$QCMS^2$ as a new method for providing insight into peptide fragmentation: The influence of the side-chain and inter-side-chain interactions

Blockhuys Frank (Orcid ID: 0000-0002-2201-6682)

QCMS^2 AS A NEW METHOD FOR PROVIDING INSIGHT INTO PEPTIDE FRAGMENTATION: THE INFLUENCE OF THE SIDE CHAIN AND INTER-SIDE-CHAIN INTERACTIONS

Julie Cautereels, Nils Van Hee, Sneha Chatterjee, Christian Van Alsenoy, Filip Lemière and Frank Blockhuys*

Department of Chemistry, University of Antwerp,
Groenenborgerlaan 171, B-2020 Antwerp, Belgium

Abstract

The identification of peptides and proteins from tandem mass spectra is a difficult task and multiple tools have been developed to aid this identification. We present a new method, called Quantum Chemical Mass Spectrometry for Materials Science (QCMS^2), which is based on quantum chemical calculations of bond orders, reaction and transition-state energies at the DFT/B3LYP/6-311+G* level of theory. The method was used to describe the fragmentation pathways of five X-His-Ser tripeptides with X = Asn, Asp, Glu, Ser and Trp, thereby focusing on the influence of the side chain and inter-side-chain interaction on the fragmentation. The main features in the mass spectra of the five tripeptides were correctly reproduced and a number of fragments were assigned to fragmentations involving the side chain and the influence of inter-side-chain interactions. Product ion spectra were recorded to evaluate the capabilities and limitations of QCMS^2 and a number of conventional tools.

*Corresponding author, frank.blockhuys@uantwerpen.be.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jms.4446

Keywords: Quantum Chemical Mass Spectrometry ($QCMS^2$), Density Functional Theory, fragmentation pathways, MS/MS, tripeptides.

1. Introduction

Tandem mass spectrometry (MS/MS) is one of the most effective tools for peptide and protein analysis due to its speed, sensitivity and versatility [1]. The analysis of peptide fragmentation in tandem mass spectrometry has, therefore, become the method of choice for protein identification: the amino acid sequence of peptides can be determined based on their fragmentation patterns and the final protein identification can be achieved using one of a variety of tools [2,3]. Among these tools, database search engines represent a broad class and were the first to be widely adopted; SEQUEST [4] and Mascot [5] are the most commonly used [6]. In general, such tools search protein sequence databases to match the unknown peptide to peptide sequences present in the databases. Based on the sequences of the matched peptides, fragment ions of the candidates are predicted and a match is evaluated using pre-defined scoring criteria related to the coincidence of the masses of precursor and fragment ions: logically, the match with the highest score is most likely the unknown peptide. SEQUEST, Mascot and many other database search engines just differ in the scoring criteria to rank the peptide matches [2,6,7]. Although database search engines are immensely popular, studies evaluating their performance have shown that 40-70% of high-signal/noise MS/MS spectra cannot be matched to predicted protein spectra or are even misidentified mostly due to the noisiness of spectra and the presence of unexpected post-translational modifications [6-12]. In this respect, ProteinProspector [13] presents itself as a promising alternative: it is a more recent, freely available web-based database search tool containing sequence database search programs and peptide/protein MS utility programs [14]. However, the major limitation remains that database search engines require the unknown peptide/protein to be available in the database since otherwise identification is impossible.

In response to the limitations of database search engines, *de novo* sequencing methods are emerging to aid identification of peptides/proteins from tandem mass spectra: they determine the peptide sequence directly from the MS/MS spectrum without the use of a database [6,15]. The algorithms for *de novo* sequencing assign (abundant) peaks in an MS/MS spectrum to certain fragments and identify series of the same fragment. The mass difference between two consecutive fragment ions can then be assigned to an amino acid sequence. In that manner and

by combining the information of different ion types, the unknown peptide/protein can be identified [16]. Multiple peptides can be identified by the algorithms as the unknown and they are ranked based on scoring schemes. Multiple *de novo* sequencing software packages, differing in the kind of algorithm used, are available of which PEAKS [17] is one of the most popular and performs better than the other packages [16,18]. Novor [19] also seems to be a promising emerging tool as it benefits from the developments of its predecessor PEAKS and is freely available [18]. Despite the usefulness of *de novo* sequencing methods they too are not without limitations: reviews and performance evaluation studies have shown that although more than 60% of amino acid residues can be correctly predicted by a number of algorithms under specific circumstances, only 30% of the peptides are correctly identified [16,18,20-23]. This is so mainly because identification by *de novo* sequencing methods requires high-quality mass spectra which can only be recorded with highly accurate mass spectrometers [18,20-23].

Alternatives for the above mentioned approaches are tools such as data-driven approaches and the kinetic model of Zhang [24]. Data-driven approaches predict spectra based on the fragmentation rules they learn from experimental tandem mass spectra using, for example, a neural network and the probabilities of those rules. Among these approaches, both PeptideART [25] and MS2PIP [26] are freely available and able to predict peak intensities; the latter is important since it has been shown that incorporating knowledge about peak intensities and amino acid composition significantly improves peptide identification rates [26-29]. The model of Zhang, on the other hand, is a kinetic model based on the Mobile Proton Model (MPM) to predict fragmentation spectra acquired on a quadrupole ion trap mass spectrometer: it includes most fragmentation pathways described in the literature plus a number of additional pathways based on the author's observations [24].

Although all these tools have been a positive force in the identification of proteins from tandem mass spectra, their common major disadvantage is that they are all limited by what is already known about the fragmentation of peptides. In any case, the complete set of fragmentations of any given peptide is never *a priori* known since peptide fragmentation in general has not been fully understood. Consequently, this lack of fully understanding peptide fragmentation results in a number of unexplained peaks in an MS/MS spectrum that are not used by the current tools which logically lead to failed and erroneous identifications. In addition, these tools only identify proteins by assigning fragments, without providing any details on the origin of each fragment and will, therefore, never be able to provide any new insights into the fragmentation mechanisms. All this leads to the conclusion that any (new) method allowing to obtain more

detailed knowledge about the fragmentation mechanisms of peptides will result in the assignment of each fragment in an MS/MS spectrum with supplementary information about their origin. This information is necessary to improve the search algorithms mentioned above since incorporation of the additional knowledge will result in fewer misidentifications and will, from there, lead to more reliable protein identification.

Some time ago, we developed a new and generally applicable *ab initio* method for the prediction and description of fragmentation routes named Quantum Chemical Mass Spectrometry for Materials Science ($QCMS^2$) [30]. Based on straightforward calculations of bond orders and energies of fragments, $QCMS^2$ was used to describe the fragmentation pathways leading to the main features in the electron ionization (EI) mass spectra of a number of organic compounds, generating considerable insight into these pathways and a clearer understanding of the provenance of peaks in the experimental spectra. More importantly, though, the method led to the discovery of new fragmentation mechanisms, which were confirmed by MS/MS measurements, and the confirmation of empirical rules from the canon of MS.

Fueled by these successes, we decided to expand the scope to beyond EI and apply $QCMS^2$ to the electrospray ionization (ESI) fragmentation pathways of tripeptides under collision-induced dissociation (CID) conditions. The focus hereby lies on the fragmentation of the side chain and the influence of hydrogen bonding in or between the side chains on the fragmentation, since a number of the peaks in the spectrum not predicted/assigned by the above mentioned tools are known to be due to fragmentations involving the side chain(s). In this paper, $QCMS^2$ is applied to the description of the ESI CID fragmentation pathways of five X-His-Ser tripeptides with X = Asn, Asp, Glu, Ser and Trp. These amino acids were chosen based on the capability of the side chain to form strong hydrogen bonds in order to explore the influence of the side chain on the fragmentation. In addition, the product ion spectra of these tripeptides were recorded: $QCMS^2$ produces fragmentation pathways leading to not only the fragments resulting from backbone fragmentations but also to those due to fragmentations of the side chain and inter-side-chain (ISC) interactions. Capabilities and limitations of both $QCMS^2$ and the well-known tools Mascot, ProteinProspector, PEAKS, and PeptideART in this context are discussed.

2. Experimental and computational details

2.1. Materials

The tripeptides DHS, EHS, NHS, SHS and WHS were purchased from Biomatik with purities of 99.1, 96.3, 99.6, 99.0 and 95.1%, respectively. The samples were prepared by adding 1 mg to a 1 mL 50/50 H_2O/CH_3OH solution and then the mixture was diluted to a 6 μM solution (50/50 H_2O/CH_3OH).

2.2. ESI Q-TOF MS/MS measurements

Electrospray mass spectra were recorded on a Q-TOF II mass spectrometer (Micromass, Manchester, UK) in positive ion mode. Capillary voltage was set to 1.60 kV for DHS and SHS and to 1.20 kV for NHS, SHS and WHS with the extractor at 10 V. The source temperature was set to 200 °C. The low-energy CID product ion scans were recorded using a collision energy ranging from 5 to 30 eV at intervals of 5 eV. The mass range was *m*/*z* 50-500. Approximately 400 scans were summed for each of the experimental conditions. Leucine-enkephalin was used to calibrate the instrument in positive ion mode. The instrument was controlled, and the data processed using Masslynx V4.1 software [31].

2.3. Details of the other tools

Mascot, PEAKS 8, Protein Prospector v5.20.00 and PeptideART v2.1 were used. For Mascot, the online version was used, namely the MS/MS ion search with all databases possible. A trial version of PEAKS was used with the following settings: the calculations were performed on the spectrum with a 15 V cone voltage and 25 eV collision energy because lower collision energies yielded too few fragments to identify the peptide correctly. "Data refinement correct precursor mass only" was selected, and "*de novo* calculations" and "precursor tolerance" were set to "9.0 min merge scans" and 0.5, respectively. For ProteinProspector the Peptide Utility Program MS-Product was used: it allows the calculation of all fragment ions resulting from the fragmentation of a peptide based on its sequence whereby the instrument type can be specified.

2.4. Computational details

Conformational analyses were performed using the simulated annealing procedure [32] implemented in AMPAC 10.1.9 [33] with the semi-empirical AM1 method [34]. Quantum chemical calculations, *i.e.*, the calculation of the molecular electrostatic potential (MEP) and the geometry optimizations of the neutral and protonated tripeptides were performed using the Gaussian 09 suite of programs [35], at the level of Density Functional Theory (DFT), using the B3LYP

functional [36], applying both the restricted and unrestricted formalisms, and the 6-311+G* basis set [37]; the functional and basis set were used as they are implemented in the program. Harmonic frequency calculations were performed to verify that the resulting structures are minima on the Potential-Energy Surface (PES). The QCMS^2 method is automated and has been implemented into the BRABO [38] and STOCK [39] software packages; the calculations of the fragmentation pathways were likewise performed at the DFT/B3LYP/6-311+G* level of theory. Bond orders were calculated using the Fractional Occupation Hirshfeld Iterative (FOHI) formalism [40], which is based on the Hirshfeld partitioning of the electron density [41-47]. Transition-state energies for the rearrangements were calculated using the Synchronous Transit-Guided Quasi-Newton (STQN) method (QST3) [48,49] implemented in Gaussian 09.

3. Results and discussion

In the following, the procedure involved in QCMS^2 and the comparison of the different tools are described. The two steps constituting the procedure when applied to the fragmentations of tripeptides are presented in Figure 1; in the following sections the ionization step and the fragmentation step will be discussed separately.

3.1. Ionization step

The ionization step naturally depends on the precise ionization technique which is used in the experiment. For ESI, the ionization technique used for peptides here, the ionization step has been further subdivided into three substeps, *i.e.*, two conformational analyses separated by the determination of the protonation site(s) (Figure 1).

(1) The first conformational analysis is performed on the neutral tripeptide and only takes one single conformer – the most stable – into account for determining the protonation site(s): other (stable) conformers will be considered in the second conformational analysis. The total energy of this conformer is obtained from a geometry optimization.

(2) The following methodology was used to determine the different protomer(s) of the tripeptide. The molecular electrostatic potential (MEP) was calculated for the most stable conformer obtained in step (1), which shows the preferred sites for protonation; this logically leads to a number of possible protomers, the total energies of which are obtained from a geometry optimization [50]. More protomers than only the most energy-favorable one have to be taken into account if the Boltzmann distribution (at the temperature of the ionization source, *i.e.*, 473-523 K) of the possible protomers reveals that the ratio of two protomers is higher than the dynamic range [51,52] of the mass spectrometer in a single spectrum (3%): this means that the mass

spectrometer is sensitive enough to differentiate between the fragments of two protomers. This leads to an energy cut-off of 15.25 kJ.mol^{-1}. The data in Table 1 show that for the five studied tripeptides the imidazole ring of the central histidine is always the most stable protonation site; taking the energy cut-off in account, only this protomer needs to be considered. Note that this approach also results in the reduction of the number of calculations.

(3) In the second conformational analysis, performed on the protonated tripeptide(s), only the lowest-energy conformers are taken into account, *i.e.*, only the conformers up to the largest jump in relative energy (see Figures S1-S5 in the Supporting Information). Conformers that only differ in rotation around the bond adjacent to an X-H bond, whereby the rotation does not result in a change in intramolecular interactions, are omitted from this selection.

The above leads to two conformers being considered for Ser-His-Ser (SHS) and Trp-His-Ser (WHS), four for Asn-His-Ser (NHS) and Glu-His-Ser (EHS), and five for Asp-His-Ser (DHS). The two conformers of SHS differ in the position of the N-terminal serine residue and the interaction between the imidazole moiety of histidine and the C-terminal carboxyl group (Figure 2). The conformers of DHS, EHS, NHS and WHS are presented in Figures S6-S9 in the Supporting Information.

3.2. Fragmentation steps

Before discussing the fragmentation pathways, the nomenclature of different major fragments along the backbone is clarified: the ions are classed systematically depending on which bond in the backbone breaks (Figure 3) [53,54]. The ion is classified as either a, b or c when the charge is retained on the N-terminal fragment and as x, y or z when the charge is retained on the C-terminal fragment. The subscript indicates the number of amino acid residues in the fragment. The observed fragments and their abundance depend on many elements including the amino acid composition, instrumentation, dissociation technique *etc.* [55,56]. In this article, the ESI CID product ion spectra are recorded on a quadrupole-time-of-flight (Q-TOF) mass spectrometer and, therefore, a-, b- and y-ions are expected to be present in the mass spectra of the five selected tripeptides.

The fragmentation of the tripeptides is simulated using QCMS^2. To the three rules of the original method [(1a), (2) and (3) below] designed to reduce the number of calculations [30], three are added which are specific for peptides [(1b), (4) and (5)].

(1a) The weakest bond breaks first, *i.e.*, the bond with the lowest bond order. If other, stronger bonds in the molecule or fragments have a bond order not more than 0.1000 a.u. higher than that of the weakest, these will also be broken. The process of breaking bonds is repeated

throughout the different charged fragments formed during fragmentation until a fragment with a molecular mass of less than 50 Da is obtained; *m*/*z* 50 is the lowest limit of the mass range of the experimental product ion spectra. Any fragmentation pathway is discontinued when the lowest-energy fragmentation requires more than 500 $kJ.mol^{-1}$.

(1b) If X–H bonds are among the weakest, they will not be considered since our observations reveal that breaking them requires too much energy.

(2) In turn, each bond selected on the basis of Rule 1 can be broken in three different ways: via homolytic cleavage, via heterolytic cleavage or via a rearrangement. Each bond can be cleaved homolytically in one or heterolytically in two ways, taking into account that in the latter the negative charge can be placed on either fragment. If the number of consecutive bonds in the precursor ion or fragments permits, each bond can also be broken via one or two 1,4-rearrangements or via a McLafferty rearrangement. The bond cleavage reactions are evaluated based on their associated ΔE whereas the rearrangements are evaluated based on their associated $\Delta E^{\ddagger}$. Thus, each fragmentation step has three, four, five or six possible pathways: for the precursor ion and each fragment the ΔE and $\Delta E^{\ddagger}$ values of all bonds considered are compared and the lowest-energy pathway is followed. However, if one or more other fragmentation mechanisms have (a) value(s) of ΔE or $\Delta E^{\ddagger}$ not more than 200 $kJ.mol^{-1}$ higher than the lowest-energy pathway, this/these additional pathway(s) is/are also considered.

(3) Only McLafferty rearrangements and 1,4-rearrangments with hydrogen are considered. Non-hydrogen rearrangements are not taken into account [57].

(4) Three computationally verified peptide fragmentation mechanisms have been directly implemented into the QCMS^2 code in the form of additional subroutines: (i) the a_1-y_x fragmentation pathway [55,58], (ii) the C-terminal residue exclusion mechanism generating the $b_{(n-1)}$ + H_2O ion [59,60], and (iii) the MPM mechanism for Arg, His, and Lys, and the MPM when there is no basic residue [55].

(5) When interactions occur between a side chain and the backbone, between two side chains (ISC) or within the backbone, the proton involved in the interaction can be transferred intramolecularly. The resulting fragment is included in the calculation of the fragmentation routes. Note that the values for the bond order and energy difference ranges can be set in function of the system under investigation: the values chosen here work well for organic compounds including tripeptides [30].

In the following, the fragmentation pathways of SHS will be discussed in detail first. Then, the general trends observed in the fragmentation routes of all five tripeptides will be presented.

3.2.1. Fragmentation pathways of SHS

The major fragmentation pathways of both conformers considered for SHS will be discussed separately. The bonds considered for fragmentation for both conformers are numbered 1-6 in Figure 2, all the predicted fragments by QCMS[2] are displayed in Table 2 and a number of the fragmentation pathways are presented in Figure 4; it should be noted that for each fragment formed in Table 2 an f+1 isotope peak can be expected.

First-generation fragments

For the lowest-energy conformer (Figure 2, left) the weakest bond which are considered for fragmentation are bonds 1 (BO 0.2893 a.u.), 2 (BO 0.2792 a.u.), 4 (BO 0.2888 a.u.), 5 (BO 0.3431 a.u.), and 6 (BO 0.3480 a.u.). The lowest-energy fragmentation routes are the homolytic cleavages of bonds 1, 4, 5, and 6, the heterolytic cleavage of bond 2, and the 1,4-rearrangment of bond 2, leading to *m*/*z* 270, 285, 299, 299, 132 and 197, respectively (Table 2). This 1,4-rearrangment results in the formation of the a_2-ion (*m*/*z* 197), which can also be formed as a result of the loss of CO from the b_2-ion.

In addition, bond 3 is broken in the MPM mechanism resulting in the formation of the b_2- (*m*/*z* 225) and y_1-ions (*m*/*z* 106). The y_2-ion (*m*/*z* 243) is formed as a result of N-terminal amide bond cleavage in the a_1-y_x fragmentation mechanism [55,58]. Furthermore, exclusion of the C-terminal Ser residue results in the formation of the b_2 + H_2O-ion, which also has *m*/*z* 243 [59,60]. The loss of NH_3 leading to *m*/*z* 313 is due to a 1,4-rearrangment involving the Ser side chain. Consequently, there are ten charged fragments in the first generation as a result of direct bond cleavages in the precursor ion. An eleventh (*m*/*z* 312) is obtained from the loss of water from the precursor ion: the interaction between the backbone NH of His and the hydroxyl group of the side chain of the N-terminal Ser causes a proton to be transferred to the hydroxyl group with subsequent loss of water (dotted line in Figure 2, left).

Second-, third- and fourth-generation fragments

Whereas *m*/*z* 313 and *m*/*z* 312 only display the loss of water as an energy-favorable fragmentation route according to Rule 1, the other ten fragments of the first generation have other energy-favorable routes. In those ten fragments, the weakest bonds are the same carbon-carbon bonds as the ones in the precursor ion which are still present in the fragments. As one may expect, breaking all these bonds in all possible manners yields many low-energy fragmentation pathways resulting in a considerable number of fragments (Table 2). Therefore, we limit our-

selves here to discussing the formation of the a_1-ion (*m*/*z* 60), the fragmentation pathways involving fragments commonly observed with His (*m*/*z* 166, 138 and 110), and the loss of small molecules from the a_2- and b_2-ions in detail.

The a_1-ion (*m*/*z* 60) results from the heterolytic cleavage of bond 1 in both the b_2-ion (*m*/*z* 225) and the fragments resulting from the homolytic cleavage of bond 4 (*m*/*z* 285) and 6 (*m*/*z* 299) from the precursor ion. *m*/*z* 166 is the result of a 1,4-rearrangement of bond 1 in the b_2-ion (Figure 4). *m*/*z* 138 is formed through the consecutive homolytic cleavages of bonds 1 and 2 or via a 1,4-rearrangment from the a_2-ion; in the former, the loss of CO from *m*/*z* 138 generates *m*/*z* 110. *m*/*z* 138 is classified as the $HisH^+$-ion and *m*/*z* 110 as the immonium ion of His according to the literature (Figure 4) [61,62]. In both the a_2- and b_2-ions, multiple losses of small molecules are observed (Table 2). The loss of CH_2O is the result of a McLafferty rearrangement involving the side chain of Ser leading to $a_2 - CH_2O$ (*m*/*z* 167, Figure 4) and $b_2 - CH_2O$ (*m*/*z* 195). The loss of NH_3 is the result of a 1,4-rearrangment involving the Ser side chain leading to $a_2 - NH_3$ (*m*/*z* 180) and $b_2 - NH_3$ (*m*/*z* 208). The loss of H_2O is the result of the interaction between the NH in the backbone and the side chain of Ser leading to $a_2 - H_2O$ (*m*/*z* 179) and $b_2 - H_2O$ (*m*/*z* 207). The loss of both H_2O and NH_3 is the result of the combination of the latter two fragmentation pathways: in *m*/*z* 208 and 180 H_2O is lost and in *m*/*z* 207 and 197 NH_3 is lost both pathways leading to $a_2 - H_2O - NH_3$ (*m*/*z* 162) and $b_2 - H_2O - NH_3$ (*m*/*z* 190). In addition, in $a_2 - H_2O$ the loss of NH_3 can also proceed due to an interaction between the N-terminal amine group and the imidazole moiety of His. This fragmentation is not predicted for the b_2-ion because the required interaction is prevented by the rigidity of the bicyclic system present in the b_2-ion.

Second lowest-energy conformer

The analysis of the fragmentation routes of the second lowest-energy conformer (Figure 2, right) reveals that the same bonds are broken as in the lowest-energy conformer and that the same fragmentation pathways are followed, except for the loss of H_2O from the precursor ion which generates *m*/*z* 312 in a different way: here the interaction between the backbone NH of the C-terminal Ser and the hydroxyl group of the side chain of the N-terminal Ser leads to *m*/*z* 312 (Figure 5, 1st fragmentation route). Also, there are four additional fragmentations (Figure 5). In the second lowest-energy conformer, H_2O is also lost due to an ISC interaction between the side chains of both Ser moieties; in this pathway the C-terminal Ser loses H_2O leading to an additional structure for *m*/*z* 312. Whereas the latter *m*/*z* 312 loses NH_3 via a 1,4-rearrangment involving the Ser side chain, the former *m*/*z* 312 loses NH_3 via a new interaction arising between the hydroxyl group of the C-terminal Ser and the N-terminal amine group, both leading

to *m*/*z* 295. In the fourth additional fragmentation the latter *m*/*z* 312 loses CH_2O via a heterolytic cleavage involving the side chain of the N-terminal Ser.

Thus, the prediction of the fragmentation routes of both SHS conformers using QCMS^2 leads to the 68 fragments which are given in Table 2. It is, however, important to note that it is quite unlikely that all of these will be observed in a given experimental spectrum, taking into account that the appearance of a mass spectrum depends to a great extent on the conditions under which it is recorded [51,63,64]. A number of predicted fragments could turn out to be low-abundant species under the precise conditions of the experimental measurement: for example, cooling of ions on their way to the mass analyzer leads to a substantial loss of energy for these ions resulting in reduced fragmentation, and labile ions, which are more likely to experience secondary fragmentation, are also less likely to be detected.

3.2.2. Fragmentation pathways of DHS, EHS, NHS and WHS involving the side chain

From the previous it can be concluded that for SHS fragmentations of the side chains themselves as well as fragmentations due to the ISC interactions are observed. Since the focus of QCMS^2 lies on those fragmentations, only these will be discussed for DHS, EHS, NHS and WHS.

These four tripeptides all display fragmentations in which weak bonds in the side chains of the peripheral amino acids are broken; fragmentation in the His side chain is never observed. For Trp these fragmentations involve breaking the carbon-carbon bonds to the backbone leading to the loss of the entire side chain as is similar to the fragmentations of Ser. On the contrary, for Asn, Asp and Glu these fragmentations involve breaking the carbon-carbon bonds adjacent to the functional groups leading to partial loss of the side chain.

Also, a number of conformers of these four tripeptides display fragmentations due to ISC interactions in both the precursor ion and the resulting fragments. As in SHS, WHS displays ISC interactions between the two peripheral amino acids leading to the loss of H_2O (Figure S9). On the other hand, DHS, EHS and NHS display ISC interactions between the N-terminal amino acid and His (Figures S6, S7 and S8): due to this interaction the proton of His can be transferred to the carbonyl oxygens of D, E and N after which the adjacent carbon-carbon bond can be easily broken resulting in the loss of CH_2=C(OH)NH_2 for NHS and the loss of CH_2=C(OH)OH for both DHS and EHS. Furthermore, in EHS there is an interaction between the imidazole moiety of His and the hydroxyl group of Glu resulting in the loss of water. For all four tripeptides there are also interactions between the side chain and the backbone leading to the loss of NH_3 and H_2O: these fragmentations proceed in the same manner as for SHS.

3.3. Assignments by the different tools

The product ion spectra recorded at a sampling cone voltage of 15 V with the six different collision energies were combined. Those signals of which the relative abundance was equal to or higher than 1.0% of the intensity of the base peak (the precursor ion) were used in the comparison; the experimental spectra of the five XHS tripeptides are collected in the Supporting Information (Figures S10, S11, S12, S13 and S14). Mascot was unable to identify the five tripeptides because they were not present in its (online) databases. Also, PEAKS was not able to unambiguously identify the peptides: one of the suggested structures was always the correct tripeptide, but also those with an incorrect amino acid sequence were suggested as a match. For QCMS^2 the fragmentations of all considered conformers were combined. In the following, the comparison of the assignments of SHS given in Table 3 is discussed in detail; the relevant data for DHS, EHS, NHS and WHS can be found in the Supporting Information (Tables S1, S2, S3 and S4).

24 fragments were identified in the combined spectrum. Since the backbone fragmentations of peptides are well known, it is obvious that all tools are more or less able to assign the fragments resulting from these, *i.e.*, *m*/*z* 243, 226, 225, 197, 166, 156, 138 and 110 (Table 3). Only PeptideArt is not able to assign *m*/*z* 225 and 197 as the b_2- and a_2-ions, respectively. However, only PeptideArt and QCMS^2 take into account isotopes since PeptideART is able to assign *m*/*z* 226 as the b_2 isotope peak and QCMS^2 is able to assign *m*/*z* 209 as an isotope peak. In the literature it is stated that the a_2-ion is formed directly from the b_2-ion by the loss of CO; QCMS^2 suggests that it is also formed directly from the precursor ion by a 1,4-rearrangment. Only QCMS^2 and ProteinProspector assign *m*/*z* 243 as both the y_2 and $b_2 + H_2O$ ions, but only QCMS^2 is able to assign *m*/*z* 156 as the y_1-ion of the $b_2 + H_2O$ ion. According to the literature, the other backbone fragments *m*/*z* 166, 138 and 110 are characteristic ions of His but ProteinProspector, PEAKS and PeptideART are not able to assign all of them.[61,62] QCMS^2 does predict the formation of these ions, as discussed above (Figure 4).

Considering the fact that QCMS^2 incorporates fragmentations of the side chains and fragmentations due to ISC interactions, it logically assigns a larger number of fragments related to these than the other tools. QCMS^2 finds thirteen fragments related to the side chains, *i.e.*, *m*/*z* 313, 312, 295, 226, 225, 208, 207, 195, 190, 180, 179, 167 and 162 (Table 3), whereas the other tools find no more than eight (*m*/*z* 313, 312, 226, 225, 208, 207, 180 and 179) and then only based on mass differences upon the loss of H_2O or NH_3. Note that ProteinProspector is able to assign the loss of NH_3 only if this functionality is present in the side chain. Only QCMS^2 is

able to assign *m*/*z* 208 as both the $b_2 - NH_3$ and $y_2 - H_2O - NH_3$ ions. PEAKS and QCMS^2 are the only tools able to predict the loss of H_2O and NH_3 from the a_2-ion (*m*/*z* 180 and 179). In addition, only QCMS^2 is able to predict both the loss of water and ammonia from the precursor, b_2- and a_2-ions (*m*/*z* 295, 190 and 162). Furthermore, since the loss of CH_2O from the b_2- and a_2-ions proceeds via a more complex mechanism than the loss of H_2O and NH_3, only QCMS^2 can assign the related peaks (*m*/*z* 195 and 167).

m/*z* 152, 150 and 122 are not assigned by any of the tools, including QCMS^2. For the latter this is most likely due to the too strict application of Rules 1 and 2: adjusting the bond order and energy cut-off values could result in the assignments of these missing fragments but would drastically increase the computational time. At this moment, it is unclear how these fragments are formed.

Overall, for SHS, QCMS^2 assigns 88% of the ions in the product ion spectra whereas the next best tool, PEAKS, assigns only 42%, and ProteinProspector and PeptideART both no more than 33% (Figure 6). Moreover, QCMS^2 produces full fragmentation pathways and provides detailed insight into fragmentation mechanisms. Similar observations regarding the backbone and side chain fragments, the fragments due to ISC interactions and the non-predicted fragments can be made for DHS, EHS, NHS and WHS and the associated assignment percentages are also presented in Figure 6.

An obvious criticism of this approach is that ProteinProspector, PEAKS, and PeptideART have not been developed to be used this way and that the above comparison is fundamentally wrong. Even though this is clearly true, it does allow the identification of some of the deficiencies of these popular tools and may present an opportunity to remedy some of them. Indeed, when just two of the most straightforward assignments provided by QCMS^2, *i.e.*, (1) the consecutive loss of NH_3 and H_2O and (2) the loss of small molecules such as CH_2O, were to be incorporated into ProteinProspector, PEAKS, and PeptideART, their assignment percentages for SHS would significantly increase to 75, 71 and 63%, respectively. Also, if PEAKS would take into account the $b_2 + H_2O$ ion, the number of matches with an incorrect amino acid sequence would probably be reduced.

4. Conclusions

The fragmentation pathways of five X-His-Ser tripeptides with X = Asn, Asp, Glu, Ser and Trp were predicted by QCMS^2 and its assignments of the experimental MS/MS spectra were compared with those of the tools Mascot, ProteinProspector, PEAKS and PeptideART. QCMS^2

correctly reproduces the main features in the product ion spectra, *i.e.*, the traditional backbone cleavages resulting in a-, b_2- and y-ions, but also the consequences of ISC interactions and the fragmentation of the side chains. Furthermore, the information it provides is detailed enough that for the majority of peaks in the experimental spectra the full fragmentation pathway becomes available.

Despite these accomplishments, which support the idea of the method's general applicability (to other ionization methods beyond EI and ESI/CID and to other systems), it is a very time-consuming approach and this is partly due to the generation of too many false-positives: for the two conformers of SHS 68 fragments are predicted of which only 24 (or about 35%) are actually found in the combined spectrum. Even though it can be argued that under different conditions more (other) fragments could be observed, it is clear that further optimization will be necessary before $QCMS^2$ could be used as a practical tool. Also, in all but one case (NHS) not all fragments observed in the experimental spectra could be assigned: their identification using more sophisticated experimental methods will be the first step in understanding why they are not predicted and how this problem can be remedied.

The comparison of the performance of $QCMS^2$ with the three conventional tools – even if not completely warranted – is instructive since it exposes some of the deficiencies of these tools and this may be used to also improve them in order to achieve more accurate peptide identification.

5. Acknowledgments

A number of calculations were performed using the Hopper HPC infrastructure at the CalcUA core facility of the University of Antwerp, a division of the Flemish Supercomputer Center VSC, funded by the Hercules Foundation, the Flemish Government (Department EWI), and the University of Antwerp. This research was funded by a Ph.D. grant (to J.C. and S.C.) of the Agency for Innovation by Science and Technology (IWT).

6. References

[1] G. Zhang, R.S. Annan, S.A. Carr and T.A. Neubert, *Curr. Protoc. Protein Sci.*, 2010, **62**, 16.1.1.

[2] A.I. Nesvizhskii, in *Methods in Molecular Biology, vol. 367: Mass Spectrometry Data Analysis in Proteomics*, ed. R. Matthiesen, Humana Press Inc., Totowa, New Jersey, 1st Edition, 2007, Chapter 6: Protein Identification by Tandem Mass Spectrometry and Sequence Database Searching, 87-119.

[3] B. Lu, A. Motoyama, C. Ruse, J. Venable and J.R. Yates III, *Anal. Chem.*, 2008, **80**, 2018.

[4] (a) J.K. Eng, A.L. McCormack and J.R. Yates III, *J. Am. Soc. Mass Spectrom.*, 1994, **5**, 976; (b) J. R. Yates III, J.K. Eng, A.L. McCormack and D. Schieltz, *Anal. Chem.*, 1995, **67**, 1426; (c) J.K. Eng, B. Fischer, J. Grossmann and M.J. Maccross, *J. Proteome Res.*, 2008, **7**, 4598.

[5] D.N. Perkins, D.J.C. Pappin, D.M. Creasy and J.S. Cottrell, *Electrophoresis*, 1999, **20**, 3551.

[6] C. Hughes, B. Ma and G.A. Lajoie, in *Methods in Molecular Biology, vol. 604: Proteome Bioinformatics*, ed. S.J. Hubbard and A.R. Jones, Humana Press, New York Dordrecht Heidelberg Londen, 1st Edition, 2010, Chapter 8: De Novo Sequencing Methods in Proteomics, 105-121.

[7] K. Boutilier, M. Ross, A.V. Podtelejnikov, C. Orsi, R. Taylor, P. Taylor and D. Figeys, *Anal. Chim Acta*, 2005, **534**, 11.

[8] E.A. Kapp, F. Schütz, L.M. Connoly, J.A. Chakel, J.E. Meza, C.A. Miller, D. Fenyo, J.K. Eng, J.N. Adkins, G.S. Omenn and R.J. Simpsom, *Proteomics*, 2005, **5**, 3475.

[9] A.W. Bell, E.W. Deutsch, C.E. Au, R.E. Kearney, R. Beavis, S. Sechi, T. Nilsson and J.J. Bergeron, *Nat. Methods*, 2009, **6**, 423.

[10] J.E. Elias, W. Hass, B.K. Faherty and S.P. Gygi, *Nat. Methods*, 2005, **2**, 667.

[11] S.E. Stein and D.R. Scott, *J. Am. Soc. Mass Spectrom.*, 1994, **5**, 859.

[12] J.M. Chick, D. Kolippakkam, D.P. Nusinow, B. Zhai, R. Rad, E.L. Huttlin and S.P. Gygi, *Nat. Biotechnol.*, 2015, **33**, 743.

[13] (a) R.J. Chalkley, P.R. Baker, L. Huang, K.C. Hansen, N.P. Allen, M. Rexach and A.L. Burlingame, *Moll. Cell Proteomics*, 2005, **4**, 1194; (b) R.J. Chalkley, P.R. Baker, K.F. Medzihradsky, A.J. Lynn and A.L. Burlingame, *Mol. Cell Proteomics*, 2008, **7**, 2386; http://prospector.ucsf.edu/prospector/mshome.htm.

[14] R.J. Chalkley, K.C. Hansen and M.A. Baldwin, in *Methods in Enzymology – vol.402: Biological Mass Spectrometry*, ed. A.L. Burlingame, Elsevier Academic Press, London, 1st Edition, 2005, Chapter 9: Bioinformatic Methods to Exploit Mass Spectrometric Data for Proteomic Applications, 289-312.

[15] C. Xu and B. Ma, *Drug Duscov. Today*, 2006, **11**, 595.

[16] J. Allmer, *Expert Rev. Proteomics*, 2011, **8**, 645.

[17] (a) B. Ma, K. Zhang, C. Hendrie, C. Liang, M. Li, A. Doherty-Kirby and G.A. Lajoie, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 2337; (b) B. Ma, K. Zhang and C. Liang, *J. Comput. Syst. Sci*, 2005, **70**, 418; (c) J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G.A. Lajoie and B.Ma, *Mol. Cell Proteomics*, 2012, **11**, M111.010587-1.

[18] T. Muth and B.Y. Renard, *Brief. Bioinform.*, 2017, DOI: 10.1093/bib/bbx033.

[19] B. Ma, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 1885.

[20] S. Pevtsov, I. Fedulova, H. Mirzaei, C. Buck and X. Zhang, *J. Proteome Res.*, 2006. **5**, 3018.

[21] E. Pitzer, A. Mascelot and J. Colinge, *Proteomics*, 2007, **7**, 3051.

[22] S. Bringans, T.S. Kendrick, J. Lui and R. Lipscombe, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 3450.

[23] V. Gorshkov, S.Y.K. Hotta, T. Verano-Braga and F. Kjeldsen, *Proteomics*, 2016, **16**, 2470.

[24] (a) Z. Zhang, *Anal. Chem.*, 2004, **76**, 3908; (b) Z. Zhang, *Anal. Chem.*, 2005, **77**, 6364.

[25] R.J. Arnold, N. Jayasankar, D. Aggarwal, H. Tang and P. Radivojac, *Pac. Symp. Biocomput.*, 2006, **11**, 219.

[26] (a) S. Degroeve and L. Martens, *Bioinformatics*, 2013, **29**, 3199; (b) S. Degroeve, D. Maddelein and L. Martens, *Nucleic Acids Res.*, 2015, **43**, W326.

[27] C. Narasimhan, D.L. Tabb, N.C. Verberkmoes, M.R. Thompson, R.L. Hettich and E.C. Uberbacher, *Anal. Chem.*, 2005, **77**, 7581.

[28] R. Sadygov, J. Wohlschlegel, S.K. Park, T. Xu and J.R. Yates III, *Anal. Chem.*, 2006, **78**, 89.

[29] D.L. Tabb, C.G. Fernando and M.C. Chambers, *J. Proteome Res.*, 2007, **6**, 654.

[30] J. Cautereels, D. Geldof, M. Claeys and F. Blockhuys, *J. Mass Spectrom.*, 2016, **51**, 602.

[31] MassLynx V4.1 SCN 803, Copyright © 2010 Waters Inc.

[32] F. Bockisch, D.A. Liotard, J.C. Rayez and B. Duguay, *Int. J. Quant. Chem.*, 1992, **44**, 619.

[33] AMPAC 10, 1992-2016 Semichem inc. 12456 W 62nd Terrace-Suite D, Shawnee, KS 66216.

[34] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.

[35] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov,

J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 09 (Revision D.01), Gaussian Inc., Wallingford CT, **2016**.

[36] (a) A.D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648; (b) C. Lee, W. Yang and R.G. Parr, *Phys. Rev. B*, 1998, **37**, 785; (c) A.D. Becke, *Phys. Rev. A*, 1988, **38**, 3098; (d) S.H. Vosko, L.Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200; (e) P.J. Stephens, F.J. Devlin, C.F Chabalowski and M.J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.

[37] R. Raghavachari, J.S. Binley, R. Seeger and J.A. Pople, *J. Chem. Phys.*, 1980, **72**, 650.

[38] C. Van Alsenoy and A. Peeters *J. Mol. Struct. (Theochem)*, 1993, **286**, 19.

[39] B. Rousseau, A. Peeters and C. Van Alsenoy, *hem. Phys. Lett.*, 2000, **324**, 189.

[40] D. Geldof, A. Krishtal, F. Blockhuys and C. Van Alsenoy, *J. Chem. Theory Comput.*, 2011, **7**, 1328.

[41] F.L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129.

[42] P. Bultinck, C. Van Alsenoy and P.W. Ayers, *J. Chem. Phys.*, 2007, **126**, 144111.

[43] A. Krishtal, P. Senet and C. Van Alsenoy, *J. Chem. Theory Comput.*, 2008, **4**, 2122.

[44] A. Krishtal, P. Senet and C. Van Alsenoy, *J. Chem. Phys.*, 2010, **133**, 154310.

[45] S. Van Damme, P. Bultinck and S. Fias, *J. Chem. Theory Comput.*, 2009, **5**, 334.

[46] A. Krishtal, K. Vannomeslaeghe, A. Olasz, T. Veszpremi, C. Van Alsenoy and P. Geerlings, *J. Chem. Phys.*, 2009, **130**, 174101.

[47] S.N. Steinmann and C. Corminboeuf, *J. Chem. Theory Comput.*, 2010, **6**, 1990.

[48] C. Peng and H.B. Schlegel, *Israel J. Chem.*, 1993, **33**, 449.

[49] C. Peng, P.Y. Ayala, H.B. Schlegel and M.J. Frisch, *J. Comput. Chem.*, 1996, **17**, 49.

[50] N. Russo, M. Toscano, A. Grand and F. Jolibois, *J. Comput. Chem.*, 1998, **19**, 989.

[51] J.H. Gross, *Mass Spectrometry: A Textbook*, Springer-Verlag, Berlin Heidelberg, 2004.

[52] G. Siuzdak, *The Expanding Role of Mass Spectrometry in Biotechnology 2nd Edition*, MCC Press, Sand Diego, 2006.

[53] P. Roepstorff and J. Fohlmann, *biomed. Mass Spectrom.*, 1984, **11**, 601.

[54] K. Biemann, *Biomed. Envir. Mass spectrom.*, 1988, **16**, 99.

[55] B. Paizs and S. Suhai, *Mass Spectrom. Rev.*, 2005, **24**, 508.

[56] V.H. Wysocki, G. Cheng, Q. Zhang, K.A. Herrmann, R.L. Beardsley and A.E. Hilderbrand, *Principles of Mass Spectrometry Applied to Biomolecules*, ed. J. Laskin and C. Lifshitz, John Wiley & Sons Inc., Hoboken New Jersey, 1st Edition, 2006, Chapter 8: Peptide Fragmentation Overview, 277-300.

[57] M. Maenhaut-Claeys and M. Vandewalle, *Bull. Soc. Chim. Belges*, 1974, **83**, 343.

[58] B. Paizs and S. Suhai, *Rapid Commun. Mass Spectrom.*, 2001, **15**, 651.

[59] B.J. Bythell, I.P. Csonka, S. Suhai, D. Barofsky and B. Paizs, *J. Phys. Chem. B*, 2010, **114**, 15092.

[60] M. Dupré, S. Cantel, J. Martinez and C. Enjalbal, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 330.

[61] A.M. Falick, W.M. Hines, K.F. Medzihradszky, M.A. Baldwin, B.W. Gibson, *J. Am. Soc. Mass Spectrom.*, 1993, **4**, 882.

[62] I.A. Papayannopoulos, *Mass Spectrom. Rev.*, 1995, **14**, 49.

[63] J.P. DeGnore and J. Qin, *J. Am. Soc. Mass Spectrom.*, 1998, **9**, 1175.

[64] F.W. McLafferty and F. Tureček, *Interpretation of Mass Spectra*, University Science Books, Sausalito California, 1993.

Table 1. Calculated energies (ΔE in kJ.mol^{-1}) of the protomers of the five studied tripeptides relative to the energy of the protomer obtained after protonation of the imidazole moiety of the central histidine; SC denotes side chain. Empty cells correspond to protomers which were not suggested by the MEP.

Protonation sites	DHS	EHS	NHS	SHS	WHS
N-terminus	40.66	87.56	53.67	60.78	30.77
C=O of X	130.38	147.59	58.47	133.45	39.68
C=O of H	84.98	124.48	131.58	142.98	94.92
C=O of S	162.03			153.42	139.96
OH of SC of X	147.20	115.16		227.96	
C=O of SC of X	205.31		127.57		
OH of SC of S	183.17	236.78	214.33	227.96	127.31

Table 2. List of the predicted fragments (*m*/*z*) of SHS by QCMS^2. The underlined masses are those of the fragments discussed in the text.

1st generation fragments	2nd generation fragments	3rd generation fragments	4th generation fragments
313	295		
312	295		
	294		
	282		
299[a]	268		
	254		
	167		
	31		
299[a]	268		
	239		
	198		
	101		
	60		
	31		
285	254		
	225		
	198		
	87		
	60		
270	239		
	225		
	138	110	
	132	101	
		87	
		45	
		31	
243[b]	226	208	
	225	208	
243[b]	226	208	
	225	208	
	156		
225	208	190	
	207	190	
	197	180	162

		179 167 138	162
	195 166 60		
197	180 179 167 138	162 162	
132	101 87 45 31		
106			

[a] *m/z* 299 is formed by the homolytic cleavage of both bonds 5 and 6 yielding two structurally different fragments with *m/z* 299 which both give rise to different fragments.

[b] *m/z* 243 is either the y_2-ion (first) or the $b_2 + H_2O$-ion (second).

Table 3. Assignments of the fragment peaks of SHS.

m/z	Protein Prospector	PEAKS	PeptideART	QCMS^2
330	MH^+	MH^+	MH^+	MH^+
313			$MH^+ - NH_3$	$MH^+ - NH_3$
312	$MH^+ - H_2O$		$MH^+ - H_2O$	$MH^+ - H_2O$
295				$MH^+ - H_2O - NH_3$
243	y_2 and $b_2 + H_2O$	y_2	y_2	y_2 and $b_2 + H_2O$
226		$y_2 - NH_3$	b_2 isotope peak and $y_2 - NH_3$	$y_2 - NH_3$ and $b_2 + H_2O - NH_3$
225	b_2 and $y_2 - H_2O$	b_2 and $y_2 - H_2O$	$y_2 - H_2O$	b_2 and $y_2 - H_2O$ and $[MH^+ - CH_2OH^\bullet - CONHCH(CH_2OH)COOH^\bullet]$
209				*m/z* 208 isotope peak
208		$b_2 - NH_3$	$y_2 - H_2O - NH_3$	$b_2 - NH_3$ and $y_2 - H_2O - NH_3$
207	$b_2 - H_2O$	$b_2 - H_2O$	$b_2 - H_2O$	$b_2 - H_2O$
197	a_2	a_2		a_2
195				$b_2 - CH_2O$
190				$b_2 - H_2O - NH_3$
180		$a_2 - NH_3$		$a_2 - NH_3$
179		$a_2 - H_2O$		$a_2 - H_2O$
167				$a_2 - CH_2O$ and $[MH^+ - COOH^\bullet - H_2NCHCH_2OH^\bullet]$
166				$b_2 - NH{=}CHCH_2OH$
162				$a_2 - H_2O - NH_3$
156				y_1 of $b_2 + H_2O$
152				
150				
138	$HisH^+$			$HisH^+$
122				
110	immonium His	immonium His		immonium His

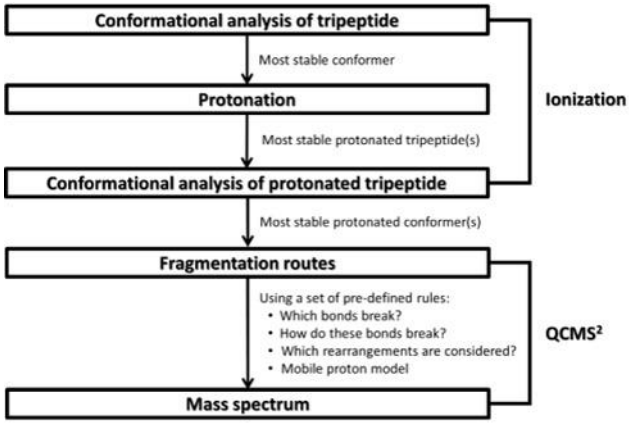

Figure 1. The two steps in the procedure involving QCMS^2 applied to tripeptides; ionization and fragmentation.

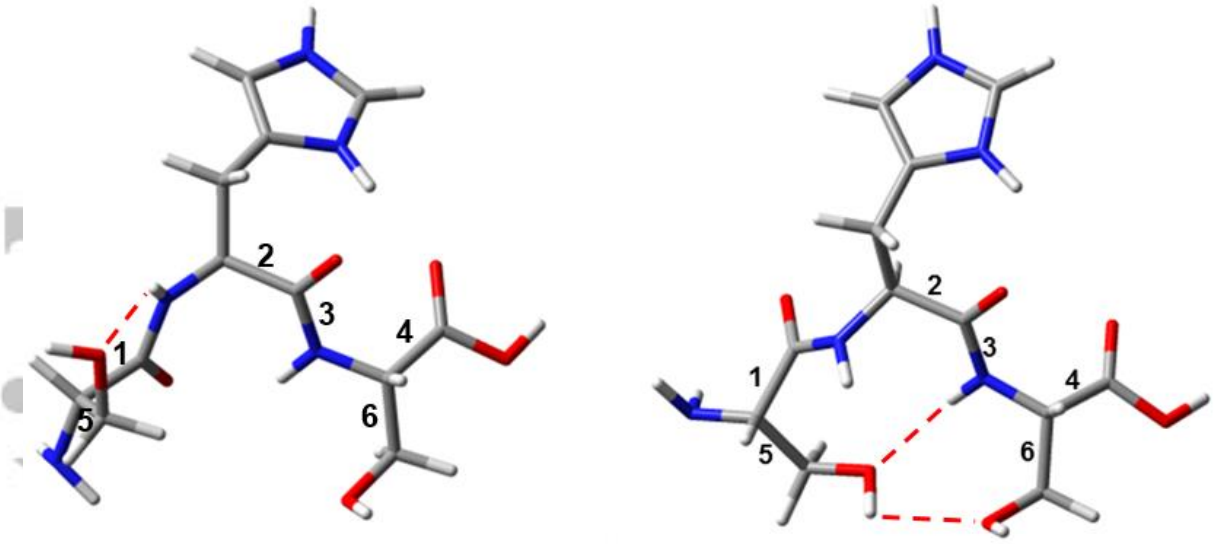

Figure 2. The lowest- (left) and second-lowest-energy conformers (right) of SHS; the six bonds considered for fragmentation have been specified in both conformers. The red dashed lines indicate the interactions discussed in the text.

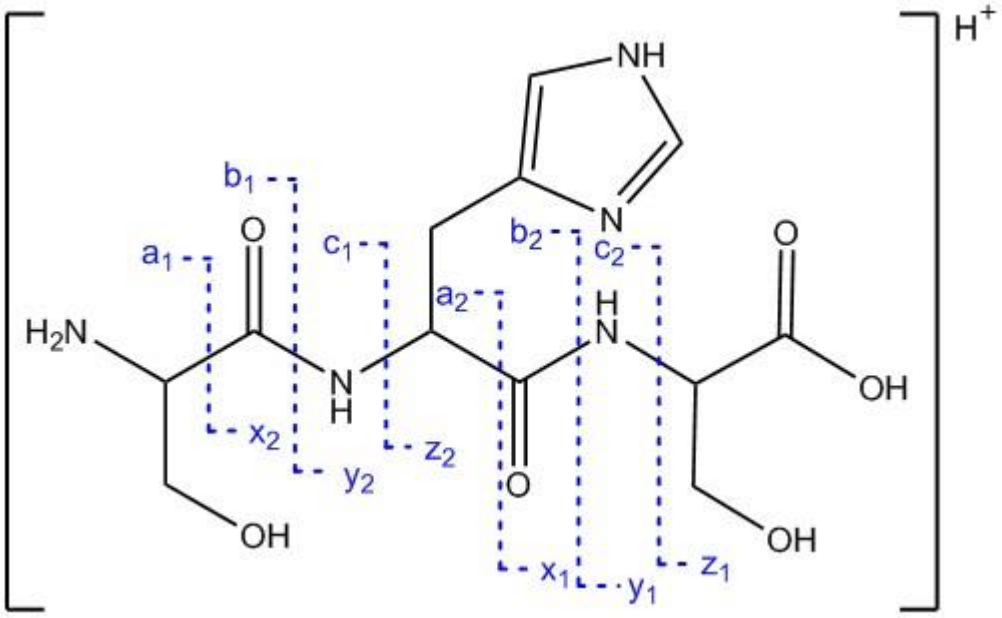

Figure 3. The nomenclature of common ion types in peptide fragmentation.

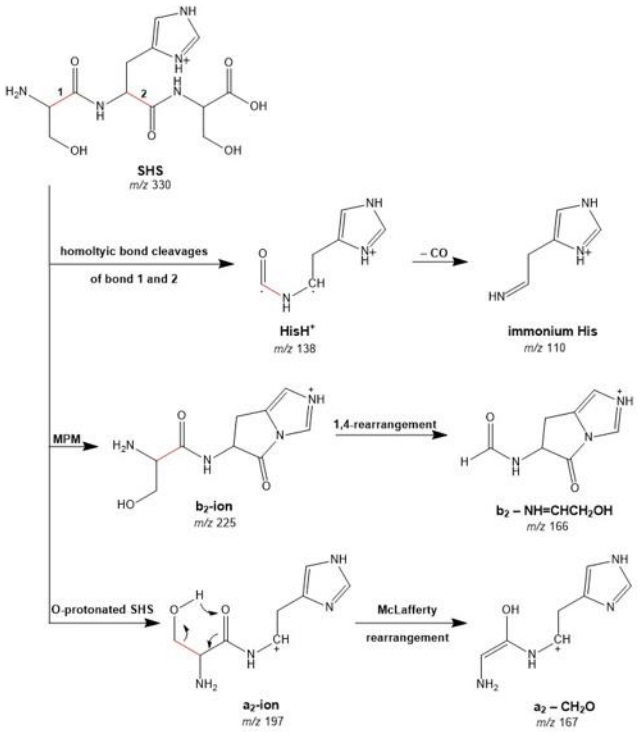

Figure 4. The most important predicted fragmentation pathways of SHS by QCMS[2].

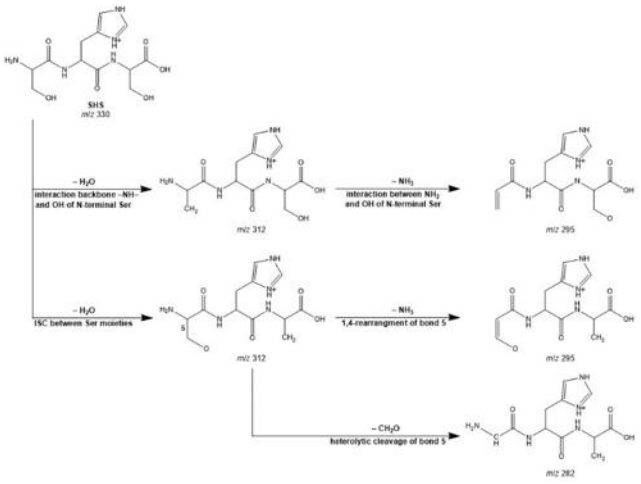

Figure 5. The different and additional fragmentations of the second lowest-energy conformer of SHS.

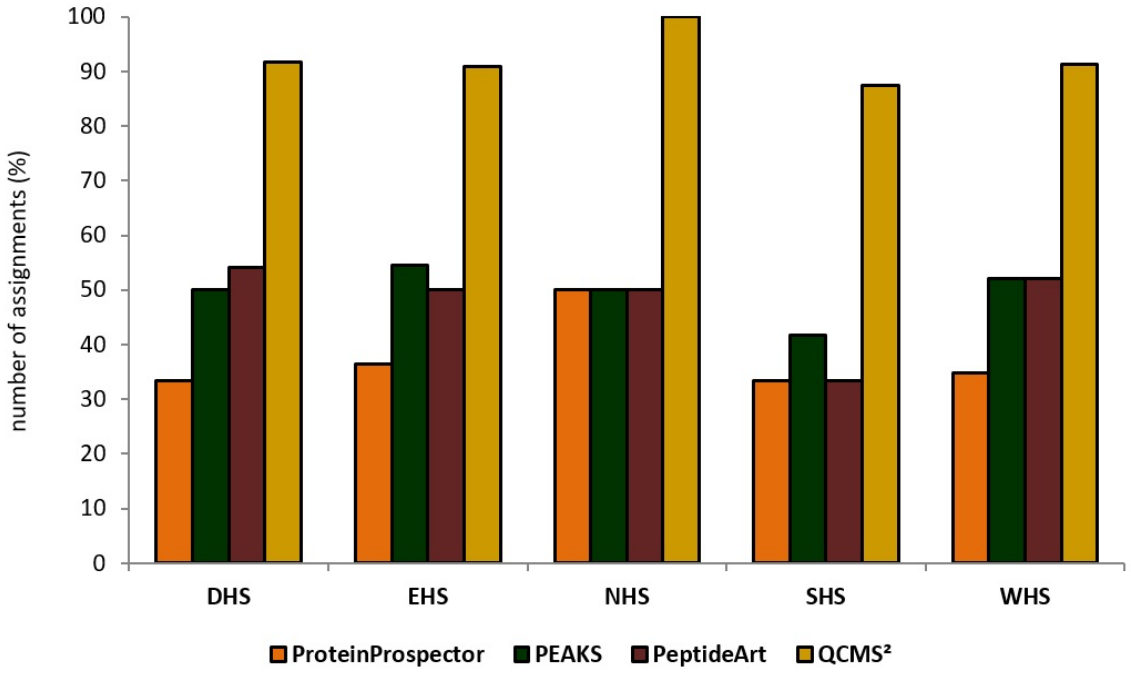

Figure 6. Comparison of the assignments by ProteinProspector, PEAKS, PeptideART and QCMS^2 for the five X-His-Ser tripeptides.