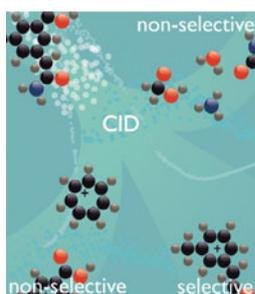


RESEARCH ARTICLE

The (Un)Certainty of Selectivity in Liquid Chromatography Tandem Mass Spectrometry

Bjorn J. A. Berendsen,¹ Linda A. M. Stolker,¹ Michel W. F. Nielen^{1,2}¹RIKILT, Wageningen University and Research Centre, Akkermaalsbos 2, 6708WB, P.O. Box 230, 6700AE Wageningen, The Netherlands²Laboratory of Organic Chemistry, Wageningen University, Dreijenplein 8, 6703 HB Wageningen, The Netherlands

Abstract. We developed a procedure to determine the “identification power” of an LC-MS/MS method operated in the MRM acquisition mode, which is related to its selectivity. The probability of any compound showing the same precursor ion, product ions, and retention time as the compound of interest is used as a measure of selectivity. This is calculated based upon empirical models constructed from three very large compound databases. Based upon the final probability estimation, additional measures to assure unambiguous identification can be taken, like the selection of different or additional product ions. The reported procedure in combination with criteria for relative ion abundances results in a powerful technique to determine the (un)certainty of the selectivity of any LC-MS/MS

analysis and thus the risk of false positive results. Furthermore, the procedure is very useful as a tool to validate method selectivity.

Key words: Liquid chromatography, Mass spectrometry, Selectivity, Identification, Confirmation, Uncertainty, Precursor ion, Product ion, Retention time

Received: 10 July 2012/Revised: 12 September 2012/Accepted: 12 September 2012/Published online: 11 December 2012

Introduction

In residue analysis of food contaminants, sports doping, and forensic sciences the qualitative aspect of a method is of prime importance: is the identity of a contaminant confirmed? There is no universally-accepted definition of confirmation or identification [1, 2]. In our view, the term “confirmation” presumes knowledge on the identity of the compound present (e.g., from previous analysis). A positive confirmation indicates that the result is not against the presumptions made: the compound shows the same characteristics as the reference compound selected following the previous analysis, but it does not exclude the possibility of the presence of another compound showing the same characteristics. In contrast, in “identification” no a priori presumption on the structure of the substance present is made and, thus, identification requires that all other substances are excluded, so that the reported substance is the only possible candidate [3].

For the confirmatory aspect of a method, the main validation parameter is selectivity, which is defined as “the power of discrimination between the analyte and closely

related substances...” [4]. Although some regulations have been established for the confirmation of the identity of a compound [4–10], these are all based on comparison of the hypothesized identity with a single reference standard, thereby ignoring that another compound on this planet might yield a similar result. No validation procedures are described to express this aspect of the uncertainty of the qualitative result in LC-MS/MS analysis. External quality control, like proficiency tests, is far too limited to adequately study selectivity of all compounds and all matrices.

Guidelines for confirmatory analysis using mass spectrometry with electron impact (EI) ionization were reported for the first time in 1978 [11]. Based on the empirical study of product ion spectra databases, using diethyl stilbestrol as a model compound, it was concluded that the minimum number of ions in an EI mass spectrum needed for identification of a compound is three. This observation became known as the “three-ion criterion” [1]. Although three ions should be sufficient, the choice of ions used for confirmatory purposes is critical to prevent false positive findings [11–14].

When using GC- or LC-MS a minimum number of identification points is required for “unequivocal identification” [4]. For regulated substances (group B [15]), three identification points are considered adequate, and for banned substances (group A [15]), four identification points are required [4, 16]. When full scan single MS is used, a minimum

Electronic supplementary material The online version of this article (doi:10.1007/s13361-012-0501-0) contains supplementary material, which is available to authorized users.

Correspondence to: Bjorn J. A. Berendsen; e-mail: bjorn.berendsen@wur.nl

number of four ions must be present in the spectrum having a relative intensity of >10 %, which is a more stringent criterion compared with the “three-ion criterion” [11]. If low resolution MS/MS is used, four identification points are earned, for example when monitoring one precursor ion and two product ions, and five identification points are earned when monitoring two precursor ions, each having one product ion. In all cases, the relative ion abundance of a set of product ions (ion ratio) must comply with the criteria established in 2002/657/EC [4] before the sample can be considered noncompliant (positive). Slightly deviating criteria for relative ion abundances have been established by other organizations [4–10]). Previously, a quantitative assessment of the probability of false positive results in EI-MS was carried out and it was concluded that more stringent criteria are needed to prevent false positive results [17]. This underlines that MS/MS criteria do not originate from statistically supported data, but are the result of experts’ judgments and, therefore, it cannot be concluded that one criterion is better than the other. Moreover, this will depend on the point of view: more stringent criteria result in an increased probability of false negative results, whereas less strict criteria result in an increased probability of a false positive result [3].

A critical review on ion ratio criteria was presented by Kaufmann et al. [18, 19] and false positive findings based upon EU criteria were reported [18, 20, 21]. It was concluded that although the current criteria provide a good starting point for confirmation of the identity of a compound, an additional procedure is needed to determine the (un)certainty of this qualitative result.

In this paper, we present a procedure to judge the (un)certainty of the selectivity of a method using LC coupled to triple-quadrupole (QqQ) MS operated in the MRM acquisition mode based on the precursor ion mass, the product ions selected, and the RT on a specific reversed phase (RP) chromatographic system. With this procedure, the uniqueness of the monitored LC-MS/MS characteristics can be estimated, and on the basis of this result, it can be decided if additional data (e.g., a third product ion or more extensive sample clean-up is needed for a trustworthy identification of the compound present.

Methods

Databases

Three databases were constructed on basis of which the (un)certainty of the selectivity of an LC-MS/MS identification procedure was determined.

Database A contains molecular mass data and was extracted from the eMolecules database [22], which contains over 5.3 million commercially available compounds. Because residue analysis primarily focusses on small molecules, only the compounds having a molecular weight of 100–1000 were taken into account. Of these, all compounds containing at least one carbon in combination with at least oxygen, nitrogen,

sulphur, and/or phosphorous were selected because these compounds can most likely be ionized by electrospray ionization. All salts and metal containing compounds were removed from the dataset, which resulted in a total of over 5.1 million compounds. Of these compounds, the mono-isotopic nominal mass was calculated.

Database B contains product ion spectra data and is constructed from the product ion spectra of 3629 compounds obtained from our own research, the Massbank database [23] and literature [24–31]. Of 3629 compounds, the positive ion mode electrospray ionization collision-induced dissociation (CID) MS spectra were converted to binary spectra with a cut-off level at 5 % relative intensity and thus information on ion abundances was removed. The combination of precursor ion mass and product ion mass, and the corresponding neutral losses, were included in the database, resulting in a total of 18314 entries.

Database C contains theoretical RT data on a specific high resolution RP LC system and is constructed from the qualitative structure-retention relationship (QSRR) model presented in this paper. For 451 compounds, the relevant physical and chemical parameters were used to calculate the theoretical RT.

Modeling the Retention Time

UPLC grade water and methanol were obtained from Biosolve (Valkenswaard, The Netherlands). Ammonium formate was obtained from Merck (Darmstadt, Germany). Reference standards of 126 model compounds consisting of veterinary drugs, growth promoters, and pesticides were obtained from Sigma-Aldrich (St. Louis, MO, USA) and Ultra Scientific (North Kingstown, RI, USA).

The LC system consisted of a vacuum degasser, autosampler, and a Waters (Milford, MA, USA) model Acquity binary pump equipped with a Waters Acquity HHS-T3 analytical column of 2.1 × 100 mm, 1.7 μm placed into a column oven at 30 °C. The gradient (solvent A, water containing 2 mM ammonium formate, pH 3.5; solvent B, methanol/water (9:1, vol/vol) containing 2 mM ammonium formate): 0–1.0 min, 0 % B, 1.0–11.0 min, linear increase to 100 % B, with a final hold of 0.5 min, operating at a flow rate of 0.4 mL min⁻¹. The injection volume was 10 μL.

Detection was carried out using a Bruker Daltonics (Billerica, MA, USA) MicroQTOF MS in the positive electrospray ionization (ESI) mode. The operating parameters were: end plate offset, -500 V; capillary voltage, 3.5 kV; nebulizer, 2.0 bar; dry gas, 8.0 L min⁻¹; dry temperature, 200 °C.

Of 126 compounds the experimental RTs in the analytical system were determined. Furthermore, of each of these compounds, 22 physical and chemical parameters were obtained from literature: BCF at pH=5.5, K_{oc} at pH=5.5, polarizability and molecular volume were obtained from ChemSpider [32]; Log P, Log D at pH=3.0 and Log D at pH=5.5 were obtained from SciFinder [33]; total energy, Connolly accessible area, Connolly molecular area, Con-

nolly solvent excluded volume, molecular mass, ovality, molecular refractivity, partitioning coefficient, Balaban index, molecular topological index, polar surface area, shape attribute, total connectivity, total valance connectivity, and Wiener index were calculated from BioChem3D Ultra 12.0 after MM2 energy optimisation [34]. Extreme correlations among parameters were studied using GenStat software [35], and high correlations were excluded before further analysis. Next, using Genstat, the optimal parameters were selected to be included in the QSSR model to obtain the best relation between the parameters and the observed RT. Using the selected parameters, multiple linear regression was carried out to establish an optimal linear model.

Using this model, the theoretical RT of an additional 325 randomly selected compounds were calculated. All compounds for which a theoretical RT below 0.6 min was calculated, being the dead time of the system, set at 0.6 min. From these data plus the theoretical RTs of the 126 compounds used to build the model, the probability distribution of theoretical RTs on the selected chromatographic system was constructed.

Proposed Procedure

A measure for selectivity is the probability of the occurrence of a compound showing the same LC-MS/MS characteristics as the compound of interest: the probability of an interference, $P(I)$. $P(I)$ is calculated based upon independent probability matching:

$$P(I) = P(M_{pc}) * \max(P(M_{pd1}), P(M_{nl1})) * \max(P(M_{pd2}), P(M_{nl2})) * P(RT), \quad (1)$$

in which $P(M_{pc})$ is the probability of the occurrence of the precursor ion, $P(M_{pd})$ the probability of the occurrence of the product ions of transition 1 and 2, $P(M_{nl})$ the probability of the occurrence of a neutral loss of transition 1 and 2, and $P(RT)$ the probability of the occurrence of a RT in the chromatographic system. $\max[P(M_{pd}), P(M_{nl})]$ is the maximum of $P(M_{pd})$ or $P(M_{nl})$ to represent a worst case situation. These parameters are determined based on the databases and empirical models presented in this paper. Next, $P(I)$ is compared to an a priori established criterion to rank the method selectivity (e.g., as sufficient or insufficient). Based on the result, corrective measures can be taken to obtain additional selectivity like the selection of another or an additional product ion. A flow chart of the proposed procedure is presented in Figure 1.

Results and Discussion

An important parameter influencing the (un)certainty of the confirmation using MS is the power of fragmentations [1, 11]. European Union (EU) legislation indicates that “the selected diagnostic ions should not exclusively originate from the same

part of the molecule” [4], but even though clear differences in the selectivity of ion transitions were reported in literature [36], in the regulations very nonselective transitions like the loss of water or ammonia during CID are considered equally selective as compound-specific fragmentations. The UK guidelines for urine drugs testing mention product ions that are considered insufficiently selective, viz. $m/z < 50, 58, 86, 91, \text{ and } 105$ [9], and the Society of Forensic Toxicology indicated the low selectivity of the loss of water [5].

McLafferty et al. [37] and Pesyna et al. [38, 39] regarded the probability of ions to occur in a mass spectrum. From a large collection of GC-EI-MS spectra, the occurrence of ions between m/z 29 and 400 were modeled indicating that ions at m/z 29, 39, 41, 43, 55, 57, 73, 91, and 105 have a high probability of being a high abundant ion in a EI mass spectrum, in contrast to ions at m/z 37, 38, 49–53, and 62–66. On the obtained data, they carried out probability matching based upon the probability theory which states that if a number of events occur with a certain probability, the probability of all these events to occur simultaneously is the multiplication of all the individual probabilities. This overall probability was considered a measure for the uniqueness of a spectrum and, thus, for selectivity.

Probability Matching of Selectivity

Selectivity is a parameter that can be expressed in terms of the uniqueness of a compounds' characteristics, in other words in terms of the probability of the occurrence of another compound possessing the same characteristics, $P(I)$. If $P(I)$ is high, the probability of an incorrect confirmation result is high and, thus, the method selectivity is considered low, whereas if $P(I)$ is low the system used can be considered as highly selective. The selectivity for identification purposes using low resolution LC-QqQ-MS/MS, and thus the probability of an incorrect identification, is primarily determined by: (1) the combination of the selected precursor and product ions, (2) the RT of the compound and therefore the probability of co-eluting interfering compounds, and (3) the selectivity of the sample preparation method [40, 41].

According to EU criteria [4], using low resolution MS/MS, at least two transitions from precursor ion to product ion should be monitored to obtain a sufficient number of identification points. The selectivity of the monitored transitions is directly related to the m/z of the precursor ion (s), M_{pc} , and the product ions, M_{pd} . In theory, the most exact calculation of the probability of the occurrence of a combination of two transitions, $P(MS)$, is given by applying dependent probability calculations. In that case, $P(MS)$ is the probability of M_{pd2} , given that the M_{pd1} occurs and that both originate from the selected precursor ion(s). This can only be determined empirically from a very large MS/MS database, and the final result will be limited by the total number of spectra included in this database. For instance, if database B, containing 3629 spectra, is used and a unique combination

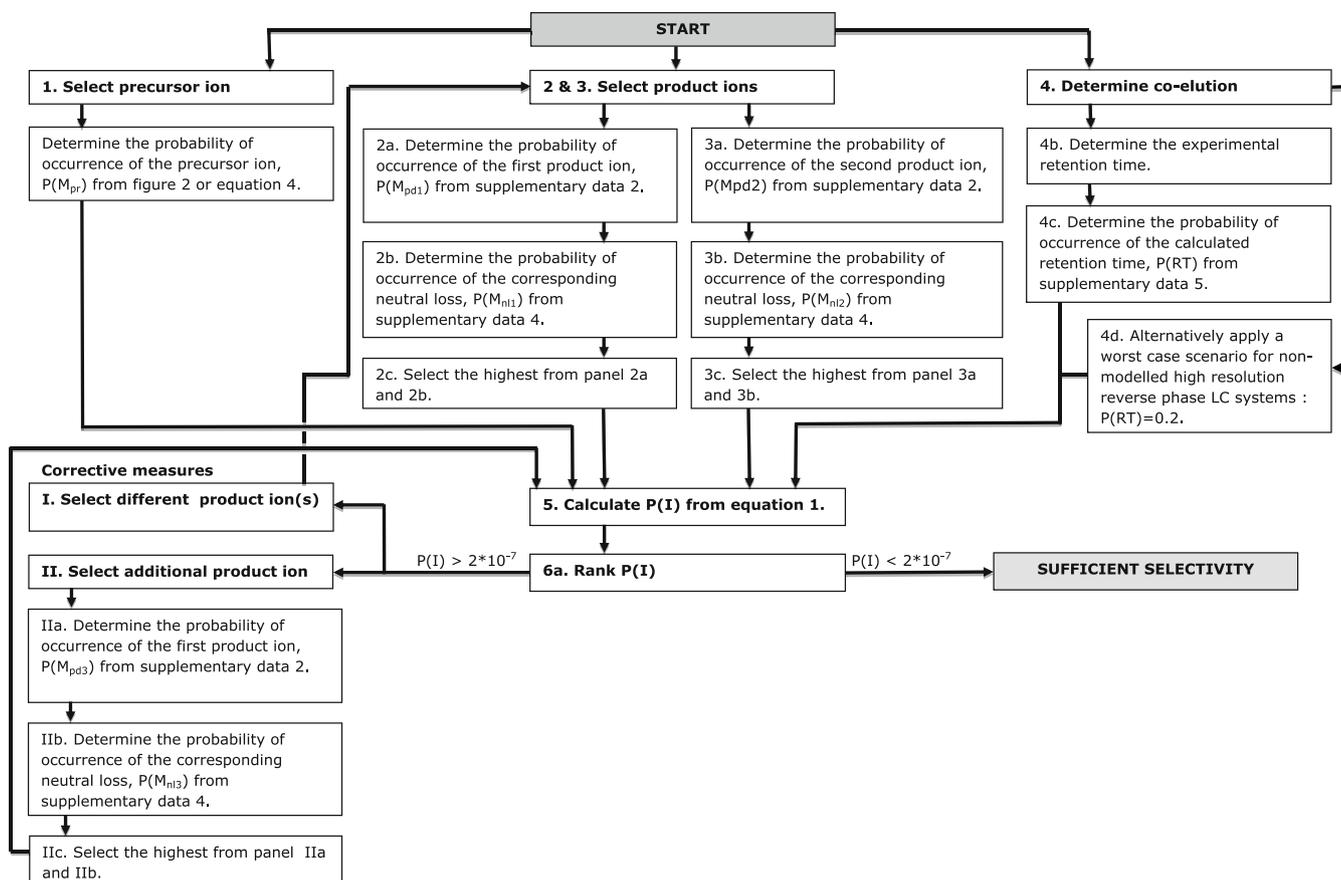


Figure 1. Flow chart of the proposed procedure including possible corrective measures

of M_{pc} , M_{pd1} , and M_{pd2} is found, it can be concluded that P (MS) is $2.8 \cdot 10^{-4}$ or, in other words, 1 in 3629. This is a severe underestimation of the selectivity of the method and can only be overcome by increasing the total number of spectra in the database. However, if a result of for example 1 in 5,000,000 is considered to be the minimum to conclude that sufficient selectivity is obtained, at least 5,000,000 product ion spectra should be included in the database, which is far off reality for the time being. Therefore, independent probability matching is proposed here, in which the individual probabilities of M_{pc} , M_{pd1} , and M_{pd2} are multiplied. The selectivity of the precursor ion(s) is given by the probability of the occurrence of compounds showing the same precursor ion mass: $P(M_{pc})$. The selectivity of the product ions is related to the product ion mass (M_{pd}) itself, but also to the resulting neutral loss ($M_{pc} - M_{pd}$). For instance, a product ion of m/z 400 has a low probability, but if it originates from a precursor ion of m/z 418 and, thus, if a neutral loss of water (18 Da) is involved, this transition should be considered nonselective. In the proposed worst case approach, the selectivity of the product ion is determined by selecting the highest of either $P(M_{pd})$ or the probability of the corresponding neutral loss, $P(M_{nl})$. As a result, the probability of the occurrence of a compound

having the same MS characteristics as the compound of interest is given by:

$$P(MS) = P(M_{pc}) * \max(P(M_{pd1}), P(M_{nl1})) * \max(P(M_{pd2}), P(M_{nl2})) \quad (2)$$

These parameters can all be derived from the constructed databases A and B.

Additional selectivity is obtained from the liquid chromatographic system. From the regression model, based on 11 physical and chemical properties, the theoretical RT of any compound, including its uncertainty, can be estimated. Based on the theoretical RT of 451 compounds included in database C, the probability of co-elution with other compounds can be determined. Then the total probability of the occurrence of the selected characteristics, $P(I)$ is calculated by:

$$P(I) = P(MS) * P(RT) \quad (3)$$

A disadvantage of the application of independent probability statistics is that dependency of the data is assumed

not to exist. Obviously, some dependencies do exist, for example, singly charged precursor ions always result in product ions at lower m/z . Therefore, no exact quantification of $P(I)$ is obtained in this work, but rather a very good estimate for a worst case scenario. The result should, therefore, not be considered as a quantitative measure, but as a ranking tool for selectivity (e.g., insufficient, sufficient, high). The calculated value of $P(I)$ in relation to the ranking of selectivity depends on fitness for purpose [1] and other factors affecting method selectivity like the preceding sample preparation procedure. In this work, the applied sample preparation procedure is not taken into account in calculating $P(I)$. A very selective sample preparation procedure (e.g., the use of bio-affinity solid phase extraction) will decrease the probability of interfering signals and thus a higher $P(I)$ could be acceptable, whereas for very generic sample preparation techniques, which are increasingly being applied [42], a lower $P(I)$ is needed. Furthermore, if a derivatization procedure is used in the sample preparation procedure (e.g., a hydrolysis using nitrobenzaldehyde as is common for nitrofurans analysis [21], all product ions or neutral losses related to nitrobenzaldehyde should be considered less selective than calculated from the method reported here, since an excessive number of compounds containing a nitrobenzaldehyde moiety will be present in the derivatized solution and, thus, database B would not be representative for this specific case. Note that in this approach, the relative ion abundances are not taken into account, whereas this would usually result in additional certainty on the identity of the compound. Therefore, the method presented here points to a worst case scenario and gives an overestimation of the true $P(I)$.

Precursor Ion Probability

The probability distribution of the m/z of the protonated molecules of all entries in database A is presented in Figure 2. Because database A includes many amino acids, peptides, sugars, fatty acids, etc., the most common matrix interferences are taken into consideration. It is observed that, in database A, an m/z of 387 is most common having a probability of 0.0052. Consequently, interfering signals are more likely for compounds in the mass range around m/z 400 compared with the lower and higher mass range. Therefore, the detection of compounds with M_{pr} around m/z 400 is considered to be less powerful than detection of compounds with an M_{pr} of for example 200 or 600. The obtained probability distribution is in good agreement with the distribution reported by Little et al. [43], who used ChemSpider [32] as a reference. Kind and Fiehn, who used the Pubchem database [44] as a reference, found the highest probability at a slightly higher mass and reported a longer tail at the high mass range [45]. This might be a result of the presence of complexes and compounds containing metal atoms present in the Pubchem database, which have been removed from database A. At around 500 Da, a non-linear

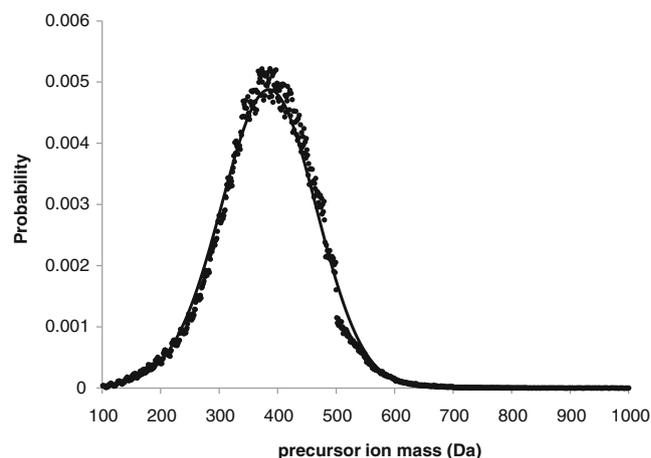


Figure 2. Probability distribution of the mass of the protonated molecules included in database A ($n > 5,000,000$). Solid line, fit according to equation 4

discontinuity is observed in the dataset. This is likely a result of the content of the eMolecules database, rather than a true lower relative occurrence of compounds at this specific mass.

$P(M_{pr})$ can be empirically determined from database A. For application of the method without having access to the original database, the probability distribution was modeled assuming binomial distributed data and using a logit link function. By doing so, $P(M_{pr})$ can be calculated by:

$$P(M_{pr}) = \frac{\exp(-19.8 + 0.067M_{pr} - 0.000084M_{pr}^2 + 405M_{pr}^{-1})}{1 + \exp(-19.8 + 0.067M_{pr} - 0.000084M_{pr}^2 + 405M_{pr}^{-1})}, \quad (4)$$

in which M_{pr} is the m/z of a precursor ion.

Product Ion Probability

From database B, the probability distribution of the product ions is presented in Figure S1 (Supplementary Data 1) for four categories of precursor ions: m/z 100–200, m/z 200–300, m/z 300–400, and >400 . As an example, the product ion data for the precursor ions category m/z 200–300 are presented in Figure 3a. These precursor ion categories were defined to yield at least 500 product ion spectra per group. From the product ion data for each precursor ion category a continuous model was constructed assuming binomial distributed data using a logit link function, to facilitate calculation of $P(M_{pd})$ (see Supplementary Data 2). Some product ions show an exceptional high or an exceptionally low probability compared to this model (high residual) and for these cases the probabilities are individually listed in Table S1 and S2 in Supplementary Data 2.

In the low mass range (M_{pr} 100–200), the probability of the occurrence of a product ion around m/z 100 is the highest, in the mass range 200–300 this is m/z 140 and at the

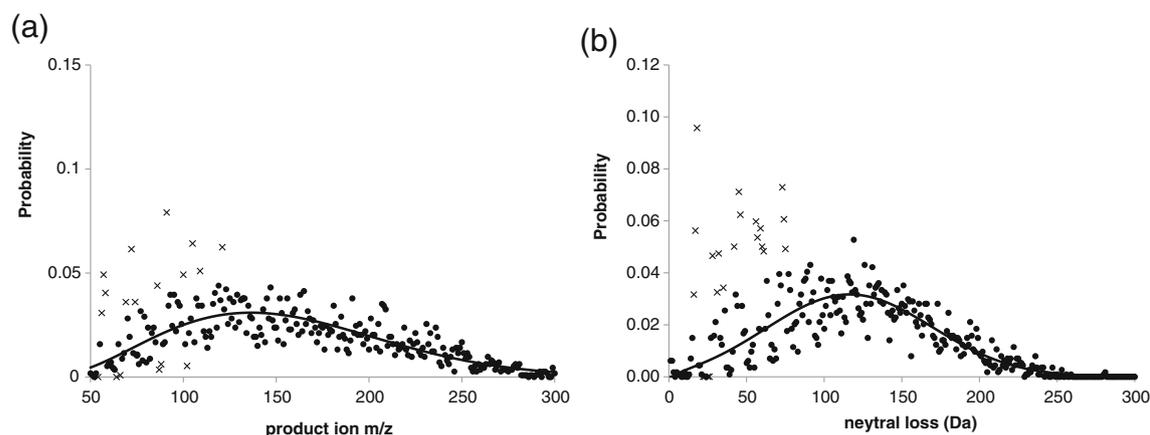


Figure 3. Probability distribution of **(a)** product ion masses and **(b)** neutral losses for precursor ions of m/z 200–300. Data points showing a high residual from the constructed model are indicated with a cross. Solid line, fit according to the equations as presented in supplementary data 2 and 4

higher mass range ($M_{pr} > 300$) this is m/z 170. Overall, a product ion at m/z 91 has the highest probability and is, therefore, the least selective, followed by 105, 121, 109, 100, and 72. From these, only the product ions at m/z 91 (tropylium cation) and 105 (benzoyl cation) were previously indicated as being insufficiently selective in LC-MS/MS by the UK guidelines for urine drugs testing [9].

The relation between the product ion probability distribution and the precursor ion mass was briefly studied. A clear decrease of the probability of the product ion at m/z 91 is observed with increasing precursor ion mass. To less extent, this decrease is observed for m/z 105. One explanation for this observation is that for large precursor ion masses, occasionally a limited product ion mass range is acquired, missing certain low mass product ions. Therefore, the probability of product ions at $m/z < 100$ originating from large precursor ions can be somewhat underestimated. However, this effect is not apparent for the continuous distribution of product ions observed for high mass precursor ions and, thus, it is concluded that this effect is actually relatively small. The second explanation is that in large molecules, more sites are available to dissipate the dissociation energy and, thus, the number of fragment stabilization options increases with the molecular mass. As a result, the probability of producing specific low weight product ions decreases.

Modeling product ion probability for the selected precursor ion mass ranges yielded large differences in models compared with the overall product ion probability model. The differentiation in the selected precursor ion mass ranges was found to be effective in coping with the obvious dependency between product and precursor ion mass probabilities (i.e., singly charged precursor ions can only yield product ions at lower m/z).

Neutral Loss Probability

From database B, the neutral losses were calculated ($M_{pr} - M_{pd}$). Although for neutral losses it is less obvious, a clear

dependency between the neutral loss and the precursor ion mass was observed. Therefore, also for the neutral losses four product ion mass categories were established. As an example, the probability distribution of the neutral losses in the precursor ion category m/z 200–300 is presented in Figure 3b, the data on all categories is presented in Figure S2 (Supplementary Data 3).

From the neutral loss data, for each precursor ion category a continuous model was constructed assuming binomial distributed data using a logit link function to facilitate calculation of $P(M_{nl})$ (see Supplementary Data 4). Again, some neutral losses show an exceptional high or an exceptionally low probability compared to this model (high residual) and for these cases the individual neutral loss probability is listed in Tables S3 and S4 (Supplementary Data 4).

Overall, a neutral loss of 18 Da has the highest probability and is, therefore, the least selective (as was discussed previously [2]), followed by 46, 17, 45, 60, 73, and 59. From experience, it was expected that the loss of water (18 Da), ammonia (17 Da), formic acid (46 Da), and acetic acid (60 Da), and for larger molecules glycoside (162 Da), have a high probability and are, thus, nonselective. However, other observations are not considered general knowledge, like the low selectivity of the neutral loss 59 and 73 Da. The neutral losses of 22–25 Da have an exceptionally low probability ($P_{nl}=0$) compared with the empirical model. This can be explained because neutral losses at these masses are chemically impossible; no molecular structure can be drawn that is in agreement with these neutral loss masses.

In a previous version of database B, the high probability of neutral loss 162, 176, 194, and 308 Da were quite dominant. These neutral losses are all to some extent related to the loss of glycoside, galactoside, and glucuronide moieties. Especially the high probability of 176 Da was caused by the inclusion of the data obtained from Wissenbach et al. [28], who presented the fragmentation of a large number of antidepressant glucuronides. After removal of this data from database B, the probability of 176 Da dropped from 0.11 to 0.03, demonstrat-

ing that originally this single paper largely biased the dataset. This effect was only limited for the other often occurring neutral losses (e.g., for 162 Da the probability dropped by 10 % only). This demonstrates that the construction of a representative product ion spectra database is crucial for correct interpretation of the selectivity.

The relation between the neutral loss probability distribution and the precursor ion mass was studied. The probability of the smaller neutral losses (e.g., 18, 17, and 46 Da) clearly decreases with increasing precursor ion mass. This might be explained by the increasing number of dissociation reactions possible with increasing molecular mass (e.g., multiple water losses. As expected, the probability of multiple dissociation reactions in a high mass molecule increases, together resulting in higher neutral loss values.

Chromatographic Co-Elution Probability

Besides the selectivity provided by the MS detection, the liquid chromatographic system adds additional selectivity. Here, LC is considered as an analytical procedure with measurable selectivity rather than as a clean-up only. To predict the probability of a co-eluting compound, the RT of an infinite number of unknown, possibly interfering compounds should be estimated. Because this is practically impossible, a theoretically predicted RT can be used. A QSRR model can be constructed to predict a compounds' RT based upon its physical and chemical properties [46, 47]. This approach has been applied by Zheng et al. [48], who obtained an excellent relation between the true and the predicted RT ($r^2 > 0.997$) for nucleic acids. However, they divided the compounds in eight groups that were modeled individually, each group containing compounds with similar properties. Baćzek et al. [49] and Ghasemi et al. [50] applied QSRR to predict the RT of compounds on a previously characterized chromatographic system, resulting in accurate RT predictions; $r^2 > 0.89$ and 0.9 , respectively.

In this work, a linear regression model was constructed on the basis of the experimentally determined RT in a RP LC system of 126 compounds and a set of 22 physical and chemical parameters for each of these compounds. From the calculations, 10 parameters proved to result in the best prediction of the RT. The empirically determined model is:

$$\begin{aligned}
 RT_t = & 4.22 + 0.20 * \log D_{pH3} + 0.61 * \log D_{pH5.5} \\
 & - 2.1 * 10^{-7} * B - 9.5 * 10^{-4} * T + 0.19 * P \\
 & - 1.9 * \frac{A_{pol}}{A_{mol}} + 0.12 * Shape - 3.1 * 10^{-3} * E_{tot} \\
 & + 6.8 * 10^{-3} * W
 \end{aligned} \quad (5)$$

in which B is Balaban index, T is molecular topological index, P is partitioning coefficient, A_{pol} is polar surface area,

A_{mol} is Connolly molecular area, E_{tot} is total energy. and W is Wiener index. The relation between the experimental RT and the calculated RT from the model is linear with a coefficient of correlation (r^2) of 0.91, which is comparable to previous QSRR models that include different compound classes [49, 50]. From this model, the prediction error was calculated, which is the combination of the residual standard error of the model and the standard error of the regression line parameters. Because the residual standard deviation of the model is significantly larger than the standard deviation of the regression line parameters, the prediction error is 0.7 min over the whole RT range. Therefore, using this model, the RT of unknown compounds on the RP system tested can be predicted on the basis of its physical and chemical properties with a prediction error of ± 1.4 min ($\alpha = 0.05$).

The probability distribution of the theoretical RT of the 126 compounds used to construct the model plus 325 randomly selected compounds is presented in Figure S3 (Supplementary Data 5). The largest probability of co-elution is observed for extremely hydrophilic compounds that do not show retention on the chromatographic system. Next, the highest probability of co-elution is observed for compounds that elute at RT 8.5 min. A high or low RT (as long as compounds undergo some retention) is less probable and therefore more selective. From this, the fraction of compounds eluting within a certain RT window can be calculated and the probability of co-elution of a target compound with other unknown compounds can be estimated. According to EU regulations, the maximum deviation of the RT of a compound compared with the reference is 5 % in case no internal standard is used [4]. Therefore, for a compound of interest that has an experimental RT of x, the best estimate of the probability of interference due to co-elution of unknown compounds is the probability of compounds with a theoretical RT of $x - 5\%$ and $x + 5\%$. The fraction of compounds eluting from the presented LC system between these theoretical RTs can be determined from database C and Figure S3 (Supplementary Data 5).

A prerequisite for obtaining a good estimate for P(RT) is that the used chromatographic system has been modeled and an RT probability distribution is constructed from a significant number of compounds. If not, a worst case approach can be applied. From our data, we found that the highest probability for co-elution of interfering compounds within the established RT interval is for compounds having an RT of 9.5 min. The corresponding P(RT) is 0.11. To challenge the suggested procedure, a second RP chromatographic system using a different analytical column and acetonitrile as the organic modifier was modeled. For this system, the relation between the experimental RT and the calculated RT from the model is linear with a coefficient of correlation (r^2) of 0.90. Co-elution showed to be most likely at RT=8.0 min and the corresponding P(RT) is 0.16. From this, we assume that for a high resolution RP system using gradient elution, a worst case value of P(RT)=0.2 can be adopted for calculating P(I) in case the used system is not modeled.

Application of the Procedure

As an illustration of the proposed procedure given in Figure 1, the calculation of $P(I)$ for sebuthylazine, 3-([2-nitrophenyl-methylene] amino-2-oxazolidinone (NPAOZ), oxolinic acid, 17 β -trenbolone, and ceftiofur are presented in Table 1. The preselected precursor mass and product ion masses for these compounds are included in Table 1. $P(M_{pr})$ is empirically determined from the probability distribution presented in Figure 2. $P(M_{pd})$ and $P(M_{nl})$ of both mass transitions are empirically determined, respectively, from the probability distributions presented in Figure S1 (Supplementary Data 1) and Table S1 and S2 (Supplementary Data 2), respectively, Figure S2 (Supplementary Data 3) and Table S3 and S4 (Supplementary Data 4), at the corresponding precursor mass category. The theoretical RT is calculated from equation 5 and $P(RT)$ is empirically determined from Figure S3 (Supplementary Data 5). The results are presented in Table 1.

The proposed procedure was applied for 200 compounds that are routinely analyzed using LC-MS/MS at the authors' laboratory, including veterinary drugs, pesticides and natural toxins. The cumulative probability distribution versus $P(I)$ is presented in Figure 4 for using one and two MS/MS transitions. It is observed that the proposed procedure supports the legal requirement of monitoring at least two product ions [4]: $P(I)$ strongly increases when using two products ions, as expected. As an example, if for ceftiofur (Table 1) only one product ion is monitored, $P(I)$ is $1.5 \cdot 10^{-6}$, and when considering two product ions, $P(I)$ is $4.2 \cdot 10^{-8}$, which is an improvement of method selectivity of approximately 300 times.

Setting a Threshold for $P(I)$

When is a method sufficiently selective? To be able to answer this question, a calculated value $P(I)$ should be ranked in terms of method selectivity (e.g., insufficient or

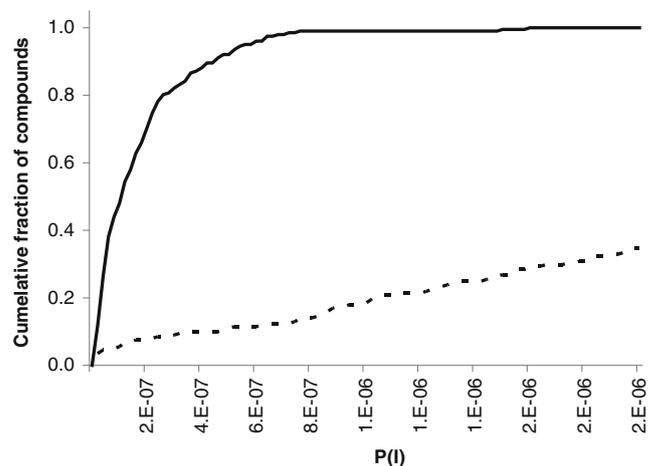


Figure 4. Cumulative distribution function of the fraction of the compounds ($n=200$) versus $P(I)$ in the authors' laboratory using (dashed line) one product ion and (solid line) two product ions

sufficient). This can be done using a threshold value for $P(I)$. However, that would suggest the existence of a sharp boundary between “selective” and “non-selective,” which is not the case. When $P(I)$ is close to any set threshold value, other factors should be considered as well (e.g., the sample preparation procedure and/or results obtained from orthogonal (screening) analysis, which might provide additional information on the compound's identity. In forensic sciences, DNA-matching techniques yield in a probability of 10^{-9} that someone would have the same DNA profile and, thus, the probability of an incorrect identification is at most 1 in a billion [51]. However, no threshold value for DNA-matching probability has been established nor for acceptance of a DNA-match in court: DNA evidence by itself is usually not considered sufficient for conviction [52]. Although it is somewhat arbitrary, we feel that a threshold for $P(I)$ can be useful in order to evaluate method selectivity. The threshold

Table 1. Precursor ion Mass, Product Ion Mass and Retention Time of Sebuthylazine, NPAOZ, Oxolinic Acid, 17 β -trenbolone, and Ceftiofur Including Calculation of $P(I)$

	Sebuthylazine [20]	NPAOZ [21]	Oxolinic acid [36]	17 β -Trenbolone [53]	Ceftiofur [54]
Precursor ion	230	335	262	271	524
$P(M_{pr})$	$7.5 \cdot 10^{-4}$	$4.0 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$	$7.6 \cdot 10^{-4}$
Product ion 1 (m/z)	174	291	244	253	241
$P(M_{pd1})^a$	0.025	0.019	0.010	0.011	0.024 ^a
Neutral loss 1	56	44	18	18	283
$P(nl_1)^a$	0.060 ^{a,b}	0.033 ^{a,b}	0.10 ^{a,b}	0.10 ^{a,b}	0.012
Product ion 2 (m/z)	104	262	216	199	125
$P(M_{pd2})^a$	0.028	0.018	0.013	0.018	0.028 ^a
Neutral loss 2	126	73	46	72	399
$P(M_{nl2})^a$	0.043 ^a	0.042 ^{a,b}	0.062 ^{a,b}	0.035 ^a	0.004
$P(MS)$	$1.9 \cdot 10^{-6}$	$5.4 \cdot 10^{-6}$	$8.7 \cdot 10^{-6}$	$5.6 \cdot 10^{-6}$	$5.1 \cdot 10^{-7}$
Theoretical RT (min)	8.4	0.6	4.8	9.3	10.3
$P(RT)$	0.09	0.06	0.04	0.11	0.08
$P(I)$	$1.7 \cdot 10^{-7}$	$3.0 \cdot 10^{-7}$	$3.7 \cdot 10^{-7}$	$5.9 \cdot 10^{-7}$	$4.2 \cdot 10^{-8}$
1 in...	5,900,000	3,300,000	2,900,000	1,700,000	24,000,000

^aThe maximum value of $P(M_{pd})$ versus $P(M_{nl})$

^bProbability with high deviation from the constructed model; empirical value taken from Supplementary Data 1 (product ions) or 2 (neutral loss)

should be fixed in such a way that it can be concluded that the method is sufficiently selective (fit for purpose) (i.e., (1) the calculated probability of the occurrence of an incorrect identification is low enough to persuade a judge or jury in court, and (2) an expert scientist would feel comfortable in defending the probability of an incorrect identification).

According to the authors, the certainty needed for satisfactory confirmation depends on fitness for purpose [1] and, thus, the threshold value can differ depending on the field of interest (e.g., food safety, sports doping, forensic sciences, and the legal framework involved). To support the establishment of a threshold value for our laboratory, which focusses on residue analysis of registered and banned veterinary drugs, pesticides, and natural toxins in food products, we studied sebuthylazine and NPAOZ, compounds for which false positive results were previously reported [20, 21]. Furthermore, we studied oxolinic acid and 17 β -trenbolone, compounds that show difficulties in confirmatory analysis [36, 53] and of which the selected neutral losses, respectively, water and formic acid for oxolinic acid and water and C₄H₈O for 17 β -trenbolone, are considered to be nonselective. The calculated P(I) for NPAOZ, oxolinic acid, and 17 β -trenbolone are, respectively, 3.0×10^{-7} , 3.7×10^{-7} , and 5.9×10^{-7} (Table 1). Because of the apparent probability of a false identification for these compounds, in our opinion, the proposed procedure should indicate insufficient method selectivity in these cases and thus we prefer $P(I) = 2 \times 10^{-7}$ as a threshold value for evaluation of the method selectivity. We challenged this threshold value on basis of the calculated P(I) of 200 compounds that are routinely analyzed at our laboratory by LC-MS/MS (Figure 4). Of these 200 compounds, 20 are routinely analyzed using neutral losses of ammonia, water, and/or formic acid. Because these neutral losses are generally considered to be non-selective, the probability of a false positive result should be apparent from the proposed procedure. Of these 20 compounds, calculated P(I) ranges from 3×10^{-8} to 1.6×10^{-6} and only 10 % are below 2×10^{-7} and would, therefore, be considered sufficiently selective, which is in agreement with expectations. This supports the choice of $P(I) \leq 2 \times 10^{-7}$ in our laboratory as a suitable criterion for sufficient method selectivity. Using this criterion, the analysis of sebuthylazine ($P(I) = 1.7 \times 10^{-7}$) would be considered just sufficient, even though false positive findings have been reported [20]. We recall that P(I) is an estimation of the true probability of interfering compounds; therefore, we suggest to critically review all compounds having a P(I) around the threshold value.

From Figure 4, it is observed that approximately 70 % of the selected 200 routinely analyzed compounds give $P(I) \leq 2 \times 10^{-7}$ and, thus, show sufficient selectivity. For the other 30 %, additional measures should be taken in case of confirmatory analysis to prevent false positive results to occur. For example, if for oxolinic acid a product ion of m/z 160 is monitored instead of m/z 244, P(I) would be 1.3×10^{-7} , which can be just sufficient, and if three product ions are monitored, P(I) would be 1.3×10^{-8} , which is quite safe. These are clear examples of how the calculation of P(I) can assist in selecting product ions and the number of product ions to be monitored

during method development to assure adequate method selectivity and, thus, a high certainty of the confirmation.

Conclusions

The probability of the co-occurrence of a compound showing the same characteristics in LC-MS/MS can be estimated from empirical models derived from three databases, which include data on the precursor ion mass, the product ion mass, and a theoretical RT. The proposed procedure is extremely useful to select sufficiently selective MRM acquisition parameters during method development. Thus, it serves as an additional tool to the established relative ion abundance criteria [4–10] and together are a very strong combination for confirmation of the identity of a compound because both relative ion abundances as well as the selectivity of the monitored product ions is taken into account. Furthermore, the procedure reported here is a useful tool to validate the selectivity of any LC-MS/MS method.

Acknowledgment

The authors acknowledge funding for this research was by the Dutch Ministry of Economic Affairs, Agriculture, and Innovation. They thank Dr. Hilko van der Voet for the discussions on the statistical approach, Olga Lushchikova for her contribution to the construction of database C, and all colleagues who have contributed MS/MS spectra to database B.

References

- Bethem, R., Boison, J., Gale, J., Heller, D., Lehotay, S., Loo, J., Musser, S., Price, P., Stein, S.: Establishing the fitness for purpose of mass spectrometric methods. *J. Am. Soc. Mass Spectrom.* **14**(5), 528–541 (2003)
- Heller, D.N., Lehotay, S.J., Martos, P.A., Hammack, W., Fernandez-Alba, A.R.: Issues in mass spectrometry between bench chemists and regulatory laboratory managers: Summary of the roundtable on mass spectrometry held at the 123rd AOAC International Annual Meeting. *J. AOAC Int.* **93**(5), 1625–1632 (2010)
- De Zeeuw, R.A.: Substance identification: the weak link in analytical toxicology. *J. Chromatogr. B* **811**, 3–12 (2004)
- Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. *Off. J. Europ. Commun.* **L221**, 8–36 (2002)
- Society of Forensic Toxicology/American Academy of Forensic Sciences. *Forensic Toxicology Laboratory Guidelines* (2002)
- Association of Official Racing Chemists. AOAC Guidelines for Identification by Chromatography and Mass Spectrometry: MS Criteria Working Group - 19 June 2002 Version. *Internal communication to AORC members* (2002)
- Division of Residue Chemistry, Office of Research, Center for Veterinary Medicine, Food Drug Administration. Final Guidance for Industry: *Mass Spectrometry for Confirmation of the Identity of Animal Drug Residues* (2003)
- IOC, International Olympic Committee. Analytical criteria for reporting low concentrations of anabolic steroids. *Internal communication to IOC accredited laboratories*, Lausanne, Switzerland (1998)
- United Kingdom Laboratory Guidelines for Legally Defensible Workplace Drug Testing: Urine Drug Testing. Version 1.0 (2001)
- WADA, The World Anti-Doping Agency. International standard for laboratories, version 3.0 (2003)
- Sphon, J.A.: Use of mass spectrometry for confirmation of animal drug residues. *J. Assoc. Off. Anal. Chem.* **81**(5), 1247–1252 (1978)

12. Lehotay, S.J., Mastovska, K., Amirav, A., Fialkov, A.B., Martos, P.A., de Kok, A., Fernández-Alba, A.R.: Identification and confirmation of chemical residues in food by chromatography-mass spectrometry and other techniques. *TrAC - Trend Anal. Chem.* **27**(11), 1070–1090 (2008)
13. Baldwin, R., Bethem, R.A., Boyd, R.K., Budde, W.L., Cairns, T., Gibbons, R.D., Henion, J.D., Kaiser, M.A., Lewis, D.L., Matusik, J.E., Sphon, J.A., Stephany, R.W., Trubey, R.K.: ASMS fall workshop: Limits to confirmation, quantitation, and detection. *J. Am. Soc. Mass Spectrom.* **1997**(8), 1180–1190 (1996)
14. National Institute of Standards and Technology. NIST '05. 2005. Available at: <http://webbook.nist.gov/chemistry/> Accessed 4 April 2012
15. Council Directive 96/23/EC on measures to monitor certain substances and residues thereof in live animals and animal products and repealing Directives 85/358/EEC and 86/469/EEC and Decisions 89/187/EEC and 91/664/EEC. *Off. J. Europ. Commun.* **L125**, 10–32 (1996)
16. Stolker, A.A.M., Stephany, R.W., van Ginkel, L.A.: Identification of residues by LC-MS. The application of new EU guidelines. *Analisis* **28** (10), 947–951 (2000)
17. Stein, S., Heller, D.: On the risk of false positive identification using multiple ion monitoring in qualitative mass spectrometry: Large-scale intercomparisons with a comprehensive mass spectral library. *J. Am. Soc. Mass Spectrom.* **17**(6), 823–835 (2006)
18. Kaufmann, A.: Validation of multiresidue methods for veterinary drug residues; related problems and possible solutions. *Anal. Chim. Acta* **637**, 144–155 (2009)
19. Kaufmann, A., Butcher, P., Maden, K., Widmer, M., Giles, K., Uria, D.: Are liquid chromatography/electrospray tandem quadrupole fragmentation ratios unequivocal confirmation criteria? *Rapid Commun. Mass Spectrom.* **23**(7), 985–998 (2009)
20. Schürmann, A., Dvorak, V., Crüzer, C., Butcher, P., Kaufmann, A.: False-positive liquid chromatography/tandem mass spectrometric confirmation of sebuthylazine residues using the identification points system according to EU directive 2002/657/EC due to a biogenic insecticide in tarragon. *Rapid Commun. Mass Spectrom.* **23**(8), 1196–1200 (2009)
21. Hurtaud-Pessel, D., Verdon, E., Blot, J., Sanders, P.: Proficiency study for the determination of nitrofuran metabolites in shrimps. *Food Add. Contam.* **23**(6), 569–578 (2006)
22. Emolecules database, 2011. Available at: <http://www.emolecules.com/doc/> Accessed 7 Oct 2011
23. Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Yokota Hirai, M., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, K., Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K., Nishioka, T.: MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **45**, 703–714 (2010)
24. Suthivaiyakit, P., Achatz, S., Lintelmann, J., Aungpradit, T., Chanwirat, R., Chumanee, S., Kettrup, A.: LC-MS/MS method for the confirmatory determination of aromatic amines and its application in textile analysis. *Anal. Bioanal. Chem.* **381**(1), 268–276 (2005)
25. Antignac, J.-P., Cariou, R., Le Bizec, B., Cravedi, J.-P., Andre, F.: Identification of phytoestrogens in bovine milk using liquid chromatography/electrospray tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**(12), 1256–1264 (2003)
26. Fernández-Alba, A.R., García-Reyes, J.F.: Large-scale multi-residue methods for pesticides and their degradation products in food by advanced LC-MS. *TrAC - Trend Anal. Chem.* **27**(11), 973–990 (2008)
27. Bonacci, G., Ascitutto, E., Woodcock, S., Salvatore, S., Freeman, B., Schopfer, F.: Gas-phase fragmentation analysis of nitro-fatty acids. *J. Am. Soc. Mass Spectrom.* **22**(9), 1534–1551 (2011)
28. Wissenbach, D., Meyer, M., Remane, D., Weber, A., Maurer, H.: Development of the first metabolite-based LC-MS(n) urine drug screening procedure-exemplified for antidepressants. *Anal. Bioanal. Chem.* **400**(1), 79–88 (2011)
29. Niessen, W.M.A.: Fragmentation of toxicologically relevant drugs in positive-ion liquid chromatography-tandem mass spectrometry. *Mass Spectrom. Rev.* **30**(4), 626–663 (2011)
30. Díaz, R., Ibáñez, M., Sancho, J.V., Hernández, F.: Building an empirical mass spectra library for screening of organic pollutants by ultra-high-pressure liquid chromatography/hybrid quadrupole time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **25**(2), 355–369 (2011)
31. Dresen, S., Kempf, J., Weinmann, W.: Electrospray-ionization MS/MS library of drugs as database for method development and drug identification. *Forensic Sci. Int.* **161**(2/3), 86–91 (2006)
32. RSC Advancing Chemical Sciences. Chemspider. 2012. Available at: www.chemspider.com. Accessed 19 Jan 2012
33. SciFinder. 2012. Available at: <https://scifinder.cas.org/scifinder/view/scifinder/scifinderExplore.jsf>. Accessed 19 Jan 2012
34. ChemBio 3D Ultra. CambridgeSoft 2012, version 12.0.2.1076
35. GenStat 14th ed. VSN International Ltd.: 2011, version 14.2.0.6297
36. Pozo, Ó.J., Sancho, J.V., Ibáñez, M., Hernández, F., Niessen, W.M.A.: Confirmation of organic micropollutants detected in environmental samples by liquid chromatography tandem mass spectrometry: Achievements and pitfalls. *TrAC - Trend Anal. Chem.* **25**(10), 1030–1042 (2006)
37. McLafferty, F.W., Hertel, R.H., Villwock, R.D.: Probability based matching of mass spectra. *Org. Mass. Spectrom.* **9**, 690–702 (1974)
38. Pesyna, G.M., McLafferty, F.W., Venkataraghavan, R., Dayringer, H.E.: Statistical occurrence of mass and abundance values in mass spectra. *Anal. Chem.* **47**(7), 1161–1164 (1975)
39. Pesyna, G.M., Venkataraghavan, R., Dayringer, H.E., McLafferty, F.W.: Probability based matching system using a large collection of reference mass spectra. *Anal. Chem.* **48**(9), 1362–1368 (1976)
40. Belcher, R.: Specific and selective reactions. *Talanta* **12**(1), 129 (1965)
41. Betteridge, D.: Selectivity index. *Talanta* **12**(1), 129–131 (1965)
42. Berendsen, B.J.A., Stolker, A.A.M., Nielen, M.W.F. Trends in sample preparation procedures for the analysis of veterinary drug residues in food products. *TrAC - Trend Anal. Chem.* (2012) unpublished (submitted).
43. Little, J.L., Williams, A.J., Pshenichnov, A., Tkachenko, V.: Identification of “Known Unknowns” utilizing accurate mass data and ChemSpider. *J. Am. Soc. Mass Spectrom.* **23**, 179–185 (2012)
44. Bolton, E., Wang, Y., Thiessen, P.A., Bryant, S.H. PubChem: Integrated Platform of Small Molecules and Biological Activities. In: R.A. Wheeler and D.C. Spellmeyer (Eds.), American Chemical Society: Washington, DC, 2008.
45. Kind, T., Fiehn, O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **8**, 105 (2007)
46. Kaliszan, R.: Structure and Retention in Chromatography. A Chemometric Approach. Harwood Academic, Amsterdam (1997)
47. Put, R., Vander Heyden, Y.: Review on modeling aspects in reversed-phase liquid chromatographic quantitative structure–retention relationships. *Anal. Chim. Acta* **602**(2), 164–172 (2007)
48. Zheng, J., Polyakova, Y., Row, K.: Prediction of retention of nucleic compounds based on a qspr model. *Chromatographia* **64**(3), 1–9 (2006)
49. Baćzek, T., Kaliszan, R.: Predictive approaches to gradient retention based on analyte structural descriptors from calculation chemistry. *J. Chromatogr. A* **987**(1/2), 29–37 (2003)
50. Ghasemi, J., Saaidpour, S.: QSRR Prediction of the chromatographic retention behavior of painkiller drugs. *J. Chromatogr. Sci.* **47**(2), 156–163 (2009)
51. National Forensic Institute, 2007. The essence of forensic DNA research, interpretation of DNA evidence I, match and calculated frequency. Available at: www.bioplein.nl/attachments/File/Humane_Sporen/EssentiesDNA_Brontekst_6_DNAbewijs_match_en_berekende_frequentie_.pdf. Accessed March 2012
52. Lynch, M.: God's signature: DNA profiling, the new gold standard in forensic science. *Endeavour* **27**(2), 93–97 (2003)
53. Blokland, M.H., Zoontjes, P.W., Sterk, S.S., Stephany, R.W., Zweigenbaum, J., van Ginkel, L.A.: Confirmatory analysis of trenbolone using accurate mass measurement with LC/TOF-MS. *Anal. Chim. Acta* **618** (1), 86–93 (2008)
54. Berendsen, B.J.A., Stolker, L.A.M., Nielen, M.W.F.: Assessment of liquid chromatography-tandem mass spectrometry approaches for the analysis of cefiofur metabolites in poultry muscle. *Food Add. Contam. A* **29**(2), 197–207 (2011)