Utilization of Fine Isotope Structures to Facilitate Unknown Identification Based on High Mass Accuracy and High Spectral Accuracy

Ming Gu, Yongdong Wang, Hongliang Xu
Cerno Bioscience, Norwalk, CT

Overview

➢Spectral accuracy was evaluated for eight small pharmaceuticals acquired with Orbitrap with resolving power of 100K.

➢Spectral accuracy as new metric has been used to successfully identify unknown compounds with high confidence.

➢Fine isotope structures obtained with resolving power of 100K by Orbitrap make huge difference for effective formula identification

Introduction

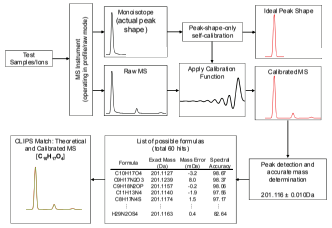
Orbitrap has become an important tool for qualitative mass spectrometry applications due to its achievable resolving power of up to 100K and easily attainable high mass accuracy of better than 3 ppm. While such performance may be sufficient to determine unknown formulas where the m/z is small (<250Da) or with additional constraints applied, e.g., limited chemical elements and their upper and lower bounds. For the identification of most real pharmaceutical impurities, however, mass accuracy of 3ppm alone is not good enough for high confidence identification. Here we describe a new method based on mass spectral peak shape calibration by utilizing isotope fine structures to achieve confident and reliable formula identification of true unknowns.

Methods

➢**Data Acquisition:** LC/MS analysis was performed in a LTQ/Orbitrap mass spectrometer at different resolving power 100K for a mixture of 8 commercially available compounds as listed in table 1.

➢ **Data Analysis:** All mass spectral data acquired from the Orbitrap were exported as ASCII data format and analyzed by **sCLIPS** (self Calibrated Line-shape Isotope Profile Search) through MassWorks (Cerno Bioscience, Norwark, CT). **sCLIPS** 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. 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 (SA).

Fig 1. MassWorks sCLIPS Formula ID



Results and Discussion

Spectral accuracy has been proved to be a powerful metric to eliminate false positive formulas during elemental component determination through peak shape calibration. The fundamental difference between current approach and others is that this method uses calibrated profile spectra for exact isotope modeling, while others employ either overly simplified centroid spectra or profile spectra with assumed peak shape function. Previous studies demonstrated that up to 99% of the incorrect formulas could be removed based on spectral accuracy for a group of compounds of molecular weights>800 amu with mass tolerance of 2 ppm. Focused on smaller molecules having formula weights less than 500 amu, current studies achieved similar results. Through the peak shape calibration, **sCLIPS** successfully identified all the "unknowns" as either number hit or number two hit based on their high spectral accuracy, which is summarized in Table 1. All the search was conducted with very generous conditions including most common elements without any restriction on their upper and lower bounds. Without the help from spectral accuracy, it will be extremely challenging to obtain positive identification of the compounds.

In many cases, the fine isotope structures of the small molecules exhibited distinct patterns to facilitate unknown identification. For example, in the formula search for Prednisone, the correct formula, C21H27O5 [M+H], was ranked as a top hit from a total of 44 candidates. The top four formulas and their fine isotope structures are illustrated in Fig 2 to demonstrate the power of spectral accuracy for the identification of true unknown compounds. The true compound, Prednisone, contains the elements of carbon, hydrogen, oxygen only. It's A+1 peak consists ^{13}C, while its A+2 peak has unresolved ^{18}O and $^{13}C^{13}C$. There has been great isotope pattern match observed between theoretically calculated (in black) and calibrated (in red) spectra for Prednisone as indicated by its high spectral accuracy of 98.1%. The second hit has slightly poor spectral accuracy than the top hit largely due to less carbons in the molecule of C19H26O3F3. Both the third and forth hits contain nitrogen atoms, which cause two mismatches to result in much poorer spectral accuracy than the top one hit. The first mismatch happened in ^{13}C peak as less carbons contained in the molecules due to the inclusion of nitrogen atoms. At the same time, ^{15}N peaks from both C18H27N2F3P and C17H23N4F4 introduced the second mismatch since the real compound does not have such a peak. While the dominant spectral errors were resulted from A+1 peak which accounted for 23.3% based on its peak height, marked differences are also observed in A+2 peak.

Compared with unresolved $^{13}C^{13}C\&^{18}O$ from C21H27O5, the A+2 peak, for example, from C17H23N4F4 has distinct $^{13}C^{15}N$ and $^{13}C^{13}C$ peaks, contributing to overall spectral errors of the fourth formulas. However, if mass accuracy is the only metric to be used for the identification of C21H27O5, it would require the mass accuracy at low ppb level to distinguish C21H27O5 (m/z = 359.1858) from C17H23N4F4 (m/z = 359.1859) since their m/z values differ by only 99.1 ppb.

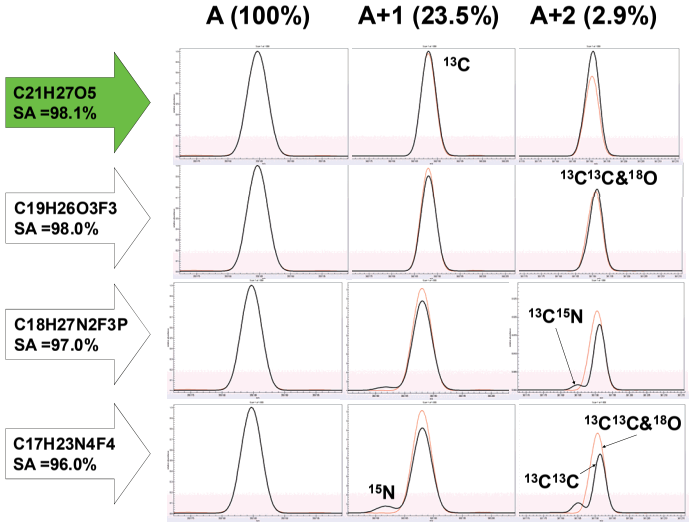
Table 1. Summary of Formula ID of Pharmaceuticals by Spectral Accuracy

Formula search was conducted with wide open conditions with mass tolerance of 5 ppm, double bond equivalent from -1 to 50, even electron state. Possible elements of C H N O S Cl F P were included and have their lower bounds set to zero and upper bounds set to the maximum which is equal to the molecular weight of searched molecules divided by the atomic weight of an element. For example, when searching for Bupropion at m/z 240, the upper bound for carbon will be 240 divided by 12 = 20

Name	Formula	Mono Isotope	Mass Error (PPM)	Spectral Accuracy	Rank	Total Formulas
Bupropion	C13H19ONCl	240.1155	3.0	95.8	1	9
Sulfamethoxazole	C10H12O3N3S	254.0599	2.9	98.9	2	25
Propranolol	C16H22O2N	260.1651	3.3	98.0	1	5
Zolmitriptan	C16H22O2N3	288.1712	2.8	98.3	1	10
Alprazolam	C17H14N4Cl	309.0907	2.9	96.0	1	36
Prednisone	C21H27O5	359.1858	3.2	98.1	1	44
Terfenadine	C32H42O2N	472.3216	3.3	98.5	2	54
Loperamide	C29H34O2N2Cl	477.2309	3.1	97.6	1	157

Fig 2. Formula Search for Prednisone Powered by Fine Isotope Structures: Top Four Hits by Spectral Accuracy

In order to closely examine the fine isotope structures of A+1 and A+2 peaks, the plots for the peaks are normalized to the largest peak in their mass range. The spectra in black are theoretically calculated and those in red are calibrated (RP=100K).



Conclusions

➢Excellent Spectral accuracy better than or close to 98% was found for most of the molecules under investigation. Spectral accuracy provides a unique and effective way to successfully identify molecular formula of unknown compounds. All the compounds were correctly identified as top one or top two hits by spectral accuracy.

➢The case studies on the identification of Prednisone clearly show that the fine isotope structures play an essential part on separating a true compound from false positives according to spectral accuracy. Mass accuracy at low ppm level alone often is not sufficient to identify true unknowns. The combination of high mass accuracy with high spectral accuracy is much more powerful way to identify unknown formulas than mass accuracy alone.

➢Future work will include the applications of the approach to the determination of true unknowns of real impurities of pharmaceuticals or the unknowns from natural products discovery.

References

(1) Y. Wang, M. Gu. The Concept of Spectral Accuracy for MS. Anal. Chem. 2010, 82, 7055.

(2) J.C.L. Erve, M.Gu, Y. Wang, W. DeMaio, R.E. Talaat. Spectral Accuracy of Molecular Ions in an LTQ/Orbitrap Mass Spectrometer and Implications for Elemental Composition Determination. J. Am. Soc. Mass Spectrom. 2009, 20, 2058.