
Philip Keyes, Joanne Rivera, and Vince Caruso
Lexicon Pharmaceuticals

Recent high profile cases of dug discovery and late stage pharmaceutical clinical candidate compound structure errors have shed light on a potentially risky situation. The current drama, as highlighted in the May 26 edition of the C&E News1, surrounding the elucidation of an active oncology agent by the Janda Group has brought into question the intellectual property protecting the claim of another group of scientists for the molecule known as TIC10 patented with an incorrect structure.2 This is the worst case scenario thus far.

Were this the only such case of an important but miss assigned structure, one might point to an acute failure of scientific rigor, however this is only the most recent in a string of relatively similar events resulting in a great deal of concern for the pharmaceutical industry. Bosutinib (supplied by distributors) and other cases have also made the headlines due to the embarrassing discovery of structural inaccuracy. Whether the structural assignment error is the result of an errant starting material or a misunderstood or misinterpreted reaction, the outcome is the same: Lost productivity, misdirection of resources and far worse, the loss of the value of an entire project!

Emerging automated methods for challenging the validity of a structure, Automated Structure Verification (ASV), have been reported previously.3, 4, 5 Enhanced validation methods including multiple challenge structure generation, incorporation of multiple spectra and systematic tightening of constraints have been assessed to determine their usefulness in preventing such nightmare scenarios that have occupied the headlines of trade journal articles lately. The real question is not whether ASV works, but how well and additionally, how well it needs to work to protect against the risk of catastrophic loss of resources, productivity and even patent protection. Is ASV enough or do further computer assisted structure elucidation (CASE) algorithms need to be put to use to provide adequate reliability of validation results. Depending on a risk assessment a trade off exists between cost of implementation and loss of value from compound structural assignment errors that will be different for each organization.
I examine a major contributing source of errors in this battle: the necessary reliance on specialty and custom chemical manufacturers to supply chemical building blocks, fragment libraries and novel investigational compounds for screening. A growing body of evidence points to a systematic failure rate of materials supplied. In general, between 2% to 8% of milligram and gram scale quantities of specialty chemicals have been reported to be incorrect. ASV uses 1H and 13C NMR chemical shift prediction comparisons to experimental data and automated assignment of atoms to spectrum features, ASV aims to protect against the consequences of errors in supply chain ordered compounds and internal custom syntheses. ASV has the potential to identify nearly 70% - 80% of incorrect isomer structures that LCMS cannot. Recent high profile case studies, such as in C&E News, highlighting incorrectly synthesized Bosutinib sold by third party vendors for research benchmark purposes, demonstrate the vulnerability of our discovery programs to supply chain induced synthesis errors. ASV may be helpful in preventing future cases.

Vendors in the library and fragment screening industry also recognize the perils of producing erroneous hit results due to incorrectly identified or mislabeled compounds citing their own rates of failure for material they use and implementing quality control measures to reduce these issues. Additionally we have conducted and published our own internal examinations of commercial building blocks and have obtained similar results.

![Figure 1: ASV Result for actual Bosutinib 1H + HSQC data against correct structure (score 88) and the 1H + HSQC data of the isomer verified against the Bosutinib actual structure (score 40)](image)

Table 1: Structure Errors found in commercial compounds purchased for benchmark studies

<table>
<thead>
<tr>
<th>All compounds (250 studied)</th>
<th>Incorrect Structure</th>
<th>Unacceptable purity</th>
<th>Registration Information Error</th>
<th>Total Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>% Error</td>
<td>2.4 %</td>
<td>2.4 %</td>
<td>1.2%</td>
<td>6%</td>
</tr>
</tbody>
</table>

We have identified these errors as a significant risk in the synthesis process that contributes to submission_posters of incorrect structures. A growing work load of compound verification is helping to build a strong case for more routine implementation of quality control checks using ASV systems as they continue to mature in the future. Additional strides in the use of additional 1D and 2D-NMR experiments applied together with 1H and HSQC ASV currently being evaluated as an Automated Multi-Spectral Structure Verification (MSV) system will be presented. Ultimately one must ask: Is ASV enough?