

Advances in structure elucidation of small molecules using mass spectrometry

Tobias Kind · Oliver Fiehn

Received: 26 February 2010 / Accepted: 3 August 2010 / Published online: 21 August 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The structural elucidation of small molecules using mass spectrometry plays an important role in modern life sciences and bioanalytical approaches. This review covers different soft and hard ionization techniques and figures of merit for modern mass spectrometers, such as mass resolving power, mass accuracy, isotopic abundance accuracy, accurate mass multiple-stage MS(n) capability, as well as hybrid mass spectrometric and orthogonal chromatographic approaches. The latter part discusses mass spectral data handling strategies, which includes background and noise subtraction, adduct formation and detection, charge state determination, accurate mass measurements, elemental composition determinations, and complex data-dependent setups with ion maps and ion trees. The importance of mass spectral library search algorithms for tandem mass spectra and multiple-stage MS(n) mass spectra as well as mass spectral tree libraries that combine multiple-stage mass spectra are outlined. The successive chapter discusses mass spectral fragmentation pathways, biotransformation reactions and drug metabolism studies, the mass spectral simulation and generation of *in silico* mass spectra, expert systems for mass spectral interpretation, and the use of computational chemistry to explain gas-phase phenomena. A single chapter discusses data handling for hyphenated approaches

including mass spectral deconvolution for clean mass spectra, cheminformatics approaches and structure retention relationships, and retention index predictions for gas and liquid chromatography. The last section reviews the current state of electronic data sharing of mass spectra and discusses the importance of software development for the advancement of structure elucidation of small molecules.

Keywords Structure elucidation · Mass spectrometry · Tandem mass spectra · Fragmentation prediction · Mass spectral interpretation · Mass spectral library search · Multistage tandem mass spectrometry

Introduction

Mass spectrometry is a standard technique for the analytical investigation of molecules and complex mixtures. It is important in determining the elemental composition of a molecule and in gaining partial structural insights using mass spectral fragmentations. The final structure confirmation of an unknown organic compound is always performed with a set of independent methods such as one- (1D) and two-dimensional (2D) nuclear magnetic resonance spectroscopy (NMR) or infrared spectroscopy and X-ray crystallography and other spectroscopic methods. The term structure elucidation usually refers to full *de novo* structure identification, and it results in a complete molecular connection table with correct stereochemical assignments. Such an identification process without any assumptions or pre-knowledge is commonly the domain of nuclear magnetic resonance spectroscopy. The term dereplication often refers to the rediscovery of known natural products by means of mass spectral library search or the interpretation of known mass spectral fragmentations.

Electronic supplementary material The online version of this article (doi:10.1007/s12566-010-0015-9) contains supplementary material, which is available to authorized users.

T. Kind (✉) · O. Fiehn
Genome Center–Metabolomics, University of California Davis,
Davis, CA 95616, USA
e-mail: tkind@ucdavis.edu

O. Fiehn
e-mail: ofiehn@ucdavis.edu

Scope of this review

This review investigates theoretical and experimental structure elucidation techniques using mass spectrometry for organic molecules with a molecular mass less than 2,000 Da. The review covers newer techniques within the last 10–15 years; if none were available, then older material was included. Hyphenated separation techniques (gas chromatography coupled to mass spectrometry (GC-MS) and liquid chromatography coupled to mass spectrometry (LC-MS)) are covered due to the close relationship of those techniques with mass spectrometry. Detailed proteomics and peptide sequencing strategies along with the structure elucidation of large biomolecules, such as RNA, DNA, and oligosaccharides/glycans, are outside the scope of this review. The term “small molecules,” used throughout this review, thus refers to all small molecules excluding peptides. Approaches for inorganic mass spectrometry as well as elemental and organometallic analysis are only sparsely covered.

Mass spectral instrumentation and ionization techniques

The history of commercial mass spectrometry instrumentation covers more than 40 years. Brunnee covers the principles of common mass analyzers in a vibrant 1987 review [1]. Gelpi discusses over 130 different mass spectrometers built since 1965 in a series of two reviews [2, 3]. Only one totally new mass spectrometer type, the Orbitrap analyzer [4, 5], has been developed lately. Nevertheless, many new hybrid approaches, among them ion mobility coupled to time-of-flight (TOF) mass spectrometers have been introduced to the market recently. A series of ionization techniques and figures of merit for mass spectrometers will be discussed in the proceeding paragraphs.

Soft and hard ionization techniques

Electron ionization (EI) at 70 eV is historically seen as the oldest ionization technique for small-molecule investigations. Because of the selected constant ionization energy, this technique results in consistent and fragment-rich mass spectra. These mass spectra can be easily used for a mass spectral library search. Electron ionization is commonly used for GC-MS setups. A major disadvantage of mass spectra obtained under EI conditions is the low abundant or missing molecular ion. An abundant molecular ion however is needed for the calculation of elemental compositions. Chemical ionization (CI) is a soft ionization technique mostly used in GC-MS setups to obtain molecular ion information [6, 7]. Supersonic molecular beam interfaces provide the ability to

obtain fragment rich electron ionization spectra together with abundant molecular ions [8].

The introduction of electrospray ionization (ESI) [9, 10] was a major breakthrough for the analysis of intact and large biomolecules. ESI is now the ionization method of choice for LC-MS in many laboratories worldwide [11]. Additionally, nanoelectrospray (nanoESI) [12] and chip-based nanoelectrospray ionization have been advanced during recent years [13–18]. The infusion of nanoliters of solvents using nanoESI allows for sustained analysis over long sample times with a minimum of sample material and increased sensitivity. These long infusions times are needed for structural identifications from data-dependent MSⁿ fragmentations obtained by ion trap mass spectrometers. The use of a new spray nozzle for each injection prevents cross-contaminations (see Fig. 1) especially when multiple compounds are infused from 396 well plates. Recently, multi-nanoelectrospray emitters (nanoESI) have been developed, which may further enhance ion production and increase the dynamic range (see Fig. 2) [19, 20].

Atmospheric pressure chemical ionization (APCI) [21–24], atmospheric pressure photoionization (APPI) [25–28], and matrix-assisted laser desorption/ionization (MALDI) [29–31] are matured soft ionization techniques. Field desorption and field ionization are also soft ionization techniques, and both produce abundant molecular ions with few fragment ions [32–34]. Direct analysis in real time (DART) [35] is an ambient ionization technique [36] and allows for the real time analysis of the sample. The DART source has been widely used in “open access/walk-up” laboratories together with robotic sample handling [37]. Techniques for sampling molecules from surfaces have been extensively reviewed as well [38]. Secondary ion mass spectrometry (SIMS) and MALDI are used for mass spectrometric imaging [39], a new and exciting technology to gain spatial and structural insights from tissues and organs [40–42]. Several new surface-based ionization

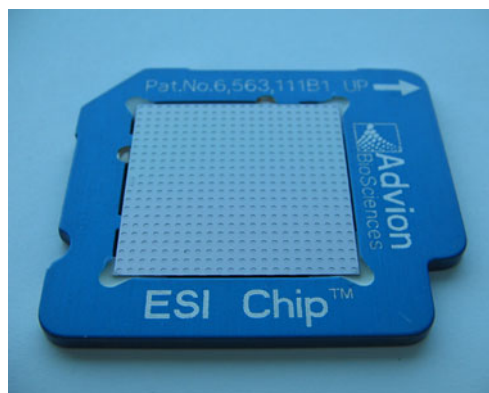


Fig. 1 Chip-based nanoelectrospray allows for sensitive and contamination-free mass spectral infusions (photo by Tobias Kind/FiehnLab)

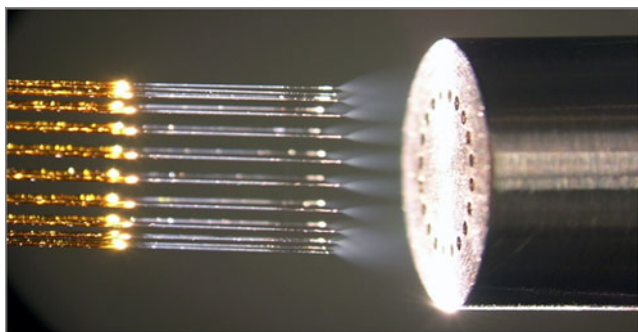


Fig. 2 Nano-electrospray emitter array for enhanced sensitivity of electrospray ionization mass spectrometry (reproduced with permission from Keqi Tang and Richard D. Smith/Pacific Northwest National Laboratory)

techniques including desorption electrospray ionization [43], desorption ionization on silicon, and nanostructure-initiator mass spectrometry [44] have been developed recently.

Multi-mode or simultaneous ion sources [45–48], as well as the pulsed and parallel use of different ionization techniques [49–52], are helpful to shorten analysis time and to obtain structural information from a wide range of different substance classes [53–55] (see Fig. 3). Although simultaneous positive and negative polarity switching is available within many ion source designs, the commercialization of dual- or multi-mode ion sources applying different ionization techniques is a more recent development [56–58].

Figures of merit of mass spectrometers

Mass spectrometers are typically designed for specific analytical aims: ion trap mass spectrometers as versatile instruments, quadrupole mass spectrometers as general work horses, triple quadrupole mass spectrometers as very sensitive instruments for targeted analysis, and Fourier transform instruments for measurements requiring high resolving power and high mass accuracy [59]. In addition

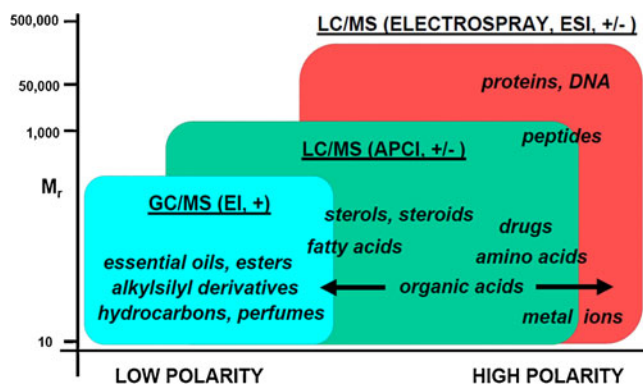


Fig. 3 Coverage of molecule classes with different ionization methods (reproduced with permission from Oxford University Press [55])

to their technical and instrumental design, mass spectrometers can be classified using specific figures of merit [60, 61] (see Table 1). These figures of merit are combinations of hardware, software, and customer experience indicators.

High mass resolving power is needed to resolve overlapping interferences by mass spectrometry only [62–64] (see Fig. 4). Up to one million resolving power can be achieved routinely with current commercially available Fourier transform ion cyclotron resonance (FT-ICR-MS) instruments [65]. A series of “world records” achieved by FT-ICR-MS [66] has been recorded. Hybrid instruments especially allow for the acquisition of high-resolution tandem mass spectra [67, 68] used for natural product structure elucidation. One drawback of FT-ICR-MS and Orbitrap instruments is the higher cycle time to acquire high-resolution broad band mass spectra [69]. At one million resolving power (FT-ICR-MS), a single scan can take up to 2 s or longer. New high-field Orbitrap analyzers can now reach resolving power in excess of 350,000 at m/z 524 (full width at half maximum) [70]. Modern TOF and Q-TOF instrument are routinely capable of higher than 10,000 mass resolving power with the latest instruments reaching up to 40,000 resolving power [71]. When coupled to ultra performance liquid chromatography and comprehensive two-dimensional GC×GC [72–74], the data acquisition rate (scan speed) and duty cycle of the mass selective detector are very important. The chromatographic peak width can be around 2–5 s or lower, and there needs to be enough time to perform additional data-dependent tandem mass spectra (MS/MS) or MS^n scans [75, 76]. Several new TOF and hybrid quadrupole-TOF and iontrap-TOF mass analyzers have been introduced into the market to obtain accurate masses at the MS^n level at a very high data acquisition rate [77, 78]. Additionally, new generation benchtop electrospray ionization time-of-flight analyzers can reach sub-ppm mass accuracy under routine conditions [79]. High mass accuracy together with high isotopic abundance accuracy is generally important to obtain only few molecular formula candidates from an accurate mass measurement [80, 81]. For structure elucidation purposes, the ability to perform multiple-stage MS^n experiments is the most important feature to obtain additional structural information from small molecules [82, 83]. The ability to obtain tandem mass spectra under positive and negative ionization in a single run [84] can speed up the identification process of unknown chemicals [85–87]. Low machine maintenance and high robustness of the instrument operating under different temperatures and humidity ranges in high-throughput manner are additional important aspects. The software as one of the cornerstones for successful compound identification is just as important as the instrument itself. Fast software bug fixes, uncomplicated software updates, easy-to-use graphical user interfaces, and responsive soft-

Table 1 Important figures of merit for modern mass spectrometric systems

Number	Figures of merit	Example ranges (multiple instruments)
1	Mass resolving power	1000–1,000,000 (at m/z 400)
2	Mass accuracy	0.1–1,000 ppm
3	Isotopic abundance accuracy	3–20%
4	Linear dynamic range	100–1,000,000,000
5	Scan speed	1–50,000 u/s
6	Mass range	>100,000 u
7	Abundance sensitivity	100–1,000,000,000
7	Accurate mass MS ⁿ capability	MS/MS or multiple-stage MS ⁿ
8	Ionizer compatibility	Pulsed or continuous
9	Positive/negative polarity switching	Fast switching within run
10	Robustness, maintenance, ease of use	Just chillin in the lab/get the hell out of here
11	Instrument and software cost	50,000–1,000,000 US\$
12	Size/weight/utility	Benchtop or room size
13	Software updates with active support	Customer involved or customer ignored
14	Open-data exchange formats supported	netCDF, mzXML, ASCII

ware support are sometimes more important than certain instrument parameters. Documented software interfaces that allow programmers to access certain software functions and the support of open mass spectral exchange formats (netCDF, mzXML, and mzData) are equally important and discussed later in the article.

Tandem mass spectrometers and modes of operation

Ion trapping instruments such as quadrupole ion traps and FT-ICR mass spectrometers can be used to create tandem mass spectra, and multiple-stage MSⁿ experiments can be

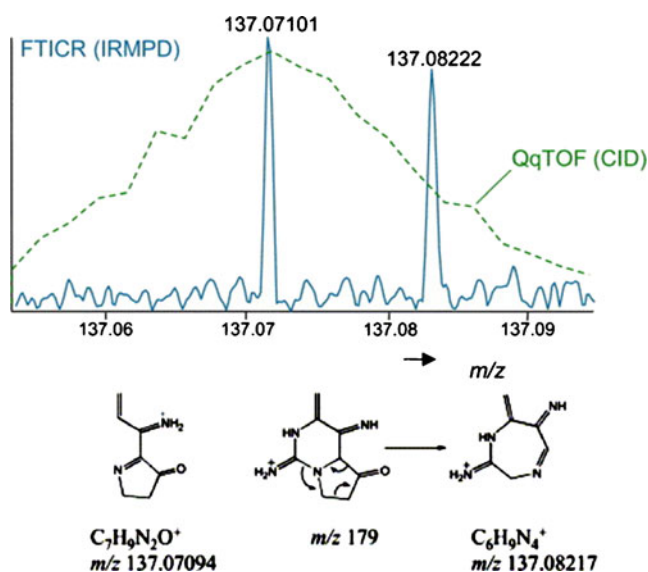


Fig. 4 The importance of mass resolving power showing a high-resolution FT-ICR-MS spectrum with lower resolution Q-TOF mass spectrum. Only the high-resolution instrument can resolve peaks with 0.0112 Da difference (reproduced with the permission from Ref. [63])

performed without instrument modification or couplings of different mass analyzers [88]. Other hybrid instrument types are discussed in Ref. [2, 3]. Orthogonal or hybrid mass spectrometers are favorable for structural elucidation because they either increase the total peak resolution or they introduce another separation dimension that can be used either to trigger or acquire additional mass spectrometric information [89, 90]. The different modes of operation, which include precursor ion scans, product ion scans, neutral loss scans and selected reaction monitoring, are discussed in De Hoffmann [91]. The MS/MS and MSⁿ scans are usually triggered via data-dependent setups. Multiple precursor ions can be manually selected or the software can acquire tandem mass spectra when a certain peak abundance or signal/noise ratio is exceeded. For example, electrospray ionization with ion mobility mass spectrometry coupled to time-of-flight mass spectrometry (ESI-IMMS-TOF-MS) was used for metabolic profiling of *Escherichia coli* metabolites [92], phospholipid [93], and drug analysis [94].

Ion activation modes

Collision-induced dissociation (CID), or collisionally activated dissociation, is the most common technique to obtain tandem mass spectra. Precursor ion stability and internal energy under CID have been previously discussed [95]. A series of new fragmentation modes are aimed at improved protein and peptide identification rates by creating more specific fragmentations. These modes include electron capture dissociation (ECD) [96–98], electron transfer dissociation [99–101], and infrared multiphoton dissociation [102]. They are not fully exploited yet for small-molecule applications outside proteomics.

Two-dimensional, three-dimensional, hybrid, and orthogonal chromatographic approaches

Multiple dimension setups are possible on the chromatographic and mass spectrometric sides. On the chromatography side, the usual aim is directed at increasing the peak resolution, which therefore provides a better separation of overlapping compound peaks. The peak capacity can be increased by using different selective chromatographic phases in a two-dimensional or multi-column setup. These approaches are known for liquid chromatography and prominently used for protein identification by coupling an ion exchange column together with a reversed phase column, which coined the term multidimensional protein identification technology [103]. The difference between simple two-dimensional connections such as GC-GC compared with truly orthogonal approaches such as comprehensive two-dimensional GC (GC×GC) [104] lies in the fact that a modulator is used to accumulate parts of the sample from the first column and pulse the sample to the second shorter column with a different polarity of the stationary phase [105]. The detector must be a fast scanning detector with a high acquisition rate and an example of this is a time-of-flight mass analyzer. Sampling rates are usually between 100 and 200 spectra per second for GC×GC-TOF-MS [106] instruments. The resulting mass spectra have a very high signal to noise ratio and therefore represent cleaner mass spectra and give better mass spectral library search scores [107]. Miniaturization and the introduction of chip-based liquid chromatography [108] play a major role in high-throughput methods.

Mass spectral data handling

The following section discusses basic steps that have to be performed to obtain clean and background free mass spectra. Charge state deconvolution, accurate mass measurements, and software algorithms for elemental composition calculations are reviewed. Certain hardware specific setups are discussed when required.

Background and noise subtraction

Automatic background and noise subtraction are standard techniques to obtain clean and interference free mass spectra. The Biller–Biemann algorithm [109] or similar algorithms by Dromey et al. [110] have been in use for more than 30 years. It is generally advisable to perform blank or solvent runs to obtain possible noise or contamination data. These infusion mass spectra or complete LC-MS and GC-MS runs must be subtracted from the real

sample data [111–113]. Most modern mass spectrometry software tools have inbuilt algorithms to perform these tasks. Many of the mentioned algorithms have been developed for EI (70 eV) mass spectra. Several approaches have been introduced with the CODA algorithm of Windig et al. [114] for ESI and LC-MS data, and similar methods have been applied in drug discovery studies [115–117]. A more detailed discussion about automated approaches is covered in the mass spectral deconvolution and biotransformation sections.

Adduct formation and detection

Ionization techniques such as CI, MALDI, ESI, or APCI show not only single adduct ions but also sets of multiple adducts [118, 119]. The process of adduct formation can be studied using heuristic and computational methods [120, 121]. Solvent and buffer constitution, pK_a , pH, substance proton donor and acceptor properties, and gas-phase acidities influence the formation of adducts [122, 123]. Different adducts also can result in different fragmentation pathways [124]. The correct adduct ion must be detected in order to obtain the accurate mass of the neutral molecule. One possible solution is to increase the concentration of specific ions in the liquid phase [125] to obtain preferably those adducts. When analyzing lipids, lithium is used as modifier [126] to obtain characteristic $[M+Li]^+$ ions. An extended list of common electrospray adducts, including $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$ and $[M-H]^-$, has been prepared [127]. In case of MALDI, metal cation adducts $[M+Na]^+$ and $[M+K]^+$ are often observed [29, 128]. Software tools such as CAMERA [129] and IntelliXtract [130], and tools for infusion spectra [131] can help detect adduct ions in mass spectra automatically. Currently, no software exists that can predict adduct probabilities based on a given compound structure for a specified ionization mode (CI, ESI, APCI, and APPI).

Charge state deconvolution

Charge state determinations play an important role in proteomics [132, 133] but are less frequently required in small-molecule investigations [132]. Many small organic molecules are usually singly charged. Certain molecule classes, such as cardiolipins, may occur as singly and doubly charged ions. The occurrence of multiply or doubly charged ions can be influenced by buffer concentration, analytes concentration, amount of organic modifier, or flow rate [134, 135]. Open-source software tools, such as Decon2LS [136], exist (see Fig. 5), which can automatically determine charge states. Most vendor mass spectrometry software has charge state determinations included.

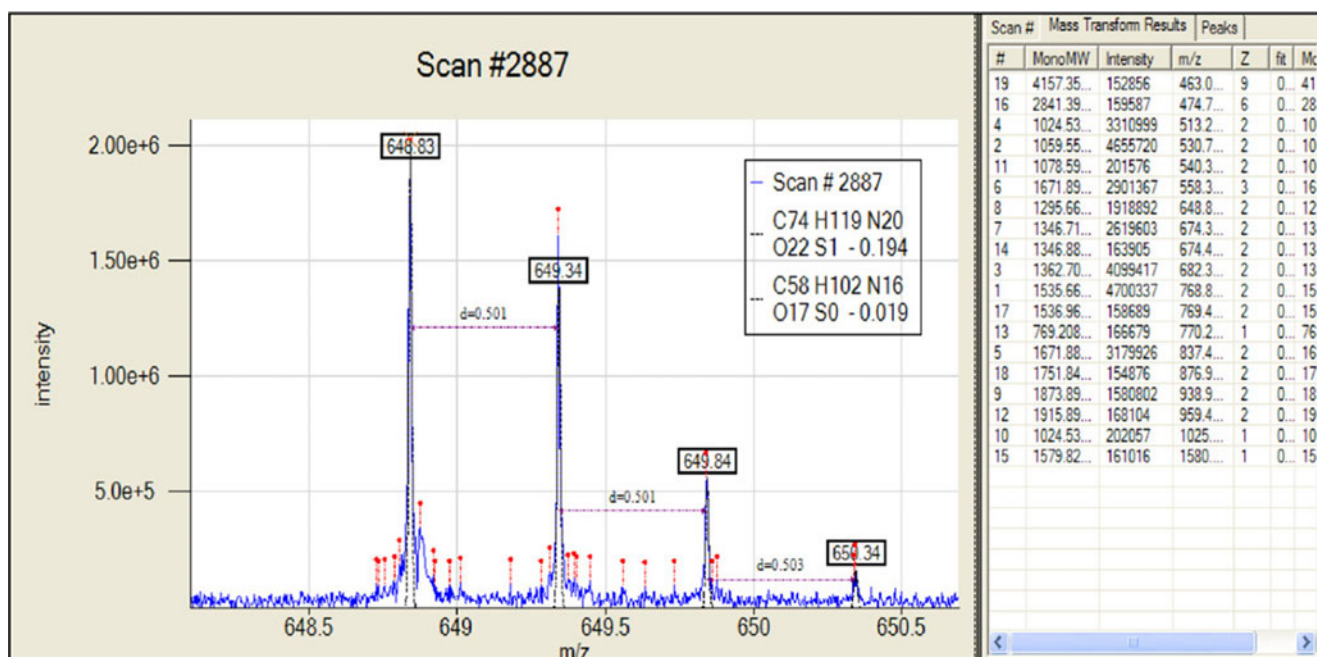


Fig. 5 Charge state deconvolution with the freely available software Decon2LS (reproduced with permission from Ref. [136])

Accurate mass measurements

Accurate masses and isotope abundances are reported in an IUPAC report [137]. The statistical evaluation of measured mass accuracies should include the proper terminology and basic statistic tests [138]. An intercomparison study from 45 laboratories [139] showed that FT-MS and magnetic sector field instruments in peak matching mode routinely achieved less than 1 ppm mass accuracy. Quadrupole-TOF, TOF, and magnetic sector field instruments in magnet scan mode achieved between 5 and 10 ppm. Newer publications reported that time-of-flight instruments can reach around 1 ppm [140] or even sub-ppm [79] mass accuracies. Orbitrap technology in hybrid mode can routinely reach sub-ppm mass accuracy and in non-hybrid mode less than 2 ppm [141]. The importance of the inclusion/exclusion of the electron mass during accurate mass measurements was discussed in Ref. [142], and the impact on LC/TOF-MS mass accuracy was further outlined in Ref. [143]. A mass error of up to 3 ppm was reported if the electron mass is not included in calculations. The mass error introduced by this calculational error can be as high as 5 ppm at 100 m/z (see Fig. 6). The red line marks 300 ppb, which can be obtained from broadband FT-ICR-MS experiments. The current accurate electron mass is reported as: $m(e^-) = 0.00054857990924$ u [144]. A recent approach used the ubiquitous presence of background ions to correct for small mass errors, and this was also used for accurate peak alignment and internal mass calibration [145]. The reported mass error on a LTQ-Orbitrap dropped from ± 1 –2 ppm to

an absolute median error of 0.21 ppm. Another research article discussed a computational method to adjust for mass errors outside the lock mass range and intensity and reported error improvements from 20 down to 1 ppm [146]. The process of selected ion monitoring (SIM) stitching was investigated [147, 148]. The authors concluded that an average mass error of 0.18 ppm could be obtained routinely on a high-resolution FT-ICR mass spectrometer. If instruments are uncalibrated or out of tune, then an automated post-calibration routine [149] can be used to remove systematic precursor mass errors. The authors' reason that in case of sample overload, the automatic gain control system (AGC) is not able to control the optimal number of ions to inject into the Orbitrap cell, which finally results in space charge effects causing noticeable systematic mass errors.

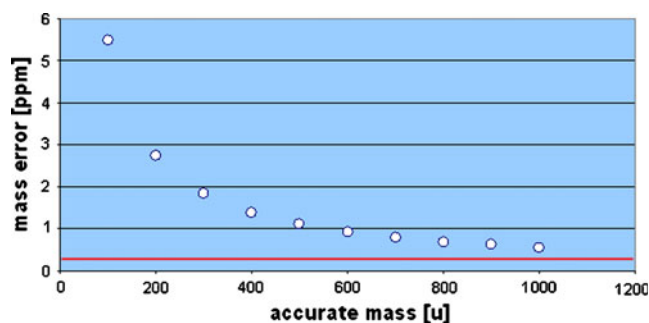


Fig. 6 A mass error of up to 5 ppm is the penalty if the electron mass is not accurately included in accurate mass calculations. The lower red line marks 0.3 ppm mass accuracy, which can be reached by FT-ICR-MS

Higher mass accuracy on unit mass resolution instruments can be obtained using post-processing peak shaping algorithms as implemented in the MassWorks software (Cerno Biosciences) [150, 151]. These algorithms use an internal calibrant that is later used for post-calibration of mass accuracy errors. Unit resolution mass spectrometers (inaccurate mass spectrometers) can be converted into accurate mass spectrometers as long as mass spectral data are obtained in profile mode, which is required to perform the spectral peak shape correction. If data are obtained in centroid mode or stick mode, then no such post-correction can be performed. A correction for spectral accuracy can also be performed with high-resolution data [152]. Artificial neural network calibration [153] in conjunction with AGC and better peak centroiding can improve the mass accuracy on FT-MS instruments to reach 100 ppb for certain experiments [154].

Several unit mass resolution instruments, including ion traps and triple quadrupole instruments [155], allow a hardware-based high-resolution or an ultra-zoom scan [156]. This zoom scan can be used for accurate mass measurements or better charge state assignments. The resolving power usually can be increased by one order of magnitude, or from 1,000 resolving power to 10,000 resolving power. However the m/z scan range is usually very limited, and the duty cycle is high for enhanced resolution scans.

Isotope abundance measurements and isotopic pattern calculations

The isotopic abundances of common monoisotopic (F, Na, P, and I) or polyisotopic (H, C, N, O, S, Cl, and Br) elements are listed [137]. Isotopic abundances are measured and have been utilized in mass spectrometric measurements since the beginning of mass spectrometry [157]. The most sensitive and accurate method for isotopic abundance measurements is accelerator mass spectrometry [158], and this method is used for age determination, forensics, and food monitoring [159]. Its precision is around 0.05% for the measurement of the $^{13}\text{C}/^{12}\text{C}$ ratio [160] requiring total combustion of the sample. The availability of commodity mass spectrometers delivering isotopic abundance errors less than $\pm 5\%$ was utilized for LC-MS-based screening approaches [161–164] and environmental screening applications [165–167].

To filter or match elemental compositions according to their experimental isotopic abundances, the high- or low-resolution isotopic envelopes of molecular formulas must be calculated. Several algorithms have been proposed to calculate the isotopic fine structures and allow the modeling of Gaussian peak shapes according to the selected resolving power of the instrument. Several of the algorithms implement either polynomial-based methods or Fourier

transform-based methods (IsoDalton, MWTWIN, Mercury, IsotopeCalculator, IsoPro, emass/qmass, libmercury++, ISOMABS, and Decon2Ls) [168–171]. Isotopic abundances from tandem mass spectra and multiple-stage MS^n can yield additional information that can help during the structure elucidation process [172–174].

Elemental composition determination

The determination of the molecular formula or elemental composition requires a clean mass spectrum with no interfering noise or coeluting compounds. A process for elemental composition determination from electrospray data was described in Ref. [175]. The algorithm includes a decision making step for proton and alkali metal adducts, automated determination of charge states and overlapping peaks, and an isotopic pattern matching. It was validated with 220 pharmaceutical compounds and yielded a success rate of 90%. Isotope-enriched metabolites can be investigated using a method that includes spectral correlation methods along with mass accuracy and isotope ratio filters [176]. Another software discusses the use of isotopic abundance ratios to confirm or reject NIST mass spectral library search results [177]. A series of papers discusses the process of isotopic pattern matching for elemental formula determination in environmental chemistry [165–167], metabolic profiling experiments [178, 179], and geochemistry [180, 181]. The freely available software SIRIUS (Sum formula Identification by Ranking Isotope patterns Using mass Spectrometry) [182] has a user-friendly graphical interface and can be used on LINUX, MAC, and Windows platforms. The newer implementation “SIRIUS Starburst” also includes features such as peak intensity, number of hetero atoms in the molecular formula, neutral losses, and tandem mass spectral information [183].

The Seven Golden Rules [81] are a set of heuristic rules for elemental composition calculations, including the Senior and Lewis rules, element ratio rules, and an isotopic abundance matching filter. The rules were developed with a set of 68,237 existing elemental compositions and validated with 6,000 molecular formulae by means of an internal database of 432,968 existing elemental compositions. The freely available software was used to calculate the molecular formula space (elements CHNSOP; $<2,000$ u) covering more than two billion elemental compositions, and it was deduced that only 623 million elemental compositions are highly probable (see Fig. 7).

The influence of spectral accuracy of molecular ions on elemental composition calculations was investigated on a high-resolution mass spectrometer [184]. The automated correction of isotope pattern abundance errors using peak shaping and correction algorithms resulted in better identification rates of the molecular formulas. An algorithm

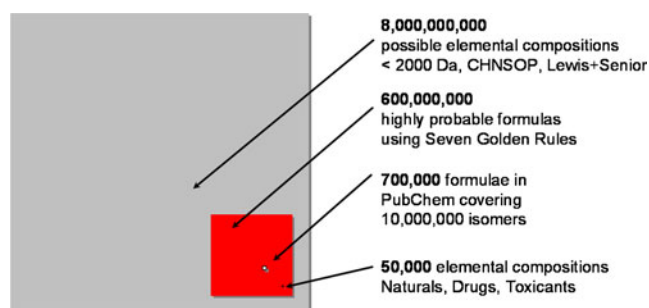


Fig. 7 The molecular formula space below 2,000 Da (elements CHNSOP) covers more than eight billion elemental compositions and can be reduced to 600 million highly probable molecular formulae using the Seven Golden Rules [81]

for isotopic pattern calculation that includes stable isotope markers (^{13}C and ^{15}N labeled) was developed [185]. Recently, an approach was developed that uses elemental formula calculations with database lookup and a subsequent in silico generation of CID mass spectra from the obtained isomer structures [186]. The obtained in silico tandem mass spectra (calculated by MassFrontier) were then compared with experimental CID spectra. This approach combined with additional filter constraints and possible MS^n fragmentation information can be used for compound annotations (compound dereplication), provided that the structure is known in compound databases. Other prerequisites such as proper validation of the in silico prediction algorithms and use of larger datasets will be discussed in a later chapter.

Algorithms for formula calculation from high-resolution MS/MS data

If the mass spectrometer is capable of obtaining accurate mass multistage product ions (MS^n), then this information should be utilized during the elemental composition determination. The possible elemental formulae for single peaks should be shown, and the algorithm should analyze if the elemental composition of the product ion could be combined to generate feasible elemental compositions of the complete molecule. Bruker (Billerica, MA, USA) developed the SmartFormula three-dimensional (3D) algorithm [187] that includes this information by using a recursive algorithm to exclude unfeasible molecular formulae from lower mass fragments (see Fig. 8). Tandem mass spectra obtained under EI can be used together with isotope abundance analysis to obtain correct elemental compositions [174]. Polynomial expansion algorithms to calculate the isotope patterns for precursor ion, neutral loss, and MS^n product ion tandem mass spectra have been discussed in Ramaley and Herrera, and Rockwood et al. [173, 188].

Another approach used accurate masses from MS/MS product ions during the investigation of fragmentation processes of some natural products [189, 190]. Sirius

Starburst [183] is a freely available software that combines MS/MS fragment and element ratio information with elemental composition determinations. A useful hardware-based approach [191], the acquisition of exact masses at high and low ionization energy MS^E , can lead to more accurate elemental formula determinations.

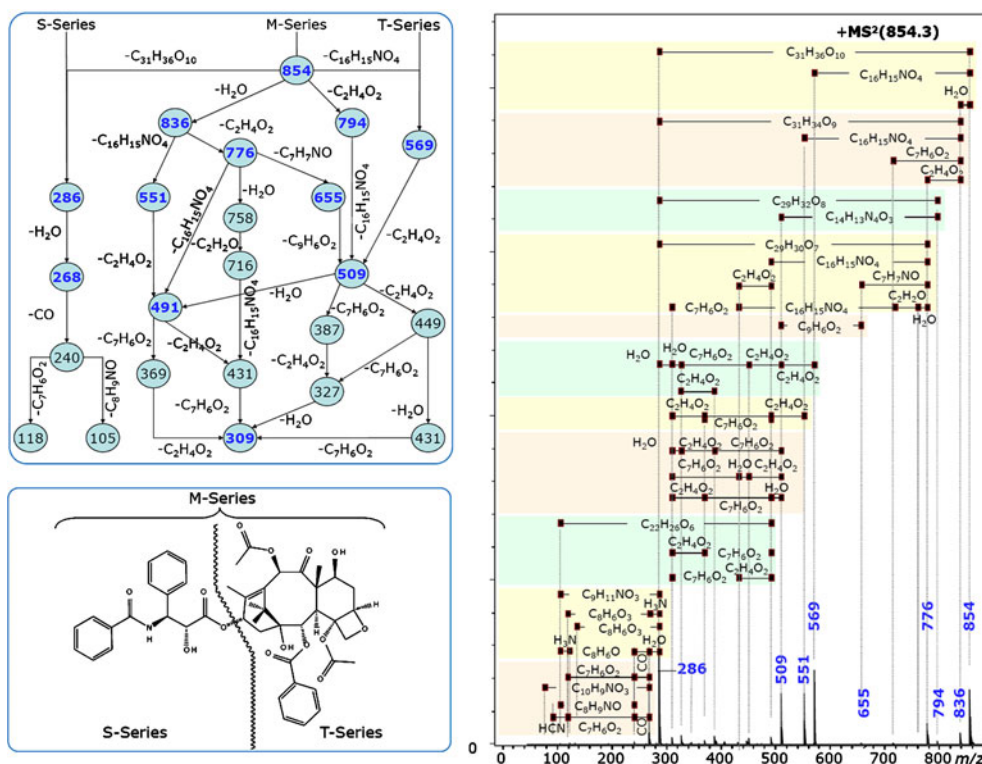
Complex data-dependent setups including ion maps and ion trees

Data-dependent acquisition methods are used in most of today's tandem mass spectrometers [87, 192–197]. The mass spectrometry software triggers MS/MS or MS^n product ion scans based on specific events. The trigger can be set on specific events such as the highest abundant peaks, manually selected masses, specific neutral losses, or specific isotopic pattern [197].

Specific data-dependent setups such as total molecule ion maps (see Fig. 9) are very powerful features for molecule fragmentation studies. The process to create ion maps has been known since more than 20 years [198]; however, it has not reached its full potential, mainly due to missing data handling options. Ion maps contain product ion mass spectra over the mass range of all precursor ions from 20 Da increasingly up to the molecular mass of the compound [199]. These ion maps can be obtained by a longer direct infusion process with autosamplers or better by nanoESI using Nanomate (Advion Inc.) robotic injections to allow long-enough scan times. The method should not be confused with spatial ion maps obtained from secondary ion mass spectrometry TOF-SIMS [200] or mass spectrometric imaging or ion maps that refer to retention time– m/z visualizations (LC-MS ion maps) [201]. The total ion map is a function of precursor m/z value versus product ion m/z value and intensity, and it can be represented in two- or three-dimensional space. The applications range from the investigation of single molecules to obtain deeper structural insights [202] to the investigation of complex petroleum mixtures [203] and natural compounds.

An even more powerful method to investigate mass spectral fragmentations and fragmentation pathways of molecules are ion tree experiments [204–208] (see Fig. 10). A data-dependent ion tree contains multiple MS^2 to MS^n product ion spectra from a single molecule and represents the ultimate mass spectral fingerprint of a molecule. The methodology has been available for many years, and in principle, any mass analyzer capable of MS^n fragmentation can make use of it. The technology is very attractive because it can be performed with inexpensive ion trap systems (tandem-in-time) using direct infusion experiments. Different ionization voltages and adduct-dependent fragmentations, as well as the use of high-resolution measurements and accurate mass MS^n spectra from hybrid

Fig. 8 Fragmentation pathway of paclitaxel and sum formulae for fragments from MS/MS and MS³ experiments calculated with the SmartFormula3D algorithm (reproduced with permission from Ilmari Krebs, Bruker Daltonik GmbH, Bremen [187])



instruments, can reveal additional fragmentation pathways. However, these complex multidimensional setups were rarely used in the past due to data handling and software issues. Application examples include fragment studies of polyphenols [196], lipids [209–213], glycans [214], and carbohydrates [215].

Mass spectral library search

Mass spectral library search is the first step in any mass spectral interpretation and therefore will be discussed in deeper detail. Mass spectral search can be performed with unit mass and high-resolution mass spectra of all stages

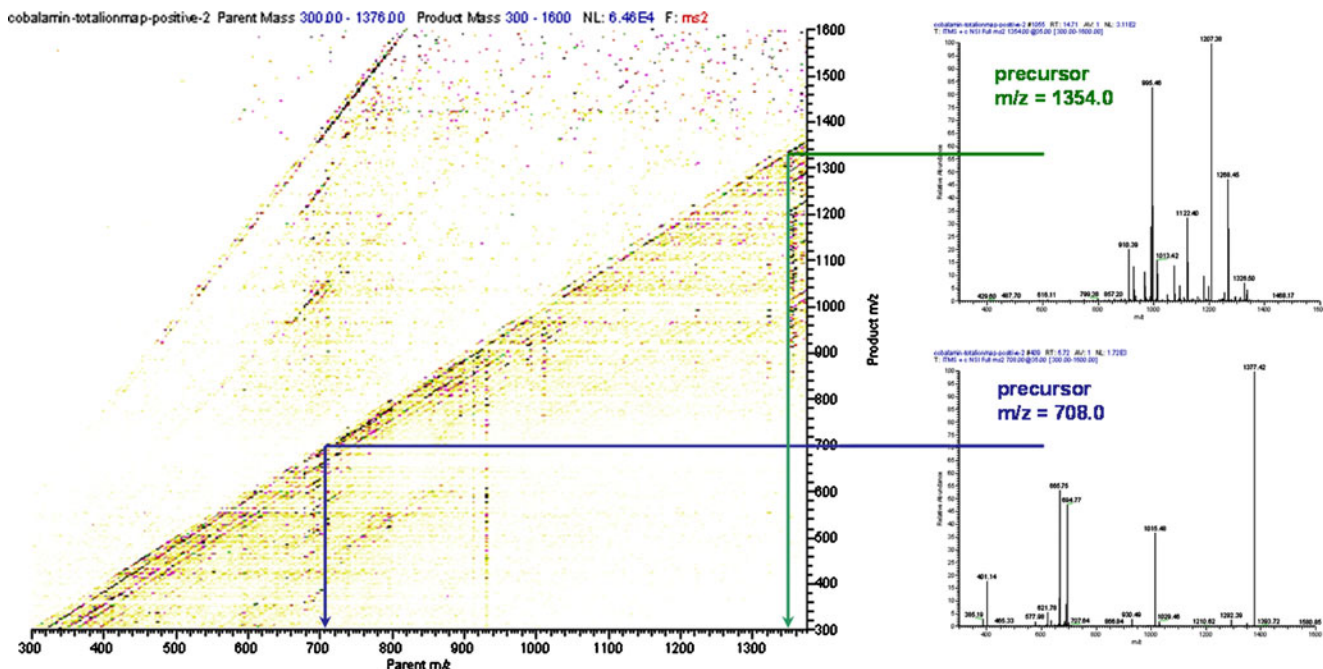


Fig. 9 A total ion map of tandem mass spectra from cobalamin (vitamin B₁₂) created by a linear ion trap mass spectrometer and visualized by the Thermo Xcalibur software. For all precursor ions in the mass range between *m/z* 300 and 1,376, one MS/MS spectrum was acquired

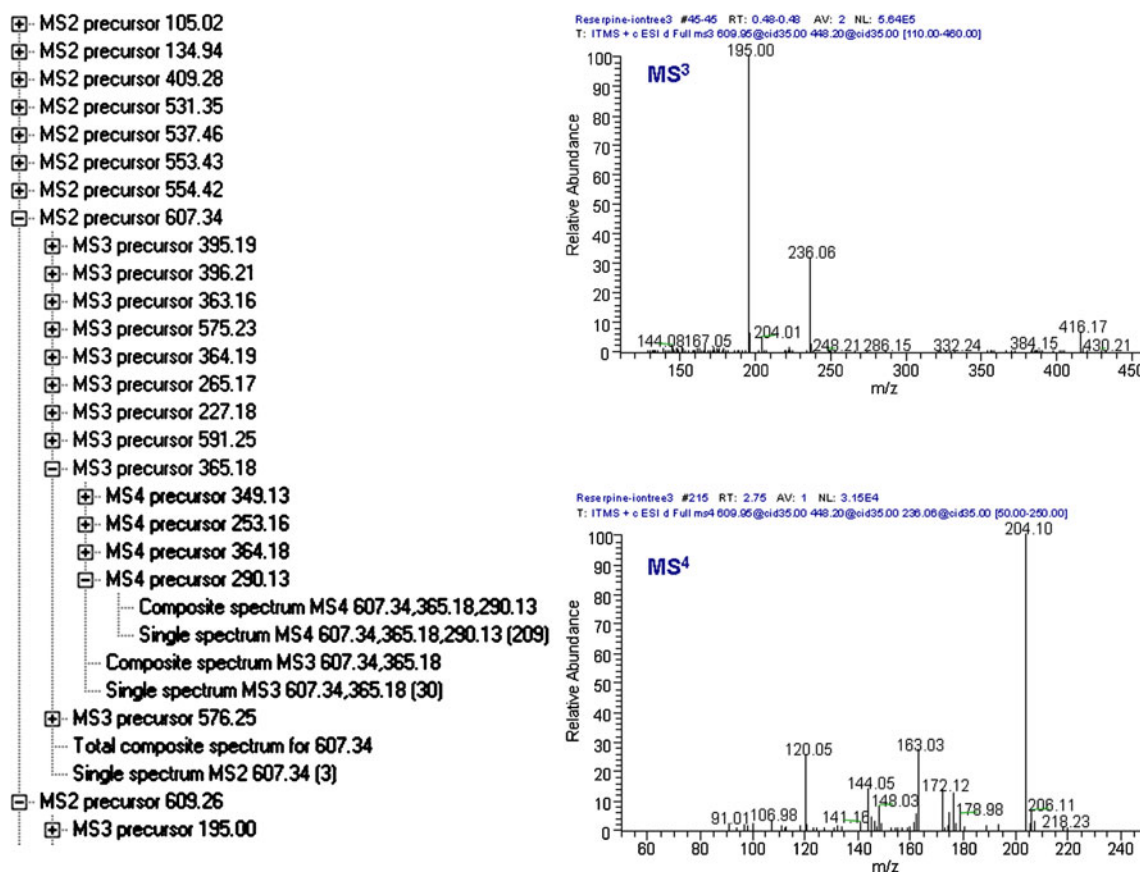


Fig. 10 An automatic data-dependent ion tree experiment with multiple-stage MS^n spectra of selected precursor ions of reserpine acquired on a linear ion trap mass spectrometer. The information rich ion tree represents the ultimate mass spectral fingerprint of a molecule

(MS to MS^n). The aim of a library search is either to obtain a correct structure hit of compounds already in the library or to obtain partial structural insights from compounds that nearly match. For that purpose, an experimental mass spectrum is searched against a large collection of already recorded mass spectra that are stored in a database. A general review of mass spectral libraries [55] and mass spectral search algorithms [216, 217] has been written.

MS and MS/MS and MS^n libraries and search algorithms

Search algorithms for electron ionization spectra were developed first [218], and these include the INCOS algorithm, probability-based matching (PBM) [216], and dot-product algorithm [217]. The size of publicly and commercially available MS/MS libraries is small compared with electron ionization libraries (Wiley and NIST) that cover several hundred thousand electron ionization mass spectra. Currently, the NIST08 MS/MS collection is a large commercially available database with 14,802 MS/MS spectra from 5,308 precursor ions. There are a variety of commercial libraries that have been generated for certain instrument types and settings. The publicly available

Massbank [219, 220] and ReSpect database (RIKEN) [221–223] are databases currently covering 24,772 mass spectra and tandem mass spectra from 13,200 compounds. An electrospray tandem mass spectrometry library (ESI-MS/MS) for forensic applications covered 5,600 spectra of 1,253 compounds acquired at different ionization voltages using a hybrid tandem mass spectrometer coupled to a linear ion trap [224]. Smaller but specialized libraries are in use for toxicological screening and drug analysis [225, 226]. An in-house library of MS/MS spectra from 1,200 natural products with the majority of entries having $[M+H]^+$ adducts and 95% of those compounds being able to ionize in positive mode was investigated in Ref. [227]. Tandem mass spectra are not as reproducible as electron ionization spectra when obtained from different instruments. However, the creation of reproducible and transferable MS/MS spectral libraries for use on multiple instrument types [228] is possible [229, 230]. A fragmentation energy index was proposed for LC-MS [231] to normalize collision energies and create reproducible spectra comparable to 70-eV electron ionization spectra. Another study compared tandem mass spectra obtained from quadrupole–quadrupole–time of flight, quadrupole–quadrupole–linear ion trap, quadrupole–quadrupole–quadrupole,

and linear ion trap–Fourier transform ion cyclotron resonance mass spectrometer and came to the conclusion that platform independent MS/MS spectra can be obtained with multiple fragmentation voltage settings [232–234].

Search algorithms for MS/MS spectra of small molecules can use similar approaches as used for EI mass spectra [55, 235]. Peptide mass spectra usually show specific fragmentations, and a series of specialized search algorithms were developed for these purposes [236, 237]. MS/MS spectra can be searched according to spectral similarity [238], probability match (PBM) [216, 239], or dot-product algorithm search [217]. If the MS/MS spectra were obtained in data-dependent mode and precursor mass information is available, this precursor mass can be used as a powerful first filter for all subsequent MS/MS matches [240, 241]. The precursor m/z search window can be selected according to the experimentally mass accuracy of the instrument. Well-calibrated unit mass resolution instruments can reach a mass accuracy of ± 0.5 Da (or better with post-calibration methods). In this case, a precursor search window of ± 0.5 Da can be set for MS/MS search. The subsequent MS/MS match uses a product ion window search tolerance that is slightly higher due to possible hydrogen shifts. Well-established dot product, PBM, and reverse search algorithms are used to match the filtered MS/MS spectra. The accuracy, recall, precision, true, and false discovery rate of the selected algorithm and all other statistical parameters are best obtained from test sets with known spectra and decoy mass spectral datasets as seen from the proteomics community [242–245]. The freely available NIST Mass Spectral Search Program contains efficient algorithms to search accurate mass tandem mass spectra, including m/z precursor and product ion filtering. Moreover, NIST MS Search can handle and search molecular structures together with their associated mass spectra, which is an obligatory prerequisite for any advanced library search program.

Mass spectral trees combine multiple-stage mass spectra

Ion traps and hybrid mass spectrometers can be used to create multiple-stage mass spectra (MS^n) by consecutively fragmenting precursor and all product ions. Usually, the abundance of the obtained product ions decreases, which sets a practical limit at MS^6 to MS^{10} . Furthermore, there must be enough time for trapping, or a direct infusion experiment has to be performed to generate enough ions. The feasibility of using MS^n data for the investigation of drugs [246], monosaccharides [247], oligosaccharides [248–250], and other molecules has been shown. The use of multistage mass spectral libraries together with precursor ion fingerprinting for structure elucidation purposes has been investigated in Sheldon et al. [205]. The authors show that similar building blocks will have similar product ion

mass spectra, and therefore, the utilization of MS^n spectra of all stages can aid in structure elucidation of the core molecule structures. For example, if a set of molecules would have different substitutions or side chains, then an accurate mass precursor search could not identify these molecules. If the side chain is cleaved off or lost in a dissociation step, then the remaining core molecules would generate similar product ion spectra and therefore could be matched among this set of similar compounds. The representation of a spectral tree (see Fig. 11) of compound mass spectra and their associated structures was obtained from MassFrontier (HighChem Ltd).

Mass spectral interpretation

Many of the developments in mass spectral interpretation are deeply rooted in the era of electron ionization mass spectrometry from the 1970s and 1980s. Hence, mass spectral fragmentation interpretation rules are best developed for EI mass spectrometry. The red book entitled “Interpretation of mass spectra” written by Turecek and McLafferty [251], the book entitled “Introduction to Mass Spectrometry” by Watson and Sparkman [252], and “Understanding mass spectra: a basic approach” by Smith [253] are standard sources for mass spectrometrists investigating electron ionization spectra. These books contain very detailed explanations of reactions and fragmentation pathways, including rearrangement reactions, homolytic or heterolytic bond cleavages, hydrogen rearrangements, electron shifts, resonance reactions, and aromatic stabilizations. Any *de novo* interpretation without any pre-knowledge is still challenging, if not totally impossible, due to the high molecular diversity and many similar compound structures.

The *even electron rule* states that usually neutral molecule fragmentations are observed from molecular ions, but radical loss can also occur in case of aromatic and nitroaromatic compounds [254, 255]. Under positive electrospray (ESI), most fragment ions were reported even electron, whereas the formation of odd electron under EI was significantly higher [256]. The *Stevenson rule* states that ions with low ionization energy are more stable and will gain high peak abundance in the mass spectrum. The *nitrogen rule* should in principle only be used for unit resolution mass spectra because high-resolution and high-accuracy mass spectrometry can always calculate the correct number of nitrogen atoms. The *Rings Plus Double Bonds Equivalent* (RDBE) should not be used with elements that allow multiple valence counts (such as phosphorus and sulfur) [257] as otherwise only possible RDBE ranges can be obtained instead of unique solutions. Mass spectral visualization techniques such as *van Krevelen* or *Kendrick plots*, and spectral mappings using dimension

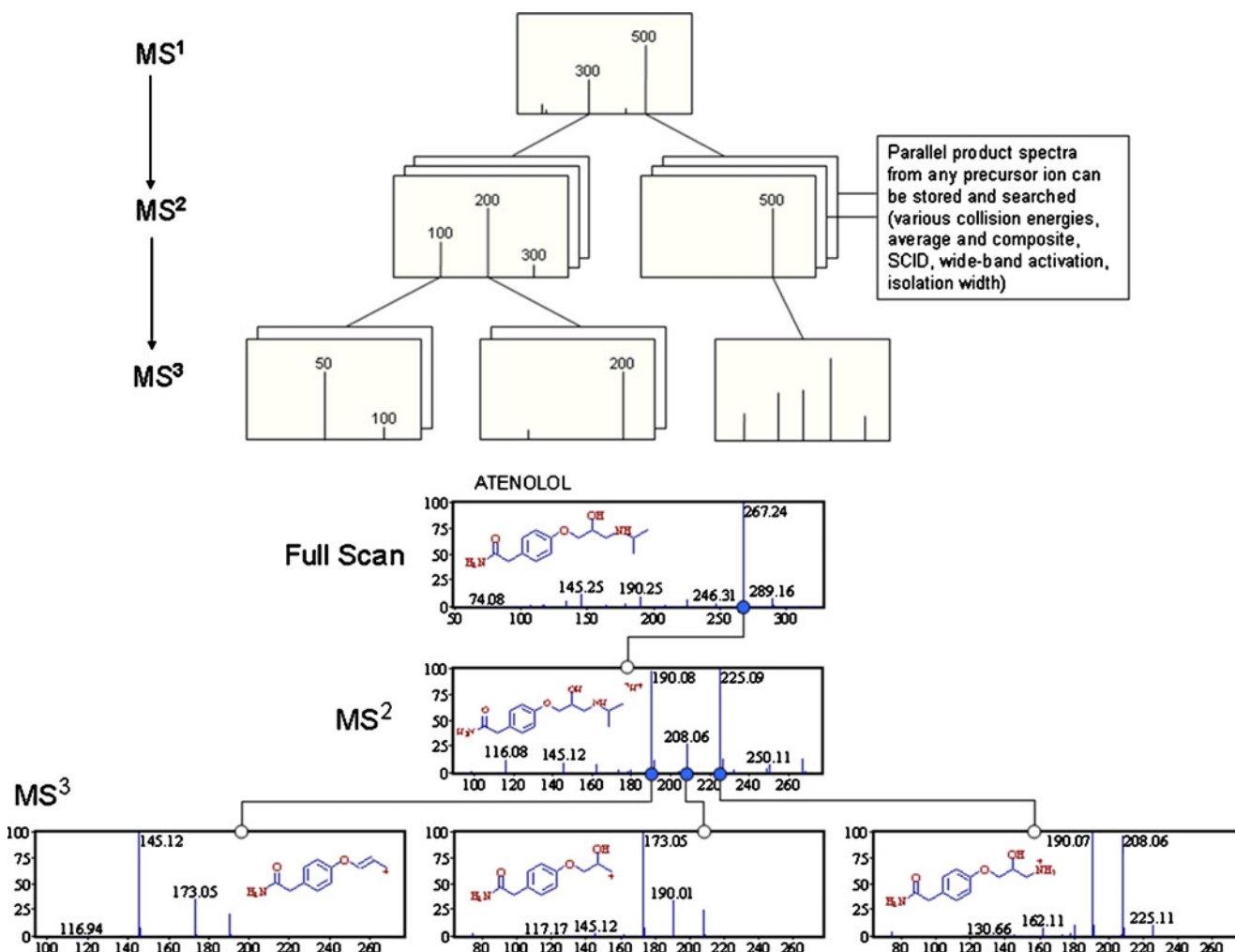


Fig. 11 A spectral tree diagram from MassFrontier representing multiple-stage MSⁿ spectra, in-source CID spectra or zoom spectra. Any stage can be searched and is logically connected with different product ion spectra (reproduced with permission from Robert Mistrík/HighChem Ltd)

reduction methods with principal component analysis [258] are helpful for the investigation of unresolved and complex organic matter (petroleum, coal, sediments, and fulvic acids) [259, 260].

Electron ionization and chemical ionization mass spectrometry

Electron ionization at 70 eV is a very hard ionization resulting in very complex rearrangements and fragmentations [261]. The EI mass spectra itself are very reproducible, which is important for a mass spectral library search. The ions in the gas phase have no “memory” where they originate from. That renders the structural interpretation of full scan EI mass spectra very complex. Electron ionization MS/MS with accurate masses may ease that problem [262]. Several book chapters discuss most important aspects of CI [6, 263]. One interesting aspect of chemical ionization is that multiple ionization gases with different proton acidities can be used,

which results in different molecular ions for correct molecular ion and elemental composition determination. Although most GC-MS instruments are capable of performing CI analysis, the use of chemical ionization GC-MS is not common anymore. One reason may be the non-existence of chemical ionization mass spectral libraries and the lower sensitivity during chemical ionization GC-MS measurements. Nevertheless, chemical ionization GC-MS remains an attractive technique for structural identifications due to the capability of obtaining abundant molecular ions.

Electrospray and atmospheric pressure chemical ionization

The study of the fragmentation behavior of compounds under electrospray conditions (ESI) [11, 264] is an important topic due to the wide availability of LC-MS devices with ESI interfaces. Using high-resolution CID data, compound substructures were ranked using a systematic bond disconnection approach [265]. In a similar approach for the

structural investigation of MS/MS product ion spectra, the authors of a freely available software used a brute-force ab initio combinatorial approach to generate possible fragment ions [266, 267], and they concluded that it is “a non-trivial task to accomplish.” Currently, only MassFrontier contains a large fragmentation reaction library as discussed in the section below. Different voltage settings should be selected for complete coverage of fragmentations. Automatic solutions such as CID voltage ramping exist [268] for obtaining maximum fragmentation patterns. A lookup table of common neutral losses during CID fragmentation has also been published [269], and typical fragmentations for atmospheric pressure ionization are discussed in Ref. [270].

Determination of stereochemistry using mass spectrometry

The determination of stereochemical (absolute) configuration usually requires a separation technique such as GC, CE, or LC with chiral columns. ESI-MS was used to determine the binding affinities of ion-molecule reactions by performing CID experiments of host-guest complexes [271]. It is possible to determine the chirality of molecules without pre-separation using chiral selector agents and ESI-MS/MS [272]. Additionally, traveling wave ion mobility spectrometry can be used to determine stereochemistry. The book titled “Applications of Mass Spectrometry to Organic Stereochemistry” [273] discusses practical approaches for stereochemical investigations of molecules.

Determination of 3D conformations using mass spectrometry

Although conformational changes of small molecules can be monitored using mass spectrometry, this approach was usually applied to high molecular weight compounds such as peptides and proteins [274] with the requirement of high resolving power. Mainly, protein folding and dynamics [275] have been studied in recent years. It has been reported that small-molecule mass spectra show differences depending on the 3D conformation of the molecule [276]. The determination of the conformational changes of small molecules is possible using ion mobility mass spectrometers or hybrids thereof. This approach requires the experimental determination of cross sections from known molecules and the use of such data for theoretical models [276, 277].

Biotransformation reactions and drug metabolism studies with mass spectrometry

Biotransformation and drug metabolism studies play a crucial role in all analytical studies targeted at drug design for phase I and phase II metabolites [278]. The tools and approaches discussed in this section are aimed to identify or

predict in vivo metabolites from cytochrome P450 (CYP) enzymes and guide through preclinical drug metabolism and pharmacokinetics, and absorption, distribution, metabolism, and excretion/Tox studies. More than 50 CYPs are known in humans, and CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4, and CYP3A5 enzymes metabolize 90% of drugs [279]. In pharmacokinetics and metabolism studies, the pathway of one single drug and all related enzymatically transformed metabolites are investigated. Levsen et al. [280] discuss the utilization of tandem mass spectrometry for the investigation of phase II metabolites. In recent years, software expert algorithms for metabolite predictions have been developed, and this includes tools such as DEREK, Catabol, LHASA, MetaboGen, METEOR, and MetabolExpert [281–284]. The software works along known metabolic transformation rules and performs an in silico prediction of possible metabolites. Those metabolite structures can be identified later either by mass accurate mass shifts or by tandem mass spectrometry [285]. Specialized mass spectrometry centric software from vendors such as Metabolite ID (AB Sciex), Metabolynx (Waters), Networks (Thermo), MassHunter/Metabolite ID (Agilent), and MetaboliteTools (Bruker) mostly use a combination of accurate mass, neutral loss, and biotransformation rules with associated accurate masses for metabolite identification.

Iontrap and triple quadrupole mass spectrometers can be used to monitor and identify common neutral losses (including methylation, acetylation, and glucuronidation). Tables with common biotransformations, and lists of metabolic changes and their accurate masses can be found in Ref. [286] and Ref. [287] (see Table 2). With the broader availability of accurate mass spectrometers, the mass defect filter rule [288–291] could be applied. Using the nominal mass shift and a mass defect window with several milliDalton (mDa) widths, the matrix influence can be separated from the analytes of interest. A recent report discussed the integration of structure-based metabolism prediction with predicted and experimental MS/MS data [292]. The information from precursor and product ion spectra can be used to find common biotransformations from possible regioisomers [293]. The approaches presented here can also be used during pesticide screening, biotechnology, and bioengineering studies with enzymatic reaction systems and metabolic profiling studies.

The availability of hybrid triple quadrupole mass spectrometers with linear ion traps (QTRAP) allows the sensitive detection of metabolites using multiple reaction monitoring (MRM) and a subsequent MS/MS (product ion) scan for metabolite identification or annotation [194, 195, 294]. A newly developed software (LightSight and ABI/Sciex) [295–297] can automatically create MRM or multiple ion monitoring transitions. This software approach, called

Table 2 Selected biotransformations for in vivo drug metabolism studies detectable by accurate mass spectrometry (reproduced from [282] with permission of Future Science Ltd)

Metabolic reaction	Mass change (Da)	Molecular formula change
Phase I: CYP, FMO		
Hydroxylation	+15.9949	+O
Dihydroxylation	+31.9898	+O ₂
N/S-oxidation	+15.9949	+O
Epoxidation	+15.9949	+O
Demethylation	-14.0157	-CH ₂
Desethylation	-28.0312	-C ₂ H ₄
Oxidative displacement of chlorine	-17.9662	-Cl+OH
Oxidative displacement of fluorine	-1.9957	-F+OH
Phase I: reductases (e.g., CYP)		
Nitro reduction	-29.9742	+H ₂ -O ₂
Reductive displacement of fluorine	-17.9906	-F+H
Phase I: dehydrogenases (ADH, ALDH), aldoketoreductases (e.g., ALK)		
Dehydrogenation (oxidation)	-2.0157	-H ₂
Hydrogenation (reduction)	+2.0157	+H ₂
Phase I: other enzymes (not easily assignable)		
Dehydration	-18.0106	-H ₂ O
Reductive displacement of chlorine	-33.9611	-Cl+H
Decarboxylation	-43.9898	-CO ₂
Loss of nitro group	-44.9851	-NO ₂ +H
Alcohol to carboxylic acid	+13.9792	-2H+O
Ketone formation	+13.9792	-2H+O
Hydration	+18.0106	+H ₂ O
Methyl to carboxylic acid	+29.9741	-2H+O ₂
Phase II		
Glucuronide conjugation	+176.0321	+C ₆ H ₈ O ₆
Sulfate conjugation	+79.9568	+SO ₃
Methylation	+14.0157	+CH ₂
Acetylation	+42.0106	+C ₂ H ₂ O
Glycine conjugation	+57.0215	+C ₂ H ₃ NO
Taurine conjugation	+107.0041	+C ₂ H ₅ NO ₂ S
Glutathione conjugation	+305.0682	+C ₁₀ H ₁₅ N ₃ O ₆ S
Cysteine conjugation	+119.0041	+C ₃ H ₅ NO ₂ S
N-Acetylcysteine conjugation	+161.0147	+C ₅ H ₇ NO ₃ S

predictive MRM, allows for a very sensitive analysis and detection of new metabolites [298].

Isotope labeling studies

Stable isotopic labeling studies [193, 299] and hydrogen/deuterium exchange reactions [300–302] are commonly applied in drug metabolism studies. Proteomics approaches use labeling studies for the quantification of peptides and proteins [303–305] as well as mass defect isotopomer studies [306, 307]. In vivo labeling with stable isotopes can be applied for metabolism studies in plants [308, 309], isotopomer-based flux balance analysis [310–315], and structural elucidation of unknown compounds [197, 316–

319]. The use of deuterated mobile phases (D₂O) or post-column infusion of D₂O has been popular over the last several years for metabolite identification studies [320–322].

Determination of impurities and contaminants

The elucidation of impurities is a reoccurring event during daily lab work. Contaminants can be avoided either by experience or better quality control sets of all reagents and solvents used. For GC-MS, LC-MS, and CE-MS, this includes the purchase of solvents and reagents in batch to obtain consistent quality and the use of quality check monitoring procedures. These chromatograms or mass spectra (solvent blanks or reagent blanks) need to be stored

long term to monitor impurities over month and years. Existing collections of fragments and ions can help during the investigation of such contaminations. Certain detergents and buffer components (Triton X) are excellently ionized in ESI mode and result in large abundant peaks that suppress the signal of other ions. A comprehensive review [323] discusses mostly ESI and MALDI interferences and contains a large EXCEL sheet in the supplementary data section that covers around 800 potential interferences and contaminant ions in positive and negative mode electrospray mode. Additionally, it also contains 40 repetitive fragments such as sodium formate clusters (NaHCO_2) and lists multiple adducts, losses, and possible replacements. A constant batch-wise monitoring of the purity of solvents and derivatization agents is important along with the removal of artifacts from datasets for GC-MS [324]. The hot injector in GC-MS can act as a small chemical reactor, and this could introduce a series of breakdown products that can lead to false analysis conclusions [325]. Many volatile compounds, among them pesticides and insecticides (DDT), easily decompose in a hot injector [326]. Using a GC cold injection system with a near zero degree Celsius injection temperature to avoid the breakdown of chemicals and an automatic liner exchange (ALEX, Gerstel Inc.) to avoid carryover can increase the level of confidence in compound identification of complex samples [327]. A chip-based nanoelectrospray system (NanoMate, Advion Inc.) can be used to avoid cross-contaminations. For each sample, a new ESI nozzle is used during direct infusion mass spectrometry experiments.

Mass spectral fragmentation reaction databases

Mass spectral fragmentation reaction databases contain chemical reactions and fragmentation mechanisms from mass spectral investigations. These are organic reactions proposals drawn by mass spectrometrists in order to explain specific fragments or mass spectral abundances. If applied to new molecules or mass spectra, they can speed up the elucidation process by using existing knowledge. Until recently, no structure searchable mass spectral fragmentation library existed. Currently, only MassFrontier (HighChem Ltd.) contains a large fragmentation library of 30,936 fragmentation schemes with 129,229 reactions and 151,762 associated structures. Direct molecule search, substructure search, similarity search, and name search can be performed, and all associated meta-data are electronically searchable. The database was manually curated from several thousand publications (see Fig. 12, reproduced with permission of HighChem Ltd.) and can be used to develop *in silico* fragmentation predictions as discussed in the next chapter.

The current practice of dissemination of chemical fragmentation reactions on paper publications (PDF) is not

keeping up with existing technological possibilities. It is impractical to search compound structures and reaction data from paper publications. Also, many data centric approaches, including the development of novel fragmentation algorithms, are actively hindered. Chemical reaction and fragmentation data should be submitted in electronic, machine-readable exchange formats to journals or external repositories. Currently, no such repository for mass spectral reaction data exists.

Mass spectral simulation and generation of *in silico* mass spectra

Chemical compound databases currently cover more than 50 million chemical structures; however, only around one million mass spectra (including duplicates) from known compounds exist. This gap could be filled by computer generation of mass spectra from large compound structure databases. An *in silico* algorithm has to predict accurate mass fragments and their abundances. Such an *in silico* generation of theoretical mass spectra could be useful because experimentally obtained mass spectra can then be matched against large *in silico* mass spectral databases. Several mass spectral simulation algorithms have been published in the literature. Many of those programs, however, were never made commercially or publicly available, which therefore prevents any possible independent scientific validation. The main problem of most algorithms is to simulate or calculate peak abundances or peak intensities [328] that reflect experimentally measured peak abundances [329–331]. This problem has not been solved for the vast majority of small molecules under different ionization modes. The success rate of any algorithm has to be determined by a validation study using unknown molecules and a library match of the *in silico* generated spectra against the experimental spectra. Furthermore, the structural diversity and the number of compounds have to be high to avoid overfitting.

Successful cases of *in silico* generation are known for molecules with certain structural scaffolds and consistent fragmentation patterns. That includes lipids (see Fig. 13), oligosaccharides [332], glycans [333], and peptides [334]. For example, compound libraries from combinatorial synthesis show common neutral mass losses when studied under electrospray conditions [335]. Another study used neural networks to simulate 70-eV electron ionization mass spectra of alkanes [336]. MASSIS/MASSIMO was a rule-based spectral simulation system for electron ionization spectra that included McLafferty rearrangements, retro-Diels-Alder reaction, neutral loss, and oxygen migration [337–339]. Another method was developed for the prediction presence of carboxylic acids using low-energy CID spectra and CO₂ (44 Da) loss in MS/MS product ions [340]. The publicly available MetFrag algorithm [341] compares *in silico* mass

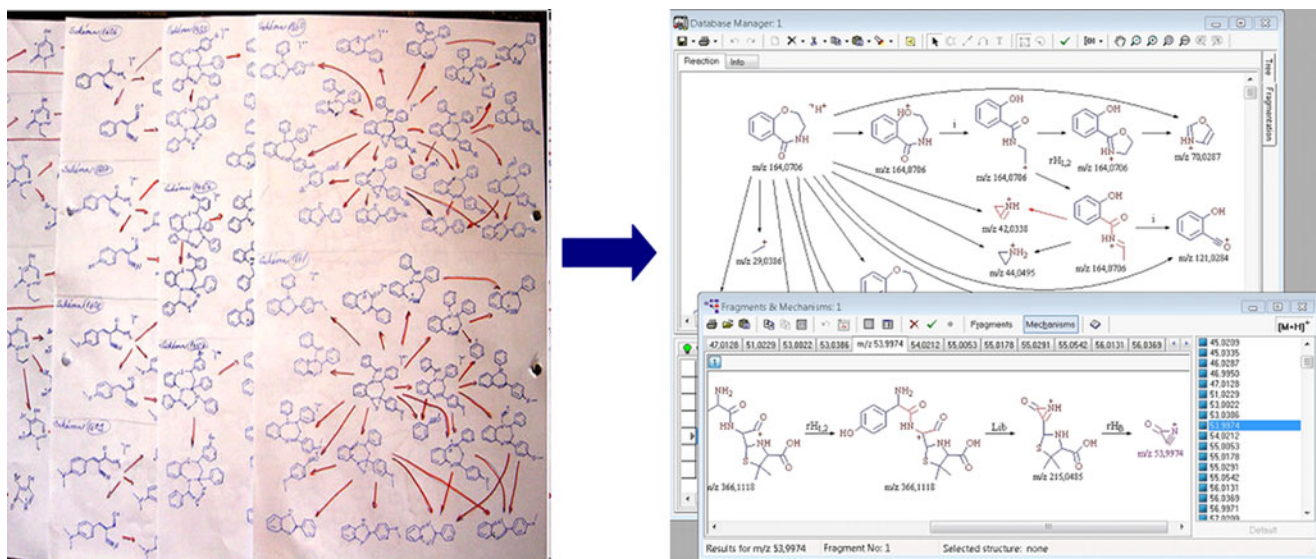


Fig. 12 A mass spectral fragmentation pathway database containing 30,936 fragmentation mechanisms. The mass spectrometry community never enthusiastically endorsed digital data sharing. Therefore, most of

the spectra and reaction data had to be captured from old paper publications (reproduced with permission from Robert Mistrík/High-Chem Ltd)

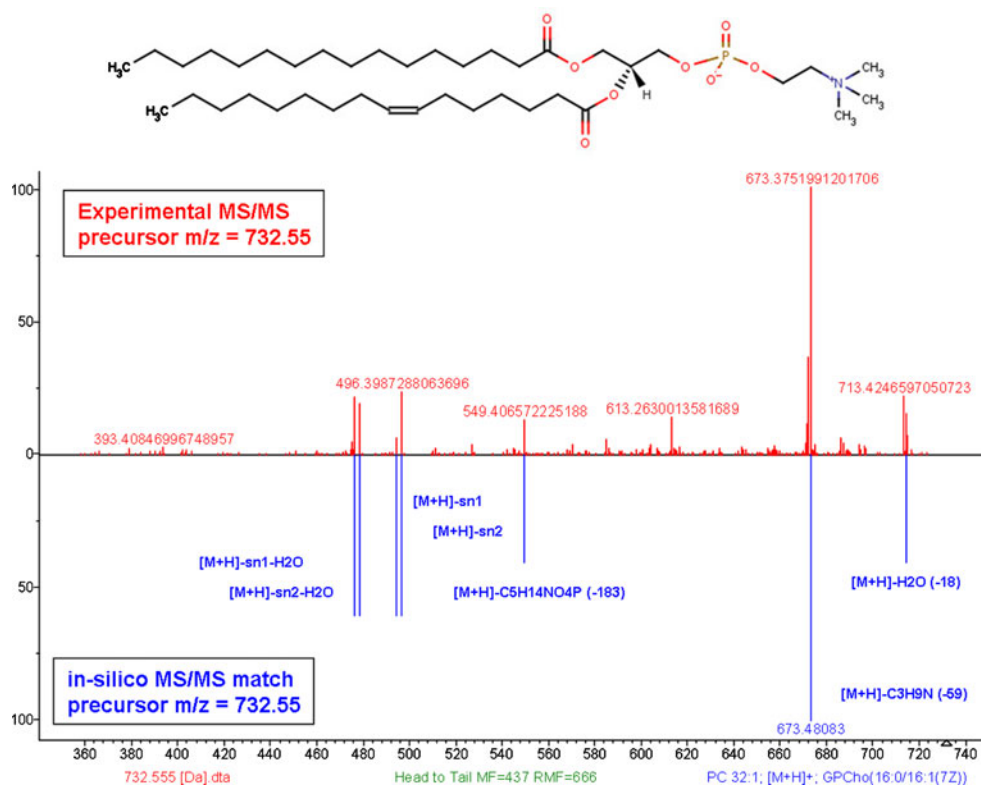
spectra, obtained by a bond dissociation approach, with experimental mass spectra and assigns a score to all results. A validation study [342] compared the success rate of three commercial programs (MOLGEN-MSF (University Bayreuth), MS Fragmenter (Advanced Chemistry Development Inc.) [343, 344], and Mass Frontier (HighChem Ltd.)) and came to the conclusion that the simulation of

mass spectral fragmentations of electron ionization spectra is still far from daily practical usability.

Expert systems for mass spectral interpretation

Computer-aided interpretation of mass spectra started in the 1960s [345, 346] when the first commercial computers

Fig. 13 An experimental phospholipid spectrum and computer generated MS/MS spectrum. Mass spectral libraries of theoretical in silico spectra can be generated from large structure databases (source: Tobias Kind/FiehnLab)



were available. The DENDRAL project pioneered approaches with the aim of predicting isomer structures from mass spectra using self-learning or artificial intelligence algorithms [347]. There are several software tools that can assist during interpretation of mass spectra, including Automated Mass Spectral Deconvolution and Identification System (AMDIS), MassFrontier, ACD/MS Manager, MASSLib [348], and the freely available NIST MS Interpreter as part of the NIST08 database search program. The NIST MS search program can generate substructure information using a nearest-neighbor approach [349] by searching unknown mass spectra against a large reference database. The algorithm will generate a good list and a bad list of substructures based on an actual hit list. If there is no mass spectrum with similar features in the database, then the algorithm fails. The tools AMDIS and MOLGEN-MS [350, 351] integrate the Varmuza feature-based classification approach [352–354]. Mass spectral classifiers for neutral loss selection using Fisher ratio and linear discriminant analysis and genetic algorithm partial least squares discriminant analysis have been investigated to distinguish alcohols and ethers [353]. A decision tree-based prediction of substructures from mass spectral features allowed the classification of unknown metabolites into different compound classes [355]. For soft ionization techniques (ESI, APCI), programs such as HighChem Mass Frontier [87, 205, 227, 356–360] or ACD/MS Manager [361] can help during data interpretation and fragmentation prediction. Older software usually works well with unit resolution data. New software should allow the handling of accurate and high-resolution mass spectral data. There is currently no software or “magic bullet” that combines mass spectral knowledge and scientific intuition and is able to present a correct compound structure from mass spectral data only.

Use of computational chemistry to explain gas-phase phenomena

There is a constant series of papers that use computational chemistry to investigate gas-phase reactions or ionization processes in regard to thermochemistry and kinetics [362]. In some cases, this approach can lead to a better understanding of fragmentation pathways. The book titled “Assigning structures to ions in mass spectrometry” covers many small-molecule-related approaches regarding thermochemistry, including potential energy curves, calculation of heats of formation, and proton affinities [363]. Quantum mechanical methods can also be used to determine bond cleavage energies and bond dissociation energies [364], and help to interpret adduct formation [365, 366]. Proton affinities have been calculated with semiempirical methods (AM1) [367] and density functional theories (DFT) on the MP2 and B3LYP level [368, 369]. Investigation of CID

cross sections can be used to determine binding affinities of cations and small molecules [370]. The kinetic method with entropy correction can be used to calculate proton and electron affinities [371, 372]. Ab initio and DFT calculations were used to elucidate the energetics of ECD [96]. A recent paper discussed the application of DFT to understand tandem mass spectrometric (MS/MS) fragmentation for non-peptidic molecules [373]. The report from three example molecules shows that protonation significantly perturbs the electron density and affects ion formation and subsequent bond fragmentation throughout the whole molecule. The fragmentation pathways for phthalates [374] were investigated using DFT. Even chirality detection of molecules is possible by means of electrospray ionization mass spectrometry and competitive binding analysis [375]. Many of the applied quantum chemical methods require a deep computational chemistry knowledge and can make use of available software tools such as GAMESS, GAUSSIAN, NWCHEM, or AMBER [376]. Moreover, just recently released Intel Xeon (Nehalem) and AMD Opteron (Magny-Cours) processor technology allows for the needed computational speed on commodity desktop computers. The performance of 200 GFlop/s (Giga floating point operations per second; double-precision mode) is comparable with speeds only reached by supercomputers 10 years ago. Both the high software and hardware barrier have made computational interpretations of mass spectra interesting for research, but they have not yet translated into easy to use software tools for mass spectrometry practitioners.

Approaches for hyphenated techniques (GC-MS and LC-MS)

Mass spectral deconvolution for clean mass spectra

Mass spectral deconvolution refers to the process of creating background- and noise-free mass spectra from GC-MS or LC-MS data. Traditionally, chromatographers would use a simple chromatographic peak detection method and would manually select a detected peak to obtain the related mass spectrum. This manual process is error prone and time consuming, and requires manual background subtraction in front and in the back of the chromatographic peak. With an automated deconvolution, routine peaks can be detected under the baseline total ion chromatogram or overlapping peaks can be resolved (see Fig. 14). Additionally, if the chromatographic resolution is not sufficient, then the process is also able to separate (deconvolute) overlapping compound mass spectra. The automated deconvolution process itself is now standard in many GC-MS investigations [377] and is mostly known from the freely available AMDIS [378]. The AMDIS process includes four sequential steps:

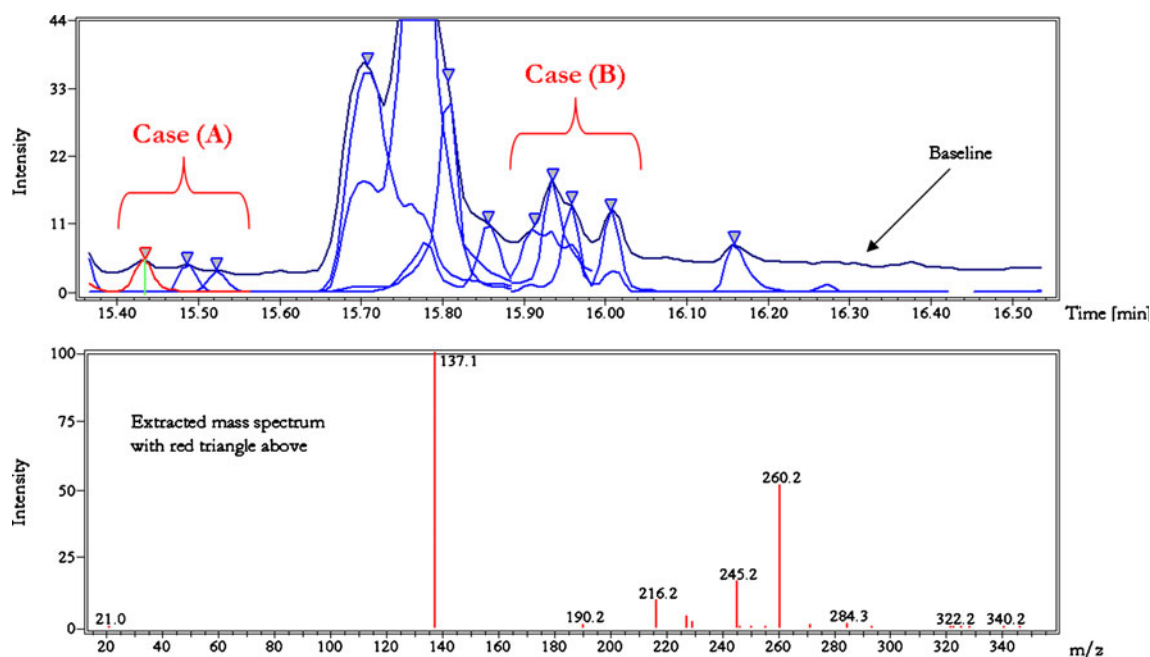


Fig. 14 Peak picking and mass spectral deconvolution. The program can automatically detect peaks under the baseline (*case A*). Overlapping (non-resolved) peaks can be detected, and clean mass spectra are extracted (*case B*) (source: Tobias Kind/FiehnLab created with MassFrontier)

(1) noise analysis, (2) component perception, (3) spectral deconvolution, and (4) compound identification. AMDIS was recently adapted to monitor air quality and identify toxic gases on board of the International Space Station [379, 380]. Multiple other software solutions for the analysis of GC-MS and LC-MS exist [381]. That includes LECO ChromaTOF, SpectralWorks AnalyzerPro, Ion Signature Quantitative Deconvolution Software, HighChem MassFrontier, and TargetSearch [382]. The use of peak picking and peak detection algorithms for LC-MS data [114, 383] is still an active field of research due to high noise ratios, broader chromatographic peaks, and mass spectra that show less fragments than electron ionization spectra. The deconvolution process itself usually performs best if it is optimized for a specific scan rate; otherwise, false-positive and false-negative peak detections may occur [384]. The detection of these deconvolution errors [385, 386] is best solved by using reference compound mixes with a known number of analytes and a subsequent optimization process to detect the correct number of compounds. Deconvoluted compound mass spectra are subsequently submitted to a mass spectral database search. If additional MS/MS spectra were extracted, then a tandem mass spectral search can be performed.

Chromatographic heart cut, column switching, and fractionation techniques

The fractionation of complex samples using liquid chromatography is an often performed technical step to obtain pure compounds or reduce the complexity of the sample. This

further allows 1D and 2D NMR investigations of complex natural products [387]. Peaks can be frozen out using a preparative fraction collector in conjunction with a low-efficiency preparative packed GC column or higher film thickness megabore traps for gas chromatography [388]. These applications are exemplified in biomarker research [389], investigation of hydrocarbons [390, 391], and entomology and pheromone studies [392, 393]. Using column switching of hydrophilic interaction chromatography (HILIC) and reversed-phase (RP) columns, complex samples can be analyzed within one single run [394]. To increase the chromatographic peak, capacity columns with different polarity can be coupled together into a 2D-LC-MS setup. Comprehensive two-dimensional liquid chromatography (LC \times LC) is currently in a developmental stage [395, 396]. The enrichment of samples using peak parking [397] or fraction collection [398] is commonly used during natural product investigations and drug research [399–402]. The combination of liquid chromatography with solid phase extraction and NMR has been applied for pharmaceutical studies [403], drug discoveries [404], food investigations [405, 406], and natural product research [407–410]. When this technique is combined with mass spectrometric detectors (LC-SPE-NMR-MS), an almost universal system for structure elucidation is created [411–414].

Cheminformatics meets mass spectrometry

Modern mass spectrometry centric approaches for structure elucidation cannot be performed without proper molecular

structure handling [415]. Many of the software tools for structural elucidation (MassFrontier, ACD/MS Manager, NIST MS Search, Sierra's APEX) also have inbuilt structure handling capabilities to either allow substructure analysis or perform structure–spectra correlations. Many drug metabolism studies also include computational chemistry approaches. It is also important to investigate how many resonance structures [416], tautomers [417–420], and stereoisomers [421] can be generated from a given structure. For the ionization processes [422], it is favorable to understand the ionization behavior by calculating charges, electronegativities, and H-bond donor and acceptor counts. It is important to calculate the distribution of microspecies and pK_a values [423] under different pH values in a given buffer system [424] to estimate the retention behavior. Software tools such as the Marvin Calculator Plugins (ChemAxon) [425] can calculate all those compound properties, including all possible tautomers and stereoisomers, with a single program. Multiple candidate structures can be investigated using the commercially available ChemAxon Instant-JChem program or the open-source BioClipse software [426].

Structure retention relationships and retention index predictions

The investigation and accurate prediction of the retention behavior of a molecule are a major cornerstone for structure elucidation using mass spectrometry. The theoretical predicted retention index or retention time can be used as a powerful orthogonal filter for hyphenated chromatographic techniques. If the elemental composition and possible substructures can be detected from the mass spectrum, then this information can be used within molecular isomer generators (MOLGEN-MS, SMOG [427], and Assemble [428]) to generate all structural isomers [429]. Those molecular isomer generators usually work with the constraint of a molecular formula and a good list and bad list of possible substructures. The retention index (RI) prediction algorithm could then be used to predict the retention index or retention time of these virtual compounds. Subsequently, these theoretical RIs can be matched against the experimental RI values, and all compounds outside a specific retention index window can be removed as false candidates. These prediction algorithms are very accurate for a small subset of structures but lack wider substance coverage, or they cover a broad range of structural classes but lack prediction accuracy. A model with a good correlation coefficient may still exercise bad prediction power. Additionally, the development datasets for comprehensive solutions should have a minimum size of 500–1,000 compounds, which are best acquired under the same conditions and the same instrumental setup. The obtained quantitative structure retention relationship (QSRR) models must be carefully validated with a large number of

external test set compounds to avoid overfitting [430]. It has to be stated that published QSRR models without an existing commercial or open software implementation are interesting scientific exercises, but they are relatively useless for the majority of practitioners because they cannot apply or use these models.

Several QSRR models for gas and liquid chromatography have been published and already reviewed in the recent literature [431]. Kaliszán wrote a series of papers regarding structure retention relationship models culminating in a single comprehensive review [432]. Both reviews cover several hundred scientific papers. Katritzky discussed the use quantum chemical descriptors for QSRR calculations [433]. HPLC retention indices were calculated from a set of 500 drug-like compounds, and molecular descriptors and neural network machine learning were applied for the prediction of the RI values [434, 435].

A large commercial database of Kovats GC retention index values was released in 2005 [436]. Table 3 lists selected GC column parameters from this database. As the column types and film thicknesses cover a wide range of possible parameters, retention indices also differ. Such variations must be included during model development and require careful statistical evaluation [437] because one single compound can have a high variability in observed experimental retention indices. A freely available software for the prediction of Kovats retention indices (based on alkanes) was released by NIST in 2007 [438]. The software was developed with RI values from 35,000 compounds and used a group contribution method of 85 different substructures for polar and nonpolar column data. The median error for polar columns was 65 RI units, which is not accurate enough to determine single structures, but can be used as a refinement filter for comprehensive structure elucidation workflows. Several standard compounds including keto-alkanes [439], alkylarylketones [440], or 1-nitroalkanes [441] were proposed as retention index markers in the past. No universal or unified HPLC retention index system for RP, normal phase, and HILIC has been developed yet. These standard compounds should cover a wide retention time range on a given LC phase, and they should be easily ionizable with electrospray ionization, non-toxic, non-reactive, inexpensive, commercially available, and outside of targeted profiling approaches. Synthetic peptides could be amicable compound structures for HPLC RI values [442].

Derivatization strategies for LC-MS and GC-MS

The detection of functional groups with the help of selective derivatization [55] is one of the oldest wet-lab techniques in chemistry. The book “A Handbook of Derivatives for Mass Spectrometry” by Zaikin and Halket

Table 3 Statistics of selected column parameters of semipolar stationary gas chromatographic phases obtained from the NIST05 retention index database

GC column parameters (semi-standard nonpolar)	Count	Percent
Column type		
DB-5	11,688	42.34
HP-5	6,310	22.86
SE-54	2,242	8.12
BPX-5	1,377	4.98
DB-5MS	1,207	4.37
SE-52	1,131	4.09
HP-5MS	845	3.06
RTX-5	720	2.60
CP Sil 8 CB	516	1.86
SPB-5	468	1.69
Column length		
30 m	18,445	66.83
25 m	2,490	9.02
50 m	2,478	8.97
60 m	2,339	8.47
Missing		
15 m	347	1.25
20 m	245	0.88
Film thickness		
0.25 μm	18,251	66.12
Missing	3,721	13.48
1 μm	1,593	5.77
0.5 μm	1,039	3.76
0.52 μm	454	1.64
0.33 μm	351	1.27
0.11 μm	281	1.01

Accurate retention index prediction algorithms must consider different column types and film thicknesses

comprehensively covers most derivatization reactions for different ionization modes, and LC-MS- and GC-MS-based mass spectrometric studies [443]. In the case of GC-MS, the aim is to increase volatility of the compound and protect reactive groups to avoid thermal breakdown or reactions with the column material. For LC-MS studies, derivatizations are performed to improve ionization characteristics for poorly ionizable compounds [444–447]. The obtained products must be hydrolysis stable as for example in the *tert*-butyldimethylchlorosilane products from a *N*-methyl-*N*-[*tert*-butyldimethyl-silyl]trifluoroacetamide derivatization [448]. Common fields of application are pesticide and environmental screening [449–451], separation of complex sugars as mono-, di-, and trisaccharides [452, 453], enantiomer analysis [454], amino acid analysis [447, 455], steroid and drug testing [456, 457], and metabolomic profiling studies [458–460].

Use of structure databases for targeted compound annotation

The availability of large public compound databases, such as PubChem [461] and ChemSpider [462], or specialized drug and metabolism databases, such as KEGG [463], HMDB [464], ChEBI [465], DrugBank [466], MZedDB [467], and the Chemical Lookup Service [468], allow for a web-based search of molecular formulae or accurate masses [469]. DrugBank has a search interface allowing an accurate mass search in positive or negative mode within the known human metabolite pool, and the results are presented with possible adducts and link to further database sources. This information can include literature, chemical taxonomy data, or other related information. If other molecular features are known from mass spectral investigations, such as the number of polar hydrogens from derivatization or H/D exchange experiments, then these molecular properties can be used as additional orthogonal filters. Additionally, theoretical retention indices can be used to match experimental RI values and remove false candidate hits. Multiple databases have advanced programming interfaces (APIs) that allow a connection of standalone programs with online databases without the need of downloading several gigabytes of the database itself.

Fields of applications—review of reviews

The use of mass spectrometric analysis in metabolomics has been reviewed in [470–473]. A comprehensive review covered the identification of known endogenous and exogenous metabolites by applying accurate mass, isotopic pattern filter, retention indices, and mass spectral fragmentation in a sequential filter cascade and combing the results with a database search [474]. A wide range of LC-MS-based methods including MRM-based approaches, precursor ion scans, and radio-labeling were discussed [475]. Multistage mass spectrometry approaches were used for the identification of drugs, metabolites, toxins, and plant and animal metabolites [124, 476–486] with their associated fragmentation pathways. The structural elucidation of flavanoids, flavonoid glycosides [487–491], and drug metabolites [294, 318, 492–498] using multiple-stage tandem mass spectrometry was reviewed in several papers. Natural product investigations usually combine mass spectral information with a de novo structure elucidation step using NMR [499–502]. Lipid and phospholipid analysis [503–512] can be performed with all ionization modes and types of mass spectrometers including triple quadrupole, ion trap, and TOF mass spectrometers. Structure elucidation of compounds from environmental samples [118, 389, 450, 513–516] is among the most complex cases of structure elucidation. The use of capillary electrophoresis coupled to mass spectrometry has been previously reviewed [517, 518].

Electronic data sharing of mass spectra

The future success of structure elucidation with mass spectrometry will largely depend on the development of new software algorithms. Similar to the success of the bioinformatics [519] and the proteomics [520] communities, which had open access to large genome associated data, mass spectrometric data must be made publicly available to enable long-term data reuse and allow data-driven research [521, 522]. Multiple software database implementations are currently in use or in development [523, 524], among them SPECTRa [525], MeltDB [526], SetupX [527], MassBank [219, 220], MMCD [528], METLIN [529, 530], GMD [458], KNApSAcK [531], and PRIME [532]. Mass spectra from a wide range of instrument types can be used in machine learning approaches for mass spectral elucidation. The future of mass spectral structure elucidation will depend on a wide array of well-described, meta-data enhanced and freely available resources. Not only high- and low-resolution mass spectra but also their suggested fragmentation pathways can be electronically collected. Mass spectra and associated molecular structure drawings need to be shared in open exchange formats and global repositories. This will create a new breed of scientists who only deal with mass spectrometric data evaluation independent from access to mass spectrometers just as in bioinformatics.

Unfortunately, there were never any data sharing policies released by the mass spectrometric community itself. The American Society for Mass Spectrometry, which is open to all scientists worldwide and is the leading mass spectrometric society worldwide, never actively pushed or developed data sharing principles for the community. Driven by community efforts, however, the proteomics community [533] and several funding agencies, including the Wellcome Trust (UK), National Cancer Institute (NCI, USA), and National Institutes of Health (NIH, USA), released the International Summit on Proteomics Data Release and Sharing Policy [534], which urges the rapid release of mass spectra, tandem MS, and liquid chromatography MS into the public domain. On the technical level [535], the data sharing problem can be solved with large repositories such as the PeptideAtlas.org or peer-to-peer (P2P) approaches as in the Tranche project [536] at ProteomeCommons.org [537]. Open exchange formats must be further developed that can store multidimensional data. This can only be done with the support of the mass spectrometry industry, which in recent years also opened up parts of their proprietary programming interfaces (APIs) to allow open-source programmers access to specific data formats.

The Human Proteome Organisation was among the leading organizations and supported the open mzData and mzML formats [538], which was later joined to the upcoming mzML format [539]. Different organizations,

such as the Institute for Systems Biology (mzXML, PepXML, and ProtXML) and the Proteomics Standard Initiative (mzData and PSIMI) started the development of those standard formats and provided vendor specific converters [537, 540, 541]. The continuous success of each of those formats requires broad support for vendor specific converter software and additional software that can visualize and manipulate such exchange data. The netCDF [542] and JCAMP-DX [543] format are widely supported within the GC-MS community. Both formats suffer from non-existing accurate mass and MSⁿ implementations and the lack of broader community development support within the mass spectrometry community.

It has been shown from the crystal structure community, that once data and exchange standards are established, no human interaction is needed anymore to collect spectral data [525]. The CrystalEye project (<http://wwmm.ch.cam.ac.uk/crystaleye/>) shows that the aggregation of crystal structures can be totally robotized using modern web technologies. The only requirement is that the spectral data must be available under open-data licenses (<http://www.opendefinition.org/>) [544]. The public availability and open mass spectral resources will allow commercial [545] as well as governmental entities (NIST) [546] that are specialized in collection of mass spectral data to focus on the expensive curation of these data. The enhancement of these spectral datasets with meta-information and compound structures will add value to those collections and allow commercial distribution due to market demands.

We have shown in a recent research publication regarding the data sharing of compound, spectral, and meta-data [521] that parsing bitmap data to obtain multidimensional high-resolution mass spectral data is not keeping up with the today's technological possibilities. There is a tremendous loss of information associated, and it is impossible to investigate (enlarge or zoom) such bitmap or paper-based mass spectra. During the spectral capturing process, many peaks disappear because their associated accurate mass values cannot be obtained (see Fig. 15). Additionally, molecular structure data capturing is an error-prone process. The scientific value of such open-access and open-data shared mass spectral collections, their structures, and associated reaction data, will outweigh initial hesitations as learned from the genomics community.

Future software and hardware advances (opinion)

The success of structure elucidation using mass spectrometric approaches depends not only on technical machine developments but also much more on the development of better software algorithms. Particularly, software for working with data output from multiple-stage mass spectrometry (MSⁿ) or

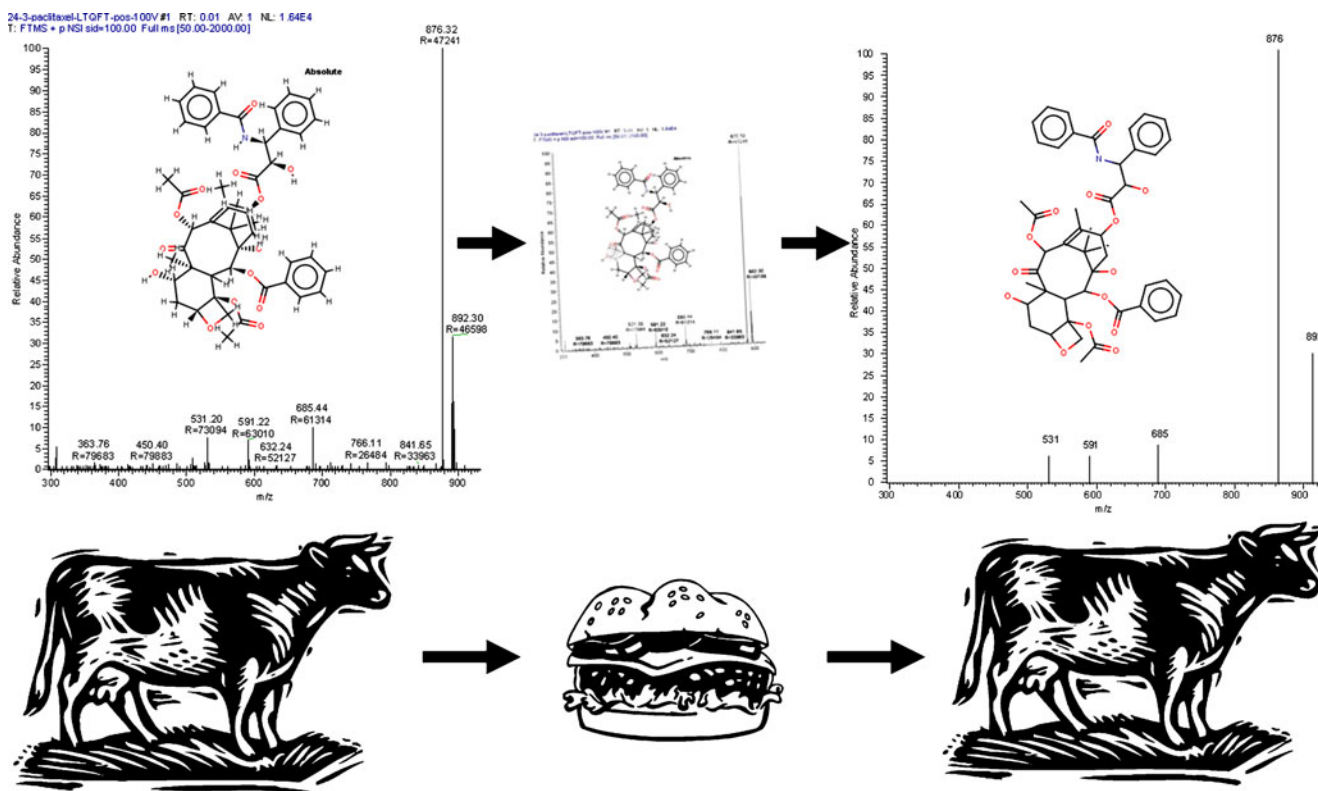


Fig. 15 Capturing high-resolution mass spectral data from paper publications is an error-prone process. The final machine readable structure usually does not represent the original spectrum (hamburger-

to-cow algorithm). New digital data-sharing principles need to be set in place (source: Tobias Kind/FiehnLab)

data from multiple orthogonal hybrids, including ion mobility, is not yet fully developed. The large proteomics software community with a very active bioinformatics development branch could be a positive example for the small-molecule community. In terms of software development, there will be always a very innovative core of commercial and open-source software developers that will develop state-of-the-art software tools. For taxpayer-funded research in universities and government-funded laboratories, the direction should go towards open-source software or at least towards freely or publicly available software with the least restrictive software licenses to allow commercial and non-commercial exploitation. Community efforts can solve many of the complex software challenges, whereas consistent software support, user help, and error fixing can be obtained from commercial services. The publication and discussion of approaches or programs that are not commercially or publicly available should be avoided because claims made within the publication cannot be independently verified. One of the most important issues is the public sharing of mass spectra and other spectral data from a wide variety of mass spectrometers. This may ultimately lead to an evolution of scientists and software developers that specialized in software development for small-molecule identification using mass spectrometry.

In terms of a technological process, it must be stated that in principle, all technological prerequisites for advancement in structure elucidation exist. The hyphenation of LC-MS with NMR seems to be a fast lane towards successful structure elucidation. Hybrid orthogonal approaches that add an additional dimension to the chromatographic side (GC×GC and LC×LC) or mass spectrometric part (ion mobility coupled to time of flight) will particularly enable the extraction of cleaner mass spectra from fully resolved compounds. Increased resolving power, better mass and isotopic abundance accuracy, and high data acquisition rate will enable a faster structure elucidation process. Accurate masses and high resolution for all multiple-stage mass spectra (MSⁿ) will subsequently allow the evolution of new software tools as discussed in this article.

Conclusions

Structure elucidation using mass spectrometry is a challenging field of research with many success stories. Mass spectrometry itself is seldom used for the de novo structure elucidation of small molecules but serves as an important building block together with NMR, IR, X-ray crystallography, and other spectroscopic techniques. Together with hyphenated

chromatographic techniques, (GC and LC) mass spectrometry serves as a powerful tool for the elucidation of drugs, pesticides, metabolites, and complex chemical mixtures. Mass spectrometry hardware is currently in a very advanced stage with many technologies not fully exploited yet. More data centric approaches have to be taken in the future. This includes the electronic publishing of investigated structures and their associated multiple-stage mass spectra with open-data licenses. The ultimate success of structure elucidation of small molecules lies in better software programs and the development of sophisticated tools for data evaluation of high-resolution and accurate mass multiple-stage (MS^n) mass spectral data.

Acknowledgements We thank Vladimir Tolstikov (UC Davis Genome Center Metabolomics CoreLab) for an internal review of the paper. We also thank Google Scholar (<http://scholar.google.com>) and Google Books for pleasant scholarly information retrieval. Additional thanks are provided to the Jane Journal Estimator (www.biosemantics.org), CAS SciFinder Scholar, and SCOPUS (www.scopus.com) for literature analysis.

Funding This work was supported by the grant R01 ES013932 of the NIEHS/US National Institutes of Health and the US National Science Foundation MCB-0520140. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests The authors declare no competing interests.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Brunnée C (1987) The ideal mass analyzer: fact or fiction? *Int J Mass Spectrom Ion Processes* 76(2):125–237
2. Gelpi E (2008) From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part I 1965–1984. *J Mass Spectrom* 43(4):419–435
3. Gelpi E (2009) From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part II 1985–2000. *J Mass Spectrom* 44(8):1137–1161
4. Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* 72(6):1156–1162
5. Perry RH, Cooks RG, Noll RJ (2008) Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom Rev* 27(6):661–699
6. Gross JH (2004) *Mass spectrometry: a textbook*. Springer, Berlin
7. Lattimer RP, Schulten HR (1989) Field ionization and field desorption mass spectrometry: past, present, and future. *Anal Chem* 61(21):1201A–1215A
8. Gordin A, Fialkov AB, Amirav A (2008) Classical electron ionization mass spectra in gas chromatography/mass spectrometry with supersonic molecular beams. *Rapid Commun Mass Spectrom* 22(17):2660–2666
9. Fenn JB (2003) Electrospray wings for molecular elephants (Nobel Lecture)*. *Angew Chem Int Ed* 42:3871–3894
10. Gaskell SJ (1997) Electrospray: principles and practice. *J Mass Spectrom* 32(7):677–688
11. Cech NB, Enke CG (2001) Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom Rev* 20(6):362–387
12. Schmidt A, Karas M, Dulcks T (2003) Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI? *J Am Soc Mass Spectrom* 14(5):492–500
13. Lazar IM, Grym J, Foret F (2006) Microfabricated devices: a new sample introduction approach to mass spectrometry. *Mass Spectrom Rev* 25(4):573–594
14. Pereira-Medrano AG, Sterling A, Snijders APL, Reardon KF, Wright PC (2007) A systematic evaluation of chip-based nano-electrospray parameters for rapid identification of proteins from a complex mixture. *J Am Soc Mass Spectrom* 18(9):1714–1725
15. Almeida R, Mosoarca C, Chirita M, Udrescu V, Dinca N, Vukeli Z, Allen M, Zamfir AD (2008) Coupling of fully automated chip-based electrospray ionization to high-capacity ion trap mass spectrometer for ganglioside analysis. *Anal Biochem* 378(1):43–52
16. Lydic TA, Busik JV, Esselman WJ, Reid GE (2009) Complementary precursor ion and neutral loss scan mode tandem mass spectrometry for the analysis of glycerophosphatidylethanolamine lipids from whole rat retina. *Anal Bioanal Chem* 394(1):267–275
17. Wickremsinhe ER, Singh G, Ackermann BL, Gillespie TA, Chaudhary AK (2006) A review of nanoelectrospray ionization applications for drug metabolism and pharmacokinetics. *Curr Drug Metab* 7(8):913–928
18. Koulman A, Cao M, Faville M, Lane G, Mace W, Rasmussen S (2009) Semi-quantitative and structural metabolic phenotyping by direct infusion ion trap mass spectrometry and its application in genetical metabolomics. *Rapid Commun Mass Spectrom: RCM* 23(15):2253–2263
19. Kelly RT, Page JS, Marginean I, Tang K, Smith RD (2008) Nanoelectrospray emitter arrays providing interemitter electric field uniformity. *Anal Chem* 80(14):5660–5665
20. Su S, Gibson GT, Mugo SM, Marecak DM, Oleschuk RD (2009) Microstructured photonic fibers as multichannel electrospray emitters. *Anal Chem* 81(17):7281–7287
21. Kostianinen R, Kotiaho T, Kuuranne T, Auriola S (2003) Liquid chromatography/atmospheric pressure ionization–mass spectrometry in drug metabolism studies. *J Mass Spectrom* 38(4):357–372
22. Rosenberg E (2003) The potential of organic (electrospray—and atmospheric pressure chemical ionisation) mass spectrometric techniques coupled to liquid-phase separation for speciation analysis. *J Chromatogr A* 1000(1–2):841–889
23. Hayen H, Karst U (2003) Strategies for the liquid chromatographic-mass spectrometric analysis of non-polar compounds. *J Chromatogr A* 1000(1–2):549–565
24. Covey TR, Thomson BA, Schneider BB (2009) Atmospheric pressure ion sources. *Mass Spectrom Rev* 28(6):870–897
25. Robb DB, Covey TR, Bruins AP (2000) Atmospheric pressure photoionization: an ionization method for liquid chromatography–mass spectrometry. *Anal Chem* 72(15):3653–3659
26. Marchi I, Rudaz S, Veuthey JL (2009) Atmospheric pressure photoionization for coupling liquid–chromatography to mass spectrometry: a review. *Talanta* 78(1):1–18
27. Robb DB, Blades MW (2008) State-of-the-art in atmospheric pressure photoionization for LC/MS. *Anal Chim Acta* 627(1):34–49
28. Bos SJ, Van Leeuwen SM, Karst U (2006) From fundamentals to applications: recent developments in atmospheric pressure photoionization mass spectrometry. *Anal Bioanal Chem* 384(1):85–99

29. Zenobi R, Knochenmuss R (1998) Ion formation in MALDI mass spectrometry. *Mass Spectrom Rev* 17(5):337–366
30. Dreisewerd K (2003) The desorption process in MALDI. *Chem Rev* 103(2):395–425
31. Knochenmuss R, Zenobi R (2003) MALDI ionization: the role of in-plume processes. *Chem Rev* 103(2):441–452
32. Miyamoto K, Fujimaki S, Ueda Y (2009) Development of a new electron ionization/field ionization ion source for gas chromatography/time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*: RCM 23(20):3350–3354
33. Qian K, Edwards KE, Siskin M, Olmstead WN, Mennito AS, Dechert GJ, Hoosain NE (2007) Desorption and ionization of heavy petroleum molecules and measurement of molecular weight distributions. *Energy Fuels* 21(2):1042–1047
34. Schaub TM, Hendrickson CL, Qian K, Quinn JP, Marshall AG (2003) High-resolution field desorption/ionization Fourier transform ion cyclotron resonance mass analysis of nonpolar molecules. *Anal Chem* 75(9):2172–2176
35. Cody RB, Laramée JA, Durst HD (2005) Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal Chem* 77(8):2297–2302
36. Weston DJ (2010) Ambient ionization mass spectrometry: current understanding of mechanistic theory; analytical performance and application areas. *Analyst* 135(4):661–668
37. Yeping Z, Michelle L, Danlin W, Rowena M (2008) Quantification of small molecules in plasma with direct analysis in real time tandem mass spectrometry, without sample preparation and liquid chromatographic separation. *Rapid Commun Mass Spectrom* 22(20):3217–3224
38. Van Berkel GJ, Pasilis SP, Ovchinnikova O (2008) Established and emerging atmospheric pressure surface sampling/ionization techniques for mass spectrometry. *J Mass Spectrom* 43(9):1161–1180
39. McDonnell LA, Heeren RMA (2007) Imaging mass spectrometry. *Mass Spectrom Rev* 26(4):606–643
40. Cornett DS, Reyzer ML, Chaurand P, Caprioli RM (2007) MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat Meth* 4(10):828–833
41. Broersen A, van Liere R, Altelaar AFM, Heeren RMA, McDonnell LA (2008) Automated, feature-based image alignment for high-resolution imaging mass spectrometry of large biological samples. *J Am Soc Mass Spectrom* 19(6):823–832
42. Rubakhin SS, Jurchen JC, Monroe EB, Sweedler JV (2005) Imaging mass spectrometry: fundamentals and applications to drug discovery. *Drug Discov Today* 10(12):823–837
43. Manicke NE, Wiseman JM, Ifa DR, Cooks RG (2008) Desorption electrospray ionization (DESI) mass spectrometry and tandem mass spectrometry (MS/MS) of phospholipids and sphingolipids: ionization, adduct formation, and fragmentation. *J Am Soc Mass Spectrom* 19(4):531–543
44. Northen TR, Yanes O, Northen MT, Marrinucci D, Uritboonthai W, Apon J, Golledge SL, Nordstrom A, Siuzdak G (2007) Clathrate nanostructures for mass spectrometry. *Nature* 449(7165):1033–1036
45. Nordstrom A, Want E, Northen T, Lehtio J, Siuzdak G (2007) Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics. *Anal Chem* 80(2):421–429
46. Kothari S, Song Q, Xia Y, Fico M, Taylor D, Amy JW, Stafford G, Cooks RG (2009) Multiplexed four-channel rectilinear ion trap mass spectrometer. *Anal Chem* 81(4):1570–1579
47. Schiewek R, Lorenz M, Giese R, Brockmann K, Benter T, Gäb S, Schmitz OJ (2008) Development of a multipurpose ion source for LC-MS and GC-API MS. *Anal Bioanal Chem* 392(1):87–96
48. Schmitt-Kopplin P, Englmann M, Rossello-Mora R, Schiewek R, Brockmann KJ, Benter T, Schmitz OJ (2008) Combining chip-ESI with APLI (cESILI) as a multimode source for analysis of complex mixtures with ultrahigh-resolution mass spectrometry. *Anal Bioanal Chem* 391(8):2803–2809
49. Byrdwell WC (2009) Dual parallel mass spectrometry for lipid and vitamin D analysis. *J Chromatogr A* 1217:3992–4003
50. Hannis JC, Muddiman DC (2000) A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 11(10):876–883
51. Nepomuceno AI, Muddiman DC, Bergen Iii HR, Craighead JR, Burke MJ, Caskey PE, Allan JA (2003) Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 75(14):3411–3418
52. Siegel MM, Tabei K, Lambert F, Candela L, Zoltan B (1998) Evaluation of a dual electrospray ionization/atmospheric pressure chemical ionization source at low flow rates (50 l/min) for the analysis of both highly and weakly polar compounds. *J Am Soc Mass Spectrom* 9(11):1196–1203
53. Liang X, Xia Y, McLuckey SA (2006) Alternately pulsed nanoelectrospray ionization/atmospheric pressure chemical ionization for ion/ion reactions in an electrodynamic ion trap. *Anal Chem* 78(9):3208–3212
54. Short LC, Hanold KA, Cai SS, Syage JA (2007) Electrospray ionization/atmospheric pressure photoionization multimode source for low-flow liquid chromatography/mass spectrometric analysis. *Rapid Commun Mass Spectrom* 21(10):1561–1566
55. Halket JM, Waterman D, Przyborowska AM, Patel RKP, Fraser PD, Bramley PM (2005) Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. *J Exp Bot* 56(410):219–243
56. Balogh MP (2004) Source design and the utility of multimode ionization. *Spectroscopy* 19(1):52–56
57. Fischer SM, Perkins PD (2005) Simultaneous electrospray and atmospheric pressure chemical ionization: the science behind the Agilent multimode ion source. In: Agilent—Technical report
58. Whitehouse C, White T, Willoughby R, Sheehan E (2006) Atmospheric pressure ion source for mass spectrometry. In: WO Patent WO/2006/107,831
59. Wang Y, Griffiths WJ (2007) Mass spectrometry for metabolite identification. In: William JG (ed) *Metabolomics, metabonomics and metabolite profiling*. Royal Society of Chemistry, London
60. McLuckey SA, Wells JM (2001) Mass analysis at the advent of the 21st century. *Chem Rev* 101(2):571–606
61. Balogh MP (2004) Debating resolution and mass accuracy. *LC GC N Am* 22(2):118–131
62. Mann M, Kelleher NL (2008) Precision proteomics: the case for high resolution and high mass accuracy. *Proc Natl Acad Sci* 105(47):18132
63. Sleno L, Volmer DA, Marshall AG (2005) Assigning product ions from complex MS/MS spectra: the importance of mass uncertainty and resolving power. *J Am Soc Mass Spectrom* 16(2):183–198
64. Kellmann M, Muenster H, Zomer P, Mol H (2009) Full scan MS in comprehensive qualitative and quantitative residue analysis in food and feed matrices: how much resolving power is required? *J Am Soc Mass Spectrom* 20(8):1464–1476
65. Marshall AG, Hendrickson CL (2008) High-resolution mass spectrometers. *Annu Rev Anal Chem* 1(1):579–599
66. Marshall AG (2000) Milestones in Fourier transform ion cyclotron resonance mass spectrometry technique development. *Int J Mass Spectrom* 200(1–3):331–356
67. Feng X, Siegel MM (2007) FTICR-MS applications for the structure determination of natural products. *Anal Bioanal Chem* 389(5):1341–1363

68. Park Y, Lebrilla CB (2005) Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides. *Mass Spectrom Rev* 24(2):232–264
69. Ackermann BL, Berna MJ, Eckstein JA, Ott LW, Chaudhary AK (2008) Current applications of liquid chromatography/mass spectrometry in pharmaceutical discovery after a decade of innovation. *Annu Rev Anal Chem* 1:357–396
70. Makarov A, Denisov E, Lange O (2009) Performance evaluation of a high-field orbitrap mass analyzer. *J Am Soc Mass Spectrom* 20(8):1391–1396
71. Ow SY, Noirel J, Salim M, Evans C, Watson R, Wright PC (2010) Balancing robust quantification and identification for iTRAQ: application of UHR-ToF MS. *Proteomics* 10(11):2205–2213
72. Mondello L, Tranchida PQ, Dugo P, Dugo P (2008) Comprehensive two-dimensional gas chromatography-mass spectrometry: a review. *Mass Spectrom Rev* 27(2):101
73. Dalluge J, Beens J, Brinkman UAT (2003) Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool. *J Chromatogr A* 1000(1–2):69–108
74. Frysinger GS, Gaines RB, Reddy CM (2002) GCxGC—a new analytical tool for environmental forensics. *Environ Forensics* 3(1):27–34
75. Bedair M, Sumner LW (2008) Current and emerging mass-spectrometry technologies for metabolomics. *Trends Anal Chem* 27(3):238–250
76. Staack RF, Hopfgartner G (2007) New analytical strategies in studying drug metabolism. *Anal Bioanal Chem* 388(7):1365–1380
77. Fjeldsted JC (2009) Accurate mass measurements with orthogonal axis time-of-flight mass spectrometry. In: Ferrer I, Thurman EM (eds) *Liquid chromatography time-of-flight mass spectrometry*. Wiley, New York
78. Glish GL, Burinsky DJ (2008) Hybrid mass spectrometers for tandem mass spectrometry. *J Am Soc Mass Spectrom* 19(2):161–172
79. Stroth JG, Petucci CJ, Brecker SJ, Huang N, Lau JM (2007) Automated sub-ppm mass accuracy on an ESI-TOF for use with drug discovery compound libraries. *J Am Soc Mass Spectrom* 18(9):1612–1616
80. Kind T, Fiehn O (2006) Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinform* 7(234):1–10
81. Kind T, Fiehn O (2007) Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinform* 8(1):105
82. Wolfender JL, Waridel P, Ndjoko K, Hobby KR, Major HJ, Hostettmann K (2000) Evaluation of Q-TOF-MS/MS and multiple stage IT-MSn for the dereplication of flavonoids and related compounds in crude plant extracts. *Analisis* 28(10):895–906
83. Jonscher KR, Yates JR (1997) The quadrupole ion trap mass spectrometer—a small solution to a big challenge. *Anal Biochem* 244(1):1–15
84. Li R, Zhou Y, Wu Z, Ding L (2006) ESI-QqTOF-MS/MS and APCI-IT-MS/MS analysis of steroid saponins from the rhizomes of *Dioscorea panthaica*. *J Mass Spectrom* 41(1):1–22
85. McCormack P, Jones P, Hetheridge MJ, Rowland SJ (2001) Analysis of oilfield produced waters and production chemicals by electrospray ionisation multi-stage mass spectrometry (ESI-MSn). *Water Res* 35(15):3567–3578
86. Dinan L (2006) Dereplication and partial identification of compounds. *Meth Biotechnol* 20:297
87. Lim HK, Chen J, Sensenhauser C, Cook K, Subrahmanyam V (2007) Metabolite identification by data-dependent accurate mass spectrometric analysis at resolving power of 60 000 in external calibration mode using an LTQ/Orbitrap. *Rapid Commun Mass Spectrom* 21(12):1821–1832
88. Payne AH, Glish GL (2005) Tandem mass spectrometry in quadrupole ion trap and ion cyclotron resonance mass spectrometers. *Meth Enzymol* 402:109–148
89. Cooke N (2003) Recent advances in mass spectrometry for drug discovery and development. *Drug Discov* 27
90. Futrell JH (2000) Development of tandem mass spectrometry: one perspective. *Int J Mass Spectrom* 200(1–3):495–508
91. De Hoffmann E (1996) Tandem mass spectrometry: a primer. *J Mass Spectrom* 31(2):129–137
92. Dwivedi P, Wu P, Klopsch SJ, Puzon GJ, Xun L, Hill HH (2008) Metabolic profiling by ion mobility mass spectrometry (IMMS). *Metabolomics* 4(1):63–80
93. Trimpin S, Tan B, Bohrer BC, O'Dell DK, Merenbloom SI, Pazos MX, Clemmer DE, Walker JM (2009) Profiling of phospholipids and related lipid structures using multidimensional ion mobility spectrometry–mass spectrometry. *Int J Mass Spectrom* 287(1–3):58–69
94. Weston DJ, Bateman R, Wilson ID, Wood TR, Creaser C (2005) Direct analysis of pharmaceutical drug formulations using ion mobility spectrometry/quadrupole-time-of-flight mass spectrometry combined with desorption electrospray ionization. *Anal Chem* 77(23):7572–7580
95. Kertesz TM, Hall LH, Hill DW, Grant DF (2009) CE50: quantifying collision induced dissociation energy for small molecule characterization and identification. *J Am Soc Mass Spectrom* 20(9):1759–1767
96. Syrstad EA, Turecek F (2005) Toward a general mechanism of electron capture dissociation. *J Am Soc Mass Spectrom* 16(2):208–224
97. Cooper HJ, Hakansson K, Marshall AG (2005) The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom Rev* 24(2):201–222
98. Zubarev RA (2004) Electron-capture dissociation tandem mass spectrometry. *Curr Opin Biotechnol* 15(1):12–16
99. Zubarev RA, Zubarev AR, Savitski MM (2008) Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? *J Am Soc Mass Spectrom* 19(6):753–761
100. Good DM, Wirtala M, McAlister GC, Coon JJ (2007) Performance characteristics of electron transfer dissociation mass spectrometry. *Mol Cell Proteomics* 6(11):1942–1951
101. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci USA* 101(26):9528–9533
102. Sleno L, Volmer DA (2004) Ion activation methods for tandem mass spectrometry. *J Mass Spectrom* 39(10):1091–1112
103. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19(3):242–247
104. Adahchour M, Beens J, Vreuls RJJ, Brinkman UAT (2006) Recent developments in comprehensive two-dimensional gas chromatography (GC×GC): I. Introduction and instrumental set-up. *TrAC Trends Anal Chem* 25(5):438–454
105. Marriott P, Shellie R (2002) Principles and applications of comprehensive two-dimensional gas chromatography. *TrAC, Trends Anal Chem* 21(9–10):573–583
106. Ryan D, Marriott P (2007) Comprehensive two-dimensional gas chromatography. *Adv Chromatogr* 46:451
107. Skoczynska E, Korytar P, Boer JD (2008) Maximizing chromatographic information from environmental extracts by GC×GC-ToF-MS. *Environ Sci Technol* 42(17):6611–6618
108. Gauthier GL, Grimm R (2006) Miniaturization: chip-based liquid chromatography and proteomics. *Drug Discov Today Technol* 3(1):59–66

109. Biller JE, Biemann K (1974) Reconstructed mass spectra, a novel approach for the utilization of gas chromatograph–mass spectrometer data. *Anal Lett* 7(7):515–528
110. Dromey RG, Stefik MJ, Rindfleisch TC, Duffield AM (1976) Extraction of mass spectra free of background and neighboring component contributions from gas chromatography/mass spectrometry data. *Anal Chem* 48(9):1368–1375
111. Phalp JM, Payne AW, Windig W (1995) The resolution of mixtures using data from automated probe mass spectrometry. *Anal Chim Acta* 318(1):43–53
112. Windig W, Payne AW (1997) Noise and background reduction method for component detection in chromatography/spectrometry. In: US Patent 5,672,869
113. Gray ZA, Abel RH (1995) Mass spectra interpretation system including spectra extraction. In: US Patent 5,453,613
114. Windig W, Phalp JM, Paynes AW (1996) A noise and background reduction method for component detection in liquid chromatography/mass spectrometry. *Anal Chem* 68(20):3602–3606
115. Zhang H, Ma L, He K, Zhu M (2008) An algorithm for thorough background subtraction from high-resolution LC/MS data: application to the detection of troglitazone metabolites in rat plasma, bile, and urine. *J Mass Spectrom* 43(9):1191–1200
116. Zhang H, Yang Y (2008) An algorithm for thorough background subtraction from high-resolution LC/MS data: application for detection of glutathione-trapped reactive metabolites. *J Mass Spectrom* 43(9):1181–1190
117. Cuyckens F, Hurkmans R, Castro-Perez JM, Leclercq L, Mortishire-Smith RJ (2009) Extracting metabolite ions out of a matrix background by combined mass defect, neutral loss and isotope filtration. *Rapid Commun Mass Spectrom* 23(2):327–332
118. Thurman EM, Ferrer I, Barcelo D (2001) Choosing between atmospheric pressure chemical ionization and electrospray ionization interfaces for the HPLC/MS analysis of pesticides. *Anal Chem* 73(22):5441–5449
119. Kauppila TJ, Wiseman JM, Ketola RA, Kotiaho T, Cooks RG, Kostiaimen R (2006) Desorption electrospray ionization mass spectrometry for the analysis of pharmaceuticals and metabolites. *Rapid Commun Mass Spectrom* 20(3):387–392
120. Zhu JH, Cole RB (2000) Formation and decompositions of chloride adduct ions, $[M+Cl]^{-}$, in negative ion electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* 11(11):932–941
121. Zhu JH, Cole RB (2001) Ranking of gas-phase acidities and chloride affinities of monosaccharides and linkage specificity in collision-induced decompositions of negative ion electrospray-generated chloride adducts of oligosaccharides. *J Am Soc Mass Spectrom* 12(11):1193–1204
122. Schug K, McNair HM (2002) Adduct formation in electrospray ionization. Part 1: common acidic pharmaceuticals. *J Sep Sci* 25(12):760–766
123. Ventola E, Vainiotalo P, Suman M, Dalcanale E (2006) ESI-FTICR mass spectrometric study of alcohol complexation properties of mono- and diphosphonate-bridged cavitands. *J Am Soc Mass Spectrom* 17(2):213–221
124. Zhong W, Irish PA, Martin GE (2009) Influence of differently ionized species on fragmentation pathways and energetics of a potential adenosine receptor antagonist using a triple quadrupole and a multistage LTQ-Orbitrap TM FTMS instrument. *J Heterocycl Chem* 46(4):591–598
125. Wujcik CE, Kadar EP (2008) Reduction of in-source collision-induced dissociation and thermolysis of sulopenem prodrugs for quantitative liquid chromatography/electrospray ionization mass spectrometric analysis by promoting sodium adduct formation. *Rapid Commun Mass Spectrom* 22(20):3195–3206
126. Hsu FF, Bohrer A, Turk J (1998) Formation of lithiated adducts of glycerophosphocholine lipids facilitates their identification by electrospray ionization tandem mass spectrometry. *J Am Soc Mass Spectrom* 9(5):516–526
127. Huang N, Siegel MM, Kruppa GH, Laukien FH (1999) Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and E-mailing of high-resolution exact-mass electrospray ionization mass spectral data. *J Am Soc Mass Spectrom* 10(11):1166–1173
128. Karas M, Kruger R (2003) Ion formation in MALDI: the cluster ionization mechanism. *Chem Rev* 103(2):427–440
129. Kuhl C, Tautenhahn R, Neumann S: CAMERA—rule based annotation of isotopes and adducts. <http://bioconductor.org/packages/devel/bioc/html/CAMERA.html>
130. Balogh MP (2006) Spectral interpretation, part II: tools of the trade, MS—the practical art. *LC GC* 24:762–769
131. Gorlach E, Richmond R (1999) Discovery of quasi-molecular ions in electrospray spectra by automated searching for simultaneous adduct mass differences. *Anal Chem* 71(24):5557–5562
132. Horn DM, Zubarev RA, McLafferty FW (2000) Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *J Am Soc Mass Spectrom* 11(4):320–332
133. Carvalho PC, Cociorva D, Wong CCL, Carvalho Mda G, Barbosa VC, Yates JR (2009) Charge prediction machine: tool for inferring precursor charge states of electron transfer dissociation tandem mass spectra. *Anal Chem* 81(5):1996–2003
134. Allard E, Troger RA, Arvidsson B, Sjoberg PJR (2010) Quantitative aspects of analyzing small molecules—monitoring singly or doubly charged ions? A case study of ximelagatran. *Rapid Commun Mass Spectrom* 24(4):429–435
135. Sterling HJ, Williams ER (2009) Origin of supercharging in electrospray ionization of noncovalent complexes from aqueous solution. *J Am Soc Mass Spectrom* 20(10):1933–1943
136. Jaitly N, Mayampurath A, Littlefield K, Adkins J, Anderson G, Smith R (2009) Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. *BMC Bioinform* 10(1):87
137. Loss RD (2003) Atomic weights of the elements 2001. *Pure Appl Chem* 75(8):1107–1122
138. Brenton AG, Godfrey AR (2010) Accurate mass measurement: terminology and treatment of data. *J Am Soc Mass Spectrom* (in press)
139. Bristow AWT, Webb KS (2003) Intercomparison study on accurate mass measurement of small molecules in mass spectrometry. *J Am Soc Mass Spectrom* 14(10):1086–1098
140. Bereman MS, Lyndon MM, Dixon RB, Muddiman DC (2008) Mass measurement accuracy comparisons between a double-focusing magnetic sector and a time-of-flight mass analyzer. *Rapid Commun Mass Spectrom* 22(10):1563–1566
141. Makarov A, Scigelova M (2010) Coupling liquid chromatography to Orbitrap mass spectrometry. *J Chromatogr A* 1217(25):3938–3945
142. Scheppele SE, Grindstaff QG, Grigsby RD, McDonald SR, Hwang CS (1983) Computer processing of mass-spectral data. II. Assignment of formulas to experimental masses illustrating program FZM; effects of electron mass and functional form of exponential scan law on accuracy of mass measurement. *Int J of Mass Spectrom Ion Phys* 49(2):179–209
143. Ferrer I, Thurman EM (2005) Measuring the mass of an electron by LC/TOF-MS: a study of “Twin Ions”. *Anal Chem* 77(10):3394–3400
144. Beier T, Häffner H, Hermanspahn N, Karshenboim SG, Kluge HJ, Quint W, Stahl S, Verdú J, Werth G (2001) New determination of the electron’s mass. *Phys Rev Lett* 88(1):011603
145. Scheltema RA, Kamleh A, Wildridge D, Ebikeme C, Watson DG, Barrett MP, Jansen RC, Breitling R (2008) Increasing the mass

- accuracy of high-resolution LC-MS data using background ions—a case study on the LTQ-Orbitrap. *Proteomics* 8(22):4647–4656
146. Mihaleva VV, Vorst O, Maliepaard C, Verhoeven HA, de Vos RCH, Hall RD, van Ham R (2008) Accurate mass error correction in liquid chromatography time-of-flight mass spectrometry based metabolomics. *Metabolomics* 4(2):171–182
 147. Southam AD, Payne TG, Cooper HJ, Arvanitis TN, Viant MR (2007) Dynamic range and mass accuracy of wide-scan direct infusion nano-electrospray Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal Chem* 79:4595–4602
 148. Payne TG, Southam AD, Arvanitis TN, Viant MR (2009) A signal filtering method for improved quantification and noise discrimination in Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics data. *J Am Soc Mass Spectrom* 20(6):1087–1095
 149. Petyuk VA, Mayampurath AM, Monroe ME, Polpitiya AD, Purvine SO, Anderson GA, Camp DG, Smith RD (2009) DtaRefinery, a software tool for elimination of systematic errors from parent ion mass measurements in tandem mass spectra data sets. *Molecular & Cellular Proteomics* 9:486–496. doi:10.1074/mcp.M900217-MCP200
 150. Kuehl D, Wang Y (2006) Peak shape calibration method improves the mass accuracy of mass spectrometers. *BioPharm Int* 19(7):32–38, +45
 151. Gu M, Wang Y, Zhao XG, Gu ZM (2006) Accurate mass filtering of ion chromatograms for metabolite identification using a unit mass resolution liquid chromatography/mass spectrometry system. *Rapid Commun Mass Spectrom* 20(5):764–770
 152. Urayama S, Zou W, Brooks K, Tolstikov V (2010) Comprehensive mass spectrometry based metabolic profiling of blood plasma reveals potent discriminatory classifiers of pancreatic cancer. *Rapid Commun Mass Spectrom* 24(5):613–620
 153. Williams DK Jr, Kovach AL, Muddiman DC, Hanck KW (2009) Utilizing artificial neural networks in MATLAB to achieve parts-per-billion mass measurement accuracy with a Fourier transform ion cyclotron resonance mass spectrometer. *J Am Soc Mass Spectrom* 20(7):1303–1310
 154. Savitski MM, Ivonin IA, Nielsen ML, Zubarev RA, Tsybin YO, Hakansson P (2004) Shifted-basis technique improves accuracy of peak position determination in Fourier transform mass spectrometry. *J Am Soc Mass Spectrom* 15(4):457–461
 155. Yang L, Amad M, Winnik WM, Schoen AE, Schweingruber H, Mylchreest I, Rudewicz PJ (2002) Investigation of an enhanced resolution triple quadrupole mass spectrometer for high-throughput liquid chromatography/tandem mass spectrometry assays. *Rapid Commun Mass Spectrom* 16(21):2060–2066
 156. Shinkawa T, Nagano K, Inomata N, Haramura M (2009) A software program for more reliable precursor ion assignment from LC-MS analysis using LTQ ultra zoom scan. *J Proteomics* 73(2):357–360
 157. Budzikiewicz H, Grigsby RD (2006) Mass spectrometry and isotopes: a century of research and discussion. *Mass Spectrom Rev* 25(1):146–157
 158. Hellborg R, Skog G (2008) Accelerator mass spectrometry. *Mass Spectrom Rev* 27(5):398–427
 159. Tuniz C, Zoppi U, Hotchkis MAC (2004) Sherlock Holmes counts the atoms. *Nucl Instrum Methods Phys Res, B Beam Interact Mater Atoms* 213:469–475
 160. Calcagnile L, Quarta G, D'Elia M (2005) High-resolution accelerator-based mass spectrometry: precision, accuracy and background. *Appl Radiat Isot* 62(4):623–629
 161. Krummen M, Hilkert AW, Juchelka D, Duhr A, Schlüter HJ, Pesch R (2004) A new concept for isotope ratio monitoring liquid chromatography/mass spectrometry. *Rapid Commun Mass Spectrom* 18(19):2260–2266
 162. Thurman EM, Ferrer I, Fernandez-Alba AR (2005) Matching unknown empirical formulas to chemical structure using LC/MS TOF accurate mass and database searching: example of unknown pesticides on tomato skins. *J Chromatogr A* 1067(1–2):127–134
 163. Ojanpera S, Pelander A, Pelzing M, Krebs I, Vuori E, Ojanpera I (2006) Isotopic pattern and accurate mass determination in urine drug screening by liquid chromatography/time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 20(7):1161–1167
 164. Zhu P, Tong W, Alton K, Chowdhury S (2009) An accurate-mass-based spectral-averaging isotope-pattern-filtering algorithm for extraction of drug metabolites possessing a distinct isotope pattern from LC-MS data. *Anal Chem* 81:5910–5917
 165. Grange AH, Genicola FA, Sovocool GW (2002) Utility of three types of mass spectrometers for determining elemental compositions of ions formed from chromatographically separated compounds. *Rapid Commun Mass Spectrom* 16(24):2356–2369
 166. Grange AH, Sovocool GW (2008) Automated determination of precursor ion, product ion, and neutral loss compositions and deconvolution of composite mass spectra using ion correlation based on exact masses and relative isotopic abundances. *Rapid Commun Mass Spectrom* 22(15):2375–2390
 167. Grange AH, Winnik W, Ferguson PL, Sovocool GW (2005) Using a triple-quadrupole mass spectrometer in accurate mass mode and an ion correlation program to identify compounds. *Rapid Commun Mass Spectrom* 19(18):2699–2715
 168. Snider RK (2007) Efficient calculation of exact mass isotopic distributions. *J Am Soc Mass Spectrom* 18(8):1511–1515
 169. Rockwood AL, Haimi P (2006) Efficient calculation of accurate masses of isotopic peaks. *J Am Soc Mass Spectrom* 17(3):415–419
 170. Li L, Kresh JA, Karabacak NM, Cobb JS, Agar JN, Hong P (2008) A hierarchical algorithm for calculating the isotopic fine structures of molecules. *J Am Soc Mass Spectrom* 19(12):1867–1874
 171. Olson MT, Yergey AL (2009) Calculation of the isotope cluster for polypeptides by probability grouping. *J Am Soc Mass Spectrom* 20(2):295–302
 172. Rockwood AL, Kushnir MM, Nelson GJ (2003) Dissociation of individual isotopic peaks: predicting isotopic distributions of product ions in MSn. *J Am Soc Mass Spectrom* 14(4):311–322
 173. Ramaley L, Herrera LC (2008) Software for the calculation of isotope patterns in tandem mass spectrometry. *Rapid Commun Mass Spectrom* 22(17):2707–2714
 174. Alon T, Amirav A (2009) Isotope abundance analysis for improved sample identification with tandem mass spectrometry. *Rapid Commun Mass Spectrom: RCM* 23(23):3668
 175. Williams JD, Weiner BE, Ormand JR, Brunner J, Thornquest AD Jr, Burinsky DJ (2001) Automated molecular weight assignment of electrospray ionization mass spectra. *Rapid Commun Mass Spectrom* 15(24):2446–2455
 176. Hobby K, Gallagher RT, Caldwell P, Wilson ID (2009) A new approach to aid the characterisation and identification of metabolites of a model drug; partial isotope enrichment combined with novel formula elucidation software. *Rapid Commun Mass Spectrom* 23(2):219–227
 177. Alon T, Amirav A (2006) Isotope abundance analysis methods and software for improved sample identification with supersonic gas chromatography/mass spectrometry. *Rapid Commun Mass Spectrom* 20(17):2579–2588
 178. Rogers S, Scheltema RA, Girolami M, Breitling R (2009) Probabilistic assignment of formulas to mass peaks in metabolomics experiments. *Bioinformatics* 25(4):512–518
 179. Werner E, Heilier J-F, Ducruix C, Ezan E, Junot C, Tabet J-C (2008) Mass spectrometry for the identification of the discriminating signals from metabolomics: current status and future trends. *J Chromatogr B* 871(2):143–163

180. Reemtsma T (2009) Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry. Status and needs. *J Chromatogr A* 1216(18):3687–3701
181. Koch BP, Ludwiczowski KU, Kattner G, Dittmar T, Witt M (2008) Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Mar Chem* 111(3–4):233–241
182. Bocker S, Letzel MC, Liptak Z, Pervukhin A (2009) SIRIUS: decomposing isotope patterns for metabolite identification. *Bioinformatics* 25(2):218
183. Bocker S, Rasche F (2008) Towards de novo identification of metabolites by analyzing tandem mass spectra. *Bioinformatics* 24(16):i49
184. Erve JCL, Gu M, Wang Y, DeMaio W, Talaat RE (2009) Spectral accuracy of molecular ions in an LTQ/Orbitrap mass spectrometer and implications for elemental composition determination. *J Am Soc Mass Spectrom* 20(11):2058–2069
185. Hegeman AD, Schulte CF, Cui Q, Lewis IA, Huttlin EL, Eghbalian H, Harms AC, Ulrich EL, Markley JL, Sussman MR (2007) Stable isotope assisted assignment of elemental compositions for metabolomics. *Anal Chem* 79(18):6912–6921
186. Hill DW, Kertesz TM, Fontaine D, Friedman R, Grant DF (2008) Mass spectral metabonomics beyond elemental formula: chemical database querying by matching experimental with computational fragmentation spectra. *Anal Chem* 80(14):5574–5582
187. Krebs I, Yates S (2008) Fully unsupervised automatic assignment and annotation of sum formulae for product ion peaks, neutral losses in MS and Product Ion Spectra. In: ASMS 2008, MPG152. Bruker Daltonics, Billerica
188. Rockwood AL, Van Orden SL, Smith RD (1995) Rapid calculation of isotope distributions. *Anal Chem* 67:2699–2704
189. Konishi Y, Kiyota T, Draghici C, Gao J-M, Yeboah F, Acoca S, Jarussophon S, Purisima E (2006) Molecular formula analysis by an MS/MS/MS technique to expedite dereplication of natural products. *Anal Chem* 79(3):1187–1197
190. Jarussophon S, Acoca S, Gao JM, Deprez C, Kiyota T, Draghici C, Purisima E, Konishi Y (2009) Automated molecular formula determination by tandem mass spectrometry (MS/MS). *Analyst* 134(4):690–700
191. Plumb RS, Johnson KA, Rainville P, Smith BW, Wilson ID, Castro-Perez JM, Nicholson JK (2006) UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun Mass Spectrom* 20(13):1989–1994
192. Harkewicz R (2002) ESI-FTICR mass spectrometry employing data-dependent external ion selection and accumulation. *J Am Soc Mass Spectrom* 13(2):144–154
193. Wind M, Grunwald H, Gebhardt K, Illig K, Spickermann J, Nuoffer C, Roussel P, Klauer D, Fullhardt P, Schmitt-Hoffmann A (2009) Investigation of the species-dependent in vitro metabolism of BAL30630 by stable isotope labeling and isotope exchange experiments analyzed by capillary liquid chromatography coupled to mass spectrometry. *J Chromatogr A* 1216(18):3946–3953
194. Yao MM, Humphreys L, Zhu WG (2008) Rapid screening and characterization of drug metabolites using a multiple ion monitoring-dependent MS/MS acquisition method on a hybrid triple quadrupole-linear ion trap mass spectrometer. *J Mass Spectrom* 43(10):1364–1375
195. Yao M, Ma L, Duchoslav E, Zhu M (2009) Rapid screening and characterization of drug metabolites using multiple ion monitoring dependent product ion scan and postacquisition data mining on a hybrid triple quadrupole-linear ion trap mass spectrometer. *Rapid Commun Mass Spectrom* 23(11):1683–1693
196. Sang S, Lee MJ, Yang I, Buckley B, Yang CS (2008) Human urinary metabolite profile of tea polyphenols analyzed by liquid chromatography/electrospray ionization tandem mass spectrometry with data-dependent acquisition. *Rapid Commun Mass Spectrom* 22(10):1567–1578
197. Lim HK, Chen J, Cook K, Sensenhauser C, Silva J, Evans DC (2008) A generic method to detect electrophilic intermediates using isotopic pattern triggered data-dependent high-resolution accurate mass spectrometry. *Rapid Commun Mass Spectrom* 22(8):1295–1311
198. Schwartz JC, Wade AP, Enke CG, Cooks RG (1990) Systematic delineation of scan modes in multidimensional mass spectrometry. *Anal Chem* 62(17):1809–1818
199. Rochfort SJ, Trenerry VC, Imsic M, Panozzo J, Jones R (2008) Class targeted metabolomics: ESI ion trap screening methods for glucosinolates based on MS_n fragmentation. *Phytochemistry* 69(8):1671–1679
200. Lloyd KG, O’Keefe DP (2004) Re-discovering surface mass spectrometry: chemical mapping from micro to macro. *Appl Surf Sci* 231–232:207–216
201. Reig N, Calaf RE, Messegueur A, Morató A, Escabros J, Gelpí E, Abian J (2007) LC-MS ion maps for the characterization of aniline derivatives of fatty acids and triglycerides in laboratory-denatured rapeseed oil. *J Mass Spectrom* 42(4):527–541
202. Pont F, Luciani B, Belmont C, Fournie JJ (2001) Characterization of phosphoantigens by high-performance anion-exchange chromatography-electrospray ionization ion trap mass spectrometry and nano-electrospray ionization ion trap mass spectrometry. *Anal Chem* 73(15):3562–3569
203. Roussis SG (2001) Automated tandem mass spectrometry by orthogonal acceleration TOF data acquisition and simultaneous magnet scanning for the characterization of petroleum mixtures. *Anal Chem* 73(15):3611–3623
204. Germanus A (2002) Method for management of daughter ion spectra over several generations. In: US Patent 6,362,473
205. Sheldon MT, Mistrik R, Croley TR (2009) Determination of ion structures in structurally related compounds using precursor ion fingerprinting. *J Am Soc Mass Spectrom* 20(3):370–376
206. Mistrik R (2007) Determination of molecular structures using tandem mass spectrometry. In: U.S. 7,197,402
207. Milgram KE, Barrett T, Evans AM (2008) Systems, methods, and computer-readable medium for determining composition of chemical constituents in a complex mixture. In: U.S. 12/327,758
208. Li Q, Yan G, Ge T (2007) A fragmentation study of podophyllotoxin and its 4'-demethyl-4b-substituted derivatives by electrospray ionization ion-trap time-of-flight tandem mass spectrometry. *Rapid Commun Mass Spectrom* 21(17):2843–2852
209. Sturt HF, Summons RE, Smith K, Elvert M, Hinrichs KU (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun Mass Spectrom* 18(6):617–628
210. Strapoc D, Picardal FW, Turich C, Schaperdoth I, Macalady JL, Lipp JS, Lin YS, Ertefai TF, Schubotz F, Hinrichs KU (2008) Methane-producing microbial community in a coal bed of the Illinois Basin. *Appl Environ Microbiol* 74(8):2424
211. Hsu FF, Turk J (2010) Electrospray ionization multiple-stage linear ion-trap mass spectrometry for structural elucidation of triacylglycerols: assignment of fatty acyl groups on the glycerol backbone and location of double bonds. *J Am Soc Mass Spectrom* 21(4):657–669
212. Hsu FF, Turk J, Rhoades ER, Russell DG, Shi Y, Groisman EA (2005) Structural characterization of cardiolipin by tandem quadrupole and multiple-stage quadrupole ion-trap mass spectrometry with electrospray ionization. *J Am Soc Mass Spectrom* 16(4):491–504
213. Serb A, Schiopu C, Flangea C, Sisu E, Zamfir AD (2009) Top-down glycolipidomics: fragmentation analysis of ganglioside oligosaccharide core and ceramido moiety by chip-nano-electrospray

- collision-induced dissociation MS2-MS6. *J Mass Spectrom* 44 (10):1434–1442
214. Saad OM, Leary JA (2005) Heparin sequencing using enzymatic digestion and ESI-MSn with HOST: a heparin/HS oligosaccharide sequencing tool. *Anal Chem* 77(18):5902–5911
215. Ashline DJ, Lapadula AJ, Liu Y-H, Lin M, Grace M, Pramanik B, Reinhold VN (2007) Carbohydrate structural isomers analyzed by sequential mass spectrometry. *Anal Chem* 79(10):3830–3842
216. McLafferty FW, Hertel RH, Villwock RD (1974) Computer identification of mass spectra: VI. Probability based matching of mass spectra: rapid identification of specific compounds in mixtures. *Org Mass Spectrom* 9(7):690–702
217. Stein SE, Scott DR (1994) Optimization and testing of mass spectral library search algorithms for compound identification. *J Am Soc Mass Spectrom* 5(9):859–866
218. David Sparkman O (1996) Evaluating electron ionization mass spectral library search results. *J Am Soc Mass Spectrom* 7 (4):313–318
219. Horai H, Arita M, Nishioka T (2008) Comparison of ESI-MS spectra in MassBank database. 2008 International Conference on BioMedical Engineering and Informatics 853–857
220. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K (2010) MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* 45(7):703–714
221. Akiyama K, Chikayama E, Yuasa H, Shimada Y, Tohge T, Shinozaki K, Hirai MY, Sakurai T, Kikuchi J, Saito K (2008) PRIME: a Web site that assembles tools for metabolomics and transcriptomics. *In Silico Biol* 8(3):339–345
222. Sawada Y, Akiyama K, Sakata A, Kuwahara A, Otsuki H, Sakurai T, Saito K, Hirai MY (2009) Widely targeted metabolomics based on large-scale MS/MS data for elucidating metabolite accumulation patterns in plants. *Plant Cell Physiol* 50(1):37
223. Matsuda F, Hirai MY, Sasaki E, Akiyama K, Yonekura-Sakakibara K, Provart NJ, Sakurai T, Shimada Y, Saito K (2010) AtMetExpress development: a phytochemical atlas of Arabidopsis development. *Plant Physiol* 152(2):566–578
224. Dresen S, Gergov M, Politi L, Halter C, Weinmann W (2009) ESI-MS/MS library of 1, 253 compounds for application in forensic and clinical toxicology. *Analytical and Bioanalytical Chemistry* 395:2521–2526
225. Schreiber A, Efer J, Engewald W (2000) Application of spectral libraries for high-performance liquid chromatography-atmospheric pressure ionisation mass spectrometry to the analysis of pesticide and explosive residues in environmental samples. *J Chromatogr A* 869(1–2):411–425
226. Mueller CA, Weinmann W, Dresen S, Schreiber A, Gergov M (2005) Development of a multi-target screening analysis for 301 drugs using a QTrap liquid chromatography/tandem mass spectrometry system and automated library searching. *Rapid Commun Mass Spectrom* 19(10):1332–1338
227. Fredenhagen A, Derrien C, Gassmann E (2005) An MS/MS library on an ion-trap instrument for efficient dereplication of natural products. Different fragmentation patterns for $[M+H]^+$ and $[M+Na]^+$ Ions†. *J Nat Prod* 68(3):385–391
228. Anthony WT, Kenneth SW, Anneke TL, John H (2004) Reproducible product-ion tandem mass spectra on various liquid chromatography/mass spectrometry instruments for the development of spectral libraries. *Rapid Commun Mass Spectrom* 18 (13):1447–1454
229. Milman BL (2005) Towards a full reference library of MSn spectra. Testing of a library containing 3126 MS2 spectra of 1743 compounds. *Rapid Commun Mass Spectrom* 19(19):2833–2839
230. Hopley C, Bristow T, Lubben A, Simpson A, Bull E, Klagkou K, Herniman J, Langley J (2008) Towards a universal product ion mass spectral library—reproducibility of product ion spectra across eleven different mass spectrometers. *Rapid Commun Mass Spectrom* 22(12):1779–1786
231. Palit M, Mallard G (2009) Fragmentation energy index for universalization of fragmentation energy in ion trap mass spectrometers for the analysis of chemical weapon convention related chemicals by atmospheric pressure ionization–tandem mass spectrometry analysis. *Anal Chem* 81(7):2477–2485
232. Oberacher H, Pavlic M, Libiseller K, Schubert B, Sulyok M, Schuhmacher R, Csaszar E, Köfeler HC (2009) On the inter-instrument and inter-laboratory transferability of a tandem mass spectral reference library: 1. Results of an Austrian multicenter study. *J Mass Spectrom* 44:485–493
233. Pavlic M, Libiseller K, Oberacher H (2006) Combined use of ESI-QqTOF-MS and ESI-QqTOF-MS/MS with mass-spectral library search for qualitative analysis of drugs. *Anal Bioanal Chem* 386(1):69–82
234. Oberacher H, Pavlic M, Libiseller K, Schubert B, Sulyok M, Schuhmacher R, Csaszar E, Köfeler HC (2009) On the inter-instrument and the inter-laboratory transferability of a tandem mass spectral reference library: 2. Optimization and characterization of the search algorithm. *J Mass Spectrom* 44(4):494–502
235. Hansen ME, Smedsgaard J (2004) A new matching algorithm for high resolution mass spectra. *J Am Soc Mass Spectrom* 15 (8):1173–1180
236. Kandasamy K, Pandey A, Molina H (2009) Evaluation of several MS/MS search algorithms for analysis of spectra derived from electron transfer dissociation experiments. *Anal Chem* 81 (17):7170–7180
237. Kapp EA, Schütz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS et al (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. *Proteomics* 5(13):3475–3490
238. Wan KX, Vidavsky I, Gross ML (2002) Comparing similar spectra: from similarity index to spectral contrast angle. *J Am Soc Mass Spectrom* 13(1):85–88
239. McLafferty FW, Zhang M-Y, Stauffer DB, Loh SY (1998) Comparison of algorithms and databases for matching unknown mass spectra. *J Am Soc Mass Spectrom* 9(1):92–95
240. Gergov M, Weinmann W, Meriluoto J, Uusitalo J, Ojanpera I (2004) Comparison of product ion spectra obtained by liquid chromatography/triple-quadrupole mass spectrometry for library search. *Rapid Commun Mass Spectrom* 18(10):1039–1046
241. Mylonas R, Mauron Y, Masselot A, Binz PA, Budin N, Fathi M, Viette V, Hochstrasser DF, Lisacek F (2009) X-Rank: a robust algorithm for small molecule identification using tandem mass spectrometry. *Anal Chem* 81(18):7604–7610
242. Samudrala R, Levitt M (2000) Decoys ‘R’Us: a database of incorrect conformations to improve protein structure prediction. *PRS* 9(07):1399–1401
243. Choi H, Nesvizhskii AI (2007) False discovery rates and related statistical concepts in mass spectrometry-based proteomics. *J Proteome Res* 7(01):47–50
244. Kim S, Gupta N, Pevzner PA (2008) Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. *J Proteome Res* 7(8):3354–3363
245. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Meth* 4(3):207–214
246. Anabel SF, Michael K, Stefan WT, Gerold K (2002) Identification of anhydroecgonine methyl ester N-oxide, a new metabolite of anhydroecgonine methyl ester, using electrospray mass spectrometry. *J Mass Spectrom* 37(5):525–532
247. Zhu XY, Sato T (2007) The distinction of underivatized monosaccharides using electrospray ionization ion trap mass spectrometry. *Rapid Commun Mass Spectrom* 21(2):191–198

248. Chen G, Pramanik BN, Bartner PL, Saksena AK, Gross ML (2002) Multiple-stage mass spectrometric analysis of complex oligosaccharide antibiotics (evernimomins) in a quadrupole ion trap. *J Am Soc Mass Spectrom* 13(11):1313–1321
249. Delaney J, Vouros P (2001) Liquid chromatography ion trap mass spectrometric analysis of oligosaccharides using permethylated derivatives. *Rapid Commun Mass Spectrom* 15(5):325–334
250. Kameyama A, Kikuchi N, Nakaya S, Ito H, Sato T, Shikanai T, Takahashi Y, Takahashi K, Narimatsu H (2005) A strategy for identification of oligosaccharide structures using observational multistage mass spectral library. *Anal Chem* 77(15):4719–4725
251. McLafferty FW, Turecek F (1993) Interpretation of mass spectra. University Science Books, Mill Valley
252. Watson JT, Sparkman OD (2007) Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation. Wiley, Chichester
253. Smith RM (2004) Understanding mass spectra: a basic approach. Wiley-Interscience, Chichester
254. Karsten L, Hans-Martin S, Johan KT, Karl JJ, Manfred E, Alfred P, Herbert T, Arnd I (2007) Even-electron ions: a systematic study of the neutral species lost in the dissociation of quasi-molecular ions. *J Mass Spectrom* 42(8):1024–1044
255. Cai Y, Mo Z, Rannulu NS, Guan B, Kannupal S, Gibba BC, Cole RB (2010) Characterization of an exception to the ‘even-electron rule’ upon low-energy collision induced decomposition in negative ion electrospray tandem mass spectrometry. *J Mass Spectrom* 45(3):235–240
256. Thurman EM, Ferrer I, Pozo OJ, Sancho JV, Hernandez F (2007) The even-electron rule in electrospray mass spectra of pesticides. *Rapid Commun Mass Spectrom* 21(23):3855–3868
257. Badertscher M, Bischofberger K, Munk ME, Pretsch E (2001) A novel formalism to characterize the degree of unsaturation of organic molecules. *J Chem Inf Comput Sci* 41(4):889–893
258. Baigorri R, Zamarreño AM, Fuentes M, Gonzalez-Gaitano G, Garcia-Mina JM, Almedros G, Gonzalez-Vila FJ (2008) Multivariate statistical analysis of mass spectra as a tool for the classification of the main humic substances according to their structural and conformational features. *J Agric Food Chem* 56(14):5480–5487
259. Meija J (2006) Mathematical tools in analytical mass spectrometry. *Anal Bioanal Chem* 385(3):486–499
260. Hertkorn N, Ruecker C, Meringer M, Gugisch R, Frommberger M, Perdue EM, Witt M, Schmitt-Kopplin P (2007) High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal Bioanal Chem* 389(5):1311–1327
261. Cooks RG (1969) Bond formation upon electron-impact. *Org Mass Spectrom* 2(5):481
262. Borth S, Hänsel W, Rösner P, Junge T (2000) Synthesis of 2,3- and 3,4-methylenedioxyphenylalkylamines and their regioisomeric differentiation by mass spectral analysis using GC-MS-MS. *Forensic Sci Int* 114(3):139–153
263. Watson JT, Sparkman OD (2008) Chemical ionization. In: Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation, 4th edition. Wiley, New York, pp 449–484
264. Richard AJ (2002) Ion chemistry and fragmentation. In: Rossi DT, Sinz MW (eds) Mass spectrometry in drug discovery. CRC, Boca Raton, p 85
265. Hill AW, Mortishire-Smith RJ (2005) Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Commun Mass Spectrom* 19(21):3111–3118
266. Heinonen M, Rantanen A, Mielikäinen T, Kokkonen J, Kiuru J, Ketola RA, Rousu J (2008) FiD: a software for ab initio structural identification of product ions from tandem mass spectrometric data. *Rapid Commun Mass Spectrom* 22(19):3043–3052
267. Heinonen M, Rantanen A, Mielikäinen T, Pitkänen E, Kokkonen J, Rousu J (2006) Ab Initio prediction of molecular fragments from tandem mass spectrometry data. Proceedings of German Conference on Bioinformatics (GCB 2006), Lecture Notes in Informatics 83:40–53
268. Goodley P (2007) Maximizing MS/MS Fragmentation in the Ion Trap Using CID Voltage Ramping. In: Agilent Report 5988-0704EN
269. Levsen K, Schiebel HM, Terlouw JK, Jobst KJ, Elend M, Preiß A, Thiele H, Ingendoh A (2007) Even-electron ions: a systematic study of the neutral species lost in the dissociation of quasi-molecular ions. *J Mass Spectrom* 42(8):1024–1044
270. Holcapek M, Jirasko R, Lisa M (2010) Basic rules for the interpretation of atmospheric pressure ionization mass spectra of small molecules. *J Chromatogr* 1217(25):3908–3921
271. Brodbelt JS (2000) Probing molecular recognition by mass spectrometry. *Int J Mass Spectrom* 200(1):57–69
272. Yao ZP, Wan TSM, Kwong KP, Che CT (2000) Chiral analysis by electrospray ionization mass spectrometry/mass spectrometry. 1. Chiral recognition of 19 common amino acids. *Anal Chem* 72(21):5383–5393
273. Splitter JS, Turecek F, Luftmann H (1994) Applications of mass spectrometry to organic stereochemistry. VCH, New York
274. Xianmei Cai AV, Dass C (2003) Conformational analysis of proteins and peptides. *Curr Org Chem* 7(18):1841–1854
275. Kaltashov IA, Eyles SJ (2002) Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom Rev* 21(1):37–71
276. Knapman TW, Berryman JT, Campuzano I, Harris SA, Ashcroft AE (2009) Considerations in experimental and theoretical collision cross-section measurements of small molecules using travelling wave ion mobility spectrometry-mass spectrometry. *Int J Mass Spectrom* (in press)
277. Valentine SJ, Clemmer DE (2009) Treatise on the measurement of molecular masses with ion mobility spectrometry. *Anal Chem* 81(14):5876–5880
278. Ramanathan R, Comezoglu SN, Humphreys WG (2009) Metabolite identification strategies and procedures. CRC, Boca Raton, 1420092200
279. Lynch T, Price A (2007) The effect of cytochrome P 450 metabolism on drug response, interactions, and adverse effects. *Am Fam Physician* 76(3):391–396
280. Levsen K, Schiebel H-M, Behnke B, Dötzer R, Dreher W, Elend M, Thiele H (2005) Structure elucidation of phase II metabolites by tandem mass spectrometry: an overview. *J Chromatogr A* 1067(1–2):55–72
281. Anari MR, Baillie TA (2005) Bridging cheminformatic metabolite prediction and tandem mass spectrometry. *Drug Discov Today* 10(10):711–717
282. Pelkonen O, Tolonen A, Korjamo T, Turpeinen M, Raunio H (2009) From known knowns to known unknowns: predicting in vivo drug metabolites. *Bioanalysis* 1(2):393–414
283. Ekins S, Andreyev S, Ryabov A, Kirillov E, Rakhmatulin EA, Bugrim A, Nikolskaya T (2005) Computational prediction of human drug metabolism. *Expert Opin Drug Metab Toxicol* 1(2):303–324
284. Marchant CA, Briggs KA, Long A (2008) In silico tools for sharing data and knowledge on toxicity and metabolism: Derek for windows, meteor, and vitic. *Toxicol Mech Methods* 18(2–3):177–187
285. Wen B, Fitch WL (2009) Analytical strategies for the screening and evaluation of chemically reactive drug metabolites. *Expert Opin Drug Metab Toxicol* 5(1):39–55
286. Anari MR, Sanchez RI, Bakhtiar R, Franklin RB, Baillie TA (2003) Integration of knowledge-based metabolic predictions with liquid chromatography data-dependent tandem mass spec-

- trometry for drug metabolism studies: application to studies on the biotransformation of indinavir. *Anal Chem* 76(3):823–832
287. Mortishire-Smith RJ, O'Connor D, Castro-Perez JM, Kirby J (2005) Accelerated throughput metabolic route screening in early drug discovery using high-resolution liquid chromatography/quadrupole time-of-flight mass spectrometry and automated data analysis. *Rapid Commun Mass Spectrom* 19(18):2659–2670
288. Zhu M, Ma L, Zhang D, Ray K, Zhao W, Humphreys WG, Skiles G, Sanders M, Zhang H (2006) Detection and characterization of metabolites in biological matrices using mass defect filtering of liquid chromatography/high resolution mass spectrometry data. *Drug Metab Dispos* 34(10):1722
289. Mortishire-Smith RJ, Castro-Perez JM, Yu K, Shockcor JP, Goshawk J, Hartshorn MJ, Hill A (2009) Generic dealkylation: a tool for increasing the hit-rate of metabolite rationalization, and automatic customization of mass defect filters. *Rapid Commun Mass Spectrom* 23(7):939–948
290. Zhang H, Zhu M, Ray KL, Ma L, Zhang D (2008) Mass defect profiles of biological matrices and the general applicability of mass defect filtering for metabolite detection. *Rapid Commun Mass Spectrom* 22(13):2082–2088
291. Bateman KP, Castro-Perez J, Wrona M, Shockcor JP, Yu K, Oballa R, Nicoll-Griffith DA (2007) MSE with mass defect filtering for in vitro and in vivo metabolite identification. *Rapid Commun Mass Spectrom* 21(9):1485–1496
292. Stranz DD, Miao S, Campbell S, Maydwell G, Ekins S (2008) Combined computational metabolite prediction and automated structure-based analysis of mass spectrometric data. *Toxicol Mech Methods* 18(2–3):243–250
293. Leclercq L, Mortishire-Smith RJ, Huisman M, Cuyckens F, Hartshorn MJ, Hill A (2009) IsoScore: automated localization of biotransformations by mass spectrometry using product ion scoring of virtual regioisomers. *Rapid Commun Mass Spectrom* 23(1):39–50
294. Hopfgartner G, Husser C, Zell M (2003) Rapid screening and characterization of drug metabolites using a new quadrupole-linear ion trap mass spectrometer. *J Mass Spectrom* 38(2):138–150
295. Duan Y, Ma X, Zou W, Wang C, Behbahan IS, Ahuja TP, Tolstikov V, Zern MA (2010) Differentiation and characterization of metabolically functioning hepatocytes from human embryonic stem cells. *Stem Cells* 28(4):674–686
296. Bell SC, Gayton-Ely M, Nida CM (2009) Bioassays for bomb-makers: proof of concept. *Anal Bioanal Chem* 395(2):401–409
297. Gorman GS, Coward L, Kerstner-Wood C, Freeman L, Hebert CD, Kapetanovic IM (2009) In-vitro and in-vivo metabolic studies of the candidate chemopreventative pentamethylchromanol using liquid chromatography/tandem mass spectrometry. *J Pharm Pharmacol* 61(10):1309–1318
298. Chen Y, Liu HF, Liu L, Nguyen K, Jones EB, Fretland AJ (2010) The in vitro metabolism of bupropion revisited: concentration dependent involvement of cytochrome P450 2C19. *Xenobiotica* 40(8):536–546
299. Wind M, Gebhardt K, Grunwald H, Spickermann J, Donzelli M, Kellenberger L, Muller M, Fullhardt P, Schmitt-Hoffmann A, Schleimer M (2007) Elucidation of the in vitro metabolic profile of stable isotope labeled BAL19403 by accurate mass capillary liquid chromatography/quadrupole time-of-flight mass spectrometry and isotope exchange. *Rapid Commun Mass Spectrom* 21(7):1093–1099
300. Liu DQ, Hop C (2005) Strategies for characterization of drug metabolites using liquid chromatography–tandem mass spectrometry in conjunction with chemical derivatization and on-line H/D exchange approaches. *J Pharm Biomed Anal* 37(1):1–18
301. Kiffe M, Graf D, Trunzer M (2007) Two-dimensional liquid chromatography/mass spectrometry set-up for structural elucidation of metabolites in complex biological matrices. *Rapid Commun Mass Spectrom* 21(6):961–970
302. Nassar A-EF, Talaat RE (2004) Strategies for dealing with metabolite elucidation in drug discovery and development. *Drug Discov Today* 9(7):317–327
303. Ong SE, Blagojev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1(5):376
304. Ramos-Fernandez A, Lopez-Ferrer D, Vazquez J (2007) Improved method for differential expression proteomics using trypsin-catalyzed ¹⁸O labeling with a correction for labeling efficiency. *Mol Cell Proteomics* 6(7):1274–1286
305. Li X-J, Zhang H, Ranish JA, Aebersold R (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal Chem* 75(23):6648–6657
306. Beynon RJ, Pratt JM (2005) Metabolic labeling of proteins for proteomics. *Mol Cell Proteomics* 4(7):857
307. Wang YK, Ma Z, Quinn DF, Fu EW (2001) Inverse ¹⁸O labeling mass spectrometry for the rapid identification of marker/target proteins. *Anal Chem* 73(15):3742–3750
308. Birkemeyer C, Luedemann A, Wagner C, Erban A, Kopka J (2005) Metabolome analysis: the potential of in vivo labeling with stable isotopes for metabolite profiling. *Trends Biotechnol* 23(1):28–33
309. Giavalisco P, Hummel J, Lisec J, Inostroza AC, Catchpole G, Willmitzer L (2008) High-resolution direct infusion-based mass spectrometry in combination with whole ¹³C metabolome isotope labeling allows unambiguous assignment of chemical sum formulas. *Anal Chem* 80(24):9417–9425
310. Garcia DE, Baidoo EE, Benke PI, Pingitore F, Tang YJ, Villa S, Keasling JD (2008) Separation and mass spectrometry in microbial metabolomics. *Curr Opin Microbiol* 11(3):233–239
311. Lane AN, Fan TWM, Higashi RM (2008) Isotopomer-based metabolomic analysis by NMR and mass spectrometry. *Methods Cell Biol* 84:541–588
312. Emmerling M, Dauner M, Ponti A, Fiaux J, Hochuli M, Szyperski T, Wuthrich K, Bailey JE, Sauer U (2002) Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol* 184(1):152–164
313. Christensen B, Nielsen J (1999) Isotopomer analysis using GC-MS. *Metab Eng* 1(4):282–290
314. Wittmann C, Heinzle E (1999) Mass spectrometry for metabolic flux analysis. *Biotechnol Bioeng* 62(6):739–750
315. Feldberg L, Venger I, Malitsky S, Rogachev I, Aharoni A (2009) Dual Labeling of Metabolites for Metabolome Analysis (DLEMMA): a new approach for the identification and relative quantification of metabolites by means of dual isotope labeling and liquid chromatography–mass spectrometry. *Anal Chem* 81(22):9257–9266
316. Liu Y, Wang M, Xue M, Li Y, Li X, Ruan J, Liu K (2008) Structural elucidation of in vivo metabolites of penethylidene in rats by the method of liquid chromatography–mass spectrometry, gas chromatography–mass spectrometry and isotope ion cluster. *J Chromatogr B Anal Technol Biomed Life Sci* 873(1):41–50
317. Kuhn F, Oehme M, Romero F, Abou-Mansour E, Tabacchi R (2003) Differentiation of isomeric flavone/isoflavone aglycones by MS² ion trap mass spectrometry and a double neutral loss of CO. *Rapid Commun Mass Spectrom* 17(17):1941–1949
318. Wickens JR, Sleeman R, Keely BJ (2006) Atmospheric pressure ionisation mass spectrometric fragmentation pathways of noscapine and papaverine revealed by multistage mass spectrometry and in-source deuterium labelling. *Rapid Commun Mass Spectrom* 20(3):473–480

319. Taylor EW, Jia W, Mark B, Dollinger GD (2002) Accelerating the drug optimization process: identification, structure elucidation, and quantification of in vivo metabolites using stable isotopes with LC/MSn and the chemiluminescent nitrogen detector. *Anal Chem* 74(13):3232–3238
320. Ohashi N, Furuuchi S, Yoshikawa M (1998) Usefulness of the hydrogen–deuterium exchange method in the study of drug metabolism using liquid chromatography–tandem mass spectrometry. *J Pharm Biomed Anal* 18(3):325–334
321. Lehane M, Sáez MJF, Magdalena AB, Cañas IR, Sierra MD, Hamilton B, Furey A, James KJ (2004) Liquid chromatography—multiple tandem mass spectrometry for the determination of ten azaspiracids, including hydroxyl analogues in shellfish. *J Chromatogr A* 1024(1–2):63–70
322. Liu DQ, Wu L, Sun M, MacGregor PA (2007) On-line H/D exchange LC–MS strategy for structural elucidation of pharmaceutical impurities. *J Pharm Biomed Anal* 44(2):320–329
323. Keller BO, Sui J, Young AB, Whittall RM (2008) Interferences and contaminants encountered in modern mass spectrometry. *Anal Chim Acta* 627(1):71–81
324. Little JL (1999) Artifacts in trimethylsilyl derivatization reactions and ways to avoid them. *J Chromatogr A* 844(1–2):1–22
325. Xu F, Zou L, Ong CN (2009) Multiorigin of chromatographic peaks in derivatized GC/MS metabolomics: a confounder that influences metabolic pathway interpretation. *J Proteome Res* 8(12):5657–5665
326. Gfrerer M, Lankmayr E (2005) DDT degradation during enhanced solid–liquid extractions A consideration. *J Chromatogr A* 1072(1):117–125
327. Fiehn O, Wohlge-muth G, Scholz M, Kind T, Lee DY, Lu Y, Moon S, Nikolau B (2008) Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* 53(4):691–704
328. Clark HA, Jurs PC (1981) Simulation of mass spectral intensities by regression analysis of calculated structural characteristics. *Anal Chim Acta* 132:75–88
329. Zhou C, Bowler LD, Feng J (2008) A machine learning approach to explore the spectra intensity pattern of peptides using tandem mass spectrometry data. *BMC Bioinform* 9(1):325
330. Timm W, Scherbart A, Böcker S, Kohlbacher O, Nattkemper TW (2008) Peak intensity prediction in MALDI-TOF mass spectrometry: a machine learning study to support quantitative proteomics. *BMC Bioinform* 9(1):443
331. Hochstrasser DF, Appel RD (2002) Peptide mass fingerprinting peak intensity prediction: extracting knowledge from spectra. *Proteomics* 2:1374–1391
332. Zhang H, Singh S, Reinhold VN (2005) Congruent strategies for carbohydrate sequencing. 2. FragLib: an MSn spectral library. *Anal Chem* 77(19):6263–6270
333. Kameyama A, Nakaya S, Ito H, Kikuchi N, Angata T, Nakamura M, Ishida HK, Narimatsu H (2006) Strategy for simulation of CID spectra of N-linked oligosaccharides toward glycomics. *J Proteome Res* 5(4):808–814
334. Chen T, Kao MY, Tepel M, Rush J, Church GM (2001) A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry. *J Comput Biol* 8(3):325–337
335. Katerina K, Frank P, Mark H, Andy O, Alistair F, Langley GJ (2003) Approaches towards the automated interpretation and prediction of electrospray tandem mass spectra of non-peptidic combinatorial compounds. *Rapid Commun Mass Spectrom* 17(11):1163–1168
336. Jalali-Heravi M, Fatemi MH (2000) Simulation of mass spectra of noncyclic alkanes and alkenes using artificial neural network. *Anal Chim Acta* 415(1–2):95–103
337. Chen H, Fan B, Xia H, Petitjean M, Yuan S, Panaye A, Doucet JP (2003) MASSIS: a mass spectrum simulation system 1. Principle and method. *Eur J Mass Spectrom* (Chichester, England) 9(3):175
338. Chen H, Fan B, Petitjean M, Panaye A, Doucet JP, Li F, Xia H, Yuan S (2003) MASSIS: a mass spectrum simulation system. 2: procedures and performance. *Eur J Mass Spectrom* (Chichester, England) 9(5):445
339. Fan B, Chen H, Petitjean M, Panaye A, Doucet JP, Xia H, Yuan S (2005) New strategy of mass spectrum simulation based on reduced and concentrated knowledge databases. *Spectrosc Lett* 38(2):145–170
340. Bandu ML, Watkins KR, Bretthauer ML, Moore CA, Desaire H (2004) Prediction of MS/MS Data. 1. A focus on pharmaceuticals containing carboxylic acids. *Anal Chem* 76(6):1746–1753
341. Wolf S, Schmidt S, Mueller-Hannemann M, Neumann S (2010) In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* 11:148
342. Schymanski EL, Meringer M, Brack W (2009) Matching structures to mass spectra using fragmentation patterns: are the results as good as they look? *Anal Chem* 81(9):3608–3617
343. Pelander A, Tyrkkö E, Ojanperä I (2009) In silico methods for predicting metabolism and mass fragmentation applied to quetiapine in liquid chromatography/time-of-flight mass spectrometry urine drug screening. *Rapid Commun Mass Spectrom* 23(4):506–514
344. Lukka PB, Kestell P, Paxton JW, Baguley BC (2008) Development and validation of a liquid chromatography–mass spectrometry (LC-MS) assay for the determination of the anti-cancer agent N-[2-(dimethylamino) ethyl]-2, 6-dimethyl-1-oxo-1, 2-dihydrobenzo [b]-1, 6-naphthyridine-4-carboxamide (SN 28049). *J Chromatogr B* 875(2):368–372
345. Eash MA, Gohlke RS (1962) Mass spectrometric analysis. A small computer program for the analysis of mass spectra. *Anal Chem* 34(6):713
346. Biemann K, McMurray W (1966) Computer-aided interpretation of high resolution mass spectra (high resolution mass spectrometry in computer-aided interpretation). *Tetrahedron Lett* 88(11):2597–2598
347. Lederberg J (1987) How DENDRAL was conceived and born. *ACM*, pp. 5–19
348. Scheurell M, Franke S, Shah RM, Hühnerfuss H (2009) Occurrence of diclofenac and its metabolites in surface water and effluent samples from Karachi, Pakistan. *Chemosphere* 77(6):870–876
349. Stein SE (1995) Chemical substructure identification by mass spectral library searching. *J Am Soc Mass Spectrom* 6(8):644–655
350. Kerber A, Meringer M, RÜcker C (2006) CASE via MS: ranking structure candidates by mass spectra. *Croat Chem Acta* 79(3):449–464
351. Schymanski EL, Meinert C, Meringer M, Brack W (2008) The use of MS classifiers and structure generation to assist in the identification of unknowns in effect-directed analysis. *Anal Chim Acta* 615(2):136–147
352. Varmuza K, Werther W (1996) Mass spectral classifiers for supporting systematic structure elucidation. *J Chem Inf Comput Sci* 36(2):323–333
353. Zhang L, Liang Y, Chen A (2009) Selection of neutral losses and characteristic ions for mass spectral classifier. *Analyst* 134(8):1717–1724
354. Xiong Q, Zhang Y, Li M (2007) Computer-assisted prediction of pesticide substructure using mass spectra. *Anal Chim Acta* 593(2):199–206
355. Hummel J, Strehmel N, Selbig J, Walther D, Kopka J (2010) Decision tree supported substructure prediction of metabolites from GC-MS profiles. *Metabolomics* 6(2):322–333
356. Mitch A, Phelps DMR, Jeffrey SJ, Katherine LF, Katie AA, John CB, Thomas SL, Michael RG, James TD (2008) Development and validation of a sensitive liquid chromatography/mass spectrometry method for quantitation of flavopiridol in plasma enables accurate estimation of pharmacokinetic parameters with a clinically active dosing schedule. *J Chromatogr B Anal Technol Biomed Life Sci* 868(1–2):110

357. Diana J, Govaerts C, Hoogmartens J, Van Schepdael A, Adams E (2006) Characterization of impurities in dirithromycin by liquid chromatography/ion trap mass spectrometry. *J Chromatogr A* 1125(1):52–66
358. Zou W, Tolstikov VV (2009) Pattern recognition and pathway analysis with genetic algorithms in mass spectrometry based metabolomics. *Algorithms* 2(2):638
359. Vafiadi C, Topakas E, Bakx EJ, Schols HA, Christakopoulos P (2007) Structural characterisation by ESI-MS of feruloylated arabino-oligosaccharides synthesised by chemoenzymatic esterification. *Molecules* 12(7):1367–1375
360. Li L, Low MY, Ge X, Bloodworth BC, Koh HL (2009) Isolation and structural elucidation of dapoxetine as an adulterant in a health supplement used for sexual performance enhancement. *J Pharm Biomed Anal* 50(5):724–728
361. Adams PE (2003) Metabolite characterization in drug discovery utilizing robotic liquid-handling, quadruple time-of-flight mass spectrometry and in-silico prediction. *Curr Drug Metab* 4(4):259–271
362. Wu R, McMahon TB (2009) Structures, energetics, and dynamics of gas phase ions studied by FTICR and HPMS. *Mass Spectrom Rev* 28(4):546–585
363. Holmes JL, Aubry C, Mayer PM (2006) Assigning structures to ions in mass spectrometry. CRC, Boca Raton, 9780849319501
364. Luo YR (2007) Comprehensive handbook of chemical bond energies. CRC, Boca Raton, 9780849373664
365. Bouchoux G (2007) Gas-phase basicities of polyfunctional molecules. Part 1: Theory and methods. *Mass Spectrom Rev* 26(6):775–835
366. Zhang J, Ha T-K, Knochenmuss R, Zenobi R (2002) Theoretical calculation of gas-phase sodium binding energies of common MALDI matrices. *J Phys Chem A* 106(28):6610–6617
367. Chen ECM, Chen ES (2000) Negative ion mass spectra, electron affinities, gas phase acidities, bond dissociation energies, and negative ion states of cytosine and thymine. *J Phys Chem B* 104(32):7835–7844
368. Wolken JK, Turecek F (2000) Proton affinity of uracil. A computational study of protonation sites. *J Am Soc Mass Spectrom* 11(12):1065–1071
369. Bourcier S, Hoppilliard Y (2009) Use of diagnostic neutral losses for structural information on unknown aromatic metabolites: an experimental and theoretical study. *Rapid Commun Mass Spectrom* 23(1):93–103
370. Heaton AL, Armentrout PB (2008) Experimental and theoretical studies of sodium cation interactions with d-arabinose, xylose, glucose, and galactose. *J Phys Chem A* 112(41):10156–10167
371. Nourse BD, Cooks RG (1991) Proton affinity determinations using the kinetic method in an ion trap mass spectrometer. *Int J Mass Spectrom Ion Processes* 106:249–272
372. Cooks RG, Wong PSH (1998) Kinetic method of making thermochemical determinations: advances and applications. *Acc Chem Res* 31(7):379–386
373. Alexander A, Sophie H, Teresa P, Frank SP, Patricia W, Jo-Anne R (2009) Can density functional theory (DFT) be used as an aid to a deeper understanding of tandem mass spectrometric fragmentation pathways? *Rapid Commun Mass Spectrom* 23(17):2619–2627
374. Jeilani YA, Cardelino BH, Ibeanusi VM (2010) Positive chemical ionization triple-quadrupole mass spectrometry and ab initio computational studies of the multi-pathway fragmentation of phthalates. *J Mass Spectrom* 45(6):678–685
375. Schug KA, Maier NM, Lindner W (2006) Chiral recognition mass spectrometry: remarkable effects observed from the relative ion abundances of ternary diastereomeric complexes using electrospray ionization. *Chem Commun* (4):414–416
376. Glotzer SC, Kim S, Cummings PT, Deshmukh A, Head-Gordon M, Kamiadakis G, Petzold L, Sagui C, Shinozuka M (2010) International assessment of research and development in simulation-based engineering and science. www.wtec.org
377. Mitrevski BS, Kouremenos KA, Marriott PJ (2009) Accelerating analysis for metabolomics, drugs and their metabolites in biological samples using multidimensional gas chromatography. *Bioanalysis* 1(2):367–391
378. Stein SE (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *J Am Soc Mass Spectrom* 10(8):770–781
379. Lee S, Mandrake L, Bornstein B, Bue B (2009) Quantification of trace chemicals using vehicle cabin atmosphere monitor. *Aerospace Conference 2009 IEEE*. doi:10.1109/AERO.2009.4839630
380. Mandrake L, Lee S, Bornstein B, Bue B (2009) Adapting AMDIS for autonomous spectral identification of hazardous compounds for ISS monitoring. . *Aerospace Conference 2009 IEEE*. doi:10.1109/AERO.2009.4839578
381. Kind T: Structure elucidation of small molecules. http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics/Structure_Elucidation/
382. Cuadros-Inostroza A, Caldana C, Redestig H, Kusano M, Lisek J, Pena-Cortes H, Willmitzer L, Hannah MA (2009) Target-Search—a bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. *BMC Bioinform* 10(1):428
383. Renard BY, Kirchner M, Steen H, Steen JAJ, Hamprecht FA (2008) NITPICK: peak identification for mass spectrometry data. *BMC Bioinformatics* 9:355
384. Lu H, Liang Y, Dunn WB, Shen H, Kell DB (2008) Comparative evaluation of software for deconvolution of metabolomics data based on GC-TOF-MS. *TrAC, Trends Anal Chem* 27(3):215–227
385. Luedemann A, Strassburg K, Erban A, Kopka J (2008) Tag-Finder for the quantitative analysis of gas chromatography–mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics* 24(5):732
386. Fiehn O (2008) Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. *Trends Anal Chem* 27(3):261–269
387. Ebada SS, Edrada RA, Lin W, Proksch P (2008) Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nat Protocols* 3(12):1820–1831
388. Nojima S, Apperson C, Schal C (2008) A simple, convenient, and efficient preparative GC system that uses a short megabore capillary column as a trap. *J Chem Ecol* 34(3):418–428
389. Sauer PE, Eglinton TI, Hayes JM, Schimmelmann A, Sessions AL (2001) Compound-specific D/H ratios of lipid biomarkers from sediments as a proxy for environmental and climatic conditions. *Geochim Cosmochim Acta* 65(2):213–222
390. Sutton PA, Lewis CA, Rowland SJ (2005) Isolation of individual hydrocarbons from the unresolved complex hydrocarbon mixture of a biodegraded crude oil using preparative capillary gas chromatography. *Org Geochem* 36(6):963–970
391. Sinninghe Damsté JS, Baas M, Geenevasen JAJ, Kenig F (2005) Structural identification of sedimentary C21 and C22 highly branched isoprenoid alkanes. *Org Geochem* 36(4):511–517
392. Zhang A, Amalin D, Shirali S, Serrano MS, Franqui RA, Oliver JE, Klun JA, Aldrich JR, Meyerdirk DE, Lapointe SL (2004) Sex pheromone of the pink hibiscus mealybug, *Maconellicoccus hirsutus*, contains an unusual cyclobutanoid monoterpene. *Proc Natl Acad Sci USA* 101(26):9601
393. Dickens JC, Oliver JE, Hollister B, Davis JC, Klun JA (2002) Breaking a paradigm: male-produced aggregation pheromone for the Colorado potato beetle. *J Exp Biol* 205(13):1925
394. Wang Y, Lu X, Xu GW (2008) Simultaneous separation of hydrophilic and hydrophobic compounds by using an online HILIC-RPLC system with two detectors. *J Sep Sci* 31(9):1564–1572

395. Francois I, Sandra K, Sandra P (2009) Comprehensive liquid chromatography: fundamental aspects and practical considerations—A review. *Anal Chim Acta* 641(1–2):14–31
396. Shellie RA, Haddad PR (2006) Comprehensive two-dimensional liquid chromatography. *Anal Bioanal Chem* 386(3):405–415
397. Okamoto M, Yamashita K, Nakai K (2006) On-line sample enrichment system coupled to electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS). *J Pharm Biomed Anal* 41(3):707–713
398. Steiner F, Mahsunah A, Arnold F, Piecha T, Huber CG (2007) Automated semipreparative purification with mass spectrometric fraction collection trigger: modeling and experimental evaluation of a setup employing passive splitting. *J Sep Sci* 30(10):1496–1508
399. Drexler DM, Tiller PR (1998) ‘Intelligent’ fraction collection to improve structural characterization by mass spectrometry. *Rapid Commun Mass Spectrom* 12(13):895–900
400. Dear GJ, Plumb RS, Sweatman BC, Ismail IM, Ayrton J (1999) Tandem mass spectrometry linked fraction collection for the isolation of drug metabolites from biological matrices. *Rapid Commun Mass Spectrom* 13(10):886–894
401. Plumb RS, Ayrton J, Dear GJ, Sweatman BC, Ismail IM (1999) The use of preparative high performance liquid chromatography with tandem mass spectrometric directed fraction collection for the isolation and characterisation of drug metabolites in urine by nuclear magnetic resonance spectroscopy and liquid chromatography/sequential mass spectrometry. *Rapid Commun Mass Spectrom* 13(10):845–854
402. de Vlieger JSB, Kolkman AJ, Ampt KAM, Commandeur JNM, Vermeulen NPE, Kool J, Wijmenga SS, Niessen WMA, Irth H, Honing M (2010) Determination and identification of estrogenic compounds generated with biosynthetic enzymes using hyphenated screening assays, high resolution mass spectrometry and off-line NMR. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences* 878(7–8):667–674
403. Alexander AJ, Xu F, Bernard C (2006) The design of a multi-dimensional LC-SPE-NMR system (LC2-SPE-NMR) for complex mixture analysis. *Magn Reson Chem* 44(1):1–6
404. Corcoran O, Spraul M (2003) LC-NMR-MS in drug discovery. *Drug Discov Today* 8(14):624–631
405. Christophoridou S, Dais P, Tseng LH, Spraul M (2005) Separation and identification of phenolic compounds in olive oil by coupling high-performance liquid chromatography with postcolumn solid-phase extraction to nuclear magnetic resonance spectroscopy (LC-SPE-NMR). *J Agric Food Chem* 53(12):4667–4679
406. Tatsis EC, Boeren S, Exarchou V, Troganis AN, Vervoort J, Gerothanassis IP (2007) Identification of the major constituents of *Hypericum perforatum* by LC/SPE/NMR and/or LC/MS. *Phytochemistry* 68(3):383–393
407. Exarchou V, Godejohann M, van Beek TA, Gerothanassis IP, Vervoort J (2003) LC-UV-solid-phase extraction-NMR-MS combined with a cryogenic flow probe and its application to the identification of compounds present in Greek oregano. *Anal Chem* 75(22):6288–6294
408. Seger C, Godejohann M, Tseng LH, Spraul M, Girtler A, Sturm S, Stuppner H (2005) LC-DAD-MS/SPE-NMR hyphenation. A tool for the analysis of pharmaceutically used plant extracts: identification of isobaric iridoid glycoside regioisomers from *Harpagophytum procumbens*. *Anal Chem* 77(3):878–885
409. Bringmann G, Wohlfarth M, Rischer H, Schlauer J, Brun R (2002) Extract screening by HPLC coupled to MS-MS, NMR, and CD: a dimeric and three monomeric naphthylisoquinoline alkaloids from *Ancistrocladus griffithii*. *Phytochemistry* 61(2):195–204
410. Lin Y, Schiavo S, Orjala J, Vouros P, Kautz R (2008) Microscale LC-MS-NMR platform applied to the identification of active cyanobacterial metabolites. *Anal Chem* 80(21):8045–8054
411. Schlotterbeck G, Ceccarelli SM (2009) LC-SPE-NMR-MS: a total analysis system for bioanalysis. *Bioanalysis* 1(3):549–559
412. Castro A, Moco S, Coll J, Vervoort J (2010) LC-MS-SPE-NMR for the isolation and characterization of neo-clerodane diterpenoids from *Teucrium luteum* subsp *flavovirens*. *J Nat Products* 73(5):962–965
413. Sorensen D, Raditsis A, Trimble LA, Blackwell BA, Sumarah MW, Miller JD (2007) Isolation and structure elucidation by LC-MS-SPE/NMR: PR toxin- and cuspidatol-related eremophilane sesquiterpenes from *Penicillium roqueforti*. *J Nat Prod* 70(1):121–123
414. Miliauskas G, Van Beek TA, De Waard P, Venskutonis RP, Sudhölter EJ (2005) Identification of radical scavenging compounds in *Rhaponticum carthamoides* by means of LC-DAD-SPE-NMR. *J Nat Prod* 68(2):168–172
415. Kind T, Leamy T, Leary JA, Fiehn O (2009) Software platform virtualization in chemistry research and university teaching. *J Cheminformatics* 1(1):18
416. Dua S, Blanksby SJ, Bowie JH (2000) Electron capture of tetracyanoethylene oxide in the gas phase. Rearrangement of the parent radical anion to form $[(NC)_3C]^-$. A joint experimental and ab initio study. *Int J Mass Spectrom* 194(2–3):165–170
417. Hofmeister GE, Zhou Z, Leary JA (1991) Linkage position determination in lithium-cationized disaccharides—tandem mass-spectrometry and semiempirical calculations. *J Am Chem Soc* 113(16):5964–5970
418. Rodgers MT, Armentrout PB (1999) Absolute alkali metal ion binding affinities of several azoles determined by threshold collision-induced dissociation. *Int J Mass Spectrom* 187:359–380
419. Fell LM, Francis JT, Holmes JL, Terlouw JK (1997) The intriguing behaviour of (ionized) oxalacetic acid investigated by tandem mass spectrometry. *Int J Mass Spectrom* 165:179–194
420. Milletti F, Storch L, Sforna G, Cross S, Cruciani G (2009) Tautomer enumeration and stability prediction for virtual screening on large chemical databases. *J Chem Inf Model* 49(1):68–75
421. Ge GB, Zhang R, Ai CZ, He YQ, Zhang YY, Liu XB, Yang L, Wang ZT, Yang L (2009) Stereochemical differentiation of C-7 hydroxyltaxane isomers by electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom* 23(3):425–432
422. Hayen H, Michels A, Franzke J (2009) Dielectric barrier discharge ionization for liquid chromatography/mass spectrometry. *Anal Chem* 81(24):10239–10245
423. Lee AC, Crippen GM (2009) Predicting pKa. *J Chem Inf Model* 49(9):2013–2033
424. Sanz-Nebot V, Toro I, Barbosa J (2001) Prediction of retention behaviour and evaluation of pKa values of peptides and quinolones in liquid chromatography. *J Chromatogr A* 933(1–2):45–56
425. Marvin Calculator Plugins—resonance structures, tautomer, stereoisomer, pKa, logP calculation. <http://www.chemaxon.com/>
426. Bioclipse cheminformatics platform; [<http://www.bioclipse.net/>]
427. Molchanova MS, Shcherbukhin VV, Zefirov NS (1996) Computer generation of molecular structures by the SMOG program. *J Chem Inf Comput Sci* 36(4):888–899
428. Fontana P, Pretsch E (2002) Automatic spectra interpretation, structure generation, and ranking. *J Chem Inf Comput Sci* 42(3):614–619
429. Faulon J-L, Visco DP, Roe D (2005) Enumerating molecules. In: Kenny B, Lipkowitz RLTRC (eds) *Reviews in computational chemistry*. Wiley, New York, pp 209–286
430. Hawkins DM (2003) The problem of overfitting. *J Chem Inf Comput Sci* 44(1):1–12
431. Heberger K (2007) Quantitative structure–(chromatographic) retention relationships. *J Chromatogr A* 1158(1–2):273–305
432. Kalisz R (2007) QSRR: quantitative structure–(chromatographic) retention relationships. *Chem Rev* 107(7):3212–3246
433. Karelson M, Lobanov VS, Katritzky AR (1996) Quantum-chemical descriptors in QSAR/QSPR studies. *Chem Rev* 96(3):1027–1044

434. Albaugh DR, Hall LM, Hill DW, Kertesz TM, Parham M, Hall LH, Grant DF (2009) Prediction of HPLC retention index using artificial neural networks and IGroup E-state indices. *J Chem Inf Model* 49(4):788–799
435. Hill DW, Kind AJ (1994) Reversed-phase solvent-gradient HPLC retention indexes of drugs. *J Anal Toxicol* 18(5):233–242
436. Babushok VI, Linstrom PJ, Reed JJ, Zenkevich IG, Brown RL, Mallard WG, Stein SE (2007) Development of a database of gas chromatographic retention properties of organic compounds. *J Chromatogr A* 1157(1–2):414–421
437. Zenkevich IG, Babushok VI, Linstrom PJ, White E, Stein SE (2009) Application of histograms in evaluation of large collections of gas chromatographic retention indices. *J Chromatogr A* 1216(38):6651–6661
438. Stein SE, Babushok VI, Brown RL, Linstrom PJ (2007) Estimation of Kovats retention indices using group contributions. *J Chem Inf Model* 47(3):975–980
439. Baker JK, Ma C-Y (1979) Retention index scale for liquid–liquid chromatography. *J Chromatogr A* 169:107–115
440. Smith RM (1982) Alkylarylketones as a retention index scale in liquid chromatography. *J Chromatogr A* 236(2):313–320
441. Bogusz M, Wu M (1991) Standardized HPLC/DAD system, based on retention indices and spectral library, applicable for systematic toxicological screening. *J Anal Toxicol* 15(4):188–197
442. Krokhin OV, Spicer V (2009) Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. *Anal Chem* 81(22):9522–9530
443. Zaikin V, Halket JM (2009) A handbook of derivatives for mass spectrometry. IM, West Sussex, 1901019098
444. Fernandes-Whaley M, Muhlberger F, Whaley A, Adam T, Zimmermann R, Rohwer E, Walte A (2005) On-line derivatization for resonance-enhanced multiphoton ionization time-of-flight mass spectrometry: detection of aliphatic aldehydes and amines via reactive coupling of aromatic photo ionization labels. *Anal Chem* 77(1):1–10
445. Eggink M, Wijtmans M, Ekkebus R, Lingeman H, de Esch IJP, Kool J, Niessen WMA, Irth H (2008) Development of a selective ESI-MS derivatization reagent: synthesis and optimization for the analysis of aldehydes in biological mixtures. *Anal Chem* 80(23):9042–9051
446. Hong H, Wang Y (2007) Derivatization with girard reagent T combined with LC- MS/MS for the sensitive detection of 5-Formyl-2-deoxyuridine in cellular DNA. *Anal Chem* 79(1):322–326
447. Yang WC, Mirzaei H, Liu X, Regnier FE (2006) Enhancement of amino acid detection and quantification by electrospray ionization mass spectrometry. *Anal Chem* 78(13):4702–4708
448. Segura J, Ventura R, Jurado C (1998) Derivatization procedures for gas chromatographic–mass spectrometric determination of xenobiotics in biological samples, with special attention to drugs of abuse and doping agents. *J Chromatogr B Biomed Sci Appl* 713(1):61–90
449. Ternes TA, Andersen H, Gilberg D, Bonerz M (2002) Determination of estrogens in sludge and sediments by liquid extraction and GC/MS/MS. *Anal Chem* 74(14):3498–3504
450. Ternes TA (2001) Analytical methods for the determination of pharmaceuticals in aqueous environmental samples. *TrAC, Trends Anal Chem* 20(8):419–434
451. Sacher F, Lange FT, Brauch HJ, Blankenhorn I (2001) Pharmaceuticals in groundwaters: analytical methods and results of a monitoring program in Baden-Württemberg, Germany. *J Chromatogr A* 938(1–2):199–210
452. Molnár-Perl I, Horváth K (1997) Simultaneous quantitation of mono-, di- and trisaccharides as their TMS ether oxime derivatives by GC-MS: I. In model solutions. *Chromatographia* 45(1):321–327
453. Horvath K, Molnár-Perl I (1997) Simultaneous quantitation of mono-, di- and trisaccharides by GC-MS of their TMS ether oxime derivatives: II. In honey. *Chromatographia* 45(1):328–335
454. Zampolli MG, Basaglia G, Dondi F, Sternberg R, Szopa C, Pietrogrande MC (2007) Gas chromatography-mass spectrometry analysis of amino acid enantiomers as methyl chloroformate derivatives: application to space analysis. *J Chromatogr A* 1150(1–2):162–172
455. Kaspar H, Dettmer K, Chan Q, Daniels S, Nimkar S, Daviglus ML, Stamler J, Elliott P, Oefner PJ (2009) Urinary amino acid analysis: a comparison of iTRAQ-LC-MS/MS, GC-MS, and amino acid analyzer. *J Chromatogr B Anal Technol Biomed Life Sci* 877(20–21):1838–1846
456. Kiouisi P, Angelis YS, Lyris E, Koupparis M, Calokerinos AC, Atta-Politou J, Georgakopoulos CG (2009) Two-step silylation procedure for the unified analysis of 190 doping control substances in human urine samples by GC-MS. *Bioanalysis* 1(7):1209–1224
457. Somuramasami J, Winger BE, Gillespie TA, Kenttämaa HI (2009) Identification and counting of carbonyl and hydroxyl functionalities in protonated bifunctional analytes by using solution derivatization prior to mass spectrometric analysis via ion-molecule reactions. *J Am Soc Mass Spectrom* 21(5):773–784
458. Hummel J, Selbig J, Walther D, Kopka J (2007) The Golm Metabolome Database: a database for GC-MS based metabolite profiling. *Topics in Current Genetics* 18:75
459. Kind T, Wohlgemuth G, Lu Y, Palazoglu M, Shahbaz S, Fiehn O (2009) FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal Chem* 81(24):10038–10048
460. Guo K, Li L (2009) Differential ¹²C-/¹³C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem* 81(10):3919–3932
461. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH (2009) PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* 37(Web Server issue):W623
462. Williams AJ (2008) A perspective of publicly accessible/open-access chemistry databases. *Drug Discov Today* 13(11–12):495–501
463. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36(Database issue):D480
464. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, Hau DD, Psychogios N, Dong E, Bouatra S (2008) HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 37:D603–D610
465. Degtyarenko K, de Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, Alcantara R, Darsow M, Guedj M, Ashburner M (2008) ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res* 36(Database issue):D344
466. Wishart DS (2008) DrugBank and its relevance to pharmacogenomics. *Pharmacogenomics* 9(8):1155–1162
467. Draper J, Enot DP, Parker D, Beckmann M, Snowdon S, Lin W, Zubair H (2009) Metabolite signal identification in accurate mass metabolomics data with MZEDB, an interactive m/z annotation tool utilising predicted ionisation behaviour ‘rules’. *BMC Bioinformatics* 10(1):227
468. Nicklaus MC, Filippov IV, Sitzmann M, Ihlenfeldt W-D, Oellien F, Bienfait B, Voigt JH, Sun G (2010) Chemical structure lookup service. <http://cactus.nci.nih.gov/>
469. Brown M, Dunn WB, Dobson P, Patel Y, Winder CL, Francis-McIntyre S, Begley P, Carroll K, Broadhurst D, Tseng A (2009) Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst* 134(7):1322–1332

470. Villas-Bôas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J (2005) Mass spectrometry in metabolome analysis. *Mass Spectrom Rev* 24(5):613–646
471. Dettmer K, Aronov PA, Hammock BD (2007) Mass spectrometry-based metabolomics. *Mass Spectrom Rev* 26(1):51
472. Haleem JI, Que NV, Timothy JW, Gary MM, Timothy DV (2009) Analytical and statistical approaches to metabolomics research. *J Sep Sci* 32(13):2183–2199
473. Wishart DS (2009) Computational strategies for metabolite identification in metabolomics. *Bioanalysis* 1(9):1579–1596
474. Kertesz TM, Hill DW, Albaugh DR, Hall LH, Hall LM, Grant DF (2009) Database searching for structural identification of metabolites in complex biofluids for mass spectrometry-based metabolomics. *Bioanalysis* 1(9):1627–1643
475. Ma S, Zhu M (2009) Recent advances in applications of liquid chromatography–tandem mass spectrometry to the analysis of reactive drug metabolites. *Chem-Biol Interact* 179(1):25–37
476. Calza P, Medana C, Carbone F, Giancotti V, Baiocchi C (2008) Characterization of intermediate compounds formed upon photoinduced degradation of quinolones by high-performance liquid chromatography/high-resolution multiple-stage mass spectrometry. *Rapid Commun Mass Spectrom* 22(10):1533–1552
477. Li Q, Yan G, Ge T (2008) A fragmentation study of two compounds related to 4'-demethylepipodophyllotoxin in negative ion electrospray ionization by MSⁿ ion-trap time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 22(3):373–378
478. Gallart-Ayala H, Moyano E, Galceran MT (2007) Liquid chromatography/multi-stage mass spectrometry of bisphenol A and its halogenated derivatives. *Rapid Commun Mass Spectrom* 21(24):4039–4048
479. Yan G, Li Q, Tan H, Ge T (2007) Electrospray ionization ion-trap time-of-flight tandem mass spectrometry of two furofurans: sesamin and gmelinol. *Rapid Commun Mass Spectrom* 21(22):3613–3620
480. Piccinelli AL, Campone L, Dal Piaz F, Cuesta-Rubio O, Rastrelli L (2009) Fragmentation pathways of polycyclic polyisoprenylated benzophenones and degradation profile of nemorosone by multiple-stage tandem mass spectrometry. *J Am Soc Mass Spectrom* 20(9):1688–1698
481. Kováčik V, Bekesová S, Tvaroska I, Kováč P (2006) Positive Electrospray ion trap multistage mass spectrometric fragmentation of synthetic analogs of saccharide part of lipopolysaccharides of vibrio cholerae O:1. *J Am Soc Mass Spectrom* 17(6):749–756
482. Kirk JM, Tarbin J, Keely BJ (2006) Analysis of androgenic steroid Girard P hydrazones using multistage tandem mass spectrometry. *Rapid Commun Mass Spectrom* 20(8):1247–1252
483. Schmidt J, Boettcher C, Kuhnt C, Kutchan TM, Zenk MH (2007) Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring-13C₆]-tyramine feeding. *Phytochemistry* 68(2):189–202
484. Cataldi TRI, Lelario F, Bufo SA (2005) Analysis of tomato glycoalkaloids by liquid chromatography coupled with electrospray ionization tandem mass spectrometry. *Rapid Commun Mass Spectrom* 19(21):3103–3110
485. Dudley E, Tuytten R, Bond A, Lemièrre F, Brenton AG, Esmans EL, Newton RP (2005) Study of the mass spectrometric fragmentation of pseudouridine: comparison of fragmentation data obtained by matrix-assisted laser desorption/ionisation post-source decay, electrospray ion trap multistage mass spectrometry, and by a method utilising electrospray quadrupole time-of-flight tandem mass spectrometry and in-source fragmentation. *Rapid Commun Mass Spectrom* 19(21):3075–3085
486. Vrkic AK, O'Hair RAJ, Foote S (2000) Fragmentation reactions of all 64 deprotonated trinucleotide and 16 mixed base tetranucleotide anions via tandem mass spectrometry in an ion trap. *Aust J Chem* 53(4):307–320
487. Vukics V, Guttman A (2010) Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. *Mass Spectrom Rev* 29:1–16
488. Cuyckens F, Claeys M (2004) Mass spectrometry in the structural analysis of flavonoids. *J Mass Spectrom* 39(1):1–15
489. de Rijke E, Out P, Niessen WMA, Ariese F, Gooijer C, Brinkman UAT (2006) Analytical separation and detection methods for flavonoids. *J Chromatogr A* 1112(1–2):31–63
490. Portet B, Fabre N, Rozenberg R, Habib-Jiwan J-L, Moulis C, Quetin-Leclercq J (2008) Analysis of minor flavonoids in Piper hostmannianum var. berbicense using liquid chromatography coupled with atmospheric pressure chemical ionization mass spectrometry. *J Chromatogr A* 1210(1):45–54
491. Nakabayashi R, Kusano M, Kobayashi M, Tohge T, Yonekura-Sakakibara K, Kogure N, Yamazaki M, Kitajima M, Saito K, Takayama H (2009) Metabolomics-oriented isolation and structure elucidation of 37 compounds including two anthocyanins from Arabidopsis thaliana. *Phytochemistry* 70(8):1017–1029
492. Prakash C, Shaffer CL, Nedderman A (2007) Analytical strategies for identifying drug metabolites. *Mass Spectrom Rev* 26(3):340–369
493. Tozuka Z, Kaneko H, Shiraga T, Mitani Y, Beppu M, Terashita S, Kawamura A, Kagayama A (2003) Strategy for structural elucidation of drugs and drug metabolites using (MS)ⁿ fragmentation in an electrospray ion trap. *J Mass Spectrom* 38(8):793–808
494. Trunzer M, Faller B, Zimmerlin A (2008) Metabolic soft spot identification and compound optimization in early discovery phases using metasite and LC-MS/MS validation. *J Med Chem* 52(2):329–335
495. Pan C, Liu F, Ji Q, Wang W, Drinkwater D, Vivilecchia R (2006) The use of LC/MS, GC/MS, and LC/NMR hyphenated techniques to identify a drug degradation product in pharmaceutical development. *J Pharm Biomed Anal* 40(3):581–590
496. Zhang J, Schubothe K, Li B, Russell S, Lebrilla CB (2004) Infrared multiphoton dissociation of O-linked mucin-type oligosaccharides. *Anal Chem* 77(1):208–214
497. Chen Y, Monshouer M, Fitch WL (2007) Analytical tools and approaches for metabolite identification in early drug discovery. *Pharm Res* 24(2):248–257
498. Habicht SC, Vinuesa NR, Duan P, Fu M, Kenttämä HI (2010) Data-dependent neutral gain MS³: toward automated identification of the N-oxide functional group in drug metabolites. *J Am Soc Mass Spectrom* 21(4):559–563
499. Elyashberg ME, Williams AJ, Martin GE (2008) Computer-assisted structure verification and elucidation tools in NMR-based structure elucidation. *Prog Nucl Magn Reson Spectrosc* 53(1–2):1–104
500. Cheng KW, Wong CC, Wang M, He QY, Chen F (2009) Identification and characterization of molecular targets of natural products by mass spectrometry. *Mass Spectrom Rev* 29(1):126–155
501. Wilson ID, Brinkman UAT (2003) Hyphenation and hyperphenation: the practice and prospects of multiple hyphenation. *J Chromatogr A* 1000(1–2):325–356
502. Laniewski K, Vagero M, Forsberg E, Formgren T, Hagman G (2004) Complementary use of gas chromatography–mass spectrometry, gas chromatography–atomic emission detection and nuclear magnetic resonance for identification of pharmaceutically related impurities of unknown structures. *J Chromatogr A* 1027(1–2):93–102
503. Schiller J, Süß R, Arnhold J, Fuchs B, Leßig J, Müller M, Petkovic M, Spalteholz H, Zschörnig O, Arnold K (2004) Matrix-assisted laser desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry in lipid and phospholipid research. *Prog Lipid Res* 43(5):449–488
504. Pulfer M, Murphy RC (2003) Electrospray mass spectrometry of phospholipids. *Mass Spectrom Rev* 22(5):332–364
505. Han X, Gross RW (2005) Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of

- cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom Rev* 24(3):367–412
506. Souady J, Dadimov D, Kirsch S, Bindila L, Peter-Katalinic J, Vakhrushev SY (2010) Software utilities for the interpretation of mass spectrometric data of glycoconjugates: application to glycosphingolipids of human serum. *Rapid Commun Mass Spectrom* 24(7):1039–1048
507. Byrdwell WC, Neff WE (2002) Dual parallel electrospray ionization and atmospheric pressure chemical ionization mass spectrometry (MS), MS/MS and MS/MS/MS for the analysis of triacylglycerols and triacylglycerol oxidation products. *Rapid Commun Mass Spectrom* 16(4):300–319
508. Forrester JS, Milne SB, Ivanova PT, Brown HA (2004) Computational lipidomics: a multiplexed analysis of dynamic changes in membrane lipid composition during signal transduction. *Mol Pharmacol* 65(4):813–821
509. Yang K, Cheng H, Gross RW, Han X (2009) Automated lipid identification and quantification by multidimensional mass spectrometry-based shotgun lipidomics. *Anal Chem* 81(11):4356–4368
510. Schwudke D, Oegema J, Burton L, Entchev E, Hannich JT, Ejsing CS, Kurzychalia T, Shevchenko A (2006) Lipid profiling by multiple precursor and neutral loss scanning driven by the data-dependent acquisition. *Anal Chem* 78(2):585–595
511. Pittenauer E, Allmaier G (2009) The renaissance of high-energy CID for structural elucidation of complex lipids: MALDI-TOF/RTOF-MS of alkali cationized triacylglycerols. *J Am Soc Mass Spectrom* 20(6):1037–1047
512. Song H, Hsu FF, Ladenson J, Turk J (2007) Algorithm for processing raw mass spectrometric data to identify and quantitate complex lipid molecular species in mixtures by data-dependent scanning and fragment ion database searching. *J Am Soc Mass Spectrom* 18(10):1848–1858
513. Ibanez M, Sancho JV, Pozo Ó, Niessen W, Hernandez F (2005) Use of quadrupole time-of-flight mass spectrometry in the elucidation of unknown compounds present in environmental water. *Rapid Commun Mass Spectrom* 19(2):169–178
514. Schymanski EL, Bataineh M, Goss KU, Brack W (2009) Integrated analytical and computer tools for structure elucidation in effect-directed analysis. *TrAC, Trends Anal Chem* 28(5):550–561
515. Niessen WMA (2009) Group-specific fragmentation of pesticides and related compounds in liquid chromatography–tandem mass spectrometry. *J Chromatogr A* 1217(25):4061–4070
516. Krauss M, Singer H, Hollender J (2010) LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. *Analytical and Bioanalytical Chemistry* 397(3):943–951
517. Smyth WF (2005) Recent applications of capillary electrophoresis–electrospray ionisation–mass spectrometry in drug analysis. *Electrophoresis* 26(7–8):1334–1357
518. Klampfl CW (2006) Recent advances in the application of capillary electrophoresis with mass spectrometric detection. *Electrophoresis* 27(1):3–34
519. Buckingham S (2003) Bioinformatics: programmed for success. *Nature* 425(6954):209–215
520. Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P, Blonder N, Billheimer DD, Blackman RK, Bunk DM, Cardasis HL (2009) Performance metrics for liquid chromatography–tandem mass spectrometry systems in proteomic analyses and evaluation by the CPTAC network. *Molecular & Cellular Proteomics* M900223
521. Kind T, Scholz M, Fiehn O (2009) How large is the metabolome? A critical analysis of data exchange practices in chemistry. *PLoS ONE* 4(5):e5440
522. Bradley JC, Lancashire RJ, Lang A, Williams AJ (2009) The spectral game: leveraging Open Data and crowdsourcing for education. *J Cheminformatics* 1(1):9
523. Go EP (2009) Database resources in metabolomics: an overview. *Journal of Neuroimmune Pharmacology* 5:18–30
524. Han J, Datla R, Chan S, Borchers CH (2009) Mass spectrometry-based technologies for high-throughput metabolomics. *Mass Spectrometry* 1(9):1665–1684
525. Downing J, Murray-Rust P, Tonge AP, Morgan P, Rzepa HS, Cotterill F, Day N, Harvey MJ (2008) SPECTRA: the deposition and validation of primary chemistry research data in digital repositories. *J Chem Inf Model* 48(8):1571
526. Neuweger H, Albaum SP, Dondrup M, Persicke M, Watt T, Niehaus K, Stoye J, Goesmann A (2008) MeltDB: a software platform for the analysis and integration of metabolomics experiment data. *Bioinformatics* 24(23):2726
527. Scholz M, Fiehn O (2007) SetupX—a public study design database for metabolomic projects. *Pac Symp Biocomput* 2007:169–180
528. Mendes P, Mendes P, Lewis IA, Lewis IA (2008) Metabolite identification via the Madison Metabolomics Consortium Database. *Nat Biotechnol* 26(2):162–164
529. Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) METLIN: a metabolite mass spectral database. *Ther Drug Monit* 27(6):747
530. Sana TR, Roark JC, Li X, Waddell K, Fischer SM (2008) Molecular formula and METLIN Personal Metabolite Database matching applied to the identification of compounds generated by LC/TOF-MS. *Journal of Biomolecular Techniques: JBT* 19(4):258–266
531. Shinbo Y, Nakamura Y, Altaf-UI-Amin M, Asahi H, Kurokawa K, Arita M, Saito K, Ohta D, Shibata D, Kanaya S (2006) KNApSACk: a comprehensive species–metabolite relationship database. *Plant Metabolomics* 57:165–181
532. Matsuda F, Yonekura-Sakakibara K, Niida R, Kuromori T, Shinozaki K, Saito K (2009) MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant J* 57(3):555
533. Hamacher M, Stephan C, Meyer HE, Eisenacher M (2009) Data handling and processing in proteomics. *Expert Rev Proteomics* 6(3):217–219
534. Rodriguez H (2008) International summit on proteomics data release and sharing policy. *J Proteome Res* 7(11):4609
535. Jennifer AM, Ian PS, Conrad B (2007) Public proteomic MS repositories and pipelines: available tools and biological applications. *Proteomics* 7(16):2769–2786
536. Falkner JA, Andrews PC (2007) Tranche: secure decentralized data storage for the proteomics community. *J Biomol Tech* 18(1):3
537. Falkner JA, Falkner JW, Andrews PC (2007) ProteomeCommons.org IO Framework: reading and writing multiple proteomics data formats. *Bioinformatics* 23(2):262
538. Pedrioli PGA, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Rought B, Pratt B, Nilsson E, Angeletti RH, Apweiler R (2004) A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* 22(11):1459–1466
539. Eric D (2008) mzML: a single, unifying data format for mass spectrometer output. *Proteomics* 8(14):2776–2777
540. Droit A, Fillon J, Morissette J, Poirier GG (2006) Bioinformatic standards for proteomics-oriented mass spectrometry. *Curr Proteomics* 3(2):119–128
541. Deutsch EW, Lam H, Aebersold R (2008) Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. *Physiol Genomics* 33(1):18

542. Davis G, Rew R, Hartnett E, Caron J, Emmerson S, Davies H (2010) netCDF exchange format description and software. <http://www.unidata.ucar.edu/software/netcdf/>
543. Lampen P, Hillig H, Davies AN, Linscheid M (1994) JCAMP-DX for mass spectrometry. *Appl Spectrosc* 48 (12):1545–1552
544. Murray-Rust P (2008) Open data in science. *Ser Rev* 34(1):52–64
545. McLafferty FW, Stauffer DA, Loh SY, Wesdemiotis C (1999) Unknown identification using reference mass spectra. Quality evaluation of databases. *J Am Soc Mass Spectrom* 10 (12):1229–1240
546. Ausloos P, Clifton CL, Lias SG, Mikaya AI, Stein SE, Tchekhovskoi DV, Sparkman OD, Zaikin V, Zhu D (1999) The critical evaluation of a comprehensive mass spectral library. *J Am Soc Mass Spectrom* 10(4):287–299