



# Enhancing data acquisition for the analysis of complex organic matter in direct-infusion Orbitrap mass spectrometry using micro-scans

Cédric Wolters, Laurène Flandinet, Chao He, Junko Isa, François-régis Orthous-daunay, Roland Thissen, Sarah Hörst, Véronique Vuitton

## ► To cite this version:

Cédric Wolters, Laurène Flandinet, Chao He, Junko Isa, François-régis Orthous-daunay, et al.. Enhancing data acquisition for the analysis of complex organic matter in direct-infusion Orbitrap mass spectrometry using micro-scans. *Rapid Communications in Mass Spectrometry*, 2020, 34 (15), <10.1002/rcm.8818>. <hal-03085119>

**HAL Id: hal-03085119**

**<https://hal.science/hal-03085119v1>**

Submitted on 4 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Wolters Cédric (Orcid ID: 0000-0002-9710-4740)

# Enhancing data acquisition for the analysis of complex organic matter in direct-infusion Orbitrap mass spectrometry by using micro-scans

Cédric Wolters<sup>1\*</sup>, Laurène Flandinet<sup>1</sup>, Chao He<sup>2</sup>, Junko Isa<sup>1</sup>, François-Régis Orthous-Daunay<sup>1</sup>, Roland Thissen<sup>3</sup>, Sarah Hörst<sup>2</sup>, Véronique Vuitton<sup>1</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France

<sup>2</sup>Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA

<sup>3</sup>Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, 91405, Orsay, France

\*Author for Correspondence : [cedric.wolters@univ-grenoble-alpes.fr](mailto:cedric.wolters@univ-grenoble-alpes.fr)

## Abstract

**RATIONALE:** Acquisition quality in analytical science is key to obtaining optimal data from a sample. In very high-resolution mass spectrometry, quality is driven by the optimization of multiple parameters, including the use of scans and micro-scans (or transients) for performing a Fourier transformation.

**METHODS:** 39 mass spectra of a single synthesized complex sample were acquired using various numbers of scan and micro-scan determined through a simple experimental design. An electrospray ion source coupled with an LTQ-Orbitrap-XL mass spectrometer was used and acquisition was performed using a single mass range. All the resulting spectra were treated in the same way to enable comparisons of assigned stoichiometric formulae between acquisitions.

**RESULTS:** Converting the number of scans into micro-scans enhances signal quality by lowering noise and reducing artifacts. This modification also increases the number of attributed stoichiometric formulae for an equivalent acquisition time, giving access to a larger molecular diversity for the analyzed complex sample.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/rcm.8818

This article is protected by copyright. All rights reserved.

**CONCLUSION:** For complex samples, the use of long acquisition times leads to optimal data quality, and the use of micro-scans instead of scans-only maximizes the number of attributed stoichiometric formulae.

## Introduction

Electrospray ionization Orbitrap mass spectrometers have been used for more than ten years for high resolution mass spectrometry analysis, either in direct-infusion or coupled to gas (GC) or liquid (LC) chromatography. Application areas include the analysis of natural and synthetic complex samples in “omics” sciences such as metabolomics (1,2), proteomics (3–6) or lipidomics (7,8), and the analysis of environment complex samples such as groundwaters (9,10) and atmospheric aerosols (11–13). High mass resolution is required for each of these fields of study due to its ability to reveal the molecular diversity of a sample by attributing unique stoichiometric formulae to detect signals.

For all these complex samples, direct-infusion Ion Cyclotron Resonance (ICR) mass spectrometers are unmatched as far as resolution is concerned, allowing the identification and attribution of mass separation values below an electron mass. On the other hand, Orbitrap mass spectrometers are mainly used due to their versatility, scan rate suitable for LC acquisition, and easy to setup parameters. Various applications use direct-infusion Orbitrap analysis especially in environmental science (14), nuclear physics (15) or planetary sciences (16–18) for the analysis of meteorites and astrophysical laboratory analogues. A development project for an Orbitrap is currently under way, the CosmOrbitrap (19–21), aiming at allowing the use of direct-infusion mode with a laser desorption ionization source onboard spacecraft to analyze molecular diversity in-situ on non-terrestrial objects.

Data quality is important in analytical techniques. This particular term could refer to multiple concepts, from the acquisition process and validation procedures (22) to the structure of the data itself. In this study, we focus on the second part by assessing that the sample can be attributed properly (resolution, overall signal intensity, artifacts) and that the attribution is representative of the sample. In order to have a reproducible spectrum with high signal to noise ratio, we need to run instrumental methods that increase accuracy, precision and sensitivity. While it is well-known that using transient averaging before performing the Fourier transformations in FT-ICR instruments gives high sensitivity and mass precision and accuracy, this option – called “micro-scans” by Thermo Scientific – is poorly known for direct-infusion Orbitrap mass spectrometry. The Fourier transformation is a linear mathematical application. Thus, doing the sum of Fourier transformed mass spectra is equivalent to doing the Fourier transformation of the summed transients. Given this, performing micro-scan acquisitions should be the equivalent of standard scan acquisitions. Nevertheless, for Orbitrap mass spectrometry, Thermo Scientific software performs the Orbitrap specific noise removal just before applying the Fourier transformation. This specific operation combined with the limited acquisition time nullifies the linearity of the Fourier transformation application, generating a non-equivalence of the Fourier transformation application depending on if you apply it on summed transients or on the individual transients.

Only a few works mention micro-scans for LC-coupling and mainly during MS/MS acquisitions (4–6,15). This specific parameter is also not advertised by Thermo Scientific on instrument specifications and technical sheets. In this work, we will focus on using a complex organic sample available in our lab to investigate the impact of micro-scans and scans on data quality and the number of attributed stoichiometric formulas.

This article is protected by copyright. All rights reserved.

## Methods

### Sample synthesis

We use a laboratory analogue of Titan's atmospheric aerosols, called tholin, which is representative of a complex organic sample with several thousands of stoichiometric formulae. This analogue was synthesized to mimic the formation of aerosols in Titan's atmosphere by using a mixture of  $N_2$ ,  $CH_4$ , and CO in proportions of 94.98%:5%:0.02% (%vol). The apparatus, sample synthesis, and sample recovery have been described extensively by He et al (23). Several mass spectrometry studies have been carried out on this kind of sample and they have revealed the presence of rich and diverse organic matter diversity (24–28). The Orbitrap mass spectra interpretation of this specific sample is ongoing and will not be discussed extensively in this paper.

### Sample preparation

1 mg of the sample was dissolved in 1 mL of methanol (Carlo Erba, Milan, Italy; UPLC grade) using an Eppendorf (Hamburg, Germany) pipette and a 1.5-mL polypropylene plastic tube. The solution was vortexed for ten minutes and then centrifugated for ten minutes at 9400 G. 500  $\mu$ L of the supernatant was dissolved with 500  $\mu$ L fresh methanol in a new 1.5-mL polypropylene plastic tube.

This preparation was performed three times to obtain all the required volumes for carrying out all the analyses.

### Instrumentation

Analyses were performed with an LTQ Orbitrap XL™ mass spectrometer equipped with an IonMax™ ESI source (Thermo Scientific, Bremen, Germany). This hybrid mass spectrometer uses an ion trap coupled to the Orbitrap cell. A C-trap is placed between the two traps to store ions coming from the source and inject them in packets directly into the Orbitrap cavity. The MS functions were controlled by LTQTune software (Thermo Scientific). Data treatment and visualization were performed with XCalibur software (Thermo Scientific) and then with a home-made software package "Attributor" (29) developed with IGOR Pro (WaveMetrics, Portland, OR, USA).

For accurate mass calibration, a mixture of Caffeine, Met-Arg-Phe-Ala peptide (MRFA) and Ultramark 1621 (a mixture of fluorinated phosphazine polymers) was used (Thermo Scientific). This calibration mixture is the Thermo Scientific default calibration mixture for LTQ-Orbitrap-XL instruments.

All analyses were performed at the maximum instrument resolution (resolution better than 100,000 at  $m/z$  400; mass accuracy of  $\pm 2$  ppm) using the AGC setting of  $5 \cdot 10^5$  ions storage. For the study needs, we used an intermediate mass range, from  $m/z$  150 to 450 as it includes the highest intensity distribution and allows a broad attribution with limited resolution issues at high masses. The sample flow rate was set at 3  $\mu$ L/min and led to a mean sample storage time around 10 ms before injection into the Orbitrap. The ESI settings were 3.5 kV for source voltage, 5 (arbitrary units) for the sheath gas flow rate and 0 (arbitrary units) for the auxiliary and sweep gas flow rates. The capillary temperature was set at 275°C, the capillary voltage at 34 V, and the tube lens voltage at 70 V. All other instrument parameters were checked and transferred from the automatic instrument calibration. Micro-scans and scans were defined before each acquisition depending on the experiment needed, as described in the paragraph "Experimental design".

### Micro-scans

The term micro-scans refers to the free parameter that can be changed inside the MS configuration panel in the LTQTune software. From the "LTQ-Orbitrap Operation Training Course Manual" (European Training Institute, Thermo Scientific). Several descriptions are given of the acquisition sequence used by the Orbitrap:

- Ion time (IT) is "the time in milliseconds that ions are allowed to accumulate in the trap. The IT time is variable and is calculated by the AGC."
- A micro-scan is "a complete cycle combination of the prescan and the analytical scan which will include the IT time (trapping and cooling of the ions) and scan out time (time taken to eject the ions)."
- A scan is the "time taken to complete the overall experiment and is dependent upon the number of micro-scans used."

At the completion of the process, the complete scan is the combination of one or several micro-scans. Each complete/combined scan includes one Fourier transformation. In the end, carrying out micro-scans or the equivalent scans-only operation results in roughly the same overall acquisition time.

The process of using transients is well-known for FT-ICR instruments where transients (micro-scans for Orbitrap) are averaged before the Fourier transformation to produce the final spectra (30,31). Unfortunately, the Orbitrap does not give access to the transient signal nor the Fourier transformation process as opposed to ICR instrument where the complete process is open access and can be used to generate the mass spectra through dedicated signal treatments. It is also common in FT-ICR direct-infusion experiments to acquire data over a long time (28,31) using a high number of transients to increase the data quality in direct infusion, whereas this process seems seldom used for Orbitrap direct-infusion measurements.

### Experimental design

To investigate the impact of scans and micro-scans, we designed a set of experiments to be able to obtain maps of numbers of stoichiometric formulae. As we aimed to investigate the impact of changing from short to long duration acquisitions and wanted to limit the number of resulting experiments, we designed an experimental plan based on asymmetric  $2^n \times 2^m$  maps with  $n=[0;3]$  the number of scans (i.e. 1, 2, 4 and 8 scans) and  $m=[0;10]$  the number of micro-scans (i.e. 1, 2, 4, 8, 16, ..., 1024 scans). This resulted in 38 individual acquisitions varying from 1 to 8 scans and from 1 to 1024 micro-scans. The reference analysis was chosen to be outside the map with 1024 scans using 1 micro-scan each. In terms of time, acquisitions varied from seconds to around 30 minutes.

The repeatability and reproducibility of the acquisition are also important factors that are not addressed by the experimental design realized for this study. A study on similar samples has shown 1% repeatability and 2% reproducibility of direct-infusion analyses. The repeatability and reproducibility were evaluated on a number of identical attributed stoichiometric formulae across the same sample prepared and acquired at different moments (hours, days and weeks). Variations are mostly observed at very low intensity and are due to signal to noise variation and slight differences in sample preparation as several solutions are done. Thus, we do not expect repeatability and reproducibility to have an impact in this study.

This article is protected by copyright. All rights reserved.

## Data treatment

A home-made software package was used to extract data from the Thermo .raw files. The extraction was based on an Igor Pro routine that uses Windows open libraries to open and read directly the data inside the Thermo-produced .raw files. After data extraction, each scan was summed to produce a single resulting mass spectrum. This mass spectrum was then loaded into Attributor, a home-made mass spectrometry software package specially developed to treat and attribute high-resolution mass spectrometry data (18). Each spectrum was converted from peak profile to centroids, normalized and internally calibrated, and peaks were assigned a molecular formula using the same routines and parameters to be able to compare the results. The internal calibration used CH<sub>2</sub> families with more than 10 members, i.e. series of attributed stoichiometric formulae that differ only by variation of a CH<sub>2</sub> group. In order to select the accurate stoichiometric formulae for the internal calibration, we selected assigned peaks whose stoichiometric formulae only included oxygen and nitrogen - 0 to up to 3. Attributions were performed for singly positive charged ions using the graph properties of the spectra, with CH<sub>2</sub>, C, NH and O as propagation parameters. Even if <sup>13</sup>C stoichiometric formulae are detected and can be attributed, they are not considered in this study. We would also have to consider the isotopic ratios to explain the absence of <sup>13</sup>C between different experiments in addition to variation induced by the acquisition method. The propagation list is then used as is, without any post-processing treatment. Typical errors and stoichiometry in the function of mass to charge ratio representation are available in Supporting Information 1.

## Results

The mass spectrum in Figure 1A highlights a striking molecular organization in the sample. We can hence distinguish a periodic pattern where maxima (or minima) are visible, separated by 13.5 u. One of the clusters constituting the pattern is highlighted with a black box in Figure 1B for illustration. This pattern structure is typical of Titan's tholins and indicates that the sample is made through a polymerization process as described in several works such as in Pernot and co-workers (24) and Hörst's thesis (25). They highlight that the major pattern is not only CH<sub>2</sub> as we can expect from complex organic matter but a combined role of CH<sub>2</sub> and HCN in the final structural composition. Given this dual polymerization pattern, each cluster exhibits a constant (C+N) value, with hydrogenation variation generating the cluster diversity (24–27).

Before doing any attribution, we check the spectra for each experiment and particularly if the micro-scans affect the signal intensity. As can be seen in Figure 1A, using the micro-scans function of the Orbitrap seems to enhance the signal after *m/z* 325 compared with the scans-only operation while at the same time, enlarging the cluster distribution at both lower and higher masses (Figure 1B). This enlargement seems to indicate a higher diversity of H content for a given (C+N) content. At first glance, we also see a difference for the minimum signal of each spectrum, at 8.10<sup>-3</sup> for micro-scans and 2.10<sup>-3</sup> for scans-only (not shown in the Figure here). This limitation of the dynamical range could appear as a problem for the sensitivity of the analysis, but we will demonstrate later in this work that the dynamic range is not the correct parameter to compare the two operation techniques.

As intensity vs  $m/z$  spectra do not show much information about a sample's molecular diversity, we use another representation that can highlight sample signals and diversity over artifacts and noise: the mass defect diagram, a variation of Kendrick diagrams using  $^{12}\text{C}$  as a reference instead of  $^{12}\text{C}^1\text{H}_2$ . The mass defect is thus defined, as shown in equation 1, by only using the exact mass measured by the instrument:

$$\text{Mass defect} = \text{Exact mass} - \text{round}(\text{Exact mass}) \quad (1)$$

We report in Figure 2 the mass defect diagrams for 1024 scans with 1 micro-scan experiment (left) and 1 scan with 1024 micro-scans experiment (right). We can see different structures and shapes inside these diagrams: (1) vertical lines, (2) cones of close points at positive and negative mass defect values and (3) individual points. The vertical lines are more frequent in scan-only and are attributed to radio artifacts, which are generated by radio wavelengths that are detected by the Orbitrap amplification system. Noise points are generated by classical electronic noise and do not present a periodic pattern. The cone structures are similar in both experiments and represent ions resulting from electrospraying the sample, with the top cone attributed to (mostly) singly-charged ions and the bottom cone to (only) doubly-charged ions.

Despite the similar acquisition time of approximately 30 minutes and the same total number of analytical scans for the analyses shown in Figure 2, all different features indicate differences directly inside the Orbitrap data generation system, depending on the way in which the experiment has been set up. Differences between scans and micro-scans concerning radio artifacts and random noise could be attributed to the way in which the Thermo Scientific software handles the automatic noise removal before doing the Fourier transformation. This process appears to be more efficient when several micro-scans are summed before the noise removal than when the noise removal is applied on each scan. In extreme cases with using scans-only, such as with diluted samples, the random points and radio artifacts can spread on the entire range of mass defect and  $m/z$  values and hide any true sample signal (see Figure 6 in Danger et al (16) for an example of radio artifacts overlapping the real sample signal). Although this kind of issue can be treated by post-treatment procedures such as noise hard cutting, using micro-scans significantly reduces the need for post-treatment procedures.

We summarize in Figure 3 the total number of attributed stoichiometric formulae per experiment. We observe that for a given number of scans and an increasing number of micro-scans and *vice-versa*, we attribute more and more stoichiometric formulae. Nevertheless, this observation hides some disparities, with an asymmetry in the number of attributed stoichiometric formulae for equivalent acquisition times. Diagonals in the table represent acquisitions carried out during an equivalent amount of time, from a second to 30 minutes. This indicates that for an identical sample, acquiring either scans, micro-scans or a combination of both does not lead to equivalent results, driving us to the search for the best acquisition setup.

This article is protected by copyright. All rights reserved.

There is also interest in checking if the attributed stoichiometric formulae are overlapping between the different analyses. We report in Supporting Information 2 individual attributed stoichiometric formulas on top of the mass defect diagram for each analysis using 1 single scan for 1 to 1024 micro-scans, respectively a second to approximately 30 minutes of acquisition time. As expected, we observe an increasing data complexity with an enlargement of the sample cloud correlated with the enlargement of the attributed stoichiometric formula cloud. This indicates that all attributed stoichiometric formulae at low micro-scans are included in the highest number of micro-scans. We also observe that attributions are strictly contained into the sample cloud and that there is no attribution out of this area, consistent with other points being either artifacts, noise or ions with a higher charge. This observation highlights that the accumulation of scans increases the signal to noise ratio, leading to more detected peaks while increasing acquisition time.

Another interesting point here is that we can observe in Figure 3 an equivalent number of stoichiometric formulae for different acquisition methods. This can be important in the case of samples with low available quantity. The most significant improvement is observed between 1024 scans and 2 scans of 128 micro-scans that attribute the same number of stoichiometric formulae, but with significant variation of acquisition time: 30 minutes and 7.5 minutes, respectively. Thus, using micro-scans allows a reduction of time and volume consumption of four times compared with the scans-only protocol for an equivalent number of stoichiometric formulae.

Aside from the number of stoichiometric formulae, we also observed from the raw mass spectra a difference of dynamic range for methods using micro-scans *versus* methods using scans-only. This can also be seen in Figure 4 that represents the histogram of intensity for stoichiometric formulae, where we have a hard intensity cut around  $10^{-2}$  using 1024 micro-scans whereas we observed intensities down to roughly  $3 \cdot 10^{-3}$  when using 1024 scans. Even if the apparent dynamic range is lower using micro-scans than using scans-only, comparing identical stoichiometric formulae between the two operations leads to the reverse conclusion. When we compare for identical stoichiometric formulae between the two methods, all identical molecules are higher than  $5 \cdot 10^{-1}$  while all detected molecules go down to  $10^{-2}$  using micro-scans. Thus, the dynamic range is not a parameter to consider when comparing scans with micro-scans operations as the intensity of identical stoichiometric formulae is not consistent between the two operations.

When focusing on identical stoichiometric formulae between the two experiments, more than 97% of the stoichiometric formulae attributed using 1024 scans are found in the stoichiometric formula attributions using 1024 micro-scans, corresponding to 2164 identical stoichiometric formulae in both datasets. A careful inspection of the remaining 3% (i.e. 41 stoichiometric formulae) indicates that they are wrong stoichiometric formulae that would be removed using a manual post-treatment work on stoichiometric formulae (for instance too high N/C or O/C ratio) and some low intensity attributed points that can be assigned to an unfortunately badly positioned noise point. The intensity comparison of the identical stoichiometric formulae shows that we are attributing more than 1300 new mid to low intensity stoichiometric formulae using micro-scans in addition to all the stoichiometric formulae detected using 1024 scans, represented in Figure 3 by the red bars. Of the new attributed molecules, some could also be misattributed and would have been removed by a manual post-treatment step, estimated to be less than 100 signals.



There are several ways to represent stoichiometric formulae to compare data. Stoichiometric ratio representation has been considered such as H/C; N/C; O/C as a function of the ratios or the mass, but the resulting information is not conclusive as the information is stretched and/or non-linear in most of the cases. We report here that the Double Bond Equivalent (DBE) as a function of the number of carbon, nitrogen, oxygen and hydrogen representation can be used to highlight new attributed stoichiometric formulae, as presented in Figure 5. In addition to identical stoichiometric formulae, we observe that the micro-scans operation detects a wider diversity with the addition of low DBE, but also detects new stoichiometric formulae at all DBE values. This is consistent with observations made on the mass spectra, where all the peak clusters inside mass spectra are denser using micro-scans than when using scans-only as discussed for Figure 1B. This can be also explained by an increase of the signal to noise ratio using micro-scans compared with the scan-only method.

## Discussion and Conclusion

In the framework of analyzing complex samples such as environmental or biological samples with several thousands of molecular formulae, it is necessary to accumulate for long enough to have the maximum possible number of formulae at the end. Nevertheless, the way used to accumulate data has a direct impact on the number of formulae obtained. Thus, converting the number of scans to micro-scans will result in the best possible acquisition methods and the maximum number of attributed stoichiometric formulae at the end. This way of using micro-scans over long periods instead of using scans-only is also identical to the data processing used in classical ICR data-processing (without specific post-treatment) where all transients are summed before the Fourier Transformation necessary to obtain the final mass spectra.

All results have been acquired using a complex organic matter sample. This means that the instrument resolving power and mass accuracy are critical as the number of probable stoichiometric formulae increases with the mass increase. The use of a high resolving power and high accuracy instrument is then mandatory to achieve sufficient separation between masses. In the case of our study, we analyze only CHNO molecules in a mass range where we know that the Orbitrap at its highest resolving power can attribute one and only one stoichiometric formula per mass, at the given  $\pm 2$  ppm mass accuracy. Based on this study, the use of micro-scans or scans operation does not affect the resolution nor the mass accuracy of the data. Micro-scans enable the detection of additional low-intensity signals by increasing their signal to noise ratio compared with the equivalent number of scans. The use of micro-scans operation results in cleaner mass spectra, in the sense that the automatic noise removal performed by the Thermo Scientific software before the Fourier transformation is more efficient using micro-scans than using the scans-only method. In the case of simple or pure samples, there is no organic diversity. In that case, the use of a limited time acquisition, i.e. several seconds to a minute using scans-only is sufficient to obtain a good mass spectrum of the analyzed sample.

## Acknowledgments

This work is supported by the French National Research Agency in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02), through the funding of the "Origin of

This article is protected by copyright. All rights reserved.

Life" project of the Univ. Grenoble-Alpes and the French Space Agency (CNES) under their Exobiology and Solar System programs. Cédric Wolters acknowledges a PhD fellowship from CNES/ANR (ANR-16-CE29-0015 2016-2021). The paper has been read by a native English speaker and many sentences have been rewritten. Authors acknowledge the reviewers for their feedbacks that helps improve the article.

## Bibliography

1. Creek DJ, Jankevics A, Breitling R, Watson DG, Barrett MP, Burgess KEV. Toward Global Metabolomics Analysis with Hydrophilic Interaction Liquid Chromatography–Mass Spectrometry: Improved Metabolite Identification by Retention Time Prediction. *Anal Chem*. 2011;83(22):8703–8710.
2. Guo X, Long P, Meng Q, Ho C-T, Zhang L. An emerging strategy for evaluating the grades of Keemun black tea by combinatory liquid chromatography-Orbitrap mass spectrometry-based untargeted metabolomics and inhibition effects on  $\alpha$ -glucosidase and  $\alpha$ -amylase. *Food Chem*. 2018;246:74–81.
3. Scigelova M, Makarov A. Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *PROTEOMICS*. 2006;6(S2):16–21.
4. Kalli A, Hess S. Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. *PROTEOMICS*. 2012;12(1):21–31.
5. Kalli A, Smith GT, Sweredoski MJ, Hess S. Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers. *Journal of Proteome Research*. 2013;12(7):3071–3086.
6. Kilpatrick LE, Kilpatrick EL. Optimizing High-Resolution Mass Spectrometry for the Identification of Low-Abundance Post-Translational Modifications of Intact Proteins. *Journal of Proteome Research*. 2017;16(9):3255–3265.
7. Taguchi R, Ishikawa M. Precise and global identification of phospholipid molecular species by an Orbitrap mass spectrometer and automated search engine Lipid Search. *J Chromatogr A*. 2010;1217(25):4229–4239.
8. Rombouts C, De Spiegeleer M, Van Meulebroek L, De Vos WH, Vanhaecke L. Validated comprehensive metabolomics and lipidomics analysis of colon tissue and cell lines. *Anal Chim Acta*. 2019;1066:79–92.
9. Sun C, Shatyck W, Cuss CW, et al. Characterization of Naphthenic Acids and Other Dissolved Organics in Natural Water from the Athabasca Oil Sands Region, Canada. *Environ Sci Technol*. 2017;51(17):9524–9532.
10. Kim C, Ryu H-D, Chung EG, Kim Y. Determination of 18 veterinary antibiotics in environmental water using high-performance liquid chromatography-q-orbitrap combined with on-line solid-phase extraction. *J Chromatogr B*. 2018;1084:158–165.
11. Paishintsev J, Vaikinen A, Lipponen K, et al. Desorption atmospheric pressure photoionization high-resolution mass spectrometry: a complementary approach for the chemical analysis of atmospheric aerosols: DAPPI-HRMS for analysis of atmospheric aerosols. *Rapid Commun Mass Spectrom*. 2015;29(13):1233–1241.
12. Riva M, Ehn M, Li D, et al. CI-Orbitrap: An Analytical Instrument To Study Atmospheric Reactive Organic Species. *Anal Chem*. 2019;91(15):9419–9423.
13. Roveretto M, Li M, Hayeck N, et al. Real-Time Detection of Gas-Phase Organohalogens from Aqueous Photochemistry Using Orbitrap Mass Spectrometry. *ACS Earth Space Chem*. 2019;3(3):329–334.
14. Urai M, Kasuga I, Kurisu F, Furumai H. Molecular characterization of dissolved

- organic matter in various urban water resources using Orbitrap Fourier transform mass spectrometry. *Water Science and Technology: Water Supply*. 2014;14(4):547–553.
15. Hoegg ED, Barinaga CJ, Hager GJ, Hart GL, Koppenaal DW, Marcus RK. Isotope ratio characteristics and sensitivity for uranium determinations using a liquid sampling-atmospheric pressure glow discharge ion source coupled to an Orbitrap mass analyzer. *J Anal At Spectrom*. 2016;31(12):2355–2362.
  16. Danger G, Orthous-Daunay F-R, de Marcellus P, et al. Characterization of laboratory analogs of interstellar/cometary organic residues using very high resolution mass spectrometry. *Geochim Cosmochim Acta*. 2013;118:184–201.
  17. Naraoka H, Hashiguchi M, Sato Y, Hamase K. New Applications of High-Resolution Analytical Methods to Study Trace Organic Compounds in Extraterrestrial Materials. *Life*. 2019;9(3):62–73.
  18. Orthous-Daunay F-R, Piani L, Flandinet L, et al. Ultraviolet-photon fingerprints on chondritic large organic molecules. *GEOCHEMICAL JOURNAL*. 2019;53(1):21–32.
  19. Briois C, Thissen R, Thirkell L, et al. Orbitrap mass analyser for in situ characterisation of planetary environments: Performance evaluation of a laboratory prototype. *Planetary and Space Science*. 2016;131:33–45.
  20. Arevalo R, Selliez L, Briois C, et al. An Orbitrap-based laser desorption/ionization mass spectrometer designed for spaceflight. *Rapid Commun Mass Spectrom*. 2018;32(21):1875–1886.
  21. Selliez L, Briois C, Carrasco N, et al. Identification of organic molecules with a laboratory prototype based on the Laser Ablation-CosmOrbitrap. *Planetary and Space Science*. 2019;170:42–51.
  22. MacDougall, Crummett WB, et al. Guidelines for data acquisition and data quality evaluation in environmental chemistry. *Anal Chem*. 1980;52(14):2242–2249.
  23. He C, Hörst SM, Riemer S, Sebree JA, Pauley N, Vuitton V. Carbon Monoxide Affecting Planetary Atmospheric Chemistry. *The Astrophysical Journal*. 2017;841(2):L31.
  24. Pémot P, Carrasco N, Thissen R, Schmitz-Afonso I. Tholinomics—Chemical Analysis of Nitrogen-Rich Polymers. *Anal Chem*. 2010;82(4):1371–1380.
  25. Hörst SM. Post-Cassini Investigations of Titan Atmospheric Chemistry. University of Arizona; 2011.
  26. Hörst SM, Yelle RV, Buch A, et al. Formation of Amino Acids and Nucleotide Bases in a Titan Atmosphere Simulation Experiment. *Astrobiology*. 2012;12(9):809–817.
  27. Somogyi Á, Smith MA, Vuitton V, Thissen R, Komáromi I. Chemical ionization in the atmosphere? A model study on negatively charged “exotic” ions generated from Titan’s tholins by ultrahigh resolution MS and MS/MS. *Int J Mass Spectrom*. 2012;316–318:157–163.
  28. Maillard J, Carrasco N, Schmitz-Afonso I, Gautier T, Afonso C. Comparison of soluble and insoluble organic matter in analogues of Titan’s aerosols. *Earth Planet Sci Lett*. 2018;495:185–191.
  29. Orthous-Daunay F-R, Thissen R, Vuitton V. Measured mass to stoichiometric formula through exhaustive search. *Proceedings LAU Symposium S350*. (Accepted). 2019;
  30. Kido Soule MC, Longnecker K, Giovannoni SJ, Kujawinski EB. Impact of instrument and experiment parameters on reproducibility of ultrahigh resolution ESI FT-ICR mass spectra of natural organic matter. *Org Geochem*. 2010;41(8):725–733.
  31. Blackburn JWT, Kew W, Graham MC, Uhrin D. Laser Desorption/Ionization Coupled to FTICR Mass Spectrometry for Studies of Natural Organic Matter. *Anal Chem*. 2017;89(8):4382–4386.

This article is protected by copyright. All rights reserved.

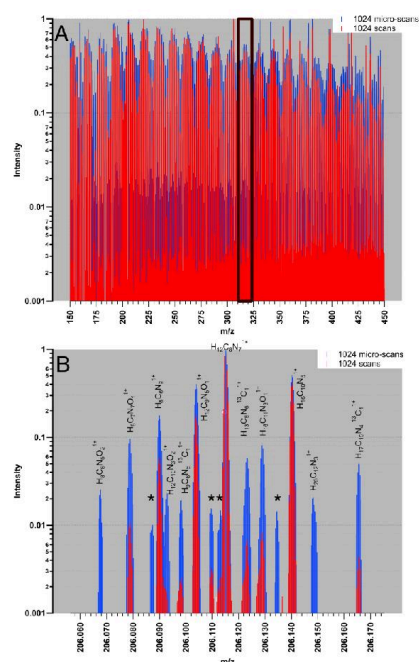


Figure 1: (A) Mass spectrum of peaks profile data for two complementary Orbitrap operations. Data are normalized around  $m/z$  206.114. The micro-scans operation of the Orbitrap present a slightly highest intensity trend at high masses. (B) Zoom on the  $m/z$  206.1 peak cluster where we observe with micro-scans an added diversity on both the left and the right of the cluster. This particular example is representative of all peak clusters. Detected but non-attributed peaks using  $\text{CHNO}^{13}\text{C}$  are highlighted with a “\*”. Two  $^{13}\text{C}$  are also attributed but are not considered afterward as with every detected  $^{13}\text{C}$  stoichiometric formula.

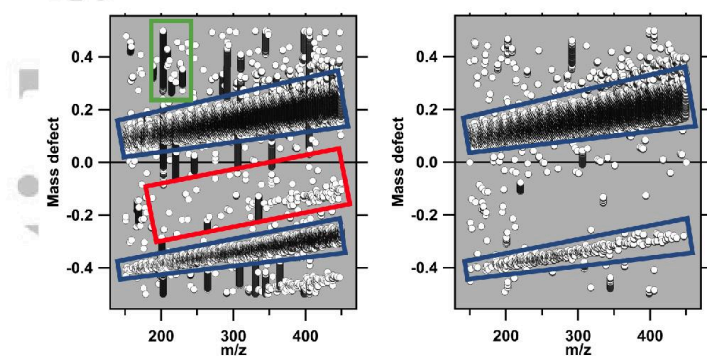
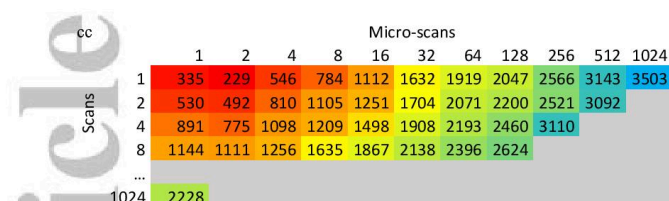


Figure 2: Left – Mass defect diagram from profile data for 1024 scans with 1 micro-scan experiment; 53,238 points. Right – Mass defect diagram from profile data for 1 scan with 1024 micro-scans experiment; 45,198 points. The blue box is sample ions, the green box is a radio artifact example and the red box is an example of random points.



**Figure 3:** Total number of stoichiometric formulae attributed per analysis; color code from red to blue for the lowest to the highest number. Grey color is an experiment that has not been acquired. The average time per micro-scans or scan is 1.5s. This average time is mostly dependent on the sample concentration and the AGC target.

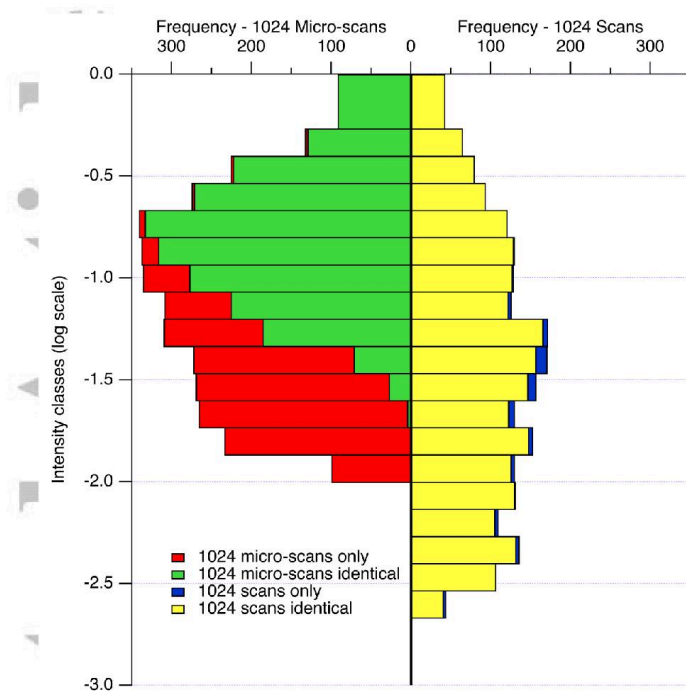


Figure 4: Intensity histogram for attributed stoichiometric formulae using 1024 scans and 1024 micro-scans. Identical stoichiometric formulae between each analysis are indicated in yellow and green, respectively.

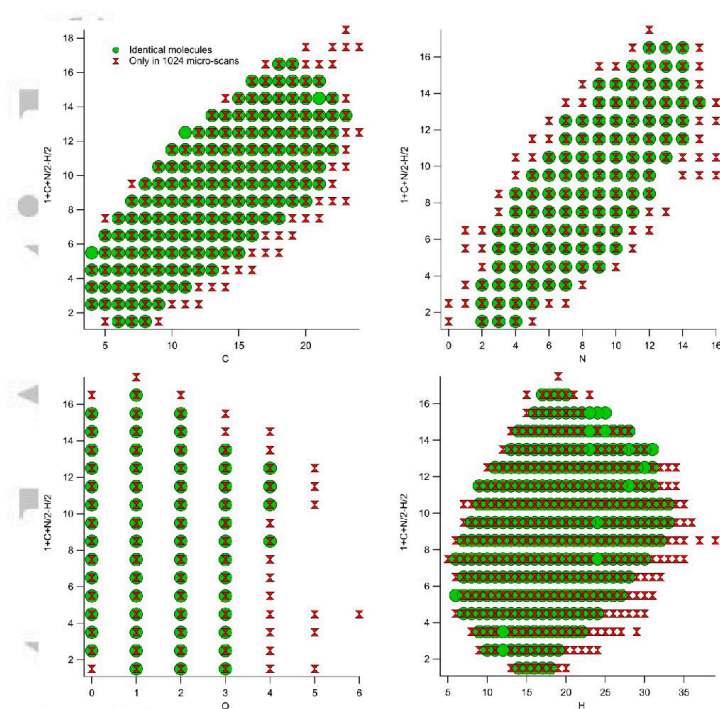


Figure 5: Double Bond Equivalent (DBE) versus the carbon, nitrogen, oxygen and hydrogen number for identical stoichiometric formulae in both operations and formulae only in 1024 micro-scans. When using the micro-scans method, the added diversity is everywhere.