Benchmarking Density Functional Theory Methods for Metalloenzyme Reactions: The Introduction of the MME55 Set

Dominique A. Wappett and Lars Goerigk*

School of Chemistry, The University of Melbourne, Victoria 3010, Australia

E-mail: lars.goerigk@unimelb.edu.au

Abstract

We present a new benchmark set of metalloenzyme model reaction energies and barrier heights, which we call MME55. The set contains eleven different enzymes, representing eight transition metals, both open and closed shell systems, and system sizes of up to 116 atoms. We use three DLPNO-CCSD(T)-based approaches to calculate reference values, against which we then benchmark the performance of a range of density functional approximations with and without dispersion corrections. Dispersion corrections improve the results across the board, and triple-ζ basis sets provide the best balance of efficiency and accuracy. While Jacob's Ladder is reproduced for the whole set based on averaged mean absolute deviations, hybrid DFT is preferred over double hybrids for copper-dependent enzymes, where MP2 correlation seriously adversely affects the results. Despite the popularity of B3LYP in computational enzymology, it is not a strong performer on our benchmark set, and we discourage its use for enzyme energetics. Instead, we recommend the range-separated hybrids ωB97M-V and ωB97X-V combined with the def2-TZVPP basis set for applications, as they are a great compromise between accuracy and efficiency and have already been shown to be robust across many other types of chemical problems.

1 Introduction

Across most areas of chemistry, Kohn-Sham Density Functional Theory[1,2] (KS-DFT, often shortened to just DFT) has become a common part of understanding reactions and structures, either alongside experiments or on its own. One particular application is the study of enzymes, where computational results can be used to understand the mechanisms of action and gain structural insights, which can then be applied to enzyme design or general catalysis applications. Generally these studies are done by treating a curated model of the active site with DFT and calculating the indirect effects of the remaining enzyme ei-

ther as an electrostatic field, as in cluster model studies, or with forcefields, as part of a hybrid quantum mechanics/molecular mechanics (QM/MM) scheme; for reviews of QM/MM approaches for enzyme modeling, see refs. 3 and 4.

While DFT is often a good choice for computational studies due to its accessibility and cost-to-accuracy ratio, there are an overwhelming number of density functional approximations (DFAs) available to users, each having been determined from different sets of data or first-principles boundary conditions, and therefore having different strengths. The choice of DFA can thus alter the conclusions drawn about possible mechanisms or structural features, so it is important to consider the choice carefully. This process is assisted by benchmarking studies that test the performance of density functionals, either on comprehensive databases of chemical reactions in order to find generally robust functionals, or test sets that are more specific to the chemical problem at hand. Some examples of broad benchmark sets that represent general main group chemistry are GMTKN55[5] and MGCDB84,[6] across which hundreds of DFAs have been tested—more than 350 by our group alone.[5,7–11] There are also smaller test sets that focus on organic biochemical reactions,[12–17] including work by us,[18] as well as sets that represent functional groups and reaction types often found in, but not exclusive to, enzyme chemistry.[19–23]

However biochemistry is not limited to main group elements. It is estimated that around one third of the human proteome requires metals to function,[24] and metalloenzymes are crucial in many other living species too. Unfortunately transition metals are harder to treat accurately with computational methods as their electronic structure is more complex, so recommendations of DFAs from tests on organic enzymes are not necessarily applicable to metalloenzymes. Due to difficulties in getting reliable reference values, there are significantly fewer benchmark studies that represent transitional metal chemistry. Most existing test sets feature inorganic dimers or small organometallic complexes,[25–38] or only include a small number of systems,[39–45] and these are all typically closed shell species. Various groups are working on filling the gaps of larger complexes and open shell species, particularly with the WCCR10,[46,47] MOR41,[48] MOBH35[49] and ROST61[50] sets. Chan et al. have also recently compiled the TMC151 database[51] from sets of organometallic reaction energies, organometallic barrier heights, and inorganic dimer bond energies, to be used as a (smaller) counterpart to the existing broad main group databases and find functionals that are robust across multiple types of transition metal chemistry.

Within computational bioinorganic chemistry, DFT has been tested for a range of properties, including spin state splitting,[52–54] metal-ligand binding and interaction energies,[55,56] excited states,[57] relative energies of Cu_2O_2 isomers,[41,58] redox potentials,[59] and bond dissociation energies.[60] Many studies of the mechanisms or properties of a specific enzyme/family of enzymes also validate the chosen DFT methodology with a comparison to ab initio methods[61–64] or other density functionals.[53,63,65–67] Most of these data, however, are small scale, both in the range of systems included and the methods tested, and were not tested against reliable references. We also note that some studies use small

simplified models,[55,59] an approach which does not reliably represent the range of interactions found in an enzyme. In our recent guide to benchmarking enzymatically catalyzed reactions,[68] we have shown that oversimplified models can have very different benchmarking outcomes compared to larger models for organic enzymes, as well as how the quality of the references will change the perceived performance of the tested functionals—a fact that has also been shown before for pericyclic and inorganic reactions.[5,69]

We have, thus, sought to create a benchmark set for metalloenzyme reactions to sit between our previous set of mostly organic enzyme models[18] and the MOBH35, MOR41 and ROST61 sets for transition metals, to explore whether the recommendations from these previous studies hold for bioinorganic systems. We additionally hope that by adding more data in the overlap of biochemical and organometallic benchmarking, this set can assist bioinorganic chemists in choosing reliable functionals for their calculations, or method developers in creating new functionals for accurate transition metal calculations.

In the following sections, we present a description of the set, which we call MME55—mechanisms of metalloenzymes, containing 55 data points, including both barrier heights (BHs) and reaction energies (REs). This is followed by an analysis of the multireference character of the included models, and details on how the reference values were calculated. Finally we test a range of density functionals as well as the still popular second-order Møller-Plesset Perturbation Theory[70] (MP2), analyze their performance, and compare the results to other related benchmarking studies. We also reiterate the importance of London dispersion corrections for DFT methods, which has been stated in many studies including for bioorganic[18,71,72] and bioinorganic[73–76] systems, although the recommendation is still not always followed.

2 Computational details

All calculations in this work were done using the ORCA[77–79] quantum chemistry package (versions 4.21, 5.0.1, 5.0.2 and 5.0.3). The Ahlrichs-type def2-nZVPP family of basis sets[80] was used, with the default def2-ECP[80,81] effective core potentials for the systems containing molybdenum and tungsten. The resolution of the identity approximation for Coulomb integrals (RI-J)[82] was used for Generalized Gradient Approximation (GGA), meta-GGA and meta-NGA (Nonseparable Gradient Approximation) functionals, while RI-J with the chain of spheres approximation for exchange integrals (RIJCOSX)[83] was used for the hybrid and double-hybrid functionals, as well as the Hartree-Fock steps of MP2 and Coupled Cluster calculations. MP2 steps—including in double hybrids—and Coupled Cluster calculations were sped up with the RI-C approximation[84] and used ORCA's default frozen core settings. These approximations were used with the def2/J[85] and def2-nZVPP/C[86] auxiliary basis sets and the "GridXS2" setting for RIJCOSX. Most calculations were done with the default grids and self-consistent-field (SCF) convergence thresholds.

Geometries of all structures were optimized at the PBEh-3c[87] level of theory in ORCA v4.2.1 with tight SCF and default geometry convergence settings, and the "grid3 finalgrid5" grids. PBEh-3c is

a PBE[88]-based hybrid functional, combined with a special double-ζ basis set and corrections for London dispersion [DFT-D3(BJ)[89,90]] and basis set superposition error (gCP[91]). For the multireference diagnostics, a test of the weighted fractional occupational electron density (FOD)[92] was done using the default settings, which are TPSS[93]/def-TZVP with TightSCF convergence and an occupation number smearing temperature of 5000 K. To calculate the A_λ diagnostic, PBE and PBE0[94,95] calculations were done with the def2-TZVPP basis set.

DLPNO-CCSD (domain based local pair natural orbital coupled cluster with singles and doubles) and DLPNO-CCSD(T) (with additional perturbative triples)[96,97] calculations were done in ORCA versions 5.0.1, 5.0.2 and 5.0.3, using the non-iterative (T0)[96] approach for the triples correction. For the initial multireference screening, def2-SVP was used with the NormalPNO thresholds.[98] For the benchmark energies, the def2-TZVPP and def2-QZVPP basis sets were used with TightPNO thresholds, and these results were extrapolated to the complete basis set (CBS) limit using the standard two point extrapolation schemes with individual extrapolations of the SCF[99] and correlation energies:[100]

$$E_{SCF}^{(\infty)} = \frac{E_{SCF}^{(X)} \cdot exp(-\alpha\sqrt{Y}) - E_{SCF}^{(Y)} \cdot exp(-\alpha\sqrt{X})}{exp(-\alpha\sqrt{Y}) - exp(-\alpha\sqrt{X})}, \quad (1)$$

and

$$E_{Corr}^{(\infty)} = \frac{X^\beta \cdot E_{Corr}^{(X)} - Y^\beta \cdot E_{Corr}^{(Y)}}{X^\beta - Y^\beta}, \quad (2)$$

where X and Y are the cardinal numbers of the basis sets, and α and β are optimized constants specific to the basis sets used. For the def2-TZVPP/def2-QZVPP CBS(3,4) extrapolation used here, $\alpha = 7.88$ and $\beta = 2.97$.[101]

All methods listed in table 1 were applied in ORCA v5.0.2 with the def2-QZVPP basis set, except for the composite DFT (-3c) methods which used their own specific basis sets. Additional calculations with B3LYP, M06, PWPB95 and ωB2PLYP were done with the def2-SVP and def2-TZVPP basis sets. The ORCA implementation of the LibXC density functional library[146] was used to run MN15-L, MN15 and MPW1B95 calculations. DOD-SCAN, revDSD-PBEP86 and revDOD-PBEP86 were run in both the DFT-D3(BJ) and DFT-D4[106,147] parameterized forms. For B97M-V, ωB97M-V and ωB97X-V, the non-local VV10 kernel was used in its post-SCF implementation, as this strategy does not impact the results but can lead to a considerable speedup in calculations.[8] The DFT-D3(BJ) dispersion correction with Becke-Johnson damping was used preferentially to the older, zero-damping DFT-D3(0)[89] variant for all functionals except M06-L, MN15-L, M06 and M062X—the Minnesota functionals are parameterized to have better descriptions of mid-range interactions, but this leads to some double-counting effects when the D3(BJ) correction is applied.[5] Specific dispersion correction damping parameters for B3LYP* have not been determined yet, so we follow the common approach of applying the standard B3LYP damping parameters.[148–154]

Table 1: Methods tested on the MME55 set. References for DFT-D3 and DFT-D4 are where the damping parameters for each functional were first presented, with D3 referring to the Becke-Johnson damping (D3(BJ)) variant unless otherwise stated.

Type	Name	References			% HF	% MP2
		Method	D3	D4		
GGA	B97-3c	102	102	–		
GGA	BLYP	103–105	90	106		
GGA	BP86	103,107,108	90	106		
GGA	OLYP	104,105,109	110	106		
GGA	OPBE	88,109	110	106		
GGA	PBE	88	90	106		
GGA	PW91	111	112	106		
GGA	revPBE	113	90	106		
meta-GGA	B97M-V [a]	114	8	10		
meta-GGA	M06L	115	110[b]	106		
meta-NGA	MN15-L	116	5[b]	–		
meta-GGA	r2SCAN	117	118	118		
meta-GGA	r2SCAN-3c	119	–	119		
meta-GGA	revTPSS	120,121	5	106		
meta-GGA	TPSS	93	90	106		
hybrid	B3LYP	122,123	89[b], 90	106	20	
hybrid	B3LYP* [c]	124	89[b], 90	106	15	
hybrid	BHLYP	125	110	106	50	
hybrid	CAM-B3LYP	126	110	106	19–65 [d]	
hybrid	M06	127	110[b]	106	27	
hybrid	M062X	127	110[b]	106	54	
hybrid	MN15	38	5	–	44	
hybrid	MPW1B95	128	110	106	31	
hybrid	PBE0	94,95	90	106	25	
hybrid	PBEh-3c	87	87		42	
hybrid	PW6B95	129	90	106	28	
hybrid	TPSS0	130	90	106	25	
hybrid	TPSSh	131	110	106	10	
hybrid	ωB97M-V [a]	132	8	10	15–100 [d]	
hybrid	ωB97X-V [a]	133	8	10	16.7–100 [d]	
double hybrid	B2PLYP	134	110	106	53	27
double hybrid	B2GP-PLYP	135	110	106	65	36
double hybrid	B2K-PLYP	135,136	69	–	72	42
double hybrid	DOD-SCAN-D3(BJ) [e]	137	137	–	66	0/62.83 [f]
double hybrid	DOD-SCAN-D4 [e]	137	–	137	66	0/63.44 [f]
double hybrid	mPW2PLYP	138	5	106	55	25
double hybrid	PBE0-DH	139	140	106	50	12.5
double hybrid	PWPB95	141	110	106	50	0/26.9 [f]
double hybrid	revDOD-PBEP86-D3(BJ) [e]	137	137	–	69	0/60.55 [f]
double hybrid	revDOD-PBEP86-D4 [e]	137	–	137	69	0/61.22 [f]
double hybrid	revDSD-PBEP86-D3(BJ) [e]	137	137	–	69	7.9/57.85 [f]
double hybrid	revDSD-PBEP86-D4 [e]	137	–	137	69	6.36/59.22 [f]
double hybrid	SOS0-PBE0-2	142	7	–	79.37	0/66 [f]
double hybrid	ωB2PLYP	143	11	11	53–100 [d]	27
double hybrid	ωB2GP-PLYP	143	11	11	65–100 [d]	36
double hybrid	ωB88PP86	144	–	–	65–100 [d]	42
double hybrid	ωPBEPP8	144	–	–	70–100 [d]	48
ab initio	MP2	70	7,145	–	100	100

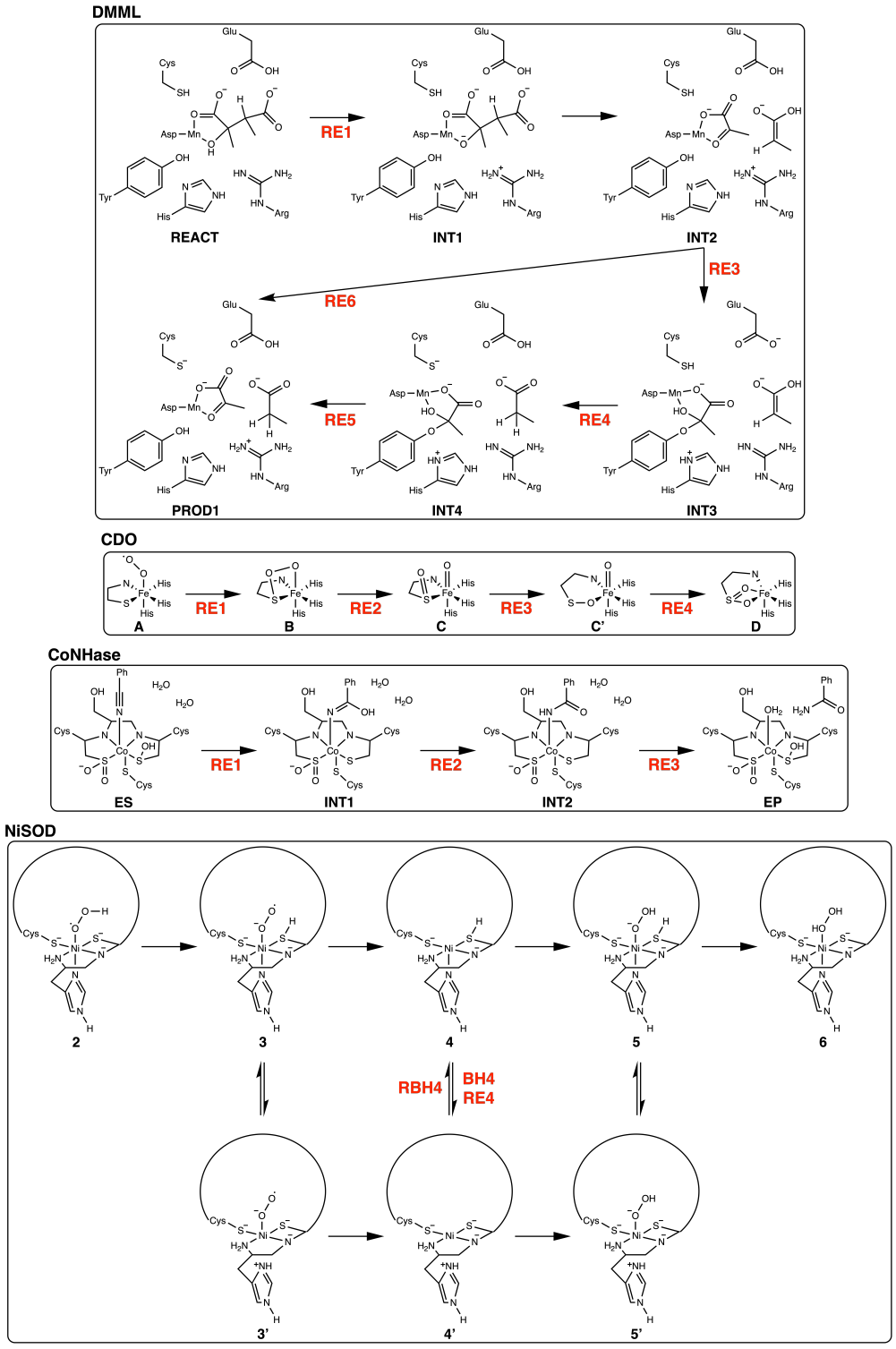
[a] The DFT-D3(BJ)/4 dispersion correction replaces the VV10 kernel in this van-der-Waals functional, while its semi-local exchange-correlation component stays the same. [b] DFT-D3(0) correction. [c] Damping parameters for B3LYP* have yet to be defined, so we use the standard B3LYP damping parameters. [d] Range-separated functionals with variable amounts of Fock exchange. [e] The underlying parameters for these functionals differ between their D3(BJ) and D4 versions. [f] Scale factors for the same spin and opposite spin contributions to the MP2 correlation energy, respectively.

3 Development of the MME55 set

3.1 Enzyme models included in the set

Two main considerations were involved in the selection of enzyme models. The first was that the set should represent a range of different transition metals, spin states, and reaction types; the second was that the systems should be of similar size to those used in typical enzyme modeling studies. The generation of reference data becomes more complicated as systems get larger or more metal ions are included; on the other hand, we have previously shown that using reduced models of enzyme active sites is inadequate for benchmarking,[68] and this is likely even more inappropriate in metalloenzymes where the coordination environments of the metal ions also need to be represented. We thus searched the literature for DFT and QM/MM studies of enzyme mechanisms, focusing on ones where the active site model or QM region contained up to 120 atoms and no more than 2 metal centers, so that the entire model could be used without needing to be simplified. The selected enzyme models, which range in size from 51-116 atoms and cover eight different transition metals, are:

- (2R,3S)-Dimethylmalate lyase[155] (DMML): a manganese-dependent enzyme that catalyzes the cleavage of dimethyl malate, forming propionate and pyruvate.
- Cysteine dioxygenase[65] (CDO): an iron-dependent enzyme that catalyzes the metabolization of cysteine. The four-step process is modeled in the singlet, triplet and quintet spin states.
- Nitrile hydratase[156] (Co-NHase): a cobalt-dependent enzyme that catalyzes the hydrolysis of organic nitriles into their amides.
- Superoxide dismutase[157] (NiSOD): a nickel-dependent enzyme that catalyzes the disproportionation of superoxide to molecular oxygen and hydrogen peroxide.
- Hemocyanin[67] (Hc): the oxygen binding process in the Cu_2O_2 core, modeled in both the singlet and triplet states.
- Aminopeptidase[158] (AAP): a zinc-dependent enzyme that catalyzes the cleavage of the N-terminal amino acid residues of polypeptides and proteins.
- Phosphotriesterase[159] (PTE): a zinc-dependent enzyme that catalyzes hydrolysis of organophosphate triesters. This model was also included in the mainly-organic enzyme set that we have previously used to benchmark DFT.[18]
- Perchlorate reductase[160] (PcrAB): a molybdenum-dependent enzyme that catalyzes the conversion of perchlorate to chlorate and subsequently chlorite.
- Acetylene hydratase[161] (AH): a tungsten-dependent enzyme that catalyzes the nonredox hydration of acetylene to acetaldehyde.
- Formaldehyde ferredoxin oxidoreductase[162] (W-FOR): a tungsten-dependent enzyme that catalyzes the reduction of formaldehyde to formic acid.


DMML
REACT
RE1
INT1
INT2
RE3
RE6
PROD1
RE5
INT4
RE4
INT3
CDO
A
RE1
B
RE2
C
RE3
C'
RE4
D
CoNHase
ES
RE1
INT1
RE2
INT2
RE3
EP
NiSOD
2
3
4
5
6
RBH4
BH4
RE4
3'
4'
5'

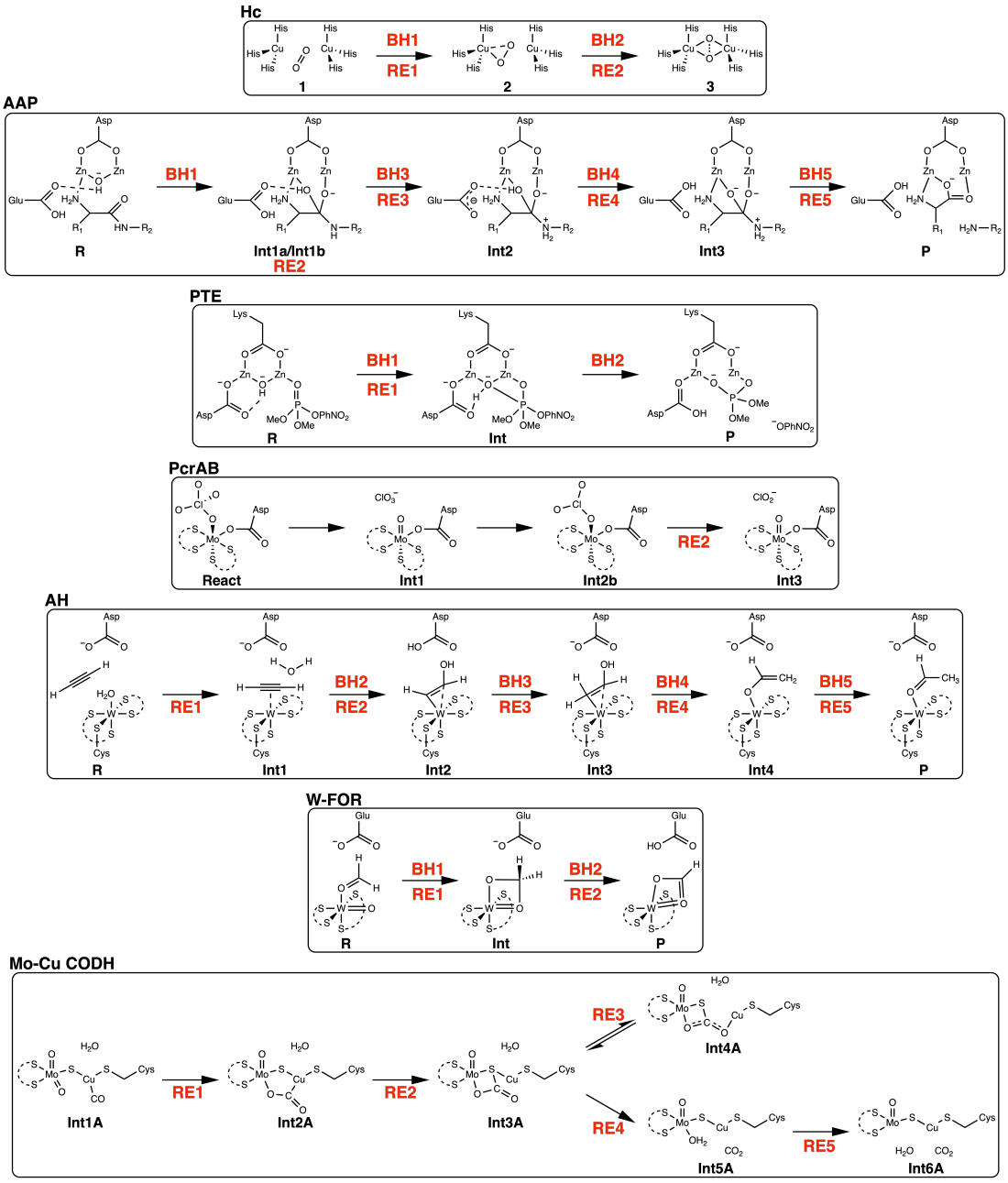


Figure 1: Reaction schemes for the enzyme models, adapted from Refs. 65,67,155–163. Steps that are included in the final MME55 set are labeled in red.

- Carbon monoxide dehydrogenase[163] (Mo-Cu CODH): a binuclear molybdenum-copper enzyme that catalyzes the oxidation of CO to CO_2.

Reaction schemes for the original eleven models are shown in fig. 1.

3.2 Geometry optimizations

After taking the published structures for each enzyme included in the set, our first step was to reoptimize the geometries to ensure that a consistent level of theory was used. Any constraints that had been applied in the previous studies (see SI for details) were maintained in our optimizations, and for some structures in the PcrAB and Hc systems, ORCA's fragment optimization feature was used to ensure that correct configurations were maintained. The original DMML structures were altered pre-optimization to replace the carbon atoms at the QM/MM crossover points with hydrogens, so that they could be optimized as cluster models. For systems where multiple spin state surfaces were included, the structures were separately optimized in each spin state. In the case of NiSOD, only the most stable state for each structure was provided in Pelmenschikov and Siegbahn's study, but in our reoptimizations, we have ensured that all reaction steps occur on the same spin surface to eliminate any spin-crossover effects from our calculated barrier heights and reaction energies. Not all transition states of each reaction were able to be successfully reoptimized, so the final set contains no barrier heights for DMML, CoNHase, PcrAB, Mo-Cu CODH, CDO and the singlet state of Hc.

The importance of dispersion corrections in geometry optimizations has been shown numerous times,[71,72] including for organometallic complexes.[164] In our previous work on enzyme benchmarking,[18] we showed some examples of how PBEh-3c structures improved the description of dispersion-supported interactions such as hydrogen bonding and aromatic π-stacking, including in the PTE model which we have also included here in the MME55 set. We see similar changes in some of the structures here, and we give two examples in fig. 2; see figure caption for dispersion-uncorrected levels of theory of the original structures. For AAP Int1, the optimized structure has the substrate rotated so that the acetyl group can interact with one of the histidine residues. For CoNHase ES, an off-center parallel π-stacking interaction between the benzonitrile substituent and tyrosine residue is only seen in the optimized structure, not the original structure.

We note that in some cases, the optimizations have converged in arrangements that do not match the original crystal structures from while the models were developed, and may be unfavorable for the overall mechanisms—for example, with amino acid side chains rotating away from where they will later react. While this is a problem when comparing to experimental data, here we use the same structures for both the theoretical reference values and the DFT tests, so there is less influence on the calculated deviations. The reoptimized systems should still represent the types and overall magnitude of non-covalent interactions that are seen in enzymes, even if they are configured differently, and therefore we believe that slight alterations in the structures are not harmful for the purpose of a benchmark study.

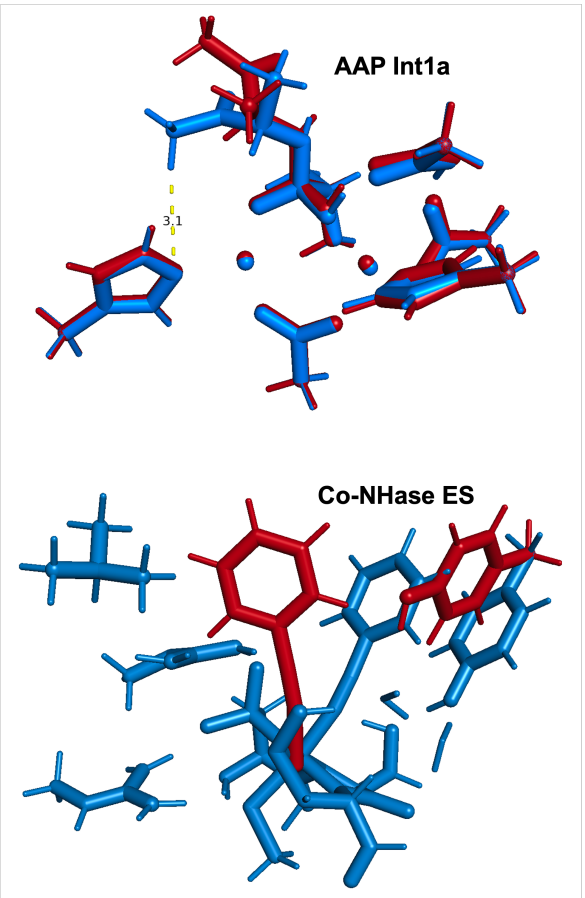


Figure 2: Comparisons between original structures (red) and reoptimized structures (blue). For clarity one histidine residue has been hidden in AAP Int1a, while for the original structure of Co-NHase ES, only the Co centre, benzonitrile substituent and tyroside side chain are shown. Original AAP structure: B3LYP/6-31G(d,p)[165] with Stuttgart-Dresden ECP[81] on Zn. Original Co-NHase structure: M06L/6-31G(d,p) with Stuttgart-Dresden ECP on Co. Reoptimized structures: PBEh-3c.

3.3 Testing multireference character

To ensure that reliable single reference benchmark energies could be calculated for the systems, the multireference character of each structure was checked with a range of diagnostics. First A_λ values[166] were obtained using the PBE and PBE0 density functionals. Then DLPNO-CCSD(T)/TightPNO/def2-SVP calculations were done to test the %TAE(T),[167] T1 diagnostic[168] and largest T_2 amplitude. The A_λ and %TAE(T) approaches test the sensitivity of the total atomization energies to the level of theory (namely Hartree-Fock and higher order coupled cluster excitations, respectively). $A_\lambda \leq 0.1$ and %TAE(T) $\leq 2\%$ generally mean that a system is dominated by dynamic correlation, while values of $A_\lambda \approx 0.15$ and %TAE(T) up to 5% indicate mild multireference character. For the T1 diagnostic, typically a threshold of 0.02 is used, but Wilson and coworkers state that this can be increased up to 0.05 for systems containing transition metals.[169] While thresholds for the largest T_2 amplitude are not as clearly defined, generally values of 0.1–0.2 have been considered mild but not necessarily problematic when other diagnostics are low,[41,166] while some say up to 0.15 can be considered low.[50,169]

The results of these diagnostic tests are summarized in table 2. %TAE(T) values could not be calculated for PcrAB and Mo-Cu CODH due to convergence issues for the Mo atom calculation, but given that no systems have %TAE(T) $\geq 2\%$, even those which are indicated to have mild multireference character by the other diagnostics, it is likely that these values would also be low. Looking at the other diagnostics, we see that none of the T1 values are greater than the revised 0.05 threshold for transition metals, and even the largest A_λ values still only indicate mild multireference character. The trends across the diagnostics,

Table 2: Summarized results of tests of the multireference character of each system, with values indicating mild or moderate multireference character shown in bold and the number of structures above the threshold for each diagnostic given in parentheses. Overall ranges for each reaction are given, while full results can be found in the Supporting Information. Any reaction steps involving structures that are flagged by three or more diagnostics are cut from the final set.

Reaction	No. of strucs	A_λ	%TAE(T)	T1 diagnostic	Largest T_2 amp.	No. of strucs cut	Remaining no. of BHs	Remaining no. of REs
DMML	6	0.10–0.10	1.18–1.20	0.012–0.013	0.054–0.056	0	0	6
CDO	15	**0.11–0.14** *(15)*	1.51–1.76	0.013–**0.039** *(5)*	0.053–**0.229** *(2)*	2	0	8
Co-NHase	4	**0.11–0.11** *(4)*	1.41–1.42	0.016–0.018	0.069–0.082	0	0	3
NiSOD	13	**0.11–0.14** *(13)*	1.45–1.70 [a]	0.014–**0.048** *(8)* [a]	0.056–**0.215** *(8)* [a]	10	2	1
Hc	8	**0.13–0.14** *(8)*	1.66–1.97	0.013–0.018	0.045–**0.196** *(2)*	0	2	4
AAP	10	0.09–0.10	1.23–1.25	0.013–0.013	0.061–0.063	0	4	5
PTE	5	**0.12–0.12** *(5)*	1.53–1.54	0.013–0.013	0.055–0.057	0	2	2
PcrAB	4	**0.16–0.19** *(4)*	–[a]	**0.023–0.031** *(4)*	0.056–**0.421** *(1)*	1	0	1
AH	10	0.09–0.09	1.38–1.45	0.013–0.015	0.057–0.070	0	4	5
W-FOR	5	0.10–0.10	1.40–1.52	0.014–0.016	0.057–0.065	0	2	2
Mo-Cu CODH	6	**0.13–0.14** *(6)*	–[a]	0.016–0.016	0.064–0.084	0	0	5

[a] The DLPNO-CCSD(T) calculation could not be converged for the Mo atom or NiSOD structures 3 and TS5, so some diagnostics could not be obtained. These two NiSOD structures were, thus, also cut from the set.

however, are a more reliable indicator of problematic behavior than any individual one. We therefore use the more conservative T1 = 0.02 and largest T_2 amplitude = 0.1 thresholds, along with $A_\lambda = 0.1$, and any structures that are above the thresholds on all three diagnostics are removed from the set.

For DMML, AAP, AH and W-FOR, all structures are shown to have low multireference character on all diagnostics. All structures for Co-NHase, PTE and Mo-Cu CODH, as well as the triplet state of Hc and quintet state of CDO, have A_λ values between 0.1 and 0.15, but are still low on the other three diagnostic tests. The A_λ values are similarly low-mild for the singlet state of Hc, while the largest T_2 amplitudes, ranging from 0.098 to 0.196, also indicate mild multireference character. PcrAB has the highest A_λ values of the set, and the T1 diagnostic values are also slightly above the general 0.02 threshold. One structure has a very high T_2 amplitude of 0.421, but all others are low (<0.075). Of the 10 CDO structures across the singlet and triplet states, all have A_λ between 0.11 and 0.14, five have T1 values between 0.02 and 0.05 and two have a largest T_2 amplitude greater than 0.1. NiSOD is the most problematic, with almost all structures being flagged as having mild to moderate multireference character by all three of the A_λ, T1 diagnostic, and largest T_2 amplitudes—only one reaction step can be considered to have low multireference character. Removal of the structures flagged on three diagnostics leads to one reaction step being cut from PcrAB, four from CDO and ten from NiSOD. The remaining number of data points in each reaction are listed in the final columns of table 2, and the steps included in the final set are labeled in fig. 1. We note that the four steps shown for CDO are modeled in three different spin states so there were initially 12 REs; two steps were cut from the singlet state (1RE3 and 1RE4) and two from the triplet state (3RE2 and 3RE3).

We briefly note that the T1 diagnostic and maximum T_2 amplitudes were also checked for the coupled cluster approaches used to calculate the reference energies. The differences between the initial screen-

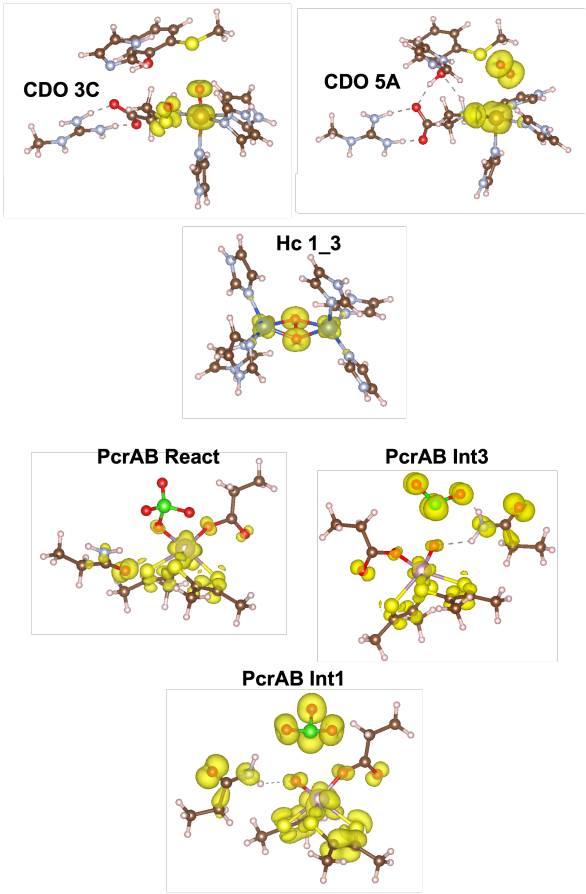


Figure 3: FOD plots of selected structures. Surfaces are plotted with σ=0.005 e bohr^{-3}.

ing NormalPNO/def2-SVP and higher level TightPNO/def2-TZVPP and def2-QZVPP values were minimal, and there was no change in which structures were flagged as mild cases. There were a few differences, however, for the calculations with modified TightPNO settings, particularly $T_{CutPNO} = 10^{-5}$. These gave T1 $\leq$ 0.02 for all four PcrAB structures and NiSOD 6, which all had T1 > 0.02 in all the other calculations. The diagnostics from all coupled cluster calculations are given in the Supporting Information.

We also calculated and visualized the weighted fractional occupational electron density (FOD). Overall the results are very similar to the other diagnostics. DMML, AAP and PTE all have virtually no FOD, while for Co-NHase, AH, W-FOR, Mo-Cu CODH, and the three safe NiSOD structures the FOD is small and mostly metal-centered. A few examples of some of the structures which were flagged as mildly multireference by at least one diagnostic are shown in fig. 3, while all others are provided in the Supporting Information (figures S1-S11). All structures for CDO have similar levels of FOD, even between ones that are flagged by all the other diagnostics and ones that are low on both the T1 and maximum T_2 amplitudes (like 3C and 5A, respectively). For Hc, it is mild but spread evenly across the whole Cu_2O_2 core. PcrAB has relatively delocalized FOD, which is mainly spread across the molybdenum, oxygen and sulfur atoms. It is large for Int1, which is cut from the set based on the other diagnostics, and mild-moderate for the other structures.

3.4 Calculation of reference values

We aimed to calculate the reference BHs and REs for this set using DLPNO-CCSD(T) with the TightPNO thresholds, which is regularly used in benchmarking as a cost effective alternative to "gold standard" CCSD(T) results. In our previous benchmark study on enzymatically catalysed reactions,[18] we did an analysis of various extrapolation and estimation schemes to determine a protocol for calculating reference values. Our first choice was a DLPNO-CCSD(T)/TightPNO/CBS(3,4) extrapolation with aug-cc-pVTZ/aug-cc-

pVQZ,[170,171] followed by the equivalent extrapolation with the ma-def2-TZVPP/ma-def2-QZVPP[172] basis sets. The third option was to use a composite scheme based on MP2, which is corrected with the difference in correlation energy between MP2 and DLPNO-CCSD(T) (CC) calculated with a small basis set, like so:

$$E = E_{MP2}^{(CBS)} + [E_{Corr(CC)}^{(TZ)} - E_{Corr(MP2)}^{(TZ)}] \quad (3)$$

Due to the size of the systems in MME55, we can already rule out the Dunning basis sets as being too expensive to their higher number of primitive Gaussian-type orbitals. The def2 basis sets gave almost identical results to their minimally augmented equivalents in our previous tests of basis sets, while they were also shown to agree well with aug-cc-pVTZ/aug-cc-pVQZ results in the ROST61 benchmark study.[50] Therefore, we consider a DLPNO-CCSD(T)/TightPNO/CBS(def2-TZVPP/def2-QZVPP) extrapolation an appropriate first choice level of theory for calculation of the reference values. For larger systems where this is still too expensive, we cannot use the MP2 based composite scheme again, as MP2 is known to be unreliable for transition metal chemistry,[41,49,50,141] so we have tested two additional alternative strategies here.

The first of these involves extrapolation of DLPNO-CCSD results and then correcting that energy with the (T) triples energy calculated at the triple-ζ level,

$$E = E_{SCF}^{(CBS)} + E_{Corr(DLPNO-CCSD)}^{(CBS)} + E_{(T)}^{(TZ)} \quad (4)$$

which cuts the computational time by eliminating the lengthy triples computation at the quadruple-ζ level. The other alternative approach, for systems where even DLPNO-CCSD/TightPNO/def2-QZVPP calculations were unfeasible, uses the PNO extrapolation scheme described by Altun et al.[173] This involves calculating results with increasing values of the threshold T_{CutPNO}, which defines the PNOs included in the final Coupled Cluster treatment, while maintaining all other cutoffs at the same values. To approximate standard TightPNO results, where $T_{CutPNO} = 10^{-7}$, calculations are done with $T_{CutPNO} = 10^{-5}$ and $T_{CutPNO} = 10^{-6}$, and the correlation energies are extrapolated as such:

$$E_{Corr} = E_{Corr}^{(5)} + 1.5 \cdot (E_{Corr}^{(6)} - E_{Corr}^{(5)}) \quad (5)$$

where $E_{Corr}^{(X)}$ is the correlation calculated with TCutPNO $= 10^{-X}$. This is done for each basis set, and then the PNO-extrapolated triple-ζ and quadruple-ζ results are extrapolated to give CBS results. We refer to the approaches described in equations 4 and 5 as "estimated CBS(3,4)", and "estimated TightPNO", respectively, and we compare both to DLPNO-CCSD(T)/TightPNO/CBS(3,4) for the three smallest systems in table 3.

We see that the estimated CBS(3,4) approach is generally a good alternative to the full extrapolation. None of the deviations are larger than 1 kcal/mol, the generally considered "chemical accuracy limit" for REs and BHs, and the deviations for the closed shell systems (Hc singlet and PcrAB) are all below 0.4 kcal/mol. Given that the triples correction is often the longest part of the DLPNO-CCSD(T) calculation, this approach cuts down the cost of the def2-QZVPP computation significantly while maintaining reasonable accuracy. The estimated TightPNO also cuts costs, but comes with a greater reduction in accuracy—the mean absolute deviation

Table 3: Tests of Coupled Cluster approaches for calculating benchmark energies. Deviations are calculated from the DLPNO-CCSD(T)/TightPNO/CBS(3,4) values. All values in kcal/mol.

System	Step	DLPNO-CCSD(T)/ TightPNO/CBS(3,4)	est. CBS(3,4)		est. TightPNO	
			Value	Deviation	Value	Deviation
NiSOD	RE4	25.484	25.770	*0.285*	25.159	*−0.326*
NiSOD	BH4	27.909	28.103	*0.193*	27.110	*−0.799*
NiSOD	RBH4	2.425	2.333	*−0.092*	1.951	*−0.474*
Hc	1RE1	−26.948	−27.053	*−0.105*	−27.605	*−0.657*
Hc	1RE2	−4.820	−4.785	*0.035*	−3.638	*1.182*
Hc	3RE1	5.085	4.210	*−0.875*	4.279	*−0.806*
Hc	3RE2	−8.022	−8.793	*−0.771*	−9.473	*−1.451*
Hc	3BH1	5.336	5.415	*0.079*	6.997	*1.661*
Hc	3BH2	4.068	4.122	*0.054*	4.792	*0.723*
PcrAB	RE2	−48.622	−48.252	*0.370*	−50.606	*−1.984*
Mean absolute deviation:				*0.286*		*1.006*
Mean deviation:				*−0.083*		*−0.293*

is just above 1 kcal/mol, and only one step (Hc 3RE1) is better with this approximation than with estimated CBS(3,4). However we note that the error seen here for this approach may be unfairly influenced by the systems we have used to test it. The DLPNO scheme involves treating only strongly interacting pairs at the Coupled Cluster level, while weakly interacting pairs are calculated with MP2, but as Liakos and Neese have pointed out,[41] the MP2 correlation energy of the Cu_2O_2 core is wildly inaccurate compared to CCSD(T) and LPNO-CCSD. In the calculations with looser T_{CutPNO} values, more pairs are considered weak pairs, and therefore MP2 has a stronger influence on the overall energy. This may lead to larger than expected changes between the $T_{CutPNO} = 10^{-5}$ and 10^{-6} results, resulting in the PNO extrapolation overshooting the standard TightPNO correlation energy. Also, the two structures involved in PcrAB RE2 have T1 diagnostic values > 0.02, which Altun et al. have shown can lead to slower convergence of the DLPNO-CCSD(T) correlation energy with increasing T_{CutPNO},[174] which may impact the accuracy of the extrapolation. For the three NiSOD energies, for which all structures have T1 < 0.015, the deviations are much smaller.

Co-NHase is the only model in the set large enough to require the estimated TightPNO approach. It is a closed shell cobalt enzyme with low T1 diagnostic values, therefore we believe that the error of this level of theory will not be unreasonable for this system. For Hc, NiSOD and PcrAB, the DLPNO-CCSD(T)/TightPNO/CBS(3,4) values were used as the references for the following benchmark study, and all other systems used the estimated CBS(3,4) approach. At this point, we have cut three additional steps from the set that had reference REs of <0.1 kcal/mol, as these give unreliably high percentage deviations for the tested DFAs. The final reference values for all 55 data points in the MME55 set are given in table 4.

Table 4: Final reference values for the MME55 set, given in kcal/mol.

Reaction	Step[f]	Ref. value	Reaction	Step[f]	Ref. value	Reaction	Step[f]	Ref. value
DMML	RE1	−0.851 [b]	Hc	1RE1[d]	−26.948 [a]	AH	RE1	−9.245 [b]
DMML	RE3	24.207 [b]	Hc	1RE2[d]	−4.820 [a]	AH	RE2	10.083 [b]
DMML	RE4	65.206 [b]	Hc	3RE1[d]	5.085 [a]	AH	RE3	−21.181 [b]
DMML	RE5	−9.256 [b]	Hc	3RE2[d]	−8.022 [a]	AH	RE4	−3.483 [b]
DMML	RE6	−31.744 [b]	Hc	3BH1[d]	5.336 [a]	AH	RE5	−3.184 [b]
CDO	1RE1[d]	−1.151 [b]	Hc	3BH2[d]	4.068 [a]	AH	BH2	16.862 [b]
CDO	1RE2[d]	−52.657 [b]	AAP	RE2	4.187 [b]	AH	BH3	14.846 [b]
CDO	3RE1[d]	−3.589 [b]	AAP	RE3	0.867 [b]	AH	BH4	8.503 [b]
CDO	3RE4[d]	−53.488 [b]	AAP	RE4	0.682 [b]	AH	BH5	17.423 [b]
CDO	5RE1[d]	−43.640 [b]	AAP	RE5	−3.480 [b]	W-FOR	RE1	−3.500 [b]
CDO	5RE2[d]	−67.777 [b]	AAP	BH1	8.830 [b]	W-FOR	RE2	−14.436 [b]
CDO	5RE3[d]	−10.256 [b]	AAP	BH3	2.484 [b]	W-FOR	BH1	14.308 [b]
CDO	5RE4[d]	−1.084 [b]	AAP	BH4	1.706 [b]	W-FOR	BH2	19.940 [b]
Co-NHase	RE1	28.079 [c]	AAP	BH5	1.479 [b]	Mo-Cu CODH	RE1	4.886 [b]
Co-NHase	RE2	−27.586 [c]	PTE	RE1	2.459 [b]	Mo-Cu CODH	RE2	−10.823 [b]
Co-NHase	RE3	8.487 [c]	PTE	BH1	6.796 [b]	Mo-Cu CODH	RE3	−3.770 [b]
NiSOD	RE4	25.484 [a]	PTE	BH2	8.256 [b]	Mo-Cu CODH	RE4	4.483 [b]
NiSOD	BH4	27.909 [a]	PcrAB	RE2	−48.622 [a]	Mo-Cu CODH	RE5	−1.684 [b]
NiSOD	RBH4[e]	2.425 [a]						

Level of theory: [a]DLPNO-CCSD(T)/TightPNO/CBS(3,4), [b]estimated CBS(3,4), [c]estimated TightPNO. [d]A number in front the label indicates multiplicity when needed for clarity. [e]Barrier of reverse reaction. [f]Specific details of each step can be found in the Supporting Information.

4 Benchmark study: results and discussion

Using the MME55 set, we have tested the density functionals listed in table 1 with and without dispersion corrections, resulting in a total of 119 unique combinations. Deviations were calculated as $RE_{method} - RE_{ref.}$ or $BH_{method} - BH_{ref.}$ for all data points in the set. Percent deviations (PDs) were calculated as $\frac{deviation}{|ref.value|} \times 100\%$, using absolute values of the reference REs and BHs so that the PDs have the same signs as the deviations. Statistical analysis was done on these values, and in table 5 we present mean absolute deviations and percent deviations (MADs and MAPDs) for each functional. Violin plots of the deviations across each rung of Jacob's Ladder[175] are shown in fig. 4, along with the mean MAD (MMAD) of each category. In this section we discuss the D3(BJ) and D3(0) results together generally under the label "D3", and only name the specific variant for individual functionals. For B3LYP, which was tested with both variants, the D3(BJ) results are used preferentially to D3(0) unless otherwise stated.

The results in table 5 and figure 4 show two key trends—the general improvement in the performance of density functionals as one goes up the rungs of Jacob's Ladder, and the improvement of the results when dispersion corrections are applied—that match the findings of many other benchmark studies.[5,18,21,23,47,48,50,176] The MMAD and MMAPD of each rung is lower than for the previous rung for both the plain and dispersion corrected functionals, although the improvements of double hybrids over

Table 5: Mean absolute deviations (kcal/mol) and mean absolute percent deviations (%) for all assessed methods, listed in alphabetical order within the rungs of Jacob's Ladder, as well as mean MADs and MAPDs for each rung of Jacob's Ladder. Results are for the def2-QZVPP basis set, except for the -3c methods which use their own modified basis sets. The lowest value for each statistic in each rung is bolded.

Type	Method	Plain		D3(0)		D3(BJ)		D4		V	
		MAD	MAPD	MAD	MAPD	MAD	MAPD	MAD	MAPD	MAD	MAPD
GGA	B97-3c					6.4	71.8				
GGA	BLYP	7.7	111.3			6.4	74.9	6.5	76.1		
GGA	BP86	7.0	91.5			6.5	80.5	6.7	81.0		
GGA	OLYP	8.1	127.9			**5.5**	71.4	6.2	85.7		
GGA	OPBE	7.9	120.3			6.6	100.2	7.3	108.7		
GGA	PBE	7.0	89.3			6.4	75.2	6.5	77.3		
GGA	PW91	6.9	87.9			6.5	77.8	6.5	77.6		
GGA	revPBE	7.5	107.5			5.9	**68.0**	6.4	77.8		
meta-GGA	B97M					4.6	60.9	4.2	56.3	4.3	**55.9**
meta-GGA	M06L	4.8	69.1	4.8	68.1			4.7	67.1		
meta-NGA	MN15-L	3.7	71.2	**3.7**	71.2						
meta-GGA	r2SCAN	5.4	69.3			5.2	65.1	5.2	66.0		
meta-GGA	r2SCAN-3c							6.0	70.1		
meta-GGA	revTPSS	6.2	77.4			6.2	77.6	6.3	78.4		
meta-GGA	TPSS	6.6	85.8			6.0	70.4	6.2	73.2		
hybrid	B3LYP	5.0	78.9	4.0	49.9	3.8	45.3	3.9	48.3		
hybrid	B3LYP*	5.5	81.9	4.5 [a]	55.7 [a]	4.3 [a]	50.9 [a]	4.4 [a]	52.1 [a]		
hybrid	BHLYP	5.3	75.1			4.3	48.9	4.1	48.0		
hybrid	CAM-B3LYP	3.6	54.5			3.0	44.1	2.9	39.1		
hybrid	M06	3.0	47.5	3.0	45.8			3.0	45.7		
hybrid	M062X	4.0	46.9	4.0	46.5			4.0	46.6		
hybrid	MN15	3.2	37.0			3.2	37.0				
hybrid	MPW1B95	2.9	37.6			2.7	31.7	2.8	34.1		
hybrid	PBE0	3.8	52.3			3.2	39.3	3.4	42.8		
hybrid	PBEh-3c					4.3	55.9				
hybrid	PW6B95	3.1	41.6			2.8	31.4	2.9	33.1		
hybrid	TPSS0	3.5	49.7			3.0	41.1	3.2	41.4		
hybrid	TPSSh	5.2	72.7			4.5	55.7	4.7	58.2		
hybrid	ωB97M					2.9	31.3	2.8	30.6	**2.5**	**24.5**
hybrid	ωB97X					3.4	39.3	2.9	31.2	2.5	27.0
double hybrid	B2PLYP	4.0	63.2			3.7	51.0	3.8	54.1		
double hybrid	B2GP-PLYP	3.3	51.5			3.1	43.1	3.1	47.0		
double hybrid	B2K-PLYP	3.5	48.7			3.4	44.1				
double hybrid	DOD-SCAN [b]					3.2	46.0	3.3	50.5		
double hybrid	mPW2PLYP	3.5	53.4			3.2	45.0	3.2	46.4		
double hybrid	PBE0-DH	3.1	41.8			3.0	39.5	2.8	37.4		
double hybrid	PWPB95	2.7	36.8			2.6	33.2	2.6	32.8		
double hybrid	revDOD-PBEP86 [b]					2.6	34.3	2.5	35.3		
double hybrid	revDSD-PBEP86 [b]					2.7	37.6	2.7	39.1		
double hybrid	SOS0-PBE0-2	2.6	36.1			**2.3**	**31.8**				
double hybrid	ωB2PLYP	2.5	37.3			2.5	37.2	2.5	37.2		
double hybrid	ωB2GP-PLYP	2.6	38.8			2.6	38.8	2.6	38.8		
double hybrid	ωB88PP86	4.4	79.1								
double hybrid	ωPBEPP86	5.6	102.8								
ab initio	MP2	7.1	89.6			7.0	87.8				

Rung	Plain		All D3		D4	
	MMAD	MMAPD	MMAD	MMAPD	MMAD	MMAPD
All GGAs	7.4	105.1	6.3	77.5	6.6	83.5
All meta-GGA/NGAs	5.3	74.6	5.1	68.9	5.5	68.5
All hybrids	4.0	56.3	3.5	42.9	3.5	42.4
All double hybrids	3.4	53.6	2.9	40.1	2.9	41.9

[a] Calculated using the damping parameters for B3LYP. [b] Separate functional parametrizations were used for the D3(BJ) and D4 versions of these functionals.

hybrids are relatively small. As mentioned earlier, MP2 is unreliable for transition metals—here it is one of the worst methods tested—which somewhat weakens the performance of double hybrid DFAs. The improvements from higher rungs are also visually obvious in the changes in shape seen in the violin plots. The density of the data narrows dramatically, the interquartile ranges get smaller, and the medians get closer to zero. The ranges of deviations (the length of each plot) for the dispersion corrected functionals also decrease, but the improvements are less than 1 kcal/mol between the top three rungs. The meta-GGA/NGAs have the smallest

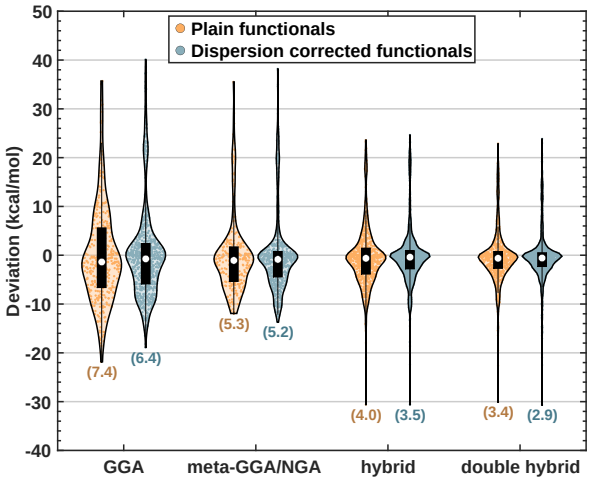


Figure 4: Violin plots by rungs of Jacob's Ladder with and without dispersion corrections, including all functionals except the -3c composite methods. The overall shape of each plot represents the density of the data across the range of deviation values, while the internal black boxes and white dots represent the interquartile range and median value, respectively. Mean MADs for each category are given in kcal/mol.

range for the plain functionals, and this would also be the case for the dispersion corrected functionals if not for r2SCAN-3c, which has the worst range in the third rung. While the maximum positive deviations for the hybrids and double hybrids are smaller than for the lower rungs, some systems are particularly sensitive to the amount of Fock exchange or MP2 correlation, leading to some larger negative deviations. Looking at the influence of the dispersion corrections, we see reductions in the MADs and MAPDs for almost all functionals when they are applied. The addition of a dispersion correction makes the most difference to GGA functionals, with the largest improvement seen for OLYP-D3(BJ)—the dispersion correction reduces the MAD by 2.6 kcal/mol, and the MAPD by 56.5 percentage points—while double hybrids receive the smallest benefit. The Minnesota functionals are also less sensitive to the addition of dispersion corrections for equilibrium structures, but the D4 correction generally improves their results slightly. The difference in the D3 and D4 MADs for each functional are shown in figure 5, with positive differences meaning the D4 MAD is lower than the D3 MAD, and vice versa for the negative differences. D3 gives better results than D4 for all GGA functionals, while for the meta-GGA/NGAs and hybrids, there is an almost equal number of functionals either side. The differences are negligible for all the double hybrids except B2PLYP and PBE0-DH, and the D3 and D4 MADs are also the same to one decimal place for PW91, r2SCAN, M06 and M062X. While D4 is clearly better than D3 for B97M, ωB97M and ωB97X, neither beat the original -V versions of these three functionals. For all DFAs except revTPSS, both D3 and D4 give MADs that are lower than or the same as the MAD of the plain functional, so either is a better choice than using no dispersion correction at all.

Given that this benchmark set is inspired by our mostly organic enzyme benchmark set[18] and the organometallic benchmark sets MOR41,[48] MOBH35[49] and ROST61,[50] we briefly compare our top-performing functionals to those. For the GGAs, OLYP-D3(BJ) has the lowest MAD, which was also the case for the organic enzymes. revPBE-D3(BJ) has the second lowest MAD and lowest MAPD, while its D4 variant was consistently among the top performing GGAs for the organometallic benchmark sets. MN15-L-D3(0) is the best third-rung (meta-GGA/NGA) functional based on MAD, as it was for MOBH35,

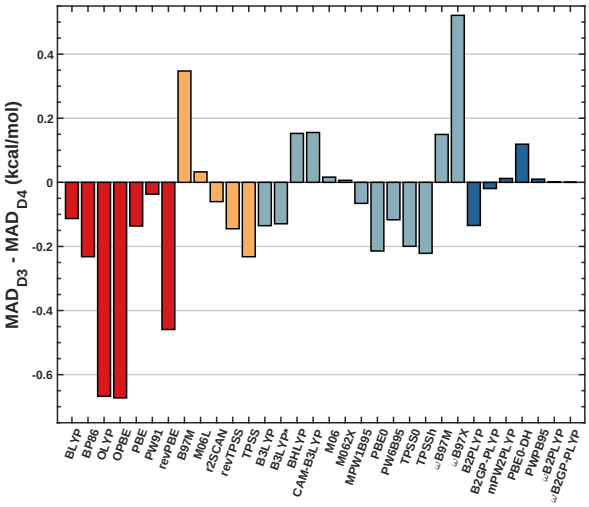


Figure 5: Differences in functional MADs between D3 and D4. A negative difference signifies that the D3 results are better than D4, and vice versa for a positive difference. GGA functionals are shown in red, meta-GGAs in orange, hybrids in light blue and double hybrids in dark blue.

while on MAPDs B97M-V and its -D3/-D4 variants are the strongest. Those functionals and the related hybrids ωB97M-V and ωB97X-V, which were the best hybrids in this study, are consistently good performers in the previously mentioned studies as well as many other benchmark sets for main group[6,8] and transition metal[42,51,177] chemistry. Finally, the best double hybrids are SOS0-PBE0-2-D3(BJ), revDOD-PBEP86-D4, ωB2PLYP-D4 and PWPB95-D4. SOS0-PBE0-2-D3(BJ) was the best double hybrid for the organic enzymes, while PWPB95-D3(BJ) was the best for ROST61 and MOR41. revDOD-PBEP86-D4 was second best for MOBH35, while the D3(BJ) version was fourth for ROST61. The good performance of ωB2PLYP-D4 on these ground state systems is interesting, as it was originally developed for excited state time-dependent DFT, but the plain functional has already been shown to perform well in predictions of UV-Vis spectra for copper complexes.[178]

Taking a closer look at the hybrids and double hybrids, we present violin plots of the best variant of selected functionals in figs. 6 and 7, with each point colored by reaction. Across the hybrids, the majority of deviations lie between −10 and 5 kcal/mol. The largest positive deviation for each functional is found for either CDO 5RE1 or PcrAB RE2, and in most cases these are significantly higher than the next positive deviation. The reference REs for these two steps, however, are −43.64 and −48.62 kcal/mol, so the average percentage deviations (PDs) for these steps across all hybrid functionals are only 39.7% and 35.4%, respectively, and thus they are not seen as outliers in the plots of PDs. In general, it appears that higher amounts of Fock exchange lead to lower REs for CDO, with TPSSh and B3LYP* (10% and 15%, respectively) having the highest positive deviations for 5RE1 of the functionals shown, and BHLYP and M062X (50% and 54%, respectively) both severely underestimating 3RE4. The PDs of CDO 3RE1 show the same trend with these functionals, with the sign of the PD changing from positive to negative with the larger amounts of Fock exchange. Looking at the spread of the deviations, ωB97M-V and ωB97X-V have very tight distributions—ωB97X-V has the smallest range of the hybrids, and this range is almost halved when CDO 5RE1 and PcrAB RE2 are ignored. The ranges of PDs for these two functionals are also impressive, with no significant tails in either direction. Despite B3LYP still being regularly used for mechanistic studies of metalloenzymes,[62,66,152,160,163,179–185] often uncorrected or with the older DFT-

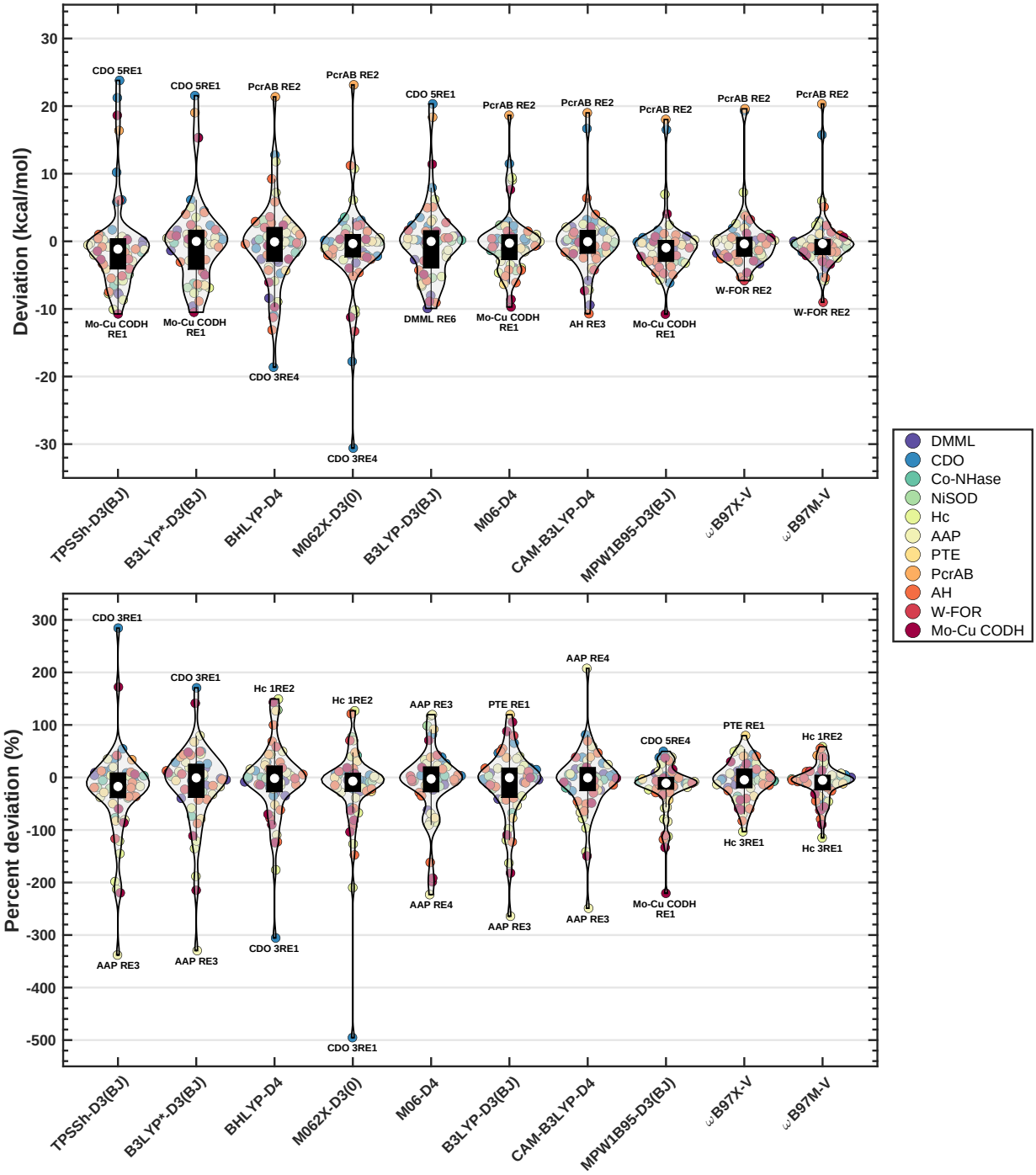


Figure 6: Violin plots of deviations (top) and percent deviations (bottom) of selected hybrid functionals. The functionals are presented in order of decreasing MAD and MAPD, respectively. The overall shape represents the density of the data along the range of (percent) deviation values, while the internal black boxes and white dots represent the interquartile range and median value, respectively. Individual points are colored by reaction, and the maxima and minima are labeled by reaction step.

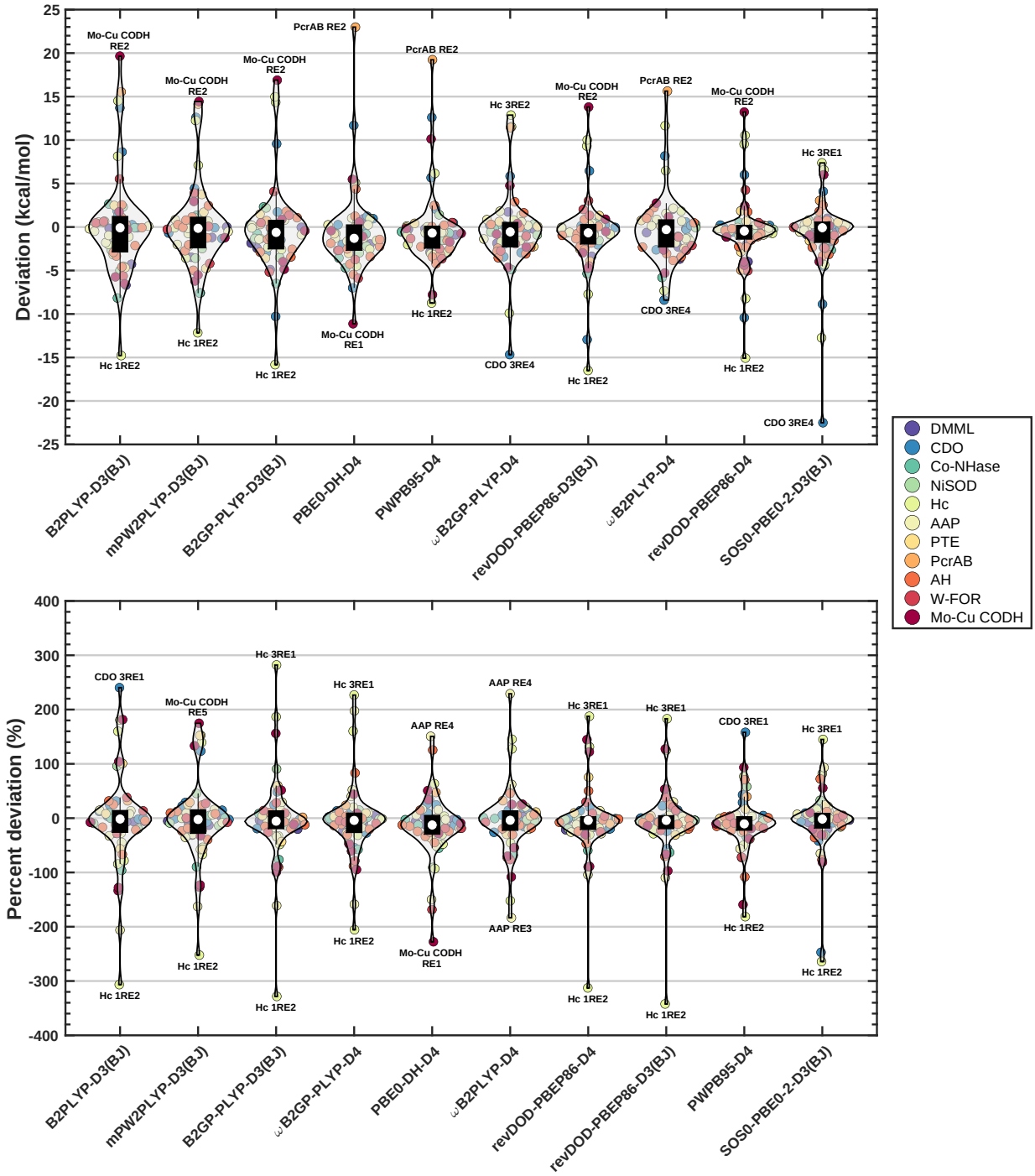


Figure 7: Violin plots of deviations (top) and percent deviations (bottom) of selected double-hybrid functionals. The functionals are presented in order of decreasing MAD and MAPD, respectively. The overall shape represents the density of the data along the range of (percent) deviation values, while the internal black boxes and white dots represent the interquartile range and median value, respectively. Individual points are colored by reaction, and the maxima and minima are labeled by reaction step.

D2[186] or -D3(0) dispersion corrections, it is not a strong performer here. Considering just the best variant of each functional, B3LYP-D3(BJ) is 10th out of 15 hybrids on MADs, and 9th on MAPDs. B3LYP* is sometimes recommended over B3LYP for metalloenzyme energetics,[152,185] but here it is even worse—B3LYP*-D3(BJ) is 13th on both MADs and MAPDs. On the other hand, CAM-B3LYP is slightly better than B3LYP, suggesting that range separation may be more effective at improving the results than changing the global amount of Fock exchange. Interestingly, CAM-B3LYP performs best with D4, instead of the D3(BJ) correction preferred for B3LYP. M06 is also widely popular, and performs similarly to B3LYP on MAPDs but much better on MADs. M06 and M062X were developed simultaneously, but the fitting set for M06 contained an additional database of transition metal bond energies,[127] and therefore should do better for organometallic chemistry. The lower amount of Fock exchange in M06 does improve the treatment of CDO, but the results are similar for the remaining systems, and the interquartile range is actually slightly smaller for M062X for both the deviations and PDs.

For the double hybrids, the deviations and PDs with the largest magnitudes mostly come from the two copper-dependent enyzmes Hc and Mo-Cu CODH, because MP2 is particularly problematic for copper complexes, as mentioned earlier. It massively overstabilizes the peroxo geometry of the Cu_2O_2 core in the singlet state,[41] leading to the large negative deviations and PDs seen for Hc 1RE2 which corresponds to the formation of that configuration—the PD of this step for MP2 is -652.9%, while the average PD across the double hybrids is -284.9%. On the other hand, the triplet REs of Hc are almost always overestimated, as is the second RE of Mo-Cu CODH, leading to large positive deviations across almost all functionals. We, thus, strongly discourage the use of double-hybrid functionals for enzymes containing copper. Conversely, higher amounts of MP2 correlation improve the deviations for PcrAB RE2, with the low-MP2 double hybrids PBE0-DH-D4 (12.5%) and PWPB95-D4 (26.9% opposite spin correlation only) having the highest deviations for this step. We again see some large negative deviations for CDO with higher-Fock exchange functionals, particularly SOS0-PBE0-2-D3(BJ) (79.4% Fock exchange). The functionals with the best ranges shown, ωB2PLYP-D4 and mPW2PLYP-D3(BJ), are therefore ones that have relatively low fractions of (short-range) Fock exchange, and around 25% MP2 correlation. SOS0-PBE0-2-D3(BJ), which has the lowest MAD and MAPD, has long tails particularly below the x-axis due to a few reaction steps, but the majority of the deviations and PDs are tightly clustered around 0 kcal/mol. revDOD-PBEP86-D4 is a robust choice, with a slightly smaller range of deviations than SOS0-PBE0-2-D3(BJ) and narrower distribution in the center than ωB2PLYP-D4. Particularly on PDs, PWPB95-D4 is also a reliable recommendation—it has the smallest range of PDs and second lowest MAPD. While PWPB95 was exclusively fitted against a small number of main-group molecules, its amount of Fock exchange was kept low for a double hybrid (50%) to also perform better for transition metals,[141] which helps explain its strength for these metalloenzyme models and other organometallic benchmark sets.[48,50] Another factor is that PWPB95,

revDOD-PBEP86-D4 and SOS0-PBE0-2 are all spin-opposite-scaled (SOS) functionals,[187] containing only opposite spin contributions to the MP2 energy. Many previous studies have shown that SCS-(spin-component-scaled[188]) and SOS-MP2 are better than standard MP2 for transition metals,[33,34,47–49,189] and here the SOS ones are particularly strong. Indeed, we see that the two versions of revDOD-PBEP86 are better than the two versions of revDSD-PBEP86. The D4 version of revDOD-PBEP86 is better on deviations than the D3(BJ) version, with a lower MAD, smaller range of deviations, and a narrower interquartile range, but revDOD-PBEP86-D3(BJ) has a slightly lower MAPD and the violin plots of PDs look very similar for the two functionals. Again, range-separated functionals give better results than their global counterparts. ωB2PLYP is significantly better than B2PLYP, and the same is seen for (ω)B2GP-PLYP. We note little impact of dispersion corrections for the two range-separated double hybrids, as reported in refs. 11 and 190.

Given that the identity of the transition metal in each enzyme appears to lead to differing functional performance, we have calculated separate MAPDs for each, grouping together molybdenum and tungsten as these exhibit similar reactivity. The best hybrid and double-hybrid functionals for each subset are presented in table 6. ωB97M-V is one of the best hybrids for the enzymes containing manganese, cobalt, nickel and zinc, while ωB97M-D3(BJ) is the best for Mo-Cu CODH. ωB97X-V and its D3/D4 variants are also in the top three for most of the transition metal groupings. The molybdenum/tungsten subset is the only one which does not have a variant of either of these functionals in the top three hybrids, but its best hybrid is PW6B95, which was the third best hybrid overall when combined with the D3(BJ) dispersion correction, and ωB97X-V comes in fifth with an MAPD of 26.4%. Similarly for the double hybrids, all subsets other than iron have a variant of SOS0-PBE0-2 and/or PWPB95 in their top three, with revDOD-PBEP86-D4 second best for iron. revDOD-PBEP86-D4 is also strong for nickel and the molybdenum/tungsten enzymes, while the D3(BJ) version is only in the top three for the zinc enzymes, despite being slightly better overall. We still see that SCS/SOS double hybrids are the best, with the only standard double hybrids in the top three for any transition metal

Table 6: Top three hybrid and double-hybrid functionals for each transition metal based on MAPDs (percentages show in parentheses). Only the best functional-dispersion correction combination is listed.

Best hybrids	Best double hybrids
Overall	
ωB97M-V (24.5)	SOS0-PBE0-2-D3(BJ) (31.8)
ωB97X-V (27.0)	PWPB95-D4 (32.8)
PW6B95-D3(BJ) (31.4)	revDOD-PBEP86-D3(BJ) (34.3)
Mn (DMML)	
ωB97X-D3(BJ) (6.5)	PWPB95-D3(BJ) (6.3)
ωB97M-V (7.5)	SOS0-PBE0-2-D3(BJ) (6.6)
PBE0-D3(BJ) (7.5)	PBE0-DH-D3(BJ) (7.3)
Fe (CDO)	
ωB97X-V (14.1)	revDSD-PBEP86-D4 (7.7)
TPSS0 (14.9)	revDOD-PBEP86-D4 (10.0)
PBE0 (15.4)	ωB2PLYP (12.9)
Co (Co-NHase)	
BHLYP-D4 (2.6)	SOS0-PBE0-2-D3(BJ) (12.2)
ωB97M-V (3.4)	PWPB95-D4 (12.9)
M06-D3(0) (4.1)	DOD-SCAN-D4 (16.7)
Ni (NiSOD)	
ωB97M-V (4.1)	ωB2PLYP-D3(BJ) (4.1)
CAM-B3LYP-D4 (6.3)	ωB2GP-PLYP (6.8)
MPW1B95-D3(BJ) (10.3)	revDOD-PBEP86-D4 (7.6)
Cu (Hc)	
MPW1B95 (47.4)	PBE0-DH (44.5)
ωB97X-D3(BJ) (49.9)	PWPB95 (63.1)
PBEh-3c (50.1)	ωB2PLYP-D4 (80.2)
Zn (AAP, PTE)	
ωB97M-V (14.9)	PWPB95-D4 (23.6)
M062X-D4 (19.9)	SOS0-PBE0-2-D3(BJ) (24.5)
ωB97X-D4 (26.8)	revDOD-PBEP86-D3(BJ) (26.1)
Mo/W (PcrAB, AH, W-FOR)	
PW6B95 (22.9)	mPW2PLYP-D3(BJ) (16.6)
TPSS0-D3(BJ) (24.8)	revDOD-PBEP86-D4 (16.6)
M06 (24.9)	SOS0-PBE0-2 (18.3)
Mo and Cu (Mo-Cu CODH)	
ωB97M-D3(BJ) (36.5)	SOS0-PBE0-2-D3(BJ) (39.6)
ωB97X-D4 (38.5)	ωB2GP-PLYP (51.3)
MN15 (45.3)	ωB2PLYP (56.2)

being PBE0-DH, mPW2PLYP, ωB2PLYP and ωB2GP-PLYP. Interestingly, despite MP2 giving bad results for copper, PBE0-DH is still slightly better than the best hybrids for Hc, but the other two double hybrids are significantly worse. If Mo-Cu CODH is grouped with Hc as an extended copper subset, the hybrid recommendations are exactly the same as for Mo-Cu CODH alone, while the top three double hybrids are PBE0-DH-D4, PWPB95-D4 and SOS0-PBE0-2-D3(BJ), a combination of the recommendations for the two systems. Similarly, the best double hybrids when Mo-Cu CODH is added into the molybdenum/tungsten subset are a balance of the two original lists: SOS0-PBE0-2-D3(BJ), followed by revDOD-PBEP86-D3(BJ) and ωB2GP-PLYP-D3(BJ). The top three hybrids are ωB97X-V, CAM-B3LYP-D4—an unexpected result, given that it was not in either original top three list—and PW6B95-D3(BJ). For the iron, copper, molybdenum and tungsten enzymes, some plain functionals are recommended over their dispersion corrected variants. When the relative dispersion energy for a reaction step has the same sign as the plain functional's deviation, the deviation of the corrected functional will have a larger magnitude, and this occurs more commonly in these model systems. Despite this, we still recommend the use of dispersion corrections in all calculations because they describe the physical behavior better, and the error cancellation that leads to lower MAPDs and MADs for uncorrected functionals should not be relied upon to get good results, as has been argued before.[5,191,192]

Finally, we look at the choice of basis set, as large transition metal complexes and open shell systems can be quite com-

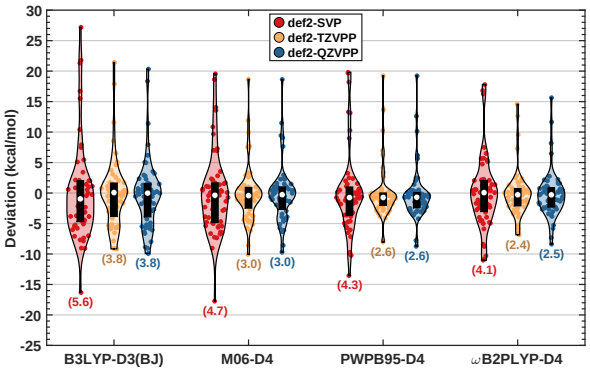


Figure 8: Violin plots of selected functionals with double-, triple- and quadruple-ζ basis sets. The overall shape represents the density of the data across the range of deviation values, while the internal black boxes and white dots represent the interquartile range and median value, respectively. Each plot is labeled with the MAD (kcal/mol) for that level of theory.

putationally demanding, so faster methods that maintain accuracy are desirable for computational studies of metalloenzymes. We do not recommend cutting costs by using GGA or meta-GGA/NGA functionals, as the results in fig. 4 and table 5 show that these are significantly less accurate. We specifically note that MN15-L-D3(0), the best third-rung functional on MADs, was the slowest converging functional overall for most structures, so it provided no cost benefit over a good hybrid or even double-hybrid functional. However this may be due to the LibXC[146] implementation evoked by ORCA, as MN15 and MPW1B95 were also slow. We also cannot recommend the -3c composite methods, as although they are fast, they perform poorly—PBEh-3c and r2SCAN-3c are some of the worst in their respective rungs. The violin plots in fig. 8 show results for B3LYP-D3(BJ), M06-D4, PWPB95-

D4 and ωB2PLYP-D4 (the best dispersion combination for each functional) with three different basis sets. Between def2-SVP and def2-TZVPP there is a significant reduction in the error range and distribution of the data, confirming that double-ζ basis sets are not adequate for calculating BHs and REs. For ωB2PLYP-D4, def2-TZVPP is slightly better than def2-QZVPP, but for the other functionals there is almost no difference between the two. PWPB95 in particular is known to have a weaker basis set dependence than other double hybrids, due to its relatively low fraction of SOS-MP2 correlation.[141] Thus, we can say that the DFT energies are adequately converged at the triple-ζ level, and it is not worth the extra cost to go up to a quadruple-ζ basis set for calculating REs and BHs in such systems.

Combining the basis set test with the functional results, our best overall recommendations are ωB97M-V/def2-TZVPP or ωB97X-V/def2-TZVPP, which are very reliable and cost effective when calculating energetics of many different types of metalloenzymes. We also recommend the double-hybrid functionals SOS0-PBE0-2-D3(BJ) and PWPB95-D4, except when studying copper, cobalt and zinc enzymes, where the results will be adversely affected by the inclusion of MP2 correlation. We do note, however, that functionals that are good for energetics are not always the best for geometries.[16] A test of small organometallic complexes have shown that ωB97M-D3(BJ) and ωB97M-D4 perform similarly to B3LYP-D3(BJ) for geometry optimizations, while B97M-V is slightly better,[10] but without further testing we can only definitively recommend these functionals for single point energy calculations.

5 Summary and conclusions

Here we have presented and used our new set, MME55, for benchmarking metalloenzyme reaction energies and barrier heights. Eleven model reactions were taken from the literature, representing 8 different transition metals, and all geometries were re-optimized at the PBEh-3c level of theory. Some reaction steps described in the original studies were omitted from the MME55 set based on an analysis of the multireference character. Three coupled cluster approaches were tested for the calculation of reference values. While DLPNO-CCSD(T)/TightPNO/CBS(def2-TZVPP/def2-QZVPP) results are preferred, for some larger systems we had to either estimate a full extrapolation of the correlation energy by adding a lower-level triples correction to DLPNO-CCSD/TightPNO/CBS results, or estimate TightPNO thresholds by extrapolating results calculated with looser values of the T_{CutPNO} parameter. The final set contains 16 barrier heights and 39 reaction energies, for a total of 55 data points, and was used to test a range of density functionals across the top four rungs of Jacob's Ladder (GGAs up to double hybrids). SOS0-PBE0-2-D3(BJ) had the lowest MAD of all functionals, while ωB97M-V had the lowest MAPD. ωB97X-V, revDOD-PBEP86-D4, and PWPB95-D4 also performed well. The best double-hybrid functionals are all spin-opposite-scaled approaches, reaffirming previous studies which show that SCS-MP2 and SOS-MP2 are significantly better for transition metals than standard MP2. For copper enzymes, however, any MP2 correlation is problematic and should

be avoided; hybrids also perform better than double hybrid for the cobalt and zinc enzymes in the set. B3LYP, despite its popularity in metalloenzyme modeling, is clearly outperformed by multiple other hybrids, even in its best dispersion-corrected form [B3LYP-D3(BJ))]. Tests of three different basis sets showed that def2-TZVPP gives almost identical results to def2-QZVPP, but double-ζ basis sets are insufficiently small. We therefore recommend ωB97M-V/def2-TZVPP and ωB97X-V/def2-TZVPP for reliable, robust and cost effective results across many different types of enzymes. We hope these recommendations help improve the results of mechanistic studies of enzymes, and help the ongoing efforts to flesh out the field of DFT benchmarking for transition metal chemistry.

Supporting Information Available

The Supporting Information is available free of charge at https://pubs.acs.org/doi/.

FOD plots for all structures (PDF)

System details; multireference diagnostics; statistics, deviations and percentage deviations of assessed methods (XLSX)

Optimized structures (ZIP)

Acknowledgement D. A. Wappett acknowledges an Australian Government Research Training Program Scholarship. We are thankful for the allocation of computing resources by the National Computational Infrastructure (NCI) National Facility within the National Computational Merit Allocation Scheme (Project No. fk5) and The University of Melbourne's Research Computing Services and the Petascale Campus Initiative (Project No. punim0094). This research was additionally supported by the Research Computing Services NCI Access scheme at The University of Melbourne.

References

(1) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev. B* **1964**, *136*, 864–871.

(2) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(3) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

(4) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

(5) 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 iInteractions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(6) 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.

(7) 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.

(8) Najibi, A.; Goerigk, L. The Nonlocal Kernel in van der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ωB97M-V Approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738.

(9) Goerigk, L.; Mehta, N. A Trip to the Density Functional Theory Zoo: Warnings and Recommendations for the User. *Aust. J. Chem.* **2019**, *72*, 563–573.

(10) Najibi, A.; Goerigk, L. DFT-D4 Counterparts of Leading Meta-Generalized-Gradient Approximation and Hybrid Density Functionals for Energetics and Geometries. *J. Comput. Chem.* **2020**, *41*, 2562–2572.

(11) Najibi, A.; Casanova-Páez, M.; Goerigk, L. Analysis of Recent BLYP- and PBE-Based Range-Separated Double-Hybrid Density Functional Approximations for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Phys. Chem. A* **2021**, *125*, 4026–4035.

(12) Brás, N. F.; Perez, M. A. S.; Fernandes, P. A.; Silva, P. J.; Ramos, M. J. Accuracy of Density Functionals in the Prediction of Electronic Proton Affinities of Amino Acid Side Chains. *J. Chem. Theory Comput.* **2011**, *7*, 3898–3908.

(13) Kříž, K.; Řezáč, J. Benchmarking of Semiempirical Quantum-Mechanical Methods on Systems Relevant to Computer-Aided Drug Design. *J. Chem. Inf. Model.* **2020**, *60*, 1453–1460.

(14) Kromann, J. C.; Christensen, A. S.; Cui, Q.; Jensen, J. H. Towards a Barrier Height Benchmark Set for Biologically Relevant Systems. *PeerJ* **2016**, *4*, e1994.

(15) Paiva, P.; Ramos, M. J.; Fernandes, P. A. Assessing the Validity of DLPNO-CCSD(T) in the Calculation of Activation and Reaction Energies of Ubiquitous Enzymatic Reactions. *J. Comput. Chem.* **2020**, *41*, 2459–2468.

(16) Kellie, J. L.; Wetmore, S. D. Selecting DFT Methods for Use in Optimizations of Enzyme Active Sites: Applications to ONIOM Treatments of DNA Glycosylases. *Can. J. Chem.* **2013**, *91*, 559–572.

(17) Joshi, R. P.; McNaughton, A.; Thomas, D. G.; Henry, C. S.; Canon, S. R.; McCue, L. A.; Kumar, N. Quantum Mechanical Methods Predict Accurate Thermodynamics of Biochemical Reactions. *ACS Omega* **2021**, *6*, 9948–9959.

(18) Wappett, D. A.; Goerigk, L. Toward a Quantum-Chemical Benchmark Set for Enzymatically Cat-

alyzed Reactions: Important Steps and Insights. *J. Phys. Chem. A* **2019**, *123*, 7057–7074.

(19) Sirirak, J.; Lawan, N.; Van der Kamp, M. W.; Harvey, J. N.; Mulholland, A. J. Benchmarking Quantum Mechanical Methods for Calculating Reaction Energies of Reactions Catalyzed by Enzymes. *PeerJ* **2020**, *2*, e8.

(20) Zev, S.; Gupta, P. K.; Pahima, E.; Major, D. T. A Benchmark Study of Quantum Mechanics and Quantum Mechanics-Molecular Mechanics Methods for Carbocation Chemistry. *J. Chem. Theory Comput.* **2022**, *18*, 167–178.

(21) 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.

(22) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

(23) Prasad, V. K.; Pei, Z.; Edelmann, S.; Otero-de-la-Roza, A.; DiLabio, G. A. BH9, a New Comprehensive Benchmark Data Set for Barrier Heights and Reaction Energies: Assessment of Density Functional Approximations and Basis Set Incompleteness Potentials. *J. Chem. Theory Comput.* **2022**, *18*, 151–166.

(24) Lothian, A.; Hare, D. J.; Grimm, R.; Ryan, T. M.; Masters, C. L.; Roberts, B. R. Metalloproteomics: Principles, Challenges and Applications to Neurodegeneration. *Front. Aging Neurosci.* **2013**, *5*, 35.

(25) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. Density Functionals for Inorganometallic and Organometallic Chemistry. *J. Phys. Chem. A* **2005**, *109*, 11127–11143.

(26) 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.

(27) Bühl, M.; Kabrede, H. Geometries of Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2006**, *2*, 1282–1290.

(28) Waller, M. P.; Braun, H.; Hojdis, N.; Bühl, M. Geometries of Second-Row Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2007**, *3*, 2234–2242.

(29) Bühl, M.; Reimann, C.; Pantazis, D. A.; Bredow, T.; Neese, F. Geometries of Third-Row Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2008**, *4*, 1449–1459.

(30) Jiménez-Hoyos, C. A.; Janesko, B. G.; Scuseria, G. E. Evaluation of Range-Separated Hybrid and Other Density Functional

Approaches on Test Sets Relevant for Transition Metal-Based Homogeneous Catalysts. *J. Phys. Chem. A* **2009**, *113*, 11742–11749.

(31) 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* **2012**, *116*, 870–885.

(32) 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.

(33) 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.

(34) Waitt, C.; Ferrara, N. M.; Eshuis, H. Thermochemistry and Geometries for Transition-Metal Chemistry from the Random Phase Approximation. *J. Chem. Theory Comput.* **2016**, *12*, 5350–5360.

(35) 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.

(36) Shee, J.; Rudshteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. On Achieving High Accuracy in Quantum Chemical Calculations of 3 *d* 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.

(37) 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.

(38) 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.

(39) 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.

(40) Kang, R.; Lai, W.; Yao, J.; Shaik, S.; Chen, H. How Accurate Can a Local Coupled Cluster Approach Be in Computing the Activation Energies of Late-Transition-Metal-Catalyzed

Reactions with Au, Pt, and Ir? *J. Chem. Theory Comput.* **2012**, *8*, 3119–3127.

(41) Liakos, D. G.; Neese, F. Interplay of Correlation and Relativistic Effects in Correlated Calculations on Transition-Metal Complexes: The $(Cu_2O_2)^{2+}$ Core Revisited. *J. Chem. Theory Comput.* **2011**, *7*, 1511–1523.

(42) Modrzejewski, M.; Chalasinski, G.; Szczesniak, M. M. Assessment of Newest Meta-GGA Hybrids for Late Transition Metal Reactivity: Fractional Charge and Fractional Spin Perspective. *J. Phys. Chem. C* **2019**, *123*, 8047–8056.

(43) Pandey, K. K.; Patidar, S. K.; Vishwakarma, R. Theoretical Insights into M–SO Bonds in Transition Metal-Sulfur Monoxide Complexes [{N(SPMe2)2}2M(SO)] (M = Fe, Ru, Os): Assessment of Density Functionals and Dispersion Interactions. *Polyhedron* **2015**, *101*, 230–238.

(44) Phung, Q. M.; Martín-Fernández, C.; Harvey, J. N.; Feldt, M. Ab Initio Calculations for Spin-Gaps of Non-Heme Iron Complexes. *J. Chem. Theory Comput.* **2019**, *15*, 4297–4304.

(45) Flöser, B. M.; Guo, Y.; Riplinger, C.; Tuczek, F.; Neese, F. Detailed Pair Natural Orbital-Based Coupled Cluster Studies of Spin Crossover Energetics. *J. Chem. Theory Comput.* **2020**, *16*, 2224–2235.

(46) 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.

(47) Husch, T.; Freitag, L.; Reiher, M. Calculation of Ligand Dissociation Energies in Large Transition-Metal Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2456–2468.

(48) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Checinski, M. P. Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions. *J. Chem. Theory Comput.* **2018**, *14*, 2596–2608.

(49) 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.

(50) Maurer, L. R.; Bursch, M.; Grimme, S.; Hansen, A. Assessing Density Functional Theory for Chemically Relevant Open-Shell Transition Metal Reactions. *J. Chem. Theory Comput.* **2021**, *17*, 6134–6151.

(51) 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.

(52) Larsson, E. D.; Dong, G.; Veryazov, V.; Ryde, U.; Hedegård, E. D. Is Density Functional Theory Accurate for Lytic Polysaccharide Monooxygenase Enzymes? *Dalton Trans.* **2020**, *49*, 1501–1512.

(53) Bushnell, E. A. C.; Gauld, J. W. An Assessment of Pure, Hybrid, Meta, and Hybrid-Meta GGA Density Functional Theory Methods for Open-Shell Systems: The Case of the Nonheme Iron Enzyme 8R-LOX. *J. Comput. Chem.* **2013**, *34*, 141–148.

(54) Vancoillie, S.; Zhao, H.; Radoń, M.; Pierloot, K. Performance of CASPT2 and DFT for Relative Spin-State Energetics of Heme Models. *J. Chem. Theory Comput.* **2010**, *6*, 576–582.

(55) Ahlstrand, E.; Spångberg, D.; Hermansson, K.; Friedman, R. Interaction Energies Between Metal Ions (Zn^{2+} and Cd^{2+}) and Biologically Relevant Ligands. *Int. J. Quantum Chem.* **2013**, *113*, 2554–2562.

(56) Boussouf, K.; Boulmene, R.; Prakash, M.; Komiha, N.; Taleb, M.; Mogren Al-Mogren, M.; Hochlaf, M. Characterization of Zn^{q+} –Imidazole (q = 0, 1, 2) Organometallic Complexes: DFT Methods vs. Standard and Explicitly Correlated Post-Hartree–Fock Methods. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14417–14426.

(57) Kornobis, K.; Kumar, N.; Wong, B. M.; Lodowski, P.; Jaworska, M.; Andruniów, T.; Ruud, K.; Kozlowski, P. M. Electronically Excited States of Vitamin B_{12}: Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated ab Initio Methods. *J. Phys. Chem. A* **2011**, *115*, 1280–1292.

(58) Stańczak, A.; Chalupský, J.; RulÃšek, L.; Straka, M. Comprehensive Theoretical View of the $[Cu_2O_2]$ Side-on-Peroxo-/Bis-μ-Oxo Equilibria. *ChemPhysChem* **2022**,

(59) Listyarini, R. V.; Gesto, D. S.; Paiva, P.; Ramos, M. J.; Fernandes, P. A. Benchmark of Density Functionals for the Calculation of the Redox Potential of Fe^{3+}/Fe^{2+} Within Protein Coordination Shells. *Front. Chem.* **2019**, *7*, 391.

(60) Wick, C. R.; Smith, D. M. Modeling the Reactions Catalyzed by Coenzyme B_{12} Dependent Enzymes: Accuracy and Cost-Quality Balance. *J. Phys. Chem. A* **2018**, *122*, 1747–1755.

(61) Borowski, T.; Wójcik, A.; Miłaczewska, A.; Georgiev, V.; Blomberg, M. R. A.; Siegbahn, P. E. M. The Alkenyl Migration Mechanism Catalyzed by Extradiol Dioxygenases: A Hybrid DFT Study. *JBIC, J. Biol. Inorg. Chem.* **2012**, *17*, 881–890.

(62) Da Silva, J. C. S.; Pennifold, R. C. R.; Harvey, J. N.; Rocha, W. R. A Radical Rebound Mechanism for the Methane Oxidation Reaction Promoted by the Dicopper Center of a pMMO Enzyme: A Computational

Perspective. *Dalton Trans.* **2016**, *45*, 2492–2504.

(63) Gupta, P.; Diefenbach, M.; Holthausen, M. C.; Fórster, M. Copper-Mediated Selective Hydroxylation of a Non-activated C-H Bond in Steroids: A DFT Study of Schönecker's Reaction. *Chem. - Eur. J.* **2017**, *23*, 1427–1435.

(64) Kozlowski, P. M.; Kumar, M.; Piecuch, P.; Li, W.; Bauman, N. P.; Hansen, J. A.; Lodowski, P.; Jaworska, M. The Cobalt–Methyl Bond Dissociation in Methylcobalamin: New Benchmark Analysis Based on Density Functional Theory and Completely Renormalized Coupled-Cluster Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 1870–1894.

(65) Kumar, D.; Thiel, W.; de Visser, S. P. Theoretical Study on the Mechanism of the Oxygen Activation Process in Cysteine Dioxygenase Enzymes. *J. Am. Chem. Soc.* **2011**, *133*, 3869–3882.

(66) Xue, J.; Lu, J.; Lai, W. Mechanistic Insights into a Non-Heme 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme: Selectivity of Ethylene-Formation versus L-Arg Hydroxylation. *Phys. Chem. Chem. Phys.* **2019**, *21*, 9957–9968.

(67) Saito, T.; Thiel, W. Quantum Mechanics/Molecular Mechanics Study of Oxygen Binding in Hemocyanin. *J. Phys. Chem. B* **2014**, *118*, 5034–5043.

(68) Wappett, D. A.; Goerigk, L. A Guide to Benchmarking Enzymatically Catalysed Reactions: The Importance of Accurate Reference Energies and the Chemical Environment. *Theor. Chem. Acc.* **2021**, *140*, 68.

(69) Karton, A.; Goerigk, L. Accurate Reaction Barrier Heights of Pericyclic Reactions: Surprisingly Large Deviations for the CBS-QB3 Composite Method and Their Consequences in DFT Benchmark Studies. *J. Comput. Chem.* **2015**, *36*, 622–632.

(70) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(71) Goerigk, L.; Reimers, J. R. Efficient Methods for the Quantum Chemical Treatment of Protein Structures: The Effects of London-Dispersion and Basis-Set Incompleteness on Peptide and Water-Cluster Geometries. *J. Chem. Theory Comput.* **2013**, *9*, 3240–3251.

(72) Goerigk, L.; Collyer, C. A.; Reimers, J. R. Recommending Hartree–Fock Theory with London-Dispersion and Basis-Set-Superposition Corrections for the Optimization or Quantum Refinement of Protein Structures. *J. Phys. Chem. B* **2014**, *118*, 14612–14626.

(73) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Inclusion of Dispersion Effects Significantly Improves Accuracy of Calculated Reaction Barriers

for Cytochrome P450 Catalyzed Reactions. *J. Phys. Chem. Lett.* **2010**, *1*, 3232–3237.

(74) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *J. Chem. Theory Comput.* **2012**, *8*, 4637–4645.

(75) Chen, S.-L.; Blomberg, M. R. A.; Siegbahn, P. E. M. How Is a Co-Methyl Intermediate Formed in the Reaction of Cobalamin-Dependent Methionine Synthase? Theoretical Evidence for a Two-Step Methyl Cation Transfer Mechanism. *J. Phys. Chem. B* **2011**, *115*, 4066–4077.

(76) Siegbahn, P. E. M.; Blomberg, M. R. A.; Chen, S.-L. Significant van Der Waals Effects in Transition Metal Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 2040–2044.

(77) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.

(78) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*.

(79) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152*, 224108.

(80) 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.

(81) Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuß, H. Energy-Adjusted *Ab Initio* Pseudopotentials for the Second and Third Row Transition Elements. *Theor. Chim. Acta* **1990**, *77*, 123–141.

(82) Vahtras, O.; Almlöf, J.; Feyereisen, M. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

(83) Izsák, R.; Neese, F. An Overlap Fitted Chain of Spheres Exchange Method. *J. Chem. Phys.* **2011**, *135*, 144105.

(84) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of Approximate Integrals in Ab Initio Theory. An Application in MP2 Energy Calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(85) Weigend, F. Accurate Coulomb-fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.

(86) 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.

(87) 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.

(88) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(89) 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.

(90) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(91) Kruse, H.; Grimme, S. A Geometrical Correction for the Inter- and Intra-Molecular Basis Set Superposition Error in Hartree-Fock and Density Functional Theory Calculations for Large Systems. *J. Chem. Phys.* **2012**, *136*, 154101.

(92) Grimme, S.; Hansen, A. A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angew. Chem. Int. Ed.* **2015**, *54*, 12308–12313.

(93) 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.

(94) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(95) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew-Burke-Ernzerhof Exchange-Correlation Functional. *J. Chem. Phys.* **1999**, *110*, 5029–5036.

(96) 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.

(97) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse Maps—A Systematic Infrastructure for Reduced-Scaling Electronic Structure Methods. II. Linear Scaling Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Phys.* **2016**, *144*, 024109.

(98) 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.

(99) Karton, A.; Martin, J. M. L. Comment on: "Estimating the Hartree-Fock Limit from Finite Basis Set Calculations" [Jensen F (2005) Theor Chem Acc 113:267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.

(100) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N_2, and H_2O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(101) 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.

(102) 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.

(103) Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(104) 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* **1988**, *37*, 785–789.

(105) 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*, 200–206.

(106) 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.

(107) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.

(108) Perdew, J. P. Erratum: Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *34*, 7406–7406.

(109) Handy, N. C.; Cohen, A. J. Left-Right Correlation Energy. *Mol. Phys.* **2001**, *99*, 403–412.

(110) Goerigk, L.; Grimme, S. A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.

(111) Perdew, J. P. *in Proceedings of the 21st Annual International Symposium on the Electronic Structure of Solids (P. Ziesche and H. Eschrig, Eds.)*; Akademie Verlag: Berlin, 1991; p 11.

(112) Reimers, J. R.; Panduwinata, D.; Visser, J.; Chin, Y.; Tang, C.; Goerigk, L.; Ford, M. J.; Sintic, M.; Sum, T.-J.; Coenen, M. J. J.; Hendriksen, B. L. M.; Elemans, J. A. A. W.; Hush, N. S.; Crossley, M. J. A Priori Calculations of the Free Energy of Formation from Solution of Polymorphic Self-Assembled Monolayers. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*.

(113) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple". *Phys. Rev. Lett.* **1998**, *80*, 890–890.

(114) 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.

(115) 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.

(116) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn-Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12*, 1280–1293.

(117) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r^{2}SCAN Meta-Generalized Gradient Approximation. *J. Phys. Chem. Lett.* **2020**, *11*, 8208–8215.

(118) Ehlert, S.; Huniar, U.; Ning, J.; Furness, J. W.; Sun, J.; Kaplan, A. D.; Perdew, J. P.; Brandenburg, J. G. r^{2}SCAN-D4: Dispersion Corrected Meta-Generalized Gradient Approximation for General Chemical Applications. *J. Chem. Phys.* **2021**, *154*, 061101.

(119) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M. r^{2}SCAN-3c: A "Swiss Army Knife" Composite Electronic-Structure Method. *J. Chem. Phys.* **2021**, *154*, 064103.

(120) 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*, 026403.

(121) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Erratum: Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry [Phys. Rev. Lett. **103** , 026403 (2009)]. *Phys. Rev. Lett.* **2011**, *106*, 179902.

(122) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(123) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(124) Reiher, M.; Salomon, O.; Artur Hess, B. Reparameterization of Hybrid Functionals Based on Energy Differences of States of Different Multiplicity. *Theor. Chem. Acc.* **2001**, *107*, 48–55.

(125) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-

Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(126) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(127) 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.* **2008**, *120*, 215–241.

(128) Zhao, Y.; Truhlar, D. G. Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van der Waals Interactions. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.

(129) 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*, 5656–5667.

(130) 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.

(131) 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.

(132) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(133) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904.

(134) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(135) Karton, A.; Tarnopolsky, A.; Lamere, 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.

(136) Tarnopolsky, A.; Karton, A.; Sertchook, R.; Vuzman, D.; Martin, J. M. L. Double-Hybrid

Functionals for Thermochemical Kinetics. *J. Phys. Chem. A* **2008**, *112*, 3–8.

(137) 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.

(138) 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.

(139) Brémond, E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-DH Model. *J. Chem. Phys.* **2011**, *135*, 024106.

(140) Goerigk, L.; Grimme, S. Double-Hybrid Density Functionals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576–600.

(141) 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.

(142) Alipour, M. Seeking for Spin-Opposite-Scaled Double-Hybrid Models Free of Fitted Parameters. *J. Phys. Chem. A* **2016**, *120*, 3726–3730.

(143) Casanova-Páez, M.; Dardis, M. B.; Goerigk, L. ωB2PLYP and ωB2GPPLYP: The First Two Double-Hybrid Density Functionals with Long-Range Correction Optimized for Excitation Energies. *J. Chem. Theory Comput.* **2019**, *15*, 4735–4744.

(144) Casanova-Páez, M.; Goerigk, L. Time-Dependent Long-Range-Corrected Double-Hybrid Density Functionals with Spin-Component and Spin-Opposite Scaling: A Comprehensive Analysis of Singlet-Singlet and Singlet-Triplet Excitation Energies. *J. Chem. Theory Comput.* **2021**, *17*, 5165–5186.

(145) L. Goerigk, Density Functional Theory Approximations: Development and Evaluation for Electronic Ground and Excited States, PhD thesis, Westfälische Wilhelms-Universität Münster, 2011.

(146) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent Developments in Libxc — A Comprehensive Library of Functionals for Density Functional Theory. *SoftwareX* **2018**, *7*, 1–5.

(147) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147*, 034112.

(148) Li, Y.-Y.; Ye, K.; Siegbahn, P. E. M.; Liao, R.-Z. Mechanism of Water Oxidation Catalyzed by a Mononuclear Manganese Complex. *ChemSusChem* **2017**, *10*, 903–911.

(149) Siig, O. S.; Kepp, K. P. Iron(II) and Iron(III) Spin Crossover: Toward an Optimal Density Functional. *J. Phys. Chem. A* **2018**, *122*, 4208–4217.

(150) Pan, H.; Duan, L.; Liao, R.-Z. Capturing the Role of Phosphate in the Ni-PY5 Catalyzed Water Oxidation. *ChemCatChem* **2020**, *12*, 219–226.

(151) Zhang, Y.-Q.; Chen, J.-Y.; Siegbahn, P. E. M.; Liao, R.-Z. Harnessing Noninnocent Porphyrin Ligand to Circumvent Fe-Hydride Formation in the Selective Fe-Catalyzed CO_2 Reduction in Aqueous Solution. *ACS Catal.* **2020**, *10*, 6332–6345.

(152) Blomberg, M. R. A. The Importance of Exact Exchange—A Methodological Investigation of NO Reduction in Heme–Copper Oxidases. *J. Chem. Phys.* **2021**, *154*, 055103.

(153) Benediktsson, B.; Bjornsson, R. Analysis of the Geometric and Electronic Structure of Spin-Coupled Iron–Sulfur Dimers with Broken-Symmetry DFT: Implications for FeMoco. *J. Chem. Theory Comput.* **2022**, *18*, 1437–1457.

(154) Ghosh, A.; Conradie, J. B12 and F430 Models: Metal- versus Ligand-Centered Redox in Cobalt and Nickel Tetradehydrocorrin Derivatives. *J. Inorg. Biochem.* **2023**, *243*, 112199.

(155) Jongkon, N.; Chotpatiwetchkul, W.; Gleeson, M. P. Probing the Catalytic Mechanism Involved in the Isocitrate Lyase Superfamily: Hybrid Quantum Mechanical/Molecular Mechanical Calculations on 2,3-Dimethylmalate Lyase. *J. Phys. Chem. B* **2015**, *119*, 11473–11484.

(156) Prejanó, M.; Marino, T.; Rizzuto, C.; Madrid Madrid, J. C.; Russo, N.; Toscano, M. Reaction Mechanism of Low-Spin Iron(III)- and Cobalt(III)-Containing Nitrile Hydratases: A Quantum Mechanics Investigation. *Inorg. Chem.* **2017**, *56*, 13390–13400.

(157) Pelmenschikov, V.; Siegbahn, P. E. M. Nickel Superoxide Dismutase Reaction Mechanism Studied by Hybrid Density Functional Methods. *J. Am. Chem. Soc.* **2006**, *128*, 7466–7475.

(158) Chen, S.-L.; Marino, T.; Fang, W.-H.; Russo, N.; Himo, F. Peptide Hydrolysis by the Binuclear Zinc Enzyme Aminopeptidase from *Aeromonas proteolytica* : A Density Functional Theory Study. *J. Phys. Chem. B* **2008**, *112*, 2494–2500.

(159) Chen, S.-L.; Fang, W.-H.; Himo, F. Theoretical Study of the Phosphotriesterase Reaction Mechanism. *J. Phys. Chem. B* **2007**, *111*, 1253–1255.

(160) Sun, S.-Q.; Chen, S.-L. How Does Mo-Dependent Perchlorate Reductase Work in the Decomposition of Oxyanions? *Dalton Trans.* **2019**, *48*, 5683–5691.

(161) Liao, R.-Z.; Yu, J.-G.; Himo, F. Mechanism of Tungsten-Dependent Acetylene Hydratase from Quantum Chemical Calculations. *Proc. Natl.*

Acad. Sci. U. S. A. **2010**, *107*, 22523–22527.

(162) Liao, R.-Z.; Yu, J.-G.; Himo, F. Tungsten-Dependent Formaldehyde Ferredoxin Oxidoreductase: Reaction Mechanism from Quantum Chemical Calculations. *J. Inorg. Biochem.* **2011**, *105*, 927–936.

(163) Xu, K.; Hirao, H. Revisiting the Catalytic Mechanism of Mo-Cu Carbon Monoxide Dehydrogenase Using QM/MM and DFT Calculations. *Phys. Chem. Chem. Phys.* **2018**, *20*, 18938–18948.

(164) 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.

(165) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(166) Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A Simple DFT-based Diagnostic for Nondynamical Correlation. *Theor. Chem. Acc.* **2013**, *132*.

(167) Karton, A.; Daon, S.; Martin, J. M. W4-11: A High-Confidence Benchmark Dataset for Computational Thermochemistry Derived from First-Principles W4 Data. *Chem. Phys. Lett.* **2011**, *510*, 165–178.

(168) Lee, T. J.; Taylor, P. R. A Diagnostic for Determining the Quality of Single-Reference Electron Correlation Methods. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.

(169) Jiang, W.; DeYonker, N. J.; Wilson, A. K. Multireference Character for 3d Transition-Metal-Containing Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 460–468.

(170) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(171) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(172) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally Augmented Karlsruhe Basis Sets. *Theor. Chem. Acc.* **2011**, *128*, 295–305.

(173) 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.

(174) Altun, A.; Riplinger, C.; Neese, F.; Bistoni, G. Exploring the Accuracy Limits of PNO-Based Local Coupled-Cluster Calculations for Transition-Metal Complexes. *J. Chem. Theory Comput.* **2023**, acs.jctc.3c00087.

(175) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conference Proceedings* **2001**, *577*, 1–20.

(176) Hu, L.; Chen, H. Assessment of DFT Methods for Computing Activation Energies of Mo/W-Mediated Reactions. *J. Chem. Theory Comput.* **2015**, *11*, 4601–4614.

(177) Bursch, M.; Hansen, A.; Pracht, P.; Kohn, J. T.; Grimme, S. Theoretical Study on Conformational Energies of Transition Metal Complexes. *Phys. Chem. Chem. Phys.* **2021**, *23*, 287–299.

(178) Curtis, K.; Panthi, D.; Odoh, S. O. Time-Dependent Density Functional Theory Study of Copper(II) Oxo Active Sites for Methane-to-Methanol Conversion in Zeolites. *Inorg. Chem.* **2021**, *60*, 1149–1159.

(179) Chen, H.; Zhou, A.; Sun, D.; Zhao, Y.; Wang, Y. Theoretical Investigation on the Elusive Reaction Mechanism of Spirooxindole Formation Mediated by Cytochrome P450s: A Nascent Feasible Charge-Shift C-O Bond Makes a Difference. *J. Phys. Chem. B* **2021**, *125*, 8419–8430.

(180) Chaturvedi, S. S.; Ramanan, R.; Hu, J.; Hausinger, R. P.; Christov, C. Z. Atomic and Electronic Structure Determinants Distinguish between Ethylene Formation and L -Arginine Hydroxylation Reaction Mechanisms in the Ethylene-Forming Enzyme. *ACS Catal.* **2021**, *11*, 1578–1592.

(181) Coleman, T.; Kirk, A. M.; Chao, R. R.; Podgorski, M. N.; Harbort, J. S.; Churchman, L. R.; Bruning, J. B.; Bernhardt, P. V.; Harmer, J. R.; Krenske, E. H.; De Voss, J. J.; Bell, S. G. Understanding the Mechanistic Requirements for Efficient and Stereoselective Alkene Epoxidation by a Cytochrome P450 Enzyme. *ACS Catal.* **2021**, *11*, 1995–2010.

(182) Zhang, Z.; Smart, T. J.; Choi, H.; Hardy, F.; Lohans, C. T.; Abboud, M. I.; Richardson, M. S. W.; Paton, R. S.; McDonough, M. A.; Schofield, C. J. Structural and Stereoelectronic Insights into Oxygenase-Catalyzed Formation of Ethylene from 2-Oxoglutarate. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 4667–4672.

(183) Summers, T. J.; DeYonker, N. J. QM-cluster Model Study of CO_2 Hydration Mechanisms in Metal-Substituted Human Carbonic Anhydrase II. *Electron. Struct.* **2023**, *5*, 014002.

(184) Siegbahn, P. E. M. Computational Modeling of Redox Enzymes. *FEBS Letters* **2022**, *597*, 38–44.

(185) Siegbahn, P. E. M. A Quantum Chemical Approach for the Mechanisms of Redox-Active Metalloenzymes. *RSC Adv.* **2021**, *11*, 3495–3508.

(186) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(187) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled Opposite-Spin Second Order Møller–Plesset Correlation Energy: An Economical Electronic Structure Method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

(188) Grimme, S. Improved Second-Order Møller–Plesset Perturbation Theory by Separate Scaling of Parallel- and Antiparallel-Spin Pair Correlation Energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

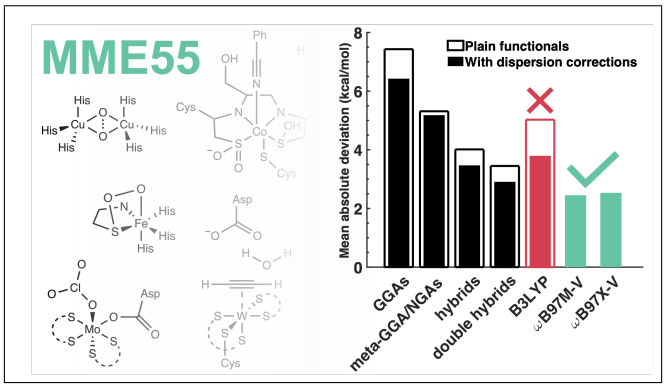
(189) Hyla-Kryspin, I.; Grimme, S. Comprehensive Study of the Thermochemistry of First-Row Transition Metal Compounds by Spin Component Scaled MP2 and MP3 Methods. *Organometallics* **2004**, *23*, 5581–5592.

(190) Hancock, A. C.; Goerigk, L. Noncovalently Bound Excited-State Dimers: a Perspective on Current Time-Dependent Density Functional Theory Approaches Applied to Aromatic Excimer Models. *RSC Adv.* **2022**, *12*, 13014–13034.

(191) Kruse, H.; Goerigk, L.; Grimme, S. Why the Standard B3LYP/6-31G* Model Chemistry Should Not Be Used in DFT Calculations of Molecular Thermochemistry: Understanding and Correcting the Problem. *J. Org. Chem.* **2012**, *77*, 10824–10834.

(192) Mehta, N.; Fellowes, T.; White, J. M.; Goerigk, L. CHAL336 Benchmark Set: How Well Do Quantum-Chemical Methods Describe Chalcogen-Bonding Interactions? *J. Chem. Theory Comput.* **2021**, *17*, 2783–2806.

TOC Graphic


MME55
Plain functionals
With dispersion corrections
Mean absolute deviation (kcal/mol)
0
2
4
6
8
GGAs
meta-GGA/NGAs
hybrids
double hybrids
B3LYP
ωB97M-V
ωB97X-V