Selective Profiling of Saponins from Gypsophila trichotoma Wend. by HILIC Separation and HRMS Detection

Reneta Gevrenova, Ventzislav Bardarov, Krum Bardarov, Laurence Voutquenne-Nazabadioko, Max Henry

Selective profiling of saponins from *Gypsophila trichotoma* Wend. by HILIC separation and HRMS detection

Reneta Gevrenova[1], Ventzislav Bardarov[2], Krum Bardarov[3], Laurence Voutquenne-Nazabadioko[4], Max Henry[5]

[1] Faculty of Pharmacy, Medical University-Sofia, 2 Dunav str., 1000 Sofia, Bulgaria
[2] Hromana Ltd, 78 Samokov str., 1113 Sofia, Bulgaria
[3]Inobiotech, 78 Samokov str., 1113 Sofia, Bulgaria
[4] Institut of Molecular Chemistry of Reims, UMR CNRS 7312, Bâtiment 18, BP 1039, 51687 Reims cedex 2, France
[5] MoBAT team – SRSMC UMR 7565 CNRS-Université de Lorraine, F-54506 Vandoeuvre les Nancy cedex, France

Corresponding author: Reneta Gevrenova
Tel.: +359 2 9236 531; fax: +359 2 9879 874
E-mail: rgevrenova@gmail.com

ABSTRACT

Introduction: Roots of *Gypsophila trichotoma* Wend. (Caryophyllaceae) are rich source of **G**lucuronide **O**leanane-type **T**riterpenoid **C**arboxylic **A**cid 3,28-*O*-**B**idesmosides (GOTCAB). These saponins have been reported to possess synergistic cytotoxicity in combination with type I ribosome-inactivating protein saporin.

Objective: To develop ultra high-performance liquid chromatography – electrospray/high resolution mass spectrometry (UHPLC-ESI/HRMS) acquisition strategy for the recognition of *Gypsophila* GOTCAB saponins.

Methodology: A highly-selective hydrophilic interaction UHPLC method (Si-HILIC UHPLC) was developed for the separation of GOTCAB saponins from the methanol-aqueous root extract of *G. trichotoma* (GTR). UHPLC was coupled to an Orbitrap mass spectrometer equipped with heated electrospray ionization (HESI) probe. ESI-HRMS and tandem MS/MS data of the separated compounds was used for saponins structure assignment.

Results: Based on the conformity of the fragmentation of eleven previously identified *G. trichotoma* saponins, 21 GOTCAB forming between 2 and 4 isobaric and positional isomers are identified with proposals for their structures. Tables with assignment of characteristic fragment ions and more than 10 newly identified saponins in GTR were described. Fragmentation rules for tentative identification of three major types of saponins from GTR were summarized and possible fragmentation pathways were proposed. Type I and II consisted of acylated and sulfated GOTCAB, respectively, while type III included acylated and sulfated saponins. The type II sulfated GOTCAB saponins were all new compounds.

Conclusions: The study demonstrates the potential of the coupling of high-selective (Si)-HILIC UHPLC with HRMS and MS/MS detection for analysis and identification of triterpenoid saponins.

KEY WORDS: GOTCAB saponins, (Si) HILIC UHPLC, ESI-HRMS, structure interpretation

INTRODUCTION

Saponins are secondary metabolites widespread distributed mainly in plant species and chemically referred to triterpenoid and steroidal glycosides (Vincken *et al.*, 2007). Indeed, saponins have potential industrial application as emulsifiers, preservatives, additives to foods and pharmaceutical products (Güçlü-Ustundag *et al.*, 2007; Piorkowski and McClements, 2014). Saponins have been ascribed a number of pharmacological activities, the most important ones being lowering of serum cholesterol levels (Francis *et al.*, 2002), cytostatic and cytotoxic effects on malignant tumor cells (Bachran *et al.*, 2008; Podolak *et al*., 2010). In addition, saponins combined with conventional chemotherapeutic agents improved inhibitory effect on tumor growth *in vitro* and *in vivo* (Fuch *et al.*, 2009). Moreover, saponin-adjuvanted particulate vaccines were reported as immunostimulatory complexes to have great potential as cancer immunotherapeutics (Scene and Sutton, 2006). Among other interesting properties of saponins are their role in permeabilization of the cell membrane (Gilabert-Oriol *et al.*, 2013) and synergisting enhancement of the toxicity of immunotoxines (Bottger *et al.*, 2013; Roger *et al.*, 2016).

Gypsophila species roots are extremely rich source of Glucuronide Oleanane-type Triterpenoid Carboxylic Acid 3,28-*O*-Bidesmosides (GOTCAB), having a glucuronic acid moiety at C-3 hydroxyl group of the aglycone (Tan *et al.*, 1999; Henry, 2005; Bottger and Melzig, 2010; Zhang *et al.*, 2013). Previous reports have shown that saponins from *Gypsophila oldhamiana* and *G. pilulifera* display cytotoxic activity against different human cancer cell lines (Bai et al, 2007; Arslan et al, 2012; Zhang et al, 2013). It has also been reported that *G. paniculata* and *G. arrostii* var. *nebulosa* saponins enhance the cytotoxicity of the type I ribosome – inactivating protein (RIP-I) saporin from *Saponaria officinalis* (Weng et al, 2008; Arslan et al, 2013).

In our previous study eleven new GOTCAB saponins were isolated from the roots of *Gypsophila trichotoma* Wend. *var. trichotoma* originated from Bulgaria (Voutquenne-Nazabadioko *et al.*, 2013). They have a commonly gypsogenin as aglycone substituted at C-3 with a branched trisaccharide chain ending with variable pentose unit (xylose/arabinose) and at C-28 with an ester-bonded oligosaccharide chain possessing methoxycinnamoyl, acetyl and (or) sulfate groups.

Recently, the synergistic cytotoxicity of nine newly isolated GOTCAB saponins from *G. trichotoma* in combination with type I ribosome-inactivating protein (RIP-I) was evaluated *in vitro* on human breast cancer cell line and quantitative structure-activity relationship (QSAR) was established (Gevrenova *et al.*, 2015).

Based on these findings, *Gypsophila* saponins are important analytical targets, needing of effective tools for their separation, identification and quantitation.

Some papers describe advantages of normal phase (Cheok *et al.*, 2014), supercritical fluid chromatography (SFC) (Huang *et al.*, 2016) and especially HILIC mode for saponins separation (Guo and Zhang, 2014) and were more selective, in comparison with the popular reverse phase separation, allowing separation and isolation of isomers in a pure form, which is a key step in the investigation of chemical structure and biological properties of individual compounds. In addition, the modern hyphenated techniques combining the better (and faster) separation with mass spectrometric detection offer additional structural information, helping to deduce the chemical structure of separated pure compounds (Zhang *et al.*, 2013; Huang *et al.*, 2016; Xing *et al.*, 2012).

Thus, the present study aims to enlarge the knowledge about the nature of *G. trichotoma* GOTCAB saponins by improvement of their separation on silica-based HILIC core-shell column (in addition to the popular RP separation) and exploring the capabilities of ESI-HRMS and tandem MS/MS data for their structural assignment.

EXPERIMENTAL

Plant material and sample preparation

Gypsophila trichotoma Wend. roots were collected in August 2004 from the Black Sea coast (Kavarna region) (43°25'60" N – 28°19'60" E) in Bulgaria and were identified by Dr. R. Gevrenova (Faculty of Pharmacy, Medical University-Sofia, Bulgaria). Voucher specimen of plant material was deposed in the Herbarium of the Faculty of Pharmacy of Nancy, Universite de Lorraine, France (HP101).

Air-dried powdered roots of the plant (1 g) were extracted with 50 ml 10% methanol (× 3) by sonication for 5 min at room temperature. The lyophilized extract constituted the crude extract 417.6 mg (GTR).

Triterpenoid saponin references

Seven saponins (compounds **A÷G**) (Figure 1, Table 1), previously isolated from *G. trichotoma*, were used as saponin references in this study. Their structure was assigned by NMR and HR-ESI-MS (Voutquenne-Nazabadioko *et al*. (2013).

Chemicals and Instruments

The LC-MS analyses were performed on a Q-Exative HESI-HRMS (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an Accela quaternary UHPLC pump and Accela autosampler. Data were processed using Xcalibur ® (Thermo Scientific Co, USA) instrument control/data handling software.

Acetonitrile (hypergrade for LC-MS and formic acid (HPLC-grade) were purchased from Merck (Darmstadt, Germany), while ammonium formate (for mass spectrometry) was provided by Sigma Aldrich (Germany).

UHPLC separation

Two separation mechanisms were tested: RP separation on Poroshell C18 150 x 3 mm, 2.7 µm column (Agilent, Santa Clara, CA, USA) and HILIC separation on Kinetex Si HILIC 100 x 3 mm, 2.6 µm column (Phenomenex, Torrance, CA,USA). The mobile phase for RP separation consisted of solution A: acetonitrile/0.08% formic acid in water 5:95 (v/v) and solution B: 0.08% formic acid in acetonitrile/water 95:5 (v/v) with gradient starting at 100% A for 0.5 min followed by a linear gradient for 18.5 min to 100% B, isocratic elution for 10 min and the mobile phase was returned to the initial conditions in 1 min and the column was equilibrated for 5 min before the next run. The flow rate was 250 µL/min

The mobile phase for HILIC separation was composed of solution A: 0.02 M ammonium formate in acetonitrile/water 95:5 (v/v) and solution B: 0.02 M ammonium formate in acetonitrile/water 60:40. The gradient program commenced at 100% A for 1 min followed by a linear gradient for 6 min to 60% A/40% B, gradient elution to 10% A/90% B in the next 8 min, 2 min cleanup at 100% B and the mobile phase was returned to the initial conditions in 2 min followed by 10 min reconditioning at 100% A. The flow rate was 400 µL/min. The solvents were filtered through 0.2 µm filterers (Millipore, Watford, Ireland).

ESI-MS and MS/MS detection

***HESI Ionization*.** Vaporizer temperature 250^{o}C, spray voltage at 3kV, ion transfer tube temperature at 300^{o}C, sheath gas pressure 35psi and auxiliary gas flow 10 (arbitrary units) were adjusted for the interphase;

***Mass spectral conditions*.** Full MS/data dependant MS^2 were performed in positive and negative ion monitoring mode with the following settings: 35000 FWHM resolution in Full

MS from 290 to 2000 *m/z*, 80msec maximal trap filling time, 17500 FWHM resolution for fragment spectra scans with 4 *m/z* quadrupole isolation window of precursor ions and 20÷40eV collision energy.
Positive and negative ion monitoring modes were tested, but in the most of experiments negative mode was used on account of its advantages, described below.
MS/MS spectra were recorded after fragmentation at different hcd (Higher Energy Collisional Dissociation) values and used for saponins structural interpretation.

RESULTS AND DISCUSSION

Optimum conditions for UHPLC-HRMS analysis

***ESI-HRMS detection*.** Positive and negative ion monitoring of GOTCAB saponins in the course of their ESI-MS detection were compared on the base of sensitivity (the areas of the chromatographic peaks), and selectivity (purity of the "molecular ion" peak).
Figure 2 indicates the ionisation abundance of the selected saponin in the negative and positive ion modes and the marked difference in the ionization responce between negative (Fig. $2a_2$, a_3) and positive (Fig. $2b_2$-b_5) ion mode of ESI-HRMS detection could be discerned. The chromatograms demonstrate the higher "selectivity" (deprotonated molecular $[M-H]^-$ ion formation predominantly) and higher sensitivity (much higher peak area AA) in negative ion mode ($2a_1$), in comparison with the detection of molecular $[M+H]^+$ ion in positive ion mode ($2b_1$). There are more intensive $[M-H]^-$ ions in negative mode as compared to positive mode. The capability of the instrument used to monitor positive and negative ions simultaneously allows to make such comparison easier. Taking into account the advantages of the negative ion mode of GOTCAB saponins detection, all next experiments were performed using negative ion monitoring SCAN mode of ESI-HRMS detection and MS/MS fragmentation of the $[M-H]^-$ ions.

Comparison of the retention and selectivity of saponins separation by RP and (Si) HILIC modes. The chromatographic behavior of GOTCAB saponin references **A**-**G** in HILIC separation mode was compared with RP mode: the order of elution in HILIC mode correlated with the number of sugar residues attached to the gypsogenin and substitution pattern of C-28 ester chain. For example, the earliest eluting saponin references **C, A, B** with $[M-H]^-$ at *m/z* 1647.637 in HILIC LC-MS are acetylated in a β-D-fucopyranose (β-D-Fucp) and sulfated in a β-D-glucopyranose (β-D-Glcp) moiety of the C-28 ester chain (Fig.1). Saponins possessing diacetylated C-28 ester chain (**E, F**) eluted earlier than monoacetylated one (**D**). The last saponin **G** contains no additional substitution in the C-28 ester chain. In HILIC LC-MS analysis, saponin references having a α-L-arabinopyranose (α-L-Arap) (**A**, **E**) eluted earlier than corresponding isomers with a β-D-xylopyranose (β-D-Xylp) (**B**, **F**). The substitution position of acetyl group in fucose moiety modified chromatographic behavior, favoring lower retention time for **C** (acetylated in β-D-Fucp C-3) in comparison with **A** (acetylated in β-D-Fucp C-4).
Aiming to obtain the best separation of GOTCAB saponins in *G. trichotoma* roots extract (GTR), RP and HILIC separation mechanisms were tested (Fig. 3). HILIC mode of UHPLC showed some advantages: more than 67 peaks of saponins were separated in less than 16 min by HILIC mode, while 40 ones were eluted in 18 min by RP UHPLC. Retention times (RT) of GOTCAB identified in GTR are described in Table 2. As shown in Figure 3, HILIC mode allowed to separate more saponins' isomers, with higher efficiency and faster. Acidity of the mobile phase in HILIC separation mode slightly influenced the selectivity and improved it with increasing of pH value. Therefore, and on account of simplification of mobile phase

preparation, the acidity was not adjusted after preparation of the mobile phase: the pH value of the ammonium formate water solution was 6.6.
More selective separation of saponins by HILIC UHPLC permits to obtain mass spectra of the pure representatives and their adequate structural assignment on the base of HRMS and MS/MS data. In the case of saponins quantitation, HILIC mode of separation and ESI with SIM detection mode of $[M-H]^-$ molecular ions of saponins could be much effective, but the aim of this work is the profiling and identification of GOTCAB saponins, not their quantitation.

Investigation of the GOTCAB references fragmentation patterns by UHPLC-ESI/HRMS

UHPLC-ESI/MS analyses of the available saponin references **A**÷**G** (Fig. 1) served to confirm their presence in GTR and to observe GOTCAB fragmentation patterns, allowing putting forward our strategy for identification of saponins in GTR. ESI/MS data of previously reported saponins (not available as reference substances) **H**-**K** were used as well (Fig. 1). The measured values for the monoisotopic masses of $[M-H]^-$ ions and their calculated values are described in Table 1. Fragment MS/MS mass spectra of the $[M-H]^-$ molecular ion of the saponin references were recorded at different values for hcd (higher-energy collisional dissociation). Figure 4 illustrates the capability to obtain an optimal MS/MS spectrum for saponins structural assignment. According to the illustrated case of saponin reference **A**, hcd values of 25÷30 could be taken as appropriate fragmentation energy giving information for the molecular $[M-H]^-$ ion and enough intensity of the fragment ions.
Figure 5 shows the fragment MS/MS spectra of the $[M-H]^-$ ions of saponin references **A**÷**G** recorded at hcd value 30. Abundant $[M-H]^-$ were first used for saponins recognition. The MS/MS spectra of the corresponding $[M-H]^-$ were explored for structural assignment of the detected new and/or isobaric compounds.
In GOTCAB saponins, carbohydrate chains were attached to the aglycone through the hydroxyl group at C-3 and carboxyl group at C-28 positions (Fig. 1). According to the structure properties of the saponins, both glycosidic bond at C-3 and ester bond at C-28 showed facile cleavage in (-) ESI-MS/MS to form deglycosylated ions. Thus, saponin references **A**-**K** were deglycosylated to form two fragment ions at *m/z* 939.458 [M-H-ester chain]$^-$ and 469.331 [aglycone-H]$^-$ (Fig. 5). The signal at *m/z* 469.331 pointed to the loss of two carbohydrate chains and, in the same time, corresponded to the deprotonated ion of the aglycone (gypsogenin) (Voutquenne-Nazabadioko et al., 2013). The fragmentation of [gypsogenin-H]$^-$ ion yielded the characteristic product ions at *m/z* 451.321 [gypsogenin-H-H_2O]$^-$, 439.321 [gypsogenin-H-CH_2O]$^-$and 423.326 [gypsogenin-H-HCO_2H]$^-$. Furthermore, *m/z* 405.316 was obtained by concomitant loss of H_2O (18 uma) and HCO_2H (46 uma) (Table 1). Formic acid (46 uma) and CO_2 (44 uma) have been reported as two possible elimination pathways of a carboxylic function of triterpenic acids in ESI-MS (Sandjo *et al.*, 2017).
At low hcd value (20) saponin references **A**÷**C** typically produced ~~several~~ abundant product ions ~~with high intensity~~ at *m/z* 707.170 [ester chain-H]$^-$ and 1177.510 [M-H-trisaccharidic chain]$^-$ (Fig.5 a,b,c, Table 1). However, the abundance of the ion at *m/z* 707.170 was higher because it is generated by a C-28 ester bond cleavage. At high HCD (35), a loss of the ester chain was observed leading to the ion at *m/z* 707.170 (acetylated and sulfated tetrasaccharide) with a relevant relative abundance, representing the base peak in **A**÷**C.**
The type of sugar moieties were determined from MS/MS spectra in which a mass difference of 162.053, 146.058 and 176.124 indicated the presence of hexose (Hex), deoxyhexose (dHex) and hexosuronic acid (HexA), respectively; and mass difference of 132.042 indicated the presence of a pentose (Pen) (arabinose/xylose). Regarding saponin references **A**÷**C**, three

mass intervals of 79.958 (sulfate group), 162.053 (glucose) and 242.011 (glucose + sulfate group) were observed at low hcd value (Table 1). In MS/MS spectra of saponin references **E** and **F** the mass differences of 42.011 (acetyl group) (*m/z* 1567.680) and 204.064 (glucose + acetyl group) (*m/z* 1405.627) indicated that the associated compounds possess acetylated glucose moiety (Table 1, Fig. 5 e, f). The position of sugars in the saccharide chain could also be ~~roughly~~ determined owing to the fact that the cleavage of the glycosidic bond usually occurred first at outermost glycosidic bond of the saccharide chain linked at C-28.

The fragmentation pathway of $Y_{0\alpha}$ [M-H-ester chain]$^-$ involved consecutive losses of hexosyl at *m/z* 777.398 [M-H-ester chain-Hex]$^-$ ($Y_{0\alpha}/Y_{1\beta}$) and pentosyl residues at *m/z* 807.421 [M-H-ester chain-Pent]$^-$ ($Y_{0\alpha}/Y_{1\beta'}$) from the trisaccharide at C-3 of gypsogenin indicating a ramification in the sugar chain. This is confirmed by the losses of water together with CO_2, resulting in abundant fragment ions at *m/z* 759.395 [M-H-ester chain-Hex-H_20]$^-$, 583.363 [M-H-ester chain-Hex-Pent-H_20-CO_2]$^-$, and 565.352 [M-H-ester chain-Hex-Pent-2H_20-CO_2]$^-$, together with ions at *m/z* 745.417 [M-H-ester chain-Pent-H_20-CO_2]$^-$, and 627.354 [M-H-ester chain-Hex-Pent-H_20]$^-$ (Table 1, Fig. X). Moreover, typical ions of the $Y_{0\alpha}$ fragmentation at *m/z* 537.359 and 519.346 resulted from both formic acid and CO_2 elimination. On the other hand, the relative abundance of the ion at *m/z* 759 ($Y_{0\alpha}/Z_{1\beta}$) could be related with the position of glycosylation (Table 1). Thus, 15% relative abundance of this ion indicated an interglycosidic linkage (1→2) of the galactose, while the occurrence of the low abundant ion at *m/z* 807 supported by 745 and 727 contributed to the assignment of the arabinose at C-3 of glucuronic acid. The ions at *m/z* 551.338 and 511.344 (references **D-K**) corresponded to the loss of pentose and hexose units together with intramolecular breakage of the glucuronic acid $^{4,5}A_{0\beta}$ and $^{0,2}A_{0\beta}$, respectively (Fig. X) The latter ion corroborated substitution at C-2 and C-3 of glucuronic acid moiety. Both trisaccharides linked at C-3 of the gypsogenin, with terminal β-D-Xylp or α-L-Arap, are identical in mass and fragmentation pattern and it was not possible to confirm the corresponding isomer. The comparison with authentic references for these saponins in GTR allowed verifying the exact configuration.

The length and type of C-28 carbohydrate chain were discovered from (-) ESI-MS fragmentation pattern as well. Thus, the loss of 708.178 amu could be attributed to the monoacetylated and sulfated C-28 tetrasaccharide (saponin references **A**÷**C**). The ester chain at *m/z* 707.170 tended to release consequently hexose, one or two deoxyhexoses and an acetyl group, yielding the most abundant fragment ions at *m/z* 519.101 [ester chain-H-dHex-Ac]$^-$ (29%) and 241.001 [ester chain-H-2dHex-Pent-Ac]$^-$ (17%). The latter ion corresponded to the sulfated glucose confirming the position of the sulfate in the C-28 ester chain. Concerning these saponins, it must be taken into account the fact that a sulfated C-28 chain provided series of peaks, the majority of them with low abundance (Table 1, Fig. Y). Some ions characterized the fucose moiety: 607, 591 and 579 resulting from the cross-ring cleavages $^{3,5}A_{0\alpha}$ (-100), $^{0,3}A_{0\alpha}$ (-116) and $^{1,3}A_{0\alpha}$ (-128). An acetyl group esterifies the hydroxyl group at C-4 of the fucose unit as observed by the fragment ions at *m/z* 503.107 [707-Pent-$^{3,4}A_{0\alpha}$]$^-$ and 527.161 [707-Pent-$^{3,5}A_{0\alpha}$]$^-$. The xylose moiety was linked at C-4 of the rhamnose unit witnessed by the ion at *m/z* 329.022 [707-Pent-(dHex+Ac)-$^{3,5}A_{1\alpha}$]$^-$. Glucose attachment at C-3 was confirmed by the ion at *m/z* 301.022 [707-Pent-(dHex+Ac)-$^{1,3}X_{1\alpha}$]$^-$supported by 371.022 [707-Pent-(dHex+Ac)-$^{3,5}A_{1\alpha}$]$^-$. The loss of 628.222 and 670.232 amu corresponded to monoacetylated C-28 tetrasaccharide (reference **D**) and diacetylated C-28 tetrasaccharide (references **E** and **F)**, respectively. The loss of 746.263 amu was in agreement with C-28 tetrasaccharide substituted with methoxycinnamoyl moiety (references **J** and **K**). This is confirmed by the loss of 178 amu at *m/z* 1507.659 [M-H-MeCin]$^-$ due to the loss of methoxycinnamoyl moiety (Table 1).

Profiling of *G. trichotoma* root extract by UHPLC-ESI/HRMS

Except for the fully characterized saponin references from *G. trichotoma*, the tentative identification of the GTR saponins was based on HRMS-measurements, comparison with mass fragmentation observed for the GOTCAB references and literature data (Chen *et al.*, 2011, Voutquenne-Nazabadioko *et al.*, 2013; Zhang *et al.*, 2013). Since GOTCAB have different ester-bonded oligosaccharides and showed different fragment ions, they were classified into 3 types: I - GOTCAB saponins with C-28 oligosaccharide substituted with acyl group; II - GOTCAB saponins with C-28 oligosaccharide substituted with sulfate group; III - GOTCAB saponins with C-28 oligosaccharide substituted with both acyl and sulfate groups.

Type I - GOTCAB saponins with C-28 oligosaccharide substituted with acyl group (acylated oligosaccharide). Compounds **1a-d,** separated in HILIC UHPLC-MS, shared the same molecular weights as saponin **D** (Table 2). **1a** displayed fragment ions at *m/z* 1097.730 $[(M-H)-Hex-Pent-HexA]^-$ and 939.461 $[(M-H)-2dHex-Pent-Hex-Ac]^-$ corresponding to the loss of a trisaccharide and ester chain, respectively. Fragment ions at *m/z* 759.397, 627.354, 583.364 and 565.353 were also consistent with the C-3 trisaccharide of saponin references **D** (Table 1). At low value of hcd **1a** produced fragment ions at *m/z* 1525.679 $[(M-H)-Ac]^-$, 1405.629 $[(M-H)-Hex]^-$ and 1141.776 $[(M-H)-Hex-2Pen]^-$. At high value of hcd, this compound displayed fragment ions at *m/z* 1363.079 $[M-H-Hex-Ac]^-$, 807.419 [M-H-ester chain-Pen]$^-$ and 469.332 [gypsogenin-H]$^-$. Peaks at *m/z* 451.322 $[aglycone-H-H_2O]^-$, 439.322 $[aglycone-H-H_2CO]^-$, 423.327 $[aglycone-H-HCO_2H]^-$ and 405.316 $[aglycone-H-HCO_2H-H_2O]^-$ are ascribed to the gypsogenin. The loss of 628.212 amu corresponded to a branched tetrasaccharide ending with both hexose and pentose. The isotopic peak profile of *m/z* 1567.673 corroborated the predicted molecular formula of $C_{72}H_{111}O_{37}$, as the relative intensities of isotopic peaks matched closely with the simulated theoretical ones. The retention time of compound **1a** is in good agreement with reference compound **D** as well. Thus compound **1a** is identified as saponins **D** or 3-*O*-(α-L-arabinopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin. Compound **1c** also displayed the same mass fingerprint as compound **D** suggesting **1c** may possess β-D-Xylp at C-3 trisaccharide, thus favoring longer retention time. Based on the identical molecular weights and MS/MS fragmentation patterns, compounds **1c** was tentatively assigned to the isomer of saponin **D** or 3-*O*-(β-D-xylopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin, previously isolated from GTR (Voutquenne-Nazabadioko *et al.*, 2013). Compounds **1b** and **1d** could be differentiated by the location of the acetyl group as observed in saponin reference **C**. They are tentatively identified in our analysis as the acetate isomers of **1a** and **1c** with fucose position 3 acetylated.

At low value of hcd the parent ion at *m/z* 1609.695 for compounds **2a-2b** yielded fragment ions at *m/z* 1567.681 $[M-H-Ac]^-$, 1477.649 [M-H-Pen]$^-$ and 1405.637 $[M-H-Hex-Ac]^-$, indicating an ester chain substituted by both terminal acetylated hexose and terminal pentose. By comparing the retention times and MS/MS spectra with those of saponin references, compounds **2a** and **2b** were ascribed to saponins **E** and its isomer **F**, respectively. Thus compound **2a** is 3-*O*-(α-L-arabinopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*6-O*-acetyl-β-D-glucopyranosyl-(1→3)-[α-L-arabinopyranosyl-(1→3)-β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin, and compound **2b** is 3-*O*-(β-D-xylopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*6-O*-acetyl-β-D-glucopyranosyl-

(1→3)-[α-L-arabinopyranosyl-(1→3)-β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin

As previously described, two GOTCAB isomers **J** and **K** with molecular weight 1686 have a methoxycinnamoyl (MeCin) moiety attached to β-D-Fucp C-4 of the C-28 ester chain (Voutquenne-Nazabadioko *et al.*, 2013). The fragmentation of **3a** gave ions at *m/z* 939.464 [(M-H)-2dHex-Pent-Hex-MeCin]$^-$ and 469.332 [gypsogenin-H]$^-$ (Table 2). Saponins **3a-3d** were identical in mass and fragmentation pattern and it was not possible to confirm which isomer corresponds to previously isolated saponin references **J/K** (not available in this study) by their retention time. However, saponins **3a-3d** probably corresponds to the saponin references **J/K** (Voutquenne-Nazabadioko *et al.*, 2013), and their isomers with xylopyranose moiety identified as compouds 4/5 in the roots of *Gypsophila perfoliata* L. (Chen *et al.*, 2011).

In the (-) ESI-MS spectrum of **4a**, the [M-H]$^-$ signal was observed at *m/z* 1727.734 indicating molecular weight of 1728 amu, 42.011 mass units higher than that of **3a-3d**. This suggested the presence of one additional acetyl group confirmed by the fragment ion at *m/z* 1667.722 [M-H-AcOH]$^-$. Furthermore, at high hcd (35), the fragment ions at *m/z* 939.468 [(M-H)-2dHex-Pent-Hex-MeCin-Ac]$^-$ and 787.272 [ester chain-H]$^-$ (**4a**) showed the presence of MeCin group which was corroborated by the peak at *m/z* 177.054 [MeCin-H]$^-$ (Table 2). The fragmentation pathway involved a loss of CH_3OH from the cleavage of the *p*-methoxy group, followed by structural rearrangement, resulting in less intensive fragment ion at *m/z* 145.028 [MeCin-H-CH_3OH]$^-$. Thus, compounds **4a-4d** were probably the acetylated derivatives of compounds **3a-3d.**

At low value of hcd **5a** ([M-H]$^-$ at *m/z* 1741.732) produced fragment ion at *m/z* 1699.725 [M-H-Ac]$^-$ and 1187.580 [M-H-Pent-Hex-HexA-2Ac]$^-$ (Table 2). MS/MS data corroborated ester chain consisting of diacetylated C-28 tetrasaccharide. Thus, we tentatively assigned compounds **5a-5d** to four isomers including previously isolated saponins **H** and **I** (Voutquenne-Nazabadioko *et al.*, 2013). It was not possible to confirm which isomer corresponds to references **H/I** (not available in this study) by their retention time. However, similarly to the results obtained for saponins references **A-C**, compounds **5c** and **5d** were attributed to compounds **H** and **I**, respectively.

At low value of hcd, the parent ion at *m/z* 1725.738 [M-H]$^-$ (**6a**) gave three significant fragments at *m/z* 1683.748 [M-H-Ac]$^-$, 1593.704 [M-H-Pen]$^-$ and 1171.584 [M-H- Pent-Hex-HexA-2Ac]$^-$. This fragmentation pattern matched with those of compound 3 reported in *Gypsophila perfoliata* or 3-*O*-β-D-galactopyranosyl-(1→2)-[β-D-xylopyranosyl-(1→3)]-β-D-glucuronopyranosyl gypsogenin α-L-arabinopyranosyl-(1→3)-β-D-xylopyranosyl-(1→4)-α-L-rhamnopyranosyl-(1→2)-[3,4-di-*O*-acetyl-β-D-quinovopyranosyl-(1→4)]-β-D-fucopyranoside (Chen *et al.*, 2011). The MS/MS data of **6a** were similar to those of **5a**, except for the appearance of an additional deoxyhexose unit, instead of hexose moiety in **H/I** (Table 2). This compound is totally different in their structure of the C-28 ester chain possessing a quinovose moiety.

The MS/MS spectra of **7a** ([M-H]$^-$ at *m/z* 1729.733) gave the fragment ion at *m/z* 939.461 [(M-H)-2dHex-2Hex-Pen-Ac]$^-$, suggesting the presence of one additional hexose unit in the C-28 ester chain in comparison with saponin reference **D**. These data suggested that this saponin is new and was note previously described in the literature, independently the location of the hexose unit.

The [M-H]$^-$ signal at *m/z* 1435.641 in compounds **8a-8d** associated to the fragment ions at *m/z* 939.450 [M-H-2dHex-Hex-Ac]$^-$ and 469.332 [gypsogenin-H]$^-$ (**8a)**, indicated the absence of a pentose moiety in the ester chain in comparison with saponin **D** (Table 2). The loss of a pentose moiety in the ester chain whatever the type of pentose unit in the C-3 chain and the

location of the acetyl group indicated that this saponin is new and was note previously described in the literature.

Type II - GOTCAB saponins with C-28 oligosaccharide substituted with sulfate group. This group was characterized by the occurrence of a base peak corresponding to the C-28 ester chain (Table 2). Moreover, typical ions of deoxyhexose moiety internal cleavage were observed: $^{3,5}X_{0\alpha}$ ([ester chain–H-58]$^-$), $^{0,3}X_{0\alpha}$ ([ester chain–H-74]$^-$) $^{1,3}X_{0\alpha}$ ([ester chain–H-86]$^-$), $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ ([ester chain–H-Pent-30]$^-$) (**9**-**15**) (Fig.Y). In addition, the fragmentation pattern afforded an abundant ion at *m/z* 519 [ester chain-H-dHex] (fragment $C_{1\alpha}$) or [ester chain-H-dHex-Pent/dHex/Hex] indicating terminal position of the additional sugar moieties (**9**-**12**). The fragment at *m/z* 329 allowed situating the pentose at C-4 of the second deoxyhexose, while a C-3 linked hexose was evidenced by the ion at *m/z* 371. This behavior was similar to the fragmentation described for the saponin references **A-C** (Table1). The sulfate group was linked at hexose unit witnessed by the ion at *m/z* 241 ($^3X_{2\alpha}$) (Fig. Y). Thus, the ester chain in **9**-**12** could be assigned as *O*-sulfate-hexosyl-(1→3)-[pentosyl-(1→4)]-deoxyhexosyl-(1→2)-deoxyhexoside.

The base peak at *m/z* 665.160 in MS/MS spectra of compounds **9a** and **9b (**[M-H]$^-$ at *m/z* 1605.626) corresponded to the aforementioned ester chain, thus indicating the loss of the acetate in comparison with the saponin references **A** and **B** (Table 2, Fig. 3). At low value of hcd, the fragment at *m/z* 1525.680 derived from the sulfate group elimination. ~~Our tentative assignment of these saponins is GOTCAB, having a tetrasacharide chain at C-28 consisting of two desoxyhexoses, pentose, hexose and a sulfate group.~~ These GOTCAB were the most abundant saponins in GTR. To the authors' knowledge there are no literature data for a saponin with matching mass spectral data.

Concerning **10a** and **10b**, there was mass difference of 14 between their C-3 trisaccharidic chains and that in saponin **D**, C-3 oligosaccharide was identified as Hex-Hex-Pent. These two compounds have a trisaccharide chain possessing hexose at the place of glucuronic acid. Peaks **11a** and **12a** showed prominent fragment ions at *m/z* 665.161 [ester chain-H] $^-$, suggesting the same ester chain as in saponin **9a** (Table 2). The mass difference of 132.042 indicated the loss of a pentose unit and that **11a** and **11b** possess a C-3 disaccharide chain consisting of hexose and hexosuronic acid. MS data of compounds **12a**-**12d**, showed the [M-H]$^-$ at *m/z* 1635.638 corresponding to 162 uma higher than that of **11a**, and indicating the presence of an additional hexose unit in C-3 carbohydrate chain. In comparison with compounds **9a**-**9b**, the pentose was replaced by a hexose in the C-3 trisaccharide chain. Saponins **10**-**12** were not described in the literature.

The fragmentation observed for **13a** ([M-H]$^-$ at *m/z* 1751.684) exhibited the loss of a sulfate group at *m/z* 1671.691 and a trisaccharidic chain at *m/z* 1281.530. At high value of hcd, the compound generated base fragment ion at *m/z* 811.198 [ester chain-H]$^-$, suggesting an additional deoxyhexose moiety in C-28 ester chain in comparison with **9a** (Table 2). In the same way, peaks **14a** ([M-H]$^-$at *m/z* 1737.668) produced a major ion at *m/z* 797.202 [ester chain-H]$^-$. Fragment ions were in good agreement with those of **9a**, except for the appearance of an additional pentose unit in the ester chain (Table 2). Parent ion at *m/z* 1767.686 [M-H]$^-$ (**15a)** afforded significant fragments at *m/z* 827.215 [ester chain-H]$^-$, 665.162 [ester chain-H-Hex]$^-$ and 503.108 [ester chain-H-2Hex]$^-$ indicating the presence of supplementary hexose unit in the ester chain in comparison with **9a**. Compounds **13**-**15** were newly identified and the location and the type of supplementary sugar unit must be determined.

Type III - GOTCAB saponins with C-28 oligosaccharide substituted with both acyl and sulfate grou. Four compounds **16a-16d** with [M-H]$^-$ at *m/z* 1647.637 were observed with

different retention times in HILIC UHPLC-MS (Table 2). At high value of hcd **16a** afforded a fragment ions at *m/z* 1177.510 [M-H-Hex-Pen-HexA)]$^-$ and 939.464 [M-H-2dHex-Pent-Hex-Ac-SO_3]$^-$ corresponding to the loss of C-3 trisaccharide chain and C-28 ester chain, respectively. In addition, **16a** displayed fragment ions at *m/z* 707.171 [ester chain-H]$^-$ (base peak) and 469.332 [gypsogenin-H]$^-$. Thus, fragmentation pattern is in agreement with the structure of compound **C** previously reported from *G. trichotoma* roots (Voutquenne-Nazabadioko *et al.*, 2013)**.** Based on the identical molecular weights, retention times and proposed fragmentation patterns observed in the UHPLC-MS analysis, compounds **16d, 16a** and **16c** were assigned to saponin references **A, B** and **C**, respectively, named 3-*O*-(α-L-arabinopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*3*-*O*-sulfate-β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin, 3-*O*-(β-D-xylopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*3*-*O*-sulfate-β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-4-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin, and 3-*O*-(α-L-arabinopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*3*-*O*-sulfate-β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-3-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin, respectively. In addition, **16b** and **16c** eluted earlier than corresponding pair of isomers **16d** and **16a** suggesting **16b** may possess acetyl group at β-D-Fucp C-3. Two pairs of isomers differ in the acetyl group position. Compound **16b** is thus a new saponin, tentatively named 3-*O*-(β-D-xylopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(*3*-*O*-sulfate-β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-3-*O*-acetyl-β-D-fucopyranosyl)-gypsogenin.

At high value of hcd **17a** (parent ion at *m/z* 1765.673 [M-H]$^-$) produced a base peak at *m/z* 825.215 [ester chain-H]$^-$ and fragment ions at *m/z* 1295.554 [M-H-470]$^-$, 647.152 [ester chain-H-MeCin-H_2O]$^-$, and 519.104 [ester chain-H-dHex-MeCin]$^-$. The formation of fragment ions at *m/z* 177.054 and 145.028 was favored for the presence of a methoxycinnamoyl moiety attached to C-4 of the terminal deoxyhexose moiety (instead of an acetyl group in saponin references **A-C**) witnessed by the fragments at *m/z* 607 and 527 (Table 2). Thus, compounds **17a-17d** corresponded probably to the sulfated derivatives of compounds **J/K** with sulfate on the glucose moiety, as deduced from the pic at *m/z* 241.002 [ester chain-H-2dHex-Pent-MeCin]$^-$, and were new compounds. Compounds **18a** displayed prominent ion at *m/z* 749.183 [ester chain-H]$^-$ indicating the presence of additional acetyl group in the ester chain in comparison with saponin references **A-C**. Thus, compounds **18a-18d** corresponded probably to the sulfated derivatives of compound references **E** and **F** with sulfate on the glucose moiety, as deduced from the pic at *m/z* 241 [$GlcSO_3$-H]$^-$, and were new compounds.

The presence of supplementary pentose in C-28 ester chain was evidenced in saponins **19a-19c,** which showed related fragmentation pattern as saponins **A-C** after the loss of this pentose. Similarly, compound **20a** with parent ion at *m/z* 1809.693 [M-H]$^-$, 162.059 mass units higher than that of saponin references **A-C,** showed prominent ions at *m/z* 1177.503 [M-H-trisaccharidic chain-Hex]$^-$, with fragments ions at *m/z* 869.228 [ester chain-H]$^-$ and 707.171 [ester chain-H-Hex]$^-$ consistent with one additional hexose unit on the ester chain (Table 2). These compounds (**19a-19d** and **20a-20d**) were not previously described.

MS spectra of the pair of isomers **21a** and **21b** at *m/z* 1525.672 indicated the absence of both acyl and sulfate groups. The fragmentation pattern consists with that of saponin reference **G** (Voutquenne-Nazabadioko *et al.*, 2013; Zhang *et al.*, 2013). Compounds **21a** was assigned to saponin reference **G**, also named 3-*O*-(α-L-arabinopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-

(1→4)]-α-L-rhamnopyranosyl-(1→2)-β-D-fucopyranosyl)-gypsogenin (Voutquenne-Nazabadioko *et al.*, 2013). This MS data could also correspond to their isomer 3-*O*-(β-D-xylopyranosyl-(1→3)-β-D-galactopyranosyl-(1→2)]-β-D-glucuronopyranosyl)-28-*O*-(β-D-glucopyranosyl-(1→3)-[β-D-xylopyranosyl-(1→4)]-α-L-rhamnopyranosyl-(1→2)-β-D-fucopyranosyl)-gypsogenin, isolated previously from the roots of *G. paniculata* (Frechet *et al.*, 1991) and *G. oldhamiana* (Zhang *et al.*, 2013).

By comparing the retention time and MS/MS of the GTR saponins with those of reference saponins, peaks **1a**, **2a**, **2b**, **16c**, **16d**, **16a**, **21a** were ascribed unambiguously to saponin references **D**, **E**, **F**, **C**, **A**, **B** and **G**, respectively. Other peaks were tentatively identified by detailed studies of their MS and MS/MS data, and by comparison with literature data.

Fragmentation patterns of *G. trichotoma* GOTCAB saponins indicated that they shared the same aglycone, gypsogenin. Saponins appeared to be 3,28-*O*-bidesmosides containing two carbohydrate chains: one attached via a glycosidic binding at C-3 and one ester-bonded at C-28 of aglycone. At hcd=25-30 a majority of GOTCAB typically produced intense ions indicating the loss of both carbohydrate chains without diminishing the ability to distinguish the $[M-H]^-$ ions. At low value of hcd, saponins produced fragment ions attributed to the loss of acyl and (or) sulfate groups, as well as ions indicating the sequential loss of sugar residues. (Figure 5, X,Y, Table 1). Our strategy for recognition of GOTCAB was based on the diagnostic ions for each type saponins in GTR. Thus, the fragment ions attributed to the loss of one acyl group, terminal acylated hexose, and trisaccharidic chain together with one (two) acyl groups, could be considered diagnostic ions for acylated GOTCAB saponins. The relative abundance of ions at 807.416, 759.395, 745.416 and 727.405 were related to the position of the glycosylation on the glucuronic acid of the C-3 trisaccharidic chain, more abundant in C-2 than in C-3. The fragmentation of the aglycone gypsogenin yielded the characteristic product ion at *m/z* 423.3258 deriving from the loss of HCO_2H. The fragment ions at *m/z* 665.159, 519.101 and 241.001 together with $[\text{ester chain-H}]^-$ (base peak), and $[M-HSO_3]^-$ (at low value of hcd) can be considered as the diagnostic ions for sulfated GOTCAB saponins. In this group, the relevant ions $^{3,5}A_{0\alpha}$ (-100), $^{0,3}A_{0\alpha}$ (-116) and $^{1,3}A_{0\alpha}$ (-128) from the cross-ring cleavages were indicative of deoxyhexose moieties (Fig.Y). The presence of low abundant ions $[\text{ester chain-Pent-(dHex+Ac)-}^{3,5}A_{1\alpha}]^-$ and $[\text{ester chain-Pent-(dHex+Ac)-}^{1,3}X_{1\alpha}]^-$ pointed to the hexose and pentose glycosylation position in C-28 ester chain, respectively. Particularly, internal cleavages $^{3,4}A0\alpha$ and $^{3,5}A_{0\alpha}$ indicated acylation at C-4 terminal deoxyhexose hydroxyl. Acylated and sulfated GOTCAB showed the same fragmentation pattern as the saponin references with sulfated C-28 oligosaccharide together with the ion at *m/z* 241.001 $[GlcSO_3\text{-}H]^-$. Both C-3 trisaccharides with terminal xylose/arabinose cause the appearance of characteristic pairs of saponins differing in the terminal sugar of C-3 trisaccharidic chain. Additional pairs of isomers resulted from the substitution with acetyl, sulfate and methoxycinnamoyl groups at C-28 ester chain. It seems that most of *G. trichotoma* saponins occur as isomeric pairs in the profiling with members of pairs differing in retention times based on their matching mass fingerprints. Their structural differences were attributed to both the pentose unit in C-3 trisaccharide and the location of acyl groups. The existence of isomers possessing acetyl group on the fucose moiety could be attributed to an isomerization by migration of the acetate from the position 3 to the position 4 and reversely.

The main finding of this study was that we put forward a strategy for the tentative identification of *Gypsophila* GOTCAB saponins by HRMS and MS/MS data. This strategy was based on the fact that GOTCAB have two oligosaccharidic chains and the aglycone part is triterpenoid carboxylic acid. In this way we tentatively identified a large number of triterpenoid saponins for the first time in HILIC UHPLC-MS, not previously reported in *G.*

trichotoma. HILIC separation indicated a much broader range of GOTCAB than these in RP UHPLC-MS analysis due to the superior selectivity and peak capacity of this stationary phase. The proposed HILIC UHPLC method enabled the separation and identification of GOTCAB saponins with advantages in terms of selectivity in comparison with the relatively time consuming isolation and structure determination procedures typically employed for saponins. The analysis of HR-MS/MS data after HILIC UHPLC separation gave 21 saponins forming between 2 and 4 isobaric and positional isomers classified into three types of compounds depending of the type of substituent on the C-28 oligosaccharide, acyl and/or sulfate group. The type I is consistent with acylated GOTCAB saponins (**1**-**8**) from which **1a, 2a, 2b, 3a, 3c, 5a, 5c** were identified as saponins references **D, E, F, J, K, H** and **I**, respectively. The type II sulfated GOTCAB saponins (**9**-**15**) were identified for the first time in this species and to our knowledge are not described in the literature. The type III - acylated and sulfated GOTCAB saponins (**16**-**20**) correspond to the saponin references **A**-**C** (**16d, 16a, 16c**) and their derivatives. Finally, compound **21a** correspond to the non acylated and sulfated saponin reference **G,** and **21b** to its isomer isolated in *G. paniculata* (Frechet *et al.*, 1991) and *G. oldhamiana* (Zhang *et al.*, 2013).

Recently, new bidesmosides with an acetyl or a methoxycinnamoyl group at C-28 ester chain, as well as 3-*O*-β-sulfated monodesmosides were reported from the roots of some *Gypsophila* species (Krasteva *et al*., 2009; Yao *et al*., 2010; Luo *et al*., 1011)**.** *G. trichotoma* is the only species with acetylated and sulfated GOTCAB saponins (Voutquenne-Nazabadioko *et al.*, 2015). Four isobars (**6a**-**6d**) matched with compound 3 isolated from *G. perfoliata* (Chen *et al.*, 2011) and its isomers. It is consistent with our results from the previous study, where the saponin composition of Bulgarian *G. trichotoma* was very similar to those observed in *G. perfoliata*, originated from China (Voutquenne-Nazabadioko *et al.*, 2015). In the UHPLC-MS analysis of GTR, two saponins at *m/z* 1725.738 (**6a-6d**) and 1525.669 $[M-H]^-$ (**21a-21d**) were related to those reported in *G. oldhamiana* (Zhang *et al.*, 2013), *G. paniculata* (Frechet *et al.*,1991) and *G. perfoliata* (Chen *et al.*, 2011).

The selective profiling of GOTCAB presented here provides a base for more comprehensive survey of saponins distributed in *Gypsophila* species. These types of chemical inventories are necessary to understand the distribution and functional significance of saponins in plants and ultimately for future efforts to promote the health benefits of these medicinally important compounds. The combination of higher selective HILIC-UHPLC with HRMS and MS/MS detection enlarges the capabilities for recognition of new saponins.

REFERENCES

Arslan I., Celik A., Chol J.H. 2012. A cytotoxic triterpenoid saponin from under-ground parts of *Gypsophila pilulifera* Boiss.& Heldr. *Fitoterapia*, **83:** 699–703.

Arslan I., Celik A., Melzig M.F. 2013. Nebulosides A–B, novel triterpene saponins from under-ground parts of *Gypsophila arrostii* Guss. var. *nebulosa*. *Bioorg. Med. Chem.* **21**: 1279–1283.

Bachran C, Bachran S, Sutherland M, Bachran D, Fuchs H. 2008. Saponins in tumor therapy. *Mini Rev Med Chem* **8**:575–584.

Bai H., Zhong Y., Xie Y.Y., Wang Y.-S., Liu L., Zhou L. **2007.** A major triterpenoid saponin from *Gypsophila oldhamiana*. *Chem. Biodivers.* **4**: 955-60.

Bottger S, Melzig MF. 2010. Triterpenoid saponins of the Caryophyllaceae and Illecebraceae family. *Phytochem Lett* **4**: 59-68.

Bottger S, Westhof E, Siems K, Melzig MF. 2013. Structure-activity relationships of saponins enhancing the cytotoxicity of ribosome-inactivating proteins type I (RIP-I). *Toxicon* **73**: 144-150.

Chen Q, Luo JG, Kong LY. 2011. New triterpenoid saponins from the roots of *Gypsophila perfoliata* Linn. *Carbohydr Res* **346**: 2206-2212.

Cheok CY, Salman HAK, Sulaiman R. 2014. Extraction and quantification of saponins: A review. *Food Res Int* **59**: 16-40.

Francis G, Kerem Z, Makkar HP, Becker K. 2002. The biological action of saponins in animal systems: a review. *Br J Nutr* **88**: 587–605.

Frechet, D., Christ, B., Monegier du Sorbier, B., Fisher, H., Vuilhorgne, M., 1991. Four triterpenoid saponins from dried roots of *Gypsophila* species. *Phytochemistry* **30**: 927-931.

Fuchs H, Bachran D, Panjideh H, Schellmann N, Weng A, Melzig MF; Sutherland M, Bachran C. 2009. Saponins as tool for improved targeted tumor therapies. *Curr Drug Targets* **10**: 140-151.

Gevrenova R, Weng A, Voutquenne-Nazabadioko L, Thakur M, Doytchinova I. 2015. QSAR study on saponins as cytotoxicity enhancers. *Lett Drug Des Discov* **12**:166-171.

Gilabert-Oriol R, Mergel K, Thakur M, Mallinckrodt B, Melzig MF, Fuchs H, Weng A. 2013. Real-time analysis of membrane permeabilizing effects of oleanane saponins. *Bioorg Med Chem.* **21**: 2387-2395.

Gilbert-Oriol R, Thakur M, Haussmann K, Niesler N, Bhargava C, Gorick C, Fuchs H, Weng A. 2016. Saponins from *Saponaria officinalis* L. augment the efficacy of rituximab-immunotoxin. Planta Med **82**: 1525-1531.

Guo X, Zhang X. 2014. Hydrophilic interaction chromatography for selective separation of isomeric saponins. *J Chromatogr A*, **1325**: 121-128.

Güçlü-Üstündag O, Mazza G. 2007. Saponins: properties, applications and processing. *Crit Rev Food Sci Nutr* **47**: 231–258.

Henry M. 2005. Saponins and phylogeny: Example of the "Gypsogenin group" saponins. *Phytochem Rev* **4**: 89-94.

Huang Y, Zhang T, Zhou H, FengY, Chunlin Fan Ch, Chen W, Jacques Crommen J, Jiang Z. 2016. Fast separation of triterpenoid saponins using SFC with single quadropole mass spectrometry. *J Pharm Biomed Analysis*, **121**: 22-29.

Krasteva I, Jenett-Siems K, Kaloga M, Nikolov S. 2009. 3-*O*-Sulfo-triterpenoid saponins from *Gypsophila trichotoma* Wend. *Z Naturforsch* B **64**: 319–322.

Luo J-G, Nie W, Kong L-Y. 2011. Three new sulfated triterpenoid from the roots of Gypsophila pacifica. *J Asia Nat Prod Res* **13**:529-533.

Piorkowski DT, McClements DJ. 2014. Beverage emulsions: Recent developments in formulation, production, and applications. *Food Hydrocoll* **42**: 5–41.

Podolak I, Galanty A, Sobolewska D. 2010. Saponins as cytotoxic agents: a review. *Phytochem Rev* **9**:425–474.

Sandjo LP, dos Santos Nascimento MVP, da Silva LAL, Munhoz ACM, Pollo LAE, Biavatti MW, Ngadjui BT, Opatz T, Fröde TS. 2017. ESI-MS^2 and Anti-inflammatory

Studies of Cyclopropanic Triterpenes. UPLC-ESI-MS and MS^2 Search of Related Metabolites from Donella ubanguiensis. *Phytochem Anal* **28**: 27-41.

Skene CD, Sutton P. 2006. Saponin-adjuvanted particulate vaccines for clinical use. *Methods* (San Diego, Calif.) **40**: 53-59.

Vincken J-P, Heng L, de Groot A, Gruppen H. 2007. Saponins, classification and occurrence in the plant kingdom. *Phytochemistry* **68**: 275-297.

Voutquenne-Nazabadioko L, Gevrenova R, Borie N, Harakat D, Sayagh Ch, Weng A, Thakur M, Zaharieva M, Henry M. 2013. Triterpenoid saponins from the roots of *Gypsophila trichotoma* Wender. *Phytochemistry* **90**: 114-127.

Weng A. Melzig M.F. Bachran C. Fuchs, H. 2008**.** Enhancement of saporin toxicity against U937 cells by *Gypsophila* saponins. *J. Immunotoxicol.* **5**: 287–292.

Xing Q, Liang T, Shen G, Wang X, Jin Y, Liang X. 2012. Comprehensive HILIC × RPLC with mass spectrometry detection for the analysis of saponins in *Panax notoginseng*. *Analyst* **137**: 2239-2249.

Yao S, Ma L, Luo J-G, Wang J-S, Kong L-Y. 2010. New Triterpenoid Saponins from the Roots of *Gypsophila paniculata* L. *Helv Chim A*cta **93**: 361-374.

Zhang W, Luo J-G, Zhang C, Kong L-Y. 2013. Different apoptotic effects of triterpenoid saponin-rich *Gypsophila oldhamiana* root extract on human hepatoma SMMC-7721 and normal human hepatic L02 cells. *Biol Pharm Bull* **36**: 1080-1087.

Figure captions

Fig. 1. Chemical structure of GOTCAB saponin references.

Fig. 2. Extracted ion chromatograms of isobaric saponins with monoisotopic mass 1648.637 at *m/z* of their $[M-H]^-$ ion at negative ion mode (a_1) and $[M+H]^+$ ion at positive ion mode (b_1) obtained by RP separation; (-) ESI-MS spectra of the peaks with retention time 13.84 (a_2) and 13.33 (a_3) on the chromatogram a_1; (+) ESI-MS spectra of the peaks with retention time 14.62 (b_2), 13.82 (b_3), 13.35 (b_4) and 12.68 (b_5).Chromatographic conditions and injection volume were the same in both experiments.

Fig. 3. Comparison of reversed phase (RP) separation (chromatograms up) and hydrophilic interaction liquid chromatography (HILIC) separation (chromatograms down) of some *G. trichotoma* GOTCAB saponins by the extracted chromatograms at *m/z* of their molecular $[M-H]^-$ ions at negative ion mode of ESI-HRMS detection.

Fig. 4. Fragment MS/MS spectra of molecular $[M-H]^-$ ion with *m/z*=1647.637 of the saponin **A** (Figure 1, Table 1) at higher energy collisional dissociation (hcd) = 20 (a), hcd = 25 (b) and hcd = 30 (c) and hcd = 35 (d).

Fig. 5. Fragment MS/MS mass spectra of the molecular $[M-H]^-$ ions of saponin references **A**÷**G** in negative ion mode (Fig. 1, Table 1).

Fig. X. The fragmentation pattern of saponin reference **D**. Experimental values for fragment ions identified by ESI-HRMS/MS are presented in Table 1.

Fig. Y. The fragmentation pattern of reference **A** ester chain. Experimental values for fragment ions identified by ESI-HRMS/MS are presented in Table 1.

Table 1. Fragmentation pattern of saponin references **A-K***.

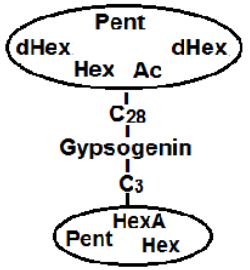
Saponin A $C_{72}H_{112}O_{40}S$	**Saponin B** $C_{72}H_{112}O_{40}S$	**Saponin C** $C_{72}H_{112}O_{40}S$	**Saponin D** $C_{72}H_{112}O_{37}$	**Saponin E** $C_{74}H_{114}O_{38}$	**Saponin F** $C_{74}H_{114}O_{38}$	**Saponin G** $C_{70}H_{110}O_{36}$	**Saponins H/I** $C_{79}H_{122}O_{42}$	**Saponins J/K** $C_{80}H_{118}O_{38}$	**MS-fragments assignment:***	**Formula**	**Theoretical mass**
							1741.733		$[M-H]^-$	$C_{79}H_{121}O_{42}$	1741.7327
							1699.725		$[M-H-Ac]^-$	$C_{77}H_{119}O_{41}$	1699.7221
								1685.723	$[M-H]^-$	$C_{80}H_{117}O_{38}$	1685.7217
1647.639	1647.6475	1647.638							$[M-H]^-$	$C_{72}H_{111}O_{40}S$	1647.6367
				1609.694	1609.693				$[M-H]^-$	$C_{74}H_{113}O_{38}$	1609.6904
							1609.693		$[M-H-Pent]^-$		
1567.671	1567.686	1567.702	1567.682(94.6)[1]						$[M-H]^-$	$C_{72}H_{111}O_{37}$	1567.6799
				1567.679	1568.693				$[M-H-SO_3]^-$ $[M-H-Ac]^-$		
								1553.680	$[M-H-Pent]^-$	$C_{75}H_{109}O_{34}$	1553.6800
						1525.670			$[M-H]^-$	$C_{70}H_{109}O_{36}$	1525.6693
			1525.677[2]						$[M-H-Ac]^- = {}^4Y_{0\alpha}{}^2$		
								1523.669	$[M-H-Hex]^-$	$C_{74}H_{107}O_{33}$	1523.6695
								1507.659	$[M-H-MeCin]^-$	$C_{70}H_{107}O_{35}$	1507.6593
1485.600	1485.6026	1485.589							$[M-H-Hex]^-$	$C_{66}H_{101}O_{35}S$	1485.5839
				1477.657	1477.667				$[M-H-Pent]^-$	$C_{69}H_{105}O_{34}$	1477.6482
			1435.650						$[M-H-Pent]^- = Y_{2\alpha'}$ or $Y_{1\beta'}$	$C_{67}H_{103}O_{33}$	1435.6376
1405.633	1405.6240		1405.632						$[M-H-Hex-SO_3]^-$ $[M-H-Hex]^- = Y_{2\alpha}$ or $Y_{1\beta}$	$C_{66}H_{101}O_{32}$	1405.6270
				1405.636	1405.637				$[M-H-Hex-Ac]^-$		
							1187.580		$[M-H-Chain1-2Ac]^-$	$C_{58}H_{91}O_{25}$	1187.5843
			1237.565						$[M-H-Pent-Hex-2\ H_2O]^- = Z_{2\alpha}/Z_{2\alpha'}$	$C_{61}H_{89}O_{26}$	1237.564
1177.511(19.1)	1177.516	1177.507							$[M-H-Chain1]^-$	$C_{55}H_{85}O_{25}S$	1177.5095
			1127.528						$[M-H-Pent-dHex-Hex]^- = Y_{1\alpha}$	$C_{55}H_{83}O_{24}$	1127.5269
			981.471						$[M-H-Pent-dHex-Hex-Ac-104]^- = Y_{1\alpha}{}^{/0,2}X_{0,\alpha}$	$C_{49}H_{73}O_{20}$	981.469
939.462	939.460	939.458	939.463(95.6)	939.461	939.460	939.461	939.460	939.464	$[M-H-Chain2]^- = Y_{0\alpha}$	$C_{47}H_{71}O_{19}$	939.4584
807.414	807.417	807.420	807.421(4.6)	807.418	807.417	807.418	807.418	807.418	$[M-H-Chain2-Pent]^- = Y_{0\alpha}/Y_{1\beta'}$	$C_{42}H_{63}O_{15}$	807.4161
				789.409		789.405			$Y_{0\alpha}{}^{/0,1}X_{1\beta}$	$C_{42}H_{61}O_{14}$	789.4056
				777.398	777.408	777.408			$[M-H-Chain2-Hex]^- = Y_{0\alpha}/Y_{1\beta}$	$C_{41}H_{61}O_{14}$	….

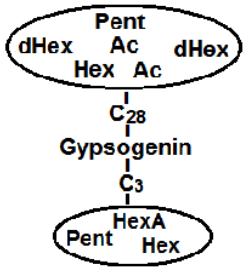
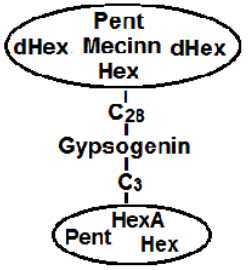
759.383	759.397	759.398	759.397 (15.2)	759.398	759.397	759.397	759.393	759.403	[M-H-Chain2-Hex-H_2O]$^-$ = $Y_{0\alpha}/Z_{1\beta}$	$C_{41}H_{59}O_{13}$	759.3950
								745.263	[Chain2-H]$^-$	$C_{33}H_{45}O_{19}$	745.2633
			745. 417 (8.4)	745.417	745.419	745.419			[M-H-Chain2-Pent-H_2O-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/CO_2$	$C_{41}H_{61}O_{12}$	745.416
			727.410	727.409	727.406	727.406			[M-H-Chain2-Pent-2H_2O-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/H_2O/CO_2$	$C_{41}H_{59}O_{11}$	727.4052
					715.411				[M-H-Chain2-Pent-H_2O-CO_2-30]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/CO_2/^{2,3}X_{1\beta}$	$C_{40}H_{59}O_{11}$	715.4052
707.172 (100)	707.1726	707.172							[Chain2-H]$^-$	$C_{25}H_{39}O_{21}S$	707.1700
					697.401				[M-H-Chain2-Pent-2H_2O-CO_2-30]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/H_2O/CO_2/^{2,3}X_{1\beta}$	$C_{40}H_{57}O_{10}$	697.3946
					673.395				[M-H-Chain2-Pent-H_2O-CO_2-72]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/CO_2/^{0,3}X_{1\beta}$	$C_{38}H_{57}O_{10}$	673.395
647.152 (11.3)	647.151	647.154							[Chain2-H-AcOH]$^-$	$C_{23}H_{35}O_{19}S$	647.1488
					645.359				[M-H-Chain2-Pent-Hex]$^-$ = $Y_{0\alpha}/Y_{1\beta}/Y_{1\beta'}$		645.363
			627.354	627.3571	627.354	627.354			[M-H-Chain2-Pent-Hex-H_2O]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Y_{1\beta}$ or $Y_{0\alpha}/Z_{1\beta}/Y_{1\beta'}$	$C_{36}H_{51}O_9$	627.3528
			609.344 (5.02)	609.343	609.343	609.344			[M-H-Chain2-Pent-Hex-2H_2O]$^-$ = $Y_{0\alpha}/Z_{1\beta}/Z_{1\beta'}$	$C_{36}H_{49}O_8$	609.3422
607.119 (10.5)	607.119								$^{3,5}X_{0\alpha}{}^3$	$C_{20}H_{31}O_{19}S$	……
591.126 (8.3)	591.123								$^{0,3}X_{0\alpha}$	$C_{20}H_{31}O_{18}S$	591.123
583.366	583.363	583.366	583.366 18.5)	583.365	583.365	583.365	583.364	583.367	[M-H-Chain2-Pent-Hex-H_2O-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Y_{1\beta}/CO_2$	$C_{35}H_{51}O_7$	583.3629
579.124	579.124								$^{1,3}X_{0\alpha}$	$C_{19}H_{31}O_{18}S$	…
565.351	565.352	565.356	565.354 (39.7)	565.354	565.354	565.354	565.354	565.354	[M-H-Chain2-Pent-Hex-2H_2O-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Z_{1\beta}/CO_2$	$C_{35}H_{49}O_6$	565.3524
			551.338 (26.0)	551.338	551.338	551.338	551.338	551.340	[M-H-Chain2-Pent-Hex-2H_2O-58]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Z_{1\beta}/^{4,5}A_{0\beta}$	$C_{35}H_{47}O_6$	551.337
			547.343(8.0)	547.342	547.342	547.343	547.343	547.343	[M-H-Chain2-Pent-Hex-3H_2O-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Z_{1\beta}/H_2O/CO_2$	$C_{35}H_{47}O_5$	547.3418
			537.360 (8.8)	537.359	537.359	537.359			[M-H-Chain2-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$ = $Y_{0\alpha}/Y_{1\beta}//H_2O/HCO_2H/CO_2$	$C_{34}H_{49}O_5$	537.3596
527.161 (3.7)	527.163								$X_{2\alpha}/^{3,5}X_{0\alpha}$	$C_{20}H_{31}O_{16}$	527.161

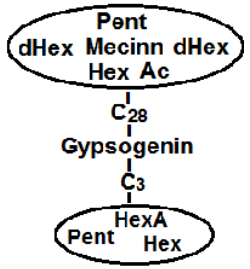
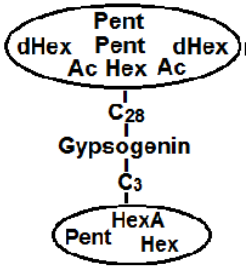
			519.348 (2.07)	519.347	519.350	519.348			[M-H-Chain2-Pent-Hex-2H_2O-HCO_2H-CO_2]$^-$ = $Y_{0\alpha}/Z_{1\beta'}/Y_{1\beta}/2H_2O/HCO_2H/CO_2$	$C_{34}H_{47}O_4$	519.3469
519.103 (29.2)	519.103	519.103							[Chain2-H-dHex-Ac]$^-$= $C_{1\alpha}$	$C_{17}H_{27}O_{16}S$	519.1014
			511.344 (4.1)	511.344	511.345	511.343			[M-H-Chain2-Pent-Hex-134]$^-$ = $Y_{0\alpha}/^{0,2}X_{0\beta}/Y_{1\beta}$	$C_{32}H_{47}O_5$	511.342
503.108 (9.8)	503.109								$Y_{2\alpha'}/^{3,4}X_{0\,\alpha}$	$C_{17}H_{27}O_{15}S$	503.108
			495.348 (2.3)	495.347		495.350			$Y_{0\alpha}/^{0,2}X_{0\beta}/Z_{1\beta}$	$C_{32}H_{47}O_4$	495.3469
469.332 (5.8)	469.333	469.332	469.332 (100)	469.333	469.333	469.332	469.332	469.332	[gypsognin-H]$^-$	$C_{30}H_{45}O_4$	469.3312
451.325	451.325	451.321	451.324 (22.1)	451.322	451.322	451.322	451.322	451.321	[gypsognin-H-H_2O]$^-$	$C_{30}H_{43}O_3$	451.3207
			439.323	439.324	439.322	439.322	439.321	439.323	[gypsogenin-H-H_2CO]$^-$	$C_{29}H_{43}O_3$	439.3207
	423.327		423.328 (13.3)	423.328	423.327	423.328	423.327	423.328	[gypsogenin-H-HCO_2H]$^-$	$C_{29}H_{43}O_2$	423.3258
			405.317	405.316	405.315	405.316			[gypsogenin-H-HCO_2H-H_2O]$^-$	$C_{29}H_{41}O$	405.3152
371.065 (4.2)	371.065	371.063							$Y_{2\alpha'}/C_{1\alpha}$	$C_{12}H_{19}O_{11}S$	371.064
329.019	329.019	329.018							$^{3,5}X_{1\alpha}/C_{1\alpha}$	$C_9H_{13}O_{11}S$	329.022
301.024	301.025								$Y_{2\alpha'}/C_{1\alpha}$	$C_8H_{13}O_{10}S$	301.022
259.013 (8.0)	259.013	259..014							$C_{2\alpha}$	$C_6H_{11}O_9S$	259.012
241.002 (17.0)	241.002	241.002							[Glc-SO_3]$^-$ = $B_{2\alpha}$	$C_6H_9O_8S$	241.0013

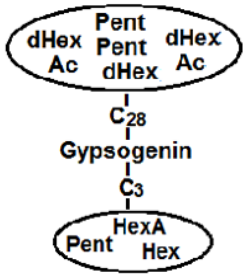
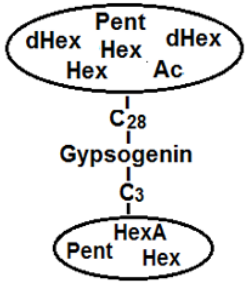
*Hex = hexose, Pent = pentose, dHex = deoxyhexose, HexA = hexosuronic acid; SO_3 = sulfate group,
Ac = acetyl group, MeCin, methoxycinnamoyl group; Chain 1 = C-3 trisaccharide chain, Chain 2 = C-28 ester chain
[1]relative abundance
[2]fragments are related to saponin reference **D (Fig. X)**
[3]fragments are related to saponin reference **A (Fig.Y)**

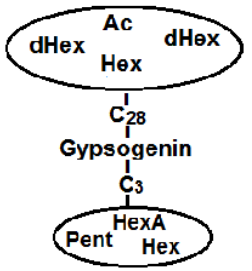
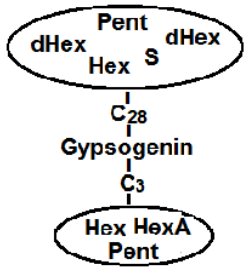
Table 2. Tentative assignment of GOTCAB saponins from *G. trichotoma* roots (GTR) by MS and MS/MS data.

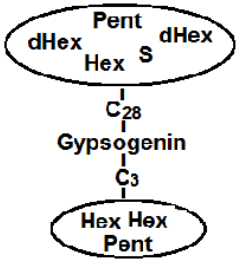
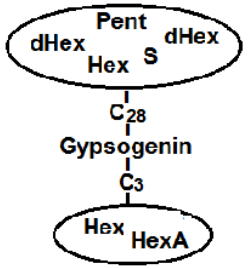
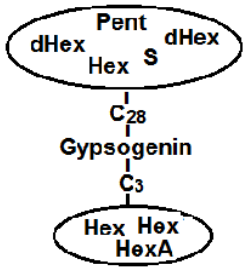
Order No of the separated saponins	[*m/z*] of the molecular $[M-H]^-$ ion (measured)	Number of the separated isobaric isomers			MS/MS fragments of the main isomer [*m/z*]	Interpretation of the structure		[*m/z*] of the molecular $[M-H]^-$ ion (calculated)	Formula and monoisotopic mass (calculated)
		by RP	by HILIC			Fragments	Tentative structure		
		total	total	t_r [min]					
1a	1567.677	2	4	11.72*	1525.679 1405.629 1363.079 1141.776 1097.730 939.461[2] 807.419 759.397 745… 727… 627.354 583.364 565.353 551.337 547.. 537.. 519... 511.343 495.345 469.332 451.322 439.322 423.327 405.316	$[M-H-Ac^1]^-$ $[M-H-Hex]^-$ $[M-H-Hex-Ac]^-$ $[M-H-Hex-2Pent]^-$ [M-H-trisaccharide chain]$^-$ [M-H-ester chain]$^-$ ($Y_{0\alpha}$) $[Y_{0\alpha}-Pent]^-$ $[Y_{0\alpha}-Hex-H_2O]^-$ $[Y_{0\alpha}-Pent-H_2O-CO_2]^-$ $[Y_{0\alpha}-Pent-2H_2O-CO_2]^-$ $[Y_{0\alpha}-Pent-Hex-H_2O]^-$ $[Y_{0\alpha}-Pent-Hex-H_2O-CO_2]^-$ $[Y_{0\alpha}-Pent-Hex-2H_2O-CO_2]^-$ $[Y_{0\alpha}-Pent-Hex-2H_2O-{}^{4,5}A_{0\beta}]^-$ $[Y_{0\alpha}-Pent-Hex-3H_2O-CO_2]^-$ $[Y_{0\alpha}-Pent-Hex-H_2O-HCO_2H-CO_2]^-$ $[Y_{0\alpha}-Pent-Hex-2H_2O-HCO_2H-CO_2]^-$ $Y_{0\alpha}/Y_{1\beta}/{}^{0,2}A_{0\beta}$ $Y_{0\alpha}/Z_{1\beta}/{}^{0,2}A_{0\beta}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$ [gypsogenin-H-HCO_2H-H_2O]$^-$		1567.6804	$C_{72}H_{112}O_{37}$ 1568.6882

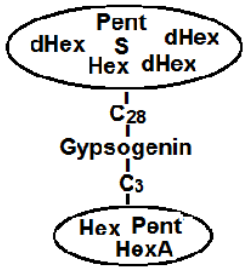
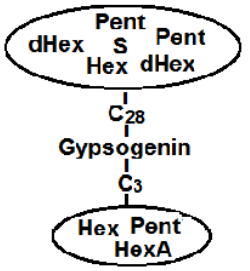
1b 1c 1d	1567.682 1567.683 1567.683			12.05 12.31 12.60	not fragmented				
2a	1609.695	2	2	11.27*	1567.681 1549.669 1477.649 1405.637 981.477 939.461[2] 807.416 759.397 627.353 583.363 565.353 551.338 537.359 469.332 451.321 423.325	[M-H-Ac]$^-$ [M-H-AcOH]$^-$ [M-H-Pent]$^-$ [M-H-Hex-Ac]$^-$ [M-H-ester chain-Pent-Hex-dHex-2Ac-$^{0,2}dHex_{1\alpha}$]$^-$ [M-H-ester chain]$^-$ [M-H-ester chain-Pent]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-$2H_2O$-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-$2H_2O$-$^{4,5}A_{0\beta}$]$^-$ [$Y_{0\alpha}$-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-HCO_2H]$^-$		1609.6910	$C_{74}H_{114}O_{38}$ 1610.6988
2b	1609.694			11.83	not fragmented				
3a	1685.723	2	4	11.42*	939.464[2] 759.403 745.424 583.367 565.354 551.341 537.357 469.332 451.321 439.323 423.327	[M-H-ester chain]$^-$ [M-H-ester chain-Pent]$^-$ [ester chain-H]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-$2H_2O$-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-$2H_2O$-$^{4,5}A_{0\beta}$]$^-$ [$Y_{0\alpha}$-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$		1685.7223	$C_{80}H_{117}O_{38}$ 1686.7301

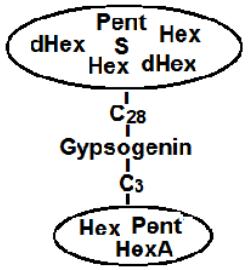
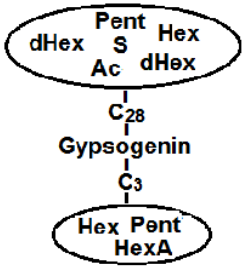
3b 3c 3d	1685.726 1685.723 1685.725			11.75 11.83 11.93	not fragmented				
4a	1727.734	2	4	10.77*	1667.722 1595.681 981.. 939.468[2] 787.272 759.394 627.364 583.362 565.356 551.337 537.358 469.331 451.319 439.320 423.327 177.054 145.028	[M-H-AcOH]$^-$ [M-H-Pent]$^-$ [M-H-ester chain-Pent-Hex-dHex-Ac-Mecinn-^{0,2}dHex$_{1\alpha}$]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ [M-H-ester chain-Hex- H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-2H_2O-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-2H_2O-$^{4,5}A_{0\beta}$]$^-$ [$Y_{0\alpha}$-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$[[gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$ [MeCin-H]$^-$ [MeCin-H-CH_3OH]$^-$		1727.7328	$C_{82}H_{120}O_{39}$ 1728.7407
4b 4c 4d	1727.733 1727.735 1727.736			10.58 10.72 11.32	not fragmented				
5a	1741.732	2	4	11.20*	1699.724 1187.580 939.454[2] 759.404 583.363 565.357 551.339 511.343 469.332 451.320 423.326	[M-H-Ac]$^-$ [M-H-trisaccharidic chain-2Ac]$^-$ [M-H-ester chain]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-2H_2O-CO_2]$^-$ $Y_{0\alpha}$-Pent-Hex-2H_2O-$^{4,5}A_{0\beta}$]$^-$ $Y_{0\alpha}$/$Y_{1\beta}$/ $^{0,2}A_{0\beta}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-HCO_2H]$^-$		1741.7332	$C_{79}H_{122}O_{42}$ 1742.7411

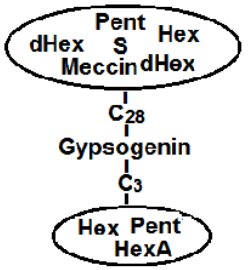
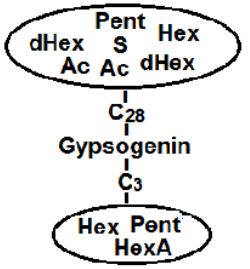
5b 5c 5d	1741.737 1741.735 1741.736			11.64 11.75 12.27	not fragmented				
6a	1725.738	2	4	10.64*	1683.748 1593.704 1171.584 939.459[2] 759.402 583.364 565.353 551.336 537.361 511.345 469.331 423.331	[M-H-Ac]$^-$ [M-H-Pent]$^-$ [M-H-trisaccharidic chain-2Ac]$^-$ [M-H-ester chain]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O- CO_2]$^-$ [M-H-ester chain-Pent-Hex-2H_2O- CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-2H_2O-$^{4,5}A_{0\beta}$]$^-$ [$Y_{0\alpha}$-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$ $Y_{0\alpha}/Y_{1\beta}/{}^{0,2}A_{0\beta}$ [gypsogenin-H]$^-$ [gypsogenin-H-HCO_2H]$^-$		1725.7378	$C_{79}H_{121}O_{42}$ 1726.7388
6b 6c 6d	1725.739 1725.742 1725.743			11.18 11.54 11.64	not fragmented				
7a	1729.733	2	4	13.81*	939.461[2] 807.421 759.396 627.351 583.363 565.351 551.338 511.341 495.347 469.332 451.322 439.331 423.326 405.316	[M-H-ester chain]$^-$ [M-H-ester chain-Pent]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-2H_2O-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-2H_2O-$^{4,5}A_{0\beta}$]$^-$ $Y_{0\alpha}/Y_{1\beta}/{}^{0,2}A_{0\beta}$ $Y_{0\alpha}/Z_{1\beta}/{}^{0,2}A_{0\beta}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$ [gypsogenin-H-HCO_2H-H_2O]$^-$		1729.7332	$C_{78}H_{122}O_{42}$ 1730.7411
7b 7c 7d	1729.734 1729.735 1729.734			13.01 13.25 13.43	Not fragmented				

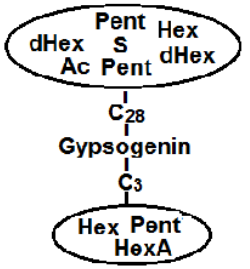
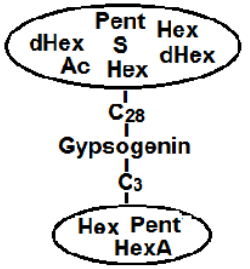
8a	1435.641	2	4	11.39*	939.450[2] 759.395 627.350 583.366 565.353 551.337 537.358 511.342 495.349 469.332 451.321 439.323 423.327 405.314	[M-H-ester chain]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-$2H_2O$-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-$2H_2O$-$^{4,5}A_{0\beta}$]$^-$ [$Y_{0\alpha}$-Pent-Hex-H_2O-HCO_2H-CO_2]$^-$ $Y_{0\alpha}/Y_{1\beta}/\,^{0,2}A_{0\beta}$ $Y_{0\alpha}/Z_{1\beta}/\,^{0,2}A_{0\beta}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$ [gypsogenin-H-HCO_2H-H_2O]$^-$		1435.6382	$C_{67}H_{104}O_{33}$ 1436.6460
8b	1435.639			11.27					
8c	1435.638			11.57	not fragmented				
8d	1435.638			11.95					
9a	1605.628	1	2	11.42*	1525.680 1135.498 939.461 665.160 607.118 579.124 527.162 519.102 503.108 469.331 451.322 423.331 371.066 329.018 301.022 259.012 241.102	[M-H-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ $X_{2\alpha}/^{3,5}X_{0\alpha}$ [ester chain-H-dHex]$^-$($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [gypsogenin-H-HCO_2H]$^-$ $Y_{2\alpha'}/C_{1\alpha}$ $^{3,5}X_{1\alpha}/C_{1\alpha}$ $Y_{2\alpha'}/C_{1\alpha}$ $C_{2\alpha}$ [ester chain-H-2dHex-Pent]$^-$ ($^3X_{2\alpha}$)		1605.6267	$C_{70}H_{110}O_{39}S$ 1606,6345
9b	1605.627			10.98	not				

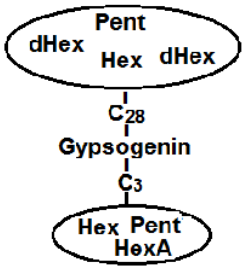
					fragmented				
10a 10b	1591.650 1591.653	2	2	10.81* 11.24	665.160[3] 607.117 579.124 519.102 503.110 241.002 not fragmented	[ester chain-H]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ [ester chain-H-dHex]$^-$($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [ester chain-H-2dHex-Pent]$^-$ ($^3X_{2\alpha}$)		1591.6474	$C_{70}H_{112}O_{38}S$ 1592.6552
11a 11b	1473.580 1473.586	1	2	10.81* 10.40	665.161[3] 607.116 579.124 519.103 503.108 469.330 241.002 not fragmented	[ester chain-H]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ [ester chain-H-dHex]$^-$($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-Pent]$^-$ ($^3X_{2\alpha}$)		1473.5844	$C_{65}H_{102}O_{35}S$ 1474.5922
12a 12b 12c 12d	1635.638 1635.643 1635.642 1635.642	1	4	10.66* 10.89 11.09 11.22	665.162[3] 519.105 469.333 259.012 241.002 not fragmented	[ester chain-H]$^-$ [ester chain-H-dHex]$^-$ [gypsogenin-H]$^-$ ($C_{2\alpha}$) [ester chain-H-2dHex-Pent]$^-$($^3X_{2\alpha}$)		1635.6372	$C_{71}H_{112}O_{40}S$ 1636.6451

13a	1751.674	1	3	10.50*	1671.691 1281.530 811.198 731… 665.145[3] 647.150 607.115 579.124 519.102 503.108 469.329 241.002	[M-H-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [ester chain-H]$^-$ [ester chain-H-SO_3]$^-$ [ester chain-H-dHex]$^-$ [ester chain-H-dHex-H_2O]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ [ester chain-H-2dHex]$^-$($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [gypsogenin-H]$^-$ [ester chain-H-3dHex-Pent]$^-$($^3X_{2\alpha}$)		1751.6846	$C_{76}H_{120}O_{43}S$ 1752.6924
13b	1751.674			10.10	not				
13c	1751.672			10.62	fragmented				
14a	1737.678	1	3	11.66*	797.202 739.157 711.166 665.159[3] 635.151 519.101 371.066 329.018 259.012 469.331 241.002	[ester chain-H]$^-$ [ester chain-H-58]$^-$ [ester chain-H-86]$^-$ [ester chain-H-Pent]$^-$ [ester chain-H-Pent-H_2O]$^-$ [ester chain-H-Pent-dHex]$^-$ ($C_{1\alpha}$) $Y_{2\alpha'}/C_{1\alpha}$ $^{3,5}X_{1\alpha}/C_{1\alpha}$ $C_{2\alpha}$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-2Pent]$^-$($^3X_{2\alpha}$)		1737.6689	$C_{75}H_{118}O_{43}S$ 1738.6768
14b	1737.670			11.48	not				
14c	1737.678			11.97	fragmented				

15a	1767.686	3	3	12.10*	1687.726 1297.556 939.444 827.215 665.162[3] 607.119 579.123 519.104 503.109 469.335 241.002	[M-H-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ [ester chain-H-Hex]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ [ester chain-H-dHex-Hex]$^-$ ($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-Pent-Hex]$^-$ ($^3X_{2\alpha}$)		1767.6795	$C_{76}H_{120}O_{44}S$ 1768.6873
15b	1767.685			12.25	not fragmented				
15c	1767.681			12.72					
16a	1647.639	3	4	11.03*	1177.510 939.464 759.395 707.171[3] 607.118 579.124 519.102 503.108 469.332 451.321 241.002	[M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [$Y_{0\alpha}$-Hex-H_2O]$^-$ [ester chain-H]$^-$ $^{3,5}X_{0\alpha}$ $^{1,3}X_{0\alpha}$ [ester chain-H-dHex-Ac]$^-$ ($C_{1\alpha}$) $Y_{2\alpha'}/^{3,4}X_{0\alpha}$ [gypsogenin-H]$^-$ [gypsogenin-H-H_2O]$^-$ [ester chain-H-2dHex-Pent-Ac]$^-$		1647.6372	$C_{72}H_{112}O_{40}S$ 1648,6451
16b	1647.639			10.43	not fragmented				
16c	1647.638			10.59					
16d	1647.639			10.87					

17a	1765.673	2	2	10.54*	1685.715 1295.554 939.463 825.215[3] 759.397 565.355 551.338 693.171 647.152 607.118 579.124 527.164 519.103 503.109 469.334 241.002 177.054 145.000	[M-H-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ [$Y_{0\alpha}$-Hex-H_2O]$^-$ [$Y_{0\alpha}$-Pent-Hex-2H_2O-CO_2]$^-$ [$Y_{0\alpha}$-Pent-Hex-2H_2O-$^{4,5}A_{0\beta}$]$^-$ [ester chain-H-Pent]$^-$ [ester chain-H-MeCin-H_2O]$^-$ [ester chain-H-(MeCinn+58)]$^-$ [ester chain-H-(MeCin+86)]$^-$ [ester chain-H-(MeCin+86)-SO_3]$^-$ [ester chain-H-(dHex+Mecin)]$^-$ [ester chain-H-Pent-(MeCin+30)]$^-$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-Pent-Mecin]$^-$ [MeCin-H]$^-$ [MeCin-H-CH_3OH]$^-$		1765.6638	$C_{76}H_{118}O_{44}S$ 1766.6717
17b	1765.677			10.12	not fragmented				
18a	1689.650	2	4	10.52*	1609.702 1219.520 939.466 749.183 689.164 649.133 621.136 569.171 561.114 545.123 469.333 371.029 343.034 301.024	[M-H-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ [ester chain-H-AcOH]$^-$ [ester chain-H-(58+Ac]$^-$ [ester chain-H-(86+Ac]$^-$ [ester chain-H-(58+Ac)-SO_3]$^-$ [ester chain-H-(dHex-Ac)]$^-$ [ester chain-H-Pent-(30+Ac]$^-$ [gypsogenin-H]$^-$ [ester chain-H-Pent-(dHex+Ac)-58]$^-$ [ester chain-H-Pent-(dHex+Ac)-86]$^-$ [ester chain-H-Pent-(dHex+Ac)-128]$^-$		1689.6478	$C_{74}H_{114}O_{41}S$ 1690.6556

					283.013 241.002	[ester chain-H-Pent-(dHex+Ac)-H_2O-86]$^-$ [ester chain-H-2dHex-Pent-2Ac]$^-$			
18b 18c 18d	1689.651 1689.651 1689.653			10.08 10.72 10.91	not fragmented				
19a	1779.679	2	3	11.27*	1699.724 1647.628 1537.645 1309.552 939.462[2] 839.212 779.196 739.161 707.171[3] 651.148 635.151 609.345 519.104 469.333 241.001	[M-H-SO_3]$^-$ [M-H-Pent]$^-$ [M-H-Hex-SO_3]$^-$ [M-H-trisaccharidic chain]$^-$ [M-H-ester chain]$^-$ [ester chain-H]$^-$ [ester chain-H-AcOH]$^-$ [ester chain-H-100]$^-$ [ester chain-H-Pent]$^-$ [ester chain-H-dHex-Ac]$^-$ [ester chain-H-Pent-72]$^-$ [M-H-ester chain-Pent-Hex-2H_2O]$^-$ [ester chain-H-Pent-dHex-Ac]$^-$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-2Pent-Ac]$^-$		1779.6795	$C_{77}H_{120}O_{44}S$ 1780.6873
19b 19c	1779.679 1779.684			10.87 11.54	not fragmented				
20a	1809.693	4	3	11.68*	1177.503 869.228 707.171[3] 681.158 607.118 579.216 527.162 519.102 503.104 469.335 241.002	[M-H-trisaccharidic chain-Hex]$^-$ [ester chain-H]$^-$ [ester chain-H-Hex]$^-$ [ester chain-H-dHex-Ac]$^-$ [ester chain-H-Hex-100]$^-$ [ester chain-H-Hex-128]$^-$ [ester chain-H-Hex-SO_3-100]$^-$ [ester chain-H-dHex-Hex-Ac]$^-$ [ester chain-H-Pent-Hex-72]$^-$ [gypsogenin-H]$^-$ [ester chain-H-2dHex-Pent-Hex-Ac]$^-$		1809.6901	$C_{78}H_{122}O_{45}S$ 1810.6979

20b 20c 20d	1809.690 1809.695 1809.691			11.60 11.87 12.36	not fragmented				
21a 21b	1525.672 1525.672	1	2	13.22* 12.68	939.4614 759.3981 627.356 583.364 565.354 469.332 451.322 439.321 423.328 405.316	[M-H-ester chain]$^-$ [M-H-ester chain-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O]$^-$ [M-H-ester chain-Pent-Hex-H_2O-CO_2]$^-$ [M-H-ester chain-Pent-Hex-2H_2O-CO_2]$^-$ [gypsogenin-H]$^-$ [gypsogenin-H- H_2O]$^-$ [gypsogenin-H-H_2CO]$^-$ [gypsogenin-H-HCO_2H]$^-$ [gypsogenin-H-HCO_2H-H_2O]$^-$		1525.6699	$C_{70}H_{110}O_{36}$ 1526.6777

*main isobar

[1]Pen, pentose; Hex, hexose; HexA, hexosuronic acid; dHex, deoxyhexose; S,sulfate group; Ac, acetyl group; MeCin, methoxycinnamoyl group

[2] Fragmentation pattern was consistent with saponin reference **D** (Table 1, Fig. X)

[3] Fragmentation pattern was consistent with saponin references **A-C** and saponin **10a ((**Table 1, Fig. Y)