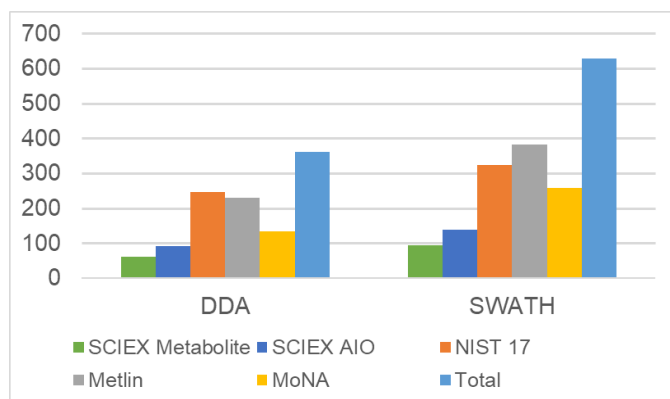# Biomarkers and Omics

SCIEX
The Power of Precision

Improved Metabolite Identification Using Data Independent Analysis for Metabolomics

Using the TripleTOF® 6600 LC-MS/MS System and SWATH® Acquisition

Robert Proos
SCIEX, USA

Comprehensive metabolite identification with MS/MS library spectral matching can be problematic for data-dependent acquisition (DDA) workflows as it often requires multiple injections for each sample to obtain all the MS/MS needed for identification and the sampling for MS/MS is stochastic. Data Independent Acquisition (DIA) workflows are of increasing interest when analyzing complex biological samples due to the comprehensive nature of the acquisition approach.[1] In addition to capturing MS/MS spectra for all detectable analytes, SWATH Acquisition also provides quantitative data at both the MS or MS/MS levels, enabling a comprehensive qualitative and quantitative analysis of metabolites in complex biological samples like plasma in a single injection. Here, the detection rates of metabolites in plasma will be explored using both the DDA and DIA workflows on the TripleTOF 6600 system.

Accurate compound identification remains a bottleneck in the field of metabolomics in the translation of experimental features to the biologic interpretation. Using library searching approaches, where MS and MS/MS information is compared to spectral libraries, is a common approach when exploring complex samples. In this study, an evaluation of several spectral libraries was performed on data from both the DDA and DIA workflows, to compare the number of metabolites accurately identified in an

extracted plasma sample in a single injection workflow.[2] The LC-MS/MS libraries evaluated in this study include NIST 17, MoNA (MassBank of North America), METLIN, and SCIEX libraries.

Figure 1. Comparison of the Number of Identified Compounds. The total number of compounds identified from extracted plasma using different libraries is compared between data dependent acquisition (DDA) and data independent acquisition (DIA) using SWATH Acquisition. In all cases, more compounds were identified using SWATH Acquisition than DDA, with over 50% more compounds identified overall using SWATH Acquisition.

Key Features of SWATH Acquisition for Metabolomics Studies

- SWATH Acquisition with variable Q1 windows provides comprehensive digital maps of the MS and MS/MS of a complex sample, allowing in-depth interrogation for identification and quantification of metabolites

- Speed of acquisition of the X500R QTOF and the TripleTOF 6600 QTOF system enables use of either the top 20 selection in DDA or 20 variable window SWATH Acquisition in DIA

- Resulting cycle time of 651 msec is easily compatible with fast chromatography

- Because of the comprehensive nature of SWATH Acquisition, library matching using broader compound libraries results in the identification of unexpected compounds

- SWATH Acquisition identifies over 50% more compounds than an optimized data-dependent acquisition (DDA) method

- Flexibility of workflow allows the use of smaller targeted libraries for focused interrogation as well as large comprehensive libraries for broad examination
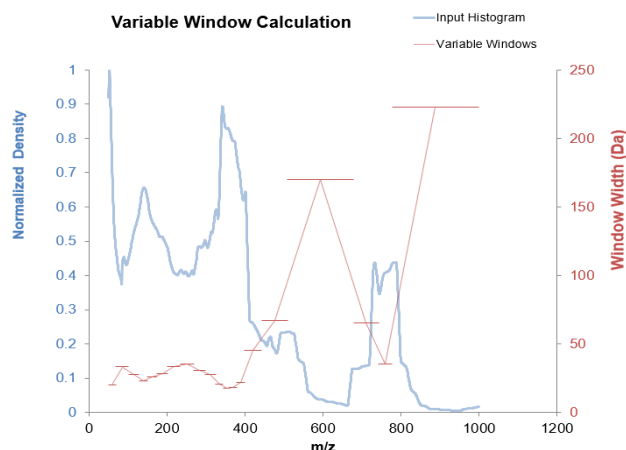
Methods

Sample Preparation: 100 µL pooled plasma (AA 45/32 Phys Control Plasma, SCIEX) was transferred into a 2 mL microtube, 800 µL of methanol was added to extract metabolites and precipitate proteins. Samples were vortexed and then centrifuged at 14,000 RPM for 10 min. An 800 µL aliquot of the supernatant, which contains the metabolites, was transferred to a 2 mL microtube. The sample was dried to a pellet using a CentriVap Concentrator using no heat. The pellet was then reconstituted in 100 µL de-ionized water, centrifuged at 14,000 RPM for 10 min, and the supernatant transferred to an HPLC vial insert.

Chromatography: An ExionLC™ AD HPLC system (SCIEX) with a Phenomenex Kinetex® F5 column (150×2.1 mm, 2.6 µm, 100 Å) at 30 ºC was used with a flow rate of 200 µL/min. Mobile phase A = water with 0.1% formic acid. Mobile phase B = acetonitrile with 0.1% formic acid. The reversed-phase gradient method is listed in Table 1. The injection volume was 3 µL.

Table 1. LC Gradient for Plasma Metabolomics Analysis

Time	%A	%B
0.0	100	0
2.1	100	0
14.0	5	95
16.0	5	95
16.1	100	0
20.0	100	0

Mass Spectrometry: A TripleTOF 6600 system with a DuoSpray™ Source (SCIEX) and Electrospray Ionization (ESI) probe with positive mode ionization was used. For both DDA and SWATH Acquisition methods, TOF MS data was collected between 50 and 1000 m/z with an accumulation time of 100 msec. DDA acquisition was performed using top 20 MS/MS with a threshold of 500 cps, dynamic background subtraction, dynamic accumulation, with a minimum accumulation time of 25 msec. No inclusion or exclusion lists were used for the DDA analysis. SWATH Acquisition was performed with 20 variable Q1 isolation windows, optimized based on the TOF MS data (Figure 2), using the SWATH Acquisition Variable Window Calcluator.[3] Each MS/MS spectrum used a 25 msec accumulation time. The cycle time was 651 msec for both the DDA and DIA methods.

Figure 2. Optimizing the SWATH Acquisition Variable Window Method. Using the TOF MS data collected on the extracted human plasma, windows are optimized such that equivalent ion densities are found within each SWATH window. The SWATH Acquisition Variable Window Calculator was used for this optimization.[3]

Data Processing: Data were processed using SCIEX OS Software for non-targeted peak finding (Table 2) with library matching to the NIST 17, METLIN, MoNA, and the SCIEX All-in-One accurate mass MS/MS spectral libraries. Only experimentally obtained LC-MS/MS spectra were used in the analysis. Compound identification required both MS level precursor mass and MS/MS level spectral matching.

- SCIEX Accurate Mass Metabolite Spectral Library (AMMSL) 2.0 – 650 compounds covering many biologically relevant metabolic pathways, including 93 new compound additions

- SCIEX All-in-One Library – includes AMMSL plus other SCIEX libraries for 4108 compounds total

- NIST 17 – 13808 compounds providing the broadest coverage of exogenous and endogenous compounds with 29507 MS/MS spectra

- METLIN – 14027 metabolites and chemical entities ranging from lipids, steroids, plant and bacterial metabolites, small peptides, central carbon metabolites and toxicants

- MoNA – A meta data-centric, auto-curating centralized collaborative repository of 13038 metabolites

Table 2. Criteria Used for Non-Targeted Peak Finding

Peak Finding Criteria	Threshold
Integration algorithm	MQ4
Minimum peak width	6 points
Minimum peak height	400 cps
S/N threshold	10

DDA Analysis

In the DDA analysis, out of the 14,573 features found in MS, a total of 4864 precursor ions selected for MS/MS analysis. Of those, 2456 were replicate spectra from the same precursor in subsequent scans, resulting in 2408 unique candidate ions fragmented. A summary of the detected features is provided in Table 3.

Table 3. Feature Detection Summary. Results for DDA and DIA acquisition summarized.

| Peak Finding Criteria	DDA Acquisition	SWATH Acquisition
Number of features$	14,573	16,799
Average number of points across peak at baseline	21.4	21.8
Average number of points across peak at half height	11.6	9.8
Total number of triggered MS/MS spectra	4864	N/A*
Number of unique features with MS/MS spectra	2408	N/A*
Number of compounds identified by mass and MS/MS library matching	362	628

$Features found by non-targeted peak finding of TOF MS data in SCIEX OS Software
*In SWATH Acquisition, MS/MS data are acquired for all detectable species

The dynamic background subtraction filter used in the selection process is used to improve the quality of the MS/MS spectra by only triggering on candidate ions near the apex of the peak. It does so by creating an XIC of the candidate ion over the last 3 data points, taking the first derivative, and determining whether the candidate ion is approaching or at the apex of the peak. This filter reduces the total number of candidate ions selected for MS/MS, eliminating background ions and not triggering on peak fronts or tails even when above the intensity threshold.

As the cycle time in the method is fixed, in cycles when less than the maximum number of candidate ions are selected, the 500 msec of MS/MS time is distributed among the selected precursors. The longer accumulation times result in decreased noise, which improves the data quality. The dynamic accumulation option also used in this method further improves data quality by allocating more time to precursor ions of lower intensity. The overall result is high-quality MS/MS data which allows for improved spectral library matching.

In the single injection DDA analysis with 2408 unique candidate ions on which MS/MS data were obtained, 362 of these features were identified by both precursor mass and MS/MS library matching using SCIEX OS Software for analysis. An example of a metabolite identification from DDA acquisition is provided in Figure 3.
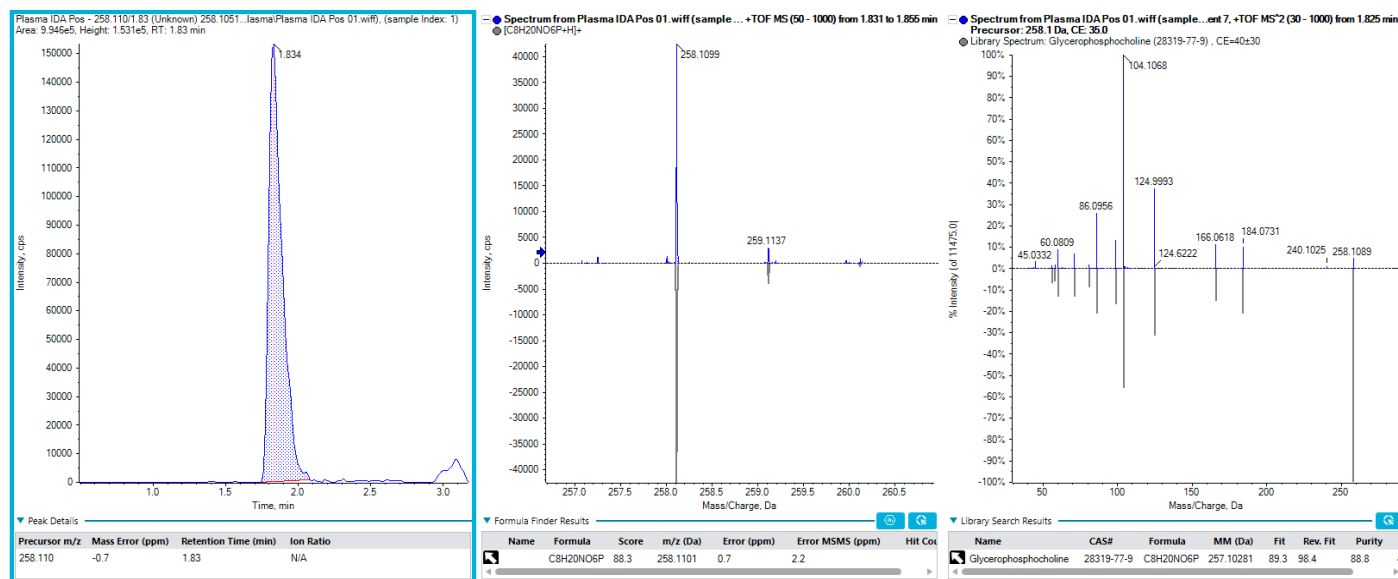
Figure 3. Example of a Compound Identified using DDA. Glycerophosphocholine was identified in extracted human plasma using DDA acquisition with library matching to the SCIEX Accurate Mass Metabolite Spectral Library 2.0. Identification is based on MS mass accuracy and isotope pattern matching, and also on matching of MS/MS to the library spectrum shown here.
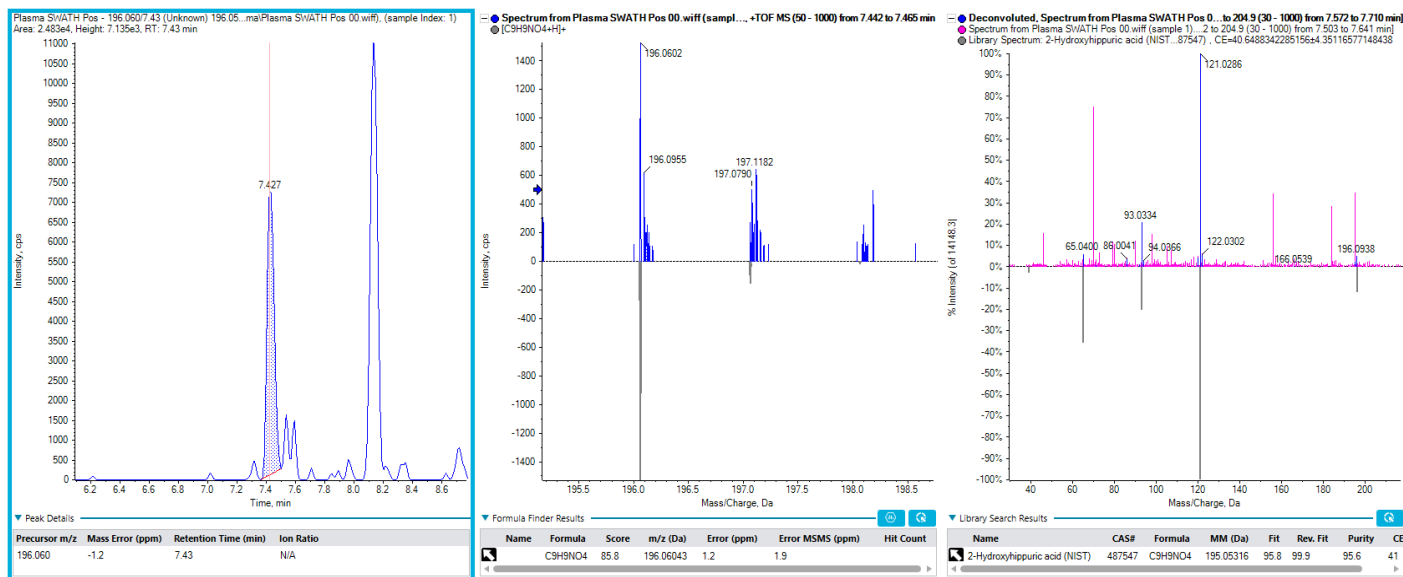
Figure 4. Example of a Compound Identified using SWATH Acquisition. 2-hydroxyhippuric acid was identified in extracted human plasma using SWATH Acquisition with library matching to the NIST 17 library. In the MS/MS pane (far right) the pink peaks in the acquired spectrum (top) were removed by the software as not being associated with the m/z 196.060 precursor by PCVG. The blue peaks which were associated with m/z 196.060 precursor were a match to the library spectrum for 2-hydroxyhippuric acid. This compound was not identified in the DDA analysis.

DIA Analysis using SWATH Acquisition

Contrary to DDA acquisition where a precursor ion is individually selected and isolated for fragmentation, in DIA, wider isolation windows are used such that MS/MS is acquired on all detectable compounds in the sample. Variable window SWATH Acquisition ensures an optimal balance between sensitivity and specificity is achieved for each sample type.[1] In this method the SWATH Acquisition Variable Window Calculator[3] was used to determine optimum window sizes by adjusting the window size to normalize the total number of ions per window. A summary of these windows is presented in Figure 2.

From the SWATH Acquisition data, all detectable features can be searched against spectral libraries for identification. The assignment of product ions to specific precursors is determined by the software using background subtraction or alignment of product ions to precursor ions by PCVG. In this example, all 16,799 features found in the SWATH Acquisition data were searched against the spectral libraries, compared to only 4,864 features with spectra from the DDA analysis.

The additional 12 thousand features not selected by DDA are likely low-intensity analytes as they were not triggered in the top 20 DDA method. The SWATH Acquisition data yielded an additional 266 compounds that were identified, increasing the total number of compounds from 362 to 628 identified by precursor match and MS/MS spectral match from a single injection of extracted human plasma. An example of a compound

identified by SWATH Acquisition that was not identified by the DDA method is shown in Figure 4.

Comparing Spectral Libraries

With SCIEX OS Software, a variety of libraries can be used to search the DDA or SWATH Acquisition data. Here, a number of different libraries varying in size and source were used to investigate the numbers of metabolites observed from DDA and SWATH Acquisition data. Table 4 and Figure 1 summarize the results of all the library searches on both datasets. Not surprisingly, the large libraries provide more identifications. Also the SWATH Acquisition data provides consistently higher identifications over DDA, also as expected due to the comprehensive nature of the data acquisition.

Table 4. Comparison of the Number of Compounds Identified. Using a variety of libraries of increasing size, the number of identified analytes from a plasma sample were compared for DDA vs SWATH acquisition.

Library	DDA Acquisition	SWATH Acquisition
SCIEX AMMSL 2.0	62	94
SCIEX All-in-One libraries	93	140
NIST 17	248	325
Metlin	230	384
MoNA	134	259
Total	362	628

Next, to understand the differences provided by using different libraries, the compounds found by each approach were compared and summarized in Figure 5. As would be expected, the larger the spectral library, the more library matches that could be obtained as seen in Table 4. As plasma often contains more than just endogenous metabolites, using libraries with broader coverage can provide more feature identifications. However if focusing on human biology is a key component of the study, using library focused on endogenous metabolites (like the AMMSL) can simplify the biological study.

Also, it was found that libraries varied quite widely in the compound identifications found, even between metabolite focused libraries. In the Venn diagram (Figure 5), there were 249 features identified using the METLIN library that were not identified using only the MoNA library. Similarly, there were 124 features identified using the MoNA library that were not identified using only the Metlin library. There were 508 compounds identified using these two libraries; however, only 135 of those compounds were identified using both libraries.

It is also shown in Figure 5 that there were an additional 120 features that were identified using the NIST 17 and SCIEX All-in-One libraries that did not have MS/MS spectral matches in either the METLIN or MoNA libraries. The NIST 17 and SCIEX all-in-One libraries include compounds from additional classes including medications, extractables, leachables, pesticides, illicit drugs, foods, natural products, PFAS, and more, which could be present in biological matrices. An example of one of these library matches is presented in Figure 6, stachydrine is found in citrus
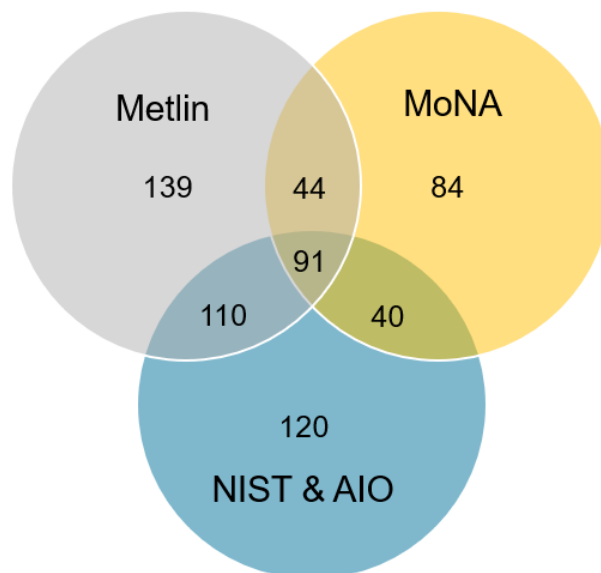
Figure 5. Library Overlap. The distribution of the 628 compounds identified using SWATH Acquisition is compared between the individual and combined libraries.

fruit and therefore not an endogenous metabolite in human plasma, but a possible biomarker for citrus fruit consumption.
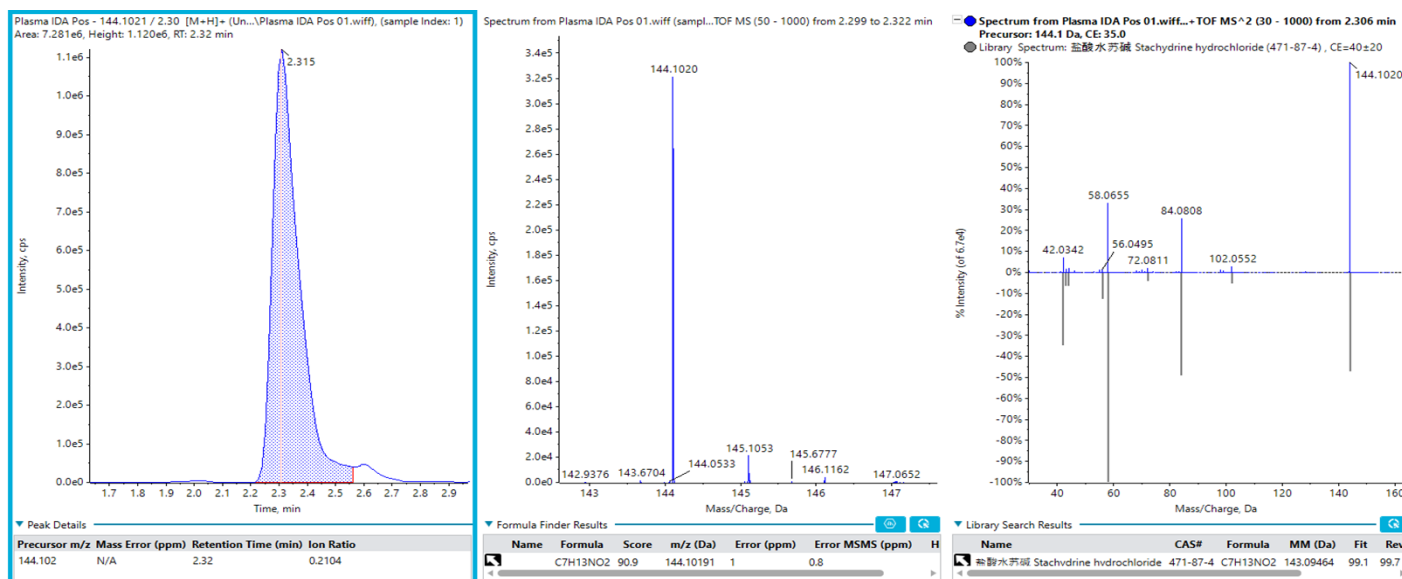
Figure 6. Identification of Stachydrine from DDA Data. Stachydrine is a constituent of citrus fruits and, therefore not included in the AMMSL library focused on core biological metabolites. This was identified in extracted plasma matching to the both the SCIEX All-in-One and NIST 17 libraries.
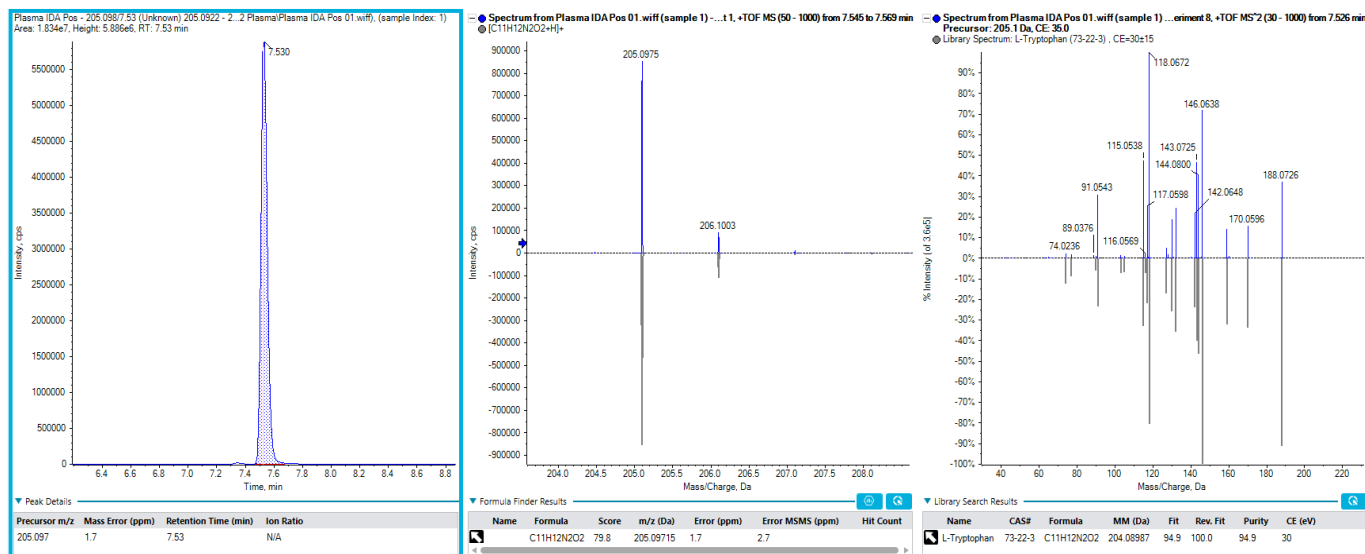
Figure 7. The Identification of Tryptophan in Extracted Plasma Matching to the SCIEX Metabolite Library.

Conclusions

A simple 20-minute RP-LC method was used to acquire DDA and SWATH acquisition data on an extracted human plasma sample using a TripleTOF 6600 System. The top 20 DDA method was able to identify 362 of the features on the basis of precursor mass and MS/MS spectral matching. The 20 variable window SWATH Acquisition method resulted in an additional 266 features being identified, for a total of 628 compounds identified in a single injection. This study has demonstrated that variable window SWATH Acquisition can be useful for improving compound identification when using untargeted acquisition approaches for metabolomics.

In global metabolomics analysis of biofluids, and especially of human plasma or urine, numerous non-metabolite analytes will be present from sources such as the environment, food, and medications. These analytes have the potential to complicate data analysis appearing as unknowns, possibly showing similar profiles to biomarkers or up or down regulation differences. This study demonstrated that the inclusion of non-metabolomics libraries can increase identification of other features in the dataset. Additionally, it was shown that even across metabolomics focused libraries, significant differences in metabolite coverage still exist.

References

1. Improved Data Quality Using Variable Q1 Window Widths in SWATH® Acquisition. SCIEX Technical note RUO-MKT-02-2879-B.

2. SWATH® Acquisition Improves Metabolite Coverage over Traditional Data Dependent Techniques for Untargeted Metabolomics, SCIEX Technical Note RUO-MKT-02-7128A

3. SWATH Acquisition Variable Window Calculator - https://sciex.com/software-support/software-downloads

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices