nature protocols

CORRECTION

https://doi.org/10.1038/s41596-020-0293-9

Addendum: A guide to small-molecule structure assignment through computation of (^{1}H and ^{13}C) NMR chemical shifts

Patrick H. Willoughby, Matthew J. Jansma and Thomas R. Hoye*

Addendum to: *Nature Protocols* https://doi.org/10.1038/nprot.2014.042, published 20 February 2014

This Addendum provides a new set of Python scripts that are compatible with current personal computer operating systems. The scripts provided in the original Protocol[1] (published in 2014) worked on all tested operating systems (Mac, Windows 7, Vista and Linux) at the time of publication, but some of these operating systems have since evolved, leading to potential sorting errors. It is recommended that the Supplementary Data 2.zip, Supplementary Data 3.zip and Supplementary Data 4.zip files provided here be used in place of the three equivalent files associated with the original Protocol.

The issue with the Python scripts was brought to our attention by the corresponding author of a study published in *Organic Letters*[2]. The authors showed that some newer personal computer operating systems may randomly sort the Gaussian optimization/frequency and NMR output files created in Step 10 of the original Protocol. Such mis-sorting would lead to inaccurate determination of conformationally averaged (i.e., Boltzmann-weighted) shielding tensors. The authors provided an improved version of the original Python script D (nmr-data_compilation.py) that addresses this glitch (see Supporting Information of the *Organic Letters* article[2]). That modified script incorporates several lines of new code to sort the files within Python as the script is being executed and then checks to ensure that the files have been properly sorted after the script has executed. These improvements effectively manage the sorting issue that is problematic in some current operating system software.

Following publication of the *Organic Letters* article, some expert programmers suggested that further revising the scripts to 'pair' (or 'match,' in computer lingo) each Gaussian optimization/frequency output file with its associated NMR output file would be even more dependable. By explicitly having the script match each pair of output files, the risk of future sorting issues could be essentially eliminated. With this in mind, we have revised the originally published Python script D to (i) pair each Gaussian optimization/frequency and NMR output files with one another, based on the matching conformation number of each, prior to determining the Boltzmann-weighted average shielding tensors; (ii) check to ensure that the conformation numbers are properly paired after the shielding tensors have been averaged; and (iii) clearly report the matched output filenames and conformation numbers in both of the newly created master .csv files.

Accordingly, we have prepared this Addendum to help ensure that the original Protocol is (and remains) compatible with all operating systems. The new Supplementary Data 2.zip contains revised versions of Python scripts A–D along with the originally published (TMS) reference, MAE, and scaling factor determination files. The new Supplementary Data 3.zip and 4.zip files contain revised versions of the master .csv files, along with the originally published Gaussian and Schrodinger (i.e., Maestro and Macromodel) files. Supplementary Data 2–4.zip files are complete and can be used as full replacements for those uploaded in the original Protocol.

Use of the new Python scripts here does not require any change in the steps described in the original Protocol. The revised scripts: (i) no longer rely on file sorting to properly pair (or 'match') the Gaussian output files; (ii) return an error alert if there happens to be improper pairing (or 'matching'); and (iii) generate a master .csv file with an updated format that allows easier inspection to ensure that proper determination of Boltzmann-weighted chemical shifts has occurred.

As software evolves, the introduction of new glitches or incompatibilities with previous code is inevitable. As such, we reemphasize our recommendation that new users benchmark the performance of the Protocol in their hands using the *cis*- and *trans*-3-methylcyclohexanol molecules, taking care to check that *all* .csv files created in Step 15c are free of errors.

1. Willoughby, P. H., Jansma, M. J. & Hoye, T. R. A guide to small-molecule structure assignment through computation of (^{1}H and ^{13}C) NMR chemical shifts. *Nat. Protoc.* **9**, 643–660, https://doi.org/10.1038/nprot.2014.042 (2014).
2. Neupane, J. B., Neupane, R. P., Luo, Y., Yoshida, W. Y., Sun, R. & Williams, P. G. Characterization of leptazolines A–D, polar oxazolines from the cyanobacterium *Leptolyngbya* sp., reveals a glitch with the "Willoughby–Hoye" scripts for calculating NMR chemical shifts. *Org. Lett.* **21**, 8449–8453, https://doi.org/10.1021/acs.orglett.9b03216 (2019).

Published online: 16 January 2020
https://doi.org/10.1038/s41596-020-0293-9