ACD/LABS [ADVANCED CHEMISTRY DEVELOPMENT, INC.]

Automated Verification of Small Molecule Structures at Sanofi

Celine Nguyen-Van-Dau
Head of Integrated Drug Discovery (IDD)

Sanofi's Structural Analysis Team works in support of synthetic chemistry, providing:

Structural/conformational analysis and structure verification for multiple sites	**Preparations of analytical descriptions for patents**	**LC/MS structural analysis** for ADC chemistry follow-up, DAR evaluation, and free drug quantitation
10,000 full analyses/year	1200 ^{1}H NMR and LC/MS descriptions/year	1800 analyses/year

When Sanofi's Structural Analysis team set out on their automated structure verification (ASV) journey more than 10 years ago, their goals were to:

- → Decrease repetitive tasks for NMR experts
- → Free up NMR experts' time and resources to focus on tasks that require NMR expertise (e.g., elucidation of unknown or complex structures)
- → Speed up preparation of ^{1}H NMR descriptions for patents

As their implementation matured over the last decade, they recognized its potential to help synthetic chemists quickly and confidently verify simple molecules based on ^{1}H NMR spectra from open access experiments, which led them to their current system that relies on ASV from NMR Workbook Suite™.

LIMS-Integrated ASV Workflow for Open-Access ^{1}H NMR Data

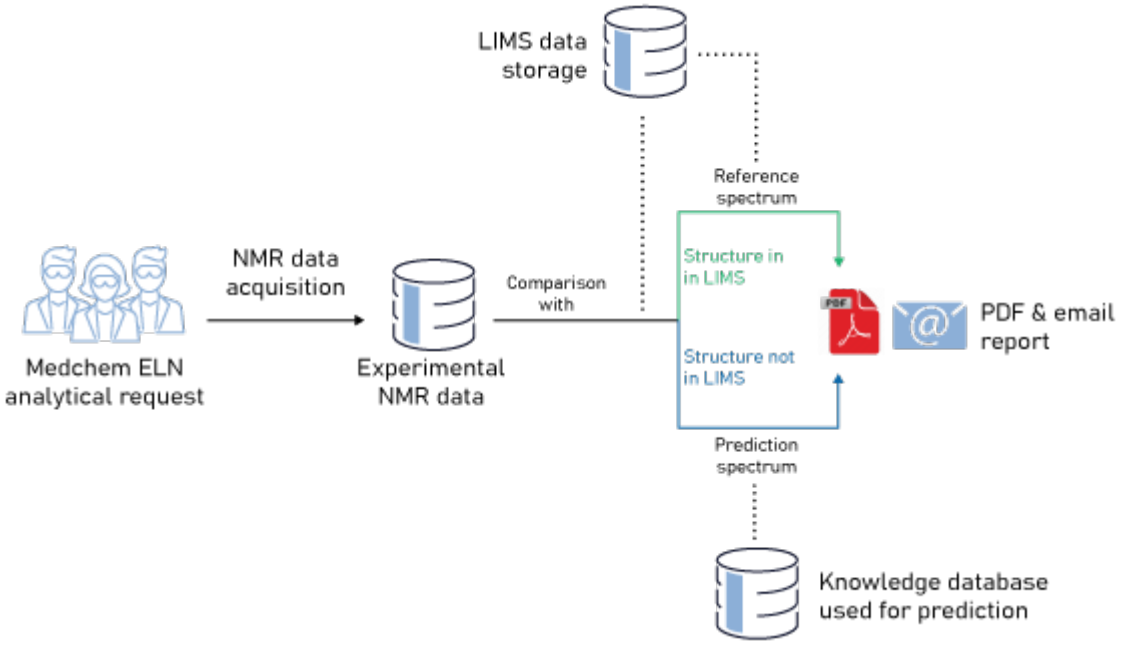

Figure 1. Sanofi's ^{1}H NMR ASV Workflow

In Sanofi's ASV workflow (Figure 1), when NMR data acquisition is initiated, whether by an electronic analytical request, or through the open access system, Sanofi's system automatically checks whether the proposed structure already exists in their LIMS.

If the structure is in the LIMS with a high-quality ^{1}H NMR spectrum from a previous batch, the ASV system will use this spectrum as a reference and compare it with the experimental ^{1}H NMR spectrum of the sample in question. This match is represented by a Hit Quality Index (HQI) between 0 and 100%, with 0% being no match between the two spectra and 100% being a perfect match. If the HQI is greater than 80%, the ^{1}H NMR spectrum is considered compatible with the proposed structure.

If, however, the proposed structure is not already in the LIMS, the ASV system will compare the experimental ^{1}H NMR spectrum with the predicted spectrum of the proposed structure. The result of this comparison is represented by the Match Factor (MF), a value between 0 and 1, where 0 indicates no match between the experimental and predicted spectrum, and 1 is a perfect match between the two. If the MF is greater than 0.80, the ^{1}H NMR spectrum is considered compatible with the proposed structure.

When the analysis is complete, the chemist or analyst automatically receives an email with the summary of the results and a PDF report. The report contains the conclusion based on the ASV result (i.e., “compatible with expected structure” or “compliance to be verified”), the MF/HQI, the ^{1}H NMR spectrum, and the multiplet report. This email also contains a link to the live data where chemists can easily see how the spectrum is assigned, make adjustments to the processing and/or assignment, change the template for the multiplet report, and see which structure(s) from the prediction database were used in the prediction, if applicable.

Reliable ASV for Small Molecules

In order for the system to meet the goals of their NMR experts and chemists, Sanofi had to ensure that the results from the ASV system were reliable. To evaluate this, a team of 6 chemists across 2 sites tested the system with spectra of more than 150 compounds.

They found that the ASV system was very reliable in cases where there is already a batch with the same structure in the LIMS. This is particularly useful for their scale-up teams.

In cases where the proposed structure was not already in the LIMS, they found the reliability of the results was highly dependent on the structures in the prediction database. When verifying novel chemical structures using only a commercial database for prediction, they were unsurprised to find that the resulting MF was very low. However, by adding only a handful of structures relevant to the proposed structure to the prediction database, they saw a substantial increase in the accuracy of the predicted spectra, and thus the MF resulting from ASV. (Figure 2) So, to ensure maximum accuracy, they regularly update the prediction database with their own confirmed structures.

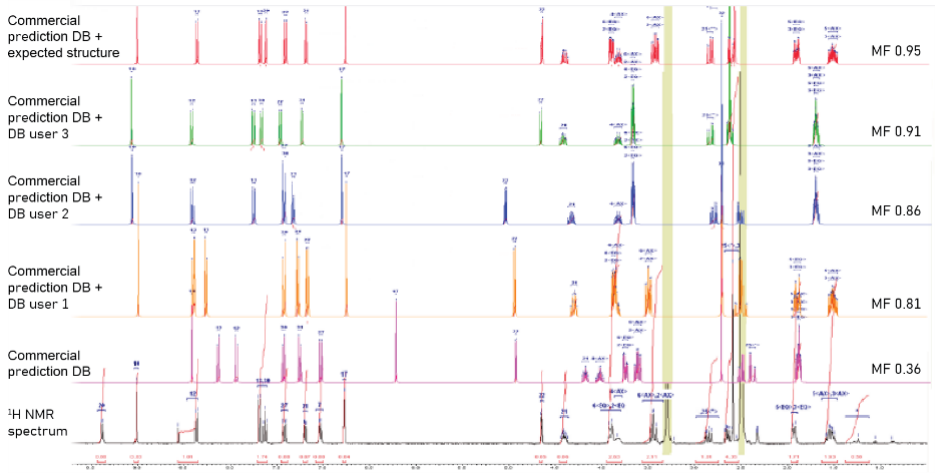

Figure 2. Spectral prediction, and thus verification, accuracy increases significantly as more relevant structures are added to the prediction database.

In this testing of the ASV system, they were also able to identify ways to reduce false negative results, or at least mitigate their impacts when they do happen. For example, they have now implemented an automatic warning on the report when the S/N in the experimental spectrum is low, to better help chemists and analysts identify false negative results for samples with low concentration or poor solvent quality.

"Accurate prediction is very important for our automated structure verification process and we've seen excellent results from training the NMR Predictors DB with our own data over the years." – Celine Nguyen-Van-Dau

Future

The ASV capabilities of NMR Workbook Suite are helping NMR experts and chemists at Sanofi quickly and confidently verify the structures of small molecules using 1H NMR data. They are working on further improving the system's reliability and expanding its benefits within the organization by. To do this, they are integrating it with a similar ACD/Labs®-based system they have been developing since 2020, which uses LC/MS data to check structure compatibility and evaluate sample purity.