Case Study

Automating Routine Structure Verification of Small Molecules by ^{1}H NMR using ChemisTwin™

Lauren Lytwak[1], Albert Farre Perez[2], Sarah Srokosz[3], Dimitris Argyropoulos[3], Shahriar Jahanbakht[3], Hans De Bie[3]
1 MilliporeSigma, Round Rock, United States, 2 Sigma-Aldrich (Merck KGaA), Buchs, Switzerland, 3 ACD/Labs, Toronto, Canada

Physical Reference Materials are a Burden for Scientists, Their Organizations, and the Environment

The identification of compounds is a fundamental task in chemical labs across all industries. This process can be accomplished through *de novo* characterization, comparison to literature references, or the use of standards to confirm compound identity. However, each of these methods are laborious to varying degrees. And in the last case, relying on physical materials as references requires the purchase and availability of these standards, along with additional sample preparation.

On-Demand Digital Reference Data

Recognizing that scientists value the corresponding data of physical reference materials more than the materials themselves, MilliporeSigma (Merck KGaA, Darmstadt, Germany, outside the US and Canada) envisioned a solution to eliminate the costly and error-prone process of handling physical materials. This led to the creation of the ChemisTwin™ portal, an online platform housing a vast database of digital reference materials (dRMs), serving as digital twins for over 25,000 physical references. Initially targeting routine structure verification workflows, ChemisTwin™ leverages NMR Workbook Suite's predictive and automated structure verification capabilities to provide high-quality, on-demand data and comprehensive reports, streamlining processes across industries.

Assessing the Reliability of Automated Structure Verification Using dRMs

The automated structure verification capabilities of ChemisTwin™ were assessed with ^{1}H spectra of 65 different molecules with structures corresponding to existing dRMs. The raw data from these spectra were compressed and uploaded to the ChemisTwin™ portal. The appropriate standard for comparison with the sample spectra was then selected, and the structure verification analysis was conducted. ChemisTwin provided the results, offering a match factor quantifying the agreement between the sample and standard spectra (Figure 1).

Figure 1. Workflow of the structure verification using ChemisTwin.

The first test (T1) was conducted on 80 spectra corresponding to 65 molecules recorded using the same solvent as that of the dRM. The second part of the test (T2) used 12 spectra recorded using a different solvent than that of the dRM. (Figure 2, b)

ACD/Labs®

ChemisTwin™ Provides Reliable Results for Common Solvents

ChemisTwin™ provided a table containing proton assignments, along with the sample and the dRM spectra (predicted to take into account the sample's experimental conditions—i.e., solvent and instrument frequency), and the match factor score between the two spectra.

For T1, ChemisTwin™ correctly verified 71 of 80 spectra (89%) with 3-star accuracy. (Figure 2a) Eight spectra (10%) received 2 stars, indicating a questionable result requiring review by the scientist. Using the results provided by ChemisTwin™ (spectra and table of assignments), the target molecule's identities were confirmed in all of these cases. One spectrum (1%) showed no match due to a mislabeling of the solvent peak from a nearby singlet to the chloroform signal. (Figure 3)

When different solvents were used, the performance of ChemisTwin™ declined slightly, with 6 of 12 spectra (50%) earning 3 stars, 5 spectra (42%) getting 2 stars, and 1 spectrum (8%) resulting in no match due to limited data with benzene. (Figure 2b) Despite this, ChemisTwin™ remains reliable overall with common solvents like DMSO, $CDCl_3$, MeOD, and CD_3CN. For best results, it is recommended to check dRM conditions before recording.

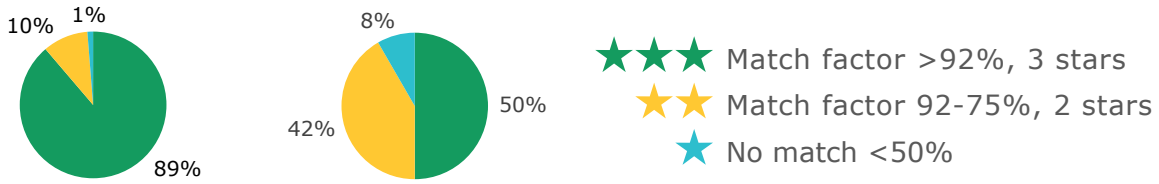

Figure 2. Match factor results of a) Sample spectra using the same solvent as the dRM was recorded (T1), b) Sample spectra using different solvents as the dRM was recorded (T2).

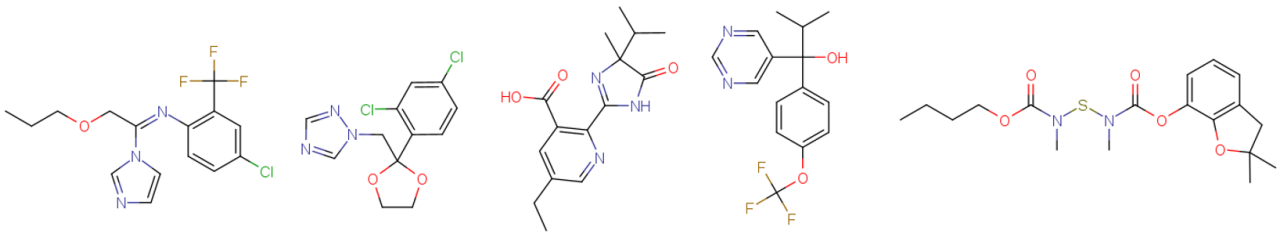

Figure 3. A selection of the molecules that were tested using the same solvent as the dRM (T1) and different solvent (T2).

A Valuable Tool for Modern Scientists

ChemisTwin™ streamlines the verification process, providing accurate results for routine structure verification workflows. Furthermore, its digital nature eliminates the need for extensive manual interpretation and literature comparison, boosting efficiency and reducing the potential for errors. The use of ChemisTwin™ eliminates the need for handling of physical reference materials, reducing the monetary and environmental cost of such analyses. Overall, ChemisTwin™ offers improved speed, accuracy, and sustainability, making it a valuable tool for routine compound identification via NMR spectroscopy.

A registered trademark of Advanced Chemistry Development, Inc. (ACD/Labs)