SUPPLEMENTARY MATERIAL

Three undescribed monoterpene rhamnosides from the aerial parts of *Vangueria agrestis*

Ahmed G. Osman[a], Omer Fantoukh[b,c], Zulfiqar Ali[a,*], Vijayasankar Raman[a], Mohamed A. Ibrahim[a] and Ikhlas A. Khan[a,b,*]

[a]*National Center for Natural Products Research, School of Pharmacy, University of Mississippi, University, Mississippi 38677, United States*

[b]*Division of Pharmacognosy, Department of BioMolecular Sciences School of Pharmacy, University of Mississippi, University, Mississippi 38677, United States*

[c]*Department of Pharmacognosy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia*

*Zulfiqar Ali: zulfiqar@olemiss.edu; Ikhlas Khan: ikhan@olemiss.edu

ABSTRACT: Phytochemical investigation of the methanolic extract of *Vangueria agrestis* aerial parts led to the isolation and characterization of four monoterpene rhamnosides, two ursane saponins, and six flavonoid glycosides. Of these constituents, are three undescribed monoterpene glycosides, vangagrestosides A-C. The reported constituents comprise the monoterpene glycoside [(2*E*,6*Z*)-2,6-dimethyl-8-[*O*-α-L-rhamnopyranosyl-(1→3)-α-L-rhamnopyranosyl)-oxy]-octadien-1-yl α-L-rhamnopyranoside], the saponins [3-*O*-[α-L-rhamnopyranosyl-(1→2)-α-L-arabinopyranosyl]pomolic acid, and 3-*O*-[α-L-rhamnopyranosyl-(1→2)-*β*-D-xylopyranosyl]pomolic acid], and the flavonoids [narcissoside, rutin, hyperoside, guajavarin, biorobin, and nicotifloroside].

Keywords: *Vangueria agrestis*, *Fadogia agrestis*, monoterpene rhamnopyranosides, vangagrestosides A-C

CONTENTS

General experimental
Determination of absolute configuration of sugars
Table S1. ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) data of compounds **1**-**3**.
Figure S1. Key COSY and HMBC correlations of vangagrestosides A-C (**1**-**3**).
Figure S2. ^{1}H NMR spectrum of compound **1.**
Figure S3. ^{13}C NMR spectrum of compound **1.**
Figure S4. ^{1}H NMR spectrum of compound **2.**
Figure S5. ^{13}C NMR spectrum of compound **2.**
Figure S6. ^{1}H NMR spectrum of compound **3.**
Figure S7. ^{13}C NMR spectrum of compound **3.**
Figure S8. HRMS of compound **1.**
Figure S9. HRMS of compound **2.**
Figure S10. HRMS of compound **3.**

General experimental

Optical rotations were measured in MeOH, using AUTOPOL IV Automatic Polarimeter (Rudolph, Hackettstown, NJ, USA). IR spectra were recorded on an Agilent Technologies Carry 630 FTIR. An Agilent Technologies 6200 series mass spectrometer was employed for ESIMS. NMR data in pyridine-d_6 or CD_3OD were obtained on Bruker Avance DRX-500 MHz or 400 MHz spectrometers for ^{1}H experiments and 100 MHz for ^{13}C experiments using standard pulse instrument software with solvent peaks as internal standards. Column chromatography (CC) was performed over flash silica gel (32-63µ, Dynamic adsorbents Inc.) and reversed-phase (RP) C_{18} (Polar bond, J. T. Baker). HPLC analysis and purification were conducted using an Agilent 1100 HPLC system equipped with a degasser (G1379A), quaternary pump (G13311A), autosampler (G1313A), column oven (G1316A), and UV-Diode detector (G1315B) controlled by Chemstation software. The chromatographic work on HPLC was carried out on RP C_{18} columns (150 × 4.6 mm; particle size 5 µm; Luna) and (250 × 10.0 mm; particle size 10 µm; Luna) with column oven temperature set at 25 °C and using the gradient system of eluent water (A) and

acetonitrile (B), acetic acid was added as a modifier to achieve a final concentration of 0.1% in each solvent. The gradient condition was as follows: 0-2 min (5% B), 2-30 min (50% B), 30-35 min (100% B). The flow rates of the solvent were 1.0 mL/min for the analytical injections while 5 mL/min for the semi-preparative ones and the injection volumes were 5.0 and 50 µL for the analytical and semi-preparative, respectively. All the analysis was carried out at wavelengths of 220 nm and 254 nm with a run time of 35 minutes. Reversed-phase HPLC [Waters Alliance 2695, equipped with photodiode array detector, and Luna C_{18} column (150 × 4.6 mm, 5 µm particle size; Phenomenex, Inc.)] was used for determination of sugar's absolute configuration. Analytical TLC was carried out on silica gel F_{254} aluminum sheet (20 x 20 cm, Fluka) or Silica 60 RP-18 $F_{254}S$ aluminum sheet (20 x 20 cm, Merck). The detection of the spots was made possible by visualization under UV-254 nm and by spraying with 1% vanillin (Sigma) in H_2SO_4-EtOH (1:9) following by heating at 110 °C for 2 min. Analytical grade solvents were used for the isolation and the purification procedures.

Determination of absolute configuration of sugars

The solution of **1** (1 mg) in 2 M HCl (1 mL) was heated at 95 °C for 2 h (Wang Y-H, Avula B, Fu X, Wang M, Khan IA. 2012. Simultaneous determination of the absolute configuration of twelve monosaccharide enantiomers from natural products in a single injection by a UPLC-UV/MS method. Planta Med. 78(08):834-837). The mixture was extracted with EtOAc after neutralizing with NH_4OH. The aqueous layer after drying was dissolved in pyridine (1 mL) and 0.1 M L-cysteine methyl ester hydrochloride in pyridine (1 mL) was added. The reaction mixture was heated at 90 °C for 1 h. Phenyl isothiocyanate in pyridine (1 mL) was added to the mixture and heated at 90 °C for an additional 1 h. The mixture was analyzed by reversed-phase HPLC using elution of water containing 0.05% formic acid (A) and a mixture of acetonitrile/methanol/isopropanol (2:1:1, v/v) with 0.05% formic acid (B) in a gradient mode: 14 % B to 16.5 % B for 22 min. and increasing B to 100% B in the next 0.5 min., at a rate of 0.30 mL/min. The response was detected at 254 nm. Similar treatment was done for compounds **2** and **3**. The standard sugar was derivatized and analyzed likewise. The sugars obtained from compounds were identified as L-rhamnose by comparison of the retention times of their derivatives with that of authentic sugar (L-rhamnose: 23.1 min.).

Table S1. ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) data of compounds **1-3**.

	1[a]		**2**[b]		**3**[b]	
position	δ_C mult.	δ_H mult. (*J* in Hz)	δ_C mult.	δ_H mult. (*J* in Hz)	δ_C mult.	δ_H mult. (*J* in Hz)
Monoterpene moiety						
1	73.51 CH_2	3.81 d (11.9) 3.96 d (11.9)	73.66 CH_2	3.83 d (11.7) 4.02 d (11.7)	73.79 CH_2	3.79 d (12.0) 3.97 d (12.0)
2	132.98 C		133.21 C		133.05 C	
3	128.56 CH	5.31 t (7.3)	128.55 CH	5.44 t (5.3)	128.92 CH	5.43 t (6.8)
4	26.88 CH_2	1.89 m, 1.98 m	27.25 CH_2	2.18[c]	27.01 CH_2	2.17[c]
5	39.97 CH_2	1.88 m	32.51 CH_2	2.18[c]	40.17 CH_2	2.12 t (7.3)
6	141.14 C		142.05 C		141.84 C	
7	121.64 CH	5.25 t (7.2)	122.18 CH	5.39 t (1.3, 6.8)	121.52 CH	5.36 t (7.3)
8	64.08 CH_2	4.01[c], 4.08[c]	63.81 CH_2	4.04[c], 4.11[c]	64.18 CH_2	4.08[c], 4.13[c]
9	14.31 CH_3	1.51 s	14.00 CH_3	1.69 s	14.14 CH_3	1.67 s
10	16.63 CH_3	1.54 s	23.53 CH_3	1.80 s	16.44 CH_3	1.72 s
Rha 1						
1	100.35 CH	4.81 d (1.7)	100.10 CH	4.67 d (1.7)	100.13 CH	4.67 d (1.7)
2	72.34 CH	3.99 dd (1.7, 3.3)	72.22 CH	3.93 dd (1.7, 3.4)	72.34 CH	3.94 dd (1.7, 3.4)
3	73.23 CH	4.05 dd (3.0, 9.5)	72.28 CH	3.64 dd (3.4, 9.5)	72.44 CH	3.62 dd (3.4, 9.5)
4	74.08 CH	3.55 t (9.5)	73.83 CH	3.34 t (9.5)	73.97 CH	3.36 t (9.5)
5	70.17 CH	3.78 m	69.68 CH	3.86 m	69.79 CH	3.87 m
6	18.50 CH_3	1.32 d (6.0)	17.87 CH_3	1.26 d (6.0)	18.02 CH_3	1.27 d (6.0)
Rha 2						
1	100.51 CH	4.84 d (1.7)	99.99 CH	4.69 d (1.5 Hz)	100.13 CH	4.70 d (1.4 Hz)
2	72.45 CH	4.10 dd (1.6, 3.0)	71.94 CH	3.85 dd (1.5, 3.3)	72.07 CH	3.87 dd (1.4, 3.4)
3	79.92 CH	4.02 dd (3.3, 9.4)	79.27 CH	3.65 dd (3.3, 9.5)	79.41 CH	3.76 dd (3.4, 9.3)
4	73.24 CH	3.66 t (9.4)	73.55 CH	3.37 t (9.5)	73.74 CH	3.37 t (9.3)
5	70.28 CH	3.77 m	70.34 CH	3.60 m	70.48 CH	3.60 m
6	18.45 CH_3	1.27 d (6.2)	17.77 CH_3	1.18 d (6.2)	17.92 CH_3	1.19 d (6.2)
Rha 3						
1	104.29 CH	5.34 d (1.6)	100.48 CH	5.06 d (1.5)	100.57 CH	5.08 brs
2	72.26 CH	4.23 dd (1.6, 3.3)	73.62 CH	5.31 dd (1.5, 3.1)	73.69 CH	5.32 dd (1.8, 3.4)
3	72.71 CH	3.91 dd (3.3, 9.4)	78.29 CH	4.06 dd (3.1, 9.6)	78.47 CH	4.09 dd (1.8, 9.6)
4	74.23 CH	3.68 t (9.4)	73.66 CH	3.51 t (9.6)	73.69 CH	3.57 t (9.6)
5	70.01 CH	3.78 m	70.05 CH	3.59 m	70.16 CH	3.58 m
6	18.50 CH_3	1.32 d (6.0)	18.01 CH_3	1.28 d (6.2)	18.19 CH_3	1.29 d (6.1)
Rha 4						
1			103.87 CH	5.00 d (1.6)	104.02 CH	5.00 d (1.6)
2			72.06 CH	3.92 dd (1.6, 3.3)	72.23 CH	3.94 dd (1.6, 3.4)
3			72.01 CH	3.58 dd (3.3, 9.6)	72.12 CH	3.66 dd (3.4, 9.5)

4			73.66 CH	3.47 t (9.6)	73.17 CH	3.52 t (9.5)
5			69.96 CH	3.84 m	70.07 CH	3.86 m
6			17.87 CH_3	1.27 d (6.0)	18.01 CH_3	1.28 d (6.0)
Acyl						
1'			168.53 C		168.62 C	
2'			128.62 C		128.96 C	
3'			143.96 CH	6.81 td (1.2, 7.3)	143.77 CH	6.81 td (1.3, 7.3)
4'			27.74 CH_2	2.39 m	27.86 CH_2	2.38 m
5'			38.91 CH_2	2.21 m	38.96 CH_2	2.22 m
6'			138.07 C		142.18 C	
7'			125.75 CH	5.40 td (1.2, 6.9)	120.61 CH	5.41 td (1.2, 7.0)
8'			59.25 CH_2	4.13 d (6.9)	62.19 CH_2	4.62 d (7.0)
9'			12.42 CH_3	1.88 s	12.59 CH_3	1.86 s
10'			16.05 CH_3	1.72 s	16.47 CH_3	1.76 s
*C*OCH_3					172.95 C	
CO*C*H_3					20.95 CH_3	2.04 s
[a]NMR data were recorded in mixture of C_5D_5N and CD_3OD. [b]NMR data were recorded in CD_3OD. [c]The chemical shifts are overlapped.						

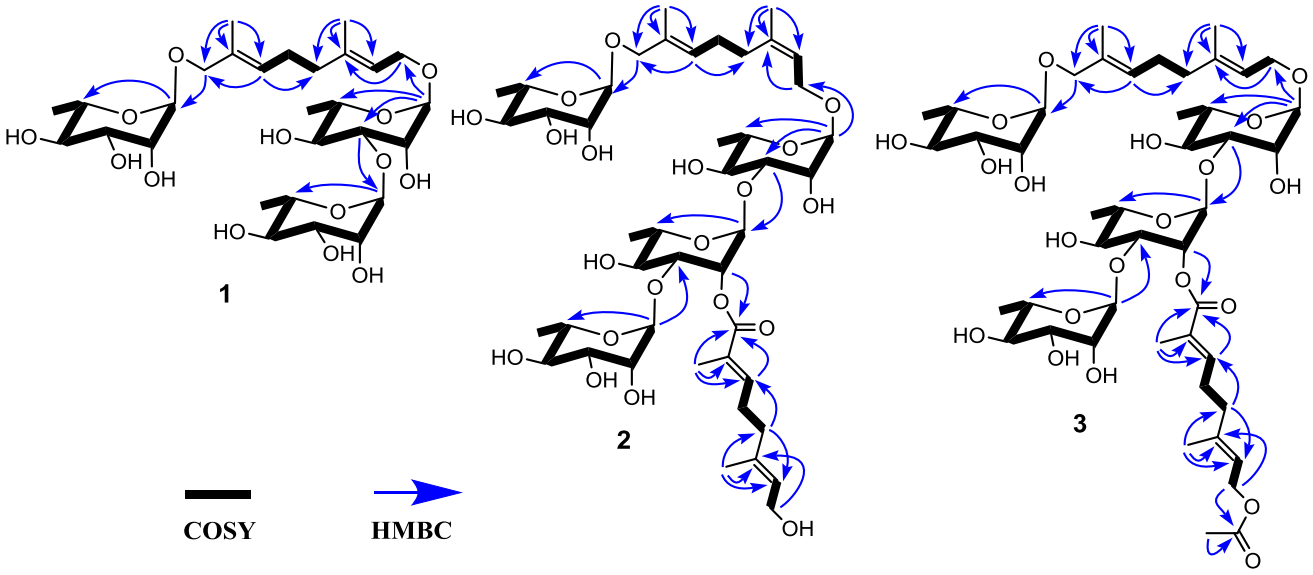

Figure S1. Key COSY and HMBC correlations of vangagrestosides A-C (**1**-**3**).

.

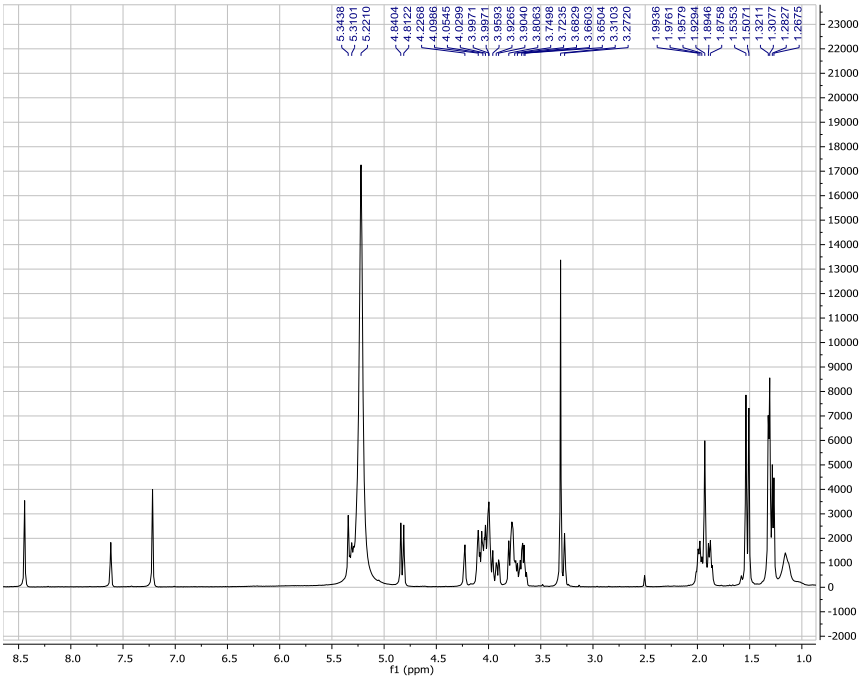

Figure S2. ^{1}H NMR spectrum of compound **1.**

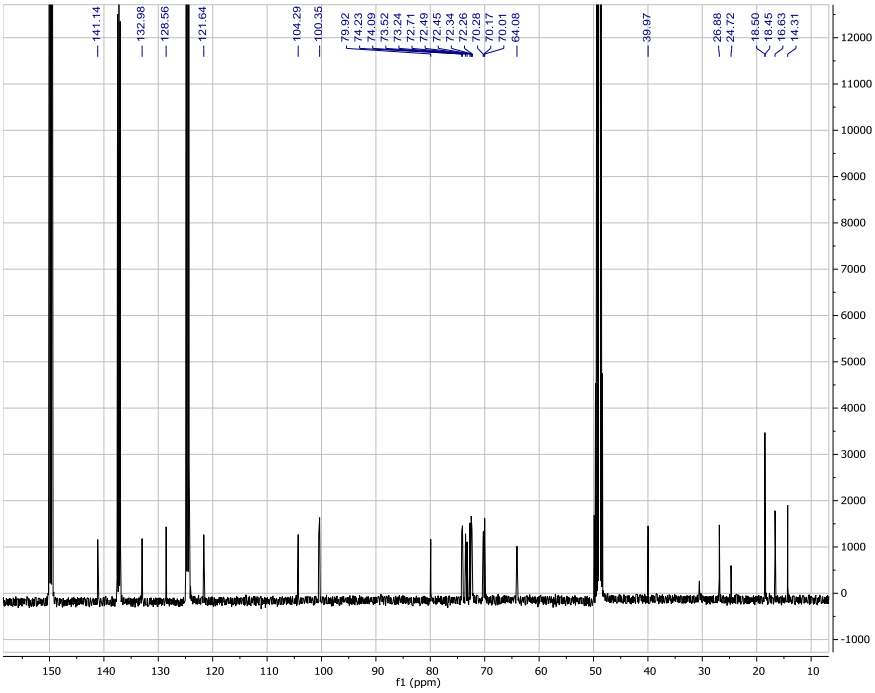

Figure S3. ^{13}C NMR spectrum of compound **1.**

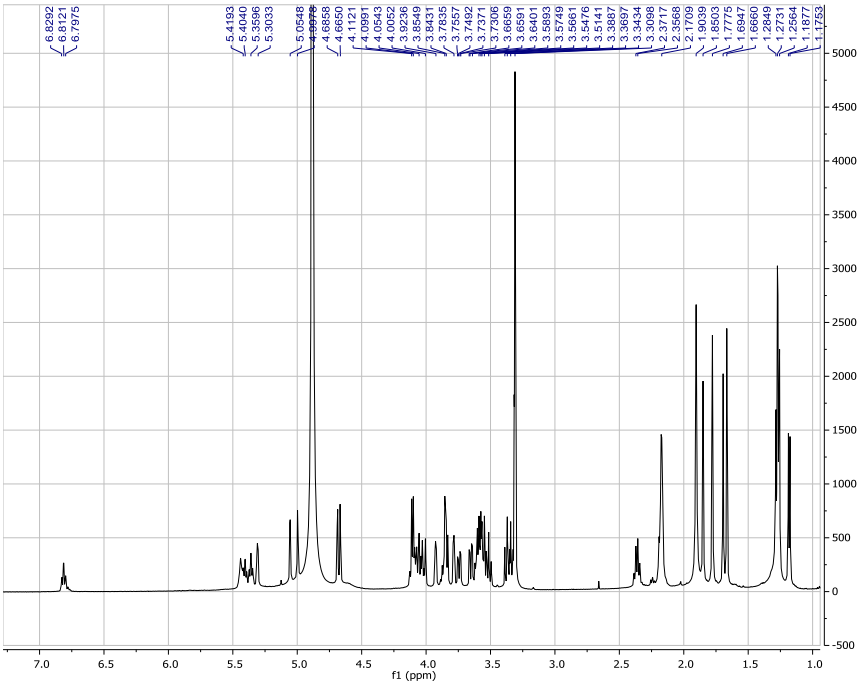

Figure S4. ^{1}H NMR spectrum of compound **2.**

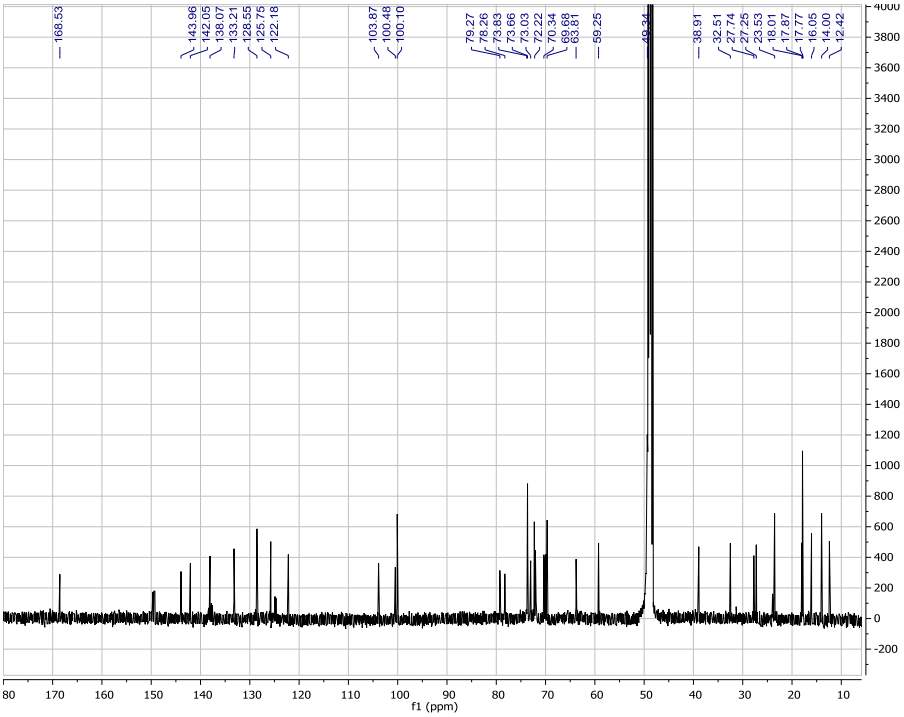

Figure S5. ^{13}C NMR spectrum of compound **2.**

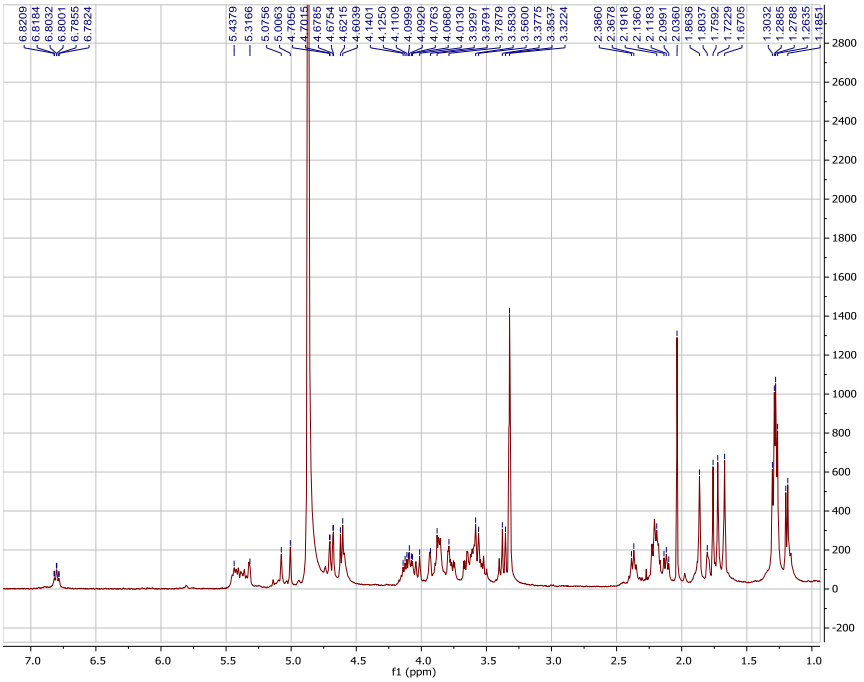

Figure S6. ^{1}H NMR spectrum of compound **3.**

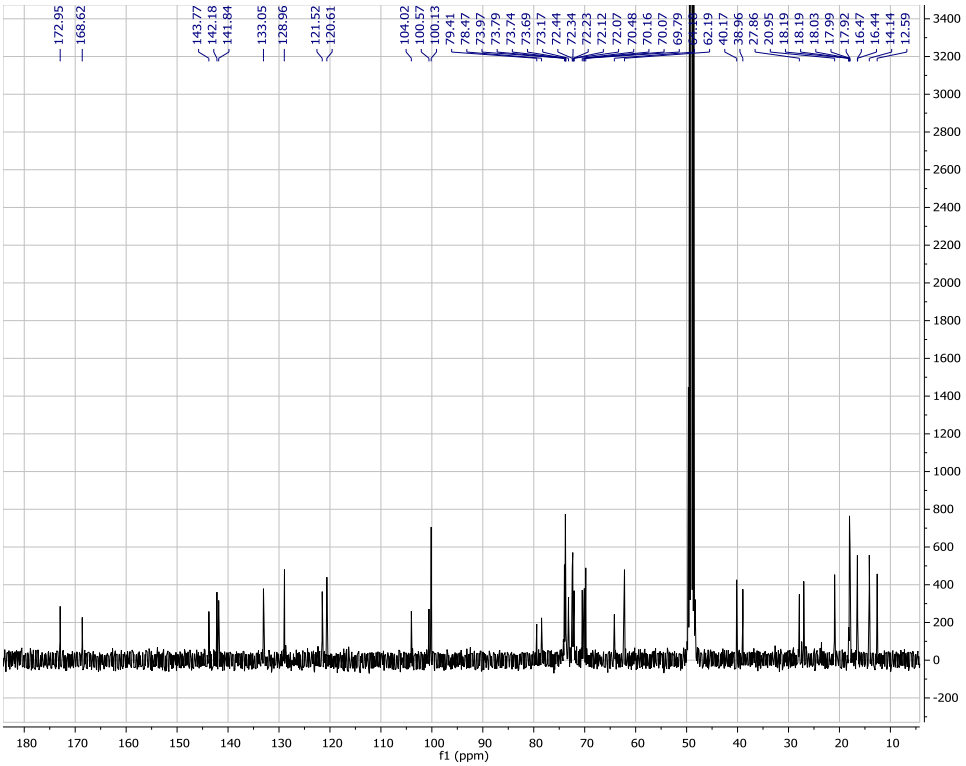

Figure S7. ^{13}C NMR spectrum of compound **3.**

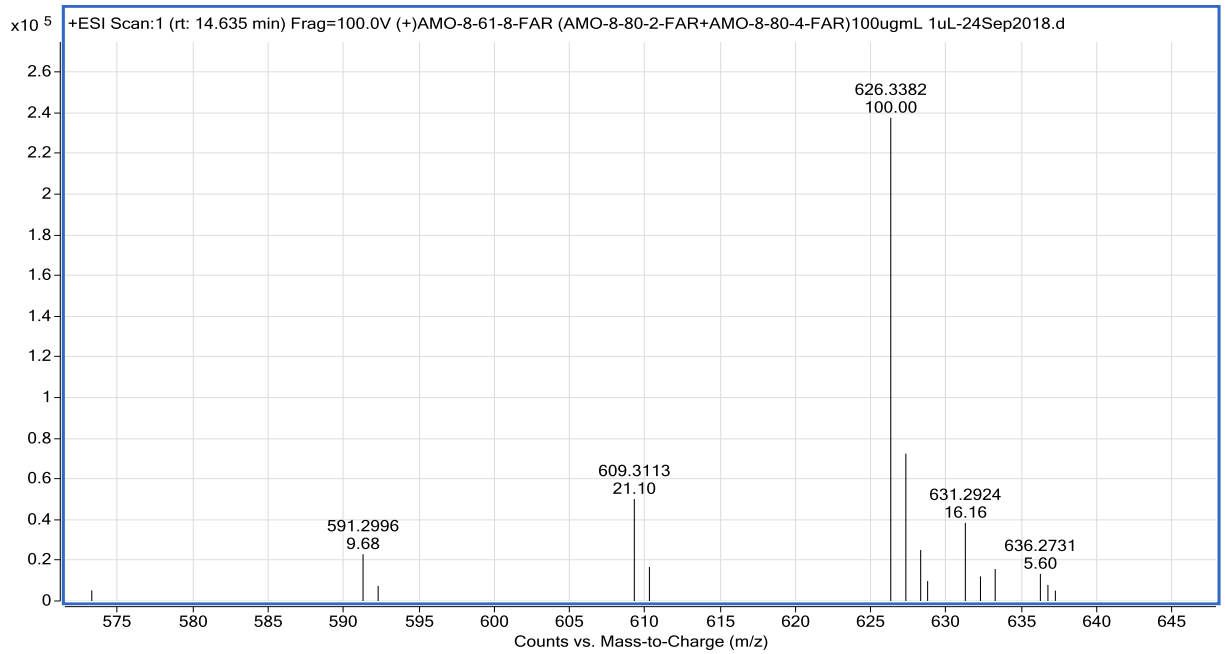

Figure S8. HRMS of compound **1.**

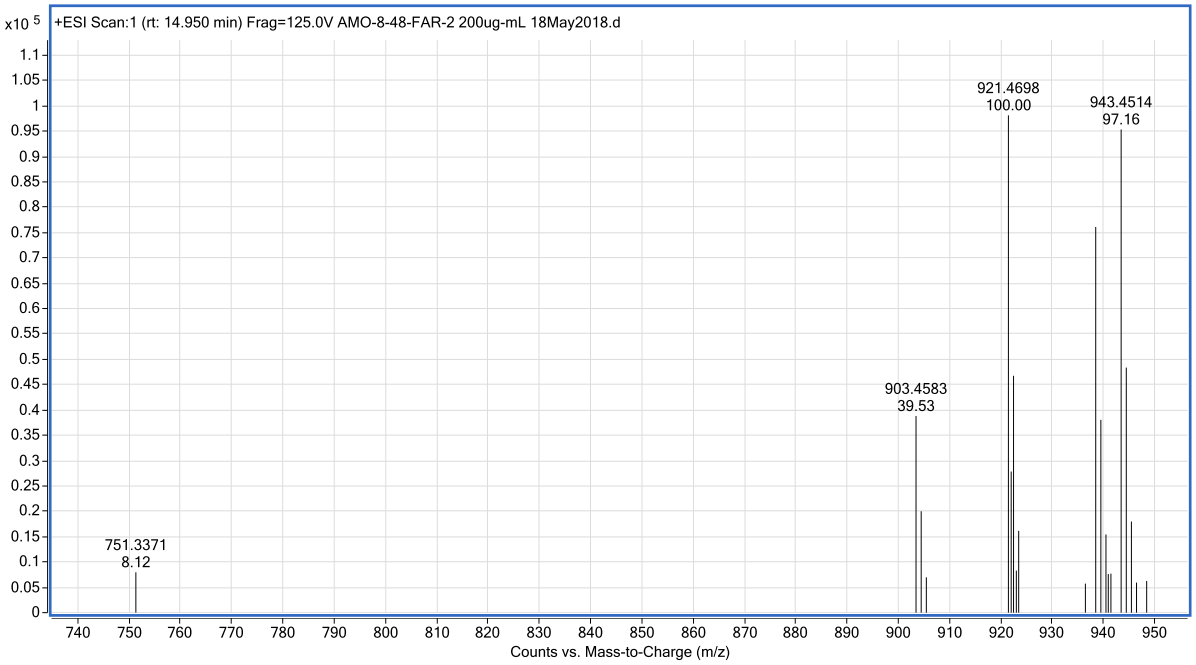

Figure S9. HRMS of compound **2.**

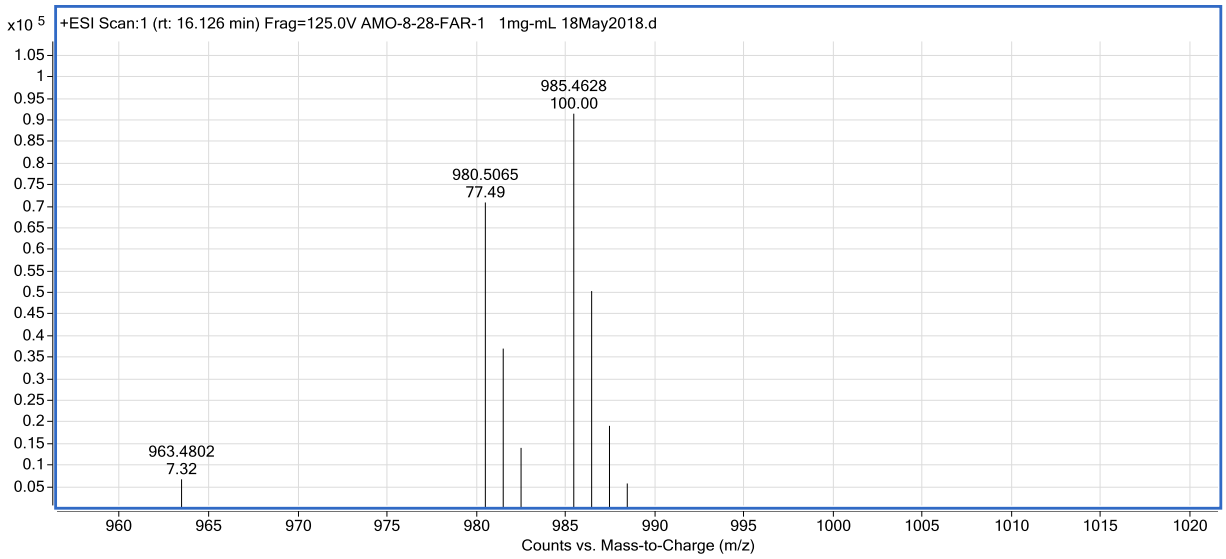

Figure S10. HRMS of compound **3.**