Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2023

Electronic Supplementary Information: ONIOM meets xtb: Efficient, Accurate, and Robust Multi-Layer Simulations Across the Periodic Table

Christoph Plett,[a] Abylay Katbashev,[a] Sebastian Ehlert,[b]
Stefan Grimme,*[a] and Markus Bursch*[c]

May 12, 2023

[a] *Mulliken Center for Theoretical Chemistry, Universität Bonn, Beringstr. 4, 53115 Bonn, Germany. E-mail:* grimme@thch.uni-bonn.de

[b] *Microsoft Research AI4Science, Evert van de Beekstraat 254, 1118 CZ Schiphol, Netherlands.*

[c] *Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany. E-mail:* bursch@kofo.mpg.de

Contents

1 Implementation

1.1 Element-specific parameters

Table 1: Element-specific parameters used to calculate the static scaling factor k in the ONIOM implementation of *xtb*. The average bond distances d^{ONIOM} are given in Å.

Elements	H	C	O	N	P	S
H	0.740	1.084	0.964	1.024	1.414	1.389
C	-	1.528	1.430	1.475	1.860	1.750
O	-	-	1.450	1.360	1.750	1.500
N	-	-	-	1.470	1.770	1.650

1.2 Derivatives for Jacobian

The general equation for the Link atom coordinates is given by:

$$\vec{R}_{\mathrm{LA}} = \vec{R}_{\mathrm{inner}} + k(\vec{R}_{\mathrm{outer}} - \vec{R}_{\mathrm{inner}}) \tag{1}$$

Herein, k factor is defined as:

$$k = \frac{d_1}{d_2} \tag{2}$$

where d_1 is the distance between the inner region atom and a hydrogen atom that is taken as corresponding parameter from Table 1. d_2 is the distance between the inner region and outer region atoms which d_2 can be chosen to be either constant or equal to the actual distances in molecular systems.

1.2.1 Constant

If d_2 is constant, the differentiation of Equation (1) with respect to the real coordinates of the full system is given by:

$$\frac{\delta R_{\mathrm{LA,\,i}}}{\delta R_{\mathrm{inner,i}}} = (1-k), \text{ with } \quad \mathrm{i{=}x,y,z} \tag{3}$$

$$\frac{\delta R_{\mathrm{LA,\,i}}}{\delta R_{\mathrm{outer,\,i}}} = k \tag{4}$$

1.2.2 Dynamic

If d_2 is set to be equal to the actual distance in the system, is can be defined as

$$d_2 = \sqrt{\sum_{i=1}^{3} (R_{\mathrm{outer,i}} - R_{\mathrm{inner,i}})^2}. \tag{5}$$

Thus, the derivatives of Equation (1) changes to

$$\frac{\delta R_{\mathrm{LA,\,i}}}{\delta R_{\mathrm{inner,\,i}}} = 1 - k\left(1 - \frac{(R_{\mathrm{outer,i}} - R_{\mathrm{inner,i}})^2}{d_2}\right), \tag{6}$$

and

$$\frac{\delta R_{\mathrm{LA,\,i}}}{\delta R_{\mathrm{outer,\,i}}} = k\frac{\sum_{i\neq j}^{3} (R_{\mathrm{outer,j}} - R_{\mathrm{inner,j}})^2}{d_2}. \tag{7}$$

2 Detailed results

2.1 Molecular structures

2.1.1 UiO-66

Table 2: Comparison of the heavy atom root-mean-square deviations (hRMSDs) (vs X-ray) and the wall-times (on 14 cores) for the cut-out of the UiO-66 metal-organic framework (MOF) with a total size of 484 atoms optimized with different GFN and DFT methods, as well as their ONIOM combinations. 1, 3, and 6 are sizes of the inner regions used for the corresponding ONIOM calculations.

Methods	size	hRMSD(vs X-ray) / Å	Wall-time (14cores)
GFN-FF	full	0.335	5s
GFN2-xTB//GFN-FF	1	0.290	12s
	3	0.233	49s
	6	0.115	1m 54s
TPSS-D4/def2-SVP//GFN-FF	1	0.288	52m
	3	0.213	5h 55m
	6	0.025	11h 1m
r²SCAN-3c//GFN-FF	1	0.292	2h 21m
	3	0.219	12h 50m
	6	0.025	20h 57m
ωB97X-3c//GFN-FF	1	0.291	16h 16m
	3	0.212	113h 19m
	6	0.031	202h 51m
GFN2-xTB	full	0.156	3m
TPSS-D4/def2-SVP//GFN2-xTB	1	0.155	1h 30m
	3	0.119	6h 1m
	6	0.059	10h 45m
r²SCAN-3c//GFN2-xTB	1	0.161	3h 31m
	3	0.124	19h 22m
	6	0.054	18h 20m
ωB97X-3c//GFN2-xTB	1	0.158	31h 35m
	3	0.126	169h 59m
	6	0.069	334h 32m
TPSS-D4/def2-SVP	full	0.031	46h 9m
r²SCAN-3c	full	0.030	190h 38m
ωB97X-3c	full	-	-

2.1.2 DALTES

Table 3: hRMSD (vs.X-Ray) and wall time values (on 28 cores) of the geometry optimizations of the DALTES polyhedron with different QC methods, ordered according to the increasing computing time.

Methods	hRMSD	Wall time
GFN-FF	0.315	3s
GFN2-xTB	0.178	2m
TPSS-D4/def2-SVP	**0.143**	**15h 42m**
TPSS-D4/def2-mSVP	0.140	37h 18m
PBEh-3c	0.167	51h 10m
B97-3c	0.162	73h 55m
r²SCAN-3c	0.152	131h 40m
TPSS-D4/def2-TZVP	0.153	147h 45m

2.2 Electronic energies

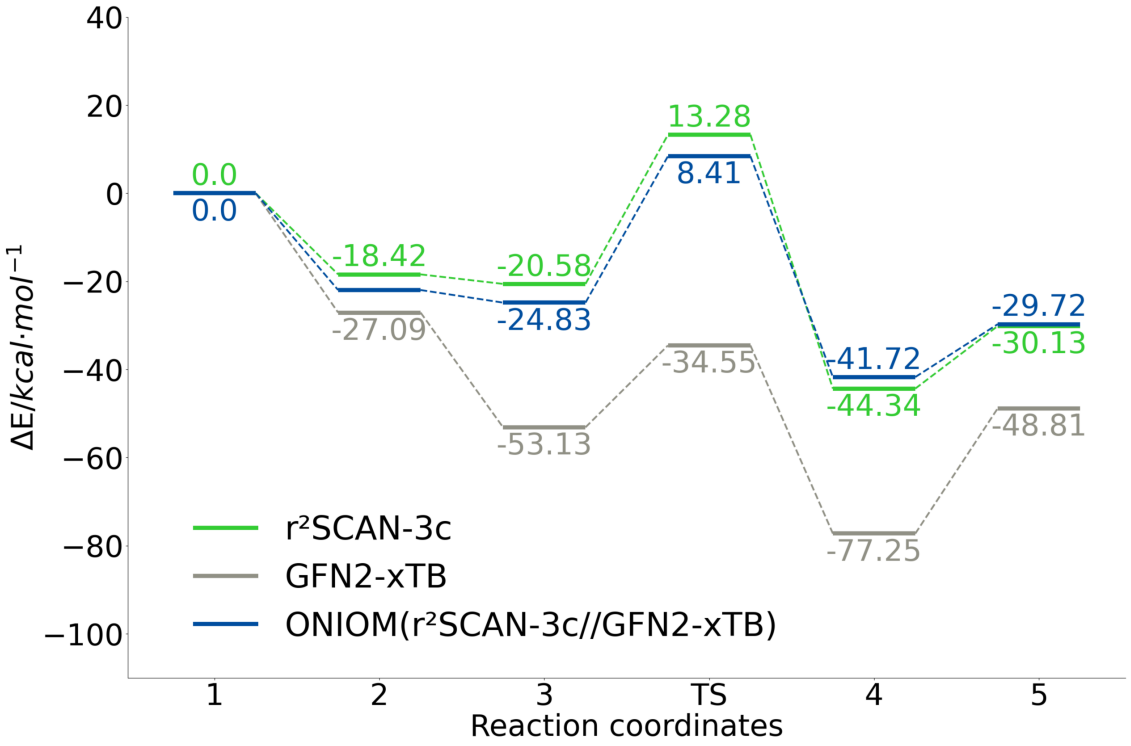

Figure 1: The relative potential energy curve of the cyanosilylation reaction computed with r²SCAN-3c, GFN2-xTB, and their r²SCAN-3c//GFN2-xTB ONIOM combination with the TPSS-D4/def2-SVP optimized geometries implicitly solvated in toluene.

Table 4: The cyanosilylation relative reaction energies at the DALTES dirhodium metal node calculated with r²SCAN-3c, TPSS-D4/def2-SVP and their ONIOM combinations with GFN2-xTB with the TPSS-D4/def2-SVP optimized geometries. The individual values are given in kcal·mol^{-1}.

Methods	Solvation model	1	2	3	TS	4	5
TPSS-D4/def2-SVP	COSMO	0.00	-25.75	-30.78	1.52	-51.88	-28.99
TPSS-D4/def2-SVP//GFN2-xTB	GBSA	0.00	-28.52	-33.98	-3.49	-51.39	-25.98
	ddCOSMO	0.00	-28.44	-32.84	-2.04	-47.80	-28.15
	COSMO/ddCOSMO	0.00	-25.69	-36.56	-7.05	-43.54	-36.85
r²SCAN-3c	COSMO	0.00	-18.42	-20.58	13.28	-44.34	-30.13
r²SCAN-3c//GFN2-xTB	GBSA	0.00	-20.28	-24.20	8.73	-43.54	-25.78
	ddCOSMO	0.00	-21.97	-24.89	8.41	-41.72	-29.78
	COSMO/ddCOSMO	0.00	-27.36	-27.58	3.29	-60.00	-37.12