High-Throughput and Automated Software Workflow Strategies for Small Molecule Identification and Characterization Using a High Resolution Q-TOF Mass Spectrometer

Alina Dindyal-Popescu; Iris Shek; Mohammad Jooyandeh; Ian Moore; Eva Duchoslav
SCIEX, 71 Four Valley Drive, Concord, ON, L4K 4V8 Canada

ABSTRACT

Transformation of accurate mass product spectra into putative structures of metabolites poses a bottleneck in early discovery studies to identify metabolic soft spots, as well as in the subsequent characterization of active metabolites. Here we present an automated software workflow for the proposal of structure analogues and its application to a variety of small molecules. Attributes affecting performance of this qualitative workflow are discussed.

INTRODUCTION

While software-predicted results for metabolite structure proposals cannot beat human expertise, software strategies can greatly aid in streamlining the process of metabolite soft-spot identification.

In this work, a high-throughput and completely automated workflow strategy is investigated for small molecule metabolite identification and characterization using a high-resolution Q-TOF mass spectrometer.

Both parent-drug tailored IDA and generic SWATH® acquisition methods have been used for data collection. The resulting files were interrogated to find drug-related material. Since the MS/MS data were collected in parallel with accurate TOF MS data, metabolite fragments in conjunction with the known parent drug structure provided grounds for ranking of putative structures and proposal of biotransformation reactions.

MATERIALS AND METHODS

10 µM microsomal incubations of various compounds (nefazodone, propranolol, bromocriptine, diclofenac, verapamil, imipramine) were analyzed using a TripleTOF® 5600 system coupled with a Shimadzu Prominence UFLC System. A 300 µL/min gradient was utilized with a Kinetex C18 (2.6 u) 2 *50 mm column and run for 10 minutes.

Table 1. Core MS data collection parameters for variety of acquisition strategies

Both IDA and SWATH® acquisition methods in either positive or negative modes were used.

	IDA positive	SWATH positive	IDA negative	SWATH negative
TOF MS mass range	100 - 1000	100 - 1000	100 - 1000	100 - 1000
TOF MS/MS mass range	100 - 1000	100 - 1000	100 - 1000	100 - 1000
TOF MS/MS CE	35 ± 15	25 to 55	35 ± 15	-4 to -40
SWATH Q1 coverage	N/A	300 - 800 (25*20)	N/A	100 - 915 (variable)
Number of MS/MS experiments	3	20	3	20
Cycle time (ms)	390	900	390	600

Resulting LC/MS data was mined with MetabolitePilot™ 2.0 software using both hypothesis-driven and generic strategies to find drug-related material. Sites of modifications for major metabolites were proposed, ranked and scored within the batch processing. Automatically generated proposed metabolite structures were then validated either manually or against published data.

DISCUSSION

Within MetabolitePilot 2.0 software structures are proposed for the following metabolite types:

- One or more dealkylation cleavage,
- One biotransformation,
- Combination of cleavage and biotransformation.

Scoring and ranking of potential modification sites is based on projection of information from the annotated MS/MS spectrum of parent drug onto the metabolite MS/MS spectrum. Subset of SOM candidate atoms is retrieved from the biotransformation database. Score for each SOM candidate atom is derived from the evidence supporting its both unchanged and modified states. Proposed structures are generated by attaching / removing biotransformation structure motif onto parent drug.

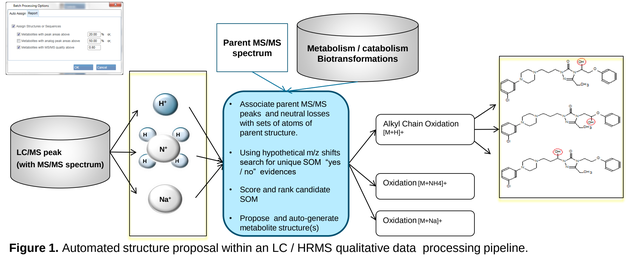
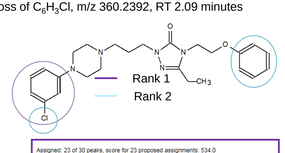


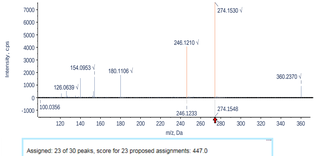
Figure 1. Automated structure proposal within an LC / HRMS qualitative data processing pipeline.

Accurate mass MS and MS/MS data were collected in parallel; complementary ions were grouped in TOF MS and MS/MS spectra of protonated / deprotonated precursor were used for structure proposal.

RESULTS

Example 1: Nefazodone dealkylation metabolite, Loss of C_6H_3Cl, m/z 360.2392, RT 2.09 minutes





Assigned: 23 of 30 peaks, score for 23 proposed assignments: 534.0

Fragments: 17 of 43 Proposed Formulae

	Use	Mass (m/z)	Ion Formula	Error (mDa)	Intensity (cps)	RDB	Proposed Structures	Score
1	✓	110.0943	C7H12N	-2.2	48.3	3.0	1	32.0
2	✓	121.0622	C8H9O	-2.6	335.7	5.0	1	40.0
3	✓	136.1005	C7H12N3	-2.1	80.6	4.0	1	33.0
4	✓	180.1106	C9H14N3O	-2.5	1808.3	5.0	1	36.0
5	✓	274.1530	C15H20N3O2	-2.0	7563.3	6.0	1	49.0
6	✓	360.2370	C19H30N3O2	-2.4	929.1	6.0	1	44.0
7	✓	149.0201	C8H5N3O2	-1.9	79.1	7.5	2	5.0
8	✓	246.1210	C13H16N3O2	-2.7	4066.6	6.0	3	22.0
9	✓	276.1601	C14H22N4O2	2.0	112.2	7.5	3	8.5
10	✓	152.0600	C7H10N3O	-1.9	156.9	5.0	6	36.0
11	✓	153.0661	C7H11N3O	-1.5	511.9	4.5	6	33.5
12	✓	154.0496	C5H8N3O	-1.0	48.3	5.0	8	21.0
13	✓	126.0639	C5H8N3O	-2.3	535.3	4.0	8	21.5
14	✓	138.0633	C6H8N3O	-2.9	96.6	5.0	8	36.0
15	✓	140.0799	C6H10N3O	-1.9	1552.7	4.0	8	36.5
16	✓	211.1535	C10H19N4O	-1.8	48.2	4.0	9	32.5

Two "Loss of C_6H_3Cl" structures were considered.

Assigned: 23 of 30 peaks, score for 23 proposed assignments: 447.0

Fragments: 13 of 43 Proposed Formulae

	Use	Mass (m/z)	Ion Formula	Error (mDa)	Intensity (cps)	RDB	Proposed Structures	Score
1	✓	135.0790	C7H9N3	-2.1	160.4	5.5	1	31.0
2	✓	136.1005	C7H12N3	-2.1	80.6	4.0	1	32.0
3	✓	180.1106	C9H14N3O	-2.5	1808.3	5.0	1	36.0
4	✓	360.2370	C19H30N5O2	-2.4	929.1	6.0	1	44.0
5	✓	110.0943	C7H12N	-2.2	48.3	3.0	3	32.0
6	✓	152.0600	C7H10N3O	-1.9	156.9	5.0	7	36.0
7	✓	153.0661	C7H11N3O	-1.5	511.9	4.5	7	33.5
8	✓	154.0953	C7H12N3O	-2.2	1652.5	4.0	7	36.5
9	✓	211.1535	C10H19N4O	-1.8	48.2	4.0	7	8.5
10	✓	124.0496	C5H6N3O	-1.0	48.3	5.0	9	29.0
11	✓	126.0639	C5H8N3O	-2.3	535.3	4.0	9	32.0
12	✓	138.0633	C6H8N3O	-2.9	96.6	5.0	9	36.0
13	✓	140.0799	C6H10N3O	-1.9	1552.7	4.0	9	36.5

Proposed structure 1 is supported by 6 unique fragment substructures including one for fragment 274.1530. No unique fragments support attachment of phenyl group to piperazine ring.

MS/MS Collection	Compound Name	Metabolite Name	RT (min)	m/z	Polarity	Candidate Structures	Structure Count* (Generated/ Expected)	Rank	Normalized Score (%)	NOTES
IDA	Simvastatin	Oxidation	4.48	435.2747	+	17	2 / 2[C]	1	100	
		Demethylation	4.79	405.2639	+	0	0 / 0	1	100	
IDA	Methocarbamol	Oxidation	6.08	258.0967	+	11	1 / 1[D]	1		
		Oxidation	7.27	258.0972	+	11	1 / 1[D]	1	100	
		Loss of CHNO + Demethylation	8.03	185.0802	+	1	manual interpretation	n/a		
IDA	Verapamil	Loss of C10H12O2 and CH2 and Oxidation	2.49	293.1858	+	36	manual interpretation	n/a		
		Loss of CH2 and CH2	3.03	427.2592	+	10	1 / 1[B]	1	100	
IDA	Diclofenac	Oxidation	1.4	310.0039	-	13	4 / 4[A]	1	100	
IDA	Benzbromarone	Oxidation	9.21	436.9021	-	10	2 / 2[B]	1	100	
IDA	Bromocriptine	Oxidation	5.7	668.2154	-	22	1 / 1[B]	1	100	
SWATH	Propranolol	Oxidation	1.87	276.1595	+	6	3 / 3[D]	1 to 2	100/94.3	
		Oxidation	2.03	276.16	+	6	3 / 3[D]	1 to 2	100/97.3	
		Glucuronidation	2.52	436.1979	+	2	1 / 1[D]	1	100	
		Glucuronidation	2.69	436.1991	+	2	1 / 1[D]	1	100	
		Loss of C3H7N+Demethylation to Carboxylic Acid	3.24	233.0822	+	1	1 / 1[D]	1	100	
SWATH	Nefazodone	Loss of C10H11ClN2+Ketone Formation	2.66	290.1503	+	7	0 / 1[B]			*
		Ethyl to Alcohol	3.43	458.1951	+	1	0 / 1[B]			**
		Loss of C6H3Cl	2.09	360.2392	+	2	1 / 1[B]	1	100	
		Loss of C6H4	2.44	394.2006	+	1	1 / 1[B]	1	100	
		Loss of C15H19N3O2+Oxidation	2.73	213.0784	+	11	1 / 1[B]	1	100	
		Loss of C15H19N3O2+Oxidation	0.57	213.0791	+	11	1 / 1[B]	1	100	
		Loss of C6H3Cl+Oxidation	1.69	376.2343	+	38	12 / 2[B]	6	99.3/98.3	
		Loss of C10H11ClN2+Oxidation	3.24	292.1657	+	15	1 / 1[B]	1	100	
		Loss of C6H3Cl+Ketone Formation	2.19	374.2179	+	18	2 / 1[B]	2	99.7	
		Loss of C6H4+Oxidation	2.17	410.1948	+	20	No reference found	n/a		
		Loss of C10H11ClN2+Demethylation to Carboxylic Acid	3.35	306.1445	+	2	1 / 1[B]	1	100	
SWATH	Diclofenac	Oxidation	1.32	310.0044	-	13	4 / 4[A]	1	100	
SWATH	Midazolam	Demethylation to Carboxylic Acid	3.67	356.0591	+	1	No reference found	n/a		
		Oxidation	2.55	342.081	+	14	2 / 2[B]	1	100	
SWATH	Bromocriptine	Oxidation	3.9	670.2254	+	22	2 / 2[B]	1	100	

* Loss of C10H11ClN2, +O, -2H (complex 3-step biotransformation, not in automated workflow scope)
** oxidation and re-arrangement (currently not represented in biotransformation database)

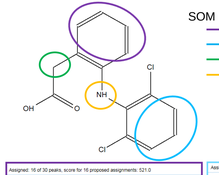
Table 2. Validation of automated structure proposal workflow.

Within the workflow validation, we found that low m/z fragments that originated from a limited specific region of the molecule provided confident insight into structure features. Also, informative MS/MS spectrum of parent drug that covered all portions of parent molecule typically enhanced contrast in candidate structure scores and confidence in putative structure assignment.

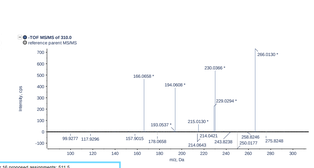
The new structure proposal and ranking workflow is complementary to existing "Interpret" functionality in MetabolitePilot™ software. The processing workflow as indicated in Figure 1 is expandable to additional cleavage and biotransformation combinations.

Compatible parameters are critical for MS/MS fragmentation annotation, in examples 1 and 2 up to 30 fragments with S/N above 3 were used, mass tolerance was 5mDa, and in-silico fragments were generated by breaking up to 4 bonds (including aromatic ring) of the parent molecule.

Example 2: Hydroxy metabolite of Diclophenac, m/z = 310.0039, RT = 1.4 minutes (negative polarity mode)



SOM regions:
Rank 1
Rank 2
N/A
N/A



Assigned: 16 of 30 peaks, score for 16 proposed assignments: 521.0

Fragments: 14 of 23 Proposed Formulae

	Use	Mass (m/z)	Ion Formula	Error (mDa)	Intensity (cps)	RDB	Proposed Structures	Score
1	✓	152.0012	C6H6O3	3.3	7.0	4.5	1	48.5
2	✓	164.0015	C7H6N	1.2	4.0	10.0	1	30.5
3	✓	166.0658	C7H8N	-0.5	465.5	9.0	1	35.5
4	✓	176.0505	C13H6N	-0.1	7.0	11.0	1	35.5
5	✓	179.0363	C12H5NO	-1.4	6.0	10.5	1	31.0
6	✓	193.0537	C13H7NO	0.4	15.4	10.5	1	36.0
7	✓	194.0600	C13H8NO	-0.4	388.2	10.0	1	39.5
8	✓	164.0796	C12H10N2	2.8	3.9	8.0	3	34.5
9	✓	140.0494	C7H6O2	1.5	5.5	5.5	4	33.0
10	✓	226.0209	C13H7ONO	-1.2	6.6	10.0	6	36.0
11	✓	229.0294	C13H5ClNO	-0.6	224.5	9.5	6	40.0
12	✓	230.0366	C13H6ClNO	-1.3	540.0	9.0	6	39.5
13	✓	266.0130	C13H10O2NO	-1.5	732.3	8.0	8	42.5
14	✓	215.0130	C12H8ONO	-1.3	66.3	9.5	22	36.0

Assigned: 16 of 30 peaks, score for 16 proposed assignments: 511.5

Fragments: 14 of 23 Proposed Formulae

	Use	Mass (m/z)	Ion Formula	Error (mDa)	Intensity (cps)	RDB	Proposed Structures	Score
1	✓	160.9939	C6H3OCO	2.2	8.5	4.0	1	51.5
2	✓	164.0015	C7H6N	1.2	4.0	10.0	1	30.5
3	✓	166.0658	C7H8N	-0.5	465.5	9.0	1	35.5
4	✓	176.0505	C13H6N	-0.1	7.0	11.0	1	35.5
5	✓	179.0363	C12H5NO	-1.4	6.0	10.5	1	31.0
6	✓	193.0537	C13H7NO	0.4	15.4	10.5	1	36.0
7	✓	194.0600	C13H8NO	-0.4	388.2	10.0	1	39.5
8	✓	140.0494	C7H6O2	-1.2	5.5	5.5	2	20.5
9	✓	164.0796	C12H10N2	2.8	3.9	8.0	3	34.5
10	✓	226.0209	C13H7ONO	-1.2	6.6	10.0	6	36.0
11	✓	229.0294	C13H5ClNO	-0.6	224.5	9.5	6	40.0
12	✓	230.0366	C13H6ClNO	-1.3	540.0	9.0	6	39.5
13	✓	266.0130	C13H10O2NO	-1.5	732.3	8.0	8	42.5
14	✓	215.0130	C12H8ONO	-1.3	66.3	9.5	22	36.0

For parent drug MS/MS annotation, 4 fragments and 3 neutral losses were mapped onto diclofenac structure.

2 SOM regions were filtered out due since they could not be rationalized due to loss of water in metabolite MS/MS spectrum.

CONCLUSIONS

A fully automated LC HRMS workflow involving either IDA or SWATH® collection and a comprehensive LC/MS data processing including metabolite structure proposal, was demonstrated and validated using variety of model compounds.

This workflow integrated within MetabolitePilot™ 2.0 software, streamlines data processing and enables routine soft-spot analysis capturing major metabolites.

REFERENCES

1 Prueksaritanont, T *et al*, Drug Metab Dispos, 1997, 25, 1191-1199.
2 Thompson R. M. *et al*, Xenobiotica, 1975, 5, 145-153.
3 Tracy T. S. *et al.*, Br. J. Clin. Pharmacol, 1999, 47, 545-552.
4 Tang *et al.*, Curr Drug Metab. 2003, 4, 319-29
5 Kitagawara Y. *et al.*, Drug Metab Dispos, 2015, 43, 1303-1306.
6 Valente D. *et al.*, JPET, 1997, 282, 1408-1424.
7 Baughman T. M. *et al.*, Rapid Commun. Mass Spectrom. 2009, 23, 2146-2150.
8 Li A. C. *et al.*, Rapid Commun. Mass Spectrom. 2007, 21, 1421-1430.
9 Gallagher R. *et al.*, Rapid Commun. Mass Spectrom. 2014, 28, 1293-1302.

TRADEMARKS/LICENSING

Document number: RUO-MKT-10-5828-A