Computational Modeling of Complex Chemical Transformations in Homogeneous Catalysis

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Ingolf Harden

aus

Oldenburg (Oldb)

Bonn, April 2023

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter: Prof. Dr. Frank Neese

Gutachter: Prof. Dr. Stefan Grimme

Tag der Promotion: 22.08.2023

Erscheinungsjahr: 2023

List of Publications

Published:

Das, S.; Mitschke, B.; De, C. K.; Harden, I.; Bistoni, G.; List, B. Harnessing the ambiphilicity of silyl nitronates in a catalytic asymmetric approach to aliphatic β^3-amino acids *Nature Catalysis* **2021**, *4* (12), 1043-1049. https://doi.org/10.1038/s41929-021-00714-x.

Chatterjee, S.; Harden, I.; Bistoni, G.; Castillo, R. G.; Chabbra, S.; van Gastel, M.; Schnegg, A.; Bill, E.; Birrell, J. A.; Morandi, B.; Neese, F.; DeBeer, S. A Combined Spectroscopic and Computational Study on the Mechanism of Iron-Catalyzed Aminofunctionalization of Olefins Using Hydroxylamine Derived N–O Reagent as the "Amino" Source and "Oxidant". *Journal of the American Chemical Society* **2022**, *144* (6), 2637-2656. https://doi.org/10.1021/jacs.1c11083.

Harden, I.; Neese, F.; Bistoni, G. An induced-fit model for asymmetric organocatalytic reactions: a case study of the activation of olefins via chiral Brønsted acid catalysts. *Chemical Science* **2022**, *13* (30), 8848-8859. https://doi.org/10.1039/D2SC02274E.

Schümann, J. M.; Ochmann, L.; Becker, J.; Altun, A.; Harden, I.; Bistoni, G.; Schreiner, P. R. Exploring the Limits of Intramolecular London Dispersion Stabilization with Bulky Dispersion Energy Donors in Alkane Solution. *Journal of the American Chemical Society* **2023**, *145* (4), 2093-2097. https://doi.org/10.1021/jacs.2c13301.

Submitted:

Harden, I.; Neese, F.; Bistoni, G. Dimerization of Confined Brønsted Acids in Enantioselective Organocatalytic Reactions. *Manuscript submitted for publication* **2023**.

Abstract

Catalysis is involved in most manufactured products and hence it is highly important for supplying the society with energy, food, pharmaceuticals and other goods. In recent years, increasing attention has been paid to the development of “green” catalytic processes for the production of valuable chemical entities that are competitive with current industrial synthetic routes while reducing at the same time the use of hazardous chemicals and waste generation. Importantly, the catalyst has to be as selective as possible in order to facilitate only those reaction pathways that lead to the desired product. Quantum-chemical calculations provide fundamental insights into the mechanism of catalytic transformations and into the role of the catalyst. Some of the key challenges of contemporary *in silico* catalytic research are: modeling of large and flexible catalysts or substrates, deciphering the influence of noncovalent interactions on the selectivity of enantioselective transformations, elucidating environmental or cooperative catalytic effects, and describing accurately reaction intermediates with complicated electronic structures. In this thesis, computational multi-level protocols (including semi-empirical, density functional theory and wave-function based methods) are developed to address these key challenges in the field of selective homogeneous catalysis. The importance of noncovalent interactions for the stereoselectivity, the shape of the catalytic pocket and the reaction rates is discussed for the intramolecular hydroalkoxylation of small and unactivated olefins catalyzed by modern organocatalysts. To elucidate cooperative catalytic effects in enantioselective organocatalysis, the general dimerization mechanism for Brønsted acid catalysts of various sizes and structural features as well as the influence of catalyst dimerization effects on the formation of aliphatic β^3-amino acid derivates from silyl nitronates as a case study is investigated and rules of thumb are provided for determining under which conditions several catalyst molecules can participate in the rate- and/or selectivity-determining reaction steps. The accurate description of solute-solvent interactions is crucial for the modeling of homogeneously catalyzed reactions. This aspect is studied for the conformational preference of molecular balances in solution. Effects of the electronic structure on the mechanism and catalytic efficiency are unveiled considering the aminofunctionalization of styrene catalyzed by small and simple iron(II) catalysts as a case study.

Table of Contents

List of Figures

List of Tables

List of Acronyms

a.u. ..... Atomic units
ACDC ..... Asymmetric counter anion directed catalysis
AI-LFT ..... Ab initio ligand field theory
BCH ..... Baker-Campbell-Hausdorff
BINOL ..... 1,1'-Bi-2-naphtol
BSSE ..... Basis set superposition error
CAS-SCF ..... complete active space self-consistent field
CC ..... Coupled-Cluster
CEPA ..... Coupled electron pair
CI ..... Configuration interaction, Climbing image
CIP ..... Chiral ion pair
COT ..... Cyclooctatetraene
CPA ..... Chiral phosphoric acid
C-PCM ..... Conductor-like polarizable continuum model
CPS ..... Complete PNO space
CREST ..... Conformer-Rotamer Ensemble Sampling Tool
CSF ..... Configuration state function
DFT ..... Density functional theory
DID ..... Dispersion interaction density
DLPNO ..... Domain-based local pair natural orbital
DSI ..... Disulfonimide
DTPA ..... Dithiophosphoric acid
e.e. ..... Enantiomeric excess
e.g. ..... Exempli gratia
e.r. ..... Enantiomeric ratio
EPR ..... Electron paramagnetic resonance
eq. ..... Equation
equiv. ..... Equivalent
ESI-MS ..... Electrospray ionization mass spectrometry
etc. ..... Et cetera
FCI ..... Full configuration interaction
gCP ..... Geometrical counterpoise correction
GGA ..... Generalized gradient approximation
GUI ..... Graphical user interface
HF ..... Hartree-Fock
i.e. ..... Id est
IDP ..... Imidodiphosphate
IDPi ..... Imidodiphosphorimidate
IDPii ..... Imidodiphosphorbis(iminosulfonylimino)-imidate
iIDP ..... Iminoimidodiphosphate
IWAE ..... Intensity weighted average energy
KS ..... Kohn-Sham
LD ..... London dispersion
LED ..... Local energy decomposition
LMO ..... Localized molecular orbital
LSD ..... Local spin density
MB ..... Molecular balance
MD ..... Molecular dynamics

MTD ..... Metadynamics
NCI ..... Noncovalent interaction
NEB ..... Nudged elastic band
NEVPT2 ..... N-electron valence perturbation theory to second order
NMR ..... Nuclear magnetic resonance
NRVS ..... Nuclear resonance vibrational spectroscopy
ODE ..... Ordinary differential equation
PAO ..... Projected atomic orbital
PNO ..... Pair natural orbital
QCG ..... Quantum cluster growth
QRO ..... Quasi-restricted orbital
RMSD ..... Root-mean-square deviation
rR ..... Resonance Raman
RRHO ..... Rigid-rotor harmonic oscillator
RS ..... Range seperation
SAPT ..... Symmetry-adapted perturbation theory
SAS ..... Solvent accessible surface
TD-DFT ..... Time-dependent DFT
TS ..... Transition state
TSFF ..... Transition state force field
UV-vis ..... Ultraviolet-visible
XAS ..... X-ray absorption spectroscopy

Chapter 1 Introduction and Motivation

The term "catalysis" was defined by Wilhelm Ostwald in 1894 as "*Catalysis is the acceleration of a slow chemical process by the presence of a foreign material*" [1]. In catalytic reactions, the catalyst reduces the activation energy by enabling an alternative reaction mechanism. Since the catalyst participates in the reaction but is not consumed, only small, non-stoichiometric amounts of the catalyst are needed [2]. In our modern societies, catalysis is a cornerstone of nearly any area of chemical industry. Catalysis is involved in approximately 80 % of all manufactured products and hence it is highly important for supplying the society with energy, food, pharmaceuticals and other goods [3]. Famous catalytic processes in industry are for example the Fischer-Tropsch process [4] for the formation of liquid hydrocarbons from syngas or the Haber-Bosch process [5] for the synthesis of ammonia. In recent years, increasing attention has been paid to the development of "green" catalytic processes for valuable chemical entities that are competitive with current industrial synthetic routes while reducing at the same time the use of hazardous chemicals and waste generation [6]. Anastas and Eghbali developed twelve guiding principles enabling sustainability of chemical processes on the molecular level [7]. Among these principles are atom economy, design for energy efficiency, usage of renewable materials or avoidance of toxic substances. Clearly, catalysis plays a key role in this context by promoting more energy efficient reactions and reducing chemical waste. Importantly, the catalyst has to be as selective as possible in order to facilitate only the reaction pathways leading to the desired product.

Thanks to the continuous increase of computational power [8] and algorithm efficiency over the last decades [9,10], highly accurate electronic structure calculations are now routinely applicable to the systems of interest in contemporary catalysis research [11]. Quantum-chemical calculations provide fundamental insights into the mechanism of catalytic transformations and into the role of the catalyst. They can therefore also assist in rational catalyst development. Nowadays, electronic structure calculations are not only used for unveiling the mechanisms of synthetic catalysts, but also for deciphering the details of how biological systems like enzymes work [12]. However, many catalytic reactions remain challenging for standard computational techniques due to the complexity of the catalytic system and of its interaction with the environment. Some of the key challenges are the modeling of large and flexible catalysts or substrates, the deciphering of the influence of noncovalent interactions on the selectivity of enantioselective transformations, the elucidation of environmental or cooperative catalytic effects, or describing accurately reaction intermediates with complicated electronic structures [13,14]. Furthermore, the

accurate quantification of many important parameters of catalytic reactions, such as reaction rates or stereoselectivities, require computational accuracy with errors that are less than one kcal mol^{-1} [13]. To tackle these challenges, it is often necessary to devise sophisticated computational protocols that are optimized ad hoc to address specific chemical questions.

In this thesis, computational protocols are developed to address these key challenges in the field of selective homogeneous catalysis. The thesis is outlined as follows. An overview of modern electronic structure methods and asymmetric organocatalysis is given in Chapter 2 and Chapter 3, respectively. In Chapter 4, the mechanism and stereo-controlling factors of the intramolecular hydroalkoxylation of small and unactivated olefins catalyzed by modern organocatalysts is investigated and the importance of noncovalent interactions for the stereoselectivity, the shape of the catalytic pocket and the reaction rates is discussed. In Chapter 5, cooperative catalytic effects in enantioselective organocatalysis are discussed and rules of thumb are provided for determining under which conditions multiple catalyst molecules can participate in the rate- and/or selectivity-determining reaction steps. The first part of Chapter 5 describes the general dimerization mechanism for Brønsted acid catalysts of various sizes and structural features, while the second part focuses on the influence of catalyst dimerization effects on the formation of aliphatic β^3-amino acid derivates from silyl nitronates as a case study. As discussed before, interactions with the solvent mark another important influence of the environment on the reaction. The accurate description of solute-solvent interactions is crucial for the modeling of homogeneously catalyzed reactions. However, this might be challenging for mainstream computational techniques. This aspect is investigated in Chapter 6, where the key solute-solvent and *intra*-solute interactions responsible for the conformational preference of molecular balances in solution are investigated. Finally, in Chapter 7, effects of the electronic structure on the mechanism and catalytic efficiency are described considering the aminofunctionalization of styrene catalyzed by small and simple iron(II) catalysts as a case study.

Chapter 2 A Journey into Modern Electronic Structure Methods

2.1 The electronic Schrödinger equation

On the microscopic scale, the properties and behavior of matter are almost entirely governed by electromagnetic interaction [15]. Since strong and weak nuclear interactions only dominate over very short distances, they are most important within the nuclear radius. On the other hand, electrons and nuclei have such low masses that gravitational interactions can be neglected, making the electromagnetic interaction the only fundamental interaction between electrons and nuclei to consider [16]. A system of *N* electrons and *K* nuclei is described by a wave function $\Psi(\vec{x}_1, \ldots, \vec{x}_N, \vec{R}_1, \ldots, \vec{R}_K, t)$ with $\vec{x}_i$ combining the position $\vec{r}_i$ and the spin σ_i of the *i*-th electron, $\vec{R}_j$ the position of the *j*-th nucleus and *t* the time. The interaction between the particles can be described by the Dirac equation [17] that is given in equation (eq.) (2.1).

$$\hat{H}_D|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k, t)\rangle = i\hbar \frac{\partial|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k, t)\rangle}{\partial t}. \tag{2.1}$$

$\hat{H}_D$ is the Dirac operator acting on the wave function, $\hbar$ is the reduced Planck constant and i is the imaginary unit. For the representation of the wave function, the so-called bra-ket notation is used. The wave function is represented in the Hilbert space of the electronic coordinates, making $|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k, t)\rangle$ dependent on the nuclear coordinates and time. The Dirac equation includes Einstein's theory of special relativity, the exact form of $\hat{H}_D$ is beyond the scope of this thesis. In the non-relativistic limit, the Dirac equation is replaced by the Schrödinger equation (eq. (2.2)).

$$\hat{H}|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k, t)\rangle = i\hbar \frac{\partial|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k, t)\rangle}{\partial t}. \tag{2.2}$$

Many properties of matter, like energies or equilibrium geometries of molecules and solids, do not depend on the time. The time dependence of the wave function can be separated out, leading to the time independent Schrödinger equation [18] (eq. (2.3)).

$$\hat{H}|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k)\rangle = E_n|\Psi_n(\vec{R}_1, \ldots, \vec{R}_k)\rangle. \tag{2.3}$$

Thus, the differential equation (2.2) has been turned into the eigenvalue equation (2.3) with E_n being the *n*-th eigenvalue. E_n is interpreted as the total energy of the system in the *n*-th excited state. In eq. (2.4), the spatial representation of $\hat{H}$ in the absence of external electromagnetic fields is given.

$$\hat{H} = \sum_{i=1}^{N} \frac{-\hbar^2}{2m_e} \vec{\nabla}_i^2 + \sum_{j=1}^{K} \frac{-\hbar^2}{2m_j} \vec{\nabla}_j^2 + \frac{1}{2} \sum_{k \neq l}^{K} \frac{e^2 Z_k Z_l}{4\pi\varepsilon_0 |\vec{R}_k - \vec{R}_l|} + \frac{1}{2} \sum_{i \neq j}^{N} \frac{e^2}{4\pi\varepsilon_0 |\vec{r}_i - \vec{r}_j|} - \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{e^2 Z_j}{4\pi\varepsilon_0 |\vec{r}_i - \vec{R}_j|}. \quad (2.4)$$

m_e is the electron mass, m_j is the mass of the *j*-th nucleus, $\vec{\nabla}$ is the Del operator, e is the elementary charge, Z_k is the charge of the *k*-th nucleus, ε_0 is the vacuum permittivity. The sum $\sum_{a \neq b}$ is a short notion for a double sum over a und b with the condition $a \neq b$. The factor $\frac{1}{2}$ accounts for double counting of each interaction term. Eq. (2.4) describes the combined kinetic and potential energy of all particles in the system. $\sum_{i=1}^{N} \frac{-\hbar^2}{2m_e} \vec{\nabla}_i^2$ is the sum of the kinetic energies of all electrons, $\sum_{j=1}^{K} \frac{-\hbar^2}{2m_e} \vec{\nabla}_j^2$ is sum of the kinetic energies of all nuclei. The next three terms describe the coulombic interaction between a) any nucleus with any other nucleus, b) any electron with any other electron and c) any electron with any nucleus.

In order to reduce the number of variables, the electron coordinates can be separated from the nuclei coordinates. The corresponding electronic Hamiltonian $\hat{H}_e$ is obtained by ignoring the kinetic energy of the nuclei in eq. (2.4).

$$\hat{H}_e = \hat{H} - \sum_{j=1}^{K} \frac{-\hbar^2}{2m_j} \vec{\nabla}_j^2. \quad (2.5)$$

The corresponding electronic Schrödinger equation (eq. (2.6)) describes the movement and interaction of the electrons in the Coulomb field of fixed nuclei.

$$\hat{H}_e |\Psi_n(\vec{R}_1, \ldots, \vec{R}_k)\rangle = E_{en} |\Psi_n(\vec{R}_1, \ldots, \vec{R}_k)\rangle. \quad (2.6)$$

Note that eq. (2.6) contains the positions of the nuclei as parameters instead of variables, which will be made implicit in the further. Finding approximate but accurate solutions for eq. (2.6) for many-particle systems (such as molecules), is one of the major targets of theoretical chemistry and will be the main topic of the subsequent sections.

2.2 The Hartree-Fock approach

For solving eq. (2.6) for a given set of nuclei, multiple approaches have risen. Since the Hartree-Fock (HF) approach is the starting point for more advanced techniques, the basic theory will be discussed here briefly. As prerequisite, the variational principle of Ritz [19] is introduced. Given a well-conditioned (*i.e.*, normalizable) trial wave function $\tilde{\Psi}_0(\vec{x}_1, \ldots, \vec{x}_N)$ for the electronic ground state, the energy $\tilde{E}_0$ of this wave function will always be higher than the "true" energy

E_0 that is associated with the true wave function $\Psi_0(\vec{x}_1, \ldots, \vec{x}_N)$ of the system, unless $\tilde{\Psi}_0(\vec{x}_1, \ldots, \vec{x}_N) = \Psi_0(\vec{x}_1, \ldots, \vec{x}_N)$.

$$\tilde{E}_0 = \frac{\langle \tilde{\Psi}_0 | \hat{H}_e | \tilde{\Psi}_0 \rangle}{\langle \tilde{\Psi}_0 | \tilde{\Psi}_0 \rangle} \geq \frac{\langle \Psi_0 | \hat{H}_e | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}. \tag{2.7}$$

A simple ansatz for $\Psi_0(\vec{x}_1, \ldots, \vec{x}_N)$ is to write the many-particle wave function as product of orthonormal one-particle states, also called spin orbitals, $\psi_i(\vec{x})$.

$$\Psi_0(\vec{x}_1, \ldots, \vec{x}_N) = \prod_{i=1}^{N} \psi_i(\vec{x}_i). \tag{2.8}$$

Eq. (2.8) is called the Hartree product, which does not satisfy the Pauli principle [20]. However, the product state can be antisymmetrized by using a Slater determinant instead [21].

$$\Psi_0(\vec{x}_1, \ldots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{x}_1) & \cdots & \psi_N(\vec{x}_1) \\ \cdots & \cdots & \cdots \\ \psi_1(\vec{x}_N) & \cdots & \psi_N(\vec{x}_N) \end{vmatrix}. \tag{2.9}$$

Exchange of two particles corresponds to exchanging two rows of the Slater determinant, resulting in a sign change. The prefactor $\frac{1}{\sqrt{N!}}$ insures normality of the many-particle wave function. To obtain working equations for the optimal one-electron states, the energy (which is the expectation value of the electronic Hamiltonian formed with the Slater determinant) has to be minimized with respect to the spin orbitals. The derivation is not shown here, but can be found in literature [22]. The expectation value of $\hat{H}_e$ with the ansatz from eq. (2.9) is the Hartree-Fock ground state energy E_{HF} given by eq. (2.10).

$$\begin{aligned} E_{HF} &= \langle \Psi_0 | \hat{H}_e | \Psi_0 \rangle = \sum_{i=1}^{N} I_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} J_{ij} - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} K_{ij} \\ I_i &= \int \psi_i^*(\vec{x}) \left(\frac{-\hbar^2}{2m_e} \vec{\nabla}_i^2 - \sum_{j=1}^{K} \frac{e^2 Z_j}{4\pi\varepsilon_0 |\vec{r} - \vec{R}_j|} \right) \psi_i(\vec{x}) d^4\vec{x} \\ J_{ij} &= \iint \psi_i^*(\vec{x}) \psi_j^*(\vec{x}') \frac{e^2}{4\pi\varepsilon_0 |\vec{r} - \vec{r}'|} \psi_i(\vec{x}) \psi_j(\vec{x}') d^4\vec{x} d^4\vec{x}' \\ K_{ij} &= \iint \psi_i^*(\vec{x}) \psi_j^*(\vec{x}') \frac{e^2}{4\pi\varepsilon_0 |\vec{r} - \vec{r}'|} \psi_i(\vec{x}') \psi_j(\vec{x}) d^4\vec{x} d^4\vec{x}'. \end{aligned} \tag{2.10}$$

The integrals I_i in eq. (2.10) are labeled one-electron integrals as they describe the kinetic energy of the electron and the attractive Coulomb interaction with all nuclei. The integrals J_{ij} are called Coulomb integrals and depend on the coordinates of two electrons. These integrals describe the classical electrostatic interaction between two charge distributions. Note that the

unphysical self-interaction of each orbital is included in the term J_{ii}. Since $J_{ii} = K_{ii}$, the self-interaction has no influence on the total Hartree-Fock energy. The integrals K_{ij} are called exchange integrals that occur because an antisymmetrized ansatz for the many-particle wave function was chosen. They do not occur in the corresponding energy expression with the ansatz from eq. (2.8) [23]. The exchange integrals are only non-zero for electrons with same spin orientation. However, they do not describe a particular interaction, but are a consequence of a fundamental symmetry principle [24] – that the many-particle wave function of indistinguishable fermions must change its sign upon exchanging two particles. Making the energy stationary with respect to the one-electron wave functions with the condition of orthonormal one-electron functions yields the Hartree-Fock equation given in eq. (2.11) in its canonical form (*i.e.*, the matrix of Lagrange multipliers is diagonal) [22].

$$\hat{f}|\psi_k\rangle = \varepsilon_k|\psi_k\rangle, \tag{2.11}$$

with ε_k being the eigenvalues of the Fock operator and

$$\begin{aligned}\hat{f}\psi_k(\vec{x}) &= \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}^2 - \sum_{j=1}^{K}\frac{e^2 Z_j}{4\pi\varepsilon_0|\vec{r} - \vec{R}_j|}\right)\psi_k(\vec{x}) \\ &+ \left(\sum_{i=1}^{N}\int\frac{e^2{\psi_i}^*(\vec{x}')\psi_i(\vec{x}')}{4\pi\varepsilon_0|\vec{r}-\vec{r}'|}d^4\vec{x}'\right)\psi_k(\vec{x}) \\ &- \sum_{i=1}^{N}\left(\int\frac{e^2{\psi_i}^*(\vec{x}')\psi_k(\vec{x}')}{4\pi\varepsilon_0|\vec{r}-\vec{r}'|}d^4\vec{x}'\right)\psi_i(\vec{x})\end{aligned} \tag{2.12}$$

being the Fock operator $\hat{f}$ acting on the one-electron function $\psi_k(\vec{x})$. The second term in eq. (2.12) describes the action of the Coulomb operator on the *k*-th spin orbital. It describes the electrostatic interaction of the electron in $\psi_k(\vec{x})$ with the charge densities of all other electrons. Note that also the self-interaction of $\psi_k(\vec{x})$ is included. The last term in eq. (2.12) is the action of the exchange operator on $\psi_k(\vec{x})$ that carries out the particle exchange between $\psi_k(\vec{x})$ and $\psi_i(\vec{x})$, making the exchange operator non-local. The physical interpretation of the eigenvalues ε_i is given by Koopmans Theorem [25], which states that in the frozen orbital approximation the energy needed for removing an electron from $|\psi_k\rangle$ is given by $-\varepsilon_k$. Since the Fock operator describes the pair-wise averaged (due to the integration over all positions $\vec{r}'$ of the second electron in eq. (2.12)) electrostatic interaction of the reference electron at position $\vec{r}$ with every other electron, E_{HF} is larger than E_0, which is the true ground state energy of the system (the

lowest eigenvalue of the electronic Hamiltonian). The difference is called the correlation energy E_{Corr}, which is negative by definition.

$$E_{Corr} = E_0 - E_{HF}. \tag{2.13}$$

The following sections are attributed to the approximate computation of E_{Corr}. Since the evaluation of the Fock operator requires the knowledge of all orbitals, the Hartree-Fock equations are solved in a self-consistent fashion: in the first step, guess orbitals are constructed. Afterwards, the Fock operator is built, the Hartree-Fock equations are solved and a new set of orbitals is obtained. This procedure is repeated until the total energy and the orbitals do not change anymore. The corresponding working equations are the Roothaan-Hall equations [26] for the spin-restricted Hartree-Fock formalism and the Pople-Nesbet equations [27] for the spin-unrestricted Hartree-Fock formalism.

It should be noted here that Slater determinants are generally eigenfunctions of the spin projection operator $\hat{S}_z$, but not eigenfunctions of the total electron spin operator $\hat{S}^2$. There are few exceptions to this: closed-shell (that is, spin-restricted with orbital occupation numbers of either 0 or 2) Slater determinants are pure singlet states. Restricted open shell determinants with all unpaired electrons having the same spin orientation are eigenfunctions of $\hat{S}^2$ as well [22]. Other Slater determinants like unrestricted determinants are not eigenfunctions of $\hat{S}^2$. However, linear combinations of a small number of determinants lead to so-called configuration state functions (CSF), which are proper eigenfunctions of $\hat{S}^2$ [28]. For a given electron configuration, the number of CSFs is much smaller than the corresponding number of Slater determinants, making CSFs interesting as expansion sets for configuration interaction or multireference self-consistent field (like complete active space self-consistent field, CAS-SCF) approaches. In the following section, however, the formalisms are introduced with Slater determinants, so it is important to keep in mind that the use of CFSs is possible as well.

2.3 Configuration interaction and Coupled-Cluster approach

2.3.1 Configuration interaction

Diagonalization of the Fock matrix (that is, the matrix representation of the Fock operator in the atomic orbital basis) yields in principle infinitely many spin orbitals. The lowest *N* (or *N/2* for spin-restricted cases) orbitals are used for building the Slater determinant and therefore the approximate ground state wave function. These orbitals are called occupied or internal. All other orbitals are called virtual orbitals. The simplest variational approach to improve the

ground state wave function and energy is called configuration interaction (CI) [29]. By replacing occupied orbitals with virtual orbitals in the Slater determinant, a distinct Slater determinant from the ground state determinant, that is, an excited Slater determinant, is obtained. The CI wave function is written as linear combination of such excited Slater determinants

$$|\Psi_{CI}\rangle \;=\; c_0|\Psi_{HF}\rangle + \sum_{k=1} c_k|\Psi_k\rangle, \tag{2.14}$$

with yet to be determined coefficients c_j. The excited Slater determinants $\left|\Psi_{j>0}\right\rangle$ can be categorized by how many occupied orbitals have been replaced by virtual orbitals, leading to singly excited determinants $|\Psi_i^a\rangle$ (the *i*-th occupied orbital is replaced by the *a*-th virtual orbital), doubly excited determinants $\left|\Psi_{ij}^{ab}\right\rangle$ (the *i*-th and *j*-th occupied orbitals are replaced by the *a*-th and *b*-th virtual orbital), triply excited determinants $\left|\Psi_{ijk}^{abc}\right\rangle$ and so on. The optimal coefficients c_j are obtained variationally, leading to another eigenvalue problem [15].

$$\begin{aligned} \boldsymbol{Hc} &= \boldsymbol{Ec} \\ H_{ij} &= \left\langle\Psi_i\middle|\hat{H}_e\middle|\Psi_j\right\rangle. \end{aligned} \tag{2.15}$$

The matrix $\boldsymbol{H}$ is the representation of the Hamiltonian in the basis of all included Slater determinants. The matrix elements H_{ij} can be obtained by means of matrix elements of the orbitals. The corresponding rules for calculating H_{ij} are the Slater-Condon rules [21,30,31] and the Brillouin theorem [32]. The CI energy is the lowest eigenvalue of the diagonal matrix $\boldsymbol{E}$. As the number of Slater determinants scales factorially with the system size, in practice restriction to few excitation types (*e.g.*, only single excitations, CIS; only double excitations, CID; single and double excitations CISD; single, double and triple excitations CISDT, etc.) is carried out. If all possible Slater determinants (for a finite set of orbitals there will be a finite number of possible Slater determinants) are taken into account, one speaks of Full CI (FCI). The truncated CI approaches, opposed to FCI, suffer from size inconsistency [33]: the energy of two non-interacting fragments A and B within one system is different from the sum of the energies of the isolated fragments. The reason is that certain excitation types occurring on each isolated fragment cannot occur simultaneously on both fragments of the combined system.

2.3.2 Coupled-Cluster approach

The problem of size inconsistency can be overcome by the so-called cluster approach. Coupled-Cluster (CC) approaches were first used in nuclear physics [34,35] and entered electronic structure theory in the late 1960s [36,37]. In the first step, the excitation operator is defined in eq. (2.16), which generates the excited Slater determinants known from the previous section 2.3.1.

$$
\begin{aligned}
\hat{T} &= \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N \\
\hat{T}_1|\Psi_0\rangle &= \sum_{i}^{N_{occ}} \sum_{a}^{N_{vir}} t_i^a \, |\Psi_i^a\rangle \\
\hat{T}_2|\Psi_0\rangle &= \frac{1}{4} \sum_{i \neq j}^{N_{occ}} \sum_{a \neq b}^{N_{vir}} t_{ij}^{ab} \, |\Psi_{ij}^{ab}\rangle,
\end{aligned}
\tag{2.16}
$$

where $\hat{T}_j$ generates all *j* times excited Slater determinants. Following the common naming conventions, the expansion coefficients *t* are called amplitudes. N_{occ} and N_{vir} are the numbers of internal and virtual orbitals, respectively. In eq. (2.16), the action of the operators $\hat{T}_1$ and $\hat{T}_2$ on the Hartree-Fock ground state Slater determinant is given exemplary. The factor $\frac{1}{4}$ for the double excitation operator accounts for double counting. Analogous prefactors arise for higher-order excitation operators. The ansatz for the Coupled-Cluster wave function $|\Psi_{CC}\rangle$ is given in eq. (2.17).

$$
|\Psi_{CC}\rangle = e^{\hat{T}}|\Psi_0\rangle. \tag{2.17}
$$

Before moving on to strategies for solving the Coupled-Cluster equations, the Taylor expansion of $e^{\hat{T}}$ (eq. (2.18)) will be discussed.

$$
e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \frac{1}{6}\hat{T}^3 + \sum_{k=4}^{\infty} \frac{1}{k!}\hat{T}^k. \tag{2.18}
$$

By inserting eq. (2.16) into eq. (2.18) and sorting the terms by the total numbers of excitations, the expression given in eq. (2.19) is obtained.

$$
e^{\hat{T}} = 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) + \left(\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right) + \cdots. \tag{2.19}
$$

Interestingly, the chosen exponential cluster operator yields different types of, say, double excitations. First, there are excitations that originate from $\hat{T}_2$ with the corresponding amplitudes t_{ij}^{ab}. These types of excitations are called connected as they describe a simultaneous, or coupled, excitation of two electrons. Second, there are disconnected double excitations originating from $\frac{1}{2}\hat{T}_1^2$ with the amplitude product $t_i^a t_j^b$. For the triple excitations, there are connected ($\hat{T}_3$) and disconnected ($\hat{T}_2\hat{T}_1$ and $\frac{1}{6}\hat{T}_1^3$) excitations, and so on. The connected excitations stem from the linear term of the Taylor expansion in eq. (2.18), while disconnected excitations stem from higher order terms. The missing higher order excitations were responsible for the truncated CI methods not to be size consistent. By including these higher order excitations, even truncated CC approaches become size consistent.

Obtaining the CC energy usually is not done variationally in most quantum chemistry codes because, due to the multiplicative nature of the cluster expansion (see eqs. (2.18) and (2.19)), the amplitudes equations become much more complicated than in CI (which is a linear variational problem) [38]. Variational CC [39] is computationally very demanding as it scales factorially [40] with the system size and can only be applied to the smallest systems. However, Robinson et al. have proposed [41] and applied [42] a quasi-variational CC approach to small model systems, which shows the same formal scaling as non-variational CCSD approaches. The Schrödinger equation using the Coupled-Cluster approach is given in eq. (2.20).

$$\hat{H}e^{\hat{T}}|\Psi_0\rangle \;=\; E_{CC}e^{\hat{T}}|\Psi_0\rangle. \tag{2.20}$$

The Coupled-Cluster energy E_{CC} can be obtained by applying a similarity transformation on the Hamiltonian and projection on the reference determinant [43].

$$\begin{aligned} e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle \;&=\; E_{CC}e^{-\hat{T}}e^{\hat{T}}|\Psi_0\rangle \\ \langle\Psi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle \;&=\; \langle\Psi_0|\hat{H}\left(1 \;+\; \hat{T}_2 \;+\; \frac{1}{2}\hat{T}_1^2\right)|\Psi_0\rangle = \; E_{CC}. \end{aligned} \tag{2.21}$$

Note that only doubly excited determinants can interact directly with the Hartree-Fock ground state wave function, which is why only the connected and disconnected double excitation operators occur in eq. (2.21). The amplitudes can be obtained by projecting the Schrödinger equation on excited Slater determinants (eq. (2.22)).

$$\begin{aligned} \langle\Psi_S|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle \;&=\; 0 \\ \langle\Psi_D|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle \;&=\; 0. \end{aligned} \tag{2.22}$$

$\langle\Psi_S|$ and $\langle\Psi_D|$ are the singly and doubly excited determinants, respectively. Higher order amplitudes are obtained from analogous sets of equations. The last missing bit is the evaluation of the similarity transformed Hamiltonian $e^{-\hat{T}}\hat{H}e^{\hat{T}}$. For this, the Baker-Campbell-Hausdorff (BCH) expansion [44] is invoked, which naturally stops after five terms, since the Hamiltonian contains only one- and two-electron interactions [15,45].

$$\begin{aligned} e^{-\hat{T}}\hat{H}e^{\hat{T}} \;=\; \hat{H} \;+\; \left[\hat{H},\hat{T}\right] \;+\; \frac{1}{2}\left[\left[\hat{H},\hat{T}\right],\hat{T}\right] \;+\; \frac{1}{6}\left[\left[\left[\hat{H},\hat{T}\right],\hat{T}\right],\hat{T}\right] \\ +\; \frac{1}{24}\left[\left[\left[\left[\hat{H},\hat{T}\right],\hat{T}\right],\hat{T}\right],\hat{T}\right], \end{aligned} \tag{2.23}$$

with $\left[\hat{H},\hat{T}\right] \;=\; \hat{H}\hat{T} \;-\; \hat{T}\hat{H}$ being the commutator of $\hat{H}$ and $\hat{T}$. Although using the similarity transformed Hamiltonian together with the Baker-Campbell-Hausdorff expansion is the

standard approach in quantum chemistry codes, in principle projection of the non-transformed Hamiltonian can be carried out, leading to the set of equations given in eq. (2.24).

$$\begin{aligned} \langle \Psi_0 | \hat{H} e^{\hat{T}} | \Psi_0 \rangle &= E_{CC} \\ \langle \Psi_i^a | \hat{H} e^{\hat{T}} | \Psi_0 \rangle &= E_{CC} \langle \Psi_i^a | e^{\hat{T}} | \Psi_0 \rangle \\ \langle \Psi_{ij}^{ab} | \hat{H} e^{\hat{T}} | \Psi_0 \rangle &= E_{CC} \langle \Psi_{ij}^{ab} | e^{\hat{T}} | \Psi_0 \rangle, \end{aligned} \tag{2.24}$$

with analogous equations for higher order amplitudes. It is noted here that for linear variational problems like CI, the projection approach and the eigenvalue approach (eq. (2.15)) are equivalent, which is not true for nonlinear problems like Coupled-Cluster. As can be seen from comparing eq. (2.22) and eq. (2.24), the formalism using the BCH expansion has decoupled the amplitudes equation from the total energy. In practice, truncated cluster operators are used, leading to methods like CCSD ($\hat{T} = \hat{T}_1 + \hat{T}_2$) or CCSDT ($\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$). If the triples amplitudes are computed in a perturbative scheme, the method is labeled CCSD(T) [46], which has been considered gold standard in computational chemistry.

2.3.3 Local correlation methods

The Coupled-Cluster methods like CCSD, CCSD(T) or CCSDT introduced in section 2.3.2 as well as (to a certain extent) truncated CI approaches like CISD provide accurate results for systems with little static correlation. However, they scale with $O(N^6)$ (CCSD), $O(N^7)$ (CCSD(T)) or $O(N^8)$ (CCSDT) with N being the number of basis functions and therefore are applicable only to small systems. Local correlation methods date back to the early 1980s and were proposed by Pulay and Sæbø [47-51] and aim to reduce the overall scaling of correlated electronic structure methods while still recovering a large fraction of the correlation energy. The basic idea behind local correlations methods is to reduce the number of important configurations by localizing the internal and virtual space. The internal orbitals usually can be localized very well. The large scaling of correlation methods, however, is because the delocalized virtual space prevents the identification of dominant pair contributions to the correlation energy. Without loss of generality, the correlation energy can be written as sum of pair correlation energies [22] (eq. (2.25)).

$$E_{corr} = \sum_{i<j} \sum_{a<b} t_{ij}^{ab} \langle \Psi_0 | \hat{H}_e | \Psi_{ij}^{ab} \rangle = \sum_{i<j} \varepsilon_{ij} \tag{2.25}$$

with ε_{ij} being the pair correlation energy between the *i*-th and *j*-th electron. If the internal orbitals are localized, the pair correlation energy falls off quickly with the distance between the orbitals [49]. By choosing a suitable cutoff for the pair correlation energy, many electron pairs

can be discarded or treated at a lower level of theory (so called weak pairs), reducing the scaling of electron pairs (strong and weak) to be taken into account from quadratic to linear [52], while it is the number of discarded pairs that scales quadratically. However, if the virtual space is delocalized, still very unfavorable scaling occurs, so proper localization of the virtual space (which is not possible with simple localization schemes) is crucial. Pulay and Sæbø used the localized internal orbitals but discarded the Hartree-Fock virtual orbitals in their approaches. Instead, they constructed the virtual space from a set of atomic basis functions $\{|\mu\rangle\}$ and projected out the occupied space. These basis functions are called projected atomic orbitals (PAOs) $\{|\tilde{\mu}\rangle\}$ [52].

$$|\tilde{\mu}\rangle \;=\; \left(1-\sum_i |\psi_i\rangle\,\langle\psi_i|\right)|\mu\rangle. \tag{2.26}$$

The PAOs are localized and located in the same region in space as the atomic orbitals they are constructed from while being orthogonal to the occupied orbital space. This setup has been used by Pulay in their local Møller-Plesset perturbation theory [53] (MP2) approach [47,49] as well as in the local MP2 [52] and local CCSD [54] approach of Schütz and Werner. Schütz and Werner also provided a local version of the perturbative computation of triple excitations (T) [55].

A further ingredient for an efficient linear-scaling Coupled-Cluster approach are pair-natural orbitals (PNOs). In 1956, Löwdin [56] introduced the term natural spin orbitals, which are the eigenfunctions of a given density matrix. The corresponding eigenvalues are called natural occupation numbers. The importance of natural orbitals (natural orbitals, opposed to natural spin orbitals, are eigenfunctions of a spin-free density matrix [57]) is that they provide faster convergence than canonical Hartree-Fock orbitals in CI calculations, as was illustrated by Shavitt et al. [58]. Pair natural orbitals were introduced in the early 1970s by Meyer, [59,60] but they have been used early on in CI and coupled electron pair (CEPA) approaches by Ahlrichs und Kutzelnigg [61-64]. They are the eigenfunctions of a given pair density matrix $\boldsymbol{D}^{ij}$ for the electron pair of the *i*-th and *j*-th electron, the eigenvalues are again the natural occupation numbers. PNOs of a given electron pair are orthogonal to each other but not orthogonal to PNOs of a different electron pair and are located in the same region in space as the electron pair *(ij)* from whose pair density matrix they were constructed.

In recent years, Neese and coworkers combined the illustrated concepts to present a set of related approaches labeled local pair natural orbital (LPNO) or domain-based local pair natural orbital (DLPNO) approaches for local CEPA [65-69], LPNO-CCSD [66,70], DLPNO-MP2 [71] and DLPNO-CCSD [72], perturbative triples corrections [73,74] and Hartree-Fock + London Dispersion

(HFLD) [75,76]. All these approaches show linear or (the earlier implementations) nearly linear scaling with the system size for the correlation part and have been implemented in the ORCA program package [77,78].

Since the DLPNO-CCSD(T) and HFLD methods have been used extensively in this thesis, the general workflow is discussed. For technical details, see refs. 72-74. In the first step, the internal orbitals are localized, typically using the Forster-Boys [79] scheme, yielding a set of localized molecular orbitals (LMOs). As discussed before, the pair correlation energy ε_{ij} falls of quickly with increasing distance between the orbitals. Electron pairs with a pair correlation energy smaller than a certain cutoff (*TCutPairs* in ORCA notion) can be excluded from further procedure. In addition, for each localized orbital a domain is defined. A domain describes a set of PAOs that is located in the same region in space as the LMO. For this, the differential overlap DOI_{ik} between the LMO $\varphi_i(\vec{r})$ and the PAOs of a given atom is used.

$$DOI_{ik} = \sqrt{\int \varphi_i^2(\vec{r})\tilde{\mu}_k^2(\vec{r})d\vec{r}}. \tag{2.27}$$

The size of a domain for a given LMO is determined by a cutoff parameter for DOI_{ik} (*TCutDO* in ORCA notion). The domain of an electron pair *(ij)* is the union of the domains of *i* and *j*. A first estimate of the pair correlation energy is obtained from some computationally cheap multipole evaluation. For the surviving pairs, a more sophisticated so-called "semi-local" MP2 approach is used, in which the virtual space for a given electron pair *(ij)* is expanded in the PAOs belonging to the domain associated with *(ij)* [80].

In the second step of the procedure, the virtual space for each surviving electron pair is constructed. For this, the unrelaxed MP2 pair density matrix $\widetilde{\boldsymbol{D}}^{\boldsymbol{ij}}$ is diagonalized, giving a set of PNOs. As discussed before, the PNOs are located in the same region in space as *(ij)*. The dimension of the pair density matrix equals the number of virtual canonical orbitals. However, the virtual space can be limited to only the most important PNOs. The selection criterium is the occupation number of that PNO. This introduces another cutoff parameter (*TCutPNO* in ORCA notion). The number of surviving PNOs is much smaller than the number of virtual canonical orbitals and correlates with the interaction strength of the electron pair. This means that for weakly interacting electron pairs, even fewer PNOs must be included. PNOs therefore provide a much more compact representation of the virtual space than canonical orbitals do.

The last step is to expand the set of PNOs $\{|\tilde{\alpha}^{ij}\rangle\}$ into the PAOs $\{|\tilde{\mu}\rangle\}$ (eq. (2.28)).

$$|\tilde{\alpha}^{ij}\rangle = a^{ij}_{\tilde{\mu}\tilde{\alpha}} \sum_{\tilde{\mu}\in\{i,j\}} |\tilde{\mu}\rangle. \tag{2.28}$$

For the expansion of the PNOs, only those PAOs belonging to the combined domain $\{i,j\}$ are included. Whether the PAOs from the domain $\{i,j\}$ belonging to a certain atom q are actually included in the expansion (eq. (2.28)) depends on the Mulliken population of the PNO for this atom. This introduces a fourth cutoff parameter (*TCutMKN* in ORCA notion). Various benchmark studies [81-84] showed great overall performance of DLPNO-CCSD(T) methods achieving chemical accuracy. Liakos et al. found great agreement between DLPNO-CCSD(T) and canonical CCSD(T) for the GMTKN55 benchmark set [81,85] while Minenkov et al. [84] compared DLPNO-CCSD(T) to experimental dissociation enthalpies of noncovalently bound iron-ligand complexes and found good overall agreement.

From the four introduced cutoff parameters, the three most important ones controlling the behavior of the DLPNO-CC calculation are *TCutPairs, TCutPNO* and *TCutMKN*. ORCA provides pre-defined keywords for controlling these thresholds (*LoosePNO, NormalPNO, TightPNO*). Tightening the thresholds increases the computational effort but also the accuracy. A real-life example of an ORCA input file for a DLPNO-CCSD(T) calculation is given in **Figure 2.1**.

```
!DLPNO-CCSD(T) def2-TZVP def2-TZVP/C def2/J RIJCOSX VeryTightSCF NormalPNO

%mdci
    TCutPairs 1e-5
end

*xyzfile 0 1 example.xyz
```

Figure 2.1 Example input file for a DLPNO-CCSD(T) calculation with ORCA. The *NormalPNO* settings were used together with a tightened *TCutPairs* value (marked in red).

Besides *TCutPairs*, the *TCutPNO* parameter is of special importance for the accuracy of DLPNO methods. Recently, Altun et al. [86] presented a two-point extrapolation scheme (Complete PNO Space limit, CPS) for the correlation energy from two different *TCutPNO* values following eq. (2.29).

$$E_{CPS} = E_{X=6} + 1.5(E_{X=7} - E_{X=6}). \tag{2.29}$$

In eq. (2.29), E_{CPS} is the correlation energy extrapolated to the CPS limit, $E_{X=6}$ is the correlation energy for $TCutPairs = 10^{-6}$, $E_{X=7}$ is the correlation energy for $TCutPairs = 10^{-7}$. This extrapolation scheme is labeled *CPS(6/7)* extrapolation. *CPS(5/6)* extrapolation is

available as well and might be applied to computationally more demanding systems. The CPS extrapolation might be done manually, however, since ORCA 5 the compound feature, which adds another layer of abstraction and logics on top of the current ORCA infrastructure and therefore allows for implementation of automatized workflows, is available that features a script for automatic CPS extrapolation. This and many additional compound scripts are deployed with the ORCA program package. Note that for DLPNO calculations, using an auxiliary basis set (def2-TZVP/C in **Figure 2.1**) is mandatory. This is because the DLPNO machinery is closely linked internally to the resolution of identity [9].

2.3.4 Local energy decomposition

From a computational chemist's point of view, accurate energies are not sufficient for understanding chemical problems, but chemical understanding should be derivable from those numbers, meaning that the numbers a quantum chemistry code produces somehow must be mapped to well-known chemical concepts [87]. In recent years, especially the importance of noncovalent interactions like London dispersion (see section 3.4) for different chemical phenomena like stereoselectivity of organocatalysts has been unveiled. The DLNPNO-CC methods discussed in section 2.3.3 provide accurate energies. For analyzing noncovalent interactions, the local energy decomposition [88,89] (LED) method is implemented in ORCA. LED allows the decomposition of total and interaction energies into chemically meaningful contributions like dispersive or non-dispersive interaction, geometric preparation or electrostatic interaction [90].

For further discussion, it is assumed that the system *(AB)* is composed of two fragments *A* and *B*. Since the internal and virtual orbitals are localized, they might be assigned to either fragment *A* or fragment *B*. The double excitations building up the correlation energy (see eq. (2.25)) therefore might be categorized based on whether the occupied and virtual orbitals are localized on the same fragment or not [88]. **Figure 2.2** shows the different classes of double excitations.

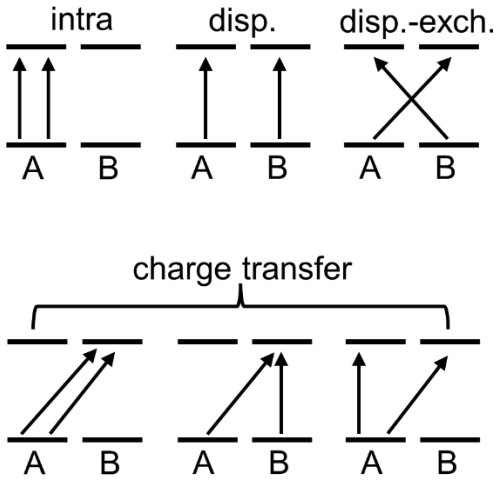

Figure 2.2 Different classes of double excitations. Disp. = dispersion, disp.-exch. = exchange-dispersion. The lower three excitations constitute charge transfer excitations.

The reference energy E_{ref} (that is, the energy of the Hartree-Fock reference wave function) of a system *(AB)* can be written as a sum of different contributions (eq. (2.30)) [90].

$$E_{ref} = E_{ref}^{(A)} + E_{ref}^{(B)} + E_{ref}^{elstat} + E_{ref}^{exch}. \tag{2.30}$$

In eq. (2.30), $E_{ref}^{(A)}$ and $E_{ref}^{(B)}$ is the respective reference energy of monomers *A* and *B* within the dimer. The last two terms E_{ref}^{elstat} and E_{ref}^{exch} describe the electrostatic and exchange interaction between the fragments *A* and *B* (based on the corresponding integrals from Hartree-Fock theory, see eq. (2.10)). An analogous decomposition can be carried out for the correlation energy and is given in eq. (2.31) [90].

$$E_C = E_C^{(A)} + E_C^{(B)} + E_{CCSD-disp}^{int} + E_{CCSD-CT}^{int} + E_{(T)}^{int}. \tag{2.31}$$

$E_C^{(A)}$ and $E_C^{(B)}$ represent the contribution from *intra*-fragment excitations to the correlation energy, $E_{CCSD-disp}^{int}$ is the dispersive interaction between fragments. $E_{CCSD-CT}^{int}$ describes the charge-transfer interaction between fragments. Note that $E_{CCSD-disp}^{int}$ and $E_{CCSD-CT}^{int}$ do not include contributions from triple excitations, this is why a triples correction $E_{(T)}^{int}$ is added to the correlation energy.

LED allows the decomposition of the binding energy ΔE_{bind} as well as the interaction energy $\Delta E_{int}(AB)$ between fragments *A* and *B*. Using a supermolecular approach, the binding energy describes the energy difference between the energy of the dimer and the sum of energies of the non-interacting monomers in their equilibrium geometries [91].

$$\Delta E_{bind} = E_{dimer} - \sum_i E_{monomer,i}. \tag{2.32}$$

The binding energy ΔE_{bind} might be decomposed using eq. (2.33) [90].

$$\begin{aligned} \Delta E_{bind} &= \Delta E_{geo-prep} + \Delta E_{int} \\ &= \Delta E_{geo-prep} + \Delta E_{el-prep} + \Delta E_{ref}^{int} + E_{disp}^{int} + \Delta E_{non-disp}^{int}. \end{aligned} \tag{2.33}$$

$\Delta E_{geo-prep}$ is called the geometric preparation. It describes the energy needed in order to distort the fragment's geometries from their equilibrium geometries to their dimer geometries. $\Delta E_{el-prep}$ describes the electronic preparation. For a given monomer, it is the energy difference between the monomer's energy within the dimer and the isolated monomer's energy in its dimer geometry. The electronic preparation contains contributions from the reference as well as the correlation wave function. For the correlation wave function, only the *intra*-fragment excitations contribute to the electronic preparation by definition. The next three terms in eq. (2.33) describe the sums of pair-wise interactions between two fragments, ΔE_{ref}^{int} is the interaction of the reference wave functions and includes electrostatic and exchange interactions, E_{disp}^{int} describes the dispersive interaction between two fragments and depends only on the correlation wave function. $\Delta E_{non-disp}^{int}$ is the non-dispersive interaction of the correlation wave functions and includes all terms but the dispersive interactions [90]. It is important to note that the decomposition scheme in eq. (2.33) is not unique – it is possible to slightly change the definition of some terms. For example, the electronic preparation $\Delta E_{el-prep}$ might be implicitly captured in ΔE_{ref}^{int} and $\Delta E_{non-disp}^{int}$ or might be explicitly split up into contributions from the reference and correlation wave function [90,92]. Also, in eq. (2.33), the interaction of triples is absorbed in the terms E_{disp}^{int} and $\Delta E_{non-disp}^{int}$, but it is possible to introduce a separate term for triples interactions, as it is was done in eq. (2.31). In eq. (2.33), the dispersive (E_{disp}^{int}) and non-dispersive ($\Delta E_{non-disp}^{int}$) interactions include interactions from the triple excitations, which are in fact computed from the *inter*-fragment triples contributions $E_{(T)}^{int}$ [93].

$$\begin{aligned} E_{disp,(T)}^{int} &= \gamma E_{(T)}^{int} \\ E_{non-disp,(T)}^{int} &= (1-\gamma) E_{(T)}^{int} \\ \gamma &= \frac{E_{disp,SP}^{int}}{E_{SP}^{int}}, \gamma \in [0,1]. \end{aligned} \tag{2.34}$$

The scaling factor γ is defined as the ratio of the dispersive energy contributed by strong pairs $E_{disp,SP}^{int}$ and the total interaction contributed by strong pairs E_{SP}^{int} between fragments *A* and *B*.

If $\gamma = 0$, this indicates that the pair of fragments does not have any dispersive interactions. Consequently, no dispersive contributions from the inter-triple excitations are assumed as well. Likewise, if $\gamma = 1$, the fragments interact entirely through dispersive interactions, which is why, in this case, the triples contribution to dispersion equals the total triples interaction.

Besides LED, which is based on a supermolecular approach, perturbative approaches such as Symmetry-Adapted Perturbation Theory (SAPT) have been proposed [94,95]. These approaches do not require calculations on the individual monomers, as is the case with the supermolecular approach. In addition, SAPT also provides a decomposition of the interaction energy into chemical meaningful terms like dispersion or electrostatic interactions.

2.4 Density functional theory

2.4.1 Foundations

A vastly different approach for finding approximate solutions to the electronic Schrödinger equation is density functional theory (DFT), whose foundation was laid by the Hohenberg-Kohn theorems [96,97]. Central observable of DFT is the electron density $\rho(\vec{r})$, which is defined *via* eq. (2.35).

$$\begin{aligned} \rho(\vec{r}) &= N \int \cdots \int \Psi^*(\vec{x}_1, \ldots, \vec{x}_N)\Psi(\vec{x}_1, \ldots, \vec{x}_N) d\sigma_1 d\vec{x}_2 \cdots \vec{x}_N \\ \int \rho(\vec{r}) d\vec{r} &= N. \end{aligned} \tag{2.35}$$

Note that in eq. (2.35), integration over the spin coordinate of the first electron σ_1 was carried out, hence making $\rho(\vec{r})$ independent of the electron spin. The electron density describes the probability of finding an electron (it does not matter which electron) at position $\vec{r}$. The prefactor N is the number of electrons, which is needed in order to ensure that the electron density does not describe the probability distribution of a particular electron, but finding *any* electron at position $\vec{r}$ [98]. Notice that the electron density is a function of three variables *(x,y,z),* independent of the total number of electrons. The first Hohenberg-Kohn theorem [96] states that the electron density $\rho(\vec{r})$ of an interacting many-particle system with an external potential $\nu(\vec{r})$ determines the external potential unambiguously up to an additive constant. With the electronic Hamiltonian from eq. (2.5), the external potential describes the electrostatic interaction of the electrons with the nuclei. From eq. (2.35), it can be seen that the electron density also determines the total number of electrons. Since $\hat{H}_e$ only varies in the number of electrons and

the external potential $\nu(\vec{r})$, $\rho(\vec{r})$ also determines the ground state energy of the system. The ground state energy E_0 of the system is given in eq. (2.36) [96].

$$E_0 \;=\; \langle \Psi_0 | \hat{T} + \hat{U}_{ee} | \Psi_0 \rangle \;+\; \langle \Psi_0 | \nu | \Psi_0 \rangle = T + \; U_{ee} + \; V_{ext}. \tag{2.36}$$

T is the kinetic energy of the electrons, U_{ee} is the exact electron-electron interaction energy, V_{ext} is the energy associated with the interaction of the electrons with the external potential $\nu(\vec{r})$ and $|\Psi_0\rangle$ is the ground state wave function. From U_{ee}, the classical Coulomb interaction of two charge densities can be separated out (eq. (2.37)).

$$U_{ee} \;=\; U_H \;+\; U_R \;=\; \frac{1}{2} \iint \frac{e^2 \rho(\vec{r}) \rho(\vec{r}')}{4\pi\varepsilon_0 |\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \; U_R. \tag{2.37}$$

U_H is called the Hartree energy and describes the classical Coulomb interaction energy of two charge distributions, while U_R describes all other effects like exchange and electron correlation [99]. Note the separation in eq. (2.36) between $\hat{T} + \hat{U}_{ee}$ on one side and $\nu(\vec{r})$ on the other. $T \;+\; U_{ee}$ is the system-independent part of the energy. The corresponding functional $F[\rho(\vec{r})] \;=\; \underset{\Psi \to \rho(\vec{r})}{min} \langle \Psi | \hat{T} + \hat{U}_{ee} | \Psi \rangle$ is called the universal functional as it is valid for any external potential. The notation given in ref. 99 is followed here. The minimization runs over all trial wave functions that yield the density $\rho(\vec{r})$. The total energy as functional of the electron density can be written as [96]

$$\begin{aligned} E[\rho(\vec{r})] \;&=\; F[\rho(\vec{r})] \;+\; V_{ext}[\rho(\vec{r})] \\ &=\; F[\rho(\vec{r})] \;+\; \int \rho(\vec{r}) \nu(\vec{r}) d\vec{r}. \end{aligned} \tag{2.38}$$

The second Hohenberg-Kohn theorem [96,99] states that the ground state energy E_0 of the system can be found variationally by minimizing $E[\rho(\vec{r})]$ with respect to $\rho(\vec{r})$.

$$E_0 \;=\; \underset{\rho(\vec{r})}{min}\, E[\rho(\vec{r})] \;=\; \underset{\rho(\vec{r})}{min}\, F[\rho(\vec{r})] \;+\; \int \rho(\vec{r}) \nu(\vec{r}) d\vec{r}. \tag{2.39}$$

If $\rho(\vec{r}) \;=\; \rho_0(\vec{r})$, it follows that $E_0 \;= E[\rho_0(\vec{r})]$.

2.4.2 Kohn-Sham formalism

For applicability of DFT to chemical systems, the so-called Kohn-Sham (KS) formalism is introduced [97], which projects the system of interacting electrons onto a system of non-interacting electrons with density $\tilde{\rho}(\vec{r})$ that equals the density $\rho(\vec{r})$ of the interacting system. The density functional $\tilde{E}[\tilde{\rho}(\vec{r})]$ of the non-interacting system is given by eq. (2.40).

$$\tilde{E}[\tilde{\rho}(\vec{r})] \;=\; T_s[\tilde{\rho}(\vec{r})] \;+\; \int \tilde{\rho}(\vec{r})\nu(\vec{r})d\vec{r}, \tag{2.40}$$

with $T_s[\tilde{\rho}(\vec{r})]$ being the kinetic energy of the non-interacting electrons, which is different from the kinetic energy $T[\rho(\vec{r})]$ of interacting electrons [100]. The universal functional and the total energy functional for the system of interacting electrons using the kinetic energy of non-interacting electrons are given in eq. (2.41) [100].

$$\begin{aligned} F[\tilde{\rho}(\vec{r})] \;&=\; T_s[\tilde{\rho}(\vec{r})] \;+\; \frac{1}{2}\iint \frac{e^2\tilde{\rho}(\vec{r})\tilde{\rho}(\vec{r}')}{4\pi\varepsilon_0|\vec{r}-\vec{r}'|}d\vec{r}d\vec{r}' +\; E_{xc}[\tilde{\rho}(\vec{r})] \\ E[\tilde{\rho}(\vec{r})] \;&=\; F[\tilde{\rho}(\vec{r})\,] \;+\; V_{ext}[\tilde{\rho}(\vec{r})] \\ &=\; T_s[\tilde{\rho}(\vec{r})] \;+\; \frac{1}{2}\iint \frac{e^2\tilde{\rho}(\vec{r})\tilde{\rho}(\vec{r}')}{4\pi\varepsilon_0|\vec{r}-\vec{r}'|}d\vec{r}d\vec{r}' +\; E_{xc}[\tilde{\rho}(\vec{r})] \\ &\quad +\; \int \tilde{\rho}(\vec{r})\nu(\vec{r})d\vec{r}. \end{aligned} \tag{2.41}$$

$E_{xc}[\tilde{\rho}(\vec{r})]$ is the so-called exchange correlation energy that combines the non-classical contributions from the electron-electron interaction (U_R from eq. (2.37)) and the difference in kinetic energy between interacting and non-interacting electrons and collects therefore all terms for which no closed expression can be given [100]. For a system of non-interacting electrons, the many-particle wave function $|\Psi\rangle$ is written as (antisymmetrized) product of orthonormal orbitals $|\psi_j\rangle$ (*i.e.*, a Slater determinant). The electron density $\tilde{\rho}(\vec{r})$ is then given by eq. (2.42).

$$\tilde{\rho}(\vec{r}) \;=\; \sum_{i=1}^{n_{occ}} |\psi_i(\vec{r})|^2. \tag{2.42}$$

The sum in eq. (2.42) runs over all n_{occ} occupied orbitals. Minimization of the energy functional in eq. (2.41) with respect to the orbitals while keeping the orbitals orthonormal yields the Kohn-Sham equations (eq. (2.43)) that determine the optimal set of orbitals [100].

$$\begin{aligned} \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}_j^2 \;+\; \nu_{eff}(\vec{r}) - \varepsilon_j\right)\psi_j(\vec{r}) \;&=\; 0 \\ \nu_{eff}(\vec{r}) \;&=\; \nu(\vec{r}) \;+\; \int \frac{e^2\tilde{\rho}(\vec{r}')}{4\pi\varepsilon_0|\vec{r}-\vec{r}'|}d\vec{r}' \;+\; \nu_{xc}(\vec{r}) \\ \nu_{xc}(\vec{r}) \;&=\; \frac{\delta E_{xc}[\tilde{\rho}(\vec{r})]}{\delta\tilde{\rho}(\vec{r})}. \end{aligned} \tag{2.43}$$

$\nu_{eff}(\vec{r})$ is an effective one-particle potential that includes the exchange correlation functional $\nu_{xc}(\vec{r})$, which is the variation of the exchange correlation energy with respect to the density. ε_j are again the diagonal elements of the Lagrange multiplier matrix. There is no closed expression

for the exchange correlation functional $\nu_{xc}(\vec{r})$ and it is not known exactly for general systems and therefore must be approximated [100]. Just like the Hartree-Fock equations, the Kohn-Sham equations are nonlinear – they must be solved iteratively by using a set of starting orbitals.

2.4.3 Jacob's ladder of functionals

As discussed in section 2.4.2, Kohn-Sham DFT is only approximate due to approximate exchange correlation functionals (xc functionals) $\nu_{xc}(\vec{r})$. There is no straightforward strategy for improving the quality of a given exchange correlation functional, but over years different classes of xc functionals were designed. In 2001, Perdew et al. [101] introduced the term Jacob's ladder that ranks the different classes of xc functionals by increasing quality. The lowest level would be the neglection of $\nu_{xc}(\vec{r})$ and therefore total neglection of exchange interaction and electron correlation, which is not suitable for chemical systems of any kind. The lowest rung on the ladder is the local spin density (LSD) approximation [96,102,103].

$$E_{xc}^{LSD}[\tilde{\rho}(\vec{r})] = \int \tilde{\rho}(\vec{r})\varepsilon_{xc}^{unif}[\tilde{\rho}^{\alpha}(\vec{r}), \tilde{\rho}^{\beta}(\vec{r})]d\vec{r}. \tag{2.44}$$

$\varepsilon_{xc}^{unif}[\tilde{\rho}^{\alpha}(\vec{r}), \tilde{\rho}^{b}(\vec{r})]$ in eq. (2.44) is the exchange correlation energy per particle of the uniform electron gas and it depends on the spin densities $\tilde{\rho}^{\alpha}(\vec{r}), \tilde{\rho}^{\beta}(\vec{r})$ (α means spin up, β means spin down). As discussed in detail by Jones and Gunnarsson [104], LSD gives much better well depths for first-row dimers than Hartree-Fock, especially for cases where correlation becomes important (like Be_2) while still overestimating binding energies.

On the next rung of Jacob's ladder [101] are generalized gradient approximated functionals (GGA functionals), for which deviation from the uniform electron gas is achieved by making the exchange correlation energy per particle dependent on the density gradient.

$$E_{xc}^{GGA}[\tilde{\rho}(\vec{r})] = \int \tilde{\rho}(\vec{r})\varepsilon_{xc}^{GGA}[\tilde{\rho}^{\alpha}(\vec{r}), \tilde{\rho}^{\beta}(\vec{r}), \nabla\tilde{\rho}^{\alpha}(\vec{r}), \nabla\tilde{\rho}^{\beta}(\vec{r})]d\vec{r}. \tag{2.45}$$

In the design of density functionals, the exchange correlation energy is often assumed to be the sum of independent contributions from exchange and correlation, as illustrated in eq. (2.46).

$$\begin{aligned} E_{xc}[\tilde{\rho}(\vec{r})] &= E_x[\tilde{\rho}(\vec{r})] + E_c[\tilde{\rho}(\vec{r})] \\ &= \int \tilde{\rho}(\vec{r})\varepsilon_x[\tilde{\rho}(\vec{r})]d\vec{r} + \int \tilde{\rho}(\vec{r})\varepsilon_c[\tilde{\rho}(\vec{r})]d\vec{r}. \end{aligned} \tag{2.46}$$

The formulation in eq. (2.46) allows for individual design, optimization and subsequent combination of exchange or correlation parts. An early functional implementing the density

gradient is the B88 functional [105] (note that the original reference does not use the name B88, though). The exchange part of the functional is given by eq. (2.47).

$$\begin{aligned} E_x^{B88}[\tilde{\rho}(\vec{r})] &= E_x^{LDA}[\tilde{\rho}(\vec{r})] - \beta \sum_\sigma \int \tilde{\rho}^\sigma(\vec{r})^{4/3} \frac{x_\sigma^2}{1 + 6\beta x_\sigma \, sinh^{-1} x_\sigma} d\vec{r} \\ x_\sigma &= \frac{|\nabla \tilde{\rho}^\sigma(\vec{r})|}{\tilde{\rho}^\sigma(\vec{r})^{4/3}}. \end{aligned} \tag{2.47}$$

In eq. (2.47), $E_x^{LDA}[\tilde{\rho}(\vec{r})]$ is the exchange correlation energy from the local density approximation (LDA), β is a constant that will be fitted, $\tilde{\rho}^\sigma(\vec{r})$ is the spin density with the general spin variable σ, x_σ is the dimensionless gradient. Over the years, many different GGA functionals were designed, most well-known are the LYP [106,107], the PBE [108] and the PW91 [109] density functionals.

The next step in improving the exchange correlation functionals is to include higher order derivatives of the density, leading to the meta-GGA functionals. Typically, these functionals include the dependence of the functional on the kinetic energy density $\tau^\sigma(\vec{r})$ of the Kohn-Sham orbitals [15], leading to functionals of the form given in eq. (2.48).

$$\begin{aligned} E_{xc}^{mGGA}[\tilde{\rho}(\vec{r})] &= \int \tilde{\rho}(\vec{r}) \varepsilon_{xc}^{mGGA}[\tilde{\rho}^\alpha(\vec{r}), \tilde{\rho}^\beta(\vec{r}), \nabla\tilde{\rho}^\alpha(\vec{r}), \nabla\tilde{\rho}^\beta(\vec{r}), \tau^\alpha(\vec{r}), \tau^\beta(\vec{r})] d\vec{r} \\ \tau^\sigma(\vec{r}) &= \frac{1}{2} \sum_{i=1}^{n_{occ,\sigma}} |\nabla \psi_i(\vec{r})|^2. \end{aligned} \tag{2.48}$$

Well-known meta-GGA functionals are TPSS [110] or M06-L [111].

The next rung on the ladder represents hybrid functionals [101]. Hybrid functionals replace the exchange energy partly by the exact exchange (represented by exchange integrals from Hartree-Fock theory).

$$E_{xc}^{hybrid} = E_{xc}^{GGA} + a(E_x^{exact} - E_x^{GGA}). \tag{2.49}$$

In eq. (2.49), E_x^{exact} represents the exact exchange energy from Hartree-Fock theory, which is weighted by the scaling factor a. Note that the GGA exchange contribution and the GGA correlation contribution do not have to be derived from the same functional. Maybe the most famous hybrid functional at all, B3LYP [112,113], uses in its original formulation Hartree-Fock and B88 exchange and PW91 correlation:

$$E_{xc}^{B3LYP} = E_{xc}^{LSD} + a_0(E_x^{exact} - E_x^{LSD}) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{PW91}. \tag{2.50}$$

Eq. (2.50) contains three fitting parameters a_0, a_x and a_c whose optimum values have been determined by Becke [112]. Note that in eq. (2.50) only the gradient correction terms of the B88 exchange energy ΔE_x^{B88} and PW91 correlation energy ΔE_c^{PW91} were used. Another widely used hybrid functional is the parameter free PBE0 [114,115] functional.

The highest rung discussed here represents functionals that include the virtual orbital space in the total energy expression [101]. A prominent representative of this functional class are double hybrid functionals. Double hybrid functionals replace parts of the DFT correlation energy by some correlation energy obtained from MP2. Since double hybrid functionals have not been used beyond test calculations in this thesis, the discussion is omitted and the interested reader is referred to a review by Goerigk and Grimme [116].

Usually not a separate rung on Jacob's ladder, but nevertheless widely used, are range-separated (RS) hybrid functionals [117,118]. In the long-range region (*i.e.*, far away from the nucleus), the exchange-correlation potentials of common hybrid functionals show the wrong asymptotic behavior, leading to larger self-interaction errors of *e.g.*, loosely bound electrons [15]. RS hybrid functionals split the Coulomb interaction into a long-range and a short-range contribution [119].

$$\begin{aligned} \frac{1}{|\vec{r}_1 - \vec{r}_2|} &= \frac{1}{r_{12}} = \frac{1 - erf(\mu r_{12})}{r_{12}} + \frac{erf(\mu r_{12})}{r_{12}} \\ erf(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \end{aligned} \tag{2.51}$$

In eq. (2.51), $erf(x)$ is the error function, which is responsible for a smooth transition from the short-range into the long-range domain. The parameter μ controls how fast the transition from short-range to long-range happens and depends on the used exchange correlation functional. In practice, the short-range part of the exchange energy is treated with a local GGA or meta-GGA functional, while the long-range part is treated exactly [117,120]. Although Jacob's ladder yields in general a good indication of the quality of a density functional, it should be kept in mind that the quality of a calculation does not depend exclusively on the density functional, but on many additional parameters such as basis set, solvation scheme, treatment of noncovalent interactions, relativistic corrections and such [121]. Over recent years, Grimme and coworkers proposed a series of composite methods [122-125] (the 3c family) that are based on different well-known density functionals and Hartree-Fock. All these methods (HF-3c, B97-3c, PBEh-3c, r^{2}SCAN-3c) feature a relatively small basis set and include three additional correction terms: a semi-empirical correction for long-range electronic effects (see section 2.4.4), the geometric

counterpoise correction [126] for the basis set superposition error (BBSE) and a short-ranged correction for basis set shortcomings [122]. These functionals are considered to be fast and robust electronic structure methods while still being more accurate than the corresponding density functionals they inherited from [123].

2.4.4 Description of London dispersion

One general shortcoming of the approximate exchange correlation functionals described in section 2.4.3 is the insufficient description of electronic long-range interactions such as London dispersion (LD) [127,128]. London dispersion, named after German-American physicist Fritz London, describes the omnipresent and temperature-independent interaction of molecules (or different groups within the same molecule) based on induced dipole moments. London dispersion forces are always attractive and fall off with $1/r^6$, with r being the distance between the interacting atoms [129]. Traditionally, London dispersion has been used to explain the boiling points of alkanes. However, the importance of LD for intramolecular interactions was exemplified by Pitzer and Catalano, who computed the intramolecular noncovalent correlation energies for a series of hydrocarbons [130]. Over recent years, the importance of LD for chemical reactivity, stability of molecular balances, π-effects or spectroscopic properties has become evident. A review was provided by Wagner and Schreiner [131]. The influence of LD on stereoselectivities was studied by Gschwind and coworkers for the chiral phosphoric acid (CPA) catalyzed transfer hydrogenation of imines [132]. The general interest in LD leads to strong interest in the approximate description of LD within DFT and various approaches were suggested and optimized [133]. It is noted here that also other approaches for the computation of LD interactions outside of DFT were suggested (see section 2.3.4). Maybe the simplest approach is the DFT-D approach, in which a simple correction term E_{disp} is added to the DFT energy (eq. (2.52)).

$$E_{tot} = E_{DFT} + E_{disp}. \tag{2.52}$$

The correction term E_{disp} consists of a sum of pair-wise atomic contributions scaled by the distance r_{AB} between atoms A and B.

$$E_{disp} = -\sum_{A}\sum_{B>A}\frac{C_6^{AB}}{r_{AB}^6}. \tag{2.53}$$

The coefficients C_6^{AB} can be obtained in various ways. A simple approach is DFT-D2, going back to Grimme [134], in which the atomic coefficients are computed from atomic ionization potentials I_p^A and static dipole polarizabilities α^A.

$$\begin{aligned} C_6^{AB} &= \sqrt{C_6^A C_6^B} \\ C_6^A &= 0.05 \bar{N} I_p^A \alpha^A. \end{aligned} \tag{2.54}$$

In eq. (2.54), $\bar{N}$ is a natural number that depends on the element row of atom A. I_p^A and α^A might be taken from high-level quantum chemical calculations. The outlined scheme is very simple and computationally cheap. The downhill, however, is that the coefficient C_6^A is the same for all atoms of the same type, irrespective of the chemical environment. Another problem is related to the simple expression for E_{disp} given in eq. (2.53). For small distances r_{AB}, the dispersive interaction diverges and becomes huge. The definition of E_{disp} is therefore changed by including a damping function f_{damp} that damps the dispersion interaction for short distances and recovers the original dispersive interaction for long distances [133]. The new expression for E_{disp} is given in eq. (2.55). Note that an overall scaling factor s_6 was introduced.

$$E_{disp} = - s_6 f_{damp}(r_{AB}, A, B) \sum_A \sum_{B>A} \frac{C_6^{AB}}{r_{AB}^6}. \tag{2.55}$$

A major improvement over the DFT-D2 scheme is the DFT-D3 [135] scheme proposed by Grimme et al. The essence of DFT-D3 is to take environmental effects into account through coordination number (CN) dependent coefficients C_6^{AB}. The dispersive interactions of a chemically bound atom usually are smaller than those of a free atom. The reason is the formation of energetically low-lying and doubly occupied molecular orbitals in contrast to the often singly occupied orbitals of the free atom. Coordination dependent coefficients shall take account of these effects. The effect of oxidation and reduction is neglected, since inclusion of charge transfer processes would require some more detailed knowledge of the density. The basis for computing C_6^{AB} in the DFT-D3 approach is the Casimir-Polder formula [136].

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty \alpha^A(i\omega) \alpha^B(i\omega) d\omega, \tag{2.56}$$

with $\alpha^A(i\omega)$ being the averaged dipole polarizability at imaginary frequency ω. The polarizabilities might be obtained from time-dependent DFT (TD-DFT) calculations. Reference values for a pair of atom types are not solely obtained from the free atoms, but from one or multiple hydrides $A_m H_n$ and $B_k H_l$ and subsequent elimination of the hydrogen contribution.

$$C_{6,ref}^{AB} = \frac{3}{\pi}\int_0^\infty \frac{1}{m}\left(\alpha^{A_m H_n}(i\omega) - \frac{n}{2}\alpha^{H_2}(i\omega)\right)\cdot\frac{1}{k}\left(\alpha^{B_k H_l}(i\omega) - \frac{l}{2}\alpha^{H_2}(i\omega)\right)d\omega. \tag{2.57}$$

For implementation of the coordination number dependency, the coordination number CN^A of atom A is defined *via* eq. (2.58).

$$CN^A = \sum_{B\neq A}\frac{1}{1+e^{-k_1(k_2(R_{A,cov}+R_{B,cov})/r_{AB}-1)}}, \tag{2.58}$$

with $R_{A,cov}$ being the covalent radius of atom A, and k_1 and k_2 being scaling factors. The computation of the CN dependent coefficients $C_6^{AB}(CN^A, CN^B)$ for a pair of atom types (AB) goes as follows. For one or more reference systems, the reference coefficients $C_{6,ref}^{AB}(CN^A, CN^B)$ are calculated using eq. (2.57). At the same time, the corresponding coordination numbers are stored. This procedure has to be done only once. For the system under investigation, the coordination numbers are calculated using eq. (2.58) and the coefficients $C_6^{AB}(CN^A, CN^B)$ are extrapolated from the set of reference coefficients using eq. (2.59).

$$\begin{aligned} C_6^{AB}(CN^A, CN^B) &= \frac{Z}{W} \\ Z &= \sum_i^{N_i}\sum_j^{N_j} C_{6,ref}^{AB}(CN_i^A, CN_j^B)L_{ij} \\ W &= \sum_i^{N_i}\sum_j^{N_j} L_{ij} \\ L_{ij} &= e^{-k_3\left((CN^A-CN_i^A)^2+(CN^B-CN_j^B)^2\right)}. \end{aligned} \tag{2.59}$$

k_3 in eq. (2.59) is another scaling constant, N_i and N_j are the total numbers of reference systems, respectively. Up to this point, only the $\frac{1}{r^6}$ dependency was covered. However, in principle higher order terms like $\frac{C_8^{AB}}{r_{AB}^8}$ or even $\frac{C_{10}^{AB}}{r_{AB}^{10}}$ might be included. This does not influence how the individual coefficients C_6^{AB} are computed. The higher-order coefficients can be computed recursively from C_6^{AB}. The DFT-D3 scheme as discussed up to this point is very often used together with the Becke-Johnson (BJ) damping function [137,138]. The DFT-D3(BJ) dispersion energy is given by eq. (2.60).

$$
\begin{aligned}
E_{disp}^{D3(BJ)} &= -\sum_{A}\sum_{B>A} s_6 \frac{C_6^{AB}}{r_{AB}^6 + f_{damp}(r_{AB}^0)^6} \\
&\quad + s_8 \frac{C_8^{AB}}{r_{AB}^8 + f_{damp}(r_{AB}^0)^8} \qquad (2.60) \\
f_{damp}(r_{AB}^0) &= a_1 r_{AB}^0 + a_2 \\
r_{AB}^0 &= \frac{C_8^{AB}}{C_6^{AB}},
\end{aligned}
$$

with some constants a_1 and a_2. In practice, s_6 is set to unity, while s_8, a_1 and a_2 are fitted for each density functional. Through these parameters, the dispersion energy $E_{disp}^{D3(BJ)}$ depends on the density functional. In total, the dispersion energy of DFT-D3 methods depends on the chemical environment through the coordination number dependency of C_6^{AB} and on the density functional through the scaling factors s_8 and the constants within the damping function $f_{damp}(r_{AB}^0)$, making the D3 methods very flexible and widely applicable. The improvements of the description of long-range interactions by including the D3 correction has been reported in the original publication [135].

In 2017, the DFT-D4 method was introduced by Grimme and coworkers [139]. The basic idea behind DFT-D4 is to define an effective atomic charge for each atom, which is the sum of the nuclear charge and the Mulliken atomic charge that itself depends on the density matrix. The dispersion coefficients still are derived from the Casimir-Polder formula and also the dependency on the coordination number is kept as discussed in eqs. (2.58) and (2.59), but the effective atomic charge enters as scaling factor for the reference coefficients [140].

A vastly different approach from the discussed DFT-D methods is to make the dispersion energy a non-local functional of the electron density. The best-known approach is the non-local VV10 density functional [141]. The non-local correlation energy E_c^{nl} is written as

$$
E_c^{nl} = \frac{\hbar}{2} \iint \rho(\vec{r}) \Theta(\vec{r}, \vec{r}') \rho(\vec{r}') d\vec{r} d\vec{r}'. \qquad (2.61)
$$

In eq. (2.61), $\Theta(\vec{r}, \vec{r}')$ is called the VV10 kernel that depends on the density as well as the density gradient. The structure of the kernel is slightly complicated and is given in eq. (2.62) for completeness. More details can be found in the literature [141-144].

$$
\begin{aligned}
\Theta(\vec{r},\vec{r}') &= \frac{3e^4}{2m^2 g g'(g+g')} \\
g &= \omega_0(\vec{r})|\vec{r}-\vec{r}'|^2 + \kappa(\vec{r}); g' = \omega_0(\vec{r}')|\vec{r}-\vec{r}'|^2 + \kappa(\vec{r}') \\
\omega_0(\vec{r}) &= \sqrt{\omega_g^2(\vec{r}) + \frac{\omega_p^2(\vec{r})}{3}} \\
\omega_p^2(\vec{r}) &= \frac{4\pi\rho(\vec{r})e^2}{m_e}; \qquad \omega_g^2(\vec{r}) = C\frac{\hbar^2}{m_e^2}\left|\frac{\nabla\rho(\vec{r})}{\rho(\vec{r})}\right|^4 \\
\kappa(\vec{r}) &= b\frac{\nu_F^2(\vec{r})}{\omega_p(\vec{r})}; \qquad \nu_F(\vec{r}) = \frac{\hbar\sqrt[3]{3\pi^2\rho(\vec{r})}}{m_e}.
\end{aligned}
\tag{2.62}
$$

In eq. (2.62), C and b are constants that control the asymptotic behavior and short-range damping, respectively, and they might be fitted for different exchange correlation functionals. In the last step, the non-local energy contribution E_c^{VV10} from the VV10 kernel is given by eq. (2.63).

$$
\begin{aligned}
E_c^{VV10} &= \int \rho(\vec{r})\left[\beta + \frac{\hbar}{2}\int \rho(\vec{r}')\Theta(\vec{r},\vec{r}')d\vec{r}'\right]d\vec{r} \\
\beta &= \frac{me^4}{32\hbar^2}\left(\frac{3}{b^2}\right)^{3/4}.
\end{aligned}
\tag{2.63}
$$

In principle, the VV10 functional can be combined with any exchange correlation functional. However, there exists a number of predefined density functionals from the B97 family that include the VV10 correction as default, such as the meta-GGA B97M-V [145], the range-separated hybrid GGA ωB97X-V [146] or the range-separated hybrid meta-GGA ωB97M-V [147] functionals. Note that these functionals also exist with corresponding D3 corrections instead of VV10 [147]. The VV10 dispersion correction might be invoked in different fashions in ORCA. First, it can be used in a post-SCF fashion (*!NL* in ORCA notation): once the self-consistency cycle with the chosen exchange correlation functional has met convergence, the converged density is used for computing the energy contribution from VV10. Alternatively, the VV10 functional can be included in the self-consistency cycle (*!SCNL* in ORCA notation). Najibi et al. [148] showed, however, that both approaches yield very comparable results while the post-SCF approach is significantly less computationally demanding. As has been discussed by Hujo and Grimme [149], D3 and VV10 dispersion yield very similar and accurate equilibrium geometries for vastly different structures. However, the D3 correction provides a better accuracy/cost ratio because D3 is computationally very cheap, while the VV10 correction produces a significant computational overhead with respect to the SCF cycle. However, several benchmarks have

demonstrated great performance of the ωB97M-V [147,148,150] density functional for isomerization energies, barrier heights, noncovalent interactions and transition metal reaction energies.

2.5 Exploring the conformational space

Molecules with the same topology can have multiple conformers, that is, the potential energy surface has a minimum for different geometries of that molecule. Many properties, such as stereoselectivity of an organocatalyst or the nuclear magnetic resonance (NMR) parameters are structure-dependent. Since the energy barriers for transforming conformers into other conformers typically are small, multiple conformers exist in equilibrium. Of course, the population of a given conformer depends on its energy relative to the ground state and on the temperature. However, for comparing computed data (such as NMR parameters) with experiment, averaging over significantly populated conformers can become necessary. In this section, common strategies for generating the conformer ensemble are introduced. Since being used extensively in this thesis, most space will be dedicated to metadynamics based approaches, as implemented in the Conformer–Rotamer Ensemble Sampling Tool [151] (CREST) software developed by Grimme.

In CREST, exploration of the conformational space is carried out by means of root-mean-square deviation of atomic positions (RMSD) based metadynamics simulations [152]. The basic idea of metadynamics simulations (MTD) is to add a history-dependent bias potential to the total energy expression [153]. This bias potential E_{bias}^{RMSD} is specified by a set of so-called collective variables $\{\Delta\}$ (see eq. (2.64)) and drives the system over energy barriers in shorter simulation times than standard molecular dynamics (MD) simulations.

$$E_{bias}^{RMSD} = \sum_{i}^{N} k_i e^{-\alpha \Delta_i^2}. \tag{2.64}$$

In eq. (2.64), N is the number of references structures. The force constants k_i are always positive, thus the bias potential is repulsive. The collective variable Δ_i is the RMSD taken between the current minimum and the *i*-th previously explored minimum. Through the repulsive bias potential, the system is therefore hindered from falling back into already explored energy minima, resulting in effective exploration of the conformational space. The metadynamics simulations are carried for a certain simulation time. Afterwards, all found minima must be optimized without the bias potential. By repeating metadynamics simulations and subsequent unbiased geometry optimizations in an iterative fashion, the conformer ensemble is generated. In the last step of the procedure, all found local minimum structures are compared pairwise in

order to distinguish between conformers, rotamers and identical structures. If the energy difference ΔE between two structures is larger than the chosen cutoff $\mathrm{E_{th}}$, the two geometries are considered conformers. If $\Delta E < E_{th}$, but the RMSD is larger than a second cutoff $RMSD_{th}$ and the difference in rotational constants ΔB is smaller than a third cutoff B_{th}, the two structures are considered rotamers. If all differences $\Delta E, \Delta B, \Delta RMSD$ are smaller than the respective thresholds, they are considered identical and one of the duplicates is discarded. A very good illustration of the detailed workflow can be found in Figure 5 of ref. 151. Metadynamics based conformer sampling as introduced in this section is independent of the electronic structure method used in the metadynamics and optimization runs. However, CREST is closely connected to the semi-empirical (GFN1-xTB [154], GFN2-xTB [155]) or force field (GFN-FF [156]) methods of the xTB family [157] developed by Grimme and coworkers. Note that conformer ensembles generated by CREST can be automatically processed further by means of the CENSO [158] tool – developed by the Grimme group as well.

In this thesis, conformer sampling is carried out entirely with the CREST program as implemented in the xTB program package. However, alternative protocols for finding low-energy conformers were published and a brief overview of alternative techniques is given in this section, starting with approaches for automatically finding transition states (TS) as these play a major role in asymmetric catalysis. A specialized approach that is used to find lowest-in energy transition states and that is therefore used in asymmetric chemistry research is the Q2MM [159] (Quantum-Guided Molecular Mechanics), and based on that the CatVS [160], approach, both by Rosales et al. The Q2MM tool is used to derive transition state force fields [161] (TSFFs) from quantum chemistry (QM) calculations. For this, the free parameters of available force fields are fitted to QM obtained data such as energies, geometric parameters, electrostatic potentials and, of course, hessians. QM hessians of TSs feature exactly one negative frequency. The TSFF, however, shall be usable for conformer samplings of transition states, which is (without geometric constraints) possible only if the transition state mode is at minimum in the MM force field. The QM hessian is therefore modified by replacing the negative eigenvalue by a large positive number, making the hessian positive definite [159]. A TSFF obtained by Q2MM is reaction specific, but might be fitted to multiple different ligands and substrates. The CatVS tool is based on the Q2MM generated TSFFs and uses them to screen a set of user-defined substrates and ligands [160]. Another available tool is AARON [162,163] by Wheeler and coworkers, which can be used for ligand screening and conformer sampling as well. AARON is connected to the Gaussian program package [164] and provides automated workflows at the DFT level. It

should be noted here, however, that exploration of the conformational space in both packages, CatVS and AARON, is carried out by rotations around certain single bonds leading to different energy minima, in contrast to CREST, which uses MTD simulations. Another force field based method for generating approximate transition states is SEAM, introduced by Jensen et al. [165-168]. Within their approach, the transition state is modeled as the minimum energy point at the intersection of the force field potential energy surfaces of the reactant and product states, meaning that the energy is minimized with respect to the geometry with the constraint that the reactant and product surfaces intersect. By specifying educt and product states, SEAM might be used for ligand screening by finding approximate low-energy transition states. If needed, fully relaxed transition states can be obtained afterwards from QM calculations.

Chapter 3 Basics of Asymmetric Organocatalysis

3.1 History

Organocatalysis describes the acceleration of chemical reactions through metal-free catalysts. Typically, organocatalysts are small molecules and therefore distinguishable from enzymes, which are large biocatalysts in living systems. Together with transition metal catalysis and enzymatic catalysis, organocatalysis forms the three pillars of homogeneous catalysis [169]. The importance of organocatalysis for humans and life in total might date back to the very beginning of chemical evolution on Earth. As discussed by Pizzarello and Weber [170], small and chiral amino acids can catalyze the stereoselective formation of small sugars from formaldehyde and glycolaldehyde. Since small amino acids were present in the prebiotic environment, they might have been important for the formation of small biomolecules as well. The first organocatalytic transformation is attributed to von Liebig, who discovered the formation of oxamide from dicyan catalyzed by an aqueous solution of formaldehyde in the second half of the nineteenth century [171]. The first observations of asymmetric catalysis were made in enzymatic systems by Pasteur, who performed kinetic measurements for the decarboxylation of ammonium tartrate by the mold *Penicillium glaucum*, which was able to decompose one enantiomer faster than the other one [172]. After the first world war, interest arose especially in nitrogen-containing chiral natural products. A very prominent class of chiral alkaloids in asymmetric catalysis are cinchona alkaloids, named after the flower genus they can be harvested from. A good overview of them has been given by Tian et al. [173]. Wegler investigated the stereoselective esterification of racemic 1-phenylethanol with different carbon acid chlorides and anhydrides catalyzed by the chiral alkaloid brucine [174]. He found a) a significant increase in reaction rates compared to the uncatalyzed reaction and b) that the reaction rate of one of the enantiomers of the alcohol is increased much more than the reaction rate of the other enantiomer. In the 1950s, asymmetric organocatalysis began to play a role in synthesis strategies. Prelog and Wilhelm [175] investigated the kinetics and reaction mechanism of the already known stereoselective cyanohydrin synthesis with aldehydes, hydrogen cyanide and chiral alkaloid bases in detail, while Pracejus [176] synthesized chiral propionates from phenylmethylketene and small alcohols catalyzed by chiral alkaloids. In addition, stereoselective aldol reactions [177] and Michael additions [178] with alkaloid catalysts were reported in the early 1970s. Enamines as structure motifs were already known by Mannich [179] in the mid-1930s and synthesis strategies like α-alkylations and α-acylations of ketones based on enamines were proposed by Stork et al. [180]. However, one of the cornerstones of asymmetric organocatalysis was the Hajos-Parrish-Eder-

Sauer-Wiechert reaction [181,182], which is an enamine-based intramolecular aldol reaction catalyzed by proline – a small heterocyclic and chiral amino acid that may be used in total syntheses of natural products. Proline-based enamine reactions are an important branch of covalent organocatalysis, they will be revisited in the next section.

3.2 Covalent catalysis

The catalytic behavior of organocatalysts can be categorized into two groups: covalent and noncovalent catalysis [172]. In covalent catalysis, the catalyst forms covalent bonds with the substrate to yield an active species that undergoes subsequent reaction steps. On opposition, in noncovalent catalysis, no covalent bonds are formed. Instead, the emulation of an enzymatic pocket is carried out. Noncovalent interactions like London dispersion, electrostatics, hydrogen bonds and steric repulsion shape the catalyst-substrate adduct and determine therefore the stereoselectivity of the transformation. Noncovalent catalysis is discussed in section 3.3. In section 3.4, an overview of typical interaction patterns determining stereoselectivity is illustrated. In this section, enamine catalysis serves as example of covalent organocatalysis.

As has been discussed in section 3.1, enamine reactions are an important type of stereoselective organocatalytic reactions. Since the carbonylic substrate reacts with a chiral secondary amine to form the enamine, they count as covalent catalysts. Besides alkaloids, proline is very often used as chiral reagent. Proline is non-toxic and commercially available in both enantiomers. It is the only secondary canonical amino acid and is therefore more basic and more nucleophilic at the nitrogen atom than other amino acids [172]. Well-known chiral auxiliaries like Enders reagent [183] are based on proline. In **Figure 3.1**, the general enamine cycle is given.

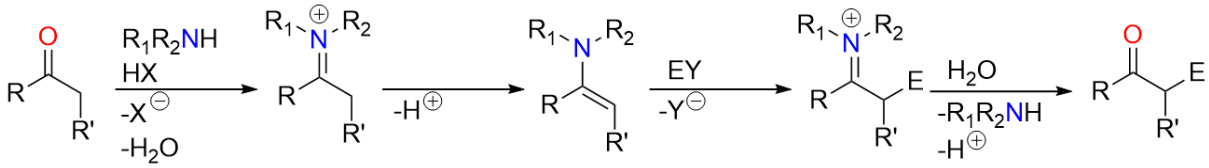

Figure 3.1 General enamine catalytic cycle. A general electrophile EY and an acid HX were introduced.

The general electrophile EY in **Figure 3.1** can be of very different kind. As discussed for the Stork reaction [180], alkyl halides or acyl halides might serve as reagents for alkylation or acylation of carbonyl compounds. However, the majority of used electrophiles are carbonyl compounds themselves. This leads to various different aldol reactions with aldehyde donors [184,185], α-oxyaldehydes [186], α-hydroxyketones [187] and intramolecular [188,189] aldol

reactions. Besides aldol reactions, proline and its derivates have been used in enantioselective Mannich reactions [190-193]. So far, only carbon electrophiles have been considered. However, proline catalyzed enamine reactions can be used to introduce heteroatomic groups in α-position as well. The interested reader is referred to ref. 194 and the literature cited therein. A related, but still different reaction is the proline catalyzed asymmetric addition of ketones to electrophilic olefins. The addition of ketone-based enamines to nitro alkenes was developed by List and coworkers [195] and refined by Enders and coworkers [196]. The general reaction cause is given in **Figure 3.2**.

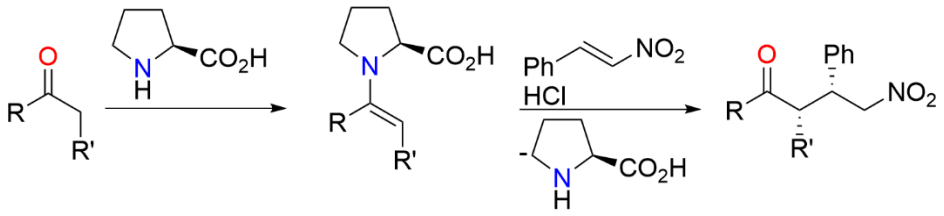

Figure 3.2 Asymmetric addition of nitro alkenes to enamines.

Using (*S*)-proline in the reaction from **Figure 3.2** leads to the *syn*-addition of the enamine. Enders et al. [196] proposed the *syn*-diastereoselectivity to origin from an intermolecular hydrogen bond between the proline carboxyl group and the olefin nitro group.

3.3 BINOL based noncovalent catalysis

In this section, noncovalent catalysis is introduced. Because of its importance for this thesis, special attention is paid to BINOL based Brønsted acid catalysts. Asymmetric Brønsted acid catalysis refers to the stereoselective induction through strong hydrogen bonds or even through acid base ion pairs [197]. Of importance for this thesis are BINOL based catalysts. 1,1'-Bi-2-naphtol (BINOL) is a diol that shows axial chirality. Its Lewis structure is given in **Figure 3.3a**. BINOL derivates have successfully been used in asymmetric Morita-Baylis-Hillman reactions [198-200]. However, BINOL can also serve as starting point for the synthesis of more acidic systems. Consequently, over the years different families of Brønsted acids have been derived from BINOL, the most important ones are introduced in this section. The first class introduced here are chiral phosphoric acids (CPAs) [201]. Their general structure is given in **Figure 3.3b**.

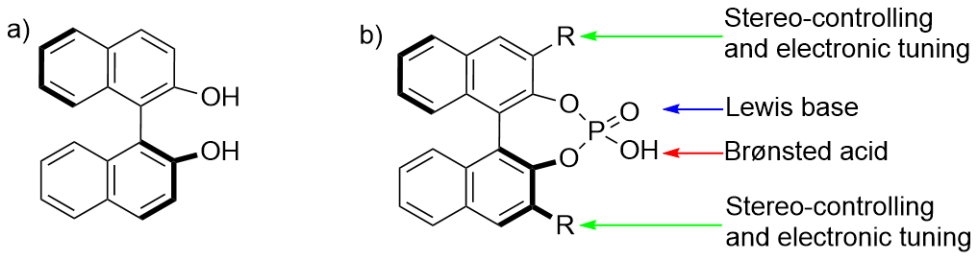

Figure 3.3 a) Lewis structure of (*S*)-BINOL. b) General Lewis structure of CPA type acids derived from (*S*)-BINOL.

The core of the catalyst is a description of the active site of a catalyst and its nearest atoms used in this thesis. In **Figure 3.3b**, the core catalyst would be the phosphoric acid group. The core catalyst of CPAs has an amphoteric nature with the acidic hydroxy group and the Lewis basic P=O group. Increased Brønsted acidity usually increases the reactivity of the catalyst. The substituents in 3,3' positions are used to introduce bulky ligands to the system, making the catalyst more rigid. They can, however, also be used to tune the electronic properties of the catalyst, *e.g.*, the ability to interact *via* London dispersion interaction. The substituents therefore critically influence the stereoselectivity of a given transformation. This point will be discussed later in the thesis (see section 4.3.4). BINOL based CPAs were introduced in 2004 independently by Akiyama et al. [202] and Terada and coworkers [203]. Maybe the best-known CPA that showed great stereoselectivity for various transformations [204-206] is TRIP, which uses the 2,4,6-triisopropylphenyl residue as ligand for the CPA [207,208]. TRIP originally was introduced by List and coworkers in 2005 for the stereoselective reduction of imines for which they achieved up to 93 % enantiomeric excess (e.e.) [204]. For a nice overview of CPA catalyzed transformations, see ref. 201.

The success of chiral phosphoric Brønsted acids in stereoselective transformations stimulated researchers to explore different types of BINOL based Brønsted acids. The considered recipe is to increase Brønsted acidity while tuning the substituents to keep stereoselective induction effective. Over the years, various different acid classes were proposed, many of them by the List group. Some important developments will be discussed now. In 2009, Pousse et al. [209] reasoned that the relatively low acidity of CPAs was responsible for the encountered difficulties in asymmetric Nazarov cyclization (the authors mention a successful asymmetric CPA catalyzed Nazarov cyclization in their publication [210] though). Furthermore, they ought to increase the acidity of CPA by replacing the oxygen sites by sulfur sites, since the sulfuric acid anions should be able to delocalize the negative charge more effectively than the oxygen sites, leading to chiral dithiophosphoric acids (DTPA). In addition, they decided not to use BINOL

as backbone, but H_8-BINOL, which is a partly hydrated derivate of BINOL. Although Pousse and coworkers could not achieve their goal of a stereoselective Nazarov cyclization, they obtained good diastereoselectivities and a generally high reactivity of their DTPA catalysts. In 2011, Shapiro et al. [211] reported the stereoselective intramolecular hydroamination with H_8-BINOL based DTPAs, which showed much higher stereoselectivities than the dehydrated BINOL based DTPA catalysts. Also in 2009, List and coworkers introduced the BINOL based Disulfonimide (DSI) catalyst class [212]. They reasoned that due to the compactness of the core functional group, the acidic center is more deeply buried in the BINOL catalyst pocket, presumably leading to better transfer of stereoinductive information. Indeed, they found great reactivity and stereoselectivity of their DSI catalysts for Mukaiyama aldolizations, for which the corresponding CPA did not react at all [212]. In addition, they considered the silylated DSI species to be very Lewis acidic and thus its general usefulness for asymmetric counter anion directed catalysis (ACDC). Before discussing further catalyst classes, the concept of ACDC is introduced. Deprotonation of CPAs leads to chiral acid anions that still can induce stereoselectivity. In 2006, Mayer and List [213] formed salts from TRIP and secondary amines. The corresponding chiral ion pairs (CIP) were then used for asymmetric hydrogen transfer reactions of α,β-unsaturated aldehydes, for which they showed great enantiomeric ratios (e.r.) of over 98% for all used substrates. List and coworkers also considered the already discussed Mukaiyama aldolization [212] catalyzed by DSI catalysts to run through an ACDC mechanism. Thus, the concept of ACDC is not limited to specific catalyst classes or reactions, but is a rather general concept. Two good reviews on ACDC have been published by Mahlau and List [214,215]. The general Lewis structures of DSIs are shown in **Figure 3.4**.

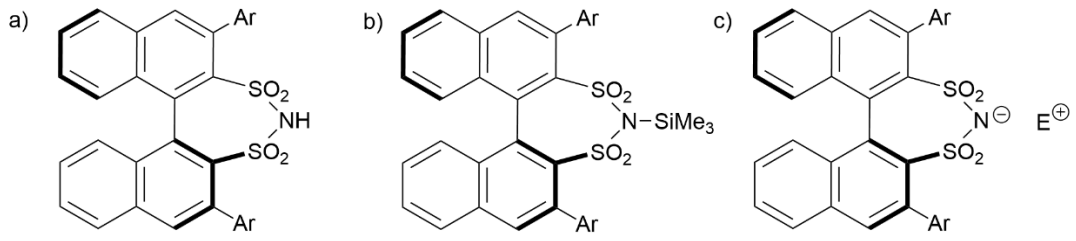

Figure 3.4 Lewis structures of DSI. a) The acidic system. b) The silylated system. c) The chiral ion pair with a general cation E^+.

In 2012, List and coworkers introduced another novel class of BINOL based acid catalysts, namely imidodiphosphates (IDP) [216]. Their synthetic problem at hand was the asymmetric spiroacetalization of hydroxy enol ethers. For this transformation, low stereoselectivity was achieved with commonly used BINOL based catalysts. It seemed that the oxocarbenium cation

that forms the chiral ion pair with the catalyst anion was too small, causing the ion pair to be too flexible, so that no asymmetric induction could be made. Their aim therefore was to introduce an even more confined and more rigid catalytic system that, on the other hand, still must be sufficiently acidic and reactive. To increase rigidness of the system, they introduced another BINOL subunit to the system. The acidic center consists of the imidodiphosphate group, which still features the amphoteric nature of CPA. The Lewis structure is given in **Figure 3.5a**.

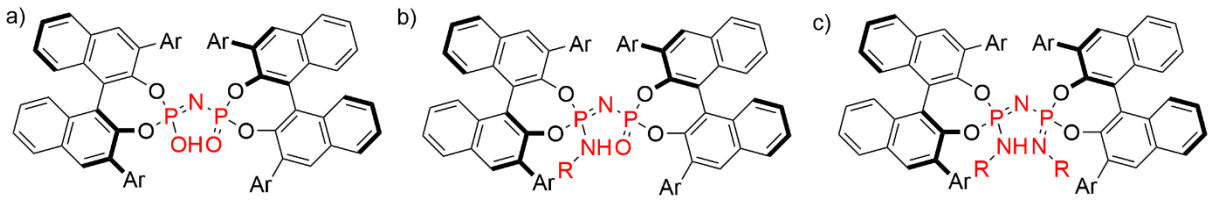

Figure 3.5 Lewis structures of a) IDP, b) iIDP, c) IDPi. The catalytically active centers are shown in red for better visualization.

In total, the IDP acid features four aryl substituents in the 3,3' positions of the BINOL subunits, leading to a much more confined reaction space. It is noted here that IDP catalysts can act *via* the ACDC mechanism as well, as was suggested in the literature [216,217]. In general, different catalyst classes can be obtained from IDP by replacing one or both of the oxygen sites of the core catalyst by the corresponding nitrogen functionality. This leads to iminoimidodiphosphates (iIDP, **Figure 3.5b**) or imidodiphosphorimidates (IDPi, **Figure 3.5c**). iIDPs were designed by the List group in 2016 when they found that IDP catalysts were not able to catalyze the stereoselective Prins reaction [218]. The basic idea is to increase the acidity of the catalyst further by replacing the OH group by an RNH group with an electron withdrawing substituent R (such as triflyl). The introduction of substituents to the core catalyst also allows for better fine tuning of the electronic properties [218]. By replacing, in addition, the keto group by the imine functionality, IDPi catalysts are obtained. IDPi catalysts were considered first in 2016 for the asymmetric addition of allyltrimethylsilane to aldehydes [219]. IDPis are even more acidic than iIDPs while still being sterically demanding due to substituted BINOL residues and the ligands at the core catalyst. Because of the massive success of IDPis, they have been used in very different organocatalytic transformations, such as Nazarov cyclizations [220], Diels-Alder reactions [221,222], intramolecular hydroalkoxylations [223], intramolecular hydroarylations [224] or aldolizations [225,226], just to name a few. Also, all organocatalytic transformations that will be studied in this thesis are catalyzed by IDPi catalysts. The most recent class of BINOL based catalyst classes are imidodiphosphorbis(iminosulfonylimino)-imidates (IDPii) that show

increased acidity with respect to IDPi catalysts by introducing $PhS(NTf)_2$ groups as electron withdrawing ligands at the core catalyst's iminyl groups [227].

The unifying concept that is behind the illustrated hierarchy of catalyst architectures developed in the List group is confinement – which List and coworkers define as the shaping of the catalyst's active site [228]. In biology, chemical reactions happen in structurally elaborated pockets of the enzymes, which constitute highly confined reaction spaces. The theme of host-guest complexes has also been applied in synthetic chemistry. For example, Hastings et al. performed Nazarov cyclizations in supermolecular hosts of tetrahedral Ga_4L_6 (L=*N,N′*-bis(2,3-dihydroxybenzoyl)-1,5-diaminonaphthalene) adducts [229]. Using these confined catalysts leads to increased reaction rates by a factor of 10^6 with respect to the uncatalyzed cyclization reaction. This significant increase in reaction rates was explained by catalyst-substrate interactions leading to a U-shaped conformation of the substrate, which facilitates the subsequent cyclization reaction. List and coworkers use this principle of confinement as guiding principle for the development of novel catalyst patterns. Typically, small substrate molecules form well-defined host-guest complexes with the catalyst, thus emulating the biological enzyme-substrate complexes. The definition of confinement, as given by the List group, is very neutral in its formulation. However, often confinement is negatively connotated. Perhaps this stems from the perception that confined reaction spaces, and therefore reaction rates and stereoselectivities, are governed by steric repulsion. This does not have to be the case. As has been discussed in section 2.4.4 and will be discussed in section 3.4 and section 4.3, attractive noncovalent interactions can be crucial for the shaping of the host-guest complex as well. This can lead to situations in which the bulkier, *i.e.*, more confined catalytic systems are actually more reactive, leading to increased reaction rates because of strong attractive noncovalent interactions [131,230]. In the author's mind, this would lead to a more positive connotation of the term "confinement", since stronger and more refined noncovalent interactions could lead to a simultaneous increase of reaction rates and stereoselectivities. The IDPi catalyst class, for example, manages to combine reactivity and selectivity while being a very confined catalyst system. Following these remarks, using confinement as a guiding principle for rational catalyst design seems reasonable.

3.4 Computational studies of catalyst-substrate interactions in organocatalysis

The developments of organocatalytic transformations outlined in the previous sections were accompanied by quantum chemical calculations, mostly DFT calculations. Over the last years, it became more and more standard to support experimental work by DFT calculations, which is, at least partially, caused by increased computational power, improved density functionals,

basis sets, dispersion corrections and such, as well as the rise of automatic workflows and black box methods. In the field of enantioselective catalysis, special interest exists in the stereo-determining interactions. In this section, some of the insights gained from computations are discussed. An extensive review including covalent catalysis, such as proline catalysis including the Houk-List model, was given by Houk and coworkers in 2011 [231]. The highly effective catalytic systems introduced by List and coworkers discussed in section 3.3 used to constitute an obstacle for computational chemistry. The systems under investigation typically consist of more than 100 atoms and therefore require a certain amount of computational power. In addition, many of the used catalysts or substrates contain very flexible substituents, thus leading to large sets of possible conformations that the system can take. One of the key quantities in stereoselective catalysis is the enantiomeric ratio which is, under kinetic control and following standard transition state theory and Curtin-Hammett principle (see ref. 232 for an overview), determined by the free energy difference between the transition states leading to the major and minor enantiomeric product, respectively [233]. However, this free energy difference usually is around 1 kcal mol^{-1} or even smaller, even for reactions that proceed essentially in an enantiopure fashion. This sets some challenging requirements for the computational protocols, even for a qualitative description of stereoselectivity. The efforts achieved in theoretical chemistry that made investigation of stereoselective transformations routinely possible originated from different research fields. First, some efficient schemes for finding the lowest-in energy transition state conformations leading to the respective enantiomeric products are mandatory for an accurate description of such transformations. Some widely used schemes like the CREST tool by Grimme were discussed in section 2.5. In addition, highly accurate methods for obtaining single point energies are required. In recent years, linearly scaling local correlation methods like DLPNO-CCSD(T) were proven to yield the desired accuracy (see section 2.3.3). As discussed before, noncovalent interactions play a crucial role in many transformations, so that tools for accurate quantification of those interactions, like LED, are required as well. The combination of semi-empirical conformer sampling, local Coupled-Cluster single point energies and LED analyses were first carried out by Yepes, Bistoni and coworkers in 2020 [234]. They investigated the Diels-Alder reaction of cinnamate esters with cyclopentadiene catalyzed by IDPi catalysts. The reaction runs *via* an ACDC-based mechanism in which the catalyst and the cinnamate esters form chiral ion pairs whose geometries and binding energies are determined by electrostatic as well as dispersive interactions. The cyclopentadiene molecule entering the catalyst pocket interacts strongly *via* LD with the CIP, which in turn changes its conformation slightly in order to allow for optimal interaction. The stereoselectivity of the

reaction is therefore governed by the balance of steric repulsion and attractive noncovalent interactions. In their journal article, they proposed a multi-step protocol that works like a filter for the candidate structures. In the first step, some initial guess structures for the respective transition states are obtained, *e.g.*, from some preliminary DFT calculations. Using CREST, a large number of different transition state conformers is generated. Generally, the conformer sampling with CREST is carried out at the semi-empirical GFN-xTB or even at the force field level of theory. For obtaining more reliable structures and, more importantly, reliable relative energies of the conformers, they are subject to constrained DFT optimizations using standard density functionals [234]. The geometric constraints are used to keep relevant modes in their transition state position. Normally, one or a few bond distances or bond angles are kept frozen during the conformer sampling and the subsequent constrained DFT geometry optimizations. The lowest-in energy conformers obtained from the constrained DFT optimizations are then further optimized without geometric constraints leading to fully relaxed transition states. Yepes et al. then performed DLPNO-CCSD(T) calculations and subsequent free energy calculations in which they added thermochemical corrections (such as zero-point vibrational energies or thermal and entropic energy corrections) and free solvation energies obtained from DFT calculations invoking the conductor-like polarizable continuum (C-PCM) [235] implicit solvation model to the electronic energy. Modified versions of this protocol have been used throughout this thesis.

However, recently the multi-level combination of metadynamics-based conformer sampling, DFT geometry optimizations on a relatively cheap level of theory and single point energy calculations on a more expensive level of theory (be it DLPNO-CC or more expensive DFT functionals like ωB97M-V) has been used by other groups as well. For example, the group of Houk used such protocols for investigations on the origin of stereoselectivity of the nickel-catalyzed stereoselective alkyl-carbamoylation of olefins [236] or the CPA-catalyzed stereoselective allylation of aldehydes with β-alkenyl allylic boronates [237]. The group of Wheeler combined constrained CREST-based conformer sampling with subsequent DFT optimizations for exploring the palladium-catalyzed alkene insertion step that determines the stereoselectivity of the addition of aryl boronic acids to chromones [238]. The group of Schreiner combined experimental and computational approaches to describe the *syn-syn* and *anti-anti* equilibrium of substituted *N,N'*-diphenylthiourea using CREST for conformer sampling and DLPNO-CCSD(T) for the quantification of noncovalent interactions [239]. They found that the sterically more demanding *syn-syn* conformer is favored over the *anti-anti* conformer. The *syn-syn* conformer becomes more stable on a relative scale with the increase of the substituents due

to increased LD interactions. In another study, Schreiner and coworkers used cyclooctatetraene-based molecular balances to measure the hydrogen bonding energy in the cyclic water dimer, which represents a transition state on the water dimer potential energy surface and is therefore not directly observable [240]. By introducing CH_2OH groups to the molecular balance, they could mimic the cyclic water dimer in the 1,6-isomer of the cyclooctatetraene. In their study, they used LED and SAPT2 for quantification of hydrogen bonding energies. Further examples of the usage of those multi-step protocols by the Schreiner group can be found in literature [241-243].

From the Grimme group, the combined experimental and computational work on catenated organocatalysts in collaboration with the Niemeyer group is mentioned here. In a series of published works, they explored the catalytic activity of interlocked CPA acid molecules [244-247]. The CPA molecules are catenated with the help of large oxa-macrocycles, allowing to study the influence of cooperative effects of confined catalytic systems on reactivity and selectivity. They compared the yield and stereoselectivity obtained with a) two catenated CPA molecules, b) one catenated CPA molecule and c) the acyclic "free" CPA catalyst for the stereoselective transfer hydrogenation of quinolines by Hantzsch esters, which consists of two reduction steps [244]. For the catalyst with two CPA molecules, they found increased stereoselectivity, which they rationalized with a sandwich-like transition state in which the substrate has strong interactions with both catalyst molecules. In contrast, for the catenanes with a single CPA molecule, they found very low but inversed stereoselectivity, which they explained with additional substrate-macrocycle interactions. In a follow-up paper, Grimme, Niemeyer and coworkers investigated cooperative effects of the non-interlocked CPA catalysts for the same catalytic reaction by means of spectroscopic, kinetic and computational techniques [246]. Their results show that monomeric and dimeric pathways compete, and the ratio between those two pathways depends on the catalyst loading. Interestingly, the stereoselectivity of the dimeric pathway, which is governed by the second reduction step, is increased with respect to the monomeric pathway. While the reaction rate of the first reduction is smaller for the dimeric pathway, it becomes larger for the second pathway, demonstrating the importance of cooperative catalytic effects for both reaction rates and stereoselectivities.

Chapter 4 An Induced-Fit Model for Asymmetric Organocatalytic Reactions

4.1 Introduction

Elements of this chapter including figures have been published in a journal article [248]:

Harden, I.; Neese, F.; Bistoni, G. An induced-fit model for asymmetric organocatalytic reactions: a case study of the activation of olefins via chiral Brønsted acid catalysts. *Chemical Science* **2022**, *13* (30), 8848-8859. https://doi.org/10.1039/D2SC02274E.

If not stated explicitly otherwise, all calculations and analyses were performed by the thesis author. Also, all figures used in this chapter have been designed by the thesis author. For each figure, the corresponding figure caption describes whether the figure is part of a publication. As has been discussed in section 3.3, enzymatic reactions happen in structurally elaborate pockets. The different confined catalysts that were introduced in the List group aim to mimic the enzyme-substrate interaction patterns with regard to the highly confined reaction space, allowing for various stereoselective transformations. The question that arises then is how the stereoselectivity actually is induced. The confinement of a reaction space itself is governed by a complex pattern of interactions that lead to a favorable conformation of the catalyst and the substrate, which subsequently determines the stereoselectivity of the transformation. Of course, these interactions can be of very different nature, either attractive or repulsive, covalent or noncovalent, etc. Especially in the past, steric repulsion was considered to be the most important interaction determining the stereoselectivity. Steric repulsion is based on the short-range repulsion of atoms, leading to stronger repulsion for increased bulkiness of two interacting groups. Stereoinduction based on steric repulsion works by blocking undesired pathways so that the reaction path leading to the major enantiomeric product is the one that minimizes steric repulsion. As has been discussed in section 3.3 and by Wagner and Schreiner [131], more confined catalysts should show reduced reactivity if the confined reaction space were actually entirely defined by steric repulsion. However, since reactivity often increases simultaneously with confinement, also attractive noncovalent interactions were considered important for the description of stereoselective transformations. Especially LD was shown to be an important interaction component in many reactions, as it is always present and increases with the size and polarizability of the interacting groups. Steric repulsion and London dispersion both increase with the size of the interacting groups while having opposite signs, so it is expected that the

interplay between these two components will play a major role for many transformations, as was shown by Yepes et al. for IDPi-catalyzed Diels-Alder reactions [234]. The aim of this project was to propose a model for the interaction patters governing the stereoselectivity of IDPi-catalyzed hydroalkoxylation reactions. Of course, the proposed model shall be of general importance. However, it should be kept in mind that for different reactions different interactions might determine stereoselectivity. This is because the reaction mechanism might be very different for different transformations. For example, for reactions that go through an ACDC-like mechanism, electrostatics is supposed to be important. The same is true for reactions that go through charged reaction intermediates. For reactions with charged species, also interactions with the solvent may become important.

The stereoselective intramolecular hydroalkoxylation reaction of small and unactivated olefins catalyzed by IDPi catalysts was used as case study. This reaction was published by the List group in 2018 [223]. The overall reaction and experimental conditions are given in **Figure 4.1**.

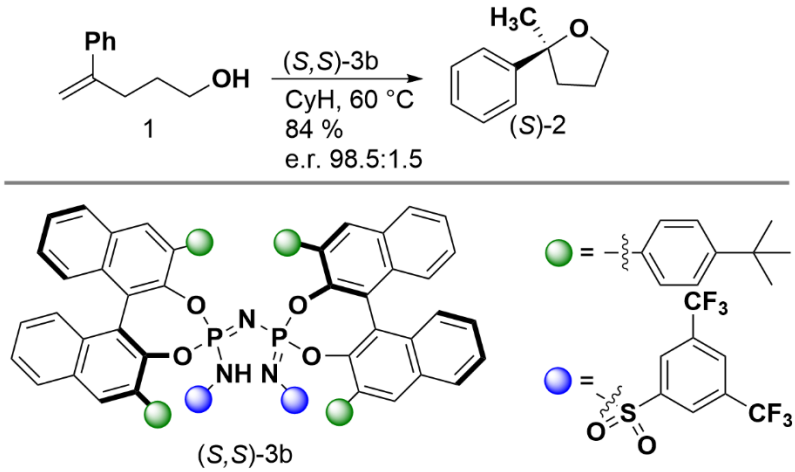

Figure 4.1 Reaction and experimental conditions for the selected case study reaction. This figure was taken from ref. 248.

As can be seen from **Figure 4.1**, the intramolecular hydroalkoxylation reaction of the substrate **1** leads to chiral tetrahydrofuran derivates in good yields and excellent stereoselectivities, even for increased temperature (60 °C). The reaction is catalyzed by the catalyst *(S,S)*-**3b** that features bulky *tert*-butylphenyl ligands. For the outlined reaction, an in-depth analysis of the key interactions determining the stereoselectivity is still missing. In the original publication by List and coworkers [223], preliminary DFT calculations were performed, however, some important aspects such as quantification of NCIs or taking into account the conformational freedom of the species were not included. Therefore, the aim of the present work was to address these issues.

4.2 Computational protocol

If not stated otherwise, the DFT and DLPNO-CC calculations reported in this chapter were performed using ORCA 4.2.1. The computational protocol used in this work is based on the protocol published in 2020 by Yepes et al. [234]. It is illustrated in **Figure 4.2**.

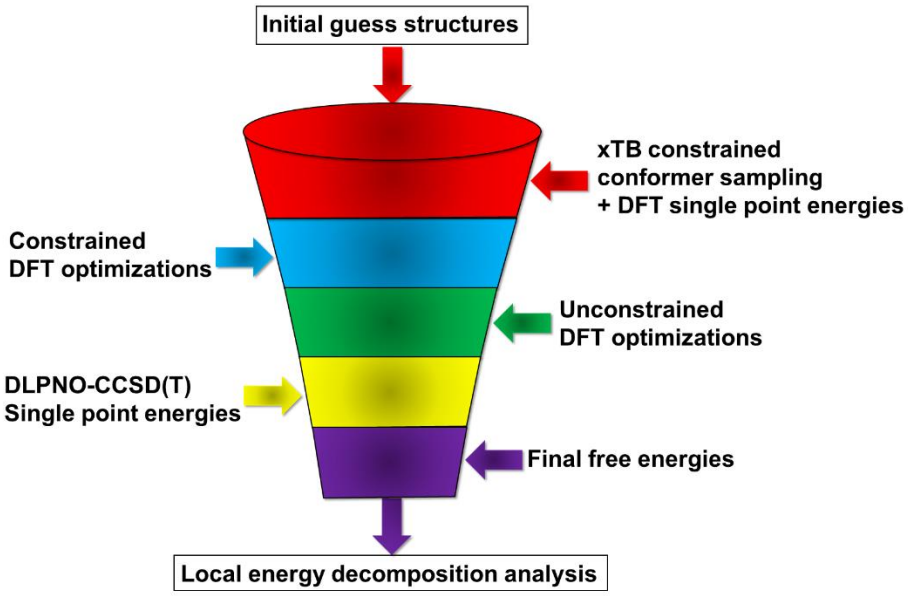

Figure 4.2 Computational protocol used for the intramolecular hydroalkoxylation reaction. This figure was taken from ref. 248.

Like the protocol by Yepes et al., the protocol presented in **Figure 4.2** works like a filter for the candidate structures, that is, the number of remaining candidate structures becomes smaller for each subsequent step. The aim of the protocol is to find the lowest-in energy transition states (TSs) leading to the respective enantiomers. In the first step, some initial guess structures for the TSs have to be defined. This might be done by some preliminary DFT optimizations. For this project, the DFT optimized transition states proposed by List and coworkers [223] were used as initial starting structures. The second step (red phase in **Figure 4.2**) is the constrained conformer sampling at the GFN-xTB level of theory using the CREST program developed by Grimme and coworkers. Geometrical constraints were used in order to keep the transition state mode of the structure intact. For this, the bond distances of the breaking N-H bond and the forming C-H and C-O bonds were restrained. It is noted here that exact constraining (*i.e.*, freezing) would make the metadynamics simulations unstable, which is why exact constraining is not used in CREST. For each transition state (major and minor pathway), multiple CREST runs were performed with different methods (GFN1-xTB or GFN2-xTB) and settings of the metadynamics runs (such as simulation times and force constants for the constraints). From

these runs, 6485 transition state conformers were obtained. Clearly, it is very demanding to optimize all structures of the conformer ensemble as it was done in the work of Yepes et al. [234]. This is why the whole ensemble was sorted by DFT single point energies. For this, the PBE [108] functional was used together with Ahlrichs double-zeta def2-SVP basis set. The D3 dispersion correction of Grimme was used together with Becke-Johnson damping. The resolution of identity in the Split-RJ variant was used together with the corresponding Coulomb-fitting basis set. This computational setup is denoted as PBE-D3/def2-SVP. The next step of the protocol (blue phase in **Figure 4.2**) consists of constrained geometry optimizations of the low-energy structures. For this, a cutoff for the relative energy of 4 kcal mol^{-1} was chosen. In addition, also selected higher-energy structures with significantly different geometrical features were included in the set of conformers as well. This was done in order to make the protocol more robust. In total, 655 TS conformers were optimized at the PBE-D3/def2-SVP level of theory. The N-H bond distance was the only used constraint during the optimizations. The next step of the protocol was to optimize the low-energy conformers with significantly different geometrical features without any constraints, leading to fully relaxed transition state structures. For the unconstrained TS optimizations, the larger def2-TZVP(-f) basis set was used. Thermochemical corrections were obtained at the same level of theory (PBE-D3/def2-TZVP(-f)) using Grimme's quasi-rigid-rotor harmonic oscillator (RRHO) approximation [249]. All transition states were proven to be first-order saddle points on the potential energy surface by the presence of a single negative frequency.

In the last step, final single point energy calculations were performed at the DLPNO-CCSD(T) [72-74] level of theory. The def2-TZVP basis set was used together with the corresponding correlation-fitting auxiliary basis set def2-TZVP/C. NormalPNO settings were used together with a tightened TCutPairs value of 10^{-5} Hartree. As mentioned earlier, the DLPNO construction is tightly linked to the resolution of identity. In addition, very tight SCF settings (VeryTightSCF in ORCA notion) and tight Grids (GRIDX7) were used. The RIJCOSX approximation to the exchange integrals was used to speed up the calculations [10,250]. From the DLPNO-CCSD(T) calculations, also dispersion interaction density [251,252] (DID) plots were generated that help to visualize the dispersive interactions between each pair of fragments. Solvation effects were captured using the solvation model based on density [253] (SMD) at the B3LYP [112,113] level together with the D3(BJ) correction and the def2-TZVP basis set and cyclohexane (CyH) as solvent. The SMD model includes, in addition to electrostatic interactions between solute and solvent, dispersive effects as well. The final Gibbs free energy G is eventually calculated *via* eq. (4.1).

$$G = E^{SP}_{DLPNO-CCSD(T)} + G^{solv}_{B3LYP} + G^{thermo}_{PBE}. \quad (4.1)$$

$E^{SP}_{DLPNO-CCSD(T)}$ is the DLPNO-CCSD(T) single point energy, G^{solv}_{B3LYP} is the implicit solvation correction, G^{thermo}_{PBE} is the thermochemical correction computed at 333 K (experimental temperature). The reaction barriers were decomposed with the help of LED using eq. (4.2). The system was divided into two fragments, namely the catalyst molecule and the substrate.

$$\Delta G^{\ddagger} = \Delta G^{\ddagger}_{corr} + \Delta E^{\ddagger}_{geo-prep} + E^{\ddagger}_{disp} + \Delta E^{\ddagger}_{non-disp}. \quad (4.2)$$

$\Delta G^{\ddagger}_{corr}$ contains the correction terms from thermochemistry (*i.e.*, zero-point vibrational energy, entropic and enthalpic effects) and solvation, $\Delta E^{\ddagger}_{geo-prep}$ is the geometric preparation, $E^{\ddagger}_{disp}$ is the dispersive interaction energy between the two fragments and $\Delta E^{\ddagger}_{non-disp}$ describes all other interaction terms, such as charge-transfer, polarization or steric effects. Note that the electronic preparation energy (see eq. (2.33)) was absorbed in the non-dispersive interaction.

4.3 Results and discussion

4.3.1 Results from the conformer sampling

In the first step of the protocol, constrained conformer sampling of the transition states leading to the major and minor enantiomeric products was carried out together with constrained DFT optimizations using the ORCA program package. As detailed in section 4.2, 655 constrained DFT optimizations were carried out at the PBE-D3/def2-SVP level of theory. From these, 459 conformers describe transition states leading directly to the stereoisomers of **2**. The other conformers describe for example transition states following alternative reaction mechanisms. The relative energies of the 459 surviving conformers are shown in **Figure 4.3**.

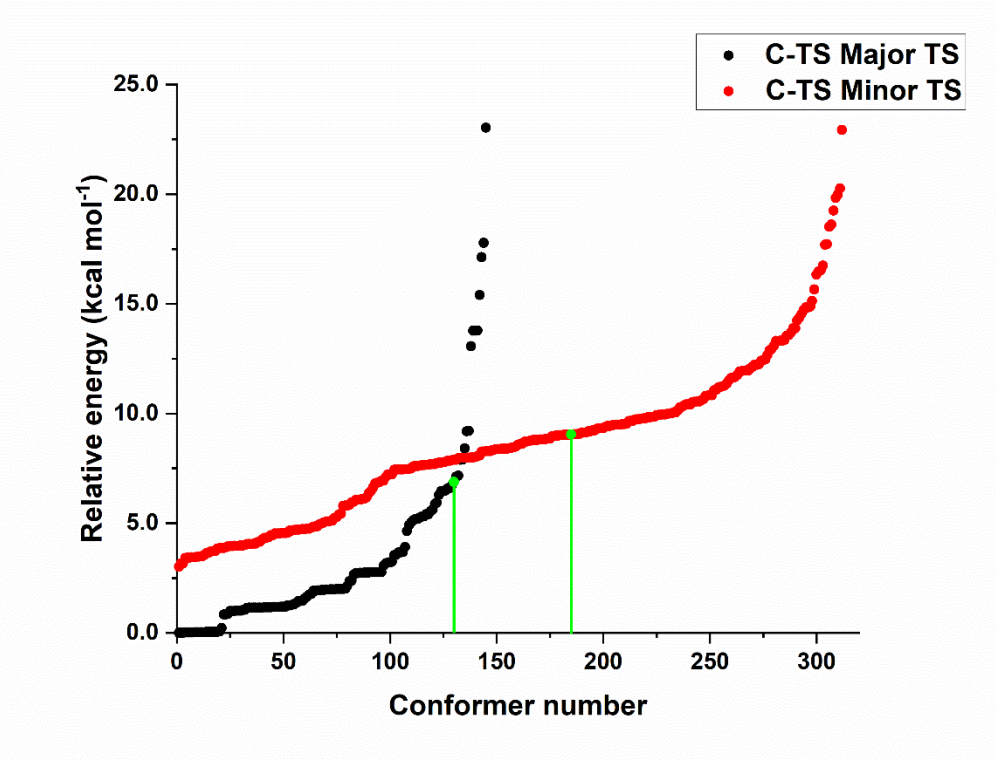

Figure 4.3 Relative energies of TS conformers leading to the major (black) or minor (red) enantiomeric product. Constrained geometries were obtained at the PBE-D3/def2-SVP level of theory. Only conformers following the concerted mechanism A (see section 4.3.2) are included. In addition, the relative energies of the initial guess structures, which are the transition states discussed in ref. 223, are shown in green. This figure was taken from ref. 248.

As can be seen from **Figure 4.3**, the conformers of the major TS are lower in energy than the corresponding minor TSs. Interestingly, for both pathways, the conformer sampling procedure identifies more than 100 conformers that are lower in energy than the corresponding initial guess structures, which are the TSs proposed in ref. 223. The initial guess TSs were re-optimized at the PBE-D3/def2-SVP level. Therefore, constrained DFT geometries were compared to fully relaxed transition states. However, the energy increase when a constrained optimized conformer is fully relaxed to a first-order saddle point is usually very small, so that the comparison still is reasonable. The results from **Figure 4.3** emphasize the importance of the conformer sampling, since the most stable major conformer is around 7 kcal mol^{-1} lower in energy than the initial guess structure. For large and flexible systems, exploring the conformational space might therefore be important to get even qualitative agreement with experiment.

4.3.2 Reaction mechanism

Before considering the stereo-controlling key interactions, the most probable reaction mechanism will be elucidated. In total, three reaction mechanisms are considered in this chapter: two concerted mechanisms "A" and "B" and step-wise mechanism "C". For both concerted mechanisms, initial guess structures for the conformer sampling were obtained from the publication by List and coworkers [223]. The three reaction mechanisms together with the free energy profiles are shown in **Figure 4.4**. Concerted mechanism "A" describes a mechanism in which the catalyst **3b** transfers its proton to the terminal carbon atom of the double bond of **1**. At the same time, the hydroxy group of **1** attacks the second carbon atom of the double bond and the oxacycle **A-P** is formed in a single step *via* transition state **A-TS1**. The concerted mechanism "B" describes a strong hydrogen bond assisted mechanism, which features an intramolecular proton transfer within **1**. In addition, a step-wise mechanism "C" was considered in this work. For this, the conformer sampling of the TSs was carried out without the geometrical constraint for the C-O bond. This leads to an unfolding of the substrate's carbon chain and thus allows for stabilization of the ion pair intermediate **C-I**. This ion pair cannot be stabilized if the substrate is folded as shown in Lewis-structure **A-E** because in this case C-O bond formation will always take place. The second step of mechanism "C" then describes the folding process of the substrate, which has to fold itself into the correct conformation in order to allow for C-O bond formation. As can be seen from the computed profiles in **Figure 4.4**, the three educt states **A-E, B-E** and **C-E** are very close in energy. Comparing the transition states of the concerted mechanisms **A-TS1** and **B-TS1,** it becomes evident that concerted mechanism "B" is highly disfavored by energy. The barrier for the hydrogen transfer of the step-wise mechanism is much smaller than the barrier of mechanism "B", it is, however, still significantly larger than the barrier of concerted mechanism "A". It can be seen that the ion pair intermediate **C-I** is only slightly lower in energy than **C-TS1**, indicating that the back reaction of the proton transfer is very fast. The second transition state **C-TS2** is very high in energy, which disfavors the step-wise mechanism as well, making concerted mechanism "A" the most probable one to occur. In the original work by List and coworkers also the non-terminal isomers of **1** were considered, however, product formation did not occur under experimental conditions [223]. If indeed a step-wise mechanism had occurred, **1** as well as the non-terminal isomers of **1** should form the exact same intermediate **C-I** and should therefore yield comparable yields and stereoselectivities.

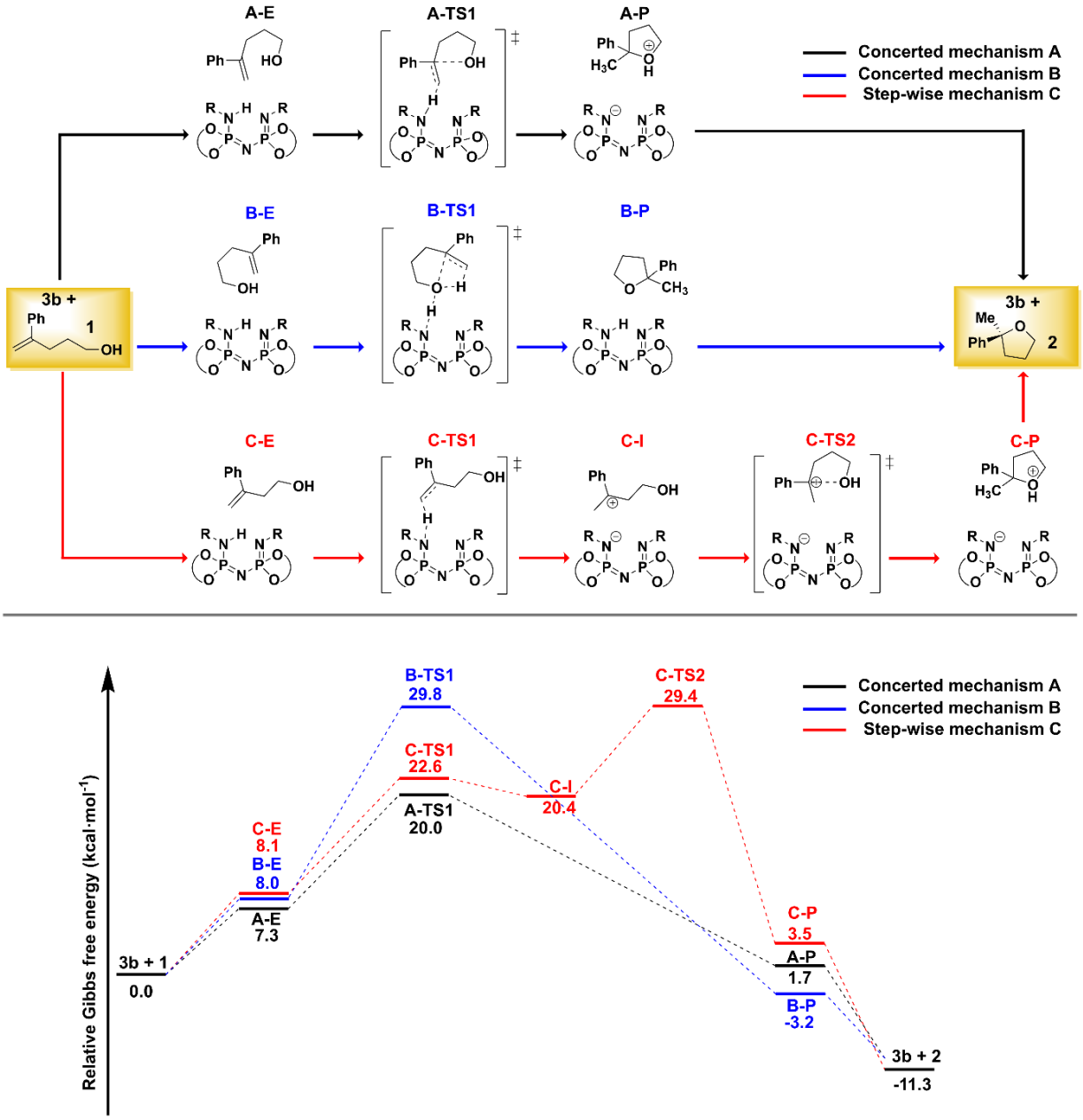

Figure 4.4 Reaction mechanisms and relative Gibbs free energies for all occurring species obtained from the outlined computational protocol. For identification of the species **3b**, see **Figure 4.1**. Only the pathways leading to the major enantiomeric products are shown. This figure was taken from ref. 248.

In **Figure 4.5**, the most stable transition states following concerted mechanism "A" leading to the major enantiomeric product (**TS1-1**) and minor enantiomeric product (**TS1-1', TS1-2'**) are illustrated. It was found that for the major pathway all low-energy conformers essentially show the same geometric features (the relative conformation of the catalyst and substrate). This is why it is sufficient to show only the most stable conformer here (**TS1-1**). Other low-energy major TSs essentially look like **TS1-1** and are therefore expected to show analogous intra- and

intermolecular interactions. For the minor pathway, the situation is different, since multiple different low-energy TSs arose that show significant different structural features. The lowest-in energy conformers are **TS1-1'** and **TS1-2'**, which are almost degenerate. Since their geometries differ significantly, all further analyzes in this chapter were carried out for both transition states. On the left side of **Figure 4.5**, the 3D structures are shown.

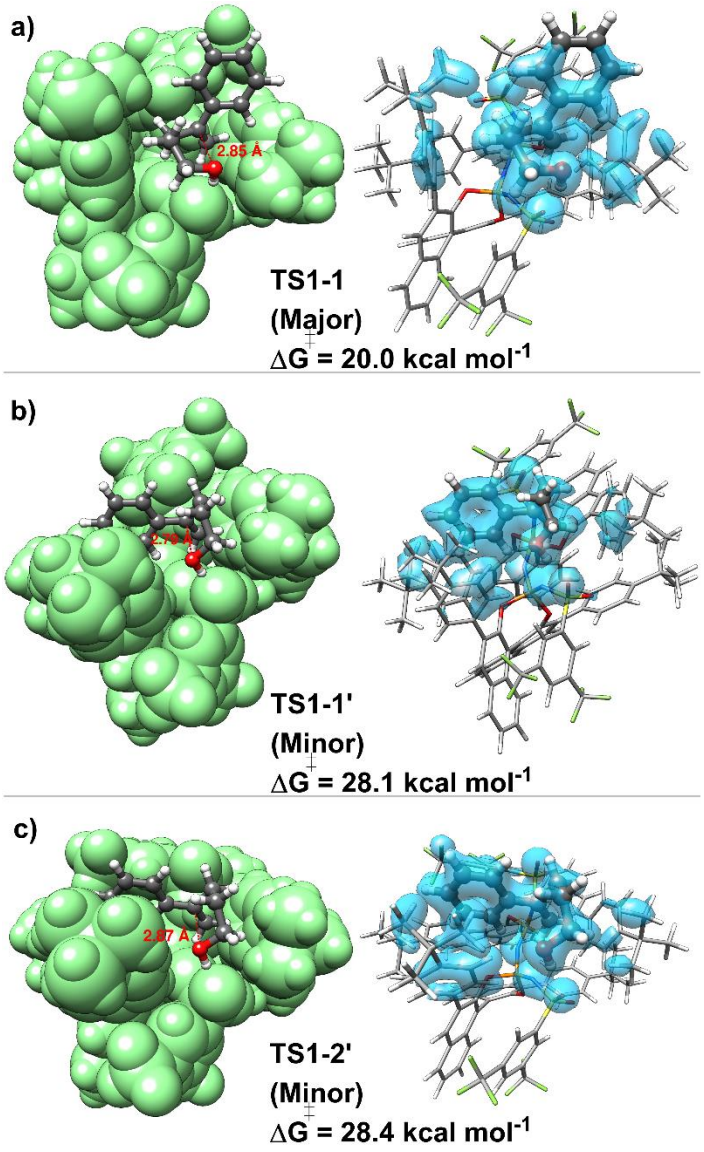

Figure 4.5 Most stable transition state structures leading to the major and minor enantiomeric product, respectively. The reaction barriers $\Delta G^{\ddagger}$ are given in kcal mol^{-1}. On the right, the dispersion interaction density plots for the interaction between the catalyst and the substrate are shown. For the isosurface, a density value of 0.01 kcal mol^{-1} bohr^{-3} was used. This figure was taken from ref. 248.

For the representation of the catalyst, a green-colored space-filling model was chosen in order to emphasize the conformation of the substrate within the catalyst pocket. For **TS-1**, it can be

seen that the substrate has a stable chair-like conformation. The C-O bond distance is 2.85 Å and is therefore very large compared to the covalent C-O bond distance. This indicates that the reaction happens concerted but asynchronous. The N-H bond distance as well as the C-H bond distance is 1.36 Å, showing that the proton that is transferred from the catalyst to the substrate is almost centered between the nitrogen and carbon site in the TS. On the other hand, the C-O bond formation is in its initial state. This is consistent with the experimental Hammett analysis indicating a partial positive charge is built at the corresponding carbon atom in the transition state [223]. The two most stable minor transition states **TS1-1'** and **TS1-2'** shown in **Figure 4.5b**,**c** belong to the same asynchronous mechanism as reflected by large C-O bond distances. The substrate, however, in both cases is rotated by around 90 ° with respect to **TS1-1**. This rotation of the substrate has been made possible by a bending of one of the aryl ligands towards the front (in direction of the reader in **Figure 4.5**). The main difference between **TS1-1'** and **TS1-2'** is the orientation of the aminyl-Ph-$(CF_3)_2$ group. On the right side of **Figure 4.5**, the dispersion interaction density plots are shown, which help to visualize the dispersive interaction between the catalyst and the substrate at the DLPNO-CCSD(T) level of theory. As can be seen, the substrate mostly interacts with the bulky *t*-Bu groups as well as the triflyl groups of the catalyst.

For the transition states **TS1-1** and **TS1-2'**, benchmark tests of the TS structure with respect to the basis set and density functional were carried out. The results are shown in **Table 4.1** and **Table 4.2**. As can be seen, the N-H and C-H bond distances are fairly independent of the chosen basis set or density functional and vary typically by less than 0.01 Å. The C-O bond distance is much larger and larger deviations are therefore expected and are observed. The data show that PBE-D3 and B3LYP-D3 almost yield identical C-O bond distances at the def2-TZVP(-f) level of theory and larger deviations only occur for the smaller def2-SVP basis set. This is true for both transition states **TS1-1** and **TS1-2'**. The deviation of the O---H hydrogen bond length between the substrate's hydroxy group and the catalyst's triflyl group is large with respect to basis set and density functional for both TSs. As quantifier for the overall difference of structures, the root-mean-square deviation of atomic positions (RMSD) was calculated, taking the PBE-D3/def2-TZVP(-f) optimized structures as reference. For computing the RMSD values, the thesis author wrote a command-line program that takes two .xyz files as input coordinates and calculates the RMSD value using either the Kabsch algorithm [254] or optionally the quaternion-based algorithm proposed by Coutsias et al. [255].

Table 4.1 Important bond distances of **TS1-1** optimized at different levels of theory. H---O denotes the hydrogen bond between the substrate's hydroxy group and the triflyl group of the catalyst. All bond distances are in Å. This table was adapted from ref. 248.

TS1-1	PBE-D3/ def2-SVP	PBE-D3/ def2-TZVP(-f)	B3LYP-D3/ def2-TZVP(-f)
N-H	1.337	1.358	1.363
C-H	1.370	1.360	1.364
C-O	2.686	2.855	2.886
H---O	1.792	1.891	1.900
RMSD	0.121	0.000	0.101

Table 4.2 Important bond distances of **TS1-2'** optimized at different levels of theory. H---O denotes the hydrogen bond between the substrate's hydroxy group and the triflyl group of the catalyst. All bond distances are in Å. This table was adapted from ref. 248.

TS1-2'	PBE-D3/ def2-SVP	PBE-D3/ def2-TZVP(-f)	B3LYP-D3/ def2-TZVP(-f)
N-H	1.367	1.370	1.368
C-H	1.330	1.329	1.337
C-O	2.787	2.868	2.858
H---O	1.800	1.924	2.004
RMSD	0.456	0.000	0.254

From the RMSD values given in **Table 4.1** and **Table 4.2**, it is seen that overall, the basis set has bigger influence on the geometry than the density functional. This is especially evident for **TS1-2'**. The usage of B3LYP-D3 increases the computational effort significantly without giving very different geometries, which seems to justify the usage of the PBE density functional together with the def2-TZVP(-f) basis set. All constrained DFT optimizations as well as fully relaxed TS optimizations were carried out *in vacuo*. For **TS1-1**, the TS optimization was repeated using C-PCM together with cyclohexane as solvent. Critical bond distances are given in **Table 4.3**.

Table 4.3 Important bond distances of **TS1-1** optimized at the PBE-D3/def2-TZVP(-f) level of theory *in vacuo* and with the C-PCM model. H---O denotes the hydrogen bond between the substrate's hydroxy group and the triflyl group of the catalyst. All bond distances are in Å. This table was adapted from ref. 248.

TS1-1	*In vacuo*	C-PCM(Cyclohexane)
N-H	1.358	1.367
C-H	1.360	1.354
C-O	2.855	2.929
H---O	1.891	1.887
RMSD	0.000	0.076

Using an implicit solvation model for the TS optimization has a fairly small impact on the geometry, as can be seen from the RSMD value of 0.076 Å. Noticeably, the influence of implicit solvation is much less important than, for example, the influence of basis set and density functional. Performing geometry optimizations *in vacuo* and using implicit solvation to compute the free solvation energy contribution to the total free energy of the system seams therefore reasonable.

From **Figure 4.4**, it can be seen that the difference in reaction barriers $\Delta\Delta G^{\ddagger}$ between the minor and major pathways amounts to 8.1 kcal mol^{-1} (**TS-1'**) and 8.4 kcal mol^{-1} (**TS1-2'**). Since the educt state is identical for both pathways, $\Delta\Delta G^{\ddagger}$ equals the free energy difference between the corresponding transition states. The experimental e.r. is 98.5:1.5 at 60 °C which translates to an energy difference of 2.8 kcal mol^{-1}. Thus, the computed energy differences are much larger than the experimental ones. In order to investigate this issue, in the first step, the electronic energy difference for five low-energy minor transition states (**TS1-1' – TS1-5'**) relative to **TS1-1** is given in **Table 4.4**. It is noted that **TS1-3' – TS1-5'** are not further discussed in this chapter, but are used here to increase the benchmark set. From **Table 4.4**, it can be seen that Hartree-Fock produces very large relative electronic energies and is therefore entirely inadequate for the description of stereoselectivity. Incorporation of electron correlation *via* DLPNO-CCSD reduces the relative energies significantly. However, especially for **TS1-1'** and **TS1-2'**, inclusion of perturbative triples excitations has major impact on the relative energies.

Table 4.4 Relative electronic energies at various levels of theory for six transition states leading to the major (**TS1-1**) or minor (**TS1-1' – TS1-5'**) enantiomeric product, respectively. All transition state structures were optimized at the PBE-D3/def2-TZVP(-f) level of theory. **TS1-4'** is the re-optimized initial guess transition state. All energies are in kcal mol^{-1}. This table was adapted from ref. 248.

	Hartree-Fock/ def2-TZVP	DLPNO-CCSD/ def2-TZVP	DLPNO-CCSD(T)/ def2-TZVP	PBE-D3/ def2-TZVP(-f)	B3LYP-D3/ def2-TZVP
TS1-1	0.0	0.0	0.0	0.0	0.0
TS1-1'	17.7	8.4	6.9	4.6	3.6
TS1-2'	17.7	10.1	8.7	5.8	4.4
TS1-3'	13.5	9.0	8.0	4.9	4.6
TS1-4'	13.3	10.7	10.1	8.3	8.9
TS1-5'	10.0	8.9	8.5	6.9	7.1

Noticeably, the PBE-D3 and B3LYP-D3 density functionals show much smaller relative energies than DLPNO-CCSD(T). This is especially true for B3LYP-D3 that yields a relative electronic energy of **TS1-1'** of 3.6 kcal mol^{-1}.

The next step was to investigate the influence of basis set and PNO space size on the relative DLPNO-CCSD(T) electronic energy. Most of the calculations in this chapter were carried out before ORCA 5.0 [78] was released and therefore were performed using ORCA 4.2.1. However, the benchmark calculations presented in the following tables were carried out using ORCA 5.0. Basis set extrapolations of the SCF and correlation energies were carried out using eqs. (4.3) and (4.4) [80], the results are shown in **Table 4.5**.

$$E_{SCF}^{(X)} = E_{SCF}^{(\infty)} + A\ exp\left(-\alpha\sqrt{X}\right). \tag{4.3}$$

$$E_{corr}^{(\infty)} = \frac{X^{\beta}E_{corr}^{(X)} - Y^{\beta}E_{corr}^{(Y)}}{X^{\beta} - Y^{\beta}}. \tag{4.4}$$

In eqs. (4.3) and (4.4), X and Y determine the cardinal number of the basis set, in this case $X=2$ and $Y=3$. The parameters α and β are fitted for each family of basis sets and each pair of cardinal numbers and were taken from ref. 80.

Table 4.5 Difference in electronic energies at the DLPNO-CCSD(T) level between the major (**TS1-1**) and minor (**TS1-1', TS1-2'**) transition states. The (2/3)-extrapolation was carried out using the def2-SVP and def2-TZVP basis sets. All energies are in kcal mol^{-1}. This table was taken from ref. 248.

	TS1-1'-TS1-1	**TS1-2'- TS1-1**
ORCA 4.2.1 def2-TZVP	6.880	8.743
ORCA 5.0.3 def2-TZVP	7.086	8.433
ORCA 5.0.3 (2/3)-extrapolated	6.228	8.299

Eq. (4.3) contains two unknown variables, so two extrapolation points are needed. $E_{SCF}^{(\infty)}$ and $E_{corr}^{(\infty)}$ are the extrapolated SCF and correlation energies, respectively. **Table 4.5** shows that ORCA 5.0.3 and ORCA 4.2.1 provide comparable relative electronic energies. The differences might be tracked down to different grid setups that were introduced with ORCA 5.0. Basis set extrapolation reduces the difference in electronic energies by around 0.8 kcal mol^{-1} for **TS1-1'** and 0.1 kcal mol^{-1} for **TS1-2'**.

In addition, CPS(6/7) extrapolation following eq. (2.29) was carried out to determine the influence of the $TCutPNO$ threshold on the difference in electronic energies. From **Table 4.6**, it becomes evident that PNO extrapolation only has marginal influence on the relative electronic energies and certainly only plays a very minor role in the comparison of computed and experimental stereoselectivities.

Table 4.6 Difference in correlation energies at the DLPNO-CCSD(T)/def2-SVP level of theory. The calculations were carried out using ORCA 5.0. All energies are in kcal mol^{-1}. This table was taken from ref. 248.

	TS1-1'-TS1-1	**TS1-2'- TS1-1**
$TcutPNO$=1e-6	-8.016	-7.120
$TCutPNO$=1e-7	-7.997	-7.263
PNO(6/7)-extrapolation	-7.987	-7.335

Another benchmark test was carried out for the thermochemical correction contribution that is obtained from the harmonic vibrational frequencies. These thermochemical corrections were originally obtained at the PBE-D3/def2-TZVP(-f) level of theory. For the benchmark, they were

recomputed using the r^{2}SCAN-3c [125] composite method by Grimme as implemented in ORCA 5.0.3.

Table 4.7 Difference in thermochemical corrections between most stable major (**TS1-1**) and different minor transition states (**TS1-1' – TS1-5'**) at the PBE-D3/def2-TZVP(-f) and r^{2}SCAN-3c levels of theory, respectively. All energies are given in kcal mol^{-1}. This table was taken from ref. 248.

	PBE-D3/def2-TZVP(-f)	**r^{2}SCAN-3c**
TS1-1' - TS1-1	1.26	1.32
TS1-2' - TS1-1	-0.03	0.04
TS1-3' - TS1-1	1.02	0.34
TS1-4' - TS1-1	-0.04	-0.38
TS1-5' - TS1-1	0.30	0.63

For the benchmark of the thermochemical corrections, again the larger benchmark set with five different minor transition states was used. It can be seen that for the most stable minor conformer **TS1-1'** the difference in thermochemical correction is the largest. If PBE-D3 and r^{2}SCAN-3c are compared, it can be seen that the largest deviation between these two methods occurs for **TS1-3'** (0.68 kcal mol^{-1}). In total, the influence of the density functional on the thermochemical correction is therefore relatively small. Taken together, it was shown that tightening the DLPNO thresholds and increasing the basis set shifts the computed enantiomeric ratios closer to the experimental ones. However, due to the challenging system size, combining quadruple-zeta basis sets and *TightPNO* settings is computationally very demanding and therefore was not carried out. In addition, also other details of the computational settings such as integration grid sizes or the details of the used solvation model may influence the computed stereoselectivity.

4.3.3 Stereo-controlling key factors

In this section, the stereo-controlling interactions are determined and discussed. For this, the catalyst geometry in its equilibrium state and in the transition states is discussed. The catalyst's TS and equilibrium structures are shown in **Figure 4.6** together with the corresponding steric maps. The aim of topographic steric maps is to provide a two-dimensional representation of the catalyst pocket [256,257]. The steric maps were created with the SambVca 2.1 web tool [258].

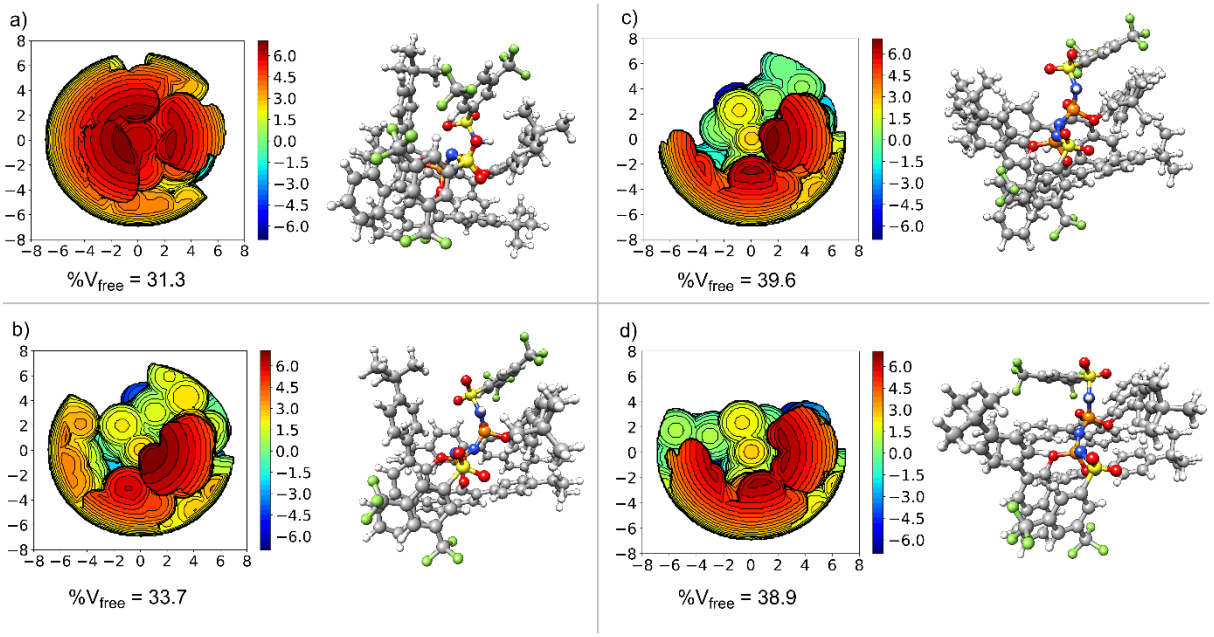

Figure 4.6 Steric maps and 3D structures of a) the catalyst in its equilibrium geometry, b) the catalyst in its major transition state geometry (**TS1-1**), c) the catalyst in the most stable minor transition state geometry (**TS1-1'**), d) the catalyst in the second most stable minor transition state geometry **(TS1-2')**. All steric maps are given in Å. This figure was taken from ref. 248.

The basic idea behind the steric maps is relatively simple. A sphere with a user-defined radius is placed at a user-defined central atom. In this case, the catalyst's nitrogen site that transfers the proton to the substrate was chosen as central atom. The sphere radius was set to 7 Å. The sphere is then divided into small voxels with a radius of 0.1 Å. In the last step, it is iterated over all voxels in the sphere. If an atom of the catalyst is within some covalent radius of the corresponding voxel, the voxel counts as "occupied", otherwise it counts as "free". The number of "free" voxels relative to the total number of voxels in the sphere yields the percentage of free volume %V_{free} within the sphere, which is used as a descriptor of the free volume within the catalyst pocket. The topography of the surrounding ligands is captured by measuring at which distance the closest atom lies. The steric maps shown in **Figure 4.6** are oriented in the same way as the corresponding 3D structures. **Figure 4.6a** shows the 3D structure of the catalyst in its equilibrium geometry. The active site (that is the catalytically active amine group) is shielded by the catalyst's aryl ligands. Thus, the catalyst's conformation can be described as "closed". For interaction of the N-H group with the arriving substrate, the catalyst has therefore to open up and the catalyst pocket into which the substrate can migrate has to be formed. This can be seen from the catalyst's geometry in **TS1-1** (**Figure 4.6b**), in which the N-H group is not shielded anymore. The free volume %V_{free} consequently is increased with respect to the

equilibrium geometry (31.1 % in **Figure 4.6a** vs. 33.7 % in **Figure 4.6b**). Interestingly, for the transition states leading to the minor enantiomeric product (**TS1-1'** and **TS1-2'**), $\%V_{free}$ increases significantly with respect to **TS1-1** (39.6 % for **TS1-1'**, 38.9 % for **TS1-2'**). This indicates that in the minor pathway the catalyst has to distort much stronger in order to allow for optimal interaction between the catalyst and the substrate. It seems to be the case that for the major pathway, a more or less optimal fit of the substrate in the catalyst pocket occurs. This is the reason why all low-energy major TS conformers feature very similar structural features. For the minor pathway, the catalyst has to distort to a larger extend, the confinement of the catalytic system is therefore decreased. This leads to a situation where different minor pathways emerge that all suffer from decreased confinement and therefore show almost degeneracy, but feature significantly different structural features. Due to the significant structural changes within the catalyst in the transition states with respect to the respective equilibrium geometries, it is expected that there is a large distortion energy which, at least partly, might be compensated by attractive noncovalent interactions. For exploring the interaction patters, the reaction barriers are decomposed using the DLPNO-CCSD(T) energies following eq. (4.2). The results are shown in **Table 4.8**.

Table 4.8 Decomposition of the reaction barriers $\Delta G^{\ddagger}$ at the DLPNO-CCSD(T) level of theory into different contributions for the major (**TS1-1**) and minor (**TS1-1', TS1-2'**) reaction pathways. For definition of the symbols, see eq. (4.2) and the affiliated text. All energies are in kcal mol^{-1}. a: contribution associated with the distortion of the catalyst. b: contribution associated with the distortion of the substrate. c: experimental difference is 2.8 kcal mol^{-1} [223]. This table has been taken from ref. 248.

	TS1-1	**TS1-1'**	**TS1-2'**	**TS1-1'-TS1-1**	**TS1-2'-TS1-1**
$\Delta G^{\ddagger}$	20.0	28.1	28.4	8.1[c]	8.4[c]
$\Delta G^{\ddagger}_{corr}$	20.0	21.2	19.6	1.2	-0.4
$\Delta E^{\ddagger}_{geo-prep}$	50.5 (41.3[a]+9.2[b])	56.6 (43.2[a]+13.4[b])	63.6 50.0[a]+13.6[b])	6.1 (1.9[a]+4.2[b])	13.1 (8.7[a]+4.4[b])
$E^{\ddagger}_{disp}$	-38.2	-43.2	-45.4	-5.0	-7.2
$\Delta E^{\ddagger}_{non-disp}$	-12.2	-6.5	-9.5	5.7	2.8

For **TS1-1**, it is found that the total barrier $\Delta G^{\ddagger}$ equals the correction contribution from thermochemistry and solvation $\Delta G^{\ddagger}_{corr}$. This indicates that the distortion energy and the

interaction energy between the catalyst and the substrate exactly have to balance each other. The geometric preparation of **TS1-1** amounts for 50.5 kcal mol^{-1}, of which the larger part (41.3 kcal mol^{-1}) is contributed by the catalyst. However, there is still a significant contribution to $\Delta E^{\ddagger}_{geo-prep}$ from the substrate. In summary, both the catalyst and the substrate have to change their conformation in order to allow for optimal interaction. The reaction is therefore best described by an induced-fit model that takes into account a certain flexibility of the catalyst to change its conformation accordingly. The total interaction between the catalyst and the substrate exactly cancels $\Delta E^{\ddagger}_{geo-prep}$. Interestingly, the larger part of the interaction stems from $E^{\ddagger}_{disp}$ (-38.2 kcal mol^{-1}), while non-dispersive contributions only amount for -12.2 kcal mol^{-1}. Therefore, it can be concluded that LD is the major stabilization interaction in the TS that counteracts the large geometric preparation and accordingly is responsible for the reaction to be observable in the first place. As can be seen from the corresponding dispersion interaction density plots (**Figure 4.5**), the strongest dispersive interactions are between the substrate and the bulky *t*-Bu and triflyl groups of the catalyst.

For the transition states leading to the minor enantiomeric product (**TS1-1', TS1-2'**), it is seen that the $\Delta G^{\ddagger}_{corr}$ correction term is only slightly different compared to **TS1-1**. Interestingly, the difference in geometric preparation $\Delta\Delta E^{\ddagger}_{geo-prep}$ is very high for both minor transition states (6.1 kcal mol^{-1} for **TS1-1'**, 13.1 kcal mol^{-1} for **TS1-2'**). This means that the difference in geometric preparation is the only significant contribution to the stereoselectivity. $\Delta\Delta E^{\ddagger}_{geo-prep}$ contains significant contributions from both the catalyst and the substrate. The substrate contribution can mostly be attributed to the partial sp³ character of the terminal carbon atom. The distortion process of the catalyst is more involved and is discussed in the next section. It is noted here that the increase in geometric preparation for the minor transition states is partly compensated by increased intermolecular interaction. Interestingly, for both minor transition states, the LD interactions become stronger compared to **TS1-1**. However, the nondispersive interactions become weaker (less negative). For **TS1-1'**, the increase in LD and decrease in nondispersive interactions cancel each other so that the overall interaction is approximately as strong as in **TS1-1**. For **TS1-2'**, the total interaction is larger (more negative) compared to **TS1-1**, which partly cancels the (even further) increased geometric preparation.

4.3.4 Catalyst distortion process

In the previous section, the geometric preparation was identified as the most important contribution to the computed stereoselectivity. In this section, the distortion process of the

catalyst is therefore studied in more detail. In its equilibrium state, the catalyst has a closed conformation and it has to open up and has to form a catalyst pocket in order to allow for optimal interaction with the substrate. In order to compute the influence of the individual ligands and substituents of the catalyst on the difference in geometric preparation $\Delta\Delta E^{\ddagger}_{geo-prep}$ between major and minor transition states, a deconstruction analysis was performed.

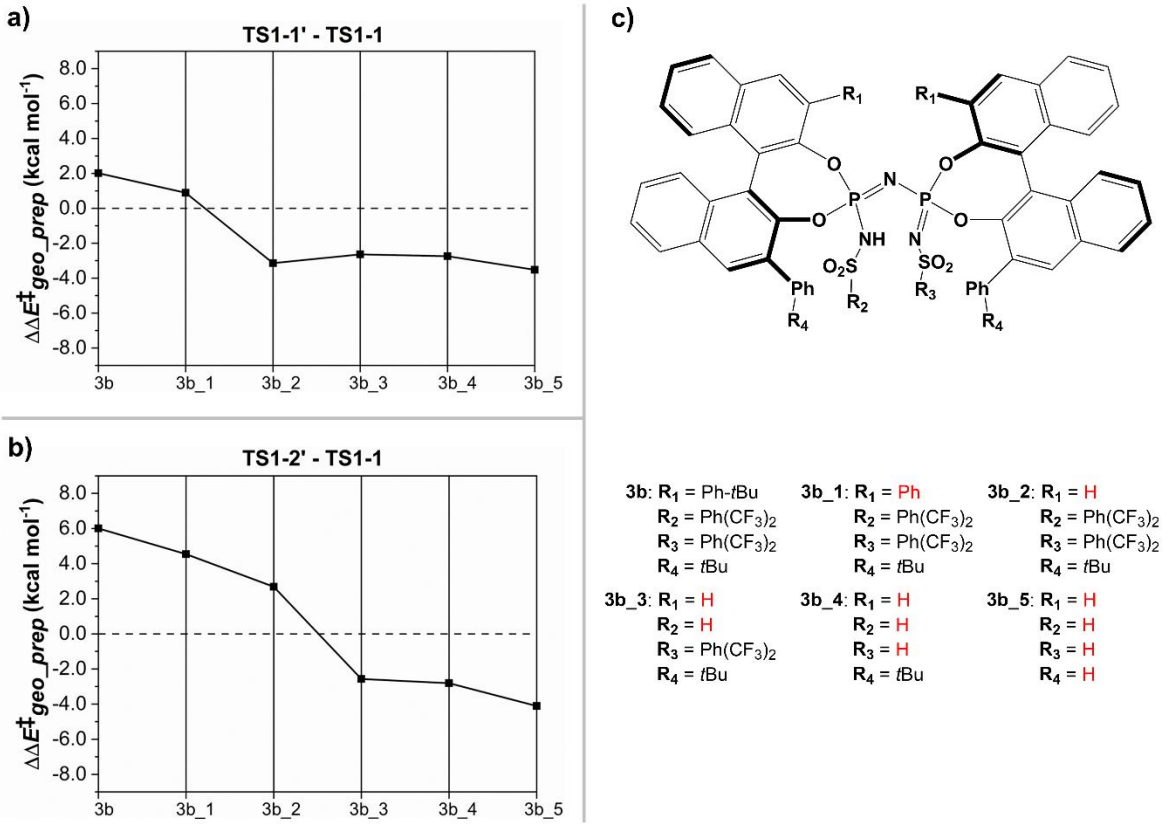

Figure 4.7 Deconstruction analysis of the catalyst. a) $\Delta\Delta E^{\ddagger}_{geo-prep}$ between **TS1-1'** and **TS1-1** for different subsystems. b) $\Delta\Delta E^{\ddagger}_{geo-prep}$ between **TS1-2'** and **TS1-1** for different subsystems. c) definition of the subsystems. The calculations were performed at the B3LYP-D3/def2-TZVP level of theory. For details, see text. This figure was taken from ref. 248.

In **Figure 4.7c**, a series of subsystems is defined. The system **3b** denotes the original catalyst. In each subsequent subsystem, substitution of one ligand by a hydrogen atom is carried out. After the substitution, the position of the newly added hydrogen atom was relaxed, while all other atoms of the system were kept frozen. In total, five different ligands were subsequently replaced by hydrogen atoms leading to the subsystems **3b_1** to **3b_5**. In **3b_5**, all aryl ligands as well as the ligands at the amine and imine groups are replaced by hydrogen atoms. This procedure was carried out for the catalyst in its equilibrium geometry as well as for the catalyst's structures in **TS1-1** (major pathway) and **TS1-1'** and **TS1-2'** (minor pathway). The calculations

were carried out at the B3LYP-D3/def2-TZVP level of theory. For each level of substitutions, the difference in geometric preparation $\Delta\Delta E^{\ddagger}_{geo-prep}$ between the corresponding subsystems of the major and minor pathway was computed to study the influence of the various ligands on the stereoselectivity. $\Delta E^{\ddagger}_{geo-prep}$ for each subsystem (with respect to the equilibrium structure) and $\Delta\Delta E^{\ddagger}_{geo-prep}$ were computed with and without D3 correction to study the influence of LD on the distortion process. It is important to note that the energy needed for the N-H bond elongation (remember that the transition states describe the proton transfer from the catalyst to the substrate) was excluded from the analysis. For this, for each transition state structure, one initial calculation was performed in which only the position of the N-H proton was optimized. The energy difference between the TS energy and that structure's energy defines the bond elongation energy. For each transition state (**TS1-1, TS1-1', TS1-2'**), the corresponding values were obtained with and without D3 correction and the values were subtracted from the total geometric preparation in order to obtain the distortion energy of the catalyst backbone solely. In **Table 4.9**, the comparison between **TS1-1'** and **TS1-1** is shown.

Table 4.9 Geometric preparation energy for the subsystems defined in **Figure 4.7** for **TS1-1** and **TS1-1'** at the B3LYP-D3/def2-TZVP level of theory. The numbers given in the table are corrected by the required energy for the N-H bond elongation to obtain the geometric preparation energy of the catalyst backbone. For details of the calculations, see text. All energies are in kcal mol^{-1}. This table was taken from ref. 248.

	$\Delta E^{\ddagger}_{geo-prep}$ (with D3BJ)			$\Delta E^{\ddagger}_{geo-prep}$ (without D3BJ)		
	TS1-1	**TS1-1'**	**TS1-1' – TS1-1**	**TS1-1**	**TS1-1'**	**TS1-1' – TS1-1**
3b	12.5	14.5	2.0	4.4	9.4	5.0
3b_1	13.3	14.2	0.9	4.7	11.0	6.3
3b_2	13.6	10.4	-3.2	4.8	9.0	4.2
3b_3	11.7	9.0	-2.7	3.4	7.0	3.6
3b_4	8.7	5.9	-2.8	2.8	5.7	2.9
3b_5	7.2	3.6	-3.6	2.3	5.2	2.9

It can be seen in **Table 4.9** that the geometric preparation energies of all subsystems of **TS1-1** and **TS1-1'** are much larger with D3 correction than without D3 correction. As detailed above, the catalyst in its equilibrium state is described by a closed conformation, so strong noncovalent interactions between the ligands are expected. During the opening process of the catalyst, these

NCIs are disrupted, which increases consequently the geometric preparation energy. For example, for **TS1-1** and the original system **3b** (*i.e.*, no substitutions so far), 65 % of the geometric preparation energy stems from LD, underpinning the importance of LD for the distortion process. For **TS1-1'**, still 35 % of the geometric preparation energy in subsystem **3b** stems from LD. The differences in geometric preparation energy $\Delta\Delta E^{\ddagger}_{geo-prep}$ between the subsystems of **TS1-1'** and **TS1-1** are discussed now. The values including the D3 correction are also graphically shown in **Figure 4.7a.** Interestingly, for the original subsystem **3b**, $\Delta\Delta E^{\ddagger}_{geo-prep}$ is smaller with D3 correction than without (2.0 kcal mol^{-1} with D3, 5.0 kcal mol^{-1} without D3), indicating that for the original catalyst, stronger disruptions of the nondispersive interactions happen in **TS1-1'**. For the subsystems, it can be seen that $\Delta\Delta E^{\ddagger}_{geo-prep}$ with D3 correction becomes smaller from step to step. Importantly, the substitution from **3b_1** to **3b_2** is most relevant in this regard, since it changes $\Delta\Delta E^{\ddagger}_{geo-prep}$ by -4.1 kcal mol^{-1} and even turns $\Delta\Delta E^{\ddagger}_{geo-prep}$ negative. Thus, starting from **3b_2**, **TS1-1'** has a smaller geometric preparation energy than **TS1-1**. At the same time, $\Delta\Delta E^{\ddagger}_{geo-prep}$ without D3 corrections also reduces over the series of subsystems. The changes are, however, much smaller and $\Delta\Delta E^{\ddagger}_{geo-prep}$ stays positive over the whole range of subsystems. Note that over the range of the first two substitutions (going from **3b** to **3b_2**), $\Delta\Delta E^{\ddagger}_{geo-prep}$ with D3 correction changes by -5.2 kcal mol^{-1}, while $\Delta\Delta E^{\ddagger}_{geo-prep}$ without D3 correction only changes by -0.8 kcal mol^{-1}. This means that there must be strong dispersive interactions between the ligands that are only present in the major transition state. If the interacting groups are then substituted by hydrogen atoms, these interactions do not longer exist and the minor transition state (which does not feature those interactions) becomes more stable on a relative scale. The negative $\Delta\Delta E^{\ddagger}_{geo-prep}$ values starting from **3b_2** indicate that there are LD interactions present at the core catalyst that actually do favor the minor transition state, and it are the interactions of the "outer" ligands like the aryl ligands that feature strong LD interactions only in the major TS. In conclusion, there is a single most important substitution (**3b_1** to **3b_2**) that changes $\Delta\Delta E^{\ddagger}_{geo-prep}$ significantly. The corresponding interactions that are only present in the major transition states have therefore major impact on the stereoselectivity. Since it was shown that mostly degenerate minor TSs arise that show significantly different structural features, the same analysis was carried out for **TS1-2'**. The results are shown in **Table 4.10**.

Table 4.10 Geometric preparation energy for the subsystems defined in **Figure 4.7** for **TS1-1** and **TS1-2'** at the B3LYP-D3/def2-TZVP level of theory. The numbers given in the table are corrected by the required energy for the N-H bond elongation to obtain the geometric preparation energy of the catalyst backbone. For details of the calculations, see text. All energies are in kcal mol^{-1}. This table was taken from ref. 248.

	$\Delta E^{\ddagger}_{geo-prep}$ (with D3BJ)			$\Delta E^{\ddagger}_{geo-prep}$ (without D3BJ)		
	TS1-1	**TS1-2'**	**TS1-2' – TS1-1**	**TS1-1**	**TS1-2'**	**TS1-2' – TS1-1**
3b	12.5	18.5	6.0	4.4	10.8	6.4
3b_1	13.3	17.9	4.6	4.7	13.2	8.5
3b_2	13.6	16.3	2.7	4.8	10.6	5.8
3b_3	11.7	9.1	-2.6	3.4	9.0	5.6
3b_4	8.7	5.9	-2.8	2.8	7.5	4.7
3b_5	7.2	3.1	-4.1	2.3	6.2	3.9

Astonishingly, an analogous situation like for **TS1-1'** arises for **TS1-2'**. $\Delta\Delta E^{\ddagger}_{geo-prep}$ between the two transition states decreases over the range of all subsystems. If D3 correction is included, the changes in $\Delta\Delta E^{\ddagger}_{geo-prep}$ are much more significant. Starting from the third substitution, $\Delta\Delta E^{\ddagger}_{geo-prep}$ becomes negative. For the substitution from **3b_2** to **3b_3**, $\Delta\Delta E^{\ddagger}_{geo-prep}$ with D3 changes by -5.3 kcal mol^{-1}, while $\Delta\Delta E^{\ddagger}_{geo-prep}$ without D3 correction changes only by -0.2 kcal mol^{-1}. The corresponding group in **TS1-1** therefore interacts almost entirely through LD. Over the range of all substitutions, $\Delta\Delta E^{\ddagger}_{geo-prep}$ with D3 corrections changes by -10.1 kcal mol^{-1} and only by -2.5 kcal mol^{-1} without D3 correction. Again, a single substitution can be identified that is most important for the decrease of $\Delta\Delta E^{\ddagger}_{geo-prep}$, in this case it is the substitution from **3b_2** to **3b_3**.

In summary, for each minor TS, one key interaction was determined that is only present in the corresponding major TS (**TS1-1**) and that determines the additional distortion energy that needs to be spent in the minor reaction pathways. For **TS1-1'**, it is the interaction of the phenyl ring bound to the BINOL backbone. For **TS1-2'**, it is the interaction of the aminyl-Ph-$(CF_3)_2$ group. For analyzing these interactions, NCI plots were obtained with the NCIPLOT program [259-261] at the B3LYP-D3/def2-TZVP level of theory. The plots are shown in **Figure 4.8**.

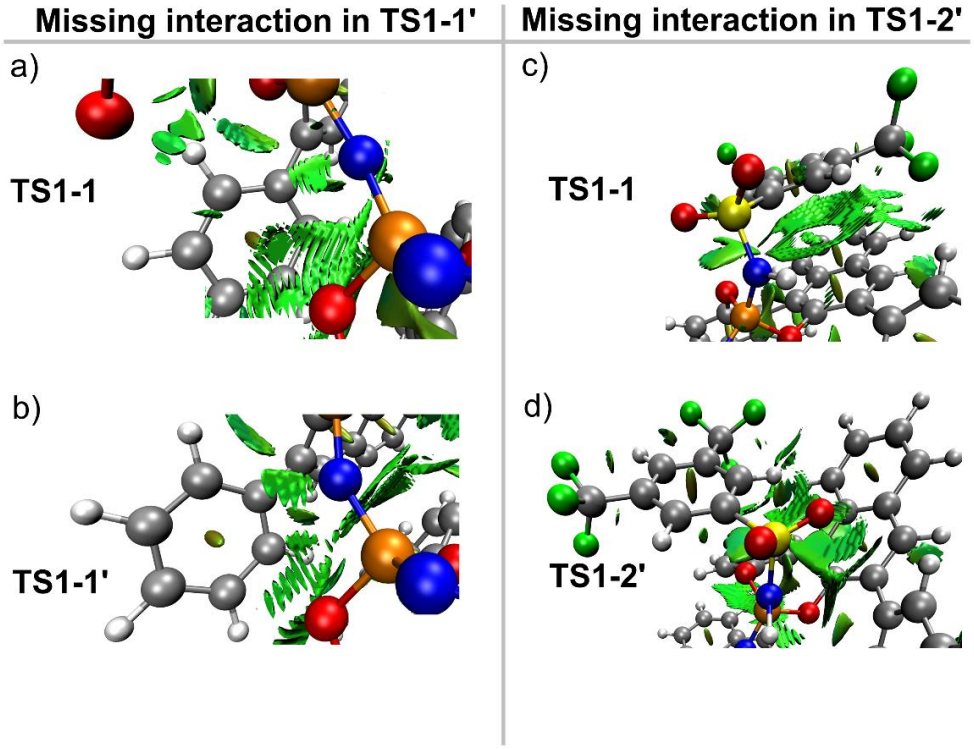

Figure 4.8 Left: NCI plots for the interaction of the Ph group and the core catalyst backbone in **TS1-1** (a) and **TS1-1'** (b). Right: NCI plots for the interaction of the Ph-$(CF_3)_2$ group and the catalyst backbone in **TS1-1** (c) and **TS1-2'** (d). The density isovalue for the NCI plots is 0.5 atomic units. This figure was taken from ref. 248.

As can be seen, the Ph ring in **TS1-1** features noncovalent interactions with the core catalyst, especially with the oxygen site. This interaction is, due to the increased distance of the Ph ring to the core catalyst, much weaker in **TS1-1'**. In addition, **TS1-1** features strong dispersive π-π interactions of the aminyl-Ph-$(CF_3)_2$ group with the BINOL backbone. Due to the rotation of the ligand, these interactions are basically absent in **TS1-2'**. As discussed before, removing these substituents during the deconstruction analysis stabilizes the minor transition states on a relative scale. It can be concluded that for the minor reaction pathways, additional distortion of the catalyst has to take place in order to allow for optimal interaction with the substrate. These distortions interrupt the intra-catalyst dispersive interactions and therefore increase the geometric preparation energy. The increase of geometric preparation energy cannot be fully compensated by additional intermolecular interactions, which explains the stereoselectivity of the studied hydroalkoxylation reaction. Thus, the proposed model might help in rational catalyst design. For example, by introducing bulky ligands that act as strong dispersion energy donors, the intra-catalyst dispersive interactions might be increased, which could lead to increased selectivity as well.

For testing the model, the equilibrium structures of closely related IDPi catalysts for which experimental stereoselectivities for the intramolecular hydroalkoxylation of 4-methylenedodecan-1-ol are known, were computed. The results are shown in **Figure 4.9**.

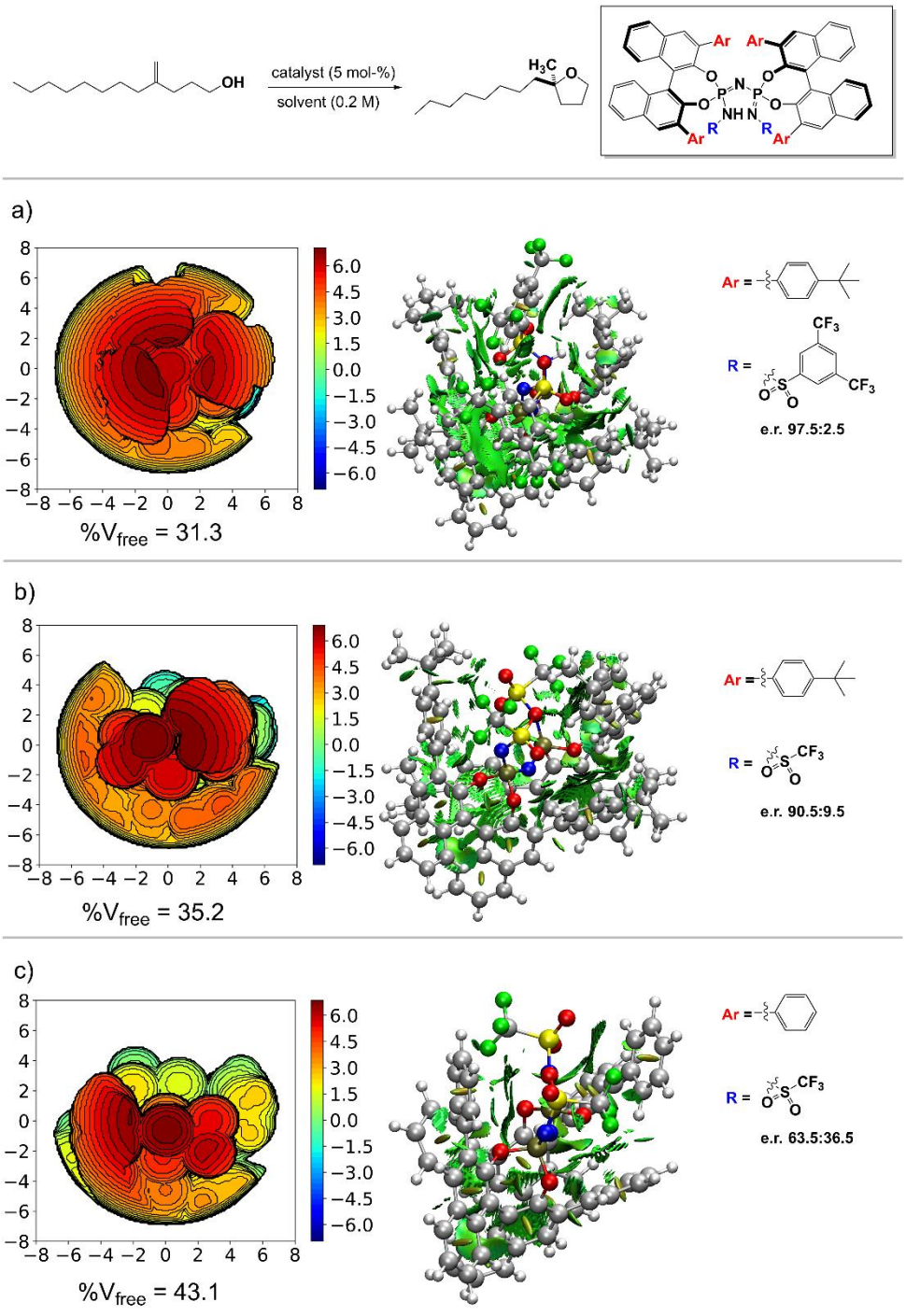

Figure 4.9 Steric maps and NCI plots together with the experimental enantiomeric rations for three different catalytic systems in their equilibrium geometries. a) catalyst **3b**. b) catalyst **3a**. c) catalyst **3d**. The isovalue for the density in the NCI plots is 0.5 atomic units. The experimental enantiomeric ratios were taken from ref. 223. This figure was taken from ref. 248.

The results in **Figure 4.9** show that the selectivity of the catalyst increases with the catalyst's confinement as quantified by $\%V_{free}$. In addition, it can be seen from the NCI plots that also the noncovalent intra-catalyst interactions increase along. Thus, the proposed induced-fit model for this reaction and the studied systems, in which catalyst distortion is governed by intra-catalyst LD forces, is supported. In conclusion, it was found that the stereoselectivity of the transformation is determined by the geometric preparation energy of the catalyst and the substrate. The catalyst structure changes drastically upon interaction with the substrate, which is best described by an induced-fit model. Interestingly, LD plays a two-fold role in this reaction, as is rationalized in **Figure 4.10**. First, the intermolecular London interaction between the catalyst and the substrate is the major stabilization interaction in the transition state that counteracts the large geometric preparation energy and is therefore responsible for the reaction to occur at the given temperature at all. Second, the intra-catalyst LD interactions play a crucial role for the catalyst distortion process. For the minor transition states, the LD interactions are much more disrupted in comparison to the major pathway, leading to an increased catalyst distortion energy. Introducing novel bulky dispersion energy donors to the system is a promising way to increase the stereoselectivity of catalysts in these types of transformations.

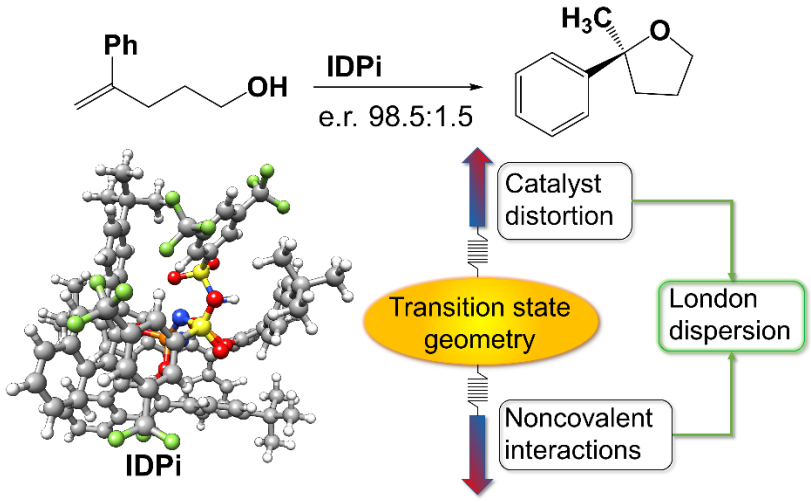

Figure 4.10 Schematic illustration of the importance of London dispersion for the transition state geometry and stereoselectivity. This figure was taken from ref. 248.

Chapter 5 Brønsted Acid Dimerization in Enantioselective Organocatalytic Reactions

5.1 Introduction

Elements of this chapter including some of the figures are part of submitted manuscripts [262] or published journal articles [263]:

Harden, I.; Neese, F.; Bistoni, G. Dimerization of Confined Brønsted Acids in Enantioselective Organocatalytic Reactions. *Manuscript submitted for publication* **2023**.

Das, S.; Mitschke, B.; De, C. K.; Harden, I.; Bistoni, G.; List, B. Harnessing the ambiphilicity of silyl nitronates in a catalytic asymmetric approach to aliphatic β^3-amino acids *Nature Catalysis* **2021**, *4* (12), 1043-1049. https://doi.org/10.1038/s41929-021-00714-x.

All figures shown in this chapter have been designed by the thesis author. For each figure, the corresponding figure caption describes whether the figure is part of a publication.

In Chapter 4, an induced-fit model for the case study of IDPi catalyzed intramolecular hydroalkoxylations was presented and critically discussed. From both experimental and computational side, a monomeric pathway was assumed [223,248]. This means that only those pathways in which a single catalyst molecule is involved in every occurring reaction step were considered. This is not only true for the presented case study, but for practically all combined experimental and computational studies that involve IDPi catalysis (see refs. 221,222,226,234,264 for examples). Thus, possible cooperative effects between two or more IDPi catalyst molecules are largely overlooked in literature. It is well known that smaller and therefore less bulky phosphoric acids can form dimers, trimers and other aggregates in solution, as was discussed for dimethyl phosphoric acid [265] or diphenyl phosphoric acid [266]. In addition, heterodimerization of CPAs with acetic acid and its influence on the kinetics and stereoselectivity of epoxide ring opening reactions was studied by the groups of List and Thiel [267]. Besides the investigation of relatively small systems, the group of Niemeyer has worked experimentally and computationally on cooperative effects of interlocked and non-interlocked CPA molecules in the two-step transfer-hydrogenation of 2-phenylquinoline in a series of papers [244-247]. For the non-interlocked CPA system, they found that for the two-step reduction of 2-phenylquinoline, monomeric and dimeric pathways are in competition, which depends on the catalyst concentration. While the reaction rate for the first reduction step is higher for the monomeric pathway, it is significantly lower for the second reaction step. Thus,

even if the dimeric pathway occurs much more rarely then the monomeric pathway, it can, in principle, still dominate the reaction by being significantly faster. In addition, the dimeric pathway yields much higher stereoselectivities than the monomeric pathway. Their work indicates that cooperative effects can be important for relatively bulky organocatalysts. Moreover, the group of Gschwind could identify CPA-imine ion pair dimers in which the two imine molecules are located between the CPA molecules [268,269]. This exemplifies that not only acidic species but ion pairs can dimerize as well. However, it is still unknown whether such effects can also occur for substantially larger IDPi systems, and if yes, under which conditions.

The aim of this chapter is to shed light on catalyst dimerization processes in organocatalysis. For achieving this, again the protocol based on conformer sampling, geometry optimizations and high-level single point energies is applied. For a more systematic approach, the dimerization of five well-known and highly relevant Brønsted acid catalysts in the absence of any substrates or co-catalysts is studied. The five used catalyst classes are CPA, DTPA, DSI, IDP and IDPi, which have already been discussed in section 3.3. The actual Lewis structures used for the dimerization study are shown in **Figure 5.1**.

In the second step, the focus is set on a recently published case study [263] for which the conditions of the possible occurrence of cooperative effects are elucidated. The case study reaction is shown in **Figure 5.2**.

5.2 Computational protocol

The computational protocol in principle follows the protocol outlined in section 4.2. However, some important changes are discussed here. For the systematic study on Brønsted acid dimerization, conformer sampling was carried out for each acid monomer and dimer. Multiple initial dimeric structures were used to cover larger parts of the conformational space. For **CPA, DTPA** and **DSI**, a single constraint was used during the conformer sampling in order to keep the hydrogen bond network intact. For **IDP** and **IDPi**, sampling runs with and without constraints were carried out. Again, the CREST program was used, however, due to the challenging system size (the largest dimer contains 400 atoms), the GFN-FF force field method [156] by Grimme was used instead of GFN-xTB.

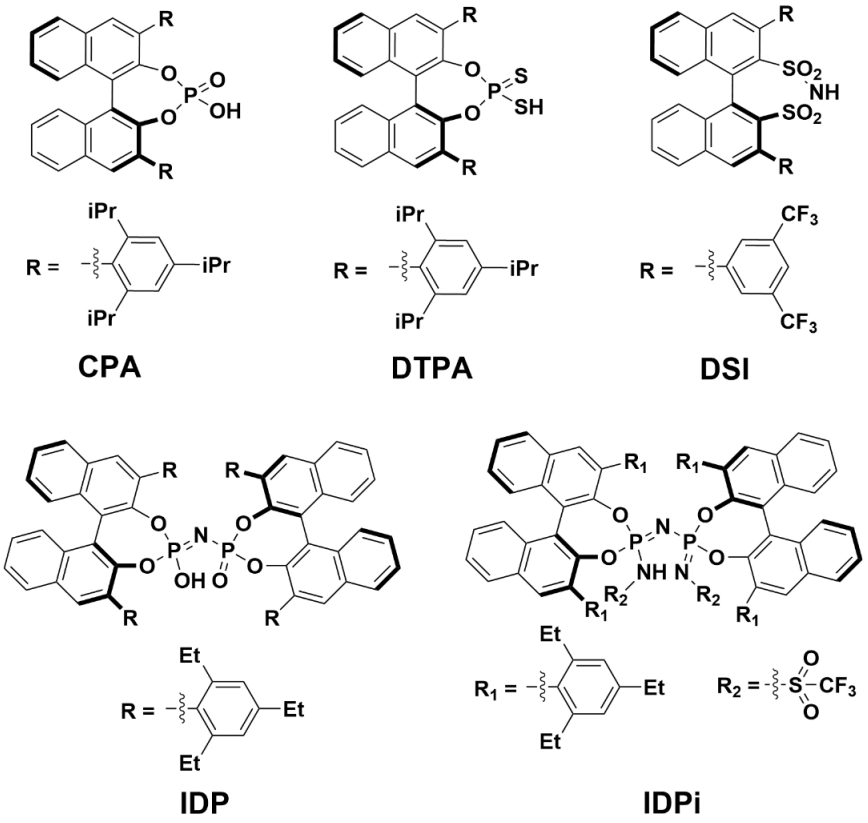

Figure 5.1 Lewis structures of Brønsted acids considered in this study. In the further, the class of the corresponding acid is used as identifier for the particular molecule. This figure was taken from ref. 262.

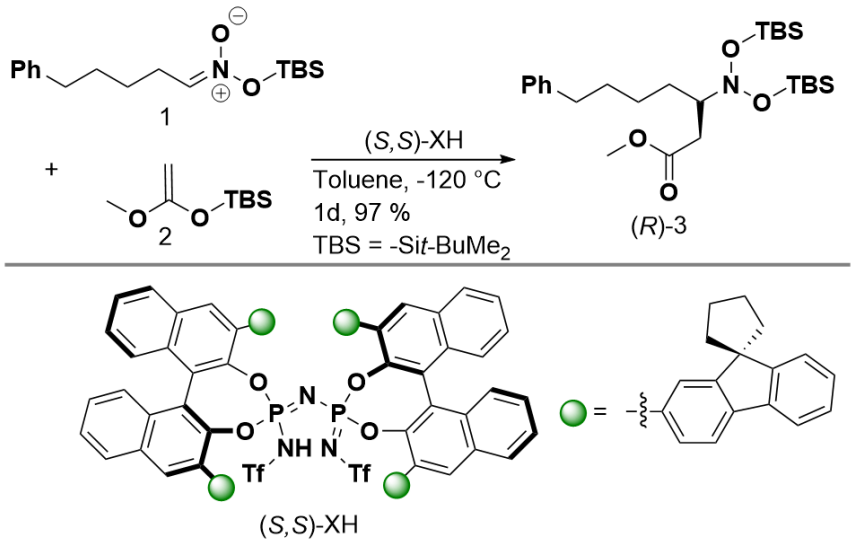

Figure 5.2 Reaction considered in this study. Details can be found in ref. 263.

The obtained conformers were optimized at the PBE-D3/def2-SVP level of theory. For the smaller systems (**CPA, DTPA** and **DSI**), all obtained conformers were optimized, while for **IDP** and **IDPi** optimization of all dimer conformers was computationally too demanding.

Instead, all dimer conformers within an energy window of 2 kcal mol^{-1} with respect to the most stable conformer were optimized, resulting in 61 (**IDP**) and 35 (**IDPi**) fully optimized unique dimer conformers. Thermochemical corrections were obtained at the same level of theory. Final single point energies were obtained with the ωB97M-V [147] functional and def2-TZVP basis set. The noncovalent interaction energy was computed in a post-SCF fashion using the VV10 functional. The RIJCOSX approximation were used for the exchange integrals. For benchmarking purposes, calculations at the B3LYP-D3 [112,113], M06-2x [270] and HFLD [75] levels of theory were carried out. For the HFLD calculations, the def2-TZVP(-f) basis set was used together with *NormalPNO* settings and a tightened *TCutPairs* value of 10^{-5}. Solvation effects were captured with the SMD model at the B3LYP-D3/def2-TZVP level of theory. Toluene was used as solvent. The free dimerization energy ΔG_{dim} was computed using a supermolecular approach [90].

$$\Delta G_{dim} = G_{dim} - 2G_{mon}. \tag{5.1}$$

In eq. (5.1), G_{dim} is the free energy of the dimer and G_{mon} is the free energy of the monomer. Regarding the computations on the C-C coupling reaction case study (**Figure 5.2**), a slightly different protocol was used. Optimization of the monomeric ion pair **I-1** and transition state **TS-1** (see **Figure 5.8** for details on the Lewis structures) was done for computations partly published in ref. 263. Conformer sampling was performed at the GFN2-xTB level of theory. The lowest-in energy conformers were optimized at the PBE-D3/def2-SVP level of theory with ORCA 4.2.1. With the release of ORCA 5, the most stable conformers of all monomeric structures were re-optimized using ORCA 5.0 and the r^{2}SCAN-3c composite method by Grimme. The dimeric ion pair **I-2** consists of **I-1** and an additional silylated catalyst molecule **X-TBS**. For the conformer sampling, the cartesian coordinates of **I-1** were constrained in order to keep the overall structure of the ion pair intact. The same was done for the conformer sampling of **TS-2** (*i.e.*, the coordinates of **TS-1** were constrained). For the subsequent geometry optimizations, an ONIOM [271] multiscale approach was used. The catalyst molecules were optimized at the GFN2-xTB level of theory, while the nitronate and the substrate were optimized at the r^{2}SCAN-3c level of theory. Solvation effects were captured with the ALPB model [272] (toluene). Of course, the corresponding monomeric structures were optimized at the same level of theory starting from the r^{2}SCAN-3c optimized geometries. The reactant and product states affiliated with the transition states were obtained from intrinsic reaction coordinate [273] (IRC) calculations with subsequent geometry optimizations. Thermochemical corrections were obtained from numerical frequency calculations. Final single point energies

were obtained at various levels of theory (see **Figure 5.9**). The SMD solvation model with CH_2Cl_2 as solvent was used. In addition, the basis set superposition error (BSSE) was treated approximately with Grimme's geometrical counterpoise correction (gCP) [126].

5.3 Results and discussion

5.3.1 Dimerization tendency of acids

First, the dimerization process of the five Brønsted acids depictured in **Figure 5.1** is studied. For this, the free dimerization energy was computed with respect to the temperature. The temperature dependence was captured by re-computing the thermochemical corrections at different temperatures. The electronic energies and the correction from implicit solvation were taken to be temperature-independent. The free dimerization energies are shown in **Figure 5.3**.

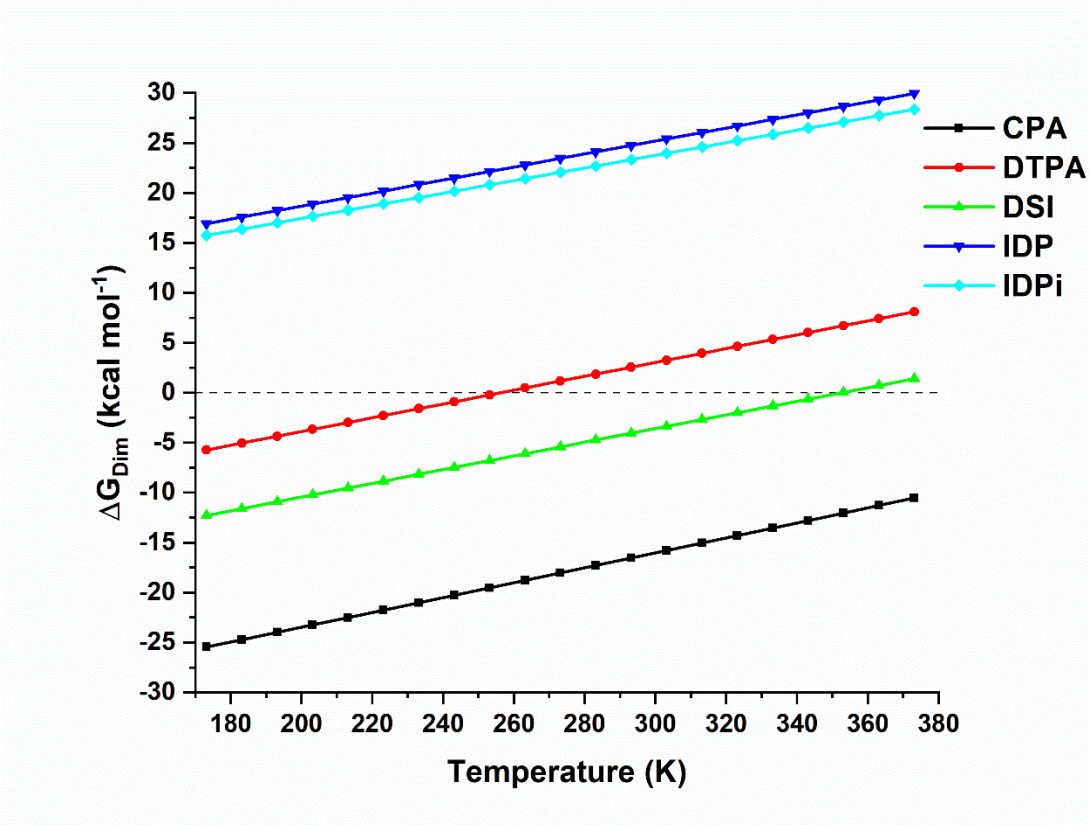

Figure 5.3 Free dimerization energies ΔG_{dim} for **CPA, DTPA, DSI, IDP** and **IDPi** between 170 K and 370 K. This figure was taken from ref. 262.

The free dimerization energies in **Figure 5.3** allow for classification of the five acids into three classes. First, there is **CPA** that shows very negative values for ΔG_{dim} over the whole temperature range, indicating its high dimerization tendency in solution. Next, there are **DTPA** and **DSI**, which show much larger (*i.e.*, closer to zero) free dimerization energies. Both systems

have a crossing temperature for which $\Delta G_{dim} = 0$. For **DTPA** it is interpolated to 256 K, for **DSI** it is interpolated to 354 K. Thus, at higher temperatures dimerization is disfavored for both systems, while under mild conditions dimerization could happen. It is noted that the crossing temperature is sensitive to the free dimerization energy – a change of ΔG_{dim} by 2 kcal mol^{-1} changes the crossing temperature by approx. $\pm$ 30 K. The substantially bulkier and more confined systems **IDP** and **IDPi** show very positive dimerization energies for all temperatures, indicating their weak tendency to dimerize. Dimerization of these systems in solution should therefore not be observable under experimental conditions.

Besides the temperature, there are other factors that can heavily influence the free dimerization energy, such as the polarity of the solvent. For real-world catalytic reactions, additional influences such as reaction rates can play a role on the actual reaction outcome.

5.3.2 Decomposition of free dimerization energies

For achieving a better understanding of the various energy contributions, ΔG_{dim} is decomposed using eq. (5.2).

$$\Delta G_{dim} = \Delta E_{geo-prep} + E_{disp} + \Delta E_{non-disp} + \Delta G_{corr}. \quad (5.2)$$

$\Delta E_{geo-prep}$ is again the geometric preparation energy, E_{disp} is the inter-fragment dispersion energy, $\Delta E_{non-disp}$ is the non-dispersive interaction energy contribution and ΔG_{corr} is a correction term that includes corrections from thermochemistry (governed by entropy) and implicit solvation. It is important to note that the dispersive contribution E_{disp} was computed using a DFT approach, while in Chapter 4 its computation was based on DLPNO-CCSD(T). In DLPNO-CCSD, dispersive interactions are well defined by means of specific double excitations (see **Figure 2.2**), while in the DFT approach used here, E_{disp} is calculated from the corresponding noncovalent interaction energies of the dimer and the monomers obtained with the non-local VV10 density functional. However, the influence of various schemes for incorporating dispersive interactions are tested below.

In **Figure 5.4**, the decomposition of ΔG_{dim} for each most stable dimer conformer is shown. The first term to discuss here is ΔG_{corr} ($\Delta G_{corr} = \Delta G_{thermo} + \Delta G_{solv}$) that is roughly constant along the series of systems. It contains two contributions, namely thermochemical (which itself contain multiple terms) and solvation corrections. The thermochemical corrections are governed by entropy, which disfavors dimerization and is temperature dependent. The

almost linear increase of ΔG_{dim} with respect to the temperature *T* (see **Figure 5.3**) is caused by the term $-T\Delta S_{dim}$, with ΔS_{dim} being the dimerization entropy.

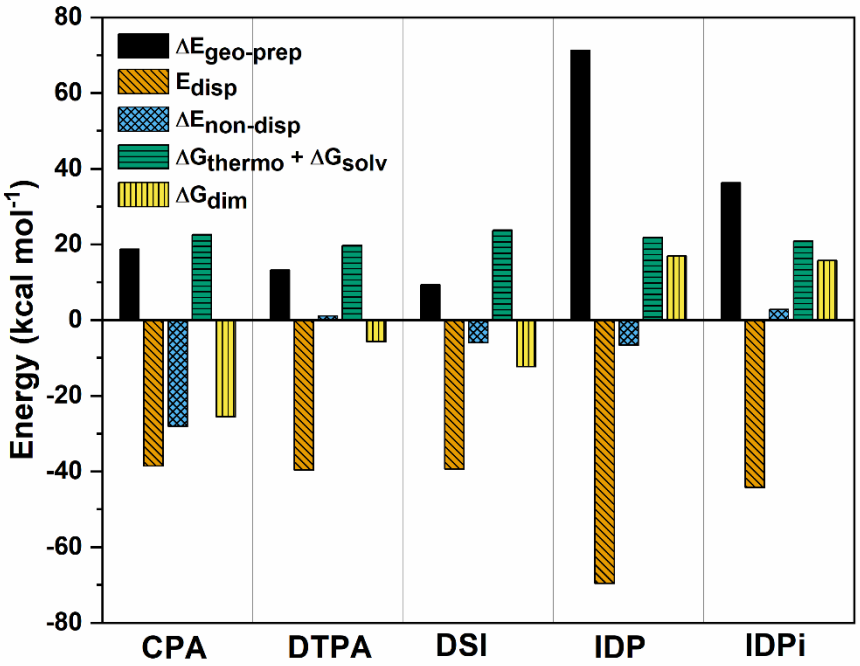

Figure 5.4 Decomposition of ΔG_{dim} into various contributions for **CPA, DTPA, DSI, IDP** and **IDPi** at the ωB97M-V/def2-TZVP level of theory at 173 K. For details see text. All energies are in kcal mol^{-1}. This figure was taken from ref. 262.

The second contribution to ΔG_{corr} are the solute-solvent interactions that are obtained with the implicit SMD solvation model. They consist of electrostatic and non-electrostatic (such as dispersive) interactions. The alternative C-PCM solvation model only captures electrostatic interactions, which stabilize the monomers. Because of electrostatic interactions, the polarity of the solvent influences the free dimerization energy significantly, with unpolar solvents favor dimerization more than polar solvents. Moreover, dispersive solute-solvent interactions favor dissociation because the solvent accessible surface (SAS) of the dimer is decreased compared to the SAS of two non-interacting monomer molecules. For testing the importance of dispersive solute-solvent interactions, the dimerization energies ΔE_{dim} were computed with C-PCM and SMD, respectively. The results are shown in **Table 5.1**.

As can be seen from **Table 5.1**, dispersive solute-solvent interactions have a large influence on the electronic dimerization energies, their contributions range from 4.5 kcal mol^{-1} (**CPA**) to 7.6 kcal mol^{-1} (**IDPi**). As expected, the dispersive solute-solvent interactions correlate with the system size and therefore with the size of the SAS.

Table 5.1 Dimerization energies ΔE_{dim} computed at the B3LYP-D3/def2-TZVP with the C-PCM and SMD solvation models, respectively. Toluene was used as solvent. All energies are in kcal mol^{-1}. This table was adapted from ref. 262.

	ΔE_{dim} (C-PCM)	ΔE_{dim} (SMD)	$\Delta\Delta E_{dim}$ (SMD – C-PCM)
CPA	-42.2	-37.7	4.5
DTPA	-24.3	-18.7	5.6
DSI	-26.5	-22.2	4.3
IDP	-0.8	6.2	7.0
IDPi	-3.0	4.6	7.6

In **Table 5.2**, the total contributions from thermochemistry ΔG_{thermo} and solvation ΔG_{solv} to ΔG_{corr} are shown. At 173 K, ΔG_{thermo} is slightly larger than ΔG_{solv} but both contribute significantly to ΔG_{corr}. For smaller temperatures, ΔG_{solv} will be the dominating contribution that even close to the zero temperature will prevent dimerization of **IDP** and **IDPi**. At high temperatures, the penalty due to entropy will dominate the solvation contribution. Note that ΔG_{solv} is taken to be temperature-independent.

Table 5.2 Contributions from thermochemical corrections ΔG_{thermo} and solvation corrections ΔG_{solv} to ΔG_{corr} at 173 K. ΔG_{thermo} was obtained at the PBE-D3/def2-SVP level of theory, ΔG_{solv} was obtained at the B3LYP-D3/def2-TZVP/SMD(toluene) level of theory. All energies are in kcal mol^{-1}. This table was adapted from ref. 262.

	ΔG_{corr}	ΔG_{solv}	ΔG_{thermo}
CPA	22.6	9.0	13.6
DTPA	19.7	8.0	11.7
DSI	23.7	10.7	13.0
IDP	21.9	10.3	11.6
IDPi	21.0	9.7	11.3

The next step is to discuss the remaining terms in eq. (5.2). They are discussed simultaneously, as they are closely connected to each other. For all systems, geometric preparation energies are

positive, since distortion of the monomer's geometries practically always happens if dimers are to be formed. For the smaller acidic systems (**CPA, DTPA, DSI**), $\Delta E_{geo-prep}$ is smaller than ΔG_{corr}, while for the confined acids (**IDP, IDPi**), $\Delta E_{geo-prep}$ is significantly larger than ΔG_{corr}. Furthermore, for all systems, LD is the most important stabilizing interaction. For all systems except **CPA**, non-dispersive interactions compensate such that their contribution to ΔG_{dim} is of very minor relevance, making LD the only stabilizing interaction in these dimers. Because of the general importance of LD for the dimerization process, the dispersive interaction density plots obtained at the HFLD/def2-TZVP(-f) level of theory are shown in **Figure 5.5**.

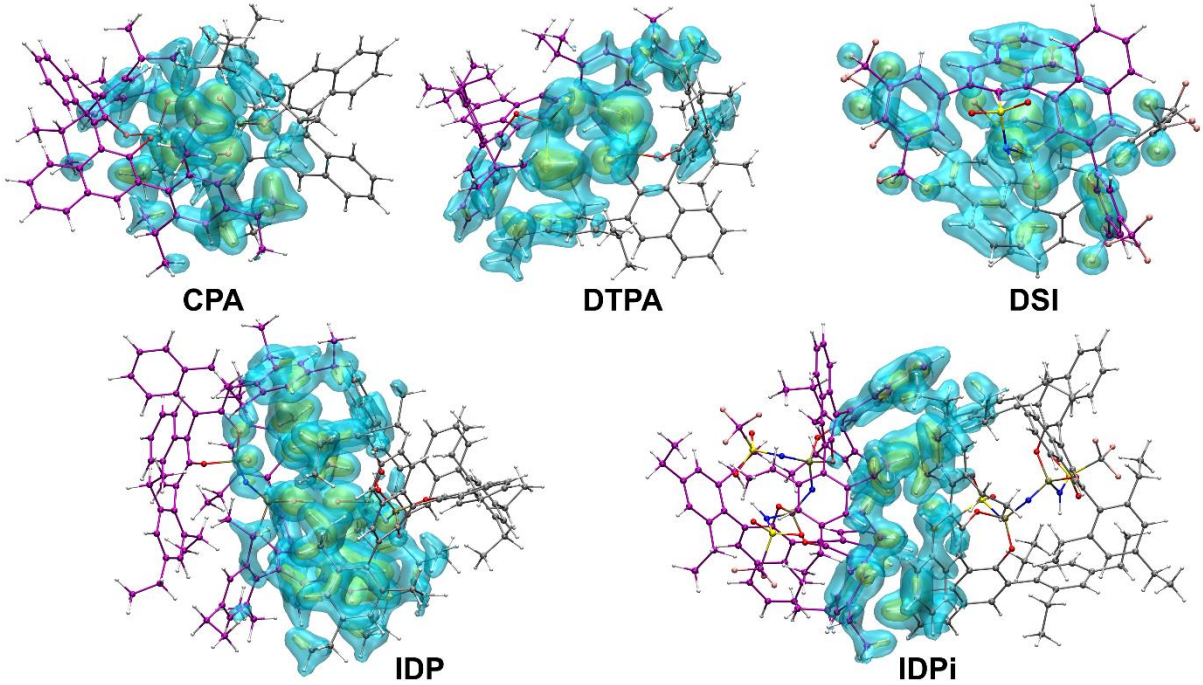

Figure 5.5 3D structures and dispersion interaction density (DID) plots obtained at HFLD/def2-TZVP(-f) level of theory for **CPA, DTPA, DSI, IDP, IDPi**. Two isosurfaces with different density values are shown in order to emphasize the key interactions. Density cutoffs: cyan surface 1.6e-05 a.u., yellow surface 7e-05 a.u. This figure was taken from ref. 262.

For **CPA** strong dispersive as well as non-dispersive interactions arise. From the corresponding 3D structure in **Figure 5.5**, it becomes evident that the most stable **CPA** dimer features two intermolecular hydrogen bonds. The strong oxygen-hydrogen bonds lead to stabilizing non-dispersive interactions. However, as can be seen from the corresponding DID plot, also dispersive interactions arise between the phosphoric acid groups, underlining the multitude of interactions that can contribute to the formation of hydrogen bonds [274]. For **DTPA** the contribution of $\Delta E_{non-disp}$ to ΔG_{dim} is almost zero. Instead of oxygen sites within the phosphoric acid group, **DTPA** contains sulfur sites. The most stable **DTPA** dimer features two

intermolecular hydrogen bonds, just like **CPA**. However, the much weaker sulfur-hydrogen bonds consequently do not lead to significant non-dispersive stabilization, which leads to larger (closer to zero) free dimerization energies. Importantly, for **DTPA** also those conformers that does not feature any hydrogen bonds at all were optimized. Their very high relative energies indicate that despite the relatively weak sulfur-hydrogen bonds, the dimers with two hydrogen bonds are preferred. **DSI** shows a single intermolecular hydrogen bond. It is also the smallest of all systems, which is reflected by the smallest $\Delta E_{geo-prep}$ of all systems. For the confined acids **IDP** and **IDPi**, significant increase in $\Delta E_{geo-prep}$ leads to very positive free dimerization energies. These systems should not dimerize under experimental conditions. Interestingly, **IDP** still features an intermolecular hydrogen bond, hence the active centers still interact directly with each other. This leads to very strong dispersive interactions that are, however, compensated by geometric preparation. For **IDPi** on the other hand, no interaction between the active center occurs, as verified by the corresponding DID plot. The molecules therefore entirely interact *via* the "outer" bulk ligands. This leads to a smaller $\Delta E_{geo-prep}$ compared to **IDP**, but also to less stabilizing LD interactions. For **IDPi**, also alternative conformations that feature at least one hydrogen bond were investigated, these were, however, always higher in relative energy. The tendency of dimerization of the different acids is illustrated in **Figure 5.6**.

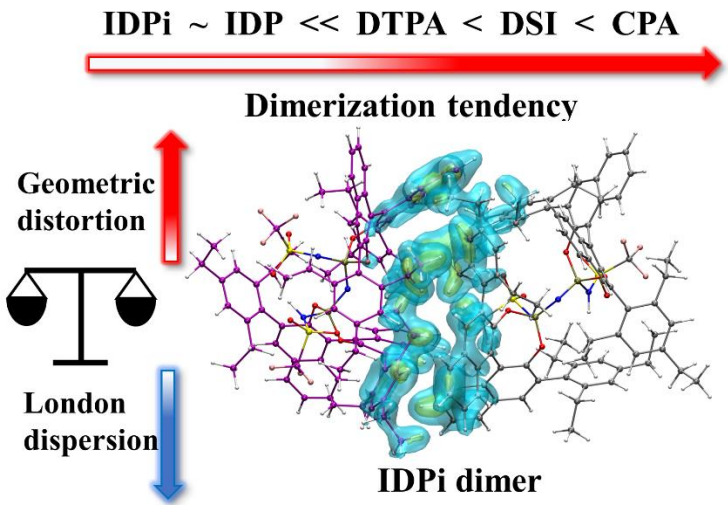

Figure 5.6 Illustration of the dimerization tendencies of **CPA, DTPA, DSI, IDP** and **IDPi**. This figure was taken from ref. 262.

For the smaller systems (**CPA, DTPA, DSI**), ΔG_{dim} mainly depends on the number and the strength of intermolecular hydrogen bonds. For the confined **IDP** and **IDPi** acid dimers, ΔG_{dim} is determined by the interplay between stabilizing LD and destabilizing distortion energy. Of

course, for all systems, a roughly constant penalty for entropy (depends on the temperature) and solvation (depends on the polarity of the solvent) has to be paid.

This section is closed with a comparison of the LD contributions obtained with various methods. As discussed before, the used ωB97M-V functional invokes the non-local VV10 functional for computation of the NCIs. It is used in a post SCF fashion, *i.e.*, the converged density obtained from the SCF cycles is used for computing the NCI energy once. The well-known DFT-D methods by Grimme (see section 2.4.4 for details and references) use semi-empirical approaches for computation of LD. The HFLD method, which originates from the DLPNO-CC family of methods, computes the LD contribution by solving approximate DLPNO-CCSD amplitudes equations in which all intra-fragment pairs are neglected [75], leading to a more systematic definition of LD. For the acid dimers, dispersive interactions were computed using various methods. The results are shown in **Figure 5.7**.

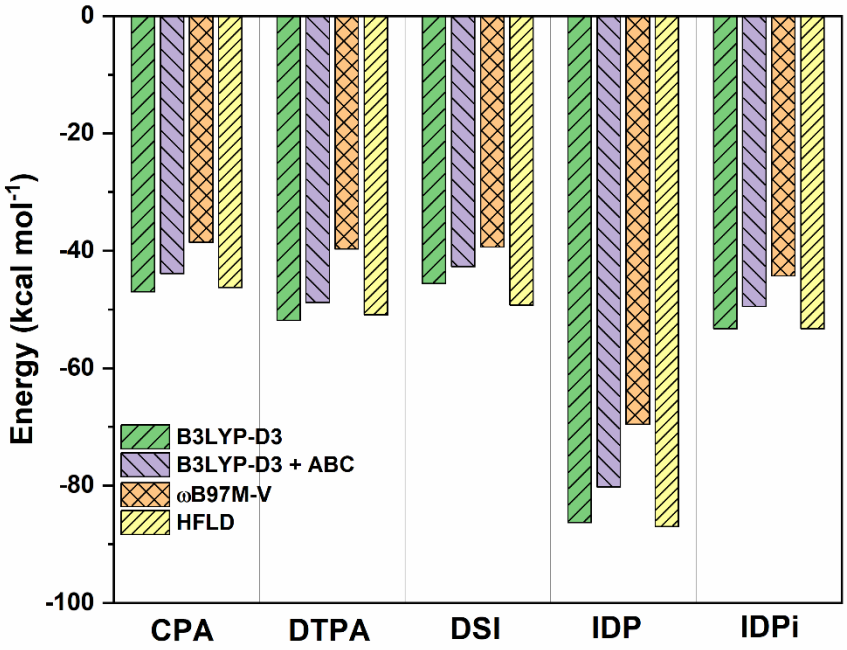

Figure 5.7 LD contributions to ΔG_{dim} obtained at different levels of theory for **CPA, DTPA, DSI, IDP, IDPi**: B3LYP-D3/def2-TZVP, B3LYP-D3 + ABC/def2-TZVP, ωB97M-V/def2-TZVP, HFLD/def2-TZVP(-f). All energies are in kcal mol^{-1}. This figure was taken from ref. 262.

From the results in **Figure 5.7**, few trends are observable. First, the three-body ABC correction [275,276] yields systematically smaller LD corrections than B3LYP-D3. Interestingly, the DFT-D obtained LD interactions match the ones obtained from HFLD very well. The VV10 functional yields slightly decreased dispersive interactions, which show, however, the very

same trends as DFT-D and HFLD. The very strong intermolecular dispersive interactions underline, regardless of the computational settings, that LD is the key stabilizing interaction in all acid dimers.

5.3.3 Towards an understanding of cooperative effects in catalytic reactions

In the previous section, it was established that relatively small Brønsted acids (**CPA, DTPA, DSI**) can form stable dimers in solution under mild conditions, while larger and more confined acids (**IDP** and **IDPi**) cannot dimerize under the given conditions. The previous section helped in understanding how two catalyst molecules interact with each other and whether these interactions are overall attractive or repulsive. In "real-world" catalytic systems, of course, the catalyst molecules do not only interact with each other. There is at least one substrate, maybe co-catalysts, solvent molecules, reaction intermediates, etc. In order to investigate cooperative effects, all these molecules besides the catalyst have to be taken into account as well, since the additional molecules can potentially influence aggregation processes heavily. In order to investigate this, the case study depicted in **Figure 5.2** is considered. The reaction describes the stereoselective C-C coupling reaction between silylated nitronates **1** and silyl ketene acetals **2** to give protected β^3-amino acids. The reaction was developed by the List group and the mechanism was partly elucidated with the help of NMR spectroscopy. In addition, computations by the thesis author were performed, as detailed in ref. 263. The monomeric pathway assumed in ref. 263 and the dimeric pathway considered in ref. 262 are depicted in **Figure 5.8**.

In the first step of the reaction, the catalytic Brønsted acid **XH** is activated by silylation. As silylating reagent, **2** is used in excess at room temperature. The silylation process leads to the formation of **X-TBS**, which is the catalytically active species that enters the catalytic cycle. The catalytic cycle "begins" with the key step of the reaction, *i.e.*, the formation of the chiral ion pair **I-1**. For this **X-TBS** transfers its TBS group to **1**. The nitronate **1** is nucleophilic at the central iminyl carbon atom. By silylation, the nitronate becomes cationic and therefore electrophilic, which is the central idea behind the reaction design [263]. The catalyst anion and the nitronate cation form a chiral ion pair, thus the reaction employs ACDC. The conformation of **I-1** determines the stereoselectivity of the reaction. In the last step, the nucleophile **2** attacks the ion pair and the corresponding product **P-1,** and after restoration of **X-TBS** the final product **3**, is formed.

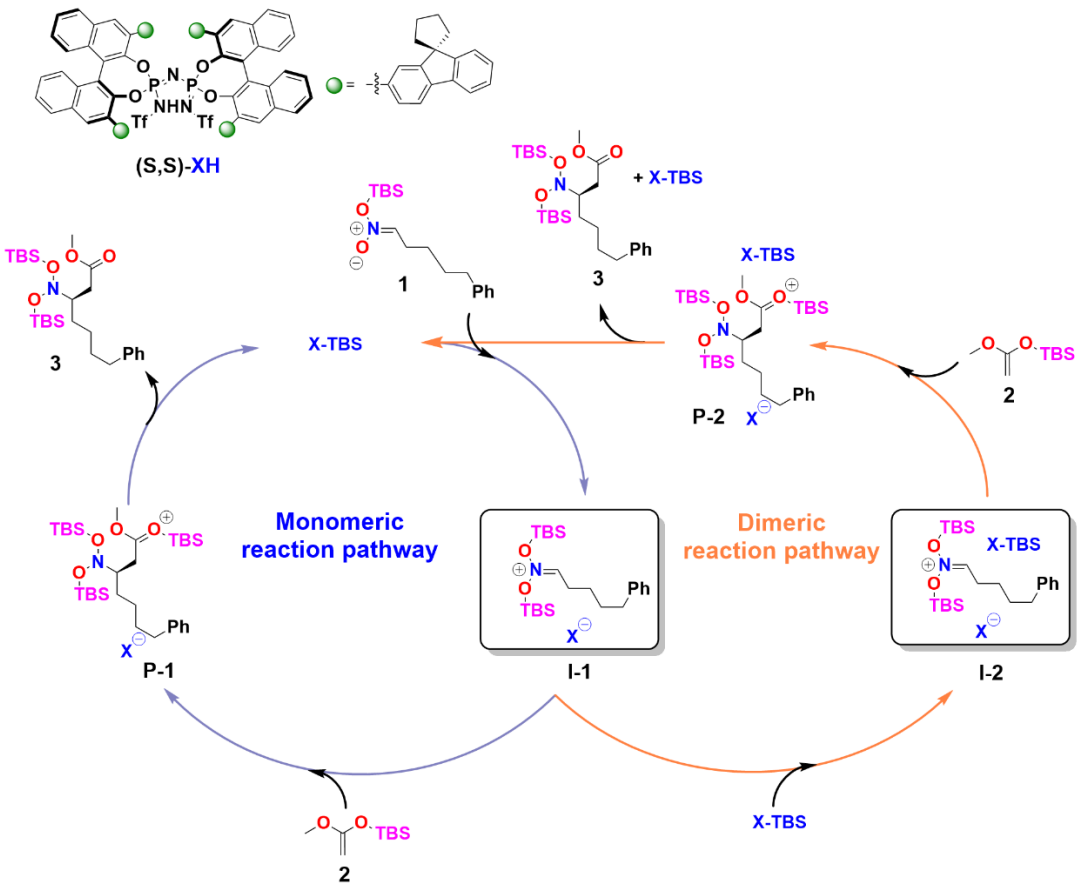

Figure 5.8 Considered monomeric and dimeric pathways for the case study reaction. Experimental details can be found in ref. 263. TBS = $SiMe_2t$-Bu. This figure was taken from ref. 262.

A plausible dimeric pathway could start from **I-1**. Association with another **X-TBS** molecule leads to the formation of **I-2**. The next step would then be the attack of **2** and product formation (*via* **P-2**), just like for the monomeric pathway.

In **Figure 5.9**, the results of the aggregation study are summarized. **Figure 5.9a** shows that the aggregation energy of **I-1** with **X-TBS** to form **I-2** is around zero. B3LYP-D3 and ωB97M-V predict slightly negative free aggregation energies (-0.8 kcal mol^{-1}), while B3LYP-D3 + ABC and ωB97M-D4 predict slightly positive free aggregation energies (2.2/4.1 kcal mol^{-1}). Direct comparison of the aggregation energy at the B3LYP-D3 and B3LYP-D3 + ABC levels of theory reveals that inclusion of many-body dispersive interactions disfavors aggregation. Moreover, the small aggregation energy indicates that the spontaneity of aggregation will depend on reaction conditions and the details of the catalyst architecture, such as choice of aryl ligands. Entropy corrections amount for 10.1 kcal mol^{-1}, indicating that the aggregation process is driven by relatively large negative aggregation energies. The results are in striking contrast to those

found for the **IDPi** acid dimer, for which large positive free dimerization energies were found for all temperatures (see **Figure 5.3**). It is noted here that the exact numbers for the **IDPi** acid dimer and **I-2** cannot be compared directly. First, for the **IDPi** acid dimer, a slightly different Lewis structure was used compared to **X-TBS** (see **Figure 5.1** and **Figure 5.2**).

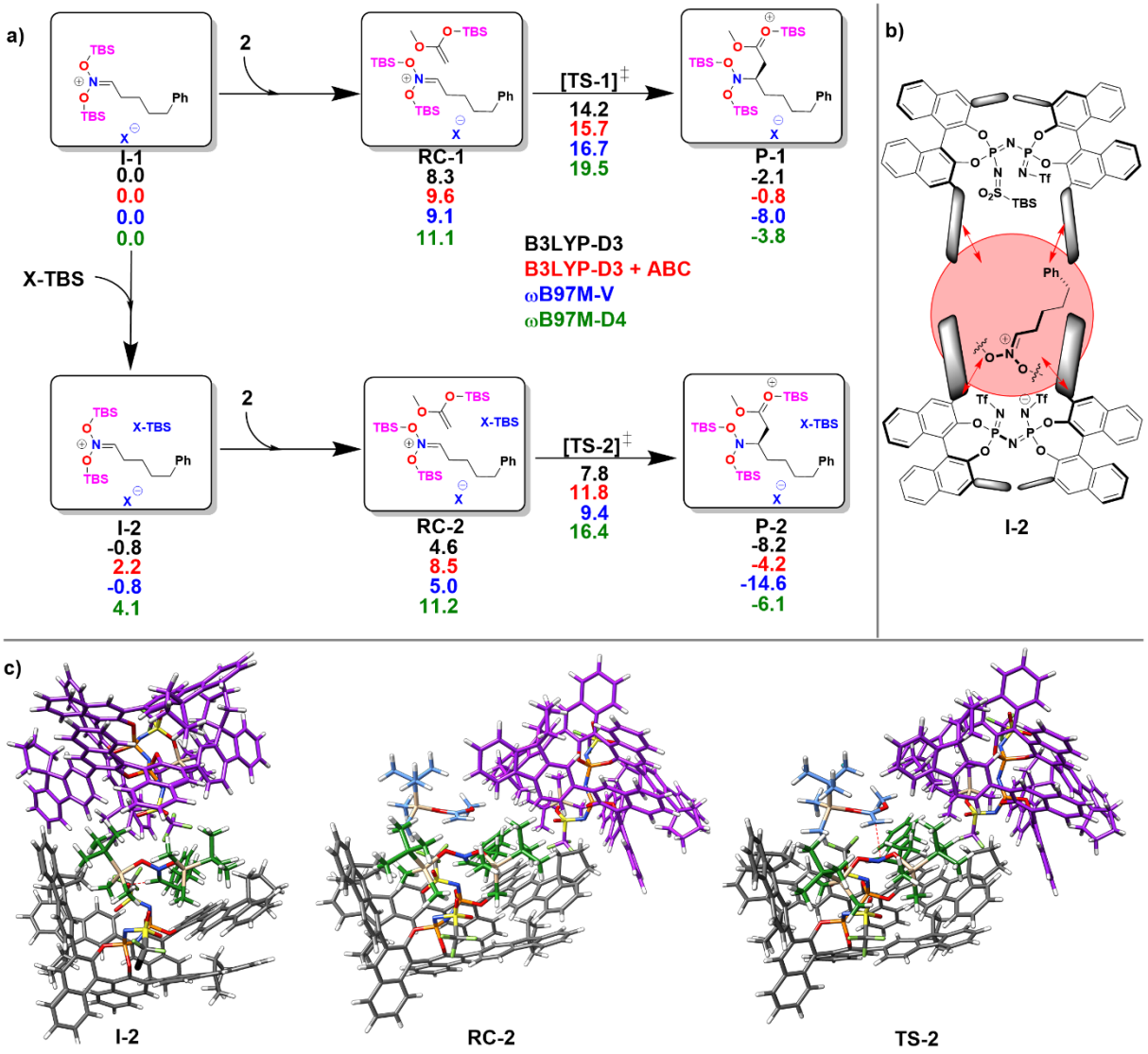

Figure 5.9 a) Reaction profiles with free energies of the individual species at different levels of theory. Structures were optimized at the r^{2}SCAN-3c/xTB level of theory. Thermochemical corrections were obtained at the same level at 153 K. b) Lewis structure of **I-2** highlighting the mediator function of the nitronate cation. c) 3D structure of **I-2**, **RC-2** and **TS-2**. The carbon atoms of the **X-TBS** molecule are shown in purple. All energies are in kcal mol^{-1}. This figure was taken from ref. 262.

In addition, in **I-2** two different IDPi species occur – the catalyst anion and the additional **X-TBS** molecule. Further, thermochemical corrections for the reaction were obtained at 153 K,

which is the experimental temperature, while for the dimerization study slightly larger temperatures (173 K) were used. Also, for studying the formation of **I-2**, dichloromethane was used as solvent, while for the dimerization study of Brønsted acids, the less polar toluene was used. However, none of these changes in structures or computational details is expected to be able to explain the significant changes noticed in the free association energy ΔG_R of **I-2**. The structural differences between the systems are small and **X-TBS** is assumed to be even bulkier than **IDPi**. In general, of course the exact free dimerization energies of the different discussed Brønsted acid classes (CPA, DTPA, DSI, IDP, IDPi) will depend on the substituents. However, the overall trends that were discussed in sections 5.3.1 and 5.3.2 should hold for any ligands that are commonly used in organocatalysis. The change in solvent also does not explain the smaller (more negative) association energy, since dichloromethane is more polar than toluene and is therefore expected to favor the dissociation. This was verified by additional calculations with toluene as solvent, for which the aggregation energy of **I-2** is 0.8 kcal mol^{-1} smaller (more negative) than for dichloromethane at the ωB97M-V level of theory. The included gCP correction favors dissociation as well (since its goal is to correct for BSSE errors that over stabilize the dimer). In addition, the most stable conformer of **I-2** was optimized at the r^{2}SCAN-3c level of theory. The association energy ΔE_R of **I-2** computed at the ωB97M-V/def2-TZVP level of theory *in vacuo* is -22.9 kcal mol^{-1} using the r^{2}SCAN-3c/xTB optimized structures and -22.2 kcal mol^{-1} using the r^{2}SCAN-3c optimized structures, indicating the minor influence of using the ONIOM approach instead of fully DFT optimized structures. This discussion demonstrates that the details of the computational settings should not be able to induce such dramatic changes in free association energies, which is why it is assumed that an alternative interaction mechanism is involved. In order to decompose the aggregation energy further, in the first step, the geometric preparation for the three fragments (catalyst anion **X**$^-$, nitronate cation **N**$^+$ and **X-TBS**) of **I-2** was computed at the ωB97M-V/def2-TZVP level of theory. For **X**$^-$ and **N**$^+$, the geometries of **I-1** were used as reference, for **X-TBS** the equilibrium geometry was used. The geometric preparation energies are given in **Table 5.3**. The total geometric preparation amounts for 8.8 kcal mol^{-1}. Notice that for **X**$^-$, a negative geometric preparation was computed. This is possible because not the equilibrium structure of **X**$^-$ was taken as reference but its geometry in **I-1**. Importantly, the total geometric preparation is much smaller than the total geometric preparation energy of the **IDPi** dimer (approx. 40 kcal mol^{-1}, see **Figure 5.4**). Since for the other methods shown in **Figure 5.9**, similarly small free aggregation energies were found, it is expected that also the geometric preparation energy computed with these methods will be in the same range. In order to understand why the distortion energies are relatively small,

interaction energies between each pair of fragments were computed at the HFLD/def2-TZVP(-f) level of theory.

Table 5.3 Geometric preparation energies of the three fragments of **I-2** at the ωB97M-V/def2-TZVP level of theory. For $\mathbf{X}^-$ and $\mathbf{N}^+$, the corresponding structures of **I-1** were taken as reference. For **X-TBS**, the equilibrium structure was taken as reference. All energies are in kcal mol^{-1}.

Fragment	$\Delta E_{geo-prep}$
$\mathbf{X}^-$	-1.2
$\mathbf{N}^+$	8.8
X-TBS	1.2

For analysis of the non-dispersive interactions, HF/def2-TZVP single point calculations were performed for each pair of fragments. Dispersive interactions were obtained, as discussed before, by solving approximate DLPNO-CCSD amplitudes equations. The results are shown in **Table 5.4**.

Table 5.4 Fragment pair-wise interaction energies of the three fragments of **I-2** at the HFLD/def2-TZVP(-f) level of theory. All energies are in kcal mol^{-1}. This table was adapted from ref. 262.

Fragment pair	E_{disp}	$\Delta E_{non-disp}$
$\mathbf{X}^-$, $\mathbf{N}^+$	-47.5	-50.6
$\mathbf{X}^-$, **X-TBS**	-5.7	7.3
$\mathbf{N}^+$, **X-TBS**	-18.4	-2.4

As expected, strong dispersive and even stronger non-dispersive interactions arise between the catalyst anion $\mathbf{X}^-$ and the nitronate cation $\mathbf{N}^+$. These strong interactions lead to the very exergonic formation of the monomeric CIP **I-1** (-19.5 kcal mol^{-1} at the ωB97M-V/def2-TZVP level of theory) and therefore are the driving force behind the reaction mechanism. The two catalyst molecules feature weak dispersive and weak repulsive non-dispersive interactions. Interestingly, strong dispersive interactions (-18.4 kcal mol^{-1}) and weak attractive non-dispersive interactions arise between the nitronate and **X-TBS**. The computed interaction

energies show that the two catalyst molecules actually interact very weakly, while the nitronate cation has strong attractive interactions with both catalyst molecules. This paves the way for the mechanism depicted in **Figure 5.9b**. The nitronate cation acts as mediator that can "glue" the two catalyst molecules together by interacting with both molecules either *via* electrostatic and/or dispersive interactions, while the catalyst molecules only interact weakly with each other. The very weak interactions between the two catalyst molecules lead to small distortion energies. For the bulky acid dimers (**IDP** and **IDPi**), it was shown that significant structural distortions must occur in order to allow for strong attractive LD interactions that, however, cannot counteract the distortion energies and penalties from entropy and solvation and therefore lead to very positive dimerization energies (see **Figure 5.4**). The proposed mechanism with mediating substrate molecules of course does not prove that indeed cooperative catalytic effects influence the reaction outcome of this particular case study. However, it provides access to cooperative catalytic effects even for very bulky confined catalysts, for which it is generally assumed that monomeric pathways determine the reaction outcome. In **Figure 5.9c**, the 3D structure of **I-2** is shown, which has a sandwich-like conformation, with the nitronate molecule being located between the catalyst molecules. Since the computed free association energy ΔG_R is relatively close to zero, whether aggregation is possible depends on the details of the experimental settings, such as temperature, polarity of the solvent, catalyst ligands, etc.

Under what circumstances association might occur? Besides the already discussed influences like temperature and solvent, it is assumed that the structure of the monomeric CIP **I-1** is important. For example, if the nitronate is very small, it might be possible that it is buried "too deep" in the catalyst pocket, making interactions with **X-TBS** much harder. Additionally, smaller nitronates lead to weaker LD interactions as long as no specific interactions such as hydrogen bonds are present. If on the other hand the nitronate is too large, formation of **I-1** already might be inhibited or interaction with **X-TBS** requires additional structural distortions.

In **Figure 5.9a**, also the energies of the monomeric (**RC-1**) and dimeric (**RC-2**) reactant complexes and transition states (**TS-1/TS-2**) for the pathway leading to the major enantiomeric product are given. Interestingly, the dimeric pathway provides lower energies for **RC-2** and **TS-2** compared to the monomeric species, indicating its rate-lowering effect. However, it can be seen that the extent is very functional-dependent. For example, ωB97M-V predicts **TS-2** to be 7.3 kcal mol^{-1} more stable than **TS-1**, while ωB97M-D4 predicts **TS-2** to be only 3.1 kcal mol^{-1} more stable. However, all four density functionals show qualitatively the same behavior. Nevertheless, the relatively large variations between the different exchange-correlation

functionals and the very complex nature of the system and its interactions limit the confidence of quantitative predictions. Since the energy difference between the two pathways is relatively small, the mechanistic details of the transformation are likely determined by the experimental parameters.

When comparing the structures of **I-2** and **TS-2** (see **Figure 5.9c**), it is seen that in **I-2** the **X-TBS** molecule and the catalyst anion form a more or less compact pocket into which the nitronate is encapsulated. In **RC-2** and **TS-2**, the **X-TBS** molecule has moved in space in order to open up the pocket and to increase the space the nucleophile can occupy. It is noted here that **TS-2** is not obtained from **I-2**, but from **TS-1** by adding another **X-TBS** molecule to the system. Moreover, it should be kept in mind that complete exploration of the conformational space was not possible due to the challenging system size.

In future, combined experimental and computational studies (*e.g.*, kinetics studies or NMR) could be carried out in order to gain deeper insights into the possibility and likeliness of cooperative catalytic effects.

Chapter 6 Explicit Solute-Solvent Interactions in Molecular Balances

6.1 Introduction

Elements of this chapter including figures have been published in a journal article [277]:

Schümann, J. M.; Ochmann, L.; Becker, J.; Altun, A.; Harden, I.; Bistoni, G.; Schreiner, P. R. Exploring the Limits of Intramolecular London Dispersion Stabilization with Bulky Dispersion Energy Donors in Alkane Solution. *Journal of the American Chemical Society* **2023**, *145* (4), 2093-2097. https://doi.org/10.1021/jacs.2c13301.

Some of the used figures in this chapter are part of the publication but were not designed by the thesis author. Therefore, for each figure the corresponding figure caption describes whether the figure is part of a publication and was designed by the thesis author.

In the last two chapters, it was shown that LD can be of critical importance in catalytic transformations. In Chapter 4, an induced-fit model for an IDPi-catalyzed intramolecular hydroalkoxylation reaction was proposed in which intra-catalyst LD determines the geometric preparation energy, while intermolecular LD is the major stabilization interaction that drives the reaction. In Chapter 5, LD was shown to be the most important attractive interaction in Brønsted acid dimers as well as in ion pair aggregates. In this chapter, a computational protocol for sampling the conformational space of solvent molecules that allows for studying explicit solute-solvent interactions is proposed, applied and discussed. For studying intramolecular LD interactions, a commonly used class of substances are molecular balances (MB) [278]. In MBs, two different isomers are in equilibrium with each other. One of the isomers features noncovalent interactions of two groups of interest, while in the other isomer the interaction is missing. The corresponding isomers often are called “folded” or “unfolded” isomers. Conversion between the two isomers might be achieved by *e.g.*, rotations around single bonds or, as in the case study to be introduced, by valence-bond isomerization. In general, the equilibrium between the isomers is governed not only by intramolecular LD effects, but also by solute-solvent interactions and entropic effects. In this chapter, the equilibrium between 1,4- and 1,6-disubstituted cyclooctatetraene (COT) is studied. Previous studies by the group of Schreiner showed that the 1,6-COT isomer (the “folded” isomer) is more stable than the corresponding 1,4-COT isomer in various solvents. As substituent, the *tert*-butyl group was

used [279]. The relative stabilization of the 1,6-COT isomer was attributed to additional NCIs between the *t*-Bu groups.

The aim of the case study discussed here is to evaluate the strength of intramolecular NCIs with different substituents (1. *Tert*-butyl (*t*-Bu), 2. Adamantyl (Ad), 3. Diamantyl (Dia)) and the interplay with the solvent. As part of the joint project together with the group of Schreiner from the Justus-Liebig-Universität Gießen, the thesis author developed and applied a protocol for treating explicit solute-solvent interactions based on constrained conformer sampling with the CREST program. As an overview, the 1,4-/1,6-COT equilibrium is shown in **Figure 6.1**. For investigating the solute-solvent interactions, a hybrid explicit and implicit solvation model was applied that will be introduced in the next section. As solvent for the computational study, *n*-dodecane will be used which is the bulkiest experimentally used solvent.

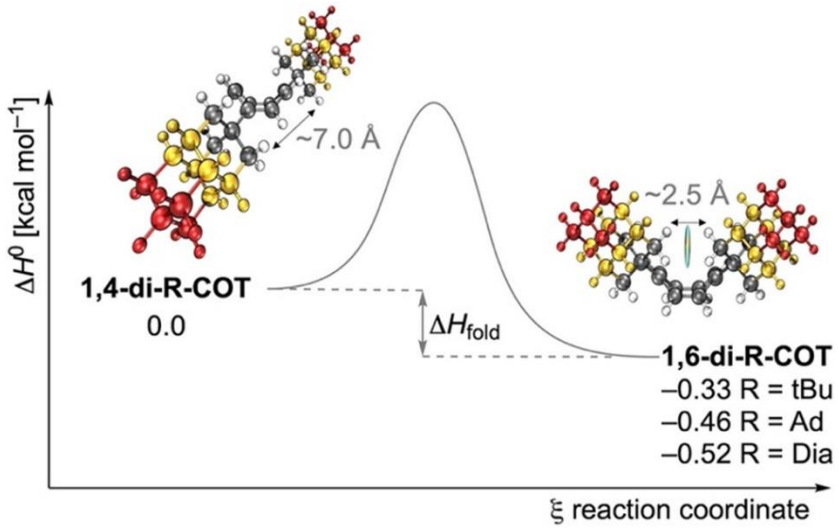

Figure 6.1 Equilibrium between disubstituted 1,4- and 1,6-COT isomers. Experimental reaction enthalpies are given for three different substituents. All energies are in kcal mol^{-1}. This figure was taken from ref. 277. This figure was not designed by the thesis author.

6.2 Computational protocol

6.2.1 Towards a space-filling model for obtaining the first solvation shell of solvent molecules

The aim of the computational protocol described in this chapter is to find the lowest-in energy conformers of a solvent network surrounding the solute. The conformers are obtained using the CREST program by Grimme. By surrounding the solute with solvent molecules, the total number of atoms in the system grows fast, especially if large molecules are to be covered. This is why only one solvation shell is placed around the solute.

For placing the solvent molecules around the solvent, a simple space-filling model was developed and implemented in a C++ program. This program is not restricted to the COT molecules, but is applicable to any solute/solvent pair. From the user, three files are required: the coordinates of the solute and solvent and the coordinates of the solute's SAS. The SAS is obtained from a single point energy calculation using the `surfacetype gepol_sas` keyword in the `%cpcm` block in ORCA together with *n*-dodecane as solvent. In addition, the solvent radius might be adjusted [280]. The subsequent workflow is actually pretty simple. The center of mass of the solvent is shifted to zero. Within the main loop, the code iterates over all SAS points. For each point, it is checked whether it is "safe" to add a solvent molecule. This is the case if the distance between each atom of the particular solvent molecule to be added and each atom of the solute is larger than a certain cutoff $d_{solute\text{-}solvent}$ and if the distance between each atom of the particular solvent molecule and each atom of every solvent molecule added before is larger than another cutoff $d_{solvent\text{-}solvent}$. If a distance is smaller than the corresponding cutoff, the surface point is skipped, otherwise the solvent molecule is added. Once all surface points were visited once, the solvent molecule is rotated around an arbitrary axis and the procedure is started again. This step is repeated with two different rotations of the solvent molecule in order to increase the surface coverage. The two cutoff parameters have default values, but user defined values can be passed *via* the command line. $d_{solvent\text{-}solvent}$ might be derived from optimization of the solvent dimer. As an example, the first solvation shell of the transition state **TS1-1** from Chapter 4 is shown in **Figure 6.2**. From the example system in **Figure 6.2**, it can be seen that both the solute and the solvent are approximately of spherical shape. In this regard, the relative orientation of a particular solvent molecule with respect to the solute is fairly unimportant. The situation is entirely different for the 1,4-/1,6-COT isomers and the used solvent (*n*-dodecane), which both are heavily extended along one space dimension.

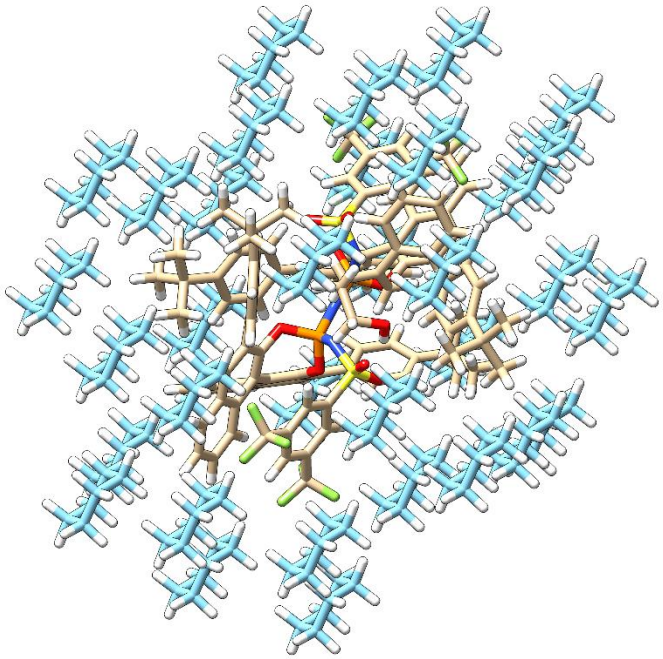

Figure 6.2 First solvation shell of cyclohexane molecules surrounding **TS1-1** from Chapter 4. This figure was prepared by the thesis author.

This is why the space-filling program is able to run multiple such runs with three initial orientations of the solvent relative to the solute. This consequently leads to three different output files, each containing a different dominant solvent orientation. It is noted here that the proposed space-filling model does not contain any physics, it is entirely based on distance measurements. Its purpose is to provide an equally covered solute surface. One disadvantage is that it is not possible to specify the number of solvent molecules to be added, nor information about favored positions and orientations of the solvent molecules can be deduced. Recently, the Grimme group developed a quantum cluster growth (QCG) model in which solvent molecules are added subsequently to the system [281]. The favored position of the next solvent molecule is determined by some docking search and once the desired number of solvent molecules has been added, a conformer ensemble is generated automatically. The advantages of such a model are that the number of solvent molecules to be added can be specified by the user and that the positioning of the solvent molecules is based on force field optimizations. On the other side, QCG models can in principle lead to solvent clusterification, meaning that instead of covering the whole SAS of the solute, new solvent molecules prefer to interact with already added solvent molecules (*i.e.*, the solvent molecules form their own cluster). However, for future studies on explicit solvation, the QCG model should be considered as alternative.

For the discussed COT isomers, three different initial structures corresponding to three different orientations of the *n*-dodecane molecules were obtained. In order to evaluate the influence of the number of surrounding solvent molecules, initial structures with a total of eight and ten solvent molecules were generated. This leads to a total of 12 initial starting structures for the subsequent conformer sampling. The starting structures are shown in **Figure 6.3**.

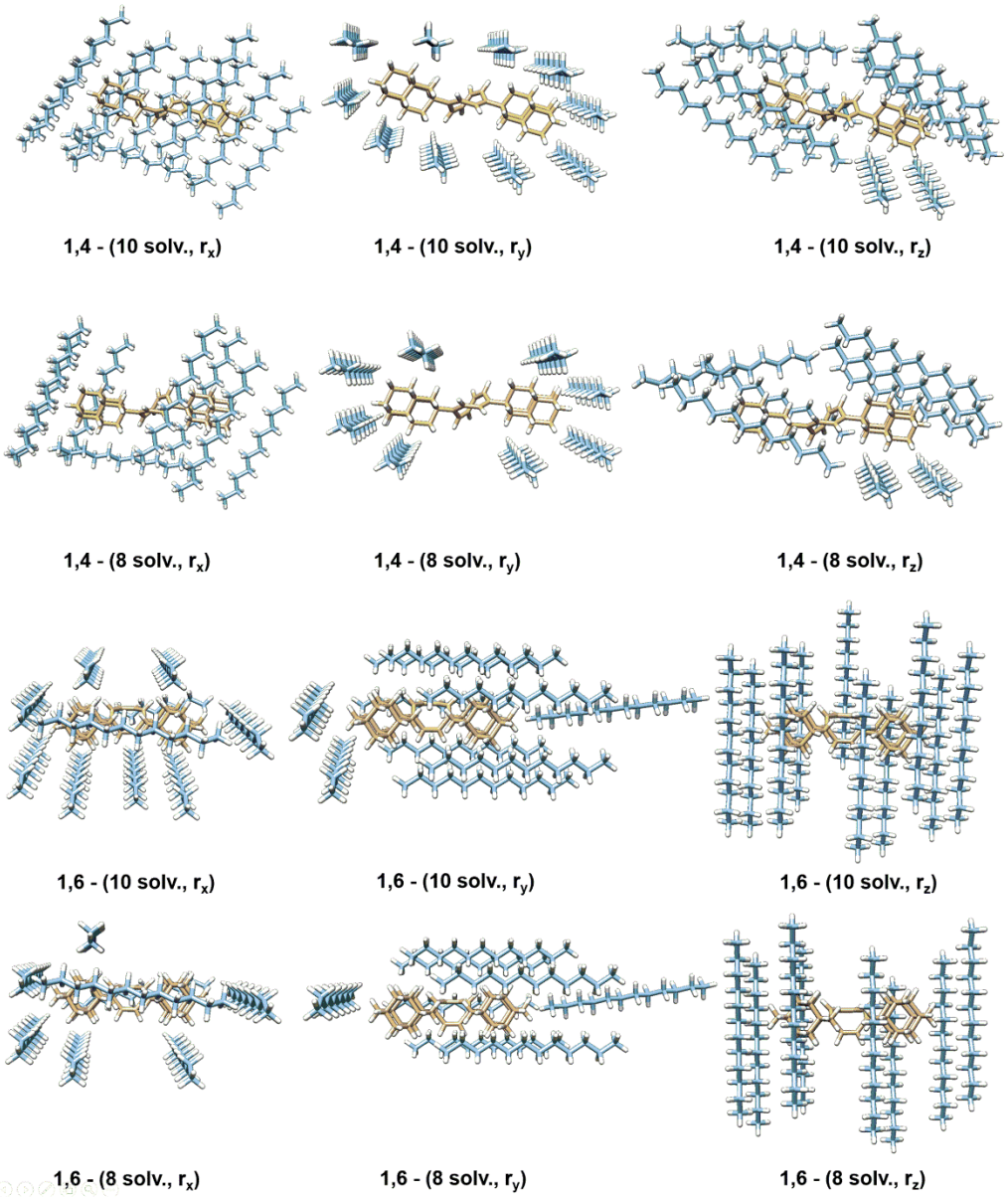

Figure 6.3 Initial starting structures for the 1,4-/1,6-bis(Dia)-COT obtained from a space-filling model for three different orientations of the solvent (*n*-dodecane). Starting structures were obtained with eight and ten solvent molecules. This figure was taken from ref. 277. This figure was prepared by the thesis author.

6.2.2 Constrained conformer sampling

For obtaining the most stable conformers, thorough conformer sampling was performed using CREST together with the GFN-FF force field. During the MTD simulations, the solute was constrained to its DFT optimized geometry, since focus is on the solvent network. For the CREST runs, the `nci` keyword in CREST was used [151]. The collective variables in the MTD simulations are the RMSD values with respect to some reference structures. The algorithm tries to maximize the RMSD with respect to previously explored local minima. For aggregates that are held together by weak NCIs, MTD simulations usually lead to dissociation of the aggregate. For the solvent covered COT isomers, this would mean that the solvent molecules would diffuse into the vacuum. The `nci` keyword sets up a bias potential that counteracts the MTD penalty potential and therefore prevents dissociation of the aggregate. For the initial conformer sampling run, an MTD simulation time of 400 ps was chosen. After the first run, additional CREST runs were performed in order to extend exploration of the conformational space. For all subsequent CREST runs, an MTD simulation time of 200 ps was chosen. As criterium for convergence, the total energy E_{tot} cannot be used. The reason is that constraining geometric parameters (in this case the cartesian coordinates of the solute's atoms) during MTD runs is never exact because this would lead to instabilities during the dynamics simulations. Instead, the geometric parameters can only be frozen approximately. This leads to very small but still measurable geometric changes of the solute during the CREST run which consequently influence the total energy of the system (because every new MTD run decreases the solute's energy as its geometry slowly approaching a local energy minimum on the GFN-FF PES). As convergence criterium, the interaction energy E_{int} between the solute and the solvent network of the most stable conformer was used.

$$E_{int} = E_{tot} - (E_{solute} + E_{solvent}). \tag{6.1}$$

E_{solute} is the energy of the solute, $E_{solvent}$ is the energy of the solvent network. With this convergence criterium, convergence of E_{int} to 0.1 kcal mol^{-1} was achieved within typically 2-3 CREST runs. This procedure was performed for all 12 initial starting structures shown in **Figure 6.3**. The most stable conformer obtained from each initial guess structure was optimized using the r^{2}SCAN-3c composite method as implemented in ORCA 5.0. No geometric constraints were used during the optimization. Final single point energies were obtained using a hybrid explicit/implicit model by placing the explicitly solvated molecules into the bulk solvent modeled at the C-PCM level. Final single point energies were hence obtained at the r^{2}SCAN-3c/C-PCM(*n*-dodecane) level of theory.

6.3 Results and discussion

The most stable r^{2}SCAN-3c optimized conformers of the solvated 1,4-/1,6-bis(Dia)-COT isomers with 10 solvent molecules are shown in **Figure 6.4**.

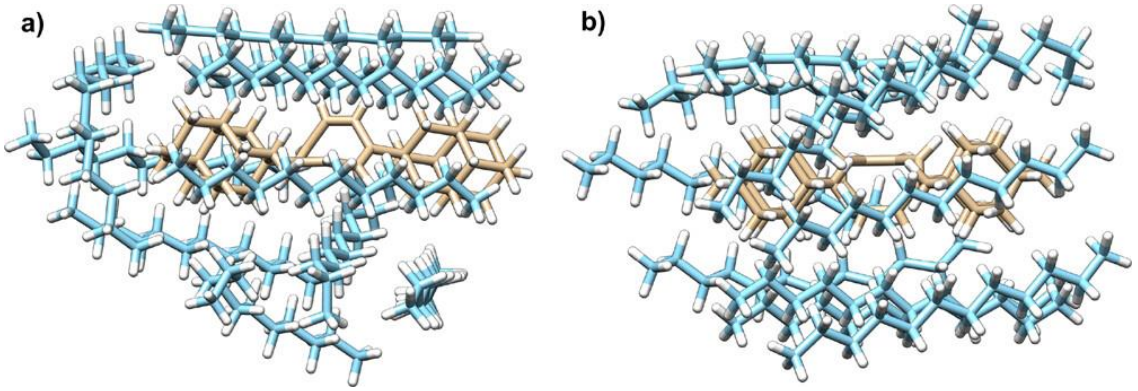

Figure 6.4 Most stable conformer obtained from the conformer sampling protocol for a) 1,4- and b) 1,6-bis(Dia)-COT. Structures with 10 *n*-dodecane molecules are shown. This figure was taken from ref. 277. This figure was prepared by the thesis author.

For both isomers it can be seen that the solvent molecules align parallel to the solute. This effect originates from the significant geometric anisotropy that both the solute and the solvent molecules show. By parallel alignment, intermolecular dispersive interactions between the solute and solvent, but presumably also between different solvent molecules, can be maximized. Interestingly, the crystal structure of 1,4-bis(Dia)-COT unveils a parallel alignment of the COT molecules in the solid state, which underpins the model based on intermolecular LD effects [277].

For analyzing the total energy difference ΔE_{tot} between the 1,4- and 1,6-COT isomers, it is decomposed into contributions from the solute and solvent.

$$\begin{aligned} \Delta E_{tot} &= E_{tot}(1,6) - E_{tot}(1,4) \\ &= [E_{int}(1,6) - E_{int}(1,4)] + [E_{solute}(1,6) - E_{solute}(1,4)] \\ &\quad + [E_{solvent}(1,6) - E_{solvent}(1,4)] \\ &= \Delta E_{int} + \Delta E_{solute} + \Delta E_{solvent}. \end{aligned} \tag{6.2}$$

For comparison with either experiment or implicit solvation calculations, only the ΔE_{tot} and E_{solute} are relevant. In **Table 6.1**, these are shown for the models with 8 and 10 solvent molecules together with a comparison to experiment and C-PCM calculations.

Table 6.1 ΔE_{tot} and ΔE_{solute} between 1,6-bis(Dia)-COT and 1,4-bis(Dia)-COT at the r^{2}SCAN-3c/C-PCM(*n*-dodecane) level of theory. Explicit/implicit hybrid models with 8 and 10 solvent molecules were considered. All energies are in kcal mol^{-1}. This table was adapted from ref. 277.

	ΔE_{solute}	ΔE_{tot}
8 solvent molecules	-0.41	-2.03
10 solvent molecules	-0.64	-0.89
Implicit solvation	-0.79 (gas phase)	-0.66 (C-PCM)
Experiment	-	-0.21 (ΔG_{298K})

It can be seen that the explicit/implicit hybrid approach predicts the 1,6-COT isomer to be more stable, irrespective of the number of explicit solvent molecules. This is in agreement with implicit solvation calculations and with experiment and shows the relative robustness of the proposed protocol. It further can be seen that the agreement of ΔE_{tot} of the smaller model (8 solvent molecules) with experiment is worse than the agreement of the larger model. The sign, however, is reproduced correctly. The larger hybrid model is able to reproduce the implicit model and the experiment remarkably well. The very small differences of ΔE_{solute} between the hybrid models and the implicitly solvated model stem from minor geometric changes of the solute within the hybrid models. The proposed space-filling model allows for fast generation of equally covered first solvation shells and can therefore safe a lot of user's time compared to manual design. Functionalities for adding more than one solvation shell and adding a user-defined number of solvent molecules are to be implemented. The drawback of such a space-filling model is the absence of any physics that can help in automatically finding preferred positions and orientations of the solvent molecules. Automatic tools like the previously discussed QCG tool or the automatic docking algorithm aISS [282], both developed by the Grimme group, can be dedicated tools for future studies. However, the here proposed conformer sampling based on three different initial structures provides profound access to the conformational space. The case study of the 1,4-/1,6-bis(Dia)-COT equilibrium showed that the larger explicit/implicit hybrid model with 10 solvent molecules represent the experimental energy differences very well. In addition, explicit solvation models allow for identification of preferred orientation of the solvent models, which becomes important, as was seen for the COT equilibrium, if both solute and solvent are geometrically anisotropic.

Chapter 7 A Combined Spectroscopic and Computational Study on the Mechanism of Iron-Catalyzed Aminofunctionalizations

7.1 Introduction

Elements of this chapter including figures have been published in a journal article [283]:

Chatterjee, S.; Harden, I.; Bistoni, G.; Castillo, R. G.; Chabbra, S.; van Gastel, M.; Schnegg, A.; Bill, E.; Birrell, J. A.; Morandi, B.; Neese, F.; DeBeer, S. A Combined Spectroscopic and Computational Study on the Mechanism of Iron-Catalyzed Aminofunctionalization of Olefins Using Hydroxylamine Derived N–O Reagent as the "Amino" Source and "Oxidant". *Journal of the American Chemical Society* **2022**, *144* (6), 2637-2656. https://doi.org/10.1021/jacs.1c11083.

For each figure, the corresponding figure caption describes whether the figure is part of a publication.

In this joint project with the DeBeer group of the Max Planck Institute for Chemical Energy Conversion, spectroscopic and computational approaches are combined to shed light on the mechanism of iron-catalyzed aminofunctionalizations. In the previous chapters, modern computational strategies were applied to tackle different aspects of homogeneous organocatalysis, such as novel models for describing catalyst-substrate interactions, cooperative catalytic effects or explicit solute-solvent interactions in molecular balances. However, alongside with conformational and environmental effects, the electronic structure of the catalyst is another important variable determining the key properties of a chemical transformation. The aim of this chapter is to apply some aspects of the previously used computational protocols (*e.g.*, the conformer sampling) to a challenging transition metal catalyzed reaction in order to unveil electronic structure effects in homogeneous catalysis. In all previous case studies, the main connection to experiment were energy differences (be it the stereoselectivities in Chapter 4, the energy difference between isomers in Chapter 6 or qualitative comparison, *i.e.*, dimerization or no dimerization in Chapter 5). Due to the potentially complicated electronic structure of transition metal intermediate species, however, the case study investigated in this chapter was subject to thorough spectroscopic analyses. Therefore, the aim is to correlate experimental spectroscopic data with computed spectroscopic parameters in order to establish the preferred models for the involved species. The case study considered in this chapter

describes the amino methoxylation of styrene catalyzed by bis(aqua)bis(acetylacetonate)iron(II). The reaction is outlined in **Figure 7.1**.

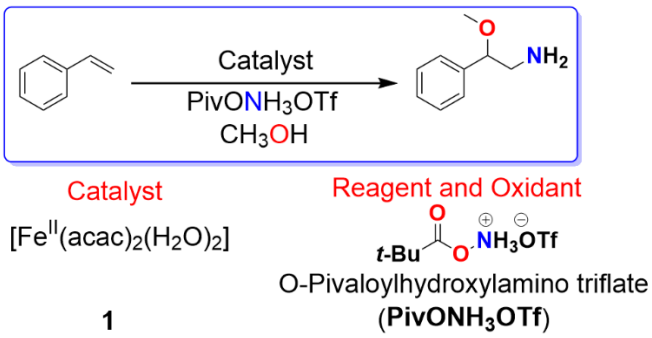

Figure 7.1 Reaction considered in this chapter together with the used catalyst **1** and aminating reagent **PivONH$_3$OTf**. This figure was adapted from ref. 283.

The amine functionality is a common structure motif that is ubiquitous in nature and industries. Aliphatic amines alone have an estimated production capacity of over 500,000 tons per year and they are used *e.g.*, as precursors of herbicides, insecticides or pharmaceuticals [284]. In addition, aromatic amines such as aniline are starting points for more than 300 different chemicals and are for example of very high importance for the synthesis of polyurethanes [285]. In biological systems, amines function for example as neurotransmitters (*e.g.*, dopamine) [286] and are part of amino acids that are needed for protein biosynthesis [287].

However, direct synthesis of unprotected amines and their derivates such as amino alcohols by addition reactions starting from simple olefins still are challenging [288]. Additionally, many of the established synthesis strategies lead to protected amines. Due to the high importance of amines in various aspects of life, developing efficient and sustainable synthesis strategies for unprotected amines is a major challenge in contemporary synthetic chemistry. As Fürstner has pointed out in a review from 2016, homogeneous iron catalysis is likely to become even more important in mainstream organic chemistry [289]. In 2016, the Morandi group started a research project that focuses on installation of unprotected amino functionalities by addition reactions *via* iron(II) catalysis [290]. As "amino source" they used O-pivaloylhydoxylamino triflate (PivONH$_3$OTf), which is air stable and non-poisonous. Using this aminating reagent, they tested various relatively simple Fe(II) catalysts for the oxyamination of styrene to regioselectively obtain 2-amino-1-phenylethanol. While they obtained decent yields with $FeSO_4$, Fe(II) phthalocyanine gave good overall yields and regioselectivities. With the illustrated scheme of simple Fe(II) catalysts and PivONH$_3$OTf as amino source, they were able to synthesize various 2-aminoethanol derivates starting from different styrene derivates. In the

following years, the Morandi group extended the reaction scope to aminochlorinations [291,292], aminoazidations [293] or amino-oxidation of thiols [294]. Sulfimidation of sulfoxides was reported by Bolm and coworkers [295].

Although the introduced paradigm was successful so far and the reaction scope was extended to several highly important aminofunctionalizations, deeper insights into the mechanisms of aforementioned transformations are still missing. Especially the efficiency of relatively simple iron(II) catalysts is still puzzling. Therefore, in the joint project discussed in this chapter, a broad spectroscopic and computational analysis of the case study given in **Figure 7.1** was carried out in order to establish the number of occurring intermediates as well as their geometric, electronic and spectroscopic properties. In addition, a plausible reaction mechanism was derived from experimental observations and computations.

7.2 Overview of experimental results

In this section, the main results from the experiments and the spectroscopic analyses are shown and discussed. Note that the experiments were carried out by Chatterjee and coworkers but not by the thesis author. Also, the figures shown in this section were not prepared by the thesis author but were taken from the joint journal article [283].

After synthesis of **1** and **PivONH$_3$OTf** (see **Figure 7.1**), optical absorption spectra of **1** and of **1** mixed with **PivONH$_3$OTf** after 15 mins and 90 mins were taken. The results are shown in **Figure 7.2**. The precursor **1** dissolved in dichloromethane yields a yellow solution (black spectrum in **Figure 7.2**). Most noticeably is the low-energy chromophore at around 23,000 cm^{-1}. After adding the aminating reagent **PivONH$_3$OTf** to the solution, the solution turns red almost instantly. The corresponding spectrum at 15 mins after mixing is shown in red in **Figure 7.2**. The low-energy feature was red-shifted to around 20,500 cm^{-1} and also shows reduced intensity compared to the precursor. After 90 mins, the solution turns violet with the low-energy chromophore again appears red-shifted, broadened and with overall lower intensity.

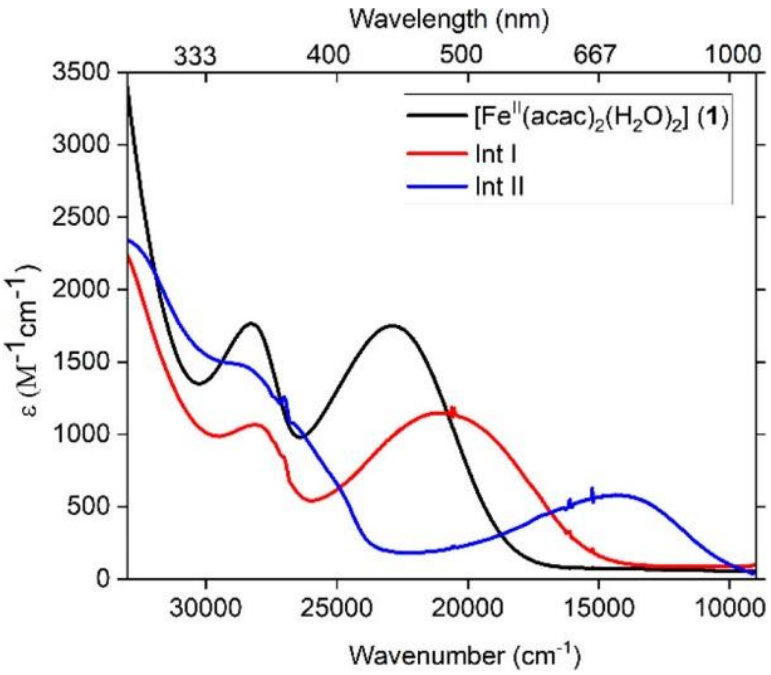

Figure 7.2 UV-vis absorption spectra of **1** (black), **1 + PivONH$_3$OTf** after 15 mins (red) and **1 + PivONH$_3$OTf** after 90 mins (blue) in CH_2Cl_2 at 293 K. This figure was taken from ref. 283.

Gaussian deconvolution was carried out by the thesis author (see **Figure A 1**), which will be discussed in the computational section. The overall time evolution of the absorbance indicates the presence of two intermediates that are labeled in the further **Int I** and **Int II**. The characteristic low-energy chromophores (480 nm for **Int I**, 700 nm for **Int II**) were further used for kinetics studies based on ultraviolet-visible (UV-vis) spectroscopy. In the absence of the substrate styrene, **Int I** shows a half-time for its decay of 35 mins. The formation of **Int II** is observed at the same time scale. The decay of **Int II** has a half-time of 260 mins. In the presence of styrene, the decay rate of **Int I** does not change, whereas the decay of **Int II** happens significantly faster compared to the self-decay in the absence of styrene. The kinetics studies therefore suggest that **Int II** is formed from **Int I** and that **Int II** is active in reaction with styrene, while **Int I** is not. In addition, reaction rates were obtained for different *para*-substituted styrene derivates. Electron-donating substituents like the methoxy group increase the reaction rates, indicating that more nucleophilic styrene derivates are more active in the nitrogen transfer reaction.

As was demonstrated, both intermediates are relatively long-living and can therefore be frozen out in solution and can be analyzed by means of various spectroscopic techniques. In **Figure 7.3**, the results from electrospray ionization mass spectrometry (ESI-MS) are shown for both intermediates **Int I** and **Int II**.

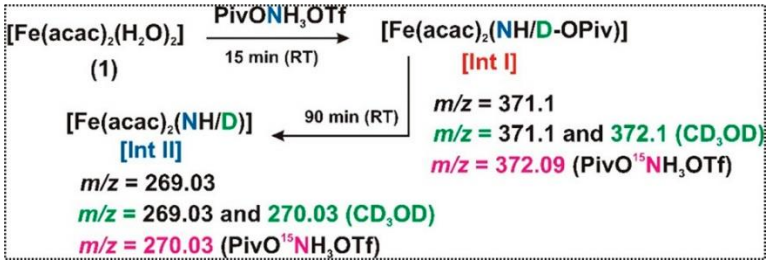

Figure 7.3 ESI-MS results of a solution of **1 + PivONH$_3$OTf** in CH_2Cl_2/CH_3OH at 15 mins and 90 mins after mixing, respectively. This figure was taken from ref. 283.

For **Int I** a dominant fragment with a mass over charge (m/z) ratio of 371.1 was found. Experiments in deuterated solvent indicate one exchangeable proton in **Int I**. The found mass corresponds to the sum formula [Fe(acac)$_2$(PivONH) + H$^+$]. Experiments with ^{15}N labeled aminating reagent **PivONH$_3$OTf** proof that the PivONH group of **Int I** originates from the aminating reagent. For **Int II** dominant fragments with m/z = 269.03 and m/z = 270.04 corresponding to the sum formulas [Fe(acac)$_2$(NH)]$^+$ and [Fe(acac)$_2$(NH) + H$^+$] were found. It becomes evident that **Int II** is formed from **Int I** by the loss of the O-pivaloyl group.

For receiving information about the total spin state of the system and the oxidation state of iron, electron paramagnetic resonance spectroscopy (EPR) measurements were carried out. In EPR, splitting of magnetic sublevels of a paramagnetic substance in an external magnetic field is used for measuring transitions between these sublevels induced by electromagnetic radiation. Usually, systems with an even number of unpaired electrons and therefore an integer total spin S do not give any signals in perpendicular EPR mode [296]. These systems are called non-Kramers systems. In a Kramers system, each state is at least doubly degenerate in the absence of magnetic fields (so-called Kramers doublet) [297]. Half-integer spin systems are Kramers systems. Therefore, the lowest magnetic sublevel is guaranteed to be degenerate without magnetic field. Each Kramers doublet will split in the presence of an external magnetic field. On the other hand, the lowest magnetic sublevel of a non-Kramers system might not be degenerate and therefore might not split up in the presence of the magnetic field. Therefore, if the zero-field splitting is much larger than the available photon energy, no resonance can be observed in these systems [296]. In EPR spectroscopy, the effective g-tensor g_{eff} is measured. It is an *effective* g-tensor because the actual energy of the magnetic sublevels is determined by a multitude of interactions such as electronic Zeeman effect, zero-field splitting, hyperfine coupling interactions, etc.

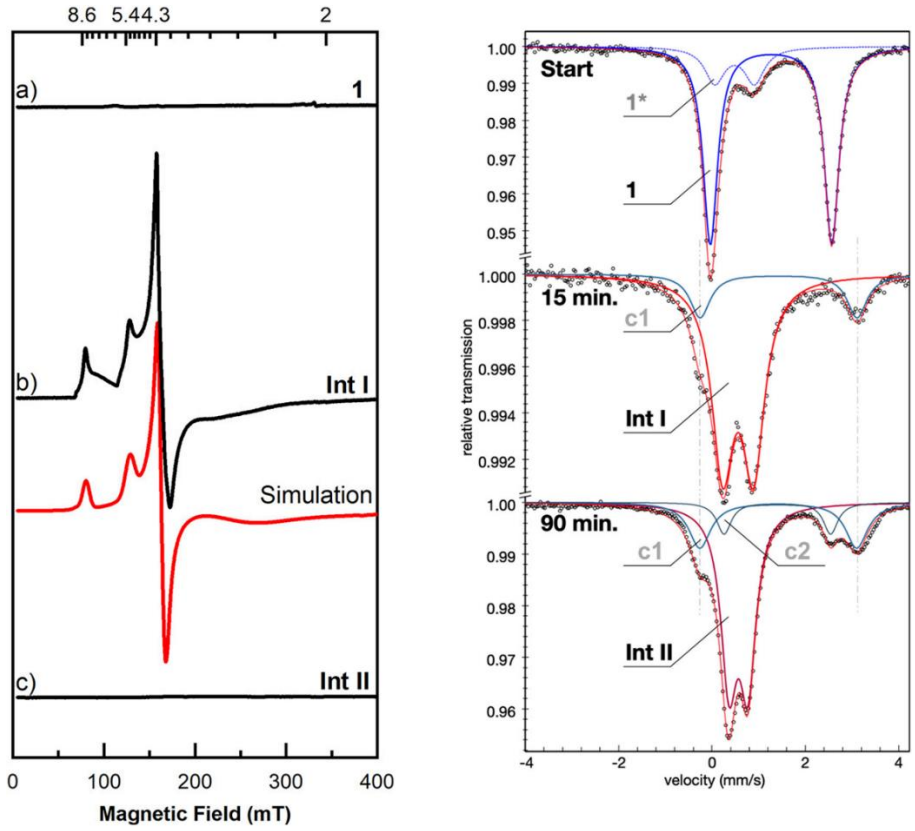

Figure 7.4 Experimental EPR and Mössbauer spectra. Left: experimental (black) and simulated (red) continuous wave EPR spectra of **1** (a), **Int I** (b) and **Int II** (c) at a frequency of 9.6 GHz. Measurements were carried out at 10 K. The simulated spectrum of **Int I** was obtained from superposition of two subspectra with a total spin $S=2.5$. Right: Mössbauer spectra of **1** (top), **Int I** (middle) and **Int II** (bottom) at 80 K. This figure was adapted from ref. 283.

As can be seen from **Figure 7.4**, no EPR signals are detected for **1** and **Int II**. **1** is known to be a ferrous complex, which is in agreement with the EPR silence of the species. **Int II** is described best as integer spin system as well. **Int I** on the other hand does show resonance and is therefore best described as half-integer spin system. The resolved peaks have effective g-values of g_{eff} = 9.3, 8.6, 5.4, 4.3. Simulations with a total spin $S=2.5$ were carried out that could reproduce the experiment reasonably well. EPR indicates that **Int I** is a high-spin iron(III) *i.e.*, ferric complex and therefore underwent oxidation relative to the precursor **1**.

Mössbauer spectroscopy is another useful technique that allows for conclusions about the oxidation state of iron. The key quantities obtained from Mössbauer spectroscopy are the isomer shift δ and the quadrupole splitting ΔE_Q. The isomer shift depends on the electron density at the iron nucleus and can be related to the formal oxidation and spin state of iron [298]. The quadrupole splitting originates from the interaction of the nuclear quadrupole moment with the

electric field gradient. The electric field gradient tensor contains the second derivatives of the electrostatic potential with respect to the position and it arises from non-spherical charge distributions in the molecule [298]. For the precursor (see **Figure 7.4** right, top), two distinct species arise. The dominant species has an isomer shift of δ = 1.26 mm s^{-1} and a quadrupole splitting of ΔE_Q = 2.60 mm s^{-1}, which strongly indicate the presence of a ferrous high-spin complex with *S=2.* The minor species **1*** (δ = 0.48 mm s^{-1}, ΔE_Q = 0.85 mm s^{-1}) was attributed to the anhydrous dimeric form of the precursor.

For the measurements carried out 15 mins after mixing **1 + PivONH$_3$OTf**, two different distinct species **Int I** and **c1** arise. **Int I** is the dominant species and has δ = 0.55 mm s^{-1} and ΔE_Q = 0.66 mm s^{-1}. Compared to the precursor, **Int I** features a significantly decreased isomer shift and quadrupole splitting and therefore indicates the presence of an iron(III) high-spin site. Note that the iron(III) nature of this intermediate was predicted by EPR as well. As was indicated by the EPR already, the precursor is oxidized, presumably by **PivONH$_3$OTf**, upon formation of **Int I**. The minor component **c1** is attributed to another ferrous high-spin species that is distinct from **1** (δ = 1.42 mm s^{-1}, ΔE_Q = 3.36 mm s^{-1}). During the activation of the precursor, several side reactions are possible, such as ligand exchange with OTf$^-$ or CH_3OH.

Mössbauer spectra were also taken at 90 mins after mixing **1 + PivONH$_3$OTf**. The dominant species has δ = 0.57 mm s^{-1} and ΔE_Q = 0.40 mm s^{-1} and is again attributed to a high-spin iron(III) site. This species is assigned to the second intermediate **Int II**. Besides the remaining species **c1** that was present in the previously measured spectrum, another minor species **c2** (δ = 1.40 mm s^{-1}, ΔE_Q = 2.28 mm s^{-1}) with a relative intensity of 11 % arises, which indicates another ferrous side product. Interestingly, Mössbauer spectroscopy predicts a high-spin iron(III) site for **Int II**, while EPR indicates an integer total spin of the system. The sum formula of **Int II** is Fe(acac)$_2$(NH) and it is therefore suggested that the iron(III) site couples with the NH iminyl group to yield an overall integer spin state.

Further spectroscopic analyses were carried out by means of X-Ray absorption spectroscopy (XAS), resonance Raman (rR) spectroscopy and nuclear resonance vibrational spectroscopy (NRVS). The results are not discussed here, the interested reader is referred to ref. 283. To summarize the experimental data, the precursor **1 (**Fe(acac)$_2$(H$_2$O)$_2$) is a ferrous high-spin complex (*S=2*) that reacts with the aminating reagent **PivONH$_3$OTf** to give the first reaction intermediate **Int I** (Fe(acac)$_2$(PivONH)), which is a ferric high-spin complex (*S=2.5*). **Int II**

($Fe(acac)_2(NH)$) is formed by decomposition of **Int I**, it has an integer total spin but features a high-spin iron(III) site as well. **Int II** is the active species that reacts with the substrate styrene.

7.3 Computational protocol

All electronic-structure calculations presented in this chapter were carried out by the thesis author. The aim of the computations is to establish the preferred model for **1**, **Int I** and **Int II** as well as to propose a plausible reaction mechanism. For the precursor and both intermediates, various different conformers are obtained from conformer sampling with CREST at the GFN-xTB level of theory. The DFT calculations were carried out with ORCA 4.2.1. Test calculations to verify some of the results have been carried out with ORCA 5.0.3. All obtained conformers were optimized at the B3LYP-D3/def2-TZVP level of theory together with the RIJCOSX approximation and the appropriate auxiliary basis set. In addition, implicit solvation effects were included with the C-PCM model with gaussian charge scheme. Toluene was used as solvent. Equilibrium structures were proven to be local energy minima or first-order saddle points (for transition states) by analysis of vibrational frequencies. Final single point energies were obtained at the B3LYP-D3/def2-TZVP level of theory. For the most stable conformers with significant different structural features, several spectroscopic properties were computed. Excited state energies were computed using the time-dependent DFT (TD-DFT) approach [96,299]. Tight grids (Grid7 and GridX7 in ORCA 4 notation) were used. The spectra were simulated with the orca_mapspc program. A line width of 3500 cm^{-1} was used. For XAS, TD-DFT was used as well for computing the pre-edge excitations. Only the 1s orbital of iron was used as donor orbital while all virtual orbitals were used as acceptor space. In addition, computation of electronic quadrupole contributions was carried out. Also, the scalar-relativistic ZORA approximation [300,301] was used together with the appropriate recontracted def2-TZVP basis set. For the spectra, a line width of 0.8 eV was used. For analysis of the nature of the excitations, natural transition orbitals (NTOs) [302] and difference densities were used. The experimental UV-vis spectra were deconvoluted using the orca_asa program [303,304] by doing a least-square fit of gaussians. One gaussian was used for each resolved feature. For easier comparison of computed and measured UV-vis spectra, also the computed spectra were deconvoluted by a least-square fit. These deconvolutions of the computed spectra were carried out using the lmfit program package for the Python programming language. The rationale behind deconvoluting the computed spectra is that often many different excitations contribute to one overall resolved feature. The dimensionless oscillator strength f_{osc} is obtained from the area of the fitted gaussian [305].

$$f_{osc} = 4.32 \cdot 10^{-9} \frac{mol\ cm^2}{L} \int \varepsilon(\nu) d\nu. \quad (7.1)$$

$\varepsilon(\nu)$ is the molar absorption given in L mol^{-1} cm^{-1} and ν is the wave number in cm^{-1}. For the simulation of the resonance Raman spectra, orca_asa was used as well. The electronic excitation energies were obtained from TD-DFT calculations. In addition, computation of the ground state hessian (FREQ keyword) and the excited state gradients (*via* the NMGRAD keyword and `%rr` block in ORCA) was carried out. From the computed excited state gradients, ORCA estimates the adiabatic excitation energies, while TD-DFT provides only vertical excitation energies. The adiabatic excitation energies for the excitations of interest were eventually used for the simulation of the rR spectra. For the simulated spectra, a line width of 10 cm^{-1} was used. Mössbauer parameters were obtained at the B3LYP-D3/def2-TZVP level of theory as well. For their computation, the density at the nucleus and the electric field gradient must be computed. For iron, the CP(PPP) basis set for core properties [306,307] was used. While the quadrupole splitting ΔE_Q is given directly in the ORCA output, the isomer shift δ must be computed from the density at the nucleus ρ_o using eq. (7.2).

$$\delta = \alpha(\rho_o - C) + \beta. \quad (7.2)$$

α, C and β are empirical parameters that must be fitted for each combination of basis set and density functionals. The corresponding values for the B3LYP functional and CP(PPP) basis set were taken from ref. 308. Note that the standard deviations for the fits are around 0.1 mm s^{-1} [308], which determines the typical error bars associated with the isomer shifts. For locating transition states, initial relaxed surfaces scans or the nudged elastic band (NEB) method [309,310] were used together with the climbing image (CI) [311,312] algorithm.

In addition, CAS-SCF calculations were performed for selected systems, together with the N-electron valence perturbation theory to second order (NEVPT2) [313,314] energy correction. For **1** and **Int I**, ab initio ligand field theory (AI-LFT) [315] analyses were performed. For these computations, all plausible spin multiplicities and for each multiplicity all roots were included. For examples of input files, see Appendix A.3.

7.4 Results and discussion

7.4.1 The precursor ($Fe^{II}(acac)_2(H_2O)_2$)

The precursor exists in two different isomers, which are the *trans* and the *cis* isomer shown in **Figure 7.5**.

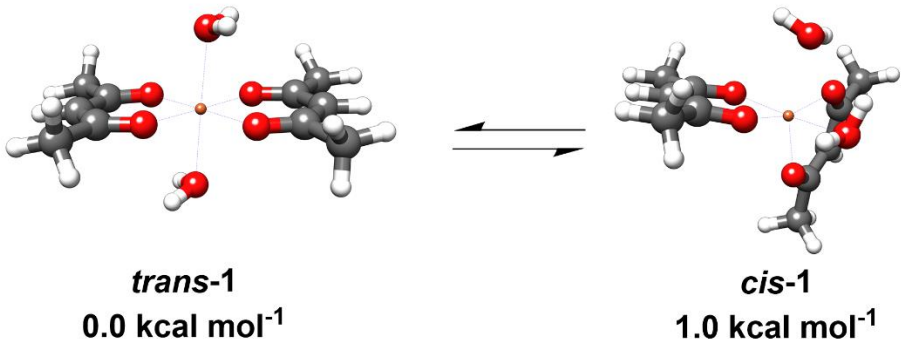

Figure 7.5 Equilibrium structures of *trans*-**1** and *cis*-**1** together with the relative Gibbs free energies. All energies are in kcal mol^{-1}. This figure was adapted from ref. [283].

In *trans*-**1**, both acetylacetonate ligands lie in one plane with the iron center while the water molecules are above and below the plane, respectively. The *trans*-**1** isomer is slightly more stable, however, existence of *cis*-**1** in solution is possible as well. To test the preferred spin state, the *trans* isomer was optimized with a total spin of $S=0,1,2.$ The corresponding relative free energies ΔG_{rel} are shown in **Table 7.1**.

Table 7.1 Relative free energies of *trans*-**1** optimized for different spin multiplicities at the B3LYP-D3/def2-TZVP level of theory. All energies are in kcal mol^{-1}. This table is adapted from ref. 283.

Total spin S	Spin multiplicity ($2S+1$)	ΔG_{rel}
0	1	28.138
1	3	15.354
2	5	0.000

Clearly, the quintet state is favored significantly over the other possible spin states, which is in agreement with the experimental Mössbauer parameters. The integer total spin of the system was also proven by EPR. In the further, only the quintet states of *trans*-**1** and *cis*-**1** are considered.

In **Figure 7.6**, the computed UV-vis spectra for both isomers are shown. The deconvoluted experimental and computed spectra are shown in Appendix A.

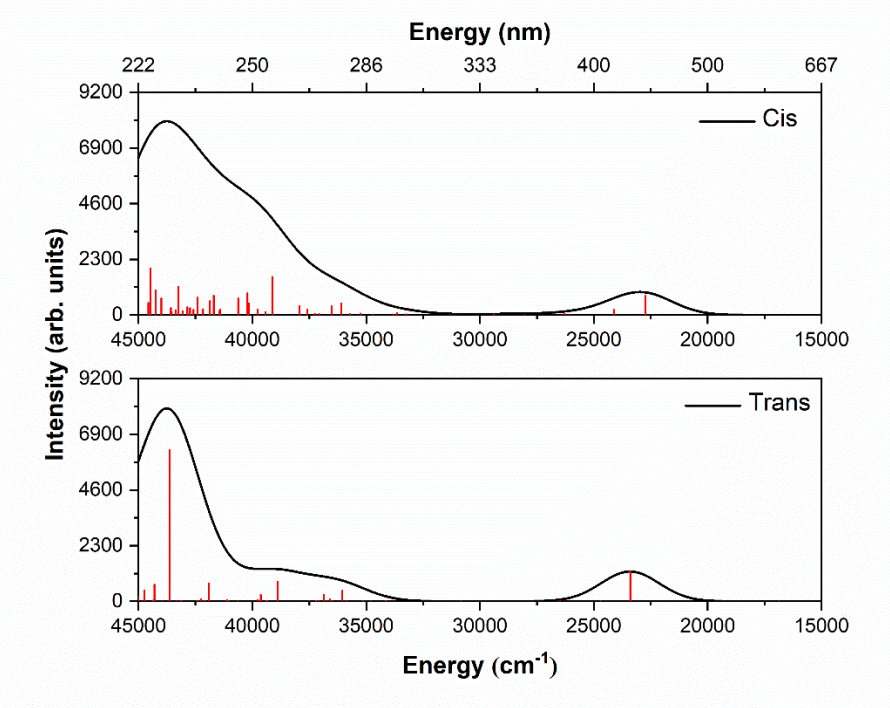

Figure 7.6 Computed UV-vis spectra of *cis*-**1** (top) and *trans*-**1** (bottom). Individual excitations are shown as red sticks. The intensity was scaled to match the experimental extinction coefficients. This figure was taken from ref. 283.

Comparison of experimental and computed excitation energies and oscillator strengths obtained from the deconvolution are given in **Table 7.2**.

Table 7.2 Experimental and computed excitation energies and oscillator strengths obtained from gaussian deconvolution. All energies are given in cm^{-1}. This table was adapted from ref. 283.

Excitation energy			Oscillator strength		
exp-**1**	*cis*-**1**	*trans*-**1**	exp-**1**	*cis*-**1**	*trans*-**1**
22806	23083	23410	0.044	0.047	0.069
28459	-	-	0.021	-	-
35702	39014	38598	0.173	0.215	0.188
41837	43881	43894	0.274	0.424	0.462

The low-energy band at 22806 cm^{-1} is reproduced very well by both isomers, with *cis*-**1** having slightly better agreement in both excitation energy and oscillator strength. Interestingly, the band at 28459 cm^{-1} is entirely missing in both computed spectra. However, as was shown by

experimental Mössbauer spectroscopy, there is a minor component **1*** with a relative intensity of 20 % which might be responsible for this band. This would be in accordance with the relatively small oscillator strength. For the high energy bands (35702 cm^{-1} and 41837 cm^{-1}), both isomers overestimate excitation energies and oscillator strengths. However, the experimental bands are relatively broad and overlap, which impedes the deconvolution.

Table 7.3 Experimental and computed isomer shifts and quadrupole splittings for the precursor **1**. All quantities are given in mm s^{-1}. This table was adapted from ref. 283.

Mössbauer parameter	exp-**1**	*cis*-**1**	*trans*-**1**
δ	1.26	1.05	1.03
ΔE_Q	2.60	2.59	2.47

The computed Mössbauer parameters are given in **Table 7.3**. The computed isomer shifts are much closer to each other than the typical maximum error of 0.1 mm s^{-1}. DFT slightly underestimates the isomer shift. However, the experimental isomer shift is already in the higher range of isomer shifts typical for iron [298], which may make it more difficult for DFT to predict them correctly. The quadrupole splitting parameters are in good agreement with experiment. However, in general, computed values for the quadrupole splitting are less reliable than the computed isomer shifts [316,317].

The basic idea behind rR spectroscopy is that the intensity of specific vibrational modes is increased dependent on the underlying electronic excitation that has been carried out. More specifically, only those modes' intensities are enhanced whose bonds or functional groups are affected by the underlying electronic excitation [318]. For **1**, the most striking band is the low-energy feature at 22806 cm^{-1}. For computing the rR spectra of *cis*-**1** and *trans*-**1**, the corresponding estimated adiabatic excitation energies were used.

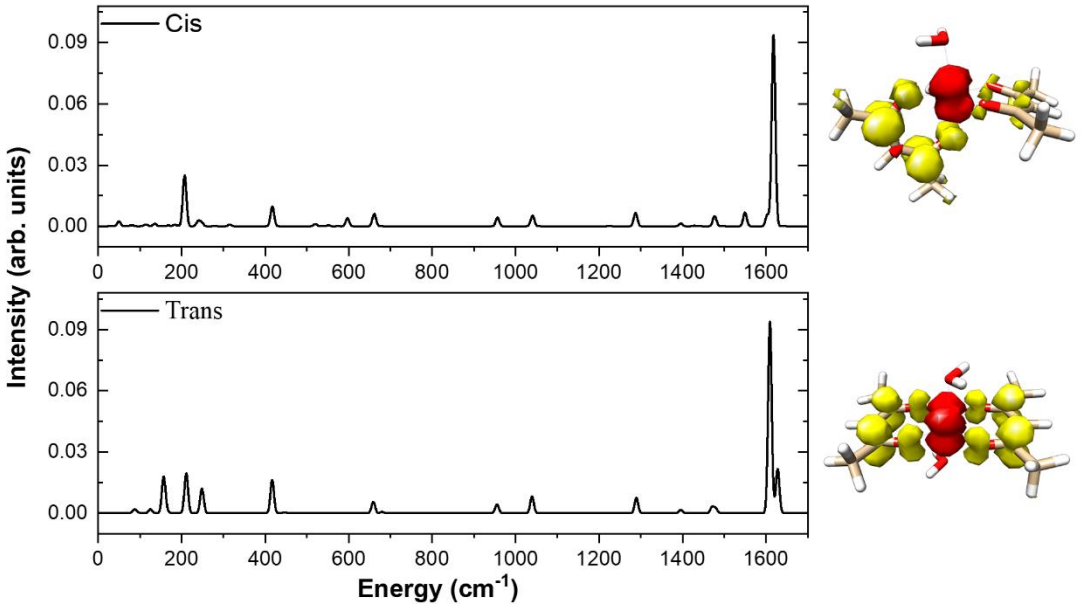

Figure 7.7 Computed rR spectra for *cis*-**1** (top) and *trans*-**1** (bottom). The computed adiabatic excitation energies were used for achieving resonance. For each spectrum, the difference density of the electronic excitation is shown. Red: donor region. Yellow: acceptor region. This figure was taken from ref. 283.

From **Figure 7.7**, it can be seen that the low-energy excitation belongs to the metal to ligand charge transfer excitation from the iron t_{2g} orbital into the empty π* orbitals of the acetylacetonate ligands. The by far most intense modes are the carbonyl stretching modes at around 1600 cm^{-1}. Other intense vibrational modes are the deformational $\delta(Fe\text{-}O_{acac})$ and the symmetric $\upsilon s(Fe\text{-}O_{acac})$ and $\upsilon(C\text{-}CH_3)$ stretching modes. Overall, *cis*-**1** has slightly better agreement with experiment then *trans*-**1**.

Table 7.4 Selected experimental and computed vibrational frequencies for **1**. All frequencies are in cm^{-1}. This table was adapted from ref. 283.

Assignment	exp-**1**	*cis*-**1**	*trans*-**1**
$\delta(Fe\text{-}O_{acac})$	265	256	262
$\upsilon s(Fe\text{-}O_{acac})$	450	420	417
$\upsilon(C\text{-}CH_3)$	1284	1289	1290
$\upsilon s(C{=}O)$	1625	1618	1610

The last spectroscopic technique discussed here is XAS. The focus is set on the pre-edge area, which is governed by the quadrupole-allowed 1s to 3d excitations into the singly occupied 3d-orbitals of iron.

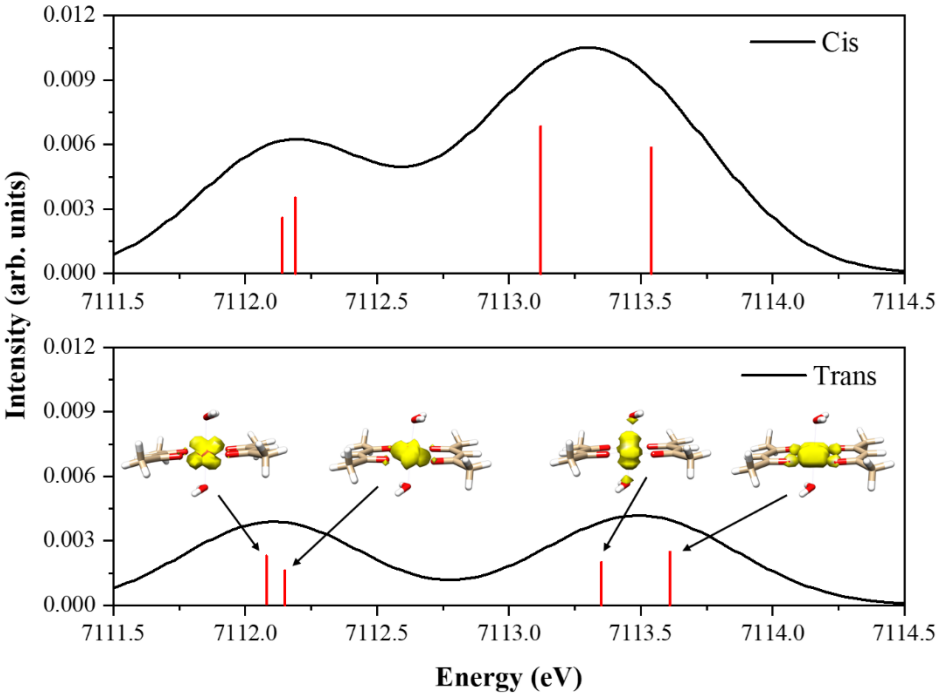

Figure 7.8 Computed XAS spectra for *cis*-**1** (top) and *trans*-**1** (bottom). The individual excitations are shown as red sticks. To match the experimental spectrum, the spectra were scaled and shifted accordingly. For *trans*-**1**, the difference densities of the excitations are given. The acceptor region is shown in yellow. The donor orbital is always the 1s orbital of iron. This figure was taken from ref. 283.

In **Figure 7.8**, the energetic separation between the excitations into the t_{2g} and e_g orbitals is clearly visible. Since the excitation energies within one group of orbitals (t_{2g} or e_g) are close to each other, only one resolved feature is observable for each group. For a more quantitative comparison, the intensity weighted average energy (IWAE) was computed for the experimental and computed spectra. The mean IWAE value of the computed spectra was used to calculate the shifting energy in order to match the experimental spectrum. Analogously, the mean value of the pre-edge areas (obtained by numerical integration of the spectra) of the computed spectra were used to scale the computed spectra to match the experimental one. The energy difference between the two resolvable peaks ΔE were computed from the energies of the corresponding intensity maxima.

Table 7.5 Computed and experimental IWAE, area and energy difference between the two intensity maxima ΔE of **1**. All energies are in eV. This table was adapted from ref. 283.

	exp-**1**	*cis*-**1**	*trans*-**1**
IWAE (eV)	7112.95	7112.84	7112.74
Area	25	42	19
ΔE (eV)	1.3	1.1	1.4

In general, *trans*-**1** has slightly better agreement with experiment than *cis*-**1** with respect to the computed area and the energy difference between the two peaks. It is noted, however, that the experimental energy resolution is around 0.2 eV. Therefore, also the expected experimental uncertainty is around 0.2 eV. The experimental spectrum shows almost identical intensities for both resolved features, which is reproduced more accurately by the *trans* isomer as well.

Taken all spectroscopic techniques together, no clear preference for one or the other computational model can be established. UV-vis and resonance Raman spectroscopy seem to slightly favor *cis*-**1**, while XAS favors *trans*-**1**. Mössbauer spectroscopy provides isomer shifts that are much closer to each other than the typical error bars and therefore does not allow for unique assignment as well. Free energy calculations slightly favor *trans*-**1** but overall, no clear assignment can be made. Given the small energy difference between the isomers, it is reasonable to assume that both isomers are in equilibrium in solution. As detailed in Appendix A.2, experimental determination of the d-d excitation energies could help distinguish between the two isomers. But that was beyond the scope of this project.

The electronic structure of *trans*-**1** is discussed in **Figure 7.9**, the main conclusions are valid for *cis*-**1** as well. Quasi-restricted orbitals [319] (QRO) were used for representation. Although *trans*-**1** belongs to the point group C_1, it can be approximated by C_{2v} point group symmetry. The highest occupied ligand orbitals are the doubly degenerate π-orbitals of the acetylacetonate ligands. Above them lies the iron d-manifold. The energetically lowest orbital (A_2) is doubly occupied. The two antibonding orbitals (A_1 and B_1) are clearly separated from the nonbonding orbitals as can be seen from the orbital energies. The covalent mixing is very small in both the nonbonding (5.9 % ligand contribution) and antibonding (10.5 % ligand contribution) orbitals. The lowest unoccupied orbitals of the system are the (again doubly degenerate) π*-orbitals of the acetylacetonate ligands. The low-energy excitation observed in UV-vis (22806 cm^{-1}) can be

assigned to the metal to ligand charge-transfer excitation from the doubly occupied metal orbital (A_2) into the π*-orbitals.

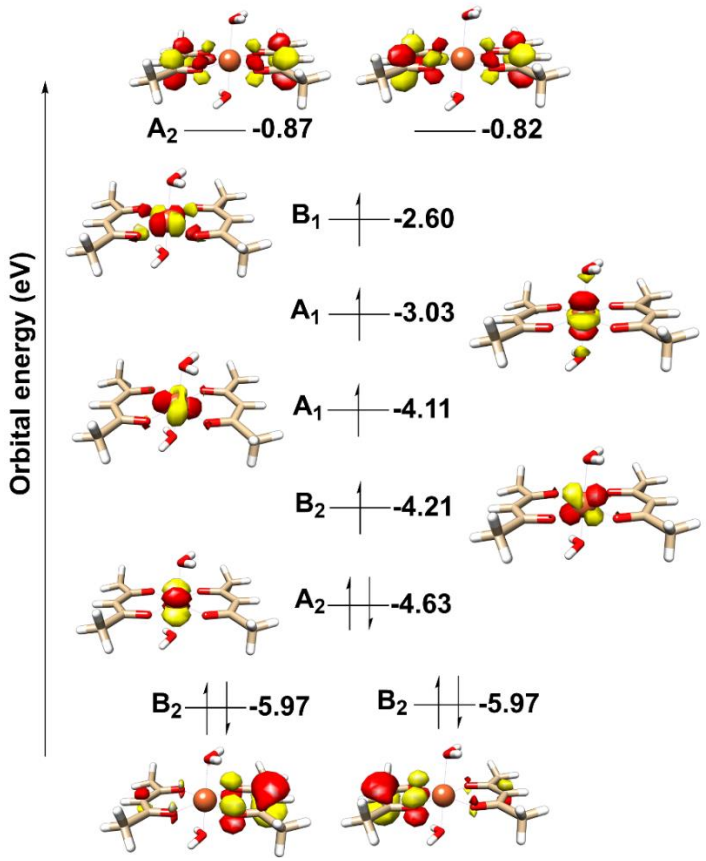

Figure 7.9 Electronic structure of *trans*-**1** represented by QROs. The QRO orbital energies are given together with the symmetry labels. For *trans*-**1**, C_{2v} symmetry is assumed. All energies are given in eV. This figure was taken from ref. 283.

7.4.2 Intermediate I ($Fe^{III}(acac)_2(NHOPiv)$)

As discussed in section 7.2, **Int I** is a half-integer spin system. The most stable isomer found assuming sextet spin multiplicity was re-optimized for the quartet and doublet spin state.

Table 7.6 Relative free energies of **Int I** optimized at different spin multiplicities at the B3LYP-D3/def2-TZVP level of theory. All energies are in kcal mol^{-1}. This table was adapted from ref. 283.

Total spin S	Spin multiplicity ($2S+1$)	ΔG_{rel}
1/2	2	12.764
3/2	4	10.370
5/2	6	0.000

The sextet state is clearly favored by free energy calculations in agreement with experimental EPR, XAS and Mössbauer spectroscopy. Two different geometric models of **Int I** were obtained from the conformer sampling procedure, which are the 6-coordinate model that features the Fe-N bond and an additional Fe-O bond with the keto group of the NHOPiv group (**Int I (FeN+FeO)**) and the 5-coodinate model that misses the additional Fe-O bond (**Int I (FeN)**). The 6-coordinate isomer of **Int I** clearly is favored by Gibbs free energy calculations.

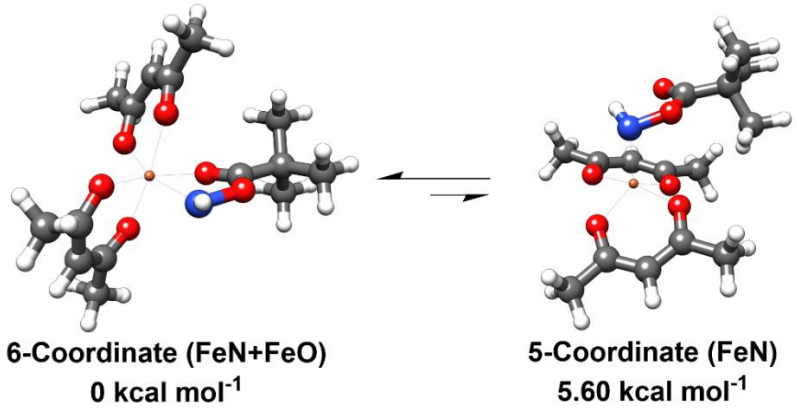

Figure 7.10 6-Coordinate and 5-coordinate models of the first intermediate **Int I** together with the relative free energies at the B3LYP-D3/def2-TZVP level of theory. All energies are in kcal mol^{-1}. This figure was adapted from ref. 283.

The computed and deconvoluted UV-vis spectra are shown in Appendix A.1. In **Table 7.7**, the obtained excitation energies and oscillator strengths are given.

Table 7.7 Experimental and computed excitation energies and oscillator strengths for **Int I** obtained from gaussian deconvolution. All energies are given in cm^{-1}. This table was adapted from ref. 283.

Excitation energy (**Int I**)			Oscillator strength (**Int I**)		
Exp.	FeN+FeO	FeN	Exp.	FeN+FeO	FeN
20656	18051, 21918	22651	0.040	0.043, 0.065	0.086
28018	25980	26560	0.009	0.043	0.082
33189	38501	36696	0.066	0.163	0.226
41070	41748	40594, 44213	0.180	0.562	0.192, 0.296

In the experimental UV-vis spectra, the most noticeable feature was the red-shift of the low-energy band for **Int I** compared to **1** (22806 cm^{-1} to 20656 cm^{-1}). This red-shift is reproduced

fairly well by 6-coordinate isomer, while the band energy for the 5-coordinate isomer remains mainly unchanged. However, for all other bands the agreement between computed and measured excitation energies is not overwhelmingly well. Noticeably, for both isomers, five gaussians were needed to fit to computed spectrum, while only four bands were extracted from experimental spectrum. However, the fit is of course not unique and changing the line width of the computed spectrum influences the number of resolvable bands.

The computed Mössbauer isomer shifts of both isomers match the experiment reasonably well. However, the experimental isomer shift lies practically in the middle between the two computed shifts, preventing unambiguous assignment. The computed quadrupole splittings are very close to each other but differ significantly from the experimental value. Therefore, Mössbauer spectroscopy does not allow for unique assignment to one or the other isomer.

Table 7.8 Experimental and computed isomer shifts and quadrupole splittings for **Int I**. All quantities are given in mm s^{-1}. This table was adapted from ref. 283.

Mössbauer parameter	Exp.	FeN+FeO	FeN
δ	0.55	0.61	0.52
ΔE_Q	0.66	1.43	1.33

In **Figure 7.11**, the computed rR spectra for **Int I** are shown. Again, the low-energy band (20656 cm^{-1} experimental excitation energy) was chosen as electronic excitation. The corresponding estimated adiabatic excitation energies were used in orca_asa. From the difference density plots, it can be seen that the low-energy band originates from a ligand to metal charge transfer excitation from the NHOPiv ligand into the singly occupied non-bonding iron d-orbitals. For **1**, the low-energy band originated from the metal to ligand charge transfer excitation. In **Table 7.9**, the most important experimental and computed vibrational frequencies together with the assignment carried out are shown. The computed spectra again feature very intense bands between 400 cm^{-1} and 450 cm^{-1} that are assigned to the υs(Fe-O_{acac}) stretching modes. The most intense band arises at 560 cm^{-1} for **Int I (FeN+FeO)** and at 629 cm^{-1} for **Int I (FeN)**. This band in both cases is assigned to the Fe-N stretching mode. For **Int I (FeN+FeO)**, this mode heavily mixes the Fe-N stretching and the Fe-O stretching due to the ring-like binding mode of the iron site with the O-pivaloyl group. The experimental frequency of this mode is 585 cm^{-1} and is therefore much better reproduced by **Int I (FeN+FeO)**.

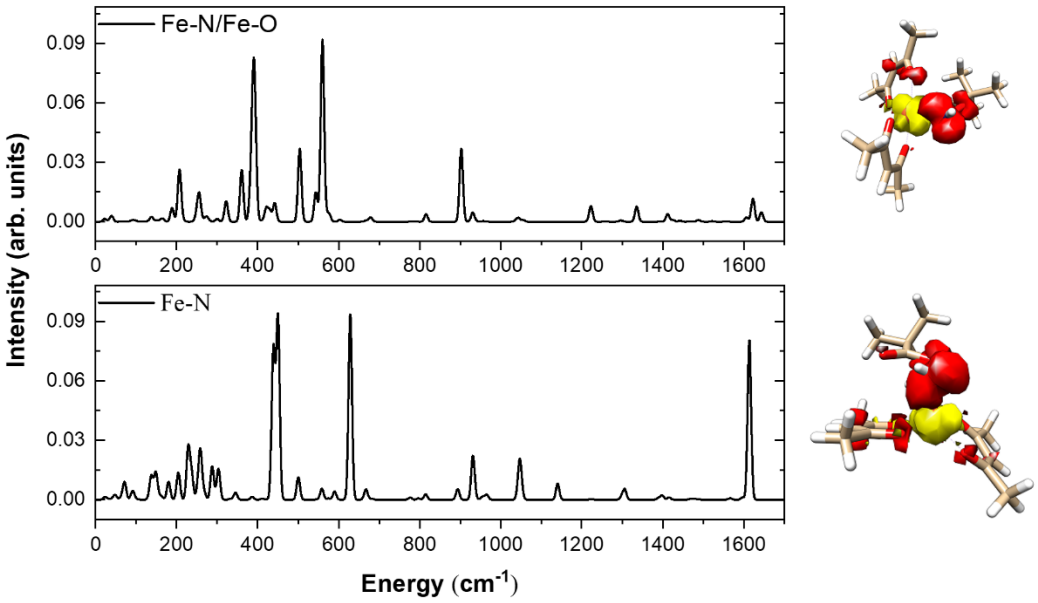

Figure 7.11 Computed rR spectra for **Int I (FeN+FeO)** (top) and **Int I (FeN)** (bottom). The computed adiabatic excitation energies were used for achieving resonance. For each spectrum, the difference density of the electronic excitation is shown. Red: donor region. Yellow: acceptor region. This figure was taken from ref. 283.

Table 7.9 Selected experimental and computed vibrational frequencies of **Int I**. All frequencies are in cm^{-1}. This table was adapted from ref. 283.

Assignment	Exp.	FeN+FeO	FeN
$\upsilon s(Fe\text{-}O_{acac})$	429	430	439
$\upsilon s(Fe\text{-}O_{acac})$	460	442	450
$\upsilon as(Fe\text{-}O_{acac})$	532	543	558
υs(Fe-N/Fe-O)	585	560	629
υ(N-O)	921	903	932
υas(C=O)	1586	1607	1597
υs(C=O)	1614	1623	1614

At 921 cm^{-1} there is the υ(N-O) stretching mode, which is slightly better reproduced by the 5-coordinate isomer. Just like UV-vis spectroscopy, rR spectroscopy is slightly favoring the 6-coordinate isomer, while Mössbauer spectroscopy is indifferent. In **Figure 7.12**, the computed XAS spectra are shown.

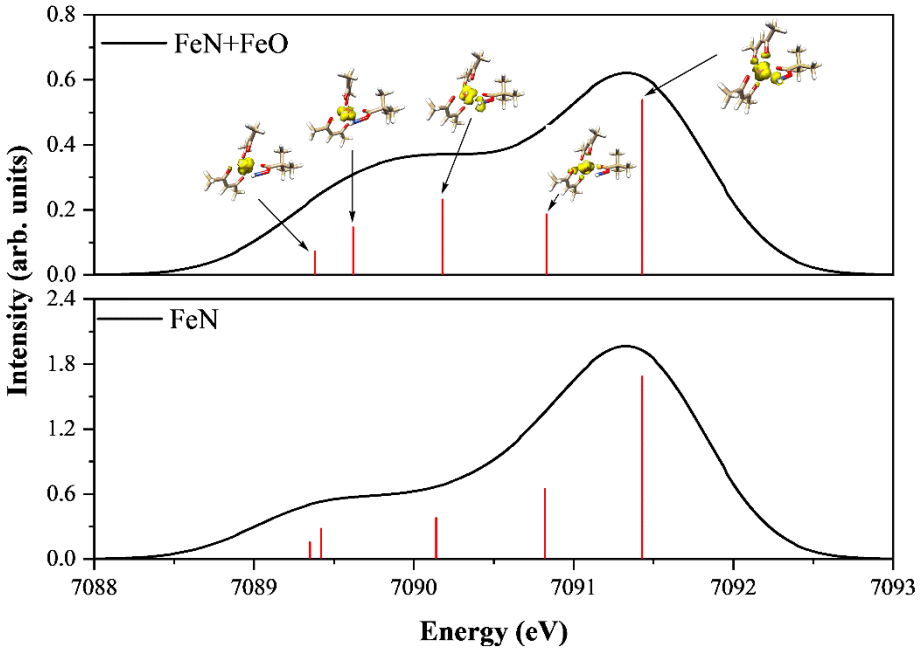

Figure 7.12 Computed XAS spectra for **Int I (FeN+FeO)** (top) and **Int I (FeN)** (bottom). The individual excitations are shown as red sticks. To match the experimental spectrum, the spectra were scaled and shifted accordingly. For **Int I (FeN+FeO)**, the difference densities of the excitations are given. The acceptor region is shown in yellow. The donor orbital is always the 1s orbital of iron. This figure was taken from ref. 283.

The computed spectra of **Int I** differ from the spectra of **1** in a couple of ways. First, no clear separation between the excitations into the non-bonding iron d-orbitals and the antibonding orbitals can be seen, so no assignment of the excitations into one of two groups is possible. Rather, it seems that the two lowest-in energy excitations constitute one resolvable feature, the two highest-in energy excitations constitute another resolvable band while the third excitation introduces a saddle-point shape between the two maxima to the spectrum. As expected, the excitations into the antibonding orbitals have higher intensities since the acceptor state orbital has significant contributions from the ligand's p-orbitals, making these excitations "less forbidden". In **Table 7.10**, the computed IWAE values, areas and ΔE values are shown. They were obtained in an analogous way as for **1**. Interestingly, **Int I (FeN+FeO)** has much better agreement with experiment for the computed pre-edge area and the energy difference between the two resolved features. Especially ΔE is reproduced very well by the computations. **Int I (FeN)** shows a significantly increased intensity caused by the reduced centrosymmetry of the

iron site due to the decreased coordination number. Also, the relative intensity of the two peaks is much better reproduced by the 6-coordinate isomer (see **Figure A 10**).

Table 7.10 Computed and experimental IWAE, area and energy difference between the two intensity maxima ΔE of **Int I**. All energies are in eV. This table was adapted from ref. 283.

	Exp.	FeN+FeO	FeN
IWAE (eV)	7112.88	7112.75	7112.87
Area	22	18	49
ΔE (eV)	1.3	1.5	1.9

Taken together, the correlation of the computed and experimental spectroscopic data together with the computed free energies of the 6-coordinate isomer and 5-coordinate isomer strongly suggest the presence of the 6-coordinate isomer **Int I (FeN+FeO)** in solution. The presence of the 6-coordinate isomer would also plausibly explain the loss of both water molecules upon activation of **1**, as was proven by ESI-MS. In accordance with experiment, the sextet state with five unpaired electrons is by far the most stable spin state. The first intermediate is therefore best described as high-spin Fe(III) complex.

7.4.3 Intermediate II ($Fe^{III}(acac)_2NH\cdot$)

The second intermediate **Int II** has the sum formula $Fe(acac)_2NH$ and is formed from the first intermediate by the loss of the OPiv group. Experimental XAS and Mössbauer spectroscopy indicate the presence of an iron(III) site, while EPR predicts an integer spin system. In this chapter, four different isomers of **Int II** are considered that are shown together with their relative free energies in **Figure 7.13**. The most stable isomer is TBP-N_{eq}, which has a distorted trigonal-bipyramidal environment of the iron site. The iminyl group is in equatorial position, meaning that the nitrogen site marks one of the corners of the pyramid base. The second most stable isomer is the square-pyramidal isomer with axial iminyl group (SQP-N_{ax}) with a relative free energy of 2.3 kcal mol^{-1}. The other two isomers are the trigonal-bipyramidal one with axial iminyl group (TBP-N_{ax}) and the square-pyramidal one with equatorial iminyl group (SQP-N_{eq}). Both are very similar in energy with relative free energies of 4.0 kcal mol^{-1} and 4.2 kcal mol^{-1}, respectively. As will be discussed later in this section, **Int II** is best described by an antiferromagnetically coupled diradical, which in turn is best described by a multireference approach. It is noted here that the energetic order of the four isomers was not only computed at the DFT level of theory but also at the DLPNO-CCSD(T) and the CAS-SCF + NEVPT2 levels

of theory. While the relative free energies change dependent on the underlying method used, the order of the isomers was the same regardless of the methodology.

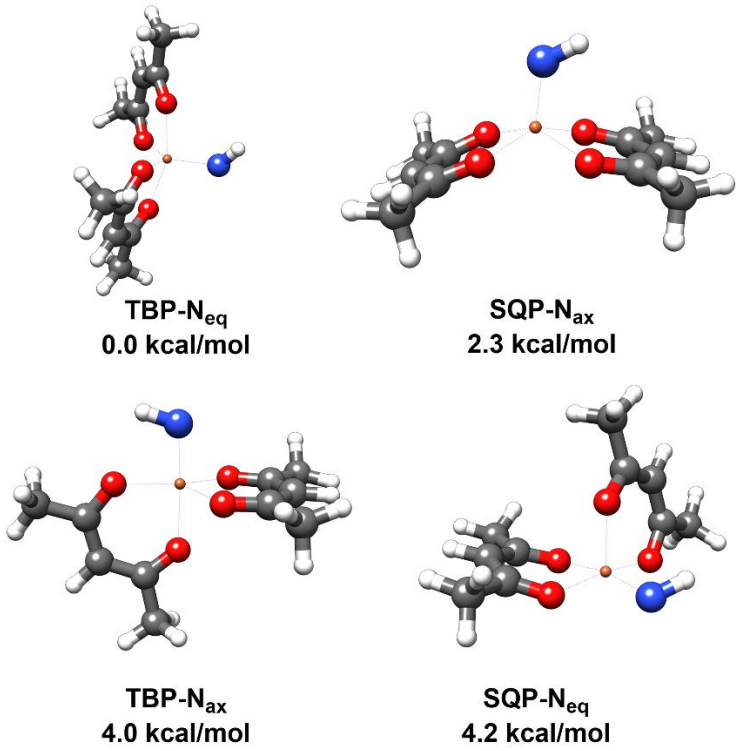

Figure 7.13 Isomers of **Int II** together with the relative free energies at the B3LYP-D3/def2-TZVP level of theory. TBP-N_{eq}: trigonal-bipyramidal coordination with equatorial iminyl group, SQP-N_{ax}: square-pyramidal coordination with axial iminyl group, TBP-N_{ax}: trigonal-bipyramidal coordination with axial iminyl group, SQP-N_{eq}: square-pyramidal coordination with equatorial iminyl group. All isomers were optimized as quintet states ($S=2$). All energies are in kcal mol^{-1}. This figure was adapted from ref. 283.

Like for the precursor **1** and the first intermediate **Int I**, the most stable isomer was re-optimized using different spin multiplicities of the system. The relative free energies of different spin multiplicities are given in **Table 7.11**. Interestingly, the different spin states are much closer in energy to each other compared to **1** or **Int I**. This is especially true for the quintet and the heptet state. However, CAS-SCF + NEVPT2 calculations predict the quintet state to be more stable as well. Therefore, **Int II** is best described as integer spin system with a total spin of $S=2$.

Table 7.11 Relative free energies of **Int II** optimized at different spin multiplicities at the B3LYP-D3/def2-TZVP level of theory. All energies are in kcal mol^{-1}. This table was adapted from ref. 283.

Total spin S	Spin multiplicity ($2S+1$)	ΔG_{rel}
0	1	13.900
1	3	8.089
2	5	0.000
3	7	4.101

In the following, the spectroscopic properties of the isomers are discussed in brief before the electronic structure of **Int II** will be elucidated in greater detail. The experimental and computed excitation energies and oscillator strengths obtained from the deconvolution of the spectra as discussed in section 7.3 are shown in **Table 7.12**.

Table 7.12 Experimental and computed excitation energies and oscillator strengths for **Int II** obtained from gaussian deconvolution. All energies are given in cm^{-1}. This table was adapted from ref. 283.

Excitation energy (**Int II**)					Oscillator strength (**Int II**)				
Exp.	TBP-N_{eq}	SQP-N_{ax}	TBP-N_{ax}	SQP-N_{eq}	Exp.	TBP-N_{eq}	SQP-N_{ax}	TBP-N_{ax}	SQP-N_{eq}
14342	21487	19000	24108	17998	0.019	0.010	0.017	0.020	0.006
27128	28149	26104	31606	29451	0.021	0.091	0.046	0.112	0.112
34606	34938	30833	36799	37784	0.112	0.112	0.065	0.177	0.176
41339	39197	39646	40987	40907	0.136	0.267	0.174	0.130	0.125
-	43429	41990	46335	47151	-	0.408	0.551	0.657	0.524

Interestingly, none of the four isomers is capable of reproducing the low-energy feature at 14342 cm^{-1}. All isomers overestimate the corresponding excitation energy significantly and underestimate the oscillator strength. This excitation is expected to be a charge-transfer excitation involving the metal d-orbitals and the iminyl group. However, as will be discussed later in this section, the electronic structure of **Int II** cannot fully adequately be described by DFT. Therefore, also the spectroscopic properties computed at the DFT level most likely will not be obtained with the same accuracy that was experienced for **1** and **Int I**. Test calculations

using various different density functionals were carried out, but no significant improvement in either excitation energy or oscillator strength was found for the low-energy excitation. The fitting procedure was carried out up to 45000 cm^{-1}. The deconvolutions of the computed spectra include a fifth gaussian with excitation energies between 42000 cm^{-1} and 48000 cm^{-1} that was not captured by experiment. However, this part of the spectrum is governed by ligand-to-ligand excitations like π to π^* excitations of the acetylacetonate ligands, which are not of special interest for exploring the electronic structure of **Int II**.

Table 7.13 Experimental and computed isomer shifts and quadrupole splittings for **Int II**. All quantities are given in mm s^{-1}. This table was adapted from ref. 283.

Mössbauer parameter	δ	ΔE_Q
TBP-N_{eq}	0.41	1.46
SQP-N_{ax}	0.47	0.78
SQP-N_{eq}	0.41	1.66
TBP-N_{ax}	0.41	2.07
Exp.	0.57	0.40

Interestingly, also the computed isomer shifts show relatively large discrepancies compared to experiment. Only the isomer shift of SQP-N_{ax} is within the expected error range of 0.1 mm s^{-1}. The other three isomers show slightly smaller isomer shifts. The overall underestimation of δ at the DFT level of theory may be induced by the blurring of the oxidation state, leading to a formal oxidation state between +3 and + 4.

As will be discussed below, DFT predicts different electronic structures for the four isomers of **Int II**, as it is not entirely capable of dealing with the multireference character of the system. This can be seen beautifully from the resonance Raman spectra. As excitation, the low-energy excitation with the largest oscillator strength was chosen. While there was one dominant low-energy excitation for **1** and **Int I**, the different isomers of **Int II** show only very low intensity excitations around 20000 cm^{-1} with very different difference densities. Interestingly, for TBP-N_{eq}, a ligand to metal charge transfer excitation occurs, in which the main density donor is the nitrogen p-orbital. For SQP-N_{ax}, excitation from the acetylacetonate orbitals into the Fe-N σ^* orbital happens. For SQP-N_{eq} and TBP-N_{ax}, ligand to metal excitations from the nitrogen site into the metal d-orbitals occur. As expected, intensity enhancement of the C=O carbonyl

stretching modes is only observed for TBP-N_{eq} and SQP-N_{ax}, since the corresponding excitations induce changes in the electron densities at the carbonyl groups.

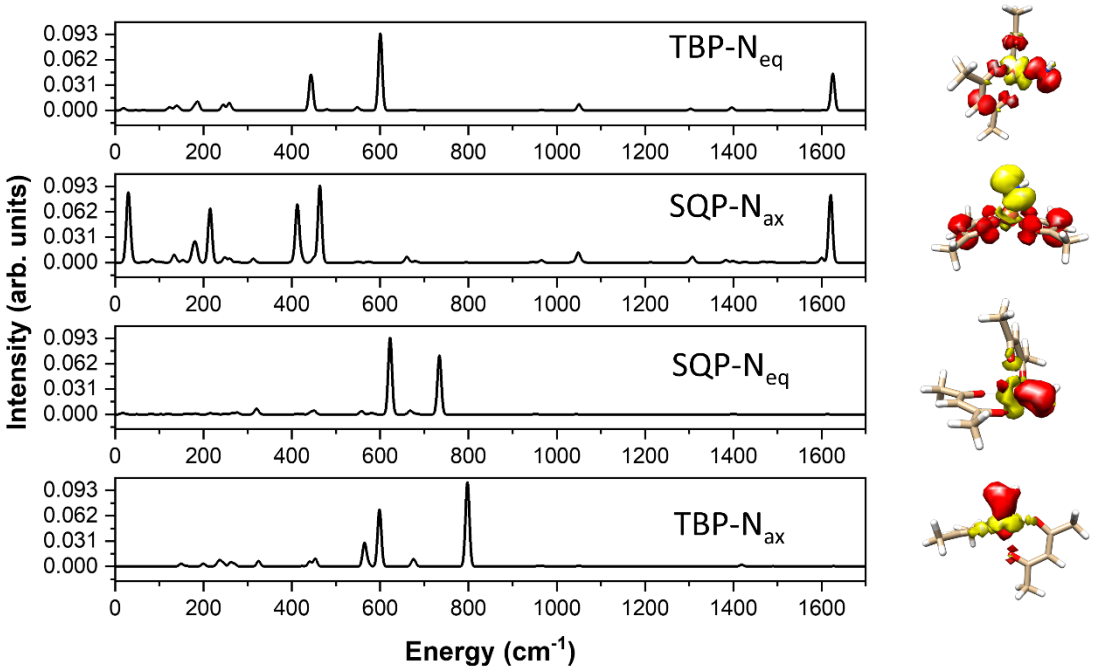

Figure 7.14 Computed rR spectra for **Int II.** The computed adiabatic excitation energies were used for achieving resonance. For each spectrum, the difference density of the electronic excitation is shown. Red: donor region. Yellow: acceptor region. This figure was adapted from ref. 283.

Some important vibrational excitation energies together with the assignment are summarized in **Table 7.14**.

Table 7.14 Selected experimental and computed vibrational frequencies for **Int II**. All frequencies are in cm^{-1}. This table was adapted from ref. 283.

Assignment	Exp.	TBP-N_{eq}	SQP-N_{ax}	TBP-N_{ax}	SQP-N_{eq}
υs(Fe-O_{acac})	433	442/444	448	440/453	443/451
δ(N-H)	462	480	465	798	734
υas(Fe-O_{acac})	528	548/559	549/560	564/574	557
υ(Fe-N)	612	600	655	598	623
δ(C-H)	802	793/795	795/800	794/797	796/797

The most intense vibrational modes in the computed rR spectra are, as expected, the ones that include the iron and nitrogen site. The iron-oxygen stretching modes υs(Fe-O_{acac}) and υas(Fe-

O_{acac}) are systematically overestimated by DFT. Very interesting is the position of the δ(N-H) deformational mode. While the experimental frequency is fairly well reproduced by TBP-N_{eq} and SQP-N_{ax}, the TBP-N_{ax} and SQP-N_{eq} isomers overestimate this mode by more than 250 cm^{-1}. That the mode with experimental frequency of 462 cm^{-1} is indeed the δ(N-H) mode was verified by ^{15}N-labeled rR spectroscopy, which predicts an isomer shift of -4.0 cm^{-1} that is very well reproduced by TBP-N_{eq} (-6.1 cm^{-1}). The experimental band at 802 cm^{-1} was assigned to δ(C-H) deformational modes of the acetylacetonate ligands. The υ(Fe-N) stretching mode was seen in experiment at 612 cm^{-1}, which is reproduced by the trigonal-bipyramidal isomers, while the square-pyramidal isomers overestimate the frequency. For this mode, ^{15}N isomer shifts indicate the participance of nitrogen as well. The υas(Fe-O_{acac}) mode shows, as expected, no ^{15}N isomer shift in experiment. For TBP-N_{ax}, the Fe-O stretches mix with the Fe-N stretches, leading to a computed ^{15}N isomer shift of -6.98 cm^{-1}. By far the best overall agreement with experiment has therefore the TBP-N_{eq} isomer.

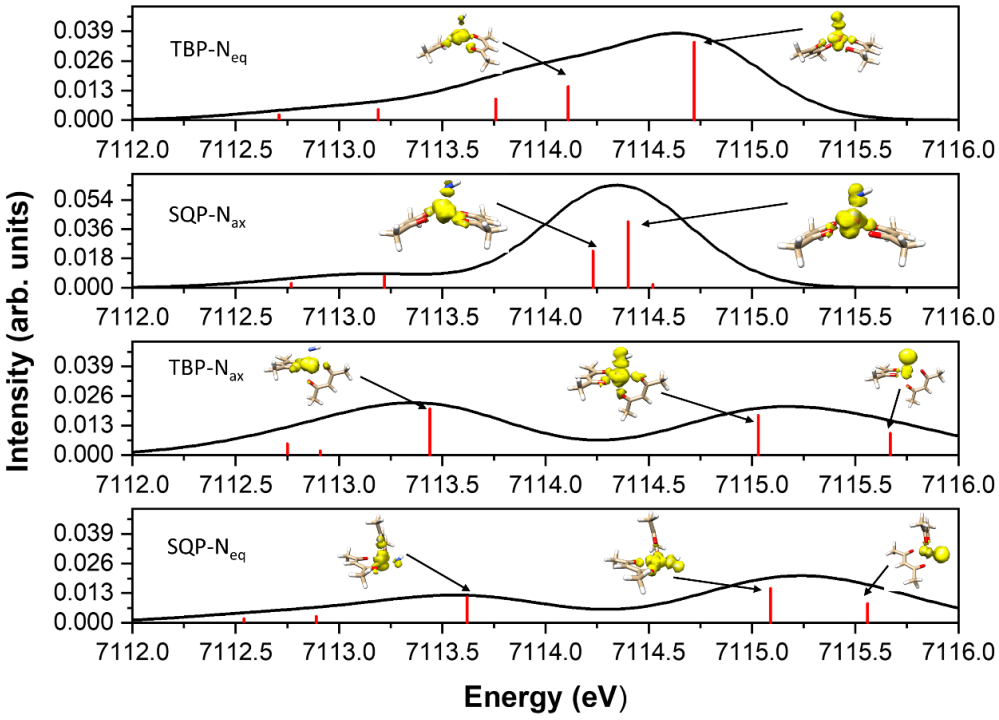

Figure 7.15 Computed XAS spectra for **Int II**. The individual excitations are shown as red sticks. To match the experimental spectrum, the spectra were scaled and shifted accordingly. The difference densities of the most intense excitations are given. The acceptor region is shown in yellow. The donor orbital is always the 1s orbital of iron. This figure was adapted from ref. 283.

In **Figure 7.15**, the computed XAS spectra for the four isomers of **Int II** are shown. Following the procedure outlined for the precursor **1**, the spectra were shifted and scaled to match the experimental spectra. Noticeably, for TBP-N_{eq} and SQP-N_{ax}, the shape of the spectra differs significantly from the shape of the experimental spectrum. For TBP-N_{eq}, only one peak is resolved, while for SQP-N_{ax} the intensity distribution of the two peaks does not match the experimental spectrum (in experiment, both peaks roughly have the same intensity, see **Figure A 10**). In general, it is seen that excitations into empty orbitals with significant contributions from the nitrogen orbitals have increased intensity. However, it should be kept in mind that the unambiguous electronic structure at the B3LYP-D3 level of theory of course influences the XAS spectra since these depend directly on the occupation numbers of the frontier orbitals. The energy difference between the two peaks of the SQP-N_{ax} isomer amounts to 1.2 eV and matches the experiment better than the other isomers.

Taking everything together, for **Int II** the quintet state is clearly the favored spin state. From the analysis of the computed and experimental spectroscopic data, no clear assignment to one or the other isomer can be made. TBP-N_{eq} is favored by free energies, while SQP-N_{ax} has a relative free energy of only 2.3 kcal mol^{-1}. TBP-N_{ax} and SQP-N_{eq} have higher relative free energies of around 4 kcal mol^{-1}. The relative energetic order of the four isomers was also corroborated by DLPNO-CCSD(T) and CAS-SCF/NEVPT2 calculations. While vibrational spectroscopy favors TBP-N_{eq} as well, Mössbauer, UV-vis and XAS spectroscopy favor SQP-N_{ax}. Like for the precursor **1**, equilibrium between TBP-N_{eq} and SQP-N_{ax} might be possible in solution at room temperature.

As was seen before, the energetically favored isomers (TBP-N_{eq} and SQP-N_{ax}) are also favored by correlation of the spectroscopic data. During the calculations, it was found that DFT predicts a large negative spin population for the nitrogen site of TBP-N_{eq}, indicating a partial radical character of the iminyl group. Therefore, broken-symmetry DFT calculations were carried out for TBP-N_{eq} and SQP-N_{ax} in order to compute the coupling constant between the iron site and the iminyl radical. For the iron site, five unpaired electrons and one unpaired electron for the iminyl group were chosen, leading to a quintet state. The coupling constant J is computed from [320,321]

$$J = -\frac{E[HS] - E[BS]}{\langle S^2 \rangle_{HS} - \langle S^2 \rangle_{BS}}. \tag{7.3}$$

$E[HS]$ is the energy of the high-spin state, $E[BS]$ is the corresponding energy of the broken-symmetry state and $\langle S^2 \rangle$ are the corresponding expectation values of S^2. For the coupling between the iron site and the iminyl radical, very negative coupling constants of -524.41 cm^{-1} (TBP-N_{eq}) and -893.41 cm^{-1} (SQP-N_{ax}) were found, indicating the antiferromagnetic orientation of the spin centers.

However, DFT predicts vastly different electronic structures for the four different isomers. In general, the Fe-N bond can be described with one σ/σ* orbital pair and one π/π* orbital pair. For all isomers, the σ and π bonding orbitals are occupied with two electrons. For TBP-N_{eq}, however, the σ* orbital is singly occupied while the π* orbital is unoccupied. For the other three isomers, the situation is reversed with the π* orbital being singly occupied and the σ* orbital being unoccupied. The different electron configurations consequently lead to different partial charges and spin occupations, as was verified from the Löwdin and NBO population analysis schemes. If the system ought to be described by a multireference (*i.e.,* CAS-SCF) wave function, the minimum orbital space should cover the Fe-N σ/σ* and π/π* orbital pairs as well as the two non-bonding d-orbitals and the σ* antibonding orbital with the acetylacetonate ligands, leading to a CAS(8,7) space. That these orbitals are indeed the most important ones was also verified by the natural occupation numbers of the MP2 natural orbitals, which differed the most from 0 or 2 for these orbitals. For a more accurate description, the active space was extended by the highest occupied ligand orbitals (two degenerate π orbitals of the acetylacetonate ligands) and the corresponding empty antibonding orbitals, leading in total to a CAS(12,11) space. The starting orbitals for the CAS-SCF calculations are the QROs obtained from DFT. The electronic structure of SQP-N_{ax} represented by the active orbitals of the CAS(12,11) calculation is shown in **Figure 7.16**. Analogous electronic structures result for the other isomers. In the CAS-SCF calculations, only one root was chosen in order to optimize the wave function for the ground state. The lowest-in energy orbital is the Fe-N σ bonding orbital, followed by the two degenerate π orbitals of the acetylacetonate ligands. These three orbitals have occupation numbers very close to 2. Surprisingly, the Fe-N π orbital has an occupation number of 1.42 and therefore cannot be considered doubly occupied. The corresponding antibonding π* orbital is the highest partially occupied orbital of the system and has an occupation number of 0.58. The π orbital contains larger contribution from the iron site, while the π* orbital is slightly more nitrogen-based. Between the π/π* orbital pair lie the singly occupied and mainly iron-based orbitals, which are two non-bonding d-orbitals and the Fe-N σ* and Fe-O_{acac} σ* antibonding orbitals.

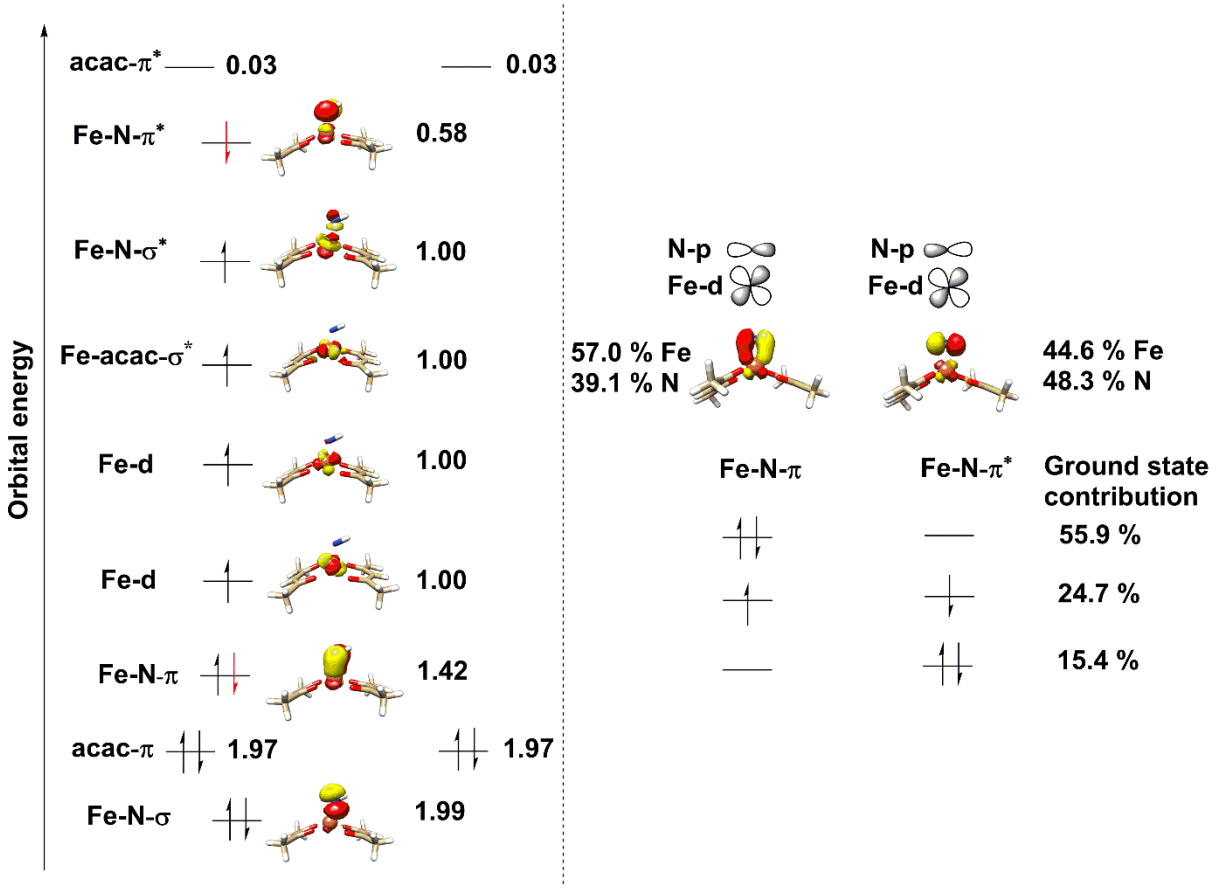

Figure 7.16 Electronic structure of **Int II** (SQP-N_{ax}) at the CAS(12,11)/def2-TZVP level of theory. Left: the active orbitals are shown together with the natural occupation numbers. Right: individual states with most significant contributions to the ground state wave function. This figure was taken from ref. 283.

Interestingly, the Fe-N σ* orbital is always singly occupied, irrespective of the isomer under investigation. The π/π* orbital pair is always occupied with two electrons, however, three different states contribute to the overall ground state, as shown in **Figure 7.16**, right panel. These three states describe the distribution of the two electrons onto the orbital pair. The largest contribution has the state with both electrons occupying the π orbital. However, the other states with one or two electrons occupying the π* orbital have significant contributions to the overall ground state as well. This observed distribution of the two electrons onto the orbitals of the corresponding orbital pair is responsible for the observed negative spin population at nitrogen and therefore the diradical character of the molecule. The predicted electronic structure could be reproduced by using MP2 natural orbitals instead of QROs as starting orbitals. In contrast to DFT, CAS(12,11) predicts very similar electronic structures for all four isomers, as can be seen from Löwdin atomic charges and spin populations given in **Table 7.15**.

Table 7.15 Löwdin atomic charges and spin populations at the CAS(12,11) level of theory for the isomers of **Int II**. Atomic charges are given in units of the elementary charge. This table is adapted from ref. 283.

	Atomic charge (e)		Spin population	
	Fe	N	Fe	N
TBP-N_{eq}	0.75	-0.38	4.31	-0.45
SQP-N_{ax}	0.71	-0.34	4.24	-0.39
TBP-N_{ax}	0.68	-0.33	4.19	-0.34
SQP-N_{eq}	0.68	-0.33	4.22	-0.36

Of course, the Löwdin scheme for portioning the density matrix is just one of many possible schemes, however, it clearly demonstrates that CAS(12,11) predicts very similar electronic structures for the four different isomers.

For further analyzing the CAS-SCF wave function, local spin analysis [322] was carried out for the TBP-N_{eq} and SQP-N_{ax} isomers. For TBP-N_{eq} an effective spin of 0.58 ħ (1.16 unpaired electrons) for the NH subsystem and an effective spin of 2.39 ħ (4.78 unpaired electrons) for the residue was obtained. For SQP-N_{ax} analogous values arise. Taking together the experimental findings and the analysis of the CAS-SCF wave function, **Int II** is best described as iron(III) site that couples antiferromagnetically with the iminyl radical to an overall quintet ground state. DFT blurs the oxidation state of iron, which presumably is responsible for the worse agreement between measured and computed spectroscopic parameters in comparison to **1** and **Int I**.

In **Figure 7.17a,** the spin density obtained from the CAS(12,11) ground state calculation is shown, illustrating the antiferromagnetic radical coupling. Another DFT-based tool for diagnosing multireference systems is the Fractional Occupation Number Weighted Electron Density (FOD) that is obtained from finite temperature DFT calculations [323]. The finite temperature leads to fractional occupation numbers, allowing the partial occupation of unoccupied orbitals with low energy. Plotting the FOD allows for visualization of the spatial origin of static correlation. In **Figure 7.17b,** the FOD plot of the TBP-N_{eq} isomer is shown that qualitatively agrees with the computed spin density at the CAS(12,11) level of theory. The iron site has smaller contributions to the FOD compared to the spin density, which is because all

singly occupied d-orbitals contribute to the spin density, while only the "hot" orbitals contribute to the FOD.

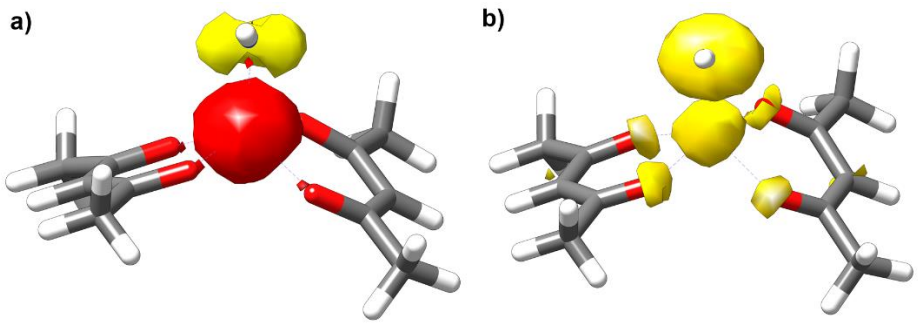

Figure 7.17 a) spin density at the CAS(12,11) and b) FOD at the B3LYP-D3 level of theory for the TBP-N_{eq} isomer. For obtaining the FOD, a smearing temperature of 9000 K was used.

The computed Fe-N Mayer bond order [324] at the CAS(12,11) level is around 1 for all isomers, illustrating the pronounced single-bond character of the Fe-N bond. The single bond character is also supported by relatively long Fe-N distances (1.85 Å for TBP-N_{eq} at the B3LYP-D3, 1.90 Å at the CAS(12,11) level of theory). For the other isomers, DFT predicts Fe-N bond distances of around 1.75 Å. This is in striking contrast to the much shorter Fe-N bond distances found in previously described iron-imido complexes. For example, Holland and coworkers [325] investigated (L)Fe^{III}=NAd (L=ligands, Ad=adamantyl) complexes and found a Fe=N bond distance of 1.67 Å. For the corresponding (L)Fe^{II}-NHAd complex, they found a Fe-N bond distance of 1.84 Å, reflecting the pronounced single bond character in the ferrous complex. In addition, Iovan et al. [326] studied (L)Fe^{III}Cl-NAr complexes and proposed an antiferromagnetic diradical coupling analogous to the coupling scheme proposed in this chapter. They found an elongated Fe-N bond distance of 1.77 Å. The high efficiency of **Int II** in nitrogen-transfer reactions most probably is a direct consequence of the pronounced Fe-N single bond, facilitating the Fe-N bond breaking compared to iron imido complexes that show significant double bond character. Betley and coworkers [327] obtained activation energies for the hydrogen abstraction reaction of iron(III)-iminyl complexes with toluene. They tested *N*-aryl and *N*-alkyl substituted iminyl complexes and found that *N*-aryl complexes have higher activation energies and therefore decreased reactivity, which was attributed to the stabilization of the *N*-aryl compound by delocalization of spin density. **Int II** does not feature any ligands at the nitrogen site, which consequently prevents delocalization of spin density. This weakens the Fe-N bond further and therefore increases reactivity of **Int II**.

7.4.4 Reaction mechanism

Activation process

The first step of the reaction is the formation of the first intermediate **Int I** from the precursor **1**. The precursor **1** is a high-spin ferrous complex, while **Int I** is a high-spin ferric complex. Therefore, **1** is oxidized upon activation. In analogy to the well-established peroxide chemistry [328-330], it is assumed that **1** is oxidized by the aminating reagent **PivONH$_3$OTf** in a sacrificial process. **PivONH$_3$OTf** therefore plays a dual role in this reaction. First, it oxidizes **1** under decomposition. Second, it was shown that **Int I** contains the PivONH ligand, so it is acting as ligand as well. This proposed activation mechanism would require that two **PivONH$_3$OTf** molecules are involved in the formation of one **Int I** molecule. This is in agreement with experimental UV-vis spectroscopy that shows that around 2 equivalents (equiv.) are needed for obtaining maximum absorption for **Int I**. In addition, stopped-flow UV-vis measurements of the reaction of **1** with **PivONH$_3$OTf** yield a reaction order of 1 with respect to the aminating reagent and a reaction order of 0.5 with respect to the precursor. The overall activation process can be formulated in two steps. First, there is the oxidation.

$Fe(acac)_2(H_2O)_2$ + *t*-$BuCOONH_3^+OTf^-$ → $[Fe(acac)_2(H_2O)_2]^+$ + *t*-BuCOO· + NH_3 + OTf^-.

Iron(II) is oxidized to iron(III) under formation of ammonia and the pivaloxy radical *t*-BuCOO·. In the second step, the coordination with a second aminating reagent molecule takes place.

$[Fe(acac)_2(H_2O)_2]^+$ + *t*-$BuCOONH_3^+OTf^-$ → $Fe(acac)_2$*t*-BuCOONH + 2 H_2O + 2 H^+ + OTf^-.

The combined reaction might be written as

$Fe(acac)_2(H_2O)_2$ + 2 *t*-$BuCOONH_3^+OTf^-$ → $Fe(acac)_2$*t*-BuCOONH + *t*-BuCOO· + NH_4OTf + HOTf + 2 H_2O.

Overall, the reaction medium should be slightly acidic because one of the **PivONH$_3$OTf** molecules loses two protons upon coordination. Since many species are involved, a complicated system of acid-base equilibria may occur. However, for the given reaction equation, a free reaction energy ΔG_R of -7.0 kcal mol^{-1} was computed.

Reaction profile

The activation of the precursor is followed by decomposition of **Int I** to give **Int II** and the subsequent reaction of **Int II** with styrene **3** to give the desired amino ether **P**. The computed reaction profile is given in **Figure 7.18**.

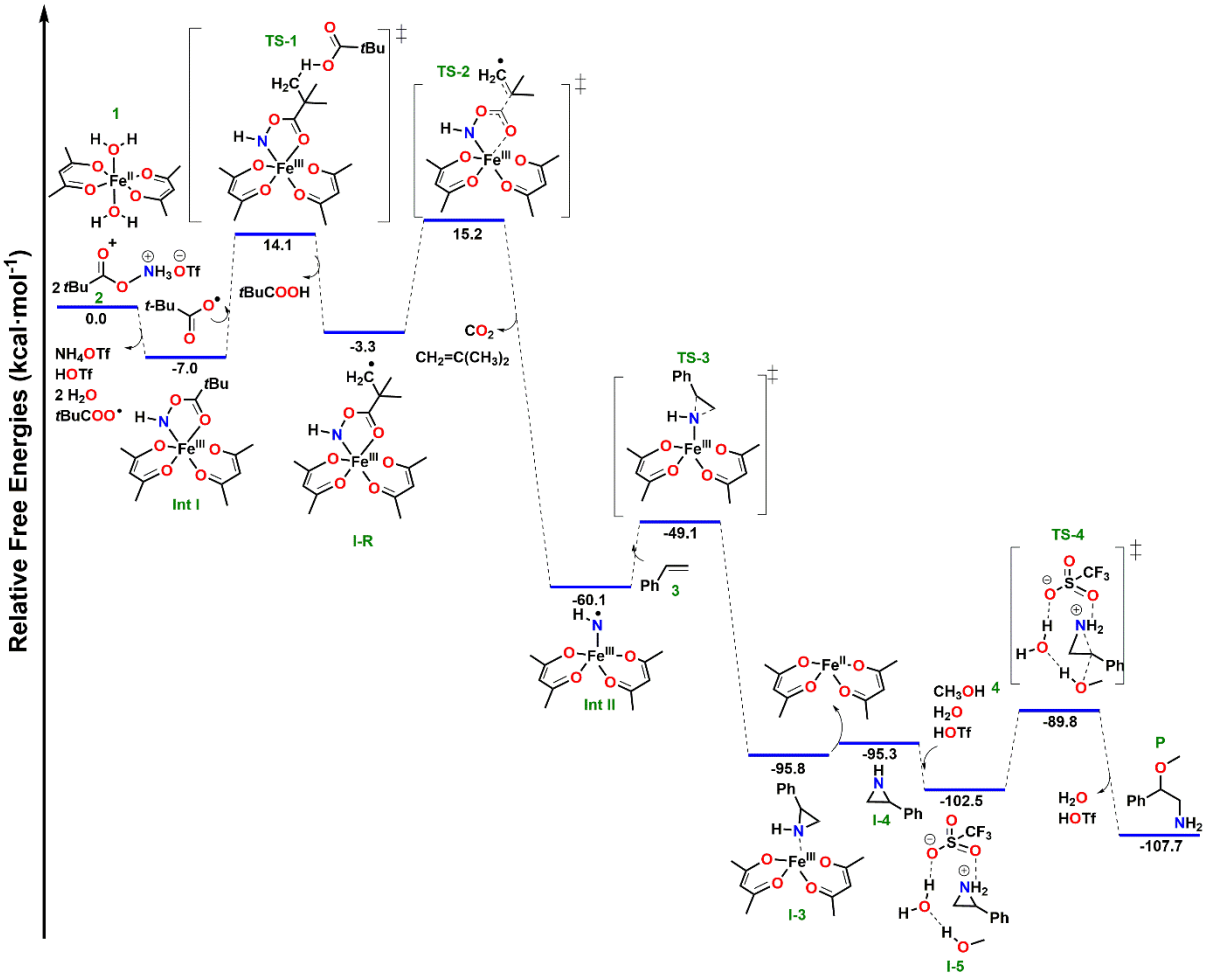

Figure 7.18 Computed reaction profile at the B3LYP-D3/def2-TZVP level of theory. All energies are given in kcal mol^{-1}. This figure was taken from ref. 283.

As discussed before, the formation of **Int I** from **1** is exergonic due to the formation of stable molecules like ammonium triflate. For the decomposition of **Int I** to give **Int II**, a two-step mechanism was proposed. The pivaloxy radical formed during the activation process abstracts a hydrogen atom from **Int I**, leading to the transient radical intermediate **I-R**. In the second step, **I-R** decomposes to **Int II** under formation of CO_2 and isobutene. Both side products were detected in gas chromatography. For the decomposition of **I-R**, three bonds break simultaneously: a) the N-O bond, b) the Fe-O bond and c) the C-C bond within the pivaloxy group, leading to **Int II** and the side products in a single step. The formation of **Int II** is highly

exergonic due to the formation of stable molecules like CO_2. The proposed radical intermediate **I-R** has a lower barrier for the decomposition to form **Int II** than **Int I**, as will be discussed below. However, it never was directly observed in experiment, indicating its short life time. As simulations of the kinetics will reveal, its concentration during the reaction is always too small to be detectable by experiment. For the conversion of **Int I** to **Int II**, the isosbestic point in the experimental UV-vis spectra is around 590 nm, but no clean isosbestic point could be identified. Multivariate curve resolution [331] of the UV-vis spectra indicates a third intermediate that prevents identification of a clear isosbestic point.

The next step is the key step of the whole reaction, namely the addition of **Int II** to the double bond of styrene **3**. The reaction has a very small barrier, indicating the high efficiency of **Int II** for nitrogen transfer reactions. The formed iron-bound aziridine **I-3** dissociates easily to give the bare aziridine **I-4** under restoration of the catalyst. In the last step of the reaction, the acid catalyzed ring opening of **I-4** by the nucleophilic methanol was computed. For this, the protonated aziridine was assumed. Also, this last reaction step has a very small reaction barrier and leads regioselectively to the formation of 2-methoxy-2-phenylethan-1-amine. Importantly, it was verified experimentally that the aziridine reacts with methanol under acidic conditions to the desired product in a very fast reaction even in the absence of any iron catalyst, which further supports the proposed iron-free ring opening of the aziridine [283]. The proposed pathway is one of many possible pathways and it is not excluded that other pathways that have the same observable outcome are competing. In the following, alternative reaction mechanisms are explored.

Alternative reaction pathways

In addition to the minimum energy path described in **Figure 7.18**, additional alternative reaction pathways were explored. The decomposition of **I-R** consists of three simultaneous bond cleavages, namely the N-O bond, the Fe-O bond and the C-C bond. However, also the step-wise mechanism in which the Fe-O bond is cleaved first, followed by simultaneous cleavage of the N-O and C-C bonds was computed. The corresponding transition states are shown in **Figure 7.19**. The step-wise TS is 9.7 kcal mol^{-1} higher in energy than the concerted one, indicating the importance of the additional Fe-O bond for the barrier height.

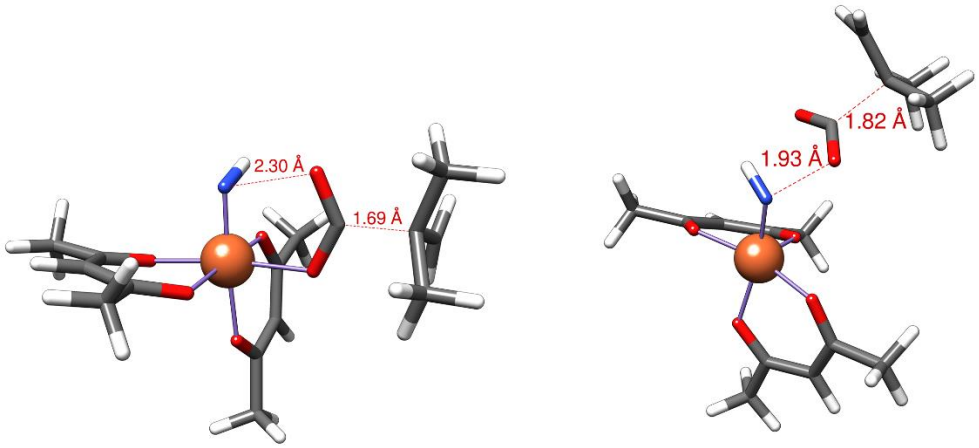

Figure 7.19 Transition states of concerted (left, **TS-2**) and step-wise (right) decomposition of **I-R**. This figure was adapted from ref. 283.

Besides the step-wise decomposition of **I-R**, the direct decomposition of **Int I** under formation of **Int II** and the *tert*-butyl radical was computed. The corresponding barrier amounts to 43.9 kcal mol^{-1}, in comparison to the barriers of 21.1 kcal mol^{-1} and 18.5 kcal mol^{-1} for the proposed decomposition in **Figure 7.18**. Therefore, this direct decomposition of **Int I** is far too high in energy and does not constitute a viable alternative to the proposed mechanism.

In the previous section, it was noted that an acidic medium is assumed. This, of course, raises the question whether **Int I** could not exist in the protonated state **[Int I-H]$^+$**. The analysis of the spectroscopic data of **[Int I-H]$^+$** was carried out analogously to the analyses carried out for **Int I** and **Int II**. The details are not discussed here, but the interested reader is referred to ref. 283. The existence of **[Int I-H]$^+$** could not be ruled out by this analysis. However, the presented reaction mechanisms (either *via* the decomposition of the radical intermediate **I-R** or *via* the direct decomposition of **Int I**) were also computed for the protonated species **[Int I-H]$^+$**. Interestingly, the direct decomposition of **[Int I-H]$^+$** has a barrier of 37.4 kcal mol^{-1} and is therefore lower than the corresponding barrier of **Int I**. However, for the two-step process, the barriers are 28.4 kcal mol^{-1} and 42.2 kcal mol^{-1}. Therefore, by far the lowest-in energy path identified leading to **Int II** is the one proposed in **Figure 7.18**.

Simulation of the reaction kinetics

A very useful approach that works like a “reality check” for any computed reaction profile is to simulate the kinetics of the reaction based on the computed barriers and to compare this kinetics with experimental data. This technique can help to identify obvious flaws in the reaction profile. For example, chronologically, the first TS found for the decomposition of **I-R** from **Figure 7.18**

actually was the TS with the already cleaved Fe-O bond (see **Figure 7.19**, right panel). Simulations of the reaction kinetics using this TS energy were not able to reproduce the experimental data at all, indicating that the profile must be fundamentally flawed. This eventually led to the discovery of the much more stable transition state **TS-2**. The idea behind these simulations is that the time dependence of the concentration of a species $\frac{dc_A(t)}{dt}$ is described by a first-order ordinary differential equation (ODE) and depends only on the educt concentrations of all elementary steps where this species is involved. For example, consider the following hypothetical elementary reaction.

$$A \xrightarrow{k} B + C. \tag{7.4}$$

In eq. (7.4), k is the corresponding rate constant. The derivatives $\frac{dc_A(t)}{dt}$, $\frac{dc_B(t)}{dt}$, and $\frac{dc_C(t)}{dt}$ only depend on the concentration of A in the following way.

$$\begin{aligned} \frac{dc_A(t)}{dt} &= -kc_A(t) \\ \frac{dc_B(t)}{dt} &= kc_A(t) = \frac{dc_C(t)}{dt}. \end{aligned} \tag{7.5}$$

Following this simple approach for the computed reaction profile in **Figure 7.18**, a system of coupled first-order differential equations can be formulated, with each species of the profile having its own differential equation. Given some initial concentrations, the concentration profiles are then obtained by integrating the system of ODEs numerically. In this chapter, the Python programming language together with the solvers provided by the SciPy package [332] was used. For an introduction to a graphical tool, see Appendix B.2. The rate constants can be obtained from the computed reaction barriers *via* the Eyring equation [333].

$$k_i = \frac{k_B T}{h} e^{-\frac{\Delta G_i^{\ddagger}}{RT}}. \tag{7.6}$$

For the simulations carried out here, always the forward and backward reaction were considered. No transition states were computed for the activation of the precursor. In order to keep the simulations stable, a small barrier was assumed in the simulations. From experiment it is known that the activation happens very fast. The experimental initial concentrations of the species (**1**, **PivONH$_3$OTf**, styrene and methanol) were used for the simulations as well.

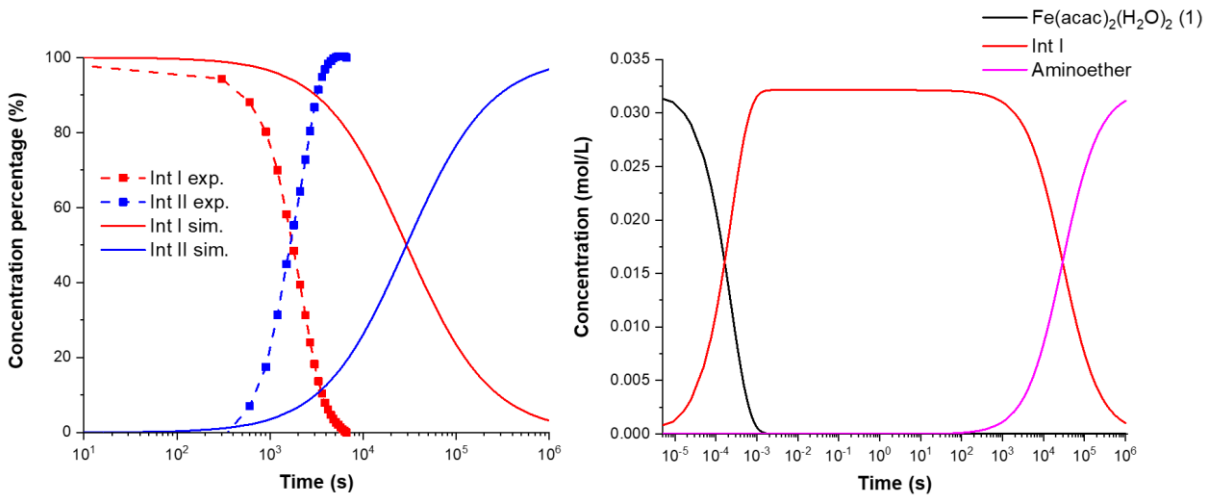

Figure 7.20 a) Time-dependence of the concentrations of **Int I** and **Int II** obtained from simulations of the kinetics in the absence of styrene. Experimental concentration profiles are given by dashed lines. Experimental concentration percentages were obtained from experimental UV-vis spectra. b) Simulated concentration profiles in the presence of styrene. This figure was taken from ref. 283.

In **Figure 7.20**, two different simulations are shown. First, the decay of **Int I** and the corresponding formation of **Int II** were simulated. For this, the initial concentration of styrene was set to zero, so that the reaction stops after the formation of **Int II** and hence **Int II** accumulates. This resembles the experiments carried out in the absence of styrene for measuring the decay rates of **Int I** and **Int II**. Since **Int II** is formed from **Int I**, the relative amount of these two species is used as measure for reaction progress. The experimental values were obtained from the corresponding time-dependence of the intensity of the characteristic chromophore. It can be seen that the simulations predict too large reaction times by around 1-2 orders of magnitude. It is important to keep in mind, however, that the computed reaction rates are very sensitive with respect to the reaction barrier. A change of the reaction barrier by 2 kcal mol^{-1} changes the rate constant by a factor of 30. Therefore, the deviations are entirely explainable with the expected errors of the chosen methodology, indicating that the computed barriers actually are in quite the right range. The second simulation was carried out in the presence of styrene, leading to the formation of the final product. During these simulations, it was noticed that **Int II** does not accumulate significantly due to the very fast consumption of **Int II** relative to its formation. The formation of **Int II** is the bottleneck of the reaction, while the nitrogen transfer reaction with styrene happens very fast. This is in agreement with experiment because also experimentally **Int II** is only observable in the absence of styrene. Importantly, no accumulation of the radical intermediate **I-R** was ever observed in any of the

simulations performed. As discussed before, also experimentally **I-R** was only observed indirectly, but never directly.

Chapter 8 Conclusion and Outlook

Accurate computational protocols based on state-of-the-art electronic structure calculations were developed and applied to highly relevant chemical transformations in the field of homogeneous catalysis. The catalytic systems investigated encompass large and flexible organic molecules for applications in enantioselective synthesis, as well as transition metal containing systems with elusive electronic structure. In principle, the approach used throughout this thesis is based on a multi-level filtering strategy for the various conformers or isomers of the system under investigation. In particular, a large number of conformers is initially generated at the semi-empirical or force field levels and then the number of candidate structures is progressively reduced while ascending a hierarchy of levels of theory. Standard DFT functionals were used for refining the geometries of the conformers and sorting them by energy, while the final electronic energies for the low energy structures were refined using accurate post-HF methods like DLPNO-CCSD(T) and HFLD or advanced exchange-correlation functionals like ωB97M-D4.

In Chapter 4, the intramolecular hydroalkoxylation of 4-phenylpent-4-ene-1-ol catalyzed by an IDPi catalyst was investigated. Three different mechanisms were computed. The preferred mechanism describes a concerted but asynchronous bond formation in which the C-H and C-O bonds are formed in a single step. It was found that the reaction is best described by an induced-fit model in which both the catalyst and the substrate change their conformation to a certain extend in order to allow for optimal interaction between the molecules. Further, it could be shown that the catalyst's distortion energy to a large extend is governed by intra-catalyst dispersion interactions. The stereoselectivity is governed by the different distortion energies of the major and minor reaction pathways, respectively. In order to realize the minor pathway, additional distortions of the catalyst and substrate have to be carried out that disrupt the intra-catalyst LD interactions and therefore increase the geometric preparation energy. Intermolecular dispersion interactions between the catalyst and the substrate on the other hand are the major stabilizing interaction component that reduces the reaction barriers and consequently are highly important for the reaction to occur on reasonable time scales. Comparison of experimental enantiomeric ratios with computed intra-catalyst NCI strengths for a series of related IDPi catalysts reveals a correlation between stereoselectivity and intra-catalyst LD interaction, which further supports the proposed model.

Chapter 5 has discussed cooperative catalytic effects for a series of modern organocatalysts (CPA, DTPA, DSI, IDP, IDPi). To this end, the first step was to investigate the dimerization process of one representative of each listed Brønsted acid class. For the less bulky systems (CPA, DTPA, DSI), the free dimerization energy is determined by the number and the strength of the intermolecular hydrogen bonds. CPA features two strong hydrogen bonds leading to negative free dimerization energies over the whole range of experimentally common temperatures. DTPA and DSI feature weaker hydrogen bonds and therefore should dimerize only under mild reaction conditions. The sterically demanding IDP and IDPi acids do not dimerize because the attractive London dispersive interactions between the monomers cannot compensate the large geometric preparation energy. In a second step, the C-C coupling reaction between silyl nitronates and silyl ketene acetals catalyzed by IDPi catalysts was investigated. It was found that catalyst aggregation is much more likely if additional charged or neutral molecules form a buffer between the catalyst molecules, which "glues" the catalyst molecules together by noncovalent interactions. The key point is that the mediator molecule sits between the two catalyst molecules and allows for attractive interactions (*e.g.*, electrostatic or dispersive) with *both* catalyst molecules without a large geometric preparation energy. For the case study reaction, aggregation of the chiral ion pair with another catalyst molecule shows free aggregation energies of around 0 at 153 K, indicating that the possibility of dimerization depends on the details of the experimental setup, such as temperature, solvent polarity or the details of the catalyst architecture. Interestingly, the dimeric pathway offers faster reaction rates compared to the monomeric pathway, indicating the potential importance of cooperative catalytic effects even for very bulky catalysts for which cooperative catalytic effects usually are not considered.

In Chapter 6, a computational protocol based on a geometric space-filling model together with extensive conformer sampling was introduced for the description of explicit solute-solvent interactions. As case study, the equilibrium of 1,4/1,6-disubstituted cyclooctatetraenes was used. Due to intramolecular LD interactions, the sterically more crowed 1,6-substituted isomer is always more stable than the 1,4-isomer, irrespective of the solvent or the size of the substituent. Two different models were tested with 8 and 10 solvent molecules, respectively. The model with 10 solvent molecules achieves quantitative agreement with experiment as well as with implicit solvation calculations for the relative energy of both isomers. The explicit solvation study allows for conclusions about the favored orientation of the solvent molecules. It was found that the solvent molecules align parallel to the solute as well as parallel to each other in order to maximize the intermolecular dispersive interactions.

Chapter 5 and Chapter 6 show how modern computational approaches can be used to study environmental effects in homogeneous catalysis. However, the efficiency of a catalyst of course critically depends on the electronic structure of the active catalytic species, especially for cases of covalent catalysis where the catalyst forms covalent bonds with the substrate. Chapter 7 focuses therefore on the amino etherification of styrene using simple and environmentally friendly iron(II) catalysts studied in a joint spectroscopic and computational project. Different models for the experimentally found intermediates **Int I** and **Int II** were proposed and the corresponding computed spectroscopic properties were used to correlate the experimental and computed data. **Int I** is a high-spin iron(III)-N-acyloxy complex that features a bidental binding mode of the N-acyloxy group. **Int II** contains an iron(III) high-spin site that couples antiferromagnetically with an iminyl (NH·) radical to overall yield a quintet ground state. The pronounced single bond character of the Fe-N bond together with the inability of the iminyl group to delocalize the spin density causes **Int II** to be highly efficient in nitrogen-transfer reactions, as was illustrated computationally by very small reaction barriers for the nitrogen-transfer reaction with styrene.

Protocols like the ones applied in this thesis have gained more and more attraction by computational as well as experimental chemists around the globe. It is expected that this trend will increase further. The increased computational power together with improved and more efficient algorithms will make application of such multi-step protocols more widespread and allows for more and more realistic modeling of complicated chemical systems. Improved multiscale approaches (such as two- or three-layered ONIOM approaches) are also likely to contribute to achieving this goal. Moreover, more and more tools are being developed to automate workflows, making these protocols easier to use and more efficient. This is not only true for *e.g.*, conformational sampling strategies, but for finding new reactions or low-energy paths connecting given reactants and products as well. The achievements made in theoretical and computational chemistry will likely find resonance in related fields of science, such as the life sciences.

Appendix A Supplementary Material

In this Appendix, supplementary material for Chapter 7 is provided.

A.1 Deconvoluted UV-vis spectra and experimental XAS spectra

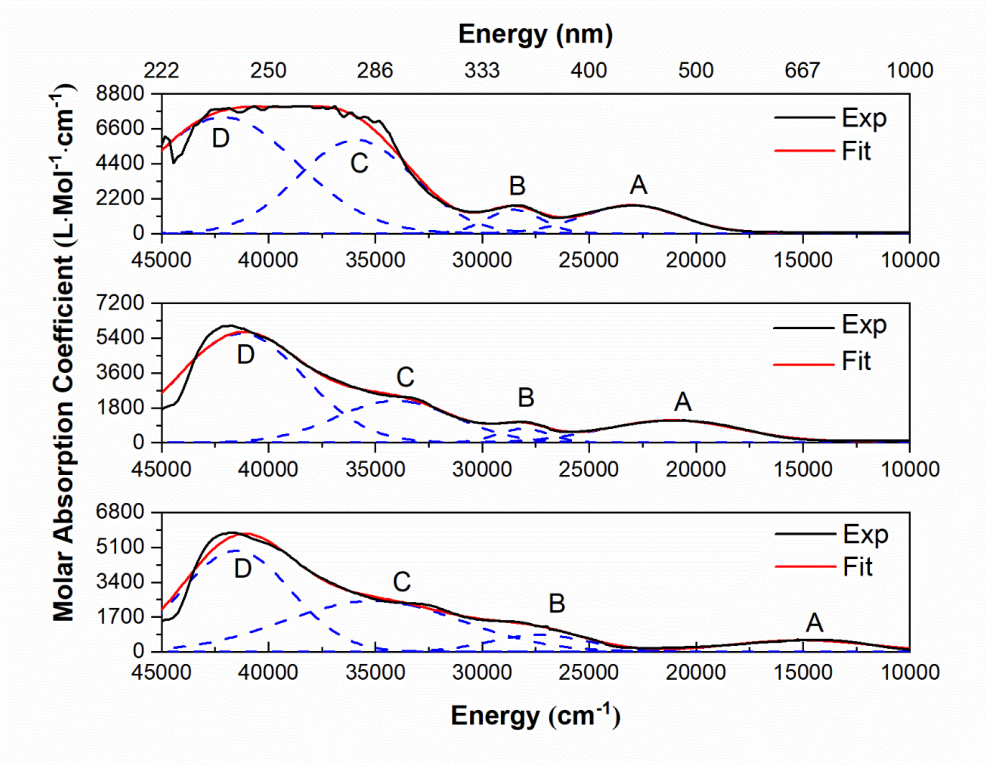

Figure A 1 Gaussian deconvolution of experimental UV-vis spectra of the precursor **1** (top), the first intermediate **Int I** (middle) and the second intermediate **Int II** (bottom). The experimental spectra are shown in black, the fitted spectra (*i.e.*, the sum of all gaussians) are shown in red. For each species, four bands (A-D) are identified. This figure is adapted from ref. 283.

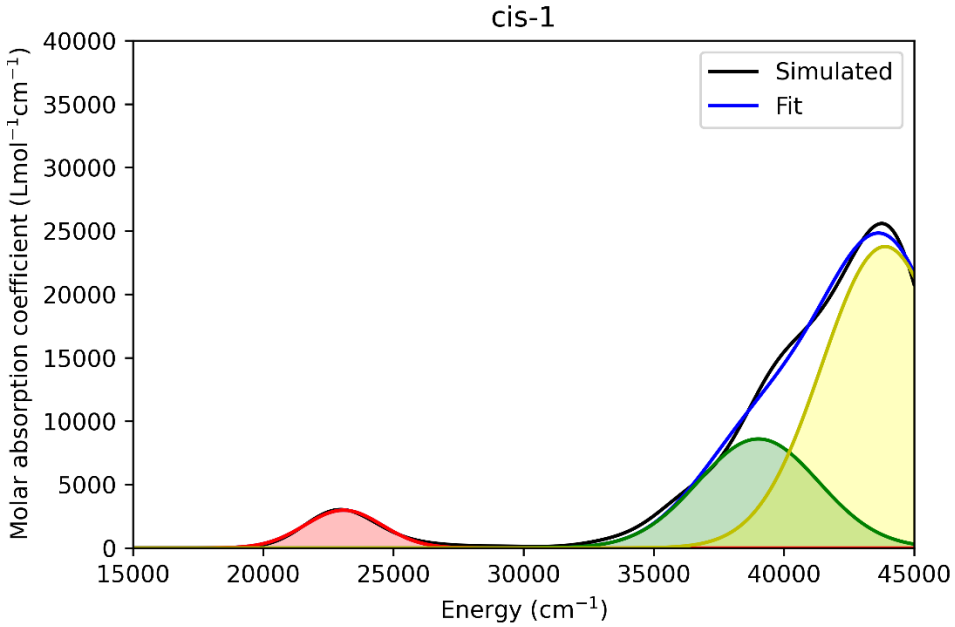

Figure A 2 Gaussian deconvolution of the computed spectrum (black) of *cis*-**1**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

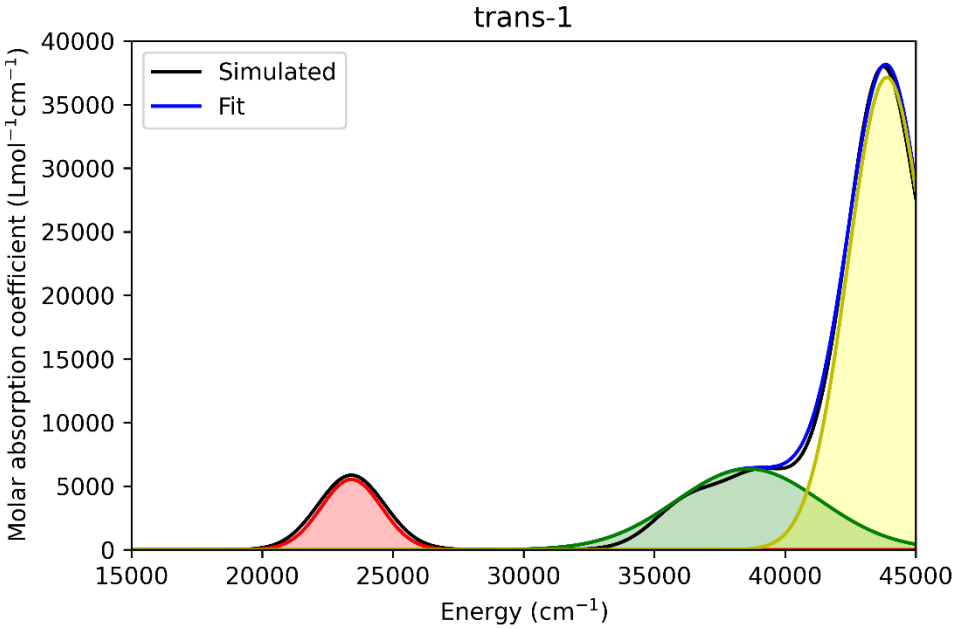

Figure A 3 Gaussian deconvolution of the computed spectrum (black) of *trans*-**1**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

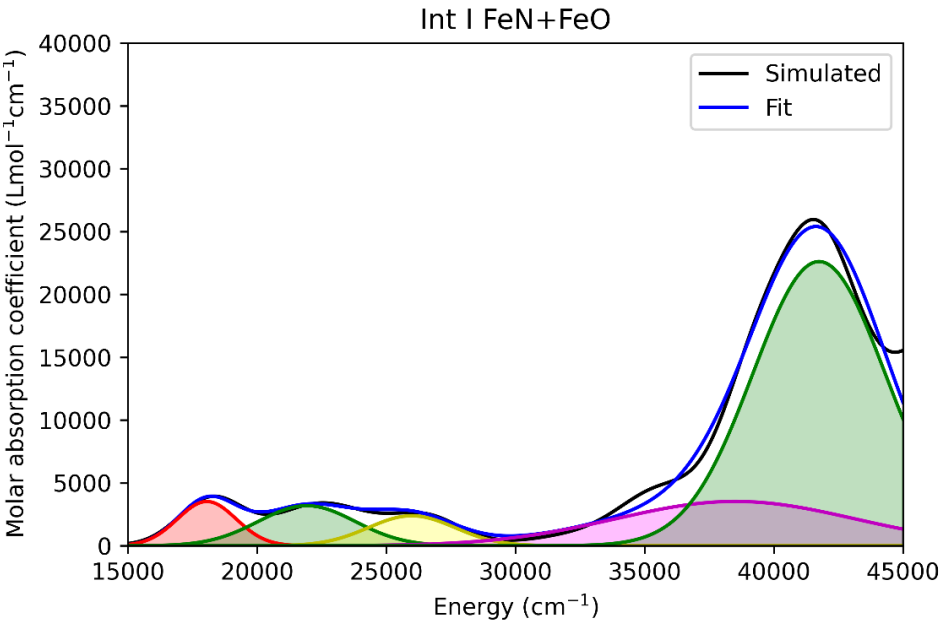

Figure A 4 Gaussian deconvolution of the computed spectrum (black) of **Int I (FeN+FeO)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

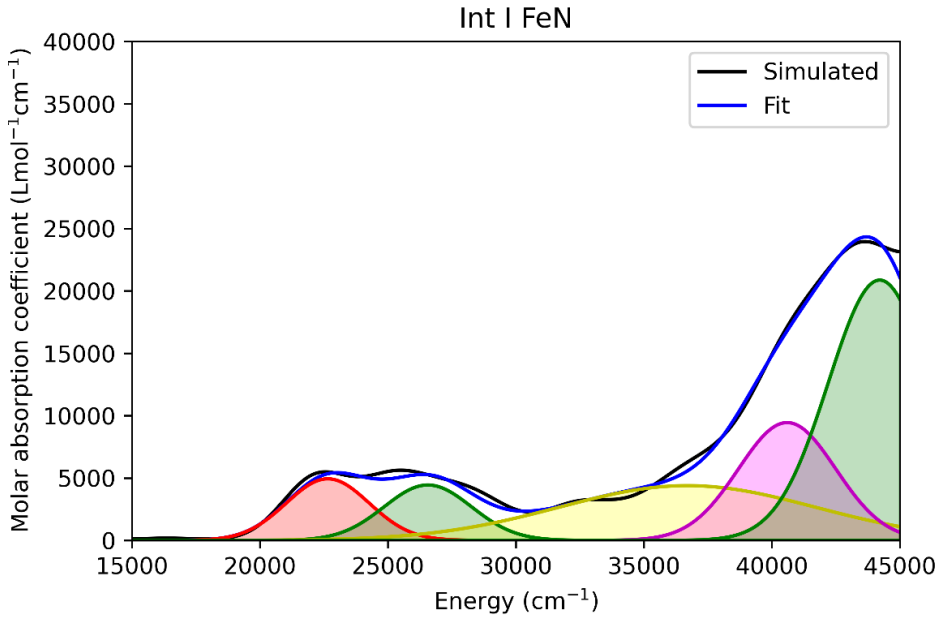

Figure A 5 Gaussian deconvolution of the computed spectrum (black) of **Int I (FeN)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

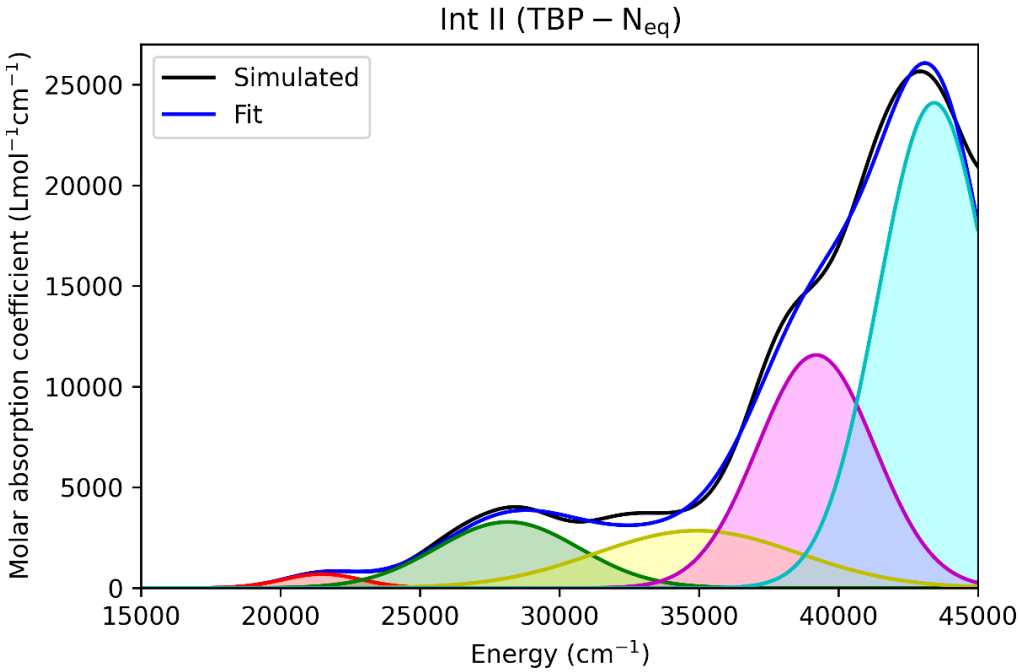

Figure A 6 Gaussian deconvolution of the computed spectrum (black) of **Int II (TBP-N$_{eq}$)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

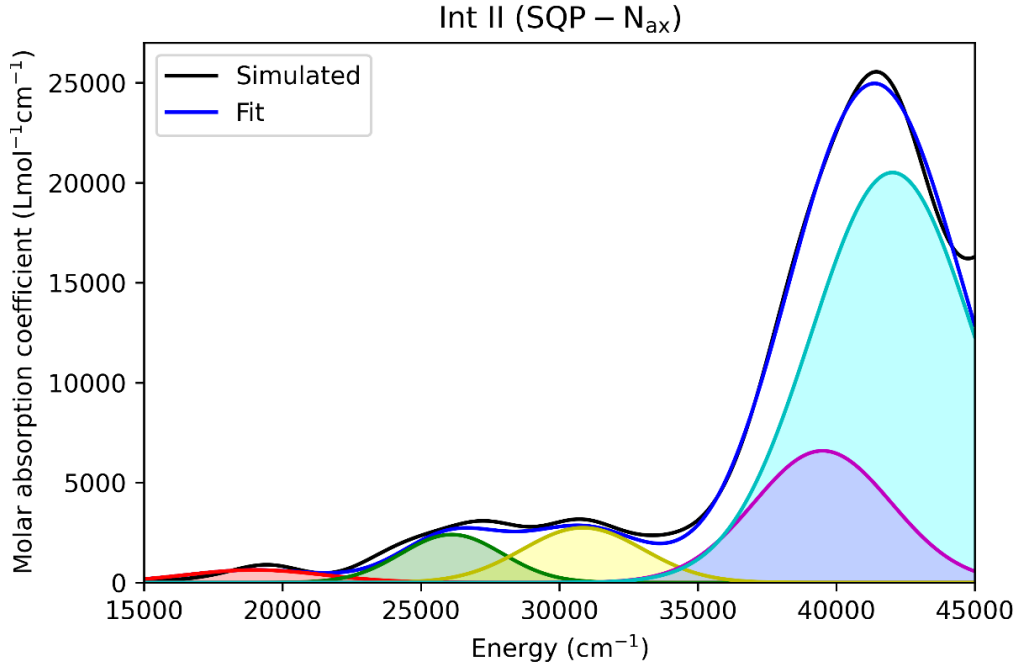

Figure A 7 Gaussian deconvolution of the computed spectrum (black) of **Int II (SQP-N$_{ax}$)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

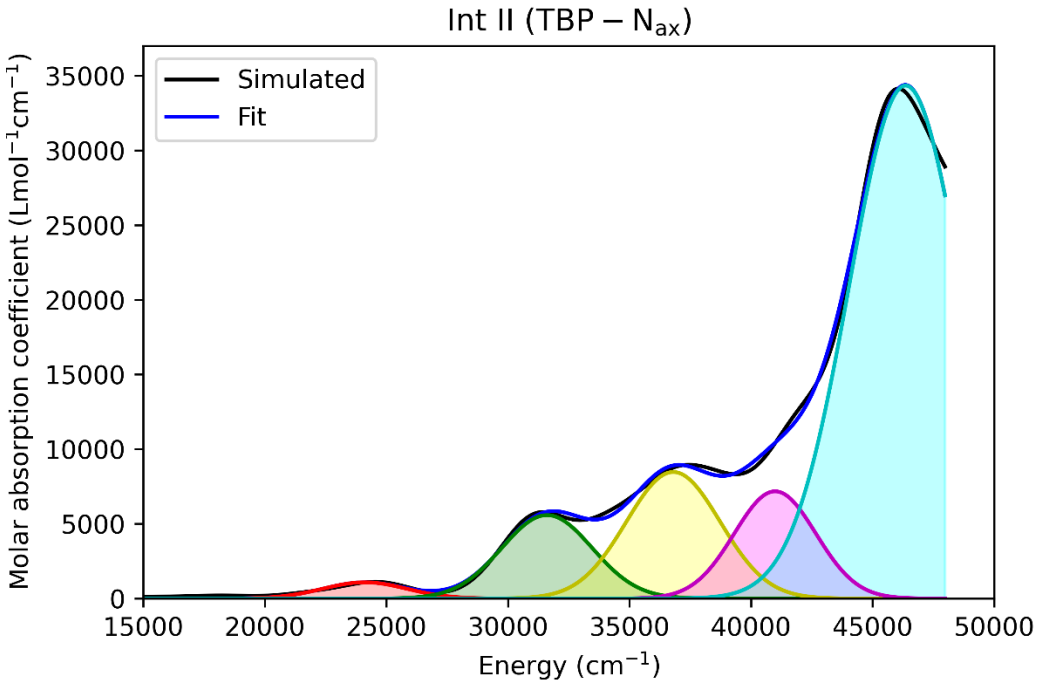

Figure A 8 Gaussian deconvolution of the computed spectrum (black) of **Int II (TBP-N$_{ax}$)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

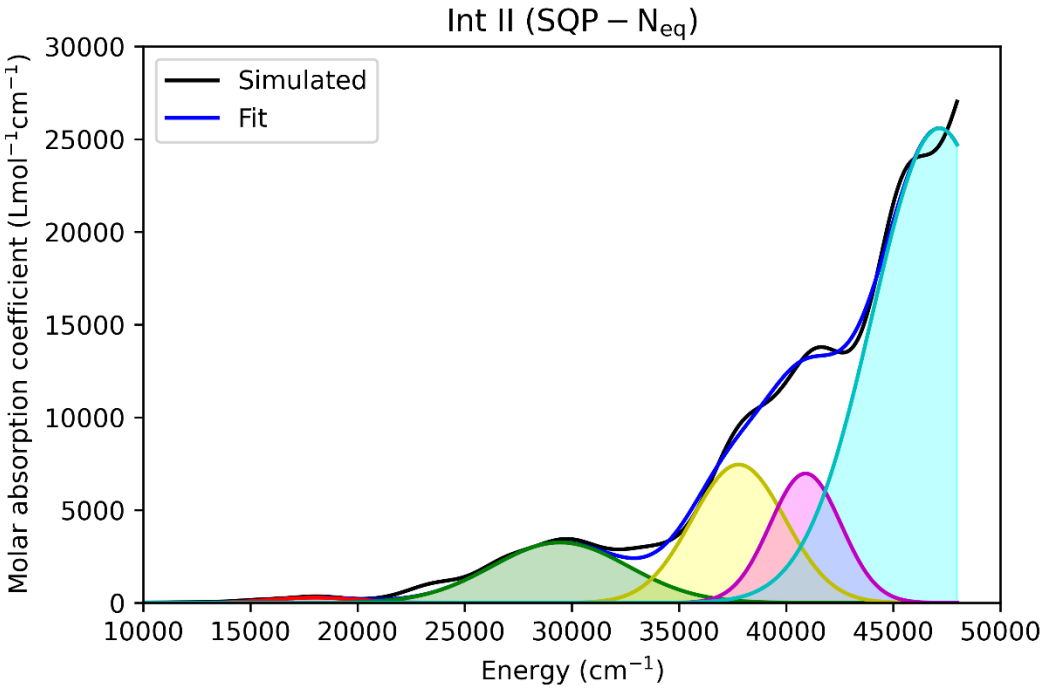

Figure A 9 Gaussian deconvolution of the computed spectrum (black) of **Int II (SQP-N$_{eq}$)**. The fitted spectrum (*i.e.*, the sum of all gaussians) is shown in blue.

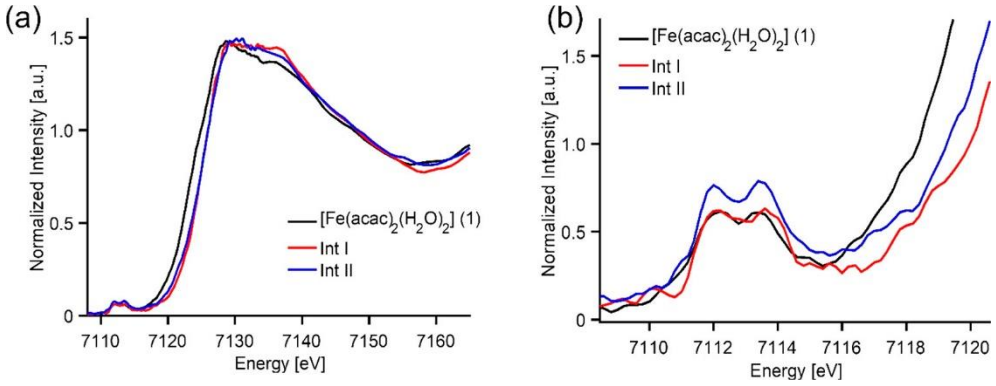

Figure A 10 Experimental XAS spectra. This figure was not prepared by the thesis author. This figure was taken from ref. 283.

A.2 Results from ab initio ligand field theory

In this section, the AI-LFT results for the precursor **1** and the first intermediate **Int I** are presented and discussed. Since the second intermediate **Int II** has a pronounced diradical character, its discussion within the AI-LFT framework is omitted here. For obtaining AI-LFT results of **1**, CAS-SCF calculations with 6 electrons (Fe(II) has 6 d-electrons) in 5 orbitals (the 5 3d-orbitals) were carried out. All multiplicities (quintet, triplet and single state) were taken into account. For each multiplicity, all roots were included (5 roots for the quintet, 45 roots for the triplet and 50 roots for the singlet state). The nephelauxetic effect of the ligands is usually described by the parameter β calculated from eq. (A.1) [334].

$$\beta = \frac{B_{complex}}{B_{free}} \leq 1. \qquad \text{(A.1)}$$

$B_{complex}$ is the Racah parameter B computed for the metal ion in the complex, while B_{free} is the computed Racah parameter of the free ion.

The results are shown in **Table A 1**. The first three orbitals are the non-bonding t_{2g} orbitals that have almost identical energies in *cis*-**1** and *trans*-**1**. However, the three t_{2g} orbitals are not degenerate because the coordination shell is not perfectly octahedral. The same effect can be seen for the two antibonding e_g orbitals which split significantly stronger than the t_{2g} orbitals. This effect is even increased for the *trans* isomer (energy difference between the two e_g orbitals is 6131.2 cm^{-1} for *trans*-**1** and 2867.7 cm^{-1} for *cis*-**1**). Especially the highest orbital energy differs significantly between the two isomers. As was shown in section 7.4.1, the two isomers of **1** show very similar spectroscopic properties for a wide range of techniques. However, they differ significantly in their d-d excitation energies. Therefore, obtaining the d-d excitation

energies experimentally (*e.g.*, by means of magnetic circular dichroism) and comparing with the computed transition energies could in principle lead to unambiguous assignment.

Table A 1 Relative orbital energies of the metal d-orbitals and Racah parameters quantifying electron repulsion obtained from AI-LFT calculations for *trans*-**1** and *cis*-**1** assuming the d^6 configuration for iron. No spin-orbit coupling was included.

Orbital number	Orbital energy (cm^{-1})	
	trans-**1**	*cis*-**1**
1	0.0	0.0
2	1194.6	1142.1
3	1958.8	2029.2
4	5688.3	6338.8
5	11819.5	9206.5
Racah parameter	Value	
A	22.490 eV	22.470 eV
B	0.142 eV	0.142 eV
C	0.528 eV	0.527 eV
C/B	3.714	3.717
β	0.963	0.961

The total root-mean square deviation of the transition energies obtained from LFT compared to the ones obtained from CAS-SCF is 224.2 cm^{-1} for *trans*-**1** and 181.6 cm^{-1} for *cis*-**1** (averaged over all multiplicities). This demonstrates that the precursor **1** can be fitted to the LFT model very well. Note that the calculations did not include the NEVPT2 energy correction. Since the NEVPT2 correction includes additional effects of dynamic correlation that go beyond the scope of LFT, it is therefore expected that LFT yields larger deviations with respect to NEVPT2 energies than with respect to CAS-SCF energies [335]. The influence of the ligands on the Racah parameter B is very small, leading to values of β that are only slightly below 1. However, it is assumed that CAS-SCF may systematically underestimate the reduction of B due to the very ionic picture CAS-SCF draws [336]. The small covalent mixing between the iron 3d orbitals and the ligand orbitals was also discussed for the QRO representation of the electronic structure of *trans*-**1** shown in **Figure 7.9**, which adds some plausibility the computed values for β.

For **Int I**, 5 electrons in 5 d-orbitals are considered. The sextet (one root), quartet (24 roots) and doublet (75 roots) states were taken into account. While **Int I (FeN+FeO)** has a distorted octahedral environment of the iron site, **Int I (FeN)** has a distorted trigonal-bipyramidal geometry. Of course, in both cases splitting of the non-bonding orbitals is observed, since the geometries are not perfectly symmetric and the ligands are different. In comparison to the precursor **1**, **Int I** shows a larger splitting between the non-bonding and antibonding orbitals. This is expected because **Int I** is a ferric complex, while **1** is a ferrous one. However, **Int I** also has different ligands than **1**, since the two water molecules of **1** are substituted by the NHOPiv group. The different splitting patters are the result of different oxidation states, ligands and geometries. The nephelauxetic parameter β is smaller for **Int I** compared to **1**. Although β is overestimated by AI-LFT, the reduced values of β for the isomers of **Int I** matches the expectations, since the covalent bonding interactions of iron(II) with the water molecules present in **1** should be much weaker than the covalency of the iron(III)-nitrogen bond of **Int I**. The RMS error for the excitation energies obtained from LFT averaged over the quartet and doublet block amounts to 229.7 cm^{-1} for **Int I (FeN+FeO)** and to 376.2 cm^{-1} **Int I (FeN)** compared to the CAS-SCF excitation energies.

Table A 2 Relative orbital energies of the metal d-orbitals and Racah parameters quantifying electron repulsion obtained from AI-LFT calculations for **Int I (FeN+FeO)** and **Int I (FeN)** assuming the d^5 configuration for iron. No spin-orbit coupling was included.

Orbital number	Orbital energy (cm^{-1})	
	Int I (FeN+ FeO)	**Int I (FeN)**
1	0.0	0.0
2	1560.2	561.6
3	2941.8	4388.3
4	9813.1	8707.8
5	13672.1	12762.0
Racah parameter	Value	
A	23.000 eV	22.703 eV
B	0.146 eV	0.143 eV
C	0.548 eV	0.543 eV
C/B	3.761	3.805
β	0.899	0.881

A.3 Example input files for Chapter 7

```
!B3LYP def2-TZVP def2/J RIJCOSX VeryTightSCF D3BJ

%basis
    newgto Fe “CP(PPP)“ end
end

### charge mult coordinates ###

%eprnmr
    nuclei = all Fe {fgrad, rho}
end
```

Figure A 11 Example input file for the computation of Mössbauer parameters.

```
!B3LYP def2-TZVP def2/J RIJCOSX VeryTightSCF D3BJ

%tddft
    nroots 200
    doNTO true
end

### charge mult coordinates ###
```

Figure A 12 Example input file for the computation of vertical excitation energies.

```
!B3LYP ZORA ZORA-def2-TZVP def2/J RIJCOSX VeryTightSCF D3BJ

%tddft
    nroots 20
    orbwin[0] = 0,0,0,0,0
    orbwin[1] = 0,0,-1,-1
    doquad true
end

### charge mult coordinates ###
```

Figure A 13 Example input file for the computation of XAS excitation energies.

```
!B3LYP def2-TZVP def2/J def2-TZVP/C RIJCOSX VeryTightSCF D3BJ NMGRAD

%tddft
    nroots 10
    doNTO true
end

%rr
    States 1,2,3,4,5
    HessName "example.hess"
    ASAInput true
end

### charge mult coordinates ###
```

Figure A 14 Example input file for computing the excited state normal mode gradient and generating the input for orca_asa. Starting from ORCA 5, the orca_esd module is recommended for computing resonance Raman spectra.

```
!UHF def2-TZVP def2/JK RI-JK conv NEVPT2 Moread
%moinp "TBP_Nax.qro"

%casscf
    nel 12
    norb 11
    mult 5
    nroots 1
    trafostep ri
    Orbstep SuperCI
    Switchstep KDIIS
end

### charge mult coordinates ###
```

Figure A 15 Example input file for the CAS-SCF ground state calculations of **Int II**.

```
!UHF def2-TZVP def2/JK RI-JK conv NEVPT2 PATOM

%casscf
    nel 6
    norb 5
    mult 5,3,1
    nroots 5,45,50
    actorbs dorbs
end

### charge mult coordinates ###
```

Figure A 16 Example input file for the AI-LFT computations of a d^6-system (the precursor **1**).

```
!B3LYP def2-TZVP def2/J RIJCOSX VeryTightSCF D3BJ

%scf
    BrokenSym 5,1
end

### charge mult coordinates ###
```

Figure A 17 Example input file for broken-symmetry DFT calculations of **Int II**. Importantly, the multiplicity of the high-spin state must be given in the coordinate section.

All example input files given in this section are adapted to ORCA 5, *e.g.*, by using the new default grids.

Appendix B Developed Software

This Appendix is used to introduce software tools that were developed by the author during the course of this thesis. The software is publicly available on GitHub and is deployed as ready-to-use binaries. The Appendix should give some insights into the design principles and basic usage of the individual programs.

B.1 Command line analysis tool for LED

Overview

The ORCA LED analysis tool `orca_led` is part of the official ORCA helpers repository [337] on GitLab that serves to provide a set of small utility programs that help in processing input/output files of ORCA calculations and might therefore be helpful for analysis. `orca_led` works entirely on ORCA output files and therefore has no interface to the ORCA program package. The main purpose of `orca_led` is to print the main results of the LED analysis of DLPNO-CC calculations. For this, the program provides two output files. `basename.led` contains the results in well-formatted and easy-to-read style, while `basename.csv` provides the data in the csv data format, which can be read in by *e.g.*, Microsoft Excel. However, `orca_led` is also able to decompose binding energies of an aggregate. For this, the user has to perform calculations on the isolated monomers in their dimer geometries (for the computation of electronic preparation) and on the monomers in their equilibrium geometries (for the computation of geometric preparation). Thus, `orca_led` takes up to $2N_{frag}+1$ ORCA output files as input.

Design principles

In principle, `orca_led` is a parsing program: it parses ORCA output files, collects the relevant data and uses them for the decomposition of binding energies. Eventually, all relevant data are printed to the standard output and to file. `orca_led` is written in C++ (C++17 is required). The basic data structures are the `Dimer` and `Monomer` classes that hold the data of the corresponding calculations for the dimer and monomers. These data consist of `double`s, `std::string`s and `std::vector<double>`s. Both classes have a constructor that takes a `std::string_view` object as parameter, which will hold the file name. This means that the reading and parsing of input files happens during the construction of the `Dimer` and `Monomer` objects. Once these objects are properly constructed, the input files are not needed anymore. It should be noted here that, since `Dimer` and `Monomer` only contain built-in data types or types from the standard

library, the compiler will generate move constructors and move assignment operator overloads for both classes, which will increase performance (*e.g.*, if `std::vector<T>::push_back()` is used). The construction of the `Monomer` objects happens in parallel making the program more efficient. For this, the C++11 task-based parallelization techniques based on `std::future` and `std::promise` are used. Eventually, the `Dimer` object and a `std::vector<Monomer>` are passed to a driver function that takes care of all computations and input/output (I/O) operations. For the formatting of the output strings and the printing, the fmt library [338] is used. Once `std::format` (C++20) and `std::print` (C++23) are sufficiently supported by compilers, those functionalities will be used instead to facilitate compilation and to increase portability.

Basic usage

Pre-compiled and statically linked executables of `orca_led` are available for Windows und Linux. However, the only used third-party library is fmt, which can be used in header-only mode, making compilation of the source code from scratch easy. `orca_led` works *via* the command line. For illustration of the basic usage, the help message is given in **Figure B 1**.

```
  -h, --help:       Print this message
  -d, --dimer:      Specify ORCA output file that contains LED results
  -m, --monomers:   Specify list of ORCA output files of monomers for electronic/geometric preparation
  -c, --compare:    Specify two .led files obtained from previous orca_led runs to compare two binding energies.
                    Works only if the decomposition of binding energies was carried out!


How you should use this program:
Start with the dimer file. Only one dimer file can be specified
Continue with the monomer files from fragment 1 to fragment N in dimer geometry (for electronic preparation)
Continue with the monomer files from fragment 1 to fragment N in equilibrium geometry (for geometric preparation)
The program will produce a DIMERNAME.led file which contains the results
It will also write a DIMERNAME.csv file which can be opened by Excel or LibreOffice
If you mess something up with the ordering, the program might help you

Command line example:
orca_led --dimer dimer.out --monomers frag1.out frag2.out frag1_opt.out frag2_opt.out

For the compare mode you have to specify exactly two .led files
Command line example:
orca_led --compare file1.led file2.led
```

Figure B 1 Help message for the `orca_led` program.

`orca_led` works entirely with output files from ORCA calculations. If only the dimer (or more accurately the multimer) output file is passed to the program, only essential results of the LED analysis will be printed, since decomposition of the binding energy is not possible if no monomers are provided. The program performs basic checks to see whether basis sets, number of electrons, etc. match. If output files of the monomers are passed to the program, it expects either exactly N_{frag} or $2N_{frag}$ monomer output files, with N_{frag} being the number of fragments

obtained from the multimer output file. If N_{frag} monomer files are passed to the program, it will use them to compute the electronic preparation. If $2N_{frag}$ monomer files are passed, electronic and geometric preparation will be computed. In addition, the program features a comparison mode which takes exactly two `.led` files (which will be generated by a usual `orca_led` run).

```
******* GENERAL INFORMATION ABOUT THE DIMER SYSTEM *******

nel:              998
nbasis:             0
nfragment:          2
HFType:           RHF

******* OVERALL ENERGIES *******

E(HF):          -7615.38715
E(Ref):         -7615.38715
E(CCSD):        -7639.00115
E(CCSD(T)):     -7640.11942
E(CORR):        -23.61399
E(Triples):     -1.11827

******* DECOMPOSITION OF HARTREE-FOCK REFERENCE ENERGY *******

E(HF) for Fragment 1:                                          -7115.43580
E(HF) for Fragment 2:                                          -499.27295

Electrostatic Interaction between fragments 2 and 1:           -0.59365

Exchange Interaction between fragments 2 and 1:                -0.08475

Consistency check:
Total Reference energy:                                        -7615.38715
Sum of the fragments HF reference and interaction energies:    -7615.38715



******* DECOMPOSITION OF INTRA-EXCITATIONS *******

                                   Fragment 1    Fragment 2

Intra strong pairs                 -21.36844     -2.03525
Intra triples                      -1.01573      -0.09157
Intra weak pairs                   -0.12855      -0.00612
Intra singles                      0.00000       -0.00000
******* INTERACTION ENERGIES BETWEEN FRAGMENTS *******

Warning: To calculate the percentages the absolute amount of the energies was used!

                                                               HFint      %     CCint      %
Interaction between fragments 2 and 1:                         -0.67840   88.7  -0.08660   11.3


Total Interaction Map:

                                   Fragment 1    Fragment 2

Fragment 1                         0.00000       -0.76500
Fragment 2                                       0.00000
```

Figure B 2 Snippet of the `orca_led` output file for an example system that consists of two fragments. The energies are given in hartrees.

This comparison mode might be interesting if, as it was done in Chapter 4, differences in binding energies shall be decomposed. In the following, some of the results obtained for the transition state **TS1-1** discussed in Chapter 4 are shown here. In the first step, some general information about the system is printed and the main results from the LED analysis of the multimer are re-printed (**Figure B 2**). Note that `orca_led` also provides an interaction map, that is, the total interaction between all pairs of fragments is printed in an upper triangular matrix. For the decomposition of the binding energies (requires calculations of the monomers), two different approaches are implemented. As discussed in section 2.3.4, the way the binding energy is decomposed is not unique.

```
******* IMPORTANT NOTE *******
There are several different ways of how to decompose the binding energy
This program uses two different approaches and it will do both approaches by default

Approach 1: Calculate dispersive and non-dispersive interactions only at the DLPNO-CCSD level and treat triples interactions separately
Ebind = E_geo_prep + E_el_prep (ref) + E_int_Coulomb (ref) + E_Int_Exchange (ref) + E_Disp (CCSD) + E_int_nondisp (CCSD) + E_int (C-(T))
Within approach 1 the non-dispersive interaction is NOT decomposed into pairwise interaction terms!

Approach 2: Decompose triples interactions into dispersive and non-dispersive contributions.
Decompose non-dispersive interactions further into electronic preparation and pairwise fragment interaction terms
Ebind = E_geo_prep + E_el_prep + E_int_Coulomb (ref) + E_Int_Exchange (ref) + E_Disp (CCSD-(T)) + E_int_nondisp (CCSD(T))
For systems with more than two fragments in addition a delocalized term from the triples excitations arises

************************************************************
*******                Approach ONE                *******
************************************************************



******* GEOMETRIC PREPARATION ENERGY *******

Geometric preparation for Fragment 1:                          41.30974
Geometric preparation for Fragment 2:                          9.17427
Sum:                                                           50.48401



******* HF-Interaction energy *******

Electronic preparation for Fragment 1:                         136.60865
Electronic preparation for Fragment 2:                         266.08582
Sum:                                                           402.69447

Electrostatic Interaction between fragments 2 and 1:           -372.51946

Exchange Interaction between fragments 2 and 1:                -53.18222

Consistency check:
Total HF interaction energy:                                   -23.00721
Sum of elec.preparation/electrostatic- and exchange Interaction:      -23.00721

******* CCSD Interaction energy *******

Warning: NonDisp (Strong Pairs) includes Singles excitations!
Total Nondispersion interaction energy:                        9.64300
Total Dispersion interaction energy:                           -33.71340

******* Triples Interaction energy *******

Total Triples interaction energy:                              -3.35707



Consistency check:
Total interaction energy (with geometric preparation):         0.04824
Sum of all mentioned interaction terms:                        0.04933
```

Figure B 3 First approach for decomposing the binding energy into various contributions. The energies are given in kcal mol^{-1}.

```
************************************************************
*******                Approach TWO                *******
************************************************************



******* GEOMETRIC PREPARATION ENERGY *******

Geometric preparation for Fragment 1:                          41.30974
Geometric preparation for Fragment 2:                          9.17427
Sum:                                                           50.48401


******* ELECTRONIC PREPARATION OF REFERENCE WAVEFUNCTION *******

Reference preparation for Fragment 1:                          136.60865
Reference preparation for Fragment 2:                          266.08582
Sum:                                                           402.69447

******* ELECTRONIC PREPARATION OF CORRELATION WAVEFUNCTION *******

Correlation preparation for Fragment 1:                        12.00462
Correlation preparation for Fragment 2:                        14.91079
Sum:                                                           26.91541



******* HF-INTERACTION ENERGY *******

Electrostatic Interaction between fragments 2 and 1:           -372.51946

Exchange Interaction between fragments 2 and 1:                -53.18222

Sum of electrostatic- and exchange Interaction:                -425.70168



******* DISPERSION INTERACTION ENERGY *******

CCSD strong pairs, weak pairs and (T) triples contribute!
Dispersion interaction between fragments 2 and 1:              -38.24473

Total Dispersion (Strong pairs):                               -26.42635
Total Dispersion (Weak pairs):                                 -7.28705
Total Dispersion (Triples):                                    -4.53133
Total Dispersion (Sum):                                        -38.24473
******* NON-DISPERSION INTERACTION ENERGY *******

CCSD strong pairs, weak pairs and (T) triples contribute!
Non-Dispersive interaction between fragments 2 and 1:        -16.09815
Sum:                                                         -16.09815



Consistency check:
Total interaction energy (with geometric preparation):       0.04824
Sum of all mentioned interaction terms:                      0.04933



************************************************************
*******                  Summary                   *******
************************************************************

E(Geo-prep)                                                  50.48401
E(El-prep)                                                   429.60988
E(ref)_int                                                   -425.70168
EDisp_int                                                    -38.24473
ENonDisp_int                                                 -16.09815
E(tot)_int                                                   0.04933


Total computation time: 427.49653 ms
```

Figure B 4 Second approach for decomposing the binding energy into various contributions. The energies are given in kcal mol^{-1}.

The program uses two different approaches for the decomposition of the binding energies, as can be seen from **Figure B 3** and **Figure B 4**. Consistency checks are carried out routinely to determine whether something went wrong during the computation.

Both approaches are always carried out, and for each approach, a summary containing the final information is printed. In addition, the main results are printed to `.xls` files that can be opened with Microsoft Excel or comparable programs.

B.2 A graphical user interface-based tool for chemical reaction kinetics

Motivation

In section 7.4.4, the kinetics, *i.e.*, the concentration course of the involved species with respect to the time were obtained by integrating the set of ordinary differential equations (ODE) numerically. For this, a program using the scripting language Python 3 was developed. For researchers without background in programming, a graphical user interface-based (GUI) tool is presented in this section. The main task of this program is, given a set of reaction equations and initial concentrations, to set up the correct system of differential equations and to integrate that system numerically. In addition, the obtained concentration profiles $c(t)$ are directly plotted for a fast and convenient overview. The program is publicly available on GitHub [339].

Theoretical background

The kinetics of chemical reactions can be formulated with the help of ordinary differential equations, since the time derivative of the concentration $\frac{dc_A(t)}{dt}$ of species A depends on the educt concentrations of the corresponding elementary step. For the elementary reaction given in eq. (B.1), the time derivatives of the concentrations are given in eq. (B.2).

$$A \xrightarrow{k} B + C. \tag{B.1}$$

$$\begin{aligned} \frac{dc_A(t)}{dt} &= -kc_A(t) \\ \frac{dc_B(t)}{dt} &= kc_A(t) = \frac{dc_C(t)}{dt}. \end{aligned} \tag{B.2}$$

A chemical reaction usually consists of multiple of these elementary reactions. From quantum chemical calculations, the corresponding reaction mechanisms might be obtained. The individual elementary steps are coupled to each other *via* the concentrations of the species. Thus, the kinetics of a chemical reaction consisting of multiple elementary steps is given by a

set of ODEs. The rate constant k_i of the *i*-th elementary reaction is computed *via* the Eyring equation.

$$k_i = \frac{k_B T}{h} e^{-\frac{\Delta G_i^{\ddagger}}{RT}} (c^{\ominus})^{1-m}. \tag{B.3}$$

In eq. (B.3), k_B is the Boltzmann constant, T is the temperature, h is the Planck constant, $\Delta G_i^{\ddagger}$ is the reaction barrier for the *i*-th elementary reaction, *R* is the ideal gas constant. $c^{\ominus}$ is the standard concentration and m is the molecularity. The term $(c^{\ominus})^{1-m}$ does not change the numerical value of k_i, but ensures that k_i has the correct dimension (the dimension depends on the molecularity of the elementary reaction). This formalism includes back reactions as ordinary elementary steps. $\Delta G_i^{\ddagger}$ might be obtained from a computed reaction profile, as was done in section 7.4.4. The rate constants can in principle also be obtained from experiment. Including all elementary steps of a reaction system leads to a system of first-order but nonlinear ODEs with constant coefficients that must be integrated numerically to obtain $c_X(t)$. The different rate constants can have very different numerical values because very fast and very slow reactions can occur within the same reaction system. For integrating these systems, special solvers, so-called stiff stable solvers, are needed that usually require the computation of the Jacobian matrix [340].

Design principles

The program is entirely written in C++. The program basically consists of two parts, namely a GUI part and the "backend" part. The idea is to keep the interaction between the individual parts as small as possible. The interaction from the GUI's side to the integrator happens *via* a single JSON object. The backend will interact with the GUI's infrastructure only at exactly one point. This will help *e.g.*, to deploy a command-line based version of the integrator engine if it is independent of the GUI. Special attention is paid to data encapsulation. The integrator's driver function that is responsible for setting up all data structures accordingly and calling the actual integration function is only 40 lines of code.

For the creation of the user interface, the Qt software package [341] version 6.3.1 was used. The task of the GUI is to interact with the user. It shall read in the reaction equations and initial conditions. It will take care of the plotting of the results and can be used to export the raw data and figures. The GUI is also used to save sessions and to load previous sessions. The save/load system is based on simple text files.

The second part of the program is responsible for the actual computations. It consists of the appropriate data structures for storing the system's information (such as coefficient matrices) and the needed functions for computing the Jacobian matrix and the right-hand side of the equation system. For the actual integration of the system, the Odeint library contained in the well-known Boost libraries is used [342]. The implicit fourth-order Rosenbrock solver is used together with an adaptive step size. The adaptive step size allows for varying step sizes between individual integration steps, depending on an error estimate. The initial step size and error tolerances have reasonable default values, but can be changed by the user *via* the GUI.

Basic usage

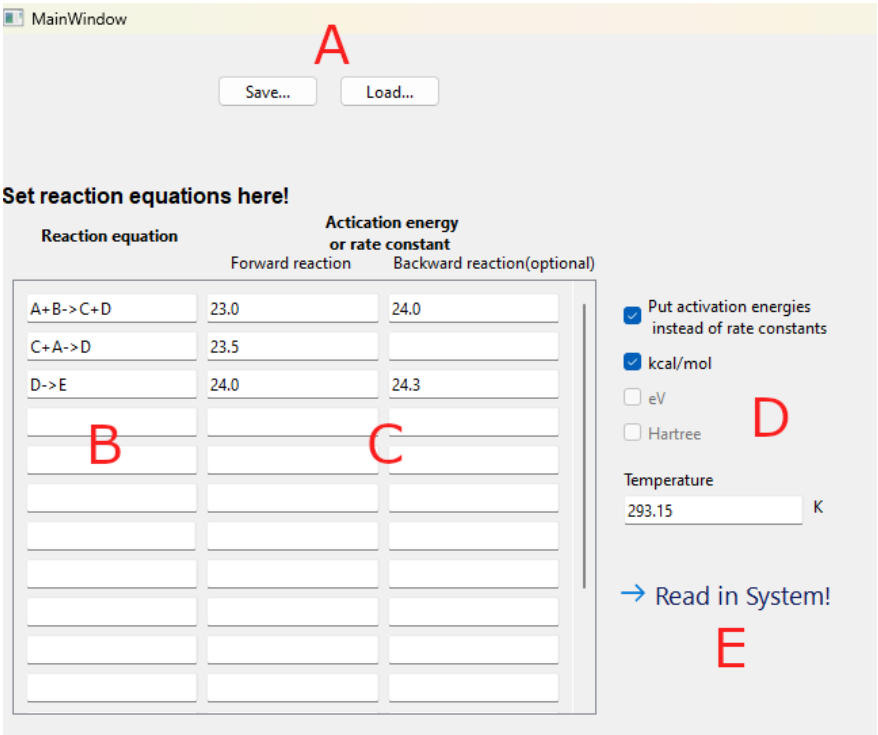

Figure B 5 Main window of the program together with the major control panels (A-E).

In **Figure B 5**, the main window is shown together with the major control panels. The different fields can be used by the user and are explained in the further.

A: the save and load buttons can be used to save the current project or to load previous projects. The save/load mechanism is based on text files that contain all necessary information.

B: in the given fields, the elementary reactions can be specified by the user. "->" is used as separator between educts and products.

C: in these fields, the activation energy or rate constant can be specified for the forward (left) and backward (right) reactions. Specifying forward rates/activation energies is mandatory.

D: if activation energies instead of experimental rate constants shall be used, the corresponding energy unit can be specified here together with the temperature.

E: once the user has specified all reactions and activation energies, the “Read in system” button must be hit.

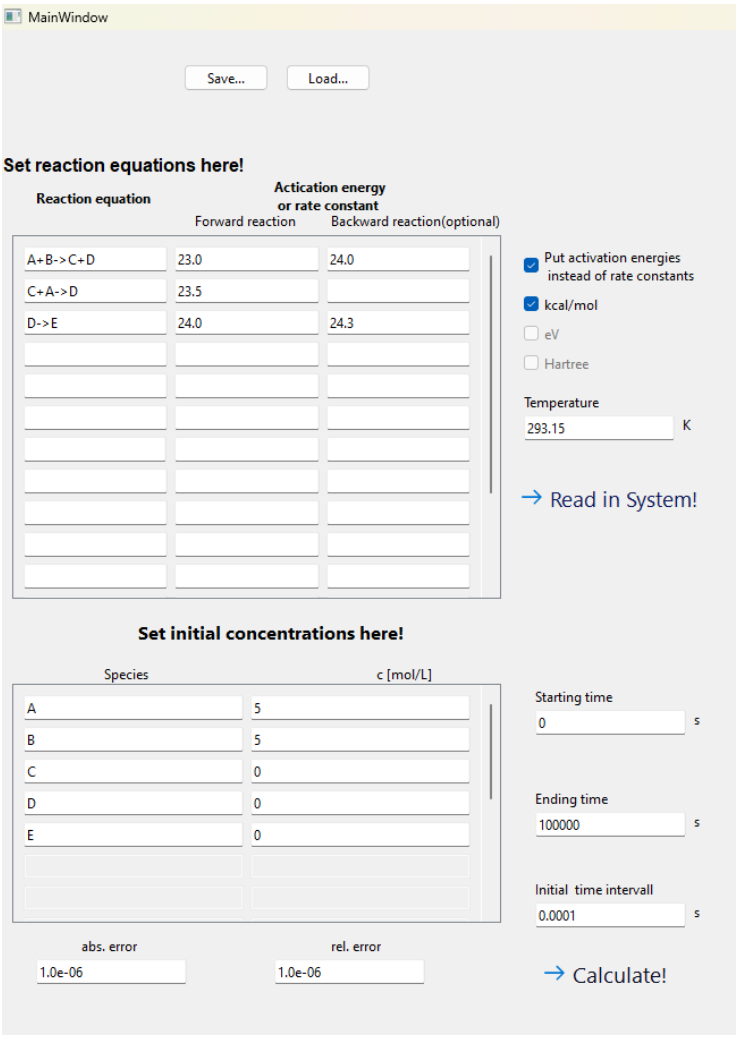

Figure B 6 Main window after the reaction equations were read in.

Once the user has specified all reactions and hit the “Read in system” button, the program identifies the unique species in the system. The user can specify the initial concentration for each species together with some technical parameters, such as the total integration time, the initial time interval or the error thresholds for the solver.

Once the user has hit the “Calculate” button, the system of ODEs is integrated and plotted. The plots are shown in **Figure B 7**. Buttons for logarithmic axes are provided. In addition, buttons

for exporting the shown figure and the raw data (*i.e.*, the concentrations at each time point) are implemented.

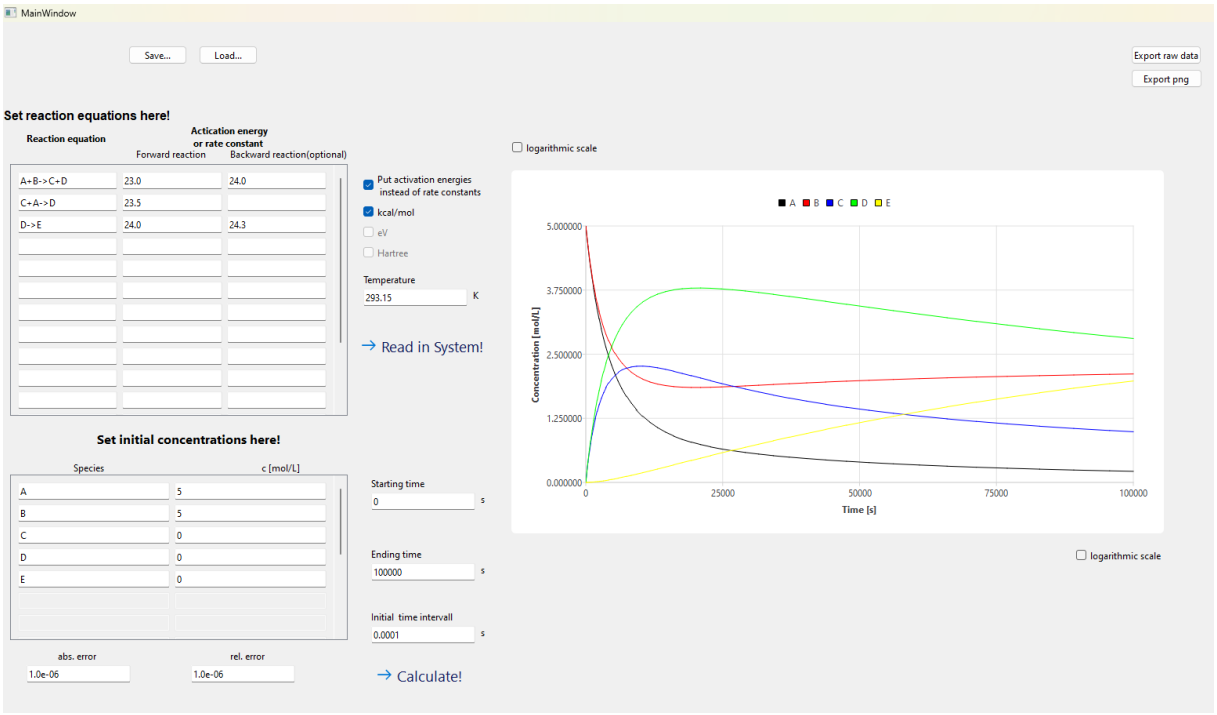

Figure B 7 Main window after the concentration profiles were computed and plotted.

References

[1] Ertl, G. Wilhelm Ostwald: Founder of Physical Chemistry and Nobel Laureate 1909. *Angewandte Chemie International Edition* **2009**, *48* (36), 6600-6606. https://doi.org/10.1002/anie.200901193.
[2] Roduner, E. Understanding catalysis. *Chemical Society Reviews* **2014**, *43* (24), 8226-8239. https://doi.org/10.1039/C4CS00210E.
[3] Catlow, C. R.; Davidson, M.; Hardacre, C.; Hutchings, G. J. Catalysis making the world a better place. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2016**, *374* (2061), 20150089. https://doi.org/10.1098/rsta.2015.0089.
[4] Dry, M. E. The Fischer–Tropsch process: 1950–2000. *Catalysis Today* **2002**, *71* (3), 227-241. https://doi.org/10.1016/S0920-5861(01)00453-9.
[5] Appl, M. Ammonia, 2. Production Processes. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2011.
[6] Sheldon, R. A. Fundamentals of green chemistry: efficiency in reaction design. *Chemical Society Reviews* **2012**, *41* (4), 1437-1451. https://doi.org/10.1039/C1CS15219J.
[7] Anastas, P.; Eghbali, N. Green Chemistry: Principles and Practice. *Chemical Society Reviews* **2010**, *39* (1), 301-312. https://doi.org/10.1039/B918763B.
[8] Shalf, J. The future of computing beyond Moore's Law. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2020**, *378* (2166), 20190061. https://doi.org/10.1098/rsta.2019.0061.
[9] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chemical Physics Letters* **1995**, *240* (4), 283-290. https://doi.org/10.1016/0009-2614(95)00621-A.
[10] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chemical Physics* **2009**, *356* (1), 98-109. https://doi.org/10.1016/j.chemphys.2008.10.036.
[11] Thiel, W. Computational Catalysis—Past, Present, and Future. *Angewandte Chemie International Edition* **2014**, *53* (33), 8605-8613. https://doi.org/10.1002/anie.201402118.
[12] Dance, I. Computational Investigations of the Chemical Mechanism of the Enzyme Nitrogenase. *ChemBioChem* **2020**, *21* (12), 1671-1709. https://doi.org/10.1002/cbic.201900636.
[13] Harvey, J. N.; Himo, F.; Maseras, F.; Perrin, L. Scope and Challenge of Computational Methods for Studying Mechanism and Reactivity in Homogeneous Catalysis. *ACS Catalysis* **2019**, *9* (8), 6803-6813. https://doi.org/10.1021/acscatal.9b01537.
[14] Funes-Ardoiz, I.; Schoenebeck, F. Established and Emerging Computational Tools to Study Homogeneous Catalysis—From Quantum Mechanics to Machine Learning. *Chem* **2020**, *6* (8), 1904-1913. https://doi.org/10.1016/j.chempr.2020.07.008.
[15] Jensen, F. *Introduction to Computational Chemistry*; Wiley, 2017.
[16] Braibant, S.; Giacomelli, G.; Spurio, M. *Particles and fundamental interactions: An Introduction to Particle Physics*; Springer, 2012.
[17] Reiher, M.; Wolf, A. Relativistic Theory of the Electron. In *Relativistic Quantum Chemistry*, 2014; pp 159-191.
[18] Schrödinger, E. Quantisierung als Eigenwertproblem. *Annalen der Physik* **1926**, *384* (4), 361-376. https://doi.org/10.1002/andp.19263840404.
[19] Ritz, W. Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. **1909**, *1909* (135), 1-61. https://doi.org/10.1515/crll.1909.135.1.

[20] Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik* **1925**, *31* (1), 765-783. https://doi.org/10.1007/BF02980631.
[21] Slater, J. C. The Theory of Complex Spectra. *Physical Review* **1929**, *34* (10), 1293-1322. https://doi.org/10.1103/PhysRev.34.1293.
[22] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, 1996.
[23] Kutzelnigg, W. Das Modell der unabhängigen Teilchen bei Mehrelektronenatomen. In *Einführung in die Theoretische Chemie*, 2001; pp 147-156.
[24] Heisenberg, W. Mehrkörperproblem und Resonanz in der Quantenmechanik. *Zeitschrift für Physik* **1926**, *38* (6), 411-426. https://doi.org/10.1007/BF01397160.
[25] Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1* (1), 104-113. https://doi.org/10.1016/S0031-8914(34)90011-2.
[26] Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23* (2), 69-89. https://doi.org/10.1103/RevModPhys.23.69.
[27] Pople, J. A.; Nesbet, R. K. Self-Consistent Orbitals for Radicals. *The Journal of Chemical Physics* **1954**, *22* (3), 571-572. https://doi.org/10.1063/1.1740120.
[28] Helgaker, T.; Jørgensen, P.; Olsen, J. Spin in Second Quantization. In *Molecular Electronic-Structure Theory*, 2000; pp 34-79.
[29] Helgaker, T.; Jørgensen, P.; Olsen, J. Configuration-Interaction Theory. In *Molecular Electronic-Structure Theory*, 2000; pp 523-597.
[30] Condon, E. U. The Theory of Complex Spectra. *Physical Review* **1930**, *36* (7), 1121-1133. https://doi.org/10.1103/PhysRev.36.1121.
[31] Löwdin, P.-O. Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction. *Physical Review* **1955**, *97* (6), 1474-1489. https://doi.org/10.1103/PhysRev.97.1474.
[32] Löwdin, P.-O. A generalization of Brillouin's theorem and the stability conditions in the quantum-mechanical variation principle in the case of general trial wave functions. *Journal of Chemical Sciences* **1986**, *96* (3), 121-126. https://doi.org/10.1007/BF02974145.
[33] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry* **1976**, *10* (S10), 1-19. https://doi.org/10.1002/qua.560100802.
[34] Coester, F. Bound states of a many-particle system. *Nuclear Physics* **1958**, *7*, 421-424. https://doi.org/10.1016/0029-5582(58)90280-3.
[35] Coester, F.; Kümmel, H. Short-range correlations in nuclear wave functions. *Nuclear Physics* **1960**, *17*, 477-485. https://doi.org/10.1016/0029-5582(60)90140-1.
[36] Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *The Journal of Chemical Physics* **1966**, *45* (11), 4256-4266. https://doi.org/10.1063/1.1727484.
[37] Čížek, J.; Paldus, J. Correlation problems in atomic and molecular systems III. Rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methodst. *International Journal of Quantum Chemistry* **1971**, *5* (4), 359-379. https://doi.org/10.1002/qua.560050402.
[38] Helgaker, T.; Jørgensen, P.; Olsen, J. Coupled-Cluster Theory. In *Molecular Electronic-Structure Theory*, 2000; pp 648-723.
[39] Kvaal, S. Variational formulations of the coupled-cluster method in quantum chemistry. *Molecular Physics* **2013**, *111* (9-11), 1100-1108. https://doi.org/10.1080/00268976.2013.812254.

[40] Voorhis, T. V.; Head-Gordon, M. Benchmark variational coupled cluster doubles results. *The Journal of Chemical Physics* **2000**, *113* (20), 8873-8879. https://doi.org/10.1063/1.1319643.
[41] Robinson, J. B.; Knowles, P. J. Quasi-variational coupled cluster theory. *The Journal of Chemical Physics* **2012**, *136* (5), 054114. https://doi.org/10.1063/1.3680560.
[42] Robinson, J. B.; Knowles, P. J. Application of the quasi-variational coupled cluster method to the nonlinear optical properties of model hydrogen systems. *The Journal of Chemical Physics* **2012**, *137* (5), 054301. https://doi.org/10.1063/1.4738758.
[43] Crawford, T. D.; Schaefer III, H. F. An Introduction to Coupled Cluster Theory for Computational Chemists. In *Reviews in Computational Chemistry*, 2000; pp 33-136.
[44] Achilles, R.; Bonfiglioli, A. The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin. *Archive for History of Exact Sciences* **2012**, *66* (3), 295-358. https://doi.org/10.1007/s00407-012-0095-8.
[45] Monkhorst, H. J. Calculation of properties with the coupled-cluster method. *International Journal of Quantum Chemistry* **1977**, *12* (S11), 421-432. https://doi.org/10.1002/qua.560120850.
[46] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters* **1989**, *157* (6), 479-483. https://doi.org/10.1016/S0009-2614(89)87395-6.
[47] Pulay, P. Localizability of dynamic electron correlation. *Chemical Physics Letters* **1983**, *100* (2), 151-154. https://doi.org/10.1016/0009-2614(83)80703-9.
[48] Pulay, P.; Sæbø, S.; Meyer, W. An efficient reformulation of the closed-shell self-consistent electron pair theory. *The Journal of Chemical Physics* **1984**, *81* (4), 1901-1905. https://doi.org/10.1063/1.447863.
[49] Sæbø, S.; Pulay, P. Fourth-order Møller–Plessett perturbation theory in the local correlation treatment. I. Method. *The Journal of Chemical Physics* **1987**, *86* (2), 914-922. https://doi.org/10.1063/1.452293.
[50] Sæbø, S.; Pulay, P. The local correlation treatment. II. Implementation and tests. *The Journal of Chemical Physics* **1988**, *88* (3), 1884-1890. https://doi.org/10.1063/1.454111.
[51] Sæbø, S.; Pulay, P. Local configuration interaction: An efficient approach for larger molecules. *Chemical Physics Letters* **1985**, *113* (1), 13-18. https://doi.org/10.1016/0009-2614(85)85003-X.
[52] Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *The Journal of Chemical Physics* **1999**, *111* (13), 5691-5705. https://doi.org/10.1063/1.479957.
[53] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review* **1934**, *46* (7), 618-622. https://doi.org/10.1103/PhysRev.46.618.
[54] Schütz, M.; Werner, H.-J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *The Journal of Chemical Physics* **2001**, *114* (2), 661-681. https://doi.org/10.1063/1.1330207.
[55] Schütz, M.; Werner, H.-J. Local perturbative triples correction (T) with linear cost scaling. *Chemical Physics Letters* **2000**, *318* (4), 370-378. https://doi.org/10.1016/S0009-2614(00)00066-X.
[56] Löwdin, P.-O.; Shull, H. Natural Orbitals in the Quantum Theory of Two-Electron Systems. *Physical Review* **1956**, *101* (6), 1730-1739. https://doi.org/10.1103/PhysRev.101.1730.
[57] McWeeny, R. Some Recent Advances in Density Matrix Theory. *Reviews of Modern Physics* **1960**, *32* (2), 335-369. https://doi.org/10.1103/RevModPhys.32.335.
[58] Shavitt, I.; Rosenberg, B. J.; Palalikit, S. Comparison of configuration interaction expansions based on different orbital transformations. *International Journal of Quantum Chemistry* **1976**, *10* (S10), 33-46. https://doi.org/10.1002/qua.560100804.

[59] Meyer, W. Ionization energies of water from PNO-CI calculations. *International Journal of Quantum Chemistry* **1971**, *5* (S5), 341-348. https://doi.org/10.1002/qua.560050839.
[60] Meyer, W. PNO–CI Studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. *The Journal of Chemical Physics* **1973**, *58* (3), 1017-1035. https://doi.org/10.1063/1.1679283.
[61] Ahlrichs, R.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. PNO–CI (pair natural orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. I. Outline of the method for closed-shell states. *The Journal of Chemical Physics* **1975**, *62* (4), 1225-1234. https://doi.org/10.1063/1.430637.
[62] Ahlrichs, R.; Driessler, F.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. PNO–CI (pair natural orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. II. The molecules BeH2, BH, BH3, CH4, CH−3, NH3 (planar and pyramidal), H2O, OH+3, HF and the Ne atom. *The Journal of Chemical Physics* **1975**, *62* (4), 1235-1247. https://doi.org/10.1063/1.430638.
[63] Ahlrichs, R.; Keil, F.; Lischka, H.; Kutzelnigg, W.; Staemmler, V. PNO–CI (pair natural-orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. III. The molecules MgH2, AlH3, SiH4, PH3 (planar and pyramidal), H2S, HCl, and the Ar atom. *The Journal of Chemical Physics* **1975**, *63* (1), 455-463. https://doi.org/10.1063/1.431073.
[64] Ahlrichs, R.; Lischka, H.; Zurawski, B.; Kutzelnigg, W. PNO–CI (pair-natural-orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. IV. The molecules N2, F2, C2H2, C2H4, and C2H6. *The Journal of Chemical Physics* **1975**, *63* (11), 4685-4694. https://doi.org/10.1063/1.431254.
[65] Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. Accurate Theoretical Chemistry with Coupled Pair Models. *Accounts of Chemical Research* **2009**, *42* (5), 641-648. https://doi.org/10.1021/ar800241t.
[66] Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *The Journal of Chemical Physics* **2009**, *131* (6), 064103. https://doi.org/10.1063/1.3173827.
[67] Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *The Journal of Chemical Physics* **2009**, *130* (11), 114108. https://doi.org/10.1063/1.3086717.
[68] Hansen, A.; Liakos, D. G.; Neese, F. Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals. *The Journal of Chemical Physics* **2011**, *135* (21), 214102. https://doi.org/10.1063/1.3663855.
[69] Liakos, D. G.; Neese, F. Improved Correlation Energy Extrapolation Schemes Based on Local Pair Natural Orbital Methods. *The Journal of Physical Chemistry A* **2012**, *116* (19), 4801-4816. https://doi.org/10.1021/jp302096v.
[70] Liakos, D. G.; Hansen, A.; Neese, F. Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods. *Journal of Chemical Theory and Computation* **2011**, *7* (1), 76-87. https://doi.org/10.1021/ct100445s.
[71] Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *The Journal of Chemical Physics* **2015**, *143* (3), 034108. https://doi.org/10.1063/1.4926879.
[72] Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based

pair natural orbital coupled cluster theory. *The Journal of Chemical Physics* **2016**, *144* (2), 024109. https://doi.org/10.1063/1.4939030.
[73] Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *The Journal of Chemical Physics* **2013**, *139* (13), 134101. https://doi.org/10.1063/1.4821834.
[74] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *The Journal of Chemical Physics* **2018**, *148* (1), 011101. https://doi.org/10.1063/1.5011798.
[75] Altun, A.; Neese, F.; Bistoni, G. HFLD: A Nonempirical London Dispersion-Corrected Hartree–Fock Method for the Quantification and Analysis of Noncovalent Interaction Energies of Large Molecular Systems. *Journal of Chemical Theory and Computation* **2019**, *15* (11), 5894-5907. https://doi.org/10.1021/acs.jctc.9b00425.
[76] Altun, A.; Neese, F.; Bistoni, G. Open-Shell Variant of the London Dispersion-Corrected Hartree–Fock Method (HFLD) for the Quantification and Analysis of Noncovalent Interaction Energies. *Journal of Chemical Theory and Computation* **2022**, *18* (4), 2292-2307. https://doi.org/10.1021/acs.jctc.1c01295.
[77] Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **2020**, *152* (22), 224108. https://doi.org/10.1063/5.0004608.
[78] Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12* (5), e1606. https://doi.org/10.1002/wcms.1606.
[79] Foster, J. M.; Boys, S. F. Canonical Configurational Interaction Procedure. *Reviews of Modern Physics* **1960**, *32* (2), 300-302. https://doi.org/10.1103/RevModPhys.32.300.
[80] Neese, F. *ORCA Manual Version 5.0.3*; 2022.
[81] Liakos, D. G.; Guo, Y.; Neese, F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *The Journal of Physical Chemistry A* **2020**, *124* (1), 90-100. https://doi.org/10.1021/acs.jpca.9b05734.
[82] Gottschalk, H. C.; Poblotzki, A.; Suhm, M. A.; Al-Mogren, M. M.; Antony, J.; Auer, A. A.; Baptista, L.; Benoit, D. M.; Bistoni, G.; Bohle, F.; et al. The furan microsolvation blind challenge for quantum chemical methods: First steps. *The Journal of Chemical Physics* **2018**, *148* (1), 014301. https://doi.org/10.1063/1.5009011.
[83] Seeger, Z. L.; Izgorodina, E. I. A DLPNO-CCSD(T) benchmarking study of intermolecular interactions of ionic liquids. *Journal of Computational Chemistry* **2022**, *43* (2), 106-120. https://doi.org/10.1002/jcc.26776.
[84] Minenkov, Y.; Chermak, E.; Cavallo, L. Accuracy of DLPNO–CCSD(T) Method for Noncovalent Bond Dissociation Enthalpies from Coinage Metal Cation Complexes. *Journal of Chemical Theory and Computation* **2015**, *11* (10), 4664-4676. https://doi.org/10.1021/acs.jctc.5b00584.
[85] Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics* **2017**, *19* (48), 32184-32215. https://doi.org/10.1039/C7CP04913G.
[86] Altun, A.; Neese, F.; Bistoni, G. Extrapolation to the Limit of a Complete Pair Natural Orbital Space in Local Coupled-Cluster Calculations. *Journal of Chemical Theory and Computation* **2020**, *16* (10), 6142-6149. https://doi.org/10.1021/acs.jctc.0c00344.
[87] Neese, F.; Atanasov, M.; Bistoni, G.; Maganas, D.; Ye, S. Chemistry and Quantum Mechanics in 2019: Give Us Insight and Numbers. *Journal of the American Chemical Society* **2019**, *141* (7), 2814-2824. https://doi.org/10.1021/jacs.8b13313.

[88] Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. Decomposition of Intermolecular Interaction Energies within the Local Pair Natural Orbital Coupled Cluster Framework. *Journal of Chemical Theory and Computation* **2016**, *12* (10), 4778-4792. https://doi.org/10.1021/acs.jctc.6b00523.
[89] Altun, A.; Saitow, M.; Neese, F.; Bistoni, G. Local Energy Decomposition of Open-Shell Molecular Systems in the Domain-Based Local Pair Natural Orbital Coupled Cluster Framework. *Journal of Chemical Theory and Computation* **2019**, *15* (3), 1616-1632. https://doi.org/10.1021/acs.jctc.8b01145.
[90] Bistoni, G. Finding chemical concepts in the Hilbert space: Coupled cluster analyses of noncovalent interactions. *WIREs Computational Molecular Science* **2020**, *10* (3), e1442. https://doi.org/10.1002/wcms.1442.
[91] Altun, A.; Neese, F.; Bistoni, G. Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study. *Beilstein Journal of Organic Chemistry* **2018**, *14*, 919-929. https://doi.org/10.3762/bjoc.14.79.
[92] Altun, A.; Izsák, R.; Bistoni, G. Local energy decomposition of coupled-cluster interaction energies: Interpretation, benchmarks, and comparison with symmetry-adapted perturbation theory. *International Journal of Quantum Chemistry* **2021**, *121* (3), e26339. https://doi.org/10.1002/qua.26339.
[93] Lu, Q.; Neese, F.; Bistoni, G. London dispersion effects in the coordination and activation of alkanes in σ-complexes: a local energy decomposition study. *Physical Chemistry Chemical Physics* **2019**, *21* (22), 11569-11577. https://doi.org/10.1039/C9CP01309A.
[94] Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chemical Reviews* **1994**, *94* (7), 1887-1930. https://doi.org/10.1021/cr00031a008.
[95] Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *The Journal of Chemical Physics* **2014**, *140* (9), 094106. https://doi.org/10.1063/1.4867135.
[96] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Physical Review* **1964**, *136* (3B), B864-B871. https://doi.org/10.1103/PhysRev.136.B864.
[97] Kohn, W. Nobel Lecture: Electronic structure of matter---wave functions and density functionals. *Reviews of Modern Physics* **1999**, *71* (5), 1253-1266. https://doi.org/10.1103/RevModPhys.71.1253.
[98] Koch, W.; Holthausen, M. C. Electron Density and Hole Functions. In *A Chemist's Guide to Density Functional Theory*, 2001; pp 19-28.
[99] Koch, W.; Holthausen, M. C. The Hohenberg-Kohn Theorems. In *A Chemist's Guide to Density Functional Theory*, 2001; pp 33-40.
[100] Koch, W.; Holthausen, M. C. The Kohn-Sham Approach. In *A Chemist's Guide to Density Functional Theory*, 2001; pp 41-64.
[101] Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings* **2001**, *577* (1), 1-20. https://doi.org/10.1063/1.1390175.
[102] Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B* **1981**, *23* (10), 5048-5079. https://doi.org/10.1103/PhysRevB.23.5048.
[103] Ortiz, G.; Ballone, P. Correlation energy, structure factor, radial distribution function, and momentum distribution of the spin-polarized uniform electron gas. *Physical Review B* **1994**, *50* (3), 1391-1405. https://doi.org/10.1103/PhysRevB.50.1391.
[104] Jones, R. O.; Gunnarsson, O. The density functional formalism, its applications and prospects. *Reviews of Modern Physics* **1989**, *61* (3), 689-746. https://doi.org/10.1103/RevModPhys.61.689.

[105] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **1988**, *38* (6), 3098-3100. https://doi.org/10.1103/PhysRevA.38.3098.
[106] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988**, *37* (2), 785-789. https://doi.org/10.1103/PhysRevB.37.785.
[107] Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chemical Physics Letters* **1989**, *157* (3), 200-206. https://doi.org/10.1016/0009-2614(89)87234-3.
[108] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77* (18), 3865-3868. https://doi.org/10.1103/PhysRevLett.77.3865.
[109] Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B* **1992**, *46* (11), 6671-6687. https://doi.org/10.1103/PhysRevB.46.6671.
[110] 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. *Physical Review Letters* **2003**, *91* (14), 146401. https://doi.org/10.1103/PhysRevLett.91.146401.
[111] Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *The Journal of Chemical Physics* **2006**, *125* (19), 194101. https://doi.org/10.1063/1.2370993.
[112] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98* (7), 5648-5652. https://doi.org/10.1063/1.464913.
[113] 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. *The Journal of Physical Chemistry* **1994**, *98* (45), 11623-11627. https://doi.org/10.1021/j100096a001.
[114] Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *The Journal of Chemical Physics* **1999**, *110* (11), 5029-5036. https://doi.org/10.1063/1.478401.
[115] Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics* **1999**, *110* (13), 6158-6170. https://doi.org/10.1063/1.478522.
[116] Goerigk, L.; Grimme, S. Double-hybrid density functionals. *WIREs Computational Molecular Science* **2014**, *4* (6), 576-600. https://doi.org/10.1002/wcms.1193.
[117] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *The Journal of Chemical Physics* **2001**, *115* (8), 3540-3544. https://doi.org/10.1063/1.1383587.
[118] Toulouse, J.; Colonna, F.; Savin, A. Long-range--short-range separation of the electron-electron interaction in density-functional theory. *Physical Review A* **2004**, *70* (6), 062505. https://doi.org/10.1103/PhysRevA.70.062505.
[119] Brémond, É.; Pérez-Jiménez, Á. J.; Sancho-García, J. C.; Adamo, C. Range-separated hybrid density functionals made simple. *The Journal of Chemical Physics* **2019**, *150* (20), 201102. https://doi.org/10.1063/1.5097164.
[120] Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. Range Separation and Local Hybridization in Density Functional Theory. *The Journal of Physical Chemistry A* **2008**, *112* (49), 12530-12542. https://doi.org/10.1021/jp806573k.

[121] Bursch, M.; Mewes, J.-M.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**. *Angewandte Chemie International Edition* **2022**, *61* (42), e202205735. https://doi.org/10.1002/anie.202205735.
[122] Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *Journal of Computational Chemistry* **2013**, *34* (19), 1672-1685. https://doi.org/10.1002/jcc.23317.
[123] Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *The Journal of Chemical Physics* **2015**, *143* (5), 054107. https://doi.org/10.1063/1.4927476.
[124] Brandenburg, J. G.; Bannwarth, C.; Hansen, A.; Grimme, S. B97-3c: A revised low-cost variant of the B97-D density functional method. *The Journal of Chemical Physics* **2018**, *148* (6), 064104. https://doi.org/10.1063/1.5012601.
[125] Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M. r2SCAN-3c: A "Swiss army knife" composite electronic-structure method. *The Journal of Chemical Physics* **2021**, *154* (6), 064103. https://doi.org/10.1063/5.0040021.
[126] Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *The Journal of Chemical Physics* **2012**, *136* (15), 154101. https://doi.org/10.1063/1.3700154.
[127] Kristyán, S.; Pulay, P. Can (semi)local density functional theory account for the London dispersion forces? *Chemical Physics Letters* **1994**, *229* (3), 175-180. https://doi.org/10.1016/0009-2614(94)01027-7.
[128] Hobza, P.; šponer, J.; Reschel, T. Density functional theory and molecular clusters. *Journal of Computational Chemistry* **1995**, *16* (11), 1315-1325. https://doi.org/10.1002/jcc.540161102.
[129] London, F. Zur Theorie und Systematik der Molekularkräfte. *Zeitschrift für Physik* **1930**, *63* (3), 245-279. https://doi.org/10.1007/BF01421741.
[130] Pitzer, K. S.; Catalano, E. Electronic Correlation in Molecules. III. The Paraffin Hydrocarbons1. *Journal of the American Chemical Society* **1956**, *78* (19), 4844-4846. https://doi.org/10.1021/ja01600a006.
[131] Wagner, J. P.; Schreiner, P. R. London Dispersion in Molecular Chemistry—Reconsidering Steric Effects. *Angewandte Chemie International Edition* **2015**, *54* (42), 12274-12296. https://doi.org/10.1002/anie.201503476.
[132] Gramüller, J.; Franta, M.; Gschwind, R. M. Tilting the Balance: London Dispersion Systematically Enhances Enantioselectivities in Brønsted Acid Catalyzed Transfer Hydrogenation of Imines. *Journal of the American Chemical Society* **2022**, *144* (43), 19861-19871. https://doi.org/10.1021/jacs.2c07563.
[133] Klimeš, J.; Michaelides, A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *The Journal of Chemical Physics* **2012**, *137* (12), 120901. https://doi.org/10.1063/1.4754130.
[134] Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006**, *27* (15), 1787-1799. https://doi.org/10.1002/jcc.20495.
[135] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.
[136] Casimir, H. B. G.; Polder, D. The Influence of Retardation on the London-van der Waals Forces. *Physical Review* **1948**, *73* (4), 360-372. https://doi.org/10.1103/PhysRev.73.360.
[137] Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32* (7), 1456-1465. https://doi.org/10.1002/jcc.21759.

[138] Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *The Journal of Chemical Physics* **2005**, *123* (15), 154101. https://doi.org/10.1063/1.2065267.
[139] Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *The Journal of Chemical Physics* **2017**, *147* (3), 034112. https://doi.org/10.1063/1.4993215.
[140] Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics* **2019**, *150* (15), 154122. https://doi.org/10.1063/1.5090222.
[141] Vydrov, O. A.; Voorhis, T. V. Nonlocal van der Waals density functional: The simpler the better. *The Journal of Chemical Physics* **2010**, *133* (24), 244103. https://doi.org/10.1063/1.3521275.
[142] Vydrov, O. A.; Voorhis, T. V. Improving the accuracy of the nonlocal van der Waals density functional with minimal empiricism. *The Journal of Chemical Physics* **2009**, *130* (10), 104105. https://doi.org/10.1063/1.3079684.
[143] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals Density Functional Made Simple. *Physical Review Letters* **2009**, *103* (6), 063004. https://doi.org/10.1103/PhysRevLett.103.063004.
[144] Vydrov, O. A.; Van Voorhis, T. Dispersion interactions from a local polarizability model. *Physical Review A* **2010**, *81* (6), 062708. https://doi.org/10.1103/PhysRevA.81.062708.
[145] Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *The Journal of Chemical Physics* **2015**, *142* (7), 074111. https://doi.org/10.1063/1.4907719.
[146] Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Physical Chemistry Chemical Physics* **2014**, *16* (21), 9904-9924. https://doi.org/10.1039/C3CP54374A.
[147] Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *The Journal of Chemical Physics* **2016**, *144* (21), 214110. https://doi.org/10.1063/1.4952647.
[148] Najibi, A.; Goerigk, L. The Nonlocal Kernel in van der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ωB97M-V Approaches. *Journal of Chemical Theory and Computation* **2018**, *14* (11), 5725-5738. https://doi.org/10.1021/acs.jctc.8b00842.
[149] Hujo, W.; Grimme, S. Performance of Non-Local and Atom-Pairwise Dispersion Corrections to DFT for Structural Parameters of Molecules with Noncovalent Interactions. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 308-315. https://doi.org/10.1021/ct300813c.
[150] Iron, M. A.; Janes, T. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *The Journal of Physical Chemistry A* **2019**, *123* (17), 3761-3781. https://doi.org/10.1021/acs.jpca.9b01546.
[151] Pracht, P.; Bohle, F.; Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics* **2020**, *22* (14), 7169-7192. https://doi.org/10.1039/C9CP06869D.
[152] Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *Journal of Chemical Theory and Computation* **2019**, *15* (5), 2847-2862. https://doi.org/10.1021/acs.jctc.9b00143.
[153] Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **2002**, *99* (20), 12562-12566. https://doi.org/10.1073/pnas.202427399.

[154] Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86). *Journal of Chemical Theory and Computation* **2017**, *13* (5), 1989-2009. https://doi.org/10.1021/acs.jctc.7b00118.
[155] Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* **2019**, *15* (3), 1652-1671. https://doi.org/10.1021/acs.jctc.8b01176.
[156] Spicher, S.; Grimme, S. Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems. *Angewandte Chemie International Edition* **2020**, *59* (36), 15665-15673. https://doi.org/10.1002/anie.202004239.
[157] Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science* **2021**, *11* (2), e1493. https://doi.org/10.1002/wcms.1493.
[158] Grimme, S.; Bohle, F.; Hansen, A.; Pracht, P.; Spicher, S.; Stahn, M. Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Nonrigid Molecules. *The Journal of Physical Chemistry A* **2021**, *125* (19), 4039-4054. https://doi.org/10.1021/acs.jpca.1c00971.
[159] Rosales, A. R.; Quinn, T. R.; Wahlers, J.; Tomberg, A.; Zhang, X.; Helquist, P.; Wiest, O.; Norrby, P.-O. Application of Q2MM to predictions in stereoselective synthesis. *Chemical Communications* **2018**, *54* (60), 8294-8311. https://doi.org/10.1039/C8CC03695K.
[160] Rosales, A. R.; Wahlers, J.; Limé, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; et al. Rapid virtual screening of enantioselective catalysts using CatVS. *Nature Catalysis* **2019**, *2* (1), 41-45. https://doi.org/10.1038/s41929-018-0193-3.
[161] Eksterowicz, J. E.; Houk, K. N. Transition-state modeling with empirical force fields. *Chemical Reviews* **1993**, *93* (7), 2439-2461. https://doi.org/10.1021/cr00023a006.
[162] Doney, A. C.; Rooks, B. J.; Lu, T.; Wheeler, S. E. Design of Organocatalysts for Asymmetric Propargylations through Computational Screening. *ACS Catalysis* **2016**, *6* (11), 7948-7955. https://doi.org/10.1021/acscatal.6b02366.
[163] Guan, Y.; Ingman, V. M.; Rooks, B. J.; Wheeler, S. E. AARON: An Automated Reaction Optimizer for New Catalysts. *Journal of Chemical Theory and Computation* **2018**, *14* (10), 5249-5261. https://doi.org/10.1021/acs.jctc.8b00578.
[164] *Gaussian 16 Rev. C.01*; Wallingford, CT, 2016. https://gaussian.com/ (accessed 12.04.2023).
[165] Jensen, F. Locating minima on seams of intersecting potential energy surfaces. An application to transition structure modeling. *Journal of the American Chemical Society* **1992**, *114* (5), 1596-1603. https://doi.org/10.1021/ja00031a009.
[166] Jensen, F. Transition structure modeling by intersecting potential energy surfaces. *Journal of Computational Chemistry* **1994**, *15* (11), 1199-1216. https://doi.org/10.1002/jcc.540151103.
[167] Jensen, F.; Norrby, P.-O. Transition states from empirical force fields. *Theoretical Chemistry Accounts* **2003**, *109* (1), 1-7. https://doi.org/10.1007/s00214-002-0382-6.
[168] Jensen, F. Using force fields methods for locating transition structures. *The Journal of Chemical Physics* **2003**, *119* (17), 8804-8808. https://doi.org/10.1063/1.1613931.
[169] Carlone, A.; Bernardi, L. Enantioselective organocatalytic approaches to active pharmaceutical ingredients – selected industrial examples. *Physical Sciences Reviews* **2019**, *4* (8). https://doi.org/10.1515/psr-2018-0097.
[170] Pizzarello, S.; Weber, A. L. Prebiotic Amino Acids as Asymmetric Catalysts. *Science* **2004**, *303* (5661), 1151-1151. https://doi.org/10.1126/science.1093057.

[171] von Liebig, J. Ueber die Bildung des Oxamids aus Cyan. *Justus Liebigs Annalen der Chemie* **1860**, *113* (2), 246-247. https://doi.org/10.1002/jlac.18601130213.
[172] Dalko, P. I. Asymmetric Organocatalysis: A New Stream in Organic Synthesis. In *Enantioselective Organocatalysis*, 2007; pp 1-17.
[173] Tian, S.-K.; Chen, Y.; Hang, J.; Tang, L.; McDaid, P.; Deng, L. Asymmetric Organic Catalysis with Modified Cinchona Alkaloids. *Accounts of Chemical Research* **2004**, *37* (8), 621-631. https://doi.org/10.1021/ar030048s.
[174] Wegler, R. Über die mit verschiedener Reaktionsgeschwindigkeit erfolgende Veresterung der optischen Antipoden eines Racemates durch opt. akt. Katalysatoren. *Justus Liebigs Annalen der Chemie* **1932**, *498* (1), 62-76. https://doi.org/10.1002/jlac.19324980105.
[175] Prelog, V.; Wilhelm, M. Untersuchungen über asymmetrische Synthesen VI. Der Reaktionsmechanismus und der sterische Verlauf der asymmetrischen Cyanhydrin-Synthese. *Helvetica Chimica Acta* **1954**, *37* (6), 1634-1660. https://doi.org/10.1002/hlca.19540370608.
[176] Pracejus, H. Organische Katalysatoren, LXI. Asymmetrische Synthesen mit Ketenen, I. Alkaloid-katalysierte asymmetrische Synthesen von α-Phenyl-propionsäureestern. *Justus Liebigs Annalen der Chemie* **1960**, *634* (1), 9-22. https://doi.org/10.1002/jlac.19606340103.
[177] Obenius, U.; Bergson, G. On an Asymmetric Aldol-Type Reaction. *Acta Chem. Scand.* **1972**, *26* (6), 2546-2547. https://doi.org/10.3891/acta.chem.scand.26-2546.
[178] Langström, B.; Bergson, G. Asymmetric Induction in a Michael-type Reaction. *Acta Chem. Scand.* **1973**, *27* (8), 3118-3119. https://doi.org/10.3891/acta.chem.scand.27-3118.
[179] Mannich, C.; Davidsen, H. Über einfache Enamine mit tertiär gebundenem Stickstoff. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1936**, *69* (9), 2106-2112. https://doi.org/10.1002/cber.19360690921.
[180] Stork, G.; Terrell, R.; Szmuszkovicz, J. A NEW SYNTHESIS OF 2-ALKYL AND 2-ACYL KETONES. *Journal of the American Chemical Society* **1954**, *76* (7), 2029-2030. https://doi.org/10.1021/ja01636a103.
[181] Hajos, Z. G.; Parrish, D. R. Asymmetric synthesis of bicyclic intermediates of natural product chemistry. *The Journal of Organic Chemistry* **1974**, *39* (12), 1615-1621. https://doi.org/10.1021/jo00925a003.
[182] Eder, U.; Sauer, G.; Wiechert, R. New Type of Asymmetric Cyclization to Optically Active Steroid CD Partial Structures. *Angewandte Chemie International Edition in English* **1971**, *10* (7), 496-497. https://doi.org/10.1002/anie.197104961.
[183] Job, A.; Janeck, C. F.; Bettray, W.; Peters, R.; Enders, D. The SAMP-/RAMP-hydrazone methodology in asymmetric synthesis. *Tetrahedron* **2002**, *58* (12), 2253-2329. https://doi.org/10.1016/S0040-4020(02)00080-7.
[184] Franzén, J.; Marigo, M.; Fielenbach, D.; Wabnitz, T. C.; Kjærsgaard, A.; Jørgensen, K. A. A General Organocatalyst for Direct α-Functionalization of Aldehydes: Stereoselective C−C, C−N, C−F, C−Br, and C−S Bond-Forming Reactions. Scope and Mechanistic Insights. *Journal of the American Chemical Society* **2005**, *127* (51), 18296-18304. https://doi.org/10.1021/ja056120u.
[185] Northrup, A. B.; MacMillan, D. W. C. The First Direct and Enantioselective Cross-Aldol Reaction of Aldehydes. *Journal of the American Chemical Society* **2002**, *124* (24), 6798-6799. https://doi.org/10.1021/ja0262378.
[186] Northrup, A. B.; Mangion, I. K.; Hettche, F.; MacMillan, D. W. C. Enantioselective Organocatalytic Direct Aldol Reactions of α-Oxyaldehydes: Step One in a Two-Step Synthesis of Carbohydrates. *Angewandte Chemie International Edition* **2004**, *43* (16), 2152-2154. https://doi.org/10.1002/anie.200453716.
[187] Notz, W.; List, B. Catalytic Asymmetric Synthesis of anti-1,2-Diols. *Journal of the American Chemical Society* **2000**, *122* (30), 7386-7387. https://doi.org/10.1021/ja001460v.

[188] Clemente, F. R.; Houk, K. N. Theoretical Studies of Stereoselectivities of Intramolecular Aldol Cyclizations Catalyzed by Amino Acids. *Journal of the American Chemical Society* **2005**, *127* (32), 11294-11302. https://doi.org/10.1021/ja0507620.
[189] Inomata, K.; Barragué, M.; Paquette, L. A. Diastereoselectivities Realized in the Amino Acid Catalyzed Aldol Cyclizations of Triketo Acetonides of Differing Ring Size. *The Journal of Organic Chemistry* **2005**, *70* (2), 533-539. https://doi.org/10.1021/jo0486084.
[190] Córdova, A.; Watanabe, S.-i.; Tanaka, F.; Notz, W.; Barbas, C. F. A Highly Enantioselective Route to Either Enantiomer of Both α- and β-Amino Acid Derivatives. *Journal of the American Chemical Society* **2002**, *124* (9), 1866-1867. https://doi.org/10.1021/ja017833p.
[191] Chowdari, N. S.; Suri, J. T.; Barbas, C. F. Asymmetric Synthesis of Quaternary α- and β-Amino Acids and β-Lactams via Proline-Catalyzed Mannich Reactions with Branched Aldehyde Donors. *Organic Letters* **2004**, *6* (15), 2507-2510. https://doi.org/10.1021/ol049248+.
[192] Mitsumori, S.; Zhang, H.; Ha-Yeon Cheong, P.; Houk, K. N.; Tanaka, F.; Barbas, C. F. Direct Asymmetric anti-Mannich-Type Reactions Catalyzed by a Designed Amino Acid. *Journal of the American Chemical Society* **2006**, *128* (4), 1040-1041. https://doi.org/10.1021/ja056984f.
[193] Notz, W.; Tanaka, F.; Watanabe, S.-i.; Chowdari, N. S.; Turner, J. M.; Thayumanavan, R.; Barbas, C. F. The Direct Organocatalytic Asymmetric Mannich Reaction: Unmodified Aldehydes as Nucleophiles. *The Journal of Organic Chemistry* **2003**, *68* (25), 9624-9634. https://doi.org/10.1021/jo0347359.
[194] Marigo, M.; Jørgensen, K. A. Enamine Catalysis: α-Heteroatom Functionalization. In *Enantioselective Organocatalysis*, 2007; pp 56-76.
[195] List, B.; Pojarliev, P.; Martin, H. J. Efficient Proline-Catalyzed Michael Additions of Unmodified Ketones to Nitro Olefins. *Organic Letters* **2001**, *3* (16), 2423-2425. https://doi.org/10.1021/ol015799d.
[196] Enders, D.; Seki, A. Proline-Catalyzed Enantioselective Michael Additions of Ketones to Nitrostyrene. *Synlett* **2002**, *2002* (01), 0026-0028. https://doi.org/10.1055/s-2002-19336.
[197] Akiyama, T.; Itoh, J.; Fuchibe, K. Recent Progress in Chiral Brønsted Acid Catalysis. *Advanced Synthesis & Catalysis* **2006**, *348* (9), 999-1010. https://doi.org/10.1002/adsc.200606074.
[198] McDougal, N. T.; Schaus, S. E. Asymmetric Morita−Baylis−Hillman Reactions Catalyzed by Chiral Brønsted Acids. *Journal of the American Chemical Society* **2003**, *125* (40), 12094-12095. https://doi.org/10.1021/ja037705w.
[199] McDougal, N. T.; Trevellini, W. L.; Rodgen, S. A.; Kliman, L. T.; Schaus, S. E. The Development of the Asymmetric Morita—Baylis—Hillman Reaction Catalyzed by Chiral Brønsted Acids. *Advanced Synthesis & Catalysis* **2004**, *346* (9-10), 1231-1240. https://doi.org/10.1002/adsc.200404122.
[200] Matsui, K.; Takizawa, S.; Sasai, H. Bifunctional Organocatalysts for Enantioselective aza-Morita−Baylis−Hillman Reaction. *Journal of the American Chemical Society* **2005**, *127* (11), 3680-3681. https://doi.org/10.1021/ja0500254.
[201] Mori, K.; Akiyama, T. Brønsted Acids: Chiral Phosphoric Acid Catalysts in Asymmetric Synthesis. In *Comprehensive Enantioselective Organocatalysis*, 2013; pp 289-314.
[202] Akiyama, T.; Itoh, J.; Yokota, K.; Fuchibe, K. Enantioselective Mannich-Type Reaction Catalyzed by a Chiral Brønsted Acid. *Angewandte Chemie International Edition* **2004**, *43* (12), 1566-1568. https://doi.org/10.1002/anie.200353240.
[203] Uraguchi, D.; Terada, M. Chiral Brønsted Acid-Catalyzed Direct Mannich Reactions via Electrophilic Activation. *Journal of the American Chemical Society* **2004**, *126* (17), 5356-5357. https://doi.org/10.1021/ja0491533.

[204] Hoffmann, S.; Seayad, A. M.; List, B. A Powerful Brønsted Acid Catalyst for the Organocatalytic Asymmetric Transfer Hydrogenation of Imines. *Angewandte Chemie International Edition* **2005**, *44* (45), 7424-7427. https://doi.org/10.1002/anie.200503062.
[205] Jia, Y.-X.; Zhong, J.; Zhu, S.-F.; Zhang, C.-M.; Zhou, Q.-L. Chiral Brønsted Acid Catalyzed Enantioselective Friedel–Crafts Reaction of Indoles and α-Aryl Enamides: Construction of Quaternary Carbon Atoms. *Angewandte Chemie International Edition* **2007**, *46* (29), 5565-5567. https://doi.org/10.1002/anie.200701067.
[206] Seayad, J.; Seayad, A. M.; List, B. Catalytic Asymmetric Pictet−Spengler Reaction. *Journal of the American Chemical Society* **2006**, *128* (4), 1086-1087. https://doi.org/10.1021/ja057444l.
[207] Adair, G.; Mukherjee, S.; List, B. TRIP - A Powerful Brønsted Acid Catalyst for Asymmetric Synthesis. *Aldrichim. Acta* **2008**, *41* (2), 31 - 39.
[208] Klussmann, M.; Ratjen, L.; Hoffmann, S.; Wakchaure, V.; Goddard, R.; List, B. Synthesis of TRIP and Analysis of Phosphate Salt Impurities. *Synlett* **2010**, *2010* (14), 2189-2192. https://doi.org/10.1055/s-0030-1258505.
[209] Pousse, G.; Devineau, A.; Dalla, V.; Humphreys, L.; Lasne, M.-C.; Rouden, J.; Blanchet, J. Synthesis of BINOL derived phosphorodithioic acids as new chiral Brønsted acids and an improved synthesis of 3,3′-disubstituted H8-BINOL derivatives. *Tetrahedron* **2009**, *65* (51), 10617-10622. https://doi.org/10.1016/j.tet.2009.10.068.
[210] Rueping, M.; Ieawsuwan, W.; Antonchick, A. P.; Nachtsheim, B. J. Chiral Brønsted Acids in the Catalytic Asymmetric Nazarov Cyclization—The First Enantioselective Organocatalytic Electrocyclic Reaction. *Angewandte Chemie International Edition* **2007**, *46* (12), 2097-2100. https://doi.org/10.1002/anie.200604809.
[211] Shapiro, N. D.; Rauniyar, V.; Hamilton, G. L.; Wu, J.; Toste, F. D. Asymmetric additions to dienes catalysed by a dithiophosphoric acid. *Nature* **2011**, *470* (7333), 245-249. https://doi.org/10.1038/nature09723.
[212] García-García, P.; Lay, F.; García-García, P.; Rabalakos, C.; List, B. A Powerful Chiral Counteranion Motif for Asymmetric Catalysis. *Angewandte Chemie International Edition* **2009**, *48* (24), 4363-4366. https://doi.org/10.1002/anie.200901768.
[213] Mayer, S.; List, B. Asymmetric Counteranion-Directed Catalysis. *Angewandte Chemie International Edition* **2006**, *45* (25), 4193-4195. https://doi.org/10.1002/anie.200600512.
[214] Mahlau, M.; List, B. Asymmetric Counteranion-Directed Catalysis (ACDC): A Remarkably General Approach to Enantioselective Synthesis. *Israel Journal of Chemistry* **2012**, *52* (7), 630-638. https://doi.org/10.1002/ijch.201200001.
[215] Mahlau, M.; List, B. Asymmetric Counteranion-Directed Catalysis: Concept, Definition, and Applications. *Angewandte Chemie International Edition* **2013**, *52* (2), 518-533. https://doi.org/10.1002/anie.201205343.
[216] Čorić, I.; List, B. Asymmetric spiroacetalization catalysed by confined Brønsted acids. *Nature* **2012**, *483* (7389), 315-319. https://doi.org/10.1038/nature10932.
[217] Kim, J. H.; Čorić, I.; Palumbo, C.; List, B. Resolution of Diols via Catalytic Asymmetric Acetalization. *Journal of the American Chemical Society* **2015**, *137* (5), 1778-1781. https://doi.org/10.1021/ja512481d.
[218] Liu, L.; Kaib, P. S. J.; Tap, A.; List, B. A General Catalytic Asymmetric Prins Cyclization. *Journal of the American Chemical Society* **2016**, *138* (34), 10822-10825. https://doi.org/10.1021/jacs.6b07240.
[219] Kaib, P. S. J.; Schreyer, L.; Lee, S.; Properzi, R.; List, B. Extremely Active Organocatalysts Enable a Highly Enantioselective Addition of Allyltrimethylsilane to Aldehydes. *Angewandte Chemie International Edition* **2016**, *55* (42), 13200-13203. https://doi.org/10.1002/anie.201607828.

[220] Ouyang, J.; Kennemur, J. L.; De, C. K.; Farès, C.; List, B. Strong and Confined Acids Enable a Catalytic Asymmetric Nazarov Cyclization of Simple Divinyl Ketones. *Journal of the American Chemical Society* **2019**, *141* (8), 3414-3418. https://doi.org/10.1021/jacs.8b13899.
[221] Ghosh, S.; Das, S.; De, C. K.; Yepes, D.; Neese, F.; Bistoni, G.; Leutzsch, M.; List, B. Strong and Confined Acids Control Five Stereogenic Centers in Catalytic Asymmetric Diels–Alder Reactions of Cyclohexadienones with Cyclopentadiene. *Angewandte Chemie International Edition* **2020**, *59* (30), 12347-12351. https://doi.org/10.1002/anie.202000307.
[222] Ghosh, S.; Erchinger, J. E.; Maji, R.; List, B. Catalytic Asymmetric Spirocyclizing Diels–Alder Reactions of Enones: Stereoselective Total and Formal Syntheses of α-Chamigrene, β-Chamigrene, Laurencenone C, Colletoic Acid, and Omphalic Acid. *Journal of the American Chemical Society* **2022**, *144* (15), 6703-6708. https://doi.org/10.1021/jacs.2c01971.
[223] Tsuji, N.; Kennemur, J. L.; Buyck, T.; Lee, S.; Prévost, S.; Kaib, P. S. J.; Bykov, D.; Farès, C.; List, B. Activation of olefins via asymmetric Brønsted acid catalysis. *Science* **2018**, *359* (6383), 1501-1505. https://doi.org/10.1126/science.aaq0445.
[224] Zhang, P.; Tsuji, N.; Ouyang, J.; List, B. Strong and Confined Acids Catalyze Asymmetric Intramolecular Hydroarylations of Unactivated Olefins with Indoles. *Journal of the American Chemical Society* **2021**, *143* (2), 675-680. https://doi.org/10.1021/jacs.0c12042.
[225] Schreyer, L.; Kaib, P. S. J.; Wakchaure, V. N.; Obradors, C.; Properzi, R.; Lee, S.; List, B. Confined acids catalyze asymmetric single aldolizations of acetaldehyde enolates. *Science* **2018**, *362* (6411), 216-219. https://doi.org/10.1126/science.aau0817.
[226] Amatov, T.; Tsuji, N.; Maji, R.; Schreyer, L.; Zhou, H.; Leutzsch, M.; List, B. Confinement-Controlled, Either syn- or anti-Selective Catalytic Asymmetric Mukaiyama Aldolizations of Propionaldehyde Enolsilanes. *Journal of the American Chemical Society* **2021**, *143* (36), 14475-14481. https://doi.org/10.1021/jacs.1c07447.
[227] Schwengers, S. A.; De, C. K.; Grossmann, O.; Grimm, J. A. A.; Sadlowski, N. R.; Gerosa, G. G.; List, B. Unified Approach to Imidodiphosphate-Type Brønsted Acids with Tunable Confinement and Acidity. *Journal of the American Chemical Society* **2021**, *143* (36), 14835-14844. https://doi.org/10.1021/jacs.1c07067.
[228] Mitschke, B.; Turberg, M.; List, B. Confinement as a Unifying Element in Selective Catalysis. *Chem* **2020**, *6* (10), 2515-2532. https://doi.org/10.1016/j.chempr.2020.09.007.
[229] Hastings, C. J.; Pluth, M. D.; Bergman, R. G.; Raymond, K. N. Enzymelike Catalysis of the Nazarov Cyclization by Supramolecular Encapsulation. *Journal of the American Chemical Society* **2010**, *132* (20), 6938-6940. https://doi.org/10.1021/ja102633e.
[230] Valente, C.; Çalimsiz, S.; Hoi, K. H.; Mallik, D.; Sayah, M.; Organ, M. G. The Development of Bulky Palladium NHC Complexes for the Most-Challenging Cross-Coupling Reactions. *Angewandte Chemie International Edition* **2012**, *51* (14), 3314-3332. https://doi.org/10.1002/anie.201106131.
[231] Cheong, P. H.-Y.; Legault, C. Y.; Um, J. M.; Çelebi-Ölçüm, N.; Houk, K. N. Quantum Mechanical Investigations of Organocatalysis: Mechanisms, Reactivities, and Selectivities. *Chemical Reviews* **2011**, *111* (8), 5042-5137. https://doi.org/10.1021/cr100212h.
[232] Seeman, J. I. Effect of conformational change on reactivity in organic chemistry. Evaluations, applications, and extensions of Curtin-Hammett Winstein-Holness kinetics. *Chemical Reviews* **1983**, *83* (2), 83-134. https://doi.org/10.1021/cr00054a001.
[233] Peng, Q.; Duarte, F.; Paton, R. S. Computing organic stereoselectivity – from concepts to quantitative calculations and predictions. *Chemical Society Reviews* **2016**, *45* (22), 6093-6107. https://doi.org/10.1039/C6CS00573J.
[234] Yepes, D.; Neese, F.; List, B.; Bistoni, G. Unveiling the Delicate Balance of Steric and Dispersion Interactions in Organocatalysis Using High-Level Computational Methods. *Journal of the American Chemical Society* **2020**, *142* (7), 3613-3625. https://doi.org/10.1021/jacs.9b13725.

[235] Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *The Journal of Physical Chemistry A* **1998**, *102* (11), 1995-2001. https://doi.org/10.1021/jp9716997.
[236] Wu, X.; Turlik, A.; Luan, B.; He, F.; Qu, J.; Houk, K. N.; Chen, Y. Nickel-Catalyzed Enantioselective Reductive Alkyl-Carbamoylation of Internal Alkenes. *Angewandte Chemie International Edition* **2022**, *61* (36), e202207536. https://doi.org/10.1002/anie.202207536.
[237] Gao, S.; Duan, M.; Andreola, L. R.; Yu, P.; Wheeler, S. E.; Houk, K. N.; Chen, M. Unusual Enantiodivergence in Chiral Brønsted Acid-Catalyzed Asymmetric Allylation with β-Alkenyl Allylic Boronates. *Angewandte Chemie International Edition* **2022**, *61* (41), e202208908. https://doi.org/10.1002/anie.202208908.
[238] Andreola, L. R.; Wheeler, S. E. Importance of favourable non-covalent contacts in the stereoselective synthesis of tetrasubstituted chromanones. *Organic Chemistry Frontiers* **2022**, *9* (11), 3027-3033. https://doi.org/10.1039/D2QO00090C.
[239] Rummel, L.; Domanski, M. H. J.; Hausmann, H.; Becker, J.; Schreiner, P. R. London Dispersion Favors Sterically Hindered Diarylthiourea Conformers in Solution. *Angewandte Chemie International Edition* **2022**, *61* (29), e202204393. https://doi.org/10.1002/anie.202204393.
[240] König, H. F.; Hausmann, H.; Schreiner, P. R. Assessing the Experimental Hydrogen Bonding Energy of the Cyclic Water Dimer Transition State with a Cyclooctatetraene-Based Molecular Balance. *Journal of the American Chemical Society* **2022**, *144* (37), 16965-16973. https://doi.org/10.1021/jacs.2c06141.
[241] Schwan, S.; Achazi, A. J.; Ziese, F.; Schreiner, P. R.; Volz, K.; Dehnen, S.; Sanna, S.; Mollenhauer, D. Insights into molecular cluster materials with adamantane-like core structures by considering dimer interactions. *Journal of Computational Chemistry* **2023**, *44* (7), 843-856. https://doi.org/10.1002/jcc.27047.
[242] Rummel, L.; Hanke, K.; Becker, J.; Schreiner, P. R. London Dispersion Stabilizes Chloro-Substituted cis-Double Bonds. *Synlett* **2022**, (EFirst). https://doi.org/10.1055/a-1928-2473.
[243] Grimm, L. M.; Spicher, S.; Tkachenko, B.; Schreiner, P. R.; Grimme, S.; Biedermann, F. The Role of Packing, Dispersion, Electrostatics, and Solvation in High-Affinity Complexes of Cucurbit[n]urils with Uncharged Polar Guests. *Chemistry – A European Journal* **2022**, *28* (38), e202200529. https://doi.org/10.1002/chem.202200529.
[244] Mitra, R.; Zhu, H.; Grimme, S.; Niemeyer, J. Functional Mechanically Interlocked Molecules: Asymmetric Organocatalysis with a Catenated Bifunctional Brønsted Acid. *Angewandte Chemie International Edition* **2017**, *56* (38), 11456-11459. https://doi.org/10.1002/anie.201704647.
[245] Thölke, S.; Zhu, H.; Jansen, D.; Octa-Smolin, F.; Thiele, M.; Kaupmees, K.; Leito, I.; Grimme, S.; Niemeyer, J. Cooperative Organocatalysis: A Systematic Investigation of Covalently Linked Organophosphoric Acids for the Stereoselective Transfer Hydrogenation of Quinolines. *European Journal of Organic Chemistry* **2019**, *2019* (31-32), 5190-5195. https://doi.org/10.1002/ejoc.201900548.
[246] Jansen, D.; Gramüller, J.; Niemeyer, F.; Schaller, T.; Letzel, M. C.; Grimme, S.; Zhu, H.; Gschwind, R. M.; Niemeyer, J. What is the role of acid–acid interactions in asymmetric phosphoric acid organocatalysis? A detailed mechanistic study using interlocked and non-interlocked catalysts. *Chemical Science* **2020**, *11* (17), 4381-4390. https://doi.org/10.1039/D0SC01026J.
[247] Thiele, M.; Rose, T.; Lõkov, M.; Stadtfeld, S.; Tshepelevitsh, S.; Parman, E.; Opara, K.; Wölper, C.; Leito, I.; Grimme, S.; et al. Multifunctional Organocatalysts - Singly-Linked and Macrocyclic Bisphosphoric Acids for Asymmetric Phase-Transfer and Brønsted-Acid Catalysis. *Chemistry – A European Journal* **2023**, *29* (2), e202202953. https://doi.org/10.1002/chem.202202953.

[248] Harden, I.; Neese, F.; Bistoni, G. An induced-fit model for asymmetric organocatalytic reactions: a case study of the activation of olefins via chiral Brønsted acid catalysts. *Chemical Science* **2022**, *13* (30), 8848-8859. https://doi.org/10.1039/D2SC02274E.
[249] Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chemistry – A European Journal* **2012**, *18* (32), 9955-9964. https://doi.org/10.1002/chem.201200497.
[250] Helmich-Paris, B.; Souza, B. d.; Neese, F.; Izsák, R. An improved chain of spheres for exchange algorithm. *The Journal of Chemical Physics* **2021**, *155* (10), 104109. https://doi.org/10.1063/5.0058766.
[251] Wuttke, A.; Mata, R. A. Visualizing dispersion interactions through the use of local orbital spaces. *Journal of Computational Chemistry* **2017**, *38* (1), 15-23. https://doi.org/10.1002/jcc.24508.
[252] Altun, A.; Neese, F.; Bistoni, G. Effect of Electron Correlation on Intermolecular Interactions: A Pair Natural Orbitals Coupled Cluster Based Local Energy Decomposition Study. *Journal of Chemical Theory and Computation* **2019**, *15* (1), 215-228. https://doi.org/10.1021/acs.jctc.8b00915.
[253] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *The Journal of Physical Chemistry B* **2009**, *113* (18), 6378-6396. https://doi.org/10.1021/jp810292n.
[254] Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* **1976**, *32* (5), 922-923. https://doi.org/10.1107/S0567739476001873.
[255] Coutsias, E. A.; Seok, C.; Dill, K. A. Using quaternions to calculate RMSD. *Journal of Computational Chemistry* **2004**, *25* (15), 1849-1857. https://doi.org/10.1002/jcc.20110.
[256] Poater, A.; Cosenza, B.; Correa, A.; Giudice, S.; Ragone, F.; Scarano, V.; Cavallo, L. SambVca: A Web Application for the Calculation of the Buried Volume of N-Heterocyclic Carbene Ligands. *European Journal of Inorganic Chemistry* **2009**, *2009* (13), 1759-1766. https://doi.org/10.1002/ejic.200801160.
[257] Falivene, L.; Credendino, R.; Poater, A.; Petta, A.; Serra, L.; Oliva, R.; Scarano, V.; Cavallo, L. SambVca 2. A Web Tool for Analyzing Catalytic Pockets with Topographic Steric Maps. *Organometallics* **2016**, *35* (13), 2286-2293. https://doi.org/10.1021/acs.organomet.6b00371.
[258] Falivene, L.; Cao, Z.; Petta, A.; Serra, L.; Poater, A.; Oliva, R.; Scarano, V.; Cavallo, L. Towards the online computer-aided design of catalytic pockets. *Nature Chemistry* **2019**, *11* (10), 872-879. https://doi.org/10.1038/s41557-019-0319-5.
[259] Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *Journal of the American Chemical Society* **2010**, *132* (18), 6498-6506. https://doi.org/10.1021/ja100936w.
[260] Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J. P.; Beratan, D. N.; Yang, W. NCIPLOT: a program for plotting non-covalent interaction regions. *J Chem Theory Comput* **2011**, *7* (3), 625-632. https://doi.org/10.1021/ct100641a. From NLM.
[261] Boto, R. A.; Peccati, F.; Laplaza, R.; Quan, C.; Carbone, A.; Piquemal, J.-P.; Maday, Y.; Contreras-Garcıa, J. NCIPLOT4: Fast, Robust, and Quantitative Analysis of Noncovalent Interactions. *Journal of Chemical Theory and Computation* **2020**, *16* (7), 4150-4158. https://doi.org/10.1021/acs.jctc.0c00063.
[262] Harden, I.; Neese, F.; Bistoni, G. Dimerization of Confined Brønsted Acids in Enantioselective Organocatalytic Reactions. *Manuscript submitted for publication* **2023**.
[263] Das, S.; Mitschke, B.; De, C. K.; Harden, I.; Bistoni, G.; List, B. Harnessing the ambiphilicity of silyl nitronates in a catalytic asymmetric approach to aliphatic β3-amino acids. *Nature Catalysis* **2021**, *4* (12), 1043-1049. https://doi.org/10.1038/s41929-021-00714-x.

[264] Ouyang, J.; Maji, R.; Leutzsch, M.; Mitschke, B.; List, B. Design of an Organocatalytic Asymmetric (4 + 3) Cycloaddition of 2-Indolylalcohols with Dienolsilanes. *Journal of the American Chemical Society* **2022**, *144* (19), 8460-8466. https://doi.org/10.1021/jacs.2c02216.
[265] Detering, C.; Tolstoy, P. M.; Golubev, N. S.; Denisov, G. S.; Limbach, H. H. Vicinal H/D Isotope Effects in NMR Spectra of Complexes with Coupled Hydrogen Bonds: Phosphoric Acids. *Doklady Physical Chemistry* **2001**, *379* (1), 191-193. https://doi.org/10.1023/A:1019290114622.
[266] Malm, C.; Kim, H.; Wagner, M.; Hunger, J. Complexity in Acid–Base Titrations: Multimer Formation Between Phosphoric Acids and Imines. *Chemistry – A European Journal* **2017**, *23* (45), 10853-10860. https://doi.org/10.1002/chem.201701576.
[267] Monaco, M. R.; Fazzi, D.; Tsuji, N.; Leutzsch, M.; Liao, S.; Thiel, W.; List, B. The Activation of Carboxylic Acids via Self-Assembly Asymmetric Organocatalysis: A Combined Experimental and Computational Investigation. *Journal of the American Chemical Society* **2016**, *138* (44), 14740-14749. https://doi.org/10.1021/jacs.6b09179.
[268] Melikian, M.; Gramüller, J.; Hioe, J.; Greindl, J.; Gschwind, R. M. Brønsted acid catalysis – the effect of 3,3′-substituents on the structural space and the stabilization of imine/phosphoric acid complexes. *Chemical Science* **2019**, *10* (20), 5226-5234. https://doi.org/10.1039/C9SC01044K.
[269] Gramüller, J.; Dullinger, P.; Horinek, D.; Gschwind, R. M. Bidentate substrate binding in Brønsted acid catalysis: structural space, hydrogen bonding and dimerization. *Chemical Science* **2022**, *13* (48), 14366-14372. https://doi.org/10.1039/D2SC05076E.
[270] 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. *Theoretical Chemistry Accounts* **2008**, *120* (1), 215-241. https://doi.org/10.1007/s00214-007-0310-x.
[271] Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels−Alder Reactions and Pt(P(t-Bu)3)2 + H2 Oxidative Addition. *The Journal of Physical Chemistry* **1996**, *100* (50), 19357-19363. https://doi.org/10.1021/jp962071j.
[272] Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. *Journal of Chemical Theory and Computation* **2021**, *17* (7), 4250-4261. https://10.1021/acs.jctc.1c00471.
[273] Ishida, K.; Morokuma, K.; Komornicki, A. The intrinsic reaction coordinate. An ab initio calculation for HNC→HCN and H−+CH4→CH4+H− *The Journal of Chemical Physics* **1977**, *66* (5), 2153-2156. https://doi.org/10.1063/1.434152.
[274] van der Lubbe, S. C. C.; Fonseca Guerra, C. The Nature of Hydrogen Bonds: A Delineation of the Role of Different Energy Components on Hydrogen Bond Strengths and Lengths. *Chemistry – An Asian Journal* **2019**, *14* (16), 2760-2769. https://doi.org/10.1002/asia.201900717.
[275] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chemical Reviews* **2016**, *116* (9), 5105-5154. https://doi.org/10.1021/acs.chemrev.5b00533.
[276] Axilrod, B. M.; Teller, E. Interaction of the van der Waals Type Between Three Atoms. *The Journal of Chemical Physics* **1943**, *11* (6), 299-300. https://doi.org/10.1063/1.1723844.
[277] Schümann, J. M.; Ochmann, L.; Becker, J.; Altun, A.; Harden, I.; Bistoni, G.; Schreiner, P. R. Exploring the Limits of Intramolecular London Dispersion Stabilization with Bulky Dispersion Energy Donors in Alkane Solution. *Journal of the American Chemical Society* **2023**, *145* (4), 2093-2097. https://doi.org/10.1021/jacs.2c13301.

[278] Mati, I. K.; Cockroft, S. L. Molecular balances for quantifying non-covalent interactions. *Chemical Society Reviews* **2010**, *39* (11), 4195-4205, 10.1039/B822665M. https://doi.org/10.1039/B822665M.
[279] Schümann, J. M.; Wagner, J. P.; Eckhardt, A. K.; Quanz, H.; Schreiner, P. R. Intramolecular London Dispersion Interactions Do Not Cancel in Solution. *Journal of the American Chemical Society* **2021**, *143* (1), 41-45. https://doi.org/10.1021/jacs.0c09597.
[280] Garcia-Ratés, M.; Neese, F. Effect of the Solute Cavity on the Solvation Energy and its Derivatives within the Framework of the Gaussian Charge Scheme. *Journal of Computational Chemistry* **2020**, *41* (9), 922-939. https://doi.org/10.1002/jcc.26139.
[281] Spicher, S.; Plett, C.; Pracht, P.; Hansen, A.; Grimme, S. Automated Molecular Cluster Growing for Explicit Solvation by Efficient Force Field and Tight Binding Methods. *Journal of Chemical Theory and Computation* **2022**, *18* (5), 3174-3189. https://doi.org/10.1021/acs.jctc.2c00239.
[282] Plett, C.; Grimme, S. Automated and Efficient Generation of General Molecular Aggregate Structures. *Angewandte Chemie International Edition* **2023**, *62* (4), e202214477. https://doi.org/10.1002/anie.202214477.
[283] Chatterjee, S.; Harden, I.; Bistoni, G.; Castillo, R. G.; Chabbra, S.; van Gastel, M.; Schnegg, A.; Bill, E.; Birrell, J. A.; Morandi, B.; et al. A Combined Spectroscopic and Computational Study on the Mechanism of Iron-Catalyzed Aminofunctionalization of Olefins Using Hydroxylamine Derived N–O Reagent as the "Amino" Source and "Oxidant". *Journal of the American Chemical Society* **2022**, *144* (6), 2637-2656. https://doi.org/10.1021/jacs.1c11083.
[284] Roose, P.; Eller, K.; Henkes, E.; Rossbacher, R.; Höke, H. Amines, Aliphatic. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2015; pp 1-55.
[285] Kahl, T.; Schröder, K.-W.; Lawrence, F. R.; Marshall, W. J.; Höke, H.; Jäckh, R. Aniline. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2011.
[286] Iversen, S. D.; Iversen, L. L. Dopamine: 50 years in perspective. *Trends in Neurosciences* **2007**, *30* (5), 188-193. https://doi.org/10.1016/j.tins.2007.03.002.
[287] Drauz, K.; Grayson, I.; Kleemann, A.; Krimmer, H.-P.; Leuchtenberger, W.; Weckbecker, C. Amino Acids. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2007.
[288] Dauban, P.; Darses, B.; Jarvis, A. 7.19 Addition Reactions with Formation of CarbonNitrogen Bonds. In *Comprehensive Organic Synthesis (Second Edition)*, Knochel, P. Ed.; Elsevier, 2014; pp 538-604.
[289] Fürstner, A. Iron Catalysis in Organic Synthesis: A Critical Assessment of What It Takes To Make This Base Metal a Multitasking Champion. *ACS Central Science* **2016**, *2* (11), 778-789. https://doi.org/10.1021/acscentsci.6b00272.
[290] Legnani, L.; Morandi, B. Direct Catalytic Synthesis of Unprotected 2-Amino-1-Phenylethanols from Alkenes by Using Iron(II) Phthalocyanine. *Angewandte Chemie International Edition* **2016**, *55* (6), 2248-2251. https://doi.org/10.1002/anie.201507630.
[291] Legnani, L.; Prina-Cerai, G.; Delcaillau, T.; Willems, S.; Morandi, B. Efficient access to unprotected primary amines by iron-catalyzed aminochlorination of alkenes. *Science* **2018**, *362* (6413), 434-439. https://doi.org/10.1126/science.aat3863.
[292] Falk, E.; Makai, S.; Delcaillau, T.; Gürtler, L.; Morandi, B. Design and Scalable Synthesis of N-Alkylhydroxylamine Reagents for the Direct Iron-Catalyzed Installation of Medicinally Relevant Amines**. *Angewandte Chemie International Edition* **2020**, *59* (47), 21064-21071. https://doi.org/10.1002/anie.202008247.
[293] Makai, S.; Falk, E.; Morandi, B. Direct Synthesis of Unprotected 2-Azidoamines from Alkenes via an Iron-Catalyzed Difunctionalization Reaction. *Journal of the American Chemical Society* **2020**, *142* (51), 21548-21555. https://doi.org/10.1021/jacs.0c11025.
[294] Chatterjee, S.; Makai, S.; Morandi, B. Hydroxylamine-Derived Reagent as a Dual Oxidant and Amino Group Donor for the Iron-Catalyzed Preparation of Unprotected

Sulfinamides from Thiols. *Angewandte Chemie International Edition* **2021**, *60* (2), 758-765. https://doi.org/10.1002/anie.202011138.
[295] Yu, H.; Li, Z.; Bolm, C. Iron(II)-Catalyzed Direct Synthesis of NH Sulfoximines from Sulfoxides. *Angewandte Chemie International Edition* **2018**, *57* (1), 324-327. https://doi.org/10.1002/anie.201710498.
[296] Atherton, N. M.; Davies, M.; Gilbert, B. *Electron spin resonance*; Royal Society of Chemistry, 1994.
[297] Rempel, R. C. 22 - PARAMAGNETIC IONS IN CRYSTALS. In *NMR and EPR Spectroscopy*, The Nmr-Epr Staff Of Varian, A. Ed.; Pergamon, 1960; pp 261-267.
[298] Greenwood, N. N. *Mössbauer spectroscopy*; Springer Science & Business Media, 2012.
[299] Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters* **1984**, *52* (12), 997-1000. https://doi.org/10.1103/PhysRevLett.52.997.
[300] Lenthe, E. v.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *The Journal of Chemical Physics* **1993**, *99* (6), 4597-4610. https://doi.org/10.1063/1.466059.
[301] Wüllen, C. v. Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations. *The Journal of Chemical Physics* **1998**, *109* (2), 392-399. https://doi.org/10.1063/1.476576.
[302] Martin, R. L. Natural transition orbitals. *The Journal of Chemical Physics* **2003**, *118* (11), 4775-4777. https://doi.org/10.1063/1.1558471.
[303] Petrenko, T.; Neese, F. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theory of electronic spectroscopy. *The Journal of Chemical Physics* **2007**, *127* (16), 164319. https://doi.org/10.1063/1.2770706.
[304] Petrenko, T.; Neese, F. Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model. *The Journal of Chemical Physics* **2012**, *137* (23), 234107. https://doi.org/10.1063/1.4771959.
[305] Belay, A. Measurement of integrated absorption cross-section, oscillator strength and number density of caffeine in coffee beans by integrated absorption coefficient technique. *Food Chemistry* **2010**, *121* (2), 585-590. https://doi.org/10.1016/j.foodchem.2009.12.052.
[306] Neese, F. Prediction and interpretation of the 57Fe isomer shift in Mössbauer spectra by density functional theory. *Inorganica Chimica Acta* **2002**, *337*, 181-192. https://doi.org/10.1016/S0020-1693(02)01031-9.
[307] Sinnecker, S.; Slep, L. D.; Bill, E.; Neese, F. Performance of Nonrelativistic and Quasi-Relativistic Hybrid DFT for the Prediction of Electric and Magnetic Hyperfine Parameters in 57Fe Mössbauer Spectra. *Inorganic Chemistry* **2005**, *44* (7), 2245-2254. https://doi.org/10.1021/ic048609e.
[308] 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. *Inorganic Chemistry* **2009**, *48* (3), 784-785. https://doi.org/10.1021/ic801535v.
[309] Mills, G.; Jónsson, H.; Schenter, G. K. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surface Science* **1995**, *324* (2), 305-337. https://doi.org/10.1016/0039-6028(94)00731-4.
[310] Henkelman, G.; Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of Chemical Physics* **2000**, *113* (22), 9978-9985. https://doi.org/10.1063/1.1323224.
[311] Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics* **2000**, *113* (22), 9901-9904. https://doi.org/10.1063/1.1329672.

[312] Ásgeirsson, V.; Birgisson, B. O.; Bjornsson, R.; Becker, U.; Neese, F.; Riplinger, C.; Jónsson, H. Nudged Elastic Band Method for Molecular Reactions Using Energy-Weighted Springs Combined with Eigenvector Following. *Journal of Chemical Theory and Computation* **2021**, *17* (8), 4929-4945. https://doi.org/10.1021/acs.jctc.1c00462.
[313] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics* **2001**, *114* (23), 10252-10264. https://doi.org/10.1063/1.1361246.
[314] Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *The Journal of Chemical Physics* **2002**, *117* (20), 9138-9153. https://doi.org/10.1063/1.1515317.
[315] Atanasov, M.; Ganyushin, D.; Sivalingam, K.; Neese, F. A Modern First-Principles View on Ligand Field Theory Through the Eyes of Correlated Multireference Wavefunctions. In *Molecular Electronic Structures of Transition Metal Complexes II*, Mingos, D. M. P., Day, P., Dahl, J. P. Eds.; Springer Berlin Heidelberg, 2012; pp 149-220.
[316] Bjornsson, R.; Neese, F.; DeBeer, S. Revisiting the Mössbauer Isomer Shifts of the FeMoco Cluster of Nitrogenase and the Cofactor Charge. *Inorganic Chemistry* **2017**, *56* (3), 1470-1477. https://doi.org/10.1021/acs.inorgchem.6b02540.
[317] Schulz, C. E.; Castillo, R. G.; Pantazis, D. A.; DeBeer, S.; Neese, F. Structure–Spectroscopy Correlations for Intermediate Q of Soluble Methane Monooxygenase: Insights from QM/MM Calculations. *Journal of the American Chemical Society* **2021**, *143* (17), 6560-6577. https://doi.org/10.1021/jacs.1c01180.
[318] Robert, B. Resonance Raman spectroscopy. *Photosynthesis Research* **2009**, *101* (2), 147-155. https://doi.org/10.1007/s11120-009-9440-4.
[319] Neese, F. Importance of Direct Spin−Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study. *Journal of the American Chemical Society* **2006**, *128* (31), 10213-10222. https://doi.org/10.1021/ja061798a.
[320] Yamaguchi, K.; Takahara, Y.; Fueno, T. Ab-Initio Molecular Orbital Studies of Structure and Reactivity of Transition Metal-OXO Compounds. In *Applied Quantum Chemistry*, Dordrecht, 1986; Smith, V. H., Schaefer, H. F., Morokuma, K., Eds.; Springer Netherlands: pp 155-184.
[321] Soda, T.; Kitagawa, Y.; Onishi, T.; Takano, Y.; Shigeta, Y.; Nagao, H.; Yoshioka, Y.; Yamaguchi, K. Ab initio computations of effective exchange integrals for H–H, H–He–H and Mn2O2 complex: comparison of broken-symmetry approaches. *Chemical Physics Letters* **2000**, *319* (3), 223-230. https://doi.org/10.1016/S0009-2614(00)00166-4.
[322] Clark, A. E.; Davidson, E. R. Local spin. *The Journal of Chemical Physics* **2001**, *115* (16), 7382-7392. https://doi.org/10.1063/1.1407276.
[323] Grimme, S.; Hansen, A. A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angewandte Chemie International Edition* **2015**, *54* (42), 12308-12313. https://doi.org/10.1002/anie.201501887.
[324] Mayer, I. Charge, bond order and valence in the AB initio SCF theory. *Chemical Physics Letters* **1983**, *97* (3), 270-274. https://doi.org/10.1016/0009-2614(83)80005-0.
[325] Cowley, R. E.; DeYonker, N. J.; Eckert, N. A.; Cundari, T. R.; DeBeer, S.; Bill, E.; Ottenwaelder, X.; Flaschenriem, C.; Holland, P. L. Three-Coordinate Terminal Imidoiron(III) Complexes: Structure, Spectroscopy, and Mechanism of Formation. *Inorganic Chemistry* **2010**, *49* (13), 6172-6187. https://doi.org/10.1021/ic100846b.
[326] Iovan, D. A.; Betley, T. A. Characterization of Iron-Imido Species Relevant for N-Group Transfer Chemistry. *Journal of the American Chemical Society* **2016**, *138* (6), 1983-1993. https://doi.org/10.1021/jacs.5b12582.
[327] Wilding, M. J. T.; Iovan, D. A.; Wrobel, A. T.; Lukens, J. T.; MacMillan, S. N.; Lancaster, K. M.; Betley, T. A. Direct Comparison of C–H Bond Amination Efficacy through

Manipulation of Nitrogen-Valence Centered Redox: Imido versus Iminyl. *Journal of the American Chemical Society* **2017**, *139* (41), 14757-14766. https://doi.org/10.1021/jacs.7b08714.
[328] Namuswe, F.; Hayashi, T.; Jiang, Y.; Kasper, G. D.; Sarjeant, A. A. N.; Moënne-Loccoz, P.; Goldberg, D. P. Influence of the Nitrogen Donors on Nonheme Iron Models of Superoxide Reductase: High-Spin FeIII−OOR Complexes. *Journal of the American Chemical Society* **2010**, *132* (1), 157-167. https://doi.org/10.1021/ja904818z.
[329] Roelfes, G.; Vrajmasu, V.; Chen, K.; Ho, R. Y. N.; Rohde, J.-U.; Zondervan, C.; la Crois, R. M.; Schudde, E. P.; Lutz, M.; Spek, A. L.; et al. End-On and Side-On Peroxo Derivatives of Non-Heme Iron Complexes with Pentadentate Ligands: Models for Putative Intermediates in Biological Iron/Dioxygen Chemistry. *Inorganic Chemistry* **2003**, *42* (8), 2639-2653. https://doi.org/10.1021/ic034065p.
[330] Kim, J.; Zang, Y.; Costas, M.; Harrison, R. G.; Wilkinson, E. C.; Que Jr, L. A nonheme iron(II) complex that models the redox cycle of lipoxygenase. *JBIC Journal of Biological Inorganic Chemistry* **2001**, *6* (3), 275-284. https://doi.org/10.1007/s007750000198.
[331] de Juan, A.; Jaumot, J.; Tauler, R. Multivariate Curve Resolution (MCR). Solving the mixture analysis problem. *Analytical Methods* **2014**, *6* (14), 4964-4976. https://doi.org/10.1039/C4AY00571F.
[332] Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **2020**, *17* (3), 261-272. https://doi.org/10.1038/s41592-019-0686-2.
[333] Cramer, C. J. *Essentials of computational chemistry: theories and models*; John Wiley & Sons, 2013.
[334] Schläfer, H. L.; Gliemann, G. *Einführung in Die Ligandenfeldtheorie.* ; Akademische Verlagsgesellschaft, 1967.
[335] Singh, S. K.; Eng, J.; Atanasov, M.; Neese, F. Covalency and chemical bonding in transition metal complexes: An ab initio based ligand field perspective. *Coordination Chemistry Reviews* **2017**, *344*, 2-25. https://doi.org/10.1016/j.ccr.2017.03.018.
[336] Aravena, D.; Atanasov, M.; Neese, F. Periodic Trends in Lanthanide Compounds through the Eyes of Multireference ab Initio Theory. *Inorganic Chemistry* **2016**, *55* (9), 4457-4469. https://doi.org/10.1021/acs.inorgchem.6b00244.
[337] *ORCA helpers repository*. 2023. https://gitlab.gwdg.de/orca-helpers/orca-helpers (accessed 16.02.2023).
[338] Zverovich, V.; Müller, J. *{fmt} A modern formatting library*. 2023. https://fmt.dev/latest/index.html (accessed 24.01.2023).
[339] Harden, I. *Chemical Reaction Integrator*. 2023. https://github.com/iharden/Chemical-Reaction-Integrator (accessed 21.03.2023).
[340] Wanner, G.; Hairer, E. *Solving ordinary differential equations II*; Springer Berlin Heidelberg New York, 1996.
[341] The Qt Company. *Qt*. 2023. https://www.qt.io/ (accessed 25.01.2023).
[342] Ahnert, K.; Mulansky, M. *Boost Odeint*. 2015. https://www.boost.org/doc/libs/1_81_0/libs/numeric/odeint/doc/html/index.html (accessed 25.01.2023).

Acknowledgment

First of all, I want to express my thanks to my supervisor Prof. Dr. Frank Neese for giving me the opportunity to work on this fascinating PhD topic, to participate in interesting collaborations and for his ever competent, patient and benevolent guidance.

To Prof. Dr. Stefan Grimme I owe thanks for taking over the second review.

Sincerely, I would like to thank Dr. Giovanni Bistoni for his untiring support, always helpful advice, elucidating discussions, and not least his motivational skills. It has been a pleasure to work with you on a wide variety of projects.

My thanks go to Burak Savci for our helpful discussions about reaction mechanisms and for encouraging me to specialize in Theoretical Chemistry during my undergraduate studies.

To all members of the MTS group, I am indebted for the pleasant working atmosphere during the last years.

Finally, my thanks go to my wonderful family and my girlfriend Nelli for the love, patience and unlimited support I have received over the past years. Without you, this work would never have been accomplished.