Communication: Accurate description of interaction energies and three-body effects in weakly bound molecular complexes by PBE-QIDH models

A. J. Pérez-Jiménez,1,a) E. Brémont,2 C. Adamo,3,4 and J. C. Sancho-García1,b)

1Department of Physical Chemistry, University of Alicante, E-03080 Alicante, Spain
2Université Paris Diderot, Sorbonne Paris Cité, ITODEM, UMR CNRS 7086, F-75013 Paris, France
3Chimie ParisTech, PSL Research University, CNRS, Institut de Recherche de Chimie Paris, F-75005 Paris, France
4Institut Universitaire de France, F-75005 Paris, France

(Received 30 May 2018; accepted 12 July 2018; published online 25 July 2018)

We apply a recently developed parameter-free double-hybrid density functional belonging to the quadratic-integrand double-hybrid model to calculate association energies (ΔE) and three-body effects (Δ³E) arising from intermolecular interactions in weakly bound supramolecular complexes (i.e., the dataset 3B-69). The model behaves very accurately for trimer association energies and is found to outperform widely used density functional approximations while approaching the accuracy of more costly ab initio methods for three-body effects. The results are further improved when we add some specific corrections for the remaining dispersion interactions, D3(BJ) or VV10 for two-body effects and Axilrod-Teller-Muto for three-body effects, leading to marginal deviations (less than 1 kcal/mol for ΔE and around 0.03–0.04 kcal/mol for Δ³E) with respect to benchmark results. Published by AIP Publishing. https://doi.org/10.1063/1.5042153

The interaction energy of a nanoaggregate of N rigidly interacting molecules is calculated as

$$\Delta E = E(\ldots XYZ) - E(Z) - E(Y) - E(X) - \cdots \quad (1)$$

where $E(M)$, $M = \ldots$, $X, Y, Z$, is the energy of individual molecules (i.e., monomers) and $E(\ldots XYZ)$ is the energy of the weakly bound nanoaggregate. Using the many-body decomposition approach, valid not only in terms of computational efficiency but also useful for providing physical insights, the energy can be decomposed as follows:

$$\Delta E = \sum_{X<Y} \Delta^2 E(XY) + \sum_{X<Y<Z} \Delta^3 E(XYZ) + \cdots \quad (2)$$

with $\Delta^n$ being the $n$th order term. The high-order terms, $n \geq 4$, are often neglected due to the fast convergence of the above expansion. The explicit form of the two-body [$\Delta^2 E(XY)$] and three-body [$\Delta^3 E(XYZ)$] terms is

$$\Delta^2 E(XY) = E(XY) - E(Y) - E(X), \quad (3)$$

$$\Delta^3 E(XYZ) = E(XYZ) - \Delta^2 E(XY) - \Delta^2 E(XZ) - \Delta^2 E(YZ)$$

$$- E(Z) - E(Y) - E(X)$$

$$= E(XYZ) - E(XY) - E(XZ) - E(YZ)$$

$$+ E(Z) + E(Y) + E(X), \quad (4)$$

after replacing in the latter expression the corresponding two-body specific terms. The energy of weakly bound trimers, $E(XYZ)$, and dimers, e.g., $E(XY)$, is needed, as well as those of monomers, e.g., $E(X)$, for calculating the three-body $\Delta^3 E(XYZ)$ term. Few datasets were developed to probe these properties, and, for that purpose, we choose the state-of-the-art 3B-69 dataset, which is composed of non-covalently bound trimer geometries (×69) extracted from the reported crystalline structures of 23 compounds. The dataset comprises a variety of compounds (see the supplementary material for further details) and packing motifs, which in fact translates into intermolecular interactions of different nature (e.g., hydrogen bonds, polarization, and/or dispersion interactions) depending on each particular molecule and supramolecular configuration. The moderate size of the compounds allowed previous computations at the counterpoise-corrected coupled-cluster single double triple [CCSD(T)] complete basis set (CBS) level, and it becomes thus possible to benchmark any theoretical method and/or computational approach for three-body effects, whose examples are still scarce. Previously applied density functional approaches are indeed reported to perform poorly, possibly due to the incomplete treatment of exchange and polarization effects even when the expressions were corrected for including dispersion energies approximately by resorting to interatomic pairwise potentials.

In this context, we will thus assess here the accuracy of the parameter-free Quadratic Integrand Double-Hybrid (QIDH) density functional, PBE-QIDH, and its recently developed dispersion-corrected extensions, PBE-QIDH-D3(BJ) and PBE-QIDH-VV10. This method linearly combines parameter-free exchange and correlation density functionals, $E_x[\rho]$ and $E_c[\rho]$, respectively, with a portion of EXact-like $e$Xchange (EXX), $E_x^{EXX}[\phi]$, and second-order Perturbation Theory (PT2), $E_x^{PT2}[\phi, \phi']$, but with the
derived before\(^5\) when applied out of the training (e.g., S130) set.\(^1\)

Having obtained these encouraging results for trimer association energies, and before calculating three-body effects\(^1\) (i.e., \(\Delta^3E\)), we bracket the accuracy of the PBE-QIDH model with respect to both \textit{ab initio} and density functional methods by introducing the three-body (and geometry-based) Axilrod-Teller-Muto (ATM) correction\(^1\) given in the form

\[
E_{\text{ATM}} = \sum_{\text{atom triples}} C_{ABC} \left( 3 \cos \theta_{AB} \cos \theta_{BC} \cos \theta_{AC} + 1 \right) \frac{R_{AB} R_{BC} R_{AC}}{(R_{AB} R_{BC} R_{AC})^3} f_n(R_{ABC}),
\]

depending on the \(R_{AB} - R_{BC} - R_{AC} \neq 0\) interatomic distances (angles), \(R_{ABC}\) is the geometric mean of the former, \(C_{ABC} = \sqrt[3]{C_6 C_6 C_6}\), and \(f_n\) is another damping function. This expression was previously applied to large supramolecular complexes,\(^1\) where its influence is expected to contribute the most to association energies due to extended polarization effects, and is known to provide a very satisfactory agreement for bulk three-body molecular effects with respect to CCSD(T) results for systems of moderate size as those contained in the 3B-69 dataset.\(^1\)

We now apply the whole set of dispersion corrections to the calculation of three-body effects \(\Delta^3E\), which would also allow us to compare with previous estimates in the literature. Table II (see also Table S2 of the supplementary material for specific values) gathers the statistical error values (MSE, MUE, and RMSE) for different methods as a function of their scaling.

Table II. Statistical errors (kcal/mol) of QIDH-based models, as compared with previous estimates, for the calculation of \(\Delta^3E\) three-body effects.

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE</th>
<th>MUE</th>
<th>RMSE</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D3(BJ)+ATM(^4)</td>
<td>0.093</td>
<td>0.122</td>
<td>0.147</td>
<td>(O(N^3))</td>
</tr>
<tr>
<td>BLYP-D3(BJ)+ATM(^4)</td>
<td>−0.084</td>
<td>0.086</td>
<td>0.111</td>
<td>(O(N^3))</td>
</tr>
<tr>
<td>PBE0-D3(BJ)+ATM(^4)</td>
<td>0.043</td>
<td>0.068</td>
<td>0.080</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>B3LYP-D3(BJ)+ATM(^6)</td>
<td>−0.046</td>
<td>0.052</td>
<td>0.069</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>BH-LYP-D3(BJ)+ATM(^6)</td>
<td>−0.018</td>
<td>0.033</td>
<td>0.045</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>PBE-QIDH</td>
<td>−0.022</td>
<td>0.035</td>
<td>0.047</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>PBE-QIDH-D3(BJ)</td>
<td>−0.022</td>
<td>0.034</td>
<td>0.047</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>PBE-QIDH-D3(BJ)+ATM</td>
<td>0.002</td>
<td>0.026</td>
<td>0.036</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>PBE-QIDH-VV10</td>
<td>−0.032</td>
<td>0.041</td>
<td>0.054</td>
<td>(O(N^5))</td>
</tr>
<tr>
<td>B2-PLYP-D3(BJ)+ATM(^6)</td>
<td>−0.028</td>
<td>0.038</td>
<td>0.050</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>MP2(^a)</td>
<td>−0.039</td>
<td>0.045</td>
<td>0.059</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>MP3(^a)</td>
<td>0.022</td>
<td>0.026</td>
<td>0.035</td>
<td>(O(N^4))</td>
</tr>
<tr>
<td>CCSD(^d)</td>
<td>0.010</td>
<td>0.014</td>
<td>0.019</td>
<td>(O(N^3))</td>
</tr>
<tr>
<td>SCS-CCSD(^d)</td>
<td>−0.001</td>
<td>0.010</td>
<td>0.015</td>
<td>(O(N^3))</td>
</tr>
<tr>
<td>MP2.5(^d)</td>
<td>−0.009</td>
<td>0.014</td>
<td>0.019</td>
<td>(O(N^3))</td>
</tr>
</tbody>
</table>

\(^\text{a}\)Values taken from Ref. 1.
extract the following general conclusions: (i) going across the approximate hierarchy of density functionals (i.e., semilocal, hybrid, and double-hybrid) systematically reduces the errors, e.g., compare PBE, PBE0, and PBE-QIDH, or BLYP, B3LYP, and B2-PLYP results, as well as it does to include a larger value of \( \lambda_1 \) in hybrid-only methods, e.g., compare B3LYP and BH-LYP values, which is related to the self-interaction \( \Delta \) effect.\(^{16}\) (ii) without including any correction for dispersion, the pristine PBE-QIDH model becomes the best density functional of all the set (PBE, PBE0, BLYP, B3LYP, BH-LYP, CAM-B3LYP, TPSS, M06-2X, M06-HF, B2-PLYP) tested up to now,\(^1\) even when previous results from the literature explicitly incorporate two- and three-body dispersion corrections and/or some functionals (e.g., M06-2X) are obtained including weak interactions into their training set; and (iii) the role of the three-body ATM correction reduces even further the error obtained with the PBE-QIDH-D3(BJ) model, achieving an accuracy competitive with methods such as MP2.5 or (SCS-)CCSD which naturally incorporates three-body interactions arising from correlation effects. Note, however, that we neglect at this level other higher-order effects (i.e., fourth-order and beyond) and many-body non-additivity of dispersion energies.\(^{17}\) Figure S3 of the supplementary material shows the impact of the ATM correction for the \( \Delta E \) values, ranging in all cases between 95% and 105% and thus indicating its predominant role here. We also note the effectiveness of the MP2+ATM coupling for \( \Delta^3 E \) estimates,\(^2\) which also retains the accuracy of MP2.5 or (SCS-)CCSD methods.

In summary, we show the high quality of the PBE-QIDH functional for describing interelectronic effects of all types arising from weakly bound interactions. The method performs well across the whole 3B-69 benchmark set, for low or highly polarizable systems, and prompts to reconsider previous findings about density functional approaches and their poor performance for three-body effects, where exchange and polarization contributions are of the same or higher importance than dispersion contributions.

See supplementary material for (i) the structures of all trimers belonging to the 3B-69 dataset; (ii) the computational details; (iii) all the \( \Delta E \) individual values for the 3B-69 dataset at the PBE-QIDH, PBE-QIDH-D3(BJ), PBE-QIDH-VV10, and PBE-QIDH-D3(BJ)+ATM levels; (iv) the contribution of full and ATM dispersion energy to the final \( \Delta E \) values at the PBE-QIDH-D3(BJ)+ATM level; (v) all the \( \Delta^3 E \) individual values for the 3B-69 dataset at the PBE-QIDH, PBE-QIDH-D3(BJ), PBE-QIDH-VV10, and PBE-QIDH-D3(BJ)+ATM levels; and (vi) the contribution of the ATM dispersion correction to the final \( \Delta^3 E \) values at the PBE-QIDH-D3(BJ)+ATM level.

We acknowledge the “Ministerio de Economía y Competitividad” of Spain and the “European Regional Development Fund” through the Project No. CTQ2014-55073-P.