How Accurate are Approximate Quantum Chemical Methods at Modelling Solute-Solvent Interactions in Solvated Clusters?

Junbo Chen,[a] Bun Chan,*[b] Yihan Shao[c] and Junming Ho*[a]

[a] School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia

[b] Graduate School of Engineering, Nagasaki University, Bunkyo-machi 1-14, Nagasaki 852-8521, Japan

[c] Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, Unites States

Abstract

In this paper, the performance of a wide range of DFT methods is assessed for the calculation of interaction energies of thermal clusters of a solute in water. Three different charge states (neutral, proton transfer transition state and zwitterion) of glycine were solvated by 1 to 40 water molecules as sampled from molecular dynamics simulations. While some *ab initio* composite methods that employ insufficiently large basis sets incurred significant errors even for a cluster containing only 5 water molecules relative to the W1X-2 benchmark, the DLPNO-CCSD(T)/CBS and DSD-PBEP86 (triple zeta basis set) levels of theory predicted very accurate interaction energies. These levels of theory were used to benchmark the performance of 16 density functionals from different rungs of Jacob's Ladder. Of the Rung 4 functionals examined, the ωB97M-V and ωB97X-V functionals stood out for predicting absolute interaction energies in 40-water clusters with mean absolute deviations (MAD) ~ 4 kJ mol^{-1}. The B3LYP-D3(BJ) functional performed exceptionally well with a MAD ~ 1.7 kJ mol^{-1} and is the overall best performing method. Calculations of relative interaction energies allow for cancellation of

systematic errors, including basis set truncation and superposition errors, and the ωB97M-V and B3LYP-D3(BJ) double zeta basis set calculations yielded relative interaction energies that are within ~ 3 kJ mol^{-1} of the benchmark. The ONIOM approximation provides another strategy for accelerating the calculation of accurate absolute interaction energies provided that the calculations have converged with respect to the size of the "high-level-layer".

*Correspondence author email: bun.chan@nagasaki-u.ac.jp; junming.ho@unsw.edu.au

Introduction

Many chemical and biological reactions occur in the solution-phase, and the solvent can often influence reaction rates, conversion efficiency and product selectivity in a variety of ways. For example, it may directly participate in the reaction (*cf.* ligand in transition metal complexes), change the relative stabilisation of the reactant, transition state or product or even alter the solubility of different components of the reaction.[1] As such, our ability to accurately predict the rate and equilibrium constants of reactions depends on how well solute-solvent interactions are described by our computational models.

The development of quantum mechanical (QM) implicit solvent models (also known as continuum solvent models)[2,3] has facilitated the study of many solution phase processes due to their ease of use, low computational cost, and moderate accuracy. Typically, these models contain parameters that have been optimised at modest levels of theory (*e.g.* DFT and double ζ basis set) to reproduce experimental data. In general, implicit solvent models are computationally very efficient, and are able to predict the free energies of solvation of typical neutral solutes to within ~ 5 kJ mol^{-1} of experiment. However, these models can easily incur errors in excess of 20 kJ mol^{-1} for charged solutes, and they are less amenable to systematic improvement.[4–10] While various hybrid implicit-explicit schemes have proposed,[10–14] *e.g.* explicit account for a few solvent molecules in the first solvation shell, this introduces other issues related to configurational sampling and the harmonic treatment of floppy vibrational modes. In this regard, explicit solvent models are considered more robust alternatives to implicit models because the solute and solvent degrees of freedom and their interactions are modelled explicitly. Accordingly, the accuracy of these models may be systematically improved through the use of increasingly accurate intermolecular potentials, *e.g.* fixed charge

force fields, polarisable force fields and hybrid quantum mechanics/molecular mechanics (QM/MM) potentials. [15–17]

There has been some success in the application of explicit solvent models to predict organic and inorganic reaction mechanisms,[18–20] electrochemistry,[21–27] and pK_a values.[28–31] However, given the size of these systems and the need for extensive configurational sampling, explicit solvent simulations are generally carried out using classical MM force fields or hybrid QM/MM potentials. An intriguing observation from these studies is that use of supposedly more accurate potentials does not necessarily translate to gains in accuracy. For example, Mei and co-workers have found that the use of QM/MM potentials did not improve the results of transfer free energies compared to the MM simulations.[32] Similarly, York and co-workers have also reported hydration free energies determined using DFT/MM methods were generally inferior to classical force field simulations.[33] This raises the question of whether the use of more robust DFT/MM potentials can indeed better describe solute-solvent interactions than classical force fields, and whether explicit solvent models can be systematically improved.

Recently, our group has applied an alternative approach that involves a separate treatment of intra- and intermolecular contributions using a thermodynamic cycle.[8] Specifically, in this approach, the gas phase free energy (ΔG_{gas}) is obtained using very high levels of theory (*e.g.* CCSD(T)/CBS) and the solvation contribution ($\Delta\Delta G_S$) to the solution phase free energy change is obtained using free energy perturbation (FEP) in conjunction with explicit solvent molecular mechanics molecular dynamics (MD) simulations (Figure 1A). The solvation contribution may subsequently be improved using end-state corrections (ΔG(MM→QM)) to mitigate the inaccuracies associated with the force field description of intermolecular interactions (Figure 1B). These end-state corrections may be obtained from FEP calculations using MM and QM solute-solvent interaction energies determined on clusters containing a sufficient number of

solvent molecules to mimic bulk solvent effects. Such a "reference potential" type approach [33–41] represents a cost-effective alternative to "one-pot" *ab initio* MD simulations, which treat both the intra- and intermolecular degrees of freedom at the same level of theory (usually DFT) and are exceedingly expensive.

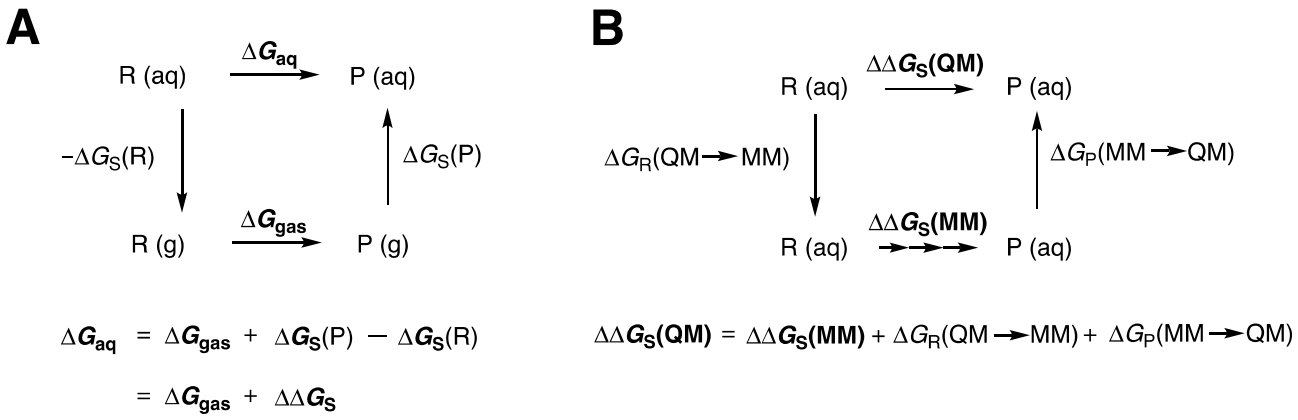

Figure 1. (A) Thermodynamic cycle for the calculation of solution phase free energy changes. (B) Thermodynamic cycle for the calculation of MM to QM end-state corrections (ΔG(MM→QM)) to the solvation contribution to the free energy change. ΔG_S, ΔG_{gas} and ΔG_{aq} refer to the free energy of solvation, gas phase and aqueous phase reaction free energy changes while R and P denote reactant and product respectively.

While our recent work demonstrated the effectiveness of this approach in the treatment of an S_N2 reaction in water, there are several issues concerning the broader applicability of this procedure. Specifically, the reliability of the end-state corrections in Figure 1B, (ΔG_R(QM→MM) and ΔG_P(MM→QM)), is rested on the assumption that the MM thermal ensembles at the reactant and product are good approximations to the ensembles that would have been obtained using accurate quantum chemistry models. This will likely depend on the nature of the system and also the choice of MM force field.[8] There are also two concerns on these end-state corrections, firstly related to their cost as they entail a myriad of QM calculations of relatively large solute-solvent clusters, and secondly what QM level of theory

to use for these calculations. The first concern has been addressed in a recent study where we demonstrated the use of the systematic molecular fragmentation approach to accelerate these QM calculations by 10-15 times while retaining a mean accuracy of ~ 5 kJ mol^{-1}.[42]

The second question is the subject of this paper, *i.e.* how accurate are approximate QM methods at describing solute-solvent interactions, particularly in thermal clusters obtained from MD simulations? Whilst there are many valuable benchmarking studies[43–49] that assessed the performance of DFT and wavefunction methods for the calculation of non-covalent interactions, the vast majority focused on zero-Kelvin structures of small dimers such as those in the S22[50] and S66[51] datasets. Recently, Martin and co-workers have benchmarked the performance of DFT and explicitly correlated *ab initio* methods against the S66x8 dataset which considers the dimers in the S66 dataset at non-equilibrium separations. The authors recommended the DSD-PBEP86, ωB97X-V, B3LYP-D3(BJ) and BLYP-D3(BJ), which performed very well with RMSD values smaller than 1 kJ mol^{-1} relative to explicitly correlated MP2 and coupled cluster methods.[46] Other datasets that contains non-equilibrium geometries include the S668a[52], which systematically scanned interfragment angles, NBC10 and HBC6, which provided frozen monomers and relaxed potential energy curves,[53,54] ACHC,[55] which examined nucleobase pair degrees of freedom and BFDb,[56] which contains numerous protein sidechain-sidechain and backbone-backbone interactions.

In another recent study, Mardirossian and Head-Gordon compiled the MGCDB84 dataset which contains *inter alia* non-covalent interaction data for 1835 dimers and 243 small water clusters of neutral and charged solutes.[45] Of the 200+ DFT methods examined in that study, the authors showed that the range-separated hybrid ωB97M-V[57] functional yielded the lowest RMSD of 1 and 2 kJ mol^{-1} for the dimer and cluster datasets respectively.[45] However, due to small size of these clusters (typically 6 or fewer water molecules) and limited variation of the

solute (F^-, Cl^- and SO_4^{2-}), it is unclear if the ωB97M-V and other similar performing functionals would perform as well in the prediction of interaction energies involving larger clusters and different solutes. With many-body van der Waals and induction effects likely to contribute more significantly to the solute-solvent interactions in larger clusters, it is of interest to examine how this affects the performance of functionals that rely on a pairwise description of dispersion.

To this end, this paper reports an extensive benchmarking study of density functionals from various rungs of Jacob's Ladder as well as approximate wavefunction methods such as DLPNO-CCSD(T) and composite methods to determine the absolute and relative glycine-water interaction energies in clusters containing up to 40 water molecules (Figure 2). The choice of system is motivated by not only the fact that it has been studied by many different groups[47,58–67] but also because it allows a systematic variation of the solute, *i.e.* neutral and zwitterionic glycine, and the proton-transfer transition state with different extent of charge separation. This will enable us to probe the performance of different theoretical methods as the magnitude and nature of the solute-solvent interactions are altered.

This paper is laid out as follows: We first sought to identify levels of theory that balance cost and accuracy based on calculations on smaller-sized clusters. These levels of theory are then used to benchmark a variety of DFT functionals against the calculation of absolute and relative interaction energies of the three states of glycine in water clusters containing up to 40 waters. Finally, we examined how the ONIOM model[68] may be used to accelerate the calculation of these interaction energies.

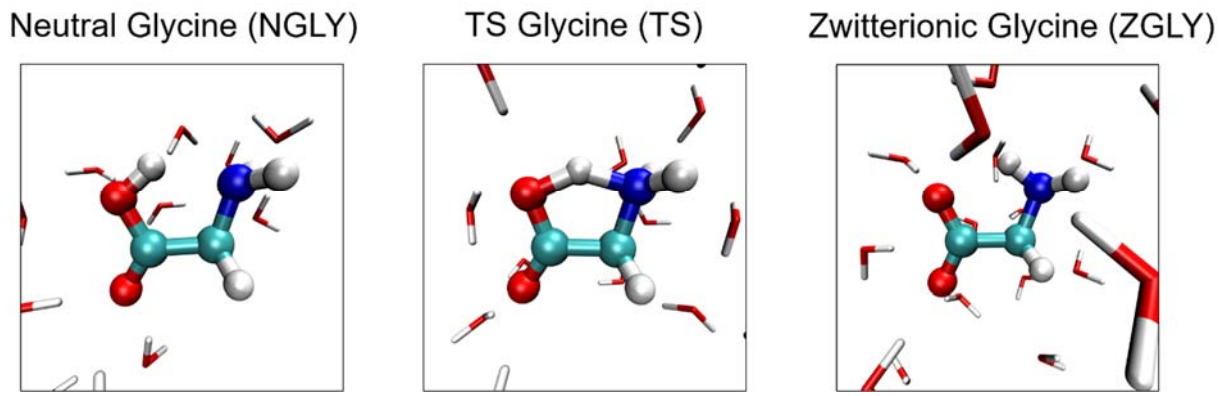

Figure 2. The three states of glycine considered in this study.

Computational Details

All electronic structure calculations were performed using the Gaussian16,[69] Molpro,[70,71] Q-CHEM (version 5.0)[72] and ORCA(version 4.1.2)[73,74] programs. All molecular dynamics (MD) simulations were performed using the CHARMM[75,76] force field and the NAMD[77] program. The geometries of glycine (neutral, zwitterionic and transition state) were calculated using the SMD[78] implicit solvent model (aqueous solvent) at the M06-2X/6-31+G(d,p) level of theory. The solute geometries were then separately embedded in a 32 × 32 × 32 Å periodic box containing 1091 TIP3P water molecules to produce the density of bulk water. During the MD simulation, the glycine solute geometry is held at a fixed geometry at the centre of the water box, and the solute-water clusters of varying sizes containing the nearest 1, 3, 5, 10, 20 and 40 water molecules were extracted from the MD trajectory. The nearest water molecules were determined by computing all pairwise distance between the atoms in the solute and the atoms in a chosen TIP3P water. The shortest atom-atom distance is recorded for that water, and the process is repeated for all water molecules within a specified cut-off radius. The distances are then ranked to determine the nearest *n* water molecules. Following a 1 ns equilibration run, a total of 20 snapshots were collected from a production run of 1 ns (spaced by 50 ps) under NPT

ensemble (25 °C and 1 atm), and the mean absolute deviations (MAD) in the calculated interaction energies were obtained from these snapshots.

The absolute interaction energies were obtained using the standard expression in Equation 1 without thermal or zero-point vibrational corrections. The number of water molecules in the cluster is denoted by n, and E_{gly} , $E_{water}(n)$, $E_{cplx}(n)$ are the gas-phase electronic energies of the solute, the water cluster, and the entire glycine-water complex, respectively.

$$\Delta E_{int}(n) = E_{cplx}(n) - E_{water}(n) - E_{gly} \qquad (1)$$

Single point calculations were performed using various levels of theory. The composite methods WGh,[79] G4(MP2)-6X,[80] CBS-QB3[81] and W1X-2[82] were calculated using the Gaussian 16 or Molpro packages. The DLPNO-CCSD(T) calculations were performed using "TightPNO" setting in ORCA, and a two-point complete basis set (CBS) extrapolation was performed based on the aug-cc-pVXZ (X=D,T) and cc-pVXZ (X=D,T) basis sets using optimised α/β values of 4.3/2.51 and 4.42/2.46 respectively.[83] These are denoted by DLPNO-CCSD(T)/aug-CBS and DLPNO-CCSD(T)/CBS respectively.

The DFT calculations were carried out using Q-Chem and ORCA and are listed according to the rungs in Perdew's DFT Jacob's Ladder:[84]

(a) Rung 5 – Double Hybrid (DH) GGA:

DuT-D3,[85] DSD-PBEP86[86,87] and DSD-PBEPBE[87]

(b) Rung 4 –Hybrid GGA/meta-GGA:

M11,[88] ωB97M-V,[57] ωB97X-V,[89] ωB97X-D,[90] ωB97X[91], M06-2X,[92] MN15,[93] PW6B95-D2,[94,95] B3LYP-D3(BJ),[96–98] B3LYP,[96,97] B97-3-D2,[95,99] B97-1[100]

(c) Rung 3 –meta-GGA:

B97M-rV,[101] mBEEF[102]

(d) Rung 2 – Generalised Gradient Approximations (GGA):

B97-D3(BJ),[97,98] GAM[103]

All DFT calculations were calculated using default settings in the respective packages. Because different programs have different exchange-correlation (XC) quadrature grids, we also verified that the effect of XC grids is less than 1 kJ mol^{-1} in terms of interaction energies. The aug'-cc-pVTZ basis sets were used for all DFT calculations. The aug'-cc-pVXZ (X=D,T) is a modified version of aug-cc-pVXZ which uses cc-pVXZ for H atom and aug-cc-pVXZ for all other elements). Semi-empirical PM7[104] calculations were carried using Gaussian16 while the tight-binding density functional with 3rd order correction (DFTB3)[105,106] was carried using the DFTB+ package.[107] The DFTB3 calculations employed the 3OB parameters and include dispersion corrections based on the Slater-Kirkwood (DFTB3-SK) polarisable atomic model[108] or Grimme's D3 with Becke-Johnson damping.[98] Energy decomposition analysis was performed using the ALMO-EDA2 method[109,109,110] as implemented in Q-CHEM.

Results and discussion

Choice of basis set and basis set superposition error (BSSE). We first considered the choice of basis sets for the DFT calculations with the view to minimising basis set superposition error (BSSE), basis set incompleteness error (BSIE) and computational cost. At the complete basis set (CBS) limit, one expects both BSSE and BSIE to approach zero so minimising BSSE is an approximate way of determining the best CBS estimate. As shown in Figure 3a, we calculated the BSSE in the interaction energy for a glycine-$(H_2O)_{10}$ cluster using the aug-cc-pVDZ and aug-cc-pVTZ and the corresponding basis sets without diffuse functions on the hydrogen atoms (denoted aug'-cc-pVDZ and aug'-cc-pVTZ), as well as several other triple and quadruple zeta basis sets. The BSSE were determined using the counterpoise method[111] and ranged from ~ 1 to 18 kJ mol^{-1}. The smallest BSSE obtained using the pc-3 and quadruple-zeta basis sets is ~ 1-2 kJ mol^{-1}, while the use of the triple-zeta basis sets incurred a slightly higher BSSE ~ 3-4 kJ mol^{-1}. Interestingly, the larger aug-cc-pVTZ basis set has a slightly higher BSSE compared to the smaller aug'-cc-pVTZ counterpart indicating that a smaller BSSE may not necessarily imply a smaller BSIE.

Figure 3b illustrates how the BSSE varies with the size of the cluster for ZGLY. As shown, BSSE increases monotonically as the number of water molecules is doubled though the change is generally minimal (~ 1 kJ mol^{-1}) for the triple-zeta basis set (see Table S1). While the use of augmented, quadruple-zeta basis set (*e.g.* def2-QZVPPD) is recommended to minimise BSSE,[57] this adds considerably to the cost of the calculations (~ 6 fold increase in wall time; Figure 3a) but with only a moderate reduction in BSSE by ~ 2 kJ mol^{-1} (See also Tables S2-4). This is consistent with several recent studies which found that the difference in RMSD between triple-and quadruple zeta basis sets for the calculation of non-covalent interactions is typically < 1 kJ mol^{-1}.[43,45]

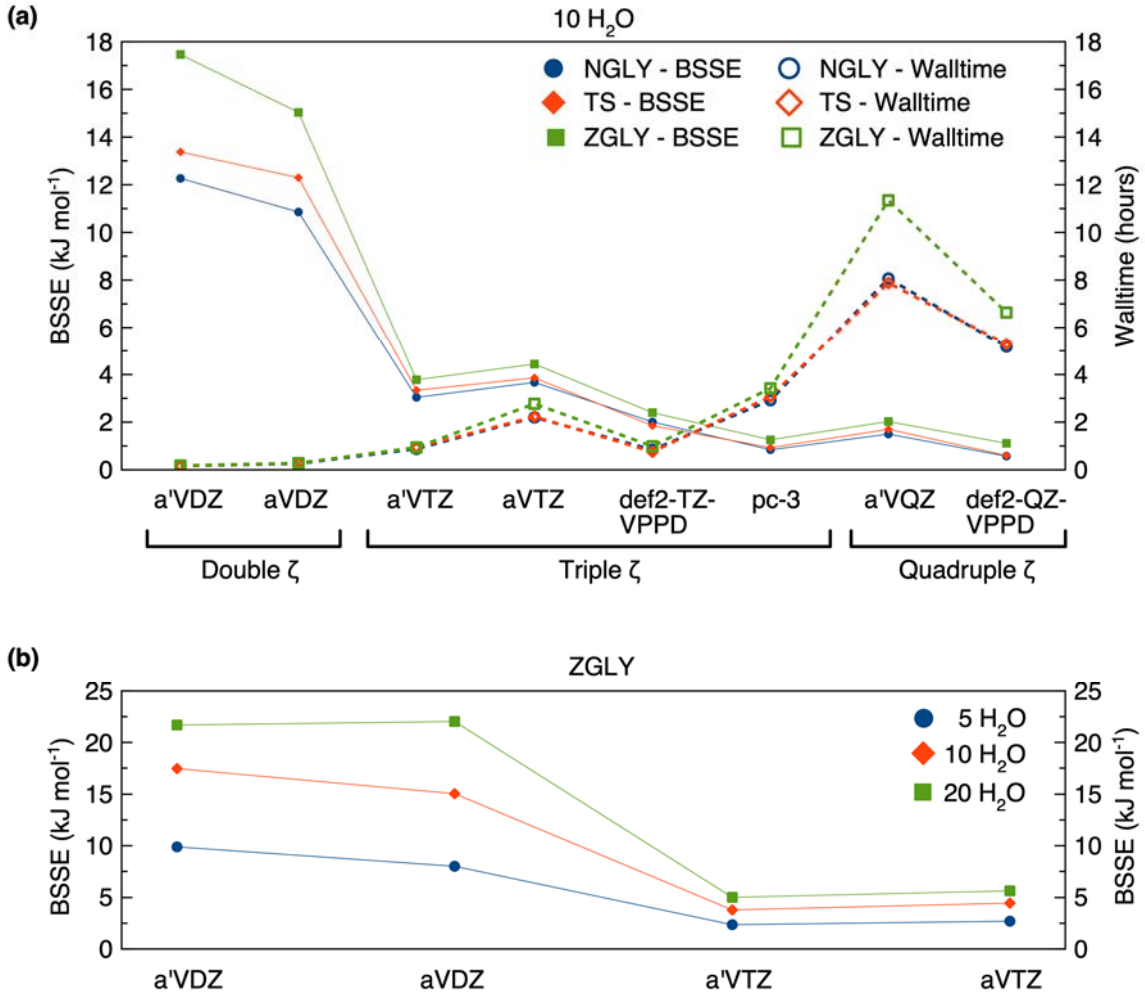

Figure 3. Comparison of basis set superposition error (BSSE) using M06-2X in conjunction with various basis sets for (a) 10 H_2O clusters with three states of solute (NGLY, TS, ZGLY) and respective timing for BSSE calculations, and (b) ZGLY with three cluster sizes (5, 10, 20 H_2O). Data shown is obtained from the average over three frames.

Further, we note that there is little difference in BSSE between the aug'-cc-pVTZ and aug-cc-pVTZ basis sets and this is true regardless of the charge state of the solute and the cluster size (Figure 3). This is encouraging because the aug'-cc-pVTZ contains 765 fewer basis functions compared to the aug-cc-pVTZ basis set for the largest 40-water cluster and is approximately 5 times faster (based on M06-2X calculations with Q-Chem version 5.2 using 8 cores Intel Xeon E5-2670 (Sandy Bridge) 2.6GHz). Accordingly, it appears that the aug'-cc-pVTZ basis set

represents a reasonable compromise in terms of cost and accuracy. All subsequent DFT calculations presented in this work are based on the aug'-cc-pVTZ basis set, unless stated otherwise. As shown in Tables S5 and S6, for a given basis set and cluster size, the combined basis set error (BSSE + BSIE) is also relatively insensitive to the choice of DFT functional and will not affect the comparison between the different DFT functionals and the DSD-PBEP86 benchmark.

For the DLPNO-CCSD(T) complete basis set extrapolated energies, it was not possible to quantify BSSE, but it is interesting to note that these calculations are more sensitive to the inclusion of diffuse functions in the basis set. As shown in Figure 4, the CBS extrapolated interaction energies based on the cc-pVXZ and aug'-cc-pVXZ basis sets are significantly different from those obtained using the aug-cc-pVXZ basis set (X=D,T). In particular, the interaction energies can differ by as much as 15 kJ mol^{-1}, and this is consistent for all charge states of the solute. For this reason, all DLPNO-CCSD(T) energies reported herein are based on the aug-cc-pVXZ (X=D,T) basis sets, *i.e.* DLPNO-CCSD(T)/aug-CBS.

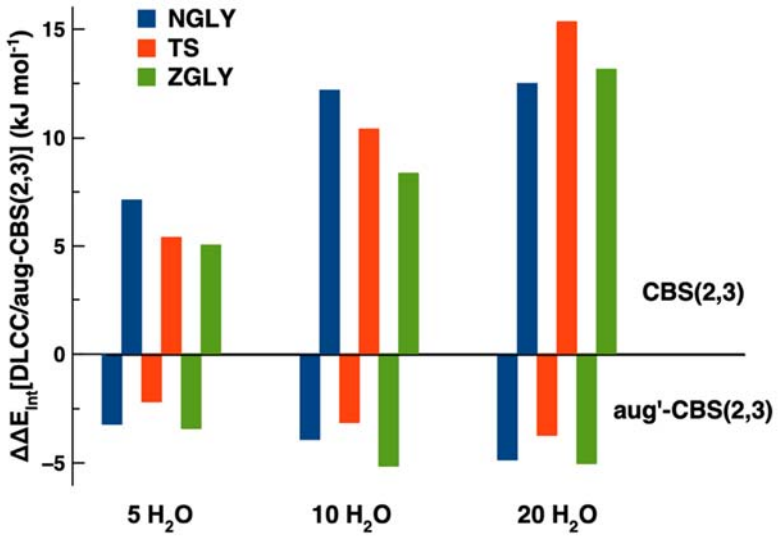

Figure 4. The difference in DLPNO-CCSD(T) interaction energies comparing aug'-CBS(2,3) and CBS(2,3) against aug-CBS(2,3), for the three states of solute and a randomly selected configuration of an increasing size of water clusters (5,10, 20 H_2O).

Choice of the benchmark. We sought to identify cost-effective methods using smaller clusters that may be used to benchmark DFT methods for larger water clusters. For this purpose, we considered the DLPNO-CCSD(T)/CBS level of theory as well as the WGh, G4(MP2)-6X and CBS-QB3 composite methods and three double-hybrid DFT methods DuT-D3, DSD-PBEP86 and DSD-PBEPBE. These methods were benchmarked against the W1X-2 composite method, an economical approximation to canonical CCSD(T)/CBS, for 1, 3 and 5 water clusters. Briefly, the W1X-2 procedure is an approximation to W1w.[112] It requires only two CCSD-F12b calculations (with the aug'-cc-pVDZ and aug'-cc-pVTZ basis sets) for the evaluation of valence CCSD(T)/CBS term, and a nearly cost-free MP2 core-correlation term. These features result in an order of magnitude reduction in computational cost (relative to W1w). The W1X-2 method has been evaluated against a diverse range of thermochemical quantities including the E2 set of 526 relative energies, the BDE261 set of 261 dissociation energies, barriers for pericyclic reactions, and heats of formation of medium-sized molecules with mean absolute deviation (MAD) below 4 kJ mol^{-1}.

Figure 5 shows their mean absolute deviation (MAD) calculated over 20 frames for ZGLY. As shown, the errors associated with all methods increase with the size of the cluster and that the errors exceed 5 kJ mol^{-1} for G4(MP2)-6X, CBS-QB3, DuT3 and DSD-PBEPBE for the 5-water cluster. Notably, the G4(MP2)-6X composite method performed poorly presumably due to the absence of diffuse basis functions in the CCSD(T) component of this method.[113] Similar behaviour was observed for the calculation of interaction energies for NGLY and TS (See Figure S1 and S2). On the other hand, WGh, DSD-PBEP86 and DLPNO-CCSD(T)/aug-CBS

errors are consistently within 4 kJ mol^{-1}. For this reason, the DSD-PBEP86 and DLPNO-CCSD(T)/aug-CBS levels of theory were considered to be our benchmark for the study of larger water clusters because of their relatively moderate computational cost.

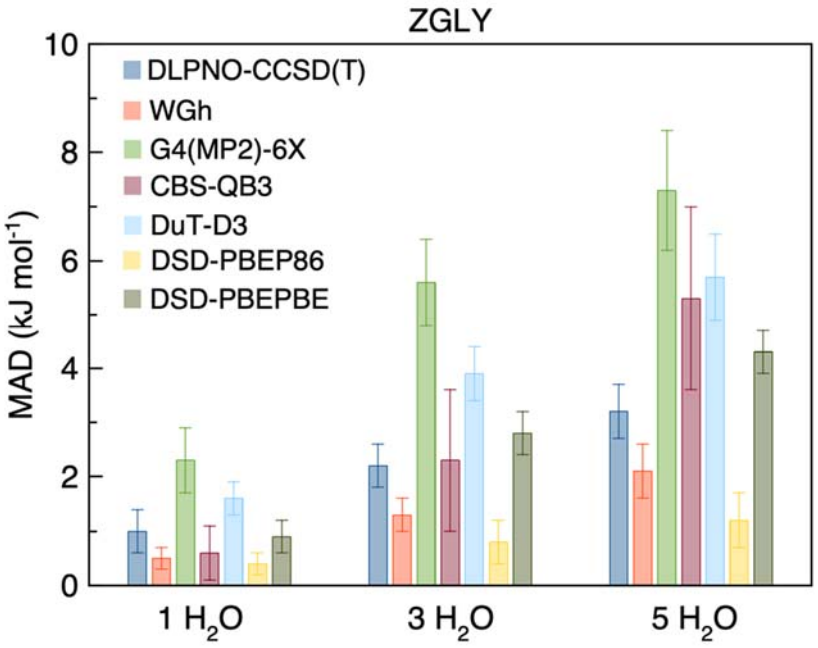

Figure 5. The mean absolute deviation (in interaction energies) from W1X-2 values for clusters containing zwitterionic glycine and 1, 3 and 5 water molecules. The error bars represent the standard deviation over 20 frames.

DFT benchmark – absolute interaction energies. We have carried out DLPNO-CCSD(T)/aug-CBS and DSD-PBEP86/aug'-cc-pVTZ calculations of clusters of NGLY, TS and ZGLY containing 1, 3, 5, 10 and 20 waters and found that the two levels of theory agree to within 2.1 kJ mol^{-1} on average (see Tables S6-S11). As it was impractical[‡] to run the DLPNO-CCSD(T)/CBS for the 40-water clusters, all the mean errors presented in this section are with respect to DSD-PBEP86/aug'-cc-pVTZ values. Figure 6 summarises the MAD of the different DFT methods arranged according to Jacob's Ladder (lowest rung starts from the right-hand side); mean absolute percent deviation (MAPD) are shown in Figure S3). The vertical

axis represents the number of waters in the cluster. The magnitude of the mean signed deviaion (MSD) is identical to the MAD values in the calculated interaction energies indicating the errors are highly systematic. With the exception of M06-2X and ωB97X, all DFT methods under-estimate the benchmark interaction energies, *i.e.* less negative compared to benchmark value.

For most of the Rung 2, Rung 3 and some of the Rung 4 functionals, there is a clear trend where the MAD values increase with the size of the cluster. The ZGLY clusters appear to be more challenging for these methods presumably because the magnitude of the interaction energies is significantly higher compared to NGLY and TS. With the exceptions of B97M-rV, B97-3-D2, B3LYP-D3(BJ) and M06-2X, all functionals in these rungs incur unacceptably significant errors as high as 84 kJ mol^{-1} for the ZGLY 40-water cluster. Interestingly, the B97M-rV performed exceptionally well for NGLY and TS with errors consistently within 5 kJ mol^{-1} for all cluster sizes; however, the MAD grows to 14 kJ mol^{-1} for ZGLY 40-H_2O clusters and is therefore not recommended.

Comparison between B3LYP and B3LYP-D3(BJ) reveal an even more dramatic effect where the MAD reduces from 84.5 to 2.2 kJ mol^{-1} for ZGLY 40-H_2O clusters upon the inclusion of dispersion corrections. This change represents about 22% of the total interaction energy, and is consistent with results from energy decomposition analysis (see Figure S4 and Tables S12-14). In particular, the B3LYP-D3(BJ) performed exceptionally well with MAD values consistently within 4 kJ mol^{-1} across all three charged states of the solute and for all cluster sizes. These results are in accord with previous studies by Grimme and Martin groups who saw a ~ 10-fold reduction in MAD for B3LYP-D3(BJ) compared to B3LYP for the S66x8 dataset that contains dimers at non-equilibrium configurations.[46,114]

Range separated hybrid GGA and meta-GGA (Rung 4) functionals perform very well with MAD values that are typically within 5 kJ mol^{-1} across the board. On this rung, the worst performing functional is M11 where it recorded the highest MAD of 9 kJ mol^{-1} (for NGLY-$(H_2O)_{40}$) and its MAD appears to grow with the size of the system. By comparison, The MAD for all range separated hybrids (ωB97 family) are relatively independent of cluster size. In accord with previous benchmarking studies of smaller dimer systems, this study demonstrates that these methods are reasonably robust and their performance are not sensitive to the cluster size or the charged nature of the solute. Accordingly, the following methods: ωB97X, ωB97X-D, M06-2X, B3LYP-D3(BJ), ωB97X-V and ωB97M-V are recommended for calculating the absolute interaction energies for water clusters and systems similar to this work.

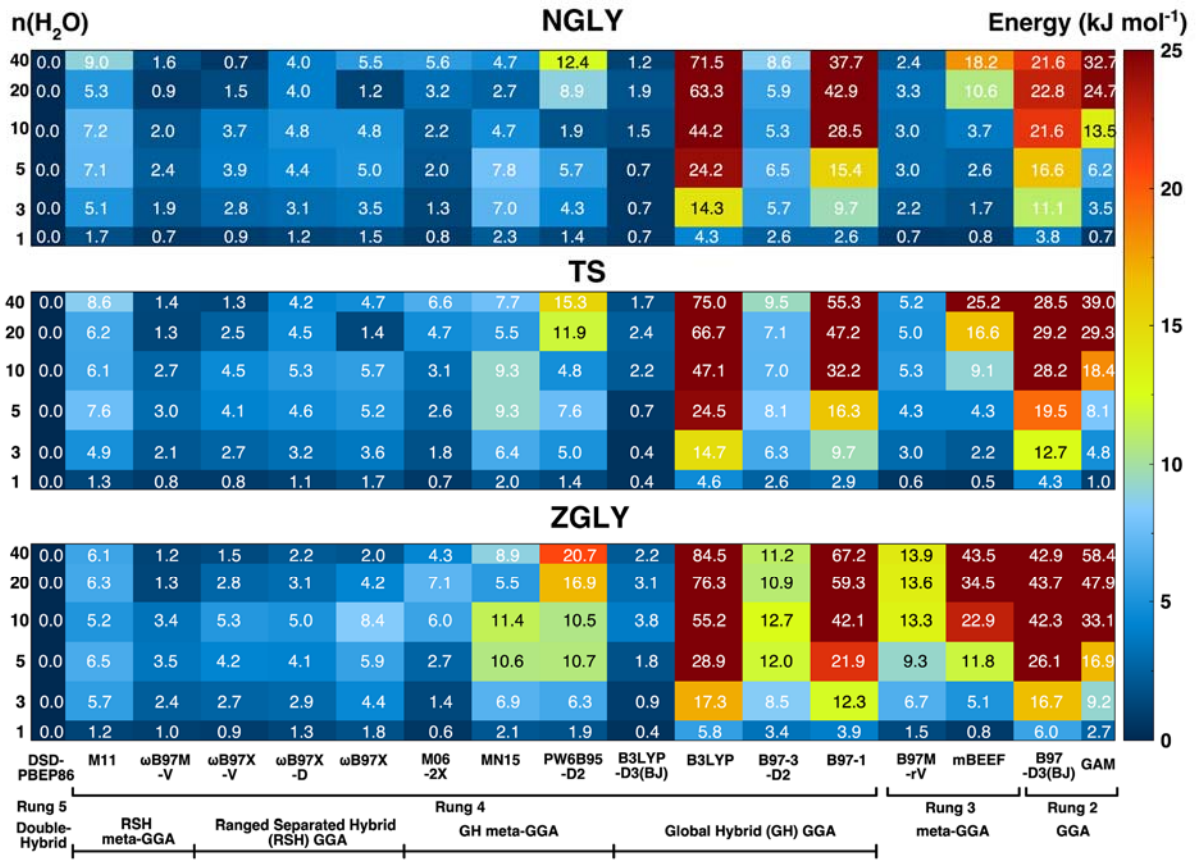

Figure 6. Mean absolute deviation (MAD) in DFT/aug'-cc-pVTZ absolute interaction energies relative to DSD-PBEP86/aug'-cc-pVTZ values averaged over 20 different solvent

configurations. n(H_2O) denotes for the number of water molecules in the glycine-water complex.

DFT benchmark – relative interaction energies. From a practical point of view, the accuracy of calculated absolute interaction energies is less critical for the calculation of reaction energies because the solvation contribution is quantified by the relative solvation free energies of the initial and final states ($\Delta\Delta G_S$). This is also true in the context of the MM→QM end-state corrections presented in Figure 1B. Accordingly, we also examined the performance of various DFT methods to describe the relative interaction energies for the following processes: NGLY → ZGLY and NGLY → TS (Figure 7). As shown, there is a significant improvement in performance across all DFT methods suggesting that there is substantial cancellation of systematic errors. In particular, nearly all DFT methods (with the exception of B97-1) can predict relative interaction energies for the NGLY → TS process to within 10 kJ mol^{-1} of the DSD-PBEP86 benchmark for all cluster sizes. The improvement is less pronounced for the NGLY → ZGLY reaction presumably because the latter reaction is accompanied by a more considerable increase in charge separation and consequently a significant change in solvation pattern.

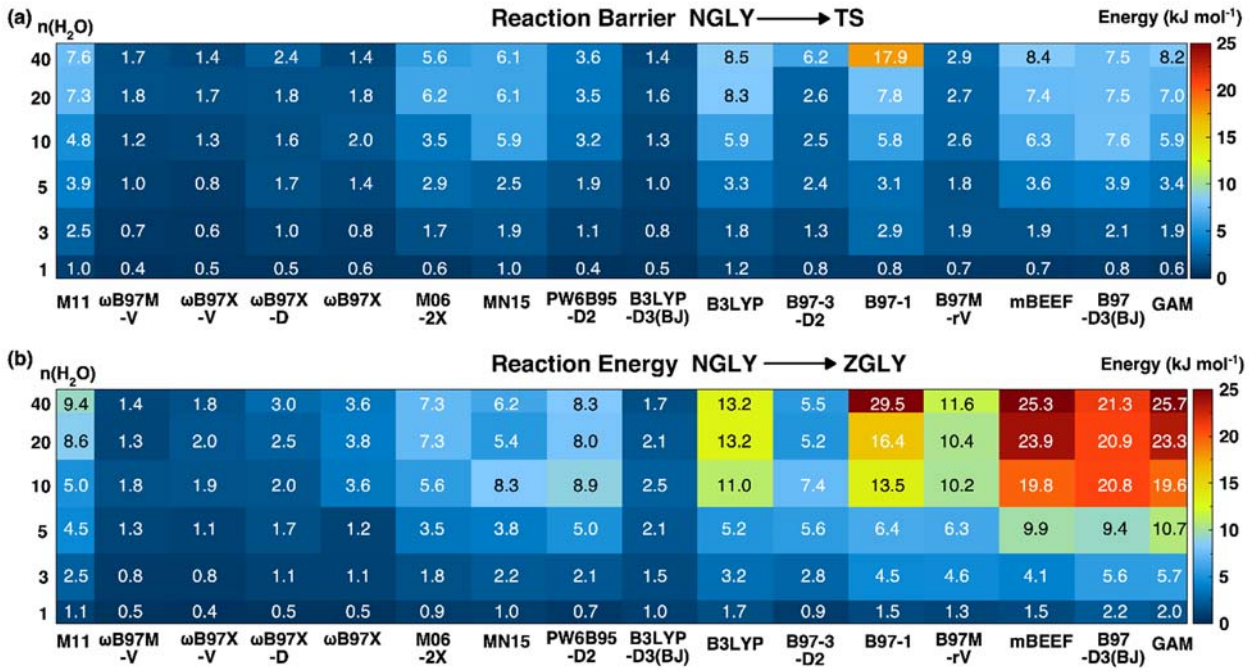

Figure 7. MAD in DFT/aug'-cc-pVTZ relative interaction energies against DSD-PBEP86/aug'-cc-pVTZ values for (a) reaction barrier, NGLY → TS, and (b) reaction energy, NGLY → ZGLY. The relative interaction energy is defined as $\Delta E_{int}^{NGLY \rightarrow X}(n) = E_{int}^{NGLY}(n) - E_{int}^{X}(n)$, with n denoting the number of water molecules, and X = TS or ZGLY.

To adequately describe the relative interaction energies for both processes, we note that with the exception of B97-1 and B3LYP, all Rung 4 global hybrid GGA functionals and above are able to predict relative interactions energies to within ~ 8 kJ mol^{-1} of the benchmark, and unsurprisingly, the previously identified best methods, *i.e.* B3LYP-D3(BJ) and ωB97 family of functionals performed the best with MAD values that are all within 4 kJ mol^{-1}. It is worth highlighting that even though the performance of some functionals such as B97-3-D2 and PW6B95-D2 is significantly improved, the errors in these models appear to grow with the size of the cluster and may not retain the same level of accuracy when applied to larger clusters.

Since the calculation of relative interaction energies benefits from systematic cancellation of errors, we have also examined if the use of smaller (double zeta) basis set calculations could

lead to a further reduction in CPU time while retaining the same level of accuracy. Using a subset of the best performing methods shown in Figure 7, namely M06-2X, ωB97M-V, ωB97X-V, B3LYP-D3(BJ), ωB97X-D and ωB97X, we repeated the calculations using the aug'-cc-pVDZ basis set. As noted in Figure 3, the use of this basis set can incur BSSE larger than 20 kJ mol^{-1} for the ZGLY-$(H_2O)_{20}$ cluster. Nevertheless, the magnitude of the BSSE also appears to be relatively insensitive to the charge state of the solute, so we expect substantial cancellation of BSSE (and presumably BSIE) in the relative interaction energy calculations. As shown in Figure 8, this is indeed the case for the six functionals considered, and the extent of error cancellation was the best for the ωB97M-V, ωB97X-V and B3LYP-D3(BJ) functionals. This is remarkable because the aug'-cc-pVDZ basis set results in ~ 4-fold speed-up for the ωB97 family of functionals when applied to the 40-water cluster and this significantly improves the cost effectiveness of the end-state correction calculations (Figure 1B).

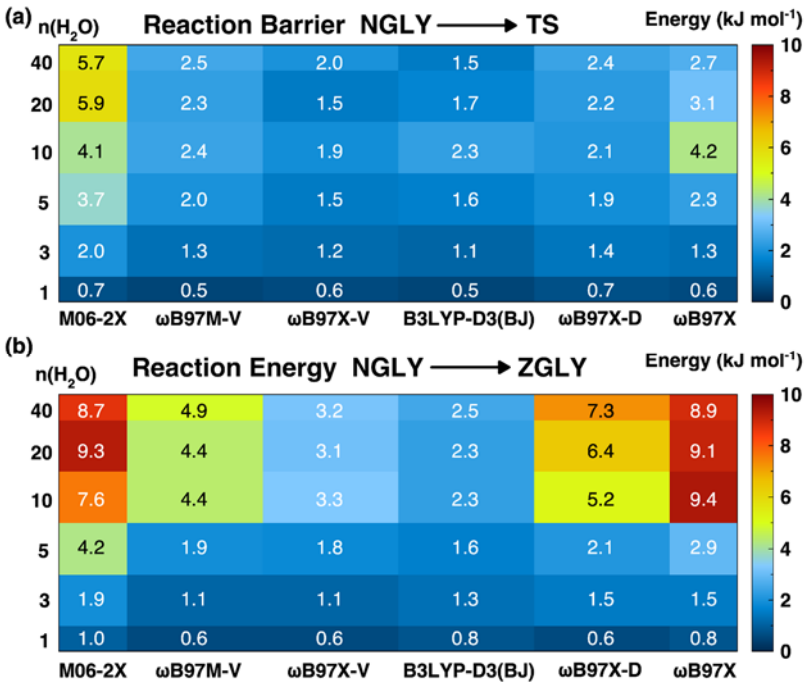

Figure 8. MAD in DFT/aug'-cc-pVDZ relative interaction energies against DSD-PBEP86/aug'-cc-pVTZ values for (a) reaction barrier, NGLY → TS, and (b) reaction energy, NGLY → ZGLY.

ONIOM. In this last section, we consider the use of the hybrid ONIOM method[68] to accelerate the calculation of absolute and relative interaction energies. As shown in Scheme 1, the interaction energy of GLY in an *N*-water cluster calculated at a higher level (HL) of theory is denoted ΔE^{HL}. In an ONIOM scheme, this calculation can be approximated as the sum of two contributions, ΔE^{HL}(core) and ΔE^{HL}(iso) where the former is the interaction energy of GLY with the nearest *n*-waters calculated at the high-level of theory, and the latter is an isodesmic reaction which is calculated at a lower-level (LL) of theory.

Recent studies of ONIOM and related hybrid (QM/MM and QM/QM') methods have found that these approximations can be quite sensitive to the size of the "high layer" and that this would also depend on the pairing of the QM and QM'(or MM) methods.[68,115] Accordingly, we have used the DSD-PBEP86/aug'-cc-pVTZ as our high-level method (HL) and paired it with several lower-level methods namely M06-2X, ωB97M-V, ωB97X-V and B3LYP-D3(BJ) using the smaller aug'-cc-pVDZ basis set as well as the DFTB3 and PM7 semi-empirical methods.

$$\mathrm{GLY^{(HL)}} + (\mathrm{H_2O})_{\mathrm{N}}^{(\mathrm{HL})} \rightarrow \mathrm{GLY}(\mathrm{H_2O})_{\mathrm{N}}^{(\mathrm{HL})} \qquad \Delta E^{HL}$$

$$\mathrm{GLY^{(HL)}} + (\mathrm{H_2O})_{\mathrm{n}}^{(\mathrm{HL})} \rightarrow \mathrm{GLY}(\mathrm{H_2O})_{\mathrm{n}}^{(\mathrm{HL})} \qquad \Delta E_{core}^{HL}$$

$$(\mathrm{H_2O})_{\mathrm{N}}^{(\mathrm{LL})} + \mathrm{GLY}(\mathrm{H_2O})_{\mathrm{n}}^{(\mathrm{LL})} \rightarrow (\mathrm{H_2O})_{\mathrm{n}}^{(\mathrm{LL})} + \mathrm{GLY}(\mathrm{H_2O})_{\mathrm{N}}^{(\mathrm{LL})} \qquad \Delta E_{iso}^{LL}$$

$$\Delta E^{HL} \sim \Delta E^{ONIOM}(n) = \Delta E_{core}^{HL}(n) + \Delta E_{iso}^{LL}(n)$$

Scheme 1. Illustration of the ONIOM approximation for the calculation of interaction energies. *N* is the total number of water molecules in the cluster and *n* refers to a subset of *n*-nearest water molecules to the solute.

Figure 9(a) shows the MAD values due to the ONIOM approximation relative to the DSD-PBEP86/aug'-cc-pVTZ absolute interaction energies for the full ZGLY-$(H_2O)_{40}$ cluster. For comparison, the DFT/aug'-cc-pVDZ interaction energies are also shown on the right hand side of Figure 9(a) (see Figure S5 for ONIOM for NGLY-$(H_2O)_{40}$). The accuracy of the ONIOM approximation was examined as the number of waters in the "high-level-layer" increases. Consistent with our expectation, the error in the ONIOM approximation decays as the "high-level-layer" expands to include an increasing number of water molecules. We note that the ONIOM(DSD-PBEP86/aug'-cc-pVTZ:M06-2X/aug'-cc-pVDZ) calculation was the best performing combination with an MAD ~ 3 kJ mol^{-1} when 10 water molecules are included in the "high-level-layer". The other ONIOM models required at least 20 waters in the "high-layer" before the errors fall below 5 kJ mol^{-1}.

As noted in Scheme 1, the error in the ONIOM approximation can be quantified by the difference in the isodesmic reaction energy computed at the high and low levels of theory, *i.e.* $\Delta E^{HL}(iso) - \Delta E^{LL}(iso)$. Further, the isodesmic reaction can be expressed in terms of the difference in glycine-water interaction energies in the $GLY(H_2O)_N$ and $GLY(H_2O)_n$ clusters so low levels of theory may still give good agreement with the high-level theory (DSD-PBEP86/aug'-cc-pVTZ) provided there is cancellation of systematic errors. Table S16 shows the isodesmic reaction energies determined at the DSD-PBEP86/aug'-cc-pVTZ level and various DFT methods in conjunction with the smaller aug'-cc-pVDZ basis set. The data indicates that M06-2X ΔE(iso) values are in best agreement with the high-level value which is consistent with the results presented in Figure 9.

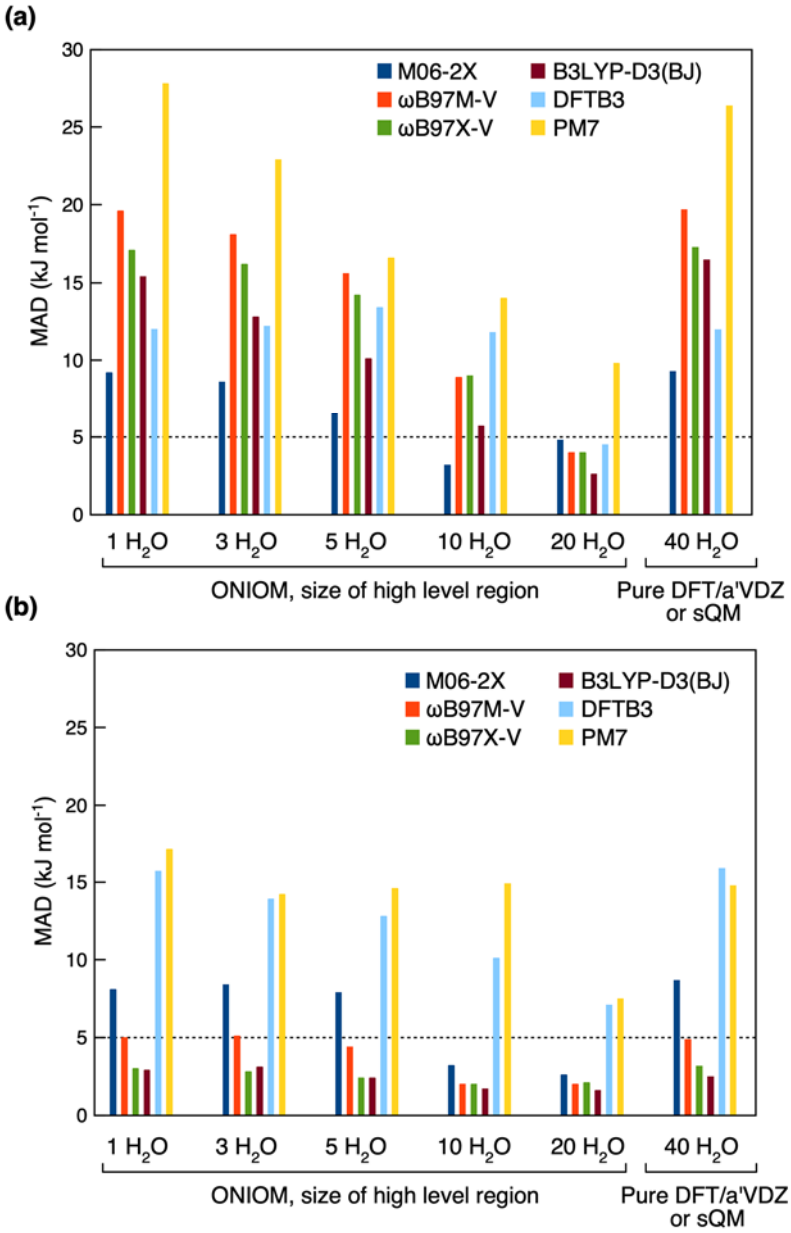

Figure 9. (a) The MAD values for the ONIOM(DSD-PBEP86/aug'-cc-pVTZ:X/aug'-cc-pVDZ) absolute interaction energies (X=M06-2X, ωB97M-V, ωB97X-V, B3LYP-D3(BJ), DFTB3 and PM7) compared to the DSD-PBEP86/aug'-cc-pVTZ benchmark values on the ZGLY-$(H_2O)_{40}$ cluster. (b) Same comparison but for relative interaction energies (NGLY→ZGLY).

Since the DSD-PBEP86/aug'-cc-pVTZ is the computational bottleneck in the ONIOM calculation, the inclusion of 10 and 20 waters in the "high-level-layer" could lead to

approximately 20 and 4-fold reduction in CPU time respectively. As such, we recommend the use of these pairings in ONIOM calculations of absolute interaction energies. The only caveat is that the optimal number of water molecules to include in the "high-level-layer" is likely to vary with solutes, solvents and cluster sizes, so some form of benchmarking is needed to ascertain that the ONIOM calculation has converged with respect to the size of the "high-level-layer".

Figure 9(b) summarises the performance of the ONIOM models for the calculation of relative interaction energies for NGLY→ZGLY. In this case, many of the lower-level methods (ωB97M-V, ωB97X-V and B3LYP-D3(BJ)/aug'-cc-pVDZ) were already predicting relative interaction energies that are in very good agreement with the DSD-PBEP86 so ONIOM is unlikely to provide any enhancement in accuracy or efficiency.

Conclusions

In this paper, we have carried out an extensive assessment of DFT, *ab initio* composite and wavefunction methods for the calculation of solute-solvent interaction energies in thermal clusters (as opposed to zero-Kelvin structures). The results indicate that both DSD-PBEP86 and DLPNO-CCSD(T)/CBS yield very accurate interaction energies in accord with W1X-2 values for small clusters containing up to five water molecules. For the calculation of absolute interaction energies, the ωB97 family of functionals are consistently the best performing methods, along with M06-2X and B3LYP-D3(BJ). These results are in accord with previous assessment studies of smaller clusters at equilibrium configurations which further validates the robustness of these methods in terms of the calculation of electrostatically dominated non-covalent interactions.

From a practical perspective, it is the relative interaction energies that are of relevance for chemical reactions, and we found that the performance of all methods, particularly lower rung DFT methods are significantly improved. However, the relative interaction energies for NGLY → ZGLY remains challenging for many methods presumably because this reaction involves a more significant change in charge separation in the solute such that there is less systematic cancellation of errors. Finally, for the prediction of absolute interaction energies, we identified specific pairings of "high" and "low" level methods that may be used in the ONIOM approximation. The main limitation of this approach is that it is not always clear how many water molecules need to be included in the "high-level-layer".

In summary, this study has identified several robust DFT methods for the calculation of solute-solvent interactions in water clusters as well as some of the strategies that one could apply to improve the cost-effectiveness of these calculations. This is particularly important for the improvement of "reference-potential" type explicit solvent simulations which rely on end-state MM (or QM') to QM corrections and typically requires many thousands of QM interaction energy calculations.

Supporting Information

Coordinates of all the 40-water clusters of NGLY, ZGLY and TS as well as the electronic energies determined at the various levels of theories and supporting figures and tables.

Acknowledgements

JH thanks the Australian Research Council for funding (DE160100807) and the Australian National Computational Infrastructure, UNSW, Intersect NSW and Pawsey Supercomputing Centre for generous allocation of computing resources. YS thanks support from the National

Institutes of Health under grant R01 GM135392 and the Oklahoma Center for the Advancement of Science and Technology under grant HR18-130.

Notes and references

‡ For a single frame of NGLY with 40 H_2O in DLPNO-CCSD(T)/aug-cc-pVTZ, the actual wall time is 286 hours with 7 cores Intel Xeon E5-2690v4 (Broadwell) 2.6GHz, 67.0 GB memory usage, and 1.27 TB disk I/O.

1 J. J. Varghese and S. H. Mushrif, *React. Chem. Eng.*, 2019, **4**, 165–206.

2 C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, 1999, **99**, 2161–2200.

3 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

4 C. J. Cramer and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 760–768.

5 J. Ho and M. Z. Ertem, *J. Phys. Chem. B*, 2016, **120**, 1319–1329.

6 E. L. M. Miguel, C. I. L. Santos, C. M. Silva and J. R. Pliego Jr, *J. Braz. Chem. Soc.*, 2016, **27**, 2055–2061.

7 R. E. Plata and D. A. Singleton, *J. Am. Chem. Soc.*, 2015, **137**, 3811–3826.

8 J. Chen, Y. Shao and J. Ho, *J. Phys. Chem. A*, 2019, **123**, 5580–5589.

9 J. Ho and M. L. Coote, *Theor. Chem. Acc.*, 2009, **125**, 3–21.

10 J. Ho, *Aust. J. Chem.*, 2014, **67**, 1441.

11 J. R. Pliego Jr and J. M. Riveros, *WIREs Comput. Mol. Sci.*, 2019, e1440.

12 W. Wu and J. Kieffer, *J. Chem. Theory Comput.*, 2019, **15**, 371–381.

13 V. S. Bryantsev, M. S. Diallo and W. A. Goddard III, *J. Phys. Chem. B*, 2008, **112**, 9709–9719.

14 D. Asthagiri, L. R. Pratt and H. S. Ashbaugh, *J. Chem. Phys.*, 2003, **119**, 2702–2708.

15 O. Acevedo and W. L. Jorgensen, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 422–435.

16 O. Acevedo and W. L. Jorgensen, *J. Phys. Chem. B*, 2010, **114**, 8425–8430.

17 O. Acevedo and W. L. Jorgensen, *Acc. Chem. Res.*, 2010, **43**, 142–151.

18 J. M. Boereboom, P. Fleurat-Lessard and R. E. Bulo, *J. Chem. Theory Comput.*, 2018,

14, 1841–1852.

19 T. Cheng, A. Fortunelli and W. A. Goddard, *Proc. Natl. Acad. Sci.*, 2019, **116**, 7718–7722.

20 M. Pu and T. Privalov, *Chem. – A Eur. J.*, 2015, **21**, 17708–17720.

21 F. Calle-Vallejo and M. T. M. Koper, *Electrochim. Acta*, 2012, **84**, 3–11.

22 L.-P. Wang and T. Van Voorhis, *J. Chem. Theory Comput.*, 2012, **8**, 610–617.

23 A. V. Marenich, J. Ho, M. L. Coote, C. J. Cramer and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2014, **16**, 15068–15106.

24 J. Blumberger, L. Bernasconi, I. Tavernelli, R. Vuilleumier and M. Sprik, *J. Am. Chem. Soc.*, 2004, **126**, 3928–3938.

25 F. H. Hodel and S. Luber, *J. Chem. Theory Comput.*, 2017, **13**, 974–981.

26 V. Vaissier and T. Van Voorhis, *J. Chem. Theory Comput.*, 2016, **12**, 5111–5116.

27 C. M. Sterling and R. Bjornsson, *J. Chem. Theory Comput.*, 2019, **15**, 52–67.

28 M. Schilling and S. Luber, *Inorganics*, 2019, **7**, 73.

29 N. Uddin, T. H. Choi and C. H. Choi, *J. Phys. Chem. B*, 2013, **117**, 6269–6275.

30 A. K. Tummanapelli and S. Vasudevan, *J. Phys. Chem. B*, 2014, **118**, 13651–13657.

31 J. Cheng, M. Sulpizi and M. Sprik, *J. Chem. Phys.*, 2009, **131**, 154504.

32 M. Wang, P. Li, X. Jia, W. Liu, Y. Shao, W. Hu, J. Zheng, B. R. Brooks and Y. Mei, *J. Chem. Inf. Model.*, 2017, **57**, 2476–2489.

33 G. König, F. Pickard, J. Huang, W. Thiel, A. MacKerell, B. Brooks and D. York, *Molecules*, 2018, **23**, 2695.

34 T. H. Rod and U. Ryde, *Phys. Rev. Lett.*, 2005, **94**, 138302.

35 T. H. Rod and U. Ryde, *J. Chem. Theory Comput.*, 2005, **1**, 1240–1251.

36 R. P. Muller and A. Warshel, *J. Phys. Chem.*, 1995, **99**, 17516–17524.

37 M. Štrajbl, G. Hong and A. Warshel, *J. Phys. Chem. B*, 2002, **106**, 13333–13343.

38 F. Duarte, B. A. Amrein, D. Blaha-Nelson and S. C. L. Kamerlin, *Biochim. Biophys. Acta - Gen. Subj.*, 2015, **1850**, 954–965.

39 T. J. Giese and D. M. York, *J. Chem. Theory Comput.*, 2019, **15**, 5543–5562.

40 X. Jia, M. Wang, Y. Shao, G. König, B. R. Brooks, J. Z. H. H. Zhang, Y. Mei, G. König, B. R. Brooks, J. Z. H. H. Zhang and Y. Mei, *J. Chem. Theory Comput.*, 2016, **12**, 499–511.

41 M. A. Olsson and U. Ryde, *J. Chem. Theory Comput.*, 2017, **13**, 2245–2253.

42 M. A. Collins and J. Ho, *J. Phys. Chem. A*, 2019, **123**, 8476–8484.

43 L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670–6688.

44 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.

45 N. Mardirossian and M. Head-Gordon, *Mol. Phys.*, 2017, **115**, 2315–2372.

46 B. Brauer, M. K. Kesharwani, S. Kozuch and J. M. L. L. Martin, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20905–20925.

47 M. K. Kesharwani, A. Karton and J. M. L. Martin, *J. Chem. Theory Comput.*, 2016, **12**, 444–454.

48 D. Manna, M. K. Kesharwani, N. Sylvetsky and J. M. L. L. Martin, *J. Chem. Theory Comput.*, 2017, **13**, 3136–3152.

49 M. Pitoňák, P. Neogrády, J. Černý, S. Grimme and P. Hobza, *ChemPhysChem*, 2009, **10**, 282–289.

50 P. Jurečka, J. Šponer, J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.

51 J. Rezác, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2011, **7**, 2427–2438.

52 J. Řezáč, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2011, **7**, 3466–3470.

53 C. D. Sherrill, T. Takatani and E. G. Hohenstein, *J. Phys. Chem. A*, 2009, **113**, 10146–10159.

54 K. S. Thanthiriwatte, E. G. Hohenstein, L. A. Burns and C. D. Sherrill, *J. Chem. Theory Comput.*, 2011, **7**, 88–96.

55 T. M. Parker and C. D. Sherrill, *J. Chem. Theory Comput.*, 2015, **11**, 4197–4204.

56 L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. A. Smith, K. Vanommeslaeghe, A. D. MacKerell, K. M. Merz and C. D. Sherrill, *J. Chem. Phys.*,

2017, **147**, 161727.

57 N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.*, 2016, **144**, 214110.

58 M. S. Gordon and J. H. Jensen, *Acc. Chem. Res.*, 1996, **29**, 536–543.

59 B. Chan, J. Rintelman, G. P. F. Wood, L. Radom and M. S. Gordon, *J. Phys. Chem. A*, 2018, **122**, 7212–7217.

60 J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, 1995, **117**, 8159–8170.

61 J. M. Mullin and M. S. Gordon, *J. Phys. Chem. B*, 2009, **113**, 14413–14420.

62 D. G. Fedorov and K. Kitaura, *J. Phys. Chem. A*, 2018, **122**, 1781–1795.

63 D. G. Fedorov, J. C. Kromann and J. H. Jensen, *Chem. Phys. Lett.*, 2018, **706**, 328–333.

64 M. S. Gordon, J. M. Mullin, S. R. Pruitt, L. B. Roskop, L. V Slipchenko and J. A. Boatz, *J. Phys. Chem. B*, 2009, **113**, 9646–9663.

65 R. Kobayashi, R. Amos and M. A. Collins, *J. Phys. Chem. A*, 2017, **121**, 334–341.

66 S. R. Pruitt, C. Bertoni, K. R. Brorsen and M. S. Gordon, *Acc. Chem. Res.*, 2014, **47**, 2786–2794.

67 C. M. Aikens and M. S. Gordon, *J. Am. Chem. Soc.*, 2006, **128**, 12835–12850.

68 L. W. Chung, W. M. C. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H. B. Li, L. Ding and K. Morokuma, *Chem. Rev.*, 2015, **115**, 5678–5796.

69 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V Marenich, J. Bloinom, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. M. Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin,

K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revis. B.01*, 2019, Gaussian Inc., Wallingford CT, 2016.

70 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, others, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang and M. Welborn, *Molpro, version 2019.2*, http://www. molpro.net.

71 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *WIREs Comput Mol Sci*, 2012, **2**, 242–253.

72 Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuš, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C. M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. Distasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T. C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S. P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. Oneill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y. C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z. Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon,

W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J. Da Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C. P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Mol. Phys.*, 2015, **113**, 184–215.

73 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.

74 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2018, **8**, e1327.

75 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell, *J. Comput. Chem.*, 2010, **31**, 671–690.

76 W. Yu, X. He, K. Vanommeslaeghe and A. D. MacKerell, *J. Comput. Chem.*, 2012, **33**, 2451–2468.

77 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.

78 A. V Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.

79 B. Chan, *J. Chem. Theory Comput.*, 2017, **13**, 2642–2649.

80 B. Chan, J. Deng and L. Radom, *J. Chem. Theory Comput.*, 2010, **7**, 112–120.

81 J. A. Montgomery Jr, M. J. Frisch, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 2000, **112**, 6532–6542.

82 B. Chan and L. Radom, *J. Chem. Theory Comput.*, 2012, **8**, 4259–4269.

83 F. Neese and E. F. Valeev, *J. Chem. Theory Comput.*, 2011, **7**, 33–43.

84 J. P. Perdew and K. Schmidt, *AIP Conf. Proc.*, 2001, **577**, 1–20.

85 B. Chan and L. Radom, *J. Chem. Theory Comput.*, 2011, **7**, 2852–2863.

86 S. Kozuch and J. M. L. Martin, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20104–20107.

87 S. Kozuch and J. M. L. Martin, *J. Comput. Chem.*, 2013, **34**, 2327–2344.

88 R. Peverati and D. G. Truhlar, *J. Phys. Chem. Lett.*, 2011, **2**, 2810–2817.

89 N. Mardirossian and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2014, **16**, 9904–

9924.

90 J. Da Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

91 J. Da Chai and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 84106.

92 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.

93 H. S. Yu, X. He, S. L. Li and D. G. Truhlar, *Chem. Sci.*, 2016, **7**, 6278–6279.

94 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 5656–5667.

95 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.

96 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

97 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

98 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

99 T. W. Keal and D. J. Tozer, *J. Chem. Phys.*, 2005, **123**, 121103.

100 F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy, *J. Chem. Phys.*, 1998, **109**, 6264–6271.

101 N. Mardirossian, L. Ruiz Pestana, J. C. Womack, C.-K. Skylaris, T. Head-Gordon and M. Head-Gordon, *J. Phys. Chem. Lett.*, 2016, **8**, 35–40.

102 J. Wellendorff, K. T. Lundgaard, K. W. Jacobsen and T. Bligaard, *J. Chem. Phys.*, 2014, **140**, 144107.

103 H. S. Yu, W. Zhang, P. Verma, X. He and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2015, **17**, 12146–12160.

104 K. Throssel and M. J. Frisch, *Manuscr. Prep.*

105 M. Gaus, Q. Cui and M. Elstner, *J. Chem. Theory Comput.*, 2011, **7**, 931–948.

106 Yang, H. Yu, D. York, Q. Cui and M. Elstner, *J. Phys. Chem. A*, 2007, **111**, 10861–10873.

107 B. Aradi, B. Hourahine and T. Frauenheim, *J. Phys. Chem. A*, 2007, **111**, 5678–5684.

108 M. Elstner, P. Hobza, T. Frauenheim, S. Suhai and E. Kaxiras, *J. Chem. Phys.*, 2001, **114**, 5149–5155.

109 P. R. Horn, Y. Mao and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2016, **18**, 23067–23079.

110 P. R. Horn and M. Head-Gordon, *J. Chem. Phys.*, 2015, **143**, 114111.

111 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.

112 A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay and J. Gauss, *J. Chem. Phys.*, 2004, **120**, 4129–4141.

113 A. Karton, R. J. O'Reilly, B. Chan and L. Radom, *J. Chem. Theory Comput.*, 2012, **8**, 3128–3136.

114 L. Goerigk, H. Kruse and S. Grimme, *ChemPhysChem*, 2011, **12**, 3421–3433.

115 J. Ho, Y. Shao and J. Kato, *Molecules*, 2018, **23**, 2466.