Supplementary Material for "Accurate and Efficient Quantum Chemistry Calculations for Non-Covalent Interactions in Many-Body Systems: The XSAPT Family of Methods"

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

(Dated: November 17, 2014)

1. BASIS SET CONVERGENCE TESTS

Table S1 shows the errors in the first-order energy components $E^{(1)}_{\text{elst}}$ and $E^{(1)}_{\text{exch}}$, with respect to SAPT2+(3)/aTZ benchmarks. The errors are averages over four representative systems: $F^-(H_2O)$, $(H_2O)_2$, and the T-shaped and parallel-displaced isomers of $(C_6H_6)_2$. Results are shown for 21 different AO basis sets, using either the dimer-centered SAPT basis or else the "projected" (pseudo-canonicalized monomer-centered) SAPT basis.

2. TUNED VALUES OF THE RANGE SEPARATION PARAMETER

Table S2 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 LRC-ωPBE/hpTZVPP method.

3. EMPIRICAL DISPERSION PARAMETERS

In addition to high-level SAPT2+(3)/aTZ, a more robust treatment of dispersion based on a coupled-cluster approach is available.[1,2] The error in the dispersion energy for the parallel-displaced benzene dimer is about 0.5 kcal/mol for SAPT2+3/aDZ as compared to SAPT2+3(CCD)/aDZ.[3] Our own tests on the S66 data set suggest that the MAE in dispersion energies between the SAPT2+(3)/aTZ and SAPT2+(3)(CCD)/aTZ methods is about 0.19 kcal/mol. The maximum discrepancy occurs for the π-stacked uracil dimer, for which the SAPT2+(3)/aTZ method overestimates the dispersion

---

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

energy by 0.8 kcal/mol as compared to SAPT2+(3)(CCD)/aTZ. Furthermore, SAPT2+(3)(CCD)/aTZ works slightly better than SAPT2+(3)/aTZ for the total binding energy. In comparison to the CCSD(T)/CBS benchmarks for S66 binding energies, the SAPT2+(3)(CCD)/aTZ method affords a MAE of 0.13 kcal/mol and the SAPT2+(3)/aTZ method affords a MAE of 0.20 kcal/mol. Due to some favorable error cancellation, however, the XSAPT(KS)+D method that is fit to SAPT2+(3) dispersion energies performs slightly better than the corresponding method whose dispersion potentials are fit to SAPT2+(3)(CCD) dispersion energies. For this reason, and in consideration of computational efficiency, we use SAPT2+(3)/aTZ as our dispersion benchmark for the purpose fo fitting the D3 dispersion potential. The sole exceptions to this protocol are the two anionic systems, $F^-(H_2O)$ and $Cl^-(H_2O)$, for which SAPT2+(3)(CCD)/aTZ is used as the benchmark.

The training set used to determine the D3 parameters consists of 74 dimers: 22 dimers from the S22 data set,[4] 22 dimers from the X40 data set,[5] 13 dimers containing divalent sulfur,[6] 8 dimers proposed by Zhao and Truhlar[7] in the KB49 data set[8] (without the H_2S dimer, which is already included as part of the divalent-sulfur data set), and 9 additional dimers optimized at the MP2/aTZ level, including $HCl\cdots H_2O$, $HCl\cdots NH_3$, $HF\cdots CH_3SH$, $HF\cdots H_2O$, $HF\cdots H_2S$, $HF\cdots HCl$, $HF\cdots NH_3$, $F^-\cdots H_2O$, and $Cl^-\cdots H_2O$. For each dimer, 5 different radial geometries corresponding to the same angular configuration were considered. For the 22 dimers from the X40 dataset and the 13 dimers containing divalent sulfur, each dimer contains five data points with relative displacements of 0.90, 1.00, 1.25, 1.50, and 2.00 with respect to the equilibrium geometry. For the rest of dimers, each dimer contains five data points with relative displacements of 0.9, 1.0, 1.2, 1.5, and 2.0 with respect to the equilibrium geometry. There are 370 training geometries in total.

The parameters of D3 were optimized using the least-squares method with deviations

$$\chi^2 = \frac{1}{370}\left[\sum_{i=1}^{370}\left(E_{\text{disp}}^{(i)}(\text{D3}) - E_{\text{disp}}^{(i)}(\text{SAPT})\right)^2\right] , \tag{S1}$$

where $E_{\text{disp}}^{(i)}$ is the dispersion energy for the ith dimer, computed either using the benchmark SAPT calculation or else the D3 dispersion potential. We used a genetic algorithm followed by simplex optimization to fit the parameters, and the final deviation was $\chi = 0.1377$ kcal/mol. Optimized values of $C_{6,i}$, $C_{8,i}$, and β_i are listed in Table S3.

4. NEW BENCHMARKS

Benchmark binding energies for $(H_2O)_6$, $(H_2O)_{20}$, $F^-(H_2O)_{n\leq 6}$, $Cl^-(H_2O)_{n\leq 6}$, and $F^-(H_2O)_{10}$ are shown in Tables S4, S5, S6, S7, and S8, respectively. The $(H_2O)_6$ geometries are available from Ref. 9, where they were optimized at the MP2/aug-cc-pVTZ level, and the $(H_2O)_{20}$ structures are available from Ref. 10, where they were optimized using the TIP4P force field. Coordinates for the remaining structures are available in this work, as a separate attachment. For $X^-(H_2O)_n$ structures with $n \leq 6$, the geometries were optimized at the RIMP2/aug-cc-pVTZ level; $F^-(H_2O)_{10}$ geometries were optimized at the B3LYP/6-31G* level.

5. SOFTWARE

All XSAPT, att-MP2, and DFT calculations were performed using a locally-modified copy of Q-Chem v. 4.2.[11] [The att-MP2 and XSAPT(KS)+D2 methods are available in the current release of Q-Chem, v. 4.2, and the XSAPT(KS)+D3 and sd-XSAPT methods will be released in v. 4.3 in 2015.] CCSD(T)-F12 calculations for $(H_2O)_6$, $F^-(H_2O)_{n\leq 6}$, and $Cl^-(H_2O)_{n\leq 6}$ were performed using Orca 3.0.2,[12] and the CCSD(T)-F12 and MP2-F12 calculations for $(H_2O)_{20}$ and $F^-(H_2O)_{10}$ were performed using Molpro 2012.1.[13] All terms in $\delta E_{\text{int}}^{\text{HF}}$, along with the CCSD(T) calculations for $X^-(H_2O)$ and the rest of the MP2 and CCSD(T) calculations, were performed using Psi4 v. beta5.[14] All supersystem calculations are counterpoise corrected, with the exception of the DFT and att-MP2 calculations, as well as the CCSD(T)-F12 and MP2-F12 calculations for $(H_2O)_{20}$ and $F^-(H_2O)_{10}$.

[1] H. L. Williams, K. Szalewicz, R. Moszynski, and B. Jeziorski, J. Chem. Phys. **103**, 4586 (1995).

[2] R. M. Parrish, E. G. Hohenstein, and C. D. Sherrill, J. Chem. Phys. **139**, 174102:1 (2013).

[3] E. G. Hohenstein and C. D. Sherrill, WIREs Comput. Mol. Sci. **2**, 304 (2012).

[4] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

[5] J. Řezác, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **8**, 4285 (2012).

[6] B. J. Mintz and J. M. Parks, J. Phys. Chem. A **116**, 1086 (2012).

[7] Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005).

[8] F. O. Kannemann and A. D. Becke, J. Chem. Theory Comput. **6**, 1081 (2010).

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

[10] J. K. Kazimirski and V. Buch, J. Phys. Chem. A **107**, 9762 (2003).

[11] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, et al., Mol. Phys. (2014).

[12] F. Neese, WIREs Comput. Mol. Sci. **2**, 73 (2012).

[13] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).

[14] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, et al., WIREs Comput. Mol. Sci. **2**, 556 (2012).

[15] K. Wang, W. Li, and S. Li, J. Chem. Theory Comput. **10**, 1546 (2014).

AO Basis Set	DCBS		Projected	
	$E^{(1)}_{\mathrm{elst}}$	$E^{(1)}_{\mathrm{exch}}$	$E^{(1)}_{\mathrm{elst}}$	$E^{(1)}_{\mathrm{exch}}$
cc-pVDZ	17.73	8.65	30.46	51.52
cc-pVTZ	10.05	3.98	18.38	29.51
cc-pVQZ	7.51	3.16	12.07	15.94
heavy-aug-cc-pVDZ	3.58	3.97	10.76	10.46
heavy-aug-cc-pVTZ	3.93	4.47	7.09	8.02
heavy-aug-cc-pVQZ	3.95	4.55	4.81	5.15
aug-cc-pVDZ	3.17	4.31	9.71	14.55
aug-cc-pVTZ	3.95	4.52	6.81	8.04
aug-cc-pVQZ	3.94	4.57	4.85	5.15
def2-SVP	16.91	8.77	30.15	50.69
def2-TZVPP	7.85	2.95	11.76	13.93
def2-QZVPP	5.40	4.01	6.80	6.99
def2-SVPD	5.46	4.74	23.53	24.21
def2-TZVPPD	3.97	4.52	9.42	7.01
def2-QZVPPD	4.05	4.56	7.11	6.50
aug-def2-TZVPP	3.90	4.50	4.63	4.69
aug-def2-QZVPP	3.94	4.55	3.37	3.86
heavy-aug-def2-TZVPP	3.91	4.48	4.52	4.53
heavy-aug-def2-QZVPP	4.00	4.55	3.64	3.98
Pople-def2-TZVPP	4.54	4.27	4.98	4.46
heavy-Pople-def2-TZVPP	4.58	4.25	5.09	4.72

TABLE S1: Mean absolute percentage errors (in kcal/mol) for the energy components $E^{(1)}_{\mathrm{elst}}$ and $E^{(1)}_{\mathrm{exch}}$ computed using SAPT(KS)/LRC-ωPBE, as compared to SAPT2+(3)/aTZ benchmarks. The test systems are $F^-(H_2O)$, $(H_2O)_2$, and the T-shaped and parallel-displaced isomers of $(C_6H_6)_2$. Results are shown for both the dimer-centered SAPT basis (DCBS) as well as the "projected" (pseudocanonicalized monomer-centered) basis, only the latter of which is appropriate for XSAPT.

Monomer	ω / a_0^{-1}	
	$\Delta\omega = 0.025\ a_0^{-1}$	$\Delta\omega = 0.005\ a_0^{-1}$
	—S22 molecules—	
adenine	0.275	—
2-aminopyridine	0.300	—
benzene	0.275	—
ethyne	0.400	—
ethene	0.350	—
methane	0.450	—
formamide	0.475	—
formic acid	0.425	—
water	0.500	0.485
HCN	0.450	—
indole	0.275	—
ammonia	0.450	—
phenol	0.275	—
pyrazine	0.375	—
2-pyridoxine	0.300	—
thymine	0.275	—
uracil	0.300	—
	—S66 molecules—	
methylamine	0.400	—
methanol	0.450	—
$AcNH_2$	0.450	—
AcOH	0.375	—
cyclopentane	0.450	—
neopentane	0.300	—
pentane	0.325	—
peptide	0.350	—
pyridine	0.325	—
	—ions—	
F^-	0.475	0.480
Cl^-	0.375	0.370

TABLE S2: Tuned values of the range separation parameter, ω, for various monomers, where the tuning was performed at the LRC-ωPBE/hpTZVPP level. In a few cases, a finer spacing of $\Delta\omega = 0.005\ a_0^{-1}$ was used to scan the $\varepsilon_{\mathrm{HOMO}}(\omega)$ and $-\mathrm{IP}(\omega)$ curves, and in these cases we used the more finely-tuned value of ω.

Element	C_6/ J nm^6 mol^{-1}	C_8/ J nm^8 mol^{-1}	β/ a_0^{-1}
H–C	0.0087	1.9024	0.3962
H–N	0.0670	2.4646	0.3245
H–O	0.0847	3.3700	0.3021
H–F	0.0272	0.8556	0.3904
H–S	0.0000	18.5511	0.2456
H–Cl	0.0001	4.1514	0.3122
C	2.4618	0.0109	4.4611
N	1.2111	0.0059	8.0263
O	0.5450	0.0010	10.4192
F	0.3468	0.0005	12.8513
S	14.1799	0.0075	8.4090
Cl	19.6013	0.0001	15.8640

TABLE S3: Fitting parameters that define the D3 dispersion potential for various elements. Note that the hydrogen parameters depend upon the atom to which it is bonded.

Isomer	Binding Energy / kcal mol^{-1}	
	CCSD(T)/CBS	MP2/CBS
prism	−48.31	−48.22
cage	−48.02	−48.13
book1	−47.61	−47.93
book2	−47.33	−47.65
cyclic chair	−46.52	−47.02
bag	−46.87	−47.19
cyclic boat1	−45.52	−46.04
cyclic boat2	−45.42	−45.91

TABLE S4: Binding energies isomers of $(H_2O)_6$, using geometries from from Ref. 9.

Isomer	Binding Energy / kcal mol^{-1}	
	CCSD(T)/CBS	MP2/CBS
1	−200.54	−197.97
2	−199.20	−196.68
3	−198.93	−196.51
4	−197.89	−195.02
5	−198.15	−195.58
6	−198.17	−195.24
7	−197.67	−194.62
8	−197.44	−194.90
9	−197.03	−194.29
10	−196.59	−193.92

TABLE S5: Binding energies isomers of $(H_2O)_{20}$, using geometries from from Ref. 15.

n	Binding Energy / kcal mol^{-1}	
	CCSD(T)/CBS	MP2/CBS
1	−32.31	−32.25
2	−52.27	−51.86
3	−70.14	−69.09
4	−85.24	−83.84
5	−101.09	−99.53
6	−116.58	−114.98

TABLE S6: Binding energies for optimized (MP2/aTZ) geometries of $F^-(H_2O)_n$.

n	Binding Energy / kcal mol^{-1}	
	CCSD(T)/CBS	MP2/CBS
1	−15.50	−15.75
2	−31.01	−31.30
3	−47.62	−47.74
4	−58.20	−58.40
5	−74.73	−74.94
6	−86.86	−87.16

TABLE S7: Binding energies for optimized (MP2/aTZ) geometries of $Cl^-(H_2O)_n$.

Isomer	Binding Energy / kcal mol^{-1}	
	CCSD(T)/CBS	MP2/CBS
1	−169.02	−166.97
2	−169.03	−166.98
3	−168.01	−165.93
4	−169.07	−167.06
5	−169.91	−167.98
6	−171.42	−169.68
7	−169.92	−167.99
8	−166.65	−164.34
9	−163.59	−161.25
10	−168.22	−166.47

TABLE S8: Binding energies for optimized (B3LYP/6-31G*) geometries of ten different isomers of $F^-(H_2O)_{10}$.