (8.6)
5'	7.38, t (7.8)	7.39, d (7.2)	7.40, ddd (8.7, 7.0, 1.6)	7.35, ddd (8.6, 6.7, 1.6)	7.43, t (7.9)	7.35, dd (9.2, 6.6)

6'	6.66, dd (8.2, 4.6)	6.67, d (8.0)	6.66, dd (8.0, 5.7)	6.55, t (7.5)	6.66, t (7.6)	6.63, t (8.7)
7'	8.46, d (8.2)	8.47, d (7.0)	8.47, dd (8.1, 1.5)	7.88, dd (8.0, 1.5)	8.22, d (6.8)	8.17, dd (8.1, 1.8)
8'	2.77, d (4.6)	2.80, d (4.8)	2.80, s	2.86, s	2.72, s	2.87, dd (14.1, 4.5)
1"	5.47, d (7.6)	4.97, d (7.7)	5.03, d (7.7)	4.59, d (7.8)		4.88, d (7.8)
2"	4.03, d (10.5)	4.14, d (7.8)	4.18, d (8.4)	3.31, m	4.13, d (10.5)	4.78, q (5.3)
3"	4.20, br s	4.32, s	4.37, d (4.9)	3.63, m	4.24, m	4.22, d(8.9)
4"	4.13, d (10.1)	4.39, t (9.0)	4.56, m	3.89, d (10.1)	4.05, d (9.3)	4.31, d (9.1)
5"	4.03, d (10.5)	4.57, t (9.4)	4.69, d (9.6)	3.59, d (10.5)	4.63, m	4.46, d (10.5)
6"	4.48, d (10.5) 4.34, br s			3.82, d (10.9) 3.60, d (10.9)		
7"		3.74, s				3.65, s
1"'		5.37, t (3.8)			5.37, br s	5.29, d (7.8)
2"'		4.08, d (10.4)			4.06, br s	3.98, br s
3"'		4.24, br s			4.25, br s	4.17, br s
4"'		4.24, br s				4.15, d (9.2)
5"'		4.02, br s			4.23, br s	3.96, br s
6"'		4.54, br s 4.33, s			4.55, br s 4.33, br s	4.45, br s 4.24, s
NH-	8.32, s	8.37, q (4.8)			8.12, s	7.90, dd (11.5, 5.4)

[a] Recorded in pyridine – d_5 at 400 MHz. [b] Recorded in methanol – d_4 at 600 MHz.

Table 23. ^{13}C NMR data of compounds **10–14, 17** from *G. inodorum*

Pos.	10 [a]	11 [a]	12 [a]	13 [b]	14 [a]	17 [a]
1	39.2, CH$_2$	39.1, CH$_2$	39.2, CH$_2$	39.6, CH$_2$	39.2, CH$_2$	39.1, CH$_2$
2	33.5, CH$_2$	27.0, CH$_2$	27.0, CH$_2$	25.6, CH$_2$	27.0, CH$_2$	27.0, CH$_2$
3	89.4, CH	89.7, CH	89.7, CH	82.7, CH	89.5, CH	89.7, CH
4	46.3, C	46.7, C	46.7, C	43.9, C	40.0, C	39.9, C
5	56.0, CH	56.0, CH	56.0, CH	56.9, CH	56.0, CH	56.0, CH
6	18.9, CH$_2$	18.8, CH$_2$	19.7, CH$_2$	18.8, CH$_2$	18.8, CH$_2$	18.8, CH$_2$
7	37.5, CH$_2$	33.3, CH$_2$	33.3, CH$_2$	33.2, CH$_2$	33.3, CH$_2$	33.2, C
8	40.7, C	40.7, C	40.7, C	41.2, C	40.7, C	40.8, C
9	47.5, CH	47.5, CH	47.5, CH	47.0, CH	47.5, CH	47.5, CH
10	37.8, C	37.6, C	37.6, C	37.5, C	37.2, C	37.1, C
11	24.2, CH$_2$	24.3, CH$_2$	24.5, CH$_2$	24.7, CH$_2$	24.3, CH$_2$	24.3, CH$_2$
12	123.5, CH	124.5, CH	124.5, CH	124.9, CH	123.1, CH	123.4, CH
13	143.0, C=C	143.0, C=C	143.0, C=C	142.8, C=C	142.8, C	142.5, C=C
14	43.7, C	43.5, C	43.5, C	43.9, C	43.0, C	43.0, C
15	37.2, CH$_2$	37.2, CH$_2$	37.2, CH$_2$	37.0, CH$_2$	36.7, CH$_2$	36.6, CH$_2$
16	66.5, CH	66.0, CH	66.0, CH	66.7, CH	68.5, CH	66.7, CH
17	40.0, C	40.0, C	40.0, C	46.5, C	40.6, C	40.0, C
18	44.2, CH	44.5, CH	44.5, CH	44.8, CH	42.5, CH	44.5, CH
19	41.7, CH$_2$	46.4, CH$_2$	46.4, CH$_2$	48.2, CH$_2$	46.8, CH$_2$	44.9, CH$_2$
20	37.8, C	32.8, C	32.8, C	33.0, C	37.2, C	32.6, C
21	35.6, CH$_2$	39.8, CH$_2$	39.8, CH$_2$	44.9, CH$_2$	80.1, CH	46.5, CH$_2$
22	73.6, CH	74.1, CH	74.1, CH	74.2, CH	71.7, CH	74.6, CH
23	17.5, CH$_3$	17.4, CH$_3$	17.4, CH$_3$	64.7, CH$_2$	17.4, CH$_3$	17.3, CH$_3$
24	28.5, CH$_3$	28.0, CH$_3$	28.0, CH$_3$	13.4, CH$_3$	28.5, CH$_3$	28.1, CH$_3$
25	16.1, CH$_3$	16.1, CH$_3$	16.1, CH$_3$	16.6, CH$_3$	16.1, CH$_3$	16.0, CH$_3$
26	17.5, CH$_3$	17.4, CH$_3$	17.4, CH$_3$	17.4, CH$_3$	17.4, CH$_3$	17.6, CH$_3$
27	27.8, CH$_3$	28.5, CH$_3$	28.5, CH$_3$	28.4, CH$_3$	27.9, CH$_3$	28.5, CH$_3$
28	61.3, CH$_2$	60.3, CH$_2$	60.3, CH$_2$	61.0, CH$_2$	58.6, CH$_2$	62.9, CH$_2$
29	67.7, CH$_2$	26.0, CH$_3$	26.0, CH$_3$	33.5, CH$_3$	20.1, CH$_3$	25.3, CH$_3$
30	28.8, CH$_3$	33.7, CH$_3$	33.7, CH$_3$	26.3, CH$_3$	30.1, CH$_3$	33.8, CH$_3$
1'	168.9, COO-	168.9, COO-	168.9, COO-	169.6, COO-	169.5, COO-	169.4, COO-
2'	112.5, C	112.6, C	112.6, C	111.9, C	111.8, C	111.1, C
3'	152.4, C-NH	152.3, C-NH	152.3, C-NH	153.0, C-NH	153.0, C-NH	153.0, C-NH
4'	111.9, CH	111.8, CH	111.9, CH	112.1, CH	111.6, CH	111.9, CH
5'	135.2, C	135.1, CH	135.1, CH	135.6, CH	135.2, CH	135.5, CH
6'	115.2, CH	115.1, CH	115.1, CH	115.3, CH	115.1, CH	115.2, CH
7'	133.1, CH	133.4, CH	133.4, CH	133.0, CH	132.4, CH	131.8, CH
8'	30.0, CH$_3$	30.1, CH$_3$	30.1, CH$_3$	29.7, CH$_3$	29.9, CH$_3$	29.8, CH$_3$
1"	106.4, CH	107.2, CH	107.2, CH	105.1, CH	106.4, CH	107.2, CH
2"	76.1, CH	74.6, CH	74.7, CH	75.4, CH	74.8, CH	70.2, CH
3"	78.7, CH	87.8, CH	87.8, CH	77.8, CH	89.4, CH	87.8, CH
4"	74.8, CH	71.9, CH	71.9, CH	71.5, CH	78.6, CH	71.9, CH
5"	79.2, CH	77.1, CH	77.2, CH	78.3, CH	72.0,	77.0, CH
6"	62.9, CH	170.7, COO-	170.7, COO-	62.6, CH$_2$	- COO-	170.7, COO-
7"		52.6, -OCH$_3$				52.6, OCH$_3$
1'''		106.4, CH			106.4, CH	106.4, CH
2'''		76.1, CH			76.0, CH	76.1, CH
3'''		78.7, CH			78.7, CH	78.7, CH
4'''		72.1, CH			79.2, C	72.1, CH
5'''		79.2, CH			72.0, CH	79.2, CH
6'''		62.9, CH$_2$			62.9, CH$_2$	62.9, CH$_2$

[a] Recorded in pyridine – d_5 at 100 MHz. [b] Recorded in methanol – d_4 at 150 MHz.

5.4. LC-MS/MS-based molecular networking

HRESI-MS/MS data were acquired from a Waters Acquity UPLC system (Waters Co.) equipped with a Waters Xevo G2 QTOF mass spectrometer (Waters Co., Manchester, UK) and electrospray ionization (ESI) interface, operating in positive-ion mode. A Waters Acquity UPLC BEH C$_{18}$ (150 mm × 2.1 mm, 1.7 μm) was used with a MeCN/H$_2$O gradient system (10:90 to 90 :10) for 14 min. The flow rate was setat 300 μL/min, and the injection volume was 2.0 μL. The temperatures of the autosampler and column oven were maintained at 15 and 40 °C, respectively. Source parameters were set as follows: a capillary voltage of 2.5 kV, cone voltage of 40 V, source temperature of 120 °C, desolvation gas temperature of 350 °C, cone gas flow of 50 L/h, and desolvation gas flow of 800 L/h. The acquisition rate was 0.2 s. Data were centroided during acquisition using an independent reference lock-mass ion via the LockSpray interface to ensure accuracy and precision. Peak picking, chromatogram deconvolution, and other data processing of MS/MS data were performed by MZmine2 software v32 (Pluskal et al., 2010). Eventually, the .mgf preclustered spectral data file and its corresponding .csv metadata file (for RT, areas, and formula integration) were exported using the dedicated "Export for GNPS" and "Export to CSV file" built-in options. A molecular network was created using the online workflow at GNPS (http://gnps. ucsd.edu). The spectra in the network were then searched against GNPS spectral libraries. All matches kept between network spectra and library spectra were required to have a score above 0.65 and at least 4 matched peaks. The molecular networking data were analyzed and visualized using Cytoscape (ver. 3.6.0) (Shannon, 2003). All of the results and parameters can be accessed with the GNPS job id for molecular networking feature-based analysis at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=924954612476435ca4a0c803a7d acd9d

5.5. Acid Hydrolysis

Compounds **10–17** (2.0 mg) was hydrolyzed with 2 mL of 1 N HCl (H$_2$O/ethylene oxide, 1:1) at 100 °C for 4 h. The solution was then dried and suspended in H$_2$O and partitioned with EtOAc. The EtOAc layers were concentrated and the residual H$_2$O

layer was dissolved in pyridine (1.0 mL), and 1.5 mg of L-cysteine methyl ester hydrochloride was added. The mixture was kept for 2 h at 60 °C, and then added 0.2 mL of trimethylsilylimidazole. After incubation for 2 h more at 60 °C, the mixture was dried and partitioned between H_2O (2.0 mL) and *n*-hexane (2.0 mL, x 3 times). The concentrated *n*-hexane layers were analyzed further by UPLC (YMC triart C_{18}, 4 x 250 mm i.d.) under condition as follows: 27% $MeCN/H_2O$ (0.1 % FA) in 60 min at flowrate of 0.6 mL/min. The sugar derivatives showed the retention times at 23.84 and 23.44 min and were identical to the trimethylsilyl-L-cysteine derivatives of authentic D-glucuronic acid and D-glucose, respectively.

5.6. Cell Viability Assay

C2C12 myoblast cells were cultured and seeded in 96-well plates at 3000 cells/well in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS). After 24 h of incubation, the tested compound was dissolved in DMSO and treated for 24 h. Cell viability was analyzed by (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide (MTT) (Sigma, MO, USA). Then, 20 μL of the MTT solution (2 mg/mL) was added to the wells. The plates were incubated at 37 °C in the dark for the MTT reaction. After 2 h of incubation, the supernatant was removed. The formazan crystals were dissolved in 100 μL of DMSO. The absorbance was obtained at 570 nm using a microplate reader (VersaMaxTM, Randor, PA, USA).

5.7. Differentiation of 3T3-L1 Adipocytes

3T3-L1 preadipocyte cells were maintained with DMEM medium (Hyclone, Logan, UT, USA) with 10% calf serum, 100 U/mL penicillin, and 100 mg/mL streptomycin (Hyclone) in 5% CO_2 at 37 °C. After 2 days of incubation, the growth medium was changed to DMEM medium with 10% fetal bovine serum (FBS) (Hyclone) containing 1 μM of dexamethasone (Sigma-Aldrich), 0.52 mM 3-isobutyl-1-methylxanthine (Sigma-Aldrich), and 1 μg/mL insulin (Roche, Germany). After 48 h, the cells were maintained in DMEM medium with 10% FBS, 1 μg/mL insulin, 100 U/mL penicillin, and 100 mg/mL streptomycin for 2 days. The medium was exchanged to DMEM supplemented with 10% FBS medium every 2 days until the induction of adipogenesis.

5.8. Glucose Uptake Assay

A glucose uptake assay was performed with a fluorescent derivative of glucose, 2-NBDG (Invitrogen, Carlsbad, CA, USA). The 3T3-L1 cells were seeded onto a 96-well plate. After the cells were induced to undergo differentiation, the medium was replaced with a glucose-free medium with insulin (100 nM) and test compounds (20 μM) with 50 μM 2-NBDG for 1 h. Then, the cells were washed two times using phosphate-buffered saline (PBS). Cell lysis was performed by treating 70 μL of 1% Triton X-100 (Nacalai Tesque, Kyoto, Japan) in PBS and 0.1 M K3PO4 (Junsei Chemical, Tokyo, Japan) for 10 min. The fluorescence signal was measured on a plate reader (VICTOR X3, PerkinElmer, Waltham, MA, USA) at 450 nm excitation and 535 nm emission.

5.9. Statistical Analysis

All data were calculated as the means $\pm$ SD of three independent experiments. The significant differences between groups were determined using one-way analysis of variance (ANOVA). Statistical significance was determined at $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

Bibliography

Abubakar, A. N. F., Achmadi, S. S., & Suparto, I. H. (2017). Triterpenoid of avocado (*Persea americana*) seed and its cytotoxic activity toward breast MCF-7 and liver HepG2 cancer cells. *Asian Pacific Journal of Tropical Biomedicine*, *7*(5), 397–400.

Aguilar-Arnal, L., Ranjit, S., Stringari, C., Orozco-Solis, R., Gratton, E., & Sassone-Corsi, P. (2016). Spatial dynamics of SIRT1 and the subnuclear distribution of NADH species. *Proceedings of the National Academy of Sciences*, *113*(45), 12715–12720.

Alberti, K. G., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabetic Medicine : A Journal of the British Diabetic Association*, *15*(7), 539–553.

Alvarez, J. A., & Ashraf, A. (2010). Role of Vitamin D in Insulin Secretion and Insulin Sensitivity for Glucose Homeostasis. *International Journal of Endocrinology*, *2010*, 1–18.

An, J.-P., Park, E. J., Ryu, B., Lee, B. W., Cho, H. M., Doan, T. P., Pham, H. T. T., & Oh, W. K. (2020). Oleanane Triterpenoids from the Leaves of *Gymnema inodorum* and Their Insulin Mimetic Activities. *Journal of Natural Products*, *83*(4), 1265–1274.

Anchalee, C., Tewin, T., & Rachana, S. (2010). Effect of *Gymnema inodorum* on postprandial peak plasma glucose levels in healthy human. *African Journal of Biotechnology*, *9*(7), 1079–1085.

Angelini, C., Micaglio, G. F., & Trevisan, C. (1980). Guanidine hydrochloride in infantile and juvenile spinal muscular atrophy. A double blind controlled study. *Acta Neurologica*, *2*(6), 460–465.

Araújo, R. G., Rodriguez-Jasso, R. M., Ruiz, H. A., Pintado, M. M. E., & Aguilar, C. N. (2018). Avocado by-products: Nutritional and functional properties. *Trends in Food Science & Technology*, *80*, 51–60.

Arlene, A. A., Prima, K. A., Utama, L., & Anggraini, S. A. (2015). The Preliminary Study of the Dye Extraction from the Avocado Seed Using Ultrasonic Assisted Extraction. *Procedia Chemistry*, *16*, 334–340.

Ashton, O. B. O., Wong, M., McGhie, T. K., Vather, R., Wang, Y., Requejo-Jackman,

C., Ramankutty, P., & Woolf, A. B. (2006). Pigments in Avocado Tissue and Oil. *Journal of Agricultural and Food Chemistry*, *54*(26), 10151–10158.

Association, A. D. (2014). Standards of Medical Care in Diabetes--2014. *Diabetes Care*, *37*(Supplement_1), S14–S80.

Association, A. D. (2021). Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes—2021. *Diabetes Care*, *44*(Supplement 1), S111–S124.

Aumeeruddy, M. Z., & Mahomoodally, M. F. (2021). Ethnomedicinal plants for the management of diabetes worldwide: A systematic review. *Current Medicinal Chemistry*, *28*.

Baba, Y. F., Gökce, H., Rodi, Y. K., Hayani, S., Chahdi, F. O., Boukir, A., Jasinski, J. P., Kaur, M., Hökelek, T., Sebbar, N. K., & Essassi, E. M. (2020). Syntheses of novel 2-oxo-1,2-dihydroquinoline derivatives: Molecular and crystal structures, spectroscopic characterizations, Hirshfeld surface analyses, molecular docking studies and density functional theory calculations. *Journal of Molecular Structure*, *1217*, 128461.

Barrosa, K., Pinto, E., Tempone, A., Martins, E., & Lago, J. (2014a). Alchornedine, a New Anti-Trypanosomal Guanidine Alkaloid from *Alchornea glandulosa*. *Planta Medica*, *80*(15), 1310–1314.

Basu, S., Yudkin, J. S., Kehlenbrink, S., Davies, J. I., Wild, S. H., Lipska, K. J., Sussman, J. B., & Beran, D. (2019). Estimation of global insulin use for type 2 diabetes, 2018–30: a microsimulation analysis. *The Lancet Diabetes & Endocrinology*, *7*(1), 25–33.

Berasategi, I., Barriuso, B., Ansorena, D., & Astiasarán, I. (2012). Stability of avocado oil during heating: Comparative study to olive oil. *Food Chemistry*, *132*(1), 439–446.

Berlinck, R. G. S. (2020). Ecological Interactions Mediated by Guanidine Secondary Metabolites. In *Comprehensive Natural Products III* (Third Edit, pp. 237–243). Elsevier.

Berlinck, R. G. S., Berlinck, R. G. S., & Berlinck, R. G. S. (1999). Natural guanidine derivatives. *Natural Product Reports*, *16*(3), 339–365.

Berlinck, R. G. S., Bernardi, D. I., Fill, T., Fernandes, A. A. G., & Jurberg, I. D. (2021). The chemistry and biology of guanidine secondary metabolites.

Natural Product Reports, *38*(3), 586–667.

Berlinck, R. G. S., Bertonha, A. F., Takaki, M., & Rodriguez, J. P. G. (2017). The chemistry and biology of guanidine natural products. *Natural Product Reports*, *34*(11), 1264–1301.

Berlinck, R. G. S., & Kossuga, M. H. (2005). Natural guanidine derivatives. *Natural Product Reports*, *22*(4), 516.

Berlinck, R. G. S., Trindade-Silva, A. E., & Santos, M. F. C. (2012). The chemistry and biology of organic guanidine derivatives. *Natural Product Reports*, *29*(12), 1382.

Bertling, I., & Bower, J. P. (2005). Sugars as energy sources – is there a link to avocado fruit quality? In *South African Avocado Growers' Association Yearbook*.

Bharath, L. P., Agrawal, M., McCambridge, G., Nicholas, D. A., Hasturk, H., Liu, J., Jiang, K., Liu, R., Guo, Z., Deeney, J., Apovian, C. M., Snyder-Cappione, J., Hawk, G. S., Fleeman, R. M., Pihl, R. M. F., Thompson, K., Belkina, A. C., Cui, L., Proctor, E. A., … Nikolajczyk, B. S. (2020). Metformin Enhances Autophagy and Normalizes Mitochondrial Function to Alleviate Aging-Associated Inflammation. *Cell Metabolism*, *32*(1), 44-55.e6.

Bodine, S. C., Latres, E., Baumhueter, S., Lai, V. K. M., Nunez, L., Clarke, B. A., Poueymirou, W. T., Panaro, F. J., Na, E., Dharmarajan, K., Pan, Z.-Q., Valenzuela, D. M., DeChiara, T. M., Stitt, T. N., Yancopoulos, G. D., & Glass, D. J. (2001). Identification of Ubiquitin Ligases Required for Skeletal Muscle Atrophy. *Science*, *294*(5547), 1704–1708.

Bommer, C., Sagalova, V., Heesemann, E., Manne-Goehler, J., Atun, R., Bärnighausen, T., Davies, J., & Vollmer, S. (2018). Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030. *Diabetes Care*, *41*(5), 963–970.

Bora, P. S., Narain, N., Rocha, R. V. M., & Queiroz Paulo, M. (2001). Characterization of the oils from the pulp and seeds of avocado (cultivar: Fuerte) fruits. *Grasas y Aceites*.

Borra, M. T., Smith, B. C., & Denu, J. M. (2005). Mechanism of Human SIRT1 Activation by Resveratrol. *Journal of Biological Chemistry*, *280*(17), 17187–17195.

Brooks, C. L., & Gu, W. (2009). How does SIRT1 affect metabolism, senescence and cancer? *Nature Reviews Cancer*, *9*(2), 123–128.

Brown, B. I. (1972). Isolation of unpleasant flavor compounds in the avocado (P*ersea americana*). *Journal of Agricultural and Food Chemistry*, *20*(4), 753–757.

Bruhn, T., Schaumlöffel, A., Hemberger, Y., & Bringmann, G. (2013a). SpecDis: Quantifying the Comparison of Calculated and Experimental Electronic Circular Dichroism Spectra. *Chirality*, *25*(4), 243–249.

Calderón-Oliver, M., Escalona-Buendía, H. B., Medina-Campos, O. N., Pedraza-Chaverri, J., Pedroza-Islas, R., & Ponce-Alquicira, E. (2016). Optimization of the antioxidant and antimicrobial response of the combined effect of nisin and avocado byproducts. *LWT - Food Science and Technology*.

Candido, L. P., Varela, R. M., Torres, A., Molinillo, J. M. G., Gualtieri, S. C. J., & Macías, F. A. (2016). Evaluation of the Allelopathic Potential of Leaf, Stem, and Root Extracts of *Ocotea pulchella* Nees et Mart. *Chemistry & Biodiversity*, *13*(8), 1058–1067.

Chae, B. S., & Shin, T. Y. (2012). Hesperidin ameliorates TNF-α-mediated insulin resistance in differentiated 3T3-L1 cells. *Natural Product Sciences*.

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., … Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920.

Chanderbali, A. S., Albert, V. A., Ashworth, V. E. T. M., Clegg, M. T., Litz, R. E., Soltis, D. E., & Soltis, P. S. (2008). *Persea americana* (avocado): bringing ancient flowers to fruit in the genomics era. *BioEssays*, *30*(4), 386–396.

Chaturvedula, V. S. P., & San Miguel, R. I. (2012). Chemical Constituents from the Polar Fraction of *Rubus suavissimus*. *Organic Chemistry: Current Research*, *1*(1).

Chen, Li, Chen, L., Wan, L., Huo, Y., Huang, J., Li, J., Lu, J., Xin, B., Yang, Q., & Guo, C. (2019). Matrine improves skeletal muscle atrophy by inhibiting E3 ubiquitin ligases and activating the Akt/mTOR/FoxO3α signaling pathway in C2C12 myotubes and mice. *Oncology Reports*.

Chen, Linlin, Xu, W., Yang, Q., Zhang, H., Wan, L., Xin, B., Zhang, J., & Guo, C. (2020). Imperatorin alleviates cancer cachexia and prevents muscle wasting via directly inhibiting STAT3. *Pharmacological Research*, *158*, 104871.

Chen, Linlin, Yang, Q., Zhang, H., Wan, L., Xin, B., Cao, Y., Zhang, J., & Guo, C. (2020). Cryptotanshinone prevents muscle wasting in CT26-induced cancer cachexia through inhibiting STAT3 signaling pathway. *Journal of Ethnopharmacology*.

Chen, X., Wu, Y., Yang, T., Wei, M., Wang, Y., Deng, X., Shen, C., Li, W., Zhang, H., Xu, W., Gou, L., Zeng, Y., Zhang, Y., Wang, Z., & Yang, J. (2016). Salidroside alleviates cachexia symptoms in mouse models of cancer cachexia via activating mTOR signalling. *Journal of Cachexia, Sarcopenia and Muscle*, *7*(2), 225–232.

Chou, T. F., Brown, S. J., Minond, D., Nordin, B. E., Li, K., Jones, A. C., Chase, P., Porubsky, P. R., Stoltz, B. M., Schoenen, F. J., Patricelli, M. P., Hodder, P., Rosen, H., & Deshaies, R. J. (2011). Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proceedings of the National Academy of Sciences of the United States of America*.

Cox, K. A., McGhie, T. K., White, A., & Woolf, A. B. (2004). Skin colour and pigment changes during ripening of "Hass" avocado fruit. *Postharvest Biology and Technology*.

Dabas, D., Shegog, R., Ziegler, G., & Lambert, J. (2013). Avocado (*Persea americana*) Seed as a Source of Bioactive Phytochemicals. *Current Pharmaceutical Design*, *19*(34), 6133–6140.

Di Stefano, V., Avellone, G., Bongiorno, D., Indelicato, S., Massenti, R., & Lo Bianco, R. (2017). Quantitative evaluation of the phenolic profile in fruits of six avocado (*Persea americana*) cultivars by ultra-high-performance liquid chromatography-heated electrospray-mass spectrometry. *International Journal of Food Properties*, *20*(6), 1302–1312.

Do, D. T. B., Bui, T. H., & Phan, D. T. A. (2022). Persea Americana Mill seed extracts: Understanding insights into the antioxidant and antityrosinase activities and effects on preserving qualities of whiteleg shrimp (*Litopenaus vannamei*) during refrigerated storage. *Food Chemistry*, *373*, 131469.

Doan, T.-P., Park, E.-J., Cho, H.-M., Ryu, B., Lee, B.-W., Thuong, P.-T., & Oh, W.-K. (2021). Rugonidines A–F, Diastereomeric 1,6-Dioxa-7,9-diazaspiro[4.5]dec-7-en-8-amines from the Leaves of *Alchornea rugosa*. *Journal of Natural Products*, *84*(12), 3055–3063.

Domergue, F., Helms, G. L., Prusky, D., & Browse, J. (2000). Antifungal compounds from idioblast cells isolated from avocado fruits. *Phytochemistry*, *54*(2), 183–189.

Dreher, M. L., Cheng, F. W., & Ford, N. A. (2021). A Comprehensive Review of Hass Avocado Clinical Trials, Observational Studies, and Biological Mechanisms. *Nutrients*, *13*(12), 4376.

Dreher, M. L., & Davenport, A. J. (2013). Hass Avocado Composition and Potential Health Effects. *Critical Reviews in Food Science and Nutrition*, *53*(7), 738–750.

Duke, J. A., & Martinez, R. V. (1994). Amazonian ethnobotanical dictionary. *Choice Reviews Online*, *32*(03), 32–1525.

Dunkhunthod, B., Talabnin, C., Murphy, M., Thumanu, K., Sittisart, P., & Eumkeb, G. (2021). Gymnema inodorum (Lour.) Decne. Extract Alleviates Oxidative Stress and Inflammatory Mediators Produced by RAW264.7 Macrophages. *Oxidative Medicine and Cellular Longevity*, *2021*, 1–20.

Dyshlovoy, S. A., Hauschild, J., Amann, K., Tabakmakher, K. M., Venz, S., Walther, R., Guzii, A. G., Makarieva, T. N., Shubina, L. K., Fedorov, S. N., Stonik, V. A., Bokemeyer, C., Balabanov, S., Honecker, F., & Amsberg, G. V. (2015). Marine alkaloid Monanchocidin a overcomes drug resistance by induction of autophagy and lysosomal membrane permeabilization. *Oncotarget*, *6*(19), 17328–17341.

Edem, D., Ekanem, I., & Ebong, P. (2009). Effect of aqueous extracts of alligator pear seed (Persea americana mill) on blood glucose and histopathology of pancreas in alloxan-induced diabetic rats. *Pakistan Journal of Pharmaceutical Sciences*, *22*(3), 272–276.

Ekanayaka, E. A. P., Celiz, M. D., & Jones, A. D. (2015). Relative Mass Defect Filtering of Mass Spectra: A Path to Discovery of Plant Specialized Metabolites. *Plant Physiology*, *167*(4), 1221–1232.

Eliakim-Ikechukwu, C. ., & Obri, A. . (2010). Histological changes in the pancreas

following administration of ethanolic extract of *Alchornea cordifolia* leaf in alloxan-induced diabetic wistar rats. *Nigerian Journal of Physiological Sciences*, *24*(2), 153–155.

Figueroa, J. G., Borrás-Linares, I., Lozano-Sánchez, J., & Segura-Carretero, A. (2018). Comprehensive identification of bioactive compounds of avocado peel by liquid chromatography coupled to ultra-high-definition accurate-mass Q-TOF. *Food Chemistry*, *245*, 707–716.

Gala, F., D'Auria, M. V., De Marino, S., Sepe, V., Zollo, F., Smith, C. D., Keller, S. N., & Zampella, A. (2009). Jaspamides M–P: new tryptophan modified jaspamide derivatives from the sponge *Jaspis splendans*. *Tetrahedron*, *65*(1), 51–56.

Gomes, M. B., Rathmann, W., Charbonnel, B., Khunti, K., Kosiborod, M., Nicolucci, A., Pocock, S. J., Shestakova, M. V., Shimomura, I., Tang, F., Watada, H., Chen, H., Cid-Ruzafa, J., Fenici, P., Hammar, N., Surmont, F., & Ji, L. (2019). Treatment of type 2 diabetes mellitus worldwide: Baseline patient characteristics in the global discover study. *Diabetes Research and Clinical Practice*, *151*, 20–32.

Gómez-Huelgas, R., Azriel, S., Puig-Domingo, M., Vidal, J., & Pablos-Velasco, P. de. (2015). Glucagon-like peptide-1 receptor agonists as insulin add-on therapy in patients with inadequate glycemic control in type 2 diabetes mellitus: lixisenatide as a new therapeutic option. *Int. Journal of Clinical Pharmacology and Therapeutics*, *53*(03), 230–240.

Governa, P., Baini, G., Borgonetti, V., Cettolin, G., Giachetti, D., Magnano, A., Miraldi, E., & Biagi, M. (2018). Phytotherapy in the Management of Diabetes: A Review. *Molecules*, *23*(1), 105.

Grimblat, N., Gavín, J. A., Hernández Daranas, A., & Sarotti, A. M. (2019). Combining the Power of J Coupling and DP4 Analysis on Stereochemical Assignments: The J -DP4 Methods. *Organic Letters*, *21*(11), 4003–4007.

Haigis, M. C., & Sinclair, D. A. (2010). Mammalian Sirtuins: Biological Insights and Disease Relevance. *Annual Review of Pathology: Mechanisms of Disease*, *5*(1), 253–295.

Hall, J. A., Dominy, J. E., Lee, Y., & Puigserver, P. (2013). The sirtuin family's role in aging and age-associated pathologies. *Journal of Clinical Investigation*,

123(3), 973–979.

Harding, J. L., Pavkov, M. E., Magliano, D. J., Shaw, J. E., & Gregg, E. W. (2019). Global trends in diabetes complications: a review of current evidence. *Diabetologia*, *62*(1), 3–16.

Hart, N. K., Johns, S. R., & Lamberton, J. A. (1969). Hexahydroimidazo-pyrimidines, a new class of alkaloids from Alchornea javanensis. *Chemical Communication*, *1969*(1484), 137–138.

Hashimura, H., Ueda, C., Kawabata, J., & Kasai, T. (2001). Acetyl-CoA Carboxylase Inhibitors from Avocado (Persea americana Mill) Fruits. *Bioscience, Biotechnology, and Biochemistry*, *65*(7), 1656–1658.

Hatzakis, E., Mazzola, E. P., Shegog, R. M., Ziegler, G. R., & Lambert, J. D. (2019). Perseorangin: A natural pigment from avocado (*Persea americana*) seed. *Food Chemistry*, *293*, 15–22.

Hehre, W., Klunzinger, P., Deppmeier, B., Driessen, A., Uchida, N., Hashimoto, M., Fukushi, E., & Takata, Y. (2019). Efficient Protocol for Accurately Calculating 13 C Chemical Shifts of Conformationally Flexible Natural Products: Scope, Assessment, and Limitations. *Journal of Natural Products*, *82*(8), 2299–2306.

Hurtado-Fernández, E., Carrasco-Pancorbo, A., & Fernández-Gutiérrez, A. (2011). Profiling LC-DAD-ESI-TOF MS Method for the Determination of Phenolic Metabolites from Avocado (*Persea americana*). *Journal of Agricultural and Food Chemistry*, *59*(6), 2255–2267.

Ibarra-Laclette, E., Méndez-Bravo, A., Pérez-Torres, C. A., Albert, V. A., Mockaitis, K., Kilaru, A., López-Gómez, R., Cervantes-Luevano, J. I., & Herrera-Estrella, L. (2015). Deep sequencing of the Mexican avocado transcriptome, an ancient angiosperm with a high content of fatty acids. *BMC Genomics*, *16*(1), 599.

Jang, S., Kang, H. T., & Hwang, E. S. (2012). Nicotinamide-induced Mitophagy. *Journal of Biological Chemistry*, *287*(23), 19304–19314.

Kang, K. Bin, Kim, H. W., Kim, J. W., Oh, W. K., Kim, J., & Sung, S. H. (2017). Catechin-Bound Ceanothane-Type Triterpenoid Derivatives from the Roots of *Zizyphus jujuba*. *Journal of Natural Products*, *80*(4), 1048–1054.

Kawagishi, H., Fukumoto, Y., Hatakeyama, M., He, P., Arimoto, H., Matsuzawa, T., Arimoto, Y., Suganuma, H., Inakuma, T., & Sugiyama, K. (2001). Liver Injury Suppressing Compounds from Avocado (*Persea americana*). *Journal of*

Agricultural and Food Chemistry, *49*(5), 2215–2221.

Khuong-Huu, F., Le Forestier, J.-P., & Goutarel, R. (1972). Alchornéine, isoalchornéine et alchornéinone, produits isolés de l'*Alchornea floribunda* Muell. Arg. *Tetrahedron*, *28*(20), 5207–5220.

Kim, H. W., Park, J., Kang, K. Bin, Kim, T. B., Oh, W. K., Kim, J., & Sung, S. H. (2016). Acylphloroglucinolated Catechin and Phenylethyl Isocoumarin Derivatives from *Agrimonia pilosa*. *Journal of Natural Products*, *79*(9), 2376–2383.

Kobayashi, K., Hashimoto, H., Kanbe, Y., & Konishi, H. (2011). One-pot synthesis of 4-substituted 4-alkoxy-1,4-dihydro-3,1-benzoxazine-2-thiones by the reaction of 2-isothiocyanatobenzoates with organolithiums. *Tetrahedron*, *67*(25), 4535–4538.

Kodali, M., Attaluri, S., Madhu, L. N., Shuai, B., Upadhya, R., Gonzalez, J. J., Rao, X., & Shetty, A. K. (2021). Metformin treatment in late middle age improves cognitive function with alleviation of microglial activation and enhancement of autophagy in the hippocampus. *Aging Cell*, *20*(2).

Kohyama, N., & Ono, H. (2013). Hordatine A β-d-Glucopyranoside from Ungerminated Barley Grains. *Journal of Agricultural and Food Chemistry*, *61*(5), 1112–1116.

Korver, O., & Wilkins, C. K. (1971). Circular dichroism spectra of flavanols. *Tetrahedron*, *27*(22), 5459–5465.

Kosińska, A., Karamać, M., Estrella, I., Hernández, T., Bartolomé, B., & Dykes, G. A. (2012). Phenolic Compound Profiles and Antioxidant Capacity of *Persea americana* Mill. Peels and Seeds of Two Varieties. *Journal of Agricultural and Food Chemistry*, *60*(18), 4613–4619.

Lamikanra, A., Ogundaini, A. O., & Ogungbamila, F. O. (1990). Antibacterial constituents of *Alchornea cordifolia* leaves. *Phytotherapy Research*, *4*(5), 198–200.

Landahl, S., Meyer, M. D., & Terry, L. A. (2009). Spatial and Temporal Analysis of Textural and Biochemical Changes of Imported Avocado cv. Hass during Fruit Ripening. *Journal of Agricultural and Food Chemistry*, *57*(15), 7039–7047.

Larin, E. M., Torelli, A., Loup, J., & Lautens, M. (2021). One-Pot, Three-Step Synthesis of Benzoxazinones via Use of the Bpin Group as a Masked

Nucleophile. *Organic Letters*, *23*(7), 2720–2725.

Lee, S.-G., Yu, M.-H., Lee, S.-P., & Lee, I.-S. (2008). Antioxidant Activities and Induction of Apoptosis by Methanol Extracts from Avocado. *Journal of the Korean Society of Food Science and Nutrition*, *37*(3), 269–275.

Leibiger, I. B., & Berggren, P.-O. (2006). Sirt1: a metabolic master switch that modulates lifespan. *Nature Medicine*, *12*(1), 34–36.

Leite, J. J. G., Brito, É. H. S., Cordeiro, R. A., Brilhante, R. S. N., Sidrim, J. J. C., Bertini, L. M., Morais, S. M. de, & Rocha, M. F. G. (2009). Chemical composition, toxicity and larvicidal and antifungal activities of *Persea americana* (avocado) seed extracts. *Revista Da Sociedade Brasileira de Medicina Tropical*, *42*(2), 110–113.

Li, H., Jiang, J., Liu, Z., Lin, S., Xia, G., Xia, X., Ding, B., He, L., Lu, Y., & She, Z. (2014). Peniphenones A–D from the Mangrove Fungus *Penicillium dipodomyicola* HN4-3A as Inhibitors of Mycobacterium tuberculosis Phosphatase MptpB. *Journal of Natural Products*, *77*(4), 800–806.

Liu, X., Robinson, P. W., Madore, M. A., Witney, G. W., & Arpaia, M. L. (1999). `Hass' Avocado Carbohydrate Fluctuations. II. Fruit Growth and Ripening. *Journal of the American Society for Horticultural Science*, *124*(6), 676–681.

Liu, X., Sievert, J., Arpaia, M. L., & Madore, M. A. (2002). Postulated Physiological Roles of the Seven-carbon Sugars, Mannoheptulose, and Perseitol in Avocado. *Journal of the American Society for Horticultural Science*, *127*(1), 108–114.

Liu, Y., Wang, X., Chen, M., Lin, S., Li, L., & Shi, J. (2016). Three pairs of alkaloid enantiomers from the root of *Isatis indigotica. Acta Pharmaceutica Sinica B*, *6*(2), 141–147.

Lu, Q.-Y., Arteaga, J. R., Zhang, Q., Huerta, S., Go, V. L. W., & Heber, D. (2005). Inhibition of prostate cancer cell growth by an avocado extract: role of lipid-soluble bioactive substances. *The Journal of Nutritional Biochemistry*, *16*(1), 23–30.

Mao, F., Zhang, L., Cai, M.-H., Guo, H., & Yuan, H.-H. (2015). Leonurine hydrochloride induces apoptosis of H292 lung cancer cell by a mitochondria-dependent pathway. *Pharmaceutical Biology*, *53*(11), 1684–1690.

Martínez, C. A., Mosquera, O. M., & Niño, J. (2017). Medicinal plants from the genus Alchornea (Euphorbiaceae): A review of their ethnopharmacology uses

and phytochemistry. *Boletín Latinoamericano y Del Caribe de Plantas Medicinales y Aromáticas*, *16*(3), 162–205.

Mavar-Manga, H., Haddad, M., Pieters, L., Baccelli, C., Penge, A., & Quetin-Leclercq, J. (2008). Anti-inflammatory compounds from leaves and root bark of *Alchornea cordifolia* (Schumach. & Thonn.) Müll. Arg. *Journal of Ethnopharmacology*, *115*(1), 25–29.

Melgar, B., Dias, M. I., Ciric, A., Sokovic, M., Garcia-Castello, E. M., Rodriguez-Lopez, A. D., Barros, L., & Ferreira, I. C. R. F. (2018). Bioactive characterization of *Persea americana* Mill. by-products: A rich source of inherent antioxidants. *Industrial Crops and Products*, *111*, 212–218.

Minot, A. S., Dood, K., & Riven, S. S. (1939). Use Of Guanidine Hydrochloride In Treatment Of Myasthenia Gravis. *Journal of the American Medical Association*, *113*(7), 553.

Mizushima, N., & Komatsu, M. (2011). Autophagy: Renovation of Cells and Tissues. *Cell*, *147*(4), 728–741.

Mohammed, R. K., Ibrahim, S., Atawodi, S. E., Eze, E. D., Suleiman, J. B., Ugwu, M. N., & Malgwi, I. S. (2012). Anti-diabetic and haematological effects of n-butanol fraction of alchornea cordifolia leaf extract in streptozotocin-induced diabetic wistar rats. *Scientific Journal of Biological Sciences*, *1*(1), 14–21.

Momken, I., Stevens, L., Bergouignan, A., Desplanches, D., Rudwill, F., Chery, I., Zahariev, A., Zahn, S., Stein, T. P., Sebedio, J. L., Pujos-Guillot, E., Falempin, M., Simon, C., Coxam, V., Andrianjafiniony, T., Gauquelin-Koch, G., Picquet, F., & Blanc, S. (2011). Resveratrol prevents the wasting disorders of mechanical unloading by acting as a physical exercise mimetic in the rat. *The FASEB Journal*, *25*(10), 3646–3660.

Monache, G. D., Volpe, A. R., Monache, F. D., Vitali, A., Botta, B., Espinal, R., De Bonnevaux, S. C., Luca, C. De, Botta, M., Corelli, F., & Carmignani, M. (1999). Further hypotensive metabolites from verbesina caracasana. *Bioorganic & Medicinal Chemistry Letters*, *9*(22), 3249–3254.

Nankar, R. P., & Doble, M. (2013). Non-peptidyl insulin mimetics as a potential antidiabetic agent. *Drug Discovery Today*, *18*(15–16), 748–755.

Nielsen, M. F., Roelsgaard, K., Keiding, S., Brodersen, K., Møller, N., Vyberg, M., & Vilstrup, H. (2015). Impaired hepatic counterregulatory response to insulin-

induced hypoglycemia in hepatic denervated pigs. *Journal of Clinical & Translational Endocrinology*, *2*(4), 131–136.

Norris, F. H., Calanchini, P. R., Fallat, R. J., Panchari, S., & Jewett, B. (1974). The administration of guanidine in amyotrophic lateral sclerosis. *Neurology*, *24*(8), 721–721.

Oberlies, N. H., Rogers, L. L., Martin, J. M., & McLaughlin, J. L. (1998). Cytotoxic and Insecticidal Constituents of the Unripe Fruit of *Persea americana*. *Journal of Natural Products*, *61*(6), 781–785.

Oboh, G., Odubanjo, V. O., Bello, F., Ademosun, A. O., Oyeleye, S. I., Nwanna, E. E., & Ademiluyi, A. O. (2016). Aqueous extracts of avocado pear (*Persea americana* Mill.) leaves and seeds exhibit anti-cholinesterases and antioxidant activities in vitro. *Journal of Basic and Clinical Physiology and Pharmacology*, *27*(2), 131–140.

Otaigbe, J. O. E., Oriji, O. G., & Ekerenam, G. E. (2016). Studies on the Paint Forming Properties of Avocado (*Persea Americana*) and African Pear (Dacryodes Edulis) Seed Oils. In *International Journal of Engineering Research and Applications* (pp. 8–15).

Owolabi, M. A., Coker, H. A. B., & Jaja, S. I. (2010). Bioactivity of the phytoconstituents of the leaves of *Persea americana*. *Journal of Medicinal Plants Research*, *4*(12), 1130–1135.

Ozdemir, F., & Topuz, A. (2004). Changes in dry matter, oil content and fatty acids composition of avocado during harvesting time and post-harvesting ripening period. *Food Chemistry*, *86*(1), 79–83.

Ozolua, R., Anaka, O., Okpo, S., & Idogun, S. (2010). Acute and sub-acute toxicological assessment of the aqueous seed extract of *Persea americana* Mill (Lauraceae) in rats. *African Journal of Traditional, Complementary and Alternative Medicines*, *6*(4), 573–578.

Pahua-Ramos, M. E., Ortiz-Moreno, A., Chamorro-Cevallos, G., Hernández-Navarro, M. D., Garduño-Siciliano, L., Necoechea-Mondragón, H., & Hernández-Ortega, M. (2012). Hypolipidemic Effect of Avocado (*Persea americana* Mill) Seed in a Hypercholesterolemic Mouse Model. *Plant Foods for Human Nutrition*, *67*(1), 10–16.

Perla, V., & Jayanty, S. S. (2013). Biguanide related compounds in traditional

antidiabetic functional foods. *Food Chemistry*, *138*(2–3), 1574–1580.

Perry, D. M., Byrnes, N. K., Heymann, H., & Hayes, J. E. (2019). Rejection of labrusca-type aromas in wine differs by wine expertise and geographic region. *Food Quality and Preference*, *74*, 147–154.

Pham, H.-T.-T., Kim, H.-W., Han, S., Ryu, B., Doan, T.-P., An, J.-P., Tran, V.-O., & Oh, W.-K. (2019). Development of a Building Block Strategy to Target the Classification, Identification, and Metabolite Profiling of Oleanane Triterpenoids in *Gymnema sylvestre* Using UHPLC-qTOF/MS. *Journal of Natural Products*, *82*(12), 3249–3266.

Pham, H. T. T., Ha, T. K. Q., Cho, H. M., Lee, B. W., An, J. P., Tran, V. O., & Oh, W. K. (2018). Insulin Mimetic Activity of 3,4- Seco and Hexanordammarane Triterpenoids Isolated from *Gynostemma longipes*. *Journal of Natural Products*, *81*(11), 2470–2482.

Pham, H. T. T., Hoang, M. C., Ha, T. K. Q., Dang, L. H., Tran, V. O., Nguyen, T. B. T., Lee, C. H., & Oh, W. K. (2018). Discrimination of different geographic varieties of *Gymnema sylvestre* , an anti-sweet plant used for the treatment of type 2 diabetes. *Phytochemistry*, *150*, 12–22.

Pieterse, Z., Jerling, J. C., Oosthuizen, W., Kruger, H. S., Hanekom, S. M., Smuts, C. M., & Schutte, A. E. (2005). Substitution of high monounsaturated fatty acid avocado for mixed dietary fats during an energy-restricted diet: Effects on weight loss, serum lipids, fibrinogen, and vascular function. *Nutrition*, *21*(1), 67–75.

Plaza, L., Sánchez-Moreno, C., de Pascual-Teresa, S., de Ancos, B., & Cano, M. P. (2009). Fatty Acids, Sterols, and Antioxidant Activity in Minimally Processed Avocados during Refrigerated Storage. *Journal of Agricultural and Food Chemistry*, *57*(8), 3204–3209.

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, *11*(1), 395.

Pretsch, E., Bühlmann, P., & Badertscher, M. (2009). Structure Determination of Organic Compounds. In *Structure Determination of Organic Compounds: Tables of Spectral Data*. Springer-Verlag.

Quanjun, Y., Lili, W., Zhiyong, Z., Yan, L., Qi, Y., Liya, L., Bin, L., & Cheng, G.

(2013). Parthenolide from *Parthenium integrifolium* reduces tumor burden and alleviate cachexia symptoms in the murine CT-26 model of colorectal carcinoma. *Phytomedicine*, *20*(11), 992–998.

R. Suárez-Castillo, O., I. Bautista-Hernández, C., Sánchez-Zavala, M., Meléndez-Rodríguez, M., Sierra-Zenteno, A., S. Morales-Ríos, M., & Joseph-Nathan, P. (2012). Microwave-Assisted Synthesis of 3,1-Benzoxazin-2-ones from 3-Hydroxyoxindoles. *Heterocycles*, *85*(9), 2147.

Rahman, S., & Islam, R. (2011). Mammalian Sirt1: insights on its biological functions. *Cell Communication and Signaling*, *9*(1), 11.

Raymond Chia, T. W., & Dykes, G. A. (2010). Antimicrobial activity of crude epicarp and seed extracts from mature avocado fruit (Persea americana) of three cultivars. *Pharmaceutical Biology*, *48*(7), 753–756.

Rodríguez-Carpena, J.-G., Morcuende, D., Andrade, M.-J., Kylli, P., & Estévez, M. (2011). Avocado (Persea americana Mill.) Phenolics, In Vitro Antioxidant and Antimicrobial Activities, and Inhibition of Lipid and Protein Oxidation in Porcine Patties. *Journal of Agricultural and Food Chemistry*, *59*(10), 5625–5635.

Rodríguez-Sánchez, D. G., Pacheco, A., García-Cruz, M. I., Gutiérrez-Uribe, J. A., Benavides-Lozano, J. A., & Hernández-Brenes, C. (2013). Isolation and Structure Elucidation of Avocado Seed (Persea americana) Lipid Derivatives That Inhibit Clostridium sporogenes Endospore Germination. *Journal of Agricultural and Food Chemistry*, *61*(30), 7403–7411.

Rodríguez-Sánchez, D., Silva-Platas, C., Rojo, R. P., García, N., Cisneros-Zevallos, L., García-Rivas, G., & Hernández-Brenes, C. (2013). Activity-guided identification of acetogenins as novel lipophilic antioxidants present in avocado pulp (*Persea americana*). *Journal of Chromatography B*, *942–943*, 37–45.

Roemmelt, S., Zimmermann, N., Rademacher, W., & Treutter, D. (2003). Formation of novel flavonoids in apple (*Malus* × *domestica*) treated with the 2-oxoglutarate-dependent dioxygenase inhibitor prohexadione-Ca. *Phytochemistry*, *64*, 709–716.

Rosenblat, G., Meretski, S., Segal, J., Tarshis, M., Schroeder, A., Zanin-Zhorov, A., Lion, G., Ingber, A., & Hochberg, M. (2011). Polyhydroxylated fatty alcohols

derived from avocado suppress inflammatory response and provide non-sunscreen protection against UV-induced damage in skin cells. *Archives of Dermatological Research*, *303*(4), 239–246.

Roslund, M. U., Tähtinen, P., Niemitz, M., & Sjöholm, R. (2008). Complete assignments of the 1H and 13C chemical shifts and JH,H coupling constants in NMR spectra of d-glucopyranose and all d-glucopyranosyl-d-glucopyranosides. *Carbohydrate Research*, *343*(1), 101–112.

Rowan, D. D., Dymock, J. J., & Brimble, M. A. (1990). Effect of fungal metabolite peramine and analogs on feeding and development of argentine stem weevil (*Listronotus bonariensis*). *Journal of Chemical Ecology*, *16*(5), 1683–1695.

Rubner, M., & Intelmann, D. (2014). The use of mass defect plots in phytochemical analysis. *Planta Medica*, *80*(16).

Ryu, B., Cho, H. M., Zhang, M., Lee, B. W., Doan, T. P., Park, E. J., Lee, H. J., & Oh, W. K. (2021). Meroterpenoids from the leaves of *Psidium guajava* (guava) cultivated in Korea using MS/MS-based molecular networking. *Phytochemistry*, *186*, 112723.

Saavedra, J., Córdova, A., Navarro, R., Díaz-Calderón, P., Fuentealba, C., Astudillo-Castro, C., Toledo, L., Enrione, J., & Galvez, L. (2017). Industrial avocado waste: Functional compounds preservation by convective drying process. *Journal of Food Engineering*, *198*, 81–90.

Sartori, R., Romanello, V., & Sandri, M. (2021). Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. *Nature Communications*, *12*(1), 330.

Schmidt, C. A., Murillo, R., Bruhn, T., Bringmann, G., Goettert, M., Heinzmann, B., Brecht, V., Laufer, S. A., & Merfort, I. (2010). Catechin Derivatives from *Parapiptadenia rigida* with in Vitro Wound-Healing Properties. *Journal of Natural Products*, *73*(12), 2035–2041.

Sciacca, L., Cassarino, M. F., Genua, M., Pandini, G., Le Moli, R., Squatrito, S., & Vigneri, R. (2010). Insulin analogues differently activate insulin receptor isoforms and post-receptor signalling. *Diabetologia*, *53*(8), 1743–1753.

Segovia, F. J., Corral-Pérez, J. J., & Almajano, M. P. (2016). Avocado seed: Modeling extraction of bioactive compounds. *Industrial Crops and Products*, *85*, 213–220.

Shannon, P. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504.

Shimizu, K., Ozeki, M., Iino, A., Nakajyo, S., Urakawa, N., & Atsuchi, M. (2001). Structure-Activity Relationships of Triterpenoid Derivatives Extracted From *Gymnema inodorum* Leaves on Glucose Absorption. *Japanese Journal of Pharmacology*, *86*(2), 223–229.

Shin, B., Ahn, S., Noh, M., Shin, J., & Oh, D.-C. (2015). Suncheonosides A–D, Benzothioate Glycosides from a Marine-Derived *Streptomyces* sp. *Journal of Natural Products*, *78*(6), 1390–1396.

Slade, D., Ferreira, D., & Marais, J. P. J. (2005). Circular dichroism, a powerful tool for the assessment of absolute configuration of flavonoids. *Phytochemistry*, *66*(18), 2177–2215.

Sleno, L. (2012). The use of mass defect in modern mass spectrometry. *Journal of Mass Spectrometry*, *47*(2), 226–236.

Sohn, E. J., & Park, H. T. (2017). Natural agents mediated autophagic signal networks in cancer. *Cancer Cell International*, *17*(1), 110.

Song, X., Wang, T., Zhang, Z., Jiang, H., Wang, W., Cao, Y., & Zhang, N. (2015). Leonurine Exerts Anti-Inflammatory Effect by Regulating Inflammatory Signaling Pathways and Cytokines in LPS-Induced Mouse Mastitis. *Inflammation*, *38*(1), 79–88.

Srinuanchai, W., Nooin, R., Pitchakarn, P., Karinchai, J., Suttisansanee, U., Chansriniyom, C., Jarussophon, S., Temviriyanukul, P., & Nuchuchua, O. (2021). Inhibitory effects of *Gymnema inodorum* (Lour.) Decne leaf extracts and its triterpene saponin on carbohydrate digestion and intestinal glucose absorption. *Journal of Ethnopharmacology*.

Tanaka, O. (1985). Application of 13C-nuclear magnetic resonance spectrometry to structural studies on glycosides: saponins of Panax spp. and natrual sweet glycosides. *Chemical Pharmaceutical Bulletin*, *105*(4), 323–351.

Tanaka, T., Nakashima, T., Ueda, T., Tomii, K., & Kouno, I. (2007). Facile

Discrimination of Aldose Enantiomers by Reversed-Phase HPLC. *Chemical and Pharmaceutical Bulletin*, *55*(6), 899–901.

Tang, S.-Q., Leloire, M., Schneider, S., Mohr, J., Bricard, J., Gizzi, P., Garnier, D., Schmitt, M., & Bihel, F. (2020). Diastereoselective Synthesis of Nonplanar 3-Amino-1,2,4-oxadiazine Scaffold: Structure Revision of Alchornedine. *The Journal of Organic Chemistry*, *85*(23), 15347–15359.

Tapondjou, L. A., Kristina, J., & Siems, K. (2016). Alchornealaxine , an Unusual Prenylguanidinyl-epicatechin Derivative from Alchornealaxine , an Unusual Prenylguanidinyl-epicatechin Derivative from *Alchornea laxiflora* (Benth) Pax and Hoffman. *Rec. Nat. Prod.*, *10*(4), 508–512.

Thompson, R. D., & Quaife, J. T. (2001). Liquid Chromatographic Determination of Methyl Anthranilate in Artificially Flavored Nonalcoholic Beverages. *Journal of AOAC International*, *84*(2), 493–497.

Trang, D. T., Yen, D. T. H., Cuong, N. T., Anh, L. T., Hoai, N. T., Tai, B. H., Doan, V. Van, Yen, P. H., Quang, T. H., Nhiem, N. X., Minh, C. Van, & Kiem, P. Van. (2021). Pregnane glycosides from *Gymnema inodorum* and their α-glucosidase inhibitory activity. *Natural Product Research*, *35*(13), 2157–2163.

Vakifahmetoglu-Norberg, H., Xia, H., & Yuan, J. (2015). Pharmacologic agents targeting autophagy. *Journal of Clinical Investigation*, *125*(1), 5–13.

Villa-Rodríguez, J. A., Molina-Corral, F. J., Ayala-Zavala, J. F., Olivas, G. I., & González-Aguilar, G. A. (2011). Effect of maturity stage on the content of fatty acids and antioxidant activity of 'Hass' avocado. *Food Research International*, *44*(5), 1231–1237.

Wang, D.-T., Yin, Y., Yang, Y.-J., Lv, P.-J., Shi, Y., Lu, L., & Wei, L.-B. (2014a). Resveratrol prevents TNF-α-induced muscle atrophy via regulation of Akt/mTOR/FoxO1 signaling in C2C12 myotubes. *International Immunopharmacology*, *19*(2), 206–213.

Wang, D., Li, G., Feng, Y., & Xu, S. (2008). Two new oleanane triterpene glycosides from *Gymnema inodorum*. *Journal of Chemical Research*, *2008*(11), 655–657.

Wang, H., Liu, T., Li, L., Wang, Q., Yu, C., Liu, X., & Li, W. (2015). Tetrandrine is a potent cell autophagy agonist via activated intracellular reactive oxygen species. *Cell & Bioscience*, *5*(1), 4.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen,

D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., … Bandeira, N. (2016a). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., … Bandeira, N. (2016b). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837.

Wang, W., Bostic, T. R., & Gu, L. (2010). Antioxidant capacities, procyanidins and pigments in avocados of different strains and cultivars. *Food Chemistry*, *122*(4), 1193–1198.

Watrous, J., Roach, P., Alexandrov, T., Heath, B. S., Yang, J. Y., Kersten, R. D., van der Voort, M., Pogliano, K., Gross, H., Raaijmakers, J. M., Moore, B. S., Laskin, J., Bandeira, N., & Dorrestein, P. C. (2012). Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences*, *109*(26), E1743–E1752.

WHO. (2007). WHO Monographs on Selected Medicinal Plants. *World Health*.

Wiart, C. (2007). Ethnopharmacology of Medicinal Plants: Asia and the Pacific. In *Angewandte Chemie International Edition, 6(11), 951–952*. Springer Science & Business Media.

World Health Organisation. (1999). *WHO Monographs on selected medicianl plants*. World Health Organisation.

World health organisation (WHO). (2020). *Fact Sheet. Detail. Diabetes*. WHO.

World Health Organization. (2009). WHO monographs on Selected medicinal plants. In *World Health organization*.

Xie, Q., Chen, Y., Tan, H., Liu, B., Zheng, L. L., & Mu, Y. (2021). Targeting Autophagy with Natural Compounds in Cancer: A Renewed Perspective from Molecular Mechanisms to Targeted Therapy. In *Frontiers in Pharmacology*.

Xu, C., Wang, L., Fozouni, P., Evjen, G., Chandra, V., Jiang, J., Lu, C., Nicastri, M., Bretz, C., Winkler, J. D., Amaravadi, R., Garcia, B. A., Adams, P. D., Ott, M.,

Tong, W., Johansen, T., Dou, Z., & Berger, S. L. (2020). SIRT1 is downregulated by autophagy in senescence and ageing. *Nature Cell Biology*, *22*(10), 1170–1179.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., de Felicio, R., Fenner, A., Wong, W. R., Linington, R. G., Zhang, L., Debonsi, H. M., Gerwick, W. H., & Dorrestein, P. C. (2013). Molecular Networking as a Dereplication Strategy. *Journal of Natural Products*, *76*(9), 1686–1699.

Yang, Z., Huang, X., Lai, W., Tang, Y., Liu, J., Wang, Y., Chu, K., Brown, J., & Hong, G. (2021). Synthesis and identification of a novel derivative of salidroside as a selective, competitive inhibitor of monoamine oxidase B with enhanced neuroprotective properties. *European Journal of Medicinal Chemistry*, *209*, 112935.

Yao, Y., & Xu, B. (2021). New insights into chemical compositions and health promoting effects of edible oils from new resources. *Food Chemistry*, *364*, 130363.

Ye, W.-C., Zhang, Q.-W., Liu, X., Che, C.-T., & Zhao, S.-X. (2000). Oleanane saponins from Gymnema sylvestre. *Phytochemistry*, *53*(8), 893–899.

Yi, J., & Luo, J. (2010). SIRT1 and p53, effect on cancer, senescence and beyond. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1804*(8), 1684–1689.

Younis, I. Y., Khattab, A. R., Selim, N. M., Sobeh, M., Elhawary, S. S., & Bishbishy, M. H. El. (2022). Metabolomics-based profiling of 4 avocado varieties using HPLC–MS/MS and GC/MS and evaluation of their antidiabetic activity. *Scientific Reports*, *12*(1), 4966.

Yu, R., Chen, J., Xu, J., Cao, J., Wang, Y., Thomas, S. S., & Hu, Z. (2017). Suppression of muscle wasting by the plant-derived compound ursolic acid in a model of chronic kidney disease. *Journal of Cachexia, Sarcopenia and Muscle*, *8*(2), 327–341.

Zhang, B., Salituro, G., Szalkowski, D., Li, Z., Zhang, Y., Royo, I., Vilella, D., Díez, M. T., Pelaez, F., Ruby, C., Kendall, R. L., Mao, X., Griffin, P., Calaycay, J., Zierath, J. R., Heck, J. V., Smith, R. G., & Moller, D. E. (1999). Discovery of a Small Molecule Insulin Mimetic with Antidiabetic Activity in Mice. *Science*,

284(5416), 974–977.

Zhang, S., Hu, D.-B., He, J.-B., Guan, K.-Y., & Zhu, H.-J. (2014). A novel tetrahydroquinoline acid and a new racemic benzofuranone from *Capparis spinosa* L., a case study of absolute configuration determination using quantum methods. *Tetrahedron*, *70*(4), 869–873.

Zhou, H.-C., Hou, Z.-W., Wang, D.-X., Ning, J.-M., & Wei, S. (2019). Large scale preparation, stress analysis, and storage of headspace volatile condensates from *Jasminum sambac* flowers. *Food Chemistry*, *286*, 170–178.

RightsLink

meaningful by sharing even rumors or scientific tales during lunch and dinner. Thank you to my dear sisters Nguyen Hai Yen, Nguyen Thi Kim Yen, Dau Thi Thanh Hao, and Bui Thi My Hang, as well as my dear friends Nguyen Thi Trang, Tran Thi Ngoc, and Nguyen Thi Ha Thu for always being there for me when I needed support, encouragement to take on difficult tasks that I had never attempted before, and for sharing happy times with me.

Last but not least, I would like to express my gratitude to my beloved family, and friends in Vietnam who always love me unconditionally and cheer me up, with my great honesty, sincerity, and wholeheartedness. Herein, I will write to them in Vietnamese:

Gửi tới những người bạn thân yêu của tôi, Huyền Trang, Hồng, Trang, Ly những lời cảm ơn chân thành nhất từ tận đáy lòng. Cảm ơn mấy đứa đã luôn lắng nghe tao lảm nhảm, than thở suốt 6 năm trời. Nếu không có mấy đứa, chắc con đường này đã bị rẽ lối giữa chừng rồi. Dù chúng ta mỗi đứa đi trên một con đường riêng, nhưng với tao, mấy đứa luôn là gia đình. Vậy nên, dù ở đâu, tao cũng luôn cảm ơn và mong tất cả chúng ta sẽ có cuộc sống vui vẻ, hạnh phúc, theo cách này hay cách khác.

Gửi bố mẹ, gia đình của con. Con cảm ơn bố mẹ vì dù có không hiểu con đường con đi vẫn luôn âm thầm ủng hộ, cổ vũ, động viên con. Với con, những giây phút khó khăn chỉ cần gọi điện về cho bố mẹ, nghe giọng mọi người, và biết mọi người vẫn luôn khỏe mạnh là con có thể yên tâm quay lại làm việc tiếp. Vậy nên, con mong bố mẹ hiểu rằng, với con, chỉ cần là bố mẹ khỏe mạnh bình an đã luôn là chỗ dựa tinh thần vững chắc cho con đi trải nghiệm thế giới bao la ngoài kia rồi. Cảm ơn sự hy sinh của bố mẹ để con có thể có điều kiện đi ra ngoài học hỏi, khám phá, trải nghiệm cuộc sống như con từng mong.

Cảm ơn tất cả mọi người vì đã hiện diện trên những chặng đường tôi đi!
Thank you everyone for your presence in my life!