

Assessing Density Functional Theory for Chemically Relevant Open-Shell Transition Metal Reactions

Leonard R. Maurer,[†] Markus Bursch,[†] Stefan Grimme, and Andreas Hansen*Cite This: *J. Chem. Theory Comput.* 2021, 17, 6134–6151

Read Online

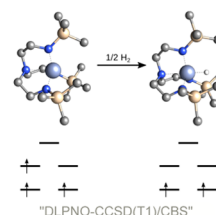
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Due to the principle lack of systematic improvement possibilities of density functional theory, careful assessment of the performance of density functional approximations (DFAs) on well-designed benchmark sets, for example, for reaction energies and barrier heights, is crucial. While main-group chemistry is well covered by several available sets, benchmark data for transition metal chemistry is sparse. This is especially the case for larger, chemically relevant molecules. Addressing this issue, we recently introduced the MOR41 benchmark which covers chemically relevant reactions of closed-shell complexes. In this work, we extend these efforts to single-reference open-shell systems and introduce the “reactions of open-shell single-reference transition metal complexes” (ROST61) benchmark set. ROST61 includes accurate coupled-cluster reference values for 61 reaction energies with a mean reaction energy of $-42.8 \text{ kcal mol}^{-1}$. Complexes with 13–93 atoms covering 20 d-block elements are included, but due to the restriction to single-reference open-shell systems, important elements such as iron or platinum could not be taken into account, or only to a small extent. We assess the performance of 31 DFAs in combination with three London dispersion (LD) correction schemes. Further, DFT-based composite methods, MP2, and a few semiempirical quantum chemical methods are evaluated. Consistent with the results for the MOR41 closed-shell benchmark, we find that the ordering of DFAs according to Jacob’s ladder is preserved and that adding an LD correction is crucial, clearly improving almost all tested methods. The recently introduced $r^2\text{SCAN-3c}$ composite method stands out with a remarkable mean absolute deviation (MAD) of only $2.9 \text{ kcal mol}^{-1}$, which is surpassed only by hybrid DFAs with low amounts of Fock exchange (e.g., $2.3 \text{ kcal mol}^{-1}$ for TPSS0-D4/def2-QZVPP) and double-hybrid (DH) DFAs but at a significantly higher computational cost. The lowest MAD of only $1.6 \text{ kcal mol}^{-1}$ is obtained with the DH DFA PWPB95-D4 in the def2-QZVPP basis set approaching the estimated accuracy of the reference method. Overall, the ROST61 set adds important reference data to a sparsely sampled but practically relevant area of chemistry. At this point, it provides valuable orientation for the application and development of new DFAs and electronic structure methods in general.

Benchmark for Realistic
Open-Shell Transition Metal
ReactionsB97-3c B97-D3 BP86
M BLYP revPBE r²SCAN
SCAN r²SCAN-3c B97M
PSS PBE revTPSS M06-L
PBEh-3c PBE0 B3LYP
MN15 TPSS0 TPSSh
ωB97X PW6B95
ωB97M M06 M06-2X
revDSD-BLYP B2GP-PLYP
mPW2PLYP PBE0-DH
PBP95 revDOD-PBEP86
r²PLYP revDSD-PBEPBE
revDSD-PBEP86
MP2 HF-3c

1. INTRODUCTION

Organometallic chemistry can be defined as the link between metal-based inorganic and synthesis-orientated organic chemistry. Over the last decades, it became an important and independent research field.^{1–7} The ongoing optimization of transition metal complexes facilitates the conversion of functional groups for synthetic applications and uncovers unprecedented binding motives.⁸

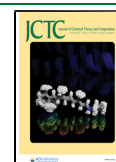
Computational chemistry is able to contribute significantly to this field, allowing for the design of new complexes and the detailed elucidation of reaction mechanisms. It particularly helps where the limits of the experiment are reached and, for example, reactive intermediates can no longer be detected. Since the size of the investigated transition metal complexes for which reliable predictions are desired grows ever larger, the choice of computational methods is limited. In this context, Kohn–Sham density functional theory (KS-DFT) has become the common workhorse of computational chemistry due to its beneficial accuracy/cost ratio.^{9,10} Consequently, a large “zoo” of density functional approximations (DFAs) was developed in

the last decades focusing on various purposes also including transition metal chemistry.¹¹ In the following, KS-DFT will only be referred to as DFT for simplicity. For explicit density functionals, the term DFA will be used.

However, none of these DFAs is free of approximations and shortcomings such as the delocalization or self-interaction error (SIE)^{12,13} and the lacking description of London dispersion (LD) interactions.¹⁴ To find a well-suited DFA for the chemical problem under investigation, careful evaluation on test sets with reliable experimental or theoretical reference data is indispensable.^{15,16} Comprehensive databases such as MGAE109,¹⁷ MGCDB84,¹⁵ GMTKN55,¹⁶ G3/05,¹⁸

Received: July 1, 2021

Published: September 21, 2021



and other sets^{19–21} are available for main-group reactions. For transition metal reactions, although ubiquitous in chemistry and biology, much less benchmark data is available. This is particularly the case for larger and open-shell complexes. Mostly, benchmark sets covering closed-shell transition metal chemistry are available which are based on measured enthalpies of formation or experimentally derived bond lengths. Noteworthy examples were compiled by Truhlar et al.^{22–27} or Li and Dixon.^{28–31} Further, Nielsen and Allendorf investigated small transition metal complexes (<30 atoms), comparing experimental data to high-level *ab initio* coupled-cluster (CC) and DFT results.^{32–34} In 2015, Truhlar et al. published the 3dMLBE20 benchmark study,³⁵ containing 20 bond dissociation energies of third-row diatomic molecules. Later, this work was revised by Cheng et al.³⁶ and Dixon et al.,³⁷ who then re-investigated the experimental values and validated the ability of complete basis set (CBS) limit extrapolated CCSD(T) to yield reasonable reference values. This was also confirmed by the work of Köhn and de Oliveira-Filho³⁸ for systems where non-dynamical correlation effects can largely be excluded.³⁹

Recently, Shee et al. investigated 44 bond dissociation energies of diatomic 3d transition metal systems.⁴⁰ However, the conclusiveness of such studies based on small molecules is limited since chemically more relevant systems with organic ligands typically have a lower degree of d-orbital degeneracy at the metal atom.⁴¹ Hence, many larger organometallic complexes may be described well by single-reference wave function theory (WFT) or DFT methods, while this is often not the case for small transition metal systems with rather complicated electronic structure. Furthermore, studies by Cundari et al. and Wilson et al.^{42–44} compare accurate experimental results to DFT and CC theory but are again restricted to small systems (<30 atoms). Wilson et al. also presented the ccCA-TM/11 benchmark set with 193 experimental enthalpies of formation for small transition metal complexes.^{45,46} The ccCA-TM/11 benchmark set includes both single-reference and multi-reference cases but is once more limited to a system size of about 20 atoms. Truhlar et al. further compiled a subset of ccCA-TM/11 termed 3dBE70⁴⁷ and assessed a much larger number of DFAs on it. Besides more DFT benchmark studies on small transition metal complexes,^{48–53} an early WFT benchmark by Hyla-Kryspin and Grimme should be mentioned as well.⁵⁴ It features small carbonyl complexes, metallocene reactions, and transition metal complexes with small organic ligands. In a study by Martin et al., the performance of DFAs for prototype reactions of small organic molecules with elemental palladium was assessed.⁵⁵ This work was later addressed by Steinmetz and Grimme, who expanded the study by palladium and nickel complexes.⁵⁶ Recently, Martin and Efremenko also evaluated DFAs for Ru(II)-catalyzed hydroarylation and oxidative coupling reactions of small molecules.⁵⁷

None of the studies mentioned so far includes transition metal complexes with more than about 30 atoms. This is mainly due to missing experimental thermochemical data (especially in the gas phase) and the difficulty to calculate highly accurate theoretical reference data for large complexes. The extensive computational cost of methods suited to treat significant amounts of dynamic and, in some systems, also non-dynamic correlation effects, prevents their application to larger systems.⁵⁸ However, many important organometallic research

areas feature the latter, for example, in the form of sterically demanding ligands.

Only a few benchmark studies on large transition metal complexes were carried out so far. In 2011, Hughes and Friesner⁵⁹ compiled a set of 57 third-row transition metal complexes with large organic ligands to examine the performance of DFAs for spin-splitting energies of closed- and open-shell complexes, all accompanied with experimental reference values. Further gas-phase reaction energies were compiled in the WCCR10 set by Chen, Reiher et al.⁶⁰ This set contains 10 ligand dissociation reactions composed of transition metal complexes with up to 174 atoms. The reference energies were derived from tandem mass spectrometry, and hence, all substrates and products are charged closed-shell species. Several DFAs were tested, leading to the conclusion that inclusion of Fock exchange can improve the results, at least for PBE0⁶¹ and TPSSH.⁶² In a follow-up study by Reiher et al., the reactions were classified by their degree of multi-reference characteristic.⁶³ Both multi-reference and single-reference methods were then examined on the WCCR10 set, showing rather poor agreement with the experiment. Minenkov et al. demonstrated the accuracy of local CC for calculating non-covalent gas phase ligand dissociation energies of coinage metal cation complexes, but a comprehensive DFA benchmark was not carried out.⁶⁴ The MOR41 benchmark study by two of the present authors comprises 41 closed-shell organometallic reaction energies of chemically relevant transition metal complexes with up to 120 atoms.⁶⁵ Another comprehensive benchmark collection by Chan et al.⁶⁶ called TMC151 covers reactions of coinage metal complexes⁶⁷ and previously mentioned benchmark sets.

Despite the reasonable number of test sets now available for small open-shell transition metal systems, there remains an ongoing need for a comprehensive benchmark set of chemically relevant organometallic reactions involving open-shell complexes of typical system sizes (30–100 atoms). At this point, it is desirable to specifically investigate whether the calculation of reaction energies for transition metal complexes with unpaired electrons is generally more challenging for commonly applied DFAs. Further, the open question of whether insights from studies on closed-shell systems are transferable to single-reference open-shell ones is of general interest for both computational chemists and method developers.

The compilation of a benchmark set including reliable reference values for the corresponding reaction energies is one of the main objectives of this work. This is particularly important as the focus in developing efficient DFT and semiempirical quantum chemical (SQM) methods is increasingly on organometallic systems. Further, the present study aims at elucidating the performance of commonly applied and recently developed DFAs for modeling single-reference open-shell organometallic reactions and to give recommendations on which DFT methods are the most suitable for this purpose. To this end, we have compiled the reactions of open-shell single-reference transition metal complex (ROST61) benchmark including a large and diverse set of open-shell organometallic reactions. Accurate local CC calculations were performed to generate reliable reference reaction energies that are subsequently used to assess various DFAs in conjunction with frequently used LD correction schemes.

In the next sections, the computational details are given, followed by a description of the ROST61 benchmark set and

details on the computation of the reference values. Furthermore, an extensive assessment of various DFAs and a few WFT and SQM methods is presented. Finally, conclusions are drawn and perspectives are given.

2. COMPUTATIONAL DETAILS

Quantum chemical calculations were carried out with the ORCA^{68–71} (V. 4.0, V. 4.1, and V. 4.2.1 for the DLPNO-CCSD(T1) computations), TURBOMOLE 7.5.1 (r²SCAN and r²SCAN-3c),^{72,73} MOPAC2016⁷⁴ (PM6-D3H4X and PM7 calculations), and xtb 6.4.1⁷⁵ (GFN2-xTB calculations) quantum chemistry packages. MN15 and revM06-L calculations were enabled by the LibXC⁷⁶ density functional library implementation in ORCA. In all non-composite WFT and DFT calculations, the Ahlrichs'-type def2-basis sets^{77,78} (def2-SVP, def2-TZVPP, and def2-QZVPP) were applied. The def2-QZVPP basis set was applied for all final DFT energy calculations. All composite DFT methods apply their inherent modified basis sets. Respective Stuttgart/Dresden effective core potentials (ECP) SD(28,MWB) and SD(60,MWB) and matching ECP basis sets for 4d and 5d systems (both obtained from the TURBOMOLE basis set library) were used. The ORCA *TightSCF* convergence criterion was employed throughout. "Grid5" numerical integration settings for the DFT calculations were applied. Molecular structures were pre-optimized with the semiempirical GFN2-xTB method.^{79,80} For final geometry optimizations, the efficient composite DFT method B97-3c was used, which is known to yield transition metal geometries at an excellent cost–accuracy ratio.⁸¹ DLPNO-CCSD(T)^{82–84} and DLPNO-CCSD(T1)^{85,86} in its sparse-maps implementation was applied for closed-shell and open-shell molecules with corresponding def2-SVP/C, def2-TZVPP/C, and def2-QZVPP/C auxiliary basis sets.⁸⁷ Both *VeryTightPNO*⁸⁸ and *TightPNO*⁸⁹ threshold settings were applied (cf. Section 3.3). Since the full local MP2 guess was not yet available for open-shell DLPNO-CCSD(T/T1) calculations in the used ORCA versions, it was also disabled for closed-shell calculations to ensure an unbiased comparison between open- and closed-shell systems. The obtained electronic energies were not corrected by vibrational zero-point or further relativistic corrections despite the already mentioned ECPs.

To reduce the basis set incompleteness (BSIE) and superposition (BSSE) errors of CC energies, a CBS limit extrapolation was performed separately for the HF and correlation part according to the scheme proposed by Neese and Valeev.⁹⁰ For all (meta-) generalized gradient approximation (GGA) functionals, the resolution of the identity (RI-J)⁹¹ approximation was employed. For the exchange part in hybrid and double-hybrid (DH) DFAs, the RI-JK method⁹² was applied. Matching def2/J and def2/JK⁹³ auxiliary basis sets were used as implemented in ORCA. The MP2 correlation part in the DH functionals was accelerated using RI with the def2-QZVPP/C auxiliary basis set.

To account for long-range LD interactions in DFAs, the D3 correction with zero damping D3(0)⁹⁴ or Becke–Johnson damping D3(BJ)⁹⁵ and Axilrod–Teller–Muto (ATM) three-body contribution^{96,97} was applied. Further, the recent D4 correction^{98,99} in its default (classical electronegativity equilibration partial charges and ATM) was used for most functionals.^{100,101} For some DFAs, also the non-local (NL) density-dependent VV10 correction method¹⁰² was employed.

3. ROST61 BENCHMARK SET

3.1. Design of the ROST61 Benchmark Set. As pointed out in the Introduction section, a comprehensive benchmark set for thermochemistry of organometallic open-shell complexes, composed of larger-sized complexes (30–100 atoms), was not yet available in the literature. The presented ROST61 set is supposed to fill this gap, and the following design criteria were defined as a framework: (i) maximum focus on open-shell complexes, (ii) only mononuclear complexes, (iii) chemically relevant and diverse reactions with "experimental background" for a large number of 3d-, 4d-, and 5d-transition metals, (iv) realistic complex size, and (v) exclusion of complexes with multi-reference characteristic.

To clearly separate open-shell from closed-shell chemistry, both substrate and product complexes have to be ground-state open-shell complexes; that is, reactions with closed-shell to open-shell transformations or excited states of open-shell complexes were not considered. Furthermore, we have searched the literature for chemically relevant reactions; that is, only experimentally characterized, postulated, or chemically reasonable transition metal complexes were taken into account. To allow for general method recommendations, reactions for the whole d-block were searched. The distribution of included elements to some extent reflects their natural abundance in chemical applications due to the initial experiment-based selection process. Therefore, frequently used elements in organometallic chemistry such as Ti and Cr are more prevalent (cf. Figure 7) in the ROST61 set. Unlike previous benchmark sets, the size of the transition metal complexes compiled in the ROST61 set resemble realistic organometallic applications, where often large organic ligands occur. Hence, medium- to large-sized complexes in the range of 30–100 atoms are covered for which accurate theoretical reference data can still be generated (cf. Section 3.3).

Another key aspect is to obtain reliable reference values with single-reference CC theory and to ensure that common DFAs are able to treat the reactions comprised in this benchmark set at least qualitatively correctly. Therefore, transition metal complexes with severe multi-reference characteristic were excluded based on the screening protocol discussed below. Last, we constrained the transition metal complexes in ROST61 to mononuclear systems. Though this also excludes important parts of organometallic chemistry, this restriction is needed to prevent multi-reference electronic structures due to, for example, complicated spin–spin couplings, which cannot be treated by single-reference CC and DFT unless cumbersome broken symmetry¹⁰³ or other special approaches are applied.^{104–106}

Based on the above-mentioned criteria, diverse reactions were found in the literature, though a major part (about two-third) was discarded in the selection process (*vide infra*). Of the 61 finally chosen reactions, 38 (62%) belong to the 3d, 16 (26%) to 4d, and 7 (11%) to the 5d block. Overall, 20 different d-block elements are covered. The reaction distribution also reflects some chemical trends in the d-block: the typically oxophilic elements of group 3 (Sc, Y, and La) often form compounds in their highest possible oxidation state +III, where typically no open-shell electronic structure is observed.⁴ Descending from 3d to 4d and 5d, higher oxidation states are generally favored and spin pairing energies are lower, thus reducing the number of observed open-shell species drastically. Further, the elements of group 10–12 often form either

closed-shell, bi- or oligonuclear compounds that are excluded from this benchmark set. Therefore, the 3d block and specifically the elements Ti and Cr are of major interest for many realistic applications in open-shell single-reference transition metal chemistry and are thus more highly represented in the ROST61 set. Nevertheless, the number of suitable 3d-reactions substantially decreased during the selection process, especially for V, Mn, Cr, Fe, Co, and Ni since complexes of these elements often exhibit multi-reference characteristic due to (nearly) degenerate orbitals.

In total, the ROST61 set comprises 61 reactions involving 150 molecules with up to 93 atoms, and the mean transition metal complex size amounts to 47 atoms. The mean reference reaction energy is $-42.8 \text{ kcal mol}^{-1}$ with all reactions being exothermic to allow a meaningful statistical evaluation of systematic errors. Five exemplary reactions are depicted in Figure 1, and a detailed list of all reactions can be found in the Supporting Information (Section S1.2). Most of the reactions compiled in the ROST61 set belong to the ligand (de)-coordination/exchange type. Other common organometallic

reactions such as oxidative additions, reductive eliminations, σ -bond-metatheses, and representatives for migratory insertions are also covered. Miscellaneous reactions such as intramolecular hydrogen abstractions, reductive couplings, or isomerizations are included as well. Hence, the ROST61 benchmark set comprises both prominent and specialized organometallic reaction types. Further, a variety of ligands regularly used in modern synthetic chemistry are represented.

3.2. Screening Protocol to Select the Final Reactions.

We conceived a multi-step protocol to further assess the suitability of the reactions found in the literature for a well-designed benchmark set. First, we pre-optimized all molecule geometries employing GFN2-xTB⁷⁹ and finally optimized them with the B97-3c DFT composite method. Further, numerical frequency calculations at the final geometry optimization level were conducted to validate the optimized structures as minima on the respective potential energy surface. In previous studies, both B97-3c and GFN2-xTB were found to be reasonably accurate and robust in geometry optimization of transition metal complexes.^{81,107} In general, the influence of the DFA choice for geometry optimization on the quality of the relative energy calculation is expected to be small ($<1 \text{ kcal mol}^{-1}$) and much smaller than for the final energy calculation (see the Supporting Information, Section S3.5, for an exemplary verification). The further selection process is exemplarily depicted for two representatives in Figure 2 (further details and more examples on the selection process can be found in the Supporting Information, Section S1.1).

The first example is the association reaction of triphenylphosphane to dichloro(cyclopentadienyl)ruthenium CpRuCl_2 (A) forming $\text{CpRuCl}_2\text{PPh}_3$ (B). The second exemplary reaction was published by Teuben et al. and refers to an insertion of 2 equiv of *tert*-butyl-nitrile into an aryne-coordinated vanadium complex C, forming the imido vanadium(III) complex D.¹⁰⁸

For the generation of reliable reference values with single-reference CC methods, it must be ensured that the molecules included in the benchmark set are not multi-reference systems. Such exhibit significant non-dynamic correlation contributions as usually indicated by a manifold of energetically low-lying, strongly coupled states. Following Hollet and Gill,¹⁰⁹ the latter can be distinguished into type A (absolute near-degeneracy correlation) and B (relative near-degeneracy correlation), even if in reality these two types are almost always present in a mixture. However, this classification is helpful in selecting suitable multi-reference indicators and interpreting their results.¹¹⁰ The fractional occupation number weighted electron density (FOD) analysis,^{111,112} a diagnostic tool to indicate and visualize the influence of primarily type A non-dynamic correlation in a molecule, was applied in the first selection step. A small (C, D) to moderate (A, B) metal-centered FOD was observed for all four complexes (see Figure 2), indicating the absence of significant non-dynamic correlation contributions. For A and B, the Cl ligands also show some FOD, and the same holds for the coordinating nitrogen atom of the imido ligand of complex D, but this is considered as less problematic wrt non-dynamical correlation.¹¹¹ Other open-shell transition metal complexes found in the literature, for example, $[(\text{N}(\text{C}(\text{CH}_3)(\text{CH}_2\text{PCy}_2))_2)\text{Fe}(\text{bipy})]$ were discarded during the screening process based on the results of FOD analysis.¹¹¹

Opposed to the complexes A–D, this complex shows a large FOD that is delocalized over the whole bipyridine ligand (cf. Supporting Information, Figure S2). In this manner, FOD

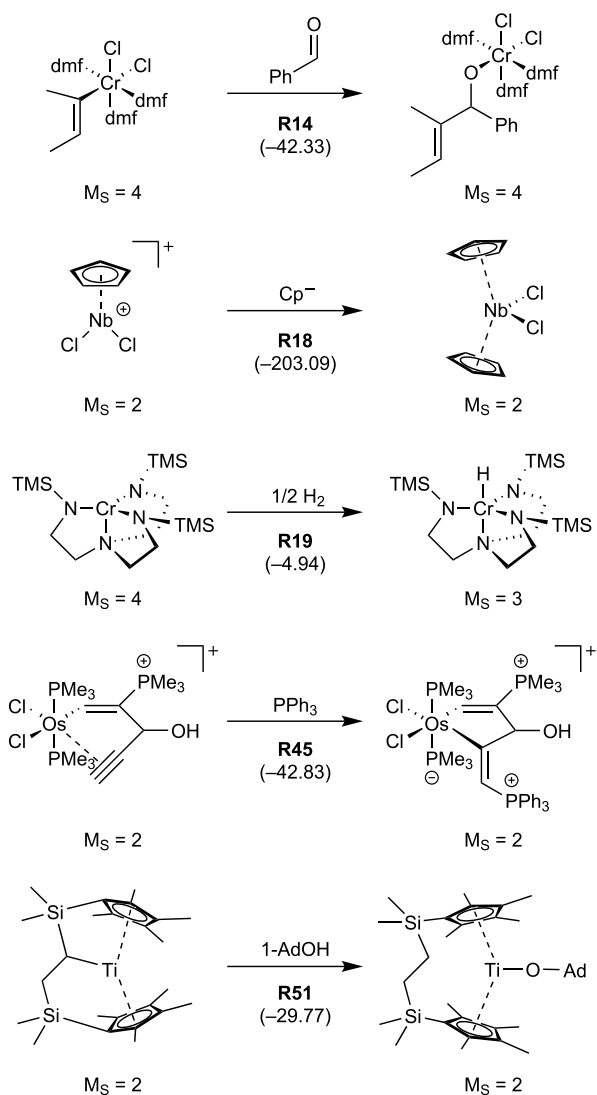


Figure 1. Selected reactions from the ROST61 benchmark set with their reference reaction energies given in parentheses in kcal mol^{-1} . M_S = multiplicity.

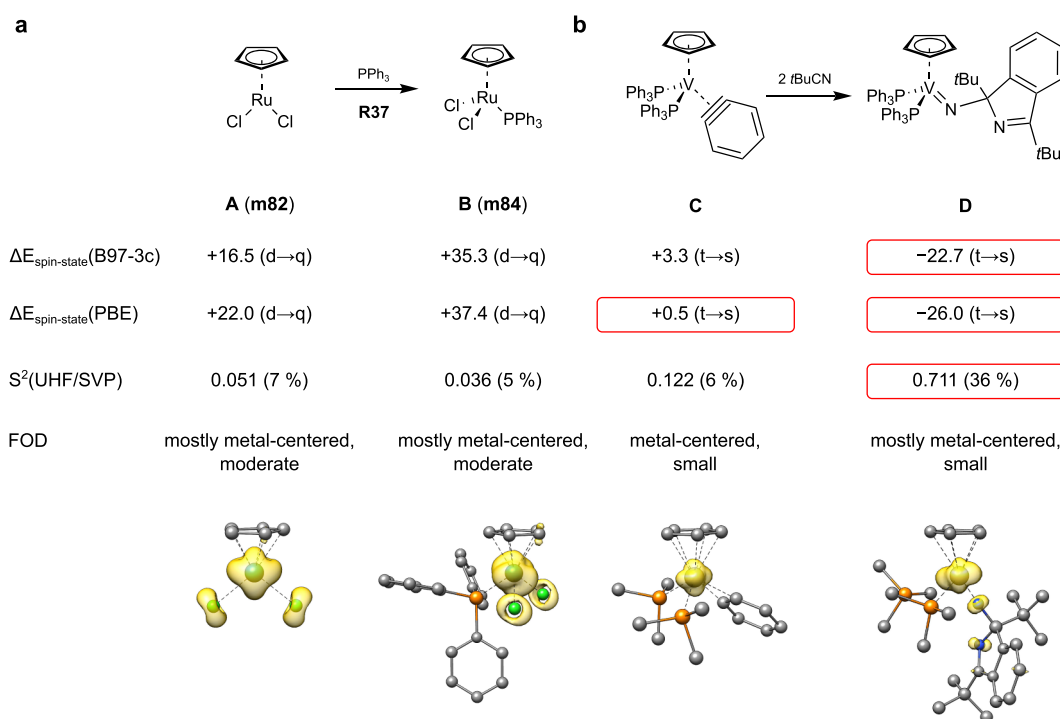


Figure 2. Exemplary selection process for a reaction included in the ROST61 set (a) and a reaction which was discarded (b). s, d, t, q = singlet, doublet, triplet, and quartet spin states. Values circled in red exceed the thresholds of the respective criterion. An isosurface value = 0.005 e bohr^{-1/2} is applied for the FOD plots. Hydrogen atoms are omitted for clarity.

analysis was performed and conservatively interpreted for all transition metal complexes potentially suitable for the benchmark that could not already be sorted out based on large N_{FOD} values.

As a second screening criterion, we calculated spin-state splitting energies since small values often indicate a multi-reference case.^{59,113} If no experimental information about the correct spin multiplicity of the ground state was available, we optimized the respective molecule in different chemically reasonable spin states employing the B97-3c method. Additionally, we compared the corresponding spin-state splitting energies to the respective PBE-D3(BJ) values (calculated at the B97-3c geometries). To verify the basic applicability of GGAs for spin-state splitting energies in such a selection process, we have additionally calculated PBE0 values for the systems discussed exemplarily in the [Supporting Information](#) (see Section S1.1 and Table S1). We chose a small threshold of 3.0 kcal mol⁻¹ for the exclusion of complexes due to energetically close spin states (if either method predicts a smaller value, the corresponding transition metal complex was no longer considered for the benchmark set) in order to also keep electronically more difficult but still dominant single-reference reactions in the benchmark set.^{114,115}

The doublet-quartet splittings of reaction a (cf. Fig. 2) favor the doublet state for both the substrate and the product, with similar B97-3c and PBE-D3(BJ) values for A and B. The situation for reaction b is different. While the B97-3c splitting for C is predicted to be slightly larger than 3.0 kcal mol⁻¹, the respective PBE-D3(BJ) value falls below the threshold energy. Moreover, the electronic minimum of D clearly refers to a closed-shell singlet state favored by more than 20 kcal mol⁻¹.

Furthermore, we analyzed the spin contamination at the UHF level of theory as large values can spoil the accuracy of our CC-based protocol for the generation of reference reaction

energies (*vide infra*).¹¹⁶ Since the basis set dependence of the spin contamination is typically small (see the comparison for the def2-SVP, def2-TZVPP, and def2-QZVPP basis sets in the [Supporting Information](#), Section S1.1), we performed these initial calculations with the def2-SVP basis set. The relative deviation from the expected $\langle S^2 \rangle$ value for A and C is acceptably small (<10%). In contrary, D shows an increased spin contamination of 36%. The high spin contamination indicates that the found SCF solution may not be the energetically lowest one and may introduce an SCF instability. Consistent with the other criteria of the screening protocol, the spin contamination also identifies reaction b as not suitable for our benchmark, and hence, it was discarded.

Based on the outlined criteria, we assessed all systems extracted from the literature. After carefully excluding all problematic molecules, the final benchmark set comprises 61 reactions (ROST61). For these reactions, we calculated CC reference reaction energies. CCSD(T) provides, at least in combination with large basis sets and for single-reference molecules, reliable reaction energies with chemical accuracy (i.e., 1 kcal mol⁻¹ residual error).¹¹⁷ This also holds true largely for open-shell transition metal complexes,^{118–120} although the error margins are typically twice as large.⁴⁵ We analyzed several indicators for non-dynamic correlation (*vide infra*) to verify that only single-reference complexes are included in the ROST61 set. Since we aim at a realistic benchmark set for chemically relevant organometallic reactions, it is essential that reliable reference values can also be calculated for large molecules, at least up to 100 atoms. However, further approximations are necessary to render CCSD(T) calculations for molecules of this size in combination with reasonably large basis sets computationally feasible. We used the open-shell variants of DLPNO-CCSD(T)⁸³ and DLPNO-CCSD(T1)⁸⁶ methods by Neese

et al. implemented in the ORCA program package, which enables open-shell CCSD(T) computations with up to about 5000 basis functions in reasonable computing times (days to weeks, depending on the threshold settings) with only minor additional errors.^{82,86}

A reasonably small spin contamination in the UHF reference wave function makes it unlikely that the UHF iterations did not converge to the lowest energy solution. Therefore, this was not considered as an exclusion criterion as CCSD(T) is less sensitive to spin contamination than UHF since it is only a second-order effect in CC theory.¹²¹ Additionally, in the open-shell DLPNO-based CC methods, quasi-restricted orbitals¹²² are employed that practically do not suffer from spin contamination without introducing a significant additional error.⁸² The actual spin contamination of the DLPNO-CCSD(T)/def2-TZVPP calculations for ROST61 further supports this statement (see the [Supporting Information](#), Table S6). The CC equations are still solved in a spin-unrestricted form, typically leading to slightly higher absolute energies compared to the respective CCSD(T) values obtained with UHF orbitals.⁸²

To further validate the reliability of single-reference CCSD(T) for the molecules in the ROST61 benchmark, the maximum values of the double-excitation amplitudes were analyzed since larger values can also indicate multi-reference cases.¹²³ Following the arguments of Wilson et al.,¹²⁴ values below 0.15 indicate a dominating single-reference characteristic for transition metal systems. Only three complexes show values slightly larger than this rough estimate. The cationic methyl(tetramethylcyclam)nickel complex $[\text{Ni}(\text{TMC})(\text{CH}_3)]^+$ with multiplicity three (**m87**) shows the largest value of 0.21 (see the [Supporting Information](#), Table S6 for the complete list). This value is considered acceptable for single-reference methods if further diagnostics and examinations can rule out a multi-reference case, which is the case for **m87** (see the [Supporting Information](#), Table S6). Additionally, the Frobenius norm of the singles amplitudes (T1 diagnostic) introduced by Lee and Taylor¹²⁵ was analyzed for all molecules of the ROST61 data set. None of them exceeded the threshold of 0.05 proposed by Wilson et al.¹²⁴ in the context of transition metal complexes. These findings also underline the reliability of the previously applied screening protocol. Only four complexes yield a T1 diagnostic larger than 0.03, whereas the cationic dibromo(cyclopentadienyl)vanadium complex $[\text{CpVBr}_2]^+$ (**m30**) has the highest value with 0.037 (see the [Supporting Information](#), Table S6 for all T1 values). Generally, we observed only a weak correlation between increased T1 diagnostic values and larger PNO doubles amplitudes. Both test a mixture of type A and B non-dynamic correlation and rather identify problematic cases than quantifying the multi-reference characteristic. Even though higher order CC contributions may be needed to generate very accurate reaction energies, the filtering of multi-reference cases worked effectively. As expected, the diagnostic values are larger on average compared to that obtained for the closed-shell MOR41 set. The key point is whether more than one multi-reference indicator points to significant non-dynamical correlation contributions.¹¹⁰ Such complexes were not considered in ROST61 to ensure that single-reference CCSD(T) can be applied to generate reliable reference reaction energies.

3.3. Generation of Accurate Reference Values. We applied the open-shell versions of DLPNO-CCSD(T) and DLPNO-CCSD(T1) (as implemented in ORCA 4.2.1, i.e.,

also including the T2T3 coupling which can be of importance for open-shell systems) to generate reference reaction energies for the ROST61 set. A CBS extrapolation employing optimized exponents as proposed by Neese and Valeev⁹⁰ with the def2-TZVPP and def2-QZVPP basis sets was performed. The remaining BSIE was estimated in their study to be 0.42 kcal mol⁻¹. For an analogous CBS extrapolation based on the cc-pVTZ/cc-pVQZ and the aug-cc-pVTZ/aug-cc-pVQZ basis sets, comparable BSIE was observed by Neese and Valeev (0.47 and 0.43 kcal mol⁻¹, respectively). We decided to use Ahlrichs' def2 basis sets because they are computationally more efficient than the standard Dunning basis sets with comparable accuracy, and Dunning basis sets with diffuse functions were not applicable for all transition metal complexes included in the ROST61 set. We tested the importance of diffuse functions exemplarily for reaction **R1** of the ROST61 set (epoxide opening of styrene oxide by a hydridobis-(cyclopentadienyl)titanium complex), which contains slightly larger but still computationally feasible complexes. We compared an aug-cc-pVTZ/aug-cc-pVQZ CBS extrapolation with the respective CBS(def2-TZVPP/def2-QZVPP) results. The two reaction energies differ by only 0.1 kcal mol⁻¹, so that it can be assumed that the influence of diffuse basis functions on the accuracy of the reference values also for the larger complexes in the ROST61 data set has no significant influence, at least not within the specified error margins (*vide infra*).

The inherent local approximations of the DLPNO-based CC methods introduce an additional error ("local error") compared to canonical CCSD(T). To keep this local error as small as possible, we employed very tight threshold settings (*VeryTightPNO*⁸⁸ for def2-TZVPP and *TightPNO*⁸⁹ for def2-QZVPP since the respective *VeryTightPNO* calculations were computationally unfeasible for the larger complexes of the ROST61 set). To avoid possible discontinuities due to the different settings in the CBS extrapolation, the def2-QZVPP correlation energies were corrected by the difference between the respective *VeryTightPNO* and *TightPNO* correlation energies obtained with the smaller def2-SVP basis set (see the [Supporting Information](#), Table S9 for details). As pointed out by Sorathia and Tew,¹²⁶ the very tight threshold values should also exclude "false basis set convergence", at least to a large extent, even if the slightly less tight values for def2-QZVPP calculations cannot completely rule out such behavior.

Another error source of our applied protocol to generate reference values for the ROST61 set is the semilocal (T0)^{84,86} approximation to compute the triples correction in DLPNO-CCSD(T), which is more problematic for open-shell systems.¹¹⁵ However, iterative (T1)⁸⁵ triples were computationally unfeasible, at least for the larger complexes in the def2-QZVPP basis set. Hence, as first suggested by Efremenko and Martin,⁵⁷ the difference between T1 and T0 triples energy was estimated with a smaller basis and added as an additive correction to the extrapolated total single-point energy (SPE). Iron and Janes showed that this difference is only slightly basis set dependent¹²⁷ and that it is a valid approximation to calculate this correction with the smaller def2-SVP basis set (see the [Supporting Information](#), Table S9 for details).

The cumulative local error of open-shell DLPNO-CCSD(T) with tight thresholds was estimated by Neese et al. to be ≈ 0.17 kcal mol⁻¹ for relative energies.⁸³ However, since the reactions in the ROST61 data set contain larger and more difficult complexes, the local error for the ROST61 complexes is very likely larger. To quantify this assumption, we investigated the

Table 1. Collected MADs (in kcal·mol^{−1}) for All Assessed Methods^a

method	class	plain	D3(0)	D3(BJ)	D4	VV10/NL
PBE ^{132,133}	GGA	7.1		4.2	4.0	
revPBE ¹³⁴	GGA	11.2		4.3	4.2	
BP86 ^{135,136}	GGA	8.7		4.6	4.2	
BLYP ^{135,137,138}	GGA	11.7		5.2	5.1	
B97-D3 ⁹⁵	GGA			5.1		
TPSS ¹³⁹	meta-GGA	7.5		3.7	3.7	
revTPSS ¹⁴⁰	meta-GGA	6.4		3.6	3.6	
M06-L ¹⁴¹	meta-GGA	4.2	4.0		3.9	
revM06-L ¹⁴²	meta-GGA	4.4				
SCAN ¹⁴³	meta-GGA	4.5		4.1	4.1	
r ² SCAN ^{144,145}	meta-GGA	4.2		3.3	3.4	3.4
B97M ¹⁴⁶	meta-GGA			3.6		3.4
TPSSH ¹⁴⁷	hybrid	6.6		2.3	2.3	
TPSSH ⁶²	hybrid	6.7		2.6	2.6	
PBE0 ⁶¹	hybrid	6.2		2.7	2.6	
B3LYP ¹⁴⁸	hybrid	10.0		3.2	3.2	3.0
M06 ¹⁴⁹	hybrid	4.4	3.9		3.8	
M06-2X ¹⁴⁹	hybrid	6.4	6.4		6.3	
MN15 ¹⁵⁰	hybrid	4.0				
PW6B95 ¹⁵¹	hybrid	5.1		2.6	2.5	
ωB97X ¹⁵²	RS-hybrid	4.9		3.8	4.2	2.8
ωB97M ¹⁵³	RS-hybrid					2.8
B2-PLYP ¹⁵⁴	DH	5.0		2.3	2.3	
PWPB95 ¹⁵⁵	DH	3.1		1.9	1.6	
B2GP-PLYP ¹⁵⁶	DH	3.7		2.0	2.3	
mPW2PLYP ¹⁵⁷	DH	4.4		2.2	2.3	
PBE0-DH ¹⁵⁸	DH	4.3		3.0	2.5	
revDOD-PBEP86 ^{159–161}	DH	3.7		2.0	2.2	
revDSD-BLYP ^{160,161}	DH	4.2		2.0	2.3	
revDSD-PBEP86 ^{160,161}	DH	3.6		2.0	2.2	
revDSD-PBEPBE ^{159–161}	DH	3.8		2.1	2.3	
B97-3c ⁸¹	composite GGA			4.0		
r ² SCAN-3c ¹⁶²	composite meta-GGA				2.9	
PBEh-3c ¹⁶³	composite hybrid			7.7		
B3LYP/SVP	low-cost hybrid	11.0				
HF-3c ¹⁶⁴	composite WFT			20.4		
MP2	WFT	13.5				
GFN2-xTB ^{79,80}	SQM			15.2		
PM6-D3H4X ¹⁶⁵	SQM		18.6			
PM7 ¹⁶⁶	SQM	23.8				

^aThe mean reference reaction energy is −42.8 kcal·mol^{−1}. Outliers with a deviation of more than 100 kcal·mol^{−1} were excluded from the statistics (HF-3c: 2; GFN2-xTB: 1; PM7: 1; and PM6-D3H4X: 2). Due to the error range of the reference method (±2.5 kcal·mol^{−1}), only differences >0.3 kcal·mol^{−1} should be regarded statistically relevant.

error of the local approximations for one of the smallest ROST61 reactions (**R17**) (association of a cyclopentadienyl anion to a cationic mono(cyclopentadienyl)tantal complex), for which comparative calculations with canonical CCSD(T) were computationally feasible, at least with smaller basis sets. Compared to the corresponding reaction energy of −201.1 kcal mol^{−1} calculated with the latter in the def2-TZVP basis set, the DLPNO-CCSD(T)/*TightPNO* and DLPNO-CCSD(T)/*VeryTightPNO* values calculated in the same basis differ only by +1.2 kcal mol^{−1} (0.6%) and +0.6 kcal mol^{−1} (0.3%), respectively. About half of the remaining local error stems from the T0 approximation as revealed by the comparison of the respective DLPNO-CCSD(T1)/*VeryTightPNO* result (−200.7 kcal mol^{−1}).

Only for nine reactions of the ROST61 set composed of the smallest complexes (see the [Supporting Information](#) for details,

Table S7), DLPNO-CCSD(T1)/*VeryTightPNO*/CBS(def2-TZVPP,def2-QZVPP) reference reaction energies could be generated. To indicate that the majority of the reference values were calculated with the previously mentioned additional approximations for the def2-QZVPP calculations, we use single quotation marks in the following (“DLPNO-CCSD(T1)/CBS”). To estimate the influence of this additional approximation, we analyzed the corresponding reaction energies for the subset composed of the nine reactions for which they were not required (see the [Supporting Information](#), Table S10). The largest deviation of −0.7 kcal mol^{−1} (3.5%) from the protocol without further approximations was found for reaction **R17**, which is still within the specified error margins (*vide infra*). However, the [mean absolute deviation (MAD) = 0.3 kcal mol^{−1}] is significantly smaller and no systematic error [mean deviation (MD) = 0.0 kcal mol^{−1}] was

detected. Therefore, we are confident that the estimation of the semilocal error of the triples energy and the small basis set correction for the slightly less tight DLPNO thresholds employed for the reactions with the larger complexes in the def2-QZVPP basis is sufficiently accurate, at least for the ROST61 set.

Based on these test calculations and the experience from other studies,^{65,128–130} the error in the reaction energies due to the local approximations was estimated to be overall smaller than 1 kcal mol^{−1}. The residual BSIE and BSSE in the CBS-extrapolated reaction energies could not be further quantified but should not exceed 0.5 kcal mol^{−1} on average for the applied protocol. The neglect of core–valence correlation effects by the applied frozen-core approximation is another source of error.¹¹⁸ To estimate the magnitude of this error for the ROST61 set, test calculations with and without the frozen core approximations were performed for the same subset for which DLPNO-CCSD(T1)/VeryTightPNO/def2-QZVPP calculations were computationally feasible, with the result that on average, an additional error of about 1 kcal/mol must be considered for the reference reaction energies (cf. Supporting Information, Table S11). Furthermore, relativistic effects not captured by the use of ECPs for the 4d and 5d transition metal complexes are neglected. However, since the conducted DFA benchmark study (*vide infra*) was also carried out with the def2-QZVPP basis set and corresponding ECPs, this approximate consideration of relativistic effects has no influence on the error statistics of the tested methods. In comparison to other recent studies of transition metal complexes conducted by Shee et al.⁴¹ and Head-Gordon et al.,³⁹ post-CCSD(T) should only play a minor role for the ROST61 set due to the relatively strict multi-reference filtering we applied previously for the selection of the reactions.

Overall, the remaining error of the “DLPNO-CCSD(T1)/CBS” reference reaction energies is estimated to be about ± 2.5 kcal mol^{−1}. Compared to the closed-shell MOR41 benchmark, where slightly smaller error margins of ± 2.0 kcal mol^{−1} for the respective reference values were assumed, even tighter threshold values and a more accurate protocol to estimate the local truncation errors were applied for the ROST61 set. However, since open-shell transition metal complexes are electronically more difficult than the closed-shell species in the MOR41 set, the estimated uncertainty of the ROST61 reference values is still larger. We are nevertheless confident that the calculated reference values are accurate enough to assess DFT- and MP2-based methods in a meaningful way, which will be the topic of the following chapter. Even if the estimated error range of the latter was exceeded for a few reactions, this would not significantly change the error statistics of the tested DFAs due to the considerable number and diversity of reactions included in ROST61. Furthermore, an error for the individual reactions is not equal to the error over the whole set, but one may assume that this uncertainty is, at least to some extent, averaged out. As an estimate for this, one may take the square root of the sum of squares of the errors divided by the number of reactions, which yields $\sqrt{61 \times 2.5^2 / 61} \approx 0.3$ kcal mol^{−1} for the ROST61 set. Accordingly, methods whose averaged errors on the ROST61 benchmark are within 0.3 kcal mol^{−1} should then be considered statistically indistinguishable (equally good).

We are convinced that the practical importance of a diverse and chemically relevant test set for which reliable reference

values could nevertheless be generated is at least as great as for a benchmark with highly accurate reference values but containing only a few smaller and less diverse organometallic reactions.

4. RESULTS AND DISCUSSION OF THE ROST61 BENCHMARK STUDY

In the following, we discuss the performance of various DFAs on the ROST61 set (Section 4.1). Section 4.2 deals with the role of LD corrections. Further, we address the composite and semiempirical methods in Section 4.3, and computation times are briefly discussed in Section 4.4.

4.1. Results for DFAs. In total, 31 representative DFAs in combination with three different LD corrections (D3, D4, and VV10/NL) were assessed with respect to their performance for the ROST61 benchmark set (Table 1). Overall, five GGAs, seven meta-GGAs, eight hybrids, two range-separated (RS) hybrids, and nine DH DFAs were considered. Further, the low-cost composite DFT-D methods B97-3c (GGA), r²SCAN-3c (mGGA), and PBEh-3c (hybrid) are covered as well as the low-cost HF-3c (WFT) method and MP2. B3LYP/def2-SVP was tested additionally as it is supposed to resemble the still frequently applied B3LYP/6-31G* level of theory¹³¹ since this Pople basis set was not available for all complexes in the applied ORCA version. Most of the DFAs were examined with different correction schemes for capturing long-range LD interactions. Besides the D3 correction with BJ or zero (0) damping in conjunction with three-body ATM corrections and the recently published D4 correction, the non-self-consistent VV10 correction was applied for some DFAs.

Figure 3 depicts the MAD for each DFA class with and without application of an LD correction. Generally, the MADs

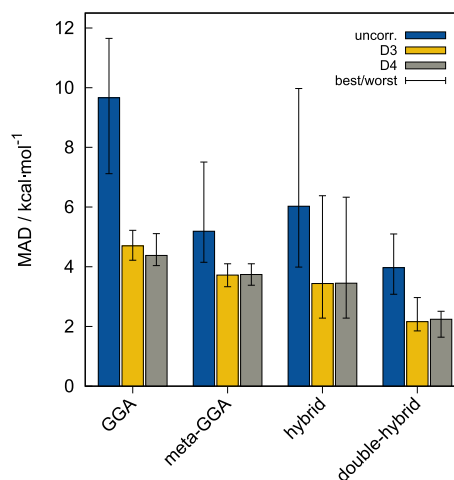


Figure 3. Averaged MAD values for different DFA classes. The error bars denote the best- and the worst-performing DFAs in the respective class. Composite and low-cost methods were excluded for comparability.

are substantially smaller if any of the tested LD correction is included, thus rendering its application essential. Therefore, mainly LD-corrected results will be discussed in the following if not stated otherwise. The overall trend of the MADs for the different functional classes follows Perdew’s “Jacob’s Ladder” scheme¹⁶⁷ (Figures 3 and 4; for further detailed statistics, see the Supporting Information, Tables S12 and S13). This is in line with the conclusion from previous main-group or

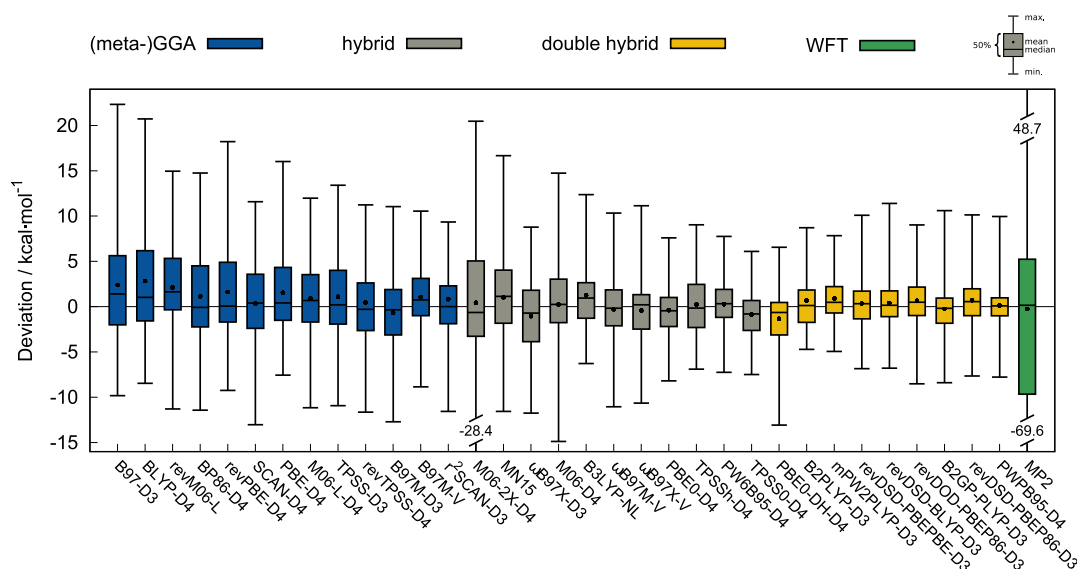


Figure 4. Boxplot of the best performing method combinations for all assessed DFAs and MP2 sorted by descending MAD in the respective class.

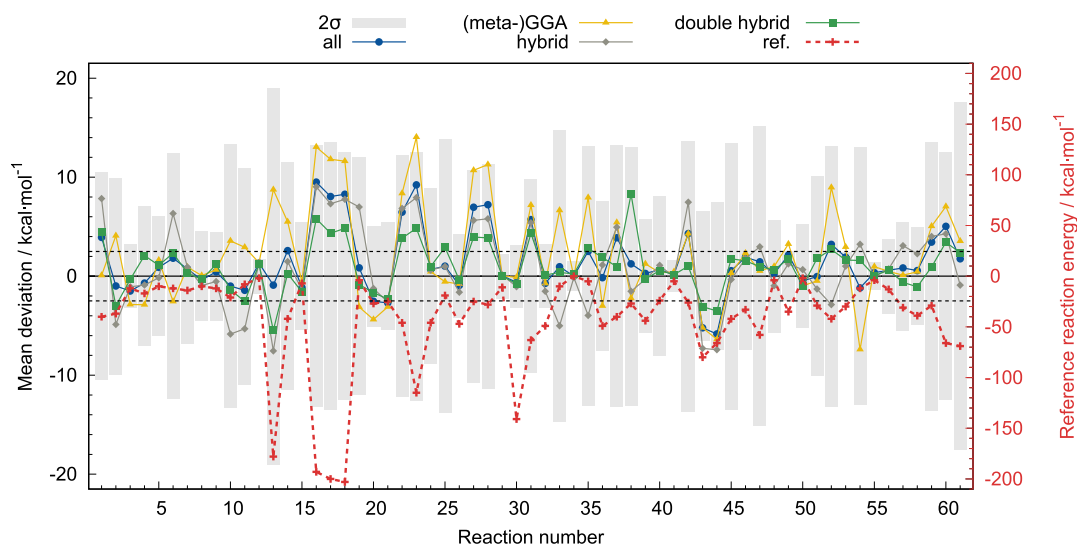


Figure 5. Evaluation of the overall method and method class MDs for reactions 1–61. The light gray bars represent the 2σ range over all assessed methods for each reaction. Dashed lines indicate the ± 2.5 kcal mol $^{-1}$ accuracy window of the reference method. The red dashed line represents the reference reaction energies in kcal mol $^{-1}$.

transition metal complex benchmark sets such as the GMTKN55,¹⁶ IONPI19,¹⁶⁸ or MOR41.⁶⁵

The GGA class yields mean MADs of 4.7 and 4.4 kcal mol $^{-1}$ and mean standard deviations (SDs) of 6.2 and 5.9 kcal mol $^{-1}$ applying the D3 and D4 correction scheme, respectively. Within the GGA class, PBE-D4 performs best with an MAD of 4.0 kcal mol $^{-1}$ and an SD of 5.3 kcal mol $^{-1}$. The worst performing GGA is BLYP-D3 with an MAD of 5.2 kcal mol $^{-1}$ and an SD of 6.8 kcal mol $^{-1}$.

The meta-GGA class shows improved results for all LD correction combinations with a class-averaged MAD and SD of 3.7 and 4.9 kcal mol $^{-1}$, respectively. In the meta-GGA class, the recently introduced r^2 SCAN-D3 performs best with an MAD of 3.3 kcal mol $^{-1}$ (3.4 kcal mol $^{-1}$ for D4) and an SD of 4.5 kcal mol $^{-1}$. B97M-V performs equally well with an MAD of 3.4 kcal mol $^{-1}$ and an SD of 4.6 kcal mol $^{-1}$. Surprisingly, the worst performing among all tested meta-GGA is SCAN-D3 with an MAD of 4.1 and an SD of 5.4 kcal mol $^{-1}$ underlining the

substantial improvement of the r^2 SCAN approach over its predecessor for the calculation of relative energies.¹⁴⁵ Overall, the small systematic error of the GGA DFAs is reflected by an MD of 1.8 kcal mol $^{-1}$. The meta-GGA class yields an even smaller MD of 0.7 kcal mol $^{-1}$. A further improvement is obtained if Fock exchange is included as well. The hybrid class of DFAs yields an MAD average of 3.4 kcal mol $^{-1}$ for D3 and 3.5 kcal mol $^{-1}$ for D4 with SDs of 4.6 kcal mol $^{-1}$. Noteworthy, the best performing hybrid DFA is TPSS0-D4 with an MAD of 2.3 kcal mol $^{-1}$ and an SD of 2.8 kcal mol $^{-1}$. The hybrid class yields an overall MD of 0.1 kcal mol $^{-1}$, indicating no significant remaining systematic error. Among all tested hybrid DFAs, M06-2X performs worst with an MAD of 6.4 kcal mol $^{-1}$ and an SD of 8.9 kcal mol $^{-1}$, probably caused by the large amount of Hartree–Fock (HF) exchange (54%). A comparably poor performance was previously observed for the MOR41⁶⁵ benchmark, while good results were obtained for main group chemistry in other studies.¹⁶ This is supported by the

comparably better performance of M06 (MAD = 4.4 kcal mol⁻¹ and SD = 5.8 kcal mol⁻¹) that only includes 27% of HF exchange while being based on the same parameterization strategy. Overall, the best performing hybrid DFAs all include less than 30% Fock exchange (TPSS0, 25%; TPSSH, 10%; PBE0, 25%; and PW6B95, 28%), indicating a general advantage of applying small admixtures of Fock exchange for the calculation of open-shell transition metal reaction energies. The excellent performance of hybrid DFAs with low amounts of Fock exchange is in accordance with previous benchmark studies on transition metal complex thermochemistry.^{65,127} The overall best results are obtained applying DH DFAs. The DH class yields an average MAD of 2.2 kcal mol⁻¹ and an average SD of 3.0 kcal mol⁻¹ (the respective D4 results perform equally well). PWPB95-D4 performs best in the DH class with a low MAD of only 1.6 kcal mol⁻¹ and an SD of 2.5 kcal mol⁻¹. This result underlines the excellent performance of LD-corrected PWPB95 in previous studies including the MOR41, GMTKN55, and MOBH35^{127,169} benchmark studies as well as the observations by Steinmetz and Grimme.⁵⁶ The worst performing DH is PBE0-DH with an MAD of 3.0 kcal mol⁻¹ and an SD of 2.5 kcal mol⁻¹ in agreement with previous studies by Goerigk et al.¹⁷⁰

Overall, all the assessed DH DFAs perform comparably well, and in particular, PWPB95-D4 is generally recommended to calculate reaction energies involving diverse (single-reference) open-shell organometallic complexes. In the context of DH DFAs, the poor performance of MP2 with an MAD of 13.5 kcal mol⁻¹ and an SD of 20.3 kcal mol⁻¹ is notable. This can be attributed to the underlying UHF orbitals that are clearly inferior to those obtained from the DFT calculation in the DH framework.

Generally, the SDs are well correlated with the MADs of the respective methods with SD to MAD ratios ranging between 1 and 1.5 kcal mol⁻¹. Except for the GGA DFAs and, to a lesser extent, the meta-GGA DFAs, which has a small yet statistically relevant systematic error, the latter is eliminated by hybrid and DH DFAs yielding MDs < 0.5 kcal mol⁻¹.

Figure 5 presents the MD per reaction for all assessed DFAs and the respective DFA classes. The DFA mean allows us to identify unsystematic errors resulting from specific reactions and to analyze DFA class-specific weaknesses. Overall, the DFA MD over all methods reveals a statistically well-balanced test set with no systematic outliers exceeding the 2 σ range. Further, an MAD to SD ratio of 1.3 indicates approximate normal error distribution.¹⁷¹ For most reactions, a behavior according to “Jacob’s ladder” is observed with the best DH DFAs approaching the error margin of the reference values. The most conspicuous deviations were observed for reaction 38, where all LD-corrected DH DFAs except PWPB95 and PBE0-DH are underestimating the reaction energy by ≈ 6 –12 kcal mol⁻¹. MP2 yields one of the largest errors on the entire benchmark set for this reaction with a deviation of 30.5 kcal mol⁻¹, and for the product of this reaction (**m87**), the corresponding diagnostics (see Section 3.2) may also suggest (small) non-dynamic correlation contributions. Hence, the large deviations of the most tested DH DFAs may also be related to this. For reactions **R16**–**R18**, large positive DFA MDs are observed for all DFA classes, with (meta-)GGAs yielding the most pronounced deviations. However, these reactions also have by far largest reaction energies in a range of -193.4 to -203.1 kcal mol⁻¹, and therefore, larger absolute deviations (ADs) are acceptable.

For some reactions, specifically **R1**, **R6**, **R10**, **R11**, and **R42**–**R44**, large MDs were observed for the hybrid class. This class-averaged deviation is dominated by the exceptionally large error of M06-2X for this reaction. Even though some of these reactions are already poorly described with M06, M06-2X introduces even larger errors, most likely caused by the doubled amount of Fock exchange.

In the GGA class, consistently large errors were obtained for reactions 13 and 54 by almost all assessed (meta-)GGAs, which may indicate a pronounced SIE^{12,172} for these reactions.

Overall, the mean DFA MD amounts of 1.4 kcal mol⁻¹ lies in the estimated error range of the reference method, indicating a well-balanced statistical behavior of the ROST61 benchmark with respect to the reference reaction energies.

A direct comparison of the results for the MOR41 (single-reference closed-shell) and the ROST61 (single-reference open-shell) benchmark sets including the methods evaluated in both studies (r^2 SCAN-3c values are taken from ref 162) (Figure 6) reveals an average MAD increase of only 0.13 kcal

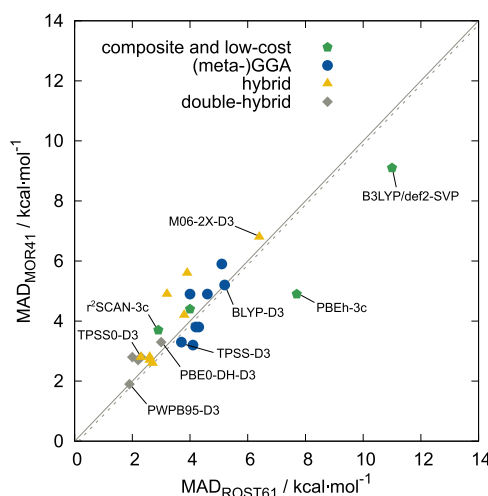


Figure 6. MAD correlation between the MOR41 and the ROST61 benchmark studies. All values in kcal mol⁻¹. The dashed line represents the MD of the presented ROST61 MADs from the MOR41 MADs.

mol⁻¹ for the ROST61 set. As this small shift is mainly caused by the worse performance of PBEh-3c for the ROST61 benchmark, we conclude that the assessed DFAs perform equally well despite the increased electronic complexity due to presence of unpaired electrons in ROST61 compared to MOR41. Further, the larger mean reaction energy of -42.8 kcal mol⁻¹ for the ROST61 set (-30.0 kcal mol⁻¹ in MOR41) has to be taken into account. Hence, we conclude that well-behaved DFAs without inherent spin-symmetry breaking issues can be safely applied to organometallic complexes bearing unpaired electrons, at least after careful exclusion of systems with multi-reference characteristic.

The robust performance of r^2 SCAN-3c, yielding good MADs for both sets, is noteworthy. This robustness can be traced down to the well-balanced application of an attractive D4 and a repulsive geometric counter-poise correction, as well as to the absence of spin-symmetry breaking issues that typically adhere to other DFAs such as B3LYP. Paired with its comparably low computational cost (cf. Figure 10), r^2 SCAN-3c is specifically suited for exploratory investigations, for

example, of unknown reaction mechanisms in the field of metal–organic chemistry.

To further investigate the trends regarding the individual periods, the averaged MAD for all assessed methods is shown for each block in Figure 7. The 3d-block mean MAD amounts

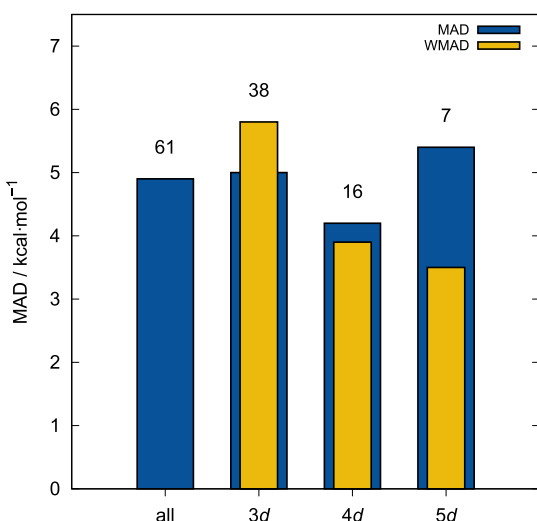


Figure 7. MAD and WMAD values averaged over all tested DFA method combinations (composite, low-cost, and MP2 excluded for comparability). The number of included reactions is depicted above the respective bar. Mean reaction energies for the 3d-, 4d-, and 5d-block are -37.1 , -46.4 , and -65.3 kcal mol $^{-1}$, respectively.

to 5.0 kcal mol $^{-1}$ for a mean reaction energy of -37.1 kcal mol $^{-1}$ (38 reactions) surpassing the mean MAD of 4.2 kcal mol $^{-1}$ for the 4d-block (16 reactions), but being lower than the 5d-block mean MAD (5.4 kcal mol $^{-1}$, seven reactions). The absolute difference between the 3d and 5d mean MADs only amounts to 0.4 kcal mol $^{-1}$. Regarding this small difference, the larger mean reaction energies of -46.4 and -65.3 kcal mol $^{-1}$ for the 4d- and 5d-block, respectively, have to be considered. Weighting the MAD with respect to the overall absolute reaction energy yields weighted MADs (WMADs) of 5.8 kcal mol $^{-1}$ (3d), 3.9 kcal mol $^{-1}$ (4d), and 3.5 kcal mol $^{-1}$ (5d), respectively. Accordingly, a slightly worse performance for the 3d-block is observed.

4.2. Effect of LD Corrections. Overall, almost all DFA results are substantially improved upon inclusion of an LD correction (cf. Figure 3). No significant performance difference regarding the applied dispersion corrections (D3, D4, and VV10) was observed. Addition of the ATM-type three-body term on top of the D3 correction does not improve the results either, even though an increasing influence is expected for larger to very large systems.⁹⁹ Therefore, the application of any of the discussed corrections is strongly recommended in line with previous studies on organometallic complexes.^{99,101} The need for dispersion corrections can further be most clearly illustrated for B3LYP (Figure 8). B3LYP as a rather repulsive DFA includes almost no implicit description of medium range dispersion. Accordingly, the effect of LD corrections is substantial, and application of the latter reduces the MAD by 6.8–7.0 kcal mol $^{-1}$, which corresponds to an improvement of $\approx 70\%$. Of all the assessed DFAs, the Minnesota-type DFAs M06-L, M06, and M06-2X profit least from an LD correction, since these functionals already indirectly capture LD interactions via their parameter fit.

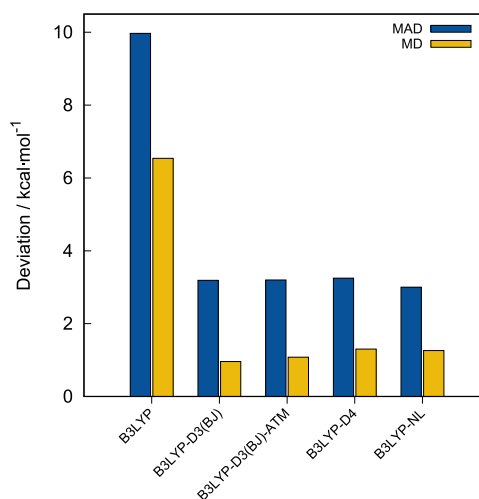


Figure 8. Comparison of LD correction schemes applied for B3LYP/def2-QZVPP on the ROST61 benchmark. All deviations are given in kcal mol $^{-1}$.

4.3. Results for Composite and Semiempirical Methods. The composite DFT methods of the “3c” family represent an approach to reach reasonable accuracy at moderate to low computational cost and were also assessed for the ROST61 benchmark (Figure 9). In general, these

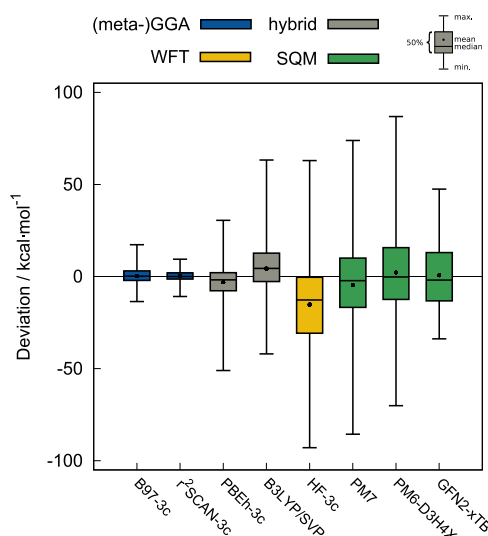


Figure 9. Boxplot of all assessed composite and SQM methods and the B3LYP/def2-SVP method. Outliers with deviations of more than 100 kcal mol $^{-1}$ (HF-3c: 2; GFN2-xTB: 1; PM7: 1; and PM6-D3H4X: 2) were excluded from the statistics.

methods apply well-balanced smaller atomic orbital basis sets and try to reduce systematic errors by application of tailored corrections. The best performing composite method is the recently introduced r²SCAN-3c functional with a small MAD of 2.9 kcal mol $^{-1}$ and an SD of 4.1 kcal mol $^{-1}$. Accordingly, r²SCAN-3c outperforms all assessed (meta-)GGAs applied together with the large def2-QZVPP basis set. r²SCAN-3c reaches hybrid quality at a fraction of computation time and therefore represents a valuable promising method. The GGA-based B97-3c also performs well in its DFA class yielding an MAD of 4.0 kcal mol $^{-1}$ and an SD of 5.6 kcal mol $^{-1}$, being on par with the GGA best performer PBE-D4. The hybrid

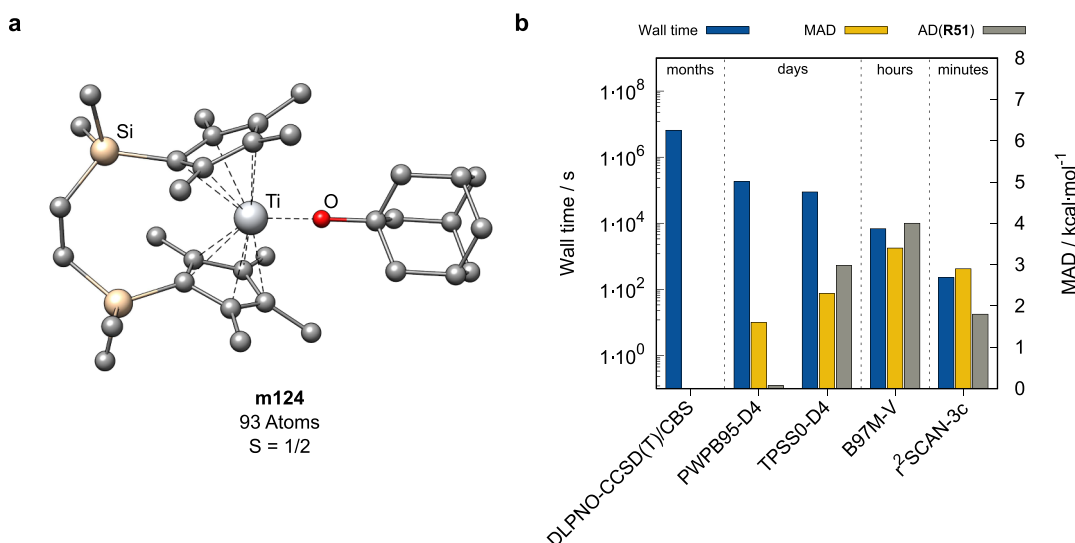


Figure 10. (a) Molecular structure of **m124**, the product of reaction **R51**. (b) Computational wall times of the SPE calculations of selected methods for molecule **m124** in seconds, their overall MADs, and the AD for reaction **R51** in kcal mol⁻¹. Note the logarithmic scale for timings. All DFT calculations were performed in parallel mode on four Intel Xeon CPU E3-1270 v5 @ 3.60 GHz cores.

representative of the “3c” methods (PBEh-3c) performs comparably poor with an MAD of 7.7 kcal mol⁻¹ and an SD of 11.1 kcal mol⁻¹. PBEh-3c suffers in this context from two shortcomings, the small def2-mSVP basis set and the high amount of HF exchange (42%). While the first may partially be compensated by the applied geometrical counter-poise correction, the large HF exchange proves problematic as already observed for M06-2X. Nevertheless, PBEh-3c still clearly outperforms B3LYP/def2-SVP that yields an even larger MAD of 11.0 kcal mol⁻¹. Not surprisingly, among the tested efficient composite methods, HF-3c with a very large MAD of 20.4 kcal mol⁻¹ and an SD of 23.9 kcal mol⁻¹ shows the worst performance due to the poor UHF orbitals, also when compared to all the assessed DFT- and WFT-based methods. Hence, HF-3c should not be used to study open-shell organometallic reactions.

For very large open-shell transition metal compounds (>500 atoms) such as metal–organic polyhedra or frameworks,¹⁷³ even efficient composite DFT methods may not be applicable anymore due to excessive computation times. Therefore, faster but also more approximate methods such as SQM methods need to be applied. However, only a few SQM methods, such as the GFN n -xTB and PM x methods, are available for the general treatment of transition metal complexes. The performance of these SQM methods for the ROST61 benchmark is briefly discussed in the following.

GFN2-xTB performs best with an MAD of 15.2 kcal mol⁻¹ and an SD of 19.3 kcal mol⁻¹. Even though these deviations are high compared to most DFAs, it is noteworthy that the observed MAD is close to that of MP2 and even smaller by 4.5 kcal mol⁻¹ compared to HF-3c. This is specifically remarkable, as results are obtained in a tiny fraction of computation time. The main competitor of the tight-binding based GFN n -xTB methods, which simulated different spin-multiplicities via Fermi smearing,⁸⁰ are the PM x methods, which are derived from HF theory including spin states explicitly. As expected, both PM6-D3H4X and PM7 perform significantly worse than GFN2-xTB with MADs of 18.6 kcal mol⁻¹ and 23.8 kcal mol⁻¹, respectively. They also yield high SDs of 25.3 and 37.5 kcal mol⁻¹. Note that all assessed SQM methods yield at least one

unreasonably large outlier with a deviation of more than 100 kcal mol⁻¹. To avoid a strong bias of error statistics, these outliers (GFN2-xTB: 1; PM7: 1; and PM6-D3H4X: 2) were excluded from the evaluation. Overall, among all tested SQM methods, only GFN2-xTB reaches a performance that may be considered sufficient for at least preliminary investigations of open-shell transition metal reactions.

4.4. Comparison of Computation Times. As the method choice for a given problem always requires careful consideration of the expendable computation time, the speed of a considered method is an important factor. Even though many of the here assessed methods yield accurate results, the computation time needed to obtain them may exceed the available resources for extensive computational studies. As an example, selected computation times of SPE calculations are provided for the titanium complex **m124** with respect to the respective method MADs (Figure 10, note the logarithmic scale). **m124** consists of 93 atoms and the SPE calculation on four parallel CPUs already takes more than 2 days for the PWPB95-D4 DH DFA, while the computation time is almost halved for TPSS0-D4. The B97M-V meta-GGA calculation requires approximately 2 h and the B97-3c composite method SPE calculation only takes 13 min. The drastically reduced computation time is offset by relatively small increases in MAD (1.6, 2.3, 3.4, and 4.0 kcal mol⁻¹ for the mentioned methods). In this context, the exceptionally good performance of the r2SCAN-3c composite method with an MAD of 2.9 kcal mol⁻¹ at comparably small computational cost (about twice as large as for B97-3c) is remarkable and underlines its value as a universally applicable method for a variety of computational chemical tasks. Further timing comparisons for a medium-sized (**m39**, 63 atoms) and a smaller complex (**m113**, 34 atoms) can be found in the Supporting Information, Section S3.4.

5. SUMMARY AND PERSPECTIVE

We introduced the ROST61 set representing the first comprehensive organometallic reaction benchmark compiled solely from larger open-shell single-reference complexes. It features 61 diverse reactions comprising 20 different d-block metals and 150 molecules in total. Due to its focus on

chemically relevant single-reference open-shell compounds, some elements such as iron or platinum are not or less considered in the benchmark set. ROST61 enables an extensive and meaningful evaluation of typically applied quantum chemical methods such as DFAs to model organometallic reaction mechanisms. A multi-step screening protocol, involving spin-state splitting energies, UHF spin contamination values, and FOD analysis was carried out. This protocol was applied to exclude possible multi-reference cases and to ensure that reliable reference values could be obtained employing single-reference CC theory. Accurate local CC reference values with an estimated error of ± 2.5 kcal mol⁻¹ were generated with a mean reaction energy of -42.8 kcal mol⁻¹.

The DFT benchmark study assessing 31 DFAs, four efficient composite methods, MP2, and four very fast SQM methods, confirmed the accuracy classification according to Perdew's Jacob's ladder picture. This finding is in line with previous benchmark studies for comprehensive main-group chemistry such as provided by the GMTKN55 data base or closed-shell organometallic thermochemistry as represented by the MOR41 set. Most DFAs perform similarly well for closed- and open-shell organometallic reactions though some DFAs, such as PBEh-3c or SCAN, yield slightly larger deviations for the open-shell ROST61 benchmark set.

As expected, the tested WFT approaches HF-3c and MP2 are not able to predict reaction energies of open-shell transition metal complexes reasonably accurate. For the latter, this can be attributed to known problems of MP2 for open-shell transition metal complexes such as poor UHF orbitals or small orbital energy differences in the denominator.¹⁷⁴ These issues may be cured, at least to some extent, by orbital-optimized^{175,176} or random-phase-approximation based approaches,^{177,178} but this is out of the scope of the present study.

According to their MADs, the best performing DFAs in each class are PBE-D4 with an MAD of 4.0 kcal mol⁻¹ (GGA), r²SCAN-D3(BJ) (3.3 kcal mol⁻¹, meta-GGA), TPSS0-D4 (2.3 kcal mol⁻¹, hybrid), and PWPB95-D4 (1.6 kcal mol⁻¹, DH). The r²SCAN-3c composite meta-GGA was found to perform best among the assessed composite methods yielding a small MAD of 2.9 kcal mol⁻¹, thus approaching hybrid DFA accuracy. The extended tight-binding SQM method GFN2-xTB turned out to be best performer among the assessed SQM methods with an MAD of 13.5 kcal mol⁻¹.

The application of LD corrections proved to be beneficial in general, with all tested schemes performing comparably well. In the course of this benchmark study, only a selection of DFAs could be assessed, but other promising candidates such as ω B97M(2),¹⁷⁹ xDSDn-PBEP86-D4,¹⁸⁰ DSD-PBEdRPA75-D4,¹⁸¹ or B-LYP-osUW12¹⁸² should also be evaluated in the future.

The current overall best performer PWPB95-D4, which hits on average the error range of the local CC reference values, can be clearly recommended to model open-shell organometallic reactions composed of larger single-reference transition metal complexes.

With the ROST61 benchmark, the first solely open-shell benchmark set with chemically relevant organometallic complexes was compiled. It facilitates the decision of (computational) chemists to reasonably choose the appropriate method to elucidate challenging organometallic reaction mechanisms. Moreover, the ROST61 set represents a valuable

resource for development of new or improved semiempirical, DFT, and approximate WFT methods.

Though also relevant in practice, transition metal complexes with a significant amount of non-dynamic correlation are not covered by the presented benchmark due to the difficulty in generating reliable reference values. This is part of the current research in our laboratory.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00659>.

Computational details; statistical measures; details on accuracy and multi-reference characteristic analyses; list of involved molecules and reactions; tabulated absolute and relative reference energies; and tabulated relative energies of assessed methods (PDF)

Optimized molecular structures in XYZ format (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Andreas Hansen – Mulliken Center for Theoretical Chemistry, Institute for Physical and Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0003-1659-8206; Email: hansen@thch.uni-bonn.de

Authors

Leonard R. Maurer – Mulliken Center for Theoretical Chemistry, Institute for Physical and Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0003-2433-1130

Markus Bursch – Mulliken Center for Theoretical Chemistry, Institute for Physical and Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0001-6711-5804

Stefan Grimme – Mulliken Center for Theoretical Chemistry, Institute for Physical and Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0002-5844-4371

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00659>

Author Contributions

[†]L.R.M. and M.B. contributed equally to this work and should be considered joint first author.

Notes

The authors declare no competing financial interest.

A table with the relative energies of all assessed methods and the molecular structures are also available here: <https://www.chemie.uni-bonn.de/pctc/mulliken-center/software/rost61>.

■ ACKNOWLEDGMENTS

The NSF and the DFG are gratefully acknowledged for their financial support in the framework of the joint NSF-DFG grant "Synergistic Experimental and Computational Approaches to Designing Electrocatalysts with Proton-Responsive Ligand Architectures" (Award ID 2055097). We thank G. Santra for technical help with the DSD/DOD DH calculations and S. Ehlert for performing r²SCAN calculations. Further proof-reading by S. Spicher and Dr. J.-M. Mewes is greatly acknowledged.

REFERENCES

- (1) Elschenbroich, C. *Organometallchemie*, Teubner Studienbücher Chemie; Vieweg + Teubner Verlag, 2008; Vol. 6.
- (2) Ho, C.-L.; Wong, W.-Y. Metal-Containing Polymers: Facile Tuning of Photophysical Traits and Emerging Applications in Organic Electronics and Photonics. *Coord. Chem. Rev.* **2011**, *255*, 2469–2502.
- (3) Warra, A. A. Transition Metal Complexes and their Application in Drugs and Cosmetics - A Review. *J. Chem. Pharm. Res.* **2011**, *3*, 951–958.
- (4) Beaumier, E. P.; Pearce, A. J.; See, X. Y.; Tonks, I. A. Modern Applications of Low-Valent Early Transition Metals in Synthesis and Catalysis. *Nat. Rev. Chem.* **2019**, *3*, 15–34.
- (5) Gandeepan, P.; Müller, T.; Zell, D.; Cera, G.; Warratz, S.; Ackermann, L. 3d Transition Metals for C-H Activation. *Chem. Rev.* **2019**, *119*, 2192–2452.
- (6) Muthusamy, S.; Raman, N. Pharmacological Activity of a Few Transition Metal Complexes: A Short Review. *J. Chem. Biol. Ther.* **2016**, *01*, 1–17.
- (7) Malinowski, J.; Zych, D.; Jacewicz, D.; Gawdzik, B.; Drzeżdżon, J. Application of Coordination Compounds with Transition Metal Ions in the Chemical Industry—A Review. *Int. J. Mol. Sci.* **2020**, *21*, 5443.
- (8) Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2018**, *119*, 2453–2523.
- (9) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2011**, *112*, 289–320.
- (10) Grimme, S.; Schreiner, P. R. Computational Chemistry: The Fate of Current Methods and Future Challenges. *Angew. Chem., Int. Ed.* **2018**, *57*, 4170–4176.
- (11) Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Chem. Phys.* **2014**, *140*, 18A301.
- (12) Bao, J. L.; Gagliardi, L.; Truhlar, D. G. Self-Interaction Error in Density Functional Theory: An Appraisal. *J. Phys. Chem. Lett.* **2018**, *9*, 2353–2358.
- (13) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Localization and Delocalization Errors in Density Functional Theory and Implications for Band-Gap Prediction. *Phys. Rev. Lett.* **2008**, *100*, 146401.
- (14) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.
- (15) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: an Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.
- (16) 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.
- (17) Peverati, R.; Truhlar, D. G. Communication: A Global Hybrid Generalized Gradient Approximation to the Exchange-Correlation Functional that Satisfies the Second-Order Density-Gradient Constraint and has Broad Applicability in Chemistry. *J. Chem. Phys.* **2011**, *135*, 191102.
- (18) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Assessment of Gaussian-3 and Density-Functional Theories on the G3/05 Test Set of Experimental Energies. *J. Chem. Phys.* **2005**, *123*, 124107.
- (19) Feller, D.; Peterson, K. A.; Dixon, D. A. A Survey of Factors Contributing to Accurate Theoretical Predictions of Atomization Energies and Molecular Structures. *J. Chem. Phys.* **2008**, *129*, 204105.
- (20) Cioslowski, J.; Schimeczek, M.; Liu, G.; Stoyanov, V. A Set of Standard Enthalpies of Formation for Benchmarking, Calibration, and Parametrization of Electronic Structure Methods. *J. Chem. Phys.* **2000**, *113*, 9377–9389.
- (21) Lynch, B. J.; Truhlar, D. G. Small Representative Benchmarks for Thermochemical Calculations. *J. Phys. Chem. A* **2003**, *107*, 8996–8999.
- (22) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. Databases for Transition Element Bonding: Metal-Metal Bond Energies and Bond Lengths and Their Use To Test Hybrid, Hybrid Meta, and Meta Density Functionals and Generalized Gradient Approximations. *J. Phys. Chem. A* **2005**, *109*, 4388–4403.
- (23) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. Density Functionals for Inorganometallic and Organometallic Chemistry. *J. Phys. Chem. A* **2005**, *109*, 11127–11143.
- (24) Peverati, R.; Truhlar, D. G. M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics. *J. Phys. Chem. Lett.* **2011**, *3*, 117–124.
- (25) Peverati, R.; Truhlar, D. G. An Improved and Broadly Accurate Local Approximation to the Exchange–Correlation Density Functional: The MN12-L Functional for Electronic Structure Calculations in Chemistry and Physics. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13171.
- (26) Amin, E. A.; Truhlar, D. G. Zn Coordination Chemistry: Development of Benchmark Suites for Geometries, Dipole Moments, and Bond Dissociation Energies and Their Use to Test and Validate Density Functionals and Molecular Orbital Theory. *J. Chem. Theory Comput.* **2008**, *4*, 75–85.
- (27) Zhao, Y.; Truhlar, D. G. Benchmark Energetic Data in a Model System for Grubbs II Metathesis Catalysis and Their Use for the Development, Assessment, and Validation of Electronic Structure Methods. *J. Chem. Theory Comput.* **2009**, *5*, 324–333.
- (28) Li, S.; Dixon, D. A. Benchmark Calculations on the Electron Detachment Energies of MO_3^- and M_2O_6^- ($\text{M} = \text{Cr}, \text{Mo}, \text{W}$). *J. Phys. Chem. A* **2007**, *111*, 11908–11921.
- (29) Li, S.; Dixon, D. A. Molecular Structures and Energetics of the $(\text{TiO}_2)_n$ ($n = 1-4$) Clusters and Their Anions. *J. Phys. Chem. A* **2008**, *112*, 6646–6666.
- (30) Zhai, H.-J.; Li, S.; Dixon, D. A.; Wang, L.-S. Probing the Electronic and Structural Properties of Chromium Oxide Clusters $(\text{CrO}_3)_n^-$ and $(\text{CrO}_3)_n$ ($n = 1-5$): Photoelectron Spectroscopy and Density Functional Calculations. *J. Am. Chem. Soc.* **2008**, *130*, 5167–5177.
- (31) Li, S.; Hennigan, J. M.; Dixon, D. A.; Peterson, K. A. Accurate Thermochemistry for Transition Metal Oxide Clusters. *J. Phys. Chem. A* **2009**, *113*, 7861–7877.
- (32) Nielsen, I. M. B.; Allendorf, M. D. High-Level ab Initio Thermochemical Data for Halides of Chromium, Manganese, and Iron. *J. Phys. Chem. A* **2005**, *109*, 928–933.
- (33) Nielsen, I. M. B.; Allendorf, M. D. Thermochemistry of the Chromium Hydroxides $\text{Cr}(\text{OH})_n$, $n = 2-6$, and the Oxyhydroxide $\text{CrO}(\text{OH})_4$: Ab Initio Predictions. *J. Phys. Chem. A* **2006**, *110*, 4093–4099.
- (34) Opila, E. J.; Myers, D. L.; Jacobson, N. S.; Nielsen, I. M. B.; Johnson, D. F.; Olminsky, J. K.; Allendorf, M. D. Theoretical and Experimental Investigation of the Thermochemistry of $\text{CrO}_2(\text{OH})_2(\text{g})$. *J. Phys. Chem. A* **2007**, *111*, 1971–1980.
- (35) Xu, X.; Zhang, W.; Tang, M.; Truhlar, D. G. Do Practical Standard Coupled Cluster Calculations Agree Better than Kohn–Sham Calculations with Currently Available Functionals When Compared to the Best Available Experimental Data for Dissociation Energies of Bonds to 3d Transition Metals? *J. Chem. Theory Comput.* **2015**, *11*, 2036–2052.
- (36) Cheng, L.; Gauss, J.; Ruscic, B.; Armentrout, P. B.; Stanton, J. F. Bond Dissociation Energies for Diatomic Molecules Containing 3d Transition Metals: Benchmark Scalar-Relativistic Coupled-Cluster Calculations for 20 Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 1044–1056.
- (37) Fang, Z.; Vasiliu, M.; Peterson, K. A.; Dixon, D. A. Prediction of Bond Dissociation Energies/Heats of Formation for Diatomic Transition Metal Compounds: CCSD(T) Works. *J. Chem. Theory Comput.* **2017**, *13*, 1057–1066.
- (38) Aoto, Y. A.; de Lima Batista, A. P.; Köhn, A.; de Oliveira-Filho, A. G. S. How To Arrive at Accurate Benchmark Values for Transition Metal Compounds: Computation or Experiment? *J. Chem. Theory Comput.* **2017**, *13*, 5291–5316.

- (39) Shee, J.; Loipersberger, M.; Hait, D.; Lee, J.; Head-Gordon, M. Revealing the Nature of Electron Correlation in Transition Metal Complexes with Symmetry Breaking and Chemical Intuition. *J. Chem. Phys.* **2021**, *154*, 194109.
- (40) Shee, J.; Rudsteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. On Achieving High Accuracy in Quantum Chemical Calculations of 3d Transition Metal-Containing Systems: A Comparison of Auxiliary-Field Quantum Monte Carlo with Coupled Cluster, Density Functional Theory, and Experiment for Diatomic Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 2346–2358.
- (41) Rudsteyn, B.; Coskun, D.; Weber, J. L.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A.; Shee, J. Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2020**, *16*, 3041–3054.
- (42) Cundari, T. R.; Arturo Ruiz Leza, H.; Grimes, T.; Steyl, G.; Waters, A.; Wilson, A. K. Calculation of the Enthalpies of Formation for Transition Metal Complexes. *Chem. Phys. Lett.* **2005**, *401*, 58–61.
- (43) DeYonker, N. J.; Williams, T. G.; Imel, A. E.; Cundari, T. R.; Wilson, A. K. Accurate Thermochemistry for Transition Metal Complexes From First-Principles Calculations. *J. Chem. Phys.* **2009**, *131*, 024106.
- (44) Tekarli, S. M.; Drummond, M. L.; Williams, T. G.; Cundari, T. R.; Wilson, A. K. Performance of Density Functional Theory for 3d Transition Metal-Containing Complexes: Utilization of the Correlation Consistent Basis Sets. *J. Phys. Chem. A* **2009**, *113*, 8607–8614.
- (45) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. Toward Accurate Theoretical Thermochemistry of First Row Transition Metal Complexes. *J. Phys. Chem. A* **2011**, *116*, 870–885.
- (46) Jiang, W.; Laury, M. L.; Powell, M.; Wilson, A. K. Comparative Study of Single and Double Hybrid Density Functionals for the Prediction of 3d Transition Metal Thermochemistry. *J. Chem. Theory Comput.* **2012**, *8*, 4102–4111.
- (47) Zhang, W.; Truhlar, D. G.; Tang, M. Tests of Exchange-Correlation Functional Approximations Against Reliable Experimental Data for Average Bond Energies of 3d Transition Metal Compounds. *J. Chem. Theory Comput.* **2013**, *9*, 3965–3977.
- (48) Mayhall, N. J.; Raghavachari, K.; Redfern, P. C.; Curtiss, L. A. Investigation of Gaussian4 Theory for Transition Metal Thermochemistry. *J. Phys. Chem. A* **2009**, *113*, 5170–5175.
- (49) Riley, K. E.; Merz, K. M. Assessment of Density Functional Theory Methods for the Computation of Heats of Formation and Ionization Potentials of Systems Containing Third Row Transition Metals. *J. Phys. Chem. A* **2007**, *111*, 6044–6053.
- (50) Jonas, V.; Thiel, W. Theoretical Study of the Vibrational Spectra of the Transition Metal Carbonyls $M(\text{CO})_6$ [$M=\text{Cr}, \text{Mo}, \text{W}$], $M(\text{CO})_5$ [$M=\text{Fe}, \text{Ru}, \text{Os}$], and $M(\text{CO})_4$ [$M=\text{Ni}, \text{Pd}, \text{Pt}$]. *J. Chem. Phys.* **1995**, *102*, 8474–8484.
- (51) Li, J.; Schreckenbach, G.; Ziegler, T. A Reassessment of the First Metal-Carbonyl Dissociation Energy in $M(\text{CO})_4$ ($M = \text{Ni}, \text{Pd}, \text{Pt}$), $M(\text{CO})_5$ ($M = \text{Fe}, \text{Ru}, \text{Os}$), and $M(\text{CO})_6$ ($M = \text{Cr}, \text{Mo}, \text{W}$) by a Quasirelativistic Density Functional Method. *J. Am. Chem. Soc.* **1995**, *117*, 486–494.
- (52) Jonas, V.; Thiel, W. Theoretical Study of the Vibrational Spectra of the Transition-Metal Carbonyl Hydrides $\text{HM}(\text{CO})_5$ ($M=\text{Mn}, \text{Re}$), $\text{H}_2\text{M}(\text{CO})_4$ ($M=\text{Fe}, \text{Ru}, \text{Os}$), and $\text{HM}(\text{CO})_4$ ($M=\text{Co}, \text{Rh}, \text{Ir}$). *J. Chem. Phys.* **1996**, *105*, 3636–3648.
- (53) Jensen, K. P.; Roos, B. O.; Ryde, U. Performance of Density Functionals for First Row Transition Metal Systems. *J. Chem. Phys.* **2007**, *126*, 014103.
- (54) Hyla-Kryspin, I.; Grimme, S. Comprehensive Study of the Thermochemistry of First-Row Transition Metal Compounds by Spin Component Scaled MP2 and MP3 Methods. *Organometallics* **2004**, *23*, 5581–5592.
- (55) Quintal, M. M.; Karton, A.; Iron, M. A.; Boese, A. D.; Martin, J. M. L. Benchmark Study of DFT Functionals for Late-Transition-Metal Reactions. *J. Phys. Chem. A* **2006**, *110*, 709–716.
- (56) Steinmetz, M.; Grimme, S. Benchmark Study of the Performance of Density Functional Theory for Bond Activations with (Ni,Pd)-Based Transition-Metal Catalysts. *ChemistryOpen* **2013**, *2*, 115–124.
- (57) Efremenko, I.; Martin, J. M. L. Coupled Cluster Benchmark of New Density Functionals and Domain Pair Natural Orbital Methods: Mechanisms of Hydroarylation and Oxidative Coupling Catalyzed by Ru(II) Chloride Carbonyls. *AIP Conf. Proc.* **2019**, *2186*, 030005.
- (58) Hait, D.; Tubman, N. M.; Levine, D. S.; Whaley, K. B.; Head-Gordon, M. What Levels of Coupled Cluster Theory Are Appropriate for Transition Metal Systems? A Study Using Near-Exact Quantum Chemical Values for 3d Transition Metal Binary Compounds. *J. Chem. Theory Comput.* **2019**, *15*, 5370–5385.
- (59) Hughes, T. F.; Friesner, R. A. Correcting Systematic Errors in DFT Spin-Splitting Energetics for Transition Metal Complexes. *J. Chem. Theory Comput.* **2011**, *7*, 19–32.
- (60) Weymuth, T.; Couzijn, E. P. A.; Chen, P.; Reiher, M. New Benchmark Set of Transition-Metal Coordination Reactions for the Assessment of Density Functionals. *J. Chem. Theory Comput.* **2014**, *10*, 3092–3103.
- (61) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (62) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137 für tpss und tpssh.
- (63) Husch, T.; Freitag, L.; Reiher, M. Calculation of Ligand Dissociation Energies in Large Transition-Metal Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2456–2468.
- (64) Minenkov, Y.; Chermak, E.; Cavallo, L. Accuracy of DLPNO-CCSD(T) Method for Noncovalent Bond Dissociation Enthalpies from Coinage Metal Cation Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 4664–4676.
- (65) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Chęcinski, M. P. Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions. *J. Chem. Theory Comput.* **2018**, *14*, 2596–2608.
- (66) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theory Comput.* **2019**, *15*, 3610–3622.
- (67) Chan, B. The CUAGAU Set of Coupled-Cluster Reference Data for Small Copper, Silver, and Gold Compounds and Assessment of DFT Methods. *J. Phys. Chem. A* **2019**, *123*, 5781–5788.
- (68) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *8*, No. e1327.
- (69) Neese, F. ORCA—An Ab Initio, Density Functional and Semiempirical Program Package, V. 4.0 (Rev. 11599); Max-Planck-Institut für Kohlenforschung: Mülheim an der Ruhr, Germany, 2018.
- (70) Neese, F. ORCA—An Ab Initio, Density Functional and Semiempirical Program Package, V. 4.1 (Rev. 12588); Max-Planck-Institut für Kohlenforschung: Mülheim an der Ruhr, Germany, 2018.
- (71) Neese, F. ORCA—An Ab Initio, Density Functional and Semiempirical Program Package, V. 4.2.1; Max-Planck-Institut für Kohlenforschung: Mülheim an der Ruhr, Germany: Mülheim an der Ruhr, Germany, 2019.
- (72) TURBOMOLE, V7.5.1; University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 2020, 1989–2007, TURBOMOLE GmbH, since 2007; available from: <http://www.turbomole.com>.
- (73) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 91–100.
- (74) MOPAC2016, Stewart Computational Chemistry, 2016, available from <http://OpenMOPAC.net>.
- (75) Semiempirical Extended Tight-Binding Program Package xtb, Version 6.4.3., 2021, <https://github.com/grimme-lab/xtb>.
- (76) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent developments in LIBXC — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.

- (77) 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.
- (78) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.
- (79) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- (80) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, No. e1493.
- (81) Brandenburg, J. G.; Bannwarth, C.; Hansen, A.; Grimme, S. B97-3c: A Revised Low-Cost Variant of the B97-D Density Functional Method. *J. Chem. Phys.* **2018**, *148*, 064104.
- (82) Hansen, A.; Liakos, D. G.; Neese, F. Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals. *J. Chem. Phys.* **2011**, *135*, 214102.
- (83) Saitow, M.; Becker, U.; Riplinger, C.; Valeev, E. F.; Neese, F. A New Near-Linear Scaling, Efficient and Accurate, Open-Shell Domain-Based Local Pair Natural Orbital Coupled Cluster Singles and Doubles Theory. *J. Chem. Phys.* **2017**, *146*, 164105.
- (84) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.
- (85) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An Improved Linear Scaling Perturbative Triples Correction for the Domain Based Local Pair-Natural Orbital Based Singles and Doubles Coupled Cluster Method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.
- (86) Guo, Y.; Riplinger, C.; Liakos, D. G.; Becker, U.; Saitow, M.; Neese, F. Linear Scaling Perturbative Triples Correction Approximations for Open-Shell Domain-Based Local Pair Natural Orbital Coupled Cluster Singles and Doubles Theory [DLPNO-CCSD(T0/T)]. *J. Chem. Phys.* **2020**, *152*, 024116.
- (87) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor. Chem. Acc.* **2007**, *117*, 587–597.
- (88) Pavošević, F.; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps—A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. *J. Chem. Phys.* **2017**, *146*, 174108.
- (89) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.
- (90) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated ab initio Methods? *J. Chem. Theory Comput.* **2011**, *7*, 33–43.
- (91) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
- (92) Weigend, F.; Kattannek, M.; Ahlrichs, R. Approximated Electron Repulsion Integrals: Cholesky Decomposition Versus Resolution of the Identity Methods. *J. Chem. Phys.* **2009**, *130*, 164106.
- (93) Weigend, F. Hartree–Fock Exchange Fitting Basis Sets for H to Rn. *J. Comput. Chem.* **2007**, *29*, 167–175.
- (94) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (95) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (96) Axilrod, B. M.; Teller, E. Interaction of the van der Waals Type Between Three Atoms. *J. Chem. Phys.* **1943**, *11*, 299–300.
- (97) Muto, Y. Force Between Nonpolar Molecules. *Proc. Phys.-Math. Soc. Jpn.* **1943**, *17*, 629.
- (98) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147*, 034112.
- (99) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.
- (100) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.
- (101) Bursch, M.; Caldeweyher, E.; Hansen, A.; Neugebauer, H.; Ehlert, S.; Grimme, S. Understanding and Quantifying London Dispersion Effects in Organometallic Complexes. *Acc. Chem. Res.* **2019**, *52*, 258–266.
- (102) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals Density Functional: The Simpler the Better. *J. Chem. Phys.* **2010**, *133*, 244103.
- (103) Noodleman, L. Valence Bond Description of Antiferromagnetic Coupling in Transition Metal Dimers. *J. Chem. Phys.* **1981**, *74*, 5737–5743.
- (104) Li, X.; Paldus, J. Reduced Multireference Couple Cluster Method. II. Application to Potential Energy Surfaces of HF, F₂, and H₂O. *J. Chem. Phys.* **1998**, *108*, 637–648.
- (105) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669–3680.
- (106) Sharma, P.; Bao, J. J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *Annu. Rev. Phys. Chem.* **2021**, *72*, 541–564.
- (107) Bursch, M.; Neugebauer, H.; Grimme, S. Structure Optimisation of Large Transition-Metal Complexes with Extended Tight-Binding Methods. *Angew. Chem., Int. Ed.* **2019**, *131*, 11195–11204.
- (108) Buijink, J. K. F.; Kloetstra, K. R.; Meetsma, A.; Teuben, J. H.; Smeets, W. J. J.; Spek, A. L. Structure and Reactivity of Mono(cyclopentadienyl)vanadium Alkynyl and Aryne Complexes. *Organometallics* **1996**, *15*, 2523–2533.
- (109) Hollett, J. W.; Gill, P. M. W. The Two Faces of Static Correlation. *J. Chem. Phys.* **2011**, *134*, 114111.
- (110) Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. A Simple DFT-Based Diagnostic for Nondynamical Correlation. *Theor. Chem. Acc.* **2013**, *132*, 1291.
- (111) Grimme, S.; Hansen, A. A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angew. Chem., Int. Ed.* **2015**, *54*, 12308–12313.
- (112) Bauer, C. A.; Hansen, A.; Grimme, S. The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure. *Chem. Eur. J.* **2017**, *23*, 6150–6164.
- (113) Pinter, B.; Chankisijjev, A.; Geerlings, P.; Harvey, J. N.; De Proft, F. Conceptual Insights into DFT Spin-State Energetics of Octahedral Transition-Metal Complexes through a Density Difference Analysis. *Chem. Eur. J.* **2018**, *24*, 5281–5292.
- (114) Roemelt, M.; Pantazis, D. A. Multireference Approaches to Spin-State Energetics of Transition Metal Complexes Utilizing the Density Matrix Renormalization Group. *Adv. Theory Simul.* **2019**, *2*, 1800201.
- (115) Flöser, B. M.; Guo, Y.; Riplinger, C.; Tuzek, F.; Neese, F. Detailed Pair Natural Orbital-Based Coupled Cluster Studies of Spin Crossover Energetics. *J. Chem. Theory Comput.* **2020**, *16*, 2224–2235.
- (116) Qiu, Y.; Henderson, T. M.; Zhao, J.; Scuseria, G. E. Projected Coupled Cluster Theory. *J. Chem. Phys.* **2017**, *147*, 64111.

- (117) Feller, D.; Peterson, K. A.; Crawford, T. D. Sources of Error in Electronic Structure Calculations on Small Chemical Systems. *J. Chem. Phys.* **2006**, *124*, 054107.
- (118) Li Manni, G.; Kats, D.; Tew, D. P.; Alavi, A. Role of Valence and Semicore Electron Correlation on Spin Gaps in Fe(II)-Porphyrins. *J. Chem. Theory Comput.* **2019**, *15*, 1492–1497.
- (119) Giner, E.; Tew, D. P.; Garniron, Y.; Alavi, A. Interplay between Electronic Correlation and Metal–Ligand Delocalization in the Spectroscopy of Transition Metal Compounds: Case Study on a Series of Planar Cu²⁺ Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 6240–6252.
- (120) Tew, D. P. Explicitly Correlated Coupled-Cluster Theory with Brueckner Orbitals. *J. Chem. Phys.* **2016**, *145*, 074103.
- (121) Sonnenberg, J. L.; Schlegel, H. B.; Hratchian, H. P. *Encyclopedia of Inorganic Chemistry*; American Cancer Society, 2009; Chapter 5, p 6.
- (122) Neese, F. Importance of Direct Spin-Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study. *J. Am. Chem. Soc.* **2006**, *128*, 10213–10222.
- (123) Nielsen, I. M. B.; Janssen, C. L. Double-Substitution-Based Diagnostics for Coupled-Cluster and Møller-Plesset Perturbation Theory. *Chem. Phys. Lett.* **1999**, *310*, 568–576.
- (124) Jiang, W.; DeYonker, N. J.; Wilson, A. K. Multireference Character for 3d Transition-Metal-Containing Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 460–468.
- (125) Lee, T. J.; Taylor, P. R. A Diagnostic for Determining the Quality of Single-Reference Electron Correlation Methods. *Int. J. Quantum Chem.* **2009**, *36*, 199–207.
- (126) Sorathia, K.; Tew, D. P. Basis set extrapolation in pair natural orbital theories. *J. Chem. Phys.* **2020**, *153*, 174112.
- (127) Iron, M. A.; Janes, T. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *J. Phys. Chem. A* **2019**, *123*, 3761–3781.
- (128) Kumar, A.; Neese, F.; Valeev, E. F. Explicitly correlated coupled cluster method for accurate treatment of open-shell molecules with hundreds of atoms. *J. Chem. Phys.* **2020**, *153*, 094105.
- (129) Altun, A.; Neese, F.; Bistoni, G. Extrapolation to the Limit of a Complete Pair Natural Orbital Space in Local Coupled-Cluster Calculations. *J. Chem. Theory Comput.* **2020**, *16*, 6142–6149.
- (130) Liakos, D. G.; Guo, Y.; Neese, F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *J. Phys. Chem. A* **2020**, *124*, 90–100.
- (131) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724.
- (132) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (133) Perdew, J. P.; Burke, K.; Ernzerhof, M. Errata: Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1997**, *78*, 1396.
- (134) Zhang, Y.; Yang, W. Comment on “Generalized Gradient Approximation Made Simple”. *Phys. Rev. Lett.* **1998**, *80*, 890.
- (135) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098.
- (136) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822.
- (137) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.
- (138) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained With the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157*, 200.
- (139) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (140) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry. *Phys. Rev. Lett.* **2009**, *103*, 26403.
- (141) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.
- (142) Wang, Y.; Jin, X.; Yu, H. S.; Truhlar, D. G.; He, X. Revised M06-L functional for Improved Accuracy on Chemical Reaction Barrier Heights, Noncovalent Interactions, and Solid-State Physics. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 8487–8492.
- (143) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (144) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r²SCAN meta-Generalized Gradient Approximation. *J. Phys. Chem. Lett.* **2020**, *11*, 8208–8215.
- (145) Ehlert, S.; Huniar, U.; Ning, J.; Furness, J. W.; Sun, J.; Kaplan, A. D.; Perdew, J. P.; Brandenburg, J. G. r²SCAN-D4: Dispersion Corrected meta-Generalized Gradient Approximation for General Chemical Applications. *J. Chem. Phys.* **2021**, *154*, 061101.
- (146) Mardirossian, N.; Head-Gordon, M. Mapping the Genome of meta-Generalized Gradient Approximation Density Functionals: The Search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.
- (147) Grimme, S. Accurate Calculation of the Heats of Formation for Large Main Group Compounds with Spin-Component Scaled MP2 Methods. *J. Phys. Chem. A* **2005**, *109*, 3067–3077.
- (148) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- (149) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2007**, *120*, 215–241.
- (150) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn-Sham Global-Hybrid Exchange-Correlation Density Functional With Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7*, 5032–5051.
- (151) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A* **2005**, *109*, 5656.
- (152) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106.
- (153) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinationally optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (154) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.
- (155) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals—Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
- (156) Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.

- (157) Schwabe, T.; Grimme, S. Towards Chemical Accuracy for the Thermodynamics of Large Molecules: New Hybrid Density Functionals Including Non-Local Correlation Effects. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398.
- (158) Brémond, E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-DH Model. *J. Chem. Phys.* **2011**, *135*, 024106.
- (159) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: In Search of the Best Double-Hybrid DFT with Spin-Component Scaled MP2 and Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104.
- (160) Kozuch, S.; Martin, J. M. Spin-Component-Scaled Double Hybrids: An Extensive Search for the Best Fifth-Rung Functionals Blending DFT and Perturbation Theory. *J. Comput. Chem.* **2013**, *34*, 2327–44.
- (161) Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: RevDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129–5143.
- (162) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M. r²SCAN-3c: A “Swiss Army Knife” Composite Electronic-Structure Method. *J. Chem. Phys.* **2021**, *154*, 064103.
- (163) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent Structures and Interactions by Density Functional Theory with Small Atomic Orbital Basis Sets. *J. Chem. Phys.* **2015**, *143*, 054107.
- (164) Sure, R.; Grimme, S. Corrected Small Basis Set Hartree-Fock Method for Large Systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- (165) Brahmshatriya, P. S.; Dobeš, P.; Fanfrlík, J.; Rezac, J.; Paruch, K.; Bronowska, A.; Lepsík, M.; Hobza, P. Quantum Mechanical Scoring: Structural and Energetic Insights into Cyclin-Dependent Kinase 2 Inhibition by Pyrazolo[1,5-a]pyrimidines. *Curr. Comput.-Aided Drug Des.* **2013**, *9*, 118–129.
- (166) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-optimization of Parameters. *J. Mol. Model.* **2013**, *19*, 1–32.
- (167) Perdew, J. P. Jacob’s Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2001**, *577*, 1.
- (168) Spicher, S.; Caldeweyher, E.; Hansen, A.; Grimme, S. Benchmarking London Dispersion Corrected Density Functional Theory for Noncovalent Ion- π Interactions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 11635–11648.
- (169) Dohm, S.; Bursch, M.; Hansen, A.; Grimme, S. Semi-automated Transition State Localization for Organometallic Complexes with Semiempirical Quantum Chemical Methods. *J. Chem. Theory Comput.* **2020**, *16*, 2002–2012.
- (170) Mehta, N.; Casanova-Páez, M.; Goerigk, L. Semi-Empirical or Non-Empirical Double-Hybrid Density Functionals: Which are More Robust? *Phys. Chem. Chem. Phys.* **2018**, *20*, 23175–23194.
- (171) Geary, R. C. The Ratio of the Mean Deviation to the Standard Deviation as a Test of Normality. *Biometrika* **1935**, *27*, 310–332.
- (172) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Many-Electron Self-Interaction Error in Approximate Density Functionals. *J. Chem. Phys.* **2006**, *125*, 201102.
- (173) Spicher, S.; Bursch, M.; Grimme, S. Efficient Calculation of Small Molecule Binding in Metal-Organic Frameworks and Porous Organic Cages. *J. Phys. Chem. C* **2020**, *124*, 27529–27541.
- (174) Harvey, J. N. On the Accuracy of Density Functional Theory in Transition Metal Chemistry. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2006**, *102*, 203–226.
- (175) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of Orbital-Optimized, Spin-Component Scaled Second-Order Many-Body Perturbation Theory for Thermochemistry and Kinetics. *J. Chem. Theor. Comput.* **2009**, *5*, 3060–3073.
- (176) Bertels, L. W.; Lee, J.; Head-Gordon, M. Third-Order Møller-Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions. *J. Phys. Chem. Lett.* **2019**, *10*, 4170–4176.
- (177) Langreth, D. C.; Perdew, J. P. The Exchange-Correlation Energy of a Metallic Surface. *Solid State Commun.* **1975**, *17*, 1425–1429.
- (178) Kállay, M. Linear-Scaling Implementation of the Direct Random-Phase Approximation. *J. Chem. Phys.* **2015**, *142*, 204105.
- (179) Mardirossian, N.; Head-Gordon, M. Survival of the Most Transferable at the Top of Jacob’s Ladder: Defining and Testing the ω B97M(2) Double Hybrid Density Functional. *J. Chem. Phys.* **2018**, *148*, 241736.
- (180) Santra, G.; Cho, M.; Martin, J. M. L. Exploring Avenues Beyond Revised DSD Functionals: I. Range Separation, with xDSD as a Special Case. *J. Phys. Chem. A* **2021**, *125*, 4614.
- (181) Santra, G.; Semidalas, E.; Martin, J. M. L. Exploring Avenues Beyond Revised DSD Functionals: II. Random-Phase Approximation and scaled MP3 corrections. *J. Phys. Chem. A* **2021**, *125*, 4628.
- (182) Williams, Z. M.; Wiles, T. C.; Manby, F. R. Accurate Hybrid Density Functionals with UW12 Correlation. *J. Chem. Theory Comput.* **2020**, *16*, 6176–6194.