Anal Bioanal Chem
DOI 10.1007/s00216-017-0252-7

RESEARCH PAPER

Tackling saponin diversity in marine animals by mass spectrometry: data acquisition and integration

Corentin Decroo[1] · Emmanuel Colson[1] · Marie Demeyer[1] · Vincent Lemaur[2] · Guillaume Caulier[3] · Igor Eeckhaut[3] · Jérôme Cornil[2] · Patrick Flammang[3] · Pascal Gerbaux[1]

Received: 29 November 2016 / Revised: 27 January 2017 / Accepted: 7 February 2017

Abstract Saponin analysis by mass spectrometry methods is nowadays progressively supplementing other analytical methods such as nuclear magnetic resonance (NMR). Indeed, saponin extracts from plant or marine animals are often constituted by a complex mixture of (slightly) different saponin molecules that requires extensive purification and separation steps to meet the requirement for NMR spectroscopy measurements. Based on its intrinsic features, mass spectrometry represents an inescapable tool to access the structures of saponins within extracts by using LC-MS, MALDI-MS, and tandem mass spectrometry experiments. The combination of different MS methods nowadays allows for a nice description of saponin structures, without extensive purification. However, the structural characterization process is based on low kinetic energy CID which cannot afford a total structure elucidation as far as stereochemistry is concerned. Moreover, the structural difference between saponins in a same extract is often so small that coelution upon LC-MS analysis is unavoidable, rendering the isomeric distinction and characterization by CID challenging or impossible. In the present paper, we introduce ion mobility in combination with liquid chromatography to better tackle the structural complexity of saponin congeners. When analyzing saponin extracts with MS-based methods, handling the data remains problematic for the comprehensive report of the results, but also for their efficient comparison. We here introduce an original schematic representation using sector diagrams that are constructed from mass spectrometry data. We strongly believe that the proposed data integration could be useful for data interpretation since it allows for a direct and fast comparison, both in terms of composition and relative proportion of the saponin contents in different extracts.

Keywords Triterpene glycosides · Echinoderms · Sea cucumbers · Mass spectrometry · Natural products · Ion mobility · MALDI-ToF · LC-MS

Electronic supplementary material The online version of this article (doi:10.1007/s00216-017-0252-7) contains supplementary material, which is available to authorized users.

✉ Patrick Flammang
Patrick.Flammang@umons.ac.be

✉ Pascal Gerbaux
Pascal.Gerbaux@umons.ac.be

[1] Organic Synthesis and Mass Spectrometry Lab, Interdisciplinary Center for Mass Spectrometry, Research Institute for Biosciences, University of Mons—UMONS, 23 Place du Parc, 7000 Mons, Belgium

[2] Laboratory for Chemistry of Novel Materials, Center of Innovation and Research in Materials and Polymers, Research Institute for Science and Engineering of Materials, University of Mons—UMONS, 23 Place du Parc, 7000 Mons, Belgium

[3] Biology of Marine Organisms and Biomimetics, Research Institute for Biosciences, University of Mons—UMONS, 23 Place du Parc, 7000 Mons, Belgium

Introduction

Saponins are secondary metabolites largely distributed in plants such as *Yucca schidigera* (Mexico), *Quillaja saponaria* (Chile), and *Sapindus mukorossi* (India) [1]. Saponins have also been discovered within the animal kingdom, particularly in different marine organisms such as sea stars [2, 3], sea cucumbers [4, 5], and sponges [6]. The extreme structural diversity of marine animal saponins makes them challenging from a structural analysis point of view since, most of the time, saponin extracts consist in a huge number of congeners presenting only subtle structural differences [7]. In a recent review, Bahrami et al. reported over 700 triterpene glycosides

Published online: 01 March 2017

identified so far from the Holothuroidea [7]. Saponins are amphiphilic molecules presenting strong membranolytic properties and their hemolytic, cytotoxic, antibacterial, antifungal, antiviral, and anti-tumor properties are already well documented [8]. Obviously, the identification of new active molecules remains of prime importance for industrial applications in pharmaceutical companies [9]. Paradoxically, their biological roles in plants and animals are still not fully understood and it is speculated that saponins could play a role in several activities, such as chemical defense [10] or interspecific chemical communication [11–13]. The identification of saponins is definitively a prerequisite to the establishment of their structure/biological activity relationships in a given animal species.

In view of their numerous pharmacological activities, saponins currently represent a hot research topic with more than 2000 published papers in 2015 (from Scopus database). The regular increase number of published papers reporting the identification of new saponins is due principally to the development of state-of-the-art analytical tools, especially mass spectrometry-based methods [14].

During the last 5 years, we extensively developed and used mass spectrometry, including MALDI imaging, for the analysis of echinoderm saponins, with a special interest in sea cucumber [15–18] and sea star [19, 20] saponins. Nevertheless, when analyzing saponin extracts with MS-based methods, the analyst immediately faces a huge number of data since saponin congeners are not only presenting different compositions, but isomeric congeners are also present in saponin extracts. Handling the data remains problematic for the comprehensive report of the results, but also for their efficient comparison on the way to the structure/activity relationship elaboration. Often, identified saponins are compiled in tables presenting all molecules together with, at least, their molecular masses, their chemical compositions, and the mass-to-charge ratios of the corresponding ions [15–20]. These tables, however, are usually difficult to read and interpret.

The objectives of the present study are multiple: (i) to describe an efficient global mass spectrometry-based protocol for saponin analysis; (ii) to introduce ion mobility (IM) experiments in the identification workflow [21, 22]; and (iii) to propose a schematic representation using sector diagrams for the data presentation and comparison. Indeed, we strongly believe that an efficient way to organize the exhaustive MS information is mandatory for interpretation of the experimental results and for a direct and fast comparison, both in terms of composition and relative proportion, of the saponin contents in different extracts.

For the present study, we selected the sea cucumber *Holothuria forskali* as a case study. Indeed, we already thoroughly investigated the complex saponin content of two organs of this animal, i.e., the body wall and the Cuvierian tubules, our results pointing to an extreme diversity of the saponin congeners providing extracts difficult to analyze [15, 16]. In addition, we will also extract saponins from the gonads (female sea cucumbers) for different reasons. First, in our previous report dedicated to the defensive role of saponins from the body wall and Cuvierian tubules [15], we identified some new saponins but some congeners previously described by Rodriguez et al. [23] were missing in our data. We suspected at that time that those saponins could be characteristic of other organs, such as the gonads. Secondly, saponins were also proposed to play a role in reproduction in echinoderms [19]. Therefore, the elucidation of the saponin content of gonads is a prerequisite on the way to decipher their biological roles. Finally, when analyzing the saponin contents of the sea stars *Asterias rubens* [15, 16, 18], we observed gonad-specific saponins enlarging the saponin diversity occurring in a single animal.

Materials and methods

Chemicals, animal sampling, and saponin extractions

Chemicals For saponin extractions and mass spectrometry analyses, technical-grade methanol, hexane, dichloromethane, chloroform and isobutanol, as well as HPLC grade water, acetonitrile, and methanol were purchased from CHEM-LAB NV (Somme-Leuze, Belgium). *N,N*-dimethylaniline (DMA) and 2,5-dihydroxybenzoic acid (DHB) were provided by Sigma-Aldrich (Diegem, Belgium).

Sampling Individuals of *H. forskali* (Delle Chiaje, 1823) were obtained from the Observatoire Océanologique of Banyuls-sur-Mer (France) during autumn 2015. They were transported to the University of Mons, where they were kept in a marine aquarium with closed circulation (20 °C, pH 7.6, 37 psu). A female holothuroid was dissected and its body wall, gonads, and Cuvierian tubules were stored separately in methanol at 4 °C. Animals used in our experiments were maintained and treated in compliance with the guidelines specified by the Belgian Ministry of Trade and Agriculture.

Saponin extractions The body wall, gonads, and Cuvierian tubules underwent the same extraction method, adapted from Van Dyck et al. [15]. The homogenized tissues are extracted with methanol followed by filtration. The extracts are diluted to 70% methanol with water. The solutions are partitioned against *n*-hexane, CH_2Cl_2, and $CHCl_3$, respectively (*v*/*v*). Finally, the hydromethanolic solution is evaporated at low pressure in a double boiler at 46 °C using a rotary evaporator. The dry extract is diluted in water in order to undergo a last partitioning against isobutanol (*v*/*v*). The butanolic phase is washed twice with water to remove salts. The organic solution contains the purified saponins.

Mass spectrometry analyses

MALDI-ToF analyses The body wall, Cuvierian tubule, and gonad saponin extracts were analyzed with a Waters Q-ToF Premier mass spectrometer in the positive ion mode. All the detected ions correspond to Na^+ adducts on saponins and are therefore detected at $m/z = M + 23$. The MALDI source is constituted of a Nd-YAG laser, operating at 355 nm with a maximum pulse energy of 104.1 μJ delivered to the sample at 200 Hz repeating rate. All samples were prepared using a mixture of 25 mg of DHB (2,5-dihydroxybenzoic acid) in water/acetonitrile (*v*/*v*) with 6 μl of DMA (*N,N*-dimethylaniline) as the matrix. The dry droplet method was selected to prepare the sample/matrix co-crystal on the target plate. In this method, the saponin extract is not premixed with the matrix. A sample droplet (1 μl) is applied on top of a fast-evaporated matrix-only bed. For the recording of the single-stage MALDI-MS spectra, the quadrupole (rf-only mode) was set to pass ions between *m/z* 250 and 1500, and all ions were transmitted into the pusher region of the time-of-flight analyzer where they were mass-analyzed with a 1-s integration time. Accurate mass measurements are also performed at 10,000 FWHM resolution to afford ion composition. To obtain reliable quantitative measurements by MALDI-MS to construct the sector diagrams (see below), it is required to average the data on a large area of the MALDI target covered by the sample/matrix preparation. In practice, we recorded about 100 spectra/acquisition using the spiral mode, allowing the laser to shoot at different spot zones.

LC-MS(MS) analyses For the on-line LC-MS analyses, a Waters Alliance 2695 liquid chromatography device was used. The HPLC device was coupled to the Waters Quattro Premier mass spectrometer and consisted of a vacuum degasser, a quaternary pump, and an autosampler. Sample volumes of 1 μl are injected. Chromatographic separation is performed on a non-polar column (Eclipse plus C18; 4.6 × 100 mm; 3.5 μm; Agilent) at 40 °C. The mobile phase is programmed with a constant flow (1 ml/min) of 60% of eluent A (water, 0.1% formic acid) and 40% of eluent B (acetonitrile) during 15 min (isocratic conditions). The ESI conditions are as follows: positive ion mode ($m/z = M + 23$); capillary voltage, 3.1 kV; cone voltage, 40 V; source temperature, 100 °C; desolvation temperature, 300 °C. Dry nitrogen is used as the ESI gas with a flow rate of 50 l/h for the gas cone and 500 l/h for the desolvation gas. The single-stage LC-MS spectra are recorded by scanning the first quadrupole analyzer between *m/z* 250 and 1500. For the LC-MS/MS experiments, the ions of interest were mass-selected by the first quadrupole. The selected ions are then submitted to collision against argon in the trap cell of the tri-wave device and the laboratory frame kinetic energy (E_{lab}), typically at 80 eV, is selected to afford intense enough product ion signals. The product ions are finally mass measured with the ToF analyzer.

LC-IMS-MS analyses For the ion mobility measurements, experiments were performed using the hybrid quadrupole (Q)–traveling wave (T-wave) ion mobility (TWIMS)–time-of-flight (TOF) mass spectrometer (Synapt G2-Si, Waters, UK) equipped with an electrospray (ESI) ionization source. Typical ion source conditions were as follows: capillary voltage, 2.5 kV; sampling cone, 40 V; source offset, 80 V; source temperature, 150 °C; and desolvation temperature, 300 °C. This mass spectrometer is used for the recording of ESI full scan mass spectrum, for the collision-induced dissociation (CID) as well as for the ion mobility experiments. Briefly, the core of the instrument is constituted by the so-called tri-wave setup that is composed of three successive T-wave elements named the trap cell, the IMS cell, and the transfer cell, in which the wave speed and amplitude are user-tunable. The trap and transfer cells are filled with argon whereas the IMS cell is filled with nitrogen. A small RF-only cell filled with helium is fitted between the trap and the IMS cell.

Molecular dynamics simulations

All sodium-cationized structures reported herein correspond to the lowest energy structures obtained using the following conformational search procedure. For each saponin, different starting geometries, differing by the localization of the sodium cation (typically three complexation sites are considered), have been built and optimized by molecular mechanics using the Dreiding force field [24], as implemented in the Materials Studio package 6.0 [25]. All default parameters have been used except the following: (i) the cutoff value for the non-bonded interactions set at 500 Å so that none of them was neglected; (ii) the conjugated gradient algorithm is selected to achieve the geometry optimizations; (iii) Gasteiger atomic charges have been used; (iv) the distance dependence of the electrostatic interactions follows a 1/r law; and (v) the van der Waals parameters for the sodium atoms have been set to 3.144 kcal/mol for D_0 (equilibrium well depth) and 0.5 Å for R_0 (equilibrium distance); i.e., the values used in the universal force field [26]. Note that new van der Waals parameters have been introduced in Dreiding since its parametrization for other alkali which will be studied later (Li^+, K^+, etc.) is missing and will also be transferred from the universal force field. Each optimized structure for a given saponin is then used as the starting point of successive 1 ns quenched molecular dynamics–quenched MD–(NVT, frames optimized every 0.1 ps) at increasing temperature ($T = 300$, 500 K ($x2$), 750 K). After each of these quenched MD, the lowest energy conformer obtained is used as the starting point of the next MD. The lowest energy structure obtained after the final 750 K quenched MD is used as the starting point of the final quenched MD at room temperature (NVT, frames optimized

every 0.1 ps) that is likely to afford the lowest energy structure on the potential energy surface. Consecutively, this structure is subjected to two successive 10 ns MD (NVT, $T = 298$ K). Whereas the first MD is used to equilibrate the system, the aim of the second run is to generate 100 conformations (0.1 ns/frame) that are finally used in the Mobcal program [21] to calculate the room temperature averaged collision cross section (CCS) using the trajectory method (TM).

Data integration process

For constructing the sector diagrams, several data are required and were obtained in the course of the MS experiments. As described in the text, we decided to limit each sector diagram to the comparison of the saponin molecules presenting the same number of monosaccharidic residues in their glycone chain. Based on our measurements, we determined the following:

(i) The *m/z* ratios (and then the corresponding elemental compositions by accurate mass measurements—see Electronic Supplementary Material (ESM) Table S1) from the MALDI-ToF and the molar ratios between elemental compositions by peak integration (all isotopic signals) in the MALDI mass spectra;
(ii) The retention times by LC-MS experiments and the molar ratios between isomers by peak integration in the extracted ion current (EIC) chromatograms;
(iii) The drift times by LC-IMS-MS experiments and the molar ratios between ions with different t_D by peak integration in the arrival time distribution (ATD) chromatograms, the so-called mobilograms.

The sector diagrams are generated using Microsoft Excel 2016 in Microsoft Office Professional Plus 2016.

Results and discussion

Holothuroid saponins are triterpene glycosides. As presented in Fig. 1, the structure of the aglycone moiety, a holostane-3β-ol, is derived from the tetracyclic triterpene lanostane-3β-ol in which the D-ring contains a γ-18(20)lactone. The oligosaccharidic chain is covalently attached to the C3 of the aglycone and may include xylose, glucose, quinovose, and 3-O-methylglucose residues [15].

Based on MS methods, we detected at least 16 saponins in the body wall and Cuvierian tubule extracts of *H. forskali* (Table 1). MALDI-ToF, LC-MS, and tandem mass spectrometry (LC-MS/MS) experiments have been used complementarily and are considered as relevant tools to detect these molecules in complex mixtures. We actually developed a stepwise protocol that start by a global overview of the saponin extracts by analyzing the sample with MALDI-ToF. This first step affords a global description of the saponin contents in terms of elemental composition following exact mass measurements (high resolution mass spectrometry) (see ESM Table S1), without attempting any structural characterization or isomeric discrimination. The outcome of this step is the measurement of the mass-to-charge ratio (*m/z*) of all saponin ions and then the determination of the elemental composition of the corresponding neutral molecules.

The expected result of the second step that involves LC-MS (liquid chromatography) analysis is the separation of saponins based on their retention times (t_R) allowing the discrimination of isomeric saponins (different t_R but same *m/z*). Structural characterization of saponin ions is further based on LC-MS/MS experiments in which mass-selected saponin ions are exposed to collision-induced dissociation (CID) experiments yielding fragment ions, useful for reconstructing the structure of the parent ions. We successfully applied this methodology to investigate the saponin content of different sea cucumber and sea star species [15–20, 27]. Here, the structural characterization process is based on low kinetic energy CID which cannot afford a total structure elucidation as far as stereochemistry is concerned. It is for instance impossible to distinguish α- from β-glycosidic bonds, or epimeric monosaccharides such as glucose and galactose. The structural difference between saponins in a same extract is often so small that coelution upon LC-MS analysis is unavoidable, rendering the isomeric distinction and characterization upon CID challenging or impossible. This led us to envisage the association between ion mobility and liquid chromatography to better tackle the complexity of saponin congeners. In the present work, we will just consider IMMS as an additional chromatographic step keeping in mind that, in the near future, we will also consider IMMS for structural analysis by measuring the collisional cross sections (CCS) of saponin ions. Actually, when starting acquiring ion mobility data (T-wave experimental setup), we observed that the usual CCS calibration procedure based on polyalanine [28] was not applicable. Indeed, for singly charged ions, the CCS values range between 89 and 228 Å^2 [28], whereas the saponin +1 ion CCS are calculated larger (see below in Fig. 4). Also, CCS data become substantial only when they could be correlated with the CCS of candidate ions generated by computational chemistry [29], and vice versa. In the absence of a reliable CCS calibration, we consequently prefer to present the experimental drift time values that are not contaminated by any approximation.

The association of the three MS-based methods, i.e., MALDI-HRMS, LC-MS/MS, and LC-IMS-MS, represents a very efficient combination for saponin identification. Nevertheless, the identification workflow can be performed using other (associations of) MS methods, provided HRMS, LC, MS/MS, and IMS experiments are accessible. In the present paper, we associate three different instruments (QToF

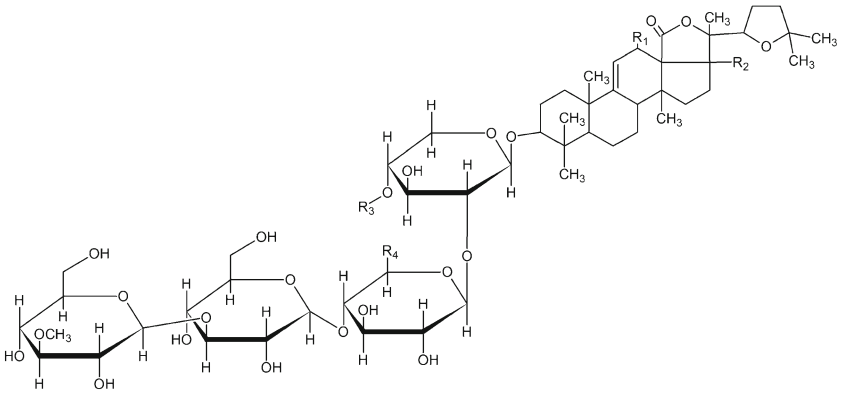

Fig. 1 Saponins of *H. forskali* identified in the body wall and Cuvierian tubules extracts [15]. See Table 1 for details

Premier, Quattro Premier, and Synapt G2—see experimental section) and two different ion sources (electrospray and MALDI). However, the saponin analysis could be also performed on a single instrument, such as the Waters Synapt G2-Si mass spectrometer in LC-IMS-MS/MS configuration.

Data acquisition—MALDI-ToF measurements

Three extracts have been prepared from the body wall, the gonads, and the Cuvierian tubules, respectively. The MALDI-ToF mass spectra of these three extracts are gathered in Fig. 2 and clearly confirm that the three organs are characterized by specific saponin contents. For the present study, we rely on previous results to attest that the observed signals clearly correspond to ionized saponins. Nevertheless, the elemental compositions of the detected ions were confirmed or obtained from accurate mass measurements (high resolution mass spectrometry in Table S1 in the ESM). The saponin nature of the detected ions was also confirmed by the MALDI-MS/MS experiments (data not shown here). Basically, the observed saponins can be classified into three different families depending on the number of monosaccharidic residues in their glycone moieties. As previously reported and schematized in Fig. 2, the tetra-, penta-, and hexaglycosidic saponins of *H. forskali* are constructed, as far as the saccharide group is concerned, on the basis of a linear tetrasaccharide chain with addition of one or two monosaccharide residues on the first monosaccharide (attached to the aglycone).

The MALDI spectra clearly highlighted the presence of the three families and, at first glance, the five-sugar saponins are clearly dominant within the saponin extracts from the gonads and the Cuvierian tubules, whereas the three groups of saponins are equally abundant in the body wall extract. The four-sugar saponins are mostly distributed between two elemental compositions detected upon MALDI at *m/z* 1125 and 1141, whereas the addition of one and two monosaccharidic residues renders the elemental compositions more diverse within the

Table 1 Saponins of *H. forskali* identified in the body wall and Cuvierian tubule extracts [15]. See Fig. 1 for the molecular structure

Holothurinoside	Mass	Molecular formula	R_1	R_2	R_3	R_4
A	1280	$C_{60}H_{96}0_{29}$	-OH	-OH	-Glc	$-CH_3$
A_1	1280	$C_{60}H_{96}0_{29}$	-OH	-H	-Glc	$-CH_2OH$
C	1102	$C_{54}H_{86}0_{23}$	-OH	-H	-H	$-CH_3$
C_1	1102	$C_{54}H_{86}0_{23}$	-H	-H	-H	$-CH_2OH$
E	1264	$C_{60}H_{96}0_{28}$	-OH	-H	-Glc	$-CH_3$
E_1	1264	$C_{60}H_{96}0_{28}$	-H	-H	-Glc	$-CH_2OH$
F	1410	$C_{66}H_{106}0_{32}$	-OH	-H	-Glc-Qui	$-CH_3$
F_1	1410	$C_{66}H_{106}0_{32}$	-H	-H	-Glc-Qui	$-CH_2OH$
G	1426	$C_{66}H_{106}0_{33}$	-OH	-OH	-Glc-Qui	$-CH_3$
G_1	1426	$C_{67}H_{108}0_{33}$	-OH	-H	-Glc-Qui	$-CH_2OH$
H	1440	$C_{67}H_{108}0_{33}$	-OH	-H	-Glc-MeOGle	$-CH_3$
H_1	1440	$C_{67}H_{108}0_{33}$	-H	-H	-Glc-MeOGle	$-CH_2OH$
I	1456	$C_{67}H_{108}0_{34}$	-OH	-OH	-Glc-MeOGle	$-CH_3$
I_1	1456	$C_{67}H_{108}0_{34}$	-OH	-H	-Glc-MeOGle	$-CH_2OH$
Desholothurin A	1118	$C_{54}H_{86}0_{24}$	-OH	-OH	-H	$-CH_3$
Desholothurin A_1	1118	$C_{54}H_{86}0_{24}$	-OH	-H	-H	$-CH_2OH$

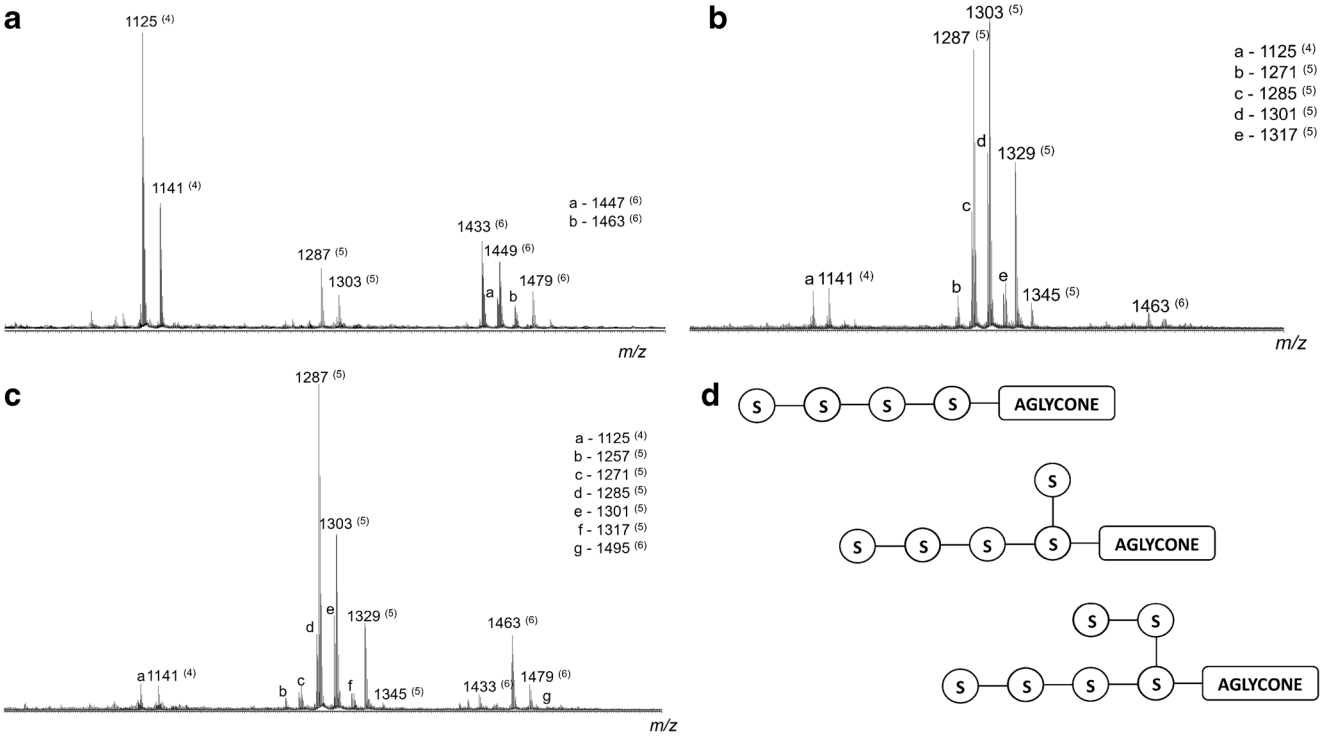

Fig. 2 MALDI-ToF mass spectra of saponin extracts from **a** the body wall; **b** the gonads; and **c** the Cuvierian tubules of *H. forskali*. The *numbers in brackets* represent the number of monosaccharidic residues. **d** Schematic representation of the tetra-, penta-, and hexaglycosidic saponins (*S* stands for saccharidic residue)

five- and six-sugar saponins. For instance, the five-sugar saponins appear in MALDI at *m/z* 1271, 1285, 1287, 1303, 1317, 1329, and 1345, see Fig. 2b for the most representative example. Trying to derive relative concentrations within the extracts from the relative intensities of the *m/z* signals in the MALDI mass spectra is highly speculative since the efficiencies of the ionization/desorption process is molecular structure dependent, i.e., the ion yield (= number of detected gaseous ions vs quantity of condensed phase molecules) is different for (even slightly) different molecules. However, we can assume that the ion yields must be reasonably similar within each of the three families, i.e., four-, five-, and six-sugar saponins. As it will be presented in the data integration section of this paper, we will then use the MALDI intensities to compare saponin compositions mostly within each family. However, keeping in mind the need for caution, MALDI intensities will be used also for comparison between families in order to estimate the relative proportion of each family in the extract (*vide infra*).

Data acquisition—LC-MS(MS) measurements

LC-MS experiments are mandatory for isomer separation. Indeed, behind all the elemental compositions established by MALDI, the presence of isomeric saponins must be considered. This issue can be solved by submitting the saponin extracts to liquid chromatography. In addition, LC-MS/MS experiments are systematically performed to (i) confirm the (primary) structure of the known saponins or (ii) try identifying new congeners based on fragmentation rules and similarities [15]. As an example, we will focus on the *m/z* 1125 ions from the gonad extract (see Fig. 2b for the MALDI spectrum). When examining the *m/z* 1125 extracted chromatogram (Fig. 3a), different peaks are detected at 3.1, 3.3, 4.7, and 5.2 min. HRMS measurements clearly identified the ions at 3.1 and 5.2 min retention time as saponin compositions, whereas the signals at 3.3 and 4.7 min arise from isobaric ions. In Fig. 3b, c, the LC-MS/MS spectra of the *m/z* 1125 ions (3.1 and 5.2 min, respectively) clearly feature saponin-type decomposition schemes [15], with monosaccharide or aglycone losses, see Fig. S1 in ESM. The CID spectrum of Fig. 3c (5.2 min) is characteristic of holothurinoside C, already detected in body wall and Cuvierian tubule extracts [15]. However, the CID spectrum in Fig. 3b that corresponds to the 3.1-min retention time has never been recorded. Based on the decomposition pathway (ESM Fig. S2), we identified a new saponin tentatively named holothurinoside C2 and presented in Fig. 3d. Although the methylglucose-glucose-quinovose-xylose oligosaccharidic chain is present in holothurinoside C (Fig. 1), holothurinoside C2 possesses the atypical methylquinovose-glucose-glucose-xylose sequence (Fig. 2d). This additional saponin was not detected in our

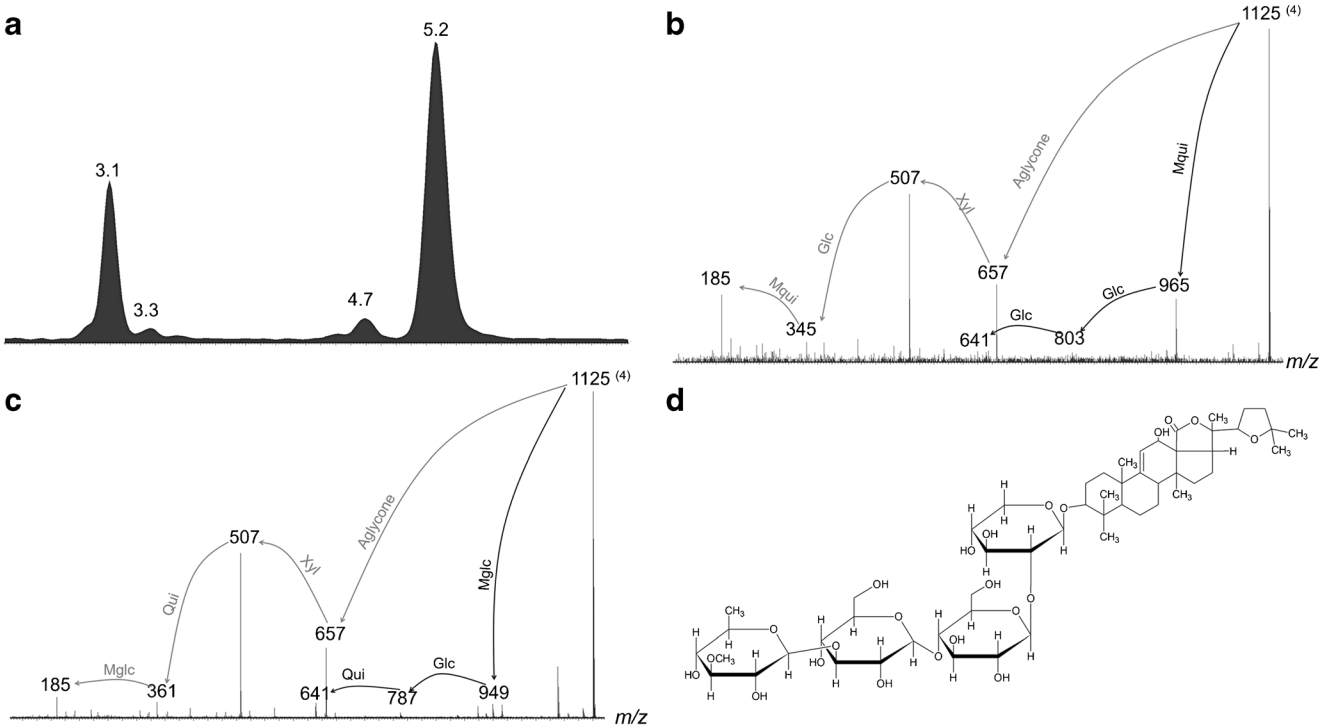

Fig. 3 LC-MS(MS) analysis of saponins from the gonad extract of *H. forskali*. **a** Extracted ion current (EIC) for the *m/z* 1125 ions. CID mass spectrum of the *m/z* 1125 ions at **b** 3.1 min and **c** 5.2 min retention time. **d** Proposed molecular structure for holothurinoside C2

previous study [9] that only considered only the body wall and Cuvierian tubule saponins.

LC-MS/MS is definitely of prime importance to make the distinction between isomeric saponins, whereas LC-(HR)MS allows rejecting isobaric compounds. Moreover, the molar ratio between isomeric saponins can be estimated on the basis of the LC-MS data by using the peak integrations (*vide infra*). For instance, we can estimate that holothurinoside C (5.2 min) and C2 (3.1 min) are roughly in a 2:1 ratio (Fig. 3a).

Data acquisition—ion mobility measurements

As the final step in the data acquisition phase, we will consider ion mobility as an additional tool for separating gaseous ionized saponins on the basis of their mobility [21]. In a recent report [22], several isomeric forms of the diterpene glycosides stevioside and rebaudioside A were successfully separated and characterized by ion mobility mass spectrometry. As already mentioned, in the present study, we will only envisage ion mobility for its capability of separating (ionized) saponins after the liquid chromatography and the ionization/desolvation steps and prior to the mass analysis. For instance, such an ion chromatography may allow separating saponins that are co-eluting in the liquid chromatography separation. The *m/z* 1125 ions from the gonad extracts are once again selected as a representative example for the LC-IMS-MS analysis.

In LC-MS, we previously identified two isomeric saponins, namely holothurinosides C and C2, with retention times at respectively 5.2 and 3.1 min (Fig. 3a). With ion mobility experiments, the corresponding *m/z* 1125 ions are characterized by single arrival times (drift time, t_D) at respectively 5.3 and 5.6 ms, see Fig. 4a, b, confirming that unique gas phase structures (or really close structures) are present. Interestingly, these two isomeric saponins, well-separated in LC, adopt quite similar gas phase structures upon ionization, since the

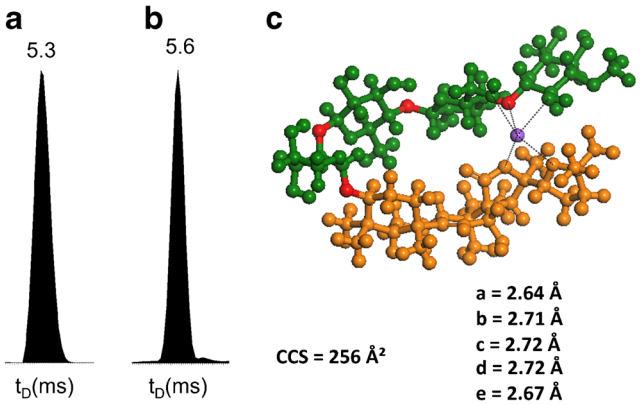

Fig. 4 LC-IMS-MS analysis of the saponins from the gonad extract of *H. forskali*. Arrival time distributions (ATD) for the *m/z* 1125 ions from holothurinoside C2 (**a**) and holothurinoside C (**b**). Lowest energy conformer of the Na^+ adduct of holothurinoside C provided by the simulations (the aglycone part is colored in *orange*, the monosaccharides in *green*, the glycoside bonds in *red*, and the cation in *purple*). As for a representative example, the averaged collision cross section (CCS) of the lowest energy conformer is calculated using the trajectory method implemented in the Mobcal software [21]

difference between the measured t_D is rather small, i.e., 0.3 ms. At this stage, we have first to remind that, upon electrospray, the saponin molecules are cationized by attachment of a Na^+ cation [15]. This positively charged molecular complex is stabilized by folding of the saponin molecule around the cation. We obtained the gas phase structure of sodium-cationized holothurinoside C by using successive quenched molecular dynamics simulations and did observe a folding of the saponin around the Na^+ ion, creating a U-shaped ion, see Fig. 4c. In this structure, the Na^+ cation is trapped in a pocket created between the aglycone and the oligosaccharidic chain. Such a folding will hide subtle structural differences and explains the small t_D difference between the isomeric holothurinosides C and C2 that are, on the other side, perfectly separated by LC in solution phase.

Another aspect that must be held into account when using ion mobility is the possibility to create different gas phase structures starting from a single molecular structure for the neutral molecule. In the present work, this was observed when submitting the *m/z* 1433 ions from the body wall extract to ion mobility. Upon LC-MS, it was confirmed that these *m/z* 1433 ions arise from a single molecule (or perfectly co-eluting molecules). However, upon ion mobility, these *m/z* 1433 ions are separated into two populations with t_D at 6.1 and 8.2 ms (Fig. 5a). This observation could be associated with two origins: (i) separation upon ion mobility of LC (perfectly) co-eluting (stereo)isomeric (same *m/z* ratio) saponins; or (ii) creation upon ESI of two different structures for a single molecule. The latter could arise for instance if two sites are available for the complexation of the cationizing particle, affording two newly termed cationers [30]. To solve this issue, we can submit the ion mobility-separated ions to collision-induced dissociation and compare their CID spectra. Such a complex experimental protocol is constituted by the following sequence: liquid chromatography, electrospray ionization, mass-selection of the parent ions (here *m/z* 1433), ion mobility separation, collision-induced dissociation, and mass analysis of the fragment ions. In the present work, we are using a Waters Synapt G2-Si HDMS mass spectrometer that is fully equipped with this elegant ion identification process. We observed that the CID spectra of the *m/z* 1433 ions (Fig. 5b) can be superimposed confirming that we are not dealing with isomeric saponins that would more than likely present (even slight) differences in their CID spectra. However, at this point of the discussion, we cannot claim that the *m/z* 1433 ion signal is not constituted by a mixture of LC co-eluting stereoisomeric saponins.

As already mentioned, two ion distributions in ion mobility experiments can also be associated with different ion structures for a single molecule. The *m/z* 1433 ions correspond to (ionized) holothurinoside F that is presented in Fig. 5c. This

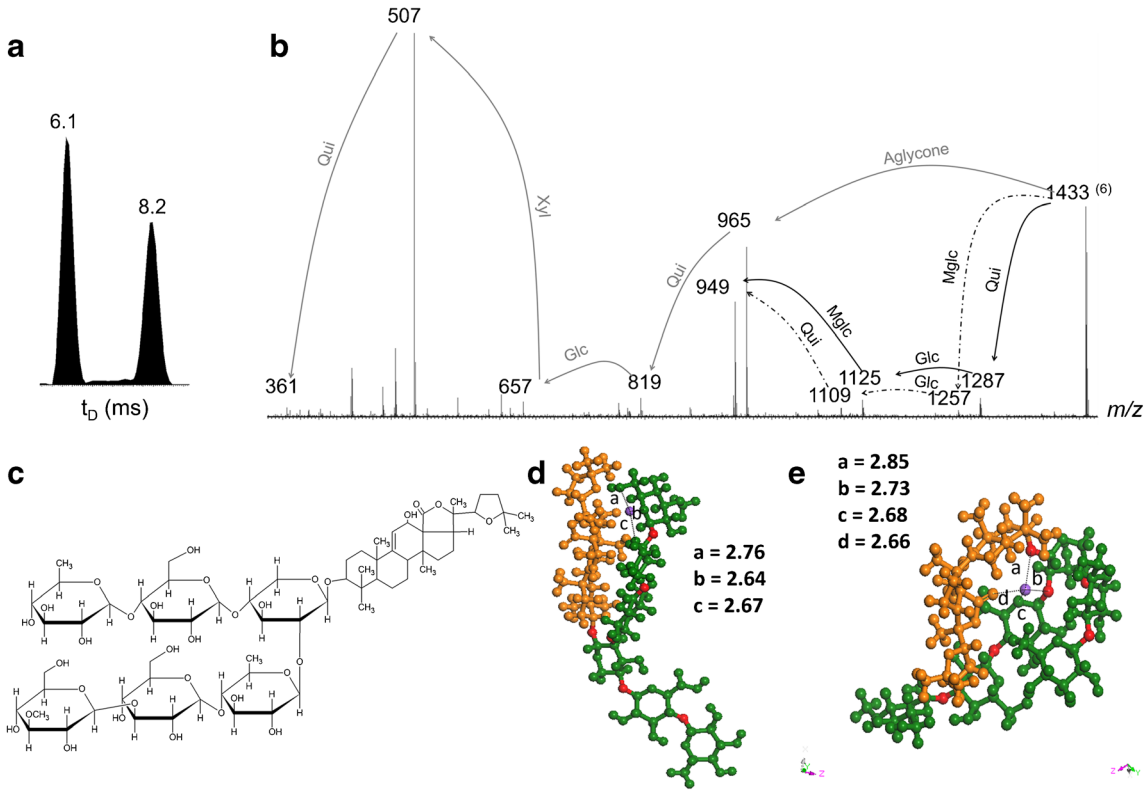

Fig. 5 LC-IMS-MS analysis of the saponins from the body wall extract of *H. forskali*. **a** Arrival time distribution (ATD) for the *m/z* 1433 ions. **b** CID spectra of the ion mobility-separated *m/z* 1433 ions. **c** Molecular structure of holothurinoside F. **d**, **e** Calculated lowest energy conformer of the Na^+ adduct of holothurinoside F. (the aglycone part is in colored in *orange*, the monosaccharides in *green*, the glycoside bonds in *red*, and the cation in *purple*)

six-sugar saponin presents a branched oligosaccharidic chain with respectively three and two monosaccharidic residues attached on the xylose residue. By performing quenched molecular dynamics simulations, we obtained two different structures for the *m/z* 1433 ions. In the most extended structure (Fig. 5d—likely to be associated with $t_D = 8.2$ ms), only one part of the oligosaccharidic chain—the three-sugar branch—is in interaction with the sodium ion, whereas in the most folded structure presented in Fig. 5e, the three parts of the saponin, i.e., the aglycone and both arms of the oligosaccharidic chain, are interacting with the sodium ion. This creates a more compact structure that is likely to be associated with the shortest t_D at 6.1 ms.

At this point of the work, when looking at the benefit of adding ion mobility in the saponin identification workflow, we must accept that hitherto ion mobility appears to be a less resolving technique than HPLC (and more for UHPLC, reasonably). In future works, we will optimize the ion mobility separation by (i) changing the nature of the cation; (ii) using doubly charged ions (that are observed, see ESM Fig. S4); and (iii) changing the nature of the collision gas [22, 31], with CO_2 instead of nitrogen for example. Indeed, being more polarizable, CO_2 changes the physico-chemistry behind the collision event by enhancing long distance interactions [22]. Moreover, we also observed, for the first time, that some IMS resolved signals can be generated from a single saponin that gives rise to different shaped ions. This could clearly appear as a drawback of IMS on the way to saponin analysis, especially if there is an ambiguity in the origin of the different ion mobility signals, i.e., (i) separation upon ion mobility of LC co-eluting isomeric saponins or (ii) creation upon ESI of two different structures for a single molecule. For the present paper, we will just consider ion mobility as an orthogonal separation tool and we will use it without further required optimizations that are under progress in our laboratory, including the development of a reliable calibration procedure.

In the following section of the paper, we will propose a way to combine, on a single graph, the data from the MALDI, LC-MS, and ion mobility experiments. We strongly believe that this will make the comparison of the saponin contents of different extracts easier. The association of all data (*m/z*, retention time, and drift time) will create the so-called fingerprint or signature of a given saponin.

Data integration

In the present study, we used different complementary mass spectrometry methods to obtain a quasi-complete description of the saponin molecules present within an organ extract. MALDI-ToF analyses (together with HRMS measurements) are used to identify the different saponin elemental compositions that are characterized by their *m/z* ratio and their relative abundances (as measured on the basis of peak intensities). For the calculation of relative abundances, we must take into account all isotopic signals (especially when comparing together four- and six-sugar saponins, given the increased number of carbon atoms). We already explained that we will mostly compare the saponins within each of the three families, namely the four-, five-, and six-sugar saponins, because the ion yield is likely to be dependent on the length of the oligosaccharidic chain, based on the expected difference in both ion formation and ion desorption. Such differences are better documented for electrospray ionization [32]. One could also consider comparing all saponins, regardless the number of monosaccharidic residues by the same approach. However, constraining the comparison inside one family already creates complex diagrams and extending the method to all saponins would lead to illegible and useless diagrams. As an intermediate solution, the relative proportion of each family calculated from MALDI analyses will be given to provide a complete overview of saponin distribution within each organ (see Fig. 7). At a second level of description, we will use the LC-MS data to characterize within each elemental composition (*m/z* ratio), the isomeric diversity. This will be done by using the retention time and the relative peak integration. Finally, for each of the retention time, we will use the ion mobility data (including the relative peak integration measured from the arrival time distribution) to identify co-eluting isomeric saponins or catiomers. All those data will be compiled within sector diagrams, one diagram for each saponin family (four to six monosaccharides) in a given organ extract. For example, we have integrated the data for the five-sugar saponins of the body wall extract (Fig. 2a) in the sector diagram presented in Fig. 6.

Based on the MALDI measurements, two elemental compositions are detected at *m/z* 1287 and 1303 with

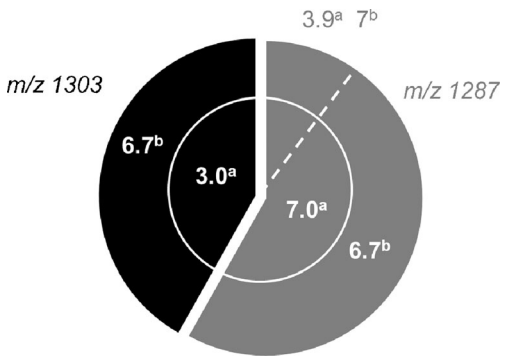

Fig. 6 Molecular diversity of the five-sugar saponins from the body wall extract of *H. forskali*. Integration of the MALDI-ToF, LC-MS/MS, and LC-IMS-MS data with [a] the retention time in LC (min) and [b] the drift time in ion mobility (ms). Section areas are associated with peak intensities in the MALDI mass spectrum, the peak integration in the extracted ion current chromatogram in LC (*retention time*), and peak integration in the arrival time distribution in ion mobility (*drift time*)

relative abundances of about 60 and 40%, respectively. These data are represented as the gray and black areas in the sector diagram. The areas are labeled according to the *m/z* ratio of the detected ions. In the diagram, the inner concentric circle represents the LC-MS data. Whereas the *m/z* 1303 ions are observed at only one retention time (3 min), two peaks are associated with *m/z* 1287 at 3.9 (15%) and 7 (85%) min. Finally, the ion mobility experiments confirm that each LC retention time corresponds to a single molecule since each retention time is associated with a single t_D.

We believe that this original representation will be easy to use when comparing saponin contents of different extracts. Indeed, from Fig. 6, we can immediately conclude that, within the detection limit of the mass spectrometer, the body wall extract of *H. forskali* contains three different five-sugar saponins, characterized by two different elemental compositions. This points to the presence of two isomers (for the *m/z* 1287 ions) and no co-eluting isomeric saponins are detected. The three molecules can be identified based on tandem mass spectrometry. In the present analysis, we identified holothurinoside A (*m/z* 1303–40% molar ratio) and holothurinoside E (*m/z* 1287—RT 7 min–51% molar ratio). The third saponin (*m/z* 1287—RT 3.9 min—9% molar ratio) is still not yet identified but we already know that it is an isomer of holothurinoside E, present in the same organ.

As already stated here above, all those data integration process can be achieved on the basis of another combination of MS methods. In particular, the MALDI-based quantitative analysis can be performed by merging, for a given elemental composition, the peak area in a LC-MS analysis.

Data comparison and analysis

In Fig. 7 (and ESM Fig. S3 for larger views), we compared the saponin contents of the three organs selected for the present study, namely the body wall (a-d-g), the gonads (b-e-h), and the Cuvierian tubules (c-f-i). As previously explained, the sector diagrams are each comparing the content of saponins (of a selected organ) presenting a specific number of monosaccharidic residues in the glycone moiety, i.e., the four-sugar (a-b-c), five-sugar (d-e-f), and six-sugar (g-h-i) saponins. The proportion of each saponin category within an organ is given by the percentages indicated next to the sector diagrams (read vertically for each organ).

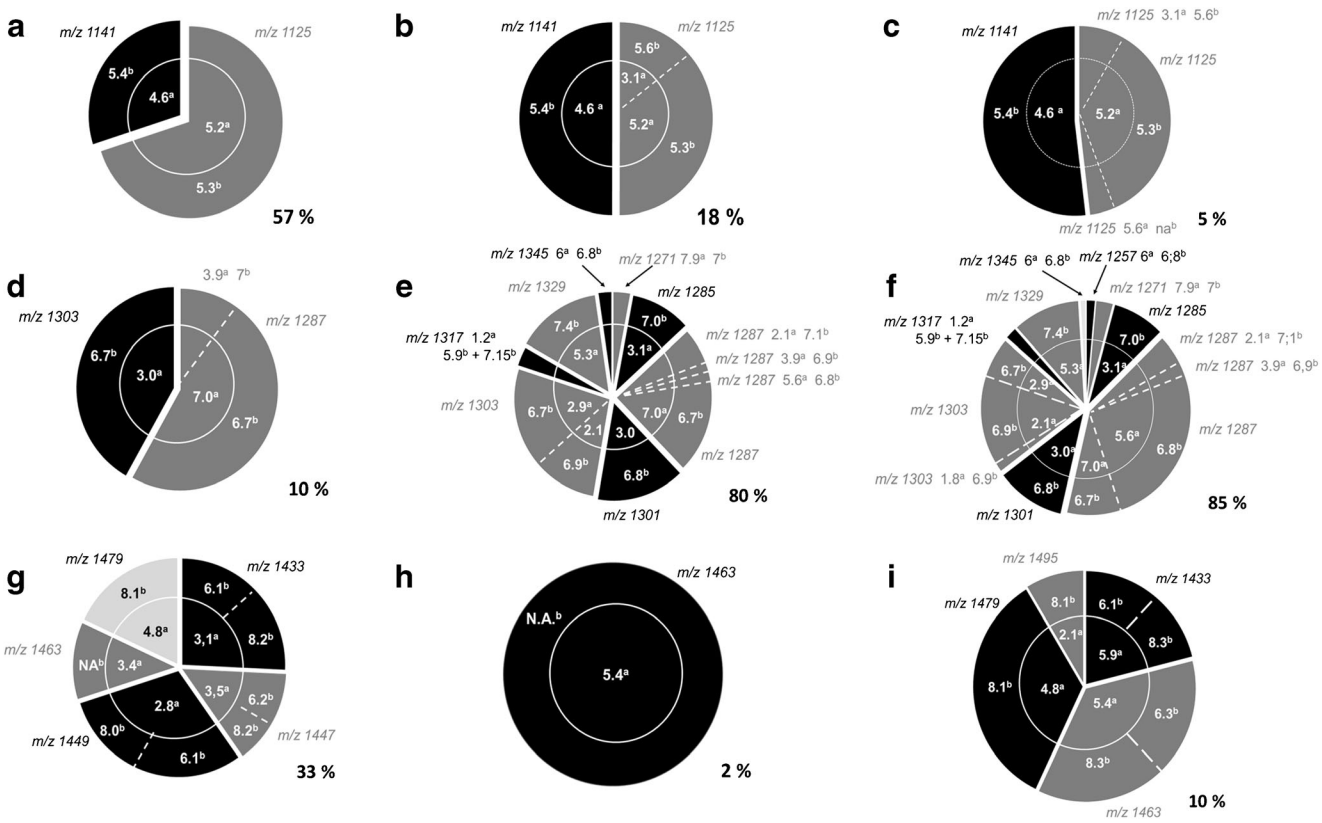

Fig. 7 Molecular diversity of saponins in the body wall (**a**-**d**-**g**), the gonads (**b**-**e**-**h**), and the Cuvierian tubules (**c**-**f**-**i**) of the sea cucumber *H. forskali*. Distribution of the four- (**a**-**b**-**c**), five-(**d**-**e**-**f**), and six-sugar (**g**-**h**-**i**) saponins in the different organs. Integration of the MALDI-ToF, LC-MS/MS, and LC-IMS-MS data with [a]the retention time in LC (min) and [b]the drift time in ion mobility (ms). *NA* not attributed. *Percentages* represent the proportion of each saponin category (for-, five-, and six-sugar saponins) within each organ (read vertically)

From this comparison, we can conclude the following:

(i) We detected at least 10, 16, and 22 different saponins within the body wall, gonads, and Cuvierian tubules of *H. forskali*. In addition to this difference in the total number of congeners identified, the body wall clearly differed from the two internal organs in the distribution of the four-, five-, and six-sugar saponins. In the body wall, the three categories of saponins are well represented and the six-sugar saponins showed the highest diversity, whereas in the gonads and Cuvierian tubules, the five-sugar saponins are the most abundant both quantitatively and in terms of diversity. Indeed, the five-sugar saponins within the internal organs are characterized by a huge diversity, with 12 to 14 different molecules including four and three isomers for the *m/z* 1287 and 1303 ions, respectively. On the other side, only one 6-sugar saponin, detected at *m/z* 1463, is detected within the gonad extract. Because of a really low S/N ratio, the t_D was not measured.

(ii) Each organ presented some specific saponins (see also ESM Table S1). Cuvierian tubules possess the highest number of specific congeners with five molecules covering the four-, five-, and six-sugar categories. In the body wall, we detected four specific congeners which interestingly all belong to the six-sugar saponins. Only one specific saponin was identified in the gonads. This five-sugar saponin, holothurinoside B, is one of the saponins previously described by Rodriguez et al. [23] but not detected in our previous study on the body wall and Cuvierian tubules [15].

(iii) Globally, for *H. forskali*, we characterized 26 different saponins, though 3 are still unidentified (see Table S2 in ESM). When compared to our previous report [15], we observed here eight additional molecules but three saponins previously detected, holothurinosides I and G_1 and desholothurin A_1, were not observed this time. This is probably related to the physiological state of the animals since we demonstrated that the saponin contents in animals, both for sea stars and sea cucumbers, is dependent on the sex, the season, and the stress state [15–20].

Saponin quantification

To provide a complete characterization of the saponin content of a given organ, information on the absolute concentration of each saponin within the extract are still missing. For the time being, we used indirect measurements such as the determination of the hemolytic activity of the extracts [15] to afford an estimation of the saponin content of an organ. We are well aware that such an estimation is far from being accurate and is definitively biased by various parameters. First, different saponins certainly present different hemolytic activities and since we are measuring the hemolytic activity of the saponin extract, we are averaging the activities over all saponins. Moreover, when preparing the saponin extracts, we cannot exclude the presence of contaminants which, if potentially active, could corrupt the calculation of the saponin concentration within the extract. To circumvent this problem, we are now isolating pure saponins to use them as internal standard, taking into account that using a single saponin or a limited set of saponins as internal standards will still remain problematic.

Conclusions

The data acquisition and integration strategy developed in the present report appears really efficient for saponin analysis in different organs. It is indeed possible to compare "at a glance" the saponin content of different organs. The association of exact mass measurement (*m/z*; either in MALDI or ESI), liquid chromatography (retention time), and ion mobility (drift time) appears really powerful for detecting new saponins, but also to confirm the presence of a known saponin within an extract. Indeed, the data set constituted by the *m/z* ratio, the retention time, and the drift time can be considered as the fingerprint of a given saponin, including the stereochemistry of the saccharidic group.

Acknowledgements The MS laboratory acknowledges the "Fonds de la Recherche Scientifique (FRS-FNRS)" for its contribution to the acquisition of the Waters QToF Premier Mass Spectrometer. P.F. and J.C. are Research Directors of the FRS-FNRS. C.D. is grateful to the F.R.I.A. for financial support. This work was supported by the FRFC research project no. T.0056.13, in part by the EU FP7-OCEAN Project "Low-toxic cost-efficient environment-friendly antifouling materials" (BYEFOULING) under Grant Agreement no. 612717, by the European Commission/Walloon Region (FEDER-BIORGEL project), and by the Interuniversity Attraction Pole program of the Belgian Federal Science Policy Office (PAI 7/05).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Vincken JP, Heng L, De Groot A, Gruppen H. Saponins, classification and occurrence in the plant kingdom. Phytochemistry. 2007;68:275–97.
2. Mackie AM, Turner AB. Partial characterization of biologically active steroid glycoside isolated from the starfish *Marthasterias glacialis*. Biochem J. 1970;117:543–50.
3. Kitagawa I, Kobayashi M. On the structure of the major saponin from the starfish *Acanthaster planci*. Tetrahedron Lett. 1977;10:859–62.

4. Nigrelli RF. The effect of holothurin on fish, and mice with sarcoma 180. Zoologica. 1952;37:89–90.
5. Yamanouchi T. On the poisonous substance contained in holothurians. Publ Seto Mar Biol Lab. 1955;4:183–203.
6. Kubanek J, Pawlik J, Eve T, Fenical W. Triterpene glycosides defend the Caribbean reef sponge *Erylus formosus* from predatory fishes. Mar Ecol Prog Ser. 2000;207:69–77.
7. Bahrami Y. Franco CMM acetylated triterpene glycosides and their biological activity from holothuroidea reported in the past six decades. Mar Drugs. 2016;14:147–85.
8. Podolak I, Galanty A, Sobolewska D. Saponins as cytotoxic agents: a review. Phytochem Rev. 2010;9:425–74.
9. Yang SW, Chan TM, Buevich A, Priestley T, Crona J, Reed J, et al. Novel steroidal saponins, Sch 725737 and Sch 725739, from a marine starfish, *Novodinia antillensis*. Bioorg Med Chem Lett. 2007;17:5543–7.
10. Mayo P, Mackie AM. Studies of avoidance reactions in several species of predatory British seastars (Echinodermata: Asteroidea). Mar Biol. 1976;38:41–9.
11. Harvey C, Garneau FX, Himmelman J. Chemodetection of predatory sea star *Leptasterias polaris* by whelk *Buccinum undatum*. Mar Ecol Prog Ser. 1987;40:79–86.
12. Caulier G, Flammang P, Gerbaux P, Eeckhaut I. When a repellent becomes an attractant: harmful saponins are kairomones attracting the symbiotic Harlequin crab. Sci Rep. 2013;2639.
13. Miyazaki S, Ichiba T, Reimer JD, Tanaka J. Chemoattraction of the pearlfish *Encheliophis vermicularis* to the sea cucumber Holothuria leucospilota. Chemoecology. 2014;24:121–6.
14. Caulier G, Van Dyck S, Gerbaux P, Eeckhaut I, Flammang P. Review of saponin diversity in sea cucumbers belonging to the family Holothuriidae. SPC Beche-de-mer Bulletin. 2011;31:48–54.
15. Van Dyck S, Gerbaux P, Flammang P. Elucidation of molecular diversity and body distribution of saponins in the sea cucumber *Holothuria forskali* (Echinodermata) by mass spectrometry. Comp Biochem Physiol B Biochem Mol Biol. 2009;152:124–34.
16. Van Dyck S, Caulier G, Gerbaux P, Fournier S, Todesco M, Wisztorski M, et al. The triterpene glycosides of *Holothuria forskali*: usefulness and efficiency as a chemical defense mechanism against predatory fish. J Exp Biol. 2011;214:1347–56.
17. Van Dyck S, Gerbaux P, Flammang P. Qualitative and quantitative study of saponin contents of five sea cucumbers of the Indian Ocean. Mar Drugs. 2010;8:173–89.
18. Van Dyck S, Flammang P, Meriaux C, Bonnel D, Salzet M, Fournier I, et al. Localization of secondary metabolites in marine invertebrates: contribution of MALDI MSI for the study of saponins in Cuvierian tubules of *H. forskali*. PLoS One. 2010;5:e13923.
19. Demeyer M, De Winter J, Caulier G, Eeckhaut I, Flammang P, Gerbaux P. Molecular diversity and body distribution of saponins in the sea star *Asterias rubens* by mass spectrometry. Comp Biochem Physiol B Biochem Mol Biol. 2014;168:1–11.
20. Demeyer M, Hennebert E, Wisztorski M, De Winter J, Caulier G, Eeckhaut I, et al. Inter- and intra-organ spatial distributions of sea star saponins by MALDI-imaging. Anal Bioanal Chem. 2015;407:8813–24.
21. Shvartsburg AA, Hudgins RR, Dugourd P, Jarrold MF. Structural elucidation of fullerene dimers by high-resolution ion mobility measurements and trajectory calculation simulations. J Phys Chem A. 1997;101:1684–8.
22. Bataglion GA, Souza GH, Heerdt G, Morgon NH, Dutra JD, Freire RO, et al. Separation of glycosidic cationers by TWIM-MS using CO_2 as a drift gas. J Mass Spectrom. 2015;50:336–43.
23. Rodrigez J, Castro R, Riguera R. Holothurinosides: new antitumour non sulphated triterpenoid glycosides from the sea cucumber *Holothuria forskali*. Tetrahedron. 1991;47:4753–62.
24. Mayo SL, Olafso BD, Goddard III WA. DREIDING: a generic force field for molecular simulations. J Phys Chem. 1990;94:8897–909.
25. MS Modeling Ver. 6.0.0.0, Accelrys Software Inc, San Diego, CA (USA). 2011.
26. Rapp AK, Casewit CJ, Colwell KS, Goddard WA, Skiff WM. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc. 1992;114:10024–35.
27. Caulier G, Mezali K, Soualili DL, Decroo C, Demeyer M, Eeckhaut I, et al. Chemical characterization of saponins contained in the body wall and the Cuvierian tubules of the sea cucumber *Holothuria* (*Platyperona*) *sanctori* (Delle Chiaje, 1823). Biochem Syst Ecol. 2016;68:119–27.
28. Bush MF, Campuzano IDG, Robinson CV. Ion mobility mass spectrometry of peptide ions: effects of drift gas and calibration strategies. Anal Chem. 2012;84:7124–30.
29. D'Atri V, Porrini M, Rosu F, Gabelica V. Linking molecular models with ion mobility experiments. Illustration with a rigid nucleic acid structure. J Mass Spectrom. 2015;50:711–26.
30. Tata A, Eberlin MN. Cationers and aniomers: unique classes of isomeric ions. Rapid Commun Mass Spectrom. 2016;30:1249–52.
31. Bleiholder C, Jonhson NR, Contreras S, Wyttenbach T, Bowers MT. Molecular structures and ion mobility cross sections: analysis of the effects of He and N_2 buffer gas. Anal Chem. 2015;87:7196–203.
32. Leize E, Jaffrezic A, Van Dorsselaer A. Correlation between solvation energies and electrospray mass spectrometric response factors. Study by electrospray mass spectrometry of supramolecular complexes in thermodynamic equilibrium in solution. J Mass Spectrom. 1996;31:537–44.