ACD/Labs Application Note

Managing Impurity Knowledge in a Process Chemistry Environment

Ryan Sasaki and Graham A. McGibbon
Advanced Chemistry Development, Inc.
Toronto, ON, Canada
www.acdlabs.com

CONTENTS

Introduction

Unified Laboratory Intelligence (ULI) is a new category of laboratory R&D informatics that provides a technology framework to improve chemical identification, characterization, and optimization for life science, chemical, energy, consumer products, industrial, and technology industries. ULI is designed to collect and unify chemical, structural, and analytical data as chemical content with context, which enables the generation of intelligence from disparate information within a laboratory environment. In order for any laboratory informatics solution to be implemented successfully and ensure productivity improvements and clear returns, it must first be properly focused. "The most successful strategies are likely to be those that address needs in the context of sub-disciplines."[1] As a result, this application note will focus on the application of ULI within the context of impurity resolution within the drug development process of the pharmaceutical industry.

Within the pharmaceutical industry, the rapid detection, identification, elucidation, and characterization of synthetic or process impurities and degradation products is an intense and comprehensive undertaking. Regulatory agencies including the FDA and ICH mandate dose dependent thresholds for reporting, identification, and qualification of impurities. For example, in the development of a formulated drug substance, the FDA requires that all impurities introduced in the proposed process above 0.1% by area percent need to be isolated and structurally characterized (the requirement can be even lower for drugs with larger doses). Impurities from synthesis, manufacture, or storage are all of interest. Furthermore, in order to develop a robust drug product, degradation products must be characterized with the intent of minimizing their presence (thus preserving the shelf life of the formulated drug product).

The identification, elucidation, and characterization of impurities and degradants in this environment requires a multi-disciplinary approach with good communication and collaboration between various groups. Through the entire drug development process, a wealth of chemical and analytical information must be generated, collected, and documented to facilitate the identification, elucidation, and

characterization of impurities. Furthermore the trend toward a paperless environment facilitates the management, dissemination, and rapid access to such large volumes of accumulated information.

Figure 1*: A summary of the groups involved in the impurity identification, elucidation, and characterization process in pharmaceutical development.*

Pharmaceutical Development groups (Pharm Dev or similar variants) are ultimately responsible for the CMC (Chemistry, Manufacturing, and Controls) sections of IND applications and NDA submissions. In general the groups involved in this process from a synthetic chemistry and data collection and analysis point of view are (also depicted in **Figure 1**):

- **Process Chemistry**—process chemists are responsible for developing the synthetic route for development candidates. They are expected to collaborate with the groups below to facilitate the identification of impurities or to refine their synthetic route design to eliminate or minimize the presence of these impurities.
- **Analytical Research and Development**—typically responsible for the development of Drug Substance (DS) and Drug Product (DP) specifications, this group will work concertedly with the process chemistry and structure characterization groups to enable impurity profiling and chemical stability evaluation with degradant identification. The “impurity profile” method is developed by Analytical R&D and this method is used to help identify all impurities through each stage of the synthetic route and to help the characterization group decide which impurities require identification. In this group, a significant amount of time, technology, and knowledge is also invested in method development, tests of robustness, and validation.
- **Structure Characterization Group**—typically responsible for performing the structure elucidation and characterization of impurities, in ideal situations, they work in close collaboration with the

process chemistry and method development groups to obtain a full sample history ('context') including the synthetic route, experimental conditions, solvents used, pH, temperature, etc. These groups most often work as interdisciplinary spectroscopy teams which bring a wealth of expertise in techniques like mass spectrometry, NMR and optical spectroscopies, and X-Ray Crystallography, amongst others. Elucidating and characterizing impurities in bulk drug substances and formulated drug products are one of their main responsibilities. (**Note**: the Structure Characterization Group is often a part of the Analytical Research and Development Group in some organizations.)

This application note is written from the perspective of the Process Chemistry Group and their focus on developing safe, practical, cost-effective, and yield maximizing synthetic routes as they scale up synthetic routes from lab to manufacturing scale. It will highlight the challenges in their environment, and the benefits of having a robust, integrated, and instrument neutral software platform with tools designed to improve collaboration with cross-functional groups and capture process knowledge.

Managing a Continuously-Evolving Synthetic Process

The process of scaling-up reactions from lab-scale to the manufacturing level is not a simple, linear process.[2] A series of problems can arise which force the chemistry group to alter and optimize steps with new materials, intermediates, solvents, and different reaction conditions at various points of drug development through the pilot scale to manufacturing scale. In the process chemistry environment the emphasis of the chemistry is shifted from the ability to simply make the molecule to the ability to make it in a safe, cost-effective, practical, and environmentally-friendly way. As a result of these priorities, through various stages of drug development the proposed synthetic route will be altered, changed, and optimized. Furthermore, the presence and identification of specific impurities in a given process need to be identified and evaluated, which can lead to changes or modifications in the master synthetic route. Finally, one of the other main priorities of a process chemistry group is to efficiently capture process knowledge that can be leveraged on successive projects down the road. The capture of this knowledge is a challenging task.

ACD/ChemAnalytical Workbook allows process chemistry groups to create/edit their synthetic routes, or import or copy/paste it from another informatics system (for example, from a record in their electronic lab notebook). This information is transferred as 'live' chemical information that can be modified, edited, or updated as the synthetic route changes and evolves over the course of the project. Within the database, the process chemist can access and manage all information related to a chemical structure, a sequence in the reaction, or the reaction itself. Related data such as biological and toxicological information, text, batch, product, and experiment IDs, links to associated documents, and most importantly the 'live' analytical data associated with each structure represented in the reaction can be accessed readily (**Figure 2**).

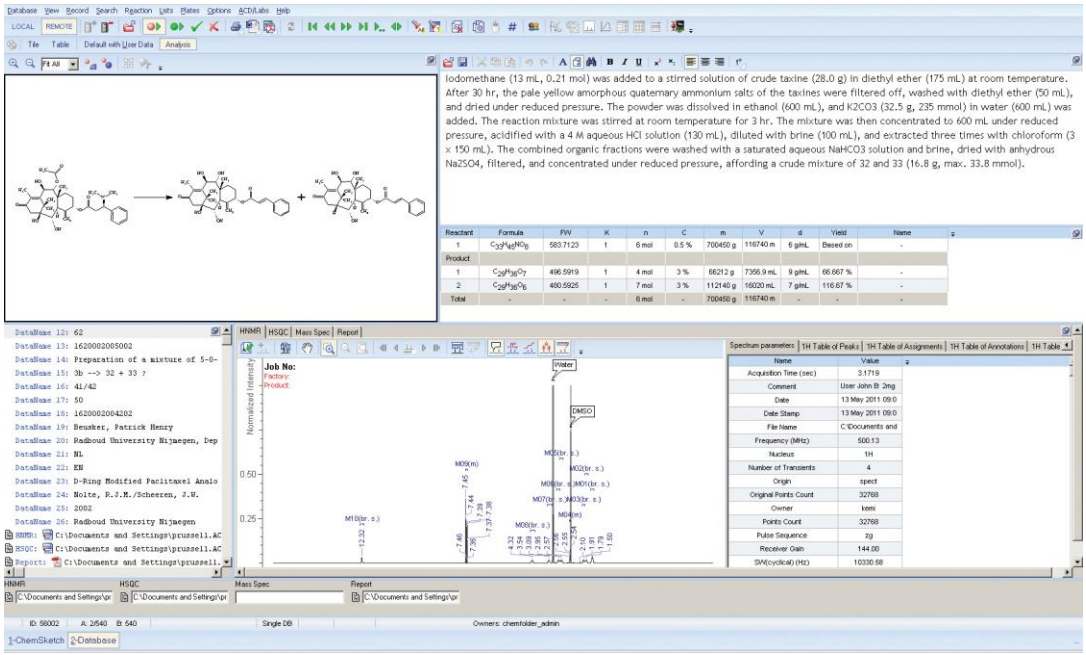

***Figure 2**: Within the ACD/ChemAnalytical Workbook environment process chemists can update/edit the synthetic route to be viewable by other cross-functional teams. Users can access a variety of data and information with the compounds in the route, a step in the route, or for the entire reaction such as biological and toxicological data, text, notebook #s, and 'live' analytical data.*

The Importance of Impurity Knowledge for Route Optimization

Throughout the long and evolving process of pharmaceutical development circumstances arising from separation, purification, and elucidation attempts, the process chemistry group may be prompted to refine and adapt their original synthetic route. As the knowledge of these impurities (and their respective fates) grow through multi-disciplinary discussion, the synthetic route can constantly evolve to help engineer these unwanted impurities out of the finished drug product. In order to understand the presence, context, and impact impurities may have on the process, it is important for all the groups involved (process chemistry, method development, and structure characterization) to participate, communicate and collaborate at a high level. With data and information being exchanged and stored in different systems and different types of media (paper, word documents, spreadsheets, etc.) this sharing of knowledge can become very challenging, and providing these groups with a collaborative platform for multi-disciplinary knowledge sharing in an electronic environment can be extremely valuable.

As mentioned above in **Figure 2**, with ACD/ChemAnalytical Workbook, Process Chemists can store the constantly evolving synthetic route as 'live' data that can be edited and modified over time. However, behind the scenes of these synthetic chemistry efforts, method development and analytical chemistry specialists are generating significant amounts of data for the separation, isolation, and characterization of key intermediates, final products, and impurities. The ACD/Spectrus Platform is capable of managing all of this data in a 'live' manner to capture the impurity profile of the drug substance. By intimately linking the 'live' analytical data from the method development and structure characterization groups to the relevant synthetic stages of the master synthetic route, real-time process knowledge can be captured (**Figure 3**). Furthermore, by making this information accessible with comprehensive searching utilities it can be an invaluable tool for providing early insights and helping guide future projects. This helps to set up an ideal collaboration platform where knowledge of the chemistry is shared with the associated data.

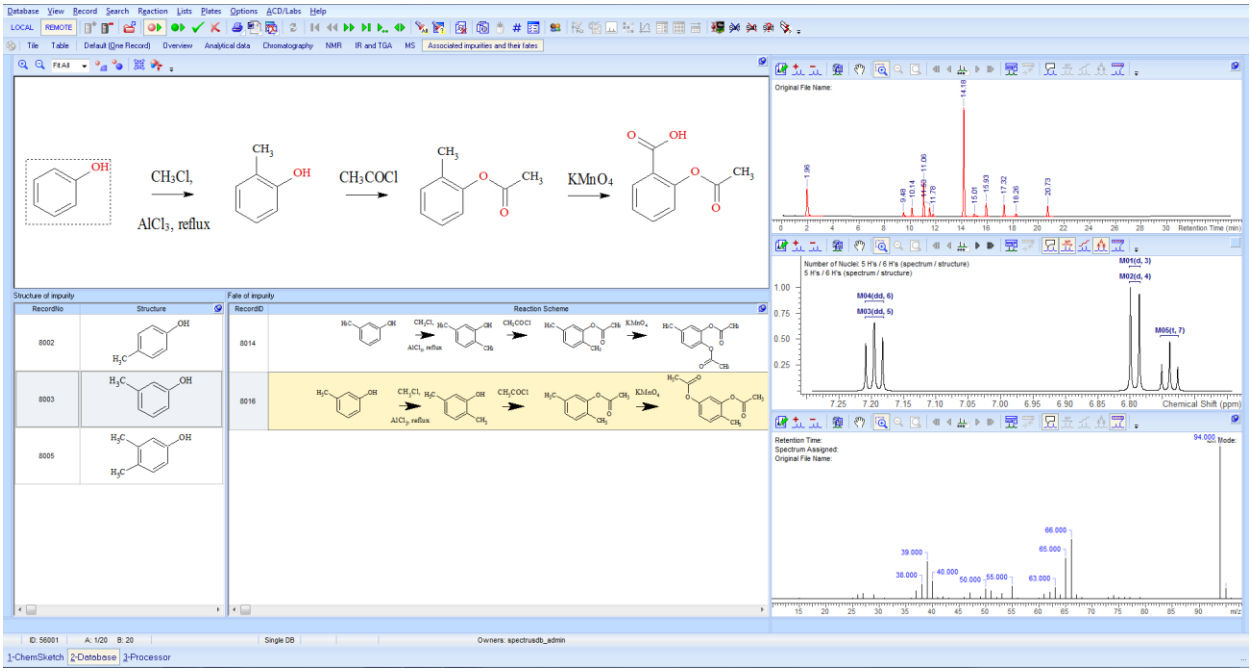

Figure 3: *The ACD/Spectrus Platform can capture live analytical data with related information and connect it with the chemical context of the development project. This view shows the master synthetic route. Clicking on individual steps will reveal a table (bottom left) of already identified impurities. Clicking on structures in the table will reveal all the relevant analytical data acquired for that impurity (right). In addition, the synthetic fate of each impurity can be tracked through the process (bottom middle).*

Offline Data Processing for Process Chemists

Historically, organizations have employed analytical chemistry groups to handle the acquisition, processing, and interpretation of routine analytical data generated in support of synthetic chemistry efforts. Now, however, thanks in large part to advancements in “Open-Access” or “Walk-up” analytical instrumentation, many routine experiments are elicited by the process chemists themselves. The advantage of this approach is twofold:

1) It reduces the sample burden placed upon the analytical support group and thus the turn-around time associated with their activities.

2) It places well-defined results or data into the hands of the individuals most familiar with the chemistry of their samples.

Unfortunately, the evolution of open-access data handling creates its own issues. The drawback of waiting for results has been replaced with the burden of processing, analysis, and interpretation of their own analytical data, typically within different software applications. This added workload has changed the priorities of synthetic chemists, requiring them to spend more time interrogating analytical data, which can have a negative impact on their overall productivity.

ACD/ChemAnalytical Workbook gives groups with open-access bottlenecks an inexpensive and easy way to provide each chemist with access to all their analytical data on their desktops in their own lab or office (**Figure 4**). This avoids queues in the instrument room and allows chemists to focus on their data at their own pace. With improved accessibility, synthetic chemists can inspect live data easily using a variety of beneficial features. The offline application also serves as a portal into the impurity knowledge base where

chemists can search past projects based on variety of comprehensive search parameters, i.e., notebook #, sample ID, structure, substructure, search by peaks, etc.

***Figure 4**: With ACD/ChemAnalytical Workbook, process chemists can access, view, and re-process all analytical data at their desktop away from the instrument room. In addition, they can quickly search and mine data from previous projects directly through this interface.*

Evaluating the Genotoxic and Carcinogenic Risks of Impurities

In the case of impurities that present a high safety or efficacy risk of the active pharmaceutical ingredient, approaches may need to be taken to reduce the presence of this impurity. This may include changing or modifying the synthetic and/or purification routes to minimize the formation and/or maximize the removal of the relevant impurity. Compounds that have been demonstrated to induce genetic mutations, chromosomal breaks, and/or chromosomal rearrangements are considered genotoxic and have the potential to cause cancer in humans. The impurity threshold limits mentioned earlier are not always deemed acceptable for genotoxic or carcinogenic impurities and thus may need to be evaluated at lower detection levels. Unfortunately, the isolation of some trace impurities is a challenge and material can be scarce. Decisions need to be made on whether to isolate more impurity material for toxicology studies, which can be a time-consuming process. The FDA has provided guidance for assessment of genotoxicity or carcinogenicity by computational methods for impurities in drug products present at levels below the ICH qualification thresholds. These *in silico* tests can be applied to quickly determine if an impurity poses a potential safety risk or to prioritize it for testing when toxicological data is limited or lacking.

Within the ACD/Spectrus Platform, users can auto-populate and access a battery of *in silico* tests (**Figure 5**) to accurately assess the genotoxic and carcinogenic potential of impurities in drug products as they appear in the knowledgebase. Developed in collaboration with the FDA, the genotoxic and carcinogenic risk prediction algorithms can help organizations remain compliant with regulatory submission requirements by assessing compounds for:

- Mutagenicity (Ames test, Mouse Lymphoma Assay, and other standard assays)
- Clastogenicity (Micronucleus test, Chromosomal Aberrations)
- DNA damage mechanisms (Unscheduled DNA Synthesis)
- Carcinogenicity (FDA rodent carcinogenicity data)
- Endocrine disruption mechanisms (estrogen receptor binding)

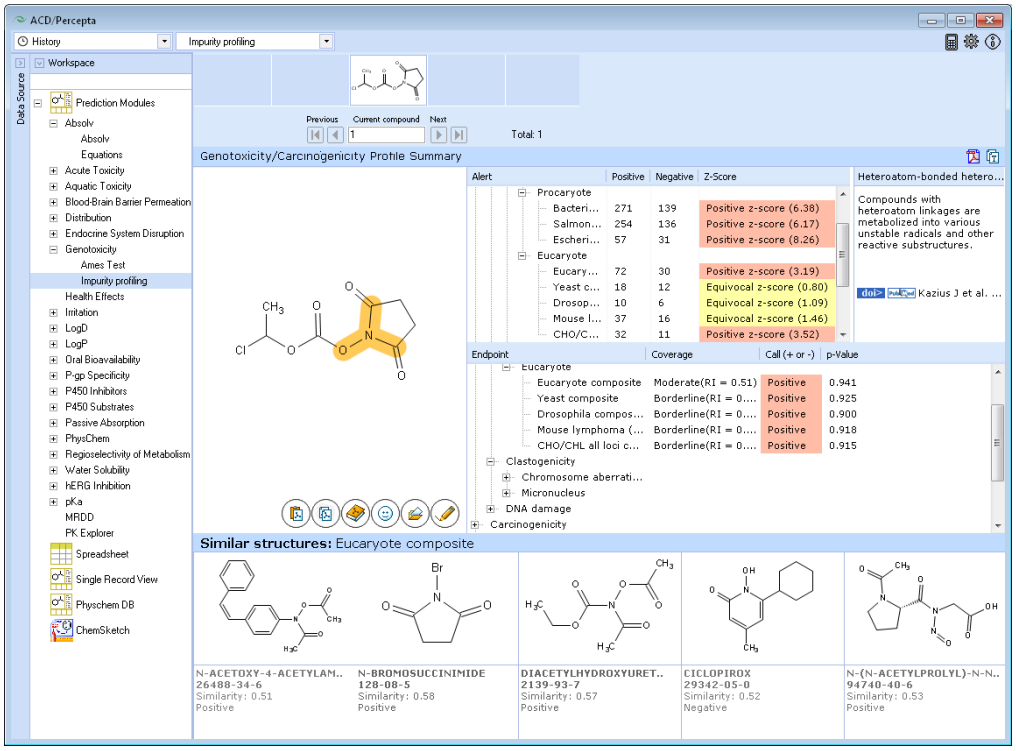

***Figure 5**: Populate your database with fields highlighting impurities with genotoxic potential.*

Bringing it All Together—An Electronic Platform for Collaboration and Reporting

As discussed in previous sections, a high degree of communication and collaboration between scientists in a multi-disciplinary group is advantageous in helping to develop a drug impurity profile. Having quick answers to questions like, "What was the intended product? What are the starting materials and reagents used? What were the reaction conditions that lead to the formation of a specific impurity? What are some related impurities we've already identified?, etc." can prove instrumental in the quick and timely optimization of synthetic routes.

Throughout the drug development process, and most importantly at the time of submission, progress reports, presentations, and final documentation of the CMC sections for IND and NDA filings are required. The preparation of this type of documentation can be very time-consuming, quite costly, and disruptive, as it requires collecting, collating, and combining representative data and chromatograms from various stages of the process. In a typical scenario, it is not unusual for hours/days to be spent searching for/compiling results of a series of specific analytical tests. Even when the desired results are located it might be difficult to understand the context surrounding the data that represents a test. A worst-case scenario involves having to re-acquire data from samples if previous test results cannot be located. In some cases this may even require re-purification of a sample or even worse re-synthesis. Clearly activities such as this can be hugely disruptive and most important costly due to delays in either NDA filing or response to an already existing submission.

The issues and time spent in the reporting process is not limited to searching for and finding relevant data. Often times, once the data is located it still involves a laborious process of copy and pasting different images from different data processing applications. Because of the vast array of analytical instrumentation used throughout the drug development process, having to interact with different data processing software applications and portals to access and re-plot data can be a very time-consuming process. The major benefit of using ACD/Spectrus Platform as a collaboration tool for data and knowledge sharing is its ability to unify the 'live' analytical data from the significant variety of heterogeneous instrument data formats in the typical laboratory informatics environment. As a result, if all the data and its context are stored in one electronic environment, key information can be compiled and combined easily and a report can be generated in the click of a button, saving days/weeks of time in report compilation/creation.

The importance of 'live' data should not go unrecognized. When data is re-accessed and re-evaluated it often will need to be re-processed, re-analyzed, or manipulated in some way. This cannot happen with 'dead' data stored as a PDF, or locked into a specific, proprietary instrument data format. In addition, a 'live' chemical and analytical data collaboration platform, ACD/Spectrus DB, helps compile all of the chemical and analytical information into the context of the life cycle of the drug being developed. As impurities arise, the Process Chemistry, Analytical R&D, and Structure Characterization groups have a platform where they can quickly visualize and assess the context of a given impurity or a particular stage in the process. The ACD/Spectrus Platform enables each scientist to dive deep into the context of the data interpretation where they can evaluate not only the data associated with the chemical structure, but how the data was assigned to the NMR, MS, and Optical data. This knowledge provides in depth chemical context, and serves as a foundation to create a more intelligent informatics environment by leveraging this knowledge to improve the algorithms being used routinely for various structure elucidation efforts.

This means that when these groups are contacted by regulatory personnel with requests for required or additional information for regulatory filings, the information can be easily searched for, accessed, and delivered in a timely fashion; avoiding any delays or re-run of experiments.

Conclusion

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.[1] Impurity Resolution Management (IRM) is an essential aspect of drug development where Unified Laboratory Intelligence (ULI) can be successfully applied to improve decision-making, reduce time, and decrease cost. As highlighted in this application note, the ACD/Spectrus Platform can serve to help process chemistry groups manage and capture the constantly evolving synthetic route and conditions from lab to pilot to manufacturing activities. Furthermore, it enables the connection of analytical data from different forms of instrumentation to starting materials, intermediates, impurities, and final products to capture process knowledge through chemical context.

---

[1] 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**).
[2] Gina Shaw, Micrograms to Kilos: The Challenges of Scaling, http://www.dddmag.com/articles/2007/09/micrograms-kilos-challenges-scaling (**Sept. 6, 2007**).