Tetrahedron 119 (2022) 132858

Contents lists available at ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Arabincosides A-D, pregnane glycosides isolated from *Caralluma arabica*

Essam A. Abdel-Sattar [a, *], Othman S.S. Al-Hawshabi [b], Akram A. Shalabi [a], Ali M. El Halawany [a], Meselhy R. Meselhy [a]

[a] Department of Pharmacognosy, Cairo University, Cairo, Egypt
[b] Department of Biology, Faculty of Science, University of Aden, Aden, Yemen

ARTICLE INFO

Article history:
Received 1 May 2022
Received in revised form
29 May 2022
Accepted 31 May 2022
Available online 2 June 2022

Keywords:
Caralluma arabica
Pregnane glycosides
Apocyanaceae
Arabincosides
NMR
MS

ABSTRACT

The aqueous remained fraction after partitioning the crude ethanol extract with ether and methylene chloride of the aerial parts of *Caralluma arabica* yielded six pregnane glycosides, four of them were found to be new pregnane glycosides and were identified using spectral analysis (MS, 1D and 2DNMR) as 3β,14β, dihydroxypregn-5-en-20-one-3-*O*-β-D-glucopyranosyl-(1 → 4)-*O*-β-D-digitalopyranoside (arabincoside A), 3β,14β-dihydroxypregn-5-en-20-one-3-*O*-β-D-glucopyranosyl-(1 → 6)-*O*-β-D-glucopyranosyl-(1 → 4)-*O*-β-D-digitalopyranoside (arabincoside B), 12-β-*O*-benzoyl-3β, 12β, 14β, 20-tetrahydroxy-(20*R*)-pregn-5-ene-3-*O*-β-D-glucopyranosyl-(1 → 4)-β-D-cymaropyranosyl-(1 → 4)-β-D-cymaropyranoside (arabincoside C) and 12-β-*O*-benzoyl-3β, 12β, 14β, 20-tetrahydroxy-(20*R*)-pregn-5-ene-3-*O*-β-D-glucopyranosyl-(1 → 4)-β-D-glucopyranosyl-(1 → 4)-β-D-cymaropyranosyl-(1 → 4)-β-D-cymaropyranoside (arabincoside D). In addition, two known pregnane glycosides were also isolated and identified as russelioside D and russelioside C, previously isolated from *C. russelliana.*

1. Introduction

Due to the topographic differences and variations in climatic factors in Yemen, its flora is considered the richest one in the Arabian Peninsula and characterized by a high diversity and density as well as including numerous plant species which are used in the traditional medicine [1,2]. Among these traditional medicinal plants, members of the genus *Caralluma* R. Br. (family Apocyanaceae; formerly Asclepiadaceae) are growing wildly and common on all habitats in Yemen. Plants of the genus *Caralluma* are widely distributed in other parts of the Arabian peninsula and parts in Africa and Asia [3]. The genus *Caralluma* includes about 26 species among them 20 species are endemic and near endemic to Yemen [4–6]. Members of this genus demonstrated antidiabetic, antirheumatic, antiobesity, antiulcer and antimicrobial activities [7,8].

C. arabica is a clump-forming succulent perennial herb with purple-red flowers in apical umbels [9]. The herb is widely distributed throughout Saudi Arabia, Yemen, Oman, United Arab Emirates, and the horn of Africa [11]. In the Arabian peninsula, a decoction of *C. arabica* is traditionally used as nutritive and tonic, and for the treatment of liver diseases, diabetes, hypertension, cuts, wounds, burns and itchy skin [7,12]. *C. arabica* is considered as an important food in the hungry period immediately after the first rains. Miller & Morris, 2004 [13] reported that the 10% ethanolic extract of *C. arabica* (at 200 and 400 mg/kg) showed good anti-inflammatory and anti-nociceptive effects in carrageenan-induced rat paw oedema and cotton pellet granuloma models. In addition, the extract reduced the gastric acidity and demonstrated cytoprotective properties [14]. The ethyl acetate fraction of the alcoholic extract of *C. arabica* showed good antioxidant and lipoxygenase inhibitory effects [15]. Luteolin-3′-*O*-α-L-rhamnoside and kaempferol-3-*O*-β-D-glucosyl-4′-*O*-α-L-rhamnoside were isolated from the alcoholic extract [16], while the phenolic acids (gallic acid, vanillic acid, syringic acid, *p*-coumaric acid, and ferulic acid) and flavonoids (epicatechin, quercetin-3-β-*D*-glucoside and rutin) were identified in the ethyl acetate fraction [15].

In continuation of our work to identify bioactive molecules from *Caralluma* species growing in the Arabian Peninsula, we report herein the isolation and structure elucidation of six pregnane glycosides from the water-soluble fraction of the alcoholic extract of

* Corresponding author. Pharmacognosy Department, Faculty of Pharmacy, Cairo University, El-Kasr El-Aini Street, 11562 Cairo, Egypt.
E-mail address: essam.abdelsattar@pharma.cu.edu.eg (E.A. Abdel-Sattar).

https://doi.org/10.1016/j.tet.2022.132858
0040-4020/

C. arabica N.E.Br. collected near Saihut in Yemen.

2. Results and discussion

Repeated column chromatography of the remaining aqueous fraction (RAF) of the alcoholic extract of *C. arabica* afforded six pregnane glycosides (**1**–**6**) four of which are new natural products. Compounds **1**–**6** gave positive tests for sterols (Liebermann-Bucrchard reagent) and sugars and/or glycosides (Molish reagent).

Analysis of the spectral data (1D and 2DNMR) of the isolated compounds (Tables 1–4) revealed the presence of a pregnane skeleton in all the isolated compounds; three methyls of C-18, C-19 and C-21 (δ_H 0.88–1.4) and one olefinic proton (δ_H 5.35–5.47) in the ^{1}HNMR spectra of compounds **1**–**6 (**Fig. 1**),** a free OH at C-20 for compounds **1** and **5,** and a ketonic group at C-20 for compounds **2**–**4,** and **6** (^{1}HNMR and ^{13}CNMR spectra).

Compounds **3** and **6** were identified as russelioside C and russelioside D, respectively, previously isolated and identified from *C. russelliana* by comparison with the reported spectral data [17].

Compound **1** was obtained as a white powder (100 mg), the ESI-MS analysis showed an adduct ion peak at 677.1 $[M+Na]^+$ (100) in positive mode, and 653.1 $[M-H]^-$ (66), 689.1 $[M+Cl]^-$ (34) in negative mode, assigned for the molecular formula $C_{34}H_{54}O_{12}$. IR spectrum of **1** showed absorption bands for hydroxyl (3418 cm^{-1}), carbonyl (1716), and olefinic double bond (1635, 960 cm^{-1}) groups.

The ^{1}H, ^{13}C, and DEPT NMR spectra of **1** showed signals for two sugar units, one of them identified as 6-deoxy sugar, as well as signals for a C_{21}-steroidal skeleton. The ^{1}H, ^{13}C NMR spectra showed the presence of four methyls, one methoxy, fifteen methines, nine methylene groups, four quaternary carbon atoms and one carbonyl with a total of 34 carbons, 21 of them were ascribed for the C_{21} steroidal moiety (Tables 1–4). Two angular methyl groups of the pregnane skeleton [18] were identified as singlets at δ_H 1.00 and 1.04 (δ_C 14.09 & 18.52, respectively), assigned for the methyls CH_3-18 and CH_3-19. The correlations between H-l to H-4, H-6 to H-12, and H-15 to H-17 of the aglycone, were clearly displayed in the TOCSY spectrum. In addition, the ^{1}H and ^{13}C NMR spectra of **1** revealed the presence of carbonyl group at C-20 (δ_C 218.35) which was confirmed by the downfield shift of CH_3-21 (δ_H 2.27 & δ_C 31.29) [18]. The presence of C-5/C-6 olefinic double bond was deduced from the proton signal at δ_H 5.43 ppm (*br d*) that was assigned to H-6 (C-6 at δ_C 121.66), and from the signal at δ_C 139.31 (^{13}C NMR) assigned to C-5. Other signals of the aglycone in ^{1}H and ^{13}C-NMR were found to be undistinguished from those of caratubergenin A [19–21]. Regarding the relative configuration at C-17, it was confirmed to be α-configuration for H-17, and β configuration for its C-17 side chain. This result was deduced by comparing the chemical shift and coupling constant of the protons H-17 and H-16 in **1** with those previously reported in the same aglycone [19–21]. In ^{1}HNMR spectrum of **1**, H-17 appeared at δ_H 2.99 as a doublet of doublet (J = 9.28 and 4.40 Hz) and showed a strong correlation contour with H-16 signal at δ 2.03, which should be assigned to H-16β,

Table 1
^{1}H -NMR (400 MHz) of aglycones of isolated compounds.

Code No.	1	2	3	4	5	6
	(CD_3OD)	(CD_3OD)	(CD_3OD)	(CD_3OD)	(DMSO-*d6*)	(CD_3OD)
1	1.11 (1H, *m*) 1.68 (1H, *m*)	1.16 (1H, *m*) 1.87 (1H, *m*)	1.77 (2H, *m*)	1.16 (1H, *m*) 1.87 (1H, *m*)	1.01 (1H, *m*) 1.80 (1H, *m*)	1.78 (2H, *m*)
2	2.27 (1H, *m*) 1.63 (1H, *m*)	1.52 (2H, *m*)	1.93 (2H, *m*)	1.90 (1H, *m*) 1.52 (1H, *m*)	1.48 (2H, *m*)	1.82 (2H, *m*)
3	3.56 (1H, *m*)	3.52 (1H, *m*)	3.55 (1H, *m*)	3.52 (1H, *m*)	3.40 (1H, *m*)	3.54 (1H, *m*)
4	1.93 (1H, *m*) 2.44 (1H, *dd*, J = 2.84, 10.48)	2.18 (1H, *m*) 2.37 (1H, *dd*, J = 2.96, 10.32)	1.85 (1H, *m*) 2.43 (1H, *dd*, J = 2.80, 13.36)	2.18 (1H, *m*) 2.37 (1H, *dd*, J = 3.00, 12.84)	1.52 (1H, *m*) 2.35 (1H, *dd*, J = 2.84, 10.48)	1.79 (1H, *m*) 2.33 (1H, *dd*, J = 2.84, 13.48)
5	–	–	–	–	–	–
6	5.43 (1H, *br d*, J = 5.24)	5.47 (1H, *br d*, J = 5.30)	5.43 (1H, *br d*, J = 5.56)	5.47 (1H, *br d*, J = 4.92)	5.35 (1H, *br d*, J = 5.24)	5.32 (1H, *br d*, J = 5.60)
7	1.85 (1H, *m*) 2.28 (1H, *m*)	1.89 (1H, *m*) 2.25 (1H, *m*)	1.87 (1H, *m*) 2.29 (1H, *m*)	1.90 (1H, *m*) 2.26 (1H, *m*)	1.74 (1H, *m*) 2.16 (1H, *m*)	1.76 (1H, *m*) 2.15 (1H, *m*)
8	1.90 (1H, *m*)	1.87 (1H, *m*)	1.67 (1H, *m*)	1.87 (1H, *m*)	1.58 (1H, *m*)	1.68 (1H, *m*)
9	1.19 (1H, *m*)	1.36 (1H, *m*)	1.09 (1H, *m*)	1.37 (1H, *m*)	1.10 (1H, *m*)	1.09 (1H, *m*)
10	–	–	–	–	–	–
11	1.48 (2H, *m*)	1.63 (2H, *m*)	1.51 (2H, *m*)	1.84 (1H, *m*) 1.65 (1H, *m*)	1.32 (1H, *m*) 1.43 (1H, *m*)	1.54 (2H, *m*)
12	1.56 (1H, *m*) 1.85 (1H, *m*)	4.82 (1H, *b.s*)	1.14 (1H, *m*) 1.68 (1H, *m*)	4.81 (1H, *b.s*)	1.34 (1H, *m*) 2.12 (1H, *m*)	1.11 (1H, *m*) 1.65 (1H, *m*)
13	–	–	–	–	–	–
14	–	–	–	–	–	–
15	2.14 (1H, *m*) 1.75 (1H, *m*)	1.94 (1H, *m*) 1.64 (1H, *m*)	1.99 (2H, *m*)	1.84 (1H, *m*) 1.68 (1H, *m*)	1.93 (1H, *m*) 1.61 (1H, *m*)	1.71 (1H, *m*) 2.08 (1H, *m*)
16	2.03 (1H, *m*) 1.91 (1H, *m*)	1.84 (1H, *m*) 1.96 (1H, *m*)	1.40 (2H, *m*)	1.96 (2H, *m*)	1.83 (1H, *m*) 1.93 (1H, *m*)	1.27 (1H, *m*) 1.39 (1H, *m*)
17	2.99 (1H, *dd*, J = 4.40, 9.28)	1.69 (1H, *m*)	1.23 (1H, *m*)	1.93 (1H, *m*)	2.72 (1H, *dd*, J = 4.8, 9.2)	1.28 (1H, *m*)
18	1.00 (3H, *s*)	1.40 (3H, *s*)	1.05 (3H, *s*)	1.40 (3H, *s*)	0.88 (3H, *s*)	0.96 (3H, *s*)
19	1.04 (3H, *s*)	1.05 (3H, *s*)	1.07 (3H, *s*)	1.06 (3H, *s*)	0.93 (3H, *s*)	0.94 (3H, *s*)
20	–	3.82 (1H, *m*)	4.01 (1H,, *q*, J = 8.28, 14.64)	3.82 (1H, *m*)	–	3.82 (1H,, *q*, J = 6.48, 13)
21	2.27 (3H, *s*)	1.20 (3H, *d*, J = 6.56)	1.10 (3H, *d*, J = 6.28)	1.20 (3H, *d*, J = 6.52)	2.20 (3H, *s*)	0.99 (3H, *d*, J = 6.20)
Bz-12						
1′′′′′		–		–		
2′′′′′		–		–		
3′′′′′,7′′′′′		8.08 (2H, *d*, J = 7.00)		8.08 (2H, *d*, J = 7.00)		
4′′′′′,6′′′′′		7.48 (2H, m)		7.48 (2H, *m*)		
5′′′′′		7.61 (1H, *m*)		7.61 (1H, *m*)		

Table 2
^{13}C -NMR (100 MHz) of aglycones of isolated compounds.

Code	1	2	3	4	5	6
No.	(CD_3OD)	(CD_3OD)	(CD_3OD)	(CD_3OD)	(DMSO-*d6*)	(CD_3OD)
1	37.07	37.01	37.12	37.00	37.18	37.13
2	29.28	29.27	29.31	29.26	29.76	29.30
3	78.60	77.32	78.66	77.63	77.58	78.62
4	38.13	38.37	38.93	38.35	38.71	38.94
5	139.31	139.09	139.39	139.09	139.66	139.40
6	121.66	121.70	121.65	121.67	122.19	121.65
7	26.97	26.84	26.72	26.82	27.40	26.72
8	36.41	36.08	36.90	36.09	37.14	36.90
9	46.03	43.31	46.33	43.32	45.89	46.34
10	36.89	36.97	36.42	36.95	36.98	36.42
11	20.40	25.47	20.73	25.46	20.82	20.55
12	38.22	78.32	38.17	78.33	38.85	38.17
13	48.77	52.47	47.61	52.45	49.07	47.62
14	85.49	85.40	84.37	85.41	84.62	84.38
15	33.71	32.08	32.35	32.04	33.85	32.36
16	23.88	25.30	20.55	25.26	23.79	19.88
17	62.45	52.18	56.32	52.20	62.99	56.32
18	14.09	9.91	13.70	9.86	15.70	13.70
19	18.52	18.48	18.54	18.44	19.66	18.55
20	218.35	70.36	64.95	70.34	215.61	64.96
21	31.29	21.69	17.37	21.64	31.94	17.37
Bz-12						
1′′′′′		166.55		166.57		
2′′′′′		130.46		130.46		
3′′′′′,7′′′′′		129.25		129.23		
4′′′′′,6′′′′′		128.21		128.19		
5′′′′′		132.87		132.85		

leaving its partner at δ 1.91 to be assigned to H-16α. In addition, the absence of spatial correlation between H-17 and Me-18 in the NOESY spectrum confirmed the α-configuration of H-17. Therefore, the aglycone **1** could be identified as 3β,14β, dihydroxypregn-5-en-20-one.

The linkage between the sugar chain and aglycone of **1** was proven at C-3 as a result of the downfield shift of C-3 (δ_C 78.60) and the upfield shifts of C-2 (δ_C 29.28) and C-4 (δ_C 38.13), which showed also correlations between H-3 (δ_H 3.56) of the aglycone and the anomeric carbon (δ_C 101.57) of the first sugar (digitalose) and between H-l′ (δ_H 4.36) and C-3 (δ_C 78.60). The sequence of the sugar units was confirmed through HMBC correlations (Fig. S1). The second sugar (β-D-glucose) was linked to C-4′ confirmed from the correlations between H-1′′ of the glucose unit (δ_H 4.62) and C-4′ of digitalose (δ_C 73.47). Therefore, the sugar chain was identified as β-D-glucopyranosyl-(1 → 4)-β-D-(3-*O*-methyl-6-deoxy) galactoside by comparing the ^{1}H, and ^{13}C- NMR of **1** with those reported for russelioside B, penicilloside A and other related glycosides [22–25]. From the above findings, compound **1** was identified as 3β,14β, dihydroxypregn-5-en-20-one-3-*O*-β-D-glucopyranosyl-(1 → 4)-*O*-β-D-digitalopyranoside (named as arabincoside A, a new natural product).

Compound **5** was obtained as a white powder (450 mg), ESI−MS analysis showed an adduct ion peak at *m*/*z* (rel. int.): 839.1 $[M+Na]^+$ (100) in positive mode, and 815.0 $[M-H]^-$ in negative mode, assigned for the molecular formula $C_{40}H_{64}O_{17}$. IR spectrum of compound **5** showed absorption bands due to hydroxyl

Table 3
^{1}H-NMR (400 MHz) of sugars moieties of isolated compounds.

Code	1	2	3	4	5	6
No.	(CD_3OD)	(CD_3OD)	(CD_3OD)	(CD_3OD)	(DMSO-*d6*)	(CD_3OD)
	Dig	**Cym**	**Dig**	**Cym**	**Dig**	**Dig**
1′	4.36 (1H, *d*, *J* = 7.72)	4.87 (1H, *d*, *J* = 7.68)	4.37 (1H, *d*, *J* = 7.72)	4.84 (1H, *d*, *J* = 7.68)	4.20 (1H, *d*, *J* = 7.60)	4.25 (1H, *d*, *J* = 7.68)
2′	3.60 (1H, *m*)	1.55 (1H, *m*)	3.61 (1H, *m*)	1.55 (1H, *m*)	3.58 (1H, *m*)	3.39 (1H, *m*)
		2.06 (1H, *m*)		2.06 (1H, *m*)		
3′	3.25 (1H, *m*)	3.85 (1H, *m*)	3.24 (1H, *m*)	3.85 (1H, *m*)	3.07 (1H, *m*)	3.22 (1H, *m*)
4′	4.19 (1H, *d*, *J* = 2.84)	3.24 (1H, *m*)	4.20 (1H, *d*, *J* = 2.80)	3.24 (1H, *m*)	4.05 (1H, *b.s*)	4.09 (1H, *d*, *J* = 2.8)
5′	3.61 (1H, *m*)	3.86 (1H, *m*)	3.64 (1H, *m*)	3.86 (1H, *m*)	3.52 (1H, *m*)	3.50 (1H, *m*)
6′	1.30 (3H, *d*, *J* = 6.3)	1.20 (3H, *d*, *J* = 6.56)	1.30 (3H, *d*, *J* = 6.40)	1.20 (3H, *d*, *J* = 6.56)	1.13 (3H, *d*, *J* = 6.2)	1.18 (3H, *d*, *J* = 6.28)
OCH_3	3.52 (3H, *s*)	3.45 (3H, *s*)	3.52 (3H, *s*)	3.45 (3H, *s*)	3.38 (3H, *s*)	3.41 (3H, *s*)
	Glc	**Cym**	**Glc**	**Cym**	**Glc**	**Glc**
1′′	4.62 (1H, *d*, *J* = 7.68)	4.82 (1H, *d*, *J* = 7.68)	4.62 (1H, *d*, *J* = 7.68)	4.80 (1H, *d*, *J* = 7.68)	4.28 (1H, *d*, *J* = 7.72)	4.49 (1H, *d*, *J* = 7.68)
2′′	3.24 (1H, *m*)	1.63 (1H, *m*)	3.27 (1H, *m*)	1.63 (1H, *m*)	2.95 (1H, *m*)	3.12 (1H, *m*)
		2.15 (1H, *m*)		2.15 (1H, *m*)		
3′′	3.39 (1H, *m*)	3.93 (1H, *m*)	3.38 (1H, *m*)	3.90 (1H, *m*)	3.05 (1H, *m*)	3.25 (1H, *m*)
4′′	3.28 (1H, *m*)	3.26 (1H, *m*)	3.28 (1H, *m*)	3.26 (1H, *m*)	3.07 (1H, *m*)	3.27 (1H, *m*)
5′′	3.20 (1H, *m*)	3.94 (1H, *m*)	3.22 (1H, *m*)	3.93 (1H, *m*)	3.31 (1H, *m*)	3.18 (1H, *m*)
6′′	3.66 (1H, *m*)	1.20 (3H, *d*, *J* = 6.56)	3.66 (1H, *m*)	1.32 (3H, *d*, *J* = 6.24)	3.97 (1H, *d*, *J* = 11.64)	3.35 (1H, *m*)
	3.90 (1H, *dd*, *J* = 4.68, 9.28)		3.90 (1H, *d*, *J* = 13.40)		3.58 (1H, *m*)	4.03 (1H, *dd*, *J* = 1.92, 11.84)
OCH_3		3.47 (3H, *s*)		3.47 (3H, *s*)		
		Glc		**Glc**	**Glc**	**Glc**
1′′′		4.35 (1H, *d*, *J* = 7.68)		4.38 (1H, *d*, *J* = 7.76)	4.36 (1H, *d*, *J* = 7.68)	4.29 (1H, *d*, *J* = 7.76)
2′′′		3.22 (1H, *m*)		3.22 (1H, *m*)	2.98 (1H, *m*)	3.09 (1H, *m*)
3′′′		3.33 (1H, *m*)		3.90 (1H, *m*)	3.05 (1H, *m*)	3.21 (1H, *m*)
4′′′		3.25 (1H, *m*)		3.50 (1H, *m*)	3.02 (1H, *m*)	3.20 (1H, *m*)
5′′′		3.33 (1H, *m*)		3.42 (1H, *m*)	3.10 (1H, *m*)	3.19 (1H, *m*)
6′′′		3.33 (1H, *m*)		3.66 (1H, *dd*, *J* = 5.4, 11.64)	3.67 (2H, *dd*, *J* = 5.6, 11.60)	3.16 (1H, *m*)
		3.66 (1H, *dd*, *J* = 5.4, 11.64)		3.84 (1H, *m*)		3.75 (1H, *dd*. 1.76, 11.76)
				Glc		
1′′′′				4.41 (1H, *d*, *J* = 7.80)		
2′′′′				3.24 (1H, *m*)		
3′′′′				3.54 (1H, *m*)		
4′′′′				3.32 (1H, *m*)		
5′′′′				3.84 (1H, *m*)		
6′′′′				3.87 (2H, *m*)		

Table 4
^{13}C -NMR (100 MHz) of sugars moieties of isolated compounds.

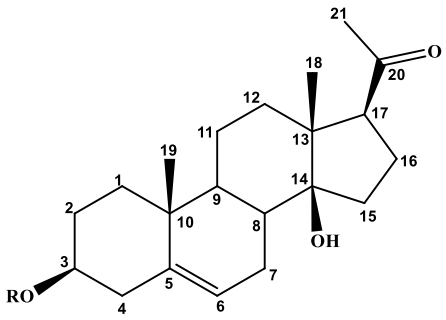
Code	1	2	3	4	5	6
No.	(CD_3OD)	(CD_3OD)	(CD_3OD)	(CD_3OD)	(DMSO-*d6*)	(CD_3OD)
	Dig	**Cym**	**Dig**	**Cym**	**Dig**	**Dig**
1′	101.57	95.87	101.58	95.86	101.75	101.55
2′	70.47	35.02	70.47	35.00	69.11	70.50
3′	84.36	77.18	84.37	77.17	84.49	84.30
4′	73.47	82.38	73.44	82.45	74.00	73.56
5′	70.23	68.72	70.23	68.69	69.75	70.00
6′	15.98	17.32	15.98	17.27	17.62	16.23
OCH_3	57.11	57.11	57.10	57.04	58.22	57.26
	Glc	**Cym**	**Glc**	**Cym**	**Glc**	**Glc**
1″	102.79	99.78	102.77	99.75	103.50	102.84
2″	74.54	35.29	74.53	35.24	74.63	73.75
3″	76.81	76.62	76.81	76.49	77.19	76.42
4″	70.06	82.45	70.06	82.51	70.92	70.25
5″	76.44	68.57	76.44	68.57	76.82	76.03
6″	61.63	17.16	61.63	17.11	69.47	68.95
OCH_3		57.18		57.15		
		Glc		**Glc**	**Glc**	**Glc**
1‴		104.84		103.21	103.92	103.68
2‴		73.89		73.59	73.96	74.43
3‴		77.63		77.25	77.19	76.68
4‴		70.38		74.91	70.52	70.25
5‴		76.54		75.11	77.36	76.61
6‴		61.61		61.04	61.54	61.38
				Glc		
1⁗				104.66		
2⁗				73.52		
3⁗				79.51		
4⁗				69.98		
5⁗				76.73		
6⁗				60.77		

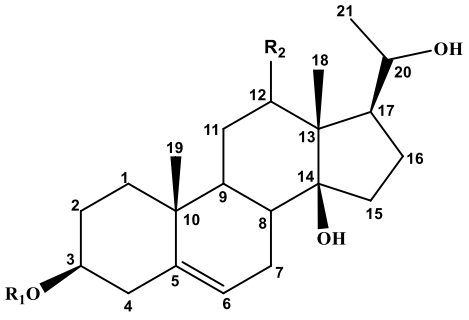
(3379 cm^{-1}), carbonyl (1681 cm^{-1}), and olefinic double bond (1639, 879 cm^{-1}) groups. The NMR spectral analysis of compound **5** revealed great similarities with the spectral data of compound **1**, except for the presence of an additional sugar moiety in the sugar chain at C-3. The presence of three anomeric protons and carbons indicated the presence of a trisaccharide. The sequential assignments of the proton and carbon resonances of the sugar moieties, as well as their connectivity to each other, were determined using HMBC and TOCSY experiments. The terminal sugar in the sugar chain at C-3 was identified as glucose based on the C–H long-range correlation observed in the HMBC spectrum (Fig. S2) between H-1‴ of the terminal β-D-glucose (δ_H 4.36) and C-6″ of the middle β-D-glucose (δ_C 69.47), and further confirmed by comparing these data with those reported by our previous work [25].

The α-configuration of H-17 and β-orientation of the side chain at C-17 was deduced in similar way as in **1** (chemical shift and coupling constant). From the above-mentioned data, compound **5** was identified as *3β*,14β-dihydroxypregn-5-en-20-one-3-*O*-β-D-glucopyranosyl-(1 → 6)-*O*-β-D-glucopyranosyl-(1 → 4)-*O*-β-D-digitalopyranoside (named as arabincoside B, a new natural product).

Compound **2** was obtained as a white powder (110 mg), ESI–MS analysis showed an adduct ion peak at *m/z* (rel. int.): 927 $[M+Na]^+$ (100) in positive mode, and 903.1 $[M-H]^-$ (88) in negative mode assigned for the molecular formula $C_{48}H_{72}O_{16}$. IR spectrum of **2** showed absorption bands due to hydroxyl (3417 cm^{-1}), carbonyl (1687 cm^{-1}), and olefinic double bond (1635, 891 cm^{-1}) groups. The ^{1}H and ^{13}C NMR spectra of compound **2** (Tables 1–4) showed the typical signals of a pregnane skeleton, and its aglycone had two major structural differences when compared to that of compounds **1** and **5**, indicated by the presence of an acyl group at C-12, and free hydroxy group at C-20 instead of the carbonyl group of C-20. The signal at δ_C 70.36 was assigned for hydroxylated C-20 with its proton appeared as mutiplet at δ_H 3.82. This was further supported by the upfield shift of the carbon signal C-17 (δ_C 62.45 in **1** to δ_C 52.18 in **2**) and the methyl signal of C-21 (from δ_C 31.29 in **1** to δ_C 21.69 in **2**). The acyl group was identified as benzoyl unit from ^{1}H and ^{13}C NMR spectra [24,25] positioned at C-12 (δ_C 78.32) and was confirmed from the downfield shift of their corresponding protons and carbons signals, and from the HMBC correlations (Fig. S3) between the signal of the ester carbonyl (δ_C 166.55) of the benzoyl moiety and H-12 (δ_H 4.82).

The α-configuration of H-17 and β-orientation of side chain at C-17 was deduced in similar way as in **1**. Regarding the absolute configuration of C-20, the ^{13}C chemical shift values of adjacent carbons in **2** were compared with those reported in literature for various 20*R*- and 20*S* pregnane compounds, which were synthesized by the reduction of 20-pregnanons and also have hydroxyl group at C-12 [26]. Remarkable differences in the ^{13}C chemical shifts values for C-16 and C-20 was observed in the two sets of epimers. Based on this observation, the configuration of C-20 of compound **2** was deduced to be *R*-configuration, as reported in





1 R= Glc-(1→4)-Dig

2 R_1= Glc-(1→4)-Cym-(1→4)-Cym, R_2= *O-Bz*

3 R_1= Glc-(1→4)-Dig, R_2= H

4 R_1= Glc-(1→4)-Glc-(1→4)-Cym-(1→4)-Cym, R_2= *O-Bz*

5 R= Glc-(1→6)-Glc-(1→4)-Dig

6 R_1= Glc-(1→6)-Glc-(1→4)-Dig, R_2= H

Fig. 1. Structures of the isolated compounds.

related pregnane glycosides [24,25,27].

Analysis of the ^{1}H and ^{13}C NMR spectra of **2** revealed the presence of three anomeric protons and carbon signals, and suggesting that **2** is a trisaccharide glycoside. The identity of sugar chain, the sequence, assignments and linkage sites were determined to be *β*-D-glucopyranosyl-(1 → 4)-*β*-D-cymaropyranosyl (1 → 4)-*β*-D-cymaropyranoside by comparing its spectral data with those of russelioside F, previously isolated by our group from *Caralluma russeliana* [28]. This assumption was clearly confirmed from TOCSY and HMBC spectra. Thus, compound **2** was identified as 12-*β*-*O*-benzoyl-3*β*, 12*β*, 14*β*, 20-tetrahydroxy-(20*R*)-pregn-5-ene-3-*O*-*β*-D-glucopyranosyl-(1 → 4)-*β*-D-cymaropyranosyl-(1 → 4)-*β*-D-cymaropyranoside (named as arabincoside C, a new natural product).

Compound **4** was obtained as a white powder (15 mg), ESI−MS analysis showed an adduct ion peak at *m*/*z* (rel. int.): 1089.1 $[M+Na]^{+}$ (100) in positive mode, and 1065.0 $[M-H]^{-}$ (100), 1101.0 $[M+Cl]^{-}$ (18) in negative mode, assigned for the molecular formula $C_{54}H_{82}O_{21}$. IR spectrum of **4** showed absorption bands due to hydroxyl (3420 cm^{-1}), carbonyl (1681 cm^{-1}), and olefinic double bond (1637, 882 cm^{-1}) groups. The ^{1}H and ^{13}C NMR data of the aglycone moiety of **4** were undistinguished from those of compound **2** and identifying the aglycone as 12-*β*-*O*-benzoyl-3*β*, 12*β*, 14*β*, 20-tetrahydroxy-(20*R*)-pregn-5-ene. The only difference was the presence of an additional sugar moiety identified as *β*-D-glucose (^{1}H and ^{13}C NMR spectra) attached to the sugar chain of **2** at C-4‴. The presence of four anomeric protons and carbons signals indicated the presence of a tetrasaccharide chain. The sequential assignments of proton and carbon resonances of the sugar moieties, as well as their connectivity to each other, were determined using HMBC experiment. The attachment of the additional glucose moiety at C-4‴ was deduced from the C−H long-range correlation observed in the HMBC spectrum (Fig. S4) between H-1⁗ of the terminal *β*-D-glucose (δ_H 4.41) and C-4‴ of the middle *β*-D-glucose (δ_C 74.91). The *α*-configuration of H-17 and *R*-configuration of C-20 were deduced in similar way as in **2**. Therefore, compound **4** could be identified as 12-*β*-*O*-benzoyl-3*β*, 12*β*, 14*β*, 20-tetrahydroxy-(20*R*)-pregn-5-ene-3-*O*-*β*-D-glucopyranosyl-(1 → 4)-*β*-D-glucopyranosyl-(1 → 4)-*β*-D-cymaropyranosyl-(1 → 4)-*β*-D-cymaropyranoside (named as arabincoside D, a new natural product).

3. Conclusion

A survey of the literature revealed that no detailed chemical investigation of *C. Arabica*, an important plant in the traditional medicine of Yemen. Therefore, the phytochemical investigation of the aqueous remaining fraction after extraction with ether and methylene chloride of the aerial parts of *C. arabica* afforded six pregnane glycosides, including four new ones (arabincosides A-D) and two known pregnane glycosides (russelioside C and D). The chemical structures of **1**–**6** were elucidated by extensive 1D NMR and 2D NMR and mass analyses. The two known pregnane glycosides were isolated previously from *C. russelliana.*

4. Experimental section

4.1. General experimental procedures

Optical rotations were measured on a Bellingham + Stanley ADP 440 + digital polarimeter (Bellingham & Stanley, Kent, UK). NMR analysis was conducted on Bruker High Performance Avance III FT-NMR spectrometer (^{1}H-NMR: 400 MHz and ^{13}C-NMR: 100 MHz) and using TMS as internal standard. IR spectrophotometer, Shimadzu FT-IR Affinity-1 was used for recording IR spectra using KBr discs. The electron spray ionization-mass spectroscopy (ESI-MS) was performed on PlateExpress TLC plate reader coupled to the Expression compact mass spectrometer (CMS) (Advion, Ithaca, NY, USA). The Advion is a single-quadrupole mass spectrometer that provides electrospray ionization (ESI) in both in the positive and negative ionization. Analytical TLC was carried out on Merck TLC plates KGF Silica gel 60 and KGF RP-18 Silica gel 60 and spots were visualized under UV light (254 and 365 nm) and after spraying with *p*-anisaldehyde/H_2SO_4 followed by heating at 110 °C. Column chromatography (CC) was carried on flash silica gel 60 (Merck, particle size 230−400 mesh), and RP-C18 (silica gel, 40−63 μm; Merck).

4.2. Plant material

The fresh aerial parts of *Caralluma arabica* N. E. Br. [Syn. *Desmidorchis arabica* (N. E. Br.) Meve & Liede; *Crenulluma arabica* (N. E. Br.) Plowes] were collected from Aqan region, Al-Musaimir District, Lahej Governorate, Southern Yemen (13° 22′ 871″ **N,** 045° 83′ 344″ **E**), in July 2020. The plant material was collected and authenticated by Dr. Othman S. S. Al-Hawshabi, Associate Professor of Plant Taxonomy and Flora, Department of Biology, Faculty of Science, Aden University, Yemen. The plant was cut into small pieces and dried in shade. A voucher specimen (# 5659) of the plant was deposited in the Department of Biology, Faculty of Science, University of Aden.

4.3. Extraction and isolation

The powdered aerial parts of *C. arabica* (300 g) was extracted with 95% ethanol (3 × 1.5 L) at room temperature by the aid of Ultraturrax. After filtration, the combined ethanolic extract was evaporated under reduced pressure to yield a solid residue (35 g). Part of the ethanolic extract (30 g) was suspended in water-MeOH (4:1 v/v, 300 mL) and partitioned successively with ether and methylene chloride (3 × 200 mL) to afford 0.6 g and 6 g, respectively, and 21 g of the remaining water fraction (RWF). Part of RWF (8 g) was chromatographed over open column chromatography using flash silica gel 60 (5 × 15 cm) and elution was performed with methylene chloride-methanol-water (10:2:0.1). Sixty fractions (150 mL each) were collected and monitored by TLC. Similar fractions were pooled together to give thirteen fractions (**Fr-A** to **Fr**−M). Fraction **Fr-E** (1 g) was chromatographed over Silica gel RP-18 column (3 × 14 cm) and elution with MeOH−MeCN−H_2O (2:1:1) gave four main subfractions (**Fr-E1 to Fr-E4**). Subfraction **Fr-E3** gave compound **1** (100 mg). Subfraction **Fr-E4** (375 mg) was purified on a silica gel 60 (3 × 14 cm) using CH_2Cl_2−MeOH, 9: 1 as eluting system and 10 mL fraction each, where compound **2** was isolated from subfractions 9−14 (110 mg). **Fr**−**F** (500 mg) was chromatographed on flash silica gel (4 × 15 cm) using CH_2Cl_2−MeOH (5:1) as eluting system and fractions 10 mL each were collected. Subfractions 46−59 (275 mg) were purified on RP-18 silica gel column (1.5 × 25 cm) using MeOH−MeCN−H_2O (5:1:1) to give compounds **3** (30 mg) and **4** (13 mg). Fraction **Fr-M** (1.03 g) was purified by precipitation to give compound **5** (450 mg). Fraction **Fr**−**N** (900 mg) was purified on silica gel 60 column using CH_2Cl_2−MeOH (9: 1) followed by EtOAc-MeOH-H_2O (100:16.5:13.5) to yield compound **6** (105 mg).

4.3.1. Compound **1** (arabincoside A)

White amorphous powder, $[\alpha]^{D}_{21}$ +108.11 (*c*. 0.07, MeOH); IR υ_{max} (KBr, cm^{-1}): 3418, 2935, 1716, 1651, 1454, 1369, 1280, 1165, 1091, 1003, 960, and 713; Tables 1 and 2 for ^{1}H and ^{13}C NMR (400 MHz, 100 MHz, CH_3OH-d_6) of aglycone moiety, respectively; and Tables 3 and 4 for ^{1}H, ^{13}C NMR of sugar moieties, respectively; ESI−MS, *m*/*z* (rel. int.): 677.1 $[M+Na]^{+}$ (100) in positive mode, and 653.1 $[M-H]^{-}$ (66), 689.1 $[M+Cl]^{-}$ (34) in negative mode.

*4.3.2. Compound **5** (arabincoside B)*

White amorphous powder, $[\alpha]^{D}_{21}$ -98.52 (*c*. 0.10, MeOH); IR υ_{max} (KBr, cm^{-1}): 3379, 2935, 1681, 1419, 1361, 1346, 1284, 1176, 1049, 964, 763 and 640; Tables 1 and 2 for ^{1}H and ^{13}C NMR (400 MHz, 100 MHz, CH_3OH-d_6) of aglycone moiety, respectively; and Tables 3 and 4 for ^{1}H, ^{13}C NMR of sugar moieties, respectively; ESI−MS, *m*/*z* (rel. int.): 839.1 $[M+Na]^+$ (100) in positive mode, and 815.0 $[M-H]^-$ in negative mode.

*4.3.3. Compound **2** (arabincoside C)*

White amorphous powder, $[\alpha]^{D}_{21}$ -39.21 (*c*. 0.04, MeOH); IR υ_{max} (KBr, cm^{-1}): 34,187, 2935, 1678, 1635, 1593, 1366, 1280, 1207, 1176, 1107, 1049, 960, 891 and 609; Tables 1 and 2 for ^{1}H and ^{13}C NMR (400 MHz, 100 MHz, CH_3OH-d_6) of aglycone moiety, respectively; and Tables 3 and 4 for ^{1}H, ^{13}C NMR of sugar moieties, respectively; ESI−MS, *m*/*z* (rel. int.): 927 $[M+Na]^+$ (100) in positive mode, and 903.1 $[M-H]^-$ (88), 939 $[M+Cl]^-$ (100) in negative mode.

*4.3.4. Compound **4** (arabincoside D)*

White amorphous powder, $[\alpha]^{D}_{21}$ -21.74 (*c*. 0.09, MeOH); IR υ_{max} (KBr, cm^{-1}): 3417, 2936, 1678, 1636, 1593, 1365, 1207, 1176, 1107, 1049, 960, 891 and 609; Tables 1 and 2 for ^{1}H and ^{13}C NMR (400 MHz, 100 MHz, CH_3OH-d_6) of aglycone moiety, respectively; and Tables 3 and 4 for ^{1}H, ^{13}C NMR of sugar moieties, respectively; ESI−MS, *m*/*z* (rel. int.): 1089.1 $[M+Na]^+$ (100) in positive mode, and 1065.0 $[M-H]^-$ (100), 1101.0 $[M+Cl]^-$ (18) in negative mode.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Essam Abdel-Sattar reports financial support was provided by Academy of Scientific Research & Technology. Partial Funding from the Academy of Scientific Research and Technology, Egypt.

Acknowledgment

This work was funded partially by the Academy of Scientific Research and Technology under the “Egypt Research and Technology Alliances (EG-KTA) Program” (No.: KTA-C2-2.10).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tet.2022.132858.

References

[1] A. Dubaie, A. El-Khulaidi, Ebadi Center for Studies and Publishing Sana'a-Yemen 127, 2005.
[2] J.R.I. Wood, H. Haig-Thomas, Handbook of the Yemen Flora, Royal Botanic Gardens, Kew, 1997.
[3] U. Meve, S. Liede, Taxon 53 (2004) 61−72.
[4] A. Al-Hubaishi, K. Müller-Hohenstein, An Introduction to the Vegetation of Yemen: Ecological Basis, Floristic Composition, Human Influence, Deutsche Gesellschaft für Technische Zusammenarbeit, 1984.
[5] A. Al-Khulaidi, Management Project (SNRMP) II, Sana'a, Yemen, 2013.
[6] P.S. Bruyns, Afr. J. Bot. 76 (2010) 249−251.
[7] M. Adnan, S. Jan, S. Mussarat, A. Tariq, S. Begum, A. Afroz, Z.K. Shinwari, J. Pharm. Pharmacol. 66 (2014) 1351−1368.
[8] N. Qayyum, H. Rani, K.B. Mir, A.Q. Khan, J.Food 2 (2018) 13.
[9] P. Bruyns, H. Jonkers, Bradleya 1993 (1993) 51−69.
[11] A. Al-Khulaidi, in: E. a. U. (Ed.), Management Project (SNRMP II), 2013. Republic of Yemen.
[12] A.G. Miller, M. Morris, Ethnoflora of the Soqotra Archipelago, Royal Botanic Garden Edinburgh, 2004.
[13] M. Zakaria, M. Islam, R. Radhakrishnan, H. Chen, M. Kamil, A. Al-Gifri, K. Chan, A. Al-Attas, J. Ethnopharmacol. 76 (2001) 155−158.
[14] M. Zakaria, M. Islam, R. Radhakrishnan, X. Liu, A. Ismail, M. Kamil, K. Chan, A. Al-Attas, Pharm. Biol. 40 (2002) 225−230.
[15] M. Khasawneh, H.M. Elwy, N.M. Fawzi, A.A. Hamza, A.R. Chevidenkandy, A.H. Hassan, Am. J. Plant Sci. 2014 (2014).
[16] M. Kamil, A. Jayaraj, F. Ahmed, C. Gunasekhar, S. Samuel, K. Chan, M. Habibullah, A.J. Attas, Pharmacol. Pharm. 52 (2005).
[17] M.A. Al-Yahya, E. Abdel-Sattar, E. Guittet, J. Nat. Prod. 63 (2000) 1451−1453.
[18] D. Deepak, A. Khare, M.P. Khare, Phytochemistry 28 (1989) 3255−3263.
[19] V.U. Ahmad, K. Usmanghani, G.H. Rizwani, J. Nat. Prod. 51 (1988) 1092−1097.
[20] L. Lee-Juian, L. Long-Ze, R.R. Gil, G.A. Cordelly, M. Ramesh, B. Srilatha, B. Reddy, A.V.N.A. Rao, Phytochemistry 35 (1994) 1549−1553.
[21] G.H. Rizwani, M. Ahmad, K. Usmanghani, V.U. Ahmad, Spectrosc. Lett. 26 (1993) 1427−1434.
[22] E. Abdel-Sattar, M.A. Al-Yahya, N. Nakamura, M. Hattori, Phytochemistry 57 (2001) 1213−1217.
[23] M.A. Al-Yahya, E. Abdel-Sattar, G.J. Eric, Nat. Prod. 63 (2000).
[24] M.A. Choucry, A.A. Shalabi, A.M. El Halawany, F.S. El-Sakhawy, A. Zaiter, H. Morita, P. Chaimbault, E. Abdel-Sattar, ACS Omega 6 (2021) 18881−18889.
[25] A.A. Shalabi, A.M. El Halawany, M.A. Choucry, F.S. El-Sakhawy, H. Morita, D.-W. Ki, E. Abdel-Sattar, Phytochem. Lett. 36 (2020) 49−57.
[26] M. Kimura, K. Hayashi, H. Narita, H. Mitsuhashi, Chem. Pharm. Bull. 30 (1982) 3932−3941.
[27] S.M. Al-Massarani, S. Bertrand, A. Nievergelt, A.M. El-Shafae, T.A. Al-Howiriny, N.M. Al-Musayeib, M. Cuendet, J.-L. Wolfender, Phytochemistry 79 (2012) 129−140.
[28] E. Abdel-Sattar, A.A. Ahmed, M.-E.F. Hegazy, M.A. Farag, M.A. Al-Yahya, Phytochemistry 68 (2007) 1459−1463.