April 2008

Application Note Number 104 Cerno Bioscience

Cerno Application Note

Extending the Limits of Mass Spectrometry

Peak Shape Calibration and Formula Determination on an AccuTOF DART System

Yongdong Wang[I]

The AccuTOF DART system[1] combines an innovative atmospheric pressure ionization technique with a high resolution Time-Of-Flight (TOF) mass spectrometer to provide a powerful tool for unknown compound identification. Despite the relatively high mass accuracy of the system, formula search based on mass accuracy alone rarely produces a unique formula. MassWorks sCLIPS peak shape calibration can substantially improve the instruments ability to uniquely identify unknowns and perform well in the presence of interferences without the need for additional calibration.

Introduction

Operating at a moderately high resolving power of 5,000 or more, Time-Of-Flight (TOF) MS systems are capable of achieving a high mass accuracy of 5ppm or better with careful experimentation and frequent calibration, including LockSpray[2] or DualSpray[3] where one or more internal standards are introduced into the mass spectrometer alternatively or simultaneously to the analyte. The invention of the DART ionization technique dramatically simplifies the introduction of these internal standards due to its versatility in accommodating different sample types, making it possible to perform accurate mass measurement in real time or on the spot with little or no sample preparation.[4]

While a high mass accuracy of 5ppm is considered necessary for the determination of elemental compositions, it has been shown that it is not sufficient for unknown formula determination. Even at sub-ppm mass accuracy, there are many chemically feasible formula candidates to consider, especially for ions with m/z larger than 400Da.[5-6]

Fortunately, there is other important information that can be brought to bear in formula determination, beyond the single point accurate mass measurement of the monoisotope peak. For a given ion, the whole isotope profile, whether fully resolved or not with respect to the individual isotopes or isobars, contains critical information about the fine structure of the underlying isotope distribution which is a function of the elemental composition. Simple measurement of the relative intensities of the isotope peaks (M,M+1,…) can be a useful additional metric for paring down the formula list, particularly for Br-, Cl-, or S- containing compounds with their pronounced isotope features.[5] Other approaches include simple computer modeling

[I] Cerno Bioscience, Danbury, CT 06810, USA

as well as more elaborate ones that assume a Gaussian or other peak shape function.[7,8,9] However, these approaches only attempt at rough approximations to the true isotope profile because the actual measured peak shape is either unknown or not available for use, resulting in modeling errors as large as a few percent, a level of error larger than the typically subtle difference from one formula to another, and largely limiting the usefulness of isotope profile information. In a TOF MS system where the ion arrival time distribution (peak shape) has a characteristically long tail, it becomes even more difficult (if not impossible) to model the peak shape function with enough accuracy sufficient for formula differentiation.

Using a real example from an AccuTOF DART system, this application note will highlight the importance of TOF MS peak shape calibration in formula determination and demonstrate the feasibility of formula determination using the MassWorks sCLIPS (self-Calibrating Line-shape Isotope Profile Search) without the use of either internal or external calibration standards.

Experimental

An unknown compound with its nominal m/z at 319Da was measured on an AccuTOF DART system along with PEG ions for possible use as internal calibration standards. An averaged mass spectrum is then exported to JCAMP data format for importing into MassWorks software (Version 2.0, released April 22, 2008, Cerno Bioscience, Danbury, CT, USA) for processing and formula determination. As is typical and highly desirable, no ion counting threshold has been applied to insure the acquisition of the true raw profile mode MS data.

Results

Figure 1 (top graph) shows the complete mass spectrum imported into MassWorks and Figure 1 (bottom graph) shows the section for the unknown ion of interest with its nominal mass at 319Da. Although there are available PEG ions that could have been used as calibration standards, we decide to ignore these calibration ions in this application note so as to demonstrate the new self-calibration feature (sCLIPS) available in MassWorks v2.0.

Figure 2 shows the general data processing flow for MassWorks sCLIPS. Compared to the CLIPS for lower resolution systems such as a single quad,[10,11] a key distinction in sCLIPS is the use of monoisotope peak of the unknown ion itself to perform a peak-shape-only calibration that does not involve m/z axis. The rationale behind this approach are three fold:

1. The monoisotope (calibration) peak of the unknown ion is as close as possible in m/z to the rest of the isotopes

2. The monoisotope (calibration) peak of the unknown ion is measured as close as possible in time to the rest of the isotopes

3. Where sCLIPS is applicable, i.e., on higher resolution MS systems with a monoisotope peak well separated from other isotopes such as M+1 etc., the m/z is typically reasonably accurate to start with and an additional m/z calibration could be avoided.

Once the peak-shape-only calibration function has been obtained, which effectively transforms the actual peak shape function into an ideal (known, symmetrical, and mathematically definable) peak shape function, the same calibration function is then applied to the complete isotope profile including the rest of the isotopes to obtain a calibrated isotope profile. Subject to the random noise and other measurement (such as detector saturation) or numerical errors, this calibrated isotope profile should be essentially the same as the theoretical mass spectrum calculated for the correct formula with the same ideal peak shape function, after proper fitting or normalization. In other words, the Spectral Accuracy as define below should approach 100.00%,

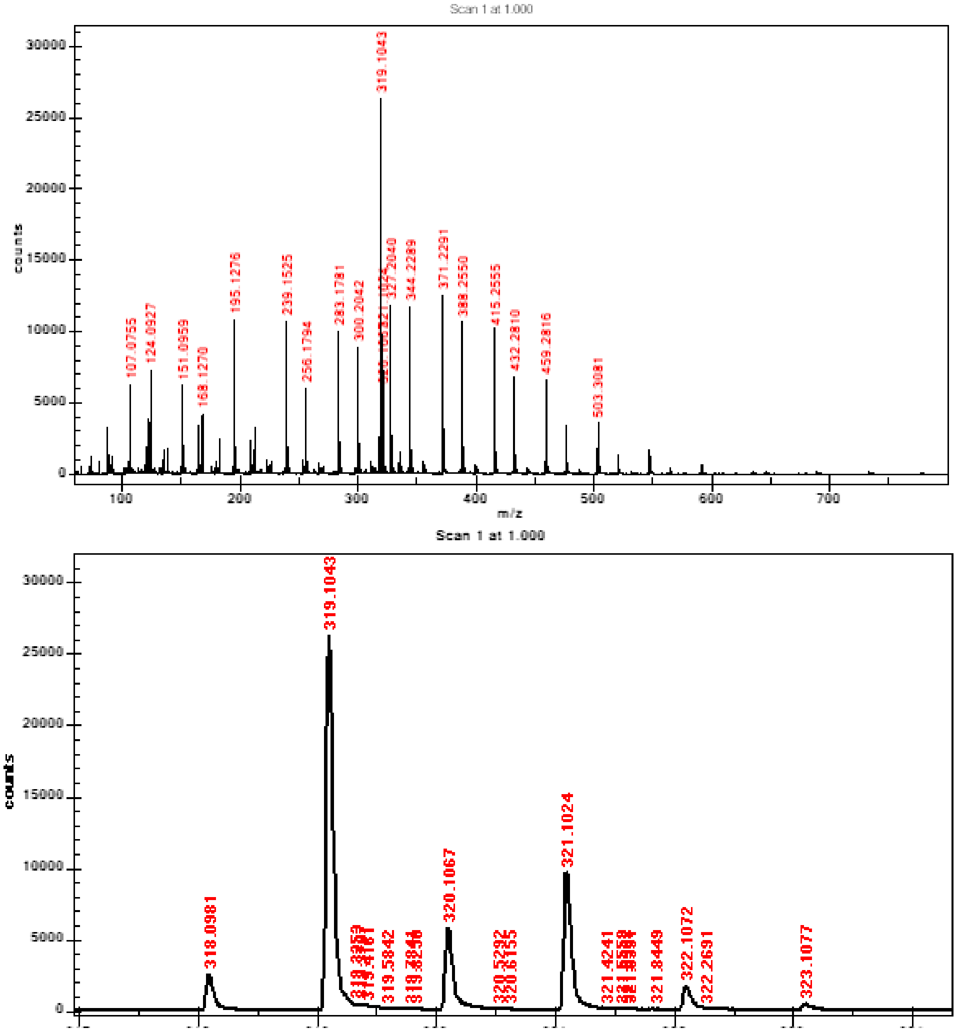
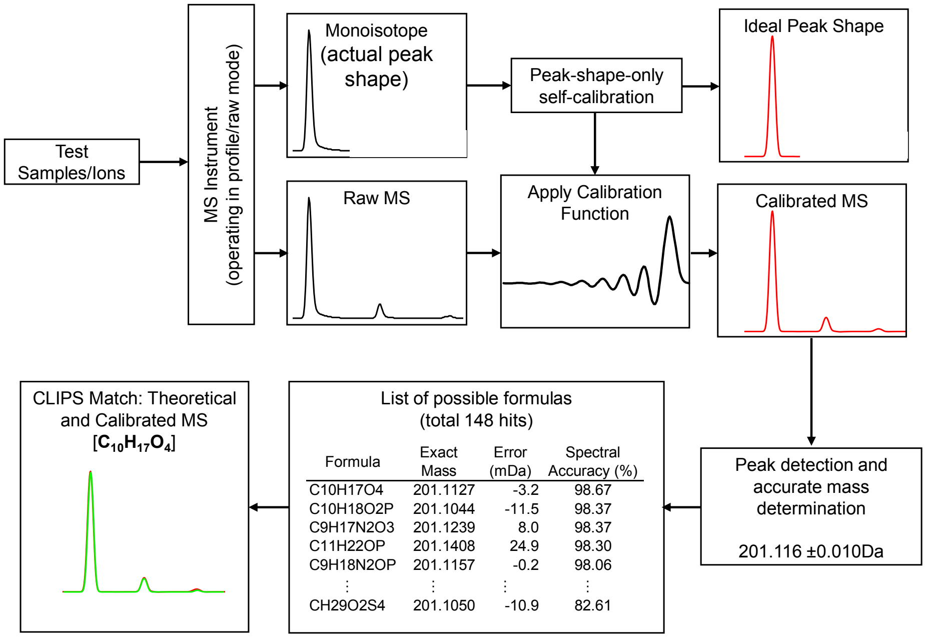


Figure 1. An averaged AccuTOF DART MS spectrum (top – complete MS spectrum, bottom – the section corresponding to the unknown ion at 319Da).



Formula	Exact Mass	Error (mDa)	Spectral Accuracy (%)
C10H17O4	201.1127	-3.2	98.67
C10H18O2P	201.1044	-11.5	98.37
C9H17N2O3	201.1239	8.0	98.37
C11H22OP	201.1408	24.9	98.30
C9H18N2OP	201.1157	-0.2	98.06
⋮	⋮	⋮	⋮
CH29O2S4	201.1050	-10.9	82.61

Figure 2. The general flow of MassWorks sCLIPS process for formula determination on higher resolution MS systems.

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

where **e** is the fitting residual vector, **r** is the calibrated isotope profile vector, and $\|\cdot\|_2$ represents the 2-norm (or square root of the sums of squares of all elements) of a vector. This Spectral Accuracy metric will be used to evaluate all possible formulas whose exact monoisotope masses come within a mass tolerance window of the reported accurate mass obtained off the actual monoisotope peak and whose elemental compositions satisfy the given chemistry constraints. The one formula with the highest Spectral Accuracy is most likely the correct formula for the unknown ion of interest.

Figure 3 shows the peak shape functions before and after sCLIPS calibration for just the monoisotope peak (top graph) and the complete isotope profile (bottom graph). The ideal peak shape after the calibration clearly compensates for the long tail in the TOF peak shape without any apparent loss of resolution. It should be noted here that there exists a potential mass spectral interference at 318Da coming most likely from the M-H ion, which will be taken into consideration as well during the sCLIPS formula determination.

Table 1 shows the sCLIPS formula determination parameters and results. Notice that the possible presence of M-H form has been entered into the mixture search, a new feature in MassWorks V2.0 capable of handling up to three mutually interfering ions. Correspondingly, the Spectral Accuracy metric is now calculated from -1.5 to 4.5Da, all relative to the reported accurate mass of 319.1043Da, so as to cover the full isotope profile with the possible interference ion. A total of 70 different formulas are found to satisfy the ±10mDa mass tolerance window and other chemistry constraints listed in Table 1. The correct formula, $C_{17}H_{20}N_2SCl$, indeed shows up as the formula with the highest Spectral Accuracy of 97.27%.

As an additional feature available in MassWorks v2.0, the user can now click through each and every formula on the list and explore interactively the mass spectral match between the calibrated and the theoretical mass spectrum calculated for a given formula. Figure 4 shows the spectral overlay between the calibrated mass spectrum and its theoretical version for the top hit (correct) formula, $C_{17}H_{20}N_2SCl$. Excellent spectral match is observed for the complete isotope profile including contributions from the M-H ion. In fact, the software additionally reports that the relative concentration for the M and M-H is at 91% and 9%. Comparing to the mass spectral difference between the raw and calibrated mass spectrum shown in Figure 3, the critical role played by the peak shape calibration becomes clear in achieving such a high level of Spectral Accuracy and helping to achieve unique formula determination. It is the added bonus that this could be accomplished without the use of external or internal calibration standards.

Figure 5 shows the spectral match between the calibrated mass spectrum and that calculated from the 2nd hit formula ($C_{15}H_{24}O_3SCl$) with only 15 carbon elements and a Spectral Accuracy of 96.52%, which shows a slight but apparently significant lack of fit on the 320Da peak, highlighting again the importance of exact isotope profile modeling with known peak shape function. It is interesting to note that this formula would have been the top hit, had element nitrogen (N) been left out of the list of possible elements in Table 1, suggesting that this interactive mass spectral overlay can also be used as a useful tool to decide on the possible elements to include in formula determination.

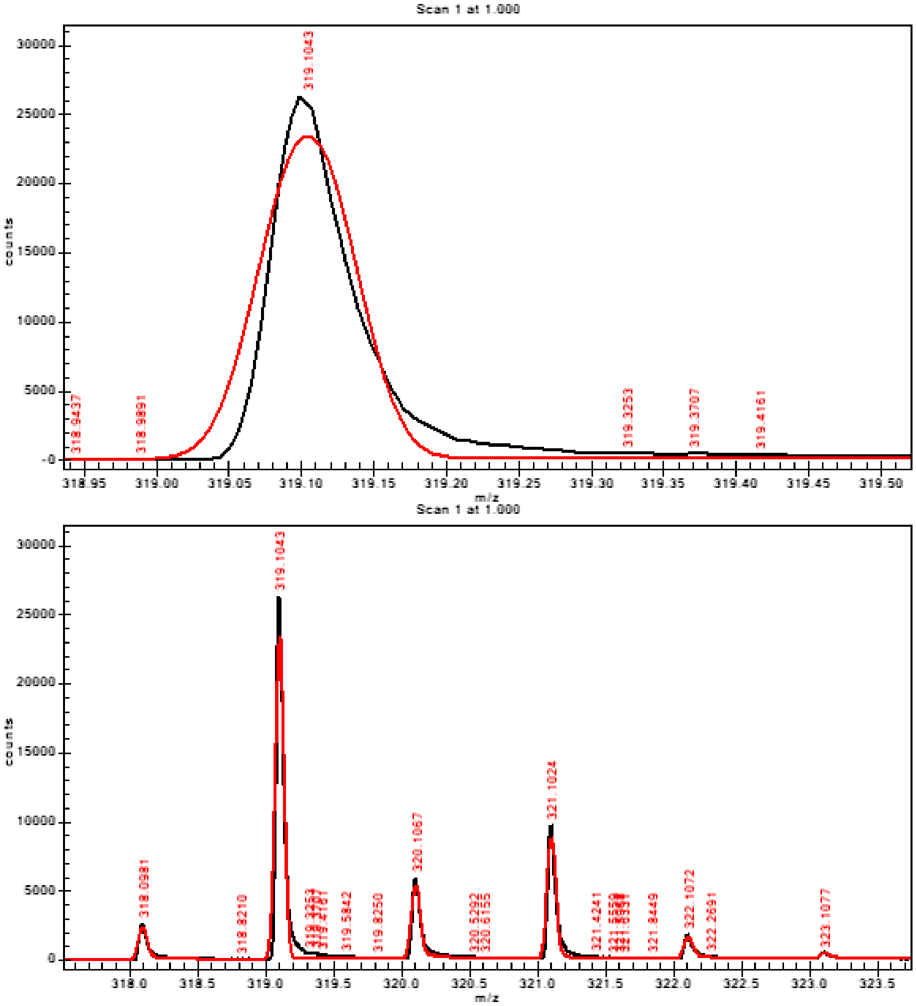


Figure 3. The raw (black) and the peak-shape-calibrated (red) mass spectrum for the ion of interest (top – monoisotope, bottom – complete isotope profile including a possible M-H).

Accurate Mass	319.1043	
Charge	1	
Mass Tolerance (mDa)	10	
Electron State	Even	
DBE (Min)	-1	
DBE (Max)	50	
Profile Mass Start (Da)	-1.5	
Profile Mass End (Da)	4.5	
Element	Minimum	Maximum
C	0	50
H	0	100
N	0	10
O	0	10
S	0	5
Cl	0	2
Mixture Search	-H	

Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy
1	$C_{17}H_{20}N_2SCl$	319.1036	-0.7279	-2.2812	97.2665
2	$C_{15}H_{24}O_3SCl$	319.1135	9.168	28.7305	96.5226
3	$C_{18}H_{20}O_3Cl$	319.1101	5.7972	18.1672	96.1848
4	$C_{20}H_{16}N_2Cl$	319.1002	-4.0988	-12.8446	95.666
5	$C_{15}H_{16}N_4O_2Cl$	319.0962	-8.1215	-25.451	95.3132
6	$C_{12}H_{20}N_4O_2SCl$	319.0995	-4.7507	-14.8876	95.3055
7	$C_{11}H_{20}N_6OSCl$	319.1108	6.4827	20.3153	94.9893
8	$C_{14}H_{16}N_6OCl$	319.1074	3.1119	9.7519	94.9518
9	$C_{14}H_{24}N_2S_2Cl$	319.1069	2.6429	8.2822	94.5673
10	$C_{15}H_{24}OS_2Cl$	319.0957	-8.5905	-26.9207	94.4946
11	$C_{14}H_{20}O_6Cl$	319.0948	-9.4589	-29.6421	94.4085
12	$C_{13}H_{20}N_2O_5Cl$	319.1061	1.7745	5.5608	94.0019
13	$C_{11}H_{24}O_6SCl$	319.0982	-6.0881	-19.0787	93.5715
14	$C_8H_{16}N_{10}SCl$	319.0969	-7.436	-23.3028	93.4716
15	$C_{10}H_{24}N_2O_5SCl$	319.1094	5.1453	16.1242	93.2443
16	$C_9H_{24}N_4O_2S_2Cl$	319.1029	-1.3799	-4.3242	92.4923

Table I. sCLIPS Formula Determination Parameters and Results. Only the top 16 of 70 formulas are shown.

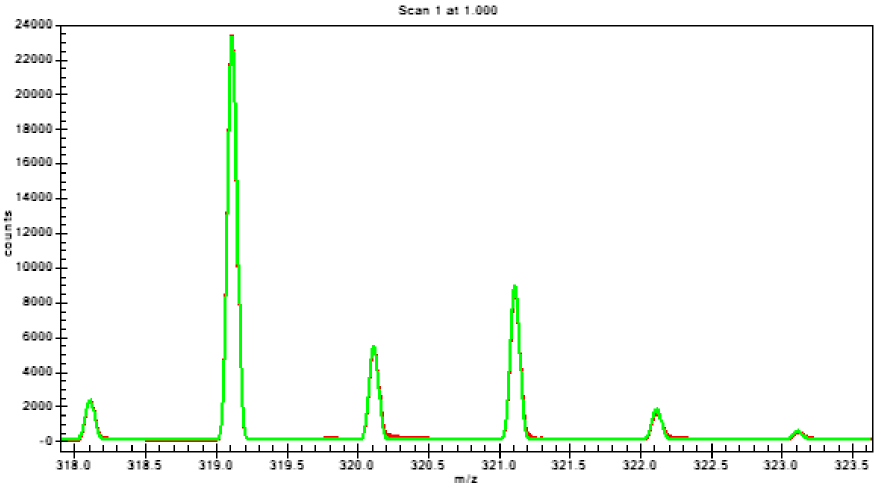


Figure 4. Overlay of calibrated (red) and theoretical mass spectrum (green) for the top hit formula ($C_{17}H_{20}N_2SCl$) with Spectral Accuracy= 97.27%.

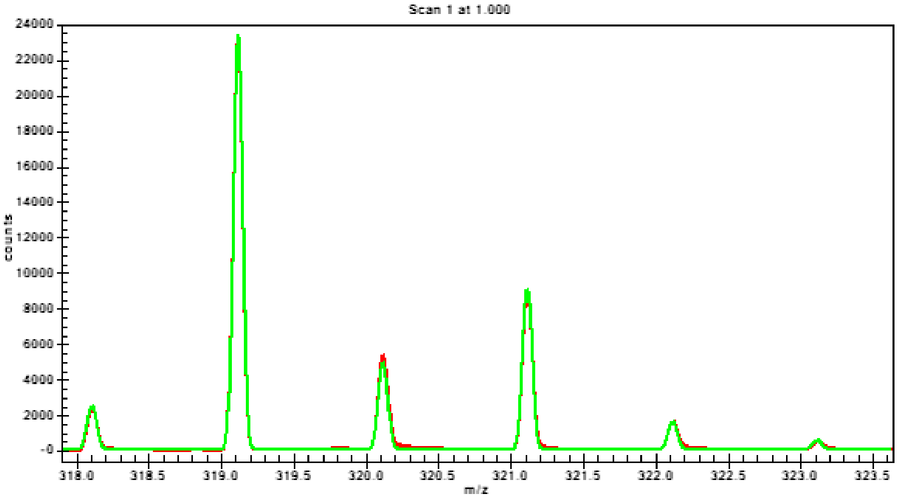


Figure 5. Overlay of calibrated (red) and theoretical mass spectrum (green) for the 2nd hit formula ($C_{15}H_{24}O_3SCl$, no N present) with Spectral Accuracy= 96.52%.

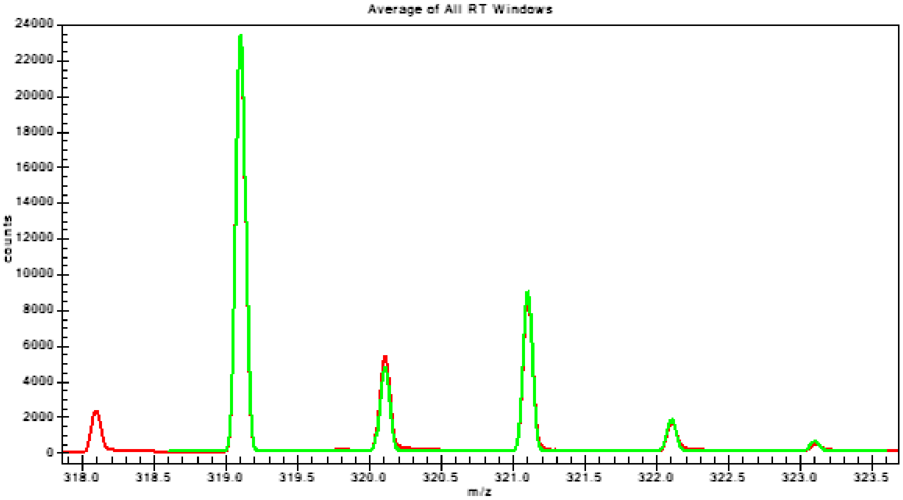


Figure 6. Overlay of calibrated (red) and theoretical mass spectrum (green) for the top hit formula ($C_{17}H_{20}N_2SCl$) without accounting for M-H with Spectral Accuracy= 96.36%.

Similarly, Figure 6 shows the sCLIPS spectral overlay without considering the M-H ion as a possible interference in a more limited relative mass spectral window of -0.5 to 4.5Da. Again, a compromise in Spectral Accuracy is observed (96.36% vs 97.27% in Table 1 or Figure 4 with M-H included). This comprise in the mass spectral fit can also be graphically observed on the 320Da ion where there is a clear lack of fit to the calibrated mass spectrum.

Conclusion

MassWorks sCLIPS requires no external or internal calibration standard to perform the critical peak shape calibration by taking advantage of the monoisotope peak of the unknown ion itself. The peak shape calibration combined with the reasonably high mass accuracy already available on an AccuTOF DART system makes achieving unique formula determination within reach while further increasing the ease of use. The added power of mixture search and interactive spectral overlay allows the user to examine the fit between the calibrated and theoretically calculated mass spectrum and also provides valuable insights with respect to the presence of any mass spectral interferences and the inclusion or exclusion of certain elements in the formula determination process.

Acknowledgement

We like to thank Dr. Robert B. Cody from JEOL USA, Inc. for providing the data sets used in this application note and for the interesting discussions.

References

[1] Jeol Ltd. Tokyo, Japan

[2] http://www.waters.com/waters/nav.htm?locale=en_US&cid=1000396

[3] Edgar Naegele, Pharma Asia, April 25, **2008**, http://www.pharmaasia.com/article-5582-statisticevaluationofmassaccuracymeasurementsbyesitofwithasampleofdegradationproductsfromtheantibiot-Asia.html.

[4] Robert B. Cody et al, Anal. Chem. **2005**, **77**, 2297-2302.

[5] Kind, T. *BMC Bioinformatics* **2006**, 7, 234.

[6] Don Kuehl, Am. Lab. Online, January **2008**, http://americanlaboratory.texterity.com/al-online/200801ol/?pg=18

[7] Evans, J. E.; Jurinski, N. B. Anal. Chem. **1975**, 47, 961-963b

[8] Tenhosaari, A. Org. Mass Spectrom. **1988**, 23, 236-239.

[9] Do Lago, C. L.; Kascheres, C. Comput. Chem. **1991**, 15, 149-155.

[10] Yongdong Wang, US Patent 6,983,213, filed October 20, 2003.

[11] Yongdong Wang, Cerno Application Note Number 103, December **2006**, http://www.cernobioscience.com/resources/ApplicationNotes/Agilent_LC-MS_Single_Quad_Infusion.pdf