SAMMI: A Direct Computational Approach to the Analysis of Multiply Charged Biomolecules and Their Modifications with Electrospray Mass Spectrometry

Ning Zhang[1,2]; Shundi Shi[2]; Shenglong Zhang[1]; David Good[3]; Don Kuehl[4]; Yongdong Wang[4]
[1]New York Institute of Technology, NY, NY; [2]Columbia University, NY, NY; [3]Covance Laboratories, Madison, WI; [4]Cerno Bioscience, Norwalk, CT

TP 662

Introduction

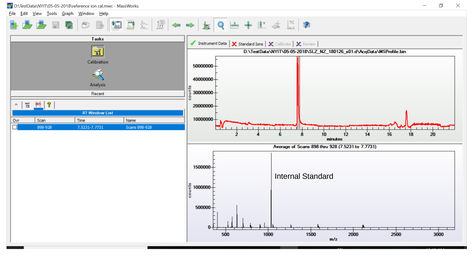
Biomolecules are typically measured as multiply charged ions with MS, requiring post-processing to convert the measured m/z into the original mass through an iterative, typically nonlinear, and often ill-understood process involving trial values for the original mass. Recognizing that MS measures not just single masses but entire distributions of all relevant masses and their abundances, we propose a different approach by eliminating the error- and artifact-prone original mass estimation step from the process altogether while achieving biomolecule identification and relative quantitation through full spectral calibration and analysis. This new approach was tested with RNA and intact proteins measured with LC-TOF and LC-MS Orbitrap system to evaluate its identification and quantitation potentials for these biomolecules and their modifications.

Method

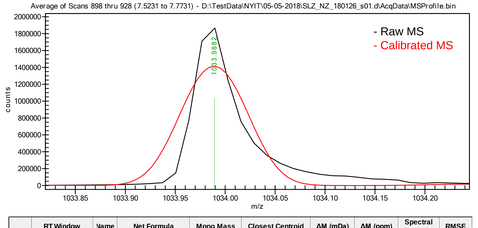
We break the process into three discrete but related steps: (1) delineate and mathematically account for the physical, chemical, and instrumental factors, namely charge state, isotope distribution, MS instrument line shape, and spectral overlaps arising from adducts or modifications; (2) apply the model along with a known standard (either singly or multiply charged) to calibrate the profile mode MS data for both mass and spectral accuracy with the MassWorks software tool[1] and, if necessary, adjust the MS operating conditions to ensure high spectral accuracy[2]; (3) run RNA and intact protein samples under high spectral accuracy conditions for both identification and relative quantitation, directly in the original mass spectral (m/z) space, utilizing all multiply charged signals but without any explicit multiple charge deconvolution step. We call this approach **SAMMI**, short for **S**pectrally **A**ccurate **M**odeling of **M**ulti-charged **I**ons.

LC-TOF Analysis of RNA Molecules

Accurate identification of RNA molecules, their fragments, and modifications are of great scientific and therapeutic importance[3].

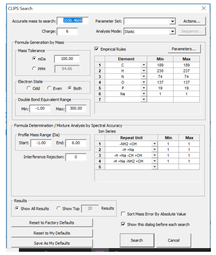


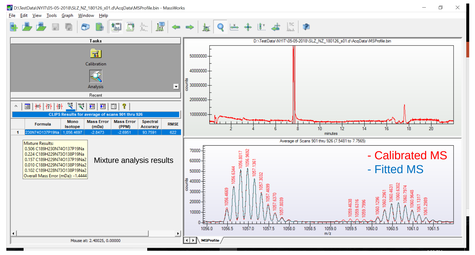
A 20-nt RNA measured on Agilent LCqTOF instrument with an internal single charge reference standard contained in the mobile phase at m/z 1034.



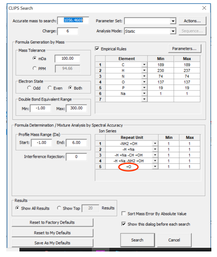
	RT Window	Name	Net Formula	Mono Mass	Closest Centroid	ΔM (mDa)	ΔM (ppm)	Spectral Accuracy	RMSE
1	Scans:898-928		C20H18N3O8F27P3	1.033.9881	1.033.9882	0.0507	0.0	99.6	4,140.09

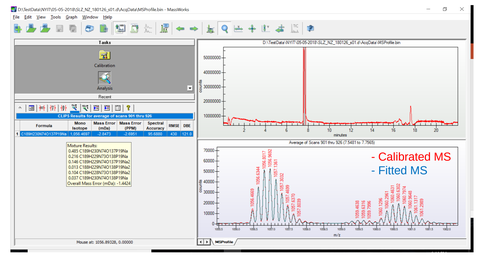
A full spectral calibration with MassWorks software tool allows for the TOF peak shape to be calibrated into a known functional form and thereby calibrating for both mass (m/z) accuracy and spectral accuracy. The spectral accuracy achieved is 98.6%, indicating the possible mass spectral residual error of about 1.4% at this (averaged) signal level.





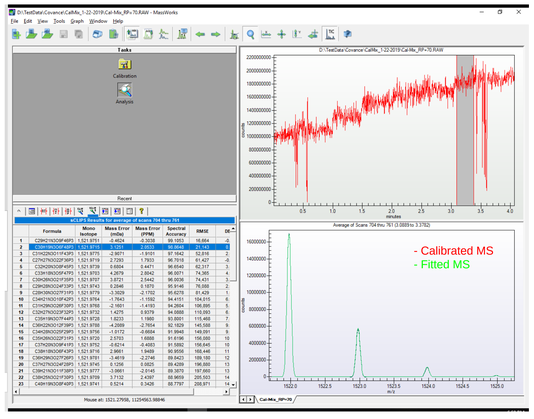
CLIPS known elemental composition search with the mixture mode analysis turned on correctly identified the 20-nt RNA sodium adduct in the presence of modifications and other adducts, to the spectral accuracy of 93.8%, accounting for a high percentage of the mass spectral signals observed in the given mass spectral range. The relative concentrations of various adducts and modifications were also obtained during the same spectral accuracy CLIPS analysis process. There however seems to be a lack of fit around m/z 1059.5.



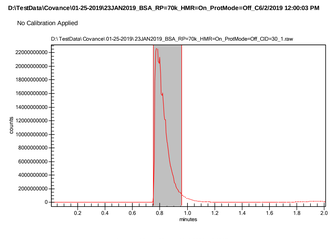


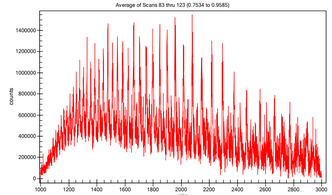
In order to account for the mass spectral signal around 1059.5 Da, it was postulated that possible oxidation may exist and the mixture analysis repeated with +O as an additional modification, resulting in improved spectral accuracy of 95.7% with the identification of this biochemically plausible modification and its relative quantitation (3.7%).

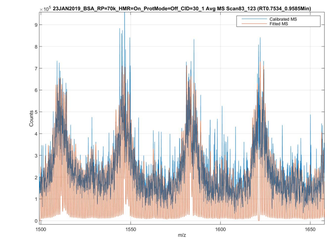
LC-Orbitrap QE Plus Analysis of Intact Protein Molecules



Earlier published work[4,5] has reported that the spectral accuracy is a significant function of both resolving power and AGC ion threshold inside the Orbitrap MS system. In order to establish the optimal operating conditions of the Orbitrap MS system, one Ultramark standard ion at m/z 1522 was observed and acquired at four resolving powers of 17.5K, 35K, 70K, and 140K while the AGC threshold was changed under Xcalibur monitor mode to the following 8 different levels: 20K, 50K, 100K, 200K, 500K, 1M, 3M, and 5M. The best spectral accuracy was observed under 70K resolving power at 3M AGC threshold (left).







Under the resolving power of 70K and 3M AGC threshold and the intact protein analysis conditions, the BSA standard was measured, showing a great many forms and species, which are further compounded by the respective multiple charges for each, as shown on the left.

A mass spectral accuracy analysis on the left shows a section taken from z = 34-45 and in consideration of the many forms of K adducts. As can be seen, only about 60% of the observed mass spectral signal was accounted for by the postulated species. This research is currently on-going with the aim of revealiing what exactly gets formed under intact conditions in order to achieve the same level of high spectral accuracy expected of a pure single charge standard at similar signal-to-noise level.

Conclusion

Accurate and scientifically transparent determination of multiply charge biomolecules is feasible, as long as the origin of mass spectral signal has been properly understood and validated through the underlying physics, chemistry/biochemistry, and careful full spectral calibration involving mass spectral peak shape.

For RNA molecules, we have achieved such full spectral understanding on a LC TOF MS systems approaching the level of actual measurement noise/error.

For larger intact proteins on an LC Orbitrap system, the work continues and the knowledge of the chemistry/biochemistry during the intact protein measurement process seems to be the key to achieve full spectral understanding and thus the measurement accuracy.

Acknowledgement

Ning Zhang, Shundi Shi, and Shenglong Zhang would like to acknowledge the support from NIH R21 grant (1R21HG009576).

References

1. MassWorks, Cerno Bioscience, Norwalk, CT USA.
2. Y. Wang, M. Gu., Anal. Chem., **2010**, *82*, 17, 7055.
3. Björkbom, A.; Lelyveld, V.S.; Zhang, S.; Zhang, W.; Tam, C. P.;Blain, J.C.;Szostak, J.W., J. Am. Chem. Soc., **2015**, *137*, 45, 14430.
4. Strife, R.J.; Wang, Y.; Kuehl, D., J. Mass Spectrom., **2018**, *53*, 10, 921.
5. Erve, J.C.L., Gu, M., Wang, Y. et al. J. Am. Soc. Mass Spectrom., **2009**, *20*, 11, 2058.