Identification of Pharmaceutical Impurities by UPLC and a Fast Scanning Quadrupole MS

The escalating need for impurity analysis in pharmaceuticals provides the basis for evaluating fast, simple, and reliable methods of analysis. UPLC provides the speed and resolving power to develop rapid separation methods. A new breed of fast scanning quadrupole MS instruments are robust, reliable, easy-to-use, moderately priced, and well matched to the demand for measuring the narrow UPLC peaks in real-time but traditionally have not been considered as tools for reliable formula ID. By combining these systems with a new calibration available through MassWorks™ CLIPS formula search, these systems can now become reliable workhorse systems for impurity determination.

> ***"Due to the significant resolving power of UPLC, the separation of impurities of simvastatin was obtained in a 9-minute chromatographic analysis."***

Author Details:

Ming Gu[1]
Cerno Bioscience,
Danbury,
CT 06810
USA
info@cernobioscience.com

INTRODUCTION

Pharmaceutical impurities are unwanted components which may come from the processes of organic synthesis, formulation, and storage of pharmaceuticals. Identification of these impurities is of great concern for regulatory agencies because even small amounts of impurities in pharmaceuticals can significantly compromise the efficacy of drug products or cause adverse drug reactions. To identify the molecular structures of these impurities, it is critical to determine their elemental composition. This is usually achieved by accurate mass measurements with high resolution mass spectrometers such as qTOF, FT-ICR, and Orbitrap. With the recent advent of innovative mass spectrometry calibration technology (1-3), high mass accuracy and elemental composition determination (ECD) can be obtained with a unit mass resolution quadrupole mass spectrometer.

This unique calibration technology not only calibrates m/z values like conventional mass calibration does, but also performs mass spectral peak shape calibration such that the calibrated spectral peak shape is symmetrical and can be described by a mathematical function. As a result, the isotope interference of unit mass resolution spectra can be mathematically removed and the center of the monoisotope peak, i.e., the accurate mass, can be measured accurately. More importantly, the same mathematical peak functions generated from the calibration apply to both experimental spectra and theoretically calculated spectra of possible formula candidates, such that exact isotope pattern comparisons can be made between the experimental spectra and the calculated spectra to determine molecular formula (Spectral Accuracy). In this application note, we will demonstrate that a Waters single quadrupole system (Waters Corporation, Milford, MA USA) coupled with a Waters UPLC can be used for impurity identification by applying this calibration technology which is available in MassWorks™ (Cerno Bioscience, Danbury, CT USA) software.

[1.] Cerno Bioscience, Danbury, CT 06810, USA

EXPERIMENTAL

Simvastatin is a hypolipidemic drug belonging to the class of pharmaceuticals called "statins". It is used to control hypercholesterolemia (elevated cholesterol levels) and to prevent cardiovascular disease. Simvastatin is a synthetic derivative of a fermentation product of Aspergillus terreus and was used in this study. The simvastatin was obtained from a commercial source (USP) and the impurities normally found in simvastatin were then separated by ultra high pressure liquid chromatography (UPLC) (Waters, Milford, MA) and detected by a Waters Acquity SQD single quadrupole mass spectrometer (Waters, Milford, MA). All mass spectral data were acquired in profile mode with a scan rate of 1000 amu/s and a mass range from 380 to 520. The mass spectral data files in MassLynx format were read directly by MassWorks software (Cerno Bioscience, Danbury) for accurate mass calibration and formula identification.

RESULTS AND DISCUSSION

Mass spectra acquired at unit mass resolution by definition have overlaps among the adjacent mass spectral isotope peaks. To obtain high mass accuracy and ECD, it is necessary to mathematically deconvolve the peaks and accurately determine their centres. Theoretical treatment of Cerno's approach to the deconvulotion is beyond the scope of this discussion but demonstrated here are the practical procedures of Cerno's patented mass spectral calibration and the results of formula identification for the impurities from the pharmaceutical compound simvastatin.

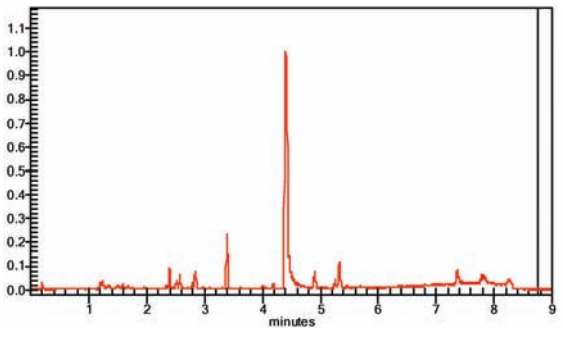

Figure 1. UPLC/MS of simvastatin (RT ~ 4.4 min) and it's impurities.

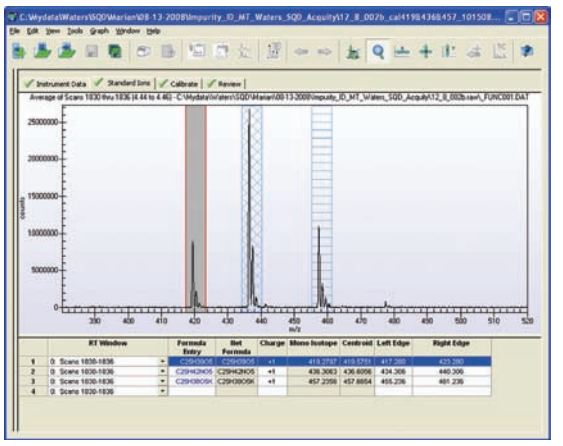

Figure 2. Calibration ions from parent drug simvastatin

Due to the significant resolving power of UPLC, the separation of impurities of simvastatin was obtained in a 9-minute chromatographic analysis. The parent drug simvastatin appears at retention time of 4.4 minute as the most abundant peak while the minor peaks that show up before or after simvastatin are the impurities related to simvastatin (*Figure 1*). As shown in *Figure 2*, simvastatin is ionised in three formations as the protonated [M+H]+, ammonium adducts [M+NH4]+, and potassium adducts [M+K]+ ions and observed at nominal m/z 419, 436, and 457 respectively. Since the elemental composition for the three ions is known, they can be used as standards to perform an internal calibration for the best mass accuracy and Spectral Accuracy.

The MassWorks' calibration requires the input of the molecular formula of the calibration standards instead of m/z values as in the classical mass only calibration. For example, the entire isotope profile of the calibration standard simvastatin is highlighted in gray and selected for mass and peak shape calibration based on the exact m/z value and theoretical isotope distribution provided by the molecular formula of C25H39O5 (*Figure 2*). After the calibration, not only the m/z value is corrected, but also the peak shape (*Figure 3*, calibrated spectra in red) is calibrated to a symmetrical and mathematically defined function. The calibration performance is measured by both mass accuracy and Spectral Accuracy. Relative mass accuracy is less than 5 ppm for all three calibration standards. The Spectral Accuracy of 98% or better is also achieved, suggesting a good match between the calibrated spectra and the theoretically calculated spectra of the standards. This high Spectral Accuracy attained from this calibration allows highly reliable and accurate isotope pattern search for ECD.

This calibration is then applied to the entire LC/MS data file of the impurities of simvastatin followed by averaging the calibrated spectra across each of the chromatographic peaks for formula search.

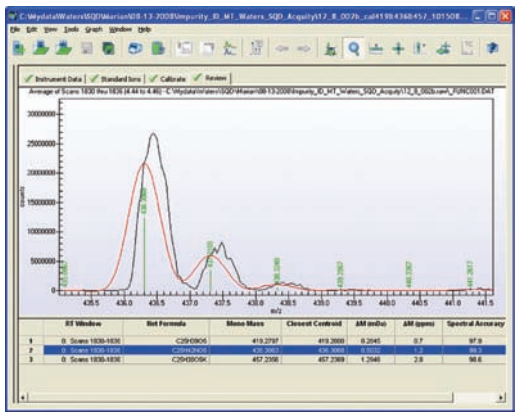

Figure 3. The calibration report summary and overlay of the uncalibrated raw data (Black) and the accurate mass and peak shape calibrated data (Red). Note the substantial improvement in signal-to-noise as well.

Without applying particular constrains, all the formula search parameters in this work are summarised in *Figure 4*. The search employed a comprehensive set of elements including C, H, N, O, K, Cl, and S with their lower limits set to zero and upper limits set very high. Prior results showed the instrument's mass accuracy to be better than 10mDa using the MassWorks calibration and as such a value of 10 mDa for mass tolerance was used for the formula search. Since the mass spectra of most impurities are similar to the parent drug's spectra having three different ion formations such as [M+H] +, [M+NH4]+, and [M+K]+, only the most abundant ions of either [M+NH4]+ or [M+H]+ were selected for formula search.

Table 1. Summary for ECD for Simvastatin Impurities.

Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	Rt	Rank	Total Formula Searched
C25H44NO5	438.3219	4.4485	10.1491	98.4	5	1	34
C25H40NO4	418.2957	5.7338	13.7077	99.1	5.3	1	28
C25H41O6	437.2903	1.814	4.1484	99.4	3	1	49
C24H40NO5	422.2906	8.3484	19.7697	94.5	4.1	4	34
C24H40NO5	422.2906	7.1484	16.928	98.1	4	3	32
C26H44NO5	450.3219	-0.3515	-0.7805	97.3	5.3	1	33
C25H40NO5	434.2906	-1.3516	-3.1122	97.2	4.2	2	36

As summarised in *Table 1*, the results show that most of impurities were identified as the number 1 or 2 hit and have Spectral Accuracy better than 97.2% except for the ions at m/z of 422 eluted at Rt = 4.0 min due to poor signal-to-noise ratio. It is evident that using Spectral Accuracy can effectively distinguish correctly the molecule formula from many possible candidates derived from mass accuracy alone. As a new concept, Spectral Accuracy simply means how closely the isotope distribution of calibrated mass spectra matches that of theoretically calculated ones. It is well known that the isotope pattern of molecules is a fingerprint of the molecules and has been used to help identify unknown compounds for a long time. For example, the 13C/12C ratios are used for determining the number of carbons in small organic molecules and the (M+2)/M peak ratio of 0.33 characterises a compound containing a single element of chlorine (4). With Cerno's comprehensive mass spectral calibration, this manual and primitive isotope pattern matching is dramatically enhanced by accurate peak shape which can be used to quantitatively evaluate search results by Spectral Accuracy.

Spectral accuracy can be conveniently expressed as define below where a perfect match should approach 100.00%,

$$SA = \left(1 - \frac{\|\mathbf{e}\|_2}{\|\mathbf{r}\|_2}\right) \times 100$$

e is the fitting residual vector, r is the calibrated isotope profile vector, and $\|.\|_2$ represents the 2-norm (or square root of the sums of squares of all elements) of a vector. This Spectral Accuracy metric will be used to evaluate all possible formulas whose exact monoisotope masses come within a mass tolerance window of the reported accurate mass obtained off the actual monoisotope peak and whose elemental compositions satisfy the given chemistry constraints. The one formula with the highest Spectral Accuracy is most likely the correct formula for the unknown ion of interest.

The results from this LC/MS analysis also demonstrate that formula identification by the combination of mass accuracy and Spectral Accuracy is more effective than by mass accuracy alone. For example, simvastatin acid, one of the impurities of simvastatin, was identified as the top hit from 49 possible candidates with a mass accuracy of 4 ppm and Spectral Accuracy of 99.4% (*Table 2a*).

Table 2a. Top 10 Formula from Search for m/z 437, ranked by Spectral Accuracy.

Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	RMSE	DBE
1	C25H41O6	437.2903	1.814	4.1484	99.4	35,330	5.5
2	C22H33N10	437.289	0.4661	1.0659	98.8	71,877	11.5
3	C21H37N6O4	437.2876	-0.8713	-1.9925	98.0	115,843	6.5
4	C26H37N4O2	437.2917	3.1515	7.2068	97.8	132,404	10.5
5	C27H37N2O3	437.2804	-8.0819	-18.4819	97.5	146,570	10.5
6	C23H41N4O2S	437.295	6.5223	14.9153	97.1	174,600	5.5
7	C24H41N2O3S	437.2838	-4.7111	-10.7735	96.9	183,226	5.5
8	C20H37N8OS	437.2811	-7.3965	-16.9144	96.7	197,386	6.5
9	C19H37N10S	437.2923	3.8369	8.7744	96.6	200,644	6.5
10	C20H41N2O8	437.2863	-2.2087	-5.0509	96.1	231,489	1.5

Table 2b. Top 10 Formula from Search for m/z 437, Ranked by Mass Accuracy.

Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	RMSE	DBE
1	C22H33N10	437.289	0.4661	1.0659	98.8	83,184	11.5
2	C29H41OS	437.2878	-0.6884	-1.5742	95.7	295,740	9.5
3	C21H37N6O4	437.2876	-0.8713	-1.9925	98.2	124,081	6.5
4	C21H42N4O3K	437.2894	0.8981	2.0539	94.0	415,178	2.5
5	C20H42N4O4Cl	437.2895	0.9586	2.1921	73.0	1,868,436	1.5
6	C21H45N2O3S2	437.2872	-1.3403	-3.065	93.0	487,894	0.5
7	C16H38N10O2Cl	437.2868	-1.7268	-3.9488	72.8	1,881,594	2.5
8	C17H38N10OK	437.2867	-1.7872	-4.087	93.1	476,649	3.5
9	C25H41O6	437.2903	1.814	4.1484	99.5	37,178	5.5
10	C20H41N2O8	437.2863	-2.2087	-5.0509	96.4	252,309	1.5

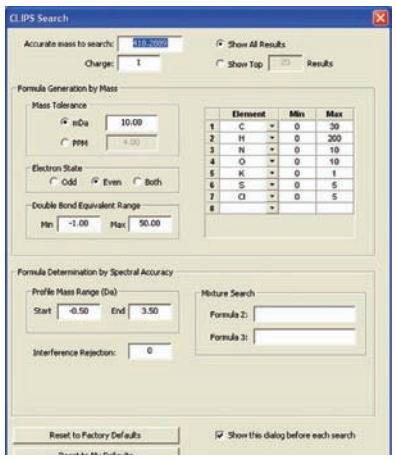

Figure 4. The formula search parameters used for all impurity searches.

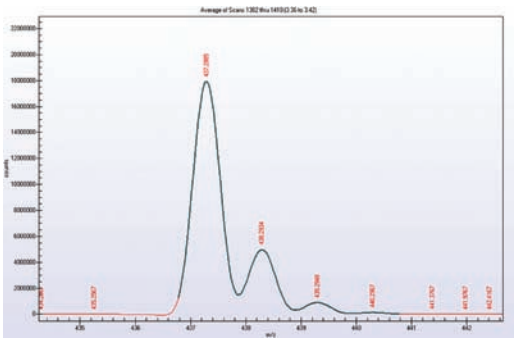

Figure 5. Overlay of calibrated spectra and theoretically calculated spectra of the impurity simvastatin acid.

This high Spectral Accuracy of 99.4% allowed the calibrated spectrum of simvastatin acid to be matched and visually overlaid almost perfectly with the theoretically calculated spectrum, leading to confident identification of the compound by the excellent match (*Figure 5*). On the other hand, with the same search parameters simvastatin acid was ranked only number 9 based on mass accuracy alone (*Table 2b*).

CONCLUSION

Most of the impurities of simvastatin were correctly identified as the top one or two hits with excellent mass accuracy and Spectral Accuracy. These results demonstrate that high mass accuracy and formula identification can be achieved with a single quadrupole instrument coupled with UPLC.

Spectral accuracy is proved to be effective to distinguish correct formula from numerous possibilities. The accurate isotope pattern recognition for formula identification by MassWorks is attributed to the novel calibration technology that provides the same peak shape functions for both calibrated and theoretical spectra. With the mathematically known functions, there is no need to make an assumption for the peak shape to generate theoretical spectra, an error prone step in conventional isotope pattern comparison.

ACKNOWLEDGEMENT

The Author wishes to thank Marian Twohig and Michael Balogh of Waters for running the experiments and providing the data used in this application note

REFERENCES

1. Wang,Y; Prest, H. Chromatography (Japan) 2006,27,135.
2. Gu, M; et al, RCM 2006, 20, 764.
3. Kuehl, D.; Wang, Y. ASMS 2008, Poster PM 013.
4. McLafferty, F; Turecek, F. Interpretation of Mass Spectra, 4th edition, USB, 1993.

Exclusive Worldwide Agreement

Agilent Technologies and **Protein Discovery, Inc** announced their distribution agreement for PPS Silent Surfactant, a specialised reagent from Protein Discovery used to prepare biological samples for mass spectrometry analysis. The exclusive agreement encompasses aspects of sales, marketing and support for PPS Silent Surfactant. PPS Silent Surfactant is an acid-cleavable, zwitterionic detergent that helps researchers produce high-quality mass spectra from biological samples. The patented reagent is particularly useful for solubilising hydrophobic proteins for extraction and preparative digestion.

"PPS Silent Surfactant is a powerful addition to Agilent's sample-preparation product offering and improves the customer experience and capability for LC/MS-based research," said Bill Molnar, Agilent Vice President and General Manager, Genomics and Bioreagents. "It can be used at any stage of the mass spec sample-preparation process without interfering with separation mechanisms, ESI or detection. It relieves detergent interference problems and improves recovery and identification of more proteins. These advantages enable researchers to produce better data and get even greater satisfaction from the other product elements of our comprehensive mass spectrometry portfolio."

"This agreement represents a significant milestone in Protein Discovery's commercialisation of specialty products for preparative proteomics," said Chuck Witkowski, President and Chief Executive Officer of Protein Discovery. "Agilent's status as an innovator in sample preparation, combined with its extensive global distribution network, makes Agilent a great partner for Protein Discovery." Effective immediately, Agilent customers can place orders for PPS Silent Surfactant through Agilent's Life Sciences and Chemical Analysis division.