thermo scientific

SMART NOTE

AcquireX Intelligent Data Acquisition Technology

Collect more meaningful data, not just more data

How can I address small-molecule study challenges while acquiring the fragment-rich spectral data I need to characterize my samples?

Based upon experimental needs, several fully automated approaches to performing precursor selection for highest-quality high-resolution accurate-mass (HRAM) MS and MS^n data acquisition, and ultimately high-confidence characterization of small-molecule samples, are provided by Thermo Scientific™ AcquireX™ intelligent data acquisition workflows.

What are the primary goals and challenges associated with small-molecule studies?

Small-molecule studies focus on two goals: 1) detecting and interrogating all compounds of interest in the samples analyzed and 2) assigning the most plausible structures to these compounds. As most sample characterization methods utilize untargeted data acquisition methods, successful compound identification and/or structural annotation relies on acquiring high-quality MS^n data as defined by high resolution and accurate mass measurements of precursors and product ions to fingerprint compounds and reduce false positives. In addition, the HRAM MS^n data must be exhaustive to ensure accurate profiling of all sample-specific compounds is accomplished.

Carrying out these studies is extremely challenging due to the large number of MS features produced by background and matrix compounds, making it difficult to identify which features are unique to the study samples, and if these unique features (the features of interest) have been exhaustively profiled. Manually creating a comprehensive exclusion list of background and matrix features can be extremely laborious.

ThermoFisher SCIENTIFIC

An additional challenge is ensuring that sufficient precursor and tandem MS data are obtained for all compounds of interest in order to sufficiently differentiate the precursor features enabling compound identification and structural elucidation. Addressing this challenge requires the use of complex data acquisition schemes. In addition, the mass spectrometer used must facilitate fast acquisition rates that maximize ion flux, resulting in high-quality MS, MS/MS, and additional MS^n data acquisition.

Setting up MS data acquisition methods to efficiently achieve small-molecule study goals while addressing these challenges often involves numerous replicate injections, and labor-intensive, iterative method creation and revisions. AcquireX intelligent MS^n data acquisition addresses these challenges with automated profiling of multiple samples and exhaustive interrogation of replicates. AcquireX intelligent data acquisition provides a fully automated method to perform exhaustive precursor selection for highest-quality HRAM MS and MS^n data acquisition. As a result, the number of compounds of interest sampled with distinguishable fragmentation spectra available for sample characterization is significantly increased. With automatic background exclusion, characterization of low-abundance analytes is substantially improved when using AcquireX data acquisition workflows.

What are the general categories of small molecule compounds that need to be addressed when making identifications?

Small-molecule compounds in experimental samples fall into three categories: “knowns”, “known unknowns”, and “unknown unknowns”. Compounds are considered “knowns” when there is prior knowledge of the compounds expected, comprised of well characterized chromatographic and mass-spectral features that define their chemical structures. The LC-MS-based features for matching knowns are found in spectral libraries.

“Known unknowns” are compounds that are not expected, but that can be identified using extended library search strategies that result in matched product ion spectra.

“Unknown unknowns” are often easily measured, but have no *a priori* information associated with them to allow a confirmatory match. Analysis of unknown unknowns is challenging because there is not a robust method to determine how many unknowns are present in a sample. In addition, unknown unknowns generally comprise a large percentage of the compounds of interest. Even when MS and MS/MS data are acquired for an unknown unknown, existing search strategies often do not result in a definitive spectral library match.

Regardless, researchers must still develop and run methods which enable efficient identification and characterization of all three categories of compounds in a sample or set of samples.

How are the features associated with compounds of interest sampled?

Small-molecule samples contain compounds associated with the background and matrix, as well as compounds of interest resulting from biological extractions, manufacturing processes, and extractables and leachables. Across the chromatographic gradient, LC-MS experiments typically measure more than 100,000 features. Features are defined as measured *m/z* values and ion intensities as a function of chromatographic retention time. Each compound may generate many features associated with isotopes, adducts (e.g., ionized by Na or K as opposed to protonation), multimers, and in-source fragmentation products. In addition, the precursor features associated with the background matrix and the compounds of interest are present across wide dynamic ranges.

Performing comprehensive compound characterization requires enhanced sampling strategies to allow post-acquisition data processing to successfully convert precursor features into compounds of interest. The most common LC-MS-based strategy is to apply data dependent acquisition (DDA) methods that use a combination of full-scan MS and MS/MS data acquisition events. A full-scan mass spectrum is used to detect precursor *m/z* values that trigger tandem MS acquisition prior to acquiring another full-scan mass spectrum (Figure 1).

This well-established method uses dynamic exclusion lists to reduce the probability of resampling the same precursor within a user-defined retention time window (typically the elution peak width). Efforts to achieve exhaustive precursor sampling have utilized replicate sample injection to leverage stoichastic sampling associated with DDA methods. However, relying solely on replicate injections can substantially increase the injections needed to acquire and process enough data files to adequately profile the sample.

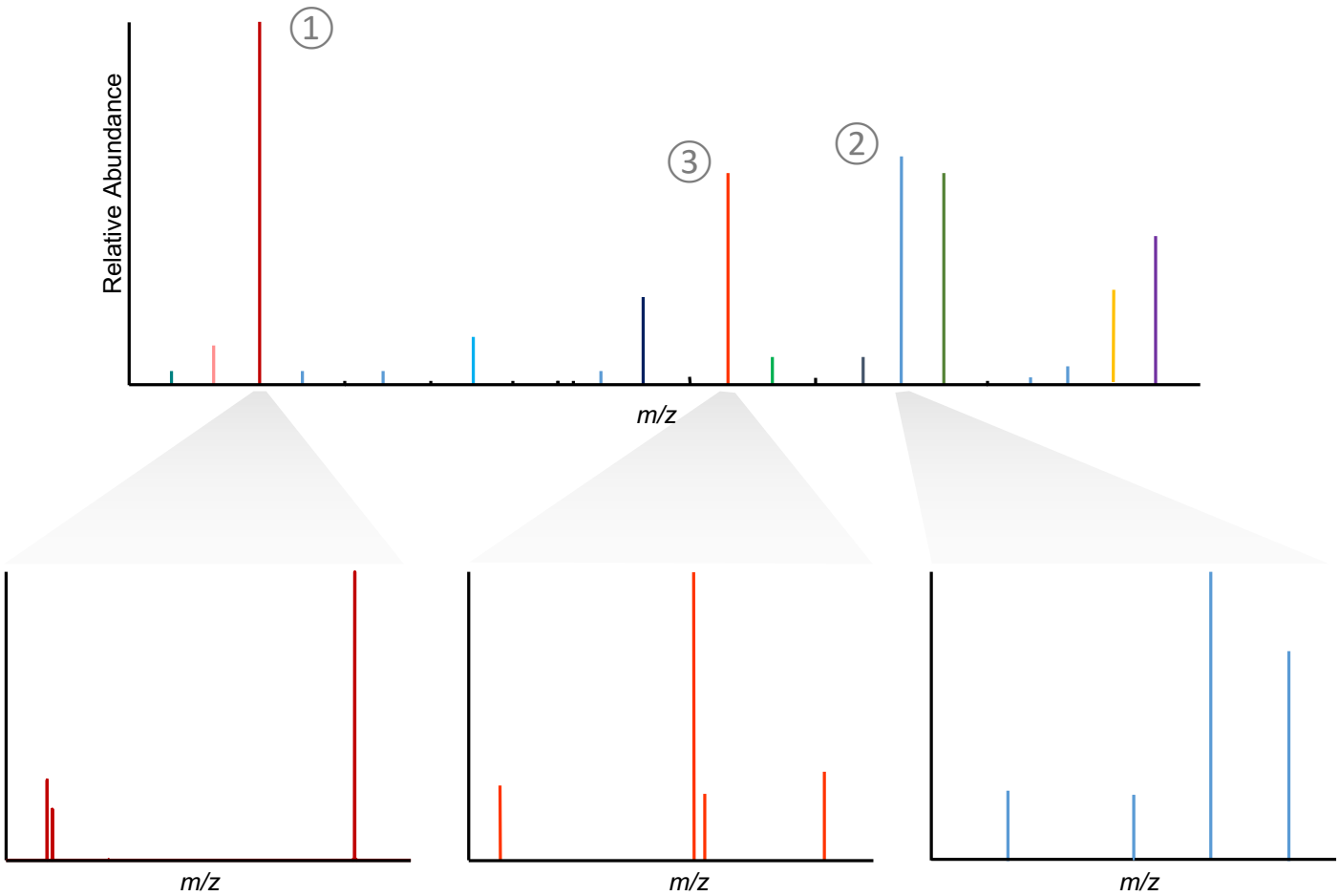

Figure 1. Representation of the DDA method where a full-scan HRAM mass spectrum is acquired and processed in real time to identify precursors to sample via MS/MS data acquisition (top). The bottom of Figure 1 shows the three product-ion spectra for the three precursors selected in the full-scan mass spectrum.

How can I increase sampling of compounds of interest?

Small-molecule analyses are usually performed as part of a study consisting of multiple samples, which include the samples of interest as well as matrix and/or background samples. LC-MS analysis of the matrix and/or background samples is used to map "irrelevant" precursor features for exclusion (Figure 2). The excluded list of features—the exclusion list—is imported into the DDA method prior to LC-MS data acquisition of the samples of interest. For background and matrix samples, there is no need to acquire MS/MS spectra, because the goal is only to map precursor *m/z* features to retention time. High-resolution accurate-mass (HRAM) MS data are of high value in defining and differentiating background and matrix features—particularly isobaric features—from the features of interest.

When used in conjunction with automated, real-time intelligent decision making, the exclusion list enables the MS data acquisition method to bypass precursor features. This capability focuses MS/MS acquisition time on sampling the features that are unique to the samples of interest. In Figure 3, the most abundant precursor features may be attributed to the background/matrix and are excluded from consideration, enabling lower-intensity precursors not on the dynamic exclusion list to be selected and sampled using DDA.

Automated, real-time intelligent decision-making using an exclusion list enables more compounds to be sampled in fewer replicate sample injections than DDA alone, but the number of features measured in the sample remain numerous and cover a wide dynamic range. Additional replicates may be required to increase unknown identification, and the approach still relies on stochastic sampling to maximize breadth and depth of unknown coverage.

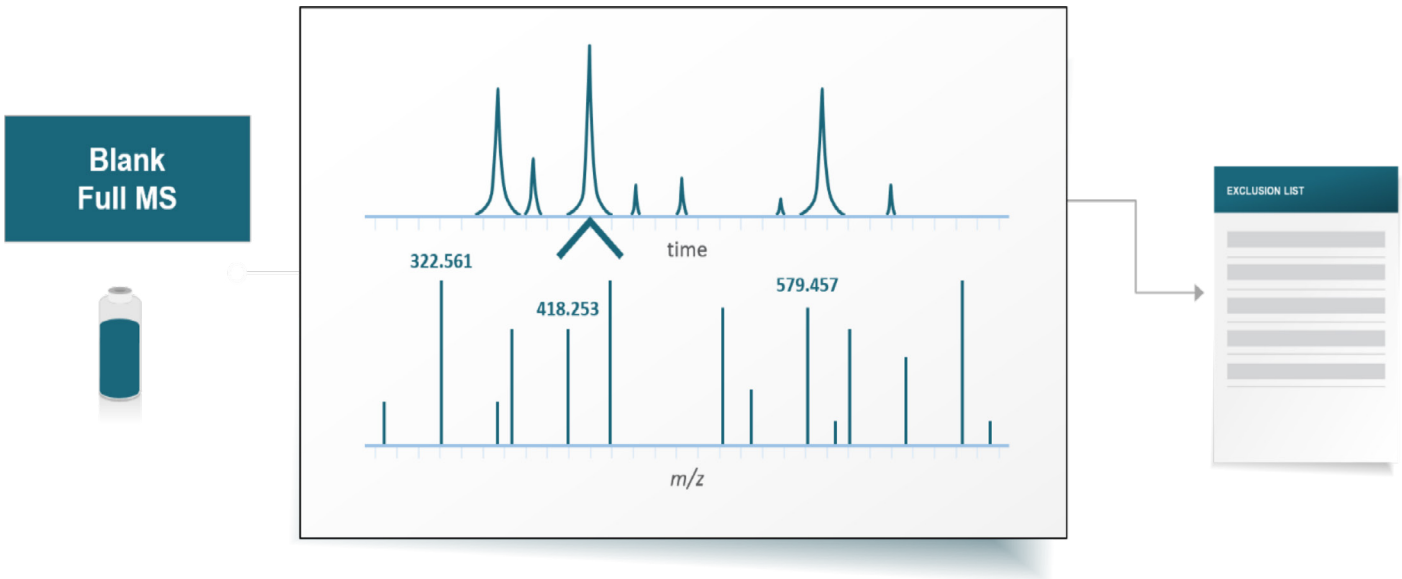

Figure 2. Representation of the precursor mapping method used to automatically create an exclusion list. The exclusion list includes *m/z* values as a function of retention time.

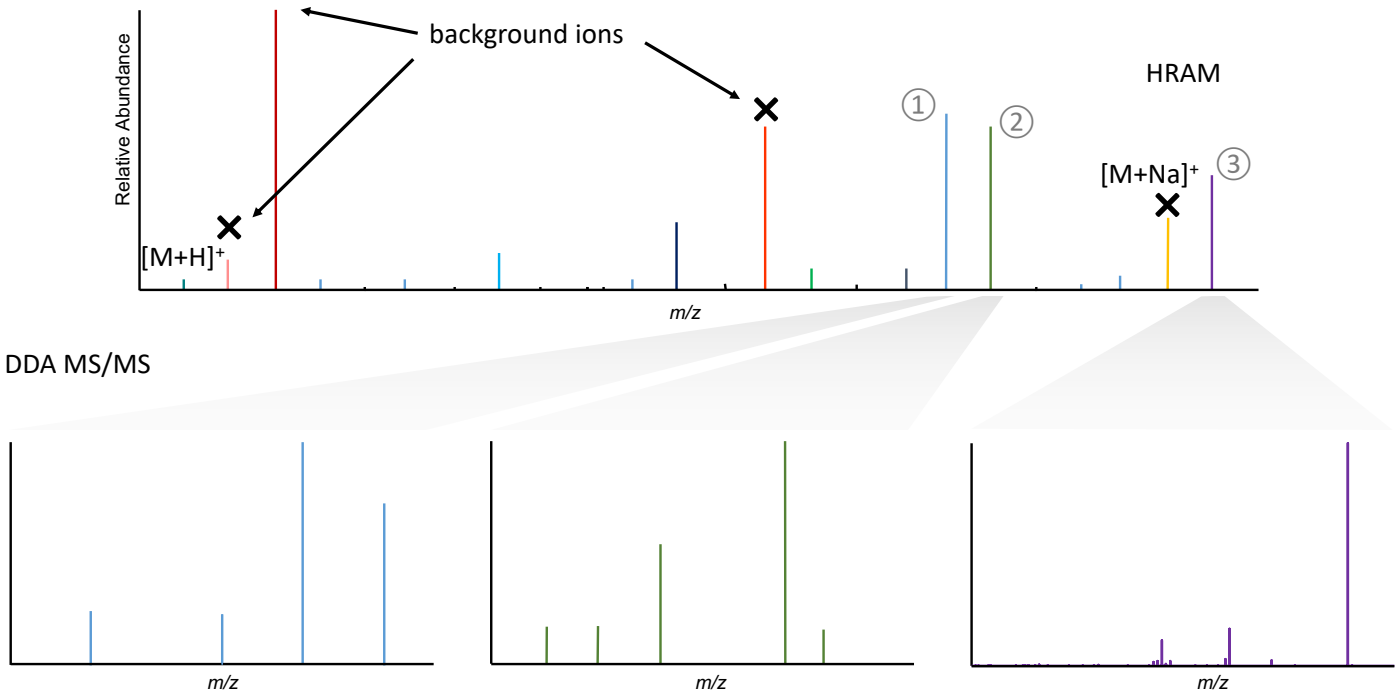

Figure 3. Representation of real-time precursor selection for subsequent tandem MS acquisition using DDA methods with an exclusion list. Precursors marked with an "X" represent features originally mapped to the background sample.

Creation of an inclusion list in addition to an exclusion list can further increase characterization of the compounds of interest. In particular, the inclusion list increases the dynamic range of precursor sampling for DDA. To create an inclusion list of precursor features of interest, representative samples are analyzed using the same LC-MS methods used to create the exclusion list. The precursor features detected from the sample are then compared to the precursor features associated with the background and/or matrix sample to create the inclusion list.

Chromatographic performance impacts the creation of inclusion and exclusion lists. Because two different samples are analyzed separately and compared, the stability and reproducibility of the chromatographic separation are critical to success. Feature mapping requires measurement of precursor *m/z* ions at a specific retention time. If the reproducibility of the gradient (reflected in retention time shifts) is poor (≥ 1%), then the associated retention-time window becomes larger, thus decreasing the selectivity of exclusions and inclusions. Consequently, isobaric compounds of interest that closely elute would be missed. In addition, adequate mass-spectral resolution is required to define precursor features, particularly isobaric species that have mass differences less than 10 ppm.

How can I automate precursor selection for highest-quality data acquisition?

AcquireX intelligent data acquisition using a Thermo Scientific™ Orbitrap™-based mass spectrometer provides a fully automated method to perform exhaustive precursor selection for the highest-quality HRAM MS and MS^n data acquisition. Collection of high-quality LC, MS, and MS^n data facilitates comprehensive compound identification and structural elucidation. By reducing the time needed for MS method development, data acquisition, and processing, AcquireX workflows enable researchers to accelerate their small-molecule studies.

Figure 4 provides an overview of the AcquireX deep scan workflow. The novel sequence acquisition editor enables the user to define the samples that will be used to automatically generate the comprehensive inclusion and exclusion lists. Background matrix or solvent blanks are analyzed first in order to map "unimportant" precursor features to the exclusion list as a function of retention time and measured ion intensity. This is followed by LC-MS analysis of a representative study sample to map all "important" precursor features as a function of retention time and intensity.

This approach identifies features corresponding to the target compound, allowing the user to select the preferred ions for each compound (commonly the protonated $[M+H]^+$ species) for populating the inclusion list, thus reducing redundancies and the number of targets for tandem mass spectral acquisition. The two precursor maps are overlaid and compared to determine which features are placed on the inclusion and exclusion lists and subsequently uploaded to the MS^n instrument method. In the case of compounds present in both the background and/or matrix and the sample(s), AcquireX data acquisition can still act on such compounds provided the measured relative precursor ion intensity ratio is greater than the user-defined threshold.

In addition to inclusion and exclusion list creation, AcquireX deep scan workflow performs exhaustive LC-MS and tandem mass-spectral acquisition for all unique features on the inclusion list. As the list is created and updated, it is automatically imported into the DDA method to drive intelligent collection of HRAM MS and MS^n data. Figure 4 shows an automated replicate analysis for profiling all unique compounds on the inclusion list. The first replicate uses the inclusion list to determine which precursors are targeted for tandem mass spectral analysis. After RAW file acquisition, the data are automatically processed to determine which compounds have been sampled as a function of retention time, and then are moved from the inclusion list to the exclusion list. The updated lists are automatically imported back into the experimental method file and the second replicate injection is performed. This process repeats for a user-defined number of injections.

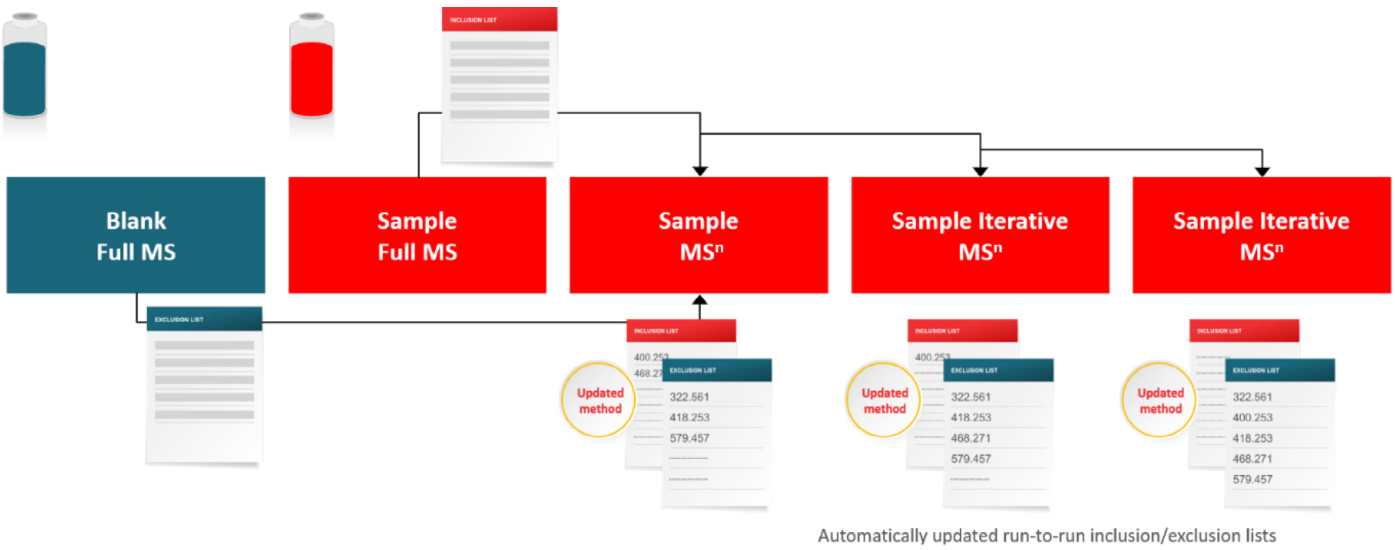

Figure 4. AcquireX deep scan acquisition workflow, one of four AcquireX workflows.

How does the enhanced data quality obtained when using AcquireX workflow facilitate small-molecule identification?

AcquireX intelligent data acquisition maximizes the capabilities of the Orbitrap mass spectrometers for exhaustive small-molecule sample characterization. LC-MS mapping with AcquireX increases data acquisition efficiency because the software predominantly identifies and samples compounds of interest in their most abundant ionized form. This significantly reduces the length of the inclusion list and thus the number of precursors sampled per DDA cycle, and the number of technical replicates analyzed. By identifying additional precursor features associated with a compound, such as $[M+Na]^+$, $[M+K]^+$, etc., and excluding them from the inclusion list, lower-level compounds of interest can be sampled in fewer replicate injections (Figure 5).

There is also some benefit to analyzing all features using re-injections. For example, if sample amount and re-injections is not an issue it may be better to get data on all features (to cover any instances of mis-assigned adducts). Because AcquireX data aquisition utilizes automation and intelligence to exhaustively identify and sample compounds of interest, more instrument time can be devoted per compound to increase data quality.

AcquireX continuously evaluates the precursor *m/z* features sampled within a replicate injection and updates the list prior to the next replicate sample injection, so more of the DDA cycle time can be spent on new compounds instead of reanalyzing the same compound or related features. Thermo Scientific™ Orbitrap™ mass analyzer resolution is directly proportional to the transient detection time; thus, acquiring MS spectra using a resolving power of 120,000 takes twice as long as a setting of 60,000 (For example: 128 vs. 64 msec.). However, a higher resolution setting increases the separation between features, improves determination of isotopic fine structure, and ensures more accurate mass measurements.

AcquireX data aquisition also increases tandem mass spectral acquisition. Because new features are continuously selected for tandem mass spectral analysis and the cycle time spent per compound can be increased, HRAM MS/MS, (and for MS^n equipped Orbitrap mass spectrometers, MS^3, and in many cases MS^4) product ion spectra can be acquired. The resulting MS^n ion tree data acquisition increases the richness of product ions and more accurately defines the unimolecular decomposition pathways that are important for structural elucidation.

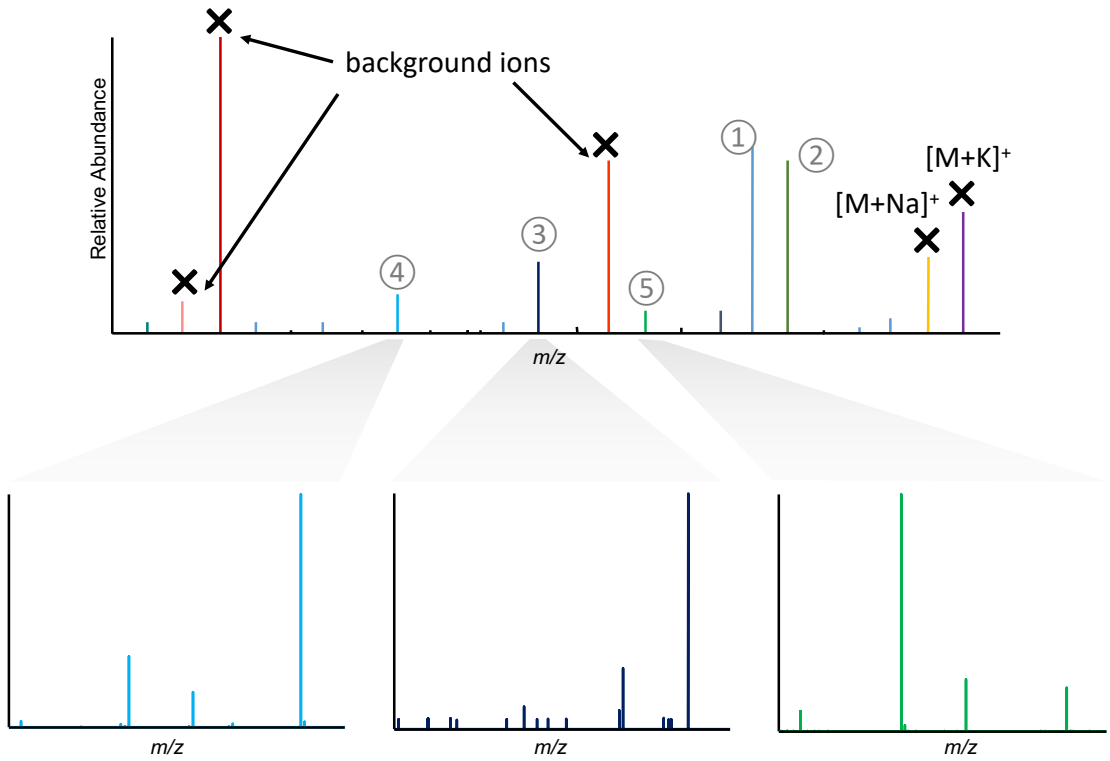

Figure 5. Precursor selection using an AcquireX DDA experimental method. Precursors are excluded based on background and matrix profiling, and precursor features are grouped to protonated compounds. Precursors labeled "1" and "2" would have been sampled in the first replicate injection. With the AcquireX workflow approach to creating the exclusion list, the DDA routine can more rapidly sample compounds of interest with lower intensities.

thermo scientific

How can I confidently identify the compounds of interest in samples?

Most small-molecule experiments are performed using methods designed to acquire data such as chromatographic retention time, precursor *m/z* values, isotopic patterns and fragmentation information. The combined data are used to differentiate between compounds and to allow automated data processing routines that sum the matched data for comparison to spectral libraries. The greater the overlap with the compound-specific attributes, the greater confidence is achieved in performing compound identification and structural annotation.

The high-quality data acquired per compound are also used to perform automated unknown unknown analysis to identify plausible structures. When exact matches to library information are not made due to limited library content, new automated search algorithms now use all experimental data to determine plausible matches. Software routines such as Thermo Scientific™ mzLogic™ Data Analysis Algorithm (included with Thermo Scientific™ Compound Discoverer™ software version 3.0 or greater and Thermo Scientific™ Mass Frontier™ software version 8.0 or greater) use both precursor and product ion information to generate a ranked list of plausible structures. The routine performs two independent searches: chemical database searching using the molecular data (precursor *m/z* values, isotopic fine structure, and distribution profiles), and similarity searching against the Thermo Scientific™ mzCloud™ mass spectral fragmentation library, which matches fragment ions from the unknown-unknown to library entries. The two independent searches result in compound lists and plausible structure rankings are based on the degree of substructure overlap. The highest ranked structures can be further evaluated using the *in silico* fragmentation analysis capability of Mass Frontier software and compared against the empirical data to strengthen confidence in plausible structure assignments.

What is important to remember when performing small-molecule studies?

Successfully identifying and characterizing compounds for small-molecule studies requires the selection of features of interest for tandem mass-spectral analysis, and the acquisition of high-quality HRAM MS and MS^n spectral data during LC-MS experiments. MS and MS^n data-acquisition methods for small-molecule compound identification have been limited by the relative abundance of compounds of interest relative to sample background matrices. A breakthrough in intelligent automation, AcquireX increases profiling efficiency by automating inclusion- and exclusion-list creation for highest quality MS^n data acquisition. AcquireX acquires these data with significantly less manual experimental setup and interpretation. Collection of highest-quality LC, MS, and MS^n data—particularly information-rich MS^n fragments—provide in-depth structural knowledge of compounds for high-confidence compound identification and structural elucidation.

How to access AcquireX intelligent data acquisition

Download Instrument Control Thermo Scientific™ Xcalibur™ 4.2 software and Tune 3.1 or newer

Find out more at **thermofisher.com/AcquireX**

ThermoFisher SCIENTIFIC