Supporting Information for:

Convergent Protocols for Computing Protein–Ligand Interaction Energies Using Fragment-Based Quantum Chemistry

Paige E. Bowling,[1,2], Dustin R. Broderick,[2] and John M. Herbert[1,2*]

[1]*Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210 USA*

[2]*Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210 USA*

December 1, 2024

List of Figures

List of Tables

*herbert@chemistry.ohio-state.edu

Table S1: Number of atoms for each supersystem.

System	Data Set	No. Atoms
181L	T4L	2,636
4W54	T4L	2,686
1LI2	T4L	2,637
3HUA	T4L	2,660
1O48	LIDS	1,781
1MMQ	LIDS	2,637
1ZP5	LIDS	2,516
1BOZ	LIDS	3,124

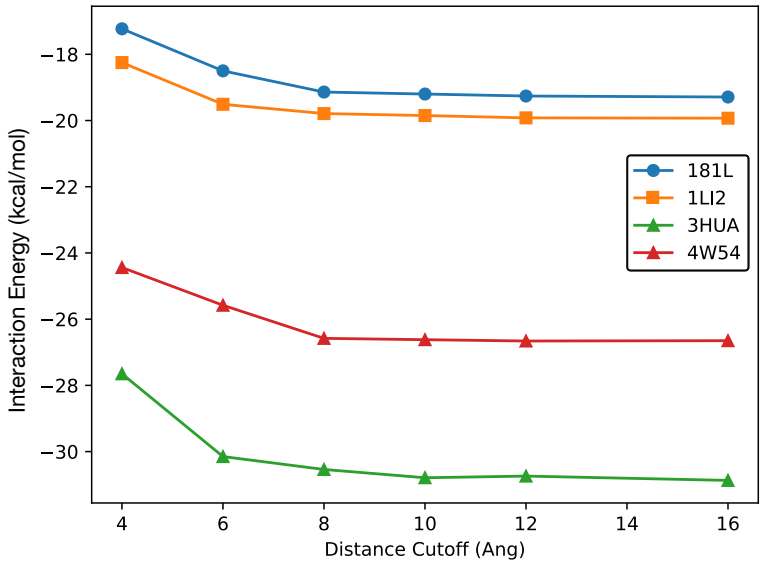

Fig. S1: Interaction energies ΔE_{int} for the T4 lysozyme complexes, computed using the MBE(2) approximation at the HF-3c level as a function of the distance cutoff R_{cut}.

Table S2: Absolute error as a function of R_{cut} for MBE(2) calculations on the T4 lysozymes, using HF-3c.

System	R_{cut} (Å)	Error[a] (kcal/mol)
181L	6	0.94
	8	0.30
	10	0.24
	12	0.18
4W54	6	0.02
	8	0.99
	10	1.02
	12	1.06
1LI2	6	0.67
	8	0.95
	10	1.01
	12	1.08
3HUA	6	0.15
	8	0.54
	10	0.79
	12	0.74

[a]With respect to a full-system calculation using HF-3c.

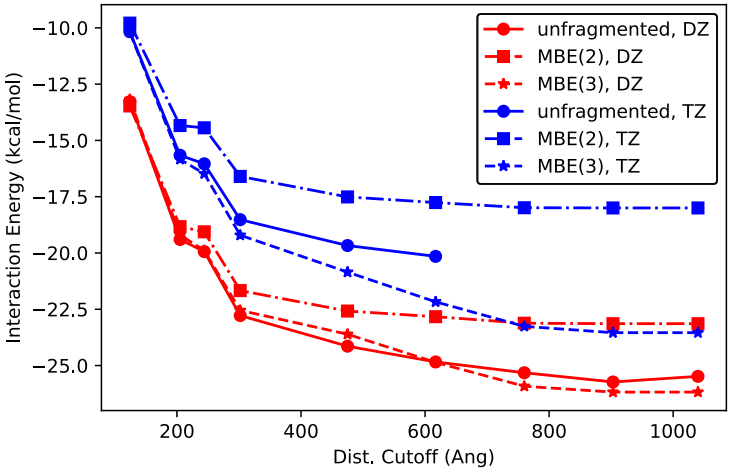

Fig. S2: Interaction energy for the benzene ligand in 1LI2, computed using radial enzyme models of increasing size. Calculations were performed using ωB97X-V in conjunction with either the def2-ma-SVP basis set (labeled "DZ" in the figure) or else the the def2-ma-TZVP basis set ("TZ"). MBE(n) calculations use $R_{\text{cut}} = 8$ Å. The "unfragmented" result is a conventional supramolecular calculation using eq. 6.

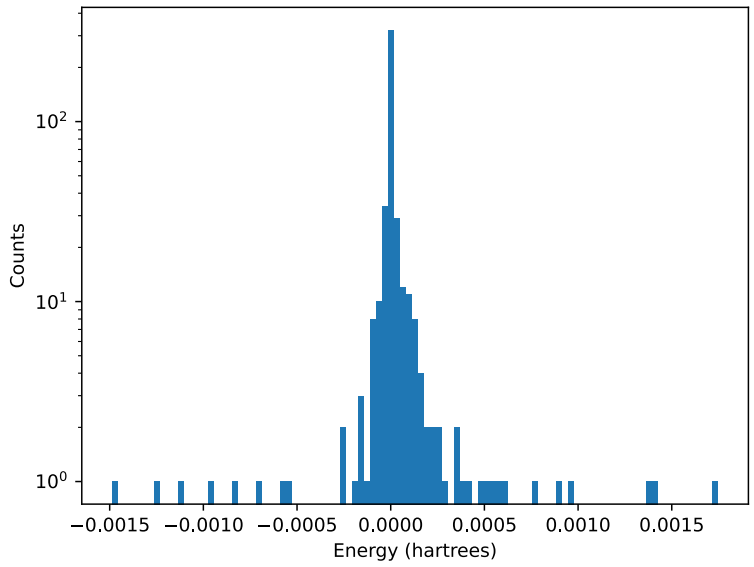

Fig. S3: Three-body corrections ΔE_{IJK} for 1O48, computed using HF-3c.

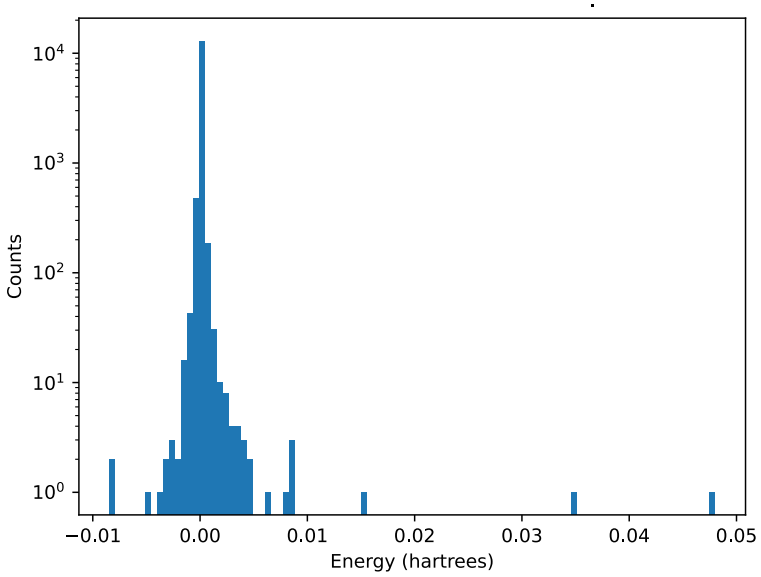

Fig. S4: Three-body corrections ΔE_{IJK} for 1ZP5, computed using HF-3c.

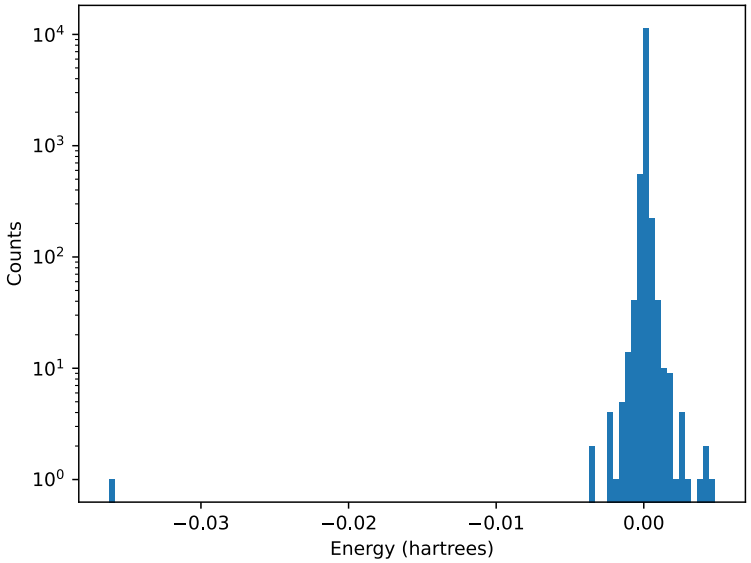

Fig. S5: Three-body corrections ΔE_{IJK} for 1MMQ, computed using HF-3c.

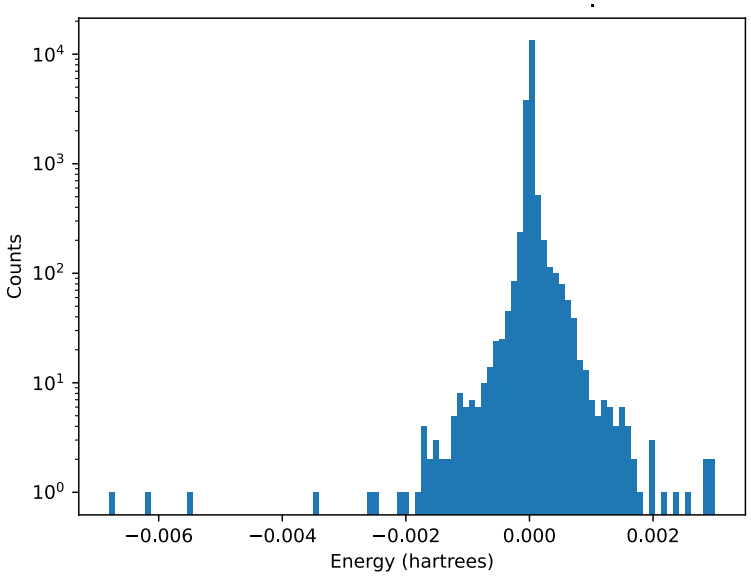

Fig. S6: Three-body corrections ΔE_{IJK} for 1BOZ, computed using HF-3c.

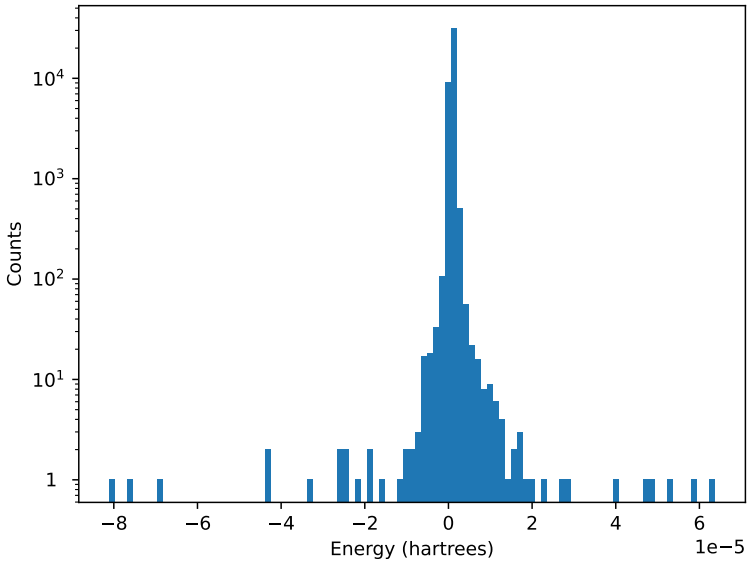

Fig. S7: Four-body corrections ΔE_{IJKL} for 1BOZ, computed using HF-3c with a 12 Å cutoff.

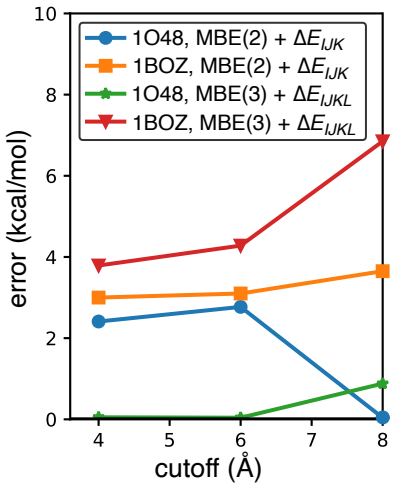

Fig. S8: Absolute errors for MBE(3) and MBE(4) calculations of ΔE_{int} at the HF-3c level, using a variable value of R_{cut} for the highest-order terms. In each case, R_{cut} = 8 Å for the two-body interactions. For the calculations labeled "MBE(2) + ΔE_{IJK}", MBE(2) is augmented with all three-body terms up to the indicated value of R_{cut}. For calculations labeled "MBE(3) + ΔE_{IJKL}", MBE(3) is performed with R_{cut} = 8 Å for both the two- and three-body terms, then augmented with four-body terms up to the indicated value of R_{cut}.

Table S3: Absolute error as a function of R_{cut} for MBE(2) calculations on the large inhibitor data set, using HF-3c.

System	R_{cut} (Å)	Error[a] (kcal/mol)
1O48	6	0.66
	8	1.29
	10	0.01
	12	0.19
1MMQ	6	1.25
	8	0.49
	10	0.35
	12	0.54
1ZP5	6	0.26
	8	0.40
	10	0.18
	12	0.34
1BOZ	6	3.48
	8	3.74
	10	2.23
	12	0.99

[a] With respect to a full-system calculation using HF-3c.

Table S4: Supersystem corrections δ_{frag} for the large inhibitor data set, evaluated at the HF-3c level.

System	Method	δ_{frag} (kcal/mol)	CPU time (hours)[a]
1O48	MBE(2)	3.77	871
	MBE(3)	−0.04	1,094
	MBE(4)	−1.34	1,475
1MMQ	MBE(2)	−21.35	3,179
	MBE(3)	1.04	3,672
	MBE(4)	−0.18	6,575
1ZP5	MBE(2)	0.09	5,664
	MBE(3)	−28.67	6,273
	MBE(4)	−3.61	10,214
1BOZ	MBE(2)	5.51	5,056
	MBE(3)	3.65	5,623
	MBE(4)	0.45	10,525

[a] Time required to evaluate $\delta_{\text{frag}} = E_{\text{low}}^{\text{super}} - E_{\text{low}}^{\text{MBE}(n)}$ (eq. 7) at the HF-3c level.