thermo scientific

THINKING BIGGER BY SEEING SMALLER

Strategies for making small-molecule unknowns known

ThermoFisher SCIENTIFIC

INTRODUCTION

Small molecules play a key role in our lives. They are a substantial part of who we are and everything that we interact with daily, including foods, drugs, and consumer products. Whether it is better understanding how our bodies interact with drugs, tracing the environmental fate of pesticides, developing new compounds, protecting brand reputations, or performing fundamental research, complete characterization and identification of small molecules is an important task.

Optimizing the small-molecule identification process is a challenge when analyzing complex samples because interferences and relatively low-abundance target compounds can prevent extraction of meaningful data. In addition, when library searching does not provide a high-confidence result, determining the identities of unknown small molecules using mass-spectral interpretation is a further challenge. Limited understanding of complex fragmentation chemistry and the difficulty of annotation, identification, and characterization of small-molecule compounds are substantial obstacles to mass-spectral structural elucidation.

Overcoming these challenges requires not only advancements in sample preparation, separation, and data-acquisition technologies, but also in computational tools to facilitate the most effective use of the high-resolution accurate-mass (HRAM) data generated by mass spectrometers.

For laboratories performing small-molecule studies using Liquid Chromatography-Mass Spectrometry (LC-MS), or that are considering adopting LC-MS, this eBook provides an overview of the trends and challenges associated with LC-MS-based small-molecule analyses, and the instrument and software solutions designed to help. For a variety of applications—including metabolomics, environmental, food safety, and clinical research, as well as forensic testing—these solutions break analytical bottlenecks to allow faster, higher-confidence profiling, screening, and quantitation of unknown small-molecule compounds.

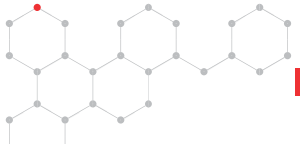


TABLE OF CONTENTS

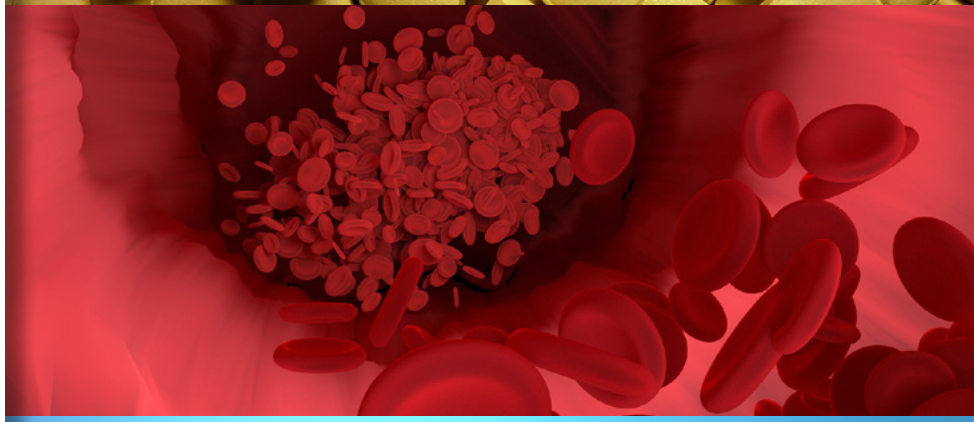
Chapter 1
APPLICATION AND INDUSTRY TRENDS

From metabolomics, drug research and development, industrial and environmental health, and food safety, to emerging drugs of abuse testing, numerous fields of work rely on successful small-molecule identification and characterization. Although traditionally isolated by their unique requirements, these industries are increasingly facing common challenges as the compounds analyzed become more similar, or even the same.

Overlapping application needs: the common ground

Common challenges point to the need for industry-wide solutions that provide comprehensive toolsets to accurately and productively profile, screen, and quantify small molecules regardless of sample type. These include:

- New approaches to identify emerging compounds of interest, often in complex sample matrices
- Improved capabilities to quantify compounds at increasingly lower limits of detection (LOD)
- Constant re-evaluation by regulatory bodies to ensure best analytical approaches
- Protecting brand and reputation with certainty in test results to keep constituents safe
- Streamlined test procedures to increase productivity and contain testing costs



INDUSTRY CHALLENGES: OVERVIEW

Metabolomics

- Potentially large and diverse sample cohorts
- Global detection and relative quantitation

Learn more in this seminar

Drug research and development

- Need for rapid, next-generation therapeutic discovery
- Exhaustive characterization of activity across multiple matrices
- Impurity, degradation, and extractables and leachables identification

Learn more in this video

Industrial and environmental health

- Extensive diversity of potential compounds requiring analysis
- Rapidly evolving regulatory guidelines

Learn more in this user meeting

Food safety testing

- Increasingly strict regulations to ensure consumer confidence and safety
- Compounds requiring characterization can include pesticides, persistent organic pollutants, veterinary drugs, mycotoxins, and others

Learn more in this eBook

Emerging Drugs of Abuse

- Continuous introduction of ‘designer drugs’ strains both law enforcement and analytical testing efforts
- Increase in ‘legal high’ use increases throughput demands for screening and novel substance identification

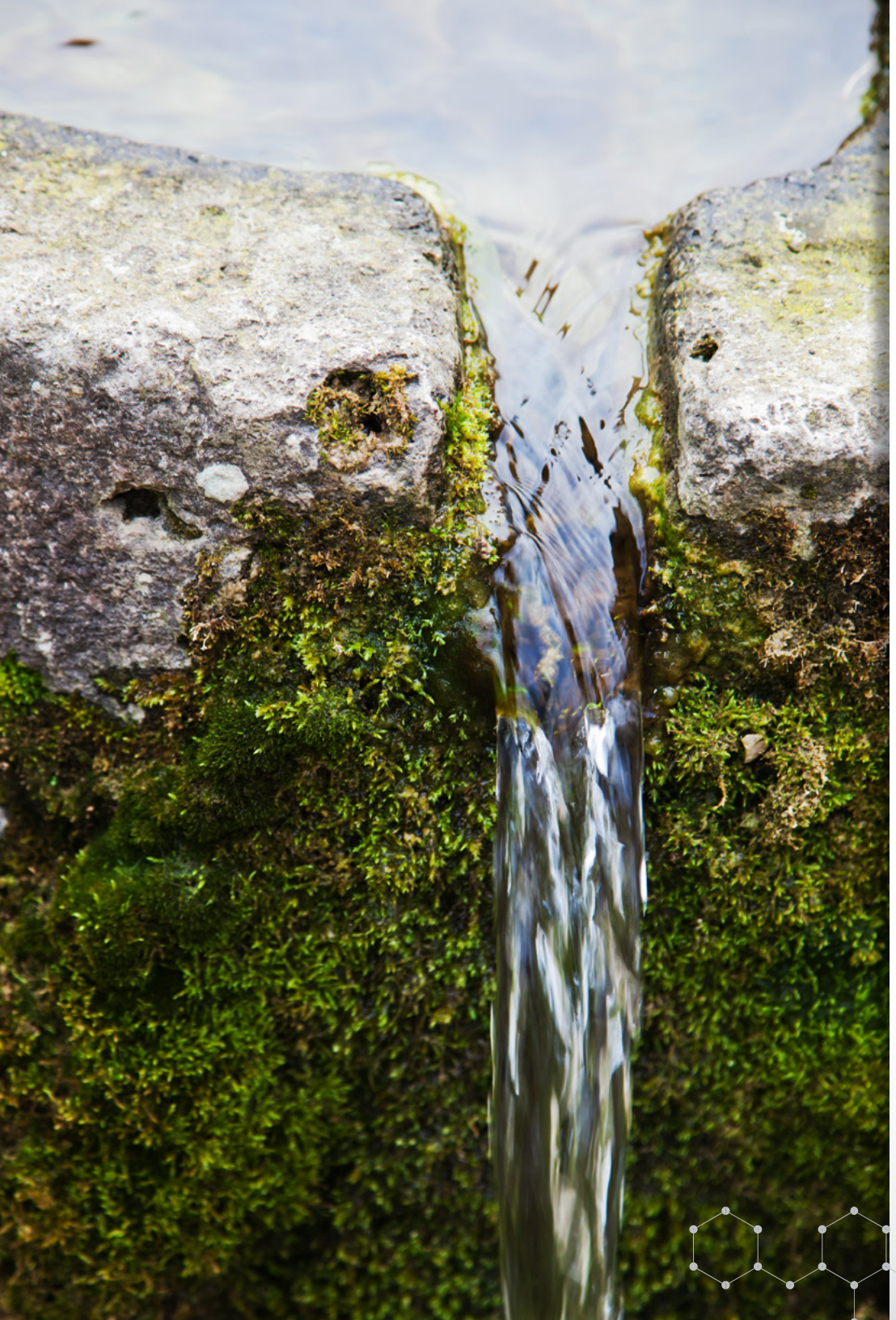
Learn more in this handbook

Chapter 2
ANALYTICAL TESTING

Small-molecule analyses strive to answer three basic questions. First, profiling asks, *"What is in the sample?"* This question can be as broad as, *"What is in the sample that we do not know about, and are those compounds harmful to the environment or human health?"* Screening asks, *"Are specific compounds in the sample?"* Targeted quantitation adds, *"If the compounds are in the samples, at what concentration levels?"*

The small-molecule compounds analyzed in samples generally fall into three categories:

- **Known-knowns** are compounds that are likely present in the sample and are known. Information about these compounds, such as their mass spectra, is in a library or database that is searched against to determine their presence.
- **Known-unknowns** are compounds present in a library or database that is used to search against, but that are not expected to be in the sample.
- **Unknown-unknowns** are compounds not expected to be present, and as true unknowns, may not be well represented within a library or database, or cannot be confidently assigned a match due to a lack of available evidence.



Profiling

The meaning of the term profiling can depend on the industry. When the term profiling is used in this eBook, it refers to comprehensive profiling of the contents of any given sample to precisely determine its contents.

Profiling bottlenecks

Profiling presents certain challenges for the analyst chartered with turning profiling data into actionable insights. Even when the potential exists for identifying previously unknown compounds such as with extractables and leachables, emerging contaminants, unexpected metabolites, degradants, designer drugs, and more, there is typically little knowledge and few resources available to quickly and confidently assign structures to many of the components.

To improve the chances of productively making confident assignments, high-quality HRAM spectral data must be generated efficiently. In addition, harnessing the power of fine isotopic fidelity, performing deep MS^n–level interrogation of each component, and ensuring high levels of mass accuracy and sufficient sensitivity, can require significant time and manual intervention.

Learn more High-resolution compound identification in metabolomics: a review of current practices

Screening

Laboratories perform untargeted screening to identify and confirm if previously profiled analytes exist in their samples, as well as to detect new compounds which may require identification and characterization. Laboratories also assess if the concentration levels of detected analytes are acceptable for human or environmental safety. Ultimately, screening experiments can inform subsequent targeted analyses.

Screening bottlenecks

As with profiling, confidence in the quality of data is of utmost importance. Sample complexity can directly impact the certainty of results, because the presence of matrix interference can give rise to false positives.

Retention time, HRAM data, and library or database searches may not be enough for unknown identification. In addition, for many unknowns, reference standards may not be readily available, leaving laboratories without a straightforward way to identify compounds with any confidence.

Learn more

High-Resolution, Accurate – Mass Orbitrap Mass Spectrometry – Definitions, Opportunities, and Advantages

Targeted quantitation

Targeted quantitation quantifies compounds that have been identified as important. The analytical tool chosen to perform quantitation assays should provide:

- **Sensitivity** to achieve the limits of detection (LODs) required for compliance testing for existing and future regulations.
- **Selectivity and specificity** to confidently identify and quantify target analytes in the presence of matrix interferences within complex samples.
- **Robustness** to minimize system downtime, increasing operational efficiency.
- **Throughput** to deliver analytical results on time.
- **Low cost per sample or test**, which includes the initial instrument investment, as well as the expense to operate and maintain it.

Learn more — When Should You Walk on the Path to More Confident Quantitation Definitions, Opportunities, and Advantages

Learn more — Ask Yourself These Three Questions When Evaluating a Triple Quadrupole Mass Spectrometer

Efficiency and confidence: goals laboratories share

With numerous samples to analyze, and in some cases sample backlogs, laboratories strive to deploy analytical workflows that efficiently provide results from complex samples without the need for extensive user expertise. This points to the need for automated, intelligent data acquisition and data processing approaches to small-molecule analyses.

Whether it is profiling, screening, or targeted quantitation, confidence in results is likewise paramount; in addition to being sure of compound assignments, analysts need precision and accuracy when quantifying targets. Decisions made based on analytical results have a profound impact, and incorrect or inaccurate results could have serious consequences. Therefore, access to the appropriate analytical tools and knowledge is key to realizing absolute productivity and confidence in results.

How can the cloud benefit **my lab**?

Cloud benefits

- Efficiency
- Scalability
- Flexibility
- Security
- Uptime
- Latest information

Download Infographic

Traditional methods

- Slow
- Time to decision
- Old information
- Not scalable
- Traceability

A problem shared is a problem half-solved

Sharing information and knowledge is extremely powerful in helping scientists rapidly answer questions, solve problems, and make decisions. As a result, laboratories are becoming increasingly connected with Cloud-based technologies that are impacting information sharing within the scientific community. Today, it is easy to access vast repositories of information, allowing large amounts of data to be processed and shared among colleagues and collaborators faster than ever before.

Considering the challenges of small-molecule characterization, and the broad chemical diversity of any potential component in a sample, the ability to access the latest information in the Cloud significantly improves the chances of making confident identifications rapidly. Data repositories are continuously updated such as is the case with mzCloud, a continually updated online mass spectral library for confident compound identification.

Instruments can also be Cloud-connected, allowing real-time online access from any location. This enables scientists to check on and diagnose system performance, to book instrument time, or to evaluate system utilization to maximize uptime and productivity, freeing up their time to focus on what's important.

Learn more Thermo Scientific™ Almanac Web-Based Monitoring and Management

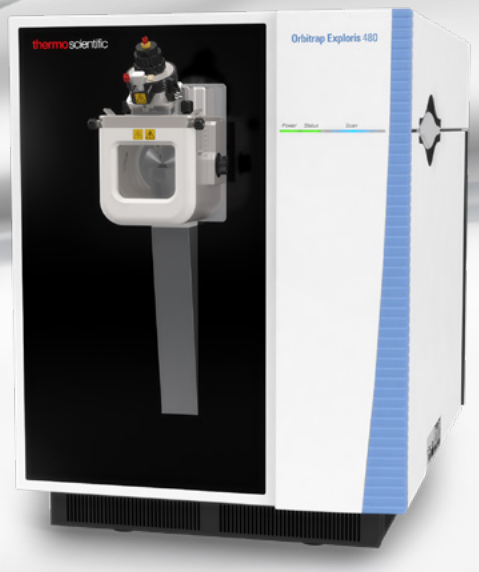
Chapter 3

SOLVING CHALLENGES: INNOVATIVE, TIME-SAVIING TECHNOLOGIES

Acquire insightful data, not just more data

Data has value because it supports decision-making, and it is important that decision-makers are confident in the data behind their decisions. Acquiring more data isn't necessarily a challenge, but acquiring data of the caliber needed to provide deep high-confidence insight is.

Capturing extensive high-quality fragmentation information for all compounds in a sample has traditionally required significant manual intervention and interpretation, especially for complex samples containing many thousands of potential known-known, known-unknown, and unknown-unknown compounds.



Highest-quality data for high-confidence identifications

Thermo Scientific™ Orbitrap™ mass spectrometers generate HRAM data that enables both untargeted and targeted screening. When operated in screening mode, the mass spectrometer analyzes retention time, isotopes, fragment ions, and exact mass. When applied to a library search or used for structural elucidation, HRAM mass spectral data with fine isotopic detail provides high-confidence identifications.

Orbitrap mass spectrometers are known for three major benefits:

- Easily distinguish between analytes of interest and interferences due to unparalleled mass resolution up to 480,000 (FWHM) at *m/z* 200
- Eliminate false positives with sub 1-ppm mass accuracy when using the optional EASY-IC ion source
- Enable interrogation of previous data for new analytes because multistage fragmentation data is collected for every component in a sample

Learn more — Which Orbitrap is right for me?

Learn more — What makes Orbitrap mass spectrometry the first choice for both routine and research applications?

Confident quantitation

Comprehensive characterization and screening of unknown substances should narrow the list of targets that need to be quantified. Moreover, qualitative screening should also specify the levels of detection needed for quantitation.

Various quantitative techniques are used including:

- LC
- LC coupled with either single quadrupole (SQ) MS, triple quadrupole (QQQ) MS, or Orbitrap mass analyzer-based MS
- LC immunoassays
- Ligand binding
- Gas Chromatography (GC), and Ion Chromatography (IC) coupled to both quadrupole and high-resolution MS

Of these techniques, LC-MS is widely used by laboratories requiring the highest sensitivity and selectivity, especially when analyzing complex matrices. LC-QQQ is known to provide ultimate sensitivity for target analytes in complex matrices. Selected reaction monitoring (SRM) has enhanced the selectivity of LC-QQQ approaches to quantitation.

Tomorrow's quantitation

Though LC-QQQ has been widely used, Orbitrap mass spectrometer HRAM analyses have gained wide popularity for their ability to provide simultaneous screening, identification, and quantitation in a single analytical run. Tremendous resolving power permits separation of ions of interest from matrix ions, enabling confident, accurate quantification of compounds in complex samples. Unlike QQQ instruments, Orbitrap mass spectrometers require minimal tuning and compound optimization due to their universal detection capabilities.

Regardless of the type of mass spectrometer chosen, Thermo Scientific™ TraceFinder™ software can be used for the high-throughput analysis of data generated by both LC- and GC-QQQ and LC-HRAM techniques. Using the same software platform, laboratories can seamlessly move analyses from method development to routine quantitation. TraceFinder software also provides extensive data visualization and flexible reporting templates to meet varied requirements.

Thermo Scientific™ Chromeleon™ Chromatography Data System (CDS) software enables laboratories to meet compliance requirements described by US FDA 21 CFR Part 11, and EU Annex 11. Additionally, Chromeleon CDS software controls both chromatography and mass spectrometry instruments in an enterprise environment, with control for multi-vendor instrumentation, further simplifying laboratory operations.

Learn more

Life Begins At 40 – A Brief History of LC-MS/MS

Learn more

Triple Quadrupole (QqQ) and Orbitrap-based high-resolution spectrometry – choose the right tools for the trade

Learn more

Orbitrap quantitation: lab of the future

Learn more

Using Chromeleon Chromatography Data System to Comply with 21 CFR Part 11

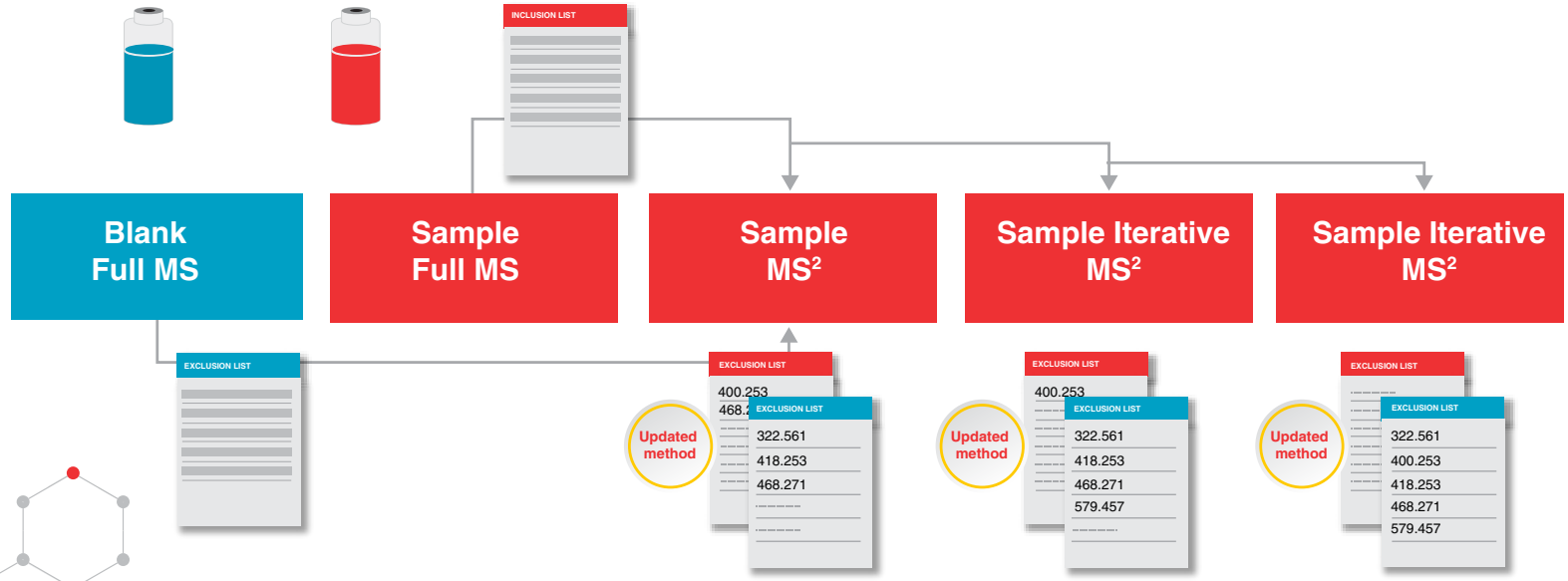
A better way: acquiring the needed data

Today, innovative instruments and data acquisition methods allow analysts to use automated LC-MS workflows to acquire extensive high-quality fragmentation data for comprehensive sample profiling. A breakthrough in intelligent automation, Thermo Scientific™ AcquireX intelligent data acquisition for Thermo Scientific™ Orbitrap Tribrid™ mass spectrometers and specifically the Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer dedicated for small molecule analysis, increases profiling efficiency by automating exclusion of background and matrix components and inclusion of components of interest. AcquireX intelligent data acquisition provides highest-quality MS^n fragment data, while acquiring these data with significantly less time-consuming manual experimental setup and interpretation.

The AcquireX intelligent data acquisition workflow begins by profiling a blank sample to identify matrix components for exclusion and adds these to an exclusion list. Next, the workflow profiles the sample in full scan mode to identify the components for inclusion, automatically creates an inclusion list, and uses the lists to automatically update the acquisition method. Dynamic MS^n fragment data acquisition is automatically triggered for every compound on the inclusion list. Once data is collected for a triggered component, that component is automatically transferred to the exclusion list in case another LC-MS run is needed to fully profile the sample.

Learn more AcquireX Intelligent Data Acquisition Technology for Orbitrap Tribrid Mass Spectrometers

Watch Transform Small Molecule Characterization and Identification in Drug Development



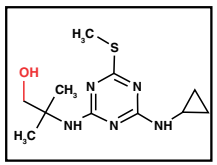
Automatically updated run-to-run inclusion/exclusion lists

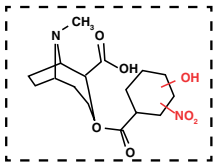
Creating and accessing knowledge

One of the challenges in creating LC-MS-based spectral libraries that can be reliably searched against is that, unlike consistent GC-MS ionization and fragmentation, there can be large differences in LC-MS ionization and fragmentation results between instruments and methods. As LC-MS instrumentation has improved to deliver highly accurate and consistent data from instrument to instrument, the ability to create LC-MS spectral libraries containing robust, reliable data to search against to make confident assignments is now a reality.

Using mzCloud™, the world's largest and most curated mass spectral fragmentation library, analysts can access high-quality spectral content containing rich metadata, spanning multiple, diverse compound classes. The contents of mzCloud are added to constantly. There are many benefits to accessing libraries contained in the Cloud: access to the most up-to-date information available, decreased infrastructure and processing costs, and—due to the extensive nature of the data—faster access to answers.

However, sometimes it isn't feasible to use the Cloud because of lack of Internet connectivity or security restrictions. Laboratories without Cloud access can obtain up to MS^2-level mzCloud offline library information using Thermo Scientific™ mzVault™ software. The MS and MS^2 information in mzCloud is added to mzVault software twice annually. This enables those who cannot access the Cloud to use the library's rich content for compound characterization and identification.





$C_6H_5N_3O_4$

192.0757

Example	*Identification confidence*	*Minimum data requirements*
	Level 1: Confirmed structure by reference standard	**MS, MS^2, RT, Reference Std.**
	Level 2: Confirmed structure a) by library spectrum b) by diagnostic evidence	**MS, MS^2, Library MS^2** **MS, MS^2, Exp. data**
	Level 3: Tentative candidate(s) structure, substituent, class	**MS, MS^2, Exp. data**
	Level 4: Unequivocal molecular formula	**MS, isotope/adduct**
	Level 5: Exact mass of interest	**MS**

Learn more

Delivering confidence for small-molecule identification

Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, 48, (4), 2097-2098.

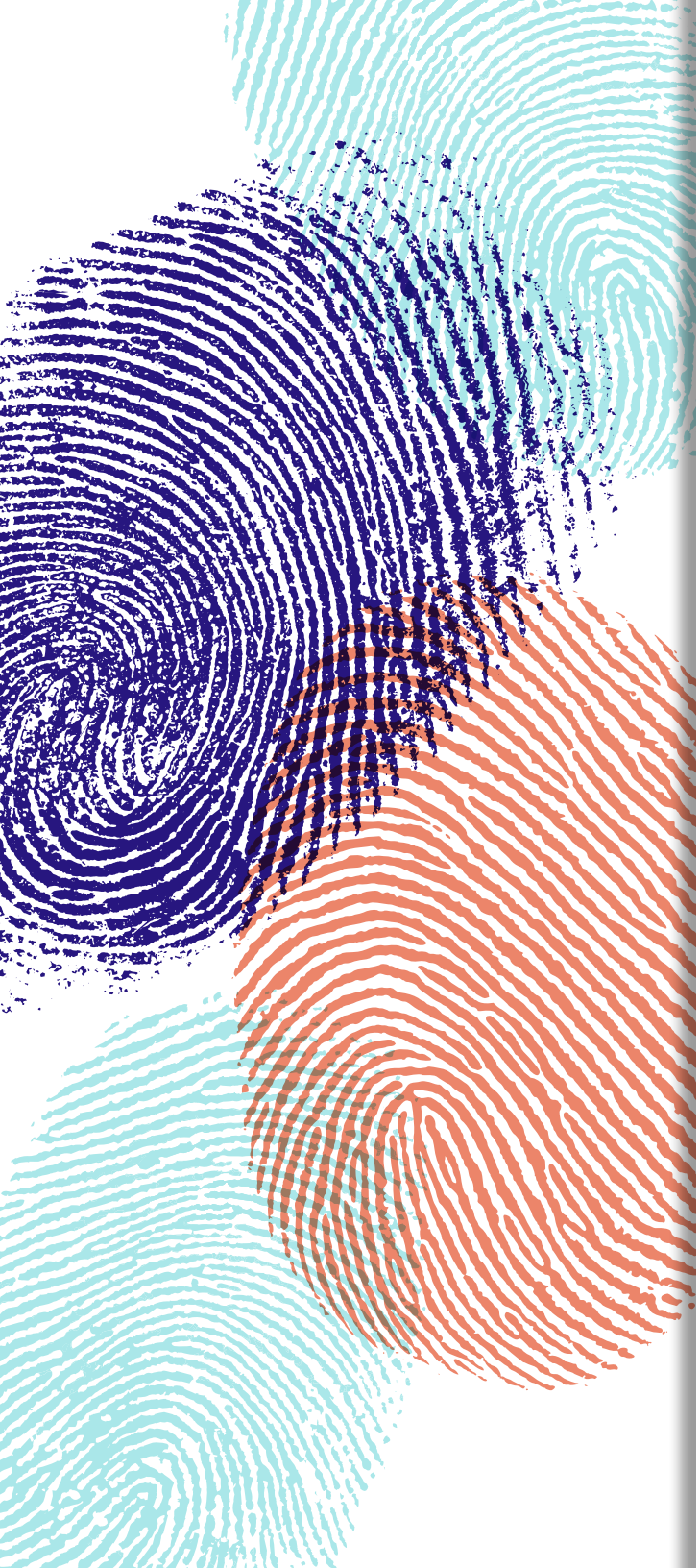
Transforming data into insight

Focusing on known-knowns and known-unknowns first, the analyst can utilize any MS^n information acquired to search against mass spectral libraries such as mzCloud.

Interrogating the extensive data in mzCloud is streamlined and can be fully automated using Thermo Scientific™ Compound Discoverer™ software. The software allows users to tailor data processing based upon their requirements. For example, analysts can choose to perform statistical analyses or data visualization, or to search against mzCloud and other library and database resources, either online or offline. Users can also rapidly view their data using smart visualization tools to boost certainty in component assignments.

Learn more

Small Molecule Characterization and Identification with Compound Discoverer



Making unknowns known

Even when sufficient experimental mass spectral data is available, it can often contain information that cannot be confidently assigned a match with any of the contents of the library searched against. HRAM MS data permits high-confidence determination of elemental composition using fine isotopic structures. When fine isotopic information is used to generate elemental compositions that are searched against chemical structural databases such as the ChemSpider database of more than 67 million structures, many hundreds to thousands of potential structures may be returned.

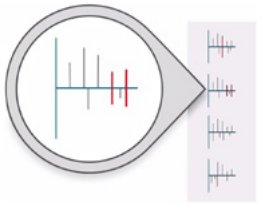
With broad compound diversity and extensive MS^2 and MS^n-level information based on real data and fragments, the fragmentation information within mzCloud can be used to effectively narrow the list of potential structures. Using mzLogic to combine spectral library similarity with chemical database searching provides a completely new way to leverage the real-world fragmentation information to propose the best candidates for true unknowns.

Potential structures are ranked in order of confidence by searching the spectral fragmentation information in mzCloud against the experimental fragmentation information to determine substructure overlap. The rankings results window presents the substructure information and overlap, making it quick and easy to review the available evidence for structural confirmation.

Thus, mzCloud can reduce many thousands of potential candidates to a short list of putative structures in a very short amount of time, eliminating key bottlenecks in the review and confident assignment of structures.

Learn more Accelerate your small molecule unknown identifications

Watch mzLogic: Maximize the use of your MS^n data for unknown identification



mzLogic algorithm

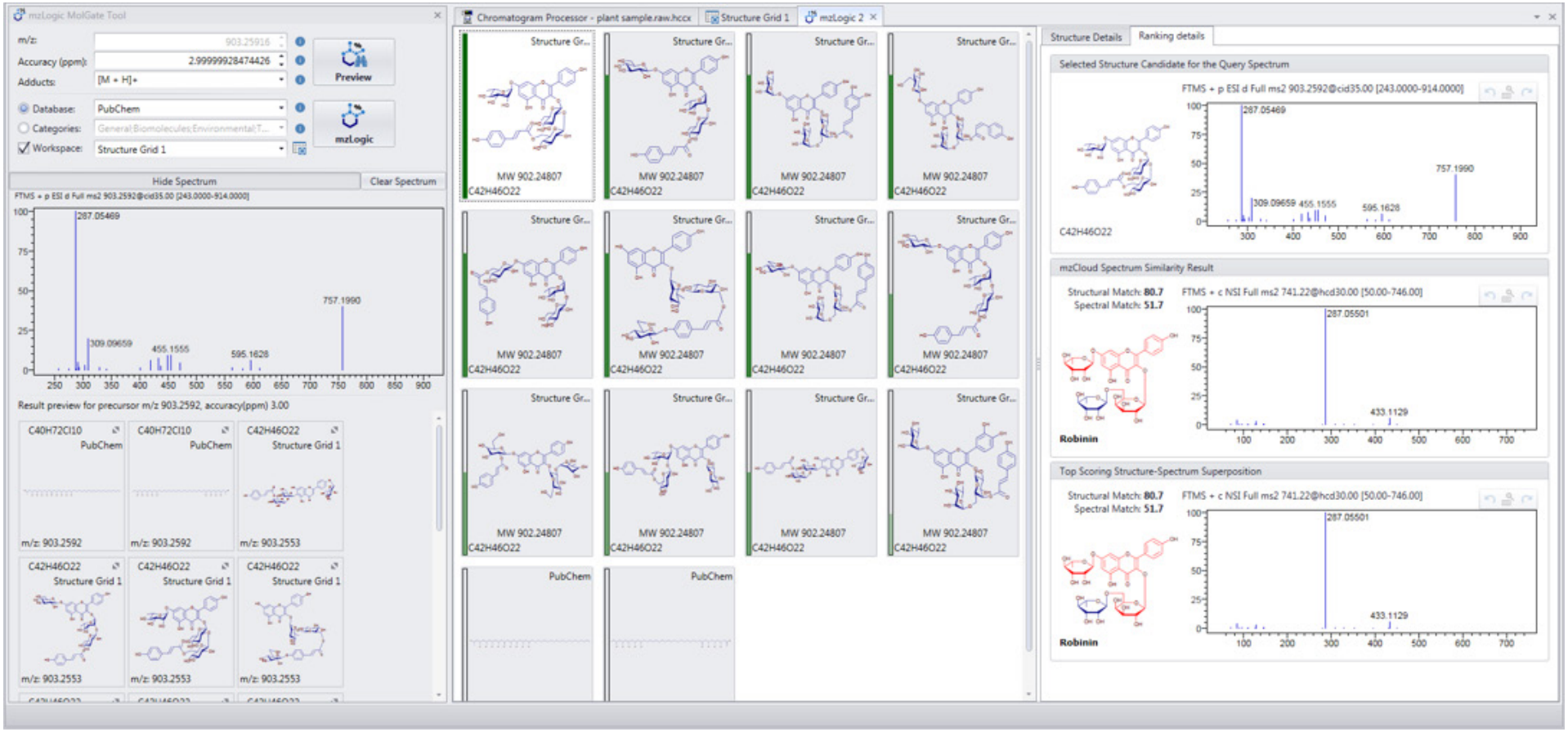
What if it's still unknown?

There can be occasions when there are no potential structural matches in any of the libraries searched. If this occurs, mzCloud can still supply enough sub-structural information to aid in putative structural assignment.

Thermo Scientific™ Mass Frontier™ Spectral Interpretation Software permits analysts to search acquired MS^2 and MS^n data against the comprehensive MS^2 and MS^n data in mzCloud. The software provides sub-structural information and spectral tree coverage for the matches generated. These capabilities are especially useful for emerging contaminants and designer drug applications, where there are common sub-structures because the unknown compound may only have a small structural difference from the known compound.

Learn more

Next-generation software tools for the identification of compounds



Customizing content, sharing knowledge

Using Mass Frontier software and Compound Discoverer software, analysts can create new libraries of proprietary compounds and identified unknown-unknowns, or add to existing mzVault libraries, which can be shared to allow others to access this knowledge.

The ability to manage and curate mzVault libraries is provided by the Mass Frontier Curator feature that allows users to access the tools used to curate mzCloud. Spectra can be mass-calibrated, with full structural fragmentation and metadata included. In addition, customized libraries can be hosted on a network for searching against, enabling organizations to develop and share combined knowledge. Once compounds are identified and added to mzVault libraries, the information can be used for subsequent targeted screening and quantitation experiments, for example with TraceFinder software which also uses mzVault.

Chapter 4

CONCLUSION

PUTTING THE PIECES TOGETHER: REMOVING SMALL-MOLECULE ANALYSIS BOTTLENECKS IN A BROAD RANGE OF APPLICATIONS

The need for certainty in test results requires end-to-end solutions that improve the analyst's ability to identify new or low-level compounds in complex sample matrices rapidly and without adding to testing costs.

Thermo Fisher Scientific is committed to providing a complete, end-to-end instrument and software solution for profiling, identifying, and quantifying unknown small molecules in complex sample matrices. The solution begins with Orbitrap mass-spectrometer-based LC/MS sample analysis to acquire rich full-scan HRAM MS data. Recent advances like the AcquireX intelligent data-acquisition workflow enable laboratories to capture more actionable—rather than just more—HRAM data.

Depending on laboratory needs, the HRAM data is processed using powerful Compound Discoverer software and Mass Frontier software with the ability to access the mzCloud online mass spectral library (or offline mzVault library) to confidently confirm the identities of known-knowns, known-unknowns, and unknown-unknowns. In addition to enabling full use of HRAM data to streamline unknown identification, Compound Discoverer software provides tools to calculate statistical differences and highlight trends between sample sets.

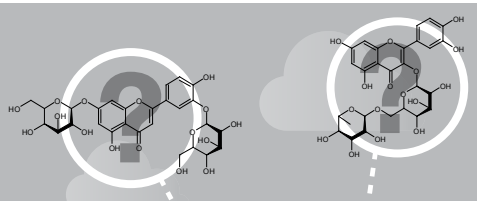
The ability to create spectral libraries of compounds allows laboratories to develop their own proprietary libraries of compounds, or to share libraries of newly identified compounds with colleagues.

From profiling and screening to quantitation, small-molecule characterization is vital to understanding disease drivers, emerging contaminants, and drugs of abuse. Regardless of application or market, Thermo Scientific small-molecule analysis solutions break analytical bottlenecks to empower high-confidence decision-making where outcomes are directly linked to the future of human health and the state of our environment.

Small molecule characterization and identification **clouding your decision making?**

Linking the cloud with hardware and software

Solving tomorrow's problems, today

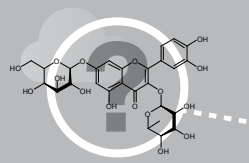


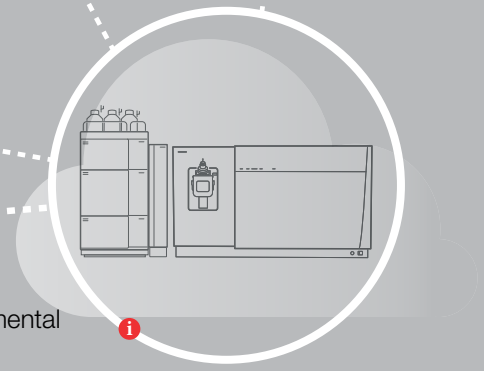
The challenge

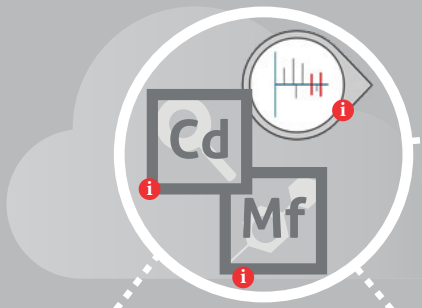
- Confident identification of compounds in challenging matrices
- Complete unknowns

Remote monitoring

- Through the cloud, monitor your instruments, view queues and acquisition information
- Powerful logbook capabilities and service tools
- Customizable utilization report to track instrument usage

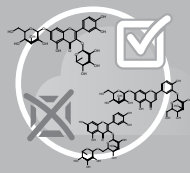






Capturing the data

- Excellent spectral clarity: separate elemental isotopes with high resolving power
- Automate complete MS^n sample profiling with Thermo Scientific™ AcquireX intelligent data acquisition



Enabling software

- Accessing the latest, comprehensive information to make fast, confident identifications based upon real data

When the cloud is a dream

- When internet connectivity is not available, utilize the power of mzCloud's MS/MS level information in an offline library

Download Infographic

Depth, breadth and completely up-to-date

- A comprehensive high-resolution spectral library, consistently updated with the broadest diversity of compound classes
- Access knowledge with speed through enabling software that can quickly and confidently identify known compounds
- From the extensive fragmentation information covering a range of collision energies and techniques, you can be sure of sub-structural matches for even the most diverse of compounds, from new chemical entities to designer drugs
- Direct access to tomorrow's knowledge, today

The right answer, confidently

- From the known components, through to the new and novel, accessing millions of fragmentation spectra allows confident assignment of component structures
- Even when something is new, identify sub-structural information and unlock knowledge

ADDITIONAL RESOURCES

Addressing the "Identity Crisis" in Small Molecule Compound Identification

This webinar describes a fundamental new approach to untargeted small-molecule analysis involving optimized mass spectrometers, powerful new data acquisition strategies, and an arsenal of new software tools to translate high quality Orbitrap mass spectrometer produced spectra into more, confidently-assigned, small molecule structures.

Watch webinar

AcquireX: Collect More Meaningful Data

Transform small molecule identification and characterization with AcquireX intelligent acquisition to generate more meaningful fragmentation spectra.

View video

Accessing the Value of Your Data: Software for the Masses

Making sure that you can access the knowledge, and therefore value, within your data is critical. Discover the range of solutions designed to meet your specific small-molecule characterization needs.

Learn more

thermo scientific

thermofisher.com/orbitrap

XX65573EN-0719S

ThermoFisher SCIENTIFIC