A 3D chemical peak finder for comprehensive small molecule identification

To receive a copy of this poster:
- Scan the code with your phone camera
- Complete the form

SCIEX The Power of Precision

Pradeep Narayanaswamy[1] and Eva Duchoslav[2]
[1]SCIEX, Singapore; [2]SCIEX, Canada

INTRODUCTION

Mass spectrometry (MS) and MS coupled with separation techniques are increasingly popular methods for non-targeted screening. Recent advances in high-resolution MS instrumentation with novel separation, ionization and comprehensive MS/MS data collection strategies require efficient tools to decrease false discovery rates and accurately identify analytes in samples. While typically thousands of ion types/features from complex samples are detected, only a fraction of them is annotated, as each analyte can give multiple features, such as in-source fragments (IF), adducts and different charge states. Accurate spectral interpretation is vital in qualitative workflows, since true molecular ions (M+H/M-H) and accurate mass determination of the analyte are prerequisites for the MS/MS-related structure proposal and confirmation.

MATERIALS AND METHODS

Metabolite standards with known concentrations were purchased from the Mass Spectrometry Metabolite Library (Sigma Aldrich, supplied by IROA Technologies), NIST SRM 1950 metabolites in human plasma (Sigma Aldrich) and mouse liver sample data from our previous publication on MetaboKit.

Standards:
Metabolite standard solutions were prepared according to the supplier protocol.

Sample:
Human plasma was extracted by applying a modified Folch method to 25 µL of human plasma, as previously described[1].
Mouse liver was extracted by applying a modified Folch method to 50 µL of tissue.
In above extraction methods organic and aqueous phases were collected and dried. Samples were reconstituted before LC-MS analysis.

HPLC:
Metabolite analysis:
For reverse phase mode, an Agilent 1290 Infinity LC system with an Agilent BC-Poroshell HPH-C18 (2.1 x 100 mm, 1.8 µm) column at 45ºC with a gradient of Eluent A was water with 10mM ammonium formate and Eluent B was 50:45:5. methanol/acetonitrile/isopropanol with10mM ammonium formate. The injection volume was set to 1 µL.
For HILIC phase mode, an Agilent 1290 Infinity LC system was used with a SeQuant ZIC-cHILIC (3 µm,100 Å, 2.1 x 100 mm) column at 40ºC with a gradient of Eluent A was water with 10mM ammonium formate and eluent B was acetonitrile with 0.1% formic acid. The injection volume was set to 1 µL

Lipid analysis:
For reverse phase mode, an Agilent 1290 Infinity LC system was used with an Agilent BC-Poroshell HPH-C18, 2.1 x 100 mm column at 45ºC with a gradient of Eluent A was 60:40 water/acetonitrile with 10mM ammonium formate and Eluent B was 90:10, isopropanol/acetonitrile with 10mM ammonium formate. The injection volume was set to 1 µL.

Mass spectrometry:
A SCIEX TripleTOF 6600 system with a Turbo V ion source and an electrospray ionization (ESI) probe was used to acquire data in both positive and negative ion modes using data-dependent acquisition (DDA) with collision energy set to 35 eV for metabolites and 45 eV for lipids.

RESULTS

A new LC-MS preprocessing method was applied to annotate TOF MS peaks with their neutral mass and to transform raw LC-TOF MS data into a list of analytes [neutral mass (M) group] signal under chromatographic peaks (Figure 1). It applies chemical knowledge to data reduction at an stage before LC-MS peak picking and allows for scoring of MS evidences for each ion type assignments in each MS spectrum of the LC run. Cycle-by-cycle assignment review pinpoints and allows to correct for assignments based on inaccurate m/z shifts and reveals structurally similar analytes with convolved chromatographic profile (Figure 2).

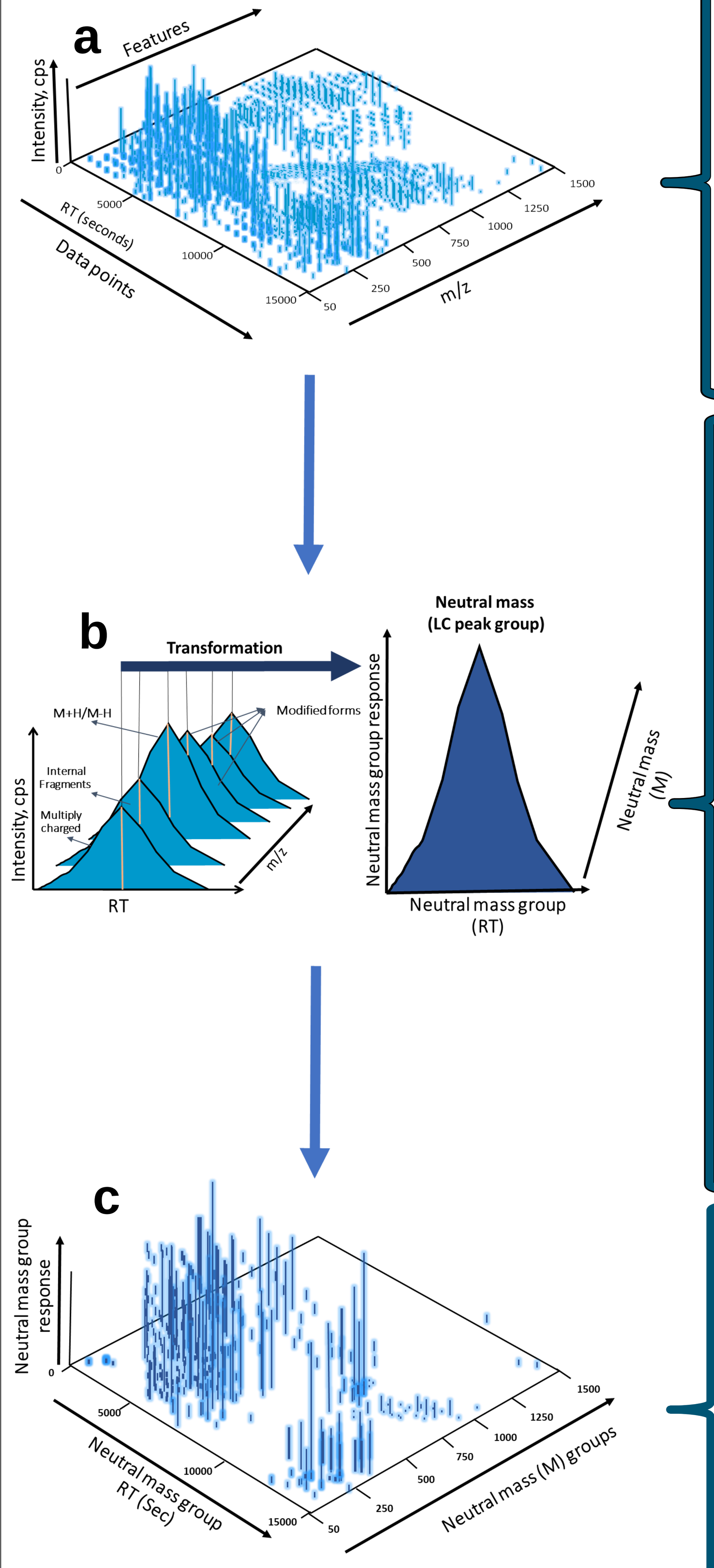
a
Features
Intensity, cps
RT (seconds)
Data points
m/z
b
Transformation
M+H/M-H
Modified forms
Internal Fragments
Multiply charged
Intensity, cps
RT
m/z
Neutral mass (LC peak group)
Neutral mass group response
Neutral mass group (RT)
Neutral mass (M)
c
Neutral mass group response
Neutral mass group RT (Sec)
Neutral mass (M) groups

Figure 1. Schematic of data reduction. (a) Shows an example of 3D m/z (RT) feature map of sample obtained from LC-TOF MS data. (b) Chemical relationship annotation of peaks in every cycle-by-cycle, to derive the neutral mass (M) LC group . (c) Simplified (Data reduction) LC-MS map of 3D M (RT) after processing with the 3D chemical peak finder

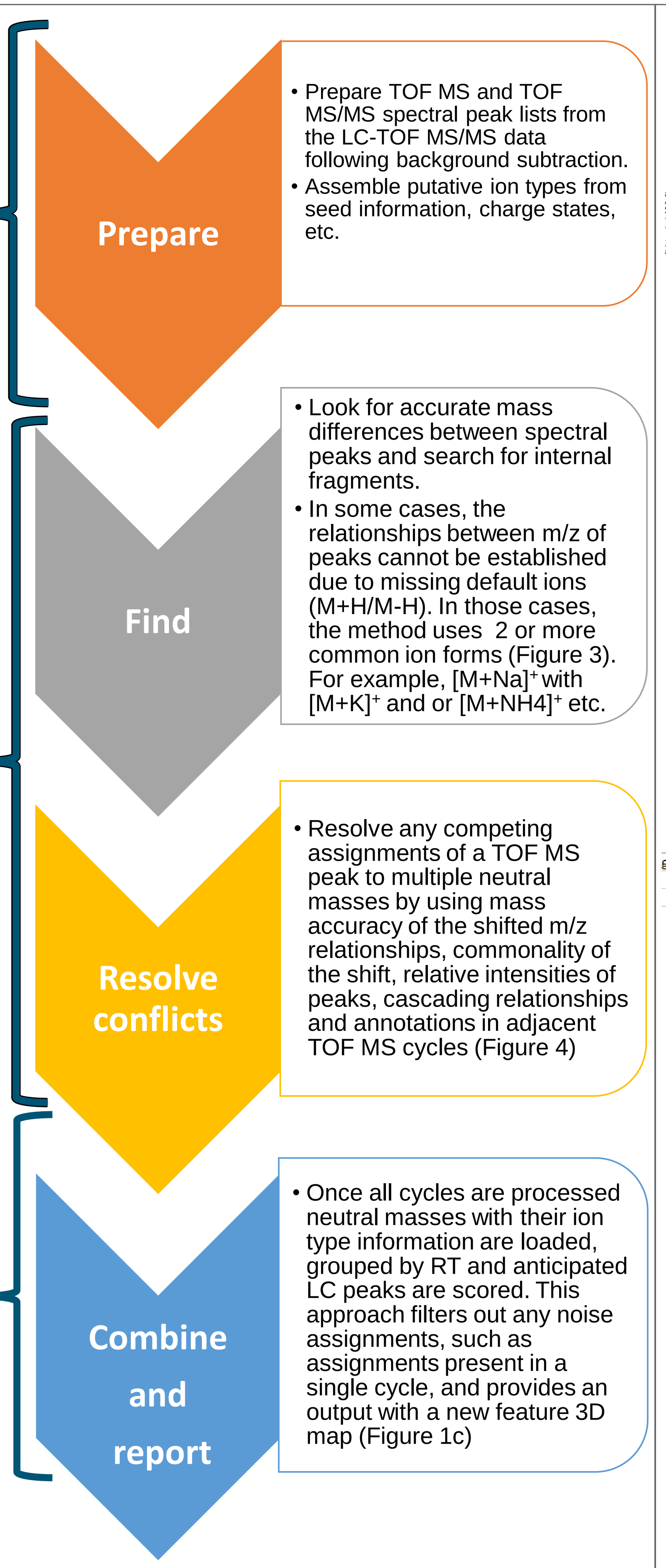
Prepare
• Prepare TOF MS and TOF MS/MS spectral peak lists from the LC-TOF MS/MS data following background subtraction.
• Assemble putative ion types from seed information, charge states, etc.
Find
• Look for accurate mass differences between spectral peaks and search for internal fragments.
• In some cases, the relationships between m/z of peaks cannot be established due to missing default ions (M+H/M-H). In those cases, the method uses 2 or more common ion forms (Figure 3). For example, [M+Na]+ with [M+K]+ and or [M+NH4]+ etc.
Resolve conflicts
• Resolve any competing assignments of a TOF MS peak to multiple neutral masses by using mass accuracy of the shifted m/z relationships, commonality of the shift, relative intensities of peaks, cascading relationships and annotations in adjacent TOF MS cycles (Figure 4)
Combine and report
• Once all cycles are processed neutral masses with their ion type information are loaded, grouped by RT and anticipated LC peaks are scored. This approach filters out any noise assignments, such as assignments present in a single cycle, and provides an output with a new feature 3D map (Figure 1c)

Figure 2. 3D chemical peak finder workflow. Brackets in figure 2 pointing to different levels of data reduction steps in figure 1.

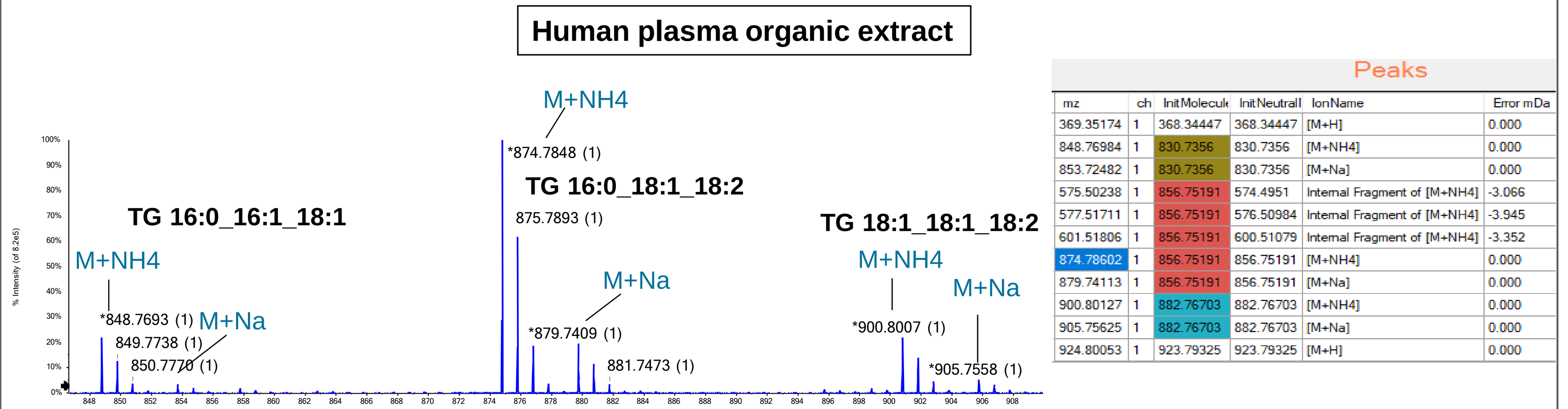

Figure 3. Human plasma organic extract. The triacylglycerol (TG) class is mainly ionized in the form of [M+NH4]+ and [M+Na]+.The 3D chemical peak finder annotates and accurately assigns ion forms to their respective neutral mass groups.

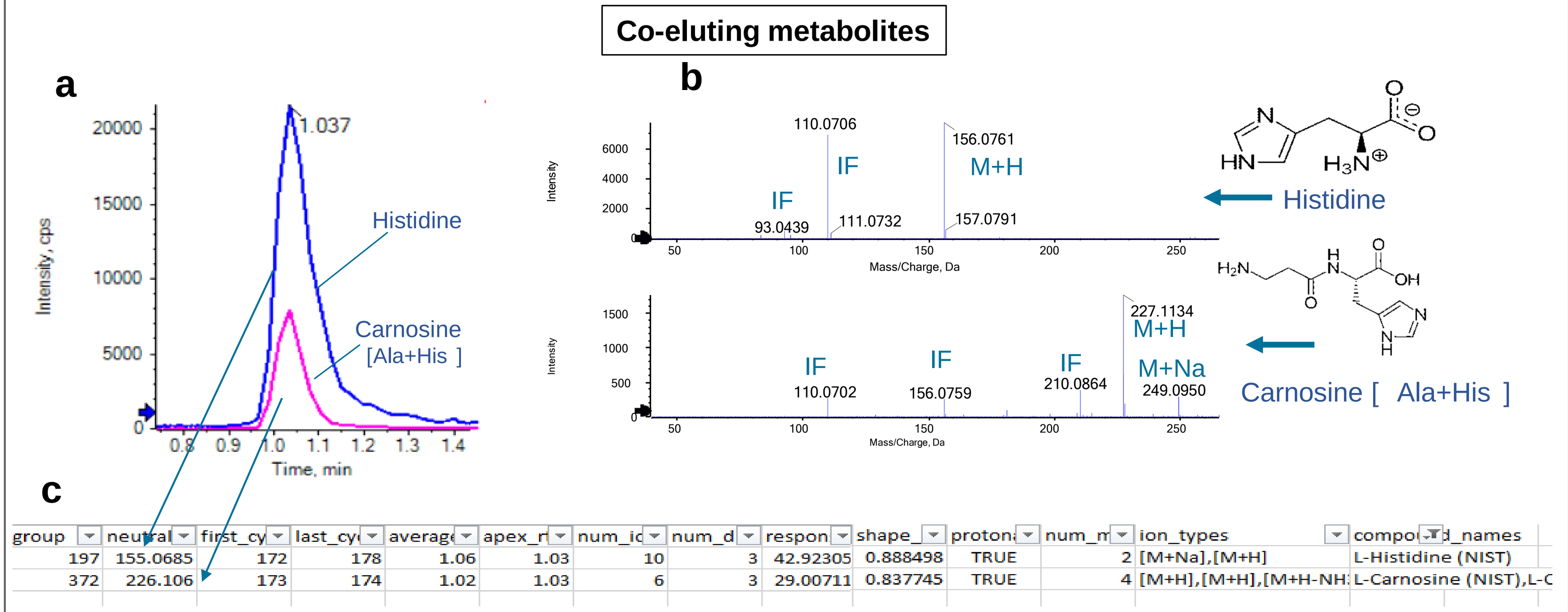

Figure 4. Co-eluting metabolites. (a) LC profile shows the co-eluting metabolites, carnosine and histidine. (b) Single injection analysis shows carnosine internal fragments mimic histidine due to a common substructure. (c) Histidine and carnosine are grouped into different neutral mass groups due to differences in their cycle-by-cycle compound fingerprint.

The application of this workflow on a small molecule omics dataset demonstrated its utility for data reduction. Approximately 75% of the mouse liver metabolome data in positive mode and 60% in negative mode shows 2 or more ion forms in the neutral mass groups (Figure 5).

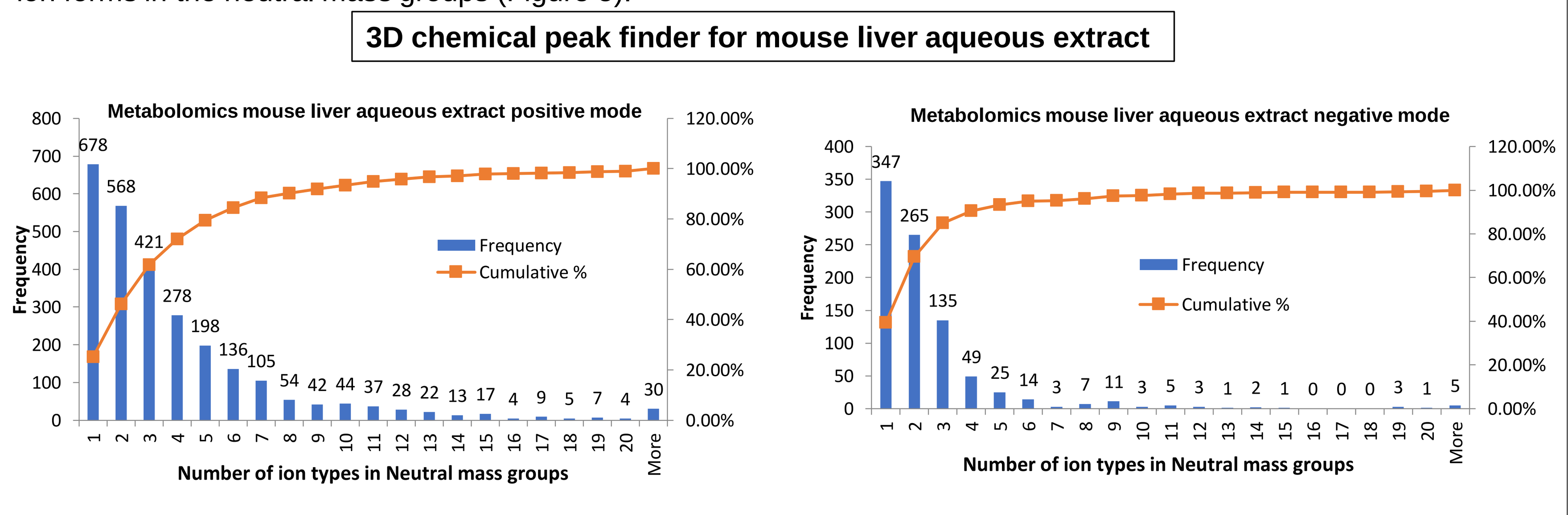

Figure 5. Ion type grouping from 3D chemical peak finder analysis for mouse liver aqueous extract DDA data. Data were reduced by 2.7-fold in positive mode and 2.2-fold in negative mode. These results were achieved by grouping of the mono-isotopic LC-MS features into neutral mass groups.

Figure 6 shows an example of nicotinamide adenine dinucleotide (NAD) that was identified in mouse liver sample showing at least 13 different ion forms grouped to neutral mass (M) 663.1088 Da. Some of these forms, such as IF mimic other analytes. It is therefore important to assign correct neutral mass for analyte identification (Figure 6).

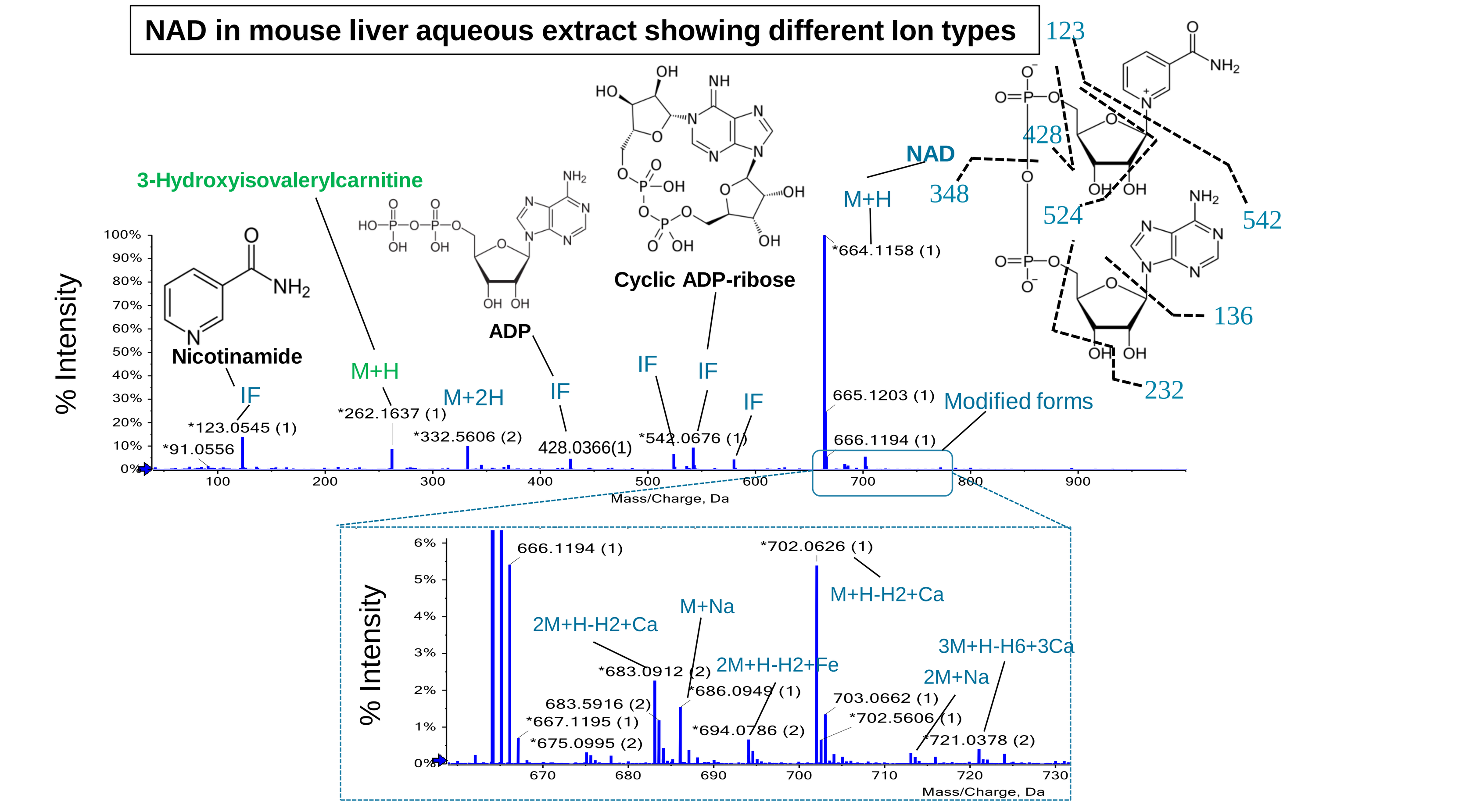

Figure 6. Mouse liver sample showing analyte NAD with different ion types, such as multiple charged, IF, different charge agents (Na, Ca, Fe) and dimers.

CONCLUSIONS

The 3D chemical peak finder was found to be effective for annotating peaks in TOF MS with data using their charge agent and using the information to consolidate the LC-MS data into LC-neutral mass groups. The ability to recognize different charge agents, oligomers and internal fragments allowed for data reduction before additional steps in qualitative and quantitative analyses. The information within neutral mass groups helped to positively identify compound analogues that were not fully resolved in LC, and to assign correct neutral masses to compounds that did not yield protonated or deprotonated peaks. Knowledge of the precursor ion type removed ambiguities in the MS/MS qualitative analysis. Pre-processing with the 3D chemical peak finder enables confident, accurate outcomes for both qualitative and quantitative high-resolution LC-MS data processing pipelines.

REFERENCES

1. Xu. J et. al *Biochimica et Biophyscia Acta* 1862 (2016) 1084-1092.
2. Mahieu NG, Patti GJ *Anal Chem* 2017 Oct 3;89(19):10397-10406.
3. Domingo-Almenara X, Montenegro-Burke JR, Benton HP, Siuzdak G. *Anal. Chem* 2018 Jan 2; 90(1): 480-489.
4. Narayanaswamy P, Teo G, Ow JR, Lau A, Kaldis P, Tate S, Choi H. *Molecular Omics* 2020 Oct 12;16(5):436-447.
5. Stricker T, Bonner R, Lisacek F, Hopfgartner G. *Anal Bioanal Chem* . 2021 Jan;413(2):503-517.
6. Yan X, Stein S Mass Spectrometry Data Centre NIST *ASMS* 2021 poster.

TRADEMARKS/LICENSING