

1 Article

2 **Flavonoid Glycosides from Endemic Bulgarian**
3 ***Astragalus aitosensis* (Ivanisch.)**4 **Hristo Vasilev**¹, **Samir Ross**², **Karel Šmejkal**^{3,*}, **Petr Maršík**⁴, **Dagmar Jankovská**^{3,*}
5 **Jaroslav Havlík**⁴, **Ondřej Veselý**⁴6 ¹ Department of Pharmacognosy; Faculty of Pharmacy, Medical University - Sofia, 2 Dunav str., 1000 Sofia,
7 Bulgaria; christo.vasilev@yahoo.com (H. V.)8 ² National Center for Natural Products Research, School of Pharmacy, University of Mississippi, 38677
9 Oxford, Mississippi, USA; sross@olemiss.edu (S. R.)10 ³ Department of Natural Drugs, Faculty of Pharmacy, University of Veterinary and Pharmaceutical Sciences
11 Brno, 61242 Brno, Czech Republic; karel.mejkal@post.cz (K. Š.); jandkovskad@vfu.cz (D. J.)12 ⁴ Department of Food Quality and Safety, Czech University of Life Sciences Prague, Kamycka 129, 165 00
13 Prague 6 – Suchbátka, Czech Republic; marsik@af.czu.cz (P. M.); havlik@af.czu.cz (J. H.); veselyo@af.czu.cz
14 (O. V.)

15 * Correspondence: karel.mejkal@post.cz, Tel.: +42-072-424-3643; jankovskad@vfu.cz, Tel. +42-054-156-2832.

16

17 **Abstract:** *Astragalus* is a very interesting plant genus well-known for content of flavonoids,
18 triterpenes and polysaccharides. Its secondary metabolites are described as biologically active
19 compounds showing a number of activities, like immunomodulating, antibacterial, antiviral and
20 hepatoprotective. This inspired us to analyze the Bulgarian endemic *A. aitosensis* (Ivanisch.) to
21 obtain deeper information about its phenolic components. We used extensive chromatographic
22 separation of *A. aitosensis* extract to obtain seven phenolic compounds (1-7), which were identified
23 using combined LC-MS and NMR spectral studies. The 1D and 2D NMR analysis and HR-MS
24 allowed us to resolve the structures of known compounds 5-7 as isorhamnetin-3-O-robinobioside,
25 isorhamnetin-3-O-(2,6-di-O- α -rhamno-pyranosyl- β -galactopyranoside), and alangiflavoside,
26 respectively, and further comparison of these spectral data with available literature helped us with
27 structural analysis of newly described flavonoid glycosides 1-4. These were described in plant
28 source for the first time.

29 **Keywords:** *Astragalus aitosensis*; flavonoid; glycoside.
3031 **1. Introduction**

32 Genus *Astragalus* comprises from more than 2500 species, which makes it the largest genus in
33 the family Fabaceae. *Astragalus* species are widely distributed, mainly in the temperate regions of the
34 Northern hemisphere [1], and they are well-known and widely used as remedies in the traditional
35 folk medicine of different countries. The pharmacological activity of flavonoids, saponins and
36 polysaccharides – found in species from genus *Astragalus* [2] – endue these plants with
37 immunomodulating, antibacterial, antiviral, hepatoprotective and other protective effects [3-5]. 31
38 species represent genus *Astragalus* in Bulgarian flora [4], and many of them are yet unstudied. In the
39 recent years Bulgarian researchers have isolated and reported new tri- and tetraglycosides of
40 flavonols, including some new compounds from the rarely-met group of flavo-alkaloids [3-5]. Many
41 species of *Astragalus* possess in nature the widely-distributed aglycones – kaempferol, quercetin and
42 methylquercetines in their free and glycosidic forms [6].

43 The object of this research - *A. aitosensis* (Ivanisch.) (syn. *Astracantha aitosensis* (Ivanisch.)
44 Podlech, *Astracantha arnacantha* (M. Bieb.) Podlech subsp. *aitosensis* (Ivanisch.) Réer & Podlech), is a
45 Bulgarian endemic plant and yet phytochemically studied very scarcely [6,7]. *A. aitosensis* is a low,

46 spiny, tussock-forming shrub with strongly branched stems (30–50 cm in height) [8]. The plant grows
 47 on dry stony places (90–550 m alt) with neutral to alkaline soil. It is distributed in the suburbs of the
 48 small Bulgarian town Aytos, which gives rise to its name.

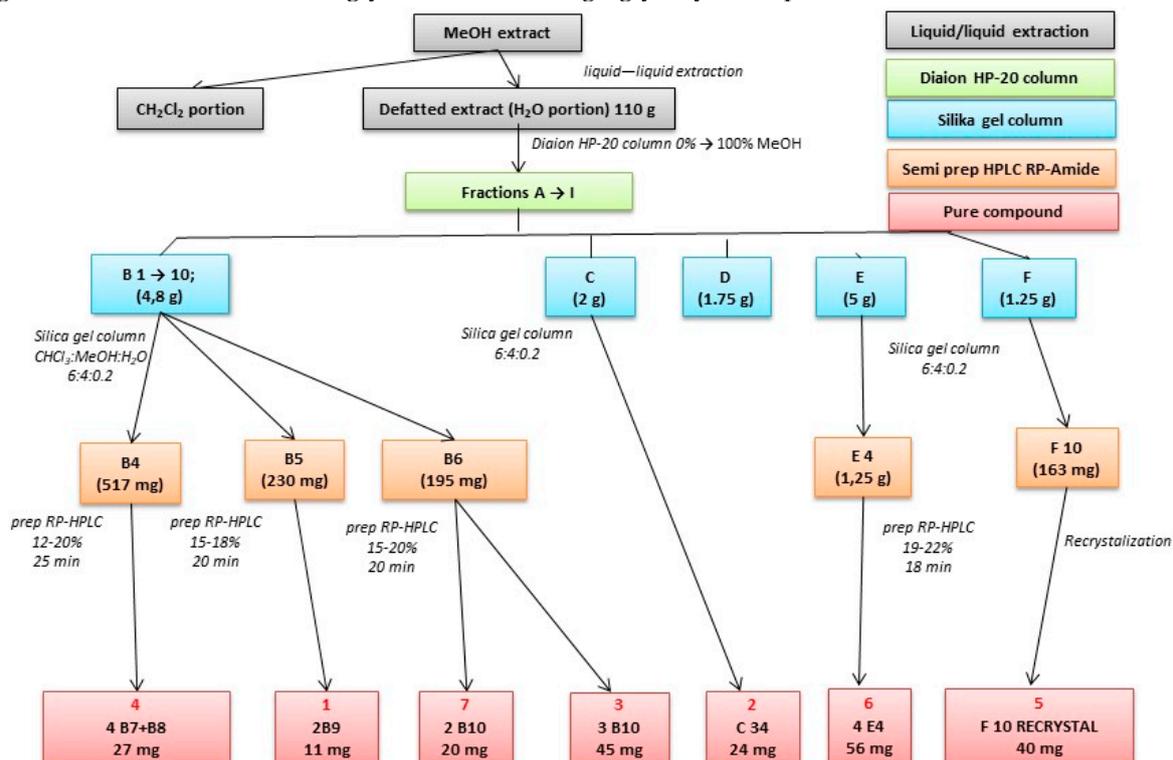
49 Aim of this study was to isolate compounds from the methanolic extract from aerial parts of *A.*
 50 *aitosensis* and perform their structural elucidation using ^1H , ^{13}C , COSY, HSQC, HMBC, NOESY and
 51 TOCSY NMR experiments. HR-ESI-MS was used for additional confirmation of the structures
 52 revealed by NMR. We report here the isolation and structural elucidation of six isorhamnetin and
 53 one kaempferol glycosides 1–7, and four of them are new natural glycosides: three tetra- (1-3) and
 54 one triglycoside (4), as well, another three already known structures were determined for the first
 55 time in *A. aitosensis*: with two (5), three (6) and four (7) sugar units, respectively.

56 2. Results

57 2.1. Isolation of compounds

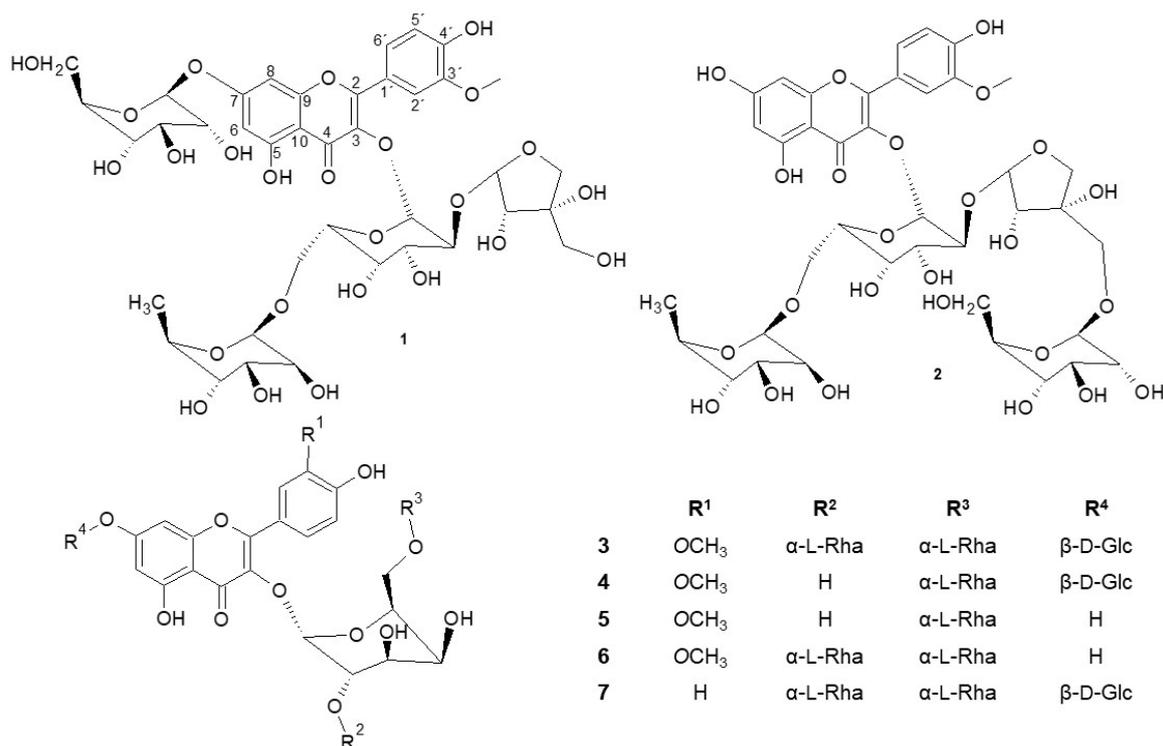
58 Aerial parts of *A. aitosensis* were extracted with 80% methanol under reflux. The preliminary
 59 chromatographic analysis showed a bunch of signals of flavonoid compounds, with retention times
 60 predicting a high degree of glycosylation.

61 The crude extract was therefore defatted by liquid–liquid partitioning with chloroform and
 62 further fractionated *via* successive column chromatography with final step of semi-preparative HPLC
 63 purification of the isolated compounds (Figure 1). Their nature (UV spectral properties) and the
 64 behavior of these compounds during separation on reversed phase – polar character of compounds –
 65 gave us the idea of flavonoid glycosides with a high glycosylation pattern.



66 **Figure 1.** Simplified scheme of separation.

67 Our separation procedures resulted in isolation of seven pure compounds (1-7) (Figure 2).
 68
 69



70
71 **Figure 2.** Structures of isolated compounds 1-7.
72

73 **2.2. Structural analysis**

74 To the best of our knowledge, spectral data of four of isolated compounds (1-4) did not
75 correspond to the data of compounds previously published in literature. Because the 1D and 2D NMR
76 analysis, HR-MS and comparison with literature allowed us to resolve the structures of compounds
77 5-7 as isorhamnetin-3-*O*-robinobioside [9], isorhamnetin-3-*O*-(2,6-di-*O*-α-rhamno-pyranosyl-β-
78 galactopyranoside) [10,11], and alangiflavoside [12], respectively, further comparison of these
79 spectral data with available literature helped us with structural analysis of newly described flavonoid
80 glycosides 1-4.

81 The aglycones for compounds (1-6) were determined based on HR-ESI-MS and NMR (¹H, ¹³C,
82 COSY, HSQC and HMBC) spectral analysis. ¹H and ¹³C spectra are shown in Table 1. HMBC spectra
83 showed the following significant correlations: proton at C-2' was a doublet with *meta* coupling and
84 in the HMBC it showed strong correlation to C-4' and weak to C-3'. The proton at C-6' was observed
85 as doublet of doublet (*ortho* and *meta* coupling) and showed strong correlation in the HMBC to C-4'.
86 The proton at C-5' was a doublet with *ortho* coupling, displaying in the HMBC the strong correlation
87 to C-3' and weak to C-4'. The methoxy group showed correlation to C-3', and therefore, the position
88 of the methoxy is at C-3' and the aglycone of compounds 1-6 was finally identified as 3'-*O*-
89 methylquercetin, syn. isorhamnetin [13].

90 Because of the identification procedure, we describe the elucidation of structures of known
91 compounds prior to the new compounds. Detailed ¹H and ¹³C chemical shifts for compound 5 are
92 listed in Table 2 and Table 3, respectively. The LC-MS analysis showed single peak (*t_R* 7.95 min)
93 chromatogram with a signal of a deprotonated molecule at *m/z* [M-H]⁻ 623.16205, which showed a
94 good correlation with the calculated value of *m/z* 623.16176 for C₂₈H₃₁O₁₆⁻ (Δ= 0.00029). The (+)HRMS
95 supported the idea of this molecular formula, the glycosylation pattern was predicted to be the
96 rhamnose and a hexose from MS/MS analysis of *m/z* [M+H]⁺ 625.1175 and observed fragments *m/z*
97 [M+H-rhamnose]⁺ 479.1184, *m/z* [M+H-rhamnose-oxygen]⁺ 463.1213, and finally *m/z* 317.0645
98 showing the aglycone. The molecular formula C₂₈H₃₂O₁₆ (calcd 624.54408) plays accord with the NMR
99 and further HR-ESI-MS spectral data and data of compound 5 are corresponding to the reported data
100 for isorhamnetin-3-*O*-robinobioside [9].

101
102**Table 1.** ^{13}C NMR δ_{C} (ppm) (100 MHz); ^1H NMR (600 MHz for 1 and 500 MHz for 2). δ_{H} (ppm), multiplicity (J in Hz)

	1		2	
	δ_{C} , type	δ_{H} (J in Hz)	δ_{C} , type	δ_{H} (J in Hz)
2	157.8, C		157.2, C	
3	133.6, C		133.3, C	
4	178.0, C		177.8, C	
5	161.3, C		161.6, C	
6	99.4, CH	6.45, d (2.02)	98.5, CH	6.19, d (2.00)
7	163.0, C		164.4, C	
8	94.4, CH	6.75, d (2.02)	93.4, CH	6.38, d (2.00)
9	156.5, C		157.0, C	
10	106.3, C		104.5, C	
1'	121.8, C		121.9, C	
2'	113.3, CH	8.05, m (1.94)	113.7, CH	8.04, d (1.97)
3'	147.0, C	-OCH ₃	147.0, C	-OCH ₃
4'	149.4, C	-OH	149.2, C	-OH
5'	114.5, CH	6.91, m (8.39)	114.5, CH	6.90, d (8.40)
6'	122.3, CH	7.62, m (1.92; 8.45)	122.1, CH	7.57, dd (1.90; 8.43)
-OCH ₃	55.7, CH	4.00 s	55.7, CH ₃	3.98, s
3-O-gal				
1	99.8, CH	5.57, d (7.84)	99.9, CH	5.58, d (7.76)
2	75.1, CH	3.96, dd	74.8, CH	3.97, dd
3	73.9, CH	3.73 dd	73.9, CH	3.74 dd
4	69.0, CH	3.77 dd	69.0, CH	3.77 dd
5	74.1, CH	3.67 dt	74.0, CH	3.64 dt
6	66.0, CH ₂	3.47/3.68 dd	65.8, CH ₂	3.47/3.71 dd
Api (1→2)				
1	109.2, CH	5.43, d (1.55)	109.1, CH	5.44, d (1.73)
2	76.6, CH	4.01, d	76.9, CH	4.07, d
3	79.5, C	-OH	78.8, C	-OH
4	74.2, CH ₂	3.64, d (9.58)	74.3, CH ₂	3.66, d (n/a)
		3.75, d (9.59)		4.07, d (9.70)
5	65.1, CH ₂	3.64, d (11.45)	61.3, CH ₂	3.67, d (11.95)
		3.75 d (11.56)		3.85 d (11.90)
Rha (1→6)				
1	100.5, CH	4.50, d (1.56)	100.5, CH	4.52, d (1.53)
2	70.7, CH	3.51, dd	70.7, CH	3.57, dd
3	70.9, CH	3.46, dd	70.9, CH	3.49, dd
4	72.4, CH	3.25, pt	72.4, CH	3.26, pt

5	68.3, CH	3.49, dq	68.3, CH	3.51, dq
6	16.5, CH ₃	1.15, d (6.17)	16.5, CH ₃	1.16, d (6.23)
Glc	7-O-Glc		Api (5→1) Glc	
1	100.1, CH	5.06, d (7.53)	103.4, CH	4.25, d (7.54)
2	73.3, CH	3.49, dd	73.6, CH	3.17, dd
3	77.0, CH	3.55, dd	76.3, CH	3.21, dd
4	69.9, CH	3.38, dt	70.4, CH	3.27, dd
5	76.4, CH	3.49, dq	76.3, CH	3.27, dt
6	61.1, CH ₂	3.69/3.94 dd	72.9, CH ₂	3.68/4.15, dd

103

104

ESI- HRAM spectrum of compound **6** showed presence of signal of deprotonated molecule m/z

[M-H]⁻ 769.21948 (calcd for C₃₄H₄₁O₂₀⁻ 769.219667 m/z ; Δ = 0.00019). The MS/MS analysis (a minimal

fragmentation using 20 eV) in a positive mode showed the presence of a protonated parental ion m/z

[M+H]⁺ 787.2284 and fragments corresponding to loss of two rhamnose units and a hexose m/z

641.1708, 479.1180 and 317.0653. The molecular formula C₃₄H₄₂O₂₀ (Mr 770.68528) is in accordance

with the NMR and HR-ESI-MS spectral data and ¹H and ¹³C chemical shifts for compound **6** are listed

in Table 2 and 3, respectively. Compound **6** was therefore identified as isorhamnetin-3-O-(2,6-di-O-

α -rhamno-pyranosyl- β -galactopyranoside) [10,11].

The molecular formula C₃₉H₅₀O₂₄ (Mr 902.7999) is in accordance the NMR and HR-ESI-MS

spectral data, where the mass of deprotonated molecule was found at m/z [M-H]⁻ 901.2624 (calcd

901.2619; Δ =0.0005). The detailed MS analysis of fragments observed in spectrum showed signal m/z

[M+H]⁺ 903.2759, m/z [M+H-rhamnose]⁺ 757.2178, m/z [M+H-rhamnose-rhamnose]⁺ 611.1601, m/z

[M+H-rhamnose-rhamnose-hexose]⁺ 449.1078, and aglycone m/z 287.0549, with pairs of fragments

showing loss of water 18 m/z [M+H-rhamnose-H₂O]⁺ 741.2245, m/z [M+H-rhamnose-rhamnose-H₂O]⁺

595.1660. Spectrum showed also fragments of sugars m/z 309.1186 (hexose-rhamnose). ¹H and ¹³C

chemical shifts for compound **7** are listed in table 2 and 3 respectively. NMR data of compound **7** and

mass spectral analysis results show a good accordance with the data already reported for

alangiflavoside [12].

Based on the previously described analysis of known compounds, we tried to identify the other

isolated compounds, which showed differences from that previously described in literature. Some of

the fragmentation MS/MS results used for the identification of compounds **1-4** are depicted in Figure

3.

126

127

Table 2: ¹³C NMR assignments (100 MHz) for compounds **3-7**, δ (ppm)

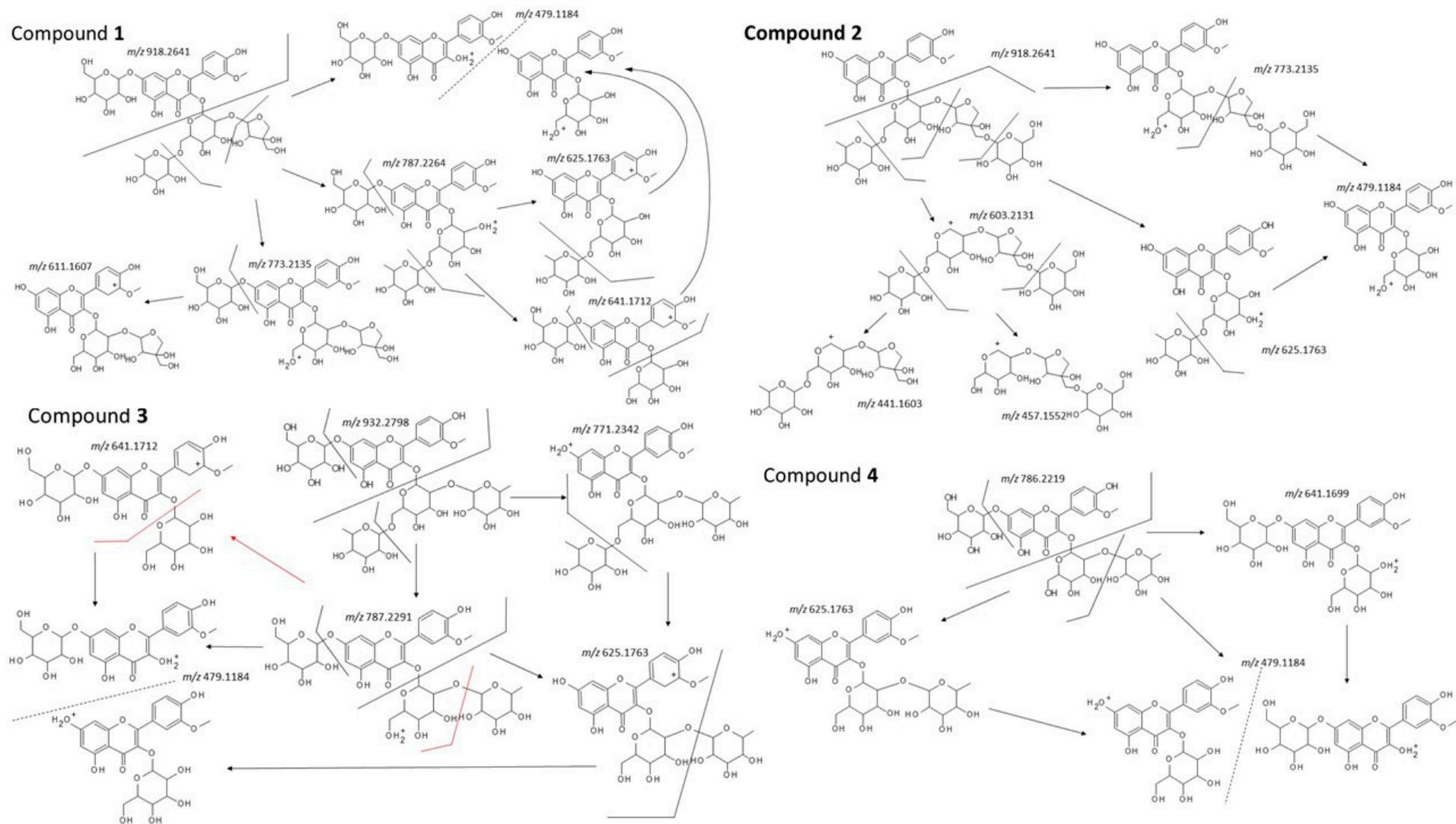
Compound/ position	3	4	5	6	7
	δ_c , type				
2	157.7, C	157.4, C	156.8, C	157.0, C	158.0, C
3	133.3, C	133.8, C	133.5, C	133.0, C	133.4, C
4	177.9, C	178.0, C	177.8, C	177.8, C	177.1, C
5	161.3, C	161.3, C	161.6, C	161.7, C	161.4, C
6	99.3, CH	99.8, CH	99.2, CH	98.4, CH	99.4, CH
7	163.0, C	163.4, C	164.7, C	164.4, C	163.0, C
8	94.4, CH	95.1, CH	94.2, CH	93.3, CH	94.3, CH

9	156.5, C	156.4, C	156.8, C	157.0, C	156.6, C
10	106.3, C	106.1, C	104.4, C	104.5, C	106.2, C
1'	121.8, C	121.3, C	121.5, C	122.0, C	121.5, C
2'	113.3, CH	113.9, CH	113.9, CH	113.3, CH	131.0, CH
3'	147.1, CH	147.5, CH	147.4, C	147.0, C	114.8, CH
4'	149.3, C	150.1, C	149.9, C	149.1, C	160.1, C
5'	114.6, CH	115.6, CH	114.6, CH	114.5, CH	114.8, CH
6'	122.1, CH	122.6, CH	122.4, CH	121.8, CH	131.0, CH
OCH ₃	56.37, CH ₃	56.39, CH	56.37, CH	55.79, CH ₃	
3-O-Gal					
1	99.3, CH	102.1, CH	102.3, CH	99.4, CH	99.4, CH
2	76.4, CH	74.1, CH	74.0, CH	76.4, CH	76.1, CH
3	74.2, CH	71.6, CH	71.5, CH	74.2, CH	74.3, CH
4	69.2, CH	68.4, CH	68.4, CH	69.1, CH	69.4, CH
5	74.1, CH	73.4, CH	73.4, CH	73.9, CH	74.1, CH
6	65.9, CH ₂	65.6, CH ₂	65.6, CH ₂	65.7, CH ₂	66.0, CH ₂
Rha (1→2)					
1	101.4, CH			101.3, CH	101.2, CH
2	71.0, CH			71.0, CH	71.0, CH
3	71.0, CH			71.0, CH	70.9, CH
4	72.5, CH			72.5, CH	72.7, CH
5	68.4, CH			68.4, CH	68.4, CH
6	16.0, CH ₃			16.0, CH ₃	16.1, CH ₃
Rha (1→6)					
1	100.6, CH	100.5, CH	100.5, CH	100.5, CH	100.5, CH
2	70.7, CH	70.9, CH	70.9, CH	70.7, CH	70.7, CH
3	70.9, CH	71.1, CH	71.1, CH	70.9, CH	70.9, CH
4	72.4, CH	72.2, CH	72.3, CH	72.6, CH	72.5, CH
5	68.3, CH	68.7, CH	68.7, CH	68.3, CH	68.3, CH
6	16.6, CH ₃	18.3, CH ₃	18.3, CH ₃	16.6, CH ₃	16.6, CH ₃
7-O-Glc					
1	100.1, CH	100.3, CH			100.1, CH
2	73.3, CH	73.6, CH			73.3, CH
3	77.0, CH	76.9, CH			77.0, CH
4	69.9, CH	70.0, CH			69.9, CH
5	76.4, CH	77.7, CH			76.4, CH
6	61.1, CH ₂	61.2, CH ₂			61.1, CH ₂

Table 3: ¹H NMR assignments (600 MHz) of compounds 3-7 δ_H (ppm), multiplicity (*J* in Hz)

	3	4	5	6	7
	δ _H (<i>J</i> in Hz)				
2					
3					
4					
5					
6	6.44, d (2.07)	6.44, d (1.91)	6.19, d (1.97)	6.15, d (1.97)	6.46, d (2.15)
7					
8	6.75, d (2.09)	6.77, d (1.97)	6.42, d (1.97)	6.37, d (1.97)	6.75, d (2.14)
9					
10					
1'					
2'	8.09, m (1.93)	8.00, d (1.84)	7.98, d (1.97)	8.07, d (1.94)	8.09, m (8.95)
3'					6.90, m (8.90)
4'					
5'	6.91, m (8.44)	6.90, d (8.39)	6.88, d (8.44)	6.90, d (8.42)	6.90, m (8.90)
6'	7.57, m (1.91; 8.42)	7.54, dd (1.82; 8.48)	7.49, dd (2.03; 8.39)	7.52, dd (1.99; 8.43)	8.09, m (8.95)
OCH3	3.83, s	3.84, s	3.83, s	4.00, s	-
3-O-Gal					
1	5.79, d (7.84)	5.47, d (7.73)	5.45, d (7.67)	5.59, d (7.83)	5.59, d (7.74)
2	3.97, dd	3.60, dd	3.57, dd	3.96, dd	3.95, dd
3	3.78, dd	3.57, dd	3.57, dd	3.76, dd	3.70, dd
4	3.80, dd	3.62, dd	3.63, dd	3.81, dd	3.75, dd
5	3.74, dt	3.42, dt	3.41, dt	3.71, dt	3.64, dt
6	3.54/3.73dd	3.30/3.60 dd	3.31/3.61 dd	3.52/3.72 dd	3.46/3.69 dd
Rha (1→2)					
1	5.16, d (1.52)			5.16, d (1.53)	5.21, d (1.32)
2	4.00, dd			4.00, dd	4.00, dd
3	3.75, dd			3.76, dd	3.78, dd
4	3.33, pt			3.32, pt	3.33, pt
5	4.04, dq			4.03, dq	4.06, dq
6	0.89, d (6.25)			0.89, d (6.23)	0.98, d (6.23)
Rha (1→6)					
1	4.53, d (1.53)	4.41, d (1.53)	4.41, d (1.53)	4.55, d (1.55)	4.49, d (1.24)
2	3.54, dd	3.37, dd	3.38, dd	3.59, dd	3.48, dd
3	3.49, dd	3.28, dd	3.28, dd	3.51, dd	3.46, dd
4	3.25, pt	3.06, pt	3.07, pt	3.26, pt	3.25, pt
5	3.52, dq	3.36, dq	3.34, dq	3.54, dq	3.49, dq
6	1.16, d (6.08)	1.04, d (6.18)	1.03, d (6.25)	1.17, d (6.22)	1.16, d (6.21)

7-O-Glc			
1	5.06, d (7.24)	5.05, d (7.89)	5.07, d (7.45)
2	3.48, dd	3.24, dd	3.48, dd
3	3.53, dd	3.27, dd	3.53, dd
4	3.39, dd	3.15, dd	3.38, dd
5	3.50, dt	3.43, dt	3.49, dt
6	3.69/3.92, dd	3.44/3.68, dd	3.69/3.91, dd



132

133

Figure 3. A fragmentation pattern of flavonoid glycosides 1-4 (shown *m/z* are calculated values according to molecular formulas of the ions).

134

135 Compound **1** was isolated as a yellow amorphous powder. The UV spectra analysis showed a
136 typical flavonoid course, with the retention time typical for glycosylated substances. The HR-ESI-MS
137 in a negative mode displayed a molecular ion m/z $[M-H]^-$ 917.2572 (calc. for $[M-H]^-$ 917.256841),
138 confirming the molecular formula $C_{39}H_{50}O_{25}$ with calculated molecular mass 918.7993 Da. HRESIMS
139 showed an adduct formation m/z $[M+HCOO]^-$ 963.26269 (confirming the analyte), and a hexose (m/z
140 162.0235) loss leading to m/z 755.21208 (possible representing cleavage of glucose from 7-*O* position
141 or galactose from 3-*O*- position of expected flavonol skeleton). Furthermore, the MS in a positive
142 mode confirmed this loss showing precursor ion m/z $[M+H]^+$ 919.2705 with product ions formed
143 either by subsequent losses of a pentose (m/z 787.2264), a rhamnose (m/z 625.1750), and a hexose
144 (479.1170) unit, either loss of rhamnose (m/z 773.2136) followed by hexose (m/z 611.1613) moieties (see
145 Figure 3). After lining up interpretation of NMR spectra (Table 1) with MS analysis, the structure of
146 the compound **1** was elucidated to be the isorhamnetin substituted by four sugars. One of them is a
147 6-deoxyhexose (a methyl group as doublet), two are hexoses (a glucose, a galactose), and one is a
148 pentose. DEPT spectrum of compound **1** showed 4 methylene groups. Altogether, two of them belong
149 to each hexose (the glucose and galactose), respectively, and since a 6-deoxyhexose (expected
150 rhamnose) does not have a methylene moiety, the two left CH_2 moieties must belong to a pentose.
151 Four anomeric protons are found in HMBC spectrum: δ_H 5.57 ppm, (1H, d, $J=7.84$ Hz), δ_H 5.43 ppm
152 (1H, d, $J=1.55$ Hz), δ_H 4.50 ppm (1H, d, $J=1.56$ Hz) and δ_H 5.06 ppm (1H, d, $J=7.53$ Hz), corresponding
153 to carbon atoms with δ_C 99.8 ppm, δ_C 109.2 ppm, δ_C 100.5 ppm and δ_C 100.1 ppm from the HSQC
154 spectrum. After complete resonance assignments, analyses of coupling constants, intensities,
155 interpretation of cross-peaks in the COSY spectrum, and ^{13}C NMR chemical shift values, one hexose
156 moiety was identified as a β -glucosyl unit, the second as a β -galactosyl moiety, the 6-deoxysugar was
157 identified to be the α -rhamnose, and the pentose was recognized as the β -apiose, which contains two
158 of the above-mentioned methylene groups. ^{13}C values of C-6 indicate that β -glucose residue is not
159 connected to other sugar unit (δ_C 61.1 ppm), while the β -galactose moiety is connected at C-6 position
160 (δ_C 66.0 ppm – shifted to a higher field). HMBC correlations allowed us to elucidate the precise
161 structure of the sugar chains and the positions of their attachment to the aglycone. The anomeric
162 proton (δ_H 5.06 ppm) of the β -glucose moiety showed a three-bond correlation to C-7 (δ_C 163.0 ppm)
163 of the aglycone, while a HMBC correlation between the anomeric proton of the galactose (δ_H 5.57
164 ppm) and the carbon at δ_C 133.6 ppm indicated that the galactosyl unit is connected at C-3 towards
165 the aglycone. The anomeric proton of α -rhamnose correlated to C-6 in the β -galactose (δ_C 66.0 ppm).
166 The anomeric atom of the last pentose sugar (apiose) δ_H 5.43 ppm bonded to δ_C 109.2 ppm showed a
167 HMBC correlation with C-2 of the β -galactose molecule (δ_C 75.1 ppm). The remaining two methylene
168 groups (δ_C 74.2 ppm and δ_C 65.1 ppm) were recognized as carbons C-4 and C-5 in the apiosyl moiety.
169 Hence, according to all interpreted spectra, we identified the compound **1** as the new natural product
170 named isorhamnetin-3-*O*-[β -D-apiofuranosyl-(1 \rightarrow 2)-[α -L-rhamnopyranosyl-(1 \rightarrow 6)]- β -D-
171 galactopyranosyl]-7-*O*- β -D-glucopyranoside.

172 Compound **2** was obtained as a yellow amorphous powder. The spectral analysis showed data
173 very similar to that observed for compound **1** (Table 1), and the only difference found was the
174 presence of a β -glucosyl residue connected to the sugar chain, attached to the 3-*O* position
175 (particularly C-5 of apiosyl residue), a connection to the 7-*O* position of the aglycone was not found.
176 Further interpretation of HRESIMS supported this suggestion by observing a lacking of the fragment
177 of a hexose loss (m/z 162.0235), which can be observed in spectra of all compounds possessing glucose
178 attached at 7-*O* position (compounds **1**, **3**, **4** and **7**). The HR-ESI-MS in negative mode displayed a
179 molecular ion m/z $[M-H]^-$ 917.25696 (calculated for m/z $[M-H]^-$ 917.256841), confirming molecular
180 formula $C_{39}H_{50}O_{25}$ with calculated molecular mass 918.7993 Da. Adduct formation $[M+HCOO]^-$ (m/z
181 953.2336) and $[M+Cl]^-$ (m/z 963.2621) was observed in a negative ESI mode. Furthermore, MS in a
182 positive mode confirmed this by $[M+H]^+$ with m/z 919.2709, and a series of corresponding losses of
183 four sugar units at a minimal fragmentation of 20 eV: m/z 773.2136 as loss of rhamnosyl unit, m/z
184 625.1759 loss of rhamnose and hexose, and 479.1185 corresponding to isorhamnetin hexoside after
185 loss of three sugar units of rhamnose, hexose and pentose. Except them, also the aglycone signal m/z

186 317.0655 and the fragment ion m/z 603.2144 interpreted as the chain of all sugar moieties with
187 corresponding fragments m/z 441.1610 and m/z 457.1552 after losses of terminal hexose and rhamnose,
188 respectively, were detected. According to differences in chemical shifts in NMR spectral data
189 (compared with **1**), the glucosyl moiety changed its bonding position from 7-*O* position of the
190 aglycone to C-5 position of the β -apiosyl residue (see Table 3). The compound was therefore identified
191 as isorhamnetin-3-*O*-[β -D-glucopyranosyl-(1 \rightarrow 5)- β -D-apiofuranosyl]-(1 \rightarrow 2)-robinobioside or [α -L-
192 rhamnopyranosyl-(1 \rightarrow 6)]- β -D-galactopyranosyl], a newly described flavonoid glycoside.

193 Compound **3** was isolated as a yellow amorphous powder. Rt of compound **3** was 4.61 minutes,
194 again slightly different from other isolates. The HR-ESI-MS in a negative mode displayed a molecular
195 ion m/z [M-H] $^-$ 931.27352 (calcd [M-H] $^-$ 931.272491), confirming molecular formula C₄₀H₅₂O₂₅ with
196 calculated molecular mass 932.82588 Da. HRESIMS showed additional adduct formation at m/z
197 [M+HCOO] $^-$ 977.27862, and a hexose loss (m/z 162.05387) leading to m/z 769.21965 (representing
198 supposed cleavage of glucose from 7-*O* position). In the HRAM ESI positive mode, [M+H] $^+$ ion with
199 m/z 933.2810 (calcd [M+H] $^+$ 933.2876) was obtained. Fragmentation with collision energy of 20 eV
200 gave ions with m/z 787.2229 and m/z 771.2324 formed by loss of hexose and rhamnose, respectively.
201 Further subsequent losses of the hexosyl and rhamnosyl moieties resulted fragments with m/z 625.1679
202 (aglycone-rhamnosyl-hexoside) and m/z 641.1652 (aglycone-hexosyl-hexoside). Finally, ions of m/z
203 479.1082 (hexosylated aglycone) and aglycone m/z 317.0526 (supposed isorhamnetin) were found.
204 Together with an interpretation of NMR spectra (Table 2 and 3), the structure was predicted to be
205 composed of isorhamnetin and four sugar moieties. Two of them were recognized as 6-deoxyhexoses
206 (two methyl groups in the form of the doublet), and the other two were hexoses (possible a glucose,
207 or a galactose). DEPT spectrum showed their two methylene groups (δ_c 61.1 ppm and 65.9 ppm) that
208 belonged to the galactose and the glucose, respectively. We observed four anomeric protons in HSQC
209 spectrum: δ_H 5.79 ppm (1H, d, J = 7.84 Hz), δ_H 5.16 ppm (1H, d, J = 1.52 Hz), δ_H 4.53 ppm (1H, d, J = 1.53
210 Hz) and δ_H 5.06 ppm (1H, d, J = 7.24 Hz), corresponding to carbon atoms δ_c 99.3 ppm, 101.4 ppm,
211 100.6 ppm and 100.1 ppm, respectively. After complete resonance assignments and analyses of
212 coupling constants, intensities of cross-peaks in the COSY spectrum, and ^{13}C NMR chemical shift
213 values, one hexose moiety was identified as a β -glucosyl unit, the other as a β -galactosyl moiety, and
214 the 6-deoxy sugars were found to be α -rhamnosyl moieties. ^{13}C values indicated that β -glucosyl
215 residue is free at C-6 (δ_c 61.1 ppm), while the β -galactose is connected to C-6 (δ_c 65.9 ppm – shifted
216 to a higher field). According to HMBC correlations, structure of the side chains and their attachment
217 to the aglycone were established. Anomeric proton (δ_H 5.06 ppm) of the glucose moiety showed a
218 three-bond correlation to C-7 (δ_c 163.0 ppm) of the aglycone, while an HMBC correlation between
219 the anomeric proton of the β -galactosyl moiety (δ_H 5.79 ppm) and the carbon at δ_c 133.3 ppm indicated
220 that the β -galactosyl unit is bonded at C-3 toward the aglycone. The anomeric proton of one of the α -
221 rhamnose residues (δ_H 5.16 ppm) was correlated to position C-2 of the β -galactose (δ_c 76.4 ppm),
222 while the anomeric proton of the other α -rhamnose showed correlation to C-6 of the β -galactose (δ_c
223 65.9 ppm). Methyl residues of α -rhamnopyranosyl residues were located in the low-field region of
224 ^1H spectrum at δ_H 0.89 ppm (3H, d, J = 6.25 Hz) for the α -rhamnosyl residue attached to C-2 of the β -
225 galactose and at δ_H 1.16 ppm (3H, d, J = 6.08 Hz) for the α -rhamnosyl residue attached to C-6 of the β -
226 galactose, respectively. Hence, we identified the compound **3** as the new natural product
227 isorhamnetin-3-*O*-(2,6-di-*O*- α -rhamnopyranosyl- β -D-galactopyranoside)-7-*O*- β -D-glucopyranoside.
228 The compound is similar to compound **7**, the difference lies in the absence of β -D-glucopyranosyl
229 moiety at 7-*O* position in compound **7**.

230 Compound **4** was isolated as a yellow amorphous powder. Chromatographic analysis showed
231 Rt 5.40 minutes. The HR-ESI-MS in negative mode displayed a molecular ion m/z [M-H] $^-$ 785.21509
232 (calcd for [M-H] $^-$ 785.21403), confirming molecular formula C₃₄H₄₂O₂₁ with calculated molecular mass
233 786.68468 Da. HRESIMS in negative mode showed additional adduct formation m/z [M+HCOO] $^-$
234 831.22089, and a hexose (m/z 162.05404) loss leading to m/z 623.16174 (representing cleavage of
235 glucose from 7-*O* position or a galactose from 3-*O*). The moderate fragmentation in positive ESI (20
236 eV) showed the presence of m/z 787.2284 for parental ion, and then the corresponding fragment with
237 a cleavage of deoxyhexose m/z 641.1714, hexose m/z 625.1753, two sugar (rhamnosyl and hexosyl)

238 units m/z 479.1178 and an aglycone at m/z 317.0658, possible isorhamnetin. After the interpretation of
239 NMR spectra (Table 2 and 3), we confirmed the presence of isorhamnetin and three sugar moieties.
240 Compound **4** possesses similar structure as compound **3**, with absence of the α -rhamnopyranosyl
241 moiety connected to C-2 of β -galactosyl residue, and we identified compound **4** as isorhamnetin-3-
242 O -robinobioside-7- O -glucoside, a new flavonoid glycoside.
243

244 3. Discussion

245 As visible from above described, we isolated 7 flavonoid compounds from *A. aitosensis* extract.
246 Their structures were elucidated by 1D (^1H , ^{13}C) and 2D NMR experiments (COSY, HSQC, HMBC,
247 NOESY and TOCSY) and confirmed by HR-ESI-MS. We report the structures of six isorhamnetin and
248 one kaempferol glycosides (**1–7**), out of which three new tetra- (**1, 2, 3**), one new tri-glycoside (**4**) and
249 three already known compounds with two (**5**), three (**6**) and four (**7**) sugar units, respectively.

250 Genus *Astragalus* is one of the largest genera of Fabaceae family. As mentioned, bioactivity of
251 *Astragalus* plants are connected with a presence of flavonoids, saponins and polysaccharides. The use
252 of *Astragalus* spp. is mainly connected with immunomodulation, antibacterial and antiviral activity,
253 and hepatoprotection [3–6]. The reviews of Gorai et al. [14], Bratkov et al. [5], and Li et al. [15] showed
254 an overview of *Astragalus* genera and flavonoids isolated, showing the presence of flavones,
255 flavonols, flavanones, flavan-4-ols, isoflavones, isoflavans, pterocarpans and others in 60 different
256 *Astragalus* species. Their review included also isorhamnetin and kaempferol glycosides. Another
257 comprehensive review of Bulgarian *Astragalus* species was published in 2016, similarly describing
258 the presence of several kaempferol and isorhamnetin glycosides, including alangiflavoside from *A.*
259 *monspessulanus* ssp. *monspessulanus* [16, 17].

260 The recent years Bulgarian researchers have isolated and reported new tri- and tetraglycosides
261 of flavonols, including some new compounds from the rarely-met group of flavo-alkaloids [3,5].
262 Many species of *Astragalus* possess in nature the widely-distributed aglycones – kaempferol,
263 quercetin and methylquercetines in their free and glycosidic forms [3]. *A. aitosensis* showed
264 previously presence of rutin, quercetin-3- O - β -D-glucoside and astragalin [7]. As visible from
265 comparison of glycosides isolated from *A. aitosensis* with literature, similar compounds – glycosides
266 were obtained for example from *A. monspessulanus* ssp. *monspessulanus*, *A. cicer* and *A. centralpinus*
267 [16–18]. This could from chemotaxonomic point of view confirm their close relationships.
268

269 4. Materials and Methods

270 4.1. General Experimental Procedures.

271 Optical rotations were measured on a JASCO P-2000 spectropolarimeter (Easton, MD, USA) at 20 °C
272 in MeOH with Spectramanager software.

273 HPLC used configuration of analytical system by Agilent 1100 Series (Degasser G1322A, Quaternary
274 Pump G1311A, Autosampler ALS G1313A, Column Compartment G1316, DAD G1315B, Loop 20 μl ,
275 UV spectrum 200–900 nm) with column Kinetex® PFP 100 A, 250 \times 4,6 mm I.D., 5 μm (Phenomenex,
276 USA), flow rate: 1 ml/min. Semi-preparative HPLC was carried out using Dionex UltiMate 3000
277 system (Pump Dionex UltiMate 3000 UPLC+ Focused, Dionex UltiMate 3000 RS Variable Wavelength
278 Detector, fraction collector Dionex UltiMate 3000 with 6 positions, LCO 101 ECOM column oven,
279 constant temperature 40°C, autosampler Dionex UltiMate 3000, loop 100 μl), column: Ascentis® RP-
280 AMIDE, 250 mm \times 10 mm, 5 μm (Supelco, USA), flow rate 5 ml/min. TLC was carried out on
281 precoated silica gel plates (Supelco Kieselgel G, F254, 60, Merck, Darmstadt, Germany) with the
282 solvent systems EtOAc:MeOH:H₂O (100:13.5:10, v/v/v). Spots were visualized under UV light (365
283 nm) after spraying with NTS/PEG reagent. Column chromatography (CC) was performed using
284 Diaion HP-20 (Supelco, USA), \varnothing =80 mm, height 70 cm ~ 700 g and Silica gel (40–63 μm , Sigma-
285 Aldrich®, USA) \varnothing =35 mm, height 60 cm.

286 287 4.2. Plant material

288 Aerial parts of *A. aitosenis* (Ivan.) Podl (Fabaceae) (syn. *Astracantha aitosenis* (Ivan.) Podl.) was
289 collected and identified by Hristo Vasilev in June 2015 in the suburbs of town Aytos, Bulgaria
290 (coordinates Google maps: 42.702191 N, 27.266976 E, UTM: NH22), voucher specimen has been
291 deposited in Herbarium of the Institute of Biodiversity and Ecosystem Research, Bulgarian Academy
292 of Sciences with Ref № SOM001362.

293

294 4.3. Extraction and isolation

295 Dried aerial parts (3600 g) of *A. aitosenis* were extracted under reflux with 80% MeOH (20 × 1.25 L,
296 40 min each) at 70 °C. The total methanol extract was evaporated to dryness (128 g, 3.5 %), and it was
297 re-suspended in H₂O (800 ml) to remove unpolar compounds by liquid/liquid extraction with
298 chloroform (200 ml × 5), giving 18 g of CHCl₃ fraction and 110 g of MeOH soluble material. This
299 defatted methanol portion was subjected to Diaion HP-20 column, and eluted with gradient system
300 of H₂O and MeOH (100:0 to 0:100, v/v), giving 10 combined fractions assigned A-J. Fraction B, C, E,
301 and F were further purified *via* open column chromatography (Silica gel, eluted with
302 CHCl₃:MeOH:H₂O 60:40:2, v/v/v). Fraction B was re-chromatographed on silica gel
303 (CHCl₃:MeOH:H₂O 6:4:0.2, v/v/v, as mobile phase,) and resulted in 6 combined fractions (B₁-B₆).
304 Fraction B₄, was further purified using a semipreparative HPLC (gradient of acetonitrile in 0.2%
305 HCOOH from 12 to 20 % in 25th minute), and this purification gave compound 4 (27 mg). Fraction B₅,
306 after semipreparative HPLC (gradient of acetonitrile in 0.2% HCOOH from 15 to 18 % in 20th minute),
307 gave compound 1 (11 mg). Fraction B₆, was purified on semipreparative HPLC (gradient of
308 acetonitrile in 0.2% HCOOH from 15 to 20 % in 20th minute), gave compounds 7 (20 mg) and 3 (45
309 mg), respectively. Fraction C was re-chromatographed on silica gel (mobile phase 6:4:0.2
310 CHCl₃:MeOH:H₂O, v/v/v) and this separation resulted in 45 fractions. Pure compound 2 (24 mg)
311 precipitated from fraction C₃₄. 300 mg of fraction E was purified with semipreparative HPLC
312 (gradient of acetonitrile in 0.2% HCOOH from 19 to 22 % in 18th minute), and gave compounds 6 (56
313 mg). Fraction F was re-chromatographed on silica gel (mobile phase 6:4:0.2 CHCl₃:MeOH:H₂O,
314 v/v/v), what resulted in 35 fractions 100 ml each. Pure compound 5 (40 mg) yielded from fraction F₁₀
315 after re-crystallization from MeOH.

316

317 4.4. Identification of isolated compounds

318 NMR spectra were recorded on a NMR Agilent DD2 600 MHz (compounds 1-7) and a NMR Agilent
319 VNMRS 500 MHz (compound 2) – equipped with four and three channels, respectively, and structure
320 elucidation was carried out using modern 1D and 2D pulse sequences, following 1D and 2D
321 experiments were carried out: ¹H, ¹³C, COSY, HMBC, HSQC, TOCSY, and NOESY. The spectra were
322 processed with MestReNova version 12.0.0. Mass spectra were recorded using a Thermo Scientific Q
323 Exactive Plus quadrupole – Orbitrap mass spectrometer coupled with a UPLC Dionex Ultimate 3000
324 RSLC system equipped with an RP-18 Kinetex column (2.10 mm × 100 mm, 2.6 μm, Phenomenex
325 Corporation, Torrance, CA, USA). Full-scan data were recorded in negative ESI mode from *m/z* 100
326 to 1500 at a resolution of 70 000 (at *m/z* 200). Full-scan dd-MS² (top 5) was performed at a resolution
327 of 17 500 (at *m/z* 200), AGC target 1^{e5} with maximum IT 30 ms. For HR-MS in positive mode Q-TOF
328 mass spectrometer with ultra-high resolution and high mass accuracy (HRAM) Impact II (Bruker
329 Daltonik, Bremen, Germany). UHPLC Dionex UltiMate 3000 (Thermo Fisher Scientific, Waltham,
330 MA, USA) was used for LC, with mobile phases: 0.1% formic acid (A) and MeOH (B), flow rate: 0.2
331 mL/min, gradient elution: 0. min 5% of B, 3.6 min 10% of B, 10th min 100% of B. Column block
332 temperature was 35°C, injection volume 5 μL. Kinetex P5, 100A, 1.7μm, 100 × 2.1 mm (Phenomenex,
333 Torrance, CA, USA) was used as column. Solution of sodium formate clusters was used as calibration
334 mixture for accurate mass calibration, with MS source settings: end plate offset: 500 V, capillary 4500
335 V, nebulizer pressure 0.3 Bar, dry gas: 4.0 L/min, dry temperature: 250°C. MS/MS spectra were
336 collected at three collision energy levels of 20, 40, 60 eV per each peak. Data acquisition were carried
337 out by Control 4.0 and HyStar 3.2 software and the results were processed using Compass
338 DataAnalysis 4.3 (all SW of Bruker Daltonik, Bremen, Germany). Chromatography was controlled by
339 Chromeleon Xpress link (Thermo Fisher Scientific, Waltham, MA, USA). For fragmentation analysis

340 Mass Frontier 7.0.5.9 SR3 (High Chem Ltd., Bratislava, Slovakia) and Thermo Excalibur 3.0.63
341 (Thermo Fisher Scientific, Waltham, MA, USA) software was used.

342 **Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Figure S1-S77: NMR
343 spectra of compounds 1-7, HRMS of compounds 1-7, HPLC chromatograms of compounds 1-7.

344 **Author Contributions:** Conceptualization, Jaroslav Havlík; Data curation, Hristo Vasilev, Samir Ross, Dagmar
345 Jankovská and Ondřej Veselý; Formal analysis, Dagmar Jankovská and Ondřej Veselý; Funding acquisition,
346 Karel Šmejkal and Petr Maršík; Investigation, Hristo Vasilev, Petr Maršík and Jaroslav Havlík; Methodology,
347 Hristo Vasilev and Karel Šmejkal; Writing – original draft, Hristo Vasilev, Karel Šmejkal and Petr Maršík.

348 **Funding:** The financial support of GACR no. 16-07193S project is gratefully acknowledged by K. Š., MŠMT no.
349 LD14079 by P. M., and CIGA project no. 20172031 to P. M., J. H., and O. V; projects of Erasmus Program and
350 project no. 30016/672 DBU by H. V.

351 **Acknowledgments:** Authors acknowledge the technical assistance of Ing. Hanka Vítková, Dr. Klaus Bergander,
352 and Dr. Paraskev Nedialkov.

353 **Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the
354 study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to
355 publish the results.

356

357 **References**

- 358 1. Heywood, V. H. et al. *Flora Europaea*. 2, (Cambridge University Press, 1972).
- 359 2. Pistelli, L. F. Secondary metabolites of genus *Astragalus*: Structure and biological activity. in *Bioactive*
- 360 *Natural Products (Part H)* (ed. Atta-ur-Rahman, B. T.-S. in N. P. C.) 27, 443–545 (Elsevier, 2002).
- 361 3. Ionkova, I., Shkondrov, A., Krasteva, I. & Ionkov, T. Recent progress in phytochemistry, pharmacology
- 362 and biotechnology of *Astragalus* saponins. *Phytochem. Rev.* 13, 343–374 (2014).
- 363 4. Andreev, M. Anchev, S. Kozucharov, M. Markova, D. Peev, A. P. *Field Guide of the Bulgarian Vascular*
- 364 *plants*. (Nauka i Izkustvo, 1992).
- 365 5. Bratkov, V., Shkondrov, A., Zdraveva, P. & Krasteva, I. Flavonoids from the genus *Astragalus*:
- 366 *Phytochemistry and biological activity*. *Pharmacogn. Rev.* 10, 11–32 (2016).
- 367 6. Ionkova I. Isolation and HPLC-TLC Analysis of the Major Flavonoids from *Astragalus aitosensis*.
- 368 *Planta Medica* 1990; 56(06): 581 - 581. doi:10.1055/s-2006-961184.
- 369 7. Ionkova I (2008) *Farmaceutichno znachimi biologichno aktivni veshchestva ot iztochnici s optimiziran*
- 370 *fitohimichen potencial (Pharmaceutically significant biologically active compounds from sources with*
- 371 *optimized phytochemical potential)*. DSci Dissertation, Medical University of Sofia.
- 372 8. Apostolova, I. *Red Data Book of the Republic of Bulgaria - Digital edition*. *Astracantha arnacantha*
- 373 Available at: www.e-codb.bas.bg/rdb/en/vol1/Astarnac.html.
- 374 9. Buschi, Carlos A.; Pomilio, Alicia B. Isorhamnetin 3-O-robinobioside from *Gomphrena martiana*.
- 375 *Journal of Natural Products* (1982), 45(5), 557-9.
- 376 10. Yasukawa, K. E. N. & Sekine, H. Two flavonol glycosides from *Lysimachia fortunei*. 28, 2215–2216
- 377 (1989).
- 378 11. Yasukawa, K.; Takido, M. A flavonol glycoside from *Lysimachia mauritiana*. *Phytochemistry* 1987, 26,
- 379 1224-1226.
- 380 12. Kijima, Hidehiko; Ide, Toshinori; Otsuka, Hideaki; Takeda, Yoshio. Alangiflavoside, a new flavonol
- 381 glycoside from the leaves of *Alangium premnifolium*. *Journal of Natural Products* (1995), 58(11), 1753-
- 382 5.
- 383 13. Zhang, Yujuan; Wang, Dongmei; Yang, Lina; Zhou, Dan; Zhang, Jingfang. Purification and
- 384 characterization of flavonoids from the leaves of *Zanthoxylum bungeanum* and correlation between
- 385 their structure and antioxidant activity. *PLoS One* (2014), 9(8), e105725/1-e105725/11, 11 pp.
- 386 14. Gorai, Dilip; Jash, Shyamal K.; Roy, Rajiv. Flavonoids from *Astragalus* genus. *International Journal of*
- 387 *Pharmaceutical Sciences and Research* (2016), 7(7), 2732-2747.
- 388 15. Li, Xiaoxia; Qu, Lu; Dong, Yongzhe; Han, Lifeng; Liu, Erwei; Fang, Shiming; Zhang, Yi; Wang, Tao. A
- 389 review of recent research progress on the *Astragalus* genus. *Molecules* (2014), 19(11), 18850-18880, 31
- 390 pp.
- 391 16. Krasteva, Iliana; Shkondrov, Aleksandar; Ionkova, Iliana; Zdraveva, Petranka. *Advances in*
- 392 *phytochemistry, pharmacology and biotechnology of Bulgarian Astragalus species*. *Phytochemistry*
- 393 *Reviews* (2016), 15(4), 567-590.
- 394 17. Krasteva I, Bratkov V, Bucar F et al (2015) Flavoalkaloids and flavonoids from *Astragalus*
- 395 *monspessulanus*. *J Nat Prod* 78(11):2565–2571.
- 396 18. Alaniya MD, Aneli DN, Patudin AV et al (1983) Flavonoidnye glikozidy iz *Astragalus cicer* (Flavonoid
- 397 glycosides of *Astragalus cicer*). *Khim Prir Soedin* 19(4):528.

398 **Sample Availability:** Samples of the compounds are not available from the authors.