Molecules-in-Molecules (MIM) Fragmentation-Based Quantum Chemical Predictions of NMR Chemical Shifts of Proteins and Nucleic acids

Sruthy K. Chandy, and Krishnan Raghavachari

Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States

Abstract

We have developed, implemented, and assessed efficient protocols for the prediction of NMR chemical shifts of large proteins and nucleic acids using our molecules-in-molecules (MIM) fragment-based quantum chemical approach. Fragmentation protocols involving a single layer, two layers (with different fragment sizes and levels of theory), as well as those involving extended-two layers are explored. MIM-NMR chemical shifts are evaluated using various density functional theory (DFT) methods with large basis sets, and the effect of solvent environment (i.e., implicit, explicit or implicit-explicit solvation) is also analyzed.

To assess the performance of our approach and demonstrate its applicability, MIM-NMR calculations are performed on a test set of proteins and nucleic acids, where the structure is derived from solution-phase NMR studies. A rigorous analysis of different structure preparation procedures is performed separately for proteins and nucleic acids, and unique protocols are implemented to obtain appropriate structures for accurate NMR chemical shift predictions. The effect of the solvent environment on the calculated NMR chemical shifts is incorporated through explicit, implicit or implicit-explicit solvation models. Overall, our target accuracy of ~0.3 ppm for ^{1}H and ~2-3 ppm for ^{13}C has been achieved for both large proteins and nucleic acids ranging from ~300-800 atoms. The proposed MIM-NMR method is accurate and computationally cost-effective and should be applicable to study a wide range of large biomolecules.