Research Article

Journal of MASS SPECTROMETRY

Received: 14 April 2008 Accepted: 9 October 2008 Published online in Wiley Interscience: 16 January 2009

(www.interscience.com) DOI 10.1002/jms.1525

On the inter-instrument and the inter-laboratory transferability of a tandem mass spectral reference library: 2. Optimization and characterization of the search algorithm

Herbert Oberacher,[a]* Marion Pavlic,[a] Kathrin Libiseller,[a] Birthe Schubert,[a] Michael Sulyok,[b] Rainer Schuhmacher,[b] Edina Csaszar[c] and Harald C. Köfeler[d]

A sophisticated matching algorithm developed for highly efficient identity search within tandem mass spectral libraries is presented. For the optimization of the search procedure a collection of 410 tandem mass spectra corresponding to 22 compounds was used. The spectra were acquired in three different laboratories on four different instruments. The following types of tandem mass spectrometric instruments were used: quadrupole-quadrupole-time-of-flight (QqTOF), quadrupole-quadrupole-linear ion trap (QqLIT), quadrupole-quadrupole-quadrupole (QqQ), and linear ion trap-Fourier transform ion cyclotron resonance mass spectrometer (LIT-FTICR). The obtained spectra were matched to an established MS/MS-spectral library that contained 3759 MS/MS-spectra corresponding to 402 different reference compounds. All 22 test compounds were part of the library. A dynamic intensity cut-off, the search for neutral losses, and optimization of the formula used to calculate the match probability were shown to significantly enhance the performance of the presented library search approach. With the aid of these features the average number of correct assignments was increased to 98%. For statistical evaluation of the match reliability the set of fragment ion spectra was extended with 300 spectra corresponding to 100 compounds not included in the reference library. Performance was checked with the aid of receiver operating characteristic (ROC) curves. Using the magnitude of the match probability as well as the precursor ion mass as benchmarks to rate the obtained top hit, overall correct classification of a compound being included or not included in the mass spectrometric library, was obtained in more than 95% of cases clearly indicating a high predictive accuracy of the established matching procedure.

Supporting information may be found in the online version of this article.

Keywords: tandem mass spectrometry; library; search algorithm

Introduction

An important field of application for mass spectrometry (MS) is structure elucidation of (bio)organic compounds.[1] For this purpose, fragment ion mass spectra are interpreted to gather the structural information inherently included in them. Except for biopolymers such as peptides[2] and oligonucleotides[3] fragment ion mass spectra are poorly predictable because of the complexity and variety of processes occurring during the collision-induced dissociation of a molecule. For this reason, spectra of compounds to be identified are often treated simply as molecular fingerprints, for which an identical or at least structural related compound has to be found within a mass spectral database via library search.

In general, there are two methods for searching spectral libraries: identity searches and similarity searches.[4] In an identity search the spectrum of the unknown compound is assumed to be part of the spectral library, and only experimental variability prevents a perfect match of unknown and reference spectra. In the more sophisticated similarity search, the spectral collection does not contain a spectrum of the unknown compound. The library search returns structures belonging to library spectra that show some degree of similarity with the unknown spectrum.[5–8]

A number of algorithms and software tools have been developed for matching two mass spectra and were reviewed.[9–12] Two of them have gained widespread use due to their availability in commercial mass spectrometric data systems. One of these is the probability-based matching system which uses peak occurrence statistics in its spectral comparison logic.[13] The other one is the dot-product-algorithm which compares unknown and reference spectra by calculating the cosine of the angle between their vector representatives.[10] These algorithms were optimized for

* *Correspondence to: Herbert Oberacher, Institute of Legal Medicine, Innsbruck Medical University, Muellerstrasse 44, 6020 Innsbruck, Austria. E-mail: herbert.oberacher@i-med.ac.at*

a *Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria*

b *Center for Analytical Chemistry, Department IFA-Tulln, University of Natural Resources and Applied Life Sciences, Vienna, Austria*

c *Max F. Perutz Laboratories, University Departments at the Vienna Biocenter, Department of Biochemistry, University Vienna, Vienna, Austria*

d *Core Facility for Mass Spectrometry, Center for Medical Research, Medical University of Graz, Graz, Austria*

Journal of MASS SPECTROMETRY

the comparison of mass spectra generated by electron ionization (EI). Because of the use of standardized conditions, EI mass spectra are very reproducible. Irrespective of the instrumental platform used for acquisition almost identical peak patterns are observed. Fragment ion mass spectra obtained from tandem mass spectrometry (MS/MS) tend to be less reproducible than EI spectra. Especially, the inter-instrument comparability of MS/MS-spectra seems to be in question.[14–17] The commonly applied search algorithms have difficulties to correctly identify a compound if the number of fragment ions and/or the corresponding signal intensities vary significantly between sample and reference spectra. Hence, libraries consisting of MS/MS-spectra have found only limited acceptance for compound identification so far.

In this report, a sophisticated matching algorithm is presented which has been developed for identity search within a tandem mass spectral library containing 3759 MS/MS-spectra of 402 compounds developed on a quadrupole-quadrupole-time-of-flight (QqTOF) instrument.[18] A set of 410 MS/MS-spectra corresponding to 22 compounds included in the reference library was used to optimize the performance of the search algorithm. The MS/MS-spectra were collected in three different laboratories using four different instruments. Several features were implemented to obtain the highest number of correct assignments. For statistical evaluation of the match reliability the dataset was extended with 300 spectra corresponding to 100 compounds not included in the reference library. Sensitivity as well as the specificity of the library search approach exceeded a value of 0.95 clearly indicating a high predictive accuracy of the established matching procedure.

Experimental Section

Reference library

The reference library was recently developed on a QqTOF instrument (Qstar XL, Applied Biosystems, Foster City, CA) and contained 3759 MS/MS-spectra of 402 compounds. The experimental setup has been described in detail elsewhere.[18] For each reference compound, product ion spectra were acquired at 10 different collision-energy values between 5 and 50 eV. Because of possible saturation effects, and to avoid false positive matching of the precursor ion with product ions from alternative compounds, all signals within a ±4.0 amu window around the m/z of the precursor ion were deleted from the reference spectra obtained. To remove unspecific noise and centroiding artifacts, reference spectra were filtered.

Test compounds

The sample set consisted of 122 substances. Compounds, 22 in number, represented the set of positive controls (reference spectra included in the reference library). Compounds, 100 in number, were designated negative controls (no reference spectrum included in the reference library). All compounds were more or less randomly selected. A complete summary of the test compounds is available as a Table in the Electronic Supporting Information. Prior to analysis, the chemical identity of the samples had been checked by GC/MS as described previously.[18] Additionally, MS/MS-spectra of 98 compounds (25 positive controls and 73 negative controls) downloaded from the METLIN Metabolite Database[19] were used to check the predictive accuracy of the matching procedure. It is important to note that all 98 compounds were different from the 22 test compounds sent to the participating laboratories. A complete summary of these compounds is also available as a Table in the Electronic Supporting Information.

Instrumentation

Tandem mass spectra of the 100 negative controls were collected on a QqTOF instrument (Qstar XL, Applied Biosystems). The experimental setup has been described in detail elsewhere.[18] For each compound, fragment ion mass spectra were collected at three different collision energy settings. Tandem mass spectra of the 22 positive controls were acquired in 3 other laboratories on the following types of tandem mass spectrometric instruments: QqTOF (Qstar Pulsar i, Applied Biosystems), quadrupole-quadrupole-linear ion trap (QqLIT, QTrap 4000, Applied Biosystems), quadrupole-quadrupole-quadrupole (QqQ, TSQ Quantum Ultra, Thermo Fisher Scientific, Waltham, MA), and linear ion trap-Fourier transform ion cyclotron resonance mass spectrometer (LIT-FTICR, LTQ-FT Ultra, Thermo Fisher Scientific, Waltham, MA). Three devices were classified as tandem-in-space instruments (QqQ, QqLIT, QqTOF), and one as tandem-in-time instrument (LIT-FTICR). The QqLIT was operated in two different scanning modes: in product ion scan (pi) as well as in enhanced product ion scan (epi) mode. In both operational modes, precursor ions were selected in the first quadrupole and fragmented in the collision cell (= second quadrupole). The third quadrupole was either operated as quadrupole (pi) or as LIT (epi) and was used to scan the fragment ions. On the LIT-FTICR instrument, product ions were generated in the LIT and were either analyzed at low resolution in the LIT or at high resolution in the FTICR. Test samples were weighed and dissolved in 0.1% aqueous acetic acid solution containing 50% (v/v) acetonitrile before analysis. Depending on the performance of the different instruments as well as on the compound-specific ionization efficiencies the concentrations of the sample solutions varied from 0.02 to 10 µg/ml. Samples were directly infused into the mass spectrometer. On each single instrumental platform, tandem mass spectra were acquired at three different collision-energy values. Additionally, on the QqLIT in epi mode, a single spectrum under collision energy-spread conditions was measured. A total number of 418 fragment ion mass spectra representing the positive controls were collected. A more detailed description of the experimental conditions can be found in the companion paper.

Data handling

In each laboratory, MS/MS-spectra were centroided and exported as txt-files. Each txt-file contains information about the precursor ion mass, a list of the observed fragment ions (mass-to-charge ratios (m/z), and the corresponding relative signal intensities. All files are available for review from the authors upon request. Before matching the collected sample spectra to the library, each of the 718 (418 positive controls and 300 negative controls) fragment ion mass spectra was compared visually with the corresponding reference spectra. Within some spectra, noticeable discrepancies to the corresponding reference spectra were uncovered that resulted either from contamination or sample mix-up. These spectra were eliminated from the sample set. The remaining 710 (410 positive controls and 300 negative controls) spectra were matched against the established library.[18] Furthermore, MS/MS-spectra of another 98 compounds were downloaded from the Internet (http://metlin.scripps.edu). Within this dataset, 25 compounds have already been included in

the reference library (positive controls), and 73 substances represented negative controls, of which no reference spectra were available in the spectral library. Automated library search was performed with a program written in ActivePerl 5.6.1 (Active State Corporation, Vancouver, Canada). Calculations, as described below, were performed on a personal computer running Windows XP operating system (1.7 GHz Pentium, 1.0 GB RAM). Receiver operating characteristic (ROC) curves were created in SPSS 14.0.1 (SPSS Incorporated, Chicago, IL).

The library search strategy

Depending on the applied experimental conditions, the number of fragment ions and/or the corresponding signal intensities can vary between compound-specific MS/MS-spectra. Common library search algorithms were developed and optimized for the comparison of highly reproducible EI spectra. Thus, they often malfunction if the identity of compounds needs to be proven via the comparison of MS/MS-spectra. Extending our recent work,[18] we present here a sophisticated procedure dedicated to the identification of an unknown by finding similarity, and/or identity, between its fragment ion spectrum and a collection of fragment ion mass spectra stored in a library. The following problems were solved during different development stages of the library search procedure: (1) Changes in experimental conditions can have a significant impact on the contribution of distinctive decomposition pathways to the overall fragmentation. Hence, the optimized library search algorithm should show a high tolerance towards differences between compound-specific fragmentation patterns measured and stored. (2) For each compound, the library contains several reference spectra collected at different collision-energy settings. A number of match probabilities are obtained for each compound. Nevertheless, each compound should be represented only once in the obtained hit-list. Thus, the reference-specific match probabilities should be combined into one compound-specific value.

The principal steps of the developed library search approach are outlined in Fig. 1. The measured product ion mass spectrum of an unknown compound represents the input for library search. The spectrum is compared with all mass spectra stored in the library (Fig. 1, step 1). In each case, the similarity is determined. The estimation of similarity starts with the identification of ions that are present in both of the two spectra compared. They are called matching fragments (*mf*), (Fig. 1, step 1a). For a match, the difference of the *m/z* values must be smaller than a user-defined value ($\Delta = 0.01-0.1$ amu). To avoid the occurrence of a false positive match between the precursor ion and product ions specific for an incorrect compound, all signals within a ± 4.0 amu window around the *m/z* of the precursor ion are excluded from matching. Next, the reference spectrum-specific match probability (*mp*) is calculated (Fig. 1, step 1b). Recently, we have proposed the following formula, herein referred to as the published formula, for calculating *mp*:[18]

$$mp = \frac{mf^4}{f_{unknown} \cdot f_{reference} \cdot \left(\sum |int\%_{unknown} - int\%_{reference}|\right)^a}$$

Where, $f_{unknown}$ and $f_{reference}$ are the number of fragments in the input and the reference spectrum, respectively, and $\Sigma|int\%_{unknown} - int\%_{reference}|$ represents the sum of the intensity differences observed for matching fragments. As the relative intensities can show some degree of variability, less weight was assigned to the peak intensities by setting the exponent, *a*, to 0.25. The *mp* value increases with increasing correlation between the two spectra compared. Although the published formula showed good performance,[18] the aim of this study was to develop an advanced formula for *mp* determination (see below).

As the mass spectral library consists of MS/MS-spectra that have been collected at several different collision-energy values, for each single reference compound a number of *mp* values are obtained. The task was to combine all the different reference compound-specific *mp* values to one value that specifies the similarity between the unknown and the reference compound. Thus, the reference compound-specific *mp* values are averaged to yield the compound-specific average match probability (*amp*), (Fig. 1, step 2). To facilitate comparison, *amp* is converted into the relative average match probability (*ramp*), (Fig. 1, step 3). Consequently, single *ramp* values range between 0 and 100. A high compound-specific *ramp* value indicates high similarity between the unknown and the reference compound. A list is gathered as output of the search algorithm that is sorted in order of decreasing *ramp* (Fig. 1, step 4). The substance with the highest *ramp* is considered to represent the unknown compound. Next, the monoisotopic mass of the best matching compound is checked for accordance with the monoisotopic mass of the precursor ion (Fig. 1, step 5). If the monoisotopic masses do not agree with each other, identity is excluded. Only the presence of some structural similarity between the unknown and the best matching reference compound can be considered. Provided that the top hit passes this final check the correct compound should have been identified with high probability. Nevertheless, a false positive match may occur with some probability, which largely depends on the magnitude of the calculated *ramp* value (see below).

Results and Discussion

Starting point for optimizing the search algorithm – evaluation of the performance of the existing library search procedure

The set of positive controls used for evaluating the efficiency of the matching procedure consisted of 410 MS/MS-spectra collected from 22 samples in 3 different laboratories on 4 different mass spectrometers applying 3–4 different collision-energy settings for fragmentation. The goal was to retrieve the correct compound as top hit in as many spectral comparisons as possible. The spectra were matched against a library containing 3759 MS/MS-spectra of 402 compounds developed on a QqTOF instrument (Qstar XL, Applied Biosystems).[18] For *mp* calculation the above-mentioned published formula was used. Δ was set to 0.1 amu for low-resolution instruments (QqQ, LIT, QqLIT), and to 0.01 amu for high-resolution instruments (QqTOF, LIT-FTICR). To characterize the quality of the matching procedure, the number of correct assignments and the average *ramp* values of the correct compounds irrespective of their position in the hit-list were determined. The obtained results are summarized in Fig. 2. For all instrumental platforms used, the percentage of correct assignments was clearly below 95%. Best results were obtained on the high-performance mass spectrometric platforms (66.7% on the QqTOF; 71.2% on the LIT-FTICR). The lowest number of correct hits was gathered from data acquired on the QqLIT in *epi* mode (26.3%). The average *ramp* values ranged between 6.1 and 50. Using the published formula, the library search procedure clearly failed the quality check. Thus, some kind of upgrade was necessary

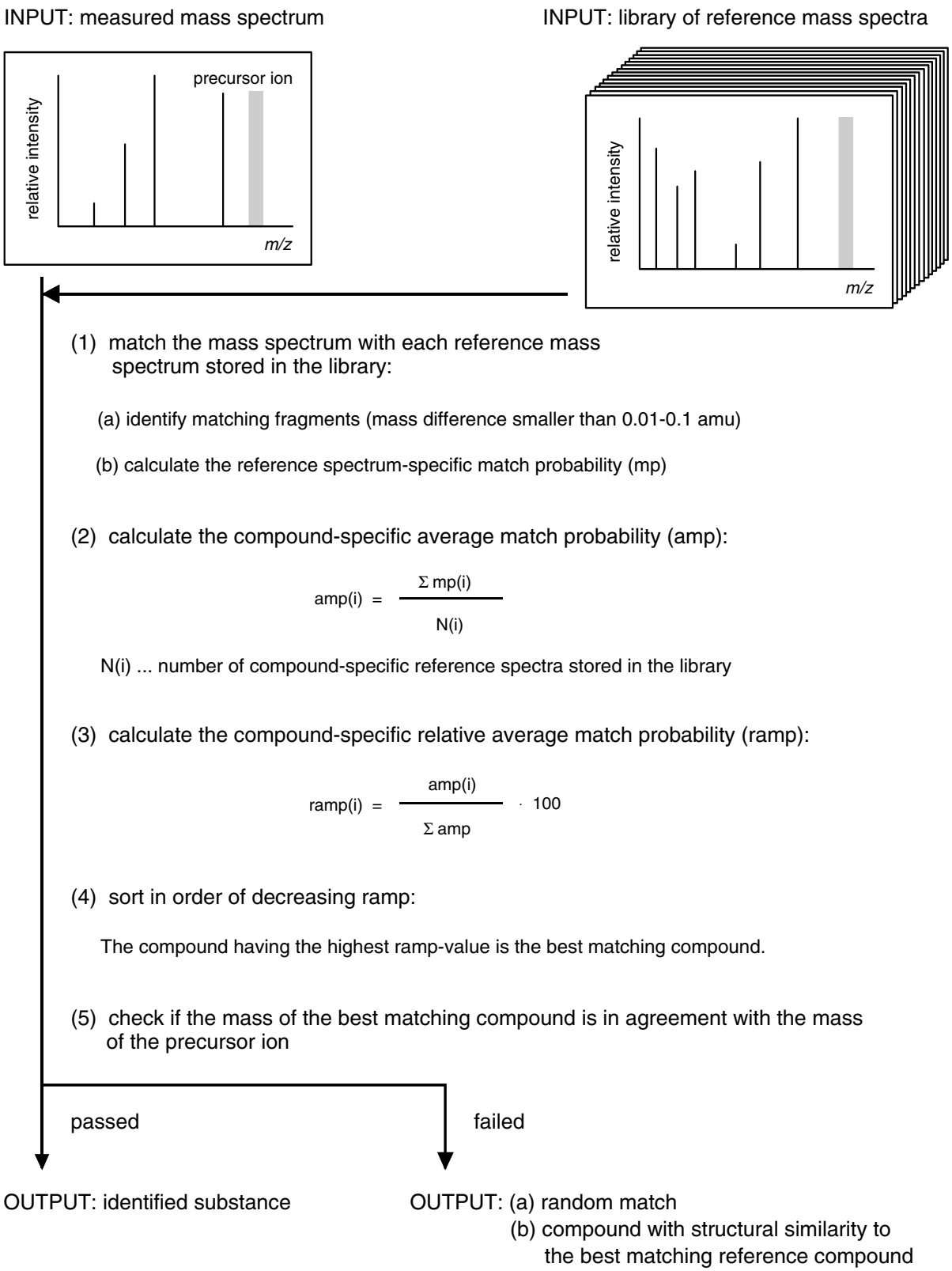

Figure 1. Schematic drawing showing the principles of the automated database search approach. Details are described in the text.

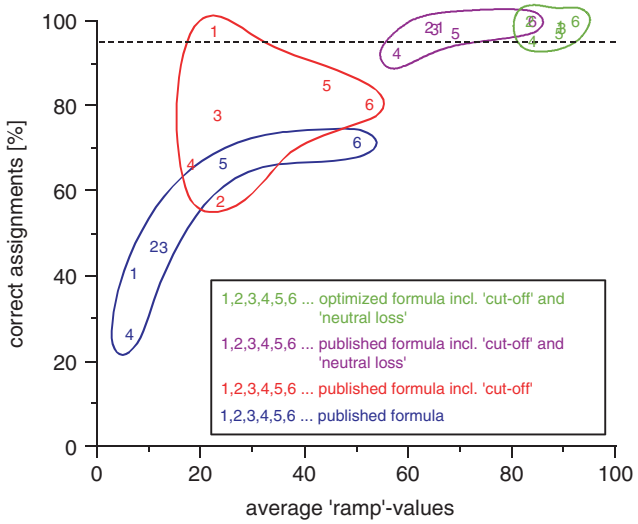

Figure 2. Correlation between the number of correct assignments and the average *ramp* values. Mass spectra collected on 6 different instrumental platforms, (1) QqQ, (2) LIT, (3) QqLIT-pi, (4) QqLIT-epi, (5) QqTOF, (6) LIT-FTICR, were matched against the database. Different strategies were used to determine the reference spectrum-specific match probabilities, and are explained in the text. The dashed line is used to indicate a success rate of 95%.

to improve the ability of the search procedure to connect a sample spectrum with the correct reference compound.

1st upgrade – introduction of an intensity cut-off

The first feature that was implemented into the search algorithm was an intensity cut-off. As signal intensities largely depend on the applied collision-energy value, a dynamic threshold was used to exclude putative noise within sample spectra from matching. The intensity threshold was set by multiplying the intensity of the most intense fragment ion by a user-defined factor. Only those signals with intensities above this threshold were considered as compound-specific and were allowed to match to signals in the reference spectra. The implementation of the intensity cut-off had a positive impact on the matching efficiency (Fig. 2). By using a factor of 0.05 to calculate the cut-off, the percentage of correct assignments as well as the average *ramp* values were more than doubled for spectra collected on the QqQ as well as on the QqLIT in *epi* mode. The lowest number of correct assignments was obtained for the LIT spectra (57.6%). The highest percentage of correct calls was gathered from the QqQ data (97.3%).

2nd upgrade – introduction of neutral loss-counting

For determining the similarity between two MS/MS-spectra, only matching ions were considered so far. To further improve the efficiency of the search algorithm, information retrieval was extended to an additional source of compound-specific information within a fragment ion mass spectrum: neutral losses. Each neutral loss implies information which is characteristic for a certain substructure of the corresponding precursor ion and can be considered a valuable source of structural information to increase the sensitivity of the search algorithm in identifying the correct compound. The query for neutral losses was integrated into the library search routine in the following way: If in the course of spectral comparison an *mf* is identified, conformity of the precursor ions is assessed. If the difference between the measured precursor ion mass, and the precursor ion mass of the reference compound is smaller than Δ the *mf* will be counted twice. As shown in Fig. 2, the counting of neutral losses clearly improved the efficiency of the search procedure. Except for the QqLIT in *epi* mode the percentages of correct answers exceeded the 95% limit. We believe that in some cases, due to the combination of low sample concentrations (0.02 µg/ml) with rather low collision-energy values, MS/MS-spectra of rather low specificity were obtained which were not correctly matched causing the overall somewhat lower efficiency of the library search approach observed for the QqLIT. The average *ramp* values ranged between 57.9 and 84.0. The implementation of the neutral loss feature significantly increased the ability of the search algorithm to identify the correct compound among structurally closely related compounds.

3rd upgrade – optimization of the formula used for *mp* calculation

Other developments concerned the optimization of the formula used to calculate the *mp* value. In a first step, the dependence of the efficiency of the library search approach using the published formula on factor, *a*, was studied. The *a* value is used to modulate the impact of the $\Sigma|int\%1 - int\%2|$-term on the *mp* value. More weight will be given to the peak intensities if larger *a* values are used. The results of the study are summarized in Fig. 3. Overall best performance was observed by setting the *a* value to 0.50. Using that distinct *a* value, the average number of correct assignments exceeded the 95% limit irrespective of the instrumental platform used. Overall 97.8% of all searches yielded a correct result. Lower and higher *a* values can have some adverse effects on the search efficiency. The obtained results clearly suggest that some kind of signal intensity information must be implemented into the formula used for *mp* calculation. The weight of the signal intensity to the overall *mp* value, however, needs to be carefully tuned. Generally, lower weighting seems to be favored (Fig. 3).

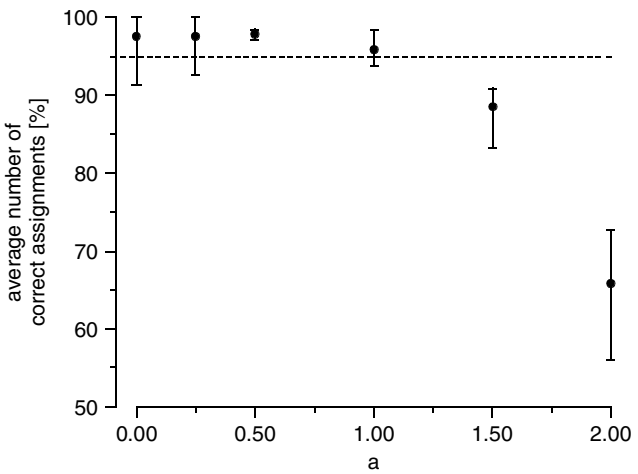

Figure 3. Dependence of the average number of correct assignments on the *a* term. The *a* term is part of the published formula[18] used for the calculation of the reference spectrum-specific *mp* and determines the impact of observed signal intensity differences on the *mp* value. Dots represent the overall average number of correct assignments obtained for the complete set of fragment ions. The bars are used to indicate the observed variability of the average number of correct assignments on different instrumental platforms. The dashed line is used to indicate a success rate of 95%.

Journal of MASS SPECTROMETRY

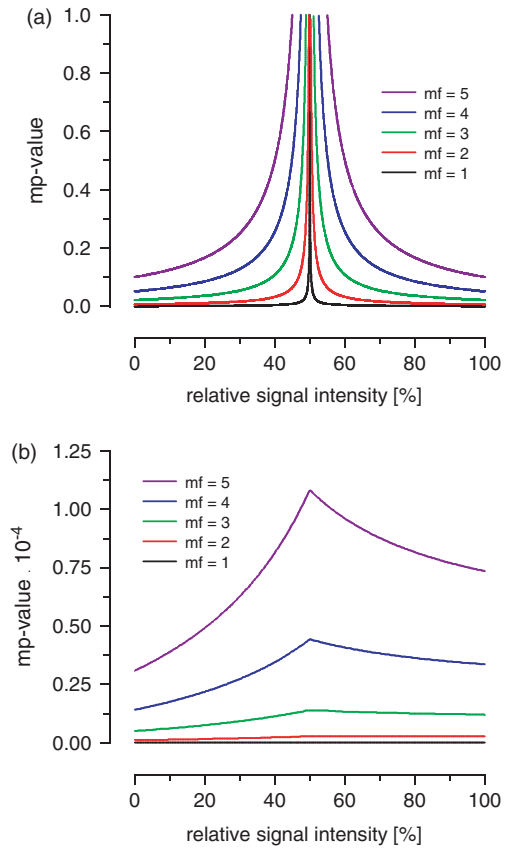

Figure 4. Effects of different terms of the published and the optimized formulae on the magnitude of the calculated *mp* value. In the depicted example, the sample spectrum consists of 5 fragment ions having identical relative signal intensities that vary between 0% and 100%. The reference spectrum consists of 5 fragment ions having identical peak heights fixed at 50%. A match of the two spectra can yield up to five *mf*. (a) The published formula (a = 1) and (b) the optimized formula (b = 4, c = 1.25, d = 2) have been used to calculate the corresponding *mp* values.

Although after the implementation of the intensity cut-off and the neutral loss, counting the search efficiency was satisfying using the published formula for calculating *mp* values, there was still some room for improvement in regard to the achieved *ramp* values. By setting the *a* value to 0.50 average *ramp* values between 60 and 90 were obtained. Obviously, the published formula used for calculating the *mp* values needed an upgrade. To find a formula with a better performance, simulations of the dependence of the *mp* value on the parameters used to calculate the *mp* value were conducted. An example of such a simulation is shown in Fig. 4. The sample spectrum consisted of 5 fragment ions having identical relative signal intensities that varied between 0 and 100%. The reference spectrum also consisted of 5 fragment ions having peak heights fixed to 50%. A match of the two spectra yielded up to five *mf*. In Fig. 4(a) the results obtained by using the published formula ($a = 1$) for *mp* calculations are depicted. The simulations clearly revealed that the major disadvantage of the published formula is the fact that a small sum of the intensity differences can give rise to an unrealistically large *mp* value even in cases where *mf* is comparably low (Fig. 4(a)). Thus, two spectra that do not show much overlap in regard of *mf* might have a higher *mp* value than two spectra perfectly matching if by chance the sum of the signal intensity differences is very small. To overcome this problem, the formula for *mp* calculation was reconfigured. Multiplication of the parameters was substituted by summation. New parameters were introduced to further increase specificity. $\Sigma int\%_{unknown}$ and $\Sigma int\%_{reference}$ represent the sum of the signal intensities in the unknown and in the reference spectra, respectively. $\Sigma|m/z_{unknown} - m/z_{reference}|$ equals to the sum of the absolute *m/z* differences observed for matching fragment ions. The following equation, referred herein as the optimized formula, was obtained:

$$mp = \frac{mf^b \cdot \left(\sum int\%_{unknown} + 2 \cdot \sum int\%_{reference}\right)^c}{(f_{unknown} + 2 \cdot f_{reference})^d + \sum |int\%_{unknown} - int\%_{reference}| + \sum |m/z_{unknown} - m/z_{reference}|}$$

A trial-and-error strategy was used for parameter optimization. The summed intensity ($\Sigma\ int\%_{reference}$) and number of fragments ($f_{reference}$) are multiplied by 2 in order to weight the reference spectrum more heavily. The exponent *b* was set to 4, *c* to 1.25, and *d* to 2. As can be deduced from Fig. 4(b), *mf* has the highest impact on the magnitude of *mp*. Nevertheless, if two sample spectra match equally well to a reference spectrum in regard of *mf*, the individual correlations of the signal intensities will decide upon the quality of the matches. Small intensity differences give high *mp* values (Fig. 4(b)). It is important to note, however, that for the optimized formula the impact of the sum of signal intensity differences on the *mp* value is in all cases much smaller than it has been observed for the published formula. Out of two spectra having the same *mf* and the same sum of signal intensity differences, a larger *mp* value will be obtained for the spectrum having an overall larger sum of the signal intensities. Furthermore, spectra showing signals of very low intensities might be disfavored over spectra having a lower *mf* but exhibiting more intense fragments (Fig. 4(b)). This feature is advantageous, as it may compensate to some extent unspecific matching with low intense signals in noisy spectra.

By applying the optimized formula the platform-specific percentages of correct assignments changed only slightly in comparison to the results obtained with the published formula and ranged between 95.0 and 100% (Fig. 2). Overall, 98.1% of all search results were correct. The average *ramp* values, however, increased significantly and were in all cases greater than 83 (Fig. 2). The observed performance suggests that the optimized formula represents a sensitive tool for unequivocal identification of the correct compound within a mass spectral library.

4th upgrade – optimization of the cut-off level

The intensity cut-off was introduced to increase the efficiency of the search algorithm. The factor to calculate the exclusion limit was arbitrarily set to 0.05. Although excellent results were obtained, an optimization of the cut-off factor setting was performed. The dependence of the search efficiency on that parameter can be deduced from the results shown in Fig. 5. For all values between 0.05 and 0.10, the percentage of correct assignments was beyond 95% irrespective of the instrumental platform used. Overall best results were obtained for a cut-off factor of 0.10. Values beyond 0.10 tend to exclude compound-specific fragment ions from matching and are not favored. Depending on the spectral quality, even cut-off factors below 0.05 might give good results. Very low cut-off levels (<0.01), however, carry the risk of matching noise to signals specific for incorrect compounds giving rise to false positive hits, and should therefore be avoided.

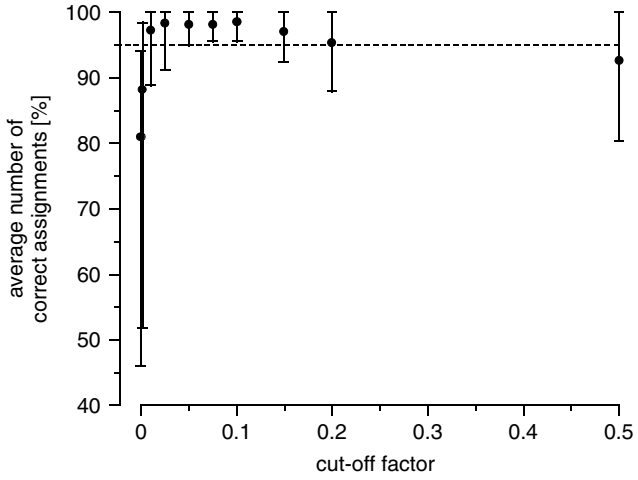

Figure 5. Optimization of the cut-off level. The cut-off level is used to exclude low-abundant fragment ions within a sample spectrum from matching with signals in the reference files. Dots represent the overall average number of correct assignments obtained for the complete set of fragment ions. The bars are used to indicate the observed variability of the average number of correct assignments on different instrumental platforms. The dashed line is used to indicate a success rate of 95%.

5th upgrade – optimization of Δ

Δ is defined as the maximum tolerated mass difference between two matching fragment ions. Arbitrarily, Δ was set to 0.1 amu for low-resolution instruments (QqQ, LIT, QqLIT), and to 0.01 amu for high-resolution instruments (QqTOF, LIT-FTICR). Although excellent results were obtained with these values, optimization of the Δ-settings was performed. The results are depicted in Fig. 6. For high-resolution instruments Δ was varied between 0.001 and 2. For all values tested between 0.0025 and 0.25 the average numbers of correct assignments were beyond 95%. Overall best results were obtained by setting Δ to 0.01. For low-resolution instruments, Δ was varied between 0.01 and 2. Best results were obtained by setting Δ to 0.1. These data show that 0.1 is a good compromise value, regardless of the instrumental platform used.

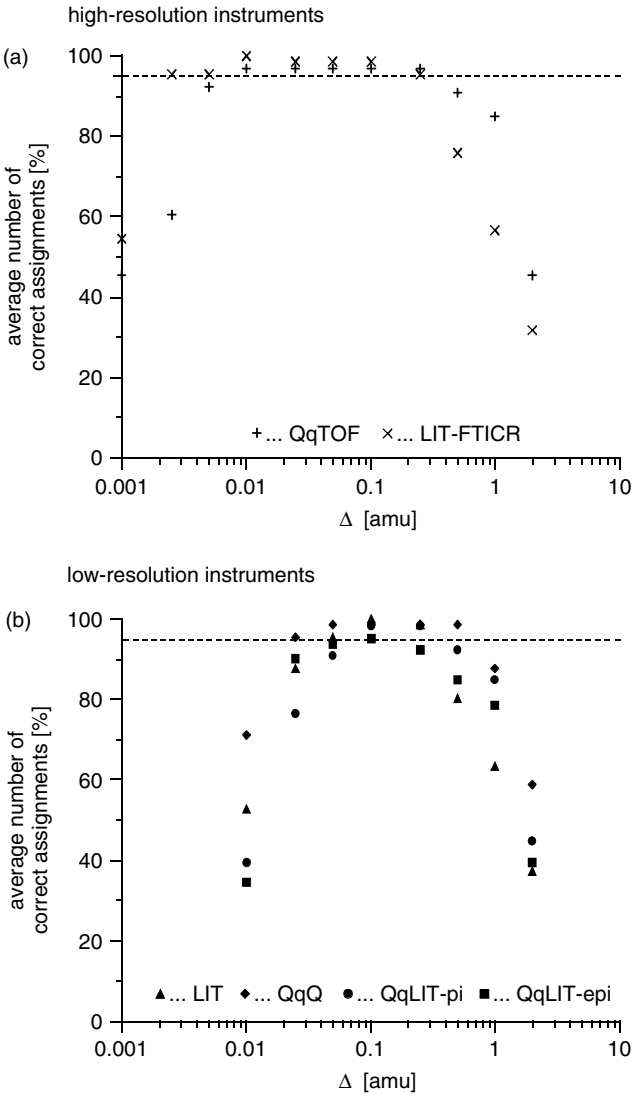

Figure 6. Optimization of Δ for (a) high-resolution, and (b) low-resolution instruments. Two fragment ions are considered as matching ions if the difference of their *m/z* values is smaller than Δ. The dashed line is used to indicate a success rate of 95%.

Statistical evaluation of the performance of the upgraded matching procedure using ROC curves

An ROC curve is a plot of the true positive fraction (= sensitivity) *versus* the false positive fraction (= 1-specificity), and therefore, represents a tool to evaluate the predictive accuracy of the established matching procedure.[20,21] For statistical evaluation of the reliability of the developed library search approach 710 spectra were used. The set of positive controls consisted of 410 spectra collected in 3 different laboratories; 300 spectra acquired on a QqTOF represented the set of negative controls. Sensitivity- and (1-specificity)-values were analyzed as a function of the *ramp* value. For the set of positive controls, the status of all matches where a wrong compound was retrieved as top hit was changed to incorrect, irrespective of the obtained ramp value. The obtained ROC curve is shown in Fig. 7. In a first set of calculations, Δ was set to 0.1 amu for matching the negative controls, and only the magnitude of the *ramp* value was used as the parameter to classify the obtained top hit as correct or incorrect. A perfectly working search algorithm would have yielded an ROC curve that coincided with the left and top sides of the plot. A search strategy that is completely useless would have given a straight line from the bottom left corner to the top right corner. The obtained ROC curve lay between these two extremes. A global assessment of the performance of the search algorithm is given by the area under the ROC curve (AUC). A perfect search algorithm would yield an AUC of 1.0. With the aforementioned parameter settings, an AUC-value of 0.967 ± 0.015 (95% confidence limits) was obtained, which indicates a high predictive accuracy of the developed search algorithm. Furthermore, the ROC curve was used to determine the optimal cut-off point for the *ramp* value at which optimal sensitivity and specificity would be achieved. For this purpose, the point on the curve closest to the left and top side of the plot was determined. An optimal cut-off for the *ramp* value of 50.0 was obtained. A sensitivity of 0.951 and a specificity of 0.937 were achieved. By implementing the m/z of the precursor ion as additional qualifying criterion to exclude putative false positive hits (Fig. 7), the specificity increased to 0.980. Moreover, after the reduction of Δ from 0.1 to 0.01 (Fig. 7), which was recommended for matching QqTOF spectra to the library, specificity reached a value of 0.990. Out of the set of 300 negative controls only 3 xanthinol-specific

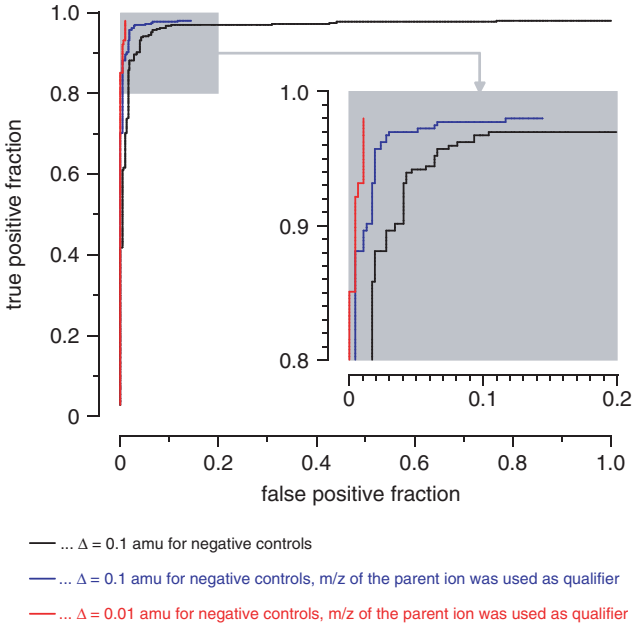

Figure 7. ROC curve used to evaluate the predictive accuracy of the established matching procedure. The set of positive controls consisted of 410 spectra corresponding to 22 compounds acquired on several different instrumental platforms; 300 spectra corresponding to 100 compounds acquired on a QqTOF instrument represented the set of negative controls.

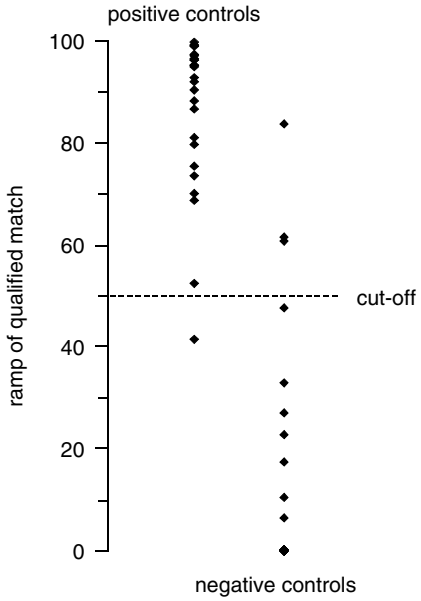

Figure 8. Dot diagram of the ramp values obtained for qualified matches. The dataset consisted of 98 MS/MS-spectra (25 positive and 73 negative controls) that were downloaded from the METLIN Metabolite Database. Only those matches where the m/z value of the best matching compound was in agreement with the m/z value of the precursor ion were called qualified matches.

spectra were incorrectly assigned to nalorphine. Out of the set of 410 positive controls only 3 spectra (0.7%) would have been incorrectly matched to a wrong compound, and for just 12 spectra (2.9%) no unequivocal identification of the correct compound would have been obtained. Overall correct inclusion/exclusion of a compound from being part of the mass spectrometric library was obtained in 97.5% of cases clearly indicating a high predictive accuracy of the established matching procedure.

Assessing the search reliability with spectra downloaded from METLIN metabolite database

To further prove the platform-independence of the presented mass spectral library, MS/MS-spectra of 98 compounds (25 positive controls and 73 negative controls) were matched against the reference library. The spectra were downloaded from the METLIN Metabolite Database,[19] which is a public, Web-based database designed for archiving, visualization, and analysis of metabolite data. The retrieved MS/MS data was converted into a format appropriate for library search before use. The Δ-value was set to 0.1 amu. An obtained search result would have been classified as putatively correct match if all the following criteria were fulfilled: (1) Only the best matching compound is considered. (2) The m/z value of the best matching compound must not differ more than $\pm\Delta$ from the m/z value of the precursor ion. (3) The ramp value obtained for the best matching compound must exceed a value of 50.0.

Out of the set of positive controls (Fig. 8) all compounds, except reserpine, were unequivocally identified. Thus, the sensitivity of the library search approach reached 0.960. The spectrum of reserpine matched equally well to reserpine and phenprocoumon. A closer inspection of the downloaded reserpine-specific mass spectral data revealed that all fragment ion masses showed a negative deviation from the expected values. The absolute deviation ranged between 0.12 and 0.86 amu, and decreased proportionally with the m/z value, which suggests that in this case improper instrument calibration was the source of the ambiguous search result. Out of the set of negative controls (Fig. 6) only three compounds passed all three selection criteria and were classified wrongly as correctly identified. Thus, the specificity of the library search approach reached 0.959. All in all, the high predictive accuracy observed for a random dataset downloaded from a public database represents a further hint for the platform-independence of the developed mass spectral library search approach.

Conclusions

A number of computer-assisted mass spectral library search procedures have been developed over the last decades. Some of these approaches have found broader applicability due to their implementation into commercially available software packages. The vast majority of search algorithms were optimized for finding similarity and/or identity between highly reproducible EI spectra. In contrast to EI spectra, MS/MS-spectra tend to be less reproducible. Changes within experimental conditions can have a significant impact on the contribution of distinctive decomposition pathways to the overall fragmentation. Research focused on the development strategies to increase the comparability of MS/MS-spectra instead of the development of advanced library search procedures. Thus, sophisticated tools optimized for identity searches within MS/MS-spectral libraries are missing, which may be the real reason why MS/MS-spectral libraries have found only limited acceptance for compound identification so far.

Among other features, the ultimate library search algorithm should show a high tolerance towards differences between compound-specific fragmentation patterns measured and stored. In the present work, several steps were taken to improve

the efficiency of a recently presented MS/MS-spectral library search approach. For performance optimization, a collection of 410 MS/MS-spectra corresponding to 22 library compounds was used. The MS/MS-spectra were collected in 3 different laboratories using 6 different instrumental platforms, and matched against a library containing 3759 MS/MS-spectra of 402 compounds developed on a QqTOF instrument. Several sophisticated features were implemented into the search algorithm to increase the reliability of the matching procedure. For statistical evaluation of the identification performance the dataset was extended with 300 spectra corresponding to 100 compounds not included in the reference library. Sensitivity as well as the specificity of the library search approach exceeded a value of 0.95, clearly indicating a high predictive accuracy of the established matching procedure. As far as we know, no other search algorithm has, so far, reached a similar level of performance in the context of MS/MS-spectral library search. Future work, however, will show if the ongoing increase in the number of database entries will have a major impact on the efficiency and transferability of the mass spectral library.

Acknowledgements

The authors wish to thank Applied Biosystems/MDS Sciex for support. Furthermore, financial support by the Austrian Research Promotion Agency (FFG, Österreichisches Sicherheitsforschungs-Förderprogramm KIRAS – eine Initiative des Bundesministeriums für Verkehr, Innovation, Technologie (BMVIT), Projekt 813786) is acknowledged.

Supporting information

Supporting information may be found in the online version of this article.

References

[1] S. A. McLuckey, J. M. Wells. Mass analysis at the advent of the 21st century. *Chemical Reviews* **2001**, *101*, 571.
[2] P. Roepstorff, J. Fohlman. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry* **1984**, *11*, 601.
[3] S. A. McLuckey, G. J. Van Berkel, G. L. Glish. Tandem mass spectrometry of small, multiply charged oligonucleotides. *Journal of the American Society for Mass Spectrometry* **1992**, *3*, 60.
[4] M. Zürcher, J. T. Clerc, M. Farkas, E. Pretsch. General theory of similarity measures for library search systems. *Analytica Chimica Acta* **1988**, *206*, 161.
[5] S. E. Stein. Chemical substructure identification by mass spectral library searching. *Journal of the American Society for Mass Spectrometry* **1995**, *6*, 644.
[6] K. Varmuza, W. Werther. Mass spectral classifiers for supporting systematic structure elucidation. *Journal of Chemical Information and Computer Sciences* **1996**, *36*, 323.
[7] K. Klagkou, F. Pullen, M. Harrison, A. Organ, A. Firth, G. J. Langley. Approaches towards the automated interpretation and prediction of electrospray tandem mass spectra of non-peptidic combinatorial compounds. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 1163.
[8] W. Demuth, M. Karlovits, K. Varmuza. Spectral similarity versus structural similarity: mass spectrometry. *Analytica Chimica Acta* **2004**, *516*, 75.
[9] F. W. McLafferty. Computer identification of mass spectra. *Pure and Applied Chemistry* **1978**, *50*, 197.
[10] S. E. Stein, D. R. Scott. Optimization and testing of mass spectral library search algorithms for compound identification. *Journal of the American Society for Mass Spectrometry* **1994**, *5*, 859.
[11] C. S. Tong, K. C. Cheng. Mass spectral search method using the neural network approach. *Chemometrics and Intelligent Laboratory Systems* **1999**, *49*, 135.
[12] M. A. Grayson. *The History and Heritage of Scientific and Technological Information Systems: Proceedings of the 2002 Conference*. American Society for Information Science and Technology: Philadelphia, **2004**.
[13] F. W. McLafferty, R. H. Hertel, R. D. Villwock. Probability based matching of mass spectra. Rapid identification of specific compounds in mixtures. *Organic Mass Spectrometry* **1974**, *9*, 690.
[14] M. J. Bogusz, R. D. Maier, K. D. Kruger, K. S. Webb, J. Romeril, M. L. Miller. Poor reproducibility of in-source collisional atmospheric pressure ionization mass spectra of toxicologically relevant drugs. *Journal of Chromatography A* **1999**, *844*, 409.
[15] A. W. Bristow, K. S. Webb, A. T. Lubben, J. Halket. Reproducible product-ion tandem mass spectra on various liquid chromatography/mass spectrometry instruments for the development of spectral libraries. *Rapid Communications in Mass Spectrometry* **2004**, *18*, 1447.
[16] M. Gergov, W. Weinmann, J. Meriluoto, J. Uusitalo, I. Ojanperä. Comparison of product ion spectra obtained by liquid chromatography/triple-quadrupole mass spectrometry for library search. *Rapid Communications in Mass Spectrometry* **2004**, *18*, 1039.
[17] R. Jansen, G. Lachatre, P. Marquet. LC-MS/MS systematic toxicological analysis: comparison of MS/MS spectra obtained with different instruments and settings. *Clinical Biochemistry* **2005**, *38*, 362.
[18] M. Pavlic, K. Libiseller, H. Oberacher. Combined use of ESI-QqTOF-MS and ESI-QqTOF-MS/MS with mass-spectral library search for qualitative analysis of drugs. *Analytical and Bioanalytical Chemistry* **2006**, *386*, 69.
[19] C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan, G. Siuzdak. METLIN: a metabolite mass spectral database. *Therapeutic Drug Monitoring* **2005**, *27*, 747.
[20] F. W. McLafferty. Performance prediction and evaluation of systems for computer identification of spectra. *Analytical Chemistry* **1977**, *49*, 1441.
[21] M. H. Zweig, G. Campbell. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* **1993**, *39*, 561.