Computational Electrochemistry of 3d Transition Metal Complexes

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Hagen Neugebauer
aus
Neuss

Bonn 2023

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Prof. Dr. Stefan Grimme
2. Gutachter:	Prof. Dr. Thomas Bredow
Tag der Promotion:	16.01.2024
Erscheinungsjahr:	2024

Affirmation in Lieu of an Oath

I, Hagen Neugebauer, hereby declare that I am the sole author of this dissertation.

The doctoral thesis I submitted is my own work and was prepared without unauthorized outside assistance. I have not included text passages, graphics or other materials from third parties or my own examination papers without identifying them. Only the sources and resources that I have indicated were used. All verbatim and non-verbatim citations from other works are identified in accordance with the citation rules for academic writing. The thesis that I submitted has not yet been published, or has been published in full or in part other than the locations indicated. The thesis that I submitted has not yet been submitted in any form as part of an examination/qualification course. I prepared the doctoral thesis that I submitted in accordance with the principles of good research practice. I am aware of the significance and criminal consequences of a false affirmation in lieu of an oath.

My statements are true to the best of my knowledge and belief.

Publications

Parts of this thesis have been published in peer-reviewed journals.

1. H. Neugebauer, B. Bädorf, S. Ehlert, A. Hansen, and S. Grimme, *High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods*, J. Comput. Chem. **44** (2023) 2120, DOI: 10.1002/jcc.27185.

2. M. Bursch, H. Neugebauer, S. Ehlert, and S. Grimme, *Dispersion corrected r^2SCAN based global hybrid functionals: r^2SCANh, r^2SCAN0, and r^2SCAN50*, J. Chem. Phys. **156** (2022) 134105, DOI: 10.1063/5.0086040.

3. H. Neugebauer, P. Pinksi, S. Grimme, F. Neese, and M. Bursch, *Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT*, J. Chem. Theory Comput. **19** (2023) 7695, DOI: 10.1021/acs.jctc.3c00896.

4. H. Neugebauer, H. T. Vuong, J. L. Weber, R. A. Friesner, J. Shee, and A. Hansen, *Toward Benchmark-Quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts: A Comparison of CCSD(T) and ph-AFQMC*, J. Chem. Theory Comput. **19** (2023) 6208, DOI: 10.1021/acs.jctc.3c00617.

5. P. Sánchez, B. Goel, H. Neugebauer, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts*, Inorg. Chem. **60** (2021) 17407, DOI: 10.1021/acs.inorgchem.1c03142.

6. B. Goel, H. Neugebauer, A. I. VanderWeide, P. Sánchez, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes*, ACS Catal. **13** (2023) 13650, DOI: 10.1021/acscatal.3c02911.

For the following articles significant contributions have been made.

7. T. Gunasekara, G. P. Abramo, A. Hansen, H. Neugebauer, M. Bursch, S. Grimme, and J. R. Norton, *TEMPO-Mediated Catalysis of the Sterically Hindered Hydrogen Atom Transfer Reaction between (C_5Ph_5)Cr(CO)$_3$H and a Trityl Radical*, J. Am. Chem. Soc. **141** (2019) 1882, DOI: 10.1021/jacs.8b12892.

8. M. Bursch, E. Caldeweyher, A. Hansen, H. Neugebauer, S. Ehlert, and S. Grimme, *Understanding and Quantifying London Dispersion Effects in Organometallic Complexes*, Acc. Chem. Res. **52** (2019) 258, DOI: 10.1021/acs.accounts.8b00505.

9. E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme, *A generally applicable atomic-charge dependent London dispersion correction*, J. Chem. Phys. **150** (2019) 154122, DOI: 10.1063/1.5090222.

10. M. Bursch, H. Neugebauer, and S. Grimme, *Structure Optimisation of Large Transition-Metal Complexes with Extended Tight-Binding Methods*, Angew Chem. Int. Ed. **58** (2019) 11078, DOI: 10.1002/anie.201904021.

11. H. Neugebauer, F. Bohle, M. Bursch, A. Hansen, and S. Grimme, *Benchmark Study of Electrochemical Redox Potentials Calculated with Semiempirical and DFT Methods*, J. Phys. Chem. A **124** (2020) 7166, DOI: 10.1021/acs.jpca.0c05052.

12. H. B. Vibbert, H. Neugebauer, J. R. Norton, A. Hansen, M. Bursch, and S. Grimme, *Hydrogen atom transfer rates from Tp-containing metal-hydrides to trityl radicals*, Can. J. Chem. **99** (2021) 216, DOI: 10.1139/cjc-2020-0392.

13. D. S. Tresp, H. Neugebauer, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Electronic Effects of Aminoindenyl Ligands Coordinated to Manganese: Structures and Properties of a Mn^0 Metalloradical and Bimetallic Mn^{-I}/Mn^{I} Adduct*, Organometallics **41** (2022) 3055, DOI: 10.1021/acs.organomet.2c00463.

Abstract

The topic of this thesis is the computational quantum chemical (QC) description of homogeneous first-row (3d) transition metal (TM) electrocatalysis. This branch of chemistry holds great potential for the use of Earth-abundant 3d TMs in renewable energy concepts. Therefore, routine predictions for the reactivity of 3d TM electrocatalysts are desirable, but due to numerous challenges, they are only possible to a limited extent. First, TM chemistry is diverse and includes many bonding motifs, some of which include challenging electron correlation effects which makes their theoretical description difficult. Also, in 3d TM chemistry, ground states with different electron spin states occur. Finding the spin state lowest in energy, the ground state, can be challenging and requires extensive calculations. Particularly critical for all QC methods is the multireference (MR) character of some 3d TM complexes that hinders the application of standard single-reference methods. These challenges also occur for general 3d TM chemistry but are severe for electrocatalysis since noninnocent ligands, open-shell species, and highly charged molecules are regularly encountered. Furthermore, the countless possibilities for combinations of different metal centers and ligands, investigated for catalyst optimization, require QC methods for different cost and accuracy regimes for efficient multilevel workflows. These range from semiempirical quantum mechanical (SQM) to density functional theory (DFT) up to correlated wave function theory (WFT) methods. The methods should not be limited in applicability to electrocatalysis with 3d TMs but should be as general as possible. This thesis describes the development, benchmarking, and application of such methods for 3d TM electrocatalysis.

In the first Chapter, an introduction to the QC treatment of 3d TM electrocatalysis is given, followed with a brief overview of the different QC methods in the second Chapter. On the first step of QC multilevel workflows are SQM methods, which are the topic of the third Chapter. Here, the inclusion of spin-polarization in the extended tight-binding Hamiltonian (xTB) is elaborated for the calculation of spin splittings, which are crucial for the correct description of TM complexes. The performance of the spin-polarized GFN*n*-xTB methods is evaluated on a newly compiled benchmark set termed TM90S, which employs DFT spin splittings as references. The next step in multilevel workflows is typically DFT, which is the topic of Chapter four. Here, the extension of the non-empirical r^2SCAN density functional approximation (DFA) to the hybrid functional level, resulting in the r^2SCANh, r^2SCAN0, and r^2SCAN50 DFAs, is described. The new DFAs are combined with semiclassical corrections for London dispersion interactions and are extensively benchmarked for thermochemistry and structural features. At the highest DFT level are the double-hybrids (DHs), which are the subject of Chapter five. Their applicability is extended with the domain-based local pair natural orbital (DLPNO) approximation for second-order Møller-Plesset theory (MP2). The resulting DLPNO-DH scheme allows the routine application of DLPNO-DHs to molecules with more than 100 atoms and enables the geometry optimization of large closed-shell molecules. On the highest step of QC

multilevel workflows are the correlated WFT methods. Their application can face severe difficulties in 3d TM electrocatalysis due to MR character, which is the subject of the sixth Chapter. Here, the recognition of MR systems and the calculation of their ionization potential (IPs) is studied for 3d TM electrocatalysts at the highest feasible WFT level. For this purpose, a new benchmark set for electrocatalysts, termed 3dTMV was compiled and coupled cluster calculations with singles, doubles, and perturbative triples (CCSD(T)), and phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC) calculations were conducted. This approach allows cross-checking the so-called gold standard, CCSD(T), with the ph-AFQMC method. A quantitative protocol for the determination of MR character is presented, together with a classification of different electron correlation regimes where CCSD(T) is still applicable.

Chapter seven deals with the application of SQM and DFT methods for the elucidation of electrocatalytic cycles with three-legged piano-stool iron complexes. An efficient workflow is presented for the calculation of Gibbs free energies yielding a free energy map that is used to propose an initial catalytic cycle. The extension of the free energy map to also include kinetics by transition state theory with the newly developed r^2SCAN0-D4 DFA is shown in Chapter eight.

Finally, in the ninth Chapter, the findings of this work are summarized, and their impact on the theoretical description of 3d TM electrocatalysis and 3d TM chemistry in general is evaluated. Novel multilevel workflows can benefit from the methods and findings presented in this work and accelerate the discovery of efficient (electro-)catalysts employing Earth-abundant 3d transition metals.

Contents

CHAPTER 1

Introduction

Homogeneous electrocatalysis with first-row (3d) transition metal (TM) complexes represents a promising approach for chemical energy conversion processes.[14] In catalysis, a catalyst participates in a specific reaction, which results in better product selectivity and a faster reaction rate compared to the plain, uncatalyzed reaction. The catalyst is recovered in the process and not consumed, therefore, small amounts of catalyst can catalyze the reaction of large amounts of reactant. Electrocatalysis is a special case of catalysis that occurs in electrochemical reactions, i.e., near an electrode with an applied external potential. The main advantage of 3d TM electrocatalysts lies in the use of Earth-abundant 3d TMs[15,16] instead of commonly used precious metals, such as palladium or platinum.[17] Considering the growing demand of chemical energy conversion for renewable energy concepts[18], such cheaper and better available alternatives are highly desirable. Applications include CO_2[19–22], N_2[23], and O_2[24] reduction, water splitting,[25,26] and other hydrogen evolution reactions.[27] The strength of the molecular (homogeneous) approach lies in the almost endless possibilities to combine various metal centers with diverse ligand frameworks to tailor TM complexes with the desired properties, such as electrocatalytic activity.[28] Due to the vast amount of combinations, computational methods are required to help with the design and optimization of electrocatalysts. In recent decades, computational simulations have become a vital tool in chemistry that gives a fundamental understanding of the properties and reactivity of molecules.[29–33] Empowered by the rapid developments in computation power and in algorithms,[34,35] computational chemistry based on quantum mechanics, i.e., quantum chemistry, nowadays finds application in catalysis research,[36–39] pharmaceutical drug design,[40] and material science.[28,41,42] Typical applications are to elucidate underlying reaction mechanisms and propose molecules with the desired reactivity and properties.[43] This way, quantum chemistry helps to reduce the amount of experiments conducted, which leads to a reduction of chemical waste and required resources.

The central quantity that is calculated for elucidation of reaction mechanisms is the Gibbs free energy G, which incorporates the electronic energy of a system and its thermostatistical contributions.[44] The change in the free energy, ΔG, tells whether a process occurs spontaneously (negative ΔG) or if it requires an additional driving force (positive ΔG) to occur[44]. With the help of transition state theory, reaction free energy barriers can be calculated, which can be directly related to the reaction kinetics via the Eyring equation.[44,45] Therefore, if the free energies of all reactants, intermediates, products, and transition states of a reaction network are known, the underlying mechanism is also known. The different reaction profiles for reactions with and without a catalyst can also be calculated.

Catalysts actively participate in reactions and reduce the reaction free energy barriers for product formation, resulting in higher reaction rates (See Figure 1.1). Ideal catalysts avoid high and low-energy intermediates and large reaction barriers, resulting in a flat reaction profile. Therefore, the performance of different catalysts can be studied by the calculation of reaction free energies and barriers, and improved catalysts can be proposed based on such calculations.

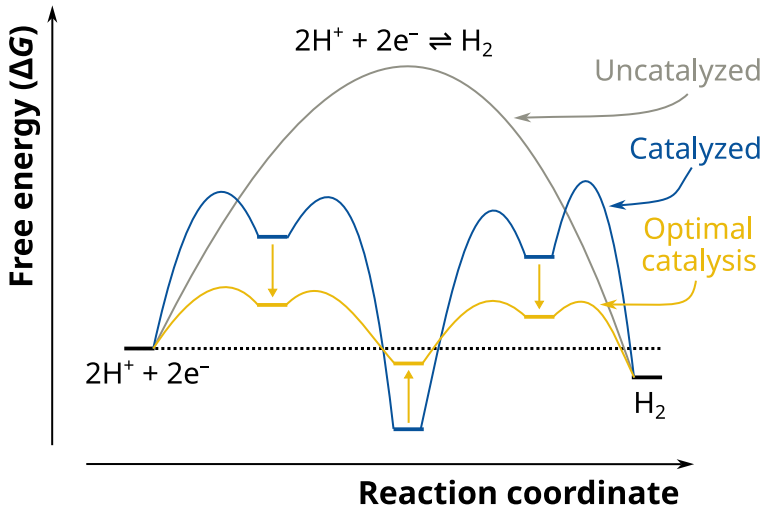

Figure 1.1: Typical energy profiles for a model reaction that is uncatalyzed, catalyzed, and catalyzed with an optimal catalyst, which avoids low- and high-energy intermediates and large barriers. Adapted from ref. [38].

Calculating accurate free energies remains a central challenge for quantum chemistry due to the many different contributions required.[31] Accurate molecular structures, gas phase electronic energies, solvation free energies, and thermostatistical corrections including the zero-point-vibrational energy have to be calculated.[44] The electronic structure and the corresponding potential energy surface (PES) are the cornerstones of these contributions. In principle, they are obtained by solving the time-independent electronic Schrödinger equation, but in practice, it can only be solved analytically for one-electron systems, and numerical approximations have to be used for systems containing multiple electrons. These approximations have to compromise between a targeted level of accuracy and computational demands, which limit the system size to which the approximations can be applied. Additionally, some approximations employ empirical parameters derived from physical or chemical insight, which can restrict their generality to specific applications. Central to attempts at solving many-electron problems is the Hartree-Fock (HF) approximation, in which the interaction between electrons is approximated as the interaction of each electron with the averaged field of all other electrons.[46] In HF, the wave function is represented by a single Slater determinant (SD), which is antisymmetric concerning the permutation of electrons and thus fulfills the Pauli Principle.[44,46] Typically, SDs consist of molecular orbitals (MOs) expanded in atomic-centered orbitals (AOs), the composition of which is determined by a self-consistent field (SCF) procedure[44]. HF is the simplest wave function theory (WFT) method, and, since it employs a mean-field ansatz (only uses one SD), it does not account for electron correlation effects. The electron correlation is roughly divided into two categories, namely dynamic and static correlation (also called nondynamic or strong correlation), where the first originates from many small contributions of multiple SDs, and the latter

from large contributions of few SDs.[44,47] Some authors introduce even further classifications for electron correlation, which will not be done here.[47] Usually, HF constitutes the first step towards more accurate approximations, so-called post-HF or correlated WFT methods, which incorporate multiple SDs and therefore electron correlation. The most accurate of the WFT methods is the full configuration interaction (FCI) procedure that yields the exact solution of the electronic Schrödinger equation for a fixed one-particle basis set.[48] The problem with this method is that its computational demands grow factorial with the system size, which severely limits its application to systems containing at maximum a few atoms. Developing modern approximations to the FCI approach is an active field of research[49] where quantum Monte Carlo (QMC) methods, such as the phaseless auxiliary-field QMC (ph-AFQMC) method, have proven to be accurate and robust over a wide range of applications.[50–55] The most prominent of the wave function methods are the coupled cluster (CC)[56] methods, where CC with single, double, and perturbative triple excitations (CCSD(T)) is the so-called gold standard of quantum chemistry.[57] For reactions of organic systems, CCSD(T) typically yields the "chemical accuracy" of 1 kcal/mol.[58] The downside of CCSD(T) is that its computational costs scale with $O(N^7)$ for the system size, which restricts the applicability of this method to small molecules. Since even the application to small molecules can be computationally expensive, CCSD(T) is often not directly employed in computational studies, and instead, it is used to generate accurate electronic reference energies for evaluation of lower-level methods in benchmark studies. Modern localized CCSD(T) schemes, such as domain-based local pair natural orbital CCSD(T) (DLPNO-CCSD(T)), can yield similar accuracy at drastically reduced computational costs and allow benchmarking on larger system sizes.[59–62] While CC methods paved the way for the accurate description of molecules by quantum chemical methods, the workhorse of modern quantum chemistry is density functional theory (DFT).[63–67] The excellent cost-to-accuracy ratio of DFT surpasses that of HF and enables the routine investigation of system sizes up to a few hundred atoms. DFT additionally coined many concepts of chemical reactivity, which are known to the general chemistry community.[68] When accurate structures of molecules are required, DFT is the method of choice, since structure optimizations with correlated WFT methods are often not feasible. The computationally cheapest quantum chemical calculations can be done with semiempirical quantum mechanical (SQM) methods[69] such as GFNn-xTB,[70–72] DFTB,[73,74] PMx,[75–77] or OMx.[78] These methods approximate either DFT or HF and apply systematic approximations to the one and two-electron integrals, leading to computational schemes that are several orders of magnitude faster. The extended tight-binding (xTB) methods, which are parameterized for geometries, frequencies, and non-covalent interactions, resulting in the acronym GFNn-xTB ($n = 0, 1, 2$), have received more and more attention over the last years. The xTB methods approximate DFT without the necessity of pair-wise parameters, which allowed their parameterization for all elements up to radon ($Z = 86$). They enable the cheap quantum chemical optimization of geometries and allow for the metadynamics-based search of conformers,[79,80] which revolutionized the way computational chemistry is conducted, as it enables the routine generation of extensive conformer ensembles, which can be refined by DFT.[81] Beyond SQM methods, the domain of quantum mechanics is left, and the domain of Newtonian mechanics is entered. Force field (FF) methods allow structural optimizations of even larger compounds by completely neglecting the electronic structure and are mentioned here only to complete the toolchain of methods for atomistic simulations. Often FF methods are parameterized for special applications, and only a few general force fields covering most of the periodic table are available, such as the UFF[82] or the GFN-FF.[83]

For homogeneous electrocatalysis with 3d TM complexes, calculating accurate free energies is particularly challenging. The challenges that limit accurate predictions are depicted in Figure 1.2 and

are elaborated in the following.

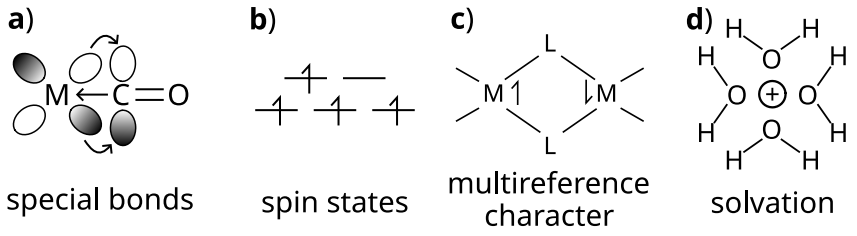

Figure 1.2: Challenges of homogeneous electrocatalysis with 3d TMs.

Compared to plain organic chemistry, organometallic chemistry offers much more diverse chemical bonding situations, which can be challenging to describe (Figure 1.2 **a**). For example, the electron correlation in carbonyl-ligands with σ-donation and π-back-bonding is not sufficiently described in a simple pair-wise additive scheme for electron correlation.[84] Additionally, it is difficult to achieve generality with a limited amount of empirical parameters in SQM methods for the diverse bonding situations that can occur. Another challenge is that in 3d TM complexes, due to the weak ligand field of 3d TMs,[85] different spin ground-states are possible, and it is not always clear whether the system at hand has a low-spin (LS) or a high-spin (HS) ground state (Figure 1.2 **b**). Calculating the energetic difference between these spin states (spin splittings) is very challenging for electronic structure methods and requires additional calculations, and therefore additional computation time. The third and major obstacle is the potential multireference or multi-configurational character of TM complexes (Figure 1.2 **c**). TM complexes with high symmetry, low coordination number, or magnetically coupled radicals can have degenerate or nearly degenerate orbitals and spin states, which may result in multireference character, also known as static, strong, or nondynamic correlation.[47,86–88] Single-reference methods, such as CCSD(T), DFT, or SQM methods, fail in such cases and yield unreliable results. These systems require treatment by multireference methods, for example, the complete active space SCF (CASSCF) method.[89] The problems with these methods are the inclusion of dynamic correlation effects, and their non-black-box nature since a limited active space has to be selected. The latter often requires expert users with insights into the system at hand. Additionally, it is a priori unknown whether a system requires the complicated multireference treatment. A unified approach for the computation of 3d TM electrocatalysts, or 3d TM complexes in general, with multireference character is yet to be found. The fourth major obstacle is the inclusion of solvation effects (Figure 1.2 **d**). In electrocatalysis, the catalysts are reduced by applying an external potential in an electrochemical step (E) and protonated by an acid in a chemical step (C). The free energy changes of the first are described by redox potentials and of the latter by acidity constants (pK_a values). Both can be obtained experimentally but can also be calculated.[38,90] Therefore, charged species necessarily occur in electrochemistry, which result in large solvation free energy contributions, particularly for polar solvents like acetonitrile (MeCN). Solvation effects are often modeled by implicit solvation models, which incorporate solvation effects implicitly as potential in the Hamiltonian. They offer the advantage of low computational costs, as the solvent molecules are not explicitly required in the calculation. However, the available implicit models have problems describing charged species.[91] Additionally, in the experiment, solvent molecules can coordinate to metal centers and directly influence the ligand-field of complexes. Such effects can not be accounted for by implicit solvation models, and explicit solvent molecules have to be added to the calculation. In practice, the calculation of redox potentials can profit from error compensation

by explicitly calculating the reference redox couple, typically the ferrocenium/ferrocene (Fc/Fc^+) couple,[90] while for the computation of pK_a values isodesmic schemes can be employed to reduce errors.[38]

Due to these challenges (**a**-**d**), the prediction of reaction free energies or barriers in 3d TM (electro-)catalysis is only possible to some extent. For systems with large multireference character, and/or solvation interactions that require explicit solvation for several solvent shells with extensive sampling for the degrees of freedom, the modern quantum chemical toolbox is quickly at its limit. To achieve accurate predictions regularly, improvements in electronic structure theory in all cost versus accuracy regimes and in solvation protocols are required. The benchmarking of protocols on gas phase references and solvation free energies separately is required, as this allows tracing down the different sources of error and potentially reveals error compensation effects. This approach is complicated by the fact that only little experimental data in the gas phase or for solvation free energies are available. Therefore, the practical approach is to benchmark gas phase electronic energies against theoretical references, and once these errors are well understood, to combine the electronic structure methods with solvation protocols and thermal corrections to compare to experimental data in solution, e.g., redox potentials or pK_a values. In practice, for the optimization of electrocatalysts, extensive screening of potential candidates, spin states, and conformers is required, and developments in all costs/accuracy regimes are necessary because it is not feasible to use the most accurate method for all candidates due to high computational demands. Instead, so-called multilevel workflows are typically employed, where in initial stages, computationally cheap but less accurate methods are employed for pre-screening steps, and with each step, candidates are sorted out, and a more accurate, but also more expensive, method is used for the subsequent step. The topic of this thesis is the development, assessment, and the application of methods for the accurate calculation of electronic energies in different cost/accuracy regimes for 3d TM electrochemistry, for the use in such multilevel workflows.

The introduction continues in Chapter 2 with an overview of the theoretical basis for the calculation of free energies and the quantum chemistry methods required for this purpose. In computational studies of 3d TM complexes, the first step is the determination of the spin multiplicity. To perform this task with the efficient SQM methods, the popular GFN*n*-xTB methods were extended by a spin-dependent energy expression in Chapter 3. The nonempirical r^2SCAN functional has been extended to the hybrid DFT level in Chapter 4 and was extensively benchmarked for organic and organometallic thermochemistry. The highest DFT rung, the double-hybrid (DH) functionals, are limited in their application due to their steep $O(N^5)$ scaling with the system size. In Chapter 5, their application has been extended to larger system sizes by combining the DH scheme with the DLPNO-MP2 approximation[92] resulting in the DLPNO-DH scheme. Identification of electron correlation regimes in 3d TM electrocatalysis in which CCSD(T) is applicable has been intensively studied in Chapter 6 by cross-validation with highly accurate localized ph-AFQMC calculations.

Finally, SQM and DFT methods are extensively used in a newly developed workflow to elucidate the hydrogen evolution reaction mechanism of iron-based three legged piano-stool electrocatalysts in Chapter 7. This workflow has been adapted for kinetics by employing the newly developed r^2SCAN0-D4 hybrid functional as shown in Chapter 8. The impact of this thesis on the field of computational chemistry with 3d TM complexes in general and with 3d TM electrocatalysts in particular is summarized in Chapter 9.

CHAPTER 2

Theoretical Background

This chapter presents an overview over the quantum chemical methods employed throughout this thesis and is based on refs. [44, 46, 72, 89]. For an in-depth description the reader is referred to the original literature. Atomic units are used throughout for electronic structure methods.[46]

2.1 Electronic structure theory

2.1.1 Molecular Hamiltonian

The cornerstone of most electronic structure methods in quantum chemistry is the time-independent non-relativistic Schrödinger equation[93] given as

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

where $\hat{H}$ is the molecular Hamiltonian operator, Ψ is the wave function, and E is the energy. The Hamiltonian operator can be separated according to

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \tag{2.2}$$

with the kinetic energy operators of the nuclei $\hat{T}_n$ and electrons $\hat{T}_e$, the nucleus-nucleus $\hat{V}_{nn}$ and the electron-electron $\hat{V}_{ee}$ Coulomb repulsion operators, and the electron-nucleus Coulomb attraction $\hat{V}_{ne}$ operator. Because the lighter electrons are much faster than the heavier nuclei, this expression is typically simplified with the Born-Oppenheimer approximation[94] resulting in

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \tag{2.3}$$

where the kinetic energy of the nuclei is neglected and the nucleus-nucleus repulsion is constant. All calculations conducted in this thesis employ the Born-Oppenheimer approximation. The kinetic energy operator of the electrons is given as

$$\hat{T}_e = -\sum_i^N \frac{1}{2}\nabla_i^2 \tag{2.4}$$

with the index i accounting for the individual electron, N the total number of electrons in the system, and the Laplace operator ∇_i^2 acting on electron i. The electron-nucleus attraction operator is given according to

$$\hat{V}_{ne} = -\sum_{i}^{N}\sum_{A}^{M}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}. \tag{2.5}$$

Here, the index A stands for the individual nucleus and M gives the total number of nuclei in the system. Z_A gives the nuclear charge of nucleus A and $\mathbf{r}_i$ and $\mathbf{R}_A$ are the coordinates of electron i and nucleus A respectively. The electron-electron and nucleus-nucleus repulsion operator are then given as

$$\hat{V}_{ee} = \sum_{i}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.6}$$

and

$$\hat{V}_{nn} = \sum_{A}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}. \tag{2.7}$$

It can be practical to collect the operators of the Hamiltonian according to the number of electron indices

$$\hat{H}_e = \sum_{i}^{N}\hat{h}_i + \sum_{i}^{N}\sum_{j>i}^{N}\hat{g}_{ij} + \hat{V}_{nn} \tag{2.8}$$

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{A}^{M}\frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} \tag{2.9}$$

$$\hat{g}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.10}$$

resulting in the one-electron operator $\hat{h}_i$ describing the motion of electron i in the field of all the nuclei and the two-electron operator $\hat{g}_{ij}$ describing the electron-electron repulsion between electron i and electron j.

2.1.2 Hartree-Fock

In order to calculate an approximate energy of the system the expectation value of the Hamiltonian $\hat{H}_e$ has to be calculated. The expectation value of the electronic Hamiltonian operator of an approximate wave function can be written in the bra-ket notation

$$E = \int \Psi^*(\mathbf{r})\hat{H}_e\Psi(\mathbf{r})d\mathbf{r} = \langle\Psi|\hat{H}_e|\Psi\rangle \tag{2.11}$$

where the integration is carried out over all coordinates. Since electrons are fermions and have a spin of 1/2, their wave function must be antisymmetric with respect to the interchange of any two electron coordinates. The antisymmetry can be realized by building the wave functions from Slater

determinants (SDs). For the general molecular case with N electrons and a single SD, the approximate wave function is given by

$$\Psi \approx \Phi_{\mathrm{SD}}(1,2,\cdots,N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_N(N) \end{vmatrix} \tag{2.12}$$

where $\phi_i(k)$ denotes the i^{th} molecular orbital (MO) occupied by the k^{th} electron $\in \{1,\cdots,N\}$. The MOs may be taken as orthonormal

$$\langle \phi_i | \phi_j \rangle = \int \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

with the Kronecker delta δ_{ij}. The MOs are given as the product of a spatial orbital ψ_i and a spin-function $\sigma_i \in \{\alpha, \beta\}$ as shown in equation 2.14.

$$\phi_i(1) = \sigma_i \psi_i(\mathbf{r}_1) \tag{2.14}$$

The energy expression for a single SD is then given by

$$E = \sum_{i=1}^{N} h_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (J_{ij} - K_{ij}) + V_{NN} \tag{2.15}$$

$$= \sum_{i}^{N} \langle \phi_i | \hat{h}_i | \phi_i \rangle + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\langle \phi_j | \hat{J}_i | \phi_j \rangle - \langle \phi_j | \hat{K}_i | \phi_j \rangle) + V_{NN} \tag{2.16}$$

with the Coloumb ($\hat{J}_i$) and exchange ($\hat{K}_i$) operators defined for convenience according to

$$\hat{J}_i \, |\phi_j(2)\rangle = \langle \phi_i(1) | \hat{g}_{12} | \phi_i(1) \rangle \, |\phi_j(2)\rangle \tag{2.17}$$

$$\hat{K}_i \, |\phi_j(2)\rangle = \langle \phi_i(1) | \hat{g}_{12} | \phi_j(1) \rangle \, |\phi_i(2)\rangle \, . \tag{2.18}$$

Here, the Coulomb operator corresponds to the averaged classical Coulomb interaction between electrons that can be represented by a local potential. The exchange operator, said to be a nonlocal operator, does not have such a classic analogon and is a direct consequence of the Pauli principle. In order to obtain a set of MOs that make the energy a minimum under the constraint that the MOs remain orthogonal and normalized, Lagrange multipliers (λ_{ij}) are employed. The condition is that the Lagrange function remains constant with orbital variation according to

$$\mathcal{L} = E - \sum_{ij}^{N} \lambda_{ij}(\langle\phi_i|\phi_j\rangle - \delta_{ij}) \tag{2.19}$$

$$\delta\mathcal{L} = \delta E - \sum_{ij}^{N} \lambda_{ij}(\langle\delta\phi_i|\phi_j\rangle - \langle\phi_i|\delta\phi_j\rangle) = 0. \tag{2.20}$$

After solving this condition and simplifying the solution by a unitary transformation resulting in a diagonal matrix of Lagrange multipliers (ϵ), the final Hartree-Fock equations are obtained as

$$\hat{f}_i\phi'_i = \epsilon_i\phi'_i \tag{2.21}$$

where ϕ'_i are the canonical MOs (the primes are dropped from now on) with the corresponding energies ϵ_i. The Fock-operator $\hat{f}_i$ is given by

$$\hat{f}_i = \hat{h}_i + \sum_{j}^{N}(\hat{J}_j - \hat{K}_j). \tag{2.22}$$

The Fock operator is an effective one-electron energy operator describing the kinetic energy of an electron and the attraction to the nuclei by $\hat{h}_i$, and the repulsion to other electrons with $\hat{J}$ and $\hat{K}$. The MO energies are the expectation value of the Fock operator in the MO basis:

$$\langle\phi_i|\hat{f}_i|\phi_i\rangle = \epsilon_i \langle\phi_i|\phi_i\rangle = \epsilon_i. \tag{2.23}$$

Because the Fock operator depends on all occupied MOs via the Coulomb and the exchange operators and a specific Fock orbital can only be determined if all the other occupied orbitals are known, iterative methods are required to solve the HF equations. Note that the HF energy is not simply the sum of the orbital energies of all occupied MOs, because the J_{ij} and K_{ij} would be counted twice this way. Therefore, the actual HF energy is given by

$$E_{\mathrm{HF}} = \sum_{i}^{N} \epsilon_i - \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N}(J_{ij} - K_{ij}) + V_{nn} \tag{2.24}$$

$$\epsilon_i = \langle\phi_i|\hat{f}_i|\phi_i\rangle = h_i + \sum_{j}^{N}(J_{ij} - K_{ij}). \tag{2.25}$$

The Hartree-Fock method is a so-called mean-field method where the interaction between electrons is approximated by the interaction of each electron with the mean-field of all the other electrons without considering any explicit correlation effects between the electrons.

2.1.3 Basis Set Approximation

In principle there are many possible choices of basis functions that can be employed for expressing the MOs, but in practice for molecular calculations, most often the linear combination of atomic orbitals

(LCAO) approach is employed where the MO ψ is expanded as a linear combination of atomic orbitals (AOs) χ according to

$$\psi_i = \sum_{\mu}^{N_{\mathrm{AO}}} C_{\mu i} \chi_\mu. \tag{2.26}$$

Here, the coefficients $C_{\mu i}$ give the contribution of χ_μ to ψ_i. Note that N_{AO} gives the number of AOs that is equal to the number of MOs (N_{MO}). Within the LCAO Ansatz the HF equations (2.21) may be written as

$$\hat{f}_i \sum_{\mu}^{N_{\mathrm{AO}}} C_{\mu i} \chi_\mu = \epsilon_i \sum_{\mu}^{N_{\mathrm{AO}}} C_{\mu i} \chi_\mu. \tag{2.27}$$

After multiplication with a specific basis function from the left and integration over all coordinates the Roothaan-Hall equation are obtained which are given here in the matrix notation

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.28}$$

with the matrix $\mathbf{S}$ containing the overlap elements $S_{ij} = \langle \chi_i | \chi_j \rangle$ between the AOs, the $\mathbf{F}$ matrix containing the elements of the Fock operator, and the $\mathbf{C}$ matrix containing the MO coefficients. The elements of the Fock operator in the MO basis are calculated according to

$$F_{\mu\nu} = \underbrace{\langle \chi_\mu | \hat{h} | \chi_\nu \rangle}_{H_{\mu\nu}} + \underbrace{\sum_{\lambda}^{N_{\mathrm{MO}}} \sum_{\sigma}^{N_{\mathrm{MO}}} P_{\lambda\sigma} [\langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle - \frac{1}{2} \langle \chi_\mu \chi_\lambda | \chi_\sigma \chi_\nu \rangle]}_{G_{\mu\nu}} \tag{2.29}$$

where the density matrix P is given by

$$P_{\mu\nu} = \sum_{i}^{N_{\mathrm{MO}}} n_i C_{\mu i}^* C_{\nu i} \tag{2.30}$$

with the occupation number n_i being 0 or 2. The two-electron integrals $\langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle$ are denoted here in the "physicists" notation and are calculated as

$$\langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle = \int \chi_\mu(1) \chi_\lambda(2) \left| \frac{1}{\mathbf{r}_1 - \mathbf{r}_2} \right| \chi_\nu(1) \chi_\sigma(2) d\mathbf{r}_1 d\mathbf{r}_2. \tag{2.31}$$

Typically calculating these two-electron integrals is the computationally most expensive step in the HF methods scaling formally with $O(N_{\mathrm{MO}}^4)$. The final restricted HF (RHF) energy is then obtained by

$$E_{\mathrm{RHF}} = \frac{1}{2} \sum_{\mu}^{N_{\mathrm{MO}}} \sum_{\nu}^{N_{\mathrm{MO}}} P_{\nu\mu} (H_{\mu\nu} + F_{\mu\nu}) + E_{NN}. \tag{2.32}$$

The AOs typically consist of linear combinations of Gaussian type orbitals (GTOs) of the form (for an s-type orbital)

$$G_A(\mathbf{r}) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \exp(-\alpha(\mathbf{r} + \mathbf{R}_A)^2), \tag{2.33}$$

where α is the exponent determining the diffuseness of the GTO and $\mathbf{R}_A$ is the aufpunkt. GTOs have the advantage that their linear combination yield again a Gaussian, which simplifies the generation of the required one and two-electron integrals.

2.1.4 Unrestricted Hartree-Fock

For the treatment of open-shell species often the unrestricted HF (UHF) method is employed. Here, α and β spinorbitals can have different spatial orbitals $\psi_i^{\alpha/\beta}$ leading to different LCAOs for α and β orbitals according to

$$\psi_i^{\alpha} = \sum_{\mu}^{N_{\mathrm{AO}}} C_{\mu i}^{\alpha} \chi_\mu \tag{2.34}$$

$$\psi_i^{\beta} = \sum_{\mu}^{N_{\mathrm{AO}}} C_{\mu i}^{\beta} \chi_\mu \tag{2.35}$$

The set of atomic orbitals χ still remains the same, but two sets of coefficients $C^{\alpha/\beta}$ are available leading to two sets of spin orbitals. This is depicted in Figure 2.1 together with the RHF and the restricted open-shell HF (ROHF) method, which do not allow breaking of the spin-symmetry.

The two sets of MO coefficients result in two density matrices calculated according to

$$P_{\mu\nu}^{\alpha} = \sum_{i}^{N_{\mathrm{MO}}} n_i^{\alpha} (C_{\mu i}^{\alpha})^* C_{\nu i}^{\alpha} \tag{2.36}$$

$$P_{\mu\nu}^{\beta} = \sum_{i}^{N_{\mathrm{MO}}} n_i^{\beta} (C_{\mu i}^{\beta})^* C_{\nu i}^{\beta} \tag{2.37}$$

with the occupation numbers $n_i^{\alpha/\beta}$ being either 0 or 1.

From the α and β density matrices the spin density matrix $\mathbf{P^S}$ and the total density matrix $\mathbf{P^T}$ can be obtained

$$\mathbf{P^S} = \mathbf{P}^{\alpha} - \mathbf{P}^{\beta} \tag{2.38}$$

$$\mathbf{P^T} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}. \tag{2.39}$$

Like α and β density matrices, there are two Fock matrices $\mathbf{F}^{\alpha/\beta}$ with their matrix elements given according to

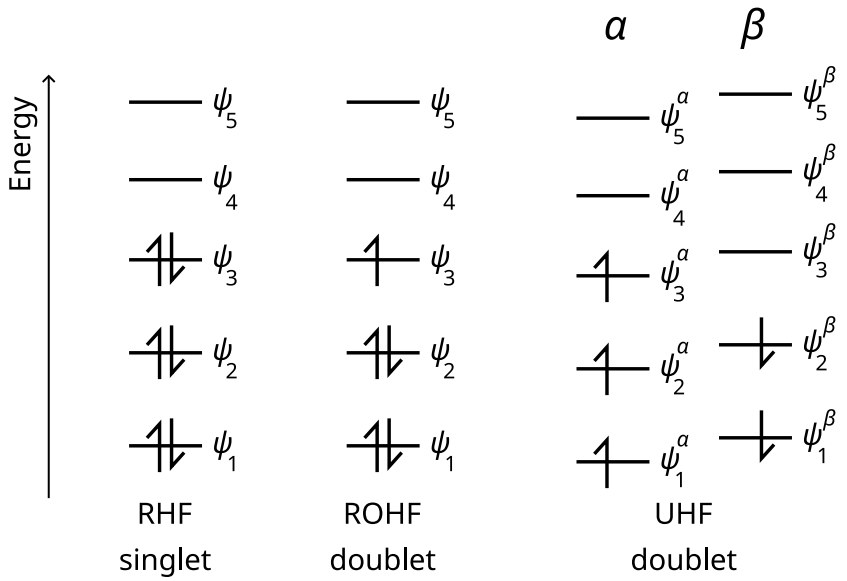

Figure 2.1: Illustration of an RHF singlet, an ROHF doublet, and an UHF doublet. Figure adapted from ref. [44].

$$F^{\alpha}_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda}^{N_{\mathrm{MO}}} \sum_{\sigma}^{N_{\mathrm{MO}}} P^{T}_{\lambda\sigma} \langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle - P^{\alpha}_{\lambda\sigma} \langle \chi_\mu \chi_\lambda | \chi_\sigma \chi_\nu \rangle \tag{2.40}$$

$$F^{\beta}_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda}^{N_{\mathrm{MO}}} \sum_{\sigma}^{N_{\mathrm{MO}}} P^{T}_{\lambda\sigma} \langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle - P^{\beta}_{\lambda\sigma} \langle \chi_\mu \chi_\lambda | \chi_\sigma \chi_\nu \rangle \tag{2.41}$$

Here, the two Fock matrices are explicitly coupled by P^T. The corresponding generalized eigenvalue problems are called the Pople-Nesbet equations given as

$$\mathbf{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\epsilon^{\alpha} \tag{2.42}$$

$$\mathbf{F}^{\beta}\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\epsilon^{\beta}. \tag{2.43}$$

Since two sets of equations need to be solved, the costs of the UHF SCF compared to the RHF SCF are doubled (under the assumption that SCF convergence is the same in both cases). The UHF energy is then calculated according to

$$E_{\mathrm{UHF}} = \frac{1}{2}\sum_{\mu}^{N_{\mathrm{MO}}}\sum_{\nu}^{N_{\mathrm{MO}}}\left[P^{T}_{\nu\mu}H_{\mu\nu} + P^{\alpha}_{\nu\mu}F^{\alpha}_{\mu\nu} + P^{\beta}_{\nu\mu}F^{\beta}_{\mu\nu}\right] + E_{NN}. \tag{2.44}$$

Unlike the RHF wave function, the UHF wave function is not necessarily an eigenfunction of the S^2 operator. Higher lying spin states can mix into the UHF solution resulting in deviation from the exact $\langle S^2\rangle$ value for the corresponding spin state. The $\langle S^2\rangle$ value can be calculated according to

$$\langle S^2\rangle = \langle S^2\rangle_{\mathrm{exact}} + N_\beta - \sum_{i}^{N}\sum_{j}^{N}|S^{\alpha\beta}_{ij}|^2 \quad \text{with} \quad N_\alpha \geq N_\beta \tag{2.45}$$

with the exact value calculated according to

$$\langle S^2\rangle_{\mathrm{exact}} = \left(\frac{N_\alpha - N_\beta}{2}\right)\left(\frac{N_\alpha - N_\beta}{2} + 1\right). \tag{2.46}$$

Since HF does not include dynamic electron correlation so-called artificial symmetry breaking[95] can occur, i.e., spin symmetry breaking (SSB) occurs and recovers some of the dynamic correlation. It is in that sense artificial that the spin symmetry is broken for systems that are adequately described by a single Slater determinant.[95] In a similar fashion, $\langle S^2\rangle$ can be calculated from the Kohn-Sham (KS) auxiliary orbitals in density functional theory (DFT). Albeit the $\langle S^2\rangle$ value is not rigorously defined in KS-DFT, because the KS orbitals are not a wave function that must be an eigenvalue of $\langle S^2\rangle$ but just a set of auxiliary orbitals, their $\langle S^2\rangle$ value can be used for the diagnostic of static correlation.[86]

2.1.5 Electron Correlation

In the following, the electron correlation energy (E_{corr}) is defined as the energetic difference between the formally exact full configuration interaction (FCI) solution (exact for the non-relativistic Schrödinger equation within the Born-Oppenheimer approximation) at the complete basis set (CBS) limit and the RHF method at the CBS limit:

$$E_{\mathrm{corr}} = E_{\mathrm{FCI}} - E_{\mathrm{RHF}}. \tag{2.47}$$

Due to the antisymmetric wave function in HF, electrons are not completely uncorrelated and electrons of the same spin repel each other because of the Pauli principle. This correlation is often called Fermi correlation while correlation between electrons of different spins is called Coulomb correlation. In this work, the term electron correlation refers to the latter. The FCI wave function is given as

$$\Psi_{FCI} = a_0\Phi_0 + \sum_{S} a_S\Phi_S + \sum_{D} a_D\Phi_D + \sum_{T} a_T\Phi_T + \cdots = \sum_{i=0} a_i\Phi_i \tag{2.48}$$

where Φ_0 is the reference SD, typically the HF ground state solution, and $\Phi_{S/D/T\ldots}$ are excited SDs (S for single excitations, D for double excitations and so on) generated from it. The coefficients $a_{0/S/D/\ldots}$ represent the corresponding weights of the SDs in the total FCI wave function. Upon introducing a Lagrange multiplier with the constraint that the FCI wave function is normalized

$$\mathcal{L} = \langle\Psi_{FCI}|\hat{H}|\Psi_{FCI}\rangle - \lambda(\langle\Psi_{FCI}|\Psi_{FCI}\rangle - 1) \tag{2.49}$$

the final FCI energy, after variation is obtained by solving the secular equations

$$(\langle \Psi_{FCI}|\hat{H}|\Psi_{FCI}\rangle - E\mathbf{I})\mathbf{a} = 0. \tag{2.50}$$

Here, $\mathbf{I}$ is the identity matrix and $\mathbf{a}$ is the matrix of the coefficients $a_{0/S/D/\ldots}$. The eigenvalue lowest in energy corresponds to the FCI energy of the ground state and the next lower to the 1$^{\text{st}}$ excited state and the next to the 2$^{\text{nd}}$ excited state and so on.

The correlation energy can be loosely separated into dynamic and static (or nondynamic) correlation

$$E_{\text{corr}} = E_{\text{dynamic}} + E_{\text{static}}. \tag{2.51}$$

In the FCI definition, dynamic correlation is a result of small contributions from many determinants, while static correlation occurs from large contributions from a few determinants. Systems that have large static correlation (more than one large contribution) can be described as multi-configurational or multireference systems.[47,86]

In practice, FCI can only be computed for small model systems and is not feasible for chemical relevant systems due to its $O(N!)$ scaling with the system size. Truncated CI methods are termed CIS, CISD, CISDT, ... and have two shortcomings. First, they are not size-consistent, meaning that for two non-interacting subsystems the sum of the subsystem energies will not be equal to the energy calculated for the total system and second they converge only slowly toward the FCI limit with the number of SDs. The size-consistency problem is not present in coupled cluster (CC) methods, where all excitations of a given type are included up to infinite order. We define an excitation operator according to

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N \tag{2.52}$$

where the $\hat{T}_1$ operator generates all single excited determinants from the HF reference and the $\hat{T}_2$ all double excited determinants and so on. With this excitation operator the CI wave function is obtained by a linear ansatz while for CC an exponential ansatz is employed according to

$$\Psi_{CI} = (1 + \hat{T})\Phi_0 = (1 + \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots)\Phi_0 \tag{2.53}$$

$$\Psi_{CC} = e^{\hat{T}}\Phi_0 = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{T}^k\Phi_0. \tag{2.54}$$

The naming of the truncated CC methods is similar to the truncated CI methods (CCSD, CCSDT, CCSDTQ, ...). The CC energy is not obtained variationally but by projection

$$E_{CC} = \langle \Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle \tag{2.55}$$

and the amplitudes are calculated iteratively according to

$$\langle \Phi_S | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle = 0 \tag{2.56}$$

$$\langle \Phi_D | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle = 0 \tag{2.57}$$

$$\langle \Phi_T | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle = 0 \tag{2.58}$$

$$\vdots$$

From the truncated CC methods, CCSD(T), for which the triple excitations are calculated perturbatively,[57] is well-known as the so-called gold-standard of quantum chemistry which typically yields errors below 1 kcal·mol^{-1} for atomization energies, ionization potentials, and reaction energies of organic molecules.

Another efficient approach for calculating the correlation energy stems from perturbation theory.[44] The idea in perturbation theory is that the approximate solution differs from the exact solution only by a small perturbation. For this purpose the Hamiltonian ($\hat{H}$) is separated into a reference ($\hat{H}_0$) and a perturbation ($\hat{H}'$) according to

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \tag{2.59}$$

where the parameter λ scales the strength of the perturbation. It is assumed that the Schrödinger equation for the reference Hamiltonian is solved. This Hamiltonian is inserted into the Schrödinger equation and a Taylor expansion in powers of the perturbation parameter λ is performed. Depending on where the Taylor expansion is truncated the corresponding orders of perturbation theory are obtained. The most common choice for the reference Hamiltonian is the sum over Fock operators, leading to Møller-Plesset (MP) perturbation theory.[44] The most popular variant of MP is MP2, for which the correlation energy is calculated according to

$$E(MP2)_{\text{corr}} = \sum_{i<j}^{N_{occ}} \sum_{a<b}^{N_{virt}} \frac{|\langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \tag{2.60}$$

The most expensive part of an MP2 calculation is to transform the two-electron integrals from the AO to the MO basis which scales with $O(N_{el}^2 N_{\text{MO}}^3)$. Compared to other post-HF methods MP2 has relatively small computational demand, which makes it a popular method that is also often combined with DFT in so-called double-hybrid (DH) methods. Nevertheless, compared to most DFT methods MP2 is costly and the application of DHs is not possible to large systems (above 100 atoms). One shortcoming of MP2 is that for small HOMO-LUMO gaps the denominator approaches zero and the MP2 correlation energy diverges. This makes the application of MP2 for TM systems critical, where small gap systems often occur.

2.1.6 Kohn-Sham Density Functional Theory

The foundation of Density Functional Theory[63,65] (DFT) are the Hohenberg-Kohn theorems.[96,97] The first Hohenberg-Kohn theorem legitimizes the use of the ground-state electron density $\rho(\mathbf{r})$ as basic variable. It states that the external potential $v(\mathbf{r})$ is determined by the electron density. Since the electron density also determines the number of electrons, it follows that it determines the ground-state

wave function and all other electronic properties of the system.[63] The second Hohenberg-Kohn theorem[96] provides a variational principle. In Kohn-Sham (KS) DFT[98] the ground state energy is formulated according to

$$E^{KS}[\rho] = T_s^{KS}[\rho] + V_{ne}[\rho] + J[\rho] + E_{XC}[\rho] \tag{2.61}$$

where the kinetic energy is calculated for the KS non-interacting reference system according to

$$T_s^{KS}[\rho] = \sum_i^{N_{\mathrm{MO}}} n_i \langle \phi_i | -\frac{1}{2}\nabla^2 | \phi_i \rangle \tag{2.62}$$

where ϕ_i are the one-electron KS orbitals and n_i are occupation numbers. The electron density of the KS reference system can be calculated according to

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} n_i |\phi_i(\mathbf{r})|^2. \tag{2.63}$$

The remaining terms in 2.61 are the potential energy between the nuclei and the electrons V_{ne}, the classical electron-electron repulsion J, and the exchange-correlation functional E_{XC}. While the first two are known, for the latter no explicit form is available. There are hundreds of approximate exchange-correlation functionals that have been developed throughout the years.[66,67,99] The "Jacob's ladder" by Perdew[100] represents the most prominent[67] attempt to categorize the different exchange-correlation functional approximations based on their physical ingredients and is depicted in Figure 2.2.

On the first rung of the ladder are the local spin density approximations (LSDA) which only depend on the density $\rho(\mathbf{r})$ and assume a uniform electron gas. For molecular calculations LSDAs yield large errors, but they have found much applications in extended systems, such as metals, where the approximation of a slowly varying electron density is quite valid. On the second rung of the ladder are the generalized gradient approximation (GGA) exchange-correlation functionals that in addition to the density, depend on the gradient of the density $\nabla\rho(\mathbf{r})$. GGA functionals are regularly employed for DFT calculations of molecules. For example the PBE[101] or BP86[102,103] functionals are prominent GGA functionals. On the third rung are the meta-GGA (mGGA) functionals that incorporate the kinetic energy density $\tau = (1/2)\sum_i^{N_{\mathrm{MO}}} |\nabla\phi_i|^2$ or the Laplacian of the density $\nabla^2\rho(\mathbf{r})$, which further improves the accuracy. Well known examples include the TPSS,[104] the SCAN[105] and the r^{2}SCAN[106,107] functionals. On the fourth rung are the hybrid functionals which incorporate Hartee-Fock exchange (HFX) according to

$$E_{XC}^{hybrid} = (1 - a_X)E_X^{GGA/mGGA} + a_X E_X^{HF} + E_C^{GGA/mGGA} \tag{2.64}$$

where the mixing parameter a_X can be determined empirically or based on the adiabatic connection.[108] On the highest rung of Jacob's ladder, virtual orbitals are incorporated in the energy expression, typically via MP2, which leads to so-call double-hybrid (DH) functionals. The exchange-correlation energy is given according to

$$E_{XC}^{double-hybrid} = (1 - a_X)E_X^{GGA/mGGA} + a_X E_X^{HF} + (1 - a_C)E_C^{GGA/mGGA} + a_C E_C^{MP2} \tag{2.65}$$

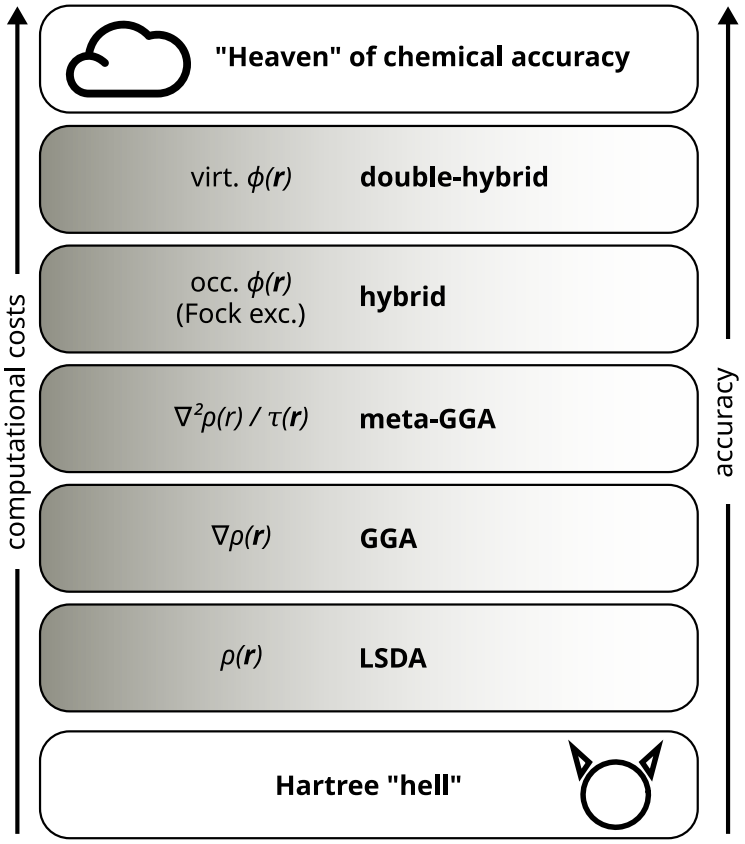

Figure 2.2: Jacobs ladder of DFT functional approximations.

where the parameter a_C can again be determined empirically or by arguments based on the adiabatic connection.[109]

For a specific exchange-correlation approximation the Kohn-Sham equations are given similar to the Roothaan-Hall equations according to

$$\hat{f}_i^{KS}\phi_i = \epsilon_i\phi_i \tag{2.66}$$

where the KS effective one-electron operator is obtained with

$$\hat{f}_i[\rho] = \hat{h}_i[\rho] + \sum_j^N \hat{J}_j[\rho] + v_{xc}[\rho]. \tag{2.67}$$

and the exchange-correlation potential is obtained according to

$$v_{xc} = \frac{\delta E_{XC}}{\delta\rho}. \tag{2.68}$$

The KS equations are solved self-consistently in an SCF procedure.

2.1.7 Semiclassical London Dispersion Corrections

Mean field electronic structure methods like HF or DFT cannot describe long-range correlation effects and hence they cannot account for London dispersion interactions.[8,110] The long-range behavior of the London dispersion can be described in a simplified way according to

$$E_{disp}(R_{AB}) \propto \frac{C_6^{AB}}{R_{AB}^6} \tag{2.69}$$

where C_6^{AB} are the pairwise dipole-dipole dispersion coefficients describing the strength of the dispersion interaction between two charge densities at atoms A/B. Because the long-range behavior is known, one can add the missing dispersion interaction in a post-HF or post-KS fashion with semiclassical dispersion corrections such as the D3[111,112] or the D4 model.[9,113,114]

The energy expression of the D4 model is given according to

$$\begin{aligned} E_{disp}^{D4} = &-\frac{1}{2}\sum_{AB}\sum_{n=6,8} s_n \frac{C_{(n)}^{AB}}{R_{AB}^{(n)}} f_{BJ}^{(n)}(R_{AB}) \\ &-\frac{1}{6}\sum_{ABC} s_9 E^{ABC} f_{damp}^{(9)}(\overline{R}_{ABC}). \end{aligned} \tag{2.70}$$

Three-body effects are included by the Axilrod-Teller-Muto (ATM) term[115,116] according to

$$E^{ABC} = \frac{C_9^{ABC}(3\cos(\theta_a)\cos(\theta_b)\cos(\theta_c)+1)}{(R_{AB}R_{BC}R_{CA})^3}. \tag{2.71}$$

Here, θ_a, θ_b, θ_c are the inertial angles of the triangle formed by R_{AB}, R_{BC}, and R_{CA}, respectively. The D4 model considers the C_6^{AB}, C_8^{AB}, and C_9^{ABC} dispersion coefficients for the dipole-dipole, dipole-quadrupole, and triple-dipole dispersion interactions scaled by the corresponding parameters (s_6, s_8, and s_9) where in the DFT context typically only the s_8 is fitted empirically and the other two are set to one, except for double-hybrids where the s_6 is adjusted as well. The corresponding Becke-Johnson $f_{BJ}^{(n)}(R_{AB})$ and zero-damping $f_{damp}^{(9)}(\overline{R}_{ABC})$ functions damp the dispersion correction in the intermediate distance regime where DFT methods can – to some extend – describe different amounts of London dispersion interactions based on the exchange correlation functional approximation employed. The Becke-Johnson damping function is given according to

$$f_{BJ}^{(n)}(R_{AB}) = \frac{R_{AB}^{(n)}}{R_{AB}^{(n)} + (a_1 R_0^{AB} + a_2)} \tag{2.72}$$

where the functional specific parameters a_1 and a_2 are fitted and the cutoff-radii R_0^{AB} is defined according to

$$R_0^{AB} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}. \tag{2.73}$$

The zero-damping scheme for the three-body effects is given by

$$f_{damp}^{(9)}(\overline{R}_{ABC}) = \frac{1}{1 + 6(\overline{R}_{ABC})^{-16}} \tag{2.74}$$

with the averaged interaction distance

$$\overline{R}_{ABC} = \left(R_{AB}R_{BC}R_{CA})/(R_0^{AB}R_0^{BC}R_0^{CA}\right)^{1/3}. \tag{2.75}$$

In the D4 model the C_6^{AB} dispersion coefficients are obtained by an on-the-fly Casimir-Polder integration[9] from precomputed reference dynamic polarizabilities at imaginary frequencies according to

$$C_6^{AB} = \frac{\pi}{3}\sum_{A,ref}^{N_{ref}}\sum_{B,ref}^{N_{ref}}\sum_{u}^{23} w_u w_A^{CN} w_B^{CN} \zeta(z^A, z^{A,ref})\zeta(z^A, z^{B,ref})\alpha_{A,ref}(i\omega)\alpha_{B,ref}(i\omega) \tag{2.76}$$

where w_u are the weights of the frequency integration grid, w_A^{CN} scales the atom-in-molecule polarizability $\alpha_{A,ref}$ based on the coordination number of A incorporating its chemical environment. The charge scaling function ζ scales the atom-in-molecule polarizability based on partial atomic charges according to

$$\zeta(z^A, z^{A,ref}) = \exp\left[\beta_1\left\{1 - \exp\left[\gamma^A\left(1 - \frac{z^{A,ref}}{z^A}\right)\right]\right\}\right] \tag{2.77}$$

with γ^A being the Hubbard parameter for the chemical hardness, β_1 being an empirical parameter set to 3 by inspection, and z_A being the effective nuclear charge calculated from the nuclear charge Z_A and the atomic partial charge q^A according to

$$z^A = Z^A + q^A. \tag{2.78}$$

The atomic partial charges q^A are obtained from a classical charge model based on electronegativity equilibration of Gaussian-type charge densities (EEQ). The C_8^{AB} and the C_9^{ABC} dispersion coefficients can then be approximated from the C_6^{AB} coefficients. For a detailed description of the D4 dispersion correction the interested reader is referred to ref. [9].

2.1.8 Semiempirical Tight-Binding Methods

Density functional tight-binding methods like the GFNn-xTB methods[70–72] or DFTB methods[73,74,117] are semiempirical quantum mechanical methods that approximate KS-DFT. Their total energy is expanded in a Taylor expansion for the electron density $\rho = \rho_0 + \delta\rho$ in terms of density fluctuations $\delta\rho$ around a superposition of (neutral) atomic reference densities $\rho_0 = \sum_A \rho_{0,A}$ according to

$$E[\rho] = E^{(0)}[\rho_0] + E^{(1)}[\rho_0, \delta\rho] + E^{(2)}[\rho_0, (\delta\rho)^2] + E^{(3)}[\rho_0, (\delta\rho)^3] + \cdots. \tag{2.79}$$

The expansion is formally exact, but in practice it has to be truncated. For the most sophisticated

methods the expansion is truncated after the third-order term.[74,117] In the case of the GFN*n*-xTB methods, two self-consistent variants have been developed, namely GFN1-xTB, and GFN2-xTB. Both include the energy terms from zeroth to third order, which are approximated by chemically motivated energy expressions and only require element-wise or global parameters, but no pair-parameters. The only integrals that have to be calculated in GFN*n*-xTB methods are the relatively cheap overlap integrals $S_{\mu\nu}$ of AOs μ and ν, keeping the computational costs three to four orders of magnitude below the costs of DFT methods. The GFN*n*-xTB wave functions are formulated in terms of a partially polarized, mostly minimal valence basis sets, consisting of spherical Gaussian-type orbitals and they are parameterized for geometries, frequencies and non-covalent interactions (GFN) against a large DFT reference fit set.[71] Since GFN1- and GFN2-xTB go beyond the first order energy contribution, the Mulliken charges in the first to third order energy expression are determined self-consistently and a pseudo-eigenvalue problem of the Roothaan-Hall-type has to be solved to determine the coefficients of the LCAO Ansatz. In the following, the energy expressions for the self-consistent GFN1-xTB and GFN2-xTB methods are discussed. The repulsion energy is given as an atom-pairwise expression:

$$E_{rep} = \frac{1}{2}\sum_{A,B}^{N_{at}} \frac{Z_A^{eff} Z_B^{eff}}{R_{AB}} \exp\left[-\sqrt{\alpha_A \alpha_B} R_{AB}^{k_f}\right] \tag{2.80}$$

where N_{at} is the number of atoms, and Z_A^{eff} the element-specific constants, which define the magnitude for the repulsion energy and loosely correspond to effective nuclear charges. k_f is a global parameter with slight adjustments for H and He in GFN2-xTB, while α_A are element-specific parameters. Covalent bond formation is allowed through an extended Hückel-type (EHT) energy given by

$$E_{EHT} = \sum_{\mu\nu} P_{\mu\nu} H_{\nu\mu}^{EHT} \tag{2.81}$$

with the valence electron density matrix $P_{\mu\nu}$. The EHT matrix elements are given by

$$H_{\mu\nu}^{EHT} = \frac{1}{2} K_{AB}^{ll'} S_{\mu\nu} (H_{\mu\mu} + H_{\nu\nu}) X(EN_A, EN_B) \Pi(R_{AB}, l, l') \Upsilon(\zeta_l^A, \zeta_{l'}^B), \forall \mu \in l(A), \nu \in l'(B). \tag{2.82}$$

Here, $K_{AB}^{ll'}$ is a shell-specific scaling constant (for some element pairs GFN1-xTB contains element pair-specific scaling parameters hence the subscript A and B), $S_{\mu\nu}$ is the overlap matrix element of the AOs and $H_{\mu\mu}/H_{\nu\nu}$ are the diagonal elements that depend on the chemical environment and differ for different GFN*n*-xTB methods (see below). The last three terms are system-specific, involve flexible scaling functions, and are not included in standard EHT. The distance-dependent adjustment of the EHT energy in addition to the distance dependence encoded in the overlap matrix is given by a polynomial scaling function according to

$$\Pi(R_{AB,l,l'}) = \left(1 + k_{A,l}^{poly} \left(\frac{R_{AB}}{R_{cov,AB}}\right)^{\frac{1}{2}}\right)\left(1 + k_{B,l'}^{poly} \left(\frac{R_{AB}}{R_{cov,AB}}\right)^{\frac{1}{2}}\right). \tag{2.83}$$

Here, $k_{A,l}^{poly}$ are element- and shell-specific parameters and $R_{cov,AB}$ are the summed covalent radii of A and B. The electronegativity-dependent term $X(EN_A, EN_B)$ is either different or not present

at all in the different GFN methods. The same holds for the last term $\Upsilon(\zeta_l^A, \zeta_{l'}^B)$ depending on the Slater-type orbital (STO) exponents ζ_l^A. The isotropic electrostatic and XC energy is the same for GFN1-xTB and GFN2-xTB and originates from the second-order term in the tight-binding expression. It is given according to

$$E_\gamma = \frac{1}{2}\sum_{A,B}^{N_{at}}\sum_{l\in A}\sum_{l'\in B} q_l q_{l'}\gamma_{AB,ll'} \tag{2.84}$$

with the partial Mulliken shell-charges q_l/q_l', and $\gamma_{AB,ll'}$, which are short-ranged damped Coulomb interactions given by

$$\gamma_{AB,ll'} = \frac{1}{\sqrt{R_{AB}^2 + \eta_{AB,ll'}^{-2}}}. \tag{2.85}$$

The short-range damping term $\eta_{AB,ll'}$ differs for GFN1-xTB and GFN2-xTB and is given below. GFNn-xTB methods employ Fermi-smearing to smear out the orbital occupation adding the expression

$$G_{Fermi} = k_B T_{el}\sum_{\sigma=\alpha,\beta}\sum_i [n_{i\sigma}\ln(n_{i\sigma}) + (1-n_{i\sigma})\ln(1-n_{i\sigma})] \tag{2.86}$$

to the energy terms. Here, k_B is the Boltzmann constant and T_{el} is the electronic temperature that is set to 300 K. The occupation numbers are determined by the Fermi-distribution

$$n_{i\sigma} = \frac{1}{\exp[(\epsilon_i - \epsilon_F^\sigma)/(k_B T_{el})] + 1} \tag{2.87}$$

with the orbital energy ϵ_i of the orbital ψ_i and ϵ_F^σ is the Fermi level within the respective spin orbital space. By default GFNn-xTB calculations are restricted because the energy expressions are spin-independent. Therefore, low-spin states are always favored. Nevertheless, geometries of high-spin states can be optimized simply by employing the respective occupation in the geometry optimization. With the inclusion of spin-polarization into the xTB Hamiltonian unrestricted calculations with different spatial orbitals for α and β spin orbitals are possible which enables the correct description of high-spin states. This has been done in Chapter 3. The energy expressions for GFN1-xTB and GFN2-xTB are given in the following.

GFN1-xTB

The final energy expression for GFN1-xTB is given according to:

$$E_{GFN1-xTB} = E_{rep} + E_{disp}^{D3} + E_{XB}^{GFN1} + E_{EHT} + E_\gamma + E_\Gamma^{GFN1}. \tag{2.88}$$

The diagonal EHT matrix elements are atomic environment dependent and given by

$$H_{\mu\mu} = h_A^l(1 + k_{CN,l}CN_A), \forall\mu \in l \in A \tag{2.89}$$

where h_A^l is a shell- and element-specific parameter, while $k_{CN,l}$ are global angular momentum-specific parameters. CN_A is the is the geometric atomic fractional coordination number taken from

the D3 dispersion model.[111] For GFN1-xTB the STO exponent depending term is $\Upsilon(\zeta_l^A, \zeta_{l'}^B) = 1$ for all cases and the electronegativity dependent term is given by

$$X(EN_A, EN_A) = (1 + k_{EN}\Delta EN_{AB}^2) \tag{2.90}$$

where ΔEN_{AB} is the difference of the standard Pauling electronegativities and $k_{EN} = -0.007$ is a global parameter. The short-range damping term for the isotropic electrostatic is given by

$$\eta_{AB,ll'} = 2\left(\frac{1}{\eta_A(1+\kappa_A^l)} + \frac{1}{\eta_B(1+\kappa_B^l)}\right)^{-1} \tag{2.91}$$

which is the harmonic mean of the effective shell hardness values, which originate from element specific atomic hardness values and a shell-dependent scaling parameter. The third order on-site electrostatic/XC correction is given by

$$E_\Gamma^{GFN1} = \frac{1}{3}\sum_A^{N_{at}} q_A^3 \Gamma_A \tag{2.92}$$

with the atomic Mulliken charges $q_A = \sum_l q_{A,l}$ and the element-specific Γ_A parameters.

The London dispersion interactions (E_{disp}^{D3}) in GFN1-xTB are described by the D3 dispersion correction with the Becke-Johnson damping function.[70,111,112]

Because weak halogen bonds are not described well with the monopole electrostatic of GFN1-xTB, a purely geometry-dependent halogen-bond (XB) correction E_{XB} is added, which is described elsewhere.[70] The monopole electrostatic in a minimal basis is also problematic with hydrogen bonds. Therefore, an additional s-AO function on hydrogen is added in GFN1-xTB to improve the description of hydrogen bonds.

GFN2-xTB

In GFN2-xTB the energy expression is given by

$$E_{GFN2-xTB} = E_{rep} + E_{disp}^{D4} + E_{EHT} + E_\gamma + E_{AES} + E_{AXC} + E_\Gamma^{GFN2} \tag{2.93}$$

and the diagonal elements of the EHT matrix are given by

$$H_{\kappa\kappa} = h_A^l - \delta h_{CN_A'}^l CN_A' \tag{2.94}$$

with the shell and element-specific parameters h_A^l and $\delta h_{CN_A'}^l$, where $\delta h_{CN_A'}^l$ is scaled by the GFN2-xTB coordination number, which is a modified long-range variant of the D3 coordination number. In GFN2-xTB the exponent-depending term is given by

$$\Upsilon(\zeta_l^A, \zeta_{l'}^B) = \left(\frac{2\sqrt{\zeta_l^A \zeta_{l'}^B}}{\zeta_l^A + \zeta_{l'}^B}\right)^{\frac{1}{2}} \tag{2.95}$$

which mimics the effects of kinetic energy integrals from ab initio theory. The expression for the

electronegativity-dependent term is the same as in GFN1-xTB, but the parameter $k_{EN} = 0.02$ is increased. The short-range damping for the second-order electrostatic energy is given by

$$\eta_{AB,ll'} = \frac{1}{2}\left[\eta_A(1+\kappa_A^l) + \eta_B(1+\kappa_B^{l'})\right] \tag{2.96}$$

where η_A/η_B are element specific fit parameters and $\kappa_A^l/\kappa_B^{l'}$ are element- and shell-specific scaling factors. In contrast to GFN1-xTB, the London dispersion interactions are treated by the D4 dispersion model in an iterative fashion where the charges are not taken from the EEQ model but are the GFN2-xTB Mulliken charges that are obtained self-consistently.

The third-order term in GFN2-xTB is as in GFN1-xTB an on-site term, formulated in a shell-specific form

$$E_\Gamma^{GFN2} = \frac{1}{3}\sum_A^{N_{at}}\sum_{l\in A}(q_l)^3 K_l^\Gamma \Gamma_A \tag{2.97}$$

with the only difference being the additional global shell-specific parameter K_l^Γ. Special to GFN2-xTB is that it includes anisotropic electrostatic (AES) and XC terms given by

$$\begin{aligned} E_{AES} =& E_{q\mu} + E_{q\Theta} + E_{\mu\mu} \\ =& \frac{1}{2}\sum_{A,B}\{f_3(R_{AB})[q_A(\boldsymbol{\mu}_B^T\mathbf{R}_{AB}) + q_B(\boldsymbol{\mu}_A^T\mathbf{R}_{AB})] \\ &+ f_5(R_{AB})[q_A\mathbf{R}_{AB}^T\boldsymbol{\Theta}_B\mathbf{R}_{AB} + q_B\mathbf{R}_{AB}^T\Theta_A\mathbf{R}_{AB} \\ &- 3(\boldsymbol{\mu}_A^T\mathbf{R}_{AB})(\boldsymbol{\mu}_B^T\mathbf{R}_{AB}) + (\boldsymbol{\mu}_A^T\boldsymbol{\mu}_B)R_{AB}^2]\} \end{aligned} \tag{2.98}$$

with $\boldsymbol{\mu}_A$ the cumulative atomic dipole moment of atom A and $\boldsymbol{\Theta}_A$ the corresponding traceless quadrupole moment. These cumulative atomic multipole moments[118] describe the local atomic multipole moment contribution in a Mulliken approximation scheme. The damping functions $f_n(R_{AB})$ damp the anisotropic electrostatic interaction and are a modification of the original damping function employed in the D3 model.[111] The second-order anisotropic XC energy in GFN2-xTB is given by

$$E_{AXC} = \sum_A (f_{XC}^{\mu_A}|\boldsymbol{\mu}_A|^2 + f_{XC}^{\Theta_A}||\boldsymbol{\Theta}_A||^2) \tag{2.99}$$

with the element specific parameters $f_{XC}^{\mu_A}$ and $f_{XC}^{\Theta_A}$. Because of the anisotropic electrostatic terms in GFN2-xTB no halogen bond or hydrogen bond corrections are necessary. Since the additional s-AO on the hydrogen atom is not required the GFN2-xTB method is computationally slightly cheaper than GFN1-xTB.

2.2 Free Energy Contributions

In order to make reliable comparisons between calculated and experimental data, electronic energies alone are not sufficient, but Gibbs free energies are required[44] considering the effect of a finite Temperature T. The difference in the free energy of a chemical reaction can be decomposed (assuming

SI units) according to:

$$\Delta G = \Delta H - T\Delta S \tag{2.100}$$
$$= \Delta E + \Delta E_{ZPVE} + \Delta H(0K \rightarrow T) + p\Delta V - T\Delta S \tag{2.101}$$
$$= \Delta E + \Delta G_{corr.}(T). \tag{2.102}$$

Here, ΔE is the electronic reaction energy corresponding to the difference in the electronic energies of all reactants and products in a chemical reaction. For example, for an association reaction of the type $A + B \rightarrow C$ it is given by $\Delta E = E(C) - E(A) - E(B)$. The contribution ΔE_{ZPVE} accounts for the quantum mechanical zero-point vibrations occurring even at a temperature of 0 K. The term $\Delta H(0K \rightarrow T)$ considers the temperature dependent contributions to the enthalpy in a chemical reaction stemming from the translational, rotational and vibrational degrees of freedom. The volume work is considered by $p\Delta V$ and the contributions from the entropy by $T\Delta S$.

Within the Born-Oppenheimer approximation these corrections are usually obtained by the rigid-rotor harmonic-oscillator (RRHO) approximation. Here, it is assumed that the electronic, rotational, and vibrational degrees of freedom are separable and their contributions together with the translational contribution are additive:[44]

$$H_{tot} = H_{trans} + H_{rot} + H_{vib} + H_{elec} \tag{2.103}$$
$$S_{tot} = S_{trans} + S_{rot} + S_{vib} + S_{elec}. \tag{2.104}$$

The electronic contribution is normally neglected because the first excitation energy is often much larger than $k_B T$ with k_B being the Boltzmann constant. Therefore, electronically excited states are not populated under standard conditions. The contributions of translational (trans), rotational (rot), and vibrational (vib) degrees of freedom to the enthalpy are calculated by

$$\begin{aligned} H_{trans} &= \frac{5}{2}RT \\ H_{rot} &= \frac{3}{2}RT \\ H_{vib} &= R \sum_{i=1}^{3N-6(7)} \left(\frac{h\nu_i}{2k_B} + \frac{h\nu_i}{k_B} \frac{1}{\exp(h\nu_i/k_B T) - 1} \right) \end{aligned} \tag{2.105}$$

with the ideal gas constant R, Planck's constant h, and the vibrational frequencies ν_i. The translational contribution to the entropy is given by

$$S_{trans} = \frac{5}{2}R + R \ln \left(\frac{V}{N_A} \left(\frac{2\pi M k_B T}{h^2} \right)^{\frac{3}{2}} \right) \tag{2.106}$$

with M being the total molecular mass, N_A being the Avogadro constant, and V being the volume

of 1 mol of ideal gas. The rotational contribution is calculated according to

$$S_{rot} = R\left(\frac{3}{2} + \ln\left(\frac{\sqrt{\pi}}{\sigma}\left(\frac{8\pi^2 k_B T}{h^2}\right)^{\frac{3}{2}}\sqrt{I_1 I_2 I_3}\right)\right) \tag{2.107}$$

with the symmetry index σ, which is the order of the rotational subgroup in the molecular point group and the moments of inertia I_1,I_2, and I_3. The vibrational contribution to the entropy in the default RRHO scheme is calculated according to

$$S_{vib} = R\sum_{i=1}^{3N-6(7)}\left(\frac{h\nu_i}{k_B T} + \frac{1}{\exp(h\nu_i/kT) - 1} - \ln(1 - \exp(-h\nu_i/k_B T))\right). \tag{2.108}$$

Since low-lying vibrational frequencies are notoriously inaccurate in the harmonic approximation[119] the modified RRHO (mRRHO) scheme is used throughout this thesis. In the mRRHO scheme,[119] low-lying vibrational modes are treated by a free-rotor approximation. They are approximated according to

$$S_{rot\,(\text{free})} = R\left[1/2 + ln\left\{\left(\frac{8\pi^3 \mu' k_B T}{h^2}\right)^{1/2}\right\}\right] \tag{2.109}$$

with the modified moment of inertia

$$\mu' = \frac{\mu B_{av}}{\mu + B_{av}} \tag{2.110}$$

where B_{av} is the average molecular moment of inertia, and μ is the is the moment of inertia for a free-rotor given by

$$\mu = \frac{h}{8\pi^2 \nu}. \tag{2.111}$$

The continuous interpolation between the rotational and harmonic vibrational approximations is given by a weighting function w of ν according to

$$S_{vib,\text{mRRHO}} = w(\nu)S_{vib} + [1 - w(\nu)]S_{rot\,(\text{free})} \tag{2.112}$$

with

$$w(\nu) = \frac{1}{1 + (\nu_0/\nu)^4}. \tag{2.113}$$

Here, ν_0 is typically chosen to be 100 or 50 cm^{-1}.

Adding up all these contributions results in the total thermostatistical correction:

$$G_{trv} = (H_{trans} + H_{rot} + H_{vib}) - T(S_{trans} + S_{rot} + S_{vib,\text{mRRHO}}). \tag{2.114}$$

2.3 Solvation Corrections

Chemical reactions, and especially 3d TM electrocatalytic reactions, typically occur in solution and modelling these solvation effects is important. Broadly, two approaches can be distinguished: Either the solvent molecules are included explicitly in the calculation or the solvent is included implicitly as a continuous medium.[91] The latter is used preferably because it has the advantage that the solvent molecules are not included in the simulation and therefore computational costs are kept low. On the downside explicit solvent-solute interactions, such as hydrogen bonds or solvent molecules coordinating to metal centers, or dynamic changes in the solvent structure cannot be described in an averaged fashion by continuous solvation models.

The contribution to the free energy of bringing the system from the gas-phase in solution can be assumed to be additive and the free energy of a reaction in the gas-phase (ΔG_{gas}) can be transformed to solution by

$$\Delta G = \Delta G_{gas} + \Delta \delta G_{solv} \tag{2.115}$$

where δG_{solv} is the solvation free energy for reactants and products and $\Delta \delta G_{solv}$ is their difference. The error for calculating solvation free energies with implicit solvation models (upon other errors) can be reduced for the calculation of redox potentials by explicitly calculating the reference electrode and benefitting from systematic error compensation. This was done in Chapter 7 and in Chapter 8.

2.4 Transition State Theory

If the rate constant k_{rate} and the initial concentrations of all species for a reaction are known the concentration of various species at any given time can be calculated.[44] This macroscopic rate constant can be obtained from transition state theory in the following way. On the potential energy surface within the Born-Oppenheimer approximation stable molecules correspond to minima and a chemical reaction can be described as nuclei moving from one minima to another. In transition state theory it is assumed that this reaction occurs along the pathway lowest in energy as shown in Figure 2.3.[44]

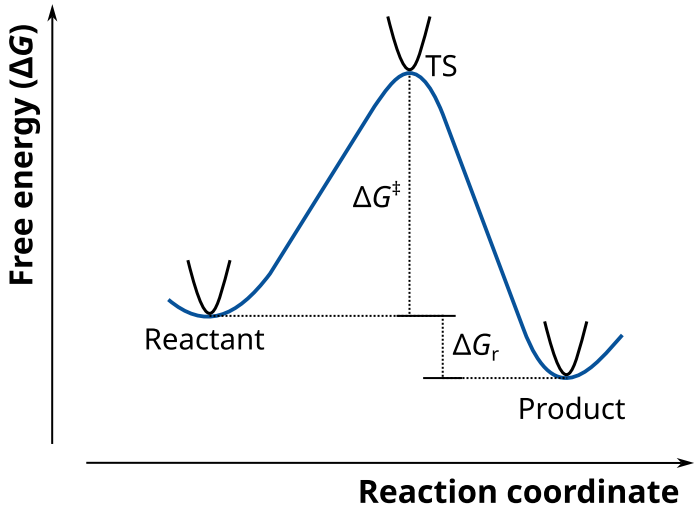

Figure 2.3: Schematic illustration of a reaction path. Figure adapted from ref. [44].

The transition state (TS) or the transition structure is the configuration that divides the reactant and the product parts of the energy surface. It corresponds to a first-order saddle point, a maximum in the reaction coordinate direction and a minimum along all other coordinates.[44] In transition state theory the dynamics along the reaction coordinate is treated classically and it is assumed that the molecules at the TS are in equilibrium with the reactant. The macroscopic rate constant can then be calculated according to the Eyring equation:

$$k_{\text{rate}} = \frac{k_B T}{h} \exp(-\Delta G^{\ddagger}/RT). \tag{2.116}$$

Here, $\Delta G^{\ddagger}$ is the Gibbs free energy difference between the TS and the reactant. Finding transition states on the potential energy surface can be challenging, algorithms such as the growing string method (GSM)[120,121] can help with this task. Rate constants were calculated this way in Chapter 8.

CHAPTER 3

High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods

Hagen Neugebauer,[†] Benedikt Bädorf,[†] Sebastian Ehlert,[‡] Andreas Hansen,[†] Stefan Grimme[†]

Received: 19 March 2023
Published online: 4 July 2023

Reprinted in Appendix A (adapted) with permission[§] from
H. Neugebauer, B. Bädorf, S. Ehlert, A. Hansen, and S. Grimme, *High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods*, J. Comput. Chem. **44** (2023) 2120, DOI: 10.1002/jcc.27185

Own contributions

- Concept and choice of methods
- performing of timing calculations
- supervision for all calculations
- interpretation of the results
- writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, University of Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]Microsoft Research AI4Science, 1118 CZ Schiphol, Netherlands
[§]Permission requests to reuse material from this chapter should be directed to Wiley Periodicals LLC.

The GFNn-xTB methods ($n = 1, 2$) are an efficient yet robust tool for modeling organic chemistry and are also reliable for modeling transition metal (TM) complexes. They can be employed for the geometry optimization of large TM complexes,[10] metadynamic-based conformer searches,[122] or high-throughput reaction mechanism exploration.[123] One shortcoming of the GFNn-xTB methods is that their energy expression is only occupation-dependent but not spin-dependent. As a result, geometry optimizations of high-spin states can be performed with the plain GFNn-xTB methods, but the low-spin state will always be the ground state. Therefore, if information about the actual spin-state of a system is desired, one has to fall back to comparatively expensive DFT methods. In this work, we have remedied this issue by introducing the concept of spin-polarization to the xTB Hamiltonian that enables spin-polarized calculations of open-shell systems. The new methods, termed spin-polarized (sp)GFNn-xTB, enable unrestricted calculations with GFNn-xTB methods and allow different spatial orbitals for α and β electrons. Due to the two different spin channels, the computational costs are around twice the costs of the unpolarized calculations under the assumption that the integral generation is negligible and the SCF convergence is similar. These methods require additional atom-wise parameters (so-called spin constants), which, unlike the original GFN parameters, were not fitted to gradients and frequencies, but obtained from numerical derivatives of spin-polarized DFT calculations on the respective atoms. For this, different DFT methods were tested and the GGA functional PBE turned out to be a robust choice.[101,124] The GFNn-xTB methods approximate an unspecified averaged GGA functional to which the non-empirical PBE functional fits well. The (sp)GFNn-xTB methods were evaluated on a set of benchmarks containing open-shell systems, namely the AC12 for singlet-triplet spin-splittings of carbenes,[125] the open-shell barrier heights of the BH76,[99] the RC21 for cationic open-shell systems relevant in the context of mass-spectrometry,[99] the YBDE18 for bond dissociation energies of Ylides,[99] and the ROST61 for open-shell organometallic reactions.[126] Here, an overall improvement was observed for both spGFN1-xTB and spGFN2-xTB. Next, the performance of the (sp)GFNn-xTB methods was compared to that of the Hartree-Fock-based SQM methods PM6-D3H4 and PM7 on a newly compiled benchmark set, termed TM90S, for the calculation of spin-splittings of TM complexes. This benchmark set utilizes DFT reference spin splittings on the TPSSh-D4/def2-QZVPP level[9,104,127] that are appropriately accurate for the evaluation of SQM methods. For this set, improvements in the mean absolute deviation (MAD) were observed for the (sp)GFN-xTB methods relative to their respective parent methods. Particularly large improvements were observed for the 3d subset containing the most complexes for which high-spin states are favored (20 out of 39), where the MADs are halved. More importantly, the qualitative correct counter that compares if the tested method gives the same sign for the spin splitting as the reference method improves for the 3d subset from 20/39 and 19/39 correct cases to 33/39 and 35/39 for spGFN1-xTB and spGFN2-xTB, respectively. In the 4d subset, only a few high-spin favoring complexes are included and the improvements are smaller. For the 5d subset, no improvements were observed. Also, vertical approaches with spGFNn-xTB methods were investigated on the TM90S set, where spGFN2-xTB slightly improved due to error compensation. To test spGFNn-xTB methods for the calculation of spin splittings of TM complexes containing two 3d metals, a divalent iron-cobalt complex[128] with an experimentally known septet ground state was calculated with spGFNn-xTB methods and low-cost DFT methods (B97-3c and TPSSh-D4/def2-SV(P)). Here, spGFNn-xTB methods obtained the qualitatively same result as DFT in only a few seconds, while DFT calculations took several hours. Overall, the spGFNn-xTB methods represent a step toward the SQM-based high-throughput screening of TM complexes. They allow a computationally inexpensive and fast check of the spin state and are valuable additions to the SQM toolbox.

CHAPTER 4

Dispersion Corrected r^{2}SCAN Based Global Hybrid Functionals: r^{2}SCANh, r^{2}SCAN0, and r^{2}SCAN50

Markus Bursch,[†] Hagen Neugebauer,[‡] Sebastian Ehlert,[‡] Stefan Grimme[‡]

Received: 21 January 2022
Published online: 4 April 2022

Reprinted in Appendix B (adapted) with permission from[§]
M. Bursch, H. Neugebauer, S. Ehlert, and S. Grimme, *Dispersion corrected r^{2}SCAN based global hybrid functionals: r^{2}SCANh, r^{2}SCAN0, and r^{2}SCAN50*, J. Chem. Phys. **156** (2022) 134105, DOI: 10.1063/5.0086040

Own contributions

- fitting of parameters for dispersion corrections
- performing all calculations on the GMTKN55 benchmark set
- interpretation of the results
- co-writing the manuscript

[†]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany
[‡]Mulliken Center for Theoretical Chemistry, Universität Bonn, Beringstr. 4, D-53115 Bonn, Germany
[§]Permission requests to reuse material from this chapter should be directed to AIP Publishing.

The strongly constrained and approximately normed (SCAN) functional[105] recovers all 17 exact constraints presently known for meta-GGA density functional approximations (DFAs), and it is one of the most accurate non-empirical DFAs. However, in real-world applications, it suffers from numerical instabilities and high sensitivity to the numerical integration grid.[129,130] Therefore, the rSCAN[129] and the r^{2}SCAN[106,107,131] regularized schemes have been proposed that improve the numerical performance at the expense of not recovering all of the 17 exact constraints (14 out of 17 for rSCAN and 16 out of 17 for r^{2}SCAN).[106,107] The r^{2}SCAN DFA has shown exceptional performance for thermochemistry and geometry optimizations of main-group and metal-organic molecules in combination with semiclassical dispersion corrections.[132] One of the largest problems of these methods is the so-called self-interaction error. To overcome this problem for the r^{2}SCAN functional, three non-empirical global hybrid DFAs with varying amounts of Hartree-Fock exchange (HFX), based on the r^{2}SCAN DFA, namely r^{2}SCANh (10 % HFX), r^{2}SCAN0 (25 % HFX), and r^{2}SCAN50 (50 % HFX), were created in this work and were provided with adjusted parameters for semiclassical London dispersion interaction corrections. The new functionals were generated in the ORCA software package, where they can be built simply via the input. Adjusted damping parameters were fitted separately for each of the three hybrid DFAs for the D4[9,113] and the D3(BJ)[111,112] dispersion corrections with the def2-QZVP and the def2-TZVPP basis set, respectively. In addition, the VV10 London dispersion correction[133] was adjusted for all three hybrids with the def2-QZVP basis set. The new DFAs with the adjusted D4 corrections were extensively tested for main-group thermochemistry, organometallic chemistry, non-covalent interactions, conformational energies, and the calculation of structural parameters (overall, about 6975 relative energies and 252 structural parameters in total). By this extensive evaluation, it was shown that a moderate admixture of 25 % HFX to r^{2}SCAN (r^{2}SCAN0) is overall most beneficial. For the main-group GMTKN55 database, r^{2}SCAN0-D4 yields an improved WTMAD-2 of 5.63 kcal·mol^{-1} compared to 7.54 kcal·mol^{-1} obtained with the r^{2}SCAN-D4 meta-GGA parent DFA.[132] For organometallic reaction energies and barrier heights, the same was found. Here, r^{2}SCAN0-D4 yields small MADs of 2.31 kcal·mol^{-1} for reactions (MOR41[134]) and 2.07 kcal·mol^{-1} for barrier height (MOBH35[135,136]) where the parent functional yields MADs of 3.32 and 3.71 kcal·mol^{-1}, respectively. Unsurprisingly, the large amount of 50% HFX was found to be beneficial for specific sets concerning the self-interaction error such as the SIE4x4 or the SIE8 subsets, but the otherwise worse performance discourages the use of r^{2}SCAN50 for general thermochemistry applications. For geometry optimizations, the hybrid DFAs with low to medium amounts of HFX show minor improvements but an overall similar performance to their parent method, while r^{2}SCAN50-D4 yields overall worse structural parameters. Since the new hybrid DFAs employ the same base DFA as r^{2}SCAN-D4 and the efficient composite method r^{2}SCAN-3c,[137] it can be expected that the potential energy surfaces of all these methods are similar making them ideal candidates for combination in multi-level workflows.[81] Additionally, because of its non-empirical nature, the r^{2}SCAN0 DFA outperforms empirical DFAs, that have only been trained for thermochemistry, for the calculation of ^{1}H and ^{13}C NMR chemical shifts.[138] The r^{2}SCAN hybrids have already been implemented by others in the TURBOMOLE package[139] and the CRYSTAL program.[140] Other programs that support hybrid DFAs and the r^{2}SCAN base functional can easily implement the r^{2}SCAN hybrids or even support building them in the input. This is the case for the ORCA[141] and the PySCF[142] program packages. Due to their overall robust performance and their relatively simple accessibility via the LibXC library,[143] the r^{2}SCAN hybrids will become more and more relevant tools in multilevel workflows and property calculations based on non-empirical functionals. Their usefulness for modelling 3d TM electrocatalysis is demonstrated in Chapter 8.

CHAPTER 5

Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT

Hagen Neugebauer[†], Peter Pinski[‡], Stefan Grimme[†], Frank Nesse[§], Markus Bursch[§]

Received: 15 August 2023
Published online: 20 October 2023

Reprinted in Appendix C (adapted) with permission[¶] from
H. Neugebauer, P. Pinksi, S. Grimme, F. Neese, and M. Bursch, *Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT*, J. Chem. Theory Comput. **19** (2023) 7695, DOI: 10.1021/acs.jctc.3c00896

Own contributions

- performing calculations on the GMTKN55 benchmark set
- performing calculations with varied amounts of Hartree-Fock exchange
- interpretation of the results
- co-writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, Universität Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]HQS Quantum Simulations GmbH, Rintheimer Straße 23, D-76131 Karlsruhe, Germany
[§]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

Double-hybrid (DH) density functionals represent the highest rung of density functional theory (DFT) methods.[67,144–146] In DH functionals, parts of the correlation energy are calculated by correlated wavefunction theory methods (WFT), typically second-order Møller-Plesset perturbation theory (MP2). The application of DH functionals is often limited for the application to large chemical systems (above 100 atoms) due to the scaling of the computational costs of $O(N^5)$ with the system size. A prominent approach to reduce this scaling is the domain-based local pair natural orbital (DLPNO) approximation that has been developed for MP2[92,147] and is well known in the context of localized coupled cluster methods (DLPNO-CCSD(T)).[59–61,148] A reduced scaling is obtained in DLPNO methods by several approximations. The most important of these approximations is the expansion of the virtual space in the truncated space of pair natural orbitals, which are combined with the domain approximation and orbital pair screening. In this work, the DLPNO-MP2 approximation has been employed in the DH scheme to yield the DLPNO-DH methods, which allows the routine application of DLPNO-DHs to system sizes above 100 atoms. To investigate the errors of the DLPNO-DHs, the prominent B2PLYP DH functional[149] was employed in the DLPNO-B2PLYP scheme. It was compared to conventional B2PLYP for 7925 thermochemistry data points and 239 structural features. Multiple thresholds are available in the DLPNO-DH methods that determine the accuracy and the computational costs. The truncation thresholds with the largest impact is the one that determines the domain size and the one that controls how many PNOs are retained for a given electron pair. For all these truncation thresholds, default keywords *loosePNO*, *normalPNO*, *tightPNO*, and *verytightPNO* are available in ORCA. The errors of DLPNO-B2PLYP for all these default settings were investigated. In addition, extrapolation schemes that are known from DLPNO-CCSD(T), which extrapolate to the complete PNO space (CPS) were investigated.[150] It was found that these schemes are also applicable to DLPNO-DHs to reduce DLPNO errors. Here, the extrapolation schemes CPS($l \rightarrow n$), CPS($n \rightarrow t$), and CPS($t \rightarrow vt$) were investigated with the nomenclature CPS($X \rightarrow Y$) with l as an abbreviation for *loosePNO*, n for *normalPNO*, t for *tightPNO*, and vt for *verytightPNO*. The evaluation revealed that the *normalPNO* setting is often sufficient to yield small errors compared to the conventional DH, except for NCIs where *tightPNO* is sometimes necessary to obtain errors below 1 kcal/mol. For geometries, the *normalPNO* settings were found to be sufficient. A look at the computational timings showed that DLPNO-B2PLYP for system sizes below 100 atoms outspeeds conventional B2PLYP in the evaluation of the molecular gradient, but not for energy calculations. This changes for larger system sizes (above 100 atoms), where DLPNO-B2PLYP also becomes faster in energy calculations. The commonly available DH functionals employ different amounts of MP2 correlation (typically around 30%) and Hartree-Fock exchange (HFX) (typically between 50 – 80%). To investigate the effects of the different admixtures in DLPNO-DHs, the non-covalent interactions of the L7[151,152] set were calculated with different B2PLYP variants, employing varied amounts of MP2 correlation and HFX (from 20 to 80% in 20% steps). It was found that the DLPNO-DH error increases linearly with the amount of MP2 correlation employed, which is not critical since relatively small amounts are typically used. A slight reduction of the DLPNO errors with increasing amounts of HFX was found. This is not surprising since the admixture of HFX reduces the self-interaction error and therefore the electron delocalization in DFT methods. DLPNO-DHs, which exploit the locality of the electron correlation, yield smaller errors for more localized electrons. In conclusion, this work presented the DLPNO-DH methods and systematically investigated their errors, which are representative for the B2PLYP functional. The errors were shown to be small and transferable to other DH admixtures. Therefore, this work paves the way for the routine application of DHs in the DLPNO-DH scheme to large system sizes above 100 atoms, which is relevant for DH calculations of 3d TM electrocatalysts with large ligand frameworks.

CHAPTER 6

Toward Benchmark-quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts - A Comparison of CCSD(T) and ph-AFQMC

Hagen Neugebauer,[†] Hung T. Vuong,[‡] John L. Weber,[‡] Richard A. Friesner,[‡] James Shee,[§] Andreas Hansen[†]

Received: 9 June 2023
Published online: 1 September 2023

Reprinted in Appendix D (adapted) with permission[¶] from
H. Neugebauer, H. T. Vuong, J. L. Weber, R. A. Friesner, J. Shee, and A. Hansen, *Toward Benchmark-Quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts: A Comparison of CCSD(T) and ph-AFQMC*, J. Chem. Theory Comput. **19** (2023) 6208, DOI: 10.1021/acs.jctc.3c00617

Own contributions

- compiling the 3dTMV benchmark set
- performing all DFT calculations and parts of the CC calculations
- statistical evaluation
- interpretation of the results
- co-writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, University of Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027
[§]Department of Chemistry, University of California, Berkeley, California 94720
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

Modeling homogeneous electrocatalysis with transition metal (TM) containing complexes by quantum chemical methods is highly desirable to support the experiment. Typical applications are the elucidation of the underlying reaction mechanisms, which enables systematic tuning of ligand structures to achieve better substrate selectivity, turnover frequencies, and lower overpotentials.[153] One step toward this is the accurate prediction of redox potentials, which is challenging for multiple reasons. First, complex bonding situations with strong dynamic correlation effects can occur in TM chemistry (e.g., carbonyl ligand with σ-donation and π-backbonding) that require a theoretical description beyond second-order perturbation theory.[84,154] Second, often highly charged systems occur, and calculations of accurate solvation free energies can be problematic for implicit and explicit solvation protocols. Last but not least, due to (nearly) degenerate orbitals and competing spin states, static correlation of different degrees can occur. In this work, the calculation of ionization potentials (IPs) for 3d TM electrocatalysts with different degrees of static correlation was investigated. For this purpose, a new benchmark set termed 3dTMV containing 28 molecules with relevance for electrocatalysis was compiled from the literature. In this set, the degree of static correlation for each initial and oxidized state was determined by multireference diagnostic tools from the literature. By investigating these diagnostics with a principal component analysis, a set of relevant diagnostics was identified, namely spin-symmetry breaking (SSB) with CCSD,[155] SSB with CCSD on PBE0[156] orbitals, SSB with PBE0, and 1-C_0^2 from CASSCF calculations. The strong correlation between spin-symmetry breaking with CCSD and PBE0 is notable, since the latter is computationally much cheaper but not well-defined. Employing SSB from DFT to identify multireference character has already been suggested in the literature,[86] and here empirical evidence for this approach is provided. With these diagnostics, three subsets were assigned, namely a single-reference (SR) set containing 12 IPs, an intermediate single-reference/multireference (SR/MR) set containing 10 IPs, and a multireference (MR) set containing 6 IPs. To cross-check the applicability of the so-called gold standard of quantum chemistry, CCSD(T),[56] for calculating the IPs, the localized-orbital (LO-)ph-AFQMC method was employed.[51–55] This method has the advantage that it is an inherent multireference method without any terms from perturbation theory. It scales with N^3 but has a large prefactor that has been significantly reduced via efficient parallel implementations on graphical processing units (GPUs). LO-ph-AFQMC strongly dependents on the (multireference) trial wavefunction that is used as a starting point. For the systems included in the SR and the SR/MR set, sufficiently large multireference trial wavefunctions were employed, while for the MR subset, huge efforts were conducted to converge away the phaseless bias introduced by not sufficiently large trial wavefunctions. With the reference LO-ph-AFQMC IPs, the CCSD(T) method with different sets of orbitals (restricted and unrestricted HF and PBE0 orbitals) was cross-checked for the 3dTMV set. As expected, significantly better agreement between CCSD(T) and LO-ph-AFQMC was observed for the SR and the SR/MR subsets, while relatively large deviations were observed for the MR subset. The best-performing set of orbitals turned out to be the UHF/RHF orbitals, where singlets were calculated with RHF and everything else with UHF orbitals. The LO-ph-AFQMC reference values also enable benchmarking of DFT methods. Larger errors were observed with DFT compared to CCSD(T), but the trend regarding the subsets stayed the same with most functionals. A promising candidate of this initial benchmark investigation is the ωB97X-V functional[157] after adapting it for TM chemistry by a system-dependent optimal tuning (OT) procedure resulting in the OT-ωB97X-V method. In conclusion, this study represents a major step toward benchmarking IPs for electrocatalysts, and it gives quantitative measures to determine degrees of static correlation for which CCSD(T) is not reliable anymore.

CHAPTER 7

Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts

Práxedes Sánchez,[§†] Bhumika Goel,[§†] Hagen Neugebauer,[§‡] Roger A. Lalancette,[†] Stefan Grimme,[‡] Andreas Hansen,[‡] Demyan E. Prokopchuk[†]

Received: 9 October 2021
Published online: 4 November 2021

Reprinted in Appendix E (adapted) with permission[¶] from
P. Sánchez, B. Goel, H. Neugebauer, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts*, Inorg. Chem. **60** (2021) 17407, DOI: 10.1021/acs.inorgchem.1c03142

Own contributions

- performing all quantum chemical calculations
- interpretation quantum chemical results
- writing and revising parts of the manuscript concerning DFT calculations
- none of the experiments or measurements were conducted by me

[†]Department of Chemistry, Rutgers University Newark, Newark, New Jersey 07102, United States
[‡]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms Universität Bonn, Bonn 53115, Germany
[§]These authors contributed equally.
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

In this work, a library of three monometallic H_2 production electrocatalysts containing electron-rich triamine-cyclopentadienyl (Cp) ligands has been synthesized and experimentally investigated using cyclic voltammetry (CV) and single-crystal X-ray diffraction by the collaboration partners. The redox potentials of the tricarbonyl precursors **1-Fe(CO)$_3^+$**, **2-Fe(CO)$_3^+$**, and **3-Fe(CO)$_3^+$** and of the corresponding electrocatalysts **1-Fe(NCME)$^+$**, **2-Fe(NCME)$^+$**, and **3-Fe(NCME)$^+$** were calculated by DFT methods. Additionally, DFT methods were employed to generate a ground-state potential energy surface map for **1-Fe(NCME)$^+$**. Based on this map, an initial ECEC (E = electrochemical step, C = chemical step) reaction mechanism for H_2 production has been proposed. A special-purpose multilevel DFT protocol was created to generate accurate free energies at computational costs that allow a large potential-energy landscape scan. In the initial step, the conformer-rotamer ensemble sampling tool (CREST)[158] was used to perform a conformation search employing the semiempirical GFN2-xTB method.[71] Because of the challenging electronic structure of the electrocatalysts with different charges and spin-multiplicities (singlet or doublet), the CREST run had to be constrained with potentials on all covalent bonds. This approach has been already suggested in ref. [122]. The conformer lowest in energy was selected and re-optimized with the efficient composite r^{2}SCAN-3c[137] DFT method in combination with the implicit DCOSMO-RS[159] solvation model for acetonitrile (MeCN). Because MeCN is a polar solvent that can potentially coordinate at the iron center of the electrocatalyst and directly influence the catalyst's ligand field, explicit solvation with a single MeCN molecule was investigated since such effects can not be described by an implicit model such as DCOSMO-RS. The acetonitrile molecule was placed at the iron and geometry optimizations were performed. If the acetonitrile coordinated to the iron center in the optimization it was kept in the simulation. More sophisticated workflows are available for the automatic generation of solvent clusters for explicit solvation,[160] but these would be computationally not feasible for extensive energy landscape scans. For the final free energies, DFT single-point energies were calculated on the optimized structures with the r^{2}SCAN-D4/def2-QZVPP[106,107,127,132] level of theory with thermostatistical corrections obtained from the single-point-hessian approach[161] computed with GFN2-xTB and the solvation contribution from COSMO-RS.[162,163] One downside of the DCOSMO-RS solvation model is that no analytic vibration frequency calculations are possible due to technical reasons. This shortcoming was overcome by employing the single-point-hessian approach that allows large computational savings for frequency calculations and yields results of satisfactory quality. The multilevel protocol is also depicted in the appendix in Figure E.7. With these free energies, redox potentials were calculated using the Nernst equation and referenced against the ferrocene/ferrocenium ($Fc^{+/0}$) redox couple, which was also calculated. Redox potentials could be calculated with decent accuracy (mean absolute deviation of 0.17 V), and the free energy map allowed to propose the following ECEC mechanism: In an initial E step **1-Fe(NCME)$^+$** is reduced to **1-Fe(NCME)**, and a haptotropic shift from η^5 to η^2 is observed. Next, the protonation occurs at the endo-Cp position resulting in **1-Fe(NCME)(*endo*−CpH)$^+$**. After the second reduction, the MeCN ligand dissociates, and a ligand-to-metal tautomerization occurs, yielding **1-FeH**. The final protonation occurs directly at **FeH**, resulting in **1-Fe(H$_2$)$^+$**. The initial catalyst is recovered, upon H_2 release and MeCN coordination. Unlike previously anticipated, the amines of the Cp ligand do not function as proton relays in this proposed mechanism, as the *endo*-Cp ring protonation proved to be the more favorable pathway. The next step is to investigate the complete reaction pathway, including transition states. For this purpose, the multilevel workflow created for the fast scanning of the free energy landscape had to be adjusted, because DFT frequency calculations are required in addition to DFT single-point energies with a hybrid functional. This has been done in Chapter 8.

CHAPTER 8

The Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes

Bhumika Goel,[†] Hagen Neugebauer,[‡] Andrew VanderWeide[†], Práxedes Sánchez,[†] Roger A. Lalancette,[†] Stefan Grimme,[‡] Andreas Hansen,[‡] Demyan E. Prokopchuk[†]

Received: 27 June 2023
Published online: 9 October 2023

Reprinted in Appendix F (adapted) with permission[¶] from
B. Goel, H. Neugebauer, A. I. VanderWeide, P. Sánchez, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes*, ACS Catal. **13** (2023) 13650, DOI: 10.1021/acscatal.3c02911

Own contributions

- performing all quantum chemical calculations
- interpretation of quantum chemical results
- writing and revising parts of the manuscript concerning DFT calculations
- none of the experiments or measurements were conducted by me

[†]Department of Chemistry, Rutgers University Newark, Newark, New Jersey 07102, United States
[‡]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms Universität Bonn, Bonn 53115, Germany
[§]These authors contributed equally.
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

The mechanism propose in the previous Chapter[†] was mainly based on a free energy map for equilibrium structures, which was generated with a computationally efficient protocol using the r^{2}SCAN-D4 functional.[5] The subsequent step after proposing the initial mechanism is the calculation of the reaction free energy barriers associated with chemical steps (C). Therefore, transition states of protonation and tautomerization reaction steps were modeled. For this purpose, frequency calculations are required to identify the transition states as saddle points on the potential energy surface (PES). The r^{2}SCAN-D4 functional is a meta-GGA and typically yields too small reaction barriers due to the self-interaction error (SIE). Therefore, it was replaced with the newly developed r^{2}SCAN0-D4 hybrid functional[2] (see Chapter 4) for the final energy evaluation. The efficient single-point Hessian scheme used in the previous study had to be replaced by the expensive analytical Hessian calculation on the r^{2}SCAN-3c level of theory. Since analytical Hessians are not available for the DCOSMO-RS solvation model, the plain COSMO solvation model had to be used for geometry optimizations.[164] Also, additional experimental data allowed further validation of the proposed reaction mechanism. A difference Fourier transform infrared (FTIR) spectra of **1-Cp^{N3}Fe(CO)$_3^+$** was recorded by our collaboration partners with an applied potential of –1.5 V vs $Fc^{+/0}$ over 180 seconds. The CO vibrational modes were compared between experimentally measured values and DFT calculated values, which allowed the identification of different species. This way, a previously not observed dimer **1-(Cp^{N3}Fe)$_2$** was identified.Calculations of the transition states revealed the second protonation step to be rate-determining and to occur directly at the hydrogen connected to the metal center in **1-Cp^{N3}FeH** to generate **1-Cp^{N3}Fe(H$_2$)$^+$**. The DFT calculated rate constants were compared to the experimentally derived rate constants for two different catalysts **1-Cp^{N3}Fe(NCMe)** and **2-Cp^{N3}Fe(NCMe)** and two different acids (Tos_2NH and $(PhSO_2)_2NH$). The change in the rate constants measured in the experiment could be reproduced by DFT with reasonable accuracy. For the first protonation step, several pathways were investigated. The direct protonation at the iron metal center could be ruled out with a free energy barrier of 27.0 kcal·mol^{-1}. It turned out that the protonation at the endo-position of the cyclopentadienyl (Cp) ligand is not possible without acetonitrile coordinated to the Fe center due to a free energy barrier of 30.0 kcal·mol^{-1}, which is reduced to 9.8 kcal·mol^{-1} by coordination of MeCN. Protonation of the -NHR group at the Cp ligand was also found to be possible, but the barrier for the ligand to metal tautomerization was found to be too high in free energy. Therefore, the endo-protonation followed by a ligand-to-metal tautomerization was determined to be the pathway to generate **1-Cp^{N3}FeH** . The intermediate **(endo-Cp^{N3}H)Fe(NCME)** could not be isolated experimentally but the isoelectronic complexes **1-(endo-Cp^{N3}H)Fe** and **2-(endo-Cp^{N3}H)Fe** could be obtained. To investigate the ligand-to-metal hydrogen transfer, the kinetics of the reaction from **2-(endo-Cp^{N3}H)Fe** to **2-Cp^{N3}FeH** could be experimentally modeled by the collaborators under pseudo-first-order conditions due to an excess of dissolved CO. A clean transition state for the ligand-to-metal transfer could not be obtained from DFT and therefore a relaxed potential surface scan was employed to estimate the calculated free energy barrier. The measured and calculated kinetics were in reasonable agreement indicating that the barrier occurs primarily due to the CO ligand dissociation. The final mechanism was put together consisting of an initial reduction step resulting in **1-Cp^{N3}Fe(NCMe)** followed by a protonation at the endo-Cp position yielding **1-(endo-Cp^{N3}H)Fe(NCMe)$^+$**. After another reduction the MeCN ligand dissociates and a ligand-to-metal hydrogen transfer occurs resulting in **1-Cp^{N3}FeH**. After another reduction step, the final protonation occurs directly at FeH resulting in **1-Cp^{N3}Fe(H$_2$)$^+$**. The catalyst is then recovered after H_2 release.

[†]The nomenclature has been updated in this publication and in the corresponding Chapters

CHAPTER 9

Summary and Outlook

The almost unlimited amount of possible combinations between transition metal (TM) centers and ligands in TM complexes allows target-oriented tailoring with specific properties.[28] An enormous chemical space has to be examined to find the combinations that yield a desired property, and computational chemistry based on quantum chemistry is a crucial tool to conduct this search efficiently. Employing complexes with Earth-abundant 3d TMs, such as iron, has become a rapidly developing field of research, especially for the catalytic reduction of CO_2[19–22] and oxygen,[24] water splitting,[25,26] and other hydrogen evolution reactions.[27] To date, the accurate quantum chemical description of these catalysts faces multiple challenges. In particular challenging is the calculation of electronic energies, and solvation free energies for highly charged complexes containing redox noninnocent ligands, which regularly occur in electrocatalysis and often lead to multireference character. Robust and accurate electronic structure methods are required in different cost/accuracy regimes ranging from semiempirical quantum mechanical (SQM) over density functional (DFT) to wave function theory (WFT) methods to perform the screening of potential candidates in reliable multilevel screening workflows.[81]

In multilevel modeling workflows, SQM methods are typically used in the first quantum chemical step. The extended tight-binding (xTB) methods are of special interest for TM chemistry, due to the availability of parameters for all elements up to Radon. Their capabilities have already been demonstrated for the structural optimization of TM complexes,[10] the generation of conformer ensembles,[122] and the exploration of reaction pathways.[123,165] Redox potentials[11] and pK_a[166] values have also been calculated with these methods. The original GFN1-xTB and GFN2-xTB methods were designed without a spin-dependent energy expression, which allows the robust treatment of TM complexes at reduced computational costs. The downside of this approach is that GFN*n*-xTB methods cannot correctly describe high-spin states. As a result, they cannot be used to scan the spin states of TM complexes, which are crucial for their correct theoretical description. With the introduction of the spin-polarized (sp)GFN*n*-xTB methods in this work, this flaw has been corrected. The spin-polarized methods have around twice the computational costs of their parent methods, but they are still several orders of magnitude faster than DFT. Therefore, they are not general replacements but special-purpose tools with their primary application in screening workflows. The spGFN*n*-xTB methods extend the xTB toolchain and allow the screening of spin-states in addition to the screening of structures and conformers. The xTB methods can also help with the critical description of solvation effects, as they have sufficiently small computational costs to describe large amounts of solvation molecules.[167]

Also, microsolvation workflows can benefit from xTB methods,[160] but a rigorous protocol for redox potentials or pK_a values in 3d TM electrocatalysis is still missing. A central problem for describing TM complexes with xTB or other SQM methods is their limited applicability to diverse bonding situations. The empirical parameters in SQM methods can only accurately describe some of the diverse bonding motifs encountered in TM chemistry for specific target properties. Part of the solution to this problem can lie in special-purpose reparametrizations of xTB or DFTB Hamiltonians, potentially combined with machine learning (ML) algorithms.[168,169] Here, reference data could be generated on the fly with relatively cheap DFT calculations, which suffer less from transferability issues. However, there are challenges for SQM methods that cannot be solved by adjusting empirical parameters, such as the treatment of negatively charged species, which require basis sets beyond the typical minimal basis set, or charge-transfer problems caused by the self-interaction error. Novel tight-binding developments employing larger basis sets[170] have recently been presented, and tight-binding methods employing range separated Hartree-Fock exchange are known.[171] The description of TM complexes by xTB methods will hopefully benefit from further developments in these directions. In summary, a general xTB method for energies, geometries, frequencies, and non-covalent interactions for TM complexes is desirable.

The next step after SQM methods in multilevel workflows typically consists of calculations with DFT methods, which are used for refinement steps and – depending on the system size – the final energy evaluation. Nonempirical DFT functional approximations have the advantage of not being designed for specific bonding motifs, which are very diverse in TM chemistry. Therefore, the strongly constrained and appropriately normed (SCAN)[105] meta-GGA, which obeys all 17 known exact constraints that a meta-GGA can, is of interest for the accurate description of TM complexes. Its regularized successor, r^{2}SCAN,[106,107] in combination with London dispersion corrections (r^{2}SCAN-D4,[132] r^{2}SCAN-rVV10[172]), and especially the composite version r^{2}SCAN-3c,[137] which was designed as a "Swiss-army knife" of quantum chemistry, are efficient tools for the description of TM chemistry. One major problem for these methods is the self-interaction error (SIE), that is inherent to meta-GGA DFT approximations. The extension of these methods to the hybrid level presented in this work partly fixes this error and resulted in the r^{2}SCANh, r^{2}SCAN0, and the r^{2}SCAN50 functionals, of which r^{2}SCAN0-D4 with 25% Hartree-Fock exchange turned out to be the best performing functional. Meanwhile, it already found application in the calculation of NMR chemical shifts,[138,173] and its robustness for modeling homogeneous TM catalysis has also been demonstrated in this work.[6] The extension to the next rung on Jacob's ladder, the double-hybrid (DH) level, has recently also been presented.[174] Initially, the "non-empirical" DHs derived from the adiabatic connection[175] were generated and optimized with minimal empiricism. Including range separated HFX has led to the promising ωPr^{2}SCAN-D4, which yields overall good performance.

In practice, the application of DHs to larger systems with more than 100 atoms is limited due to the steep $O(N^5)$ scaling with the system size in the MP2 correlation part. The domain-based local pair natural orbital (DLPNO) approximation can be employed to drastically reduce the computational scaling of MP2. A thorough investigation of the DLPNO-MP2 approximation[92,147] in DHs, resulting in DLPNO-DHs, has been presented in this work. To demonstrate its usefulness, the prominent B2PLYP DH functional was compared in the DLPNO formalism (DLPNO-B2PLYP) for different PNO-accuracy settings to the conventional B2PLYP functional, as a representative DH functional.[3] The accuracy and speed-up of DLPNO-B2PLYP can be transferred to other DH functionals since the linear dependence of the DLPNO error on the MP2 contribution and the HFX admixture in DH functionals has been demonstrated.

The DLPNO approximation also allows the routine application of DLPNO-CCSD(T) to chemical problems and can yield highly accurate results if appropriately converged.[176] Such WFT methods are only required at the end of multilevel workflows when the highest level of accuracy is necessary. To achieve this accuracy for 3d TM electrocatalysis the suitability of CCSD(T) for the system at hand has to be proven first. An elaborate investigation of the issue when CCSD(T) is not applicable due to multireference character has been conducted in this work. Additionally, the effects of different reference orbitals have been studied. For this purpose, the 3dTMV set, containing vertical ionization energies, has been compiled from the literature with a focus on relevance for 3d TM electrocatalysis, and accurate LO-ph-AFQMC calculations[51–55,177] were conducted to cross-check the performance of CCSD(T). A quantitative protocol to conduct reliable CCSD(T) calculations on this set has been proposed. The classification scheme discriminates between single-reference (SR) and multireference (MR) cases as well as an intermediate set (SR/MR). Within the SR and the SR/MR set, the errors of CCSD(T) with regard to ph-AFQMC are around 2 kcal/mol. The UHF/RHF orbitals, i.e., RHF is employed for singlet spin multiplicities and UHF is applied for other multiplicities, were found to be better performing than pure UHF, ROHF, or PBE0 orbitals. With this classification of MR character and the question of applicability answered for CCSD(T), the next step will be to study the complete basis set limit for the SR and SR/MR set where CCSD(T) was found to yield a reasonable agreement. To achieve this, localized CCSD(T) methods will be required due to the otherwise unfeasible computational demands. Converging the localization procedure will be challenging for such complicated systems, which has been shown recently by Altun *et al.*[176] By establishing these references, much insight into localization errors in CCSD(T) can be gained, and the final references will advance the development of DFT and SQM methods for 3d TM electrocatalysis since there are few experimental references available in the gas phase. For the MR subset, CCSD(T) is not applicable and alternative benchmark methods are required. A promising candidate in this regard is the LO-ph-AFQMC method with its inherent treatment of multireference character. Although it formally scales only with $O(N^3)$, its huge pre-factor restricts the application of this method to benchmarking purposes only. Additionally, the calculations are technically demanding and require much insight from expert users. To verify the accuracy of the LO-ph-AFQMC method for MR systems, calculations at the complete basis set limit will be required, and further verification by comparison to the experiment is a necessity. Once accurate electronic energies for all three correlation regimes, SR, SR/MR, and MR are available, thermostatistical and solvation contributions can be systematically investigated by comparison to experimental data in the condensed phase such as redox potentials. In particular, the solvation of charged species is a critical issue and its further investigation would be enabled by this approach.

From these developments, the computational description of catalytic cycles will benefit greatly. At the moment, the elucidation of catalytic cycles is possible by extensive SQM and DFT calculations, as demonstrated in Chapter 7 and Chapter 8. However, cross-checking the experiment and the calculations as often as possible is a necessity. Improved electronic energies and solvation contributions will reduce this dependency on experimental data and even enable routine predictions in the future. Another challenge to overcome is the extensive computational modeling that has to be performed manually by computational chemists to investigate catalytic cycles. Building different tautomers and in particular modeling transition states is still extensively time-demanding. There are efforts to automate such steps,[123,178–182] but fully automated workflows are still not established in standard applications. Also, automatic workflows are often based on SQM methods, which can yield flawed potential energy surfaces for TM complexes, especially for non-equilibrium structures. Here, automation with

the help of ML algorithms might help in the future. While ML protocols have gained more and more momentum over recent years,[183–187] QC methods remain the methodology of choice for the computational description of 3d TM complexes and electrocatalysts. Their main advantage over ML protocols is that they can be used to produce accurate reference data, while ML models themselves require reference data, which is critical for TM chemistry in which only few data is available. In this regard, QC and ML methods are not in competition with each other but complementary.

Further developments of computational workflows for electrocatalysis with 3d TM catalysts should include the routine screening of spin-states, potentially by the spGFN*n*-xTB methods or a pre-screening with these methods, the evaluation of multireference character by cheap diagnostic tools such as the spin-symmetry breaking with the PBE0 functional, explicit solvation workflows supported by xTB methods, and a self-interaction error corrected composite DFT method since this is one of the major remaining issues for the otherwise excellent performing r^2SCAN-3c composite method. Here, the ωB97X-3c method is promising, but the large-core effective core potentials employed in the basis set sometimes result in large deviations for TM thermochemistry,[188] which could be problematic. For clear MR cases, routine protocols potentially based on NEVPT2[189,190] or CASPT2[191,192] are required to achieve reasonable results without the computational costs of benchmark level calculations, such as LO-ph-AFQMC. The application of novel MR approaches like DLPNO-NEVPT2[193] and CASPT2-K[194] in combination with semi-automatic active space selection methods, such as AVAS,[195] should be investigated for the inclusion in standard protocols for cases that are not suitable for SR methods. Evaluation of such approaches can be done on the 3dTMV set as it covers a wide range of electron correlation regimes, and therefore, it can guide the development of the next generation of methods for 3d TM electrocatalysis.

Once accurate electronic and solvation free energies can be generated in all electron correlation regimes by standard protocols the special purpose tailoring of 3d TM complexes for electrocatalysis, catalysis in general, and material design will be accelerated enormously, boosting the discovery of competitive electrocatalysts with Earth-abundant 3d TMs.

APPENDIX A

High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods

Hagen Neugebauer,[†] Benedikt Bädorf,[†] Sebastian Ehlert,[‡] Andreas Hansen,[†] Stefan Grimme[†]

Received: 19 March 2023
Published online: 4 July 2023

Reprinted in Appendix A (adapted) with permission[§] from
H. Neugebauer, B. Bädorf, S. Ehlert, A. Hansen, and S. Grimme, *High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods*, J. Comput. Chem. **44** (2023) 2120, DOI: 10.1002/jcc.27185

Own contributions

- Concept and choice of methods
- performing of timing calculations
- supervision for all calculations
- interpretation of the results
- writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, University of Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]Microsoft Research AI4Science, 1118 CZ Schiphol, Netherlands
[§]Permission requests to reuse material from this chapter should be directed to Wiley Periodicals LLC.

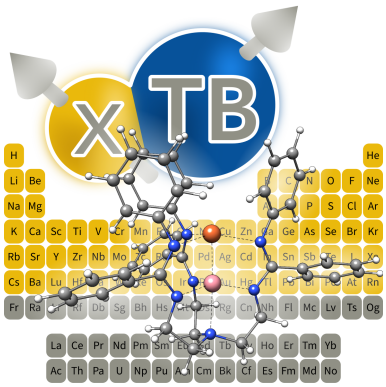

Figure A.1: Table of content graphic (ToC).

Abstract The semiempirical GFN*n*-xTB ($n = 1, 2$) tight-binding methods are extended with a spin-dependent energy term (spin-polarization), enabling the fast and efficient screening of different spin states for transition metal complexes. While GFN*n*-xTB methods inherently can not differentiate properly between high-spin (HS) and low-spin (LS) states, this shortcoming is corrected with the presented methods termed spGFN*n*-xTB. The performance of spGFN*n*-xTB methods for spin state energy splittings is evaluated on a newly compiled benchmark set of 90 complexes (27 HS and 63 LS complexes) containing 3d, 4d, and 5d transition metals (termed TM90S) employing DFT references at the TPSSh-D4/def2-QZVPP level of theory. The challenging TM90S set contains complexes with charges between −4 and +3, spin multiplicities between 1 and 6, and spin-splitting energies that range from −47.8 to 146.6 kcal/mol with a mean average of 32.2 kcal/mol. On this set the (sp)GFN*n*-xTB methods, the PM6-D3H4 method, and the PM7 method are evaluated with spGFN1-xTB yielding the lowest MAD of 19.6 kcal/mol followed by spGFN2-xTB with 24.8 kcal/mol. While for the 4d and 5d subsets small or no improvements are observed with spin-polarization, large improvements are obtained for the 3d subset with spGFN1-xTB yielding the smallest MAD of 14.2 kcal/mol followed by spGFN2-xTB with 17.9 kcal/mol and PM6-D3H4 with 28.4 kcal/mol. The correct sign of the spin state splittings is obtained with spGFN2-xTB in 89% of all cases closely followed by spGFN1-xTB with 88%. On the full set, a pure semiempirical vertical spGFN2-xTB//GFN2-xTB-based workflow for screening purposes yields a slightly better MAD of 22.2 kcal/mol due to error compensation, while being qualitative correct for one additional case. In combination with their low computational cost (scanning spin states in seconds), the spGFN*n*-xTB methods represent robust tools for pre-screening steps of spin state calculations and high-throughput workflows.

A.1 Introduction

Properties of transition metal complexes can be widely varied through ligand design and metal selection for catalysis or functional material design.[28,196] Due to the wide field of possible candidates to choose from, theoretical screening methods for the electronic spin state (multiplicity) are highly desirable. Here, correlated wave function theory (WFT)[197–203] provides systematically improvable quantum chemistry (QC) methods, which in principle can yield an accuracy of about 1-2 kcal/mol for the relative (excitation) energies. On the downside, WFT methods suffer in typical applications from their non-black-box character and their high computational costs, which makes them unfeasible for

screening purposes. On the other hand, Density functional theory (DFT) methods[67,204–206] offer a good compromise between computational costs and accuracy, but for electronically complicated transition metal complexes, DFT can give even qualitatively incorrect results. Also, machine learning (ML) models can be employed to characterize spin states, even in seconds of computation time,[207] but prior training with data points similar to the investigated system is required. The performance of QC methods is inherently much weaker system dependent. Therefore, so-called multilevel screening workflows, where at each level more expensive methods are employed and candidates are sorted out, are an attractive alternative approach. Such an approach has been successfully applied for example to conformer rankings.[81] For the initial selection stage in such multilevel workflows, semiempirical quantum mechanical (SQM) methods are of interest as they represent the cheapest class of quantum-mechanics-based methods. Of special interest are the GFNn-xTB ($n = 1, 2$) methods,[70–72] which have been consistently parameterized for all elements up to radon (Z= 86) as special purpose tools with a focus on geometries, (vibrational) frequencies, and noncovalent interactions resulting in the GFN acronym. These methods showed robustness for the fast geometry optimization of transition metal complexes and large metal-organic frameworks,[10,208] for conformer searches in transition metal complexes,[79,80,122] and for transition state localization.[165] They have also been employed successfully for off-target properties such as redox potentials[11] and pKa values.[166] We refer the interested reader to the review article ref. [72] for detailed theory on GFN methods. An important weak point of the current GFNn-xTB methods is that they do not employ a spin-dependent energy expression and thus they can not differentiate properly between high-spin (HS) and low-spin (LS) states. This disqualifies them for the screening of spin states for transition metal complexes.

To circumvent this issue we here introduce the so-called spin-polarization terms into the xTB Hamiltonian, which are known in density functional tight-binding (DFTB) methods.[209–211] We implemented the spin-polarization energy term and its nuclear gradient into the *tblite* library[212] through which it is also available in the *xtb* program package.[213] Furthermore, we calculated the required spin constants for most elements of the periodic table (1–56 and 72–86) by DFT. Similar work for spin-orbit coupling has been presented recently by Heine et al.[214] The new methods, termed spin-polarized (sp)GFNn-xTB ($n = 1, 2$), were tested for spin state-energy splittings of a newly compiled benchmark set termed TM90S containing 3d, 4d, and 5d transition metal complexes, and compared to their unpolarized parent methods as well as to the PM6-D3H4 and PM7 Hartree-Fock based SQM methods.[76,77,215]

A.2 Methodology

A.2.1 Theory

In the unpolarized formalism of the GFNn-xTB ($n = 1, 2$) methods, HS states can be calculated and differentiated from LS states by selecting the corresponding molecular orbital occupations, i.e., the number of singly and doubly occupied orbitals. However, although the GFNn-xTB energy expression is occupation-dependent, its spin independence is usually dominating and the LS state will be lower in energy. The geometries of open-shell high-spin-states can be optimized using the unpolarized formalism, however, for the calculation of spin-splittings this treatment is often not sufficient and spin-polarization is required. This is especially the case for transition metal complexes. Spin-polarized tight-binding theory is known for more than two decades.[209–211] The spin-dependent

energy contribution is calculated according to:

$$E_{\text{spin-pol}} = \frac{1}{2} \sum_{A} \sum_{l \in A} \sum_{l' \in A} p_{A,l} p_{A,l'} W_{A,ll'} \tag{A.1}$$

Here, $p_{A,l}$ is the magnetization Mulliken population on atom A for an orbital shell of angular momentum l calculated from the Mulliken spin populations q as

$$p_{A,l} = q_{A,l,\uparrow} - q_{A,l,\downarrow}. \tag{A.2}$$

Due to the dependence on the Mulliken populations, $E_{\text{spin-pol}}$ enters the spGFNn-xTB energy expression self-consistently. The constants $W_{A,ll'}$ are the so-called spin constants, which are obtained from atomic DFT calculations by numerical differentiation according to

$$W_{A,ll'} = \frac{1}{2} \left(\frac{\partial \varepsilon_{l\uparrow}}{\partial n_{l'\uparrow}} - \frac{\partial \varepsilon_{l\uparrow}}{\partial n_{l'\downarrow}} \right). \tag{A.3}$$

Here, $\varepsilon_{l\uparrow}$ are the eigenvalues and $n_{l'}$ are the occupation numbers of the atomic orbitals obtained from an unrestricted (spin-polarized) DFT treatment. These derivatives were obtained with the PBE functional[101] in the *Turbomole* 7.5.1 program package[216,217] (see SI for details and tests for spGFN2-xTB with other density functionals). The spin constants are not fitted specifically to any of the GFNn-xTB methods and can therefore be applied to all self-consistent methods of the family, such as GFN1-xTB and GFN2-xTB. In this work, the spin-polarized GFNn-xTB methods for n= 1, 2 (spGFN1-xTB and spGFN2-xTB) are evaluated in detail. In the following the term (sp)GFNn-xTB methods refers to the unpolarized and the spin-polarized GFN1-xTB and GFN2-xTB methods. All empirical parameters of the methods are unchanged.

A.2.2 Computational Details

All (sp)GFNn-xTB calculations were performed with the *tblite* v. 0.3.0 development version either as a standalone or in the *xtb* v. 6.5.1 development version. All B97-3c calculations[218] were performed in the *ORCA* 5.0.3 software package[141] employing the *DEFGRID3*, *TightSCF*, and *TightOpt* keywords for the TM90S set, while for the timing example on **FeCo**, the default grid and convergence settings were employed. For the reference single-point energy calculations on the TM90S the TPSSh functional[104] with the D4 dispersion correction model[9] and the large def2-QZVPP basis set[127] was employed (abbreviated in the following as TPSSh-D4). The D4 correction was evaluated with the *dftd4* v. 3.5.0 standalone program.[219] Robust SCF convergence for DFT calculations was ensured by employing the TRAH SCF solver.[220] The split-RI-J approximation was used to speed up DFT calculations[221] using the corresponding auxiliary basis set.[222] Hybrid DFT calculations were sped up with the RIJCOSX approximation.[223–225] PM6-D3H4 and PM7 calculations were conducted with the *MOPAC* program,[226] version 19.179.

A.2.3 The TM90S Benchmark Set

In order to evaluate the performance of spGFNn-xTB methods for the screening of spin states of transition metal complexes, a test set of 90 complexes containing 3d, 4d, and 5d transition metal atoms

termed TM90S was compiled. The charges of the molecules in the set vary from +3 to −4 and the spin multiplicities from 1 to 6. On this set, spin-splittings between high-spin (HS) and low-spin (LS) states were calculated according to

$$\Delta E_{\text{HS-LS}} = E_{\text{HS}} - E_{\text{LS}}, \tag{A.4}$$

where E refers to the total energy of the corresponding state in its optimum geometry (adiabatic energy splitting). For screening purposes, the qualitatively correct result (sign of ΔE) is especially important and therefore the number of correct signs compared to the TPSSh-D4 DFT reference is counted. The structures of the 3d subset are shown in Figure A.2.

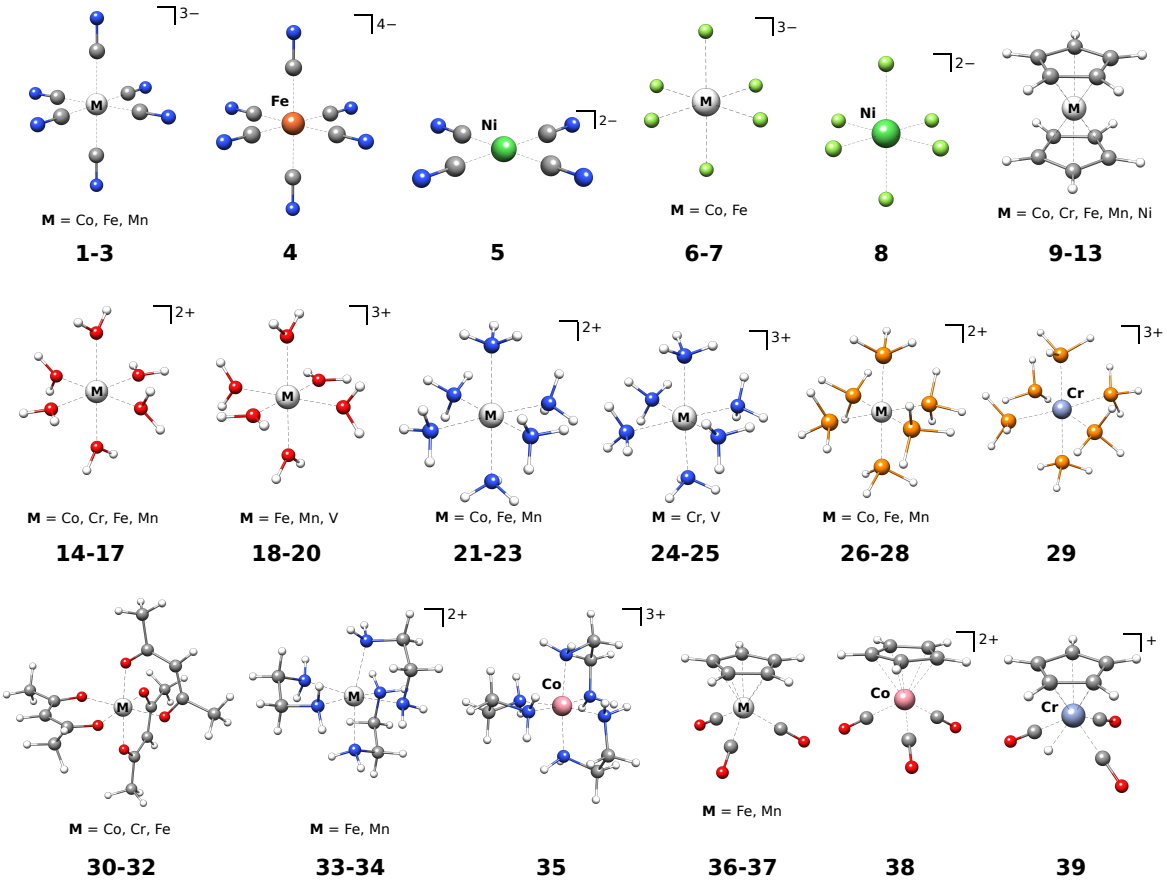

Figure A.2: Structures of the 3d subset from the TM90S benchmark set optimized with B97-3c depicted for the spin state lower in energy. Atoms are colored by element (see SI for all complexes included).

For the 4d and 5d subsets, the same ligand motifs have been employed (see SI for details). The geometries were optimized on the B97-3c level of theory and reference spin-splittings were calculated on the TPSSh-D4/def2-QZVPP level of theory. DFT errors for spin-splittings can vary largely for different systems and hybrid functionals can favor the LS or the HS state based on the amount of Fock-exchange included. The TPSSh functional was selected based on the robust performance over multiple sets in the literature[198,200,201,205] and the balanced amount of 10% Fock exchange.

Nevertheless, it should be clear that the DFT references presented here are meant for benchmarking SQM methods and not other DFT or even correlated WFT methods. In order to estimate the typical DFT differences for the TM90S we also evaluated the PBE-D4/def2-QZVPP GGA method[101] on the TM90S and obtained an MAD of 4.7 kcal/mol together with an expected positive MD of 2.6 kcal/mol (no Fock exchange favours the LS state) with respect to the TPSSh-D4 data. The statistical evaluation for (sp)GFN*n*-xTB methods for the whole set with PBE-D4/def2-QZVPP as reference is shown in Table S3. This evaluation is qualitatively the same (within one case difference) as with TPSSh-D4.

A.3 Results

First, the performance of the (sp)GFN*n*-xTB methods was evaluated on several benchmark sets containing open-shell species (with transition metal atoms in one set) to evaluate their performance for electronically relatively simple cases. Next, PM6-D3H4, PM7, GFN1-xTB, GFN2-xTB, spGFN1-xTB, and spGFN2-xTB were evaluated on the TM90S benchmark set. Initial tests showed that spGFN*n*-xTB methods do not yield improved geometries and therefore screening workflows with spGFN*n*-xTB on GFN*n*-xTB geometries were tested. Finally, computational wall times were compared for (sp)GFN*n*-xTB methods, B97-3c, and TPSSh-D4.

A.3.1 Standard Benchmarks

As an initial evaluation of the spGFN1-xTB and spGFN2-xTB methods, several benchmark sets including open-shell radicals for which high level WFT references are available for spin-splittings (AC12), reaction energies (RC21, YBDE18, ROST61) and reaction barrier heights (BH76), were used for testing. The statistical performance evaluation is shown in Table A.1 together with the parent methods GFN1-xTB, GFN2-xTB, and TPSSh-D4.

For closed-shell benchmark sets the results for spGFN*n*-xTB and GFN*n*-xTB are identical and no testing is required. On a set of 12 singlet-triplet splittings for aryl carbenes, termed AC12,[125] spGFN1-xTB yields a reduced MAD of 5.2 kcal/mol in comparison to GFN1-xTB with 11.0 kcal/mol. Similar improvements are observed for spGFN2-xTB with an MAD of 9.3 kcal/mol for GFN2-xTB and 6.5 kcal/mol for spGFN2-xTB. But more importantly, GFN*n*-xTB methods calculate all molecules to be in a low-spin singlet ground state, which is not correct for seven molecules of this test set. spGFN1-xTB and spGFN2-xTB give the correct result for all these high-spin cases and only yield one qualitatively wrong spin-splitting with a high-spin-state as the ground state (see Figure S2 in the SI). Next, three benchmark sets from the GMTKN55 benchmark database[99] were selected for testing. Namely, the BH76 set for barrier heights (only barriers involving open-shell species were included), the RC21 set containing fragmentations and rearrangements in organic radical cations, and the YBDE18 containing bond-dissociation energies of ylides. For all three sets improvements in the MAD as well as in the RMSD from GFN*n*-xTB to spGFN*n*-xTB are observed, ranging from 1-14 kcal/mol in the MAD and RMSD, showing the overall robustness of the spGFN*n*-xTB approach. In order to test the thermochemistry of open-shell transition metal-containing molecules, the ROST61 benchmark set[126] was evaluated. Here, spGFN1-xTB yields a similar result to GFN1-xTB with MADs of 14.0 and 14.1 kcal/mol respectively while spGFN2-xTB yields slightly better results than GFN2-xTB (about 1.5 kcal/mol improvement in the MAD).

Table A.1: Mean deviation (MD), mean absolute deviation (MAD), and root mean square deviation (RMSD) of (sp)GFN*n*-xTB methods and TPSSh-D4/def2-QZVPP for several benchmark sets containing open-shell molecules. One system of the ROST61 benchmark set was excluded for (sp)GFN2-xTB and spGFN1-xTB from the evaluation due to SCF convergence failure. The TPSSh-D4 Data for the ROST61 was taken from ref [126]. For the BH76 set only barrier heights involving radicals were considered.

	Method	AC12	BH76	RC21	YBDE18	ROST61
MD	**GFN1-xTB**	11.0	−19.5	20.9	11.2	1.4
	GFN2-xTB	9.3	−15.7	22.7	9.2	0.7
	spGFN1-xTB	−5.2	−13.1	15.6	−4.6	1.4
	spGFN2-xTB	−6.5	−9.4	17.4	−6.6	0.9
	TPSSh-D4	−5.8	−7.6	3.3	−4.0	0.2
MAD	**GFN1-xTB**	11.0	21.0	22.0	23.2	14.1
	GFN2-xTB	9.3	20.4	23.7	24.6	15.2
	spGFN1-xTB	5.2	15.8	17.0	9.6	14.0
	spGFN2-xTB	6.5	16.8	18.8	13.3	13.6
	TPSSh-D4	5.8	7.7	4.1	4.1	2.6
RMSD	**GFN1-xTB**	11.8	26.7	25.4	27.2	18.7
	GFN2-xTB	9.9	26.8	26.7	27. 2	19.1
	spGFN1-xTB	6.0	21.7	20.4	13.0	18.5
	spGFN2-xTB	7.0	24.2	22.6	16.2	17.1
	TPSSh-D4	5.9	8.4	4.4	4.5	3.3

A.3.2 TM90S Results

The statistic result for the TM90S benchmark set with GFN1-xTB, GFN2-xTB, spGFN1-xTB, spGFN2-xTB, PM6-D3H4, and PM7 is shown in Table A.2.

Note that GFN2-xTB yields in all cases the LS state to be lower in energy (i.e., positive spin-splitting) which is particularly noticeable for the 3d subset, because here the number of TPSSh-D4 spin-splittings with a negative sign (i.e., more favourable HS state) is higher (20 out of 39) than for the 4d (4 out of 25) or 5d (3 out of 26) set. For GFN1-xTB this is also the case, but with one exception. In the 3d subset one slightly (−1.37 kcal/mol) negative spin-splitting is obtained, that is because the evaluation is performed on B97-3c geometries and not on structures optimized with the respective method. Both GFN1-xTB and GFN2-xTB yield enormous MADs of 48.6 kcal/mol and 47.5 kcal/mol on the 3d subset respectively. On the other hand, the unrestricted PM6-D3H4 method is at least in principle able to differentiate between HS and LS states and consequently yields a better MAD of 28.4 kcal/mol, while giving the right sign for 30 out of 39 spin-splittings. For the unrestricted PM7 method, a large MAD of 45.4 kcal/mol is obtained with 27 out of 39 spin-splittings being qualitatively correct. The spGFN1-xTB and the spGFN2-xTB methods show clear improvements compared to their parent methods, with spGFN1-xTB being the best performer achieving an MAD of 14.2 kcal/mol and 33 out of the 39 spin-splittings qualitative correct closely followed by spGFN2-xTB which obtained a slightly worse MAD of 17.9 kcal/mol but 35 out of 39 cases qualitative correct. For the 4d subset spGFN2-xTB and its parent method GFN2-xTB yield the smallest MAD of 20.2 kcal/mol followed

Table A.2: Performance of GFN1-xTB, spGFN1-xTB, GFN2-xTB, spGFN2-xTB, PM6-D3H4, and PM7 on the spin splittings of the TM90S benchmark set (and its 3d, 4d, and 5d subsets). The MD, MAD, and RMSD are given in kcal/mol while the qualitative correctness tracks how often the given method obtained a qualitatively right result compared to the TPSSh-D4 reference value. Suffixes (-xTB and -D3H4) are neglected for clarity.

3d metals (20 HS)	**GFN1**	**GFN2**	**spGFN1**	**spGFN2**	**PM6**	**PM7**
MD	47.6	46.7	−4.3	−10.9	19.1	12.0
MAD	48.6	47.5	14.2	17.9	28.4	45.4
RMSD	55.8	55.6	19.5	23.5	38.8	83.8
Qualitative correct	20/39	19/39	33/39	35/39	30/39	27/39
4d metals (4 HS)						
MD	11.7	11.8	−15.3	−15.8	−11.7	−26.4
MAD	20.6	20.2	21.8	20.2	23.0	33.6
RMSD	26.1	25.8	26.5	26.8	28.9	42.8
Qualitative correct	21/25	21/25	22/25	23/25	23/25	19/25
5d metals (3 HS)						
MD	10.1	−5.7	−20.2	−39.1	−41.8	−26.0
MAD	20.8	26.3	25.6	39.5	48.7	45.3
RMSD	28.5	30.8	31.6	51.7	55.6	56.5
Qualitative correct	23/26	23/26	24/26	22/26	18/26	18/26
all metals (27 HS)						
MD	26.8	21.8	−11.9	−20.4	−7.1	−9.7
MAD	32.8	33.8	19.6	24.8	32.8	42.1
RMSD	41.9	42.2	25.1	34.3	41.6	66.5
Qualitative correct	64/90	63/90	79/90	80/90	71/90	64/90

by GFN1-xTB with 20.6 kcal/mol and spGFN1-xTB with 21.8 kcal/mol. PM6-D3H4 performs only slightly worse with an MAD of 23.0 kcal/mol and PM7 still remains the worst performer with an MAD of 33.6 and a poor qualitative correct counter (19 out of 25 cases correct) while the other methods get between 21/25 to 23/25 cases qualitative correct. In this subset only for four cases the HS state is favoured hence the difference between the six methods diminishes. This is also observed for the 5d subset where GFN1-xTB yields the smallest MAD with 20.8 kcal/mol followed by spGFN1-xTB with 25.6 kcal/mol. Compared to their (sp)GFN1 counterparts GFN2-xTB and spGFN2-xTB perform considerably worse on this subset yielding MADs of 26.3 and 39.5 kcal/mol respectively. Although spGFN1-xTB does not yield the lowest MAD, it provides the highest qualitative correct counter with 24 out of 26 cases correct, hence one more correct case than GFN1-xTB. The PM6-D3H4 and the PM7 method over stabilize the HS states and yield MADs of 48.7 kcal/mol and 45.3 kcal/mol. Also, both get only 18 out of the 26 cases qualitatively correct. In the MAD over the whole TM90S set spGFN1-xTB (MAD = 19.6 kcal/mol) outperforms spGFN2-xTB (MAD = 24.8 kcal/mol), PM6-D3H4 (MAD = 32.8 kcal/mol), GFN1-xTB (MAD = 32.8 kcal/mol), GFN2-xTB (MAD = 33.8 kcal/mol), and PM7 (MAD = 42.1 kcal/mol) and gives the correct sign for 88 % of all cases. Even though the MAD is worse with spGFN2-xTB, it achieves the qualitatively correct

result in one additional case (89 % overall). Considering this divergence in the MAD and qualitative correct counter in combination with the somewhat limited accuracy of DFT references, it can not be concluded that any of the two spGFN methods clearly outperforms the other. Problematic systems for the spGFN*n*-xTB methods include the metallocenes where the HS states are often overstabilized. We attribute this behavior to the Fermi smearing,[227] which is employed in GFN*n*-xTB methods to enable fractional orbital occupations to describe static (near-degeneracy) correlation effects and to improve the convergence of the SCF procedure.[71] In HS states of metallocenes, it smears out the beta occupations over the cyclopentadienyl (Cp) ligands and yields artificially larger magnetization Mulliken populations. Unfortunately, for these cases, the Fermi smearing can not be turned off because this results in severe SCF convergence problems. This issue seems to be less prone for piano stool complexes, which only have one Cp ligand. Another issue observed with spGFN*n*-xTB methods is that for a few HS acetylacetonate (acac) complexes the LS state is still over stabilized, however, the difference between the two states is drastically reduced compared to their parent methods. Overall the performance of spGFN*n*-xTB methods for 5d metals is somewhat worse than for 3d or 4d metals as HS states are over stabilized.

A.3.3 Screening Workflow

In the previous section, the robust performance of spGFN1-xTB and spGFN2-xTB on the TM90S set with B97-3c structures was demonstrated. For actual screening workflows, the computationally expensive geometry optimization step should as well be performed by SQM methods. Since no improvements were observed for employing spGFN*n*-xTB instead of GFN*n*-xTB for geometry optimizations, the unpolarized GFN*n*-xTB methods were used to optimize the TM90S structures without any information about the spin state given (yielding LS states). On these structures, vertical spin-splittings were calculated employing the corresponding spGFN*n*-xTB method. This workflow is the spGFN*n*-xTB-based screening approach with the smallest computational costs and it is of interest how the results deteriorate with it. The results for spGFN1-xTB and spGFN2-xTB on fully relaxed B97-3c structures and on the corresponding GFN*n*-xTB LS structures are shown in Figure A.3.

With the vertical approach, spGFN*n*-xTB methods yield more positive spin-splittings as the geometries of the HS states are not optimized. For spGFN1-xTB this results in a significantly increased MAD of 25.1 kcal/mol in comparison to that of 19.6 kcal/mol on B97-3c structures and the overall qualitative correctness decreases (73 out of 90 cases correct). However, for spGFN2-xTB the resulting MAD of 22.2 kcal/mol on GFN2-xTB structures is even an improvement compared to that of 24.8 kcal/mol on B97-3c structures and the overall qualitative correctness even improves with one case (81 out of 90 signs correct). This is because, for the 4d and the 5d subset, the vertical approach improves results in some cases, as there the high-spin-states are too low in energy with spGFN2-xTB. Slightly worse results are obtained for the 3d subset, especially for three LS cases (**1,2,4**) where spGFN2-xTB calculates too positive spin-splittings by more than 150 kcal/mol. For these cases, important geometrical changes for the HS state (elongation of M-C bond length) are neglected in the vertical approach and already an HS optimization on the GFN2-xTB level of theory significantly reduces the error. With spGFN1-xTB//GFN1-xTB this is observed only for one of these three complexes (complex **1**). Nevertheless, the vertical approach still yields qualitatively the correct sign for these spin-splittings.

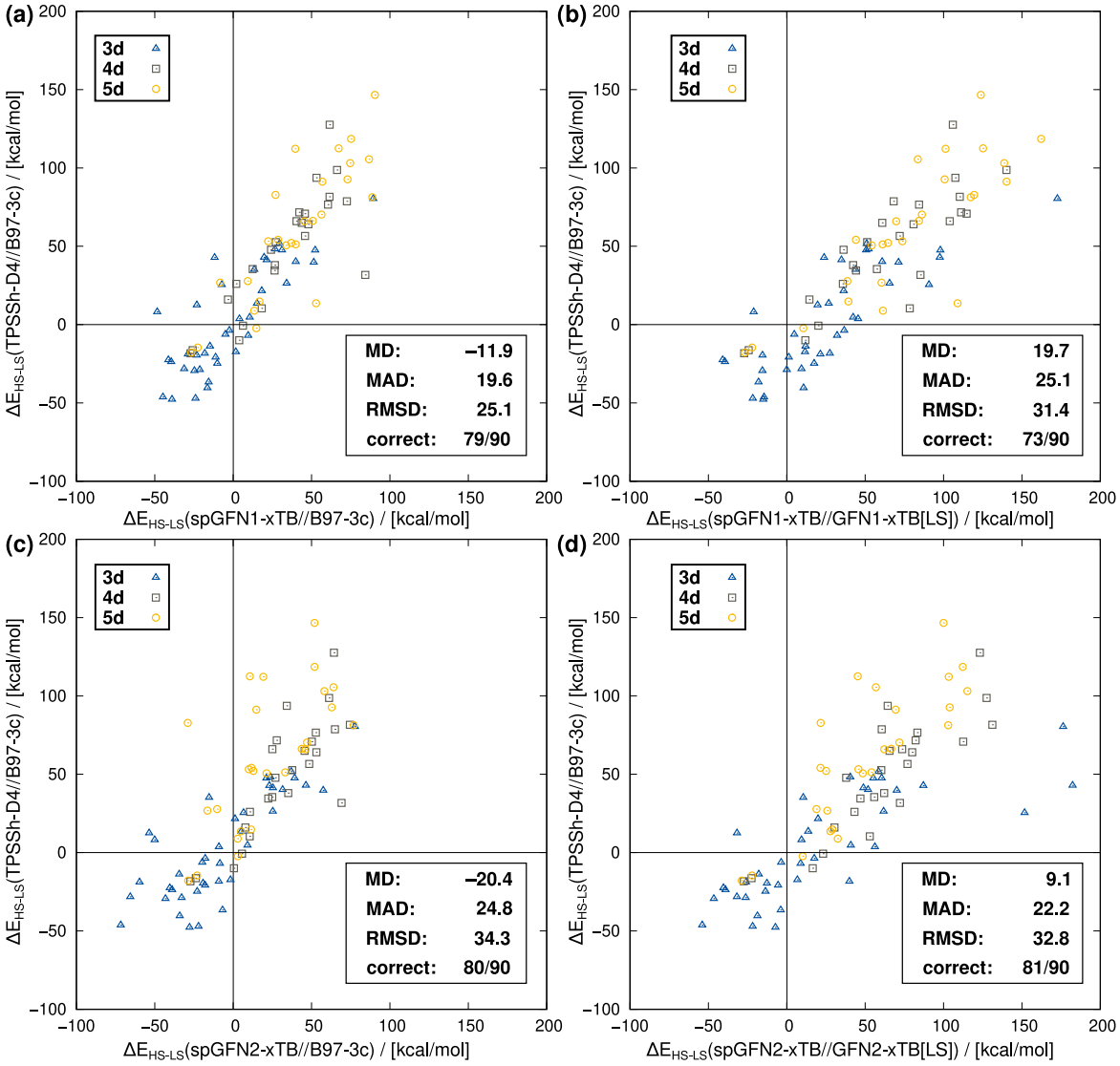

Figure A.3: Correlation plot for (**a**,**c**) the performance of spGFN1-xTB and spGFN2-xTB on B97-3c structures (spGFN*n*-xTB//B97-3c) and (**b**,**d**) vertical spin-splittings with spGFN*n*-xTB on the respective GFN*n*-xTB LS structures (spGFN*n*-xTB//GFN*n*-xTB). Both depicted against the fully relaxed TPSSh-D4/def2-QZVPP//B97-3c spin-splittings. The term "correct" corresponds to the qualitative correct spin state counter compared to the reference. Points of qualitative disagreement between the reference and the respective evaluated approach can be found in the upper left and bottom right quadrants of the plots.

A.3.4 Timings

In order to demonstrate how spGFN*n*-xTB methods can be employed to rapidly scan spin states, even in complexes with two 3d transition metals, a mixed-valence iron-cobalt complex[128] (**FeCo**) was optimized with GFN2-xTB in the septet ground state and single-point energies were calculated for the multiplicities 1, 3, 5, 7, and 9 with GFN1-xTB, GFN2-xTB, spGFN1-xTB, spGFN2-xTB, B97-3c, and TPSSh-D4/def2-SV(P) to determine the lowest lying spin state. The composite B97-3c DFT method and the smaller def2-SV(P) basis set for TPSSh-D4 were used for comparison because as computationally favourable DFT methods they could potentially be employed instead of spGFN*n*-xTB or in a follow up step in multilevel workflows. The results are shown in Table A.3 and the respective timings are shown in Figure A.4. GFN1- and GFN2-xTB yield, as expected, the singlet LS state as the ground state while spGFN1-xTB, spGFN2-xTB, B97-3c, and TPSSh-D4 correctly identify the septet to be the spin state lowest in energy. While spGFN*n*-xTB methods only take a few seconds for the five single-point energy calculations, B97-3c takes around 16.5 hours on a single CPU core (see caption for details) and TPSSh-D4 takes around 20.7 hours. spGFN1-xTB shows convergence problems with the spin multiplicity of 3, but the problem is resolved upon starting the calculation from the converged SCF solution of spin multiplicity 5. The spGFN*n*-xTB methods are both able to yield the same qualitative results and spGFN2-xTB even yields semi-quantitative spin-splittings in comparison to B97-3c and TPSSh-D4 at computational costs four orders of magnitude smaller. (sp)GFN1-xTB is generally computationally more costly than (sp)GFN2-xTB because GFN1-xTB employs an additional s-function for hydrogen to improve the description of hydrogen bonding.[70] In general calculations with spGFN*n*-xTB methods are expected to have a twofold increased computational cost compared to their unpolarized counterparts because an additional spin channel (matrix diagonalization) is introduced. This is observed for (sp)GFN2-xTB. The comparatively long computation time with spGFN1-xTB is due to difficulties in SCF convergence.

Table A.3: Spin-splittings calculated with (sp)GFN*n*-xTB methods, B97-3c, and TPSSh-D4/def2-SV(P) for the **FeCO** complex with respect to the septet ground state given in kcal/mol. The suffixes (-xTB) and the basis sets are left out for clarity.

multiplicity	GFN1	GFN2	spGFN1	spGFN2	B97-3c	TPSSh-D4
1	−32.7	−28.3	40.1	55.2	64.5	66.4
3	−30.1	−26.4	8.2	48.1	19.7	31.5
5	−21.3	−18.5	5.4	12.1	12.8	13.8
7	0.0	0.0	0.0	0.0	0.0	0.0
9	62.7	57.8	55.8	43.6	45.2	44.6

A.4 Conclusion

In this work, spin-polarization was introduced into the GFN*n*-xTB Hamiltonians yielding the spGFN*n*-xTB methods, which are available with analytical nuclear gradients in the *tblite* library through which they can also be used in the *xtb* program package. The spin constants calculated in this work can also be employed in other programs implementing xTB methods such as the *DFTB+* code.[228] The spGFN*n*-xTB ($n = 1, 2$) methods were benchmarked on main-group thermochemistry sets containing open-shell

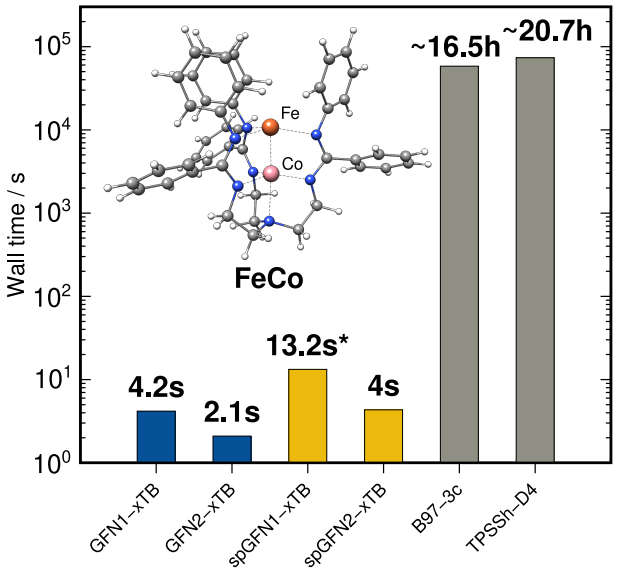

Figure A.4: Wall-time comparison with (sp)GFN*n*-xTB methods, B97-3c, and TPSSh-D4/def2-SV(P) for the single-point energy calculations shown in Table A.3. All calculations were performed on a single Intel Xeon E3-1270 v5 @ 3.60 GHz CPU core. For spGFN1-xTB the computation for spin multiplicity 3 was started from the converged result of multiplicity 5.

molecules and demonstrated overall robustness. For closed-shell systems, they yield the same results as their parent GFN*n*-xTB methods. Improvements were also observed for the ROST61 benchmark set with spGFN2-xTB for open-shell transition metal reaction energies. To evaluate the capability of spGFN*n*-xTB methods to screen spin states for transition metal complexes the new TM90S benchmark set was compiled. With its appropriately accurate DFT references the purpose of the TM90S set is to benchmark SQM methods. On this challenging set, both spin-polarized methods perform better than their respective parent methods. spGFN2-xTB improves for the 3d subset ($MAD_{GFN2\text{-}xTB}$ = 47.5 kcal/mol, $MAD_{spGFN2\text{-}xTB}$ = 17.9 kcal/mol), and stays the same for the 4d subset ($MAD_{GFN2\text{-}xTB}$ = 20.2 kcal/mol, $MAD_{spGFN2\text{-}xTB}$ = 20.2 kcal/mol), while the performance deteriorates for the 5d subset ($MAD_{GFN2\text{-}xTB}$ = 26.3 kcal/mol, $MAD_{spGFN2\text{-}xTB}$ = 39.5 kcal/mol). spGFN1-xTB also shows improvements for the 3d subset ($MAD_{GFN1\text{-}xTB}$ = 48.6 kcal/mol, $MAD_{spGFN1\text{-}xTB}$ = 14.2 kcal/mol), and yields slightly worse results for the 4d ($MAD_{GFN1\text{-}xTB}$ = 20.6 kcal/mol, $MAD_{spGFN1\text{-}xTB}$ = 21.8 kcal/mol) and the 5d subset ($MAD_{GFN1\text{-}xTB}$ = 20.8 kcal/mol, $MAD_{spGFN1\text{-}xTB}$ = 25.6 kcal/mol). This behavior is attributed to the relatively large number of HS-favoured cases in the 3d subset (20 out of 39) and the relatively low number of HS-favoured cases in the 4d (4 out of 25) and 5d subset (3 out of 26). In the MAD the spGFN1-xTB method outperformed all other SQM methods tested on the TM90S set with 19.6 kcal/mol, while spGFN2-xTB got with 80 out of 90 cases the highest score on the qualitative correct counter (89 % correct). Screening workflows where the geometries were optimized with GFN*n*-xTB ($n = 1, 2$) and vertical spin-splitting calculations were performed with the corresponding spGFN*n*-xTB method were also tested on the TM90S set obtaining a worse MAD of 25.1 kcal/mol for spGFN1-xTB and even an improved MAD of 22.2 kcal/mol for spGFN2-xTB.

To test spGFN*n*-xTB methods for a complex containing two transition metals, they were employed on an iron-cobalt complex (**FeCo**) where they determined the correct spin-ground-state in a few

seconds of computation time, while B97-3c took around 16.5 hours and TPSSh-D4 in a small basis set around 20.7 hours. Due to the relatively small performance difference between spGFN1-xTB and spGFN2-xTB, no general recommendation for one of the methods over the other can be made. Instead, the decision on which of these methods to employ should be based on the smoothness of the SCF convergence for the system at hand. For 4d and 5d complexes HS ground states are rare[229] and unless screening for HS states is desired we recommend employing the unpolarized GFN*n*-xTB methods for calculations of the spin-splittings. The benefits of spGFN*n*-xTB lie in large-scale high-throughput screening of spin states for multilevel workflows, while it also enables a quick and easy way for computational chemists to check the multiplicity of systems they investigate. Due to the higher computational costs, we do not recommend using spGFN*n*-xTB as a general replacement for GFN*n*-xTB, but as an additional tool when information about the spin state is explicitly desired. Exclusion of the spin-dependent energy term can sometimes be advantageous because smoother SCF convergence is observed without spin-polarization. Also, the spin independence of GFN*n*-xTB enables special applications such as screening of excited state minimal energy crossing points as recently demonstrated by Pracht et al.[230] Interesting future applications could also involve using the spGFN*n*-xTB wavefunctions, or properties derived from them, as an input feature vector for machine learning models or further optimization of spin constants by such techniques.[168]

A.5 Acknowledgments

The German Science Foundation (DFG) is gratefully acknowledged for financial support (Grant 1927/16-1). The authors also thank Dr. Bálint Aradi from the University of Bremen and Dr. Benjamin Hourahine from the University of Strathclyde for helpful comments on the calculation of spin constants.

A.6 Data Availability Statement

The data that support the findings of this study are available within the article and its supplementary material (available at https://doi.org/10.1002/jcc.27185). Any further information is available from the corresponding author upon reasonable request.

APPENDIX B

Dispersion Corrected r^{2}SCAN Based Global Hybrid Functionals: r^{2}SCANh, r^{2}SCAN0, and r^{2}SCAN50

Markus Bursch,[†] Hagen Neugebauer,[‡] Sebastian Ehlert,[‡] Stefan Grimme[‡]

Received: 21 January 2022
Published online: 4 April 2022

Reprinted in Appendix B (adapted) with permission from[§]
M. Bursch, H. Neugebauer, S. Ehlert, and S. Grimme, *Dispersion corrected r^{2}SCAN based global hybrid functionals: r^{2}SCANh, r^{2}SCAN0, and r^{2}SCAN50*, J. Chem. Phys. **156** (2022) 134105, DOI: 10.1063/5.0086040

Own contributions

- fitting of parameters for dispersion corrections
- performing all calculations on the GMTKN55 benchmark set
- interpretation of the results
- co-writing the manuscript

[†]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany
[‡]Mulliken Center for Theoretical Chemistry, Universität Bonn, Beringstr. 4, D-53115 Bonn, Germany
[§]Permission requests to reuse material from this chapter should be directed to AIP Publishing.

Abstract The regularized and restored semilocal meta generalized gradient approximation (meta-GGA) exchange-correlation functional r^{2}SCAN [J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, J. Phys. Chem. Lett. 11, 8208–8215 (2020)] is used to create three global hybrid functionals with varying admixtures of Hartree–Fock "exact" exchange (HFX). The resulting functionals r^{2}SCANh (10% HFX), r^{2}SCAN0 (25% HFX), and r^{2}SCAN50 (50% HFX) are combined with the semiclassical D4 London dispersion correction. The new functionals are assessed for the calculation of molecular geometries, main-group, and metalorganic thermochemistry at 26 comprehensive benchmark sets. These include the extensive GMTKN55 database, ROST61, and IONPI19 sets. It is shown that a moderate admixture of HFX leads to relative improvements of the mean absolute deviations (MADs) for thermochemistry of 11% (r^{2}SCANh-D4), 16% (r^{2}SCAN0-D4), and 1% (r^{2}SCAN50-D4) compared to the parental semi-local meta-GGA. For organometallic reaction energies and barriers, r^{2}SCAN0-D4 yields an even larger mean improvement of 35%. The computation of structural parameters (geometry optimization) does not systematically profit from HFX admixture. Overall, the best variant r^{2}SCAN0-D4 performs well for both main-group and organometallic thermochemistry and is better or on par with well-established global hybrid functionals such as PW6B95-D4 or PBE0-D4. Regarding systems prone to self-interaction errors (SIE4x4), r^{2}SCAN0-D4 shows reasonable performance, reaching the quality of the range-separated ωB97X-V functional. Accordingly, r^{2}SCAN0-D4 in combination with a sufficiently converged basis set (def2-QZVP(P)) represents a robust and reliable choice for general use in the calculation of thermochemical properties of both, main-group and organometallic chemistry.

B.1 Introduction

In the last decades, Kohn–Sham (KS) density functional theory (DFT) has emerged as a versatile and powerful tool in quantum chemistry.[97] DFT has proven to provide broad applicability towards a large variety of chemical problems at a typically excellent computational cost accuracy ratio. This has led to its status as the "working horse" of quantum chemistry and caused a massive impulse in the development of new density functional approximations (DFAs). These are usually categorized according to the "Jacob's ladder" hierarchy coined by Perdew and Schmidt in 2001.[100] The introduced rungs reflect the respective methodological improvement of DFAs, resulting in categories of local (spin-)density approximations (LDAs), generalized gradient approximations (GGAs), meta-GGAs, hybrid functionals, and double-hybrid functionals. Even though the expected accuracy of the DFAs improves ascending the rungs of Jacob's ladder, this also results in increasing computational demand. While (meta-)GGA functionals formally scale cubic with the system size (N^3), hybrid DFAs already have a formal scaling of N^4 due to the admixture of Hartree-Fock "exact" exchange (HFX) into the energy calculation. Nevertheless, for the less computationally demanding (meta-)GGA functionals the self-interaction error (SIE) is specifically problematic for the calculation of, e.g., reaction barriers. The admixture of HFX in hybrid functionals reduces the impact from SIE and therefore typically improves results for systems prone to this kind of error. Here, hybrid functionals can be classified into global hybrid functionals, applying a fixed HFX parameter, and range-separated hybrid (RSH) functionals that divide the Coulomb operator into short- and long-ranged regimes that apply different fractions of HFX. Even though RSH functionals address overdelocalization effects in the long-ranged regime more accurately compared to global hybrid functionals,[231] the robustness and computational efficiency of the latter render them still highly valuable in most quantum chemical applications. Successful variants are the well-known PBE0, PW6B95, and TPSSh hybrid functionals. An interesting

starting point for the development of new global hybrid functionals is the strongly constrained and appropriately normed (SCAN) functional[105] as it is constructed to rigorously satisfy all known exact constraints applicable to a meta-GGA. Previously proposed global hybrid functionals, like SCANh[232] (10% HFX) and SCAN0[233] (25% HFX), were developed without any correction for London dispersion interactions, which cannot be included by semilocal functionals,[110] and therefore are not competitive for real chemical applications. Martin and Santra[234] developed dispersion corrected global hybrid SCAN functionals with 10%, 25%, 37.5%, and 50% by utilizing the D4 dispersion correction.[9] One major shortcoming inherited from the SCAN functional for those hybrid functionals are the severe numerical instabilities and the resulting need to use dense computationally costly integration grids,[106,129,130] which impedes their application for many computational studies. This issue is resolved with the regularized SCAN (rSCAN)[129] and the subsequent r^{2}SCAN functional.[106,107,131] Inspired by the excellent performance of r^{2}SCAN, its London dispersion corrected variants,[132] and the composite DFT method r^{2}SCAN-3c,[137] we present three global hybrid functional variants of r^{2}SCAN termed r^{2}SCANh, r^{2}SCAN0, and r^{2}SCAN50 with 10%, 25%, and 50% of HFX admixture, respectively. Matching parameters for the D4,[9,113] the D3(BJ),[111,112] and the non-self consistent VV10[133] London dispersion correction are provided. To provide a clear picture of the capabilities of the new functionals, their performance is assessed for a variety of state-of-the-art benchmark data. These include thermochemistry, kinetics, non-covalent interactions, and molecular geometries of main-group elements (e.g., GMTKN55,[99] IONPI19[235]) and transition metal compounds (e.g., MOR41,[134] ROST61[126]).

B.2 Methods

B.2.1 Hartree–Fock exchange admixture

The recently proposed regularized and restored SCAN meta-GGA exchange-correlation functional is modified by admixture of varied amounts of HFX. The obtained global hybrid exchange-correlation functionals are constructed according to equation B.1 with a denoting the factor of HFX.

$$E_{xc}^{\mathrm{r^2SCAN}x} = (1-a)E_X^{\mathrm{r^2SCAN}} + aE_X^{\mathrm{HF}} + E_C^{\mathrm{r^2SCAN}} \tag{B.1}$$

In this work we present three variants with increasing amounts of HF exchange admixture. Referring to the well known TPSSh/0, PBE0, and BHLYP hybrid functionals, the proposed exchange correlation functionals are r^{2}SCANh (10%), r^{2}SCAN0 (25%), and r^{2}SCAN50 (50%). No re-parameterization of the original r^{2}SCAN functional was conducted.

B.2.2 Dispersion corrections

In general, the application of London dispersion corrections has proven indispensable.[8,110,236] Therefore, the recently developed atomic-charge dependent London dispersion correction D4[9,113] was parameterized for the new hybrid functionals. In the D4 correction scheme, the dispersion energy is calculated according to equation (B.2) including an Axilrod–Teller–Muto (ATM) type three-body

energy correction.

$$E_{disp}^{D4} = -\frac{1}{2}\sum_{AB}\sum_{n=6,8} s_n \frac{C_{(n)}^{AB}}{R_{AB}^{(n)}} f_{damp}^{(n)}(R_{AB}) \tag{B.2a}$$

$$-\frac{1}{6}\sum_{ABC} s_9 \frac{C_{(9)}^{ABC}}{R_{ABC}^{(9)}} f_{damp}^{(9)}(R_{ABC}, \theta_{ABC}) \tag{B.2b}$$

where $A/B/C$ are the atomic indices, R_{AB} their distance, $C_{(n)}^{AB}$ is the nth dispersion coefficient and θ_{ABC} the angle dependent term of the atomic triple. The default Becke–Johnson (BJ) damping function $f_{BJ}^{(n)}(R_{AB})$ is applied as described in equation (B.3).

$$f_{BJ}^{(n)}(R_{AB}) = \frac{R_{AB}^{(n)}}{R_{AB}^{(n)} + (a_1 R_0^{AB} + a_2)^{(n)}} \tag{B.3}$$

For a detailed description of the D4 correction see references [9] and [113]. Accordingly, the D4 model requires three functional specific parameters s_8, a_1, and a_2. The newly determined parameters for D4 as well as D3(BJ) and the parameter b for a non-self-consistent VV10 dispersion correction are are presented in table B.1.

Table B.1: Presented hybrid exchange-correlation functionals, HFX admixture, and determined parameters for the D4, D3(BJ), and VV10 London dispersion corrections fit for def2-QZVP (QZ) and def2-TZVPP (TZ) basis sets.

				D4 / QZ					D3(BJ) / QZ				VV10 / QZ
Functional	HFX	s_6	s_8	s_9	a_1	a_2	s_6	s_8	s_9	a_1	a_2	b
r^{2}SCANh	10%	1	0.8324	1	0.4944	5.9019	1	1.1236	1	0.4709	5.9157	11.9
r^{2}SCAN0	25%	1	0.8992	1	0.4778	5.8779	1	1.1846	1	0.4534	5.8972	11.4
r^{2}SCAN50	50%	1	1.0471	1	0.4574	5.8969	1	1.3294	1	0.4311	5.9240	10.8
				D4 / TZ					D3(BJ) / TZ				VV10 / TZ
Functional	HFX	s_6	s_8	s_9	a_1	a_2	s_6	s_8	s_9	a_1	a_2	b
r^{2}SCANh	10%	1	0.9119	1	0.4832	6.2073	1	1.1493	1	0.4761	6.0947	-
r^{2}SCAN0	25%	1	0.9397	1	0.4578	6.1864	1	1.1859	1	0.4567	6.0583	-
r^{2}SCAN50	50%	1	1.0576	1	0.4232	6.2378	1	1.2980	1	0.4314	6.0662	-

B.2.3 Parameterization strategy

Independently from this work, Santra and Martin[237] proposed a number of density corrected hybrids based on r^{2}SCAN and have provided D4 parameters for those methods. The major difference in the parameterization strategy employed by Santra and Martin is the use of the GMTKN55 set and the WTMAD-2 as objective function, while we followed the proven original parameterization strategy proposed in ref. [9], which minimizes the mean-square deviation for non-covalent interaction curves. Furthermore, since the interaction curves are not counter–poise corrected, the residual basis set superposition error can be accounted for in the damping parameters, especially for the triple-ζ basis sets. We prefer this strategy to provide a targeted correction of London dispersion effects and to

avoid overfitting by including systems prone to various other error sources such as the self-interaction error (SIE) in the SIE4x4 subset of the GMTKN55. The resulting parameters differ significantly, especially in the value of the s_8 dipole–quadrupole scaling. While Santra and Martin reported values around three for the s_8, we found smaller and more physical values < 1, which is in line with previous parameterizations of the rational damping function for SCAN[130] and r^{2}SCAN.[132] To investigate the effect of this discrepancy we evaluated all benchmark sets tested for this work with the parameters proposed by Santra and Martin as well.

Unsurprisingly, we find that the damping parameters proposed by Santra and Martin perform slightly better by 0.2 kcal·mol^{-1} for the GMTKN55, as this was the objective function for optimizing the damping parameters. However, we find that damping parameters optimized on the GMTKN55 will produce spurious interaction energies for large molecule benchmark sets with large dispersion contributions such as the S30L set increasing the MAD by 2.4 kcal·mol^{-1} for r^{2}SCAN0-D4 with the damping parameters by Santra and Martin. For most sets tested in this study both damping parameters yield similar results in the MAD. However, the benefit in the improved WTMAD-2 is out-weighted by the deterioration in interaction energies for large systems. Further, the tendentially smaller mean deviations using our parameterizations hint on a more systematic correction of the descriptions of London dispersion effects for the r^{2}SCAN*x* hybrid functionals. For completeness we included a full statistics of the Santra and Martin damping parameters for all tested sets in the supporting information.

B.2.4 Computational details

All quantum chemical calculations were conducted with the `ORCA` 5.0.1 program package.[141,238] DFT calculations were generally accelerated using the resolution-of-the-identity approximation[239,240] for Coulomb and exchange integrals (RIJK) applying matching auxiliary basis sets[222,241] (*def2/JK* option). If not stated else, Ahlrichs' type large quadruple-ζ def2-QZVP and def2-QZVPP basis sets[127] were applied in the following abbreviated as "QZ". Triple-ζ basis set calculations employ the def2-TZVPP basis set abbreviated by "TZ." For all basis sets, the default Stuttgart–Dresden small-core effective core potentials[242,243] (ECPs) were used for the respective elements. The numerical quadrature grid option *DefGrid3* and *TightSCF* convergence criteria were generally applied as implemented in `ORCA`. D4 London dispersion corrections were calculated with the `dftd4` 3.3.0 stand-alone program.

B.3 Results and discussion

B.3.1 Main-group thermochemistry

To evaluate the performance of the proposed global hybrid DFAs for general main group chemistry, we employ the general thermochemistry, kinetics and noncovalent interactions (GMTKN55) database.[99] The GMTKN55 database is a compilation of 55 benchmark sets and comprises 1505 relative energies divided into five categories, namely, basic properties and reactions of small systems (basic properties), isomerisations and reactions of large systems (reactions), barrier heights (barriers), intermolecular noncovalent interactions (intermol. NCIs), and intramolecular noncovalent interactions (intramol. NCIs). The comparison between r^{2}SCAN and its hybrid variants, as well as the comparison of r^{2}SCAN0-D4 to other very well performing hybrid DFAs over the five categories as well as their weighted MAD (WTMAD-2) are shown in Figure B.1. The global hybrid r^{2}SCAN DFAs yield smaller WTMAD-2 values than the meta-GGA with r^{2}SCAN0-D4 as their best performer (WTMAD-2 =

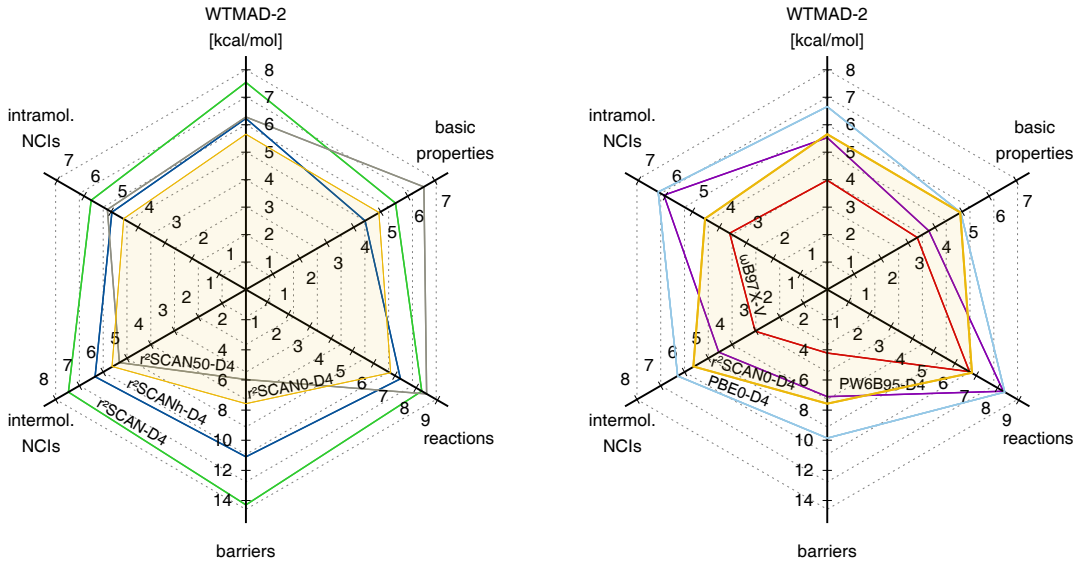

Figure B.1: Weighted mean absolute deviation of r^{2}SCANx-D4 hybrids compared to other very well performing DFAs for the large database of general main group thermochemistry, kinetics and non-covalent interactions GMTKN55. On the left side panel the different r^{2}SCANx-D4 hybrids are compared against the meta-GGA r^{2}SCAN-D4.

5.63 kcal·mol^{-1}), which is an improvement of almost 2 kcal·mol^{-1} over r^{2}SCAN-D4. The other two global hybrid DFAs r^{2}SCANh-D4 and r^{2}SCAN50-D4 perform slightly worse with WTMAD-2s of 6.22 kcal·mol^{-1} and 6.27 kcal·mol^{-1} respectively. The largest improvements with the inclusion of HFX, and therefore the reduction of SIE, are observed for barriers, while the remaining four categories benefit moderately from HFX. For the basic properties the improvements for the self-interaction error related problems (SIE4x4) set with HFX are compensated by the worse performance for total atomisation energies (W4-11). In comparison of r^{2}SCAN0-D4 with the well-performing hybrid functionals PBE0-D4, PW6B95-D4, and ωB97X-V, the r^{2}SCAN0-D4 outperforms PBE0-D4 (WTMAD-2 = 6.66 kcal·mol^{-1}), is on par with PW6B95-D4 (WTMAD-2 = 5.53 kcal·mol^{-1}), and performs moderately worse than the computationally more demanding RSH ωB97X-V (WTMAD-2 = 3.98 kcal·mol^{-1}).

Reducing the applied basis set to triple-ζ quality (def2-TZVPP) leads to moderate increases in the WTMAD-2 values of 4 (r^{2}SCAN50-D4) to 9% (r^{2}SCANh-D4, r^{2}SCAN0-D4).

B.3.2 Non-covalent interactions

Noncovalent interactions are of crucial importance in many chemical systems. Therefore, all hybrids were assessed on various benchmark sets representing diverse NCI patterns. These include NCIs of large systems (S30L, L7), ion-π interactions (IONPI19[235]), halogen bonds (X40x10[244]), hydrogen bonds (HB300SPX[245]), chalcogen bonds (CHAL336[246]), and NCIs in repulsive regions (R160x6[247,248]). Further, the subsets of the GMTKN55[99] that involve significant influence of intramolecular (IDISP, ICONF, ACONF, Amino20x4, PCONF21, MCONF, SCONF, UPU23, BUT14DIOL) and intermolecular NCIs (RG18, ADIM6, S22, S66, HEAVY28, WATER27, CARBHB12, PNICO23, HAL59, AHB21,

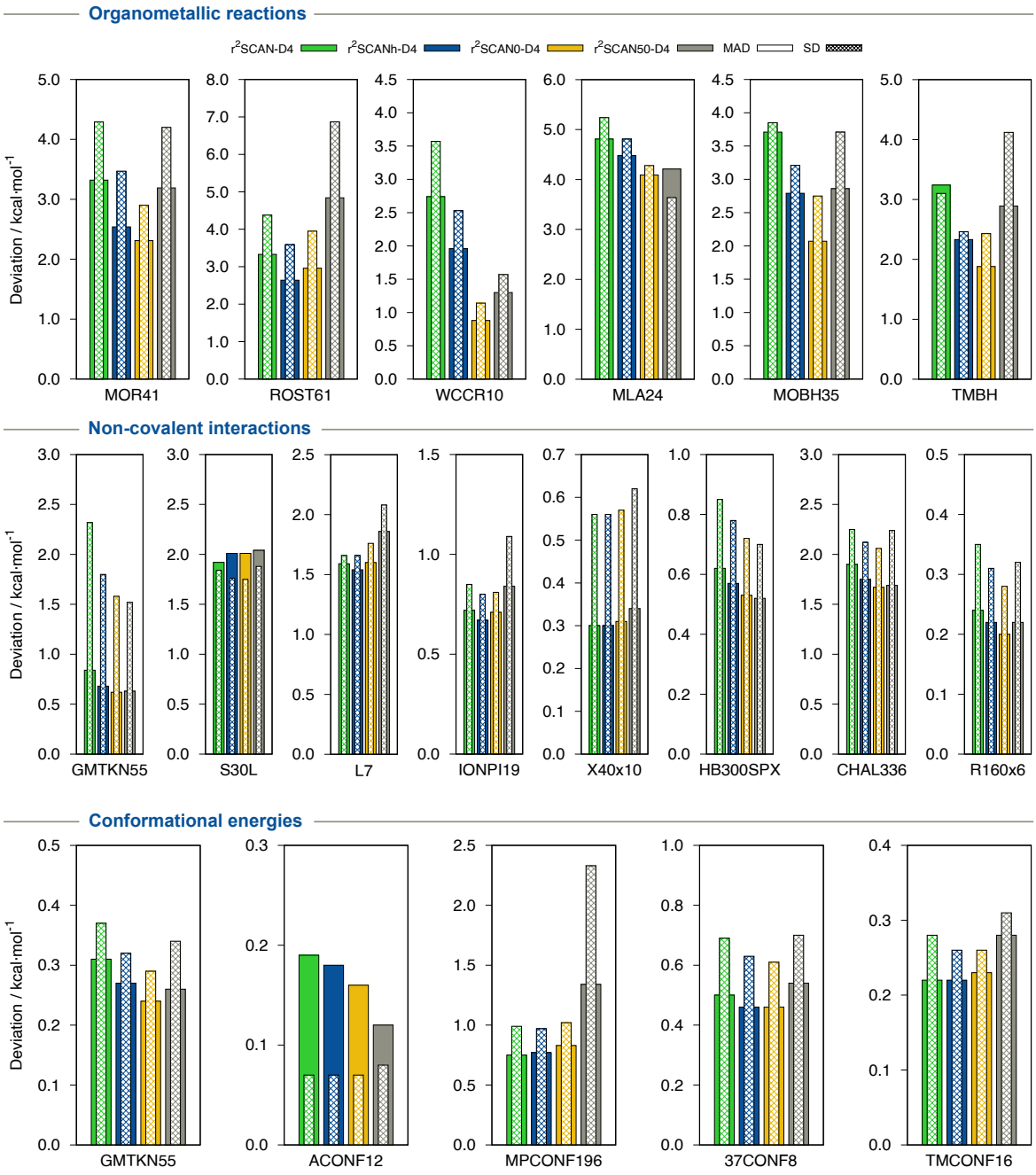

Figure B.2: Bar-plots of deviations (MAD, SD) for organometallic reactions, non-covalent interactions, and conformational energies obtained with r^{2}SCAN-D4/QZ and its r^{2}SCANx-D4/QZ hybrid (x = h, 0, 50) variants. All deviations are given in kcal·mol^{-1}.

CHB6, IL16) were assessed. The results of the proposed hybrid DFAs is depicted in Figure B.2. Finally, a comparison of r^{2}SCAN0-D4 to other very well performing DFAs is shown in Figure B.3.

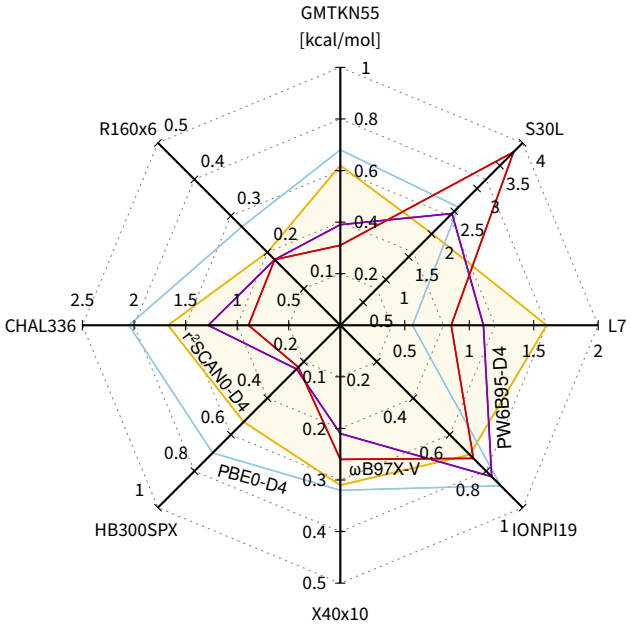

Figure B.3: Radar chart of the MADs for r^{2}SCAN0-D4 in comparison to other very well performing (RS-)hybrid DFAs. For different noncovalent interaction benchmark sets.

For the calculation of NCIs, the correct description of London dispersion effects is indispensable.[8,110] Accordingly, the parameterization of sophisticated dispersion corrections such as the D4 method to systematically correct DFT and in this case the r^{2}SCANx hybrid functionals is specifically crucial. The r^{2}SCANx-D4 hybrid functionals with moderate amounts of HFX (10, 25%) yield systematically improved results for NCIs compared to the pure meta-GGA. For r^{2}SCAN0-D4, the MAD for the NCI subsets of the GMTKN55, is reduced from 0.83 kcal·mol^{-1} to 0.62 kcal·mol^{-1}, which is an improvement of 25%. Further, the description of hydrogen bonds (HB300SPX) is improved, reducing the MAD from 0.62 kcal·mol^{-1} to 0.53 kcal·mol^{-1}. On average r^{2}SCAN0-D4 yields an improvement of 8% for NCIs. Tentatively, these positive effects of HFX admixture can be attributed to a reduced SIE leading to smaller artificial intramolecular charge-transfer. Large amounts of HFX on the other hand, lead to increased errors for some subsets such as the L7 and IONPI19 benchmark sets with increased MAD by 17% compared to r^{2}SCAN-D4. No improvement for any of the tested hybrid functionals is observed for the S30L and X40x10 (halogen bonds) benchmark sets, yet retaining the already excellent performance of the meta-GGA. Overall, the description of NCIs by the r^{2}SCANx hybrid functionals in combination with the D4 correction is on par with other very well performing global hybrid functionals such as PW6B95-D4. The excellent results obtained with the RSH ωB97X-V are not reached by any of the assessed hybrid variants. Also for NCIs the def2-TZVPP yields comparably good results for most subsets compared to a QZ quality basis set. Nevertheless, NCIs proved to be the most basis set sensitive subgroup assessed with percent increases in the MADs of up to 70% (r^{2}SCANh-D4, IONPI19) relative to the QZ value. All MAD increases due to basis set size reduction are still below 0.7 kcal·mol^{-1}. Surprisingly, for the S30L even a small improvement of the MADs

applying the def2-TZVPP basis set was observed (approx. -0.15 kcal·mol^{-1}) which is within the uncertainty of the reference values. Nevertheless, this good performance indicates that optimized hybrid/TZ level damping parameters efficiently account for some residual basis set incompleteness and superposition errors.

B.3.3 Organometallic thermochemistry

All hybrid functionals were assessed in the context of the thermochemistry of organometallic complexes. Reaction energies were assessed for closed-shell transition metal complexes on the comprehensive MOR41[134] and WCCR10[249,250] benchmark sets and for open-shell transition metal complexes on the ROST61[126] benchmark set. Reaction barrier heights of transition metal complex reactions are represented by the MOBH35[135,136] and a collection of 34 barrier heights computed by Chen and co-workers termed TMBH.[251–254] Further, the MLA24 represents a collection of alkylchains linked by a (earth) alkaline or transition metal.[255] A comparison of the obtained results to the parent r^{2}SCAN meta-GGA functional is depicted in Figure B.2. For all sets, small to moderate admixture of HFX (10%, 25%) results in smaller deviations and reduced error scattering. While 25% HFX admixture (r^{2}SCAN0) yields the best overall results, a further HFX increase to 50% (r^{2}SCAN50) increases the deviations significantly. This is in line with previous observations on increased errors upon inclusion of high amounts of HFX in the context of transition metal thermochemistry.[126,134] The frequently used M06-2X functional applying 54% HFX, shows a similar behavior compared to its sibling M06 (27%). Also, admixture of 25% HFX has proven very successful in the PBE0 functional.[156] A large improvement of the hybrid approach over the original meta-GGA is observed for the reaction barrier height subsets MOBH35 and TMBH overall decreasing the respective MADs by 44 and 42% for r^{2}SCAN0-D4. In general, specifically the r^{2}SCAN0-D4 functional reliably yields good results for transition metal complex thermochemistry (Figure B.4. For the MOR41 benchmark set, r^{2}SCAN0-D4 is only outperformed by the range-separated ωB97X-V (MAD = 2.21 kcal·mol^{-1}) and the PWPB95-D3(BJ) (MAD = 1.85 kcal·mol^{-1}) double-hybrid functional.[134] For the ROST61, containing challenging open-shell single-reference complexes, it yields a good MAD of 2.96 kcal·mol^{-1}which is close to that of ωB97X-V (MAD = 2.8).[126] An even better performance is observed for the WCCR10 benchmark set, where r^{2}SCAN0-D4 yields a very small MAD of only 0.88 kcal·mol^{-1}, which is similar to the best tested DFA PBE0 that yields a MAD of 0.83 kcal·mol^{-1} in combination with the D4 dispersion correction. For all subsets covering organometallic chemistry, also reasonable MADs can be obtained from applying the smaller def2-TZVPP basis set. The resulting MAD increases are typically below 0.2 kcal·mol^{-1}.

B.3.4 Conformational energies

The reliable computation of conformational energies is of high importance in many quantum chemical applications as finding the energetically most favored conformer is fundamental.[137] Moreover, many properties require the consideration of well described conformer-ensembles that may be generated by sophisticated conformer-ensemble sampling algorithms such as CREST.[158] The final energetic ranking of conformers often requires more accurate methods, that still keep a beneficial cost-accuracy ratio. Accordingly, the application of global hybrid functionals may be desired. The performance of the r^{2}SCANx functionals for computation of conformational energies was assessed for various conformational energy benchmarks sets. These include the MPCONF196, the 37CONF8, the

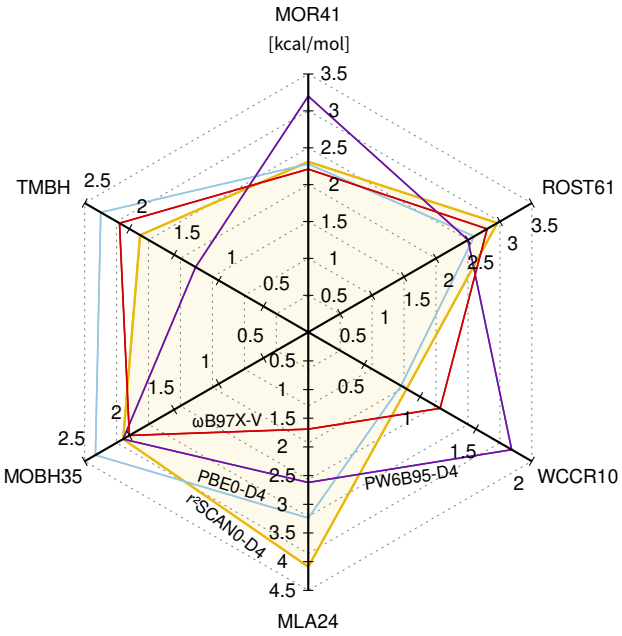

Figure B.4: Radar chart of the MADs for r^{2}SCAN0-D4 in comparison to other very well performing (RS-)hybrid DFAs for different metalorganic benchmark sets.

ACONF12, and the corresponding subsets of the GMTKN55 for main group conformers. Further, the TMCONF16 addresses conformational energies in transition metal complexes. Results are depicted in Figure B.2. A comparison of r^{2}SCAN0-D4 to other very well performing DFAs is shown in Figure B.5.

While reaction energies and barriers were shown to profit greatly from moderate HFX admixture in the framework of r^{2}SCANx global hybrid functionals, no significant improvement is observed for conformational energies. While conformational energies of alkanes (ACONF12) are improved by 5, 16, and 37%, respectively, no systematic improvement is observed for (bio-)chemically relevant molecules covered by the MPCONF196 and 37CONF8 or the transition metal complexes of the TMCONF16. Nevertheless, it is to note, that the differences in the MADs compared to r^{2}SCAN-D4 are typically very small and below 0.1 kcal·mol^{-1}. Overall specifically r^{2}SCAN0-D4 yields excellent conformational energies mostly on par with the RSH ωB97X-V. Except for the ACONF12 benchmark set, the results for conformational energies are quite insensitive to the reduction in basis set size from QZ to TZ.

B.3.5 Self-interaction Error

The artificial interaction of an electron with its own mean field is one of the major shortcomings of common KS-DFT. The so-called self-interaction error (SIE)[256–258] is specifically problematic in any local density functional approximation and in part also in various hybrid functionals. Accordingly, also the parental r^{2}SCAN meta-GGA functional is prone to SIE. Even though, several approaches to correct the SIE are known[259] the most common one is to introduce HFX in the hybrid DFT scheme. Therefore, the hybrid variants of r^{2}SCAN should show an improved performance for SIE prone systems and properties. The improvement of HFX inclusion regarding the SIE was assessed for the

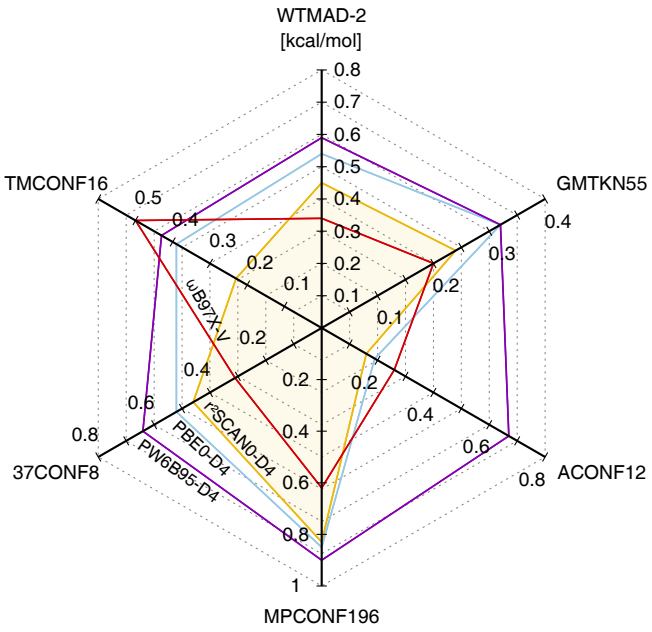

Figure B.5: Radar chart of the MADs for r^{2}SCAN0-D4 in comparison to other very well performing (RS-)hybrid DFAs for different conformational benchmark sets.

SIE4x4 and SIE8 subsets. The SIE4x4 subset of the GMTKN55 includes dissociation energies of small open-shell cationic dimers and the SIE8 consists of the remaining mostly neutral systems of the original SIE11 subset presented in the GMTKN24[260] database. The results for the r^{2}SCANx-D4 hybrid functionals are depicted in Figure B.6. For all three hybrids, a substantial improvement is observed for both subsets dependent on the amount of HFX admixture. With respect to the meta-GGA r^{2}SCAN-D4 (MAD = 18.1 kcal·mol^{-1}), the MAD is reduced to 15.2 kcal·mol^{-1}for r^{2}SCANh-D4 (10% HFX), to 10.9 kcal·mol^{-1} for r^{2}SCAN0-D4 (25% HFX), and to 4.6 kcal·mol^{-1} for r^{2}SCAN50 (50% HFX). A comparable improvement is observed for the SIE8 subset, where the MAD is reduced from 9.8 kcal·mol^{-1} (r^{2}SCAN-D4), to 8.3, 7.5, and 5.6 kcal·mol^{-1}, respectively. Even though the 50% HFX variant yields the smallest MADs for both subsets and also outperforms sophisticated DFAs such as the range-separated ωB97X-V functional,[157] the high amount of HFX has proven not beneficial for the overall performance as discussed in the previous sections. However, the best tested compromise r^{2}SCAN0-D4 with 25% HFX yields reasonable results on par with ωB97X-V and even outperforms the prominent PBE0-D4 global hybrid functional. The influence of SIE was further evaluated for a system of the IONPI19[235] benchmark set, involving the non-covalent interaction energy of the cyclopropylium cation and anthracene (Figure B.7). The interaction energy scan of both unrelaxed fragments along with the center-of-mass distance (R_{CMA}) was analyzed with reference to W1-F12 data. Here, a similar trend regarding the HFX admixture is observed as the interaction energy curve increasingly approaches the reference data. Nevertheless, for r^{2}SCAN50-D4, a beginning shift of the minimum value to a larger R_{CMA} is observed. While this is only indicated by a slight change in the shape of the curve for r^{2}SCAN50-D4, ωB97X-V already yields a different minimum R_{CMA}. Further, ωB97X-V systematically underestimates the interaction energy at shorter distances. Overall, it is shown that the HFX admixture to r^{2}SCAN-D4 significantly reduces the SIE as already indicated by

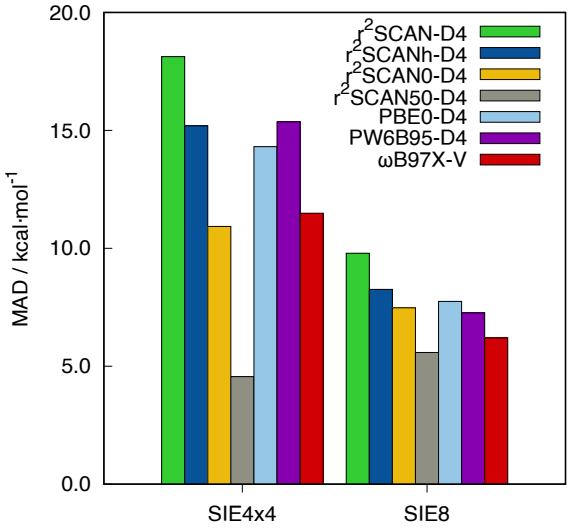

Figure B.6: MADs for the SIE4x4 and SIE8 sets. All DFT data calculated with the def2-QZVPP basis set.

the much improved performance for reaction barriers (*vide infra*). 25% HFX can be considered as the most promising compromise regarding the overall performance.

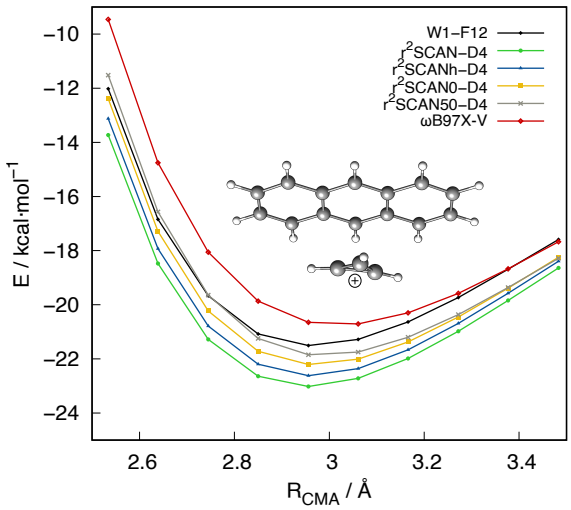

Figure B.7: Interaction energy scan of the cyclopropenyl cation and anthracene along the center-of-mass distance R_{CMA}. All DFT data calculated with the def2-QZVPP basis set.

B.3.6 Geometries

In addition to thermochemical properties, the correct description of molecular geometries is of major interest. Specifically, covalent bond lengths and angles are key structural features. Accordingly, we assessed ground-state equilibrium distances (R_e) for transition metal complexes (TMC32[261]) and heavy and light main group compounds (HMGB11,[262] LMGB35,[262] LB12[262]). Further, distances and angles in organic molecules are compared to semi-experimental reference data (CCse21[263,264]). A comparison of r^{2}SCAN-D4/QZ and its hybrid variants is depicted in Figure B.8. For geometries, an

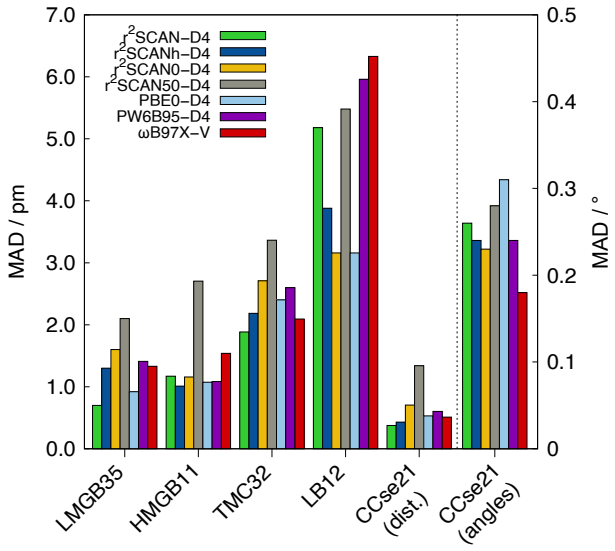

Figure B.8: Mean absolute deviations for geometries obtained with r^{2}SCAN-D4/QZ and its r^{2}SCANx-D4/QZ hybrid (x = h, 0, 50) variants compared to well-performing (RS-)hybrids. All deviations for bond lengths are given in pm, all for angles in degrees.

admixture of HFX did not prove beneficial regarding the reproduction of structural parameters. Mostly, the r^{2}SCANx-D4 hybrids yield slightly worse results compared to the already very well performing r^{2}SCAN-D4.[132] In the context of the higher computational demand of the hybrid functionals, geometry optimizations using such may not be recommended if no strong SIE effects are expected. Alternatively, the original r^{2}SCAN-D4 or its even more efficient composite variant r^{2}SCAN-3c[137] may be applied instead.

All collected MADs for all assessed thermochemistry and geometry benchmark sets are depicted in Table B.2.

B.4 Conclusion

In this study global hybrid variants of the r^{2}SCAN meta-GGA functional are assessed on a large collection of comprehensive benchmark sets such as the GMTKN55, MOR41, HB300SPX, and CHAL336 data collections (overall about 6975 relative energies and 252 structural parameters in total). The used benchmark sets cover main-group and transition metal thermochemistry, non-covalent interactions and conformational energies. Starting from r^{2}SCAN, three different hybrid functionals

Table B.2: Mean absolute deviations (MADs) / standard deviations (SDs) for the r^{2}SCAN-D4 and its hybrid variants. Energies in kcal·mol^{-1}, distances in pm, angles in °, rotational constants in MHz. All values for def2-QZVP(P) basis set with D4 dispersion correction if not noted otherwise.

Benchmark	r^{2}SCAN	r^{2}SCANh	r^{2}SCAN0	r^{2}SCAN50	ωB97X-V	PBE0	PW6B95
Organometallic chemistry							
MOR41	3.32	2.54 / 3.47	2.31 / 2.90	3.19 / 4.20	2.21	2.28	3.20
ROST61	3.33	2.64 / 3.59	2.96 / 3.95	4.84 / 6.87	2.8	2.6	2.5
WCCR10	2.74	1.96 / 2.53	0.88 / 1.14	1.30 / 1.57	1.18	0.83	1.82
MOBH35	3.71	2.79 / 3.21	2.07 / 2.75	2.86 / 3.71	2.00	2.38	2.07
TMBH	3.24	2.33 / 2.46	1.88 / 2.43	2.89 / 4.12	2.11	2.32	1.26
MLA24	4.81	4.48 / 4.81	4.09 / 4.28	4.21 / 3.64	1.69	3.24	2.62
Non-covalent interactions							
GMTKN55[†]	7.54	6.22 / —	5.63 / —	6.27 / —	3.98	6.66	5.53
L7	1.59	1.54 / 1.66	1.60 / 1.76	1.86 / 2.08	0.86	1.56	1.11
S30L	1.92	2.01 / 1.76	2.01 / 1.75	2.04 / 1.88	3.83	2.59	2.45
IONPI19	0.72	0.67 / 0.80	0.71 / 0.81	0.84 / 1.09	0.73	0.88	0.83
X40x10	0.30	0.30 / 0.56	0.31 / 0.57	0.34 / 0.62	0.26	0.32	0.21
R160x6	0.24	0.22 / 0.31	0.20 / 0.28	0.22 / 0.32	0.18	0.27	0.18
HB300SPX	0.62	0.57 / 0.78	0.53 / 0.72	0.52 / 0.70	0.23	0.70	0.24
SIE8[‡]	9.79	8.26 / 8.15	7.48 / 7.23	5.58 / 7.06	6.20	7.75	7.26
CHAL336	1.90	1.75 / 2.12	1.67 / 2.06	1.69 / 2.24	0.89	2.05	1.28
Conformational energies							
ACONF12	0.19	0.18 / 0.07	0.16 / 0.07	0.12 / 0.08	0.26	0.19	0.67
MPCONF196	0.75	0.77 / 0.97	0.83 / 1.02	1.34 / 2.33	0.62	0.85	0.90
37CONF8	0.50	0.46 / 0.63	0.46 / 0.61	0.54 / 0.70	0.31	0.52	0.64
TMCONF16	0.22	0.22 / 0.26	0.23 / 0.26	0.28 / 0.31	0.50	0.39	0.43
Structural parameters							
LMGB35	0.68	1.29 / 2.30	1.59 / 2.28	2.09 / 2.45	1.33	0.92	1.41
HMGB11	1.17	1.01 / 1.21	1.16 / 1.06	2.71 / 1.15	1.54	1.07	1.09
TMC32	1.89	2.18 / 1.80	2.71 / 1.96	3.37 / 2.73	2.09	2.40	2.60
ROT34 (MHz)	4.64	3.00 / 3.94	10.46 / 7.40	24.15 /16.28	7.71	3.86	11.71
LB12	3.57	3.88 / 5.77	3.16 / 5.05	5.48 / 9.14	6.33	3.16	5.96
CCse21 (dist.)	0.38	0.43 / 0.57	0.70 / 0.71	1.34 / 1.02	0.51	0.53	0.60
CCse21 (angles)	0.26	0.24 / 0.29	0.23 / 0.28	0.28 / 0.38	0.18	0.31	0.24

with varying amounts of HFX admixture are constructed. The new global hybrid functionals are termed r^{2}SCANh (10% HFX), r^{2}SCAN0 (25%), and r^{2}SCAN50 (50%). The sophisticated D4 London dispersion correction is parameterized for all three hybrid functionals and an additional parameter set for use with a smaller triple-ζ basis set (def2-TZVPP) is provided. Additionally, we extensively tested the D4 damping parameters proposed by Santra and Martin for the r^{2}SCAN hybrids, which perform similarly well over a wide range of assessed benchmark sets. Nevertheless, the herein presented parameterization yields more systematic and consistent improvements indicated by reduced mean deviations and the absence of crucial outliers, specifically for dispersion dominated test sets such as the S30L.

It is shown that admixture of moderate amounts of HFX to r^{2}SCAN is beneficial for most assessed properties. For the GMTKN55 database, r^{2}SCAN0-D4/QZ yields a WTMAD-2 of 5.63 kcal·mol^{-1} compared to 7.54 kcal·mol^{-1} obtained with the parent r^{2}SCAN-D4/QZ meta-GGA. The same holds for organometallic reactions and reaction barrier heights, where small MADs of 2.31 kcal·mol^{-1} (MOR41) and 2.07 kcal·mol^{-1} (MOBH35) are obtained with r^{2}SCAN0-D4/QZ. Large amounts of 50% HFX were found to be beneficial only in very specific cases such as SIE prone systems covered by the SIE4x4 and SIE8 subsets.

For geometry optimization, no benefit of applying any r^{2}SCANx-D4 hybrid variant over the parent r^{2}SCAN-D4 meta-GGA was observed. In general, while the admixture of pure HFX proved beneficial, it yields comparably minor improvements for an already excellent performing and robust meta-GGA functional such as r^{2}SCAN. Therefore, the global r^{2}SCAN0-D4 hybrid functional applying 25% of HFX has proven to perform best regarding its broad applicability. It performs robustly for a variety of properties on par with other excellent performing functionals such as PW6B95-D4 and typically outperforms the prominent non-empirical PBE0-D4 functional. Because the underlying r^{2}SCAN functional remains unchanged, it can be expected that the potential energy surface is similar to that of r^{2}SCAN-D4 and r^{2}SCAN-3c making it a robust choice for multi-level protocols based on those functionals. Accordingly, the assessed r^{2}SCAN0-D4 global hybrid functional represent an efficient alternative to the still slightly more accurate RSH functionals such as ωB97X-V. It may be applied whenever the RSH cannot be utilized due to technical reasons. Overall, r^{2}SCAN0-D4 can be considered as robust and reliable choice for a variety of computational chemistry applications.

B.5 Supplementary material

See the supplementary material for statistical data, functional availability information, and sample inputs for ORCA. All data is also openly available at https://github.com/haneug/r2scanx-D4.

B.6 Conflict of interest

There are no conflicts of interest to disclose.

B.7 Acknowledgments

The German Science Foundation (DFG) is gratefully acknowledged for financial support (Grant 1927/16-1). Further, SG and MB gratefully acknowledge financial support of the Max Planck Society through the Max Planck fellow program. We further thank Martin Blaško and Prof. Miroslav Urban for providing support with the MLA24 benchmark.

APPENDIX C

Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT

Hagen Neugebauer[†], Peter Pinski[‡], Stefan Grimme[†], Frank Nesse[§], Markus Bursch[§]

Received: 15 August 2023
Published online: 20 October 2023

Reprinted in Appendix C (adapted) with permission[¶] from
H. Neugebauer, P. Pinksi, S. Grimme, F. Neese, and M. Bursch, *Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT*, J. Chem. Theory Comput. **19** (2023) 7695, DOI: 10.1021/acs.jctc.3c00896

Own contributions

- performing calculations on the GMTKN55 benchmark set
- performing calculations with varied amounts of Hartree-Fock exchange
- interpretation of the results
- co-writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, Universität Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]HQS Quantum Simulations GmbH, Rintheimer Straße 23, D-76131 Karlsruhe, Germany
[§]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

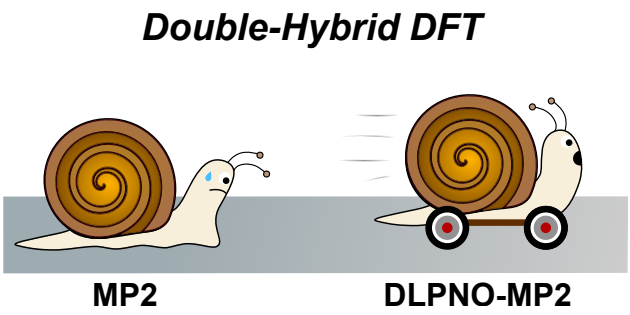

Figure C.1: Table of content graphic (ToC).

Abstract The unfavorable scaling (N^5) of conventional second-order Møller-Plesset theory (MP2) typically prevents the application of double-hybrid (DH) density functionals to large systems with more than 100 atoms. A prominent approach to reduce the computational demand of electron correlation methods is the domain-based local pair natural orbital (DLPNO) approximation that is successfully used in the framework of DLPNO-CCSD(T). Its extension to MP2 [P. Pinski, C. Riplinger, E. F. Valeev and F. Neese, *J. Chem. Phys.* **143**, 034108 (2015)] paved the way for DLPNO-based DH (DLPNO-DH) methods. In this work, we assess the accuracy of the DLPNO-DH approximation compared to conventional DHs on a large number of 7925 data points for thermochemistry and 239 data points for structural features, including main-group and transition-metal systems. It is shown, that DLPNO-DH-DFT can be applied successfully to perform energy calculations and geometry optimizations for large molecules at a drastically reduced computational cost. Furthermore, PNO space extrapolation is shown to be applicable, similar to its DLPNO-CCSD(T) counterpart, to reduce the remaining error.

C.1 Introduction

Kohn-Sham density functional theory (DFT) is widely considered the work-horse of modern computational chemistry. Within the zoo of density functionals available, double-hybrid (DH) functionals typically represent the most accurate approaches.[67,144–146] The most common DH functionals employ an admixture of the correlation energy with a fraction a_C computed using second-order perturbation theory (PT2) into the correlation energy expression of the respective density functional (eq C.1) according to

$$E_{XC}^{DH} = (1 - a_X)E_X^{DFT} + a_X E_X^{HF} + (1 - a_C)E_C^{DFT} + a_C E_C^{PT2}. \tag{C.1}$$

One of the first and most prominent DH functionals is Grimme's B2PLYP functional[149] that employs an 27% ($a_C = 0.27$) admixture of second-order Møller-Plesset perturbation theory (MP2) correlation energy and 53% ($a_X = 0.53$) of "exact" Hartree-Fock exchange (HFX).

A critical downside of the MP2-based DH approach is its comparably high computational demand as common MP2 formally scales with $O(N^5)$ of the system size. Accordingly, approaches to reduce the computational cost of the MP2 part of the DH calculation without losing significant accuracy

are desirable. Local wave-function based correlation methods have proven highly successful in this respect. They exploit the spatial locality of electron correlation by truncation of the virtual orbital space thus drastically reducing the number of considered orbitals. The most prominent representative of this class is the domain-based local pair natural orbital (DLPNO) approach that is frequently used in the framework of coupled cluster calculations (e.g., DLPNO-CCSD(T)).[59–61,148,265] The DLPNO approach can also be applied to MP2 calculations which renders DLPNO-MP2 a promising candidate to use in the context of DH-DFT.[92] The resulting DLPNO-DH scheme is available for energies, geometric gradients for closed-shell systems,[147] polarizabilities, and NMR shieldings.[266]

The efficiency of local methods in the context of DHs has already been demonstrated for main-group thermochemistry for localized pair natural orbitals in combination with F12 explicit correlation by Mehta and Martin.[267] But thorough studies for DLPNO-DHs that investigate the chemical space beyond the GMTKN55 and also consider organometallic compounds are missing. In the following, the DLPNO-MP2 implementation in the ORCA quantum chemistry software package[141,268] is employed for B2PLYP as a representative DH functional resulting in the DLPNO-DH method DLPNO-B2PLYP. Its performance is evaluated against the conventional MP2-based B2PLYP functional for a selection of comprehensive benchmark sets for thermochemistry and molecule geometries.

C.2 Methods

C.2.1 DLPNO Accuracy Settings

DLPNO correlation methods are based on the decomposition of the total correlation energy into contributions from electron pairs (pair correlation energies). Two principle approximations lead to linear scaling and high efficiency: (1) elimination of negligible electron pairs based on a highly efficient pre-screening process that is based on the asymptotic expansion of the pair correlation energy. (2) Restriction of the virtual space to a local space spanned by projected atomic orbitals (PAOs) as well as compaction of this space through the natural orbital expansion for each electron pair separately. The accuracy of the first approximation is determined by a domain threshold T_{CutDO} and the second by the pair natural orbital threshold T_{CutPNO}. The values of these two thresholds balance the accuracy of the approximation versus the computational cost. At default thresholds, typically more than 99.9% of the canonical correlation energy are recovered. For a detailed description of the DLPNO approximation in the context of MP2, we refer the interested reader to the literature.[92,147,265] Similar to DLPNO-CCSD(T),[269] default accuracy settings for DLPNO-MP2 are available in ORCA. These settings are employed for DLPNO-DH calculations as well, and the key truncation thresholds are shown in Table C.1 (see Table S1 and S2 in the SI for all truncation thresholds). Albeit *loosePNO* is not meant for accurate DLPNO-MP2 or DLPNO-DH calculations, but rather for exploratory calculations, it was tested here because it is available in ORCA via a simple keyword and is relevant in the context of PNO-space extrapolation. In contrast to DLPNO-CCSD(T), the accuracy thresholds are generally tighter. Additionally, compared to restricted references (RHF/RKS) tighter settings are required for unrestricted calculations (UHF/UKS). Therefore, in benchmark sets involving open-shell systems the tighter thresholds were used for all systems including closed-shell systems.

For a fair assessment of the error introduced by the DLPNO-DH approximation, only errors with reference to the conventional MP2-based DH functional are discussed in the following. This means that no deviations from the original reference data of the investigated benchmark sets are discussed. The error is calculated according to eq C.2:

$$\Delta x^{B2PLYP} = x^{B2PLYP}_{DLPNO-MP2} - x^{B2PLYP}_{MP2}. \tag{C.2}$$

The resulting mean absolute deviation with regard to the conventional DH (MAD_C) is calculated as

$$MAD_C = \frac{1}{n}\sum_{i}^{n}(\left|\Delta x_i^{B2PLYP}\right|). \tag{C.3}$$

The MAD_C values are then employed to calculate the weighted mean absolute deviation (WTMAD-2_C) according to

$$\text{WTMAD-}2_C = \frac{56.17\ \text{kcal}\cdot\text{mol}^{-1}}{\sum_i^{55} N_i}\sum_{i}^{55} N_i \frac{MAD_{C,i}}{\overline{|\Delta E|}_i}. \tag{C.4}$$

Here, 56.17 kcal·mol^{-1} is the average of the average absolute energies $|\Delta E|_i$ with the reference (B2PLYP) over all 55 sets of the GMTKN55, and N_i is the number of reactions with the $MAD_{C,i}$ for the corresponding set i (See SI for details).

Table C.1: PNO key accuracy settings for DLPNO-DHs.

PNO-Settings	T_{CutDO} RKS/UKS	T_{CutPNO} RKS	T_{CutPNO} UKS
loosePNO	$2 \cdot 10^{-2}$	10^{-7}	10^{-8}
normalPNO	$1 \cdot 10^{-2}$	10^{-8}	10^{-9}
tightPNO	$5 \cdot 10^{-3}$	10^{-9}	10^{-10}
verytightPNO	$2.5 \cdot 10^{-3}$	10^{-10}	10^{-11}

C.2.2 PNO-Space Extrapolation

The computational cost of any DLPNO-MP2 or DLPNO-DH calculation increases drastically upon tightening the T_{CutPNO} threshold. Accordingly, an extrapolation of the PNO space is desirable to obtain high accuracy at reduced computational cost. The extrapolation to the complete PNO space (CPS) was successfully applied in the framework of local coupled cluster following eq C.5.[150] Here, E^X and E^Y are the energies (or properties) obtained with the respective T_{CutPNO} thresholds (e.g., $X = 8$ for *normalPNO* with $T_{CutPNO} = 10^{-8}$ and $Y = 9$ for *tightPNO* with $T_{CutPNO} = 10^{-9}$), F is an empirical scaling parameter and E^{XY} is the extrapolated energy:

$$E^{XY} = E^X + F \cdot (E^Y - E^X). \tag{C.5}$$

Furthermore, it has been shown that the CPS extrapolation reduces the size dependency of the DLPNO error in the context of DLPNO-CCSD(T).[270] In a recent study by Kubas *et al.* a DLPNO-MP2 based extrapolation scheme for DLPNO-CCSD(T) has been proposed.[271] Its good performance suggests that the DLPNO errors for MP2 and CCSD(T) are rather similar and that CPS extrapolation with a similar F parameter should be beneficial for DLPNO-MP2 and DLPNO-DHs as well. Therefore, in the following, the same F parameter ($F = 1.5$) that has been used for the DLPNO-CCSD(T) CPS extrapolation[150] was assessed for CPS extrapolation in DLPNO-B2PLYP. In this work, $F = 1.5$ proved

suitable also for DLPNO-DH calculations supporting the findings of Kubas *et al.* In the following, the nomenclature for CPS extrapolation will be CPS($X \rightarrow Y$) with *l* as abbreviation for *loosePNO*, *n* for *normalPNO*, *t* for *tightPNO*, and *vt* for *verytightPNO*.

C.2.3 Computational Details

All calculations were performed with ORCA version 5.0.4[141,268] employing the B2PLYP DH functional[149] either with the DLPNO approximation (DLPNO-B2PLYP) or with the conventional resolution of the identity (RI)-B2PLYP method[272,273] in combination with the def2-TZVPP triple-ζ basis[127,222] with the corresponding def2-TZVPP/C auxiliary basis. As integration grid for the DFT calculations the large *DEFGRID3* was employed and for the SCF *TightSCF* settings were selected. Additionally, the Split-RI-J[221] and RIJCOSX[225] approximations were used to speed up the calculations. The frozen core approximation with default settings was used throughout.

C.3 Results and Discussion

Table C.2: WTMAD-2_C of DLPNO-B2PLYP on the GMTKN55 database in kcal·mol^{-1}. and CPS extrapolation is indicated by an arrow.

Set	#	*loose*	*normal*	*tight*	*verytight*	*l*→*n*	*n*→*t*	*t*→*vt*
basic	473	0.06	0.03	0.02	0.02	0.03	0.02	0.02
reactions	243	0.47	0.15	0.07	0.05	0.09	0.05	0.04
barriers	194	0.10	0.04	0.04	0.05	0.05	0.05	0.05
inter NCIs	304	1.28	0.47	0.21	0.13	0.25	0.15	0.14
intra. NCIs	291	0.95	0.32	0.15	0.08	0.30	0.11	0.09
GMTKN55	1505	0.55	0.20	0.09	0.06	0.14	0.07	0.07

C.3.1 Thermochemistry

The general main-group thermochemistry, kinetics, and non-covalent interactions (NCIs) database (GMTKN55)[99] was employed to investigate the influence of the DLPNO-DH approximation on general main-group thermochemistry. The WTMAD-2_C for different PNO thresholds on the whole GMTKN55 and on the respective subsets with reference to conventional B2PLYP are shown in Table C.2. The WTMAD-2_C for the whole GMTKN55 is also depicted in Figure C.2.

For the whole GMTKN55 database, the largest WTMAD-2_C decrease is observed from *loosePNO* (0.55 kcal·mol^{-1}) to *normalPNO* (0.20 kcal·mol^{-1}) settings and smaller further reductions are obtained with *tightPNO* (0.09 kcal·mol^{-1}), and *verytightPNO* (0.06 kcal·mol^{-1}) settings. In none of the subsets the WTMAD-2_C is above 1 kcal·mol^{-1} except for the intermolecular NCIs when *loosePNO* is employed. For the basic properties subset, *loosePNO* only yields a tiny WTMAD-2_C (0.06 kcal·mol^{-1}). Here, only minor improvements in the WTMAD-2_C can be obtained by going up to *verytightPNO* (0.02 kcal·mol^{-1}). This is because the basic property subset mostly contains small molecules. For the reactions subset, the WTMAD-2_C with *loosePNO* is larger

Table C.3: Benchmark sets included in the assessment of DLPNO-B2PLYP with the respective MAD_C in kcal·mol^{-1}. PNO settings are abbreviated (*l*, *n*, *t*, *vt*) and CPS extrapolation is indicated by an arrow. $\overline{|E_{ref.}|}$ is the original mean absolute reference energy of the respective benchmark sets.

Set	#	$\overline{\lvert E_{ref.}\rvert}$	*l*	*n*	*t*	*vt*	*l*→*n*	*n*→*t*	*t*→*vt*
IONPI19[235]	19	20.87	0.23	0.10	0.04	0.02	0.04	0.02	0.01
R160x6[247,248]	960	2.04	0.02	0.01	0.00	0.00	0.01	0.00	0.00
X40x10[244]	400	2.73	0.05	0.02	0.01	0.00	0.01	0.01	0.00
CHAL336[246]	336	14.09	0.12	0.05	0.02	0.01	0.02	0.01	0.01
ACONF-L[274]	50	4.62	0.21	0.08	0.04	0.02	0.03	0.01	0.01
HB300SPX[245]	3000	3.18	0.04	0.02	0.01	0.01	0.01	0.01	0.01
revBH9$_{BH}$[275,276]	898	20.37	0.31	0.12	0.05	0.04	0.05	0.02	0.04
revBH9$_{RE}$[275,276]	449	11.08	0.19	0.08	0.03	0.02	0.04	0.02	0.03
LP14[277]	14	23.33	1.04	0.48	0.23	0.10	0.21	0.11	0.04
L7[151,152]	7	16.27	1.34	0.60	0.24	0.08	0.23	0.08	0.02
S30L[278]	30	37.51	2.52	1.22	0.60	0.26	0.57	0.28	0.10
HS13L[a279]	13	45.82	2.29	1.14	0.57	0.28	0.57	0.29	0.13
MOR41[134]	41	31.20	0.60	0.26	0.11	0.05	0.10	0.04	0.02
ROST61[126]	61	42.78	0.40	0.18	0.09	0.05	0.08	0.05	0.03
WCCR10[249,250]	10	48.72	1.12	0.51	0.22	0.09	0.21	0.10	0.04
TMCONF16[122]	16	3.15	0.04	0.02	0.01	0.00	0.03	0.01	0.00
TMBH[251–254]	40	14.47	0.14	0.05	0.02	0.01	0.02	0.01	0.01
MOBH35[135,136,165]	70	20.89	0.25	0.09	0.04	0.02	0.04	0.02	0.01
TMIP[280]	11	95.62	0.58	0.28	0.13	0.06	0.14	0.06	0.03

Table C.4: Geometry benchmark sets included in the assessment of DLPNO-B2PLYP. Mean absolute deviations (MAD_C), and mean deviations (MD_C) in pm, °, or MHz. All deviations are given relative to the conventional MP2-based DH functional. $\overline{|x_{ref.}|}$ is the original mean absolute reference structural property of the respective benchmark sets.

			loose		*normal*		*tight*		*verytight*	
Set	#	$\overline{\lvert x_{ref.}\rvert}$	MAD_C	MD_C	MAD_C	MD_C	MAD_C	MD_C	MAD_C	MD_C
CCse21$_{bonds}$[263,264]	68	122.33	0.003	0.001	0.002	0.001	0.002	0.000	0.002	0.000
CCse21$_{angles}$[263,264] [°]	42	116.03	0.004	0.000	0.004	0.000	0.004	0.000	0.004	0.000
ROT34[a281] [MHz]	34	1411.72	0.526	−0.379	0.224	−0.129	0.051	−0.050	0.013	−0.009
HMGB11[262]	11	243.40	0.055	0.055	0.019	0.019	0.006	0.006	0.004	0.001
LMGB35[a262]	26	114.01	0.003	0.000	0.003	0.000	0.003	0.000	0.003	0.000
LB12[262]	12	299.26	0.701	0.689	0.339	0.323	0.164	0.152	0.083	0.059
TMC32[a,b261]	46	189.47	0.298	0.047	0.070	0.033	0.035	0.020	0.024	0.015

[a] Open-shell systems were excluded as no gradient is yet available for them. [b] $Fe(CO)_2(NO)_2$ was excluded due to convergence problems with B2PLYP.

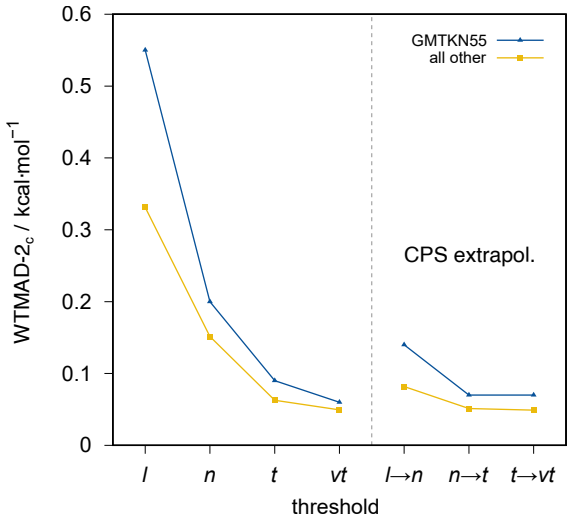

Figure C.2: WTMAD-2$_C$ with reference to conventional B2PLYP in kcal·mol^{-1} for the GMTKN55 benchmark set collection and all other thermochemistry benchmark sets assessed (cf. Table C.3)

(0.47 kcal·mol^{-1}) and still present with *normalPNO* (0.15 kcal·mol^{-1}), but becomes negligible with *tightPNO* (0.07 kcal·mol^{-1}) and *verytightPNO* (0.05 kcal·mol^{-1}) settings. Small errors are also observed for barriers with the WTMAD-2$_C$ for *loosePNO* being already tiny (0.10 kcal·mol^{-1}) with small improvements with *normalPNO* (0.04 kcal·mol^{-1}), but no further improvements with even tighter settings. Larger deviations are observed for the inter and intramolecular NCI subsets where *loosePNO* yields WTMAD-2$_C$ values around 1 kcal·mol^{-1}. These WTMAD-2$_C$ values are reduced to a third by employing *normalPNO* (0.47 kcal·mol^{-1} and 0.32 kcal·mol^{-1}) and further halved by using *tightPNO* (0.21 kcal·mol^{-1} and 0.15 kcal·mol^{-1}) and *verytightPNO* (0.13 kcal·mol^{-1} and 0.08 kcal·mol^{-1}) settings. The CPS extrapolation general reduces the WTMAD-2$_C$ for CPS(*l*→*n*) and CPS(*n*→*t*) but no improvement is observed for CPS(*t*→*vt*). Since the errors with *tightPNO* are almost converged with regard to the PNO thresholds, no further improvement is obtained by CPS(*t*→*vt*) extrapolation in this case. The improvement from *normalPNO* to CPS(*l*→*n*) is larger (from 0.20 kcal·mol^{-1} to 0.14 kcal·mol^{-1}) than from *tightPNO* to CPS(*n*→*t*) (from 0.09 kcal·mol^{-1} to 0.07 kcal·mol^{-1}). For the GMTKN55 *tightPNO* and tighter settings and CPS(*n*→*t*) and higher can be considered as converged, because WTMAD-2$_C$ values smaller than 0.1 kcal·mol^{-1} are obtained. Such errors are negligible for practical applications in comparison to the overall DH errors.

In addition to the GMTKN55, several benchmark sets were considered. The results are shown in Table C.3 and an overall weighted MAD$_C$ in Figure C.2. These include sets for NCIs of large systems (L7,[151,152] S30L,[278] and HS13L[279]), ion-π interactions (IONPI19[235]), halogen bonds (X40x10[244]), hydrogen bonds (HB300SPX[245]), chalcogen bonds (CHAL336[246]), frustrated Lewis pairs (LP14), conformational energies of alkanes (ACONF-L[274]), and repulsive NCIs (R160x6[247]). For barrier heights and reaction energies, the revBH9[275,276] set is included. Also included are sets containing transition metal complexes for closed-shell reaction energies (MOR41[134] and WCCR10[249,250]), open-shell reaction energies (ROST61[126]), conformational energies (TMCONF16[67]), barrier heights (MOBH35[135,136] and TMBH[251–254]), and ionization energies (TMIP[280]).

The largest errors are obtained for the NCI sets containing large systems. For the S30L *loosePNO* yields an MAD_C of 2.52 kcal·mol^{-1} that is larger than the MAD_Cs of the best performing DFT methods for this set (around 2 kcal·mol^{-1}). Tightening the PNO settings successively halves the MAD_C for this set from *normalPNO* (1.22 kcal·mol^{-1}) to *tightPNO* (0.60 kcal·mol^{-1}) to *verytightPNO* (0.26 kcal·mol^{-1}). Here, CPS($t\rightarrow vt$) yields a basically converged MAD_C of 0.10 kcal·mol^{-1}, but due to the many $\pi-\pi$ interactions in the S30L (as for the HS13L, L7, and LP14) the application of DH functionals to this set is questionable in the first place. Similar behavior as for the S30L is observed for the HS13L. Less prone but still severe are the MAD_Cs of the L7 and the LP14 sets where the MAD_Cs (as the average interaction energies) are basically halved compared to the S30L and HS13L. Much smaller are the MAD_Cs for the IONPI19, the revBH9, and the ACONF-L set where MAD_Cs of 0.1 kcal·mol^{-1} are already reached with *normalPNO* except for the barriers of the revBH9 by 0.02 kcal·mol^{-1}. The MAD_Cs of the CHAL336, X40x10, HB300SPX, and the R160x6 sets are already small with *loosePNO* and become vanishing small with tighter settings. This may again be attributed to the relatively small system size of the molecules in these sets. For the transition metal containing sets the WCCR10 shows the largest MAD_Cs (1.12 kcal·mol^{-1} with *loosePNO*) followed by the MOR41, the TMIP, and the ROST61 (between 0.4 – 0.6 kcal·mol^{-1} with *loosePNO*). Smaller errors are observed for the MOBH35 and the TMBH, where all settings tighter than *loosePNO* yield MAD_Cs smaller than 0.1 kcal·mol^{-1}. Surprising are the vanishing MAD_Cs for the TMCONF16 set. In conclusion, the errors for the organometallic sets are larger than for typical organic reactions but with *tightPNO* settings or CPS($l\rightarrow n$) the MAD_Cs are around 0.1 kcal·mol^{-1} (with one exception). This error is negligible compared to the errors of the corresponding DH.

C.3.2 MP2 and HFX Contribution

As DH functionals typically include different amounts of MP2 correlation in their energy expression (cf. eq C.1) the estimated error introduced by the DLPNO-DH approximation can vary as well. Nevertheless, the introduced error behaves linearly with the amount of MP2 correlation which is demonstrated for DLPNO-B2PLYP variants with varying amounts of MP2 of 20, 40, 60, and 80% on the L7 set (Figure C.3**a**). Accordingly, tripling the amount of MP2 correlation triples the DLPNO-DH error with respect to conventional B2PLYP. Nevertheless, most robust and well-behaved DHs employ values of around 30% MP2 correlation, allowing for a reasonable error estimate based on the results for B2PLYP (27% MP2). In addition to different amounts of MP2, different amounts of HFX are employed as well in DH functionals, which can influence the DLPNO error. This is demonstrated for DLPNO-B2PLYP variants with varying amounts of HFX of 20, 40, 60, and 80% on the L7 set (Figure C.3**b**). With increasing amounts of HFX, the self-interaction error in the DH functionals is reduced, which results in reduced electron delocalization and a slightly reduced DLPNO error. The most robust and well-behaved DHs employ between 50 – 80% HFX, and can therefore benefit from a slightly reduced DLPNO error compared to DHs with less HFX.

C.3.3 Size dependence of correlation energy error

In line with previous findings on PNO errors in DLPNO-CCSD(T) approaches, the correlation energy error with respect to conventional B2PLYP behaves almost linearly with the size of the system. This is demonstrated for a polyalanine chain (Figure C.4) where a clear decrease in the size dependence upon tightening the PNO thresholds is observed. Further, even a CPS($l\rightarrow n$) PNO space extrapolation can

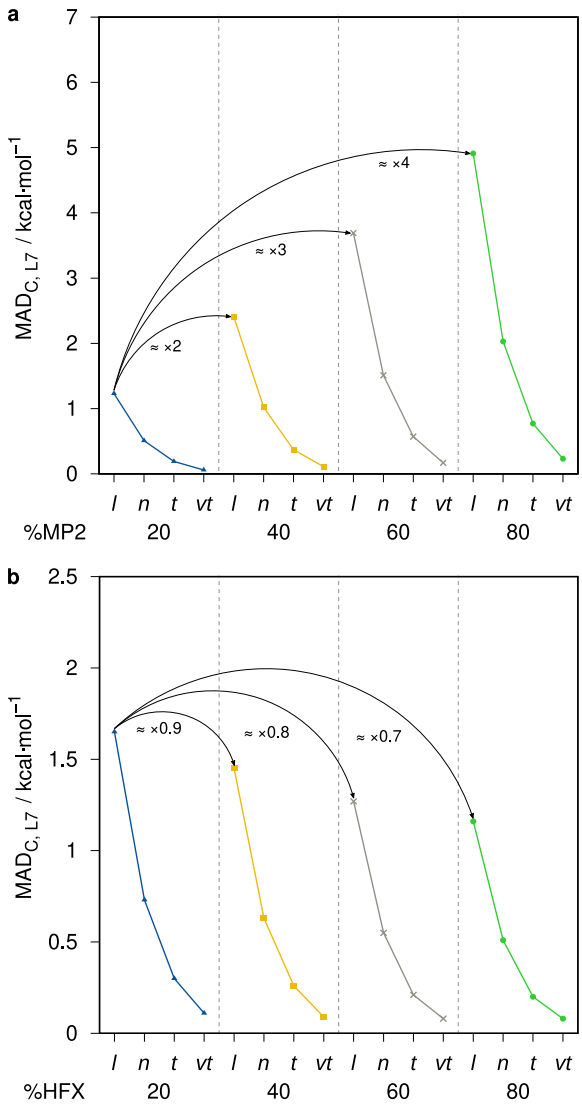

Figure C.3: (a) MAD$_C$s for the L7 benchmark set for B2PLYP variants with varying amounts of MP2 correlation. (b) MAD$_C$s for B2PLYP variant with varying amounts of HFX. *l* = *loosePNO*, *n* = *normalPNO*, *t* = *tightPNO*, *vt* = *verytightPNO*.

eliminate most of the size-dependent correlation energy error for this case.

C.3.4 Geometry Optimizations

As even energy calculations on a high theoretical level such as DH-DFT are computationally demanding, geometry optimizations requiring many energy and gradient evaluations are typically unfeasible. Nevertheless, highly accurate geometry optimizations are desirable for critical cases and specifically benchmarking more approximate methods such as semi-empirical quantum mechanics (SQM) or force-fields (FF). By employing DLPNO-DH, respective DH functionals become feasible again

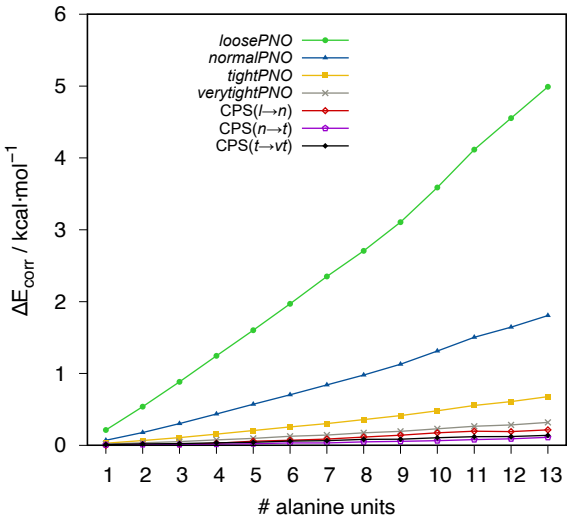

Figure C.4: Error in MP2 correlation energy with reference to conventional B2PLYP in kcal·mol^{-1} for polyalanines.

for geometry optimizations of medium-sized to large molecules. To estimate the influence of the DLPNO-DH threshold settings, DLPNO-B2PLYP geometry optimizations were performed for various established geometry optimization benchmark sets. The resulting geometries were compared to the conventional MP2-based B2PLYP results (Figure C.5 and Table C.4). The following geometric features were investigated: Rotational constants for small to medium-sized organic molecules were compared (ROT34[281]). Bond lengths were compared for 3d transition metal complexes (TMC32[261]) and light (LMGB35[262]) and heavy (HMGB11[262]) main-group compounds as well as a mixed set containing unusually long bonds (LB12[262]). Additionally, for the CCe21 set,[263,264] containing semi-experimental structures of organic molecules, bond distances and angles were compared. No effect of the accuracy settings is observed for the LMGB35 and the conventional B2PLYP bond lengths are almost obtained with a vanishing MAD_C of 0.003 pm because the molecules in this test set are very small. Similar errors are observed for the CCse21 and the HMGB11 set. Although here small differences between the PNO settings are observed. Larger errors that are still below 1 pm are found for the TMC32 and the LB12 set with *loosePNO*. For the TMC32 *normalPNO* is already sufficient, while for the LB12 set errors below 0.1 pm are only obtained with *verytightPNO*. For the ROT34 small MAD_Cs are observed with *loosePNO* (0.526 MHz) and *normalPNO* (0.224 MHz) and basically vanish with *tightPNO* (0.051 MHz). In general, the introduced errors of the DLPNO approximation are very small compared to the B2PLYP result. In all cases MAD_Cs below 1 pm, 1 MHz, or 1 [°] were obtained. The very small differences between the structures obtained using varying PNO thresholds and conventional B2PLYP can also be seen for the large frustrated Lewis-pair (FLP) system of the LB12 benchmark set. An overlay of all optimized structures shows no significant difference in the optimized structures (Figure C.6), underlining the value of using less tight PNO thresholds for geometry optimizations. Overall, the errors introduced by the DLPNO-DH approximation are generally much less pronounced for geometrical features. This renders the efficient *normalPNO* settings already suitable for DLPNO-DH geometry optimizations of large systems.

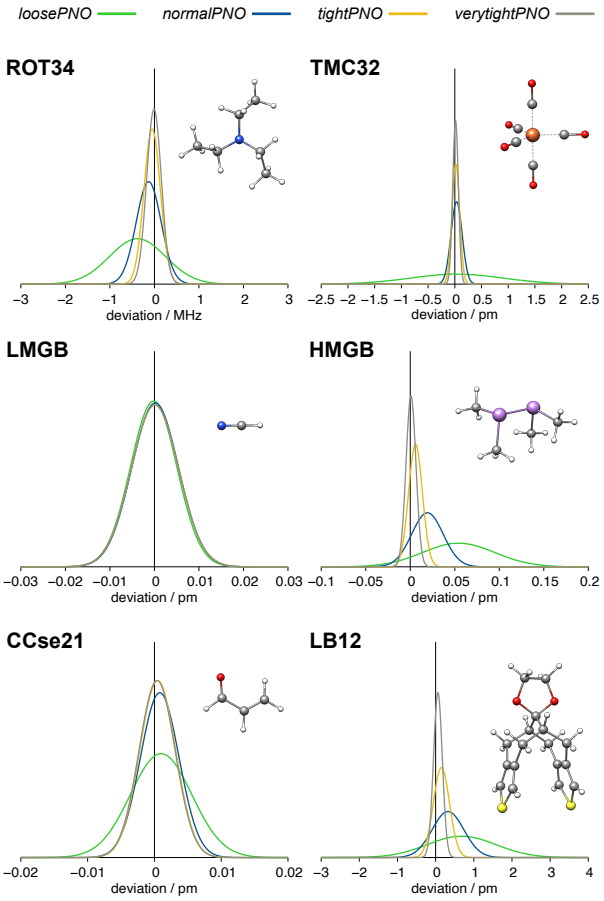

Figure C.5: Gaussian error distributions for selected bond length benchmark sets with reference to conventional B2PLYP results. Negative mean deviations indicate overall too short bond lengths compared to the canocial result.

C.3.5 Timing Comparisons

The computational demand of energy and gradient evaluations typically determines the feasibility of a geometry optimization. Therefore, the computational wall-time reduction of a subsequent energy and gradient calculation is assessed for various PNO thresholds for polyalanine with varying chain length (Figure C.7). At a crossing point of about five alanine units (53 atoms), the DLPNO-DH approximation begins to drastically reduce the computation time of the combined energy and gradient compared to the conventional DH. The steep scaling of the latter causes a drastic increase in computation time while the DLPNO-DH approach yields a flat, almost linear, scaling with the size of the system. A comparison of the energy and gradient computation time contributions for the parental hybrid functional and the DH variants is depicted in Figure C.8. The scaling of the conventional and the DLPNO-DH with respect to the system size (number of basis functions) for these selected molecules is shown in Figure S1 in the SI. In line with the results shown for the polyalanine chain, the gradient evaluation profits significantly from the DLPNO-DH approximation, even for medium sized molecules such as the depicted molybdosilylidine complex with 56 atoms. Nevertheless, for these molecule sizes

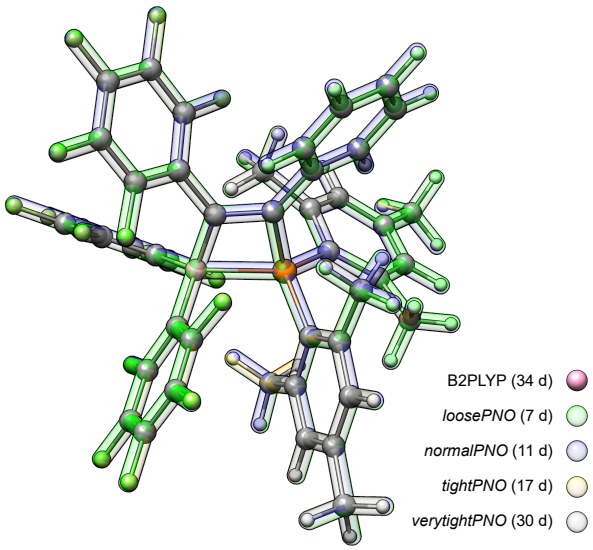

Figure C.6: Structure overlay of the FLP (88 atoms) from the LB12 benchmark set optimized at various PNO threshold settings. All optimizations were performed on 4 CPUs using an Intel® Xeon® CPU E3-1270 v5 @ 3.60GHz machine.

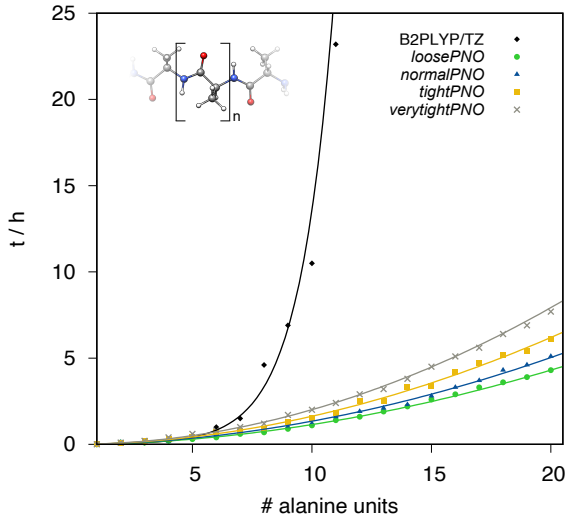

Figure C.7: Computation wall-times in h for energy and gradient evaluation of polyalanine chains with up to 20 alanine units (203 atoms) for conventional B2PLYP/def2-TZVP and DLPNO-B2PLYP/def2-TZVP with different PNO thresholds. All calculations were performed on 14 CPUs using an Intel® Xeon® CPU E5-2660 v4 @ 2.00GHz machine.

the overhead of the DLPNO space construction causes more costly energy evaluations compared to the conventional DH, thus resulting in a higher overall computation time. With increasing size, the energy computation using the DLPNO approximation becomes increasingly faster, and the benefit for the gradient evaluation is even more drastic underlining the value of DLPNO-DH calculations for

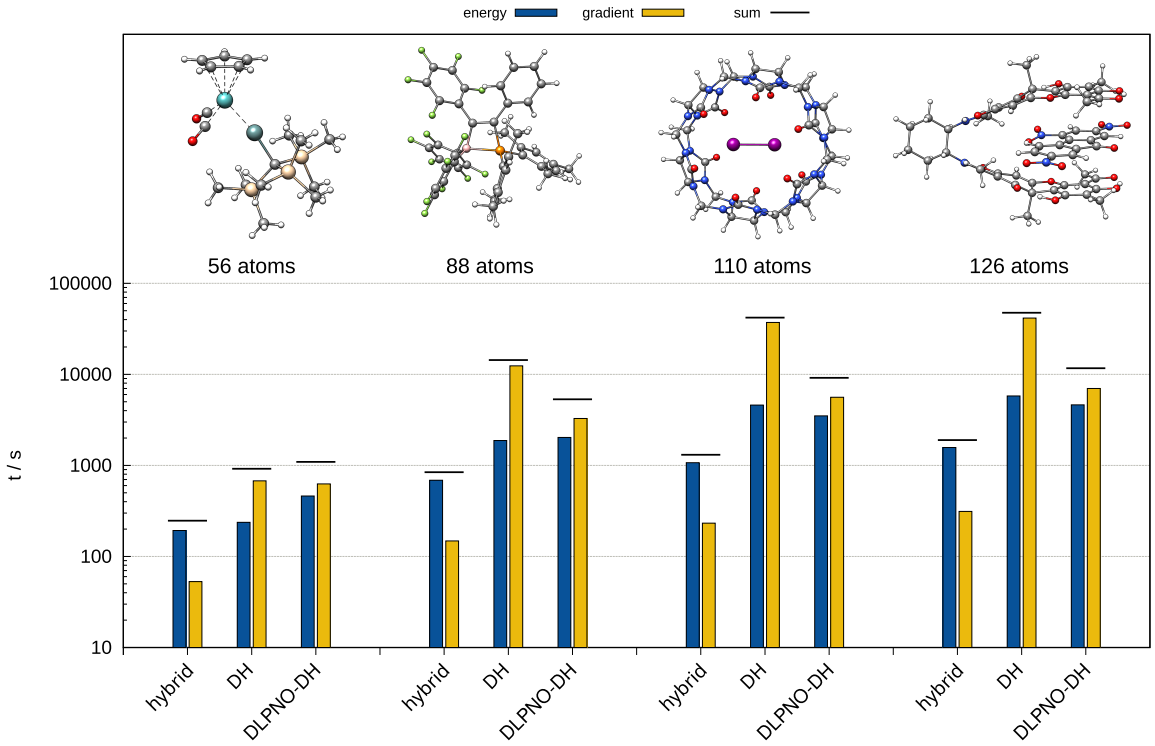

Figure C.8: Computation times in s for energy and gradient evaluation of selected molecules in the range of 56 to 126 atoms. Hybrid = BLYP with 53% HFX; DH = B2PLYP; DLPNO-DH = DLPNO-B2PLYP with *normalPNO* thresholds. The def2-TZVP(-f) basis was used throughout. All calculations were performed on 14 CPUs using an Intel® Xeon® CPU E5-2660 v4 @ 2.00GHz machine. Note the logarithmic scale.

molecules with more than 100 atoms.

C.3.6 General Recommendations

Finally, as a good compromise between computational cost and accuracy, we recommend employing *normalPNO* for the gradient calculation of medium-sized organic compounds (up to 100 atoms) and conventional DHs for the energy evaluation as the latter does not profit from the DLPNO approximation speedup. In the regime of 100 atoms and larger, DLPNO-DHs yield increasing speedup and may be employed with *normalPNO* for most systems. In this context, the DLPNO approximation starts to generally enable DH calculations that would be unfeasible for such large systems due to computation time and memory issues. In terms of PNO-space extrapolation, we recommend CPS($l \rightarrow n$) for reactions and CPS($n \rightarrow t$) for non-covalent interactions.

C.4 Conclusion

In this work, the application of the DLPNO-MP2 approximation in the DH-DFT framework was assessed. The performance of different PNO thresholds as well as PNO space extrapolations was tested for the prominent B2PLYP functional on various benchmark sets for the thermochemistry of

main-group molecules and transition metal complexes. It was demonstrated that *tightPNO* settings yield reliably small deviations from conventional B2PLYP at a drastically reduced computational cost for large systems (WTMAD-$2_C{}^{\text{all}}$=0.06 kcal·mol^{-1}, WTMAD-$2_C{}^{\text{GMTKN55}}$=0.09 kcal·mol^{-1}).

In general we expect that the observed DLPNO error is transferable to other DH functionals as the error behaves linearly with the amount of the MP2 correlation and HFX admixture. The errors for geometry optimizations were found to be even smaller and in many cases negligibly small even at moderately tight PNO thresholds. *normalPNO* yields already satisfactory agreement with geometries optimized with conventional B2PLYP. The CPS extrapolation scheme introduced in the DLPNO-CCSD(T) framework was successfully applied to DLPNO-DH calculations, with CPS($n\rightarrow t$) typically yielding accurate results with very small residual errors compared to the conventional DHs (WTMAD-$2_C{}^{\text{all}}$=0.05 kcal·mol^{-1}, WTMAD-$2_C{}^{\text{GMTKN55}}$=0.07 kcal·mol^{-1}). The CPS parameter $F = 1.5$ was found to be also suitable in the DLPNO-DH-DFT framework.

The performance of DLPNO-DH-DFT can potentially also benefit from employing the so-called tightened semicore settings as proposed by Altun*et al.*[176] for transition metal complexes or modified PNO settings as proposed by Werner and Hansen.[282]

Overall, it is demonstrated that DLPNO-DH-DFT represents a valuable alternative to conventional DH-DFT for large systems where the unfavorable N^5 scaling of MP2 prevents its application. DLPNO-DH-DFT may be applied to enable highly accurate geometry optimizations and energy calculations of large molecules that are unfeasible with conventional DH functionals. Further, the technical implementation of the underlying DLPNO-MP2 is much easier compared to its DLPNO-CCSD(T) counterpart, increasing its potential availability in common quantum chemistry programs.

C.5 Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00896.

- Detailed DLPNO thresholds, statistical measures, and additional statistics for all benchmark sets (PDF)

C.6 Acknowledgements

The German Science Foundation (DFG) is gratefully acknowledged for financial support (Grant 1927/16-1). Further, SG and MB gratefully acknowledge financial support of the Max Planck Society through the Max Planck fellow program.

APPENDIX D

Toward Benchmark-quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts - A Comparison of CCSD(T) and ph-AFQMC

Hagen Neugebauer,[†] Hung T. Vuong,[‡] John L. Weber,[‡] Richard A. Friesner,[‡] James Shee,[§] Andreas Hansen[†]

Received: 9 June 2023
Published online: 1 September 2023

Reprinted in Appendix D (adapted) with permission[¶] from
H. Neugebauer, H. T. Vuong, J. L. Weber, R. A. Friesner, J. Shee, and A. Hansen, *Toward Benchmark-Quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts: A Comparison of CCSD(T) and ph-AFQMC*, J. Chem. Theory Comput. **19** (2023) 6208, DOI: 10.1021/acs.jctc.3c00617

Own contributions

- compiling the 3dTMV benchmark set
- performing all DFT calculations and parts of the CC calculations
- statistical evaluation
- interpretation of the results
- co-writing of the manuscript

[†]Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, University of Bonn, Beringstr. 4, D-53115 Bonn, Germany
[‡]Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027
[§]Department of Chemistry, University of California, Berkeley, California 94720
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

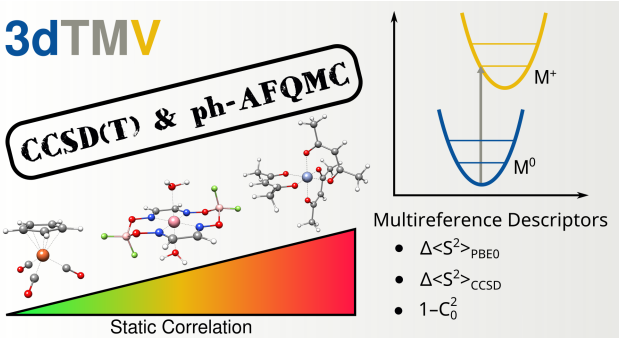

Figure D.1: Table of content graphic (ToC).

Abstract Generating accurate *ab initio* ionization energies for transition metal complexes is an important step toward the accurate computational description of their electrocatalytic reactions. Benchmark-quality data is required for testing existing theoretical methods and developing new ones but is complicated to obtain for many transition metal compounds due to the potential presence of both strong dynamical and static electron correlation. In this regime, it is questionable whether the so-called gold standard, coupled cluster with singles, doubles, and perturbative triples (CCSD(T)), provides the desired level of accuracy – roughly 1 – 3 kcal·mol^{-1}. In this work, we compiled a test set of 28 3d metal-containing molecules relevant to homogeneous electrocatalysis (termed 3dTMV) and computed their vertical ionization energies (ionization potentials) with CCSD(T) and phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC) in the def2-SVP basis set. A substantial effort has been made to converge away the phaseless bias in the ph-AFQMC reference values. We assess a wide variety of multireference diagnostics, and find that spin-symmetry breaking of the CCSD wave function and in the PBE0 density functional correlate well with our analysis of multiconfigurational wave functions. We propose quantitative criteria based on symmetry breaking to delineate correlation regimes inside of which appropriately-performed CCSD(T) can produce mean absolute deviations from the ph-AFQMC reference values of roughly 2 kcal·mol^{-1} or less, and outside of which CCSD(T) is expected to fail. We also present a preliminary assessment of density functional theory (DFT) functionals on the 3dTMV set.

D.1 Introduction

Molecular electrocatalysis based on 3d transition metal compounds is an important tool for the synthesis of complex molecules[283,284] and a promising approach for CO_2 reduction,[19–22] water splitting,[25,26] oxygen reduction,[24] and other hydrogen evolution reactions[27] employing earth abundant metals.[16] Quantum chemistry (QC) can potentially guide the development of novel catalysts. Elucidating the underlying mechanisms and tuning ligand structures to achieve lower overpotentials, higher turnover frequencies, and better substrate selectivity are typical applications.[153]

However, the choice of a reliable QC method – e.g., density functional theory (DFT) or wave function theory (WFT) – for first-row (3d) transition metal electrocatalysts is more challenging than in the case of organic molecules, as the target systems can be large and often exhibit relatively complicated electronic structures. Transition metal complexes with low coordination number, high symmetry, and/or multiple magnetically-coupled radical sites can have more degenerate or nearly-degenerate orbitals

and spin-states, which can result in multireference (MR) character (also known as static or strong correlation).[86–88] The degree of static correlation increases in many open-shell species or complexes with redox noninnocent ligands that often appear in electrocatalysis.[86,204] Various MR descriptors are available in the literature for detecting such cases with significant static correlation.[285,286] Dynamic correlation effects between multiple electron pairs (for example in carbonyl ligands with σ-donation and π-backbonding) can occur as well, which requires a theoretical description beyond second-order perturbation theory.[84,154]

The first step towards robust modeling of experimentally relevant electrocatalysts is the accurate prediction of redox potentials, which besides accurate solvation free energies, requires reliable ionization energies.[90] To test QC methods, multiple benchmarks of transition metal complexes with experimental redox potentials were compiled.[11,90,287–290] One shortcoming when employing solely experimental redox potentials as reference is that the electronic structure problem cannot be investigated independently from solvation effects and error compensation hinders the analysis of individual error contributions. Reference ionization energies allow a separation of the error sources from the electronic structure and solvation contributions and enable precise error tracing. For example, this was done in a study by Isegawa *et al.*, who compared experimental ionization and redox potentials of organic compounds to DFT and CCSD(T) calculated potentials.[291] Here, the favorable performance of CCSD(T) for ionization energies diminished for redox potentials due to missing solvation contributions. Sterling *et al.* applied an explicit solvation workflow to include the neglected solvation contributions and obtained excellent agreement with experiment.[167] Another interesting approach that tackles both problems at once is the explicit inclusion of solvent molecules in combination with an implicit solvation model. In a multilayer DLPNO-CCSD(T) study performed by Bhattacharjee *et al.* for first-row transition metals in water,[292] accurate reduction potentials were obtained.

However, only rarely are experimental reference ionization energies available for 3d electrocatalysts; computational protocols that can generate these accurate ionization energies are highly desired. For main group chemistry, benchmark-quality reference ionization potentials and electron affinities can be generated with CCSD(T)[56] as done for medium-sized acceptor molecules by Richard *et al.*[293] Because of the varying (and sometimes large) degree of static correlation encountered in 3d transition metal electrocatalysis, the expected accuracy of CCSD(T) is unclear. For transition metal atoms CCSD(T) can yield accurate results for ionization energies,[52] but for many diatomic molecular bond dissociation energies the predictive power of CCSD(T) deteriorates.[294–299] However, for experimentally derived spin splittings of nonheme iron complexes as well as metallocenes, CCSD(T), especially when based on Kohn−Sham (KS) orbitals, yielded accurate results.[198,201] Good performance for CCSD(T) with KS orbitals could also be observed for calculating vibrational frequencies.[300] Another issue is that molecules of medium to large sizes cannot be treated by canonical CCSD(T). This is especially the case at the complete basis set (CBS) limit, which is required for benchmark-quality data to be comparable to those of experimental measurements. Therefore, less expensive but potentially less accurate localized coupled cluster schemes are often employed. One of the most prominent of these schemes is DLPNO-CCSD(T).[59–61] For reactions involving single reference (SR), large-gap states, DLPNO-CCSD(T) can be employed as a reliable reference method for benchmarking closed-shell (MOR41)[134] and open-shell (ROST61)[126] organometallic reactions. The efficiency of the DLPNO-CCSD(T) model has been demonstrated via applications to very large systems, including metalloenzymes.[301] Iron *et al.* used the method to compute transition metal barrier heights in the MOBH35 study.[135,136] Later, a revision to their DLPNO-CCSD(T) references was suggested on the grounds of static correlation

effects,[302] but a study by Altun *et al.* traced the DLPNO error to correlation effects from the 3s and 3p semicore orbitals and to dynamical correlation-induced orbital relaxation effects.[176] In another study the DLPNO approach performed well for challenging spin-splitting energies of iron complexes and reproduced CASPT2/CC results with a two-point PNO extrapolation and improved full iterative triples.[203] Results of similar quality were found for the ionization energy of cobaltocene.[303]

Nevertheless, it would be highly desirable to use accurate and scalable methods to cross-check CCSD(T) and its localized schemes. A promising candidate for this task is the phaseless auxiliary field quantum Monte Carlo (ph-AFQMC) method,[51–55] which was originally developed in the physics community.[304,305] Recently, a localized orbital version (LO-ph-AFQMC) has been developed.[177] Especially relevant for transition metal chemistry is the nonperturbative and inherently MR nature of this method. In principle, the phaseless bias can be converged away by systematically improving the trial wave function toward the exact wave function. Once a trial wave function is obtained, the cubic scaling computational cost with system size and near-perfect parallel efficiency of LO-ph-AFQMC enables applications to relatively large molecular systems, at least when compact trail wave functions are required. However, due to the stochastic nature of the method, one must typically sample thousands of trajectories to obtain adequate statistics. Encouragingly, the scaling prefactor has been reduced significantly via efficient implementations on graphical processing units (GPUs),[52,306] and the branching random walks are amenable to massively parallel computing systems that can efficiently utilize hundreds of GPUs at a time. ph-AFQMC has been successfully used to calculate accurate ionization potentials for transition metal atoms,[52] dissociation energies of transition metal containing diatomics,[298] ligand dissociation energies[307] and ionization potentials[280] of transition metal complexes. Alternative accurate MR methods include the recently developed CASPT2+δMRCI approach for spin-splittings,[202,308] but this approach is not feasible for the system sizes at hand.

In this work, we compile a set of medium-sized 3d transition metal complexes relevant to electrocatalysis. We compare CCSD(T) and ph-AFQMC predictions in the def2-SVP basis set, and investigate the degree of MR character present using a variety of diagnostics. We propose a quantitative classification protocol, involving the dominant coefficient in multiconfigurational wave functions and the degree of spin-symmetry breaking or restoration from coupled cluster wave functions and the PBE0 density functional, to gauge the expected accuracy of various coupled cluster approaches. For vertical ionization involving predominantly SR systems, average errors of less than 2.3 kcal·mol^{-1} can be achieved with a specific choice of orbitals and spin restriction. In contrast, large deviations between all CCSD(T) protocols and reference ph-AFQMC values were observed for ionizations involving states exhibiting strong static correlation. In contrast to other studies in which many coupled cluster approaches were tested for a specific, smaller system,[309] we investigate only select CC protocols for which i) the required computational demands exceed those of CCSD(T) by at most a factor of 2, and ii) local implementations are available for subsequent studies with extended basis sets.

D.1.1 Design of the 3dTMV Benchmark set

The 3dTMV benchmark was generated with the following design criteria in mind. We chose complexes that (i) have experimental relevance to electrocatalysis (e.g., hydrogen evolution, CO_2 reduction), (ii) are of medium size ($20-40$ atoms), and (iii) contain one or two transition metal atoms. The molecules included in the 3dTMV set are depicted in Figure D.2. Their charges, multiplicities, and experimental references are given in Table D.1. The vertical ionization energies, which we will refer to in the following as ionization potentials (IPs), were calculated. The calculation of vertical ionization energies

is more challenging than that for adiabatic ionization potentials because the oxidized species are not in their optimal ground state geometry, which can lead to more challenging electronic structures. The classification of the IPs into SR, SR/MR, and MR subsets is discussed in detail in Section D.3.2.

Complexes **1** and **2** are two of the few studied complexes for fuel generation with titanium.[310,311] Complexes **3-5**, **13**, and **14** are all metallocenes of which ferrocene is of special interest to electrochemistry due to its ideal reversible oxidation.[312–316] The IPs have already been investigated in ref. [280], and there is renewed interest in the isolation and electronic structures of the manganocene, ferrocene and cobaltocene anions.[317,318] Systems **6**, **7**, and **19** are examples of iron-based three-legged piano-stool catalysts[5,319–321] and **8-10** are models of three-legged piano-stool catalysts with cobalt.[322] System **11** is a cyclam model[323], while **12** is a model for the DuBois type catalyst.[324] The carbonyl bipyridine complexes **15-18** are model catalysts for more elaborate pyridine based catalysts.[325–328] The cobalt complexes **20** and **23** are examples of cobaloximes that have been employed for H_2 generation.[329–332] System **21** is a nitrogen reduction catalyst that can reduce dinitrogen into ammonia and hydrazine.[333] Complex **22** is a cutout from a hydrogenase model.[334] The chromium complex **24** and the iron complex **25** are acetylacetone (acac) complexes that are widely studied as models for other tris-β diketonate complexes and have also been investigated in ref. [280]. $Fe(acac)_3$ is also of interest for photoredox catalysis.[335] Complex **26** is a hydrogenase mimic[336] and the only system with two metal atoms in the set. The planar complexes **27** and **28**, relevant for hydrogen generation[337,338], are especially challenging due to their redox-active ligands.[339]

Table D.1: Molecules included in the 3dTMV set with corresponding charges and spin multiplicities for the oxidized (ox) and initial (in) state as well as experimental studies where these complexes were investigated.

#	charge		multiplicity		ref.	#	charge		multiplicity		ref.
	ox	in	ox	in			ox	in	ox	in	
1	1	0	2	1	[[310]]	15	1	0	2	1	[[325]]
2	1	0	1	2	[[311]]	16	1	0	2	1	[[326]]
3	1	0	4	3	[[340]]	17	1	0	2	1	[[327]]
4	1	0	2	1	[[313]]	18	1	0	2	1	[[328]]
5	1	0	2	1	[[316]]	19	2	1	2	1	[[320]]
6	2	1	2	1	[[319]]	20	1	0	3	2	[[329]]
7	1	0	2	1	[[321]]	21	1	0	3	2	[[333]]
8	2	1	2	1	[[322]]	22	1	0	2	1	[[334]]
9	2	1	2	1	[[322]]	23	1	0	2	1	[[330]]
10	2	1	2	1	[[322]]	24	1	0	3	4	[[341]]
11	2	1	1	2	[[323]]	25	1	0	3	6	[[341]]
12	2	1	1	2	[[324]]	26	1	0	2	1	[[336]]
13	1	0	1	2	[[340]]	27	0	-1	2	3	[[337]]
14	1	0	2	3	[[340]]	28	0	-1	1	2	[[338]]

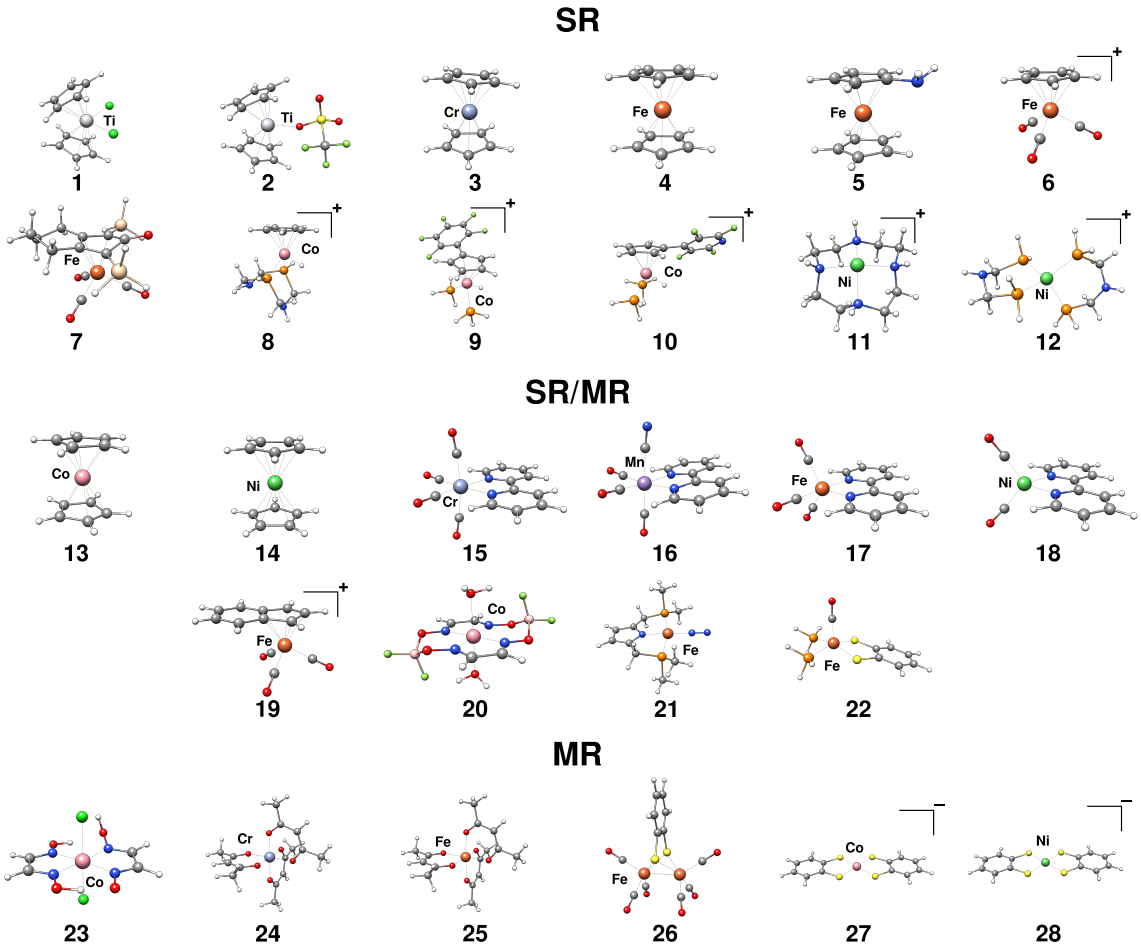

Figure D.2: Structures included in the 3dTMV benchmark set divided into single reference (**SR**), an intermediate category (**SR/MR**), and multireference (**MR**) subsets.

D.2 Computational Methods

D.2.1 DFT and Coupled Cluster Calculations

Geometries were optimized with the r^{2}SCAN-3c level of theory[137] in TURBOMOLE 7.5.1.[342,343] As shown in Section D.3.2, the initial charge states (even for IPs in the MR subset) are of single-reference character, so DFT methods should be suitable for the geometry optimizations relevant to calculating vertical IPs. To investigate the influence of different methods for the geometries, the geometries of **12** and **13** were reoptimized on the CCSD/def2-SVP level of theory and UHF CCSD(T) IPs were calculated for comparison (see Table S18). The differences are moderate and well within the uncertainty of the reference IPs (*vide infra*). Hence, r^{2}SCAN-3c is a reasonable choice for the geometry optimizations given that CCSD optimizations for the whole 3dTMV set would be computationally very demanding.

All single-point DFT calculations, if not stated otherwise, were performed with the double-ζ

def2-SVP basis set[127] in ORCA 5.0.3[141] employing the *DEFGRID3* and the *TightSCF* settings. Namely, the GGA functional PBE[101,124], the meta-GGA functionals r^{2}SCAN[106,107] and M06-L[344], the hybrid functionals PBE0[156], r^{2}SCAN0[67], B3LYP[345,346], PW6B95[347], and M06-2X[348] as well as the range separated hybrid ωB97X-V[157], and the double hybrid PWPB95[349] were used. Robust SCF convergence for DFT calculations was ensured by employing the TRAH SCF solver.[220] The split-RI-J approximation was used to speed up DFT calculations[221] using the corresponding auxiliary basis set[222] and hybrid DFT calculations were additionally sped up with the RIJCOSX approximation.[223–225] To account for London Dispersion effects the D4 dispersion correction[9] was employed in the dftd4 v.3.4.0 standalone program.[219] The local-hybrid LH20t DFT functional[350] was employed in TURBOMOLE 7.5.1. and the hybrid KP16 functional[351] was employed in the xTron program package.[352] CCSD(T) calculations were performed in the Q-Chem 5.4 program package[353] with the def2-SVP basis set and with the cutoff for neglecting two electron integrals set to 10^{-14} and the SCF convergence set to 10^{-8}. To ensure robust SCF convergence, the GDM algorithm was employed as SCF solver[354] and the internal stability analysis was used to verify SCF solutions as minima in orbital space. The LIBPT library was used for the evaluation of the triples contribution. DFT calculations for follow up CCSD(T) calculations in Q-Chem were performed with the PBE0 functional[156] and the SG-1 grid.[355] To account for correlation effects in the 3s and 3p shells of 3d transition metals, the frozen core settings in the correlated calculations were adjusted to freeze only the 1s, 2s, and 2p shells of 3d transition metal atoms. The importance of these core-valence correlation contributions has been emphasized recently for 3d transition metals in the context of DLPNO-CCSD(T)[176] and CASSCF[356] calculations.

Note that for accurate quantitative IPs that are comparable to experimental measurements (apart from thermostatistical corrections) larger basis sets are necessary. However, the comparisons made between various high-level correlated WFT methods (CC vs. ph-AFQMC) evaluated with the same basis set is justified given the comparable degree of basis set incompleteness error (BSIE) for both methods. Also, the MR diagnostics is still reliable with the applied double-ζ basis set, which has already been shown in several other works.[86,302,357,358] For completeness, we confirmed this exemplary by comparing the def2-SVP and the def2-QZVPP $\langle S^2 \rangle$ values obtained with PBE0 employing the ORCA program package. We observed only small differences between the small and large basis set values and the conclusions are virtually the same (see SI Table S7).

In the following UHF/RHF means that singlets were calculated with restricted Hartree-Fock (RHF) and all other spin multiplicities with unrestricted Hartree-Fock (UHF). Similarly, orbital protocols such as RPBE0/UPBE0 indicate that R and U are used for singlets and non-singlets, respectively. The ROHF/RHF CR-CC(2,3)[359,360] calculations were performed with a serial implementation from GAMESS 2021 R2 Patch 2.[361,362]

D.2.2 ph-AFQMC Calculations

Trial wave functions and the phaseless bias in AFQMC

In principle, it should be possible to systematically converge AFQMC calculations, even for strongly MR transition metal containing species, by systematically increasing the quality of the trial wave function until a stable value of the observable (in this case the ionization potential of the complex) is obtained. In practice, the question of how to best optimize a multideterminental wave function for AFQMC for transition metal containing systems is a challenging basic research problem at present. In this Letter, we utilize two different approaches based on the degree of MR character, which will be

outlined below. While we do not claim to have a universally valid, rigorously converged protocol (which would require studying substantially larger and more diverse data sets, as well as a considerably larger investment of computational resources), we do believe that the results reported here represent progress in controlling the error in the calculations for these challenging cases as compared to our prior efforts. The significant fluctuations in the CCSD(T) results, seen even in the test cases judged to be squarely in the SR regime, are indicative of the difficulty of obtaining ionization potentials that are robustly of chemical accuracy.

Our initial work using AFQMC methods to study transition metal containing species investigated three types of trial functions: Hartree Fock (both restricted and unrestricted), DFT (primarily unrestricted), and CASSCF. For difficult cases, CASSCF trials were generally required, although computational limitations significantly restricted the size of the active space that could be employed. The lack of scalability of CASSCF (with regard to both active space dimension and overall molecular size) motivated us to explore the use of selected CI approaches, which enable the investigation of substantially larger active spaces than what is accessible through CASSCF. The trial wave functions used for the SR and SR/MR subsets were from a variational selected CI procedure, which we refer to as HCISCF.[363,364] Configurations were selected using $\epsilon_1 = 10^{-4}$ a.u. Active spaces were chosen as follows: First, the eigenvalues (orbital energies) of the Fock or Kohn−Sham operator were obtained, and the sequential energy differences were plotted. We made cuts between orbitals that have relatively large energetic separation, such that 24 – 43 active orbitals are kept in a subsequent optimization of both selected CI coefficients and orbital coefficients. Between 88 and 92 % of the HCISCF CI weight was retained in the ph-AFQMC trial wave functions, which corresponds to between 5 and 1199 determinants (a relatively small number compared to the full Hilbert space). In some cases (see the SI), B3LYP orbitals were used as an initial guess for the HCISCF optimization. This was done when a lower-energy solution at the HCISCF level could be obtained, when the initial HCISCF calculations using HF orbitals as a starting point did not converge, or when the HF orbitals were found to be qualitatively wrong (*vide infra*).

In order to get a sense of the sensitivity of our results for the SR and SR/MR subsets to the trial wave function employed, we also experimented with two alternative protocols which will be presented in a future work. First, a CASSCF optimization in a small active space is followed by a single-shot selected CI calculation in a larger active space. Second, we use an initial selected CI calculation in a very large active space (with a relatively loose selection threshold) and use natural orbital occupation number cutoffs to define a smaller active space which can be treated with a tighter cutoff. We find mean absolute deviations from the present predictions with HCISCF trials of roughly 1.4 and 2.1 kcal·mol^{-1}, which are a little larger than what would traditionally be considered "chemical accuracy" (1 kcal·mol^{-1}) but well within the looser criterion of "transition metal chemical accuracy" (1-3 kcal·mol^{-1}),[294] which we believe is an appropriate target for the current state of the art in electronic structure technology for these systems, and which is capable of providing a very reasonable picture of reaction thermochemistry for complex problems in materials and biology. To conservatively account for any residual error due to the possibility of unconverged phaseless bias in the SR and SR/MR subsets, we report (in the Figures to follow) a total uncertainty on the predicted IPs that represents the statistical error of each energy difference plus 1.5 kcal·mol^{-1} for the SR and SR/MR subsets.

For the MR set, we attempted to converge our ph-AFQMC predictions with respect to three different dimensions: the size of the active space, the value of ϵ_1 in selected CI, and the % CI weight retained in the trial wave function (which we abbreviate hereafter as % CI). Regarding active spaces, three different active space sizes were chosen: a small space of $\leq$ 18 orbitals, a medium-sized space with

typically 40 to 50 orbitals, and a large-sized active space with some 60 − 80 orbitals. These active spaces were selected by making cuts based on the orbital energies obtained from a PBE0 calculation, similar to the SR and SR/MR subsets. We generated a CASSCF trial for the smallest active space, using PBE0 orbitals as a starting point, and then assessed qualitatively the MR character using the natural orbital occupation numbers (NOONs) and the $1 - C_0^2$ values. The medium and large active space trials were then generated from HCI (no SCF) calculations using the orbitals optimized from CASSCF. Three different choices of ϵ_1 were investigated: 10^{-3}, 10^{-4}, and 5×10^{-5} a.u. We note that in every case, except for complex **23**, the ph-AFQMC IPs with ϵ_1 values of 10^{-4} and 5×10^{-5}, which represent the addition of 294-581 determinants, are converged to within statistical error bars (the difference in the IP is 4.6 ± 1.6 kcal·mol^{-1} for **23**, which will be discussed further below). Finally, ph-AFQMC calculations were performed with at least three different %CI values in the largest active space with $\epsilon_1 = 5 \times 10^{-5}$ trials. All MR trials for ph-AFQMC were in the natural orbital (NO) basis. Additional details, energies, and ionization potentials are given in the Supporting Information.

All ph-AFQMC calculations utilized a population control (PC) scheme, in which walkers with large overlaps with the trials are duplicated while those with small weights are purged periodically. However, the above three-dimensional convergence could not be unambiguously shown for complex **25**. In this case, keeping 89, 91, and 93% of the CI weight in the trial wave functions produced ph-AFQMC IPs of 225.6 ± 1.3, 224.8 ± 1.1, and 221.4 ± 1.1 (see Table S23). Keeping a higher %CI in the trials is currently too computationally demanding with our current implementation, in an active space of 70 electrons in 78 orbitals with $\epsilon_1 = 5 \times 10^{-5}$ a.u. Therefore, we used a correlated sampling (CS) approach in this case, which produced an IP of 215.8 ±1.3 kcal·mol^{-1}. Interestingly, this value is not far from what one would obtain from extrapolating the three PC results above to the 100% CI weight limit. Indeed, CS has been shown to produce results that are less sensitive to the trial wave function employed (and thus closer to the exact, unconstrained result).[51] This method has previously been shown to be in agreement with exact/experimental values for vertical ionization potentials of metallocenes[280]; in difficult cases, CS ph-AFQMC has yielded superior results vs the PC approach for a fixed trial wave function. However, we did not use CS for the entire 3dTMV set because without the removal of walkers with small or vanishing weights it can be more expensive than the PC ph-AFQMC algorithm, and we preferred the ability to check total energies, and their convergence for MR systems.

Taken together, while we have made our best effort to produce accurate reference values with ph-AFQMC, we cannot claim to have achieved exact IPs especially for the cases in the MR subset. For the MR subset we report a total uncertainty which is the statistical error plus 3 kcal·mol^{-1}. We think this is a reasonable estimate given that other possible error metrics, e.g., the difference in IPs with ϵ_1 of 10^{-4} vs 5×10^{-5}, suggest an average difference of less than 1.5 kcal·mol^{-1}. While most of the IPs appear to be converged with respect to the trial wave function, as can be seen in the SI, some of the cases are strikingly sensitive to the quality of the trial wave function. For instance, for complex **23** the ph-AFQMC IP goes from 209.6 to 209.8 to 204.7 kcal·mol^{-1} in small, medium, and large active space sizes, and in the latter size the IP is 204.7, 203.2, 200.1, and 198.9 kcal·mol^{-1} with 85, 87, 89, and 90 %CI (all of these IPs have associated statistical error bars in the range of 1-1.2 kcal·mol^{-1}). Admittedly, the added uncertainty of 3 kcal·mol^{-1} for the MR cases was chosen somewhat arbitrarily. However, we note that the level of accuracy that we target is "transition metal chemistry accuracy" and that the spread of the various CC approaches for complex **23** easily exceeds 20 kcal·mol^{-1}.

Computational Details of the ph-AFQMC Calculations

All electrons were correlated in the ph-AFQMC calculations. Electron repulsion integrals and the trial wave functions were generated with PySCF.[142] We utilized a localized orbital implementation of ph-AFQMC[177] with a threshold of 5×10^{-5} a.u., in which occupied orbitals outside of the active space are localized, and the half-rotated Cholesky matrices are "compressed" using singular value decomposition.[177] Typical compression rates for the systems in the set were 60 – 80%. ph-AFQMC calculations used a mixed-precision scheme, wherein floating point operations carrying out the imaginary-time propagation were performed with double-precision, while two-electron integrals were stored in single-precision. Test calculations are shown for complex **1** in the SI, which shows that the Cholesky and localization thresholds along with the use of our approximate mixed-precision scheme, have negligible effects on the reported ionization energies, given the statistical error bars.

D.2.3 Multireference Diagnostics

Multiple proposed diagnostics for static correlation have been considered, and are detailed in the SI. The principal component analysis of the MR diagnostics was performed in the R statistical environment[365] (version 4.2.0).

For practical purposes, one can distinguish between diagnostics obtained from relatively cheap HF or DFT calculations and diagnostics obtained from higher-scaling correlated WFT methods. While, e.g., CCSD calculations, which formally scale with the sixth power of the system size in most canonical implementations, can often be performed for small to medium-sized molecules in a minimal or double-ζ basis set, in our experience, for systems with ≥50 atoms, even this is not computationally feasible. In this regime, not to mention when predictions must be extrapolated to the CBS limit, MR diagnostics that use SCF-level information are often the only option.

The HF- and DFT-based diagnostics include the deviation of the total spin expectation value of UHF or a determinant made from UPBE0 orbitals from the exact value:

$$\Delta\langle S^2\rangle = \langle S^2\rangle_{\text{method}} - \langle S^2\rangle_{\text{exact}}. \tag{D.1}$$

Spin-symmetry breaking (SSB) in approximate electronic structure theories which incorporate some degree of dynamic correlation has been put forth as a diagnostic of static correlation, since it encodes the physical effects (as can be seen in open-shell singlets) when low-lying excited states of higher multiplicity approach near-degeneracy and mix into a SR wave function.[86,95,154] In this work, we use a regularized percentage quantity:

$$\text{reg. } \Delta\langle S^2\rangle = \frac{100 \cdot |\Delta\langle S^2\rangle|}{MAX(\langle S^2\rangle_{\text{exact}}, 0.75)}. \tag{D.2}$$

where the regularization parameter of 0.75 was selected to weight the spin contamination in singlets equivalently as in doublets. The number of SCF iterations required to achieve convergence is also considered. The fractional occupation density (FOD)[366] was employed with the r^{2}SCAN[106,107] and the hybrid r^{2}SCAN50[67] functionals. In the FOD formalism, finite-temperature DFT is employed to enable fractional orbital occupations yielding the N_{FOD} value upon integration of their respective density. Because the N_{FOD} value is not size-consistent, the fractional occupation numbers (FON) from FOD calculations were expressed as Matitio's nondynamical correlation index,[367,368] which is a

size-consistent form suggested by Martin *et al.* ($r_{nd}(r^2SCAN/r^2SCAN50)$).[369]

With regard to diagnostics from correlated WFT methods, the well-known T_1 diagnostic,[370] the number of required iterations for the coupled-cluster amplitudes (#CC) and the spin contamination at the CCSD level were considered, with UHF and UKS reference determinants. Also, the leading coefficients of the trials (specified in the Supporting Information) in the NO basis for the ph-AFQMC calculations were also employed as static correlation indicators ($1 - C_0^2$). For systems with large spin-symmetry breaking (using cluster 1 $\langle S^2 \rangle$ diagnostics), we used HCI trial wave functions, i.e., without orbital optimization, to compute the reported C_0 values; otherwise, HCISCF wave functions were used.

D.3 Results

D.3.1 Analysis of Multireference Diagnostics

We perform a statistical analysis of the myriad MR diagnostics mentioned above, in an effort to correlate their predictions regarding the presence of static correlation, which will guide our classification of the 3dTMV complexes into subsets. In order to cluster the different diagnostics, a principal component analysis of the Pearson correlation matrix (visualized in Figure D.3**a**) was performed and diagnostics were clustered according to their contribution to principal components as shown in Figure D.3**b**. The first two principal components accounted for 73.3 % of the variance and were used to cluster the variables. Cluster 1 contains the WFT based diagnostics $\Delta\langle S^2\rangle_{CCSD}$, $1 - C_0^2$, and the $\Delta\langle S^2\rangle_{PBE0}$. Cluster 2 contains only $\Delta\langle S^2\rangle_{UHF}$. Cluster 3 contains the number of SCF iterations with HF and PBE0, the T_1(HF) diagnostic, the number of CC iterations for an HF reference, and the finite-temperature DFT based diagnostics (r_{ND}). Cluster 4 contains the T_1 diagnostic and the number of CC iterations with the PBE0 reference orbitals.

In cluster 1, $1 - C_0^2$ is directly related to the definition of MR character, in the sense that it indicates when more than one configuration has a large weight in the wave function. The $\Delta\langle S^2\rangle$ diagnostics in this cluster, which are computed from CCSD wave functions or PBE0 orbitals, have been demonstrated to reveal “essential" spin-symmetry breaking, i.e., the spin-contamination in theories which include dynamical correlation provides a better physical description of the static correlation. $\Delta\langle S^2\rangle_{UHF}$ is not part of this cluster, as one might expect, because HF (which does not formally include any Coulomb correlation) artificially stabilizes high spin states relative to low-spin states, and thus its spin-symmetry breaking behavior is not a reliable indicator of MR character.

The relatively good correlation between the T_1 diagnostic and the number of CC iterations in Clusters 3 and 4 can be interpreted in the following way: both indicate that CCSD requires many orbital rotations to remedy shortcomings in the reference wave function. We emphasize that large values do not necessarily indicate MR character but rather that the reference wave function is inadequate. Similar observations regarding the T_1 diagnostic have also been made by others.[285,297]

To obtain a subset of k principal variables that span a space similar to the variables in cluster 1, the variables included in this cluster were further analyzed with the "subselect" module developed by Cadima et al.[371] Within the module the "eleaps" algorithm with the GCD (generalized coefficient of determination) as the objective function was selected. The following subsets of increasing size were obtained:
$k = 1$: $\Delta\langle S^2\rangle_{PBE0}$, GCD=91.4%
$k = 2$: $(1 - C_0^2) + \Delta\langle S^2\rangle_{CCSD(PBE0)}$, GCD=97.5%

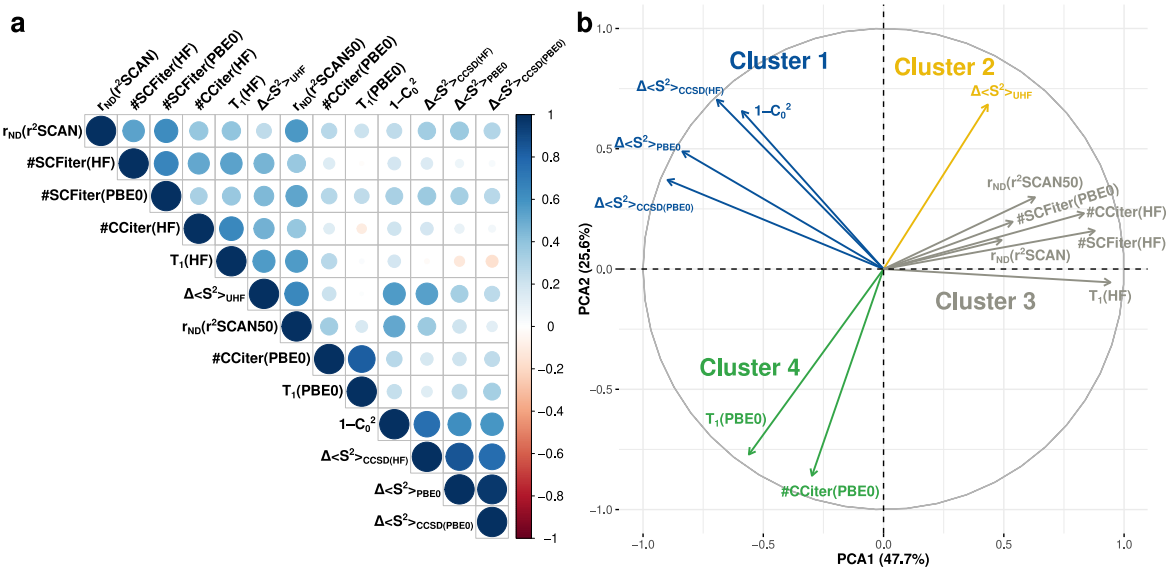

Figure D.3: (**a**) Visualized correlation matrix for the static correlation diagnostics. (**b**) The plot of the first two principal components (clusters are indicated by color).

$k = 3$: $(1 - C_0^2) + \Delta\langle S^2\rangle_{\text{CCSD(HF)}} + \Delta\langle S^2\rangle_{\text{CCSD(PBE0)}}$, GCD=99.6%

The large GCD of 91.4% for $k = 1$ with $\Delta\langle S^2\rangle_{\text{PBE0}}$ demonstrates that cluster 1 can be represented well by $\langle S^2\rangle_{\text{PBE0}}$ alone. This empirically validates the usefulness of $\langle S^2\rangle_{\text{PBE0}}$, especially since it is the variable of cluster 1 that is obtained with the least computational effort.

D.3.2 Classification into Subsets

Cluster 1 was employed to partition the 3dTMV IPs into subsets as depicted in Figure D.4. We note that although $1 - C_0^2$ reports directly on how dominant the configuration of the largest weight is in the linear superposition (and thus is the most physically transparent quantity), the values depend on the choice of active space and may be biased due to missing dynamical correlation effects. On the other hand, while physically justifiable, the degree of spin-symmetry breaking depends on the degree of inclusion of dynamical correlation (e.g., exhibits xc functional dependence) and in the CC context is derived from a perturbative analysis.[155] Therefore, in what follows, our classification will incorporate four quantities: $1 - C_0^2$ and spin-symmetry breaking or restoration from CC/UHF, CC/UPBE0, and UPBE0.

The partitioning was performed with the following criteria and the IP was assigned if either the initial or the oxidized state fulfilled them. IPs where neither the initial nor the oxidized species exceed a deviation for the regularized $\langle S^2\rangle$ value of 5% and a $1 - C_0^2$ value of 0.28 were put in the SR subset. Cases where only one or two of the regularized $\langle S^2\rangle$ values exceed the 5% deviation threshold and $1 - C_0^2$ is still below 0.28 were put in an intermediate category, which we denote SR/MR. If all three $\langle S^2\rangle$ diagnostics (in cluster 1) exceed the 5% deviation threshold, and $1 - C_0^2$ is above 0.28, the IP was assigned to the MR subset. There are some exceptions to the threshold based classification. For complexes **11** and **12** the oxidized species exhibit large $\langle S^2\rangle$ deviations but only small $1 - C_0^2$ values. At the SCF level the triplet state is predicted to be the ground state, resulting in what has been

referred to as "variational collapse"[86] when targeting the singlet state (see SI Section 4.4 for details). Subsequent CCSD calculations on these SCF solutions are not able to completely restore the SSB introduced. The large $\langle S^2 \rangle$ deviations for these complexes therefore do not indicate MR character and the corresponding IPs are sorted into the SR subset. Separately, complex **14** is a borderline case and was sorted into the SR/MR subset.

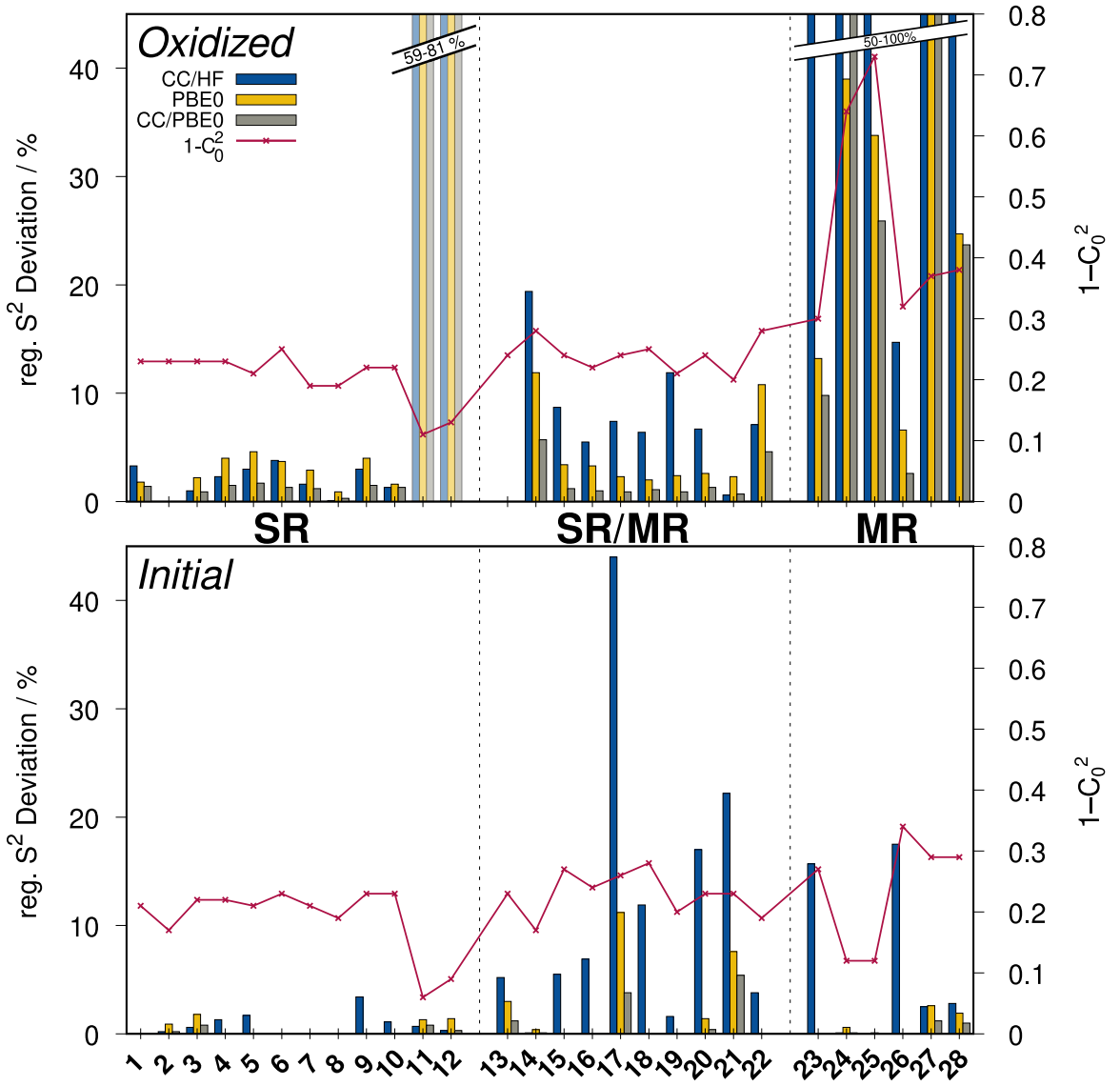

Figure D.4: Values of the cluster 1 diagnostics for species in the 3dTMV set. Oxidized species are shown at the top and initial species at the bottom. The left vertical axis shows the regularized $\langle S^2 \rangle$ deviations quantified as a percentage, i.e., $(|\langle S^2 \rangle - \langle S^2_{\text{exact}} \rangle|) * 100/MAX(\langle S^2 \rangle_{\text{exact}}, 0.75)$. C_0 values are from the multiconfigurational trial wave functions used in ph-AFQMC, as specified in the SI.

The values of the four diagnostics in cluster 1 are shown in Figure D.4 for the IPs classified into SR, SR/MR, and MR subsets. While nearly all of the complexes broke spin symmetry at the UHF level (see Table S2 in the SI), this symmetry breaking is artificial, certainly in the SR subset. In the

SR/MR subset, in almost all cases the SSB from CC/UHF is significantly reduced with CC/UPBE0, and UPBE0. In other words, there is a set of non-HF orbitals that reduces the spincontamination in the reference, which typically enables substantial spin symmetry restoration in the CCSD/UPBE0 wave function. It is noticeable that in almost all cases except **17** and **21**, one state, typically the oxidized one, shows significantly larger SSB than the other, which motivated us to test the EOM-IP-CC methods. Finally, the MR subset on average shows large SSB behavior vs the other two subsets. In many cases the $1 - C_0^2$ value is significantly above 0.28, in the range of $0.6 - 0.7$ for the oxidized states of the Cr- and Fe-centered acac complexes.

D.3.3 Comparison of CCSD(T) and LO-ph-AFQMC

Figure D.5 compares the deviation in the IPs calculated by three flavors of CCSD(T) vs. the reference ph-AFQMC values. Statistics for each of the three subsets are listed in Table D.2.

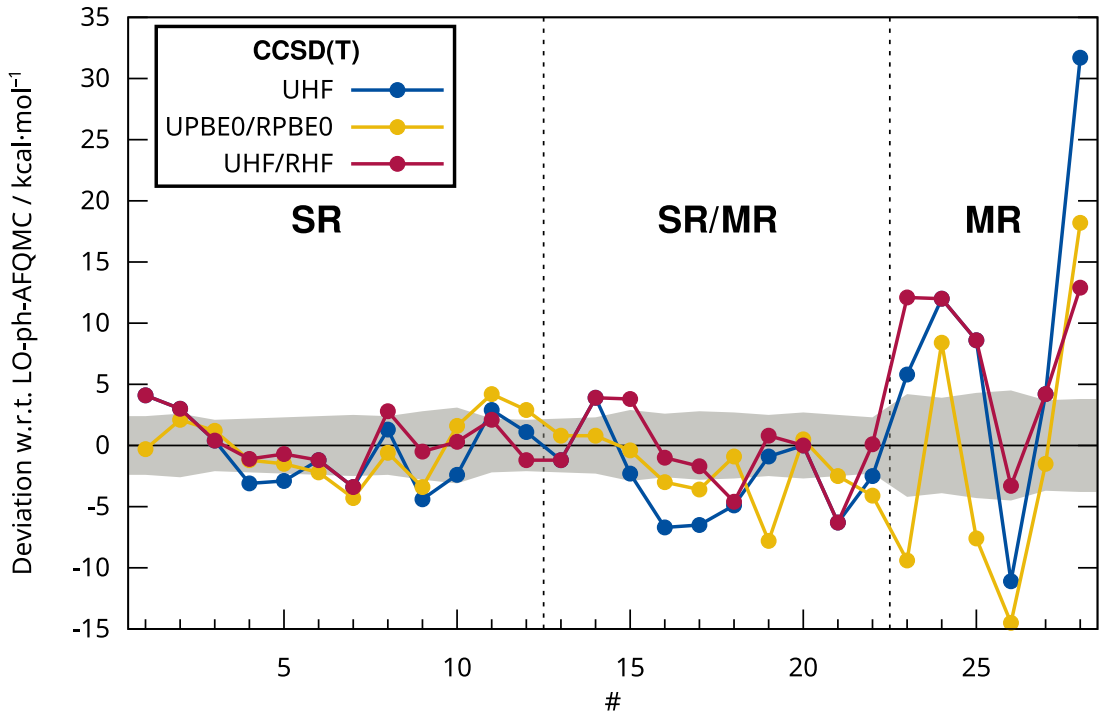

Figure D.5: Comparison of CCSD(T) with different orbitals with respect to LO-ph-AFQMC. All IPs are given in the SI in Table S9. The mean IP with LO-ph-AFQMC is 179.2 kcal·mol^{-1}. Dots are connected by lines to guide the eye. The uncertainty of LO-ph-AFQMC is depicted in grey.

For the SR subset, CCSD(T) yields accurate IPs, irrespective of the orbital set employed. All MADs are less than 3 kcal·mol^{-1}. The CCSD(T)/(UHF/RHF) protocol performs the best, with MAD of 1.7 kcal·mol^{-1} and an absolute max error of 4.1 kcal·mol^{-1}. The MADs are only slightly worsened for the SR/MR subset – all except the UHF and the ROHF/RHF orbital choices are still sub 3 kcal·mol^{-1} – though the maximum errors from UPBE0 orbitals are worsened by 2-3 kcal·mol^{-1}. The unreasonably large errors, especially as reflected in the MR subset, which result from CCSD(T) with ROHF/RHF references can be traced to a few cases with relatively high multiplicity states (**24**,**25**), which are

Table D.2: Statistical comparison between CCSD(T) with UHF, UPBE0, UHF/RHF, and ROHF/RHF orbitals (see section D.2.1 for details) with respect to LO-ph-AFQMC. The statistical quantities are mean deviation (MD), mean absolute deviation (MAD), standard deviation (SD), and absolute maximum deviations (AMAX), defined in the SI. Statistical evaluation with ROHF/RHF orbitals over the whole set without **24** and **25** from the MR set are given in parenthesis.

Set	CCSD(T)	MD	MAD	SD	AMAX
SR	UHF	−0.4	2.5	2.9	4.4
	UPBE0	0.0	2.3	2.8	4.6
	UHF/RHF	0.4	1.7	2.2	4.1
	UPBE0/RPBE0	−0.1	2.1	2.6	4.3
	ROHF/RHF	−1.6	2.7	3.3	9.7
SR/MR	UHF	−2.7	3.5	3.4	6.7
	UPBE0	−2.1	2.5	2.8	7.8
	UHF/RHF	−0.6	2.3	3.2	6.3
	UPBE0/RPBE0	−2.0	2.4	2.7	7.8
	ROHF/RHF	−3.3	3.9	4.1	10.5
MR	UHF	8.5	12.2	13.9	31.7
	UPBE0	−0.7	10.3	12.9	20.3
	UHF/RHF	7.8	8.9	6.3	12.9
	UPBE0/RPBE0	−1.1	9.9	12.3	18.2
	ROHF/RHF	−14.6	36.6	63.8	131.9
3dTMV	UHF	0.7	5.0	7.8	31.7
	UPBE0	−0.9	4.1	6.1	20.3
	UHF/RHF	1.6	3.5	4.9	12.9
	UPBE0/RPBE0	−1.0	3.9	5.8	18.2
	ROHF/RHF	−5.0 (−2.7)	10.4 (3.7)	28.1 (4.3)	131.9 (14.1)

improperly described at the ROHF level. In fact, this is a salient reason we consider the UHF/RHF protocol, which uses RHF for singlets and UHF otherwise.

As expected, CCSD(T) is clearly unreliable for the MR subset, with MADs ranging from 8.9 kcal·mol^{-1} with the UHF/RHF protocol to 36.6 with ROHF/RHF orbitals. The maximum absolute errors range from 12.9 to 131.9 kcal·mol^{-1}. Not only are the majority of IPs far from the ph-AFQMC references, but the sensitivity of the CCSD(T) predictions to the orbital set employed is also dramatically increased (Figure D.5). Thus, one clear takeaway from this study is actually in line with the common wisdom that computationally feasible SR methods such as CCSD(T) should not be used for MR systems. Our study confirms this statement, and qualifies it – as we have defined a concrete classification protocol that relies on relatively inexpensive diagnostics.

In all subsets, indeed for the whole 3dTMV set as well, the use of UPBE0 orbitals in CCSD(T) leads to slightly more accurate results than those of UHF orbitals. This is unsurprising given the large and pervasive artificial SSB at the UHF level (see the SI Table S2). Using restricted orbitals for the singlet states is the simplest way to prevent any spin symmetry breaking (for better or worse), and we find that the CCSD(T)/(UHF/RHF) protocol always leads to more accurate IPs. This is slightly

surprising in the context of the MR subset, since one might have expected spin-symmetry breaking, especially at the UPBE0 level, to lead to physically more appropriate electron densities (occupied orbitals). Nevertheless, the data presented here suggest that the UHF/RHF orbital choice leads to higher accuracy for all 3dTMV subsets.

It is interesting to note that all metal complexes with bipyridine ligands are classified in the SR/MR set (**15**, **16**, **17**, and **18**). Many aromatic rings such as benzene and naphthalene are substantially spin-contaminated at the UHF level,[197] as are these bipyridine ligand complexes. The SSB from UPBE0 is much reduced versus UHF, and that from CCSD/UPBE0 is much reduced versus CCSD/UHF (this can generally be seen in the SR/MR subset). For the above systems, the calculated IPs from CCSD(T)/UPBE0 are much closer to the ph-AFQMC values than those from CCSD(T)/UHF. However, there are a small number of cases in the SR/MR set for which CC/UPBE0 reduces the SSB versus CC/UHF but leads to a worse IP. **19** is one such case, and an analysis of the atomic spin densities proves illuminating. For this doublet oxidized state, UHF, although strongly spin-contaminated (see Figure D.6**a**), shows in principle the correct electron configuration, i.e., the unpaired electron is metal-centered. Upon introducing dynamic correlation when going to CCSD/UHF (Figure D.6**b**), the SSB is significantly reduced and the metal-centered radical can still be observed. In contrast, UPBE0, although almost without any SSB, implies that the radical is centered on the indenyl ligand (Figure D.6**c**). Apparently, CC based on this qualitatively incorrect reference cannot recover the expected electron configuration (cf. Figure D.5).

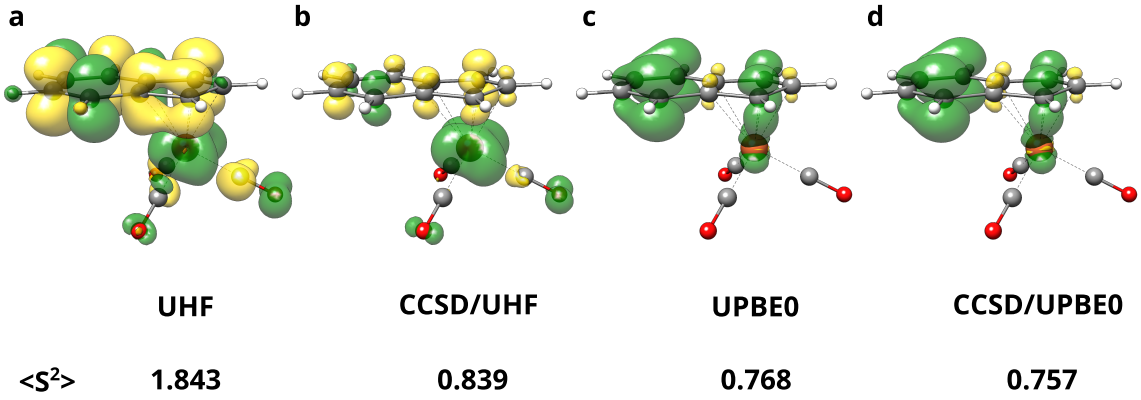

Figure D.6: Spin density plots with (**a**) UHF, (**b**) CCSD/UHF, (**c**) UPBE0, and (**d**) CCSD/UPBE0 for the the doublet oxidized state of **19**. α density is depicted in green and β density in yellow with an isovalue of 0.005 a.u. Mulliken CCSD charge and spin populations are shown in the SI in Table S10.

The effect of the Yamaguchi spin projection[372–375] on UCCSD(T) total energies was systematically investigated for CCSD(T) calculations with UHF, UPBE0, UHF/RHF, and UPBE0/RPBE0 orbitals. The formula for spin-projected energy of a low-spin state is:

$$E_{LS} = \frac{E_{BS} - (1-\alpha)E_{HS}}{\alpha}, \tag{D.3}$$

where LS indicates the spin-pure low-spin state, BS indicates the broken symmetry low-spin state, and

HS indicates the high-spin state. The spin-coupling coefficient α is calculated according to

$$\alpha = \frac{\langle S^2 \rangle_{HS} - \langle S^2 \rangle_{BS}}{\langle S^2 \rangle_{HS} - \langle S^2 \rangle_{LS}}. \tag{D.4}$$

We note that this expression is exact when the BS state can be written as a linear combination of exactly two states – the LS and HS eigenstates. This condition is not generally true, however it is very likely to be applicable when the HS state is spin-pure.

For cases where the CCSD regularized $\langle S^2 \rangle$ deviation is greater than or equal to 2%, the next higher-spin state (HS) (i.e., the original spin multiplicity plus two) was calculated and used in the projection if the HS regularized $\langle S^2 \rangle$ deviation was smaller than or equal to 2% (results utilizing 1, 3, and 4% are similar, and are shown in Section 5.5 of the SI). For IPs in which either the oxidized or the initial state did not agree with these criteria the Yamaguchi spin-projection was not applied for both of these states. The statistical results for spin-projected CCSD(T) with various choices of orbitals, vs the LO-ph-AFQMC reference values, are shown in Table D.3 (depicted in Figure S3 in the SI). We find that the Yamaguchi projection protocol does not lead to substantial changes in the accuracy statistics. For example, the MAD of CCSD(T) with UHF/RHF orbitals remains at 1.7 kcal·mol^{-1} for the SR set (although the AMAX is reduced from 4.1 to 3.6) and is reduced from 2.3 to 2.2 kcal·mol^{-1} in the SR/MR subset. Given the statistical error bars in the LO-ph-AFQMC reference calculations, these changes are not significant. For the MR set, the MAD of 8.9 with UHF/RHF orbitals is slightly reduced to 8.3 kcal·mol^{-1}, with rather large maximum errors still.

Table D.3: Statistical comparison of Yamaguchi corrected CCSD(T) methods.

Set	CCSD(T)	MD	MAD	SD	AMAX
SR	UHF+Y	−0.5	2.4	2.8	4.4
	UPBE0+Y	−0.1	2.2	2.7	4.3
	UHF+Y/RHF	0.3	1.7	2.2	3.6
	UPBE0+Y/RPBE0	−0.1	2.1	2.6	4.3
SR/MR	UHF+Y	−3.0	3.4	2.9	6.5
	UPBE0+Y	−1.5	2.3	3.0	7.8
	UHF+Y/RHF	−1.4	2.2	2.8	6.3
	UPBE0+Y/RPBE0	−2.1	2.4	2.7	7.8
MR	UHF+Y	8.0	11.1	13.7	31.7
	UPBE0+Y	−1.6	10.3	12.6	19.4
	UHF+Y/RHF	6.6	8.3	7.4	12.9
	UPBE0+Y/RPBE0	−1.8	10.1	12.2	18.2
3dTMV	UHF+Y	0.4	4.6	7.6	31.7
	UPBE0+Y	−0.9	4.0	6.0	19.4
	UHF+Y/RHF	1.1	3.3	4.9	12.9
	UPBE0+Y/RPBE0	−1.2	3.9	5.8	18.2

The observation that the oxidized states appear to be more MR than the initial states, on average,

motivated us to investigate the accuracy of EOM-IP-CCSD. Despite a few technical difficulties (**2**, **4**, and **20** which are discussed in the SI in Section 5.3), we performed EOM-IP-CCSD calculations with UHF/RHF reference orbitals. Deviations of the predicted vertical IPs vs AFQMC, along with those from CCSD and CCSD(T) with the same UHF/RHF orbital choice, are shown in Figure D.7**a**, and a statistical summary is shown in Table D.4. While EOM-IP-CCSD is slightly more accurate than CCSD (both with UHF/RHF orbitals) for both the SR and SR/MR subsets, we find the opposite for the MR subset. Figure D.7**b** shows the (T) contribution to the IP for each complex, and indeed the cases in the SR and SR/MR subsets in which this is large exhibit notable improvements going from CCSD to CCSD(T) (e.g., complexes **2**, **8**, **14**, **15**-**18**, **21**). In almost every case, the perturbative triples correction is larger with UKS orbitals than with UHF orbitals (cf. also Figure S2 and Table S12 of the SI). The KS orbitals are recanonicalized prior to subsequent CC calculations, therefore, this finding cannot simply be ascribed to smaller KS eigenvalue differences. While one might expect that the contribution of triples should correlate with MR character, it appears that it need not, in the sense that there are cases in the SR and SR/MR subsets that have large (T) contributions. Complex **2** shows a (T) contribution in excess of 10 kcal·mol^{-1}, which appears to be due to the bonded triflate anion. The bipyridines, **15**-**18**, also have large (T) contributions, as does complex **21** with the bound N_2. The latter is consistent with the finding that CCSD makes large errors in the bond dissociation of a system involving a triple-bonded ligand (c.f. $Cu(CO)_4^+$ from ref. [86]). Finally, we find that (T) corrections often play a large role in MR systems. This can be understood by considering the case of a diradicaloid (such as ozone). Double excitations are needed to produce a qualitatively correct open-shell singlet reference, while triples and higher-order excitations are needed to describe the dynamical correlation required for a quantitative description. In these MR transition metal cases, (T) is clearly not enough to describe the relatively large dynamical correlation.

To improve the description of correlation effects beyond CCSD, more sophisticated approximations for connected triples excitations such as the fully renormalized triple correction to CCSD (CR-CC(2,3)) developed by Piecuch *et al.*[359,360] may be needed. The advantage of this method over other approaches that resort to MR concepts to improve the conventional CCSD(T) approach is that it is only at most twice as expensive as the latter. In this work, due to technical restrictions (partly ROHF convergence problems), we computed only four systems (**1**, **2**, **7**, and **9**) from the SR set using the ROHF/RHF based CR-CC(2,3) implementation in GAMESS. As expected, the results for these SR systems are very similar to the ROHF/RHF CCSD(T) results (see Table S17 in the SI). It would be of interest to examine the SR/MR and MR subsets with methods such as CR-CC(2,3) (which would require a robust unrestricted implementation).

D.3.4 Preliminary DFT Evaluation

A statistical comparison between various DFT functionals with respect to LO-ph-AFQMC, all evaluated in the def2-SVP basis, is given in Figure D.8. Although this comparison is not completely justified because of the different degrees of BSIE between DFT and orbital-space methods (CC and AFQMC), we present these results as a preliminary indication of how the various DFT functionals are likely to perform.

The MDs of different functionals range from −0.9 kcal·mol^{-1} for PBE-D4 to 5.9 kcal·mol^{-1} for ωB97X-V (see SI Table S3-S5 for details). Thus, this preliminary comparison seems to be reasonably appropriate, although a slight trend to positive MDs is observed. The hybrid functional M06-2X yields the largest MAD, which is not surprising because it has been mainly designed for main group

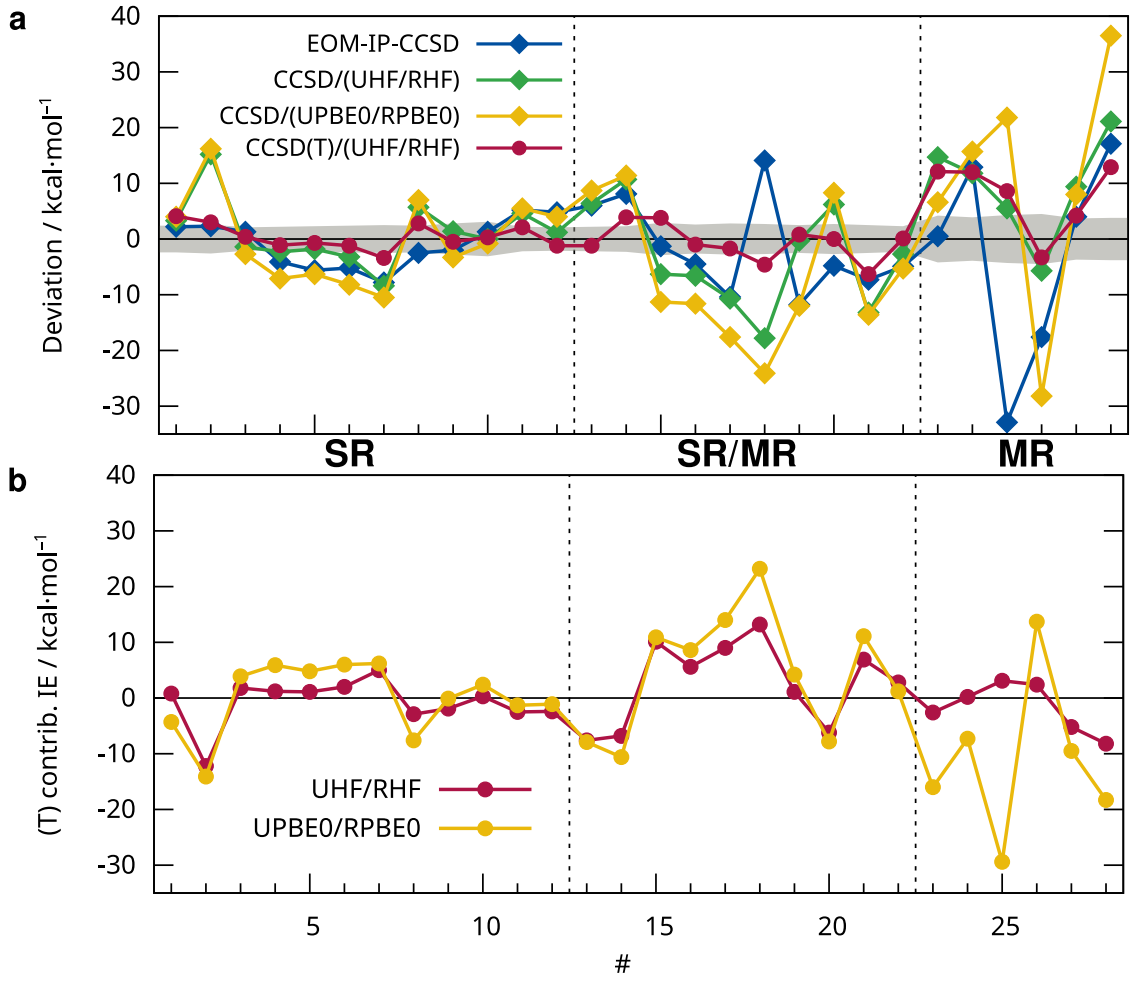

Figure D.7: (**a**) Comparison between EOM-IP-CCSD, CCSD, and CCSD(T) on UHF/RHF orbitals with respect to LO-ph-AFQMC. (**b**) Triples contribution (T) for the IPs in CCSD(T) with UHF/RHF and UPBE0/RPBE0 orbitals.

Table D.4: Statistical comparison between EOM-IP-CCSD, CCSD and CCSD(T) on UHF/RHF orbitals and CCSD on UPBE0/RPBE0 orbitals with respect to LO-ph-AFQMC.

Set	Method	MD	MAD	SD	AMAX
SR	EOM-IP-CCSD/(UHF/RHF)	-0.8	3.7	4.3	7.8
	CCSD/(UHF/RHF)	1.2	4.0	5.8	15.2
	CCSD/(UPBE0/RPBE0)	−0.2	6.3	7.7	16.2
	CCSD(T)/(UHF/RHF)	0.4	1.7	2.2	4.1
SR/MR	EOM-IP-CCSD/(UHF/RHF)	−1.7	7.3	8.4	14.1
	CCSD/(UHF/RHF)	−3.4	8.1	9.3	17.8
	CCSD/(UPBE0/RPBE0)	−6.7	12.4	12.2	24.1
	CCSD(T)/(UHF/RHF)	−0.6	2.3	3.2	6.3
MR	EOM-IP-CCSD/(UHF/RHF)	−2.7	14.2	19.1	32.9
	CCSD/(UHF/RHF)	9.5	11.4	9.1	21.1
	CCSD/(UPBE0/RPBE0)	10.1	19.5	21.7	36.5
	CCSD(T)/(UHF/RHF)	7.8	8.9	6.3	12.9
3dTMV	EOM-IP-CCSD/(UHF/RHF)	−1.5	7.2	10.0	32.9
	CCSD/(UHF/RHF)	1.3	7.1	9.0	21.1
	CCSD/(UPBE0/RPBE0)	−0.3	11.3	14.1	36.5
	CCSD(T)/(UHF/RHF)	1.6	3.5	4.9	12.9

chemistry applications. It is followed by the GGA functional PBE-D4 (MAD = 11.3 kcal·mol^{-1}) and the KP16 hybrid functional (MAD = 10.8 kcal·mol^{-1}). The meta-GGAs r^{2}SCAN-D4 and M06-L show an improvement over PBE-D4 with MADs of 8.7 kcal·mol^{-1} and 7.9 kcal·mol^{-1} respectively. The hybrid functionals PBE0-D4, r^{2}SCAN0-D4, B3LYP-D4, and PW6B95-D4 all perform relatively well with overall MADs between 7.8 kcal·mol^{-1} and 6.8 kcal·mol^{-1} with B3LYP being the best performer. It is noticeable that from these hybrid functionals B3LYP has the smallest amount of Fock exchange (20%). A worse MAD of 9.1 kcal·mol^{-1} is obtained with ωB97X-V, but an improvement to an MAD of 6.1 kcal·mol^{-1} is observed when an optimal tuning (OT) procedure is applied for each IP separately (see SI section 4.3 for details), yielding the overall best performing DFT method OT-ωB97X-V. Typical OT-ω values on this set are around 0.15 and therefore $\omega = 0.15$ as global parameter was also tested yielding the ωB97X15-V method, which yields almost the same overall MAD of 6.2 kcal·mol^{-1}. The local hybrid functional LH20t-D4 yields an MAD of 7.3 kcal·mol^{-1} and the double-hybrid PWPB95 an MAD of 7.4 kcal·mol^{-1}.

On average, as with CCSD(T) methods, the MAD for the SR subset is the smallest followed by those of the SR/MR and MR subsets. While the differences in MAD between the SR and SR/MR subsets are small for most functionals, the MAD for the MR subset is between 2 and 3 times larger than that of the SR/MR subset in most cases. There are two exceptions to this: for the KP16 functional the MR set has an MAD of 14.0 kcal·mol^{-1} compared to 9.3 kcal·mol^{-1} for the SR/MR subset, and the difference is small but the overall performance is already bad compared to the other hybrid functionals. This could be due to the nonself-consistent evaluation of B3LYP orbitals. The other exception is the PWPB95-D4 functional where the SR/MR set has an relatively large MAD of 8.5 kcal·mol^{-1} and the MR subset has an MAD of 10.9 kcal·mol^{-1}. In the SR/MR set relatively large errors for the

bpy-complexes **15-18** on the order of 10-15 kcal·mol^{-1} are observed. This is in accordance with the large (T) contribution for these complexes. Apparently higher-order correlation effects that are not covered by the pairwise-additive MP2 part[84,86] in PWPB95-D4 are crucial for these complexes and therefore it fails to produce reliable results.

The Yamaguchi spin projection was also evaluated for DFT with the ωB97X15-V functional. Here, the same criteria as for UCCSD(T) were applied and resulted in a small reduction in the overall MAD from 6.2 to 6.0 kcal·mol^{-1}.

In order to estimate the remaining BSIE, the PBE0 and the PWPB95 functionals were evaluated with the def2-QZVPP basis set and the def2-SVP IPs were compared to these results (see Table S6 in the SI for details). For PBE0 this yields an MD of -0.9 kcal·mol^{-1} and an MAD of 1.7 kcal·mol^{-1} relative to the QZ basis with the largest deviation of -3.7 kcal·mol^{-1} for complex **23**. The double hybrid PWPB95 is expected to be particularly sensitive to the basis set, as a part of its correlation energy is computed by DFT, while the other part is from MP2. Indeed, with PWPB95 this difference is larger and an MD of -3.7 kcal·mol^{-1} and an MAD of 4.0 kcal·mol^{-1} is obtained. The largest deviation of -9.3 kcal·mol^{-1} is obtained for complex **5**. Novel DFT developments such as local hybrids incorporating a strong correlation factor[376] and double-hybrid functionals with regularized MP2[84,377] contribution are interesting candidates for testing on this newly compiled set.

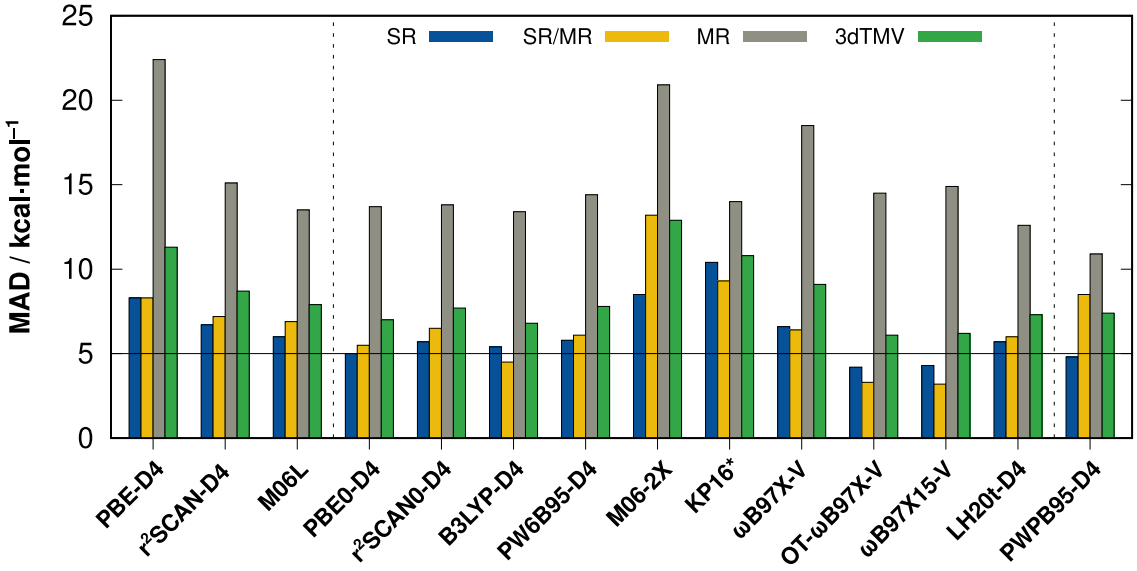

Figure D.8: MAD with respect to the LO-ph-AFQMC reference values for the three subsets and the whole 3dTMV set. The KP16 functional could not be evaluated self consistently and was therefore evaluated on B3LYP orbitals.

D.4 Conclusions

Electron transfers involving transition metal catalysts are ubiquitous in chemistry, but are difficult to model accurately with approximate quantum chemical methods due to the presence of both dynamic and static electron correlation. In this regime, the appropriateness of SR CCSD(T) – the “gold

standard" computational level for the majority of chemically relevant systems – is still debatable. In the absence of gas-phase experimental values, we leverage the unique scalability and high accuracy of auxiliary-field quantum Monte Carlo to provide reference vertical ionization energies. We compile a set of 28 3d complexes relevant to homogeneous electrocatalysis (which we refer to as 3dTMV), classify them into subsets based on the degree of multireference character in the involved states, and assess various CC protocols. Mean absolute deviations roughly equal or less than 2 kcal·mol^{-1} can be achieved for the predominately SR subsets – namely, with unrestricted/restricted reference orbitals for non-singlet/singlet spin states – and we confirm that CCSD(T) is an inappropriate model for strongly correlated transition metal complexes.

This work demonstrates that one valuable application of AFQMC, which in recent years has undergone rapid development and optimization, is to produce reference values for transition metal thermochemistry. Compared to the CCSD(T) calculations performed in this work, which required wall-times ranging from 2 to 48 h with 8 cores, LO-ph-AFQMC calculations are relatively more expensive yet can be trivially parallelized. For example, complexes in the SR and SR/MR subsets required between 40 and 150 GPU-node hours, while some MR molecules using trials with the largest active space size took up to 1200 node hours, which corresponds to a wall-time of 12 h on 96 OLCF Summit nodes. While challenging to converge away the phaseless bias for multireference states, in this work, we demonstrate that this is possible for realistic monometal electrocatalysts, at least in a double-ζ basis set. Current work is underway to approach the complete basis set limit, which will be necessary to properly assess and develop approximate density functionals and to compare with gas-phase experiments. We acknowledge that a truly robust and predictive computational protocol for realistic electrocatalysis must incorporate solvation free-energies and finite-temperature effects – from this perspective, the present work is a promising first step.

Another notable achievement is that we have pinpointed quantitative metrics based on symmetry breaking and the largest coefficient in selected CI multideterminant expansions, which are meaningful and effective in diagnosing MR character. Specifically, we propose thresholds for spin-symmetry breaking from unrestricted CCSD and KS-PBE0 along with the metric $1 - C_0^2$ from ph-AFQMC trial wave functions, which can delineate regimes inside of which appropriately-performed CCSD(T) can produce < 2.3 kcal·mol^{-1} accuracy, and outside of which CCSD(T) with the investigated choices of orbitals can be expected to fail. Our analysis of the many MR diagnostics proposed in the literature suggests that $\langle S^2 \rangle_{PBE0}$ is a computationally-inexpensive proxy which, while admittedly not rigorous, is herein found empirically to be practically useful in assessing the regime of applicability of CCSD(T) methods.

Among the many implications of this work on best-practices in quantum chemistry for transition metal systems, we propose that for target molecules that can be classified as SR or SR/MR, the agreement of ph-AFQMC and CCSD(T) with UHF/RHF orbitals can be expected, and consensus predictions ought to be more reliable than predictions from one of the two methods alone. In fact, occasionally in this work CCSD(T) IPs were used to guide the choice of CASSCF or HCISCF trial wave function used in ph-AFQMC; indeed, while an MCSCF optimization may converge to a qualitatively incorrect local minimum that is "closest" to the reference state used to initialize the calculation – thus inheriting an electronic state with, e.g., unphysical spatial symmetry – the $\hat{T}_1$ operator in the CC ansatz makes CCSD(T) relatively less sensitive to the reference used.

The exciting development of localized orbital approximations such as, e.g., the PNO-LCCSD(T)-F12[62] and DLPNO-CCSD(T) implementations,[59–61] when appropriately converged,[176] can also be readily used in these regimes to provide reliable reference values. With these localized coupled cluster

implementations, the use of extended basis sets is possible and thus a more realistic evaluation of DFT methods as well as a comparison with experimentally measured ionization energies. Our (preliminary) evaluation of selected DFT functionals in a rather smaller basis set revealed that ωB97X-V with a lowered ω value of 0.15 may be well suited for application in computational studies involving 3d transition metal electrocatalysts.

Finally, we remark that large, orbital-dependent triples contributions for complexes with, e.g., triflate and fluorine atoms, bipyridine, and N_2 (triple bond) provide new opportunities to assess alternate and develop improved approximate triples variants in the SR and SR/MR correlation regime, such as the renormalized coupled cluster methods.[359,360]

In summary, we have taken important first steps toward the reliable modeling of chemically relevant 3d transition metal electrocatalysts. We envision that future improvements in CC and AFQMC methods, used in combination with effective MR diagnostics and solvation models, can be used to predict reference-quality thermochemical values involving electronic states spanning a wide variety of correlation regimes. The curation of new transition metal data sets will also accelerate the development of faster quantum chemical or data-driven methods as well.

D.5 Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00617.

- DFT IPs, MR descriptors, OT-omega values (XLSX)
- Cartesian coordinates of the 3dTMV (ZIP)
- IPs of all coupled cluster and ph-AFQMC calculations; contribution of (T) to CCSD(T) correlation energy; IPs and statistics for Yamaguchi projection with CCSD(T); basis set study and statistics with DFT; and additional details of LO-ph-AFQMC calculations (PDF)

D.6 Acknowledgements

The German Science Foundation (DFG) is gratefully acknowledged for financial support (Grant 1927/16-1). This research used resources from the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC05-00OR22725. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. Calculations used the XSEDE resource Expanse at the SDSC through allocation ID COL151. J.S. acknowledges funding from the National Institute of General Medical Sciences of the National Institutes of Health under award number F32GM142231. This research used resources from the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory. The authors are thankful to Achintya Kumar Dutta for stimulating discussion and helpful advice about EOM-CCSD methods and to Stefan Grimme for proofreading. The authors greatly acknowledge the help from Piotr Piecuch, Jun Shen, and Jorge Deustua concerning the CR-CC(2,3) calculations. Furthermore, the authors thank Demyan Prokopchuk for helpful literature recommendations, Jing Kong for providing the xTron code and for technical support with setting it up, and James Smith for useful discussions about HCISCF. The authors acknowledge the Bill and Melinda Gates Foundations.

APPENDIX E

Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts

Práxedes Sánchez,[§†] Bhumika Goel,[§†] Hagen Neugebauer,[§‡] Roger A. Lalancette,[†] Stefan Grimme,[‡] Andreas Hansen,[‡] Demyan E. Prokopchuk[†]

Received: 9 October 2021
Published online: 4 November 2021

Reprinted in Appendix E (adapted) with permission[¶] from
P. Sánchez, B. Goel, H. Neugebauer, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts*, Inorg. Chem. **60** (2021) 17407, DOI: 10.1021/acs.inorgchem.1c03142

Own contributions

- performing all quantum chemical calculations
- interpretation quantum chemical results
- writing and revising parts of the manuscript concerning DFT calculations
- none of the experiments or measurements were conducted by me

[†]Department of Chemistry, Rutgers University Newark, Newark, New Jersey 07102, United States
[‡]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms Universität Bonn, Bonn 53115, Germany
[§]These authors contributed equally.
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

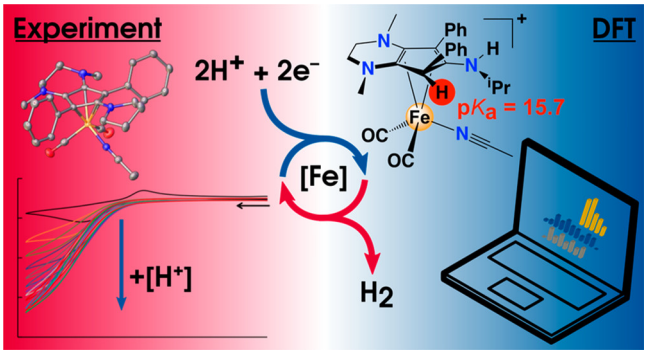

Figure E.1: Table of content graphic (ToC).

Abstract We present monometallic H_2 production electrocatalysts containing electron rich triamine-cyclopentadienyl (Cp) ligands coordinated to iron. After selective CO extrusion from the iron tricarbonyl precursors, electrocatalysis is observed via cyclic voltammetry in the presence of exogenous acid. Contrary to the fact that amines in the secondary coordination sphere are often protonated during electrocatalysis, comprehensive quantum chemical calculations indicate that the amines likely do not function as proton relays; instead, endo-Cp ring protonation is most favorable after $1e^-$ reduction. This unusual mechanistic pathway emphasizes the need to consider a broad domain of H^+/e^- addition products by synergistically combining experimental and theoretical resources.

E.1 Introduction

Earth-abundant metals (EAMs) are essential to life, catalyzing complex redox reactions involving multiple protons and electrons. In Nature, [Fe-Fe] and [Ni-Fe] hydrogenases use azadithiolate and arginine moieties, respectively, to shuttle protons between the metal's primary coordination sphere and these nitrogen bases.[378–380] Consequently, this behavior has inspired the development of synthetically tractable electrocatalyst biomimics containing amine moieties capable of facilitating proton movement for H_2 production and/or oxidation using EAMs.[16,324,381–384] As ligand designs for fuel-forming electrocatalysis are refined to incorporate secondary coordination sphere interactions,[385–388] mechanistic scenarios become increasingly complicated due to multiple donor/acceptor moieties mediating the movement of protons and electrons. Thus, synergistic cooperation between experimentalists and theoreticians is essential to identify plausible mechanistic pathways and avoid energy pitfalls that could stifle catalytic activity.[36,38,382]

While η^5-C_5H_5 (Cp) and electron-rich η^5-C_5Me_5 (Cp*) often function as ancillary ligands, they can also participate in C-H bond activation reactions.[389–391] Mechanistic studies have proposed the existence of η^4-CpH or η^4-Cp*H intermediates using Co,[392] Rh,[393], Re,[394] and f-block[395] complexes, and there is synthetic precedent for η^4-Cp ligation at Fe,[396,397] Co,[398,399] and Rh.[398,400] Experimental and computational analyses have demonstrated that active participation of the Cp/Cp* ligand can be essential in the context of proton-coupled electron transfer (PCET),[392,401–403] driving the need to evaluate new Cp derivatives for small molecule activation and catalysis.

In contrast, amine-functionalized Cp ligands remain underexplored for applications in catalysis.[404–406] Notably, the electron rich triamine-cyclopentadienylium ligand $^{\mathbf{en}}\mathbf{Cp}^{\mathbf{NH}^{\mathbf{i}}\mathbf{Pr}+}$ (**1**, Scheme E.2; en = N,N'-dimethylethylenediamine) was used to synthesize a piano-stool iron carbonyl complex which

catalyzes the reductive amination of secondary amines.[406] Computational studies suggested that a pendant amine on ligand 1 facilitates proton transfer during substrate reduction, inspiring us to synthesize a library of $^{en}\mathbf{Cp}^{R+}$ ligands for applications in H_2 production electrocatalysis (Figure E.2). Due to the multitude of potential protonation sites, we rationalized that comprehensive multi-level modeling methods[81,122] considering conformational flexibility, tautomers, and solvent/solute interactions should be utilized to understand the experimental results.

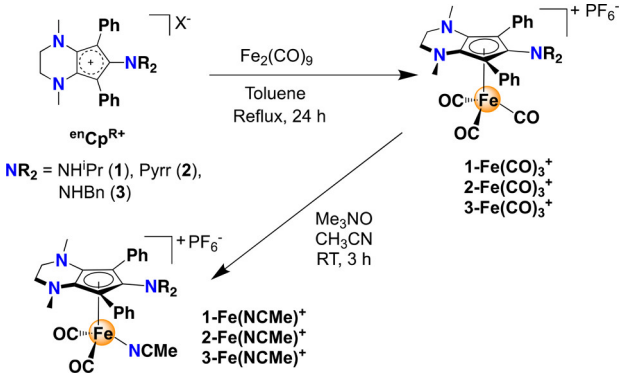

Figure E.2: Synthesis of Iron Complexes

We describe the synthesis, electrochemistry, H_2 production electrocatalysis, and quantum chemical modeling of Fe complexes with the electron rich triamine-cyclopentadienyl ligand $^{en}\mathbf{Cp}^{R}$ (R = NHiPr, pyrrlolidinyl (Pyrr), NHBn). Surprisingly, thermochemical landscape analysis reveals that direct protonation of the Cp ring is most favorable, generating a key endo-Fe(η^4-CpH) intermediate. Therefore, the experimental and computational data collectively suggest that the pendant amines behave as *ancillary* electron-rich functional groups, providing evidence that amines may not always function as proton relays due to pK_a mismatches between the exogenous proton source and amine functionalities.

E.2 Results and Discussion

The amine-functionalized cyclopentadienylium salts **1** and **2** ($^{en}\mathbf{Cp}^{R+}$; R = NH^{i}Pr, Pyr) were prepared as previously reported[407] while ligand **3** (R = NHBn) was prepared using a similar protocol (Figure E.2). Reaction of these salts in refluxing toluene with $Fe_2(CO)_9$ yields air stable Fe(II) tricarbonyl complexes **1-Fe(CO)$_3^+$**,[406] **2-Fe(CO)$_3^+$**, and **3-Fe(CO)$_3^+$** which were isolated as PF_6^- salts after workup. X-ray crystallography reveals isostructural motifs for all three complexes, with nearly equidistant C-C bond lengths connecting the carbon atoms within the Cp ring (Figure S11). When compared to the CO stretching frequencies of Cp*Fe(CO)$_3^+$ (ν_{CO} = 2130, 2078 cm^{-1}; Nujol),[408] solid-state IR data for **1-Fe(CO)$_3^+$** exemplifies the potent electron donor ability of these $^{en}\mathbf{Cp}^{R}$ ligands (ν_{CO} = 2048, 1969 cm^{-1}; KBr).

Cyclic voltammetry (CV) curves for complexes **1-Fe(CO)$_3^+$**, **2-Fe(CO)$_3^+$**, and **3-Fe(CO)$_3^+$** are shown in Figure E.3, and data are summarized in Table E.1. All three complexes exhibit partially reversible or irreversible redox features, whose currents and potentials remain constant with repeated linear potential sweeps. Although the pendant amines vary with respect to sterics and electronics, there is a reduction peak around E_{pc} = −1.6 V in all three cases, attributed to a diffusion-controlled 1e$^-$

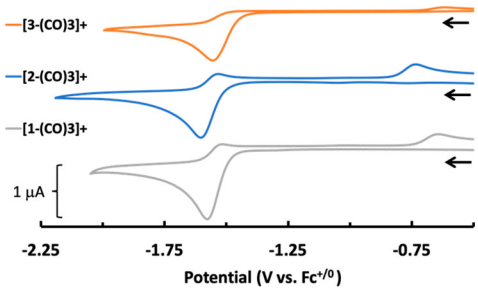

Figure E.3: CV curves of **1-Fe(CO)$_3^+$**, **2-Fe(CO)$_3^+$**, and **3-Fe(CO)$_3^+$**. Conditions: CH3CN, 0.1 M [Bu_4N][PF_6], 1.0 mM analyte, 100 mV/s.

Table E.1: Electrochemical Data for the **Fe(CO)$_3^+$** Complexes. DFT-calculated redox potentials are given in parentheses.

Complex	$E_{1/2}$ (V)	E_{pa} (V)	k_{CO} (s^{-1})
1-Fe(CO)$_3^+$	–1.59 (–1.25)	–0.62 (–0.61)	2.0 ± 1.1
2-Fe(CO)$_3^+$	–1.57 (–1.22)	–0.72 (–0.59)	6.0 ± 1.4
3-Fe(CO)$_3^+$	–1.55 (–1.23)	–0.60 (–0.60)	11.9 ± 3.6

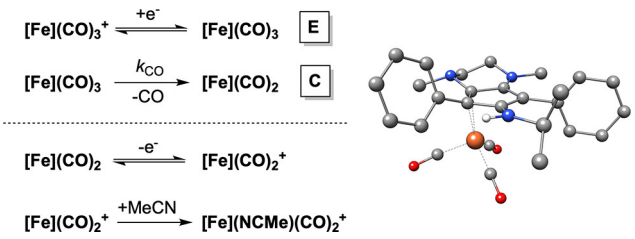

Figure E.4: Elementary Steps Showing Irreversible CO Loss by an EC Mechanism Followed by Solvent Uptake after Reoxidation (Left) with Computed η^2-($^{en}Cp^R$) Intermediate 1-Fe(CO)$_3$ (Right). [Fe] = Fe(**$^{en}Cp^R$**), where R = NH^iPr, Pyrr, Bn.

reduction event which becomes more reversible with faster scan rates (Figure S16). In all three cases, this behavior is generally described by an EC reaction mechanism where the rapid, reversible reduction of **[Fe](CO)$_3^+$** is followed by irreversible CO loss to generate the neutral iron dicarbonyl intermediate **[Fe](CO)$_2$** (Figure E.4, top). The rate of CO dissociation (k_{CO}) was calculated by plotting the change in cathodic peak potential (E_{pc}) versus the natural logarithm of inverse scan rate (ln(1/(ν)), (Eqn S1).[409,410] Complexes **1-Fe(CO)$_3^+$** and **2-Fe(CO)$_3^+$** have similar CO dissociation rate constants (k_{CO} = 2.0 ± 1.1 and 6.0 ± 1.4 s^{-1} respectively) while **3-Fe(CO)$_3^+$** is significantly faster (k_{CO} = 11.9 ± 3.6 s^{-1}). Since the redox potentials for all three complexes are extremely similar, the bulkier benzyl group on **3-Fe(CO)$_3^+$** suggests that ligand sterics may enhance the rate of CO dissociation. These k_{CO} values are all much lower than observed for the rate of CO loss from tricarbonyl complex $CpFe(CO)_3$ ($k_{CO} \geq 10^3$ s^{-1}), which undergoes a strictly dissociative mechanism upon reduction (i.e., no $\eta^5 \rightarrow \eta^3$ ring slippage).[320] For these new Fe complexes, computational data suggest that a formally $19e^-$ intermediate is avoided by a ring slip (haptotropic shift) upon reduction, generating an unusual Fe(η^2-**$^{en}Cp^R$**) intermediate (Figure E.4, right and Figure S25). However, the DFT-calculated potentials for $1e^-$ reduction of **1-Fe(CO)$_3^+$**, **2-Fe(CO)$_3^+$**, and **3-Fe(CO)$_3^+$** are anodically shifted by ca. +0.3 V

compared to experiment. If rapid ring slippage were to occur upon $1e^-$ reduction, the decrease in coordination number would anodically shift the observed redox potential as predicted by the DFT calculations (see SI for discussion). Therefore, the experimental data suggests that electron transfer is rapid and subsequent ring slippage is slow, leading to more negative observed redox potentials and retention of the η^5-$^{en}\mathbf{Cp}^{R}$ coordination mode on the electrochemical timescale.[411] Finally, an irreversible oxidation observed at E_{pa} = −0.6 to −0.7 V is postulated to be re-oxidation of the neutral Fe(I) dicarbonyl intermediate **[Fe](CO)$_2$** followed by irreversible coordination of MeCN, which is in excellent agreement with experiment (Table E.1).

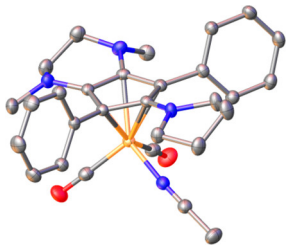

Figure E.5: X-ray crystallographic structure of **2-Fe(NCMe)$^+$** with 50% probability ellipsoids. Hydrogen atoms and PF_6^- are omitted for clarity.

On the basis of previous electrochemical studies using $[CpFeL(CO)_2]^+$ and $CpFeX(CO)_2$ complexes,[319,412,413] loss of L or X from these coordinatively saturated piano stool complexes is a prerequisite for H_2 production. Thus, a CO ligand is extruded by treatment with 1 equiv CH_3NO to generate the $18e^-$ adducts **1-Fe(NCMe)$^+$**, **2-Fe(NCMe)$^+$**, and **3-Fe(NCMe)$^+$** (Figure E.2). The structural assignment of **2-Fe(NCMe)$^+$** was confirmed by single crystal X-ray diffraction (Figure E.5). CV curves of these three complexes in acetonitrile show broadened and partially reversible redox waves, attribute to an EC mechanism with reversible solvent binding at iron (Figure S18 (black trace) in the SI).[410] This behavior has been studied in depth using $[CpCo(diphosphine)(NCMe)]^+$ salts, where the NCMe dissociation rate constant is on the order of 104-107 s^{-1} in non-coordinating solvents.[392,409] CV curves of **1-Fe(NCMe)$^+$**, **2-Fe(NCMe)$^+$**, and **3-Fe(NCMe)$^+$** in fluorobenzene at 100 mV/s reveal loss of the anodic feature, qualitatively indicating that solvent loss is facile upon electrochemical reduction (Figure S22 in the SI). In addition, the observed redox potentials in acetonitrile are in good agreement with theory (Table E.2).

Table E.2: Electrocatalysis Data for the **Fe(NCMe)$^+$** Complexes. Calculated redox potentials ($E_{1/2}$) are given in parentheses.

complex	$E_{1/2}$ (V)	$E_{cat/2}$ (V)	kcat (s^{-1})	η (V)	FE (%)
1-Fe(NCMe)$^+$	−1.61 (−1.76)	−1.72	29	0.98	65 ± 3
2-Fe(NCMe)$^+$	−1.61 (−1.67)	−1.76	35	1.02	65 ± 3
3-Fe(NCMe)$^+$	−1.58 (−1.74)	−1.72	45	0.98	64 ± 1

Next, these dicarbonyl solvento adducts were tested for H_2 production electrocatalysis with the exogenous acid 4-methyl-N-tosylbenzenesulfonamide (Tos_2NH; pK_a^{MeCN} = 11.97).[414] Initial screening with 20 mM acid reveals an S-shaped current response at cathodic potentials, whose plateau current (i_{cat}) becomes independent of scan rate at ≥ 800 mV/s (Figure S17 in the SI).[410] The relatively shallow slopes may be due to slow electron transfer from the electrode to the metal complex (Figure

S18 in the SI).[415] For **1-Fe(NCMe)$^+$**, successive additions of acid trigger the appearance of increased peak currents, attributed to H_2 production at a half-wave catalytic potential ($E_{cat/2}$) of –1.72 V (Figure E.6). When Tos_2NH < 70 mM, there is a linear dependence between k_{obs} and [acid], indicating a first-order dependence on [acid]. When [Tos2NH] > 70 mM, the observed rate constant k_{obs} becomes independent of [acid], with the maximum observed turnover frequency for H_2 production calculated to be 29 s^{-1}.

The reaction is first order in catalyst, displaying a linear dependence on [Fe] with constant [acid] (Figure S19 in the SI). Importantly, CV control experiments comparing 20 mM Tos_2NH solutions in the absence and presence of **1-Fe(NCMe)$^+$** clearly indicate that the iron catalyst is responsible for the observed current enhancement (Figure S21). Using a weaker acid such as $HNEt_3^+$ (pK_a^{MeCN} = 18.82)[416] results in no current enhancement in CV traces, indicating that stronger acids are essential for H_2 production to occur in the applied potential window (Figure S21).

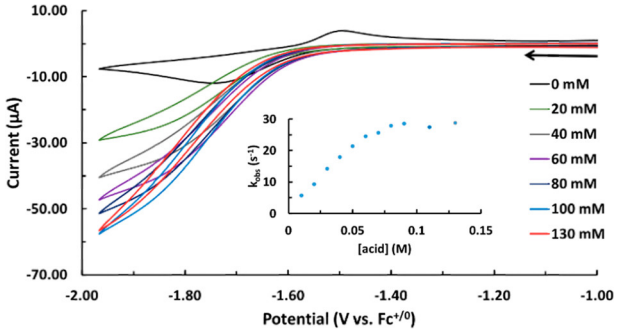

Figure E.6: CV curves of **1-Fe(NCMe)$^+$** with increasing concentrations of Tos_2NH. Conditions: MeCN, 0.1 M $[Bu_4N][PF_6]$, 1.0 mM analyte, 800 mV s^{-1} . Inset: Dependence of k_{obs} on the acid concentration.

Table E.2 summarizes the electrocatalysis data. Similar turnover frequencies, modest Faradaic efficiencies,[417] and similar overpotentials (η)[418] are found for this triad, suggesting that the amine variation has a minimal effect on catalytic activity. "Rinse test" electrolysis following a typical controlled potential electrolysis trial indicates that the iron complexes are mainly responsible for H_2 production (Figure S24 and Table S2).[417]

For the DFT calculations a multi-level protocol was employed, where different levels of theory were used at different steps to obtain reliable free energy estimates with best efficiency (Figure E.7).

Structure ensembles were generated by CREST[79] with the semi empirical quantum mechanical (SQM) GFN2-xTB[71,72] method. The low computational costs of GFN2-xTB allow extensive exploration of the chemical space on a quantum mechanical level. The structures lowest in energy were picked and optimized with the efficient composite r^2SCAN-3c[137] DFT method in combination with the implicit solvation model DCOSMO-RS[159] to account for solvation effects in the geometry. Employing these optimized structures, higher level single point energy calculations were performed with r^2SCAN-D4[106,107,132] and the extensive def2-QZVPP[127] basis set. Secondly, the recently developed single-point Hessian approach[161] was applied in combination with GFN2-xTB to calculate vibrational frequencies and obtain modified rigid-rotor-harmonic-oscillator (mRRHO) thermostatistical contributions at an accurate and cost-efficient level. Adding the solvation contribution calculated with the implicit COSMO-RS[162,163] solvation model results in the total free energy calculated with excellent efficiency[11] (see SI for details).

To assess the range of possible reaction pathways, a ground-state free energy landscape was surveyed.

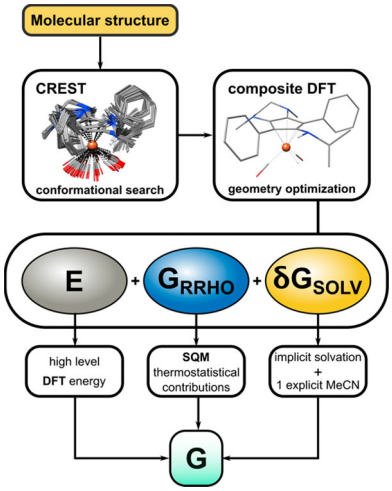

Figure E.7: Computational multilevel protocol employed to generate accurate free energies

Complex **1-Fe(NCMe)$^+$** is used as a test case to rationalize the observed electrocatalytic behavior under the assumption that the thermochemically preferred protonation sites are also kinetically facile (Figure 5, top).[36,38] The 3D graphical representation of the ground state thermochemical data plots the overall charge, number of hydrogen atoms, and ground state free energies for selected H^+/e^- addition products, with a total of 44 possible intermediates considered in this study. The SI contains a more detailed 3D graph along with free energy data for all intermediates (Figure S28-S31).[418] Inspecting structures with a +2 charge (yellow bars) indicates that protonation prior to reduction is unfavorable (+20.2 to +64.9 kcal·mol^{-1}), with the first step instead being 1e$^-$ reduction to generate the solvento adduct 1-Fe(NCMe) (+23.7 kcal·mol^{-1}) as denoted by a red arrow. By interrogating the lowest energy pathway for addition of H^+/e^- prior to releasing H_2, the catalyst likely operates by an ECEC mechanism (E = electrochemical step, C = chemical (H^+ addition) step) and is distilled into a 2D energy profile at the bottom of Figure 5. After initial reduction, solvent loss to produce 1-Fe is exergonic (+15.4 kcal·mol^{-1}; $\Delta G = -8.3$ kcal·mol^{-1}), consistent with the irreversible dissociation of acetonitrile from **1-Fe(NCMe)** when non-coordinating solvents are used (vide infra).

A structural summary of the proposed ECEC mechanism is shown in Figure E.9, where redox potentials and pK_a values are calculated from the computed free energies (Figure S28-S30). Reduction of **1-Fe(NCMe)$^+$** generates a mixture of **1-Fe** and **1-Fe(NCMe)**, the former of which is thermodynamically favorable; however both are considered plausible intermediates since these thermochemical landscape analyses do not consider reaction kinetics.[419] Next, direct protonation of the metal center is disfavored because it is endergonic by +3.3 kcal·mol^{-1} in the presence of Tos2NH (pK_a = 9.6; Figure S28). Instead, endo-CH protonation of the Cp carbon atoms containing the phenyl moieties is exergonic by 5.1 kcal·mol^{-1}, generating **1-Fe(NCMe)(endo−CpH)$^+$** (pK_a = 15.7), which is significantly more favorable than protonation at any of the amines or protonation at any other Cp ring position (Figure S29). Next, reduction and ligand-to-metal tautomerization is exergonic by 9.5 kcal·mol^{-1}, generating the neutral hydride complex **1-FeH** (pK_a = 33.4) with subsequent loss of acetonitrile (Figure E.9, bottom). This computed acidity is significantly lower than the experimentally derived acidity of $FeHCp\cdot(CO)_3$ (pK_a^{MeCN} = 29.7)[420], reflecting the markedly stronger donor character of $^{en}Cp^{NHiPr}$ vs. the Cp* ligand.[421] Importantly, protonation of the metal center to generate an

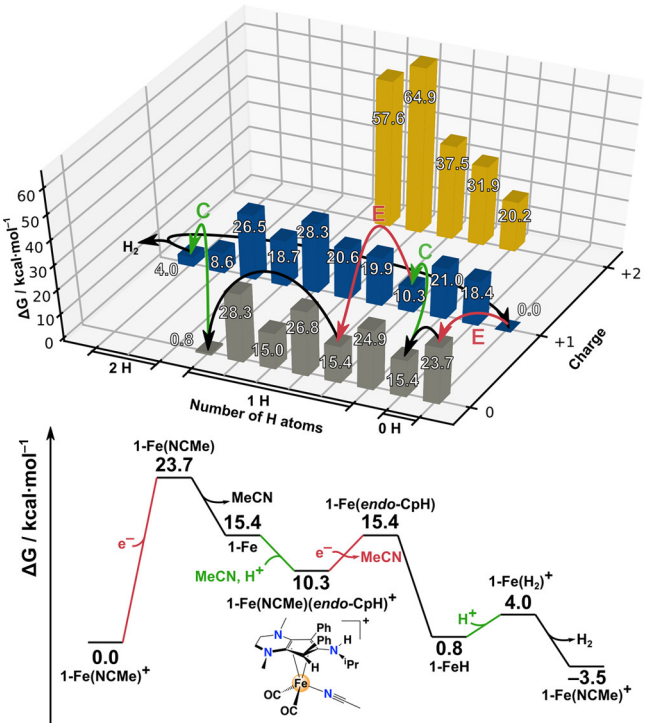

Figure E.8: Top: Partial ground state 3D free energy landscape (kcal·mol^{-1}) for electrocatalytic H_2 production with arrows representing most favorable reduction and protonation steps. Free energies are all referenced to complex **1-Fe(NCMe)$^+$** (0.0 kcal·mol^{-1}), where the acid source is incorporated by adding the electrochemical free energy term $\Delta G = -nFE$ where n = number of electrons, F = Faraday constant, and E is the standard state thermodynamic potential in acetonitrile for the reaction $2Tos_2NH + 2e^- \rightleftharpoons H_2 + 2Tos_2N^-$ ($E^\circ_{H^+/H_2} = -0.736$ V vs. $Fc^{+/0}$; i.e., zero overpotential). Bottom: 2D ground state free energy plot of ECEC mechanism based on arrows derived from the 3D plot.

Fe(IV) dihydride intermediate can also be excluded based on its high acidity ($pK_a = -0.8$). Direct protonation of the hydride produces **1-Fe(H$_2$)$^+$** ($pK_a = 9.6$), more favorable than protonation of the amine to yield **1-FeH(NH)$^+$** by three orders of magnitude ($pK_a = 6.3$). Consequently, this disfavors a classic bifunctional/heterolytic H_2 evolution pathway[422–424] on thermochemical grounds due to a pK_a mismatch between the proton source and pendant amine. Replacing H_2 for NCMe is exergonic by 4.0 kcal·mol^{-1}, regenerating **1-Fe(NCMe)$^+$** (Figure E.8).

E.3 Conclusion

In summary, we present a series of novel amine-rich monometallic iron electrocatalysts. Due to slow CO dissociation from the coordinatively saturated iron tricarbonyl precursors, one CO ligand is selectively extruded to afford electrocatalysts with H_2 production rates of up to 45 s^{-1} at room temperature. High-level computational investigation reveals that the commonly accepted paradigm of amine-assisted proton delivery and removal is likely not operative, underpinning the need to use comprehensive thermochemical landscape analyses for this mechanistically divergent electrocatalyst family.

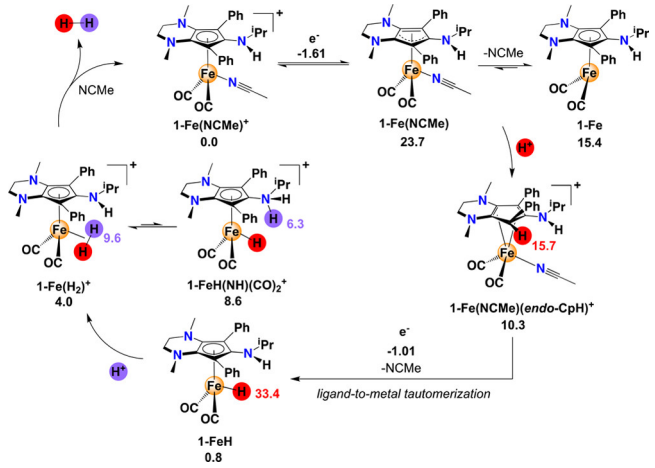

Figure E.9: Associated pK_a^{MeCN} values (red, purple) reduction potentials (black, V relative to $Fc^{+/0}$) are also shown. Ground state free energies are shown in black relative to **1-Fe(NCMe)$^+$** (0.0 kcal·mol^{-1}) at zero overpotential for the reaction $2Tos_2NH + 2e^- \rightleftharpoons H_2 + 2Tos_2N^-$ ($E^\circ_{H^+/H_2} = -0.736$ V vs $Fc^{+/0}$). See Figures S28-S30 for tabulated free energies and pK_a^{MeCN} values of all other species considered.

E.4 Supporting Information

The Supporting Information is available free of charge at:
https://pubs.acs.org/doi/10.1021/acs.inorgchem.1c03142.

- Experimental details, syntheses, NMR spectra, electrochemistry, and DFT computational workflow (PDF)
- Computed Cartesian coordinates (XYZ)

E.5 Acknowledgements

D.E.P. thanks the National Science Foundation (NSF) forsupport under Grant 2055097, and S.G. thanks the Deutsche Forschungsgemeinschaft under Grant 1927/16-1. This research was also supported by a Rutgers University–Newark startup grant (to D.E.P.). X-ray structural solutions were partly supported by the NSF under Grant 2018753. D.E.P. thanks Prof. Robert Morris, Dr. R. Morris Bullock, and Prof. Mark Lipke for constructive suggestions.

APPENDIX F

The Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes

Bhumika Goel,[†] Hagen Neugebauer,[‡] Andrew VanderWeide[†], Práxedes Sánchez,[†] Roger A. Lalancette,[†] Stefan Grimme,[‡] Andreas Hansen,[‡] Demyan E. Prokopchuk[†]

Received: 27 June 2023
Published online: 9 October 2023

Reprinted in Appendix F (adapted) with permission[¶] from
B. Goel, H. Neugebauer, A. I. VanderWeide, P. Sánchez, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes*, ACS Catal. **13** (2023) 13650, DOI: 10.1021/acscatal.3c02911

Own contributions

- performing all quantum chemical calculations
- interpretation of quantum chemical results
- writing and revising parts of the manuscript concerning DFT calculations
- none of the experiments or measurements were conducted by me

[†]Department of Chemistry, Rutgers University Newark, Newark, New Jersey 07102, United States
[‡]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms Universität Bonn, Bonn 53115, Germany
[§]These authors contributed equally.
[¶]Permission requests to reuse material from this chapter should be directed to the American Chemical Society.

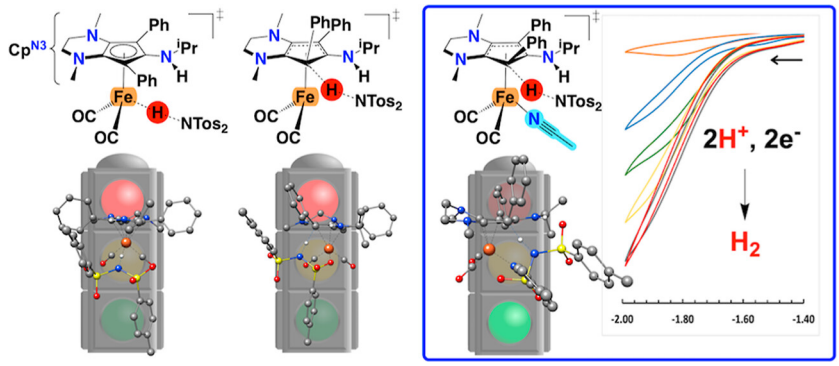

Figure F.1: Table of content graphic (ToC).

Abstract Cyclopentadienyl (Cp), a classic ancillary ligand platform, can be chemically noninnocent in electrocatalytic H−H bond formation reactions via protonation of coordinated η^5-Cp ligands to form η^4-CpH moieties. However, the kinetics of η^5-Cp ring protonation, ligand-to-metal (or metal-to-ligand) proton transfer, and the influence of solvent during H_2 production electrocatalysis remain poorly understood. We report in-depth kinetic details for electrocatalytic H_2 production with Fe complexes containing amine-functionalized $\mathbf{Cp^{N3}}$ ligands that are protonated via exogenous acid to generate via η^4-$\mathbf{Cp^{N3}}$H intermediates ($\mathbf{Cp^{N3}}$ = 6-amino-1,4-dimethyl-5,7-diphenyl-2,3,4,6-tetrahydrocyclopenta[b]pyrazin-6-yl). Under reducing conditions, state-of-the-art DFT calculations reveal that coordinated solvent plays a crucial role in mediating stereo- and regioselective proton transfer to generate **(endo-Cp^{N3}H)Fe(CO)$_2$(NCMe)**, with other protonation pathways being kinetically insurmountable. To demonstrate regioselective **endo-Cp^{N3}H** formation, the isoelectronic model complex **(endo-Cp^{N3}H)Fe(CO)$_3$** is independently prepared, and kinetic studies with the on-cycle hydride intermediate **Cp^{N3}FeH(CO)$_2$** under CO cleanly furnish the ring-activated complex **(endo-Cp^{N3}H)Fe(CO)$_3$** via metal-to-ligand proton migration. The on-cycle complex **Cp^{N3}FeH(CO)$_2$** reacts with acid to release H_2 and regenerate **[Cp^{N3}Fe(CO)$_2$(NCMe)]$^+$**, which was found to be the TOF-determining step via DFT. Collectively, these experimental and computational results underscore the emerging importance of Cp ring activation, inner-sphere solvation, and metal-ligand cooperativity to perform proton-coupled electron transfer catalysis for chemical fuel synthesis.

F.1 Introduction

There is sustained interest in developing Earth-abundant electrocatalysts for chemical energy conversion that rival the activity, robustness, and efficiency of precious metals,[16] particularly since heterogenized platinum group metals are most commonly employed in H_2 electrolysis and fuel cell systems.[17,425,426] As an alternative to heterogeneous electrocatalyst design, molecular (homogeneous) transition metal complexes enable the atomically precise tuning of primary, secondary, and outer coordination spheres, allowing for the interrogation of reaction mechanisms using solution-phase mechanistic tools and state-of-the-art computational protocols. For example, molecular electrocatalysts with proton-responsive ligand features, as found in naturally occurring hydrogenases,[378–380] have been used for H_2 production/oxidation[324,382–384,427] and CO_2 reduction[428–435] in combination with Earth-abundant metals. This concept has been extended to a wide range of ligand scaffolds for electrocatalytic small molecule activation reactions, and it has been firmly established that ligand frameworks containing

appended amines or alcohols facilitate cooperativity between the metal center and ligand, which increase reaction rates and/or offer new bond activation pathways.[427,435,436] In many cases, density functional theory (DFT) calculations have also been used to complement experimental findings and predict electrocatalyst properties by accurately computing free energy landscapes.[38,67,437] In contrast to ligand scaffolds containing proton-responsive O-H and N-H bonds, using ligand-based C-H bonds for proton-coupled electron transfer (PCET) applications with Earth-abundant metals remains scant. Only recently, C-H bonds within the classic "ancillary" η^5-cyclopentadienyl (Cp) ligand family have been utilized for PCET chemistry, where the η^5-Cp ligand is transformed into a dearomatized η^4-cyclopentadiene (η^4-CpH) moiety via nucleophilic or electrophilic activation.[438] Peters and co-workers reported that reactions of decamethylcobaltocene (Cp^*_2Co) and decamethylcobatocenium ($[Cp^*_2Co]^+$) with Brønsted acids and $NaBH_4$, respectively, produce (η^4-Cp*H) intermediates that release potent $H^\bullet$ or H^- moieties during the Fe-catalyzed reduction of N_2 to NH_3.[399,402,403] Dempsey and co-workers recently investigated the stoichiometric proton transfer kinetics with piano-stool complexes $[CpCo^{III}(diphosphine)(NCMe)]_2^+$ under reducing conditions to understand the relationship between the kinetics of Co^{III}-H formation and the strength of exogenous acid.[392] Interestingly, the rate constant became pK_a-independent with stronger acids, and in this regime the protonation rate was gated by the rate of initial MeCN dissociation from the 1e$^-$ reduction intermediate $[CpCo^{II}(diphosphine)(NCMe)]^+$. Not only was solvent dissociation a key rate-limiting factor under these conditions, but Cp ligand protonation followed by rapid metal-to-ligand proton migration via $[(\eta^4\text{-CpH})Co^{I}(diphosphine)]^+$ was kinetically preferred over direct protonation of the Co^I center to yield the cationic Co^{III}-H adduct. Blakemore and co-workers recently reported detailed kinetic studies of the elementary reaction steps for the H_2 production electrocatalyst $Cp^*Rh^I(bpy)$ ($Cp^* = C_5Me_5$; bpy = 2,2-bipyridyl) and studied the metal-ligand tautomerization behavior for $[Cp^*Rh^{III}H(bpy)]^+ \rightleftharpoons [(\eta^4\text{-Cp*H})Rh^{I}(bpy)]^+$.[439] In contrast to the aforementioned Co systems, direct metal protonation with exogenous acid is preferred ($k = 6400\ M^{-1}s^{-1}$) instead of Cp* protonation. These divergent results suggest that ligand protonation via exogenous acid is kinetically preferable with Earth-abundant metals, however the scarcity of kinetic studies in this area of research makes it difficult to draw any broader conclusions. Our foray into the chemically noninnocent behavior of Cp ligands began with the investigation of H_2 production electrocatalyis using monometallic Fe complexes containing amine-rich $\mathbf{Cp^{N3}}$ ligands.[5] Electrochemical studies revealed that a series of piano-stool iron complexes $\mathbf{Cp^{N3}Fe(NCMe)^+}$ are active for electrocatalytic H_2 production, reaching turnover frequencies of up to 45 s^{-1} (Cp^{N3} = 6-amino-1,4-dimethyl-5,7-diphenyl-2,3,4,6-tetrahydrocyclopenta[b]pyrazin-6-yl); Figure F.2). Based solely on an exhaustive free energy landscape analysis of computationally-derived pK_a and electrochemical measurements, our key mechanistic insight was that the amines were not directly involved in proton transfer and stereospecific endo-Cp ring protonation occurs after 1e$^-$ reduction, breaking the η^5 hapticity to generate an $\mathbf{(endo\text{-}Cp^{N3}H)Fe(NCMe)^+}$ intermediate. Using acid-base arguments, solvent coordination was posited at this stage, with $\mathbf{(endo\text{-}Cp^{N3}H)Fe(NCMe)^+}$ being the only plausible option (pK_a^{MeCN} = 15.7) given the acidity of the exogenous acid (Tos_2NH; pK_a^{MeCN} = 11.97)[414] during catalysis.[5] Next, ligand-to-metal proton migration was proposed to occur after a second reduction to generate the hydride complex $\mathbf{Cp^{N3}FeH}$, followed by direct protonation of $\mathbf{Cp^{N3}FeH}$, liberating H_2 and regenerating $\mathbf{Cp^{N3}Fe(NCMe)^+}$. However, no experimental evidence for Cp ring activation was presented, and kinetic analyses were not undertaken to further support this unexpected reaction pathway.

In this article, we expand significantly on our initial mechanistic understanding of this reaction by using a suite of spectroscopic and computational techniques that rationalize the observed kinetics data,

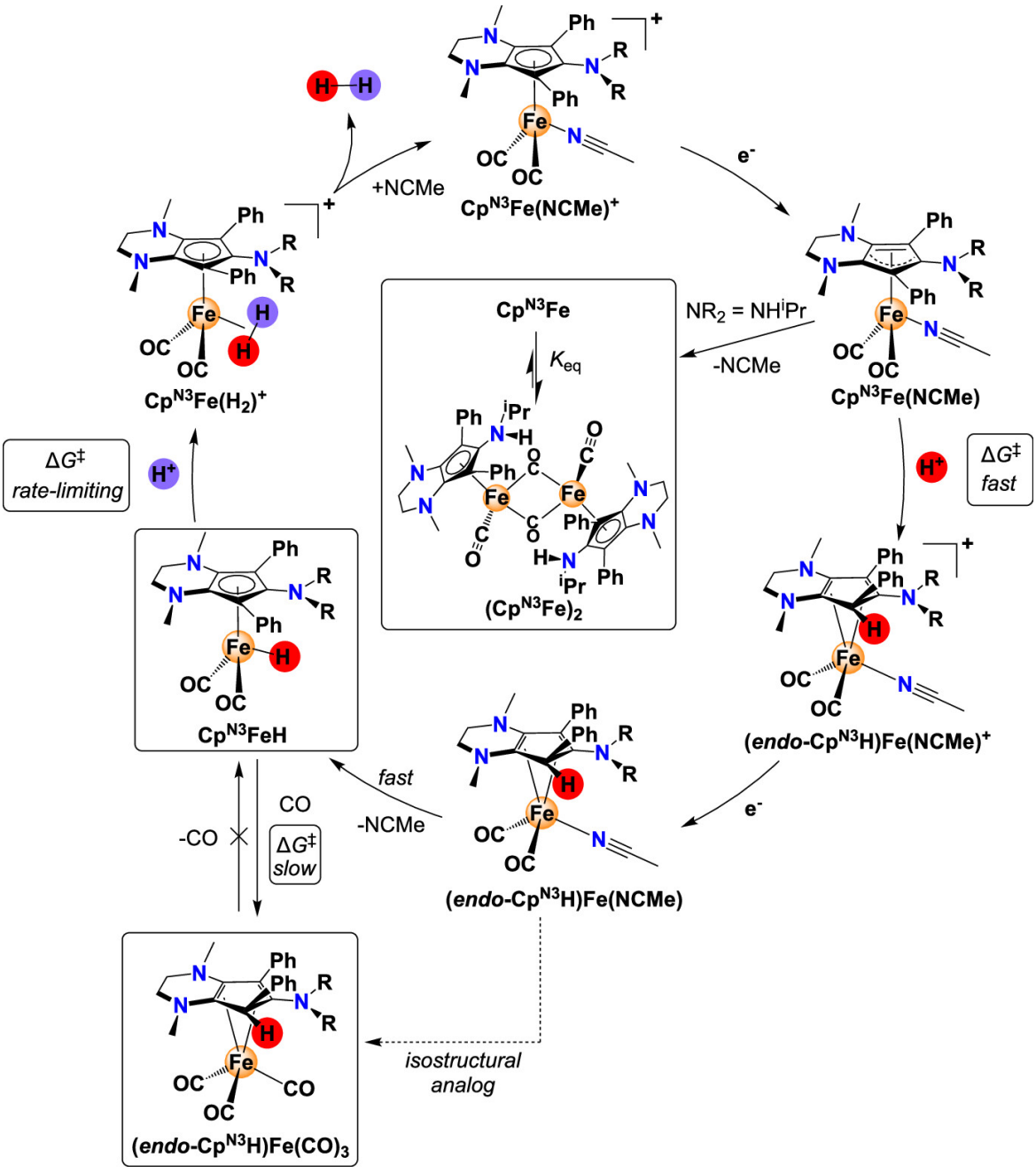

Figure F.2: Proposed electrocatalytic H_2 production mechanism including newly obtained mechanistic data (solid boxes) through experimental and computational analysis (NR_2 = NH^iPr, Pyrrolidynyl).

provide strong support for Cp ring activation, and highlight the crucial role of solvent coordination during electrocatalysis (Figure F.2, solid boxes). We show that using a slightly stronger acid in acetonitrile boosts electrocatalytic H_2 production by fivefold (up to 266 s^{-1}), however switching the reaction medium to isobutyronitrile (^{i}PrCN) or *N*, *N*-dimethylformamide (DMF) suppresses catalysis, indicating that solvent (acetonitrile) association in the primary coordination sphere is essential to facilitate catalytic turnovers. The hydride intermediate **Cp^{N3}FeH** is independently prepared and structurally validated, and DFT calculations indicate that direct protonation of **Cp^{N3}FeH** with exogenous acid is rate-limiting during electrocatalysis. State-of-the-art computational workflows demonstrate that stereo- and regioselective ligand protonation is kinetically facile to generate **(endo-Cp^{N3}H)Fe(NCMe)** during catalysis, followed by solvent dissociation and highly exothermic ligand-to-metal proton transfer to generate **Cp^{N3}FeH**. Computed pathways involving direct metal protonation or ligand protonation in the absence of coordinated solvent are kinetically insurmountable. To indirectly support the ligand-based protonation behavior during catalysis, the isoelectronic analogues **(endo-Cp^{N3}H)Fe(CO)$_3$** are synthesized (NR_2 = NH^{i}Pr, Pyrr), both of which are structurally authenticated via single crystal X-ray diffraction. DFT-based kinetic data show that substitution of CO for NCMe dramatically decreases the rate of ligand-to-metal proton transfer, going from $\Delta G^{\ddagger} \approx 5$ kcal·mol^{-1} (NCMe) to $\Delta G^{\ddagger} \approx 34$ kcal·mol^{-1} (CO), providing a clear rationale for the rapid ligand-to-metal proton migration behavior during catalysis. Experimental kinetic data show that in the reverse reaction to produce **(endo-Cp^{N3}H)Fe(CO)$_3$** from **Cp^{N3}FeH** is quantitative in the presence of CO, supporting the endo-selective proton migration behavior. Infrared spectroelectrochemistry (IR-SEC) and stoichiometric reduction reactions reveal that an off-cycle CO-bridged dimer **(Cp^{N3}Fe)$_2$** is formed in the absence of exogenous acid, which is characterized via single crystal X-ray diffraction and solution-phase equilibrium measurements. Collectively, these results underscore the novelty and importance of Cp ring activation, inner-sphere solvation, and metal-ligand cooperativity to perform PCET catalysis for chemical fuel synthesis.

F.2 Results and Discussion

F.2.1 Influence of Acid and Solvent on Electrocatalysis

Our initial report[5] surveyed H_2 production electrocatalysis using Tos_2NH (pK_a^{MeCN} = 11.97)[414] as the exogenous acid since many other commonly used exogenous acids produced significant amounts of background current at potentials beyond −1.6 V.[440] We found that the slightly stronger acid dibenzenesulfonamide ($(PhSO_2)_2NH$; pK_a^{MeCN} = 11.35)[414] is also compatible in the applied potential window (Figure S16). Thus, complexes **1-Cp^{N3}Fe(NCMe)$^+$** (NR_2 = NH^{i}Pr) and **2-Cp^{N3}Fe(NCMe)$^+$** (NR_2 = Pyrrolidinyl) were tested with $(PhSO_2)_2NH$ for H_2 production electrocatalysis using cyclic voltammetry (CV) at ≥ 800 mV/s to observe current responses that are independent of scan rate.[5] When **1-Cp^{N3}Fe(NCMe)$^+$** is treated with increasing concentrations of $(PhSO_2)_2NH$, the plateau current increases when $[(PhSO_2)_2NH] \leq 100$ mM (Figure F.3, inset). In the kinetic regime where current response is independent of $[(PhSO_2)_2NH]$, the observed rate constant k_{obs} = 118 s^{-1} for H_2 production at an estimated overpotential of 1.04 V (Table F.1).[441] Next, **2-Cp^{N3}Fe(NCMe)$^+$** was tested for catalysis with $(PhSO_2)_2NH$, resulting in k_{obs} = 266 s^{-1} for H_2 production at an estimated overpotential of 1.12 V (Figure F.3 and Table F.1). Therefore, an increase is acidity by 0.6 pK_a units results in a fourfold increase in activity for **1-Cp^{N3}Fe(NCMe)$^+$** and eightfold increase for **2-Cp^{N3}Fe(NCMe)$^+$**, respectively. Controlled potential electrolysis (CPE) with exogenous acid and

electrocatalyst over the course of 30 minutes shows sustained current output and a persistent yellow color, however Faradaic efficiencies (FE) below unity are found in both cases (Table F.1). Based on control experiments in the absence of Fe complex, we attribute this to slow decomposition of the exogenous acid during CPE (see the SI). Although rinse tests alone do not conclusively exclude the formation of heterogeneously adsorbed metal catalyst,[442,443] performing a rinse test using a freshly prepared 20 mM $(PhSO_2)_2NH$ solution reveals a similar current response and FE in comparison to control CPE experiments with a freshly polished electrode dipped into solutions containing 20 mM $(PhSO_2)_2NH$ in the absence of electrocatalyst (Figure S25 and Table S1).

To probe the effect of isotopic labeling on the kinetics of H_2 production, deuterium substituted 4-methyl-N-tosylbenzenesulfonamide (Tos_2ND) was synthesized (Figure S12) and used as the exogenous acid with **1-Cp^{N3}Fe(NCMe)$^+$**. Successive additions of acid showed significantly lower current enhancements with k_{obs} = 22 s^{-1} when [Tos_2ND] = 60 mM (Figure S17). Compared to the activity with Tos_2NH, k_H/k_D = 1.33, with the normal KIE suggesting that H^+ transfer is involved in the rate determining step.[444] Using DFT (see the SI for details), the computed free energy barriers of direct protonation of **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** are consistent with this finding (Table F.1) and in good agreement with the observed rates ($\Delta\Delta G^{\ddagger}$ = 0.7 to 1.9 kcal·mol^{-1}; Figure F.4). DFT rate constants were obtained from conventional transition-state theory with the r^{2}SCAN0-D4/def2-QZVPP level of theory employing the Eyring equation under the assumption that the transmission coefficient equals one. Although the computed rate constants are lower than the experimentally observed values, the general trend has been reproduced, showing that the stronger acid $(PhSO_2)_2NH$ lowers the free energy barrier for protonation. Considering that small errors for reaction barriers of 1 kcal·mol^{-1} lead to a factor around 10 in the rate constants[45] and typical DFT errors for transition metal barrier heights are around 2 kcal·mol^{-1} with the best performing DFT methods,[135,136] the rate constants from the experiment and DFT in Table F.1 are in reasonable agreement.[2]

Table F.1: Electrocatalysis data with **1-Cp^{N3}Fe(NCMe)$^+$** and **2-Cp^{N3}Fe(NCMe)$^+$**.

Catalyst	Acid / pK_a^{MeCN}	k^a_{obs} (k^b_{DFT}) / s^{-1}	Overpotential (η) / V[a]	FE / %
1-Cp^{N3}Fe(NCMe)$^+$	Tos_2NH (11.97)	29[c] (9)	0.98[c]	65 ±3[c]
2-Cp^{N3}Fe(NCMe)$^+$		35[c] (4)	1.02[c]	65 ± 3[c]
1-Cp^{N3}Fe(NCMe)$^+$	$(PhSO_2)_2NH$ (11.35)[a]	118 (34)	1.04	77 ± 2
2-Cp^{N3}Fe(NCMe)$^+$		266 (12)	1.12	77 ± 3

[a] Calculated using standard protocols for catalytic plateau current and overpotential analysis.[418,445]
[b] Computed values from $\Delta G^{\ddagger}$ for the protonation of **Cp^{N3}FeH** to form **[Cp^{N3}Fe(H$_2$)]$^+$**.
[c] taken from [5]

Next, we explored the effect of solvent on current response in the presence and absence of exogenous acid by dissolving **1-Cp^{N3}Fe(NCMe)$^+$** in *N*, *N*-dimethylformamide (DMF) or isobutyronitrile (^{i}PrCN). We presume that the coordinated NCMe ligand of **1-Cp^{N3}Fe(NCMe)$^+$** is immediately replaced by DMF or ^{i}PrCN in solution. In the absence of acid, CVs of **1-Cp^{N3}Fe(NCMe)$^+$** in DMF and ^{i}PrCN show irreversible reduction waves at E_{pc} = −1.72 V and −1.83 V, respectively (Figure S18), suggesting that solvent dissociation from the metal's primary coordination sphere is rapid since an anodic wave appears in MeCN. Compared to the CV of **1-Cp^{N3}Fe(NCMe)$^+$** in MeCN (E_{pc} = −1.72 V), the cathodic peak potential in DMF is nearly identical while the potential in ^{i}PrCN is 110 mV more negative due to its slightly increased donicity (basicity). Next, CVs were conducted

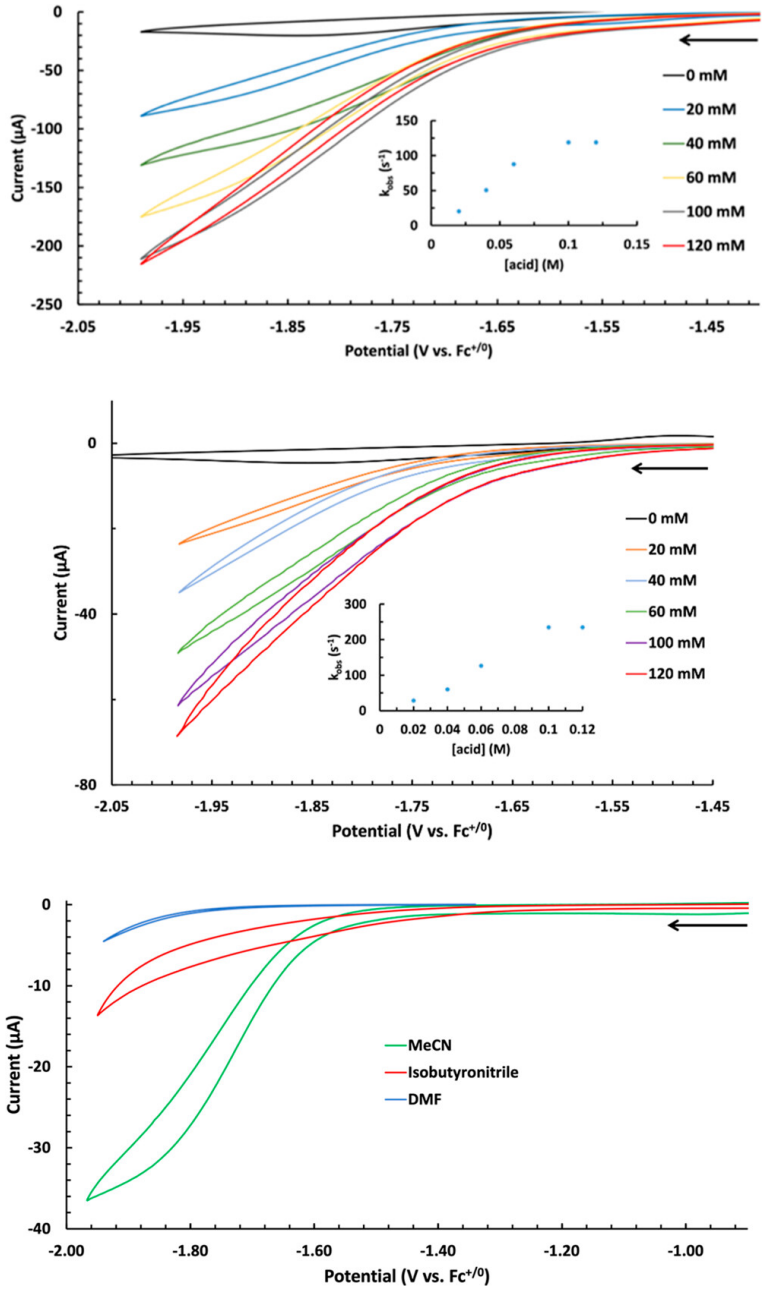

Figure F.3: IUPAC plotted, iR uncompensated CV traces of **1-Cp^{N3}Fe(NCMe)$^{+}$** (top) and **2-Cp^{N3}Fe(NCMe)$^{+}$** (middle) with increasing concentrations of $(PhSO_2)_2NH$ at 800 mV/s with inset plots showing dependence of k_{obs} on acid concentration. Bottom: comparison of electrocatalytic current responses with 1 mM **1-Cp^{N3}Fe(NCMe)$^{+}$** dissolved in MeCN, *N*, *N*-dimethylformamide (DMF), or isobutyronitrile in the presence of 20 mM Tos_2NH at 800 mV/s. Conditions: Ar, MeCN solvent, 0.1 M $[Bu_4N][PF_6]$, 1.0 mM analyte, PEEK-encased glassy carbon working electrode, Type 2 glassy carbon rod counter electrode, Ag/AgCl pseudoreference electrode in a frit-separated (Coralpor®) glass compartment containing solvent and electrolyte. Initial scan direction and starting position indicated with a black arrow.

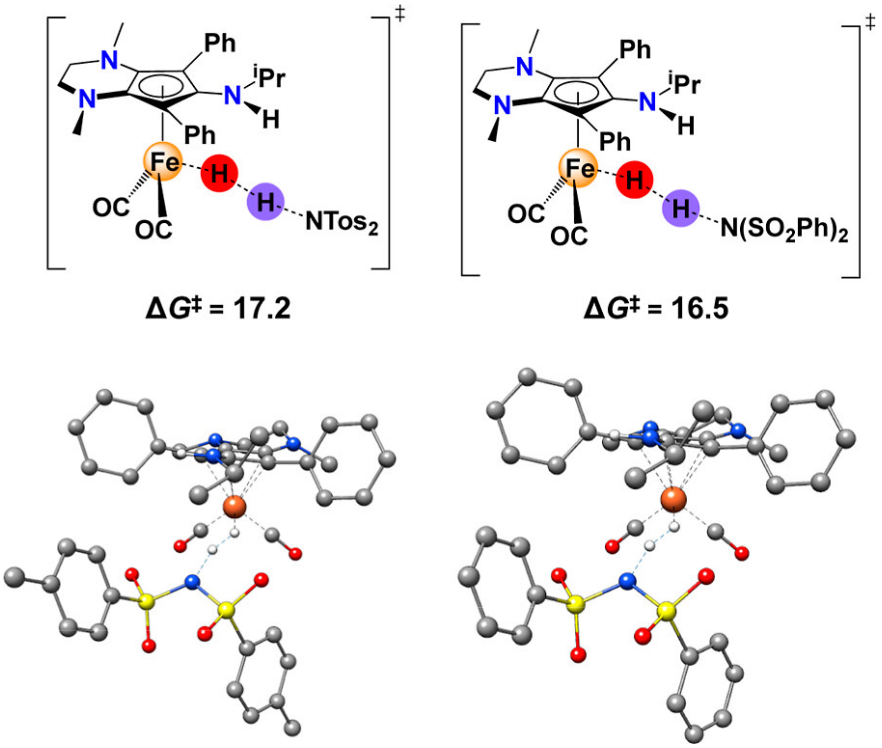

Figure F.4: Computed transition states for the direct protonation of **1-Cp^{N3}FeH** to generate **1-Cp^{N3}Fe(H$_2$)$^+$** via Tos_2NH (left) and $(PhSO_2)_2NH$ (right). Energies are given in kcal·mol^{-1}.

with **1-Cp^{N3}Fe(NCMe)$^+$** in the presence of 20 mM Tos_2NH and almost no current enhancement is observed in DMF while a modest increase is seen using ^{i}PrCN (Figure F.3 bottom and Figure S19). This strongly indicates that the presence of MeCN in the primary coordination sphere of Fe is crucial for rapid electrocatalysis. Using DFT, we investigated the protonation free energy barriers under three different scenarios: (1) protonation at the endo-CpH position of **1-Cp^{N3}Fe(NCMe)** with $\Delta G^{\ddagger}$ = 9.8 kcal·mol^{-1} (**TS1**); (2) dissociation of MeCN and protonation at the endo-CpH position of **1-Cp^{N3}Fe** with $\Delta G^{\ddagger}$ = 30.0 kcal·mol^{-1} (**TS3**); (3) dissociation of MeCN and protonation at the metal center of **1-Cp^{N3}Fe**, formally generating an Fe^{III} hydride with $\Delta G^{\ddagger}$ = 27.0 kcal·mol^{-1} (**TS4**; Figure F.5). Clearly, **TS3** and **TS4** are kinetically inaccessible, which is consistent with the experimentally observed solvent dependence on rapid electrocatalysis as described above.

Despite many attempts, DFT modeling of similar ligand protonation transition states for complexes with the general formula **1-Cp^{N3}Fe(L)** (L = DMF, ^{i}PrCN) were unsuccessful, possibly due to the increased steric congestion which makes the carbon atom inaccessible on the **Cp^{N3}** ring. A low barrier pathway involving amine protonation at **1-Cp^{N3}Fe** was also found, however the proton transfer barrier to form **1-Cp^{N3}FeH** is insurmountable (**TS6**, $\Delta G^{\ddagger}$ = 23.6 kcal·mol^{-1}; Figure S29).

F.2.2 Synthesis and Reactivity of Cp^{N3}FeH complexes

To demonstrate the release of H_2 in the final stages of the cycle shown in Figure F.2, we successfully synthesized iron hydride complexes by taking **1-Cp^{N3}Fe(NCMe)$^+$** or **2-Cp^{N3}Fe(NCMe)$^+$** and stirring

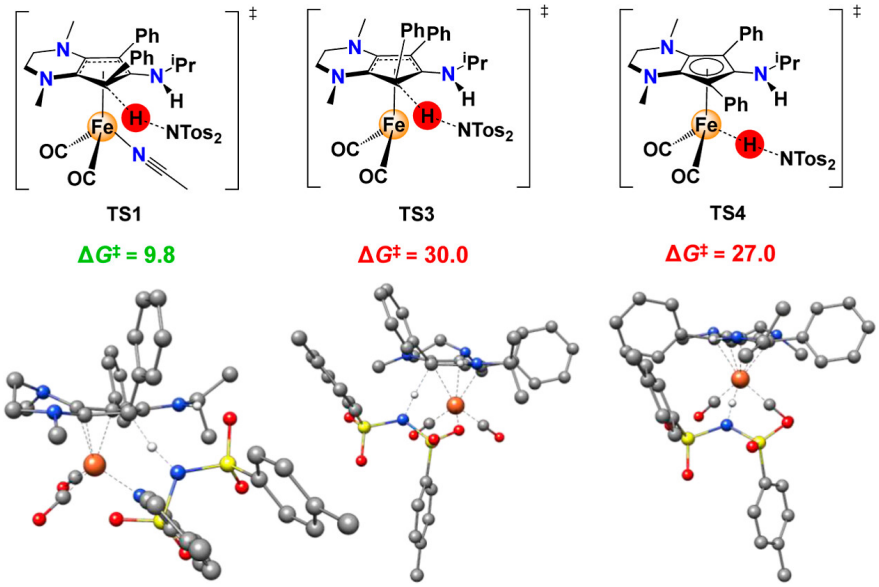

Figure F.5: Free energies for three different protonation scenarios starting with **1-Cp^{N3}Fe(NCMe)**. Energies are given in kcal·mol^{-1}.

with $LiAlH_4$ at room temperature, yielding **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** in good yield (Figure F.6).

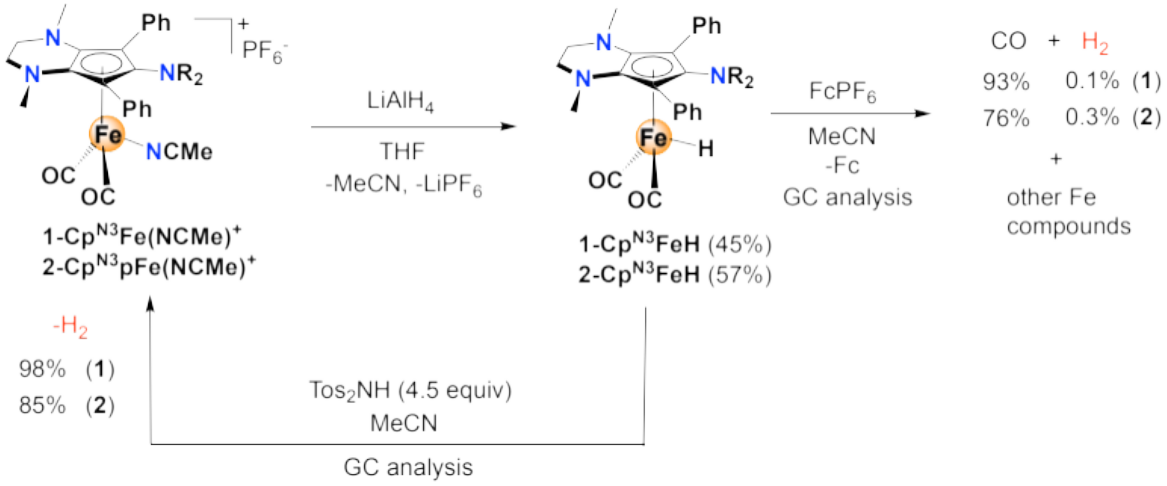

Figure F.6: Synthesis and reactivity of **Cp^{N3}FeH**.

The symmetric and asymmetric CO stretches for the carbonyl ligands appear at 1968 and 1907 cm^{-1} for **1-Cp^{N3}FeH** and 1968 and 1900 cm^{-1} for **2-Cp^{N3}FeH**. The ^{1}H NMR spectra show characteristic singlets for the **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** at −10.75 and −10.72 ppm, respectively, while ^{13}C NMR spectra show singlets at 218.47 and 218.11 ppm. Single crystals suitable for X-ray diffraction reveal the molecular structures of **1-Cp^{N3}FeH** (Figure F.7A) and **2-Cp^{N3}FeH** (Figure F.7B). For **1-Cp^{N3}FeH**, the hydride ligand is freely refined for one of the molecules in the asymmetric unit (Fe1-H27 = 1.47(2) Å), however the hydride could not be located for the second Fe complex due to

two-site positional disorder of the hydride and CO ligand in the asymmetric unit. The measured Fe-H bond distance is comparable to other piano-stool iron hydride complexes that have been characterized by X-ray diffraction.[446] In the Fourier difference map of **2-Cp^{N3}FeH**, the hydride could not be located (Figure F.7B).

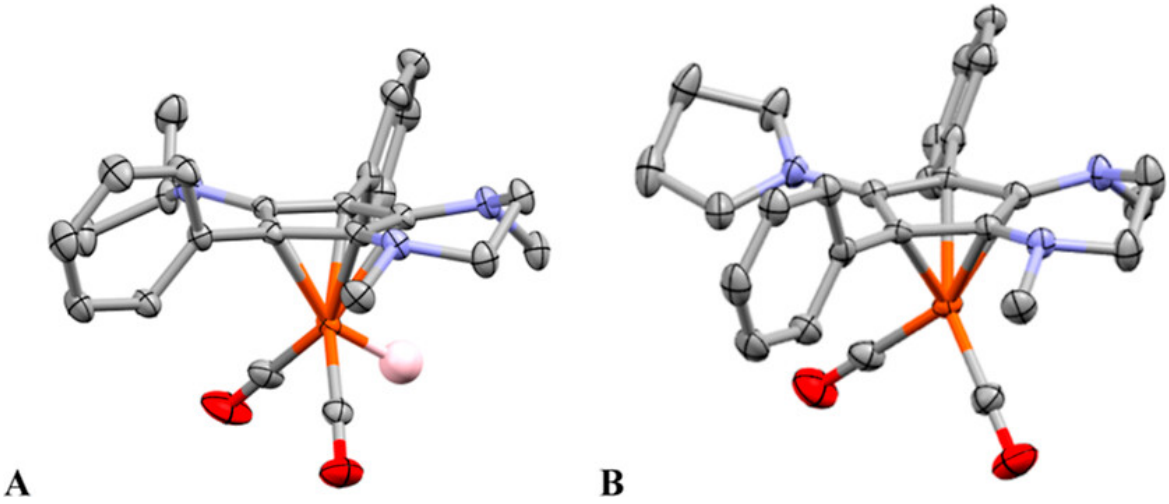

Figure F.7: X-ray crystallographic structures of **1-Cp^{N3}FeH** (A) and **2-Cp^{N3}FeH** (B) with 50% probability ellipsoids. Hydrogens removed for clarity except for the hydride ligand of **1-Cp^{N3}FeH**. For **1-Cp^{N3}FeH**, the hydride could not be located for the second Fe complex in the asymmetric unit due to two-site positional disorder of the hydride and CO ligand. For **2-Cp^{N3}FeH**, the hydride could not be located in the Fourier difference map.

The reactivity of **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** were interrogated. The catalytic cycle in Figure F.2 indicates that the H_2 production loop is closed by direct protonation of **1-Cp^{N3}FeH** by Tos_2NH to generate a transient dihydrogen complex followed by H_2 release and MeCN binding to Fe. Thus, complex **1-Cp^{N3}FeH** or **2-Cp^{N3}FeH** was reacted with 4.5 equiv Tos_2NH in a septum-sealed NMR tube under N_2 and the headspace was analyzed by gas chromatography (Figure F.6). The chromatographs reveal that **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** smoothly release H_2 in 98% and 85% yield, respectively, with no detectable formation of gaseous CO.

Next, the reactivity of **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** were evaluated under oxidizing conditions. The ligand-to-metal proton migration to form **Cp^{N3}FeH** in Figure F.2 is proposed to occur via **1-(endo-Cp^{N3}H)Fe(NCMe)**, however, there are no experimental data ruling out ligand-to-metal proton migration from the putative radical cation **1-(endo-Cp^{N3}H)Fe(NCMe)$^+$**, which would generate the intermediate **1-Cp^{N3}FeH$^+$**.[5] Thus, the anodic electrochemistry of **1-Cp^{N3}FeH** in acetonitrile was investigated. An irreversible oxidation peak is observed (E_{pc} = 0.0 V), consistent with theory (E° = 0.04 V),[5] however the peak current diminishes significantly after each CV sweep which suggests electrode fouling (Figure S20). Irreversible CVs were conducted at 10 and 20 V/s with repolishing in between each run, however no return oxidation is observed, indicating rapid degradation of electrogenerated **1-Cp^{N3}FeH$^+$** (Figure S21). Using a weakly coordinating solvent such as fluorobenzene engenders partially reversible redox waves with **1-Cp^{N3}FeH** at scan rates greater than 5 V/s, suggesting that MeCN coordination may play a role during product decomposition (Figure S22). Finally, the oxidative degradation of **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** was further supported by treating each complex with 5 equiv ferrocenium hexafluorophosphate ($FcPF_6$) in acetonitrile under an inert

atmosphere in a septum-sealed NMR tube. GC headspace analysis reveals CO gas evolution in 93% and 76% yields for **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH**, respectively, with only trace amounts of H_2 (0.1-0.3%) relative to initial [Fe]; NMR analysis of the solution phase reaction mixtures show intractable mixtures of products while IR analysis reveals several low intensity stretches in the carbonyl region. Although many stable transition metal hydride radical cations are known,[447] we posit that decomposition of **1-Cp^{N3}FeH$^+$** is facile and this species does not play a role in H_2 production electrocatalysis.

F.2.3 Synthesis and Proton Migration Kinetics of (endo-Cp^{N3}H)Fe(CO)$_3$ Complexes

Thus far, our evidence indicates that the formation **1-Cp^{N3}FeH** likely occurs by intramolecular ligand-to-metal proton migration via **(endo-Cp^{N3}H)Fe(NCMe)**, formally an 18e$^-$ Fe^0 complex. This endo stereoisomer is energetically preferred over the exo-Cp^{N3}H adduct by 9.6 kcal·mol^{-1},[5] presumably due to the decreased steric congestion by having the phenyl ring pointing away from the metal center and CO ligands. Although we have been unable to isolate **(endo-Cp^{N3}H)Fe(NCMe)**, exposure of **1-Cp^{N3}Fe(CO)$_3^+$** or **2-Cp^{N3}Fe(CO)$_3^+$** to $NaHBEt_3$ under 2.5 atm of CO in cold THF results in regio- and stereoselective endo-Cp ring activation to generate complexes **1-(endo-Cp^{N3}H)Fe** and **2-(endo-Cp^{N3}H)Fe** (Figure F.8, top), which are isoelectronic variants of the key catalytic intermediate **(endo-Cp^{N3}H)Fe(NCMe)**. These complexes are very stable under an inert atmosphere, but higher pressures of $CO_{(g)}$ are necessary to minimize CO dissociation and subsequent formation of **1-Cp^{N3}FeH** or **2-Cp^{N3}FeH**. Diagnostic ^{1}H NMR spectroscopic data for **1-(endo-Cp^{N3}H)Fe** and **2-(endo-Cp^{N3}H)Fe** include singlets at 4.84 and 4.77 ppm, respectively, corresponding to the **endo-Cp^{N3}H** proton. The presence of three CO stretches from 1978-1886 cm^{-1} in the IR spectra are also indicative of an increase of electron density at the metal center, as the starting complexes exhibit CO stretches from 1999-1940 cm^{-1}.[5] Crystallization from ether/pentane solutions provides X-ray quality crystals, validating the molecular structures of **1-(endo-Cp^{N3}H)Fe** (Figure F.8A) and **2-(endo-Cp^{N3}H)Fe** (Figure F.8B). Structural analyses indicate there is a high degree of electronic delocalization throughout the unsaturated carbon atoms of the η^4-cyclopentadiene moiety due to the formal 2e$^-$ reduction of the metal center and greater degree of π backbonding into the butadiene motif.[448] For example, in **1-Fe(endo-Cp^{N3}H)**, all three C-C bonds in the diene portion of the ring are similar in length (C6-C7 = 1.428(2) Å; C7-C8 = 1.442(2) Å; C8-C4 = 1.445(2) Å) while the C-C bonds connected to the tetrahedral C5 are substantially elongated (C4-C5 = 1.530(2) Å; C6-C5 = 1.512(2) Å). To the best of our knowledge, nucleophilic activation at an η^5-Cp ring to generate an isolable (η^4-CpH)Fe^0 complex is rare, with only three other examples reported in the literature.[396,397,449]

Since experimental data for the kinetics of ligand-to-metal proton migration via η^4-CpH intermediates remains scant,[392] we attempted to convert **2-(endo-Cp^{N3}H)Fe** into **2-Cp^{N3}FeH** to emulate the proton migration behavior during electrocatalysis via **(endo-Cp^{N3}H)Fe(NCMe)**. Complex **2-(endo-Cp^{N3}H)Fe** is very stable under inert gas at room temperature, therefore the compound was heated in toluene-d_8 for prolonged periods of time and monitored by ^{1}H NMR. However, even after 16.5 hours at 105 °C, only a trace amount of **2-Cp^{N3}FeH** was observed (< 1%) along with some signal broadening, suggestive of gradual decomposition (Figure S26). Since the hydride complexes **1-Cp^{N3}FeH** and **2-Cp^{N3}FeH** are also formed as minor products during synthesis of the endo-CpH using $NaHBEt_3$, we hypothesized that the Lewis acidic BEt_3 might facilitate proton migration; however, addition of BEt_3 to a pure sample of **2-(endo-Cp^{N3}H)Fe** in CD_3CN elicited no reaction after 24 h at room temperature (Figure S27). To obtain a clearer understanding of why **2-(endo-Cp^{N3}H)Fe** is so unreactive, relaxed potential energy surface scans via DFT were performed, providing the electronic

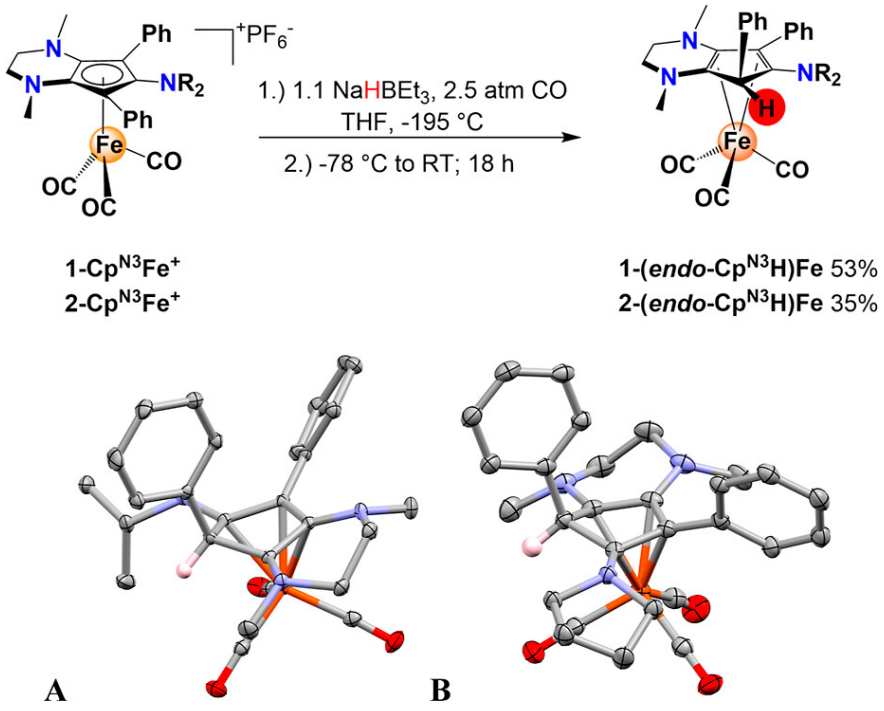

Figure F.8: Top: Synthesis of ring activated compounds **1-(endo-Cp^{N3}H)Fe** and 2-(endo-Cp^{N3}H)Fe. Bottom: X-Ray crystallographic structures of **1-(endo-Cp^{N3}H)Fe** (A) and **2-(endo-Cp^{N3}H)Fe** (B) with 50% probability ellipsoids (most hydrogens omitted for clarity).

energies (ΔE, kcal·mol^{-1}) as a function of Fe-CO distance (Δd, Å; Figure F.9). For selected points on this scan, approximate thermostatistical corrections were added because no clean transition state could be obtained on going from **2-Fe(endo-Cp^{N3}H)** to **2-Cp^{N3}FeH** (see the SI for details). Thus, all free energies discussed in this section refer to these corrected free energies. Calculations indicate that the barrier to removing CO, a strong π-acceptor, is substantial at room temperature ($\Delta G^{\ddagger}$ = 34.2 kcal·mol^{-1}). In contrast, substituting CO with NCMe dramatically lowers the free energy barrier for ligand dissociation in **(endo-Cp^{N3}H)Fe(NCMe)**, making the reaction very rapid under standard-state conditions ($\Delta G^{\ddagger}$ = 5.3 kcal·mol^{-1}) and is therefore consistent with the observed kinetics of H_2 production electrocatalysis. Again, these data underscore the importance of acetonitrile coordination, which facilitates facile proton migration from the Cp ligand to iron. A very similar reactivity profile was also modeled for **1-Fe(endo-Cp^{N3}H)** (Figure S30).

The large energy difference between **2-(endo-Cp^{N3}H)Fe** and **2-Cp^{N3}FeH** in Figure F.9 (blue trace) suggests that the reverse reaction (i.e., metal-to-ligand proton migration) should be very exothermic under relatively mild conditions. Thus, we prepared an 18 mM solution of **2-Cp^{N3}FeH** in 0.50 mL CD_3CN and pressurized an NMR tube with 6.5 atm $CO_{(g)}$, ensuring there was an excess (about 5 equiv) of dissolved CO relative to **2-Cp^{N3}FeH**, enabling us to model the kinetics under pseudo first-order conditions (see the SI for details). Room temperature monitoring via ^{1}H NMR revealed that the reaction was sluggish, with only ca. 1% conversion after 16 hours. Thus, the reaction was conducted in the NMR probe at 45 °C and ^{1}H NMR spectra were automatically collected every 20 minutes over the course of 8 hours. Very clean conversion from **2-Cp^{N3}FeH** to **2-(endo-Cp^{N3}H)Fe** was observed

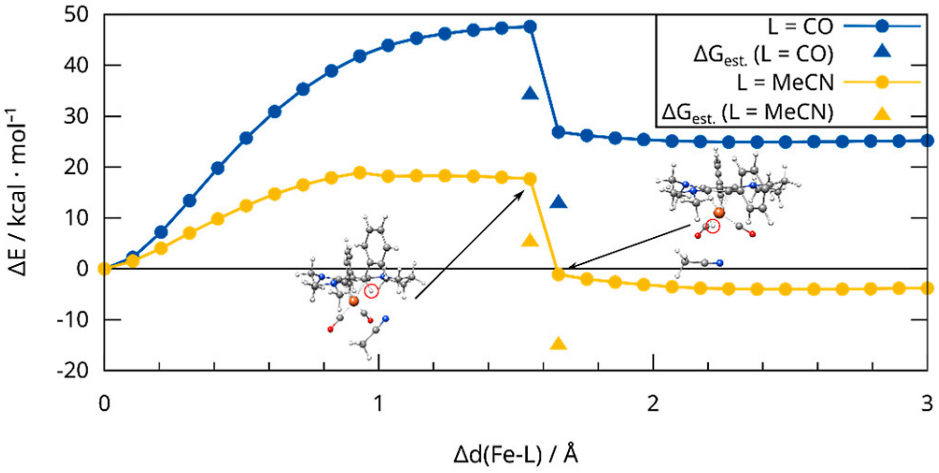

Figure F.9: Relaxed surface scan performed in ORCA[141] on the PBE0-D3(BJ)/def2-SVP (def2-TZVP for Fe) + CPCM(MeCN) level of theory[111,112,156,450] going from **2-(endo-Cp^{N3}H)Fe** to **2-Cp^{N3}FeH** + CO (blue trace) and **2-(endo-Cp^{N3}H)Fe(NCMe)** to **2-Cp^{N3}FeH** + NCMe (yellow trace). The (free) energies are given relative to Δd(Fe-L) = 0. Thermostatistical corrections were added for selected nodes of the scan (blue and yellow triangles).

by monitoring the loss of the hydride signal at −10.72 ppm and growth of the CpH resonance at 4.84 ppm (Figure S28). Thus, the first-order rate constant for the disappearance of **2-Cp^{N3}FeH** (1.1×10^{-4} s^{-1}) and appearance of **2-(endo-Cp^{N3}H)Fe** (8.8×10^{-5} s^{-1}) yields an average rate constant k_{obs} = 9.9 ± 1.5 x 10^{-4} s^{-1} for metal-to-ligand proton migration (Figure F.10), which is in reasonable agreement with the computed barrier ($\Delta G^{\ddagger}$ = 21.2 kcal·mol^{-1}; $k_{DFT} = 1.1 \times 10^{-2}$ s^{-1}) considering the approximate nature of the relaxed surface scan. Therefore, it appears that the kinetic barrier impeding ligand-to-metal proton migration in **2-(endo-Cp^{N3}H)Fe** is almost solely due to CO ligand dissociation, implying that a once vacant site is available at Fe the proton transfer reaction is practically barrierless.

F.2.4 IR Spectroscopy and an Off-cycle Fe-Fe Dimer

We sought to study the speciation of precatalyst mixtures via infrared sprectroelectrochemistry (IR-SEC) using an optically transparent thin layer electrochemical (OTTLE-type) cell[445,451] in the absence of exogenous acid to identify any catalytically relevant intermediates prior to the first protonation step. Initial IR-SEC experiments on the dicarbonyl adduct **1-Cp^{N3}Fe(NCMe)$^{+}$** under reducing conditions in an MeCN solution containing 0.25 M [nBu$_4$N][PF$_6$] resulted in no discernable CO stretches, suggesting rapid decomposition in the absence of acid. Next, we turned our attention to IR-SEC studies on the previously reported tricarbonyl cation **1-Cp^{N3}Fe(CO)$_3^{+}$** which has been shown to lose one equiv CO during cathodic CV scans at an onset potential of ca. −1.45 V vs Fc$^{+/0}$.[5] Prior to applying a cathodic potential, CO stretches at 2058 and 1992 cm^{-1} are observed in an MeCN solution containing 0.25 M [nBu$_4$N][PF$_6$] (Figure F.11). After applying a potential of −1.5 V, the **1-Cp^{N3}Fe(CO)$_3^{+}$** bands decrease in intensity and four new stretches appear at 1713, 1901, 1910 (sh) and 1969 cm^{-1}. Comparison of these new peaks with computed IR spectral data indicate that **1-Cp^{N3}Fe(NCMe)** and **1-Cp^{N3}Fe** are absent since the regions at 1829 and 1888 cm^{-1} are silent (Table F.2). However, the CO stretching peaks for the independently prepared iron hydride complex

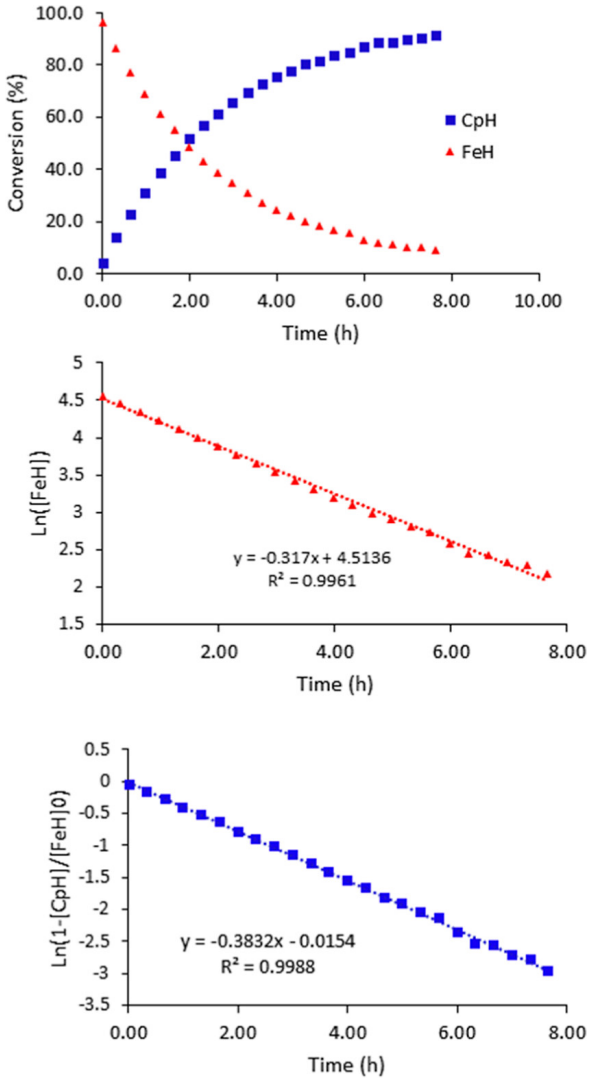

Figure F.10: Kinetic traces for the conversion of **2-Cp^{N3}FeH** to **2-(endo-Cp^{N3}H)Fe** in CD_3CN at 45 °C. Top: Percent conversion for the appearance of **2-(endo-Cp^{N3}H)Fe** (blue squares) and disappearance of **2-Cp^{N3}FeH** (red triangles). Middle: First-order fit for the disappearance of **2-Cp^{N3}FeH**. Bottom: First-order fit for the appearance of 2-(endo-Cp^{N3}H)Fe.

1-Cp^{N3}FeH (1901,1969 cm^{-1}; KBr), coincide with the CO stretches at 1907 and 1968 cm^{-1} produced via IR-SEC. Even though our reaction conditions rigorously exclude moisture and exogenous acid, these data indicate that the transient reduction intermediates **1-Cp^{N3}Fe(NCMe)** and/or **1-Cp^{N3}Fe** are highly reactive with trace amounts of H^+ in solution to generate **1-Cp^{N3}FeH**. Interestingly, the major product contains a strong CO band at 1713 cm^{-1} which is a characteristic stretching region for CO-bridged (μ_2-CO) metal complexes. Based on a reasonable match between experiment (1731, 1928 cm^{-1}) and computation (1741, 1908 cm^{-1}), theory predicts that this new product is the CO-bridged dimer **1-(Cp^{N3}Fe)$_2$** which contains two bridging and two terminal CO ligands, reminiscent of the well-known CO-bridged dimers [Fe(CO)$_2$Cp]$_2$ (Fp$_2$) and [Fe(CO)$_2$Cp$\cdot$]$_2$ (Fp*$_2$).[452,453]

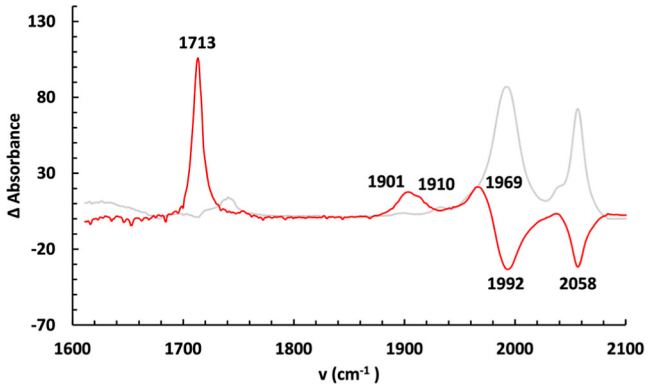

Figure F.11: Difference FTIR spectra of **1-Cp^{N3}Fe(CO)$_3^+$** before (gray) and after an applied potential (red) . Conditions: 1 mM **1-Cp^{N3}Fe(CO)$_3^+$** ; 0.25 M [Bu$_4$][PF$_6$], MeCN solvent, set potential = −1.5 V vs Fc$^{+/0}$, 180 s.

Table F.2: Experimental and Computational IR Data for selected Fe complexes. Values provided in cm^{-1} with computational data provided in parentheses.

1-Cp^{N3}Fe(CO)$_3^+$	**1-Cp^{N3}Fe(NCMe)**	**1-Cp^{N3}Fe**	**1-(Cp^{N3}Fe)$_2$**	**1-Cp^{N3}FeH**	**1-(endo-Cp^{N3}H)Fe(CO)$_3$**
1992, 2058	-	1895, 1986	1731, 1928	1907, 1968	1975, 1896, 1886
(1997, 2065)	(1825, 1893)	(1883, 1955)	(1741, 1908)	(1910, 1978)	(1985,1906, 1892)

To further validate these spectral observations, we reacted **1-Cp^{N3}Fe(CO)$_3^+$** with 2 equiv decamethylchromocene ($E_{1/2}$(Cp*$_2$Cr) = −1.53 vs. Fc$^{+/0}$, MeCN; Figure F.12, top). The nonpolar products were extracted with pentane and analyzed by ^{1}H NMR, confirming the generation of **1-Cp^{N3}FeH** while an IR spectrum (KBr) of the polar extracts (CH_3CN) show CO stretches at 1731 and 1928 cm^{-1}, corresponding to the proposed **1-(Cp^{N3}Fe)$_2$** as observed by IR-SEC. After working up the polar extracts, dark red-brown X-ray quality crystals were grown from a deep green solution, verifying the structure of the carbonyl-bridged dimer **1-(Cp^{N3}Fe)$_2$** (Figure F.12, bottom). Although the complex is unambiguously dimeric in the crystalline state, a solid-state IR spectrum of the crystals (KBr pellet) shows two additional sets of CO stretches (1895, 1980 cm^{-1}; Figure S2) which correlate with the DFT-computed values for monomeric **1-Cp^{N3}Fe** (Table F.2). The bulkier nature of the amine-rich ligand framework likely promotes dissociation to monomeric **1-Cp^{N3}Fe**, since it is known that the bulky peralkylated ligand C$_5$^{i}Pr$_5$ yields [Fe(CO)$_2$Cp^{i}Pr]$_2$ in the solid state but fully dissociates into paramagnetic Fe(CO)$_2$Cp^{i}Pr in solution.[454] To investigate the possibility of a monomer-dimer equilibrium, a known mass of **1-(Cp^{N3}Fe)$_2$** was dissolved in 0.5 mL THF-d_8 in the presence of a

known concentration of the internal standard (1,3,5-trimethoxybenzene; Figure S13). Based on the well-resolved signals corresponding to diamagnetic **1-($Cp^{N3}Fe)_2$**, the actual concentration was indeed much lower than expected, indicating that the NMR-silent product (**1-$Cp^{N3}Fe$**) is present. Using these data, a monomer-dimer equilibrium constant was calculated (K_{eq} = 0.11), indicating that the dimer is slightly favored at room temperature (Figure F.12). Variable temperature UV-vis spectra in THF were conducted between 0 and 60 °C, revealing decreases in intensity at 478 and 591 nm, suggesting that the equilibrium is being shifted towards the monomer at elevated temperatures (Figure S15). Preliminary EPR data show no signals at 77 K due to exclusive formation of the dimer while a broad anisotropic signal is observed at room temperature (deep green solution), qualitatively consistent with the NMR-derived equilibrium constant and UV-vis spectroscopic data.

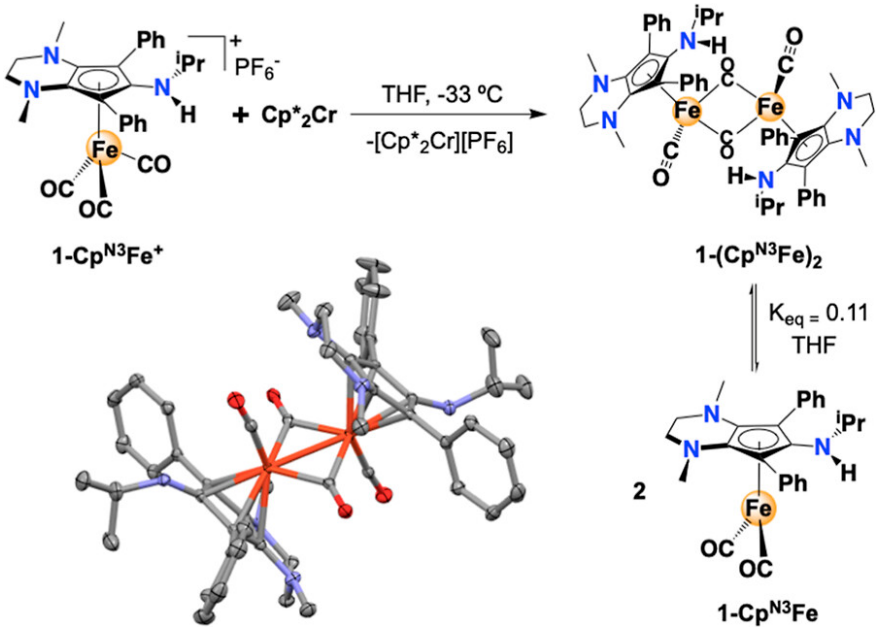

Figure F.12: Synthesis of **1-($Cp^{N3}Fe)_2$** with observed monomer-dimer equilibrium in solution and X-ray crystallographic structure of **1-($Cp^{N3}Fe)_2$** with 50% probability ellipsoids (bottom) (hydrogens omitted for clarity).

Importantly, electrochemical experiments were conducted to determine if the equilibrium mixture or **1-($Cp^{N3}Fe)_2$/1-$Cp^{N3}Fe$** is electrocatalytically active for H_2 production in the presence of exogenous acid. First, crystals of **1-($Cp^{N3}Fe)_2$** were dissolved in acetonitrile for CV analysis, showing partially reversible redox behavior which becomes reversible at higher scan rates (Figure S23), with a redox potential at $E_{1/2}$ = −1.53 V (Figure F.13, black trace, inset). We ascribe this to be the **1-$Cp^{N3}Fe$/1-$Cp^{N3}Fe^-$** redox couple which is in excellent agreement with theory ($E_{1/2}$ = −1.48 V). Consistent with our earlier report,[5] the observed irreversible oxidation peak at E_{pa} = −0.63 V in the CV trace is consistent with irreversible MeCN coordination after oxidation of **1-$Cp^{N3}Fe$** to **1-$Cp^{N3}Fe^+$**, generating **1-$Cp^{N3}Fe(NCMe)^+$** (Figure S24). The addition of 20 mM Tos_2NH to **1-($Cp^{N3}Fe)_2$** shows a very modest current enhancement at ca. −1.6 V (Figure F.13, blue trace). When compared with **1-$Cp^{N3}Fe(NCMe)^+$** in the presence of 20 mM Tos_2NH (Figure F.13, purple trace), **1-$Cp^{N3}Fe$/1-$Cp^{N3}Fe^-$** are incapable of directly entering the catalytic cycle and enabling rapid H_2 production. Therefore, **1-($Cp^{N3}Fe)_2$**

and **1-Cp^{N3}Fe** are off-cycle intermediates that may form to some extent during the electrocatalytic production of H_2 but are not directly responsible for rapid catalysis.

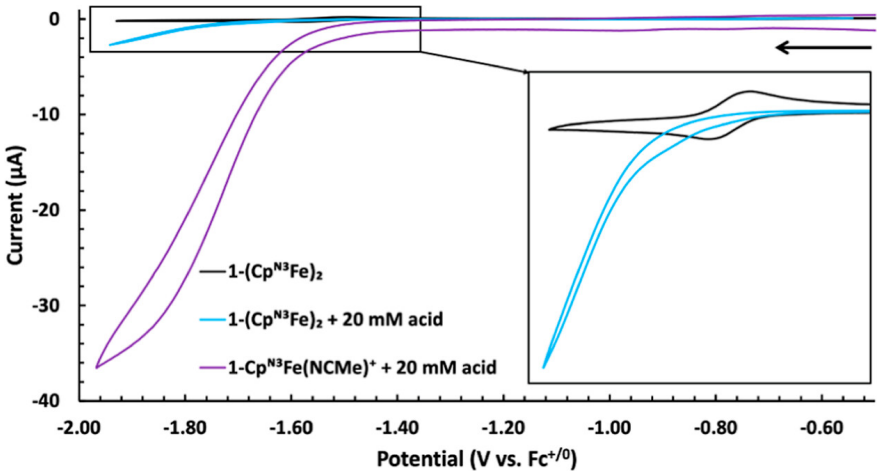

Figure F.13: IUPAC plotted, iR uncompensated CV traces of **1-(Cp^{N3}Fe)$_2$** (black trace), **1-(Cp^{N3}Fe)$_2$** with 20 mM Tos_2NH (blue trace), and **1-Cp^{N3}Fe(NCMe)$^+$** with 20 mM Tos_2NH (purple) at 1000 mV/s. Inset: CV traces of **1-(Cp^{N3}Fe)$_2$** (black) and **1-(Cp^{N3}Fe)$_2$** with 20 mM Tos_2NH (blue). Conditions: Ar, MeCN solvent, 0.1 M $[Bu_4N][PF_6]$, 1.0 mM analyte, PEEK-encased glassy carbon working electrode, Type 2 glassy carbon rod counter electrode, Ag/AgCl pseudoreference electrode in a frit-separated (Coralpor®) glass compartment containing solvent and electrolyte. Initial scan direction and starting position indicated with a black arrow.

F.2.5 Putting Together the Computed Mechanism

The free energies for the computed H_2 electrocatalytic mechanism are computed with an external applied potential of −1.49 V (the calculated redox potential of **1-Cp^{N3}Fe(NCMe)$^{+/0}$**), enabling the 1e$^-$ reduction of **1-Cp^{N3}Fe(NCMe)$^+$** and setting **1-Cp^{N3}Fe(NCMe)** to 0.0 kcal·mol^{-1} (Figure F.14). Geometries were optimized in Turbomole[342] V.7.5.1 program package on the efficient r^{2}SCAN-3c[137] level of theory with the implicit COSMO[164] solvation model for acetonitrile. With this level of theory initial guesses for the transition states were obtained with the DE-GSM[120,121] method and then optimized in Turbomole. Single-point energy calculations were conducted in ORCA[141] 5.0.3 with the r^{2}SCAN0-D4[9,67,113] functional employing the large def2-QZVPP basis set.[127] This method was selected based on its robust performance for benchmark sets of organometallic reaction energies (MOR41[134] and ROST61[126]) and barrier heights (MOBH35[135,136]). Thermostatistical corrections were obtained with the mRRHO[119] scheme based on analytical frequencies and solvation corrections to the final free energy were obtained with COSMO-RS[162,163] (see SI for details). After 1e$^-$ reduction of **1-Cp^{N3}Fe(NCMe)$^+$** to yield **1-Cp^{N3}Fe(NCMe)**, the initial protonation occurs with a reaction free energy barrier of 9.8 kcal·mol^{-1} (**TS1**) yielding the protonated species **1-(endo-Cp^{N3}H)Fe(NCMe)$^+$** in an exergonic reaction (ΔG = −17.9 kcal·mol^{-1}). This species is then reduced to yield **1-(endo-Cp^{N3}H)Fe(NCMe)** (ΔG = −25.8 kcal·mol^{-1}, with an external potential of −1.49 V) and in a follow-up step dissociates the MeCN ligand followed by ligand-to-metal proton transfer with a barrier of ΔG = 4.6 kcal·mol^{-1} estimated by a relaxed surface scan (see Figure F.9) yielding **1-Cp^{N3}FeH** (ΔG = −47.2 kcal·mol^{-1}). The second protonation step occurs directly at **1-Cp^{N3}FeH** with a barrier of

17.2 kcal·mol^{-1} (**TS2**) producing **1-Cp^{N3}Fe(H$_2$)$^+$** ($\Delta G = -38.4$ kcal·mol^{-1}), which releases H_2 to give **1-Cp^{N3}Fe(NCMe)$^+$**. Protonation of the amine is possible for **1-Cp^{N3}Fe** and **1-Cp^{N3}FeH** but the subsequent barriers for intramolecular proton transfer are too high in free energy (Figure S29 and Figure S30).

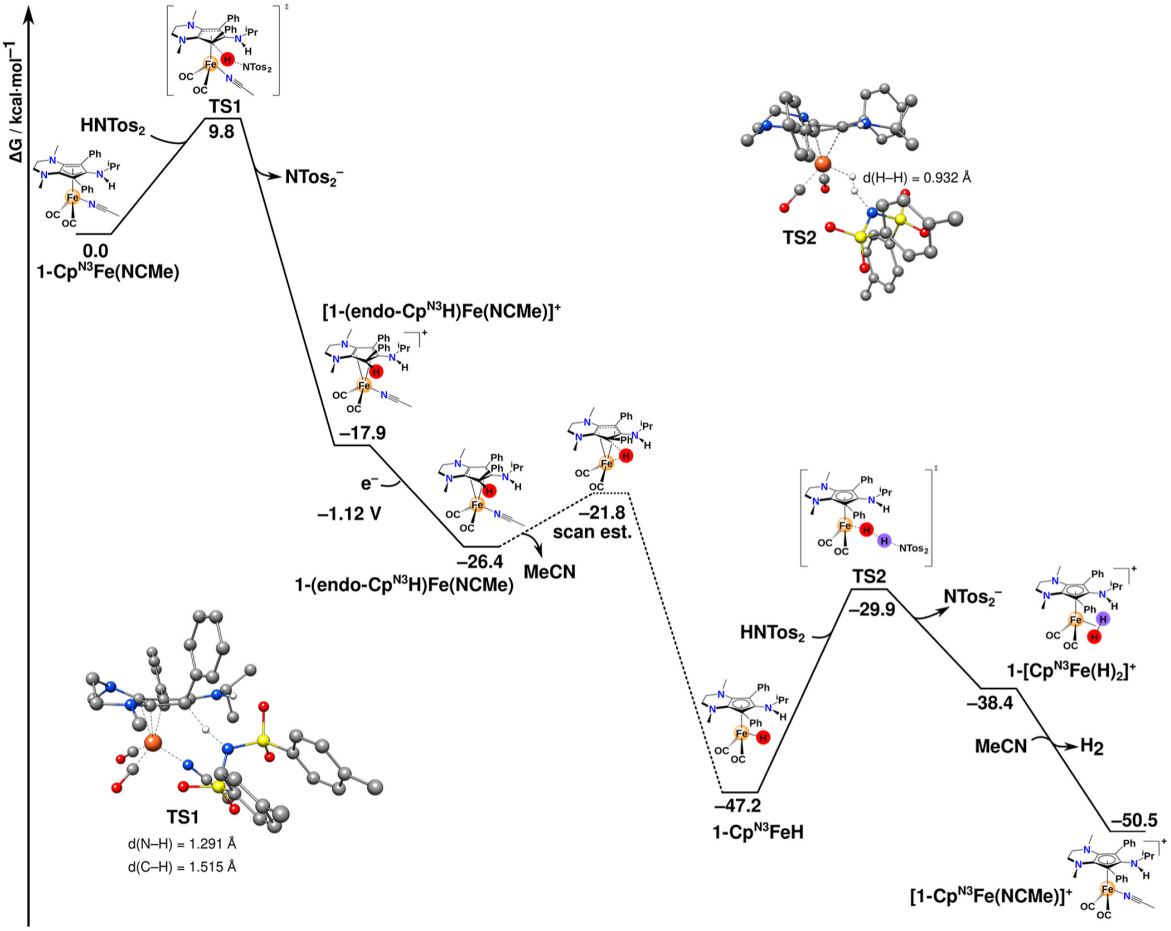

Figure F.14: Free energy profile for H_2 production relative to **1-Cp^{N3}Fe(NCMe)** given in kcal·mol^{-1} with an applied external potential of −1.49 V (calculated redox potential of **[1-Cp^{N3}Fe(NCMe)]$^{+/0}$**). The barrier for the MeCN dissociation followed by proton transfer from **1-(endo- Cp^{N3}H)Fe(NCMe)** was estimated by a relaxed potential energy scan (see Figure F.9) because no transition state could be obtained.

F.3 Conclusion

We have presented a detailed mechanistic analysis of electrocatalytic H_2 production using Fe complexes containing amine-rich Cp^{N3} ligands by using a blend of experimental techniques and state-of-the-art DFT calculations. Although it is still relatively uncommon for Cp ligands to behave as chemically noninnocent moieties, we demonstrate that regio- and stereoselective ligand protonation occurs at the Cp^{N3} ring, breaking its η^5-Cp^{N3} hapticity to generate a coordinated η^4-$Cp^{N3}H$ moiety. Surprisingly, the coordination of inner-sphere solvent (acetonitrile) to the Fe center is crucial for electrocatalysis,

effectively lowering the Cp ring protonation barrier and enabling rapid ligand-to-metal proton migration to generate an FeH complex. Using other nitrogen-based donor solvents inhibits catalysis. Synthesis of an isoelectronic (η^4-$Cp^{N3}H$)Fe complex containing an inner-sphere CO ligand in place of acetonitrile also renders the system inert, as the ligand-to-metal proton transfer is kinetically insurmountable due to the high CO ligand dissociation energy. Independent synthesis of on-cycle Cp^{N3}FeH intermediates shows that protonation in presence of exogenous acid furnishes high yields of H_2; DFT calculations indicate that direct Cp^{N3}FeH protonatiom is the TOF-determining step during catalysis. As more Cp-based systems are reported that demonstrate chemical non-innocence in proton-coupled electron transfer (PCET) chemistry, we hope that broader trends will be uncovered to better understand the kinetic origins of metal versus Cp ligand protonation for the efficient reduction of protons and other chemical fuel precursors.

F.4 Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.3c02911.

- General comments; syntheses; IR, NMR, and UV-vis spectra; electrochemistry; NMR kinetics; and computational details (PDF)
- Computed structures and energies (ZIP)
- Crystallographic data for CCDC 2271971 - 2271975 (CIF)

F.5 Acknowledgements

D.E.P.thanks the National Science Foundation (NSF) for support under Grant 2055097, and S.G.thanks the Deutsche Forschungsgemeinschaft under Grant 1927/16-1. This research was also supported by a Rutgers University-Newark startup grant (to D.E.P.). X-ray structural solutions were partly supported by the NSF under Grant 2018753. Supplemental funding for this project was provided by the Rutgers University-Newark Chancellor's Research Office.

Bibliography

[1] H. Neugebauer, B. Bädorf, S. Ehlert, A. Hansen, and S. Grimme, *High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods*, J. Comput. Chem. **44** (2023) 2120, DOI: 10.1002/jcc.27185.

[2] M. Bursch, H. Neugebauer, S. Ehlert, and S. Grimme, *Dispersion corrected r^2SCAN based global hybrid functionals: r^2SCANh, r^2SCAN0, and r^2SCAN50*, J. Chem. Phys. **156** (2022) 134105, DOI: 10.1063/5.0086040.

[3] H. Neugebauer, P. Pinksi, S. Grimme, F. Neese, and M. Bursch, *Assessment of DLPNO-MP2 Approximations in Double-Hybrid DFT*, J. Chem. Theory Comput. **19** (2023) 7695, DOI: 10.1021/acs.jctc.3c00896.

[4] H. Neugebauer, H. T. Vuong, J. L. Weber, R. A. Friesner, J. Shee, and A. Hansen, *Toward Benchmark-Quality Ab Initio Predictions for 3d Transition Metal Electrocatalysts: A Comparison of CCSD(T) and ph-AFQMC*, J. Chem. Theory Comput. **19** (2023) 6208, DOI: 10.1021/acs.jctc.3c00617.

[5] P. Sánchez, B. Goel, H. Neugebauer, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Ligand Protonation at Carbon, not Nitrogen, during H_2 Production with Amine-Rich Iron Electrocatalysts*, Inorg. Chem. **60** (2021) 17407, DOI: 10.1021/acs.inorgchem.1c03142.

[6] B. Goel, H. Neugebauer, A. I. VanderWeide, P. Sánchez, R. A. Lalancette, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Essential Roles of Cp Ring Activation and Coordinated Solvent During Electrocatalytic H_2 Production with Fe(Cp^{N3}) Complexes*, ACS Catal. **13** (2023) 13650, DOI: 10.1021/acscatal.3c02911.

[7] T. Gunasekara, G. P. Abramo, A. Hansen, H. Neugebauer, M. Bursch, S. Grimme, and J. R. Norton, *TEMPO-Mediated Catalysis of the Sterically Hindered Hydrogen Atom Transfer Reaction between $(C_5Ph_5)Cr(CO)_3H$ and a Trityl Radical*, J. Am. Chem. Soc. **141** (2019) 1882, DOI: 10.1021/jacs.8b12892.

[8] M. Bursch, E. Caldeweyher, A. Hansen, H. Neugebauer, S. Ehlert, and S. Grimme, *Understanding and Quantifying London Dispersion Effects in Organometallic Complexes*, Acc. Chem. Res. **52** (2019) 258, DOI: 10.1021/acs.accounts.8b00505.

[9] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme, *A generally applicable atomic-charge dependent London dispersion correction*, J. Chem. Phys. **150** (2019) 154122, DOI: 10.1063/1.5090222.

[10] M. Bursch, H. Neugebauer, and S. Grimme, *Structure Optimisation of Large Transition-Metal Complexes with Extended Tight-Binding Methods*, Angew Chem. Int. Ed. **58** (2019) 11078, DOI: 10.1002/anie.201904021.

[11] H. Neugebauer, F. Bohle, M. Bursch, A. Hansen, and S. Grimme, *Benchmark Study of Electrochemical Redox Potentials Calculated with Semiempirical and DFT Methods*, J. Phys. Chem. A **124** (2020) 7166, DOI: 10.1021/acs.jpca.0c05052.

[12] H. B. Vibbert, H. Neugebauer, J. R. Norton, A. Hansen, M. Bursch, and S. Grimme, *Hydrogen atom transfer rates from Tp-containing metal-hydrides to trityl radicals*, Can. J. Chem. **99** (2021) 216, DOI: 10.1139/cjc-2020-0392.

[13] D. S. Tresp, H. Neugebauer, S. Grimme, A. Hansen, and D. E. Prokopchuk, *Electronic Effects of Aminoindenyl Ligands Coordinated to Manganese: Structures and Properties of a Mn^0 Metalloradical and Bimetallic Mn^{-I}/Mn^{I} Adduct*, Organometallics **41** (2022) 3055, DOI: 10.1021/acs.organomet.2c00463.

[14] D. L. DuBois, *Development of Molecular Electrocatalysts for Energy Storage*, Inorg. Chem. **53** (2014) 3935, DOI: 10.1021/ic4026969.

[15] P. Chirik and R. Morris, *Getting Down to Earth: The Renaissance of Catalysis with Abundant Metals*, Acc. Chem. Res. **48** (2015) 2495, DOI: 10.1021/acs.accounts.5b00385.

[16] R. M. Bullock et al., *Using nature's blueprint to expand catalysis with Earth-abundant metals*, Science **369** (2020) eabc3183, DOI: 10.1126/science.abc3183.

[17] J. N. Hansen, H. Prats, K. K. Toudahl, N. Mørch Secher, K. Chan, J. Kibsgaard, and I. Chorkendorff, *Is There Anything Better than Pt for HER?* ACS Energy Lett. **6** (2021) 1175, DOI: 10.1021/acsenergylett.1c00246.

[18] R. Schlögl, *Chemical energy storage enables the transformation of fossil energy systems to sustainability*, Green Chem. **23** (4 2021) 1584, DOI: 10.1039/D0GC03171B.

[19] E. E. Benson, C. P. Kubiak, A. J. Sathrum, and J. M. Smieja, *Electrocatalytic and homogeneous approaches to conversion of CO_2 to liquid fuels*, Chem. Soc. Rev. **38** (1 2009) 89, DOI: 10.1039/B804323J.

[20] R. Francke, B. Schille, and M. Roemelt, *Homogeneously Catalyzed Electroreduction of Carbon Dioxide—Methods, Mechanisms, and Catalysts*, Chem. Rev. **118** (2018) 4631, DOI: 10.1021/acs.chemrev.7b00459.

[21] K. E. Dalle, J. Warnan, J. J. Leung, B. Reuillard, I. S. Karmel, and E. Reisner, *Electro- and Solar-Driven Fuel Synthesis with First Row Transition Metal Complexes*, Chem. Rev. **119** (2019) 2752, DOI: 10.1021/acs.chemrev.8b00392.

[22] N. W. Kinzel, C. Werlé, and W. Leitner, *Transition Metal Complexes as Catalysts for the Electroconversion of CO_2: An Organometallic Perspective*, Angew. Chem. Int. Ed. **60** (2021) 11628, DOI: 10.1002/anie.202006988.

[23] H. Liu, L. Wei, F. Liu, Z. Pei, J. Shi, Z.-j. Wang, D. He, and Y. Chen, *Homogeneous, Heterogeneous, and Biological Catalysts for Electrochemical N_2 Reduction toward NH_3 under Ambient Conditions*, ACS Catal. **9** (2019) 5245, DOI: 10.1021/acscatal.9b00994.

[24] M. L. Pegis, C. F. Wise, D. J. Martin, and J. M. Mayer, *Oxygen Reduction by Homogeneous Molecular Catalysts and Electrocatalysts*, Chem. Rev. **118** (2018) 2340, DOI: 10.1021/acs.chemrev.7b00542.

[25] V. Artero, M. Chavarot-Kerlidou, and M. Fontecave, *Splitting Water with Cobalt*, Angew. Chem. Int. Ed. **50** (2011) 7238, DOI: 10.1002/anie.201007987.

[26] P. Du and R. Eisenberg, *Catalysts made of earth-abundant elements (Co, Ni, Fe) for water splitting: Recent progress and future challenges*, Energy Environ. Sci. **5** (3 2012) 6012, DOI: 10.1039/C2EE03250C.

[27] G.-G. Luo, H.-L. Zhang, Y.-W. Tao, Q.-Y. Wu, D. Tian, and Q. Zhang, *Recent progress in ligand-centered homogeneous electrocatalysts for hydrogen evolution reaction*, Inorg. Chem. Front. **6** (2 2019) 343, DOI: 10.1039/C8QI01220B.

[28] A. Nandy, C. Duan, M. G. Taylor, F. Liu, A. H. Steeves, and H. J. Kulik, *Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning*, Chem. Rev. **121** (2021) 9927, DOI: 10.1021/acs.chemrev.1c00347.

[29] W. Thiel, *Theoretical Chemistry—Quo Vadis?* Angew. Chem. Int. Ed. **50** (2011) 9216, DOI: 10.1002/anie.201105305.

[30] M. Cavalleri, *Quantum chemistry reloaded*, Int. J. Quantum Chem. **113** (2013) 1, DOI: 10.1002/qua.24364.

[31] S. Grimme and P. R. Schreiner, *Computational Chemistry: The Fate of Current Methods and Future Challenges*, Angew. Chem. Int. Ed. **57** (2018) 4170, DOI: 10.1002/anie.201709943.

[32] K. N. Houk and F. Liu, *Holy Grails for Computational Organic Chemistry and Biochemistry*, Acc. Chem. Res. **50** (2017) 539, DOI: 10.1021/acs.accounts.6b00532.

[33] A. Krylov et al., *Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science*, J. Chem. Phys. **149** (2018) 180901, DOI: 10.1063/1.5052551.

[34] M. S. Gordon and T. L. Windus, *Editorial: Modern Architectures and Their Impact on Electronic Structure Theory*, Chem. Rev. **120** (2020) 9015, DOI: 10.1021/acs.chemrev.0c00700.

[35] R. Di Felice et al., *A Perspective on Sustainable Computational Chemistry Software Development and Integration*, J. Chem. Theory Comput. **19** (2023) 7056, DOI: 10.1021/acs.jctc.3c00419.

[36] S. Chen, M.-H. Ho, R. M. Bullock, D. L. DuBois, M. Dupuis, R. Rousseau, and S. Raugei, *Computing Free Energy Landscapes: Application to Ni-based Electrocatalysts with Pendant Amines for H_2 Production and Oxidation*, ACS Catal. **4** (2014) 229, DOI: 10.1021/cs401104w.

[37] T. Sperger, I. A. Sanhueza, I. Kalvet, and F. Schoenebeck, *Computational Studies of Synthetically Relevant Homogeneous Organometallic Catalysis Involving Ni, Pd, Ir, and Rh: An Overview of Commonly Employed DFT Methods and Mechanistic Insights*, Chem. Rev. **115** (2015) 9532, DOI: 10.1021/acs.chemrev.5b00163.

[38] S. Raugei, D. L. DuBois, R. Rousseau, S. Chen, M.-H. Ho, R. M. Bullock, and M. Dupuis, *Toward Molecular Catalysts by Computer*, Acc. Chem. Res. **48** (2015) 248, DOI: 10.1021/ar500342g.

[39] J. N. Harvey, F. Himo, F. Maseras, and L. Perrin, *Scope and Challenge of Computational Methods for Studying Mechanism and Reactivity in Homogeneous Catalysis*, ACS Catal. **9** (2019) 6803, DOI: 10.1021/acscatal.9b01537.

[40] Y.-h. Lam et al., *Applications of Quantum Chemistry in Pharmaceutical Process Development: Current State and Opportunities*, Org. Process Res. Dev. **24** (2020) 1496, DOI: 10.1021/acs.oprd.0c00222.

[41] J. Neugebauer and T. Hickel, *Density functional theory in materials science*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **3** (2013) 438, DOI: 10.1002/wcms.1125.

[42] B. W. J. Chen, L. Xu, and M. Mavrikakis, *Computational Methods in Heterogeneous Catalysis*, Chem. Rev. **121** (2021) 1007, DOI: 10.1021/acs.chemrev.0c01060.

[43] S. Ahn, M. Hong, M. Sundararajan, D. H. Ess, and M.-H. Baik, *Design and Optimization of Catalysts Based on Mechanistic Insights Derived from Quantum Chemical Reaction Modeling*, Chem. Rev. **119** (2019) 6509, DOI: 10.1021/acs.chemrev.9b00073.

[44] F. Jensen, *Introduction to Computational Chemistry*, Vol. 2, Wiley, 2007.

[45] N. Fey and J. M. Lynam, *Computational mechanistic study in organometallic catalysis: Why prediction is still a challenge*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **12** (2022) e1590, DOI: 10.1002/wcms.1590.

[46] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, First, Dover Publications, 1996.

[47] R. Izsák, A. V. Ivanov, N. S. Blunt, N. Holzmann, and F. Neese, *Measuring Electron Correlation: The Impact of Symmetry and Orbital Transformations*, J. Chem. Theory Comput. **19** (2023) 2703, DOI: 10.1021/acs.jctc.3c00122.

[48] C. David Sherrill and H. F. Schaefer III., "The Configuration Interaction Method: Advances in Highly Correlated Approaches," ed. by P.-O. Löwdin, J. R. Sabin, M. C. Zerner, and E. Brändas, vol. 34, Advances in Quantum Chemistry, Academic Press, 1999 143, DOI: 10.1016/S0065-3276(08)60532-8.

[49] J. J. Eriksen, *The Shape of Full Configuration Interaction to Come*, J. Phys. Chem. Lett. **12** (2021) 418, DOI: 10.1021/acs.jpclett.0c03225.

[50] M. Motta and S. Zhang, *Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8** (2018) e1364, DOI: 10.1002/wcms.1364.

[51] J. Shee, S. Zhang, D. R. Reichman, and R. A. Friesner, *Chemical Transformations Approaching Chemical Accuracy via Correlated Sampling in Auxiliary-Field Quantum Monte Carlo*, J. Chem. Theory Comput. **13** (2017) 2667, DOI: 10.1021/acs.jctc.7b00224.

[52] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, *Phaseless Auxiliary-Field Quantum Monte Carlo on Graphical Processing Units*, J. Chem. Theory Comput. **14** (2018) 4109, DOI: 10.1021/acs.jctc.8b00342.

[53] H. Shi and S. Zhang, *Some recent developments in auxiliary-field quantum Monte Carlo for real materials*, J. Chem. Phys. **154** (2021) 024107, DOI: 10.1063/5.0031024.

[54] J. Lee, H. Q. Pham, and D. R. Reichman, *Twenty Years of Auxiliary-Field Quantum Monte Carlo in Quantum Chemistry: An Overview and Assessment on Main Group Chemistry and Bond-Breaking*, J. Chem. Theory Comput. **18** (2022) 7024, DOI: 10.1021/acs.jctc.2c00802.

[55] J. Shee, J. L. Weber, D. R. Reichman, R. A. Friesner, and S. Zhang, *On the potentially transformative role of auxiliary-field quantum Monte Carlo in quantum chemistry: A highly accurate method for transition metals and beyond*, J. Chem. Phys. **158** (2023) 140901, DOI: 10.1063/5.0134009.

[56] R. J. Bartlett and M. Musiał, *Coupled-cluster theory in quantum chemistry*, Rev. Mod. Phys. **79** (1 2007) 291, DOI: 10.1103/RevModPhys.79.291.

[57] J. A. Pople, M. Head-Gordon, and K. Raghavachari, *Quadratic configuration interaction. A general technique for determining electron correlation energies*, J. Chem. Phys. **87** (1987) 5968, DOI: 10.1063/1.453520.

[58] J. A. Pople, *Nobel Lecture: Quantum chemical models*, Rev. Mod. Phys. **71** (5 1999) 1267, DOI: 10.1103/RevModPhys.71.1267.

[59] A. Hansen, D. G. Liakos, and F. Neese, *Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals*, J. Chem. Phys. **135** (2011) 214102, DOI: 10.1063/1.3663855.

[60] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, and F. Neese, *A new near-linear scaling, efficient and accurate, open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory*, J. Chem. Phys. **146** (2017) 164105, DOI: 10.1063/1.4981521.

[61] C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, *Natural triple excitations in local coupled cluster calculations with pair natural orbitals*, J. Chem. Phys. **139** (2013) 134101, DOI: 10.1063/1.4821834.

[62] Q. Ma and H.-J. Werner, *Explicitly correlated local coupled-cluster methods using pair natural orbitals*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8** (2018) e1371, DOI: 10.1002/wcms.1371.

[63] R. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1994.

[64] W. Kohn, A. D. Becke, and R. G. Parr, *Density Functional Theory of Electronic Structure*, J. Phys. Chem. **100** (1996) 12974, DOI: 10.1021/jp9606691.

[65] A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Challenges for Density Functional Theory*, Chem. Rev. **112** (2012) 289, DOI: 10.1021/cr200107z.

[66] N. Mardirossian and M. Head-Gordon, *Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals*, Mol. Phys. **115** (2017) 2315, DOI: 10.1080/00268976.2017.1333644.

[67] M. Bursch, J.-M. Mewes, A. Hansen, and S. Grimme, *Best-Practice DFT Protocols for Basic Molecular Computational Chemistry*, Angew Chem. Int. Ed. **61** (2022) e202205735, DOI: 10.1002/anie.202205735.

[68] P. Geerlings, F. De Proft, and W. Langenaeker, *Conceptual Density Functional Theory*, Chem. Rev. **103** (2003) 1793, DOI: 10.1021/cr990029p.

[69] A. S. Christensen, T. Kubař, Q. Cui, and M. Elstner, *Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications*, Chem. Rev. **116** (2016) 5301, DOI: 10.1021/acs.chemrev.5b00584.

[70] S. Grimme, C. Bannwarth, and P. Shushkov, *A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86)*, J. Chem. Theory Comput. **13** (2017) 1989, DOI: 10.1021/acs.jctc.7b00118.

[71] C. Bannwarth, S. Ehlert, and S. Grimme, *GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions*, J. Chem. Theory Comput. **15** (2019) 1652, DOI: 10.1021/acs.jctc.8b01176.

[72] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, and S. Grimme, *Extended tight-binding quantum chemistry methods*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **11** (2021) e1493, DOI: 10.1002/wcms.1493.

[73] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, Phys. Rev. B **58** (11 1998) 7260, DOI: 10.1103/PhysRevB.58.7260.

[74] M. Gaus, Q. Cui, and M. Elstner, *DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB)*, J. Chem. Theory Comput. **7** (2011) 931, DOI: 10.1021/ct100684s.

[75] J. J. P. Stewart, *Optimization of parameters for semiempirical methods I. Method*, J. Comput. Chem. **10** (1989) 209, DOI: 10.1002/jcc.540100208.

[76] J. J. P. Stewart, *Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements*, J. Mol. Model. **13** (2007) 1173, DOI: 10.1007/s00894-007-0233-4.

[77] J. J. P. Stewart, *Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters*, J. Mol. Model. **19** (2013) 1, DOI: 10.1007/s00894-012-1667-x.

[78] P. Dral, X. Wu, L. Spörkel, A. Koslowski, W. Weber, R. Steiger, M. Scholten, and W. Thiel, *Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Theory, Implementation, and Parameters*, J. Chem. Theory Comput. **12** (2016) 1082, DOI: 10.1021/acs.jctc.5b01046.

[79] P. Pracht, F. Bohle, and S. Grimme, *Automated exploration of the low-energy chemical space with fast quantum chemical methods*, Phys. Chem. Chem. Phys. **22** (14 2020) 7169, DOI: 10.1039/C9CP06869D.

[80] S. Grimme, *Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations*, J. Chem. Theory Comput. **15** (2019) 2847, DOI: 10.1021/acs.jctc.9b00143.

[81] S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher, and M. Stahn, *Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Nonrigid Molecules*, J. Phys. Chem. A **125** (2021) 4039, DOI: 10.1021/acs.jpca.1c00971.

[82] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. I. Goddard, and W. M. Skiff, *UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations*, J. Am. Chem. Soc. **114** (1992) 10024, DOI: 10.1021/ja00051a040.

[83] S. Spicher and S. Grimme, *Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems*, Angew. Chem. Int. Ed. **59** (2020) 15665, DOI: 10.1002/anie.202004239.

[84] J. Shee, M. Loipersberger, A. Rettig, J. Lee, and M. Head-Gordon, *Regularized Second-Order Møller–Plesset Theory: A More Accurate Alternative to Conventional MP2 for Noncovalent Interactions and Transition Metal Thermochemistry for the Same Computational Cost*, J. Phys. Chem. Lett. **12** (2021) 12084, DOI: 10.1021/acs.jpclett.1c03468.

[85] R. Hoffmann, S. Alvarez, C. Mealli, A. Falceto, T. J. I. Cahill, T. Zeng, and G. Manca, *From Widely Accepted Concepts in Coordination Chemistry to Inverted Ligand Fields*, Chem. Rev. **116** (2016) 8173, DOI: 10.1021/acs.chemrev.6b00251.

[86] J. Shee, M. Loipersberger, D. Hait, J. Lee, and M. Head-Gordon, *Revealing the nature of electron correlation in transition metal complexes with symmetry breaking and chemical intuition*, J. Chem. Phys. **154** (2021) 194109, DOI: 10.1063/5.0047386.

[87] W. Jiang, N. J. DeYonker, and A. K. Wilson, *Multireference Character for 3d Transition-Metal-Containing Molecules*, J. Chem. Theory Comput. **8** (2012) 460, DOI: 10.1021/ct2006852.

[88] A. Khedkar and M. Roemelt, *Modern multireference methods and their application in transition metal chemistry*, Phys. Chem. Chem. Phys. **23** (32 2021) 17097, DOI: 10.1039/D1CP02640B.

[89] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, 2000.

[90] A. V. Marenich, J. Ho, M. L. Coote, C. J. Cramer, and D. G. Truhlar, *Computational electrochemistry: prediction of liquid-phase reduction potentials*, Phys. Chem. Chem. Phys. **16** (29 2014) 15068, DOI: 10.1039/C4CP01572J.

[91] C. J. Cramer and D. G. Truhlar, *Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics*, Chem. Rev. **99** (1999) 2161, DOI: 10.1021/cr960149m.

[92] P. Pinski, C. Riplinger, E. F. Valeev, and F. Neese, *Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals*, J. Chem. Phys. **143** (2015), 034108, DOI: 10.1063/1.4926879.

[93] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **384** (1926) 361, DOI: 10.1002/andp.19263840404.

[94] M. Born and R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Ann. Phys. **389** (1927) 457, DOI: 10.1002/andp.19273892002.

[95] J. Lee and M. Head-Gordon, *Distinguishing artificial and essential symmetry breaking in a single determinant: approach and application to the C_{60}, C_{36}, and C_{20} fullerenes*, Phys. Chem. Chem. Phys. **21** (9 2019) 4763, DOI: 10.1039/C8CP07613H.

[96] P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev. **136** (3B 1964) B864, DOI: 10.1103/PhysRev.136.B864.

[97] W. Kohn, *Nobel Lecture: Electronic structure of matter—wave functions and density functionals*, Rev. Mod. Phys. **71** (1999) 1253, DOI: 10.1103/RevModPhys.71.1253.

[98] W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev. **140** (4A 1965) A1133, DOI: 10.1103/PhysRev.140.A1133.

[99] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, *A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions*, Phys. Chem. Chem. Phys. **19** (48 2017) 32184, DOI: 10.1039/C7CP04913G.

[100] J. P. Perdew and K. Schmidt, *Jacob's ladder of density functional approximations for the exchange-correlation energy*, AIP Conf. Proc. **577** (2001) 1, DOI: 10.1063/1.1390175.

[101] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77** (18 1996) 3865, DOI: 10.1103/PhysRevLett.77.3865.

[102] A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A **38** (6 1988) 3098, DOI: 10.1103/PhysRevA.38.3098.

[103] J. P. Perdew, *Density-functional approximation for the correlation energy of the inhomogeneous electron gas*, Phys. Rev. B **33** (12 1986) 8822, DOI: 10.1103/PhysRevB.33.8822.

[104] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, *Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes*, J. Chem. Phys. **119** (2003) 12129, DOI: 10.1063/1.1626543.

[105] J. Sun, A. Ruzsinszky, and J. P. Perdew, *Strongly Constrained and Appropriately Normed Semilocal Density Functional*, Phys. Rev. Lett. **115** (3 2015) 036402, DOI: 10.1103/PhysRevLett.115.036402.

[106] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, *Accurate and Numerically Efficient r^2SCAN Meta-Generalized Gradient Approximation*, J. Phys. Chem. Lett. **11** (2020) 8208, DOI: 10.1021/acs.jpclett.0c02405.

[107] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, *Correction to: 'Accurate and Numerically Efficient r^2SCAN Meta-Generalized Gradient Approximation' (J. Phys. Chem. Lett. (2020) 11:19 (8208-8215) DOI: 10.1021/acs.jpclett.0c02405)*, J. Phys. Chem. Lett. (2020) 9248, DOI: 10.1021/acs.jpclett.0c03077.

[108] J. Harris, *Adiabatic-connection approach to Kohn-Sham theory*, Phys. Rev. A **29** (1984) 1648, DOI: 10.1103/PhysRevA.29.1648.

[109] E. Brémond, I. Ciofini, J. C. Sancho-García, and C. Adamo, *Nonempirical Double-Hybrid Functionals: An Effective Tool for Chemists*, Acc. Chem. Res. **49** (2016) 1503, DOI: 10.1021/acs.accounts.6b00232.

[110] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, *Dispersion-Corrected Mean-Field Electronic Structure Methods*, Chem. Rev. **116** (2016) 5105, DOI: 10.1021/acs.chemrev.5b00533.

[111] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, J. Chem. Phys. **132** (2010) 154104, DOI: 10.1063/1.3382344.

[112] S. Grimme, S. Ehrlich, and L. Goerigk, *Effect of the Damping Function in Dispersion Corrected Density Functional Theory*, J. Comput. Chem. **32** (2011) 1456, DOI: 10.1002/jcc.21759.

[113] E. Caldeweyher, C. Bannwarth, and S. Grimme, *Extension of the D3 dispersion coefficient model*, J. Chem. Phys. **147** (2017) 034112, DOI: 10.1063/1.4993215.

[114] E. Caldeweyher, J.-M. Mewes, S. Ehlert, and S. Grimme, *Extension and evaluation of the D4 London-dispersion model for periodic systems*, Phys. Chem. Chem. Phys. **22** (16 2020) 8499, DOI: 10.1039/D0CP00502A.

[115] B. M. Axilrod and E. Teller, *Interaction of the van der Waals Type Between Three Atoms*, J. Chem. Phys. **11** (1943) 299.

[116] Y. Muto, *Force between nonpolar molecules*, Proc. Phys. Soc. Jpn. **17** (1943) 629.

[117] Y. Yang, H. Yu, D. York, Q. Cui, and M. Elstner, *Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method: Third-Order Expansion of the Density Functional Theory Total Energy and Introduction of a Modified Effective Coulomb Interaction*, J. Phys. Chem. A **111** (2007) 10861, DOI: 10.1021/jp074167r.

[118] A. M. Köster, M. Leboeuf, and D. R. Salahub,
"Molecular electrostatic potentials from density functional theory,"
Molecular Electrostatic Potentials Concepts and Applications, vol. 3, 1996 105.

[119] S. Grimme, *Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory*, Chem. Eur. J. **18** (2012) 9955, DOI: 10.1002/chem.201200497.

[120] P. M. Zimmerman, *Growing string method with interpolation and optimization in internal coordinates: Method and examples*, J. Chem. Phys. **138** (2013) 184102,
DOI: 10.1063/1.4804162.

[121] P. Zimmerman,
Reliable Transition State Searches Integrated with the Growing String Method,
J. Chem. Theory Comput. **9** (2013) 3043, DOI: 10.1021/ct400319w.

[122] M. Bursch, A. Hansen, P. Pracht, J. T. Kohn, and S. Grimme,
Theoretical study on conformational energies of transition metal complexes,
Phys. Chem. Chem. Phys. **23** (1 2021) 287, DOI: 10.1039/D0CP04696E.

[123] J. P. Unsleber et al., *High-throughput ab initio reaction mechanism exploration in the cloud with automated multi-reference validation*, J. Chem. Phys. **158** (2023) 084803,
DOI: 10.1063/5.0136526.

[124] J. P. Perdew, K. Burke, and M. Ernzerhof,
Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)],
Phys. Rev. Lett. **78** (7 1997) 1396, DOI: 10.1103/PhysRevLett.78.1396.

[125] R. Ghafarian Shirazi, F. Neese, and D. A. Pantazis,
Accurate Spin-State Energetics for Aryl Carbenes, J. Chem. Theory Comput. **14** (2018) 4733,
DOI: 10.1021/acs.jctc.8b00587.

[126] L. R. Maurer, M. Bursch, S. Grimme, and A. Hansen, *Assessing Density Functional Theory for Chemically Relevant Open-Shell Transition Metal Reactions*,
J. Chem. Theory Comput **17** (2021) 6134, DOI: 10.1021/acs.jctc.1c00659.

[127] F. Weigend and R. Ahlrichs, *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. **7** (18 2005) 3297, DOI: 10.1039/B508541A.

[128] C. M. Zall, L. J. Clouston, V. G. J. Young, K. Ding, H. J. Kim, D. Zherebetskyy, Y.-S. Chen, E. Bill, L. Gagliardi, and C. C. Lu, *Mixed-Valent Dicobalt and Iron–Cobalt Complexes with High-Spin Configurations and Short Metal–Metal Bonds*, Inorg. Chem. **52** (2013) 9216,
DOI: 10.1021/ic400292g.

[129] A. P. Bartók and J. R. Yates, *Regularized SCAN functional*,
J. Chem. Phys. **150** (2019) 161101, DOI: 10.1063/1.5094646.

[130] J. G. Brandenburg, J. E. Bates, J. Sun, and J. P. Perdew, *Benchmark tests of a strongly constrained semilocal functional with a long-range dispersion correction*,
Phys. Rev. B **94** (11 2016) 115144, DOI: 10.1103/PhysRevB.94.115144.

[131] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun,
Construction of meta-GGA functionals through restoration of exact constraint adherence to regularized SCAN functionals, J. Chem. Phys. **156** (2022) 034109,
DOI: 10.1063/5.0073623.

[132] S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew, and J. G. Brandenburg, *r^2SCAN-D4: Dispersion corrected meta-generalized gradient approximation for general chemical applications*, J. Chem. Phys. **154** (2021) 061101,
DOI: 10.1063/5.0041008.

[133] O. A. Vydrov and T. Van Voorhis,
Nonlocal van der Waals density functional: The simpler the better,
J. Chem. Phys. **133** (2010) 244103, DOI: 10.1063/1.3521275.

[134] S. Dohm, A. Hansen, M. Steinmetz, S. Grimme, and M. P. Checinski, *Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions*,
J. Chem. Theory Comput **14** (2018) 2596, DOI: 10.1021/acs.jctc.7b01183.

[135] M. A. Iron and T. Janes, *Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database*,
J. Phys. Chem. A **123** (2019) 3761, DOI: 10.1021/acs.jpca.9b01546.

[136] M. A. Iron and T. Janes,
Correction to "Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database",
J. Phys. Chem. A **123** (2019) 6379, DOI: 10.1021/acs.jpca.9b06135.

[137] S. Grimme, A. Hansen, S. Ehlert, and J.-M. Mewes,
r^2SCAN-3c: A "Swiss army knife" composite electronic-structure method,
J. Chem. Phys. **154** (2021) 064103, DOI: 10.1063/5.0040021.

[138] J. B. Kleine Büning and S. Grimme,
Computation of CCSD(T)-Quality NMR Chemical Shifts via Δ-Machine Learning from DFT,
J. Chem. Theory Comput. **19** (2023) 3601, DOI: 10.1021/acs.jctc.3c00165.

[139] Y. J. Franzke et al., *TURBOMOLE: Today and Tomorrow*,
J. Chem. Theory Comput. **19** (2023) 6859, DOI: 10.1021/acs.jctc.3c00347.

[140] A. Erba et al., *CRYSTAL23: A Program for Computational Solid State Physics and Chemistry*,
J. Chem. Theory Comput. **19** (2023) 6891, DOI: 10.1021/acs.jctc.2c00958.

[141] F. Neese, *Software update: The ORCA program system—Version 5.0*,
Wiley Interdiscip. Rev. Comput. Mol. Sci. **12** (2022) e1606, DOI: 10.1002/wcms.1606.

[142] Q. Sun et al., *Recent developments in the PySCF program package*,
J. Chem. Phys. **153** (2020) 024109, DOI: 10.1063/5.0006074.

[143] S. Lehtola, C. Steigemann, M. J. Oliveira, and M. A. Marques, *Recent developments in libxc — A comprehensive library of functionals for density functional theory*, SoftwareX **7** (2018) 1,
DOI: 10.1016/j.softx.2017.11.002.

[144] L. Goerigk and S. Grimme, *Double-hybrid density functionals*,
Wiley Interdiscip. Rev. Comput. Mol. Sci. **4** (2014) 576, DOI: 10.1002/wcms.1193.

[145] N. Mehta, M. Casanova-Páez, and L. Goerigk,
Semi-empirical or non-empirical double-hybrid density functionals: Which are more robust?
Phys. Chem. Chem. Phys. **20** (2018) 23175, DOI: 10.1039/c8cp03852j.

[146] J. M. L. Martin and G. Santra, *Empirical Double-Hybrid Density Functional Theory: A 'Third Way' in Between WFT and DFT*, Isr. J. Chem. **60** (2020) 787,
DOI: 10.1002/ijch.201900114.

[147] P. Pinski and F. Neese, *Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2)*,
J. Chem. Phys. **150** (2019), 164102, DOI: 10.1063/1.5086544.

[148] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese,
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. **144** (2016), 024109, DOI: 10.1063/1.4939030.

[149] S. Grimme,
Semiempirical hybrid density functional with perturbative second-order correlation,
J. Chem. Phys. **124** (2006) 034108, DOI: 10.1063/1.2148954.

[150] A. Altun, F. Neese, and G. Bistoni, *Extrapolation to the Limit of a Complete Pair Natural Orbital Space in Local Coupled-Cluster Calculations*,
J. Chem. Theory Comput. **16** (2020) 6142, DOI: 10.1021/acs.jctc.0c00344.

[151] R. Sedlak, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay, and P. Hobza,
Accuracy of quantum chemical methods for large noncovalent complexes,
J. Chem. Theory Comput. **9** (2013) 3364, DOI: 10.1021/ct400036b.

[152] Y. S. Al-Hamdani, P. R. Nagy, A. Zen, D. Barton, M. Kállay, J. G. Brandenburg, and A. Tkatchenko, *Interactions between large molecules pose a puzzle for reference quantum mechanical methods*, Nat. Commun. **12** (2021) 3927,
DOI: 10.1038/s41467-021-24119-3.

[153] K. D. Vogiatzis, M. V. Polynski, J. K. Kirkland, J. Townsend, A. Hashemi, C. Liu, and E. A. Pidko, *Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities*, Chem. Rev. **119** (2019) 2453,
DOI: 10.1021/acs.chemrev.8b00361.

[154] A. Rettig, J. Shee, J. Lee, and M. Head-Gordon, *Revisiting the Orbital Energy-Dependent Regularization of Orbital-Optimized Second-Order Møller–Plesset Theory*,
J. Chem. Theory Comput. **18** (2022) 5382, DOI: 10.1021/acs.jctc.2c00641.

[155] H. Yuan and D. Cremer,
The expectation value of the spin operator $\hat{S}^2$ as a diagnostic tool in coupled cluster theory: The advantages of using UHF-CCSD theory for the description of homolytic dissociation,
Chem. Phys. Lett. **324** (2000) 389, DOI: 10.1016/S0009-2614(00)00644-8.

[156] C. Adamo and V. Barone,
Toward reliable density functional methods without adjustable parameters: The PBE0 model,
J. Chem. Phys. **110** (1999) 6158, DOI: 10.1063/1.478522.

[157] N. Mardirossian and M. Head-Gordon,
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. **16** (2014) 9904, DOI: 10.1039/C3CP54374A.

[158] P. Pracht, F. Bohle, and S. Grimme,
Automated exploration of the low-energy chemical space with fast quantum chemical methods,
Phys. Chem. Chem. Phys. **22** (2020) 7169, DOI: 10.1039/c9cp06869d.

[159] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, and F. Neese,
Calculation of Solvent Shifts on Electronic g-Tensors with the Conductor-Like Screening Model (COSMO) and Its Self-Consistent Generalization to Real Solvents (Direct COSMO-RS),
J. Phys. Chem. A **110** (2006) 2235, DOI: 10.1021/jp056016z.

[160] S. Spicher, C. Plett, P. Pracht, A. Hansen, and S. Grimme, *Automated Molecular Cluster Growing for Explicit Solvation by Efficient Force Field and Tight Binding Methods*,
J. Chem. Theory Comput. **18** (2022) 3174, DOI: 10.1021/acs.jctc.2c00239.

[161] S. Spicher and S. Grimme, *Single-Point Hessian Calculations for Improved Vibrational Frequencies and Rigid-Rotor-Harmonic-Oscillator Thermodynamics*,
J. Chem. Theory Comput. **17** (2021) 1701, DOI: 10.1021/acs.jctc.0c01306.

[162] A. Klamt, *Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena*, J. Phys. Chem. **99** (1995) 2224,
DOI: 10.1021/j100007a062.

[163] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz,
Refinement and Parametrization of COSMO-RS, J. Phys. Chem. A **102** (1998) 5074,
DOI: 10.1021/jp980017s.

[164] A. Klamt and G. Schüürmann, *COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient*,
J. Chem. Soc., Perkin Trans. 2 (5 1993) 799, DOI: 10.1039/P29930000799.

[165] S. Dohm, M. Bursch, A. Hansen, and S. Grimme, *Semiautomated Transition State Localization for Organometallic Complexes with Semiempirical Quantum Chemical Methods*,
J. Chem. Theory Comput. **16** (2020) 2002, DOI: 10.1021/acs.jctc.9b01266.

[166] P. Pracht and S. Grimme, *Efficient Quantum-Chemical Calculations of Acid Dissociation Constants from Free-Energy Relationships*, J. Phys. Chem. A **125** (2021) 5681,
DOI: 10.1021/acs.jpca.1c03463.

[167] C. M. Sterling and R. Bjornsson,
Multistep Explicit Solvation Protocol for Calculation of Redox Potentials,
J. Chem. Theory Comput. **15** (2019) 52, DOI: 10.1021/acs.jctc.8b00982.

[168] A. McSloy, G. Fan, W. Sun, C. Hölzer, M. Friede, S. Ehlert, N.-E. Schütte, S. Grimme, T. Frauenheim, and B. Aradi,
TBMaLT, a flexible toolkit for combining tight-binding and machine learning,
J. Chem. Phys. **158** (2023) 034801, DOI: 10.1063/5.0132892.

[169] F. Hu, F. He, and D. J. Yaron, *Treating Semiempirical Hamiltonians as Flexible Machine Learning Models Yields Accurate and Interpretable Results*, J. Chem. Theory Comput. **19** (2023) 6185, DOI: 10.1021/acs.jctc.3c00491.

[170] S. Grimme, M. Müller, and A. Hansen, *A non-self-consistent tight-binding electronic structure potential in a polarized double-ζ basis set for all spd-block elements up to $Z = 86$*, J. Chem. Phys. **158** (2023) 124111, DOI: 10.1063/5.0137838.

[171] T. A. Niehaus and F. Della Sala, *Range separated functionals in the density functional based tight-binding method: Formalism*, physica status solidi (b) **249** (2012) 237, DOI: 10.1002/pssb.201100694.

[172] J. Ning, M. Kothakonda, J. W. Furness, A. D. Kaplan, S. Ehlert, J. G. Brandenburg, J. P. Perdew, and J. Sun, *Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems:* r^{2}SCAN + rVV10, Phys. Rev. B **106** (7 2022) 075422, DOI: 10.1103/PhysRevB.106.075422.

[173] D. Kalle, J. Bahr, T. J. Keller, J. Kleine Büning, S. Grimme, M. Bursch, S.-S. Jester, and S. Höger, *Modular Bicyclophane-Based Molecular Platforms*, Chem. Eur. J. () e202302662, DOI: 10.1002/chem.202302662.

[174] L. Wittmann, H. Neugebauer, S. Grimme, and M. Bursch, *Dispersion-corrected r^2SCAN based double-hybrid functionals*, J. Chem. Phys. **159** (2023) 224103, DOI: 10.1063/5.0174988.

[175] É. Brémond, M. Savarese, Á. J. Pérez-Jiménez, J. C. Sancho-García, and C. Adamo, *Range-Separated Double-Hybrid Functional from Nonempirical Constraints*, J. Chem. Theory Comput. **14** (2018) 4052, DOI: 10.1021/acs.jctc.8b00261.

[176] A. Altun, C. Riplinger, F. Neese, and G. Bistoni, *Exploring the Accuracy Limits of PNO-Based Local Coupled-Cluster Calculations for Transition-Metal Complexes*, J. Chem. Theory Comput. **19** (2023) 2039, DOI: 10.1021/acs.jctc.3c00087.

[177] J. L. Weber, H. Vuong, P. A. Devlaminck, J. Shee, J. Lee, D. R. Reichman, and R. A. Friesner, *A Localized-Orbital Energy Evaluation for Auxiliary-Field Quantum Monte Carlo*, J. Chem. Theory Comput. **18** (2022) 3447, DOI: 10.1021/acs.jctc.2c00111.

[178] E. I. Ioannidis, T. Z. H. Gani, and H. J. Kulik, *molSimplify: A toolkit for automating discovery in Inorg. Chem.*, J. Comput. Chem. **37** (2016) 2106, DOI: 10.1002/jcc.24437.

[179] T. A. Young, J. J. Silcock, A. J. Sterling, and F. Duarte, *autodE: Automated Calculation of Reaction Energy Profiles— Application to Organic and Organometallic Reactions*, Angew. Chem. Int. Ed. **60** (2021) 4266, DOI: 10.1002/anie.202011941.

[180] Q. Zhao and B. M. Savoie, *Algorithmic Explorations of Unimolecular and Bimolecular Reaction Spaces*, Angew. Chem. Int. Ed. **61** (2022) e202210693, DOI: 10.1002/anie.202210693.

[181] J. P. Unsleber, S. A. Grimmel, and M. Reiher, *Chemoton 2.0: Autonomous Exploration of Chemical Reaction Networks*, J. Chem. Theory Comput. **18** (2022) 5393, DOI: 10.1021/acs.jctc.2c00193.

[182] A. Hashemi, S. Bougueroua, M.-P. Gaigeot, and E. A. Pidko, *ReNeGate: A Reaction Network Graph-Theoretical Tool for Automated Mechanistic Studies in Computational Homogeneous Catalysis*, J. Chem. Theory Comput. **18** (2022) 7470, DOI: 10.1021/acs.jctc.2c00404.

[183] J. P. Janet and H. J. Kulik, *Predicting electronic structure properties of transition metal complexes with neural networks*, Chem. Sci. **8** (7 2017) 5137, DOI: 10.1039/C7SC01247K.

[184] J. P. Janet and H. J. Kulik, *Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure–Property Relationships*, J. Phys. Chem. A **121** (2017) 8939, DOI: 10.1021/acs.jpca.7b08750.

[185] M. Meuwly, *Machine Learning for Chemical Reactions*, Chem. Rev. **121** (2021) 10218, DOI: 10.1021/acs.chemrev.1c00033.

[186] J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, and A. Tkatchenko, *Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems*, Chem. Rev. **121** (2021) 9816, DOI: 10.1021/acs.chemrev.1c00107.

[187] A. Hagg and K. N. Kirschner, *Open-Source Machine Learning in Computational Chemistry*, J. Chem. Inf. Model. **63** (2023) 4505, DOI: 10.1021/acs.jcim.3c00643.

[188] M. Müller, A. Hansen, and S. Grimme, *ωB97X-3c: A composite range-separated hybrid DFT method with a molecule-optimized polarized valence double-ζ basis set*, J. Chem. Phys. **158** (2023) 014103, DOI: 10.1063/5.0133026.

[189] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, *Introduction of n-electron valence states for multireference perturbation theory*, J. Chem. Phys. **114** (2001) 10252, DOI: 10.1063/1.1361246.

[190] C. Angeli, R. Cimiraglia, and J.-P. Malrieu, *n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants*, J. Chem. Phys. **117** (2002) 9138, DOI: 10.1063/1.1515317.

[191] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *Second-order perturbation theory with a CASSCF reference function*, The Journal of Physical Chemistry **94** (1990) 5483, DOI: 10.1021/j100377a012.

[192] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, *Second-order perturbation theory with a complete active space self-consistent field reference function*, J. Chem. Phys. **96** (1992) 1218, DOI: 10.1063/1.462209.

[193] Y. Guo, K. Sivalingam, E. F. Valeev, and F. Neese, *SparseMaps-A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital N-electron valence perturbation theory*, J. Chem. Phys. **144** (2016) 094111, DOI: 10.1063/1.4942769.

[194] C. Kollmar, K. Sivalingam, and F. Neese, *An alternative choice of the zeroth-order Hamiltonian in CASPT2 theory*, J. Chem. Phys. **152** (2020) 214110, DOI: 10.1063/5.0010019.

[195] E. R. Sayfutyarova, Q. Sun, G. K.-L. Chan, and G. Knizia, *Automated Construction of Molecular Active Spaces from Atomic Valence Orbitals*, J. Chem. Theory Comput. **13** (2017) 4063, DOI: 10.1021/acs.jctc.7b00128.

[196] M. Foscato and V. R. Jensen, *Automated in Silico Design of Homogeneous Catalysts*, ACS Catal. **10** (2020) 2354, DOI: 10.1021/acscatal.9b04952.

[197] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, *Singlet–Triplet Energy Gaps of Organic Biradicals and Polyacenes with Auxiliary-Field Quantum Monte Carlo*, J. Chem. Theory Comput. **15** (2019) 4924, DOI: 10.1021/acs.jctc.9b00534.

[198] M. Radoń, *Benchmarking quantum chemistry methods for spin-state energetics of iron complexes against quantitative experimental data*, Phys. Chem. Chem. Phys. **21** (9 2019) 4854, DOI: 10.1039/C9CP00105K.

[199] B. M. Flöser, Y. Guo, C. Riplinger, F. Tuczek, and F. Neese, *Detailed Pair Natural Orbital-Based Coupled Cluster Studies of Spin Crossover Energetics*, J. Chem. Theory Comput. **16** (2020) 2224, DOI: 10.1021/acs.jctc.9b01109.

[200] S. E. Neale, D. A. Pantazis, and S. A. Macgregor, *Accurate computed spin-state energetics for Co(III) complexes: implications for modelling homogeneous catalysis*, Dalton Trans. **49** (19 2020) 6478, DOI: 10.1039/D0DT00993H.

[201] G. Drabik, J. Szklarzewicz, and M. Radoń, *Spin-state energetics of metallocenes: How do best wave function and density functional theory results compare with the experimental data?* Phys. Chem. Chem. Phys. **23** (1 2021) 151, DOI: 10.1039/D0CP04727A.

[202] M. Reimann and M. Kaupp, *Spin-State Splittings in 3d Transition-Metal Complexes Revisited: Benchmarking Approximate Methods for Adiabatic Spin-State Energy Differences in Fe(II) Complexes*, J. Chem. Theory Comput. **18** (2022) 7442, DOI: 10.1021/acs.jctc.2c00924.

[203] M. Drosou, C. A. Mitsopoulou, and D. A. Pantazis, *Reconciling Local Coupled Cluster with Multireference Approaches for Transition Metal Spin-State Energetics*, J. Chem. Theory Comput. **18** (2022) 3538, DOI: 10.1021/acs.jctc.2c00265.

[204] U. Rastetter, A. Jacobi von Wangelin, and C. Herrmann, *Redox-active ligands as a challenge for electronic structure methods*, J. Comput. Chem. **44** (2023) 468, DOI: 10.1002/jcc.27013.

[205] K. P. Kepp, *Theoretical Study of Spin Crossover in 30 Iron Complexes*, Inorg. Chem. **55** (2016) 2717, DOI: 10.1021/acs.inorgchem.5b02371.

[206] T. F. Hughes and R. A. Friesner, *Correcting Systematic Errors in DFT Spin-Splitting Energetics for Transition Metal Complexes*, J. Chem. Theory Comput. **7** (2011) 19, DOI: 10.1021/ct100359x.

[207] J. P. Janet, L. Chan, and H. J. Kulik, *Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network*, J. Phys. Chem. Lett. **9** (2018) 1064, DOI: 10.1021/acs.jpclett.8b00170.

[208] M. Nurhuda, C. C. Perry, and M. A. Addicoat,
Performance of GFN1-xTB for periodic optimization of metal organic frameworks,
Phys. Chem. Chem. Phys. **24** (18 2022) 10906, DOI: 10.1039/D2CP00184E.

[209] T. Frauenheim, G. Seifert, M. Elsterner, Z. Hajnal, G. Jungnickel, D. Porezag, S. Suhai, and R. Scholz, *A Self-Consistent Charge Density-Functional Based Tight-Binding Method for Predictive Materials Simulations in Physics, Chemistry and Biology*,
Phys. Status Solidi B **217** (2000) 41, DOI:
10.1002/(SICI)1521-3951(200001)217:1%3C41::AID-PSSB41%3E3.0.CO;2-V.

[210] C. Köhler, G. Seifert, and T. Frauenheim,
Density functional based calculations for Fe_n (n≤32), Chem. Phys. **309** (2005) 23,
DOI: 10.1016/j.chemphys.2004.03.034.

[211] C. Köhler, T. Frauenheim, B. Hourahine, G. Seifert, and M. Sternberg, *Treatment of Collinear and Noncollinear Electron Spin within an Approximate Density Functional Based Method*,
J. Phys. Chem. A **111** (2007) 5622, DOI: 10.1021/jp068802p.

[212] Light-weight tight-binding framework tblite (https://github.com/tblite/tblite). For help getting started, visit https://tblite.readthedocs.io/en/latest/.

[213] Semiempirical Extended Tight-Binding Program Package xtb
(https://github.com/grimme-lab/xtb). For help getting started, visit
https://xtb-docs.readthedocs.io/en/latest/contents.html.

[214] G. Jha and T. Heine, *Spin–orbit coupling corrections for the GFN-xTB method*,
J. Chem. Phys. **158** (2023) 044120, DOI: 10.1063/5.0129071.

[215] J. Řezáč and P. Hobza, *Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods*, J. Chem. Theory Comput. **8** (2012) 141,
DOI: 10.1021/ct200751e.

[216] *TURBOMOLE V7.5 2020, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from https://www.turbomole.org.*

[217] S. G. Balasubramani et al., *TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations*, J. Chem. Phys. **152** (2020) 184107,
DOI: 10.1063/5.0004635.

[218] J. G. Brandenburg, C. Bannwarth, A. Hansen, and S. Grimme,
B97-3c: A revised low-cost variant of the B97-D density functional method,
J. Chem. Phys. **148** (2018) 064104, DOI: 10.1063/1.5012601.

[219] Generally Applicable Atomic-Charge Dependent London Dispersion Correction
(https://github.com/dftd4/dftd4).

[220] B. Helmich-Paris, *A trust-region augmented Hessian implementation for restricted and unrestricted Hartree–Fock and Kohn–Sham methods*, J. Chem. Phys. **154** (2021) 164104,
DOI: 10.1063/5.0040798.

[221] F. Neese, *An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix*, J. Comput. Chem. **24** (2003) 1740, DOI: 10.1002/jcc.10318.

[222] F. Weigend, *Accurate Coulomb-fitting basis sets for H to Rn*, Phys. Chem. Chem. Phys. **8** (9 2006) 1057, DOI: 10.1039/B515623H.

[223] F. Neese, F. Wennmohs, A. Hansen, and U. Becker, *Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange*, Chem. Phys. **356** (1-3 2009) 98, DOI: 10.1016/2008.10.036.

[224] R. Izsák and F. Neese, *An overlap fitted chain of spheres exchange method*, J. Chem. Phys. **135** (2011) 144105, DOI: 10.1063/1.3646921.

[225] B. Helmich-Paris, B. de Souza, F. Neese, and R. Izsák, *An improved chain of spheres for exchange algorithm*, J. Chem. Phys. **155** (2021) 104109, DOI: 10.1063/5.0058766.

[226] J. J. P. Stewart, *MOPAC2016, Version: 19.179L, Computational Chemistry, Colorado Springs*, Accessed: 15-02-2023.

[227] N. D. Mermin, *Thermal Properties of the Inhomogeneous Electron Gas*, Phys. Rev. **137** (5A 1965) A1441, DOI: 10.1103/PhysRev.137.A1441.

[228] B. Hourahine et al., *DFTB+, a software package for efficient approximate density functional theory based atomistic simulations*, J. Chem. Phys. **152** (2020) 124101, DOI: 10.1063/1.5143190.

[229] A. Nandy, D. B. K. Chu, D. R. Harper, C. Duan, N. Arunachalam, Y. Cytter, and H. J. Kulik, *Large-scale comparison of 3d and 4d transition metal complexes illuminates the reduced effect of exchange on second-row spin-state energetics*, Phys. Chem. Chem. Phys. **22** (34 2020) 19326, DOI: 10.1039/D0CP02977G.

[230] P. Pracht and C. Bannwarth, *Fast Screening of Minimum Energy Crossing Points with Semiempirical Tight-Binding Methods*, J. Chem. Theory Comput. **18** (2022) 6370, DOI: 10.1021/acs.jctc.2c00578.

[231] É. Brémond, Á. J. Pérez-Jiménez, J. C. SancJCTCho-García, and C. Adamo, *Range-separated hybrid density functionals made simple*, J. Chem. Phys. **150** (2019) 201102, DOI: 10.1063/1.5097164.

[232] B. Chan, *Assessment and development of DFT with the expanded CUAGAU-2 set of group-11 cluster systems*, Int. J. Quantum Chem. **121** (2021) e26453, DOI: 10.1002/qua.26453.

[233] K. Hui and J.-D. Chai, *SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters*, J. Chem. Phys. **144** (2016) 044114, DOI: 10.1063/1.4940734.

[234] G. Santra and J. M. L. Martin, *What Types of Chemical Problems Benefit from Density-Corrected DFT? A Probe Using an Extensive and Chemically Diverse Test Suite*, J. Chem. Theory Comput. **17** (2021) 1368, DOI: 10.1021/acs.jctc.0c01055.

[235] S. Spicher, E. Caldeweyher, A. Hansen, and S. Grimme, *Benchmarking London dispersion corrected density functional theory for noncovalent ion-π interactions*, Phys. Chem. Chem. Phys. **23** (2021) 11635, DOI: 10.1039/d1cp01333e.

[236] D. J. Liptrot and P. P. Power,
London dispersion forces in sterically crowded inorganic and organometallic molecules,
Nat. Rev. Chem. **1** (2017) 0004, DOI: 10.1038/s41570-016-0004.

[237] G. Santra and J. M. Martin,
Pure and Hybrid SCAN, rSCAN, and r^2SCAN: Which One Is Preferred in KS-and HF-DFT Calculations, and How Does D4 Dispersion Correction Affect This Ranking?
Molecules **27** (2022) 141, DOI: 10.3390/molecules27010141.

[238] ORCA – an ab initio, density functional and semiempirical program package, V. 5.0.1, F. Neese, MPI für Kohlenforschung, Mülheim a. d. Ruhr (Germany), **2021**.

[239] O. Vahtras, J. Almlöf, and M. W. Feyereisen,
Integral approximations for LCAO-SCF calculations, Chem. Phys. Lett. **213** (1993) 514,
DOI: 10.1016/0009-2614(93)89151-7.

[240] R. A. Kendall and H. A. Früchtl, *The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development*,
Theor. Chem. Acc. **97** (1997) 158, DOI: 10.1007/s002140050249.

[241] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, *Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials*,
Theor. Chem. Acc. **97** (1997) 119, DOI: 10.1007/s002140050244.

[242] D. Andrae, U. Häußermann, M. Dolg, H. Stoll, and H. Preuß, *Energy-Adjusted Ab Initio Pseudopotentials for the Second and Third Row Transition Elements*,
Theor. Chim. Acta **77** (1990) 123, DOI: 10.1007/BF01114537.

[243] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg,
Systematically Convergent Basis Sets with Relativistic Pseudopotentials. II. Small-Core Pseudopotentials and Correlation Consistent Basis Sets for the Post-d Group 16–18 Elements,
J. Chem. Phys. **119** (2003) 11113, DOI: 10.1063/1.1622924.

[244] M. K. Kesharwani, D. Manna, N. Sylvetsky, and J. M. Martin, *The X40×10 Halogen Bonding Benchmark Revisited: Surprising Importance of (n-1)d Subvalence Correlation*,
J. Phys. Chem. A **122** (2018) 2184, DOI: 10.1021/acs.jpca.7b10958.

[245] J. Řezáč, *Non-Covalent Interactions Atlas Benchmark Data Sets 2: Hydrogen Bonding in an Extended Chemical Space*, J. Chem. Theory Comput. **16** (2020) 6305,
DOI: 10.1021/acs.jctc.0c00715.

[246] N. Mehta, T. Fellowes, J. M. White, and L. Goerigk, *CHAL336 Benchmark Set: How Well Do Quantum-Chemical Methods Describe Chalcogen-Bonding Interactions?*
J. Chem. Theory Comput. **17** (2021) 2783, DOI: 10.1021/acs.jctc.1c00006.

[247] V. M. Miriyala and J. Řezáč, *Testing Semiempirical Quantum Mechanical Methods on a Data Set of Interaction Energies Mapping Repulsive Contacts in Organic Molecules*,
J. Phys. Chem. A **122** (2018) 2801, DOI: 10.1021/acs.jpca.8b00260.

[248] V. M. Miriyala and J. Řezáč, *Correction to: Testing semiempirical QM methods on a data set of interaction energies mapping repulsive contacts in organic molecules (Journal of Physical Chemistry A (2018) 122 :10 (2801–2808) DOI: 10.1021/acs.jpca.8b00260)*,
J. Phys. Chem. A **122** (2018) 9585, DOI: 10.1021/acs.jpca.8b11266.

[249] T. Husch, L. Freitag, and M. Reiher,
Calculation of Ligand Dissociation Energies in Large Transition-Metal Complexes,
J. Chem. Theory Comput. **14** (2018) 2456, DOI: 10.1021/acs.jctc.8b00061.

[250] T. Husch, L. Freitag, and M. Reiher, *Correction to: 'Calculation of ligand dissociation energies in large transition-metal complexes' (J. Chem. Theory Comput. (2018) 14:5 (2456–2468) DOI: 10.1021/acs.jctc.8b00061)*, J. Chem. Theory Comput. **15** (2019) 4295,
DOI: 10.1021/acs.jctc.9b00502.

[251] Y. Sun and H. Chen, *Performance of Density Functionals for Activation Energies of Re-Catalyzed Organic Reactions*, J. Chem. Theory Comput. **10** (2014) 579,
DOI: 10.1021/CT4010855.

[252] Y. Sun and H. Chen,
Performance of density functionals for activation energies of Zr-mediated reactions,
J. Chem. Theory Comput. **9** (2013) 4735, DOI: 10.1021/ct400432x.

[253] Y. Sun, L. Hu, and H. Chen,
Comparative Assessment of DFT Performances in Ru- and Rh-Promoted σ-Bond Activations,
J. Chem. Theory Comput **11** (2015) 1428, DOI: 10.1021/CT5009119.

[254] L. Hu and H. Chen,
Assessment of DFT Methods for Computing Activation Energies of Mo/W-Mediated Reactions,
J. Chem. Theory Comput. **11** (2015) 4601, DOI: 10.1021/acs.jctc.5b00373.

[255] M. Blaško, L. F. Pašteka, and M. Urban, *DFT Functionals for Modeling of Polyethylene Chains Cross-Linked by Metal Atoms. DLPNO–CCSD(T) Benchmark Calculations*,
J. Phys. Chem. A **125** (2021) 7382, DOI: 10.1021/acs.jpca.1c04793.

[256] J. P. Perdew and A. Zunger,
Self-interaction correction to density-functional approximations for many-electron systems,
Phys. Rev. B **23** (1981) 5048, DOI: 10.1103/PhysRevB.23.5048.

[257] P. Mori-Sánchez, A. J. Cohen, and W. Yang,
Many-electron self-interaction error in approximate density functionals,
J. Chem. Phys. **125** (2006) 201102, DOI: 10.1063/1.2403848.

[258] J. L. Bao, L. Gagliardi, and D. G. Truhlar,
Self-Interaction Error in Density Functional Theory: An Appraisal,
J. Phys. Chem. Lett. **9** (2018) 2353, DOI: 10.1021/acs.jpclett.8b00242.

[259] T. Tsuneda and K. Hirao, *Self-interaction corrections in density functional theory*,
J. Chem. Phys. **140** (2014) 18A513, DOI: 10.1063/1.4866996.

[260] L. Goerigk and S. Grimme,
A general database for main group thermochemistry, kinetics, and noncovalent interactions - Assessment of common and reparameterized (meta-)GGA density functionals,
J. Chem. Theory Comput. **6** (2010) 107, DOI: 10.1021/ct900489g.

[261] M. Bühl and H. Kabrede,
Geometries of transition-metal complexes from density-functional theory,
J. Chem. Theory Comput. **2** (2006) 1282, DOI: 10.1021/ct6001187.

[262] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen, *Consistent structures and interactions by density functional theory with small atomic orbital basis sets*, J. Chem. Phys. **143** (2015) 054107, DOI: 10.1063/1.4927476.

[263] M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko, and V. Barone, *Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules*, J. Phys. Chem. A **119** (2015) 2058, DOI: 10.1021/jp511432m.

[264] É. Brémond, M. Savarese, N. Q. Su, Á. J. Pérez-Jiménez, X. Xu, J. C. Sancho-García, and C. Adamo, *Benchmarking Density Functionals on Structural Parameters of Small-/Medium-Sized Organic Molecules*, J. Chem. Theory Comput. **12** (2016) 459, DOI: 10.1021/acs.jctc.5b01144.

[265] C. Riplinger and F. Neese, *An efficient and near linear scaling pair natural orbital based local coupled cluster method*, J. Chem. Phys. **138** (2013) 034106, DOI: 10.1063/1.4773581.

[266] G. L. Stoychev, A. A. Auer, J. Gauss, and F. Neese, *DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings*, J. Chem. Phys. **154** (2021), 164110, DOI: 10.1063/5.0047125.

[267] N. Mehta and J. M. L. Martin, *Reduced-Scaling Double Hybrid Density Functional Theory with Rapid Basis Set Convergence through Localized Pair Natural Orbital F12*, J. Phys. Chem. Lett. **13** (2022) 9332, DOI: 10.1021/acs.jpclett.2c02620.

[268] ORCA – an ab initio, density functional and semiempirical program package, V. 5.0.4, F. Neese, MPI für Kohlenforschung, Mülheim a. d. Ruhr (Germany), **2023**.

[269] D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin, and F. Neese, *Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory*, J. Chem. Theory Comput. **11** (2015) 1525, DOI: 10.1021/ct501129s.

[270] A. Altun, S. Ghosh, C. Riplinger, F. Neese, and G. Bistoni, *Addressing the System-Size Dependence of the Local Approximation Error in Coupled-Cluster Calculations*, J. Phys. Chem. A **125** (2021) 9932, DOI: 10.1021/acs.jpca.1c09106.

[271] J. Pogrebetsky, A. Siklitskaya, and A. Kubas, *MP2-Based Correction Scheme to Approach the Limit of a Complete Pair Natural Orbitals Space in DLPNO-CCSD(T) Calculations*, J. Chem. Theory Comput. **19** (2023) 4023, DOI: 10.1021/acs.jctc.3c00444.

[272] M. Feyereisen, G. Fitzgerald, and A. Komornicki, *Use of approximate integrals in ab initio theory. An application in MP2 energy calculations*, Chem. Phys. Lett. **208** (1993) 359, DOI: 10.1016/0009-2614(93)87156-W.

[273] D. E. Bernholdt and R. J. Harrison, *Large-scale correlated electronic structure calculations: the RI-MP2 method on parallel computers*, Chem. Phys. Lett. **250** (1996) 477, DOI: 10.1016/0009-2614(96)00054-1.

[274] S. Ehlert, S. Grimme, and A. Hansen, *Conformational Energy Benchmark for Longer n-Alkane Chains*, J. Phys. Chem. A **126** (2022) 3521, DOI: 10.1021/acs.jpca.2c02439.

[275] V. K. Prasad, Z. Pei, S. Edelmann, A. Otero-De-La-Roza, and G. A. Dilabio, *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. **18** (2022) 151, DOI: 10.1021/acs.jctc.1c00694.

[276] V. K. Prasad, Z. Pei, S. Edelmann, A. Otero-De-La-Roza, and G. A. Dilabio, *Correction to "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. **18** (2022) 4041, DOI: 10.1021/acs.jctc.2c00362.

[277] G. Bistoni, A. A. Auer, and F. Neese, *Understanding the Role of Dispersion in Frustrated Lewis Pairs and Classical Lewis Adducts: A Domain-Based Local Pair Natural Orbital Coupled Cluster Study*, Chem. - A Eur. J. **23** (2017) 865, DOI: 10.1002/chem.201604127.

[278] R. Sure and S. Grimme, *Comprehensive Benchmark of Association (Free) Energies of Realistic Host–Guest Complexes*, J. Chem. Theory Comput. **11** (2015) 3785, DOI: 10.1021/acs.jctc.5b00296.

[279] J. Gorges, S. Grimme, and A. Hansen, *Reliable prediction of association (free) energies of supramolecular complexes with heavy main group elements - the HS13L benchmark set*, Phys. Chem. Chem. Phys. **24** (2022) 28831, DOI: 10.1039/d2cp04049b.

[280] B. Rudshteyn, J. L. Weber, D. Coskun, P. A. Devlaminck, S. Zhang, D. R. Reichman, J. Shee, and R. A. Friesner, *Calculation of Metallocene Ionization Potentials via Auxiliary Field Quantum Monte Carlo: Toward Benchmark Quantum Chemistry for Transition Metals*, J. Chem. Theory Comput. **18** (2022) 2845, DOI: 10.1021/acs.jctc.1c01071.

[281] T. Risthaus, M. Steinmetz, and S. Grimme, *Implementation of nuclear gradients of range-separated hybrid density functionals and benchmarking on rotational constants for organic molecules*, J. Comput. Chem. **35** (2014) 1509, DOI: 10.1002/jcc.23649.

[282] H.-J. Werner and A. Hansen, *Accurate Calculation of Isomerization and Conformational Energies of Larger Molecules Using Explicitly Correlated Local Coupled Cluster Methods in Molpro and ORCA*, J. Chem. Theory Comput. (2023), *accepted*, DOI: 10.1021/acs.jctc.3c00270, DOI: 10.1021/acs.jctc.3c00270.

[283] R. Francke and R. D. Little, *Redox catalysis in organic electrosynthesis: basic principles and recent developments*, Chem. Soc. Rev. **43** (8 2014) 2492, DOI: 10.1039/C3CS60464K.

[284] J. C. Siu, N. Fu, and S. Lin, *Catalyzing Electrosynthesis: A Homogeneous Electrocatalytic Approach to Reaction Discovery*, Acc. Chem. Res. **53** (2020) 547, DOI: 10.1021/acs.accounts.9b00529.

[285] D. Süß, S. E. Huber, and A. Mauracher, *On the impact of multi-reference character of small transition metal compounds on their bond dissociation energies*, J. Chem. Phys. **152** (2020) 114104, DOI: 10.1063/1.5143495.

[286] C. Duan, D. B. K. Chu, A. Nandy, and H. J. Kulik,
Detection of multi-reference character imbalances enables a transfer learning approach for virtual high throughput screening with coupled cluster accuracy at DFT cost,
Chem. Sci. **13** (17 2022) 4962, DOI: 10.1039/D2SC00393G.

[287] A. Galstyan and E.-W. Knapp,
Accurate redox potentials of mononuclear iron, manganese, and nickel model complexes,
J. Comput. Chem. **30** (2009) 203, DOI: 10.1002/jcc.21029.

[288] L. E. Roy, E. Jakubikova, M. G. Guthrie, and E. R. Batista,
Calculation of One-Electron Redox Potentials Revisited. Is It Possible to Calculate Accurate Potentials with Density Functional Methods? J. Phys. Chem. A **113** (2009) 6745,
DOI: 10.1021/jp811388w.

[289] S. J. Konezny, M. D. Doherty, O. R. Luca, R. H. Crabtree, G. L. Soloveichik, and V. S. Batista,
Reduction of Systematic Uncertainty in DFT Redox Potentials of Transition-Metal Complexes,
J. Phys. Chem. C **116** (2012) 6349, DOI: 10.1021/jp300485t.

[290] D. Coskun, S. V. Jerome, and R. A. Friesner, *Evaluation of the Performance of the B3LYP, PBE0, and M06 DFT Functionals, and DBLOC-Corrected Versions, in the Calculation of Redox Potentials and Spin Splittings for Transition Metal Containing Systems*,
J. Chem. Theory Comput. **12** (2016) 1121, DOI: 10.1021/acs.jctc.5b00782.

[291] M. Isegawa, F. Neese, and D. A. Pantazis,
Ionization Energies and Aqueous Redox Potentials of Organic Molecules: Comparison of DFT, Correlated ab Initio Theory and Pair Natural Orbital Approaches,
J. Chem. Theory Comput. **12** (2016) 2272, DOI: 10.1021/acs.jctc.6b00252.

[292] S. Bhattacharjee, M. Isegawa, M. Garcia-Ratés, F. Neese, and D. A. Pantazis,
Ionization Energies and Redox Potentials of Hydrated Transition Metal Ions: Evaluation of Domain-Based Local Pair Natural Orbital Coupled Cluster Approaches,
J. Chem. Theory Comput. **18** (2022) 1619, DOI: 10.1021/acs.jctc.1c01267.

[293] R. M. Richard, M. S. Marshall, O. Dolgounitcheva, J. V. Ortiz, J.-L. Brédas, N. Marom, and C. D. Sherrill, *Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules I. Reference Data at the CCSD(T) Complete Basis Set Limit*,
J. Chem. Theory Comput. **12** (2016) 595, DOI: 10.1021/acs.jctc.5b00875.

[294] W. Jiang, N. J. DeYonker, J. J. Determan, and A. K. Wilson,
Toward Accurate Theoretical Thermochemistry of First Row Transition Metal Complexes,
J. Phys. Chem. A **116** (2012) 870, DOI: 10.1021/jp205710e.

[295] Z. Fang, M. Vasiliu, K. A. Peterson, and D. A. Dixon, *Prediction of Bond Dissociation Energies/Heats of Formation for Diatomic Transition Metal Compounds: CCSD(T) Works*,
J. Chem. Theory Comput. **13** (2017) 1057, DOI: 10.1021/acs.jctc.6b00971.

[296] L. Cheng, J. Gauss, B. Ruscic, P. B. Armentrout, and J. F. Stanton,
Bond Dissociation Energies for Diatomic Molecules Containing 3d Transition Metals: Benchmark Scalar-Relativistic Coupled-Cluster Calculations for 20 Molecules,
J. Chem. Theory Comput. **13** (2017) 1044, DOI: 10.1021/acs.jctc.6b00970.

[297] Y. A. Aoto, A. P. de Lima Batista, A. Köhn, and A. G. S. de Oliveira-Filho, *How To Arrive at Accurate Benchmark Values for Transition Metal Compounds: Computation or Experiment?* J. Chem. Theory Comput. **13** (2017) 5291, DOI: 10.1021/acs.jctc.7b00688.

[298] J. Shee, B. Rudshteyn, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, *On Achieving High Accuracy in Quantum Chemical Calculations of 3d Transition Metal-Containing Systems: A Comparison of Auxiliary-Field Quantum Monte Carlo with Coupled Cluster, Density Functional Theory, and Experiment for Diatomic Molecules*, J. Chem. Theory Comput. **15** (2019) 2346, DOI: 10.1021/acs.jctc.9b00083.

[299] D. Hait, N. M. Tubman, D. S. Levine, K. B. Whaley, and M. Head-Gordon, *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. **15** (2019) 5370, DOI: 10.1021/acs.jctc.9b00674.

[300] L. W. Bertels, J. Lee, and M. Head-Gordon, *Polishing the Gold Standard: The Role of Orbital Choice in CCSD(T) Vibrational Frequency Prediction*, J. Chem. Theory Comput. **17** (2021) 742, DOI: 10.1021/acs.jctc.0c00746.

[301] *Wappett D, Goerigk L. Benchmarking Density Functional Theory Methods for Metalloenzyme Reactions: The Introduction of the MME55 Set. ChemRxiv. Cambridge: Cambridge Open Engage; 2023; This content is a preprint and has not been peer-reviewed.*

[302] E. Semidalas and J. M. Martin, *The MOBH35 Metal–Organic Barrier Heights Reconsidered: Performance of Local-Orbital Coupled Cluster Approaches in Different Static Correlation Regimes*, J. Chem. Theory Comput. **18** (2022) 883, DOI: 10.1021/acs.jctc.1c01126.

[303] H. M. Aðalsteinsson and R. Bjornsson, *Ionization energies of metallocenes: A coupled cluster study of cobaltocene*, Phys. Chem. Chem. Phys. (2023) 4570, DOI: 10.1039/D2CP04715B.

[304] S. Zhang and H. Krakauer, *Quantum Monte Carlo Method using Phase-Free Random Walks with Slater Determinants*, Phys. Rev. Lett. **90** (13 2003) 136401, DOI: 10.1103/PhysRevLett.90.136401.

[305] W. A. Al-Saidi, S. Zhang, and H. Krakauer, *Auxiliary-field quantum Monte Carlo calculations of molecular systems with a Gaussian basis*, J. Chem. Phys. **124** (2006) 224101, DOI: 10.1063/1.2200885.

[306] F. D. Malone, S. Zhang, and M. A. Morales, *Accelerating Auxiliary-Field Quantum Monte Carlo Simulations of Solids with Graphical Processing Units*, J. Chem. Theory Comput. **16** (2020) 4286, DOI: 10.1021/acs.jctc.0c00262.

[307] B. Rudshteyn, D. Coskun, J. L. Weber, E. J. Arthur, S. Zhang, D. R. Reichman, R. A. Friesner, and J. Shee, *Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo*, J. Chem. Theory Comput. **16** (2020) 3041, DOI: 10.1021/acs.jctc.0c00070.

[308] M. Reimann and M. Kaupp, *Spin-State Splittings in 3d Transition-Metal Complexes Revisited: Toward a Reliable Theory Benchmark*, J. Chem. Theory Comput. **19** (2023) 97, DOI: 10.1021/acs.jctc.2c00925.

[309] H. Verplancke, M. Diefenbach, J. N. Lienert, M. Ugandi, M.-P. Kitsaras, M. Roemelt, S. Stopkowicz, and M. C. Holthausen,
Another Torture Track for Quantum Chemistry: Reinvestigation of the Benzaldehyde Amidation by Nitrogen-Atom Transfer from Platinum(II) and Palladium(II) Metallonitrenes,
Isr. J. Chem. **63** (2023) e202300060, DOI: 10.1002/ijch.202300060.

[310] K. A. Grice, C. Saucedo, M. A. Sovereign, and A. P. Cho,
The Electrochemical Behavior of Early Metal Metallocene Cp_2MCl_2 Complexes under CO_2,
Electrochim. Acta **218** (2016) 110, DOI: 10.1016/j.electacta.2016.09.095.

[311] M. Kessler, S. Hansen, D. Hollmann, M. Klahn, T. Beweries, A. Spannenberg, A. Brückner, and U. Rosenthal, *Synthesis of $Cp^*_2Ti(OTf)$ and Its Reaction with Water*,
Eur. J. Inorg. Chem **2011** (2011) 627, DOI: 10.1002/ejic.201001225.

[312] R. R. Gagne, C. A. Koval, and G. C. Lisensky,
Ferrocene as an internal standard for electrochemical measurements,
Inorg. Chem. **19** (1980) 2854, DOI: 10.1021/ic50211a080.

[313] A. M. Bond, K. B. Oldham, and G. A. Snook,
Use of the Ferrocene Oxidation Process To Provide Both Reference Electrode Potential Calibration and a Simple Measurement (via Semiintegration) of the Uncompensated Resistance in Cyclic Voltammetric Studies in High-Resistance Organic Solvents,
Anal. Chem. **72** (2000) 3492, DOI: 10.1021/ac000020j.

[314] M. Namazian, C. Y. Lin, and M. L. Coote, *Benchmark Calculations of Absolute Reduction Potential of Ferricinium/Ferrocene Couple in Nonaqueous Solutions*,
J. Chem. Theory Comput. **6** (2010) 2721, DOI: 10.1021/ct1003252.

[315] D. Astruc, *Why is Ferrocene so Exceptional?* Eur. J. Inorg. Chem. **2017** (2017) 6,
DOI: 10.1002/ejic.201600983.

[316] S. Sethi, P. K. Das, and N. Behera, *The chemistry of aminoferrocene, $Fe(\eta^5\text{-}C_5H_4NH_2)(\eta^5\text{-}Cp)$: Synthesis, reactivity and applications*,
J. Organomet. Chem. **824** (2016) 140, DOI: 10.1016/j.jorganchem.2016.10.014.

[317] C. A. P. Goodwin, M. J. Giansiracusa, S. M. Greer, H. M. Nicholas, P. Evans, M. Vonci, S. Hill, N. F. Chilton, and D. P. Mills, *Isolation and electronic structures of derivatized manganocene, ferrocene and cobaltocene anions*, Nat. Chem. **13** (2021) 243,
DOI: 10.1038/s41557-020-00595-w.

[318] M. G. Walawalkar, P. Pandey, and R. Murugavel,
The Redox Journey of Iconic Ferrocene: Ferrocenium Dications and Ferrocenate Anions,
Angew. Chem. Int. Ed. **60** (2021) 12632, DOI: 10.1002/anie.202101770.

[319] V. Artero and M. Fontecave, *Hydrogen evolution catalyzed by $CpFe(CO)_2$-based complexes*,
C. R. Chim. **11** (2008) 926, DOI: 10.1016/j.crci.2008.03.006.

[320] K. A. Pevear, M. M. B. Holl, G. B. Carpenter, A. L. Rieger, P. H. Rieger, and D. A. Sweigart, *Ligand Substitution at 19-Electron Centers and the Indenyl Effect in Organometallic Radicals. Electrocatalytic CO Substitution in (cyclopentadienyl)$Fe(CO)_3^+$ and (indenyl)$Fe(CO)_3^+$*,
Organometallics **14** (1995) 512, DOI: 10.1021/om00001a070.

[321] A. Rosas-Hernández, H. Junge, M. Beller, M. Roemelt, and R. Francke, *Cyclopentadienone iron complexes as efficient and selective catalysts for the electroreduction of CO_2 to CO*, Catal. Sci. Technol. **7** (2 2017) 459, DOI: 10.1039/C6CY02352E.

[322] M. Fang, E. S. Wiedner, W. G. Dougherty, W. S. Kassel, T. Liu, D. L. DuBois, and R. M. Bullock, *Cobalt Complexes Containing Pendant Amines in the Second Coordination Sphere as Electrocatalysts for H_2 Production*, Organometallics **33** (2014) 5820, DOI: 10.1021/om5004607.

[323] C. A. Craig, L. O. Spreer, J. W. Otvos, and M. Calvin, *Photochemical reduction of carbon dioxide using nickel tetraazamacrocycles*, J. Phys. Chem. **94** (1990) 7957, DOI: 10.1021/j100383a038.

[324] M. L. Helm, M. P. Stewart, R. M. Bullock, M. R. DuBois, and D. L. DuBois, *A Synthetic Nickel Electrocatalyst with a Turnover Frequency Above 100,000 s^{-1} for H_2 Production*, Science **333** (2011) 863, DOI: 10.1126/science.1205864.

[325] J. Tory, B. Setterfield-Price, R. A. W. Dryfe, and F. Hartl, *[M(CO)$_4$(2,2'-bipyridine)] (M=Cr, Mo, W) Complexes as Efficient Catalysts for Electrochemical Reduction of CO_2 at a Gold Electrode*, ChemElectroChem **2** (2015) 213, DOI: 10.1002/celc.201402282.

[326] C. W. Machan, C. J. Stanton, J. E. Vandezande, G. F. Majetich, H. F. Schaefer, C. P. Kubiak, and J. Agarwal, *Electrocatalytic Reduction of Carbon Dioxide by Mn(CN)(2,2'-bipyridine)(CO)$_3$: CN Coordination Alters Mechanism*, Inorg. Chem. **54** (2015) 8849, DOI: 10.1021/acs.inorgchem.5b01715.

[327] P. G. Alsabeh, A. Rosas-Hernández, E. Barsch, H. Junge, R. Ludwig, and M. Beller, *Iron-catalyzed photoreduction of carbon dioxide to synthesis gas*, Catal. Sci. Technol. **6** (2016) 3623, DOI: 10.1039/C5CY01129A.

[328] L. Garnier, Y. Rollin, and J. Périchon, *Electrosynthesis of symmetrical ketones from organic halides and carbon dioxide catalysed by 2,2'-bipyridine-nickel complexes*, J. Organomet. Chem. **367** (1989) 347, DOI: 10.1016/0022-328X(89)87056-1.

[329] M. Razavet, V. Artero, and M. Fontecave, *Proton Electroreduction Catalyzed by Cobaloximes: Functional Models for Hydrogenases*, Inorg. Chem. **44** (2005) 4786, DOI: 10.1021/ic050167z.

[330] P. Du, J. Schneider, G. Luo, W. W. Brennessel, and R. Eisenberg, *Visible Light-Driven Hydrogen Production from Aqueous Protons Catalyzed by Molecular Cobaloxime Catalysts*, Inorg. Chem. **48** (2009) 4952, DOI: 10.1021/ic900389z.

[331] J. L. Dempsey, B. S. Brunschwig, J. R. Winkler, and H. B. Gray, *Hydrogen Evolution Catalyzed by Cobaloximes*, Acc. Chem. Res. **42** (2009) 1995, DOI: 10.1021/ar900253e.

[332] D. Dolui, S. Khandelwal, P. Majumder, and A. Dutta, *The odyssey of cobaloximes for catalytic H_2 production and their recent revival with enzyme-inspired design*, Chem. Commun. **56** (59 2020) 8166, DOI: 10.1039/D0CC03103H.

[333] S. Kuriyama, K. Arashiba, K. Nakajima, Y. Matsuo, H. Tanaka, K. Ishii, K. Yoshizawa, and Y. Nishibayashi, *Catalytic transformation of dinitrogen into ammonia and hydrazine by iron-dinitrogen complexes bearing pincer ligand*, Nat. Commun. **7** (2016) 12181, DOI: 10.1038/ncomms12181.

[334] M. Beyler, S. Ezzaher, M. Karnahl, M.-P. Santoni, R. Lomoth, and S. Ott, *Pentacoordinate iron complexes as functional models of the distal iron in [FeFe] hydrogenases*, Chem. Commun. **47** (42 2011) 11662, DOI: 10.1039/C1CC14449A.

[335] R. Lindroth, A. Ondrejková, and C.-J. Wallentin, *Visible-Light Mediated Oxidative Fragmentation of Ethers and Acetals by Means of Fe(III) Catalysis*, Organic Letters **24** (2022) 1662, DOI: 10.1021/acs.orglett.2c00231.

[336] G. A. N. Felton, A. K. Vannucci, J. Chen, L. T. Lockett, N. Okumura, B. J. Petro, U. I. Zakai, D. H. Evans, R. S. Glass, and D. L. Lichtenberger, *Hydrogen Generation from Weak Acids: Electrochemical and Computational Studies of a Diiron Hydrogenase Mimic*, J. Am. Chem. Soc. **129** (2007) 12521, DOI: 10.1021/ja073886g.

[337] W. R. McNamara, Z. Han, P. J. Alperin, W. W. Brennessel, P. L. Holland, and R. Eisenberg, *A Cobalt–Dithiolene Complex for the Photocatalytic and Electrocatalytic Reduction of Protons*, J. Am. Chem. Soc. **133** (2011) 15368, DOI: 10.1021/ja207842r.

[338] A. Das, Z. Han, W. W. Brennessel, P. L. Holland, and R. Eisenberg, *Nickel Complexes for Robust Light-Driven and Electrocatalytic Hydrogen Production from Water*, ACS Catal. **5** (2015) 1397, DOI: 10.1021/acscatal.5b00045.

[339] K. Ray, S. DeBeer George, E. I. Solomon, K. Wieghardt, and F. Neese, *Description of the Ground-State Covalencies of the Bis(dithiolato) Transition-Metal Complexes from X-ray Absorption Spectroscopy and Time-Dependent Density-Functional Calculations*, Chem. Eur. J. **13** (2007) 2783, DOI: 10.1002/chem.200601425.

[340] A. J. Bard, E. Garcia, S. Kukharenko, and V. V. Strelets, *Electrochemistry of metallocenes at very negative and very positive potentials. Electrogeneration of 17-electron Cp_2Co^{2+}, 21-electron Cp_2Co^{2-}, and 22-electron Cp_2Ni^{2-} species*, Inorg. Chem. **32** (1993) 3528, DOI: 10.1021/ic00068a024.

[341] A. B. P. Lever, *Electrochemical parametrization of metal complex redox potentials, using the ruthenium(III)/ruthenium(II) couple to generate a ligand electrochemical series*, Inorg. Chem. **29** (1990) 1271, DOI: 10.1021/ic00331a030.

[342] *TURBOMOLE V7.5.1 2021, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from https://www.turbomole.org.*

[343] S. G. Balasubramani et al., *TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations*, J. Chem. Phys. **152** (2020) 184107, DOI: 10.1063/5.0004635.

[344] Y. Zhao and D. G. Truhlar, *A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions*, J. Chem. Phys. **125** (2006) 194101, DOI: 10.1063/1.2370993.

[345] A. D. Becke, *Density-functional thermochemistry. III. The role of exact exchange*, J. Chem. Phys. **98** (1993) 5648, DOI: 10.1063/1.464913.

[346] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields*, J. Phys. Chem. **98** (1994) 11623, DOI: 10.1021/j100096a001.

[347] Y. Zhao and D. G. Truhlar, *Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions*, J. Phys. Chem. A **109** (2005) 5656, DOI: 10.1021/jp050536c.

[348] Y. Zhao and D. G. Truhlar, *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. **120** (2008) 215, DOI: 10.1007/s00214-007-0401-8.

[349] L. Goerigk and S. Grimme, *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. **7** (2011) 291, DOI: 10.1021/ct100466k.

[350] M. Haasler, T. M. Maier, R. Grotjahn, S. Gückel, A. V. Arbuznikov, and M. Kaupp, *A Local Hybrid Functional with Wide Applicability and Good Balance between (De)Localization and Left–Right Correlation*, J. Chem. Theory Comput. **16** (2020) 5645, DOI: 10.1021/acs.jctc.0c00498.

[351] J. Kong and E. Proynov, *Density Functional Model for Nondynamic and Strong Correlation*, J. Chem. Theory Comput. **12** (2016) 133, DOI: 10.1021/acs.jctc.5b00801.

[352] *The xTron code has been developed by Jing Kong's group at Middle Tennessee State University and is available on Github https://github.com/jingkongmtsu/xTron.mKP16.*

[353] E. Epifanovsky et al., *Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package*, J. Chem. Phys. **155** (2021), 084801, DOI: 10.1063/5.0055522.

[354] T. van Voorhis and M. Head-Gordon, *A geometric approach to direct minimization*, Mol. Phys. **100** (2002) 1713, DOI: 10.1080/00268970110103642.

[355] P. M. Gill, B. G. Johnson, and J. A. Pople, *A standard grid for density functional calculations*, Chem. Phys. Lett. **209** (1993) 506, DOI: 10.1016/0009-2614(93)80125-9.

[356] G. Li Manni, D. Kats, D. P. Tew, and A. Alavi, *Role of Valence and Semicore Electron Correlation on Spin Gaps in Fe(II)-Porphyrins*, J. Chem. Theory Comput. **15** (2019) 1492, DOI: 10.1021/acs.jctc.8b01277.

[357] D. L. Crittenden, *A Hierarchy of Static Correlation Models*, J. Phys. Chem. A **117** (2013) 3852, DOI: 10.1021/jp400669p.

[358] U. R. Fogueri, S. Kozuch, A. Karton, and J. M. L. Martin, *A simple DFT-based diagnostic for nondynamical correlation*, Theor. Chem. Acc. **132** (2012) 1291, DOI: 10.1007/s00214-012-1291-y.

[359] P. Piecuch and M. Wloch, *Renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed Hamiltonian*, J. Chem. Phys. **123** (2005), DOI: 10.1063/1.2137318.

[360] M. Wloch, J. R. Gour, and P. Piecuch, *Extension of the Renormalized Coupled-Cluster Methods Exploiting Left Eigenstates of the Similarity-Transformed Hamiltonian to Open-Shell Systems: A Benchmark Study*, J. Phys. Chem. A **111** (2007) 11359, DOI: 10.1021/jp072535l.

[361] G. M. J. Barca et al., *Recent developments in the general atomic and molecular electronic structure system*, J. Chem. Phys. **152** (2020) 154102, DOI: 10.1063/5.0005188.

[362] *A performant parallel implementation is available in GAMESS, according to the developers, but this did not run on the hardware available to us, which significantly limited the application of the CR-CC(2,3) method.*

[363] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, *Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory*, J. Chem. Theory Comput. **13** (2017) 1595, DOI: 10.1021/acs.jctc.6b01028.

[364] J. E. T. Smith, B. Mussard, A. A. Holmes, and S. Sharma, *Cheap and Near Exact CASSCF with Large Active Spaces*, J. Chem. Theory Comput. **13** (2017) 5468, DOI: 10.1021/acs.jctc.7b00900.

[365] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022.

[366] S. Grimme and A. Hansen, *A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects*, Angew. Chem. Int. Ed. **54** (2015) 12308, DOI: 10.1002/anie.201501887.

[367] E. Ramos-Cordoba, P. Salvador, and E. Matito, *Separation of dynamic and nondynamic correlation*, Phys. Chem. Chem. Phys. **18** (34 2016) 24015, DOI: 10.1039/C6CP03072F.

[368] E. Ramos-Cordoba and E. Matito, *Local Descriptors of Dynamic and Nondynamic Correlation*, J. Chem. Theory Comput. **13** (2017) 2705, DOI: 10.1021/acs.jctc.7b00293.

[369] M. K. Kesharwani, N. Sylvetsky, A. Köhn, D. P. Tew, and J. M. L. Martin, *Do CCSD and approximate CCSD-F12 variants converge to the same basis set limits? The case of atomization energies*, J. Chem. Phys. **149** (2018) 154109, DOI: 10.1063/1.5048665.

[370] T. J. Lee and P. R. Taylor, *A diagnostic for determining the quality of single-reference electron correlation methods*, Int. J. Quantum Chem. **36** (1989) 199, DOI: 10.1002/qua.560360824.

[371] J. Cadima, J. Cerdeira, and M. Minhoto, *Computational aspects of algorithms for variable selection in the context of principal components*, Comput. Stat. Data. Anal. **47** (2004) 225, DOI: 10.1016/j.csda.2003.11.001.

[372] K. Yamaguchi, F. Jensen, A. Dorigo, and K. Houk,
A spin correction procedure for unrestricted Hartree-Fock and Møller-Plesset wavefunctions for singlet diradicals and polyradicals, Chem. Phys. Lett. **149** (1988) 537, DOI: 10.1016/0009-2614(88)80378-6.

[373] T. Saito, N. Yasuda, Y. Kataoka, Y. Nakanishi, Y. Kitagawa, T. Kawakami, S. Yamanaka, M. Okumura, and K. Yamaguchi, *Potential Energy Curve for Ring-Opening Reactions: Comparison Between Broken-Symmetry and Multireference Coupled Cluster Methods*, J. Phys. Chem. A **115** (2011) 5625, DOI: 10.1021/jp201463h.

[374] J. T. Margraf, A. Perera, J. J. Lutz, and R. J. Bartlett,
Single-reference coupled cluster theory for multi-reference problems, J. Chem. Phys. **147** (2017), 184101, DOI: 10.1063/1.5003128.

[375] H. Schurkus, D.-T. Chen, H.-P. Cheng, G. Chan, and J. Stanton, *Theoretical prediction of magnetic exchange coupling constants from broken-symmetry coupled cluster calculations*, J. Chem. Phys. **152** (2020), 234115, DOI: 10.1063/1.5144696.

[376] A. Wodyński and M. Kaupp,
Local Hybrid Functional Applicable to Weakly and Strongly Correlated Systems, J. Chem. Theory Comput. **18** (2022) 6111, DOI: 10.1021/acs.jctc.2c00795.

[377] G. Santra and J. M. L. Martin, *Do Double-Hybrid Functionals Benefit from Regularization in the PT2 Term? Observations from an Extensive Benchmark*, J. Phys. Chem. Lett. **13** (2022) 3499, DOI: 10.1021/acs.jpclett.2c00718.

[378] W. Lubitz, H. Ogata, O. Rüdiger, and E. Reijerse, *Hydrogenases*, Chem. Rev. **114** (2014) 4081, DOI: 10.1021/cr4005814.

[379] R. M. Evans, E. J. Brooke, S. A. M. Wehlin, E. Nomerotskaia, F. Sargent, S. B. Carr, S. E. V. Phillips, and F. A. Armstrong,
Mechanism of hydrogen activation by [NiFe] hydrogenases, Nat. Chem. Biol. **12** (2016) 46, DOI: 10.1038/nchembio.1976.

[380] M. D. Wodrich and X. Hu, *Natural inspirations for metal-ligand cooperative catalysis*, Nat. Rev. Chem. **2** (2018) 0099, DOI: 10.1038/s41570-017-0099.

[381] J. Y. Yang, R. M. Bullock, M. R. DuBois, and D. L. DuBois, *Fast and efficient molecular electrocatalysts for H_2 production: Using hydrogenase enzymes as guides*, MRS Bull. **36** (2011) 39, DOI: 10.1557/mrs.2010.8.

[382] M. O'Hagan, M.-H. Ho, J. Y. Yang, A. M. Appel, M. R. DuBois, S. Raugei, W. J. Shaw, D. L. DuBois, and R. M. Bullock, *Proton Delivery and Removal in $[Ni(P^R{}_2N^{R'}{}_2)_2]^{2+}$ Hydrogen Production and Oxidation Catalysts*, J. Am. Chem. Soc. **134** (2012) 19409, DOI: 10.1021/ja307413x.

[383] T. B. Rauchfuss, *Diiron Azadithiolates as Models for the [FeFe]-Hydrogenase Active Site and Paradigm for the Role of the Second Coordination Sphere*, Acc. Chem. Res. **48** (2015) 2107, DOI: 10.1021/acs.accounts.5b00177.

[384] D. Schilter, J. M. Camara, M. T. Huynh, S. Hammes-Schiffer, and T. B. Rauchfuss, *Hydrogenase Enzymes and Their Synthetic Models: The Role of Metal Hydrides*, Chem. Rev. **116** (2016) 8693, DOI: 10.1021/acs.chemrev.6b00180.

[385] J. Y. Yang, S. E. Smith, T. Liu, W. G. Dougherty, W. A. Hoffert, W. S. Kassel, M. R. DuBois, D. L. DuBois, and R. M. Bullock,
Two Pathways for Electrocatalytic Oxidation of Hydrogen by a Nickel Bis(diphosphine) Complex with Pendant Amines in the Second Coordination Sphere,
J. Am. Chem. Soc. **135** (2013) 9700, DOI: 10.1021/ja400705a.

[386] D. Dolui, S. Khandelwal, A. Shaik, D. Gaat, V. Thiruvenkatam, and A. Dutta,
Enzyme-Inspired Synthetic Proton Relays Generate Fast and Acid-Stable Cobalt-Based H_2 Production Electrocatalysts, ACS Catal. **9** (2019) 10115,
DOI: 10.1021/acscatal.9b02953.

[387] J. W. Slater, S. C. Marguet, M. E. Gray, H. A. Monaco, M. Sotomayor, and H. S. Shafaat,
Power of the Secondary Sphere: Modulating Hydrogenase Activity in Nickel-Substituted Rubredoxin, ACS Catal. **9** (2019) 8928, DOI: 10.1021/acscatal.9b01720.

[388] P. Ghosh, S. de Vos, M. Lutz, F. Gloaguen, P. Schollhammer, M.-E. Moret, and R. J. M. Klein Gebbink,
Electrocatalytic Proton Reduction by a Cobalt Complex Containing a Proton-Responsive Bis(alkylimdazole)methane Ligand: Involvement of a C-H Bond in H_2 Formation,
Chem. - Eur. J. **26** (2020) 12560, DOI: 10.1002/chem.201905746.

[389] S. I. Johnson, H. B. Gray, J. D. Blakemore, and W. A. Goddard, *Role of Ligand Protonation in Dihydrogen Evolution from a Pentamethylcyclopentadienyl Rhodium Catalyst*,
Inorg. Chem. **56** (2017) 11375, DOI: 10.1021/acs.inorgchem.7b01698.

[390] M. G. Alférez, J. J. Moreno, N. Hidalgo, and J. Campos, *Reversible Hydride Migration from C_5Me_5 to Rh^I Revealed by a Cooperative Bimetallic Approach*,
Angew. Chem. **132** (2020) 21049, DOI: 10.1002/anie.202008442.

[391] J. J. Moreno, M. F. Espada, J. Campos, J. López-Serrano, S. A. Macgregor, and E. Carmona, *Base-Promoted, Remote C-H Activation at a Cationic (η^5-C_5Me_5)Ir(III) Center Involving Reversible C-C Bond Formation of Bound C_5Me_5*, J. Am. Chem. Soc. **141** (2019) 2205,
DOI: 10.1021/jacs.8b11752.

[392] D. A. Kurtz, D. Dhar, N. Elgrishi, B. Kandemir, S. F. McWilliams, W. C. Howland, C.-H. Chen, and J. L. Dempsey, *Redox-Induced Structural Reorganization Dictates Kinetics of Cobalt(III) Hydride Formation via Proton-Coupled Electron Transfer*,
J. Am. Chem. Soc. **143** (2021) 3393, DOI: 10.1021/jacs.0c11992.

[393] W. D. Jones, V. L. Kuykendall, and A. D. Selmeczy, *Ring migration reactions of (C_5Me_5)Rh(PMe_3)H_2. Evidence for η^3 slippage and metal-to-ring hydride migration*,
Organometallics **10** (1991) 1577, DOI: 10.1021/om00051a057.

[394] W. D. Jones, G. P. Rosini, and J. A. Maguire, *Photochemical C-H Activation and Ligand Exchange Reactions of CpRe(PPh_3)$_2H_2$. Phosphine Dissociation Is Not Involved*,
Organometallics **18** (1999) 1754, DOI: 10.1021/om9809872.

[395] C. E. Kefalidis, L. Perrin, C. J. Burns, D. J. Berg, L. Maron, and R. A. Andersen,
Can a pentamethylcyclopentadienyl ligand act as a proton-relay in f-element chemistry? Insights from a joint experimental/theoretical study, Dalton Trans. **44** (2015) 2575,
DOI: 10.1039/C4DT02387K.

[396] A. Davison, M. L. H. Green, and G. Wilkinson,
620. π-Cyclopentadienyl- and cyclopentadiene-iron carbonyl complexes,
J. Chem. Soc. **0** (1961) 3172, DOI: 10.1039/JR9610003172.

[397] D. J. Schild, M. W. Drover, P. H. Oyala, and J. C. Peters,
Generating Potent C-H PCET Donors: Ligand-Induced Fe-to-Ring Proton Migration from a CpFe^{III}-H Complex Demonstrates a Promising Strategy*,
J. Am. Chem. Soc. **142** (2020) 18963, DOI: 10.1021/jacs.0c09363.

[398] M. L. H. Green, L. Pratt, and G. Wilkinson,
760. A new type of transition metal-cyclopentadiene compound, J. Chem. Soc. **0** (1959) 3753,
DOI: 10.1039/JR9590003753.

[399] M. J. Chalkley, P. H. Oyala, and J. C. Peters,
Cp Noninnocence Leads to a Remarkably Weak C-H Bond via Metallocene Protonation*,
J. Am. Chem. Soc. **141** (2019) 4721, DOI: 10.1021/jacs.9b00193.

[400] O. V. Gusev, L. I. Denisovich, M. G. Peterleitner, A. Z. Rubezhov, N. A. Ustynyuk, and P. M. Maitlis, *Electrochemical generation of 19- and 20-electron rhodocenium complexes and their properties*, J. Organomet. Chem. **452** (1993) 219,
DOI: 10.1016/0022-328X(93)83193-Y.

[401] W. C. Henke, D. Lionetti, W. N. G. Moore, J. A. Hopkins, V. W. Day, and J. D. Blakemore, *Ligand Substituents Govern the Efficiency and Mechanistic Path of Hydrogen Production with [Cp*Rh] Catalysts*, ChemSusChem **10** (2017) 4589, DOI: 10.1002/cssc.201701416.

[402] M. J. Chalkley, T. J. Del Castillo, B. D. Matson, and J. C. Peters,
Fe-Mediated Nitrogen Fixation with a Metallocene Mediator: Exploring pKa Effects and Demonstrating Electrocatalysis, J. Am. Chem. Soc. **140** (2018) 6122,
DOI: 10.1021/jacs.8b02335.

[403] M. J. Chalkley, T. J. Del Castillo, B. D. Matson, J. P. Roddy, and J. C. Peters,
Catalytic N_2-to-NH_3 Conversion by Fe at Lower Driving Force: A Proposed Role for Metallocene-Mediated PCET, ACS Cent. Sci. **3** (2017) 217,
DOI: 10.1021/acscentsci.7b00014.

[404] J. H. Choi, Y. H. Kim, S. H. Nam, S. T. Shin, M.-J. Kim, and J. Park,
Aminocyclopentadienyl Ruthenium Chloride: Catalytic Racemization and Dynamic Kinetic Resolution of Alcohols at Ambient Temperature, Angew. Chem., Int. Ed. **41** (2002) 2373,
DOI: 10.1002/1521-3773(20020703)41:13<2373::AID-ANIE2373>3.0.CO;2-7.

[405] C. P. Casey, T. E. Vos, S. W. Singer, and I. A. Guzei,
Protonated Aminocyclopentadienyl Ruthenium Hydride Reduction of Benzaldehyde and the Conversion of the Resulting Ruthenium Triflate to a Ruthenium Hydride with H_2 and Base,
Organometallics **21** (2002) 5038, DOI: 10.1021/om020507d.

[406] A. Lator, Q. G. Gaillard, D. S. Mérel, J.-F. Lohier, S. Gaillard, A. Poater, and J.-L. Renaud,
Room-Temperature Chemoselective Reductive Alkylation of Amines Catalyzed by a Well-Defined Iron(II) Complex Using Hydrogen, J. Org. Chem. **84** (2019) 6813,
DOI: 10.1021/acs.joc.9b00581.

[407] R. Gompper and H. Glöckner, *Stable Cyclopentadienylium Salts*,
Angew. Chem., Int. Ed. Engl. **23** (1984) 53, DOI: 10.1002/anie.198400532.

[408] D. Catheline and D. Astruc,
Synthesis and characterization of $C_5(CH_3)_5Fe(CO)_3^+$ PF_6^- and $C_5(CH_3)_5Fe(CO)_2^-K^+$,
J. Organomet. Chem. **226** (1982) C52, DOI: 10.1016/S0022-328X(00)83414-2.

[409] N. Elgrishi, D. A. Kurtz, and J. L. Dempsey, *Reaction Parameters Influencing Cobalt Hydride Formation Kinetics: Implications for Benchmarking H_2-Evolution Catalysts*,
J. Am. Chem. Soc. **139** (2017) 239, DOI: 10.1021/jacs.6b10148.

[410] C. Jean-Michel Savéant Costentin, *Elements of Molecular and Biomolecular Electrochemistry: An Electrochemical Approach to Electron Transfer Chemistry*, 2019,
DOI: 10.1002/9781119292364.

[411] K. T. Nguyen, E. E. Lane, C. D. McMillen, J. A. Pienkos, and P. S. Wagenknecht,
Is Indenyl a Stronger or Weaker Electron Donor Ligand than Cyclopentadienyl? Opposing Effects of Indenyl Electron Density and Ring Slipping on Electrochemical Potentials,
Organometallics **39** (2020) 670, DOI: 10.1021/acs.organomet.9b00818.

[412] G. A. N. Felton, A. K. Vannucci, N. Okumura, L. T. Lockett, D. H. Evans, R. S. Glass, and D. L. Lichtenberger, *Hydrogen Generation from Weak Acids: Electrochemical and Computational Studies in the $[(\eta^5\text{-}C_5H_5)Fe(CO)_2]_2$ System*, Organometallics **27** (2008) 4671,
DOI: 10.1021/om800366h.

[413] E. B. Hemming, B. Chan, P. Turner, L. Corcilius, J. R. Price, M. G. Gardiner, A. F. Masters, and T. Maschmeyer, *$[Fe(C_5Ar_5)(CO)_2Br]$ complexes as hydrogenase mimics for the catalytic hydrogen evolution reaction*, Appl. Catal., B **223** (2018) 234,
DOI: 10.1016/j.apcatb.2017.04.053.

[414] A. Kütt, S. Tshepelevitsh, J. Saame, M. Lõkov, I. Kaljurand, S. Selberg, and I. Leito, *Strengths of Acids in Acetonitrile*, Eur. J. Org. Chem. **2021** (2021) 1407,
DOI: 10.1002/ejoc.202001649.

[415] A. Dutta, A. M. Appel, and W. J. Shaw,
Designing electrochemically reversible H_2 oxidation and production catalysts,
Nat. Rev. Chem. **2** (2018) 244, DOI: 10.1038/s41570-018-0032-8.

[416] I. Kaljurand, A. Kütt, L. Sooväli, T. Rodima, V. Mäemets, I. Leito, and I. A. Koppel, *Extension of the Self-Consistent Spectrophotometric Basicity Scale in Acetonitrile to a Full Span of 28 pKa Units: Unification of Different Basicity Scales*, J. Org. Chem. **70** (2005) 1019,
DOI: 10.1021/jo048252w.

[417] Although the origins of these modest Faradaic efficiencies remain unclear, preliminary data suggest that slow reductive decomposition of exogenous acid is occurring. Work is ongoing to identify the postelectrolysis product(s).

[418] A. M. Appel and M. L. Helm, *Determining the Overpotential for a Molecular Electrocatalyst*,
ACS Catal. **4** (2014) 630, DOI: 10.1021/cs401013v.

[419] It may also be possible that a concerted proton-electron transfer directly converts $FeCpN^+$ to an endo-protonated intermediate, circumventing these two compounds entirely.

[420] D. P. Estes, A. K. Vannucci, A. R. Hall, D. L. Lichtenberger, and J. R. Norton,
Thermodynamics of the Metal-Hydrogen Bonds in $(\eta^5\text{-}C_5H_5)M(CO)_2H$ (M = Fe, Ru, Os),
Organometallics **30** (2011) 3444, DOI: 10.1021/om2001519.

[421] R. Morris, *Brønsted-Lowry Acid Strength of Metal Hydride and Dihydrogen Complexes*, Chem. Rev. **116** (2016) 8588, DOI: 10.1021/acs.chemrev.5b00695.

[422] C. P. Casey and H. Guan, *An Efficient and Chemoselective Iron Catalyst for the Hydrogenation of Ketones*, J. Am. Chem. Soc. **129** (2007) 5816, DOI: 10.1021/ja071159f.

[423] C. P. Casey and H. Guan, *Cyclopentadienone Iron Alcohol Complexes: Synthesis, Reactivity, and Implications for the Mechanism of Iron-Catalyzed Hydrogenation of Aldehydes*, J. Am. Chem. Soc. **131** (2009) 2499, DOI: 10.1021/ja808683z.

[424] B. L. Conley, M. K. Pennington-Boggio, E. Boz, and T. J. Williams, *Discovery, Applications, and Catalytic Mechanisms of Shvo's Catalyst*, Chem. Rev. **110** (2010) 2294, DOI: 10.1021/cr9003133.

[425] L. Chong, J. Wen, J. Kubal, F. G. Sen, J. Zou, J. Greeley, M. Chan, H. Barkholtz, W. Ding, and D.-J. Liu, *Ultralow-loading platinum-cobalt fuel cell catalysts derived from imidazolate frameworks*, Science **362** (2018) 1276, DOI: 10.1126/science.aau0630.

[426] D. J. S. Sandbeck, N. M. Secher, F. D. Speck, J. E. Sørensen, J. Kibsgaard, I. Chorkendorff, and S. Cherevko, *Particle Size Effect on Platinum Dissolution: Considerations for Accelerated Stability Testing of Fuel Cell Catalysts*, ACS Catal. **10** (2020) 6281, DOI: 10.1021/acscatal.0c00779.

[427] E. S. Wiedner, A. M. Appel, S. Raugei, W. J. Shaw, and R. M. Bullock, *Molecular Catalysts with Diphosphine Ligands Containing Pendant Amines*, Chem. Rev. **122** (2022) 12427, DOI: 10.1021/acs.chemrev.1c01001.

[428] A. W. Nichols, S. L. Hooe, J. S. Kuehner, D. A. Dickie, and C. W. Machan, *Electrocatalytic CO_2 Reduction to Formate with Molecular Fe(III) Complexes Containing Pendent Proton Relays*, Inorg. Chem. **59** (2020) 5854, DOI: 10.1021/acs.inorgchem.9b03341.

[429] M. H. Rønne et al., *Ligand-Controlled Product Selectivity in Electrochemical Carbon Dioxide Reduction Using Manganese Bipyridine Catalysts*, J. Am. Chem. Soc. **142** (2020) 4265, DOI: 10.1021/jacs.9b11806.

[430] A. Chapovetsky, M. Welborn, J. M. Luna, R. Haiges, T. F. I. Miller, and S. C. Marinescu, *Pendant Hydrogen-Bond Donors in Cobalt Catalysts Independently Enhance CO_2 Reduction*, ACS Cent. Sci. **4** (2018) 397, DOI: 10.1021/acscentsci.7b00607.

[431] A. Chapovetsky, T. H. Do, R. Haiges, M. K. Takase, and S. C. Marinescu, *Proton-Assisted Reduction of CO_2 by Cobalt Aminopyridine Macrocycles*, J. Am. Chem. Soc. **138** (2016) 5765, DOI: 10.1021/jacs.6b01980.

[432] N. Queyriaux, D. Sun, J. Fize, J. Pécaut, M. J. Field, M. Chavarot-Kerlidou, and V. Artero, *Electrocatalytic Hydrogen Evolution with a Cobalt Complex Bearing Pendant Proton Relays: Acid Strength and Applied Potential Govern Mechanism and Stability*, J. Am. Chem. Soc. **142** (2020) 274, DOI: 10.1021/jacs.9b10407.

[433] S. Roy, B. Sharma, J. Pécaut, P. Simon, M. Fontecave, P. D. Tran, E. Derat, and V. Artero, *Molecular Cobalt Complexes with Pendant Amines for Selective Electrocatalytic Reduction of Carbon Dioxide to Formic Acid*, J. Am. Chem. Soc. **139** (2017) 3685, DOI: 10.1021/jacs.6b11474.

[434] S. Roy, B. Sharma, J. Pécaut, P. Simon, M. Fontecave, P. D. Tran, E. Derat, and V. Artero, *Correction to "Molecular Cobalt Complexes with Pendant Amines for Selective Electrocatalytic Reduction of Carbon Dioxide to Formic Acid"*, J. Am. Chem. Soc. **139** (2017) 8778, DOI: 10.1021/jacs.7b05505.

[435] A. W. Nichols and C. W. Machan, *Secondary-Sphere Effects in Molecular Electrocatalytic CO_2 Reduction*, Front. Chem. **7** (2019), DOI: 10.3389/fchem.2019.00397.

[436] J.-M. Savéant, *Proton Relays in Molecular Catalysis of Electrochemical Reactions: Origin and Limitations of the Boosting Effect*, Angew. Chem. Int. Ed. **58** (2019) 2125, DOI: 10.1002/anie.201812375.

[437] R. E. Warburton, A. V. Soudackov, and S. Hammes-Schiffer, *Theoretical Modeling of Electrochemical Proton-Coupled Electron Transfer*, Chem. Rev. **122** (2022) 10599, DOI: 10.1021/acs.chemrev.1c00929.

[438] A. VanderWeide and D. E. Prokopchuk, *Cyclopentadienyl ring activation in organometallic chemistry and catalysis*, Nat. Rev. Chem. **7** (2023) 561, DOI: 10.1038/s41570-023-00501-1.

[439] W. C. Henke, Y. Peng, A. A. Meier, E. Fujita, D. C. Grills, D. E. Polyansky, and J. D. Blakemore, *Mechanistic roles of metal- and ligand-protonated species in hydrogen evolution with [Cp∗Rh] complexes*, Proc. Natl. Acad. Sci. U.S.A. **120** (2023), DOI: 10.1073/pnas.2217189120.

[440] B. D. McCarthy, D. J. Martin, E. S. Rountree, A. C. Ullman, and J. L. Dempsey, *Electrochemical Reduction of Brønsted Acids by Glassy Carbon in Acetonitrile—Implications for Electrocatalytic Hydrogen Evolution*, Inorg. Chem. **53** (2014) 8350, DOI: 10.1021/ic500770k.

[441] To the best of our knowledge, the homoconjugation effects of bis(arylsulfonyl)imide acids in acetonitrile have not been investigated.

[442] N. Kaeffer, A. Morozan, J. Fize, E. Martinez, L. Guetaz, and V. Artero, *The Dark Side of Molecular Catalysis: Diimine-Dioxime Cobalt Complexes Are Not the Actual Hydrogen Evolution Electrocatalyst in Acidic Aqueous Solutions*, ACS Catal. **6** (2016) 3727, DOI: 10.1021/acscatal.6b00378.

[443] K. J. Lee, B. D. McCarthy, and J. L. Dempsey, *On decomposition, degradation, and voltammetric deviation: the electrochemist's field guide to identifying precatalyst transformation*, Chem. Soc. Rev. **48** (11 2019) 2927, DOI: 10.1039/C8CS00851E.

[444] M. Gómez-Gallego and M. A. Sierra, *Kinetic Isotope Effects in the Study of Organometallic Reaction Mechanisms*, Chem. Rev. **111** (2011) 4857, DOI: 10.1021/cr100436k.

[445] K. J. Lee, N. Elgrishi, B. Kandemir, and J. L. Dempsey,
Electrochemical and spectroscopic methods for evaluating molecular electrocatalysts,
Nat. Rev. Chem. **1** (2017), DOI: 10.1038/s41570-017-0039.

[446] D. Patel, A. Wooles, A. D. Cornish, L. Steven, E. S. Davies, D. J. Evans, J. McMaster, W. Lewis, A. J. Blake, and S. T. Liddle, *Synthesis and characterisation of halide, separated ion pair, and hydride cyclopentadienyl iron bis(diphenylphosphino)ethane derivatives*,
Dalton Trans. **44** (31 2015) 14159, DOI: 10.1039/C5DT00704F.

[447] Y. Hu, A. P. Shaw, D. P. Estes, and J. R. Norton, *Transition-Metal Hydride Radical Cations*,
Chem. Rev. **116** (2016) 8427, DOI: 10.1021/acs.chemrev.5b00532.

[448] Y. Peng, M. V. Ramos-Garcés, D. Lionetti, and J. D. Blakemore,
Structural and Electrochemical Consequences of [Cp] Ligand Protonation*,
Inorg. Chem. **56** (2017) 10824, DOI: 10.1021/acs.inorgchem.7b01895.

[449] P. Hamon, J.-R. Hamon, and C. Lapinte,
Isolation and characterization of a cationic 19-electron iron(III) hydride complex; electron transfer induced hydride migration by carbon monoxide at an iron(III) centre,
J. Chem. Soc., Chem. Commun. (21 1992) 1602, DOI: 10.1039/C39920001602.

[450] M. Cossi, N. Rega, G. Scalmani, and V. Barone, *Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model*,
J. Comput. Chem. **24** (2003) 669, DOI: 10.1002/jcc.10189.

[451] M. Krejčik, M. Daněk, and F. Hartl, *Simple construction of an infrared optically transparent thin-layer electrochemical cell: Applications to the redox reactions of ferrocene, $Mn_2(CO)_{10}$ and $Mn(CO)_3$(3,5-di-t-butyl-catecholate)$^-$*,
J. electroanal. chem. interfacial electrochem. **317** (1991) 179,
DOI: 10.1016/0022-0728(91)85012-E.

[452] R. G. Teller and J. M. Williams,
Crystal and molecular structure of bis[dicarbonyl(π-pentamethylcyclopentadienyl)iron], (η^5-C_5Me_5)$_2$$Fe_2(CO)_4$, and structural comparisons with the nonmethylated analog,
Inorg. Chem. **19** (1980) 2770, DOI: 10.1021/ic50211a055.

[453] O. S. Mills, *Studies of some carbon compounds of the transition metals. I. The crystal structure of dicyclopentadienyldiiron tetracarbonyl*, Acta Crystallogr. **11** (1958) 620,
DOI: 10.1107/S0365110X58001663.

[454] H. Sitzmann, T. Dezember, W. Kaim, F. Baumann, D. Stalke, J. Kärcher, E. Dormann, H. Winter, C. Wachter, and M. Kelemen, *Synthesis and Characterization of the Stable Dicarbonyl(cyclopentadienyl)iron Radical [(C_5R_5)$Fe(CO)_2$] (R = $CHMe_2$)*,
Angew. Chem., Int. Ed. Engl. **35** (1996) 2872, DOI: 10.1002/anie.199628721.

List of Figures

List of Tables

Acknowledgements

Through my Ph.D. studies, many people have accompanied, supported, and inspired me, and I want to take the opportunity to thank them. First, I want to thank my supervisor, Stefan Grimme, who has been very supportive throughout my studies. I have been in his group since my bachelor thesis, and I have learned an incredible amount from him and the people around him. In particular, I want to thank him for the opportunity to visit the European Summer School in Quantum Chemistry (ESQC) in Sicily in 2019 and the 10th triennial conference on Molecular Quantum Mechanics (MQM) in Blacksburg in the USA in 2022. At both of these events, I have learned a lot on a professional and a personal level while also having a lot of fun. Also, I want to thank Thomas Bredow for being so kind and agreeing to be my second reviewer. I want to thank Andreas Hansen, who supported me throughout my doctoral studies. He initiated and led exciting collaborations with Demyan Prokopchuk (Rutgers SASN) and James Shee (Rice), whom I also want to thank. The interesting scientific exchanges and collaborations with Demyan, James, and their students were very productive and enlightening. Many thanks to Benedikt Bädorf, Sebastian Ehlert, Thomas Gasevic, and Julius Kleine Büning for proofreading this thesis. Special thanks go to my long-term mentor, Markus Bursch, who supported me since my bachelor thesis in the Grimme group and continued to do so throughout my master's and Ph.D. studies. I also want to thank Sebastian Ehlert, my fellow throughout the bachelor´s and master´s program, long-term office colleague, and good friend. My special thanks also go to Jeroen Koopman and Julius Kleine Büning for their great support and friendship. During my time at the Mulliken Center, I had the privilege of mentoring the talented student Benedikt Bädorf for a short time, whom I also want to thank for this opportunity. Over the years, there have been so many more people with whom I had interesting scientific and non-scientific discussions and who supported me scientifically and mentally. I want to thank: Christoph Bannwarth, Christoph Bauer, Fabian Bohle, Eike Caldeweyher, Sebastian Dohm, Marvin Friede, Thomas Froitzheim, Johannes Gorges, Thomas Gasevic, Christian Hölzer, Julia Kohn, Abylay Katbashev, Lukas Kunze, Joachim Laun, Sarah Löffelsender, Jan-Michael Mewes, Marcel Müller, Christoph Plett, Philipp Pracht, Thomas Rose, Jakob Seibert, Sebastian Spicher, Marcel Stahn, and Lukas Wittmann. For administrative and technical support, I want to thank Claudia Kronz and Jens Mekelburger.

Outside the scientific community, I want to thank my family and Inken for their loving support, without which this work would have never been possible.