ASSIGNING STEREO AND REGIOISOMERS FROM CALCULATED CHEMICAL SHIFTS

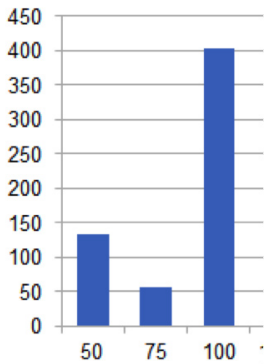
Both protocols described in a previous topic allow calculated chemical shifts to be compared with measured values in order to choose to among possible stereoisomers. The simpler protocol introduced in ***Spartan'20*** is more practical for larger molecules with multiple degrees of conformational freedom and for which more than a few isomers need to be considered. Following the work of Goodman [reference], selection is based on DP4 score, using either the Boltzmann average of chemical shifts over accessible conformers for each of the isomers (original protocol) or the conformer for each isomer with the highest DP4 score (new protocol).

Spartan'20 provides a mechanism for isomer generation, requiring the user to designate which stereocenters are to be inverted (**Generate Isomers** under the **Geometry** menu; **Chapter 19**). Duplicate stereoisomers as well as stereoisomers that differ in absolute configuration are automatically removed. Isomer identification can be done with a single job submission, requiring input of a single conformer of one of the stereoisomers, designation of rotatable single bonds and flexible rings for this isomer (which are in turn applied to all isomers) and designation of stereocenters that are to be inverted. A more cautious approach, requiring two sequential jobs may also be followed: generate a list of stereoisomers and remove any that are high-energy or otherwise "unrealistic" and run the NMR analysis on each in turn.

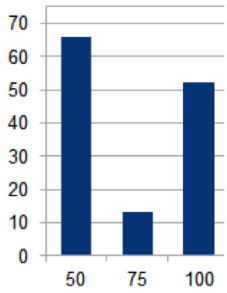
As before, assessment draws on structures from the natural products literature that have been confirmed by X-ray crystallography or by independent synthesis. This involves more than 600 molecules, each of which yields two or more distinct isomers (most commonly either four or eight isomers), leading to over 4000 distinct stereoisomers.

The histogram below documents the performance of the NMR protocol introduced in ***Spartan'20***. ~70% of the molecules have DP4 scores above 75%, suggesting that the protocol successfully identifies the isomer that is observed. On the other hand, ~20% of the molecules have DP4 scores below 50%, suggesting that protocol

is not successful. The remaining 10% of the molecules with DP4 scores between 50 and 75% may be interpreted to mean that while the calculations identify the correct isomer, one or more alternative isomers are competitive. These results are consistent with those obtained by Goodman based on a different and somewhat more computationally demanding procedure and applied to proton as well as ^{13}C chemical shifts for a different and smaller set of molecules.



There are at least two reasons for the failure of the calculations to correctly identify the isomer corresponding to the observed NMR spectrum. For one, it is likely that in some cases ^{13}C shifts for different isomers will be indistinguishable within the error limits of the calculations. Indeed, the majority of the "failures" involve change in chirality of just a single center. A second source of error may arise from the assumption that chemical shifts for the conformer that best fits the experimental ^{13}C data, as opposed to that from a properly Boltzmann-weighted average of chemical shifts over all conformers. This is of course what distinguishes the new protocol in ***Spartan'20*** from both the original protocol in ***Spartan'18*** as well as from Goodman's work. It is also directly responsible for the large saving in computation cost. While it is not practical to obtain DP4 statistics over isomers for the full set of molecules using the previous protocol, we have obtained them for the 20% of the molecules with DP4 scores less than 50%. Data for this subset are provided below as a histogram.



In only about a third of the cases does the "correct" isomer exhibit a DP4 score of 75% or higher, and half of the problem cases have DP4 scores below 50% are satisfactorily resolved. This suggests that the previous protocol which aimed to establish ^{13}C shifts that were Boltzmann-averaged over accessible conformers is unlikely to be more successful than the simplified procedure available in ***Spartan'20***.