Supporting Information for: "Detection and correction of delocalization errors for electron- and hole-polarons using density-corrected density functional theory"

Bhaskar Rana, Marc P. Coons, and John M. Herbert*
Department of Chemistry & Biochemistry
The Ohio State University, Columbus, Ohio 43210 USA

June 1, 2022

List of Tables

*herbert@chemistry.ohio-state.edu

List of Figures

Tests of the analytic gradient implementations for DC-DFT are shown in Tables S1–S4, where we provide an element-by-element comparison between analytic and finite-difference gradients for several different systems. The finite-difference values were computed using a five-point stencil, *i.e.*, using displacements $\pm\Delta x$ and $\pm 2\Delta x$. We have found this to be necessary in order to reduce the finite-difference errors to the point where good agreement with analytic results can be obtained.[3] The mean absolute deviation between the analytic and finite-different gradient components is 1.84×10^{-7} a.u., demonstrating the validity of the implementation.

As a further test, we performed a short *ab initio* molecular dynamics (MD) simulation of $(H_2O)_5^+$ (Fig. S1). Energy conservation in MD simulations is extremely sensitive to consistency between energy and forces, and small errors in the gradient generally lead to rapid energy drift in simulations.[4] In contrast, energy fluctuations in Fig. S1 are quite stable at the level of $< 1.5 \times 10^{-4}$ Ha, even while the cluster is equilibrating.

For the further testing of this methodology we computed the barrier heights for the same set of reactions that was used for testing DC-DFT barrier heights in Ref. 1; see Table S5. For DC-BLYP, the mean absolute error (MAE) with respect to CCSD(T) barrier heights is 3.0 kcal/mol, which is a significant improvement over both self-consistent BLYP (MAE = 9.8 kcal/mol with all barrier heights underestimated) and also HF theory (MAE = 11.9 kcal/mol). These DC-BLYP results are in good agreement with the analogous data presented in Ref. 1.

All MP2 calculations reported herein were performed within the resolution-of-identity (RI) approximation. The auxiliary (density fitting) basis set designed for def2-SVP[5] is used in conjunction with the target basis set, 6-31G*.

Table S1: Differences (in hartree/bohr) between analytic and finite-difference gradient components computed at two different levels of theory, for a hemibonded complex $(H_2O)_2^+$. Finite difference calculations were performed using a five-point stencil. The maximum absolute deviation is 3.13×10^{-7} a.u..

Atom	BLYP/6-31G*			B3LYP/6-31G*		
	x	y	z	x	y	z
O	−3.135e-07	7.893e-08	−3.517e-08	−2.623e-07	−6.314e-09	−1.259e-08
H	8.310e-08	−1.222e-07	8.276e-08	7.196e-08	−1.077e-07	4.892e-08
H	6.253e-08	−8.962e-08	3.587e-09	7.190e-08	1.687e-08	−6.261e-08
O	2.266e-07	−1.725e-07	−1.723e-08	2.181e-07	−1.994e-07	−1.286e-08
H	−6.210e-08	8.154e-08	7.674e-08	−7.718e-08	3.852e-08	2.127e-08
H	−7.891e-09	1.285e-07	−1.562e-07	−3.414e-08	6.667e-08	−1.226e-07

.

Table S2: Differences (in hartree/bohr) between analytic and finite-difference gradient components computed at two different levels of theory, for a hydrogen-bonded isomer of $(H_2O)_2^+$. Finite difference calculations were performed using a five-point stencil. The maximum absolute deviation is 9.77×10^{-8} a.u..

Atom	BLYP/6-31G*			B3LYP/6-31G*		
	x	y	z	x	y	z
O	9.771e-08	1.738e-09	−5.018e-08	4.449e-08	1.660e-08	−4.340e-08
H	−6.878e-08	−4.817e-08	−6.990e-08	−6.992e-08	−2.034e-08	2.299e-09
H	8.668e-08	4.343e-08	−1.898e-08	−2.426e-08	1.518e-08	−4.875e-08
O	5.735e-08	−3.157e-09	−2.113e-08	−3.478e-08	1.157e-08	1.992e-10
H	1.117e-08	2.421e-08	−7.824e-08	−2.371e-08	2.788e-08	−7.292e-09
H	−2.230e-08	−4.363e-08	−2.608e-08	−2.945e-08	−7.423e-08	3.412e-08

.

Table S3: Differences (in hartree/bohr) between analytic and finite-difference gradient components computed at two different levels of theory, for $(H_2O)_2^+$. Finite difference calculations were performed using a five-point stencil. The maximum absolute deviation is 9.57×10^{-7} a.u..

Atom	BLYP/6-31G*			B3LYP/6-31G*		
	x	y	z	x	y	z
O	−3.244e-07	3.129e-07	6.751e-07	2.641e-07	2.771e-07	5.934e-07
H	3.605e-07	−4.290e-07	−3.741e-07	3.082e-07	−3.977e-07	−2.984e-07
H	−2.100e-07	−4.170e-07	−3.512e-07	−1.868e-07	−3.568e-07	−2.913e-07
O	8.716e-08	−1.689e-07	2.646e-08	8.522e-08	−1.525e-07	3.499e-08
H	−1.852e-07	3.660e-07	2.948e-07	−1.778e-07	3.430e-07	2.618e-07
H	7.547e-08	1.641e-07	−2.154e-07	1.071e-07	1.187e-07	−1.882e-07
O	4.128e-08	−3.663e-07	9.838e-08	1.123e-07	−2.161e-07	9.244e-08
H	1.423e-07	4.878e-07	−2.589e-07	1.230e-07	3.094e-07	−2.442e-07
H	2.325e-07	−1.620e-08	2.529e-07	2.182e-07	−5.651e-08	1.033e-07
O	−5.208e-08	2.038e-07	−1.550e-07	−5.315e-08	2.280e-07	−1.220e-07
H	−2.086e-07	−1.651e-07	−1.620e-07	−2.379e-07	−1.355e-07	−1.344e-07
H	4.829e-07	−8.912e-08	−3.987e-07	2.886e-07	−6.868e-09	−3.934e-07
O	−1.050e-09	−3.431e-07	7.041e-07	−7.767e-08	−3.371e-07	5.428e-07
H	3.270e-07	5.810e-07	−2.420e-07	2.248e-07	3.850e-07	−2.881e-07
H	−9.568e-07	−3.410e-07	−5.931e-08	−8.993e-07	−3.046e-07	−3.160e-08
O	−9.057e-08	−3.265e-07	−2.990e-07	−4.757e-08	−2.821e-07	−3.024e-07
H	−3.115e-07	−1.014e-07	5.742e-09	−1.740e-07	−8.389e-08	5.420e-09
H	6.784e-08	3.329e-07	1.379e-07	−1.200e-08	3.719e-07	1.291e-07

.

Table S4: Differences (in hartree/bohr) between analytic and finite-difference gradient components computed at two different levels of theory, for the $Al(OSiH_3)_4$ cluster examined in this work. The maximum absolute deviation is 1.31×10^{-6} a.u..

Atom	BLYP/6-31G*			B3LYP/6-31G*		
	x	y	z	x	y	z
Si	4.936e-08	−7.837e-09	−8.067e-08	1.571e-08	3.548e-08	−1.155e-07
H	−2.627e-07	−3.505e-07	−1.034e-07	−2.227e-07	−2.818e-07	−1.045e-08
H	−2.614e-08	−1.935e-07	4.838e-07	4.020e-08	−1.090e-07	4.970e-07
H	2.068e-07	−4.529e-07	−2.699e-07	2.239e-07	−4.223e-07	−2.900e-07
H	2.159e-08	−1.969e-07	−1.055e-07	2.497e-08	−1.840e-07	−1.201e-07
Si	2.360e-08	7.413e-07	1.311e-06	6.014e-08	6.843e-07	1.238e-06
O	1.200e-06	−2.786e-07	−1.182e-06	1.156e-06	−2.017e-07	−1.173e-06
Al	−9.434e-07	−5.565e-07	1.532e-07	−9.688e-07	−5.065e-07	1.517e-07
H	1.779e-07	−3.961e-07	−1.527e-07	2.035e-07	−3.956e-07	−2.098e-07
H	−1.853e-07	1.656e-07	−2.369e-09	−1.631e-07	1.122e-07	−5.434e-08
O	1.984e-07	6.578e-07	−4.272e-07	1.841e-07	7.034e-07	−3.721e-07
Si	−1.067e-06	−6.141e-07	−4.590e-07	−1.190e-06	−6.639e-07	−4.794e-07
H	−3.594e-07	3.533e-07	8.758e-08	−3.238e-07	3.792e-07	1.012e-07
O	3.680e-07	1.262e-07	4.692e-07	4.822e-07	1.341e-07	4.403e-07
O	−1.955e-07	2.562e-07	−3.844e-07	−1.246e-07	1.517e-07	−3.583e-07
H	−7.658e-08	2.761e-07	1.242e-07	−1.335e-07	3.157e-07	8.427e-08
Si	8.973e-07	−8.443e-07	−5.329e-07	9.322e-07	−7.670e-07	−5.215e-07
H	2.476e-07	4.322e-08	2.164e-07	1.921e-07	−2.925e-08	2.056e-07
H	−4.249e-07	1.191e-07	−6.588e-08	−4.745e-07	1.310e-07	−1.208e-07
H	−3.334e-09	−9.330e-08	1.357e-07	4.714e-08	−9.196e-08	1.472e-07
H	4.422e-07	−6.737e-09	−1.853e-07	4.288e-07	−2.478e-08	−1.092e-07

.

Table S5: Barrier heights for the data set of reactions in Ref. 1. Reactants and product structures were optimized with the 6-31+G** basis set for all of the methods but transition state (TS) structures are taken from Ref. 2, where they were optimized at the QCISD/6-31+G** level. Error is significantly reduced by DC-BLYP as compared to the self-consistent BLYP or HF methods.

Reaction	Direction	CCSD(T)	BLYP	DC-BLYP	HF
$Cl + H_2 \rightarrow HCl + H$	forward	9.8	4.8	8.3	22.6
	backward	5.7	−3.3	0.2	13.6
$OH + H_2 \rightarrow H + H_2O$	forward	6.4	−2.9	5.2	22.9
	backward	21.7	8.5	16.4	26.0
$CH_3 + H_2 \rightarrow H + CH_4$	forward	12.5	7.4	10.8	22.2
	backward	15.7	7.4	10.7	23.7
$H + H_2 \rightarrow H_2 + H$	forward	10.5	2.7	5.2	17.9
	backward	10.5	2.7	5.2	17.9
$OH + NH_3 \rightarrow H_2O + NH_2$	forward	4.9	−9.9	2.6	28.9
	backward	14.6	−0.2	11.7	32.2
$F + H_2 \rightarrow H + HF$	forward	2.6	−11.7	1.9	14..0
	backward	33.1	17.6	30.1	26.3
$H + PH_3 \rightarrow PH_2 + H_2$	forward	3.5	−3.1	0.6	10.3
	backward	26.1	22.8	26.8	13.9
$H + H_2S \rightarrow H_2 + HS$	forward	4.5	−2.9	0.6	11.7
	backward	19.2	15.9	19.6	28.3
$OH + H \rightarrow H_2 + O$	forward	10.8	−0.0	5.6	17.3
	backward	14.9	1.9	8.9	33.0
$O + HCl \rightarrow OH + Cl$	forward	11.6	−9.1	9.5	33.4
	backward	11.6	−3.1	14.3	25.8
$H + t\text{-}N_2H_2 \rightarrow H_2 + N_2H$	forward	3.9	−7.1	9.5	12.2
	backward	41.1	35.2	44.8	49.3
$CH_4 + NH \rightarrow NH_2 + CH_3$	forward	23.5	14.1	19.9	40.1
	backward	8.9	2.4	7.9	20.0
$CH_4 + NH_2 \rightarrow NH_3 + CH_3$	forward	14.9	8.0	13.4	31.9
	backward	17.2	9.8	15.5	30.1
Average Error			−9.8	−2.0	10.3
Average Absolute Error			9.8	3.0	11.9

.

Table S6: Density sensitivities S [Eq. (3)] for Al-doped silica clusters, computed using the BLYP and PBE functionals. The localized and delocalized structures correspond to geometries optimized using MP2 and BLYP, respectively.

Complex	Hole	S (kcal/mol)		S/N_{atoms} (kcal/mol)	
		BLYP	PBE	BLYP	PBE
$Al(OH)_4$	Localized	37.0	33.7	4.1	3.7
	Delocalized	67.5	64.0	7.5	7.1
$Al(OSiH_3)_4$	Localized	63.5	66.2	3.0	3.2
	Delocalized	53.7	56.4	2.6	2.7
$Al[OSi(OH)_3]_4$	Localized	129.4	121.9	3.9	3.7
	Delocalized	156.7	148.0	4.7	4.5

Table S7: Mulliken spin charges[a] (in a.u.) around the Al dopant in silica clusters. This is a more complete version of Table 2.

Complex	Geometry	HF density				BLYP density			
		O_1	O_2	O_3	O_4	O_1	O_2	O_3	O_4
$Al(OH)_4$	HF	1.07	0.00	0.00	0.00	0.65	0.13	0.11	0.13
	BLYP	1.06	0.01	0.00	0.01	0.35	0.24	0.23	0.26
	PBE	1.06	0.01	0.00	0.01	0.35	0.23	0.23	0.26
	B3LYP	0.63	0.00	0.00	0.46	0.50	0.01	0.02	0.52
	PBE0	0.64	−0.00	−0.00	0.45	0.51	0.01	0.01	0.52
	DC-BLYP	1.07	0.00	0.00	0.00	0.70	0.12	0.12	0.12
	DC-PBE	1.07	0.00	0.00	0.00	0.70	0.12	0.12	0.12
	DC-B3LYP	1.07	0.00	0.00	0.00	0.69	0.12	0.12	0.13
	DC-PBE0	0.01	0.01	−0.01	1.02	0.70	0.12	0.12	0.12
$Al(OSiH_3)_4$	HF	0.01	0.01	−0.01	1.02	0.11	0.11	0.05	0.57
	BLYP	0.25	0.24	0.25	0.25	0.19	0.19	0.19	0.19
	PBE	0.24	0.24	0.25	0.25	0.19	0.19	0.19	0.19
	B3LYP	0.47	−0.00	−0.00	0.50	0.40	0.03	0.03	0.40
	PBE0	0.49	−0.00	−0.00	0.49	0.43	0.01	0.01	0.43
	DC-BLYP	0.01	0.00	−0.01	1.02	0.13	0.09	0.05	0.57
	DC-PBE	0.01	0.01	−0.01	1.02	0.12	0.11	0.04	0.57
	DC-B3LYP	0.00	0.01	−0.01	1.02	0.10	0.13	0.04	0.57
	DC-PBE0	0.01	0.01	−0.01	1.02	0.12	0.11	0.04	0.57
$Al[OSi(OH)_3]_4$[b]	HF	0.99	−0.01	0.01	0.00	0.69	0.00	0.05	0.05
	DC-BLYP	1.02	−0.01	−0.01	0.01	0.64	0.01	0.03	0.11
	DC-PBE	0.99	0.00	0.00	0.00	0.67	−0.00	0.02	0.06
	DC-B3LYP	1.01	−0.01	−0.00	0.01	0.64	0.01	0.03	0.11
	DC-PBE0	0.98	−0.01	0.00	0.00	0.67	0.00	0.02	0.05

[a]6-31++G* basis set. [b]Data for PBE, BLYP, PBE0, and B3LYP are not shown due to significant delocalization (see Fig. S3).

Table S8: Mulliken spin charges (in a.u.) around the Al dopant in the $Al(OSiH_3)_4$ complex computed with def2-TZVPD basis set.

Complex	Geometry	HF density				BLYP density			
		O_1	O_2	O_3	O_4	O_1	O_2	O_3	O_4
$Al(OSiH_3)_4$	HF	−0.00	0.00	0.01	1.06	0.11	0.11	0.05	0.58
	BLYP	0.28	0.29	0.29	0.29	0.19	0.19	0.19	0.19
	PBE	0.27	0.28	0.28	0.31	0.18	0.19	0.19	0.19
	B3LYP	0.55	−0.01	−0.01	0.58	0.41	0.02	0.02	0.41
	PBE0	0.57	−0.01	−0.01	0.57	0.44	0.01	0.01	0.44
	DC-BLYP	−0.00	0.00	0.01	1.06	0.13	0.10	0.05	0.58
	DC-PBE	−0.00	0.00	0.01	1.06	0.12	0.11	0.04	0.59
	DC-B3LYP	−0.00	0.00	0.01	1.06	0.10	0.13	0.05	0.58
	DC-PBE0	−0.00	0.00	0.01	1.05	0.12	0.11	0.04	0.58

Table S9: Density sensitivities S [Eq. (3)] for HF-optimized geometries of $[Ti_{21}O_{70}H_{56}]$ and $[Ti_{21}O_{70}H_{56}]^-$ using the BLYP and PBE functionals.

Complex	S (kcal/mol)		S/N_{atoms} (kcal/mol)	
	BLYP	PBE	BLYP	PBE
$[Ti_{21}O_{70}H_{56}]$	1229.12	1199.83	8.36	8.16
$[Ti_{21}O_{70}H_{56}]^-$	1236.69	1205.83	8.41	8.20

Table S10: Average value of $\langle \hat{S}^2 \rangle / \hbar^2$ (with standard deviations in parenthesis) for structures along the PBE or PBE0 optimization pathway of the TiO_2 polaron, computed at different levels of theory. These results demonstrate that spin contamination is not significant in this system.

Single-Point	PBE pathway	PBE0 pathway
HF	0.752 (0.0010)	0.753 (0.0000)
PBE	0.753 (0.0005)	0.754 (0.0001)
PBE0	0.763 (0.0035)	0.753 (0.0001)
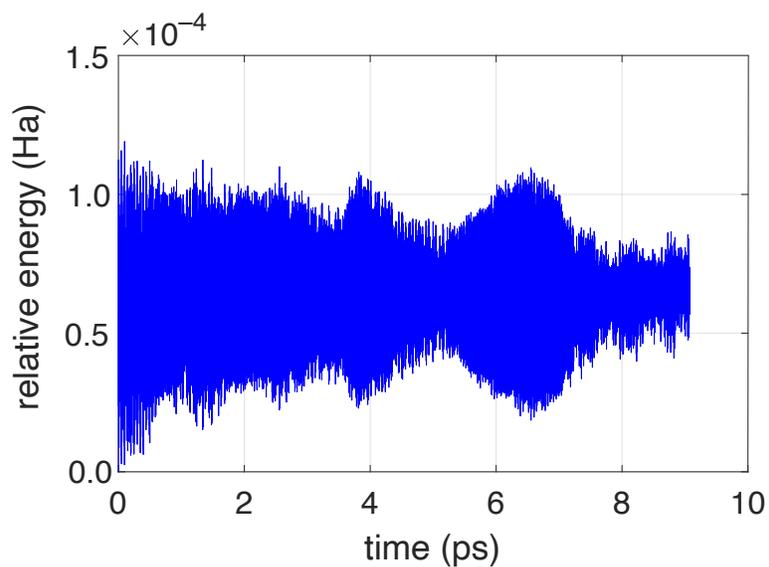
Figure S1: Energy fluctuations in a *ab initio* MD simulation of $(H_2O)_5^+$ at the DC-B3LYP/6-31G* level of theory in the microcanonical ensemble.
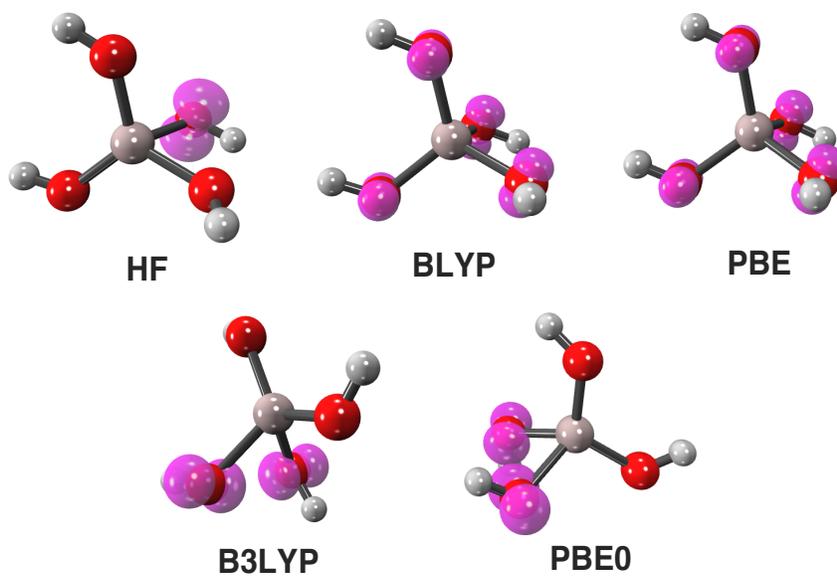
Figure S2: Spin densities for the $Al(OH)_4$ cluster at different levels of theory. (Each structure corresponds to the optimized geometry at the indicated level of theory and the 6-31++G* basis set is used for all calculations.) The MP2 and DC-DFT results are not shown because they are each quite similar to the HF result.
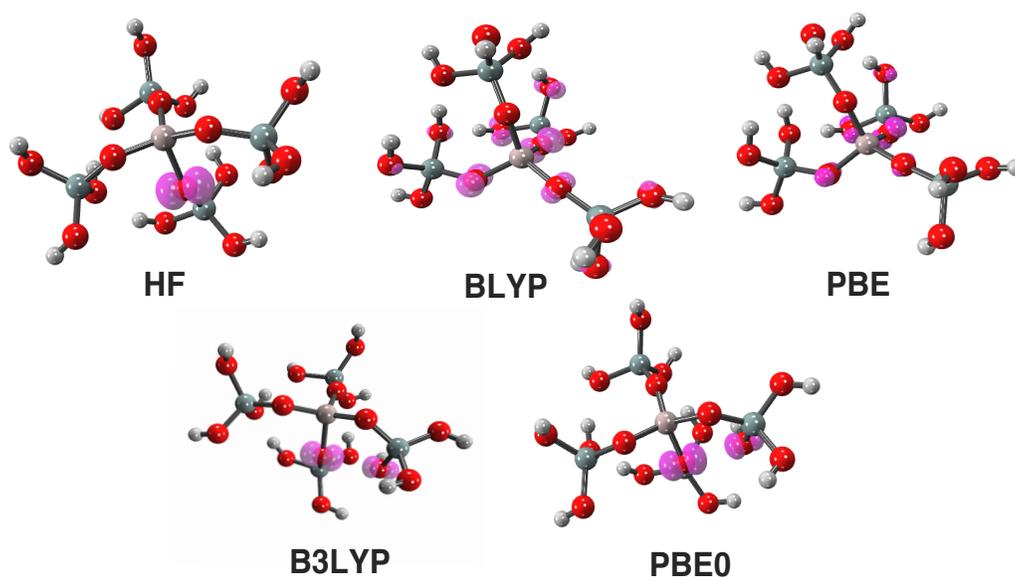
Figure S3: Spin densities for the Al[OSi(OH)$_3$]$_4$ cluster at different levels of theory. (Each structure corresponds to the optimized geometry at the indicated level of theory and the 6-31++G* basis set is used for all calculations.) The MP2 and DC-DFT results are not shown because they are each quite similar to the HF result.
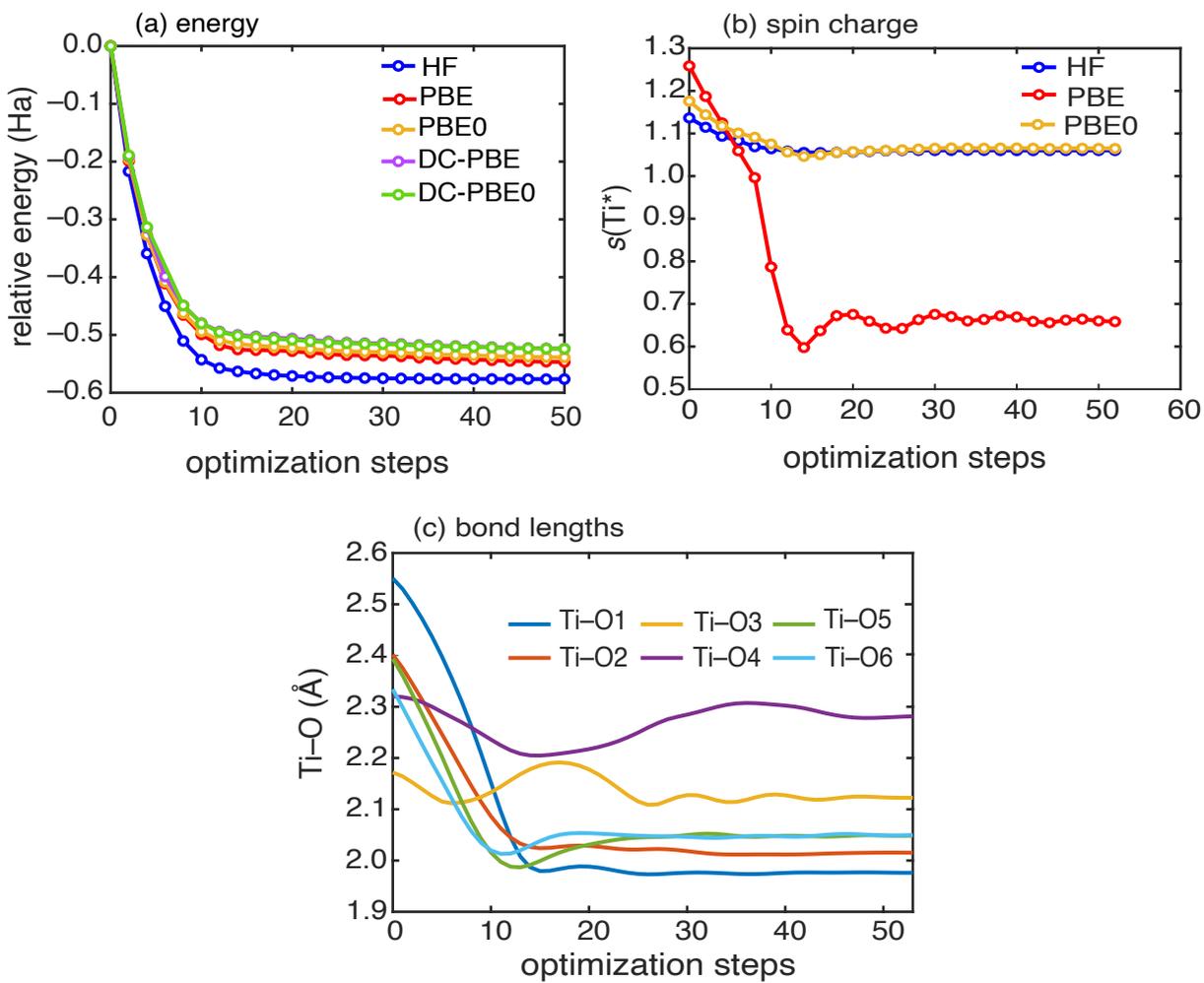
Figure S4: Behavior of the localized electron-polaron in TiO_2 along the HF/LANL2DZ optimization pathway, including variations in (a) the relative energy computed at different level of theory, (b) the Mulliken spin charge $s(Ti^*)$ on the central Ti atom, and (c) the Ti^*–O bond lengths around the central Ti atom.
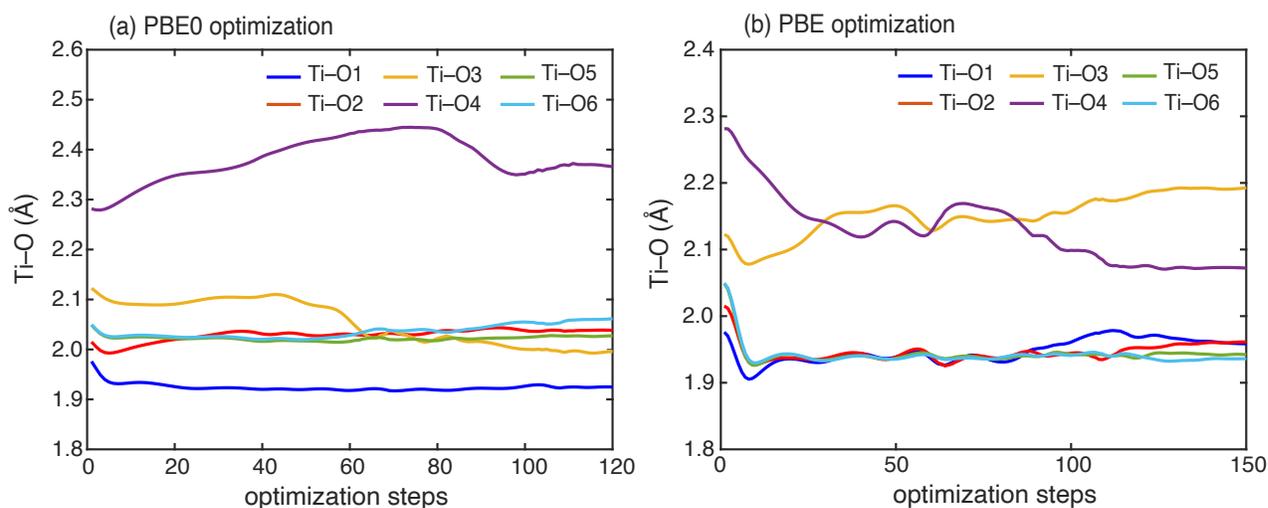
Figure S5: Variation in Ti*–O distances around the central Ti atom, for optimizations of an initially-localized electron-polaron in TiO$_2$. Optimizations were performed at (a) the PBE0/LANL2DZ level and (b) the PBE/ LANL2DZ level, starting from a HF-optimized structure.
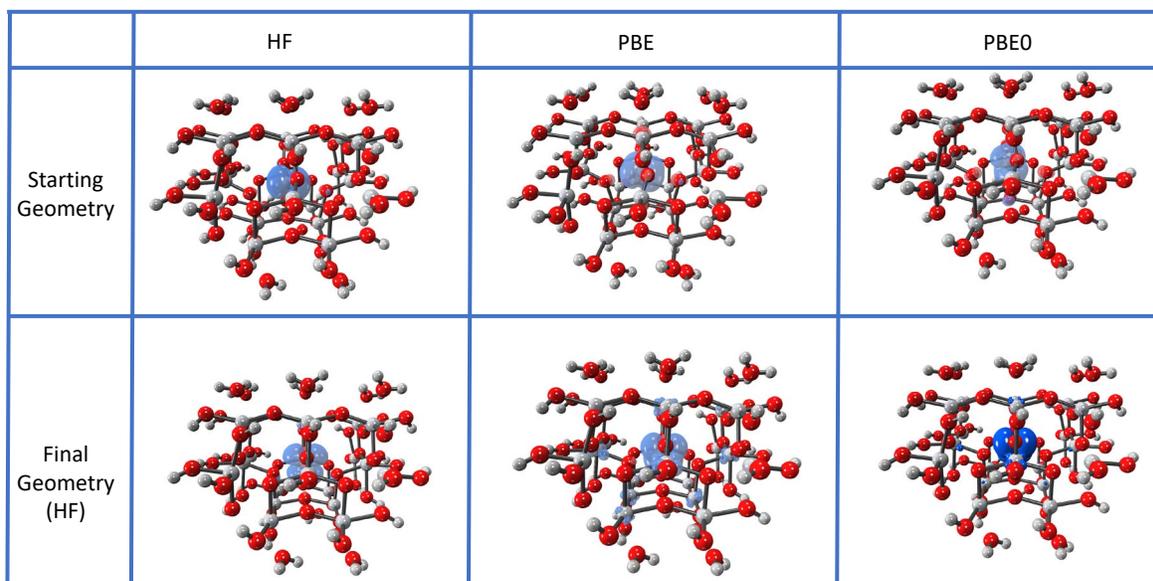
Figure S6: Spin densities for the initial and final geometries along the HF optimization pathway for $[\text{Ti}_{21}\text{O}_{70}\text{H}_{56}]^-$, computed at different levels of theory. All calculations employed the LANL2DZ basis set for Ti and the 6-31G* basis set for O and H.
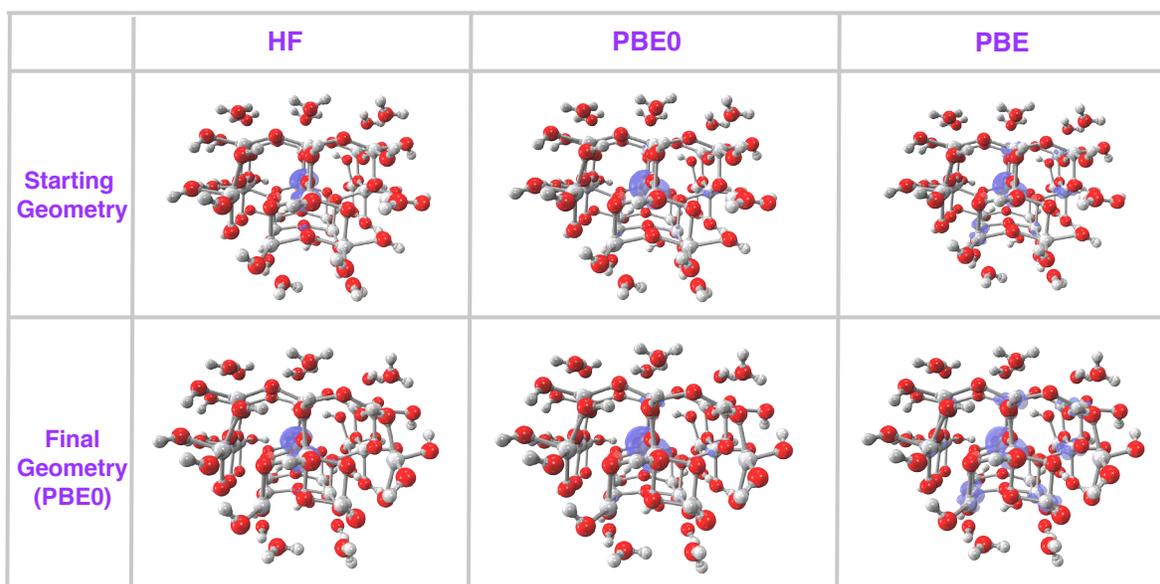
Figure S7: Spin densities for the initial and final geometries along the PBE0 optimization pathway for $[\text{Ti}_{21}\text{O}_{70}\text{H}_{56}]^{-}$, computed at different levels of theory. All calculations employed the LANL2DZ basis set for Ti and the 6-31G* basis set for O and H. The starting structure is the HF-optimized geometry.
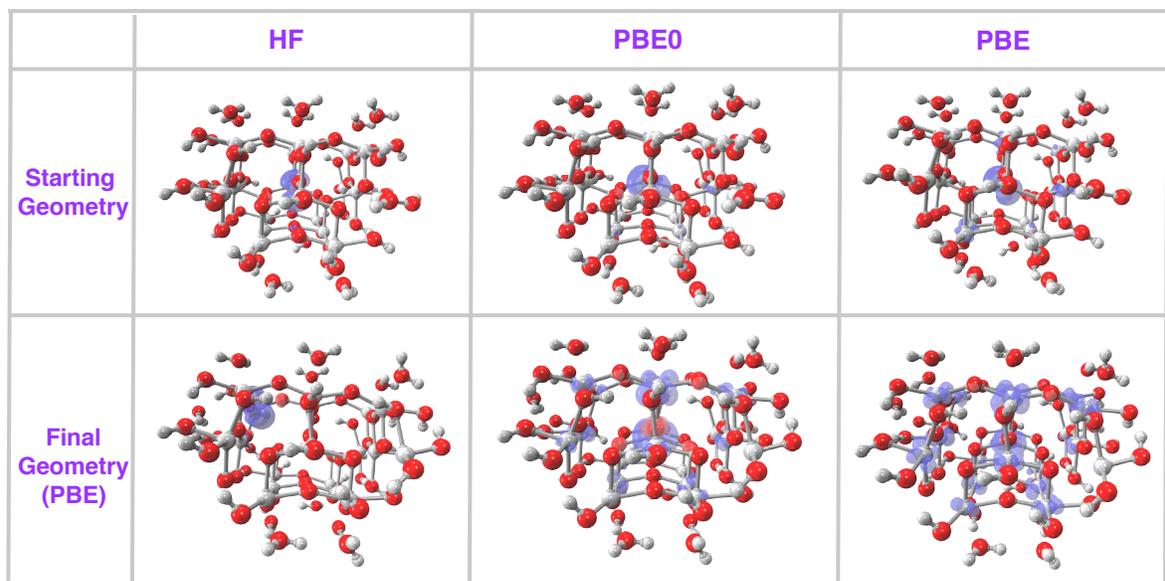
Figure S8: Spin densities for the initial and final geometries along the PBE optimization pathway for $[\text{Ti}_{21}\text{O}_{70}\text{H}_{56}]^{-}$, computed at different levels of theory. All calculations employed the LANL2DZ basis set for Ti and the 6-31G* basis set for O and H. The starting structure is the HF-optimized geometry.
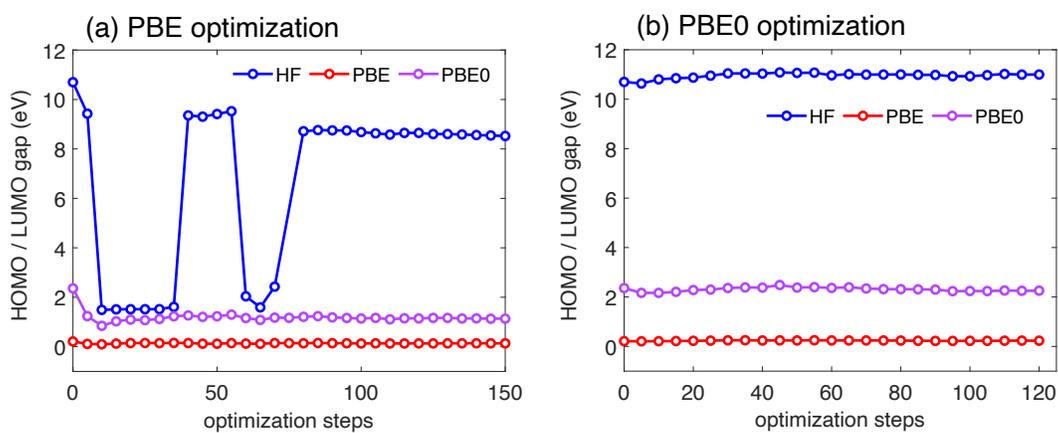
Figure S9: HOMO/LUMO gap for $[Ti_{21}O_{70}H_{56}]^-$ computed at various levels of theory, along (a) the PBE optimization pathway, versus (b) the PBE0 optimization pathway. All calculations employed the LANL2DZ basis set for Ti and the 6-31G* basis set for O and H. The starting structure is the HF-optimized geometry. In (a), the sudden jumps in the gap that are obtained at the HF level should be compared to the jumps in the energy that are observed in Fig. 6 when HF, MP2, or hybrid functionals are used.
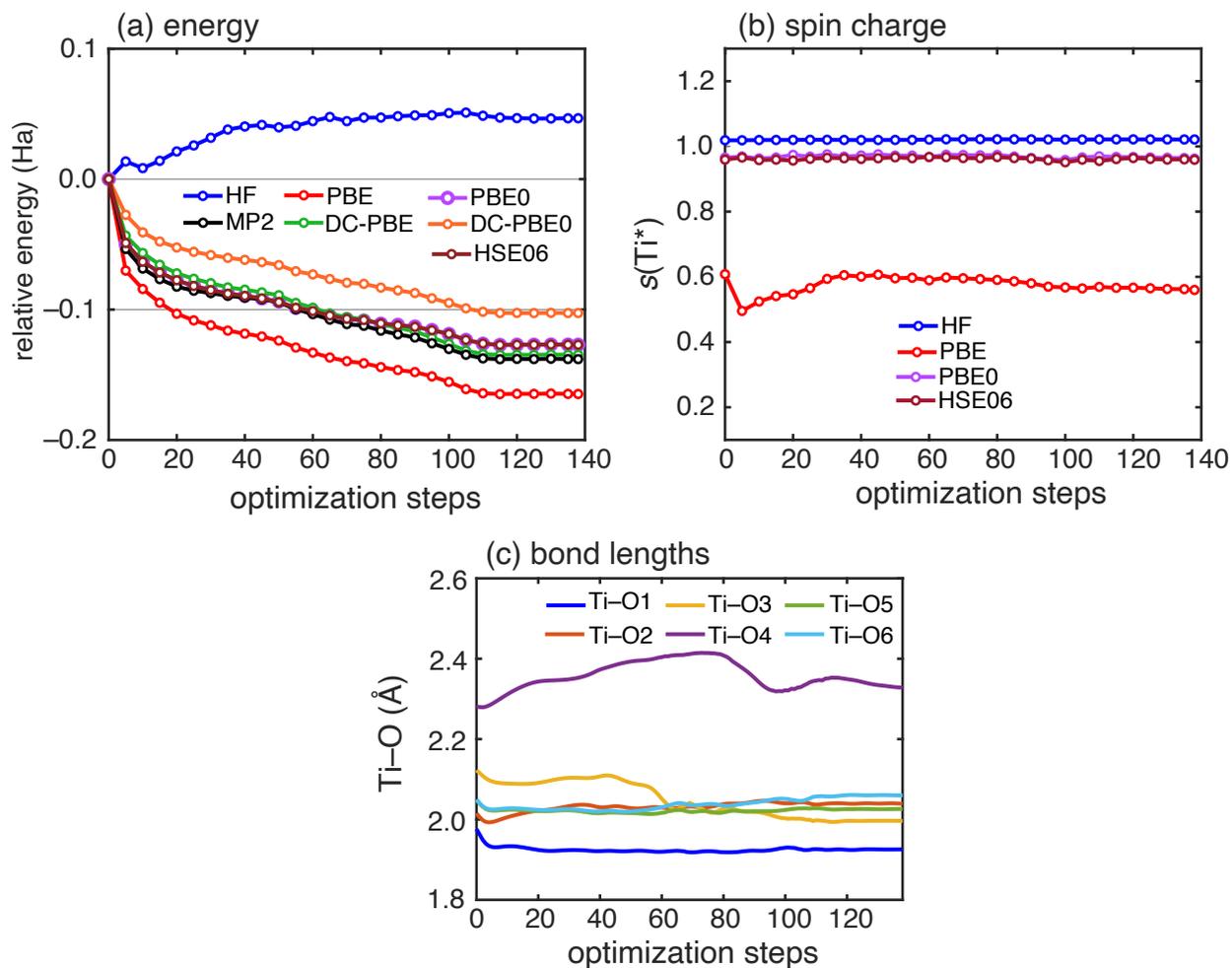
Figure S10: Behavior of the localized electron-polaron in TiO_2 along the HSE06/LANL2DZ optimization pathway, including variations in (a) the relative energy computed at different level of theory, (b) the Mulliken spin charge $s(Ti^*)$ on the central Ti atom, and (c) the Ti^*–O bond lengths around the central Ti atom.
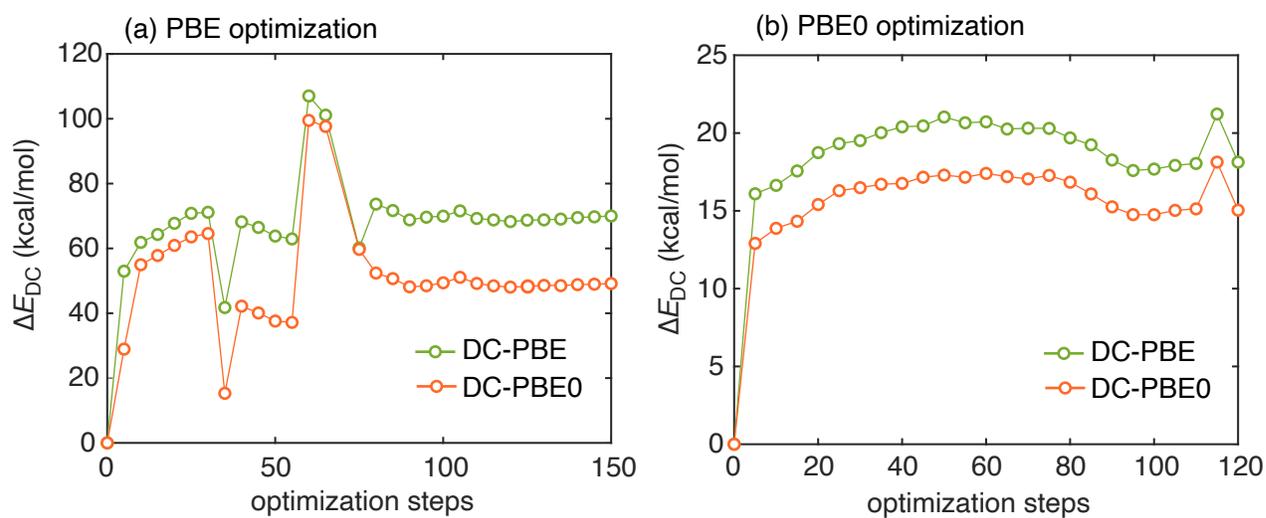
Figure S11: Differences ΔE_{DC} between energies computed using DFT and DC-DFT (for DFT = PBE or DFT = PBE0), along the (a) the PBE optimization pathway versus (b) the PBE0 optimization pathway. All calculations used the def2-SVP basis set.
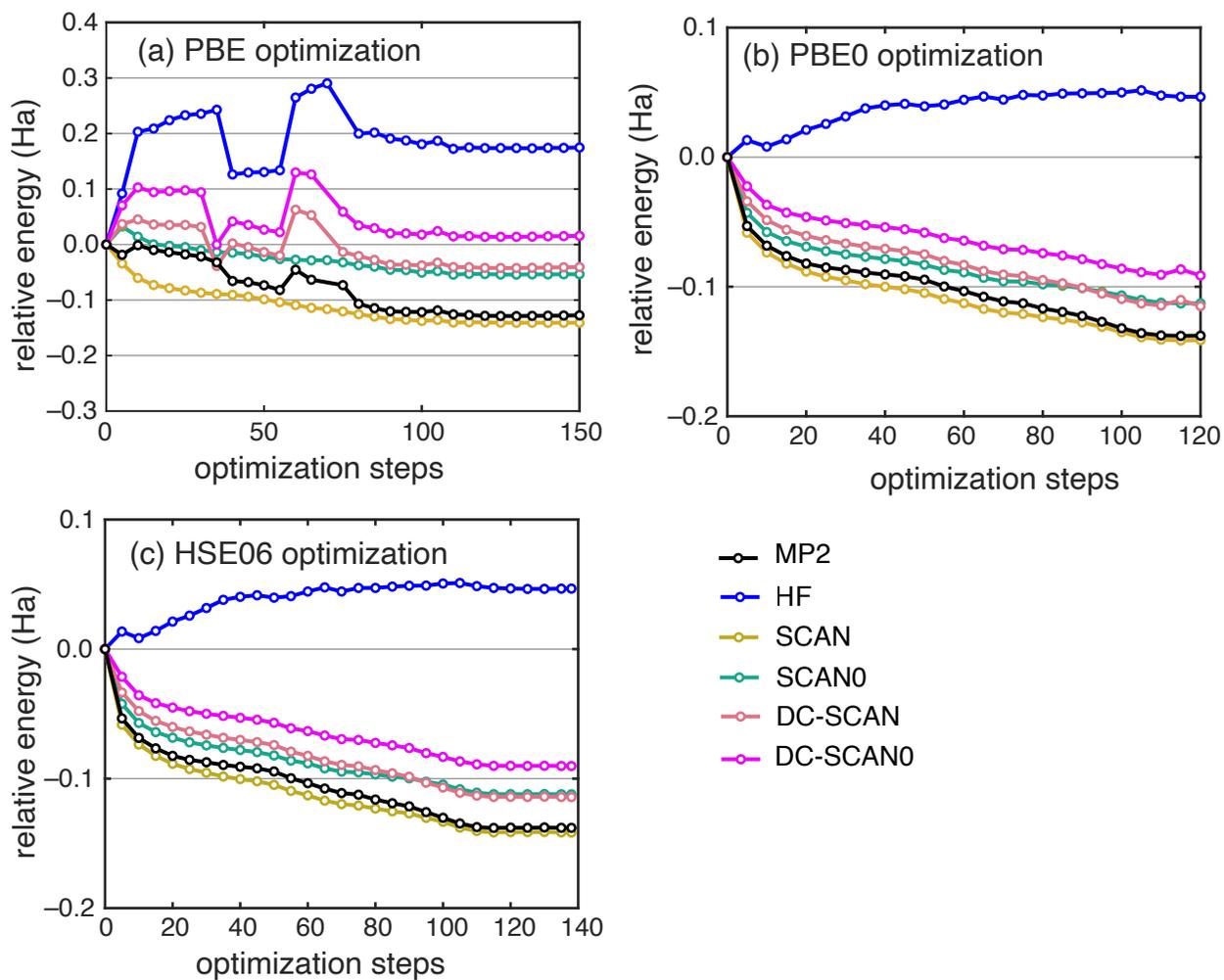
Figure S12: Relative energies computed with DFT and DC-DFT along optimization pathways computed using the (a) PBE, (b) PBE0, and (c) HSE06 functionals. All calculations use the def2-SVP basis set.

References

[1] Verma, P.; Perera, A.; Bartlett, R. J. Increasing the applicability of DFT I: Non-variational correlation corrections from Hartree–Fock DFT for predicting transition states. *Chem. Phys. Lett.* **2012,** *524,* 10–15.

[2] Lynch, B. J.; Truhlar, D. G. How well can hybrid density functional methods predict transition state geometries and barrier heights? *J. Phys. Chem. A* **2001,** *105,* 2936–2941.

[3] Holden, Z. C.; Rana, B.; Herbert, J. M. Analytic energy gradients for the QM/MM-Ewald method using atomic charges derived from the electrostatic potential: Theory, implementation, and application to *ab initio* molecular dynamics of the aqueous electron. *J. Chem. Phys.* **2019,** *150,* 144115:1–20.

[4] Liu, J.; Rana, B.; Liu, K.-Y.; Herbert, J. M. Variational formulation of the generalized many-body expansion with self-consistent embedding charges: Simple and correct analytic energy gradient for fragment-based *ab initio* molecular dynamics. *J. Phys. Chem. Lett.* **2019,** *10,* 3877–3886.

[5] Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998,** *294,* 143–152.