Supplementary Material for "Self-consistent charge embedding at very low cost, with application to symmetry-adapted perturbation theory"

Kuan-Yu Liu, Kevin Carter-Fenk, and John M. Herbert*
Department of Chemistry and Biochemistry
The Ohio State University, Columbus, Ohio 43210 USA

June 27, 2019

1 Reparameterization of Many-Body Dispersion for XSAPT + MBD

We recently published[1] a dispersion correction for SAPT based on a modified version of the self-consistent, many-body dispersion (MBD) method of Tkatchenko and coworkers.[2] The MBD procedure begins with a volume-based scaling of the vacuum atomic polarizabilities, using Hirshfeld atom-in-molecule atomic volumes, which accounts for changes due to the local molecular environment. These volume-adjusted polarizabilities are then self-consistently screened at short range, resolving the nonadditive interaction of each atomic polarizability with the dynamic electric field of the rest of the system. We have since corrected an error in our original parameterization[1] that resulted in the use of volume-adjusted quantities corresponding to dipole polarizabilities, oscillator frequencies, and van der Waals radii in the effectively-screened dipole-quadrupole (esDQ) correction that is described in Ref. 1, rather than the self-consistently-screened versions of these quantities. The latter should be used for maximum compatibility of the esDQ and MBD corrections. To reflect this change, we have reparameterized the MBD+esDQ potential using only self-consistently screened quantities.

Our modification to MBD incorporates esDQ response functions by deriving the C_8 coefficients from the self-consistently screened C_6 coefficients of the MBD procedure. The esDQ potential,

$$E_{disp}^{esDQ} = s_8 \sum_{i>j}^{\text{atoms}} f(R_{ij}) \frac{C_{8,es}^{ij}}{R_{ij}^8} , \tag{1}$$

acts as a short-range correction to the MBD energy. The Tang-Toennies damping function,

$$f(R_{ij}) = 1 - e^{-(a_1 R_{vdW}^{ij} + a_2)R_{ij}} \sum_{k=1}^{8} \frac{\left[(a_1 R_{vdW}^{ij} + a_2)R_{ij}\right]^k}{k!} , \tag{2}$$

depends on parameters a_1 and a_2, and the self-consistently screened van der Waals contact distance $R_{vdW}^{ij} = R_{vdW}^{i} + R_{vdW}^{j}$.

In our original implementation, the pairwise C_8^{AB} coefficients were derived from their atomic counterparts using the quantum Drude oscillator (QDO) expression,

$$C_{8,\text{es}}^{AB} = \frac{3}{2}\left[\frac{C_{8,\text{es}}^{AA} C_{8,\text{es}}^{BB}}{(\alpha_1^B/\alpha_1^A) C_{8,\text{es}}^{AA} + 2(\alpha_2^A/\alpha_2^B) C_{8,\text{es}}^{BB}} + \frac{C_{8,\text{es}}^{AA} C_{8,\text{es}}^{BB}}{(\alpha_1^A/\alpha_1^B) C_{8,\text{es}}^{BB} + 2(\alpha_2^B/\alpha_2^A) C_{8,\text{es}}^{AA}}\right] , \tag{3}$$

---

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

Table S1: Results of the MBD parameterization.

Data	Parameters				Error (kcal/mol)[a]		
Set	β	a_1	a_2	s_8	MAD	(%)	RMSD
SXπ	0.497	0.198	0.428	0.798	0.184	(5.0%)	0.287

[a] Mean absolute deviation (MAD) and root-mean square deviation (RMSD) are in kcal/mol. Mean absolute percent deviation is given in parenthesis.

where α_1 and α_2 are atomic dipole and quadrupole polarizabilities, respectively. Using the relationship

$$C_8^{AA} = 5\alpha_1^A \alpha_2^A \omega_A \ , \tag{4}$$

Eq. (3) can also be written in terms of the oscillator frequencies,

$$C_{8,\mathrm{es}}^{AB} = \frac{15}{2}\left[\alpha_1^A \alpha_2^B \frac{\omega_A \omega_B}{(\omega_A + 2\omega_B)} + \alpha_2^A \alpha_1^B \frac{\omega_A \omega_B}{(2\omega_A + \omega_B)}\right] . \tag{5}$$

As Eqs. (3) and (5) are equivalent, and the relationship between the C_8 coefficients and the oscillator frequencies is set by Eq. (4), accurately representing the quadrupole polarizabilities becomes the key driving force in reducing errors in the parameterization.

In the new parameterization, the quadrupole polarizabilities are directly obtained from the QDO relationship:

$$\alpha_2^A = \frac{3}{4}\frac{\gamma \alpha_1^A}{\omega_A^2} \ , \tag{6}$$

Here, the scaling factor $\gamma = s_h \mathrm{erf}(Z_A - 1) + s_l$ is capable of treating the light elements separately, reducing the errors for hydrogen and helium significantly. The parameters $s_h = 2.03$ and $s_l = 0.8141$ were fit to minimize percentage error of quadrupole polarizabilities for hydrogen and the noble gases. The resultant quadrupole polarizabilities are used directly in Eq. (5).

The new parameters are given in Table S1. As in previous work, [1] these were fit to minimize the mean absolute percentage error (MAPE) against the SXπ data set that is described in Ref. 1. We use the same form of switching function as in previous work,

$$f_{\mathrm{MBD}}(R_{ij}) = \frac{1}{1 + e^{-6(R_{ij}/\beta R_{vdW}^{ij} - 1)}} \ , \tag{7}$$

where β is a range-separation parameter to be determined.

The new parameterization was tested against three of the large molecule datasets from Ref. 1, with nearly the same results. The error statistics are given in detail in Table S2. In general, the method performs very nearly the same with the new parameters, but these parameters are now more consistent with the underlying MBD model.

Table S2: Benchmark results for the new MBD parameters in large-molecule datasets.

Data	Error (kcal/mol)			
Set	MAD	Max	ΔMAD[a]	ΔMax[a]
L7[b]	1.04	1.56	−0.01	−0.59
S30L∩S12L[c,d]	4.34	11.00	0.18	0.64
S30L[c,d]	6.34		0.44	

[a] Changes in MAD and maximum error, relative to the parameterization reported in Ref. 1. Negative values indicate that the error is smaller for the new parameterization. [b] Relative to DLPNO-CCSD(T)/CBS `TightPNO` benchmarks from Ref. 1. [c] Relative to back-corrected experimental binding energies from Ref. 3. [d] Includes SCS-MP2/CBS geometric deformation energies from Ref. 1.

2 Additional Data

Table S3: Errors in interaction energies for the S22 data set, at the XSAPT(KS)/hpTZVPP level.

	error (kcal/mol)	
	CM5	ChElPG
1 2-pyridoxine 2-aminopyridine complex	0.22	0.20
2 Adenine thymine complex stack	−0.19	−0.19
3 Adenine thymine Watson-Crick complex	0.94	0.92
4 Ammonia dimer	0.15	0.15
5 Benzene - Methane complex	0.08	0.08
6 Benzene ammonia complex	−0.11	−0.11
7 Benzene dimer parallel displaced	0.10	0.10
8 Benzene dimer T-shaped	0.02	0.02
9 Benzene HCN complex	−0.82	−0.82
10 Benzene water complex	−0.31	−0.31
11 Ethene dimer	0.10	0.07
12 Ethene ethyne complex	0.42	0.42
13 Formamide dimer	−1.12	−1.15
14 Formic acid dimer	−1.11	−1.09
15 Indole benzene complex stack	0.60	0.58
16 Indole benzene T-shape complex	−0.19	−0.19
17 Methane dimer	0.14	0.14
18 Phenol dimer	0.36	0.35
19 Pyrazine dimer	−0.08	−0.08
20 Uracil dimer h-bonded	0.97	0.97
21 Uracil dimer stack	−0.28	−0.29
22 Water dimer	−0.01	−0.01
MUE	0.38	0.38

Table S4: Errors in interaction energies for the S66 data set, at the XSAPT(KS)/hpTZVPP level.

	error (kcal/mol)	
	CM5	ChElPG
1 Water ... Water	−0.14	−0.18
2 Water ... MeOH	−0.10	−0.11
3 Water ... MeNH2	0.13	0.09
4 Water ... Peptide	−0.01	−0.02
5 MeOH ... MeOH	−0.07	−0.07
6 MeOH ... MeNH2	0.01	0.02
7 MeOH ... Peptide	0.00	0.06
8 MeOH ... Water	−0.06	−0.14
9 MeNH2 ... MeOH	0.08	0.08
10 MeNH2 ... MeNH2	−0.04	−0.04
11 MeNH2 ... Peptide	−0.03	0.02
12 MeNH2 ... Water	−0.03	−0.01
13 Peptide ... MeOH	−0.01	0.01
14 Peptide ... MeNH2	−0.09	−0.06
15 Peptide ... Peptide	0.19	0.25
16 Peptide ... Water	0.10	0.07
17 Uracil ... Uracil (BP)	1.04	1.15
18 Water ... Pyridine	−0.10	−0.12
19 MeOH ... Pyridine	−0.15	−0.08
20 AcOH ... AcOH	−0.28	−0.22
21 AcNH2 ... AcNH2	−0.78	−0.63
22 AcOH ... Uracil	0.38	0.39
23 AcNH2 ... Uracil	−0.19	−0.19
24 Benzene ... Benzene (pi-pi)	−0.07	−0.07
25 Pyridine ... Pyridine (pi-pi)	−0.44	−0.43
26 Uracil ... Uracil (pi-pi)	−0.46	−0.40
27 Benzene ... Pyridine (pi-pi)	−0.20	−0.20
28 Benzene ... Uracil (pi-pi)	0.23	0.21
29 Pyridine ... Uracil (pi-pi)	0.09	0.08
30 Benzene ... Ethene	0.36	0.39
31 Uracil ... Ethene	0.07	0.04
32 Uracil ... Ethyne	0.29	0.23
33 Pyridine ... Ethene	0.23	0.25

Table S5: Continuation of Table S9.

	error (kcal/mol)	
	CM5	ChElPG
34 Pentane ... Pentane	0.81	0.85
35 Neopentane ... Pentane	0.82	0.86
36 Neopentane ... Neopentane	0.79	0.83
37 Cyclopentane ... Neopentane	0.55	0.59
38 Cyclopentane ... Cyclopentane	0.22	0.25
39 Benzene ... Cyclopentane	−0.30	−0.28
40 Benzene ... Neopentane	0.09	0.13
41 Uracil ... Pentane	−1.03	−0.95
42 Uracil ... Cyclopentane	−1.11	−1.07
43 Uracil ... Neopentane	−0.54	−0.49
44 Ethene ... Pentane	0.57	0.55
45 Ethyne ... Pentane	0.15	0.14
46 Peptide ... Pentane	0.11	0.16
47 Benzene ... Benzene (TS)	0.06	0.07
48 Pyridine ... Pyridine (TS)	−0.44	−0.42
49 Benzene ... Pyridine (TS)	−0.10	−0.11
50 Benzene ... Ethyne (CH-pi)	−0.23	−0.26
51 Ethyne ... Ethyne (TS)	0.12	0.05
52 Benzene ... AcOH (OH-pi)	0.13	0.15
53 Benzene ... AcNH2 (NH-pi)	0.15	0.22
54 Benzene ... Water (OH-pi)	−0.31	−0.30
55 Benzene ... MeOH (OH-pi)	−0.16	−0.18
56 Benzene ... MeNH2 (NH-pi)	−0.10	−0.11
57 Benzene ... Peptide (NH-pi)	−0.31	−0.30
58 Pyridine ... Pyridine (CH-N)	0.26	0.28
59 Ethyne ... Water (CH-O)	0.02	−0.05
60 Ethyne ... AcOH (OH-pi)	0.65	0.64
61 Pentane ... AcOH	0.18	0.23
62 Pentane ... AcNH2	0.00	0.04
63 Benzene ... AcOH	−0.02	−0.01
64 Peptide ... Ethene	0.18	0.18
65 Pyridine ... Ethyne	−0.04	−0.04
66 MeNH2 ... Pyridine	−0.50	−0.47
MUE	0.26	0.27

Table S6: Errors in interaction energies for the IHB data set, at the XSAPT(KS)/def2-TZVPPD level.

	error (kcal/mol)	
	CM5	ChElPG
1 acetate ... methanol	0.82	−1.54
2 acetate ... water	1.52	−3.34
3 acetate ... methylamine	1.24	0.50
4 methylammonium ... formaldehyde	−0.95	−1.48
5 methylammonium ... methylamine	−3.29	−4.22
6 methylammonium ... methanol	−0.90	−1.49
7 methylammonium ... water	−0.22	−0.56
8 guanidinium ... formaldehyde	0.72	0.39
9 guanidinium ... methylamine	−0.30	−0.31
10 guanidinium ... methanol	1.48	1.29
11 guanidinium ... water	1.20	1.28
12 imidazolium ... formaldehyde	−0.45	−1.48
13 imidazolium ... methylamine	−2.39	−5.27
14 imidazolium ... methanol	−0.61	−1.53
15 imidazolium ... water	−0.03	−0.69
MUE	1.08	1.69

Table S7: Errors in interaction energies for the AHB21 data set, at the XSAPT(KS)/def2-TZVPPD level.

	error (kcal/mol)	
	CM5	ChElPG
1	1.87	1.13
2	0.75	0.09
3	−2.41	−2.72
4	0.71	0.55
5	0.24	0.02
6	1.09	−0.25
7	1.71	−0.28
8	−0.50	−3.37
9	2.71	2.46
10	1.33	−1.17
11	0.14	−0.38
12	−0.26	−0.91
13	−0.39	−2.50
14	−1.06	−2.01
15	1.10	0.60
16	0.98	0.08
17	1.28	0.91
18	1.29	0.33
19	1.38	0.47
20	1.53	0.37
21	1.53	−1.85
MUE	1.15	1.07

Table S8: Errors in interaction energies for the CHB6 data set, at the XSAPT(KS)/def2-TZVPPD level.

	error (kcal/mol)	
	CM5	ChElPG
22	0.70	1.91
23	0.89	1.29
24	0.75	0.80
25	−2.93	−1.83
26	−0.27	0.20
27	−0.28	−0.68
MUE	0.97	1.12

Table S9: Errors in interaction energies for the IL16 data set, at the XSAPT(KS)/def2-TZVPPD level.

	error (kcal/mol)	
	CM5	ChElPG
008	−0.11	−6.80
144	−1.30	−4.70
147	−1.15	−6.68
148	−3.72	−11.57
150	−0.93	−6.17
152	−0.98	−6.73
187	−2.90	−13.83
202	−3.79	−15.06
212	−3.71	−13.62
213	−1.52	−6.26
214	−1.38	−5.09
227	−2.94	−9.95
228	−0.36	−3.71
229	−0.86	−4.84
230	−2.00	−2.74
231	−2.81	−5.58
MUE	1.90	7.71

Table S10: Errors in interaction energies for the S30L data set, at the XSAPT(KS)/def2-TZVPPD level.

	error (kcal/mol)	
	CM5	ChElPG
1 TCNQ@tweezer	−3.47	−3.21
2 DCB@tweezer	−1.14	−0.98
3 TCNB@pincer	8.11	8.00
4 NBD@pincer	4.57	4.10
5 TNF@tweezer2	−2.23	−2.23
6 TCNQ@tweezer2	0.76	0.87
7 5CPPA@8CPPA	−4.13	−6.55
8 6CPPA@9CPPA	−7.96	−9.76
9 C60@catcher	−5.69	2.48
10 C70@catcher	−4.29	5.23
11 C60@CA10	−3.76	−3.61
12 C70@CA10	−3.44	−1.16
13 morpholine@RA4	3.01	1.03
14 tioxane@RA4	1.44	−1.09
15 TMPDA@XB-donor	−2.60	−0.90
16 HHTAP@XB-donor	9.00	7.94
17 BQ@mcycle	0.87	0.84
18 GLH@mcycle	1.65	1.82
19 C5H9OH@β-CD	1.77	2.17
20 C8H15OH@β-CD	−0.30	0.22
21 AdOH@CB7	−8.19	−5.70
22 DAAD@ADDA	15.30	15.63
23 AAAA@DDDD$^+$	6.33	8.26
24 $Ad_2(NMe_3)_2$@CB7	−6.39	−11.03
25 tetraphene@Ex^{2}Box	−6.43	−6.42
26 chrysene@Ex^{2}Box	−6.57	−7.75
27 BuNH4$^+$@CB6	1.41	2.42
28 PrNH4$^+$@CB6	2.80	3.66
29 acetate@CP4	1.01	3.21
30 benzoate@CP4	1.76	3.65
MUE	4.21	4.68

References

[1] K. Carter-Fenk, K. U. Lao, K.-Y. Liu, and J. M. Herbert, “Accurate and efficient *ab initio* calculations for supramolecular complexes: Symmetry-adapted perturbation theory with many-body dispersion”, *J. Phys. Chem. Lett.*, **10**, 2706–2714 (2019).

[2] J. Hermann, R. A. DiStasio Jr., and A. Tkatchenko, “First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications”, *Chem. Rev.*, **117**, 4714–4758 (2017).

[3] R. Sure and S. Grimme, “Comprehensive benchmark of association (free) energies of realistic host–guest complexes”, *J. Chem. Theory Comput.*, **11**, 3785–3801 (2015). Erratum: *ibid.* **11**, 5990 (2015).