

Unusual Applications of Kendrick Plots: Recalibration and Tolerance

Thierry N. J. Fouquet* and Orlando Cabarcos

Surgical Analytical & Materials Science (SAMS), Bausch & Lomb, 1400 N Goodman St, Rochester NY 14609, USA

Received November 16, 2023, Revised December 3, 2023, Accepted December 14, 2023

First published on the web December 31, 2023; DOI: 10.5478/MSL.2023.14.4.173

Abstract : Kendrick plots offer an alternative visualization of mass spectral data which reveals ion series and patterning by turning a mass spectrum into a map, plotting the fractional mass (wrongly called mass defect) as a function of mass-to-charge ratios and ion abundances. Although routinely used for polymer mass spectrometry, two unreported applications of these Kendrick plots are proposed using the program “kendo2”: the graphical recalibration of a mass spectrum via the simulation of a theoretical fractional mass and a multi-segment fit; and the rapid evaluation of scan-to-scan variation of accurate mass measurements used as tolerances for the blank subtraction of UPLC-MS data files. Both applications are compatible with any type of high-resolution MS data including LC/GC-MS(/MS).

Keywords : Kendrick analysis, polymer mass spectrometry, LC-MS, calibration, blank subtraction, data processing

Introduction

A mass spectrum is often difficult to interpret due to the overwhelming number of possible overlapping features that can be introduced via different physical, chemical and analytical effects. Elemental isotopic patterns, in-source fragmentation, the formation of adducts and dimers, isobaric species, ion series (e.g. salt clusters) and multiple charge states generate innumerable peaks. High-resolution mass analyzers (e.g. reflectron or multi-turn Time-of-Flight, Orbitrap, Fourier-Transform Ion Cyclotron Resonance) do not mitigate the difficulty of interpreting a complex mass spectrum but increase it by increasing the number of features thanks to their intrinsic resolving power capable of differentiating extremely small variations of m/z (e.g. resolving power = $m/\Delta m = 60,000$ at m/z 200 $\rightarrow \Delta m \sim 3.3$ mDa).

A Kendrick plot (named after E. Kendrick¹) turns the two-dimensional mass spectrum displaying peaks (mass-to-charge ratios m/z along the x -axis, abundance along the y -axis) into a three-dimensional map displaying bubbles, with the fractional mass of m/z ratios (Equation 1,[‡] incorrectly known

as the mass defect or mass excess) as a new y -axis, and peak abundances being retained via the size of the bubbles.²

$$\text{fractional } m/z = m/z - \text{round}(m/z) \quad (1)$$

By definition from the International Union of Pure and Applied Chemistry (IUPAC), mass of atoms, molecules and ions are calculated based on the mass of isotope 12 of carbon, arbitrarily set at 12 Da / amu.³ However and if clearly stated, any exact mass noted R can be used as a new reference mass (or basis), depending on the expected / possible repeating moiety that an analyst may expect in a mass spectrum (such as repeating units of a (co)polymer, an isotopic pattern such as ³⁷Cl-³⁵Cl, the loss of a neutral fragment, a salt adduct, etc). The classical Kendrick analysis uses $R = 14.01565$ (mass of CH₂) as the new reference mass set at 14, the closest integer from its accurate mass in the IUPAC scale.⁴ The rescaled mass for all isotopes and molecules is calculated using a simple cross-multiplication (Equation 2):

$$m/z_{\text{Kendrick}} = m/z_{\text{IUPAC}} \cdot \frac{\text{round}(R)}{R} \quad (2)$$

In this mass scale, an entire $-(\text{CH}_2)_n-$ ion series has the same fractional mass regardless of n , and is thus displayed as a horizontal line in the associated Kendrick plot.

Recently, a “resolution-enhanced” or “dynamic” Kendrick analysis has been introduced using a refined definition as defined in Equation 3.⁵

$$m/z_{\text{Kendrick}} = m/z_{\text{IUPAC}} \cdot Z \cdot \frac{X}{R} \quad (3)$$

with $Z \in \mathbb{N}$, and $x \in \mathbb{R}$ so $Z \cdot x \in \mathbb{N}$.

Open Access

*Reprint requests to Thierry N. J. Fouquet

<https://orcid.org/0000-0002-9473-9425>

E-mail: thierry.fouquet@bausch.com

All the content in Mass Spectrometry Letters (MSL) is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All MSL content is published and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

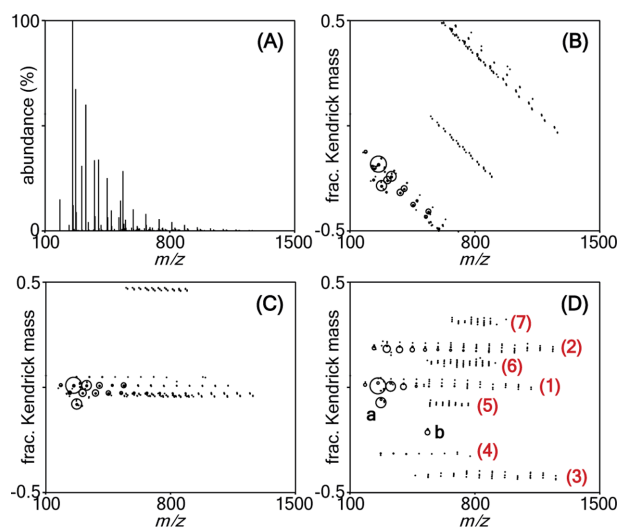


Figure 1. A) ESI-HRMS of polyTHF. B) IUPAC Kendrick plot (basis: ^{12}C , $R=12$). C) Static Kendrick plot (basis: $\text{C}_4\text{H}_8\text{O}$, $R=72.0575 \rightarrow 72$). D) dynamic Kendrick plot (basis: $\text{C}_4\text{H}_8\text{O}$, $R=72.0575 \rightarrow 430$) Series (1)–(7) are linear oligomers of polyTHF $\text{OH}(\text{C}_4\text{H}_8\text{O})_n\text{H}$ adducted with H^+ , NH_4^+ , Na^+ , K^+ , 2H^+ , H^+NH_4^+ and 2NH_4^+ , respectively.

This new definition which sets the mass of the reference to *some* integer $Z \cdot x$ and not *the closest* integer round (R) yields twice the advantage of the original Kendrick equation by allowing us to a) take into account the charge state of the ions in a mass spectrum via the parameter Z ,⁶ and b) generate an *infinite* number of Kendrick plots with a user-defined separation/clustering capabilities along the y -axis by varying $Z \cdot x$, whereas the classical definition of a Kendrick analysis generates only one plot per R . Figure 1 depicts the infusion electrospray ionization (ESI) high-resolution mass spectrum of a poly(tetrahydrofuran) sample (polyTHF, repeating moiety $\text{C}_4\text{H}_8\text{O}$), the Kendrick plot in the IUPAC mass scale displaying oblique line and aliasing, the static Kendrick plot using $R = 72.0575 \rightarrow 72$ with horizontal alignments but poor separation along the y -axis, and *one* possibility of a dynamic Kendrick analysis using $R = 72.0575 \rightarrow 430$ (i.e. $Z = 3$ and $x = 143.333\dots$ in Equation 3) showing at a glance different polyTHF ion series (1–7) as horizontal lines spreading along the entire y -axis, as well as two unexpected additives as outliers (a, b) with an evident gain in terms of visualization, rapid assignment of series, and potential extraction of a targeted series as compared to the mass spectrum.

Kendrick analyses have become an invaluable tool for the visualization and interpretation of mass spectral data from polymeric samples or any samples with repeating moieties such as salts or metallic clusters, generated by infusion or chromatography – mass spectrometry with a variety of ion sources, single stage or tandem mass analyses, etc. However, unusual and undescribed applications still

exist, as reported here for the graphical recalibration of mass spectral data and the evaluation of the best tolerance for blank subtraction based on the instant visualization of scan-to-scan dispersion of mass measurements.

Experimental

PolyTHF M_n 250 g mol^{-1} standard purchased from Sigma Aldrich was dissolved in tetrahydrofuran at $\sim 10 \text{ mg mL}^{-1}$, then diluted in acetonitrile to $\sim 10 \text{ mg mL}^{-1}$. Infusion ESI-MS and ultra high performance liquid chromatography UPLC-MS data were recorded using a 6560 IM QTOF mass spectrometer coupled to an Infinity 1290 chromatography system (Agilent, resolving power at m/z 200 \sim 25k). Data files were converted from Agilent's proprietary format .d to .mzML using MSConvert 3.0.23060-c109104 (filters: peak-picking “vendor, MS1” and threshold “absolute intensity” = 200 cts). Kendrick analyses were computed using “kendo2” v 2.1.2, a standalone program written by T. N. J. Fouquet with Visual Studio 2022 v. 17.3.0 available online for download.⁷

Results and Discussion

Recalibration of mass spectral data

The theoretical m/z values for the sodiated linear polyTHF 3-mer $\text{OH}(\text{C}_4\text{H}_8\text{O})_3\text{H}$ and the protonated polyTHF 11-mer $\text{OH}(\text{C}_4\text{H}_8\text{O})_{11}\text{H}$ are 257.1723 and 811.6505, although the ions are detected at 257.1732 and 811.6532, i.e. small yet existing errors of 0.9 and 2.7 mDa in the ESI mass spectrum of polyTHF on top of Figure 2. Instead of computing *each* theoretical m/z for *each* congener of a polymer series, plotting a Kendrick plot using its repeating unit (s) for the change of basis instantly reveals the quality of mass measurements and/or the extent of miscalibration via the deviation from horizontality. As illustrated in Figure 2, the red dots corresponding to the monoisotopic peaks of $[\text{OH}(\text{C}_4\text{H}_8\text{O})_n\text{H} + \text{H}]^+$ in the raw mass spectrum clearly line up obliquely, deviating from the theoretical fractional mass of H_3O^+ which is the *only* value to be simulated to describe the entire series (= end-groups and adducted ion). The average error for the entire series as compared to the simulation is instantly computed by the program kendo2 at ~ 1.4 mDa. Beyond the visualization of the miscalibration, the plots can be used to *recalibrate* the mass spectrum. A simple multi-segment curve $[F_i, F_{i+1}]$ is drawn in the Kendrick plot using the program to approximately fit the targeted ion series. For each segment $[F_i, F_{i+1}]$ of the fit, the correction of all m/z values within this segment is as follows:

$$m/z_{\text{recalibrated}} = m/z_{\text{initial}} - \text{correction} \quad (4)$$

$$\begin{aligned} \text{with correction} = & \frac{R}{Z \cdot x} \cdot (\text{frac.mass}_{\text{simulation}} - \text{frac.mass}_{F_i}) \\ & + \text{slope}_{[F_i, F_{i+1}]} \cdot (m/z_{F_i} - m/z_{\text{initial}}) \end{aligned}$$

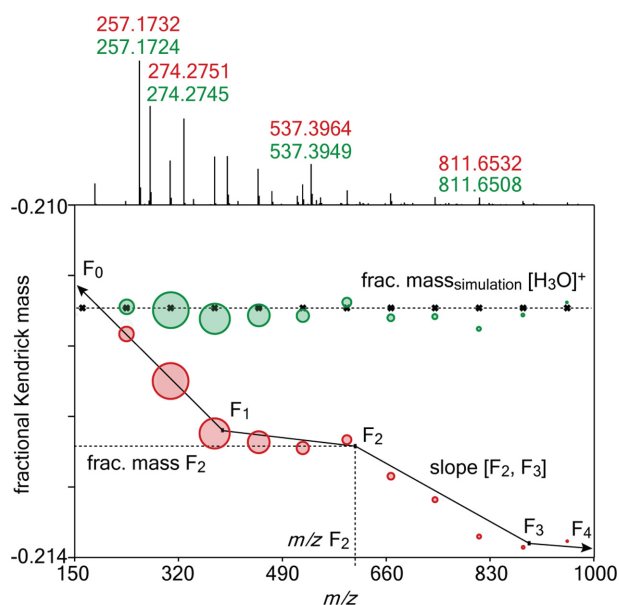


Figure 2. Kendrick plot from the ESI mass spectrum of polyTHF using $R = 72.0575$, $Z = 1$ and $x = 69$ zoomed around the monoisotopic peaks of $[\text{H}(\text{C}_4\text{H}_8\text{O})_n\text{OH} + \text{H}]^+$ series. Red and green dots represent raw and recalibrated ions. The crosses represent the theoretical fractional Kendrick mass of $[\text{HOH} + \text{H}]^+$. A simple fit for recalibration is displayed with segments $[F_i, F_{i+1}]$. Top: corresponding mass spectrum with raw (red) and recalibrated (green) accurate mass measurements.

Applying this formula to all the segments provides a recalibrated Kendrick plot / mass spectrum with the oligomers of polyTHF being realigned horizontally along the simulation depicted as green dots in Figure 2. The average error of the entire series is now 0.1 mDa. The quality of the recalibration can be improved to near perfection by increasing the number of segments of the fit at the user's discretion. Obviously, all the m/z values of the mass spectrum are recalibrated as revealed by the change of m/z values for two additives found in the mass spectrum. Lauryldiethanolamine ($\text{C}_{16}\text{H}_{35}\text{NO}_2$) and dilauryl thiodipropionate ($\text{C}_{30}\text{H}_{58}\text{O}_4\text{S}$) are detected at m/z 274.2751 and m/z 537.3964 in the raw mass spectrum, i.e. errors of 1.0 and 1.6 mDa. After the graphical recalibration, their m/z values are measured at m/z 274.2745 (error: 0.4 mDa) and m/z 537.3949 (error: 0.1 mDa, top of Figure 2). Such a recalibration procedure can be applied to a single mass spectrum, an entire LC/GC-MS run or a batch of data files.[†]

Evaluation of scan-to-scan variability used as tolerance for blank subtraction

Commercial programs generate centroid mass spectra with one averaged m/z value per ion, which corresponds to one bubble per ion in a Kendrick plot (Figure 1, Figure 2, red and green bubbles and Figure 3B, black bubbles). However, mass spectra are initially composed of a succession of

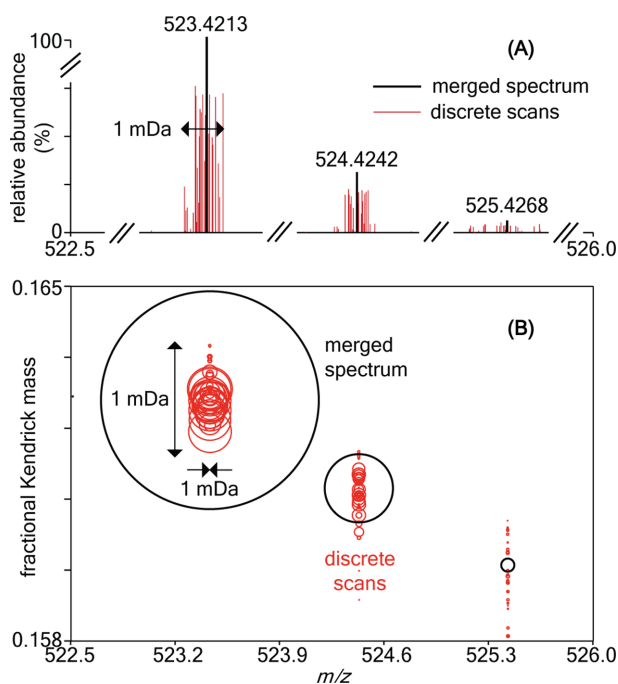


Figure 3. (A) Merged ESI mass spectrum from 30 scans zoomed around the isotopic pattern of $[\text{H}(\text{C}_4\text{H}_8\text{O})_7\text{OH} + \text{H}]^+$ (black line) and original data displaying peaks from each discrete scan (red lines); (B) Kendrick plot from (A) using $R = 72.0575$, $Z = 1$ and $x = 72$.

scans (e.g. with the QTOF used in this study, acquisition rate of 10 Hz, 3 seconds of acquisition = 30 scans = 30 values of m/z per ion). It may be useful to keep all the peak-picked scans rather than one average value to visualize the extent of 'scan-to-scan' dispersion, revealing the instrument accuracy and stability. Constituting scans used to generate the merged spectrum in Figure 3A are displayed in red. This representation is very inconvenient due to the need for peak-by-peak magnification and multi-scale along the x-axis. A Kendrick plot solves this issue as shown in Figure 3B, with the constituting scans seen as bubbles overlapping along the x-axis but clearly spreading along the y-axis, which reveals the scan-to-scan dispersion of ~1 mDa at a glance.

That graphical representation and the obvious possibility for a program to measure it are of high interest for any data processing involving a tolerance parameter, such as blank subtraction or binary comparison. Figure 4A and 4B show the total ion chromatogram (TIC) from the UPLC-MS analysis of a blank and the same polyTHF sample. Two types of "background signal" are clearly seen in both TICs: discrete LC peaks with defined retention times corresponding to impurities injected with the sample on the column at $t = 0$; and a diffuse signal seen as a baseline rise due to impurities in the mobile phases constantly eluting through the column. Using the base peak chromatogram (BPC) instead of the

TIC eliminates the diffuse background noise but does not remove the discrete peaks as depicted in Figure 4C. The subtraction of the blank TIC to the sample TIC is a better solution to get a clean LC-MS fingerprint that may reveal minor hidden signals (either co-eluting with the discrete background signals or lost in the diffuse background noise). The program kendo2 a) loads both blank and sample mzML files, b) decompresses and subtracts one or multiple scans combined into one from the blank file to each scan of the sample file, c) calculates a subtracted TIC and d) store the raw and blank-subtracted data in a new mzML file. The process requires three parameters: a drift time to account for the variation of retention time between the blank and the sample chromatograms (= the number of scans to be combined), and two tolerance in mDa to account for a) the variation scan-to-scan of the m/z values due to the instrument

variability *within* the blank or the sample file, and b) the variation of m/z *between* the blank and the sample due to possible drift of calibration over time. Choosing the best values for the last two parameters can be either arbitrary or based on a proper evaluation from the data relying on the capability of the Kendrick plots to reveal the minuscule variations of m/z ratios (Figure 3).

Figure 4D shows a Kendrick plot generated from the scans 5670-5720 from the TIC of the polyTHF sample (black bubbles) and blank (red bubbles) corresponding to one discrete background peak. The extent of separation along the y-axis can be adjusted at will by changing the values of Z and x (Equation 3) to separate the bubbles/ions or interest. The scans are *not* averaged, so every m/z value from every scan is displayed as a bubble (50 scans, ~ 1200 points/scan, $\sim 60,000$ points). Zooming in around the largest

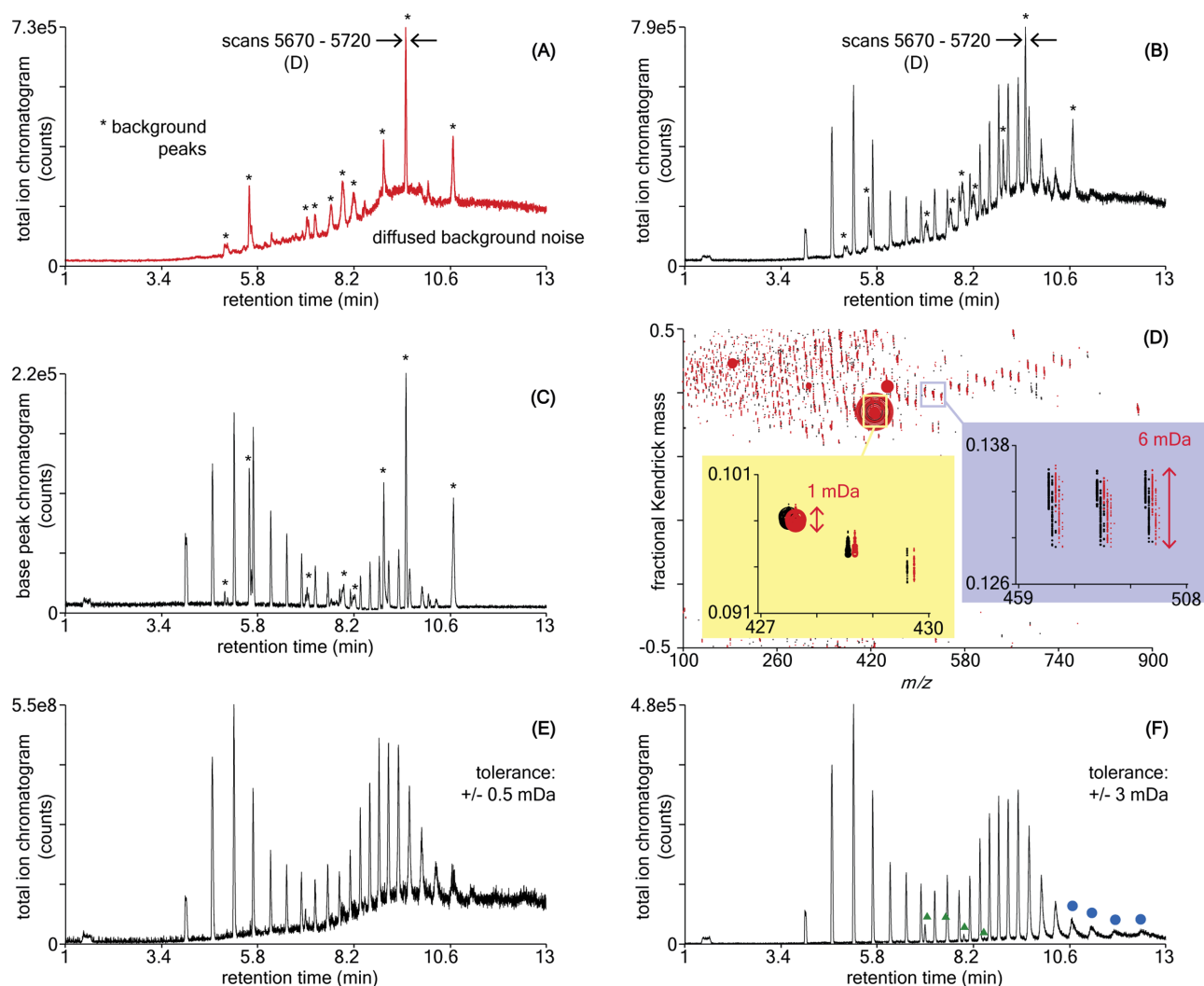


Figure 4. UPLC-MS(TIC) of (A) a blank and (B) polyTHF; (C) BPC from (B); (D) Kendrick plot from the scans 5670-5720 of (A) (red dots) and (B) (black dots) using $R = 72.0575$, $Z = 1$ and $x = 144$. Insets: magnified plot around a background peak (yellow) and diffused background signal (purple); (E, F) Blank-subtracted UPLC-MS(TIC) from (B) using a tolerance of ± 0.5 mDa and ± 3 mDa.

bubbles (yellow shape, most intense ion in the mass spectrum), a dispersion of ~ 1 mDa for the m/z values of the blank and of the sample is readily visualized along the y-axis. At the same time, the variation of m/z values from blank to sample is evaluated at less than 0.1 mDa. Using these two tolerances of ± 0.5 mDa and ± 0.05 mDa (instantly computed by the program), all the discrete background peaks are removed in the blank-subtracted TIC (Figure 4E). However, the diffuse background noise has not been reduced with a baseline still rising during the elution. Back to the Kendrick plot and zooming in around a minor signal, the greater dispersion of m/z values scan-to-scan is instantly evaluated at ~ 6 mDa by the program, and sent as a tolerance of ± 3 mDa for the blank subtraction. A new TIC is generated (Figure 4F) free of both the discrete background peaks and the diffuse background noise. The baseline of the chromatogram is near 0 throughout the run, a minor cyclic polyTHF series is revealed (green triangles), and the longest polyTHF chains eluting in the last minutes are clearly visible (blue dots). Interestingly, these late eluters were lost in the BPC (Figure 4C) due to the fact that the background signal is more intense than the signal of interest in that region. From Figure 4B to Figure 4F, a blank subtraction with rationally rather than arbitrarily chosen parameters provides an incredibly cleaner LC-MS profile, containing considerably less points (from 2M to 250k) that will speed up the search for knowns via library search, grouping of features and deconvolutions.

Conclusions

Dynamic Kendrick plots are shown to be a valuable graphical tools for the (re)calibration of mass spectral data and the evaluation of scan-to-scan variability which can be conveniently used as a well-defined input tolerance for blank subtraction or other processing relying on tolerances in place of arbitrarily chosen values. Through this letter, the author strongly encourages the implementation of blank subtraction (the procedure is compatible with GC-MS, GC/LC-MS/MS, MSⁿ and data independent analyses such as SWATH or MS^E for any type of HR mass spectrometer[†]),

and militates for the incorporation of real-time Kendrick plots capabilities in the *acquisition* programs for a user-friendly visualization of data.

Acknowledgments

TNJ Fouquet warmly acknowledges Bausch & Lomb for the privileged access to the UPLC-QTOF and Orbitrap instruments.

Notes

‡Other definitions exist using “floor” or “ceiling”, of no other consequence than the range of the y-axis (0-1 or -1-0 instead of -0.5-0.5).

†Electronic Supplementary Information (ESI) available: [List of abbreviations, blank subtraction with evaluation of tolerance for UPLC-MS data using an Orbitrap with resolving power of $\sim 120k$, two brief video recordings of the recalibration procedure and the blank subtraction following the evaluation of the tolerance via Kendrick plots in real-time using kendo2]. See DOI: 10.5478/MSL.2023.14.4.160

References

1. Hsu C. S.; Qian K. N.; Chen Y. N. C. *Anal Chim Acta*. **1992**, 264, 79-89. DOI: 10.1016/0003-2670(92)85299-L
2. Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. *Anal. Chem.* **2001**, 73, 19, 4676–4681. <https://doi.org/10.1021/ac010560w>
3. Cohen, E.R.; Cvitas, T.; Frey, G. J. et al. Quantities, units and symbols in physical chemistry 3rd edition. IUPAC RSC Publishing, **2007**
4. Kendrick, E. *Anal Chem.* **1963**, 35, 2146-2154. <https://doi.org/10.1021/ac60206a048>
5. Fouquet, T. N. J. *J. Mass Spectrom.* **2019**, 54, 933-947. <https://doi.org/10.1002/jms.4480>
6. Fouquet T.; Cody R. B.; Ozeki Y. et al. *J. Am. Soc. Mass Spectrom.* **2018**, 29, 1611-1626. <https://doi.org/10.1007/s13361-018-1972-4>
7. Fouquet, T.; Nakamura, S.; Sato, H. Kendo and Kendo2; See <https://kendo2.wixsite.com/kendo2> (October 2023).