
Towards Predicting Spin States for Understanding Catalytic Activity of Iron Complexes

Dissertation

with the aim of achieving a doctoral degree at the Faculty of Mathematics, Informatics and
Natural Sciences, Department of Chemistry
University of Hamburg

submitted by
Ursula Christine Rastetter
2024

The presented work was carried out in the period from April 2020 to June 2024 in the Institute of Inorganic and Applied Chemistry at the University of Hamburg in the group of Prof. Dr. Carmen Herrmann and in the group of Prof. Dr. Axel Jacobi von Wangelin.

Date of Submission: 19.08.2024
Board of examiners: Prof. Dr. Carmen Herrmann
Prof. Dr. Axel Jacobi von Wangelin

Date of Defense: 06.09.2024
Board of examiners: Prof. Dr. Carmen Herrmann
Prof. Dr. Andrew Torda
Prof. Dr. Wolfgang Maison
Prof. Dr. Gabriel Bester
Prof. Dr. Axel Jacobi von Wangelin

Approval of dissertation: 04.04.2025

Contents

1	Abstract	1
2	Zusammenfassung	3
3	Introduction	5
I	Computational Studies	7
4	Theoretical Background	9
4.1	Kohn–Sham Density Functional Theory	9
4.1.1	Approximate Exchange–Correlation Functionals	10
4.2	Supervised Machine Learning	11
4.2.1	Validation of Supervised Machine Learning Models	11
4.2.2	Nearest Neighbors Regression	12
4.2.3	Gaussian Process Regression	13
4.2.4	Smooth Overlap of Atomic Positions Descriptor	14
4.2.5	Many-Body Tensor Representation Descriptor	14
4.2.6	Atom-centered Symmetry Functions Descriptor	16
4.2.7	Autocorrelation Functions as Descriptors	17
4.3	Error Measures	17
5	Redox-Active Ligands as a Challenge for Electronic Structure Methods	19
5.1	Introduction	20
5.2	Computational Details and Data Set Construction	23
5.3	Spin-State Energetics in Tris(diimine) Iron(II) Complexes	25
5.4	SCF Convergence to Local Minima	28
5.5	Reduced Tris(diimine) Iron Complexes as a Challenging Test Case	31
5.6	Conclusions	34
6	Machine Learning Guided Strategies Towards Spin-Gaps in Carbenes	37
6.1	Introduction	38
6.2	Computational Details	41
6.3	Data Set Construction	41
6.4	Computing Times	43
6.5	Influence of Molecular Structures on DFT Calculated Adiabatic and Vertical Spin Gaps	45
6.5.1	Differences in Molecular Structures	50
6.6	Influence of Molecular Structures on the Supervised Learning of Adiabatic Spin Gaps	52
6.6.1	Performance Analysis of Models	55
6.6.2	Performance Analysis of Descriptors	56
6.6.3	Dependency between Performance and Molecular Structure	57
6.6.4	Descriptors Based on Electronic Structure Properties	59
6.7	Delta-Machine Learning	61
6.7.1	Adiabatic Spin Gaps with CASSCF- and B3LYP-Optimized Structures	62
6.7.2	Spin Gaps with B3LYP-Optimized Structures	64
6.7.3	Spin Gaps with PM6-Optimized Structures	68
6.8	Conclusions	71

II Synthesis and Characterization	73
7 Synthesis and Characterization of Dinuclear Iron Complexes	75
7.1 Introduction to Polynuclear Hydrido-Bridged Iron Complexes	76
7.2 Catalyzed Hydrogenation of Alkenes with ^{Dipp} BIAN Iron Complexes	82
7.3 Synthesis and Characterization of Reduced ^{Dipp} BIAN Iron Complexes	84
7.4 Reactivity of Ironhydrides	95
7.5 Conclusions	99
8 Experimental Section	101
8.1 General	101
8.2 Synthesis	102
8.2.1 General Procedure for Catalysed Alkene Hydrogenation	102
8.2.2 General Procedure for Reactivity Tests of Hydridoferrates	102
8.2.3 General Procedure for Catalysed CO ₂ Hydrogenation	102
8.2.4 General Procedure for Reactions of CO ₂ with Hydridometallates	103
8.2.5 Synthesis of ^{Dipp} BIAN L	103
8.2.6 Synthesis of (^{Dipp} BIAN)FeCl ₂ 1	103
8.2.7 Synthesis of 2 , 5 , 3a and 4	104
8.2.8 Synthesis of [Na(thf) ₆] ₂ {[Fe(^{Dipp} BIAN)] ₂ (μ-H) ₂ } 3b	105
8.2.9 Synthesis of [Fe(L)(tol) 6]	106
8.2.10 Synthesis of (^{Dipp} BIAN)CoBr ₂ 7	107
8.2.11 Synthesis of [{Co(^{Dipp} BIAN)] ₂ (μ-H) ₃ } ⁻ 8	107
8.2.12 Synthesis of [{Co(^{Dipp} BIAN)] ₂ (formato) ₂ } 9	108
8.2.13 Reaction of [{Co(^{Dipp} BIAN)] ₂ (μ-H) ₃ } ⁻ 8 with DBU	108
8.2.14 Reaction of (^{Dipp} BIAN)CoBr ₂ 7 with DBU	109
8.2.15 Synthesis of (^{Dipp} BDI)CoBr ₂ 11	109
8.2.16 Synthesis of [{Co(^{Dipp} BDI)] ₂ (μ-H) ₂ } ⁻ 12a	110
8.2.17 Synthesis of [{Co(^{Dipp} BDI)] ₂ (μ-H) ₂ } ⁻ 12b	110
9 Conclusions	113
A List of Publications	115
B Supporting Information for Redox-Active Ligands as a Challenge for Electronic Structure Methods	117
B.1 Additional Data on Spin-State Energy Splittings of Trisdiimineiron(II) Complexes	117
B.2 Additional Data on Spin-State Energy Splittings of Trisdiimineiron(I) Complexes	121
B.3 Spin Contamination and FOD Analysis	126
B.4 Broken Symmetry Calculations	133
B.5 Studies on Correlations With Redox-Activity	134
B.6 Characteristic Structural Features of Optimized Structures	139
B.7 Zero-Point Vibrational Energies and Dispersion Corrections	146
C Supporting Information for Machine Learning Guided Strategies Towards Spin Gaps in Carbenes	149
C.1 Additional Data on Computing Times	150
C.2 Additional Data for DFT Calculations	154
C.3 Additional Data on ML	159
D Supporting Information for Synthesis and Characterization	165
D.1 IR Spectra	165
D.2 NMR Spectra	166
D.3 Mass Spectra	173
D.4 Crystal Data	178

D.5 Cyclic Voltammetry	187
D.6 Additional DFT Calculations	189
E Hazardous Materials	197
Bibliography	203

1 Abstract

Redox active ligands are a promising tool to improve the catalytic activity of 3d transition metal catalysts. These ligands influence the oxidation state of the metal center as well as the ground spin-state and make the determination of both properties by experiment alone challenging. Therefore, this thesis aims to advance the understanding and optimization of these complexes for catalytic applications through experiment complimented by the use of first-principles calculations, particularly density functional theory (DFT).

For employing DFT, a proper choice of exchange–correlation (xc) functional is crucial. 43 homoleptic octahedral tris(diimine) iron(II) complexes are investigated to evaluate the influence of redox-active ligands on spin-state energetics. Despite the significant effect of varying redox activity on the electronic structure, the sensitivity of spin-state energy splittings to the exact exchange admixture in exchange–correlation (xc) functionals is found to be unexpectedly small. However, for iron(II) complexes with highly redox-active ligands and for reduced tris(diimine) iron(I) complexes, self-consistent field (SCF) convergence to local minima is observed, leading to incorrect redox states and molecular structures. Protocols for detecting and addressing these convergence issues, particularly in large-scale calculations, are proposed and discussed.

To evaluate computational methods for the prediction of spin state energetics, this work also focuses on organic carbenes. A data set of 2841 carbenes from the QMspin data set[1] was utilized to highlighting the impact of molecular structure not only on DFT calculations and xc functional choice but also on machine learning (ML) and Δ -ML as a combination of both methods. The study reveals that B3LYP-optimized structures, rather than a combination of CASSCF- and B3LYP-optimized molecular structures, often yield better performance for DFT calculations due to error compensation effects. Additionally, Δ -ML-based methods can significantly improve predictive accuracy. For the latter method CASSCF- and B3LYP-optimized molecular structures proved to be better suited.

The final part of the thesis presents the synthesis and characterization of iron complexes with redox-active bis(imino)acenaphthene (BIAN) ligands. Novel hydrogen-bridged dinuclear ferrate complexes have been synthesized and a neutral dinuclear iron complex $\text{Fe}_2(\text{BIAN})_2$ with a unique coordination sphere and a very short Fe–Fe distance of 2.3068(8) Å has been isolated. The hydridoferrate complexes are shown to exhibit unique electronic structures and reactivity, including catalytic activity in hydrogenation reactions. The findings provide valuable insights into the design of new catalysts based on similar ligand frameworks.

Overall, this thesis contributes to the development of more accurate and reliable computational methods for studying transition metal complexes and their catalytic properties. It also opens avenues for future research, including the refinement of DFT functionals, the integration of machine learning with quantum chemistry, and the exploration of novel ligand systems for catalytic applications.

2 Zusammenfassung

Redox-aktive Liganden sind ein vielversprechendes Werkzeug zur Verbesserung der katalytischen Aktivität von $3d$ -Übergangsmetallkatalysatoren. Diese Liganden beeinflussen den Oxidationszustand des Metallzentrums sowie den Spingrundzustand und erschweren die Bestimmung beider Eigenschaften allein durch Experimente. Daher zielt diese Dissertation darauf ab, das Verständnis und die Optimierung dieser Komplexe für katalytische Anwendungen durch Experimente, ergänzt durch erstprinzipielle Berechnungen, insbesondere der Dichtefunktionaltheorie (DFT), voranzutreiben.

Für den Einsatz von DFT ist eine sorgfältige Auswahl des Austausch-Korrelations-Funktionalentscheidend. 43 homoleptische, oktaedrische Tris(diimin)-Eisen(II)-Komplexe wurden untersucht, um den Einfluss von redox-aktiven Liganden auf die Spin-Zustands-Energien zu bewerten. Trotz des signifikanten Einflusses der variierenden Redox-Aktivität auf die elektronische Struktur zeigt sich, dass die Empfindlichkeit der Spin-Zustands-Energieaufspaltung gegenüber dem exakten Austauschanteil in Austausch-Korrelations-Funktionalen unerwartet gering ist. Allerdings wird bei Eisen(II)-Komplexen mit stark redox-aktiven Liganden und bei reduzierten Tris(diimin)-Eisen(I)-Komplexen eine Konvergenz des selbstkonsistenten Feldes (SCF) zu lokalen Minima beobachtet, was zu falschen Redox-Zuständen und molekularen Strukturen führt. Protokolle zur Erkennung und Bewältigung dieser Konvergenzprobleme, insbesondere bei groß angelegten Berechnungen, werden vorgeschlagen und diskutiert.

Zur Evaluierung von Rechenmethoden zur Vorhersage von Spin-Zustands-Energien konzentriert sich diese Arbeit auch auf organische Carbene. Ein Datensatz von 2841 Carbenen aus dem QMspin-Datensatz [1] wurde verwendet, um die Auswirkungen der molekularen Struktur nicht nur auf DFT-Berechnungen und die Wahl des Austausch-Korrelations-Funktional, sondern auch auf maschinelles Lernen (ML) und Δ -ML als Kombination beider Methoden hervorzuheben. Die Studie zeigt, dass B3LYP-optimierte Strukturen, anstelle einer Kombination aus CASSCF- und B3LYP-optimierten molekularen Strukturen, oft eine bessere Leistung für DFT-Berechnungen aufgrund von Fehlerkompensationseffekten liefern. Darüber hinaus können auf Δ -ML basierende Methoden die Vorhersagegenauigkeit erheblich verbessern. Für letztere Methode erwiesen sich CASSCF- und B3LYP-optimierte molekulare Strukturen als besser geeignet.

Der letzte Teil der Dissertation stellt die Synthese und Charakterisierung von Eisenkomplexen mit redox-aktiven Bis(imino)acenaphthen (BIAN)-Liganden vor. Neuartige hydridverbrückte dinukleare Ferrat-Komplexe wurden synthetisiert, und ein neutraler dinuklearer Eisenkomplex $\text{Fe}_2(\text{BIAN})_2$ mit einer einzigartigen Koordinationssphäre und einem sehr kurzen Fe-Fe-Abstand von $2.3068(8) \text{ \AA}$ wurde isoliert. Es wird gezeigt, dass die Hydridoferrat-Komplexe einzigartige elektronische Strukturen und Reaktivitäten aufweisen, einschließlich katalytischer Aktivität in Hydrierungsreaktionen. Die Ergebnisse liefern wertvolle Einblicke in das Design neuer Katalysatoren auf Basis ähnlicher Ligandengerüste.

Insgesamt leistet diese Dissertation einen Beitrag zur Entwicklung genauerer und zuverlässigerer Rechenmethoden für das Studium von Übergangsmetallkomplexen und deren katalytischen Eigenschaften. Sie eröffnet auch Perspektiven für zukünftige Forschungen, einschließlich der Verfeinerung von DFT-Funktionalen, der Integration von maschinellem Lernen mit Quantenchemie und der Erforschung neuartiger Ligandensysteme für katalytische Anwendungen.

3 Introduction

Understanding catalytic intermediates through both experimental and computational approaches is crucial for improving the performance of chemical transformations. Well-known catalysts for key transformations, such as hydrogenation reactions, are noble $4d$ and $5d$ transition metal (TM) complexes. In contrast, $3d$ TMs, which are more abundant and often less toxic than their heavier counterparts, present a more sustainable alternative. Among these, iron has garnered significant attention as a substitute for established noble metal catalysts, aligning with the global push for sustainability.[2–10] Unlike noble TM complexes, $3d$ TMs can access a wide range of spin ground states, opening up new reaction pathways. One strategy to mimic $4d$ and $5d$ TMs is to use redox-active ligands with $3d$ TMs. These ligands can accept electrons from the metal center, thereby facilitating redox steps in catalytic processes.

The synthesis and experimental characterization of these often paramagnetic $3d$ TM complexes can be highly challenging. As a result, computational approaches are frequently employed to complement experimental work. Among the various electronic structure methods, density functional theory (DFT) remains the most popular choice due to its balance between accuracy and efficiency. However, accurately predicting spin-dependent properties remains a notorious challenge for DFT.[11, 12] Although methods such as coupled cluster (CC) theory, second-order N -electron valence state perturbation theory (NEVPT2), or complete active space self-consistent field (CASSCF) offer higher accuracy, DFT is significantly more computationally efficient.[13, 14] Additionally, DFT’s performance can be tailored by selecting appropriate exchange-correlation (xc) functionals. Several studies have investigated the influence of different xc functionals on spin-state energy splittings in $3d$ TM complexes, revealing a strong, mostly linear dependence on the amount of exact exchange admixture in the xc functional.[15–24] This dependence is further influenced by the electronic properties of the ligand.[24, 25]

Given the significant impact of redox-active ligands on the electronic structure of TM complexes, one aim of this work is to investigate this influence. To focus specifically on the redox properties of the ligand, a test set of tris(diimine) iron(II) complexes was selected. The uniform coordination geometry and the use of the same metal across all complexes eliminate variability due to these factors. This benchmark study is presented in Chapter 5.

The significance of spin-dependent properties extends beyond TM complexes. Carbenes, which are key intermediates in numerous organic reactions, are characterized by a divalent carbon atom with an electron sextet.[26, 27] The reaction pathways of these highly reactive species are heavily influenced by their spin state,[28, 29] and computational studies have become essential for exploring these reactive intermediates.[30–33] The group of Lilienfeld *et al.* conducted a large-scale study on the electronic structures of carbenes, resulting in the QMspin data set,[1] which encompasses a broader chemical space compared to the tris(diimine) iron(II) complexes used in this work.

This extensive data set also facilitates comparisons between DFT and emerging approaches like machine learning (ML), which has gained attention for its potential to reduce computational time while maintaining or even improving accuracy. ML has been applied to predict molecular properties, including spin state energies.[34–41] One promising approach is Δ -ML, which augments calculated values with machine-learned corrections rather than predicting molecular properties directly.[42] While molecular representation and descriptor selection

for ML have been discussed extensively in the literature,[43–45] the impact of the origin of molecular structures on ML predictions remains underexplored.

This work aims to evaluate the performance of DFT with various xc functionals, ML, and Δ -ML in predicting spin gaps in carbenes. A particular focus is placed on the impact of the origin of molecular structures—whether they are optimized using DFT or the more computationally efficient semiempirical PM6 method. Additionally, the use of only singlet structures or both singlet and triplet structures significantly influences computational cost, and this aspect is also discussed in Chapter 6.

As previously mentioned, 3d TM catalysis is of significant interest, particularly iron. Homogeneously catalyzed hydrogenation reactions are a cornerstone of organic synthesis, essential for producing a wide range of chemical products. In these catalytic processes, hydride species often serve as crucial intermediates.[46–51] Despite their importance, there are relatively few examples of polynuclear iron complexes featuring bridging hydrides, highlighting an area with significant potential for further exploration.

A dimeric hydridocobaltate complex with the redox-active ligand bis(imino)acenaphthene (BIAN) has been reported as an efficient precatalyst for the hydrogenation of alkenes under mild conditions.[52] BIANs are readily synthesized from commercial precursors on a multigram scale and are highly redox-active, capable of accepting up to four electrons.[53, 54] Preliminary work has successfully adapted the synthetic approach used for the hydridocobaltate to create a dimeric hydrido-bridged ferrate complex, which also serves as an active precatalyst for alkene hydrogenation.[55] As mentioned earlier, the characterization of this paramagnetic and highly air- and moisture-sensitive compound has proven to be challenging.

Consequently, another focus of this work is to further characterize this hydridoferrate, employing computational approaches informed by the knowledge gained from the studies mentioned above. Additionally, this work explores the reactivity profile of such compounds. The approaches to synthesize highly reduced iron complexes with BIAN ligands, as well as their experimental and computational characterization, are presented in Chapter II.

Part I

Computational Studies

4 Theoretical Background

4.1 Kohn–Sham Density Functional Theory

In contrast to other quantum mechanical theories that rely on wave functions, density functional theory (DFT) calculates energy and other quantum mechanical properties directly from the ground-state electron density (Equation (4.1)).[56]

$$E = E[\rho]. \quad (4.1)$$

As a result, only three spatial coordinates need to be considered instead of $4N$, where N is the number of electrons in the system. The foundation of DFT lies in the Hohenberg–Kohn theorems.[57] The first theorem establishes a one-to-one correspondence between the ground-state density and the external potential (Equation (4.2)), which further implies a unique relationship between the ground-state density and the ground-state wave function (Equation (4.3)).

$$\rho_0(\vec{r}) \leftrightarrow V_{ext}(\vec{r}) \quad (4.2)$$

$$\rho_0(\vec{r}) \leftrightarrow \Psi_0(\vec{r}_1, \dots, \vec{r}_N). \quad (4.3)$$

According to the variational principle, the ground-state wave function is the one that minimizes the energy. Due to the correlation between ground-state density and ground-state wave function, the ground-state energy E_0 can also be obtained by minimizing the energy functional $E[\rho]$ with respect to variations in the density ρ (second Hohenberg–Kohn theorem).

The energy functional is composed of several energy contributions:

$$E[\rho] = T_n[\rho] + T_e[\rho] + V_{ee}[\rho] + V_{en}[\rho] + V_{nn}[\rho]. \quad (4.4)$$

The kinetic energy of the nuclei $T_n[\rho]$ and the internuclear interaction $V_{nn}[\rho]$ can be simplified using the Born–Oppenheimer approximation,[58]

$$T_n[\rho] = 0 \quad (4.5)$$

$$V_{nn}[\rho] = \text{const.} \quad (4.6)$$

For the electron-nuclei interaction $V_{en}[\rho]$, a term can be derived from classical electrostatics (Equation (4.7)):

$$V_{en}[\rho] = \int V_{ext}(\vec{r})\rho(\vec{r}) \, d\vec{r}. \quad (4.7)$$

The same principle applies to the Coulomb term $J[\rho]$, which represents the classical part of the electron-electron interaction $V_{ee}[\rho]$. However, no classical analog exists for the remaining part of the electron-electron interaction, $E_{nonclass}[\rho]$, which arises from exchange and correlation, as well as the kinetic energy of the electrons $T_e[\rho]$ due to electron correlation.

$$V_{ee}[\rho] = J[\rho] + E_{nonclass}[\rho] \quad (4.8)$$

$$T_e[\rho] = T_s[\rho] + \Delta T[\rho] \quad (4.9)$$

To address the challenge posed by interacting particles, Kohn and Sham introduced a reference system of non-interacting fermions that share the same ground-state density as the original system (Equation (4.10)).[59]

$$\rho(\vec{r}) = \sum_{i=1}^N |\psi_i(\vec{r})|^2. \quad (4.10)$$

For such a system, the exact ground-state wave function is a single Slater determinant containing the N Kohn–Sham orbitals $\psi_i(\vec{r})$. The kinetic energy $T_s[\rho]$ of the non-interacting system can therefore be calculated exactly. The minor error introduced by this approximation is encapsulated, along with exchange and correlation terms, in the exchange–correlation functional $E_{xc}[\rho]$:

$$E[\rho] = T_s[\rho] + V_{eK}[\rho] + J[\rho] + E_{XC}[\rho] \quad (4.11)$$

$$\text{with } E_{XC}[\rho] = E_X[\rho] + E_C[\rho] + \Delta T[\rho]. \quad (4.12)$$

For practical applications of KS-DFT, approximations must be made for the exchange–correlation functional.

4.1.1 Approximate Exchange–Correlation Functionals

Approximate exchange–correlation (xc) functionals can be divided by different types. “Pure” functionals, also known as local or semi-local functionals, solely rely on the local electron density (Local Density Approximation (LDA)) and additionally on the gradient of the electron density in case of the generalized gradient approximation (GGA). An example for a GGA xc functional is BLYP. Its the exchange–correlation energy E_{XC}^{BLYP} is given by (4.13):

$$E_{XC}^{BLYP} = E_X^{LDA} + \Delta E_X^{B88} + E_C^{LDA} + \Delta E_C^{LYP} \quad (4.13)$$

Contributions are E_X^{LDA} the exchange energy in the LDA), ΔE_X^{B88} the gradient-corrected exchange energy according to Becke’s 1988 functional (B88), E_C^{LDA} the correlation energy in LDA and ΔE_C^{LYP} the correlation energy correction according to the Lee–Yang–Parr (LYP) functional.[60]

Hybrid xc functionals also include an amount a_{HF} of the exact exchange energy calculated using Hartree-Fock theory E_X^{HF} which is defined by (4.14).

$$E_X^{HF} = -\frac{1}{2} \sum_{i,j} \int \int \psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (4.14)$$

In (4.14) $\psi_i(\mathbf{r})$ is the spin-orbital wavefunction for the i -th orbital and $\psi_i^*(\mathbf{r})$ is the complex conjugate of the spin-orbital wavefunction. $|\mathbf{r}_1 - \mathbf{r}_2|$ is the distance between two electrons at positions \mathbf{r}_1 and \mathbf{r}_2 .

An example for a hybrid functional is B3LYP functional. (4.15) shows the exchange-correlation energy calculated using the B3LYP functional E_{XC}^{B3LYP} [61]:

$$E_{XC}^{B3LYP} = E_{XC}^{LDA} + a_{HF}(E_X^{HF} - E_X^{LDA}) + a_X \Delta E_X^{B88} + a_C \Delta E_C^{LYP} \quad (4.15)$$

a_{HF} , a_X , and a_C are empirically determined parameters, with the values $a_{HF} = 0.20$, $a_X = 0.72$ and $a_C = 0.81$

Range separated xc functionals are designed to improve the accuracy of the exchange part of the functional by separating the treatment of short-range (SR) and long-range (LR) electron-electron interactions. This allows different admixtures of exact HF exchange depended on the range of the interaction. An example is the CAM-B3LYP functional, that at short ranges uses 19 % E_X^{HF} that smoothly increases to 65 % at long ranges.[62]

4.2 Supervised Machine Learning

Supervised machine learning (ML) is a type of data analysis where a model is trained to predict an output based on a set of input features. This approach is called “supervised” because the model learns from a labeled data set, where both the input features and the corresponding output labels or values (sometimes also referred to as targets) are known.

In supervised ML, the data set used for training consists of pairs $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, where \mathbf{x}_i represents the input feature vector and y_i denotes the corresponding output value. The goal is to learn a mapping from the input features to the output, such that the model can make accurate predictions on new, unseen data. The model is typically represented as a function $f(\mathbf{x}; \theta)$, where θ represents the parameters of the model. During training, the model’s parameters are adjusted to minimize a loss function $L(y, f(\mathbf{x}; \theta))$, which measures the difference between the predicted output $f(\mathbf{x}; \theta)$ and the true output y .

While for electronic structure methods the representation in nuclear charges and atomic coordinates is sufficient, they are not a suitable representation for ML as they are not invariant to rotation and translation. Therefore, descriptors which represent the molecule in a more suitable feature space are designed. Some examples are explained in Sections 4.2.4, 4.2.5, 4.2.6 and 4.2.7. Typically, the descriptors are based on structural features but it is also common ad other molecular properties.[63]

4.2.1 Validation of Supervised Machine Learning Models

Validation is a critical step in the development and evaluation of supervised machine learning models, including the tuning of their hyperparameters. Hyperparameters define the configuration of a machine learning model and are not adjusted during the learning procedure itself.

Validation involves assessing the model’s performance on unseen data, also referred to as the test set, to ensure that the model generalizes well beyond the specific data set used for training. The primary goal of validation is to detect and mitigate issues such as overfitting, where a model performs well on the training data but poorly on new data.

A common validation technique is the train-test split, where the data set is divided into two distinct parts: a training set and a test set. The model is trained on the training set and evaluated on the test set. This approach provides a straightforward measure of model performance, but it is important to ensure that the split is representative of the overall data

set to avoid biased results. A simple approach to achieve this is to repeat the validation with a random train-test split. An advantage of this validation technique is that the size of the training set can be easily adjusted, allowing for an analysis of the dependence of the model's performance on the training set size (learning curves).

One of the most commonly used techniques for validation is cross-validation. In k -fold cross-validation, the data set is divided into k subsets of equal size. The model is trained on $k-1$ of these subsets (training set) and validated on the remaining subset (test set). This process is repeated k times, with each subset being used exactly once as the validation data. The results are averaged to provide an estimate of the model's performance. However, in contrast to the random train-test split, learning curves cannot easily be created since the number of repetitions also depends on k , which controls the size of the training and test sets.

4.2.2 Nearest Neighbors Regression

The simple k -nearest neighbors regression model involves mainly two hyperparameters. The number of neighbors k and a “weights” parameter which can be set to either “uniform” or “distance”. As illustrated in Figure 4.1, distances for a new data point which is described by the feature vector \mathbf{x}_{new} are calculated to the points (\mathbf{x}_i) in the training data. In the case that the weights are uniform, the predicted value \hat{y}_{new} for this new point is then calculated from the mean of target values y_i of the k nearest points ((4.16)). Each of the k neighboring data points contribute equally. Closer points will have a greater influence than points further away from the new point if a weight based on distance is used. Again, only the k closest points are considered for calculating the predicted value. But the target values of the training set are in this case multiplied by the inverse distance as shown in (4.17)

$$\hat{y}_{\text{new}} = \frac{1}{k} \sum_{i=1}^k y_i \quad (4.16)$$

$$\hat{y}_{\text{new}} = \frac{\sum_{i=1}^k \frac{1}{d_i} \cdot y_i}{\sum_{i=1}^k \frac{1}{d_i}} \quad (4.17)$$

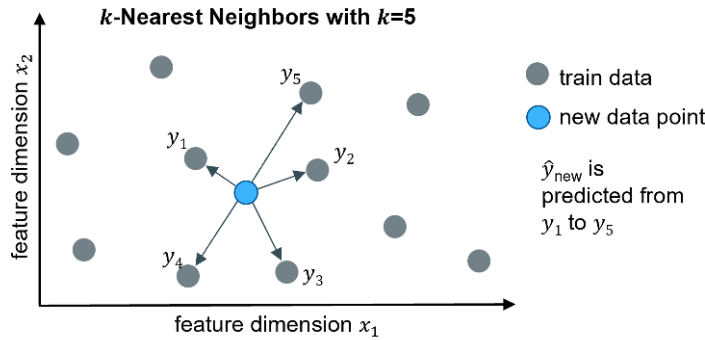


Figure 4.1: Illustration of the k -nearest neighbors prediction for 2 dimensional feature space.

4.2.3 Gaussian Process Regression

The Gaussian process regression (GPR) is a non-parametric, probabilistic model. It is particularly powerful for small to medium-sized data sets and provides a measure of uncertainty in the predictions. GPR models the distribution over possible functions that fit the data, making it well-suited for problems where the underlying relationship between input and output is complex and non-linear.[64]

The predictive distribution of the target value y_* of a new input feature vector \mathbf{x}_* for a training set with input matrix \mathbf{X} and the target feature vector \mathbf{y} is a normal distribution with the predictive mean \bar{f}_* and the predictive variance σ_* (4.18). \mathbf{x}_* has the size $N \times D$, where N is the number of training data and D the dimensionality of the features.[65]

$$p(y_* | \mathbf{x}_*, \mathbf{X}, \mathbf{y}) = \mathcal{N}(\mu_*, \sigma_*^2) \quad (4.18)$$

The predictive mean and variance can be calculated according to (4.19) and (4.20), respectively. The covariance matrix of the training data \mathbf{K} is defined by covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$. \mathbf{k}_* is the vector of covariances between the new point and all N training points.[65]

$$\bar{f}_* = \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} \quad (4.19)$$

The noise variance σ_n^2 includes noise of the training data and is 0 for noise-free data.[65]

$$\sigma_*^2 = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_* \quad (4.20)$$

With $\alpha = (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}$ the mean prediction (4.20) can also be written as a linear combination of the observations \mathbf{y} :[65]

$$\bar{f}(\mathbf{x}_*) = \sum_i^N \alpha_i k(\mathbf{x}_i, \mathbf{x}_*) \quad (4.21)$$

The covariance function in GPR is also referred to as the kernel, which is a key component that measures measuring similarities between two data points. Common kernels involve parameters that control the shape and smoothness of the predicted function. These parameters are fitted during the training of the data. The radial basis function (RBF) kernel, also known as the “squared exponential” or Gaussian kernel, is given by (4.22) and is parameterized by the length scale parameter l . As the name suggests it controls the length scale for which two data points are considered similar. It can be either a scalar or a vector with D dimensions, that controls each feature dimension separately.

$$k_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2l^2}} \quad (4.22)$$

The Matérn kernel is a generalization of the RBF kernel ($\nu = 0.5$) that involves an additional parameter ν that controls the smoothness of the predicted function $\Gamma(\nu)$ and $K_\nu()$ are a modified Bessel function and the gamma function, respectively.[65]

$$k_{\text{Matern}}(\mathbf{x}_i, \mathbf{x}_j) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} \|\mathbf{x}_i - \mathbf{x}_j\|}{l} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu} \|\mathbf{x}_i - \mathbf{x}_j\|}{l} \right) \quad (4.23)$$

4.2.4 Smooth Overlap of Atomic Positions Descriptor

The smooth overlap of atomic positions (SOAP) method[43] describes the local environment within a atomic system. The SOAP representation is invariant to translation and rotation. The local atomic surroundings is represented by expanding a gaussian smoothed atomic density using a combination of spherical harmonics Y_{lm} and radial basis functions g_n .

The number of feature dimensions of the SOAP descriptor depends on the number of pairs of different atom species (α and β), the number n_{max} of the radial basis functions and the number l_{max} of spherical harmonics. The features are elements $p_{nn'l}^{\alpha\beta}$ of the partial power spectrum vector \mathbf{p} which are defined as:

$$p_{nn'l}^{\alpha\beta} = \pi \sqrt{\frac{8}{2l+1}} \sum_m c_{nlm}^{\alpha} c_{n'l m}^{\beta} \quad (4.24)$$

n and n' are indices for radial basis functions that correspond to atom species α and β , respectively. l is the angular degree of the spherical harmonics. The coefficients in equation 4.24 can be calculated as followed:

$$c_{nlm}^{\alpha} = \iiint_{\mathcal{R}^3} dV g_n(r) Y_{lm}(\theta, \phi) \rho^{\alpha}(\mathbf{r}) \quad (4.25)$$

The atomic density of atom species α , $\rho^{\alpha}(\mathbf{r})$, is obtained by a sum over un-normalized gaussians centered on each atom of the species α :

$$\rho^{\alpha}(\mathbf{r}) = \sum_{i \in \alpha} e^{-\frac{1}{2\sigma^2}|\mathbf{r}-\mathbf{R}_i|^2} \quad (4.26)$$

The SOAP descriptor can be tuned by several key parameters: apart from l_{max} and n_{max} prescribed above, the radial cutoff defines the extent of the atomic environment considered. The atomic density can be varied by the standard deviation σ of the Gaussian functions (equation 4.26). Additionally, the choice of radial basis functions can vary, with common options including spherical Gaussian orbitals or polynomial bases.[63]

4.2.5 Many-Body Tensor Representation Descriptor

The many-body tensor representation (MBTR)[66] encodes a structure by breaking them down into distributions \mathcal{D}_k of structural motifs that involve k atoms. These distributions are grouped by the chemical spaces. Typically k values up to 3 are used (Figure 4.2).

The distributions are generated by firstly applying a geometry function g_k on the k atoms of the structural motive to obtain a single value. The geometry function g_1 is the atomic number. For g_2 either the distance or the inverse distance is used and g_3 is the angle or cosine of the angle. Broadening these scalar values with gaussian kernels with the standard deviation σ_k leads to the distributions defined in (4.27), (4.28) and (4.29):

$$\mathcal{D}_1^l(x) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-g_1(Z_1))^2}{2\sigma_1^2}} \quad (4.27)$$

$$\mathcal{D}_2^{l,m}(x) = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-g_2(\mathbf{R}_l, \mathbf{R}_m))^2}{2\sigma_2^2}} \quad (4.28)$$

$$\mathcal{D}_3^{l,m,n}(x) = \frac{1}{\sigma_3 \sqrt{2\pi}} e^{-\frac{(x-g_3(\mathbf{R}_l, \mathbf{R}_m, \mathbf{R}_n))^2}{2\sigma_3^2}} \quad (4.29)$$

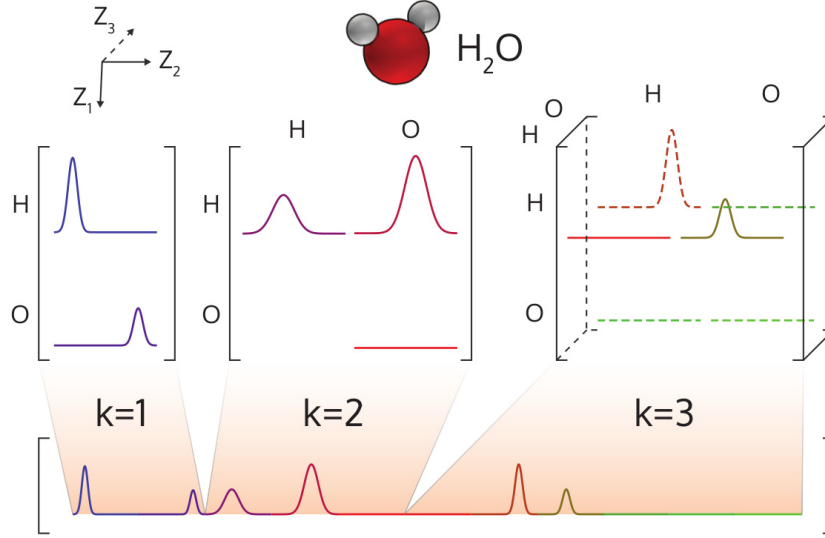


Figure 4.2: MBTR output for a water molecule showing the distributions MBTR_k for $k = 1, 2, 3$ with different combinations of chemical elements. For each term, the distributions can be arranged into a k -dimensional grid, resulting in a $k+1$ dimensional tensor. If a flattened one-dimensional vector is needed by the learning model, the distributions may be concatenated together, possibly with some weighting, as shown in the lower panel. Reproduced from reference [63].

The values x ensure that all possible values for g_k are covered. The combined weighted sums over each of the distributions defined in (4.30), (4.31) and (4.32) then represents the MBTR feature vector.

$$\text{MBTR}_1^{Z_1}(x) = \sum_l^{|Z_1|} \omega_1^l \mathcal{D}_1^l(x) \quad (4.30)$$

$$\text{MBTR}_2^{Z_1, Z_2}(x) = \sum_l^{|Z_1|} \sum_m^{|Z_2|} \omega_2^{l,m} \mathcal{D}_2^{l,m}(x) \quad (4.31)$$

$$\text{MBTR}_3^{Z_1, Z_2, Z_3}(x) = \sum_l^{|Z_1|} \sum_m^{|Z_2|} \sum_n^{|Z_3|} \omega_3^{l,m,n} \mathcal{D}_3^{l,m,n}(x) \quad (4.32)$$

The indexes l , m , and n run over all atoms with the atomic number Z_1 , Z_2 and Z_3 respectively. The weighting functions ω are used to give more importance to structural motives were the contributing atoms are closer together. Hence, no weighting is typically needed for $k = 1$ as only one atom is involved in that structure motive. Weighting functions for $k = 2$ and $k = 3$ are given in (4.33) and (4.34), respectively.

$$\omega_2^{l,m} = e^{-s_k |\mathbf{R}_l - \mathbf{R}_m|} \quad (4.33)$$

$$\omega_3^{l,m,n} = e^{-s_k (|\mathbf{R}_l - \mathbf{R}_m| + |\mathbf{R}_m - \mathbf{R}_n| + |\mathbf{R}_l - \mathbf{R}_n|)} \quad (4.34)$$

An important parameter to tune the MBTR descriptor is the cutoff distance s_k .

By using only geometry functions with degree $k > 1$ the local many-body tensor representation (LMBTR) can be created as a modification of MBTR.[63]

4.2.6 Atom-centered Symmetry Functions Descriptor

Atom-centered Symmetry Functions (ACSFs)[67] are utilized to represent the local atomic environment through a fingerprint composed of the outputs of multiple two- and three-body functions. These functions are customizable, allowing for the detection of specific structural features. ACSFs encode the configuration of atoms around a central atom, indexed as i , by employing symmetry functions. The presence of atoms neighboring the central atom is captured by three distinct two-body symmetry functions, defined in equations (4.35), (4.36), and (4.37).[63]

$$G_i^{1,Z_1} = \sum_j^{|Z_1|} f_c(R_{ij}) \quad (4.35)$$

$$G_i^{2,Z_1} = \sum_j^{|Z_1|} e^{\eta(R_{ij}-R_s)^2} f_c(R_{ij}) \quad (4.36)$$

$$G_i^{3,Z_1} = \sum_j^{|Z_1|} \cos(\kappa R_{ij}) f_c(R_{ij}) \quad (4.37)$$

In these symmetry functions, the sums are taken over all atoms with atomic number Z_1 , and R_{ij} represents the distance between atoms i and j . The smooth cutoff function f_c , with cutoff radius r_{cut} , is defined as follows:

$$f_c(r) = \frac{1}{2} \left[\cos \left(\pi \frac{r}{r_{\text{cut}}} \right) + 1 \right] \quad (4.38)$$

The parameters η , R_s , κ , and r_{cut} can be adjusted to customize the two-body components of the ACSF descriptor.

Additionally, three-body functions, shown in equations (4.39) and (4.40), can be used to detect specific motifs involving three atoms, with one being the central atom. These functions account for both the angles between atom triplets and their mutual distances within the cutoff radius.

$$G_i^{4,Z_1,Z_2} = 2^{1-\zeta} \sum_{j \neq i}^{|Z_1|} \sum_{k \neq i}^{|Z_2|} (1 + \lambda \cos(\theta))^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \times f_c(R_{ij}) f_c(R_{ik}) f_c(R_{kj}) \quad (4.39)$$

$$G_i^{5,Z_1,Z_2} = 2^{1-\zeta} \sum_{j \neq i}^{|Z_1|} \sum_{k \neq i}^{|Z_2|} (1 + \lambda \cos(\theta))^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2)} \times f_c(R_{ij}) f_c(R_{ik}) \quad (4.40)$$

Similar to the two-body functions, the sums run over all atoms j and k of types Z_1 and Z_2 , respectively, with θ representing the angle between the three atoms i , j , and k . The three-body functions can also be tailored using the user-defined parameters ζ , λ , η , R_s , κ , and r_{cut} . [63]

4.2.7 Autocorrelation Functions as Descriptors

Autocorrelation functions (ACs) as descriptors are solely based on the connectivity of atoms in a molecule and do not require any information on Cartesian coordinates. They are formed by summing the product of atomic properties of atom pairs, as shown in (4.41)[68]:

$$P_d = \sum_i \sum_j P_i P_j \delta(d_{ij}, d) \quad (4.41)$$

P_d is the AC for property P at depth d . The depth describes how many bonds separate the atom pairs. A depth of 0 thus results in a sum over squared properties. δ is the Dirac delta function, and d_{ij} is the bond path between atoms i and j . [69] The five atomic properties used for the AC descriptor, as constructed by Kulik *et al.*, are the nuclear charge, Pauling electronegativity, the atom’s coordination number (the number of neighbors), the identity (which is 1 for any atom), and the covalent atomic radius. [69] The dimensionality of the AC is the product of the number of depths considered and the number of atomic properties used. Therefore, with the five properties mentioned and a maximum depth of 3 (depths 0 to 3 are considered), the resulting descriptor has 20 dimensions. This construction ensures that AC feature vectors are invariant with respect to system size and composition.

4.3 Error Measures

In order to evaluate and compare the performance of a method, a measure—ideally a single number—is required.

The mean signed error (MSE)¹ provides the mean of an error distribution, as shown in (4.42), where y_i is the prediction for point i and \hat{y}_i is its true value. While the MSE reflects the accuracy of a model, it provides no information on its precision (the spread of the errors). Positive and negative errors can cancel each other out.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (4.42)$$

Error cancellation is avoided with the root-mean-square error (RMSE) (see (4.43)), where the square root of the averaged squared errors is taken. As a result, this error measure has the same unit as the predicted property. Because errors contribute by their square, larger errors have a more significant impact, making the RMSE very sensitive to outliers. The RMSE and the root-mean-square deviation (RMSD) are technically evaluated in the same way.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4.43)$$

The mean absolute error (MAE) ((4.44)) is less sensitive to outliers compared to the RMSE. It is the mean of the absolute errors, so no error cancellation occurs as with the MSE.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4.44)$$

¹Note that in some cases the abbreviation MSE is used for the mean squared error. In this work, however, MSE always refers to the mean signed error.

Errors can also be measured using the Q_{95} error measure ((4.45)). Here, the absolute errors are calculated, and the 95th percentile of their distribution is determined. This is the value where 95% of the data points have an absolute error smaller than this value. As a result, outliers are excluded, and both precision and accuracy are considered. Unlike the MAE, Q_{95} communicates probabilistic information: by definition, the probability that an absolute error exceeds Q_{95} is 5%. For the MAE, such statements cannot be made, and the probability that an absolute error exceeds the MAE typically varies between 20% and 50% [70]. While Q_{95} often provides similar rankings compared to the MAE, it also offers more useful information. Moreover, the percentile used for this error measure can be easily tailored to a required risk level [70].

$$Q_{95} = \text{percentile}(|y_i - \hat{y}_i|, 95 \%) \quad (4.45)$$

5 Redox-Active Ligands as a Challenge for Electronic Structure Methods

This chapter has already been published.[71] Unless otherwise noted, all calculations, data analysis and writing have been performed by the main author of this thesis. Carmen Herrmann has supervised and proofread and Axel Jacobi von Wangelin has co-supervised and proofread the work in this chapter.

5.1 Introduction

Following the example of nature, where enzymes containing 3d transition metals (TM) effectively catalyze biochemical transformation reactions,[72] 3d TM catalysis, especially with iron and cobalt complexes, is the subject of intense investigation.[3, 73] In comparison to their established heavier homologues, 3d TMs are more abundant in the earth’s crust, easier to extract, often less toxic and therefore less costly.[74] While precious metal catalysts surrounded by a strong ligand field undergo predictable two-electron processes, first-row transition metals have a weaker ligand field, resulting in a higher tendency for high-spin (HS) complexes.[3] This causes a preference for single electron transfers (SETs), opening up new pathways of radical reaction mechanisms and causing the mechanistic landscape to be more complex than for second- and third-row TM catalysis. Redox-active or redox-non-innocent ligands can equip 3d TM complexes with similar reactivity profiles as the heavier 4d and 5d metal complexes by cooperative electronic effects of ligand and metal ions.[75] Well-known examples of this type of ligands are bis(imino)pyridines, quinones and α -diimines.[4, 76, 77] A huge variety of accessible oxidation and spin-states, paired with reduced stability of the 3d TM complexes compared to their 4d and 5d TM counterparts, makes experimental preparation and analysis more challenging. First-principles calculations become therefore even more important to uncover catalytic properties and the operating reaction mechanisms for 3d TM complexes.[78–81]

Kohn–Sham density functional theory (DFT) is one of the most important methods for this purpose, due to its reasonable accuracy combined with its efficiency.[82] While DFT is in principle exact, in practice its accuracy is limited by a proper choice of an approximate exchange–correlation (xc) functional. Therefore, the development of new xc functionals with larger universal accuracy as well as xc functionals that provide high accuracy for specific properties at the cost of being less transferable are an ongoing area development.[83] Validation studies or error measures when transferring DFT protocols to new classes of systems are necessary.[11, 84–88]

Catalytic reaction mechanisms of 3d TM complexes may involve two or more spin-states during a reaction.[89] Also, the change of spin-state by external stimuli in spin-crossover (SCO) systems may be exploited for applications such as functional materials.[90] Therefore, the determination of accurate spin-state energy splittings is fundamental to a deeper understanding in these areas. Spin-dependent properties are a notorious challenge for DFT.[11, 12] Various studies have investigated the influence of different xc functionals on spin-state energy splittings of 3d TM complexes.[15–22] A strong, mostly linear dependence on the amount of exact exchange admixture in the xc functional was reported.[16, 23, 24] This dependence was investigated for σ -donor and π -acceptor ligands and was found to originate from different metal–ligand interactions in HS and low-spin (LS) complexes.[24, 25]

Spin-state energy splittings ($\Delta E^{\text{HS-LS}}$) in first-principles simulations are the difference between energy of the HS state for the optimized HS molecular structure ($E^{\text{HS}}(\mathbf{R}^{\text{HS}})$) and the energy of the LS state for the optimized LS structure ($E^{\text{LS}}(\mathbf{R}^{\text{LS}})$), $\Delta E^{\text{HS-LS}} = E^{\text{HS}}(\mathbf{R}^{\text{HS}}) - E^{\text{LS}}(\mathbf{R}^{\text{LS}})$. Therefore, positive splittings correspond to the LS state being more stable. Spin-state energy splittings have to be corrected by the zero-point vibrational energies (ZPVE) for the comparison with experimental data. Investigations by Mortensen *et al.* showed that zero-point energies in general favor HS states, with the magnitude varying depending on the ligand.[17] Fortunately, studies by Reiher *et al.* showed that between different functionals, only small changes of $\Delta \text{ZPVE}^{\text{HS-LS}}$ were observed.[91]

In contrast to “pure” xc functionals based on the generalized gradient approximation (GGA), hybrid xc functionals include an admixture of (exact) Hartree Fock (HF) exchange (E_X^{HF}), $E_{\text{XC}}^{\text{hybrid}} = E_{\text{XC}}^{\text{DFT}} + a^{\text{HF}}(E_X^{\text{HF}} - E_X^{\text{DFT}})$. As mentioned above, the spin-state energy splittings depend mostly linearly on the amount of this admixture (a^{HF}). The slope $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a^{\text{HF}}}$ was introduced by Kulik *et al.* as a measure of the sensitivity towards a^{HF} . [16] Generally, pure xc functionals over-stabilize LS states; hybrid functionals with larger amounts of exact exchange a^{HF} highly favor HS states, and therefore slopes are typically negative. This has been suggested to originate from the difference in dealing with the self interaction error (SIE) and with static correlation as described by pure and hybrid xc functionals. [24, 25] The occupied $\sigma(\text{M-L})$ and unoccupied $\sigma^*(\text{M-L})$ orbitals (Figure 5.1) in the LS complex lead to static correlation in LS states. For HS states, $\sigma^*(\text{M-L})$ is already occupied. Therefore, there is no static correlation originating from $\sigma(\text{M-L})$ and $\sigma^*(\text{M-L})$ in the HS state. Pure xc functionals tend to over-stabilize LS states, which has been attributed to them containing some correction to this type of correlation. [24, 25, 92, 93] The energies of $\sigma(\text{M-L})$ and $\sigma^*(\text{M-L})$ are closer for complexes with stronger σ -donor ligands compared to complexes with weaker σ -donors. Therefore, static correlation in LS complexes is increased when stronger σ -donor ligands are present, and spin-state energy splittings have a stronger sensitivity towards a^{HF} (Figure 5.1). [24] In the HS state, the occupation of the anti-bonding $\sigma^*(\text{M-L})$ decreases the M-L bond order, such that the bond length is increased and the covalent character of the M-L bond is decreased. Charge is less delocalized in the M-L bond and more localized on the ligand and on the metal center. In full agreement with this, Kulik *et al.* found the local charges on the metal center to be increased in HS complexes compared to LS complexes. [16] The differences in local charges on the metal between HS and LS states, $\Delta q^{\text{HS-LS}}(\text{M})$, were found to correlate with the sensitivity of $\Delta E^{\text{HS-LS}}$ towards the amount of exact exchange a^{HF} . The decrease of covalency and the accompanying increase of electron localization increases the SIE for HS structures. [25] In contrast to pure xc functionals, SIEs are partly canceled out by the admixture of HF exchange in hybrid functionals. [94] Therefore, hybrid xc functionals stabilize HS states. π -Acceptor ligands cause the formerly non-bonding d^{nb} orbitals to become bonding ($\pi_{\text{M-L}}$) and increase the covalent character of the M-L bond. This should result in an increase of a^{HF} dependence. An example supporting this hypothesis is the reversal of the dependence on a^{HF} for spin-state transitions in a porphin-containing iron(II) complex. Due to the rigidity of the cyclic ligand, M-L bond elongation for HS complexes is restricted and decreases the sensitivity on the amount of exact exchange. For one spin-state transition, a positive slope $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a^{\text{HF}}}$ was found. [25]

Redox-active ligands contain low-lying π -orbitals (π_{L}). The lower in energy these orbitals, the more redox-active the ligand. In comparison to π -acceptor ligands, the π_{L} orbitals of redox-active ligands are more centered on the ligand backbone, and the overlap with the metals d^{nb} is smaller. In contrast to π -back-bonding, the electron is transferred from the metal center and delocalized on ligand. This situation is expected to reduce SIE and to be increasingly favored by lower amounts of exact exchange. As shown in Figure 5.1, this is expected to happen for HS states more likely than for LS states. A change in redox state upon spin-state transition should largely influence the dependence of $\Delta E^{\text{HS-LS}}$ on a^{HF} and change its slope towards less negative or even positive values. For the tris(diimine) iron(II) complexes studied here, the ligands are neutral in the ground state. [95, 96] In their LS state, additional static correlation might occur involving π_{L} and d^{nb} . This should result in more negative slopes for larger redox activity of the ligand.

The considerations discussed above suggest that the metal-ligand interaction with redox-active ligands may cause a strong sensitivity of the spin-state energy splittings toward the amount of exact exchange. While for σ -donor ligands this sensitivity depends, apart from the metal center, mostly on the nature of the donor atom, for redox-active ligands, an additional correlation of the slope with the strength of redox activity would be expected, resulting in a potentially larger spread of slopes than for comparable complexes with redox-innocent ligands. Following Shatruck *et al.*, [97] we investigated a diverse set of

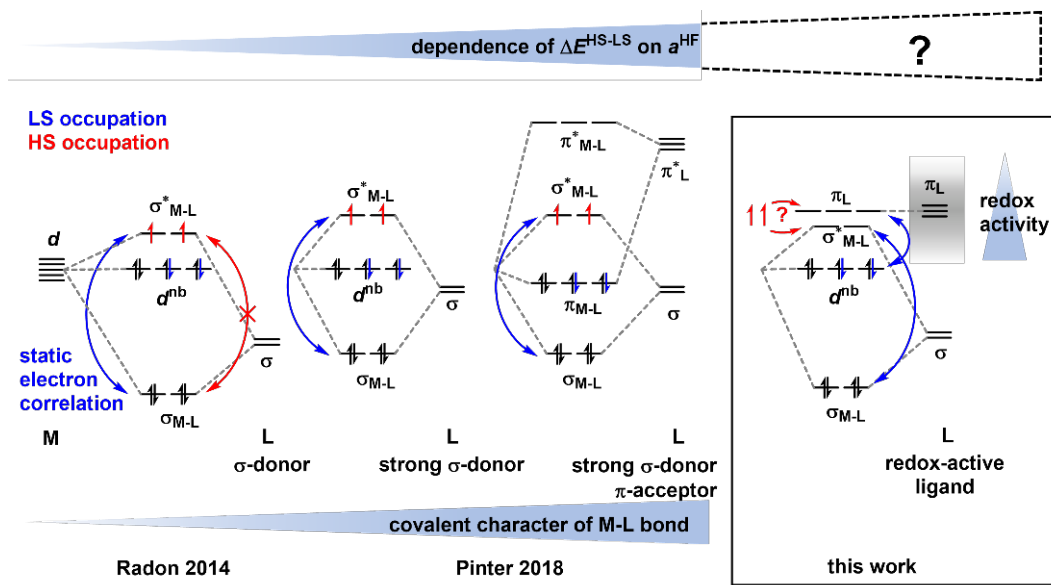


Figure 5.1: Increase of static correlation (blue arrow) for LS complexes compared to HS states (no static correlation, red arrow) with increasing sigma donor strength has been suggested to increase the dependence of spin-state energy splittings on exact exchange admixture. This could imply that for complexes with redox-active ligands, this dependence is even more pronounced than for analogous complexes without such ligands. Reproduced from reference [71].

43 structurally related octahedral iron(II) complexes coordinated by three α -diimine ligands (see Figure 5.3). Ligands of very low (2,2'-bipyridin) to very high redox capacity (bis(imino)acenaphthenes (BIAN)) were included. For these complexes, the individual spin-state preferences (stable LS, stable HS or SCO) are known from the experiment. Homoleptic complexes with uniform metal centers and fixed coordination geometries were chosen to exclude effects by different combinations of ligand environments and metal ions. The uniform diimine bond pattern allowed easy comparability between the ligands.

The uniform bond pattern, together with the correlation of N–N distances in the free ligand with the complexes' ground spin-states[97] also make these systems a promising tool for developing and testing machine learning models. Such approaches have been under intense investigation lately,[98] especially in the prediction of spin-state energy splittings,[35] as quantum chemical calculations like DFT are too computationally expensive to scan large numbers of molecular structures for optimal properties. Not only for machine learning, but also for molecular dynamics simulations and for validation studies, it is important to perform calculations of TM complexes in an automated and robust manner. We show that complexes with redox-active ligands can pose a challenge toward such objectives, as the self-consistent field (SCF) algorithm can be prone to converging to energetically higher-lying solutions in these cases. This also may result in the possible convergence to molecular structures which do not correspond to the lowest-energy one. We discuss signs and the underlying reasons for such problems.

5.2 Computational Details and Data Set Construction

All KS-DFT calculations were performed with the GAUSSIAN 16[99] program package. Molecular structure optimizations were employing the BP86[100, 101] exchange–correlation functional and the D3 version of Grimme’s empirical dispersion corrections with Becke–Johnson damping[102], combined with the def2-TZVP[103, 104] basis set (even though established empirical dispersion corrections do not always improve the theoretical descriptions of transition metal systems [91]). Density fitting with the W06[103, 104] fitting set was employed for the BP86 exchange–correlation functional. Subsequent vibrational frequency calculations were used to confirm the stationary point to be a minimum as well as to evaluate ZPVE corrections. Single-point calculations on optimized structures were carried out with the B3LYP[105], M06[106] TPSSh[107, 108], ω B97XD[109], CAM-B3LYP[62], PBE0[110], BP86[100, 101], PBE[111, 112], TPSS[107], M06L[113] and LC- ω PBE[114, 115] exchange–correlation functionals combined with Ahlrichs’ def2-TZVP[103, 104] basis set. SCF convergence criteria were set to “tight” (change in the root mean square of the density matrix $<10^{-8}$, change in the root mean square of the density matrix $<10^{-8}$, maximal change in the density matrix $<10^{-6}$ and change in the root mean square of the energy $<10^{-6}$ au). Stability analyses of the wave function including relaxing symmetry constraints of orbitals and checking for internal instabilities have been carried out as implemented in GAUSSIAN for selected single-point calculations.[116] Fractional occupation number weighted electron density (FOD) analyses have been performed with the ORCA package version 5.0.2[117–120] (TPSS/def2-TZVP with a smearing temperature of 5000 K).

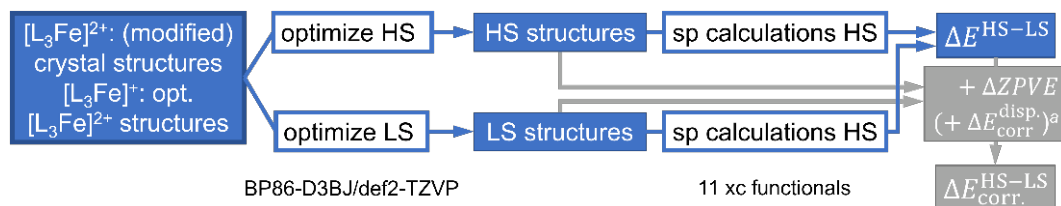


Figure 5.2: Workflow for evaluating spin-state energy splittings ($\Delta E^{\text{HS-LS}}$) and corrected spin-state energy splittings ($\Delta E_{\text{corr.}}^{\text{HS-LS}}$). ^a Dispersion corrections were not added to ω B97XD. Reproduced from reference [71].

Natural Population Analysis using the NBO version 3.1 as implemented in GAUSSIAN 16 was used to calculate atomic partial charges. Atomic (or local) numbers of unpaired electrons ($n^{\alpha-\beta} = n^{\alpha} - n^{\beta}$) are extracted based on Mulliken population analysis. In GAUSSIAN output files, these are (somewhat misleadingly) referred to as atomic “spin densities” (which would imply their multiplication by a factor of 0.5). Raw data for all calculations are available in the NOMAD online repository [121] (<https://nomad-lab.eu>; DOI: 10.17172/NOMAD/2022.07.27-1).

The data set of 43 tris(diimine) iron(II)¹ complexes (Figure 5.3) is based on a collection by Shatruck and coworkers[97] and on the redox-active complex $[(^{\text{Ph}}\text{BIAN})_3\text{Fe}]^{2+}$ (TEDCAJ) from our own research activities.[95] Literature ground spin-states have been determined experimentally by temperature-dependent magnetic measurements.[97] Input structures were taken from the Cambridge Structural Database (CSD) and named after the identifier from this database. If structures were not available in the CSD, input structures were obtained by manual manipulation of similar structures with available crystal structures. This is indicated by “mod” for modified or “simp” for simplified in the structure name.

¹For the reader’s convenience, we refer to complexes of the type $[\text{L}_3\text{Fe}]^{2+}$ as “iron(II)” complexes and those of the type $[\text{L}_3\text{Fe}]^{+}$ as “iron(I)” or “reduced” complexes based on their total charge, even though this might not correctly describe the metal centers’ oxidation state due to possible ligand reduction.

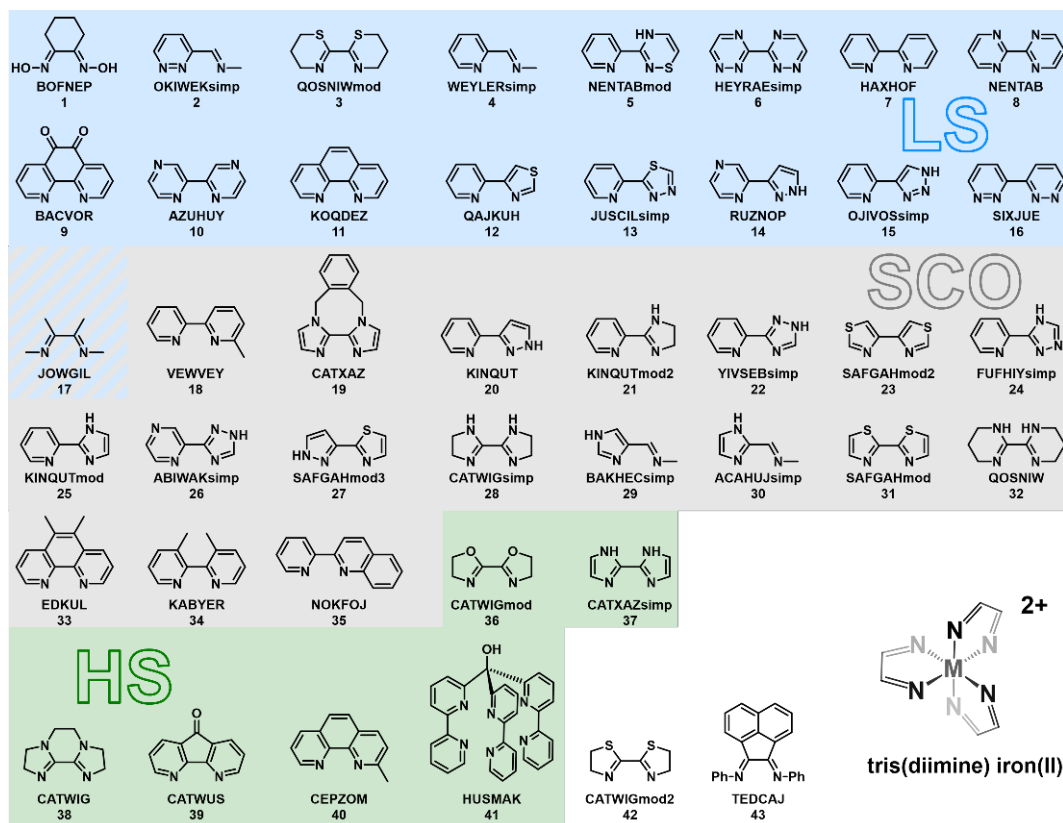


Figure 5.3: Lewis structures of all ligands under study, forming high-spin (HS), low-spin (LS) or spin-crossover (SCO) complexes with Fe(II), and an example for the structure of an octahedral tris(diimine) iron(II) complex with a simplified version of the ligands studied here. For complexes CATWIGmod2 and TEDCAJ, not enough experimental data is available for an unambiguous assignment of a ground spin state. For complex JOWGIL, the SCO is accompanied by the reduction of one ligand by the counter anion.[96] Therefore, a differentiation between LS and SCO was not possible experimentally. Based on the findings by Shatruk *et al.*[97], within the same ground spin-state the complexes are ordered by the N–N distance in the free ligand. Reproduced from reference [71].

The complex NENTABmod contains an asymmetric ligand resulting in two possible diastereomers. Calculation of the second isomer revealed to be not essential for the investigated properties (see SI). Spin-state energy splittings were obtained as shown in Figure 5.2. Molecular structure optimizations of the reduced complexes $[L_3Fe]^+$ were performed starting from the optimized Fe(II) structures.

5.3 Spin-State Energetics in Tris(diimine) Iron(II) Complexes

To investigate the dependence of the spin-state energy splittings of the α -diimine iron(II)¹ complexes on the xc functional, 11 common xc functionals were employed. ZPVE corrections ($\Delta\text{ZPVE}^{\text{HS-LS}}$) are the difference between the ZPVE of the HS state and the LS state. $\Delta\text{ZPVE}^{\text{HS-LS}}$ and empirical dispersion corrections (D3BJ) were evaluated with BP86/def2-TZVP to compare the calculated with the experimental ground state. As the deviations of $\Delta\text{ZPVE}^{\text{HS-LS}}$ among different functionals were found to be considerably small,[91] the BP86 corrections were applied to all xc functionals. The ωB97XD functional already features dispersion corrections. Hence, only ZPVE corrections were added to the energies calculated with this xc functional. The corrected spin-state energy splitting ($\Delta E_{\text{corr.}}^{\text{HS-LS}} = \Delta E^{\text{HS-LS}} + \Delta\text{ZPVE}^{\text{HS-LS}} + \Delta E_{\text{disp}}^{\text{HS-LS}}$) of the tris(diimine) iron(II) complexes are shown in Figure 5.4. Except for one case (NENTABmod (**5**)), the orderings of the spin-state energy splittings calculated with the different xc functionals are the same for each complex. As discussed in Section 5.4, the exceptional behavior was also observed for BACVOR **B** which is the HS structure of BACVOR (**9**) converged to a local minimum.

We will now compare the calculated spin-state energy splitting energies with the experimentally determined ground spin-states. We consider HS states predicted correctly if $\Delta E_{\text{corr.}}^{\text{HS-LS}} < 0$ and LS and SCO² predicted correctly if $\Delta E_{\text{corr.}}^{\text{HS-LS}} > 0$. As observed in previous studies[16, 24, 25, 92, 93], the pure xc functionals favor LS ground states, whereas hybrid xc functionals favor HS states. Accordingly, TPSS, BP86, and PBE show always positive values for the spin-state energy splittings. Also, the hybrid functional TPSSH (10 % HF exchange) over-stabilizes LS states. In contrast to this, the hybrid functionals with a large amount of exact exchange, PBE0 and M06, with 25 % and 27 % HF exchange, respectively, strongly favor HS states for the majority of LS complexes. The best agreement with the experiment was achieved by the highly parameterized M06L xc functional, which only predicts 4 of the SCO complexes to be HS and complex HUSMAK to be LS instead of HS. Also the LC- ωPBE xc functional, which only predicts complex RUZNOP to be HS instead of LS but struggles with the harder to predict SCO complexes, performs well. It should be mentioned that the data set is a little biased towards LS and SCO complexes, which might result in xc functionals with smaller amounts of exact exchange performing better. Despite great uncertainty due to the large spread of $\Delta E^{\text{HS-LS}}$ between the chosen xc functionals, trends among the complexes are described well by DFT.

As the thermodynamic and dispersion corrections are a constant for each complex, they do not influence the spread among the xc functionals and are neglected in further discussions. Only the dispersion already included in ωB97XD is kept. As reported by others,[16, 23, 24] linear dependence of $\Delta E^{\text{HS-LS}}$ on a^{HF} was found (Figure 5.5 (a)). While TPSS, BP86, and PBE gave similar spin-state energy splittings, the highly parameterized M06L xc functional (with $a^{\text{HF}} = 0$) gave significantly lower spin-state energy splittings and therefore worsened the linear fit of $\Delta E^{\text{HS-LS}}$ as a function of a^{HF} and was therefore not considered. Except for NENTABmod and BACVOR **B** (with a local minimal molecular HS structure), negative slopes between $-4.75 \frac{\text{kJ mol}^{-1}}{\%}$ and $-6.75 \frac{\text{kJ mol}^{-1}}{\%}$ were observed for our set of complexes, with $R^2 = 0.996 \pm 0.002$. This implies spin-state energy splittings spread by averagely 145 kJ mol^{-1} among the common xc functionals tested here in the moderate regime of 0 % to 27 % HF exchange. Although the range-separated xc functionals do not have a fixed amount of HF exchange, they, based on their $\Delta E^{\text{HS-LS}}$, behave like hybrid xc functionals

²SCO implies that the ground state has low spin multiplicity, but the energy difference to the high-spin-state is so small that the entropic factors favoring this high-spin-state can make it the dominantly populated variant at higher temperatures accessible with standard experimental techniques. Predicting the temperature at which SCO occurs is beyond the scope of this work.

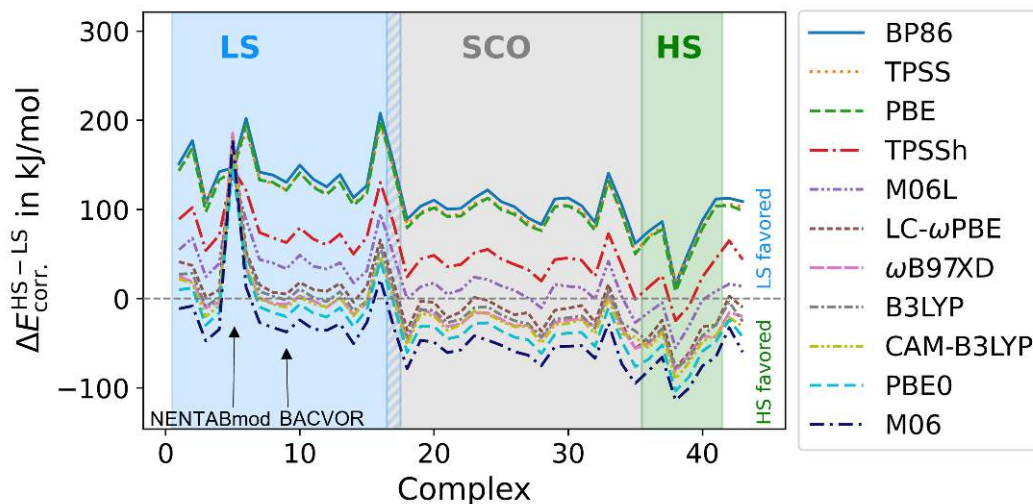


Figure 5.4: Spin-state energy splittings including ZPVE and dispersion corrections of tris(diimine) iron(II) complexes evaluated from single-point calculations with different xc functionals (see workflow in Figure 5.2). The labels of the xc functionals are ordered in the same manner as their spin-state energy splittings. For each complex, the molecular HS and LS structures were optimized with BP86-D3BJ/def2TZVP. For BACVOR, preoptimization of the molecular HS structure with M06-D3/def2-TZVP was necessary. Molecular structure optimization of the HS structure starting from the crystal structure led to the local minimum, structure BACVOR **B**, which is not included in this plot. Reproduced from reference [71].

with an a^{HF} of 19.7(6) % (LC- ω PBE), 22.2(8) % (CAM-B3LYP) and 19.7(9) % (ω B97XD). M06L behaves like a hybrid xc functional with 15.3(6) % exact exchange. Interestingly, the positive slopes for the two outliers, NENTABmod ($0.77 \frac{\text{kJ mol}^{-1}}{\%}$, $R^2 = 0.561$) and BACVOR **B** ($4.83 \frac{\text{kJ mol}^{-1}}{\%}$, $R^2 = 0.972$) imply the LS state is increasingly preferred at larger amounts of exact exchange. Stability analyses were performed on the single-point calculations of NENTABmod and BACVOR (**B**) and confirmed all LS and HS wave functions to have no internal instability. $\langle S^2 \rangle$ values are provided in Section B.3 and show no significant deviations from their ideal values, i.e., no significant multi-reference character of the wave functions. However, diagnosing static correlation by fractional occupation number weighted electron density (FOD)[120] of all HS and LS complexes revealed that the number of strongly correlated electrons ($N(\text{FOD})$) correlates with the size of the HOMO–LUMO gaps of the free ligands (see Figure B.14). An increased redox activity (smaller HOMO–LUMO gaps) of the ligand is accompanied with an increased number of strongly correlated electrons. The largest amount of strongly correlated electrons has been found for NENTABmod and BACVOR. For these complexes with strongly redox-active ligands, strongly correlated electrons are also delocalized on the ligand backbone, while they are mostly localized on the metal center for complexes with redox-innocent ligands like HAXHOF (see Figure B.9).

We expected the redox activity of the ligands to have an impact on the slope, which, as reported by Kulik and coworkers,[16] depends on changes in local charges on the iron center between HS and LS structure. For the complexes discussed, the linear fit of the slopes as a function of $\Delta q^{\text{HS-LS}}$ showed, at first sight, a correlation with $R^2 = 0.880$ (see Figure B.13). However, excluding NENTABmod, which showed much lower $\Delta q^{\text{HS-LS}}$ than the other complexes, resulted in a rather poor correlation with $R^2 = 0.165$. This suggests the structures are too similar to have a large enough variety of $\Delta q^{\text{HS-LS}}$ to show a correlation

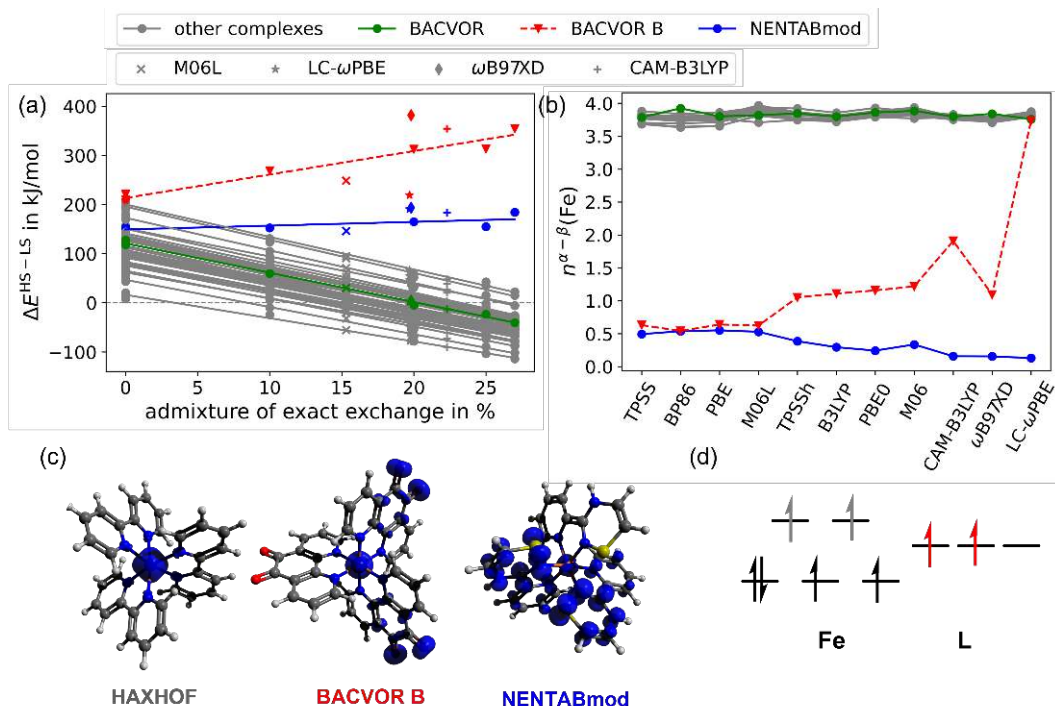


Figure 5.5: (a) Spin-state energy splittings of BP86-D3BJ/def2TZVP-optimized tris(diimine) iron(II) complexes as function of exact exchange admixture a^{HF} . For linear regression used xc functionals with $a^{\text{HF}}/\%$ given in parenthesis: BP86 (0), TPSS (0), PBE (0), TPSSH (10), B3LYP (20), PBE0 (25), M06 (27) (all three values for $a^{\text{HF}}=0$ were treated individually and no averaging was performed). The spin-state energy splitting for M06L (x), LC- ω PBE (star), ω B97XD (rhombus) and CAM-B3LYP (+) are given at admixtures of exact exchange obtained from the linear regression. (b) Mulliken numbers of unpaired electrons on iron centers calculated with different xc functionals and the BP86-D3BJ/def2TZVP-optimized HS tris(diimine) iron(II) complexes; (c) Spin density of selected HS complexes (B3LYP/def2-TZVP, iso surface value = $0.01 e a_0^{-3}$); (d) qualitative MO scheme for Fe(II) HS complexes localizing electrons on the ligands (red) or on the iron atom (grey). Reproduced from reference [71].

to the slope, which supports Kulik's suggestion for the slope to be mainly influenced by the type and oxidation state of the metal center and the type of the donor atom.[16] In addition, as a measure for the redox activity of the ligands, we evaluated electron affinities (EA) and HOMO-LUMO gaps of the free ligands. Electron affinities were evaluated as energy differences between the neutral ligand and the radical anion, $\text{EA} = E(\text{Ligand}^-) - E(\text{Ligand}^0)$, optimized in their respective charge states.[122, 123] No systematic correlation of the slope ($\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a^{\text{HF}}}$) or the spin-state energy splitting with the electron affinity or with the HOMO-LUMO gap of the free ligand was found (see Figure B.13). The ligand of the outlier BACVOR has the largest electron affinity with -2.32 eV . But, the ligand of the other outlier complex NENTABmod shows just an average electron affinity of 0.66 kJ mol^{-1} . Nevertheless, the two outliers NENTABmod and BACVOR exhibit the lowest HOMO-LUMO gaps (1.57 eV and 1.92 eV , respectively). This indicates that the redox activity could be connected to the conspicuous behavior of both complexes.

5.4 SCF Convergence to Local Minima

As we will discuss in this chapter, the positive slope $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial \alpha_{\text{HF}}}$ is caused by the reduction of the ligands upon spin-state transition. For BACVOR **B**, this is associated with SCF convergence of the HS structure to a local minimum. However, for NENTABmod no data suggesting the reduction of the ligands being the incorrect description of the HS state has been found. Although this might seem like a trivial technical problem, we consider it noteworthy as on the one hand, such artifacts may go unnoticed when predicting properties for individual structures, and on the other hand, this poses a problem for evaluating large amounts of structures for purposes ranging from machine learning to method validation, in which it is unrealistic to check every calculation by hand.

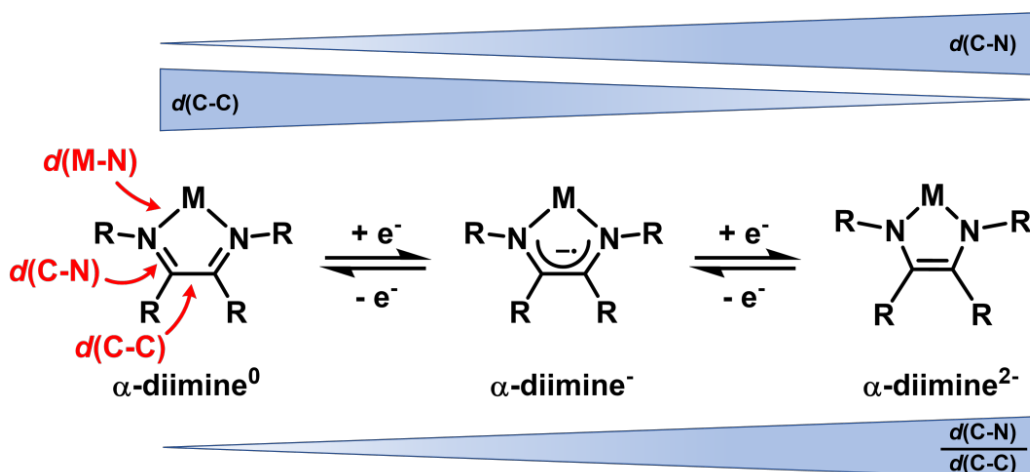


Figure 5.6: Characteristic bond length trends in reduced α -diimine ligands. Reproduced from reference [71].

In this section, we will discuss the observed abnormalities of the spin-state energy splitting of NENTABmod and BACVOR **B** by first looking at characteristic bond lengths and bond patterns, and then confirm involvement of redox-activity of the ligand by analyzing atomic numbers of unpaired electrons. As discussed in Section 5.1, HS states are destabilized by pure xc functionals due to increased SIE by reduced covalency of the metal–ligand bond of HS structures in comparison to LS structures.[25] The metal–ligand distance $d(\text{Fe-N})$ (see Figure 5.6) is expected to be larger for HS complexes than for LS complexes, and thus $\Delta d^{\text{HS-LS}}(\text{Fe-N})$ should be positive. This is observed for all complexes except for NENTABmod and BACVOR **B**. On average, Fe–N bonds are elongated by 0.17 Å to 0.22 Å in the HS structure compared to the LS structures. For BACVOR **B** and NENTABmod, there is almost no change in bond length comparing the HS and LS structures: Their $\Delta d^{\text{HS-LS}}(\text{Fe-N})$ are 0.2 pm and –0.3 pm. Small changes in Fe–N bond length were also considered the reason for the positive slope of a spin-state transition of the porphyrin complex discussed in Section 5.1.[25] This suggests a comparatively high covalent character of the Fe–N bonds in the HS and LS structures, and therefore no increase of SIE in the HS structure. This explains the untypical dependence of $\Delta E^{\text{HS-LS}}$ on HF admixture.[25] Crystal structures of Fe(II) HS complexes with anionic α -diimine ligands have also shown similar short Fe–N bond length.[124, 125]

Another important way in which the electronic structure is reflected in the molecular structure are bond-length patterns in the α -diimine ligands: Upon reduction of these ligands, the imine bonds increase and the C–C bond decreases (Figure 5.6). These trends are captured in the ratio of the imine to the C–C bond length ($\frac{d(\text{C-N})}{d(\text{C-C})}$), which increases

upon reduction. For all structures, the average value of this ratio ranges from 0.862 (TEDCAJ HS) to 0.967 (EDKUL LS). Among the HS structures, the average ratio $\frac{d(\text{C-N})}{d(\text{C-C})}$ is largest for the two outliers, 0.955 (BACVOR **B**) and 0.957 (NENTABmod). The two complexes are the only ones showing a larger average ratio for the HS structure than for LS structure. Remarkably, for these two structures, two out of the three ligands have a larger ratio than the remaining one. This implies that two of them are more reduced than the other. This again suggests the redox activity of the ligands to be key to the different sensitivity of $\Delta E^{\text{HS-LS}}$ to the amount of exact exchange.

Another way of analyzing the redox state of the ligands in the HS structures is to look at the numbers of unpaired electrons on the atoms. In comparison to atomic partial charges, these have the advantage of being much less dependent on the choice of the local projector (see Figure B.3).[126] A one-electron reduction of the ligand results in a radical anion. Therefore, spin density should be found on the ligand, and the local spin density on the iron center should be lower in the case of ligand reduction. Without redox-active ligands, a number of around 4 unpaired electrons on the metal center is expected for d^6 HS complexes (Figure 5.5 (d) grey). This is indeed the case for the HS state of all complexes, except for the two outliers BACVOR **B** and NENTABmod. Their HS states have less than 2 unpaired electrons on the iron center (Figure 5.5 (b)). In agreement with the structural characteristics of the HS structures, unpaired electrons are localized on two of the ligands (Figure 5.5 (c) and (d) red). As an exception, for BACVOR **B**, LC- ω PBE gives a large number of 3.74 unpaired electrons on Fe, in contrast to the other xc functionals.

Taking the molecular orbitals (MOs) of LC- ω PBE as an initial guess for the single-point calculations of the HS structure with the other xc functionals not only leads to a lower SCF energy for M06L, the hybrid, and range-separated xc functionals, but also changes the dependency of $\Delta E^{\text{HS-LS}}$ on a^{HF} (Table 5.1). The corrected $\Delta E^{\text{HS-LS}}$ is now decreasing with a^{HF} as it does for the others. For TPSS, the SCF did not converge, and for the other two BP86 and PBE, the same solution as before was obtained for both types of initial guess. However, stability analysis of all HS wave functions with reduced ligands revealed them to be stable. This suggests that for BACVOR **B**'s HS state, two SCF minima close in energy are present: one with spin density localized mainly on the iron center (**A**), and one with spin density partially localized on the ligands (**B**) (Figure 5.7 (a, red dashed line)). The presence of a second minimum leads, in this case, to SCF convergence to a local rather than the global minimum. This might be a general problem for complexes with highly redox-active ligands.

Table 5.1: Mulliken numbers of unpaired electrons on iron center $n^{\alpha-\beta}(\text{Fe})$ of complex BACVOR **B**'s Fe(II) HS structure and spin-state energy splittings $\Delta E^{\text{HS-LS}}$ of complex BACVOR **B** determined with the xc functionals converging to different SCF minima with the default initial guess^a (**1**) and with electron density of LC- ω PBE as initial guess (**2**).

	$\Delta E^{\text{HS-LS}}$ in kJ mol ⁻¹ (1)	$\Delta E^{\text{HS-LS}}$ in kJ mol ⁻¹ (2)	$n^{\alpha-\beta}(\text{Fe})$ (1)	$n^{\alpha-\beta}(\text{Fe})$ (2)
M06L	248.7	216.8	0.54	3.91
TPSSh	268.7	251.8	1.05	3.78
B3LYP	312.4	225.1	1.11	3.73
PBE0	313.5	180.4	1.16	3.81
M06	354.2	151.8	1.22	3.94
CAM-B3LYP	354.3	212.9	1.91	3.70
ω B97XD	382.0	216.1	1.09	3.66

^a Diagonalization of the Harris functional[127]

Finding a lower SCF energy for M06L, TPSSh, B3LYP, PBE0, M06, CAM-B3LYP and

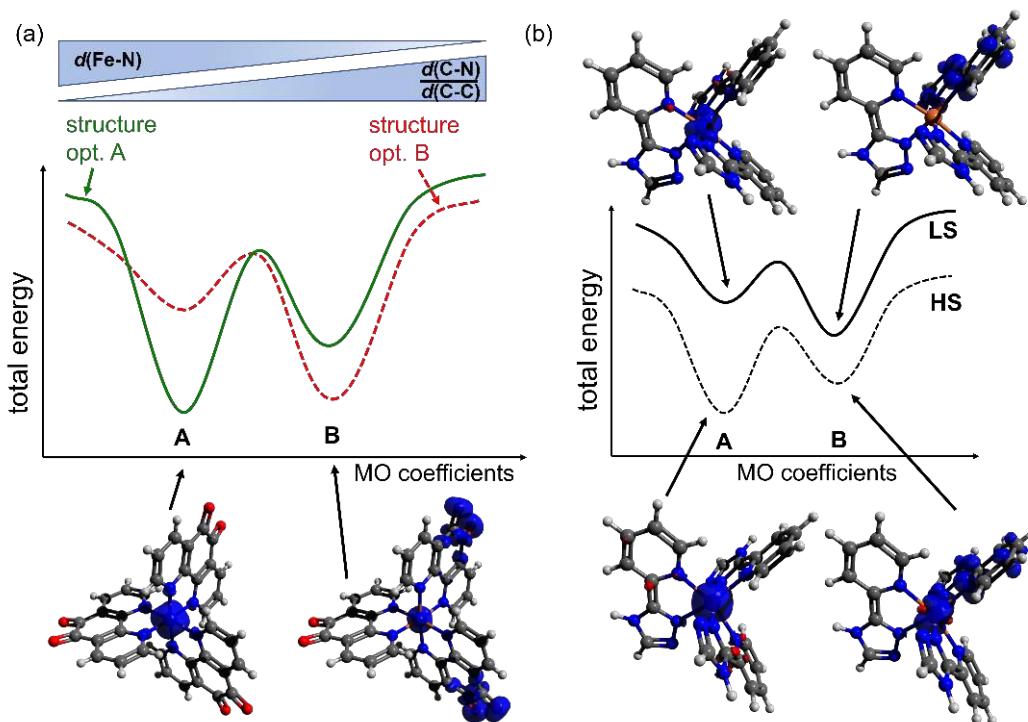


Figure 5.7: Exemplary illustration of a local and global minima for the SCF total energy as a function of the MO coefficients: (a) dependence of electronic-structure minima for BACVOR's $[\text{L}_3\text{Fe}]^{2+}$ on molecular HS structure: Molecular structure optimization from electronic-structure minimum **A** (with all spin density on the metal center) leads to molecular structure **A** (green, solid); Molecular structure optimization from electronic-structure minimum **B** (with two ligands reduced and thus carrying spin density) leads to molecular structure **B** (red, dashed). (b) minima for HS (dashed line) and LS (solid line) structures of $[\text{L}_3\text{Fe}]^+$. Spin densities from FUFHIYsimp (**24**) were evaluated with $\omega\text{B97XD}/\text{def-2TZVP}$ (iso surface value = 0.01). Reproduced from reference [71].

ωB97XD proves that the minimum **B** is a local minimum for these approximate xc functionals. For TPSS, no clear statement is possible due convergence issues and for BP86, PBE, and TPSS, convergence to minimum **B** in all cases suggests that, in contrast to the other xc functionals, these xc functionals consider this minimum a global minimum of the total SCF energy. The molecular structure is expected to influence the energy difference between both minima **A** and **B**, and vice versa, molecular structure optimization will be influenced by SCF convergence to either **A** or **B** (Figure 5.7 (a)). Similar considerations are relevant in other contexts, for example for structural diradical character [128] and for mixed-valence systems [129, 130]).

As discussed above, our standard protocol the molecular structure optimization of BACVOR with BP86 starting from the crystal structure, yields a HS structure with larger $\frac{d(\text{C-N})}{d(\text{C-C})}$ for two of the three ligands and shorter Fe–N bond length compared to the other HS structures (BACVOR **B**, Figure 5.7 (a) red dashed line). This fits to the SCF minimum **B**, where the ligands are more reduced compared to **A**. Preoptimization of the molecular structure using M06-D3/def2-TZVP (a single-point calculation with this xc functional on the crystal structure leads to the electronic structure **A**) and subsequent optimization with BP86-D3BJ/def2TZVP results in a structure with an average Fe–L bond length of 2.16 Å and an average $\frac{d(\text{C-N})}{d(\text{C-C})}$ of 0.92 that fits the SCF minimum **A** (Figure 5.7 (a) green

solid line). A subsequent single-point calculation using BP86-D3BJ/def2-TZVP revealed a SCF convergence to minimum **A**. In contrast to the HS structure obtained by BP86-D3BJ optimization, the energy of this new molecular HS structure is lower by 67.3 kJ mol⁻¹ (as evaluated with BP86-D3BJ). This shows that molecular structure optimization might lead to a non-minimum molecular structure due to the presence of the two SCF minima. A test using the ORCA package version 5.0.2[117–119] for molecular structure optimization of BACVOR (HS) starting from the crystal structure resulted in the identical structure as obtained with GAUSSIAN and excluded the observed structure resulting from an artifact of the quantum chemistry program package.

The difference in the energy between the two HS structures BACVOR **A** and BACVOR **B** is larger for larger amounts of exact exchange (Figure 5.5 (a) red and red dashed line). Therefore, the pure xc functional BP86 is more prone to erroneous convergence of the SCF algorithm compared to hybrid xc functionals. However, a single-point calculation with TPSSh on the M06-D3/def2-TZVP-preoptimized BACVOR HS molecular structure **A** still converged to a local SCF minimum, indicated by the reduced number of electrons on the iron center. Thus, despite the improved HS structure, both SCF minima **A** and **B** are still close and can lead to erroneous SCF convergence.

5.5 Reduced Tris(diimine) Iron Complexes as a Challenging Test Case

So far, we have seen erroneous SCF convergence and molecular structure optimization towards a local minimum for strongly redox-active ligands in Fe(II) complexes. As we have discussed in Section 5.1, those ligands have low-lying π -orbitals, and electrons from a metal center might be localized on these. This reduction of the ligand can also be favored by a more reduced metal center with higher-lying d -orbitals. Therefore, reduction of the complexes should lead to the errors discussed above for an increased amount of complexes. Formal Fe(I) complexes $[\text{L}_3\text{Fe}]^+$ were optimized in HS (quartet) and LS (doublet) states, and spin-state energy splittings were evaluated to test this hypothesis (Figure 5.8(a)). For complex NOKFOJ (**35**) the coordination number dropped from 6 to 5 upon molecular structure optimization for the HS as well as the LS structure. As these complexes were no longer octahedral, NOKFOJ has been left out for the $[\text{L}_3\text{Fe}]^+$ complexes. If the additional electron is placed on a ligand instead of on the iron center, the resulting HS complex $[\text{L}_2\text{L}^{\bullet}\text{Fe}^{\text{II}}]^+$ represents an antiferromagnetically coupled quartet (see Figure B.4). For this reason, broken symmetry calculations with an fragment based initial guess (fragment 1: Fe (+2, quintet), fragment 2: all ligand atoms (-1, triplet)) were performed on the BACVOR, NENTABmod, and HAXHOF HS complexes (B3LYP/def2-TZVP) as exemplary test cases. However, no lower energies were obtained for these cases. In contrast to the $[\text{L}_3\text{Fe}]^{+2}$ complexes, spin contamination and the number of strongly correlated electrons evaluated by FOD analysis are considerably increased for the reduced complexes (see Figure B.14). This is consistent with studies by Head-Gordon *et al.*, according to which TM complexes in low oxidation states tend to show this spin symmetry breaking and multi-reference character.[131]

SCF convergence to saddle points in MO coefficient space can be identified by an internal instability of the wave function. Stability analyses have been performed for all single point calculations on the reduced complexes. Convergence to such saddle points (168 cases) as well as local SCF minima (27 cases) poses a challenge for the calculation of these reduced systems (Figure 5.8 (b)). Especially the range-separated xc functionals are prone to this behavior. While standardized stability analysis can identify convergence to saddle points, local minima are not recognized. Potentially problematic cases can be identified by large positive spin-state energy splittings deviating from the above observed ordering among the

xc functionals, combined with decreased numbers of unpaired electrons on the iron center (Figure B.5) for the HS structures. Taking orbitals from other xc functionals (BP86 (HS) or B3LYP (LS)) as an initial guess does not only yield a lower energy for such cases but also poses a more cost efficient alternative to the optimization of the wave function in cases where SCF convergence to saddle points occurred.

As for the Fe(II) complexes ($[\text{L}_3\text{Fe}]^{2+}$), the ordering of the spin-state energy splittings as a function of exact exchange admixture is reversed for the reduced complex of NENTABmod as well as BACVOR. Comparing their HS and LS structures reveals that the Fe–N bond lengths are almost the same for their HS and LS structures. Also, they are the only complexes with larger average $\frac{d(\text{C}-\text{N})}{d(\text{C}-\text{C})}$ in the HS structures.

Largely deviating numbers of unpaired electrons from the ideal values for iron (II) centers are indicators for convergence to local minima or saddle point that have large impact on the total SCF energy. They indicate a potentially wrong description of the redox state of the iron center and the ligand. Properly defining the metal oxidation state is a topic of great interest in current research and could be a helpful means for further analysis of such complexes.[132, 133] Numbers of unpaired electrons on iron of around 0 for LS structures and around 3.5 for HS structures suggest that the additional electron is localized on the ligands, while the iron centers remain in oxidation state two for all complexes despite their formal Fe(I) character. At the example of spin densities of FUFHIYsimp (**24**) in Figure 5.7 (b) (dashed line), it can be seen that the global SCF minimum of the HS structures is described best by minimum **A**. The local minimum has much less spin density localized on the metal center and can be described by minimum **B** (Figure 5.7 (b)). In contrast to that, for the LS structure, the local minimum localizes more spin density on the metal center (**A**), and the global minimum localizes spin density on the ligands (**B**) (Figure 5.7 (b) (solid line)). This implies the ligands are less reduced at the local minimum. As expected for the reduced complexes, the increased involvement of the ligands’ redox activity makes convergence to the global SCF minimum more challenging.

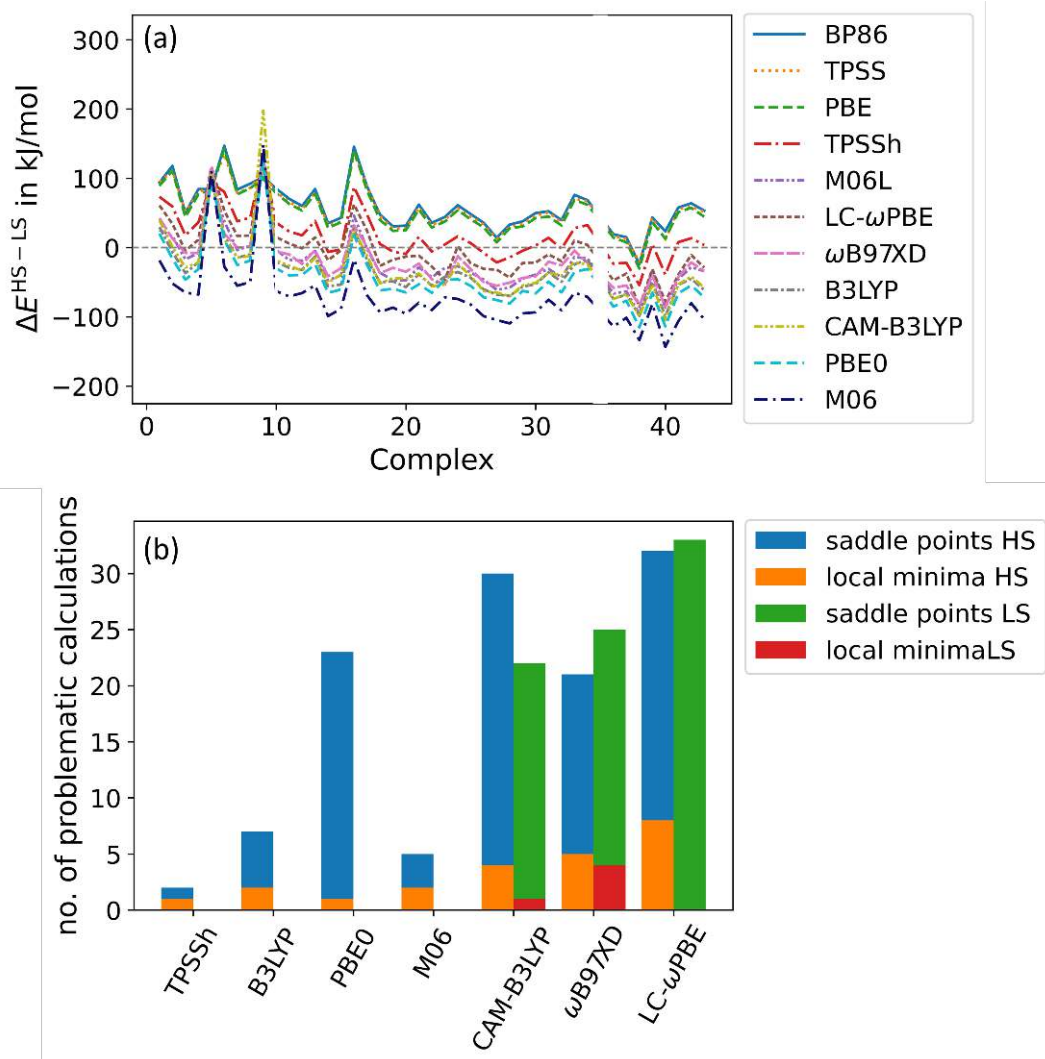


Figure 5.8: (a) Spin-state energy splittings of reduced tris(diimine) complexes $[\text{L}_3\text{Fe}]^+$, evaluated from single-point calculations with different xc functionals; Complex NOKFOJ (**35**) has been left out due to decoordination of one of the ligands. The labels of the xc functionals are ordered in the same manner as their spin-state energy splittings. For each complex, the molecular HS and LS structures were optimized with BP86-D3BJ/def2TZVP. (b) Numbers of single point calculation where the SCF convergence towards local minima or saddle points has been observed for HS and LS $[\text{L}_3\text{Fe}]^+$ complexes. 924 single point calculations have been performed for the reduced complexes. Of these 169 converged to saddle points and 28 to local minima. Reproduced from reference [71].

5.6 Conclusions

The sensitivity of spin-state energy splittings on the exact exchange admixture ($\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$) is known to be dependent on metal–ligand interactions. The redox activity of ligands is related to the energies of their unoccupied π -orbitals and is expected to influence the strength of this metal–ligand interaction. In contrast to what one might expect based on these considerations, redox activity of the ligands does not correlate much with $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$ for our set of homoleptic octahedral tris(diimine) iron(II) complexes. $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$ is generally highly negative for all of them. The spread of the spin-state energy splittings between pure xc functionals (0% exact exchange) and M06 (27% exact exchange) ranges for the $[\text{L}_3\text{Fe}]^{2+}$ complexes from 134.4 to 189.5 kJ mol^{−1}, with an average of 145 kJ mol^{−1}, and from 91.2 to 174.6 kJ mol^{−1} for the reduced $[\text{L}_3\text{Fe}]^+$ complexes, with an average of 116 kJ mol^{−1}. Regardless of the large spread among the xc functionals, with the right choice of functional, the experimentally determined ground spin-states are reproduced well for the investigated complexes. The best performing functionals here are M06L and LC- ω PBE, which correspond to an exact exchange admixture of around 15 to 19 percent.

In two cases, we find a change of the oxidation state of the ligand when going from one spin-state to the other. This is accompanied by a positive slope $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$. The two cases are the iron(II) complexes BACVOR and NENTABmod. Such transitions are in principle chemically reasonable, as HS states are expected to reduce the ligand more likely than LS states (Figure 5.1). However, we found that for BACVOR, the reduction of the ligand originates from SCF convergence towards a local minimum which is around 35 kJ mol^{−1} higher in energy than the global one. Starting from the “wrong” SCF solution leads to an optimized molecular structure for BACVOR’s HS state which was not the global minimum. Conversely, a “good” molecular structure does not guarantee convergence to the matching lowest-energy SCF solution, as we observed for the HS structure of complex BACVOR for some functionals. For NENTABmod, only the “outlier” positive-slope SCF minimum could be converged, so that the question of chemically reasonable structure vs. SCF artifact could not be answered. In the reduced complexes $[\text{L}_3\text{Fe}]^+$, where the redox activity of the ligand is more important, we observed a strong increase of instances of SCF convergence towards local minima in our data set. One might argue that this is associated with stronger static correlation in complexes with redox-active ligands, and that approximate DFT is simply not suitable for such systems and multireference methods should be employed instead. However, at the time being DFT is the only method which can optimize molecular structures for such systems with reasonable accuracy and in a routine fashion.

Correlated electronic structure methods such as density matrix renormalization group (DMRG) might be suitable to avoid such troubles.[134–136] However, these methods are too expensive for dealing with large amounts of complexes or for routine structure optimizations. Fortunately, critical cases can be detected by deviations from the majority of complexes in a data set, in particular (1) in spin-state bond lengths ratios (e.g., in our case, Fe–N bond lengths in HS and LS structures are very similar for the two critical cases, while they deviate considerably for the majority of cases), (2) in numbers of unpaired electrons on the metal center and (3) in the sign of $\frac{\partial \Delta E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$. Of course, such deviations could always point to an interesting chemical structure rather than an SCF artifact. But first, for a set of chemically related systems as studied here, artifacts would not appear too unlikely an explanation, and second, if they were “chemically real”, it would still be interesting to automatically scan for such interesting observations in data sets too large to scrutinize by hand. In such automated protocols, (1) and (2) can be implemented directly as cost-efficient tests. The computational cost for the check of (3) could be achieved by a quick screen of $\Delta E^{\text{HS-LS}}$ by single-point calculations for several amounts of exact exchange, without optimizing molecular orbitals. For potentially critical cases, we also recommend testing several initial guesses with reduced and not reduced ligands based on, e.g., fragment guesses

as implemented in JAGUAR[137] or employing restrained optimizations.[138] Also, varying characteristic structural features such as Fe-L bond lengths in starting structures can push molecular structure optimization as well as SCF convergence towards different potentially relevant minima.

6 Machine Learning Guided Strategies Towards Spin-Gaps in Carbenes

6.1 Introduction

Carbenes are pivotal intermediates in numerous organic chemistry reactions[26, 27], distinguished by a divalent carbon atom with two non-bonding orbitals containing two electrons. Depending on the spin configuration of these electrons, carbenes can exist either as singlet states, where the electrons are paired, or as triplet states, where the electrons remain unpaired. The spin state of a reaction intermediate or reactant significantly influences the reaction pathway and outcome [28, 29]. While stable singlet carbenes have been synthesized [29, 139, 140], carbenes are typically highly reactive and exhibit short lifetimes, complicating their experimental characterization. Consequently, computational methods have become essential for studying these reactive species [30–33].

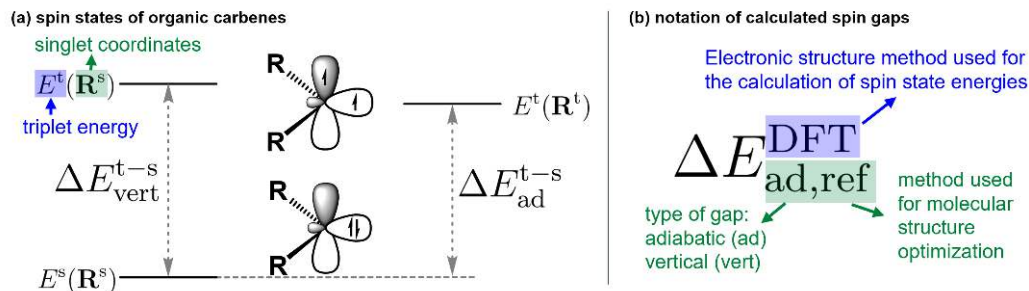


Figure 6.1: a) Illustration of adiabatic (ad) and vertical (vert) spin gaps. Adiabatic spin gaps are the difference between energies of the triplet state for the optimized triplet molecular structure ($E^t(\mathbf{R}^t)$) and the energy of the singlet state for the optimized singlet structure ($E^s(\mathbf{R}^s)$), $\Delta E_{\text{ad}}^{t-s} = E^t(\mathbf{R}^t) - E^s(\mathbf{R}^s)$. Vertical spin gaps are the difference between energies of the triplet state for the optimized singlet molecular structure ($E^t(\mathbf{R}^s)$) and $E^s(\mathbf{R}^s)$ ($\Delta E_{\text{vert}}^{t-s} = E^t(\mathbf{R}^s) - E^s(\mathbf{R}^s)$). b) Explanation of the notation of the differently calculated spin gaps used in this work. ΔE always refers to the energy difference between a triplet and a singlet spin set (a spin gap). The superscript (s-t) as used in a) is omitted for clarity. The superscript refers to the method used for the energy calculation of the triplet, as well as singlet state (single point calculations). No superscript refers to the reference spin gaps. The index contains information on the method used for the optimization of molecular structures. Vertical spin gaps are calculated on singlet structures only. Adiabatic spin gaps use molecular structures optimized in singlet as well as triplet states. Reference structures (ref) are optimized by CASSCF (singlet) and B3LYP (triplet). Blue highlighting is used for calculations used for the spin state energies, and green highlighting is used when referring to molecular structure properties.

The importance of carbenes prompted the group of Lilienfeld *et al.* to undertake a large-scale study on the electronic structures of carbenes, resulting in the creation of the QMspin dataset [1]. Derived from the QM9 dataset [141], QMspin includes organic molecules containing up to nine heavy atoms (C, N, O, F) and provides multireference level spin gaps of carbenes, which were categorized based on their neighboring groups. These groups range from σ -electron-withdrawing and π -electron-donating neighbors, to aliphatic and aromatic α -substituents, and even those with σ - and π -electron-withdrawing characters, such as cyano and carbonyl groups. The dataset spans a wide range of adiabatic spin gaps, from -70 kJ mol^{-1} to 300 kJ mol^{-1} . These gaps are defined as the energy difference between the triplet state, with its corresponding optimized molecular structure, and the singlet state, with its own optimized structure ($\Delta E_{\text{ad}}^{t-s} = E^t(\mathbf{R}^t) - E^s(\mathbf{R}^s)$, Figure 6.1 a). For the QMspin dataset CASSCF and DFT with the B3LYP exchange-correlation(xc) functional were employed for these optimizations. In contrast, vertical spin gaps are defined as the energy difference between the singlet and triplet states calculated at the same molecular geometry, typically

the singlet structure ($\Delta E_{\text{vert}}^{\text{t-s}} = E^{\text{t}}(\mathbf{R}^{\text{s}}) - E^{\text{s}}(\mathbf{R}^{\text{s}})$, Figure 6.1 a).

The molecular structure optimization is a key factor of the computational time in terms of the calculation of adiabatic spin gaps (see Section 6.3). Moreover, a linear relationship between vertical and adiabatic spin gaps calculated with three xc functionals was reported for a set of 14 halocarbenes.[142] The authors recognized an almost constant energy contribution from the geometrical relaxation ($E^{\text{t}}(\mathbf{R}^{\text{s}})$ vs. $E^{\text{t}}(\mathbf{R}^{\text{t}})$). This indicates that vertical spin gaps could serve as a computationally less demanding alternative to adiabatic gaps. One of the goals of this work is to challenge this with the QMspin dataset that provides a significantly greater coverage of chemical space compared to the halocarbenes.

Over the past three decades, Kohn–Sham DFT has become the most widely used electronic structure method in computational chemistry, due to its balance of reasonable accuracy and computational efficiency.[143, 144] Ongoing research continues to refine exchange-correlation functionals for diverse applications, including spin gap calculations (see also Chapter 5). Meanwhile, machine learning approaches have gained prominence for their potential to significantly reduce computational time while maintaining or improving accuracy. This has led to increased focus on machine learning (ML) in computational chemistry [38, 69, 145–152], particularly for predicting molecular properties such as spin state energies.[34–41] In a conceptually different approach, ML has been applied as tool to recommend the best xc functional for calculating vertical spin splitting energies of transition metal complexes.[153] Also combining the possibilities of ML and electronic structure methods is the Δ -ML approach that instead of predicting molecular properties directly adds machine learned corrections to the calculated value.[42]

While molecular representation and descriptor selection for ML have been extensively discussed in the literature [43–45], the impact of the origin of molecular structures on ML predictions remains underexplored. Hence, this work aims to investigate trade-offs between accuracy and computational cost based on three aspects, using the reference data from QMspin:

1. The number of optimized molecular structures needed: Molecular graph-based descriptors for ML do not require a 3D molecular structure. For vertical and adiabatic spin gaps, one and two optimized molecular structures are required, respectively.
2. The method used for molecular structure optimization: DFT (B3LYP) and the semiempirical PM6 method are used in this work.
3. The method used for the calculation/prediction of the spin gaps: DFT, ML, and Δ -ML as a hybrid approach are considered.

The combination of these three aspects leads to a variety of ways to obtain spin gaps, as illustrated in Figure 6.2.

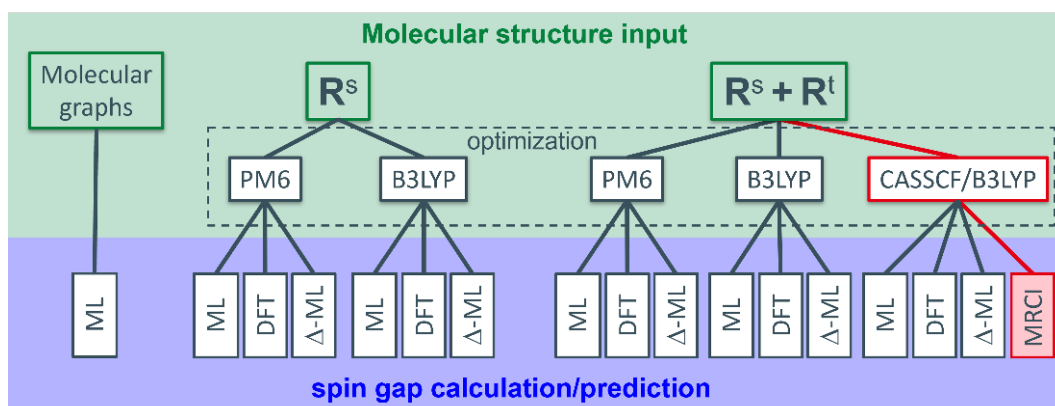


Figure 6.2: Overview of methods to obtain spin gaps. ML can be applied to molecular graph-based descriptors, and no optimized molecular structures are needed. For vertical spin gaps, only singlet molecular structures (R^s) are required, while for adiabatic spin gaps, both singlet and triplet molecular structures ($R^s + R^t$) are needed. These molecular structures can be obtained by optimization with different methods. Based on the different structures, different methods can be applied to obtain the spin gaps. The path used for the QMspin dataset is highlighted in red.

6.2 Computational Details

All KS-DFT calculations were performed using the GAUSSIAN 16 program package[99]. Molecular structure optimizations were carried out with the B3LYP[105] exchange-correlation functional combined with the def2-TZVP[103, 104] basis set or with the semi-empirical PM6[154] method.

Single-point calculations on optimized molecular structures were conducted with various exchange-correlation (xc) functionals, including B3LYP[105], M06[106], TPSSH[107, 108], ω B97XD[109], CAM-B3LYP[62], PBE0[110], BP86[100, 101], PBE[111, 112], TPSS[107], M06L[113], and LC- ω PBE[114, 115], all combined with Ahlrichs’ def2-TZVP[103, 104] basis set. SCF convergence criteria were set to “tight” (change in RMS density matrix $< 10^{-8}$, change in RMS MAX density matrix $< 10^{-6}$, and change in RMS energy $< 10^{-6}$ au). Atomic numbers of unpaired electrons ($n^{\alpha-\beta} = n^{\alpha} - n^{\beta}$) were extracted based on Mulliken population analysis. In GAUSSIAN output files, these are referred to as atomic “spin densities.” Similarities between two molecular structures were evaluated using the minimal root-mean-square deviation (RMSD) of the Cartesian coordinates with the RMSD package[155], which implements Kabsch[156] and Quaternion[157] algorithms for aligning the molecules.

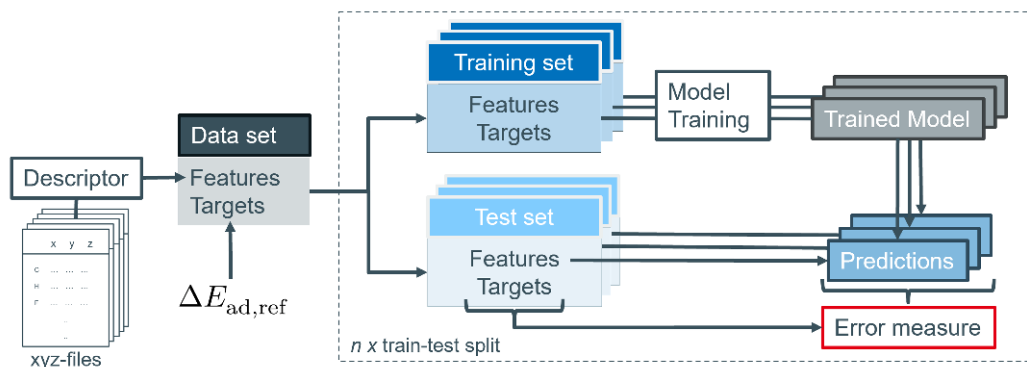


Figure 6.3: Workflow of the machine learning process and model validation used for the prediction of spin gaps in carbenes.

Machine learning models were implemented using SCIKIT-LEARN[64]. The model setup including hyperparameters are given sample code in Listing C.1. Molecular descriptors LMBTR, SOAP and ACSF were constructed using DSCRIBE[63]. Parameters can be taken from sample code in Listing C.2 An autocorrelation function descriptor (AC) with a maximum depth of 3 was constructed as implemented in MOLSIMPLIFY[158, 159]. The validation of ML models was achieved by repeated random train-test splits as depicted in Figure 6.3.

international chemical identifier (InChI)[160] strings were extracted using OPEN BABEL Version 3.0.0[161]. Unless otherwise noted, for scaling of the feature vectors the standard scaler as implemented in SCIKIT-LEARN has been used on all 2841 carbenes prior to the test-train split. In this standard scaler, each feature is reported relative the mean of the corresponding feature dimension for the full data set and divided by its variance.

6.3 Data Set Construction

The QMspin dataset[1] is constructed from the QM9 data set[141] and provides 3743 carbenes with MRCISD-F12+Q/cc-pVDZ-F12 calculated adiabatic spin gaps $\Delta E_{\text{ad,ref}}$ [1], which we take as reference. The molecular structures in the QMspin dataset have been optimized with CASSCF(2e,2o)/cc-pVDZ-F12 and B3LYP/def2-TZVP for singlet and triplet carbenes, respectively.[1] From this carbenes, those that exhibited rearrangements during molecular

structure optimization with B3LYP (singlet carbenes) or with PM6 (singlet and triplet carbenes) were excluded. This resulted in a final data set of 2841 carbenes.

Rearrangements result in more stable molecules compared to the artificially constructed carbenes. Thus, spin gaps calculated with these structures result in spin gaps that are outside the expected range of -70 kJ mol^{-1} to 300 kJ mol^{-1} . Rearrangements in the singlet molecular structure result in a lowering of ΔE^s and thus in spin gaps above that range ($\Delta E^{t-s} = \Delta E^t - \Delta E^s$). In contrast to that, rearranged triplet molecular structures cause spin gaps below that range. This can be visualized by a plot of kernel density estimation (KDE)¹ of the adiabatic spin gaps (Figure 6.4 a). An explanation of the used notation for the differently calculated spin gaps can be found in Figure 6.1 (b). In the top plot spin gaps of the uncut dataset (3743 carbenes) is shown. $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ (green) shows numerous spin gaps larger than 300 kJ mol^{-1} as highlighted by the green up pointing arrow. These spin gaps are calculated with different singlet molecular structures compared to $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ (orange curve). The latter are calculated with the molecular structures provided by QMspin. This is indicative for the above mentioned rearrangements during the B3LYP optimization of the singlet carbenes. On the other hand, $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ show fewer spin gaps around 0 kJ mol^{-1} compared to $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ (down pointing arrow). This small spin gaps are mostly observed for carbenes with aliphatic neighboring groups, which seem to be prone to undergo ring-opening rearrangements or α H-shifts (Figure 6.4b top). Also for the spin gaps calculated with PM6-optimized molecular structures $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$ (red curve) a few spin gaps outside -70 kJ mol^{-1} to 300 kJ mol^{-1} have been observed.

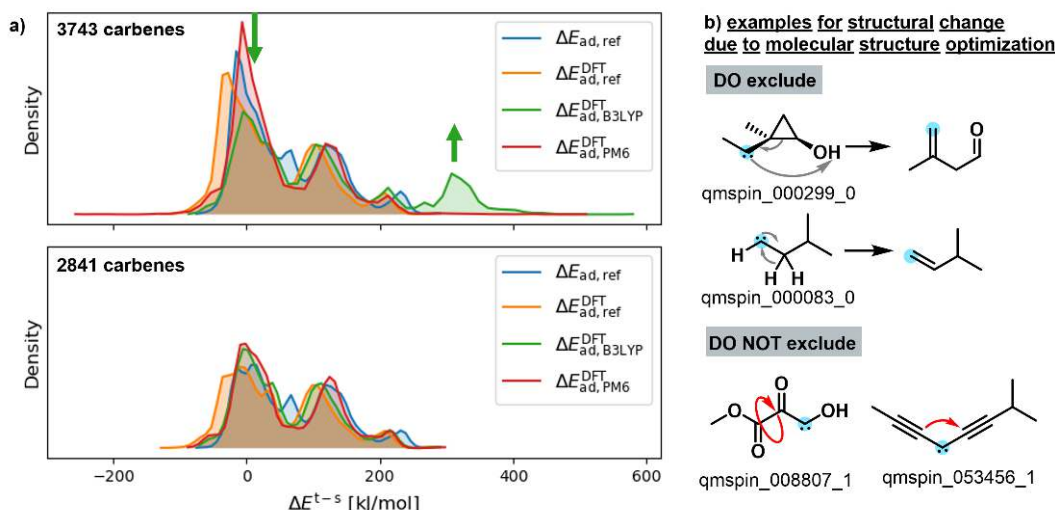


Figure 6.4: (a) Kernel density estimation (KDE) (gaussian kernel, width=5) of adiabatic spin gaps for the full data set (top) and the cut data set (bottom). (b) Examples of molecular structure changes during molecular structure optimization. Rearrangements that change the chemical identity of the molecules are excluded from the data set (top). Changes in bond angles and rotamers result in large RMSD but do not result in a different chemical identity and are kept in the data set (bottom).

Rearrangements change the chemical identity of the molecule (Figure 6.4b top). Thus, calculating spin gaps with these structures and comparing them to carbenes is not useful. Early attempts to exclude rearranged molecules were based on RMSD and evaluating the neighboring atoms of the carbene carbon. This method failed as on one hand the number of neighbors does not change for all rearranged carbenes (e.g. cumulenes). Therefore, some rearranged carbenes remained undetected. On the other hand large RMSD can also

¹The KDE can be considered as a smoothed histogram.

be caused by rotations around σ -bonds or changes in bond angles (Figure 6.4 b). These changes do not influence the chemical identity and should therefore not be excluded. The comparison of InChI[160] strings generated from the xyz-files before and after molecular structure optimization proved to be a simple and straightforward way to exclude molecules with a changed chemical identity. This can be seen from the distribution of spin gaps after the exclusion of rearrangements by this method (Figure 6.4 a, bottom). For $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ and $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$ no spin gaps outside the expected range are observed. Additionally, the distribution of $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ within the expected range does not change much compared to uncut dataset. Hence, no carbenes have been falsely removed.

6.4 Computing Times

For all DFT single-point calculations and the molecular structure optimization of the singlet carbenes with B3LYP ($\mathbf{R}_{\text{B3LYP}}^{\text{s}}$), the molecular structure optimization with PM6 of singlet ($\mathbf{R}_{\text{PM6}}^{\text{s}}$) and triplet ($\mathbf{R}_{\text{PM6}}^{\text{t}}$) carbenes, the CPU time has been analyzed (Figures C.1 and C.2). On average, DFT single-point calculations took 13.8 min, the optimization with B3LYP took 119 min, and the optimization with PM6 took 2.9 min. For the single-point MRCI calculation of QMspin, a mean CPU time of 7.5 h was reported by Heinen *et al.*[162] The CPU time for the CASSCF optimizations of QMspin was reported to be 4.1 h.[162] As different machines have been used for the calculations in this work and for the generation of the QMspin dataset, these are not directly comparable. Nonetheless, we want to compare the computational cost of an ML or Δ -ML approach to pure DFT calculations. Therefore, we assume the CPU timings in this work and those from the QMspin dataset are comparable. Additionally, we assume that the train-test split does not influence the mean computational time of the train and test sets. Lastly, it is assumed that the time for model fitting and prediction of the test set is negligible for all ML processes compared to the electronic structure calculations. Except for some combinations of Gaussian process regression (GPR) with very large descriptors (LMBTR or SOAP, see Figure C.9), the last statement is valid, and including the model fitting, the prediction per molecule costs less than 1 s. Unlike DFT calculations, which are independent of the reference calculations, with all these assumptions, ML processes depend linearly on the training set size, as shown in Figure 6.5. ML processes are only more efficient compared to DFT when the molecular structure quality used for ML is lower than the one for DFT and training sets are small. With a training set size larger than 800 carbenes, DFT calculations are more time-efficient. Due to the very small CPU times of the PM6 optimization, it is almost negligible compared to the reference calculation. Hence, ML with $\mathbf{R}_{\text{PM6}}^{\text{s+t}}$, $\mathbf{R}_{\text{ad,PM6}}^{\text{s}}$, as well as molecular graph-based descriptors, do not differ significantly. The last two cases are therefore not shown in Figure 6.5.

While B3LYP optimization timings for QMspin are not provided by Heinen *et al.*, a mean CPU time for B3LYP optimizations of the QM9 dataset is presented (4.4 min).[162] Although the QM9 dataset is different from QMspin, the size of the molecules is similar. Therefore, the CPU times for the B3LYP optimization of QM9 might be a realistic approximation. Based on that, CPU times for B3LYP optimizations are 56 times faster compared to CASSCF optimizations. With this updated approximation, relative computing times almost exclusively dependent on the amount of CASSCF optimizations and MRCI single-point calculations (Figure C.3). Under these conditions, any ML that needs training sets larger than 600 carbenes has longer CPU times compared to DFT calculations based on the reference molecular structures. Whether ML or Δ -ML are still useful compared to DFT calculations depends on their performance in predicting the spin gaps. To evaluate this, in Section 6.5, the performance of DFT based on different molecular structures will be analyzed first. Afterwards, the performance of ML and Δ -ML will be analyzed in Sections 6.6 and 6.7, respectively.

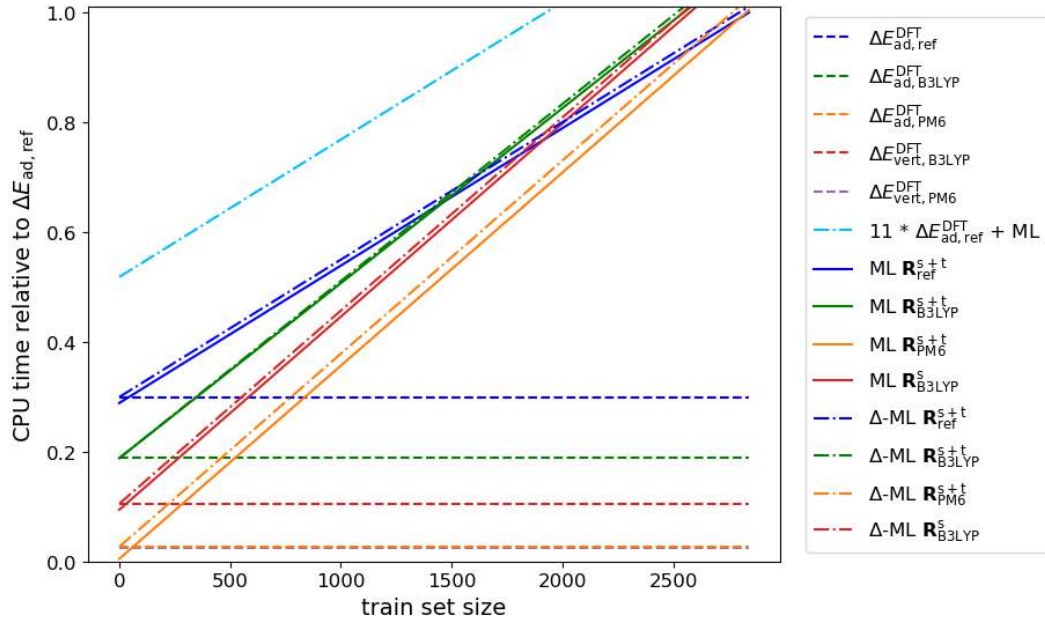


Figure 6.5: Approximation of the CPU timing for the DFT calculations, ML, and Δ -ML of spin gaps in carbenes relative to the CPU times of the reference spin gaps $\Delta E_{\text{ad,ref}}$ (1 CASSCF optimization, 1 B3LYP optimization, and 2 single-point MRCI calculations per carbene). It is assumed that the time needed for the ML process is negligible in comparison to the electronic structure calculations. Further, it is assumed that $t(\text{optimization CASSCF})$ is approximately 2 $t(\text{optimization B3LYP})$.

6.5 Influence of Molecular Structures on DFT Calculated Adiabatic and Vertical Spin Gaps

Before looking into the ML and Δ -ML approach for predicting spin gaps in carbenes, in this section, the performance of DFT for the calculation of spin gaps in organic carbenes is examined based on different molecular structures. This includes various exchange–correlation (xc) functionals such as pure, hybrid, and range-separated ones. While the reference molecular structures employ the accurate yet computationally costly CASSCF method for the optimization of the singlet carbenes ($\mathbf{R}_{\text{ref}}^s$), purely DFT-optimized molecular structures (B3LYP xc functional) as well as the computationally very efficient semi-empirical PM6 method are also considered here. An explanation of the notation used to differentiate between these differently obtained spin gaps is shown in Figure 6.1 b.

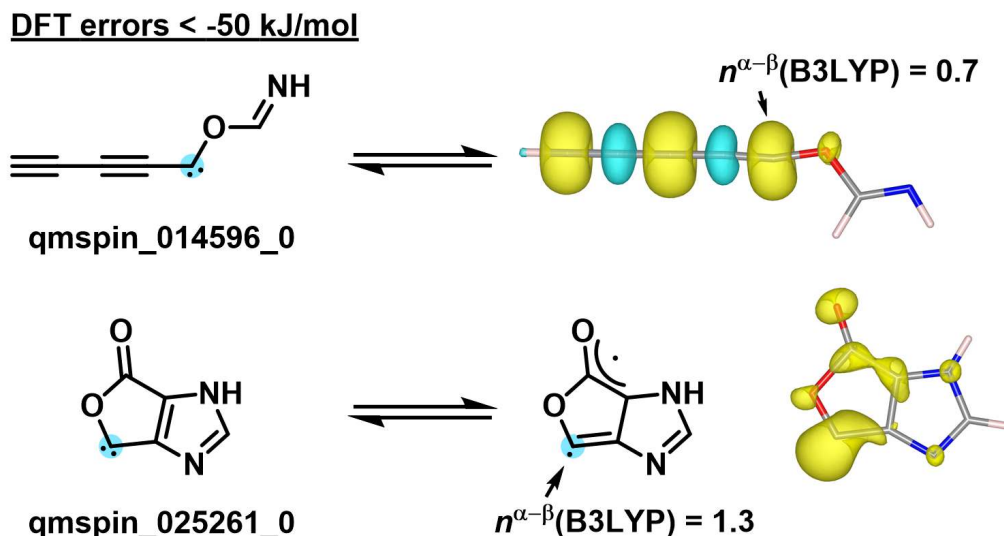


Figure 6.6: Examples of carbenes where all 11 DFT xc functionals give errors < -50 kJ mol^{-1} . The number of unpaired electrons on the carbene carbon has been evaluated using Mulliken population analysis. Spin densities are given with an isosurface value of $0.01 e a_0^{-3}$.

Figure 6.8 provides an overview of the error distributions for spin gaps calculated with each xc functional on the different molecular structures. Errors are given relative to the reference spin gaps ($\text{Error} = \Delta E^{\text{DFT}} - \Delta E_{\text{ad,ref}}$). The rows in the Figure use the same structures, while the columns group pure, hybrid, and range-separated xc functionals.

First, the adiabatic DFT spin gaps calculated with the same structures as the reference calculation ($\Delta E_{\text{ad,ref}}^{\text{DFT}}$) will be analyzed. Their error distributions are depicted in the first row of Figure 6.8. Additionally, each of these error distributions for individual xc functionals is described by MSE, MAE, and the Q_{95} error measure. Those are depicted in the top, middle, and bottom of Figure 6.7, respectively. For the MSE error bars show the standard deviation of the errors. While the MSE reflects the accuracy of a prediction the standard deviation is a measure of its precision. For the MAE the standard deviation of absolute errors is shown by the error bar. The combination of low accuracy and good precision are generally indicative of systematical errors. There is a high chance that these can be corrected by additionally applying an ML approach. High precision at low accuracy on the other hand might not easily be corrected.

$\Delta E_{\text{ad,ref}}^{\text{DFT}}$ show the highest precision compared to spin gaps calculated with different structures. This is evident from the small standard deviation of the MSE in Figure 6.7 (a more

detailed view can be found in the supporting information in Figure C.4). However, the error distribution for $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ is shifted towards negative errors most strongly (most negative MSE), and the differences among the xc functionals are the largest. This can be clearly seen when comparing the MAE or MSE with the other methods. Negative errors result from triplet spin states being more favored.

Interestingly, spin gaps of the pure xc functionals BP86, TPSS, and PBE are most negative among the functionals. That is, pure functionals favor the triplet (high-spin) state more than hybrid functionals. That is the opposite behavior observed when calculating spin gaps in transition metals, as discussed in more detail in Chapter 5. Although the spin gaps calculated with M06, which has the largest amount of exact exchange ($a_{\text{HF}} = 27\%$) among the xc functionals tested in this study are the least negative, the dependence on the exact exchange is not simply inverted if compared to the TM complexes. For carbenes, there is no clear linear dependence on a_{HF} , as evidenced by the strong negative shift of the spin gaps of PBE0 ($a_{\text{HF}} = 25\%$). The functionals working well for carbenes are different from those working well for TM complexes discussed in Chapter 5. The best-performing xc functional is M06 (27% exact exchange) with a MAE of 7.6 kJ mol^{-1} and Q_{95} of 25.5 kJ mol^{-1} , making it the best functional choice for 89% of the carbenes when using the reference molecular structures.

All distributions of the adiabatic spin gap errors of $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ evaluated for the reference structures (first row Figure 6.8) are unsymmetrical and show a tail at the negative side of the distribution. These very large negative errors have been observed for carbenes conjugated to unsaturated or aromatic systems (Figure 6.6). For these systems, DFT calculations quite often converge to a solution where the spin density is delocalized onto the conjugated system, instead of it being localized on the carbene carbon. In accordance with this, an analysis of the number of unpaired electrons on the carbene carbon ($n_{\text{C}}^{\alpha-\beta}$) shows that there is a correlation with the absolute error of $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ for most xc functionals (see Figure C.6). The largest correlation was observed for B3LYP (linear fit: $R^2 = 0.77$, Figure C.7). However, the linear correlation for LC- ω PBE is very weak ($R^2 = 0.17$, Figure C.8).

Interestingly, all DFT xc functionals discussed here outperform $\Delta E_{\text{ad,ref}}^{\text{CASSCF}}$ precisionwise (see error bar on yellow bar in Figure 6.7). Due to additionally better accuracy of the four best xc functionals, those perform better than CASSCF by any error measure tested here.

For $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, where both triplet and singlet molecular structures were optimized with B3LYP, errors are more symmetrically distributed compared to $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ (Figure 6.8, second row). The shift towards negative errors is much less pronounced, resulting in better overall performance for each xc functional. Due to using molecular structures of the same quality for singlet and triplet carbenes error compensation might occur. ω B97XD, CAM-B3LYP, M06, and B3LYP perform best with Q_{95} values of 23.2 kJ mol^{-1} , 24.0 kJ mol^{-1} , 24.5 kJ mol^{-1} , and 24.7 kJ mol^{-1} , respectively (Figure 6.7, bottom row). While Q_{95} values are improved compared to $\Delta E_{\text{ad,ref}}^{\text{M06}}$, the MAEs (8.0 kJ mol^{-1} , 8.1 kJ mol^{-1} , 8.7 kJ mol^{-1} , and 9.1 kJ mol^{-1} , respectively) are not.

Although the quantitative differences for the MAEs and Q_{95} values compared to $\Delta E_{\text{ad,ref}}^{\text{M06}}$ are small, this demonstrates the importance of the choice of error measure for determining the quantitatively best choice.

The difference in the performance between $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ and $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ stems only from the differently optimized molecular singlet structures. Although the computationally more expensive CASSCF-optimized singlet molecular structures are considered more accurate compared to the B3LYP-optimized ones, the latter are energetically preferred by not only the B3LYP functional but all xc functionals tested. This explains the systematic favoring of the B3LYP-optimized triplet carbenes observed for $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ and shows that more accurate molecular structures do not necessarily result in more accurate predictions.

For $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$, which is determined based on PM6-optimized singlet and triplet structures, the accuracy (represented by the MSE) is very similar to $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ (Figure 6.7, top).

However, the precision (standard deviation) of $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$ is much worse, resulting in poorer performance compared to $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ for all xc functionals.

Now, the focus will be on to the vertical spin gaps, which are calculated using the molecular structure of the singlet carbene only. These errors are generally expected to shift towards more positive values due to a systematic favoring of the singlet carbene, for which the molecular structure has been optimized.

This shift is evident in the case of $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$ ², which exhibits very large errors and a broad error distribution across all xc functionals. The performance of these vertical spin gap predictions is even worse than simply predicting the spin gap of each carbene as 69 kJ mol^{-1} , which is the mean of all $\Delta E_{\text{ad,ref}}$. This comparison is visualized in Figure 6.7, where the standard deviation (MSE) and error measures calculated from this prediction are indicated by black horizontal lines.

In contrast, $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$ performs surprisingly well, with errors more evenly distributed around 0 kJ mol^{-1} . Some xc functionals even show better performance with these vertical gaps compared to the adiabatic spin gaps. Nevertheless, the distribution of errors remains broad. When compared to spin gaps calculated based on other molecular structures, the differences in performance among the xc functionals are relatively small. This indicates that the choice of molecular structure optimization method plays a significant role in the accuracy of the predicted spin gaps.

Vertical spin gaps calculated on the CASSCF-optimized singlet structures have been provided for CASSCF and the reference MRCI method.[1] For both methods, the expected preference of singlet carbenes has been observed. Interestingly, this preference is larger for MRCI than for CASSCF (see Figure C.4). This results for CASSCF in a comparable performance to $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$ and a worse performance for $\Delta E_{\text{vert,CASSCF}}^{\text{MRCI}}$.

²All single-point calculations terminated normally, and no convergence issues were observed to cause the large errors.

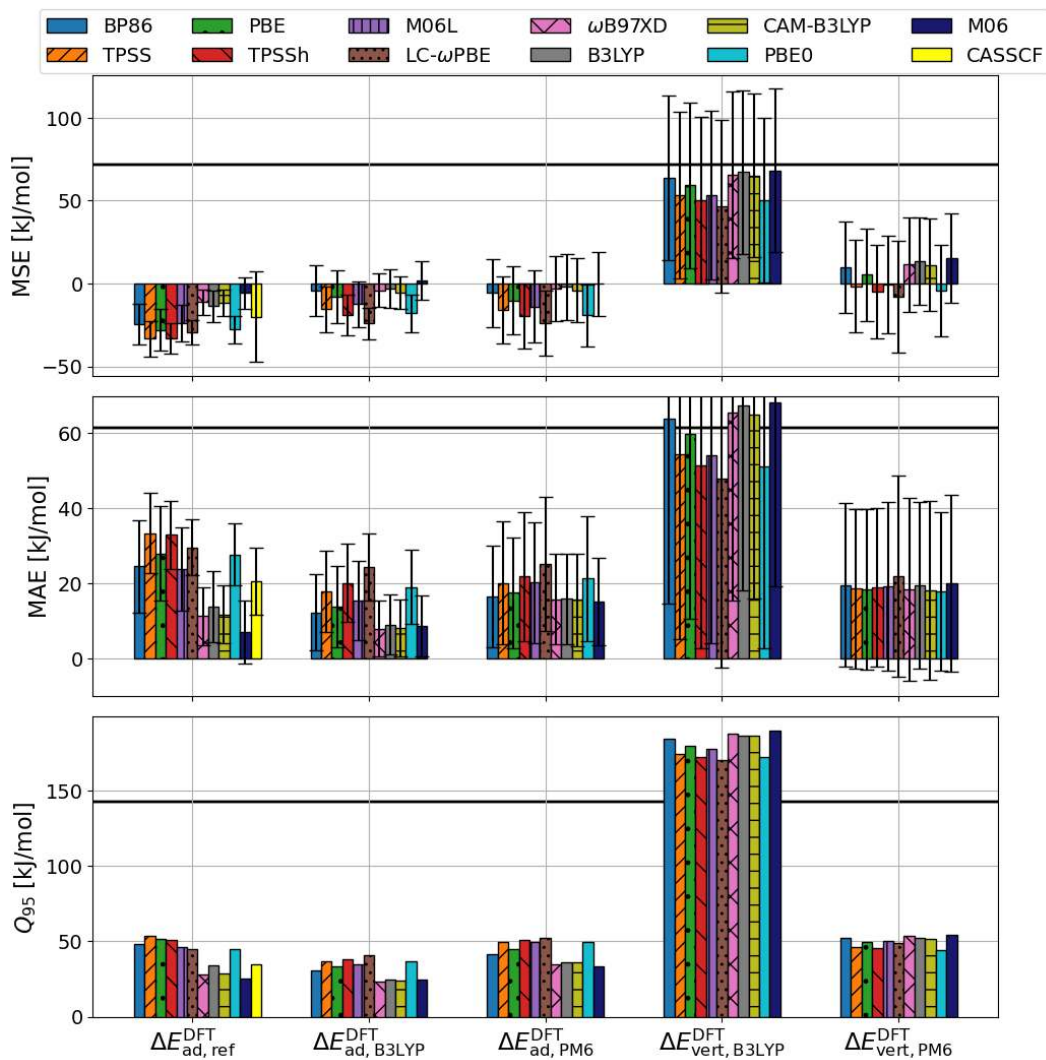


Figure 6.7: Errors of DFT predictions for spin gaps relative to MRCI calculated adiabatic spin gaps ($\Delta E^{DFT} - \Delta E_{ad,ref}$) evaluated by MSE, MAS and Q_{95} . Black error bars indicate standard deviation of signed errors (top) and absolute errors (middle). The black horizontal line is the error measure resulting from predicting the constant 69.1 kJ mol^{-1} , the mean of $\Delta E_{ad,ref}$, for each carbene.

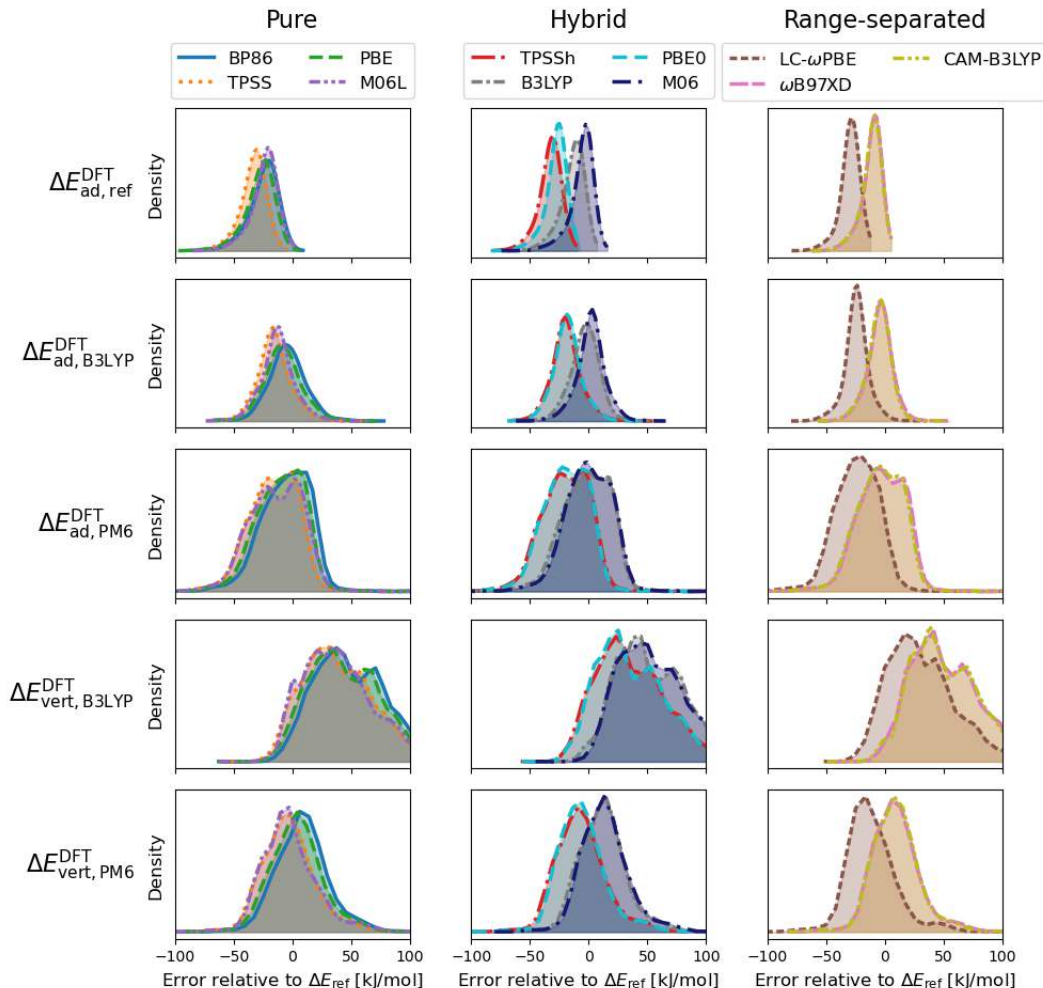


Figure 6.8: Kernel density estimate of DFT spin gap errors (kernel='gaussian', width=0.5) between -100 kJ mol^{-1} and 100 kJ mol^{-1} . Errors are given relative to the MRCI spin gaps ($\Delta E^{DFT} - \Delta E_{ad,ref}$). For each row the same molecular structures have been used. For the first three rows, adiabatic spin gaps with optimized singlet and triplet molecular structures have been calculated. The bottom two rows show the errors of vertical spin gaps that only employ optimizes singlet molecular structures. Columns contain the error distributions of the spin gaps calculated with sets of the same xc functionals. In the columns, pure, hybrid and range separated xc functionals are grouped together from left to right.

6.5.1 Differences in Molecular Structures

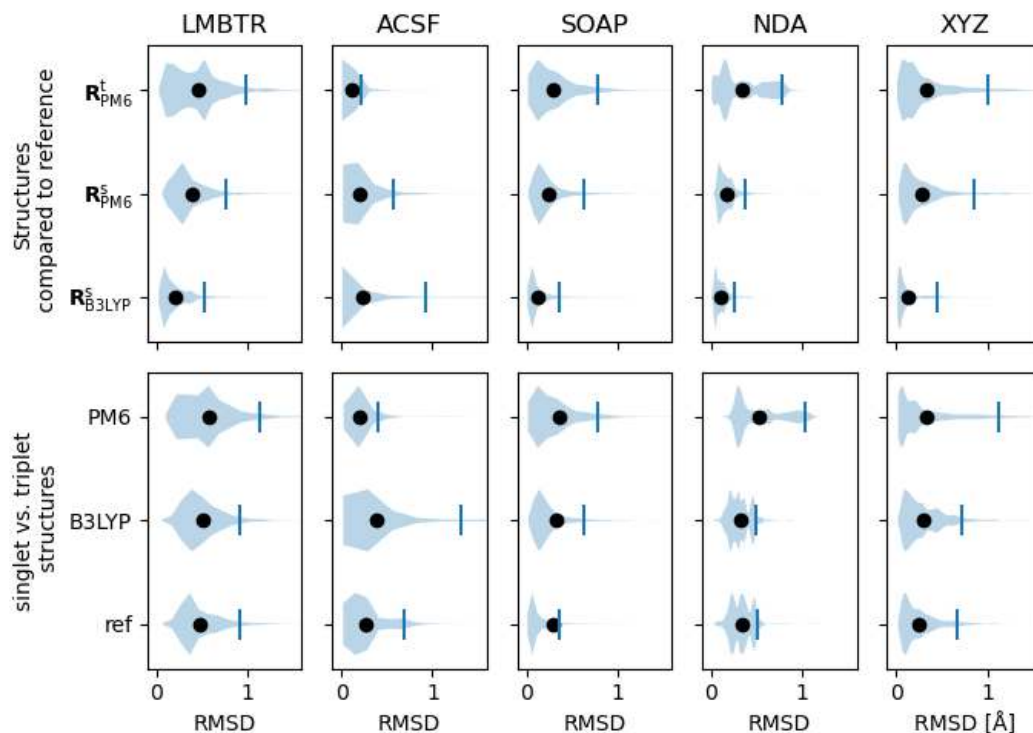


Figure 6.9: Violin plots of RMSD between structures evaluated for different descriptors and xyz coordinates. Blue shaded areas show the distribution of RMSDs for individual carbenes (KDE with gaussian kernel). The Mean of each distribution is represented by black dots and Q_{95} values are given by the blue vertical lines. Features have been scaled before evaluating the RMSD. comparison of B3LYP- and PM6-optimized structures with the reference structures. Singlet structures are compared to singlet CASSCF-optimized molecular structures and triplet PM6 structures are compared to triplet B3LYP structures. Bottom: Structural comparison between singlet and triplet carbenes.

Differences in performance discussed in this Section are primarily based on variations in molecular structures. To investigate these differences, structural changes resulting from the three types of molecular structure optimizations were analyzed. Structural differences between two molecules can be reduced to a single value by calculating the root mean squared deviation (RMSD) of the atomic coordinates. However, this method does not distinguish where the structural changes occur, and, for example, a rotation of an alkyl group far away from the carbene carbon could result in a larger RMSD than a structural change near the carbene carbon. But, when looking at spin gaps, especially the latter case should have a greater impact.

For this reason, we also use the descriptors that are employed for the machine learning in Section 6.6 to compare structural details. As local descriptors are used, they mostly reflect the surroundings of the carbene. More details on the descriptors can be found in Sections 4.2 and 6.6. Molecular changes are visualized in Figure 6.9 using violin plots, which depict the root mean squared deviation (RMSD) between reference and newly optimized structures (top row). In the columns, the RMSD, for the different descriptors as well as the atomic coordinates (XYZ) are displayed. The RMSD of the descriptors is dimensionless due to feature scaling, while for coordinates, no scaling was applied as all dimensions are atomic distances. In the violin plots, the blue-shaded areas represent the RMSD distribution for each carbene, the black dot indicates the mean, and the vertical blue line shows the Q_{95} .

Optimization of singlet carbenes using B3LYP resulted in relatively small changes for the majority of carbenes, as seen in the changes in Cartesian coordinates as well as LMBTR, SOAP, and NDA descriptors. This is in good agreement with the good performance of DFT for the adiabatic spin gaps calculated with the B3LYP-optimized molecular structures. However, the ACSF descriptor showed some more outliers. The spin gaps calculated with $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ were also most similar to the reference values. In contrast, optimization with PM6 resulted in more significant changes, with the triplet structures showing even more changes than the singlet structures. Interestingly, for the ACSF descriptor, these trends were inverted, raising the question of whether this will also reflect in the machine learning performance for this descriptor.

To compare vertical and adiabatic gaps, structural differences between singlet and triplet structures were analyzed similarly (Figure 6.9 bottom). The structural differences between singlet and triplet structures are generally larger than the changes made by B3LYP optimization. These differences increase progressively from reference structures to purely B3LYP-optimized to purely PM6-optimized structures. This leads to the conclusion that the significant differences between singlet and triplet molecular structures are essential for precise calculations.

6.6 Influence of Molecular Structures on the Supervised Learning of Adiabatic Spin Gaps

In Chapter 6.5, we observed how sensitive electronic structure properties are on how the underlying molecular structures were generated, a phenomenon that is recognized widely. Now, we investigate the influence of these structures on the performance of machine learning approaches for predicting the MRCI-calculated adiabatic spin gaps ($\Delta E_{\text{ad,ref}}$). For carbenes, where spins are mainly localized on the carbene carbon, the local environment is expected to have the most significant impact on the spin gap. Therefore, we have used *local* variants of the descriptors in this work. One example is the ACSF descriptor, where symmetry functions have been exclusively generated for the carbene carbon. In preliminary tests, this has lead to a better performance.

Additionally, a very simple descriptor, Neighbors + Distances + Angles (NDA), has been developed. A similar approach of constructing such use specific descriptors for a compound class has been reported for the prediction of spin exchange couplings in dicopper complexes.[37] The carbene specific NDA descriptor includes the atomic numbers of both neighboring atoms of the carbene carbon, and the corresponding distances and angle as illustrated in Figure 6.10. The “adiabatic” version of the NDA descriptor considers distances and angles from both singlet and triplet carbene structures. Therefor, the feature dimensions increase from 5 to 8 compared to a variant based solely on one molecular structure. A consistent notation, in line with the spin gaps, has been applied to these descriptors. Subscripts “s”, “t”, or “ad” indicate their origin from singlet, triplet, or both molecular structures, respectively. The method used for molecular structure optimization is also denoted with subscripts.

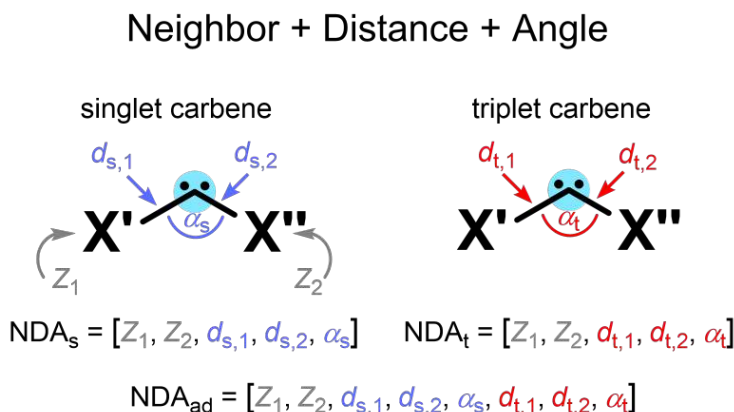


Figure 6.10: Construction of the Neighbors + Distances + Angle (NDA) descriptor. The distance + angle (DA) descriptor does not include the atomic number of the neighboring atoms.

Figure 6.11 provides an overview of the performance for predicting $\Delta E_{\text{ad,ref}}$ measured by Q_{95} for each combination of ML approach and descriptor. Both scaled and unscaled descriptors have been tested. For better comparability, descriptors in Figure 6.11 have been grouped as follows: all descriptors solely based on molecular structures are found in the **R** groups. In each **R** group, the same molecular structures are used. For example the group $\mathbf{R}_{\text{B3LYP}}^s$ employs B3LYP-optimized molecular structures of the singlet carbenes and $\mathbf{R}_{\text{ref}}^{s+t}$ combines CASSCF optimized structures of the singlet carbenes with B3LYP-optimized triplet molecular structures. with the descriptors NDA, DA (reduced version of NDA that only includes Distances + Angle, not used for adiabatic versions of the descriptors), ACSF, LMBTR, and SOAP.

As DFT-calculated spin gaps ΔE^{DFT} have shown to systematically favor triplet errors (Section 6.5), these have also been used as features (11 feature dimensions, one for each xc functional) to check the extent to which machine learning can compensate these errors. Feature vectors of this type have been created for each set of spin gaps that are calculated with different molecular structures and are grouped under ΔE^{DFT} in Figure 6.11 in the order $E_{\text{ad,ref}}^{\text{DFT}}$, $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$, $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$ and $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$. This is an approach that combines ML and DFT calculations like Δ -ML. The difference lies with the used targets. While for Δ -ML the errors of DFT calculated spin gaps ($\Delta E^{\text{DFT}} - \Delta E_{\text{ad,ref}}$) are the targets in this Section ΔE^{DFT} containing feature vectors are processed the same as the structure based descriptors and the adiabatic spin gaps $\Delta E_{\text{ad,ref}}$ are the target directly. Descriptors By computational time using the spin gaps calculated with all 11 xc time is the most costly approach discussed here (see also Figure 6.5 discussed in Section 6.4). Using only properties calculated by single point calculations with one xc functional is timewise identical to the Δ -ML approach.

In a similar fashion to the ΔE^{DFT} based descriptors, properties calculated with DFT have been used as features. A HOMO–LUMO gaps descriptor (HL^{DFT}) has been created with the HOMO–LUMO gaps calculated with the 11 xc functionals for the singlet carbenes with the singlet carbenes molecular structure $E_{\text{s,CASSCF}}^{\text{s,DFT}}$, $E_{\text{s,B3LYP}}^{\text{s,DFT}}$ and $E_{\text{s,PM6}}^{\text{s,DFT}}$. In contrast to the descriptor ΔE^{DFT} only one single point calculation is needed for each feature instead of the two single point calculation needed to calculate the spin gaps. Additionally, the Charges+Spins (CS) descriptor was created. It involves Mulliken charges on the carbene carbon and both neighboring atoms for the singlet as well as triplet state. And for the triplet state also the Mulliken number of unpaired electrons on the carbene carbon as well as both neighboring atoms is added. Here, only the M06 calculated properties were considered (CS^{M06}) which results in 9 dimensional feature vectors. Again those have been constructed based on the different molecular structures discussed in this work ($\text{CS}_{\text{ad,ref}}^{\text{M06}}$, $\text{CS}_{\text{ad,B3LYP}}^{\text{M06}}$, $\text{CS}_{\text{ad,PM6}}^{\text{M06}}$, $\text{CS}_{\text{vert,B3LYP}}^{\text{M06}}$ and $\text{CS}_{\text{vert,PM6}}^{\text{M06}}$).

From the overview Figure 6.11 can be seen that feature scaling generally improves the performance. Because of this, we will focus on scaled features in the following. The performances of the ML models and the different descriptors that are given are depicted combined in Figure 6.11 will be discussed separately in the following. In Section 6.6.1 the performance will be analyzed based on the ML models. Afterwards the influence of the descriptor will be analyzed by the different types of descriptors in Section 6.6.2 and by the different underlying molecular structures in Section 6.6.3. Lastly, the performance of descriptors involving electronic structure properties will be discussed in Section 6.6.4. When analyzing the performances it is important to keep in mind that with a training set size of 1500 carbenes the computational cost for any ML process validated in Figure 6.11 are at least 60 % larger than any DFT calculated spin gap discussed in this work. Hence, in order consider ML to be useful in practical application a Q_{95} considerably smaller 23.2 kJ mol^{-1} which was achieved by DFT calculations is required.

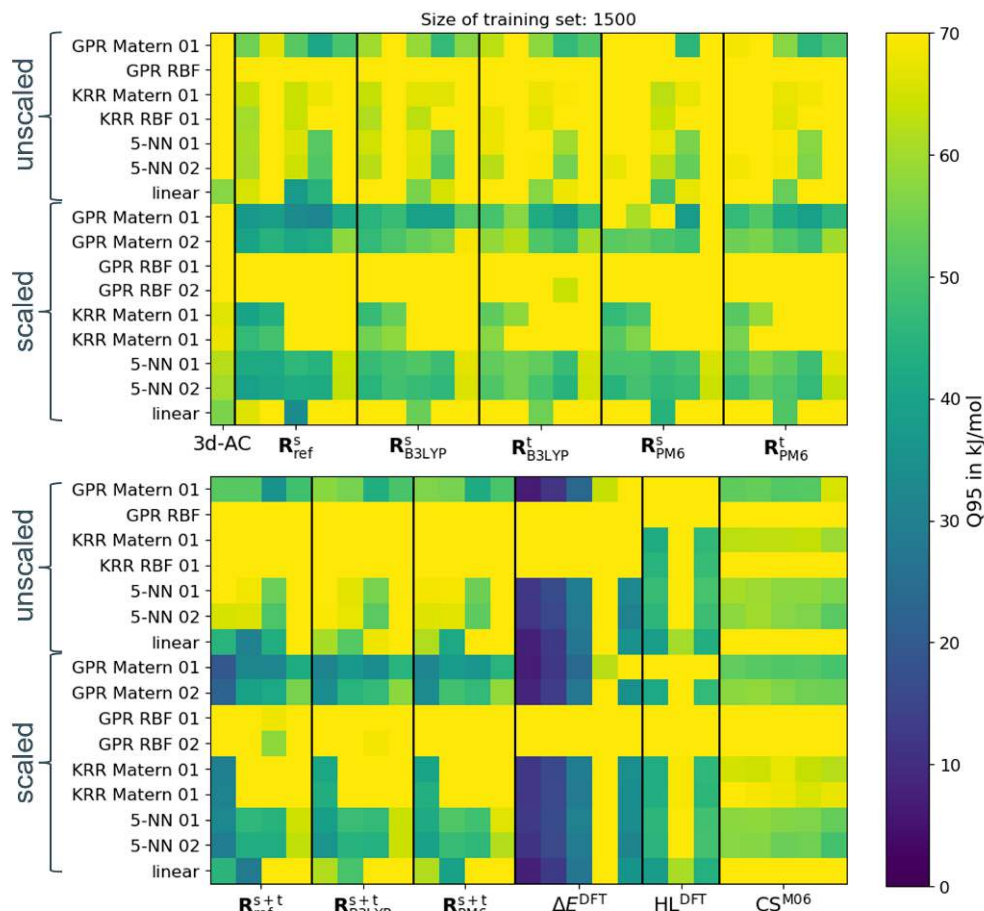


Figure 6.11: Overview of Q_{95} error measures for supervised machine learning with different models (y-axis) and descriptors (x-axis). The train-test split has been performed randomly 10 times with training set size of 1500. The squares are colored by the mean Q_{95} of the performance of each trained model with the corresponding test set. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Descriptors grouped under the **R** label employ only singlet or singlet and triplet molecular structures indicated by superscribed “s” or “s+t”, respectively. The method used for molecular structure optimization is denoted with subscripts. Within the groups, descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the “s+t” groups. For the group ΔE^{DFT} , DFT spin gaps calculated with all 11 xc functionals are used as features. The group contains $\Delta E^{\text{DFT}}_{\text{ad,ref}}$, $\Delta E^{\text{DFT}}_{\text{ad,B3LYP}}$, $\Delta E^{\text{DFT}}_{\text{ad,PM6}}$, $\Delta E^{\text{DFT}}_{\text{vert,B3LYP}}$, $\Delta E^{\text{DFT}}_{\text{vert,PM6}}$. For the HL^{DFT} group HOMO–LUMO, gaps calculated for the singlet carbenes with the 11 xc functionals are used as features. CASSCF, B3LYP- and PM6-optimized singlet structures are used to calculate the HOMO–LUMO gaps. The CS^{M06} descriptor contains Mulliken charges on the carbene carbon and the two neighboring atoms for the singlet and triplet carbene as well as the Mulliken number of unpaired electrons on those three atoms for the triplet carbene (9 feature dimensions). For these calculations the M06 functionals has been used. Different molecular structures have been used and the CS group contains $\text{CS}^{\text{M06}}_{\text{ad,ref}}$, $\text{CS}^{\text{M06}}_{\text{ad,B3LYP}}$, $\text{CS}^{\text{M06}}_{\text{ad,PM6}}$, $\text{CS}^{\text{M06}}_{\text{vert,B3LYP}}$ and $\text{CS}^{\text{M06}}_{\text{vert,PM6}}$.

6.6.1 Performance Analysis of Models

The models tested for the ML are Gaussian process regression (GPR) and kernel ridge regression (KRR), each with different kernels. The k -nearest neighbors (k -NN) method has been used with $k = 5$ and no weights (5-NN 01) and with weights based on the distance (5NN 02) (see also Section 4.2.2). Additionally, a simple linear regression has been performed for each descriptor (see Appendix C.1 for technical details).

When focusing on the ML models (y-axis in Figure 6.11) at first glance it is noticeable that GPR in combination with the RBF kernel performs very bad with Q_{95} almost always larger than 70 kJ mol^{-1} . Therefore, it is not considered further. Similarly, KRR with any structure based descriptor except NDA or DA performs bad. The GPR with Matern kernel is the only model considered here, that results in Q_{95} lower than 40 kJ mol^{-1} for almost any structure based descriptor.

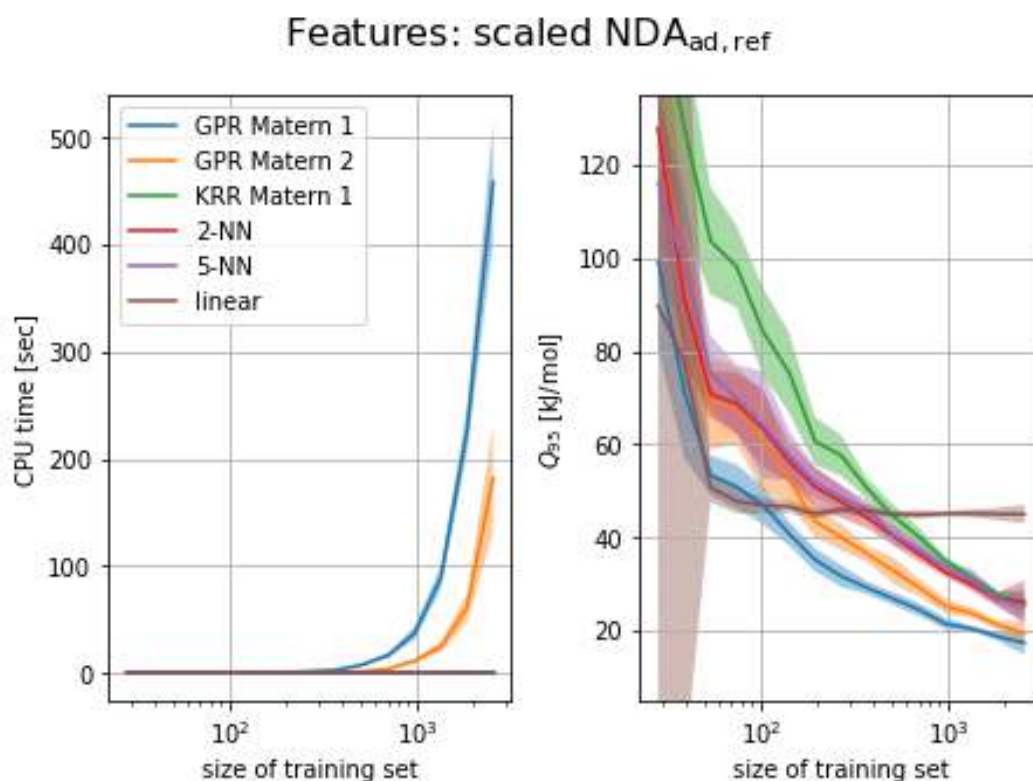


Figure 6.12: CPU times (left) and learning curves (right) for machine learning of adiabatic reference spin gaps. Various models have been tested on the scaled $\text{NDA}_{\text{ad,ref}}$ descriptor. The learning has been repeated 10 times with random train-test splits. Solid lines display the mean and shaded areas display the standard deviation of the Q_{95} error measure (right) and training times (left) for the 10 repetitions.

To get a more detailed comparison of the different ML models learning curves with the scaled version of the $\text{NDA}_{\text{ad,ref}}$ descriptor have been generated. This is the structure based descriptor that performs best for all of models. And the only one for which Q_{95} values similar to those achieved by DFT are obtained. With learning curves the dependence of performance on the size of the training set is analyzed. The learning curve in Figure 6.12 (right) shows that the performance (Q_{95}) of linear regression is mostly independent of the size of the training set as soon as the size exceeds 50 carbenes. The performance of all other methods continuously improves with the training set size. The KRR with the Matern 01

kernel performed worst and only beats the linear regression at training set sizes larger than 50 carbenes.

For the k -NN model, whether the number is 2 or 5 neighbors has no significant influence on the performance. Only at large training set sizes of 2500 carbenes is a performance comparable to calculations with DFT achieved ($Q_{95} = 23 \text{ kJ mol}^{-1}$).

GPR with a Matern kernels performs best. The Matern 01 kernel, which has an additional parameter to be fitted compared to Matern 02, is superior to the Matern 02 kernel. For training set sizes larger than 35 % (1000 carbenes), the Q_{95} are smaller than 23 kJ mol^{-1} and all DFT calculations are outperformed. The increased performance of the GPR with the Matern kernel goes hand in hand with increased CPU time needed for the fit of the data (Figure 6.12 left). However, even for training set sizes of 2500 within 500 s the model is trained and 341 spin gaps are predicted. Per molecule this results in a CPU time of less than 2 s, which is negligible compared to all electronic structure calculations as discussed in Section 6.4. However, the computational cost for optimizing the molecular structures as well as for calculating the targets for the training set result in an increased CPU time by a factor of at least 2.5.

6.6.2 Performance Analysis of Descriptors

Analyzing the ML models revealed that GPR and k -NN perform best for the structure based descriptors. As the relative performance of the descriptors might depend on the ML model, both will be to analyze the different types of structure based descriptors.

Figure 6.13 shows learning curves for all structure-based descriptors for the GPR Matern 01 and 5-NN 02 models. The adiabatic reference structures, considered to be the best description for the singlet and triplet carbenes, are used for this purpose. While GPR generally performs better compared to k -NN, as has been found in Section 6.6.1, both models show similar trends. The SOAP descriptor, with 8200 feature dimensions (adiabatic version), the largest descriptor tested in this work, performs the worst. The simple NDA descriptor performs best among the structure-based descriptors, regardless of the size of the training set in the case of GPR. This highlights the advantage of problem-driven thinking and simplicity over much more complex general-purpose descriptors. ACSF and LMBTR perform similarly for training sets larger than 300. For smaller set sizes, ACSF performs more similarly to NDA.

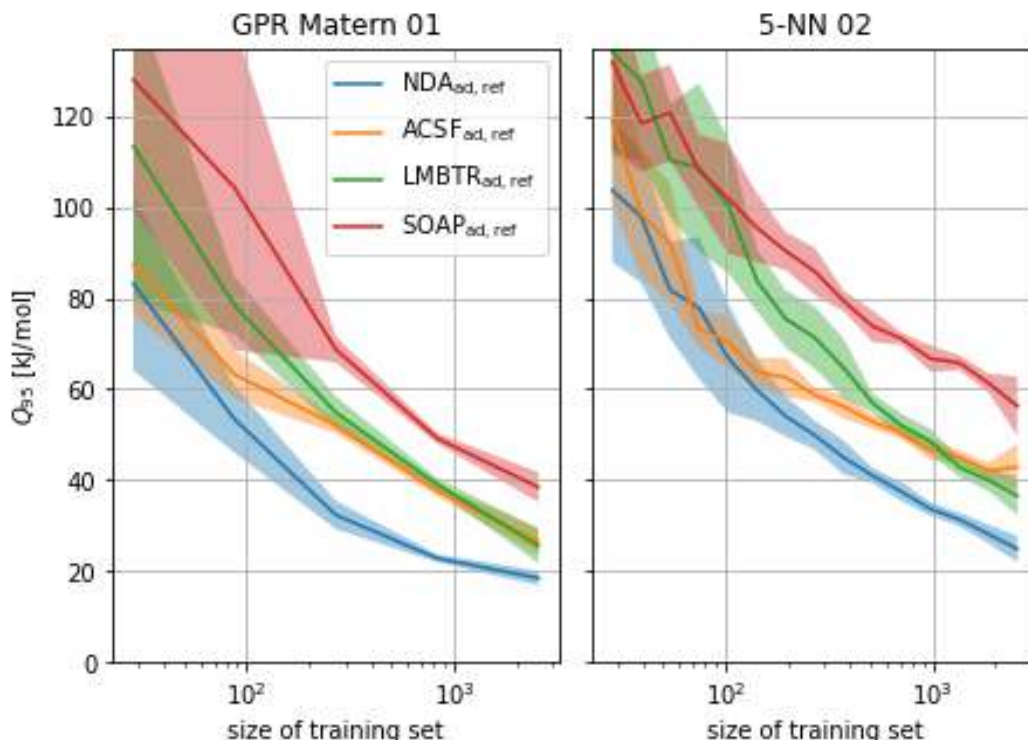


Figure 6.13: Learning curves for machine learning of adiabatic reference spin gaps for different structure based descriptors. Various scaled features based on molecular structures optimized separately for each spin state structures have been tested. The learning has been repeated 10 times with random train-test splits. Shaded areas display the standard deviation of the Q_{95} error measure for these 10 repetitions. As machine learning model Gaussian process regression with the Matern 01 ($1.0 * \text{Matern}(1.0, \nu = 0.5)$) kernel (left) and 5-nearest neighbors (weights='distance') (right) have been used.

6.6.3 Dependency between Performance and Molecular Structure

After comparing different descriptors with one another, the dependence on the how the molecular structure were generated will now be the focus for each of them. The dependence on the molecular structure is strongest for the NDA descriptor (Figure 6.14 and C.11, left). As originally expected, the adiabatic descriptor performs better than the vertical ones; reference structures perform best, and B3LYP-optimized structures performs slightly better than (5-NN singlet) or equal to PM6.

For GPR Matern 01 and $\text{NDA}_{s,\text{ref}}$, $\text{NDA}_{s,\text{B3LYP}}$ and $\text{NDA}_{s,\text{PM6}}$, large deviations of the predictions for the repeated ML processes are observed at some training set sizes. This indicates a significant dependence of the performance on the train-test split and underfitting due to the low number of feature dimensions (5). In Section 6.5 the structural changes by reoptimizing the molecular structures provided by QMspin with PM6 and B3LYP was analyzed by comparing the molecular structures in the representation of the descriptors. In the representation of ACSF, the changes for $\mathbf{R}_{\text{PM6}}^t$ were smallest and the changes for $\mathbf{R}_{\text{B3LYP}}^s$ were largest. For all other representations this trend was inverted. For that reason, it was expected that the performance of ACSF depends differently on the underlying molecular structures compared to the other descriptors. Compared to NDA it is less dependent on the molecular structure and shows different trends. Fitting problems have been observed for $\text{ACSF}_{\text{ad,ref}}$, which performs the worst, and the $\text{ACSF}_{s,\text{B3LYP}}$ structure. Among the ACSF descriptors, the best performance is achieved with $\text{ACSF}_{s,\text{CASSCF}}$ and $\text{ACSF}_{\text{ad,B3LYP}}$. For

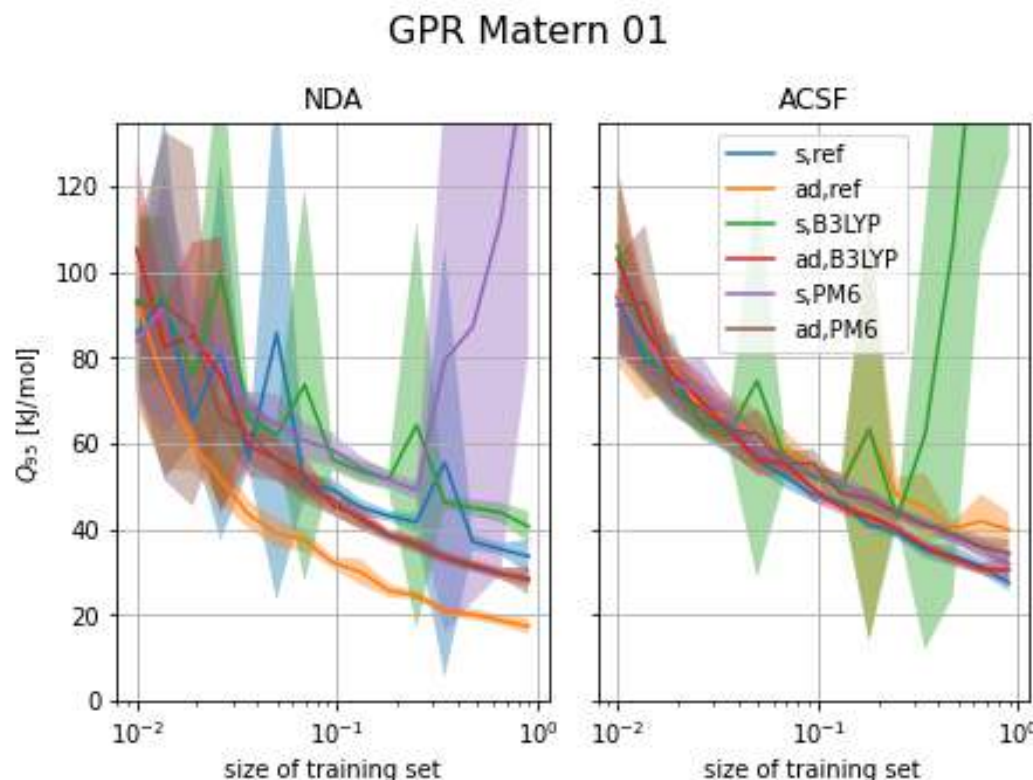


Figure 6.14: Learning curve for machine learning of adiabatic reference spin gaps. The GPR with Matern 01 kernel has been used on scaled features. The NDA (left) and ACSF (right) descriptor generated from different structures have been used. The learning has been repeated 10 times with random train-test splits. Shaded areas display the standard deviation of the Q_{95} error measure for these 10 repetitions.

the B3LYP-optimized structures by ACSF representation the largest RMSD was observed between singlet and triplet state. Even for the best ACSF versions, training sets larger than 2500 are needed to achieve performances similar to DFT.

As for ACSF, for LMBTR the dependence of the performance on the underlying molecular structures is small (Figure 6.15, left). For LMBTR, there is a larger dependence on the train-test split for train sizes smaller than 250. Surprisingly, the adiabatic reference structures perform worse than all other molecular structures in LMBTR representation. As already mentioned, the high dimensional SOAP descriptor does not perform well in general with GPR. The performance of single ML runs varies significantly with different train-test splits (Figure 6.15, left). This effect only decreases at large training set sizes (> 1000), which hints that the combination of large feature dimensions and comparable smaller training set sizes cause the observed fitting issues. With the k -NN model those fitting issues are not observed. For k -NN the general dependence on how the molecular structure was obtained is smaller compared to GPR (See SI Figure C.11 and C.13). Apparently, the identity of the closest neighbors does not change much with differently optimized molecular structures.

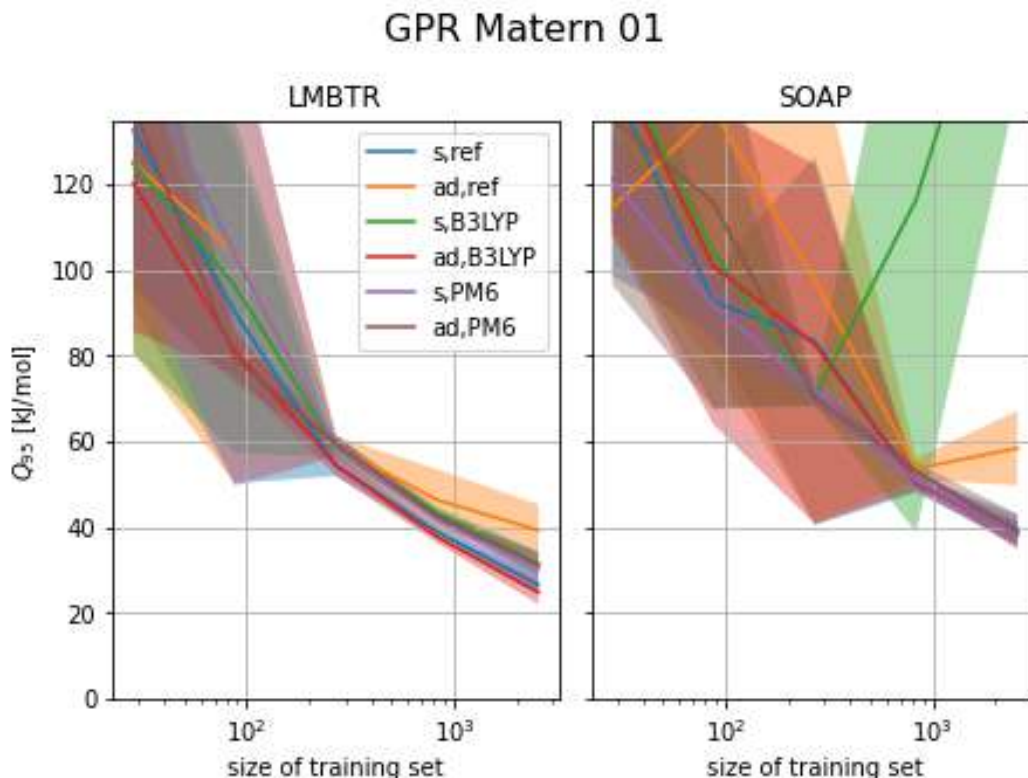


Figure 6.15: Learning curve for machine learning of adiabatic reference spin gaps. The GPR with Matern 01 kernel has been used on scaled features. The LMBTR (left) and SOAP (right) descriptor generated from different structures have been used. The learning has been repeated 15 times. Shaded areas display the standard deviation of the Q_{95} error measure.

6.6.4 Descriptors Based on Electronic Structure Properties

As mentioned above, using DFT-calculated properties as features in ML can combine both methods. In the overview plot (Figure 6.11) these type of descriptors are displayed under the groups ΔE^{DFT} , HL^{DFT} and CS^{M06} with DFT calculated spin gaps, HOMO-LUMO gaps and Mulliken CHarges and numbers of unpaired electrons as features, respectively. As this method is the most computationally expensive one discussed in this work, it is desirable to keep the number of computationally expensive reference calculations, and thus the training set size, small. If only single point calculations of one xc functional are involved in the generation of the features, this methods computational cost is identical to one of Δ -ML. If more xc functionals are employed, this method is computationally more expensive respectively.

For ΔE^{DFT} , HL^{DFT} and CS descriptors feature dimensions either have the same unit or values are compatible in magnitude. Hence, scaling of the features has no significant influence on the performance (Q_{95}). Both HL^{DFT} and CS descriptors do not yield any improvement compared to those only based on molecular structures. On the contrary, $\text{HL}_{\text{s,B3LYP}}^{\text{DFT}}$ performed very bad with $Q_{95} > 60 \text{ kJ mol}^{-1}$. The CS descriptor is mostly independent on how molecular structures are obtained.

As has been expected by the slim distributions of errors for the DFT calculated adiabatic gaps $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ and $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, using those as features results in very good performances $Q_{95} < 10 \text{ kJ mol}^{-1}$ for almost any ML model. The learning curve for different ML models using $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ as features shows that a simple linear regression works best for training

set sizes smaller than 1000 carbenes (Figure C.14). 100 carbenes are already enough to fit the linear regressor properly and there is no further improvement when increasing the training set size more. But already even smaller training set sizes of only 40 carbenes yields $Q_{95} < 10 \text{ kJ mol}^{-1}$. A comparison of the spin gap calculated with different molecular structures is shown in Figure 6.16. For all of them training set sizes of 100 carbenes are sufficient. The linear regression model includes the addition of a constant value. Therefore, if the training set is large enough to accurately represent the mean error, a considerable improvement compared to the DFT calculations is expected. Using $\Delta E_{\text{ad,ref}}^{\text{DFT}}$, that showed the most narrow error distributions, as features works best ($Q_{95} = 8 \text{ kJ mol}^{-1}$, training set size: 100). $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ as features also results in a very low Q_{95} of 14 kJ mol^{-1} . For $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$, the correction effect is much smaller. The Q_{95} value at a training set size of 100 is 29 kJ mol^{-1} while $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$ has a Q_{95} of 35 kJ mol^{-1} . For $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$, which showed Q_{95} values around 50 kJ mol^{-1} (see Figure 6.7), improvements to 38 kJ mol^{-1} could be made. For $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$, which showed very broadly distributed large errors with $Q_{95} > 160 \text{ kJ mol}^{-1}$, still have $Q_{95} > 80 \text{ kJ mol}^{-1}$ when used as features for ML.

Using only the spin gaps calculated with B3LYP or M06 as features (1-dimensional feature vector) still yields an improvement compared to the DFT calculations. However, the correction effect is considerably smaller compared to using spin gaps calculated with all 11 xc functionals as descriptors (Figure 6.16 right). This applies most notably to $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ and $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, as the largest differences in performance of individual xc functionals are observed for these. Future investigations might clarify which of the xc functionals are most important for good ML performance. Using only the most different performing xc functionals in the feature vector could, for instance, reduce the dimensionality of the feature vector and thus the number of necessary DFT calculations for its construction, while maintaining the good performance of the feature vector containing all 11 xc functionals.

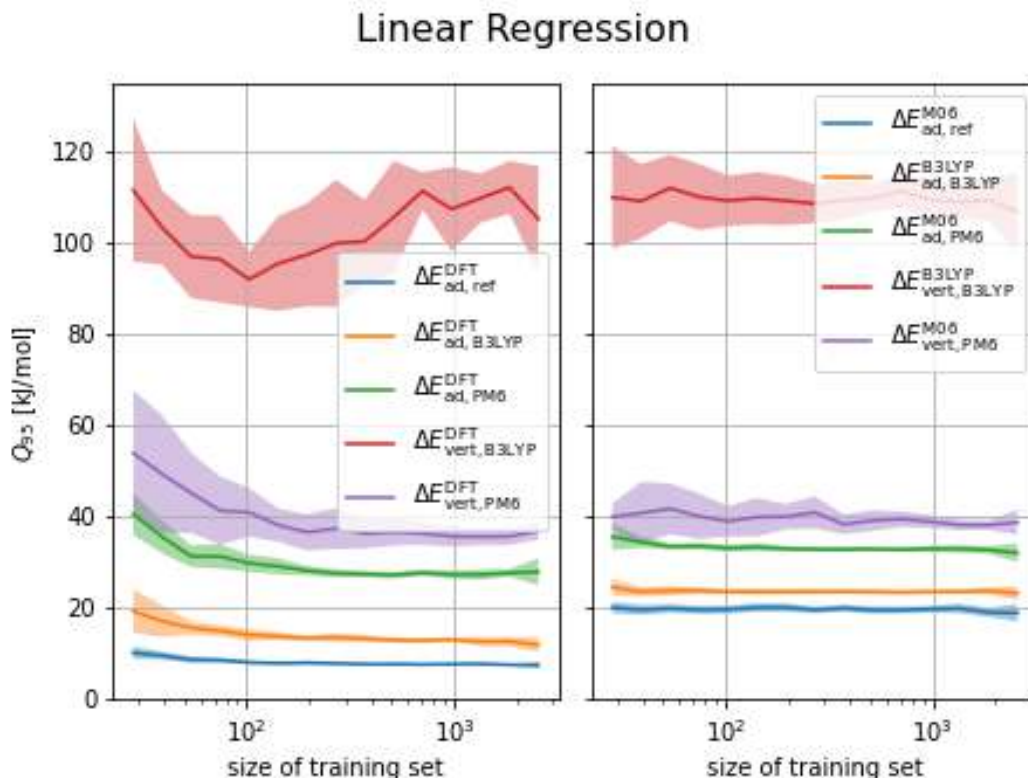


Figure 6.16: Learning curve for machine learning of adiabatic reference spin gaps. A linear regression has been performed together with spin gaps calculated with the 11 xc functionals (left) or only one xc functional (right) as features. The learning has been repeated 20 times. Shaded areas display the standard deviation of the Q_{95} error measure.

6.7 Delta-Machine Learning

The concept of combining less accurate method (here DFT) compared to a reference method (here MRCI), as we have applied by applying a linear regression on DFT calculated spin gaps, is also followed by Δ -ML. In contrast to the ML discussed in Section 6.6.4, where the adiabatic reference spin gaps are used the targets, for Δ -ML method errors are used as targets. Hence, in this Section the DFT errors ($\Delta\Delta E^{\text{DFT}} = \Delta E^{\text{DFT}} - \Delta E_{\text{ad,ref}}$) are used as the target for the machine learning process.

Δ -ML has been performed with $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$, $\Delta\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$, $\Delta\Delta E_{\text{ad,PM6}}^{\text{M06}}$, $\Delta\Delta E_{\text{vert,PM6}}^{\text{M06}}$, and $\Delta\Delta E_{\text{vert,B3LYP}}^{\text{B3LYP}}$ ³ used as targets. The overview of the performances is depicted in Figures 6.17, 6.19, 6.22, 6.23, and 6.20, respectively.

The plots contain the same ML models used in Section 6.6 on the y-axis, which are used with scaled and unscaled features as indicated on the y-axis. On the x-axis are the descriptors. For each of them, a version that includes the DFT-calculated spin gap is generated (right part of x-axis). The features are generated based only on molecular structures that have also been used for the DFT calculation of the spin gaps for which the errors are learned. Singlet and triplet molecular structures are used for adiabatic spin gaps, while for the vertical gaps only the singlet structures are used.

³For all spin gaps calculated only on B3LYP-optimized structures, the B3LYP xc functional was also chosen for the determination of the spin gap as after molecular structure optimization no further single point calculation would be necessary. In all other cases, the M06 xc functional was chosen because it showed to be the best performing one.

For comparability, in addition to Δ -ML (bottom plot), each figure contains an overview of the standard ML results (targets are reference gaps, top plot) using identical descriptors. Δ -ML should be an improvement compared the DFT calculated spin gaps. To visualize that, the scaling for the color code of Q_{95} values was adapted compared to the one used in Figure 6.11. A yellow color is indicative of no improvement compared to DFT calculations.

6.7.1 Adiabatic Spin Gaps with CASSCF- and B3LYP-Optimized Structures

Among the investigated xc functionals M06 showed the best performance for calculating the adiabatic spin gaps with the reference molecular structures (see Section 6.5). Hence, $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$ serves as target for the Δ -ML discussed in this section.

A first glance at the overview plot 6.17 (bottom) shows that the dependence on feature scaling is much smaller than for the ML models discussed in Section 6.6 as long as the spin gap $\Delta E_{\text{ad,ref}}^{\text{M06}}$ is not included in the feature vector. For the features that include the spin gaps or for the NDA descriptor, scaling results in an improvement of the performance by up to 9 kJ mol^{-1} . This probably originates from the different scale of values for the angle between the carbene carbon and its neighbors as well as the spin gaps compared to the other feature dimensions. Similar to the standard ML models, GPR with a RBF kernel generally does not perform well, while GPR with the Matern 01 kernel performs best and achieves Q_{95} values around 10 kJ mol^{-1} for almost any descriptor. Exceptions are only the AC, $\text{NDA}_{\text{s,ref}}$, $\text{DA}_{\text{s,ref}}$ and as well as the 1 dimensional descriptor $\Delta E_{\text{ad,ref}}^{\text{M06}}$. However, when combining any of $\text{NDA}_{\text{s,ref}}$ or $\text{DA}_{\text{s,ref}}$ with $\Delta E_{\text{ad,ref}}^{\text{M06}}$ good performances with $Q_{95} = 10 \text{ kJ mol}^{-1}$ are observed.

At the training set size of 1500 carbenes $\text{NDA}_{\text{ad,ref}}$ is the best performing descriptor with GPR ($Q_{95} = 6.9 \text{ kJ mol}^{-1}$). Therefore, descriptors based on the singlet and triplet molecular structures ($\mathbf{R}_{\text{ref}}^{\text{s+t}}$) have been used for comparing the dependence of the different descriptor types on the training set size (Figure 6.18 left). The learning curves for GPR Matern 01 with scaled features shows, that the addition of $\Delta E_{\text{ad,ref}}^{\text{M06}}$ to the feature vector does not influence the performance considerably. Also, the dependence on the choice of descriptor is low. As for the standard ML the NDA descriptor performs best. At training set sizes larger than 400 carbenes the fluctuation Q_{95} for different train-test splits becomes smaller than 1 kJ mol^{-1} and NDA shows a great performance of with $Q_{95} < 9 \text{ kJ mol}^{-1}$. A further increase of the training set to 1000 carbenes is necessary to lower Q_{95} by 1 kJ mol^{-1} . This is the best performance observed so far and still reduces the CPU time to a third (400 carbenes in test set) or half (1000 carbenes in test set) compared to the MRCI reference calculations.

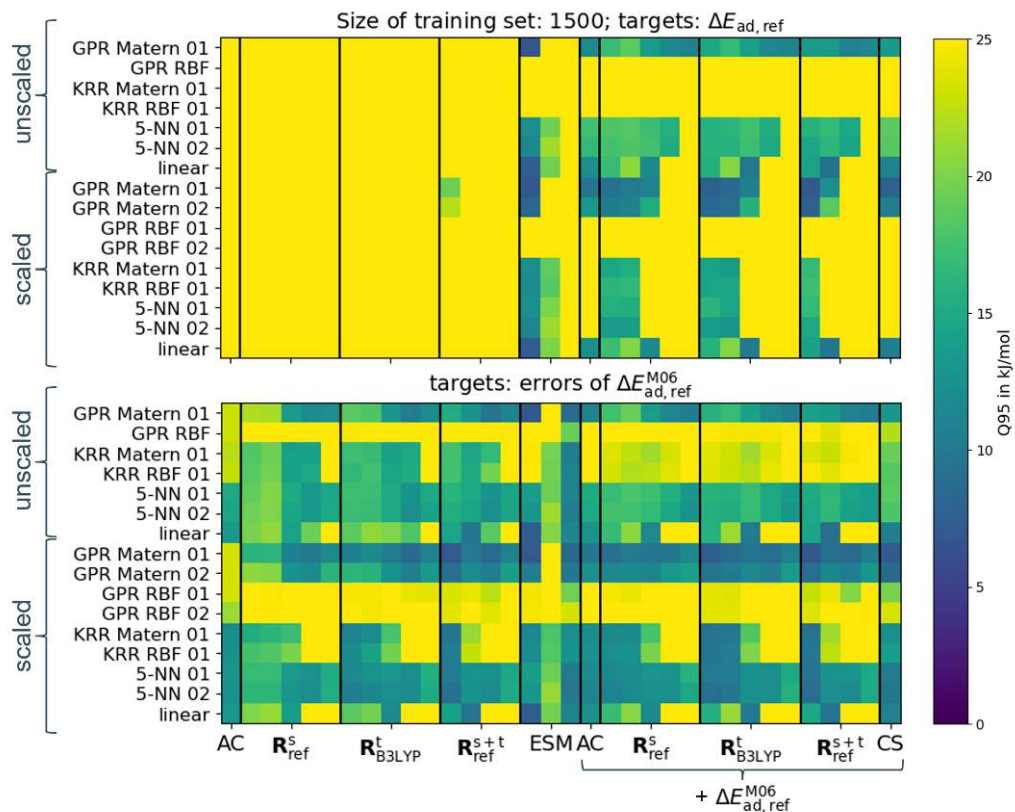


Figure 6.17: Overview on Q_{95} error measures for supervised machine learning (top, targets are the reference spin gaps) and Δ -machine learning (bottom, targets are errors of $\Delta E_{\text{ad,ref}}^{\text{M06}}$) with different models (y-axis) and descriptors (x-axis). The size of the training set is 1500 and the machine learning has been performed 10 times each. The squares are colored by the mean Q_{95} of the 10 runs for each combination of model and descriptor. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Descriptors grouped under the **R** label employ only singlet, triplet or singlet and triplet molecular structures indicated by superscribed 's', 't' or 's+t', respectively. The method used for molecular structure optimization is denoted by subscripts. Within this groups descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the 's+t' groups. The ESM group contains features obtained by electronic structure calculations. It includes $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ (11 feature dimensions), $\Delta E_{\text{ad,ref}}^{\text{M06}}$ (1 feature dimension) and $\text{CS}_{\text{ad,ref}}^{\text{M06}}$ (9 feature dimensions). To the feature vectors on the right side $\Delta E_{\text{ad,ref}}^{\text{M06}}$ has been added as feature dimension.

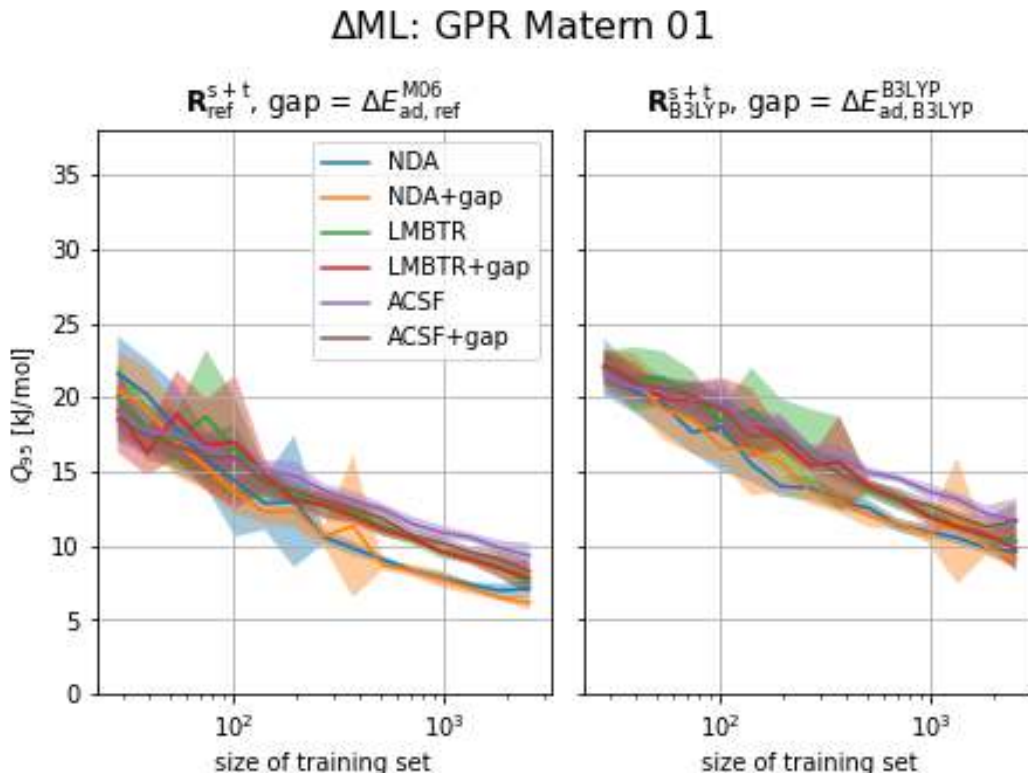


Figure 6.18: Learning curve for Δ -machine learning of $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$ (left) and $\Delta\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ (right). The features have been scaled prior to the test train split. The learning has been repeated 10 times. Shaded areas display the standard deviation of the Q_{95} error measure.

6.7.2 Spin Gaps with B3LYP-Optimized Structures

For the Δ -ML of $\Delta\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ ⁴, trends are very similar compared to the Δ -ML of $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$ (Figure 6.19). Feature scaling also has the largest effect when $\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ is included in the feature vector and GPR Matern 01 performs best among the ML models. The performance is generally worse compared to $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$. For the GPR Matern 01 the Q_{95} is worse by 5 kJ mol^{-1} for the Δ -ML of $\Delta\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ compared to $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$. Interestingly, and also different from the Δ -ML of $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$, the NDA descriptor only performs best compared to the other descriptors for the scaled version with $\mathbf{R}_{\text{B3LYP}}^{s+t}$ and the addition of $\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ (GPR Matern 01).

As can be seen from the learning curves in Figure 6.18 (right), in comparison to learning $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$, the performance is even less dependent on the descriptor choice for $\Delta\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$. Like for $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$, performance is strongly dependent on the training set size. The best performing descriptor $\text{NDA}_{\text{ad,B3LYP}}$ has a Q_{95} of 11 kJ mol^{-1} at a training set size of 1000 carbenes. Although, the CASSCF optimization is not required for the whole data set like for $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$, due to the larger training set needed there is also no advantage in terms of computational time compared to the Δ -ML of $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$. Depending on the model used to approximate the CPU time for the CASSCF optimization and MRCI single point calculation, the CPU time is reduced to 51 % or 36 % at training set sizes of 1000 carbenes.

⁴The different choice of xc functional ($\Delta E_{\text{ad,B3LYP}}^{\text{B3LYP}}$ vs. $\Delta E_{\text{ad,B3LYP}}^{\text{M06}}$) for the single point calculations used to determine the spin gaps does not have a large influence on the performance. The differences observed mostly result from the different molecular structures for the singlet carbenes.

Figure 6.20 shows the overview on the Δ -ML of $\Delta\Delta E_{\text{vert,B3LYP}}^{\text{B3LYP}}$. For the DFT-calculated vertical spin gaps $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$ very broad error distributions were observed (see Section 6.5). As expected this results in still large errors for Δ -ML of $\Delta\Delta E_{\text{vert,B3LYP}}^{\text{B3LYP}}$. The best performance is achieved with GPR Matern 01 and $\text{NDA}_{\text{s,B3LYP}} + \text{gap}$ descriptor at a training set size of 1500 with Q_{95} of 30 kJ mol⁻¹. At that training set size the Δ -ML of $\Delta\Delta E_{\text{vert,B3LYP}}^{\text{B3LYP}}$ is less efficient by computational time as well as by performance compared to calculating the spin gaps by DFT.

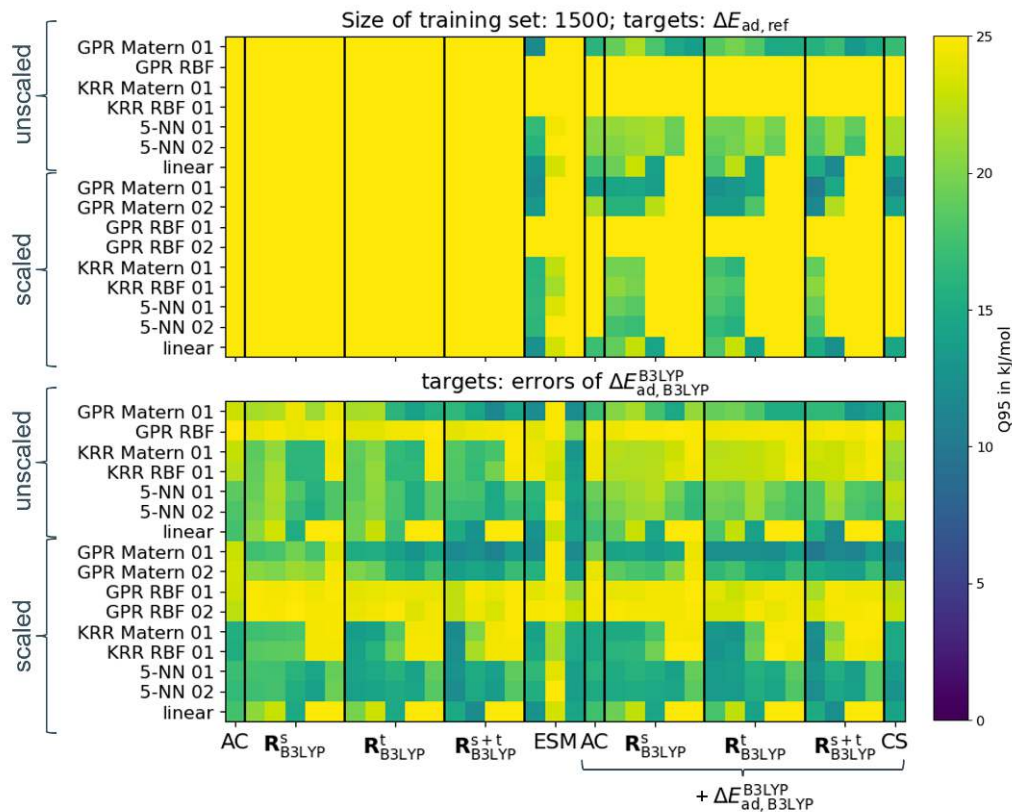


Figure 6.19: Overview of Q_{95} error measures for supervised machine learning (top, targets are the reference spin gaps) and Δ -machine learning (bottom, targets are errors of $\Delta E_{ad,B3LYP}^{B3LYP}$) with different models (y-axis) and descriptors (x-axis). The size of the training set is 1500 and the machine learning has been performed 10 times each. The squares are colored by the mean Q_{95} of the 10 runs for each combination of model and descriptor. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Descriptors grouped under the R_{B3LYP} label employ only singlet, triplet or singlet and triplet B3LYP-optimized molecular structures indicated by superscripted 's', 't' or 's+t', respectively. Within this groups, descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the 's+t' groups. The ESM group contains features obtained by electronic structure calculations. It includes $\Delta E_{ad,ref}^{DFT}$ (11 feature dimensions), $\Delta E_{ad,B3LYP}^{B3LYP}$ (1 feature dimension) and $CS_{ad,B3LYP}^{B3LYP}$ (9 feature dimensions). To the feature vectors on the right side $\Delta E_{ad,B3LYP}^{B3LYP}$ has been added as feature dimension.

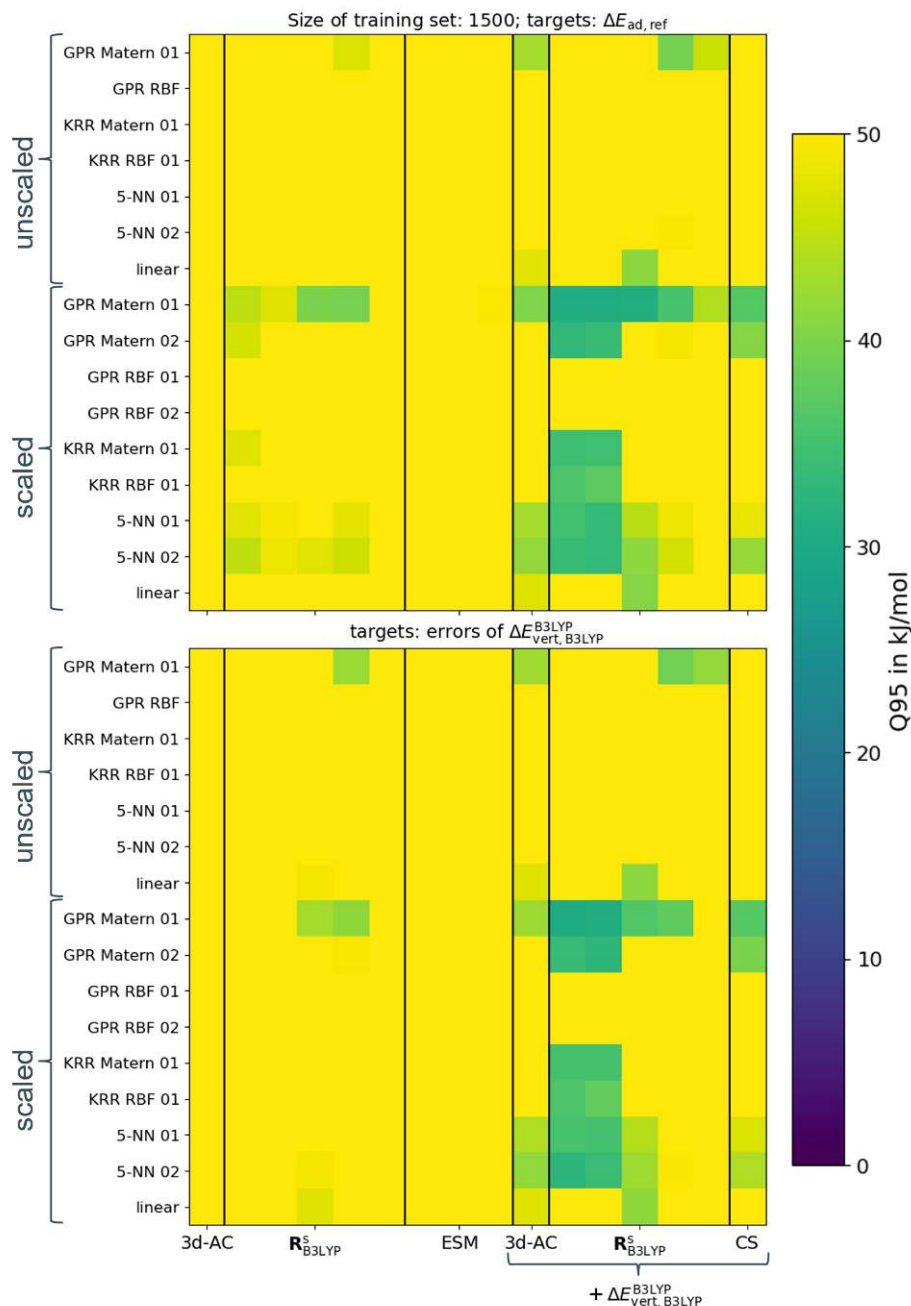


Figure 6.20: Overview of Q_{95} error measures for supervised machine learning (top, targets are the reference spin gaps) and Δ -machine learning (bottom, targets are errors of $\Delta E_{\text{vert}, \text{B3LYP}}^{\text{B3LYP}}$) with different models (y-axis) and descriptors (x-axis). The size of the training set is 1500 and the machine learning has been performed 10 times each. The squares are colored by the mean Q_{95} of the 10 runs for each combination of model and descriptor. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Structure based descriptors are grouped under the $\mathbf{R}_{\text{B3LYP}}^s$ label and employ B3LYP-optimized singlet molecular structure. Within this groups descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The ESM group contains features obtained by electronic structure calculations. It includes $\Delta E_{\text{vert}, \text{B3LYP}}^{\text{DFT}}$ (11 feature dimensions), $\Delta E_{\text{vert}, \text{B3LYP}}^{\text{B3LYP}}$ (1 feature dimension) and $\text{CS}_{\text{vert}, \text{B3LYP}}^{\text{B3LYP}}$ (9 feature dimensions). As indicated by the bracket, to the feature vectors on the right side, $\Delta E_{\text{vert}, \text{B3LYP}}^{\text{B3LYP}}$ has been added as feature dimension.

6.7.3 Spin Gaps with PM6-Optimized Structures

Lastly, we will look into Δ -ML of spin gaps based on PM6-optimized molecular structures, which are the computationally least costly in this study. The difference in computational cost between Δ -ML based on optimized molecular structures for one and for both spin states is negligible, due to in comparison much larger cost for the reference calculations needed for the training set.

As can be seen from the overview plots 6.22 and 6.22, even at large training set sizes of 1500 carbenes, no combination of model and descriptor results in better Q_{95} than 15 kJ mol^{-1} . The learning curves depicted in Figure 6.21 show that neither the Δ -ML of $\Delta\Delta E_{\text{ad,PM6}}^{\text{M06}}$ (left) nor the Δ -ML of $\Delta\Delta E_{\text{vert,PM6}}^{\text{M06}}$ (right) are strongly dependent on the descriptor type chosen. The best performance for Δ -ML of $\Delta E_{\text{ad,PM6}}^{\text{M06}}$ is observed for $\text{NDA}_{\text{ad,PM6}} + \text{gap}$ with GPR Matern 01. At a training set size of 1000 a Q_{95} of 18 kJ mol^{-1} is obtained.

Unfortunately, that neither results in satisfying errors smaller nor do the methods discussed in this section reduce computational time sufficiently at that training set sizes compared to the Δ -ML of $\Delta\Delta E_{\text{ad,ref}}^{\text{M06}}$ which achieved Q_{95} smaller than 10 kJ mol^{-1} at much smaller training set sizes.

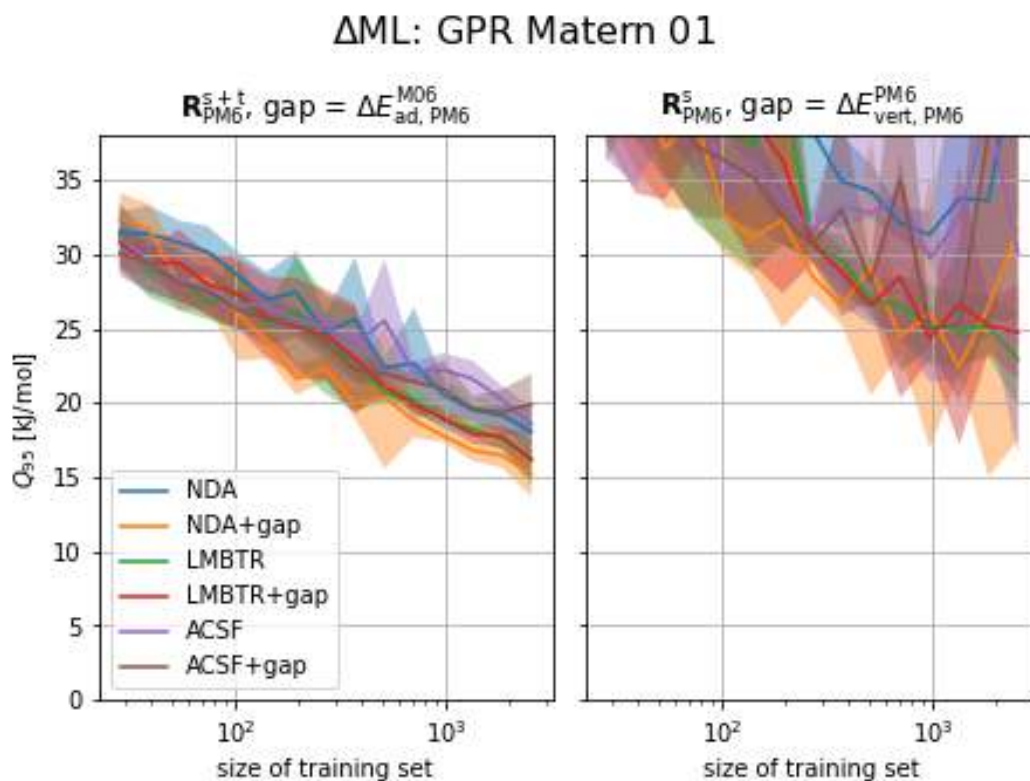


Figure 6.21: Learning curve for Δ -machine learning of $\Delta\Delta E_{\text{ad,PM6}}^{\text{M06}}$ (left) and $\Delta\Delta E_{\text{vert,PM6}}^{\text{M06}}$ (right). The features have been scaled prior to the test train split. The learning has been repeated 10 times. Shaded areas display the standard deviation of the Q_{95} error measure.

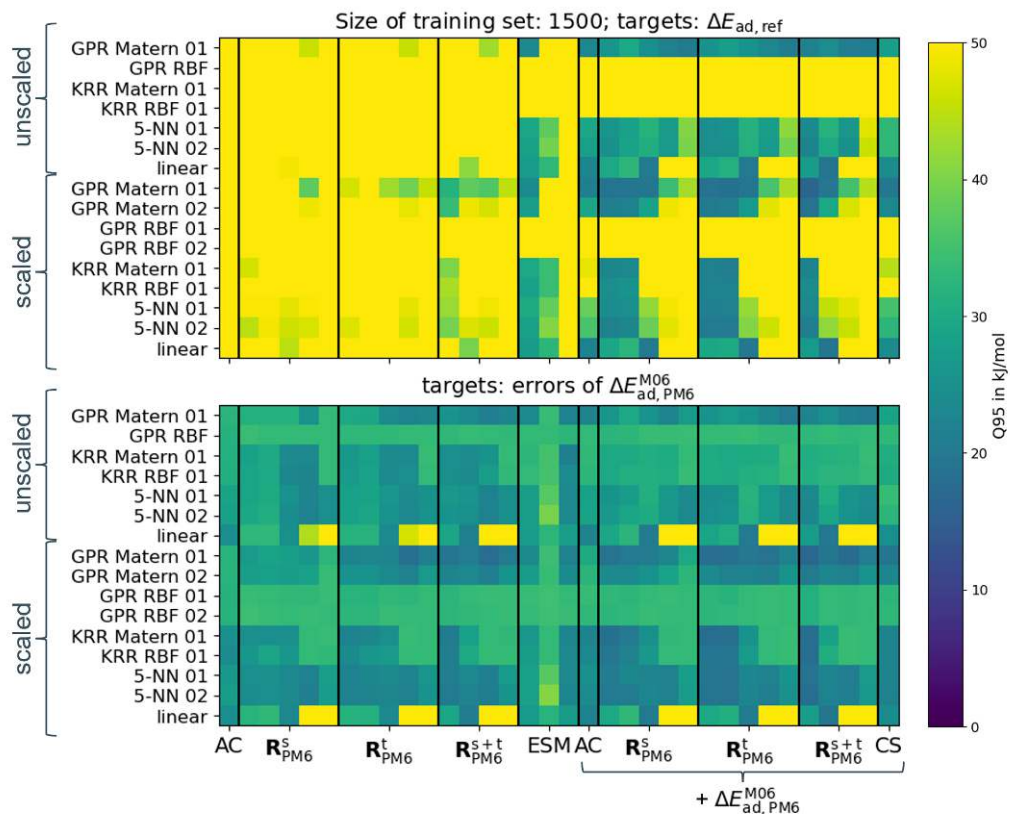


Figure 6.22: Overview of Q_{95} error measures for supervised machine learning (top, targets are the reference spin gaps) and Δ -machine learning (bottom, targets are errors of $\Delta E_{\text{ad},\text{PM6}}^{\text{M06}}$) with different models (y-axis) and descriptors (x-axis). The size of the training set is 1500 and the machine learning has been performed 10 times each. The squares are colored by the mean Q_{95} of the 10 runs for each combination of model and descriptor. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Descriptors grouped under the \mathbf{R}_{PM6} label employ only singlet, triplet or singlet and triplet PM6-optimized molecular structures indicated by superscripted 's', 't' or 's+t', respectively. Within this groups descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the 's+t' groups. The ESM group contains features obtained by electronic structure calculations. It includes $\Delta E_{\text{ad},\text{PM6}}^{\text{DFT}}$ (11 feature dimensions), $\Delta E_{\text{ad},\text{PM6}}^{\text{M06}}$ (1 feature dimension) and $\text{CS}_{\text{ad},\text{PM6}}^{\text{M06}}$ (9 feature dimensions). To the feature vectors on the right side $\Delta E_{\text{ad},\text{PM6}}^{\text{M06}}$ has been added as feature dimension.

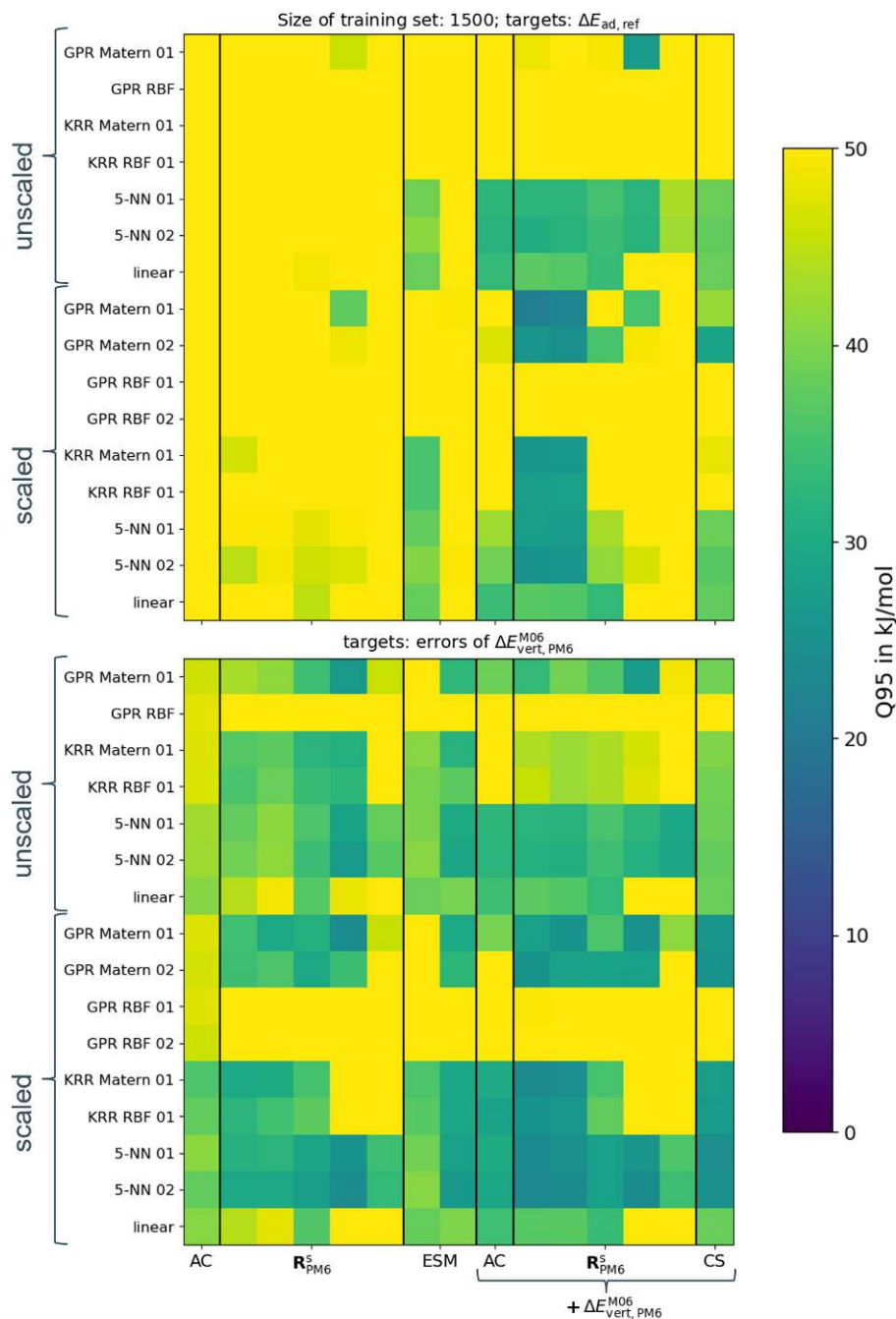


Figure 6.23: Overview of Q_{95} error measures for supervised machine learning (top, targets are the reference spin gaps) and Δ -machine learning (bottom, targets are errors of $\Delta E_{vert,PM6}^{M06}$) with different models (y-axis) and descriptors (x-axis). The size of the training set is 1500 and the machine learning has been performed 10 times each. The squares are colored by the mean Q_{95} of the 10 runs for each combination of model and descriptor. Features have been scaled for the top 7 models prior to the test-train split using the standard scaler. Structure-based descriptors are grouped under the R_{PM6}^s label and employ PM6-optimized singlet molecular structure. Within this groups descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The ESM group contains features obtained by electronic structure calculations. It includes $\Delta E_{vert,PM6}^{M06}$ (1 feature dimension) and $CS_{vert,PM6}^{M06}$ (9 feature dimensions). To the feature vectors on the right side $\Delta E_{vert,PM6}^{M06}$ has been added as feature dimension.

6.8 Conclusions

The QMspin data set utilized in this work provides CASSCF-optimized singlet carbene molecular structures as well as B3LYP-optimized triplet carbene molecular structures. Additionally, it includes MRCI-calculated adiabatic spin gaps ($\Delta E^{\text{ad,ref}}$) based on these structures [1]. In this study, we have generated additional sets of B3LYP-optimized singlet, as well as PM6-optimized singlet and triplet carbene molecular structures.

Using these differently optimized molecular structures, five sets of spin gaps were calculated using DFT, employing 11 xc functionals for each set. Adiabatic spin gaps were calculated based on the molecular structures provided by QMspin ($\Delta E_{\text{ad,ref}}^{\text{DFT}}$), purely B3LYP-optimized molecular structures ($\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$), and purely PM6-optimized molecular structures ($\Delta E_{\text{ad,PM6}}^{\text{DFT}}$). Vertical spin gaps were calculated based on B3LYP-optimized singlet structures ($\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$) and PM6-optimized singlet structures ($\Delta E_{\text{vert,PM6}}^{\text{DFT}}$).

In addition to DFT calculations, the same sets of molecular structures were used with different descriptors for machine learning (ML) of the reference spin gaps ($\Delta E^{\text{ad,ref}}$). Furthermore, the errors for each of the DFT-calculated spin gap sets were predicted using the Δ -ML approach.

The generated data allowed for a comprehensive analysis of different approaches to predicting spin gaps in carbenes and how these predictions depend on the origin of the molecular structures used.

The performance of DFT calculations strongly depends on the quality of the molecular structures and the chosen xc functionals. Surprisingly, the best performance was observed not with the reference molecular structures ($Q_{95} = 25.5 \text{ kJ mol}^{-1}$, $\Delta E_{\text{ad,ref}}^{\text{M06}}$), but with B3LYP-optimized molecular structures ($Q_{95} = 23.2 \text{ kJ mol}^{-1}$, $\Delta E_{\text{ad,B3LYP}}^{\omega\text{B97XD}}$). This is likely due to error compensation when using B3LYP-optimized singlet structures, which reduces the strong favoring of triplet carbenes by DFT. The better accuracy (smaller spreads of the error distributions) of $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ compared to $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ supports this conclusion. Calculating $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ reduces computational time by at least 80 % compared to MRCI-calculated spin gaps. However, a Q_{95} of 23.2 kJ mol^{-1} still represents a significant error for a spin gap, where values less than 10 kJ mol^{-1} would be desirable.

It is also worth noting that the dependence of spin gaps on the xc functional, or more precisely, the amount of exact exchange, observed for transition metal complexes as discussed in Chapter 5, differs for carbenes. For transition metal complexes, pure xc functionals favor lower spin states, whereas for carbenes, the same functionals most strongly favor the triplet state.

In comparison to the DFT-calculated spin gaps, the performance of ML predictions generally depends much less on how the molecular structures used for feature generation were obtained. A simple descriptor specifically designed for carbenes performed best compared to the other general-purpose descriptors ACSF, LMBTR, and SOAP. This descriptor, which includes the atom type of the neighbors of the carbene carbon as well as their bond lengths and angles (NDA), highlights the importance of problem-driven thinking.

Unlike the other descriptors, NDA depends more strongly on the origin of the molecular structure. Using both singlet and triplet molecular structures from the QMspin data set ($\text{NDA}_{\text{ad,ref}}$) is required for a Q_{95} below 23 kJ mol^{-1} at training set sizes of 1000 carbenes (ML model: GPR with Matern kernel). However, the computational costs for this prediction exceed those of $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$ by a factor of 2.5. Additionally, even at larger training set sizes, the performance is not significantly improved, making this method less suitable for practical application.

Δ -ML, which targets the error of DFT-calculated spin gaps, was performed separately for spin gaps calculated based on each set of molecular structures. Δ -ML performs best for $\Delta E_{\text{ad,ref}}^{\text{M06}}$, which showed the greatest accuracy among the DFT-calculated spin gaps. Again,

the NDA descriptor showed the best results in combination with GPR (Matern kernel). At small training set sizes of 400 carbenes, good performances of $Q_{95} = 9 \text{ kJ/mol}$ were observed.

Although this method has longer computational times compared to $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, its performance is greatly improved. Given that the CPU time is still reduced to a third compared to MRCI-calculated spin gaps, this method is worth considering, especially for large data sets.

Additionally, the performance might be further improved by optimizing the descriptor. A promising starting point could be combining NDA with atomic spins on the carbene carbon, for which a correlation to spin gap errors was observed. Furthermore, automated feature selection algorithms could be a valuable tool for exploring feature dimensions.

Part II

Synthesis and Characterization

7 Synthesis and Characterization of Dinuclear Iron Complexes

This chapter is based on a project started by Sebastian Sandl.[55] Therefore, catalytic studies, a first report of complexes **2**, **3a** and **4** and some reactivity studies with **2** have already been published in his thesis. Furthermore, following authors contributed to this chapter: The measurement of Mössbauer spectra and SQUID magnetometry have been performed by Serhiy Demeshko in the group of Frank Meyer in the University of Göttingen. ESI-MS measurements have been performed in cooperation with Konrad Koszinowski from the University of Göttingen. Catalytic studies (Section 7.2) have been performed by Matteo Villa and Sebastian Sandl. Reactivity studies have been performed by Sebastian Sandl (Figure 7.17 (a) A and Figure 7.18 (a)) and Ursula Rastetter. CV measurements, DFT calculations and writing have been done by Ursula Rastetter.

7.1 Introduction to Polynuclear Hydrido-Bridged Iron Complexes

The replacement of noble metal catalysts with more abundant, affordable, and less toxic 3d transition metals, especially iron, has garnered significant interest in recent years.[2–10] Homogeneously catalyzed hydrogenation reactions represents a key transformation in organic synthesis, pivotal for the production of a wide range of chemical products. In these catalytic processes, hydride species often serve as crucial intermediates.[46–51] Despite the importance of these intermediates, there are relatively few examples of polynuclear iron complexes featuring bridging hydrides, highlighting an area with significant potential for further exploration and development.

An early example reported by Sacconi and coworkers in 1974 is the cationic $[\text{Fe}_2\text{H}_3(\text{p}_3)_2]^+$, which was obtained by reacting an iron(II) salt with the ligand in the presence of NaBH_4 . [163] The compound was analyzed by activation energy (E_a) and UV-vis spectroscopy. Single crystal XRD was used to identify the molecular structure, revealing an Fe-Fe distance of 2.332(3) Å. In the ^1H -NMR spectrum of the diamagnetic compound, the bridging hydrides produce a septet at $\delta = 21.13$ ppm due to coupling with the phosphorus nuclei of the ligand. Another indicator for bridging hydrido ligands is the IR band at 1048 cm^{-1} , which shifts to 790 cm^{-1} for the deuterido complex. The hydride content was additionally quantified by reaction with an excess of hydrochloric acid, that led to the evolution of 3 equivalents of H_2 . In 1992, Shilov *et al.* reported anionic dinuclear hydridoferrates with aryl ligands and coordinated counter cations.[164] In diethyl ether, $\text{Li}_4[\text{FePh}_4]$ was synthesized *in situ* by the reaction of FeCl_3 with PhLi . Subsequent addition of H_2 gas resulted in the mononuclear dihydridoferrate $\text{Li}_4[\text{FePh}_4\text{H}_2]$, which was recrystallized from THF to yield the dinuclear trihydridoferrate $\text{Li}_5[\text{Ph}_3\text{Fe}(\mu\text{-H})_3\text{FePh}_3]$. In contrast to the mononuclear hydride, the dimer reacts with N_2 and forms hydrazine upon subsequent addition of HCl .

Several neutral, dinuclear tetrahydridoiron complexes with cyclopentadienyl (Cp) derivatives have been reported by Suzuki ($\text{Cp}^* = \text{C}_5\text{Me}_5$), White ($\text{Cp}' = 1,2,4\text{-(Me}_3\text{C)}_3\text{C}_5\text{H}_2$), Qu ($\text{Cp}'' = \text{C}_5\text{iPr}_4\text{H}$), and coworkers.[165–168] The syntheses involve the reduction of monomeric (Cp^*) or dimeric (Cp' and Cp'') precursor complexes with LiAlH_4 , KBET_3H , or reduction under H_2 pressure. The crystal structures reveal the Fe-Fe distances to be very short (2.1989(5) Å to 2.2034(8) Å). These reported tetrahydrides have been analyzed by Mössbauer spectroscopy, which shows well-resolved doublets and supports the presence of two equivalent low-spin Fe^{II} . [166, 168, 169] This is also in agreement with the NMR spectra, which show diamagnetic signals. The bridging hydrides exhibit a singlet with a chemical shift around -22 ppm in the ^1H -NMR spectra. Variable temperature NMR spectroscopy of $[(\text{cp}^*)\text{Fe}(\mu\text{-H})_4\text{Fe}(\text{cp}^*)]$ revealed the complex as a classical metal hydride with no bonding interaction between the H atoms.[165, 167] In contrast to analogous ruthenium complexes, $[(\text{cp}^*)\text{Fe}(\mu\text{-H})_4\text{Fe}(\text{cp}^*)]$ is more reactive: C-H activation was observed in C_6D_6 , with complete H/D exchange achieved after 75 h.[165] Also, the H_2 activation was faster than in the ruthenium complex, and D/H exchange of $[(\text{cp}^*)\text{Fe}(\mu\text{-D})_4\text{Fe}(\text{cp}^*)]$ is completed in 9 h. Further reactivities involve Si-H, P-H, and B-H activation,[165, 170, 171] the reaction with P_4 , [167] and S_8 . [168] Addition of MeCN leads to the evolution of H_2 and the formation of dimeric MeCN complexes.[168]

Holland and coworkers synthesized dihydride-bridged diiron complexes supported by β -diketiminato ligands.[172–177] The neutral dihydride complexes were synthesized by reduction of the precursor $[\text{Fe}^{(\text{R},\text{R}')}\text{Nacnac}^{\text{Ar}}\text{Cl}]$ (where $\text{Nacnac} = \beta$ -diketiminato) with KBET_3H . An equilibrium between the dimer and the monomeric three-coordinate hydride complex LFeH in solution was determined using ^1H -NMR spectroscopy. While the magnetic moment for the monomer is consistent with a high-spin Fe^{II} , the dimer shows a lower magnetic moment resulting from the antiferromagnetic coupling of both iron centers.[172] Additionally, the monomer was trapped with 4-*tert*-butylpyridine. Characteristic Fe-H bands in IR and Raman spectra have not been observed. A structurally closely related dihydrido iron complex bearing unsymmetrical enamido-phosphinimine (NpN) ligands was reported by Fryzuk

and coworkers.[178] Both the Nacnac and the NpN hydrido iron complexes show insertion reactivity towards 3-hexyne, azobenzene, and 1-azidoadamantane[172, 174, 178] and are active in the catalytic hydrodefluorination of perfluorinated arenes in the presence of silanes.[173, 178] Murray and coworkers used a macrocyclic tris(β -diketiminate)cyclophanate ligand to synthesize a trimeric hydrido complex with a $[\text{Fe}_3(\mu\text{-H})_3]^{3+}$ core, which forms a planar cycle. The Fe-Fe distances, ranging from 3.2570(6) Å to 3.3561(5) Å, are much larger compared to the dinuclear iron complexes.[179] The iron centers are each in a tetrahedral coordination environment and are assigned as $\text{Fe}^{\text{II,hs}}$ by Mössbauer spectroscopy (one doublet at 80 K, $\delta = 0.79 \text{ mm s}^{-1}$, $\Delta E_{\text{Q}} = 2.34 \text{ mm s}^{-1}$).

Using a combination of X-ray diffraction, Mössbauer spectroscopy, and magnetic measurements at variable temperatures as well as neutron diffraction, Holland and coworkers were able to capture the dynamic nature of the iron sites and illustrate the motion of the bridging hydrides of complex $[(^{\text{Me}}\text{Nacnac}^{\text{Ar}})\text{Fe}(\mu\text{-H})_2\text{Fe}(^{\text{Me}}\text{Nacnac}^{\text{Ar}})]$ in a unique way.[177] While at room temperature the Mössbauer spectrum and magnetic measurements are consistent with two equivalent high-spin Fe^{II} centers, similar to the previously described $[(^{\text{R,R'}}\text{Nacnac}^{\text{Ar}})\text{Fe}(\mu\text{-H})_2\text{Fe}(^{\text{R,R'}}\text{Nacnac}^{\text{Ar}})]$ complexes, cooling the crystals below 170 K leads to splitting in the X-ray diffraction data and in the Mössbauer spectrum (Figure 7.2). A phase change to a lower-symmetry space group occurs at lower temperatures, resulting in two inequivalent iron centers. One was found to be tetrahedrally coordinated and of high-spin, and the other has a square planar structure and intermediate spin. Rotation of both hydrido ligands leads to the change of the coordination sphere from tetrahedral to square planar and vice versa for each iron center. A similar coordination pattern was observed for the bis(carbene)borate ligand-bearing dihydride reported by Smith and coworkers in 2019.[180] With the help of magnetic measurements and Mössbauer spectrometry, they described the metal centers as iron(II) S=1 (square planar) and iron(II) S=2 (tetrahedral) with an S=3 ground state.

Reduction of $[(^{\text{tBu,H}}\text{Nacnac}^{\text{Dipp}})\text{Fe}(\mu\text{-H})_2\text{Fe}(^{\text{tBu,H}}\text{Nacnac}^{\text{Dipp}})]$ with potassium graphite produces a dianionic dihydridoferrate. Unlike the other hydrido iron complexes discussed here, this dimer is stabilized by potassium cations that interact with the aryl groups of the β -diketiminate ligands.[176] Analysis of Mössbauer spectra and magnetic measurements suggests that both iron centers are weakly antiferromagnetically coupled high-spin iron(I). Introducing potassium chelating agents into a solution of this hydridoferrate led to the formation of the monomeric hydridoferrate, which also consists of high-spin Fe^{I} .

Gomez and coworkers investigated a system comprising a precursor iron(II) complex with an iminopyrrolyl ligand and KBet_3 for the catalytic hydroboration of olefins.[181] Mechanistic studies of this system revealed a mixture of a monomeric iron hydrido complex, a dihydrido-bridged dimer, and a neutral toluene complex, demonstrating the dynamic nature of this system. In the dimer, the two distinct iron centers can be observed through ^{57}Fe Mössbauer spectroscopy.

Examples with PNP pincer ligands differ structurally from the previous examples, where each iron center is coordinated by one ligand and the Fe-Fe axis lies mostly in the plane with the ligand planes (Figure 7.1). Thompson and coworkers synthesized a dihydrido iron dimer with a pyrrole-based pincer ligand. Uniquely, each ligand in this complex coordinates both iron centers. Measurements of the magnetic moment by NMR spectroscopy indicates the presence of $\text{Fe}^{\text{II,hs}}$ centers.[182] Similarly, Nishibayashi developed a hydride with a derivative of the pyrrole-based pincer ligand and utilizing it in boron-catalyzed sp^3 C-H borylation of anisole derivatives.[183] Carbazole-based PNP pincer ligands were synthesized by reducing the chlorido precursor with KBet_3 or treating monomeric alkyl complexes with H_2 . Each iron atom's coordination sphere is highly distorted square-pyramidal, with twisted $[\text{Fe}(\text{PNP})]$ fragments. Magnetic susceptibility measurements based on Evan's method suggest the iron centers as $\text{Fe}^{\text{II,hs}}$ with significant antiferromagnetic coupling. This complex also catalyzes the hydrogenation of alkenes.[184]

Table 7.1: Overview on Fe-Fe distances and Mössbauer parameters for polynuclear iron hydrides.

Reference	$d(\text{Fe-Fe})$ in Å	Mössbauer parameters $\delta(\Delta E_Q)$ in mm s^{-1}
<i>Sacconi</i> 1974 [163]	2.332(3)	-
<i>Shilov</i> 1992 [164]	2.389(1), 2.379(1)	-
<i>Suzuki</i> 2000 [165, 166]	2.202(2)	90 K: 0.271(1.801)
β -diketiminates [172, 175, 177]	2.46 to 2.62	80 K: 0.60(0.74) to 0.43(2.16)
<i>Fryzuk</i> 2017 [178]	2.6762(4)	-
<i>Murray</i> 2015 [179]	-	80 K: 0.79(2.34)
<i>Bill, Hoffman, Holland</i> 2012 [176]	-	80 K: 0.48(1.90)
<i>Smith</i> 2019 [180]	2.5247(4)	80 K: 0.35(3.77), 0.51(1.92)
[185]	2.5719(5)	-
<i>Tonzetich</i> 2017 [182]	2.4796(7)	-
<i>Gade</i> 2018 [184, 189]	2.5880(7)	-
[190]	2.86	-
<i>Gomes</i> 2019 [181]	2.4874(5)	4 K: 0.69(3.08), 0.96(1.92)
[185]	2.5719(5)	-
<i>Ohki</i> 2017 [Fe_4][188]	2.513 to 2.769	78 K: 0.178(1.189), 0.592(1.811)
<i>Ohki</i> 2017 [Fe_6][188]	2.434 to 2.623	78 K: -0.259 (0.391), 0.482(0.650)

A recent example reported by Tamm and coworkers features an unsymmetrical amino-imidazolin-2-imine ligand. Again, antiferromagnetically coupled high-spin Fe(II) centers are determined based on solid-state magnetic measurements. This complex serves as pre-catalysts in H/D exchange reactions with deuterium (D_2) in hydrosilanes.[185]

Apart from the dimeric structures discussed above, the bis(hexamethyldisilazide) ligand gives rise to polynuclear iron clusters in a broad variety of structures (Figure 7.1). Jacobi von Wangelin and coworkers have reported the reduction of $[\text{Fe}(\text{HMDS})_2]$ using Dibal-H which resulted in the isolation of Fe_4 , Fe_6 , and Fe_7 clusters. These clusters were characterized by crystallographic structure analysis and LIFDI-MS and they are capable of hydrogenating a wide range of tri- and tetrasubstituted alkenes under mild conditions.[186, 187] Ohki's group utilized the same precursor but employed pinacolborane as the reducing agent. By adding phosphines, they accessed various clusters with different oxidation and spin states of the iron centers, which were characterized by Mössbauer spectroscopy. The hydride bridges were analyzed by deuteration and the comparison of IR spectra between hydrido and deuterido clusters, similar to the approach used by Sacconi and coworkers.[188] These clusters were also employed for the catalytic silylation of N_2 in the presence of sodium and Me_3SiCl .

The examples above have demonstrated that the synthesis of the variously applied polynuclear hydrides in catalysis can be achieved through reduction using hydride sources such as alkali tetraethylborohydrides or the direct addition of H_2 . Their isolation and characterization pose challenges due to their high sensitivity and paramagnetic nature. Key methods for characterization include Mössbauer spectrometry, magnetic measurements, deuteration coupled with IR spectroscopy comparison, mass spectrometry, and single crystal X-ray diffraction. An overview of the Fe-Fe distances and typical Mössbauer parameters for the iron hydrides can be found in Table 7.1

In this work, the stabilization of highly reduced ferrates by the redox-active ligand bis(imino)acenaphthene (BIAN) was considered. BIANs are easily synthesized from commercial precursors on multigram scales and are highly redox-active, capable of harboring up to four electrons.[53, 54] Recently, BIAN was employed for the synthesis of a related dimeric cobaltate with three bridging hydrides, which served as an active precatalyst for the hydrogenation of alkenes.[52] Consequently, the focus of this study is on the catalytic reactivity

of a system involving the analogous iron precursor as well as the isolation, characterization, and reactivity studies of the derived ferrate complexes.

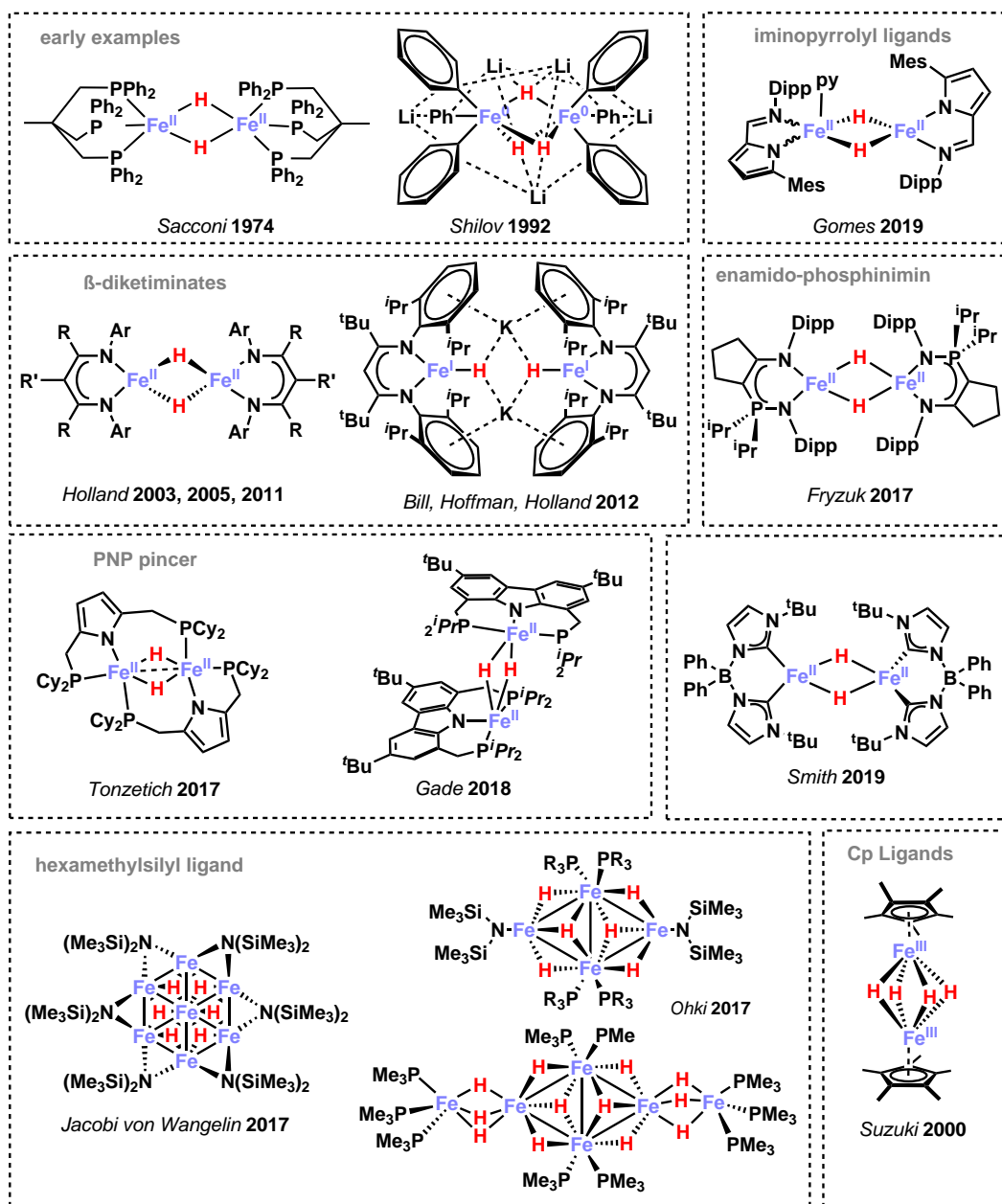


Figure 7.1: Reported polynuclear ironhydride complexes.

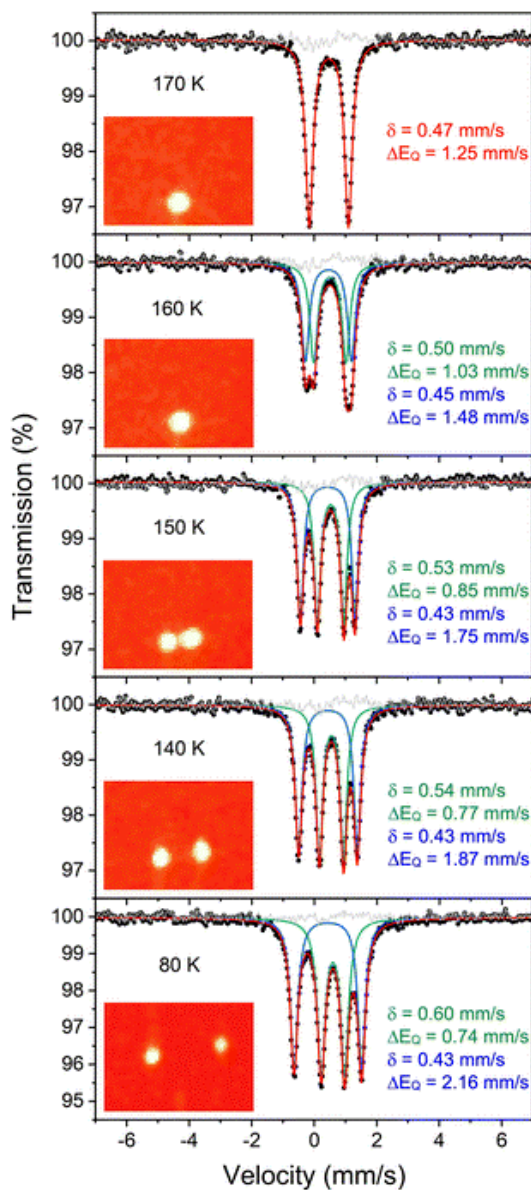


Figure 7.2: Mössbauer data of crystalline $[(\text{Me,MeNacnac}^{\text{Ar}})\text{Fe}(\mu\text{-H})_2\text{Fe}(\text{Me,MeNacnac}^{\text{Ar}})]$ (with Ar = 2,6-dimethylphenyl, toluene-free crystal). Each orange inset shows an equivalent section of the diffraction pattern, corresponding to the $(-8 \ -4 \ 6)$ reflection in the high-temperature indexing in C2/c, and the splitting of the reflection at low temperature is a result of conversion to the twinned structure. Each main panel shows the Mössbauer data (black circles) with the associated fits (red) that are the sum of components for the tetrahedral site (green) and the square-planar site (blue). The residuals are shown in grey. Reproduced from Ref. [177] with permission from the Royal Society of Chemistry.

7.2 Catalyzed Hydrogenation of Alkenes with ^{Dipp}BIAN Iron Complexes

The catalytic studies shown in this section have been performed in preliminary work.[55, 95]

As for the ^{Dipp}BIAN (**L**) cobalt system, optimizations of the reducing agent in previously reported catalytic systems of [(**L**)FeCl₂] **1**/ 3 *n*-BuLi[95] revealed that LiBEt₃H is even more suitable (Tab. 7.2).[55]

Table 7.2: Catalytic optimization studies on reducing agent. 0.25 mmol triphenylethylene were hydrogenated. Yields and conversions are determined by quantitative GC-FID vs. *n*-pentadecane as internal standard.[55]

Entry	Reductant	Yield (Conversion)
1	<i>n</i> -BuLi	31 % (31 %)
2	<i>i</i> -PrMgCl	51 (51)
3	LiBEt ₃ H	77 (79)
4	NaBEt ₃ H	3 (3)
5	KBEt ₃ H	2 (2)
6	L-selectride	62 (62)
7	N-selectride	1 (1)
8	DiBAL-H	25 (25)

Similar to related cobalt-containing catalytic systems, a strong alkali cation effect was observed.[52, 191] Quantitative hydrogenation of alkenes under mild conditions was achieved (20–60° C, 2–10 bar) (Figure. 7.3). Even challenging alkenes such as triphenylethylene were hydrogenated quantitatively under these conditions. [Li][Fe(**L**)(toluene)] is assumed to be the active species in the hydrogenation catalysis. It was detected in the catalyst solution by LIFDI-MS (*m/z* = 655) alongside the neutral [Fe(**L**)(toluene)] **6**, which was previously reported by Findlater *et al.*[192] but found to be inactive as a catalyst for alkene hydrogenation under the conditions described above.[55, 95] On the other hand, the tetrahydridoferrate **2**, which will be discussed in greater detail in section 7.3, proved to be an active precatalyst (see Figure. 7.3).[55] Notably, **1**/3 LiBEt₃H and **2** produced a very similar product ratio of the product from competing alkene hydrogenation and the product from hydrogenative ring-opening of α -cyclopropyl styrene (Figure 7.3 bottom). This key experiment suggests a closely related mechanism for the hydrogenation with the *in situ* reduced precursor and the isolated hydridoferrate **2**.[55]

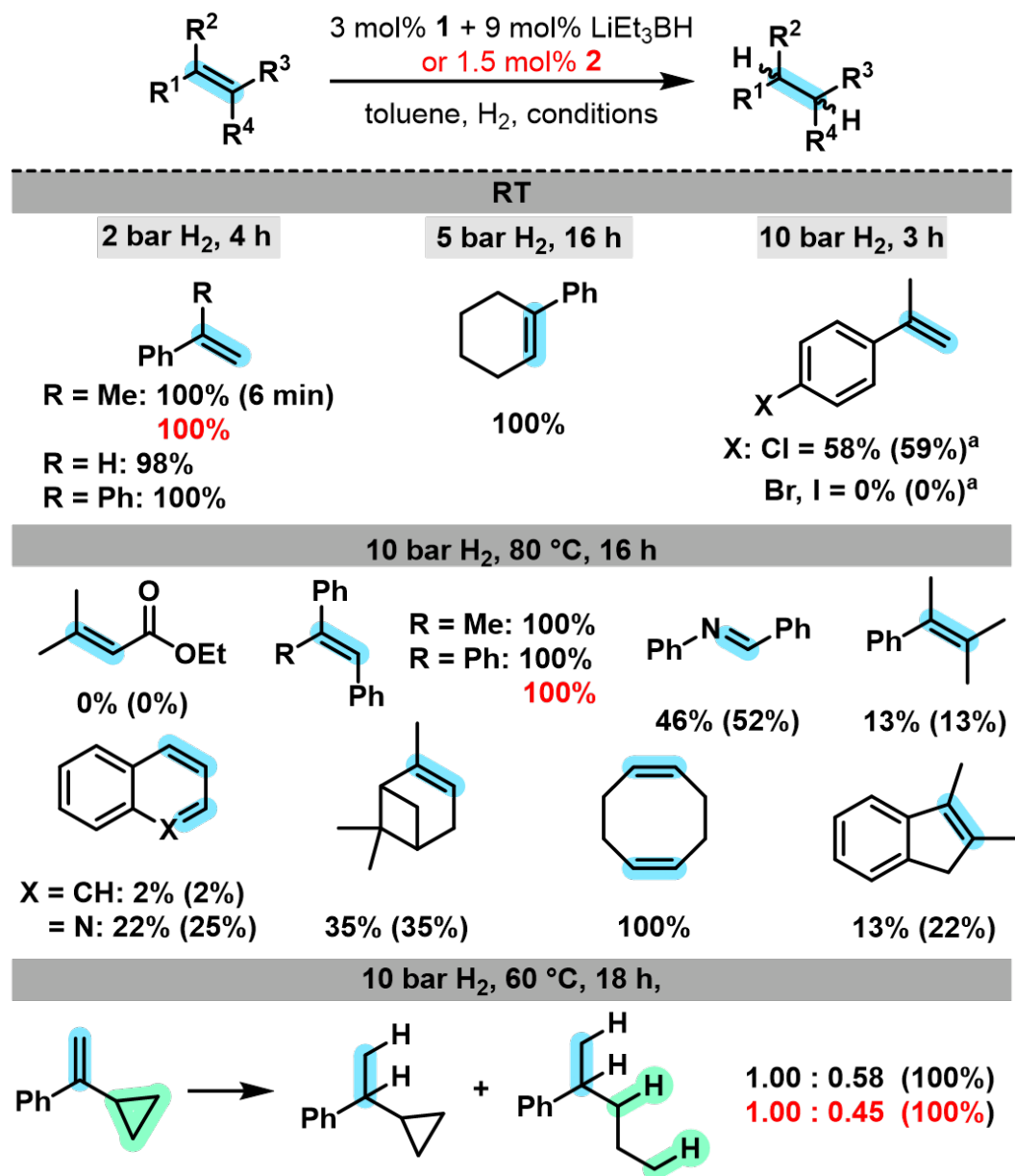


Figure 7.3: Catalytic hydrogenation of alkenes. Yields are determined by quantitative GC-FID vs. *n*-pentadecane. Conversions are given in parenthesis if <95%. Blue highlighting indicates π -bonds being hydrogenated. Green highlighting indicates sites for hydrogenative ring-opening. ^a traces of hydrodehalogenation.

7.3 Synthesis and Characterization of Reduced DippBIAN Iron Complexes

To synthesize hydridoferrates, the stable precursor $[\text{LFeCl}_2]$ was reduced with 3 equivalents of an alkali metal triethylborohydride (AMBET_3H , $\text{AM} = \text{Li}, \text{Na}^1$) in Et_2O in a closed reaction vessel.[193] Gas evolution was observed during the reduction, presumably due to the formation of H_2 . The products were precipitated with *n*-hexane. Afterwards subsequent extraction with Et_2O and THF afforded the dianionic tetra- and dihydridoferrate **2** and **3**, respectively, as the main products with yields up to 42 % [55] (Figure 7.4).

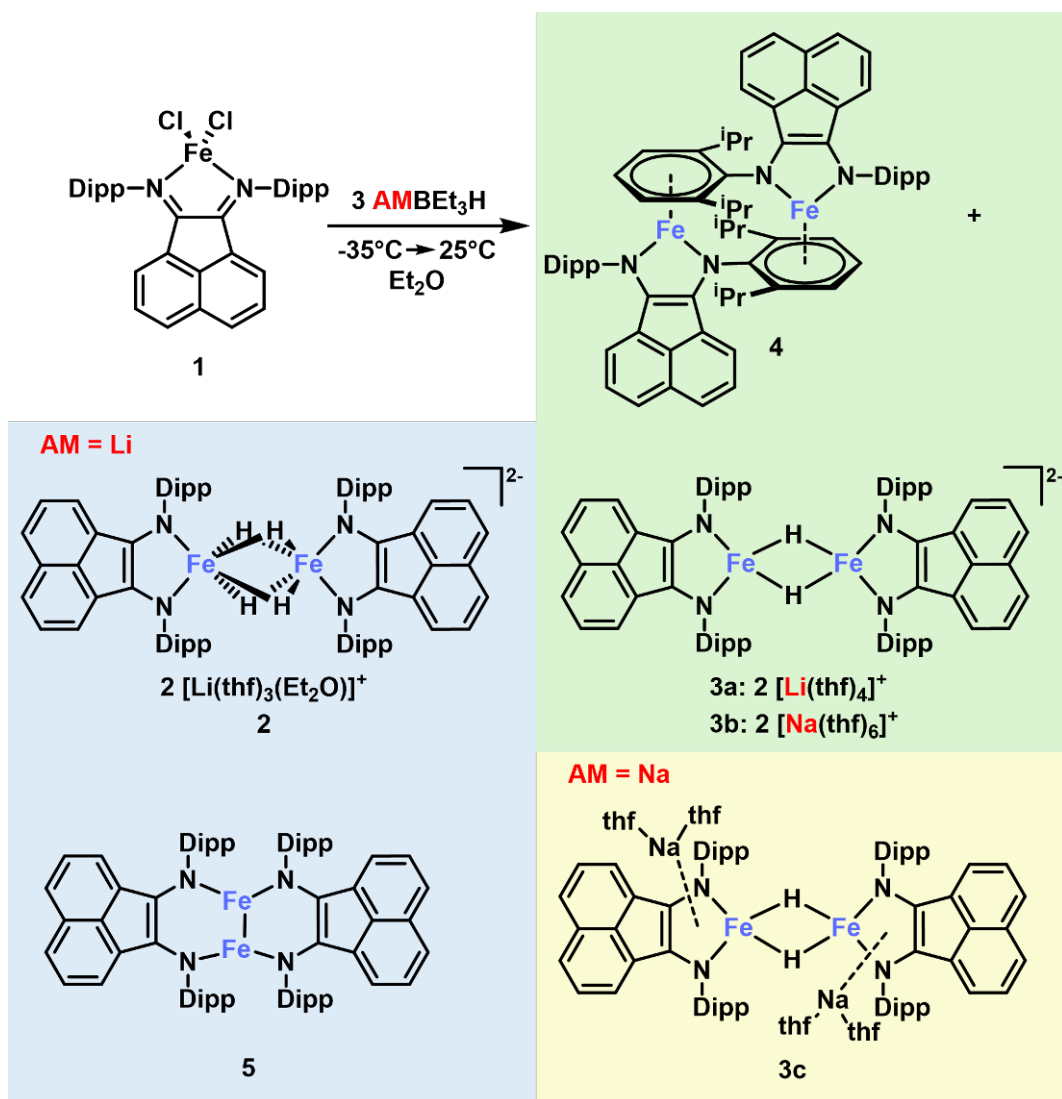


Figure 7.4: Complexes synthesized by reduction of $[\text{LFeCl}_2]$ with AMBET_3H ($\text{AM} = \text{Li}$ and Na). Complexes isolated from both reactions are highlighted green, those isolated from $\text{AM} = \text{Li}$ are highlighted blue and those isolated from $\text{AM} = \text{Na}$ are highlighted yellow.

In different fractions, isomeric $[\text{Fe}_2\text{L}_2]$ complexes **4** and **5** were isolated. While complexes **2**, **3**, and **5** form dark green block-shaped crystals and solutions in THF, complex **4** crystallizes

¹Synthesis attempts with KBET_3H resulted in oily residues, and no crystals suitable for sc-XRD could be obtained so far.

as green plates and forms red solutions in THF or toluene, similar in appearance to the complex **6**[192].

Complexes **2**[55], **3a**[55], **3b**, and **4**[55] have been analyzed by elemental analysis (EA). Unfortunately, individual measurements for complex **5** varied and no satisfactory EA results were obtained, likely due to the decomposition of the highly air-sensitive compound before measurement.

Like the ^1H -NMR spectra of **2** and **3a**,[55] also the ^1H -NMR spectrum of **3b** (Figure,D.8) shows only solvent signals from the solvated cation in the range from -120 ppm to 120 ppm which is indicative of the paramagnetic nature of this complex. **5** displays broad signals that are significantly shifted, suggesting that it is also paramagnetic. In contrast, the isomeric complex **4** shows a diamagnetic ^1H -NMR spectrum with sharp signals.[55] Additional sets of doublets for the *i*-Pr groups indicate coordination to one of each *Dipp* group. Similar to the closely related $[\text{FeL}(\text{tol})]$ **6** complex,[192] the *p*- and *m*-aryl protons of the coordinated *Dipp* are shifted upfield (5.62 ppm to 4.76 ppm, Figure,D.7).

Single crystal X-ray diffraction analysis of **2** shows four hydride ligands that bridge two **L**Fe units (Figure 7.5).[55] The molecule exhibits pseudo D_{2h} symmetry and lies across an inversion center located between the iron atoms. The $[\text{Li}(\text{thf})_3(\text{Et}_2\text{O})]^+$ cation is solvent-separated. A short Fe–Fe distance ($2.5286(7)$ Å) is observed due to the presence of four bridging hydrides and may indicate metal-metal interaction. This bond is longer than in the related complexes $[\text{Cp} \cdot \text{Fe}_2(\mu\text{-H})_4]$ ($2.202(2)$ Å)[165] and $\text{Li}_5[\text{Ph}_3\text{Fe}_2(\mu\text{H})_3]$ ($2.389(1)$ and $2.379(1)$ Å),[164] but shorter than $[(\text{NacNac})\text{Fe}(\mu\text{-H})_2]$ ($2.624(2)$ Å).[172] The Fe–H bond distances are $1.60(7)$ Å and $1.73(11)$ Å. The NC=CN bond length of BIAN is significantly shortened in comparison to the neutral ligand, indicative of a dianionic BIAN (C–N $1.383(4)$ Å and $1.374(4)$ Å; C–C $1.397(4)$ Å; Fe–N: $1.981(2)$ Å; N–Fe–N: $88.26(10)^\circ$). Similar bond lengths have been recently reported by Wolf and coworkers for a related ferate $[\text{LFe}^{\text{I}}(1,5\text{-cod})]^-$ ($1,5\text{-cod}$ = 1,5-cyclooctadiene, C–N $1.385(4)$ Å, C–C $1.388(5)$ Å, Fe–N $1.971(2)$ Å, N–Fe–N $82.1(1)^\circ$).[194]

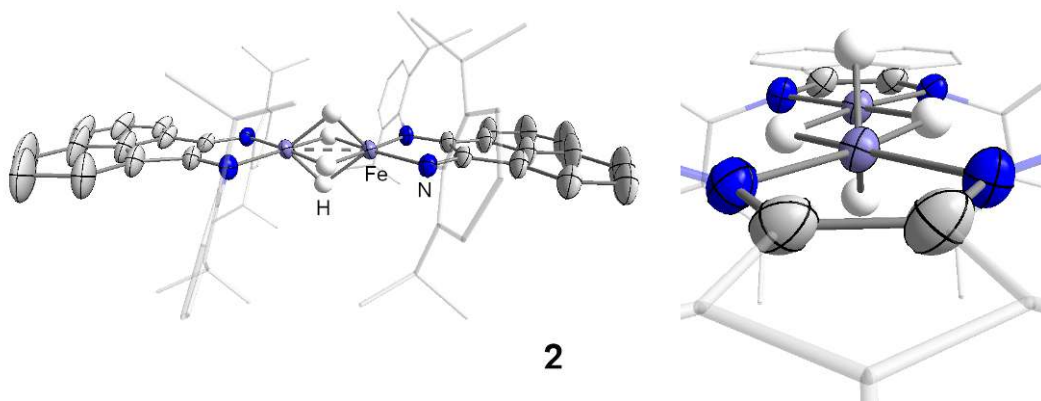


Figure 7.5: Molecular structures of **2**. Thermal ellipsoids at the 50 % probability level; minor disordered parts, non-coordinated solvents, selected H atoms, and cations are omitted for clarity. Adapted from reference [55].

Next to the main product **2**, the related dihydridoferrate **3a** was isolated in small amounts from a THF fraction. Single crystal X-ray diffraction analysis shows a very similar structure to **2**, which formally differs by the loss of H_2 . Like for **2**, bond lengths are indicative of a dianionic BIAN (C–C $1.387(3)$ Å; C–N: $1.382(2)$ and $1.393(2)$ Å; Fe–Fe: $2.5122(5)$ Å; Fe–N: $1.9863(16)$ and $1.9890(16)$ Å; Fe–H $1.84(3)$ Å; N–Fe–N $89.00(6)^\circ$).[55]

When using NaBEt_3 as the reducing agent, the analogous dihydridoferrate **3b** was isolated as

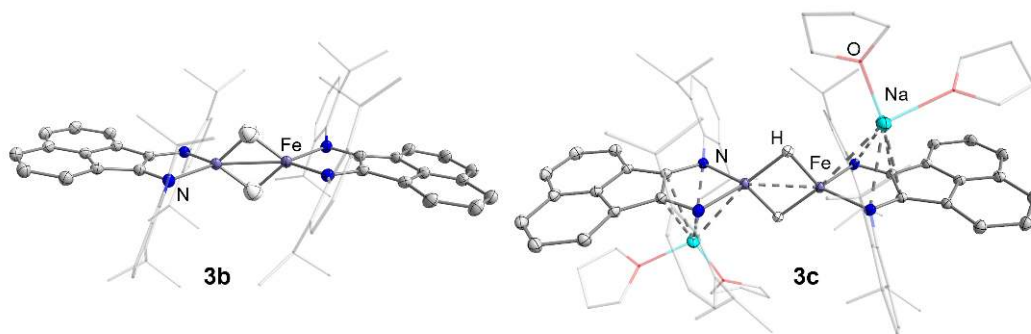


Figure 7.6: Molecular structures of **3b** and **3c**. Thermal ellipsoids at the 50 % probability level; non-coordinated solvents, selected H atoms, and cations are omitted for clarity.

the main product. So far, no sodium variant of the dihydridoferrate was isolated. The structural parameters obtained from single crystal X-ray diffraction analysis are again close to those of **2** and **3a**, and the BIAN ligand is presumably in its dianionic state (C-C 1.397(2) Å; C-N: 1.381(2) and 1.377(2) Å; Fe-Fe: 2.5360(6) Å; Fe-N: 1.9874(13) and 1.9927(14) Å; Fe-H 1.68(3) Å; N-Fe-N 87.79(6)°, Figure 7.6 left).

In contrast to the ion pair separated **3b**, the dihydridoferrate **3c** with coordinated Na⁺ cations shows better solubility in nonpolar solvents, and a few crystals were obtained from a *n*-hexane fraction. The distance between the sodium ion and the coordinated plane through the iron diimine moiety is 2.4506(9) Å. Again, bond lengths and angles are closely related to the above-described tetra- and dihydridoferrates (C-C 1.392(2) Å; C-N: 1.3926(19) and 1.3907(19) Å; Fe-Fe: 2.5660(5) Å; Fe-N: 1.9882(12) and 1.9983(12) Å; Fe-H 1.86(2) Å; N-Fe-N 86.59(5)°, Figure 7.6 right). Among the hydridoferrates described here, **3c** shows the longest Fe-Fe distance (Table 7.3).

For all reductions performed with AMBE₃H, the neutral complex **4** could be crystallized from the filtrate of the reaction mixture. In this dinuclear complex, each iron coordinates to one diisopropylphenyl moiety of the ligand of the other iron center (Figure 7.7). The NCCN bond lengths of BIAN (1.405(4) Å for C-C, 1.345(4) Å and 1.336(4) Å for C-N) are in good agreement with a related complex [LFe(*i*PrC₆H₅)] (1.405(3) Å for C-C, 1.341(3) Å and 1.343(3) Å for C-N)[55] and [LFe(tol)] (1.400(4) Å for C-C, 1.343(3) Å for C-N) as reported by Findlater and coworkers.[192, 195] The authors suggested a monoanionic BIAN strongly and antiferromagnetically coupled to a Fe^I center (*S* = 0).[195] However, Fedushkin and coworkers reported a monoanionic BIAN antiferromagnetically coupled to a high-spin Fe^{II} in [(L₂Fe)] (1.4234(18) Å for C-C, 1.3367(15) Å and 1.3393(15) Å for C-N).[196] Note that the C-C bond length of this monoanionic BIAN is much longer than in **3**. The electronic structure (*S* = 2 - 2 · $\frac{1}{2}$ = 1 ground state) is assigned by SQUID magnetometry.[196] Hence, an alternative description for **3** would involve a dianionic BIAN with a low-spin Fe^{II} (*S* = 0).[55]

The structurally unique complex **5** is isomeric to complex **13**. Both BIAN ligands coordinate both iron centers, each with one of the N atoms ($1\kappa^1 N, 2\kappa^1 N'$), forming a diferradi-cyclic moiety (Figure 7.8). Each Fe₂NC₂N cycle is almost planar with a sum of internal angles of 717.9° (N-C-C: 125.86(19)°, 125.12(18)°; C-N-Fe: 130.35(14)°, 135.76(13)°; N-Fe-Fe: 104.04(6)°, 96.78(5)°), which is only slightly smaller than the 720° of a plane hexagon.

Complex **5** has a remarkably short Fe-Fe bond length of 2.3068(8) Å. Apart from iron carbonyl clusters, only a small number of complexes with such short Fe-Fe bond distances have been reported so far. The shortest Fe-Fe bonds described are found in two iron complexes with two Fe^{I, hs} centers and two bulky guanidinato ligands (2.127(7) Å and 2.1516(5) Å).[197, 198] Like **5**, their inner Fe₂N₄C₂ or Fe₂N₆ entities are essentially planar.

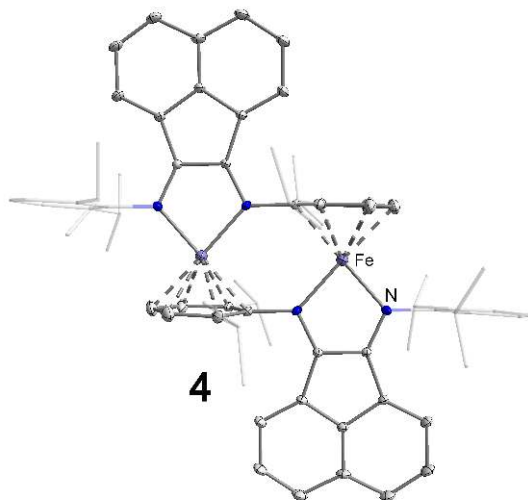


Figure 7.7: Molecular structures of **4**. Thermal ellipsoids at the 50 % probability level; selected H atoms are omitted for clarity. Reproduced from reference [55]

The NCCN bond lengths of the ligands are similar to those of the hydridoferrates (Table 7.3), which is characteristic of a dianionic BIAN ligand. For tri- and tetragonal lantern-type iron(II) dimers with three amidinato and four triazenido ligands, Fe–Fe distances of 2.198(2) Å and 2.167(8) Å have been reported.[199, 200] Other examples involve large multi-dentate ligand systems[201–204] such as the mixed-valence $[\text{Fe}^{\text{III}}\text{Fe}^{\text{I}}\text{Cl}(\text{py}_3\text{tren})]$ [201] (with py_3tren being the triply deprotonated form of *N,N,N*-tris(2-(2-pyridylamino)ethyl)amine, Fe–Fe: 2.2867(5) Å) reported by Lu and coworkers and $[\text{Fe}^{\text{II,hs}}_2(\text{NDI})(\text{C}_6\text{H}_6)]$ with a rigid naphthyridine–diimine (NDI) ligand (Fe–Fe: 2.287(1) Å).[203]

Table 7.3: Overview of the characteristic bond lengths in the described dinuclear iron complexes in this work.

Complex	d(Fe–Fe)/Å	d(Fe–N)/Å	d(C–C)/Å	d(C–N)/Å
2	2.5286(7)	1.981(2)	1.397(4)	1.383(4), 1.374(4)
3a	2.5122(5)	1.9863(16), 1.9890(16)	1.387(3)	1.382(2), 1.393(2)
3b	2.5360(6)	1.9874(13), 1.9927(14)	1.397(2)	1.381(2), 1.377(2)
3c	2.5660(5)	1.9882(12), 1.9983(12)	1.392(2)	1.3926(19), 1.3907(19)
4	-	1.928(3), 1.909(3)	1.405(4)	1.345(4), 1.336(4)
5	2.3068(8)	1.9019(17), 1.8895(13)	1.391(3)	1.380(3), 1.377(3)

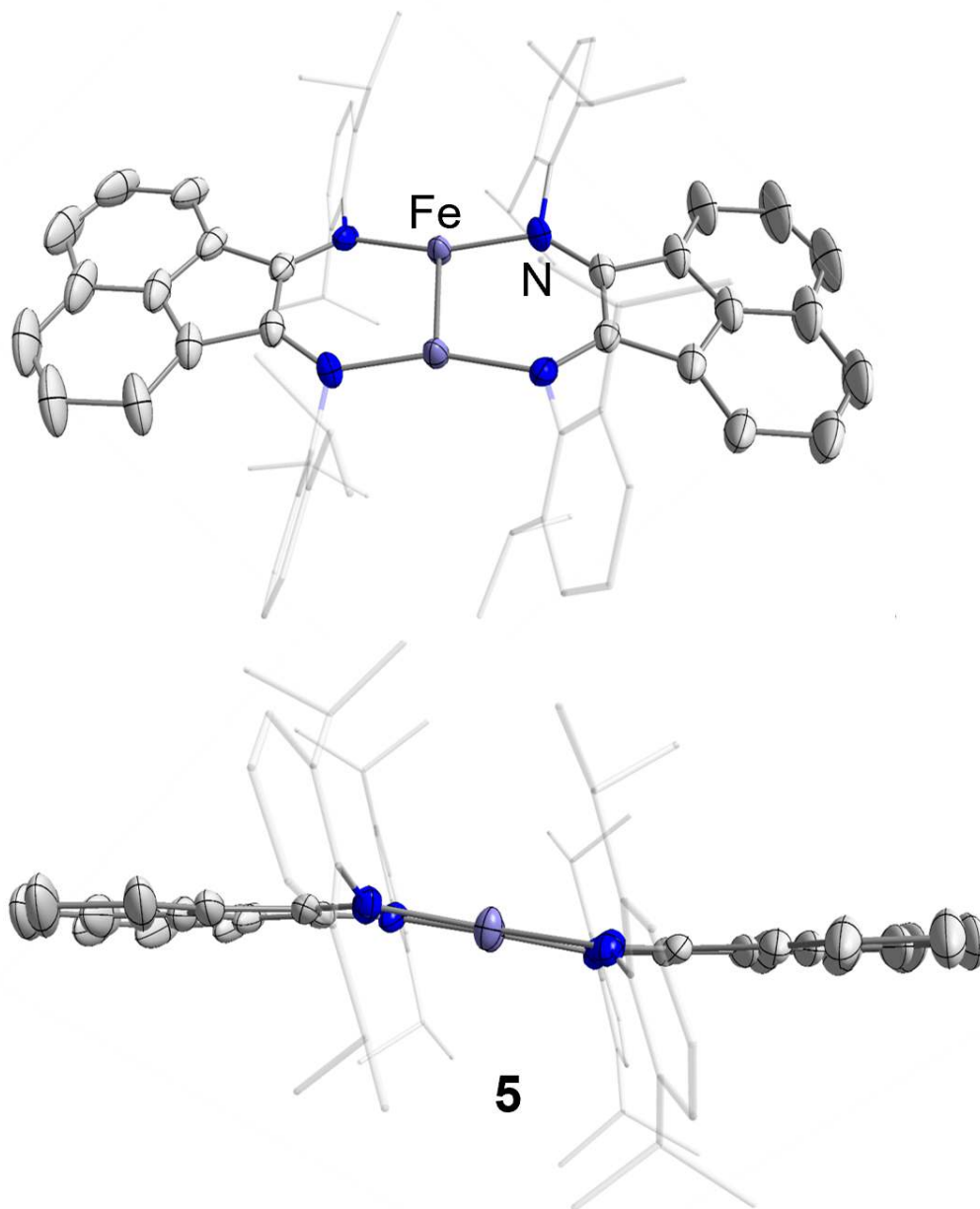


Figure 7.8: Molecular structure of **5** (top: top view, bottom: side view). Thermal ellipsoids at the 50 % probability level; selected H atoms are omitted for clarity.

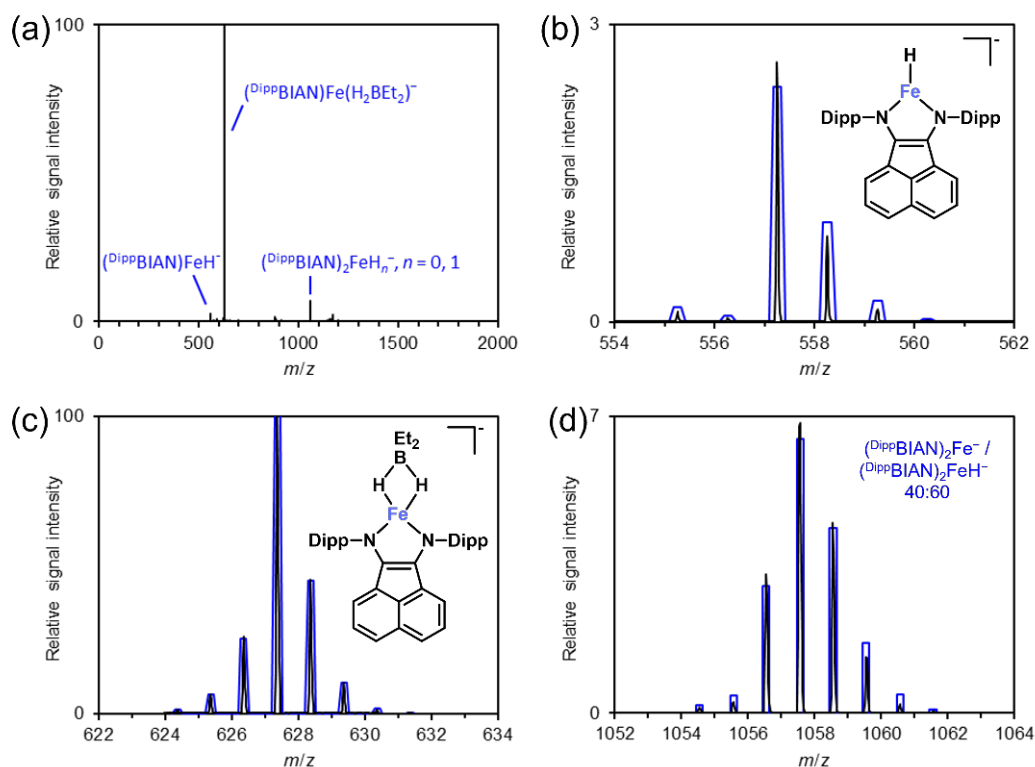


Figure 7.9: (a): Negative-ion mode ESI mass spectrum of a solution of **2** ($c \approx 6$ mM) in THF. (b): Section from the negative-ion mode ESI mass spectrum (black) together with the theoretical isotope pattern of $[\text{LFeH}]^-$ (blue). (c): Section from the negative-ion mode ESI mass spectrum (black) together with the theoretical isotope pattern of $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ (blue). (d): Section from the negative-ion mode ESI mass spectrum (black) together with the theoretical isotope pattern for a 40:60 mixture of $[\text{L}_2\text{Fe}]^-$ and $[\text{L}_2\text{FeH}]^-$ (blue).

Utilizing negative-ion mode ESI-mass spectrometric analysis under mild conditions, a solution of **2** in THF afforded an ion of the composition of $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ as the main species, $[\text{LFeH}]^-$ and $[\text{L}_2\text{FeH}_n]^-$ ($n = 0$ and 1) (Figure 7.9, Table D.1). The presence of $[\text{LFeH}]^-$ is quite expected as it could supposedly form from **2** by dehydrogenation and dissociation. The occurrence of these processes despite the use of mild ESI conditions suggests that **2** readily loses H_2 and easily dissociates into its mono-nuclear components. Upon gas-phase fragmentation, $[\text{LFeH}]^-$ underwent extensive dehydrogenation with the loss of up to 4 H_2 molecules (Figure D.17). Although coordinatively more saturated than $[\text{LFeH}]^-$, its $[\text{L}_2\text{FeH}_n]^-$ counterparts also underwent dehydrogenation, together with the elimination of one *Dipp*BIAN ligand (Figure D.19). Similar behavior has previously been observed for the related dinuclear hydridocobaltate complex LCo_2H_3^- .^[52] Most likely, this high propensity toward dehydrogenation reflects not just the unsaturated character of $[\text{LFeH}]^-$, but also its intrinsically high reactivity for such reactions. Under catalytic conditions, the principle of microscopic reversibility then implies that $[\text{LFeH}]^-$ and related complexes, such as **2**, are also active in mediating hydrogenation reactions. During measurement the intensity of the $[\text{LFeH}]^-$ peak decreased, indicating the decomposition of **2** inside the syringe with the measured solution. $[\text{LFeH}]^-$ and $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ were also observed for the reaction mixture of LFeCl_2 and 3 LiBEt_3H in THF (Figure D.20 and Table D.2). Intuitively, $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ is closely related to **2** and can be assumed to form from the dinuclear complex by dehydrogenation, dissociation, and the exchange of the hydride ligand for H_2BEt_2^- , the latter apparently remaining from the synthesis. Gas-phase fragmentation of mass-selected $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ resulted in the selective and facile loss of HBEt_2 and the formation of the hydridoferrate $[\text{LFeH}]^-$ (Figure D.18).

To investigate the electronic structure, DFT calculations were employed. Given the involvement of the redox-active ligand, which is most likely either an anion ($S = \frac{1}{2}$) or a closed-shell dianion, several oxidation and spin states are plausible (Figure D.25). Optimization of the molecular structure encountered convergence issues, as well as the dissociation of the complex **2** for some spin states. Only using optimized molecular structures for some spin states would result in a strong preference for these optimized spin states, as has been discussed in chapter 6.5. Initial attempts revealed that spin state energetics are highly dependent on the choice of functional (Figure D.26). A benchmark study on iron complexes with redox-active α -diimine ligands, discussed in Chapter 5, shows that hybrid functionals with 15 % exact exchange are suitable for spin state energetics of iron complexes with this type of ligand[71]. For the aforementioned reasons, crystal structures were used for the calculations, and the B3LYP* functional, which has generally been shown to be suitable for spin state energetics of iron complexes, was chosen. For **2**, the septet (two ferromagnetically coupled $\text{Fe}^{\text{III, is}}$) is lowest in energy, while for **3**, the nonet (two ferromagnetically coupled $\text{Fe}^{\text{II, hs}}$) is lowest in energy (Figure 7.10). The plot of the spin density shows nearly no spin density localized on the ligand, which is consistent with a closed-shell dianion as suggested by the crystal structure. Using the OPBE functional yielded similar results for **2** (Figure D.27). Analyzing the number of unpaired electrons on each iron center shows that **2** has two inequivalent iron centers with 2.6 and 3.4 unpaired electrons each, which aligns with the assumption of $\text{Fe}^{\text{III, is}}$ (Figure 7.12). However, the calculated number of unpaired electrons is usually smaller than the theoretical one. With some spin density of opposite sign localized on two of the bridging hydrides, the complex could also be described as $\text{Fe}^{\text{II, hs}}$ antiferromagnetically coupled to two H^\bullet (Figure 7.10). A direct comparison of **3** and **2** by adding the energy of H_2 to **3** resulted in an unreasonably large energy difference (Figure D.28), likely due to the insufficient description of entropic effects.

For complex **4**, as expected from the diamagnetic NMR spectrum, the singlet state ($\text{Fe}^{\text{II, ls}}$) is much preferred compared to the quintet and nonet (two $\text{Fe}^{\text{II, hs}}$) (Figure 7.11). In contrast, spin state energies of isomer **5** are similar. In agreement with the paramagnetic NMR spectrum, the nonet is predicted to be in the ground state. Unlike the singlet of **4**, where no unpaired electrons are localized on the iron center, **5** has an open-shell singlet with 3.4 unpaired electrons of opposite spins localized on each iron center, consistent with two antiferromagnetically coupled $\text{Fe}^{\text{II, hs}}$. As **4** and **5** are isomers a direct comparison of their energy levels is possible. Both complexes are energetically close with **5** being favored by 6.9 kJ/mol. This is in agreement with isolating both complexes from the same reaction.

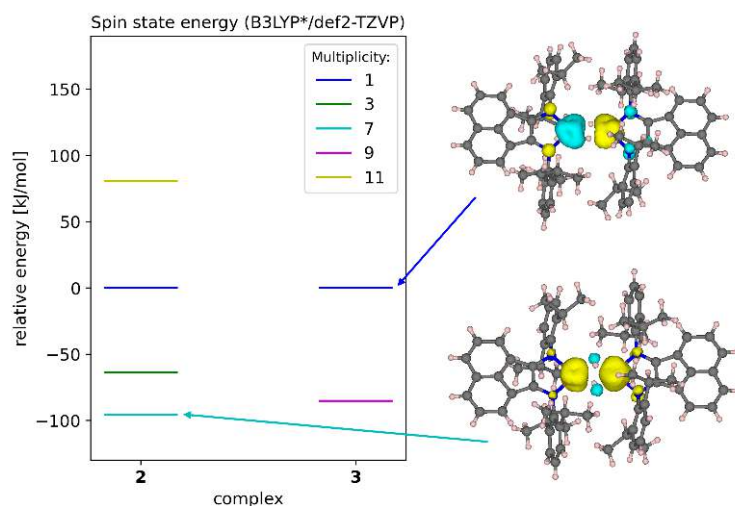


Figure 7.10: Relative energies of spin states calculated with B3LYP*/def2-TZVP with PCM(THF) on crystal structures. Spin densities for selected spin states are calculated with an iso surface value of $0.005 e a_0^{-3}$.

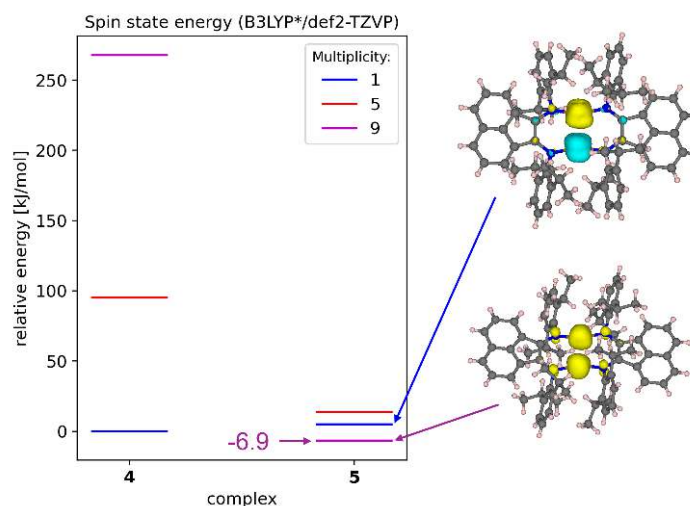


Figure 7.11: Relative energies of spin states calculated with B3LYP*/def2-TZVP with PCM(THF) on crystal structures. Spin densities for selected spin states are calculated with an iso surface value of $0.005\ e a_0^{-3}$.

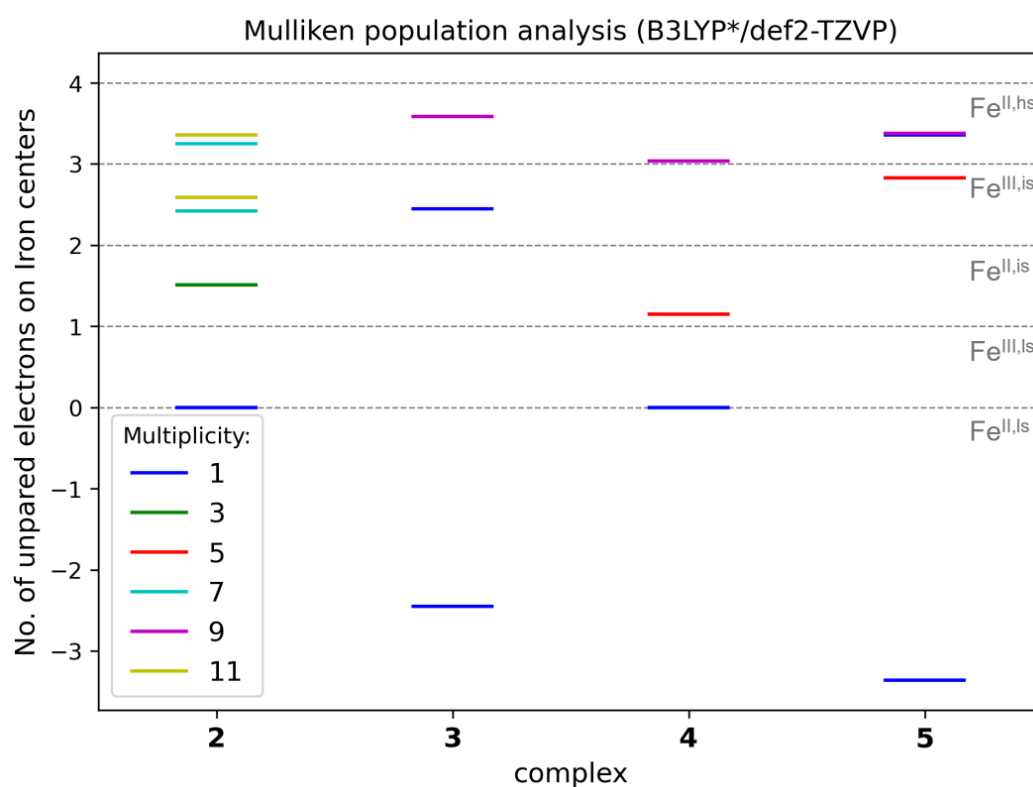


Figure 7.12: Number of unpaired electrons on iron atoms calculated by Mulliken population analysis for different spin states (B3LYP*/def2-TZVP with PCM(THF)). The grey dashed lines indicate ideal values for different oxidation and spin states of the iron centers.

As mentioned before, Mössbauer spectroscopy is crucial for the analysis of the oxidation and spin state of iron centers. The ^{57}Fe Mössbauer spectra of **2** and **3** show the same isomer shift (0.53 mm s^{-1}) and very large quadrupole splittings of 3.98 mm s^{-1} and 4.21 mm s^{-1} , respectively (Figure 7.13 top). These values do not fit the smaller quadrupole splitting observed in the literature for $\text{Fe}^{\text{II,hs}}$ complexes (see Table 7.1)[179, 181]. A better fit might be an Fe^{II} or Fe^{III} intermediate spin state, as described for the square planar coordinated Fe^{II} reported by Smith et al. ($\delta=0.35\text{ mm s}^{-1}$, $\Delta E_{\text{Q}}=0.35\text{ mm s}^{-1}$). To gain more insight, Mössbauer spectra were calculated with DFT according to Neese and coworkers[205, 206] (Figure 7.13 bottom). The calculated isomer shifts do not depend much on the spin state and are similar for complexes **2** and **3**. They are slightly smaller than the experimental values, which might result from the use of unoptimized molecular structures.[207]

The calculation for **2** with $\text{Fe}(\text{III})$ intermediate spin (septet), predicted to be the ground state by DFT, results in a very small QS of 0.50 mm s^{-1} and therefore constituting the furthest deviation from the experimental data. In contrast, the calculation of $\text{Fe}^{\text{II,hs}}$ (nonet) for **3** results in a much larger QS of 3.63 mm s^{-1} , thus describing the experiment much better. A possible explanation for these findings is that **2** loses H_2 during drying. The main species in the Mössbauer spectrum is then **3** with $\text{Fe}^{\text{II,hs}}$. The impurities observed in the spectrum of **2**, which fit the calculated QS of $\text{Fe}^{\text{II,hs}}$ much better, actually result from the tetrahydride **2**. Although $\text{Fe}^{\text{II,hs}}$ is predicted as the ground state of complex **3**, such a large quadrupole splitting is atypical for $\text{Fe}^{\text{II,hs}}$. Nevertheless, this scenario is also supported by solid-state magnetic measurements, which can be fitted to two $S = 2$ iron centers (Figures 7.14).

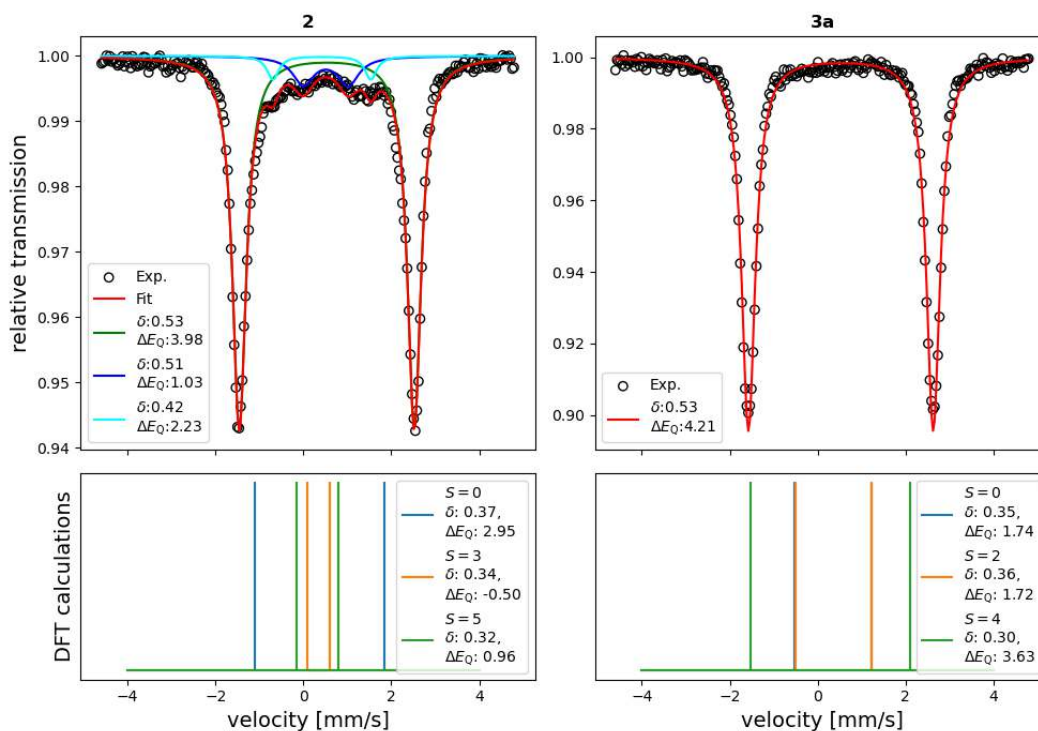


Figure 7.13: Mössbauer spectra of **2** and **3a** measured at 80 K (top) mössbauer parameters calculated with TPSSH and CP(PPP) basis set on iron atoms and def2-TZVP on all other atoms (bottom).

The Mössbauer spectra for the isomeric complexes **4** and **5** are clearly distinguishable (Figure 7.15 top). Neither of these compounds appears to be the impurity observed in the Mössbauer spectrum of **2**. The parameters of **4** are in good agreement with those of the related toluene complex, which has an isomer shift (δ) of 0.45 mm s^{-1} and a quadrupole splitting (ΔE_{Q}) of 0.41 mm s^{-1} . [192] The authors of the related toluene complex are hesitant to assign an

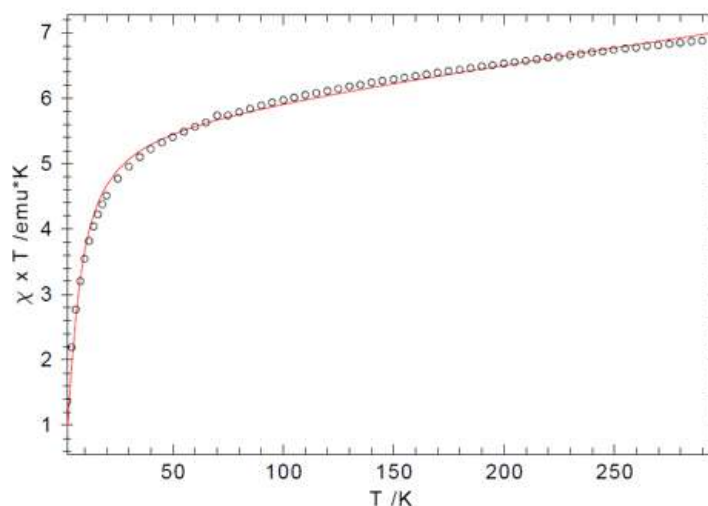


Figure 7.14: Temperature-dependence of the product $\chi_M T$ of **2**. The red curve is the fit for $2 \cdot S = 2$.

oxidation state due to missing measurements. Similarly, for **4**, magnetic measurements cannot be provided due to iron nanoparticle impurities in the measured sample. However, the isomeric shift of **4** is well fitted by the DFT calculations.

The Mössbauer spectrum of **5** reveals two observed species in a 1:2 ratio (green fit vs. blue fit in Figure 7.15 top right). The calculated spectra of the quintet ($\text{Fe}^{\text{II, is}}$) fit the green curve better. Nevertheless, it is unclear which species is the impurity, and there is a lack of comparable precedence in the literature due to the unique coordination sphere. Additionally, no further measurements of the magnetic moment are available and a definitive statement on the oxidation and spin state cannot yet be made.

The electrochemical properties of the complexes were investigated using cyclic voltammetry (CV) measurements. For the formally more reduced complex **2**, oxidative processes were examined. The CV of tetrahydridoferrate **2** shows a reversible oxidation at $E_0 = -2.53$ V (THF/ $[\text{nBu}_4\text{N}]\text{PF}_6$, $\Delta E_p = 84$ mV) and another likely reversible oxidation at $E_0 = -2.19$ V ($\Delta E_p = 123$ mV) vs. Fc/Fc^+ .^[55]

For the formally less reduced **4**, reductive processes were investigated. Two very close reversible reductions were observed, with a shoulder formation at -2.48 V and -2.60 V (Figure 7.16 left). The differences between the two peak potentials (ΔE_p) for **4** are 55 mV and 91 mV, which are close to the expected value of 57 mV for a reversible $1 e^-$ process. The small deviations likely result from the overlap of both reductions. These findings strongly support that **4** is also a dimer in solution. The related toluene complex **6** exhibits a reversible reduction at -2.56 V (Figure 7.16 right), which is between the reduction potentials of **4**, thereby highlighting the electrochemical similarity of both compounds. The similarity of the reduction potentials of **4** and **6** with the oxidation potential of **3** supports the assumption that a $1 e^-$ reduction of **4** or **6** under a hydrogen atmosphere could lead to the formation of **2**.

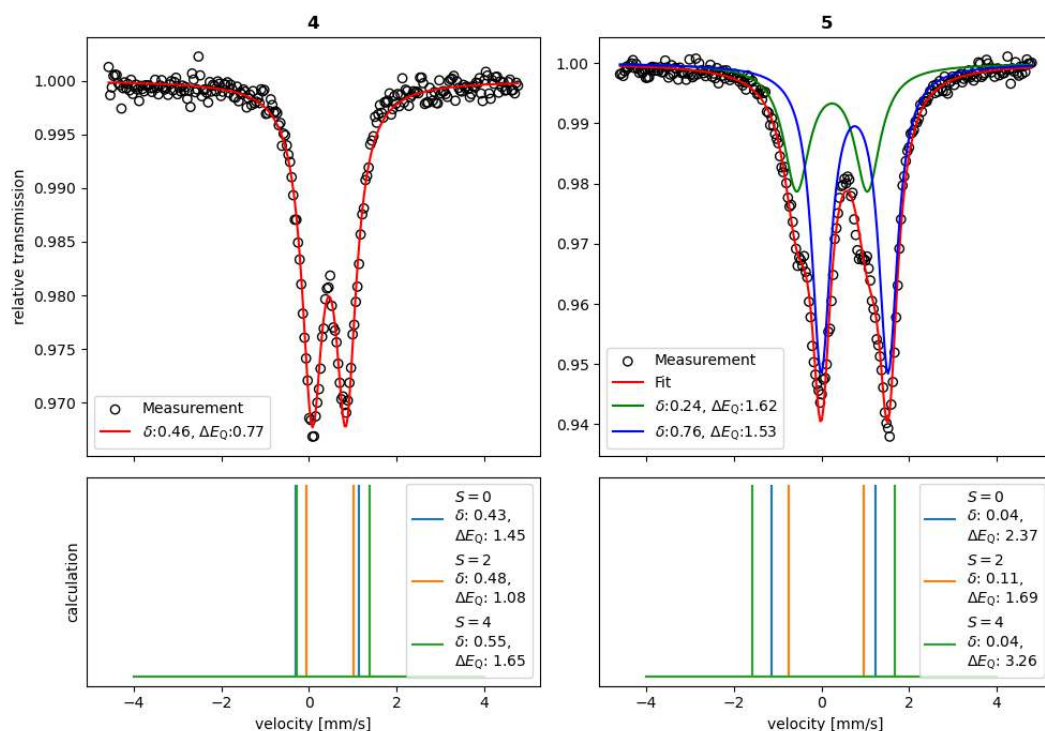


Figure 7.15: Mössbauer spectra of **4** and **5** measured at 80 K (top). Mössbauer parameters calculated with TPSSH and CP(PPP) basis set on iron atoms and def2-TZVP on all other atoms (bottom)

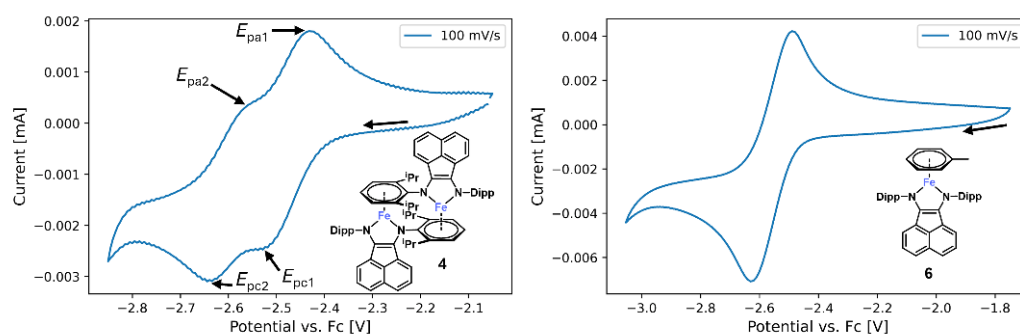


Figure 7.16: Cyclic voltammograms in THF/ $[\text{nBu}_4\text{N}]\text{PF}_6$ at scan rate 100 mV s^{-1} . left: complex **4** $E_{0,1} = -2.48 \text{ V}$, $E_{\text{pc}1} = -2.53 \text{ V}$, $E_{\text{pa}1} = -2.44 \text{ V}$, $\Delta E_{\text{p}1} = 91 \text{ mV}$, $E_{0,1} = -2.60 \text{ V}$, $E_{\text{pc}2} = -2.62 \text{ V}$, $E_{\text{pa}2} = -2.57 \text{ V}$, $\Delta E_{\text{p}2} = 55 \text{ mV}$; right: complex **6** $E_{\text{pc}} = -2.62 \text{ V}$, $E_{\text{pa}} = -2.49 \text{ V}$, $\Delta E_{\text{p}} = 74 \text{ mV}$.

7.4 Reactivity of Ironhydrides

To gain insight into the possible applications of hydridoferrates **2**[55] and **3b**, a reactivity assessment was conducted. The reaction with acetophenone (Figure 7.17 (a)) yielded 1-phenylethanol and hydrobenzoin, with hydrobenzoin present in slightly larger amounts for both hydrides. The formation of 1-phenylethanol indicates hydride transfer, while hydrobenzoin formation suggests a single-electron transfer (SET) process. This competition between hydride transfer and SET was also observed during catalytic hydrogenation (see Section 7.2). Similarly, the reaction with benzyl bromide produced both hydride transfer (toluene) and SET (1,2-diphenylethane) products (Figure 7.17 (b)). The hydrodebromination of 4-bromotoluene was observed with both complexes. Additionally, under mild conditions (room temperature and 1 atm pressure), hydride transfer to CO₂ was successfully demonstrated.

The isomerization of *Z*-stilbene to *E*-stilbene indicates olefin coordination to the iron complexes (Figure 7.18 (a))[55]. For the terminal alkene allylbenzene, isomerization products were only observed as side products, while the main product was propylbenzene (Figure 7.18 (b)). The workup was performed in D₂O, and since no deuteration was found in the product, it is likely that of the added hydrogen atoms originate from the hydridoferrates. The reactivity of **2** towards phenylacetylene was poor, resulting in only 10 % cyclotrimerization and no hydrogenation.[55] Due to the acidic hydrogen in phenylacetylene, hydrogen gas formation could be a potential reason for this. However, experiments targeting the detection of formed hydrogen gas showed no formation of hydrogen during the reaction of **2** or **3** with phenylacetylene (Table 7.4 entry 5).

These experiments were conducted by placing a smaller vial containing a solution of styrene and a Pd/C catalyst inside a vial with the hydridoferrate solution. After the addition of the reactant to the outer vial with the hydridoferrate, the vial was quickly closed. In this setup, any formed hydrogen could diffuse through the gas phase to the inner vial, where it would be consumed to hydrogenate styrene. The formed ethylbenzene could then be quantified by GC-FID. A blank sample with an empty outer vial resulted in no conversion of styrene (Table 7.4 entry 1), proving that the formed hydrogen resulted from the reaction and not from H₂ contamination in the glovebox atmosphere.

Table 7.4: Hydrogen evolution reactions with hydridoferrates.

Entry	[Fe]	solvent	additives	Yield PhEt
1	blank	-	-	0 μmol
2	0.5 mL "fresh" solution of 2	Et ₂ O	-	16 μmol
3	0.5 mL "fresh" solution of 2	Et ₂ O	anthracene	24 μmol
4	65 μmol 2	THF	anthracene	8.5 μmol
5	6 μmol 2 or 3b	THF	phenylacetylene	0 μmol

With this setup, a solution of freshly synthesized **2** was also investigated. In this context, a fresh solution refers to the extract obtained in the synthesis, which would usually be cooled for crystallization (see Section 8.2.7). It was suspected that the hydrides easily lose H₂ upon drying. Therefore, amounts of the fresh solution of the hydride in ether were added to the outer vial before quickly sealing it. The ether would then evaporate and dissolve in the THF solution in the inner vial. During this process, a color change from the green solution to a brown residue was observed in the outer vial, and hydrogen formation was verified (Table 7.4

entry 2). A higher volume of hydrogen gas evolution could be induced by adding anthracene to the outer vial (Table 7.4 entry 3). Hydrogen formation was also observed using the prior isolated complex **2** without the drying effect (Table 7.4 entry 4). The addition of anthracene to a solution of **2** resulted in a rapid color change from green to blue. Thus suggesting the formation of an anionic anthracene complex, $[\text{LFe}(\text{anthracene})]^-$.

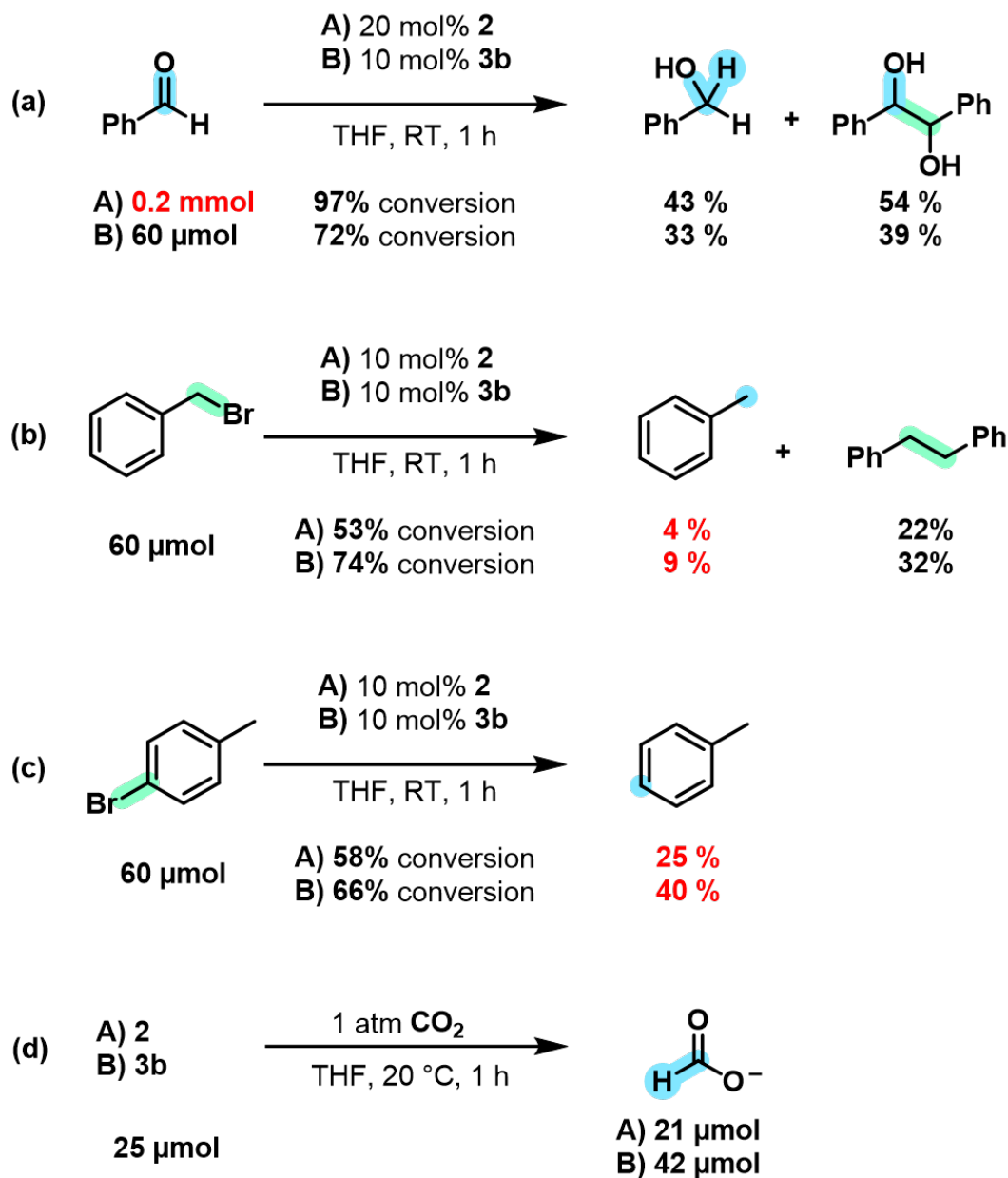


Figure 7.17: Reactivity of hydridoferrates **2** and **3a** towards electrophiles. Yields are determined by quantitative GC-FID vs. *n*-pentadecane after aqueous workup. Due to its volatility, the yields for toluene might be higher than the actual detected amount. The yield of formate has been determined by quantitative ¹H-NMR vs. internal DMF standard.

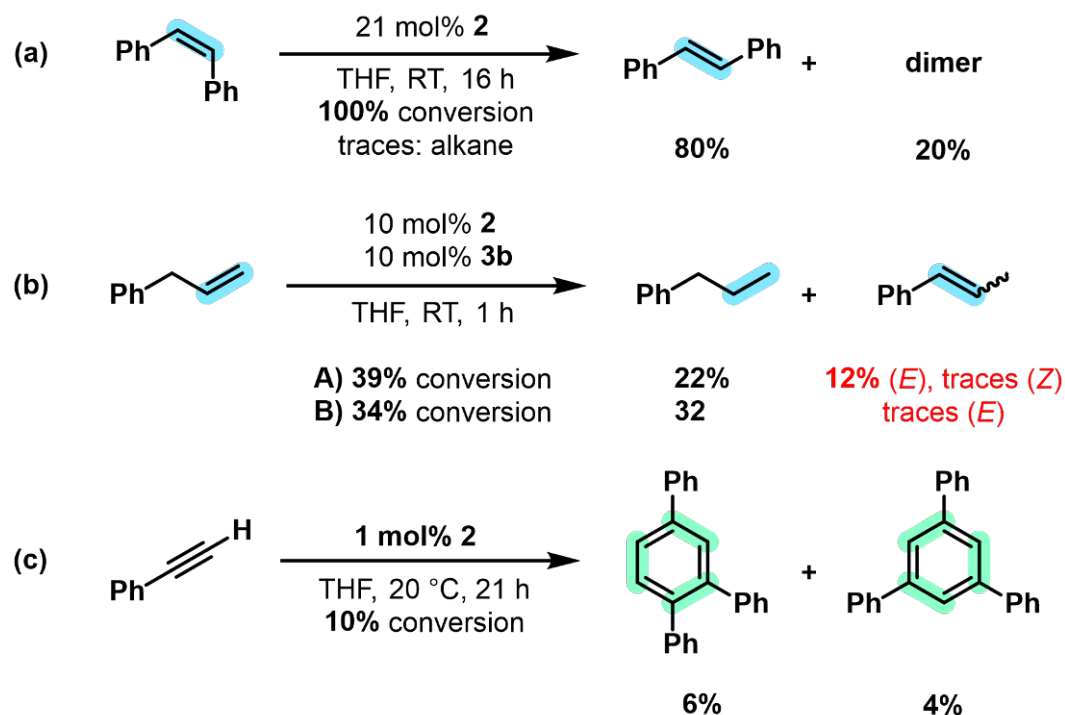


Figure 7.18: Reactivity of hydridoferrates **2** and **3a** towards alkenes and alkynes. Yields are determined by quantitative GC-FID vs. *n*-pentadecane after aqueous workup. Reactivity of hydridoferrates **2** and **3a** with alkenes and alkynes. Yields are determined by quantitative GC-FID vs. *n*-pentadecane. Yields shown in red are quantified by GC-MS. Although GC-MS quantification is less accurate, it provides better peak separation compared to GC-FID in these cases.

7.5 Conclusions

The bis(imino)acenaphthene ligand has demonstrated its ability to stabilize highly reduced iron complexes by maintaining the iron centers in stable oxidation states through the uptake of two electrons. Consequently, several rare hydrogen-bridged dinuclear iron complexes, including the tetrahydridoferrate **2** and the dihydridoferrates **3a**, **3b**, and **3c**, have been successfully synthesized and characterized.

Investigations into the electronic structure of these complexes involved Mössbauer spectroscopy, which revealed unusually large quadrupole splittings for all hydridoferrate complexes, along with SQUID magnetometry and DFT calculations. The most fitting description of the iron centers is $\text{Fe}^{\text{III,ls}}$ for the tetrahydridoferrate **2** and $\text{Fe}^{\text{II,hs}}$ for the dihydridoferrates.

Reactivity studies indicated that **2** easily loses H_2 and suggesting an equilibrium between the tetrahydridoferrate **2** and the dihydridoferrate, facilitated by the elimination or addition of H_2 . This dynamic system is further supported by the presence of the complex **4**, which is related to **2** through a reductive process, as demonstrated by cyclic voltammetric measurements.

Additionally, the unique complex **5**, an isomer of **4**, has been isolated. As supported by DFT calculations, both complexes are energetically similar. The electronic structure of the diamagnetic **4** is closely related to the arene complex $[\text{LFe}(\text{tol})]$ (**6**), likely with $\text{Fe}^{\text{II,ls}}$ centers, while **5** is paramagnetic with $\text{Fe}^{\text{II,hs}}$ centers.

The system $[\text{LFeCl}_2]$ **1**/3 LiBEt_3H , as well as the dihydridoferrates, have proven to be active precatalysts for the hydrogenation of challenging alkenes even under mild conditions. Mechanistic experiments and reactivity studies have demonstrated that the presented hydridoferrates can undergo both single-electron transfer and hydride transfer reactions.

In summary, this work has provided valuable insights into the electronic structures and reactivity of these unique iron complexes.

8 Experimental Section

8.1 General

All experiments involving air- and moisture-sensitive compounds were performed under an atmosphere of dry argon or nitrogen using standard Schlenk and glove box techniques.

Chemicals and Solvents: Anhydrous dichloromethane, toluene and hexane were obtained from solvent purification system by M. Braun. Tetrahydrofuran and diethylether were distilled from sodium benzophenone ketyl. All anhydrous solvents and liquid alkenes were stored over molecular sieves (3 Å). Ethyl acetate and commercially available olefins were distilled under reduced pressure before use.

Elemental analyses (CHN) were performed with a vario ELIII by Fa. elementar.

IR measurements were carried out inside an nitrogen filled glove box with an Agilent Cary 630 FT-IR-spektrometer with ATR-element.

The ^1H -NMR spectra were measured with an Bruker AV3400 (400 MHz) and Bruker FourierHD (300 MHz) spectrometer.

Hydrogenation reactions were carried out in 300 mL high pressure reactors (Parr). The reactors were loaded under argon, purged with H_2 (1 min, H_2 : 99.9992 % from Linde), sealed, and the internal pressure was adjusted.

Gas chromatography with Flame Ionization Detector (GC-FID): HP6890 GC-FID with injector 7683B and Agilent 7820A, carrier gas: N_2 . Calibration was performed with analytically pure samples vs. internal standard *n*-pentadecane). **Gas chromatography with mass-selective detector (GC-MS):** Agilent 6890N Network GC-System, mass detector 5975 MS. Column: HP-5MS (30 m \times 0.25 mm \times 0.25 μm , 5 % phenylmethylsiloxane, carrier gas: H_2 . Standard heating procedure: 50 $^\circ\text{C}$ (2 min), 25 $^\circ\text{C min}^{-1}$ to 300 $^\circ\text{C}$ (5 min)

Liquid injection field desorption mass spectrometry (LIFDI-MS): The spectra were recorded by the Central Analytics Lab at the Department of Chemistry, University of Regensburg, on a LIFDI-MS from Linden connected to an AccuTOF GCX from Jeol.

ESI mass spectrometry: Sample solutions were transferred into a gas-tight syringe and fed into the ESI source of a microTOF-Q II mass spectrometer (Bruker Daltonik) at a flow rate of 8 L min $^{-1}$. The ESI source was operated at a voltage of 3500 V with nitrogen as nebulizer (10 L min $^{-1}$ flow rate) and drying gas (heated to 333K and held at 0.7 bar backing pressure). The thus produced ions with 50 m/z 3000 were then allowed to pass the instrument's quadrupole mass filter and collision cell before entering the time-of-flight (TOF) mass analyzer. In gas-phase fragmentation experiments, the ions were mass-selected in the quadrupole mass filter (with isolation widths of 6 or 8 u), accelerated to kinetic energies ELAB, and subjected to collisions with N_2 gas present in the collision cell. The residual precursor ions and resulting fragment ions were then also injected into the TOF analyzer. Ions were identified on the basis of their m/z ratio, their isotope pattern, and their fragmentation behavior. On average, measured m/z accuracies of < 10 ppm were obtained with an external calibration with a mixture of CF_3COOH and phosphazenes in $\text{H}_2\text{O}/\text{MeCN}$. Theoretical m/z ratios and isotope patterns were calculated with the DataAnalysis software (Bruker Daltonik).

Cyclovoltammetry: Electrochemical measurements were carried out under an atmosphere of nitrogen using 0.5 mM solutions of the analyte and $[\text{N}(\text{n-Bu})_4][\text{PF}_6]$ (0.1 M/0.1 mol L $^{-1}$) as supporting electrolyte in tetrahydrofuran. The redox potentials are given against the ferrocene/ferrocenium redox couple according to Gritzner and Kuta.[208] An Autolab PG-STAT101 setup was used with a platinum electrode (by Ω Metrohm 6.0301.100), a glassy carbon electrode (by Ω Metrohm 6.1241.060) and an Ag/AgCl electrode (by Ω Metrohm

6.0724.140) were used as working, counter and reference electrode, respectively.

DFT Calculations: All KS-DFT calculations were carried out with ORCA program package^{117–119}. B3LYP*[23, 61, 105] or OPBE[111, 209] xc functionals with atom-pairwise dispersion correction with the Becke-Johnson damping scheme (D3BJ)[102] were employed together with the def2-TZVPP[103] basis set and the auxiliary basis def2/J[104] for single point calculations. Mössbauer parameters were calculated according to Neese[205, 206] *at al.* using TPSSh xc functional and CP(PPP) basis function for iron atoms and def2-TZVPP basis set for all other atoms. Isomer shifts have been fitted with parameters $\alpha = -0.376$, $\beta = 4.130$ and $C = 11810$ by $(\delta = \alpha(\rho - C) + \beta)$. [206] The conductor-like Polarizable Continuum Model (C-PCM) with THF as solvent has been used for all calculations. [210] Convergence criteria were set to TightSCF.

8.2 Synthesis

8.2.1 General Procedure for Catalysed Alkene Hydrogenation

A 4 mL vial was charged with the precatalyst (7.5 μ mol, 3 mol%) in toluene (2 mL). LiEt₃BH (22.5 μ mol, 1.0 M in THF, 9 mol%) was added dropwise. After stirring for 10 min, the alkene (0.25 mmol, 1 equiv.) was dissolved in in toluene (0.5 mL) and added to the catalyst solution. After a cannula was placed in the septum of the vial, it was transferred to a high-pressure reactor. The reactor was purged with H₂ for 1 min. The pressure was then adjusted. After the reaction time the pressure was released. The reaction was quenched with saturated aqueous NH₄Cl solution (1 mL) and extracted with ethyl acetate (2 x 1 mL). The organic phases were dried over Na₂SO₄ and analyzed by quantitative GC-FID analysis against internal standard (*n*-pentadecane).

8.2.2 General Procedure for Reactivity Tests of Hydridoferrates

A 4 mL vial was charged with the hydridoferrate (6 μ mol, 10 mol%). The substrate (60 μ mol, 1 equiv.) was dissolved in THF 1 mL and added to the hydridoferrate. After stirring for 10 min at room temperature, the reaction mixture was quenched with D₂O, extracted with ethyl acetate (2 x 1 mL) and the extracts dried over Na₂SO₄. Reaction products were analyzed by GC-MS and quantitative GC-FID analysis against internal standard (*n*-pentadecane).

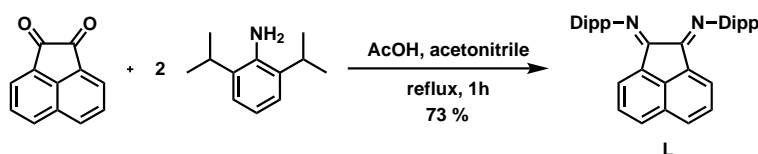
8.2.3 General Procedure for Catalysed CO₂ Hydrogenation

A 4 mL vial was charged with the precatalyst (6 μ mol, 3 mol%) in THF (2 mL). The reducing agent (18 μ mol, 9 mol%) was added dropwise. After stirring for 10 min, the base (0.20 mmol, 1 equiv.) was added to the catalyst solution. After a cannula was placed on the septum of the vial, it was transferred to a high-pressure reactor. The reactor was purged with H₂ for 1 min. Then, the pressure was adjusted to 20 bar. Afterwards, CO₂ was added until the total pressure of 40 bar was reached. After heating to 80 °C the pressure increased to 46 bar. After the reaction time, the reactor was cooled to room temperature and the pressure was released. The reaction was quenched with D₂O and the solvent removed under reduced pressure. The residue was dissolved in D₂O and DMF (15.5 μ L, 0.20 mmol, 1 equiv.) was added as internal standard. After filtration the formed formate salt was quantified by ¹H-NMR.

8.2.4 General Procedure for Reactions of CO₂ with Hydridometallates

In a 25 mL Schlenk tube the hydridometallate (25 μ mol, 1 equiv.) was dissolved in 5 mL THF. The solution was frozen in liquid nitrogen and the gas phase exchanged with CO₂ (1 bar). The solution was allowed to warm to room temperature. Afterwards it was stirred for 1 h. The solvent was removed in vacuo and the residue dissolved in 1 mL D₂O. DMF (7.7 μ L, 0.100 mmol, 4 equiv.) was added as internal standard. After filtration the formed formate salt was quantified by ¹H-NMR.¹

8.2.5 Synthesis of ^{Dipp}BIAN L



Scheme 1: Synthesis of ^{Dipp}BIAN (L).

^{Dipp}BIAN was synthesized according to literature procedure.²¹¹

Acenaphthenquinone (3.65 g, 20.0 mmol, 1.00 equiv.) were suspended in 100 mL acetonitrile and 30 mL acetic acid was added. After was stirred at 90 °C for 15 min, 2,6-Diisopropylaniline (8.5 mL, 45 mmol, 2.3 equiv.) were added dropwise. The reaction mixture was refluxed at 90 °C for 5 h during which the acenaphthenquinone was first dissolved and then an orange-yellow precipitate formed. After cooling to room temperature the suspension was filtered. After washing with *n*-pentane (70 mL) the orange-red solid was dried in vacuo.

Yield: 7.3 g (14.6 mmol, 73 %).

¹H-NMR (300 MHz, 27 °C, CDCl₃): δ [ppm] = 7.87 (d, ³*J*(HH) = 8.3 Hz, 2 H, CH_{BIAN}); 7.36 (t, ³*J*(HH) = 7.8 Hz, 2 H, CH_{BIAN}); 7.27 (m, 6 H, CH_{Dipp}); 6.64 (d, ³*J*(HH) = 7.2 Hz, 2 H, CH_{BIAN}); 3.03 (hept, ³*J*(HH) = 6.9 Hz, 6 H, CH(CH₃)₂); 1.24 (d, ³*J*(HH) = 6.9 Hz, 12 H, CH(CH₃)₂); 0.97 (d, ³*J*(HH) = 6.8 Hz, 12 H, CH(CH₃)₂).

¹³C-NMR (101 MHz, 27 °C, CDCl₃): δ [ppm] = 161.11 (C=N); 147.65 (*ipso*-C_{Dipp}); 140.94 (C_{BIAN}); 135.57 (*ortho*-C_{Dipp}); 131.26 (C_{BIAN}); 129.65 (C_{BIAN}); 129.00 (C_{BIAN}); 128.01 (C_{BIAN}); 124.43 (*ortho*-C_{Dipp}); 123.61 (*meta*-C_{Dipp}); 123.49 (C_{BIAN}); 28.76 (CH(CH₃)₂); 23.56 (CH(CH₃)₂); 23.27 (CH(CH₃)₂).

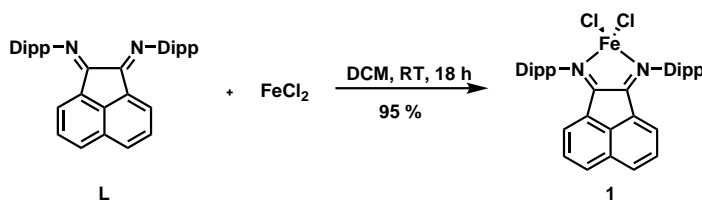
IR: (ATR) $\tilde{\nu}$ [cm⁻¹] = 3064, 2961, 1668 ($\tilde{\nu}_{\text{C=N}}$), 1640 ($\tilde{\nu}_{\text{C=N}}$), 1430, 1383, 787, 751.

8.2.6 Synthesis of (^{Dipp}BIAN)FeCl₂ 1

Synthesis was performed following an adapted procedure by M. Villa *et al.*[95]

In a Schlenk flask FeCl₂ (2.46 g, 19.4 mmol, 1.07 equiv.) and ^{Dipp}BIAN (9.08 g, 18.1 mmol, 1.00 equiv.) were suspended in 60 mL DCM. The mixture was stirred for 18 h at room temperature, during which the solution turned green. The solution was concentrated in vacuo to 80 mL. The residual FeCl₂ was separated by filtration. The solvent was removed in vacuo and the green residue was washed with toluene (3 x 10 mL) and dried in vacuo. The product was obtained as green solid.

¹Remaining paramagnetic impurities can be removed by an additional filtration or addition of EDTA followed by filtration.

Scheme 2: Synthesis of $(^{\text{Dipp}}\text{BIAN})\text{FeCl}_2$.

Yield: 10.8 g (17.2 mmol, 95 %).

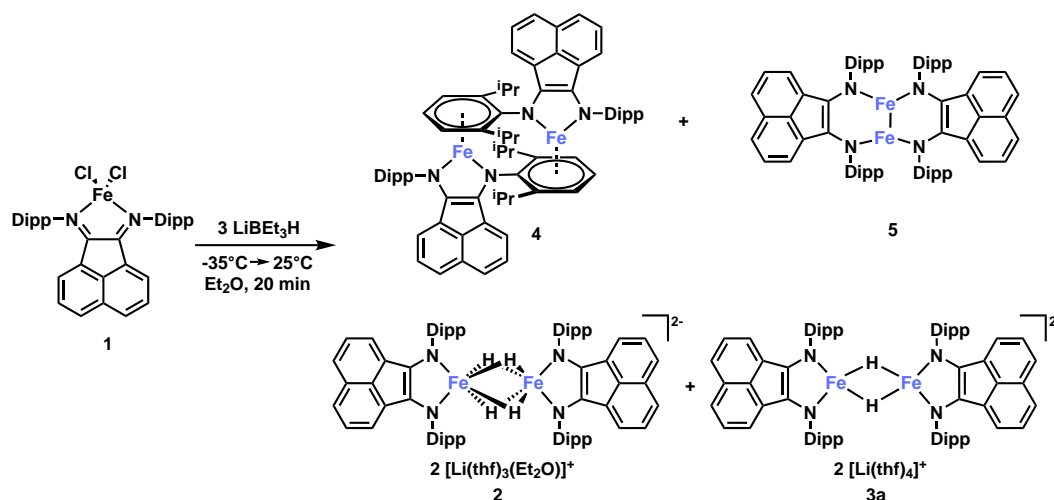
Elemental Analysis: Found (calc. for): C: 69.00 (68.91); H: 6.41 (6.43); N: 4.46 (4.46).

Magnetic Moment (Evans): $\frac{\mu_{\text{eff}}}{\mu_{\text{B}}}$ (THF- d_8 , 293 K): 6.78.

HR-MS (ESI⁺, m/z): calc.: 501.327 $[\text{M}+\text{H}]^+$, found: 501.324 $[\text{M}+\text{H}]^+$.

IR: (KBr) $\tilde{\nu}$ [cm^{-1}] = 3064, 2961, 1668, 1640, 1430, 1383, 787, 751.

8.2.7 Synthesis of 2, 5, 3a and 4

Scheme 3: Reduction of **1** with LiEt_3H .

Inside a Schlenk flask $(^{\text{Dipp}}\text{BIAN})\text{FeCl}_2$ (512 mg, 0.80 mmol, 1.0 equiv.) was suspended in 20 mL diethylether. The reaction vessel was closed with an unused septum and cooled to -35°C . A solution of LiEt_3BH (2.3 mmol, 1.05 M in THF, 3.0 equiv.) was put in a syringe closed with rubber stopper and also cooled to -35°C . The LiEt_3BH solution was added dropwise to the suspension of $(^{\text{Dipp}}\text{BIAN})\text{FeCl}_2$. During the addition a color change from pale green to a dark green solution was observed. The reaction mixture was allowed to warm to room temperature and stirred for 35 min. The product was precipitated by the addition of 20 mL *n*-hexane and filtered with a frit (P3). The filtrate **V** had a maroon color and was stored at room temperature for three days. After washing with *n*-hexane (4 mL) and diethylether (6 mL) the residue was extracted with diethylether (10 mL) and filtrated through the P3 frit (fraction **A**). Afterwards the residue was extracted with 6 mL THF and the filtrate was collected as fraction **B**. Fraction **A** and **B** had a dark green color and were stored at -35°C for crystallization. After 3 days green, plate shaped crystals of **4** had formed from fraction **V**. The tetrahydridoferrate **2** crystallized as dark green blocks from fraction **A** and the dihydridoferrate **3a** crystallized as dark green blocks from fraction **B**. The mother liquor of **B** was decanted and stored at -35°C for at least a three months. A few crystals of **5** were obtained as dark blocks from that fraction. After decanting the mother liquor and washing with hexane (2 x 2 mL) all crystals were dried via exposure to

the glovebox atmosphere.

Yield 5: 20 mg (0.018 mmol, 4.4 %).

$^1\text{H-NMR}$ (400 MHz, 300 K, THF- d_8): δ [ppm] = 47.23 (bs); 46.71 (bs); 14.10 (bs); 11.76 (bs); 7.14 (bs); 5.29 (bs); 2.86 (bs); 2.35 (bs); -0.94 (bs); -3.34 (bs); -3.93 (bs); -10.92 (bs).

IR: (ATR) $\tilde{\nu}$ [cm^{-1}] = 2983, 2954, 2863, 1584, 1489, 1458, 1423, 1379, 1308, 1252, 1206, 1175, 1100, 1035, 918, 885, 867, 814, 798, 758, 677, 663.

Yield 4: 26 mg (0.023 mmol, 6 %).

$^1\text{H-NMR}$ (300 MHz, 300 K, THF- d_8): δ [ppm] = 7.60–7.43 (m, 6 H); 7.28–7.15 (m, 6 H); 6.91 (dd, $^3J(\text{HH}) = 8.3 \text{ Hz}$, $^3J(\text{HH}) = 7.0 \text{ Hz}$, 1 H); 6.77 (dd, $^3J(\text{HH}) = 8.3 \text{ Hz}$, $^3J(\text{HH}) = 7.0 \text{ Hz}$, 2 H); 6.61 (dd, $^3J(\text{HH}) = 8.2 \text{ Hz}$, $^3J(\text{HH}) = 7.1 \text{ Hz}$, 2 H); 6.15 (d, $^3J(\text{HH}) = 6.4 \text{ Hz}$, 4 H); 6.07 (d, $^3J(\text{HH}) = 6.9 \text{ Hz}$, 1 H); 5.62 (d, $^3J(\text{HH}) = 7.0 \text{ Hz}$, 2 H); 5.43 (d, $^3J(\text{HH}) = 7.1 \text{ Hz}$, 2 H); 4.76 (t, $^3J(\text{HH}) = 6.3 \text{ Hz}$, 2 H); 4.40 (sept, $^3J(\text{HH}) = 6.8 \text{ Hz}$, 2 H); 4.26 (sept, $^3J(\text{HH}) = 6.6 \text{ Hz}$, 2 H); 1.87 (d, $^3J(\text{HH}) = 6.5 \text{ Hz}$, 12 H); 1.22 (d, $^3J(\text{HH}) = 6.9 \text{ Hz}$, 6 H); 1.06 (d, $^3J(\text{HH}) = 6.8 \text{ Hz}$, 6 H); 0.82 (d, $^3J(\text{HH}) = 6.8 \text{ Hz}$, 12 H); 0.64 (d, $^3J(\text{HH}) = 6.6 \text{ Hz}$, 12 H).

Yield 3: 67 mg (0.040 mmol, 10 %).

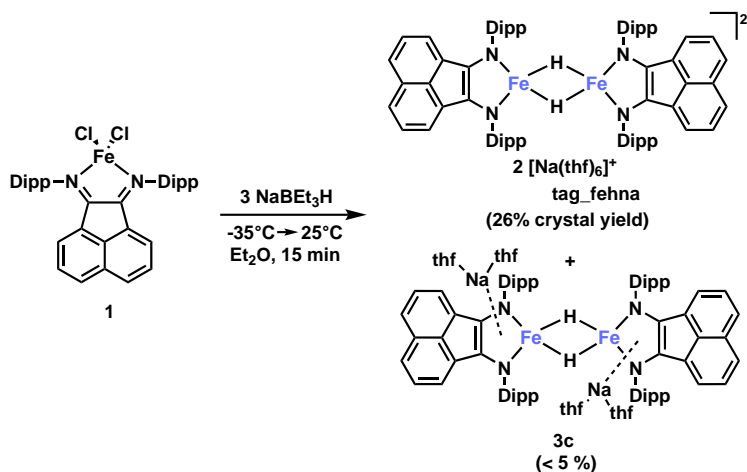
Yield 2: 122 mg (0.071 mmol, 17 %).

$^1\text{H-NMR}$ (400 MHz, 27°C, THF- d_8): δ [ppm] = 8.33 (d, $^3J(\text{HH}) = 8.2 \text{ Hz}$, 4 H, CH_{BIAN}); 7.23 (d, $^3J(\text{HH}) = 7.0 \text{ Hz}$, 4 H, CH_{BIAN}); 6.59 (s, 12 H, CH_{Ar}); 6.50 (dd, $^3J(\text{HH}) = 8.2 \text{ Hz}$, $^3J(\text{HH}) = 6.9 \text{ Hz}$, 4 H, CH_{BIAN}); 3.52–3.34 (br, 8 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.87 (d, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.13–0.03 (br, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); -74.53 (s, 3 H, Co-H).

MS (ESI, m/z): ber.: 186.01 $[\text{M}]^+$, gef.: 186.08 $[\text{M}]^+$.

IR: (ATR) $\tilde{\nu}$ [cm^{-1}] = 3048, 2983, 2953, 2861, 1608, 1582, 1488, 1458, 1419, 1378, 1313, 1253, 1172, 1101, 1036, 998, 918, 884, 865, 806, 798, 757, 681.

8.2.8 Synthesis of $[\text{Na}(\text{thf})_6]_2[\{\text{Fe}(\text{Dip}^{\text{P}}\text{BIAN})\}_2(\mu\text{-H})_2]$ 3b



Scheme 4: Synthesis of $[\text{Na}(\text{thf})_6]_2[\{\text{Fe}(\text{Dip}^{\text{P}}\text{BIAN})\}_2(\mu\text{-H})_2]$.

Inside a Schlenk flask $\text{Dip}^{\text{P}}\text{BIAN})\text{FeCl}_2$ (502 mg, 0.80 mmol, 1.0 equiv.) was suspended in 20 mL diethylether. The reaction vessel was closed with an unused septum and cooled to -35°C . A solution of NaEt_3BH (2.4 mmol, 1.0 M in THF, 3.0 equiv.) in a syringe closed with rubber stopper was also cooled to -35°C . The NaEt_3BH solution was added dropwise

to the suspension of $\text{DippBIAN})\text{FeCl}_2$. During the addition a color change from pale green to a dark green solution was observed. The reaction mixture was allowed to warm to room temperature and stirred for 15 min. The product was precipitated by the addition of 20 mL *n*-hexane and filtered with a frit (P3). After washing with *n*-hexane (6 mL), diethylether (4 mL) and THF/*n*-hexane 2:1 (5 mL) the residue was extracted with THF (25 mL) and filtrated through a P3 frit. The product was crystallized at -35°C and the solvent decanted. After washing with hexane (2 x 2 mL) it was dried via exposure to the glovebox atmosphere. The product was obtained as black, block shaped crystals.

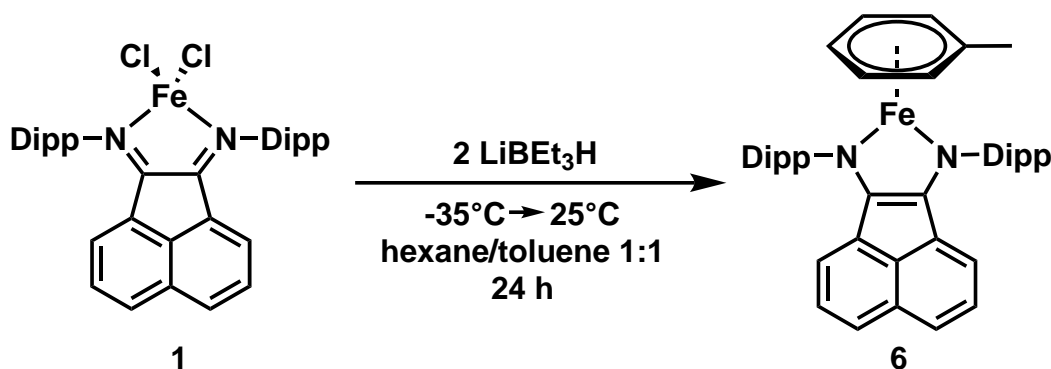
Yield: 227 mg (0.104 mmol, 26 %).

Elemental Analysis: Found (calc. for $\text{Na}_2\text{L}_2\text{Fe}_2\text{H}_2 + 4.55 \text{ THF}$): C: 72.75 (72.75); H: 8.17 (8.01); N: 3.92 (3.76).

Magnetic Moment (Evans): $\frac{\mu_{eff}}{\mu_B}$ (THF- d_8 , 293 K): 6.78.

IR: (ATR) $\tilde{\nu} [\text{cm}^{-1}] = 2983, 2955, 1584, 1496, 1457, 1425, 1314, 1253, 1209, 1180, 1099, 1045, 918, 886, 865, 812, 799, 756, 683$.

8.2.9 Synthesis of $[\text{Fe}(\text{L})(\text{tol})]$ (6)



Scheme 5: Synthesis of $[\text{Fe}(\text{L})(\text{tol})]$ (6)

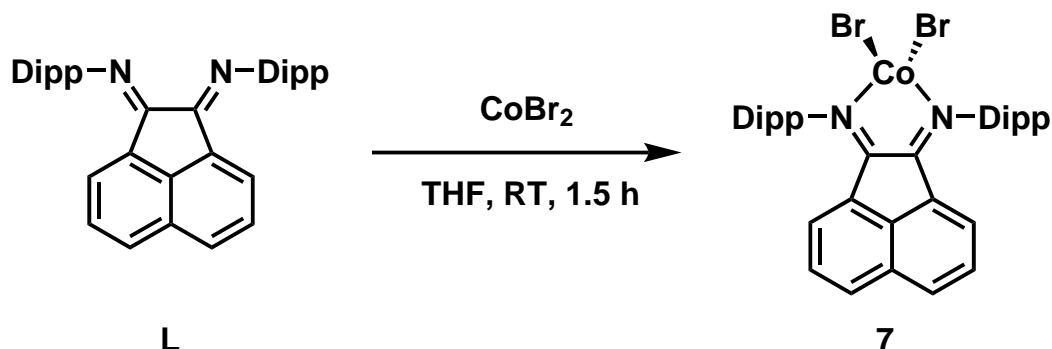
The synthesis was carried out inspired by the literature protocol.[192]

Inside a Schlenk flask $\text{DippBIAN})\text{FeCl}_2$ (504 mg, 0.80 mmol, 1.0 equiv.) was suspended in 20 mL *n*-hexane/toluene 1:1. The reaction vessel was cooled to -35°C . A solution of LiEt_3BH (1.6 mmol, 1.0 M in THF, 2.0 equiv.) was added dropwise to the suspension of $\text{DippBIAN})\text{FeCl}_2$. During the addition a color change from pale green to red was observed. The solvent was removed under reduced pressure and red residue extracted with hexane. At room temperature the product crystallized as green plates.

Yield : 60 mg (0.092 mmol, 12 %)².

$^1\text{H-NMR}$ (400 MHz, 300 K, C_6D_6): δ [ppm] = 7.47 (t, $^3J(\text{HH}) = 7.6 \text{ Hz}$, 2 H); 7.37 (d, $^3J(\text{HH}) = 7.6 \text{ Hz}$, 4 H); 7.10 (d, $^3J(\text{HH}) = 8.2 \text{ Hz}$, 2 H); 6.69 (t, $^3J(\text{HH}) = 7.6 \text{ Hz}$, 2 H); 6.09 (d, $^3J(\text{HH}) = 7.0 \text{ Hz}$, 2 H); 5.88 (t, $^3J(\text{HH}) = 5.6 \text{ Hz}$, 1 H, *p*-**H**(tol)); 4.99 (d, $^3J(\text{HH}) = 6.0 \text{ Hz}$, 2 H, *o*-**H**(tol)); 4.87 (t, $^3J(\text{HH}) = 5.8 \text{ Hz}$, 2 H, *m*-**H**(tol)); 4.04 (sept, $^3J(\text{HH}) = 6.8 \text{ Hz}$, 4 H); 2.35 (s, 3 H, PhCH_3); 1.54 (d, $^3J(\text{HH}) = 6.9 \text{ Hz}$, 12 H); 0.93 (d, $^3J(\text{HH}) = 6.8 \text{ Hz}$, 12 H).

²Only the crystal yield has been considered. The yield could be increased through further work-up.

8.2.10 Synthesis of $(^{\text{Dipp}}\text{BIAN})\text{CoBr}_2$ **7****Scheme 6:** Synthesis of $(^{\text{Dipp}}\text{BIAN})\text{CoBr}_2$ **7**.

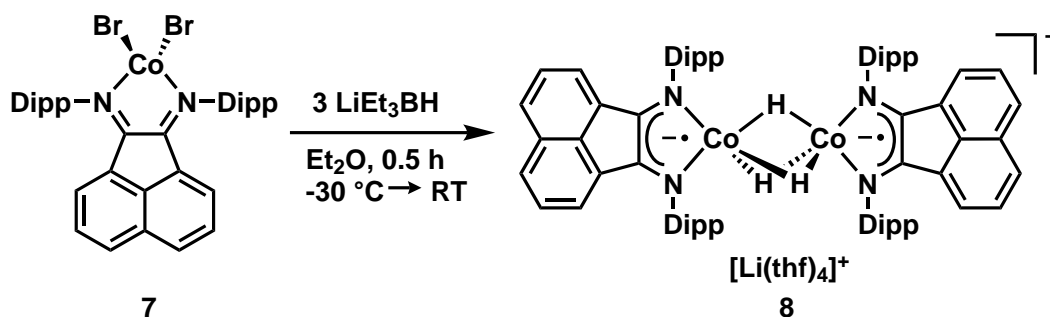
The synthesis was performed following the literature procedure.[52, 212]

A Schlenk flask was charged with CoBr_2 (2.1 g, 9.5 mmol, 1.00 equiv.) and $^{\text{Dipp}}\text{BIAN}$ (5.0 g, 10 mmol, 1.1 equiv.) in THF (120 mL). The mixture was stirred for 1.5 h, during which the color changed from orange to brown. The solvent was removed in vacuo, the residue was washed with toluene (50 mL) and dissolved in DCM (130 mL). The solution was filtered with a frit (P3) and concentrated in vacuo (90 mL). The raw product was recrystallized by carefully adding a layer of hexane (40 mL) on top of the filtrate. Black needles were isolated by decantation and washed with toluene (3 x 15 mL).

Yield: 4.5 g (7.52 mmol, 66 %).

Elemental Analysis: Found (calc. for $\text{C}_{36}\text{H}_{40}\text{Br}_2\text{CoN}_2$): C: 60.19 (60.10); H: 5.67 (5.60); N: 3.72 (3.89).

Magnetic Moment (Evans): $\frac{\mu_{\text{eff}}}{\mu_{\text{B}}}$ (THF- d_8 , 299 K): 4.91.

8.2.11 Synthesis of $[\{\text{Co}(^{\text{Dipp}}\text{BIAN})\}_2(\mu\text{-H})_3]^-$ **8****Scheme 7:** Synthesis of $[\{\text{Co}(^{\text{Dipp}}\text{BIAN})\}_2(\mu\text{-H})_3]^-$ **8**.

The synthesis of $[\{\text{Co}(^{\text{Dipp}}\text{BIAN})\}_2(\mu\text{-H})_3]^-$ **8** was carried out based on literature protocol.[52]

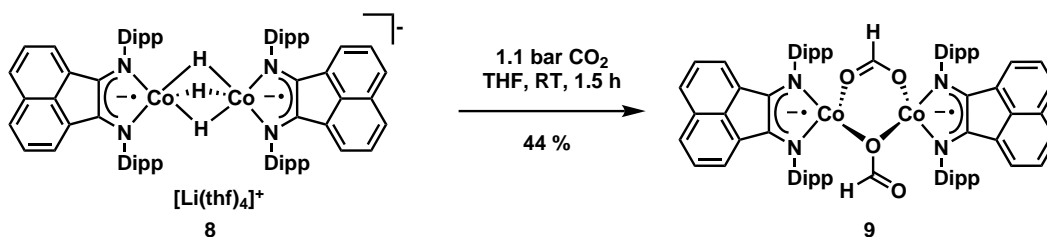
Inside a 100 mL Schlenk tube **7** (1.0 g, 1.4 mmol, 1 equiv.) was suspended in 10 mL diethylether. The reaction vessel was closed with an unused septum and cooled to -35°C

as well as the LiEt_3BH solution (4.2 mmol, 1.05 M in THF, 3 equiv.). It was added dropwise to the suspension of **7**. During the addition a color change from pale brown to a dark green solution was observed. The reaction mixture was allowed to warm to room temperature and was stirred for 30 min. The product was precipitated by the addition of 10 mL hexane and filtered with a frit (P3). After washing with hexane (3 x 8 mL) and diethylether (3 x 8 mL) the filter cake was extracted subsequently with THF/hexane 1:1 (2 x 8 mL), THF/hexane 2:1 (3 x 8 mL). The product was crystallized at -35°C and the solvent decanted. After washing with hexane (2 x 2 mL) it was dried via exposure to the glovebox atmosphere.

Yield: 321 mg (0.227 mmol, 33 %).

$^1\text{H-NMR}$ (400 MHz, 27°C , $\text{THF}-d_8$): δ [ppm] = 8.33 (d, $^3J(\text{HH}) = 8.2\text{ Hz}$, 4H, CH_{BIAN}); 7.23 (d, $^3J(\text{HH}) = 7.0\text{ Hz}$, 4H, CH_{BIAN}); 6.59 (s, 12H, CH_{Ar}); 6.50 (dd, $^3J(\text{HH}) = 8.2\text{ Hz}$, $^3J(\text{HH}) = 6.9\text{ Hz}$, 4H, $\text{CH}_{(\text{BIAN})}$); 3.52–3.34 (br, 8H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.87 (d, 24H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.13–0.03 (br, 24H, $-\text{C}(\text{CH}_3)_2\text{H}$); -74.53 (s, 3H, Co-H).

8.2.12 Synthesis of $[\{\text{Co}(\text{DippBIAN})\}_2(\text{formato})_2]$ **9**



Scheme 8: Synthesis of $[\{\text{Co}(\text{DippBIAN})\}_2(\text{formato})_2]$ **9**.

In a 50 mL Schlenk tube hydridocobaltate **8** (210 mg, 148 μmol) were dissolved in 20 mL THF. The solution was frozen in liquid nitrogen and the gas phase exchanged with CO_2 . The solution was allowed to warm to room temperature and stirred for 1.5 h. During that time the color changed from dark green-blue to brown to olive green. The solvent was removed in vacuo and the green residue was extracted with *n*-hexane. A cannula was placed in the septum of the vial to allow slow evaporation of the solvent. After 4 d the mother liquor was decanted and the crystals were dried via exposure to the glovebox atmosphere.

Yield: 79 mg (65 μmol , 44 %).

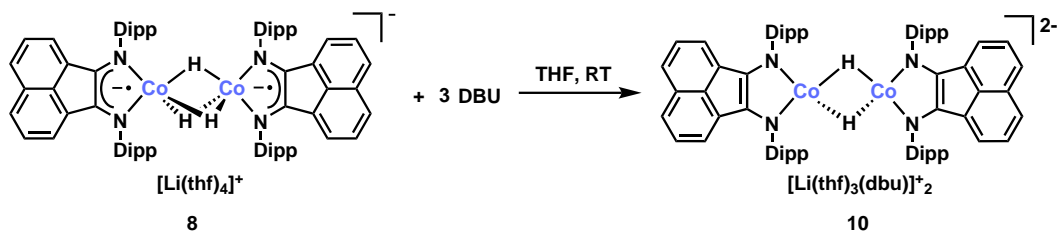
Elemental Analysis: Found (calc. for $\text{Co}_2\text{C}_{74}\text{H}_{82}\text{N}_4\text{O}_4$): C: 73.50 (73.49); H: 6.91 (6.83); N: 4.64 (4.63).

$^1\text{H-NMR}$ (400 MHz, 27°C , $\text{THF}-d_8$): δ [ppm] = 78.75 (bs); 75.69 (bs); 65.95 (bs); 7.20 (s); 2.57 (bs); 2.09 (bs); 1.23 (d, $^3J(\text{HH}) = 6.8\text{ Hz}$); 1.07 (d, $^3J(\text{HH}) = 7.5\text{ Hz}$); 0.12 (s); 78.75 (bs); -0.24 (d, $^3J(\text{HH}) = 7.5\text{ Hz}$); -0.84 (s); -4.76 (bs); -5.10 (bs); -5.90 (bs); -6.14 (d, $^3J(\text{HH}) = 7.2\text{ Hz}$); -13.27 (bs); -16.00 (bs); -19.32 (bs).

Magnetic Moment (Evans): $\frac{\mu_{\text{eff}}}{\mu_{\text{B}}}$ ($\text{THF}-d_8$, 299 K): 4.59.

MS (LIFDI (FD+), m/z): calc.: 1208.4995 $[\text{M}]^+$, found: 1208.4948 $[\text{M}]^+$.

8.2.13 Reaction of $[\{\text{Co}(\text{DippBIAN})\}_2(\mu\text{-H})_3]^-$ **8** with DBU

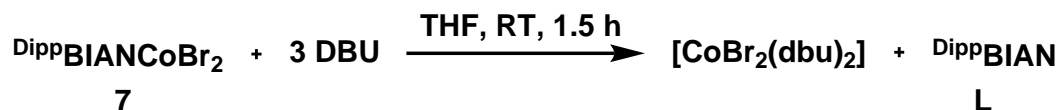


Scheme 9: Deprotonation of $[\{\text{Co}^{\text{DippBIAN}}\}_2(\mu\text{-H})_3]^-$ **8** with DBU.

$[\{\text{Co}^{\text{DippBIAN}}\}_2(\mu\text{-H})_3]^-$ **8** (21 mg, 15 μmol , 1 equiv.) was dissolved in 0.7 mL THF- d_8 and DBU (9 μL , 60 μmol , 4 equiv.) was added dropwise. A ^1H -NMR measurement of the reaction mixture showed complete conversion after 20 min. The reaction mixture was layered with *n*-hexane (0.5 mL). Dihydridocobaltate **10** was obtained as dark green-blue crystals.

Yield: 11 mg (6 μmol , 40 %).

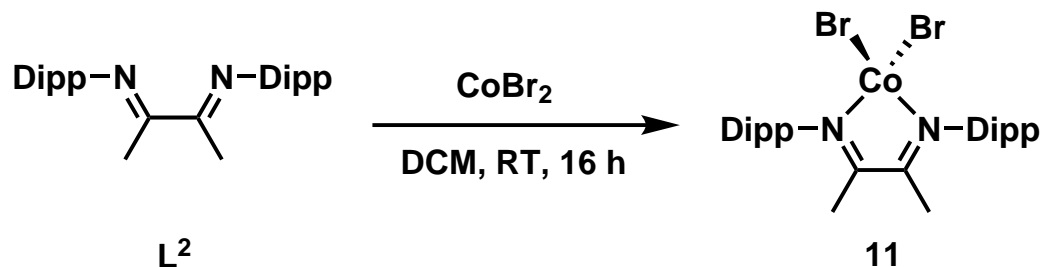
8.2.14 Reaction of $(^{\text{DippBIAN}})\text{CoBr}_2$ **7** with DBU



Scheme 10: Reaction of $(^{\text{DippBIAN}})\text{CoBr}_2$ **7** with DBU.

$(^{\text{DippBIAN}})\text{CoBr}_2$ **7** (144 mg, 0.20 mmol, 1 equiv.) was dissolved in 2 mL THF and DBU (90 μL , 0.60 mmol, 3 equiv.) was added dropwise. An immediate color change of the solution from brown to green was observed and a yellow precipitant formed. The mixture was filtrated and the filtrate layered with hexane. Yellow and blue crystals formed. The yellow precipitant precipitate was determined to be free $^{\text{DippBIAN}}$ ligand. Single crystal XRD revealed the blue crystals to be complex $[\text{Co}(\text{dbu})_2\text{Br}_2]^3$.

8.2.15 Synthesis of $(^{\text{DippBDI}})\text{CoBr}_2$ **11**



Scheme 11: Synthesis of $(^{\text{DippBDI}})\text{CoBr}_2$ **7**.

The synthesis was performed following the literature procedure.[52, 213]

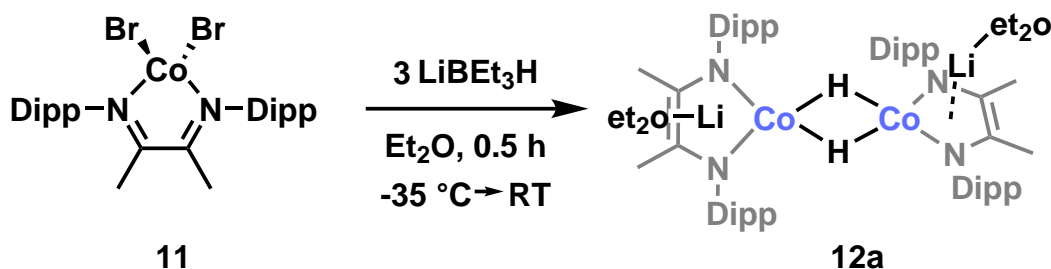
³The quality of the measurement is low but sufficient to identify the compound.

A Schlenk flask was charged with CoBr_2 (2.2 g, 10 mmol, 1.00 equiv.) and DippBDII (4.2 g, 11 mmol, 1.1 equiv.) in DCM (150 mL). The mixture was stirred for 16 h, during which the color changed from yellow to green. The solvent was removed in vacuo, the residue was washed with hexane (3 x 10 mL). The product was obtained as green solid.

Yield: 6.0 g (9.6 mmol, 96 %).

Elemental Analysis: Found (calc. for $\text{C}_{28}\text{H}_{40}\text{Br}_2\text{CoN}_2$): C: 54.07 (53.95); H: 6.40 (6.47); N: 4.51 (4.49).

8.2.16 Synthesis of $[\{\text{Co}(\text{DippBDI})\}_2(\mu\text{-H})_2]^-$ **12a**



Scheme 12: Synthesis of $[\{\text{Co}(\text{DippBDI})\}_2(\mu\text{-H})_2]^-$ **12a**.

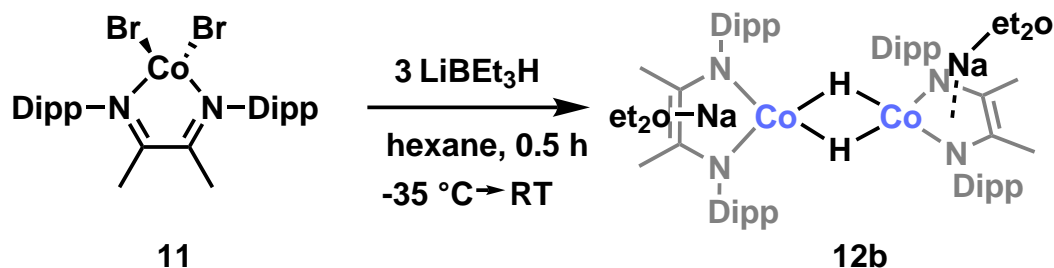
Inside a 100 mL Schlenk tube **11** (679 mg, 1.09 mmol, 1 equiv.) was suspended in 10 mL diethylether. The reaction vessel was closed with an unused septum and cooled to -35°C . A LiEt_3BH solution (3.1 mmol, 1.05 M in THF, 3 equiv.) cooled to -35°C was added dropwise to the suspension of **7**. During the addition a color change from green to a dark green solution was observed. The reaction mixture was allowed to warm to room temperature and was stirred for 40 min. 35 mL hexane were added and filtered with a frit (P3). At -35°C the crude product was precipitated and extracted with hexane (3 x 2 mL). The product was crystallized at -35°C and the solvent decanted. After washing with hexane (0.1 mL) the dark green blocks were dried via exposure to the glovebox atmosphere.

Yield: 11 mg (10 μmol , 1.8 %).

^1H -NMR (400 MHz, 27°C , $\text{THF}-d_8$): δ [ppm] = 6.94 (t, $^3J(\text{HH}) = 7.3 \text{ Hz}$, 4 H, *p*-H(Dipp)); 6.85 (d, $^3J(\text{HH}) = 7.5 \text{ Hz}$, 8 H, Cm_{BIAN}); 3.49–3.29 (m, 4 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 1.28–1.14 (m, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.20–0.07 (m, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); -0.87 (s, 12 H, $-\text{CH}_3$); -45.95 (s, 2 H, Co-H).

8.2.17 Synthesis of $[\{\text{Co}(\text{DippBDI})\}_2(\mu\text{-H})_2]^-$ **12b**

Inside a 100 mL Schlenk tube **11** (932 mg, 1.5 mmol, 1 equiv.) was suspended in 10 mL hexane. The reaction vessel was closed with an unused septum and cooled to -35°C as well as the NaEt_3BH solution (4.5 mmol, 1.0 M in THF, 3 equiv.). It was added dropwise to the suspension of **7**. During the addition a color change from green to a dark green solution was observed. The reaction mixture was allowed to warm to room temperature and was stirred for 25 min. After filtration with a frit (P3) the filtrate the solvent was removed in vacuo and the residue co-evaporated with hexane (3 x 2 mL). The residue was washed with



Scheme 13: Synthesis of $[\{\text{Co}(\text{DippBDI})\}_2(\mu\text{-H})_2]^-$ **12b**.

hexane (10 mL) and extracted with hexane/diethylether 1:1. The green extract was stored at $-35\text{ }^\circ\text{C}$ for 4 d. After washing with hexane (0.1 mL) the dark green blocks were dried under reduced pressure.

Yield: 15 mg (13 μmol , 1.6 %).

^1H -NMR (400 MHz, 27°C , $\text{THF}-d_8$): δ [ppm] = 6.95 (t, $^3J(\text{HH}) = 7.3\text{ Hz}$, 4 H, *p*-H(Dipp)); 6.85 (d, $^3J(\text{HH}) = 7.5\text{ Hz}$, 8 H, Cm_{BIAN}); 3.39 (sept, $^3J(\text{HH}) = 6.9\text{ Hz}$, 4 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 1.23 (d, $^3J(\text{HH}) = 6.0\text{ Hz}$, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); 0.77 (d, $^3J(\text{HH}) = 6.8\text{ Hz}$, 24 H, $-\text{C}(\text{CH}_3)_2\text{H}$); -0.83 (s, 12 H, $-\text{CH}_3$); -45.93 (s, 2 H, Co-H).

9 Conclusions

This thesis has explored the complex interplay between electronic structure, molecular structure, and reactivity in transition metal complexes, particularly homoleptic tris(diimine) iron(II) complexes, bis(imino)acenaphthene-stabilized iron complexes, and organic carbenes. Across these studies, density functional theory (DFT) played a pivotal role in characterizing spin-state energy splittings, and the electronic structures of various iron species, while also revealing the limitations and challenges associated with these computational methods.

The investigation into the sensitivity of spin-state energy splittings in iron(II) complexes highlighted the significant dependence of these properties on the choice of exchange-correlation functionals (see Chapter 5). Despite the variability in performance among functionals, careful selection allowed for accurate reproduction of experimental ground spin states. This work also underscored the importance of addressing SCF convergence issues, particularly in complexes with redox-active ligands, where static correlation effects may demand more advanced electronic structure methods.

In the study of carbenes, the analysis of spin gap predictions emphasized the crucial role of molecular structure and functional selection in determining DFT accuracy (see Chapter 6). The successful application of Δ -machine learning (ML) techniques demonstrated the potential for improving spin gap predictions through error correction. The promising results obtained from Δ -ML approaches suggest that future work could focus on the refinement of ML descriptors and feature selection to enhance predictive accuracy and computational efficiency further.

Finally, the synthesis and characterization of bis(imino)acenaphthene-stabilized iron complexes revealed the remarkable ability of these ligands to stabilize iron in unusual oxidation states, leading to the discovery of novel hydrogen-bridged dinuclear complexes (see Chapter II). These findings contribute to a deeper understanding of the electronic structures and reactivity of such complexes, with implications for their application in catalysis. Further investigations could explore the scope of these complexes in catalyzing more challenging transformations and their potential to be tuned through ligand modification for enhanced activity and selectivity.

Overall, this thesis advances our knowledge of the factors influencing the electronic structure and reactivity of transition metal complexes, providing valuable insights into the development of more accurate computational methods and the design of novel catalytic systems. The integration of experimental and theoretical approaches throughout this work highlights the importance of interdisciplinary research in pushing the boundaries of modern chemistry.

A List of Publications

- Rastetter, U., Jacobi von Wangelin, A. & Herrmann, C. Redox-active ligands as a challenge for electronic structure methods. *J. Comput. Chem.* **44**, 468–479. doi:10.1002/jcc.27013 (2023).

B Supporting Information for Redox-Active Ligands as a Challenge for Electronic Structure Methods

B.1 Additional Data on Spin-State Energy Splittings of Trisdiimineiron(II) Complexes

As shown in Figure B.1, there are two possible diastereomers for complex NENTABmod (**5** and **5a**). A positive slope ($\frac{\partial E^{\text{HS-LS}}}{\partial a^{\text{HF}}}$) were observed for both isomers (Figure B.2 and the same change of bond pattern was observed for the optimization of HS structures (see Section B.6)).

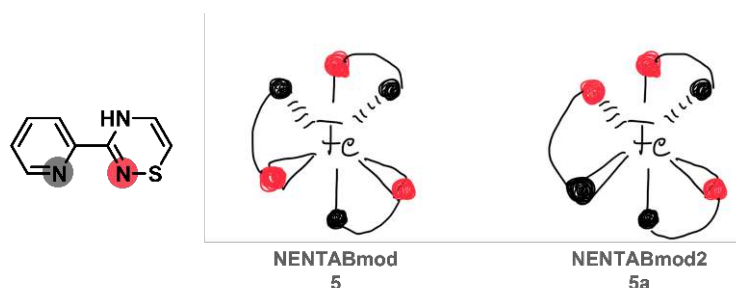


Figure B.1: Possible diastereomers for complex NENTABmod. Reproduced from reference [71].

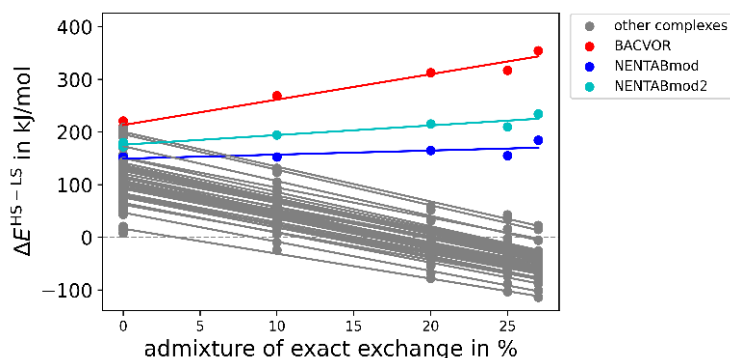


Figure B.2: Spin splitting energy splittings of BP86-D3BJ/def2TZVP-optimized trisdiimineiron(II) complexes including NENTABmod2 as function of exact exchange admixture a^{HF} . Used xc functionals with $a^{\text{HF}}/\%$ given in parenthesis: BP86 (0), TPSS (0), PBE (0), TPSSh (10), B3LYP (20), PBE0 (25), M06 (27). Reproduced from reference [71].

The influence of the local projector on the local spin densities were investigated for BACVOR's HS structure using BP86-D3BJ/def2-TZVP Figure B.3. Local Hirshfield, AIM and Becke spin densities have been calculated using MULTIWFN Version 3.5[214] (medium grid quality).

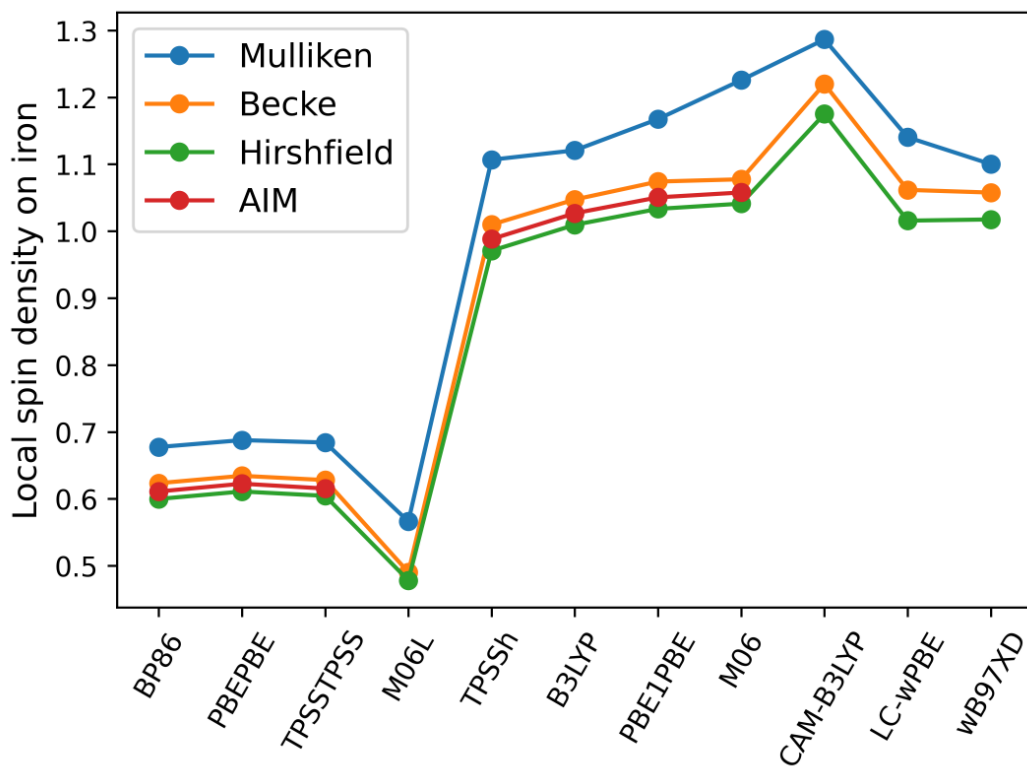


Figure B.3: Local spin densities on Iron center of BP86-D3BJ/def2-TZVP optimized HS structure BACVOR using different local projectors. Reproduced from reference [71].

Table B.1: Prediction of ground spin-state for Fe(II) complexes with BP86/def2-TZVP optimized structures

No.	BP86	TPSS	PBE	TPSSH	M06L	LC- ω PBE	wB97XD	B3LYP	CAM-B3LYP	PBE0	M06	average	lit_spin
1	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	HS	LS	LS
2	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	HS	LS	LS
3	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	LS
4	LS	LS	LS	LS	LS	LS	HS	LS	HS	HS	HS	LS	LS
5	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
6	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
7	LS	LS	LS	LS	LS	LS	LS	LS	HS	HS	HS	LS	LS
8	LS	LS	LS	LS	LS	LS	HS	LS	HS	HS	HS	LS	LS
9	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	LS
9B	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
10	LS	LS	LS	LS	LS	LS	LS	LS	LS	HS	HS	LS	LS
11	LS	LS	LS	LS	LS	LS	HS	LS	HS	HS	HS	LS	LS
12	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	LS
13	LS	LS	LS	LS	LS	LS	LS	LS	LS	HS	HS	LS	LS
14	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	LS
15	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	LS
16	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
17	LS	LS	LS	LS	LS	LS	LS	LS	HS	HS	HS	LS	LS or SCO
18	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	SCO
19	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
20	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
21	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
22	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
23	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	SCO
24	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
25	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
26	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
27	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
28	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	SCO

Continued on next page

No.	BP86	TPSS	PBE	TPSSH	M06L	LC- ω PBE	wB97XD	B3LYP	CAM-B3LYP	PBE0	M06	average	lit_spin
29	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
30	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
31	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
32	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	LS	SCO
33	LS	LS	LS	LS	LS	LS	LS	LS	LS	HS	HS	LS	SCO
34	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	SCO
35	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	HS	SCO
36	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	HS
37	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	LS	HS
38	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS
39	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	HS	HS
40	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	HS	HS	HS
41	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	HS
42	LS	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	LS	NaN
43	LS	LS	LS	LS	LS	HS	HS	HS	HS	HS	HS	LS	NaN

B.2 Additional Data on Spin-State Energy Splittings of Trisdiimineiron(I) Complexes

Reduced complexes of $[\text{L}_3\text{Fe}]^+$ were obtained by adding an electron to the optimized (BP86-D3BJ/def2-TZVP) iron(II) structures $[\text{L}_3\text{Fe}]^{2+}$. Geometry optimizations were then performed for the LS (doublet) and HS (quartet) spin state using the same methodology. For the HS structure of NENTABmod taking the unoptimized structure (manipulated crystal structure) did not result in significant changes of the bond pattern compared to the preoptimized structure (B.10).

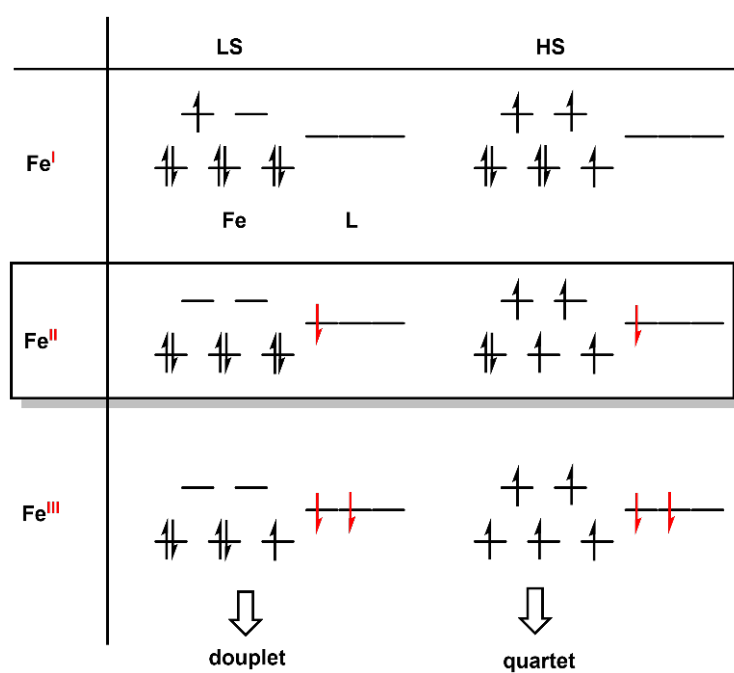


Figure B.4: Possible occupation of metal d and ligand centered π^* orbitals of reduced $[\text{L}_3\text{Fe}]^+$ complexes for HS and LS cases. Reproduced from reference [71].

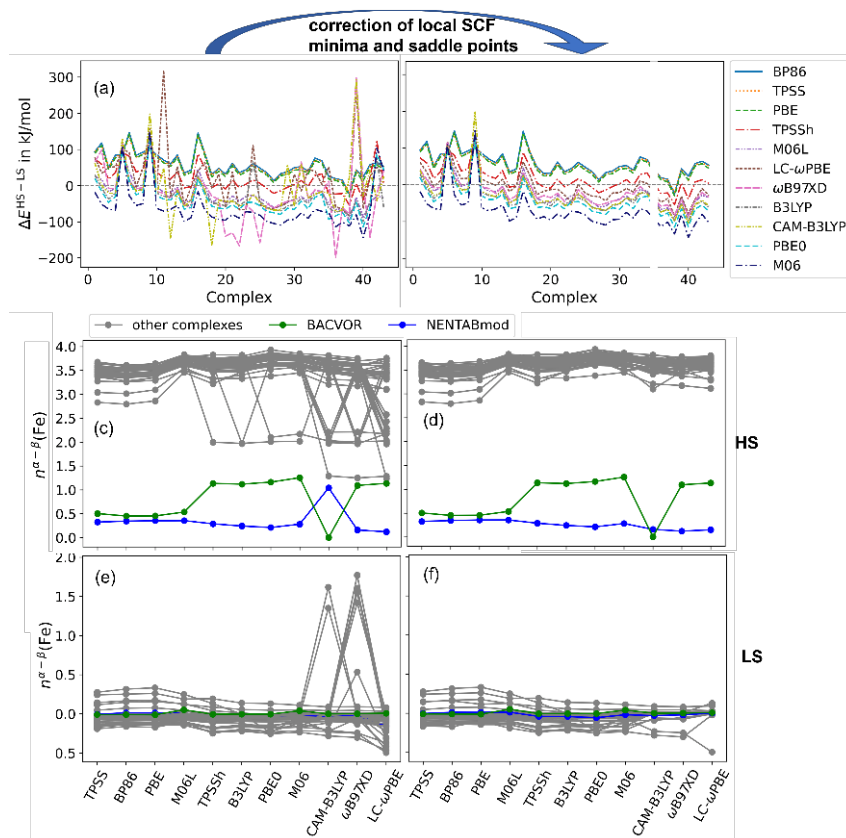


Figure B.5: Spin-state energy splittings of reduced tris(diimine) complexes $[\text{L}_3\text{Fe}]^+$, evaluated from single-point calculations with different xc functionals before (a) and after (b) correction of convergence issues by using orbitals from BP86 (HS) or B3LYP (LS) as an initial guess and optimization of wave function with internal instability (“stable=opt” keyword in the GAUSSIAN program package). The labels of the xc functionals are ordered in the same manner as their spin-state energy splittings. For each complex, the molecular HS and LS structures were optimized with BP86-D3BJ/def2TZVP. Mulliken numbers of unpaired electrons on Fe for the HS complexes before (c) and after (d) the correction of SCF convergence issues as well as for LS complexes before (e) and after (f) these corrections. Reproduced from reference [71]

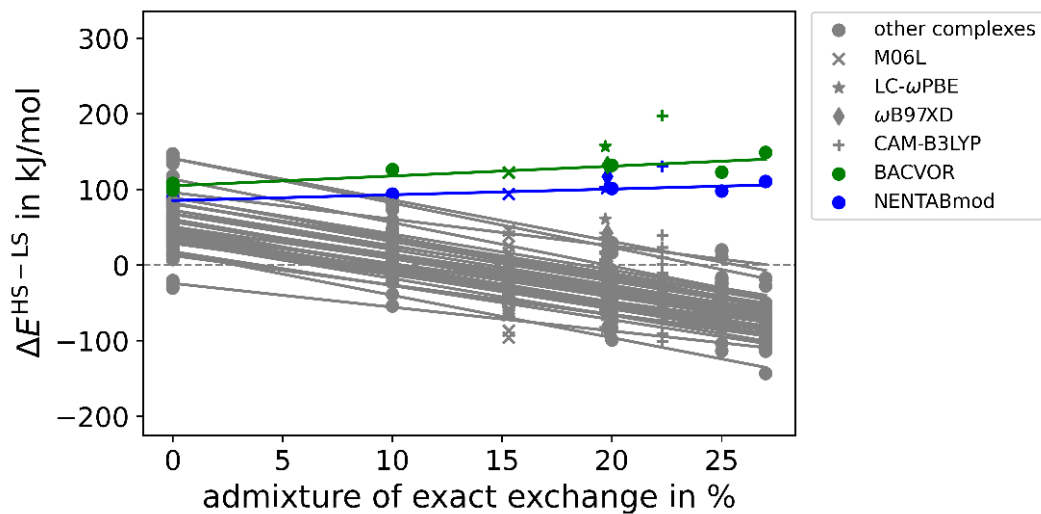


Figure B.6: Spin-state energy splittings of reduced complexes between quartet (HS) and doublet (LS) spin state of reduced $[\text{L}_3\text{Fe}]^+$ complexes over the amount of exact exchange. Structures have been optimized with BP86-D3BJ/def2-TZVP. Reproduced from reference [71].

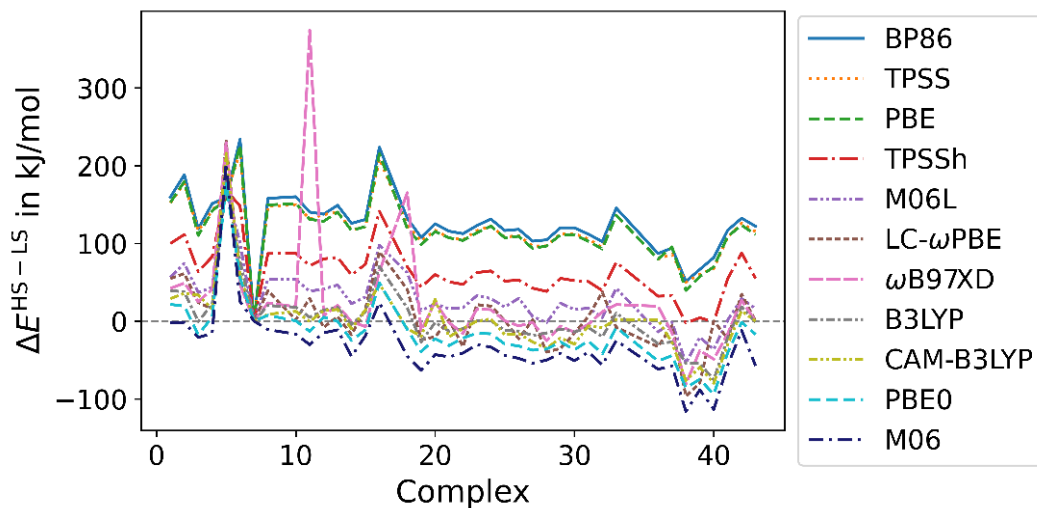


Figure B.7: Spin-state energy splittings of reduced complexes between sextet (HS) and doublet (LS) spin state of reduced $[\text{L}_3\text{Fe}]^+$ complexes. Structures have been optimized with BP86-D3BJ/def2-TZVP. Reproduced from reference [71].

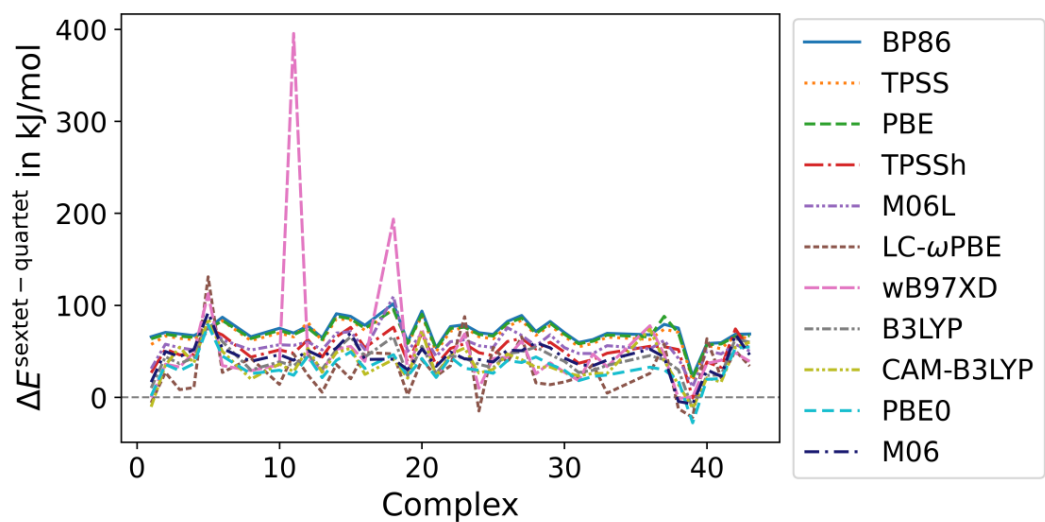


Figure B.8: Spin-state energy splittings of reduced complexes between sextet (HS) and doublet (LS) spin state of reduced $[\text{L}_3\text{Fe}]^+$ complexes. Structures have been optimized with BP86-D3BJ/def2-TZVP. Reproduced from reference [71].

B.3 Spin Contamination and FOD Analysis

Table B.2: $\langle S^2 \rangle$ values of $[\text{L}_3\text{Fe}]^{+2}$ HS complexes (quintet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	TPSS	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
1	6.03	6.10	6.04	6.05	6.05	6.03	6.02	6.04	6.06	6.07	6.14
2	6.02	6.08	6.03	6.04	6.06	6.02	6.01	6.04	6.07	6.08	6.14
3	6.01	6.04	6.02	6.02	6.02	6.01	6.01	6.02	6.03	6.03	6.06
4	6.04	6.18	6.08	6.11	6.14	6.04	6.03	6.09	6.16	6.18	6.27
5	6.01	6.04	6.02	6.03	6.03	6.01	6.01	6.02	6.04	6.05	6.07
6	6.02	6.04	6.01	6.02	6.07	6.01	6.01	6.01	6.08	6.09	6.05
7	6.03	6.12	6.05	6.07	6.09	6.03	6.02	6.05	6.11	6.12	6.20
8	6.01	6.06	6.03	6.04	6.06	6.02	6.01	6.03	6.07	6.08	6.14
9	6.02	6.06	6.03	6.04	6.05	6.02	6.01	6.03	6.07	6.08	6.13
9B	6.05	6.09	6.06	6.05	6.02	6.08	6.21	6.05	6.02	6.02	6.03
10	6.02	6.08	6.03	6.04	6.06	6.02	6.01	6.04	6.07	6.08	6.14
11	6.07	6.04	6.05	6.04	6.02	6.06	6.13	6.04	6.02	6.02	6.03
12	6.03	6.09	6.04	6.05	6.06	6.03	6.02	6.04	6.09	6.10	6.16
13	6.01	6.05	6.02	6.02	6.03	6.02	6.01	6.02	6.04	6.05	6.08
14	6.01	6.04	6.02	6.02	6.03	6.01	6.01	6.02	6.04	6.04	6.08
15	6.01	6.05	6.02	6.02	6.03	6.01	6.01	6.02	6.04	6.05	6.08
16	6.01	6.04	6.02	6.02	6.03	6.01	6.01	6.02	6.04	6.04	6.07
17	6.02	6.07	6.03	6.04	6.05	6.02	6.01	6.03	6.07	6.08	6.14
18	6.01	6.06	6.02	6.03	6.05	6.01	6.01	6.02	6.06	6.07	6.11
19	6.01	6.05	6.02	6.03	6.04	6.02	6.01	6.02	6.04	6.05	6.09
20	6.02	6.06	6.03	6.03	6.04	6.02	6.01	6.03	6.06	6.07	6.11
21	6.02	6.06	6.03	6.03	6.05	6.02	6.01	6.03	6.06	6.07	6.12
22	6.03	6.11	6.05	6.06	6.08	6.03	6.02	6.05	6.10	6.11	6.18
23	6.03	6.14	6.06	6.08	6.10	6.03	6.02	6.06	6.12	6.13	6.22
24	6.02	6.07	6.03	6.04	6.05	6.02	6.01	6.03	6.07	6.08	6.13
25	6.02	6.07	6.03	6.04	6.05	6.02	6.01	6.03	6.07	6.08	6.13
26	6.01	6.05	6.02	6.02	6.04	6.01	6.01	6.02	6.05	6.05	6.09
27	6.01	6.04	6.02	6.03	6.05	6.01	6.01	6.02	6.06	6.07	6.11
28	6.02	6.09	6.04	6.05	6.08	6.02	6.01	6.04	6.09	6.11	6.17

Continued on next page

Table B.2: $\langle S^2 \rangle$ values of $[\text{L}_3\text{Fe}]^{+2}$ HS complexes (quintet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	TPSS	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
29	6.01	6.05	6.02	6.03	6.04	6.02	6.01	6.02	6.05	6.06	6.09
30	6.02	6.09	6.04	6.05	6.07	6.03	6.02	6.04	6.08	6.09	6.15
31	6.02	6.06	6.03	6.03	6.05	6.02	6.01	6.03	6.06	6.07	6.12
32	6.01	6.05	6.02	6.03	6.04	6.02	6.01	6.02	6.05	6.06	6.10
33	6.02	6.07	6.03	6.04	6.05	6.02	6.01	6.03	6.06	6.07	6.12
34	6.01	6.05	6.02	6.03	6.04	6.01	6.01	6.02	6.05	6.05	6.09
35	6.03	6.14	6.06	6.08	6.10	6.04	6.02	6.07	6.11	6.13	6.21
36	6.03	6.03	6.01	6.01	6.11	6.04	6.01	6.01	6.13	6.14	6.22
37	6.01	6.04	6.02	6.02	6.03	6.01	6.01	6.02	6.03	6.04	6.07
38	6.01	6.02	6.01	6.01	6.04	6.01	6.01	6.01	6.06	6.07	6.07
39	6.01	6.03	6.01	6.02	6.03	6.01	6.01	6.02	6.04	6.05	6.07
40	6.02	6.07	6.03	6.04	6.06	6.02	6.01	6.03	6.07	6.08	6.14
41	6.04	6.18	6.08	6.10	6.11	6.05	6.03	6.08	6.12	6.13	6.24
42	6.02	6.06	6.03	6.04	6.05	6.02	6.01	6.03	6.07	6.08	6.12
43	6.02	6.09	6.04	6.05	6.07	6.02	6.01	6.04	6.09	6.10	6.17

Table B.3: $\langle S^2 \rangle$ values of reduced $[\text{L}_3\text{Fe}]^+$ LS complexes (doublet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	TPSS	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
1	0.78	0.76	0.78	0.77	0.76	0.78	0.79	0.77	0.76	0.76	0.76
2	0.77	0.79	0.79	0.78	0.76	0.80	0.78	0.78	0.76	0.76	0.77
3	0.86	0.80	0.81	0.78	0.76	0.88	1.10	0.79	0.76	0.76	0.76
4	0.77	0.76	0.77	0.76	0.75	0.77	0.89	0.76	0.75	0.75	0.75
5	0.77	0.77	0.77	0.76	0.75	0.77	0.84	0.76	0.75	0.75	0.76
6	0.77	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.76
7	0.76	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.76
8	0.76	0.76	0.76	0.76	0.75	0.76	0.78	0.76	0.75	0.75	0.76
9	0.77	0.76	0.76	0.75	0.75	0.78	0.81	0.76	0.75	0.75	0.75

Continued on next page

Table B.3: $\langle S^2 \rangle$ values of reduced $[\text{L}_3\text{Fe}]^+$ LS complexes (doublet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	TPSS	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
10	0.77	0.76	0.76	0.76	0.75	0.77	0.79	0.76	0.75	0.75	0.76
11	0.76	0.76	0.76	0.76	0.75	0.76	0.77	0.76	0.75	0.75	0.76
12	0.77	0.76	0.76	0.76	0.76	0.77	0.79	0.76	0.75	0.75	0.76
13	0.77	0.76	0.76	0.76	0.75	0.77	0.79	0.76	0.75	0.75	0.76
14	0.77	0.76	0.76	0.76	0.76	0.77	0.79	0.76	0.75	0.75	0.76
15	0.77	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.76
16	0.77	0.77	0.77	0.76	0.76	0.78	0.80	0.76	0.75	0.75	0.76
17	0.78	0.79	0.90	0.76	0.75	0.77	1.11	0.78	0.75	0.75	0.75
18	0.77	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.75
19	0.77	0.76	0.76	0.76	0.75	0.76	0.78	0.76	0.75	0.75	0.76
20	0.77	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.76
21	0.77	0.77	0.77	0.77	0.76	0.79	0.83	0.77	0.76	0.76	0.76
22	0.77	0.76	0.76	0.76	0.75	0.76	0.77	0.76	0.75	0.75	0.76
23	0.76	0.76	0.76	0.76	0.75	0.76	0.78	0.76	0.75	0.75	0.75
24	0.77	0.76	0.76	0.76	0.75	0.76	0.78	0.76	0.75	0.75	0.76
25	0.77	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.76
26	0.77	0.76	0.76	0.76	0.76	0.77	0.80	0.76	0.75	0.75	0.76
27	0.77	0.76	0.76	0.76	0.75	0.77	0.79	0.76	0.75	0.75	0.76
28	0.84	0.82	0.82	0.80	0.77	0.81	1.00	0.80	0.76	0.76	0.78
29	0.77	0.76	0.77	0.76	0.76	0.77	0.82	0.76	0.75	0.75	0.76
30	0.81	0.79	0.79	0.78	0.77	0.82	0.90	0.78	0.76	0.76	0.77
31	0.77	0.76	0.76	0.76	0.76	0.78	0.81	0.76	0.75	0.75	0.76
32	0.76	0.79	0.79	0.78	0.77	0.82	0.76	0.78	0.76	0.76	0.77
33	0.76	0.76	0.76	0.76	0.75	0.77	0.77	0.76	0.75	0.75	0.76
34	0.78	0.76	0.77	0.76	0.76	0.78	0.79	0.76	0.75	0.75	0.76
35	1.34	1.61	1.35	0.97	0.84	1.30	1.43	1.20	0.84	0.88	0.81
36	0.82	0.80	0.80	0.79	0.77	0.84	0.93	0.79	0.76	0.76	0.77
37	0.77	0.76	0.76	0.76	0.75	0.77	0.79	0.76	0.75	0.75	0.76
38	0.78	0.77	0.77	0.77	0.76	0.78	0.82	0.76	0.76	0.76	0.76

Continued on next page

Table B.3: $\langle S^2 \rangle$ values of reduced $[L_3Fe]^+$ LS complexes (doublet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
39	0.76	0.76	0.76	0.75	0.77	0.78	0.76	0.75	0.75	0.75
40	0.77	0.76	0.76	0.76	0.77	0.80	0.76	0.75	0.75	0.75
41	0.77	0.76	0.76	0.75	0.77	0.79	0.76	0.75	0.75	0.75
42	0.86	0.83	0.82	0.80	0.88	1.02	0.80	0.76	0.76	0.78
43	0.82	0.82	0.80	0.80	0.84	0.96	0.79	0.76	0.77	0.78

Table B.4: $\langle S^2 \rangle$ values of reduced $[L_3Fe]^+$ HS complexes (quartet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
1	4.52	4.54	3.85	4.08	4.49	4.55	3.84	3.89	3.91	4.33
2	4.67	4.83	4.00	4.56	4.66	3.88	3.90	4.34	4.39	4.71
3	4.66	4.81	4.71	4.49	4.65	4.58	4.62	4.19	4.25	4.67
4	4.03	3.84	4.05	4.53	3.98	4.54	3.91	4.31	4.36	4.69
5	3.80	3.78	3.79	3.77	3.79	3.85	3.78	3.76	3.76	3.77
6	4.52	4.77	4.65	4.48	4.51	4.40	4.56	4.32	4.37	4.67
7	4.45	4.08	4.63	4.41	4.44	4.30	4.50	4.23	4.29	4.61
8	4.39	4.66	4.59	4.39	4.39	4.23	3.93	4.25	4.30	4.60
9	3.80	3.83	3.81	3.80	3.83	3.88	3.79	3.76	3.76	3.77
10	4.45	4.05	4.59	4.41	4.44	4.30	3.98	4.21	4.26	4.61
11	4.42	4.68	4.57	4.42	4.41	4.27	4.49	4.30	4.34	4.61
12	4.32	4.50	4.44	4.21	4.26	4.14	4.31	4.13	4.16	4.43
13	4.50	4.73	4.62	4.44	4.49	4.38	4.53	4.30	4.34	4.64
14	4.32	4.58	4.45	4.29	4.32	4.24	4.37	4.16	4.20	4.50
15	4.33	4.50	4.42	4.23	4.29	4.38	4.31	4.14	4.18	4.44
16	4.58	3.93	4.70	4.49	4.57	4.47	3.87	4.23	4.29	4.68
17	3.86	4.89	3.89	3.88	4.27	4.72	3.83	4.17	4.24	4.74
18	4.39	4.65	4.59	4.36	4.44	4.32	4.48	4.22	4.27	4.55
19	4.49	4.61	4.47	4.32	4.35	4.22	4.41	4.21	4.25	4.53

Continued on next page

Table B.4: $\langle S^2 \rangle$ values of reduced $[\text{L}_3\text{Fe}]^+$ HS complexes (quartet). Structures have been optimized with BP86-D3BJ/def2-TZVP.

No.	CAM-B3LYP	M06	PBE0	TPSSH	TPSS	ω B97XD	LC- ω PBE	B3LYP	BP86	PBE	M06L
20	4.27	4.55	4.41	4.25	4.14	4.29	4.48	4.33	4.14	4.18	4.47
21	4.65	3.82	4.71	4.55	4.40	4.60	4.52	3.90	4.37	4.42	4.71
22	4.38	4.49	4.47	4.24	4.15	4.32	4.51	4.33	4.16	4.20	4.45
23	4.01	4.34	4.17	4.08	4.03	4.01	3.90	4.13	4.04	4.07	4.33
24	4.58	3.79	4.59	4.39	4.27	4.48	4.43	4.49	4.26	4.31	4.58
25	4.00	4.65	4.59	4.39	4.27	4.47	4.67	4.48	4.27	4.31	4.58
26	4.43	3.80	3.96	4.30	4.19	4.39	4.06	3.91	4.19	4.23	4.49
27	4.26	4.52	4.38	4.23	4.13	4.26	4.12	4.31	4.13	4.17	4.45
28	4.68	4.83	4.76	4.58	4.42	4.66	4.64	4.66	4.39	4.43	4.73
29	4.37	4.61	4.49	4.33	4.21	3.86	4.27	4.41	4.21	4.25	4.53
30	3.97	4.75	4.66	4.48	4.33	4.58	4.57	4.57	4.30	4.34	4.65
31	4.54	4.73	4.63	4.47	4.35	4.52	4.44	4.55	4.33	4.37	4.65
32	4.63	4.80	4.69	4.49	4.29	4.62	4.56	4.61	4.26	4.32	4.67
33	4.43	4.70	4.58	4.42	4.31	4.42	4.29	4.50	4.30	4.35	4.62
34	4.19	4.53	4.39	4.17	4.00	4.20	3.96	4.29	4.00	4.05	4.45
35	4.60	4.71	4.64	4.43	4.23	4.57	4.59	4.56	4.20	4.27	4.62
36	4.59	4.75	4.68	4.51	4.36	4.58	4.50	4.59	4.33	4.38	4.66
37	4.29	4.54	4.40	4.28	4.19	4.28	4.15	4.35	4.19	4.23	4.48
38	4.49	4.63	4.64	4.48	4.39	4.49	4.58	4.59	4.38	4.42	4.64
39	4.21	4.52	4.41	4.48	4.34	4.21	4.06	4.36	4.36	4.40	4.57
40	4.39	4.66	4.59	4.38	4.26	4.53	4.43	4.49	4.27	4.31	4.57
41	4.57	4.60	4.57	4.40	4.64	4.59	4.47	4.50	4.34	4.38	4.66
42	4.72	4.88	4.80	4.61	4.45	4.70	4.65	4.70	4.42	4.47	4.76
43	4.79	4.94	4.84	4.67	4.49	4.77	4.77	4.75	4.45	4.49	4.81

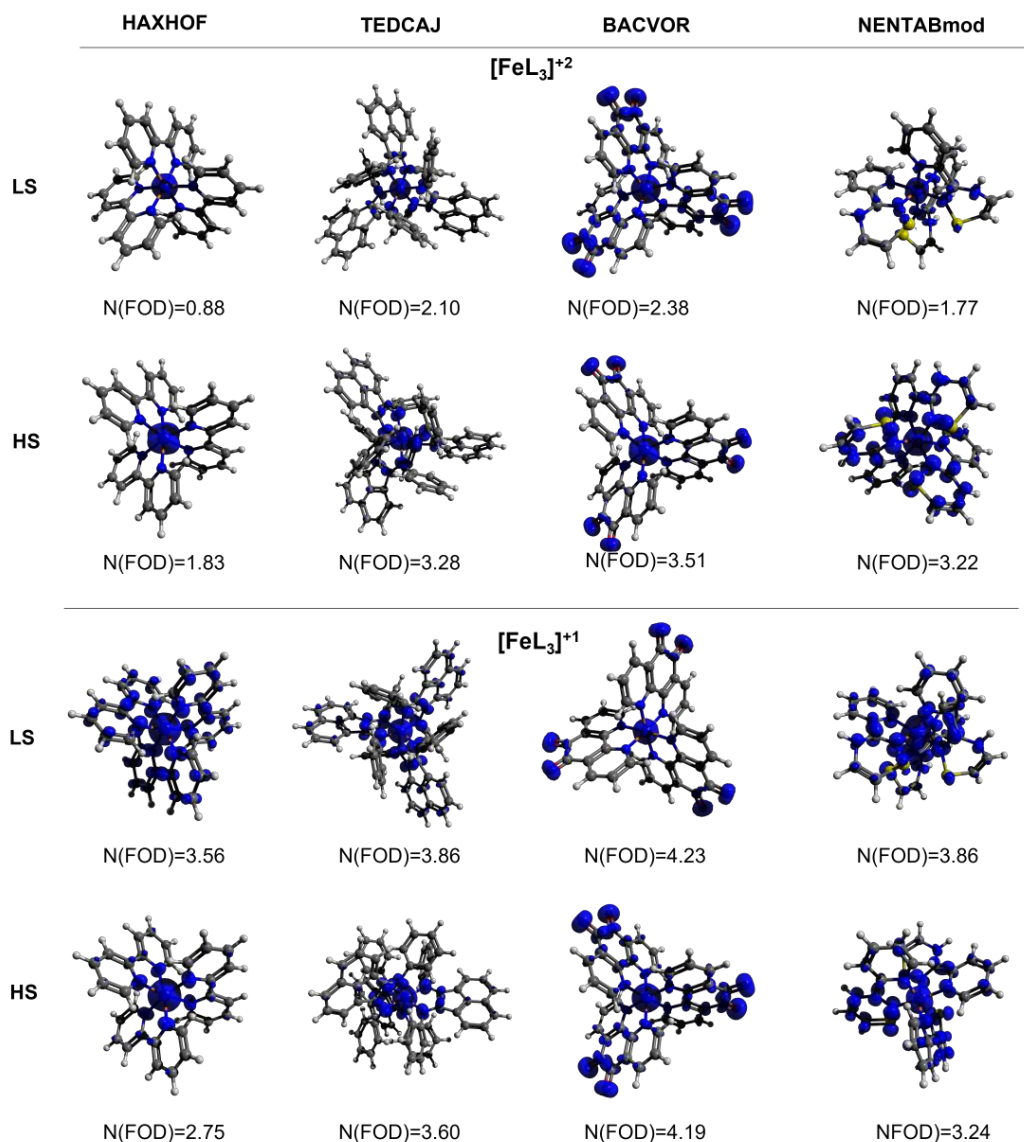


Figure B.9: FOD plots of selected $[\text{L}_3\text{Fe}]\equiv_2$ (top) and $[\text{L}_3\text{Fe}]^+$ (bottom) HS and LS complexes (isosurface value=0.005 e Bohr $^{-3}$, TPSS/def2-TZVP (T=5000 K) level). Reproduced from reference [71].

Table B.5: Fractional occupation number weighted electron density (FOD) analysis (TPSS/def2-TZVP with a smearing temperature of 5000 K). Structures have been optimized with BP86-D3BJ/def2-TZVP

Complex	No.	N(FOD) $[\text{L}_3\text{Fe}]^{+2}$		N(FOD) $[\text{L}_3\text{Fe}]^{+1}$	
		css	quintet	doublet	quartet
BOFNEP	1	0.76	1.67	1.37	1.93
OKIWEKsimp	2	0.92	1.99	2.03	2.85
QOSNIWmod	3	1.26	1.97	2.05	2.49
WEYLERsimp	4	0.90	1.78	1.99	2.66
NENTABmod	5	1.78	3.22	2.49	3.25
HEYRAEsimp	6	1.29	2.73	2.52	3.42
HAXHOF	7	0.88	1.83	2.16	2.75
NENTAB	8	1.04	2.09	2.24	2.91

Continued on next page

Table B.5: Fractional occupation number weighted electron density (FOD) analysis (TPSS/def2-TZVP with a smearing temperature of 5000 K). Structures have been optimized with BP86-D3BJ/def2-TZVP

Complex	No.	N(FOD) $[\text{L}_3\text{Fe}]^{+2}$		N(FOD) $[\text{L}_3\text{Fe}]^{+1}$	
		csc	quintet	doublet	quartet
BACVOR	9	2.38	3.51	3.17	4.19
AZUHUY	10	1.11	2.28	2.20	2.83
KOQDEZ	11	1.01	2.00	2.39	2.89
QAJKUH	12	0.77	1.71	2.13	2.60
JUSCILsimp	13	0.89	1.81	2.08	2.68
RUZNOP	14	0.84	1.83	2.01	2.29
OJIVOSsimp	15	0.64	1.58	2.02	2.35
SIXJUE	16	0.97	2.25	2.26	2.99
JOWGIL	17	0.97	1.82	1.95	2.45
VEWVEY	18	0.94	1.68	2.13	2.50
CATXAZ	19	0.83	1.73	2.20	2.53
KINQUT	20	0.71	1.62	1.97	2.30
KINQUTmod2	21	1.12	1.79	2.22	2.70
YIVSEBsimp	22	0.71	1.60	2.03	2.37
SAFGAHmod2	23	0.68	1.69	2.01	2.43
FUFHIYsimp	24	0.79	1.66	1.98	2.51
KINQUTmod	25	0.85	1.70	2.09	2.50
ABIWAKsimp	26	0.82	1.76	2.03	2.35
SAFGAHmod3	27	0.74	1.63	1.94	2.18
CATWIGsimp	28	1.12	1.76	2.08	2.46
BAKHECsimp	29	0.70	1.58	1.87	2.19
ACAHUJsimp	30	0.87	1.65	1.94	2.31
SAFGAHmod	31	1.03	1.90	2.10	2.63
QOSNIW	32	1.07	1.70	2.13	2.46
EDUKUL	33	1.08	2.12	2.39	2.93
KABYER	34	0.95	1.76	2.13	2.60
NOKFOJ	35	1.32	1.97	2.15	2.74
CATWIGmod	36	0.96	1.63	2.00	2.27
CATXAZsimp	37	0.73	1.60	1.95	2.26
CATWIG	38	1.19	1.76	2.27	2.44
CATWUS	39	1.42	2.27	2.56	3.01
CEPZOM	40	1.06	1.86	2.41	2.65
HUSMAK	41	0.89	1.65	2.14	2.61
CATWIGmod2	42	1.35	2.16	2.16	2.75
TEDCAJ	43	2.10	3.26	2.85	3.51

Continued on next page

B.4 Broken Symmetry Calculations

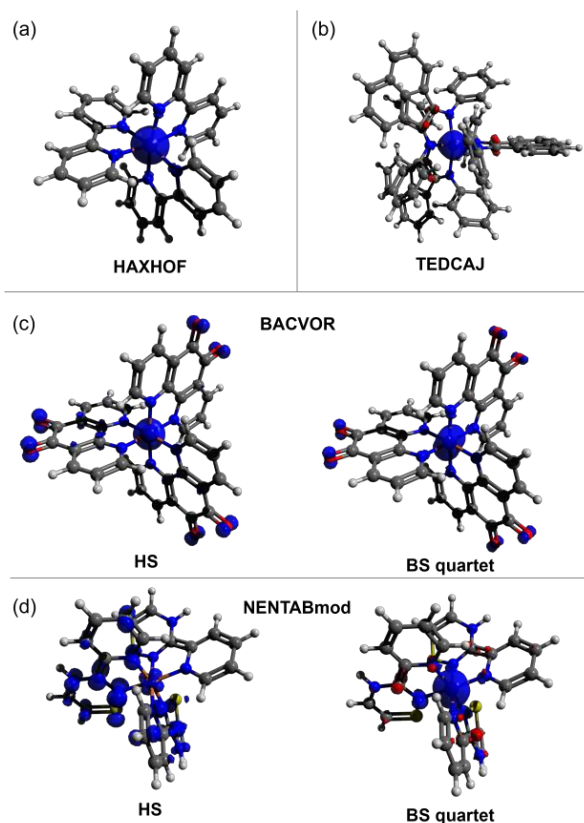


Figure B.10: Spin density of selected $[L_3Fe]^+$ complexes in HS (quartet) and BS (quartet) (isosurface value=0.01 e Bohr⁻³, B3LYP/def2-TZVP). For HAXHOF (a) and TEDCAJ (b) identical solutions were obtained for the HS and BS calculation (fragment based initial guess: fragment 1: Fe (+2, quintet), fragment 2: all ligand atoms (-1, triplet)). For BACVOR (c) the HS solution is lower in energy by 22.1 kJ mol⁻¹ compared to the BS solution and for NENTABmod (d) the HS solution is lower in energy by 91.8 kJ mol⁻¹. Reproduced from reference [71].

B.5 Studies on Correlations With Redox-Activity

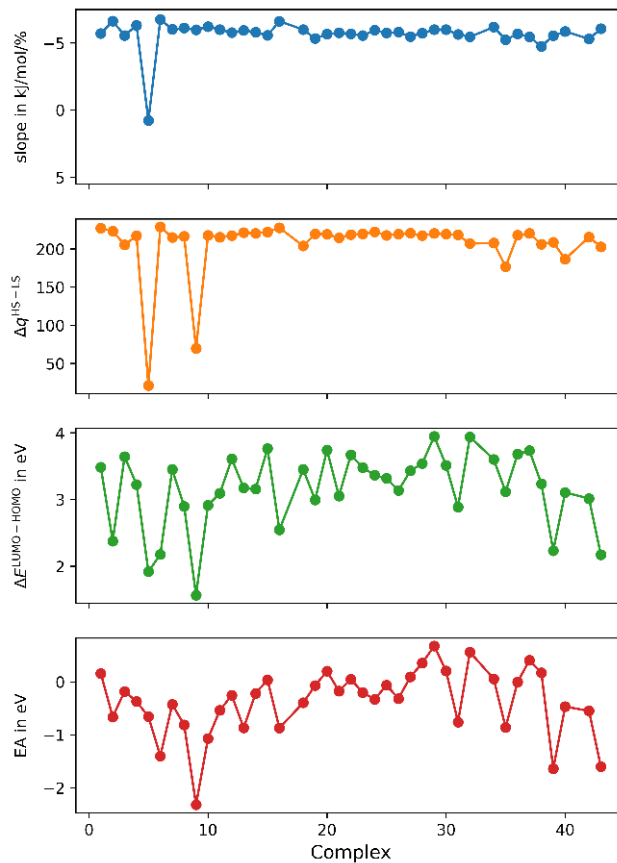


Figure B.11: HOMO–LUMO gaps of the neutral free ligands and electron affinities (EA) of the free ligands as energy difference of neutral and negatively charged ligand. Ligands have been optimized with BP86-D3BJ/def2-TZVP in neutral (singlet) and anionic (doublet) charge are calculated on optimized ligand structures using BP86-D3BJ/def2-TZVP. Slopes ($\frac{\partial E^{\text{HS-LS}}}{\partial a_{\text{HF}}}$) and difference in local charges on iron centers ($q^{\text{HS-LS}}(\text{NPA})$) have been evaluated by single-point calculations with all xc functionals and BP86-D3BJ/def2-TZVP optimized Fe(II) complexes. Reproduced from reference [71].

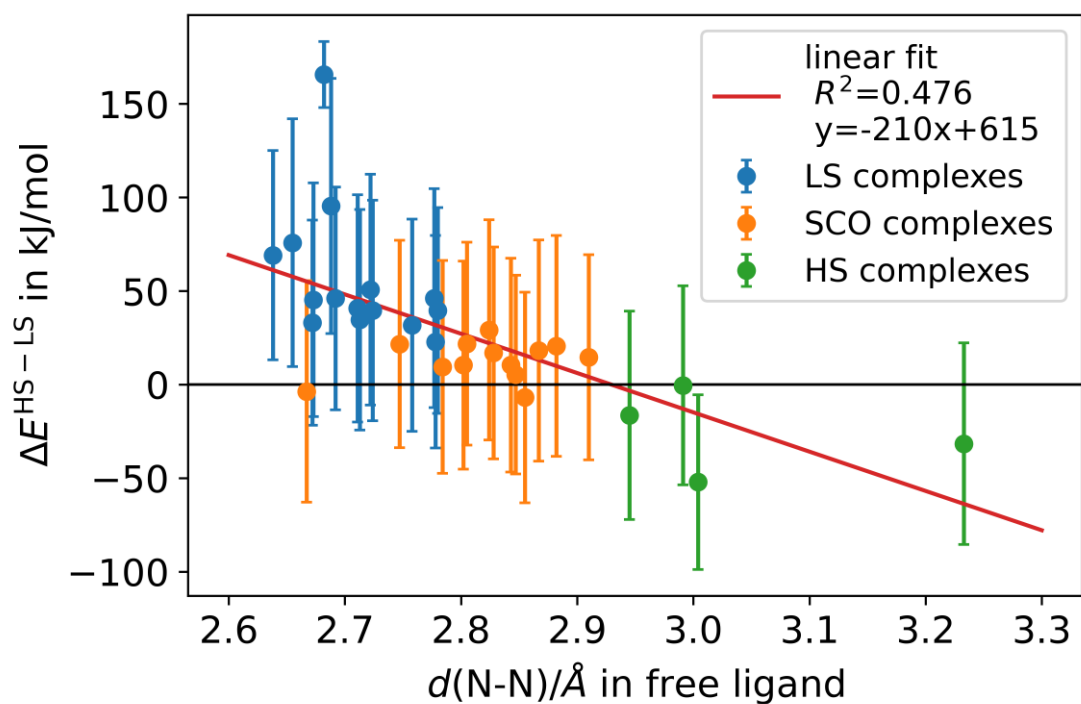


Figure B.12: Dependence of spin-state energy splittings of Fe(II) complexes on the N–N bond length in the free ligand (crystal structure). Structures have been optimized with BP86-D3BJ/def2-TZVP. Reproduced from reference [71].

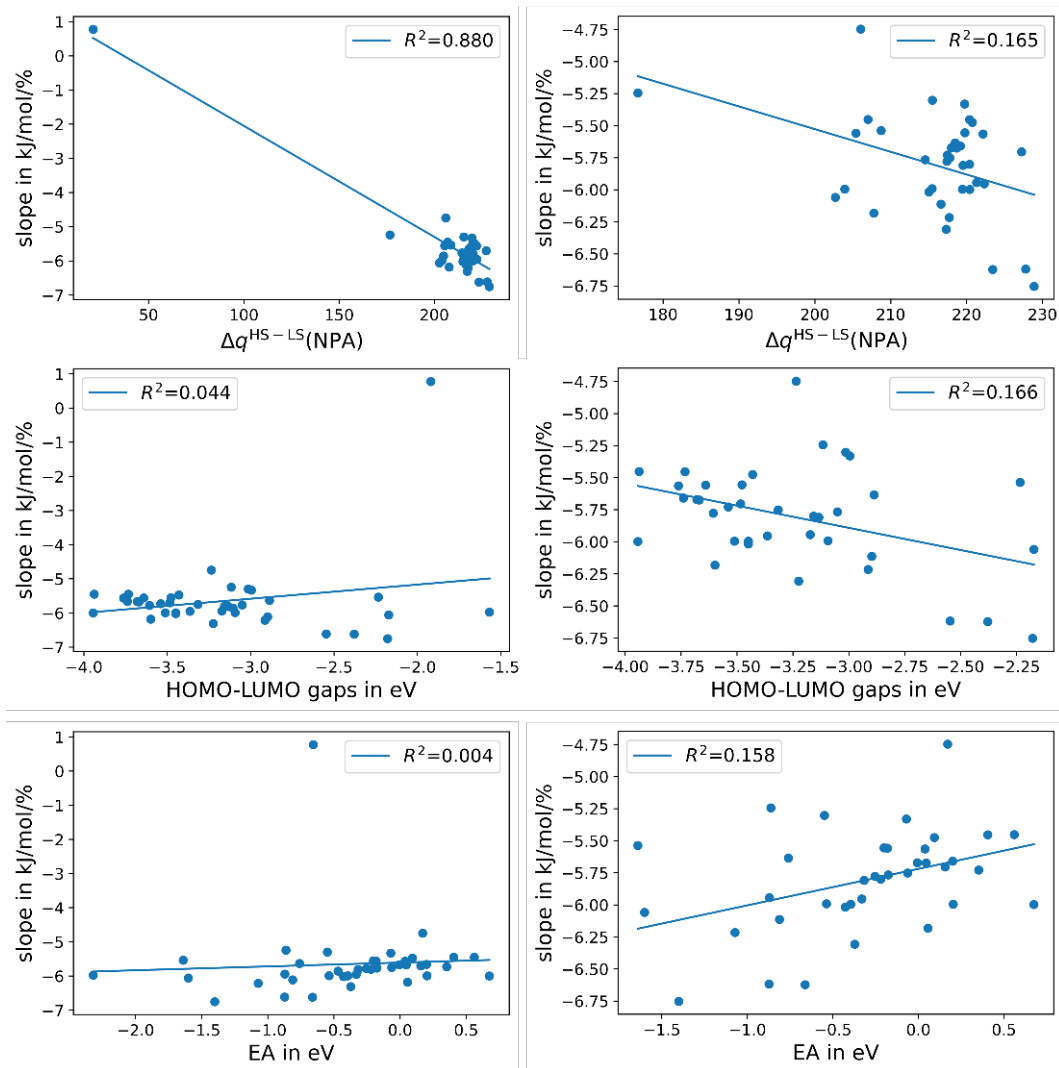


Figure B.13: Correlations between slope ($\frac{\partial E^{\text{HS-LS}}}{\partial a^{\text{HF}}}$), difference in local charges on iron centers ($q^{\text{HS-LS}}(\text{NPA})$) of BP86-D3BJ/def2-TZVP optimized Fe(II) complexes and HOMO-LUMO gaps as well as electron affinities (EA) of the free ligands including NENTABmod (left) and excluding NENTABmod (right). Ligands have been optimized with BP86-D3BJ/def2-TZVP in neutral (singlet) and anionic (doublet) charge. Reproduced from reference [71]

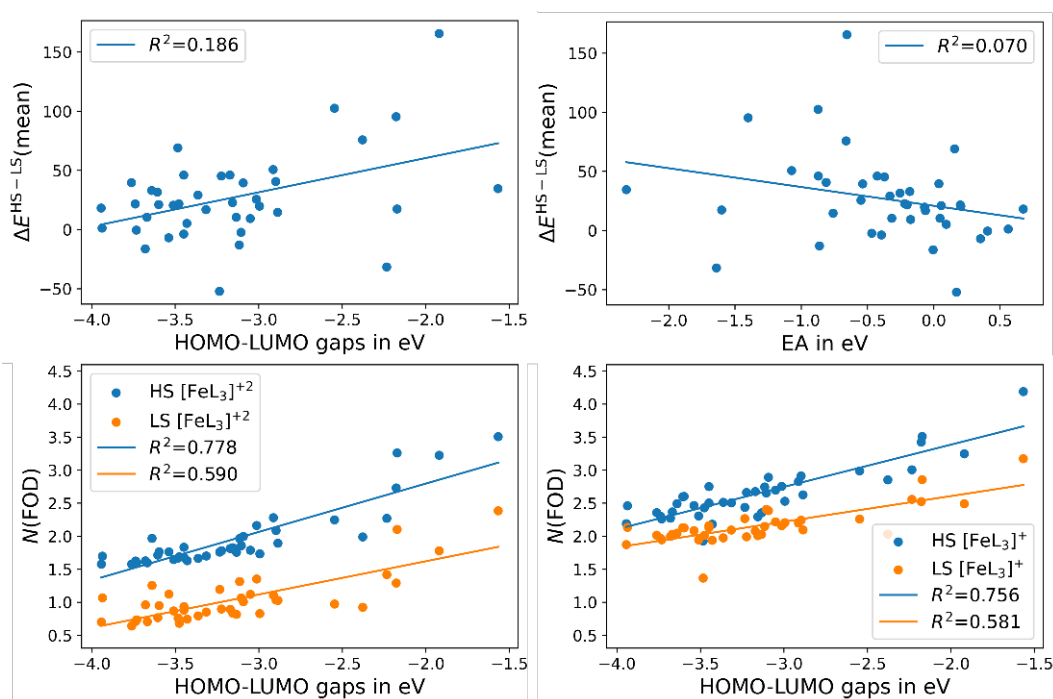


Figure B.14: Correlations between the mean spin-state energy splitting of all functionals and HOMO–LUMO gaps as well as electron affinities (EA) of the free ligands (top). Correlations between the number of strongly correlated electrons ($N(\text{FOD})$) and the HOMO–LUMO gaps of the free ligands (bottom). Ligands have been optimized with BP86-D3BJ/def2-TZVP in neutral (singlet) and anionic (doublet) charge. Reproduced from reference [71].

B.6 Characteristic Structural Features of Optimized Structures

All KS-DFT calculations have been performed with GAUSSIAN 16[99]. Geometry optimizations were performed with the BP86[100, 101] exchange-correlation functional, the D3 version of Grimme’s dispersion with Becke-Johnson damping[102] and the def2-TZVP[103, 104] basis set. Density fitting with the W06[103, 104] fitting set was applied. Optimization of the $[\text{L}_3\text{Fe}]^{2+}$ complexes using BP86 were performed on crystal structures or manually altered crystal structures. These structures were used for optimizations of the $[\text{L}_3\text{Fe}]^+$ complexes.

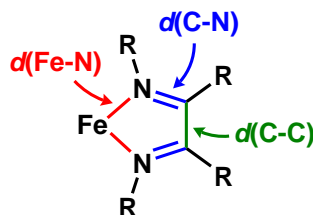


Figure B.15: Characteristic bond lengths. Reproduced from reference [71].

Table B.6: Comparison of crystal structures with trisdiimineiron(II) HS structures optimized with BP86-D3BJ/def2-TZVP (opt.-crystal). Bond lengths are given in Å.

Complex	$\Delta d(\text{Fe}-\text{N})$	$\Delta d(\text{C}-\text{N})$	$\Delta d(\text{C}-\text{C})$	$\Delta \frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})}$
BOFNEP	0.193213	0.031408	0.034026	0.000731
HAXHOF	0.164181	-0.004172	0.018208	-0.014335
NENTAB	0.185792	0.008418	0.020288	-0.006880
BACVOR	0.173849	0.005427	-0.000730	-0.003924
BACVOR B	-0.020691	0.008404	-0.031873	0.026632
AZUHUY	0.191366	0.001581	0.019629	-0.011478
KOQDEZ	0.181716	0.001775	0.014206	-0.008239
QAJKUH	0.191708	0.003528	0.014877	-0.007362
RUZNOP	0.186096	0.005106	0.017069	-0.007507
SIXJUE	0.206236	0.011612	-0.003330	0.009922
JOWGIL	0.179637	0.008710	0.012505	-0.001465
CATXAZ	0.174869	0.018347	0.000827	0.012163
VEWVEY	0.175048	-0.065655	-0.025863	-0.027809
KINQUT	0.186853	0.017106	0.051077	-0.021672
QOSNIW	-0.027676	0.012905	-0.009722	0.014253
EDUKUL	0.189234	-0.002676	0.012726	-0.010397
KABYER	0.082123	0.007935	-0.000431	0.005599
NOKFOJ	-0.039210	0.009946	-0.007526	0.011268
CATWIG	-0.027757	0.011426	-0.008089	0.012842
CATWUS	0.000514	0.015229	-0.001961	0.011690
CEPZOM	-0.021470	0.010407	0.027953	-0.011476
HUSMAK	-0.031664	0.011359	-0.005492	0.011101
TEDCAJ	0.152636	0.005745	0.027481	-0.012120

Table B.7: Comparison of crystal structures with trisdiimineiron(II) LS structures optimized with BP86-D3BJ/def2-TZVP (opt.-crystal). Bond lengths are given in Å.

Complex	$\Delta d(\text{Fe-N})$	$\Delta d(\text{C-N})$	$\Delta d(\text{C-C})$	$\Delta \frac{d(\text{N-C})}{d(\text{C-C})}$
BOFNPE	0.001136	0.034049	0.020789	0.010625
HAXHOF	-0.030455	0.004406	0.002898	0.001161
NENTAB	-0.012487	0.017281	0.004799	0.008720
BACVOR	-0.022858	0.005920	-0.013367	0.012643
AZUHUY	-0.009372	0.010035	0.005697	0.003176
KOQDEZ	-0.014344	0.007450	-0.002760	0.007110
QAJKUH	-0.010454	0.010983	0.003633	0.005014
RUZNOP	-0.017821	0.010293	0.007697	0.002111
SIXJUE	-0.002825	0.020559	-0.019097	0.025942
JOWGIL	-0.008674	0.019214	-0.008594	0.018202
CATXAZ	-0.022791	0.021884	-0.016454	0.025925
VEWVEY	-0.024913	-0.056243	-0.040648	-0.012075
KINQUT	-0.013671	0.022507	0.039732	-0.010674
QOSNIW	-0.207727	0.019884	-0.029486	0.030714
EDUKUL	-0.005349	0.003470	-0.003929	0.005116
KABYER	-0.109746	0.018720	-0.012584	0.020497
NOKFOJ	-0.118645	0.017672	-0.015483	0.021516
CATWIG	-0.218660	0.011067	-0.034840	0.029486
CATWUS	-0.203319	0.015572	-0.029153	0.029426
CEPZOM	-0.215963	0.017687	0.012283	0.004104
HUSMAK	-0.236500	0.022730	-0.021094	0.028735
TEDCAJ	-0.014584	0.013816	-0.000292	0.009512

Table B.8: Characteristic average distances in the trisdiimineiron(II) HS structures optimized with BP86-D3BJ/def2-TZVP. Bond lengths are given in Å.

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
BOFNPE	1	2.126	1.470	1.309	0.891
OKIWEKsimp	2	2.149	1.463	1.322	0.904
QOSNIWmod	3	2.140	1.499	1.302	0.869
WEYLERsimp	4	2.156	1.460	1.326	0.908
NENTABmod	5	1.945	1.433	1.371	0.957
NENTABmod2	5a	1.947	1.433	1.371	0.957
HEYRAEsimp	6	2.139	1.486	1.351	0.909
HAXHOF	7	2.148	1.478	1.360	0.920
NENTAB	8	2.156	1.489	1.359	0.913
BACVOR	9	2.159	1.468	1.357	0.924
BACVOR B	9	1.964	1.431	1.366	0.955
AZUHUY	10	2.153	1.470	1.359	0.925
KOQDEZ	11	2.157	1.436	1.364	0.950
QAJKUH	12	2.155	1.465	1.374	0.938
JUSCILsimp	13	2.149	1.459	1.346	0.923
RUZNOP	14	2.157	1.457	1.359	0.933
OJIVOSsimp	15	2.159	1.458	1.368	0.938
SIXJUE	16	2.133	1.476	1.354	0.917
JOWGIL	17	2.135	1.490	1.301	0.873
VEWVEY	18	2.186	1.478	1.360	0.920
CATXAZ	19	2.151	1.446	1.350	0.934
KINQUT	20	2.153	1.461	1.359	0.930

Continued on next page

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
KINQUTmod2	21	2.149	1.471	1.337	0.909
YIVSEBsimp	22	2.172	1.460	1.370	0.938
SAFGAHmod2	23	2.158	1.458	1.390	0.953
FUFHIYsimp	24	2.161	1.455	1.348	0.926
KINQUTmod	25	2.162	1.453	1.352	0.931
ABIWAKsimp	26	2.172	1.457	1.369	0.940
SAFGAHmod3	27	2.161	1.446	1.348	0.932
CATWIGsimp	28	2.153	1.467	1.308	0.891
BAKHECsimp	29	2.174	1.442	1.344	0.932
ACAHUJsimp	30	2.167	1.441	1.320	0.917
SAFGAHmod	31	2.160	1.441	1.339	0.929
QOSNIW	32	2.144	1.493	1.306	0.875
EDUKUL	33	2.152	1.434	1.364	0.951
KABYER	34	2.145	1.485	1.364	0.918
NOKFOJ	35	2.177	1.475	1.351	0.916
CATWIGmod	36	2.172	1.469	1.297	0.883
CATXAZsimp	37	2.171	1.438	1.344	0.934
CATWIG	38	2.201	1.449	1.302	0.899
CATWUS	39	2.218	1.455	1.336	0.919
CEPZOM	40	2.186	1.438	1.366	0.950
HUSMAK	41	2.154	1.475	1.358	0.921
CATWIGmod2	42	2.145	1.469	1.304	0.887
TEDCAJ	43	2.146	1.505	1.297	0.862

Table B.9: Characteristic average distances in the trisdiimineiron(II) LS structures optimized with BP86/def2-TZVP. Bond lengths are given in Å.

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
BOFNEP	1	1.934	1.457	1.312	0.901
OKIWEKsimp	2	1.941	1.445	1.332	0.922
QOSNIWmod	3	1.959	1.480	1.311	0.886
WEYLERsimp	4	1.955	1.444	1.336	0.926
NENTABmod	5	1.948	1.461	1.341	0.918
NENTABmod2	5a	1.948	1.461	1.341	0.918
HEYRAEsimp	6	1.928	1.471	1.361	0.925
HAXHOF	7	1.954	1.462	1.368	0.936
NENTAB	8	1.958	1.474	1.368	0.928
BACVOR	9	1.962	1.450	1.364	0.941
AZUHUY	10	1.952	1.456	1.367	0.939
KOQDEZ	11	1.961	1.419	1.370	0.965
QAJKUH	12	1.953	1.454	1.382	0.950
JUSCILsimp	13	1.945	1.446	1.355	0.937
RUZNOP	14	1.953	1.448	1.364	0.942
OJIVOSsimp	15	1.949	1.448	1.375	0.949
SIXJUE	16	1.924	1.460	1.363	0.933
JOWGIL	17	1.947	1.469	1.311	0.893
VEWVEY	18	1.986	1.463	1.370	0.936
CATXAZ	19	1.953	1.429	1.354	0.947
KINQUT	20	1.953	1.450	1.364	0.941
KINQUTmod2	21	1.953	1.454	1.344	0.924
YIVSEBsimp	22	1.963	1.448	1.376	0.950
SAFGAHmod2	23	1.951	1.449	1.397	0.964
FUFHIYsimp	24	1.949	1.442	1.355	0.940

Continued on next page

Complex	No.	$d(\text{Fe}-\text{N})$	$d(\text{C}-\text{C})$	$d(\text{C}-\text{N})$	$\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})}$
KINQUTmod	25	1.957	1.440	1.360	0.944
ABIWAKsimp	26	1.962	1.446	1.375	0.951
SAFGAHmod3	27	1.954	1.438	1.354	0.941
CATWIGsimp	28	1.952	1.452	1.315	0.906
BAKHECsimp	29	1.959	1.433	1.351	0.943
ACAHUJsimp	30	1.957	1.427	1.329	0.931
SAFGAHmod	31	1.955	1.431	1.348	0.943
QOSNIW	32	1.964	1.473	1.313	0.891
EDUKUL	33	1.958	1.417	1.370	0.967
KABYER	34	1.954	1.473	1.375	0.933
NOKFOJ	35	2.098	1.467	1.359	0.926
CATWIGmod	36	1.963	1.453	1.305	0.898
CATXAZsimp	37	1.959	1.428	1.351	0.946
CATWIG	38	2.010	1.422	1.301	0.915
CATWUS	39	2.014	1.428	1.337	0.936
CEPZOM	40	1.992	1.422	1.373	0.965
HUSMAK	41	1.949	1.459	1.369	0.938
CATWIGmod2	42	1.948	1.453	1.313	0.904
TEDCAJ	43	1.979	1.478	1.305	0.883

Table B.10: Characteristic average distances in the trisdiimineiron(I) HS structures (quartet) optimized with BP86/def2-TZVP. ^a Manipulated crystal structure or crystal structure was used as input structure for the geometry optimization. Bond lengths are given in Å.

Complex	No.	$d(\text{Fe}-\text{N})$	$d(\text{C}-\text{C})$	$d(\text{C}-\text{N})$	$\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})}$
BOFNEP	1	2.103	1.445	1.323	0.915
OKIWEKsimp	2	2.130	1.447	1.332	0.921
QOSNIWmod	3	2.135	1.472	1.310	0.890
WEYLERsimp	4	2.132	1.445	1.336	0.924
NENTABmod	5	1.941	1.430	1.365	0.955
NENTABmod ^a	-	1.943	1.429	1.365	0.955
NENTABmod2	5a	1.940	1.431	1.365	0.955
HEYRAEsimp	6	2.125	1.474	1.360	0.923
HAXHOF	7	2.131	1.466	1.366	0.932
NENTAB	8	2.131	1.474	1.367	0.927
BACVOR	9	1.958	1.430	1.366	0.955
BACVOR	-	1.958	1.430	1.366	0.955
AZUHUY	10	2.130	1.460	1.364	0.934
KOQDEZ	11	2.138	1.429	1.369	0.958
QAJKUH	12	2.135	1.460	1.378	0.944
JUSCILsimp	13	2.131	1.444	1.354	0.938
RUZNOP	14	2.130	1.450	1.363	0.940
OJIVOSsimp	15	2.139	1.450	1.373	0.947
SIXJUE	16	2.127	1.462	1.363	0.932
JOWGIL	17	2.129	1.458	1.317	0.903
VEWVEY	18	2.164	1.467	1.366	0.931
CATXAZ	19	2.139	1.432	1.356	0.947
KINQUT	20	2.130	1.451	1.364	0.940
KINQUTmod2	21	2.130	1.451	1.343	0.925
YIVSEBsimp	22	2.153	1.452	1.375	0.947
SAFGAHmod2	23	2.141	1.456	1.390	0.954
FUFHIYsimp	24	2.139	1.441	1.355	0.940

Continued on next page

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
KINQUTmod	25	2.142	1.442	1.358	0.942
ABIWAKsimp	26	2.150	1.451	1.374	0.947
SAFGAHmod3	27	2.141	1.434	1.353	0.943
CATWIGsimp	28	2.136	1.440	1.316	0.914
BAKHECsimp	29	2.149	1.435	1.352	0.942
ACAHUJsimp	30	2.140	1.429	1.329	0.931
SAFGAHmod	31	2.141	1.426	1.345	0.943
QOSNIW	32	2.133	1.464	1.314	0.897
EDUKUL	33	2.135	1.426	1.369	0.960
KABYER	34	2.137	1.477	1.370	0.927
NOKFOJ	35	2.154	1.463	1.357	0.927
CATWIGmod	36	2.144	1.440	1.308	0.908
CATXAZsimp	37	2.159	1.424	1.350	0.948
CATWIG	38	2.196	1.423	1.308	0.919
CATWUS	39	2.207	1.454	1.335	0.918
CEPZOM	40	2.169	1.432	1.370	0.957
HUSMAK	41	2.134	1.462	1.365	0.934
CATWIGmod2	42	2.132	1.444	1.312	0.908
TEDCAJ	43	2.141	1.480	1.306	0.882

Table B.11: Characteristic average distances in the trisdiimineiron(I) LS structures optimized with BP86/def2-TZVP. Bond lengths are given in Å.

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
BOFNPE	1	1.933	1.436	1.328	0.925
OKIWEKsimp	2	1.935	1.426	1.347	0.944
QOSNIWmod	3	1.948	1.452	1.323	0.911
WEYLERsimp	4	1.947	1.430	1.348	0.943
NENTABmod	5	1.940	1.442	1.352	0.938
NENTABmod2	5a	1.942	1.441	1.351	0.938
HEYRAEsimp	6	1.919	1.452	1.371	0.944
HAXHOF	7	1.946	1.445	1.378	0.953
NENTAB	8	1.945	1.454	1.378	0.947
BACVOR	9	1.954	1.441	1.365	0.947
AZUHUY	10	1.944	1.442	1.376	0.954
KOQDEZ	11	1.955	1.410	1.378	0.977
QAJKUH	12	1.943	1.442	1.385	0.961
JUSCILsimp	13	1.938	1.430	1.365	0.955
RUZNOP	14	1.942	1.438	1.372	0.954
OJIVOSsimp	15	1.943	1.434	1.386	0.967
SIXJUE	16	1.920	1.442	1.375	0.953
JOWGIL	17	1.940	1.445	1.328	0.919
VEWVEY	18	1.975	1.447	1.379	0.954
CATXAZ	19	1.948	1.416	1.363	0.963
KINQUT	20	1.942	1.437	1.375	0.957
KINQUTmod2	21	1.945	1.433	1.355	0.945
YIVSEBsimp	22	1.951	1.435	1.385	0.965
SAFGAHmod2	23	1.940	1.445	1.392	0.963
FUFHIYsimp	24	1.944	1.425	1.366	0.958
KINQUTmod	25	1.950	1.426	1.370	0.961
ABIWAKsimp	26	1.949	1.436	1.386	0.965
SAFGAHmod3	27	1.944	1.425	1.363	0.956
CATWIGsimp	28	1.942	1.424	1.329	0.933

Continued on next page

Complex	No.	$d(\text{Fe}-\text{N})$	$d(\text{C}-\text{C})$	$d(\text{C}-\text{N})$	$\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})}$
BAKHECsimp	29	1.950	1.425	1.365	0.958
ACAHUJsimp	30	1.949	1.415	1.342	0.949
SAFGAHmod	31	1.947	1.417	1.357	0.958
QOSNIW	32	1.958	1.443	1.326	0.919
EDUKUL	33	1.952	1.407	1.378	0.979
KABYER	34	1.950	1.460	1.383	0.947
NOKFOJ	35	2.129	1.460	1.363	0.934
CATWIGmod	36	1.949	1.424	1.321	0.928
CATXAZsimp	37	1.952	1.413	1.361	0.963
CATWIG	38	2.000	1.399	1.312	0.938
CATWUS	39	2.004	1.418	1.338	0.943
CEPZOM	40	1.978	1.424	1.379	0.968
HUSMAK	41	1.941	1.442	1.380	0.957
CATWIGmod2	42	1.939	1.427	1.327	0.930
TEDCAJ	43	1.975	1.458	1.317	0.903

Table B.12: Differences in bond lengths and bond length ratios $\Delta(\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})})^{\text{HS-LS}}$ (HS-LS) for Fe(II) and Fe(I) complexes. Bond lengths are given in Å.

No.	Fe(II)		Fe(I)	
	$\Delta d^{\text{HS-LS}}(\text{Fe}-\text{N})$	$\Delta(\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})})^{\text{HS-LS}}$	$\Delta d^{\text{HS-LS}}(\text{Fe}-\text{N})$	$\Delta(\frac{d(\text{N}-\text{C})}{d(\text{C}-\text{C})})^{\text{HS-LS}}$
1	0.192	-0.010	0.170	-0.010
2	0.208	-0.018	0.195	-0.023
3	0.181	-0.017	0.186	-0.021
4	0.201	-0.017	0.185	-0.018
5	-0.003	0.039	0.001	0.017
5a	-0.001	0.039	-0.002	0.017
6	0.211	-0.016	0.207	-0.021
7	0.195	-0.015	0.185	-0.022
8	0.198	-0.016	0.186	-0.020
9	0.197	-0.017	0.004	0.008
10	0.201	-0.015	0.186	-0.020
11	0.196	-0.015	0.183	-0.019
12	0.202	-0.012	0.191	-0.017
13	0.204	-0.014	0.192	-0.017
14	0.204	-0.010	0.188	-0.014
15	0.210	-0.011	0.197	-0.020
16	0.209	-0.016	0.207	-0.021
17	0.188	-0.020	0.189	-0.016
18	0.200	-0.016	0.190	-0.022
19	0.198	-0.014	0.191	-0.016
20	0.201	-0.011	0.188	-0.017
21	0.196	-0.015	0.186	-0.020
22	0.208	-0.011	0.202	-0.018
23	0.207	-0.011	0.201	-0.009
24	0.211	-0.013	0.195	-0.018
25	0.204	-0.014	0.192	-0.019
26	0.210	-0.011	0.201	-0.018
27	0.206	-0.009	0.197	-0.013
28	0.201	-0.014	0.194	-0.019
29	0.215	-0.011	0.198	-0.015
30	0.210	-0.015	0.191	-0.018

Continued on next page

No.	Fe(II)		Fe(II)	
	$\Delta d^{\text{HS-LS}}(\text{Fe-N})$	$\Delta(\frac{d(\text{N-C})}{d(\text{C-C})})^{\text{HS-LS}}$	$\Delta d^{\text{HS-LS}}(\text{Fe-N})$	$\Delta(\frac{d(\text{N-C})}{d(\text{C-C})})^{\text{HS-LS}}$
31	0.205	-0.013	0.194	-0.015
32	0.180	-0.016	0.175	-0.022
33	0.195	-0.016	0.184	-0.020
34	0.192	-0.015	0.187	-0.020
35	0.079	-0.010	0.025	-0.007
36	0.209	-0.015	0.194	-0.019
37	0.213	-0.011	0.208	-0.015
38	0.191	-0.017	0.196	-0.019
39	0.204	-0.018	0.203	-0.025
40	0.194	-0.016	0.191	-0.011
41	0.205	-0.018	0.193	-0.023
42	0.197	-0.016	0.193	-0.022
43	0.167	-0.022	0.166	-0.021

Table B.13: Characteristic bond pattern of BP86/def2-TZVP optimized $[(\alpha\text{-diimine})_3\text{Fe}]^+$ complexes in sextet spin state. Bond lengths are given in Å.

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
BOFNEP	1	2.133759	1.447710	1.324909	0.915478
OKIWEKsimp	2	2.134664	1.443610	1.336479	0.925821
QOSNIWmod	3	2.138215	1.469664	1.314651	0.894588
WEYLERsimp	4	2.137989	1.441974	1.340138	0.929389
NENTABmod	5	1.942462	1.417395	1.379584	0.973323
NENTABmod2	5a	1.944745	1.417489	1.380324	0.973781
HEYRAEsimp	6	2.129923	1.470629	1.364335	0.927728
FUFHIYsimp	7	2.151677	1.437329	1.359251	0.945694
NENTAB	8	2.140825	1.467042	1.371502	0.934876
AZUHUY	10	2.138326	1.457150	1.367505	0.938480
KOQDEZ	11	2.146017	1.426536	1.372216	0.961922
QAJKUH	12	2.142202	1.451579	1.379648	0.950447
JUSCILsimp	13	2.140454	1.440744	1.357866	0.942476
RUZNOP	14	2.140721	1.448898	1.367185	0.943604
OJIVOSsimp	15	2.148641	1.445515	1.379126	0.954082
SIXJUE	16	2.129991	1.460771	1.365586	0.934845
JOWGIL	17	2.132119	1.462445	1.318971	0.901950
VEWVEY	18	2.174404	1.461189	1.370795	0.938137
CATWUS	19	2.201989	1.443739	1.338442	0.927068
KINQUT	20	2.139158	1.447281	1.370000	0.946617
KINQUTmod2	21	2.135750	1.447612	1.347819	0.931067
YIVSEBsimp	22	2.156982	1.445560	1.380764	0.955176
SAFGAHmod2	23	2.141881	1.452937	1.384605	0.952970
EDUKUL	24	2.142436	1.423513	1.372128	0.963903
KINQUTmod	25	2.151006	1.438700	1.363042	0.947413
ABIWAKsimp	26	2.154452	1.447537	1.379746	0.953168
SAFGAHmod3	27	2.148698	1.431737	1.358337	0.948735
CATWIGsimp	28	2.139513	1.438060	1.322568	0.919745
BAKHECsimp	29	2.155264	1.431933	1.357891	0.948299
ACAHUJsimp	30	2.150041	1.428096	1.334000	0.934146
SAFGAHmod	31	2.149513	1.424741	1.348835	0.946723
QOSNIW	32	2.137361	1.462568	1.318850	0.901754
CEPZOM	33	2.175266	1.428553	1.373416	0.961405

Continued on next page

Complex	No.	$d(\text{Fe-N})$	$d(\text{C-C})$	$d(\text{C-N})$	$\frac{d(\text{N-C})}{d(\text{C-C})}$
KABYER	34	2.140217	1.473610	1.374354	0.932650
CATWIGmod	36	2.149877	1.437687	1.314417	0.914318
CATXAZ	37	2.149287	1.430378	1.361558	0.951896
CATWIG	38	2.207819	1.428090	1.313869	0.920041
CATWOM	39	2.142425	1.423511	1.372133	0.963908
CATXAZsimp	40	2.161913	1.422505	1.354245	0.952015
HUSMAK	41	2.144077	1.457142	1.368629	0.939284
CATWIGmod2	42	2.138882	1.442637	1.316576	0.912626
TEDCAJ	43	2.146400	1.478141	1.312177	0.887723

B.7 Zero-Point Vibrational Energies and Dispersion Corrections

Table B.14: Zero-point vibrational energies (ZPVE), thermal corrections at $T = 298$ K and dispersion (D3BJ) corrections to Fe(II) complexes calculated with BP86/def2-TZVP. All energies are given in kJ mol^{-1} . Thermal and entropic corrections have been determined by vibrational analysis using the harmonic approximation.

No.	ΔZPVE	$\Delta\Delta H^{\text{HS-LS}}$	$T\Delta\Delta S^{\text{HS-LS}}$	$\Delta\Delta G^{\text{HS-LS}}$	$\Delta E_{\text{disp}}^{\text{corr}}$
1	-9.08	-4.42	21.57	-25.99	3.57
2	-16.06	-8.30	28.27	-36.57	13.08
3	-8.72	-5.03	13.48	-18.51	0.93
4	-14.73	-7.36	27.36	-34.72	13.24
5a	-7.93	-5.95	9.62	-15.56	1.40
5	-9.65	-7.35	10.74	-18.09	-2.83
6	-15.36	-7.43	33.80	-41.23	13.18
7	-11.94	-5.88	26.18	-32.06	13.93
8	-11.83	-5.55	26.20	-31.75	13.82
9	-10.82	-5.33	24.87	-30.20	13.81
10	-12.68	-6.26	27.49	-33.76	14.47
11	-11.30	-5.74	24.24	-29.97	13.06
12	-11.06	-5.11	26.78	-31.88	15.56
13	-12.43	-6.26	27.90	-34.16	13.07
14	-10.41	-4.61	25.70	-30.31	11.92
15	-11.64	-5.49	26.70	-32.19	12.17
16	-14.68	-7.23	34.87	-42.10	14.21
17	-15.89	-9.54	23.13	-32.67	12.28
18	-13.17	-6.04	31.04	-37.09	14.18
19	-8.83	-4.02	24.39	-28.41	11.19
20	-10.28	-4.49	26.35	-30.84	12.33
21	-12.68	-5.81	29.61	-35.42	14.11
22	-9.90	-4.23	24.76	-28.99	13.58
23	-10.99	-5.34	26.27	-31.61	14.97
24	-11.16	-5.24	25.97	-31.21	10.82
25	-10.53	-4.82	25.25	-30.08	13.34
26	-10.55	-4.64	25.64	-30.28	13.50
27	-10.45	-4.64	27.28	-31.92	16.54
28	-12.06	-5.47	33.19	-38.67	14.29
29	-13.19	-6.08	27.44	-33.52	13.59
30	-13.52	-6.50	26.96	-33.46	12.68

Continued on next page

B.7 Zero-Point Vibrational Energies and Dispersion Corrections

No.	$\Delta ZPVE$	$\Delta\Delta H^{\text{HS-LS}}$	$T\Delta\Delta S^{\text{HS-LS}}$	$\Delta\Delta G^{\text{HS-LS}}$	$\Delta E_{\text{disp.}}^{\text{corr.}}$
31	-9.56	-4.09	25.31	-29.40	12.44
32	-9.44	-4.69	21.32	-26.01	10.42
33	-10.15	-5.20	22.62	-27.82	12.83
34	-11.58	-5.22	24.73	-29.94	-1.88
35	-15.24	-11.51	14.91	-26.42	8.48
36	-9.18	-4.17	21.70	-25.88	14.60
37	-9.75	-4.13	25.76	-29.89	12.82
38	-7.28	-2.65	22.22	-24.87	7.20
39	-8.01	-3.36	22.69	-26.06	9.20
40	-12.59	-6.02	27.13	-33.15	13.70
41	-9.74	-4.42	21.47	-25.88	11.29
42	-12.68	-6.60	29.82	-36.43	16.84
43	-13.06	-7.37	28.36	-35.73	13.03

C Supporting Information for Machine Learning Guided Strategies Towards Spin Gaps in Carbenes

Listing C.1: Machine Learning Models

```
1  from sklearn.linear_model import LinearRegression
3  from sklearn.neighbors import KNeighborsRegressor
4  from sklearn.kernel_ridge import KernelRidge
5  from sklearn.gaussian_process import GaussianProcessRegressor
6  from sklearn.gaussian_process.kernels import Matern
7
8  #kernels
9  matern_01 = 1.0 * Matern(1.0, nu = 0.5) # Matern 01 kernel
10 matern_02 = Matern(1.0, nu = 0.5) # Matern 02 kernel
11 rbf_01 = RBF(1.0) # RBF 01 kernel
12 rbf_02 = 1.0*RBF(1.0) # RBF 02 kernel
13
14 #GPR
15 gpr_matern_01 = GaussianProcessRegressor(kernel = matern_01,
16                                         n_restarts_optimizer = 10)
17 gpr_matern_02 = GaussianProcessRegressor(kernel = matern_02,
18                                         n_restarts_optimizer = 10)
19 gpr_rbf_01 = GaussianProcessRegressor(kernel = rbf_01,
20                                     n_restarts_optimizer = 10)
21 gpr_rbf_02 = GaussianProcessRegressor(kernel = rbf_02,
22                                     n_restarts_optimizer = 10)
23
24 #KRR
25 krr_matern_01 = KernelRidge(kernel = matern_01)
26 krr_rbf_01 = KernelRidge(kernel = rbf_01)
27
28 #k-NN
29 knn_01 = KNeighborsRegressor(weights='uniform', n_neighbors=5)
30 knn_02 = KNeighborsRegressor(weights='distance', n_neighbors=5)
31
32 #linear regression
33 lin = LinearRegression()
```

Listing C.2: Descriptors for ML

```
2  from dscribe.descriptors import ACSF, SOAP, LMBTR
3
4  '''
5  variables
6  mol: ase.Atoms object
7  carbene_position: position of carbene carbon
8  '''
9
10 species = [H,C,F,N,O]
```

```

12 ##### ACSF #####
   acsf = ACSF(species = species,
14             rcut = 6.0,
             g2_params = [[1, 1], [1, 2], [1, 3]],
16             g3_params = None,
             g4_params = [[1, 1, 1], [1, 2, 1], [1, 1, -1], [1, 2, -1]],
18             g5_params = None)

20 features_acsf = acsf.create(system = mol, centers = [carbene_position])

22 ##### SOAP #####
   soap = SOAP(species = species, periodic = False, rcut = 14, nmax = 8,
               lmax = 4, average = 'outer', rbf = 'gto')
24
   features_soap = soap.create(system = mol, centers = [carbene_position])
26
28 ##### LMBTR #####
   lmbtr = LMBTR(species = species, periodic = False,
                 k2 = {"geometry": {"function": "inverse_distance"},
30                     "grid": {"min": 0.1, "max": 2, "sigma": 0.1, "n":
                               50},
                     "weighting": {"function": "exp", "scale": 0.75, "
                                   threshold": 1e-2}},
                 k3 = {"geometry": {"function": "angle"},
32                     "grid": {"min": 0, "max": 180, "sigma": 5, "n": 50},
                     "weighting": {"function": "exp", "scale": 0.5, "
34                                   threshold": 1e-3}},
                 normalize_gaussians = normalize_gaussians,
36                 normalization = 'none', flatten = True)

38 features_lmbtr = lmbtr.create(system = mol, centers = [carbene_position])

```

C.1 Additional Data on Computing Times

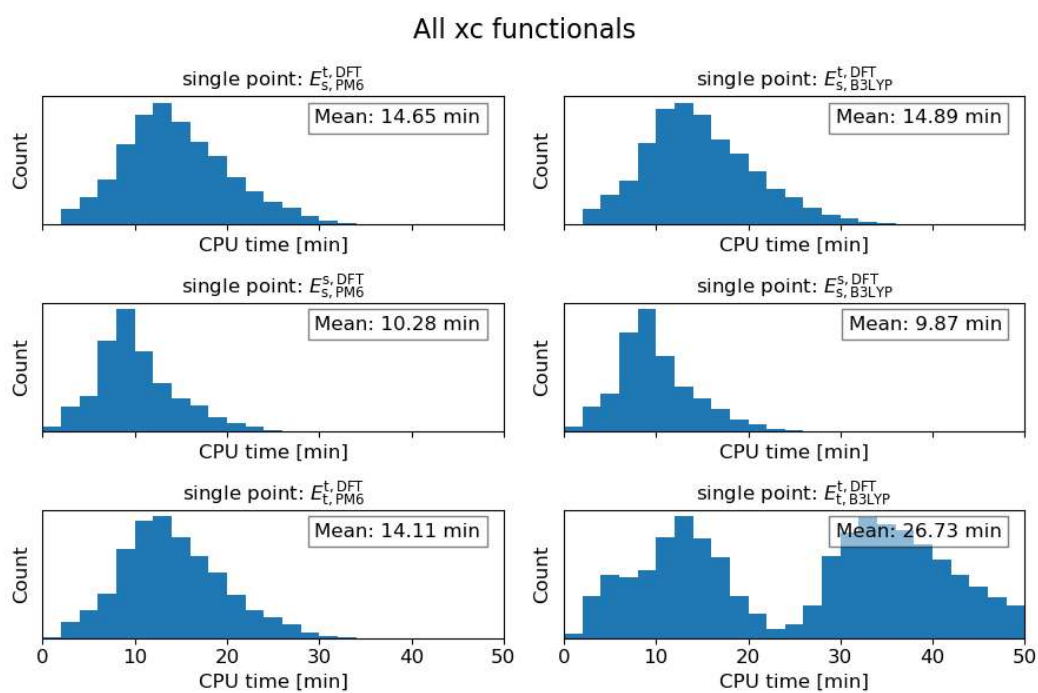


Figure C.1: Distribution of CPU times for DFT single point calculations for 11 xc functionals.

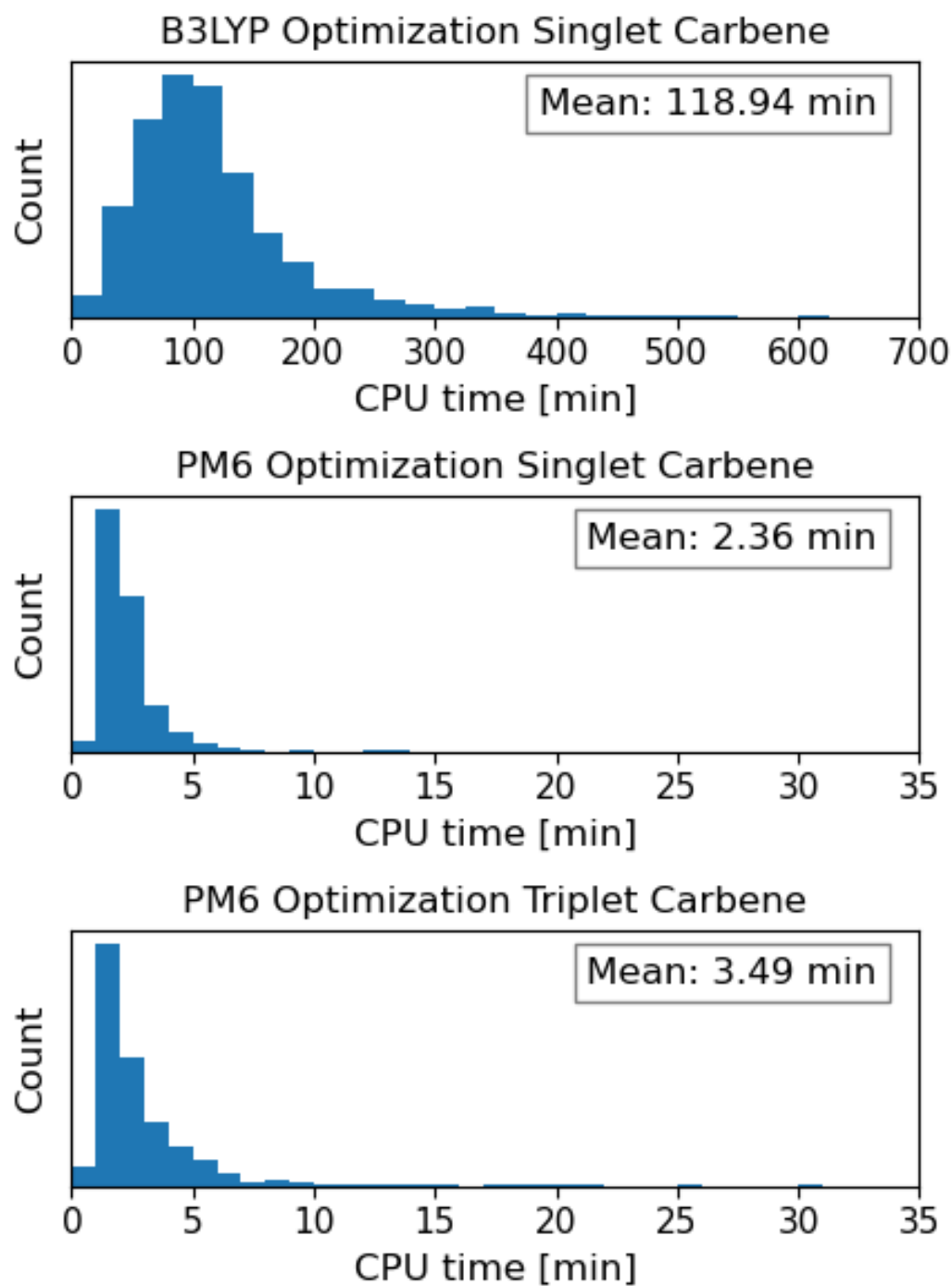


Figure C.2: Distribution of CPU times for the molecular structure .optimization of carbenes.

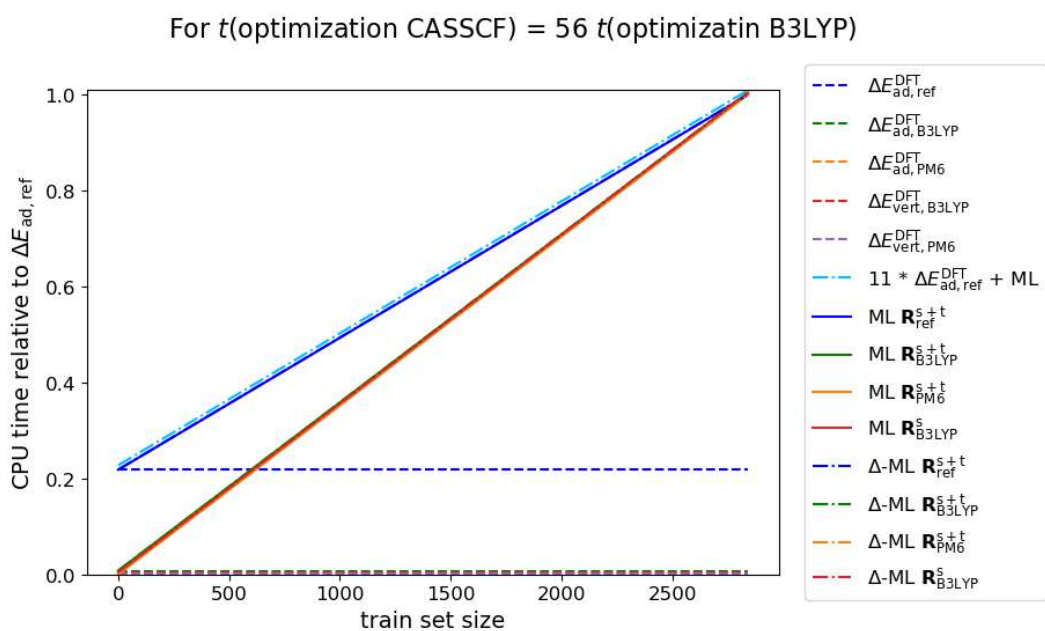


Figure C.3: Distribution of CPU times for DFT single point calculations for 11 xc functionals.

C.2 Additional Data for DFT Calculations

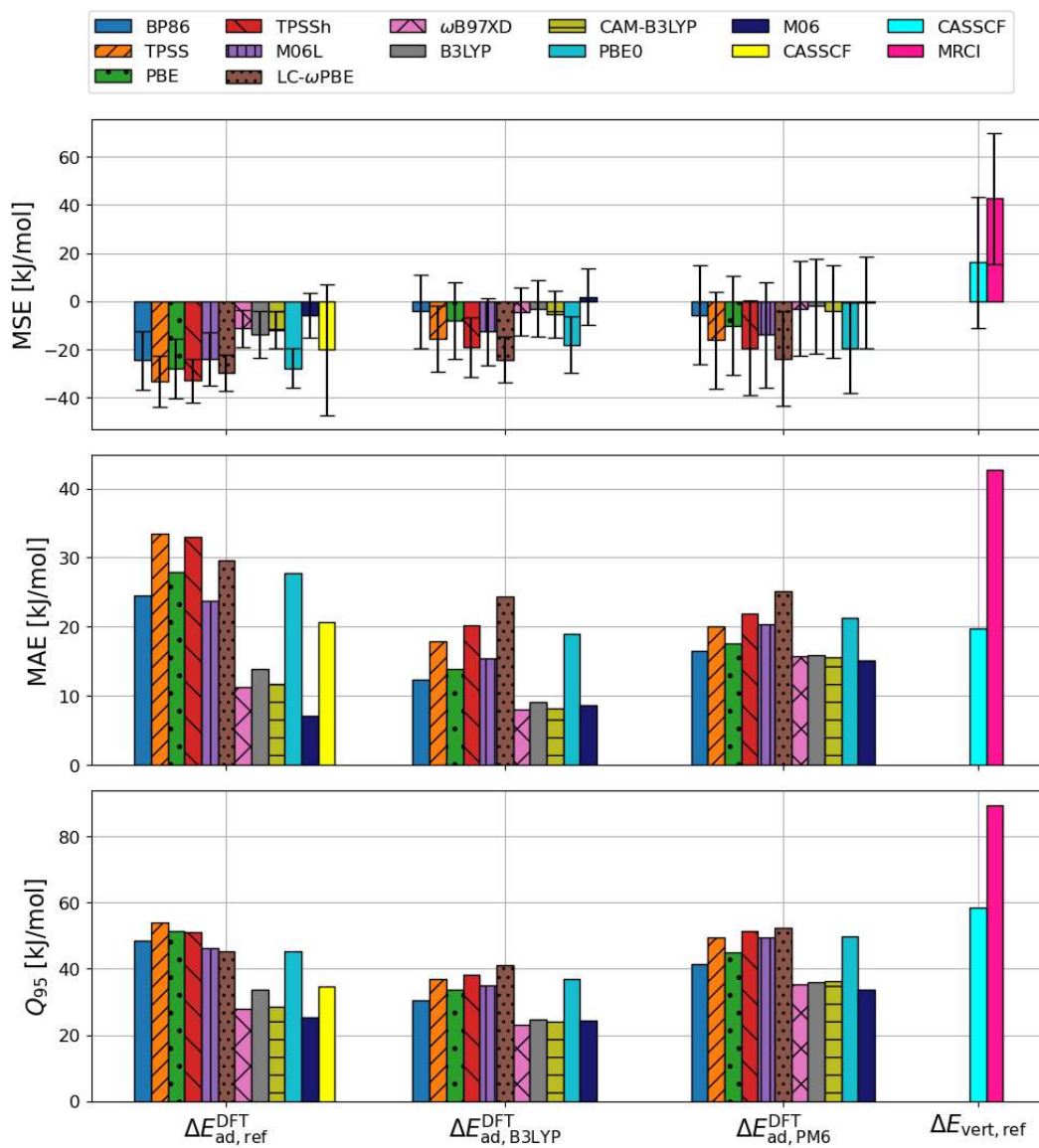


Figure C.4: Errors are given relative to the MRCI spin gaps ($\Delta E^{\text{DFT}} - \Delta E_{\text{ad,ref}}$).

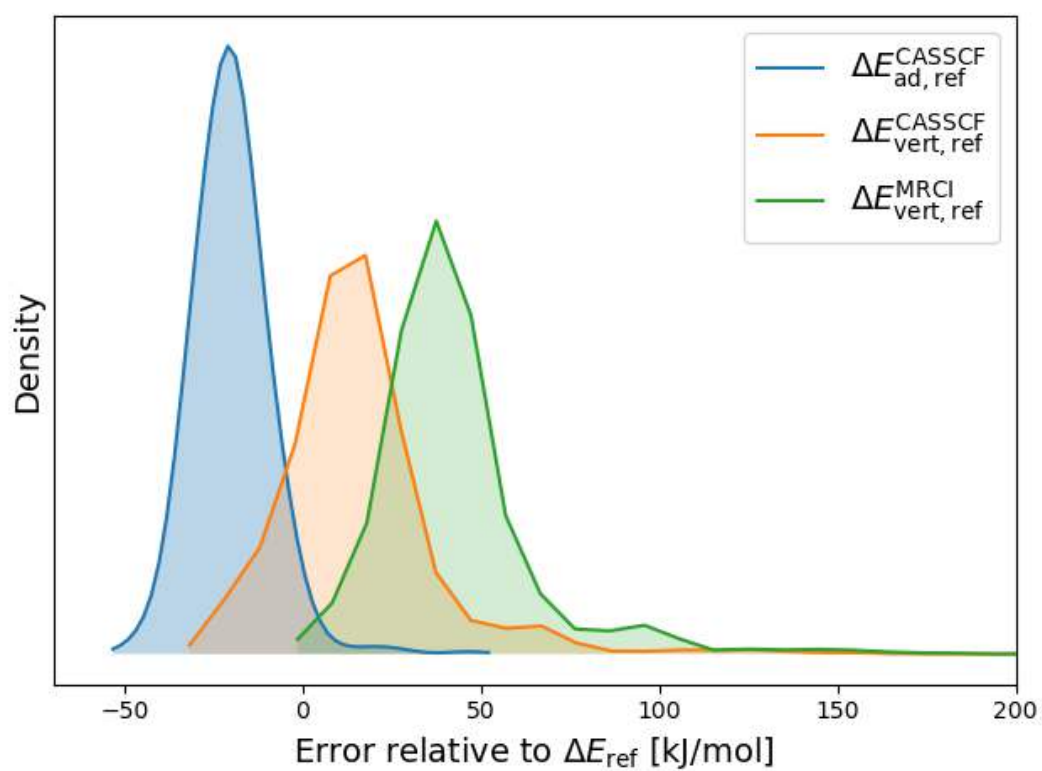


Figure C.5: Kernel density estimate of adiabatic and vertical spin gap errors (kernel='gaussian', width=0.5). Vertical spin gaps are calculated on CASSCF-optimized singlet structures.[1]

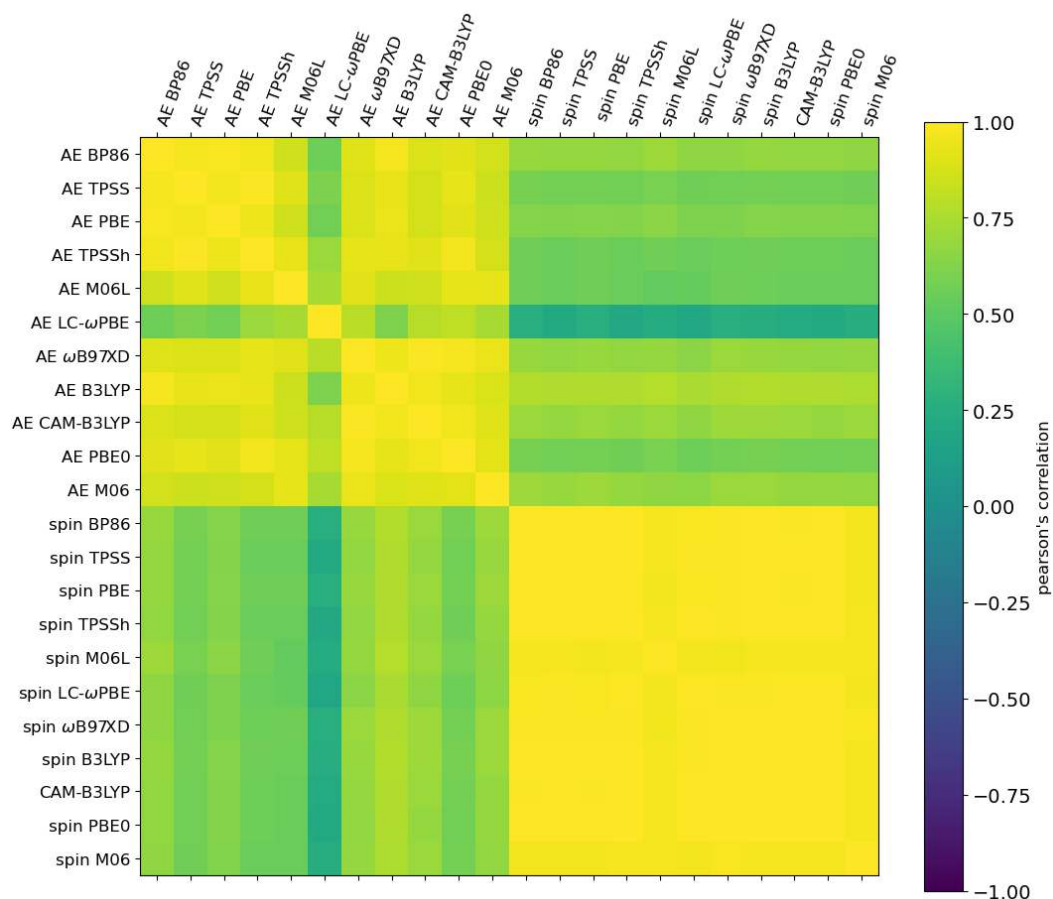


Figure C.6: Correlation matrix for all absolute errors (AE) of $\Delta E^{DFT} - \Delta E_{ad,ref}$ calculated with 11 different xc functionals and number of unpaired electrons localized on the carbene carbon for each of the triplet carbenes $n_C^{\alpha-\beta}$.

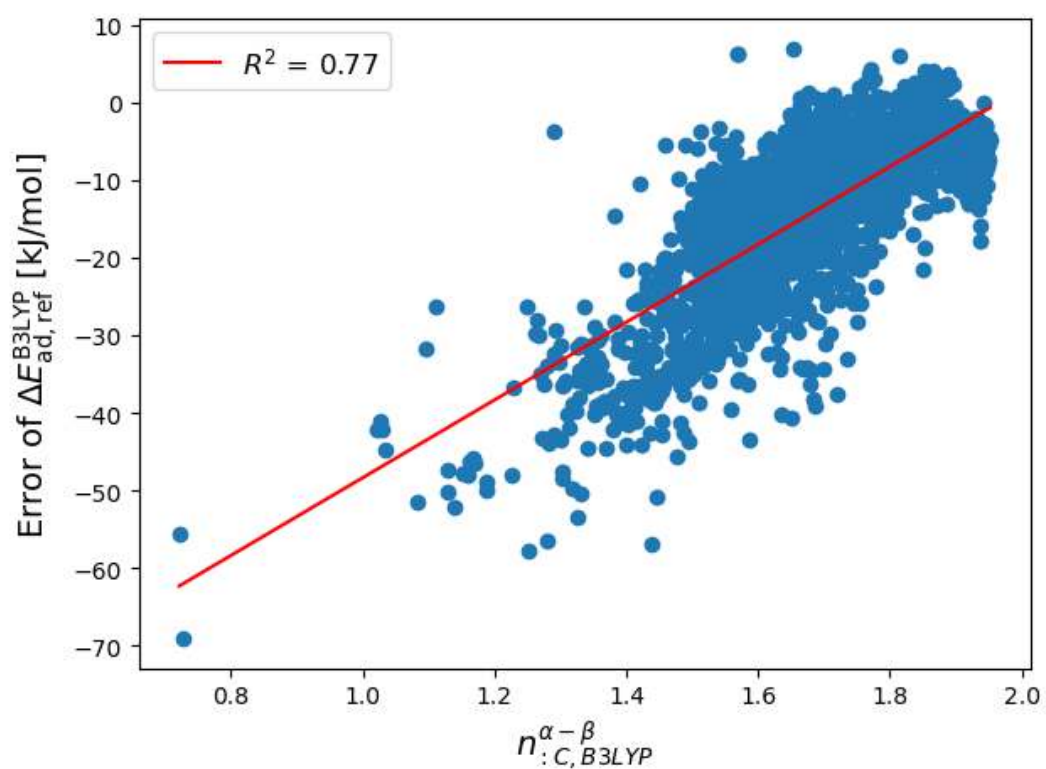


Figure C.7: Fit for all absolute errors (AE) of $\Delta E^{B3LYP} - \Delta E_{ad,ref}$ number of unpaired electrons localized on the carbene carbon for each of the triplet carbenes $n_{:C,B3LYP}^{\alpha-\beta}$.

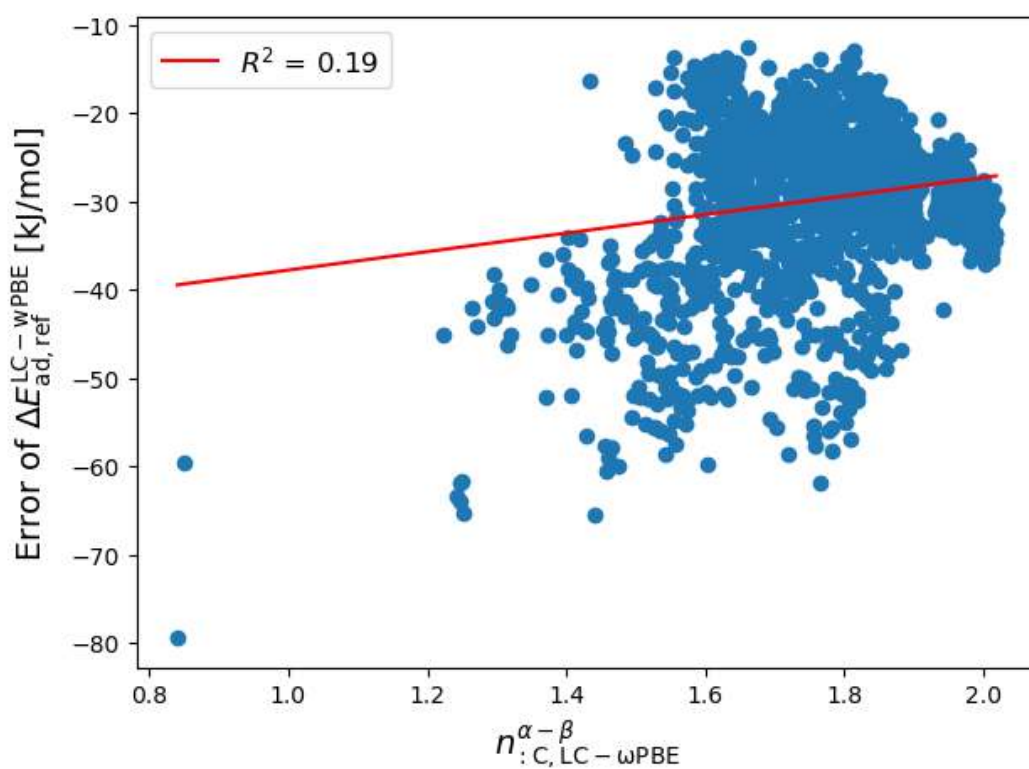


Figure C.8: Fit for all absolute errors (AE) of $\Delta E^{\text{LC}\omega\text{-PBE}} - \Delta E_{\text{ad,ref}}$ number of unpaired electrons localized on the carbene carbon for each of the triplet carbenes $n^{\alpha-\beta}_{\text{:C,LC}\omega\text{-PBE}}$.

C.3 Additional Data on ML

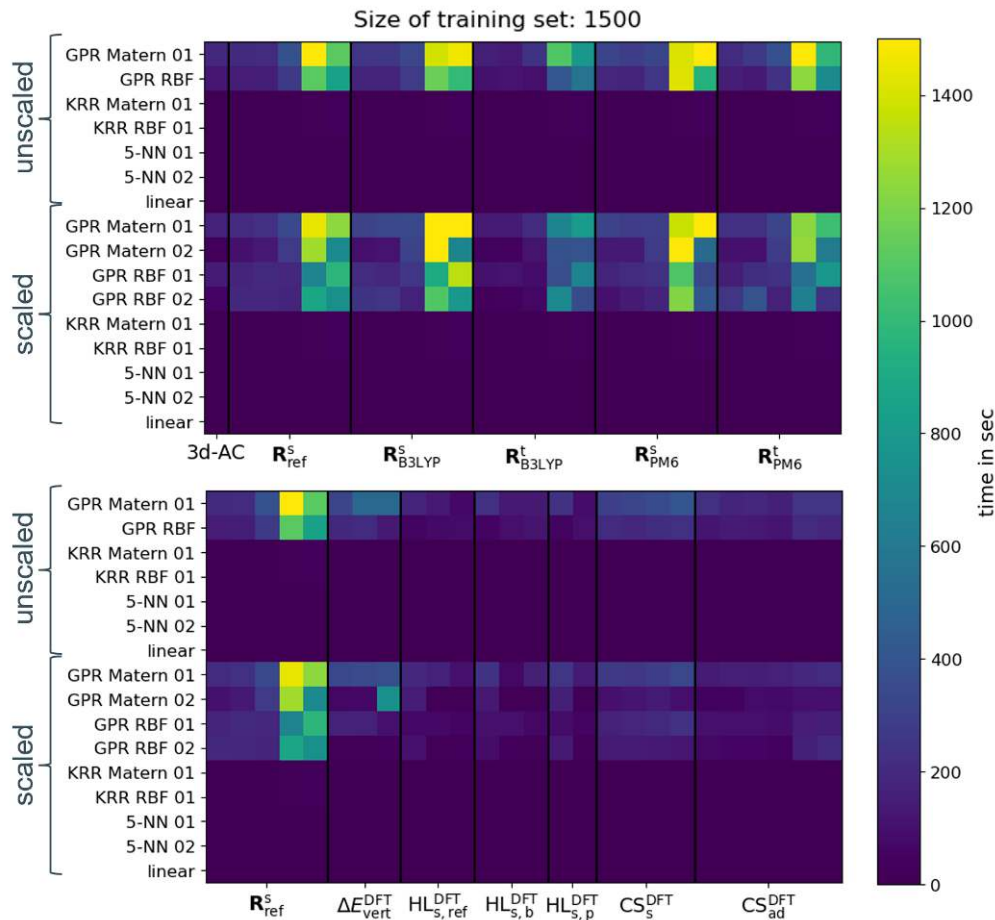


Figure C.9: Overview of CPU time for supervised machine learning with different models (y-axis) and descriptors (x-axis). The train-test split has been performed randomly 10 times with training set size of 1500. Models have been used with unscaled and scaled as indicated by the brackets. Descriptors grouped under the **R** label employ only singlet or singlet and triplet molecular structures indicated by superscribed “s” or “s+t”, respectively. The method used for molecular structure optimization is denoted with subscripts. Within the groups, descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the “s+t” groups. For the group ΔE^{DFT} , DFT spin gaps calculated with all 11 xc functionals are used as features. The group contains $\Delta E_{\text{ad,ref}}^{\text{DFT}}$, $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$, $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$, $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$. For the HL^{DFT} group HOMO–LUMO, gaps calculated for the singlet carbenes with the 11 xc functionals are used as features. CASSCF, B3LYP- and PM6-optimized singlet structures are used to calculate the HOMO–LUMO gaps. The CS^{M06} descriptor contains Mulliken charges on the carbene carbon and the two neighboring atoms for the singlet and triplet carbene as well as the Mulliken number of unpaired electrons on those three atoms for the triplet carbene (9 feature dimensions). For these calculations the M06 functionals has been used. Different molecular structures have been used and the CS group contains $\text{CS}_{\text{ad,ref}}^{\text{M06}}$, $\text{CS}_{\text{ad,B3LYP}}^{\text{M06}}$, $\text{CS}_{\text{ad,PM6}}^{\text{M06}}$, $\text{CS}_{\text{vert,B3LYP}}^{\text{M06}}$ and $\text{CS}_{\text{vert,PM6}}^{\text{M06}}$.

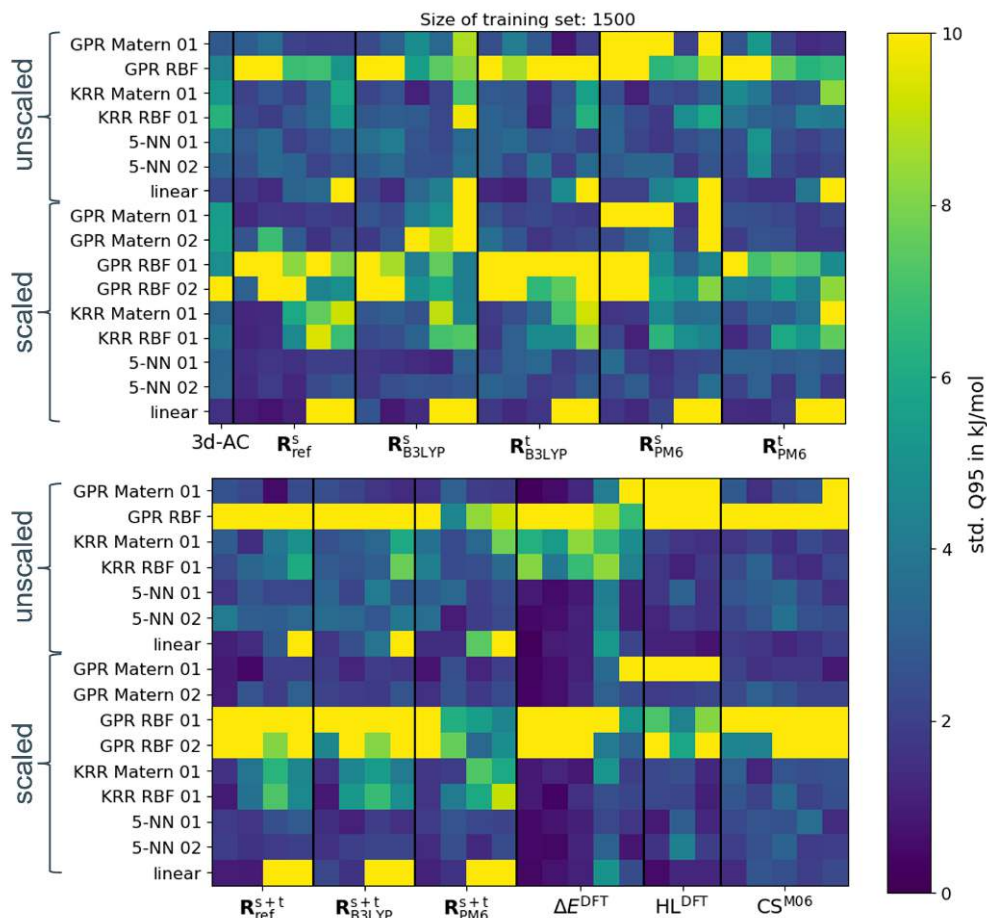


Figure C.10: Overview of standard deviation of Q_{95} for supervised machine learning with different models (y-axis) and descriptors (x-axis). The train-test split has been performed randomly 10 times with training set size of 1500. Models have been used with unscaled and scaled as indicated by the brackets. Descriptors grouped under the **R** label employ only singlet or singlet and triplet molecular structures indicated by superscribed “s” or “s+t”, respectively. The method used for molecular structure optimization is denoted with subscripts. Within the groups, descriptors are ordered: NDA, DA, ACSF, LMBTR and SOAP. The DA descriptor is omitted for the “s+t” groups. For the group ΔE^{DFT} , DFT spin gaps calculated with all 11 xc functionals are used as features. The group contains $\Delta E_{\text{ad,ref}}^{\text{DFT}}$, $\Delta E_{\text{ad,B3LYP}}^{\text{DFT}}$, $\Delta E_{\text{ad,PM6}}^{\text{DFT}}$, $\Delta E_{\text{vert,B3LYP}}^{\text{DFT}}$, $\Delta E_{\text{vert,PM6}}^{\text{DFT}}$. For the HL^{DFT} group HOMO–LUMO, gaps calculated for the singlet carbenes with the 11 xc functionals are used as features. CASSCF, B3LYP- and PM6-optimized singlet structures are used to calculate the HOMO–LUMO gaps. The CS^{M06} descriptor contains Mulliken charges on the carbene carbon and the two neighboring atoms for the singlet and triplet carbene as well as the Mulliken number of unpaired electrons on those three atoms for the triplet carbene (9 feature dimensions). For these calculations the M06 functionals has been used. Different molecular structures have been used and the CS group contains $\text{CS}_{\text{ad,ref}}^{\text{M06}}$, $\text{CS}_{\text{ad,B3LYP}}^{\text{M06}}$, $\text{CS}_{\text{ad,PM6}}^{\text{M06}}$, $\text{CS}_{\text{vert,B3LYP}}^{\text{M06}}$ and $\text{CS}_{\text{vert,PM6}}^{\text{M06}}$.

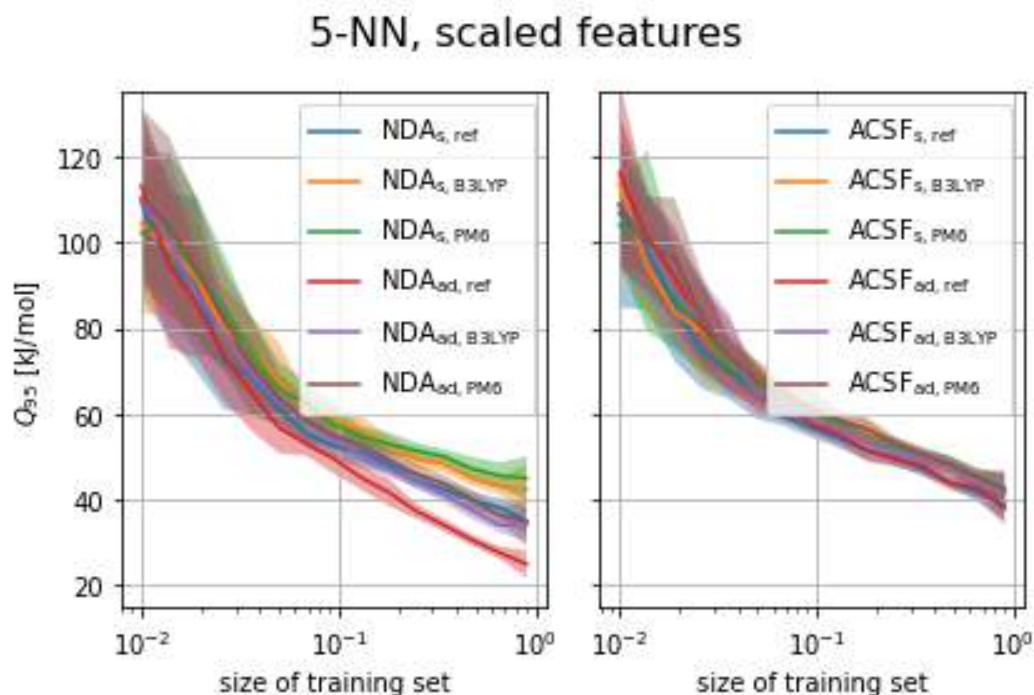


Figure C.11: Learning curve for machine learning of adiabatic reference spin gaps. The GPR with Matern 01 kernel has been used on scaled features. The NDA (left) and ACSF (right) descriptor generated from different structures have been used. The learning has been repeated 10 times. Shaded areas display the standard deviation of the Q_{95} error measure.

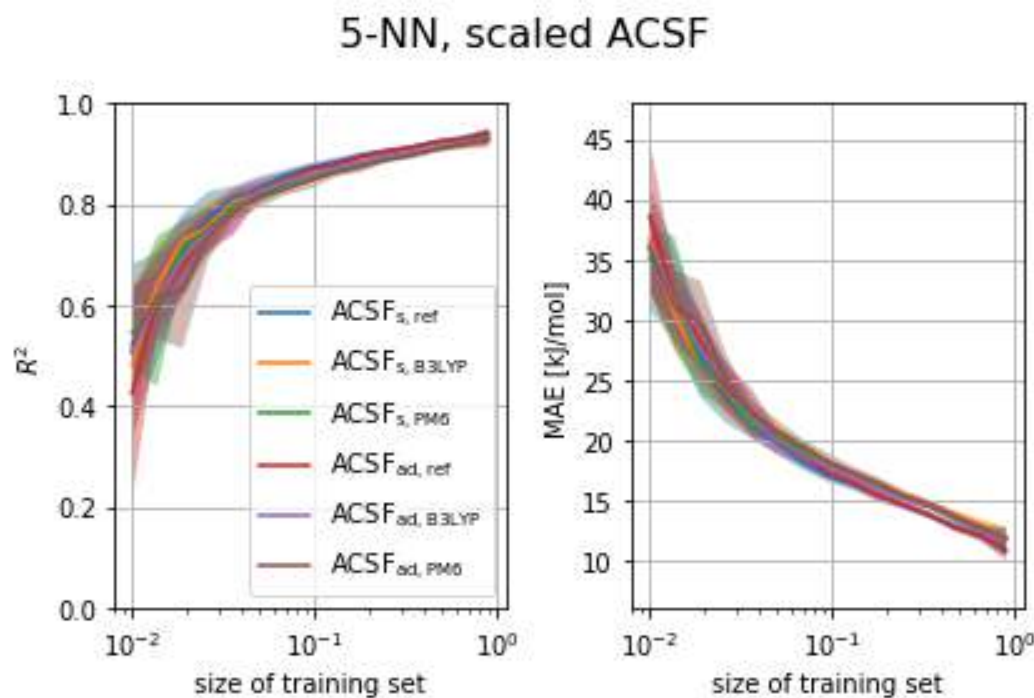


Figure C.12: Learning curve with various scaled ACSF descriptors. Repetitions: 10

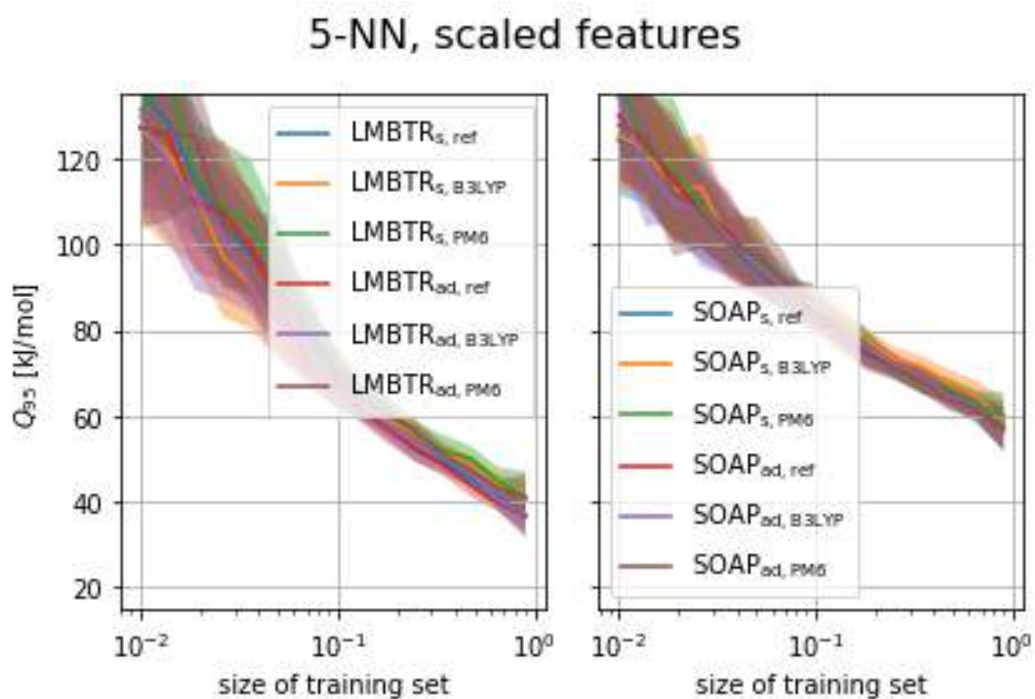


Figure C.13: Learning curve for machine learning of adiabatic reference spin gaps. The k -nearest neighbors methods with $k = 5$ and distance based weights has been used on scaled features. The NDA (left) and ACSF (right) descriptor generated from different structures have been used. The learning has been repeated 15 times. Shaded areas display the standard deviation of the Q_{95} error measure.

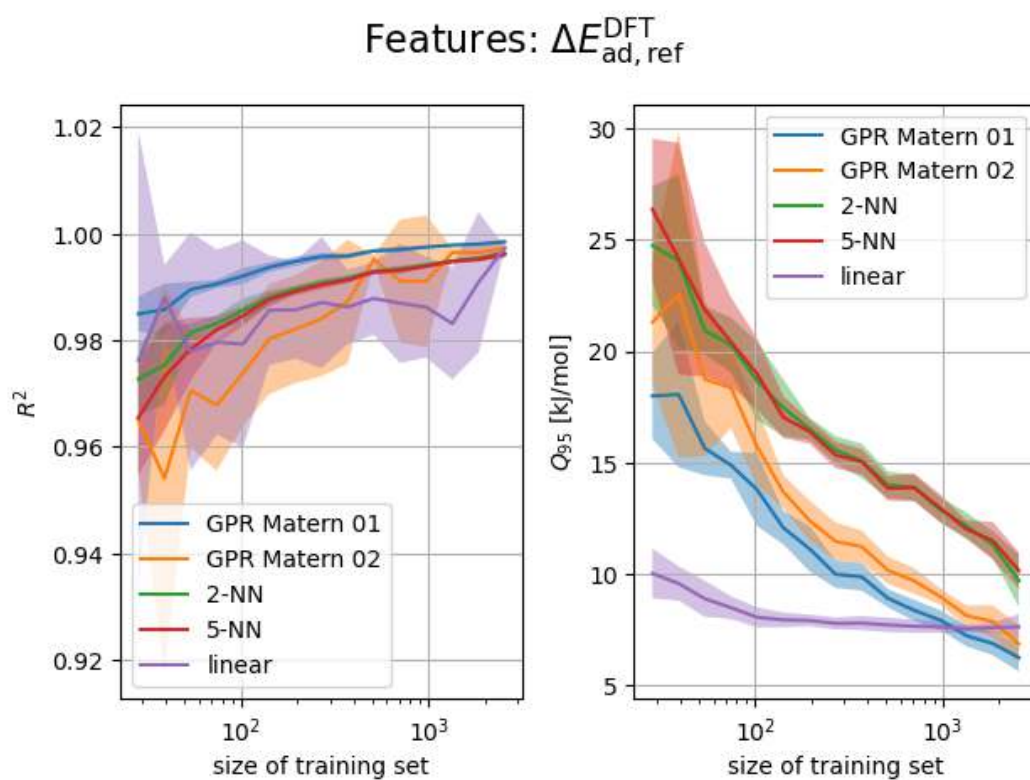


Figure C.14: Learning curve for machine learning of adiabatic reference spin gaps. Different models have been used together with $\Delta E_{\text{ad,ref}}^{\text{DFT}}$ as features. The learning has been repeated 20 times. Shaded areas display the standard deviation of the Q_{95} error measure.

D Supporting Information for Synthesis and Characterization

D.1 IR Spectra



Figure D.1: IR spectrum of **5**.

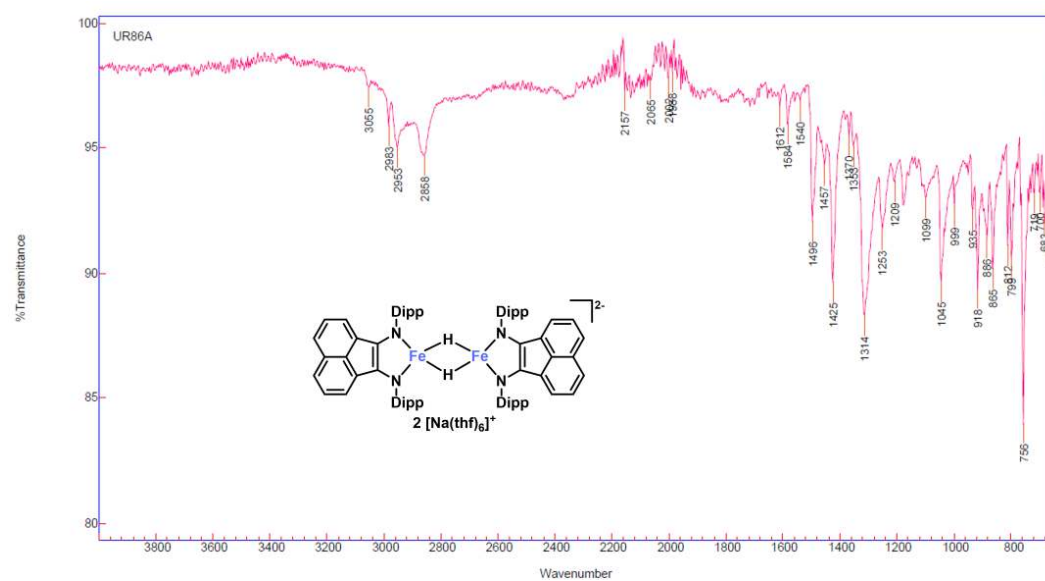


Figure D.2: IR spectrum of **3b**

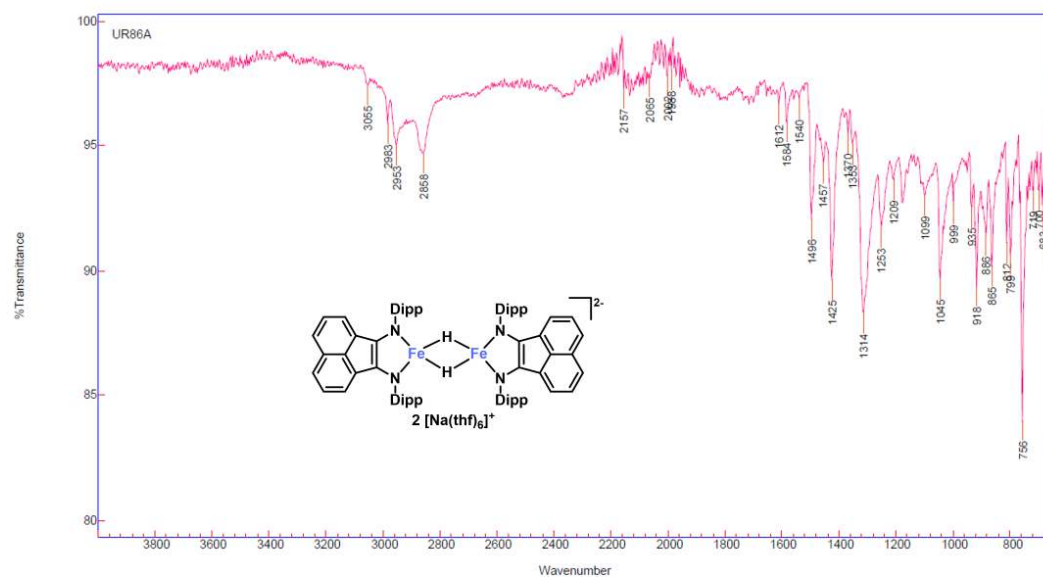


Figure D.3: IR spectrum of 2

D.2 NMR Spectra

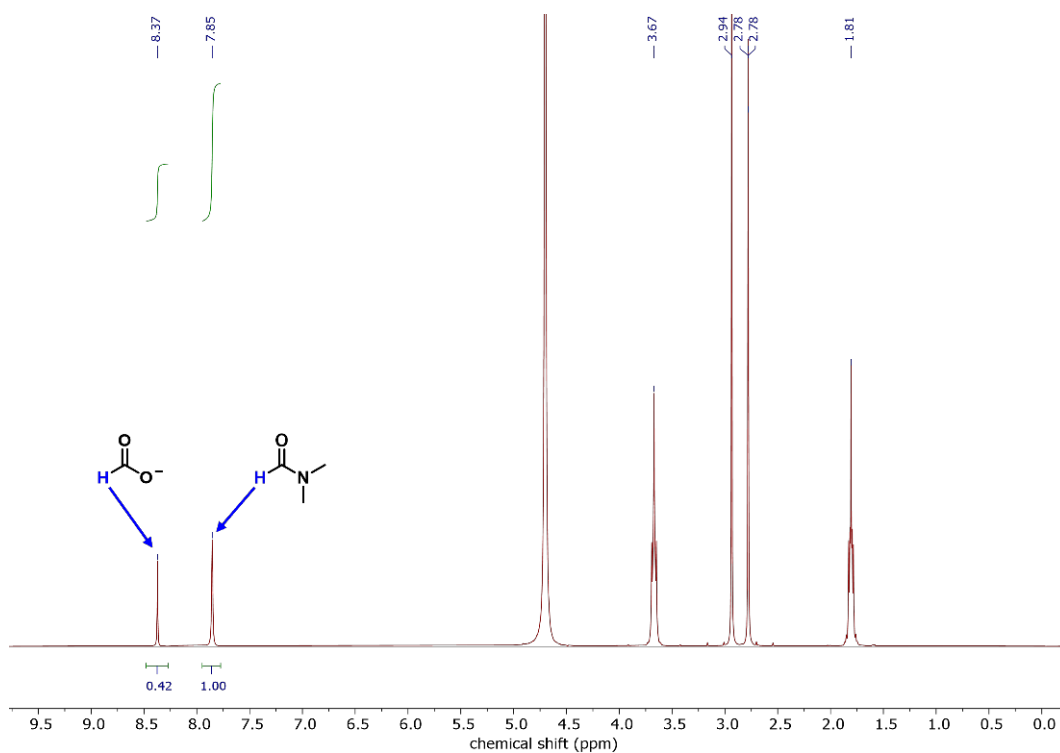


Figure D.4: Example ^1H -NMR spectrum of quantifying formate salt. Measured in D_2O .

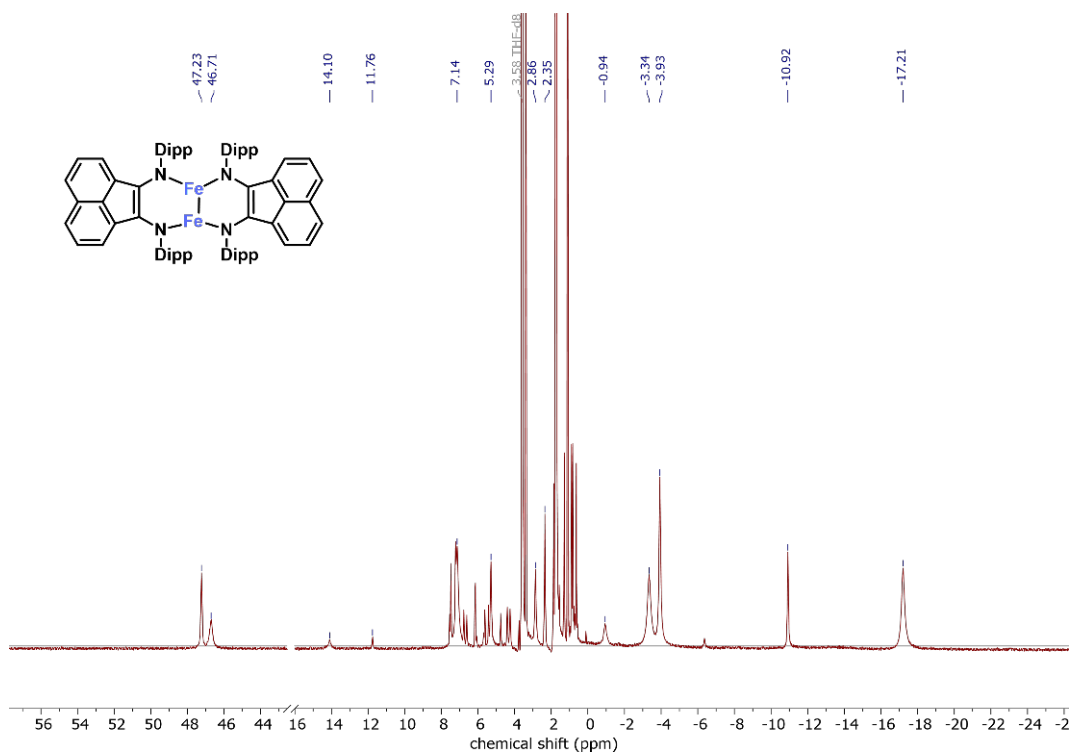


Figure D.5: ^1H -NMR spectrum of **5** measured in $\text{THF-}d_8$.

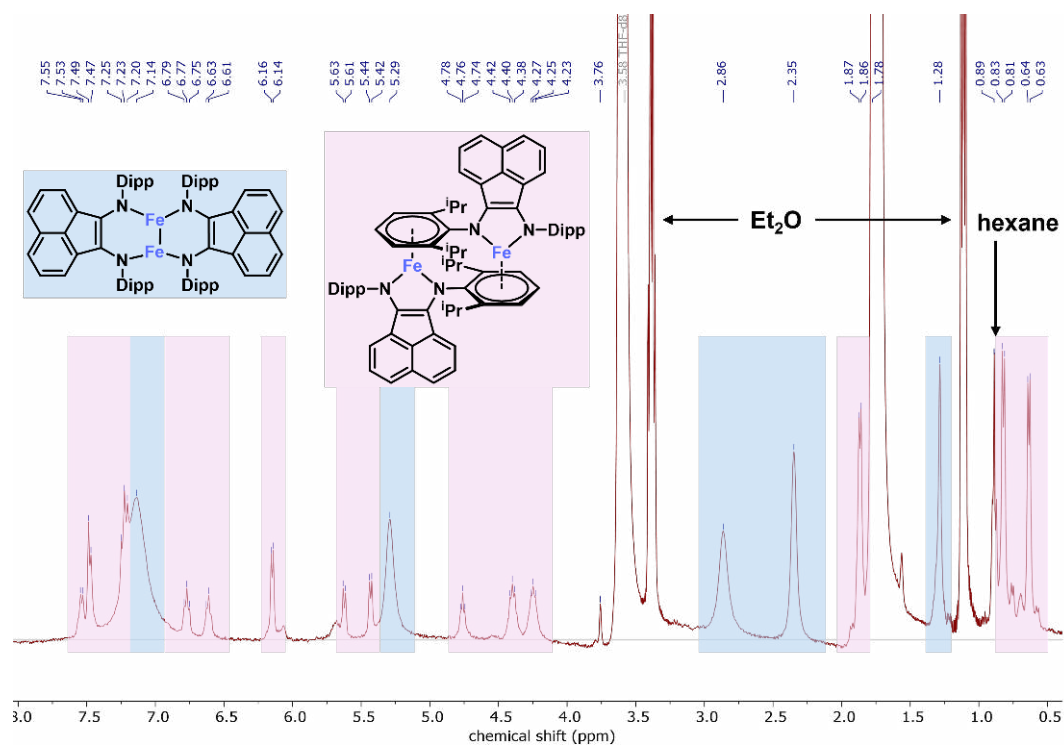


Figure D.6: Detailed view of ^1H -NMR spectrum of **5** measured in $\text{THF-}d_8$.

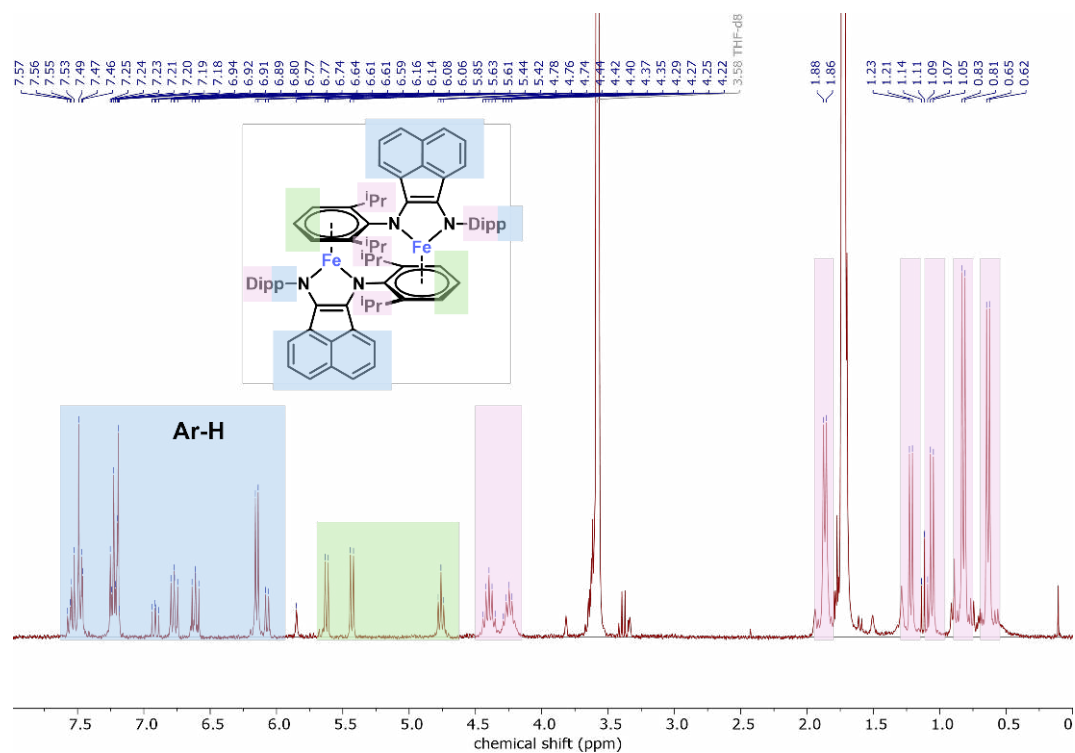


Figure D.7: ^1H -NMR spectrum of **4** measured in $\text{THF-}d_8$.

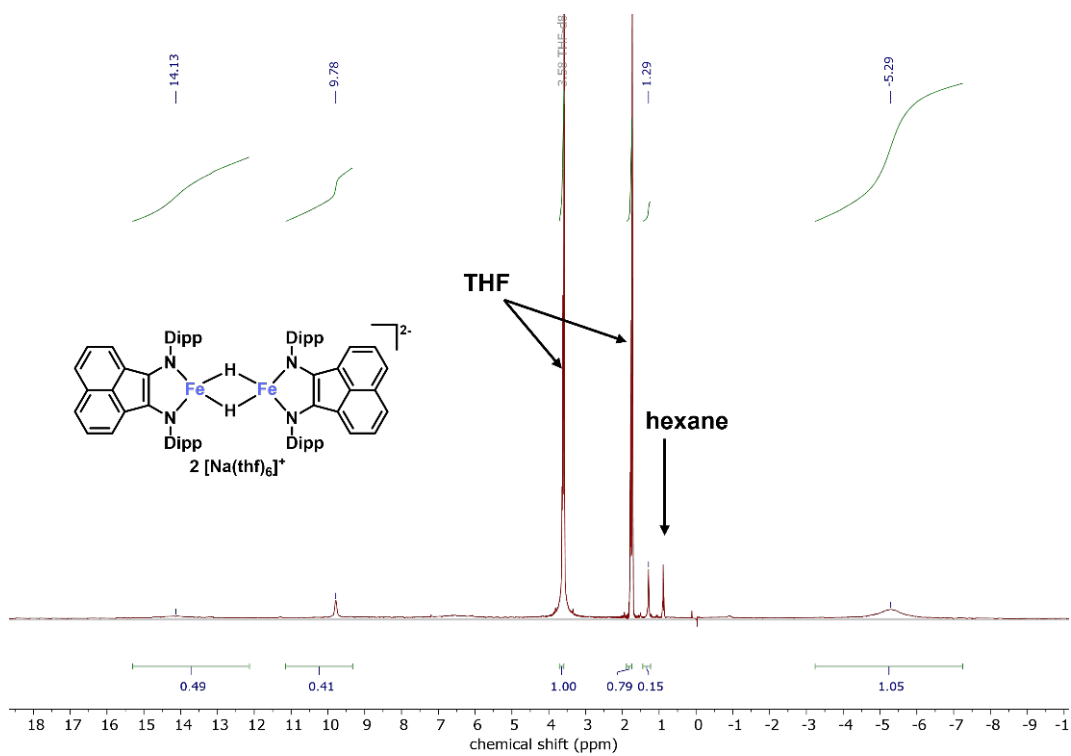


Figure D.8: ^1H -NMR spectrum of **3b** measured in $\text{THF-}d_8$.

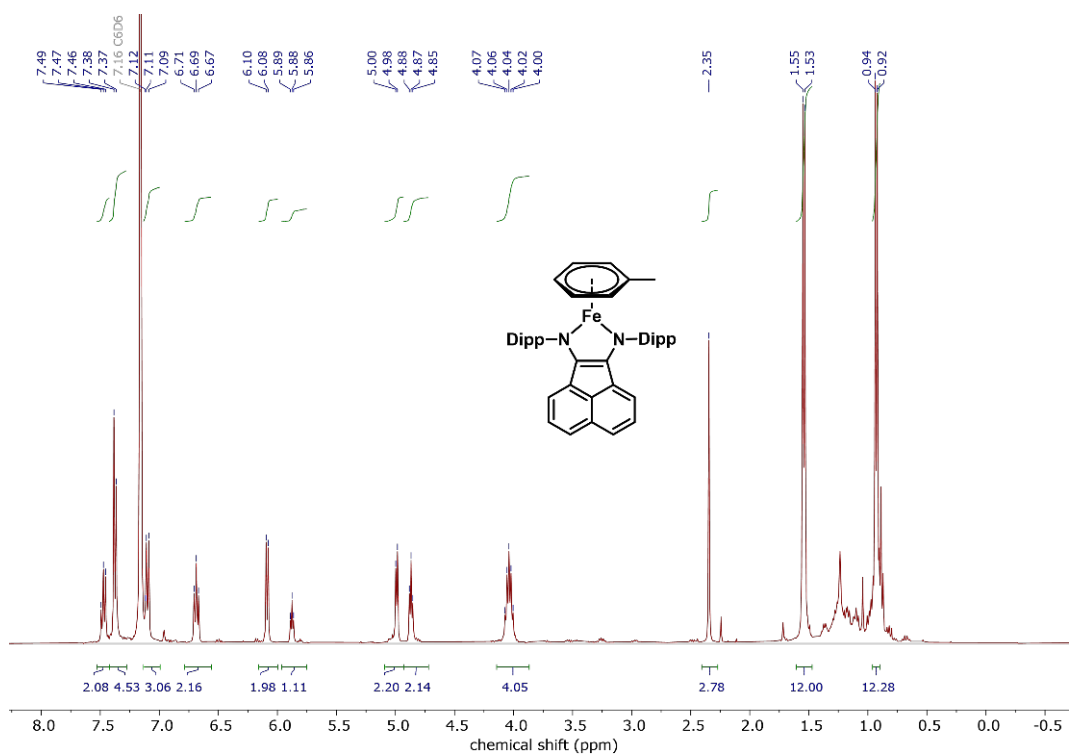


Figure D.9: ^1H -NMR spectrum of **6** measured in C_6D_6 .

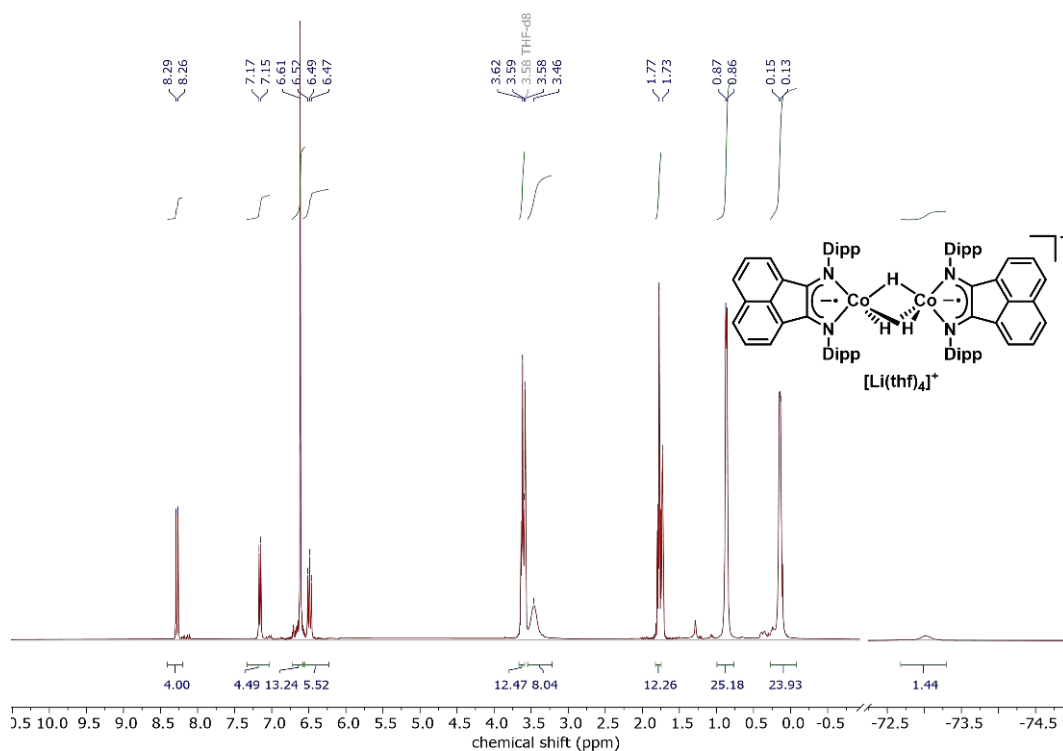


Figure D.10: ^1H -NMR spectrum of **8** measured in $\text{THF-}d_8$.

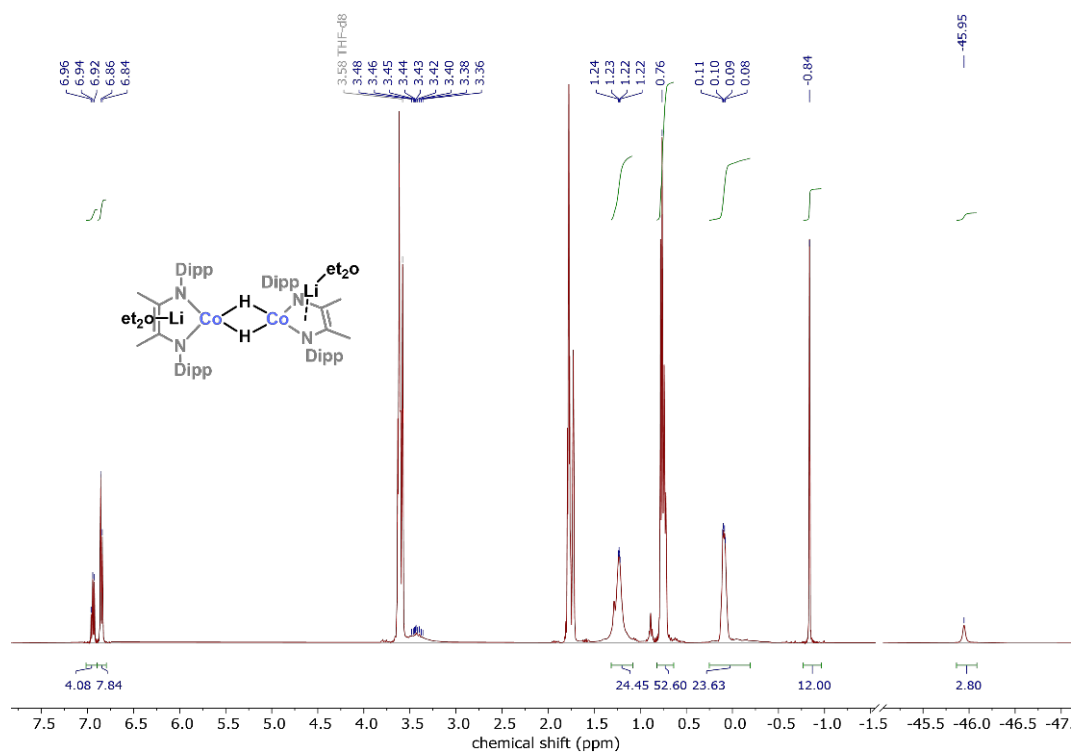


Figure D.11: ^1H -NMR spectrum of **12a** measured in $\text{THF-}d_8$.

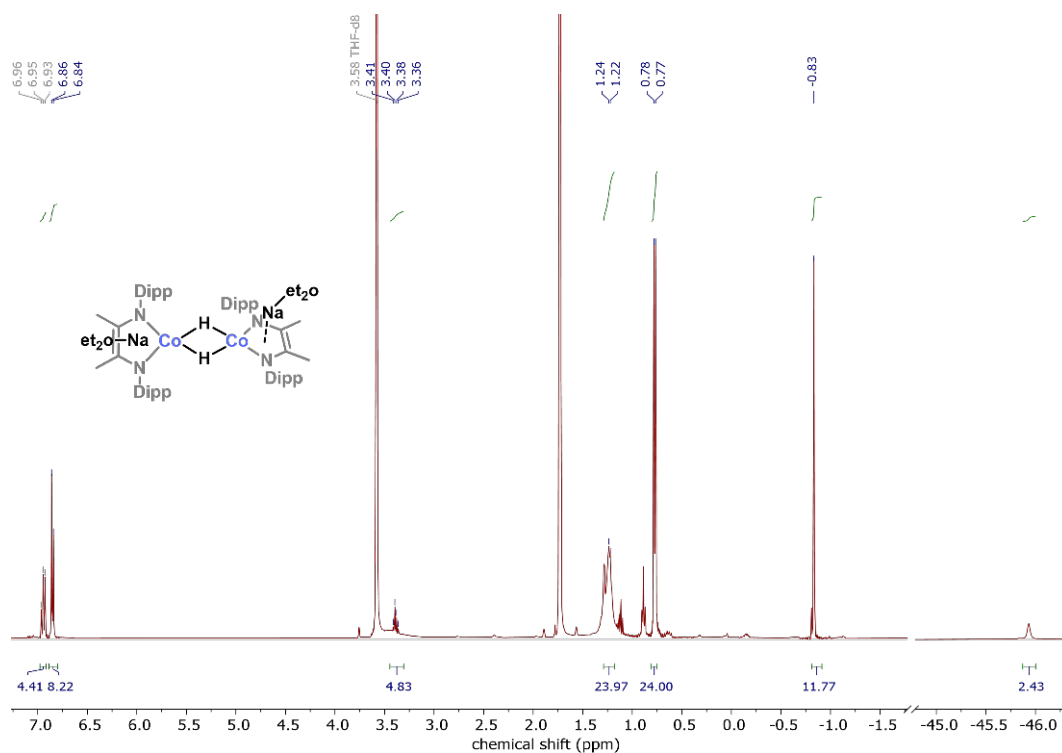


Figure D.12: ^1H -NMR spectrum of **12b** measured in $\text{THF-}d_8$.

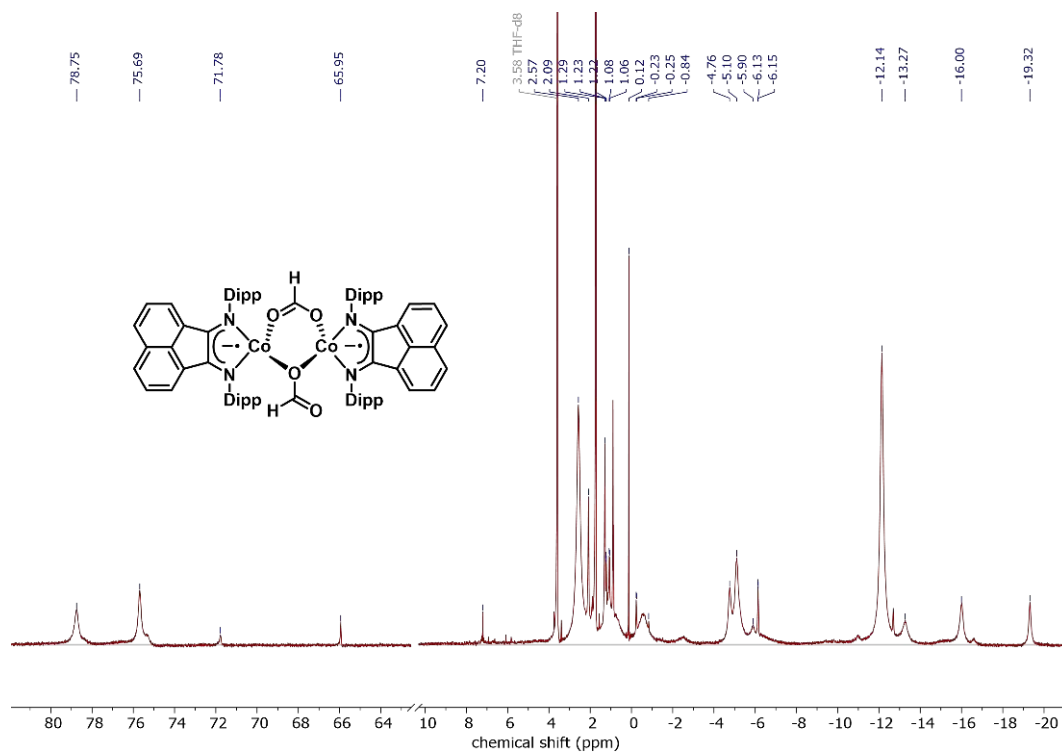


Figure D.13: ^1H -NMR spectrum of **9** measured in $\text{THF-}d_8$.

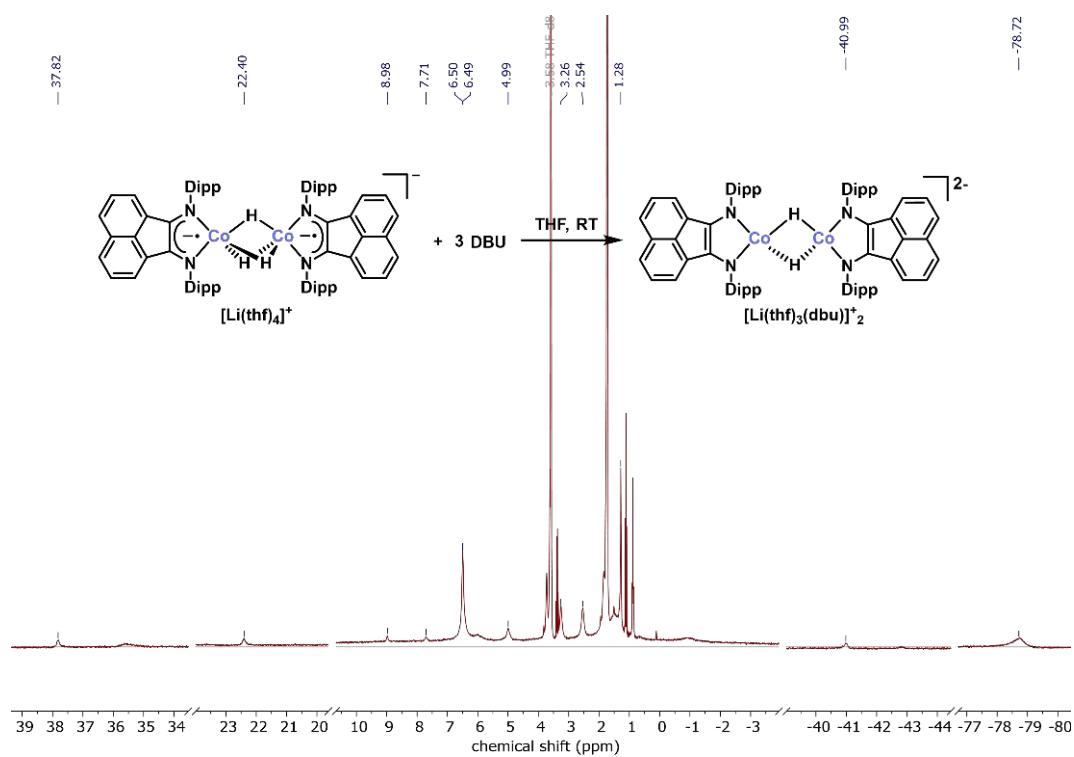


Figure D.14: ^1H -NMR spectrum of the reaction mixture of the deprotonation of **8** with DBU measured in $\text{THF-}d_8$.

D.3 Mass Spectra

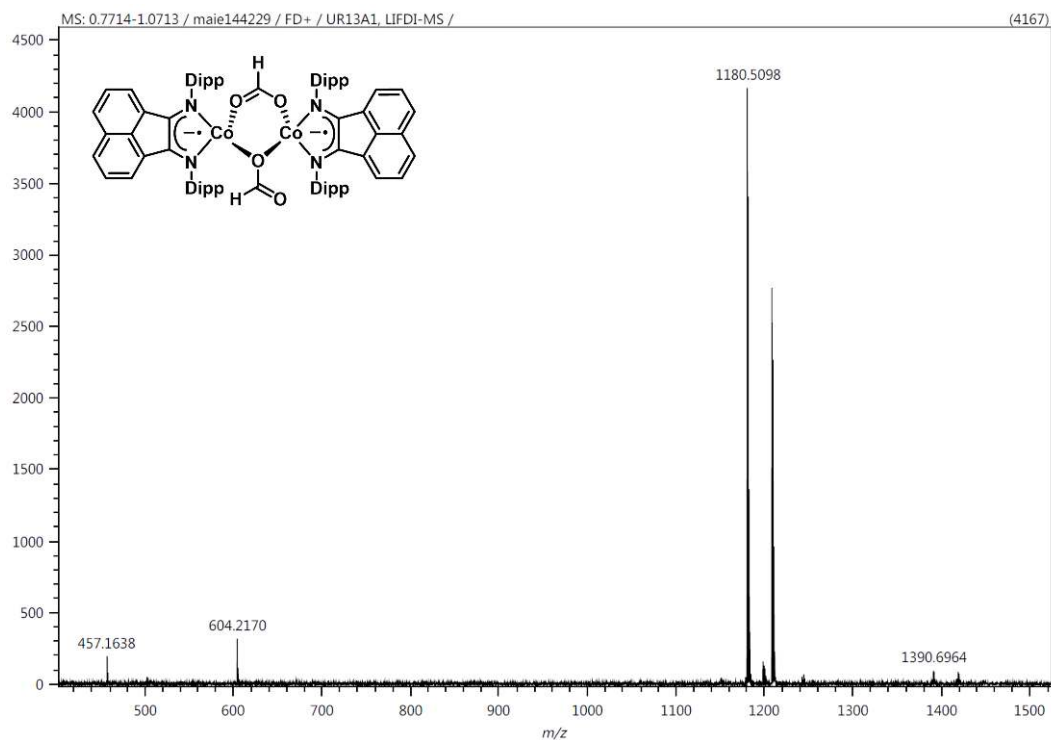
Figure D.15: LIFDI-MS spectrum of **9**.

Table D.1: Measured and theoretical m/z ratios of the ions observed upon negative-ion mode ESI of a solution of **2** in THF ($c \approx 6$ mM). The given values refer to the most abundant isotopologue of each species.

Assignment	m/z (measured)	m/z (theoretical)	Relative error in ppm
$[\text{LFeH}]^-$	557.255	557.263	-14
$[\text{LFe}(\text{H}_2\text{BET}_2)]^-$	627.353	627.359	-10
$[\text{L}_2\text{FeH}]^-$	1057.565	1057.582	-16

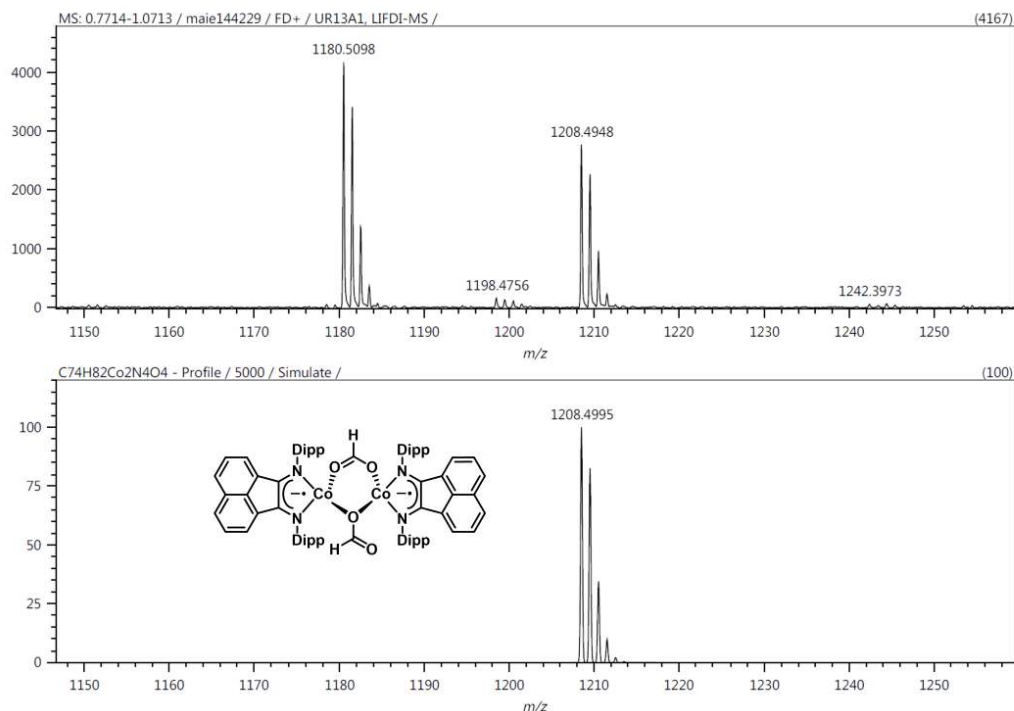


Figure D.16: Zoom of LIFDI-MS spectrum of **9** (top) simulation for $\text{Co}_2\text{C}_{74}\text{H}_{82}\text{N}_4\text{O}_4$ (bottom).

Table D.2: Measured and theoretical m/z ratios of the ions observed upon negative-ion mode ESI of the reaction mixture of $\text{LFeCl}_2/3\text{LiBEt}_3\text{H}$ in THF ($c \approx 8 \text{ mM}$). The given values refer to the most abundant isotopologue of each species.

Assignment	m/z (measured)	m/z (theoretical)	Relative error in ppm
$[\text{LFeH}]^-$	557.260	557.263	-5
$[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$	627.358	627.359	-2
$[\text{LFe}(\text{HBEt}_2)(\text{MeCN})]^-$	696.413	696.417	16

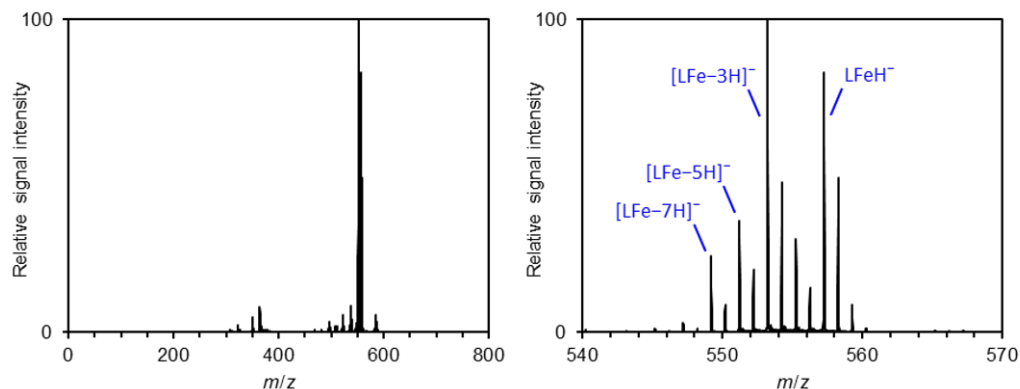


Figure D.17: Mass spectrum of mass-selected $[\text{LFeH}]^-$ and its fragment ions formed upon collision-induced dissociation at $E_{\text{LAB}} = 30 \text{ eV}$. Left: whole spectrum, right: enlarged section

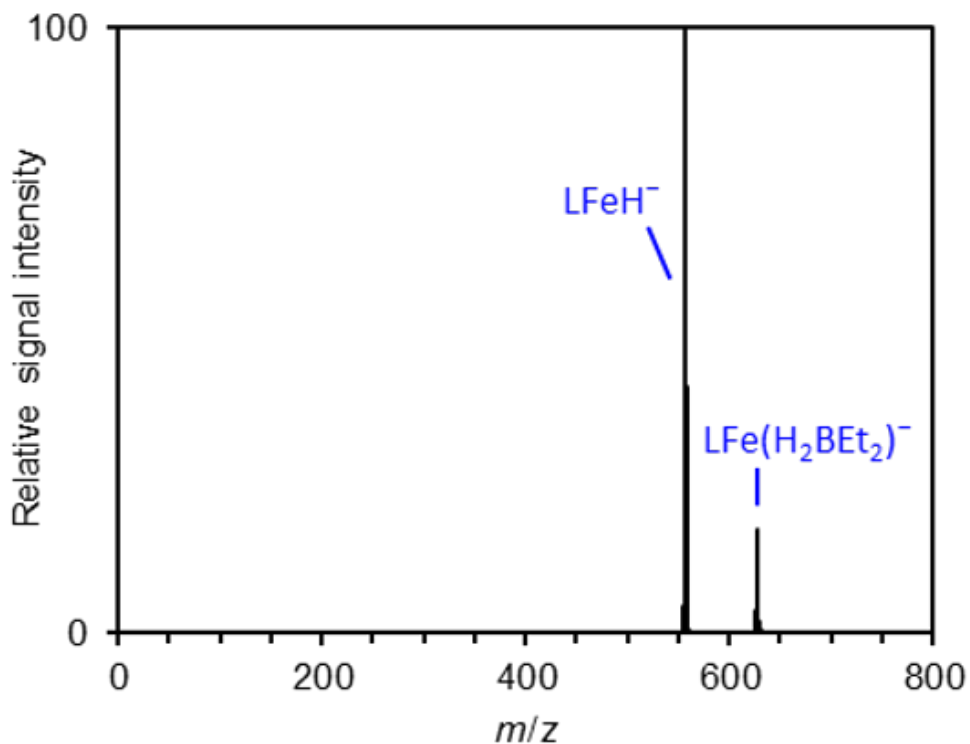


Figure D.18: Mass spectrum of mass-selected $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ and its fragment ions formed upon collision-induced dissociation at $E_{\text{LAB}} = 12$ eV. Left: whole spectrum, right: enlarged section

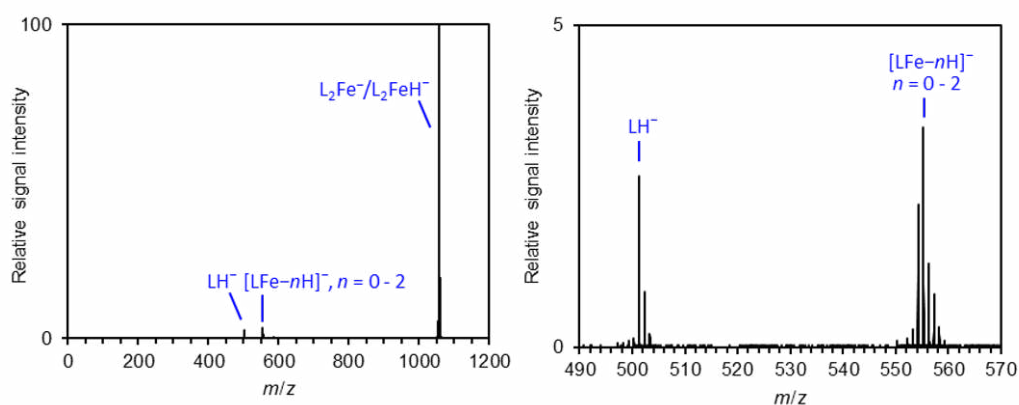


Figure D.19: Mass spectrum of mass-selected $[\text{L}_2\text{FeH}_n]^-$ ($n = 0$ and 1) and its fragment ions formed upon collision-induced dissociation at $E_{\text{LAB}} = 30$ eV. Left: whole spectrum, right: enlarged section

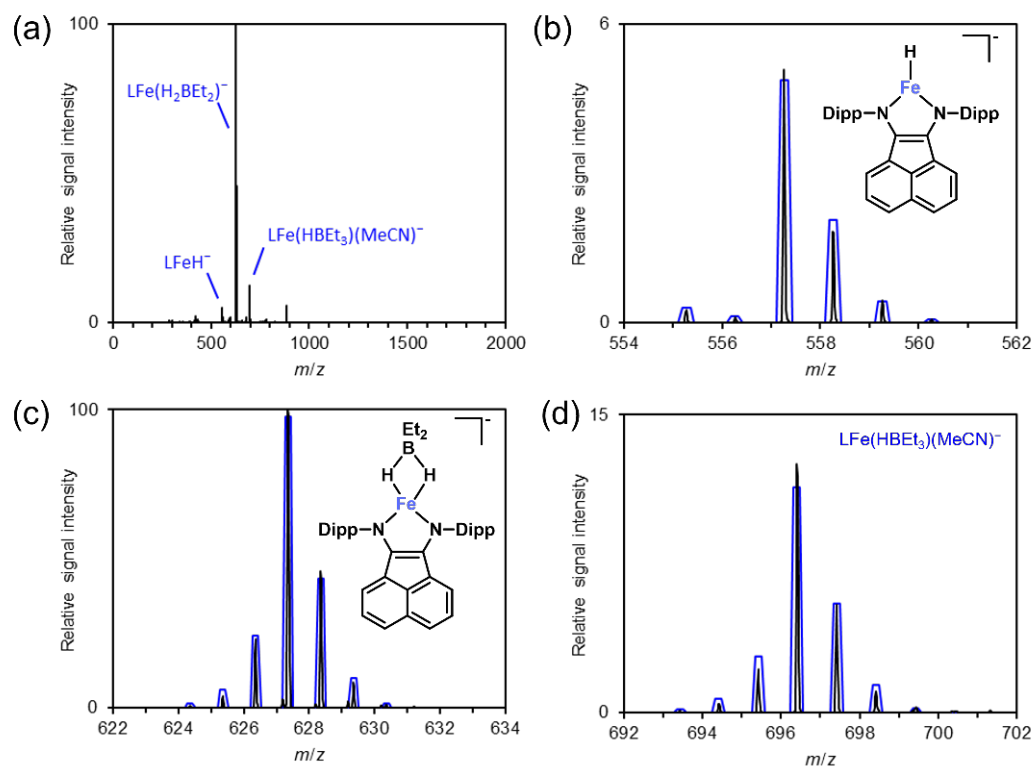


Figure D.20: (a): Negative-ion mode ESI mass spectrum of a solution of the reaction mixture of $\text{LFeCl}_2/3\text{LiBEt}_3\text{H}$ in THF ($c \approx 8\text{ mM}$). (b): Section from the negative-ion mode ESI mass spectrum (black) together with theoretical isotope pattern of $[\text{LFeH}]^-$ (blue). (c): Section from the negative-ion mode ESI mass spectrum (black) together with theoretical isotope pattern of $[\text{LFe}(\text{H}_2\text{BEt}_2)]^-$ (blue). (d): Section from the negative-ion mode ESI mass spectrum (black) together with theoretical isotope pattern of $[\text{LFe}(\text{HBEt}_2)(\text{MeCN})]^-$ (blue).

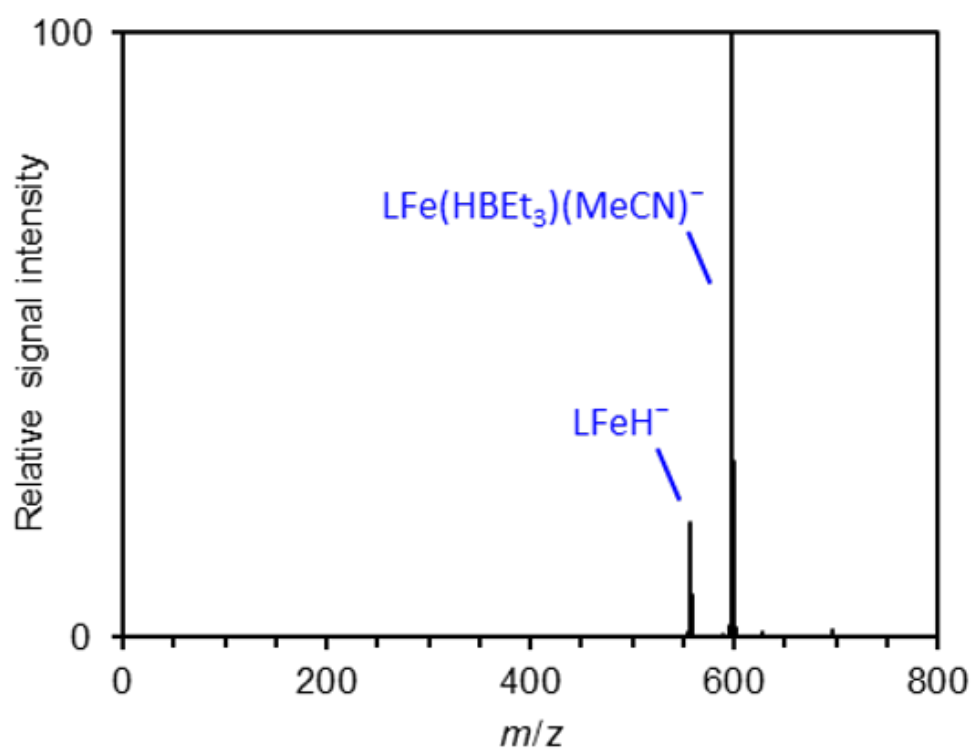


Figure D.21: Mass spectrum of mass-selected $[\text{LFe(HBEt}_2\text{)(MeCN)}]^-$ and its fragment ions formed upon collision-induced dissociation at $E_{\text{LAB}} = 10 \text{ eV}$. Left: whole spectrum, right: enlarged section

D.4 Crystal Data

Table D.3: Crystal data and structure refinement for **3b**

Identification code	UR65B
Empirical formula	C ₁₂₈ H ₁₉₄ Fe ₂ N ₄ Na ₂ O ₁₄
Formula weight	2170.54
Temperature/K	100.0(7)
Crystal system	monoclinic
Space group	P2 ₁ /c
a/Å	12.77279(12)
b/Å	29.8307(3)
c/Å	17.35368(18)
$\alpha/^\circ$	90.0
$\beta/^\circ$	111.4258(11)
$\gamma/^\circ$	90.0
Volume/Å ³	6155.17(10)
Z	2
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.171
μ/mm^{-1}	2.433
F(000)	2352.0
Crystal size/mm ³	0.25 × 0.25 × 0.05
Radiation	Cu K α (λ = 1.54184)
2 Θ range for data collection/ $^\circ$	6.222 to 152.818
Index ranges	-15 ≤ h ≤ 16, -37 ≤ k ≤ 37, -21 ≤ l ≤ 21
Reflections collected	138081
Independent reflections	12823 [R_{int} = 0.0579, R_{sigma} = 0.0240]
Data/restraints/parameters	12823/648/698
Goodness-of-fit on F ²	1.036
Final R indexes [$I \geq 2\sigma(I)$]	R_1 = 0.0467, wR_2 = 0.1203
Final R indexes [all data]	R_1 = 0.0510, wR_2 = 0.1239
Largest diff. peak/hole / e Å ³	0.64/-0.52

Table D.4: Crystal data and structure refinement for **3c**

Identification code	UR86V_2
Empirical formula	$\text{C}_{88}\text{H}_{114}\text{Fe}_2\text{N}_4\text{Na}_2\text{O}_4$
Formula weight	1449.51
Temperature/K	200(140)
Crystal system	monoclinic
Space group	$\text{P2}_1/\text{n}$
a/Å	14.4333(3)
b/Å	13.7450(2)
c/Å	19.3612(4)
$\alpha/^\circ$	90.0
$\beta/^\circ$	90.332(2)
$\gamma/^\circ$	90.0
Volume/Å ³	3840.92(13)
Z	2
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.253
μ/mm^{-1}	0.443
F(000)	1552.0
Crystal size/mm ³	$0.24 \times 0.22 \times 0.18$
Radiation	Mo K_α ($\lambda = 0.71073$)
2Θ range for data collection/ $^\circ$	4.594 to 59.13
Index ranges	$-19 \leq h \leq 19, -18 \leq k \leq 18, -26 \leq l \leq 24$
Reflections collected	114684
Independent reflections	10274 [$R_{\text{int}} = 0.0479, R_{\text{sigma}} = 0.0240$]
Data/restraints/parameters	10274/0/463
Goodness-of-fit on F^2	1.040
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0414, wR_2 = 0.1023$
Final R indexes [all data]	$R_1 = 0.0489, wR_2 = 0.1071$
Largest diff. peak/hole / e Å ³	0.96/-0.50

Table D.5: Crystal data and structure refinement for **4**

Identification code	UR33H-2
Empirical formula	C ₇₂ H ₈₀ Fe ₂ N ₄
Formula weight	1113.10
Temperature/K	99.96(18)
Crystal system	monoclinic
Space group	P2 ₁ /n
a/Å	13.1317(2)
b/Å	13.23430(10)
c/Å	17.5846(2)
$\alpha/^\circ$	90.0
$\beta/^\circ$	107.1180(10)
$\gamma/^\circ$	90.0
Volume/Å ³	2920.63(6)
Z	2
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.266
μ/mm^{-1}	0.543
F(000)	1184.0
Crystal size/mm ³	0.452 × 0.243 × 0.142
Radiation	Mo K $_{\alpha}$ (λ = 0.71073)
2 Θ range for data collection/ $^\circ$	5.742 to 65.928
Index ranges	-19 ≤ h ≤ 19, -20 ≤ k ≤ 20, -26 ≤ l ≤ 26
Reflections collected	183454
Independent reflections	10637 [R_{int} = 0.0341, R_{sigma} = 0.0126]
Data/restraints/parameters	10637/0/360
Goodness-of-fit on F ²	1.071
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0586$, $wR_2 = 0.1818$
Final R indexes [all data]	$R_1 = 0.0631$, $wR_2 = 0.1868$
Largest diff. peak/hole / e Å ³	1.52/-1.51

Table D.6: Crystal data and structure refinement for **5**

Identification code	UR25b1
Empirical formula	C ₇₂ H ₈₀ Fe ₂ N ₄
Formula weight	1113.10
Temperature/K	99.96(18)
Crystal system	monoclinic
Space group	C2/c
a/Å	19.1589(4)
b/Å	18.1992(3)
c/Å	24.2834(5)
$\alpha/^\circ$	90.0
$\beta/^\circ$	99.617(2)
$\gamma/^\circ$	90.0
Volume/Å ³	8348.1(3)
Z	4
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	0.886
μ/mm^{-1}	0.380
F(000)	2368.0
Crystal size/mm ³	0.3 × 0.28 × 0.06
Radiation	Mo K $_{\alpha}$ ($\lambda = 0.71073$)
2 Θ range for data collection/ $^\circ$	5.924 to 59.202
Index ranges	-24 ≤ h ≤ 26, -24 ≤ k ≤ 23, -31 ≤ l ≤ 32
Reflections collected	109592
Independent reflections	11141 [$R_{\text{int}} = 0.0479$, $R_{\text{sigma}} = 0.0270$]
Data/restraints/parameters	11141/0/360
Goodness-of-fit on F ²	1.021
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0590$, $wR_2 = 0.1497$
Final R indexes [all data]	$R_1 = 0.0746$, $wR_2 = 0.1592$
Largest diff. peak/hole / e Å ³	0.53/-0.34

Table D.7: Crystal data and structure refinement for **9**

Identification code	UR13a1
Empirical formula	C ₁₄₈ H ₁₆₄ Co ₄ N ₈ O ₈
Formula weight	2418.58
Temperature/K	99.9(3)
Crystal system	monoclinic
Space group	P2 ₁ /n
a/Å	10.8709(2)
b/Å	14.0289(2)
c/Å	21.2653(4)
$\alpha/^\circ$	90.0
$\beta/^\circ$	93.7880(10)
$\gamma/^\circ$	90.0
Volume/Å ³	3236.02(10)
Z	1
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.241
μ/mm^{-1}	0.564
F(000)	1280.0
Crystal size/mm ³	0.3 × 0.2 × 0.1
Radiation	Mo K $_{\alpha}$ (λ = 0.71073)
2 Θ range for data collection/ $^\circ$	5.948 to 59.228
Index ranges	-14 ≤ h ≤ 15, -19 ≤ k ≤ 19, -29 ≤ l ≤ 29
Reflections collected	140115
Independent reflections	8834 [R _{int} = 0.0335, R _{sigma} = 0.0153]
Data/restraints/parameters	8834/813/671
Goodness-of-fit on F ²	1.121
Final R indexes [I ≥ 2 σ (I)]	R ₁ = 0.0534, wR ₂ = 0.1229
Final R indexes [all data]	R ₁ = 0.0617, wR ₂ = 0.1271
Largest diff. peak/hole / e Å ³	0.49/-0.32

Table D.8: Crystal data and structure refinement for **12b**

Identification code	UR59A2
Empirical formula	$\text{C}_{68}\text{H}_{112}\text{Co}_2\text{N}_4\text{Na}_2\text{O}_3$
Formula weight	1197.45
Temperature/K	99.97(18)
Crystal system	triclinic
Space group	P-1
a/Å	12.5304(3)
b/Å	13.0331(5)
c/Å	13.0416(3)
$\alpha/^\circ$	61.799(3)
$\beta/^\circ$	81.560(2)
$\gamma/^\circ$	64.995(3)
Volume/Å ³	1697.33(10)
Z	1
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.171
μ/mm^{-1}	0.547
F(000)	648.0
Crystal size/mm ³	$0.32 \times 0.18 \times 0.1$
Radiation	Mo K $_{\alpha}$ ($\lambda = 0.71073$)
2 Θ range for data collection/ $^\circ$	6.132 to 60.048
Index ranges	$-17 \leq h \leq 17, -17 \leq k \leq 17, -17 \leq l \leq 18$
Reflections collected	75834
Independent reflections	9017 [$R_{\text{int}} = 0.0439, R_{\text{sigma}} = 0.0285$]
Data/restraints/parameters	9017/54/397
Goodness-of-fit on F ²	1.052
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0426, wR_2 = 0.1036$
Final R indexes [all data]	$R_1 = 0.0533, wR_2 = 0.1094$
Largest diff. peak/hole / e Å ³	0.96/-0.44

Table D.9: Crystal data and structure refinement for **10**

Identification code	UR69
Empirical formula	C ₁₁₄ H ₁₆₂ Co ₂ Li ₂ N ₈ O ₆
Formula weight	1872.25
Temperature/K	201(140)
Crystal system	triclinic
Space group	P-1
a/Å	13.0903(4)
b/Å	14.9456(4)
c/Å	15.0317(4)
$\alpha/^\circ$	96.566(2)
$\beta/^\circ$	95.032(2)
$\gamma/^\circ$	105.366(2)
Volume/Å ³	2795.65(14)
Z	1
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.112
μ/mm^{-1}	2.732
F(000)	1010.0
Crystal size/mm ³	0.16 × 0.16 × 0.08
Radiation	Cu K $_{\alpha}$ ($\lambda = 1.54184$)
2 Θ range for data collection/ $^\circ$	6.196 to 153.458
Index ranges	-16 ≤ h ≤ 16, -18 ≤ k ≤ 18, -18 ≤ l ≤ 18
Reflections collected	54573
Independent reflections	11464 [$R_{\text{int}} = 0.0495$, $R_{\text{sigma}} = 0.0371$]
Data/restraints/parameters	11464/582/635
Goodness-of-fit on F ²	1.083
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0490$, $wR_2 = 0.1275$
Final R indexes [all data]	$R_1 = 0.0580$, $wR_2 = 0.0580$
Largest diff. peak/hole / e Å ³	0.38/-0.44

Table D.10: Crystal data and structure refinement for **12a**

Identification code	UR52B
Empirical formula	$\text{C}_{64}\text{H}_{102}\text{Co}_2\text{Li}_2\text{N}_4\text{O}_2$
Formula weight	1091.23
Temperature/K	99.9(2)
Crystal system	monoclinic
Space group	P2/c
a/Å	22.5648(3)
b/Å	14.45820(10)
c/Å	20.3155(2)
$\alpha/^\circ$	90
$\beta/^\circ$	107.3920(10)
$\gamma/^\circ$	90
Volume/Å ³	6324.84(12)
Z	4
$\rho_{\text{calc}}/\text{g}/\text{cm}^3$	1.146
μ/mm^{-1}	4.423
F(000)	2360.0
Crystal size/mm ³	$0.3 \times 0.15 \times 0.12$
Radiation	Cu K $_{\alpha}$ ($\lambda = 1.54184$)
2 Θ range for data collection/ $^\circ$	7.364 to 152.236
Index ranges	$-27 \leq h \leq 28$, $-18 \leq k \leq 18$, $-5 \leq l \leq 25$
Reflections collected	137981
Independent reflections	13139 [$R_{\text{int}} = 0.0573$, $R_{\text{sigma}} = 0.0253$]
Data/restraints/parameters	13139/697/702
Goodness-of-fit on F ²	1.031
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0854$, $wR_2 = 0.2256$
Final R indexes [all data]	$R_1 = 0.0895$, $wR_2 = 0.2287$
Largest diff. peak/hole / e Å ³	0.98/-0.74

Table D.11: Crystal data and structure refinement for $[\text{CoBr}_2\text{dbu}]_2$

Identification code	UR60C
Empirical formula	$\text{C}_{18}\text{H}_{31}\text{Br}_2\text{CoN}_4$
Formula weight	522.22
Temperature/K	99.9(2)
Crystal system	monoclinic
Space group	$\text{P}2_1/\text{n}$
$a/\text{\AA}$	12.32(4)
$b/\text{\AA}$	14.69(3)
$c/\text{\AA}$	14.41(3)
$\alpha/^\circ$	90
$\beta/^\circ$	90
$\gamma/^\circ$	90
Volume/ \AA^3	2610(11)
Z	4
$\rho_{\text{calc}}/\text{g cm}^{-3}$	1.329
μ/mm^{-1}	3.727
$F(000)$	1056.0
Crystal size/ mm^3	$0.32 \times 0.26 \times 0.21$
Radiation	Mo K_α ($\lambda = 0.71073$)
2θ range for data collection/ $^\circ$	5.158 to 70.52
Index ranges	$-16 \leq h \leq 15, -20 \leq k \leq 20, -19 \leq l \leq 20$
Reflections collected	32244
Independent reflections	6405 [$R_{\text{int}} = 0.0280, R_{\text{sigma}} = 0.0233$]
Data/restraints/parameters	6405/0/226
Goodness-of-fit on F^2	1.112
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0694, wR_2 = 0.1805$
Final R indexes [all data]	$R_1 = 0.0901, wR_2 = 0.2011$
Largest diff. peak/hole / $e \text{\AA}^{-3}$	1.27/-4.00

D.5 Cyclic Voltammetry

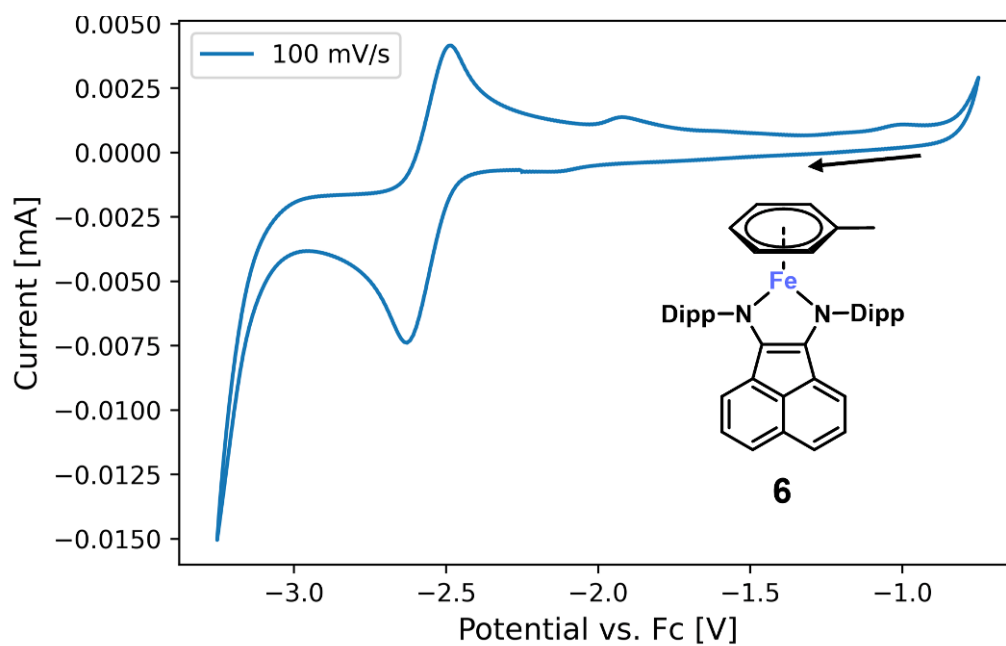


Figure D.22: Cyclic voltammograms of complex **6**.

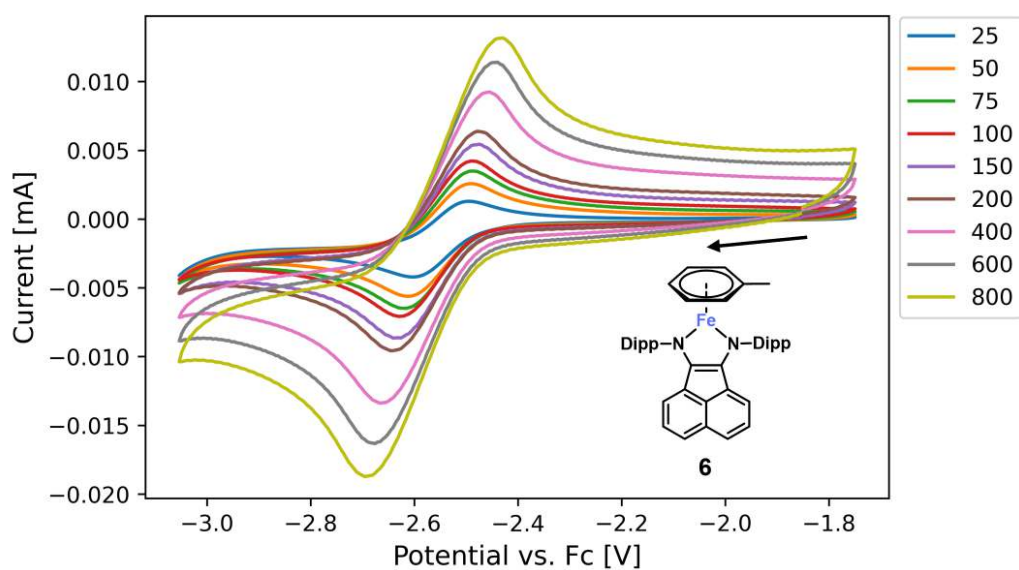


Figure D.23: Cyclic voltammograms of complex **6** at different scan rates (in m V s⁻¹).

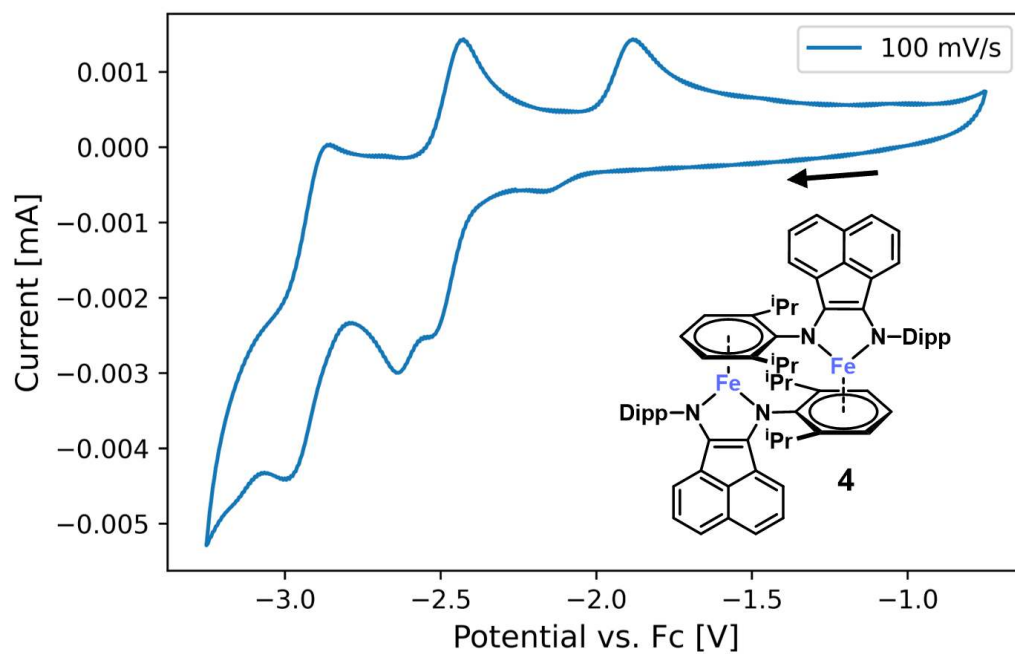


Figure D.24: Cyclic voltammograms of complex **4**.

D.6 Additional DFT Calculations

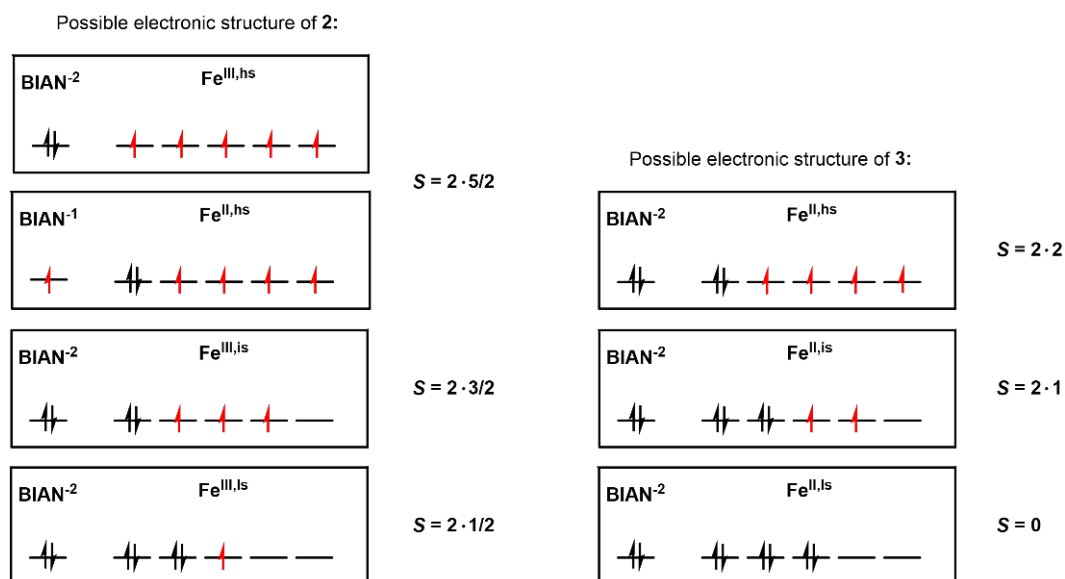


Figure D.25: Possible electronic structures for each L-Fe fragment in complex **2** (left) and complex **3**. Resulting total spin S for each possibility is given for the whole complex. Cases, where the ligand is charged -1 does not influence the total spin. In the case of antiferromagnetic coupling $S = 0$ is also plausible for the complex **2**.

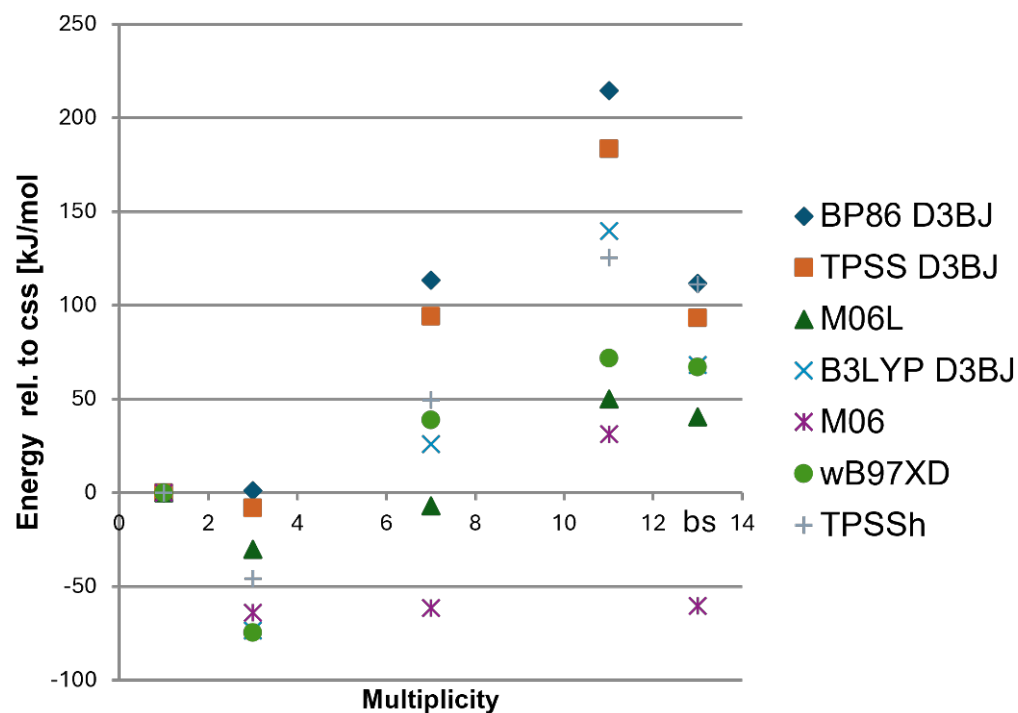


Figure D.26: Energies of spin states calculated with different xc functionals and def2-TZVP basis set and with PCM(THF) on TPSS D3BJ optimized molecular structures of **2**. The undecet molecular structure was used for all septet calculations due to the failed convergence of the septet molecular structure optimization. For the triplet molecular structure only the equatorial hydrides are bridging both iron centers. The remaining two hydrides coordinate one iron center each.

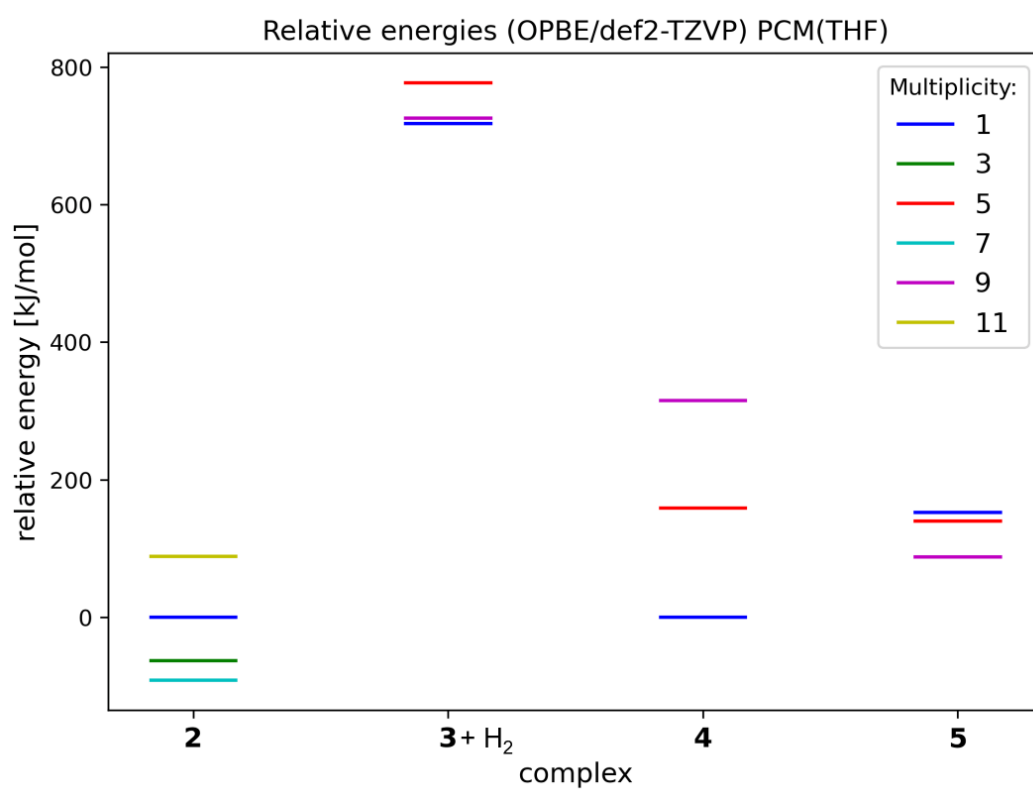


Figure D.27: Energies of spin states calculated with OPBE/def2-TZVP with PCM(THF) on crystal structures relative to oss of complex **2** for complexes **2** and **3** and relative to the oss of **14** for complexes **14** and **5**.

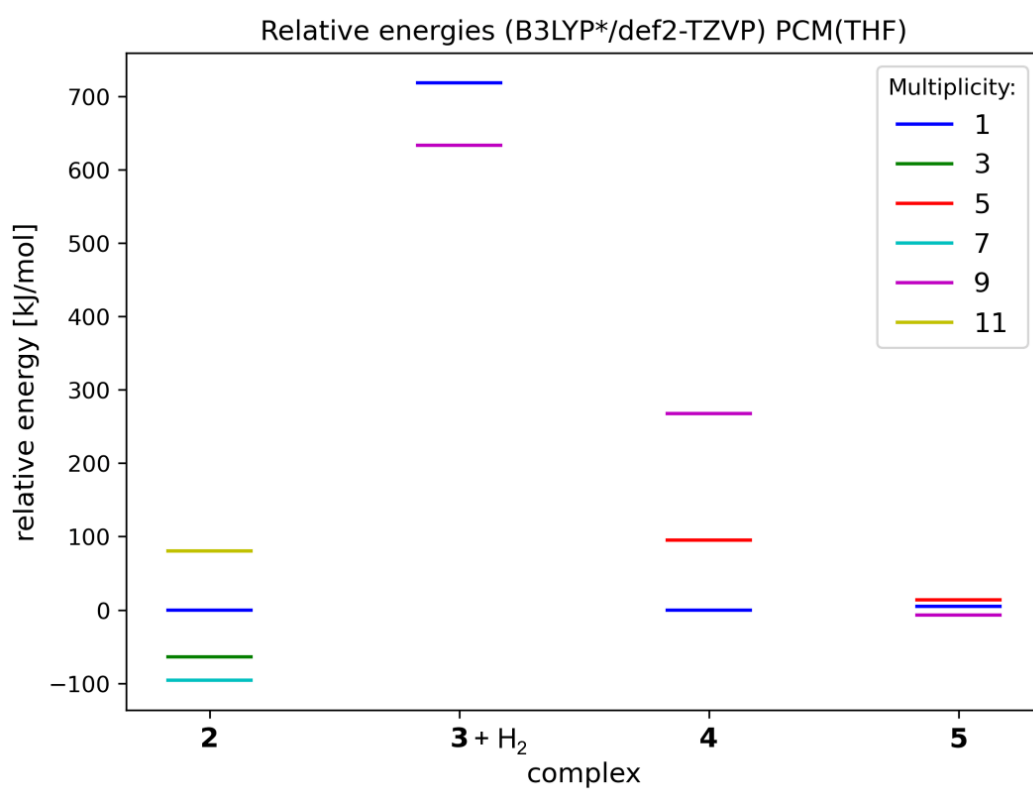


Figure D.28: Energies of spin states calculated with B3LYP*/def2-TZVP with PCM(THF) on crystal structures relative to loss of complex **2** for complexes **2** and **3** and relative to the loss of **14** for complexes **14** and **5**.

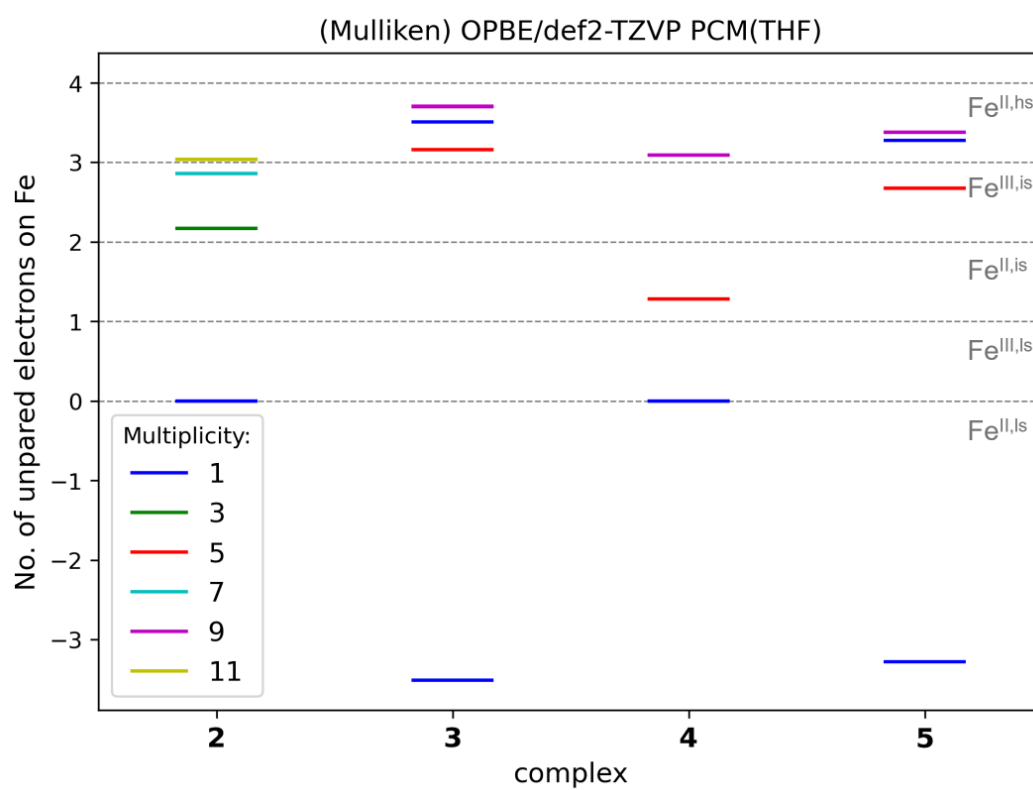


Figure D.29: Number of unpaired electrons on iron atoms calculated by Mulliken population analysis for different spin states (OPBE/def2-TZVP with PCM(THF)). The grey dashed lines indicate ideal values for different oxidation and spin states of the iron centers.

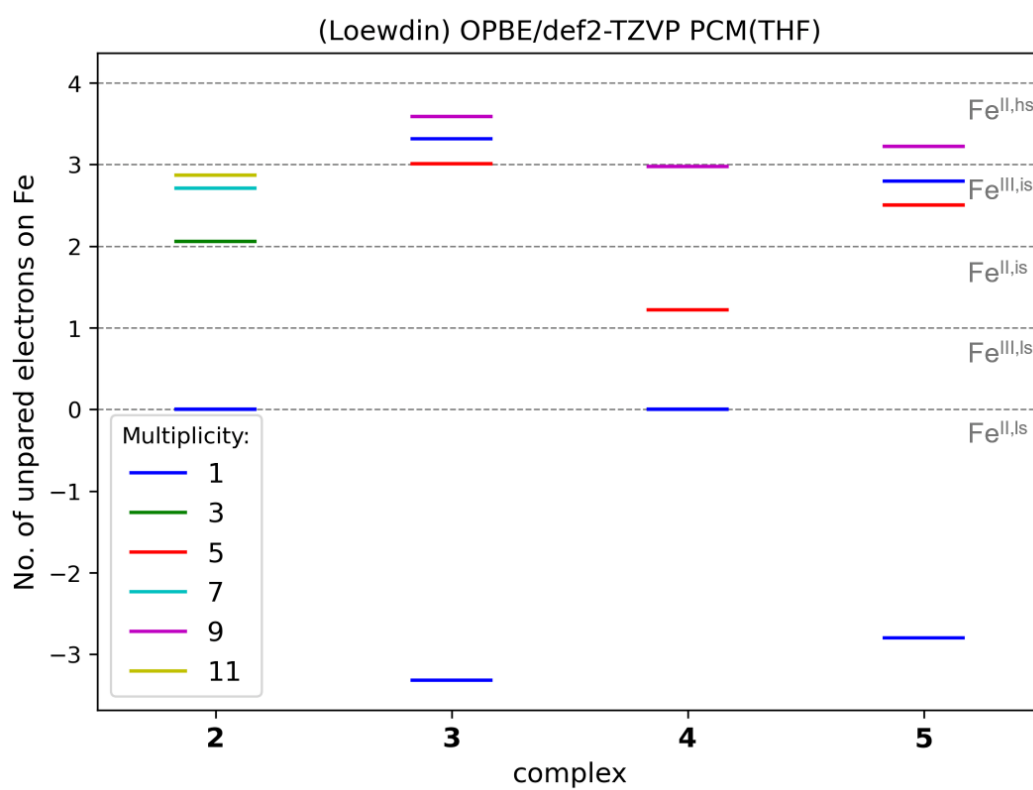


Figure D.30: Number of unpaired electrons on iron atoms calculated by Löwdin population analysis for different spin states (OPBE/def2-TZVP with PCM(THF)). The grey dashed lines indicate ideal values for different oxidation and spin states of the iron centers.

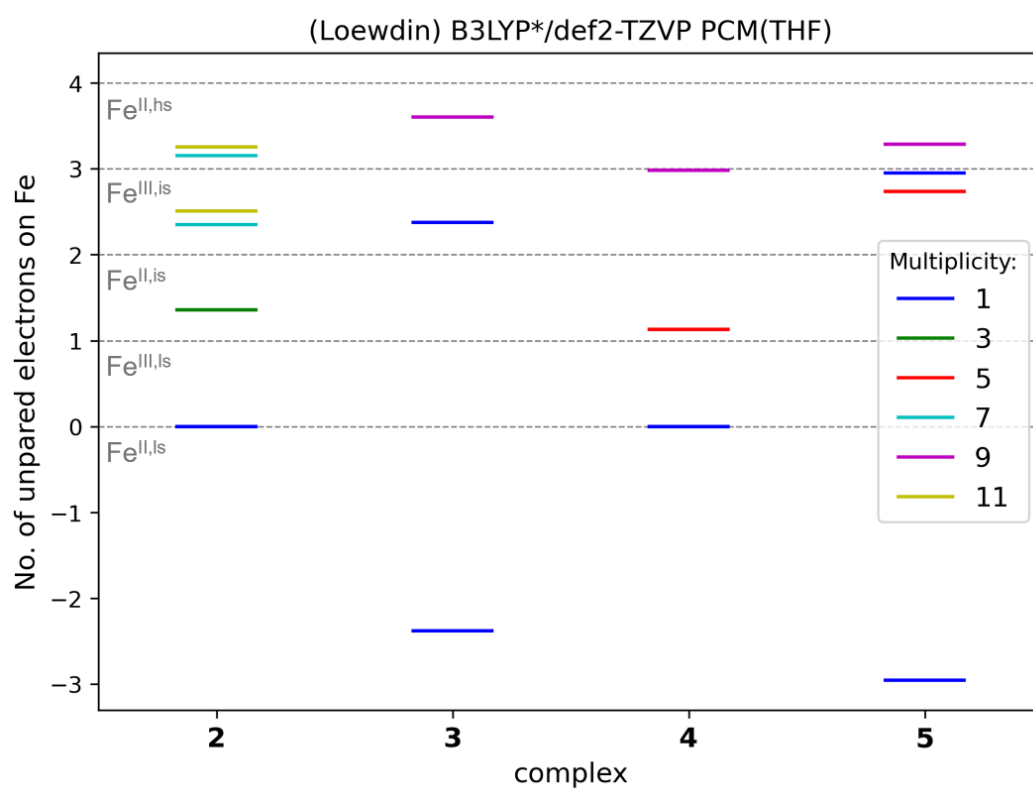














Figure D.31: Number of unpaired electrons on iron atoms calculated by Löwdin population analysis for different spin states (B3LYP*/def2-TZVP with PCM(THF)). The grey dashed lines indicate ideal values for different oxidation and spin states of the iron centers.

E Hazardous Materials

Material	GHS-Symbol	H-Statements	P-Statements
Acenaphthenequinone		315, 319, 335	261, 305+351+338
Acetonitrile		225, 302+312+332	210, 240, 302+352, 305+351+338, 403+233
Acetic acid		226, 290, 314	210, 280, 301+330+331, 305+351+338, 308+310
Benzene		225, 304, 315, 319, 340, 350, 372, 412	201, 210, 280, 308+313, 307+378, 403+233
Carbondioxide		280	403
Cobalt(II)-bromide		302, 317, 334, 341, 350, 410	201, 261, 280, 284, 304+340, 308+313
Dichlormethane		315, 319, 335, 336, 351, 373	261, 281, 305+351+338
Diethylether		225, 304, 315, 319, 340, 350, 372, 412	201, 210, 280, 308+313, 307+378, 403+235
Hydrogen		220, 280	210, 377, 381, 403
α -Methylstyrene		226, 319, 335, 411	261, 273, 305+351+338
Nitrogen (refrigerated, liquid)		281	282, 336+315, 403
Propan-2-ol		225, 319, 336	210, 233, 240, 305+351+338, 403+235












Material	GHS-Symbol	H-Statements	P-Statements
Trichlormethane- <i>d</i>		302, 315, 319, 331, 336, 351, 361d, 372	261, 281, 305+351+338, 311
Triphenylethylene		373, 302	260, 301+312
n-Hexane		225, 304, 361f, 373, 315, 315, 336, 411	210, 240, 273, 301+310, 331, 302+352, 403+235
Super-Hydride®-solution (1 M in THF)		225, 260, 302, 314, 335, 351, EUH014, EUH019	210, 231+232, 280, 370+378, 402+404, 403+235
Sodium triethylborohydride-solution (1 M in THF)		225, 261, 304, 314, 336, 361d, 373, 412	210, 231+232, 280, 301+330+331, 303+361+353, 304+340+310
Tetrahydrofuran		225, 302, 319, 335, 351, EUH019	210, 280, 301+312+330, 305+351+338, 370+378, 403+235
2,6-Diisopropylylaniline		319, 412	237, 305+351+338
Iron(II)chloride		302, 318	264, 270, 280, 301+312, 305+351+338, 501
1,8-Diazabicyclo[5.4.0]undec-7-ene		290, 301, 314, 412	234, 273, 280, 303+361+353, 304+340+310, 301+312, 305+351+338, 501
Carbon dioxide		280	410+403
Dimethylformamide		226, 312+332, 319, 360d	210, 280, 303+361+353, 304+340+310, 301+312, 305+351+338, 308+313

Table E.2: Used CMR-compounds of category 1a und 1b (GHS).

Cas-number	Name (IUPAC)	Procedure and used amount	Category (GHS)
865-49-6	Trichlormethan- <i>d</i>	Analysis 40 x 0.6 mL	K: 1B R: 1B
7789-43-7	Cobalt(II)- bromide	Synthesis 2 x 5 g	K: 1B
68-12-2	<i>N,N</i> -Dimethyl- formamide	Analysis 20 x 5 mg	R: 1B

Acronyms

AC	autocorrelation
AIM	atoms in molecules
ACSF	atom-centered symmetry functions
BIAN	bis(imino)acenaphthene
bp.	boiling point
calc.	calculated
cat.	catalyst
CASSCF	complete active space self consistent field
css	closed shell singlet
DCM	dichlormethane
DFT	density functional theory
Dipp	2,6-diisopropylphenyl
DME	1,2-dimethoxyethane
E_a	activation energy
EPR	electron paramagnetic resonance
equiv.	equivalent
ESI	electron spray ionization
Et	ethyl
FID	flame ionization detector
GC	gas chromatography
HF	Hartree–Fock
HOMO	highest occupied molecular orbital
IR	infrared
InChI	international chemical identifier
KDE	kernel density estimation
KRR	kernel ridge regression
KS	Kohn-Sham
LMBTR	local many-body tensor representation
LUMO	lowest unoccupied molecular orbital
ML	machine learning
MO	molecular orbital
MAE	mean absolute error

Acronyms

MBTR	many-body tensor representation
MS	mass spectrometry
MSE	mean signed error
MRCI	multireference configuration interaction
NMR	nuclear magnetic resonance
oss	open shell singlet
PCA	principle component analysis
RMSD	root-mean-square deviation
RMSE	root-mean-square error
RT	room temperature
s	singlet
SCF	self consistent field
SCO	spin crossover
SET	single electron transfer
SIE	self-interaction error
SOAP	smooth overlap of atomic positions
SQUID	superconducting quantum interference device
THF	tetrahydrofuran
TM	transition metal

Bibliography

1. Schwilk, M., Tahchieva, D. N. & von Lilienfeld, O. A. *Large yet bounded: Spin gap ranges in carbenes* 2020.
2. Czaplik, W. M., Mayer, M., Cvengros, J. & von Wangelin, A. J. Coming of age: sustainable iron-catalyzed cross-coupling reactions. *ChemSusChem* **2**, 396–417. ISSN: 1864-5631. doi:10.1002/cssc.200900055 (2009).
3. Chirik, P. J. Iron- and Cobalt-Catalyzed Alkene Hydrogenation: Catalysis with Both Redox-Active and Strong Field Ligands. *Acc. Chem. Res.* **48**, 1687–1695. doi:10.1021/acs.accounts.5b00134 (2015).
4. Bauer, I. & Knölker, H.-J. Iron catalysis in organic synthesis. *Chem. Rev.* **115**, 3170–3387. ISSN: 0009-2665. doi:10.1021/cr500425u (2015).
5. Holzwarth, M. S. & Plietker, B. Biorelevant Metals in Sustainable Metal Catalysis-A Survey. *ChemCatChem* **5**, 1650–1679. ISSN: 18673880. doi:10.1002/cctc.201200592 (2013).
6. Filonenko, G. A., van Putten, R., Schulpen, E. N., Hensen, E. J. M. & Pidko, E. A. Highly Efficient Reversible Hydrogenation of Carbon Dioxide to Formates Using a Ruthenium PNP-Pincer Catalyst. *ChemCatChem* **6**, 1526–1530. ISSN: 18673880. doi:10.1002/cctc.201402119 (2014).
7. Cassani, C., Bergonzini, G. & Wallentin, C.-J. Active Species and Mechanistic Pathways in Iron-Catalyzed C–C Bond-Forming Cross-Coupling Reactions. *ACS Catal.* **6**, 1640–1648. ISSN: 2155-5435. doi:10.1021/acscatal.5b02441 (2016).
8. Mako, T. L. & Byers, J. A. Recent advances in iron-catalysed cross coupling reactions and their mechanistic underpinning. *Inorg. Chem. Front.* **3**, 766–790. doi:10.1039/C5QI00295H (2016).
9. Piontek, A., Bisz, E. & Szostak, M. Iron-Catalyzed Cross-Couplings in the Synthesis of Pharmaceuticals: In Pursuit of Sustainability. *Angew. Chem. Int. Ed.* **57**, 11116–11128. ISSN: 0570-0833. doi:10.1002/anie.201800364 (2018).
10. Alig, L., Fritz, M. & Schneider, S. First-Row Transition Metal (De)Hydrogenation Catalysis Based On Functional Pincer Ligands. *Chem. Rev.* **119**, 2681–2751. ISSN: 0009-2665. doi:10.1021/acs.chemrev.8b00555 (2019).
11. Jacob, C. R. & Reiher, M. Spin in density-functional theory. *Int. J. Quantum Chem.* **112**, 3661–3684. ISSN: 00207608. doi:10.1002/qua.24309 (2012).
12. Brüggemann, J. & Jacob, C. R. Spin-state dependence of exchange-correlation holes. *Faraday Discuss.* doi:10.1039/D0FD00060D (2020).
13. Hehn, L., Deglmann, P. & Kühn, M. Chelate Complexes of 3d Transition Metal Ions A Challenge for Electronic-Structure Methods? *J. Chem. Theory. Comput.* **20**, 4545–4568. doi:10.1021/acs.jctc.3c01375 (2024).
14. Drabik, G. & Radoń, M. Approaching the Complete Basis Set Limit for Spin-State Energetics of Mononuclear First-Row Transition Metal Complexes. *J. Chem. Theory. Comput.* **20**, 3199–3217. doi:10.1021/acs.jctc.4c00092 (2024).
15. Jensen, K. P. & Cirera, J. Accurate computed enthalpies of spin crossover in iron and cobalt complexes. *J. Phys. Chem. A* **113**, 10033–10039. ISSN: 1089-5639. doi:10.1021/jp900654j (2009).

16. Ioannidis, E. I. & Kulik, H. J. Towards quantifying the role of exact exchange in predictions of transition metal complex properties. *J. Chem. Phys.* **143**, 034104. ISSN: 0021-9606. doi:10.1063/1.4926836 (2015).
17. Mortensen, S. R. & Kepp, K. P. Spin propensities of octahedral complexes from density functional theory. *J. Phys. Chem. A* **119**, 4041–4050. ISSN: 1089-5639. doi:10.1021/acs.jpca.5b01626 (2015).
18. Ioannidis, E. I. & Kulik, H. J. Ligand-Field-Dependent Behavior of Meta-GGA Exchange in Transition-Metal Complex Spin-State Ordering. *J. Phys. Chem. A* **121**, 874–884. ISSN: 1089-5639. doi:10.1021/acs.jpca.6b11930 (2017).
19. Siig, O. S. & Kepp, K. P. Iron(II) and Iron(III) Spin Crossover: Toward an Optimal Density Functional. *J. Phys. Chem. A* **122**, 4208–4217. ISSN: 1089-5639. doi:10.1021/acs.jpca.8b02027 (2018).
20. Radoń, M. Benchmarking quantum chemistry methods for spin-state energetics of iron complexes against quantitative experimental data. *Phys. Chem. Chem. Phys.* **21**, 4854–4870. ISSN: 1463-9076. doi:10.1039/c9cp00105k (2019).
21. Römer, A., Hasecke, L., Blöchl, P. & Mata, R. A. A Review of Density Functional Models for the Description of Fe(II) Spin-Crossover Complexes. *Molecules* **25**, 5176. ISSN: 1420-3049. doi:10.3390/molecules25215176 (2020).
22. Neale, S. E., Pantazis, D. A. & Macgregor, S. A. Accurate computed spin-state energetics for Co(III) complexes: implications for modelling homogeneous catalysis. *Dalton Trans.* doi:10.1039/d0dt00993h (2020).
23. Reiher, M., Salomon, O. & Artur Hess, B. Reparameterization of hybrid functionals based on energy differences of states of different multiplicity. *Theor. Chem. Acc.* **107**, 48–55. ISSN: 1432-881X. doi:10.1007/s00214-001-0300-3 (2001).
24. Pinter, B., Chankisjiev, A., Geerlings, P., Harvey, J. N. & de Proft, F. Conceptual Insights into DFT Spin-State Energetics of Octahedral Transition-Metal Complexes through a Density Difference Analysis. *Chem. Eur. J.* **24**, 5281–5292. ISSN: 0947-6539. doi:10.1002/chem.201704657 (2018).
25. Radoń, M. Revisiting the role of exact exchange in DFT spin-state energetics of transition metal complexes. *Phys. Chem. Chem. Phys.* **16**, 14479–14488. ISSN: 1463-9076. doi:10.1039/C3CP55506B (2014).
26. Woodworth, R. C. & Skell, P. S. Methylene, CH₂. Stereospecific Reaction with cis- and trans-2-Butene. *J. Am. Chem. Soc.* **81**, 3383–3386. doi:10.1021/ja01522a058 (1959).
27. Csizmadia, I. G., Font, J. & Strausz, O. P. Mechanism of the Wolff rearrangement. *J. Am. Chem. Soc.* **90**, 7360–7361. doi:10.1021/ja01028a041 (1968).
28. Schuster, G. B. in (eds Gold, V. & Bethell, D.) 311–361 (Academic Press, 1986). doi:[https://doi.org/10.1016/S0065-3160\(08\)60170-7](https://doi.org/10.1016/S0065-3160(08)60170-7).
29. Bourissou, D., Guerret, O., Gabbai, F. P. & Bertrand, G. Stable Carbenes. *Chem. Rev.* **100**, 39–92. ISSN: 0009-2665. doi:10.1021/cr940472u (2000).
30. Freeman, F., Lau, D. J., Patel, A. R., Pavia, P. R. & Willey, J. D. Substituent effects on singlet-triplet gaps and mechanisms of 1,2-rearrangements of 1,3-oxazol-2-ylidenes to 1,3-oxazoles. *J. Phys. Chem. A* **112**, 8775–8784. ISSN: 1089-5639. doi:10.1021/jp8030286 (2008).
31. Chu, J., Munz, D., Jazzar, R., Melaimi, M. & Bertrand, G. Synthesis of Hemilabile Cyclic (Alkyl)(amino)carbenes (CAACs) and Applications in Organometallic Chemistry. *J. Am. Chem. Soc.* **138**, 7884–7887. doi:10.1021/jacs.6b05221 (2016).
32. Modarelli, D. A. & Platz, M. S. Interception of dimethylcarbene with pyridine: a laser flash photolysis study. *J. Am. Chem. Soc.* **113**, 8985–8986. doi:10.1021/ja00023a079 (1991).

33. Dröge, T. & Glorius, F. The measure of all rings–N-heterocyclic carbenes. *Angew. Chem. Int. Ed.* **49**, 6940–6952. ISSN: 0570-0833. doi:10.1002/anie.201001865 (2010).
34. Duan, C., Janet, J. P., Liu, F., Nandy, A. & Kulik, H. J. Learning from Failure: Predicting Electronic Structure Calculation Outcomes with Machine Learning Models. *J. Chem. Theory. Comput.* **15**, 2331–2345. doi:10.1021/acs.jctc.9b00057 (2019).
35. Janet, J. P., Chan, L. & Kulik, H. J. Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network. *J. Phys. Chem. Lett.* **9**, 1064–1071. doi:10.1021/acs.jpcllett.8b00170 (2018).
36. Janet, J. P. *et al.* Designing in the Face of Uncertainty: Exploiting Electronic Structure and Machine Learning Models for Discovery in Inorganic Chemistry. *Inorg. Chem.* **58**, 10592–10606. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.9b00109 (2019).
37. Bahlke, M. P., Mogos, N., Proppe, J. & Herrmann, C. Exchange Spin Coupling from Gaussian Process Regression. *J. Phys. Chem. A* **124**, 8708–8723. ISSN: 1089-5639. doi:10.1021/acs.jpca.0c05983 (2020).
38. Taylor, M. G. *et al.* Seeing Is Believing: Experimental Spin States from Machine Learning Model Structure Predictions. *J. Phys. Chem. A* **124**, 3286–3299. ISSN: 1089-5639. doi:10.1021/acs.jpca.0c01458 (2020).
39. Eckhoff, M., Lausch, K. N., Blöchl, P. E. & Behler, J. Predicting oxidation and spin states by high-dimensional neural networks: Applications to lithium manganese oxide spinels. *J. Chem. Phys.* **153**, 164107. ISSN: 0021-9606. doi:10.1063/5.0021452 (2020).
40. Kulik, H. J. Making machine learning a useful tool in the accelerated discovery of transition metal complexes. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **10**. ISSN: 1759-0876. doi:10.1002/wcms.1439 (2020).
41. Cheng, L., Sun, J., Deustua, J. E., Bhethanabotla, V. C. & Miller, T. F. Molecular-orbital-based machine learning for open-shell and multi-reference systems with kernel addition Gaussian process regression. *J. Chem. Phys.* **157**, 154105. ISSN: 0021-9606. doi:10.1063/5.0110886 (2022).
42. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The Delta-Machine Learning Approach. *J. Chem. Theory. Comput.* **11**, 2087–2096. doi:10.1021/acs.jctc.5b00099 (2015).
43. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115. ISSN: 1098-0121. doi:10.1103/PhysRevB.87.184115 (2013).
44. Huang, B. & von Lilienfeld, O. A. Communication: Understanding molecular representations in machine learning: The role of uniqueness and target similarity. *J. Chem. Phys.* **145**, 161102. ISSN: 0021-9606. doi:10.1063/1.4964627 (2016).
45. Ghiringhelli, L. M., Vybiral, J., Levchenko, S. V., Draxl, C. & Scheffler, M. Big data of materials science: critical role of the descriptor. *Phys. Rev. Lett.* **114**, 105503. ISSN: 0031-9007. doi:10.1103/PhysRevLett.114.105503 (2015).
46. Darensbourg, M. Y. & Ash, C. E. in *Advances in organometallic chemistry* (eds Stone, F. G. A. & West, R.) 1–50 (Academic Press, San Diego, Calif. and London, 1987). ISBN: 9780120311279. doi:10.1016/S0065-3055(08)60025-8.
47. Nishimura, S. *Handbook of heterogeneous catalytic hydrogenation for organic synthesis* ISBN: 0471396982 (Wiley, New York, 2001).
48. De Vries, J. G. & Elsevier, C. J. *The Handbook of Homogeneous Hydrogenation* ISBN: 9783527311613. doi:10.1002/9783527619382 (Wiley, 2006).
49. Robinson, S. J. C. & Heinekey, D. M. Hydride & dihydrogen complexes of earth abundant metals: structure, reactivity, and applications to catalysis. *Chem. Commun.* **53**, 669–676. doi:10.1039/c6cc07529k (2017).

50. Dai, H. & Guan, H. Iron Dihydride Complexes: Synthesis, Reactivity, and Catalytic Applications. *Isr. J. Chem.* **57**, 1170–1203. ISSN: 0021-2148. doi:10.1002/ijch.201700101 (2017).
51. Yu, R. P., Darmon, J. M., Semproni, S. P., Turner, Z. R. & Chirik, P. J. Synthesis of Iron Hydride Complexes Relevant to Hydrogen Isotope Exchange in Pharmaceuticals. *Organometallics* **36**, 4341–4343. ISSN: 0276-7333. doi:10.1021/acs.organomet.7b00398 (2017).
52. Sandl, S. *et al.* Cobalt-Catalyzed Hydrogenations via Olefin Cobaltate and Hydride Intermediates. *ACS Catal.* **9**, 7596–7606. ISSN: 2155-5435. doi:10.1021/acscatal.9b01584 (2019).
53. Fedushkin, I. L., Skatova, A. A., Chudakova, V. A. & Fukin, G. K. Four-step reduction of dpp-bian with sodium metal: crystal structures of the sodium salts of the mono-, di-, tri- and tetraanions of dpp-bian. *Angew. Chem.* **42**, 3294–3298. doi:10.1002/ange.200351408 (2003).
54. Hill, N. J., Vargas-Baca, I. & Cowley, A. H. Recent developments in the coordination chemistry of bis(imino)acenaphthene (BIAN) ligands with s- and p-block elements. *Dalton Trans.*, 240–253. doi:10.1039/B815079F (2009).
55. Sandl, S. *Intermediates in Iron Group Metal-Catalyzed Hydrogenation Reactions* Dissertation (University of Regensburg, 2020).
56. Reinhold, J. *Quantentheorie der Moleküle: Eine Einführung* 4., überarb. u. erw. Aufl. 2013. ISBN: 9783834806307. doi:10.1007/978-3-8348-2037-2 (Springer, Wiesbaden, 2013).
57. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **136**, B864–B871. ISSN: 0031-899X. doi:10.1103/PhysRev.136.B864 (1964).
58. Born, M. & Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. d. Phys.* **389**, 457–484 (1927).
59. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, A1133–A1138. ISSN: 0031-899X. doi:10.1103/PhysRev.140.A1133 (1965).
60. Lee, Yang & Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens. Matter* **37**, 785–789. ISSN: 0163-1829. doi:10.1103/physrevb.37.785 (1988).
61. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **98**, 11623–11627. ISSN: 0022-3654. doi:10.1021/j100096a001 (1994).
62. Yanai, T., Tew, D. P. & Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393**, 51–57. ISSN: 00092614. doi:10.1016/j.cplett.2004.06.011 (2004).
63. Himanen, L. *et al.* DScibe: Library of descriptors for machine learning in materials science. *Comput. Phys. Commun.* **247**, 106949. ISSN: 0010-4655. doi:10.1016/j.cpc.2019.106949 (2020).
64. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
65. Rasmussen, C. E. & Williams, C. K. I. *Gaussian processes for machine learning* 3. print. ISBN: 9780262182539 (MIT Press, Cambridge, Mass., 2008).
66. Huo, H. & Rupp, M. Unified representation of molecules and crystals for machine learning. *Mach. Learn.: Sci. Technol.* **3**, 045017. doi:10.1088/2632-2153/aca005 (2022).

67. Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106. ISSN: 0021-9606. doi:10.1063/1.3553717 (2011).
68. Broto, P., Moreau, G. & Vandycke, C. Molecular structures: perception, autocorrelation descriptor and sar studies: system of atomic contributions for the calculation of the n-octanol/water partition coefficients. *Eur. J. Med. Chem.* **19**, 71–78 (1984).
69. Janet, J. P. & Kulik, H. J. Predicting electronic structure properties of transition metal complexes with neural networks. *Chem. Sci.* **8**, 5137–5152. ISSN: 2041-6520. doi:10.1039/c7sc01247k (2017).
70. Pernot, P. The parameter uncertainty inflation fallacy. *J. Chem. Phys.* **147**, 104102. ISSN: 0021-9606. doi:10.1063/1.4994654 (2017).
71. Rastetter, U., Jacobi von Wangelin, A. & Herrmann, C. Redox-active ligands as a challenge for electronic structure methods. *J. Comput. Chem.* **44**, 468–479. doi:10.1002/jcc.27013 (2023).
72. Kakar, S., Hoffman, F. G., Storz, J. F., Fabian, M. & Hargrove, M. S. Structure and reactivity of hexacoordinate hemoglobins. *Biophys. Chem.* **152**, 1–14. doi:10.1016/j.bpc.2010.08.008 (2010).
73. Kyne, S. H. *et al.* Iron and cobalt catalysis: new perspectives in synthetic radical chemistry. *Chem. Soc. Rev.* **49**, 8501–8542. ISSN: 0306-0012. doi:10.1039/D0CS00969E (2020).
74. Bolm, C. A new iron age. *Nat. Chem.* **1**, 420. ISSN: 1755-4349. doi:10.1038/nchem.315 (2009).
75. Chirik, P. J. & Wieghardt, K. Chemistry. Radical ligands confer nobility on base-metal catalysts. *Science* **327**, 794–795. doi:10.1126/science.1183281 (2010).
76. Nikolaevskaya, E. N., Druzhkov, N. O., Syroeshkin, M. A. & Egorov, M. P. Chemistry of diazadiene type ligands with extra coordination groups. Prospects of reactivity. *Coord. Chem. Rev.* **417**, 213353. ISSN: 00108545. doi:10.1016/j.ccr.2020.213353 (2020).
77. Bernauer, J., Pölker, J. & Jacobi von Wangelin, A. Redox-active BIAN-based Diimine Ligands in Metal-Catalyzed Small Molecule Syntheses**. *ChemCatChem* **14**. ISSN: 18673880. doi:10.1002/cctc.202101182 (2022).
78. Cramer, C. J. & Truhlar, D. G. Density functional theory for transition metals and transition metal chemistry. *Phys. Chem. Chem. Phys.* **11**, 10757–10816. ISSN: 1463-9076. doi:10.1039/b907148b (2009).
79. Sameera, W. M. C. & Maseras, F. Transition metal catalysis by density functional theory and density functional theory/molecular mechanics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 375–385. ISSN: 1759-0876. doi:10.1002/wcms.1092 (2012).
80. Ray, K., Petrenko, T., Wieghardt, K. & Neese, F. Joint spectroscopic and theoretical investigations of transition metal complexes involving non-innocent ligands. *Dalton Trans.*, 1552–1566. doi:10.1039/b700096k (2007).
81. Remenyi, C. & Kaupp, M. Where is the spin? Understanding electronic structure and g-tensors for ruthenium complexes with redox-active quinonoid ligands. *J. Am. Chem. Soc.* **127**, 11399–11413. doi:10.1021/ja051811b (2005).
82. Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **136**, 150901. ISSN: 0021-9606. doi:10.1063/1.4704546 (2012).
83. Verma, P. & Truhlar, D. G. Status and Challenges of Density Functional Theory. *Trends in Chemistry* **2**, 302–318. ISSN: 25895974. doi:10.1016/j.trechm.2020.02.005 (2020).
84. Simm, G. N., Proppe, J. & Reiher, M. Error Assessment of Computational Models in Chemistry. *Chimia* **71**, 202–208. ISSN: 0009-4293. doi:10.2533/chimia.2017.202 (2017).

85. Mortensen, J. J. *et al.* Bayesian error estimation in density-functional theory. *Phys. Rev. Lett.* **95**, 216401. ISSN: 0031-9007. doi:10.1103/PhysRevLett.95.216401 (2005).
86. Bergmann, T. G., Welzel, M. O. & Jacob, C. R. Towards theoretical spectroscopy with error bars: systematic quantification of the structural sensitivity of calculated spectra. *Chem. Sci.* **11**, 1862–1877. ISSN: 2041-6520. doi:10.1039/C9SC05103A (2019).
87. Weymuth, T. & Reiher, M. The transferability limits of static benchmarks. *Phys. Chem. Chem. Phys.* **24**, 14692–14698. ISSN: 1463-9076. doi:10.1039/D2CP01725C (2022).
88. Reiher, M. Molecule-specific Uncertainty Quantification in Quantum Chemical Studies. *Isr. J. Chem.* **62**. ISSN: 0021-2148. doi:10.1002/ijch.202100101 (2022).
89. Sun, Y. *et al.* Two-State Reactivity in Low-Valent Iron-Mediated C-H Activation and the Implications for Other First-Row Transition Metals. *J. Am. Chem. Soc.* **138**, 3715–3730. doi:10.1021/jacs.5b12150 (2016).
90. NIHEI, M., SHIGA, T., MAEDA, Y. & OSHIO, H. Spin crossover iron(III) complexes. *Coord. Chem. Rev.* **251**, 2606–2621. ISSN: 00108545. doi:10.1016/j.ccr.2007.08.007 (2007).
91. Weymuth, T., Couzijn, E. P. A., Chen, P. & Reiher, M. New Benchmark Set of Transition-Metal Coordination Reactions for the Assessment of Density Functionals. *J. Chem. Theory. Comput.* **10**, 3092–3103. doi:10.1021/ct500248h (2014).
92. Gritsenko, O. V., Schipper, P. R. T. & Baerends, E. J. Exchange and correlation energy in density functional theory: Comparison of accurate density functional theory quantities with traditional Hartree–Fock based ones and generalized gradient approximations for the molecules Li₂, N₂, F₂. *J. Chem. Phys.* **107**, 5007–5015. ISSN: 0021-9606. doi:10.1063/1.474864 (1997).
93. Handy, N. C. & Cohen, A. J. Left-right correlation energy. *Mol. Phys.* **99**, 403–412. ISSN: 0026-8976. doi:10.1080/00268970010018431 (2001).
94. Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem. Rev.* **112**, 289–320. ISSN: 0009-2665. doi:10.1021/cr200107z (2012).
95. Villa, M. *et al.* Synthesis and Catalysis of Redox-Active Bis(imino)acenaphthene (BIAN) Iron Complexes. *ChemCatChem* **9**, 3203–3209. ISSN: 18673880. doi:10.1002/cctc.201700144 (2017).
96. Khusniyarov, M. M., Weyhermüller, T., Bill, E. & Wieghardt, K. Reversible Electron Transfer Coupled to Spin Crossover in an Iron Coordination Salt in the Solid State. *Angew. Chem.* **120**, 1248–1251. doi:10.1002/ange.200704951 (2008).
97. Phan, H., Hrudka, J. J., Igimbayeva, D., Lawson Daku, L. M. & Shatruk, M. A Simple Approach for Predicting the Spin State of Homoleptic Fe(II) Tris-diimine Complexes. *J. Am. Chem. Soc.* **139**, 6437–6447. doi:10.1021/jacs.7b02098 (2017).
98. Nandy, A. *et al.* Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning. *Chem. Rev.* **121**, 9927–10000. ISSN: 0009-2665. doi:10.1021/acs.chemrev.1c00347 (2021).
99. Frisch, M. J. *et al.* *Gaussian 16 Revision B.01* Gaussian Inc. Wallingford CT. 2016.
100. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A Gen. Phys.* **38**, 3098–3100. ISSN: 0556-2791. doi:10.1103/physreva.38.3098 (1988).
101. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B Condens. Matter* **33**, 8822–8824. ISSN: 0163-1829. doi:10.1103/physrevb.33.8822 (1986).
102. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465. doi:10.1002/jcc.21759 (2011).

103. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305. ISSN: 1463-9076. doi:10.1039/b508541a (2005).
104. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057–1065. ISSN: 1463-9076. doi:10.1039/B515623H (2006).
105. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652. ISSN: 0021-9606. doi:10.1063/1.464913 (1993).
106. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241. ISSN: 1432-881X. doi:10.1007/s00214-007-0310-x (2008).
107. Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **91**, 146401. ISSN: 0031-9007. doi:10.1103/PhysRevLett.91.146401 (2003).
108. Staroverov, V. N., Scuseria, G. E., Tao, J. & Perdew, J. P. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **119**, 12129–12137. ISSN: 0021-9606. doi:10.1063/1.1626543 (2003).
109. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620. ISSN: 1463-9076. doi:10.1039/b810189b (2008).
110. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158–6170. ISSN: 0021-9606. doi:10.1063/1.478522 (1999).
111. Perdew, Burke & Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868. ISSN: 0031-9007. doi:10.1103/PhysRevLett.77.3865 (1996).
112. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **78**, 1396. ISSN: 0031-9007. doi:10.1103/PhysRevLett.78.1396 (1997).
113. Zhao, Y. & Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **125**, 194101. ISSN: 0021-9606. doi:10.1063/1.2370993 (2006).
114. Vydrov, O. A. & Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.* **125**, 234109. ISSN: 0021-9606. doi:10.1063/1.2409292 (2006).
115. Vydrov, O. A., Scuseria, G. E. & Perdew, J. P. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.* **126**, 154109. ISSN: 0021-9606. doi:10.1063/1.2723119 (2007).
116. *Computational Advances in Organic Chemistry: Molecular Structure and Reactivity* (eds Ögretir, C. & Csizmadia, I. G.) ISBN: 978-94-010-5439-3. doi:10.1007/978-94-011-3262-6 (Springer Netherlands, Dordrecht, 1991).
117. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 73–78. ISSN: 1759-0876. doi:10.1002/wcms.81 (2012).
118. Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**. ISSN: 1759-0876. doi:10.1002/wcms.1327 (2018).
119. Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **152**, 224108. ISSN: 0021-9606. doi:10.1063/5.0004608 (2020).

120. Grimme, S. & Hansen, A. A practicable real-space measure and visualization of static electron-correlation effects. *Angew. Chem. Int. Ed.* **54**, 12308–12313. ISSN: 0570-0833. doi:10.1002/anie.201501887 (2015).
121. Draxl, C. & Scheffler, M. The NOMAD laboratory: from data sharing to artificial intelligence. *J. Phys. Mater.* **2**, 036001 (2019).
122. Baranovski, V. I., Denisova, A. S. & Kuklo, L. I. Quantum chemical study of the electron affinity of the diimino derivatives of mono- and polycyclic organic molecules. *J. Mol. Struct.-Theochem* **759**, 111–115. ISSN: 01661280. doi:10.1016/j.theochem.2005.10.056 (2006).
123. Li, P., Wang, W.-h., Sun, H.-t. & Bi, S.-w. A DFT study on the electron affinity of tetrachloro-p-benzoquinone: Toward to understanding its electron-accepting ability in solution. *Comput. Theor. Chem.* **1006**, 127–132. ISSN: 2210271X. doi:10.1016/j.comptc.2012.11.022 (2013).
124. Bart, S. C., Hawrelak, E. J., Lobkovsky, E. & Chirik, P. J. Low-Valent alpha-Diimine Iron Complexes for Catalytic Olefin Hydrogenation. *Organometallics* **24**, 5518–5527. ISSN: 0276-7333. doi:10.1021/om050625b (2005).
125. Muresan, N. *et al.* Bis(alpha-diimine)iron complexes: electronic structure determination by spectroscopy and broken symmetry density functional theoretical calculations. *Inorg. Chem.* **47**, 4579–4590. ISSN: 0020-1669. doi:10.1021/ic7022693 (2008).
126. Herrmann, C., Reiher, M. & Hess, B. A. Comparative analysis of local spin definitions. *J. Chem. Phys.* **122**, 34102. ISSN: 0021-9606. doi:10.1063/1.1829050 (2005).
127. Harris. Simplified method for calculating the energy of weakly interacting fragments. *Phys. Rev. B Condens. Matter* **31**, 1770–1779. ISSN: 0163-1829. doi:10.1103/PhysRevB.31.1770 (1985).
128. Voigt, B. A., Steenbock, T. & Herrmann, C. *Structural diradical character* 2018.
129. Kröncke, S. & Herrmann, C. Designing Long-Range Charge Delocalization from First-Principles. *J. Chem. Theory. Comput.* **15**, 165–177. doi:10.1021/acs.jctc.8b00872 (2019).
130. Parthey, M. & Kaupp, M. Quantum-chemical insights into mixed-valence systems: within and beyond the Robin–Day scheme. *Chem. Soc. Rev.* **43**, 5067–5088 (2014).
131. Shee, J., Loipersberger, M., Hait, D., Lee, J. & Head-Gordon, M. Revealing the nature of electron correlation in transition metal complexes with symmetry breaking and chemical intuition. *J. Chem. Phys.* **154**, 194109. ISSN: 0021-9606. doi:10.1063/5.0047386 (2021).
132. Gimferrer, M., van der Mynsbrugge, J., Bell, A. T., Salvador, P. & Head-Gordon, M. Facing the Challenges of Borderline Oxidation State Assignments Using State-of-the-Art Computational Methods. *Inorg. Chem.* **59**, 15410–15420. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.0c02405 (2020).
133. Gimferrer, M., Aldossary, A., Salvador, P. & Head-Gordon, M. Oxidation State Localized Orbitals: A Method for Assigning Oxidation States Using Optimally Fragment-Localized Orbitals and a Fragment Orbital Localization Index. *J. Chem. Theory. Comput.* **18**, 309–322. doi:10.1021/acs.jctc.1c01011 (2022).
134. Hedegård, E. D., Knecht, S., Kielberg, J. S., Jensen, H. J. A. & Reiher, M. Density matrix renormalization group with efficient dynamical electron correlation through range separation. *J. Chem. Phys.* **142**, 224108. doi:10.1063/1.4922295 (2015).
135. Roemelt, M. & Pantazis, D. A. Multireference Approaches to Spin–State Energetics of Transition Metal Complexes Utilizing the Density Matrix Renormalization Group. *Adv. Theory Simul.* **2**, 1800201. ISSN: 2513-0390. doi:10.1002/adts.201800201 (2019).
136. Sharma, S., Sivalingam, K., Neese, F. & Chan, G. K.-L. Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics. *Nature Chem.* **6**, 927–933. doi:10.1038/nchem.2041 (2014).

137. Bochevarov, A. D. *et al.* Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem.* **113**, 2110–2142. ISSN: 00207608. doi:10.1002/qua.24481 (2013).
138. Herrmann, C., Podewitz, M. & Reiher, M. Restrained optimization of broken-symmetry determinants. *Int. J. Quantum Chem.* **109**, 2430–2446. ISSN: 00207608. doi:10.1002/qua.21997 (2009).
139. Arduengo, A. J., Harlow, R. L. & Kline, M. A stable crystalline carbene. *J. Am. Chem. Soc.* **113**, 361–363. doi:10.1021/ja00001a054 (1991).
140. Arduengo III, A. J., Goerlich, J. R., Krafczyk, R. & Marshall, W. J. 1,3,4,5-Tetraphenylimidazol-2-ylidene: The Realization of Wanzlick’s Dream. *Angew. Chem. Int. Ed.* **37**, 1963–1965. ISSN: 0570-0833. doi:10.1002/(SICI)1521-3773(19980803)37:13/14<1963::AID-ANIE1963>3.0.CO;2-M (1998).
141. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Sci. Data* **1**, 140022. ISSN: 2052-4463. doi:10.1038/sdata.2014.22 (2014).
142. Vargas, R., Galván, M. & Vela, A. Singlet-Triplet Gaps and Spin Potentials. *J. Phys. Chem. A* **102**, 3134–3140. ISSN: 1089-5639. doi:10.1021/jp972984t (1998).
143. Goerigk, L. *et al.* 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**, 32184–32215. ISSN: 1463-9076. doi:10.1039/C7CP04913G (2017).
144. Mardirossian, N. & Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **115**, 2315–2372. ISSN: 0026-8976. doi:10.1080/00268976.2017.1333644 (2017).
145. Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **108**, 058301. ISSN: 0031-9007. doi:10.1103/PhysRevLett.108.058301 (2012).
146. Hansen, K. *et al.* Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies. *J. Chem. Theory Comput.* **9**, 3404–3419. ISSN: 1549-9618. doi:10.1021/ct400195d (2013).
147. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555. doi:10.1038/s41586-018-0337-2 (2018).
148. Wu, Z. *et al.* MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.* **9**, 513–530. ISSN: 2041-6520. doi:10.1039/C7SC02664A (2018).
149. Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *J. Phys. Chem. Lett.* **11**, 2336–2347. doi:10.1021/acs.jpclett.9b03664 (2020).
150. Von Lilienfeld, O. A., Müller, K.-R. & Tkatchenko, A. Exploring chemical compound space with quantum-based machine learning. *Nat. Rev. Chem.* **4**, 347–358. doi:10.1038/s41570-020-0189-9 (2020).
151. Noé, F., Tkatchenko, A., Müller, K.-R. & Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **71**, 361–390. doi:10.1146/annurev-physchem-042018-052331 (2020).
152. Kulik, H. J. What’s Left for a Computational Chemist To Do in the Age of Machine Learning? *Isr. J. Chem.* **62**. ISSN: 0021-2148. doi:10.1002/ijch.202100016 (2022).
153. Duan, C., Nandy, A., Meyer, R., Arunachalam, N. & Kulik, H. J. A transferable recommender approach for selecting the best density functional approximations in chemical discovery. *Nat. Comput. Sci.* **3**, 38–47. doi:10.1038/s43588-022-00384-0 (2023).

154. Stewart, J. J. P. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **13**, 1173–1213. doi:10.1007/s00894-007-0233-4 (2007).
155. Kromann, J. C. *Calculate Root-mean-square deviation (RMSD) of Two Molecules Using Rotation, GitHub* <https://github.com/charnley/rmsd>. Version 1.4. Accessed: April 21, 2024. 2022.
156. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Cryst A* **32**, 922–923. ISSN: 0567-7394. doi:10.1107/S0567739476001873 (1976).
157. Walker, M. W., Shao, L. & Volz, R. A. Estimating 3-D location parameters using dual number quaternions. **54**, 358–367. ISSN: 10499660. doi:10.1016/1049-9660(91)90036-o (1991).
158. Ioannidis, E. I., Gani, T. Z. H. & Kulik, H. J. molSimplify: A toolkit for automating discovery in inorganic chemistry. *J. Comput. Chem.* **37**, 2106–2117. doi:10.1002/jcc.24437 (2016).
159. Nandy, A., Duan, C., Janet, J. P., Gugler, S. & Kulik, H. J. Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry. *Industrial & Engineering Chemistry Research* **57**, 13973–13986. ISSN: 0888-5885. doi:10.1021/acs.iecr.8b04015 (2018).
160. Heller, S. R., McNaught, A., Pletnev, I., Stein, S. & Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *J. Cheminform* **7**, 23. ISSN: 1758-2946. doi:10.1186/s13321-015-0068-4 (2015).
161. O’Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminform* **3**, 33. ISSN: 1758-2946. doi:10.1186/1758-2946-3-33 (2011).
162. Heinen, S., Schwilk, M., von Rudorff, G. F. & von Lilienfeld, O. A. Machine learning the computational cost of quantum chemistry. *Mach. Learn.: Sci. Technol.* **1**, 025002. doi:10.1088/2632-2153/ab6ac4 (2020).
163. Dapporto, P., Midollini, S. & Sacconi, L. Synthesis, characterization, and x-ray structure of confacial-bioctahedral iron(II)- and cobalt(II)-hydrido complexes with tridentate tripod ligands containing phosphorus and arsenic as the donor atoms. *Inorg. Chem.* **14**, 1643–1650. ISSN: 0020-1669. doi:10.1021/ic50149a043 (1975).
164. Bazhenova, T. A., Kachapina, L. M., Shilov, A. E., Antipin, M. & Struchkov, Y. Mono- and binuclear sigma-aryl iron-lithium hydrides; synthesis and molecular structure. *J. Organomet. Chem.* **428**, 107–123. ISSN: 0022328X. doi:10.1016/0022-328X(92)83223-5 (1992).
165. Ohki, Y. & Suzuki, H. [(5-C5Me5)Fe2(μ -H)4]: A Novel Dinuclear Iron Tetrahydrido Complex. *Angew. Chem. Int. Ed.* **39**, 3120–3122. ISSN: 0570-0833. doi:10.1002/1521-3773(20000901)39:17<3120::AID-ANIE3120>3.0.CO;2-K (2000).
166. Nowik, I. & Herber, R. H. Metal atom motional anisotropy in iron organometallics: the Gol’danskii–Karyagin effect in a dinuclear iron tetrahydrido complex. **64**, 313–317. ISSN: 00223697. doi:10.1016/S0022-3697(02)00305-0 (2003).
167. Walter, M. D., Grunenberg, J. & White, P. S. Reactivity studies on [Cp’FeI]2: From iron hydrides to P4-activation. *Chem. Sci.* **2**, 2120. ISSN: 2041-6520. doi:10.1039/c1sc00413a (2011).
168. Bai, X. *et al.* Synthesis, characterization and reactivity toward small molecules of a diiron tetrahydrido bridged complex. **122**, 108286. ISSN: 13877003. doi:10.1016/j.inoche.2020.108286 (2020).
169. Reiners, M. *et al.* Monomeric Fe(iii) half-sandwich complexes Cp’FeX2 - synthesis, properties and electronic structure. *Dalton Trans.* **47**, 10517–10526. doi:10.1039/c8dt01570h (2018).
170. Ohki, Y., Kojima, T., Oshima, M. & Suzuki, H. (5-C5Me5)Fe2(μ -H)2(μ -eta2:eta2-H2Si t Bu2), a Versatile Precursor for Bimetallic Active Species. *Organometallics* **20**, 2654–2656. ISSN: 0276-7333. doi:10.1021/om0101238 (2001).

171. Peldo, M. A., Beatty, A. M. & Fehlner, T. P. Routes to Compounds Containing M–B Bonds. Reaction of $[\text{Cp}^*\text{FeH}_2]_2$ with $\text{BH}_3 \cdot \text{THF}$, Yielding the Hydrogen-Rich arachno -Ferrapentaborane 1- $\text{Cp}^*\text{FeB}_4\text{H}_{11}$ ($\text{Cp}^* = \eta^5\text{-C}_5\text{Me}_5$). *Organometallics* **21**, 2821–2823. ISSN: 0276-7333. doi:10.1021/om020273y (2002).
172. Smith, J. M., Lachicotte, R. J. & Holland, P. L. N=N bond cleavage by a low-coordinate iron(II) hydride complex. *J. Am. Chem. Soc.* **125**, 15752–15753. doi:10.1021/ja038152s (2003).
173. Vela, J. *et al.* Synthesis and reactivity of low-coordinate iron(II) fluoride complexes and their use in the catalytic hydrodefluorination of fluorocarbons. *J. Am. Chem. Soc.* **127**, 7857–7870. doi:10.1021/ja0426721 (2005).
174. Yu, Y. *et al.* The reactivity patterns of low-coordinate iron-hydride complexes. *J. Am. Chem. Soc.* **130**, 6624–6638. doi:10.1021/ja710669w (2008).
175. Rodriguez, M. M., Bill, E., Brennessel, W. W. & Holland, P. L. N_2 reduction and hydrogenation to ammonia by a molecular iron-potassium complex. *Science* **334**, 780–783. doi:10.1126/science.1211906 (2011).
176. Chiang, K. P. *et al.* Characterization of the Fe–H bond in a three-coordinate terminal hydride complex of iron(I). *Angew. Chem. Int. Ed.* **51**, 3658–3662. ISSN: 0570-0833. doi:10.1002/anie.201109204 (2012).
177. McWilliams, S. F. *et al.* Dynamic effects on ligand field from rapid hydride motion in an iron(ii) dimer with an $S = 3$ ground state. *Chem. Sci.* **14**, 2303–2312. ISSN: 2041-6520. doi:10.1039/d2sc06412j (2023).
178. Hein, N. M., Pick, F. S. & Fryzuk, M. Synthesis and Reactivity of a Low-Coordinate Iron(II) Hydride Complex: Applications in Catalytic Hydrodefluorination. *Inorg. Chem.* **56**, 14513–14523. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.7b02199 (2017).
179. Lee, Y. *et al.* Reactivity of Hydride Bridges in High-Spin $3\text{M}-3(\mu\text{-H})$ Clusters ($\text{M} = \text{FeII}, \text{CoII}$). *J. Am. Chem. Soc.* **137**, 10610–10617. doi:10.1021/jacs.5b05204 (2015).
180. Hickey, A. K. *et al.* A Dimeric Hydride-Bridged Complex with Geometrically Distinct Iron Centers Giving Rise to an $S = 3$ Ground State. *J. Am. Chem. Soc.* **141**, 11970–11975. doi:10.1021/jacs.9b04389 (2019).
181. Cruz, T. F. C., Pereira, L. C. J., Waerenborgh, J. C., Veiros, L. F. & Gomes, P. T. Hydroboration of terminal olefins with pinacolborane catalyzed by new 2-iminopyrrolyl iron(II) complexes. *Catal. Sci. Technol.* **9**, 3347–3360. ISSN: 2044-4753. doi:10.1039/C8CY02319K (2019).
182. Thompson, C. V., Arman, H. D. & Tonzetich, Z. J. A Pyrrole-Based Pincer Ligand and Permits Access to Three Oxidation States of Iron in Organometallic Complexes. *Organometallics* **36**, 1795–1802. ISSN: 0276-7333. doi:10.1021/acs.organomet.7b00144 (2017).
183. Kato, T., Kuriyama, S., Nakajima, K. & Nishibayashi, Y. Catalytic C–H Borylation Using Iron Complexes Bearing 4,5,6,7-Tetrahydroisindol-2-ylidene-Based PNP-Type Pincer Ligand. *Chem. Asian J.* **14**, 2097–2101. ISSN: 1861-4728. doi:10.1002/asia.201900501 (2019).
184. Ott, J. C., Blasius, C. K., Wadepohl, H. & Gade, L. H. Synthesis, Characterization, and Reactivity of a High-Spin Iron(II) Hydrido Complex Supported by a PNP Pincer Ligand and Its Application as a Homogenous Catalyst for the Hydrogenation of Alkenes. *Inorg. Chem.* **57**, 3183–3191. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.7b03227 (2018).
185. Din Reshi, N. U. *et al.* Iron(I) and Iron(II) Amido-imidazolin-2-imine Complexes as Catalysts for H/D Exchange in Hydrosilanes. *ACS Catal.* **14**, 1759–1772. ISSN: 2155-5435. doi:10.1021/acscatal.3c05355 (2024).

186. Gieshoff, T. N., Chakraborty, U., Villa, M. & Jacobi von Wangelin, A. Alkene Hydrogenations by Soluble Iron Nanocluster Catalysts. *Angew. Chem. Int. Ed.* **56**, 3585–3589. ISSN: 0570-0833. doi:10.1002/anie.201612548 (2017).
187. Chakraborty, U. *et al.* Planar Iron Hydride Nanoclusters: Combined Spectroscopic and Theoretical Insights into Structures and Building Principles. **10**, 265–271. doi:10.1002/open.202000307 (2021).
188. Araake, R., Sakadani, K., Tada, M., Sakai, Y. & Ohki, Y. Fe₄ and Fe₆ Hydride Clusters Supported by Phosphines: Synthesis, Characterization, and Application in N₂ Reduction. *J. Am. Chem. Soc.* **139**, 5596–5606. doi:10.1021/jacs.7b01965 (2017).
189. Ott, J. C., Wadepohl, H., Enders, M. & Gade, L. H. Taking Solution Proton NMR to Its Extreme: Prediction and Detection of a Hydride Resonance in an Intermediate-Spin Iron Complex. *J. Am. Chem. Soc.* **140**, 17413–17417. doi:10.1021/jacs.8b11330 (2018).
190. Liu, T., Murphy, R. P., Carroll, P. J., Gau, M. R. & Tomson, N. C. C-C sigma-Bond Oxidative Addition and Hydrofunctionalization by a Macrocycle-Supported Diiron Complex. *J. Am. Chem. Soc.* **144**, 14037–14041. doi:10.1021/jacs.2c06266 (2022).
191. Gawron, M. *et al.* Counterion Effect in Cobaltate-Catalyzed Alkene Hydrogenation. *Angew. Chem. Int. Ed.* **63**, e202315381. ISSN: 0570-0833. doi:10.1002/anie.202315381 (2024).
192. Wekesa, F. S. *et al.* Iron-Catalyzed Hydrosilylation of Aldehydes and Ketones under Solvent-Free Conditions. *Organometallics* **34**, 5051–5056. ISSN: 0276-7333. doi:10.1021/acs.organomet.5b00630 (2015).
193. Laxson, W. W., Özkar, S., Folkman, S. & Finke, R. G. The story of a mechanism-based solution to an irreproducible synthesis resulting in an unexpected closed-system requirement for the LiBEt₃H-based reduction: The case of the novel subnanometer cluster, [Ir(1,5-COD)(μ-H)]₄, and the resulting improved, independently repeatable, reliable synthesis. *Inorganica Chim. Acta* **432**, 250–257. ISSN: 00201693. doi:10.1016/j.ica.2015.04.015 (2015).
194. Maier, T. M. *et al.* Amine-Borane Dehydrogenation and Transfer Hydrogenation Catalyzed by alpha-Diimine Cobaltates. *Chem. Eur. J.* **25**, 238–245. ISSN: 0947-6539. doi:10.1002/chem.201804811 (2019).
195. Larson, P. J. *et al.* Synthesis, characterization, electrochemical properties and theoretical calculations of (BIAN) iron complexes. *Polyhedron* **159**, 365–374. ISSN: 02775387. doi:10.1016/j.poly.2018.11.060 (2019).
196. Fedushkin, I. L. *et al.* New high-spin iron complexes based on bis(imino)acenaphthenes (BIAN): synthesis, structure, and magnetic properties. *Russ. Chem. Bull.* **62**, 2122–2131. ISSN: 1066-5285. doi:10.1007/s11172-013-0311-y (2013).
197. Fohlmeister, L. *et al.* Low-coordinate iron(I) and manganese(I) dimers: kinetic stabilization of an exceptionally short Fe-Fe multiple bond. *Angew. Chem. Int. Ed.* **51**, 8294–8298. ISSN: 0570-0833. doi:10.1002/anie.201203711 (2012).
198. Maity, A. K., Metta-Magaña, A. J. & Fortier, S. Donor Properties of a New Class of Guanidinate Ligands Possessing Ketimine Backbones: A Comparative Study Using Iron. *Inorg. Chem.* **54**, 10030–10041. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.5b01815 (2015).
199. Cotton, F., Daniels, L. M., Matonic, J. H. & Murillo, C. A. Highly distorted diiron(II, II) complexes containing four amidinate ligands. A long and short metal-metal distance. *Inorganica Chim. Acta* **256**, 277–282. ISSN: 00201693. doi:10.1016/S0020-1693(96)05470-9 (1997).
200. Soussi, K. *et al.* N-Alkyl substituted triazenide-bridged homoleptic iron(II) dimers with an exceptionally short Fe-Fe bond. *Dalton Trans.* **53**, 1439–1444. doi:10.1039/d3dt03132b (2024).

201. Tereniak, S. J. *et al.* Role of the metal in the bonding and properties of bimetallic complexes involving manganese, iron, and cobalt. *J. Am. Chem. Soc.* **136**, 1842–1855. doi:10.1021/ja409016w (2014).
202. Miller, D. L. *et al.* Redox Pairs of Diiron and Iron-Cobalt Complexes with High-Spin Ground States. *Inorg. Chem.* **55**, 9725–9735. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.6b01487 (2016).
203. Behlen, M. J. *et al.* Dinuclear oxidative addition reactions using an isostructural series of Ni₂, Co₂, and Fe₂ complexes. *Dalton Trans.* **46**, 5493–5497. doi:10.1039/c6dt04465d (2017).
204. Spentzos, A. Z. *et al.* Investigating Metal-Metal Bond Polarization in a Heteroleptic Tris-Ylide Diiron System. *Inorg. Chem.* **62**, 11487–11499. ISSN: 0020-1669. doi:10.1021/acs.inorgchem.3c01068 (2023).
205. Neese, F. Prediction and interpretation of the ⁵⁷Fe isomer shift in Mössbauer spectra by density functional theory. *Inorganica Chim. Acta* **337**, 181–192. ISSN: 00201693. doi:10.1016/S0020-1693(02)01031-9 (2002).
206. Römelt, M., Ye, S. & Neese, F. Calibration of modern density functional theory methods for the prediction of ⁵⁷Fe Mössbauer isomer shifts: meta-GGA and double-hybrid functionals. *Inorg. Chem.* **48**, 784–785. ISSN: 0020-1669. doi:10.1021/ic801535v (2009).
207. McWilliams, S. F., Brennan-Wydra, E., MacLeod, K. C. & Holland, P. L. Density Functional Calculations for Prediction of ⁵⁷Fe Mössbauer Isomer Shifts and Quadrupole Splittings in beta-Diketiminato Complexes. *ACS Omega* **2**, 2594–2606. ISSN: 2470-1343. doi:10.1021/acsomega.7b00595 (2017).
208. Gritzner, G. & Kuta, J. Recommendations on reporting electrode potentials in non-aqueous solvents (Recommendations 1983). *Pure Appl. Chem.* **56**, 461–466. ISSN: 0033-4545. doi:10.1351/pac198456040461 (1984).
209. Hoe, W.-M., Cohen, A. J. & Handy, N. C. Assessment of a new local exchange functional OPTX. *Chem. Phys. Lett.* **341**, 319–328. ISSN: 00092614. doi:10.1016/S0009-2614(01)00581-4 (2001).
210. Barone, V. & Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **102**, 1995–2001. ISSN: 1089-5639. doi:10.1021/jp9716997 (1998).
211. Dastgir, S., Coleman, K. S., Cowley, A. R. & Green, M. L. H. Synthesis, Structure, and Temperature-Dependent Dynamics of Neutral Palladium Allyl Complexes of Annulated Diaminocarbenes and Their Catalytic Application for C–C and C–N Bond Formation Reactions †. *Organometallics* **29**, 4858–4870. ISSN: 0276-7333. doi:10.1021/om1000327 (2010).
212. Rosa, V. *et al.* Synthesis, Solid-State Structures, and EPR Spectroscopic Studies on Polycrystalline and Single-Crystal Samples of alpha-Diimine Cobalt(II) Complexes. *Eur. J. Inorg. Chem.* **2006**, 4761–4769. ISSN: 1434-1948. doi:10.1002/ejic.200600448 (2006).
213. Rosa, V. *et al.* Synthesis, characterisation and solid state structures of alpha-diimine cobalt(II) complexes: Ethylene polymerisation tests. *J. Organomet. Chem.* **693**, 769–775. ISSN: 0022328X. doi:10.1016/j.jorganchem.2007.12.007 (2008).
214. Lu, T. & Chen, F. Multiwfn: a multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592. doi:10.1002/jcc.22885 (2012).

Acknowledgments

First and foremost, I would like to thank my supervisors, Carmen Herrman and Axel Jacobi von Wangelin. Both gave me the opportunity to conduct my research, combining experiment and theory. Thank you for all the guidance and support you provided in scientific as well as personal matters throughout the sometimes challenging times.

I am grateful to Michael Deffner for all the advice he shared on machine learning and Python-related matters. Thank you also for proofreading my thesis.

I would like to express my gratitude to Jie Qi Ng (Jackie) for the support with GC measurements and for proofreading.

I would like to thank Sebastian Sandl for sharing his research project with me. Thank you for all the discussions.

In the same manner, I would like to thank all my coworkers for always lending a helping hand. Thank you, Andrey Fedulin, Ben Bimberg, Bernhard Gregori, Constanze von Meyenn, Dieter Schaarschmidt, Eduard Hahn, Haitao Zhang, Jennifer Pölker, Johannes Strauch, Karen Schaefer, Leah Luxenberger, Luana Cardinale, Mattis Schmotz, Michaela Schneeberger, Nathalie Smith, Nils Nun, Ömer Sahin, Philipp Baltruschat, Philipp Weise, Robin Stuhr, Roderich Meißner, Sebastian Sandl, Sumit Naskar, Susanne Köncke, Thilo Baggendorf, Till Jahny, Tim-Oliver Kindler, Ulrich Potoschnig, and Yannik Apparius for the great time!

I would also like to thank all the service departments of the Chemistry department. Thank you, Frank Hoffmann and Isabelle Nevoigt, for your help with my crystals. I would like to thank the High-Performance Computing Center at the University of Hamburg for providing computational resources and the IT-service of the Chemistry department at the University of Hamburg for technical support.

For financial support, I would like to thank the Fonds der Chemischen Industrie (FCI).

Finally, I would like to thank my friends and family for all their support and encouragement. A special thanks to my husband, Tobias Rastetter, for always believing in me and being by my side.

Declaration and Oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Pinneberg den 19.08.2024, Ursula Rastetter