Journal of Chromatography A, 1217 (2010) 6281–6289

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Complementary fragmentation pattern analysis by gas chromatography–mass spectrometry and liquid chromatography tandem mass spectrometry confirmed the precious lignan content of *Cirsium* weeds

I. Boldizsár[a], M. Kraszni[b], F. Tóth[c], B. Noszál[b], I. Molnár-Perl[d,*]

[a] *Institute of Biology, Department of Plant Anatomy, L. Eötvös University, Budapest 1117, Pázmány Péter sétány 1/C, Hungary*
[b] *Semmelweis University, Department of Pharmaceutical Chemistry, Budapest 1092, Hőgyes Endre utca 7, Hungary*
[c] *Central Service for Plant Protection and Soil Conservation, Chemical Department, Velence 2481, Hungary*
[d] *Institute of Chemistry, Department of Analytical Chemistry, L. Eötvös University, Budapest 112, H-1518, P.O. Box 32, Hungary*

ARTICLE INFO

Article history:
Received 7 June 2010
Received in revised form 5 August 2010
Accepted 6 August 2010
Available online 13 August 2010

Keywords:
Gas chromatography–mass spectrometry
Trimethylsilyl derivatives
Liquid chromatography tandem mass spectrometry
Mass fragmentation
Cirsium species
Lignan composition

ABSTRACT

In this paper, as novelties to the field, it is confirmed at first, that the fruits of *Cirsium* species, regarded as injurious weeds, do contain lignans, two, different butyrolactone-type glycoside/aglycone pairs: the well known arctiin/arctigenin and the particularly rare tracheloside/trachelogenin species. These experiences were supported by gas chromatography–mass spectrometry (GC–MS), by liquid chromatography tandem mass spectrometry (LC–MS/(MS)) and by nuclear magnetic resonance (NMR) spectroscopy. The study reflects the powerful impact of the complementary chromatographic mass fragmentation evidences resulting in the identification and quantification, the extremely rare, with on line technique not yet identified and described, tracheloside/trachelogenin pair lignans, without authentic standard compounds. Fragmentation pattern analysis of the trimethylsilyl (TMS) derivative of trachelogenin, based on GC–MS, via two different fragmentation pathways confirmed the detailed structure of the trachelogenin molecule. The complementary chromatographic evidences have been unambiguously confirmed, by ^{1}H and ^{13}C NMR analysis of trachelogenin, isolated by preparative chromatography. Identification and quantification of the fruit extracts of four *Cirsium* (*C.*) species (*C. arvense*, *C. canum*, *C. oleraceum*, and *C. palustre*), revealed that (i) all four species do accumulate the tracheloside/trachelogenin or the arctiin/arctigenin butyrolactone-type glycoside/aglycone pairs, (ii) the overwhelming part of lignans are present as glycosides (tracheloside 9.1–14.5 mg/g, arctiin 28.6–39.3 mg/g, expressed on dry fruit basis), (iii) their acidic and enzymatic hydrolyses to the corresponding aglycones, to trachelogenin and arctigenin are fast and quantitative and (iv) the many-sided beneficial trachelogenin and arctigenin can be prepared separately, without impurities, excellent for medicinal purposes.

1. Introduction

Cirsium species are characterized as injurious weeds [1]: since, for the time being, it was unknown that their fruits are particularly rich in lignan glycosides.

According to a review on the chemistry and biological activity of 67 *Cirsium* species, it was shown that *Cirsiums*' main extractable constituents are 78 different flavonoids; lignans were not found [2].

Even the latest reports [3–7] confirmed that all the beneficial impacts of different *Cirsium* species are associated with their various flavonoid contents: like aldose reductase inhibitory [3], antimicrobial [4], anti-inflammatory [5], antioxidant [6] and anticancer [7] activities.

Based on our last 4 year experiences on the gas chromatography–mass spectrometry (GC–MS) and liquid chromatography mass spectrometry (LC–MS) of the anthraquinone [8], flavonoid [9] and lignan [10,11] constituents of plant extracts we turned our attention to the analysis of the biologically active compounds of *Cirsium* species: available in unlimited amounts in the plant kingdom, as injurious weeds, foredoomed to eradication.

Introductory results revealed that the fruits of *Cirsium* plants, in contrary to flavonoids, are abundant sources of the special butyrolactone-type glycoside/aglycone pairs of tracheloside/trachelogenin and arctiin/arctigenin lignans. This finding proved to be of particular importance for four reasons (a–d), since

(a) the arctiin/arctigenin containing plants are ubiquitous in the plant kingdom [10,11],
(b) the occurrence of the tracheloside/trachelogenin containing plants are extremely rare [12–23], and their extractable lignan

* Corresponding author. Tel.: +36 1372 26 16; fax: +36 1 372 25 92.
E-mail address: perlne@chem.elte.hu (I. Molnár-Perl).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.08.018

contents, if indicated at all [14,16,17,19,21–23] proved to be low: varying between 3.6×10^{-4} mg/g [19] and 7.1×10^{-1} mg/g [17], respectively. As examples

(c) tracheloside was obtained from the dried leaves and stem of *Trachelospermum asiaticum* (3.2×10^{-2} mg/g) [14] and from the aerial parts of *Saussurea salicifolia* (1.4×10^{-1} mg/g) [22], while,

(d) trachelogenin was isolated from the aerial parts of *Cnicus benedictus* (2.5×10^{-2} mg/g) [16], *Centaurea moesiaca* (1.2×10^{-2} mg/g) [21] and *Saussurea salicifolia* (5.6×10^{-2} mg/g) [22], from the leaves of *Glycydendron amazonicum* (7.1×10^{-1} mg/g) [17] and *Torreya jackii* (3.6×10^{-4} mg/g) [19], as well as from the seeds of *Carthamus tinctorius* (2.2×10^{-1} mg/g) [23]. (*Note:* In order to be comparable, all literature data are given on dry matter basis, in parentheses.)

The beneficial characteristics of arctiin and arctigenin were recently compiled [11]. Tracheloside was described as an anti-estrogenic lignan [20], while trachelogenin exerts numerous beneficial impacts [17,18,23]: it was confirmed to be effective as a Ca^{2+} antagonist [17], as antihypertensitive [17], as in vivo inhibitor of HIV-1 [18], and, as a cancer chemopreventive [23] agent.

These tracheoleside/trachelogenin lignans are not commercially available, and therefore their identification and structural elucidation was performed upon their extraction from various plants using nuclear magnetic resonance spectroscopy (NMR) [12–23]. A single exception [19], the lignans of *Torreya jackii* were identified by stopped-flow high-performance liquid chromatography-^{1H}NMR. To our knowledge, for the time being, no attention was paid on their identification and quantification by on-line chromatography mass spectrometry.

Chromatographic analysis of lignans has been reviewed recently [24], focusing, in general on the extraction/separation, identification and quantification of several lignan containing samples, with special attention on the analysis of tree lignans, determined by gas chromatography isotope dilution mass spectrometry.

The aim of this study was

(a) to prove that the fruits of *Cirsium* species do contain lignans,
(b) to examine their enzymatic and acidic hydrolyses,
(c) to identify and quantify these lignans by on-line techniques, like GC–MS and/or LC–MS/(MS),
(d) to prove the similar fragmentation behavior of the butyrolactone-type lignans by GC–MS and by HPLC–MS/(MS),
(e) to confirm the chemical structure of trachelogenin isolated from *C. arvense* by NMR spectroscopy, as well as
(f) to show the utility of our multicomponent analysis principle in the identification and quantification of further, relevant constituents of the fruit extracts of *Cirscium* species.

2. Experimental

2.1. Materials

Materials and reagents were all analytical reagent grade of the highest purity available, except solvents used for the HPLC-UV–MS: all of them were of HPLC grade chemicals.

Methyl alcohol, pyridine, hydroxylamine hydrochloride, acetonitrile were from Reanal (Budapest, Hungary), hexamethyldisilazane (HMDS) and trifluoroacetic acid (TFA) from Serva (Heidelberg, Germany). Model lignans, such as arctiin, arctigenin, as well as dimethyl sulfoxide-d_6 (DMSO-d_6) were from Sigma–Aldrich (St. Louis, MO, USA). Tracheloside and trachelogenin were isolated from *Cirsium arvense.* Plant samples (fruits of *Cirsium arvense*, further on, *C. arvense*, *C. canum*, *C. oleraceum* and *C. palustre*) were collected from different Hungarian areas, indicated as, *C. arvense*-1-3, *C. canum*-1,2, *C. oleraceum*-1,2 (samples' joint designation: plant samples).

2.2. Chromatography

2.2.1. Gas chromatography–mass spectrometry (GC–MS)

The instrument was the Saturn II GC–MS system of Varian (Walnut Creek, CA, USA), supplied with an ion trap detector (ITD) system, a Varian 8200 autosampler and with a septum-equipped programmable injector (SPI). A SGE BPX5 (30 m, 0.25 mm id, 0.25 μm film thickness) capillary column was used (SGE Incorporated, Austin, TX, USA).

Separation of the TMS (oxime) ether/ester derivatives:

Injector program: injections were made at 150 °C, held at 150 °C for 2 min, heated up to 330 °C within 1 min and hold at 330 °C for 5 min.
Gradient program: separations under temperature gradient conditions were made at 150 °C, held at 150 °C for 4.5 min, heated up to 330 °C within 22.5 min and hold at 330 °C for 7 min (total elution time: 34 min).
MS conditions: electron energy was 70 eV; the temperatures of the transfer line and that of the manifold were 280 °C and 80 °C, respectively. Multiplier offset was 250 eV. The actual parameters of the ITD were defined by the automatic set up mode.
Actual automatic set-up conditions: mass range: 40–650 amu; the scan rate was 1 scan/s acquisition time: 34 min; solvent delay was 420 s (omitting the acquisition of reagent peaks); peak threshold 100 count; mass defect: 100 mmu/100 u; background mass: 50 u.

2.2.2. Preparation of the trimethylsilyl (oxime) ether/ester derivatives

The mixtures of various amounts of standards (arctiin, arctigenin, aliphatic and aromatic carboxylic, saturated and unsaturated fatty acids and sugars), or isolated compounds (tracheloside and trachelogenin), as well the corresponding amounts containing stock solutions of plant extracts were evaporated to dryness on a vacuum evaporator, at 30–40 °C, in 2–4 mL screw capped vials. The dry residues were treated with 250 μL hydroxylamine hydrochloride containing pyridine (2.5 g hydroxylamine hydrochloride/100 mL), at 70 °C, for 30 min. Thereafter silylation was performed with 450 μL HMDS + 50 μL TFA, and heated at 100 °C for 60 min. After dilution with HMDS, 1 μL of the diluted solutions was injected into the GC–MS system.

2.2.3. High-performance liquid chromatography-photo diode array and mass detections (HPLC-UV–MS), applying electrospray ionization (ESI) in the positive mode (HPLC–ESI-MS)

The instrument was the TSQ Quantum AM Triple Quadrupole Mass Spectrometer of Thermo Finnigan (River Oaks Parkway, San Jose, CA, USA), equipped with the Surveyor MS Pump (quaternary), with Surveyor Liquid Chromatography, with autosampler and with Surveyor PDA5 Photo Diode Array Detector.

For the HPLC-UV and HPLC–ESI-MS analyses the same column, the same eluents and the same gradient were used.

Column: GraceSmart RP18 (5 μm), 150 mm × 4.6 mm (Grace Davison Discovery Sciences Lokeren, Belgium);
Eluents: eluent A, ACN: 0.07 M aqueous acetic acid 15:85 (v/v), eluent B, ACN: 0.07 M aqueous acetic acid 85:15 (v/v).
The following solvent gradient was used: eluent A was changed for eluent B. Program: 0 min, 15% B; 5 min, 30% B; 12 min, 44% B and 20 min, 60% B. Total elution time: 20 min; flow rate: 1 mL/min, detection at 280 nm; injected volumes 20 μL. MS parameters: chamber, ES (electrospray); ion mode: positive; scan range of

Table 1
Lignan composition of the nonhydrolyzed (intact), by enzyme (HydrEnz) and by trifluoroacetic acid hydrolyzed (HydrTFA) fruit extracts of *Cirsium* (*C.*) species, analyzed as trimethylsilyl ether derivatives by GC–MS (GC), and without derivatization with HPLC-UV (LC), recorded at 280 nm.

Lignans	Method of analysis, t_R	Concentrations[a] of the lignans in fruits											
		C. arvense-1						*C. oleraceum*-1					
		Intact		HydrEnz		HydrTFA		Intact		HydrEnz		HydrTFA	
Tracheloside	GC, 30.91	13.9	**14.0**	–		–		–		–		–	
	LC, 3.99	14.1	(1.0)	–		–		–		–		–	
Trachelogenin	GC, 23.45	0.70	**0.77**	9.64	**9.70**	9.96	**9.37**	–				–	–
	LC, 7.71	0.83	(12.0)	9.75	(0.8)	8.77	(9.0)	–				–	–
Arctiin	GC, 32.92	–		–		–		36.5	**37.9**	–		–	
	LC, 5.43	–		–		–		39.3	(5.2)	–		–	
Arctigenin	GC, 24.10	–		–		–		1.31	**1.23**	31.4	**30.0**	29.3	**30.2**
	LC, 10.01	–		–		–		1.15	(9.2)	28.5	(6.8)	31.0	(4.0)
Lignans in total, mM/100 g[b]		2.74		2.50		2.41		7.4		8.1		8.1	

Indications: t_R = retention times in Figs. 2 and 3.
[a] Calculated on dried fruit basis, expressed in mg/g; bold printed data represent the average of GC and LC analyses; ( ) = in parentheses: relative standard deviation percentages of the GC and LC data.
[b] Calculated on the basis of the average of GC and LC.

140–700 *m/z*; drying gas N_2: 8 L/min (≈1.5 Torr), used at ambient temperature; capillary voltage: 3000 V; fragmentation voltage (collision energy): 80 V; collision gas was Argon. MS/MS conditions: optimization was based on the fragmentation of aqueous adducts of molecular ions; optimum voltage, in all four cases, proved to be 25 eV.

2.3. Isolation of tracheloside and trachelogenin

For this purpose the analytical HPLC instrument was connected to a preparative HPLC column: Nucleosil100, C18 (10 μm), 15 cm × 1 cm (Teknokroma, Sant Cugat del Vallès, Barcelona, Spain); eluents were the same as described above, while gradient was altered for: 0 min, 15% B; 10 min, 40% B; and 15 min, 100% B; flow rate: 3 mL/min; detection at 280 nm.

2.3.1. Isolation of tracheloside

Lyophilized, pulverized and homogenized fruits of *C. arvense* (2.0 g) were extracted, twice, with 25–25 mL 80% (v/v) methyl alcohol, at 60 °C, applying a reflux condenser. The unified extracts were evaporated to dryness on a vacuum evaporator at 30–40 °C. The residue was dissolved in 2.5 mL 80% (v/v) methyl alcohol, using its 100 μL aliquots for separation.

2.3.2. Isolation of trachelogenin

Lyophilized, pulverized and homogenized fruits of *C. arvense* (2.0 g) was enzymatically hydrolyzed, it means stirred with 5 mL distilled water (40 °C, 60 min). The hydrolyzed sample was vacuum evaporated to dryness (30–40 °C), and residue (from the extraction step), was treated according to Section 2.3.1.

2.4. Preparation of plant extracts

Lyophilized, pulverized and homogenized samples (0.02 g weighed with analytical precision) was extracted three times: first, with 2 mL 80% (v/v) methyl alcohol, applying a reflux condenser (1 h). Thereafter, the insoluble, centrifuged material was extracted for the second time (2 mL) and for the third time (1 mL), as before. The combined supernatants were completed to 5.0 mL stock solutions.

2.5. Hydrolysis of extracts

2.5.1. Performing enzymatic hydrolysis

Lyophilized, pulverized and homogenized 0.02 g fruit samples, weighed with analytical precision, were stirred with 0.5 mL distilled water at 40 °C, for 10 min, 30 min and 60 min. Thereafter, the sample was dried on a vacuum evaporator at 30–40 °C and extracted according to Section 2.4. Dried samples were measured subsequently to their derivatization by GC–MS or without derivatization with HPLC-UV and HPLC-UV–MS (further on: indicated as HydrEnz samples).

2.5.2. Performing acidic hydrolyses

Aliquots of the stock solutions (0.25–1.00 mL, weighed with analytical precision) were evaporated to dryness on a vacuum evaporator at 30–40 °C in 2–4 mL screw capped vials. Hydrolyses were performed with 200 μL 2 M TFA, for 60 min, at 100 °C. Samples were dried on a vacuum evaporator, at 30–40 °C. Dried samples were analyzed subsequently to their derivatization by GC–MS, or without derivatization using HPLC-UV and HPLC–ESI-MS (further on: indicated as HydrTFA samples).

2.6. Nuclear magnetic resonance spectroscopy

NMR spectra of the isolated, both by GC–MS and LC–MS identified trachelogenin were recorded in DMSO-d_6 at 25 °C on a Varian VNMRS spectrometer (599.9 MHz for ^{1}H and 150.9 MHz for ^{13}C) equipped with a dual 5-mm inverse-detection gradient (IDPFG) probehead. Standard pulse sequences and parameters were used to obtain 1D ^{1}H, 1D ^{13}C, COSY, TOCSY, HSQC and HMBC spectra. For the appropriate assignation of the methoxy signals and their position on the aromatic ring band selective HSQC (bsgHSQC) and HMBC (bsgHMBC) spectra were recorded. ^{1}H and ^{13}C chemical shifts were referenced relative to the solvent resonances (δ_H = 2.50 ppm, δ_C = 39.5 ppm).

3. Results and discussion

3.1. Hydrolysis; introductory enzymatic hydrolysis of the fruit extracts of C. arvense-1 and Cirsium oleraceum-1

Enzymatic hydrolyses of the fruit extracts of *C. arvense* and that of *C. oleraceum*, performed as a function of time, was the first

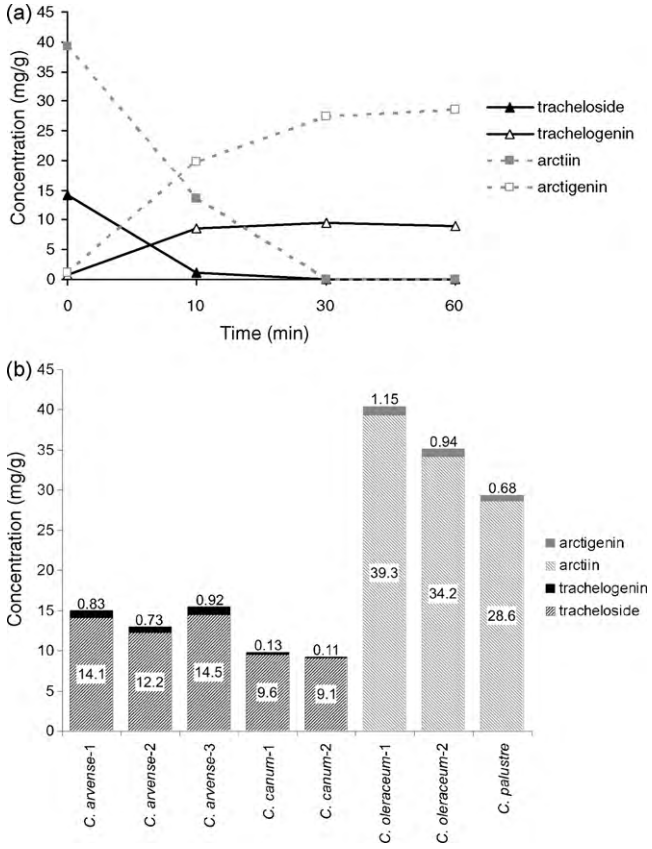

Fig. 1. (a) The yield of enzymatic hydrolyses of the fruit extracts of *C. arvense* (continuous line) and *C. oleraceum* (dashed line), performed as a function of time, and followed by HPLC-UV, at 280 nm. (b) Comparison of the lignan composition analyzed in the fruit extracts of *Cirsium* species, in their initial conditions, followed by HPLC-UV, at 280 nm. Concentrations expressed in mg/g dry basis.

approach of these studies: the yield of active compounds were followed by HPLC-UV (Fig. 1a). Results were unexpected: since, instead of flavonoids, one unknown glycoside/aglycone pair was identified and quantified in the fruit extract of *C. arvense* while the well known arctiin/arctigenin lignans were determined in the fruit extract of *C. oleraceum*.

Due to the fact that the arctiin/arctigenin standards are available, even their quantification could be confirmed and expressed as mg/g dry fruit basis, according to our earlier experiences [10,11]. In the first step, the unknown glycoside/aglycone pair of *C. arvense* was quantified with the arctiin/arctigenin responses, taking into account their – by our GC–MS fragmentation pattern analysis – assumed molecular weights (Section 3.3). (*Note:* the isolated and by ^{1}H and ^{13}C NMR confirmed tracheloside/trachelogenin lignans reassured our assumption: the authentic tracheloside/trachelogenin lignans provided the calculated response values, no correction was needed.) The quantitative transformations both of tracheloside to trachelogenin (Fig. 1a: *C. arvense*-1) and the arctiin to arctigenin (Fig. 1a: *C. oleraceum*-1) proved to be fast, and within 30 min hydrolysis time, quantitative.

Introductory comparisons performed with HPLC-UV and by GC–MS (Table 1), were convincing and promising:

(a) demonstrating the stability of the aglycones even upon acidic hydrolyses,
(b) proving the fact that upon enzymatic hydrolyses both trachelogenin and arctigenin can be prepared easily, in pure form, without loss (last horizontal line: mM/100 g values) and without any disturbing extracts' impurity, and
(c) in relatively high concentrations. The first two peculiarities {points (a), (b)} are of primary importance: revealing the possibility to isolate these lignans in pure form, immediately suitable for preparation of medicinal purposes.

Table 2
Fragmentation pattern analysis of the trimethylsilyl derivatives of lignans.

Lignans (Mw)	Selective fragment ion (SFI), *m/z*	Relative abundance of SFI-s (in percentages of TIC)	
		Average^{n}	RSD%
Tracheloside (520)	151	12.8	3.2
	361	11.3	3.6
	532	8.1	3.3
	245	8.0	4.0
	451	6.8	4.4
	209	6.3	4.8
	SFI in total:	53.3	
Trachelogenin (358)	209	27.9	3.5
	245	9.8	3.9
	151	7.0	0.83
	532	1.91	4.7
	179	1.72	2.9
	SFI in total:	48.3	
Arctiin (534)	444	17.6	1.2
	361	9.7	7.8
	151	3.59	2.8
	451	3.39	5.4
	209	0.95	4.6
	SFI in total:	35.3	
Arctigenin (372)	444	16.6	0.9
	209	16.2	2.8
	151	5.1	3.0
	179	3.18	3.4
	SFI in total:	41.0	

Indications: TIC = total ion current; (Mw) = molecular weight; average^{n} = at least of three separate derivatization and two injections of 2.5, 5 and 10 ng of each.

3.2. Composition of the fruit extracts of additional Cirsium species

Results demonstrated that in all cases investigated (*C. arvense*-2,3, *C. oleraceum*-2, *C. canum*-1,2 and *C. palustre*) the overwhelming part of *Cirsium* lignans is present initially as lignan glycosides (Fig. 1b: expressed as mg/g on dried fruit basis).

(a) Arctiin contents of *C. oleraceum* and *C. palustre* varied between 28.6 mg/g and 39.3 mg/g, while their free arctigenin content between a 0.68 mg/g and 1.15 mg/g.
(b) The tracheloside/trachelogenin contents of *C. arvense* and *C. canum* revealed the same trend: providing from 9.1 mg/g to 14.5 mg/g tracheloside and from 0.11 mg/g to 0.92 mg/g trachelogenin, respectively.

3.3. Identification studies performed by gas chromatography–mass spectrometry

Due to our earlier experiences, relating to the GC–MS retention and mass fragmentation properties of the butyrolactone-type lignan glycosides/aglycones [10,11], their presence in the fruit extracts' derivatives of the *C. arvense and C. oleraceum* seemed to be unambiguous (Fig. 2, traces 1–4, Spectra 1A, 1B, 2A, 2B; *Note:* C atom numbering of the butyrolactone-type structure is shown in detail in Fig. 4).

The (initially) unknown glycoside/aglycone (tracheloside/trachelogenin) pair (Fig. 2, unhydrolyzed extract: trace 1, spectrum 1A; enzyme hydrolyzed extract: trace 2, spectrum 1B) was assumed to be an additional free hydroxyl group containing version of the dibenzylbutyrolactone-type arctiin/arctigenin lig-

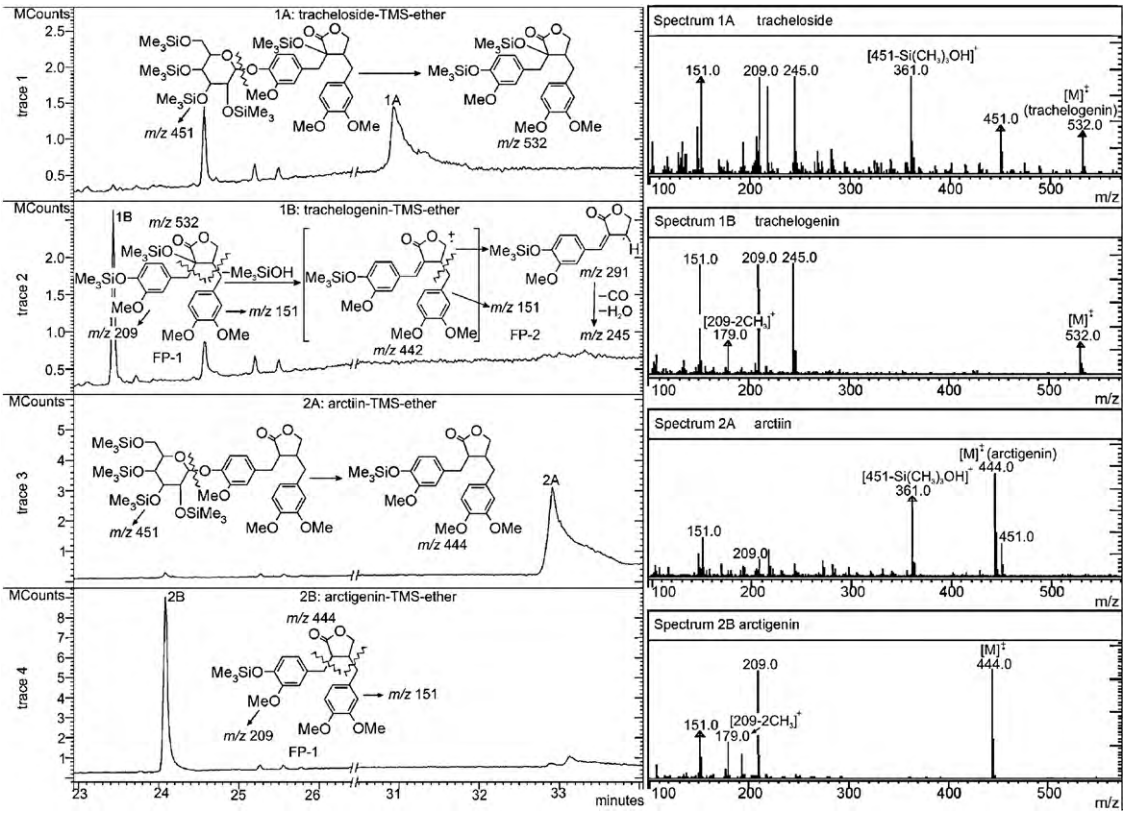

Fig. 2. Peak profile, fragmentation patterns and mass spectra of the trimethylsilylated fruit extracts of *Cirsium* species, followed by gas chromatography–mass spectrometry; obtained from *C. arvense*-1, from its intact (trace 1, spectrum 1A) and from its enzyme hydrolyzed fruit extracts (trace 2, spectrum 1B), as well as, from *C. oleraceum*-1, from its intact (trace 3, spectrum 2A) and from its enzyme hydrolyzed extracts (trace 4, spectrum 2B). Indications: traces 2 and 4: FP-1, fragmentation pathway-1; FP-2, fragmentation pathway-2.

nans. The reasons for this assumption could be characterized by the facts that

(a) the difference in their molecular masses (trachelogenin $[M]^{\ddagger}$ = *m*/*z* 532, arctigenin $[M]^{\ddagger}$ = *m*/*z* 444) do reflect in trachelogenin the presence of an additional free hydroxyl, i.e., one more TMS-O group,
(b) the selective fragment ions (SFIs), formed in fragmentation process-1 (FP-1) proved to be identical, due to the same fragmentation behavior of their corresponding C^1–C^7 and $C^{1'}$–$C^{7'}$ benzylic moieties of trachelogenin and artigenin, characteristic to the dibenzylbutyrolactone-type structure (Fig. 2, traces 2, 4, FP-1, spectrum 1B, 2B: *m*/*z* 151, *m*/*z* 209), albeit, present in various abundances (Table 2), and,
(c) as considerable difference in case of trachelogenin, in addition to their common fragment ions (*m*/*z* 151, *m*/*z* 209), the mass of *m*/*z* 245 emerges with significant intensity. This mass (*m*/*z* 245) corresponds to the product of a specific molecular rearrangement (Fig. 2, trace 2, spectra 1A, 1B), described for the analogous nortrachelogenin [25], consisting of the same hydroxyl group containing lactone ring, predestinated to the same molecular rearrangement.

Fragment ion *m*/*z* 245 originates in the fragmentation process-2 (Fig. 2, trace 2: FP-2) via the transient ion (*m*/*z* 442) formed from the molecular ion by the loss of TMS-OH (*m*/*z* 442 = 532–90). Transient product's (*m*/*z* 442) further transformation (accompanied with the elimination of one molecule H_2O and one molecule CO: Fig. 2, trace 2, FP-2, Spectra 1A, 1B), provides the stable ion *m*/*z* 245.

On the basis of abundances of SFIs (Table 2, Fig. 2 spectra 1A, 1B), it is of interest to estimate the probability of fragmentation pattern pathways: immediately from the molecular ion (*m*/*z* 532: FP-1) and after the loss of TMS-OH followed by molecular rearrangement (from the ion *m*/*z* 442: FP-2). Assuming that all the end product ions (*m*/*z* 151, *m*/*z* 209, *m*/*z* 245) manifest the same stability the fragment intensity of the ion *m*/*z* 151 should be the total of the ions *m*/*z* 209 and *m*/*z* 245. Although, this is not the case. The intensity of *m*/*z* 151 both in comparison to the mass *m*/*z* 209 (FP-1) and to the total of *m*/*z* 209 and *m*/*z* 245 (FP-2) is lower that the calculated one, and dependent where it is originated from (Table 2: relative abundances).

Regarding the 12.8% relative abundance of the fragment ion *m*/*z* 151 obtained from tracheloside, its intensity proved to be somewhat lower only, compared to the expected total, i.e., to the 14.3% abundant intensity (6.3% for *m*/*z* 209 and 8.0% for *m*/*z* 245).

Originating the fragment ion *m*/*z* 151 from trachelogenin – compared to the total intensities of fragment ions *m*/*z* 209 (27.9%) and *m*/*z* 245 (9.8%) – it is infinitesimally scarce (7.0%), less than the one fifth (27.9 + 9.8% = 37.7%) of the calculated one. This phenomenon can be explained with the direct loss of mass *m*/*z* 151 from the TMS glycoside derivative, prior to its special behavior [11], which means its fragmentation followed by immediate derivatization of the released trachelogenin resulting in its fully silylated species (Fig. 2, trace 1, spectrum 1A).

Arctiin and arctigenin were identified and quantified in the fruit extracts of *C. oleraceum* (Fig. 2, unhydrolyzed extract: trace 3, enzyme hydrolyzed extract: trace 4). On the basis of their retention and fragmentation properties, the arctiin/arctigenin pair could be confirmed by their molecular ions $[M]^{\ddagger}$ = *m*/*z* 444 and by their SFIs (*m*/*z* 151, *m*/*z* 209) specific to the process according to the FP-1, both of their C^1–C^7 and $C^{1'}$–$C^{7'}$ benzylic moieties (Fig. 2, traces 3, 4 spectra 2A, 2B).

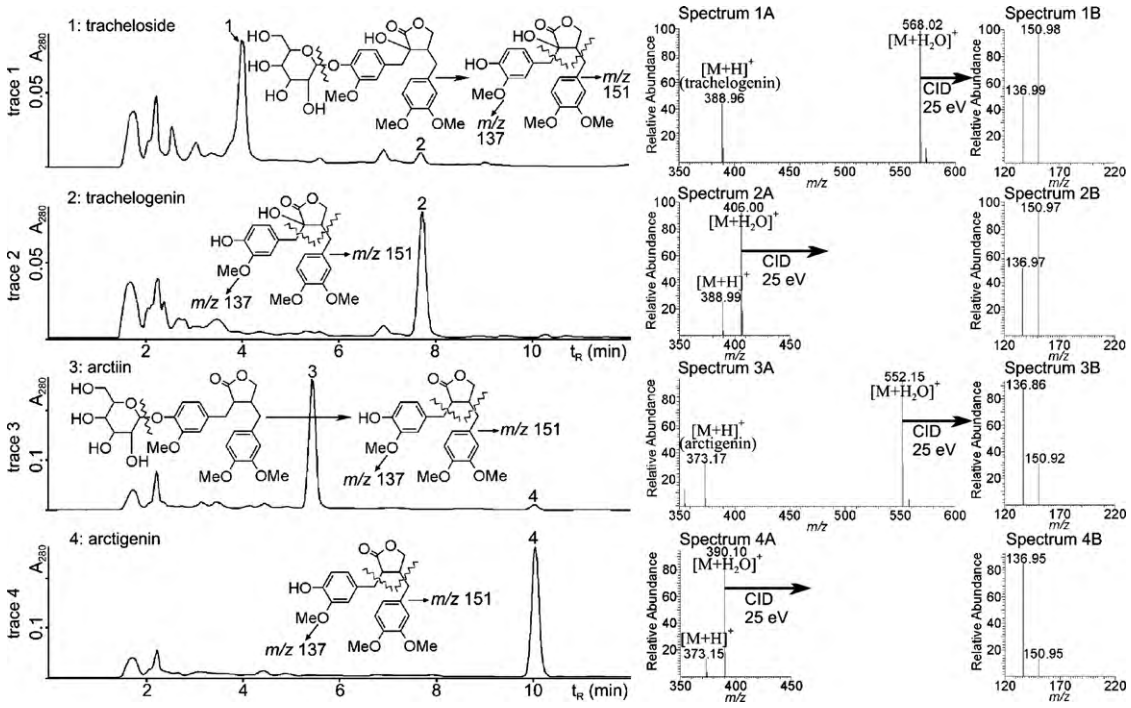

Fig. 3. Peak profile, fragmentation patterns and mass spectra of underivatized lignans obtained from the fruit extracts of *Cirsium* species, followed by HPLC–ESI-MS/(MS). Prepared from *C. arvense*-1, from its intact (trace 1, spectra 1A, 1B) and from its enzyme hydrolyzed fruit extracts (trace 2, spectra 2A, 2B), as well as, from *C. oleraceum*-1, from its intact (trace 3, spectra 3A, 3B) and from its enzyme hydrolyzed extracts (trace 4, spectra 4A, 4B). Spectra 1A–4A represent the LC–ESI-MS spectra, while spectra 1B–4B the LC–ESI-MS/(MS) fragmentation patterns obtained under optimized (25 eV) collision induced dissociation.

Table 3
Assigned ^{1}H and ^{13}C chemical shifts for trachelogenin isolated from *C. arvense*.

Atom	δ ^{1}H (J_{HH} in Hz)	δ ^{13}C
1		131.7
2	6.68 *d* (1.8)	112.5
3		148.7
4		147.2
5	6.83 *d* (8.1)	111.9
6	6.64 *dd* (8.1, 1.8)	120.4
7	2.62 *dd* (13.5, 4.0)	30.7
	2.42 *dd* (13.5, 10.1)	
8	2.36 *m*	42.8
9	3.94 *d* (8.3)	69.9
1′		126.4
2′	6.74 *d* (1.7)	114.5
3′		147.2
4′		145.4
5′	6.67 *d* (8.0)	115.2
6′	6.58 *dd* (8.0, 1.7)	122.6
7′	2.98 *d* (13.6)	40.2
	2.82 *d* (13.6)	
8′		75.4
9′		177.9
3-OMe	3.71 *s*	55.33
4-OMe	3.70 *s*	55.49
3′-OMe	3.69 *s*	55.54
4′-OH	8.87 *br s*	
8′-OH	6.21 *br s*	

Indications: trachelogenin's structure: Fig. 4; completed explanations: Section 3.5.

It is worth to highlight, the recently published [11] and in this study repeatedly confirmed phenomenon related the TMS derivatives of lignan glycosides: this means that under their gas chromatographic mass spectrometric elution condition (at 330 °C), at their expected t_R values (tracheloside at 30.91 min: Fig. 2, trace 1, spectrum 1A; arctiin at 32.92 min: Fig. 2, trace 3, spectrum 2A), their fragmentation into the corresponding free lignans and glucose is going on. The fragment species of the TMS derivatives of lignan glycosides, in the presence of excess hexamethyldisilazane leaded to the instantaneous derivatization of the released aglycones resulting in the fully silylated trachelogenin (spectrum 1A: *m*/*z* 532) and arctigenin (spectrum 2A: *m*/*z* 444). The ring structure of the fully silylated TMS glucose derivative (*m*/*z* 451) and its fragment ion formed by the loss of one TMS-OH group (*m*/*z* 361), are characteristic masses in the spectra of both glycosides (Fig. 2, spectra 1A and 2A).

Based on the above detailed experiences the presence of lignans was unambiguously confirmed, along with the structure elucidation of tracheloside and trachelogenin.

3.4. Identification studies performed by liquid chromatography tandem mass spectrometry (HPLC–ESI-MS/(MS))

For this study the same extracts were used as for the GC–MS fragmentation pattern analysis (Section 3.3).

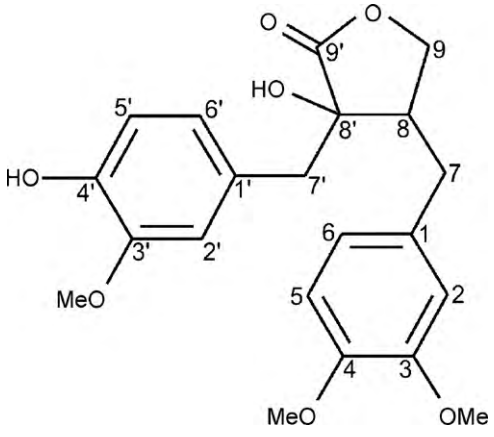

Fig. 4. Molecular structure of trachelogenin confirmed by ^{1}H and ^{13}C NMR spectroscopy (details in Table 3).

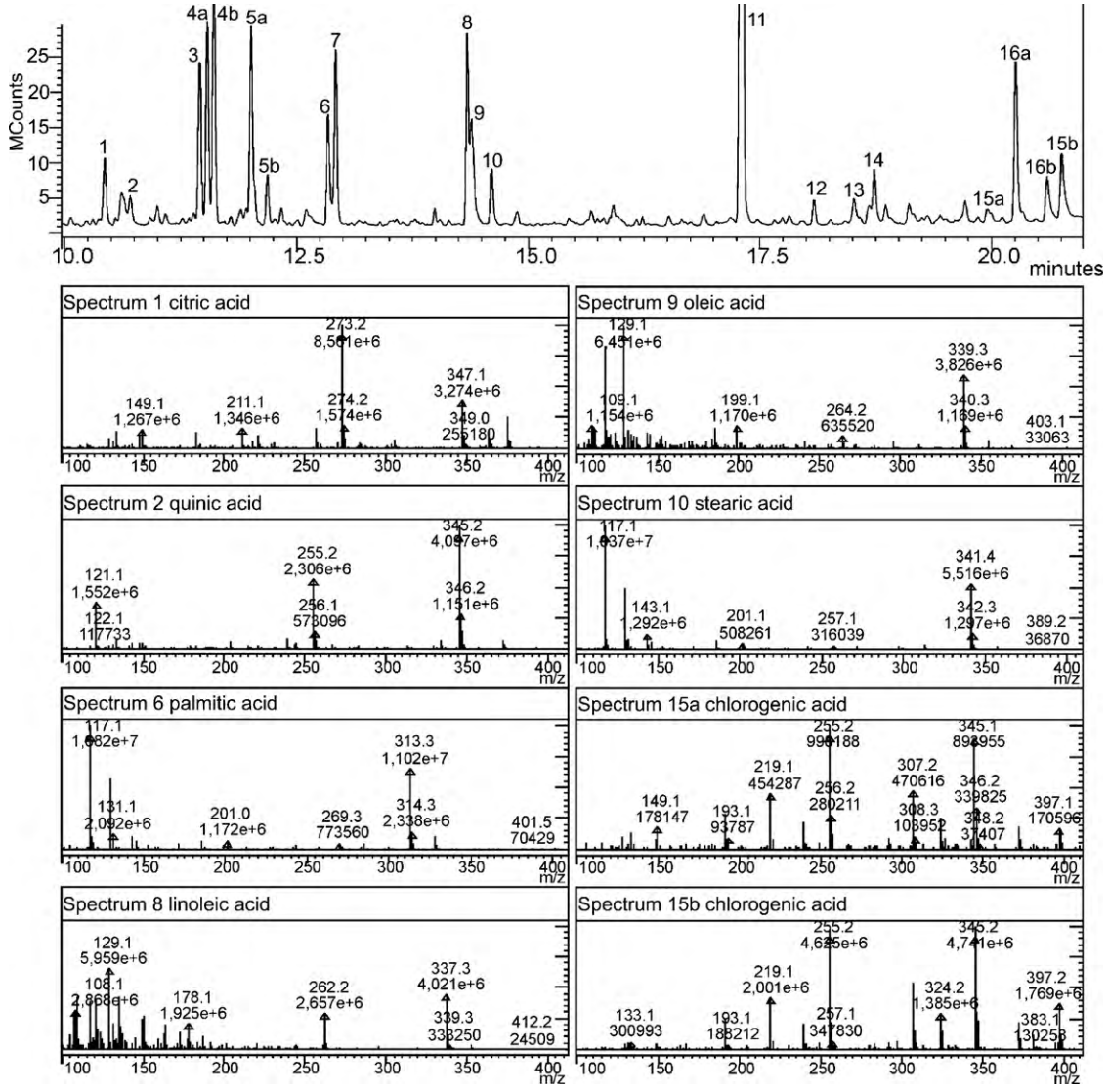

Fig. 5. Elution profile and mass spectra of the trimethylsilyl (oxime) ether/ester derivatives obtained from the intact fruit extract of *C. arvense*-1, determined by GC–MS of in the range between t_R, 10 and 21 min; indications: identity of peak numbers (1–20), listed according to their retention times (including lignan glycosides and free lignans, shown at Fig. 2), are detailed in Table 4.

3.4.1. HPLC–ESI-MS

The HPLC–ESI-MS investigations (Fig. 3, traces 1–4, Spectra 1A–4A) enabled confirmation of GC–MS identification (Fig. 2) and also for the prepared lignans' identity, shown by GC–MS.

Evaluating first the fragment ions originating from the HPLC–ESI-MS fragmentations, it turned out that our GC–MS evidences have been confirmed in all cases investigated.

Abundant ions, without exception, in all spectra (Fig. 3 spectra 1A–4A) were the hydrated molecular ions ($[M+H_2O]^+$, Fig. 3: spectrum 1A, *m*/*z* 568.0 for tracheloside, spectrum 2A, *m*/*z* 406.0 for trachelogenin, spectrum 3A, *m*/*z* 552.1 for arctiin and spectrum 4A, *m*/*z* 390.1 for arctigenin). The second abundant masses were the protonated molecular ions of free lignans ($[M+H]^+$, Fig. 3: spectrum 1A, *m*/*z* 388.96 and spectrum 2A, *m*/*z* 388.99 for trachelogenin, spectrum 3A, *m*/*z* 373.17 and spectrum 4A, *m*/*z* 373.15 for arctigenin). Glycosides do not provide protonated species.

3.4.2. HPLC–ESI-MS/(MS)

Further fragmentations of the hydrated molecular ions, selected as precursors (Fig. 3, spectra 1B–4B) have been performed in all four cases under optimized collision induced dissociation (CID: 25 eV). The corresponding product ions as obtained from the MS/MS fragmentations were formed as follows:

(a) the two methoxy groups containing fragment ions (*m*/*z* 151), originating from the C^1–C^7 benzylic moiety, as expected, should be identical with those obtained by GC–MS, while,
(b) the other fragment ion, formed from the free hydroxyl and methoxy group bearing $C^{1'}$–$C^{7'}$ benzylic moiety showed up in their initial forms, as *m*/*z* 137, in agreement with its silylated version obtained as *m*/*z* 209 (137 + 72) by GC–MS.
(c) The analogous presence of the fragment ion *m*/*z* 245, originating from the lactone ring, and being associated with the intrinsic feature of the rearrangement of the TMS derivatives exclusively does not have any alternative by the underivatized sample analyzed by LC–MS.

As concordant consequences of the GC–MS and that of the LC–ESI-MS/(MS) fragmentation pattern analyses it was clear that the additional hydroxy group containing lignans are the trachelo-

Table 4
Composition of the intact, and by enzyme hydrolyzed (HydrEnz) fruit extracts of *Cirsium* (*C.*) species, analyzed as trimethylsilyl (oxime) ether/ester derivatives, by GC–MS.

Compounds	Concentrations[a] in the fruit extracts				
Peak number/name	t_R, min	*C. arvense*-1		*C. oleraceum*-1	
		Intact	HydrEnz	Intact	HydrEnz
1/Citric acid	10.43	2.53	1.58	2.12	1.74
2/Quinic acid	10.71	0.56	0.73	1.01	1.73
3/Sorbitol	11.45	2.80	7.6	1.47	11.3
4/Fructose[b]	11.46; 11.54	6.1	16.3	3.32	24.9
5/Glucose[b]	12.02; 12.20	4.22	25.8	3.03	53.6
6/Palmitic acid	12.84	11.2	8.5	8.7	9.6
7/Inositol	12.92	2.83	3.38	2.26	2.80
8/Linoleic acid	14.34	22.6	18.6	8.6	9.6
9/Oleic acid	14.39	17.8	15.0	9.8	11.4
10/Stearic acid	14.60	7.1	5.4	7.1	6.6
11/Sucrose	17.29	26.0	12.3	52.9	29.2
12/Trehalose	18.09	0.50	0.26	0.050	0.045
13/, 14/Disaccharides[ni]	18,52; 18.74	2.14	2.49	3.08	3.34
15a/, 15b/Chlorogenic acid	19.96; 20.76	1.27	0.021	2.30	0.42
16a/, 16b/Benzoic acid[ni]	20.25; 20.60	3.14	3.11	0.28	0.30
17/Trachelogenin	23.45	0.70	9.64	–	–
18/Arctigenin	24.10	–	–	1.31	31.4
19/Tracheloside	30.91	13.9	–	–	–
20/Arctiin	32.92	–	–	36.5	–
Found in the fruits, in total		125.4	130.7	143.8	198.0
Identified in the extract, %		55.3	54.2	63.4	66.9

[a] Calculated on the dried fruit basis, expressed in mg/g, performed in parallel tests (relative standard deviation percentages varied between 1.51% and 6.3%); peak numbers, indicate constituents obtained from the same chromatogram in Fig. 5, peaks 1–16 from the elution range of 10–21 min) and in Fig. 2, peaks 17–20 from the elution range of 23–34 min, listed according to their retention time (t_R, min).
[b] The E/Z trimethylsilyl (oxime) ether derivatives of fructose and glucose; 15a, 15b, 16a, 16b stereoisomers of chlorogenic acid (15) and not identified[ni], probably benzoic acid containing (16) derivatives with identical fragment ions (*m/z* 223, 267, 297).

side/trachelogenin species. However, remaining on the safe side, as final confirmation the isolated lignan obtained from the enzyme hydrolyzed extract of *C. arvense* was subjected to a detailed NMR study.

3.5. Structural identification of the lignans of C. arvense-1 applying ^1H and ^{13}C NMR spectroscopy

The NMR spectroscopy provided supportive evidence for the tentative identifications by the GC–MS and reassured by the LC–MS and LC–MS/(MS) analyses.

From the HSQC and TOCSY spectra it was obvious that there is a spin system of five aliphatic hydrogens in the molecule, corresponding to a CH_2–CH–CH_2 substructure. Thus, the only possible position for –OH is $C^{8'}$, which was also confirmed by the multiplicity pattern of the hydrogens and the higher chemical shift of $C^{8'}$ containing the hydroxyl group (Fig. 4).

The second question to be solved was the alignment of the dimethoxy-, methoxy- and hydroxyl-substituted benzyl rings, and also the substituent positions. The multiplicity pattern of the aromatic protons (Table 3) showed *meta* and *para* substitution in both rings. From the bsgHMBC spectra we could find the long range correlation between the methyl protons and the aromatic carbons attached to the methoxy groups (3.71–148.7; 3.70–147.2; 3.69–147.2). From the HMBC spectrum the long range correlation of the H6 and H6′ protons unambiguously assigned the carbons in *para* position which were the methoxy-substituted carbon at 147.2 ppm and the hydroxyl-substituted carbon at 145.4 ppm.

These results confirmed that the unknown aglycone was the trachelogenin (Fig. 4) whose complete ^1H and ^{13}C NMR assignation (Table 3) is in good agreement with literature data [17].

3.6. Identification and quantification of the constituents in the fruit extracts of C. arvense-1 and C. oleraceum-1 determined as their TMS (oxime) ether/ester derivatives

The practical utility of our multicomponent analysis system [8–11] was shown by the identification and quantification of the precious fruit compounds of *C. arvense*-1 and *C. oleraceum*-1, from a single solution by one injection (Table 4, Fig. 5). It was published recently [26] that different species of the genus *Cirsium* are known to be appetizer and tonic; in addition they exert antiviral activity and cytotoxicity, assumed to be associated with their fatty acid contents. In agreement with this paper we do find several saturated and unsaturated fatty acids, citric and quinic acids, the very special, free radical scavenger chlorogenic acid, poly alcohols and sugars (Table 4: compounds 1–20, Fig. 5: selected spectra of acids).

It is worth to remark that the identifiable part of the extracts proved to be in average of intact and enzyme hydrolyzed extracts 54.8 (*C. arvense*-1) and 65.2% (*C. oleraceum*-1), respectively.

4. Conclusion

1. Complementary chromatographic studies {GC–MS, HPLC-UV–MS, LC–MS/(MS)}, supplemented by ^1H and ^{13}C NMR investigations, it has been proved that the injurious weeds, native throughout the globe, everywhere in the roadside, contain precious lignans: the well known butyrolactone-type arctiin/arctigenin and the scarcely found, as authentic standards not available, tracheloside/trachelogenin, glycoside/aglycone pairs.
2. The characteristic behavior and specific fragmentation pathways of the TMS ether derivatives of the tracheloside/trachelogenin pair were shown by GC–MS and as underivatized species by HPLC-UV–MS and LC–MS/(MS). These two chromatographic techniques, in a synergistic combination proved the structure of the tracheloside/trachelogenin lignans.

3. The NMR spectroscopy provided supportive evidence for the tentative identifications by the GC–MS and reassured by the LC–MS and LC–MS/(MS) analyses.
4. As to the practical utility of the study, it has been confirmed that
   (a) the overwhelming parts of *Cirsium* lignans are present initially as lignan glycosides,
   (b) glycosides' hydrolyses both acidic and enzymatic are fast, resulting in stable free lignans (trachelogenin, arctigenin), in high concentrations;
   (c) one of the main advantages of *Cirsiums*' lignans can be attributed to the fact that upon enzymatic hydrolyses both trachelogenin and arctigenin can be prepared in pure form, without loss, separately. These peculiarities are of primary importance, revealing the possibility to isolate these free lignans, immediately suitable for preparation of medicinal purposes.
5. As a completing practical utility of our multicomponent analysis system, together with lignans, further, relevant constituents, have been identified and quantified, as their trimethysilyl (oxime) ether/ester derivatives, by GC–MS: citric, quinic, chlorogenic acids, C_{16} and C_{18} fatty acids, sugars, sugar alcohols, in total 20 compounds, simultaneously.

References

[1] PLANTS Profile for Cirsium vulgare (bull thistle) USDA PLANTS, USDA Plant Database, http://plants.usda.gov/java/profile-symbol=CIVU. 2008-07-15.
[2] I.E. Jordon-Thaden, S.M. Louda, Biochem. Syst. Ecol. 31 (2003) 1353.
[3] H.A. Jung, Y.S. Kim, J.S. Choi, Food Chem. Toxicol. 47 (2009) 2790.
[4] J. Nazaruk, S.K. Czechowska, R. Markiewicz, M.H. Borawska, Nat. Prod. Res. 22 (2008) 1583.
[5] H. Lim, K.H. Son, H.W. Chang, K. Bae, S.S. Kang, H.P. Kim, Biol. Pharm. Bull. 31 (2008) 2063.
[6] Y.M. Yoo, J.H. Nam, M.Y. Kim, J. Choi, H.J. Park, Biol. Pharm. Bull. 31 (2008) 760.
[7] S. Liu, J. Zhang, D. Li, W. Liu, X. Luo, R. Zhang, L. Li, J. Zhao, Nat. Prod. Res. 21 (2007) 915.
[8] I. Boldizsár, Z. Szűcs, Zs. Füzfai, I. Molnár-Perl, J. Chromatogr. A 1133 (2006) 259.
[9] Zs. Füzfai, I. Molnár-Perl, J. Chromatogr. A 1149 (2007) 88.
[10] É. Sedlák, I. Boldizsár, L. Borsodi, Zs. Füzfai, I. Molnár-Perl, É. Preininger, I. Gyurján, Chromatographia 68 (2008) S35.
[11] I. Boldizsár, Zs. Füzfai, F. Tóth, É. Sedlák, L. Borsodi, I. Molnár-Perl, J. Chromatogr. A 1217 (2010) 1674.
[12] I. Inagaki, S. Hisada, S. Nishibe, Phytochemistry 10 (1971) 211.
[13] S. Nishibe, S. Hisada, I. Inagaki, Phytochemistry 10 (1971) 2231.
[14] I. Inagaki, S. Hisada, S. Nishibe, Chem. Pharm. Bull. 20 (1972) 2710.
[15] S. Nishibe, S. Hisada, I. Inagaki, Chem. Pharm. Bull. 21 (1973) 1108.
[16] M. Vanhaelen, R. Vanhaelen-Fastré, Phytochemistry 14 (1975) 2709.
[17] L.M.D. John, W.F. Tinto, J. Nat. Prod. 55 (1992) 1313.
[18] J.L. Charlton, J. Nat. Prod. 61 (1998) 1447.
[19] Y. Zhao, A. Nookandeh, B. Schneider, X. Sun, B. Schmitt, J. Stöckigt, J. Chromatogr. A 837 (1999) 83.
[20] H.H. Yoo, J.H. Park, S.W. Kwon, Biosci. Biotechnol. Biochem. 70 (2006) 2783.
[21] A. Trendafilova, M. Todorova, S. Bancheva, Biochem. System. Ecol. 35 (2007) 544.
[22] K. Kang, H.J. Lee, C.Y. Kim, S.B. Lee, J. Tunsag, D. Batsuren, C.W. Nho, Biol. Pharm. Bull. 30 (2007) 2352.
[23] E.O. Kim, K.S. Kim, W.J. Lee, S.W. Choi, Food Sci. Biotechnol. 18 (2009) 689.
[24] S.M. Willför, A.I. Smeds, B.R. Holmbom, J. Chromatogr. A 1112 (2006) 64.
[25] T. Okunishi, N. Takaku, P. Wattanawikkit, N. Sakakibara, S. Suzuki, F. Sakai, T. Umezawa, M. Shimada, J. Wood Sci. 48 (2002) 237.
[26] I. Orhan, D. Deliorman-Orhan, B. Özçelik, Food Chem. 115 (2009) 701.