Inconsistencies in *ab initio* evaluations of non-additive contributions of DNA stacking energies

S1620406 Qin Ken

The main research methods on biomolecules are experiments at present. As with other materials science studies, experimental methods are difficult to describe the microscopic details of biomolecules, including molecular-level dynamics and basic functional states. For drug development, such as cancer-targeted drugs, precise matching of biological targets is required. Generally, screening of target compounds requires high-throughput experiments, but this is costly. Therefore, computer simulation plays an important role in understanding and studying the structure and biological function of biomolecules compared to traditional experimental methods. In many-body systems, the empirical force field based on pairwise method cannot consider the non-additivity beyond force field superposition. Unlike classical mechanics, the force field superposition of quantum systems cannot be added by the instantaneous polarization. Therefore, the non-additivity in the interactions is expected in inter-molecular bindings due to the induced polarization by quantum fluctuations, such as van der Waals (vdW) forces.

The description and reproduce of the bonding itself, due to intermolecular forces, is a huge challenge for *ab initio* methods. Non-additivity is a more difficult subject that has long been far from the mainstream research field and has not been well analyzed yet. The *ab initio* quantum chemistry theory is a good description of the natural stacking energy, which allows reliable energy to be found on any base structure. Calculations, in any case, need to be done at a sufficient theoretical level. For example, standard Density Functional Theory (DFT), Hartree-Fock (HF), and semi-empirical methods all fail in the description of base stacking because they cannot correctly capture the dispersion effect. Of course, the high-precision method coupled-cluster with singles, doubles, and perturbative triples with complete basis set correction (CCSD(T)/CBS) is a widely used method called "gold standard" for quantum chemistry. It is impossible to calculates large systems due to its exponential growth computational cost. Recent advances in accurate calculation methods, especially through Quantum Monte Carlo (QMC) calculations, make it possible to handle larger systems. However, some work applied to systems consisting of weakly constrained subsystems shows that non-additivity is much larger than we expected. Although there is non-additivity in larger molecules, if the non-additive contribution is positive, then there is no research significance. If so, it will only make minor corrections to the C_6 (the coefficient of $1/R^6$ deciation interactions) force without any qualitative impact.

DNA molecules are the basis of biological, genetic variation. The most typical ten kinds of B-DNA molecular structures are composed of two kinds of purines and pyrimidines. For DNA molecules, the number of atoms is around 10, and the weak force between many-body molecules itself is the challenge in the quantum chemistry. We evaluated the non-additive contribution of the inter-molecular interactions in B-DNA stacking by using Fixed-Node Diffusion Monte Carlo (FNDMC) methods. In the previous calculation methods, the sign of the non-additive contribution is positive and tiny. Recent studies have shown that in the calculation using FNDMC method, and negative values appear in the results of non-additive contributions. First of all, the evaluation result in dispersion interaction by the standard SCF methods was proved failed due to the lack of dispersion term. Second, even the CCSD(T) method still evaluates dispersion interaction as the SCF-level, because of the practical handling of CBS at the feasible level with second order Møller–Plesset perturbation theory (MP2). Finally, although the calculation is trustworthy for the DMC method itself, the FN method cannot be ruled out to cancel the approximation error when the sign problem occurs, because we still divide the system according to the H-bond in the non-additive evaluation. While the SCF-level non-additivity is mostly positive, the non-additive contributions described by FNDMC are both positive and negative signs. The negative sign is found to be reasonable, which might be supported by a simple model analysis based on the London theory. It would, however, be premature to draw a conclusion that the FNDMC non-additivity reveals the truth. This is because the Watson-Crick base-pair involves the charge transfer caused by the H-bonds, but we could not verify if the error cancellations of the FN errors were successful for the H-bonds, as in the case of complexation energies. However, we cannot ignore the FN approximation which is used to solve the QMC sign problem. We analyzed the possible errors in this approach and discussed it.

Keywords: B-DNA, Stacking energy, Non-additivity, Quantum Monte Carlo, *ab initio* methods