Spectral Accuracy of Molecular Ions in an LTQ/Orbitrap Mass Spectrometer and Implications for Elemental Composition Determination

John C. L. Erve,[a] Ming Gu,[b] Yongdong Wang,[b] William DeMaio,[a] and Rasmy E. Talaat[a]

[a] Drug Safety and Metabolism, Wyeth Research, Collegeville, Pennsylvania, USA
[b] Cerno Bioscience, Danbury, Connecticut, USA

In addition to mass accuracy, the ability of a mass spectrometer to faithfully measure the isotopic distribution of an ion, defined as spectral accuracy, is also important. Although time-of-flight mass spectrometers are reported to possess high spectral accuracy capability compared with other mass spectrometers, the Orbitrap has not yet been investigated. Ten natural products (moxidectin, erythromycin, digoxin, rifampicin, amphotericin B, rapamycin, gramicidin S, cyclosporin A, vancomycin, and thiostrepton) ranging in molecular weight from 639 to 1663 Da were measured on an LTQ/Orbitrap mass spectrometer with resolving power settings of 7.5, 15, 30, 60, and 100 K. The difference in the observed profile isotope pattern compared with the theoretical calculation after peak shape calibration, denoted spectral error, was calculated using the program MassWorks (Cerno Bioscience, Danbury, CT, USA). Spectral errors were least at 7.5 K resolving power (≤3%) but exceeded 10% for some compounds at 100 K. The increasing spectral error observed at higher resolving power for compounds with complex fine structure might be explained by the phenomena of isotopic beat patterns as observed in FTICR. Several compounds with prominent doubly charged ions allowed comparison of spectral accuracies of singly- versus doubly-charged ions. When using spectral error to rank elemental compositions with formula constraints ($C_{0\text{-}100}H_{0\text{-}200}N_{0\text{-}50}O_{0\text{-}50}Cl_{0\text{-}5}S_{0\text{-}5}$) and a mass tolerance ≤2 parts-per-million, the correct formula was ranked first 35% of the time. However, spectral error considerations eliminated >99% of possible elemental formulas for compounds with molecular weight >900 Da. (J Am Soc Mass Spectrom 2009, 20, 2058–2069) © 2009 American Society for Mass Spectrometry

Despite significant advances in both mass spectrometry hardware and software tools, identification and structure elucidation of unknowns remains challenging. Although high mass accuracy and precision is of great value to help establish the elemental composition of small molecules, when confronted with complex elemental compositions at larger masses (>600 Da), high mass accuracy (<1 ppm) alone cannot exclude enough potential candidates to be truly useful [1, 2]. Traditionally, the list of compound candidates can be reduced by limiting the possible elements and applying other chemical constraints, but the list can still easily contain dozens of compounds. However, high-resolution mass spectra also contain peak intensity information generated by the natural abundances of the various isotopes contained in the compound. This isotope information can be used to eliminate elemental compositions that despite an excellent match based on mass are not consistent with the observed isotope pattern [3]. Simple measurement of the relative intensities of the isotope peaks (M, M + 1, M + 2 . . .) has long been recognized as a powerful additional metric for reducing the possible elemental compositions for Br, Cl, or S containing compounds with their distinctive isotope patterns. For the subtle isotope differences generated by C, H, N, and O, however, more sophisticated computational approaches are required for effective elemental composition elucidation.

There are many published methods for utilizing isotope information in the context of elemental composition determination. Based on methodology by which isotopic information is processed, they can be classified into three categories:

1. Utilizing mass spectral centroid data, i.e., a limited list of typically 2 to 5 prominent isotopes for small molecules. This class of methods include early work by Grange and coworkers [4], a commercially available tool called iFIT through MassLynx software (Waters, Milford, MA, USA), and patented approaches by Zweigenbaum, who utilized isotope abundances and mass defects of the prominent

Address reprint requests to Dr. J. C. L. Erve, Drug Safety and Metabolism, Wyeth Research, 500 Arcola Road, Collegeville, PA 19426, USA. E-mail: john_erve@hotmail.com

1044-0305/09/$32.00
doi:10.1016/j.jasms.2009.07.014

Published online July 30, 2009
Received May 13, 2009
Revised July 18, 2009
Accepted July 18, 2009

isotope peaks [5], and Amirav, who used isotope abundances in combination with supersonic molecular beam or electron ionization [6]. A recent addition to this type of methodology, the "FuzzyFit" process, is described by Hobby and coworkers [7]. Since the fine and ultra-fine structures of higher isotopes are typically overlapped with each other due to the very close proximity of their *m/z* values, even on FT ICR MS instruments with R ≈ 1,000,000, centroid isotope data are only marginally useful due to the gross approximations of these overlapped fine structures in the mass spectral centroiding process, as pointed out by Wang and Gu [8]. It is in these overlapped fine structures that the information most relevant to elemental composition is contained, even when these fine structures are under-resolved or unresolved mass spectrally, a point to be discussed further later in this paper.

2. Taking advantage of the fact that, for a given compound (ion), the isotope pattern is unique and theoretically predictable even if the individual isotopes and isobars are not fully resolved, and using profile mass spectral data without centroiding for elemental composition determination. This may include the SigmaFit approach in Bruker Daltonics software, of which the exact details are unknown to the authors due to the proprietary nature of the algorithms, although a reference to a patent for the purpose of precise mass positioning is available [9]. An extensive evaluation of the SigmaFit approach was recently published by Bristow et al. [10]. A similar approach was independently proposed by Fernandez-de-Cossio et al. [11]. Although superior to the first class of methods because centroiding is avoided, this approach also entails significant error, arising from the following fact: all mass spectral profile mode data come with a given but not accurately known or defined mass spectral peak shape function. This function describes the statistical distribution of a population of ions of a given isotope along the *m/z* axis, and the full width at half maximum (FWHM) of this distribution approximates instrument resolution. Since the theoretical isotope distribution is a discrete distribution representing mass spectral data measured on a mass spectrometer of both infinite resolving power and linear dynamic range, an assumed peak shape function is required to convert the theoretical isotope distribution into profile mode mass spectral data before it can be compared with experimentally observed mass spectral data. In this conversion process, the assumed peak shape function is superimposed onto a theoretical isotope abundance distribution for comparison. Grange observed that such an assumption could lead to as much as 2.5% error [4], which is too large to differentiate closely related elemental compositions, thereby limiting the power of this methodology. It should be noted here that although some mass spectral systems are known to have a certain type of peak shape function, these peak shape functions, as measured from an actual system, typically vary from one instrument to another or even from one tune or data acquisition to another, giving rise to a modeling error of a few (up to more than 10) percent and significantly hamper the ability to differentiate closely related elemental compositions, whose theoretical isotope profiles may differ by far less than this modeling error. For example, quadrupole and magnetic sector mass spectrometers are typically considered as having Gaussian-type peak shape functions when the actually measured peak shape function may be modified Gaussians with unsymmetrical peak shapes. Similarly, in the case of FT ICR or Orbitrap mass spectrometers, the peak shape may nominally be Lorentzian but specifically dependent upon the processing method used. For a Fourier transform experiment, this typically involves various apodization functions and zero-filling, which can impact on quantitative peak height and area measurements [12].
3. Performing a mass spectral calibration, including the peak shape function, before elemental composition determination with profile mass spectral data. After the mass spectral peak shape calibration [13], the peak shape becomes accurately known and mathematically defined. When the most likely elemental composition needs to be determined from numerous possible formula candidates, the same mass spectral peak shape function is superimposed onto a discrete isotope abundance distribution to form a theoretical profile mass spectrum for each possible candidate formula. Since no assumption is made about the peak shape function used for both the calibrated and theoretical mass spectrum, the theoretical mass spectrum **t** should exactly match the calibrated experimental mass spectrum **r**, both of which are expressed and compared in vector form to preserve all isotope fine structural information even if these fine structures are under-resolved or unresolved, except for a difference in scaling between the two vectors due to ion abundance, assuming (1) no systematic or random measurement error and (2) a given candidate formula is indeed the correct formula. The scaling factor c can be easily established through a linear fitting process such as least-squares regression between vector **t** and **r**,

$$\mathbf{r} = c\mathbf{t} + \mathbf{e} \tag{1}$$

As a result of this fitting process, a metric called Spectral Accuracy (SA%) [14] can be readily calculated,

$$SA\% = \left(1 - \frac{\|\mathbf{e}\|_2}{\|\mathbf{r}\|_2}\right) \times 100 \tag{2}$$

to describe the congruence between the calibrated and theoretical isotope profile data vectors, where **e** is the

fitting residual or spectral error vector, **r** is the calibrated isotope profile vector from the experimental measurement, and $\|.\|_2$ represents the 2-norm (or square root of the sums of squares of all elements) of a vector. This spectral accuracy metric is used to evaluate all possible formulas whose exact monoisotope masses come within a user-defined mass tolerance window of the reported accurate mass from the actual monoisotope peak and whose elemental compositions satisfy the user-defined chemical constraints. A related metric, called Spectral Error (SE%), can also be similarly defined,

$$SE\% = \frac{\|\mathbf{e}\|_2}{\|\mathbf{r}\|_2} \times 100 \quad (3)$$

The formulas with the higher spectral accuracy (or lower spectral error) are ranked higher and considered as the more likely candidate formulas for the unknown ion of interest. It should be noted that the matrix computational approach outlined above avoids the issue of having to choose between peak height and/or peak area for comparison by keeping all the information, whether mass spectrally resolved or not, in the form of a continuously sampled mass spectral vector array in a mass spectral fitting process so as to preserve all isotope information, regardless of the actual resolving power used. This is a key advantage in addition to the peak shape function calibration. The interested reader can find the algorithmic details for this approach described fully in the patent applications [15, 16].

There are several published papers on the efficient calculation of theoretical isotope distributions, including both discrete distributions and the superimposition or convolution of a peak shape function for continuum distributions, notably by Yergey [17] and most recently by Rockwood et al. [18, 19]. However, these programs do not provide a numerical assessment comparing the theoretical spectrum with the experimentally observed spectrum as does the program self-Calibrated Lineshape Isotope Profile Search (sCLIPS) described here.

Commercially available LTQ/Orbitrap mass spectrometers allow the user to choose the resolving power with which to acquire a spectrum, at 7.5, 15, 30, 60, and 100 K (defined at *m/z* 400) [20]. In the Orbitrap, resolving power is inversely proportional to the measurement time so that higher resolution data require longer measurement times [21]. This can limit the use of the high-resolution capabilities when conducting LC analysis due to finite chromatographic peak widths. The commercially available Triversa NanoMate with its electrospray ionization (ESI)-chip employed in the direct infusion mode removes any time constraints to high-resolution measurements because the ESI-chip, which contains 5 μm diameter nozzles, leads to flow rates measured in nL per minute so that μL volumes can be analyzed for an hour or more. In addition to high-resolution capabilities, the mass accuracies of the Orbitrap are often within 2 ppm when employing internal calibration [22] and sub-ppm mass accuracies have also been reported [23].

Previous reports have investigated how isotope patterns can be used as a tool to help identify unknowns on various mass spectrometers. To date, no such investigations have been undertaken on the newest class of mass spectrometer, the Orbitrap. Thus, the aim of the present study was to evaluate Orbitrap spectral accuracy at various resolving power settings and its relative importance in elemental composition determinations when compared to mass accuracy. To this end, ten commercially available natural products with masses between 639 Da and 1664 Da were obtained (Figure 1 and Table 1). Molecular ions were recorded in selected ion monitoring (SIM) mode in an Orbitrap instrument with both mass and spectral accuracies determined.

Experimental

Chemicals

Vancomycin, cyclosporin A, erythromycin, rifampicin. and moxidectin were purchased from Fluka (Stenheim, Switzerland). Rapamycin and digoxin were obtained from Wyeth Research (Pearl River, NY, USA). Caffeine, verapamil, amphotericin B, thiostrepton, gramicidin S, sodium dodecyl sulfate (SDS), and sodium taurocholate were from Sigma-Aldrich (St. Louis, MO, USA). High-performance liquid chromatography (HPLC) grade methanol and dimethyl sulfoxide (DMSO) were obtained from EMD Chemicals (Gibbstown, NJ, USA); ethanol and acetic acid were obtained from Mallinckrodt Baker (Phillipsburg, NJ, USA). Formic acid was obtained from Fisher Scientific (Fair Lawn, NJ, USA). MRFA peptide was purchased from Research Plus, Inc, (Manasquan, NJ, USA), and Ultramark 1621 was obtained from Alfa Aesar (Ward Hill, MA, USA). ES-TOF tuning mix was from Agilent (New Castle, DE, USA). ESI-chips (type A) and pre-cleaned V-bottomed 96-well storage plates were purchased from Advion BioSciences, Inc (Ithaca, NY, USA).

Nanoelectrospray Ionization Mass Spectrometry

Stock solutions were prepared in ethanol for rifampicin (2.3 mM), cyclosporin A (1.1 mM), moxidectin (1.4 mM), and rapamycin (3.3 mM); water for vancomycin (1.0 mM), amphotericin B (0.27 mM), and erythromycin (1.4 mM); DMSO for thiostrepton (1.5 mM), digoxin (20 mM), and gramicidin S (0.88 mM). Stock solutions were diluted in 50:50 MeOH:water containing 0.1% formic or 0.1% acetic acid for positive or negative ESI, respectively, and diluted to ~2–10 μM for MS analysis. An LTQ-Orbitrap hybrid mass spectrometer (Thermo

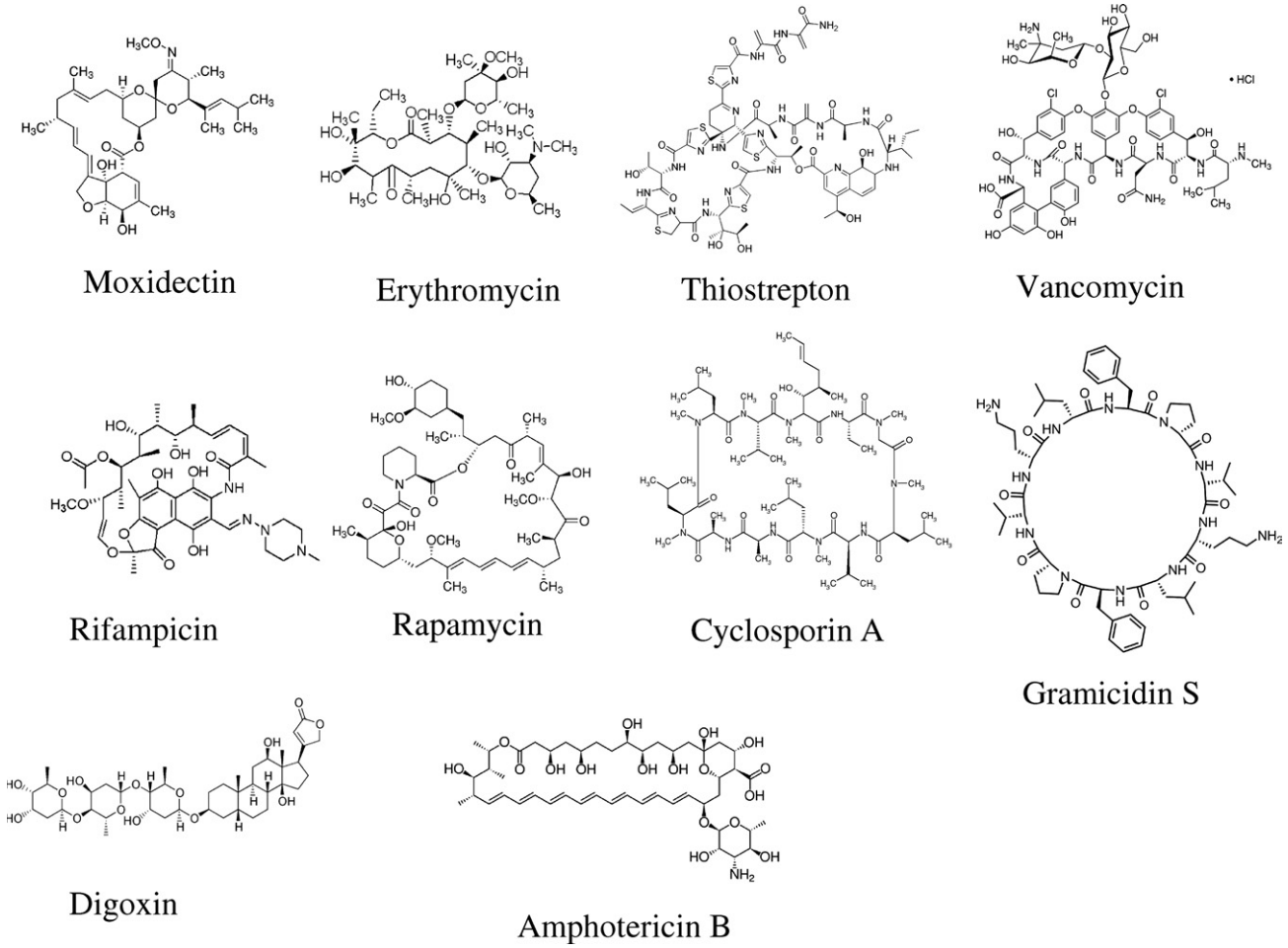

Figure 1. Structures of the compounds used in the study.

Fisher Scientific, Bremen, Germany) equipped with a TriVersa NanoMate chip-based electrospray ionization system (Advion BioSciences Inc., Ithaca, NY, USA) was operated in the positive ionization mode with a spray voltage of 1.4–1.7 kV or in the negative ionization mode with a spray voltage of 1.65–1.85 kV. No sheath or auxiliary gas was used and the ion transfer tube was maintained at 125 °C. Samples (5 to 8 μL) were introduced to the mass spectrometer by flow infusion using a head pressure of 40–50 psi at ~50–150 nL/min via the ESI chip (5 μm nozzle diameter). Pre-cleaned V-bottomed 96-well storage plates were used to hold samples before introduction into the mass spectrometer.

Table 1. Molecular formulas and exact masses of the protonated or deprotonated molecules investigated in the study

Compound	Molecular formula	Polarity	Monoisotopic *m/z*
Moxidectin	C37H53NO8	(−)	638.3698
Erythromycin	C37H67NO13	(−)	732.4539
Digoxin	C41H64O14	(−)	779.4223
Rifampicin	C43H58N4O12	(+)	823.4124
Amphotericin B	C47H73NO17	(+)	924.4951
Rapamycin	C51H79NO13	(+)	936.5444*
Gramicidin S	C60H92N12O10	(+)	1141.7132
Cyclosporin A	C62H111N11O12	(+)	1202.8486
Vancomycin	C66H75Cl2N9O24	(+)	1448.4375
Thiostrepton	C72H85N19O18S5	(+)	1664.4996
Agilent Ion	C30H19N3O6P3F48	(+)	1521.9715

*As sodium adduct.

The Orbitrap mass analyzer was calibrated at least bimonthly according to the manufacturer's directions using a mixture of caffeine, MRFA peptide, and Ultramark for positive ionization mode or a mixture of MRFA peptide, Ultramark, SDS, and sodium taurocholate for negative ionization mode. Selected ion monitoring (SIM) profile MS data (20–25 Da window and 40 scans) were acquired in the Orbitrap with resolving power settings of 7.5, 15, 30, 60, and 100 K. The injection time was from 50 to 350 ms with a target of 100,000 ions in the Orbitrap adjusted by automatic gain control (AGC). To achieve the highest mass accuracy possible, the lock mass function was enabled and either verapamil (*m/z* 455.29043) in positive ionization mode or SDS (*m/z* 265.147903) in negative ionization mode were used for real time internal recalibration. Mass spectral data acquisition and processing were performed using

Xcalibur software (version 2.0.7; Thermo Fisher Scientific, San Jose, CA, USA).

Spectral accuracy was calculated by MassWorks software (version 2.0, Cerno Bioscience, Danbury, CT, USA) using sCLIPS, which is a formula determination tool that performs peak shape calibration and matches calibrated experimental isotope pattern against possible theoretical ones using the spectral accuracy metric discussed above. The peak shape calibration is created using the monoisotope peak of the ion itself as the peak shape standard. A screen shot of the sCLIPS interface is shown in Figure 2 showing the elemental constraints and other required inputs. The profile mass range determines the relative mass spectral range used for the isotope profile comparison and the spectral accuracy calculation. The calibration range, determined by the approximate peak width of the monoisotopic peak, was ±0.5, ±0.25, ±0.1, ±0.075, and ±0.05 Da for singly charged ions or ±0.25, ±0.13, ±0.05, ±0.038, and ±0.025 Da for doubly charged ions at resolving power settings of (R =) 7.5, 15, 30, 60, and 100 K, respectively. The mass tolerance for the sCLIPS searches was 2 ppm. Spectral error (%) was obtained by subtracting the calculated spectral accuracy value from 100%, as described in eq 3.

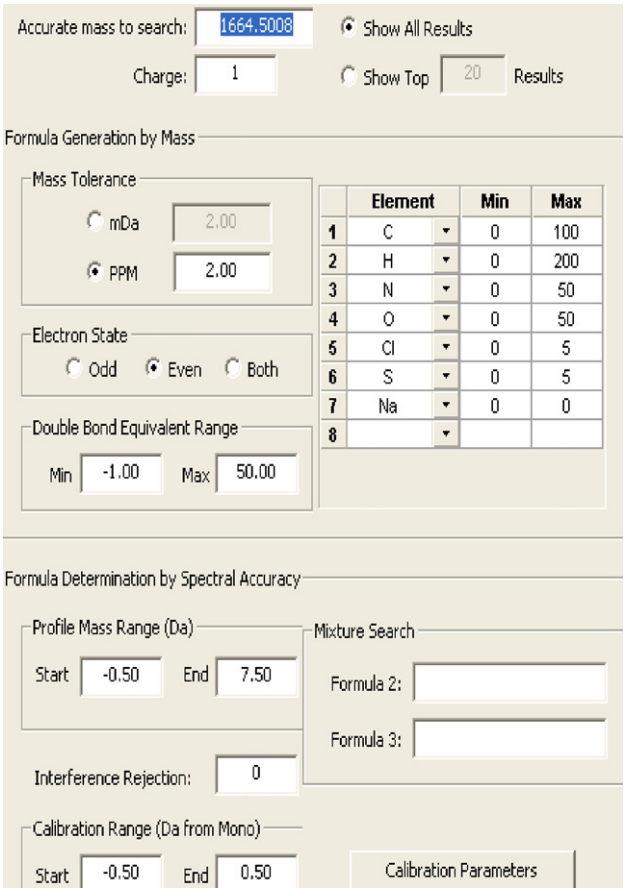

Figure 2. Screen shot of sCLIPS interface.

Results and Discussion

With few exceptions, spectral error (%) increased as a function of resolving power (see Table 2). Estimated uncertainties for spectral errors are based on quadruplicate measurements on a single day. The corresponding mass errors are also indicated in Table 2 and are classified as ≤2.0, ≤1.5, ≤1.0, or ≤0.5 ppm. Mass accuracy did not significantly improve when resolving power was increased from 7.5 to 100 K, likely reflecting the purity of the sample being analyzed and the fact the remaining systematic error in mass calibration dominates the mass error. Had random error from ion measurement statistics been the dominating source of mass error, the mass error should have decreased systematically as the resolving power increased, even with AGC.

Table 3 shows the rank of the correct elemental formula when using spectral error or mass error to sort the possible formulas. The rank of a compound based on mass error was determined by the number of elemental formulas with mass errors equal to or less than that for the correct compound, e.g., if the observed mass error was 1 ppm, the number of elemental compositions with mass errors ≤1 ppm were counted and reported as the rank for this compound. Results for each individual compound are briefly discussed below.

Moxidectin

Spectral error was 1.95% at R = 7.5 K and increased to 4.95% at R = 100 K. Ranking based on spectral error placed moxidectin in second place for all resolution settings out of between 31 and 42 possible elemental compositions. When using mass accuracy to rank compounds, moxidectin (with mass error <0.5 ppm at all settings) was ranked between 4 and 8 depending on resolving power setting.

Erythromycin

Spectral errors increased monotonically from 1.60% at R = 7.5 K to 2.60% at R = 100 K. Ranking based on spectral error placed erythromycin in first place at all resolving power settings out of ~48 possible elemental compositions. When ranked based on mass error, erythromycin was ranked between 16th and 23rd place depending on resolving power setting.

Digoxin

Spectral errors ranged from 1.65% at R = 7.5 K to 3.86% at R = 100 K. Ranking based on spectral error placed digoxin between first place at R = 7.5 K and fourth place at R = 100 K, out of ~100 possible elemental compositions. When ranked based on mass error, digoxin was ranked between 10th and 37th place depending on resolving power setting.

Table 2. Mass and spectral errors for measurements made at 7.5, 15, 30, 60, and 100 K resolving power on the Orbitrap

Name	Error	Orbitrap resolving power setting (K)				
		7.5	15	30	60	100
Moxidectin	Mass (ppm)	<2	<1	<0.5	<0.5	<0.5
	Spectral (%)	1.95 ± 0.04	1.58 ± 0.02	2.11 ± 0.02	2.94 ± 0.02	4.95 ± 0.02
Erythromycin	Mass (ppm)	<2	<2	<2	<2	<2
	Spectral (%)	1.60 ± 0.03	1.69 ± 0.07	2.15 ± 0.02	2.33 ± 0.03	2.60 ± 0.01
Digoxin	Mass (ppm)	<2	<1.5	<1.5	<1.5	<1.5
	Spectral (%)	1.65 ± 0.18	3.09 ± 0.26	2.78 ± 0.09	3.32 ± 0.05	3.86 ± 0.18
Rifampicin	Mass (ppm)	<1	<0.5	<0.5	<1	<1
	Spectral (%)	1.36 ± 0.15	1.33 ± 0.03	1.68 ± 0.23	2.55 ± 0.20	2.64 ± 0.70
Amphotericin B	Mass (ppm)	<0.5	<1.5	<1.5	<1	<1.5
	Spectral (%)	1.58 ± 0.14	1.73 ± 0.03	2.15 ± 0.10	2.67 ± 0.24	3.15 ± 0.40
Rapamycin	Mass (ppm)	<0.5	<2	<1	<1	<1
	Spectral (%)	1.11 ± 0.06	1.72 ± 0.02	1.89 ± 0.06	1.95 ± 0.02	2.20 ± 0.04
Gramicidin S	Mass (ppm)	<1	<2	<1	<1.5	<1.5
	Spectral (%)	0.94 ± 0.12	1.06 ± 0.06	1.62 ± 0.03	1.97 ± 0.07	3.04 ± 0.20
Gramicidin S*	Mass (ppm)	<0.5	<0.5	<0.5	<0.5	<0.5
	Spectral (%)	1.62 ± 0.01	1.84 ± 0.03	2.57 ± 0.02	3.71 ± 0.01	6.48 ± 0.04
Cyclosporin A	Mass (ppm)	<1	<1	<1.5	<1.5	<1.5
	Spectral (%)	0.63 ± 0.17	1.67 ± 0.16	1.74 ± 0.27	2.13 ± 0.20	3.37 ± 0.10
Vancomycin	Mass (ppm)	<1	<1.5	<1	<1	<1
	Spectral (%)	2.23 ± 0.14	2.75 ± 0.22	3.80 ± 0.17	8.83 ± 0.13	12.81 ± 0.34
Vancomycin (fragment)	Mass (ppm)	<1	<1.5	<1	<1	<1
	Spectral (%)	1.46 ± 0.04	1.52 ± 0.06	2.13 ± 0.03	3.49 ± 0.02	5.99 ± 0.08
Vancomycin*	Mass (ppm)	<1.5	<1	<1	<1	<1
	Spectral (%)	3.00 ± 0.01	3.03 ± 0.03	3.27 ± 0.03	4.54 ± 0.04	4.51 ± 0.04
Thiostrepton	Mass (ppm)	<2	<1	<1.5	<1.5	<1.5
	Spectral (%)	2.38 ± 0.17	1.41 ± 0.13	3.18 ± 0.10	3.59 ± 0.13	5.64 ± 0.39
Thiostrepton*	Mass (ppm)	<1	<1	<0.5	<1	<1
	Spectral (%)	1.25 ± 0.29	1.71 ± 0.27	2.64 ± 0.17	4.05 ± 0.23	6.48 ± 0.33
Agilent 1521	Mass (ppm)	<1	<0.5	<0.5	<0.5	<1
	Spectral (%)	1.39 ± 0.16	1.85 ± 0.04	1.91 ± 0.08	2.14 ± 0.06	2.65 ± 0.11

*Doubly charged.

Rifampicin

Spectral errors increased from 1.36% at R = 7.5 K to 2.64% at R = 100 K. Ranking based on spectral error placed rifampicin in either third place (R = 7.5–60 K) or first place at the 100 K resolution setting, out of ~150 possible elemental compositions. Rifampicin can readily oxidize through the loss of two hydrogen atoms to rifampicin-quinone [24] and, indeed, this species was visible in the spectrum (Figure 3), which increased even upon storage at −70 °C (data not shown). MassWorks can calculate the spectral error of a spectrum representing a mixture of related compounds whose compositional relationships are known. If the spectrum is treated as pure rifampicin, the contribution of the M + 2, M + 3, and M + 4 peaks from rifampicin-quinone will distort the M, M + 1, and M + 2 peaks of rifampicin, resulting in a higher spectral error. However, when treating the spectrum as a mixture of rifampicin and its −2 Da oxidation product, spectral errors decreased by ~0.4% to 1.1% compared with treating the spectrum as pure rifampicin, in turn resulting in a higher ranking at 7.5–30 K resolving power settings. When ranked based on mass error, rifampicin was ranked between 13th and 60th place depending on resolving power settings.

Amphotericin B

Spectral errors increased monotonically from 1.58% at R = 7.5 K to 3.15% at R = 100 K. Ranking based on spectral error placed amphotericin B in first place for all resolving power settings except 100 K (second place) out of ~190 possible elemental compositions. When ranked based on mass error, amphotericin B was ranked between 26th and 45th place depending on resolving power settings.

Rapamycin

Spectral errors increased monotonically from 1.11% at R = 7.5 K to 2.20% at R = 100 K. Ranking based on spectral error placed rapamycin in first or second place for all resolving power settings out of ~280 possible elemental compositions. When ranked based on mass error, rapamycin was ranked between 12th and 67th place depending on resolving power settings.

Gramicidin S

Both singly and doubly charged ions were investigated. Spectral errors increased monotonically from 0.94% at R = 7.5 K to 3.04% at R = 100 K, and from 1.62% at R =

Table 3. Ranking of elemental formulas based on spectral error or mass error†

Compound	Rank	Orbitrap resolving power (K)				
		7.5	15	30	60	100
Moxidectin	Spectral error	2	2	2	2	2
	Mass error	4 (42)	8 (34)	8 (33)	7 (31)	4 (32)
Erythromycin	Spectral error	1	1	1	1	1
	Mass error	23 (48)	16 (45)	16 (45)	16 (45)	16 (46)
Digoxin	Spectral error	1	3	3	3	4
	Mass error	10 (97)	16 (99)	22 (96)	37 (99)	22 (99)
Rifampicin	Spectral error	1^	3/2	3/1	3/3	1/1
	Mass error	60 (147)	13 (147)	46 (144)	34 (151)	24 (148)
Amphotericin B	Spectral error	3/2	1	1	1	2
	Mass error	30 (186)	45 (188)	40 (187)	26 (191)	28 (189)
Rapamycin	Spectral error	1	2	1	2	1
	Mass error	67 (281)	30 (275)	12 (278)	12 (278)	21 (278)
Gramicidin S	Spectral error	1	1	1	3	4
	Mass error	31 (570)	194 (577)	79 (584)	108 (577)	80 (584)
Gramicidin S*	Spectral error	1	1	1	3	3
	Mass error	183 (369)	76 (303)	97 (297)	97 (297)	54 (302)
Cyclosporin A	Spectral error	1	2	2	3	1
	Mass Error	134 (1,089)	200 (1,089)	245 (1,089)	222 (1,099)	270 (1,090)
Vancomycin	Spectral error	6	6	7	8	31
	Mass error	400 (1,523)	86 (1,515)	81 (1,529)	392 (1,529)	391 (1,529)
Vancomycin*	Spectral error	6	5	6	7	7
	Mass error	314 (769)	315 (769)	314 (769)	234 (774)	234 (774)
Thiostrepton	Spectral error	5	2	3	6	7
	Mass error	193 (1,908)	529 (1,919)	218 (1,908)	448 (1,912)	356 (1,912)
Thiostrepton*	Spectral Error	1	2	4	4	3
	Mass error	309 (971)	309 (971)	71 (978)	135 (973)	135 (973)

*Doubly charged ion.
†Based on number of formulas with mass errors in ppm less than or equal to the mass error of the correct formula.
^Second number shows rank when treating rifampicin as a mixture. See text for further explanation.
Number in ( ) are the total number of possible formulas consistent with search criteria.

7.5 K to 6.48% at R = 100 K for the singly and doubly charged ions, respectively. Despite these higher spectral errors, rankings of the doubly charged ion placed gramicidin S at either first or third place out of between 297 and 369 possible elemental compositions, almost identical to the observed rankings for the singly charged ion. Ranking based on mass error placed gramicidin S between 31st and 183rd place depending on resolution and charge state.

Cyclosporin A

Spectral errors increased monotonically from 0.63% at R = 7.5 K to 3.37% at R = 100 K. Ranking based on spectral error placed cyclosporin A in the top three depending on resolving power settings out of ~1100 possible elemental compositions. When ranked based on mass error, cyclosporin A was ranked between 134th and 270th place depending on resolution settings.

Vancomycin

Both singly and doubly charged ions were investigated. Spectral errors increased monotonically from 2.23% at R = 7.5 K to 12.81% at R = 100 K for the singly charged ion, but only from 3.0% at R = 7.5 K to 4.51% at R = 100 K for the doubly charged ions. Ranking based on spectral error placed singly charged vancomycin in sixth or eighth place for 7.5 to 60 K resolving power settings but 31st place at R = 100 K, out of ~1500 possible elemental compositions. When ranked based on mass error, singly charged vancomycin was ranked between 81st and 400th place depending on resolution setting. Due to the lower spectral error of the doubly charged ion at all resolving power settings including 100 K, doubly charged vancomycin was ranked no less than seventh out of 774 possible elemental compositions in all cases. When ranked based on mass error, doubly charged vancomycin was ranked between 234th and 314th place depending on resolving power settings, with no significant change in mass accuracy between the singly and doubly charged ions. In-source fragmentation of vancomycin, in which a sugar ($-C_7H_{14}NO_3$) group was lost from the molecular ion to give *m/z* 1305, was also examined because there was less background interference present compared to the parent (Figure 4d–f versus 4a–c). It has been reported that vancomycin and related glycopeptide antibiotics can undergo dimerization [25], and the formation of such dimers may be the source of the interference that we observed. Indeed, spectral errors for the fragment ion were significantly less than those for the molecular ion itself at all resolving power settings, increasing monotonically from 1.46% at R = 7.5 K to 5.99% at R = 100 K,

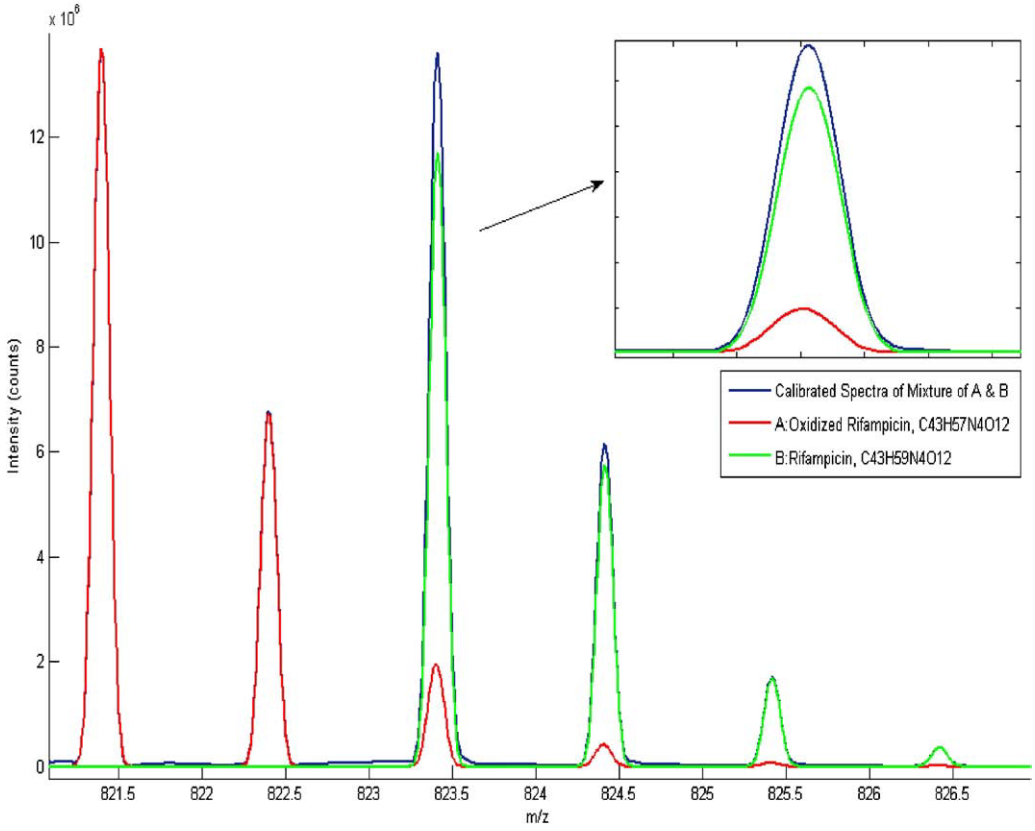

Figure 3. Spectrum of rifampicin at 15 K resolution. Rifampicin-quinone (oxidized rifampicin) is also present. When treating the spectrum as a mixture, spectral error is less than treating this as a pure compound.

ostensibly due to lack of dimer interference. Comparing the lower (7.5 K) and higher (100 K) resolution mass spectra of the parent ion (Figure 4a–c), one can see that the peak intensity for the M + 4 isotope cluster decreased from ~35% relative intensity to ~20%. This intensity decrease appears to become relatively larger for higher isotopes, e.g., M + 5 and M + 6. In contrast, for the fragment ion (Figure 4d–f), the decrease in M +

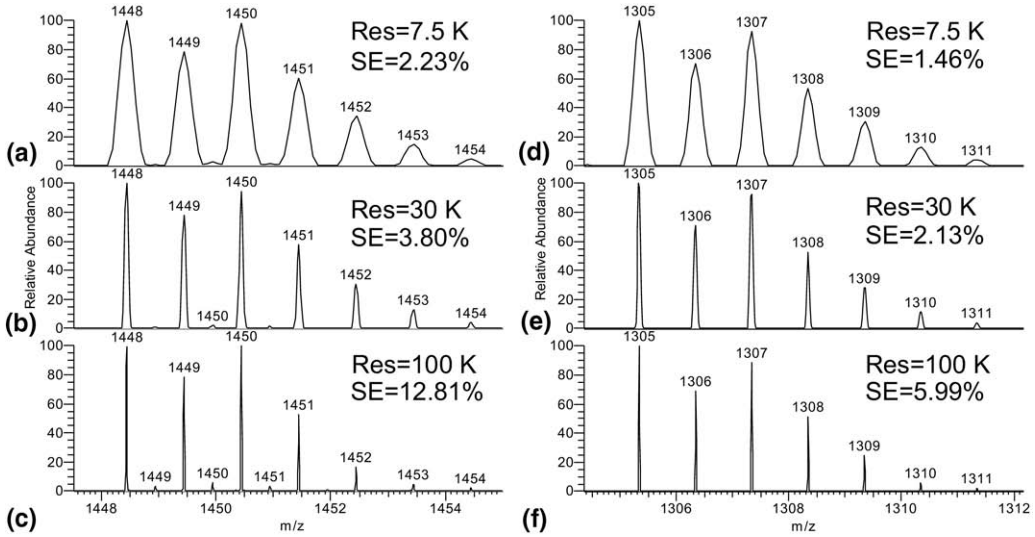

Figure 4. Spectra of singly charged vancomycin taken at 7.5 K (**a**), 30 K (**b**), and 100 K(**c**) resolution illustrating increasingly poorer spectral accuracy with increasing resolution. In contrast, spectra of a fragment of vancomycin taken at 7.5 K (**d**), 30 K (**e**), and 100 K (**f**) resolution have spectral errors less than 6%.

4 is less obvious and only for M + 5 and M + 6 are intensity deficits noticeable.

Thiostrepton

Both singly and doubly charged ions were investigated. Spectral errors were 2.38% at R = 7.5 K and increased to 5.64% at R = 100 K, and from 1.25% at R = 7.5 K to 6.48% at R = 100 K for the singly and doubly charged ions, respectively. Unlike gramicidin S, where rankings of singly and doubly charged ions were similar, ranking based on spectral error placed thiostrepton between second and seventh place for the singly charged ion (out of ~1900 possible elemental compositions), but systematically better between first and fourth for the doubly charged ion (out of ~975 possible elemental compositions) indicating that the improved spectral accuracy from the doubly charged ion could indeed improve elemental composition determination for this ion with higher *m/z*. Ranking based on mass error placed thiostrepton between 71st and 448th place depending on resolving power and charge state. For purposes of benchmarking, the computational efficiency of sCLIPS can be illustrated with thiostrepton. Even with the 2 ppm mass accuracy cutoff, nearly 2000 possible formula candidates had to be evaluated. sCLIPS, using Yergey's [17] algorithm for isotope calculation, took about 0.2 s on a 1 GHz laptop to evaluate one possible formula candidate, or 400 s to evaluate 2000 formula candidates.

Effect of Resolution on Spectral Accuracy

One general finding of our work is the observation of decreasing spectral accuracy with increasing resolving power as is evident in Table 2. With its richer isotope fine structure, thiostrepton illustrates the impact of resolving power on spectral accuracy. Figures 5a and b compare the calibrated and theoretical mass spectra for singly charged thiostrepton at resolving power of 15 K

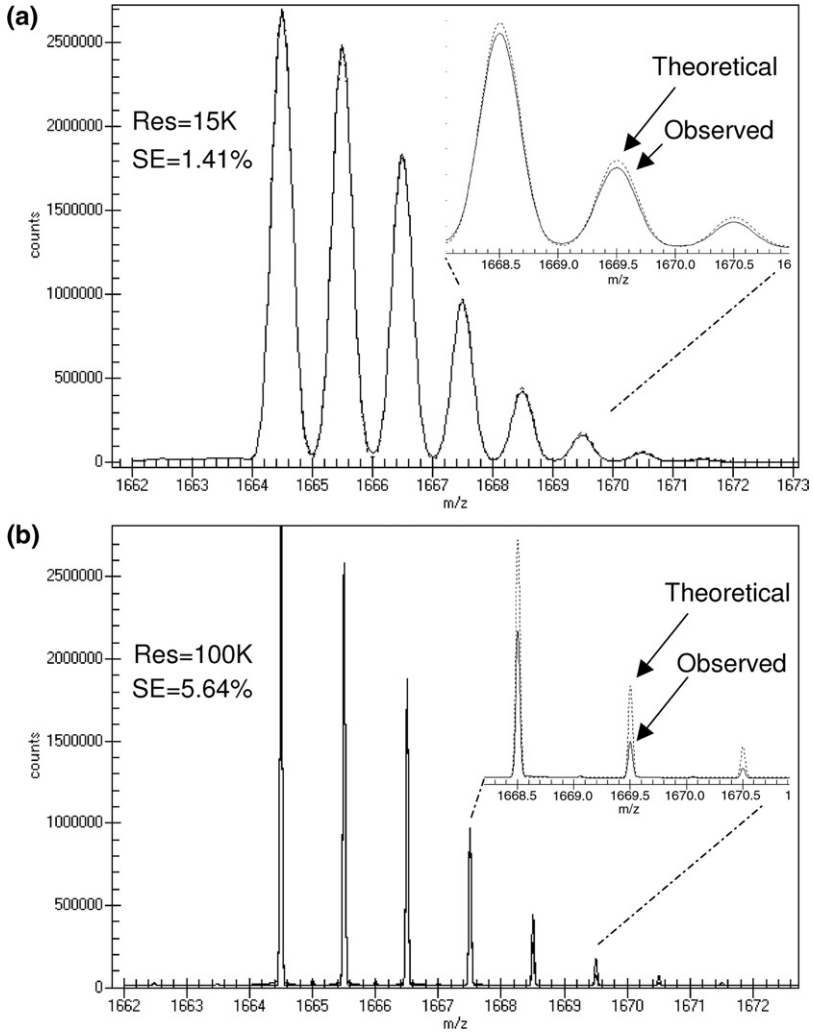

Figure 5. Spectrum of thiostrepton at 15 K (**a**) and 100 K (**b**). The inserts illustrate the intensity deficiency from the experimental data on ions below 1000,000 counts which is greater at 100 K than at 7.5 K.

and 100 K. At 15 K, the spectral error is 1.41%, (Figure 5a), while at 100 K the spectral error is 5.64%, (Figure 5b). As seen in the respective inserts, for isotopes M + 3 and higher at 100 K resolving power the experimentally observed intensity deviates by as much as 50% from the theoretical, while at 15 K this deviation is quite small. Abundance deficits in higher isotope clusters at higher resolving power have been observed in FT-ICR MS due to the effects of ion population on ion-cloud stability and transient duration [26]. As a result of columbic interactions, the ion cloud of a large isotope was claimed to perturb a smaller ion cloud representing a less abundant isotope peak leading to transient damping of the latter which manifests as an intensity deficit. However, this is not likely to be occurring in the Orbitrap as this phenomena requires longer time to manifest (>5 s) than the scan duration (1.52 s) required for 100 K resolving power measurements. Rather, the phenomena of isotopic beat patterns in which closely spaced yet unresolved frequencies cause constructive (or destructive) interference [27] and has been shown to effect measured relative abundances [28] may explain the abundance shortfalls for the smaller isotope peaks observed in our work (Makarov, personal communication). Indeed, our observation that the fluorophosphazine ion $[C_{30}H_{19}O_6N_3P_3F_{48}]^+$ (*m/z* 1521.9715) from the Agilent tune mix, which has limited fine structure due to its low oxygen and high fluorine content, does not display a large difference in spectral accuracy between low and high resolving powers as does thiostrepton and the other natural products with significant fine structure that we studied with similarly high mass (see Table 2), is consistent with these considerations although additional effort may be needed to more rigorously explain these phenomena. Although in practice we were not able to resolve the fine structure of the compounds studied here, as this would require resolving power of about at least 400 K, the ability of high-resolution mass spectrometry to resolve isotopic fine structure and the resulting impact on mass spectra has been investigated by earlier researchers [19, 29]. However, isotope fine structures contain important elemental composition information regardless of the resolving power, i.e., whether the isotope fine structure is visually resolved or not. Due to the matrix computational approach taken here for the spectral accuracy calculation, the lack of mass spectral resolution on these fine isotope structures does not necessarily lead to the loss of important mass spectral information contained in these fine structures, as the fine structural information is simply encoded in the continuously sampled vector arrays **t, r,** and **e** in eq 1–3. This is the fundamental reason why it is feasible to determine elemental compositions even on a single quadrupole GC/MS [30], provided that there are no spectral interferences arising from other unknown ions.

Closing Remarks and Summary

One objective of the work presented here was to determine the spectral error of molecular ions for 10 high molecular weight (>600 Da) natural products measured in the LTQ/Orbitrap at the five available resolving power settings. A general trend, with minor exceptions, was the observation that spectral error increased with increasing resolving power. A plot of the measured spectral error versus mass for each resolution setting is depicted in Figure 6. One may conclude that there is no strong dependence of spectral error on mass in the mass range investigated, consistent with the explanation that spectral error depends more on isotope fine structures and resolving power settings than on mass. Spectral error appears especially pronounced at higher resolving power in higher isotope clusters containing fine isotope structures and manifests with less-than-expected abundance from theoretical calculations.

A second objective of our work was to evaluate the performance of spectral error compared to mass error when ranking possible elemental compositions. The theoretical considerations of Kind and Fiehn demon-

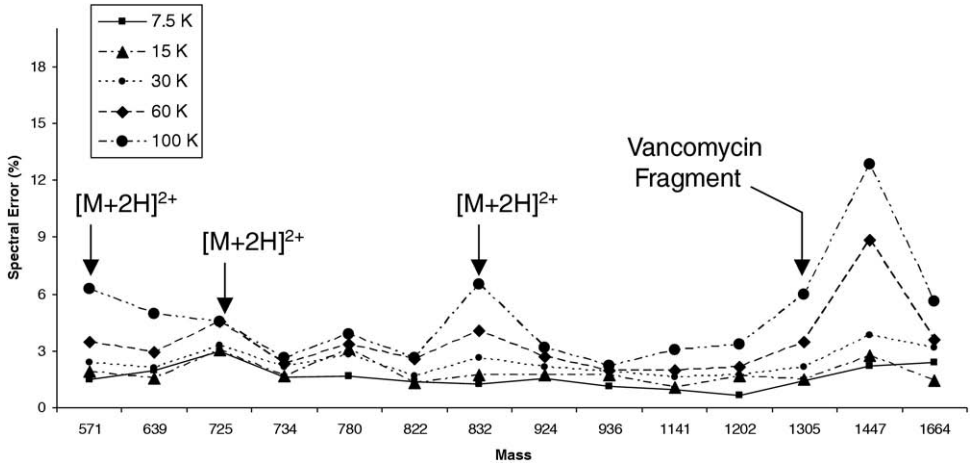

Figure 6. Plot of spectral error versus mass for each of the five different resolution settings.

strated that the use of isotopic abundance information (i.e., spectral error) can eliminate >95% of false candidates [1]. In our experimental work reported here, spectral accuracy proved to be a powerful discriminator for formulas that come within the 2 ppm mass tolerance range. For example, spectral accuracy ranked moxidectin above ~94% of possible candidates. However, for higher molecular weight compounds, such as cyclosporin A (1202 Da) or thiostrepton (1664 Da), spectral accuracy ranked the correct elemental composition within the top 0.5% of compounds, thereby eliminating >99% of false candidates. Not only does spectral accuracy help eliminate incorrect formula candidates, it also adds confidence to the formula candidates included for consideration. Thus, a higher spectral accuracy provides greater support to an elemental formula compared to the same elemental formula associated with a lower spectral accuracy. This reasoning, however, cannot be made for mass accuracy, since it is observed that a correct elemental formula can often have a lower mass accuracy compared with incorrect elemental formulas with higher mass accuracy. Additionally, spectral accuracy can provide additional information about the presence (or absence) of ion interferences from, e.g., dimer formation or oxidation, which can then be taken into consideration during formula identification. From the perspective of MS instrument design and operations, spectral accuracy can provide an objective criterion to evaluate how the instrument measurement reflects the true isotope response of a given ion, and can help with optimization of MS operating conditions and/or signal processing. The extensive studies on the behavior of ions in ion cyclotron resonance cells [27, 31] as well as work on signal processing strategies [32] may illustrate the types of studies that may lead to a better understanding of Orbitrap fundamentals and how to further improve its already impressive spectral accuracy characteristics.

Finally, it is important to recognize that spectral accuracy also has limitations. In our work, the correct elemental composition was ranked first only 23/65 (35%) of the time, even with our elemental constraints. If additional elements were included, this would result in an even lower percentage of top hits and would probably not be offset by further marginal improvement in mass accuracy, e.g., at a lower mass tolerance of 1 ppm. What is also clear from this study is that attaining better spectral accuracy for a given ion can improve the chance of arriving at a unique elemental composition. Achieving this goal may require a better understanding of how isotopes of different abundances behave inside the Orbitrap and how best to process the resulting analog and digital signals [27]. Without high enough spectral accuracy, additional selection criteria may need to be employed, such as some empirical chemical rules suggested by Kind and Fiehn to further eliminate possible formulas [1]. In the end, however, additional experimentation, including chemical derivatization, tandem mass spectrometry, hydrogen-deuterium exchange, among others, may be employed to help determine the correct elemental composition of a true unknown.

Acknowledgments

The authors thank their Wyeth colleagues Drs. Steven Unger and Chris Petucci for their critical review of the manuscript and Dr. Alexander Makarov for his insights and helpful suggestions. The authors also acknowledge the constructive reviewer comments of Dr. Alan Rockwood.

References

1. Kind, T.; Fiehn, O. Seven Golden Rules for Heuristic Filtering of Molecular Formulas Obtained by Accurate Mass Spectrometry. *BMC Bioinformatics* **2007,** *8,* 105.
2. Kind, T.; Fiehn, O. Metabolomic Database Annotations via Query of Elemental Compositions: Mass Accuracy is Insufficient Even at Less Than 1 ppm. *BMC Bioinformatics* **2006,** *7,* 234.
3. Stoll, N.; Schmidt, E.; Thurow, K. Isotope Pattern Evaluation for the Reduction of Elemental Compositions Assigned to High-Resolution Mass Spectral Data from Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2006,** *17,* 1692–1699.
4. Grange, A. H.; Donnelly, J. R.; Sovocool, G. W.; Brumley, W. C. Determination of Elemental Compositions from Mass Peak Profiles of the Molecular Ion (M) and the M + 1 and M + 2 Ions. *Anal. Chem.* **1996,** *68,* 553–560.
5. Zweigenbaum, J. A.; Thurman, E. M.; Ferrer, I. (Agilent Technologies, Inc., USA). Application: Determination of Chemical Empirical Formulas of Unknown Compounds Using Accurate Ion Mass Measurement of All Isotopes. USA **2007,** 1–16.
6. Amirav, A.; Alon, T. (Israel). Application: Mass Spectrometric Based Method for Sample Identification. USA **2006,** 1–11.
7. Hobby, K.; Gallagher, R. T.; Caldwell, P.; Wilson, I. D. A New Approach to Aid the Characterization and Identification of Metabolites of a Model Drug; Partial Isotope Enrichment Combined with Novel Formula Elucidation Software. *Rapid Commun. Mass Spectrom.* **2009,** *23,* 219–227.
8. Wang, Y.; Gu, M. *PCT Int. Appl.* (Cerno Bioscience LLC, USA). WO 2007; Vol. U.S. Patent 6,983,213, p 35.
9. Koester, C. (Bruker Daltonik GmbH, Germany). Application: GB GB, 1999, p 15.
10. Bristow, T.; Constantine, J.; Harrison, M.; Cavoit, F. Performance Optimization of a New-Generation Orthogonal-Acceleration Quadrupole-Time-of-Flight Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **2008,** *22,* 1213–1222.
11. Fernandez-de-Cossio, J.; Gonzalez, L. J.; Satomi, Y.; Betancourt, L.; Ramos, Y.; Huerta, V.; Besada, V.; Padron, G.; Minamino, N.; Takao, T. Automated Interpretation of Mass Spectra of Complex Mixtures by Matching of Isotope Peak Distributions. *Rapid Commun. Mass Spectrom.* **2004,** *18,* 2465–2472.
12. Goodner, K. L.; Milgram, K. E.; Williams, K. R.; Watson, C. H.; Eyler, J. R. Quantitation of Ion Abundances in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1998,** *9,* 1204–1212.
13. Gu, M.; Wang, Y.; Zhao, X. G.; Gu, Z. Accurate Mass Filtering of Ion Chromatograms for Metabolite Identification Using a Unit Mass Resolution Liquid Chromatography/Mass Spectrometry System. *Rapid Commun. Mass Spectrom.* **2006,** *20,* 764–770.
14. Kuehl, D.; Wang, Y. The Role of Spectral Accuracy in Mass Spectrometry. *LCGC North Am.* **2007,** *10,* 12–16.
15. Wang, Y. (Cerno Bioscience LLC, USA). Computational Method and System for Mass Spectral Analysis. WO 2004/097581 A2, **2004,** 1–63.
16. Wang, Y.; Kuehl, D. (Cerno Bioscience LLC, USA). A Self Calibration Approach for Mass Spectrometry. WO 2008/151153 A1, **2008,** 1–28.
17. Yergey, J. A. A General Approach to Calculating Isotopic Distributions for Mass Spectrometry. *Int. J. Mass Spectrom. Ion. Phys.* **1983,** *52,* 337–349.
18. Rockwood, A. L.; Van Orden, S. L.; Smith, R. D. Rapid Calculation of Isotope Distributions. *Anal. Chem.* **1995,** *67,* 2699–2704.
19. Rockwood, A. L.; Van Orden, S. L.; Smith, R. D. Ultrahigh Resolution Isotope Distribution Calculations. *Rapid Commun. Mass Spectrom.* **1996,** *10,* 54–59.
20. Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal. Chem.* **2006,** *78,* 2113–2120.
21. Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrom. Rev.* **2008,** *27,* 661–699.
22. Olsen, J. V.; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts Per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-Trap. *Mol. Cell. Proteom.* **2005,** *4,* 2010–2021.

23. Erve, J. C. L.; DeMaio, W.; Talaat, R. E. Rapid Metabolite Identification with Sub Parts-Per-Million Mass Accuracy from Biological Matrices by Direct Infusion Nanoelectrospray Ionization After Clean-Up on a Zip-Tip and LTQ/Orbitrap Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2008,** *22,* 3015–3026.
24. Bolt, H. M.; Remmer, H. Implication of Rifampicin-Quinone in the Irreversible Binding of Rifampicin to Macromolecules. *Xenobiotica* **1976,** *6,* 21–32.
25. Mackay, J. P.; Gerhard, U.; Beauregard, D. A.; Maplestone, R. A.; Williams, D. H. Dissection of the Contributions toward Dimerization of Glycopeptide Antibiotics. *J. Am. Chem. Soc.* **1994,** *116,* 4573–4580.
26. Bresson, J. A.; Anderson, G. A.; Bruce, J. E.; Smith, R. D. Improved Isotopic Abundance Measurements for High Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectra via Time-Domain Data Extraction. *J. Am. Soc. Mass Spectrom.* **1998,** *9,* 799–804.
27. Hofstadler, S. A.; Bruce, J. E.; Rockwood, A. L.; Anderson, G. A.; Winger, B. E.; Smith, R. D. Isotopic Beat Patterns in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Implications for High Resolution Mass Measurements of Large Biopolymers. *Int. J. Mass Spectrom. Ion Processes* **1994,** *132,* 109–127.
28. Easterling, M. L.; Amster, I. J.; van Rooij, G. J.; Heeren, R. M. A. Isotope Beating Effects in the Analysis of Polymer Distributions by Fourier Transform Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1999,** *10,* 1074–1082.
29. Werlen, R. C. Effect of Resolution on the Shape of Mass Spectra of Proteins: Some Theoretical Considerations. *Rapid Commun. Mass Spectrom.* **1994,** *8,* 976–980.
30. Wang, Y.; Prest, H. Accurate Mass Measurement on Real Chromatographic Time Scale with a Single Quadrupole Mass Spectrometer. *Chromatography* **2006,** *27,* 135–140.
31. Gordon, E. F.; Muddiman, D. C. Impact of Ion Cloud Densities on the Measurement of Relative Ion Abundances in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Experimental Observations of Coulombically Induced Cyclotron Radius Perturbations and Ion Cloud Dephasing Rates. *J. Mass Spectrom.* **2001,** *36,* 195–203.
32. Farrar, T. C.; Elling, J. W.; Krahling, M. D. Application of Linear Prediction to Fourier Transform Ion Cyclotron Resonance Signals for Accurate Relative Ion Abundance Measurements. *Anal. Chem.* **1992,** *64,* 2770–2774.