Investigation of the glucosinolates in *Hesperis matronalis* L. and *Hesperis laciniata* All.: unveiling 4'-O- β -D-apiofuranosylglucomatronalin

Sabine Montaut^{a,*}, Sharayah Read^a, Ivica Blažević^b, Jean-Marc Nuzillard^c, Marin Roje^d,
Dominique Harakat^c, Patrick Rollin^e

^a *Department of Chemistry and Biochemistry, Biomolecular Sciences Programme, Laurentian University, 935 Ramsey Lake Road, Sudbury, ON P3E 2C6, Canada*

^b *Department of Organic Chemistry, Faculty of Chemistry and Technology, University of Split, Ruđera Boškovića 35, 21000 Split, Croatia*

^c *Université de Reims Champagne Ardenne, CNRS, ICMR, UMR 7312, Reims, France*

^d *Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia*

^e *Université d'Orléans et CNRS, ICOA, UMR 7311, BP 6759, F-45067 Orléans, France*

* Corresponding author. E-mail address: smontaut@laurentian.ca (S. Montaut). Tel: +1(705)675-1151 ext. 2185. Fax: +1(705)675-4844. Department of Chemistry and Biochemistry, Biomolecular Sciences Programme, Laurentian University, 935 Ramsey Lake Road, Sudbury, ON P3E 2C6, Canada

ABSTRACT

The glucosinolate (GSL) profiles of wild-growing plants from the genus *Hesperis*, i.e. *Hesperis laciniata* All. (leaf, stem, flower, and root) from Croatia and *Hesperis matronalis* L. (leaf, stem, flower, seed, and root) from Canada, were established by LC-MS. During this investigation, 5-(methylsulfanyl)pentyl- (**3**), 6-(methylsulfanyl)hexyl- (**4**), 6-(methylsulfinyl)hexyl- (**6**), and 4'- α -L-rhamnopyranosyloxybenzyl- (**17**) GSLs were identified. In addition, the presence of 7-(methylsulfinyl)heptyl GSL (**18**), hydroxy-(α -L-rhamnopyranosyloxy)benzyl GSL, and of one D-apiosylated analogue of **17** were suggested. Moreover, one new GSL, 4'-O- β -D-apiofuranosylglucomatronalin (**19**) was isolated from *H. laciniata* (flower, stem and leaf) and characterized by spectroscopic data interpretation. Finally, we report the presence of **3**, **4**, **6**, **19**, glucosinalbin (**12**), and 4-hydroxyglucobrassicin (**20**) in *H. matronalis* and hypothesize the presence of glucomatronalin (**13**) and 3-hydroxy-6-(methylsulfanyl)hexyl GSL (**21**).

Keywords: *Hesperis laciniata* All., *Hesperis matronalis* L., Brassicaceae, glucosinolate, NMR, LC-MS, 4'-O- β -D-apiofuranosylglucomatronalin.

1. Introduction

Glucosinolates (GSLs) are secondary metabolites whose structures are highly diverse (Fig. 1) [1]. A review by Fahey *et al.* (2001) reported *ca* 120 natural of natural GSLs including poorly resolved ones [2]. Since then, the number of accepted GSL structures has increased to around 130, although all of the structures have not yet been established using the proper techniques, such as MS and NMR [1,3]. In other respects, Clarke (2010) suggested in his review that there exist approximately 200 GSLs without regarding the documentation and natural occurrence of each structure [1]. Each new series of homologous GSL structures was proposed using extrapolation, and an additional 180 GSLs were predicted to be found in Nature; the structures, formulae and accurate masses were also provided for use in mass spectrometry [4]. A more recent review established that sufficiently characterized GSLs by modern spectroscopic methods (NMR and MS), were 88 by mid-2018. Moreover, 49 partially characterized structures with highly variable evidence exist, including a few detected in genetically manipulated plants [5].

According to the structure-based classification, GSLs are generally regarded as being aliphatic, arylaliphatic, or indole-type. Arylaliphatic and indolyl GSLs have been identified together with extra-glycosylated GSLs, *e.g.* containing L-rhamnose, L-arabinose or D-apiose as additional sugar units linked to the side chain [3] - the prefixes “intra” and “extra” are respectively indicative of functionalizations (namely glycosylation or esterification) either on the GSL thioglucosyl unit or on the GSL side chain. Only *Hesperis matronalis* L. and *Noccaea caerulea* (J. Presl & C. Presl) F. K. Mey. from the Brassicaceae family, *Moringa stenopetala* (Baker f.) Cufod. and *Moringa oleifera* Lam. from the Moringaceae family, and *Reseda lutea* L. from the Resedaceae family are known to contain such uniquely extra-glycosylated phenolic

GSLs [3,6,7]. In other respects, intra-acylated (on the glucosyl unit) GSLs and/or extra-acylated (on the side chain) GSLs have also been detected among Brassicaceae. In most cases however, these esters derive from benzoic, cinnamic, *p*-coumaric, isoferulic, or sinapic acids, being frequently conjugated at the C-6 position of the glucose moiety. Such GSLs, which are mainly present in the seeds, are found in either very low or not quantified amounts [3, 8-15]. Some of the most unusual GSLs originate from four species: *Barbarea vulgaris* R. Br., *Arabidopsis thaliana* (L.) Heynh., *Eruca sativa* Mill., and *Isatis tinctoria* L. [1]. Mithen *et al.* also reported that New World Capparidaceae contain several distinctive and perhaps unique GSLs with complex and unresolved structures, indicating continued diversification in GSL biosynthesis [16].

Hesperis is a genus of flowering plants in the mustard family (Brassicaceae) which comprises almost 60 species; it is especially well represented with many taxa at the junctions of the Irano-Turanian, Mediterranean, and Euro-Siberian phytogeographic regions. Among the 14 *Hesperis* species registered in the Flora of Europe [17], 4 species and 5 subspecies are wild-growing in Croatia [18]. Many of these plants bear showy, fragrant flowers in shades of purple and white. One of the more widely known and investigated species is the common garden flower Dame's Rocket (*H. matronalis*), that grows in most parts of the U.S. and Canada [19]. The GSL profile of this plant includes 5 groups of molecules as previously reported by many research groups (Fig. 1): 1) thiofunctionalized GSLs [3-(methylsulfanyl)propyl GSL (glucoibervirin, **1**), 4-(methylsulfanyl)butyl GSL (glucoerucin, **2**), 5-(methylsulfanyl)pentyl GSL (glucoberteroin, **3**), 6-(methylsulfanyl)hexyl GSL (glucolesquerellin, **4**), 5-(methylsulfinyl)pentyl GSL (glucoalyssin, **5**), and 6-(methylsulfinyl)hexyl GSL (glucohesperin, **6**); 2) branched alkyl GSLs [1-methylpropyl GSL (glucocochlearin, **7**); 3) alkenyl GSLs [but-3-enyl GSL (gluconapin, **8**), (*R*)-

2-hydroxybut-3-enyl GSL (progoitrin, **9**), and (*S*)-2-hydroxybut-3-enyl GSL (epiprogoitrin, **10**); 4) benzyl-type GSLs [benzyl GSL (glucotropaeolin, **11**), 4-hydroxybenzyl GSL (glucosinalbin, **12**), and 3,4-dihydroxybenzyl GSL (glucomatronalin, **13**)] [2,6,20-23]; and 5) extra-glycosylated GSLs [3-*O*-apiosylglucomatronalin, **14**], its 3,4-dihydroxybenzoyl ester (**15**), and 3,4-dimethoxybenzoyl ester (**16**) derivatives] [2,6,22]. Compound **2** was identified in 8-week-old plants [20], whereas **1** and **3-14** were found in seeds [2,6,21-23].

Extra-glycosylated GSLs produced by *H. matronalis*, containing D-apiofuranosyl- or acylated D-apiofuranosyl moieties connected to a benzyl-type side chain are unique to this species [24]. However, the identification of such unusual apiosylated GSLs, claimed by Larsen *et al.* (1992), only referenced unpublished work [22]. Later on, Bennett *et al.* (2004) reported **14** and **16** with only an ion-pairing LC-MS method [6]. In the reviews by Bellostas *et al.* (2007) and Clarke *et al.* (2010), four and five apiosylated GSL structures were illustrated, respectively; however, none referenced peer-reviewed papers [1,4,25]. The roles of compounds **14-16** in herbivore deterrence are unknown, but they have been regarded as attractants to a monophagous herbivore that specializes on *H. matronalis* [22]. Larsen *et al.* (1992) showed that the Euro-Siberian weevil *Ceutorhynchus inaeffectatus* Gyllenhal was monophagous on this plant and found these three specific apiosylated GSLs to act as powerful feeding stimulants for the weevil [22]. Those GSLs were thus suggested to play a key role in the specificity of monophagous insects, a property which could be considered for biological control [22]. *Hesperis pendula* DC. was also investigated, but only **6** and **12** were reported in this plant [2,20].

Thus, the previously not investigated species *Hesperis laciniata* All., wild-growing in Croatia, was chosen mostly as it was hypothesized that this plant can biosynthesize unique apiose-containing GSLs. In parallel, we decided to reexamine the GSLs in *H. matronalis* wild-

growing in Canada. The investigations were based on a well-established method for intact GSLs, involving HPLC-ESI-MS coupled to a photodiode array detector. Besides, in order to establish the structure of compound **19**, the intact GSL was isolated from *H. laciniata* and *H. matronalis*, and characterized using spectroscopic techniques.

2. Results and discussion

The flower, leaf, stem, and root of *H. laciniata* were harvested on Marjan hill in Split, Croatia. The plant parts were extracted and analyzed by LC-MS for intact GSLs. The t_R , mass, and UV data of the products were compared with the values for standards from our GSL library (see experimental section). The structures of the GSLs not previously detected in the genus *Hesperis* are given in Fig. 2. In the root (Fig. S1a), the minor compound at t_R 21.4 min had similar t_R , mass ($M = 570$ u), and UV spectra to those of an authenticated sample of 4'- α -L-rhamnopyranosyloxybenzyl GSL (glucomoringin, **17**) [26]. The major compound found at t_R 21.8 min (464 u) and the minor compounds found at t_R 27.5 min ($M = 434$ u) and t_R 31.6 min ($M = 448$ u) were determined to be glucohesperin (**6**), glucoberteroin (**3**), and glucolesquerellin (**4**), respectively, by comparison of t_R , mass, and UV spectra to data obtained from previous isolation of **3**, **4**, and **6** in our laboratory [27]. The minor compound at t_R 25.1 min was hypothesized to be 7-(methylsulfinyl)heptyl GSL (**18**) by comparing its mass spectrum ($M = 478$ u) to the mass spectrometry data reported by Clarke (2010) [4]. In the root (Fig. S1a), flower (Fig. S1b), leaf (Fig. S1c), and stem (Fig. S1d), the compound at t_R 20.1 min was hypothesized to be 4'-hydroxy-3'- or 3'-hydroxy-4'- or 2'-hydroxy-4'-(α -L-rhamnopyranosyloxy)benzyl GSL by comparing its mass spectrum ($M = 586$ u) to the mass spectrometry data reported by Clarke (2010) [4]. Peaks 1 ($t_R = 19.9$ min) and 4 ($t_R = 21.6$ min) in Fig. S1a did not match any GSLs in our library. Therefore, the identification of peak 1, in flower, stem, leaf, and root, as well as peak 4 in flower,

stem, and leaf required isolation (see experimental section). Performed on a combined mass of 11.4 g of flower, stem, and leaf, the isolation produced 68 mg of peak 1 and 33 mg of peak 4, the structures of which were elucidated using UV, NMR (^1H , ^{13}C , HMBC, HSQC, and ROESY), and HRMS data. The peak 4 at $t_R = 21.6$ min contained:

- some minor impurity of glucomoringin (**17**),
- one major apiosylated analogue of **17** with a mass of 556 u and a molecular formula of $\text{C}_{19}\text{H}_{26}\text{NO}_{14}\text{S}_2$ established by HRMS,
- another major arylaliphatic compound, whose structure was not firmly established.

We were not able to purify this apiosylated analogue of **17** further in order to assign with certainty the position of the apiosyl moiety on the aromatic ring. In fact, in the aromatic proton area, the impurities display signals which overlap with the signals of other products of peak 4 (data not presented).

The UV spectrum of peak 1 ($t_R = 19.9$ min, Fig. S1) indicated that the compound belongs to the arylalkyl GSL class [28]. The mass of the compound was 572 u, and the molecular formula $\text{C}_{19}\text{H}_{26}\text{NO}_{15}\text{S}_2$ was established by HRMS (Fig S3). The ^1H (Fig. S4-S5) and ^{13}C spectra (Fig. S6) of the compound, recorded in D_2O , show the typical resonances of a β -thioglucopyranosyl unit within a benzylic GSL, such as glucomoringin (**17**) (Table 1) [26,29]. The chemical shifts of H-2 and H-3 are very close, thus giving to the signals of H-1 and H-4 a complex appearance caused by the virtual coupling of H-1 with H-3 and of H-4 with H-2. The benzylic methylene group was identified by the shielding of C-8 in ^{13}C NMR (δ_{C} 40.5) and by the H-8 signals appearing as an AB system with a high coupling constant value ($J = 16.4$ Hz). The aromatic ring is trisubstituted, as revealed by the signals of three aromatic protons whose coupling pattern is typical of a 1',3',4'-substitution. The H-8 signals correlate in the HMBC spectrum (Fig. S7-S10) with a

quaternary carbon at δ 133.5 (C-1') and with two methine carbons at δ_C 123.2 (C-6') and 118.8 (C-2'). The signal of H-6' is a dd, $J_{H-6'-H-2'} = 2.1$ Hz and $J_{H-6'-H-5'} = 8.3$ Hz. The chemical shift of C-5' at δ 120.8 was given by the HSQC spectrum (Fig. S11). Two ^{13}C NMR resonances at δ_C 146.2 and 149.3 are those of two oxygen-bearing aromatic carbons; the one at δ_C 146.2 strongly correlates in the HMBC spectrum with the 3 aromatic ^1H resonances, thus leading to assign it at position 4' and therefore to assign the one at δ_C 149.3 to position 3'. Five ^{13}C NMR resonances at δ_C 110.5 (CH), 79.9 (CH), 82.3 (C), 77.3 (CH₂) and 66.2 (CH₂) fit with a pentose unit that could substitute an oxygen atom at position 4' or 3'. The HMBC spectrum showed a correlation between the anomeric proton signal of H-1'' at δ_H 5.69 with the signal of C-4', thus indicating the presence of a hydroxyl group at position 3'. The coupling partner of H-1'' is H-2'' (δ_H 4.36, $J = 3.2$ Hz). The absence of any coupling partner of H-2'', H-1'' apart, incited us to place the quaternary carbon at position 3''. The HMBC correlation of the CH₂ at δ_C 77.3 with H-2'' lead to its placement at position 4''; the HMBC correlation of the same C-4'' with H-1'' lead to the closure of a furanose ring. The remaining CH₂OH group was finally connected to C-3'', as supported by the HMBC correlations of the H-5'' protons with C-3'', C-2'' and C-4''. The atom connectivity within the pentose unit is compatible with the structure of apiose, a hypothesis that was comforted by a ROESY correlation (Fig. S12) between the H-2'' and H-5'' signals that indicated that the hydroxyl groups at positions 2'' and 3'' are located on the same side of the ring. The β -anomeric configuration for the apiofuranosyl moiety was supported by the anomeric signal at H-1'' (δ_C 110.5, δ_H 5.69, $J = 3.2$ Hz) [30,31]. The ROESY correlation between the H-1'' and H-5' signals also contributed to establish the substitution pattern of the aromatic ring. Therefore, the NMR data reported here suggested that peak 1 (Fig. S1) corresponds to 4'-O- β -D-

apiofuranosylglucomatronalin (**19**) which represents a GSL not previously reported in the plant kingdom. Its structure is given in Fig. 2.

These findings brought us to reexamine the GSL profile of *H. matronalis*. In fact, 3'-*O*-apiosylglucomatronalin (**14**), a regio-isomer of GSL **19**, was previously reported in seeds of *H. matronalis* by Larsen *et al.* (1992) [22]. However, the spectroscopic data supporting Larsen *et al.*'s claim were never published. Therefore, in order to verify whether **14** was really present in *H. matronalis* or if it was **19** instead as we suspected it, we harvested several plant organs (leaf, stem, flower, seed, and root) of *H. matronalis* growing in Sudbury (Ontario, Canada). The plant parts were extracted and analyzed by LC-MS for intact GSLs following the same experimental procedure used for the analyses of *H. laciniata* extracts. The only GSL present in the leaf, stem and fruit of *H. matronalis* (Fig. S2a-c) had t_R (19.9 min), mass ($M = 572$ u), and UV spectrum similar to those of **19** previously isolated from *H. laciniata*. In the root (Fig. 2d), the same major peak at $t_R = 19.9$ min with identical mass ($M = 572$ u), and UV spectrum compared with **19** detected in *H. laciniata* was present. In addition, the presence of glucoberteroin (**3**), glucolesquerellin (**4**), and glucohesperin (**6**) was confirmed by comparison of t_R , UV, and mass spectra to data obtained from authenticated samples previously isolated in our laboratory [27]. Moreover, the compound at t_R 18.1 min had t_R , mass, and UV spectra identical to those of a commercial standard of glucosinalbin (**12**). Furthermore, the peak at t_R 20.9 min was identified by comparison of t_R , mass, and UV spectra with those of 4-hydroxyglucobrassicin (**20**), previously identified in our laboratory [32]. The compounds at $t_R = 13.1$ min ($M = 440$ u) and t_R 24.1 min ($M = 464$ u) were hypothesized to be glucomatronalin (**13**) and 3-hydroxy-6-(methylsulfanyl)hexyl GSL (**21**), respectively, by comparing their mass spectra to the mass data reported by Clarke (2010) [4]. Additionally, in the flower (Fig. S2d), the same major peak at $t_R =$

19.9 min with identical mass, and UV spectrum compared to 4'-O- β -D-apiofuranosylglucomatronalin (**19**) isolated from *H. laciniata* was also observed, along with some traces of **20**.

In order to ascertain the presence of GSL **19** in *H. matronalis*, the isolation of the peak at t_R = 19.9 min was performed from 5.2 g of the flower producing 24.4 mg of pure product (see experimental section). The UV, NMR (^1H , ^{13}C , HMBC, HSQC, COSY, and ROESY) (Fig. S4-S15), and HRMS mass spectrometric data of this product confirmed that the peak t_R = 19.9 min isolated from *H. matronalis* was also 4'-O- β -D-apiofuranosylglucomatronalin (**19**). Together with the NMR data, the obtained CD spectra (Fig. S16) suggest that the peak t_R = 19.9 min isolated from *H. matronalis* and from *H. laciniata* correspond to the same compound. However, examined solutions are probably not of the same purity considering slight differences in spectra at 231 nm.

We have reported here a first investigation of the GSLs in *H. laciniata* which resulted in the identification of four known GSLs (**3**, **4**, **6**, and **17**) and the isolation of one new GSL (**19**). In our study, we have confirmed that *H. matronalis* contains four GSLs (**3**, **4**, **6**, and **12**) previously described in the literature [2,6,21-23]. Besides, we proved the presence in *H. matronalis* of a new extra-glycosylated GSL **19**. Contrary to previous reports, we did not detect any **1**, **2**, **5**, **7-11**, or **14-16** in *H. matronalis*. Our study has focused on leaf, stem, fruit, and root of *H. matronalis* separately, whereas previous studies were performed on seeds and 8-week-old plants. Thus, the differences in GSL profiles may be due to genetic and ecological contexts.

Apiosylated GSLs seem to have a narrow distribution in the family Brassicaceae and to be specific to the genus *Hesperis* and in particular to *H. matronalis* and *H. laciniata*. Thus, it would be worth examining or reexamining other *Hesperis* species. In fact, the gas chromatographic

analysis of the GSL hydrolysis products (isothiocyanates or oxazolidine-2-thiones) obtained from the seeds of *Hesperis pendula* DC. showed that the plant biosynthesized GSLs **6** and **12** [2,21]. However, isothiocyanates generated from apiosylated GSLs are not volatile compounds and therefore they cannot be detected by gas chromatography, which means that the presence of apiosylated GSLs in *H. pendula* cannot be *a priori* ruled out.

Apiosylated GSLs are not the only naturally occurring D-apiosides. Some examples of apiosylated phenolics (e.g. phenylpropanoids lignans, flavonoids including isoflavonoids, coumarins, and anthocyanins), terpenes, terpenoids, saponins, cyanogenic glucosides, aliphatic and aromatic alcohols, lactones, and alkaloids have been reported [33]. Such secondary metabolites are found particularly in ferns, gymnosperms, and angiosperms (Fabaceae, Lamiaceae, and Asteraceae families). Apiose-containing glycosides have also been reported in lichens and in one fungus as well [33].

In conclusion, further investigation of the biological potential of the newly characterized 4'-*O*- β -D-apiofuranosylglucomatronalin (**19**) and its myrosinase-degradation products would be of interest to estimate their possible health benefits. Additionally, structural characterization and comparison of biological activities of the minor apiosylated GSLs from *Hesperis* sp. also constitute a valuable research project.

3. Experimental

3.1. General experimental procedures

All solvents were ACS grade and used as such, except CHCl₃, which was redistilled. D₂O was purchased from Euriso-Top (Saint-Aubin, France). Formic acid was purchased from BDH

(Toronto, ON, Canada). Glucosinabin was purchased from Apin Chemicals Ltd. (Abingdon, UK). HPLC-grade MeOH, Et₃N (reagent grade), and thymol were purchased from Fisher Scientific (Whitby, ON, Canada). HPLC-grade H₂O was generated in the laboratory through a Nanopure Diamond Ultrapure water system provided by Barnstead (Dubuque, IA, USA). Kieselgel F₂₅₄ analytical TLC aluminum sheets were purchased from EM Science (Gibbstown, NJ, USA); compounds were visualized under UV light and by dipping the plates into a 95% ethanol solution containing 1% (w/v) thymol and 10% (v/v) H₂SO₄ followed by heating. Flash column chromatography was carried out using SPE bulk sorbent large-pore C-18 from Alltech (State College, PA, USA). C₁₈ silica gel cartridges (Mega Bond EluteFlash, 10 g sorbent mass, 60 mL volume) were obtained from Varian, Inc. (Mississauga, ON, Canada). UV spectra were determined on a Cary 60 UV/visible spectrophotometer from Agilent Technologies (Santa Clara, CA, USA). NMR spectra were recorded in D₂O at 600 MHz (¹H) and 150 MHz (¹³C) on a Bruker Avance III 600 spectrometer equipped with a TCI cryoprobe at the Institut de Chimie Moléculaire de Reims (Reims, France); δ values (ppm) are referenced to sodium 3-(trimethylsilyl)propanoate-2,2,3,3-*d*₄ (Sigma-Aldrich, Saint Quentin Fallavier, France), and coupling constants *J* are given in Hz. HRESIMS measurements were recorded at the Institut de Chimie Moléculaire de Reims (Reims, France). An aliquot of isolated natural products was also solubilized in MeOH and directly infused in a quadrupole time-of-flight hybrid mass spectrometer (QTOF micro[®], Waters, Manchester, UK) equipped with an electrospray source. The mass range of the instrument was set at *m/z* 100-1500 and scan duration was set at 2 s in the negative ion mode. The main conditions were: capillary voltage, 3500 V; extraction cone voltage varying between 30 and 40 V (flow of injection 5 μ L/min); source temperature, 80 °C; desolvation temperature, 100 °C. The CD spectra (mdeg) were taken on a UV-CD instrument

Jasco J-815, in a 1 mm (optical path) cuvette. The parameters were the following: band width 2 nm, response 1 sec, sensitivity standard, measurement range 600 - 200 nm, data pitch 0.2 nm, scanning speed 200 nm/min, accumulation 2, cell length 1 mm, and temperature 25 °C. All the spectra were corrected by subtracting baseline value. The samples were dissolved in MeOH (c = 1 mg/mL) using an ultrasound bath at room temperature.

3.2. Plant material

Hesperis laciniata All. was collected in April 2017 on Marjan hill in Split, Croatia. The plant parts were dried in the shade for several days. The plant was identified by Dr. M. Ruščić at the Department of Botany, University of Split, Split, Croatia using Flora Europea [34]. A voucher specimen (DBHL001) is kept at the Department of Biology, University of Split, Split, Croatia. *H. matronalis* L. was collected on July 12, 2017 on the Laurentian University Campus in Sudbury, Ontario, Canada. The plant was identified by Dr. S. Montaut at the Department of Chemistry and Biochemistry, Laurentian University, Sudbury, Ontario, Canada. Voucher specimens (23690, 23591 and 23692) are kept at the Herbarium at Laurentian University Department of Biology, Sudbury, Ontario, Canada.

3.3. LC-MS analysis of glucosinolates

LC-MS analysis was performed by injecting a 10 µL aliquot of the solution of crude extract or fraction into an Agilent Technologies HP 1100 (New Castle, DE, USA) high-performance liquid chromatograph equipped with a quaternary pump, automatic injector, diode-array detector (wavelength range 190-600 nm), vacuum degasser, and a Hypersil ODS column (5 µm, 4.6 × 200 mm). The two mobile phase solvents, MeOH and H₂O, were prepared with 0.15% Et₃N and

0.18% HCO₂H, added as ion-pairing reagents. Both solutions were filtered using 0.45 µm nylon membranes. The initial mobile phase was 100% HPLC-grade H₂O. At 10 min, the mobile phase was switched to a linear gradient of 100% H₂O to 100% MeOH over 60 min. After each run, the initial mobile phase conditions were set and the system was allowed to equilibrate. The flow rate was kept constant at 1 mL/min. The column temperature was held at room temperature [35]. The HPLC was interfaced to an Agilent model 6120 single quadrupole mass spectrometer (Toronto, ON, Canada) with a Chemstation data system LC-MSD B.03.01. The electrospray interface was a standard ES source operating with a capillary voltage of 4 kV and temperature of 350 °C. Spectra were obtained with a fragmentation voltage of 200 eV. The system was operated in the negative and positive ion electrospray modes. Nitrogen was used as nebulizing and drying gas at a flow rate of 12 L/min (60 psig). The mass spectrometer was programmed to perform full scans between *m/z* 100 and 1,500 u.

3.4. Extraction and isolation of 4'-O-β-D-apiofuranosylglucomatronalin (19) and peak 4 (Fig. S1) from H. laciniata

The leaves (3.9 g), stems (4.6 g), and flowers (2.9 g) were frozen in liquid N₂, ground in a mortar, and immediately extracted three times with boiling EtOH-H₂O (60 mL, 7:3 v/v) for 5 min. The solution was filtered and concentrated to dryness (4.1 g). This extract was dissolved in H₂O (10 mL) and submitted to liquid-liquid extraction with EtOAc. The organic layer (EtOAc fraction 576 mg) and the aqueous layer (3.0 g) were concentrated to dryness. The aqueous layer was separated by flash column chromatography (C18 phase, 40 × 90 mm, gradient H₂O/MeOH, 100:0 and 0:100 v/v, 20 mL fractions). Fraction 5 (283 mg) eluted with H₂O was submitted to a solid-phase separation (C18 cartridge, H₂O/MeOH, 100:0 and 0:100 v/v, 3 mL fractions). The combined fractions 13 to 38, obtained from the solid-phase separation and eluted with H₂O,

yielded **19** (68 mg). Fractions 12-24 from the flash column chromatography were combined (94 mg) and submitted to a solid-phase separation (C₁₈ cartridge, H₂O/MeOH, 100:0 and 0:100 v/v, 3 mL fractions). The combined fractions 21 to 47, obtained from the solid-phase separation and eluted with H₂O, yielded a mixture of three compounds (**17**, an apiosylated analogue of **17**, and another compound) gathered under peak 4 (33 mg) (Fig. S1).

4'-O-β-D-Apiofuranosylglucomatronalin (19): white, amorphous powder; UV (H₂O) λ_{max} (log ϵ) 219 (4.0), 277 (3.3) nm; ¹H NMR (D₂O, 600 MHz) and ¹³C NMR (D₂O, 150 MHz) data see Table 1; HRESIMS m/z 572.0750 [M-H]⁻ (calcd for C₁₉H₂₆NO₁₅S₂, 572.0744); HPLC, t_R = 19.9 min.

Apiosylated analog of (17) (peak 4): white, amorphous powder; UV (HPLC) λ_{max} 226, 274 nm; HRESIMS m/z 556.0787 [M-H]⁻ (calcd for C₁₉H₂₆NO₁₄S₂, 556.0795); t_R = 21.6 min.

3.5. Extraction and isolation of 4'-O-β-D-apiofuranosylglucomatronalin (**19**) from *H. matronalis*

The fresh flowers (5.2 g) were frozen in liquid N₂, ground in a mortar, and immediately extracted three times with boiling EtOH-H₂O (5 mL, 7:3 v/v) for 5 min. The solution was filtered and concentrated to dryness (322.7 mg). The extract was submitted to a solid-phase separation (C₁₈ cartridge, H₂O/MeOH, 100:0 and 0:100 v/v, 3 mL fractions). The combined fractions 6 to 11, obtained from the solid-phase separation and eluted with H₂O, yielded compound **19** (24.4 mg).

4'-O-β-D-Apiofuranosylglucomatronalin (19): white, amorphous powder; UV (H₂O) λ_{max} (log ϵ) 219 (4.0), 277 (3.3) nm; ¹H NMR (D₂O, 600 MHz) and ¹³C NMR (D₂O, 150 MHz) data see

Table 1; HRESIMS m/z 572.0738 [M-H]⁻ (calcd for C₁₉H₂₆NO₁₅S₂, 572.0744); HPLC, t_R = 19.9 min.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

Financial support from Natural Sciences and Engineering Research Council of Canada (NSERC Research Tool and Instruments, grant 315095-05), Canadian Foundation for Innovation (Leaders Opportunity Fund) – Ontario Research Fund (Grant 11666), and Laurentian University (Career and Employment Centre) is gratefully acknowledged by S.M. Financial support from CNRS, Conseil Régional Champagne Ardenne, Conseil Général de la Marne, Ministry of Higher Education and Research (MESR), and EU-programme FEDER to the P1AneT CPER project is gratefully acknowledged. Financial support by the Croatian Science Foundation (Grant IP-2016-06-1316) is gratefully acknowledged by I.B.

Appendix A. Supplementary data

Supplementary data related to this article include the reference to the NMR raw data files and of the NMReDATA file [36] for compound **19** and can be found at ...

References

- [1] N. Agerbirk, C.E. Olsen, Glucosinolate structures in evolution, *Phytochemistry* 77 (2012) 16–45.
- [2] J.W. Fahey, A.T. Zalcmann, P. Talalay, The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56 (2001) 5–51.

- [3] I. Blažević, S. Montaut, F. Burčul, P. Rollin, 2017. Glucosinolates: Novel Sources and Biological Potential, in: Ramawat, K. G., Mérillon, J.-M. (Eds.), Glucosinolates. Springer International Publishing, Cham, Switzerland, pp. 3-60.
- [4] D.B. Clarke, Glucosinolates, structures and analysis in food. Anal. Methods 2 (2010) 310–325.
- [5] I. Blažević, S. Montaut, F. Burčul, C.E. Olsen, M. Burow, P. Rollin, N. Agerbirk, Glucosinolate structural diversity, identification, chemical synthesis and metabolism by plants. Phytochemistry (2019) In press.
- [6] R.N. Bennett, F.A. Mellon, P.A. Kroon, Screening crucifer seeds as sources of specific intact glucosinolates using ion-pair high-performance liquid chromatography negative ion electrospray mass spectrometry. J. Agric. Food Chem. 52 (2004) 428–438.
- [7] R.M. de Graaf, S. Krosse, A.E.M. Swolfs, E. te Brinke, N. Prill, R. Leimu, P. M. van Galen, Y. Wang, M.G.M. Aarts, N.M. van Dam, Isolation and identification of 4- α -rhamnosyloxy benzyl glucosinolate in *Noccaea caerulea* showing intraspecific variation. Phytochemistry 110 (2015) 166–171.
- [8] D.J. Kliebenstein, J.C. D’Auria, A.S. Behere, J.H. Kim, K.L. Gunderson, J.N. Breen, G. Lee, J. Gershenzon, R.L. Last, G. Jander, Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. Plant J. 51 (2007) 1062–1076.
- [9] T. Mohn, Hamburger, M., Glucosinolate pattern in *Isatis tinctoria* and *I. indigotica* seeds. Planta Med. 74 (2008) 885–888.

- [10] N.S. Survey, B. Kumar, C.P. Upadhyaya, E. Ko, C. Lee, J.N. Choi, D.-Y. Yoon, Y.-S. Jung, S.W. Park, Characterization of a cinnamoyl derivative from broccoli (*Brassica oleracea* L. var. *italica*) florets. *Fitoterapia* 81 (2010) 1062–1066.
- [11] N. Agerbirk, C.E. Olsen, Isoferuloyl derivatives of five seed glucosinolates in the crucifer genus *Barbarea*. *Phytochemistry* 72 (2011) 610–623.
- [12] G. Bianco, F. Lelario, F.G. Battista, S.A. Bufo, T.R.I. Cataldi, Identification of glucosinolates in capers by LC-ESI-hybrid linear ion trap with Fourier transform ion cyclotron resonance mass spectrometry (LC-ESI-LTQ-FTICR MS) and infrared multiphoton dissociation. *J. Mass. Spectrom.* 47 (2012) 1160–1169.
- [13] N. Agerbirk, C.E. Olsen, D. Cipollini, M. Ørgaard, I. Linde-Laursen, F.S. Chew, Specific glucosinolate analysis reveals variable levels of epimeric glucobarbarins, dietary precursors of 5-phenyloxazolidine-2-thiones, in watercress types with contrasting chromosome numbers. *J. Agric. Food Chem.* 62 (2014) 9586–9596.
- [14] G. Bianco, N. Agerbirk, I. Losito, T.R.I. Cataldi, Acylated glucosinolates with diverse acyl groups investigated by high resolution mass spectrometry and infrared multiphoton dissociation. *Phytochemistry* 100 (2014) 92–102.
- [15] N. Agerbirk, C.E. Olsen, C. Heimes, S. Christensen, S. Bak, T.P. Hauser, Multiple hydroxyphenethyl glucosinolate isomers and their tandem mass spectrometric distinction in a geographically structured polymorphism in the crucifer *Barbarea vulgaris*. *Phytochemistry* 115 (2015) 130–142.
- [16] R. Mithen, R. Bennett, J. Marquez, Glucosinolate biochemical diversity and innovation in the Brassicales. *Phytochemistry* 71 (2010) 2074–2086.

- [17] A. Duran, *Hesperis ozcelikii* (Brassicaceae), a new species from Turkey. Ann. Bot. Fennici 46 (2009) 577–584.
- [18] T. Nikolić, Flora Croatica Database. <http://hirc.botanic.hr/fcd>, 2012 (accessed 4 February 2018).
- [19] USDA, NRCS, The PLANTS Database. <http://plants.usda.gov>, 2018 (accessed 4 February 2018).
- [20] R.A. Cole, Isothiocyanates, nitriles and thiocyanates as products of autolysis of glucosinolates in *Cruciferae*. Phytochemistry 15 (1976) 759–762.
- [21] M.E. Daxenbichler, G.F. Spencer, D.G. Carlson, G.B. Rose, A.M. Brinker, R.G. Powell, Glucosinolate composition of seeds from 297 species of wild plants. Phytochemistry 30 (1991) 2623–2638.
- [22] L.M. Larsen, J.K. Nielsen, H. Sørensen, Host plant recognition in monophagous weevils: Specialization of *Ceutorhynchus inaeffectatus* to glucosinolates from its host plant *Hesperis matronalis*. Entomol. Exp. Appl. 64 (1992) 49–55.
- [23] S.F. Vaughn, M.A. Berhow, Glucosinolate hydrolysis products from various plant sources: pH effects, isolation, and purification. Ind. Crops Prod. 21 (2005) 193–202.
- [24] J.A.A. Renwick, The chemical world of crucivores: lures, treats and traps. Entomol. Exp. Appl. 104 (2002) 35–42.
- [25] N. Bellostas, A.D. Sørensen, J.C. Sørensen, H. Sørensen, Genetic variation and metabolism of glucosinolates. Adv. Bot. Res. 45 (2007) 369–415.

- [26] D. Gueyrard, J. Barillari, R. Iori, S. Palmieri, P. Rollin, First synthesis of an *O*-glycosylated glucosinolate isolated from *Moringa oleifera*. *Tetrahedron Lett.* 41 (2000) 8307-8309.
- [27] S. Montaut, J. Grandbois, L. Righetti, J. Barillari, I. Renato, P. Rollin, Updated glucosinolate profile of *Dithyrea wislizenii*. *J. Nat. Prod.* 72 (2009) 889-893.
- [28] J.-P. Wathelet, R. Iori, O. Leoni, P. Rollin, A. Quinsac, S. Palmieri, Guidelines for glucosinolate analysis in green tissues used for biofumigation. *Agroindustria* 3 (2004) 257-266.
- [29] N. Ibrahim, I. Allart-Simon, G.R. De Nicola, R. Iori, J.-H. Renault, P. Rollin, J.-M. Nuzillard, Advanced NMR-based structural investigation of glucosinolates and desulfoglucosinolates. *J. Nat. Prod.* 81 (2018) 323-334.
- [30] Y. Lei, L.-J. Wu, H.-M. Shi, P.-F. Tu, Three new glycosides from the stems of *Milusa balansae*. *Helv. Chim. Acta* 91 (2008) 495-500.
- [31] W.-S. Feng, Y.-Y. Pei, X.-K. Zheng, C.-G. Li, Y.-Y. Ke, Y.-Y. Lv, Y.-L. Zhang, A new kaempferol trioside from *Silphium perfoliatum*. *J. Asian Nat. Prod. Res.* 16 (2014) 393-399.
- [32] S. Montaut, I. Blažević, M. Ruščić, P. Rollin, LC–MS profiling of glucosinolates in the seeds of *Brassica elongata* Ehrh., and of the two stenoendemic *B. botteri* Vis and *B. cazzae* Ginz. & Teyber. *Nat. Prod. Res.* 31 (2017) 58-62.
- [33] M. Pičmanová, B.M. Møller, Apiose: one of nature's witty games. *Glycobiology* 26 (2016) 430-442.

- [34] T. G. Tutin, V.H. Heywood, N.A. Burges, D.M. Moore, D.H. Valentine, S.M. Walter, D.A. Webb, *Flora Europaea*, Cambridge University Press, Cambridge (U.K.), Vol. 1-5, 1968-1980.
- [35] C.L. Zrybko, E.K. Fukuda, R.T. Rosen, Determination of glucosinolates in domestic and wild mustard by high-performance liquid chromatography with confirmation by electrospray mass spectrometry and photodiode-array detection. *J. Chromatogr. A.* 767 (1997) 43-52.
- [36] M. Pupier, J.-M. Nuzillard, J. Wist, N. E. Schlörer, S. Kuhn, M. Erdelyi, C. Steinbeck, A. J. Williams, C. Butts, T. D. W. Claridge, B. Mikhova, W. Robien, H. Dashti, H. R. Eghbalian, C. Farès, C. Adam, P. Kessler, F. Moriaud, M. Elyashberg, D. Argyropoulos, M. Pérez, P. Giraudeau, R. R. Gil, P. Trevorow, D. Jeannerat, NMReDATA, a standard to report the NMR assignment and parameters of organic compounds. *Magn. Res. Chem.* 56 (2018) 703–715.

Table 1NMR data (600 MHz, D₂O) for 4'-O- β -D-apiofuranosylglucomatronalin (**19**).

	$\delta(^{13}\text{C})$	$\delta(^1\text{H})$	$J^1\text{H} - ^1\text{H}$ (in Hz)	HMBC (H $\rightarrow$ C)	COSY	ROESY
1	84.3	4.73	d, 9.4	2, 3, 5(w), 7	2	5, 8a/b
2	74.7	3.33	m	1, 3	1, 4(w)	
3	79.8	3.35	m	1, 2, 4	4	
4	71.5	3.42	t, 9.3	3, 5, 6	2(w), 3, 5	
5	82.7	3.21	ddd, 2.6, 4.6, 9.8	1, 3, 4, 6	4, 6a/b	1, 6a/b
6	63.1	a: 3.67	m	4, 5	5	5
		b: 3.64	m	4, 5	5	5
7	165.5					
8	40.50	a: 4.07	AB, 16.4	7, 1', 2', 6'	2'(w)	1, 2', 6'
		b: 4.05	AB, 16.4	7, 1', 2', 6'	2'(w)	1, 2', 6'
1'	133.5					
2'	118.8	6.97	d, 2.1	8, 3'(w), 4', 6'	8a/b(w), 6'	8a/b
3'	149.3					
4'	146.2					
5'	120.8	7.15	d, 8.3	1', 2'(w), 3', 4'	6'	1''
6'	123.2	6.92	dd, 2.1, 8.3	2'(w), 4'	2', 5'	8a/b
1''	110.5	5.69	d, 3.2	4', 2'', 3'', 4''	2''	5', 2''
2''	79.8	4.36	d, 3.2	1'', 3'', 4'', 5''	1'', 4''b(w)	1'', 5''
3''	82.3					
4''	77.3	a: 4.22	AB, 10.2	1'', 2'', 3'', 5''	4''b	
		b: 3.99	AB, 10.2	1'', 2'', 3'', 5''	2'' (w), 4''a	
5''	66.2	a/b: 3.76	s (2H)	2'', 3'', 4''		2''