Feasibility and Reliability of Low and High-Resolution MS Approaches for Accurate Mass and Molecular Formula Determination in Drug Discovery.

Vladimír Čápka[1], Ming Gu[2]

[1]AstraZeneca R&D Boston, Infection Discovery, 35 Gatehouse Dr., Waltham, MA 02451

[2]Cerno Bioscience, 14 Commerce Dr., Danbury, CT 06810

Overview

- Accurate mass provides information regarding compound identity in confirmation of reaction products, drug candidates, and metabolite identification.
- Traditionally, high mass accuracy was required for meaningful molecular formula determination.
- Isotopic MS pattern of an analyte bears valuable information about its elemental composition and can be used for molecular formula determination.
- The current work shows evaluation of isotopic pattern information for molecular formula ID by low and high resolution instruments.

Introduction

- Each unique molecular formula has a unique mass.
- Each molecular formula has a unique isotopic pattern.
- Even with <5ppm mass accuracy, molecular formula determination using TOF data can yield ambiguous data in the absence of chemical knowledge of analytes.
- Isotopic pattern interpretation greatly increases confidence in molecular formula determination for accurate mass data (TOF, <5ppm mass accuracy), and enables molecular formula determination for low resolution data (quadruple-based instruments).
- Isotopic pattern interpretation and spectral accuracy calculation was done using Mass Works software and compared with MassLynx (TOF only).
- Molecular formula ranks determined from a single quadrupole and a TOF instrument were compared for 15 model compounds.
- The aim of this work was to simplify the molecular formula determination using a more cost effective low-resolution instrument.

Results

Mass Works software (Figure 1)

Quadrupole Instrument

- Mass Works generates a peak shape calibration function using a known standard that is applied to the raw MS data for the unknown analyte. Mass accuracy is also calibrated.
- The calibrated spectrum is compared with the theoretical one for a molecular formula candidate. Spectral accuracy is calculated. Formula rank is based on spectral accuracy.

TOF Instrument

- Because mass accuracy is assumed, no standards are required.

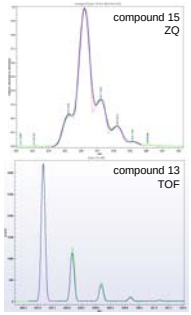

Figure 1

Mass Works processed mass spectra.
Spectral Accuracy (%) = 100*(1-RMSE)
TOP: SA = 98.9 %, rank 1
Bottom: SA = 94.6 %, rank 32

Saturation effects can be seen in TOF spectrum resulting in low spectral accuracy. Thus, lower formula rank.

Red – raw data
Green – calibrated measured data
Blue – theoretical match

Proof of Concept (Quadrupole Instrument)

- Nine standards were used for spectral accuracy calibration.
- Second set of the same standards were used as "unknowns".

Results (Table I)

Mass accuracy: < 8mDa, 5-17 ppm (within range only).

Excellent molecular formula rank even 200 Da out of calibrated range.

Table I. Molecular formula ID of standards via self-calibration.

	Formula	Rank	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy
acetaminophen	C8H10NO2	3	152.0712	98.4536	647.8372	94.8309
terbutaline	C12H20NO3	1	226.1443	15.7185	69.5114	94.8913
imipramine	C19H25N2	2	281.2018	-27.3262	-97.167	95.1169
buspirone	C21H32O2N5	1	386.2556	4.6697	17.346	99.4
Phe-Gly-Gly-Phe	C22H27O5N4	2	427.1981	-2.755	-6.449	99.3491
Loperamide	C29H34O2N2Cl	2	477.2309	-7.7189	-16.1742	99.1139
Tyr-Tyr-Tyr	C27H30N3O7	2	508.2084	-3.4247	-10.674	99.1199
leucine enkephalin	C28H41O7N5	1	609.2812	-3.7641	-6.2271	99.4988
erythromycin	C37H68NO13	2	734.4691	6.6663	9.0764	99.029

Green – calibration standards covering the mass range of unknowns.
Yellow – standards used for testing extrapolation of formula ID beyond the calibrated range.

Table II. Molecular formula ID of model unknowns compounds.

	Formula	Accurate Mass [M+H]+	Mass Works ZQ	Mass Works TOF	MassLynx TOF	mass accuracy MassLynx, ppm	Mass Works TOF	MassLynx TOF	MassLynx i-FIT	mass accuracy MassLynx, ppm
Compound 1	C18H18N4O3S	371.1178	2	2	n/f	25.3	1	5	1	1.3
Compound 2	C19H19FN4O6	*419.1367*	3	10	n/f	20.5	1	15	1	4.8
Compound 3	C19H19FN4O4S	*419.1189*	2	6	n/f	19.6	1	16	1	3.3
Compound 4	C21H19Cl2N3O3S	548.0787	5	10	n/f	14.4	1	21	1	4.9
Compound 5	C23H19F3N6O5S	*549.1168*	9	43	n/f	15.5	4	10	4	0.5
Compound 6	C25H21ClN8O3S	*549.1224*	4	15	n/f	19.3	1	9	1	1.8
Compound 7	C24H29ClN6O5S	*549.1687*	4	5	n/f	14.4	1	6	1	0.9
Compound 8	C24H22ClN9O3S	552.1333	3	5	n/f	16.8	1	13	1	1.3
Compound 9	C23H29ClN6O6S	553.1636	1	20	n/f	18.6	1	10	1	2.0
Compound 10	C23H27Cl2N7O5S	*584.1250*	6	16	n/f	13.9	1	20	1	2.2
Compound 11	C24H23Cl2N1IOS	*584.1263*	1	6	n/f	14.2	2	8	3	0.7
Compound 12	C26H26F3N6O7PS	655.1352	1	70	n/f	13.1	7	12	9	-1.1
Compound 13	C29H32F3N6O7PS	*697.1821*	15	215	n/f	14.8	32	29	31	2.9
Compound 14	C33H35F3N8O4S	*697.2532*	2	59	n/f	16.6	12	19	17	1.7
Compound 15	C26H28BF2N7O4S2	*615.1820*	1	30	n/f	4.2	5	45	8	4.2

Red – isobaric compounds
Blue – unique isotopic pattern (monoisotopic mass with low abundance)

No lock mass, n/f – not feasible

With lock mass

Methods

	Single Quadrupole, ZQ	TOF, LCT Premier
Instrument	Waters ZQ + HP1100	Waters LCT premier + Waters Acquity UPLC
Column	Varian Polaris C18A, 3μm, 50 mm x 2.0 mm	
Flow Rate	1 mL/min	0.5 mL/min
Injection Volume	2 μL	1-4 μL (to achieve 100-300 cps signal)
Mobile Phase	A: 0.1% formic acid in H_hO B: 0.1% formic acid in ACN	
Gradient	5 to 95% B in 4.5 min (unknowns) 5 to 50% B in 4.5 min (standards)	5 to 95% B in 4.5 min
MS detection, ESI+	100 – 800 amu in 1 s, inter-scan delay 0.03 s	200 – 800 amu in 0.5 s, inter scan delay 0.05 s
Resolution	Unit, ~0.6 amu peak width half hight	~6000 PWHM
Mass accuracy (based on MassLLynx)	N/A, nominal mass, ± 0.2 amu	<5 ppm, lock spray (5 μL/min) 4-25 ppm no lock spray
Lock Mass	N/A	buspirone ([M+H] 386.2556), 1 ng/mL leucine enkephalin ([M+H] 556.2771), 10 ng/mL Reserpine ([M+H] 609.2812), 5 ng/mL
Spectra accuracy calibration	Required	Not required

Single Quadrupole

- Due to low mass accuracy, spectral accuracy calibration is required. 9 commercially available calibration standards used.

TOF

- No spectral accuracy standards required.

Mass Works Data Processing Criteria

C: 1-50	F: 1-5
H: 1-100	Cl: 1-2*
N: 1-25	B: 1-2*
O: 1-25	P: 1-2
S: 1-11	

- Only elements present in each molecule were used in the molecular formula search.

* Easily recognizable pattern, more restrictive criteria used.

Conclusions

- Isotopic pattern interpretation using MassWorks software enables molecular formula ID using a low resolution quadrupole-based instrumentation. < 20mDa mass accuracy was achieved.

  -Quadrupole instruments – cost effective, wide dynamic range, suitable for inexperienced user. Due to dynamic range, able to interpret unique isotopic patterns (see compound 15).
- Molecular formula ID using TOF instrument requires high mass accuracy. Mass accuracy alone is insufficient for unambiguous molecular formula ID.
- Isobaric compounds were unambiguously differentiated by quadrupole and TOF instruments, i.e. formula list did not contain the other isobaric compounds or they were present with a poor rank.
- Quadrupole data calibrated by Mass Works software is superior to TOF data without internal calibration and to the accurate mass data (<5ppm) without isotopic pattern interpretation.
- Ultimate molecular formula ID requires high mass accuracy and the isotopic interpretation.
- TOF molecular formula ID is problematic for compounds with monoisotopic elements that lack isotopic signature (F, P, I) resulting in lower rank.

Acknowledgments

AstraZeneca Infection and Cancer Chemistry

Cerno Bioscience