Metallomics

Water soluble selenometabolome of Cardamine violifolia

Journal:	*Metallomics*
Manuscript ID	Draft
Article Type:	Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Ouerdane, Laurent; Université de Pau et des Pays de l'Adour, e2s UPPA, CNRS, IPREM-UMR5254 Both, Eszter; Szent István University, Faculty of Food Science, Department of Applied Chemistry Xiang, Jiqian; Enshi Autonomous Prefecture Academy of Agriculture Sciences Yin, Hongqing; Enshi Autonomous Prefecture Academy of Agriculture Sciences Yu, Kang; Enshi Autonomous Prefecture Academy of Agriculture Sciences Shao, Shuxun; Institute of Geochemistry Chinese Academy of Sciences Kiszelák, Katalin; Szent István University, Faculty of Food Science, Department of Applied Chemistry Jókainé, Zsuzsanna; Szent István University, Faculty of Food Science, Department of Applied Chemistry Dernovics, Mihály; Centre for Agricultural Research, Dept. Plant Physiology

Water soluble selenometabolome of Cardamine violifolia

Laurent Ouerdane[a], Eszter Borbála Both[b], Jiqian Xiang[c], Hongqing Yin[c], Kang Yu[c], Shuxun Shao[d], Katalin Kiszelák[b], Zsuzsa Jókai[b], Mihály Dernovics[e*]

[a] Université de Pau et des Pays de l'Adour, e2s UPPA, CNRS, IPREM-UMR5254, Hélioparc, 2, Av. Pr. Angot, 64053 Pau, France

[b] Department of Applied Chemistry, Szent István University, Villányi út 29-43., 1118 Budapest, Hungary

[c] Enshi Autonomous Prefecture Academy of Agriculture Sciences, 517 Shizhou Road, Enshi, Hubei Province 445002, China

[d] State Key Laboratory of Ore Deposit Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, 99 Lincheng West Road, Guanshanhu District, Guiyang, Guizhou Province 550081, China

[e] Department of Plant Physiology, Agricultural Institute, Centre for Agricultural Research, Brunszvik u. 2., 2462 Martonvásár, Hungary

*corresponding author

Abstract

Low molecular weight selenium containing metabolites in the leaves of the selenium hyperaccumulator Cardamine violifolia (261 mg total Se / kg d.w.) were targeted in this study. One dimensional cation exchange chromatography coupled to ICP-MS was used for purification and fractionation purposes prior to LC-Unispray-QTOF-MS analysis. Seeking for selenium species in full scan spectra was assisted with an automated mass defect based filtering approach. Besides selenocystathionine, selenohomocystine and its polyselenide derivative, a total number of 35 water soluble selenium metabolites other than selenolanthionine were encountered, including 30 previously unreported compounds. High occurrence of selenium containing hexoses was observed, together with the first assignment of N-glycoside derivatives of selenolanthionine. Quantification of the most abundant selenium species, selenolanthionine, was carried out with an ion pairing LC – post column isotope dilution ICP-MS setup, which revealed that this selenoamino acid accounted for 30% of the total selenium content of the leaf (78 mg (as Se) /kg d.w.).

Keywords

selenosugar, speciation, structure elucidation, IDA

Introduction

Exploration of the largest possible set of selenium species in a sample has not been an evident goal for decades. One of the main driving forces has been to identify one or only a few selenium species that had been adequately separated with chromatography from matrix constituents and the intensity of which had been high enough to be monitored during purification and identification. Clearly, such species could be obtained after following a given sample preparation protocol and a given (often orthogonal) chromatographic set-up (if there was any), which might have disclosed different species or might have undisclosed the same species even from highly similar matrices. Also, as identification has been mostly based on electrospray mass spectrometry, spotting the target compound in the spectra has been a prerequisite that considerably depends on personal experience and related

software features [1-3]. Finally, the ultimate characterization could be mostly achieved in case the structure of the novel compound was to be elucidated without NMR data, that is, the obtained MS/MS fragments, accurate mass, elemental composition and chromatographic information were adequate to support a presumed molecule [4].

Studies on selenised yeast mirror all the options listed above. For example, selenodiglutathione was detected in Se-yeast in 2002 as one out of only two, relatively abundant selenium species [5] and it was re-found as a minor compound out of dozens [6, 7] or almost 200 of selenocompounds [8] – but it wasn't included in the list of >100 species detected in Se-yeast by Gilbert-López et al.[9], though water-soluble Se species were studied in all cases and selenodiglutathione is a common metabolite in Se metabolism. However, thioredoxin reductases bypass selenodiglutathione by directly reducing selenite into selenide, and the activity of these enzymes are highly depending on oxidative stress in yeast (Pinson 2002, Moreno 2012). Therefore, fermentation parameters indirectly influence the abundance of selenodiglutathione in Se yeast batches, as demonstrated by Casal et al. [10] by the comparison of selenometabolomes of Se yeast products of six suppliers: in their study, selenodiglutathione was found in two out of seven samples.

Description of selenometabolome of Se yeast has already covered yeast strains other than Saccharomyces cerevisiae (e.g., Candida utilis [11, 12]) and has practically reached its quasi-industrial application through the characterization of dietary supplements. On the other hand, the actual state-of-the-art knowledge gained on Se yeast samples offered an almost ready-to-use instrumental approach and a database of potential selenium species for plant selenometabolome oriented studies that had been conducted parallel to yeast analysis. While both low and high selenium plant samples became addressed [13, 14] and the list of identified Se species has been still increasing dramatically, it is to mention that the expansion of this metabolite list was not always the function of already characterized Se yeast metabolites. Indeed, several compounds including γ-glutamyl-Se-methyl-selenocysteine [15], γ-glutamyl-selenocystathionine [16] and selenohomolanthionine [17] were first identified in plants before their observation in yeast samples. This indicates a kind of interaction that clearly advances both fields to narrow the "selenium gap", that is, the amount of undetected and/or unidentified selenium compounds [18].

Description of large sets of selenometabolites can help to reveal new metabolic pathways in plants [19-21], and also calls attention to the careful selection of sample preparation protocols [22] in order to avoid artefacts that are often formed in selenium speciation analysis [23-25]. Additionally, the more selenium species are accounted for, the more compounds must be assigned with adequate chromatographic and mass spectrometric resolution. As an example, one-dimensional HPLC separation is hardly enough for unambiguous identification in case only ICP-MS is used with retention time matching as a single tool for compound assignment. This is especially important to consider when a sample is described as bearing one main or only a few characteristic selenium species. Actually, this has been the case with the first selenium hyperaccumulator plant species of the genus Cardamine, C. violifolia. After facing controversial studies on its main water soluble selenometabolite, i.e., selenocystine vs. selenolanthionine, it has been proven that selenocystine cannot account for a considerable part of its selenium content [26-29].

Accordingly, C. violifolia possesses a unique metabolism of selenium that ends up in the accumulation of selenolanthionine, the synthesis and metabolic pathway of which haven't been elucidated yet. Such missing information might be at least partially filled up by discovering the surrogating selenium species that might refer both to the enzymes involved and to important intermediates that serve as precursors or act as landmarks towards selenolanthionine accumulation. As an example, the discovery of Se-methyl-selenomethionine was an important contribution to explain the formation of dimethyl-

selenide in Brassica plants [30]. More recently, in the case of another selenium hyperaccumulator, monkeypot nut (Lecythis minor), the appearance of polyselenides could be attributed to the high abundance of selenohomocysteine and its derivatives [31], which opened up a novel direction in selenium depletion.

Basic goal of our study was to map the water soluble selenometabolome of C. violifolia with the help of a cation exchange liquid chromatography-ICP-MS based purification protocol and LC-Unispray-MS derived identification. Additionally, as even semi-quantitative determination of the main selenium species can considerably contribute to the assignment of metabolic pathways [6], post-column species-unspecific isotope dilution (ID) LC-ICP-MS was also applied. However, this technique was first optimised for selenium species with commercially available standards (e.g., for wheat samples by Huerta et al. [32]), it is still theoretically the most suited method to quantify unknown[33] or standardless species, including artefacts, such as oxidised selenomethionine [34].

Materials and methods

Plant sample

Leaves of Cardamine violifolia (registered by the Wuhan Botanical Garden, Chinese Academy of Sciences; Wuhan, China) were collected in the springtime of 2017 in the natural seleniferous region Yutangba, Enshi (Hubei Province, China), cleaned with deionised water, lyophilised and milled (total Se content: 261 mg Se /kg d.w.).

Reagents and standards

Certified ^{82}Se isotopic abundance solution (10 µg/ml ^{82}Se in 5% nitric acid) was bought from Inorganic Ventures (Christiansburg, VA, USA). Heptafluorobutyric acid (HFBA; ≥99%) and Se-(methyl)selenocysteine hydrochloride (≥95%) were supplied by the Merck-Sigma group (Schnelldorf, Germany). Deionised water (18.2 MΩ cm) was obtained from a Millipore purification system (Merck-Millipore; Darmstadt, Germany). Acetonitrile (UPLC-MS grade) and formic acid (~98% for LC-MS) were supplied by VWR (Radnor, Pennsylvania, USA), while nitric acid (a.r., 65 ≥ m/m%) and formic acid (puriss; 98–100%) were purchased from Scharlau (Barcelona, Spain). Pyridine (99.5%) was obtained from Carlo Elba (Peypin, France), and methanol (HPLC Gradient Grade) was a Fisher Scientific product (Loughborough, UK).

Water extraction

0.5 g of the C. violifolia leaf sample was extracted with an ultrasonic probe (UP100H, Hielscher Ultrasound Technology, Teltow, Germany) at ambient temperature with 10.0 ml deionised water for 1 min. Supernatant was recovered by centrifugation (10 min at 4000*g*), filtered (0.45 µm, cellulose acetate syringe filter) and lyophilised in four aliquots.

LC-ICP-MS setups

Both (strong cation exchange /SCX/ and ion-pairing reversed phase /IP-RP/) chromatographic set-ups were achieved by using an Agilent 1200 HPLC system connected to an Agilent 7500cs ICP-MS for the element-specific detection of ^{76}Se, ^{77}Se, ^{78}Se, ^{79}Br, ^{80}Se, ^{81}Br, ^{82}Se and ^{83}Kr. H_2 was used as collision/reaction gas in the flow rate of 2.5 ml/min. In the case of IP-RP hyphenation, oxygen (40 ml/min) was used as optional gas.

SCX-ICP-MS chromatography

A Zorbax 300-SCX column (150 mm × 4.6 mm × 5 µm; Agilent Technologies, Santa Clara, CA, USA) equipped with a matching guard column was used. Gradient elution was done with pyridine formate (pH 2.2; buffer A: 1 mM; buffer B: 40 mM) delivered at 1.2 ml/min. The program was as follows: 0–2 min, 100% A; 2–15 min, up to 30% B; 15–16 min, up to 100% B; 16–20 min, 100% B; 20–21 min, 100% A. First aliquot of the lyophilized water extract was dissolved in 2 ml deionised water, diluted 15x with eluent buffer. The injection volume was 5 µl (for mapping purposes) or 10 µl (for fraction collection). Selected peaks with high selenium abundance were repeatedly (15x) collected, frozen and lyophilized, then dissolved in the starting eluent of either the IP-RP-ICP-MS or the LC-Unispray-QTOF-MS setups for further analyses.

IP-RP-ICP-MS chromatography and post-column isotope dilution analysis (IDA)

XTerra MS-C_{18} (250 mm × 4.6 mm × 5 µm; Waters, Milford, MA, USA) column was used. The mobile phase consisted of deionised water (eluent A) and methanol (eluent B) both containing 0.05 v/v% HFBA for the analysis of SCX fractions and 0.1 v/v% HFBA in case of IDA. The flow rate was 0.6 ml/min and the gradient elution program was: 0–2 min, 5% B; 2–10 min, up to 65% B; 10–15 min, 65% B; 15–16 min, down to 5% B; 16–19 min 5% B. Injection volume was 40 µl.

For the IDA experiment, the outlet of the RP column was connected through a T-piece where the ^{82}Se standard solution was continuously added in the concentration of 20 ng/g in 0.1% nitric acid with a flow rate of 0.1 mL/min with the help of a peristaltic pump. Mass bias was determined with the repeated injections of 20 µl of Se-methylselenocysteine standard solution (0.5 mg L^{-1} as Se). Second aliquot of the lyophilized water extract of C. violifolia leaf sample was dissolved in 2 ml deionised water, diluted 15x with deionised water containing 0.1 v/v% HFBA. Sample injection volume was 40 µl. IDA calculations were based on the equations published by Rodríguez-González et al. and Koellensperger et al. [35, 36].

LC-Unispray-QTOF-MS set-up

A Vion ion mobility quadrupole time-of-flight mass spectrometer (Waters) equipped with a UniSpray (Waters) ion spray source was applied. Chromatographic elution was provided by an Acquity UPLC I-Class system (Waters) using a BEH-C_{18} reversed phase (RP) UPLC column (100 mm × 2.1 mm × 1.7 µm; Waters). Gradient elution with eluent A (deionised water with 0.1 v/v% formic acid) and eluent B (acetonitrile with 0.1 v/v% formic acid) was carried out at 0.4 ml/min as follows: 0-1 min 10% B, 1-4 min up to 80% B, 4-4.5 min 80% B, 4.5-5 min down to 10% B, 5-7 min 10% B. The UniSpray ion source was used both in positive and negative ionisation modes either in MS^E or $MS^E \rightarrow$ MSMS /DDA/ functions. The related instrumental parameters are described in the Supplementary material (SM Table S1). Data evaluation was carried out with the help of the Unifi software (version 1.9.4; Waters).

Mass defect based Se species filtering was carried out in the frame of an iterative approach. Default settings (mass padding: 15 Da; defect padding: 40 mDa; isotope defect: ^{80}Se-^{78}Se; minimal compound response: 2000 counts) were tested whether they allowed for the detection of known Se species (selenolanthionine and selenocystathionine among others) that had been detected in the fractions by database derived searching. Afterwards, the list of Se species was step-by-step completed with newly detected species resulting from the in-source selenohomocysteine fragment search (m/z 135.9660 /$C_3H_6NSe^+$/ and 181.9715 /$C_4H_8O_2NSe^+$/) and from the manual Se isotopologue pattern search processes; with this, the mass padding and defect padding settings were recursively tuned to include the newly introduced species in the result list. After each tuning step, the updated mass defect based Se species filtering was run on the LC-QTOFMS data of the four fractions to look for still undiscovered

selenium species that met the actual search settings. All positive hits were manually checked and individually validated to remove false positives.

Results and discussion

Characterization of the water extract of C. violifolia with SCX- and IP-RP-ICP-MS chromatographic set-ups

As the water soluble part (approximately 60% of total selenium content) is dominated by compounds with cationic properties at physiological pH, SCX-ICP-MS is the most suited and robust method for the separation of extracted selenium species of C. violifolia [26]. It has also been proven that the main species is selenolanthionine, accounting for about 68% of water soluble selenium content, determined with the help of an in-house synthesised selenolanthionine standard [27]. In Fig. 1a, selenolanthionine elutes at 3.04 min as the most intense peak from the SCX column, and it is surrounded by several less intense peaks containing unknown selenium compounds. These peaks (indicated with Fr1 to Fr4) were fraction-collected for further characterization.

Generally, an orthogonal chromatographic approach is addressed when the complexity of a given chromatographic peak should be assessed [37, 38]. For cationic compounds, ion pairing reversed phase (IP-RP) setups used with anionic ion pairing agents, especially HFBA, have been proven a suitable technique [39, 40]. Fig. 1b presents the overlaid IP-RP-ICP-MS chromatograms of the four fractions. However, the original peaks were slightly separated on SCX, three fractions out of four (Fr1, 2 and 3) could not be clearly differentiated by IP-RP and were eluted in the range of 4.8-6.6 min. These overlaps might be also explained by the overcharged SCX system that could not provide adequate (baseline) separation for the first three fractions that were anyway eluted close to the void volume ($k<2$). Only one fraction (Fr4) showed a well retained and distinct peak at 6.9 min that had no considerable overlap with any of the other fractions. It is also to note that each fraction consisted of a few abundant and several less abundant peaks, which indicated that numerous selenium species should be expected/discovered in the LC-Unispray-MS acquisitions.

In the case of complex mixtures of selenium species three-dimensional orthogonal chromatographic purification/acquisition can even be successfully addressed to achieve adequate analyte purity for organic mass spectrometry based identification, mostly by (Q)TOF-MS or Orbitrap MS. Although such an approach multiplies the number of fractions to be individually analysed, there is usually a higher chance to assign selenized species out of the sample matrix. In our case, the most intense peaks eluting between 4.5-10.5 min in each fraction in the second dimension (IP-RP) were repeatedly collected, concentrated by lyophilisation and subjected to LC-Unispray-QTOF-MS. Contrary to previous results, no selenium species could be discovered this case after the two dimensional (SCX+IP-RP) purification, which might be explained by unstable retention time parameters on IP-RP (i.e., hampering a precise fraction collection procedure). Accordingly, LC-Unispray-QTOF-MS acquisitions were finally carried out from the SCX-ICP-MS derived four fractions.

Screening for selenium species in the LC-Unispray-QTOF-MS acquisitions

Localising unprecedented selenium compounds in full scan mass spectrometric data can be basically achieved by manual pattern exploration, extracting diagnostic ion source fragments and matching retention time data through complementary LC-ICP-MS acquisitions (if available), as summarized by Németh et al. [31]. Manufacturer dependent possibility for the automated data screening of selenium species in LC-ESI-MS analyses was first presented by Preud'homme et al. [7] for the Mass Frontier software (Thermo Fisher Scientific) on the basis of isotopic ratio and intra-isotope mass defect. For Waters-Vion applications, the Unifi software offers the possibility also to introduce an intra-isotope

mass defect based extraction of compounds containing multi-isotopic elements, e.g., chlorine or selenium.

As an effective approach, these techniques must complete the database dependent search, i.e., when previously reported selenium species are individually extracted within a compromised mass accuracy range (generally up to 20 ppm) to end up in extracted ion chromatograms (EICs). Finally, each arising peak in the EICs of all the result lists, that is, those of diagnostic ion source fragments, automated data screening and known selenium species, must be individually and manually checked by verifying at least the selenium specific isotopologue pattern and the mass defect values, in order to avoid false positive hits.

Table 1 presents the list of all detected selenium compounds in the four fractions collected from SCX chromatographic runs. However the first three fractions were evidently co-eluting during fraction collection, their selenium species could be assorted on the basis of intensities, therefore they are positioned into the fraction where they showed the highest abundance. Although most of the species were localised/discovered by more than one search methods, the most contributing technique to provide the first positive hit on the given species is indicated in Table 1. Clearly, manual pattern exploration was crucial for discovering the relatively high (m/z > 370) molecular mass species, while the other methods could successfully localise the lower molecular mass species. As full scan spectra under m/z 300 are of relatively high density ("crowded"), the most visible mass spectral features of selenium containing molecules (i.e., characteristic /7-12 isotopologue wide/ base, unique isotope ratio of ^{80}Se and ^{78}Se, high /0.1-0.2 m/z/ absolute mass defect compared to neighbouring ions) can usually be spotted only in the less crowded parts of the spectra. In other words, any method that helps to localise lower molecular mass selenium species without time consuming manual pattern exploration is evidently useful to provide a lower miss ratio of selenium species.

Definitely, the four methods must be harmonized as they complete and amend each other: as presented for the example of selenocystathionine in FigS1 in the Supplementary Material, one species can provide several selenium patterns in full scan spectra that might be detected independently by the different search methods. Moreover, the individual extraction of all discovered selenium species and compounds must be carried out: EICs must be matched for correct in-source fragment/adduct alignment (see FigS2 in the Supplementary Material for SeHCys in-source fragments) and the presence of isomers should be carefully verified. Also, matching an in-source fragment with its hypothetic mother molecule on the basis of chromatographic data (that is, through retention time and peak shape matching) must be always completed with MS/MS acquisition as closely eluting non-related selenium species might be undiscovered without correctly assigning which in-source fragment belongs to a given parent molecule.

Extracting another characteristic (diagnostic) ion source fragment at m/z 183.9871 ($C_4H_{10}O_2NSe^+$) did not reveal any novel hit, which indicates no derivatives of selenohomolanthionine were present in these fractions in high concentration. Although, any missing hit of a diagnostic ion doesn't automatically exclude the presence of related selenium species, as in-source and MSE fragmentation events usually run under lower fragmentation energy settings compared to directed MS/MS (DDA) analyses. Indeed, this might be a reason why the selenocysteinyl moiety ($C_3H_6O_2NSe^+$; m/z 167.9558) could only be found in MS/MS datasets.

In order to assess the capability of software-based discovery of selenium species, all the 35 compounds detected in the C. violifolia water extract were used to optimise the intra-isotope mass defect filter settings. Any increase in the mass padding or in the defect padding parameters helps to increase the lists of both the true positive and the false positive hits, hitherto a trade-off must be set between high

species recovery and the labour need for selenium pattern/species validation. As presented in Tables S2-S4 in the supplementary material, the setting of 80 Da of mass padding and 200 mDa for defect padding helped to recover 16 selenium species out of 35 involved in the optimization process at the detector response level of 2000 counts, the intensity which is enough to provide adequately high signals for MS/MS acquisition. At these settings, the total number of compounds assigned for pattern validation as potentially selenium carrying molecules ranged between 206-457, which represented 4.5-6.3% of the total number of detected compounds. However these data still seem to indicate too many high false positive hits with only 6-8 true hits per chromatographic fractions, the practical implementation is reasonable: indeed, the outcome of some hundreds of compounds (out of several thousands) can be easily shortlisted on the basis of absolute mass defect and isotopologue redundancy. Moreover, as presented in Table S4, the most frequent reason of species exclusion was low intensity, which anyhow hampers species identification because of the lack of MS/MS data. Also, a considerable part of the species detected by mass defect filtering eluted further from the chromatographic void volume, where the generally hydrophilic selenium species are less evidently occurring (e.g., m/z 197 and m/z 282), which indicates an additional benefit of this search method.

Consequently, the mass defect based species filtering can be regarded as a useful tool for the discovery of previously unknown selenium compounds. This is especially true when no hint on retention time parameters is available, e.g., when no parallel LC-ICP-MS runs can be conducted to decrease the width of the searching window [13] or when HILIC separation is chosen where selenium species are not usually distributed close to the void volume as it is observed in unmodified (that is, non ion-pairing assisted) reversed phase separations.

Although six accurate mass values out of the 35 analytes matched those of previously detected selenium species, only four species could be finally marked as known in Table 1. In two cases, at m/z 254.02899 and m/z 313.02972, the MS/MS fragments of neither of the formerly reported selenium species with the same elemental compositions (that is, N-acetyl-Se-methylselenomethionine [31] and gamma-Glu-Se-methylselenocysteine [41] or N-acetyl-selenocystathionine [42], respectively) matched the fragments detected in our study. This event calls the attention to the fact that accurate mass information in itself is a necessary but not sufficient condition for the unambiguous identification of a selenium species. Moreover, the detection of two selenium species bearing the same accurate mass (m/z 241.99261) but possessing highly different MS/MS spectra also highlights this warning.

No elemental composition could be provided for eleven analytes, including even abundant species with MS/MS spectra rich in fragments. Indeed, unravelling the most possible elemental composition not only requires high mass accuracy (< 2.0 ppm) but the presence of characteristic fragments the composition of which can help to limit the number and kind of contributing elements.

Structural elucidation of novel selenium species

A considerable part of the species listed in Table 1 were either detected in too low abundancy, exposed to isobaric interferences, or featuring low quality MS/MS spectra, which hampered their structural elucidation. For example, the species with the elemental composition of $C9H18NO3Se^+$ (as $[M+H]^+$) showed only one, low intense selenium containing fragment at m/z 133.00 even at low collision energy settings (see the Supplementary Material).

Out of the 35 selenium species detected in C. violifolia water extracts, 13 species were proven with MS/MS analysis to contain the selenohomocysteine moiety. Among these, selenocystathionine [43] is a compound that is a subsequent metabolite with selenohomocysteine in selenium metabolism through the action of cystathionine beta lyase, while selenohomocystine [44] and its polyselenide derivatives [31] might be formed due to oxidation. These three species have been reported to act as non-

proteinaceous storage (deposit) molecules that contribute to the elimination of selenium beside selenolanthionine, the most abundant selenium species in C. violifolia. However not all the other, previously undetected selenohomocysteine derived species could be structurally elucidated, five analytes might be at least tentatively assigned – even if the unambiguous identification does require more elaborative (NMR based) approaches.

(i) the compound detected in Fraction #3 at the theoretical m/z 241.99261 possesses a C2H3O2 moiety with one double bond over the selenohomocysteine fragment. Two -OH groups in vicinal positions should present a considerable neutral loss of -18 Da in the MS/MS spectrum, but no fragment is seen at m/z 224. Accordingly, an end-chain carboxylic group can be regarded more possible, to propose the structure of Se-carboxymethyl-selenohomocysteine (Fig. 2 and Fig. S3). This species has already been hypothetically foreseen in a selenium accumulating plant [45], but it is presented first here with ESI-MS data.

(ii) the compound detected in Fraction #4 at the theoretical m/z 254.02899 has got a C4H7O moiety with one double bond over the selenohomocysteine fragment. The most characteristic features in the relevant MS/MS spectrum are the high intense fragments, namely, m/z 132.96 and 134.97, before the m/z 135.97 (selenohomocysteine with a formic acid loss) fragment that cannot be all assigned as its isotopologues. Accordingly, these fragments must arrive from the other part of the molecule, through the loss of an oxygen, which can only happen in case an –OH group goes during the neutral loss of 18 Da (H_2O). The different ways of rearrangement (differing in two hydrogen atoms) of the leftover moiety can be similar to the case of selenohomocysteine (C4H8NO2Se+ vs. C4H10NO2Se+, with the latter one from the fragmentation of selenohomolanthionine [17]. A tentative structure is therefore presented in Fig. 3 and Fig. S4; definitely, the position of the –OH group might differ.

(iii) the compound detected in Fraction #4 at the theoretical m/z 282.06029 belongs to the relatively hydrophobic species. It shows a C6H11O moiety with one double bond over the selenohomocysteine fragment and the molecule has got a characteristic MS/MS fragment at m/z 163.00 that can only be determined as $C6H11Se^+$. The most feasible way to arrive at a fragment without any oxygen would presume an –OH group that might leave during fragmentation in the form of a neutral loss (18 Da; H_2O). Accordingly, one possible structure is presented in Fig. 4 and Fig. S5.

(iv) in Fraction #4, there was a pair of molecules detected at the theoretical m/z 284.03956 ($C9H18NO4Se^+$). The MS/MS spectra of the two molecules differed in the relative intensities of the m/z 146.97 and 164.98 fragment and the non-selenised m/z 102.05 and 124.04 fragments. None of the spectra showed the sign of a 46 Da loss; this observation might indicate there might not be an end-chain carboxylic group over the selenohomocysteine fragment, therefore this moiety (C5H9O2) contains either a ketone group or two –OH groups with a double bond. However, there are no fragments that would refer to a ketone (serving for a spot of fragmentation). On the other hand, the fragments m/z 146.97 and 164.98 might arrive from subsequent neutral losses of H_2O, which supports the theory that two –OH groups take place in these reactions. The origin of the non-selenised, nitrogen containing fragment at m/z 102.05 (C4H8NO2+) should be connected to the selenohomocysteine moiety, and it may denote to the close (non end-chain) position of the –OH groups which may weaken the C-S-C bonds through the electrophilicity of oxygens in at least one of the isomers. Taking into account all these observations, two possible structures are presented in Fig. 5 and Fig. S6. It must be noted

that the origin of the m/z 124.04 fragment (possibly C6H6NO2+) can be explained only with a post-fragmentation rearrangement, as the abundant intensity of selenohomocysteine derived fragments excludes the possibility of other options.

(v) in Fraction #4, there was a pair of molecules detected at the theoretical m/z 312.03447 ($C_{10}H_{18}NO_5Se^+$). Apart from the selenohomocysteine related fragments, the m/z 146.97, 164.98 and 192.97 should be taken into account that all originate from the other side of the molecule, possessing C6H9O3 elemental composition. None of the three latter fragments possesses more than two oxygens which indicates there should be an –OH group in the structure to give rise a water loss event; also, two out of the three fragments have only five carbon atoms, which might refer to a formic acid loss (-46 Da). Accordingly, an end-chain carboxylic group, a C=C bond and an –OH group should be featured in the isomer molecules; two possible structures are drawn in Fig. 6 and Fig. S7.

In case there is no selenohomocysteine (or other established, selenium containing) moiety in the molecule, the assignment of species becomes more problematic as the position of the Se atom cannot be evidently localised in the structure. In such cases, fragment affiliation should follow an indirect way; clearly, any proposed structure can only be regarded as one of the possibilities.

(i) the lowest molecular mass selenized compound was detected at the theoretical mass of m/z 197.00753 in Fraction #3. As presented in Fig. 7 and Fig. S8, this molecule was surrogated by a lower intense + 18 Da analyte at m/z 215.01720. As the fragmentation of the m/z 215 molecule hardly provided MS/MS fragments, it can be taken as the in-source water adduct ($[M+H_2O+H]^+$) of the m/z 197 molecule. The two main MS/MS fragments of this molecule have the SeCH3 and the SeC2H5 compositions (m/z 94.94 and 108.95, respectively), which supports the theory that a ketone borders the SeC2H5 moiety, providing a spot of fragmentation. Taking into account that the composition of C6H13O2Se+ allows for only one double bond, and an ester or ether bound would result in other fragments too, a structure with an –OH group might be suggested, as shown in Fig. 7.

(ii) in Fraction #2, there was a pair of molecules detected at the theoretical m/z 285.05996 ($C_{10}H_{21}O_4Se^+$) with similar MS/MS spectra, differing in fragment ion ratios especially in terms of the m/z 85.06, 134.97, 164.98, 182.99 and 239.06 fragments (Fig. 8 and Fig. S9). This latter fragment refers to a formic acid loss, which usually arrives from an end-chain carboxylic group. The most intense, oxygen free and selenium containing fragments at m/z 134.97 (C4H7Se+) and m/z 132.96 (C4H5Se+) might refer to a structure where these fragment can be created from the end of the molecule, through a formic acid loss. Both isomers show the m/z 106.94 (C2H3Se+) fragment as well. The other main selenium containing fragment has got one oxygen and one double bond in its structure (m/z 164.98; C5H9OSe+), which practically excludes the presence of a ketone, and limits the structure to a C=C double bond and an –OH group; the double bond, in turn, might originate in the neutral loss of a H_2O molecule. This fragment can be affiliated to the fragment at m/z 85.06 (C5H9O+) after the loss of the Se atom and can be also linked to the fragment at m/z 182.99 (C5H11O2Se+), which still contains two oxygen atoms. The fact that the most intense fragments contain either four or five carbon atoms might indicate that the Se atom is localised in the middle of the molecule. The difference in the intensities of the m/z 164.98 fragment can be the result of the different position of two –OH groups, e.g., vicinal vs. non-vicinal position. Accordingly, two possible structures are presented in Fig. 8.

Eleven species out of 35 showed the evident characteristic loss of hexoses (162.05 Da) during MS/MS fragmentation and three of them could be analogously ($[M-H]^-$) detected in negative ion mode as well. These molecules, accounting one-third of all selenium species detected in C. violifolia and containing highly abundant ones too, are presumed selenium containing sugars: this group of selenium species in plants was presented first in 2012 [46]. One of them (at m/z 407.04510) might already be considered a ubiquitous selenium species, still without a structural elucidation, as it has been detected in several samples including crops and vegetables [46, 47].

Identification of selenosugars cannot be carried out without possessing at least mg-sized amount of their purified form for NMR studies. However, some features can be noted on the basis of their characteristic MS/MS fragments:

(i) Selenosugars often share the same specific fragments, creating groups with the assumption of common core structures; that is, the species with m/z 419 and m/z 581, share the selenium containing fragments of 167.95, 257.00, 401.04 and non-selenized fragments of 214.07 and 232.08 as well. The fragment at m/z 401.04 appears as an ion-source fragment too. The MS/MS spectra of the two species feature a series of -18 Da losses that is also characteristic to hexoses. Indeed, the appearance of the m/z 167.95 fragment is in itself is a specific reference to a selenocysteinyl moiety, which – together with the m/z 257.00 fragment – indicates these molecules can be assigned as the mono- or di-N-glycosides of selenolanthionine (Fig. 9, Fig. S10 and S11). Such N-glycosyl conjugates of (seleno)amino acids and carbohydrates can be formed by their direct reactions, or at alkaline pH, and through condensation in the Maillard reaction scheme when exposed to elevated temperature [48]. It is unclear whether these selenium species were formed ex vivo during the freeze-drying process or spontaneously in vivo; however, due to the high availability of non-proteinaceous selenolanthionine, their appearance calls the attention to this new class of selenium species as other non-proteinaceous selenoamino acids, especially Se-methlyselenocysteine, might be involved in such reactions in many selenium accumulator plants.

(ii) Beside the hexose loss (162.05 Da), the loss of a deoxyhexose moiety (146.06 Da) is also a characteristic event in the fragmentation of oligosaccharides. In Table 1, there are two unknown selenohexoses at m/z 377.07 and m/z 391.09 that can be linked to a couple of molecules sharing the core molecules together with an additional + 146.06 Da sized moiety (at m/z 523.13 and m/z 537.15, respectively). Both selenohexoses and both selenohexose-deoxyhexose molecules were found to possess several isomers, which also reinforces their affiliation. However the neither the core molecules, nor the aglycones can be unambiguously assigned, the discovery of selenium-containing disaccharides opens a new class of molecules that can serve as selenium depository in the exclusion of excess selenium.

(iii) Additionally, in order to propose elemental composition data for unknown selenium compounds, all detected species possessing either selenohomocysteine or selenocysteine moieties (or, in other words, a free primary amine group) were checked for an N-glycated hexose couple, analogously to N-glycosyl selenolanthionine. Practically, this process involved the accurate mass matching between the unassigned selenium species and the assigned species completed with the C6H10O5 conjugated hexose composition (i.e., C6H12O6-H2O). Finally, taking into account extra water addition on unsaturated sidechains, three molecules (indicated with '(x)' in Table 1) could be selected and presented with a hypothetical elemental composition (m/z 416 /254+162/, 434 /254+18+162/ and 446 /284+162/). None of the three species showed the evident loss of

162.05 Da in their MS/MS spectra, therefore their selenosugar property can only be presumed. However the MS/MS fragments of these compounds were of low or medium abundance, two out of the three species might be featured with a possible structure according to the hypothetical N-glycosylation and due to the matching of some of the fragments (Fig. S12, S13 and S14).

Quantification of selenolanthionine with post-column IP-RP-IDA-ICP-MS technique

Apart from the main species, selenolanthionine, the water extract of C. violifolia contained more than 30 selenium species with highly different intensities and concentration. Facing that none of the species are available commercially as standards, no direct quantification with LC-Unispray-QTOF or other (triple quadrupole) MS setup could be carried out. Also, for LC-ICP-MS, the baseline separation of all the compounds could not be done with the available chromatographic methods; moreover, some of the moderately abundant species showed highly similar retention characteristics both on SCX and RP columns, which would have resulted in a biased quantification even if isocratic elution method would have been addressed [49]. In such cases, dilution helps to clean up the chromatogram from low(er) intensity selenium species, which finally ends up in only a few species whose purity can be regarded adequate for quantification through post-column isotope dilution analysis (IDA).

Taking into account the dilution steps, the final dilution ratio of the water extract was 300 ml for 1.0 g lyophilised sample to achieve adequate resolution on the IP-RP setup (Fig. S15). The first abundant peak was not quantified as peaks eluting in the void may contain several unresolved species. Although the peak eluting at 9.4 min was found to contain mostly selenocystathionine, several SeHCys-derived compounds were also detected (Fig. S16). Therefore, only the selenolanthionine peak eluting at 6.8 min was evaluated. The concentration of this species as Se was 77.6 mg/kg that accounts for approximately 30% of the total selenium content of the leaf. This value can be slightly biased by minor selenium species that still remained unresolved/undiscovered in the chromatographic peak. This contribution from selenolanthionine to total Se content is 25% less than the value determined in a highly selenised [27] pooled (stem + leaves) C. violifolia biomass with the help of an in-house synthesised selenolanthionine standard, which might indicate that the relative concentration of this non-proteinaceous selenoamino acid increases by the higher selenium accumulation rate. This behaviour – together with the higher complexity of selenium speciation with dozens of selenium species – might indicate that C. violifolia selenium biotransformation capacity cannot be exhausted even at high (several grams of Se kg^{-1} plant biomass, d.w.) selenium load. The transcriptomical – biochemical background of this feature is still to be revealed [50].

Conclusions

This study focused on a plant sample that accumulated selenium under natural conditions, without reaching extreme (> 1 g Se kg^{-1} d.w.) selenium concentration, and apart from possible oxidation, the plant biomass was not exposed to any sample preparation step (e.g., drying at high temperature, alkaline extraction, etc.) that would result in considerably altered speciation. Still, dozens of novel selenium species could be detected beside the major (selenolanthionine) and previously described compounds (i.e., selenocystathionine, selenohomocystine). The high abundance of hexose and deoxyhexose containing selenium species reflects this group of selenium compounds should be moved into the focus of selenium speciation as they might represent a significant part of total selenium content in selenised plant biomass.

Definitely, as Cardamine genus is not a highly studied family in the Brassicaceae family in terms of selenium speciation, there is hardly enough information to declare whether the higher reactivity of selenium (compared to sulphur) is solely responsible for the formation of these species, or there is a dedicated route in the selenium detoxification pathway, or a modified scheme (e.g., C_X-sulphonate biosynthesis, where X=3 or higher) contributes to such a complex selenium speciation. It is however important to highlight that overall and comprehensive description of selenium compounds in plant species can be provided only through the combination of inorganic (LC-ICP-MS) and high resolution organic mass spectrometry set-ups. Moreover, the list of selenium species to expect in any plant sample is far from complete, which can be considered a call for analysts to be prepared for unprecedented species at definitely non-minor abundancies.

Acknowledgements

E. B. Both acknowledges the support of the Doctoral School of Food Science SZIU. The support of the Science and Technology Department of Hubei Province (project code: 2020BJH024, Research on High-efficiency Utilization of Medicine and Food Homologous Plant Resources and Biologically Active Selenopeptide) is acknowledged. S. Shao acknowledges the support of the National Science Foundation of China (grant no. 40971287) M. Dernovics acknowledges Bolyai János Research Scholarship of the Hungarian Academy of Sciences. This research was supported by the Hungarian Government and the European Union, with the co-funding of the European Regional Development Fund in the frame of Széchenyi 2020 Program GINOP-2.3.2-15-2016-00029 project. The authors thank the kind professional contribution of S. Hann (BOKU, Vienna, Austria) and J.R. Encinar (University of Oviedo, Spain) to the IDA calculation process.

Table caption

Table 1

Summary of selenium species detected in the water soluble selenometabolome of C. violifolia. 'n.p.' denotes to 'not provided'; 'n.a.' denotes to 'not available'. Relative intensity scaling indicate high abundance with good quality MS/MS spectrum obtained ('+++'), moderate abundance with moderate quality MS/MS spectrum ('++') and low abundance with missing or low quality MS/MS spectrum ('+'). The signs '(x)' indicate hypothetical selenosugar species.

Figure captions

Figure 1: (a) Strong cation exchange (SCX) – ICP-MS chromatogram of water soluble extract of C. violifolia, recorded on the ^{82}Se isotope. Fractions labelled #1 - #4 were collected for further characterization on LC-Unispray-QTOFMS. (b) Ion-pairing reversed phase (IP-RP) – ICP-MS chromatograms of the fractions #1 - #4 collected from SCX – ICP-MS, recorded on the ^{77}Se isotope.

Figure 2: Full scan spectrum (a), zoomed full scan spectrum (b) and MS/MS spectrum (c) of the compound with the experimental m/z 241.99302, presented and putatively identified as Se-carboxymethyl-selenohomocysteine in the inset.

Figure 3: Full scan spectrum (a), zoomed full scan spectrum (b) and MS/MS spectrum (c) of the selenohomocysteine-derived compound with the experimental m/z 254.02883, presented with a putative structure in the inset.

Figure 4: Full scan spectrum (a), zoomed full scan spectrum (b) and MS/MS spectrum (c) of the selenohomocysteine-derived compound with the experimental m/z 282.06027, presented with a putative structure in the inset.

Figure 5: Full scan spectrum (a), zoomed full scan spectrum (b), MS/MS spectrum of the more hydrophobic isomer (c) and MS/MS spectrum of the less hydrophobic isomer (d) of the selenohomocysteine-derived compounds with the experimental m/z 284.03997, presented with putative structures in the inset.

Figure 6: Full scan spectrum (a), zoomed full scan spectrum (b), and MS/MS spectrum of the less hydrophobic isomer (c) of the selenohomocysteine-derived compounds with the experimental m/z 312.03454, presented with putative structures in the inset. MS/MS spectrum of the more hydrophobic isomer couldn't be recorded because of low abundance.

Figure 7: Full scan spectrum (a), zoomed full scan spectrum (b), low fragmentation energy MS/MS spectrum (c) and high fragmentation energy MS/MS spectrum (d) of the compound with the experimental m/z 197.00725, presented with a putative structure in the inset.

Figure 8: Full scan spectrum (a), zoomed full scan spectrum (b), MS/MS spectrum of the less hydrophobic isomer (c) and MS/MS spectrum of the more hydrophobic isomer (d) of the compounds with the experimental m/z 285.05998, presented with putative structures in the inset.

Figure 9: Putative mono- and di-N-glycosides of selenolanthionine. Full scan spectrum (a), zoomed full scan spectrum (b), MS/MS spectrum (c) and structure of selenolanthionine-di-N-glycoside (d). Full scan spectrum (e), zoomed full scan spectrum (f), MS/MS spectrum (g) and structure of selenolanthionine-N-glycoside (h). Additional MS/MS spectrum of selenolanthionine-di-N-glycoside can be found in the Supplementary Material.

References

1. Y. Anan, G. Nakajima and Y. Ogra, Complementary use of LC-ICP-MS and LC-ESI-Q-TOF-MS for selenium speciation, *Anal Sci*, 2015, **31**, 561-564.
2. K. Bierla, S. Godin, R. Lobinski and J. Szpunar, Advances in electrospray mass spectrometry for the selenium speciation: Focus on Se-rich yeast, *TRAC-Trend. Anal. Chem.*, 2018, **104**, 87-94.
3. B. Gammelgaard, C. Gabel-Jensen, S. Stürup and H. R. Hansen, Complementary use of molecular and element-specific mass spectrometry for identification of selenium compounds related to human selenium metabolism, *Anal. Bioanal. Chem.*, 2008, **390**, 1691-1706.
4. E. Kurek, M. Michalska-Kacymirow, A. Konopka, O. Kosciuczuk, A. Tomiak and E. Bulska, Searching for low molecular weight seleno-compounds in sprouts by mass spectrometry, *Molecules*, 2020, **25**.
5. T. Lindemann and H. Hintelmann, Identification of selenium-containing glutathione S-conjugates in a yeast extract by two-dimensional liquid chromatography with inductively coupled plasma MS and nanoelectrospray MS/MS detection, *Anal. Chem.*, 2002, **74**, 4602-4610.
6. C. Arnaudguilhem, K. Bierla, L. Ouerdane, H. Preud'homme, A. Yiannikouris and R. Lobinski, Selenium metabolomics in yeast using complementary reversed-phase/hydrophilic ion interaction (HILIC) liquid chromatography-electrospray hybrid quadrupole trap/Orbitrap mass spectrometry, *Anal. Chim. Acta*, 2012, **757**, 26-38.
7. H. Preud'Homme, J. Far, S. Gil-Casal and R. Lobinski, Large-scale identification of selenium metabolites by online size-exclusion-reversed phase liquid chromatography with combined inductively coupled plasma (ICP-MS) and electrospray ionization linear trap-Orbitrap mass spectrometry (ESI-MSn), *Metallomics*, 2012, **4**, 422-432.
8. P. Ward, M. Chadha, C. Connolly, A. Stalcup and R. Murphy, A comparative assessment of water-soluble selenium metabolites in commercial selenised yeast supplements by liquid chromatography-electrospray ionisation QTOF-MS, *Int. J. Mass Spectrom.*, 2019, **439**, 42-52.
9. B. Gilbert-López, M. Dernovics, D. Moreno-González, A. Molina-Díaz and J. F. García-Reyes, Detection of over 100 selenium metabolites in selenized yeast by liquid chromatography electrospray time-of-flight mass spectrometry, *J Chromatogr B*, 2017, **1060**, 84-90.
10. S. G. Casal, J. Far, K. Bierla, L. Ouerdane and J. Szpunar, Study of the Se-containing metabolomes in Se-rich yeast by size-exclusion - Cation-exchange HPLC with the parallel ICP MS and electrospray orbital ion trap detection, *Metallomics*, 2010, **2**, 535-548.
11. K. Bierła, N. Suzuki, Y. Ogra, J. Szpunar and R. Łobiński, Identification and determination of selenohomolanthionine – The major selenium compound in Torula yeast, *Food Chem.*, 2017, **237**, 1196-1201.
12. M. Kieliszek and S. Błazejak, Speciation analysis of selenium in Candida utilis yeast cells using HPLC-ICP-MS and UHPLC-ESI-Orbitrap MS techniques, *Appl Sci-Switzerland*, 2018, **8**.
13. F. A. Aborode, A. Raab, S. Foster, E. Lombi, W. Maher, E. M. Krupp and J. Feldmann, Selenopeptides and elemental selenium in Thunbergia alata after exposure to selenite: quantification method for elemental selenium, *Metallomics*, 2015, **7**, 1056-1066.
14. K. M. Kubachka, J. Meija, D. L. Leduc, N. Terry and J. A. Caruso, Selenium volatiles as proxy to the metabolic pathways of selenium in genetically modified Brassica juncea, *Envir Sci Tech*, 2007, **41**, 1863-1869.
15. S. N. Nigam and W. B. McConnell, Seleno amino compounds from Astragalus bisulcatus isolation and identification of gamma-L-glutamyl-Se-methyl-seleno-L-cysteine and Se-methylseleno-L-cysteine, *BBA - General Subjects*, 1969, **192**, 185-190.
16. S. N. Nigam and W. B. McConnell, Isolation and identification of two isomeric glutamylselenocystathionines from the seeds of Astragalus pectinatus, *Biochim. Biophys. Acta*, 1976, **437**, 116-121.

17. Y. Ogra, T. Kitaguchi, K. Ishiwata, N. Suzuki, Y. Iwashita and K. T. Suzuki, Identification of selenohomolanthionine in selenium-enriched Japanese pungent radish, *J. Anal. At. Spectrom.*, 2007, **22**, 1390-1396.
18. B. Lajin, D. Kuehnelt and K. A. Francesconi, Exploring the urinary selenometabolome following a multi-phase selenite administration regimen in humans, *Metallomics*, 2016, **8**, 774-781.
19. K. Wrobel, M. Guerrero Esperanza, E. Yanez Barrientos, A. R. Corrales Escobosa and K. Wrobel, Different approaches in metabolomic analysis of plants exposed to selenium: a comprehensive review, *Acta Physiol Plant*, 2020, **42**.
20. P. J. White, Selenium metabolism in plants, *Biochim. Biophys. Acta - General Subjects*, 2018, **1862**, 2333-2342.
21. L. W. Lima, E. A. H. Pilon-Smits and M. Schiavon, Mechanisms of selenium hyperaccumulation in plants: A survey of molecular, biochemical and ecological cues, *Biochim. Biophys. Acta - General Subjects*, 2018, **1862**, 2343-2353.
22. L. Ouerdane, F. Aureli, P. Flis, K. Bierla, H. Preud'homme, F. Cubadda and J. Szpunar, Comprehensive speciation of low-molecular weight selenium metabolites in mustard seeds using HPLC - electrospray linear trap/orbitrap tandem mass spectrometry, *Metallomics*, 2013, **5**, 1294-1304.
23. M. Dernovics, L. Ouerdane, L. Tastet, P. Giusti, H. Preud'homme and R. Lobinski, Detection and characterization of artefact compounds during selenium speciation analysis in yeast by ICP-MS-assisted MALDI MS, oMALDI MS/MS and LC-ES-MS/MS, *J. Anal. At. Spectrom.*, 2006, **21**, 703-707.
24. E. Lipiec, G. Siara, K. Bierla, L. Ouerdane and J. Szpunar, Determination of selenomethionine, selenocysteine, and inorganic selenium in eggs by HPLC-inductively coupled plasma mass spectrometry, *Anal. Bioanal. Chem.*, 2010, **397**, 731-741.
25. A. Polatajko, B. Banas, J. R. Encinar and J. Szpunar, Investigation of the recovery of selenomethionine from selenized yeast by two-dimensional LC-ICP MS, *Anal. Bioanal. Chem.*, 2005, **381**, 844-849.
26. E. B. Both, G. C. Stonehouse, L. W. Lima, S. C. Fakra, B. Aguirre, A. L. Wangeline, J. Xiang, H. Yin, Z. Jókai, Á. Soós, M. Dernovics and E. A. H. Pilon-Smits, Selenium tolerance, accumulation, localization and speciation in a Cardamine hyperaccumulator and a non-hyperaccumulator, *Sci Total Environ*, 2020, **703**.
27. E. B. Both, S. Shao, J. Xiang, Z. Jókai, H. Yin, Y. Liu, A. Magyar and M. Dernovics, Selenolanthionine is the major water-soluble selenium compound in the selenium tolerant plant Cardamine violifolia, *Biochim. Biophys. Acta - General Subjects*, 2018, **1862**, 2354-2362.
28. L. Yuan, Y. Zhu, Z. Q. Lin, G. Banuelos, W. Li and X. Yin, A Novel Selenocystine-Accumulating Plant in Selenium-Mine Drainage Area in Enshi, China, *PLoS ONE*, 2013, **8**.
29. L. Cui, J. Zhao, J. Chen, W. Zhang, Y. Gao, B. Li and Y. F. Li, Translocation and transformation of selenium in hyperaccumulator plant Cardamine enshiensis from Enshi, Hubei, China, *Plant Soil*, 2018, **425**, 577-588.
30. B. G. Lewis, C. M. Johnson and T. C. Broyer, Volatile selenium in higher plants the production of dimethyl selenide in cabbage leaves by enzymatic cleavage of Se-methyl selenomethionine selenonium salt, *Plant Soil*, 1974, **40**, 107-118.
31. A. Németh, J. F. García Reyes, J. Kosáry and M. Dernovics, The relationship of selenium tolerance and speciation in Lecythidaceae species, *Metallomics*, 2013, **5**, 1663-1673.
32. V. D. Huerta, L. H. Reyes, J. M. Marchante-Gayón, M. L. F. Sánchez and A. Sanz-Medel, Total determination and quantitative speciation analysis of selenium in yeast and wheat flour by isotope dilution analysis ICP-MS, *J. Anal. At. Spectrom.*, 2003, **18**, 1243-1247.
33. M. Sánchez-Martínez, T. Pérez-Corona, C. Martínez-Villaluenga, J. Frías, E. Peñas, J. M. Porres, G. Urbano, C. Cámara and Y. Madrid, Synthesis of [77 Se]-methylselenocysteine when preparing sauerkraut in the presence of [77 Se]-selenite. Metabolic transformation of [77 Se]-methylselenocysteine in Wistar rats determined by LC-IDA-ICP-MS, *Anal. Bioanal. Chem.*, 2014, **406**, 7949-7958.

34. A. A. Krata, M. Wojciechowski, J. Karasinski and E. Bulska, Comparative study of high performance liquid chromatography species-specific and species-unspecific isotope dilution inductively coupled plasma mass spectrometry. A case study of selenomethionine and the origin of its oxidized form, *Microchem. J.*, 2018, **143**, 416-422.
35. P. Rodríguez-González, J. M. Marchante-Gayón, J. I. García Alonso and A. Sanz-Medel, Isotope dilution analysis for elemental speciation: A tutorial review, *Spectrochim Acta B*, 2005, **60**, 151-207.
36. G. Koellensperger, S. Hann, J. Nurmi, T. Prohaska and G. Stingeder, Uncertainty of species unspecific quantification strategies in hyphenated ICP-MS analysis, *J. Anal. At. Spectrom.*, 2003, **18**, 1047-1055.
37. S. McSheehy, P. Pohl, J. Szpunar, M. Potin-Gautier and R. ŁObiński, Analysis for selenium speciation in selenized yeast extracts by two-dimensional liquid chromatography with ICP-MS and electrospray MS-MS detection, *J. Anal. At. Spectrom.*, 2001, **16**, 68-73.
38. L. H. Reyes, J. R. Encinar, J. M. Marchante-Gayón, J. I. G. Alonso and A. Sanz-Medel, Selenium bioaccessibility assessment in selenized yeast after "in vitro" gastrointestinal digestion using two-dimensional chromatography and mass spectrometry, *J. Chromatogr. A*, 2006, **1110**, 108-116.
39. O. Egressy-Molnár, L. Ouerdane, J. Gyorfi and M. Dernovics, Analogy in selenium enrichment and selenium speciation between selenized yeast Saccharomyces cerevisiae and Hericium erinaceus (lion's mane mushroom), *LWT - Food Sci Technol*, 2016, **68**, 306-312.
40. M. Kotrebai, M. Birringer, J. F. Tyson, E. Block and P. C. Uden, Selenium speciation in enriched and natural samples by HPLC-ICP-MS and HPLC-ESI-MS with perfluorinated carboxylic acid ion-pairing agents, *Analyst*, 2000, **125**, 71-78.
41. S. McSheehy, W. Yang, F. Pannier, J. Szpunar, R. Lobinski, J. Auger and M. Potin-Gautier, Speciation analysis of selenium in garlic by two-dimensional high-performance liquid chromatography with parallel inductively coupled plasma mass spectrometric and electrospray tandem mass spectrometric detection, *Anal. Chim. Acta*, 2000, **421**, 147-153.
42. M. Dernovics, J. Far and R. Lobinski, Identification of anionic selenium species in Se-rich yeast by electrospray QTOF MS/MS and hybrid linear ion trap/orbitrap MSn, *Metallomics*, 2009, **1**, 317-329.
43. T. K. Virupaksha and A. Shrift, Biosynthesis of selenocystathionine from selenate in Stanleya pinnata, *BBA - Biochimica et Biophysica Acta*, 1963, **74**, 791-793.
44. J. W. Hamilton, Chemical Examination of Seleniferous Cabbage Brassica oleracea capitata, *J. Agric. Food Chem.*, 1975, **23**, 1150-1152.
45. S. Shao, G. Deng, X. Mi, S. Long, J. Zhang and J. Tang, Accumulation and speciation of selenium in Cardamine sp. In Yutangba Se Mining Field, Enshi, China, *Chinese J Geochem*, 2014, **33**, 357-364.
46. F. Aureli, L. Ouerdane, K. Bierla, J. Szpunar, N. T. Prakash and F. Cubadda, Identification of selenosugars and other low-molecular weight selenium metabolites in high-selenium cereal crops, *Metallomics*, 2012, **4**, 968-978.
47. A. Ruszczyńska, A. Konopka, E. Kurek, J. C. Torres Elguera and E. Bulska, Investigation of biotransformation of selenium in plants using spectrometric methods, *Spectrochim Acta B*, 2017, **130**, 7-16.
48. J. P. Danehy and W. W. Pigman, Reactions between Sugars and Nitrogenous Compounds and Their Relationship to Certain Food Problems.*Journal*, 1951, **3**, 241-290.
49. J. Far, H. Preud'homme and R. Lobinski, Detection and identification of hydrophilic selenium compounds in selenium-rich yeast by size exclusion-microbore normal-phase HPLC with the on-line ICP-MS and electrospray Q-TOF-MS detection, *Anal. Chim. Acta*, 2010, **657**, 175-190.
50. Y. Zhou, Q. Tang, M. Wu, D. Mou, H. Liu, S. Wang, C. Zhang, L. Ding and J. Luo, Comparative transcriptomics provides novel insights into the mechanisms of selenium tolerance in the hyperaccumulator plant Cardamine hupingshanensis, *Sci Rep-UK*, 2018, **8**.

Experimental m/z	Elemental composition, $[M+H]^+$	Theoretical m/z	Difference ppm	Rt, min	Detected in fraction				Isomers detected	Detected also in negative ion mode	Contains one or two (deoxy)hexose moiety(ies)	Detected first on the basis of				SeHCys derivative	Characteristic non-SeHCys fragment on ^{80}Se	Characteristic non-Se fragment	Relative intensity	Reference
					#1	#2	#3	#4				manual screening	mass defect	database search	SeHCys in-source fragments					
197.00725	C6H13O2Se+	197.00753	-1.42	2.13			x						x			-	94.94, 108.95	-	+++	
241.99302	C6H12NO4Se+	241.99261	1.69	0.74			x								x	x	108.95	-	++	
241.99296	C6H12NO4Se+	241.99261	1.45	1.48	x							x				-	108.95, 138.98, 152.94, 195.98	130.05	+	
254.02883	C8H16NO3Se+	254.02899	-0.63	0.93, 1.04				x	x						x	x	106.94, 134.97	-	++	
268.04482	C9H18NO3Se+	268.04464	0.67	1.36			x						x			-	133.00	86.06	++	
270.02341	C8H16NO4Se+	270.02391	-1.85	0.73			x					x				n.a.	-	-	+	
271.02005	C7H15N2O4Se+	271.01916	3.28	0.60				x						x		x	-	109.08	++	Virupaksha, Shrift [43]
282.06027	C10H20NO3Se+	282.06029	-0.07	2.34				x							x	x	163.00	81.07	+++	
284.03997	C9H18NO4Se+	284.03956	1.44	0.66, 0.79				x	x						x	x	-	85.06, 102.06, 124.04	+++	
285.05998	C10H21O4Se+	285.05996	0.07	1.63, 1.69		x			x				x			-	132.96, 134.97, 168.98	85.06	++	
301.09063	C11H25O4Se+	301.09126	-2.09	2.95, 3.09		x			x			x				n.a.	n.a.	n.a.	+	
312.03454	C10H18NO5Se+	312.03447	0.22	1.17, 1.36				x	x						x	x	146.97, 164.98, 192.98	102.10	++	
313.02949	C9H17N2O5Se+	313.02972	-0.73	0.76				x						x		x	-	108.05, 154.09	++	
364.95162	C8H17N2O4Se2+	364.95133	0.79	0.70				x						x		x	-	-	+	Hamilton[44]
377.07037	C12H25O8Se+	377.07092	-1.46	1.76, 2.09, 2.34	x				x		x	x				-	215.02	85.06, 103.07	++	
391.08672	C13H27O8Se+	391.08657	0.38	2.33, 2.42	x				x		x	x				-	229.03	85.03, 99.08	++	
405.06603	C13H25O9Se+	405.06583	0.49	0.60		x					x	x				-	243.01	97,03, 145.05	++	
407.04499	C12H23O10Se+	407.04510	-0.27	1.04	x					x	x			x		-	244.99	89.02	+++	Aureli et al.[46]
416.08051	C14H26NO8Se+	416.08182	-3.15	0.85		x					(x)	x				x	134.97, 168.98, 193.97, 326.02	114.05	+	
419.05681	C12H23N2O9Se+	419.05633	1.15	0.59		x				x	x	x				-	167.95, 257.00, 401.04	96.05, 104.11, 196.06, 214.07, 232.08, 258.11	+++	
434.09228	C14H28NO9Se+	434.09238	-0.23	0.76		x					(x)	x				-	134.97, 168.98, 293.98, 312.00, 416.08	130.05	+	
435.04004	C13H23O11Se+	435.04001	0.07	2.01	x						x	x				-	244.99	108.05	+	
444.86739	C8H17N2O4Se3+	444.86785	-1.03	1.03				x						x		x	244.86, 347.84	80.05	+	Németh et al. [31]
446.09233	C15H28NO9Se+	446.09238	-0.11	0.65, 0.72			x		x		(x)				x	x	193.97, 326.02	114.05, 152.06, 172.06, 218.06	++	
447.11604	n.p.	n.p.	n.p.	1.50, 1.54	x				x			x				n.a.	n.a.	n.a.	+	
448.10670	n.p.	n.p.	n.p.	0.76, 1.32	x				x			x				x	326.02	-	+	
460.10880	n.p.	n.p.	n.p.	0.84, 0.90	x				x			x				x	326.02	99.08	+	
463.02100	n.p.	n.p.	n.p.	0.60			x				x	x				-	196.91, 211.92, 300.97, 304.02	160.03	++	
482.99120	n.p.	n.p.	n.p.	1.02	x							x				n.a.	n.a.	n.a.	+	
489.05794	n.p.	n.p.	n.p.	0.72				x				x				x	-	76.02, 116.01, 162.02, 179.05, 233.06, 308.09	+	
523.12784	C18H35O12Se+	523.12883	-1.89	2.08, 2.18	x				x		x	x				-	215.02, 377.07	-	+	
537.14451	C19H37O12Se+	537.14448	0.06	2.33, 2.39		x			x		x	x				-	229.03, 391.08	99.08	+	
552.06109	n.p.	n.p.	n.p.	1.02	x							x				-	284.97	153.06, 177.05	+	
581.10958	C18H33N2O14Se+	581.10915	0.74	0.60		x				x	x	x				-	167.95, 257.00, 296.00, 314.01, 332.03, 401.04, 545.08	110.06, 166.05, 170.08, 214.07, 232.08	+++	
614.14680	n.p.	n.p.	n.p.	0.77	x						x	x				-	434.08, 596.14	116.04, 235.11	++	

Figure 1

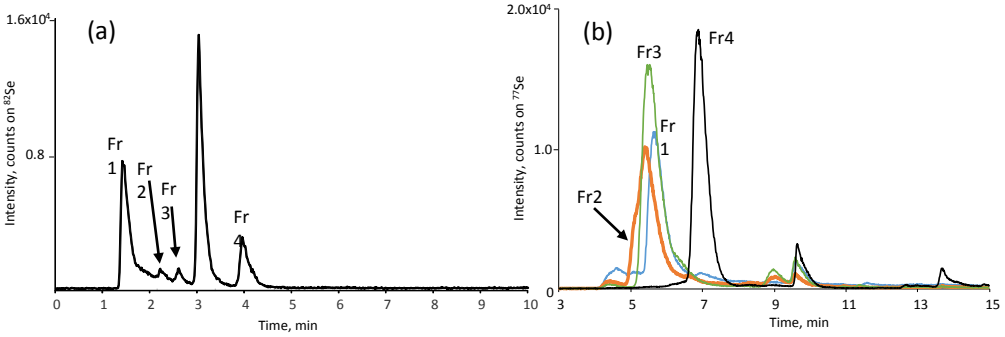

(a)
1.6x10⁴
0.8
Intensity, counts on ⁸²Se
Fr 1
Fr 2
Fr 3
Fr 4
0 1 2 3 4 5 6 7 8 9 10
Time, min
(b)
2.0x10⁴
1.0
0
Intensity, counts on ⁷⁷Se
Fr3
Fr4
Fr 1
Fr2
3 5 7 9 11 13 15
Time, min

Figure 2

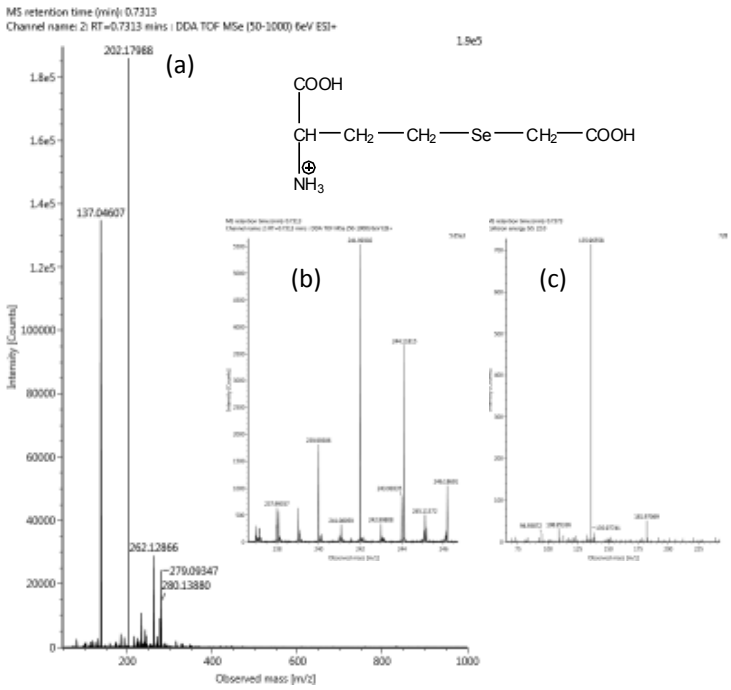

MS retention time (min): 0.7313
Channel name: 2: RT=0.7313 mins : DDA TOF MSe (50-1000) 6eV ESI+
1.9e5
202.17988
(a)
137.04607
262.12866
279.09347
280.13880
Intensity [Counts]
Observed mass [m/z]
COOH
CH
CH2
CH2
Se
CH2
COOH
NH3
(b)
(c)

Figure 3

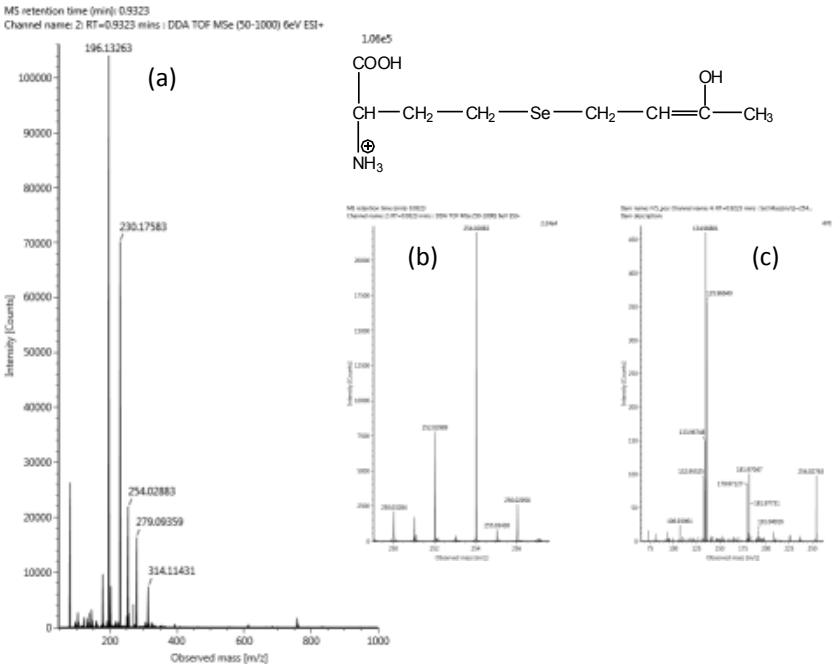

MS retention time (min): 0.9323
Channel name: 2: RT=0.9323 mins : DDA TOF MSe (50-1000) 6eV ESI+
196.13263
230.17583
254.02883
279.09359
314.11431
Intensity [Counts]
Observed mass [m/z]
(a)
(b)
(c)
1.06e5
COOH
OH
CH
CH2
CH2
Se
CH2
CH
C
CH3
NH3

Figure 4

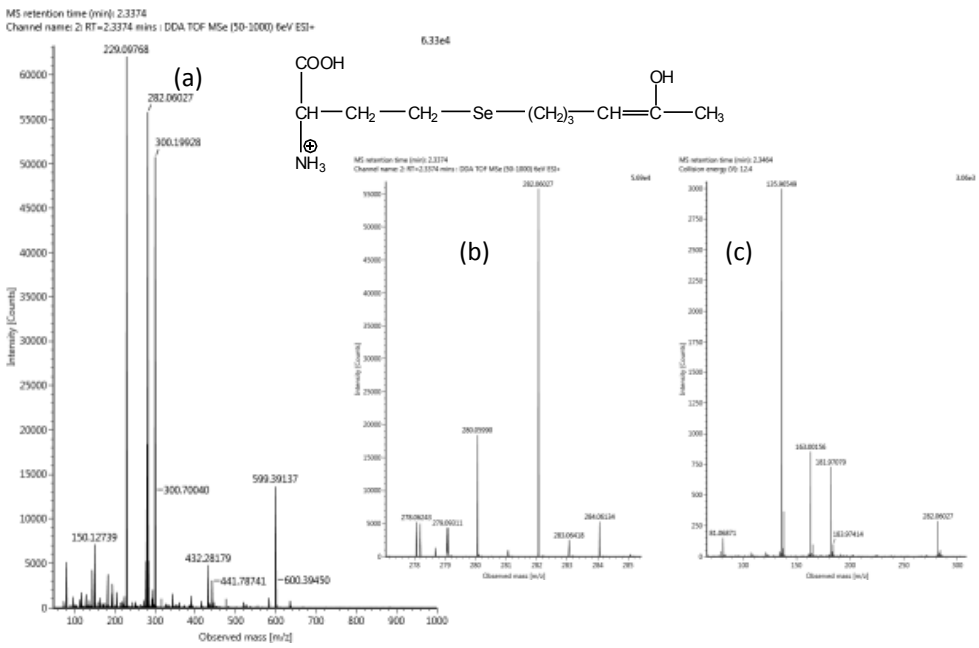

MS retention time (min): 2.3374
Channel name: 2: RT=2.3374 mins : DDA TOF MSe (50-1000) 6eV ESI+
6.33e4
229.09768
282.09027
300.19928
300.70040
599.39137
150.12739
432.28179
441.78741
600.39450
Intensity [Counts]
Observed mass [m/z]
(a)
COOH
CH
CH2
CH2
Se
(CH2)3
CH
C
OH
CH3
NH3
(b)
(c)

Figure 5

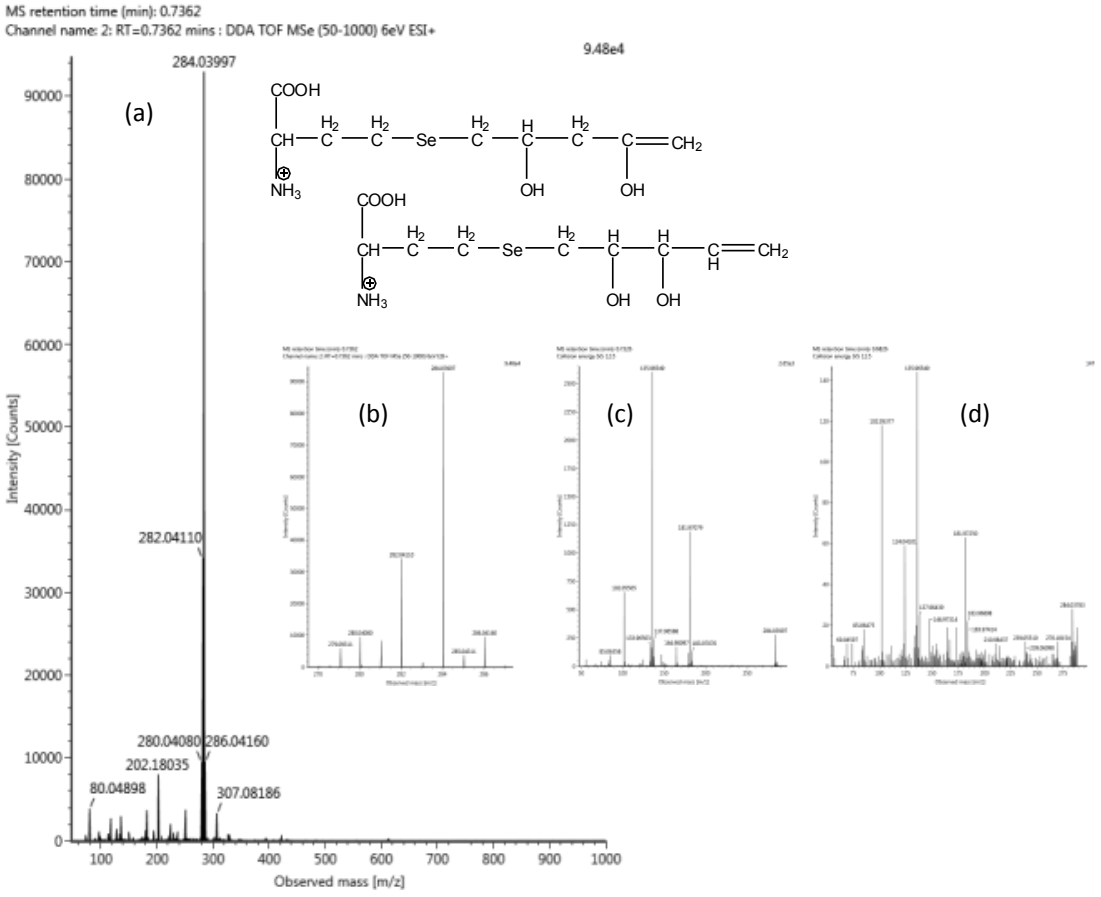

MS retention time (min): 0.7362
Channel name: 2: RT=0.7362 mins : DDA TOF MSe (50-1000) 6eV ESI+
9.48e4
284.03997
(a)
282.04110
280.04080
286.04160
202.18035
80.04898
307.08186
Intensity [Counts]
Observed mass [m/z]
(b)
(c)
(d)

Figure 6

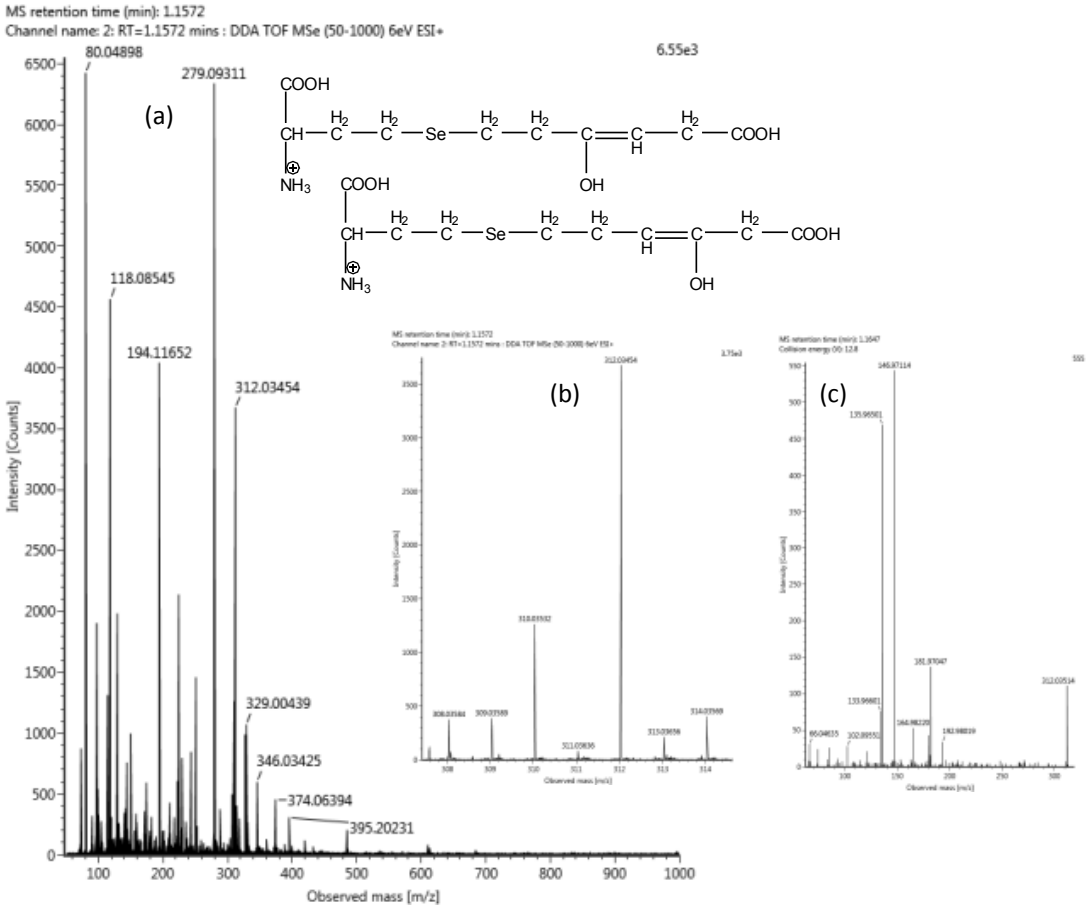

MS retention time (min): 1.1572
Channel name: 2: RT=1.1572 mins : DDA TOF MSe (50-1000) 6eV ESI+
6.55e3
80.04898
279.09311
118.08545
194.11652
312.03454
329.00439
346.03425
374.06394
395.20231
(a)
(b)
(c)
Intensity [Counts]
Observed mass [m/z]
COOH
NH3
OH
Se

Figure 7

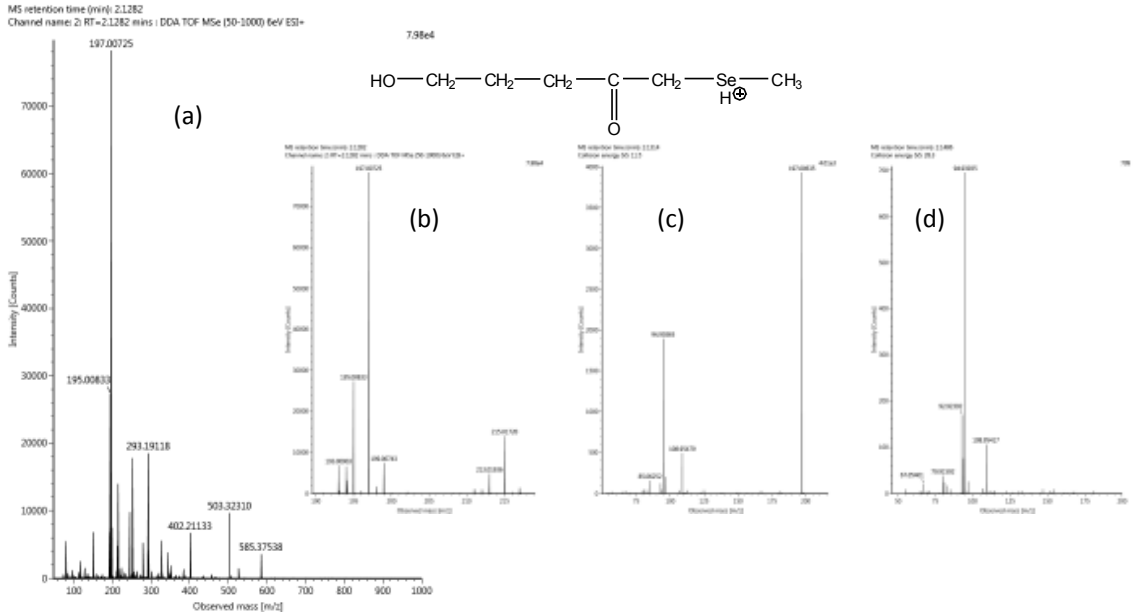

MS retention time (min): 2.1282
Channel name: 2: RT=2.1282 mins : DDA TOF MSe (50-1000) 6eV ESI+
197.00725
195.00833
293.19118
503.32310
402.21133
585.37538
Intensity [Counts]
Observed mass [m/z]
(a)
(b)
(c)
(d)
HO—CH2—CH2—CH2—C—CH2—Se—CH3
O
H⊕

Figure 8

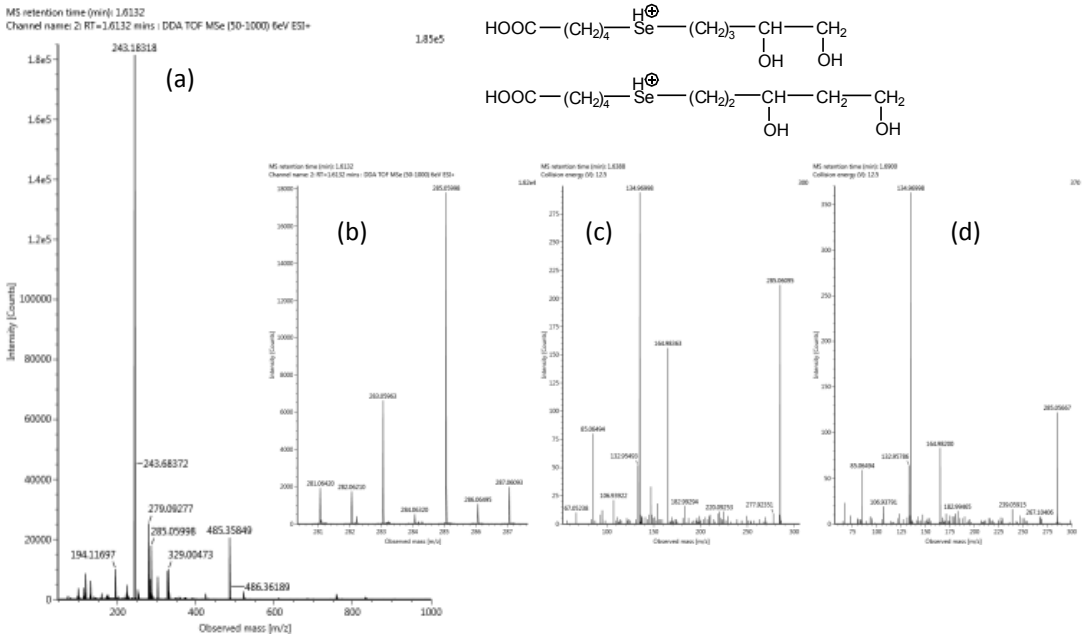

(a)
(b)
(c)
(d)
243.18318
279.09277
285.05998
485.35849
329.00473
194.11697
243.68372
486.36189
Observed mass [m/z]
Intensity [Counts]
HOOC — (CH2)4 — Se — (CH2)3 — CH — CH2
HOOC — (CH2)4 — Se — (CH2)2 — CH — CH2 — CH2
OH

Figure 9

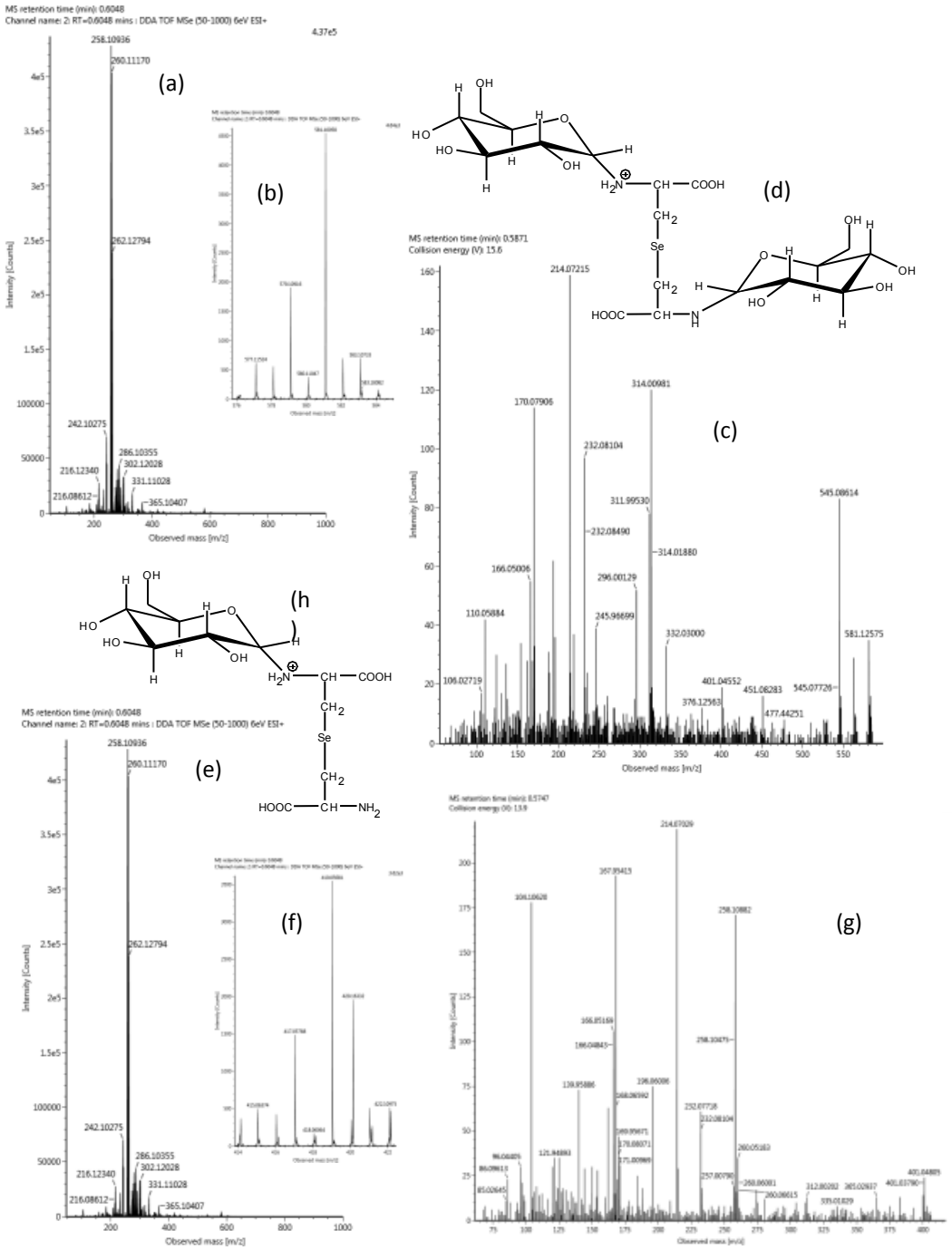

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
MS retention time (min): 0.6048
Channel name: 2: RT=0.6048 mins : DDA TOF MSe (50-1000) 6eV ESI+
Intensity [Counts]
Observed mass [m/z]
258.10936
260.11170
262.12794
242.10275
286.10355
302.12028
331.11028
216.12340
216.08612
365.10407
MS retention time (min): 0.5871
Collision energy (V): 15.6
214.07215
314.00981
170.07906
232.08104
311.99530
232.08490
314.01880
166.05006
296.00129
110.05884
245.96699
332.03000
106.02719
401.04552
451.08283
545.07726
376.12563
477.44251
545.08614
581.12575
MS retention time (min): 0.5747
Collision energy (V): 13.9
214.07929
167.85415
166.10628
258.10882
166.85169
186.04845
139.95886
196.06006
252.07728
212.08104
260.05163
257.80790
401.04889
401.03790

Table S1. UPLC-Unispray-QTOFMS instrumental setup parameters

Acquity I-Class UPLC		**Vion IMS Unispray (+/-) -QTOF-MS**	
UPLC column	Acquity BEH C_{18}, 2.1*100 mm; 1.7 µm	Source temperature	120°C
Eluent "A"	water with 0.1 v/v% formic acid	Desolvation temperature	550°C
Eluent "B"	acetonitrile with 0.1 v/v% formic acid	Capillary voltage	300 V
Flow rate	0.4 ml/min	Desolvation gas	1000 L/h
Column temperature	25°C	Cone gas	100 L/h
Gradient	0 – 1.0 min 10% „B" 1.0 – 4.0 min ↑ 80% „B" 4.0 – 4.5 min 80% „B" 4.5 – 5.0 min ↓ 10% „B" 5.0 – 7.0 min 10% „B"	IMS MS scan MS scan time Lock mass MS/MS scan	OFF 100 – 1000 m/z 0.2 s ON 50 – 1000 m/z
Injection volume	3.0 µl	Low mass ramp	20 – 30 eV
Sample temperature	8°C	High mass ramp	30 – 80 eV

Table S2. Effect of defect padding settings on the efficiency of automatic selenium pattern recognition.

Fraction	m/z	Minimally required mass padding (Da) for successful detection			
		at 20 mDa defect padding	at 40 mDa defect padding	at 60 mDa defect padding	at 80 mDa defect padding
#1	282	>1000	17	11	10
	284	89	12	11	11
	407	>1000	121	106	71
	242	27	21	21	21
	391	>1000	>1000	191	116
	391	>1000	>1000	190	116
#2	419	>1000	130	115	86
	581	>1000	>1000	>1000	437
	419	>1000	173	125	115
	401	>1000	120	103	67
	285	>1000	21	13	12
	405	106	98	85	41
	285	>1000	21	13	12
#3	441	391	138	128	100
	419	>1000	175	125	115
	446	>1000	>1000	457	180
	268	>1000	2	2	2
	197	53	52	52	52
	215	40	40	40	39
#4	254	15	12	12	12
	312	53	33	31	31
	271	1	1	1	1
	282	>1000	17	11	10
	284	89	12	11	11

Table S3. Efficiency of automatic selenium pattern recognition. It is to note that the compounds with detected theoretical selenium mass defect include false positive hits and the different isotopologues of the same compound as well.

Fraction	Number of compounds detected	Number of compounds with detected theoretical selenium mass defect	Relative amount of compounds detected with theoretical selenium mass defect	Number of compounds with detected theoretical selenium mass defect	Relative amount of compounds detected with theoretical selenium mass defect
		Settings: defect paddig, 60 mDa; mass padding, 200 Da; minimum intensity, 2000 cps		Settings: defect paddig, 80 mDa; mass padding, 200 Da; minimum intensity, 2000 cps	
#1	7293	350	4.8%	457	6.3%
#2	5062	227	4.5%	277	5.5%
#3	4838	229	4.7%	269	5.6%
#4	4572	178	3.9%	206	4.5%

Table S4. Efficiency of automatic selenium pattern recognition at optimised settings (mass padding: 200 Da; defect padding: 80 mDa). Values in bold indicate successful detection.

FR1		FR2		FR3	FR4	
m/z	background of unsuccessful automatic detection	m/z	background of unsuccessful automatic detection	m/z	m/z	background of unsuccessful automatic detection
241.99302	-	**285.05998**	-	**197.00725**	**254.02883**	-
377.07037	isobaric interference	301.09063	isobaric interference	**241.99302**	**271.02005**	-
391.08672	isobaric interference	405.06603	low intensity	**268.04482**	**282.06027**	-
407.04499	-	416.08220	low intensity	**270.02341**	**284.03997**	-
435.04004	low intensity	**419.05681**	-	**446.09233**	**312.03454**	-
447.11604	low intensity	434.09228	low intensity	**463.02100**	313.02949	low intensity
448.10670	low intensity	537.14451	low intensity		**364.95162**	-
460.10880	low intensity	581.10958	low mass padding		444.86739	low intensity + triple selenium
482.99120	low intensity				489.05794	low intensity
523.12784	low intensity					
552.06109	low intensity					
614.14680	low intensity					

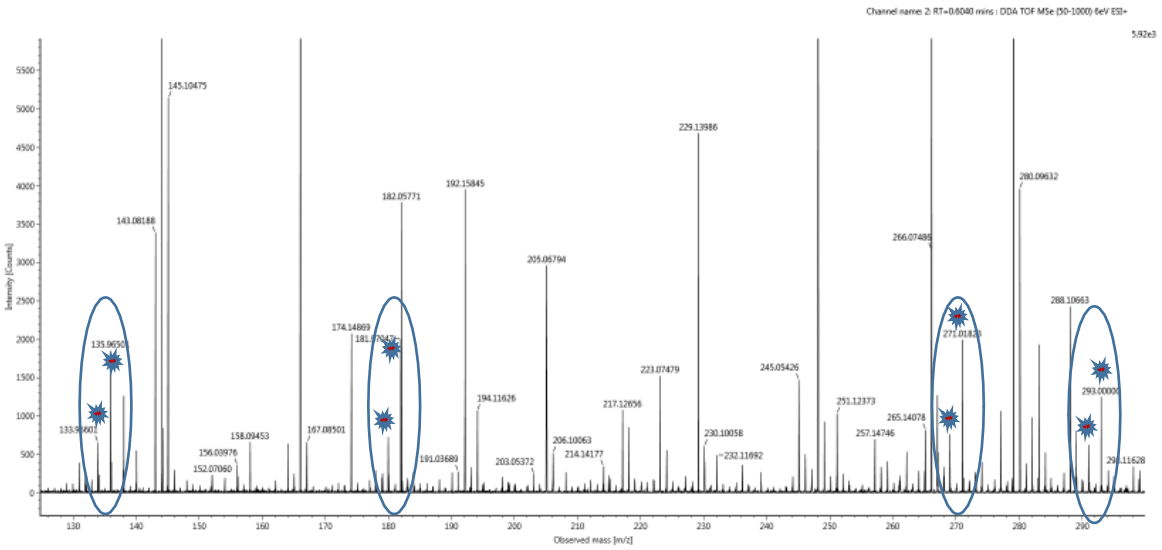


Fig. S1: Full scan spectrum from Fr4, showing the selenocystathionine related selenium patterns (in-source fragments, parent molecule and sodium adduct). Stars indicate the ^{78}Se-^{80}Se isotopologues.

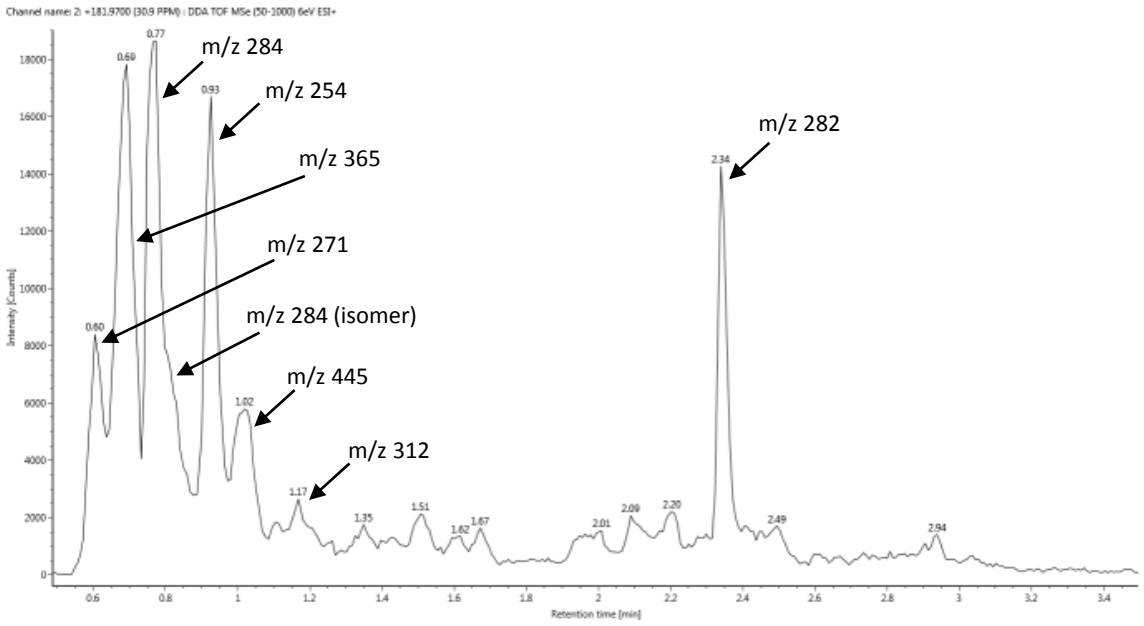
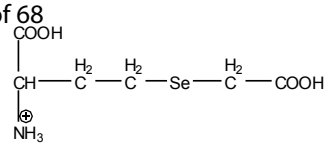
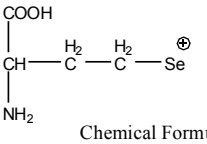


Fig. S2: Extracted ion chromatogram of selenohomocysteine ion-source fragment (m/z 181.97). Arrows and values indicate the corresponding and detected selenium species (see Table 1 for further information).

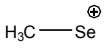


Chemical Formula: $C_6H_{12}NO_4Se^+$
Exact Mass: 241.9926

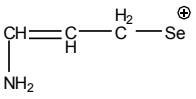
Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
181.97069	C4H8NO2Se+	181.97150	-4.45
135.96558	C3H6NSe+	135.96600	-3.09
108.95336	C2H5Se+	108.95510	-15.97
94.93872	CH3Se+	94.93940	-7.16



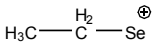
Chemical Formula: $C_4H_8NO_2Se^+$
Exact Mass: 181.9715



Chemical Formula: CH_3Se^+
Exact Mass: 94.9394

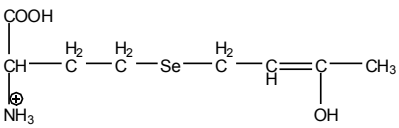


Chemical Formula: $C_3H_6NSe^+$
Exact Mass: 135.9660



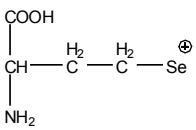
Chemical Formula: $C_2H_5Se^+$
Exact Mass: 108.9551

Fig. S3: Proposed structures of the m/z 241.9926 compound and its MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 2).

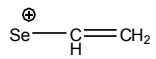


Chemical Formula: $C_8H_{16}NO_3Se^+$
Exact Mass: 254.0290

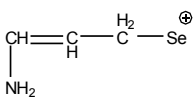
Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
181.97047	C4H8NO2Se+	181.9715	-5.66
135.96649	C3H6NSe+	135.966	3.60
134.96881	C4H7Se+	134.97070	-14.00
132.95525	C4H5Se+	132.95510	1.13
106.93961	C2H3Se+	106.93940	1.96



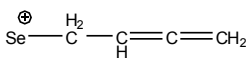
Chemical Formula: $C_4H_8NO_2Se^+$
Exact Mass: 181.9715



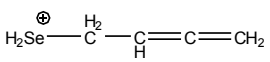
Chemical Formula: $C_2H_3Se^+$
Exact Mass: 106.9394



Chemical Formula: $C_3H_6NSe^+$
Exact Mass: 135.9660

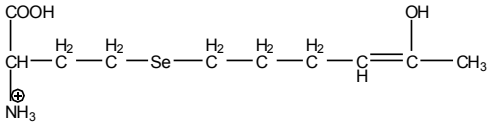


Chemical Formula: $C_4H_5Se^+$
Exact Mass: 132.9551



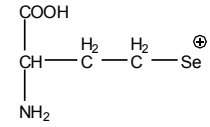
Chemical Formula: $C_4H_7Se^+$
Exact Mass: 134.9707

Fig. S4: Proposed structures of the m/z 254.0289 compound and its MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 3).

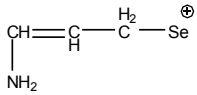


Chemical Formula: $C_{10}H_{20}NO_3Se^+$
Exact Mass: 282.0603

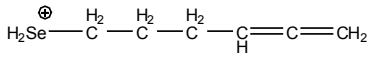
Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
181.97079	C4H8NO2Se+	181.97150	-3.90
163.00156	C6H11Se+	163.00200	-2.70
135.96549	C3H6NSe+	135.96600	-3.75



Chemical Formula: $C_4H_8NO_2Se^+$
Exact Mass: 181.9715

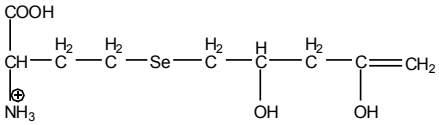


Chemical Formula: $C_3H_6NSe^+$
Exact Mass: 135.9660

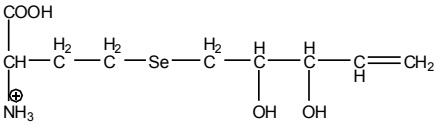


Chemical Formula: $C_6H_{11}Se^+$
Exact Mass: 163.0020

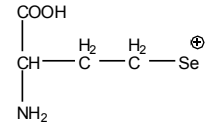
Fig. S5: Proposed structures of the m/z 282.0603 compound and its MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 4).



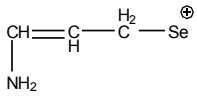
Chemical Formula: $C_9H_{18}NO_4Se^+$
Exact Mass: 284.0396



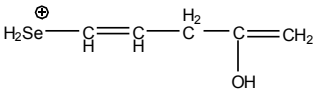
Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
181.97079	C4H8NO2Se+	181.97150	-3.90
164.98097	C5H9OSe+	164.98130	-2.00
146.97314	C5H7Se+	146.97070	16.60
135.96549	C3H6NSe+	135.96600	-3.75
102.05505	C4H8NO2+	102.05500	0.49



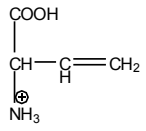
Chemical Formula: $C_4H_8NO_2Se^+$
Exact Mass: 181.9715



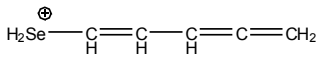
Chemical Formula: $C_3H_6NSe^+$
Exact Mass: 135.9660



Chemical Formula: $C_5H_9OSe^+$
Exact Mass: 164.9813

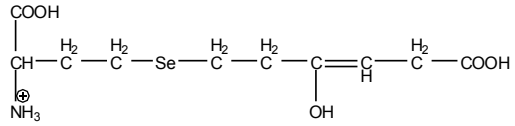


Chemical Formula: $C_4H_8NO_2^+$
Exact Mass: 102.0550



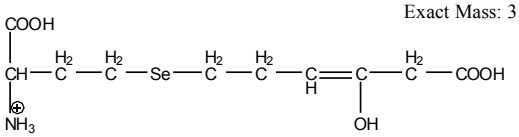
Chemical Formula: $C_5H_7Se^+$
Exact Mass: 146.9707

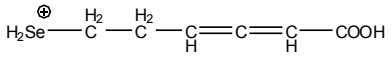
Fig. S6: Proposed structures of the m/z 284.0396 compounds and their MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 5).



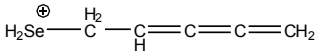
Chemical Formula: $C_{10}H_{18}NO_5Se^+$
Exact Mass: 312.0345

Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
192.98019	C6H9O2Se+	192.97620	20.68
181.97047	C4H8NO2Se+	181.97150	-5.66
164.98220	C5H9OSe+	164.98130	5.46
146.97114	C5H7Se+	146.97070	2.99
135.96501	C3H6NSe+	135.96600	-7.28

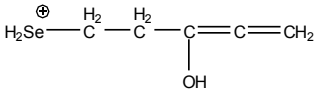




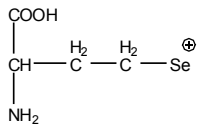
Chemical Formula: $C_6H_9O_2Se^+$
Exact Mass: 192.9762



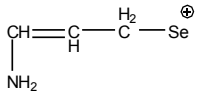
Chemical Formula: $C_5H_7Se^+$
Exact Mass: 146.9707



Chemical Formula: $C_5H_9OSe^+$
Exact Mass: 164.9813

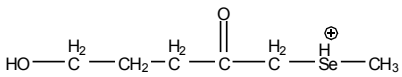


Chemical Formula: $C_4H_8NO_2Se^+$
Exact Mass: 181.9715



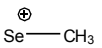
Chemical Formula: $C_3H_6NSe^+$
Exact Mass: 135.9660

Fig. S7: Proposed structures of the m/z 312.0345 compounds and their MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 6).

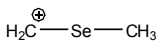


Chemical Formula: $C_6H_{13}O_2Se^+$
Exact Mass: 197.0075

Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
108.95479	C2H5Se+	108.9551	-2.85
94.93865	CH3Se+	94.9394	-7.90

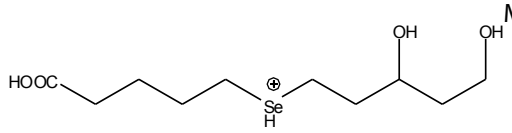


Chemical Formula: CH_3Se^+
Exact Mass: 94.9394

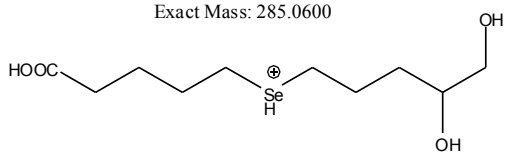


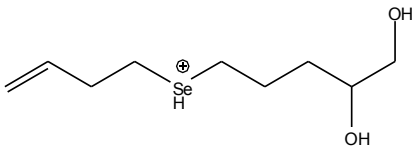
Chemical Formula: $C_2H_5Se^+$
Exact Mass: 108.9551

Fig. S8: Proposed structures of the m/z 197.0075 compound and its MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 7).



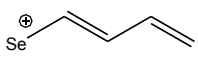
Chemical Formula: $C_{10}H_{21}O_4Se^+$
Exact Mass: 285.0600



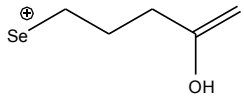


Chemical Formula: $C_9H_{19}O_2Se^+$
Exact Mass: 239.0545

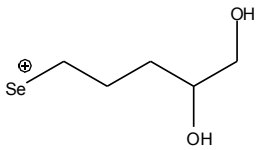
Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
239.05915	C9H19O2Se+	239.05450	19.45
182.99294	C5H11O2Se+	182.99190	5.68
164.98200	C5H9OSe+	164.98130	4.24
134.96998	C4H7Se+	134.97070	-5.33
132.95493	C4H5Se+	132.95510	-1.28
106.93922	C2H3Se+	106.93940	-1.68
85.06494	C5H9O+	85.06480	1.65



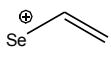
Chemical Formula: $C_4H_5Se^+$
Exact Mass: 132.9551



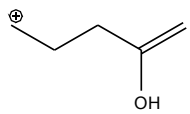
Chemical Formula: $C_5H_9OSe^+$
Exact Mass: 164.9813



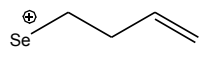
Chemical Formula: $C_5H_{11}O_2Se^+$
Exact Mass: 182.9919



Chemical Formula: $C_2H_3Se^+$
Exact Mass: 106.9394



Chemical Formula: $C_5H_9O^+$
Exact Mass: 85.0648



Chemical Formula: $C_4H_7Se^+$
Exact Mass: 134.9707

Fig. S9: Proposed structures of the m/z 285.0600 compounds and their MS/MS fragments, together with mass accuracy data of the fragments (see Fig. 8).

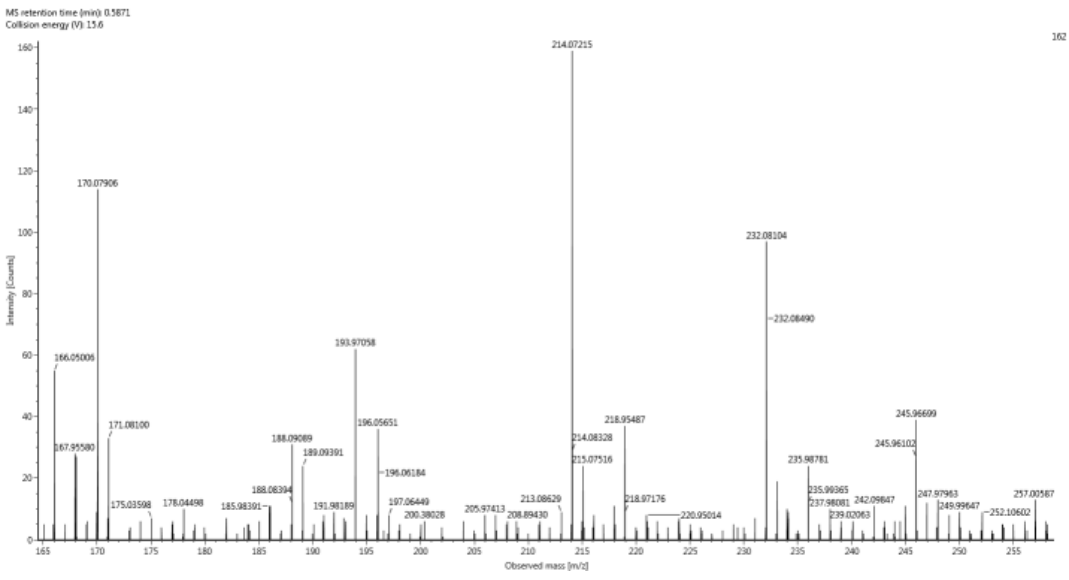
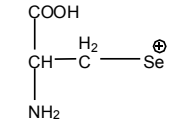
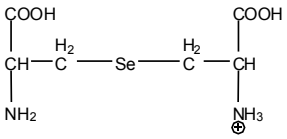


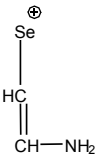
Fig. S10: Additional (zoomed) MS/MS spectrum of the m/z 581.1092 compound (see Fig. 9 and S11).



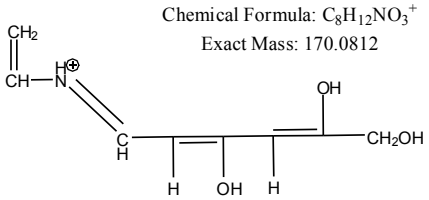
Chemical Formula: $C_3H_6NO_2Se^+$
Exact Mass: 167.9558



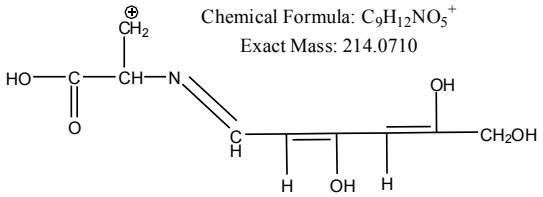
Chemical Formula: $C_6H_{13}N_2O_4Se^+$
Exact Mass: 257.0035



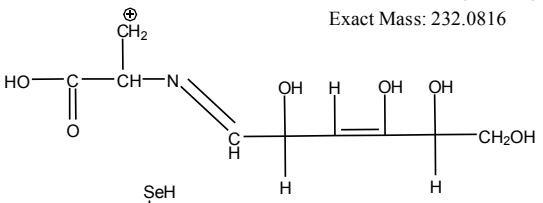
Chemical Formula: $C_2H_4NSe^+$
Exact Mass: 121.9503



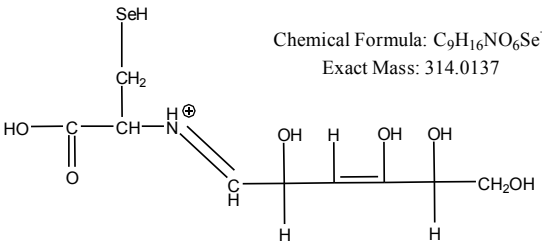
Chemical Formula: $C_8H_{12}NO_3^+$
Exact Mass: 170.0812



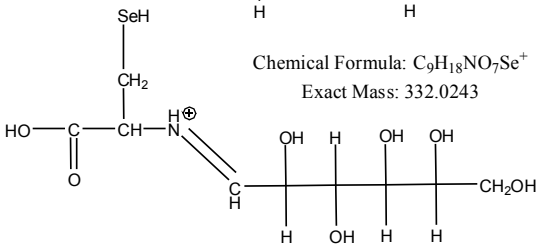
Chemical Formula: $C_9H_{12}NO_5^+$
Exact Mass: 214.0710



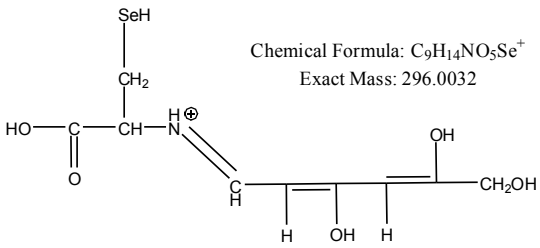
Chemical Formula: $C_9H_{14}NO_6^+$
Exact Mass: 232.0816



Chemical Formula: $C_9H_{16}NO_6Se^+$
Exact Mass: 314.0137

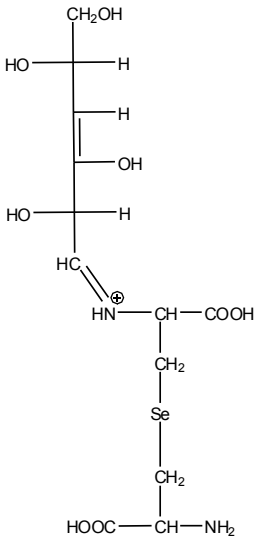


Chemical Formula: $C_9H_{18}NO_7Se^+$
Exact Mass: 332.0243

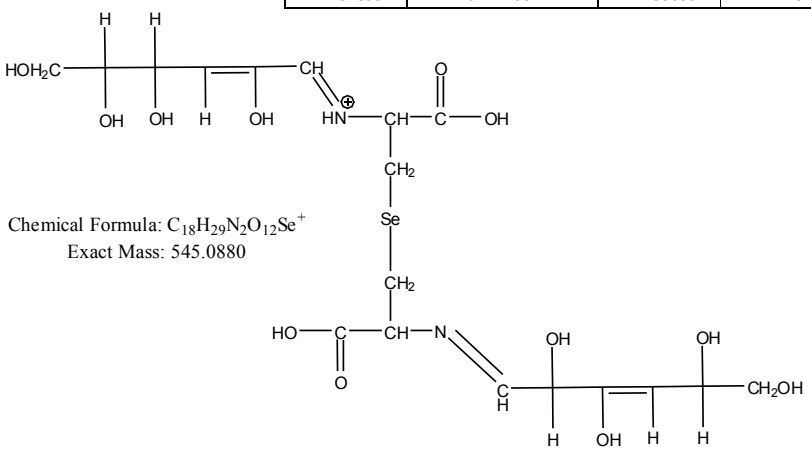


Chemical Formula: $C_9H_{14}NO_5Se^+$
Exact Mass: 296.0032

Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
545.08614	C18H29N2O12Se+	545.08800	-3.41
401.04552	C12H21N2O8Se+	401.04580	-0.70
332.03000	C9H18NO7Se+	332.02430	17.17
314.00981	C9H16NO6Se+	314.01370	-12.39
296.00129	C9H14NO5Se+	296.00320	-6.45
257.0079	C6H13N2O4Se+	257.00350	17.12
232.08104	C9H14NO6+	232.08160	-2.41
214.07215	C9H12NO5+	214.07100	5.37
170.07906	C8H12NO3+	170.08120	-12.58
167.95415	C3H6NO2Se+	167.95580	-9.82
121.94893	C2H4NSe+	121.95030	-11.23

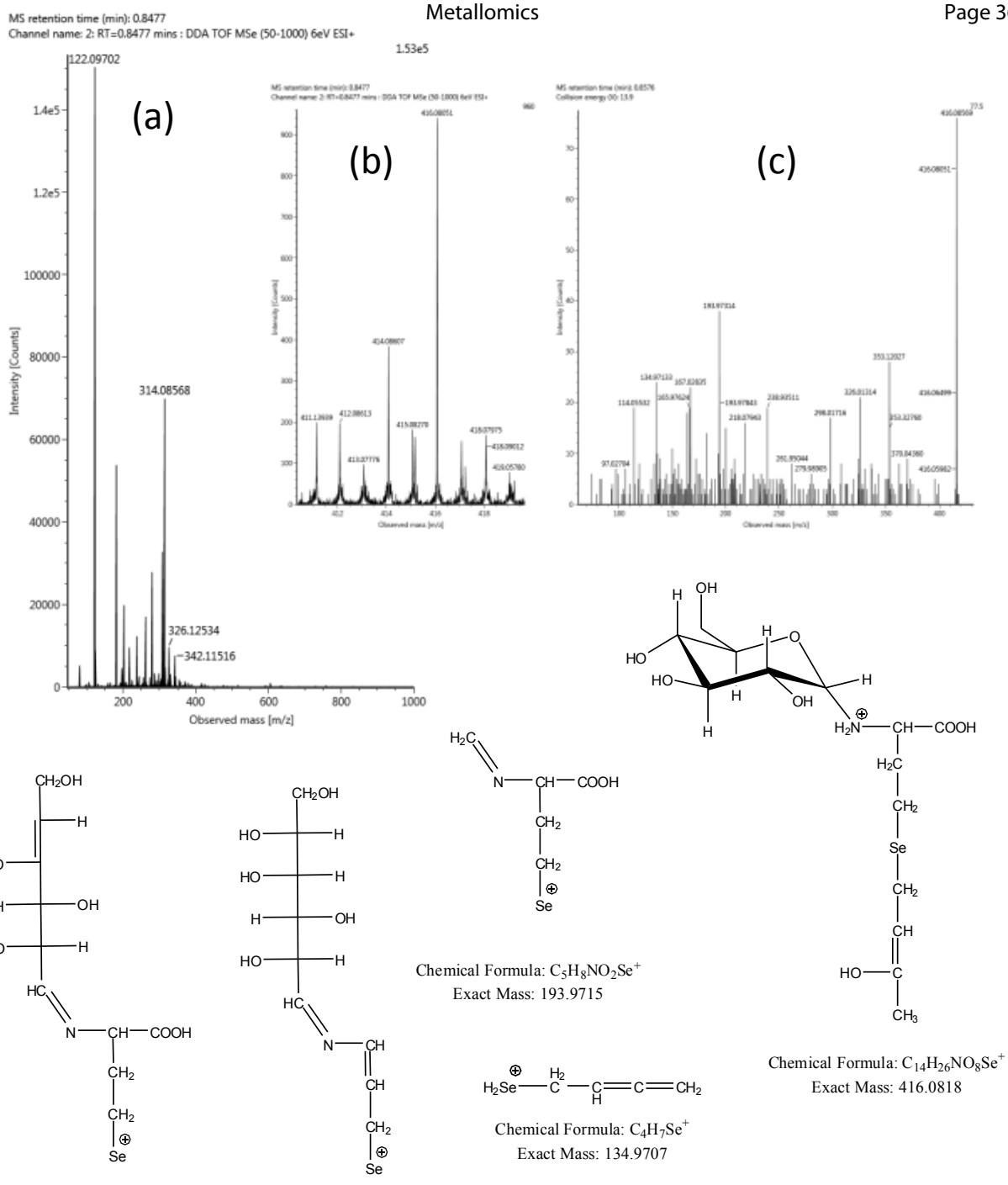


Chemical Formula: $C_{12}H_{21}N_2O_8Se^+$
Exact Mass: 401.0458



Chemical Formula: $C_{18}H_{29}N_2O_{12}Se^+$
Exact Mass: 545.0880

Fig. S11: Proposed structures of the MS/MS fragments of the m/z 419.0568 and 581.1092 compounds, together with mass accuracy data of the fragments (see Fig. 9).



Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
326.01314	C10H16NO6Se+	326.01370	-1.72
298.01716	C9H16NO5Se+	298.01880	-5.50
193.97314	C5H8NO2Se+	193.97150	8.45
181.97515	C4H8NO2Se+	181.97150	20.06
135.96461	C3H6NSe+	135.96600	-10.2
134.97133	C4H7Se+	134.97070	4.67

Fig. S12: Compound at the experimental m/z 416.08051 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum, together with the proposed structures and the mass accuracy data of the MS/MS fragments.

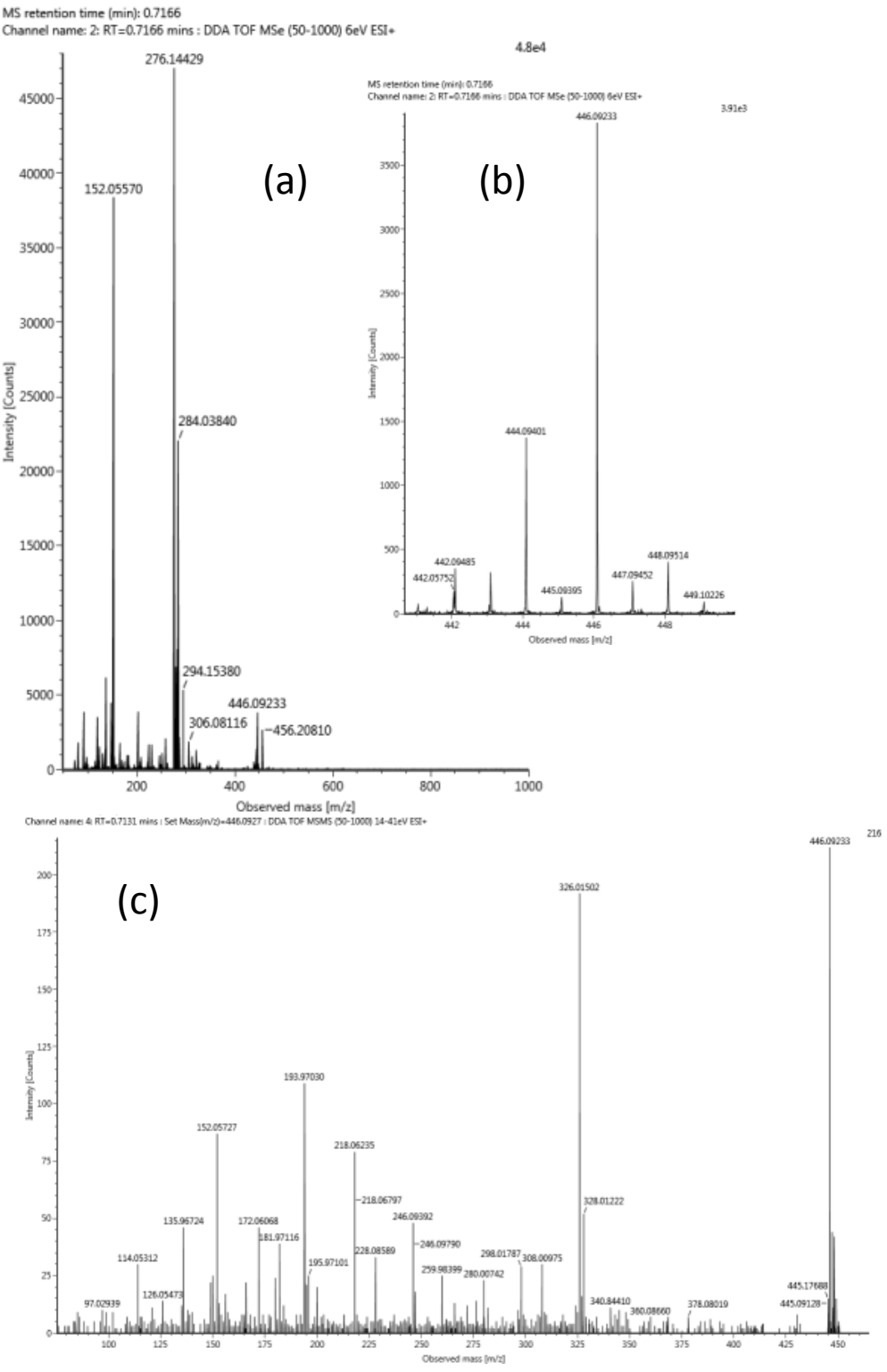
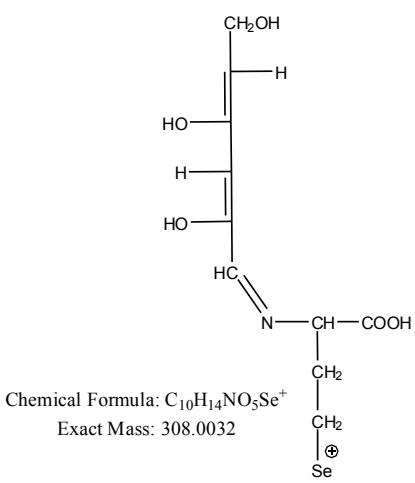
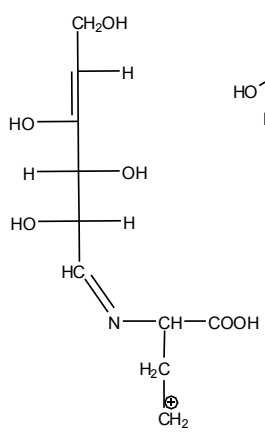
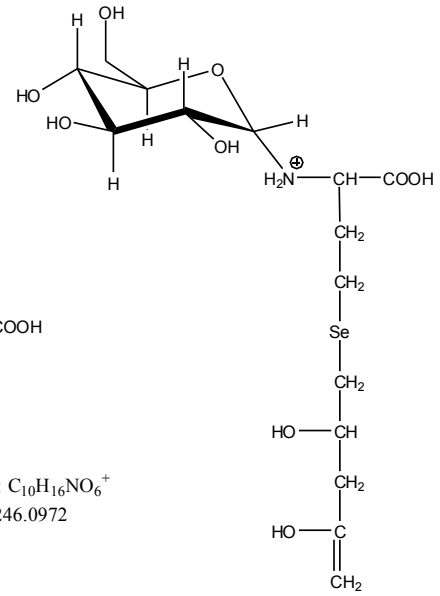
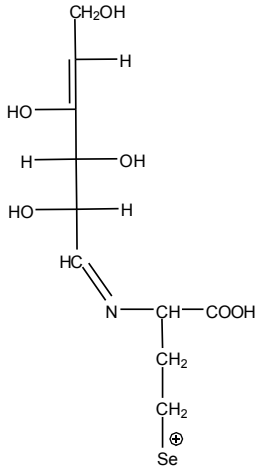


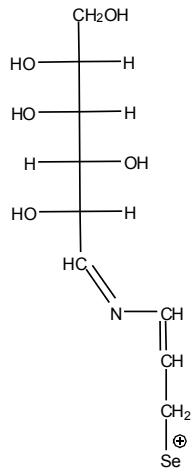
Fig. S13: Compound at the experimental m/z 446.09233 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum, together with the proposed structures and the mass accuracy data of the MS/MS fragments.

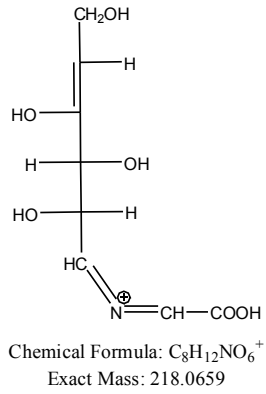


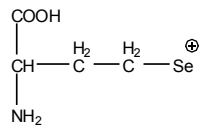


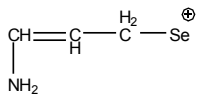












Experimental m/z	Elemental composition, [M+H]+	Theoretical m/z	Difference, ppm
326.01314	C10H16NO6Se+	326.01370	-1.72
308.00975	C10H14NO5Se+	308.00320	21.27
298.01787	C9H16NO5Se+	298.01880	-3.12
246.09392	C10H16NO6+	246.09720	-13.33
218.06235	C8H12NO6+	218.06590	-16.28
181.97116	C4H8NO2Se+	181.97150	-1.87
135.96724	C3H6NSe+	135.96600	9.12

Fig. S14: Compound at the experimental m/z 446.09233 (for details, see Table 1); proposed structures and the mass accuracy data of the MS/MS fragments.

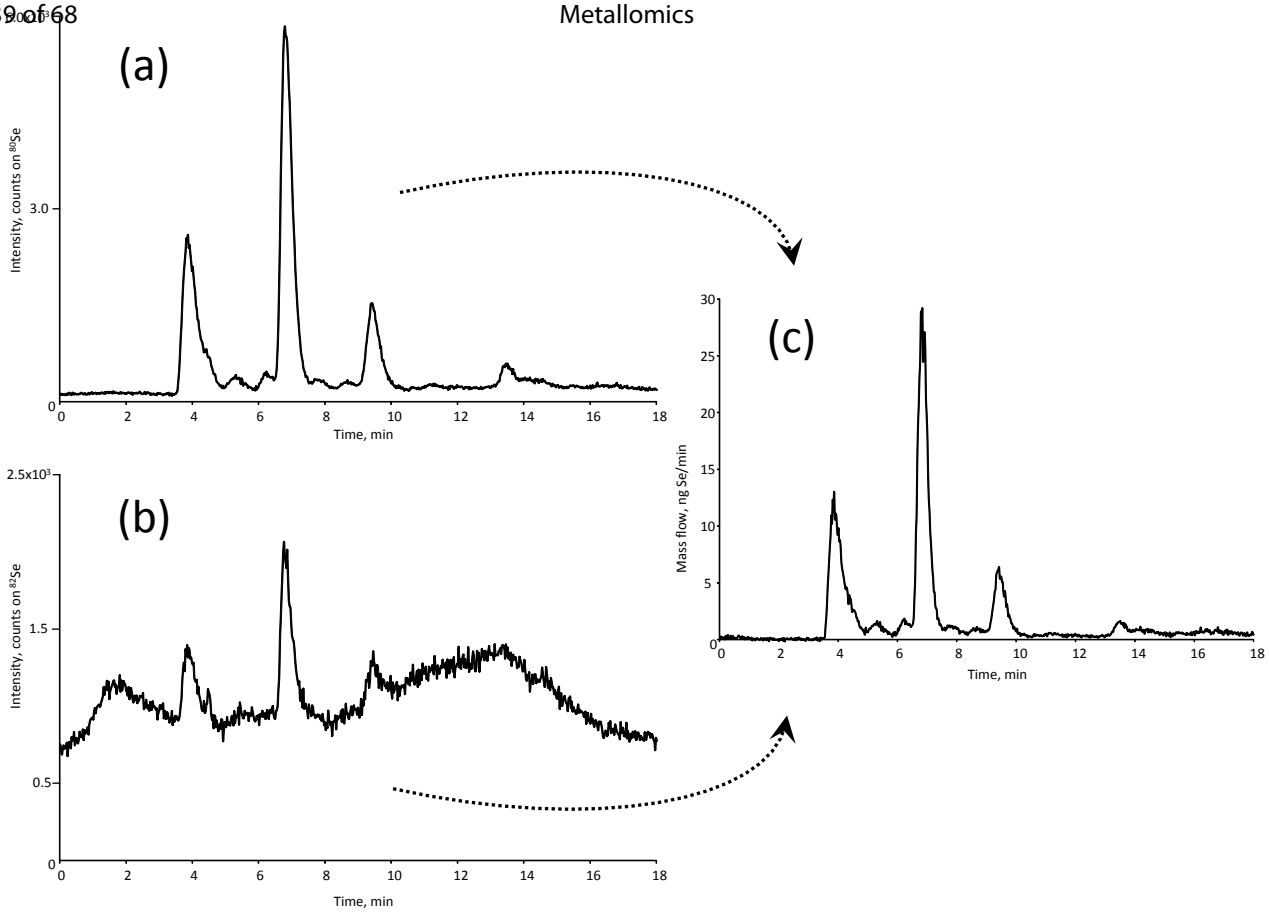


Fig. S15: IP-RP-IDA-ICP-MS based quantification of selenolanthionine (eluting at 6.8 min) extracted from C. violifolia leaves. (a) LC-ICP-MS chromatogram recorded on ^{80}Se; (b) LC-ICP-MS chromatogram recorded on ^{82}Se used for isotope dilution; (c) selenium mass flow calculated from the $^{80}Se/^{82}Se$ data.

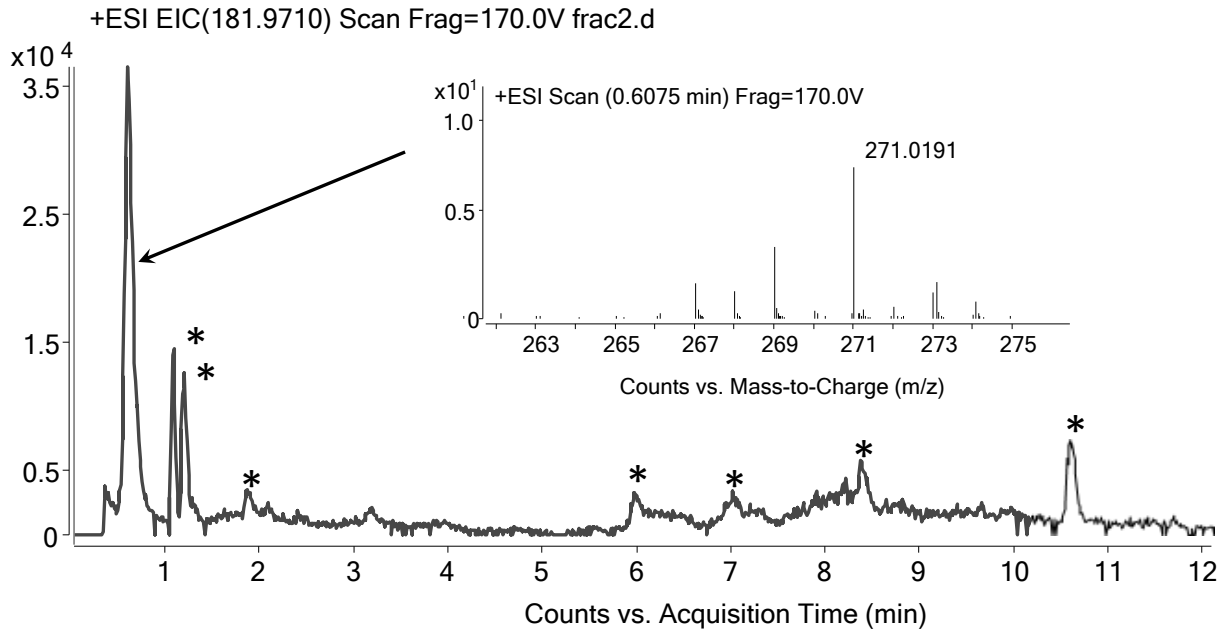
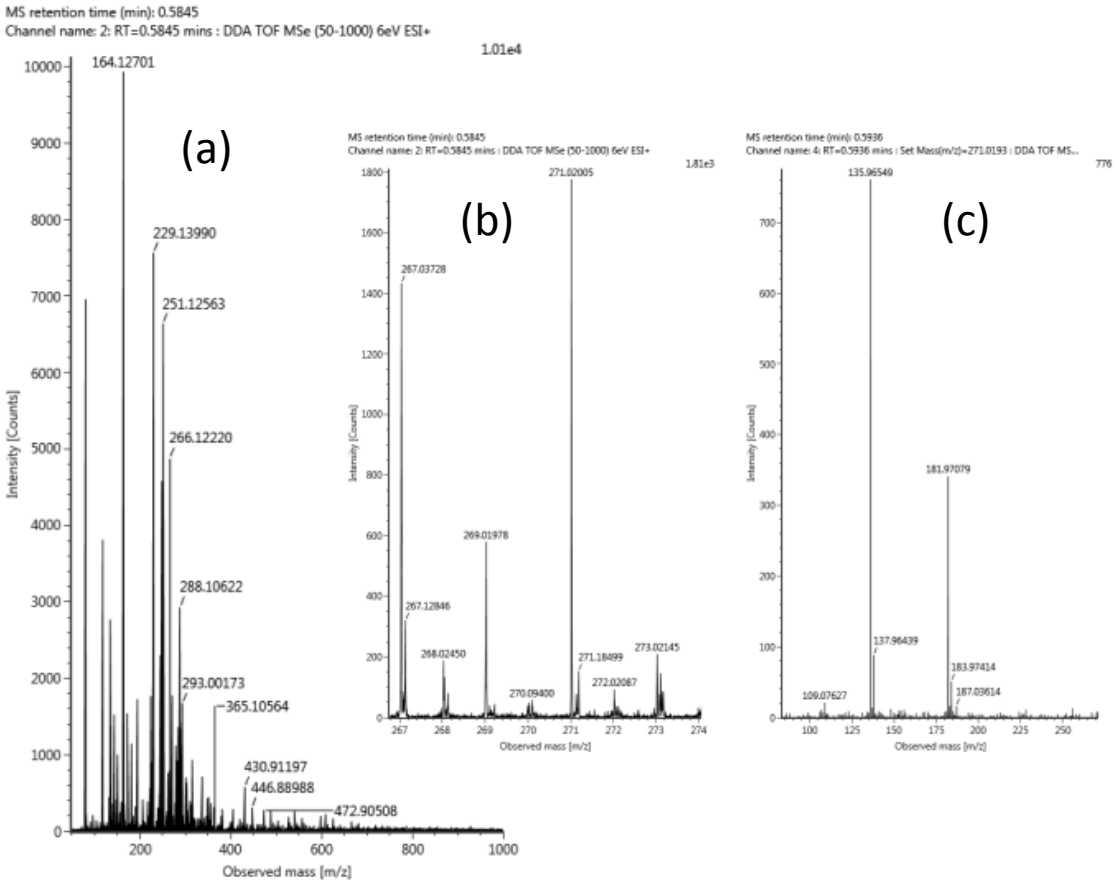
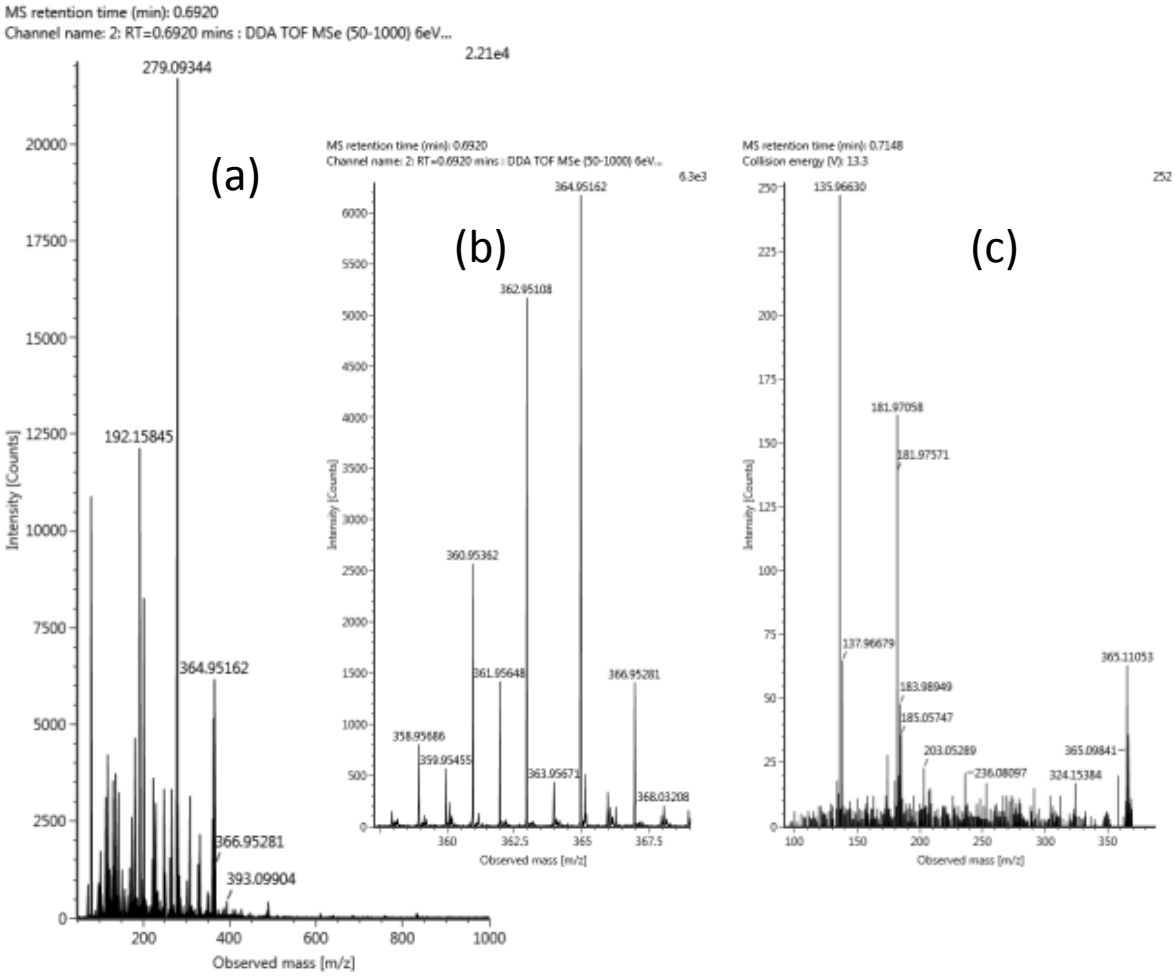


Fig. S16: In-source selenohomocysteine fragment (m/z 181.971) of selenocystathionine from Fraction #4, together with several other minor appearance of this fragment (indicated with '*'). Data were recorded on an Agilent 6530 ESI-QTOFMS instrument.

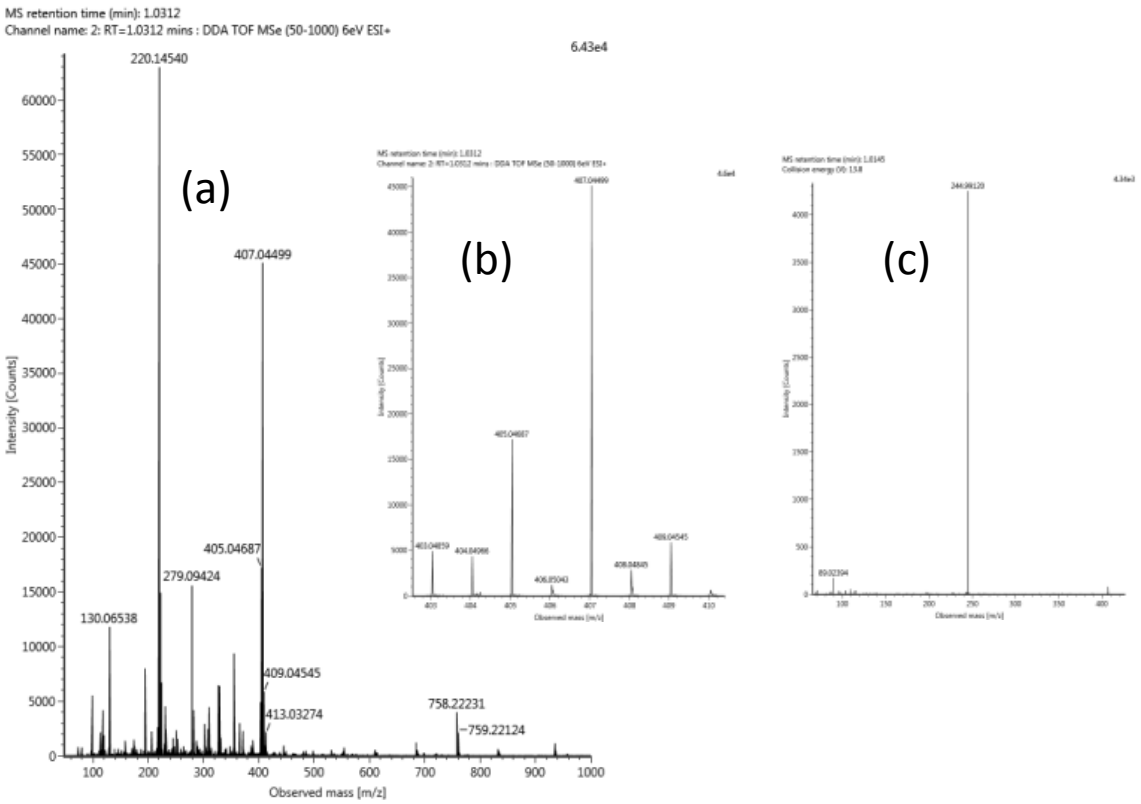
Known selenium species (other than selenolanthionine) detected in the water soluble selenometabolome of C. violifolia



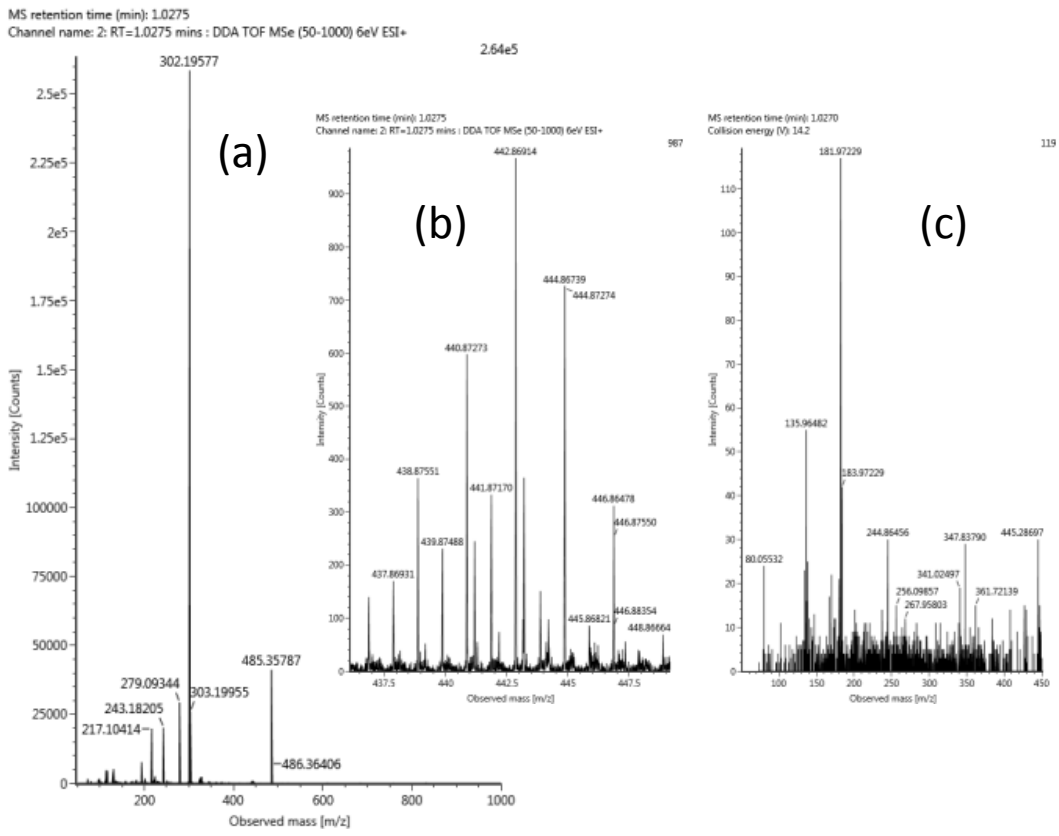
Selenocystathionine (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



Selenohomocystine (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.

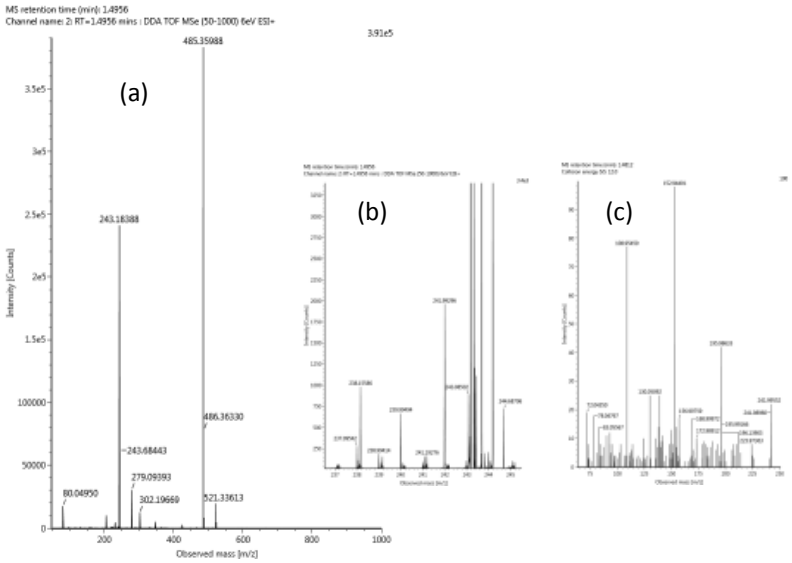


Selenosugar at m/z 407 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.

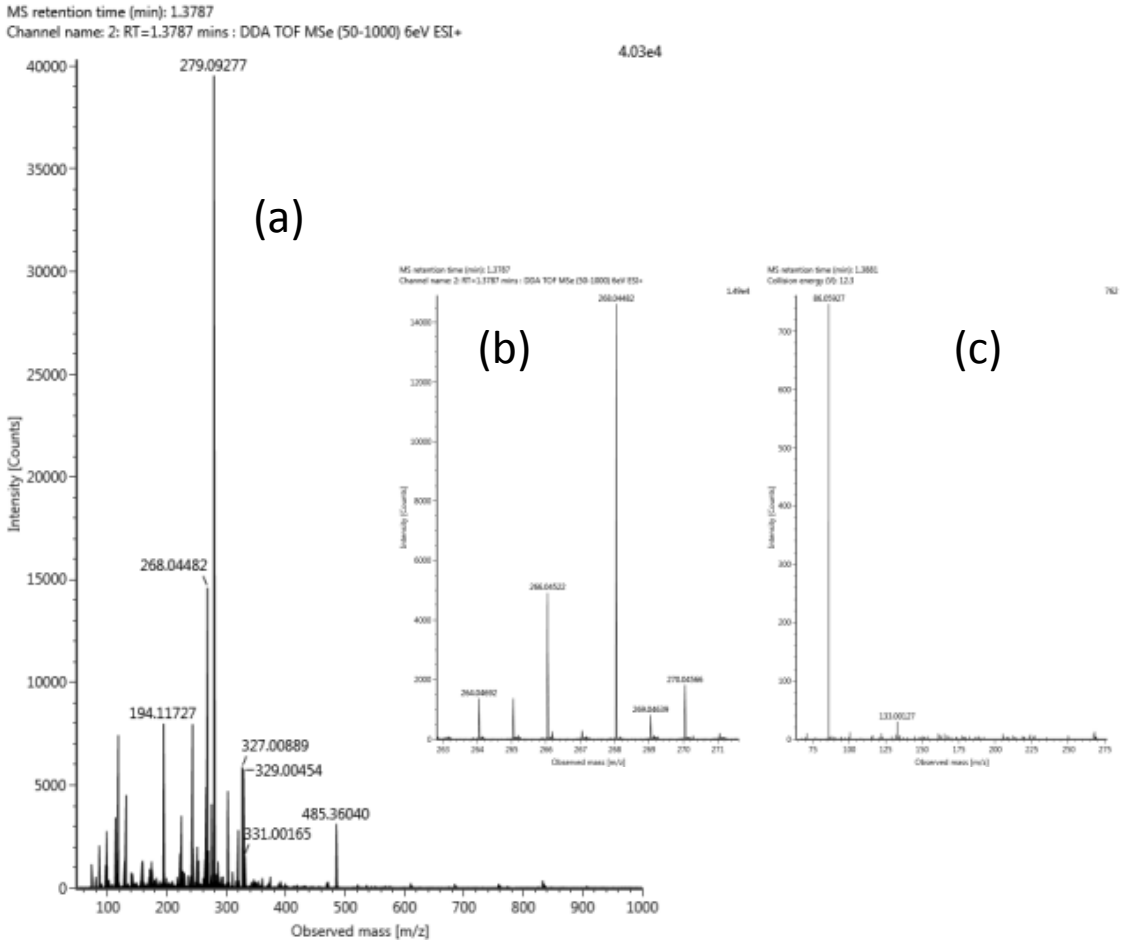


Se-selenohomocysteinyl-diseleno-homocysteine (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.

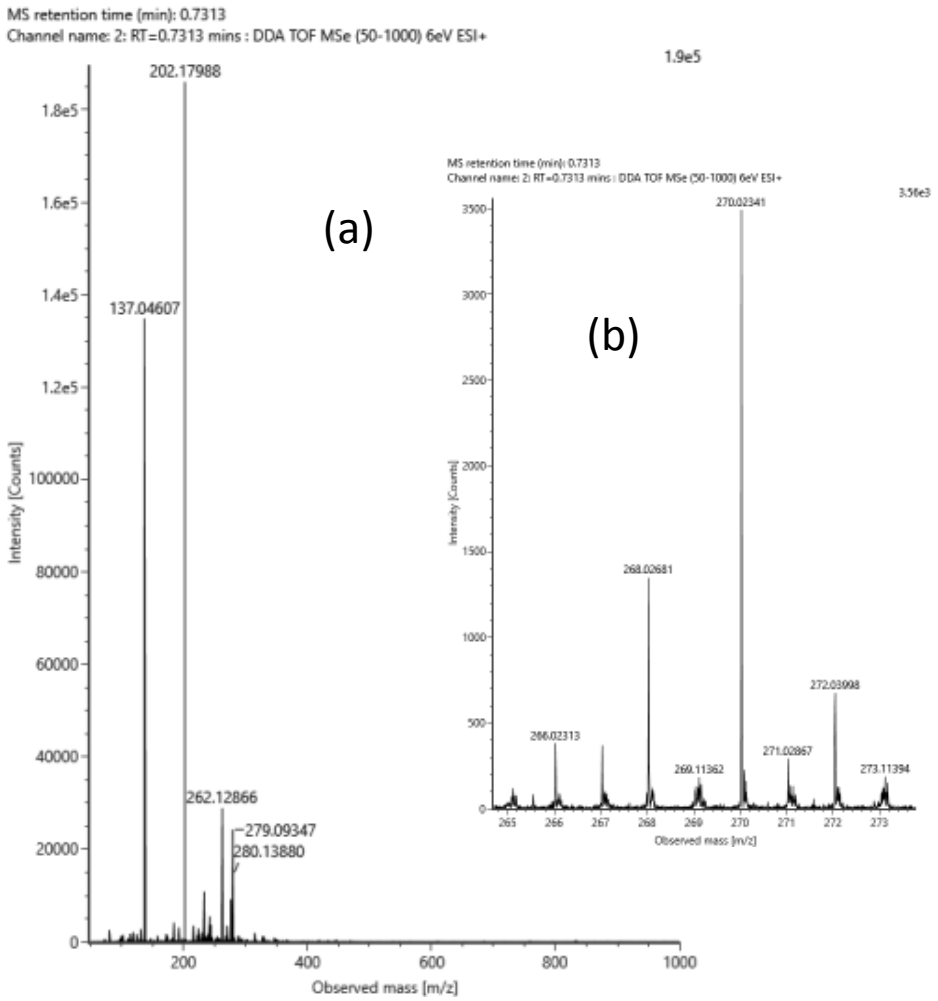
Unknown selenium species detected in the water soluble selenometabolome of C. violifolia without structure assignment (in the order of m/z values; see Table 1)



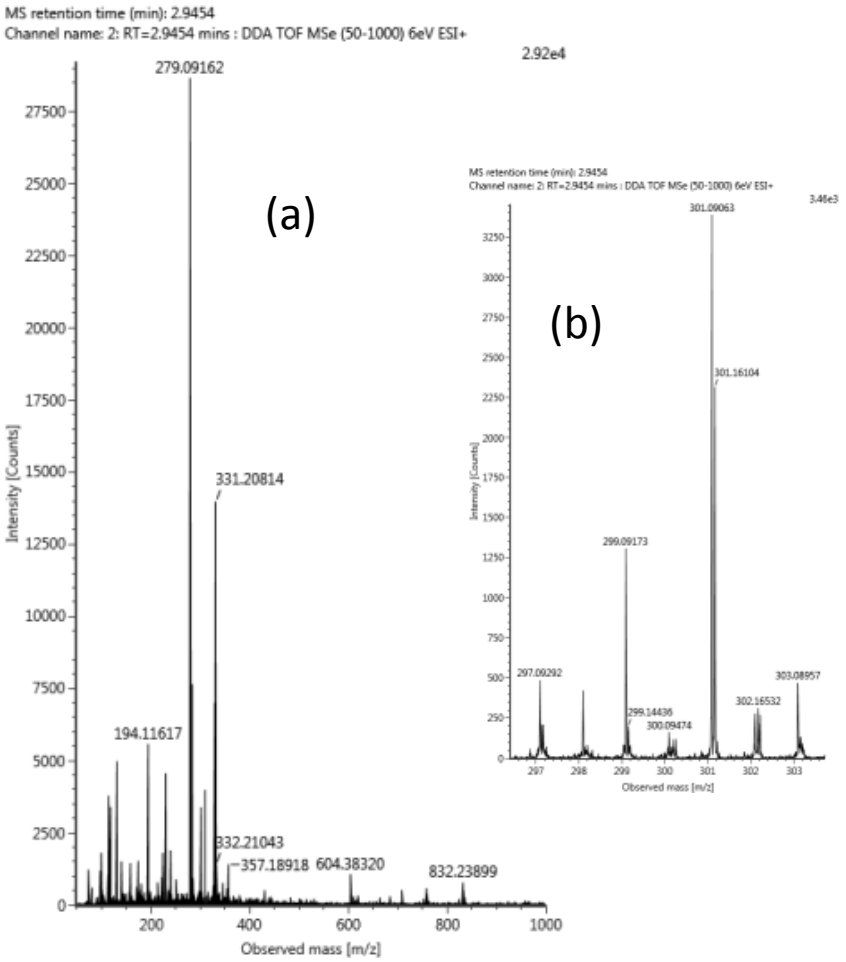
Compound at the experimental m/z 241.99296 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



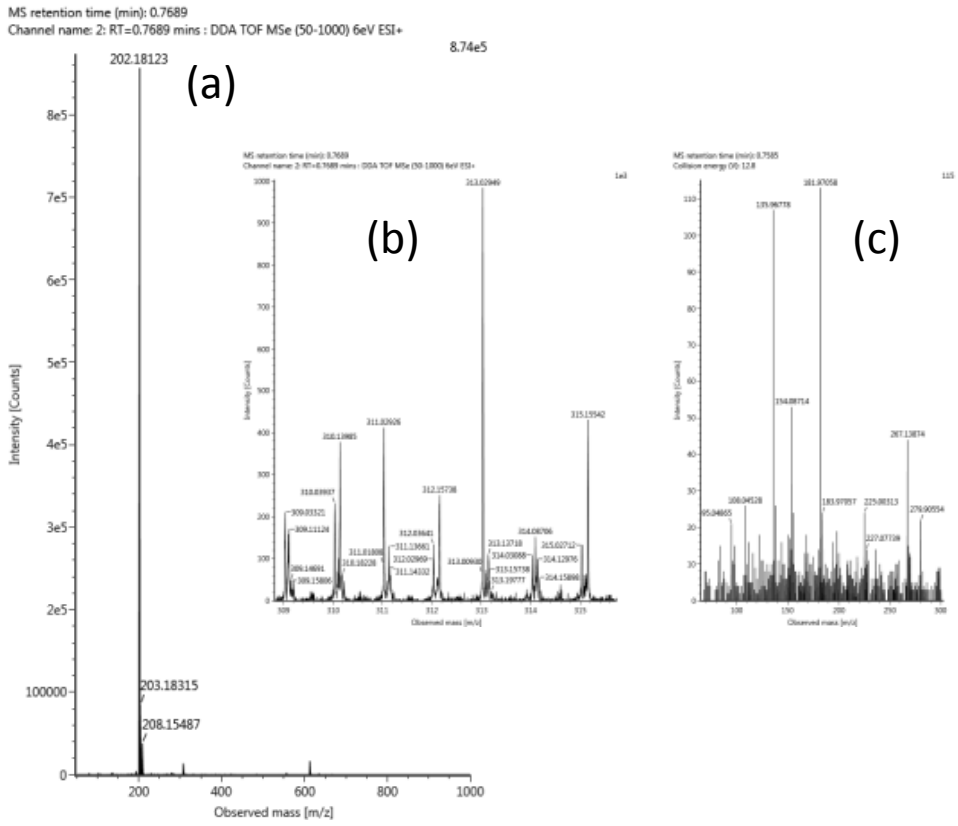
Compound at the experimental m/z 268.04482 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



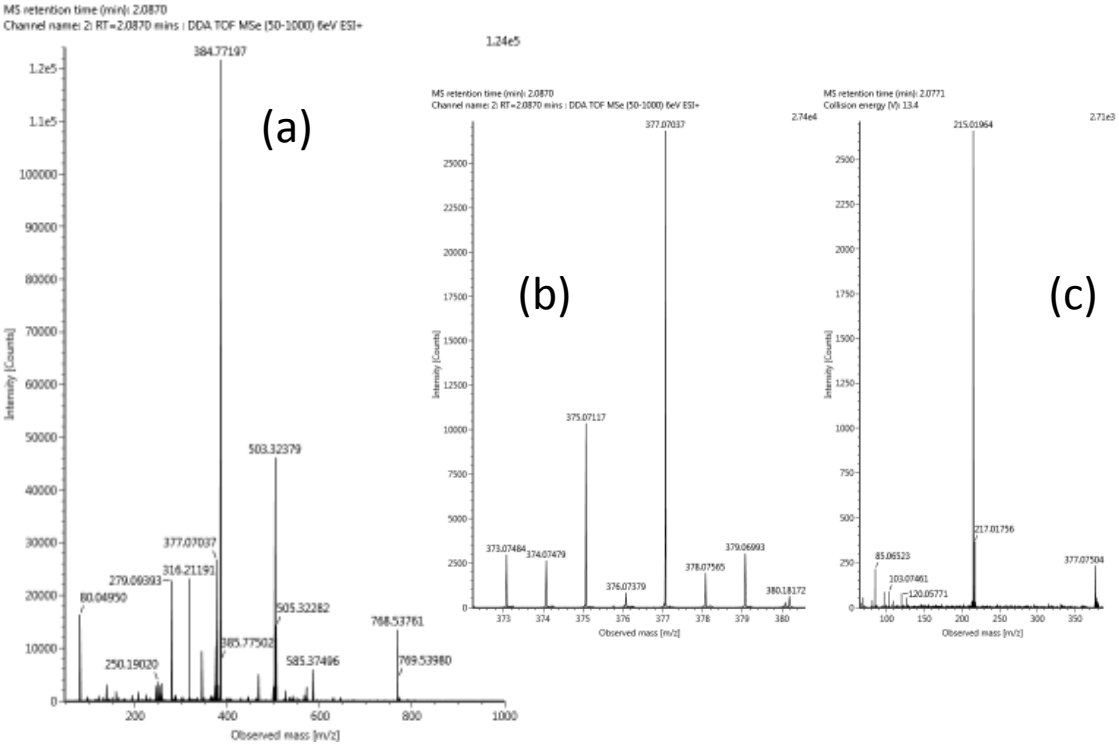
Compound at the experimental m/z 270.02341 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/. MS/MS spectrum couldn't be recorded because of low abundance.



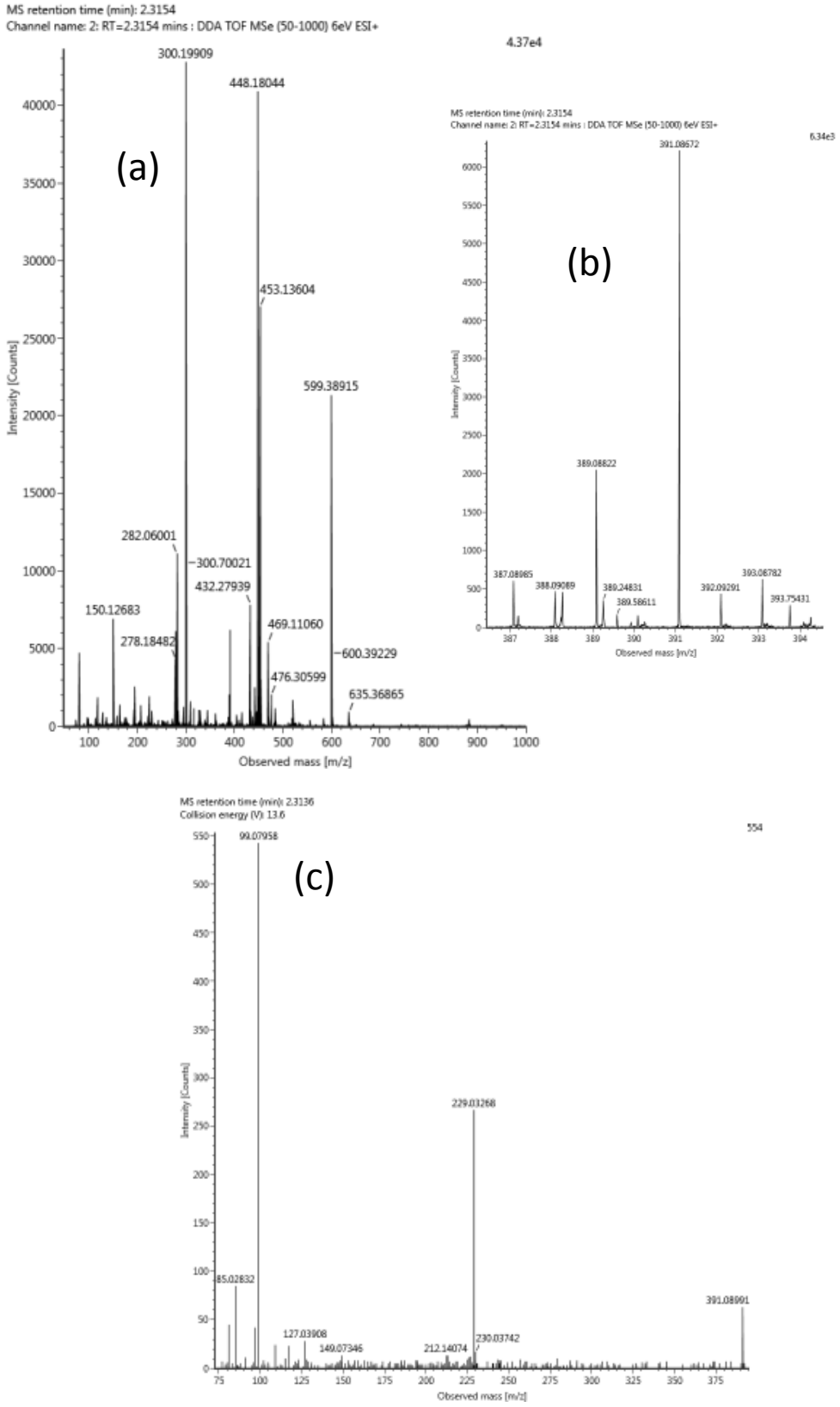
Compound at the experimental m/z 301.09063 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/. MS/MS spectrum couldn't be recorded because of spectral interference.



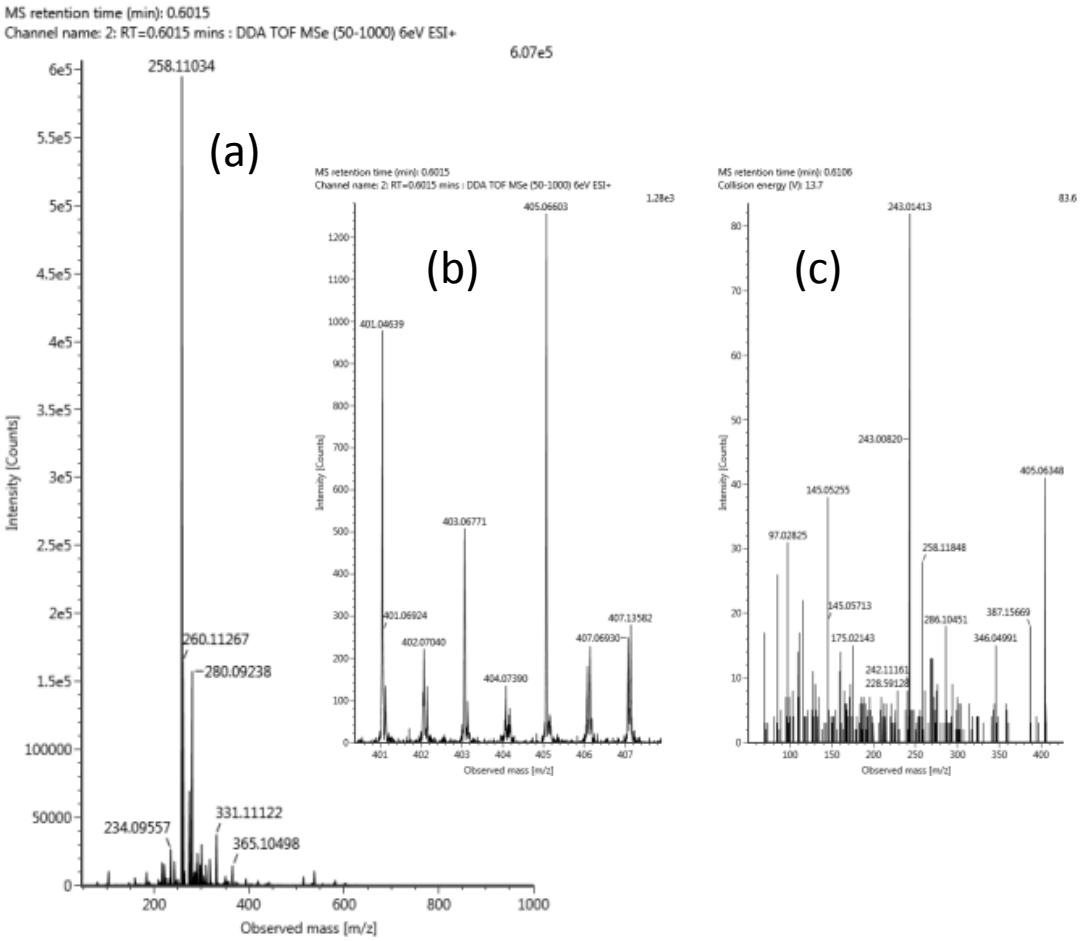
Compound at the experimental m/z 313.02949 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



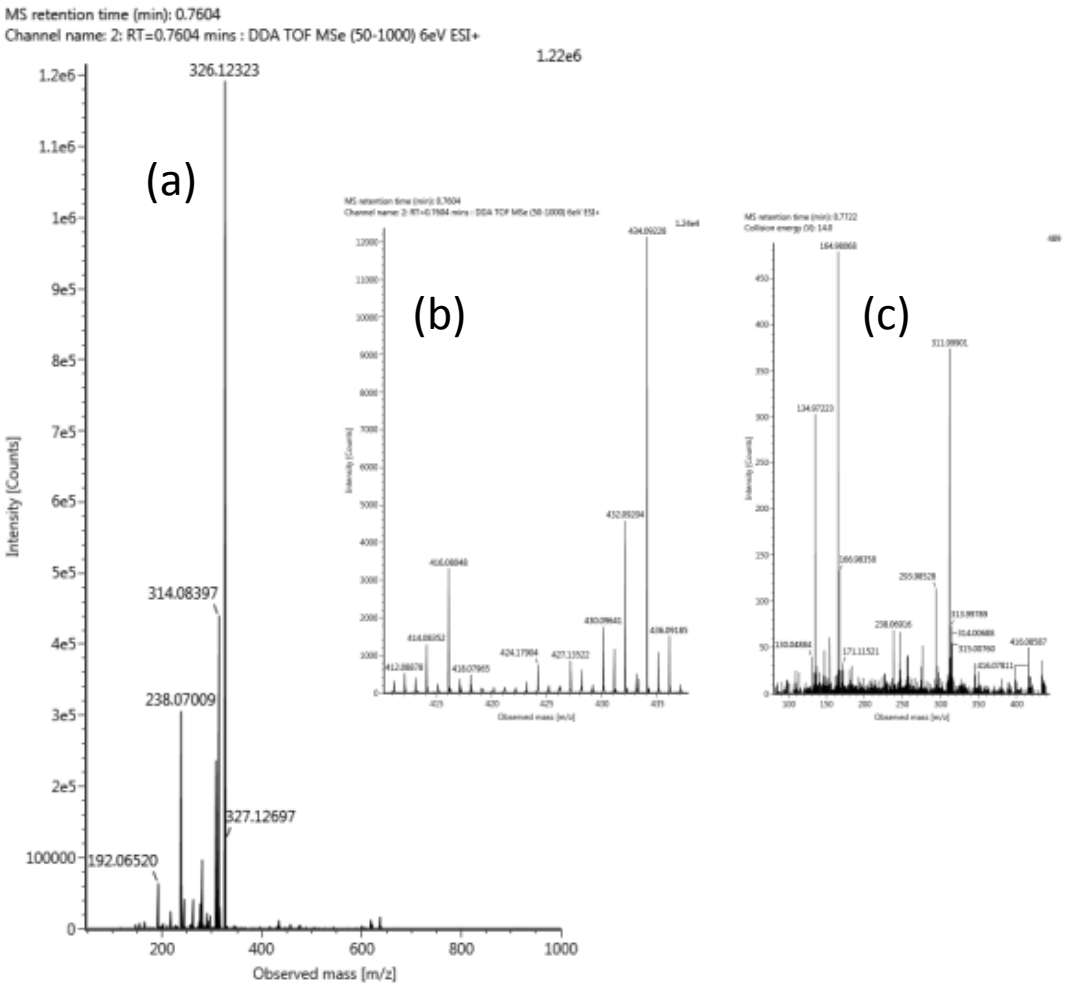
Compound at the experimental m/z 377.07037 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



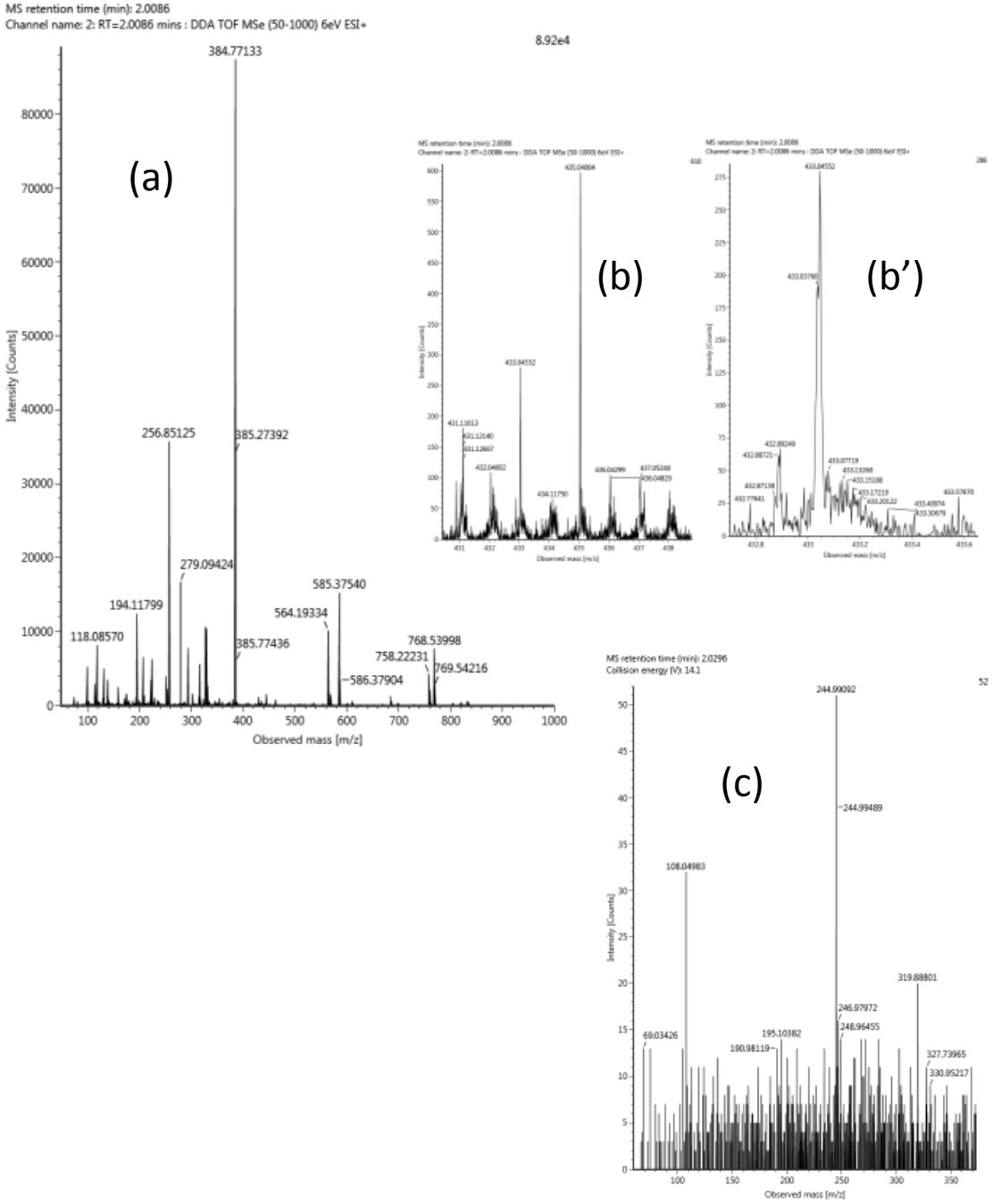
Compound at the experimental m/z 391.08672 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



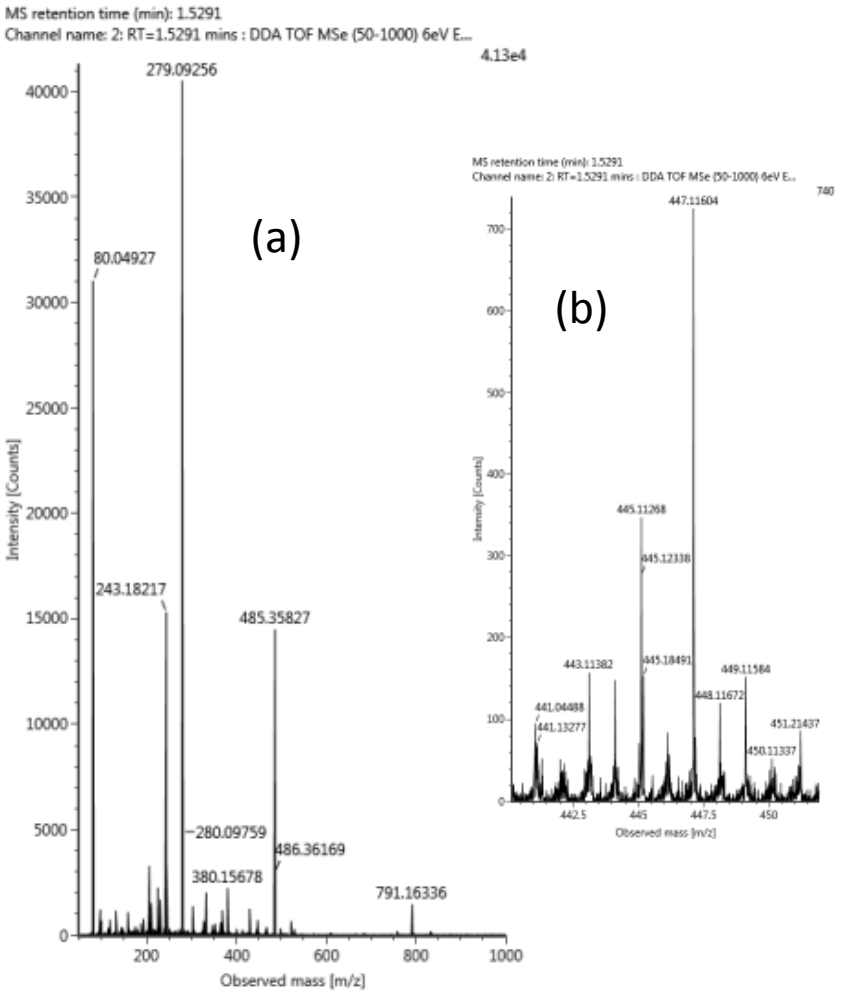
Compound at the experimental m/z 405.06603 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



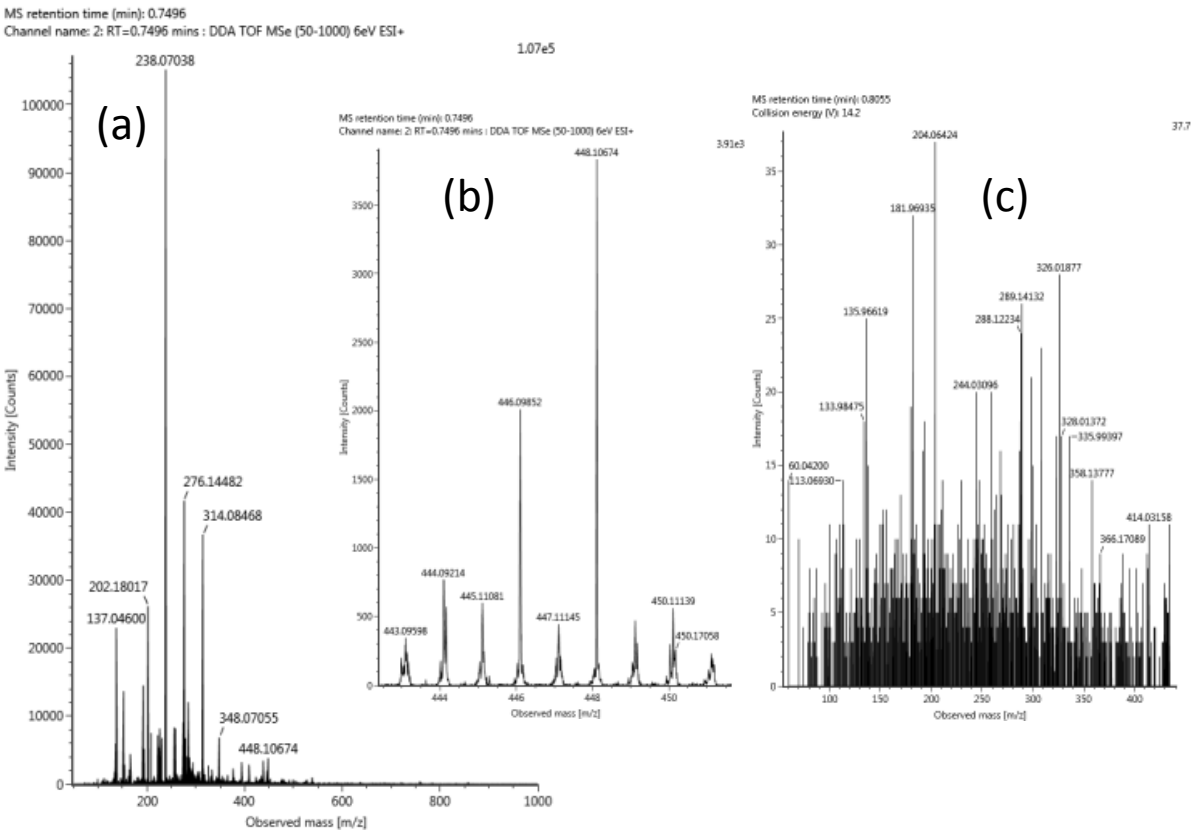
Compound at the experimental m/z 434.09228 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



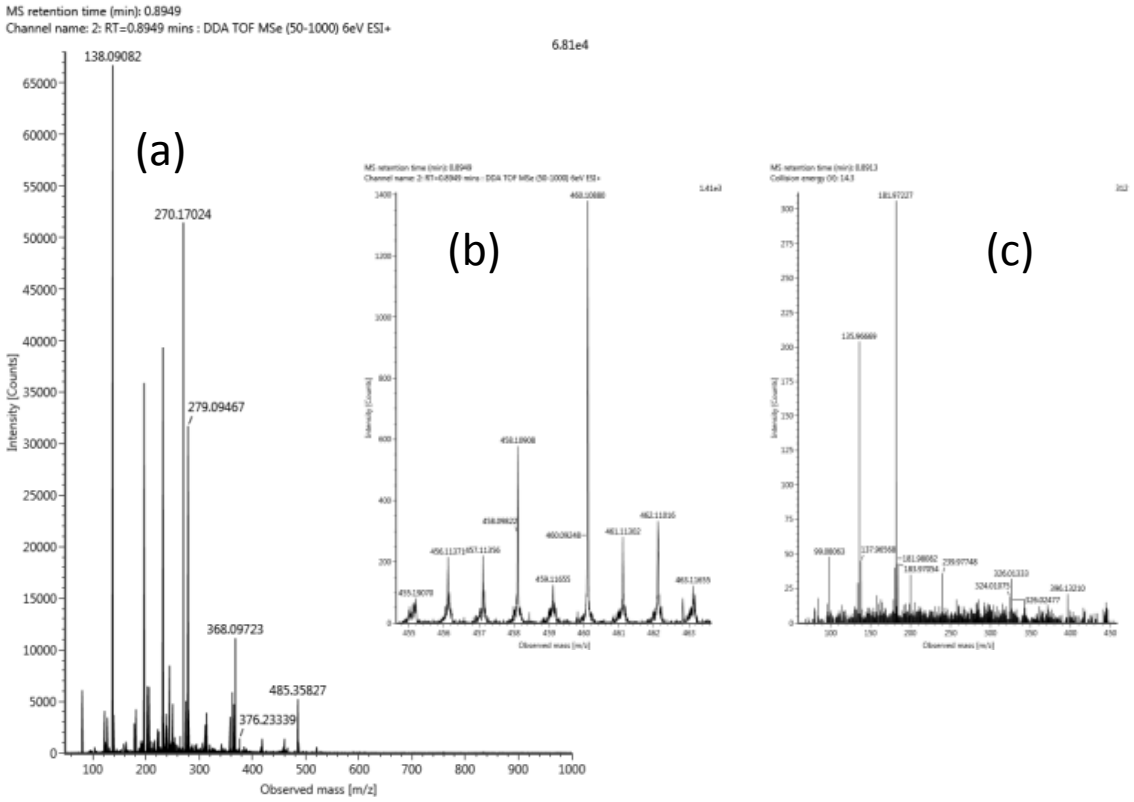
Compound at the experimental m/z 435.04004 (for details, see Table 1); (a) full scan spectrum; (b) and (b') full scan spectra /zoomed/; (c) MS/MS spectrum.



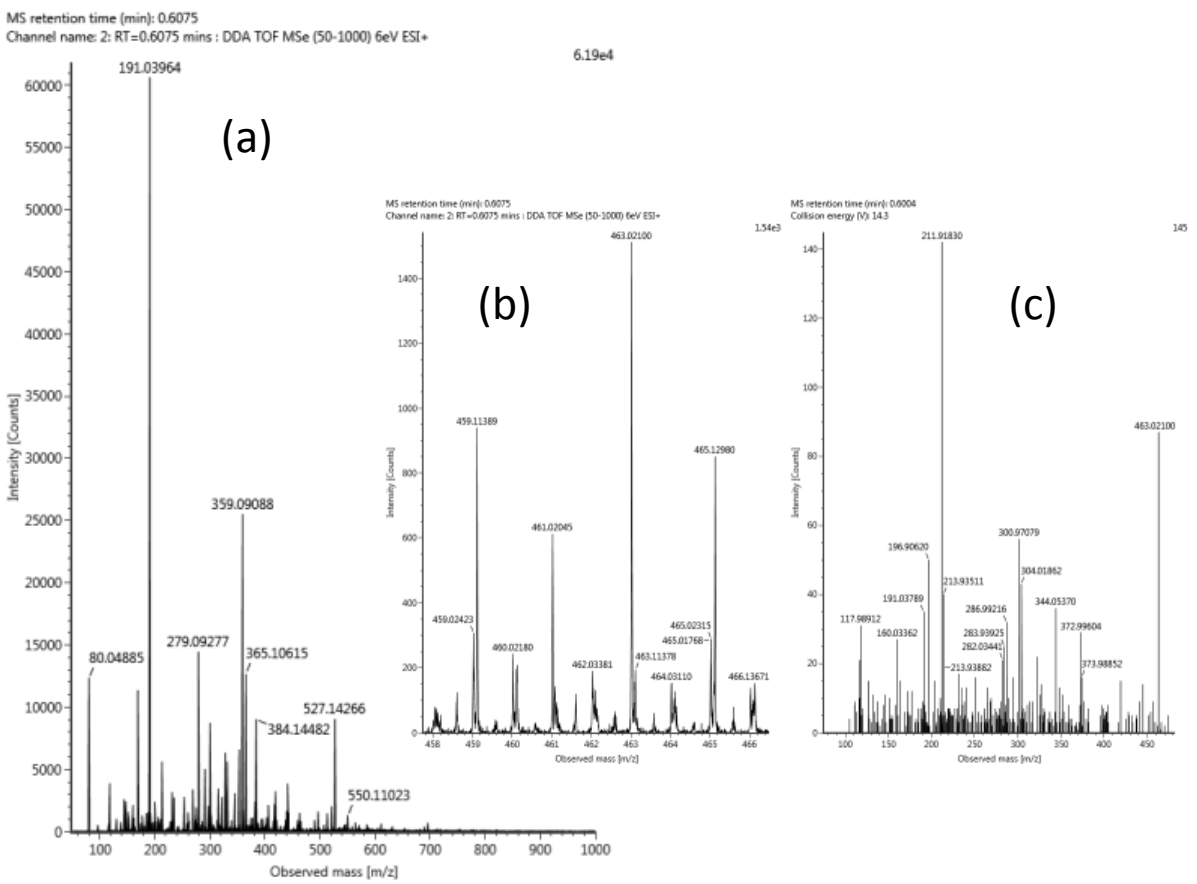
Compound at the experimental m/z 447.11604 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/. MS/MS spectrum couldn't be recorded because of low abundance.



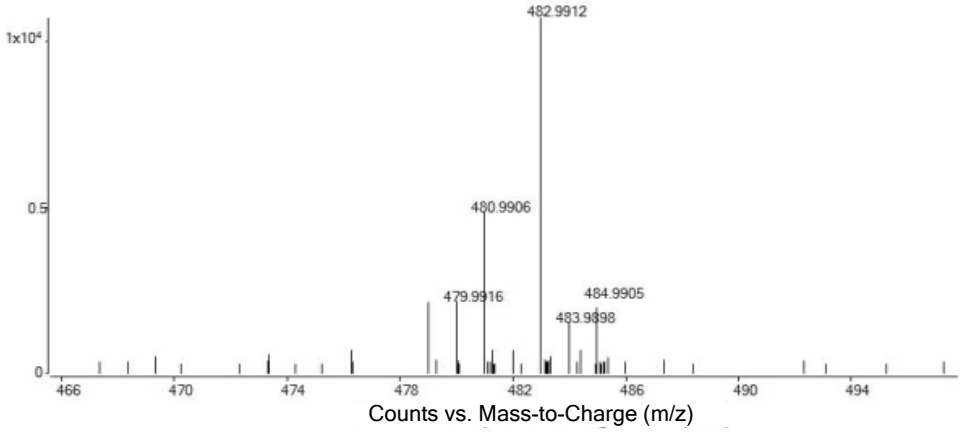
Compound at the experimental m/z 448.10674 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



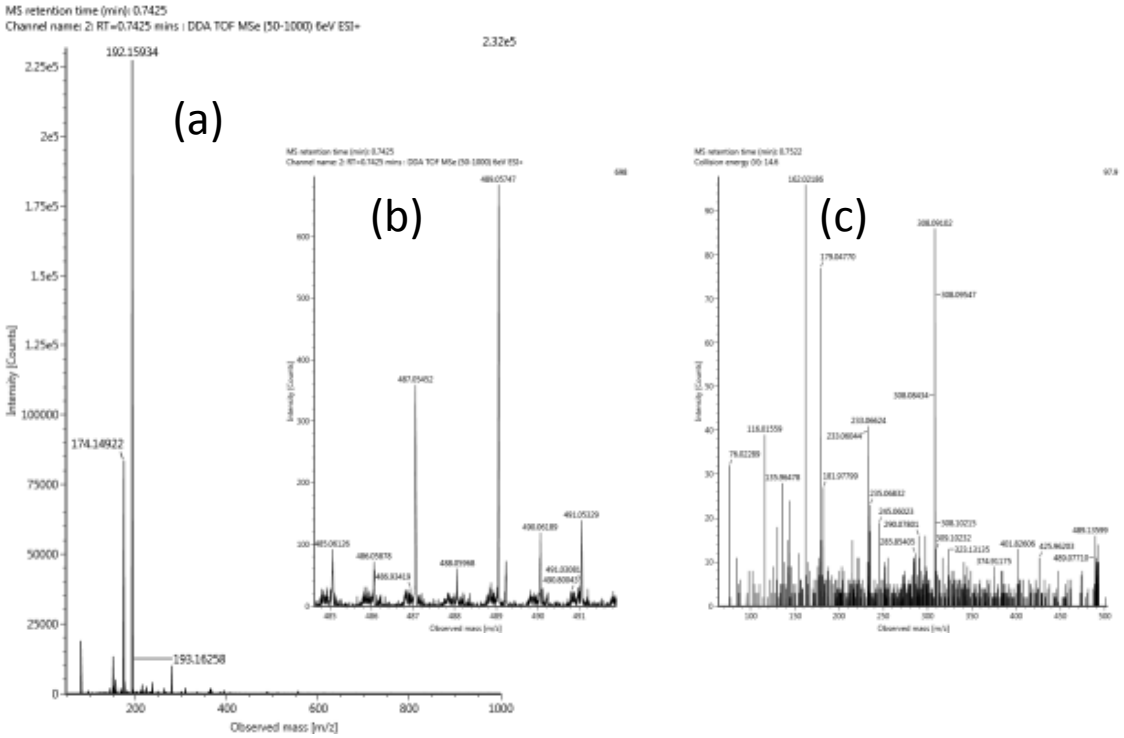
Compound at the experimental m/z 460.10880 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



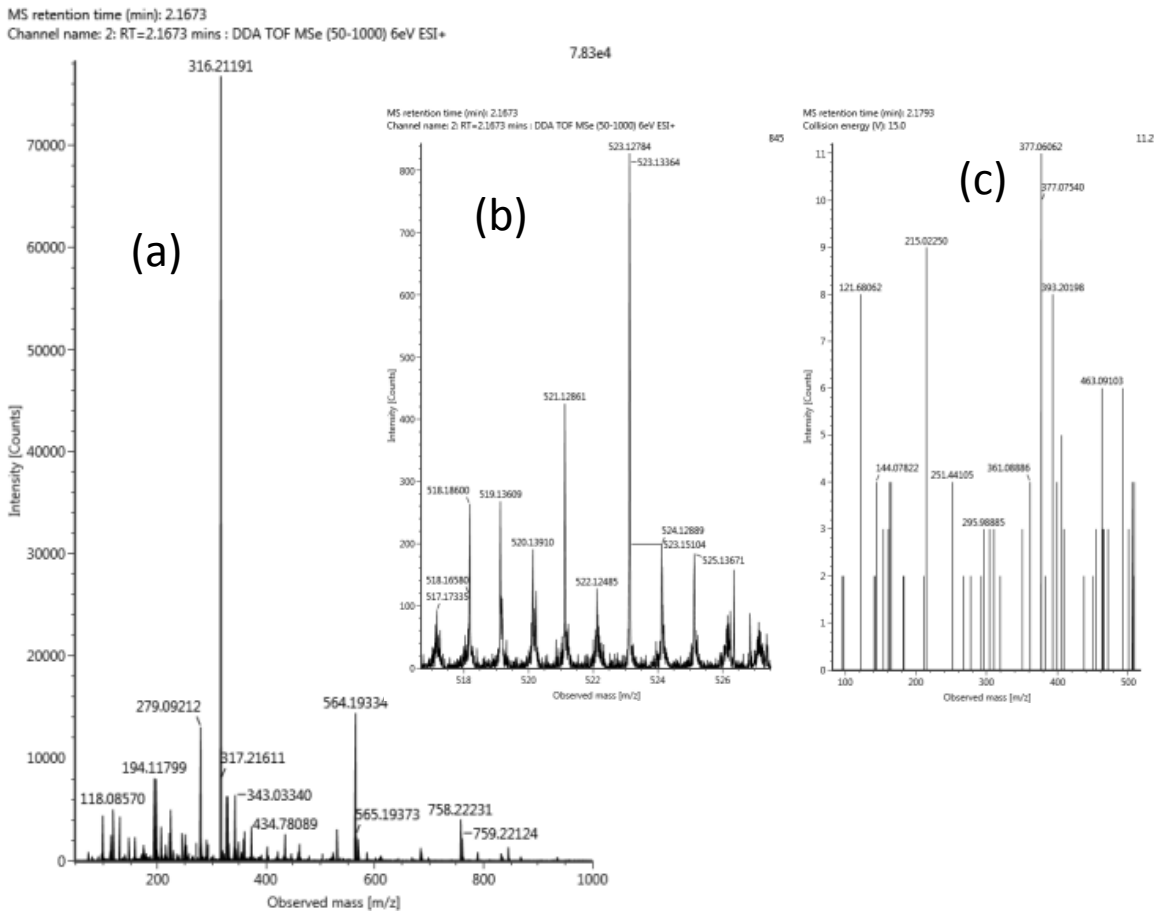
Compound at the experimental m/z 463.02100 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



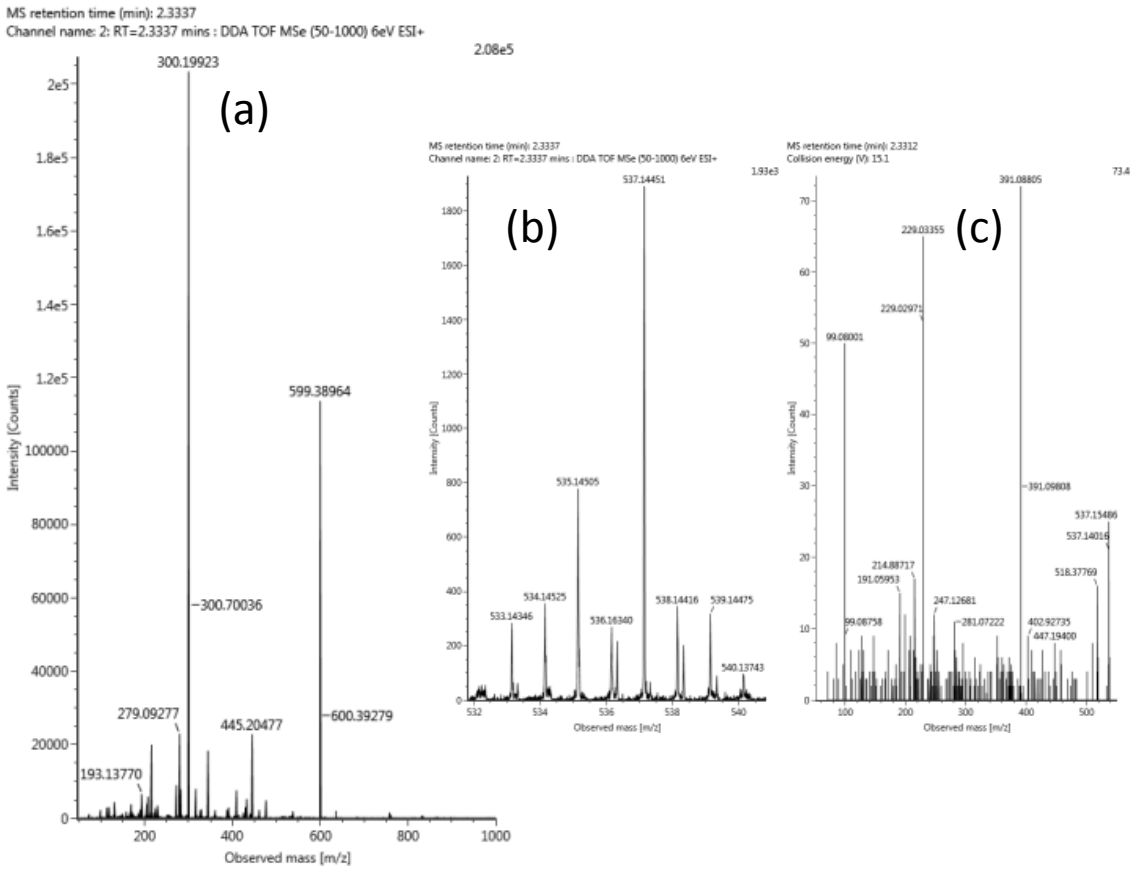
Compound at the experimental m/z 482.9912 (for details, see Table 1). MS/MS spectrum couldn't be recorded because of low abundance. Data obtained with an Agilent 6530 ESI-QTOFMS system.



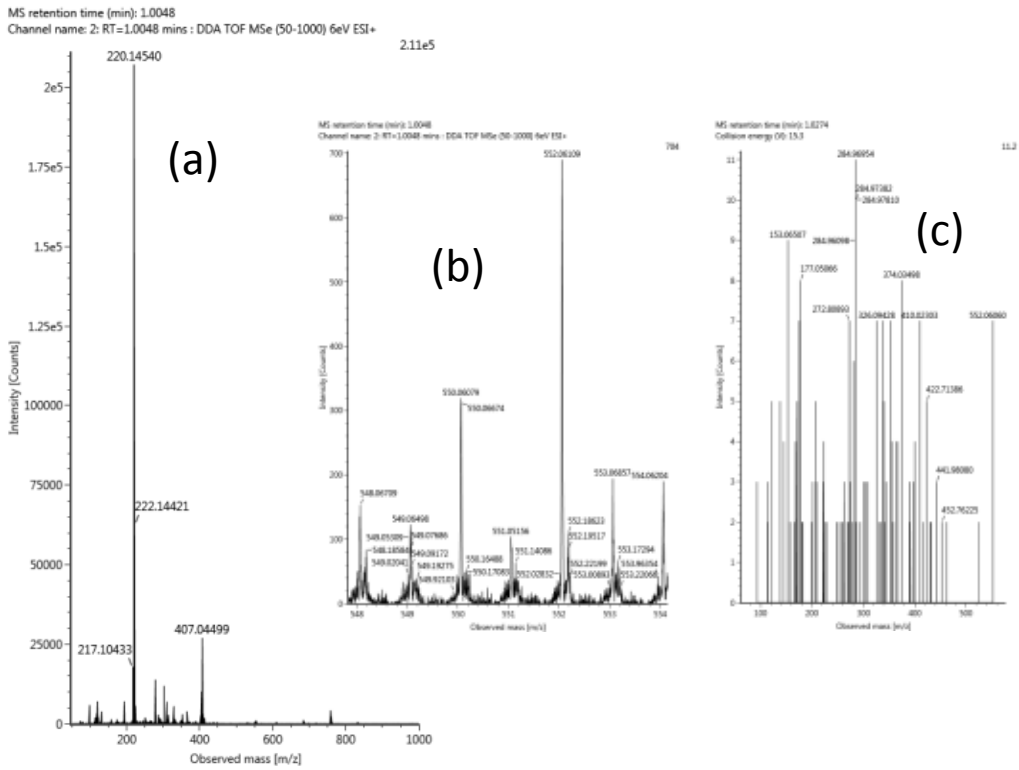
Compound at the experimental m/z 489.05747 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



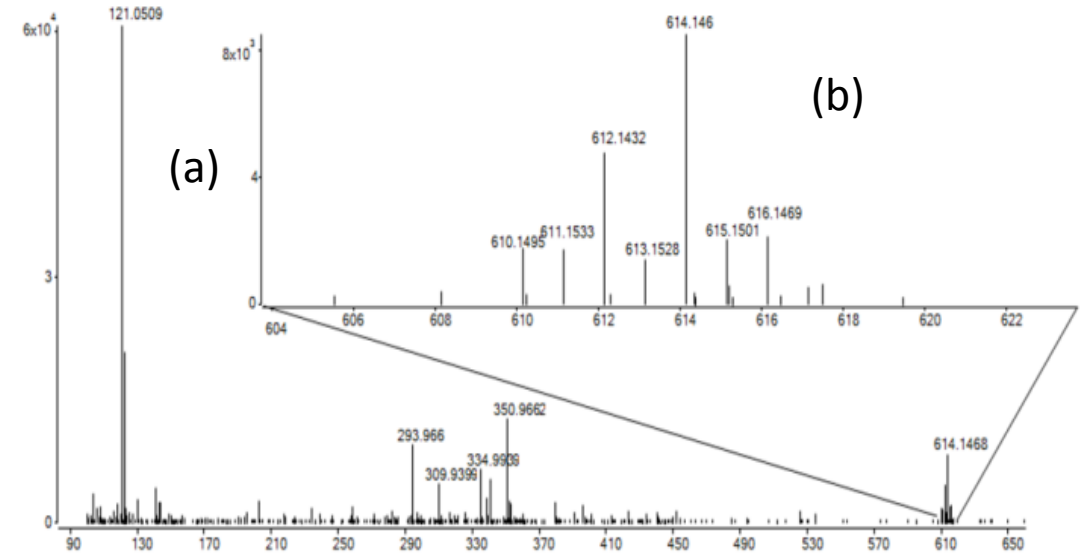
Compound at the experimental m/z 523.12784 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.

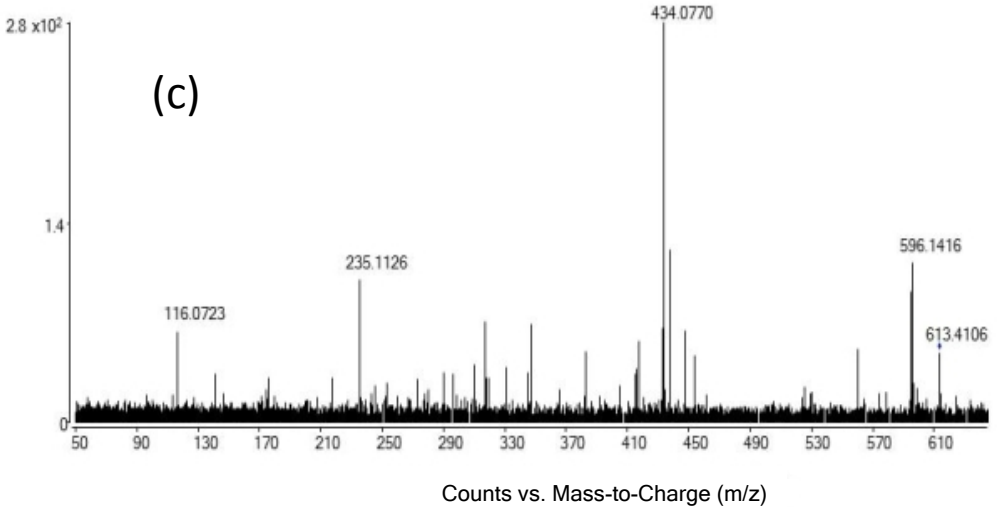


Compound at the experimental m/z 537.14451 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.



Compound at the experimental m/z 552.06109 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum.





Compound at the experimental m/z 614.1468 (for details, see Table 1); (a) full scan spectrum; (b) full scan spectrum /zoomed/; (c) MS/MS spectrum. Data obtained with an Agilent 6530 ESI-QTOFMS system.

From **Centre for Agricultural Research**
Agricultural Institute
2462 Martonvásár, Brunszvik u.2.
Hungary

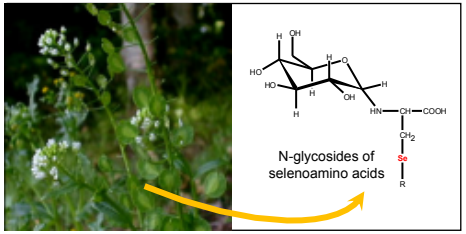
e-mail dernovics.mihaly@atk.hu

Subject "Significance to Metallomics" statement

Date 16/09/2020

A completely novel class of selenium species, namely, N-glycoside derivatives of selenoamino acids are presented in the manuscript on the basis of LC-Unispray-QTOF-MS datasets.

This is the most comprehensive selenometabolome study up to now not only in the genus Cardamine but in the whole Brassicaceae family, including structure elucidation and a detailed and stepwise method development for software based metabolite filtering.



Appearance of selenium containing sugars, including N-glycosylated selenoamino acids, has been observed in a selenium hyperaccumulator plant from the Cardamine genus.

From **Centre for Agricultural Research**
Agricultural Institute
2462 Martonvásár, Brunszvik u.2.
Hungary
e-mail dernovics.mihaly@atk.hu

Subject Cover letter/Letter to referees

Date 16/09/2020

To whom it concerns

Dear Referees,

Please find enclosed our manuscript that presents the study on the selenometabolome of the hyperaccumulator plant Cardamine species, C. violifolia. This study has been carried out in the cooperation of five institutes from three countries, namely, France, China and Hungary. We believe this manuscript contains significant novelties in the field of selenium speciation studies:

- this is the most comprehensive selenometabolome study up to now not only in the genus Cardamine but in the whole Brassicaceae family, including the overall characterization of water soluble selenium species, method development for software based metabolite filtering, precise (isotope dilution assisted) quantification of the main selenium metabolite, etc.;
- a completely novel class of selenium species, namely, N-glycoside derivatives of selenoamino acids are presented in the manuscript;
- structural identification section has been completed with an extensive and detailed dataset on high resolution MS and MS/MS information and a stepwise protocol about selenium metabolite detection.

We hope you will find our revised manuscript of high quality enough to be considered for publication.

Yours sincerely,

Martonvásár (Hungary), 16/09/2020

Mihály Dernovics, PhD, dr. habil.