Supporting Information for:
Comment on "Benchmarking Basis Sets for Density Functional Theory Thermochemistry Calculations: Why Unpolarized Basis Sets and the Polarized 6-311G Family Should Be Avoided"

Montgomery Gray,[1] Paige E. Bowling,[1,2] and John M. Herbert[1,2*]

[1]*Department of Chemistry & Biochemistry*
The Ohio State University, Columbus, Ohio 43210 USA
[2]*Biophysics Graduate Program*
The Ohio State University, Columbus, Ohio 43210 USA

June 13, 2024

S1 Computational details

All calculations use Cartesian Gaussian basis functions. The SCF process is initiated started from a superposition of atomic densities (SAD) guess, where the atomic densities were computed at the Hartree-Fock level in the target basis set. Direct inversion in the iterative subspace (DIIS)[1] was used as a convergence accelerator. The convergence criterion was set to $\tau_{\text{SCF}} = 10^{-8}\ E_\text{h}$ for all of the GMTKN55 calculations and all timing benchmarks. For the GMTKN55 calculations, we used $\tau = 10^{-14}$ a.u., where $\tau \equiv \tau_{\text{ints}} = \tau_{\text{shlpr}}$. Timing benchmarks are reported for both $\tau = 10^{-12}$ a.u. and $\tau = 10^{-14}$ a.u.. Looser values than $\tau = 10^{-12}$ a.u. were used only to ascertain that certain calculations fail to converge unless $\tau \leq 10^{-12}$ a.u..

S2 Data sets

Of the 55 data sets that constitute the GMTKN55 database,[2] the following ones were used in this study: AL2X6, BH76, BHDIV10, BHROT27, C60ISO, DARC, DIPCS10, G21EA, G2RC, ISO34, MB16-43, PA26, PX13, SIE4x4, W4-11, YBDE18, ALK8, BH76RC, BHPERI, BSR36, CDIE20, DC13, FH51, G21IP, INV24, ISOL24, NBPRC, PArel, RSE43, TAUT15, and WCPT18. Other data sets contained within GMTKN55 either pertain to noncovalent interactions, or else included elements that are not supported by the 6-311G family of basis sets, and were excluded. We exclude these data sets, as did McKemmish and co-workers also.[3] The aggregate performance of various methods on all of these data sets is shown in Fig. S1, and Fig. 1 presents a subset of these functional and basis set combinations. Error statistics for all basis sets tested are presented in Table S1, from which Table 1 is drawn.

The aforementioned data sets were then grouped into three categories and the statistical performance was plotted separately for each (Figs. S2–S4). The three categories are:

- **"Easy" thermochemistry** (Fig. S2): AL2X6, DIPCS10, G21EA, G2RC, PA26, SIE4x4, W4-11, YBDE18, ALK8, BH76RC, DC13, FH51, G21IP, NBPRC, and TAUT15.

---

[*] herbert@chemistry.ohio-state.edu

- **"Hard" thermochemistry** (Fig. S3): C60ISO, DARC, ISO34, MB16-43, BSR36, CDIE20, ISOL24, PArel, and RSE43.
- **Barrier heights** (Fig. S4): BH76, BHDIV10, BHROT27, PX13, BHPERI, INV24, and WCPT18.

S3 Performance of different basis sets

Statistical summaries of the errors for each functional can be found in Table S2 (B3LYP+D3), Table S3 (M06-2X), and Table S4 (ωB97M-V). A subset of these statistics is presented in Table S1, organized in such a way that the three functionals can be compared. All three functionals behave in a nearly identical way with regard to the choice of basis set, so only ωB97M-V will be discussed in detail.

Of the basis sets that we tested, 6-311G(d), 6-311+G(d), and 6-311+G(d,p) perform quite poorly, as reported previously.[3] (We include these primarily as a means to make direct contact with the calculations in Ref. 3.) However, the performance of other basis sets is rather comparable. This includes Pople basis sets with additional polarization functions [6-311G(2df,p) and better] along with Karlsruhe basis sets of triple- and quadruple-ζ quality. Each of these basis sets exhibits a mean absolute error (MAE) around 3 kcal/mol. There is somewhat more variation in the root-mean-square error (RMSEs), which vary from 5–8 kcal/mol depending on the choice of basis set.

For the Pople basis sets, adding polarization functions on the heavy atoms, beyond those present in 6-311G(d), reduces the MAE by 2–3 kcal/mol, but additional polarization functions on the hydrogen atoms have a nearly negligible effect: $\sim$ 0.1 kcal/mol reduction in the MAE and $\sim$ 0.2 kcal/mol reduction in the RSMD. Addition of sp diffuse functions to the heavy atoms ("+") results in a further reduction by $\approx$ 0.5 kcal/mol but diffuse functions added hydrogen atoms ("++") have no statistical effect. Upon adding a full suite of polarization functions, the 6-311G(3df,2pd) family of basis sets, the summary statistics actually become slightly worse in the case of ωB97M-V, suggesting some error cancellation in basis sets with fewer polarization functions, although the effect is small ($\lesssim$ 0.5 kcal/mol in the MAE and $\approx$ 1.5 kcal/mol in the RMSE) and is not observed for either B3LYP+D3 or M06-2X.

Karlsruhe basis sets ranging from def2-TZVP,[4] def-TZVPD,[5] and the partially-augmented versions def2-ma-TZVP[6] and def2-ha-TZVP,[6] all exhibit MAEs in the range 3.0–3.4 kcal/mol, which should be compared to 3.6 kcal/mol for 6-311G(2df,p) and 3.5 kcal/mol for 6-311G(2df,2p). RMSEs for these triple-ζ Karlsruhe basis sets lie in the range 7.0–7.2 kcal/mol as compared to 6.6 kcal/mol for 6-311G(2df,p) and 6.4 kcal/mol for 6-311G(2df,2p).

Notably, the aforementioned values are lie near to the basis set limit, in the sense that the def2-QZVPD basis (MAE = 2.8 kcal/mol and RMSE = 7.3 kcal/mol) affords only slightly smaller errors. For the high-quality triple-ζ basis sets, meaning any of the Karlsruhe choices (def2-TZVP or larger) and Pople basis sets starting with 6-311G(2df,p), the statistical differences among the basis sets are comparable to differences with respect to basis-set limit (def2-QZVPD) results and much smaller than the overall spread of the errors. This suggests that there is no statistically meaningful way, based on these particular (but rather comprehensive) data, to make a recommendation for def2-TZVP over 6-311G(2df,p).

S4 Timings

All calculations were performed on dedicated compute nodes equipped with Dual Intel Xeon 8268s Cascade Lakes processors with 48 cores operating at 2.9 GHz. (All calculations were multithreaded across all 48 cores.) Timing data presented below are averages over two runs to account for minor differences on the order of a few seconds. Even for the most expensive calculations (using def2-QZVP), the difference between the two runs amounts to $< 0.6\%$ of the total time.

Figure S5 presents timing data, analogous to those in Fig. 2 but including a few more basis sets, and setting $\tau = 10^{-14}$ a.u. (Fig. S5) rather than $\tau = 10^{-12}$ a.u. (Fig. 2). Trends amongst the various basis sets are essentially the same for either choice of threshold.

Figure S6 presents timing data (using ωB97M-V) for various Pople-style basis sets, comparing the result obtained using standard composite sp shells to that obtained by separating these s and p functions, despite the fact that they share the same orbital exponent. This comparison is relevant because not all integrals codes are structured to take advantage of sp shells, and a code that does not do so will give a sub-optimal impression of the performance of Pople basis sets. For these calculations, we start from an automatically generated SAD guess, using the ωB97M-V functional and the basis set in question to generate each atomic density. For the 6-311G(2df,p) and 6-311G+(2df,p) basis sets, the calculation using sp shells is more than 30% faster than the corresponding calculation with separated shells.

References

[1] Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980,** *73,* 393–398.

[2] Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017,** *19,* 32184–32215.

[3] Pitman, S. J.; Evans, A. K.; Ireland, R. T.; Lempriere, F.; McKemmish, L. K. Benchmarking basis sets for density functional theory thermochemistry calculations: Why unpolarized basis sets and the polarized 6-311G family should be avoided. *J. Phys. Chem. A* **2023,** *127,* 10295–10306.

[4] Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7,* 3297–3305.

[5] Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010,** *133,* 134105:1–11.

[6] Gray, M.; Herbert, J. M. Comprehensive basis-set testing of extended symmetry-adapted perturbation theory and assessment of mixed-basis combinations to reduce cost. *J. Chem. Theory Comput.* **2022,** *18,* 2308–2330.

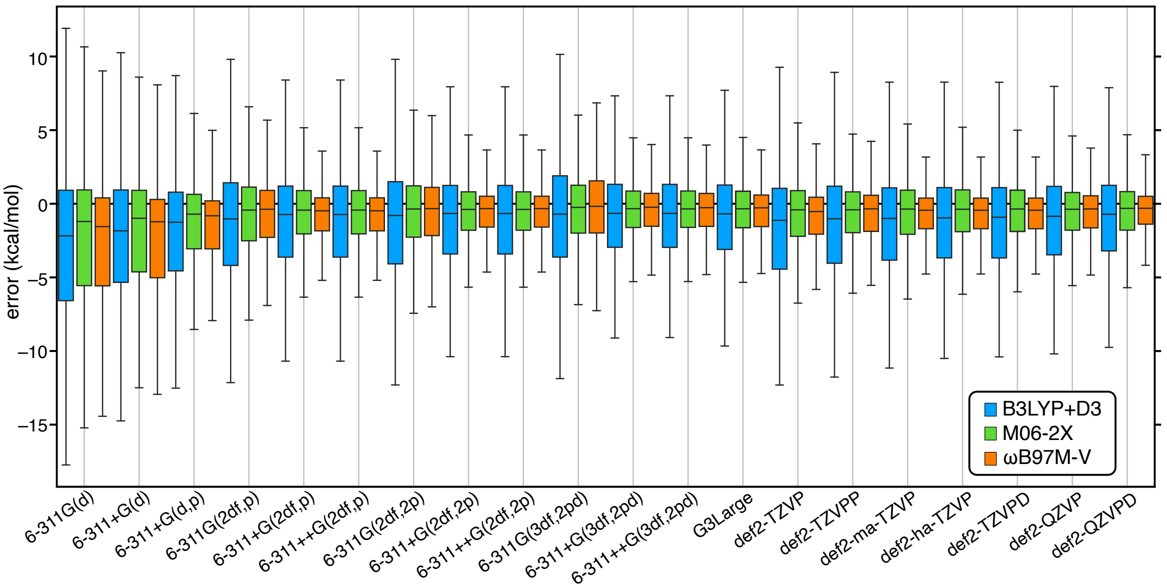

Fig. S1: Statistical summary of signed errors for all of the GMTKN55 data used in this work. Each colored box contains the middle 50% of the data points and the median error (with respect to the benchmark value) is indicated by a horizontal line Whiskers represent 1.5× times the interquartile range, which is 2.7σ for a normal distribution or a 99.3% confidence interval.

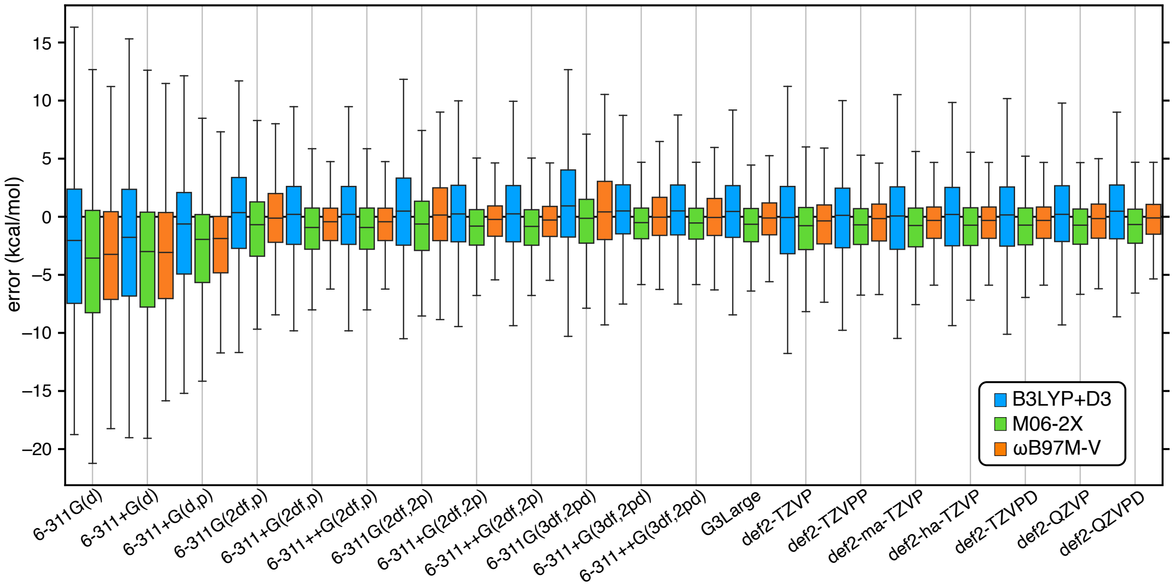

Fig. S2: Statistical summary of signed errors for the “easy thermochemistry” subset of GMTKN55. Symbols mean the same as in Fig. S1.

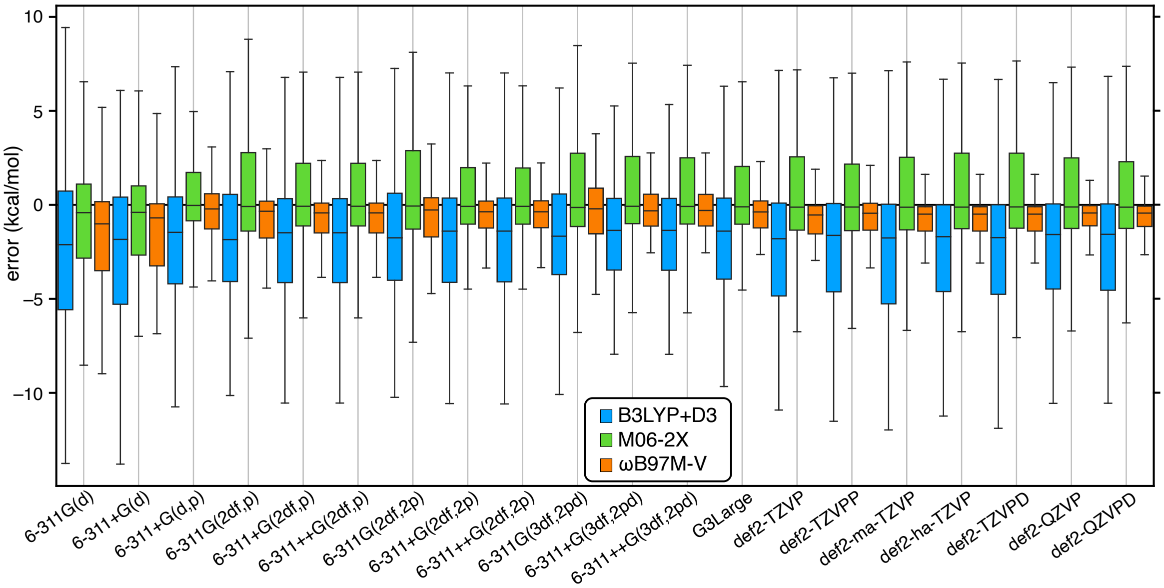

Fig. S3: Statistical summary of signed errors for the “hard thermochemistry” subset of GMTKN55. Symbols mean the same as in Fig. S1.

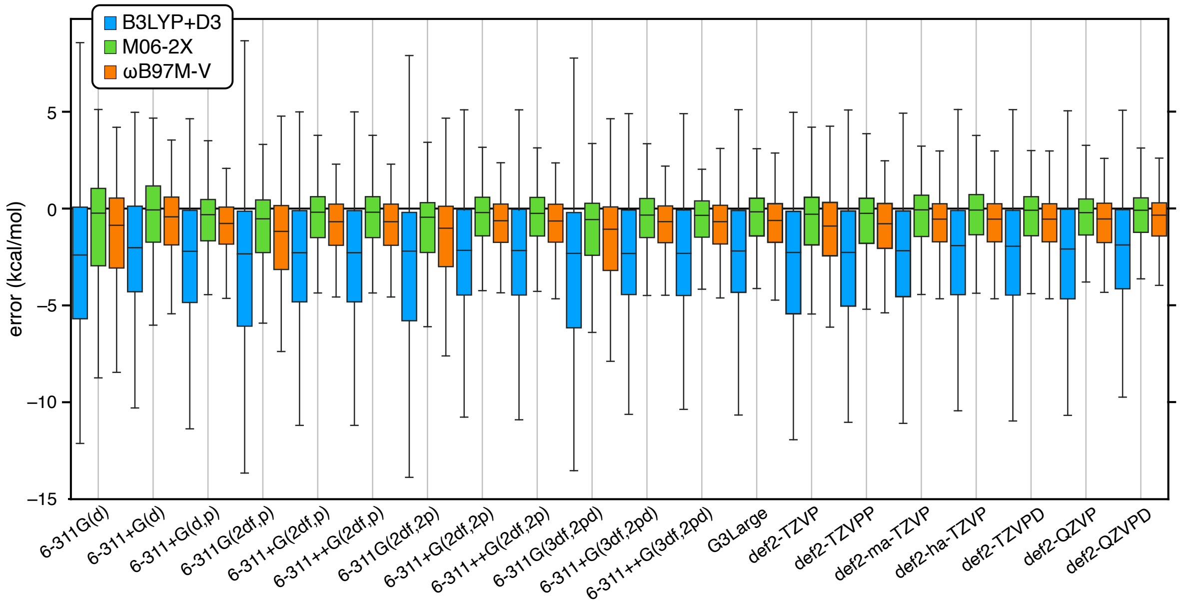

Fig. S4: Statistical summary of signed errors for barrier heights in GMTKN55. Symbols mean the same as in Fig. S1.

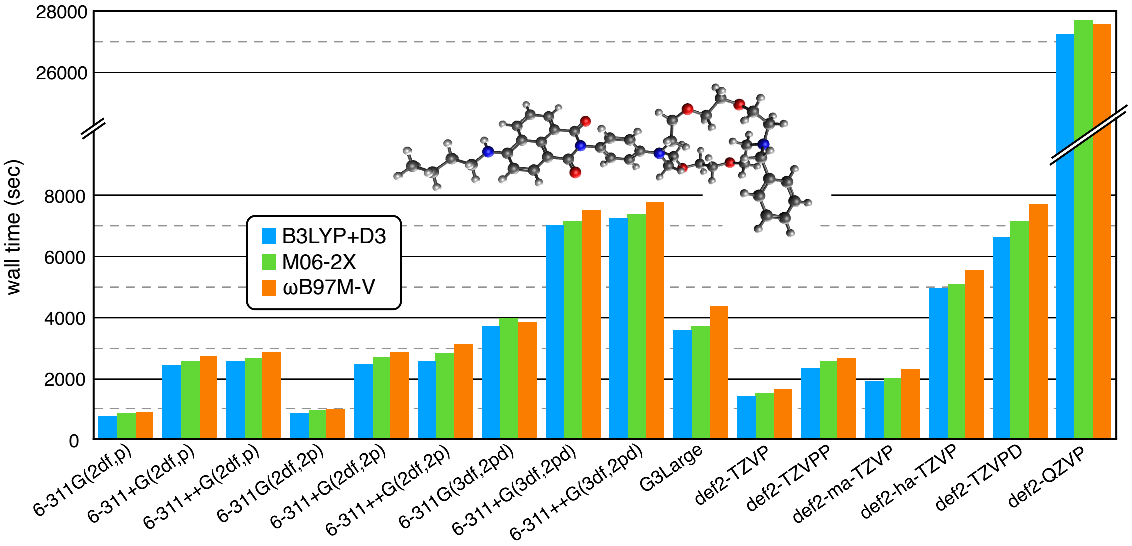

Fig. S5: Wall times (on a single 48-core node) for a single-point energy calculation on the isomer of $C_{41}H_{50}O_6N_4$ that is shown. All calculations use $\tau_{\rm SCF} = 10^{-8}$ Ha and $\tau_{\rm ints} = \tau_{\rm shlpr} = 10^{-14}$ a.u., whereas the timing data in Fig. 2 use $\tau_{\rm ints} = \tau_{\rm shlpr} = 10^{-12}$ a.u.. Each calculation converged in 14–16 SCF iterations.

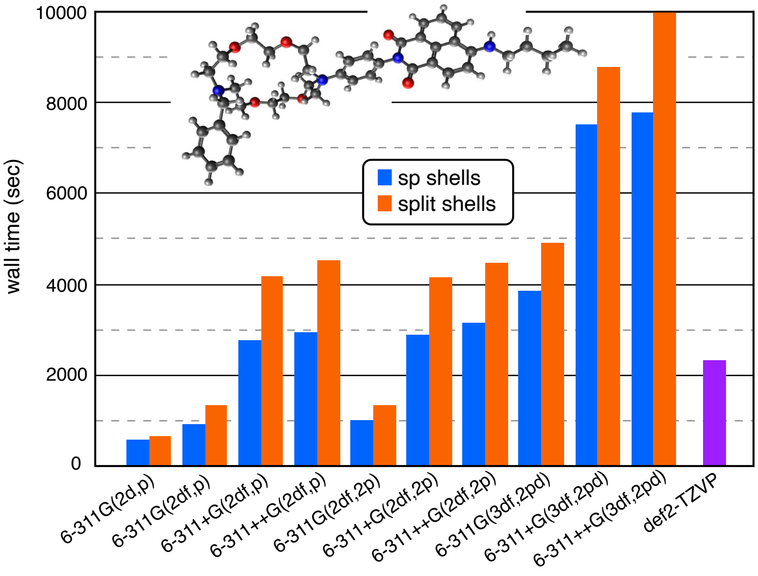

Fig. S6: Wall times (on a single 48-core node) for a single-point energy calculation on the isomer of $C_{41}H_{50}O_6N_4$ that is shown, using the ωB97M-V functional in various Pople-style basis sets and comparing the use of composite "sp" shells with split shells in which the relevant s- and p-functions have been separated even though they have the same exponent. All calculations use $\tau_{\text{SCF}} = 10^{-8}$ Ha and $\tau_{\text{ints}} = \tau_{\text{shlpr}} = 10^{-14}$ a.u., whereas the timing data in Fig. 2 use $\tau_{\text{ints}} = \tau_{\text{shlpr}} = 10^{-12}$ a.u.. Each calculation converged in either 14 or 15 SCF iterations. The timing for the split-shells 6-311++G(3df,2pd) calculation is almost exactly 10^4 s (*i.e.*, it is not cut off on this scale).

Table S1: Error statistics (in kcal/mol) for the thermochemical and kinetics subsets of the GMTKN55 database, relative to benchmark values

Basis	B3LYP+D3			M06-2X			ωB97M-V		
	median abs.[a]	MAE[b]	RMSE[c]	median abs.[a]	MAE[b]	RMSE[c]	median abs.[a]	MAE[b]	RMSE[c]
6-311G(d)	3.8	7.6	12.9	3.2	5.7	10.1	3.1	6.0	10.2
6-311+G(d)	3.6	7.1	12.8	2.7	5.0	9.4	2.7	5.3	9.7
6-311+G(d,p)	2.6	5.9	10.9	1.8	4.6	10.0	1.6	4.1	8.6
6-311G(2df,p)	2.8	5.5	9.4	1.8	4.0	7.3	1.7	3.6	6.6
6-311+G(2df,p)	2.4	4.9	9.5	1.5	3.0	5.6	1.2	2.6	5.4
6-311++G(2df,p)	2.4	4.9	9.5	1.5	3.0	5.6	1.2	2.6	5.4
6-311G(2df,2p)	2.4	5.3	9.0	1.8	3.9	7.2	1.8	3.5	6.4
6-311+G(2df,2p)	2.3	4.8	9.2	1.4	2.9	5.5	1.1	2.4	5.1
6-311++G(2df,2p)	2.3	4.8	9.2	1.4	3.0	5.5	1.1	2.5	5.1
6-311G(3df,2pd)	2.7	5.1	8.8	1.6	4.4	9.8	1.8	4.0	8.1
6-311+G(3df,2pd)	2.2	4.5	8.7	1.3	3.3	8.2	1.2	2.8	6.6
6-311++G(3df,2pd)	2.1	4.5	8.7	1.3	3.3	8.3	1.2	2.8	6.6
G3Large	2.2	4.6	8.8	1.3	2.8	5.5	1.1	2.4	4.9
def2-TZVP	2.7	5.3	10.0	1.6	3.4	7.6	1.3	3.4	7.2
def2-TZVPP	2.5	5.1	9.8	1.4	3.3	7.5	1.2	3.2	7.0
def2-ma-TZVP	2.5	5.2	10.1	1.5	3.3	7.4	1.1	3.0	7.0
def2-ha-TZVP	2.4	5.0	9.9	1.4	3.2	7.5	1.1	3.0	7.0
def2-TZVPD	2.4	5.0	9.9	1.4	3.2	7.5	1.1	3.0	7.0
def2-QZVP	2.3	4.9	9.5	1.4	3.2	8.2	1.0	2.9	7.3
def2-QZVPD	2.2	4.9	9.6	1.4	3.2	8.1	1.0	2.8	7.3

[a] Median absolute error, which is the primary error statistic in Ref. 3. [b] Mean absolute error. [c] Root-mean-square error or standard deviation.

Table S2: Error statistics (in kcal/mol) for the B3LYP+D3 functional, relative to benchmark values, for the thermochemical and kinetics subsets of the GMTKN55 database.

Basis Set	Median Abs. Error	Averages		Outliers		RMSE[c]
		MSE[a]	MAE[b]	min	max	
6-311G(d)	3.84	−4.25	7.59	−101.58	39.19	12.91
6-311+G(d)	3.59	−3.99	7.07	−102.19	39.19	12.84
6-311+G(d,p)	2.65	−3.04	5.87	−88.43	39.96	10.94
6-311G(2df,p)	2.83	−2.13	5.48	−81.66	39.96	9.36
6-311+G(2df,p)	2.38	−2.08	4.92	−83.81	39.63	9.47
6-311++G(2df,p)	2.38	−2.08	4.92	−83.81	39.63	9.47
6-311G(2df,2p)	2.38	−1.92	5.29	−77.01	40.26	9.04
6-311+G(2df,2p)	2.29	−1.86	4.78	−78.94	40.26	9.16
6-311++G(2df,2p)	2.31	−1.88	4.77	−79.14	39.92	9.15
6-311G(3df,2pd)	2.73	−1.58	5.14	−77.56	40.27	8.76
6-311+G(3df,2pd)	2.15	−1.58	4.49	−80.00	40.27	8.69
6-311++G(3df,2pd)	2.14	−1.60	4.49	−80.28	39.93	8.68
G3Large	2.19	−1.69	4.57	−78.17	39.95	8.78
def2-TZVP	2.66	−2.59	5.31	−86.12	39.73	10.02
def2-ma-TZVP	2.53	−2.47	5.16	−85.82	39.73	10.12
def2-ha-TZVP	2.42	−2.34	5.02	−84.93	39.73	9.87
def2-TZVPP	2.46	−2.38	5.12	−83.95	39.97	9.77
def2-TZVPD	2.40	−2.32	5.02	−85.51	40.19	9.90
def2-QZVP	2.27	−2.10	4.88	−83.23	39.91	9.48
def2-QZVPD	2.25	−2.04	4.86	−83.50	39.97	9.63

[a]Mean (signed) error. [b]Mean absolute error. [c]Root-mean-square error.

Table S3: Error statistics (in kcal/mol) for the M06-2X functional, relative to benchmark values, for the thermochemical and kinetics subsets of the GMTKN55 database.

Basis Set	Median Abs. Error	Averages		Outliers		RMSE[c]
		MSE[a]	MAE[b]	min	max	
6-311G(d)	3.17	−2.08	5.72	−58.96	89.24	10.08
6-311+G(d)	2.72	−1.78	5.01	−61.81	90.05	9.42
6-311+G(d,p)	1.84	−0.56	4.61	−61.81	91.98	9.95
6-311G(2df,p)	1.82	−0.11	4.04	−36.54	41.84	7.34
6-311+G(2df,p)	1.48	−0.02	3.05	−29.26	33.63	5.56
6-311G++(2df,p)	1.48	−0.02	3.05	−29.26	33.63	5.56
6-311G(2df,2p)	1.76	−0.02	3.92	−36.54	41.57	7.25
6-311+G(2df,2p)	1.38	0.08	2.94	−29.26	32.06	5.46
6-311++G(2df,2p)	1.39	0.07	2.95	−29.26	31.97	5.47
6-311G(3df,2pd)	1.63	0.70	4.35	−37.16	92.65	9.85
6-311+G(3df,2pd)	1.28	0.76	3.32	−29.32	88.59	8.25
6-311++G(3df,2pd)	1.27	0.75	3.33	−29.32	88.56	8.26
G3Large	1.34	−0.24	2.85	−29.16	35.75	5.48
def2-TZVP	1.56	0.12	3.44	−28.91	79.66	7.62
def2-ma-TZVP	1.50	0.23	3.27	−28.92	79.11	7.45
def2-ha-TZVP	1.39	0.35	3.25	−28.94	78.76	7.54
def2-TZVPP	1.43	0.20	3.32	−28.91	79.27	7.53
def2-TZVPD	1.36	0.36	3.25	−28.94	78.73	7.54
def2-QZVP	1.43	0.43	3.25	−29.26	92.38	8.18
def2-QZVPD	1.38	0.48	3.16	−29.27	92.43	8.08

[a] Mean (signed) error. [b] Mean absolute error. [c] Root-mean-square error.

Table S4: Error statistics (in kcal/mol) for the ωB97M-V functional, relative to benchmark values, for the thermochemical and kinetics subsets of the GMTKN55 database.

Basis Set	Median	Averages		Outliers		RMSE[c]
	Abs. Error	MSE[a]	MAE[b]	min	max	
6-311G(d)	3.12	−2.93	6.00	−76.95	66.06	10.24
6-311+G(d)	2.67	−2.66	5.30	−77.78	65.51	9.68
6-311+G(d,p)	1.61	−1.74	4.07	−65.68	67.29	8.55
6-311G(2df,p)	1.71	−1.09	3.62	−59.49	30.24	6.61
6-311+G(2df,p)	1.17	−1.09	2.64	−61.77	29.81	5.45
6-311++G(2df,p)	1.17	−1.09	2.64	−61.77	29.81	5.45
6-311G(2df,2p)	1.75	−0.83	3.52	−53.13	30.64	6.40
6-311+G(2df,2p)	1.09	−0.81	2.49	−55.29	30.64	5.11
6-311++G(2df,2p)	1.09	−0.84	2.49	−55.64	30.20	5.11
6-311G(3df,2pd)	1.81	−0.20	4.00	−54.39	68.30	8.09
6-311+G(3df,2pd)	1.22	−0.25	2.79	−57.15	63.44	6.62
6-311++G(3df,2pd)	1.22	−0.29	2.77	−57.44	63.41	6.61
G3Large	1.11	−0.64	2.37	−54.72	30.23	4.86
def2-TZVP	1.28	−1.27	3.37	−64.30	53.00	7.21
def2-ma-TZVP	1.09	−1.04	3.00	−63.41	51.95	6.95
def2-ha-TZVP	1.09	−1.04	3.00	−63.41	51.95	6.95
def2-TZVPP	1.16	−1.04	3.18	−61.10	52.40	6.97
def2-TZVPD	1.09	−1.04	3.00	−63.41	51.95	6.95
def2-QZVP	1.05	−0.65	2.87	−57.74	71.51	7.26
def2-QZVPD	0.98	−0.62	2.83	−58.96	71.52	7.34

[a] Mean (signed) error. [b] Mean absolute error. [c] Root-mean-square error.