Accelerating the Metabolite Identification Process Using High Resolution Q-TOF Data and Mass-MetaSite Software

Application Note

Drug discovery and development: Metabolite Identification

Authors

Yuqin Dai, Stephen Madden, and Lester Taylor
Agilent Technologies, Inc.
Santa Clara, CA, USA

Gabriele Cruciani and Ismael Zamora
Molecular Discovery Ltd,
London, UK

Abstract

This application note demonstrates use of the Agilent 6550 iFunnel Q-TOF together with Mass-MetaSite software, to rapidly identify drug related metabolites and to efficiently assign chemical structures to each of the metabolites found. This process is illustrated with reference to the *in vitro* metabolism of the anxiolytic drug buspirone.

Agilent Technologies

Introduction

LC/MS plays a key role in drug metabolite identification in drug discovery and development. Modern Q-TOF instruments are capable of providing accurate mass information for parent species together with diagnostic fragment ion information. This enables collection of data that can distinguish drug metabolites from most endogenous interferences, as well as can determine elemental compositions of metabolite precursors and their fragment ions. The current bottlenecks in metabolite identification include assay analysis throughput and data interpretation, together with the need for sophisticated data processing software tools to help transform the huge volume of data into useful information that can be shared in an efficient manner.

This application note demonstrates use of the Agilent 6550 iFunnel Q-TOF together with Mass-MetaSite software, to rapidly identify drug related metabolites and to efficiently assign chemical structures to each of the metabolites found. This process is illustrated with reference to the *in vitro* metabolism of the anxiolytic drug buspirone, Figure 1.

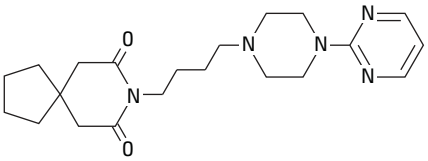

Figure 1. Structure of buspirone ($C_{21}H_{31}N_5O_2$).

Experimental

Methodology

The Mass-MetaSite methodology consists of two steps. The first step identifies drug related material using the data obtained from the Q-TOF MS and MS/MS analysis in order to find chromatographic peaks that are related to the parent compound. This identifies two types of peaks:

- Those that correspond to a mass shift from the substrate corresponding to one or more metabolic reactions (phase I and II) that are within the software
- Unknown metabolites, these are observed peaks but do not correspond to any of the list of known bio-transformations.

In the second step, the chemical structures giving rise to the peaks are assigned. To perform this operation, the software produces a set of theoretical fragments (structure and mass) for the parent and the metabolites. The masses for this fragment list are compared with the actual fragment ions found in the MS and MS/MS spectra. From the comparison of the predicted metabolites and high resolution accurate mass data the software is able to localize the region in the molecule where the metabolic reaction has occurred. In cases where the fragment analysis cannot specify a single atom in the parent that gives rise to the metabolic product, a Markush representation of the metabolites compatible with the mass spectra data is presented to the user and in order to prioritize among the different potential structural solutions, the MetaSite Site of Metabolism (SoM) algorithm[1,2] is applied. The SoM prediction is used to identify the regioselectivity of the metabolic reaction only for phase I metabolism, and it is based on an analysis of the potential interactions of the compounds in the cytochrome P450 cavity and their chemical reactivity towards oxidation.

In vitro metabolism in rat liver microsomes (RLM)

Substrate	1 µM buspirone
Incubation	In RLM at 37 °C for 0, 5, 10, 20, and 30 minutes
Reaction	Initiated by adding NADPH-regenerating system and quenched by adding 1:2 (v/v) iced-cold acetonitrile

Instrumentation

LC/MS system

An Agilent 1290 Infinity UHPLC system comprised of a binary pump with an integrated degasser, a high performance autosampler with a thermostat and a thermostatted column compartment, and an Agilent 6550 iFunnel Q-TOF with a dual Agilent Jet Stream source.

Results and Discussion

Computing the probability of site of metabolism

Three driving forces control the site of metabolism for substrates of human CYP enzymes. The site of metabolism can be described by a probability function P_{SM}:

$$P_{SM}i = Ei * Ri * Mi \qquad (\sim\text{free energy})$$

Where:

$P_{SM}i$ = The probability of an atom i being the site of metabolism catalyzed by the CYP-Heme group

Ei = The accessibility of atom i to the CYP-Heme

Ri = The reactivity of atom i in the actual mechanism of reaction

Mi = The relative probability of a reaction mechanism under consideration occurring

Buspirone is well absorbed, but is subject to first-pass metabolism. The mean systemic availability is approximately 4 %. Buspirone is eliminated primarily by oxidative metabolism, which produces several hydroxylated metabolites, including 5-hydroxy-buspirone and 1-pyrimidinylpiperazine.

Figure 2 shows buspirone (substrate) and several metabolites determined by the peak finding algorithm in Mass-MetaSite. Table 1 shows a summary of metabolites with retention time (RT), observed *m/z*, *m/z* shift with respect to the substrate, the abundance, and the best matching ion formula. Using Mass-MetaSite software, a total of 24 Phase I metabolites were quickly identified from the 10 minute RLM incubation sample. As shown in Figure 2 and Table 1, the major metabolites identified from this incubation sample are mono-hydroxylation metabolites (M10, M12 and M23) and an N-dealkylated metabolite (M1).

Chromatographic conditions

LC conditions		
Column	Agilent ZORBAX Eclipse Plus C18, 2.1 × 100 mm, 1.8 µm (p/n 959758-902)	
Mobile phase	A = 0.1 % formic acid in water, B = 0.1 % formic acid in 95 % acetonitrile/water	
Injection volume	3 µL	
Column temperature	50 °C	
Flow rate	0.4 mL/min	
Gradient	Time (min)	%B
	0.5	8.0
	2.5	50
	2.8	98
	3.2	98
	3.4	8.0
Post time	1 minute	
MS conditions		
Ion mode	positive	
Drying gas	250 °C at 12 L/min	
Sheath gas	400 °C at 12 L/min	
Capillary voltage	3,500 V	
Nozzle voltage	300 V	
Instrument mode	Extended Dynamic Range (2 GHz)	
Scan mode	full MS (10 Hz) followed by three data dependent auto MS/MS (10 Hz) scans	
Mass range	100–1,000 *m/z*	
Collision energy	32 V	
Reference ions	118.086255 and 922.009798	

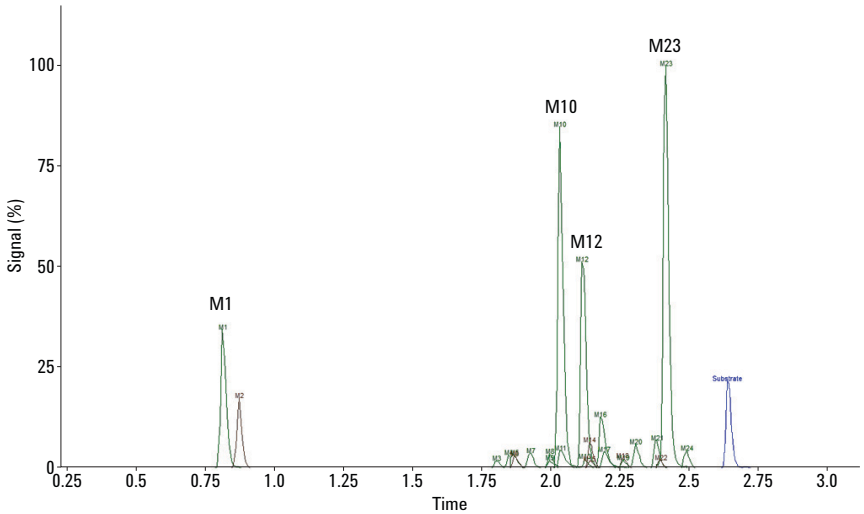

Figure 2. Peak identification from Mass-MetaSite.

Site of metabolism (SoM) prediction

Mass-MetaSite offers SoM prediction based on the molecular structure of the substrate without any MS data. Figure 3 shows the SoM prediction results for buspirone. Atoms in the graphic depiction (right panel) and bars in the graph (left panel) are in 1-to-1 relationship and their selection is synchronized. The information reported in both panels corresponds to the site of metabolism prediction.

Table 1. Summary of metabolites identified from a 10 minute RLM incubation sample using Mass-MetaSite.

Selection	RT	m/z shift	m/z obs.	Area %	UV Area %	Area Abs	UV Area Abs	ion formula	m/z calc.
Substrate	2.64	-0	386.2553	5.65		4e+05		[C21H31N5O2 + H]+	386.2551
M1	0.81	-221	165.1132	7.52		5.3e+05		[C8H12N4 + H]+	165.1135
M2	0.87	-214	172.0960	3.19		2.2e+05		[C8H13NO3 + H]+	172.0968
M3	1.80	+32	418.2438	0.62		4.4e+04		[C21H31N5O4 + H]+	418.2449
M4	1.85	+32	418.2435	0.77		5.4e+04		[C21H31N5O4 + H]+	418.2449
M5	1.87	+32	418.2442	1.10		7.7e+04		[C21H31N5O4 + H]+	418.2449
M6	1.87	-10	376.2330	0.65		4.6e+04		[C19H29N5O3 + H]+	376.2343
M7	1.93	+32	418.2439	0.78		5.5e+04		[C21H31N5O4 + H]+	418.2449
M8	2.00	+32	418.2443	1.03		7.2e+04		[C21H31N5O4 + H]+	418.2449
M9	2.00	+14	400.2329	0.52		3.6e+04		[C21H29N5O3 + H]+	400.2343
M10	2.03	+16	402.2494	17.19		1.2e+06		[C21H31N5O3 + H]+	402.2500
M11	2.03	+14	400.2329	1.36		9.5e+04		[C21H29N5O3 + H]+	400.2343
M12	2.11	+16	402.2491	16.73		1.2e+06		[C21H31N5O3 + H]+	402.2500
M13	2.12	+32	418.2440	0.62		4.4e+04		[C21H31N5O4 + H]+	418.2449
M14	2.14	-10	376.2336	1.46		1e+05		[C19H29N5O3 + H]+	376.2343
M15	2.14	+12	398.2145	0.43		3e+04		[C21H28N5O3 + H]+	398.2187
M16	2.18	+32	418.2447	2.97		2.1e+05		[C21H31N5O4 + H]+	418.2449
M17	2.19	+14	400.2332	1.09		7.7e+04		[C21H29N5O3 + H]+	400.2343
M18	2.26	-10	376.2331	0.59		4.1e+04		[C19H29N5O3 + H]+	376.2343
M19	2.26	+32	418.2435	0.50		3.5e+04		[C21H31N5O4 + H]+	418.2449
M20	2.31	+32	418.2441	1.89		1.3e+05		[C21H31N5O4 + H]+	418.2449
M21	2.38	+32	418.2442	2.11		1.5e+05		[C21H31N5O4 + H]+	418.2449
M22	2.40	-10	376.2332	0.57		4e+04		[C19H29N5O3 + H]+	376.2343
M23	2.42	+16	402.2492	29.51		2.1e+06		[C21H31N5O3 + H]+	402.2500
M24	2.49	+32	418.2435	1.10		7.7e+04		[C21H31N5O4 + H]+	418.2449

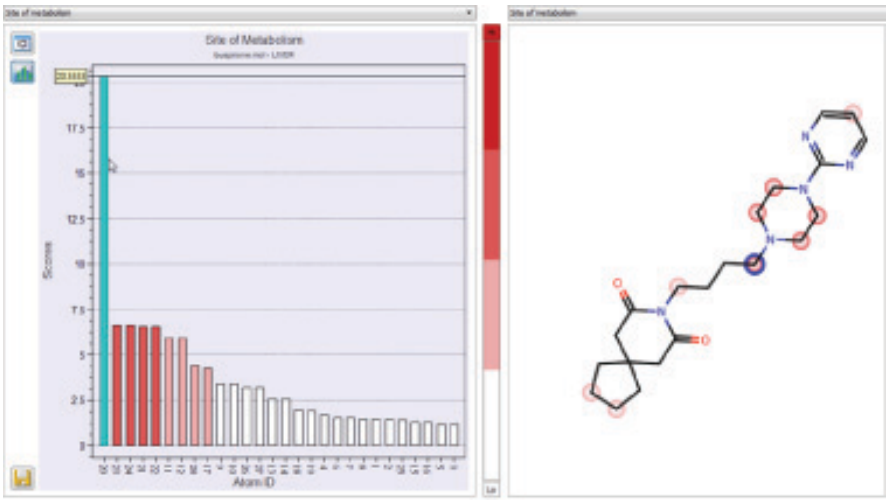

Figure 3. Site of Metabolism (SoM) prediction results.

Metabolite Identification and Structure Assignment using Q-TOF MS and MS/MS Data

With Q-TOF mass data, Mass-MetaSite performs metabolite identification and structure assignment in two steps:

1. Identify drug related metabolites using the chromatograms obtained from HRAM Q-TOF MS and MS/MS method (Figure 2).

2. Assign Markush structures to metabolites based on the predicted metabolites and HRAM Q-TOF MS and MS/MS data (Figures 4 and 5).

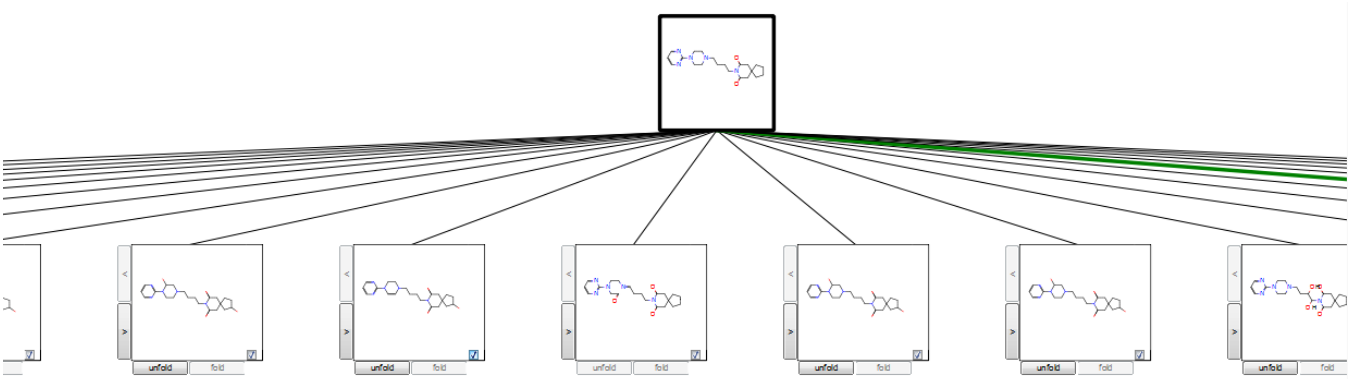

Figure 4. Partial visualization of metabolite structures based on parent drug using Mass-MetaSite. This window allows all metabolites to be viewed interactively. Each of these structures can then be selected for further information.

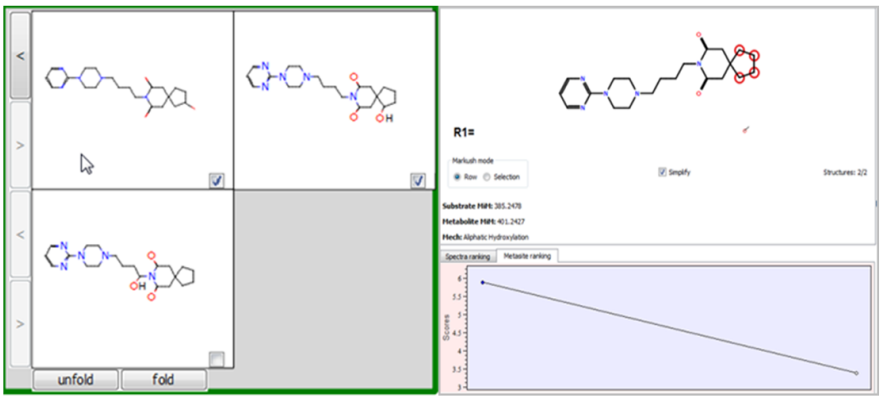

Figure 5. Proposed metabolite structures for M10 with spectra ranking scores and matches of fragment ions.

Figure 6 shows MS/MS spectra of the parent (substrate) and the metabolite M10; the structures for the fragment ions are proposed on the right panel. The software can localize the metabolic site in the molecule through automatic fragment analysis and MS/MS spectra correlation. Molecular ion shift can also be seen using MS spectra on the MS tab.

Conclusions

This application note demonstrates the integrated use of high resolution accurate mass Q-TOF data together with Mass-MetaSite metabolite identification software to readily identify and assign the metabolites of buspirone from *in vitro* RLM incubations. Mass-MetaSite gives predicted metabolites based on probability of catalytic reactivity at the CYP Heme group. Combined with high resolution accurate mass data from Q-TOF analysis, this gives a high degree of confidence to the assignment of metabolite structures. Putative metabolite candidate structures may be proposed for each peak found in the LC chromatogram and these are ranked based on the probability of reaction at the site of metabolism. This integrated solution minimizes bottlenecks in data analysis and interpretation, thus increasing the efficiency of metabolite identification.

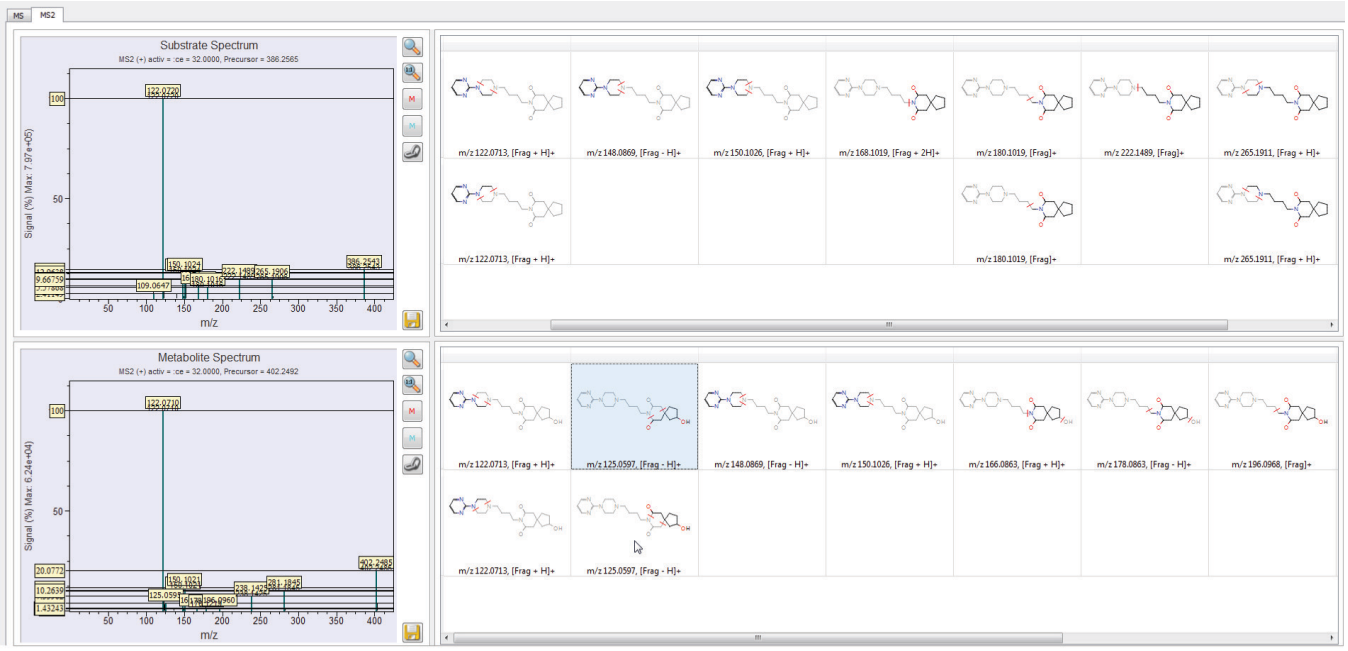

Figure 6. Automatic fragment analysis displays the fragment ions that are used for M10 structure elucidation.

References

1. Cruciani G., *et al*. MetaSite: "Understanding Metabolism in Human Cytochromes from the Perspective of the Chemist", *J. Med. Chem*., **2005,** 48, 6970-6979.

2. B. Bonn, C. *et al*. "Enhanced metabolite identification with MS(E) and a semi-automated software for structural elucidation", *Rapid Commun. Mass Spectrom*. **2010** Nov 15; 24 (21):3127-38.

3. R.E. Gammans, R.F. Mayol, and J.A. LaBudde, *Am. J. Med*., **1986** Mar 31;80 (3B):41-51.

www.agilent.com/chem/QQQ

Published in the USA, July 17, 2013
5991-2710EN

Agilent Technologies