Technical Overview

Agilent
Trusted Answers

Creation of a High-Confidence Lipidomics Personal Compound Database and Library

For targeted data mining and annotation of untargeted high-resolution lipidomics data

Introduction

Lipidomics profiling through mass spectrometry (MS) is complex due to lipid diversity, complex biological matrices, and the presence of isomers and isobars. Despite advances in liquid chromatography (LC) for separating these species, MS-based lipid identification remains difficult. It typically requires detailed analysis of lipid class, composition, and structural features such as double bond characteristics.

To enhance identification accuracy, authentic standards are used to match retention times (RT), accurate mass, and fragmentation patterns. High-resolution MS/MS experiments compare spectral data against libraries for annotation. However, in silico libraries often lack complete structural details, and authentic standards for all lipids are not available, requiring alternative high-confidence annotation strategies.

Previous work details a robust method using a triple quadrupole LC/MS (LC/TQ) with 665 transitions to monitor 763 lipids, covering all major lipid classes.[2] From these data, a personal compound database and library (PCDL) was created, which included accurate mass and RT information for each lipid. By transferring this method to a quadrupole time-of-flight (LC/Q-TOF) instrument, MS/MS spectra were generated for the PCDL, aiding in lipid identification.

This study also incorporated collision cross section (CCS) values from ion mobility experiments on an ion mobility LC/Q-TOF. These values add an extra dimension of structural information, improving discovery-based annotations when combined with accurate mass, RT, and MS/MS data.

This technical overview outlines the processes and experiments underpinning the high-confidence lipid PCDL, ensuring that each entry is thoroughly curated for accurate lipid annotation.

Generating the PCDL

The steps used to build the high-confidence lipid PCDL are summarized.

Step 1. Use Innovative Omics LipidPioneer software version 1.0 to generate molecular formulas from compound names.[3] LipidPioneer software has interactive, class-specific templates that can be used to generate exact masses and molecular formulas of lipid species. Lipids that are not available in LipidPioneer software are searched on the LIPID MAPS database to generate their formulas.

Step 2. Export the .CSV file with lipid formulas generated in Step 1 and RT from the LC/TQ method into the PCDL.

Step 3. Update the database with MS/MS spectra from the Agilent Revident LC/Q-TOF.

Step 4. Add CCS values from the Agilent 6560 ion mobility LC/Q-TOF to the database.

A schematic diagram showing the workflow for creating the high-confidence PCDL is shown in Figure 1. The export of the MS/MS spectra from the LC/Q-TOF into the PCDL is illustrated in Figure 2.

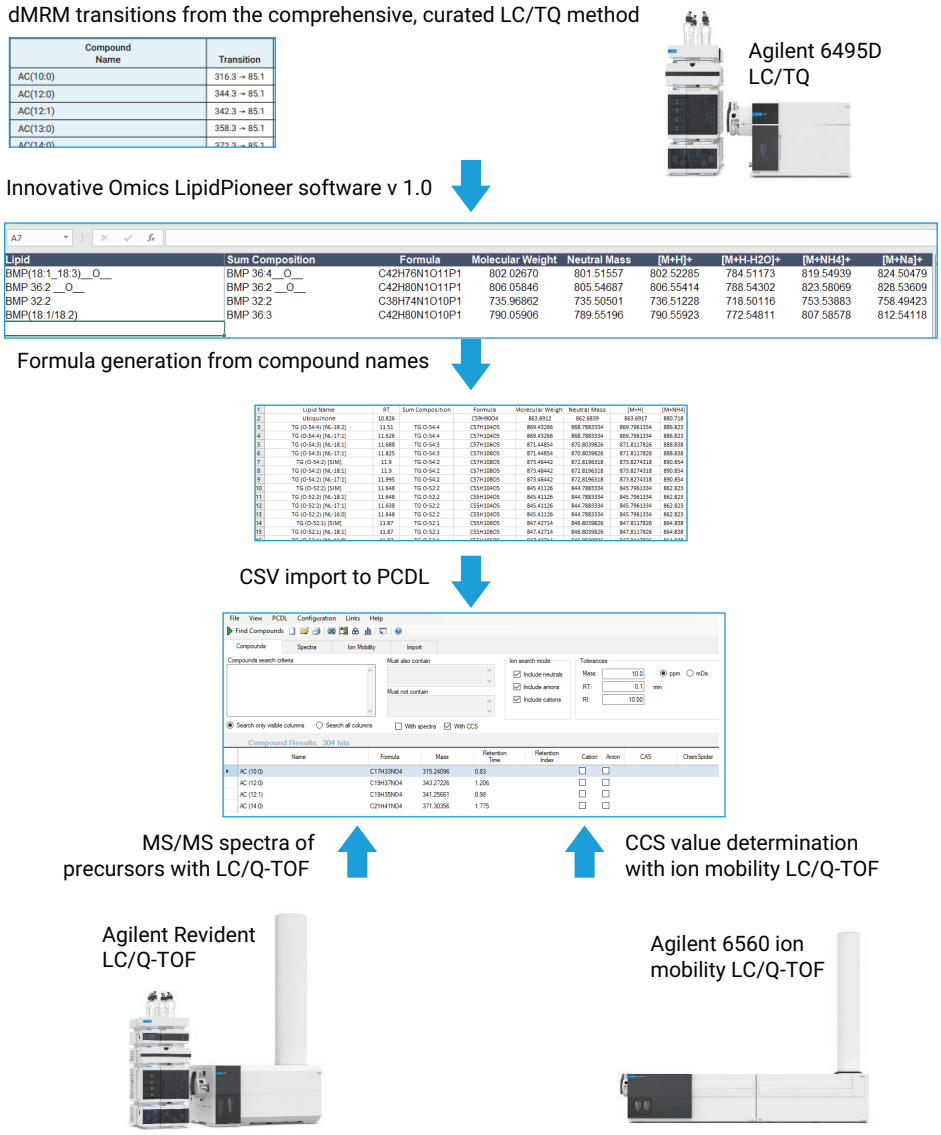
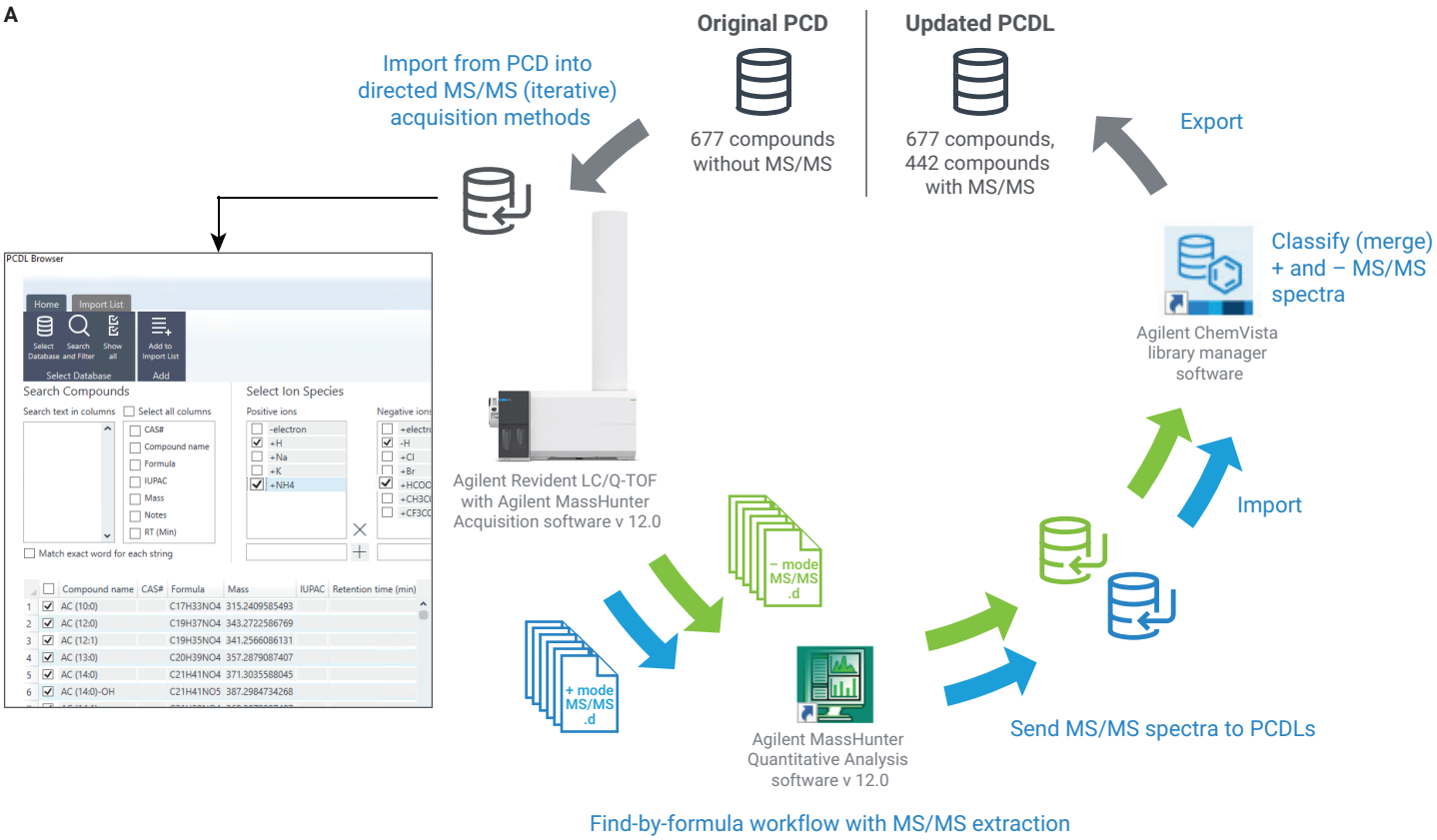


Figure 1. Schematic diagram showing the workflow for creating the high-confidence PCDL.



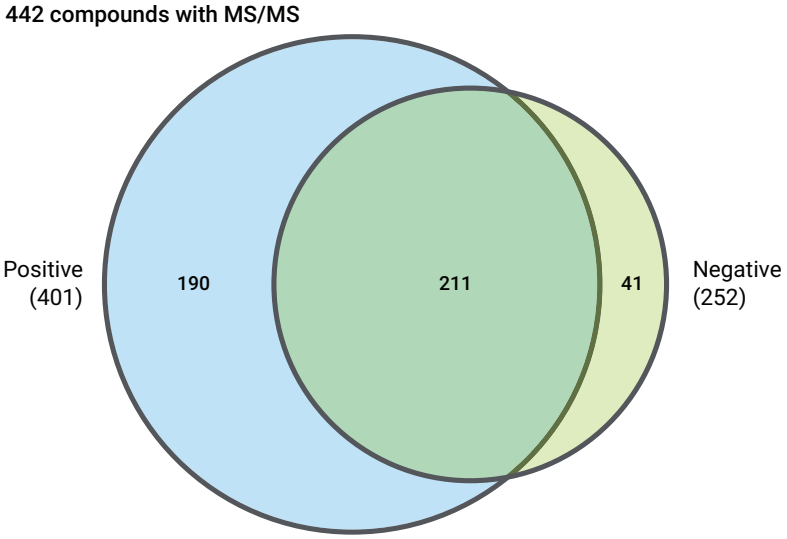


Figure 2. Curation of a lipid PCDL with MS/MS spectra. (A) Workflow illustrating the steps to acquire Q-TOF MS/MS spectra on the NIST SRM 1950 plasma sample, extract MS/MS, send spectra to PCDLs, and to merge the positive and negative libraries with Agilent ChemVista library manager software. (B) Venn diagram showing overlap of compounds with positive and negative MS/MS spectra.

Ensuring high confidence for the PCDL

The high confidence for the PCDL of lipids comes from three especially important experimental projects.

Project 1

The previously established LC/TQ plasma lipidomics method consists of a 16-minute separation that covers 44 lipid classes for a total of 763 lipids in 0.1 µL of plasma. The method was developed by Dr. Peter Meikle's group at the Baker Heart and Diabetes Institute in Australia, and was published as an Agilent application note (publication number 5994-3747EN) in October 2022.[2] The method includes a high level of curation with orthogonal experiments to confirm the annotation level of many MRM transitions, allowing for more accurate lipid specificity.[1] For example, some lipids were characterized with high-resolution Q-TOF experiments and inspection of fragmentation of deprotonated precursors in negative mode and/or lithium adducts in positive ion mode to determine the nature of the esterified fatty acids. Several lipid isomers differed based on the position of the double bond. For example, a phospholipid with a 22:5 fatty acid exists in plasma as either an omega-3 or omega-6 isoform. Standards with the two isoforms were synthesized and run with the chromatographic method for identification.

The synthesis of standards also helped characterize the retention time differences between branched and straight acyl chain isoforms when run alongside a plasma sample. In the case of plasmalogens PC(O) isomeric with PC(P) species, a simple approach was used to rapidly confirm plasmalogen species. Lipid extracts were hydrolyzed with HCl fumes, which selectively hydrolyze species with the vinyl-ether bond running two sequential samples, where one is hydrolyzed, confirming the identity of these species. These and many more studies to determine the identity of lipid isomers are described in detail in references 1 and 2.

Project 2

Relative standard deviations (RSDs) of RTs from interlaboratory studies at four different sites were conducted to determine day-to-day and site-to-site variability of RTs.[4] This is critical, as RT is one of the most important parameters in annotating lipids with high confidence. RSDs for retention time for > 600 lipids were found to be < 0.2%.

Project 3

CCS values were established for the lipid species with a high degree of precision. Repeat injections (six) of the serum extract show RSD for the CCS measurement of each lipid to be < 0.2% for most of the lipids shown in Figure 3. This precision is consistent with an interlaboratory study that showed that $^{DT}CCS_{N2}$ RSD of 0.29% is achievable with the current DTIM-MS technology on the 6560 ion mobility LC/Q-TOF.[5] Because of the high precision in the measurement of CCS, CCS can be used as a molecular identifier in untargeted screening workflows. The precision of the CCS measurement increases the confidence in the measurement of CCS and in using CCS values as an additional filter to remove isobaric interferences, especially in lipidomics studies.

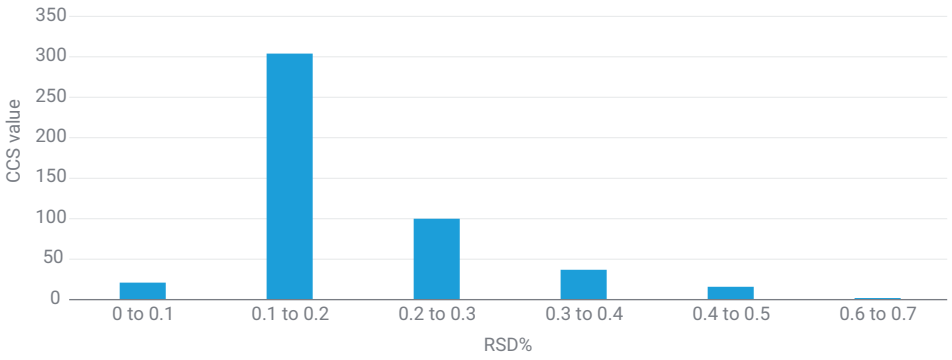


Figure 3. Plot of RSDs for the experimentally determined CCS values for lipids in serum; most lipids have RSD < 0.2%.

The spread of CCS values with *m/z* is shown in Figure 4. Untargeted lipidomics analysis detected approximately 600 lipids across commonly analyzed lipid classes, including phosphatidylcholines, ceramides, diacylglycerols, phosphatidylethanolamines, phosphatidylinositol, sphingomyelins, and triacylglycerols.

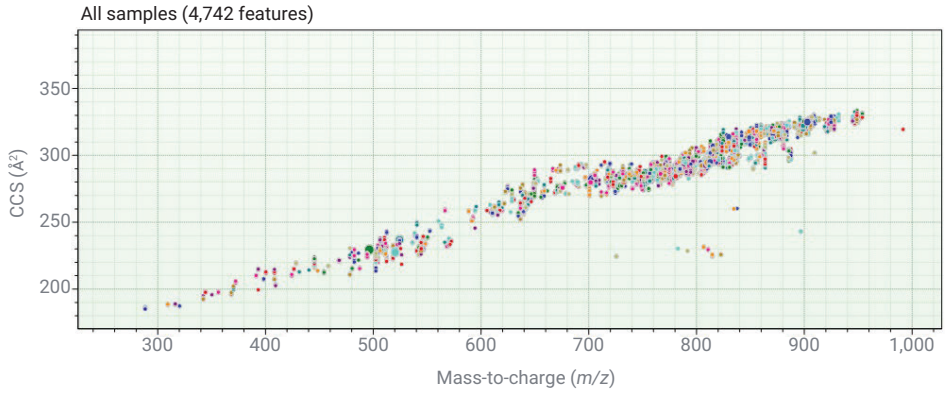


Figure 4. Plot of CCS values versus *m/z* for lipids in serum.

Using the high-confidence PCDL for untargeted lipidomics

A demonstration of how this PCDL can be used to achieve complete and unambiguous characterization of lipids in untargeted lipidomics approaches is detailed in Agilent publication number 5994-7588EN.[6]

Conclusion

- A high-quality, high-confidence database was created for the annotation of lipids.
- The database has accurate masses for 677 lipids, including MS/MS spectra, retention time, and collision cross section (CCS) values, to increase the confidence of annotation in targeted and untargeted workflows.
- CCS is highly reproducible with relative standard deviations < 0.2% for most of the lipids identified. The CCS value improves the accuracy of lipid annotation.

References

1. Huynh, K.; *et al.* High-Throughput Plasma Lipidomics: Detailed Mapping of the Associations with Cardiometabolic Risk Factors. *Cell Chem. Biol.* **2019**, *26(1)*, 71–84. DOI: 10.1016/j.chembiol.2018.10.008
2. Huynh, K.; Sartain, M.; *et al.* A Comprehensive, Curated, High-Throughput Method for the Detailed Analysis of the Plasma Lipidome. *Agilent Technologies application note*, publication number 5994-3747EN, **2021**.
3. Ulmer, C.; *et al.* LipidPioneer: A Comprehensive User-Generated Exact Mass Template for Lipidomics. *J. Am. Soc. Mass Spectrom.* **2017**, *28(3)*, 562–565. DIO: 10.1007/s13361-016-1579-6
4. Sartain, M.; *et al.* An Interlaboratory Evaluation of a Targeted Lipidomics Method in Plasma. *Agilent Technologies application note*, publication number 5994-6830EN, **2023**.
5. Stow, S.; Causon, T.; *et al.* An Interlaboratory Evaluation of Drift Tube Ion Mobility-Mass Spectrometry Collision Cross Section Measurements. *Anal. Chem.* **2017**, *89(17)*, 9048–9055. DOI: 10.1021/acs.analchem.7b01729
6. Mohsin, S.; *et al.* Targeted Data Mining and Annotation of Untargeted High-Resolution Lipidomics Data – A Comprehensive, High-Confidence Workflow. *Agilent Technologies application note*, publication number 5994-7588EN, **2024**.

www.agilent.com

DE80026151

Printed in the USA, July 30, 2024
5994-7627EN

Agilent
Trusted Answers