Further Improving Elemental Composition Determination with FT ICR MS

Scott Pennino[1], Fenghe Qiu[1], and Yongdong Wang[2]

[1]Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, CT; [2]Cerno Bioscience, Danbury, CT

Introduction

Even with the unparallel resolving power of an FT ICR MS reaching as high as 1,000,000 and its highly accurate mass measurement capability with sub-ppm mass error, unique formula determination remains a daunting task. This is especially true for ions with m/z values larger than 400Da or for the identification of truly unknown compounds encountered in applications such as natural product research or pharmaceutical impurity identification. While the use of isotope distribution does provide additional useful information (Ref 1), it proves to be quite a challenge to reliably measure the typically small (less than 1%) difference in isotope distribution so as to differentiate among similar formula candidates. This paper presents an approach for exact isotope modeling to further improve formula ID on FT ICR MS and to explore a set of experimental conditions where unique formula determination of truly unknown ions becomes feasible.

Methods

A few commercially available organic compounds are selected as unknown ions for the accurate mass measurement and formula determination test. These compounds will be measured at different concentrations and resolving powers. Accurate masses and their standard deviations will be reported to form a basis for the selection of mass tolerance window. All commonly encountered elements would be included to generate a list of possible formulas as if these were truly unknown organic compounds. After the peak shape calibration, each formula candidate will be evaluated for its Spectral Accuracy. The absolute level of Spectral Accuracy and its relative ranking among all formula candidates serves to indicate the likelihood of unique formula determination.

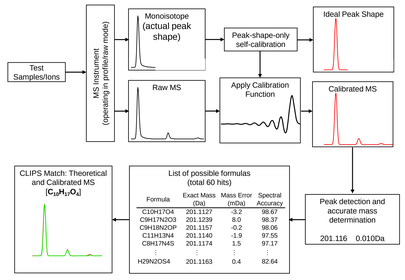
Experiments and Data Analysis

➢Sample information: A solution containing Loperamide (1 ng/μL), Reserpine (10 ng/μL), and Erythromycin (10 ng/μL) is prepared with 0.1% formic acid in 1:1 ACN:H_2O as the test sample for LC infusion into the FT ICR MS system at a flow rate of 2μL/min. An 1:2 dilution is needed in order for the highest signal ion to stay within the linear dynamic range.

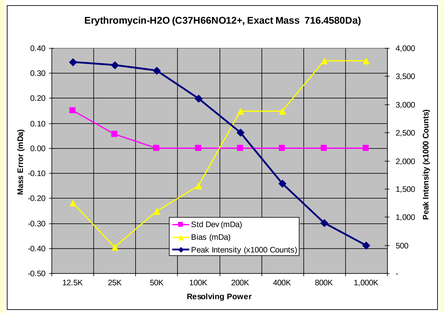
➢MS conditions: A Thermo LTQ/FT ICR MS Ultra is calibrated once according to the standard calibration procedure at the frequently used resolving power of 100K. The data acquisition is performed in profile mode at 8 different resolving powers: 12.5K, 25K, 50K, 100K, 200K, 400K, 750K, and 1,000K, all for a duration of 2.0min sample infusion, resulting in different number of scans at different resolving powers. The 1st 9 scans in each acquisition will be used to determine the standard deviations of reported accurate masses while the average mass spectrum of all available scans in each acquisition will be used for Spectral Accuracy formula ID described below. The Automatic Gain Control (AGC) is always on to automatically control the number of ions through the change of ion injection times.

➢sCLIPS in MassWorks performs a peak-shape-only calibration to transform the actual peak shape function into a known mathematical function by using the measured monoisotope peak itself as a calibration standard (Ref 2). When applied to the whole isotope profile of an unknown ion, it is transformed into a calibrated isotope profile with the same known peak shape function, which is then used in the calculation of the theoretical mass spectrum for any given formula candidate to achieve exact isotope modeling with high Spectral Accuracy.

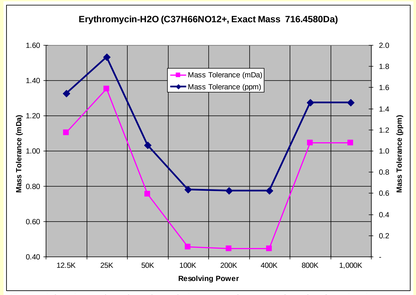
MassWorks sCLIPS Formula ID



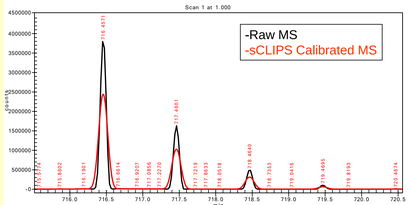
Results and Discussion



Above graph shows that the peak intensity linearly decreases with the increase in resolving power, due to the automatic gain control (AGC) implemented in this system. Consistent with the theoretical predictions (Ref 3), the mass measurement precision (standard deviation) does improve linearly as the resolving power increases, especially when the peak intensity is not significantly decreased as a result of the gain control, i.e., when the resolving power is less than 50K. It is interesting to note that there is a systematic mass measurement bias which reaches a minimum between 100K and 200K resolving power, reflecting the fact that the system has been calibrated at 100K resolving power.



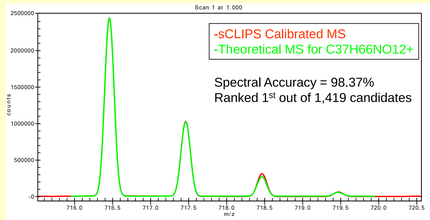
Considering both the bias and standard deviation, a mass tolerance window of 3 times the sum of the absolute mass bias and standard deviation will be used for the formula determination in order to insure ~99% probability of including the correct formula within the mass window. This translates to ~2ppm mass tolerance applicable to all resolving powers. It should be mentioned that at 100K-400K resolving power, there seems to be a sweat region where a smaller mass tolerance of 0.7ppm could be used to reduce the number of possible formula candidates.



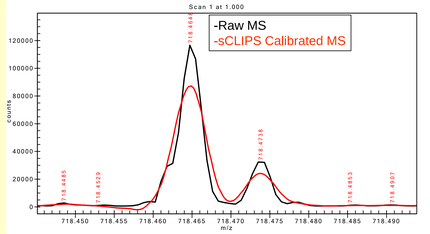
The above graph shows the mass spectral data at 12.5K resolving power before and after MassWorks sCLIPS peak-shape-only calibration. There is no adjustment in the mass axis but the peak shape after the calibration is now a known mathematical function. Using the accurate mass reported and the formula determination parameters listed in the table below for totally unknown organic compound ID, a total of **1,419** possible formulas are found, of which the correct formula is ranked as the 1st hit based on Spectral Accuracy.

Accurate Mass	716.4571	
Charge	1	
Mass Tolerance (PPM)	2	
Electron State	Even	
Double Bond Equivalent Minimum	-30	
Double Bond Equivalent Minimum	50	
Profile Mass Start (Da)	-0.5	
Profile Mass End (Da)	3.5	
Element	Minimum	Maximum
C	0	50
H	0	100
N	0	20
O	0	20
S	0	5
Cl	0	1
F	0	5
P	0	1
Na	0	1

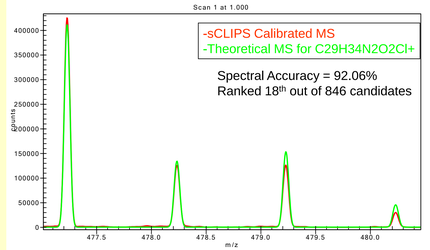
Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	RMSE	DBE
1	C37H66NO12	716.4574	0.20	1.90	98.37	7,149	5.5
2	C35H67N3O3SF4P	716.4566	-0.50	0.82	98.29	7,510	2.5
3	C33H64N9O3SFP	716.4563	-0.80	0.48	98.26	7,652	6.5
4	C35H69N3O5SFPNa	716.4566	-0.50	0.88	98.11	8,301	2.5
5	C34H63N5O3SF5	716.4561	-1.00	0.11	98.08	8,414	3.5
6	C32H60N11O3SF2	716.4558	-1.30	-0.23	98.05	8,548	7.5
7	C34H65N5O5SF2Na	716.4561	-1.00	0.16	97.96	8,965	3.5
...	...	...	...	...	...	...	...
1419	CH81N19O7S5ClFP	716.4547	-2.40	1.27	54.50	199,575	-29.5



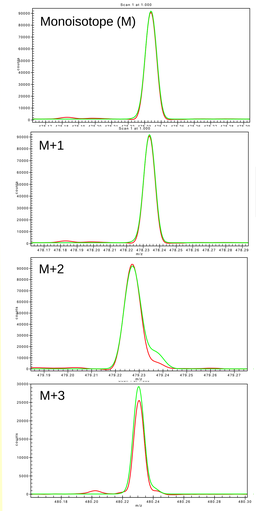
The spectral overlay between the sCLIPS calibrated mass spectrum and the theoretically calculated version for the correct formula using the same peak shape function shows a great match between the two, except for a minor lack of fit on the M+2 ion. When acquired at 400K resolving power, it is revealed that there is an unresolved interference ion (see below) at 718.4738Da.



While the above example serves to show that it is entirely possible to perform truly unknown formula determination using the FT ICR MS combined with MassWorks sCLIPS, it is imperative that good Spectral Accuracy of 98% or even 99% be achieved. For example, the determination for Loperamide ($C_{29}H_{34}N_2O_2Cl^+$, exact mass 477.2304Da) from the same data acquisition run was less successful (18th hit out of a total **846** candidates w/3ppm mass tolerance), due possibly to the ion threshold applied to this weaker ion, leading to the over-fit of the M+1, M+2, and M+3 peaks from the theoretical mass spectrum. Fortunately, this is clearly reflected in the poor Spectral Accuracy (92.06%), which can be used as a good formula determination diagnostic.



When operating at a higher resolving power of 100K, even with the low end ion threshold cutoff in place, unique formula determination can still be achieved for this weaker signal, due to the spectral separation of weaker isobars that are more susceptible to the cutoff. The compromise in Spectral Accuracy does not impact the formula ranking due to the limited number of candidate formulas at this lower mass.



Conclusions

➢Sub-ppm mass accuracy can be readily achieved on the FT ICR MS used. Since a mass tolerance window of at least 3 times the mass error (standard deviation plus bias) is needed for reliable formula search, even sub-ppm mass accuracy is not sufficient for the unique determination of truly unknown formulas.

➢The use of exact isotope modeling is key to unique formula ID and profile mode peak shape calibration is needed to achieve the required Spectral Accuracy.

➢For larger ions (m/z>700Da), a Spectral Accuracy >98.30% is required for the unique determination of truly unknown formulas.

➢While ion thresholding does compromise Spectral Accuracy, its impact can be reduced by acquiring the data at a higher resolving power, especially for ions of smaller m/z.

References

1. Kind, T., *BMC Bioinformatics* **2006**, 7, 234.
2. Kuehl, D., *Am. Lab. Online*, January **2008**, http://americanlaboratory.texterity.com/al-online/200801ol/?pg=18
3. Blom, K. F. *Anal. Chem.* **2001**, 73, 715.