Reprinted from *American Laboratory* January 2012

Pharmaceutical

by Philip Keyes, Michael Messinger, and Gonzalo Hernandez

Automated Structure Verification by NMR, Part 1: Lead Optimization Support in Drug Discovery

The pharmaceutical industry has always relied on testing and evaluating novel compounds in a wide range of chemical space to feed its pipelines. Physical properties and the ability of the compound to hit its intended target are crucial in determining the success of a medicinal chemistry campaign. The chemist's interpretation of the structure/activity relationship (SAR) of candidate molecules is at the very core of lead optimization and successful nomination of a specific compound as a clinical candidate. Nothing, then, can be more confounding or wasteful than counterintuitive or misleading structure/activity relationships that can derail or misguide the efforts of a team while selecting the next compounds to make for testing.

Therapeutic area teams must trust the integrity of the molecules being tested; yet mistakes happen. Molecules with incorrect chemical structures are submitted every day across the industry. Sometimes these mistakes are subtle and are due to administrative errors, such as misplacing a functional group on a ring or omitting a double bond when drawing the structure electronically. Peer-reviewed journal articles are subject to the occasional error in interpretation or drawing of reported molecules.[1] Certain Web logs chronicle the misdrawn and misrepresented molecules that are published; one in particular has numerous alerts regarding structure errors.[2] The more detrimental cases occur when an incorrect structure is the result of an unexpected or undetected outcome of a reaction, resulting in a product that is different from its expected structure.

It is critical to the outcome of a project that samples submitted for testing are correct. Incorrect structures have both a short- and long-term impact on research costs. Chemists find nothing more frustrating than beginning a discovery program with a series of leads only to find that, upon resynthesis, the leads are not active. This is especially true if a lead appeared to be exceptionally potent. In instances in which sufficient material remains for analysis, a great deal of effort can be expended to perform structure elucidation.

The first part of this two-part series will focus on the scientific and technical benefits of system implementation. Part 2 will cover the return on investment and financial analysis of system implementation.

Use of mass spectrometry (MS) and nuclear magnetic resonance (NMR) to characterize and validate compounds

Analytical techniques such as mass spectrometry and nuclear magnetic resonance are generally utilized to characterize and validate the compounds submitted for testing. Before the advent of high-throughput automation systems such as sample changers and autoinjectors in the 1980s and '90s, it was typical for chemists to submit compounds to staff analytical chemists. Lately, this dependence has shifted to chemists themselves as organizations have continued to minimize support staff. The fundamental problem occurs in the interpretation. A higher dependency has been placed on MS data to characterize compound integrity due to the simplicity of the method and interpretation of the data. Heavy dependency on this single analytical technique is not without risk, however, since many aspects of MS can lead to overly optimistic interpretation.

Some molecules are not easily ionized and may be invisible to MS, UV, or even evaporative light scattering (ELS) detectors. More easily ionized compounds, on the other hand, may actually be the less abundant species in such a sample, and may be overrepresented in characterization. It is common for most pharmaceutical drug discovery teams to require proton-NMR, MS, HPLC, or other auxiliary analytical methods as evidence of successful syntheses of proposed target molecules. A scant percentage of the content of the collected analytical data is used to verify or confirm the identity of proposed molecules. The main reason for this is that it is not the primary function of a medicinal chemist to interpret spectral data, but rather to synthesize compounds. The more time chemists spend characterizing their molecules, the less time they spend making them. Therefore, management and researchers must strike a delicate balance between quality and quantity of compounds produced.

Implementing automated chemistry

Automating chemistry is often used as a means to increase productivity,[3–5] and typically results in reduced requirements for analytical characterization as well. This can lessen the workload for medicinal and analytical chemists. Industry requirements for purity are dependent on the purpose of the compound. In general, individual custom synthetic molecules are required to meet a minimum 95% purity criterion, while compounds for libraries synthesized under automation are typically accepted within a 65–80% purity range.

In many cases, characterization of compounds synthesized using automation

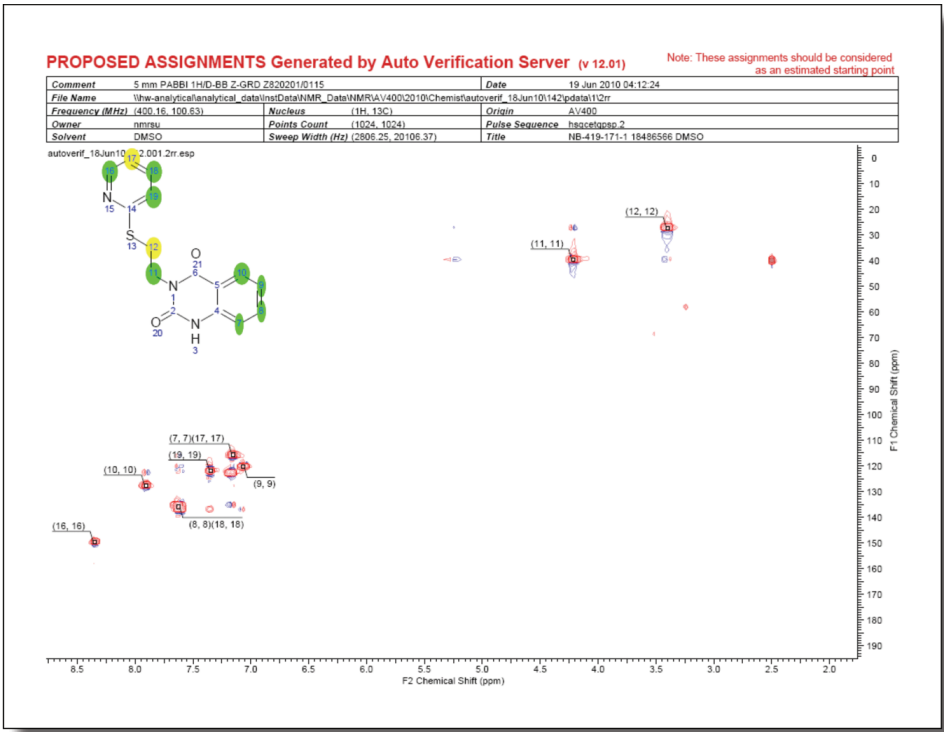

Figure 1 *Typical autoverification result on commercial compound using HSQC experiment.*

requires only a passing LC-MS. This is often done for practical reasons related to sensitivity. While these types of methods can be employed to increase "shots on goal" for hit/lead identification, they cannot meet the quality requirements and amount of material necessary in the lead optimization process, which is the focus of this discussion. Here, the higher requirements for purity allow us to more readily investigate the consistency of a compound's spectral parameters (NMR or LC-MS) relative to derived or calculated values. For NMR, this would be chemical shifts, coupling constants, and correlations. For MS, it would be associated parent ions and adducts.

To further reduce the chemist's data interpretation workload, spectral interpretation can be automated. Only recently has work been published demonstrating this principle. When automated interpretation is run as a background process, greater value can be extracted from the data that have already been collected (NMR, LC-MS, etc.), which may otherwise have been overlooked due to the limited time available for manual interpretation. *Figure 1* depicts an example of the automated interpretation output. The key to successfully accomplishing this goal in the authors' laboratory was to build the work flow into existing activities and provide the results with no extra burden to the chemists.

NMR verification system description

The automated NMR verification system that has been implemented is designed to eliminate additional work for the chemist. Use of a simple convention for lab notebook number reference at sample login on the spectrometer is all that is necessary to establish the basic structure-to-spectrum relationship. In addition to the proton-NMR experiment, the 2-D ($^{1}H/^{13}C$) heteronuclear single quantum coherence (HSQC) experiment, which is typically run for 15 min, is also collected.[6,7] With sophisticated higher magnetic field and more sensitive instrumentation, the HSQC data can be collected in 1 min. All other aspects of the work flow remain the same, as depicted in *Figure 2*. Automated functions occur completely in the background. These functions are triggered automatically upon completion of compound registration. Automated interpretation and subsequent generation of a verification score requires no additional effort on the part of the chemist, spectroscopist, or supervisor, and are conveniently reported back to the LIMS and compound registration database, where they are available for final review by supervisors prior to approval of the compound (see Figure 1).

System components

The current implementation of the system utilizes the NMR Expert, Automation Server, and C+H NMR predictors (version 12)[8] (**ACD/Labs**, Toronto, Ontario, Canada).

System justification

Budgetary approval for implementing this verification strategy is needed for a project of this scope and cost. In this case, the latter includes software license expenses and integration efforts; the converse is the cost of an incorrect structure. While an inactive compound will provide a negative result, an incorrect structure, whether active or not, will provide a misleading result. Either the SAR will not be supported or will be misdirected, or resources will be wasted running tests on a compound that otherwise would not have been tested. The worst-case scenario is that the correct compound (the one the chemist wanted to make but did not) would have answered the SAR questions correctly. This information cannot be gained without the correct intended molecule.

So, what is the cost of an incorrect structure? In the extreme, it is the cost of an entire program or project. When the cost of research on incorrect structures and the counterproductive, faulty, and misleading information they produce is compared to that of the cost of a verification system and all the auxiliary benefits it provides, a strong case can be made for investing in these systems. The cost benefits result from the following:

1. Enhanced integrity and quality of compounds submitted for testing

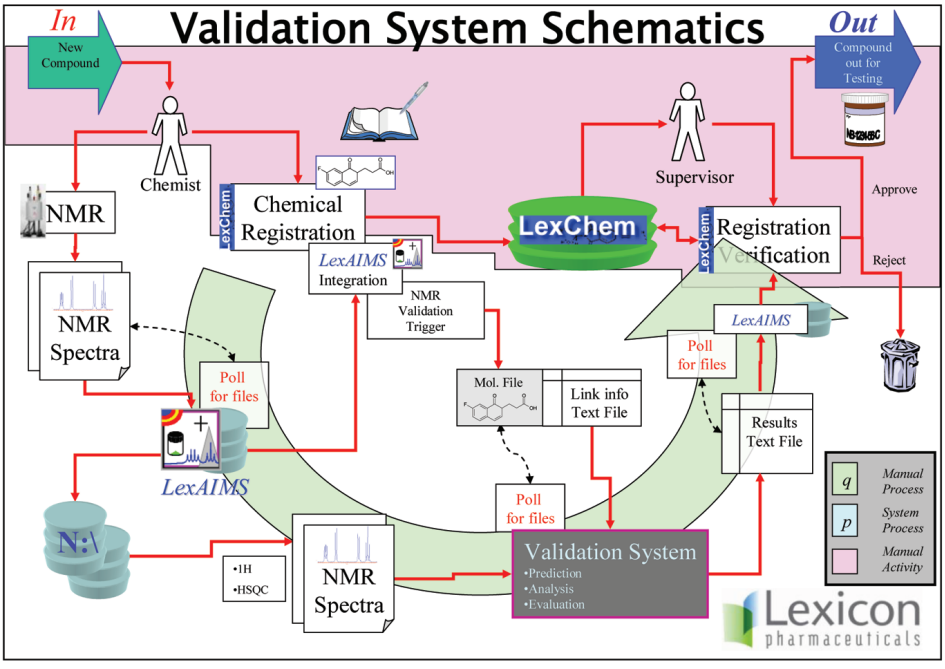

Figure 2 *Standard work flow diagram.*

2. Properly interpretable and actionable SAR, leading to potent molecules
3. Correct structures for compounds archived in compound libraries and production of interesting lead candidates during subsequent high-throughput screening (HTS) campaigns for future projects.

The key question is, "How many incorrect molecules are being submitted?" Even with the system currently in place in the authors' laboratory, this question cannot be fully answered since the team is still not completely confident that every compound identified is incorrect. However, enough errant compounds have been identified and corrected that the team is supportive of the process and wishes to continue.

Results

The accuracy and efficiency of an automated NMR verification system in conjunction with its counterpart techniques (e.g., LC-MS) are critical factors in reaching or achieving payoff for system installation. These results have been previously reported and continue to improve. Currently, the authors' laboratory has established the baseline level of performance of the autoverification system based on use in a production environment against over 3000 compounds evaluated, as well as from calibration and testing against four different commercially purchased reference compound sets totaling more than 250 compounds, with over 700 evaluations run against them under varying conditions using different NMR solvents and HSQC pulse sequences. The commercial compound reference sets were used to validate the system and provide a benchmark to establish the performance criteria, which are principally the true and false positive and negative rates that can be discriminated from scoring results. These compounds were produced and sold by reputable fine chemicals suppliers. As a result of benchmarking and detailed failure analysis, the error results indicated in *Table 1* were discovered by evaluating the negative results.

Each benchmark compound set negative result was analyzed exhaustively, and a full set of NMR spectra were collected, sufficient to determine to a great degree of certainty that the compounds were the incorrect structure, and in most cases to identify the actual structure. For purity, a very generous threshold of 70% was used to characterize unacceptable purity. In general, the prerequisite filter for running NMR autoverification is 95%. Registration errors were due to a combination of both internal receiving and structure record entry errors or a result of incorrect structures extracted from external databases via CAS#, such as the available chemical database (ACD Finder) or chemicals available to purchase (CAP) database and others. A total error of 5.2%, while seemingly high, is beyond the scope of this justification, since poor purity, while routinely failing verification as expected, will be detected by either HPLC or LC-MS as well. The more conservative figure of incorrect structures and registration errors was used and was found to be 3.2%. By arbitrarily selecting an even more conservative figure of 1%, little contention for the premise of return on investment should result and allow for a more academic discussion.

Performance metrics

The benchmark reference serves as a useful tool to calibrate and recognize system performance. We have observed a 22% false-positive rate using incorrect challenge structures against the data. The results from the production environment are of greater real-world interest. Due to the scale of the effort, however, the team presently does not have the ability to benchmark challenge structures for its larger set of production data. To date, we have observed a 15% false-negative rate of verification results in the production environment (see *Figure 3*, where

Table 1 **Structure errors found in commercial compounds purchased for benchmarking studies**

All compounds	Incorrect structure	Unacceptable purity	Registration error	Total error
250	5	5	3	13
% Error	2.0%	2.0%	1.2%	5.2%

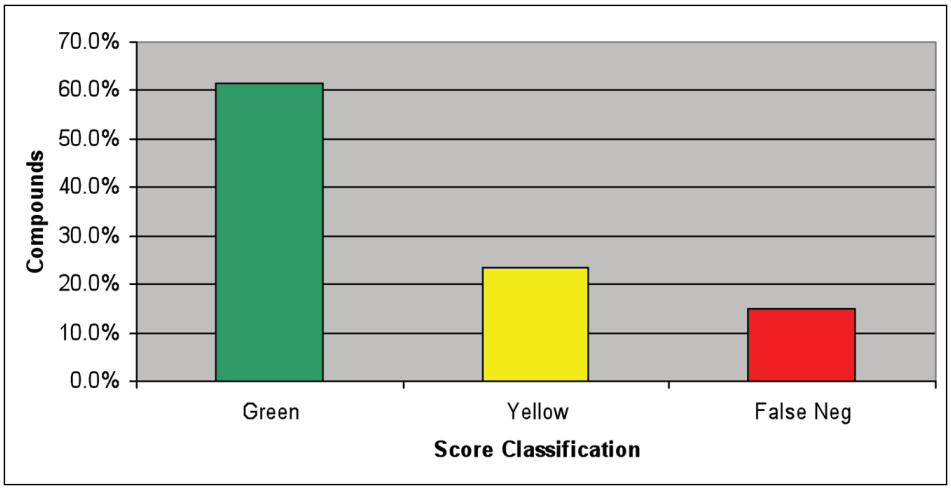

Figure 3 *NMR autoverification score classification of 2600 compounds.*

green is passing, red is failing, and yellow is caution for results reporting). In all cases, a failed result is reviewed by not only the supervisor, but by NMR staff as well. The system therefore focuses interest on the more potentially suspicious compounds while saving staff the effort of looking at more routine or easily differentiated compounds. This is a tremendous time savings, and is inherently why implementation provides a return on investment.

Summary: Performance data

Integration of automated compound verification using multidisciplinary inspection by NMR, LC-MS, and HPLC is practical in a research environment. The results also demonstrate the effectiveness of the system for a large number of compound submissions.

Part 2 of this article will show how financial and technical payback are achievable for modest-size organizations, from biotech through big pharma, in which 30 or more chemists are engaged in drug discovery for full implementation and verification of >50% of compounds. The article will explain how, under typical operating conditions—where a portion of submitted compounds are evaluated (~25%) and an error rate of 1–2% of errant compounds are identified—using simple error cost (SEC) and return on investment (ROI) models, ROI can clearly and easily be achieved in organizations with 40–60 chemists in typically 2–5 years.

References

1. Elyashberg, M.; Williams, A.J.; Blinov, K. *Nat. Prod. Rep.* **2010**, *27*, 1296–328.
2. http://pipeline.corante.com/archives/2010/05/24/great_moments_in_heterocyclic_chemistry.php.
3. Burbaum, J.J.; Ohlmeyer, M.H. et al. A paradigm for drug discovery employing encoded combinatorial libraries. *Proc. Natl. Acad. Sci. USA* June **1995**, *92*, 6027–31.
4. Ohlmeyer, M.H.J.; Swanson, R.N. et al. Complex synthetic chemical libraries indexed with molecular tags. *Proc. Natl. Acad. Sci. USA* Dec **1993**, *90*, 10922–10926.
5. Clark, M.A. et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nature Chem. Biol.* **2009**, *5*, 647–54.
6. Bodenhausen, D.R. *Chem. Phys. Lett.* **1980**, *69*, 185–9.
7. Davis, A.L.; Keeler, J. et al. *J. Magn. Reson.* **1992**, *98*, 207.
8. Sergey, S.; Golotvin, E. et al. *Magn. Reson. Chem.* **2007**, *45*, 803–13.

Philip Keyes is Director, Site Operations and Analytical Chemistry, and Michael Messinger is Director, R&D Finance, ***Lexicon Pharmaceuticals****, 350 Carter Rd., Princeton, NJ 08540, U.S.A.; tel.: 609-466-5596; fax: 609-466-3562; e-mail: PKeyes@LexPharma.com. Gonzalo Hernandez, Ph.D., is an NMR consultant at* ***Vis Magnetica****, Montevideo, Uruguay.*