da Silva, M.P., **Kaesler, J.M.**, **Reemtsma, T.**, **Lechtenfeld, O.J.** (2020):
Absorption mode spectral processing improves data quality of natural organic matter analysis by Fourier-transform ion cyclotron resonance mass spectrometry
J. Am. Soc. Mass Spectrom. **31** (7), 1615 – 1618

http://dx.doi.org/10.1021/jasms.0c00138

Absorption mode spectral processing improves data quality of natural organic matter analysis by FT-ICR-MS

Maria P. da Silva[1], Jan M. Kaesler[1], Thorsten Reemtsma[1], Oliver J. Lechtenfeld[1,2]

[1]Helmholtz Centre for Environmental Research - UFZ, Department of Analytical Chemistry, Permoserstraße 15, 04318 Leipzig, Germany.

[2]ProVIS – Centre for Chemical Microscopy, Permoserstraße 15, 04318 Leipzig, Germany.

ABSTRACT: Natural organic matter (NOM) plays an important role in elemental cycles and ecology. Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is an ultra-high resolution technique used to molecularly resolve the complexity of NOM mixtures. However, even the very high mass resolution of FT-ICR-MS may result in multiple formula assignments to peaks in an NOM spectrum, especially at the high mass-to-charge ratio (m/z). Absorption mode is one option to process raw FT-ICR-MS data that can further increase the resolution of the peaks and has not been widely applied in NOM studies. In this study we show the advantages of using absorption mode for the analysis of NOM samples using a reference sample (Suwannee River fulvic acid). Absorption mode increased the precision of peak detection as well as the number (+23%) and accuracy of formula assignment (by 28%) when compared to magnitude mode, besides achieving three times higher resolution. The results presented here highlight the potential to reduce the error threshold used during molecular formula assignment. In conclusion, absorption mode shows advantages in the processing of NOM samples and other complex mixtures and should be promoted in the NOM community.

---

1. INTRODUCTION

Natural organic matter (NOM) represents the largest pool of organic carbon in all environments, such as inland and marine waters, soils or aerosols.[1-3] NOM is composed of different classes of naturally occurring organic compounds with differing ecological and biogeochemical roles that may vary as a function of environment and individual molecular reactivity.[4] High mass resolution is essential to molecularly resolve the complexity of NOM mixtures. State-of-the art technique to study the molecular chemical composition of NOM is Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS).[5,6] However, even the ultra-high mass resolution and mass accuracy of FT-ICR-MS may result in multiple molecular formula assignments to individual peaks in an NOM spectrum. Especially at higher mass-to-charge ratio (m/z) and using typical elemental boundaries in NOM studies (e.g. $CHON_4S_1$) the number of possible formulas within a pre-defined error threshold non-linearly increases while at the same time the resolution of FT-ICR-MS decreases.[7,8]

Magnitude and absorption processing are two different modes available to convert raw FT-ICR-MS data from the time-domain to the frequency-domain. For magnitude mode frequency is plotted versus the magnitude of the complex output of the Fourier transformation. While for absorption mode knowing of the phase of the ions is necessary to plot the frequency against the real part of the Fourier transformation. However, ions with different m/z show different initial phase angles in the ICR cell when the detection starts.[9] Thus, despite absorption mode spectral processing is expected to improve peak resolution and signal to noise ratio (SNR) as compared to magnitude mode, the challenges related to the correction of the phase of the ions resulted in magnitude mode being prevalent in the processing of FT-ICR-MS spectra up to now.[10] The recent development of spectral processing tools such as phase algorithms, apodization functions and baseline corrections promoted absorption mode [11,12]

Several studies have shown an increase in the mass resolution of absorption mode spectra and the improvement in mass accuracy, studying petroleum samples.[10,11,13] Nevertheless, the reported significant improvements outside the scope of NOM studies (e.g. m/z above 1000) and the extra step in the processing of absorption spectrum resulted in magnitude mode being still routinely applied in the field of NOM. However, no studies have been published testing the precision and reproducibility of absorption mode, which also represent an important and decisive point in the choice of the spectral processing mode. In this study we evaluate absorption mode for the analysis of NOM samples via comparison with magnitude mode using analytical figures of merit such as accuracy, precision and reproducibility.

2. METHODS

Suwannee River fulvic acid (SRFA) samples were measured using a solariX XR 12 Tesla FT-ICR-MS (Bruker Daltonik GmbH, Bremen, Germany) with dynamically harmonized analyzer cell and electrospray ionization (ESI) in negative mode, 4 megaword time-domain and 256 co-added scans in the mass range of 150-1000 m/z. FT-ICR-MS data were initially recorded in magnitude mode and subsequently transformed to absorption mode using FTMSprocessing (v 2.2.0). One measurement of SRFA was done to calibrate the phase equation that was applied in the following triplicate measurements. To minimize the intensity of the side-lobes of the peaks,[14] full sine apodiza-

Table 1. Comparison of absorption and magnitude mode spectrum of triplicate measurements of SRFA.

Processing mode	SNR threshold	# peaks	Peaks shared between triplicates (%)	Avg. resolution at 400 m/z	RMSE of calibration [ppm]	# assignments	Assignments shared between triplicates (%)
Absorption	2	20,172 ± 1,085	69	1,202,493 ± 6,874	0.036 ± 0.002	7,530 ± 298	79
Magnitude	4	19,113 ± 724	63	497,124 ± 10,900	0.038 ± 0.002	6,344 ± 222	78

tion function was applied to processes the magnitude mode spectra. For absorption mode, the full sine apodization function was applied by choosing the Kilgour function with the maximum occurring at half of the original transient length (F = 0.5). All spectra were internally calibrated with a reference mass list of known NOM masses (n = 188; 250 < m/z < 640). To determine the number of peaks shared between replicate measurements alignment of the peak list was done by applying an initial window width of 1 ppm.[15] Molecular formulas were then assigned to mass peaks within ±0.5 ppm in the range of 150-1000 m/z according to published rules using element ranges $C_{1\text{-}80}H_{1\text{-}198}N_{0\text{-}4}O_{0\text{-}40}S_{0\text{-}1}$.[7,16]

3. RESULTS AND DISCUSSION

To better compare mass spectral parameters between the modes, SNR threshold was initially set to two for absorption mode and to four for magnitude mode. This resulted in mass spectra with similar number of peaks (absorption: 20172, magnitude: 19113). The different SNR is due to the different way to process the signal. In a phased spectrum the random distributions of the errors is expected to increase the SNR by $\sqrt{2}$, but the reduction can be even greater due to the use of different apodization functions.[17] In this study, the use of Kilgour proved to be suitable with no baseline deviation or imperfect baseline observed in any spectrum (see SI for details). Regarding resolution, the average mass resolution at 400 m/z was 3 times higher in absorption mode than magnitude mode (Table 1). A similar improvement in mass resolution could also be achieved in FT-ICR-MS by increasing the magnetic field strength or transient time. Next to costs and space requirements, a higher magnetic field strength demands also highest shimming precision. Similarly, longer transients are possible but often demand sample-specific cell tuning due to the risk of ion cloud phase coherence loss. Moreover, available sample volume and instrument time of FT-ICR-MS analysis often limit the number of possible measurements (and replicates thereof). Thus, mass precision and accuracy are essential for the quality of a mass spectrum and subsequent reliable assignment of formulas.

Peak precision was tested as reproducibility of peak m/z detection between triplicates. Both modes resulted in similar reproducibility of peak detection being within the threshold for acceptable reproducibility suggested by Sleighter, et al. [18] Note that the increase by 6% in the number of shared peaks for absorption mode corresponded to more than 1800 additional shared peaks as compared to magnitude mode. This increase in peak detection reproducibility is equivalent to a lower probability of false positive peaks, when repeat measurements cannot be performed. Due to random errors in peak detection between measurements, a window width of 1 ppm is usually applied to search for common peaks in replicates.[18] When narrowing the search window from 1 ppm to 0.1 ppm, the number of shared peaks decreases faster in magnitude than in absorption mode (see SI for details). The higher mass resolution in absorption mode thus translates into an increase in the precision of peak detection of FT-ICR-MS measurements without changing magnetic field strength or transient length.

Mass accuracy of both modes was tested after assigning molecular formulas to the mass peaks. Although both spectra had some peaks assigned with more than one formula, absorption mode had 20% less multiple assignments than magnitude mode (see SI for details). In the further analysis, the peaks assigned with more than one formula were removed. Absorption mode spectra resulted in >1000 additionally assigned formulas, indicating that new assignments could be made which were not possible in magnitude mode. The comparably high percentage of shared formulas in both modes indicates that the higher number of formulas assigned to absorption spectra are true positives and not the result of including spectral noise due to the decrease in SNR threshold. Out of the 1186 new assignments, 1159 (98%) were present in all three absorption mode spectra. New assignments were mainly above m/z 450 and of CHO (44%), CHNO (24%) and CHOS (19%) compound class (Figure 1).

Comparing only the molecular formulas present in all six spectra (triplicates of both modes, n = 4829), absorption mode resulted in an overall root mean square error (RMSE) of formula assignment being 28% lower than in magnitude mode (0.088 ± 0.000 vs 0.122 ± 0.002). A pronounced improvement in mass accuracy for molecular formulas containing nitrogen (N) and sulfur (S) was observed (Figure 2). This resulted in RMSE of formula assignment being more similar among compound classes in absorption mode. In NOM spectra, N- and S-containing compounds have a broader distribution of mass errors and thus larger RMSE values due to the overall lower SNR as compared to CHO formulas (see SI for details). It was recently demonstrated that spectral averaging across multiple samples and, hence, reduction of random errors can also be applied to improve mass accuracy of shared peaks.[19] However, absorption mode processing improves the mass accuracy by increasing peak resolution and shape of each individual spectrum.

Lastly also the peak intensity reproducibility between replicates was assessed. Again, only peaks assigned to molecular formulas present in both modes were selected and the coefficient of variation (CV) was calculated from the normalized intensities for each mode. The mean CV was 6% for both magnitude and absorption mode, with 75% of peaks varying in intensity less than 10% between triplicates. No systematic bias in relative intensities was observed between magnitude and absorption mode (see SI for details), recommending absorption mode also for semi-quantitative applications using relative peak intensities.

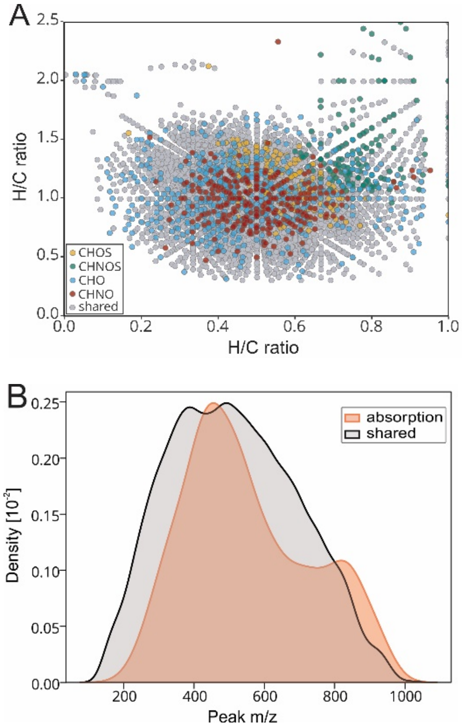

Figure 1. (A) Van Krevelen diagram showing the compound class of new molecular formulas assigned to shared peaks in triplicate measurements of absorption mode (peaks shared between modes are shown in grey); (B) m/z density distribution of the peaks shared between the modes and additionally assigned in absorption mode.

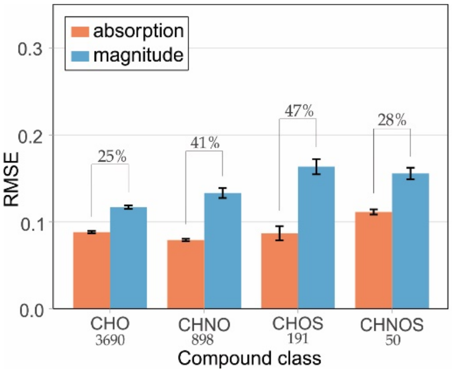

Figure 2. RMSE of formula assignment for different compound classes (n = 4829). The number below the bars indicates the number of formulas assigned in each class.

4. CONCLUSIONS

The higher resolution in absorption mode also corresponds to an increase in peak precision, reproducibility and accuracy, recommending its utilization where sample volume is limited and/or when fast acquisition times are crucial, e.g. for LC hyphenation. Furthermore, the increase in mass accuracy of absorption mode highlights the possibility of confining the window width for formula assignments. While the search range is often set to ± 0.5 ppm in NOM studies, our results indicate that ± 0.25 ppm still covers > 99% of possible assignments in absorption mode on a 12T FT-ICR-MS instrument. The small additional effort of spectrum phasing in absorption mode already pays off from the first measurement and can even be applied to previously acquired magnitude mode spectra. We recommend the use of absorption mode FT-ICR-MS to study NOM and other highly complex mixtures where easy-to-apply improvements in data quality are highly demanded.

ASSOCIATED CONTENT

Supporting Information

Precision of peak detection, example of a peak, SNR of formulas containing nitrogen and sulfur, intensity variation between runs and comparison of apodization functions in absorption mode (PDF).

AUTHOR INFORMATION

Corresponding Author

* Oliver J. Lechtenfeld − Helmholtz Centre for Environmental Research - UFZ, Department of Analytical Chemistry, Permoserstraße 15, 04318 Leipzig;
Email: oliver.lechtenfeld@ufz.de

Author Contributions

MPS, JK, TR, and OJL designed the analysis. MPS coded and conducted the statistical tests with input from OJL, JK and TR. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

This work was supported by Rudolf und Helene Glaser-Stiftung in the frame of project “Influence of natural factors on concentration, quality and impact of dissolved organic carbon in the Bavarian Forest National Park” (project No. T0083\30771\2017\kg). The authors are grateful for using the analytical facilities of ProVIS Center for Chemical Microscopy within the Helmholtz Center for Environmental Research Leipzig, which is supported by European regional Development Funds (EFRE - Europe funds Saxony) and the Helmholtz Association. We also would like to thank Kai Franze for software development.

REFERENCES

(1) Cole, J. J.; Prairie, Y. T.; Caraco, N. F.; McDowell, W. H.; Tranvik, L. J.; Striegl, R. G.; Duarte, C. M.; Kortelainen, P.; Downing, J. A.; Middelburg, J. J.; Melack, J. *Ecosystems* **2007**, *10*, 172-185.

(2) Aitkenhead, J. A.; Hope, D.; Billett, M. F. *Hydrological Processes* **1999**, *13*, 1289-1302.

(3) Simoneit, B. R. T.; Mazurek, M. A. *Atmospheric Environment (1967)* **1982**, *16*, 2139-2159.

(4) Jaffé, R.; McKnight, D.; Maie, N.; Cory, R.; McDowell, W. H.; Campbell, J. L. *Journal of Geophysical Research: Biogeosciences* **2008**, *113*.

(5) Lu, Y.; Li, X.; Mesfioui, R.; Bauer, J. E.; Chambers, R. M.; Canuel, E. A.; Hatcher, P. G. *PLoS One* **2015**, *10*.

(6) Riedel, T.; Zak, D.; Biester, H.; Dittmar, T. *Proceedings of the National Academy of Sciences* **2013**, *110*, 10101-10105.

(7) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. *Anal. Chem.* **2007**, *79*, 1758-1763.

(8) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(9) Kilgour, D. P. A.; Orden, S. L. V. *Rapid Commun. Mass Spectrom.* **2015**, *29*, 1009-1018.

(10) Qi, Y.; Barrow, M. P.; Orden, S. L. V.; Thompson, C. J.; Li, H.; Perez-Hurtado, P.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 8477-8483.

(11) Qi, Y.; Li, H.; Wills, R. H.; Perez-Hurtado, P.; Yu, X.; Kilgour, D. P. A.; Barrow, M. P.; Lin, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 828-834.

(12) Xian, F.; Hendrickson, C. L.; Blakney, G. T.; Beu, S. C.; Marshall, A. G. *Anal. Chem.* **2010**, *82*, 8807-8812.

(13) Qi, Y.; Barrow, M. P.; Li, H.; Meier, J. E.; Van Orden, S. L.; Thompson, C. J.; O'Connor, P. B. *Anal. Chem.* **2012**, *84*, 2923-2929.

(14) Zhang, L. Z.; Zhang, Y. H.; Zhao, S. Q.; Chung, K. H.; Xu, C. M.; Shi, Q. *Int. J. Mass Spectrom.* **2014**, *373*, 27-33.

(15) Kido Soule, M. C.; Longnecker, K.; Giovannoni, S. J.; Kujawinski, E. B. *Org. Geochem.* **2010**, *41*, 725-733.

(16) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2007**, *8*, 105.

(17) Marshall, A. G. *The Journal of Chemical Physics* **1971**, *55*, 1343-1354.

(18) Sleighter, R. L.; Chen, H. M.; Wozniak, A. S.; Willoughby, A. S.; Caricasole, P.; Hatcher, P. G. *Anal. Chem.* **2012**, *84*, 9184-9191.

(19) Merder, J.; Freund, J. A.; Feudel, U.; Niggemann, J.; Singer, G.; Dittmar, T. *Anal. Chem.* **2019**.

Graphic abstract