Supplementary Material for "An improved treatment of empirical dispersion and a many-body energy decomposition scheme for XPol+SAPT (XSAPT)"

Ka Un Lao and John M. Herbert*
Department of Chemistry and Biochemistry,
The Ohio State University, Columbus, OH 43210

(Dated: June 23, 2013)

1. PARAMETERS

Table S1 lists the "tuned" value of the range separation parameter, ω, for each of the monomers considered in this work. These values were tuned using the haTZVPP basis set (a "heavy augmented" version of def2-TZVPP as described in the paper), and optimized values for both the LRC-ωPBE[1] and LRC-ωPBEh[2] functionals are listed. The "h" in LRC-ωPBEh indicates that the short-range exchange functional is a hybrid, and in our previous paper on XSAPT(KS)+D,[3] we used 60% Hartree-Fock exchange in the short-range ωPBEh component of this functional. This choice afforded good binding energies, but such a large value leads to errors in the exchange components of the interaction energy, as discussed in the present paper. In the second-generation (D2) version of the method, we use 20% Hartree-Fock exchange with LRC-ωPBEh (and 0% with LRC-ωPBE), which is consistent with the original parameterization of these functionals.[1,2] Notably, whereas in Ref. 3 we encountered seven cases in the S22 data set where the ω-tuning condition

$$\varepsilon_{\mathrm{HOMO}}(\omega) = -\mathrm{IP}(\omega) \tag{S1}$$

could not be met, because the $\varepsilon_{\mathrm{HOMO}}(\omega)$ and $-\mathrm{IP}(\omega)$ curves did not cross, we encounter no such cases in the present work.

Table S2 lists the fitted values of the s_β parameter for use in Hesselmann's dispersion potential,[4] which is the one used in our previous paper.[3] This is the potential called "+D" (as opposed to "+D2") in this work. Values of s_β in Table S2 were optimized to reproduce S22A binding energies, in a basis-set-specific way.

---

* `herbert@chemistry.ohio-state.edu`

2. WATER HEXAMER BENCHMARKS

We next wish to explain our choice of binding energy benchmarks for isomers of $(H_2O)_6$. As mentioned in the paper, a primary goal in examining $(H_2O)_6$ is to make comparison to energy decomposition analysis (EDA) results in Ref. 5, as a means to validate XSAPT-based EDA. These EDA benchmarks were computed at the MP2/a5Z-h//MP2/aTZ level and *do not* include monomer relaxation energy. On the other hand, prediction of binding energies is a major focus of this work and the most accurate binding energy benchmarks for $(H_2O)_6$ are probably the ones in Ref. 6, which were computed at the CCSD(T)/CBS//MP2/haTZ level. (The customized basis sets used in these benchmarks are defined as follows: a5Z-h means aug-cc-pV5Z with h functions omitted, whereas haTZ is means aug-cc-pVTZ for oxygen and cc-pVTZ for hydrogen.)

Benchmark binding energies for eight isomers of $(H_2O)_6$, taken from the aforementioned references, are listed in Table S3. The same eight isomers of $(H_2O)_6$ are considered in both papers, differing only in whether the aTZ of haTZ basis was used for the MP2 geometry optimizations. The presence or absence of diffuse functions on the hydrogen atoms likely makes very little difference in the actual geometries, as confirmed in Ref. 5. Three sets of binding energies are taken from Ref. 5, with each set computed at the MP2/a5Z-h//MP2/aTZ level of theory but differing based on whether counterpoise correction is included and whether monomer relaxation is included. In Ref. 6, monomer relaxation is always included but counterpoise correction is not (because all results are extrapolated to the CBS limit); MP2/CBS//MP2/haTZ and CCSD(T)/CBS//MP2/haTZ results are shown. Table S4 lists some energy differences between these benchmarks.

Our strategy for benchmarking XSAPT(KS)+D2 in the case of $(H_2O)_6$ is as follows. We will use the MP2/aTZ geometries from Ref. 5, because those are the geometries at which the EDA has been performed, and for comparing energy components to XSAPT(KS)+D2 results we will use the MP2/a5Z-h EDA results in Ref. 5. On the other hand, we should examine binding energies as well (since we examine binding energies for all of the other systems considered in this paper), and we wish to do this at the same set of geometries, namely, the MP2/aTZ ones. To obtain approximate CCSD(T)/CBS //MP2/aTZ binding energies benchmarks, we add a triples correction

$$\delta_{\mathrm{CCSD(T)}} = \mathrm{BE}_{\mathrm{CCSD(T)}} - \mathrm{BE}_{\mathrm{MP2}} \tag{S2}$$

to the counterpoise-corrected MP2/a5Z-h//MP2/aTZ binding energies from Ref. 5. This correction is computed using the CCSD(T)/CBS and MP2/CBS binding energies (BEs) from Ref. 6. Finally, because we have computed $(H_2O)_6$ binding energies at a variety of levels of theory *without* considering monomer relaxation (*i.e.*, the binding energy is computed relative to monomers that retain their cluster geometries), we subtract the monomer relaxation energies δ_{relax} that are computed at the MP2/a5Z-h level in Ref. 5 and listed in Table S4.

The resulting "corrected" binding energies are listed in Table S5, under the column titled "+ triples − relaxation". (In the paper, this is the data set labeled "MP2/a5Z-h + $\delta_{\mathrm{CCSD(T)}}$" in Fig. 4.) The primary source of error in these data, relative to a true CCSD(T)/CBS //MP2/aTZ benchmark, is likely the fact that we have not extrapolated the MP2 energy to the basis-set limit, but rather have simply used MP2/a5Z-h values. Table S4 shows the difference, δ_{MP2}, between the MP2/a5Z-h //MP2/aTZ binding energies from Ref. 5 and the MP2/CBS//MP2/haTZ results from Ref. 6; this difference is consistently 0.7–0.8 kcal/mol in magnitude for the eight isomers that are considered. The differences in cluster geometries likely have a very minor impact on this difference, so most of δ_{MP2} probably arises from basis-set incompleteness at the a5Z-h level. (The MP2/CBS results in Ref. 6 are converged to within 0.1 kcal/mol of MP2-R12 calculations performed in a modified aug-cc-pV5Z basis,[6] and are therefore likely within 0.1 kcal/mol of the actual basis-set limit.) Residual basis-set superposition error is unlikely a major contribution to δ_{MP2}, since the MP2/CBS binding energies are larger than the counterpoise-corrected MP2/a5Z-h values (see Table S5), and counterpoise-corrected binding energies converge from below in this system.[7]

Lastly, we may consider simply using the CCSD(T)/CBS//MP2/haTZ binding energy benchmarks from Ref. 6 at the MP2/aTZ geometries from Ref. 5. To compare with other results in the paper, we want binding energies using unrelaxed monomer geometries, so the CCSD(T)/CBS binding energies in Ref. 6 have been corrected using the MP2/a5Z-h values of δ_{relax} in Table S4. These results constitute the final column of data in Table S5, which is the data set labeled "CCSD(T)/ CBS//MP2/haTZ" in Fig. 4 of the paper.

3. HALIDE–WATER BENCHMARKS

Figure S1 shows binding energies for the ten $F^-(H_2O)_{10}$ isomers that were considered in the paper, computed at various levels of theory including the benchmark RI-CCSD(T)/CBS results. (Geometries of these isomers have been optimized at the B3LYP/6-31G* level and Cartesian coordinates, in Angstroms, are provided as part of this Supplementary Material, in a separate text file.) This figure also serves to document that, for the most part, the various XSAPT and DFT methods considered here parallel the benchmark results, meaning that relative isomer energies are predicted more accurately than are absolute binding energies, for which there are some systematic errors.

RI-MP2/CBS and RI-CCSD(T)/CBS binding energies for $F^-(H_2O)_n$ and $Cl^-(H_2O)_n$, up to $n = 6$, are shown in Tables S6 and S7. These geometries were optimized at the RI-MP2/aug-cc-pVTZ level and are available elsewhere in this Supplementary Material. Both two- and three-body approximations to the triples correction, $\delta_{\mathrm{RI-CCSD(T)}}$, are shown for comparison; the three-body approximation is used as the benchmark in this work. Table S8 provides the analogous information for $F^-(H_2O)_{10}$ isomers whose geometries were optimized at the B3LYP/6-31G* level.

---

[1] A. W. Lange and J. M. Herbert, J. Am. Chem. Soc. **131**, 3913 (2009).

[2] M. A. Rohrdanz, K. M. Martins, and J. M. Herbert, J. Chem. Phys. **130**, 054112 (2009).

[3] K. U. Lao and J. M. Herbert, J. Phys. Chem. Lett. **3**, 3241 (2012).

[4] A. Hesselmann, J. Phys. Chem. A **115**, 11321 (2011).

[5] Y. Chen and H. Li, J. Phys. Chem. A **114**, 11719 (2010).

[6] D. M. Bates and G. S. Tschumper, J. Phys. Chem. A **113**, 3555 (2009).

[7] R. M. Richard, K. U. Lao, and J. M. Herbert (in preparation).

	$\omega/\mathrm{bohr}^{-1}$	
Monomer	LRC-ωPBE	LRC-ωPBEh
—S22 molecules—		
adenine	0.275	0.225
2-aminopyridine	0.300	0.225
benzene	0.275	0.225
ethyne	0.400	0.350
ethene	0.350	0.300
methane	0.450	0.400
formamide	0.375	0.300
formic acid	0.425	0.325
water	0.500	0.400
HCN	0.550	0.500
indole	0.275	0.225
ammonia	0.450	0.350
phenol	0.300	0.250
pyrazine	0.375	0.300
2-pyridoxine	0.300	0.250
thymine	0.625	0.525
uracil	0.475	0.375
—S66 molecules—		
methylamine	0.400	0.325
methanol	0.425	0.350
$\mathrm{AcNH_2}$	0.350	0.275
AcOH	0.375	0.300
cyclopentane	0.425	0.375
neopentane	0.300	0.250
pentane	0.300	0.250
peptide	0.325	0.250
pyridine	0.300	0.225
—Misc. molecules & ions—		
Ar	0.550	0.475
Ne	0.800	0.650
$\mathrm{F^-}$	0.475	0.375
$\mathrm{Cl^-}$	0.375	0.300

TABLE S1: Tuned range separation parameters for various monomers.

Basis set	s_β
aug-cc-pVDZ′	0.7267[a]
TZVPP	0.8439[a]
haTZVPP	0.8587[b]
haTZVPP (AC)	0.8547[b]

[a] From Ref. 3.

[b] Introduced in this work.

TABLE S2: Values of the dimensionless parameter s_β used in the "first generation" XSAPT(KS)+D method. These values have been optimized against binding energies in the S22A data set, using a KS functional equal to LRC-ωPBEh with 60% short-range Hartree-Fock exchange and a tuned value of ω. The exact AC scheme uses monomer-specific values of ω; in other cases, ω is tuned according to Eq. (S1) using the lowest supersystem IP.

	Chen & Li (Ref. 5)			Bates & Tschumper (Ref. 6)	
	MP2/a5Z-h			MP2/CBS	CCSD(T)/CBS
Includes BSSE correction?	Yes	Yes	No	No	No
Includes relaxation energy?	No	Yes	Yes	Yes	Yes
prism	−47.41	−45.05	−46.09	−45.86	−45.92
cage	−47.33	−44.98	−46.05	−45.80	−45.67
book1	−47.15	−44.73	−45.79	−45.53	−45.20
book2	−46.87	−44.41	−45.46	−45.22	−44.90
cyclic chair	−46.27	−43.94	−44.95	−44.65	−44.12
bag	−46.41	−43.83	−44.89	−44.63	−44.30
cyclic boat1	−45.29	−42.95	−43.96	−43.66	−43.13
cyclic boat2	−45.17	−42.87	−43.87	−43.58	−43.07

TABLE S3: Comparison of binding energies (in kcal/mol) for isomers of $(H_2O)_6$. All three sets of calculations from Ref. 5 are performed at the MP2/a5Z-h//MP2/aTZ level but differ depending on whether or not the Boys- Bernardi counterpoise correction was applied, and whether or not monomer relaxation is included in the binding energy. The calculations from Ref. 6 use MP2/haTZ geometries, and single-point energies at both the MP2/CBS and CCSD(T)/CBS level are listed here. All data come from Refs. 5 and 6.

isomer	correction / kcal mol^{-1}			
	δ_{MP2}[a]	$\delta_{\mathrm{CCSD(T)}}$[b]	$\delta_{\mathrm{MP2}}+\delta_{\mathrm{CCSD(T)}}$[c]	δ_{relax}[d]
prism	−0.81	−0.06	−0.87	2.36
cage	−0.82	+0.13	−0.69	2.35
book1	−0.80	+0.33	−0.47	2.42
book2	−0.81	+0.32	−0.49	2.46
cyclic chair	−0.71	+0.53	−0.18	2.33
bag	−0.80	+0.33	−0.47	2.58
cyclic boat1	−0.71	+0.53	−0.18	2.34
cyclic boat2	−0.71	+0.51	−0.20	2.30

[a] Difference between MP2/CBS and MP2/a5Z-h binding energies.

[b] Difference between CCSD(T)/CBS and MP2/CBS binding energies.

[c] The sum of these two corrections equals the difference between CCSD(T)/CBS and MP2/a5Z-h binding energies.

[d] Monomer relaxation energy, estimated from counterpoise-corrected MP2/a5Z-h results.

TABLE S4: Various differences in binding energies that can be extracted from the data in Table S3. See the footnotes and the text for an explanation of each quantity. Note that δ_{MP2} includes the effects of the slightly different geometries used in Ref. 5 versus those in Ref. 6.

isomer	binding energy / kcal mol^{-1}			
	MP2/a5Z-h results			CCSD(T)/CBS
	MP2/a5Z-h// MP2/aTZ[a]	+triples correction[b]	+triples −relaxation[c]	−relaxation[d]
prism	−45.05	−45.11	−47.47	−48.28
cage	−44.98	−44.85	−47.20	−48.02
book1	−44.73	−44.40	−46.82	−47.62
book2	−44.41	−44.09	−46.55	−47.36
cyclic chair	−43.94	−43.41	−45.74	−46.45
bag	−43.83	−43.50	−46.08	−46.88
cyclic boat1	−42.95	−42.42	−44.76	−45.47
cyclic boat2	−42.87	−42.36	−44.66	−45.37

[a] Counterpoise-corrected result including monomer relaxation (see Table S3).

[b] MP2/a5Z-h result plus $\delta_{\mathrm{CCSD(T)}}$ from Table S4.

[c] MP2/a5Z-h result plus $\delta_{\mathrm{CCSD(T)}}$ minus δ_{relax} from Table S4.

[d] CCSD(T)/CBS//MP2/haTZ result from Ref. 6, corrected to remove monomer relaxation using MP2/a5Z-h values of δ_{relax}.

TABLE S5: MP2/a5Z-h//MP2/aTZ binding energies from Ref. 5 and "corrected" values based on CCSD(T) results in Ref. 6. See the footnotes and the text for an explanation of the corrections.

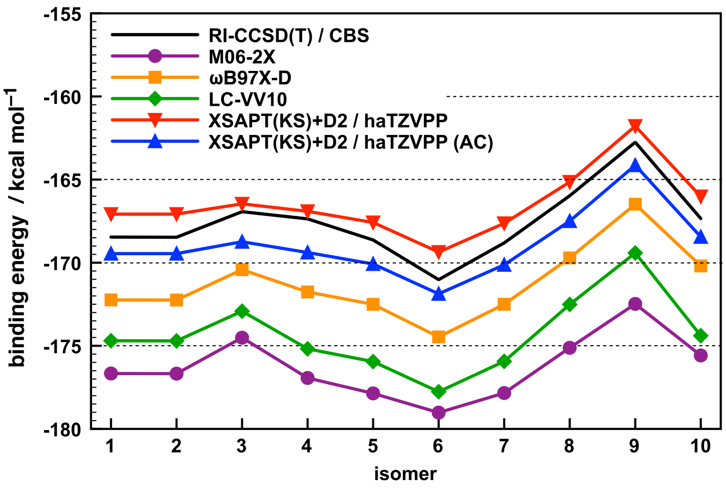

FIG. S1: Binding energies for ten isomers of $F^-(H_2O)_{10}$ isomers optimized at the B3LYP/6-31G* level. The RI-CCSD(T)/CBS benchmarks are compared to DFT/def2-QZVP results (computed using Boys-Bernardi counterpoise correction) and to XSAPT(KS)+D2/results (computed using the $\delta E_{\rm int}^{\rm HF}$ correction).

n	binding energy / kcal mol^{-1}		$\delta_{\mathrm{RI-CCSD(T)}}$ / kcal mol^{-1}	
	RI-MP2/CBS	RI-CCSD(T)/CBS	two-body	three-body
1	−32.25	−32.32	−0.07	—
2	−51.86	−52.22	−0.44	−0.36
3	−69.08	−70.07	−1.27	−0.99
4	−83.82	−85.11	−1.53	−1.29
5	−99.52	−100.97	−1.82	−1.45
6	−114.75	−116.21	−1.83	−1.46

TABLE S6: Binding energies for isomers of $F^-(H_2O)_n$, computed at the RI-MP2/CBS and RI-CCSD(T)/CBS levels using RI-MP2/aug-cc-pVTZ geometries. Also shown is the triples correction, $\delta_{\mathrm{RI-CCSD(T)}}$, computed using either a two-body or a three-body approximation. and the corresponding two-body and three-body triples corrections $\delta_{\mathrm{RI-CCSD(T)}}$ are also provided. The RI-CCSD(T)/CBS binding energy is the sum of RI-MP2/CBS binding energy (obtained by extrapolation) and the three-body triples correction.

n	binding energy / kcal mol^{-1}		$\delta_{\mathrm{RI-CCSD(T)}}$ / kcal mol^{-1}	
	RI-MP2/CBS	RI-CCSD(T)/CBS	two-body	three-body
1	−15.75	−15.48	0.27	—
2	−31.28	−30.94	0.30	0.34
3	−47.72	−47.50	0.02	0.22
4	−58.38	−58.10	0.16	0.28
5	−74.92	−74.58	0.10	0.34
6	−86.91	−86.45	0.21	0.46

TABLE S7: Binding energies for isomers of $Cl^-(H_2O)_n$, computed at the RI-MP2/CBS and RI-CCSD(T)/CBS levels using RI-MP2/aug-cc-pVTZ geometries. Also shown is the triples correction, $\delta_{\mathrm{RI-CCSD(T)}}$, computed using either a two-body or a three-body approximation. and the corresponding two-body and three-body triples corrections $\delta_{\mathrm{RI-CCSD(T)}}$ are also provided. The RI-CCSD(T)/CBS binding energy is the sum of RI-MP2/CBS binding energy (obtained by extrapolation) and the three-body triples correction.

n	binding energy / kcal mol^{-1}		$\delta_{\mathrm{RI-CCSD(T)}}$ / kcal mol^{-1}	
	RI-MP2/CBS	RI-CCSD(T)/CBS	two-body	three-body
1	−166.67	−168.46	−2.47	−1.79
2	−166.68	−168.47	−2.47	−1.79
3	−165.15	−166.93	−2.29	−1.79
4	−165.54	−167.36	−2.38	−1.82
5	−166.89	−168.64	−2.33	−1.74
6	−169.50	−171.02	−2.17	−1.52
7	−167.06	−168.81	−2.33	−1.75
8	−163.91	−165.97	−2.73	−2.07
9	−160.73	−162.75	−2.81	−2.03
10	−165.72	−167.34	−2.06	−1.63

TABLE S8: RI-MP2/CBS and RI-CCSD(T)/CBS binding energies for ten different isomers of $F^-(H_2O)_{10}$. Geometries were optimized at the B3LYP/6-31G* level and the numbering scheme reflects the order of these isomers in the Cartesian coordinate file that is provided as part of this Supplementary Material. The corresponding two-body and three-body triples corrections $\delta_{\mathrm{RI-CCSD(T)}}$ are also provided. RI-CCSD(T)/CBS binding energy are the sum of the RI-MP2/CBS binding energy (obtained by extrapolation) and the three-body triples correction.