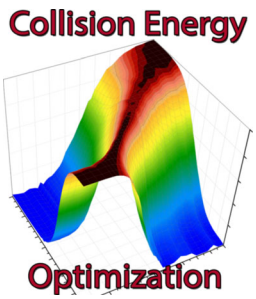
J. Am. Soc. Mass Spectrom. (2016) 27:1357–1365
DOI: 10.1007/s13361-016-1408-y

RESEARCH ARTICLE

Quantitative Comparison of Tandem Mass Spectra Obtained on Various Instruments

Fanni Laura Bazsó,[1] Oliver Ozohanics,[2] Gitta Schlosser,[3] Krisztina Ludányi,[4] Károly Vékey,[1,5] László Drahos[1,2]

[1]MS Proteomics Research Group, Research Center for Natural Sciences, Hungarian Academy of Sciences, H-1117, Magyar tudósok krt. 2, Budapest, Hungary
[2]MTA-TTK NAP B MS Neuroproteomics Research Group, Research Center for Natural Sciences, Hungarian Academy of Sciences, H-1117, Magyar tudósok krt. 2, Budapest, Hungary
[3]MTA-ELTE Research Group of Peptide Chemistry, Hungarian Academy of Sciences, Eötvös Loránd University, 1117, Budapest, Hungary
[4]Department of Pharmaceutics, Semmelweis University, Hőgyes E. Street 7-9, H-1092, Budapest, Hungary
[5]Core Technologies Center, Research Center for Natural Sciences, Hungarian Academy of Sciences, H-1117, Magyar tudosok krt. 2, Budapest, Hungary



Abstract. The similarity between two tandem mass spectra, which were measured on different instruments, was compared quantitatively using the similarity index (SI), defined as the dot product of the square root of peak intensities in the respective spectra. This function was found to be useful for comparing energy-dependent tandem mass spectra obtained on various instruments. Spectral comparisons show the similarity index in a 2D "heat map", indicating which collision energy combinations result in similar spectra, and how good this agreement is. The results and methodology can be used in the pharma industry to design experiments and equipment well suited for good reproducibility. We suggest that to get good long-term reproducibility, it is best to adjust the collision energy to yield a spectrum very similar to a reference spectrum. It is likely to yield better results than using the same tuning file, which, for example, does not take into account that contamination of the ion source due to extended use may influence instrument tuning. The methodology may be used to characterize energy dependence on various instrument types, to optimize instrumentation, and to study the influence or correlation between various experimental parameters.

Keywords: Tandem mass spectrometry, Similarity, Instrument comparison

Received: 22 December 2015/Revised: 8 April 2016/Accepted: 13 April 2016/Published Online: 20 May 2016

Introduction

Quantitative comparison of (tandem) mass spectra is often needed. A common case is determination of reproducibility and repeatability, in which case spectra of the same compound are compared. A related issue is deciding whether the spectrum of an unknown is identical (within certain tolerance) to the spectrum of a reference compound, which is an important step in establishing the identity of two compounds. These issues are particularly important for pharmaceutical applications and, in general, for well-regulated [good laboratory practice (GLP), good manufacturing practice (GMP)] environments. Similarity (repeatability or reproducibility) in such cases is typically measured as the average (relative) standard deviation of intensities of various peaks in the spectra.

A different issue is structure analysis based on mass spectrometry. Traditionally, manual spectrum evaluation was commonly used. Now emphasis is shifting to automated procedures: (1) comparison with and search of data banks, and (2) comparing experimentally obtained and theoretically expected mass spectra. Data bank search has long been in use, especially for electron impact spectra [1–3]. Comparison to expected (theoretical) tandem mass spectra is the basis of proteomics

Electronic supplementary material The online version of this article (doi:10.1007/s13361-016-1408-y) contains supplementary material, which is available to authorized users.

Correspondence to: László Drahos; *e-mail:* drahos.laszlo@ttk.mta.hu

[4], even if comparisons use predominantly fragment ion masses only, with no or limited use of ion abundance.

Quantitative comparison of ion intensities in tandem mass spectra is not widespread, although it is a subject of active research [5]. Tandem MS databases (containing ion intensities) would be highly useful for structure analysis, especially in proteomics [6]. The main difficulty is that tandem mass spectra strongly depend on experimental conditions, which are difficult to standardize [7]. The most important experimental parameter in tandem mass spectra is energetics [8], in most cases controlled by collision energy. Note that throughout the manuscript, collision energy refers to the laboratory frame collision energy. (In ion trap instruments it is often described as the fragmentation amplitude or tickling voltage or normalized collision energy.) There are other parameters influencing energetics, like mass of the collision gas (e.g., He or N_2, influencing the center of mass collision energy [8]), collision gas pressure, and the residence time (in traps). Various tuning parameters, ion-molecule reactions (mainly in ion traps), and contaminations in the ion source or ion optics may also influence relative ion intensities, although these have usually minor effect compared with the effect of collision energy. Studying energy-dependent (or energy resolved) mass spectra it is typical to keep all instrument parameters constant, and vary the collision energy only [9–11]. Note that in single stage mass spectra using electrospray ionization, similar changes may be induced varying the skimmer/cone voltage [12].

A further parameter, which is important for inter-laboratory comparisons, but not easy to vary in practice, is the type of instrument used. It is known that ion trap instruments often yield different tandem mass spectra from quadrupole type instruments, but quantitative comparisons are relatively rare. A recent study on leucine enkephalin (YGGFL [13]) has shown that setting the collision energy to a value producing 50% survival yield [10, 11, 14] (when the total abundance of fragments is equal to that of the protonated molecule), tandem mass spectra obtained on various instrument types are qualitatively similar [13], although "similarity" was not defined in a quantitative manner.

In the present article, we shall quantitatively compare ion intensities in tandem mass spectra as a function of collision energy using various instrument types. Our main objective is to describe similarity of tandem mass spectra obtained on various instruments, and to determine what degree of similarity can be obtained by varying the collision energy. For example, if some experiments in the pharma industry were run on a certain mass spectrometer, and this needs to be transferred to another instrument, will the results be acceptable for the regulatory body? In other words, can a tandem MS spectrum obtained on one instrument be tuned to be sufficiently similar to that obtained on another instrument?

Experimental

Mass Spectrometry

Experiments were performed using a Waters Micromass Quattro type QQQ, a Waters QTOF Premier (Manchester, UK), an Agilent 6460 QQQ (Santa Clara, CA, US), and a Bruker Esquire 3000+ (Bremen, Germany) ion trap type mass spectrometer in positive electrospray ionization mode (subsequently abbreviated as Waters QQQ, Waters QTOF, Agilent QQQ, and Bruker IT). The samples were infused with a syringe pump into the electrospray source at the rate of 10 μL/min using 1:1 water:acetonitrile +0.1% formic acid as solvent. With the exception of the collision energy experimental conditions were kept constant during the experiments. The source conditions were as follows:

- Micromass Quattro: voltage of the capillary was 3.5 kV, the voltage of the cone was 10 V, and the temperature of the source was 363 K. The collision gas was argon.
- QTOF Premier: voltage of the capillary was 2.8 kV, voltage of sampling cone was 15 V, voltage of extraction was 3 V, and the temperature of the source was 363 K. The collision gas was argon.
- Agilent 6460: voltage of the capillary was 3.5 kV, fragmentor voltage was 50 V, and the temperature of the source was 350 K. The collision gas was N_2.
- Bruker Esquire 3000+: capillary voltage was 4000 V, nebulizer gas pressure was 10 psi, drying gas flow was 4 L/min, and the heated capillary temperature was 523 K. The buffer gas was He.

The collision energy was varied in the 1–110 eV range on the quadrupole type instruments. We have used between 1 and 10 eV, 2 eV steps, between 10 and 30 eV, 1 eV steps, between 30 and 50 eV, 2 eV steps, and after 50 eV until 110 eV, 10 eV steps. In case of Bruker ion trap we have used between 0.1 and 0.3 V, 0.1 V steps, between 0.3 and 0.4 V, 0.02 V steps, between 0.4 and 0.5 V, 0.01 V steps, between 0.5 and 0.7 V, 0.02 V steps, and after 0.7 V until 0.9 V, 0.1 V steps.

Samples

Leucine enkephalin (amino acid sequence is YGGFL), adenosine, α-aminoadipic acid, and aminocaproic acid samples have been studied. All chemicals were purchased from Sigma-Aldrich (Budapest, Hungary).

Spectral Comparisons

Spectral comparisons were made between two selected spectra. Several mathematical functions were tested for spectral comparison; most gave analogous results. In the present paper, we have decided to use the square root of spectral intensities, as this enhances the significance of small peaks, which was deemed advantageous [15, 16]. Comparison between two spectra was based on the dot product (of the square root of the ion intensity in the spectra), and this is referred to as the similarity index

(SI). This method is often used for spectral comparisons [17, 18]. The mathematical formula therefore is:

$$\text{Similarity index, } SI = \cos\theta = \frac{\sum_i \sqrt{x_i}\cdot\sqrt{y_i}}{\sqrt{\sum_i x_i \cdot \sum_i y_i}} \quad (1)$$

Here x_i are intensities of the peaks in one spectrum and y_i are intensities in the other spectrum, and the sum goes over all peaks in the spectra. The similarity index varies between 1 (in this case the spectral contrast angle $\theta = 0°$, $\cos\theta = 1$, when the spectra are identical) and zero ($\theta = 90°$, $\cos\theta = 0$, when the spectra are completely different, and even do not contain common fragments). Spectra shown in this paper were normalized to the sum of peak intensities, but the similarity index (and also the reproducibility described below) is independent of normalization.

The similarity index is a good and often used mathematical expression for spectral comparisons, especially for library search [17, 18], but its use is not common in the pharmaceutical field (where spectral variabilities are often expressed as reproducibility) or in the proteomics field, where quantitative differences are typically described as "fold-difference." In order to relate to these applications, and to provide approximate numerical correspondence between SI, reproducibility, and fold-difference, we describe these expressions as well. For testing reproducibility the spectral difference is often measured by the formalism of relative standard deviation. For two spectra this is described by Equation (2):

$$\text{Reproducibility(rsd \%)} = \sqrt{2}\cdot 100\cdot \frac{\sum_{i=1}^{n} \frac{|x_i - y_i|}{x_i + y_i}}{n} \quad (2)$$

Here x_i and y_i are the same as described above, whereas n is the number of peaks in the mass spectrum. In the measurement of reproducibility, very small peaks are usually excluded; here we used an intensity cutoff at 1%.

In the proteomics field, spectral differences (and the difference between the amount/concentration of various proteins in complex mixtures) are often characterized by the "fold-difference" [19]. This is the ratio of protein concentration/amount (typically measured by the ratio of selected ion intensities) in two samples. This (averaged over all peaks or all compounds present) is also a measure of the difference between two samples or two spectra. Here we use the average fold difference to compare its magnitude with the similarity index.

$$\text{Fold difference} = \frac{\sum_{i=1}^{n} \frac{x_i}{y_i}}{n} \quad (3)$$

Here x_i and y_i are the intensities of the i^{th} peak in the two spectra compared, and the sum goes over all peaks in the spectra. Note that if one peak was missing in one of the spectra, it was considered to be equal to the cutoff value (1%). Note, Equations (1–3) have been used before (Refs [14–18]). Here we adopt them for our purpose.

Results and Discussion

It is probably easiest to start spectral comparisons with simple examples. Various tandem MS spectra of leucine encephalin (YGGFL) were compared pairwise, and only peaks with over 1% relative abundance were considered. When spectra were measured in 1 day, the similarity index (SI, Equation (1)) between spectra (using 20 eV collision energy on the Waters QTOF instrument) was better than 0.9999, while repeatability was ca. 1%. The long-term reproducibility was worse (as expected. One such case is shown in Figure 1a. These were also taken on the Waters QTOF instrument at 20 eV collision energy, but with 1 year distance in time (which means somewhat different tuning and source conditions). In this case, reproducibility is 4% (average rsd); the similarity index is 0.9995. Figure 1b shows the difference between tandem mass spectra taken on two QQQ type instruments (Waters and Agilent), both obtained at 13 eV collision energy. Although the two spectra share most fragment ions, relative intensities are quite different. Peak intensities differ by a factor of 2–8 times (for various fragment peaks the smallest fold-difference is 2, the biggest is 8, while the average fold-difference is 5.4), and the spectral difference is clearly out of the range, where reproducibility is a useful measure. The similarity index in this case is 0.9008. Such spectra would not even support the idea that the two samples are identical. Spectral similarity can be improved if we consider one spectrum as reference (Agilent QQQ at 13 eV) and tune the collision energy on the Waters QQQ in order to maximize the similarity index. This was obtained using 17 eV on the Waters, and the two spectra are shown in Figure 1c. These show good agreement (in contrast to that shown in Figure 1b), reflected by the high similarity index, which has improved to 0.9959. Reproducibility is 14%, the fold-difference is 1.5. This is a relatively high value and is caused by a systematic difference in the intensity of some small, high energy fragments, like the F, Y, and b_3 ions.

The last example is comparison of a 20 eV spectrum taken on the Waters QQQ with a spectrum taken on a different instrument type, the Bruker ion trap. On the Bruker instrument, we have selected the collision energy (0.62 V amplitude), which gave the best similarity index to the QQQ spectrum. The two spectra are shown in Figure 1d. Although most ions are present in both spectra, the intensities are significantly different. The similarity index is 0.8799 and the average fold-difference is 5.1. Spectra differing to such a degree may support structural similarity but not structural identity.

Based on the examples shown above, when the similarity index (Equation (1)) is higher than ca. 0.99, reproducibility

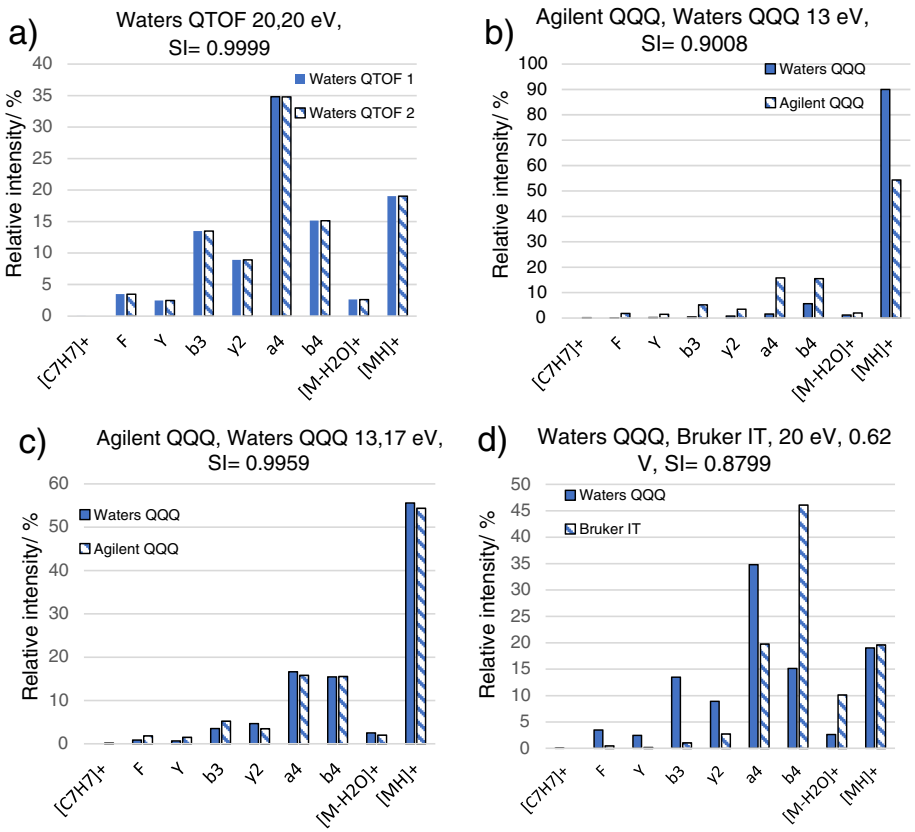


Figure 1. Similarity of tandem mass spectra of leucine enkephalin (YGGFL) measured (**a**) on the same QTOF type instrument, but in 1 year distance in time, (**b**) on two QQQ type instruments (Waters and Agilent), both obtained at 13 eV collision energy, (**c**) on two different QQQ instruments using collision energies that give the most similar spectra (collision energies are 13 and 17 eV, respectively), (**d**) on QQQ and ion trap instruments (the collision voltage on IT were tuned to give the best similarity index)

(Equation (2)) is also a good measure to compare spectral differences. Spectra with SI higher than 0.99 or when the spectral difference measured by Equation (2) is better than 10%, the spectra may be used to confirm structural identity of a compound, and are adequate to use for comparisons in the pharma industry. Although the SI is a good quantitative measure of spectral differences, as a rough guide we would suggest some qualitative limits as well. When the similarity index is lower than ca. 0.99 but higher than ca. 0.90, the spectra show major differences in peak intensities, but most fragments are present in both spectra. In this range, reproducibility is inadequate to characterize the differences. It is better to use the "fold-difference": the average ratio of peak intensities in the two spectra. Approximately 3- to 10-fold differences will correspond to spectra with a similarity index around 0.90. Using the average fold-difference is common terminology in the proteomics field [20]. Fold-differences around 2–3 are typically the range that is considered to be acceptable instrumental and biological variability in proteomics [20, 21]. Spectra with similarity indices above 0.90 are also adequate for most library search algorithms. When the similarity index is below ca. 0.90 but above ca. 0.50, the spectra will have some common fragment ions but the similarity may not even be adequate for library search algorithms.

Similarity of energy-dependent tandem mass spectra on two instruments can be determined, in a general case, in the following manner: the collision energy on one instrument (Agilent QQQ) is set to a given value (e.g., 1 eV), and the collision energy on the other instrument (Waters QQQ) is scanned over the full energy range. The similarity indices between the 1 eV Agilent spectrum and the Waters spectra taken at various energies are calculated. In the next step the whole process is repeated using an increased collision energy on the Agilent. In this manner, similarity indexes for all combinations of collision energies are determined, and the results are shown in a 3D contour map (Figure 2a). The same data are also shown in a 2D color coded "heat map" (Figure 2b). Figure 2 illustrates the generally most useful 1–50 eV collision energy range; data for the full, 1–100 eV range studied are given as Supplementary Material (Supplementary Figure S1).

Figure 2 gives a lot of information on energetics and on the comparison of the two instruments. The ridge in Figure 2a represents the combination of collision energies, which yields the most similar spectra on the two instruments. The position of the "ridge" (which may be called the conversion function, Figure 3a) indicates the combination of collision energies yielding the best similarities on the two instruments. Figure 3a shows a good linear correlation between the collision energy

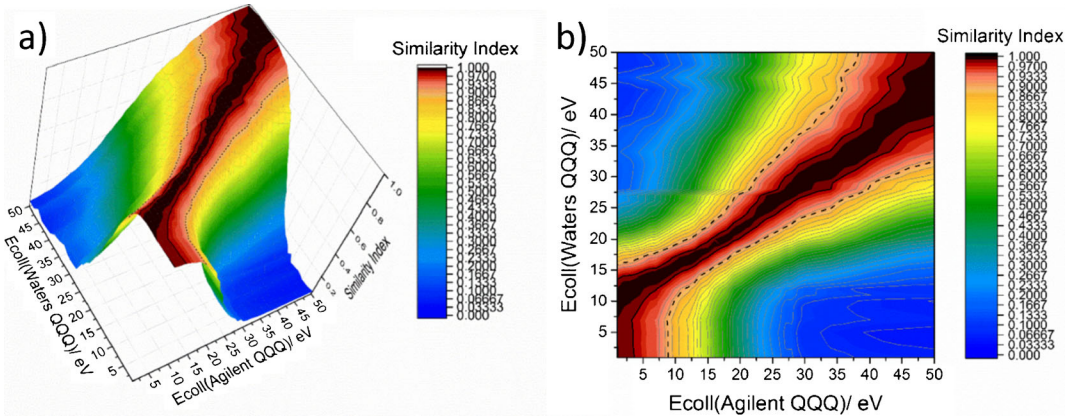


Figure 2. Similarity indices show (**a**) as a 3D contour map, and (**b**) as a "heat map" for all combinations of collision energies determined on Waters and Agilent QQQ type instruments

pairs (with a correlation coefficient $R^2 = 0.987$), as might be expected on two similar instrument types. However, the slope is less than unity (0.7), indicating that 1 eV collision energy increase on the Waters instrument increases the internal energy of the ions much more than 1 eV on the Agilent instrument. The y intercept is at 10 eV, indicating that 1 eV collision energy on the Agilent instrument produces similar internal energy as 11 eV on the Waters instrument. In other words, the Agilent instrument produces relatively "hot" protonated molecules even at low collision energy; but the internal energy increases with collision energy at a much slower rate on the Agilent than on the Waters QQQ. Note that in all cases the protonated, unfragmented YGGFL ion was mass-selected for the tandem MS experiment.

The two instruments have been compared not only using YGGFL but other compounds as well. Fragmentation of protonated adenosine, α-aminoadipic acid, and aminocaproic acid were also studied on the Waters and Agilent QQQ instruments, and the two instruments were compared in the case of these compounds as well. The results are shown in 3D contour maps in Supplementary Figure S2, and are analogous to that found for YGGFL: (1) all compounds show an approximately linear conversion function (ridge in the contour map); (2) the collision energy dependence is stronger on the Waters QQQ (the slopes of the conversion function were between 0.6 and 0.7); and (3) the Agilent QQQ produced relatively hot ions even at low collision energy.

We have performed various experiments to understand the reason for the difference between the two instruments. We have varied the fragmentor voltage on the Agilent QQQ (in the 1–120 eV range, which is the same or lower than generally recommended); we have changed the collision gas on the Waters (Ar to N_2); and we have changed the pressure on the collision gas (the pressure was reduced by 40% on the Waters). Changing these parameters did not change the qualitative appearance of Figure 2 or Figure 3a, only the slope of the ridge got closer to unity (i.e., closer to 45° slope), when the same collision gas was used on both instruments. Based on these results, we suggest that the likely reason for the "hot" ions on the Agilent QQQ may be the type of ion source (heated capillary on the Agilent, whereas skimmer-cone on the Waters). The main reason for the rate of collision energy increase

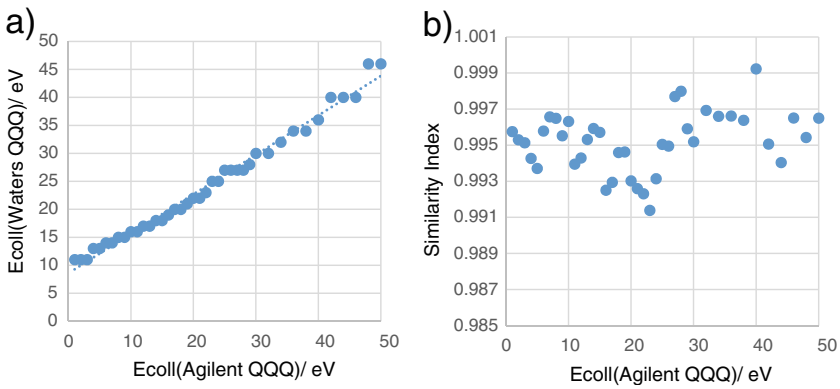


Figure 3. (**a**) Combination of collision energies that yield the most similar spectra on the two QQQ type instruments (Waters and Agilent); (**b**) the calculated similarity indices at these combination of collision energies plotted as a function of collision energy on Agilent QQQ instrument

is likely connected to the collision gas. We believe that a 2D or 3D similarity index plot is a good technique to pinpoint such differences between instruments (and also between various sets of experimental conditions).

Returning to Figure 2, another feature is the height of the ridge, which indicates how similar spectra can be obtained on the two instruments. Figure 2 shows that the ridge is high in the full collision energy range. Figure 3b indicates this in a more quantitative manner, showing that by tuning the collision energy, good agreement (SI better than 0.99) can be obtained between the two QQQ instruments in the full collision energy range.

A related feature in Figure 2a is the steepness and narrowness of the ridge, which indicates how much the spectra are changed by varying the collision energy. The ridge is steepest and narrowest around 20 eV, measured on Waters and Agilent QQQ instruments. This is the energy range, where the collision energy needs to be tuned very accurately (better than 1 eV) to obtain the best agreement between two instruments (or to get good reproducibility).

Summarizing results on the two QQQ type instruments, the following can be established:

(1) Good agreement between spectra taken on different QQQ instruments can be obtained by adjusting the collision energy. This works well in the full collision energy range (Figure 3b).
(2) The internal energy content of the precursor ions depends significantly on the instrument. The Agilent QQQ produces relatively "hot" ions even at low collision energy (Figure 3a). This may be related to the design of the heated capillary type ESI source, and might be a disadvantage for studying molecules that fragment easily (e.g., glyco- or phosphopeptides).
(3) A certain increase in collision energy does not increase fragmentation to the same degree on the two instruments compared (Figure 3a). This is not a practical problem as long as it is taken into account.
(4) Variation of the collision energy changes the spectra to a different degree in various collision energy ranges, and this is reflected by the steepness of the slope in Figure 2a. The change is largest in the medium collision energy range. In this range, the collision energy needs to be tuned very precisely to get good reproducibility.
(5) Figures 1, 2, and 3 are obtained for protonated leucine enkephalin. Other molecules show analogous behavior (see some examples in Supplementary Figure S2). The amount of information and the range where the collision energy influences the spectra most depend on the compound studied. For molecules that fragment easily, these are shifted to lower energies.

Having compared two QQQ type instruments produced by different companies, it is also important to compare two different instrument types produced by the same company. For this comparison, we have chosen a Waters QQQ and a Waters QTOF instrument. The 3D diagram showing the similarity indices and the corresponding 2D heat map are presented in Figure 4. This shows a quite symmetrical ridge and heat map. This means that the Waters QQQ and QTOF instruments behave very similarly with respect to tandem MS fragmentation. The "conversion function" is linear, with unit slope (meaning that 1 eV increase in collision energy leads to the same change in both instruments). The y intercept is –5 eV, indicating that at very low energy the QTOF instrument produces slightly hotter ions than the QQQ, although these are still less hot than on the Agilent QQQ (the difference there was 10 eV). The height of the ridge indicates the best similarity available between the two instruments at a certain collision energy. It is higher than 0.99 at all collision energies in Figure 2 (Waters and Agilent QQQ), and at most collision energies in the QTOF-QQQ comparison (Figure 4). It drops slightly to 0.988 in Figure 4 at around 30–40 eV collision energy (measured on the Waters QQQ). We have checked the data and this drop in spectral similarity is not due to random errors but to a small but systematic difference in the appearance of some low intensity fragments.

The Waters QTOF and QQQ instruments were compared for other compounds, protonated adenosine, α-aminoadipic acid, and aminocaproic acid as well, and the 3D contour maps

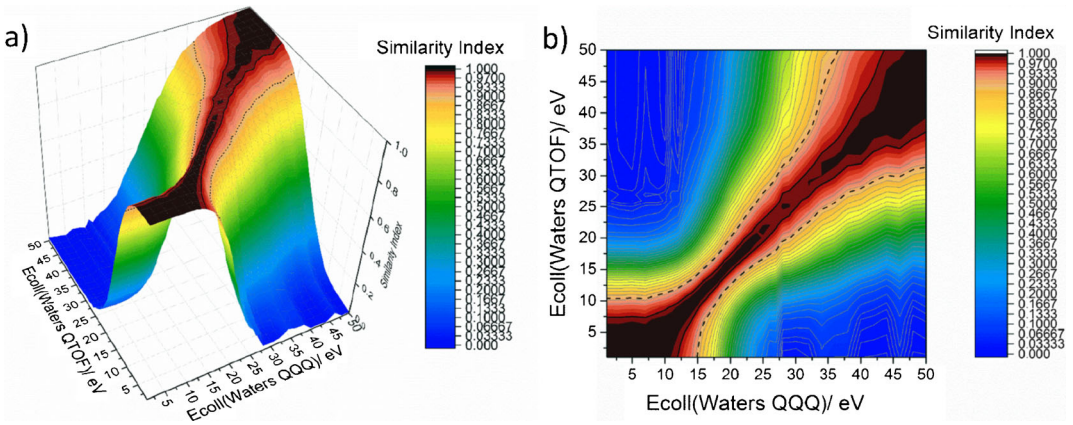


Figure 4. Similarity indices shown (**a**) as a 3D contour map, and (**b**) as a "heat map" for all combinations of collision energies determined on Waters QTOF and Waters QQQ instruments

are shown in Supplementary Figure S3. These show much worse agreement at high collision energies than that observed for leucine encephalin. The results show that for these molecules, the QQQ instrument can reach higher internal energies (capable of producing high energy fragment ions) than the QTOF. The likely reason is the special "traveling wave" type collision cell on the Waters QTOF instrument. The results seem to suggest that lower mass compounds (like protonated adenosine, α-aminoadipic acid, and aminocaproic acid) at high "collision energy" are not trapped efficiently by the traveling wave and, therefore, travel at a lower velocity (converting to lower collision energy) than a fully trapped ion. However, study of the collision mechanism in the traveling wave was outside our current interest, and was not studied in detail.

We have studied doubly protonated tryptic peptides derived from bovine serum albumin as well (m/z = 582.3189, sequence: LVNELTEFAK, m/z = 653.3617, sequence: HLVDEPQNLIK, m/z = 740.4013, sequence: LGEYGFQNALIVR). Two-dimensional heat maps corresponding to these peptides on the Waters QTOF and QQQ instruments are shown in Supplementary Figure S4. These illustrate qualitatively similar behavior as discussed above: a relatively narrow ridge in the medium energy range and approximately linear correlation between collision energies on the two instruments needed to obtain similar spectra. Physicochemical studies, like comparing the behavior of singly and multiply charged ions, investigating charge separation processes, or the influence of structure or size of the studied compounds, were outside the scope of the present study.

The last comparison is between a QQQ instrument (Waters) and an ion trap (Bruker). It is well known that collisional activation is significantly different on ion traps and on quadrupole-type instruments. In quadrupole instruments the (laboratory frame) collision energy is in the 1–100 eV range, there are relatively few (5–100) collisions, and the timeframe for the collision regime (the time necessary for the ions to pass through the collision cell) is 1 ms or less. In contrast, in ion trap instruments collisions occur at much lower energies (less than 1 eV laboratory frame) but the number of collisions is much higher (thousands or even millions of collisions may occur); and the timeframe for collisions is in the 100 ms range. Ion trap excitation is therefore often called "slow heating" [22]. For these reasons, we expect significant differences in the similarity profile determined for the QQQ–ion trap comparison.

The 3D contour map and the 2D heat map for the Waters QQQ and Bruker ion trap comparison is shown in Figure 5, which differs markedly from those shown in Figures 2 and 4. At low collision energy (called collision amplitude on the Bruker), the Bruker instrument is even "colder" than the Waters QQQ (which was coldest among the quadrupole type instruments), and up to 0.4 V amplitude there is no fragmentation. At higher but still low collision energy (0.4–0.5 V amplitude on the Bruker and 5–15 eV on the Waters QQQ) fragmentation starts, and spectra on the two instruments can be matched well. The similarity index in this range is better than 0.99, comparable to that observed among the quadrupole type instruments. This collision energy range corresponds to the beginning of fragmentation; the survival yield [10, 11, 14] is 50% or higher (i.e., the sum of fragment ion intensities is less than the intensity of the protonated molecule). This energy range may be useful for identifying low energy fragments (e.g., for identifying glycoforms). However, in practice, because of the low intensity fragments, tandem mass spectra in this collision energy range are of limited usefulness for analytical purposes. At higher collision energies the 3D ridge curves and becomes parallel with the collision amplitude, and loses altitude. This means that the ion trap is incapable of getting a spectrum that is similar to a medium (or high) energy spectrum obtained on a QQQ or QTOF instrument. As a rough guide, ion trap and QQQ spectra show qualitative similarities (SI > 0.90), whereas the protonated molecule has at least a few percent relative intensity. When all protonated molecules are decomposed by fragmentation, the MSMS spectra will be very different on quadrupole and ion trap instruments.

This is further illustrated by comparing the breakdown diagrams (energy resolved mass spectra) on the Waters QQQ and Bruker ion trap instruments in Figure 6. At low energy (15 eV collision energy and 0.5 V amplitude, respectively) the

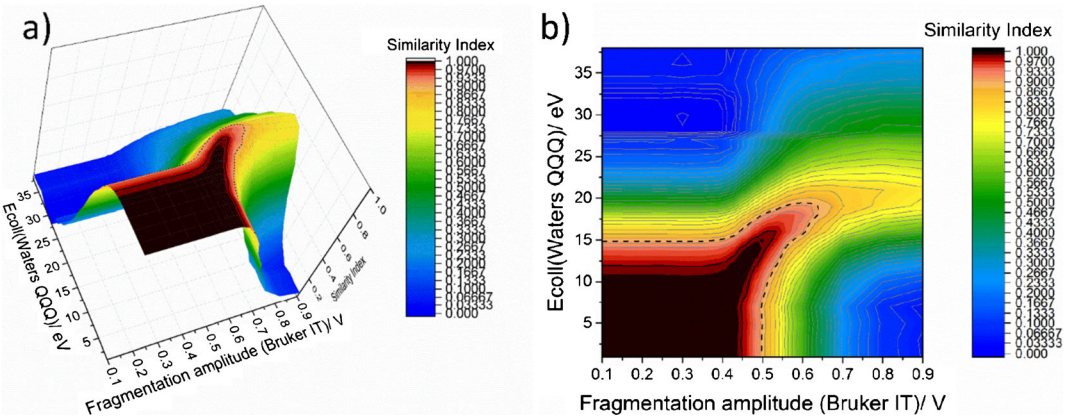


Figure 5. Similarity indices shown (**a**) as a 3D contour map, and (**b**) as a "heat map" for all combinations of collision energies determined on Waters QQQ and Bruker ion trap instruments

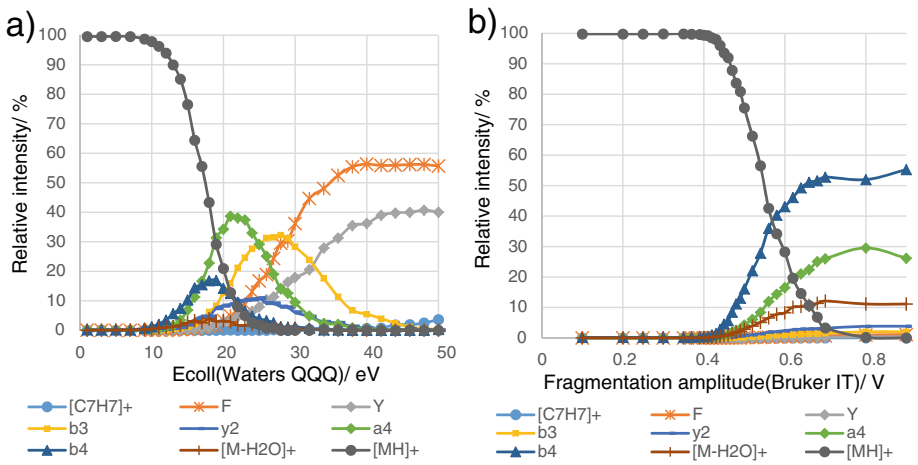


Figure 6. Breakdown curves of leucine enkephalin measured on (**a**) Waters QQQ and (**b**) Bruker ion trap instruments

major fragment on both instruments is the b_4 ion, and a minor low energy fragment is $[MH\text{-}H_2O]^+$ ion. At a slightly higher energy the a_4 fragment appears on both instruments. Further increase in the collision amplitude fragments all protonated molecules, but other fragments do not appear, and the product ratio changes only very slowly on the ion trap. In contrast, increasing the collision energy on the QQQ instrument creates a new set of fragments, and the spectra keep changing significantly up to ca. 35 eV collision energy.

This can be explained by the fundamental difference between QQQ (MS/MS in space) and ion traps (MS/MS in time) and not by mass discrimination effect. In QQQ instruments the parent ions could become more excited than in ion traps because of the limits of "slow heating" in the latter case [22]. Furthermore, in QQQ instruments the primary product ions also undergo CID, whereas in ion traps only the parent ion is excited. These differences make high activation energy and consecutive fragmentation processes possible in QQQ instruments, whereas they are nearly impossible in ion traps. These are the main reasons why QQQ and IT data are comparable only at low energy, when the survival yield is higher than ca. 50%.

Conclusions

In the present paper we have compared energy-dependent mass spectra on various instruments. The aim was to determine if, how, and to what degree is it possible to get a tandem mass spectrum, which is closely similar to another one taken on a different instrument. Although reproducibility is a commonly used and good measure to define small random errors, it is less adapted to compare similarity and differences among spectra, which may differ systematically. For this reason, we have used the dot product of the square root of peak intensities in two spectra to characterize spectral similarity (Equation (1)). This function is often used for comparing spectra; using the square root gives more weight for low intensity peaks, which are often important for databank search [15, 16].

It is well known that the collision energy is the most important single parameter, which influences the appearance of tandem mass spectra. We have measured collision energy-dependent tandem mass spectra of various compounds on several instruments. The similarity index (Equation (1)) between spectra taken on two instruments at all combinations of collision energies was calculated, and the results were shown in 3D contour maps and on 2D color-coded heat maps. Discussion is centered on leucine encephalin, as it is a peptide standard in mass spectrometry [13], has various fragments in a wide energy range, and spectra have well-described energy dependence. All other studied compounds yielded analogous results.

Comparison of the Waters QQQ and QTOF instruments (Figure 4) shows a fairly symmetric ridge on the 3D map, the top of the ridge indicating those collision energy combinations, which give the best spectral similarities. The top of the ridge shows high (>0.99) similarity indices in most of the collision energy range. This indicates that optimizing the collision energy is usually sufficient to get very similar spectra on the two instruments. The similarity in most cases is comparable to the long-term reproducibility. However, fragmentation of leucine encephalin is slightly, but systematically, different at around 30 eV collision energy, and this decreases the SI at the top of the ridge to 0.988. For lower mass compounds, the Waters QTOF and QQQ show larger difference in the high collision energy range, which is likely to be connected to the ion optic design of the "traveling wave."

Comparison between the two QQQ instruments (that of Waters and Agilent, Figure 2), somewhat to our surprise, showed significant differences: the direction of the ridge was not diagonal, and the Agilent QQQ produced relatively "hot" ions even at low collision energy (Figures 2 and 3a). These indicate major differences between the two instruments; some

likely due to the ESI source design. In spite of these, it was possible to get very good agreement between spectra taken on the two instruments by adjusting the collision energies. Spectral similarities better than 0.99 were possible to obtain in the full collision energy range.

In contrast, tandem mass spectra on the QQQ and ion trap instruments showed good similarity (SI > 0.99) only at low collision energy. This corresponds to the range where there is only little fragmentation (when the survival yield is higher than 50%, Figure 5.) For YGGFL on the Waters QQQ this means up to 15 eV; the actual collision energy depends both on the instrument and on the sample. The ridge in the 3D similarity index map curves and its height goes down at higher energies. In practice, this means that quadrupole and ion trap spectra will be significantly different when the survival yield is less than 50% (which is typical in most MS/MS studies). When the relative intensity of the precursor ion falls below a few percent, the spectra will be very different, with a few common fragments only (Figure 6).

We have found that the similarity index (square root dot product, Equation (1)) was a good way to compare spectra taken under different conditions. The 3D plots and 2D heat maps (Figures 2, 4, and 5) showing the similarity index for all combinations of collision energies is an excellent method for comparing two mass spectrometers. In the pharmaceutical industry, where reproducibility is a prime concern, this may be an excellent way to test if data obtained on two instruments may be comparable or not. This may be a key issue for selecting the best alternative, if an old instrument needs to be replaced.

If long-term reproducibility is an issue, we suggest scanning the collision energy, and determining the optimum using the similarity index (comparing the new, energy-dependent spectra to an old reference spectrum). We also suggest using YGGFL as a quality control standard [13] for energy-resolved studies, determining the tandem MS spectrum at the selected collision energy. This spectrum may be used later as a reference spectrum. In future experiments, the mass spectrometer should be tuned using YGGFL (by varying the collision energy) to get the best similarity index with the reference spectrum. This will result in better reproducibility than for example using the same tuning file on an instrument because it takes into account possible deposits in the ion source, misalignment of impurities on the quadrupole rods etc., which may vary in time. The use of YGGFL is advantageous, as it has fragment ions in a wide energy range, and its fragmentation characteristics are well known.

Acknowledgment

The authors gratefully acknowledge the Hungarian Research Fund (OTKA K109006 and OTKA NK83857) and Research and Technology Innovation Fund (VKSZ-12-1-2013-0001) for financial support. G.S. acknowledges the support by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

1. Crawford, L.R., Morrison, J.D.: Computer methods in analytical mass spectrometry. Identification of an unknown compound in a catalog. Anal. Chem. **40**, 1464–1469 (1968)
2. McLafferty, F.W., Zhang, M.Y., Stauffer, D.B., Loh, S.Y.: Comparison of algorithms and databases for matching unknown mass spectra. J. Am. Soc. Mass Spectrom. **9**, 92–95 (1998)
3. Stein, S.: Mass Spectral reference libraries: an ever-expanding resource for chemical identification. Anal. Chem. **84**, 7274–7282 (2012)
4. Yates III, J.R.: Pivotal role of computers and software in mass spectrometry—SEQUEST and 20 years of tandem MS database searching. J. Am. Soc. Mass Spectrom. **26**, 1804–1813 (2015)
5. Yang, X.Y., Neta, P., Stein, S.E.: Quality control for building libraries from electrospray ionization tandem mass spectra. Anal. Chem. **86**, 6393–6400 (2014)
6. Ji, C., Arnold, R.J., Sokoloski, K.J., Hardy, R.W., Tang, H.X., Radivojac, P.: Extending the coverage of spectral libraries: a neighbor-based approach to predicting intensities of peptide fragmentation spectra. Proteomics **13**, 75–765 (2013)
7. Dubey, R., Hill, D.W., Lai, S., Chen, M.H., Grant, D.F.: Correction of precursor and product ion relative abundances in order to standardize CID spectra and improve Ecom(50) accuracy for non-targeted metabolomics. Metabolomics **11**, 753–763 (2015)
8. Vékey, K.: Internal energy effects in mass spectrometry. J. Mass Spectrom. **31**, 445–463 (1996)
9. Josse, T., De Winter, J., Dubois, P., Coulembier, O., Gerbaux, P., Memboeuf, A.: A tandem mass spectrometry-based method to assess the architectural purity of synthetic polymers: a case of a cyclic polylactide obtained by click chemistry. Polym. Chem. **6**, 64–69 (2015)
10. Memboeuf, A., Jullien, L., Lartia, R., Brasme, B., Gimbert, Y.: Tandem mass spectrometric analysis of a mixture of isobars using the survival yield technique. J. Am. Soc. Mass Spectrom. **22**, 1744–1752 (2011)
11. Memboeuf, A., Nasioudis, A., Indelicato, S., Pollreisz, F., Kuki, Á., Kéki, S., van den Brink, O.F., Vékey, K., Drahos, L.: Size effect on fragmentation in tandem mass spectrometry. Anal. Chem. **82**, 2294–2302 (2010)
12. Schneider, B.B., Chen, D.D.Y.: Collision-induced dissociation of ions within the orifice-skimmer region of an electrospray mass spectrometer. Anal. Chem. **72**, 791–799 (2000)
13. Sztaray, J., Memboeuf, A., Drahos, L., Vekey, K.: Leucine enkephalin—a mass spectrometry standard. Mass Spectrom. Rev. **30**, 298–320 (2011)
14. Collette, C., Drahos, L., Pauw, E.D., Vékey, K.: Comparison of the internal energy distributions of ions produced by different electrospray sources. Rapid Commun. Mass Spectrom. **12**, 1673–1678 (1998)
15. Stein, S.E., Scott, D.R.: Optimization and testing of mass spectral library search algorithms for compound identification. J. Am. Soc. Mass Spectrom. **5**, 859–866 (1994)
16. Liu, J., Bell, A., Bergeron, J., Yanofsky, C., Carrillo, B., Beaudrie, C., Kearney, R.: Methods for peptide identification by spectral comparison. Proteome Sci. **5**, 3 (2007)
17. Tabb, D.L., MacCoss, M.J., Wu, C.C., Anderson, S.D., Yates, J.R.: Similarity among tandem mass spectra from proteomic experiments: detection, significance, and utility. Anal. Chem. **75**, 2470–2477 (2003)
18. Wan, K., Vidavsky, I., Gross, M.: Comparing similar spectra: From similarity index to spectral contrast angle. J. Am. Soc. Mass Spectrom. **13**, 85–88 (2002)
19. Li, Q.: Assigning significance in label-free quantitative proteomics to include single-peptide-hit proteins with low replicates. Int. J. Proteomics. **2010**, 15 (2010)
20. Lambert, J.-P., Ivosev, G., Couzens, A.L., Larsen, B., Taipale, M., Lin, Z.-Y., Zhong, Q., Lindquist, S., Vidal, M., Aebersold, R., Pawson, T., Bonner, R., Tate, S., Gingras, A.-C.: Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. Nat. Methods **10**, 1239–1245 (2013)
21. Mann, M., Kelleher, N.L.: Precision proteomics: The case for high resolution and high mass accuracy. Proc. Natl. Acad. Sci. **105**, 18132–18138 (2008)
22. McLuckey, S.A., Goeringer, D.E.: Slow heating methods in mass spectrometry. J. Mass Spectrom. **32**, 461–474 (1997)