White Paper

Unified Laboratory Intelligence for Impurity Resolution Management

Identify and characterize pharmaceutical impurities in less time with greater certainty.

Authors: Ryan Sasaki and Bruce Pharr

ACD/Labs

Impurity Identification and Characterization

The identification and characterization of impurities in pharmaceutical drugs is critically important to ensure that their presence will not evoke any form of adverse response, either pharmacological or toxicological, in a patient taking the medication. Furthermore, attention must be made to the impact impurities can have on the production process and formulation. As a result, it is a "frequent, highly specialized, and labor-intensive undertaking."[i]

Regulatory bodies such as the U.S. Food and Drug Administration (FDA) and the International Congress on Harmonization (ICH) have established rigorous guidelines for the identification of extraneous compounds in pharmaceutical agents throughout the development process and post-marketing. Generally, for drugs dosed at ≤2g/day, impurities present at ≥0.10% require isolation and structural characterization. For drugs dosed at >2g/day, the threshold for isolation and identification is lower at 0.05%.[ii]

A 2011 FDA letter to all U.S. drug manufacturers expressed concern about the rising incidence of drug shortages. The letter emphasized that, "the number of drug shortages annually has tripled from 61 in 2005 to 178 in 2010," and that, "54% of drug shortages in 2010 were due to quality issues."[iii] Impurity identification and characterization is usually a time-critical undertaking often linked to project milestones, such as the first human clinical trials or production schedules. In addition, groups tasked with the responsibility of isolating and identifying impurities typically require a capital investment of millions of dollars in hardware, as well as individuals highly skilled in the utilization of the equipment and the interpretation of the data generated.[iv]

In the pharmaceutical industry, time absolutely equals money; and maximizing the effective patent life of a marketed drug is one of the main priorities. According to a March 2010 article in Nature, the estimated average cost for pharmaceutical companies to bring a new chemical entity (NCE) to market is approximately $1.8 billion, and requires about 13.5 years on average.[v] A report published the same year by McKinsey & Company concludes that accelerating the time to market by one month equates to a $10.5 million improvement in net present value.[vi]

"While the overall elapsed time for the isolation and identification of a new impurity can vary widely, the average time required seems to be about two weeks."

"Identifying and characterizing an impurity in a couple of days less time can literally equate to a gain of a million dollars in net present value."

While the overall elapsed time for the isolation and identification of a new impurity can vary widely, the average time required seems to be about two weeks.[vii] Identifying and characterizing an impurity in a couple of days less time can *literally* equate to a gain of a million dollars in net present value. Being able to do this repeatedly as impurities are encountered throughout the development and acceptance of a drug substance can shave weeks, or even months, off the time

to market; increasing returns by millions of dollars.

Improving the overall productivity of drug development and acceptance requires a combination of strategies, but even moderate improvements can substantially increase returns.[viii] Knowledge of the synthetic route and the experimental details of an impurity profile can prove invaluable in speeding the isolation and subsequent identification of an impurity, and *in silico* screening may be employed to assess the potential impact various impurities may have on the efficacy of a pharmaceutical agent to prioritize further testing when toxicology data is limited or lacking. In some instances, an *in silico* system alone can determine the molecular structure of a complex unknown, like a natural product.[ix] Collaboration can enhance scientific productivity when collaborators bring special expertise and knowledge crucial to the research outcome, and in situations where there is a joint use of specialized equipment.[x] In order for each of these strategies to help deliver a drug to market in less time—and, thereby, prolong its effective market life—scientists must be able to obtain the highest quality data possible in the shortest amount of time.[xi]

Unified Laboratory Intelligence

Unified Laboratory Intelligence (ULI) is a category of R&D informatics that combines software, algorithmic tools, and databases to form a scalable platform for the collection and unification of chemical, structural, and analytical data. ULI accumulates information and knowledge from successive projects across different chemistry disciplines to create a one-to-many chemical intelligence-from-information 'live' cycle. With ULI, scientists can easily search and quickly retrieve 'live' data to gain and apply intelligence and insight that improves decision-making.[xii]

Scientists can gain and apply scientific insights through the management and analysis of new and existing data because ULI meets several requirements.

· Data are digital

· Data are structured

· Data are standardized in terms of metadata[xiii]

Data cultures in life science are very heterogeneous, and no single approach can suit the needs of everyone. The most successful strategies are those that address needs in the context of sub-disciplines.[xiv] Impurity Resolution Management (IRM) is an essential aspect of drug development and acceptance where Unified Laboratory Intelligence can be successfully applied to improve decision-making, reduce time, and decrease cost.

Impurity Resolution Management (IRM)

Impurity Resolution Management (IRM) is a specific application of Unified Laboratory Intelligence (ULI) to identify and characterize impurities for resolution and reporting during development and acceptance of a new drug substance. Specifically, FDA and ICH guidelines state that, "The specification for a new drug substance should include a list of impurities."[xv] Furthermore, "the applicant should summarize the laboratory studies conducted to detect impurities in the new drug substance," and "a registration application should include documented evidence that the analytical procedures are validated and suitable for the detection and quantification of impurities."[xvi]

Complying with the guidelines requires the collaboration of three drug development chemistry functions in addition to regulatory bodies, stability and forced degradation groups, and toxicology groups:

- Process chemistry–scaling up the development of an optimal synthetic route that maximizes yields, is cost-effective, safe, environmentally aware, and minimizes the presence of impurities for the drug substance from laboratory scale to manufacturing scale.
- Analytical method development–development of specifications and optimization of analytical method (or series of analytical methods) for the identification and isolation of the drug substance and related impurities.
- Structure characterization, identification, and elucidation–identification and characterization of impurities throughout the development and acceptance cycle of the drug substance.

Figure 1 illustrates how structured data from the three chemistry functions are continually collected from successive projects and unified to create a one-to-many chemical information-to-intelligence 'live' cycle. Scientists can easily search and quickly retrieve relevant 'live' chemical and analytical data for effective cross-functional collaboration, and to gain and apply intelligence and insight for Impurity Resolution Management (IRM).

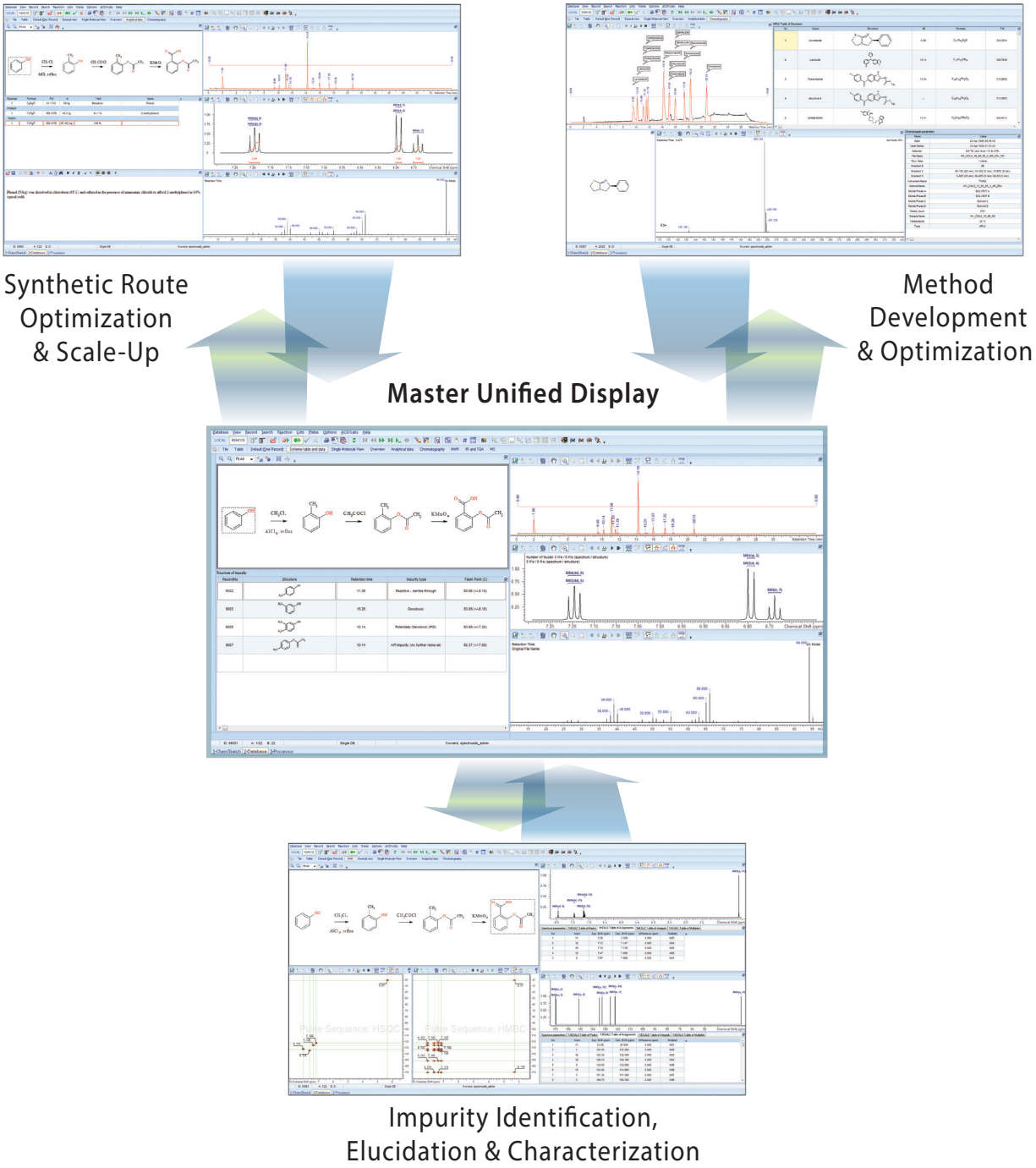


Figure 2 illustrates how the data collected and unified as 'live' data can be visualized as needed by the three chemistry functions—process chemistry, analytical method development, and structure characterization, identification, and elucidation. The ability to visualize structural, analytical, and chemical data in different contexts enables cross-functional scientific collaboration to gain and apply intelligence and insight for Impurity Resolution Management (IRM).

The Nature of Scientific Collaboration

Effective collaboration is essential to cross-functional success. Say the word 'collaboration' in a business setting, and most people will envision a diverse group engaged in a face-to-face or electronic conference actively discussing issues in order to resolve them. However, this is not indicative of most scientific collaboration.

Most scientists are naturally inclined to begin the search for a solution individually, rather than collectively. A scientist more often follows a process such as this:

- Individually investigates relevant data and publications
- Develops a hypothesis
- Discusses the hypothesis with a few trusted peers
- If necessary, modifies the hypothesis and discusses it again with peers
- When satisfied with the hypothesis, presents it to a larger group
- Tests the hypothesis to determine its validity

IRM enables specific functional groups—process chemistry, analytical method development, and structure characterization, identification, and elucidation—to visualize 'live' data in *all* the ways required to enable effective scientific collaboration, including:

- A master display with multiple functional views
- Quick access to any specific functional view and easy drill-down to detailed data
- The ability to arrange more than one functional view for comparison

Comprehensive Data Access and Visualization

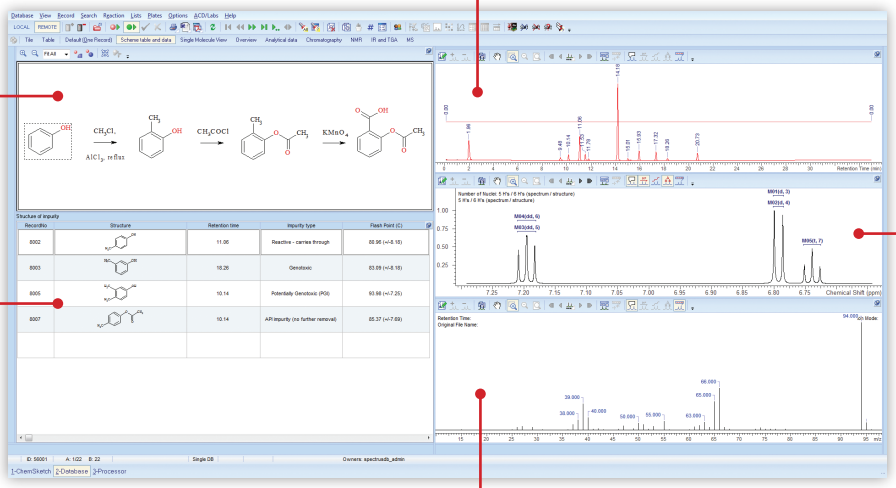
Collecting and unifying 'live' chemical, structural, and analytical data for quick and easy search, retrieval, and visualization is the key to IRM; accelerating time to market and increasing total returns.

Providing analytical content (what) with chemical context (why) in a drug development environment serves as the foundation for creating intelligence-from-information which is the goal of Unified Laboratory Intelligence. In order to create intelligence-from-information the data must be 'live', meaning it can be easily searched on, quickly retrieved, compared with other relevant 'live' data, and efficiently re-applied.

This is the starting point that enables all of the scientists involved in IRM to access and visualize all of the relevant 'live' data in complete context. While each scientific discipline has specific expertise, the ability to quickly and easily access and view all functional data in context is essential to effective cross-functional collaboration. Furthermore, the 'live' nature of the data continually updates functional views to reflect the most recent and relevant knowledge so that each discipline is working efficiently with correct, applicable data.

A high-level view of the synthetic route and reaction scheme developed by the process chemist.

A master composite chromatogram showing the composition of the drug substance and all impurities.



A table of all identified impurities with notes and comments within a specific step in the process.

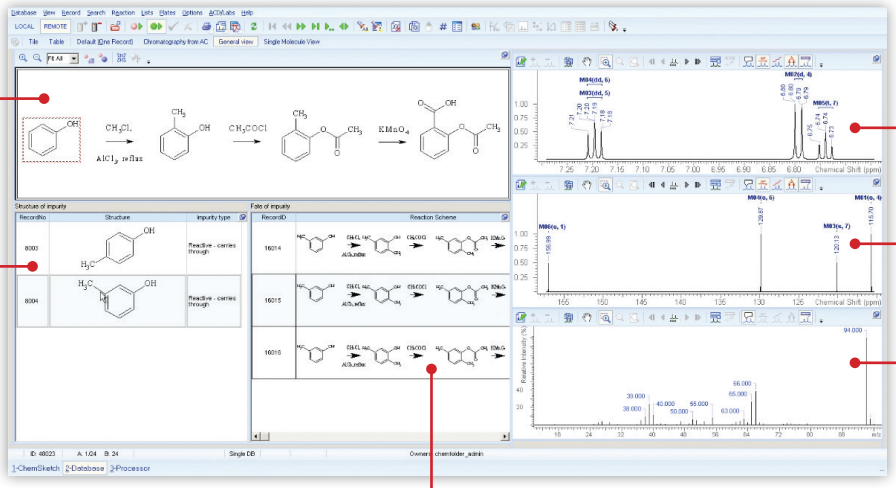
Mass spectrum for the highlighted compound

NMR spectrum for the highlighted compound

Figure 3 illustrates a master view of all the 'live' data for IRM.

The first structure in the synthetic route and reaction scheme is selected.

All analytical data associated with the selected impurity is shown.



Impurities identified in the first step of the reaction are listed.

The synthetic fate of the selected impurity from the bottom left view is shown.

Figure 4 illustrates how specific information in one view can be isolated and magnified for more detailed analysis.

Optimizing Analytical Methods

Modern drug development delivers a broad array of analytical tools and techniques for the chemist to analyze and specify drug substance, including the impurity identification and characterization required by the FDA and other regulatory organizations. This includes chromatography, mass spectrometry, optical spectroscopy, and nuclear magnetic resonance (NMR) spectroscopy. While the most expedient course of action might be to employ all available analytical tools and techniques, this approach is prohibitively expensive and time-consuming. Rather, the key to reducing the time and decreasing the cost of analytical chemistry for IRM is optimizing analytical method development.

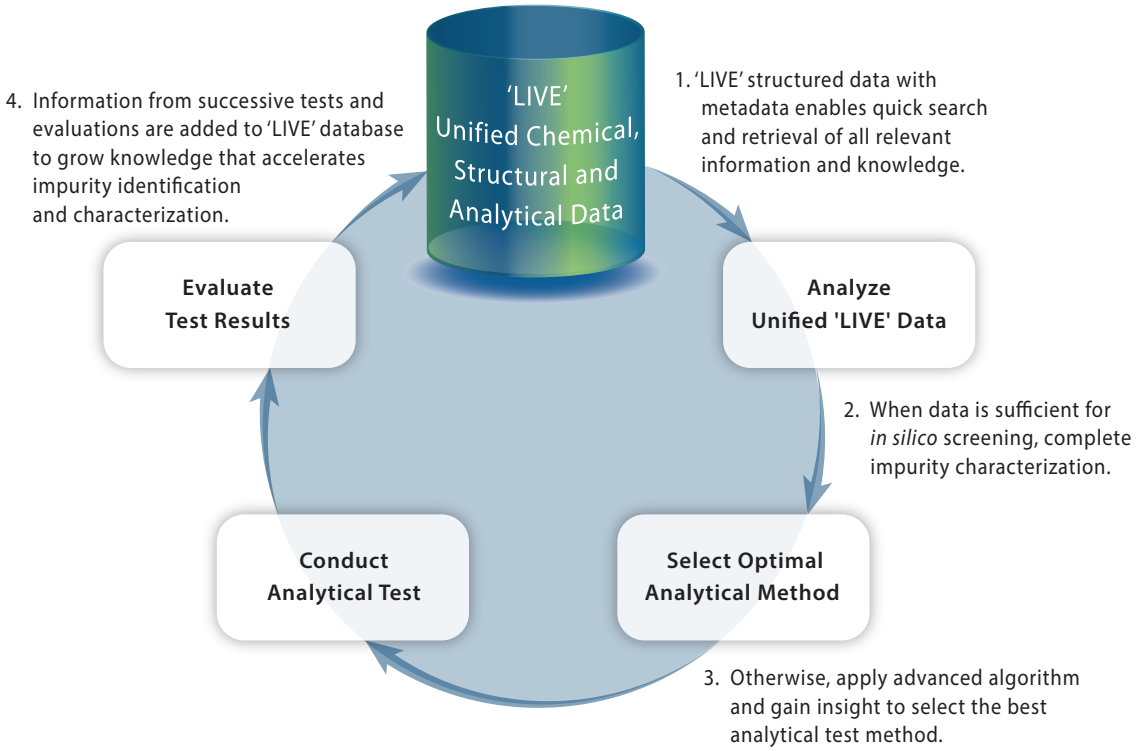


Figure 5 illustrates how the collection and unification of 'live' structural, chemical, and analytical data throughout the development process provides a rich repository of information and knowledge that can be applied to data-driven method development. The application of specialized software and advanced algorithmic tools to this 'live' data enables *in silico* analysis and evaluation for optimal method design and development.

Software-Aided Decision-Making in Method Development

An article in the February-March 2013 issue of *Chromatography Today* describes how software systems designed for the creation of chromatographic knowledge and preservation of intelligence, as an extension of everyday scientific activities, help to streamline method development, method selection, and data interpretation. The article includes examples from different scientists and organizations.[xx]

For over a decade, Rudy Sneyers, a senior scientist in the small molecule method development department at Janssen Pharmaceuticals, has made a concerted effort to improve the return on investment (ROI) on ultra high performance liquid chromatography (UHPLC) method development systems. Through a combination of hardware improvements, the ability to learn from experimentation and re-use intelligence, and the appropriate application of commercial software for organizing, visualizing, and tracking data, Sneyers' lab realized an 80% reduction in development time, a 95% reduction in cost, and extensive improvements in the quality of methods over traditional trial and error methodology.

At Eli Lilly, commercial software enables scientists to set up a method development strategy. Once the method development strategy has been decided, the software interacts directly with an analytical instrument to acquire the data and import it back into the project. According to John Stafford, the cycle time for data analysis has been reduced from more than a week to within a day.

Accelerating Impurity Structure Verification and Elucidation

The verification or elucidation of an impurity structure can be facilitated by the retrieval of reference data in combination with logical-combinatorial processing of data. An increase in computing power, databases, and software supporting the processes accompanying structure verification and elucidation has enabled scientists to increase their throughput and to solve complex problems that have previously been daunting. Computer assisted structure elucidation (CASE) has proved itself capable of solving structural problems that were previously solved by skilled spectroscopists, as well as providing a significant acceleration of the structure elucidation process in the case of more challenging problems. Modern systems are capable of not only solving common problems more quickly, but also solving problems that cannot be solved by a human.[xvii]

Computer Assisted Structure Elucidation for Multiple Impurities

In the early 21st century, computer assisted structure elucidation (CASE) first surpassed human capabilities, when the structure of an unknown alkaloid, quindolinocryptotackieine, was elucidated. NMR data were collected starting in 1991, but a solution remained elusive for a decade, in part due to extensive spectral overlap in both the ^{1}H and ^{13}C spectra. Manual analysis did not present a conclusive structure while computer-assisted elucidation produced a series of conceivable structures. A single structure consistent with all of the spectral data was selected from the list of computer-generated structures. According to volume 53 of *Progress in Nuclear Magnetic Resonance Spectroscopy*, this appears to be the first case of the application of an *in silico* system "to determine the molecular structure of a complex natural product which was not amenable to identification by competent spectroscopists."[xviii]

An article in the peer-reviewed research journal Analytical Chemistry by Anna Codina, Robert W. Ryan, Richard Joyce, and Don S. Richards of Pfizer Global Research and Development describes how quality data can enable multiple impurities structure determination semi-automatically by using a CASE approach. Commercial software that allows rapid peak picking and molecular formula (MF) determination combined with CASE significantly improves the speed of the elucidation of structures. The article concludes that, "The computational approach is not only strikingly fast, enabling the elucidation of the structures from raw MS and NMR data in minutes, but also is very reliable giving unique structures for each of the impurities with high confidence."

Dynamically Uncover Comprehensive Analytical Information

Once an appropriate analytical method is identified, applicable tests can be conducted. A fundamental aspect of ULI is its ability to collect raw and processed data from different instruments within and across laboratories, and to convert heterogeneous data to homogeneous structured data with metadata for storage, retrieval, and visualization.[xxi]

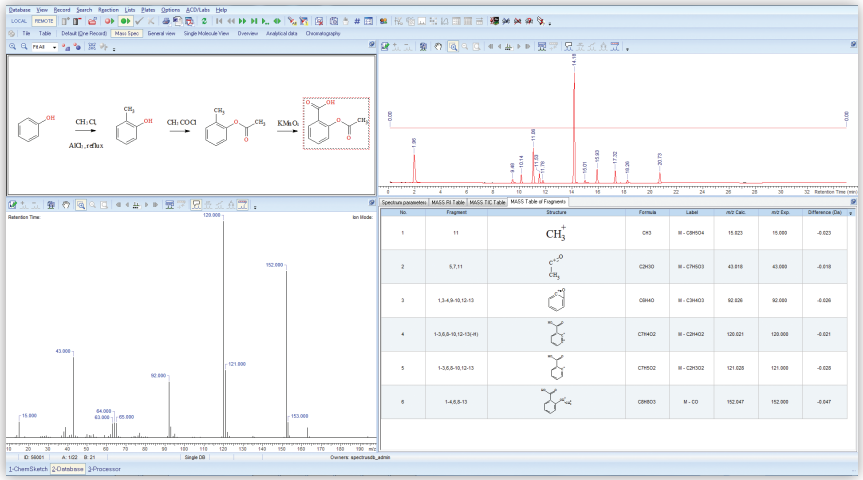


Figure 6 Shows an interactive view of 'live' chromatographic data. Moving a cursor over the various peaks reveals key data in the bottom left window and the associated structure in the reaction scheme in the top left window. This interactive view of 'live' structural and analytical data is an important feature of IRM, not only for chromatography, but also for mass spectrometry, optical spectroscopy, and nuclear magnetic resonance (NMR) spectroscopy. In addition, the value of 'live' data is realized through the ability to re-process, manipulate, and compare the data associated with different experiments run throughout the process.

Quick and Easy Reporting

Progress reports and presentations regarding impurity identification and characterization are required throughout the drug development process. And if all goes as planned, ultimately, a new drug application (NDA) is submitted with a list of impurities and a summary of the laboratory studies conducted to detect impurities. Specifically,

- Representative chromatograms should be provided.
- Chromatograms of representative batches from analytical validation studies showing separation and detectability of impurities (e.g., on spiked samples), along with any other impurity tests routinely performed, can serve as the representative impurity profiles.
- The applicant should ensure that complete impurity profiles (e.g., chromatograms) of individual batches are available, if requested."[xxii]

Preparation of reports, presentations, and, especially, an NDA can be very time-consuming and, as a result, quite costly. It is not unusual for days to be spent searching for the results of a specific analytical test. Even when the desired results are located, it is often difficult and time-consuming to incorporate the information into a master document. Therefore, IRM should incorporate quick and easy reporting capability, including the ability to:

- Quickly search and retrieve desired results by structure or metadata.
- Copy and paste chromatograms, spectra, structures, and tables directly into word documents, worksheets, and presentations.
- Quickly and easily reformat chromatograms, spectra, structures, and tables to change size, color, style, and font.
- Automate the creation of standard reports with customizable templates.

Conclusion

The identification and characterization of impurities in pharmaceutical drugs is critically important to ensure that their presence will not evoke any form of adverse response, either pharmacological or toxicological, in a patient taking the medication. It is also a frequent, highly specialized, and labor-intensive undertaking. Identifying and characterizing impurities in less time as they are encountered throughout the development and acceptance of a drug substance can shave weeks, or even months, off the time to market; increasing returns by millions of dollars. In addition, proper retention of this information and context through the lifetime of a drug in its development environment can provide insight to help resolve quality issues downstream when/if post-marketing and manufacturing issues arise that can lead to delays and drug shortages. To meet these requirements scientists must be able to obtain the highest quality data possible in the shortest amount of time. Impurity Resolution Management (IRM) is an essential aspect of drug development and acceptance that combines software, algorithmic tools, and databases to collect and unify chemical, structural, and analytical data for quick and easy search, retrieval, and visualization. This enables cross-functional scientific collaboration among the process chemistry, analytical chemistry, and structure characterization, identification, and elucidation groups through comprehensive and dynamic visualization, optimal method development, and quick and easy reporting. With Impurity Resolution Management, scientists can quickly and easily search and retrieve all relevant data to gain and apply insight that improves decision-making thereby reducing the time and decreasing the cost of impurity identification and characterization.

Notes

[i] Gary E. Martin, A Systematic Approach to Impurity Identification, in Analysis of Drug Impurities (Richard J. Smith and Michael L. Webb, ed.), Wiley-Blackwell; 1st edition (April 30, 2007), 124.

[ii] Guidance for Industry: Q3A Impurities in New Drug Substances, Revision 2, Attachment 1: Thresholds, U.S. Department of HHS: FDA, CDER, CBER and ICH (June 2008), 11.

[iii] Margaret A. Hamburg, M.D., FDA Letter to Industry, http://www.fda.gov/Drugs/DrugSafety/DrugShortages/ucm277675.htm, (October 31, 2011).

[iv] Martin, 152.

[v] Steven M. Paul, Daniel S. Mytelka, Christopher T. Dunwiddie, Charles C. Persinger, Bernard H. Munos, Stacy R. Lindborg and Aaron L. Schancht, How to improve R&D Productivity: the pharmaceutical industry's grand challenge, *Nature*, (March 2010), 204–205.

[vi] Eric David, Tony Tramontin and Rodney Zemmel, The Road to Positive Returns, Invention Reinvented: McKinsey perspectives on pharmaceutical R&D, McKinsey & Company (2010), 4.

[vii] Martin, 129.

[viii] David, 4.

[ix] M.E. Elyashberg, A.J. Williams and G.E. Martin, "Computer-assisted structure verification and elucidation tools in NMR-based structure elucidation", *Progress in Nuclear Magnetic Resonance Spectroscopy*, Elsevier, (8 July 2008), 53:84.

[x] Sooho Lee and Barry Bozeman, The Impact of Research Collaboration on Scientific Productivity, *Social Studies of Science*, (October 2005) 35: 673–702.

[xi] Martin, 152.

[xii] Ryan Sasaki and Bruce Pharr, Unified Laboratory Intelligence, ACD/Labs, http://www.acdlabs.com/unified_lab_intelligence/, (April 2013), 2.

[xiii] Anne E. Thessen and David J. Patterson, Data issues in life sciences, PMC (NIH/NLM), http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234430/, (November 28, 2011).

[xiv] Thessen.

[xv] Guidance for Industry: Q3A Impurities in New Drug Substances, Revision 2, U.S. Department of HHS: FDA, CDER, CBER and ICH (June 2008), 5.

[xvi] Guidance, 4.

[xvii] M.E. Elyashberg, 83.

[xviii] M.E. Elyashberg, 84.

[xix] Anna Codina, Robert W. Ryan, Richard Joyce, and Don S. Richards, Identification of Multiple Impurities in a Pharmaceutical Matrix Using Preparative Gas Chromatography and Computer-Assisted Structure Elucidation, *Analytical Chemistry*, (October 13, 2010), 9127–9133.

[xx] David C. Adams and Sanjivanjit K. Bhal, How Software Can Aid Decision-Making in Chromatographic Method Development, *Chromatography Today*, (February/March 2013), 22–24.

[xi] Sasaki, 9.

[xxii] Guidance for Industry: Q3A Impurities in New Drug Substances, Revision 2, U.S. Department of HHS: FDA, CDER, CBER and ICH (June 2008), 5.

Bruce Pharr is a product and marketing consultant with over 25-years of experience in technology, life sciences and healthcare. He has held executive positions with Symyx Technologies (hardware, software and services for chemical, energy and life sciences) and GenoLogics Life Sciences Software. Bruce has authored articles on R&D informatics and data management for Bio-IT World, Drug Discovery World and Drug Discovery News, and he has contributed to articles in Laboratory Equipment, Lab Manager and Scientific Computing World.

Ryan Sasaki is the Director of Global Strategy at ACD/Labs, Inc. It is his responsibility to liase with influential industry personnel and authorities in the chemistry and laboratory informatics industries. He has presented at several analytical chemistry and informatics events worldwide and has authored a series of publications on analytical chemistry and laboratory informatics. Most notably, he is the co-author of the white paper introducing the concept of Unified Laboratory Intelligence (ULI). Prior to his career at ACD/Labs, he served as a Process Chemist for a major pharmaceutical company in Canada.