

Supporting Information for “Implementation of restricted and unrestricted versions of extended symmetry-adapted perturbation theory (XSAPT) in the atomic orbital basis, and benchmark calculations for large supramolecular complexes”

Ka Un Lao and John M. Herbert*

*Department of Chemistry and Biochemistry,
The Ohio State University, Columbus, OH 43210*

(Dated: December 29, 2017)

A. Tuned Values of the Range-Separation Parameter

Tuned values of ω for complexes in L7 and S12L, using both the ω_{IP} and ω_{GDD} tuning procedures, are shown in Tables S1 and S2. For the ω_{GDD} tuning procedure there is an additional parameter μ , which is a distance cutoff to define the region of localized valence molecular orbitals; see Ref. 1. Following that work, we use $\omega = 0.3 a_0^{-1}$ as an initial guess, then determine μ to normalize $\langle d_x^2 \rangle$. A series of molecules was used to determine the constant C in

$$\omega_{\text{GDD}} = C \langle d_x^2 \rangle^{-1/2}, \quad (\text{S1})$$

and subsequently ω_{GDD} and μ were determined for each molecule of interest (with C fixed), with results in Tables S1 and S2.

B. Basis Set Dependence

See Table S3. A pseudocanonicalized² triple- ζ basis affords results within 5% of SAPT2+(3)/aTZ benchmarks for the electrostatic and exchange energies of several challenging dimers, which is not significantly worse than quadruple- ζ results, and actually *better* than results using the traditional dimer-centered basis, which is ill-defined anyway for systems composed of more than two monomers. For reasons of cost, we therefore use the triple- ζ basis for all (X)SAPT calculations in this work.

C. Energies and Other Properties

Listed in Table S4 are binding energies for complexes in L7 and S12L, computed using XSAPT(KS)+D3(ω_{IP}) and XSAPT(KS)+D3(ω_{GDD}). Also shown are binding energies computed

* herbert@chemistry.ohio-state.edu

TABLE S1: Tuned values of the range separation parameter, ω (in units of a_0^{-1}) for the L7 data set, at two different LRC-DFT levels of theory. For ω_{GDD} tuning there is also a cutoff, μ (see Ref. 1) that is listed in parentheses. A spacing of $\Delta\omega = 0.025 a_0^{-1}$ was used to scan the $\varepsilon_{\text{HOMO}}(\omega)$ and $-\text{IP}(\omega)$ curves for ω_{IP} .

Complex	ω_{IP}^a	$\omega_{\text{GDD}} (\mu)^a$	ω_{IP}^b	$\omega_{\text{GDD}} (\mu)^b$
guanine	0.275	0.316 (0.072)	0.250	0.338 (0.075)
guanine dimer	0.250	0.289 (0.047)	0.225	0.313 (0.051)
cytosine	0.300	0.328 (0.082)	0.250	0.359 (0.083)
adenine	0.275	0.316 (0.078)	0.250	0.338 (0.080)
Watson-Crick GC	0.250	0.291 (0.051)	0.225	0.314 (0.055)
circumcoronene	0.175	0.246 (0.033)	0.150	0.266 (0.036)
coronene	0.225	0.271 (0.053)	0.175	0.289 (0.055)
octadecane	0.250	0.260 (0.038)	0.200	0.269 (0.040)
phenylalanine	0.250	0.290 (0.054)	0.200	0.308 (0.057)
phenylalanine dimer	0.225	0.268 (0.037)	0.200	0.285 (0.037)

^aLRC- ω PBE/hpTZVPP. ^bLRC- ω PBEh/6-31G(d,2p).

at the B97M-V/(h)aTZ levels of theory.

Components of the interaction energies for L7 and S12L are shown in Table S5, computed using the XSAPT(KS)+D3+ $E_{\text{disp},3\text{B}}^{\text{ATM(TS)}}$ /hpTZVPP(ω_{GDD}) method.

Deformation energies for the S12L complexes, computed at various levels of theory, are shown in Table S6.

Dipole moments for the polarized and unpolarized wave functions for molecules in the L7 and S12L data sets are listed in Tables S7 (for L7) and S8 (for S12L), where they are computed using XSAPT(KS)(ω_{GDD}).

¹ Modrzejewski, M.; Rajchel, L.; Chalasinski, G.; Szczesniak, M. M. *J. Phys. Chem. A* **2013**, *117*, 11580–11586.

² Jacobson, L. D.; Herbert, J. M. *J. Chem. Phys.* **2011**, *134*, 094118:1–17.

³ Heßelmann, A.; Korona, T. *J. Chem. Phys.* **2014**, *141*, 094107:1–17.

TABLE S2: Tuned values of the range separation parameter, ω (in units of a_0^{-1}) for the S12L data set, calculated at the level of LRC- ω BPBE/hpTZVPP. For ω_{GDD} tuning there is also a cutoff, μ (see Ref. 1) that is listed in parentheses. A spacing of $\Delta\omega = 0.025 a_0^{-1}$ was used to scan the $\varepsilon_{\text{HOMO}}(\omega)$ and $-\text{IP}(\omega)$ curves for ω_{IP} .

System	Complex	ω_{IP}	$\omega_{\text{GDD}} (\mu)$
2a/2b	host	0.175	0.248 (0.032)
2a	guest	0.225	0.296 (0.057)
2b	guest	0.275	0.313 (0.080)
3a/3b	host	0.175	0.251 (0.027)
3a	guest	0.200	0.296 (0.048)
3b	guest	0.250	0.315 (0.063)
4a/4b	host	0.225	0.234 (0.029)
4a	guest	0.200	0.243 (0.038)
4b	guest	0.200	0.237 (0.035)
5a/5b	host	0.300	0.236 (0.023)
5a	guest	0.600	0.332 (0.079)
5b	guest	0.400	0.336 (0.082)
6a/6b	host	0.425	0.256 (0.027)
6a	guest	0.375	0.339 (0.085)
6b	guest	0.375	0.360 (0.091)
7a/7b	host	0.375	0.248 (0.024)
7a	guest	0.200	0.288 (0.045)
7b	guest	0.275	0.304 (0.067)

TABLE S3: Mean absolute percentage errors^a for XSAPT(KS) calculations on $F^-(H_2O)$, $(H_2O)_2$, and the T-shaped and parallel-displaced isomers of $(C_6H_6)_2$, using the IP-tuned LRC- ω PBE functional.

Basis Set	MAPE (%)	
	$E_{\text{elst}}^{(1)}$	$E_{\text{exch}}^{(1)}$
dimer-centered TZ	10.0	4.0
pseudocanonical ^b hpTZVPP ^c	5.1	4.7
pseudocanonical ^b aQZVPP ^d	3.4	3.9

^aWith respect to SAPT2+(3)/aTZ benchmarks.

^bEquivalent to the “projected” basis of Ref. 2.

^cBasis set is def2-TZVPP augmented with diffuse functions on non-hydrogen atoms, taken from Pople basis sets.

^dBasis is def2-QZVPP augmented with diffuse functions taken from aug-cc-pVQZ.

TABLE S4: Binding energies (in kcal/mol) for L7 and S12L complexes computed at various XSAPT and DFT levels of theory. The hpTZVPP basis set is used for XSAPT(KS)+D3. The aTZ basis set is used for DFT calculations except the B97M-V calculations use the haTZ basis set for **4a**, **4b**, **7a**, and **7b**. Deformation energies for DFT methods are computed using the corresponding functional. Deformation energies for XSAPT-based methods are computed using the NLDFT functional.

System	XSAPT(KS)+D3		B97M-V	B97M-V (CP)
	(ω_{IP})	(ω_{GDD})		
—L7—				
GGG	-2.50	-2.60	-2.33	-1.94
C2C2PD	-26.06	-26.39	-22.32	-21.54
GCGC	-15.68	-16.06	-15.46	-14.51
PHE	-23.36	-23.96	-25.87	-25.21
CBH	-9.45	-9.55	-12.40	-11.72
C3A	-20.11	-20.46	-17.47	-16.56
C3GC	-33.03	-33.56	-31.05	-29.27
—S12L—				
2a	-39.24	-40.92	-32.85	-31.23
2b	-25.80	-26.73	-21.88	-20.58
3a	-27.11	-28.80	-29.85	-27.11
3b	-23.02	-24.28	-22.88	-20.95
4a	-49.96	-50.18	-40.54	-37.60
4b	-51.35	-51.66	-42.58	-39.38
5a	-44.86	-38.21	-36.70	-35.16
5b	-26.60	-24.67	-25.67	-24.30
6a	-94.47	-90.35	-87.07	-85.79
6b	-90.19	-86.49	-82.58	-81.49
7a	-146.54	-142.60	-142.09	-139.85
7b	-38.51	-37.72	-34.74	-33.05

TABLE S5: Interaction energy components for L7 and S12L complexes computed using the XSAPT(KS)+D3+ $E_{\text{disp},3\text{B}}^{\text{ATM(TS)}}$ /hpTZVPP(ω_{GDD}) method, in kcal/mol. The interaction types are dispersion-dominated (“D”), electrostatics-dominated (“E”), or of mixed influence (dispersion and electrostatics, “M”).

System	Elst	Exch	Ind	Disp	Total ^a	Type
—L7—						
GGG	2.70	6.52	-1.29	-9.76	-1.84	D
C2C2PD	-7.85	26.65	-2.46	-36.54	-20.20	D
GCGC	-8.94	24.62	-1.93	-27.50	-13.74	D
PHE	-27.56	31.93	-11.01	-16.58	-23.22	M
CBH	-3.54	14.75	-1.24	-18.07	-8.09	D
C3A	-7.58	19.95	-1.96	-27.03	-16.62	D
C3GC	-13.08	37.49	-4.42	-46.95	-26.96	D
—S12L—						
2a	-27.23	53.17	-11.84	-52.56	-38.45	M
2b	-17.18	36.60	-6.01	-37.76	-24.35	D
3a	-22.41	58.02	-7.29	-62.68	-34.36	D
3b	-18.77	38.05	-5.83	-44.39	-30.93	D
4a	-26.95	69.02	-7.23	-75.02	-40.17	D
4b	-28.11	73.25	-8.17	-78.65	-41.68	D
5a	-48.50	69.21	-20.36	-43.10	-42.75	M
5b	-33.53	54.73	-12.43	-37.04	-28.27	M
6a	-68.15	39.00	-24.80	-37.54	-91.49	M
6b	-65.78	33.26	-24.69	-31.78	-88.99	E
7a	-104.05	57.97	-24.43	-69.33	-139.84	M
7b	-10.07	29.28	-3.05	-48.33	-32.18	D

^aRelaxation energies are not included.

TABLE S6: Deformation energies for complexes in S12L datasets computed at the NLDFT/def2-QZVP, MP2/CBS, SCS-MP2/CBS, and B97M-V/aTZ levels. The B97M-V results are new; the rest are from Ref. 3.

System	deformation energy (kcal/mol)			
	NLDFT	MP2	SCS-MP2	B97M-V
2a	3.08	3.25	3.12	3.56
2b	1.66	2.03	1.82	2.08
3a	12.70	23.63	17.67	15.11
3b	11.06	21.34	15.60	13.86
4a	1.24	1.24	1.21	0.99
4b	2.24	2.56	2.55	1.72
5a	6.58	8.05	7.21	7.60
5b	5.72	6.49	5.66	6.21
6a	4.90	4.09	4.37	5.01
6b	5.45	4.53	4.86	5.62
7a	7.30	0.48	2.49	6.08
7b	1.08	0.14	0.33	0.21

TABLE S7: Dipole moments (in Debye) for molecules in the L7 data set computed at the XSAPT(KS)/hpTZVPP(ω_{GDD}) level.

Complex	Fragment	Unpolarized	Polarized	Difference
GGG	guanine	6.98	6.29	-0.68
	guanine dimer	13.31	12.51	-0.80
C2C2PD	coronene	0.00	0.72	0.72
	coronene	0.00	0.72	0.72
GCGC	Watson-Crick guanine-cytosine	5.94	6.57	0.63
	Watson-Crick guanine-cytosine	5.94	6.57	0.62
PHE	phenylalanine	2.12	3.25	1.13
	phenylalanine dimer	2.46	3.76	1.29
CBH	octadecane	0.00	0.05	0.05
	octadecane	0.00	0.05	0.05
C3A	adenine	2.62	2.79	0.17
	circumcoronene	0.36	1.11	0.76
C3GC	Watson-Crick guanine-cytosine	6.27	6.55	0.28
	circumcoronene	0.09	2.48	2.39

TABLE S8: Dipole moments (in Debye) for molecules in the L7 data set computed at the XSAPT(KS)/hpTZVPP(ω_{GDD}) level.

	Complex Fragment	Unpolarized	Polarized	Difference
2a	host	1.36	1.66	0.30
	guest	0.51	0.60	0.08
2b	host	1.22	1.67	0.45
	guest	0.10	0.48	0.38
3a	host	15.44	16.43	1.00
	guest	1.22	0.95	-0.27
3b	host	15.71	17.35	1.64
	guest	4.92	6.29	1.37
4a	host	1.48	1.52	0.04
	guest	0.02	1.30	1.28
4b	host	1.54	1.49	-0.05
	guest	0.01	1.39	1.38
5a	host	0.71	0.79	0.08
	guest	0.23	0.28	0.05
5b	host	0.71	0.82	0.11
	guest	0.02	0.05	0.03
6a	host	0.35	6.58	6.24
	guest	13.28	14.29	1.01
6b	host	0.26	6.62	6.36
	guest	12.78	13.94	1.16
7a	host	0.11	0.29	0.17
	guest	1.25	1.50	0.25
7b	host	0.08	0.12	0.04
	guest	1.74	1.66	-0.07