DISS ETH NO. 22465

Towards Automated Quantum Chemical Explorations of Chemical Mechanisms

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

Maike Bergeler

MSc Chemistry, ETH Zurich

born on 31.10.1989

citizen of Germany

accepted on the recommendation of

Prof. Dr. Markus Reiher, examiner
Prof. Dr. Sereina Riniker, co-examiner
Prof. Dr. Carmen Herrmann, co-examiner

2015

Maike Bergeler: *Towards Automated Quantum Chemical Explorations of Chemical Mechanisms,* Dissertation ETH Zurich No. 22465, 2015

to my family

Contents

Abstract ix

Zusammenfassung xi

1 Setting the Stage 1

2 Case Studies 1: Multitude of Stable Intermediate Structures 7

2.1 Catalytic activity differences in [FeFe] hydrogenases . . 7

2.1.1 Single coordination energies 10

2.1.2 Coordination energy differences 13

2.1.3 Structural alignments 15

2.1.4 Molecular orbital visualizations 15

2.1.5 Calculated infrared vibrational spectra 18

2.1.6 Catalytic cycle of H_2 formation 21

2.1.7 Consequence for the chemistry of [FeFe] hydrogenases 23

2.2 Structure–reactivity analysis of Fe_4S_4 clusters 25

2.2.1 A generalized anchoring scheme 27

2.2.2 Structural strain and antiferromagnetic coupling 31

2.2.3 Tolerance of Fe_4S_4 clusters against structural distortions 36

2.2.4 Reactivity changes of distorted Fe_4S_4 clusters . 39

2.2.5 Effect of structural distortions on electron transfer properties 44

2.2.6 Consequence for the chemistry of Fe_4S_4 cubanes 57

3 Case Studies 2: Transition-State Optimizations 61

3.1 Conformer analysis of a double-decker-type rotamer . 61

3.1.1 Relative stability of isolated conformers 62

3.1.2 Transition-state searches 63

3.1.3 Intrinsic reaction coordinates 66

3.1.4 Comparison of the transition-state structures . 67

3.1.5 Transition pathways 68

3.1.6 Dependence on functional, basis set, dispersion correction and solvation 69
3.1.7 Entropy contribution of imaginary frequency . 71
3.1.8 Stability of the two solvated conformers 72
3.1.9 Consequences for the double-decker-type rotamer 76
3.2 Investigation of an organic phosphaketene 78
3.2.1 Comparison of different reaction routes 79
3.2.2 Consequences for the chemistry of the phosphaketene 81

4 General Conclusions from The Case Studies 87

5 Mode-Tracking based Transition-State Optimizations 89
5.1 Theoretical Foundations of MTSEARCH 92
5.2 The MTSEARCH program 96
5.3 Raw data generation 98
5.4 Reference calculations 99
5.5 Test systems for MTSEARCH 100
5.5.1 Benchmark Example: C_2H_6 rotation 101
5.5.2 Explorative Example: Isomerization of H_2CO . . 102
5.5.3 Intramolecular proton-transfer reaction in a hydrazine Mo complex 104

6 Heuristic Reaction Network Explorations 109
6.1 Structure of HEUREX 112
6.2 Structure generator and optimizer 115
6.3 Elementary-step finder 118
6.4 Automated transition-state searches between two minimum-energy structures 118
6.5 Verification of the TSs 119
6.6 Visualization of the results in reaction networks 120
6.7 Generation of raw data 120
6.8 Reaction Networks for Intermediates of the Chatt–Schrock cycle 121

7 Conclusions and Outlook 133

Appendix 137
A [FeFe] hydrogenase calculations 137
B Supplementary material for Fe_4S_4 calculations 139

B.1 Analysis of S(Cys)-S(Cys)/C$_\beta$-C$_\beta$ distances in selected crystal structures 141
B.2 Comparison of the two D$_{2d}$-symmetric structures 143
B.3 Results for the S-fixed model system 145
C Conformer analysis of double-decker-type rotamers . 153
D Reaction pathways of an organic phosphaketene . . . 154
E Mode-Tracking based Transition-State Optimizations . 155
E.1 Choice of the Mode-Tracking guess modes . . 155
E.2 Newton–Raphson step length and maximum number of orthogonal optimization steps . . . 157
E.3 Convergence criteria for MTsearch 159
E.4 TS-search examples 160
F Additional Reaction Networks 161

Bibliography 175

Abbreviations 191

Publications 193

Acknowledgement 195

Abstract

The elucidation of mechanisms is essential for understanding chemical reactions. A chemical system can be described by the number and type of atoms as well as the connectivities between the atoms. During a chemical reaction, atom connectivities change, which leads to a change in the energy of the system. To evaluate whether a certain reaction can be accomplished and to compare different competing reaction pathways with each other, the detailed understanding of the energy changes during a reaction is required. Potential energy surfaces (PES) relate the (free) energy of a system to its (molecular) structure. A Born–Oppenheimer PES is a special type of PES, in which the electronic energy of the system is related to its nuclear coordinates, which are assumed to be fixed, because of the large mass of the nuclei compared to the electrons. Several points on a Born–Oppenheimer PES are of special interest for mechanistic studies, because these structures correspond to minimum- or maximum-energy structures of a specific reaction pathway. The latter ones are saddle-points on the Born-Oppenheimer PES and referred to as transition-state structures.

Experimentally, transition-state structures are hard to characterize. Quantum chemical calculations can in principle access all points on the Born-Oppenheimer PES, which makes them indispensable for unravelling reaction energetics. Among the quantum chemical methods, density functional theory is one of the most popular tools, because it scales most favorable with system-size. Out of the chemically interesting stationary points on the PES, minimum-energy structures can be located in a straightforward manner. However, the identification of transition states is significantly more elaborate and requires additional chemical knowledge. Thus, the improvement of TS optimization methods is an important goal in computational chemistry.

To identify the shortcomings of a computational method, one may first study systems which are presumably challenging. Based on the observations from these studies, improved algorithms can be developed. In the first two Chapters of this work, two case studies are carried out, from which we derive current obstacles for automated chemical mechanism studies. The first case study is about transition-metal complexes, which occur in the active center of bioinorganic systems of, e.g, hydrogenases and a subsystem of it, namely the iron-sulfur cubane. We analyze structure–property relationships, that are vital for the function of the bioinorganic system. In the second case study transition-state calculations are carried out to compare several competing reaction pathways with each other.

Based on an analysis of the strengths and weaknesses of current quantum chemical methods for the exploration of chemical mechanisms, we propose two new algorithms, MTsearch and HeurEx. The first algorithm combines the efficient Mode-Tracking for the specific calculation of one vibrational frequency with an eigenvector following type algorithm to circumvent a full Hessian calculation. The latter allows for an automated reaction network construction of reactions with many possible side products. Here, we have chosen unspecific protonation and electron reductions. We show the results obtained with HeurEx for the nitrogen-fixating Schrock molybdenum complex.

Zusammenfassung

Die Aufklärung von Mechanismen ist essentiell für das Verständnis chemischer Reaktionen. Ein chemisches System kann durch die Anzahl und den Typ der Atome sowie deren Bindungen untereinander beschrieben werden. Während einer chemischen Reaktion ändern sich Bindungen zwischen Atomen, was zu einer Änderung der Energie des Systems führt. Um zu evaluieren, ob eine spezifische Reaktion stattfindet und, um verschiedene mit einander konkurrierende Reaktionen zu vergleichen, muss man die möglichen Reaktionswege des Systems kennen. Potential-Energie-Flächen geben die Energien eines Systems bezogen auf die molekulare Struktur wieder. Eine Born–Oppenheimer Potential-Energie-Fläche ist eine spezielle Potential-Fläche, die die elektronische Energie eines chemischen Systems in Bezug auf dessen Kern-Koordinaten wiedergibt. Die Kern-Koordinaten werden als stationär angesehen, da die Masse der Kerne im Vergleich zu den Elektronen sehr gross ist. Einige Punkte auf der Born–Oppenheimer Potential-Energie-Fläche sind von speziellem Interesse für mechanistische Studien, weil diese Strukturen den energetisch niedrigsten und höchsten Strukturen von Reaktionspfaden entsprechen. Die zweite Struktur entspricht einem Sattelpunkt auf der Born-Oppenheimer Potential-Energie-Fläche und wird auch als Übergangszustand bezeichnet.

Experimentell sind Übergangszustände schwer zu charakterisieren. Quantenchemische Rechnungen können im Prinzip alle stationären Punkte der Born–Oppenheimer Potentialfläche erreichen und sind daher notwendig für mechanistische Studien. Von den quantenchemischen Methoden ist die Dichte-Funktional-Theorie eine der populärsten Methoden geworden, weil sie am Besten mit der System-Grösse skaliert. Von den stationären Punkten auf der Potentialfläche, die chemisch relevant sind, sind die Minima auf unkomplizierte Weise erreichbar. Das Lokalisieren von Übergangszuständen ist allerdings deutlich komplizierter und erfordert häufig chemisches Wissen beziehungsweise Intuition. Daher ist das

Verbessern von Methoden zur Lokalisierung von Übergangszuständen ein wichtiges Ziel der computergestützten Chemie.

Um die Defizite einer Methode zu finden, kann man zunächst Test-Systeme studieren, um dann basierend auf den Beobachtungen verbesserte Algorithmen zu entwickeln. In diesem Sinne werden in den ersten zwei Kapiteln dieser Arbeit zwei Fallstudien beschrieben, bei denen es sich um bisher ungelöste chemische Probleme handelt. Die erste Fallstudie handelt von Übergangsmetallkomplexen, wie sie in den aktiven Zentren einiger bio-anorganischer Systeme, z.B. Hydrogenasen, vorkommen und einem Untersystem davon, nämlich den Eisen-Schwefel Kubanen. Durch systematische Methoden haben wir einige Struktur-Eigenschafts-Beziehungen analysiert, die entscheidend für die Funktion der bioanorganischen Systeme sind. In der zweiten Fallstudie werden Reaktionspfade einiger chemisch interessanter Systeme berechnet, die zeigen, welche Reaktionen möglich sind und welche unter den gegebenen Reaktionsbedingungen stattfinden.

Basierend auf der Analyse der Stärken und Schwächen der aktuellen quantenchemischen Methoden zur Aufklärung chemischer Mechanismen, stellen wir in den letzten zwei Kapiteln zwei neue Methoden vor, MTsearch und HeurEx. Die erste Methode verbindet das effiziente "Mode-Following"-Protokoll für die spezifische Berechnung einer Schwingungs-Frequenz mit einem Optimierungs-Algorithmus, der dem Eigenvektor folgt, um der zeitintensiven Berechnung der vollständigen Hesse-Matrix zu umgehen. Die zweite Methode erlaubt das automatische Konstruieren von Reaktionsnetzwerken und dient insbesondere der Studie von Reaktionen mit einer Vielzahl von möglichen Nebenprodukten, wie unspezifischen Protonierungen und Reduktionen, da in diesem Fall eine manuelle Berechnung der einzelnen Reaktionspfade nahezu unmöglich wird. Am Ende dieser Arbeit werden die Resultate vorgestellt, die mittels HeurEx zum Stickstoff-fixierenden Schrock-Komplex erzielt wurden.

1

Setting the Stage

A chemical system consists of one or more molecules, i.e., atoms that are connected by bonds according to a specific molecular structure. Such a system has a specific energy. If the system changes its molecular structure, which may occur within one molecule or between several molecules, the energy of the system changes. The relation between a (molecular) structures and its (free) energy has been documented the first time by Eyring and Polanyi in 1931 [1]. They refer to this relation as potential energy surface (PES). Structural distortions defined by two degrees of freedom span a two-dimensional surface and in the third dimension the potential energy of the respective structure is displayed. Hence, a PES helps to visualize the energy changes of a system with respect to structural distortions. The Born-Oppenheimer potential energy surface is a special type of PES, in which the potential energies are related to the nuclear coordinate positions. In 1927 Born and Oppenheimer showed that the nuclei are approximately static compared to the electrons [2]. Because of the significantly larger mass, the nuclei move much slower than the electrons.

Certain points on a Born-Oppenheimer PES are of special interest for chemical reaction studies. These are local energy-minimum structures (valleys of the PES) and first-order saddle-points (transition-state structures). Also higher-order saddle-points exist, but they are less relevant for chemical studies. Stationary points on the PES are characterized by their vanishing gradient, which is the derivative of the energy with respect to the positions of the atoms.

To understand a chemical reaction, the reaction pathway between two stable structures (i.e., the reactants) has to be uncovered. In principle, many pathways connect two stable structures on the Born-Oppenheimer PES. However, the chemically relevant one is the minimum (reaction) energy path, which is the lowest-energy path out of all possible reaction pathways. The highest energy structure on this path is the TS. A TS has exactly one imaginary frequency. The mode which corresponds to this imaginary frequency (the transition frequency) is the transition mode. If one follows this mode downhill along both directions, one reaches the two minimum-energy structures. These coordinates are referred to as intrinsic reaction coordinates [3, 4].

The energy differences between a TS and the two minimum-energy structures which are connected by the TS are the activation energy barriers for the forward and backward reactions, respectively. In 1889, Arrhenius derived an equation for the rate constant of a reaction [5] and thereby introduced the concept of the activation energy. The rate constant is described by

$$k = A \exp^{-\Delta E_a/RT}, \tag{1.1}$$

where A is a material-dependent prefactor, E_a is the activation energy, R the universal gas constant and T the temperature.

In 1935, Eyring, Polanyi, Wigner and Evans developed the Eyring equation [6], which describes the material constant A in the Arrhenius equation by the absolute rate theory. The rate constant k can be obtained by

$$k = \frac{k_B T}{h} \exp^{-\Delta G^\ddagger/RT}, \tag{1.2}$$

with the Boltzmann constant k_B, the Planck constant h, the universal gas constant R, the activation free energy of the reaction $\Delta G^\ddagger$ and the temperature T.

The $\Delta G^\ddagger$'s are evaluated from the Gibbs free energies of the minimum-energy structure and the TS of the reaction. The free energies may be approximated by the electronic energies at zero Kelvin (neglecting temperature and entropy contributions). The necessary electronic energies needed to calculate $\Delta E_\text{el}^\ddagger$ can be obtained by solving the electronic Schrödinger equation for the minimum-energy structure as well as for the TS (Refs. [7, 8] have been utilized in the following),

$$\hat{H}_{\mathrm{el}}\Psi_{\mathrm{el}} = E_{\mathrm{el}}\Psi_{\mathrm{el}}, \tag{1.3}$$

where $\hat{H}_{\mathrm{el}}$ is the electronic Hamiltonian, Ψ_{el} is the electronic wave function and E_{el} the electronic energy.

The electronic Hamiltonian at fixed nuclei positions reads

$$\hat{H}_{\mathrm{el}} = \hat{T}_{\mathrm{el}} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{ext}} + \hat{V}_{\mathrm{NN}}, \tag{1.4}$$

with $\hat{T}_{\mathrm{el}}$ being the operator for the kinetic energy, $\hat{V}_{\mathrm{ee}}$ the operator for the electron-electron interaction, $\hat{V}_{\mathrm{ext}} = \sum_i v(\mathbf{r_i})$ $(i = 1, .., N_{\mathrm{el}})$ the operator for an external potential, and $\hat{V}_{\mathrm{NN}}$ the operator for the nuclear-nuclear repulsion of the system.

Several quantum chemical methods exist, which all aim at an accurate description of the electronic structure of a system. Among them density functional theory (DFT [9]) is the method which scales most favorable with system-size. By contrast to wave function-based methods which depend on the N_{el} electrons in the system, in DFT the ground-state electronic energy is described by the electron density $\rho(\mathbf{r})$, which only depends on three spatial coordinates ($\mathbf{r}$),

$$\rho(\mathbf{r}) = N_{\mathrm{el}} \int \dots \int \left|\Psi_0(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\mathrm{el}}})\right|^2 \mathrm{d}\mathbf{r}_2 \dots \mathrm{d}\mathbf{r}_{N_{\mathrm{el}}}. \tag{1.5}$$

The Hohenberg–Kohn theorems [10] connect the ground-state wave function with the ground-state electronic density. To obtain the electronic ground-state energy $E_{\mathrm{el},0}$, the functional $E_{\mathrm{el},0}[\rho]$ is minimized under the condition, that the integral of the density over the whole space is equal to the number of electrons (N_{el}) by employing the Lagrange multiplier method. The functional $E_{\mathrm{el},0}[\rho]$ reads

$$E_{\mathrm{el},0}[\rho] = F[\rho] + V_{\mathrm{ext}}[\rho], \tag{1.6}$$

where $V_{\mathrm{ext}}[\rho]$ is the functional of the external potential and $F[\rho]$ is the so-called Hohenberg–Kohn functional. The Hohenberg–Kohn functional can be written in the following form,

$$F[\rho] = T[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho], \tag{1.7}$$

with $T[\rho]$ being the kinetic energy functional, $J[\rho]$ being the Coulomb repulsion energy functional and $E_{\mathrm{xc}}[\rho]$ the exchange–correlation energy functional. Unfortunately, no analytical form of the Hohenberg–Kohn functional has been found. In contrast to $J[\rho]$, which can be calculated exactly, $T[\rho]$ and $E_{\mathrm{xc}}[\rho]$ are not known. Therefore, the Hohenberg–Kohn functional has to be approximated.

Kohn and Sham [11] had the idea to split the kinetic energy term $T[\rho]$ of a system of interacting particles into one part, which describes the kinetic energy of non-interacting particles (which can be calculated exactly) and one part, which corrects the kinetic energy term of the non-interacting particles such that the energy approaches the one of a system of interacting particles. The subsequent equations are called Kohn–Sham (KS) equations. The ground-state density of the non-interacting reference system and the interacting system is the same.

KS-DFT takes advantage of the fact that the solution of the Schrödinger equation for a non-interacting particle system is a Slater determinant. The correction energy term is referred to as exchange–correlation energy. Several approximations for this term exist: e.g., the local density approximation, generalized gradient approximation, and hybrid functionals, which contain an admixture of Hartree-Fock exchange.

The correct wave functions for systems with several transition-metal centers, in which the electron spins couple, is difficult to describe by a one-determinant wave function. Therefore, Noodleman et al. [12–14] introduced a procedure, which allows for the calculation of such systems by means of DFT or Hartree–Fock equations. The main idea is to split the total electron density into two contributions: the α-spin contribution and the β-spin contribution. Different α-and β-spin contributions on the atom centers lead to a broken-spin symmetry (BS) solution. In BS-DFT it is assumed that the Hamiltonian can be approximated by interactions of the single atom centers. In this work, we employ BS-DFT [12, 15–19] to model the couplings within Fe_4S_4 cubanes and the Fe-S clusters of [FeFe] hydrogenases. The excess of α- and β-spin density described by the local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value [20–22] is calculated for selected iron atoms to characterize the broken-spin symmetry wave function (determinant). Since the BS wave functions are not eigenfunctions of $\hat{S}^2$ anymore, a spin contamination occurs.

Summarizing, for mechanistic studies, one needs a quantum chemical method to obtain the energies for a given structure and optimization protocols for the localization of TSs and minimum-energy structures. The electronic structure calculations may be difficult for, e.g., complexes with several transition-metal centers, in which broken-symmetry spin-state solutions are required, or cases in which the surrounding of the quantum mechanically treated part has a strong effect. Thus, in our first case study two bioinorganic systems, in which these difficulties may occur, are studied. These are: the active center of a [FeFe] hydrogenase, which forms hydrogen from protons and electrons, and single Fe_4S_4 cubanes, which are involved in electron transport reactions. By systematically scanning structural distortions in the first ligand sphere around Fe_4S_4 cubanes, the properties of the potential energy surface of these clusters have been elucidated.

Having achieved the first step, i.e., obtaining reliable energies, the next step is the search for minimum-energy structures and transition-state structures on the Born-Oppenheimer PES. Minimum-energy structures can be obtained in a straightforward manner and are thus not further discussed here. However, current TS-optimization procedures are not completely straightforward and cannot guarantee that the lowest-energy TS will be found. Therefore, the improvement of TS-optimization algorithms is highly desirable. To figure out an efficient way of automated transition state searches, we investigate commonly employed TS optimizers by performing case study 2.

In the second part of this thesis (Chapters 5 and 6) we develop a mode-following TS-optimization procedure, called MTsearch, and a reaction-network-exploration algorithm based on a heuristic structure generation and a subsequent TS-search between structure pairs related by an elementary reaction, called HeurEx. Our focus is on the systematic investigation of protonation and reduction reactions. To make sure that all side-reactions are captured, all sites to which protons might be added are considered in a combinatorial way. The number of protons to be added varies between zero and the number of actual proton sites in the substrate under consideration. With this program at hand, the Chatt–Schrock cycle of nitrogen fixation will be studied in detail.

Since parts of this thesis are already published, a list of papers that originated from this work is given in Appendix 3. Parts of this introduction have been published in Refs. [23–25].

2

Case Studies 1: Multitude of Stable Intermediate Structures

2.1 Catalytic activity differences in [FeFe] hydrogenases

[FeFe] hydrogenases catalyze the formation of H_2 from protons and electrons. They are the most efficient ones compared to [NiFe] hydrogenases and [Fe] hydrogenases, which catalyze the same reaction [26–30]. Our experimental coworkers (D. Krawietz, M. Winkler and T. Happe, University Bochum) observed differences in the catalytic activity of the [FeFe] hydrogenase HydA1 *Chlamydomonas reinhardtii* (wild type, WT) and its Arg187Asp (R187D) mutant. The R187D mutant is more efficient in catalyzing hydrogen than the wild type. This computational study is performed to provide structural and energetic information to better understand the experimentally observed differences in the catalytic activity of the WT and its Arg187Asp mutant. The active center of the [FeFe] hydrogenase is referred to as H-cluster, which is well characterized [28–31] and consists of an Fe_4S_4 cubane connected by a cysteine-sulfur bridge to a $[2Fe]_H$ subcluster. In the catalytic cycle of H_2 formation the H-cluster occurs in two or three different oxidation states (which is currently debated [32, 33]): the oxidized form (H^{ox}) and the reduced from (H^{red}), which are definitely involved in the catalytic H_2 formation, and a super-reduced species (H^{sred}), which may also play a role. We study the coordination energies of H^+, H_2, CO and O_2 to the $[2Fe]_H$ subcluster and of $HOO^\bullet$ to the Fe_4S_4 cluster of the H-cluster of both systems. The H^+ and H_2 coordination energies may indicate differences in the catalytic formation of H_2 from protons and electrons accomplished by the respective hydrogenase system. The O_2 and $HOO^\bullet$ coordination energies are calculated, because in our previous theoretical work, these reactive oxygen species turned out to be important for the H-cluster decomposition [34–38]. Furthermore, we analyze the energy gaps between the highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO) of the oxidized and reduced H-cluster in the different model systems, because a small HOMO-LUMO gap of the reduced (and oxidized) H-cluster has been considered important for the catalytic activity of the hydrogenase [39, 40].

In the crystal structures of the [FeFe] hydrogenase HydA1 *Chlamydomonas reinhardtii* and its Arg187Asp mutant, the $[2Fe]_H$ subcluster next to the Fe_4S_4 cluster is absent, because the protein structure with the $[2Fe]_H$ subcluster could not yet be crystallized. Therefore we structurally align the WT and mutant X-ray structures with the crystal structure of the [FeFe] hydrogenase of *C. pasteurianum* (3C8Y), which has a sequence identity to the hydrogenase *C. reinhardtii* sequence of 25% and contains the $[2Fe]_H$ subcluster. In the newly generated WT structure with inserted $[2Fe]_H$ subcluster, the distance between the bridging cysteine sulfur atom of the $[2Fe]_H$ subcluster and the neighboring iron atom of the $[2Fe]_H$ subcluster is significantly larger than in 3C8Y (3.1 Å instead of 2.3 Å). Thus, this distance is optimized to 2.3 Å by slightly manually moving the whole $[2Fe]_H$ subcluster within the cavity.

Out of the WT and mutant protein structures with manually inserted $[2Fe]_H$ subcluster, we select model structures of two different sizes (referred to as large and small model system in the following) for the quantum mechanical study. The choice of the two sizes allows us to investigate the effect of the immediate protein environment on the calculated properties in the small model system, where only amino acids in close proximity of the $[2Fe]_H$ subcluster are taken into account, whereas in the large model system a significantly larger number of amino acids close to the H-cluster is considered. The resulting model systems are depicted in Figures 2.1 and 2.2. The choice of the amino acids, the cutting of the protein backbone, and the fixation in the large and the small model system are described in detail in Appendix A.

Since the R187D mutation introduces a charge difference of two elementary units, we also create a reference structure with a charge difference of one elementary unit with respect to the WT and the R187D mutant, i.e., there is no charge at amino acid position 187. For the large model system we generate an R187A model. In the small model system we remove amino acid 187 (labeled R187–).

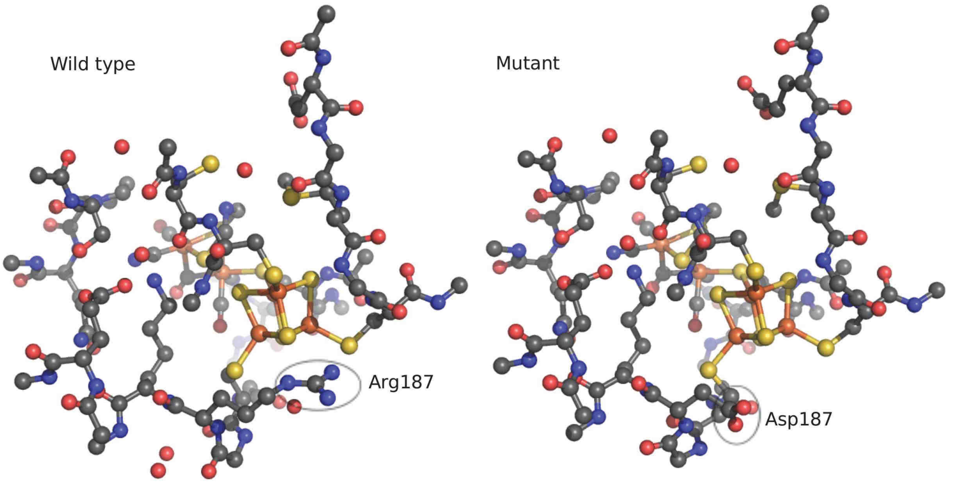


Figure 2.1: Large model systems of the H-cluster in the wild type hydrogenase of *C. reinhardtii* (left) and its R187D mutant (right). The proximal Fe_4S_4 cluster is included in both models together with residues Glu373, Met375, Cys129 (protonated), Cys377, Cys381, Cys130, Cys185, Lys188, Glu191, Ser153, Gln155, Pro154 and Ala54 (labeling according to protein sequence of HydA1). We fixed the backbone C, N and O atoms at their crystal positions. The H atoms are left out for clarity.

First of all, we perform a constrained optimization of the small and large model systems of the WT and the R187D mutant, which preserves the anchoring of the models in the full protein. The computational details are described in Appendix A. Noteworthy, the $[2Fe]_H$ subcluster stays within the cavity after TPSS/RI/def2-SVP/COSMO(ϵ = 4) structure optimizations in all model systems. Hydrogen bonding interactions between CN^- bound to the distal Fe atom of the $[2Fe]_H$ subcluster and amino acids Glu191, Ser153 and the backbone of Gln155 and Pro154 are found. In a recent computational study of hydrogenase *C. pasteurianum* in our group it has been shown that these hydrogen bonding interactions lead to the kinetic hindrance of the μ-H formation [41].

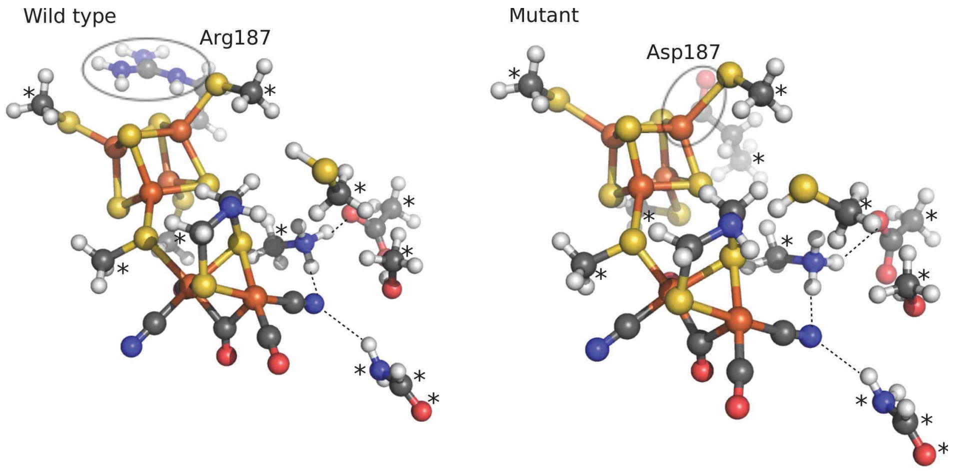


Figure 2.2: Small model systems of the H-cluster in the wild type hydrogenase of *C. reinhardtii* (left) and its Arg187Asp mutant (right). The proximal Fe_4S_4 cluster is included in both models together with parts of amino acids Lys188, Glu191, Ser153, Gln155, Pro154, Cys377, Cys381, Cys130, Cys185 and protonated Cys129 (labeling according to protein sequence of HydA1). Atoms labeled by an asterisk were fixed at their crystal positions.

2.1.1 Single coordination energies

In Tables 2.1 and 2.2 the TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) coordination energies for H_2, O_2 and $HOO^\bullet$ to the H-clusters of the WT, the R187A and R187D mutant are reported. The coordination energies obtained for the small and the large model system are qualitatively similar, which demonstrates that the small model system already resembles the protein embedding reasonably. Nevertheless, the electronic energy differences between H^{ox} and H^{red} vary by up to 20 kcal/mol. Also the $HOO^\bullet$ coordination energies to the Fe_4S_4 cubane differ significantly, about 10 kcal/mol. This effect might be due to the fact that in the small model system we did not take into account any additional amino acids around the cubane besides the four cysteines and amino acid 187, but focused on the $[2Fe]_H$ subcluster, which is in contrast to the large model system.

To investigate the effect of the fixation, which represents the anchoring of the small model system in the full protein, on the properties studied, we probe two different types of fixations at the example of the small model system. Until now, we have only kept the C atoms of the side chains fixed (as shown in Figure 2.2). In a second step, we also constrain one hydrogen atom next to each fixed C atom at a distance of 1.1 Å. In Table 2.3 we show the coordination energies obtained for the small model system with the more restrictive fixation. We observe that the more restrictive fixation does only lead to small differences in the coordination energies, as for example for the H_2 coordination energy to the $[2Fe]_H$ subcluster.

Table 2.1: Coordination energies: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) electronic energy differences (ΔE_{el} in kcal/mol) between H^{ox}/H^{red}, H^+, H_2, O_2 and CO coordination energies to the distal Fe atom of the H-cluster and $HOO^\bullet$ coordination energies to the Fe_4S_4 cubane for the wild type (WT), R187– and R187D **small model systems**. $\Delta E_{el}(H^+)_{int}$ is the energy difference between the system protonated at the NH bridge and the one protonated at the distal Fe atom, therefore it is labeled "int" (internal). -* means that the respective molecule dissociates.

	WT (+) $H^{ox}/H^{red}/H^{sred}$	R187–(0) $H^{ox}/H^{red}/H^{sred}$	R187D (−) $H^{ox}/H^{red}/H^{sred}$
$\Delta E_{el}(H^{ox}/H^{i}red)/$	−35.1/	−22.8/	−11.5/
$\Delta E_{el}(H^{red}/H^{sred})$	−9.1	+11.5	+21.3
$\Delta E_{el}(H^+)_{int}$	+1.4/−8.5	n.a./n.a.	−3.2/−12.6
$\Delta E_{el}(H_2)$	−4.9/+7.2	+1.6/-*	+3.7/+11.2
$\Delta E_{el}(O_2)$	−21.6/−24.0	−23.1/−17.3	−21.1/−18.8
$\Delta E_{el}(HOO^\bullet)$	-*/−27.1	−18.2/−20.9	−18.1/−21.9
$\Delta E_{el}(CO)$	−29.8/−23.3/−21.0	n.a./n.a./−3.8	−26.3/−18.4/−18.7

In Table 2.4 the charges (in elementary units) and unpaired electron numbers of the WT and R187D H^{ox}, H^{red} and H^{sred} clusters in the large model systems are reported. The distal (Fe_d) and proximal (Fe_p) Fe atom of H^{sred} do not have unpaired electrons, instead unpaired electrons can be found on the cubane S atoms, which are furthest away from the $[2Fe]_H$ subcluster. From the charges as well as the unpaired electrons on the iron and sulfur, one can see that between H^{ox} and H^{red} the biggest differences occur on Fe_d and Fe_p, whereas the charges and unpaired electron numbers stay relatively constant for the cubane iron and sulfur atoms. If we compare H^{red} with H^{sred}, we observe that the charge on the H^{sred} cubane iron and sulfur atoms is twice as large as the charge on the H^{red} cubane. This may indicate that the first electron accepted by H^{ox} is mainly localized on the $[2Fe]_H$ subcluster, whereas a further reduction of H^{red} mainly affects the cubane.

Table 2.2: Coordination energies: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) electronic energy differences (ΔE_{el} in kcal/mol) between H^{ox} and H^{red}, H_2 and O_2 coordination energies to the distal Fe center of the H-cluster and $HOO^{\bullet}$ coordination energies to the Fe_4S_4 cubane for the wild type (WT), R187A and R187D **large model systems.** -* means that the respective molecule dissociates.

	WT (+)	R187A(0)	R187D (−)
	$H^{ox}/H^{red}/H^{sred}$	$H^{ox}/H^{red}/H^{sred}$	$H^{ox}/H^{red}/H^{sred}$
$\Delta E_{\mathrm{el}}(H^{ox}/H^{red})/$	−63.8/	−26.3	−27.2/
$\Delta E_{\mathrm{el}}(H^{red}/H^{sred})$	−39.3		+0.7
$\Delta E_{\mathrm{el}}(H^{+})_{\mathrm{int}}$	+1.0/−11.8/−14.5		−6.6/−19.5/−28.7
$\Delta E_{\mathrm{el}}(H)_{\mathrm{int}}$	−3.1/-/		−2.7/-/
$\Delta E_{\mathrm{el}}(H_2)$	−1.0/+7.8		+0.6/-*
$\Delta E_{\mathrm{el}}(O_2)$	−15.7/−34.9		−17.7/−36.3
$\Delta E_{\mathrm{el}}(HOO^{\bullet})$	−6.1/−10.8		−1.1/−9.0
$\Delta E_{\mathrm{el}}(CO)$	−27.9/−18.7/−16.0		−26.1/−17.7/−16.0

Table 2.3: Coordination energies: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) electronic energy differences (ΔE_{el} in kcal/mol) between H^{ox}/H^{red}, H_2 and O_2 coordination energies to the distal Fe center of the H-cluster and $HOO^{\bullet}$ coordination energies to the Fe_4S_4 cubane for the WT and R187– **small model systems** with CH's fixed at a distance of 1.1 Å. -* means that the respective molecule dissociates.

	WT (+)	R187–(0)
	H^{ox}/H^{red}	H^{ox}/H^{red}
$\Delta E_{\mathrm{el}}(H^{ox}/H^{red})$	−33.7	−20.3
$\Delta E_{\mathrm{el}}(H_2)$	−1.8/n.a.	
$\Delta E_{\mathrm{el}}(O_2)$	−21.2/-19.9	
$\Delta E_{\mathrm{el}}(HOO^{\bullet})$	-*/−26.9	

Table 2.4: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) charges (in elementary units) and unpaired electron numbers on selected atoms. Fe_d is the distal and Fe_p the proximal Fe atom.

		total charge $Fe_d/Fe_p/(Fe+S)_{cubane}$	unpaired electrons Fe_d/Fe_p
WT	H^{ox}	-0.28/-0.38/-0.25	0.91/0.17
	H^{red}	-0.47/-0.30/-0.3	0.14/-0.08
	H^{sred}	-0.48/-0.28/-0.6	0.01/0.01
R187D	H^{ox}	-0.30/-0.38/-0.20	0.94/0.18
	H^{red}	-0.51/-0.26/-0.22	0.02/0.02
	H^{sred}	-0.51/-0.27/-0.56	0.0/0.01

2.1.2 Coordination energy differences

In Table 2.5 the TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) energy differences between the coordination energies for H^+, H_2, CO, O_2 and $HOO^\bullet$ to the H-clusters of the WT and the R187D mutant are reported for the small and the large model system. Since the coordination energies of the small and the large model systems are relatively similar, we assume that the small model system resembles the protein embedding reasonably well. Nevertheless, we observe that the electronic energy differences between H^{ox} and H^{red} as well as the $HOO^\bullet$ coordination energies differ significantly for the large and small model systems (by up to 30 kcal/mol and 20 kcal/mol, respectively). This effect might be due to the fact that in the small model system we did not take any additional amino acids around the cubane besides the four cysteines and amino acid 187 into account, which is in contrast to the large model system.

Overall, the coordination energies for the WT and the R187D mutant do not differ strongly, but there are several trends, that indicate a higher H_2 formation activity of the R187D mutant compared to the WT. First of all, the internal protonation energy from the bridging NH of the $[2Fe]_H$ subcluster to the distal Fe atom of the R187D mutant is by 4–14 kcal/mol more exothermic than of the WT H^{ox}, H^{red} and H^{sred} clusters, respectively. Secondly, the H_2 coordination energy is by 2–10 kcal/mol more endothermic for the R187D mutant compared to the WT, which indicates

Table 2.5: Comparison of WT and R187D mutant: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) electronic energy differences (ΔE_{el} in kcal/mol) between H^{ox} and H^{red} and between H^{red} and H^{sred}, respectively, HOMO-LUMO gaps ($\Delta\epsilon_{\mathrm{HOMO-LUMO}}$), H_2 and O_2 coordination energies to the distal Fe center of the H-cluster and $HOO^{\bullet}$ coordination energies to the Fe_4S_4 cubane for the wild type (WT) and R187D **large and small model systems.** $\Delta E_{\mathrm{el}}(H^+)_{\mathrm{int,1}}$ is the energy difference between the system protonated at the NH bridge and the one protonated at the distal Fe atom, therefore it is labeled "int" (internal). In a second internal proton transfer step from the NH bridge to the distal iron atom a H_2 molecule is formed. The corresponding energy difference is listed under $\Delta E_{\mathrm{el}}(H^+)_{\mathrm{int,2}}$. The H^{red} and H^{sred} clusters of the large model system are protonated at the NH bridge of the $[2Fe]_H$ subcluster leading to an NH_2^+ group.

	WT (+) → R187D (−)		
	Large model	Small model	Results
	$H^{ox}/H^{red}/H^{sred}$	$H^{ox}/H^{red}/H^{sred}$	
$\Delta E_{\mathrm{el}}(H^+)_{\mathrm{int,1}}$	−7.6/−7.7/−14.2	−4.6/−4.1/–	The H^+ shift from NH_2^+ to Fe_d is more exothermic for R187D than for the WT.
$\Delta E_{\mathrm{el}}(H^+)_{\mathrm{int,2}}$	+0.4/–/–	–/–/–	The energy of the H_2 formation from an NH_2^+ bridge and a Fe_d with a hydride ion is approx. the same for WT and mutant.
$\Delta E_{\mathrm{el}}(H_2)$	+1.6/n.a./–	+2.9/+9.9/–	The H_2 coordination energy of R187D is more endothermic compared to the WT.
$\Delta E_{\mathrm{el}}(O_2)$	−2.0/−1.4/–	+0.5/+5.2/–	R187D is slightly more sensitive towards O_2 compared to the WT.
$\Delta E_{\mathrm{el}}(HOO^{\bullet})$	+5.0/+1.8/–	n.a./+5.2/–	H^{ox}/H^{red} of R187D is less reactive towards $HOO^{\bullet}$ compared to the WT.
$\Delta E_{\mathrm{el}}(CO)$	+2.1/+1.0/+0.0	+3.5/+11.2/–	The CO coordination energy of R187D H^{ox}/H^{red} is more endothermic for R187D compared to the WT.
$\Delta E_{\mathrm{HOMO/LUMO}}$	−10.0/+12.2/–	−11.7/+7.0/–	$\Delta E_{\mathrm{HOMO/LUMO}}$ decreases for H^{ox} of R187D and increases for H^{red} compared to the WT.
$\Delta E_{\mathrm{el}}(H^{ox/red})$/ $\Delta E_{\mathrm{el}}(H^{red/sred})$	+36.6/+40.0 +1.1/+17.9 (same charges for each redox pair)	+23.6/+30.4	The reduction of R187D is more endothermic than of the WT → H^{sred} gets unstable, but: the charge artifact is large!

that the H_2 affinity to the distal Fe atom is lower than in the WT. Moreover, the $HOO^\bullet$ coordination energies are slightly less exothermic (by 2–5 kcal/mol) for the mutant than for the WT. This observation might be taken as an indication of a lower susceptibility towards reactive oxygen species of the R187D mutant. Although the O_2 coordination energies are slightly more exothermic (by 1–2 kcal/mol) for the mutant than for the WT in the large model system, this difference is still relatively small. Hence, the calculated bonding energies for H^+, H_2, O_2 and $HOO^\bullet$ are in accord with the experimental observations.

2.1.3 Structural alignments

In Figure 2.3 the structural alignments of the TPSS optimized H-clusters (H^{ox}, H^{red} and H^{sred}) in the wild type hydrogenase of *C. reinhardtii* (left) and its R187D mutant (right) are shown.

Only a small variation in the terminal CO ligands of the H^{ox}, H^{red} and H^{sred} clusters is observed. In general, the clusters for the differently charged species are relatively similar.

2.1.4 Molecular orbital visualizations

It has been found by Bruschi et al. [39, 40] that a small energy gap between the highest-occupied (HOMO) and lowest-unoccupied molecular (LUMO) orbital of the H^{red} cluster of [FeFe] hydrogenase is an important diagnostic for its functionality. Since electrons donated by the Fe_4S_4 clusters have to be accepted by the H-cluster, a small HOMO-LUMO gap would facilitate this step. Hence, the HOMO-LUMO gaps ($\Delta\epsilon_{\text{HOMO-LUMO}}$) are utilized as a qualitative descriptor for the electronic energy difference between the reduced and oxidized H-clusters in the different model systems. Baerends and Parr have shown that the Kohn–Sham (KS) orbitals obtained from DFT calculations (with commonly employed density functionals, like TPSS, BP86, B3LYP) are appropriate for the investigation of chemical properties [42–46]. For this, it is important to compare the order, symmetry and shape of the KS orbitals [46], to make sure that the corresponding molecular orbitals are chosen for the analysis of different structural models.

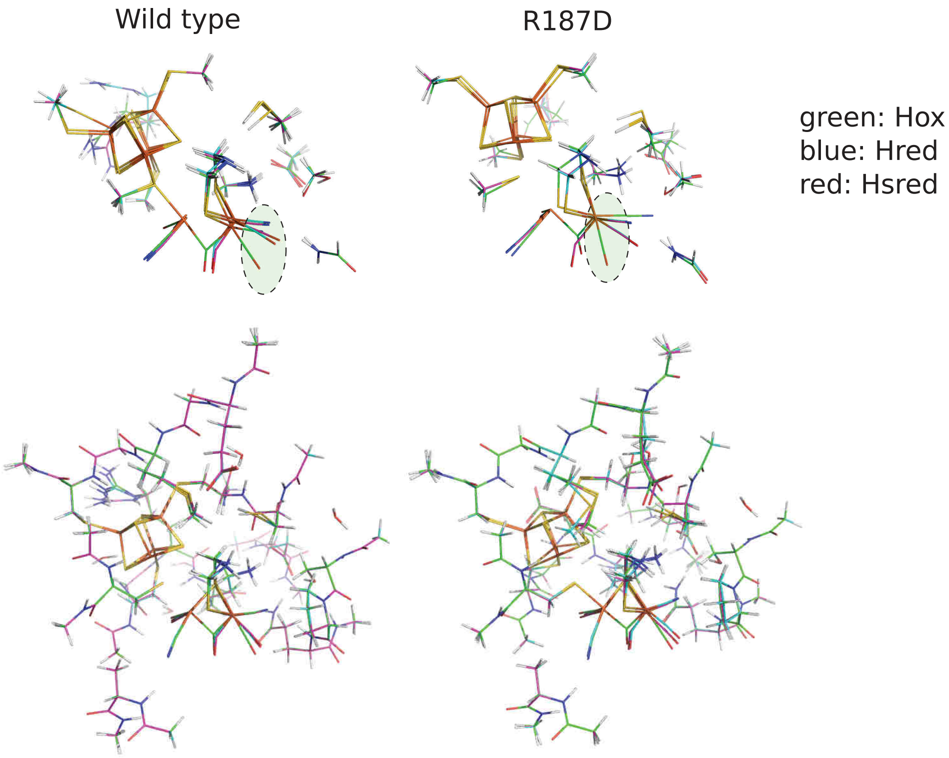


Figure 2.3: Structural alignments of the TPSS optimized H-clusters (H^{ox}, H^{red} and H^{sred}) in the wild type hydrogenase of *C. reinhardtii* (left) and its R187D mutant (right). The minimal (top) and the large model systems (bottom) are shown.

The molecular orbitals of H^{ox} and H^{red} of the small model system are displayed in Figures 2.4 and 2.5. It can be seen that the TPSS/TPSSh and B3LYP molecular orbitals are clearly located on either the Fe_4S_4 cubane or the $[2Fe]_H$ subunit, besides the HOMO of the H^{ox} cluster of the R187D mutant. Still, the HOMO's and LUMO's of the WT, R187– and R187D structures are not fully identical in shape and the ordering of the molecular orbitals in the different model systems is not the same. As a first result, we observe that the HOMO's and LUMO's of R187D and R187– are very similar in contrast to the wild type H^{ox}, where the HOMO and LUMO are both located on the cubane.

To obtain the energy gaps that correspond to the electronic energy difference between the reduced and oxidized H-cluster, we compare the molecular orbitals of one model system with the HOMO and LUMO shapes obtained for the other systems and calculate all respective energy

Table 2.6: TPSS/RI, TPSSh and B3LYP def2-SVP/COSMO($\epsilon = 4.0$) **HOMO-LUMO gaps** ($\Delta\epsilon_{\text{HOMO-LUMO}}$) in kcal/mol for H^{ox} and H^{red} of the wild type (WT), R187– and R187D **small model system**. In every column the molecular orbital energy differences of the system labeled above (WT, R187– and R187D) are calculated with respect to the type of HOMO and LUMO molecular orbitals of the system labeled at the beginning of each row.

	WT(+)	R187–(0)	R187D(−)
H^{ox}	TPSS/TPSSh/B3LYP	TPSS	TPSS/TPSSh/B3LYP
WT	20.8/45.3/55.2	22.0	20.3
R187–	23.5	13.3	15.3
R187D	-	-	9.1/37.4/63.8
H^{red}			
WT	3.8/18.4/30.9(20.9)	14.1	-
R187–	3.8	11.6	-
R187D	17.3	-	10.8/30.4/45.2

gaps. The $\Delta\epsilon_{\text{HOMO-LUMO}}$'s for the H^{ox} and H^{red} clusters are shown in Table 2.6. In every column the molecular orbital energy differences of the systems (WT, R187– and R187D) are given. $\Delta\epsilon_{\text{HOMO-LUMO}}$ of the WT H^{red} cluster is small compared to $\Delta\epsilon_{\text{HOMO-LUMO}}$ of the reference and mutant system, only 3.8 kcal/mol compared to 11.6 kcal/mol and 10.8 kcal/mol, and increases in the mutant and reference system to 10–11 kcal/mol. To investigate the dependence of $\Delta\epsilon_{\text{HOMO-LUMO}}$ on the density functional chosen (especially on the exact Hartree–Fock-type (HF) exchange admixture in the density functional), we perform TPSSh (10% HF exchange) and B3LYP (20% HF exchange) single-point calculations on the TPSS optimized structures of the WT and mutant small model systems. The TPSSh and B3LYP $\Delta\epsilon_{\text{HOMO-LUMO}}$ data, which are also listed in Table 2.6, show the same qualitative trend as the TPSS $\Delta\epsilon_{\text{HOMO-LUMO}}$ data. Nevertheless, it must be emphasized that we can only use the KS orbitals as a qualitative descriptor.

The corresponding $\Delta\epsilon_{\text{HOMO-LUMO}}$ gaps of the large model systems are listed in Table 2.7.

In total, the differences between the WT and the mutant $\Delta\epsilon_{\text{HOMO-LUMO}}$'s of the H^{ox} and H^{red} clusters amount to approximately −10 kcal/mol and +10 kcal/mol, respectively. The same trend is observed for the small

Table 2.7: TPSS/RI/def2-SVP/COSMO($\epsilon = 4.0$) **HOMO-LUMO gaps** ($\Delta\epsilon_{\mathrm{HOMO-LUMO}}$) in kcal/mol for H^{ox} and H^{red} of the **large model system**. In every column the molecular orbital energy differences of the system labeled above (WT and R187D) are calculated with respect to the type of HOMO and LUMO molecular orbitals of the system labeled at the beginning of each row. WT(+), dist. means that the positively charged arginine amino acid was manually distorted by a dihedral backbone angle such that the positive charge is further away from the Fe_4S_4 cubane.

	WT(+)	WT(+), dist.	R187D(−)
H^{ox}			
WT	12.5	-	-
WT(+), dist.	-	19.9	-
R187D	-	-	2.5
H^{red}			
WT	1.1	-	-
R187D	-	-	13.3

and the large model systems. Hence, independent of the larger protein environment taken into account in the large model system, a significant difference in the $\Delta\epsilon_{\mathrm{HOMO-LUMO}}$s of the WT and the mutant is observed.

2.1.5 Calculated infrared vibrational spectra

In Figure 2.6 the calculated infrared (IR) spectra of the mutant (R187D) and WT H^{ox} and H^{red} in the small model system are reported.

The largest difference in the IR intensities of the WT and the R187D mutant H-clusters is observed for the CO stretching vibration of the terminal CO at the $[2Fe]_{H^{red}}$ subcluster. In the WT, the strong CO vibration is mostly coupled to the stretching vibration of the CO that is closest to the Fe-S cubane, whereas in the mutant the strong CO vibration is mostly coupled to the bridging CO. Besides this CO stretching vibration, two vibrations in the backbone (a NH_3^+ and OH vibration) of the protein environment show large IR intensities. Noteworthy, the IR intensities obtained for the small model system might deviate from the ones in the protein due to the restricted model size. However, the IR spectra cal-

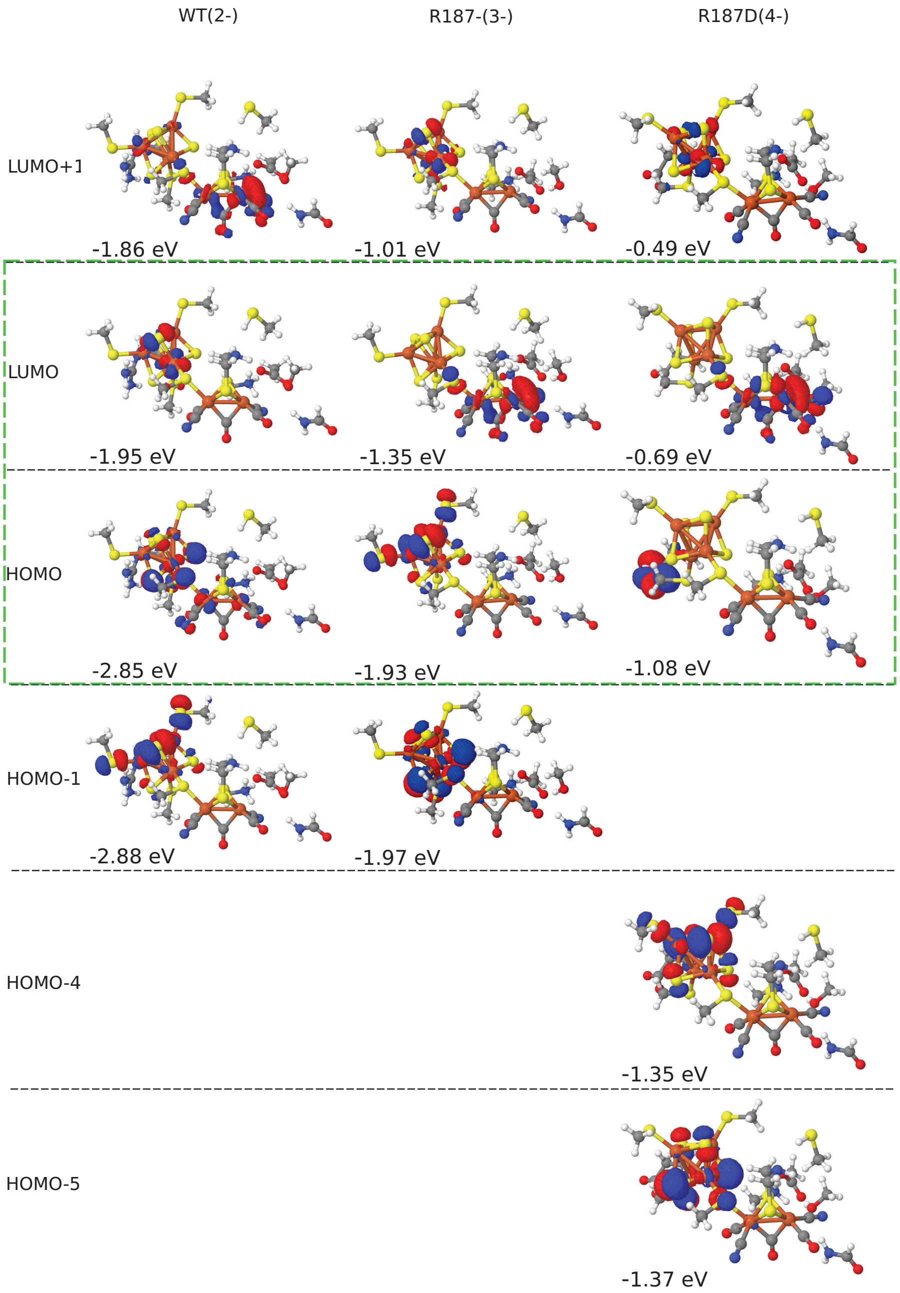

WT(2-)
R187-(3-)
R187D(4-)
LUMO+1
-1.86 eV
-1.01 eV
-0.49 eV
LUMO
-1.95 eV
-1.35 eV
-0.69 eV
HOMO
-2.85 eV
-1.93 eV
-1.08 eV
HOMO-1
-2.88 eV
-1.97 eV
HOMO-4
-1.35 eV
HOMO-5
-1.37 eV

Figure 2.4: Highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO) visualization of the TPSS/RI/def2-SVP/COSMO($\epsilon = 4$) optimized wild type, R187– and R187D $\mathbf{H}^{\mathrm{ox}}$ structures in the **small model system**. The energies of the molecular orbitals are given in eV.

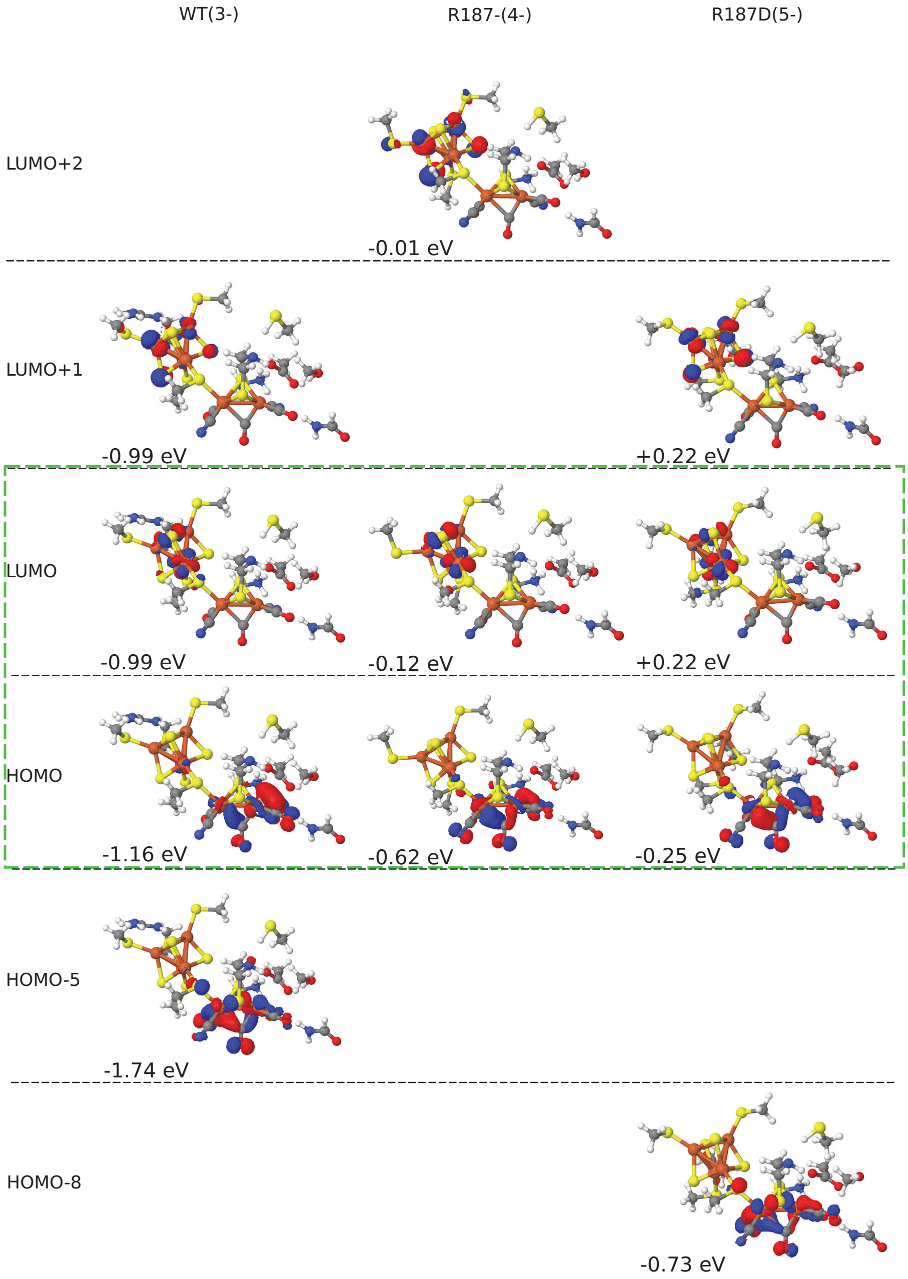

WT(3-)
R187-(4-)
R187D(5-)
LUMO+2
-0.01 eV
LUMO+1
-0.99 eV
+0.22 eV
LUMO
-0.99 eV
-0.12 eV
+0.22 eV
HOMO
-1.16 eV
-0.62 eV
-0.25 eV
HOMO-5
-1.74 eV
HOMO-8
-0.73 eV

Figure 2.5: Highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO) visualization of the TPSS/RI/def2-SVP/COSMO($\epsilon = 4$) optimized wild type, R187– and R187D $\mathbf{H}^{\text{red}}$ structures in the **small model system**. The energies of the molecular orbitals are given in eV.

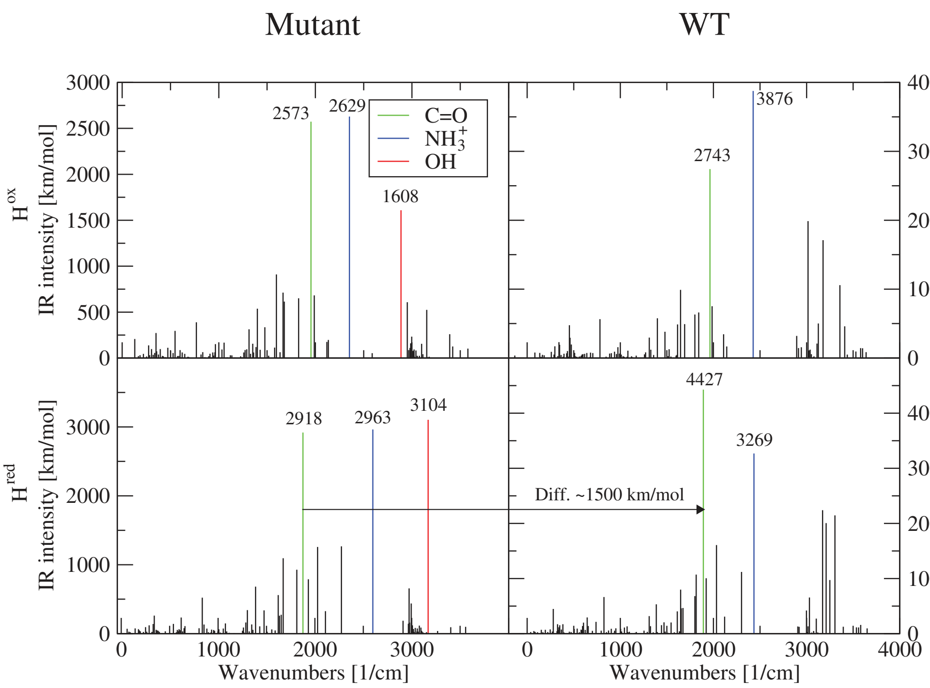


Figure 2.6: TPSS/RI/def2-SVP/COSMO($\epsilon = 4$) IR spectra of the H^{ox} and H^{red} of the mutant (R187D) and WT in the **small model system.**

culations for the large model system are computationally much more demanding.

2.1.6 Catalytic cycle of H_2 formation

Several computational studies on the mechanism of H_2 formation at [FeFe] hydrogenases have already been performed [41, 47–51]. In this Section, we calculate the energy differences for all reaction steps of the catalytic H_2 production cycle of the specific system, i.e., the WT and the R187D mutant. As a reference, we also report the ΔE values for the WT.

In the $\Delta\Delta E$ (which is ΔE_{mutant}-ΔE_{WT}), the energies of the solvated protons drops out and the $\Delta\Delta E_{\text{cycle}}$ for the whole cycle is very close to zero. The ΔE_{cycle} values of the mutant and the WT, respectively, must correspond to the energy difference between the H_2 and the two protons and two electrons, which does not depend on the catalyst. The energy of the solvated H^+ ion is fitted according to the condition $\Delta E(\text{WT})_{\text{cycle}} = -740$

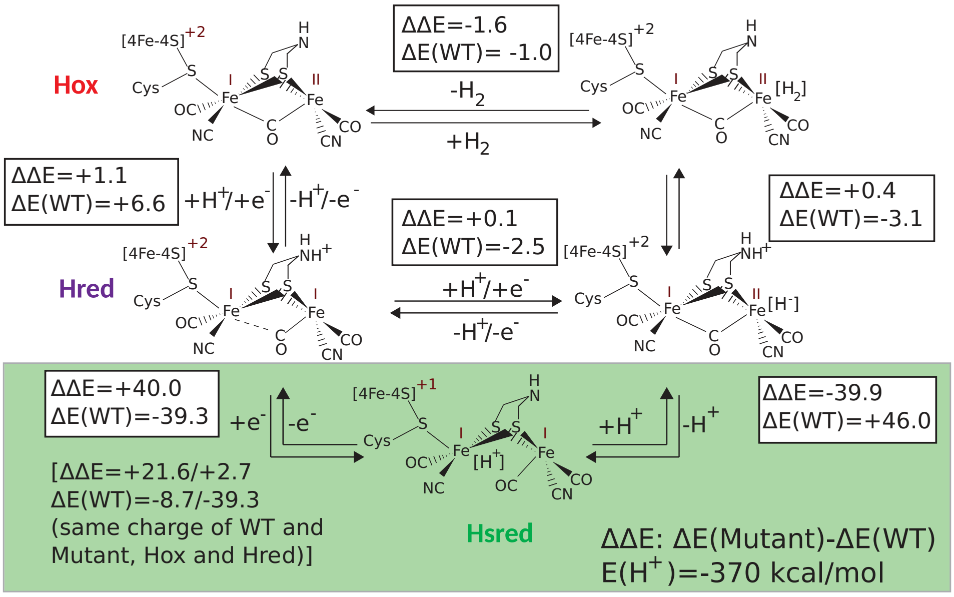

Hox
[4Fe-4S]+2
ΔΔE=-1.6
ΔE(WT)= -1.0
-H2
+H2
ΔΔE=+1.1
ΔE(WT)=+6.6
+H+/+e-
-H+/-e-
ΔΔE=+0.1
ΔE(WT)=-2.5
ΔΔE=+0.4
ΔE(WT)=-3.1
[H2]
Hred
+H+/+e-
-H+/-e-
[H-]
ΔΔE=+40.0
ΔE(WT)=-39.3
+e-
-e-
[4Fe-4S]+1
[H+]
+H+
-H+
ΔΔE=-39.9
ΔE(WT)=+46.0
[ΔΔE=+21.6/+2.7
ΔE(WT)=-8.7/-39.3
(same charge of WT and
Mutant, Hox and Hred)]
Hsred
ΔΔE: ΔE(Mutant)-ΔE(WT)
E(H+)=-370 kcal/mol
Cys
S
Fe
OC
NC
CO
CN
H
N
NH+
I
II

Figure 2.7: TPSS/RI/def2-SVP/COSMO($\epsilon = 4$) ΔE and $\Delta\Delta E$ values (in kcal/mol) for all steps in the catalytic cycle of the hydrogen production of the mutant (R187D) and WT in the **large model system.**

kcal/mol (H_2 formation energy) and amounts to −370 kcal/mol, if the energy of the electron is neglected. To evaluate the reaction kinetics of the two systems, one would have to calculate also the activation barriers in addition to the energy minima. Nevertheless, an analysis of the potential energy differences already gives an estimate of the thermodynamics of the catalytic cycle.

Two different catalytic pathways for the H_2 formation are debated, which are on the one hand, the direct proton-coupled electron transfer from H^{red} to the iron-hydride species and, on the other hand, the reduction of H^{red} to H^{sred} and subsequent protonation. Only in the lower part of the diagram in Figure 2.7, in which the H^{sred} species is involved (highlighted in green) $\Delta\Delta E$ is large, about 40 kcal/mol. In all other steps, $\Delta\Delta E$ is smaller than 1.6 kcal/mol, which means that the reaction energies of the WT and mutant are very similar. However, the one-electron reduction from H^{red} to H^{sred} is energetically strongly favored for the WT, exothermic by up to 40 kcal/mol, and slightly endothermic for the mutant, +0.7 kcal/mol. Thus, if we assume that the reaction barrier of the electron reduction is not very large, the H^{red} state is like a sink of the catalytic cycle of the WT. This may explain the increased catalytic activity of the mutant compared to the WT. Furthermore, it has been experimentally observed that the H^{sred} state of the mutant is not stable, which is in agreement with our computational results, because the WT H^{sred} probably reacts further with H^+ significantly faster than in the mutant.

2.1.7 Consequence for the chemistry of [FeFe] hydrogenases

Structure optimizations of the [FeFe] hydrogenase HydA1 *Chlamydomonas reinhardtii* and R187D mutant model systems with manually inserted $[2Fe]_H$ subcluster demonstrate that the H-cluster is stable in the protein environment as present in the crystal structures, if decomposition of the structure or significant rearrangement is taken as a criterion, which we did not observe. A similar hydrogen bonding network between the $[2Fe]_H$ subcluster and the protein environment as observed in hydrogenase *C. pasteurianum* is formed.

The H^+, H_2, CO, O_2 and $HOO^\bullet$ coordination energies to the H cluster in the wild type and the R187D mutant differ by up to 12 kcal/mol, indicating a stronger catalytic activity of the R187D mutant compared to the WT, especially due to the significantly more exothermic energy corresponding to the internal first proton transfer from the NH bridge to the distal iron atom.

According to Bruschi et al. [39, 40] a small HOMO-LUMO gap of the H^{red} cluster of [FeFe] hydrogenase is important for its functionality. The calculated HOMO-LUMO gaps of the wild type H^{red} cluster in the small and large model systems are by approximately 7.0 kcal/mol smaller than the $\Delta\epsilon_{\text{HOMO-LUMO}}$ gaps of the R187D mutant H^{red} cluster, which means that electron transfer to the H cluster is facilitated. Although the DFT molecular orbitals are simply a qualitative descriptor, our results indicate that the R187D mutation has a significant effect on the $\Delta\epsilon_{\text{HOMO-LUMO}}$ gaps and thus on the electronic energy difference between the oxidized and reduced H-cluster.

From the calculations of all single steps in the catalytic H_2-formation cycles of the WT and the mutant, we observed that the largest differences occur, if the H^{sred} species is involved in the catalytic cycle. The electron reduction from H^{red} to H^{sred} species is significantly more exothermic for the WT than for the mutant. Hence, the occurrence of the H^{sred} state may hinder the catalytic H_2 production.

2.2 Structure–reactivity analysis of Fe_4S_4 clusters

Iron-sulfur cluster containing proteins play a key role in electron transfer processes, metabolism, signal transduction and activation of low-reactivity molecules such as N_2 and H_2 [52–54]. The structural variability of iron-sulfur complexes ranges from single Fe ions coordinated by four cysteines over Fe_2S_2 units to cubane-type Fe_4S_4 clusters and the complex structures one encounters in the active sites of, for instance, nitrogenases and [FeFe] hydrogenases [55, 56].

Here, we focus on Fe_4S_4 clusters. In contrast to synthetic Fe_4S_4 cubanes [57, 58], enzymatic cubanes are affected by their particular protein environment. The protein environment determines charges in the vicinity of the cluster, hydrogen bonds, solvent accessibility as well as the structural conformation of the first coordination sphere. Cluster properties — like relative energies, spin couplings, reduction potentials and reactivity — provide insights into cluster stability and function. Since differences in properties have been observed for clusters of different proteins [59–61], an understanding of the effect of the protein environment is highly desirable. For example, it is known that reduction potentials and reactivities of cubanes differ in various protein embeddings. Fe_4S_4 clusters in high-potential iron proteins (HiPIP) and ferredoxins operate in different oxidation states [59]. Also within the ferredoxin family remarkable variations of about 550 mV in the reduction potentials have been observed [60]. Moreover, because reactive oxygen species (ROS) damage Fe_4S_4 clusters [56], understanding the reactivity of these clusters is decisive. The influence of charged or polar groups in close vicinity to the Fe-S cluster [62–66], hydrogen-bonding effects [67, 68], ligand conformations [68] as well as solvent effects [61] have been discussed in the literature. Changes in the reactivity with respect to different ligands have also been studied [69–71]. Furthermore, theoretical investigations on synthetic as well as protein-embedded Fe_4S_4 cubanes considered spin topology and structural conformation of differently charged cubanes as well as their redox properties [61, 67, 68, 72–85]. For example, it has been found that reduction potentials are mainly affected by hydrogen bonding [61] and electrostatic effects [68, 81, 85] whereas the effect of the conformation of the ligated cubane plays a minor role. As a consequence, theoretical studies on reduction potentials of protein-embedded Fe-S clusters often neglect structural changes of the redox-site. However, differences in the distances between the sulfur atoms of two coordinating cysteine residues have been observed in HiPIP and ferredoxin crystal structures [76, 80].

The problem that emerges when the structures of enzymatic clusters are compared is that multiple structural features (e.g., distances and angles) are changed simultaneously. In order to be able to deduce structure–property relationships, the clusters must be analyzed in a systematic way. Fee et al. have addressed this question by analyzing differences in the cluster geometry with respect to an idealized tetrahedral geometry [76]. The variations they found in the S(Cys)-S(Cys) distances of HiPIP and ferredoxin structures might explain why different oxidation states are favored in the two kinds of iron-sulfur clusters and therefore why HiPIP and ferredoxins have different reduction potentials. These observations suggest that the spatial arrangement of the first ligand shell can have a significant effect on cluster properties.

The question how sensitive cluster properties are with respect to ligand distortions still requires a detailed theoretical investigation performed such that general conclusions can be drawn. For this, the structural embedding of the cluster in the protein has to be studied in detail. Cubane-type Fe_4S_4 clusters are most often ligated by four cysteines and share a consensus motif of CXXCXXC or CXXC (C denotes a cysteine and X an arbitrary amino acid) and one (or two, respectively) cysteine(s) further apart in sequence space [60]. By aligning several highly resolved crystal structures from the protein data bank (PDB), we have seen (cf. section 2.2.1.2) that the geometry of the cluster embedding is similar, but deviations in the positions of the cysteine sulfur and of the C_β atoms of up to one Å occur (C_β is the first carbon atom of the side chain and bound to the C_α atom of the protein backbone).

In this theoretical study, we systematically investigate the sensitivity of cluster properties on structural distortions in terms of correlation diagrams. We create a generalized anchoring scheme of an idealized model system and scan the properties of interest as functions of the distorted structures. Following this approach, structure–property relationships can be systematically investigated. Conclusions can be drawn on the stability of the clusters as well as on the effect of the first coordination sphere on specific cluster properties. The content of this Chapter has been published in Ref. [23].

2.2.1 A generalized anchoring scheme

The idea to investigate structure–property relationships by means of correlation diagrams has a long tradition in chemistry, but became somewhat old-fashioned in view of the capabilities of modern numerical quantum chemistry (examples are Walsh diagrams [86], Shaik-Pross models [87, 88] and Tanabe-Sugano diagrams [89, 90]). We have recently argued that correlation diagrams based on first-principle calculations can be very useful to systematize observations on related metalloenzymes in bioinorganic chemistry [91].

2.2.1.1 Symmetrized Fe_4S_4 model clusters

As model system we choose a cubane-type Fe_4S_4 cluster ligated by four HS^- groups, which are spatially fixed to mimic the constraints imposed by the protein environment. The total number of structural degrees of freedom in this model is 42 ($3N-6$, where N = number of atoms). This number is too large to systematically explore the potential energy hypersurface. It is therefore advantageous to start with an idealized symmetric cluster, which significantly reduces the number of degrees of freedom.

In X-ray crystallography and Density Functional Theory (DFT) calculations it was found that Fe_4S_4 clusters adopt a geometry with four contracted Fe-S bonds (by about 0.1 Å compared to elongated Fe-S bonds) and eight elongated Fe-S bonds [73, 92]. This so called $\bar{4}$ contraction (shown in Figure 2.11) can be explained by a Jahn–Teller-like effect [72]. Two Fe_2S_2 rhombs with ferromagnetic spin coupling are antiferromagnetically coupled with each other leading to a contraction along the axis perpendicular to the rhombs [72]. The $\bar{4}$ contraction lowers the symmetry of the cubane from T_d to D_{2d}. Hence, we assume D_{2d} symmetry for our model as the highest reasonable symmetry. For the given cluster configuration, consisting of the Fe_4S_4 cluster and four HS^- ligands, two D_{2d}-symmetric structures are possible, referred to as D_{2d}-1 and D_{2d}-2 (see Figure 2.8).

We have chosen the distances between the hydrogen atoms and the distances between the sulfur atoms of the HS^- ligands, respectively, as constraints to mimic the structural embedding of the cluster. In D_{2d}-symmetric clusters only two different H-H distances and S(H)-S(H) distances, respectively, exist (S(H) denotes the sulfur atom of the HS^- ligand).

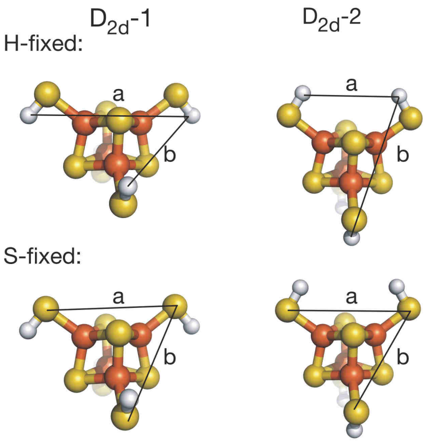


Figure 2.8: Graphical representation of the two distance constraints *a* and *b* in the two D_{2d}-symmetric clusters for fixed sulfur atoms (S-fixed) and for fixed hydrogen atoms (H-fixed). Element color code: yellow, S; orange, Fe; white, H.

This leads to a set of only two constraints, denoted *a* and *b* throughout the text. The *a* distance occurs twice in the model system and the *b* distance four times. The symmetry-related distances are connected by an S_4 symmetry operation. The two structural constraints determine all cluster structures, which are then subject to structure optimization with no point-group symmetry constraints (see Figure 2.8).

The choice of a generalized symmetric model system allows us to conveniently analyze and visualize structure–property relationships. For potential energy surfaces (PES's) and plots of reaction energies, the two distances *a* and four distances *b* are changed in steps of 0.10 and 0.20 Å, respectively, as illustrated in Figure 5.4. The distortions are performed in such a way that the lowest-energy ligand arrangement is located approximately in the center of the plot and the range of distortions has been chosen according to enzymatic cluster structures. Note that the distortions do not break the point group symmetry, but the subsequent constrained structure optimization may do so. Cluster properties can then be deduced for a specific set of constraints *a* and *b*.

Since the two different D_{2d}-symmetric starting structures have very similar properties, only the D_{2d}-2 form will be discussed below (a detailed comparison to the D_{2d}-1 structure is provided in Appendix B.2).

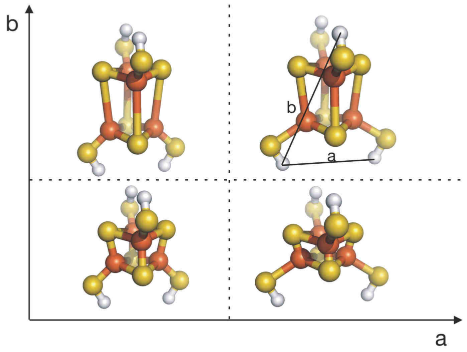


Figure 2.9: Graphical representation of the distorted starting structures for the constrained structure optimizations. The two distances a and the four distances b are changed stepwise to obtain a grid of distorted structures. Element color code: yellow, S; orange, Fe; white, H.

2.2.1.2 Comparison with protein-embedded clusters

In order to generalize our findings for the model clusters to iron-sulfur clusters occurring in proteins, we have to make sure that the structures share similar geometric features (like $FeSC_\beta$ angles and Fe-Fe and Fe-S(Cys) distances). Our distorted model structures have to be within a relevant range of structural parameters. Structures of largely increased electronic energy may be considered irrelevant and have to be avoided. For this reason, we have investigated different ferredoxin and HiPIP structures [PDB entries: 2FGO, 2Z8Q, 6FD1, 1A6L (T14C mutant of 6FD1), 2ZVS, 1ISU, 2FDN, 1B0Y, 1HLQ, 1HPI, 3H31, 2AMS, 1CP2 [93–105]] as well as a [FeFe] hydrogenase structure [3C8Y [106]], an acetylene hydratase [2E7Z [107]] and two oxidoreductases [2C42, 2YIV [108, 109]] with resolutions higher than 2.1 Å with respect to the occurring S(Cys)-S(Cys) and C_β-C_β distances of the ligated iron-sulfur clusters. Also the $FeSC_\beta$ angles and Fe-Fe and Fe-S(Cys) distances have been analyzed (cf. Appendix B.1). In Figure 2.10 clusters from the selected proteins are structurally aligned with respect to the iron atoms using ClusterGeom [110]. It can be clearly seen that the cluster embedding is quite similar (especially for the ferredoxins and HiPIPs among each other, Figure 2.10 on the left and right), but deviations in the positions of the cysteine sulfur and C_β atoms occur.

Apparently, there is a correlation between cluster structure and protein sequence. All of the ferredoxins studied feature the consensus motifs

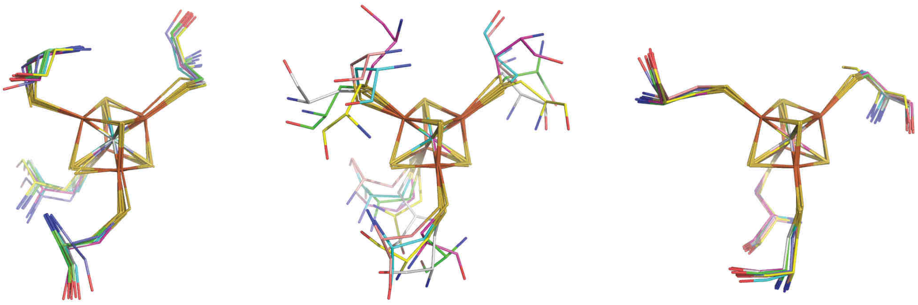


Figure 2.10: Structural alignment of selected cysteine-ligated iron-sulfur cubanes from the PDB with respect to cubane Fe atoms. Ferredoxin clusters with very similar ligand spheres are shown on the left (motif CXXCXXC), clusters with very different ligand spheres are shown in the middle, and HiPIP clusters with very similar ligand spheres are depicted on the right.

CXXCXXC or CXXC, where three or two of the cluster-coordinating cysteine residues are separated by two other amino acids (X) in sequence space. The distance between the C_β of the last cysteine of the motif and the C_β of the cysteine residue which does not belong to the motif is shortened compared to the other C_β-C_β distances. In the HiPIPs there is also a consensus motif of two cysteines separated by two arbitrary amino acids (CXXC).

The C_β atoms of the cysteines correspond to the hydrogen atoms in our model system. Both structural constraints, *a* and *b*, are in the range of the distances observed in proteins. Also the internal cluster and cluster-ligand angles and distances are very similar to the protein-embedded clusters. Hence, even though the distorted clusters of our model system do not fully match the natural ones, they share important features, which allows us to draw general conclusions from the structure–property correlation diagrams.

2.2.2 Structural strain and antiferromagnetic coupling

2.2.2.1 Analysis of possible spin couplings in $[Fe_4S_4(SR)_4]^{2-}$

By Mössbauer and electron paramagnetic resonance spectroscopy it was found that in iron-sulfur cubanes high-spin iron atoms exist which are significantly spin coupled [111]. We employ broken-symmetry DFT [12, 15–19] to model these couplings within the cluster. The antiferromagnetically coupled low-spin state (open-shell singlet with total spin quantum number $S = 0$) has been identified as the lowest one for the $[Fe_4S_4(SH)_4]^{2-}$ cluster [72, 73, 111, 112]. The excess of α- and β-spin density described by the local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value [20–22] is calculated for selected iron atoms to characterize the broken-spin symmetry (BS) wave function (determinant). Note that twice this atomic expectation value is the difference in α- minus β-occupation of the atom under study.

If the cubane is ligated by an arbitrary rest R (usually a cysteine residue), which reduces the point group symmetry, three different coupling schemes are possible (see Figure 2.11). Here, ”coupling scheme” only denotes the distribution of α-spin density excess and β-spin density excess on the iron atoms as described by the BS determinant. In the open-shell singlet BS coupling scheme two iron atoms show an excess of α-spin density and two iron atoms an excess of β-spin density. It must be noted that this spin polarization is artificial [17, 19], but crucial in order to classify different BS solutions and to guarantee an artifact-free, intrinsically consistent BS study. I.e, the energy of the BS solution may depend on the density functional chosen, but energy jumps because of a (sudden and often artificial) change in the coupling scheme must be carefully tracked. In all three coupling schemes, two ferromagnetically coupled Fe_2S_2 subclusters are antiferromagnetically coupled and the Fe-S bonds between the Fe_2S_2 layers are shortened by about 0.1 Å (BP86) in comparison to the Fe-S distances within the ferromagnetically coupled subclusters.

In the D_{2d}-symmetric model system with HS^- ligands, two of the three possible coupling schemes are energetically identical, because they are mirror images. In Figure 2.12 the BP86 optimized structures (see the Computational Methodology described in Appendix B for details) for the three possible broken-symmetry coupling schemes of $[Fe_4S_4(SH)_4]^{2-}$ are shown for the lowest-energy D_{2d}-symmetric ligand arrangement. The different coupling schemes are denoted as A, B, and C. Since B and C

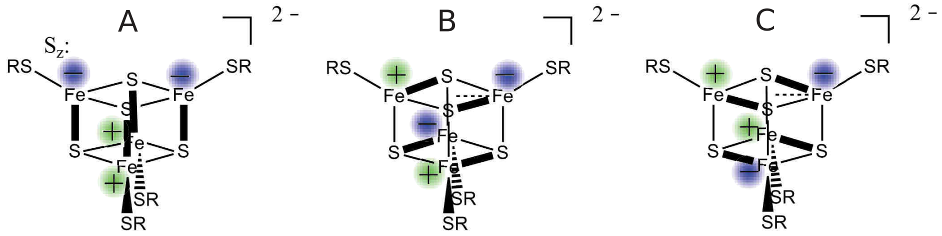


Figure 2.11: Illustration of the three broken-spin symmetry solutions for the $[Fe_4S_4(SR)_4]^{2-}$ cluster, which define the three coupling schemes A, B, and C. R denotes an arbitrary rest. The sign of the local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value is indicated by + (green) and – (blue), which corresponds to an excess of α- and β-spin density, respectively. In each BS state four Fe-S bonds are shortened (thick lines).

are energetically identical, we will only refer to A and B in the following for the D_{2d}-symmetric model system. Note that if the system is not D_{2d}-symmetric, as in the case of the cysteine-ligated clusters from metalloprotein crystal structures, the three coupling schemes differ and A, B, and C have to be defined with respect to a given orientation of the ligated clusters.

In our model system, the coupling scheme on the left in Figure 2.12 (A), where the shortened side of the Fe_4S_4 cluster is approximately parallel to the main rotation axis (S_4, C_2), represents the energetically lowest spin distribution. The electronic energy difference ΔE_{el} of coupling scheme B to coupling scheme A amounts to 3.7 kcal/mol for the lowest-energy structure with H-fixation (cf. Figure 2.12). A change from HS^- ligands to CH_3S^- or cysteine-like $CH_3CH_2S^-$ ligands, does not significantly affect ΔE_{el} as calculated with BP86: With CH_3S^- ligands ΔE_{el} is 4.3 kcal/mol and with $CH_3CH_2S^-$ ligands ΔE_{el} is 3.8 kcal/mol. For $a = 4.05$ Å and $b = 8.77$ Å, ΔE_{el} changes from 4.9 kcal/mol (HS^-) to 7.7 kcal/mol (CH_3S^-). In comparison, if we explore the structure, where a is large and b small ($a = 5.85$ Å and $b = 7.37$ Å), we obtain a ΔE_{el} of −1.0 kcal/mol (coupling scheme B being favored now) for HS^- and 0.2 kcal/mol for CH_3S^-. Hence, qualitatively the changes in ΔE_{el} follow the same trend for the different ligands, which justifies the choice of HS^- ligands in this study.

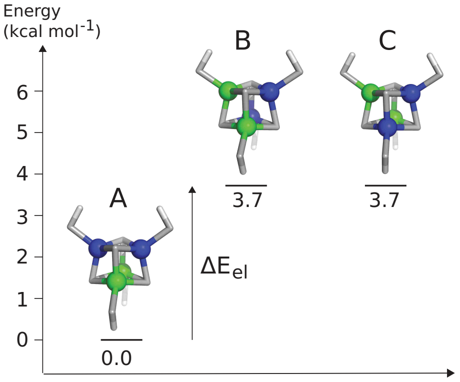


Figure 2.12: BP86/RI/def2-TZVP/COSMO optimized structures for the three BS solutions A, B, and C of $[Fe_4S_4(SH)_4]^{2-}$ with hydrogen atoms fixed at $a = 5.05$ Å and $b = 7.77$ Å. Element color code: gray, S; white, H; blue, Fe with negative local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value; green, Fe with positive local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value.

2.2.2.2 Comparison of differently charged model clusters

To systematically understand the dependence between a structural constraint and the open-shell singlet or doublet spin distribution, we investigate ΔE_{el} for the whole grid of distorted structures in Figure 2.13.

If the $[Fe_4S_4(SH)_4]^{2-}$ cluster is reduced or oxidized by one electron, several spin states can be adopted. We have chosen the doublet ($S = 0.5$) spin state here, since we obtained the lowest BP86 energies for this spin state solution. The extra electron/hole is mainly delocalized over one Fe_2S_2 subcluster leading to larger (absolute) local spin expectation values on two iron atoms and slightly smaller local spin expectation values on the other two iron atoms. Still, three different coupling schemes (as shown in Figure 2.11), which describe only the qualitative distribution of α-spin density excess and β-spin density excess on the iron atoms, are possible.

In Figure 2.13 the BP86 correlation diagrams of ΔE_{el} are presented for $[Fe_4S_4(SH)_4]^{3-}$, $[Fe_4S_4(SH)_4]^{2-}$ and $[Fe_4S_4(SH)_4]^{1-}$. Large distances b and small distances a increase ΔE_{el} until at one point an increase in b leads again to a decrease of ΔE_{el}. On the other hand, small distances b and large distances a decrease ΔE_{el}. If ΔE_{el} drops below zero, coupling scheme B is favored. This means that the shortened axis of the cubane (antiferromagnetic coupling) is preferred along the direction in which the S-S distances (or H-H distances) are large. Moreover, decreasing the S-S and H-H distances, respectively, of the HS^- ligands on one Fe_2S_2

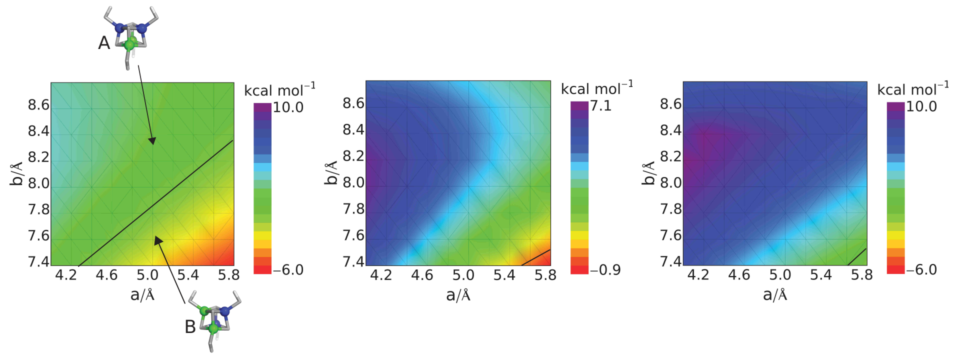


Figure 2.13: BP86/RI/def2-TZVP/COSMO electronic energy difference for the coupling schemes A and B of structurally constrained, originally D$_{2d}$-2 symmetric $[Fe_4S_4(SH)_4]^{3-}$ (doublet, left), $[Fe_4S_4(SH)_4]^{2-}$ (open-shell singlet, middle) and $[Fe_4S_4(SH)_4]^{1-}$ (doublet, right) clusters, $\Delta E_{\text{el}} = E_{\text{el}}^{(\text{B})} - E_{\text{el}}^{(\text{A})}$. The hydrogen atoms of the ligands were kept fixed and define distances a and b. The black line separates the regimes where the favored coupling scheme is different, i.e., at the black line $\Delta E_{\text{el}} = 0$.

subcluster leads to a favorization of ferromagnetic coupling within this subcluster.

The line which indicates the change in the coupling scheme favored by the ground state, i.e., $\Delta E_{\text{el}} = 0$, shifts to smaller distances b and larger distances a upon oxidation of the cluster. Hence, the energetically favored coupling scheme depends on the charge of the cubane. Although the qualitative behavior of the favored coupling scheme is similar for the three differently charged model systems, the quantitative changes are considerable. Spin-coupling scheme A is preferred for almost all $[Fe_4S_4(SH)_4]^{1-}$ and $[Fe_4S_4(SH)_4]^{2-}$ structures, but in the $[Fe_4S_4(SH)_4]^{3-}$ cluster the regions where either coupling scheme A or B is favored are almost equally large.

It is important to examine the dependence of these results on the exchange-correlation functional. In addition to the BP86 functional we employed the hybrid functional B3LYP to study the sensitivity of our results with respect to the admixture of exact exchange (with BP86 containing no exact exchange at all). With B3LYP containing 20% exact exchange admixture, coupling scheme B is favored in $[Fe_4S_4(SH)_4]^{2-}$ for structures with shortened b and elongated a distances (see Figure 2.14, left), which is the same trend as observed with BP86. Still, the absolute

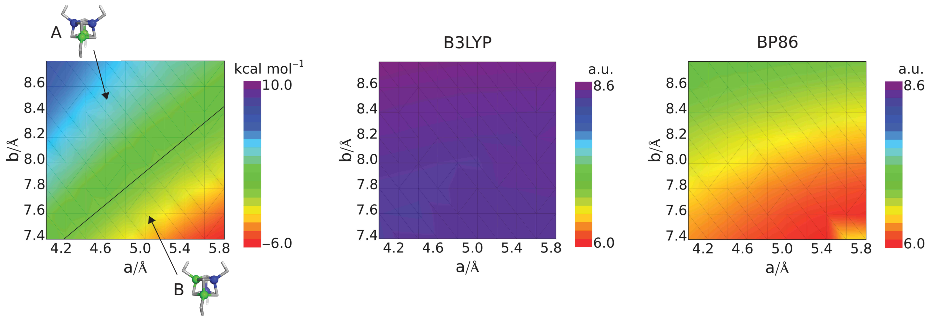


Figure 2.14: Left panel: B3LYP/def2-TZVP/COSMO plot of $\Delta E_{\mathrm{el}} = E_{\mathrm{el}}^{(\mathrm{B})} - E_{\mathrm{el}}^{(\mathrm{A})}$ of structurally constrained $[Fe_4S_4(SH)_4]^{2-}$ clusters (in singlet spin state). The hydrogen atoms of the ligands were kept fixed at distances a and b. The black line indicates $\Delta E_{\mathrm{el}} = 0$, i.e. where the coupling scheme of the lowest-energy state changes. The $\langle \hat{S}^2 \rangle$ expectation values are plotted for $[Fe_4S_4(SH)_4]^{2-}$ clusters optimized with B3LYP (middle) and BP86 (right). Element color code: gray, S; white, H; blue, Fe with negative local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value; green, Fe with positive local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value.

values differ and the line which points to the change in the coupling scheme of the ground state shifts significantly. In Figure 2.14 we also plotted the expectation values of the squared total spin operator, $\langle \hat{S}^2 \rangle$, for the grid of BP86 and B3LYP optimized clusters. As expected, the spin contamination of the broken-symmetry determinant is significant in both cases (one would expect zero for a properly spin-coupled wave function). However, such spin contamination is natural in BS calculations [19] and the important point to make is that the spin contamination is significantly larger in the B3LYP results.

Furthermore, it is important to study the effect of the fixation type in order to understand whether the chosen constraint leads to artificial changes in the cluster properties. For the S-fixed $[Fe_4S_4(SH)_4]^{2-}$ clusters a preference of coupling scheme B still exists for large a and small b, but for large distances b at small distances a, ΔE_{el} decreases significantly and even a change in the coupling scheme favored by the lowest-energy solution can be observed in the BP86 results. When comparing $[Fe_4S_4(SH)_4]^{3-}$ and $[Fe_4S_4(SH)_4]^{2-}$ clusters, coupling scheme B is considerably favored for the reduced cluster, as already observed for the H-fixed model system (the detailed results can be found in Appendix B.3). In summary, the S-fixation affects ΔE_{el} stronger than the H-fixation and it is thus advisable to employ constraints on the hydrogen atoms instead in order not to impose too strict (potentially artificial) constraints.

2.2.3 Tolerance of Fe_4S_4 clusters against structural distortions

In this section we analyze the robustness of Fe_4S_4 cluster properties with respect to structural distortions. In Figure 2.15 (top, left) the BP86 potential energy surface of $[Fe_4S_4(SH)_4]^{2-}$ with fixed hydrogen atoms is presented. For each fixation all relevant BS solutions have been calculated to make sure that the lowest-energy solution has been found. The structure with the lowest energy has been arbitrarily set to zero and 12 kcal/mol has been chosen as a cut-off threshold. Within a range of about 0.8 Å in *a* and 0.4 Å in *b* from the global energy minimum conformation, the potential energy increases only by about 3 kcal/mol. Therefore, we conclude that the influence of the ligand sphere arrangement on the cluster stability is moderate within this range. These distortions in *a* and *b* correspond to simultaneous distortions of up to 0.1 Å in the six Fe-Fe distances, 0.01 Å in the 12 cluster-internal Fe-S distances and 0.02 Å in the four Fe-S(H) distances from the lowest-energy structure (all calculated with BP86). For comparison, if we calculate the classical turning points for the zero-point vibrational energy in an harmonic potential along one mode, which elongates (or compresses) the Fe-S or Fe-Fe distances, we obtain changes of the same order. If the Fe-S or Fe-S(H) distances are more significantly distorted due to the applied structural constraint, the potential energy rapidly increases to 12 kcal/mol and up to 20 kcal/mol for the most distorted structures. In these structures we found deviations of up to 0.2 Å in the Fe-Fe distances, 0.05 Å in the cluster-internal Fe-S distances and 0.1 Å in the Fe-S(H) distances from the lowest-energy structure. In this latter case, the FeSH and S(H)FeFe angles change by about 15° and the cluster internal SFeS angles by 5–10°.

On the right panel of Figure 2.15 the correlation diagrams of local $\langle \hat{S}_{z,\mathrm{A}} \rangle$ expectation values (exemplarily depicted for one of the iron atoms, i.e. A = Fe) are presented to demonstrate the change in spin density under the influence of structural distortions. This allows us to confirm whether always the same spin configuration has been obtained. The absolute local BP86 $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation values are in the range from 1.4 to 1.5 a.u. for

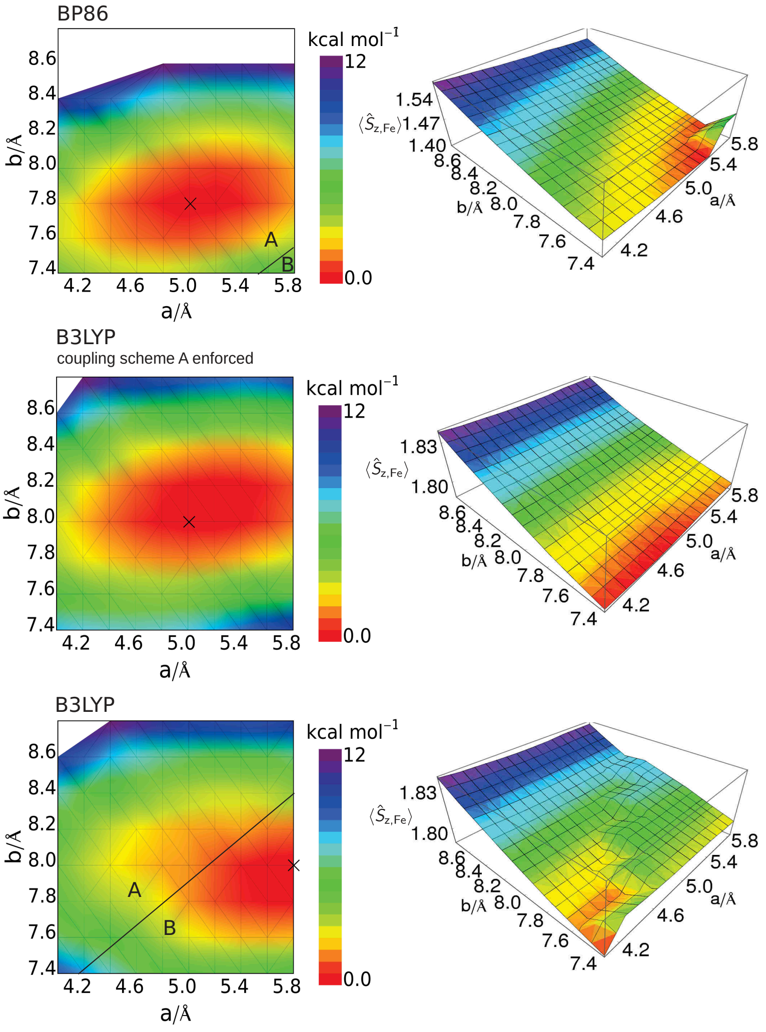


Figure 2.15: Potential energy surfaces (left) and local spin $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation values (right, exemplarily for one iron atom with excess of α-spin density) of constrained optimized $[Fe_4S_4(SH)_4]^{2-}$ clusters (in open-shell singlet spin state) with fixed hydrogen atoms. On the top, the PES of BP86/RI/def2-TZVP/COSMO optimized structures with the lowest-energy coupling scheme is plotted. In the middle, the B3LYP/def2-TZVP/COSMO results are depicted for clusters in coupling scheme A, and at the bottom the B3LYP PES for the lowest-energy coupling schemes is plotted. The crosses indicate the minimum energy structures.

the investigated region of distorted structures. Hence, the BP86 Slater determinant was, in almost all cases, converged to the same electronic structure for all distorted Fe_4S_4 clusters. The smaller the absolute local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value, the more pronounced is the $\bar{4}$-compression. The BP86 plot of local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation values shows an irregularity for two distorted structures with large a and small b distances, because the energetically favored coupling scheme changes from A to B. Noteworthy, this does not cause significant differences in the potential energy (Figure 2.15, left panel).

In order to understand to what extent the choice of a particular density functional affects these results, we are advised to also consider hybrid functionals. Because of the fact that B3LYP favors coupling scheme B for several structural distortions (see Figure 2.14), the potential energy is plotted once for all cubanes in the same coupling scheme (namely, A) and once for the lowest-energy coupling scheme (Figure 2.15 left panel, middle and bottom, respectively). The energy minimum of the B3LYP potential energy surface is shifted to slightly larger distances a and b compared to the BP86 results. The qualitative shape of the potential energy surface is very similar, if the same coupling scheme is assumed.

Interestingly, the BP86 and B3LYP PES's increase considerably for large b distances, whereas distortions in a hardly change the energy of the cluster. We analyzed the optimized cluster structures and found that the FeSH angle, the Fe-S(H) distance and the cluster-internal Fe-S and Fe-Fe(2) distances are mainly affected by a change of b, whereas the FeFeS angle and Fe-Fe(1) distances are affected by changes in both a and b.

The PES's discussed so far are constrained by a fixation of the hydrogen atoms of the HS^- ligands. By constraining the sulfur atoms instead, a more direct coupling between structural and energy changes can be achieved due to more restrictive structural adaptation possibilities. However, the qualitative changes in the PES's are very similar.

2.2.4 Reactivity changes of distorted Fe_4S_4 clusters

2.2.4.1 Starting structures for $HOO^\bullet$ coordination

Recently, we studied the reactivity of the Fe_4S_4 cluster in the H-cluster of [FeFe] hydrogenase with respect to an attack of various ROS [35, 38]. It has been found that $HOO^\bullet$ binds to the Fe-S cubane subcluster by about −19 kcal/mol and initiates degradation processes of the H-cluster. Triplet oxygen and superoxide anion do not bind to the $[Fe_4S_4]^{2+}$ cubane [35]. Hydrogen peroxide also damages the H-cluster, but it attacks the sulfur atoms of the cysteine ligands.

To investigate the effect of structural distortions on the reactivity of Fe_4S_4 clusters, we study exemplarily the coordination of the $HOO^\bullet$ radical. For this purpose, the PES's of the cluster before and after the attack of an $HOO^\bullet$ radical has been calculated and the reaction energy is obtained by the difference of the electronic energies of the products and educts.

The $HOO^\bullet$ radical can in principle attack the cluster through four different reaction channels, along one of the three Fe_2S_2 subsites or along the SH^- ligand attached to the iron atom as shown in Figure 2.16. The first two attack sites are identical by symmetry and thus yield the same energy. The third possibility is sterically unfavored and therefore only starting structures 1 and 4 of Figure 2.16 have to be considered for structure optimization. We focus on starting structure 1 in the following, as we obtained slightly lower energies for the optimized structure than for product 4 after structure optimization.

2.2.4.2 Spin coupling schemes in $[Fe_4S_4(SH)_4HOO^\bullet]^{2-}$

The coupling between the attacked iron atom and the oxygen atom of the ROS has to be taken into account. Hence, the number of possible coupling schemes increases to six. In our calculations we found that the ferromagnetic coupling between iron and oxygen is favored over the antiferromagnetic one, but this strongly depends on the density functional chosen. Therefore, the number of coupling schemes is reduced back to three (see Figure 2.17). Two of these three coupling schemes are preferred. In the high-energy spin solution the compressed Fe-S side is the one opposite to the Fe-O bond (about 180° angle). The binding of $HOO^\bullet$ to the iron atom elongates the Fe-S bond opposite to the Fe-O bond and therefore this spin state is unfavored. Depending on the spin-coupling

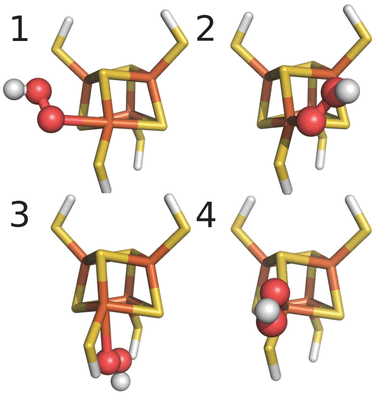


Figure 2.16: Graphical representation of the four different starting structures for the optimization of the reaction product of HOO• with the D_{2d} symmetric cluster with fixed hydrogen atoms at $a = 4.25$ Å and $b = 7.37$ Å. Element color code: red, O; yellow, S; orange, Fe; white, H.

scheme, we observe potential energy differences of about 8 kcal/mol for BP86 (see Figure 2.17).

Furthermore we observe that HOO• forms H-bridges to sulfur atoms bound to the attacked iron atom as well as to sulfur atoms of the ligands and therefore affects the stability of the cluster (see Figure 2.17). The formation of hydrogen bridges might stabilize the cluster in certain ligand arrangements. In order to assess their magnitude, we applied the shared electron number (SEN) approach [113, 114], which allows to systematically evaluate the existence and strength of hydrogen bonds. Exemplarily we calculated the hydrogen bond energies in $[Fe_4S_4(SH)_4(HO_2)]^{2-}$ fixed at $a = 5.05$ Å and $b = 7.77$ Å with def1-TZVP basis set, for which the SEN parameter had been parametrized [113]. We obtained two strong hydrogen bonds of 17.4 kcal/mol and 12.2 kcal/mol between the hydrogen atom of the HOO• and the two closest sulfur atoms in the cubane. These energies are large for hydrogen bonds to sulfur atoms, but might be due to the double negative charge of the ligated cluster, which could lead to too large SEN's and thus to artificially too large hydrogen bond energies. In the more stretched clusters slightly weaker H-bonds are found than in the global energy-minimum cluster. In order to extract a second estimate for the intramolecular hydrogen bond energies, we rotated the ligated HOO• radical to obtain a structure in which the distance between the hydrogen atom and the sulfur atoms is maximized. This structure is 7.3

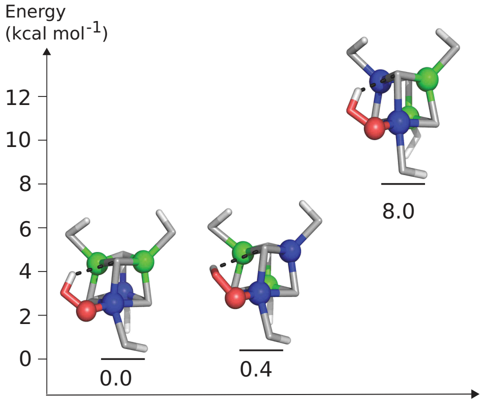


Figure 2.17: Comparison of the BP86/RI/def2-TZVP/COSMO electronic energies for optimized $[Fe_4S_4(SH)_4]^{2-}$ clusters attacked by $HOO^\bullet$ in three different coupling schemes (doublet spin state). Exemplarily the relative energies of a structure with a fixation of hydrogen atoms at $a = 4.25$ Å and $b = 7.37$ Å are depicted. Element color code: red, O ; gray, S; white, H; blue, Fe with negative local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value; green, Fe with positive local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value. The oxygen of the $HOO^\bullet$ radical is ferromagnetically coupled with the iron atom (negative local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value). The dashed lines indicate hydrogen bonds.

kcal/mol higher in energy than the global minimum structure. Hence, the hydrogen bond energy can be expected to be on this order and is thus remarkably large, as already indicated by the SEN approach.

2.2.4.3 Energies for the reaction of $[Fe_4S_4(SH)_4]^{2-}$ with $HOO^\bullet$

The BP86 $HOO^\bullet$ coordination energies of H-fixed $[Fe_4S_4(SH)_4]^{2-}$ clusters are presented in Figure 2.18. Within a distortion range of ± 0.5 Å in *a* and *b* starting from the structure with lowest binding energy, the coordination energies increase by about 3–5 kcal/mol. The binding energy depends on the stability of the cluster itself. Hence, in stretched and compressed clusters the coordination energy is higher than in clusters with a ligand arrangement close to the minimal energy structure. Still, the differences in potential energy of the investigated structures (up to 20 kcal/mol for the most distorted ones compared to the global minimum) are significantly larger than the changes in coordination energy.

Hence, structural deviations in the chosen range do not lead to significant changes in the binding energies so that the coordination of $HOO^\bullet$ would eventually become endothermic. Within the natural range of lig-

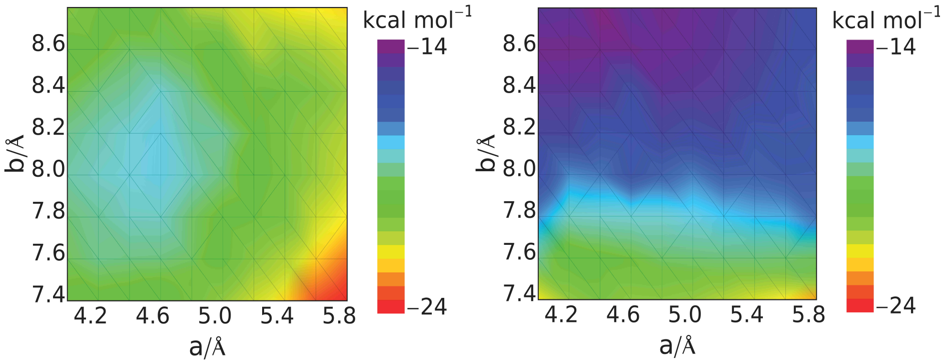


Figure 2.18: Coordination energy correlation diagram of $[Fe_4S_4(SH)_4]^{2-}$ clusters with $HOO^\bullet$ (doublet state) calculated with BP86/RI/def2-TZVP/COSMO (left) and B3LYP/def2-TZVP/COSMO (right). The hydrogen atoms of the ligands were kept fixed at distances *a* and *b* during structure optimization.

and distortions (see Section 2.2.1.2), the coordination energy changes are small. From this result, it is clear that small distortions of the ligand sphere itself cannot make the cluster tolerant against $HOO^\bullet$. The same trend can be observed for the S-fixed anchoring. If B3LYP is employed (Figure 2.18, right), the ligand surrounding, for which the cluster shows a minimal $HOO^\bullet$ coordination energy, is shifted to larger distances *b* and shorter distances *a* compared to the BP86 results.

In order to understand why some distorted structures are more reactive than others with respect to ROS, it is necessary to investigate the contribution of the educts and products on the reaction energy in more detail [38].

2.2.4.4 Rotational distortion along an S_4 axis

Besides symmetry retaining distortions, like the stretching of the Fe-S clusters in *a* and *b* direction, we distort the cluster such that the cluster symmetry is reduced, for example by a rotation of the two Fe_2S_2 subclusters along the main rotation axis of the D_{2d}-symmetric starting structure. Here, the spin state pattern is of special interest. We expect the two Fe_2S_2 subclusters to rotate back to adopt the preferred arrangement found in the BS spin state pattern of the structure with no rotational distortion. Due to the (twisted) fixation of the H atoms, which imposes the rotational distortion on the cluster, there will be strain created in the system leading

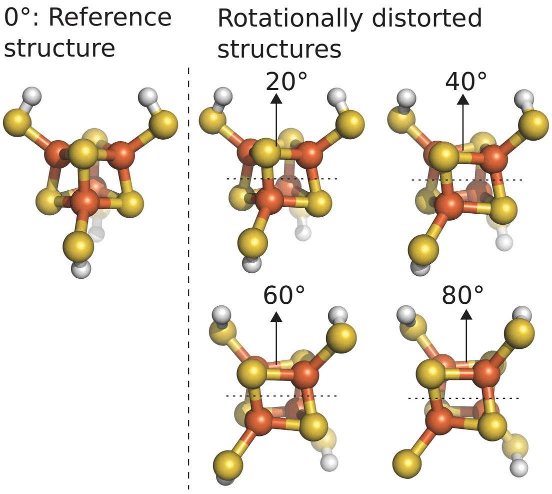


Figure 2.19: BP86/RI/def2-TZVP/COSMO optimized structures for rotationally distorted $[Fe_4S_4(SH)_4]^{2-}$ clusters (in open-shell singlet spin state). The pairs of hydrogen atoms were rotated, but kept fixed at distances $a = 5.05$ Å and $b = 7.77$ Å which corresponds to the energetically favored structure. Element color code: yellow, S; orange, Fe; white, H.

to a higher energy and the reactivity of the cluster should be increased. The BP86 optimized structures for the rotationally distorted clusters are shown in Figure 2.19.

In Table 2.8 the BP86 and B3LYP coordination energies of $[Fe_4S_4(SH)_4]^{2-}$ clusters with $HOO^\bullet$ are shown for various angles of distortions compared to the reference cluster (see Figure 2.19, 0°). We figure out that, due to the rotation, the $HOO^\bullet$ coordination energy on the $[Fe_4S_4(SH)_4]^{2-}$ clusters increases. The increase in potential energy is larger with BP86 than with B3LYP. The largest BP86 coordination energy is 7 kcal/mol higher than the coordination energy of the reference structure. For comparison, the B3LYP coordination energy increases by only 0.3 kcal/mol upon rotation by 40° (compared to 2.1 kcal/mol with BP86). The attack of one $HOO^\bullet$ radical on structures rotated by 60° and 80° does not lead to reasonable coordination products in B3LYP structure optimizations, which is the reason for the two missing data points in Table 2.8.

Table 2.8: Reaction energies (Rct.) in kcal/mol for $[Fe_4S_4(SH)_4]^{2-}$ clusters distorted by a rotation along the main rotation axis with $HOO^\bullet$ and potential energy differences (Pot.) in kcal/mol of $[Fe_4S_4(SH)_4]^{2-}$ clusters with respect to the reference structure (0°). The hydrogen atoms of the ligands were rotated against each other, but kept fixed during BP86/RI/def2-TZVP/COSMO structure optimizations.

	BP86		B3LYP	
	Rct.	Pot.	Rct.	Pot.
0°	−19.6	0.0	−17.4	0.0
20°	−20.2	0.7	−17.7	0.1
40°	−21.7	2.7	−17.7	0.8
60°	−23.9	8.1	—	3.4
80°	−26.9	18.3	—	10.4

2.2.5 Effect of structural distortions on electron transfer properties

2.2.5.1 Potential energy surfaces of $[Fe_4S_4(SH)_4]^{3-/1-}$ clusters

The most important property of iron-sulfur cubanes is their ability to serve in electron transfer processes. Several groups have studied the reduction potentials of protein-embedded iron-sulfur clusters [64, 115–118]. In contrast to these studies, we focus on differences in the PES's of $[Fe_4S_4(SH)_4]^n$ clusters (with n = 1−, 2− and 3−) caused by symmetric distortions as elaborated in the previous sections. The BP86 PES's of cubanes of total charges n = 1− and n = 3− with fixed hydrogen atoms are shown in Figure 2.20. The BP86 and B3LYP PES's of $[Fe_4S_4(SH)_4]^{2-}$ have already been presented in Figure 2.15. The shape of the PES's of the three charged states is similar, but the minimum in the PES of $[Fe_4S_4(SH)_4]^{3-}$ is shifted to larger distances a compared to the PES's of $[Fe_4S_4(SH)_4]^{1-}$ and $[Fe_4S_4(SH)_4]^{2-}$.

The oxidation or reduction of the $[Fe_4S_4(SH)_4]^{2-}$ cluster is thus accompanied by changes in the cluster structure. Sigfridsson et al. [75] observed in B3LYP calculations of methyl thiolate ligated cubanes that upon oxidation from charged state 3− to 2− the Fe-S(CH_3) bond lengths contract. If the $[Fe_4S_4(SCH_3)_4]^{2-}$ cluster is oxidized, the Fe-S(CH_3) bond lengths decrease further and two of the Fe-Fe distances increase significantly (in

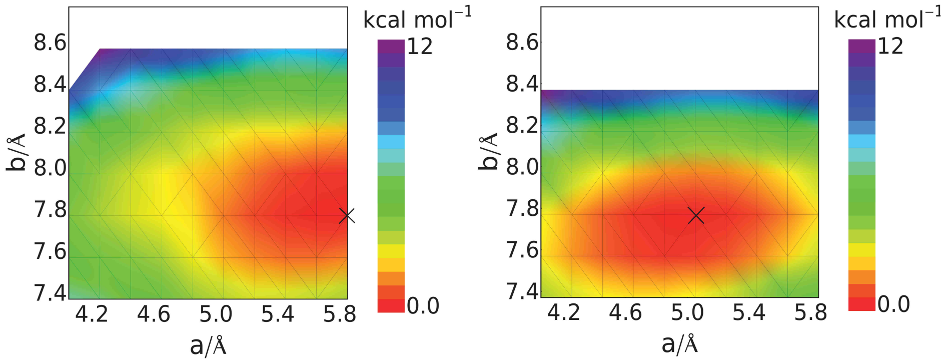


Figure 2.20: BP86/RI/def2-TZVP/COSMO potential energy surfaces of structurally constrained, originally D_{2d}-2 symmetric $[Fe_4S_4(SH)_4]^{3-}$ (left) and $[Fe_4S_4(SH)_4]^{1-}$ clusters (right), both in doublet spin state. The hydrogen atoms of the ligands were kept fixed at distances a and b during structure optimization.

our BP86 calculations one Fe-Fe distance increases and one slightly decreases compared to all Fe-Fe distances). To compare Sigfridsson's results with our symmetrically distorted clusters, we analyzed several structural features of our model system with charges n = 3−, 2− and 1−. Interestingly, most of the distances and angles are very similar for all three charged states. We have observed slight differences in the Fe-S(H) and Fe-S distances (contraction upon oxidation) and more considerable variations in one of the Fe-Fe distances. In a D_{2d}-symmetric cluster two different Fe-Fe distances exist as shown in Figure 2.21. Here, they are denoted as Fe-Fe(1) and Fe-Fe(2). We further divide the Fe-Fe(1) distances into Fe-Fe(1)_1 and Fe-Fe(1)_2 (in Figure 2.21), as they differ in the doublet spin state solutions of the 1− and 3− charged states. Although the differences in the two Fe-Fe(1) distances reduce the symmetry of the cluster, the overall structure is still approximately D_{2d}-symmetric.

Since for the distorted $[Fe_4S_4(SH)_4]^n$ clusters with charges n = 2− and n = 1− coupling scheme A is energetically favored for almost all grid points, the optimized structures are approximately D_{2d}-symmetric and the Fe-Fe(1) distances correspond to the ferromagnetically coupled iron centers and the Fe-Fe(2) distances to the antiferromagnetically coupled ones. The Fe-Fe(1) distance of the reduced ferromagnetically coupled subcluster is shortened compared to the Fe-Fe(1) distance of the oxidized ferromagnetically coupled subcluster (Figure 2.21). If coupling scheme B or C are

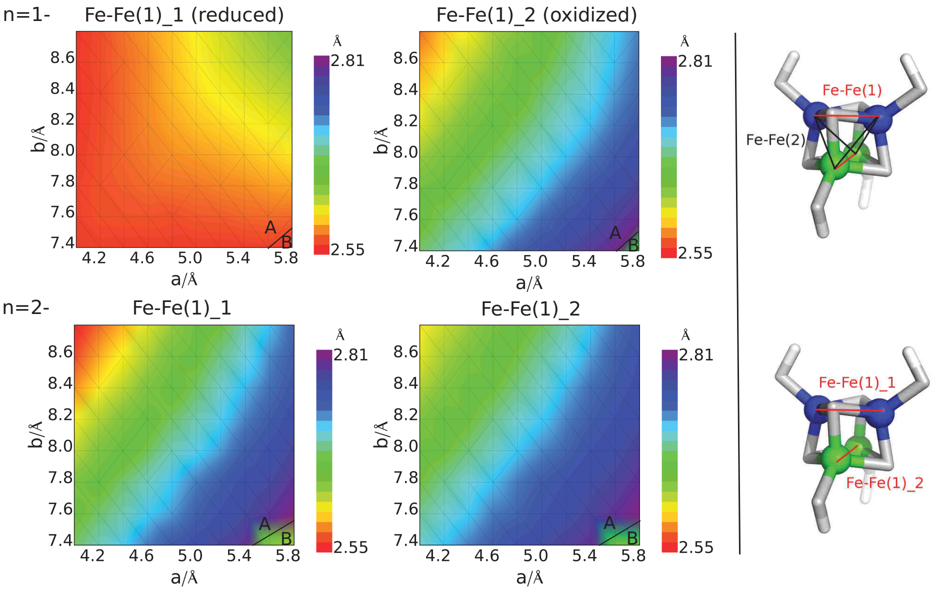


Figure 2.21: Fe-Fe(1) distances of BP86/RI/def2-TZVP/COSMO optimized $[Fe_4S_4(SH)_4]^n$ (n = 1−, 2−) clusters. The hydrogen atoms of the ligands were kept fixed according to the values a and b. The black lines indicate $\Delta E_{\mathrm{el}} = 0$, i.e. where the spin-coupling scheme of the lowest-energy state changes. On the right, the Fe-Fe distance labels in an approximately D_{2d}-symmetric cubane are shown. Element color code: gray, S; white, H; blue, Fe with negative local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value; green, Fe with positive local $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation value.

energetically favored, the Fe-Fe(1) distance also refers to the ferromagnetically coupled subclusters, whereas the Fe-Fe(2) distances denote the iron distances between antiferromagnetically coupled iron atoms, though the clusters are not approximately D_{2d}-symmetric anymore.

A fixation of the sulfur atoms instead of hydrogen atoms leads to the same conclusions, namely that the energy minimum of the PES of $[Fe_4S_4(SH)_4]^{3-}$ is shifted to slightly larger a distances (b is about the same) compared to the energy minima of the PES's of $[Fe_4S_4(SH)_4]^{2-}$ and $[Fe_4S_4(SH)_4]^{1-}$. Because of the stronger structural constraints exerted by the S-fixation, the Fe-Fe(1) and Fe-Fe(2) distances change directly with the S(H)-S(H) distance. At small distances a and b, all six Fe-Fe distances are therefore short. Since in the H-fixed model system there is a clear distinction between the Fe-Fe(1)_1, Fe-Fe(1)_2 and Fe-Fe(2) distances, the sulfur fixation appears to constrain the system too strongly. It is thus more sensible to investigate the H-fixed system.

In the crystal structures of HiPIPs (n = 2−/1− redox couple), the C_β-C_β distances of the cysteines, which are modeled here by the distance *b*, are rather large. Concerning the two distances which resemble *a*, one C_β-C_β is very short and the other one is the longest. This longest C_β-C_β distance usually corresponds to the cysteines with shortest S-S distance. We investigated which spin-coupling scheme is preferred for $[Fe_4S_4(SEt)_4]^{1-}$ clusters with ligand anchorings according to the cysteines in the crystallographic data. For this, the nitrogen and carbonyl C atom of the cysteines were changed to hydrogen atoms and fixed at a distance of 1.1 Å. Interestingly, in the energetically favored spin-coupling scheme, the reduced Fe_2S_2 subcluster is most often coordinated by the cysteines with either shortest S(Et)-S(Et) distance or shortest $C(H_2)$-$C(H_2)$ distance. This corresponds to spin-coupling scheme A, as defined in Figure 2.11. We also calculated the energy for the cluster with the same $\bar{4}$ contraction (according to A), but interchanged reduced and oxidized subclusters. In the symmetric model system, these two spin-coupling patterns are identical, because of the underlying symmetry, which is not the case in the crystal structures. The energy difference between these two spin-state solutions is very small (most often below 1 kcal/mol). Spin-coupling scheme A is preferred over B and C by about 2 kcal/mol. This favored orientation of the ferromagnetically coupled subclusters agrees with the results obtained for the $[Fe_4S_4(SH)_4]^{1-}$ clusters, where spin-coupling scheme A is favored for small distances *a* and large distances *b* (Figure 2.13).

Dey et al. [78] also investigated the orientation of redox-active, ferromagnetically coupled Fe_2S_2 subclusters in Fe-S cubanes. Depending on the Fe-Fe and Fe-S distances in highly resolved crystal structures, they could assign a certain spin-coupling pattern to the iron atoms in the Fe-S cubanes. They suggest that the orientation of these Fe_2S_2 subclusters stays the same upon oxidation or reduction. For comparison to their study, we investigated the effect of the ligand surroundings of Fe-S cubanes on the energetically favored spin-coupling scheme. The spin-coupling scheme assignment in our study leads to the same results for the HiPIP [1CKU [100]] compared to the results of Dey et al. [78] and of Bertini et al. from 2D-NMR experiments [119]. By contrast, the spin assignment in cluster 56 of 1FDN [120] obtained by Dey et al. and by the 2D-NMR experiments is different from ours. The experimental ground-state spin-coupling scheme resembles C, whereas we obtain spin-coupling scheme B with BP86 and almost the same energies for B and C with B3LYP. Since the energy difference between the two lowest energy spin-coupling schemes is significantly larger in HiPIP clusters than in the ferredoxin clusters, it is not surprising that we obtained the same spin assignment in the HiPIP cluster but not in the ferredoxin cluster. In other words, we may consider the DFT results for our ferredoxin-type clusters as a match with experiment considering the method-inherent approximations in the density functionals and the fact that we study a model cluster.

2.2.5.2 Adiabatic detachment energies

Next, we discuss the electronic energy differences between the lowest-energy structures of the reduced and oxidized states. These differences resemble the adiabatic detachment energies (ADE's) [68],

$$\mathrm{ADE} = E_{\mathrm{el}}^{\mathrm{A}}(\mathrm{A}) - E_{\mathrm{el}}^{\mathrm{A}^-}(\mathrm{A}^-) + (E_{\mathrm{vib},0}^{\mathrm{A}}(\mathrm{A}) - E_{\mathrm{vib},0}^{\mathrm{A}^-}(\mathrm{A}^-)), \qquad (2.1)$$

without vibrational zero-point energy difference.

We calculated $E_{\mathrm{el}}([\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-}) - E_{\mathrm{el}}([\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{3-})$, denoted as ADE(3−/2−), and $E_{\mathrm{el}}([\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{1-}) - E_{\mathrm{el}}([\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-})$, denoted as ADE(2−/1−). We have evaluated the zero-point vibrational energy contributions in harmonic approximation and neglecting the fact that the cluster structures should have been fully relaxed for several extremal distortions and found that they do not affect the results significantly.

The BP86 ADE(3−/2−) (Figure 2.22, left panel) vary between 11 kcal/mol and 20 kcal/mol, which means that the oxidation of $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{3-}$ is endothermic for all distorted structures. Hence, the one-electron reduction of $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-}$ is exothermic. For the ADE(1−/2−) (Figure 2.22, right panel) we find 65 kcal/mol to 71 kcal/mol. The oxidation from $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-}$ to $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{1-}$ is therefore very endothermic, and reversely, the reduction is exothermic. This result correlates well with the properties of ferredoxins, which operate between charges n = 2− and n = 3−. In HiPIPs, the oxidation of $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-}$ to $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{1-}$ should be preferred compared to the ferredoxin type clusters.

The overall ADE changes for both redox couples are in the range of 6–9 kcal/mol for the chosen structural constraints. Variations in the ADE's mainly stem from a distortion in b. To understand which structural features affect the ADE's, we compared the qualitative changes in the correlation diagrams with structural changes in the distorted cubanes. The Fe-S(H) bond length as well as the Fe-Fe(2) and Fe-S distances change in a similar way as the ADE's.

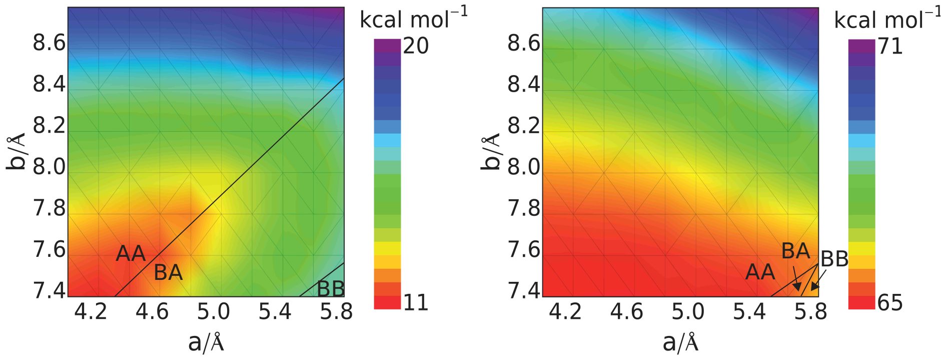


Figure 2.22: Correlation diagrams of ADE(3−/2−) (left) and ADE(2−/1−) (right). [Fe$_4$S$_4$(SH)$_4$]$^{3-/1-}$ are in doublet spin state, [Fe$_4$S$_4$(SH)$_4$]$^{2-}$ possesses a singlet spin state. The hydrogen atoms of the ligands were kept fixed at distances a and b during BP86/RI/def2-TZVP/COSMO structure optimization. The black lines indicate $\Delta E_{\mathrm{el}} = 0$, i.e. where the spin-coupling scheme of the lowest-energy state changes, for both charged states.

Niu et al. [68] found that ligand dihedral angles affect the reduction potentials of iron-sulfur cubanes by about 100 mV, because they induce changes in the Fe-S$_{\mathrm{ligand}}$ covalency. These changes in the covalency could be determined by natural bond orbital analysis [68] and by X-ray absorption spectroscopy (XAS) [61]. Although Niu et al. found that the major effect on redox properties is the electrostatic effect of hydrogen bonds (about 700 mV), the influence of the ligand conformation might be important to explain the differences between HiPIP and ferredoxin clusters.

Moreover, Dey et al. [78, 121] observed that a reduction in Fe-S$_{\mathrm{ligand}}$ covalency (measured by XAS) increases the reduction potential of Fe-S clusters approximately linearly. They investigated compressed, non-H-bonded and elongated, H-bonded Fe-S clusters and found that the H-bonding and structural distortions (elongation) of the clusters lead to a decrease in Fe-S$_{\mathrm{ligand}}$ covalency. The favored spin-coupling scheme in the compressed cluster corresponds to A as defined here and the spin-coupling scheme of the elongated clusters resembles B. Due to two shorter Fe-S bonds in one Fe$_2$S$_2$ subcluster of the elongated cluster, the Fe-S$_{\mathrm{ligand}}$ covalency decreases. These results indicate that structural changes within the Fe-S cluster might lead to differences in the reduction potentials. Symmetric distortions of our model system do not significantly change the ligand dihedral angles, which were discussed by Niu et al. [68] but the Fe-S and Fe-Fe bond lengths. We observed an increase of ADE(3−/2−) and ADE(2−/1−) by 6–9 kcal/mol for an elongation of *b* by 1.4 Å. This elongation corresponds to distortions of 0.2 Å in the Fe-Fe distances and of 0.05 Å in the Fe-S distances.

Next, we compare the results of $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ with clusters ligated by ethyl thiolates, which are fixed according to the cysteine backbone atoms in the crystal structures of selected proteins. The BP86 and B3LYP ADE(3−/2−) and ADE(2−/1−) are listed in Table 2.9.

For all ligand conformations — ferredoxin-type and HiPIP — we obtained almost the same ADE's with BP86, although we calculated potential energy differences of up to 8.8 kcal/mol for the doubly negatively charged clusters. The only variation is that the ADE(3−/2−) is slightly larger for HiPIPs than for ferredoxins. The ADE(3−/2−) are in the range of 11.9 kcal/mol to 14.9 kcal/mol and the ADE(2−/1−) change between 59.3 kcal/mol and 60.3 kcal/mol. With B3LYP the ADE's are larger than with BP86, but the qualitative result is the same, namely that ligand sphere distortions as observed in the crystallographic data do not lead to significant changes in the ADE's.

2.2.5.3 Inner-sphere reorganization energies

In order to understand the effect of adding or removing one electron on the cluster structure, Sigfridsson et al. [75] have calculated reorganization energies of iron-sulfur clusters with the B3LYP density functional. Reorganization energies λ describe the ability of a system to change between different oxidation states and can be split into inner-sphere λ_{in} and outer-sphere λ_{out} contributions. The outer-sphere reorganization energy describes structural changes of protein and solvent upon oxidation or reduction of the cluster. The inner-sphere reorganization energy λ_{in} considers changes of the cluster structure and its ligands. The latter one can be calculated by the free energy difference between the oxidized complex at its optimal geometry and the reduced complex at the optimal geometry of the oxidized complex plus the corresponding difference for the reduced complex,

$$\lambda_{\text{in}} = E_{\text{el}}^{\text{A}}(\text{A}^-) - E_{\text{el}}^{\text{A}}(\text{A}) + E_{\text{el}}^{\text{A}^-}(\text{A}) - E_{\text{el}}^{\text{A}^-}(\text{A}^-), \tag{2.2}$$

Table 2.9: ADE's (in kcal/mol) for $[Fe_4S_4(SEt)]^n$ clusters with charges n = 3−/2− and n = 2−/1−. The ethyl thiolate ligands, SEt, were kept fixed at the two hydrogen atoms corresponding to the carbonyl C atom and N atom positions of the cysteine ligands.

PDB entry	ADE(3−/2−)	ADE(2−/1−)
BP86/RI/def2-TZVP/COSMO		
Ferredoxins		
1A6L	13.0	59.5
2FDN	13.0	59.3
6FD1	13.4	59.6
2FGO	12.7	59.5
2Z8Q	12.2	60.3
2ZVS-1	14.2	59.5
2ZVS-2	13.8	59.8
HiPIPs		
1HPI	14.9	60.1
1B0Y	14.9	60.0
1ISU	14.5	60.0
Others		
1CP2	11.9	59.5
2C42	12.5	60.2
3C8Y	13.9	59.6
2E7Z	11.9	59.8
2YIV	12.8	59.6
B3LYP/def2-TZVP/COSMO		
Ferredoxins		
1A6L	24.1	68.9
2FDN	24.2	67.3
HiPIPs		
1B0Y	25.1	67.8
1ISU	23.7	67.7

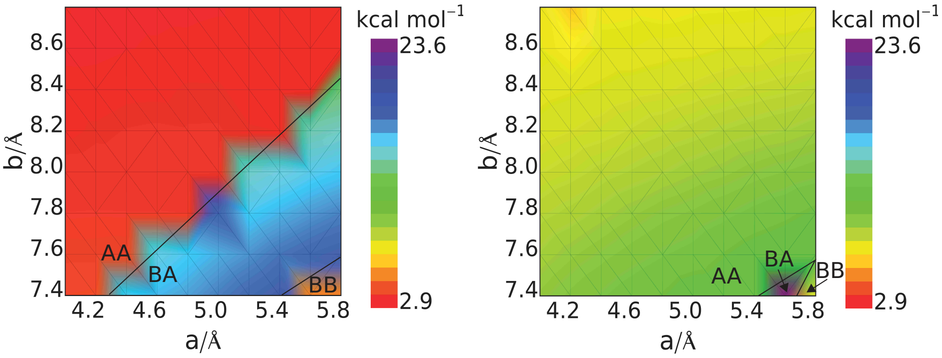


Figure 2.23: Inner-sphere reorganization energies λ_{in} of structurally constrained, originally D_{2d}-2 symmetric $[Fe_4S_4(SH)_4]^n$ clusters for the n = 3−/2− (left) and the n = 2−/1− (right) redox couples. The hydrogen atoms of the ligands were kept fixed at distances a and b during BP86/RI/def2-TZVP/COSMO structure optimization. The black lines indicate $\Delta E_{\text{el}} = 0$, i.e. where the spin-coupling scheme of the lowest-energy state changes, for both charged states. The kinks in the correlation diagram occur due to the grid size.

where e.g. $E_{\text{el}}^{\text{A}}(\text{A}^-)$ is the electronic energy of the oxidized complex (A) at the optimized structure of the reduced complex (A^-). As λ_{in} does not depend on the detailed structure of the protein environment, a minimal model system like ours is sufficient. For our model system $[Fe_4S_4(SH)_4]$, we may assume that the enthalpic and entropic contribution at finite temperature cancel, because the masses of the reduced and oxidized complexes are the same and the vibrational frequencies and moments of inertia are similar (especially in the harmonic approximation). Hence, the calculation of electronic energies E_{el} should be sufficient. Sigfridsson et al. [75] analyzed inner-sphere reorganization energies of mono-nuclear iron complexes, Fe_2S_2, and Fe_4S_4 clusters. For the latter ones these authors considered CH_3S^- as ligands and obtained inner-sphere reorganization energies of 14.8 kcal/mol for ferredoxin-type and of 10.3 kcal/mol for HiPIP-type cluster models.

To understand how symmetric distortions affect λ_{in}, we have calculated this property for the grid of distorted structures (Figure 2.23).

The reorganization energy plot for the n = 3−/2− redox couple (Figure 2.23, left panel) shows a division in cluster structures of low reorganization energies around 3 kcal/mol and cluster structures of significantly increased reorganization energies (14–18 kcal/mol). Also for the n = 2−/1−

redox couple (Figure 2.23, right panel) λ_{in} adopts small values around 6–12 kcal/mol for almost the whole correlation diagram, but a significantly larger value of about 24 kcal/mol for one structure. The reason for this result is a change in the ground-state spin-coupling scheme, since λ_{in} is much lower for clusters favoring the same spin-coupling scheme before and after electron transfer than for clusters with different spin-coupling schemes independent of the charged state. In total, λ_{in} of the HiPIP couple ($n = 2-/1-$, Figure 2.23, right panel) is larger than λ_{in} of the ferredoxin (3−/2−) couple, if the same spin-coupling scheme is assumed for both redox partners. One explanation could be the more significant difference in the cluster structure of $[Fe_4S_4(SH)_4]^{1-}$ in comparison to $[Fe_4S_4(SH)_4]^{3-}$ and $[Fe_4S_4(SH)_4]^{2-}$, since one Fe-Fe distance of the reduced, ferromagnetically coupled Fe_2S_2 subcluster is decreased compared to the Fe-Fe distances in the ferromagnetically coupled Fe_2S_2 subclusters of $[Fe_4S_4(SH)_4]^{2-}$.

Noodleman et al. [122] compared the experimental total reorganization energies obtained by Kümmerle et al. [123] for *Chromatium vinosum* ferredoxin and by Babini et al. [124] for HiPIPs with the theoretical results by Sigfridsson et al. [75]. In the experiments, higher total reorganization energies have been observed for the HiPIPs (13.8–20.8 kcal/mol) than for the ferredoxin cluster (4.6–11.5 kcal/mol), which is contradicting the trend in inner-sphere reorganization energies observed by Sigfridsson et al. The experimental results qualitatively agree with our BP86 results, if we assume that the ground-state spin-coupling scheme of the redox partners does not change upon oxidation or reduction. Therefore, the discrepancies might be explained by the dependence of the calculated inner-sphere reorganization energies on the spin-coupling scheme. On the other hand, in experimentally obtained reorganization energies the outer-sphere contribution is included, which we did not study here. However, our results indicate that it is important to understand the relationship between a structure and the energetically favored spin-coupling scheme in order to reliably calculate inner-sphere reorganization energies, because of the dependence of λ_{in} on the chosen spin-coupling scheme.

To investigate whether similar results can be obtained for Fe-S cubanes from crystal structures, we calculated BP86 and B3LYP inner-sphere reorganization energies for several ferredoxin-type and HiPIP-type clusters as well as Fe-S clusters occurring in other proteins (Table 2.10). Hereby, we assume, that the backbone arrangement of the cysteine ligands does not change significantly upon oxidation or reduction. We exchanged the nitrogen and carbonyl carbon atom of the cysteines obtained from the crystal structure of a certain charged state by hydrogen atoms and clamped them. For the structure optimizations of the reduced and oxidized complexes we chose the same fixation. Since the oxidation or reduction of the Fe-S cubane might lead to changes in the protein backbone arrangement, we also calculated λ_{in} after we released the structural constraint (H-fixation) for selected $[Fe_4S_4(SEt)]^n$ clusters (Table 2.10, squared brackets). We found that the differences between these two inner-sphere reorganization energies are small.

Following this procedure, we obtained for example for 1A6L (ferredoxin) $\lambda_{\text{in}} = 2.9$ kcal/mol for the n = 3−/2− couple (same spin-coupling scheme) and $\lambda_{\text{in}} = 9.6$ kcal/mol for the n = 2−/1− couple (same spin-coupling scheme). These results agree with the λ_{in}'s obtained for the HS^- ligated model system. For 1ISU, a HiPIP, the inner-sphere reorganization energy of the n = 3−/2− couple is also very small (2.7 kcal/mol) as observed for the ferredoxin clusters, if the same coupling scheme is enforced. The inner-sphere reorganization energy of the n = 2−/1− couple amounts to 9.0 kcal/mol (same spin-coupling scheme). Note that the HiPIP cluster actually operates only between n = 2−/1− and the ferredoxin clusters between n = 3−/2−. Nevertheless, we calculated λ_{in} for both redox couples in order to be able to compare the ferredoxin-type and HiPIP-type clusters with each other. A comparison of the inner-sphere reorganization energies of ferredoxin-type structures with HiPIPs indicates, that the observed structural differences between ferredoxin and HiPIP clusters do not significantly change the reorganization energies.

If we compare the reorganization energies of several more ferredoxins, HiPIPs and other Fe-S cubanes, we can conclude that the reorganization energy of the 2−/1− couple is always around 9–10 kcal/mol, if the same spin-coupling scheme is assumed and that λ_{in} of the 3−/2− couple is either around 2–3 kcal/mol (same spin-coupling scheme) or 13–17 kcal/mol (different spin-coupling schemes). This means that for HiPIPs and ferredoxins the $[Fe_4S_4(SEt)_4]^{3-}$ and $[Fe_4S_4(SEt)_4]^{2-}$ cluster structures in the same spin-coupling scheme are more similar than the $[Fe_4S_4(SEt)_4]^{2-}$ and $[Fe_4S_4(SEt)_4]^{1-}$ cluster structures. Therefore, the reduction of $[Fe_4S_4(SEt)_4]^{2-}$ should be kinetically preferred over the oxidation.

Table 2.10: Energetically preferred spin-coupling schemes (SCS) and inner-sphere reorganization energies λ_{in} (in kcal/mol) of $[Fe_4S_4(SEt)]^n$ clusters of charges n = 3−/2− and n = 2−/1−. The ethyl thiolate ligands were kept fixed at the two hydrogen atoms corresponding to the carbonyl C atom and N atom positions of the cysteine ligands in the crystal structures. Constrained structure optimizations have been performed with BP86/RI or B3LYP. ΔE_{el} is the electronic energy difference between the two energetically lowest spin-coupling schemes (A/B/C) in kcal/mol. For some clusters λ_{in} was computed between structures optimized without structural constraints starting from the converged broken-symmetry solutions. These energies are given in squared brackets.

PDB entry	SCS (ΔE_{el})			λ_{in} (3−/2−)	λ_{in} (2−/1−)
	3−	2−	1−		
BP86/def2-TZVP/COSMO					
Ferredoxins					
1A6L	B (0.9)	B (0.1)	B (0.2)	2.9 [6.5]	9.6 [9.8]
2FDN	B (1.9)	B (0.2)	B (0.3)	2.6	9.5
6FD1	A (0.7)	A (0.3)	A (0.6)	3.3	10.0
2FGO	C (0.5)	A (0.7)	A (1.5)	16.8	10.2
2Z8Q	C (0.8)	B (0.8)	A (0.7)	16.6 [19.8]	20.4 [27.2]
2ZVS-1	C (0.8)	A (0.4)	A (0.2)	14.1	10.4
2ZVS-2	C (1.6)	B (0.5)	B (0.5)	13.8	9.0
HiPIPs					
2AMS	A (1.7)	A (2.2)	A (2.2)	2.6 [3.8]	9.2 [9.6]
1B0Y	A (4.0)	A (1.9)	A (2.1)	2.7	9.1
1HPI	A (1.5)	A (2.3)	A (2.3)	2.7	9.0
3H31	A (1.9)	A (1.5)	A (1.5)	2.6	8.9
1ISU	A (1.6)	A (2.0)	A (1.5)	2.7	9.0
Others					
1CP2	A (0.6)	A (1.9)	A (3.3)	3.5	9.6
2C42	A (0.6)	A (1.7)	A (2.6)	3.6	9.0
3C8Y	B (0.7)	A (1.9)	A (2.1)	14.3	9.7
2E7Z	B (0.1)	A (2.1)	A (2.8)	13.8	8.6
2YIV	B (1.0)	A (2.0)	A (2.8)	13.8	9.0
B3LYP/def2-TZVP/COSMO					
Ferredoxins					
1A6L	C (0.2)	C (0.7)	A (0.8)	8.4 [10.1]	10.2 [13.3]
2FDN	C/A (0.0)	C/B (0.0)	A/C (0.0)	11.2	10.6
HiPIPs					
1B0Y	A (1.7)	A (2.0)	A (2.7)	9.4 [9.8]	8.3 [8.7]
1ISU	A (1.1)	A (2.9)	A (4.1)	9.0	8.5

With B3LYP, the λ_{in}'s of the HiPIP and ferredoxin clusters are very similar (8-11 kcal/mol). For some crystal structures λ_{in} of the HiPIP is even slightly smaller than λ_{in} of ferredoxins-type clusters. Moreover, the difference between inner-sphere reorganization energies of Fe-S clusters favoring different or the same spin-coupling scheme is much smaller than with BP86 (about 3 kcal/mol instead of 10–15 kcal/mol) as can be seen in Table 2.10. The reason for this observation is that the optimized Fe-S and Fe-Fe distances depend strongly on the chosen density functional. The $\bar{4}$ contraction of the antiferromagnetically coupled Fe_2S_2 subclusters is significantly less pronounced with B3LYP than with BP86. In the $[Fe_4S_4(SEt)_4]^{2-}$ cluster the compressed Fe-S bonds are only by about 0.05 Å smaller than the elongated Fe-S bonds, instead of 0.1 Å as calculated with BP86. Noteworthy, the $\langle \hat{S}^2 \rangle$ expectation value is larger for density functional B3LYP than for BP86 (cf. Figure 2.14).

To visualize the dependence of λ_{in} on the favored spin-coupling scheme and on the density functional chosen, BP86 and B3LYP potential energy curves are depicted for several redox couples of Fe-S clusters from ferredoxins and HiPIPs in Figure 2.24. If we compare the BP86 potential energies of ferredoxin clusters with the same and different ground-state spin-coupling schemes, we see that the potential energy increases significantly, if the spin-coupling scheme changes from one charged state to the other. The BP86 and B3LYP results for a HiPIP cluster are very similar for clusters with the same spin-coupling scheme. If a change of the spin-coupling scheme is enforced upon electron transfer, the BP86 λ_{in} is significantly larger than the B3LYP result. As can be seen in Figure 2.24, the slope of the potential energy curve is larger with BP86 than with B3LYP. This shall indicate, that the structural variations between $[Fe_4S_4(SEt)_4]^{2-}$ and $[Fe_4S_4(SEt)_4]^{1-}$ clusters, which are optimized in different spin-coupling schemes, are smaller for B3LYP than for BP86.

It should be noted, that the energy differences between the spin-coupling schemes are small, often below 1 kcal/mol (Table 2.10). However, it is remarkable that for almost all structures investigated the same spin-coupling scheme is energetically favored for $[Fe_4S_4(SEt)_4]^{1-}$ and $[Fe_4S_4(SEt)_4]^{2-}$ clusters, whereas for $[Fe_4S_4(SEt)_4]^{3-}$ and $[Fe_4S_4(SEt)_4]^{2-}$ clusters, the ground-state spin-coupling schemes most often differ. However, because of the small energy differences between the spin-coupling schemes, it is a reasonable recommendation to always compare clusters of the same spin-coupling scheme for the calculation of λ_{in}.

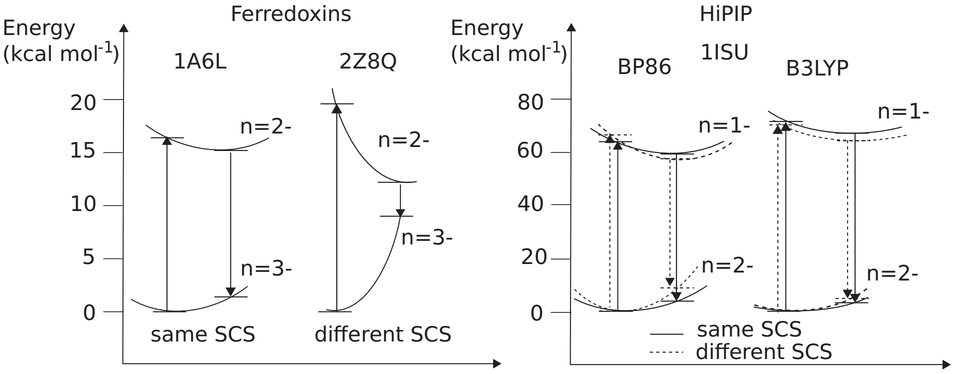


Figure 2.24: Graphical representation of the potential energy curves of two ferredoxin and one HiPIP derived $[Fe_4S_4(SEt)_4]^n$ cluster for the n = 3−/2− and the n = 2−/1− redox couples with the same and different spin-coupling schemes. The potential energy curves have been sketched on the basis of data presented in Table 2.10. The lowest potential energies of every redox couple were set to zero. SCS means spin-coupling scheme.

2.2.6 Consequence for the chemistry of Fe_4S_4 cubanes

We have studied correlation diagrams in order to systematically uncover dependencies of cluster properties on cluster structures. The range of structures covered in these diagrams has been chosen such that the angles and distances match the ones measured in crystal structures of Fe_4S_4-containing proteins. As a model system we have chosen quasi-D_{2d}-symmetric $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ cubanes with constrained hydrogen atoms to mimic the anchoring of the cluster in the protein environment. Differences in adiabatic detachment energy, inner-sphere reorganization energy and reactivity of Fe-S cubanes can be related to specific structural distortions determined by the hydrogen atom positions.

We first investigated the topology and role of the broken-symmetry spin-coupling scheme of the $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ model cubanes. A clear structural division in eight elongated Fe-S distances within the two ferromagnetically coupled Fe_2S_2 subclusters and four compressed Fe-S distances between these Fe_2S_2 subclusters was observed with BP86. With B3LYP these differences in the Fe-S distances are less pronounced. Still, the spin-coupling scheme adopted in the ground state can be identified by inspection of the Fe-S distances.

The energetically favored spin-coupling scheme depends on the anchoring and the charge of the clusters. Spin-coupling scheme A, which almost retains D_{2d} symmetry, is most often lowest in energy for clusters with −2 and −1 elementary charges. The cysteine ligand arrangement found in HiPIP crystal structures also leads to a preference of spin-coupling scheme A over B and C (by about 2 kcal/mol) for elementary charges −1 and −2. For the triply negatively charged clusters spin-coupling schemes A and B occur equally likely within the chosen range of structural distortions. For all three charges, the preference of spin-coupling scheme A over B can be increased, if the ligands of the two ferromagnetically coupled Fe_2S_2 subclusters are narrowed and the others are departed (see Figure 2.13).

BP86 and B3LYP potential energy surfaces of $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ cubanes show that the ligated cubane structures are quite flexible (Figure 2.15). Within the applied distortions, there is a large range of structures with potential energies increased by less than 3 kcal/mol compared to the global minimum structure (± 0.5 Å in H-H distances, which is equivalent to 0.1 Å in Fe-Fe distances). If the same spin-coupling scheme is enforced for all structures, the BP86 and B3LYP potential energy surfaces are qualitatively very similar. If the lowest energy spin-coupling scheme is assumed, the minima of the PES's differ slightly. Interestingly, the potential energy surfaces of $[Fe_4S_4(SH)_4]^{2-}$ and $[Fe_4S_4(SH)_4]^{3-}$ are more similar than the one of $[Fe_4S_4(SH)_4]^{1-}$, for which a more compressed anchoring is favored (Figure 2.20). Furthermore, we observed a contraction of the Fe-Fe distance in the reduced Fe_2S_2 subcluster of $[Fe_4S_4(SH)_4]^{1-}$.

Since Fe-S cubanes can be attacked by $HOO^\bullet$ radicals [35, 38], we also studied the effect of structural distortions on the coordination energies of one $HOO^\bullet$ radical to $[Fe_4S_4(SH)_4]^{2-}$ as an example for cubane ligation. We observe that symmetric distortions of Fe_4S_4 clusters in D_{2d} symmetry do not lead to significant changes in the coordination energies (Figure 2.18). The differences in the coordination energies are small in the chosen range of structural distortions (about 6 kcal/mol), though, in addition, the potential energy changes by up to 20 kcal/mol. Hence, small structural distortions of the ligand sphere imposed, for instance, by the protein backbone, are not likely to produce an anchored cluster tolerant against $HOO^\bullet$.

Furthermore, we studied the electronic energy differences between differently charged states of the clusters. The electronic energy differences of $[Fe_4S_4(SH)_4]^n$ clusters for redox couples n = 2−/3− and n = 1−/2− change by 6–9 kcal/mol upon the chosen structural distortion (Figure 2.22). For the $[Fe_4S_4(SEt)_4]^n$ clusters with fixations of the ligands according to the crystal structures these electronic energy differences do not change significantly.

Inner-sphere reorganization energies have been calculated to find out how large the structural changes of the ligated clusters are upon oxidation or reduction. The BP86 inner-sphere reorganization energy is low, if the same spin-coupling scheme is enforced for both oxidation states. If the ground-state spin-coupling schemes change upon oxidation or reduction, the reorganization energy becomes significantly larger (Figure 2.23). With B3LYP the dependence of the inner-sphere reorganization energy on the enforced spin-coupling scheme is less pronounced. The ground-state spin-coupling scheme strongly determines the cluster structure. Therefore, changes in the spin-coupling scheme of the oxidized and reduced clusters considerably increase the reorganization energy. Moreover, the optimized Fe-S and Fe-Fe distances depend on the employed density-functional, which explains the observed differences in the BP83 and B3LYP inner-sphere reorganization energies.

In total, lower BP86 reorganization energies have been obtained for the ferredoxin couple than for the HiPIP redox couple, if the same spin-coupling scheme is enforced. With B3LYP the ferredoxin and HiPIP cluster reorganization energies are almost the same. Sigfridsson et al. obtained lower inner-sphere reorganization energies for HiPIPs than for ferredoxins in B3LYP calculations. In experiments the opposite has been observed by Kümmerle [123] and Babini [124]. The explanation for the discrepancies might be the differences in the spin-coupling schemes of the redox partners.

We conclude that the ground-state spin-coupling scheme is significantly determined by deviations from the ideal cluster structure as enforced by the anchoring within a protein. Therefore, a careful analysis of the BS spin-coupling schemes is necessary for any theoretical investigation of cluster properties, but especially for the inner-sphere reorganization energies, which may depend strongly on the chosen spin-coupling scheme. This observation is particularly critical considering that orbital optimization in self-consistent field procedures may converge to local energy minima and hence finding the lowest-energy coupling scheme is not guaranteed.

3

Case Studies 2: Transition-State Optimizations

3.1 Conformer analysis of a double-decker-type rotamer

The facile synthesis of a rotor-shaped compound with two stacked triazine units, which are symmetrically connected by three anthracene blades through oxygen linkers (double-decker 1), has been accomplished by our experimental coworkers (M. Kory, A. D. Schlüter, D-MATL, ETHZ)) [125, 126]. The content of this Chapter has been published in Ref. [125].

This computational analysis is performed to predict the conformational stability of double-decker 1 (see Figure 3.1). The intrinsic bent shape of its oxygen bond can, in principle, lead to various conformers. We focus on three (as shown in Figure 3.1): the C_{3h} symmetric rotor-shaped form (**1A**), a non-C_3 symmetric form (**1B**), which is accessible from **1A** through a flipping motion of one of the three blades, hereby twisting the two overlaying oxygen linkers to the opposite side, and another C_3 symmetric screw-shaped form (**1C**), in which the three oxygen bonds connecting one triazine core face in opposite direction compared to the other ones.

To investigate the stability of these conformers, as well as potential rearrangement pathways, density functional theory (DFT) calculations are performed. A detailed Computational Methodology can be found in Appendix C. First structures **1A**, **B**, and **C** are optimized. It is important to note that standard DFT calculations usually completely neglect attractive

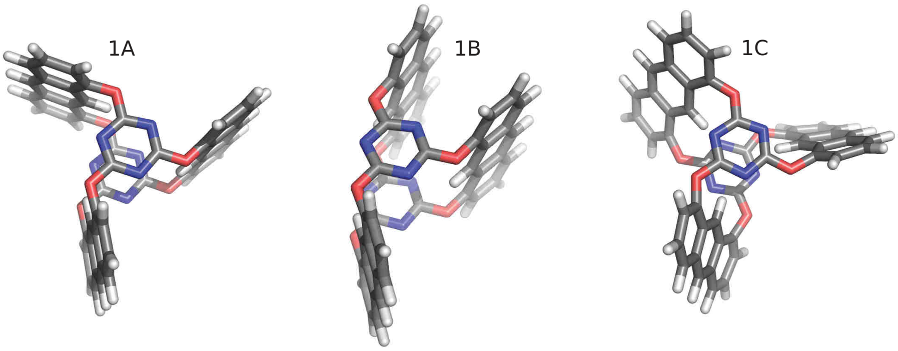


Figure 3.1: BP86/RI/def2-TZVP optimized conformations of compound 1.

long-range dispersion effects. They can be efficiently included through a dispersion correction for atom pairs at sufficiently large distance. Because of the interactions between the three anthracene moieties, we expect that dispersion interactions play an important role and thus employ Grimme's dispersion corrections D and D3 [127, 128].

3.1.1 Relative stability of isolated conformers

The electronic energy (ΔE_{el}) and free energy (ΔG) differences between conformers **1A**, **1B** and **1C** are reported in Table 3.1. We observe that **1B** is energetically favored over **1A** by about −3 to −4 kcal/mol, if we take dispersion interaction energies into account. By contrast, without dispersion interactions, **1B** is disfavored over **1A** by around 1 kcal/mol. Attractive dispersion interactions also affect the molecular structures. Thus, in **1B**, the dispersion-affected face-to-face-stacking blades are by 1.7 Å closer together than in the structure for which dispersion is not taken into account in the structure optimization. Comparing dispersion corrected, as well as dispersion-free results, structures **1A** and **B** are more similar in energy (ΔE_{el}(**1A**→**1B**) ranges from −10 to 1 kcal/mol) than structures **1A** and **C** (ΔE_{el}(**1A**→**1C**) ranges from 14 to 19 kcal/mol). Because structure **1C** has a significantly higher electronic energy than structure **1A**, it is not likely that **1C** will be formed from **1A**. The free energy differences are qualitatively similar to the electronic energy differences.

Table 3.1: Electronic energy (ΔE_{el}) and free energy (ΔG) differences (in kcal/mol) between the optimized structures **1A**, **1B** and **1C**. For the free energy calculations T = 298.15 K and p = 0.1 MPa has been chosen. The double slash denotes that a single-point calculation has been performed on a structure which was optimized with the density functional and basis set defined after the double slash.

density functional	$\Delta E_{\mathrm{el}}/\Delta G(\mathbf{1A}\rightarrow\mathbf{1B})$	$\Delta E_{\mathrm{el}}/\Delta G(\mathbf{1A}\rightarrow\mathbf{1C})$
BP86	+1.2/+1.1	+17.5/+16.4
BP86-D3//BP86	−1.2	+16.3
MP2/def2-TZVPP [129]//BP86	−2.0	+18.9
BP86-D3	−3.7/−2.9	+15.1/+14.2
B97-D	−3.0	+14.2
TPSS-D3	−2.8/−3.1	+15.7/+14.9
MP2/def2-TZVPP//B97-D	−7.2	+19.0
MP2/def2-TZVPP//BP86-D3	−10.0	+19.1
MP2/def2-QZVPP [130]//BP86-D3	−10.4	+18.9

3.1.2 Transition-state searches

To identify the transition state (TS) between structures **1A** and **1B**, we consider two different search strategies. On the one hand, we perform a quadratic synchronous transit (QST3) transition state search with Gaussian, and on the other hand, we search for the transition states with Turbomole by eigenvector following starting from a reasonable guess structure. A reasonable guess structure means that the most negative eigenvalue of its geometry Hessian should describe the transition between **1A** and **1B**.

For the transition state search with Turbomole (TPSS-D3/SVP and def2-TZVP, BP86-D3/def2-TZVP) we choose two different starting structures as shown on the left in Figures 3.2 and 3.3, respectively. Because there is always only one vector on one of the two oxygen atoms involved in the transition from **1A** to **1B** and because of the fact that we found a minimum structure between **1A** and **1B** (denoted as **1AB**), we expect to have two transition states between **1A** and **1B**. In agreement with this assumption, the Turbomole calculations converge to two transition states, denoted as **TS-2** and **TS-1** (on the right hand side of Figures 3.2 and 3.3, respectively).

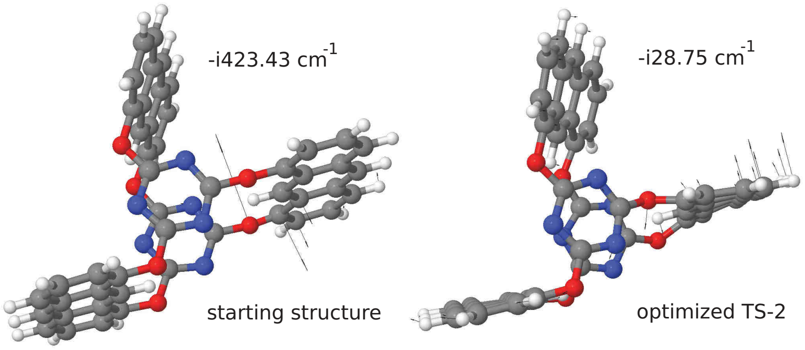


Figure 3.2: On the left, the first guess structure for a TURBOMOLE transition state search (TPSS-D3/SVP) and the mode with the most negative eigenvalue are shown. On the right, the optimized **TS-2** structure is depicted with its imaginary frequency. Element color code: gray, C; white, H; blue, N; red, O.

The quadratic synchronous transit QST3 calculation performed with GAUSSIAN (B97-D/TZVP) relies on three input structures, namely two minima and a transition state guess structure that should connect the two minima. In this study, structure **1A** or structure **1B** and **1AB** are the two minima (see Figure 3.4). For the transition state guesses we adapt the two COC angles that change during the transition from **1A** to **1B**. One COC angle is enlarged to 180° and the other one is kept at 118°. The smaller COC angle of 118° is either oriented as in structure **1A** or as in **1B**, depending on the TS we are looking for. Note that we employ a different density functional compared to the TURBOMOLE calculations, namely B97-D instead of TPSS-D3. Both functionals are dispersion-corrected ones and have been tested to yield reliable results for similar molecules [127, 128]. Since the D3 correction needed for the TPSS-D3 calculations is not implemented in our current GAUSSIAN version, we choose the B97-D functional here.

The calculations converge to a transition state between **1A** and the minimum (**TS-1**), with one imaginary frequency of $-i34.59\ \mathrm{cm}^{-1}$ and a transition state between the minimum and **1B** (**TS-2**), with one imaginary frequency of $-i22.69\ \mathrm{cm}^{-1}$. These structures are very similar to the corresponding ones obtained with TURBOMOLE.

The structure optimization of **TS-1** leads to structure **1A** and the one of **TS-2** to structure **1B**, which already indicates that **TS-1** and **TS-2** are tran-

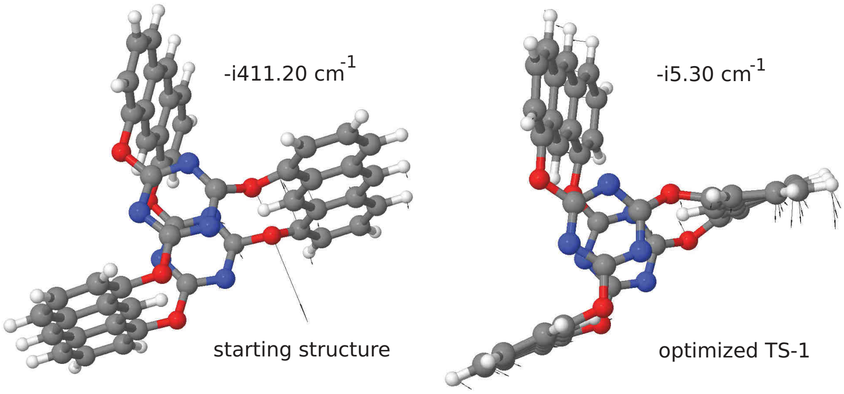


Figure 3.3: On the left, the second guess structure for a Turbomole transition state search (TPSS-D3/SVP) and the mode with the most negative eigenvalue are shown. On the right, the optimized **TS-1** structure is depicted with its imaginary frequency. Element color code: gray, C; white, H; blue, N; red, O.

sition states between **1A** and **1B**. However, the calculation of an intrinsic reaction coordinate is necessary and will be discussed in the next section.

In order to figure out whether the default convergence criteria in Gaussian are tight enough, we calculate the **TS-1** transition state again with parameters opt = verytight and grid = ultrafine and the result is the same ($E_{\mathrm{el}} = -2622.370007$ hartree in comparison to E_{el} (opt = default, grid = fine) = -2622.370054 hartree).

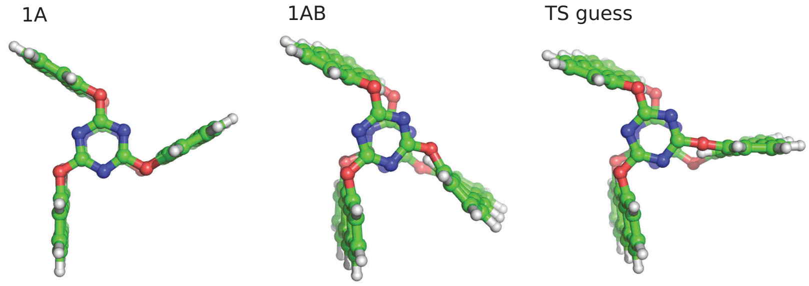


Figure 3.4: Input structures for the QST3 calculation of the first transition state, **TS-1** (between **1A** and **1AB**), with Gaussian. Element color code: green, C; white, H; blue, N; red, O.

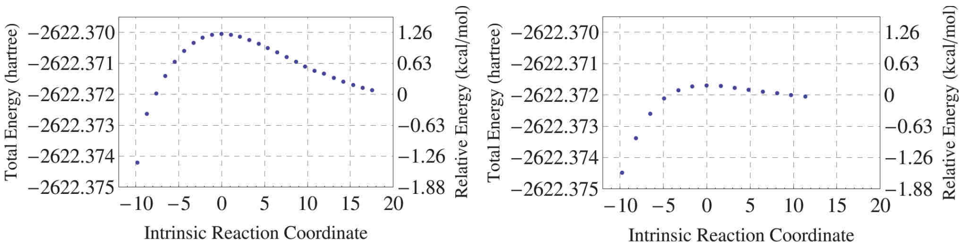


Figure 3.5: Intrinsic reaction coordinates starting from **TS-1** (left) and **TS-2** (right) calculated with GAUSSIAN (B97-D/TZVP).

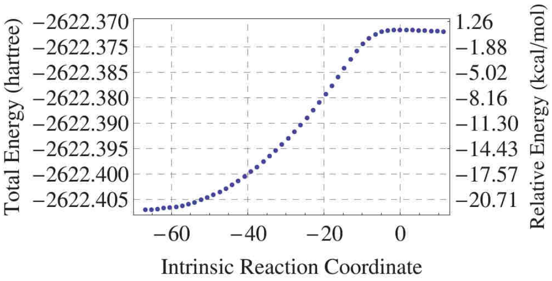
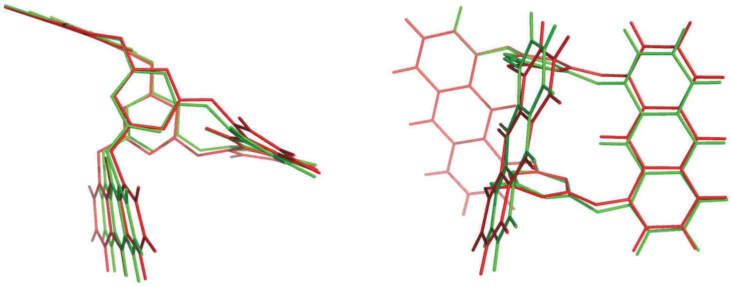


Figure 3.6: Complete intrinsic reaction coordinates from structure **1B** to **1AB** starting from **TS-2** calculated with GAUSSIAN (B97-D/TZVP).

3.1.3 Intrinsic reaction coordinates

To investigate whether the stationary points found in the transition state searches are really transition states between structures **1A** and **1B**, we calculate the intrinsic reaction coordinates (IRC) with GAUSSIAN (see Figures 3.5 and 3.6). If we follow the eigenvector of the imaginary frequency in one direction (negative intrinsic reaction coordinate), the energy decreases rapidly and the (previous) **TS-1** and **TS-2** starting structures approach structures **1A** and **1B**, respectively.

By following the eigenvector in the other direction (positive intrinsic reaction coordinate), a potential energy minimum is detected in both IRC (starting from **TS-1** and from **TS-2**), which should be the same minimum, since it connects **TS-1** and **TS-2** (see right hand sides of both graphs in Figure 3.5). The energy of the two minima found in the IRC differs by only 0.1 kcal/mol, which indicates the accuracy of the IRC approach. If we optimize these two minima and the **TS-1** and **TS-2** structures with



green: minimum structure in IRC
red: optimized structure starting from TS-1/TS-2 with distortion along EV or from minimum structure in IRC (green structure)

Figure 3.7: Alignment of the minimum obtained from the IRC (starting from **TS-1**, green) and the optimized minimum structure (red) calculated with Gaussian (B97-D/TZVP).

a small distortion along the eigenvector with imaginary frequency in direction of **1AB**, we obtain the same minimum structure, which is 0.2 kcal/mol lower in energy than the IRC minimum. The alignment of the minimum structure and the IRC "minimum" structure (Figure 3.7) shows that the structures are slightly different (especially with respect to the distortion out of the C_3N_3 plane). In the IRC calculation we do not use opt = (gediis,verytight) and grid = ultrafine, as in the structure optimization. This might explain the small structural differences.

3.1.4 Comparison of the transition-state structures

The two transition states, **TS-1** and **TS-2**, and the minimum structure **1AB** are relatively similar. The root-mean-square deviations (RMSD) between **TS-1** and **1AB** are 0.56 Å (Gaussian) and 0.19 Å (Turbomole) and between **TS-2** and **1AB**, the RMSD are 0.40 Å (Gaussian) and 0.30 Å (Turbomole). Still, in the minimum structure the two COC angles, which change during the transition, are more symmetrically arranged and in the TS structures always one COC angle approaches a 180° angle from the top, and the other one is bend in one direction (either **1A** or **1B**) according to the transition state between **1A** and **1AB** or **1B** and **1AB**. The transition states and **1AB** structures calculated with Gaussian and Turbomole are like diastereomers, since the distortion of the COC

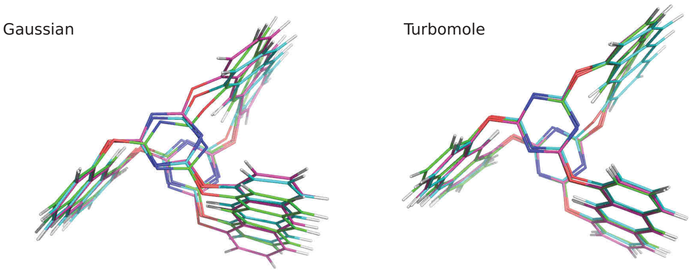


Figure 3.8: Alignment of the **TS-1** (green), **TS-2** (blue) and **1AB** (pink) structures calculated with Gaussian/B97-D/TZVP (on the left) and Turbomole/TPSS-D3/SVP (on the right).

angle out of the C_3N_3 plane accidentally occurs in different directions. In principle, there are two additional **1AB**, **TS-1** and **TS-2** structures, which are enantiomers of the two diastereomers we found.

3.1.5 Transition pathways

To draw conclusions about transition probabilities from one minimum to the other, it is necessary to analyze transition barriers between the minima. The TS optimizations reveal that for the transformation from structure **1A** to **1B**, at least 19 kcal/mol have to be provided to reach the first transition state, **TS-1**, and after a shallow energy minimum (**1AB**), a second transition state, **TS-2**, of about the same electronic energy as **TS-1**, exists, which connects **1AB** with structure **1B**.

In Figure 3.9, we show the rearrangement energy path calculated with BP86, BP86-D3 and TPSS-D3 [131–136]. The electronic energy profiles are very similar, and the addition of the D3 dispersion correction to the BP86 functional does not change the electronic energy profile significantly. Due to the double-decker nature of **1**, it is reasonable that the transition from **1A** to **1B** occurs in two steps and requires two transition states. This implies that the C-O-C angles change one after the other with an energetically shallow intermediate, in which the C-O-C angles are of the same size, but the upper and the lower part of the blades are twisted in different directions (**1AB**). The electronic- energy difference between the minimum **1AB** and the two transition states is very small (about 1

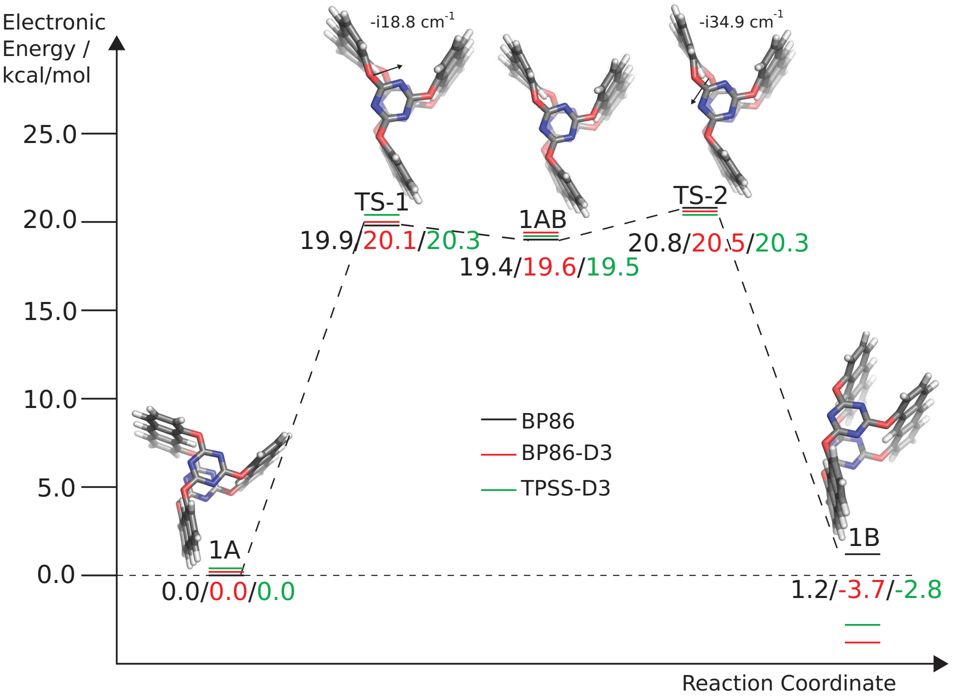


Figure 3.9: Electronic-energy transition path from structure **1A** to **1B** calculated with density functionals BP86, BP86-D3, and TPSS-D3 and with the def2-TZVP basis set. The BP86/def2-TZVP-optimized structures are depicted. For TS-1 and TS-2, the mode with imaginary frequency is shown. Element color code: gray, C; white, H; blue, N; red, O.

kcal/mol). By carefully probing our results on technical artifacts (such as optimization thresholds or numerical grid sizes), we confirm the existence of a small dip in the potential energy path from **1A** to **1B**.

3.1.6 Dependence on functional, basis set, dispersion correction and solvation

To investigate the influence of the basis set, the dispersion correction and COSMO solvation (with $\epsilon = 46.7$ for DMSO) on the reaction energy profile, we perform single-point calculations on the Turbomole/TPSS-D3/SVP optimized structures. The results are shown in Figure 3.10.

Qualitatively, the results of the dispersion-corrected calculations are similar among themselves. Without dispersion correction the barriers

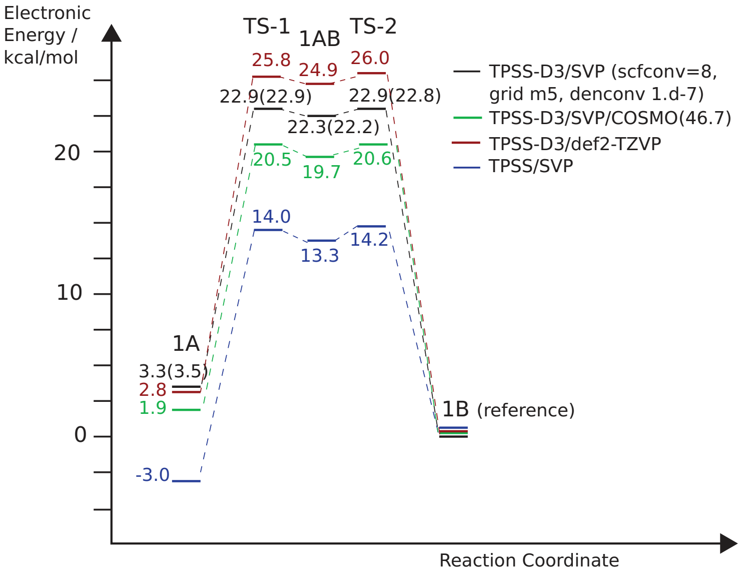


Figure 3.10: Results of various DFT and HF single-point calculations on the TPSS-D3/SVP optimized structures in the reaction energy profile from **1A** to **1B** (calculated with Turbomole). Structure **1B** has been arbitrarily chosen as the zero energy reference in each of the calculations.

from **1A** to **TS-1** or **1B** to **TS-2** are significantly smaller than the results with dispersion correction (14 kcal/mol instead of 23 kcal/mol). The electronic energy difference between the minimum **1AB** and the two transition states is small (about 1 kcal/mol). Therefore, we probe our results on technical artifacts by employing single-point calculations with a more accurate grid size of m5, a tighter SCF convergence criterion where the energy difference between two cycles was less than $1{\times}10^{-8}$ hartree, and a density convergence criterion of $1{\times}10^{-7}$ (see values in parentheses). The deviations from the results with smaller grid size (m3), energy convergence criterion of $1{\times}10^{-7}$ and no density convergence criterion are very small (less than 0.2 kcal/mol).

We also perform transition state searches with BP86/def2-TZVP, BP86-D3/def2-TZVP and TPSS-D3/def2-TZVP besides TPSS-D3/SVP to verify that our results are not affected by a particular choice of density functional, dispersion corrections and the basis set (see Table 3.2). The transition state optimizations are carried out with Turbomole starting from the TPSS-D3/SVP transition structures or from structures close to these two TSs, if no imaginary frequency is obtained for the TPSS-D3/SVP opti-

Table 3.2: Free energy differences ΔG's (in kcal/mol) for the reaction profile from **1A** to **1B**. T = 298.15 K and p = 0.1 MPa has been chosen.

functional/basis set	ΔG			
	1A→TS-1	**TS-1→1AB**	**1AB→TS-2**	**TS-2→1B**
BP86/def2-TZVP	+19.8	−1.8	+2.5	−19.4
BP86-D3/def2-TZVP	+20.0	−1.8	+2.2	−23.3
TPSS-D3/SVP	+19.1	−2.0	+2.0	−21.8
TPSS-D3/def2-TZVP	+20.1	−2.0	+2.0	−23.1

mized structure with the respective employed density functional and basis set. The minimum structures is obtained by optimizations of the TPSS-D3/SVP transition structures.

The BP86 and TPSS free energy profiles are very similar and the application of the D3 dispersion correction in the BP86 functional does not change the free energy profile significantly. Moreover, the larger def2-TZVP basis set yields similar results compared to the SVP basis set. Hence, we conclude that the parameters, which we have initially chosen, are sufficient. However, in our work we report the results obtained with def2-TZVP basis set.

Recall that the imaginary frequency at a transition state is not included in the free energy calculation. The approximation of this entropy contribution by translational and vibrational partition functions, leads to a free energy change of about 4–5 kcal/mol.

3.1.7 Entropy contribution of imaginary frequency

We approximate the entropy contribution of the imaginary frequency at the transition states by the vibrational and translational partition functions. For the first approximation, we have choose two small positive frequencies, $\epsilon_1 = 1\ \text{cm}^{-1} = 1.99\times10^{-23}$ J and $\epsilon_2 = 0.01\ \text{cm}^{-1} = 1.99\times10^{-25}$ J and obtain with

$$S_{\text{vib}} = R\left(\frac{\epsilon/kT}{e^{\epsilon/kT}-1} - \ln\left(1 - e^{-\epsilon/kT}\right)\right) \tag{3.1}$$

at T = 298.15 K an entropy of $S_1 = 52.65$ J/(mol K), which corresponds to a free energy contribution of $G_1 = 3.8$ kcal/mol (T = 298.15 K). For the second frequency, we obtain $S_2 = 90.94$ J/(mol K) and $G_2 = 6.5$ kcal/mol.

If we assume that the transition frequency can be approximated by a translation of the particle in a one-dimensional box (note that we intend to approximate only one degree of freedom), we obtain from the translational partition function,

$$q_{\text{trans}} = L \times \left(\frac{2\pi mkT}{h^2}\right)^{1/2}, \qquad (3.2)$$

the entropy

$$S_{\text{trans}} = R(\ln(q_{\text{trans}}e) + T\frac{\partial \ln q}{\partial T}) = R\left(\ln q_{\text{trans}} + 3/2\right), \qquad (3.3)$$

which amounts to $S_{\text{trans}} = 69.515$ J/(mol K) (with T = 298.15 K, p = 0.1 MPa, $m = 1.296\times10^{-24}$ kg and $q_{\text{trans}} = 954.044$). This entropy corresponds to a free energy of G = 5.0 kcal/mol.

3.1.8 Stability of the two solvated conformers

The relative stability of structures **1A** and **1B** is analyzed by performing Born-Oppenheimer molecular dynamics (BOMD) simulations of both conformers embedded in a box of 32 DMSO molecules under periodic boundary conditions. Out of the trajectories, 34 snapshots of the solute and 19 DMSO molecules are chosen and subjected to constrained BP86/def2-SVP(DMSO)/def2-TZVP(solute) optimizations in order to obtain the energy minimum for each solvent arrangement.

After about 4 ps simulation time, the energy and temperature of the system consisting of the monomer in 32 DMSO molecules is approximately converged and we therefore assume that our system is equilibrated.

By averaging the dispersion-free BP86 energies E_{BP86} of the first 24 snapshots for each conformation (**1A** and **1B**, respectively) over 12 ps simulation time after the equilibration, we obtain an energy difference of 4.4 kcal/mol, where solvated **1A** is energetically lower (see Table 3.3).

Since standard DFT calculations usually neglect attractive long-range dispersion effects, we calculate the dispersion interaction energies separately. If we switch on dispersion interactions in the single-point calculations, the electronic energy difference between **1A** and **1B** amounts to

−3.4 kcal/mol. This result is in contrast to the results obtained for the isolated molecules, where we found that **1A** is disfavored over **1B** by 1–4 kcal/mol.

To uncover which interactions contribute most to the observed energy differences, we analyze the individual contributions of the BP86 dispersion-free energies, E_{BP86}, and dispersion energies, E_{disp}.

First, we calculate the solvent and solute contributions of E_{BP86} separately. The BP86 dispersion-free energy for the solvent-solute interactions is then obtained as the difference between the overall electronic energy minus the solvent contribution and the solute contribution. The calculations reveal that the BP86 dispersion-free energy between the solvent atoms and conformation **1A** is by 3.3 kcal/mol larger than the BP86 dispersion-free energy between the solvent atoms and conformation **1B**. $E_{\mathrm{BP86,M}}$ amounts to −1.2 kcal/mol (the structure optimization has not been performed with dispersion correction) and $E_{\mathrm{BP86,S}}$ amounts to only +0.2 kcal/mol. Although $\Delta E_{\mathrm{BP86,S}}$ does not have to be zero, this result indicates that the contribution of the solvent arrangement is relatively small. Still, the standard deviation of ± 6–7 kcal/mol is large.

Next, to understand the difference in the dispersion energy of the two solvated conformers, we split E_{disp} into dispersion energies between the solute and all solvent molecules and E_{disp} calculated for solute only and solvent molecules only, respectively (see Table 3.4). We observe that the dispersion interaction energy between structure **1A** and all DMSO molecules is by 3.7 kcal/mol larger than for **1B**. The solvent-solvent dispersion interaction energies of solvated conformer **1A** is by 2.1 kcal/mol less than that of **1B** and the solute-solute dispersion interaction of **1A** is by 2.6 kcal/mol less than that of **1B**.

In the isolated conformers, $E_{\mathrm{disp,M}}$ amounts to −74.1 kcal/mol for **1A** and −76.4 kcal/mol for **1B**, which results in a $\Delta E_{\mathrm{disp,MM}}$(**1A**-**1B**) of +2.3 kcal/mol. In the solvent, the $\Delta E_{\mathrm{disp,MM}}$ value (+2.6 kcal/mol) is very similar to the gas-phase results.

The BP86 dispersion-free solute-solvent energy of −3.3 kcal/mol and the solute-solvent dispersion interaction energy of −3.7 kcal/mol lead to the inversion of the energy order of the solvated structures with respect to the isolated ones. Concluding, the solvent-solute interaction energy dominates the intramolecular interaction energy. It must be noted though that the simulation time is rather short (and the box size is small and the

Table 3.3: Average of the relative electronic energies, E_{BP86}, the electronic energies of the solvent, $E_{\text{BP86,S}}$, the electronic energy of the monomer, $E_{\text{BP86,M}}$, and the electronic energy between solute and solvent, $E_{\text{BP86,SM}}$, (in kcal/mol) over 25 snapshots every 0.5 ps after the equilibration period of about 4 ps of structures **1A** and **1B** in 19 DMSO molecules. The def2-SVP basis set has been chosen for all DMSO molecules and the def2-TZVP basis set has been chosen for the monomer. M denotes monomer, S solvent and SM solvated monomer.

conformer	E_{BP86}	$E_{\text{BP86,S}}$	$E_{\text{BP86,M}}$	$E_{\text{BP86,SM}}$
1A	−4.4 (± 7.6)	+0.2 (± 5.8)	−1.2 (± 0.9)	−22.1 (± 5.4)
1B	0.0 (± 8.9)	0.0 (± 7.4)	0.0 (± 1.2)	−18.7 (± 4.4)
ΔE_{BP86}(**1A**-**1B**)	−4.4	+0.2	−1.2	−3.3

number of DMSO molecules selected is even smaller). An elongation of the simulation might change the mean values significantly.

Table 3.4: Average of the dispersion interaction energies between all atoms, E_{disp}, the intramolecular $E_{\text{disp,SS}}$ and $E_{\text{disp,MM}}$ dispersion interaction energies and the intermolecular $E_{\text{disp,SM}}$ dispersion energies between solute and solvent (in kcal/mol) over 25 snapshots of every 0.5 ps after the equilibration period of about 4 ps structures **1A** and **1B** in 19 DMSO molecules. M denotes monomer, S solvent and SM solvated monomer.

conformer	E_{disp}	$E_{\text{disp,SS}}$	$E_{\text{disp,MM}}$	$E_{\text{disp,SM}}$
1A	−231.2 (± 7.6)	−108.5 (± 4.2)	−74.0 (± 0.2)	−48.7 (± 6.3)
1B	−232.2 (± 10.3)	−110.6 (± 7.2)	−76.6 (± 0.7)	−45.0(± 5.6)
ΔE_{disp}(**1A**-**1B**)	+1.0	+2.1	+2.6	−3.7
ΔE_{el}(**1A**-**1B**)	−3.4			

To investigate the role of the simulation time, we continued the BOMD simulations and took 34 snapshots every 0.5 ps after the equilibration period of about 4 ps. The averaged BP86 dispersion-free and dispersion energies are listed in Tables 3.5 and 3.6. The only values that differ significantly from the average values over 25 snapshots are the BP86 dispersion-free and dispersion contributions from the solvent molecules themselves. Thus, on average the arrangement of the DMSO molecules around conformer **1B** appears to be energetically preferred compared to the arrangement of DMSO molecules around **1A** in the additional

9 snapshots. Due to these stabilizing solvent-solvent interactions the dispersion-corrected BP86 electronic energy difference between **1A** and **1B** in 19 DMSO molecules decreases from 3.4 kcal/mol to 1.7 kcal/mol.

It should be stressed that this exploratory study has been performed in order to obtain a qualitative estimate of the inter- and intramolecular interaction energies of the solvated conformers. Although it cannot be guaranteed that the simulation time is long enough and the standard deviations are relatively large, we observe that the solvent-solute interaction energies ($E_{\mathrm{BP86,SM}}$ and $E_{\mathrm{disp,SM}}$) of conformer **1A** are still more exothermic than the corresponding values of conformer **1B**.

Table 3.5: Average of the relative BP86 dispersion-free energies of the whole system, E_{BP86}, of the solvent only, $E_{\mathrm{BP86,S}}$, of the monomer, $E_{\mathrm{BP86,M}}$, and between solute and solvent, $E_{\mathrm{BP86,SM}}$, (in kcal/mol) over 34 snapshots every 0.5 ps after the equilibration period of about 4 ps of structures **1A** and **1B** in 19 DMSO molecules calculated with BP86. The def2-SVP basis set has been chosen for all DMSO molecules and the def2-TZVP basis set has been chosen for the monomer. M denotes monomer, S solvent and SM solvated monomer.

conformer	E_{BP86}	$E_{\mathrm{BP86,S}}$	$E_{\mathrm{BP86,M}}$	$E_{\mathrm{BP86,SM}}$
1A	−2.8 (± 7.0)	+2.5 (± 6.1)	−1.2 (± 0.9)	−22.1 (± 5.0)
1B	0.0 (± 9.8)	0.0 (± 8.7)	0.0 (± 1.1)	−18.0 (± 4.5)
ΔE_{BP86}(**1A**-**1B**)	−2.8	+2.5	−1.2	−4.1

Table 3.6: Average of the dispersion interaction energies between all atoms, E_{disp}, the intramolecular $E_{\mathrm{disp,SS}}$ and $E_{\mathrm{disp,MM}}$ dispersion interaction energies and the intermolecular $E_{\mathrm{disp,SM}}$ dispersion energies between solute and solvent (in kcal/mol) over 34 snapshots of every 0.5 ps after the equilibration period of about 4 ps structures **1A** and **1B** in 19 DMSO molecules. M denotes monomer, S solvent and SM solvated monomer.

conformer	E_{disp}	$E_{\mathrm{disp,SS}}$	$E_{\mathrm{disp,MM}}$	$E_{\mathrm{disp,SM}}$
1A	−231.0 (± 8.0)	−107.9 (± 4.8)	−74.0 (± 0.2)	−49.1 (± 6.0)
1B	−232.1 (± 9.1)	−111.9 (± 7.1)	−76.6 (± 0.7)	−43.7(± 5.7)
ΔE_{disp}(**1A**-**1B**)	+1.1	+4.0	+2.6	−5.4
ΔE_{el}(**1A**-**1B**)	−1.7			

In Figure 3.11 the positions of three DMSO molecules adopted in the snapshots obtained from the BOMD simulation of conformer **1A** are depicted in order to show the range of solvent-molecule orientations that have been sampled.

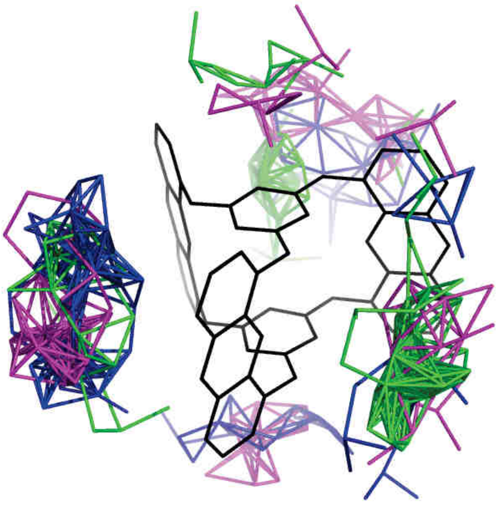


Figure 3.11: Superposition of the movement of three arbitrarily chosen DMSO molecules (shown in green, blue and purple) in the chosen snapshots obtained from the BOMD simulation of conformer **1A** (black). The hydrogen atoms are not shown.

3.1.9 Consequences for the double-decker-type rotamer

The calculated energy barrier between the two conformers **1A** and **1B** theoretically permits transformations between them at elevated temperatures. However, under experimental conditions, also solvent-solute interactions have to be taken into account. DMSO is known for its general strong solvation and was therefore chosen as an example solvent for BOMD simulations of explicitly solvated structures **1A** and **1B** to predict whether these interactions would play a role. The analysis of BOMD snapshot structures of the solvated conformers indicates that conformer **1A** is slightly more stable compared to **1B**. This result is in contrast to the theoretical results obtained for the isolated conformers. To understand why the energy of structure **1A** in DMSO is lower than the one of structure **1B** in DMSO, the inter- and intramolecular contributions to the electronic dispersion-free and dispersion-affected interaction energies were analyzed. We observed that on average, the dispersion interaction energy between structure **1A** and all solvating DMSO molecules in a given number of snapshots (intermolecular contribution) is by 5.4 kcal/mol more exothermic than the dispersion interaction energy between structure **1B** and its DMSO solvent shell. The BP86 dispersion free intrinsic interaction energy of **1A** with its DMSO solvent shell is, on average, by 4.1 kcal/mol larger than for **1B**: the intrinsic interaction energy was calculated from two single-point DFT calculations for the solute and the solvent-shell, respectively, subtracted from the total energy of a snapshot. Because in **1B** two anthracene blades are stacked ftf to each other with no DMSO molecules between them, less solvent-shell-solute interactions occur than for **1A**. From this analysis, we therefore deduce a preference for the C_{3h} symmetric conformation **1A** in solution. However, the formation of **1B** is not ruled out. It should be noted that the fact that X-ray crystal-structure analysis of **1**, revealed solely the C_{3h} symmetric conformation **1A** might be due to possible unfavorable packing effects of conformation **1B** in the crystal, or the fact that only small amounts of **1B** are formed upon heating.

3.2 Investigation of an organic phosphaketene

The content of this Chapter has been published in Ref. [137]. Our experimental coworkers Z. Li and Prof. Grützmacher (D-CHAB, ETHZ) synthesized a stable organic phosphaketene, which is a "salt" between $[P(NR)_2C_2H_2]^+$ (R = diisopropylphenyl) and the anion OCP^- [137]. The phosphor atom of OCP^- binds to the phosphor atom of $[P(NR)_2C_2H_2]^+$, which is denoted as structure **A** in the following. In principle, a stable product with an O-P bond (**OCP_changed**) also appears reasonable. To investigate the relative energy difference between these two products, we first calculate the BP86 energy difference between the OCP^- and PCO^--bound products of a model system with R = methyl. A detailed Computational Methodology can be found in Appendix D. The BP86 energy difference amounts to 18 kcal/mol and thus reveals a clear preference of **A** over **OCP_changed** (see Figure 3.12).

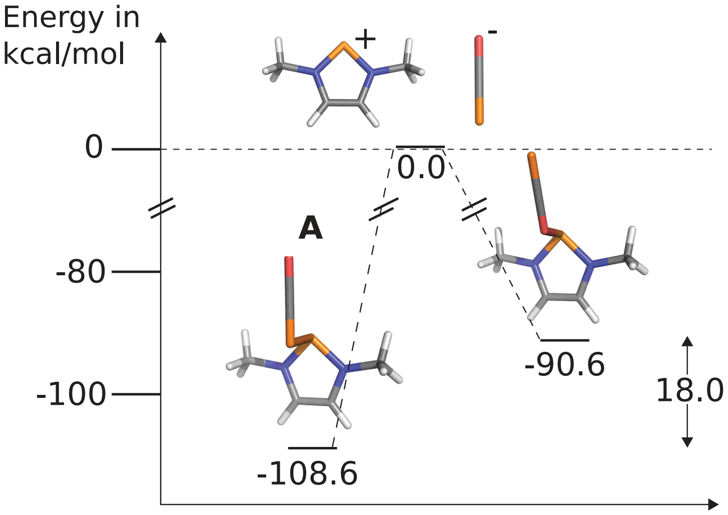


Figure 3.12: The BP86/def2-TZVP energy difference between structure $(HC)_2(NMe)_2P\text{-}P\text{=}C\text{=}O$ (**A**) and the oxy-phosphaalkyne isomer $(HC)_2(NMe)_2P\text{-}O\text{-}C\equiv P$ **OCP_changed**. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

3.2.1 Comparison of different reaction routes

Experimentally it has been observed, that a dimer between **A** and an intermediate with almost the same stoichiometry as **A** (but one carbon atom and one oxygen atom less) is formed. Our coworkers suggested a transformation pathway from **A** into an intermediate **I**, which then reacts with another molecule **A** under loss of CO to the observed dimer. Based on these conceptual ideas, we calculate BP86 and B3LYP reaction energy pathways from structure **A** to the suggested **I** (shown in Figure 3.13). By performing a constrained optimization scan along the C-C bond, which is formed in the transition from $\mathbf{A}_{\text{rot}}$ to **I**, as well as a quadratic synchronous transit approach (QST3), we obtained two different TSs, $\mathbf{TS}^1$ and $\mathbf{TS}^2$. The intrinsic reaction coordinates calculated for both, $\mathbf{TS}^1$ and $\mathbf{TS}^2$, confirmed that the transition between $\mathbf{A}_{\text{rot}}$ and **I** proceeds directly through one transition state and not through several TSs, i.e., either over $\mathbf{TS}^1$ or $\mathbf{TS}^2$. $\mathbf{TS}^1$ appears to be slightly lower in energy. Although it cannot be excluded that more reaction pathways than the reported two exist, the energy of $\mathbf{TS}^1$ may serve as an upper limit for the transition barrier.

The optimization of structure **A** leads to two local minimum energy structures, where the PCO group is rotated around the P-P bond (relative to each other). One conformer is by 1–2 kcal/mol higher in energy (called $\mathbf{A}_{\text{rot}}$) than the other one (**A**). Structure $\mathbf{A}_{\text{rot}}$ is structurally much closer to the transition state structure and the energy difference between **A** and $\mathbf{A}_{\text{rot}}$ is not very large. Noteworthy, the CO dissociation energy from **A** to **G** is by 4.6 kcal/mol more endothermic than the reaction from **A** to **I**. The activation energy for the CO dissociation from **A** to **G** amounts to 46.8 kcal/mol. Therefore, we assume that under the given reaction conditions no **G** is formed.

The dimerization pathway of structures **A** and **I** to **C** is reported in Figure 3.14.

Besides the reaction in which the three-membered phosphor ring is formed, two other reaction pathways are studied, which are a Diels-Alder reaction of **I** with dimethylbutadiene (DMB, chosen as solvent), and a 5-ring formation between **A** and a cyclopentadienone. In this latter reaction, CO is lost, which makes the back reaction very unlikely and, thus, drives the reaction towards the product side.

The reaction pathway of the Diels-Alder reaction between **I** and DMB is reported in Figure 3.15. From the analysis of the highest-occupied and

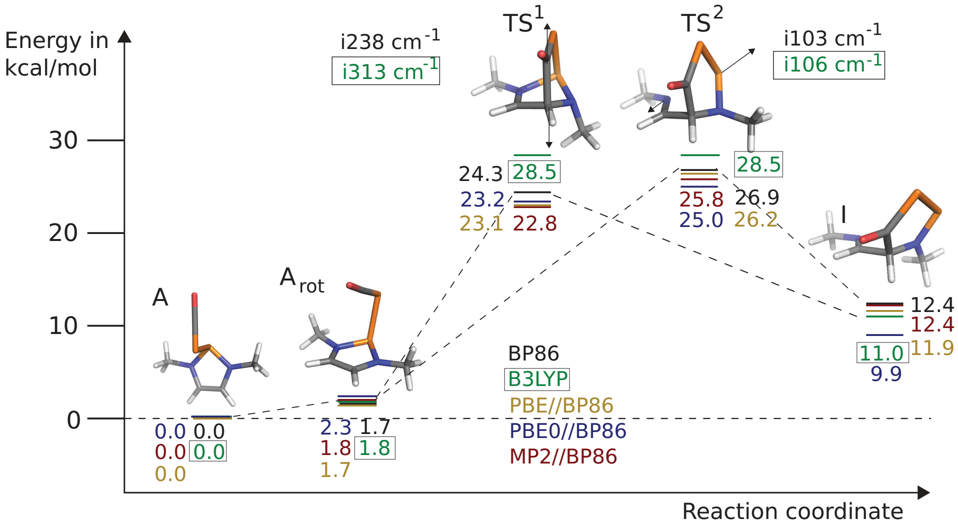


Figure 3.13: BP86/def2-TZVP and B3LYP/def2-TZVP transition path from structure $\mathbf{A}_{\mathrm{rot}}$ to **I**. The imaginary frequency of the transition states, $\mathbf{TS}^1$ and $\mathbf{TS}^2$, are reported and the BP86 optimized structures are shown. The PBE, PBE0 and MP2 single-point energies on the BP86/def2-TZVP optimized structures (denoted by e.g., PBE//BP86) are also reported. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

lowest-unoccupied molecular orbitals of the reactants (cf. Figure 3.16), we conclude that a normal Diels-Alder reaction occurs and not an inverse one [138].

In Figure 3.17 and 3.18 possible BP86/def2-TZVP transition paths of structure **A** with tetramethylcyclopentadienone **M** (and tetraphenylcyclopentadienone **P**) are displayed. The potential energy surface of this system (**A** and **M**) is probably quite flat, because several minimum-energy structures have been found, which are very close on the Born-Oppenheimer PES. Several side reactions are possible, because the PCO fragment can migrate to different positions of the dienone 5-ring.

3.2.2 Consequences for the chemistry of the phosphaketene

First of all, the product, in which a P-P bond is formed, is by 18 kcal/mol more stable than the P-O bond product. From the transition pathway calculations we observed that the Diels-Alder reaction is the most exothermic one with the lowest activation energy (only 4.4 kcal/mol). However, also the back reaction with an activation energy of 24 kcal/mol is feasible under the given conditions. Therefore, we assume that a certain amount of **A** is always present in the system. The conversion from **A** to **I** proceeds over a TS which is 23 kcal/mol higher in energy than $\mathbf{A}_{\mathrm{rot}}$. Two reaction pathways have been uncovered, which are very similar in their energy profile. The forward and backward reactions will occur at room temperature. The formation of a phosphor three-membered ring under the loss of CO is only by −1.7 kcal/mol exothermic. However, due to the CO loss, the reaction occurs only in one direction and the product will accumulate over time. The phosphaketene forms a 5-ring with methyl- and phenyl-substituted cyclopentadienones. Several stable intermediates have been obtained, which demonstrate how shallow the potential energy surface of the system is. Most of the stable intermediates have been formed without the loss of CO. However, one product, which is endothermic by +6.4 kcal/mol compared to the reactants **A** and tetramethylcyclopentadienone, lost a CO molecule. Therefore, we assume that after a certain time only the two products which have lost a CO molecule during the reaction pathway can be found, although, in principle, energetically lower reaction pathways exist.

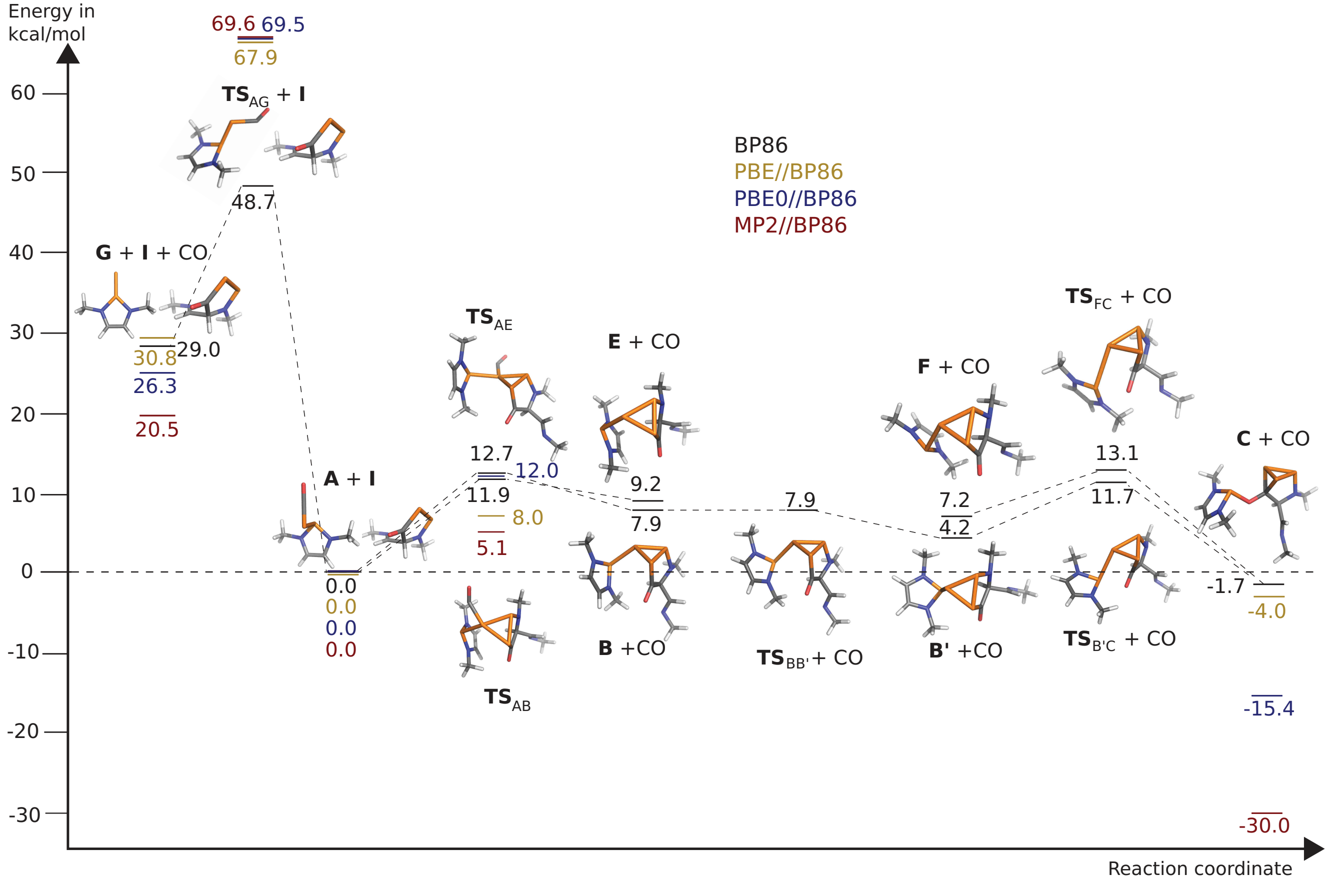


Figure 3.14: BP86/def2-TZVP transition path from structures **A** and **I** to the dimer (**C**) and CO. For comparison, the transition path for the dissociation reaction of **A** to **C** and CO is reported. Furthermore, the PBE, PBE0 and MP2 single-point energies on the BP86/def2-TZVP optimized structures (denoted by e.g., PEB//BP86) are reported for several intermediates. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

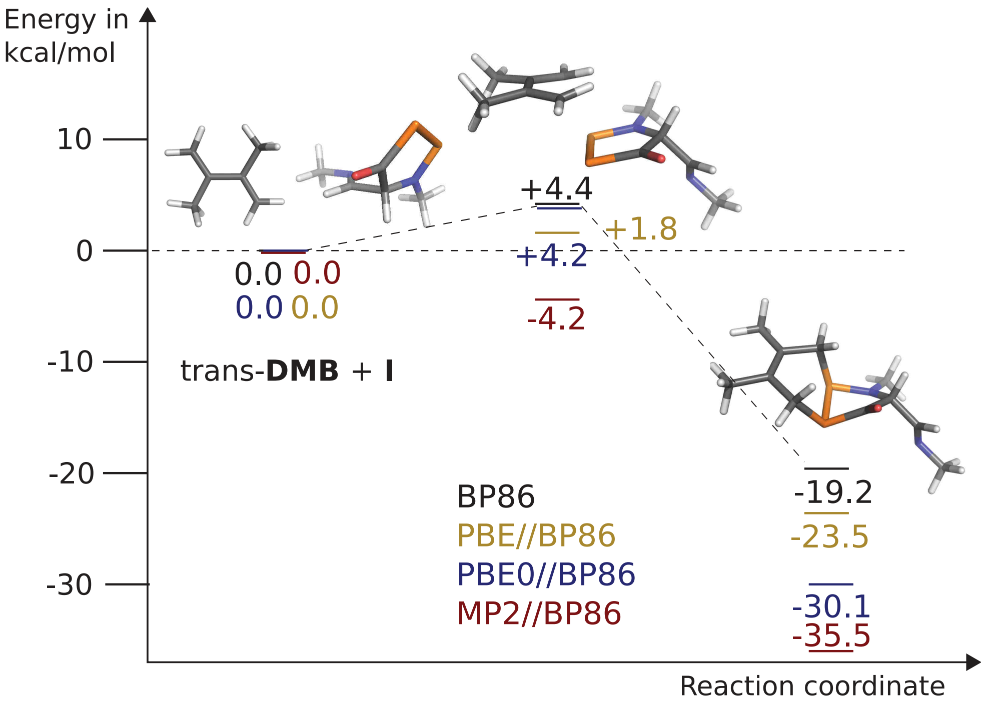


Figure 3.15: BP86/def2-TZVP Diels-Alder reaction path of **I** with trans-dimethylbutadiene (**DMB**). The PBE, PBEo and MP2 single-point energies on the BP86/def2-TZVP optimized structures (denoted by e.g., PEB//BP86) are also reported. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

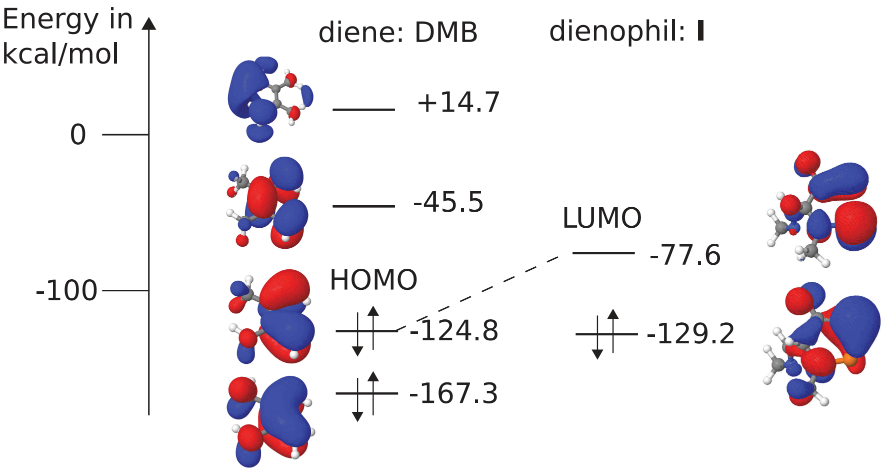

Energy in
kcal/mol
0
-100
diene: DMB
dienophil: **I**
+14.7
-45.5
HOMO
-124.8
-167.3
LUMO
-77.6
-129.2

Figure 3.16: BP86/def2-TZVP energies of the highest occupied and lowest unoccupied molecular orbitals of **I** and cis-dimethylbutadiene (**DMB**).

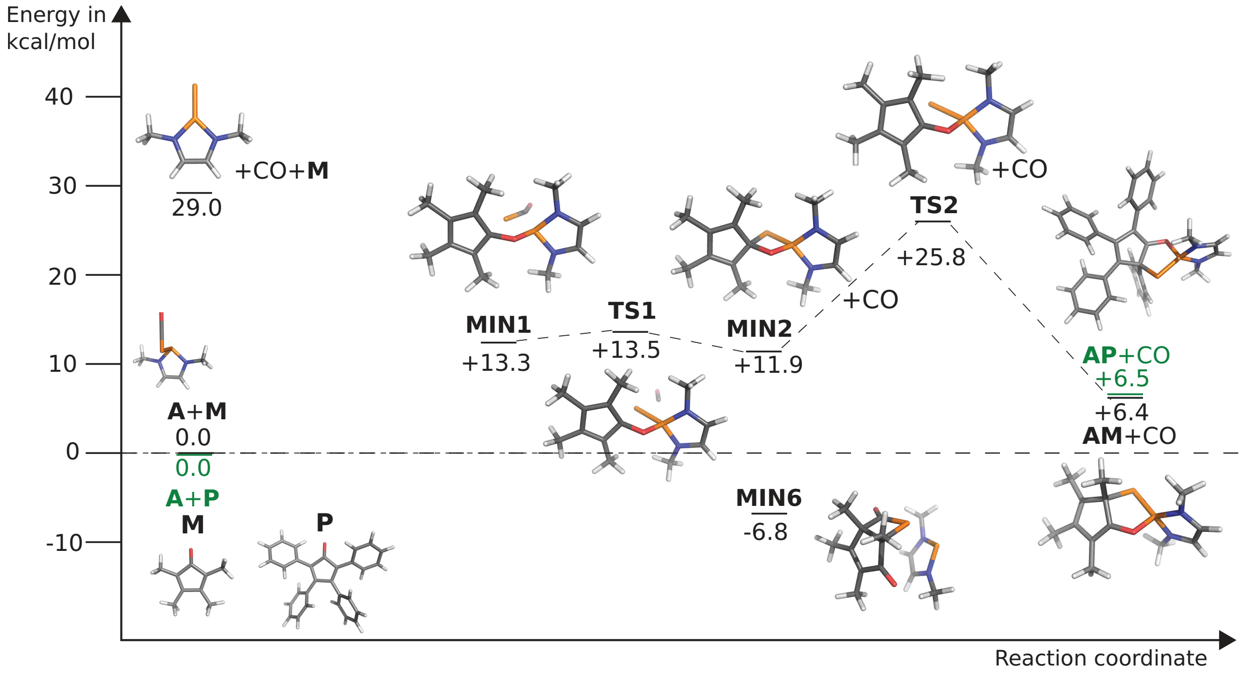


Figure 3.17: **Part 1:** Possible intermediates of the transition path from structures **A** and tetramethylcyclopentadienone **M** (and tetraphenylcyclopentadienone **P**) to the 5-ring addition products, **AM** (and **AP**), and CO calculated with BP86/def2-TZVP. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

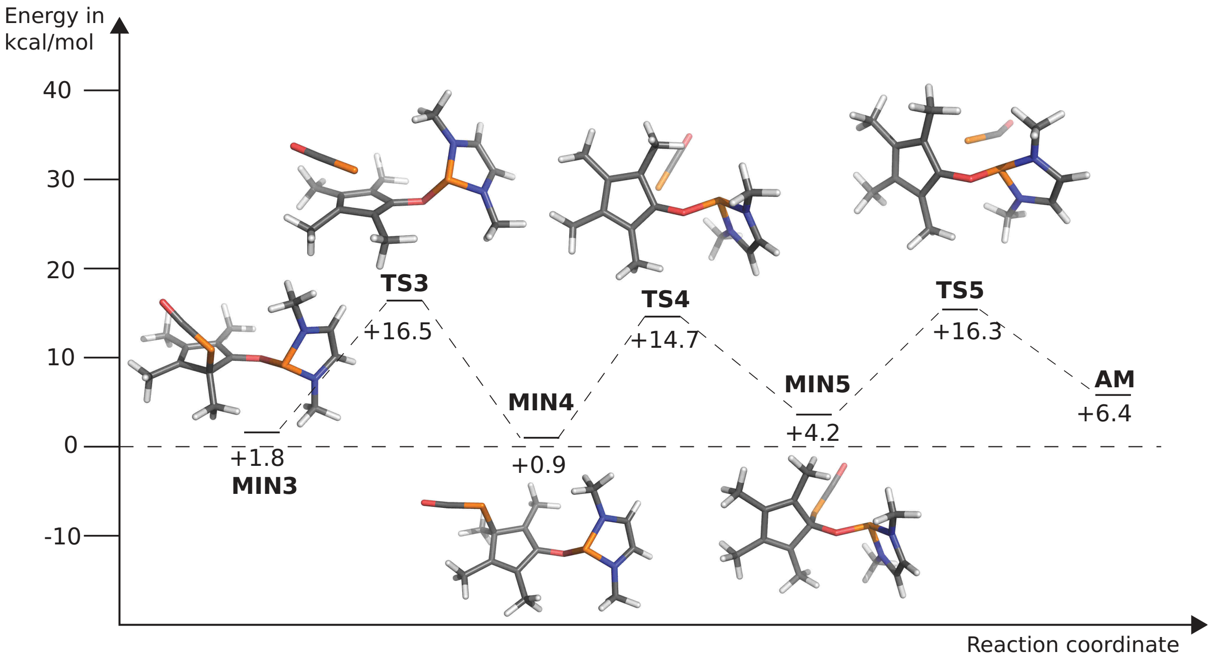


Figure 3.18: **Part 2:** Possible intermediates of the transition path from structures **A** and **M** to **AM** and CO calculated with BP86/def2-TZVP. Element color code: gray, C; red, O; blue, N; white, H; orange, P.

4

General Conclusions from The Case Studies

The first case study addressed the active centers of a [FeFe] hydrogenase and its R187D mutant, as well as a subcluster of the hydrogenase system, namely the single Fe_4S_4 cubane. These clusters are embedded in a protein environment, which affects its properties in many regards. For example, such effects may be caused by structural distortions in the ligand-sphere, charged or polar amino acids in the vicinity of the cluster, water molecules within the protein or the flexibility of the protein backbones. Moreover, these studies are rather difficult due to the complex structure of the clusters themselves. Several iron centers couple ferromagnetically and antiferromagnetically with each other and one has to ensure that for each optimization the lowest-energy spin-coupling scheme has been obtained. We observed that DFT yields reliable electronic energies even for such complex broken-symmetry spin-states as found in the Fe-S clusters. An intrinsic difficulty in the computational study of protein-embedded bioinorganic systems is that the system size should, in principle, be as large as possible to represent the protein most accurately. A larger system size is equivalent to a longer computational time, which makes these calculations often very time-consuming. Besides all these difficulties, one of the biggest challenges is the combinatorial increase of structures and spin-coupling schemes that have to be taken into account. The manual setup, submission and analysis of these calculations can become very tedious. An automated program which performs all the required steps and only reports relevant structures and energies is thus highly desirable. To facilitate computational studies on Fe-S-cluster-containing proteins, we wrote a program for the automated input structure generation as well as a scanning tool for different distortions in the (symmetrized) first ligandsphere of the active center (e.g. cysteines). With the scanning tool we investigated the sensitivity of Fe_4S_4 cluster properties such as potential energies, spin couplings, adiabatic detachment energies, inner-sphere reorganization energies and reactivities on structural distortions in the first ligandsphere. We observed a dependence of the ground-state spin-coupling scheme on the structural distortion. For inner-sphere reorganization energies the spin-coupling scheme adopted by the broken-symmetry wave function plays an important role since it can change the reorganization energies by up to 13 kcal/mol.

In the second case study we localized several TSs to compare the relevance of the corresponding reaction pathways. We found that one of the biggest challenges in reaction path analyses is the localization of TSs on a PES with many shallow minima. Moreover, the knowledge about the system and its chemical properties facilitates the identification of reaction pathways. Computational time can be saved by searching for analogies between not yet investigated reaction paths and known ones. Especially for the investigation of differently charged systems, a TS search from an already converged TS structure of a differently charged system is a promising TS-guess structure. If large systems are studied, the calculation of all vibrational modes is computationally demanding. However, most transition-state optimizers require the full Hessian at least once in the beginning and then update an approximate Hessian.

The knowledge of all minimum-energy structures on a PES would simplify the transition path calculation significantly, because elementary-steps could be identified more easily. Chemical heuristics, i.e., experience-based structure transformations, may enable us to obtain all relevant minimum-energy structures. Subsequently, for the localization of TSs, all minimum-energy structures have to be compared with each other to obtain those structure pairs between which a TS search shall be carried out.

In Chapters 5 and 6 two computational algorithms, that have been implemented based on the experience made from these two case studies, are presented.

5

Mode-Tracking based Transition-State Optimizations

The optimization of transition-state structures (TSs) is key to the understanding of mechanisms and kinetics of chemical reactions on a computational basis. Transition states are defined as first-order saddle-point structures located on the minimum (reaction) energy path between reactants and products. First-order saddle points are characterized by one negative eigenvalue of the matrix of second partial derivatives of the electronic energy with respect to the Cartesian nuclear coordinates, i.e. of the Hessian. Reactants and products are local minima on the Born–Oppenheimer potential energy surface (PES). The energy differences between a TS and two minima of an elementary reaction are the activation energy barriers. They should in principle be evaluated from the Gibbs free energy, but are approximated here, as in most quantum-chemical studies, by the electronic energy at zero Kelvin (neglecting temperature and entropy contributions).

Numerous methods have been developed to efficiently find TSs. Examples are interpolation methods [139], eigenvector following (EVF) [140–144], string methods [145], and the scaled hypersphere search method [146]. The existing TS search methods can be divided into those that start from one structure (often called single-ended methods) or those that require at least two starting structures, usually reactant and product structures (double-ended methods). Double-ended TS search algorithms are often based on interpolation methods such as linear (LST [139, 147]) or quadratic synchronous transit (QST [139]), string methods or nudged elastic band (NEB [145, 148]) algorithms. Since the double-ended methods usually show slow convergence near a TS [7], they are mainly employed to find a guess structure close to the TS, which is then refined by a more efficient single-ended method, such as EVF. Hence, it is beneficial to combine single-ended and double-ended methods for TS searches.

In most of the EVF-based methods, the full Hessian of the transition-state guess structure is calculated to obtain the exact vibrational mode to follow. For large molecules, the complete Hessian calculation is computationally demanding as the calculation of the elements of the Hessian matrix is very time consuming within a first-principles electronic-structure description. Therefore, several algorithms have been developed to circumvent the calculation of the full Hessian in structure-optimization algorithms. A quasi-Newton–Raphson method has been introduced by Broyden [149, 150]. In this method, an approximate Hessian is built from gradients only and then updated (according to Bofill [151] and Powell [152]) by the gradients of intermediate points obtained during the optimization. These methods reduce the computational effort significantly, but for large molecules a further reduction of the computational cost is desirable.

Recently, Sharada et al. [153] introduced an approximate-Hessian approach based on the tangent of the transition-state guess structure determined by an interpolation between reactant and product structures and by local curvature information. This approximate Hessian approach combined with the growing string method turned out to be computationally less expensive than previous Hessian approximations [154, 155].

Since the efficiency of a TS search depends strongly on the initial Hessian, a main goal is to set up an approximate Hessian matrix that resembles the exact one as closely as possible. In 2002, we proposed an algorithm based on Davidson subspace diagonalization for the selective calculation of eigenvectors of the mass-weighted Hessian based on predefined molecular distortions [156]. This so-called Mode-Tracking scheme turned out to be very efficient in vibrational spectroscopy [156–164]. Because of the straightforward and flexible implementation, Mode-Tracking was implemented in a semi-numerical fashion [156]. At the same time, Deglmann and Furche [165] presented an implementation of a fully analytical Davidson subspace diagonalization of the Hessian for the optimization of its lowest eigenvalue required for the identification of stationary points.

Very recently, Sharada et al. [166] described a semi-numerical Davidson subspace iteration method to obtain selected information of the Hessian spectrum, which is identical to Mode-Tracking [167]. For transition-state optimizations, Sharada et al. [166] extract the guess mode from the coordinates along the pathway obtained from the freezing-string method (FSM). In contrast to the Hessian approach presented by Sharada et al. [166], we here develop a Mode-Tracking-based TS and minimum localization algorithm that can iteratively refine a specific eigenvector of interest, which does not have to be the one with lowest eigenvalue. Our algorithm can be executed in an explorative fashion as we can circumvent the NEB or FSM calculation by starting from only one minimum-energy structure and by following several eigenvectors in one optimization in parallel. We will demonstrate these capabilities at the example of the isomerization reactions of formaldehyde, which has been studied as a benchmark system for automated transition-state search algorithms [168–171]. Subsequently, we investigate an internal proton-transfer reaction in a hydrazine complex, which is an intermediate in the Schrock N_2-fixation catalytic cycle [172–174]. At this example, we examine the applicability of our algorithm for finding TSs for large molecules (the Schrock catalyst contains 284 atoms). Although smaller model complexes can be generated, the smallest ones, which resemble the structure of the original catalyst, still comprise 41 atoms.

We choose these examples to highlight the capabilities of the Mode-Tracking-based approach to TS searches, which improves on existing methods rather than proposing a new TS search algorithm. Hence, validating the performance of our Mode-Tracking version of existing TS search algorithms at standard TS test sets [175–178] is neither needed nor necessary.

This Chapter is organized as follows: In the next section, the Mode-Tracking algorithm and the theoretical background of transition-state optimizations are described. After the subsequent Computational Methodology section, results are reported for our benchmark reactions.

The content of this Chapter has been published in Ref. [24].

5.1 Theoretical Foundations of MTsearch

The main idea of the algorithm to be described is to find transition-state structures by following only certain eigenvectors of the Hessian matrix selectively calculated by Mode-Tracking. For stationary structures, the harmonic vibrational frequencies and the corresponding eigenvectors of a system can be obtained by solving the following eigensystem,

$$\mathbf{H}\mathbf{Q}_k = \lambda_k \mathbf{Q}_k, \tag{5.1}$$

where $\mathbf{H}$ is the mass-weighted Cartesian Hessian, λ_k are the eigenvalues, and the eigenvectors $\mathbf{Q}_k$ are the mass-weighted vibrational normal modes. For non-stationary structures, for which the length of the geometry gradient is nonzero, Eq. (5.1) cannot be related to the vibrational properties of a molecule, but the eigenpairs (eigenvalues and eigenvectors) of $\mathbf{H}$ still characterize the PES.

In Mode-Tracking, the eigenpairs of interest are obtained through a Davidson-type subspace iteration method [156, 159], in which the original Hessian matrix $\mathbf{H}$ is transformed to the (reduced-dimensional) Davidson matrix $\tilde{\mathbf{H}}^i$,

$$\tilde{\mathbf{H}}^i = (\mathbf{B}^i)^T \mathbf{H}\mathbf{B}^i \equiv (\mathbf{B}^i)^T \Sigma^i, \tag{5.2}$$

where i denotes the i-th iteration step. $\mathbf{B}^i$ is a matrix whose columns contain collective displacement vectors $\mathbf{b}^l$ ($l = 1,...,i$) along the $3M$ (mass-weighted) nuclear Cartesian basis vectors (M is the number of atoms). In our semi-numerical implementation, Σ^i contains all vectors $\boldsymbol{\sigma}^l$, which collect the (numerical) derivatives of the (analytical) Cartesian gradient $\mathbf{g}$ of the total electronic energy with respect to the corresponding collective displacement vector $\mathbf{b}^l$,

$$\boldsymbol{\sigma}^l = \mathbf{H}\mathbf{b}^l = \begin{pmatrix} \sum_n H_{1,n} b_n^l \\ \sum_n H_{2,n} b_n^l \\ \cdots \\ \sum_n H_{3M,n} b_n^l \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \mathbf{b}^l} \frac{\partial E_{\mathrm{el}}}{\partial R_1} \\ \frac{\partial}{\partial \mathbf{b}^l} \frac{\partial E_{\mathrm{el}}}{\partial R_2} \\ \cdots \\ \frac{\partial}{\partial \mathbf{b}^l} \frac{\partial E_{\mathrm{el}}}{\partial R_{3M}} \end{pmatrix} = \frac{\partial}{\partial \mathbf{b}^l} \mathbf{g}. \tag{5.3}$$

By solving

$$\tilde{\mathbf{H}}^i \mathbf{c}_k^i = \lambda_k^i \mathbf{c}_k^i, \tag{5.4}$$

for the eigenvectors $\mathbf{c}_k^i$ and eigenvalues λ_k^i. In the i-th iteration step, Mode-Tracking calculates the approximate k-th normal mode $\mathbf{Q}_k^i$ as

$$\mathbf{Q}_k^i = \sum_{l=1}^{i} c_{k,l}^i \mathbf{b}^l. \tag{5.5}$$

New basis vectors $\mathbf{b}^{i+1}$ are generated from the residuum vector,

$$\mathbf{r}_k^i = [\tilde{\mathbf{H}}^i - \lambda_k^i]\mathbf{Q}_k^i, \tag{5.6}$$

after applying a preconditioner $\mathbf{X}^i$ to it [159],

$$\mathbf{b}^{i+1} = \mathbf{X}^i \mathbf{r}_k^i. \tag{5.7}$$

The initial guess mode $\mathbf{b}^1$ can be obtained from the LST, which linearly interpolates between the reactant and product structures, or from other path methods such as NEB. Let $\mathbf{R}_j^{(\mathrm{nmw})}$ be the non-mass-weighted '(nmw)' Cartesian coordinates of a structure j on the PES, then the initial normalized, non-mass-weighted mode is constructed from the coordinate differences between this point and each of its neighboring points, $\mathbf{R}_{j+1}^{(\mathrm{nmw})} = \{R_{k,j+1}^{(\mathrm{nmw})}\}$ and $\mathbf{R}_{j-1}^{(\mathrm{nmw})} = \{R_{k,j-1}^{(\mathrm{nmw})}\}$,

$$\mathbf{b}_j^{(\mathrm{nmw}),1} = \frac{1}{2}\left[\frac{\left(\mathbf{R}_{j+1}^{(\mathrm{nmw})} - \mathbf{R}_j^{(\mathrm{nmw})}\right)}{\left|\mathbf{R}_{j+1}^{(\mathrm{nmw})} - \mathbf{R}_j^{(\mathrm{nmw})}\right|} + \frac{\left(\mathbf{R}_j^{(\mathrm{nmw})} - \mathbf{R}_{j-1}^{(\mathrm{nmw})}\right)}{\left|\mathbf{R}_j^{(\mathrm{nmw})} - \mathbf{R}_{j-1}^{(\mathrm{nmw})}\right|}\right] = \left\{b_{k,j}^{(\mathrm{nmw})}\right\}, \quad k = 1, \ldots, 3M. \tag{5.8}$$

This mode is then mass-weighted,

$$\mathbf{b}_j^1 = \left\{\frac{b_{k,j}^{(\mathrm{nmw})}\sqrt{m_k}}{\sum_{k=1}^{3M}\left(b_{k,j}^{(\mathrm{nmw})}\sqrt{m_k}\right)^2}\right\}, \qquad k = 1, \ldots, 3M, \tag{5.9}$$

where m_k is the mass of the k-th atomic nucleus.

In general, Mode-Tracking can either optimize the mode with largest overlap with the initial guess vector or the one with largest overlap with the approximate eigenvector chosen in the last iteration (root-homing). If the initial guess vector differs strongly from the normal mode of the transition-state structure, the second option might be more suited to find a TS.

The eigenvector following algorithm [144] is then employed to steer the optimization into the direction of the TS and to finally locate it. Newton–Raphson steps along the converged Mode-Tracking eigenvector, which is referred to as transition vector, are carried out to maximize the energy in this direction, while in all directions orthogonal to the transition direction the structure is relaxed [179]. For this, we project out the gradient along the transition vector, $\mathbf{g}_{\mathrm{TS}}$, from the total molecular gradient,

$$\mathbf{g} = \{g_k\} = \left\{\frac{\partial E}{\partial R_k}\right\}, \qquad k = 1, \ldots, 3M. \tag{5.10}$$

To obtain the components of the molecular gradient that are orthogonal to the eigenvector, $\mathbf{g}_{\mathrm{ort}}$, we subtract the gradient part along the transition vector from the original molecular gradient and obtain

$$\mathbf{g}_{\mathrm{ort}}^{(\mathrm{nmw})} = \mathbf{g}^{(\mathrm{nmw})} - \underbrace{\mathbf{Q}_{\mathrm{TS}}^{(\mathrm{nmw})}\mathbf{Q}_{\mathrm{TS}}^{(\mathrm{nmw}),T}\mathbf{g}^{(\mathrm{nmw})}}_{\mathbf{g}_{\mathrm{TS}}^{(\mathrm{nmw})}} = (1 - \mathbf{Q}_{\mathrm{TS}}^{(\mathrm{nmw})}\mathbf{Q}_{\mathrm{TS}}^{(\mathrm{nmw}),T})\mathbf{g}^{(\mathrm{nmw})}, \tag{5.11}$$

where $\mathbf{Q}_{\mathrm{TS}}^{(\mathrm{nmw})}$ is the selected eigenvector calculated with Mode-Tracking, which approximately points into the direction of the TS. This is done in no-mass-weighted coordinates. The corresponding eigenvalue is λ_{TS}.

Let $\mathbf{R}_0$ be the coordinates of the targeted stationary point, for which $\mathbf{g}_0 \equiv \{(\partial E/\partial R_k)_{R_k=R_{0,k}}\}$ vanishes component-wise, and the Hessian is described by $\mathbf{H}_0 \equiv \{(\partial^2 E/\partial R_k \partial R_l)_{R_k=R_{0,k},R_l=R_{0,l}}\}$. From a truncated Taylor series expansion of the potential energy around $E_0 = E(\mathbf{R}_0)$ on the PES,

$$E(\mathbf{R}) = E_0 + \mathbf{g}_0^T\mathbf{R} + \frac{1}{2}\mathbf{R}^T\mathbf{H}_0\mathbf{R} + O(\mathbf{R}^3). \tag{5.12}$$

the coordinate displacement $\mathbf{R} \equiv \mathbf{R} - \mathbf{R}_0$ that leads to a stationary point ($dE(\mathbf{R})/d\mathbf{R} = \mathbf{0}$),

$$\mathbf{R} = -\frac{\mathbf{g}_0}{\mathbf{H}_0}, \tag{5.13}$$

can be derived. $\mathbf{R}_0$ is the position of a stationary structure, $\mathbf{g}_0$ its gradient and $\mathbf{H}_0$ its Hessian. $\mathbf{R}$ can be split into a direction parallel to the transition vector, $\mathbf{R}_{\mathrm{TS}}$, and into all other directions. The step in the direction of the transition vector reads

$$\mathbf{R}_{\mathrm{TS}} = -\frac{\mathbf{g}_{\mathrm{TS}}}{\lambda_{\mathrm{TS}}}. \tag{5.14}$$

The energy in direction of the selected mode is maximized if λ_{TS} is negative. If we do not start the EVF procedure from a structure close to the TS, but, for instance, from a minimum structure, we must ensure that the transition vector is still followed uphill. This can either be accomplished by employing the absolute value of λ_{TS}

$$\mathbf{R}_{\text{TS}} = \frac{\mathbf{g}_{\text{TS}}}{|\lambda_{\text{TS}}|}, \tag{5.15}$$

or by employing Eq. (5.20) described below.

To improve on the convergence of the EVF optimization, Wales [179] defined a Lagrangian with Lagrangian multipliers κ_k for each degree of freedom l:

$$L = -E_0 - \mathbf{g}_0^T \Delta\mathbf{R} - \frac{1}{2}\Delta\mathbf{R}^T \mathbf{H}_0 \Delta\mathbf{R} + \frac{1}{2}\sum_{l=1}^{3M} \kappa_l \left(\Delta R_l^2 - c_l^2\right). \tag{5.16}$$

Wales’ method employs the rational function by Banerjee [141, 180, 181], in which the Lagrangian multipliers are defined by the eigenvalues λ_k and the gradient components $\mathbf{g}_k$ along the eigenvectors,

$$\kappa_l = \frac{1}{2}\left(\lambda_l \pm \sqrt{\lambda_l^2 + 4\mathbf{g}_l^2}\right). \tag{5.17}$$

It appears to be more efficient [182] to modify the equation to the following one:

$$\kappa_k = \lambda_l \pm \frac{1}{2}|\lambda_l|\left(1 + \sqrt{1 + \frac{4\mathbf{g}_l^2}{\lambda_l^2}}\right) \tag{5.18}$$

where ’+’ is for maximization and ’–’ for minimization.

Wales arrived at the following equation that describes the steps to be made along all degrees of freedom l,

$$\Delta\mathbf{R}_l = \frac{\pm 2\mathbf{g}_l}{|\lambda_l|\left(1 + \sqrt{1 + \frac{4\mathbf{g}_l^2}{\lambda_l^2}}\right)}, \tag{5.19}$$

where ’+’ leads to an uphill and ’–’ to a downhill energy step. For a TS search, an uphill step along the desired mode (i.e., the approximate transition vector) is required,

$$\Delta\mathbf{R}_{\mathrm{TS}} = \frac{+2\mathbf{g}_{\mathrm{TS}}}{|\lambda_{\mathrm{TS}}|\left(1+\sqrt{1+\frac{4\mathbf{g}_{\mathrm{TS}}^2}{\lambda_{\mathrm{TS}}^2}}\right)}. \tag{5.20}$$

5.2 The MTsearch program

We implemented the theory presented in the previous section in a computer program called MTsearch. The program is based on the original Mode-Tracking program [156, 159, 183], which is currently available in its latest release as part of the MoViPac package [184]. MTsearch is a parallelized meta-program that accesses standard quantum-chemical programs for the calculation of gradients and electronic energies. The computational methodology for the generation of these raw data is described in detail in the next subsection. The algorithmic structure of MTsearch is sketched in Figure 5.1.

A set-up tool, called tsdefine, creates the necessary input files for an MTsearch calculation. With tsdefine we read in initial guess structures and, if available, initial modes. The initial guess modes and structures can either be created within MTsearch, from a LST or an NEB path, or by an external program, which provides guess structures and modes, e.g., based on a constrained optimization scan. The LST or NEB path consists of six to twelve nodes (that is, molecular structures on an (approximate) reaction path, including reactants and products), which we found to be a reasonable number. The spring forces in an NEB calculation are set to 0.02 a.u., and such a calculation is considered converged when the difference between the gradient norm of the actual iteration and the previous one drops below 1×10^{-3} a.u.

The first step of the TS optimization procedure is the Mode-Tracking optimization of the initial guess mode to produce the corresponding minimal mode. Mode-Tracking is assumed to have converged when the maximum element of the residuum vector drops below 5×10^{-3} a.u. and the change in the length of the residuum vector drops below 5×10^{-6} a.u. One may also choose the convergence criteria corresponding to the last-added basis vector contribution to the selected eigenvector or to the change in the eigenvalue between the last iterations.

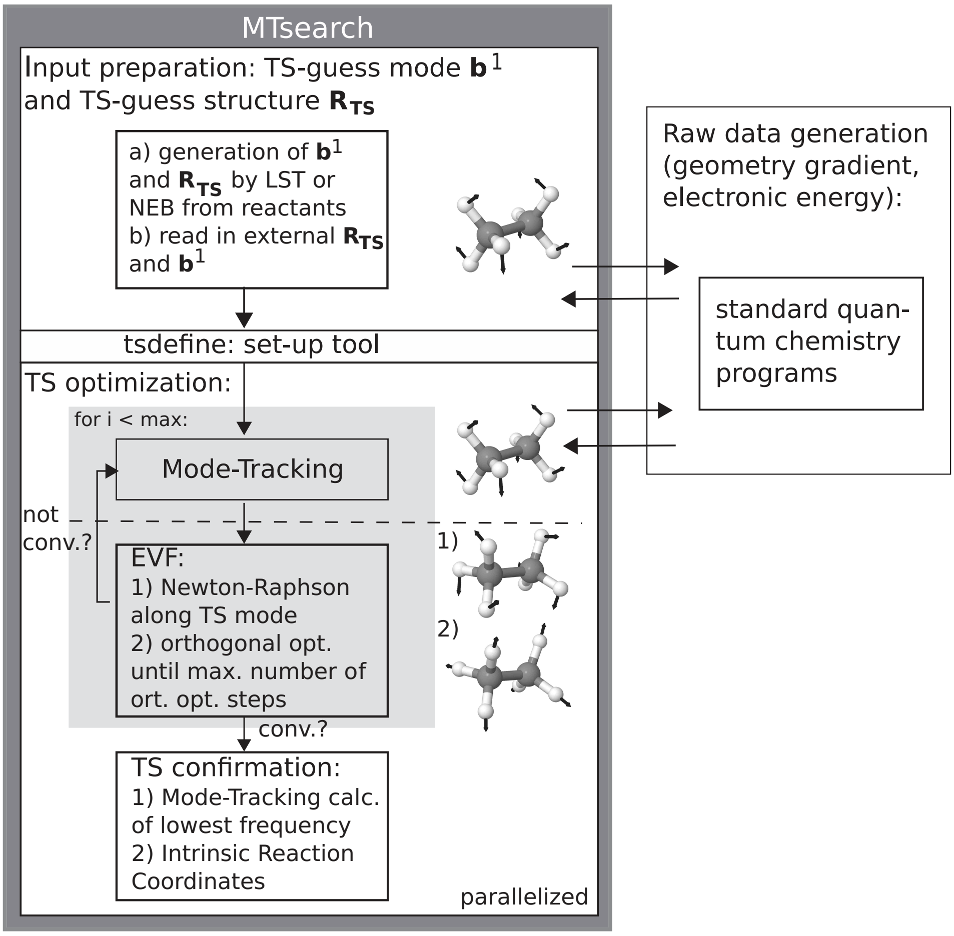

MTsearch
Input preparation: TS-guess mode $\mathbf{b}^1$ and TS-guess structure $\mathbf{R_{TS}}$
a) generation of $\mathbf{b}^1$ and $\mathbf{R_{TS}}$ by LST or NEB from reactants
b) read in external $\mathbf{R_{TS}}$ and $\mathbf{b}^1$
tsdefine: set-up tool
TS optimization:
for i < max:
Mode-Tracking
not conv.?
EVF:
1) Newton-Raphson along TS mode
2) orthogonal opt. until max. number of ort. opt. steps
1)
2)
conv.?
TS confirmation:
1) Mode-Tracking calc. of lowest frequency
2) Intrinsic Reaction Coordinates
parallelized
Raw data generation (geometry gradient, electronic energy):
standard quantum chemistry programs

Figure 5.1: Overview of the MTsearch meta-program structure.

After the calculation of a specific mode with Mode-Tracking, an EVF step is performed based on this converged mode. For the Newton–Raphson step along the transition vector, $\mathbf{R}_{\text{TS}}$, we define a maximum step size of 0.2 Å/$\sqrt{\text{amu}}$, which is decreased to 0.1 Å/$\sqrt{\text{amu}}$ when the norm of $\mathbf{g}_{\text{TS}}$ drops below $3{\times}10^{-2}$ a.u., and to 0.05 Å/$\sqrt{\text{amu}}$ when the norm of $\mathbf{g}_{\text{TS}}$ drops below $1{\times}10^{-2}$ a.u. For TS searches starting from minimum-energy structures, the first four Newton–Raphson steps are set to a maximum length of 1.0 Å/$\sqrt{\text{amu}}$, whenever the Hessian eigenvalue is positive or close to zero, i.e., no imaginary frequency with a large absolute magnitude is obtained.

After each Newton–Raphson step a predefined number of optimization steps orthogonal to the eigenvector is performed. As default, a maximum of three iterations is chosen, if not otherwise mentioned.

If the norm of the total gradient is still above the threshold (default is $1{\times}10^{-3}$ a.u.) after the predefined number of orthogonal optimization steps, another Mode-Tracking calculation is launched, for which the last converged Mode-Tracking eigenvector is chosen as default guess vector. By default, a root-homing scheme selects the eigenvectors during the Mode-Tracking calculation with respect to the largest overlap with the initial one. For comparison, we also employed a root-homing scheme in which the eigenvector is always compared to the previous one.

It would also be possible to reuse the same eigenvector for a predefined number of EVF steps, and/or to perform more than one EVF step between orthogonal optimizations. This has not been explored in this study. It is also possible to supply more information about the transition path direction to MTSEARCH than only the first eigenvector (e.g., a sequence of structures which can for example easily be generated by a haptic device [185–187]). The guess vectors for the first few Mode-Tracking calculations are then chosen according to the predefined sequence of structures. It has to be specified how many times the initial guess structure path shall be taken as reference for creating a guess mode, which is then refined by Mode-Tracking.

The structures and normal modes were visualized with PYMOL [188] and JMOL [189], respectively.

5.3 Raw data generation

All energies and gradients which are read as raw data by MTSEARCH were calculated with density functional theory employing the TURBOMOLE program package (version 6.3.1) [190] with Ahlrichs' def2-SV(P), def2-SVP and def2-TZVP basis sets [129]. MTSEARCH launches these calculations by system calls. Restricted and unrestricted BP86 [133, 134] all-electron Kohn–Sham calculations in combination with the resolution-of-the-identity technique were carried out. Self-consistent-field single-point calculations are considered to be converged when the total electronic energy difference between two iteration steps is less then 10^{-7} Hartree, if not otherwise indicated. Molecular structure minimizations are considered converged when the norm of the geometry gradient is below 10^{-4} a.u. For the optimization of transition-state structures a geometry-gradient threshold of 10^{-3} a.u. is chosen.

5.4 Reference calculations

For comparison, we performed Turbomole (version 6.3.1) [190] EVF calculations for comparison with the MTsearch results. Starting structures were chosen from the LST, NEB, or constrained optimization paths. We carry out a single-point calculation on the starting structure and continue with a calculation of all vibrational modes with Turbomole. Then, we employ the trust-radius imaging method (the maximum radius and the trust radius are chosen between 0.1 Å/$\sqrt{\text{amu}}$ and 0.2 Å/$\sqrt{\text{amu}}$) to follow the lowest eigenvalue. We refer to this procedure in the following as "standard EVF method". In these Turbomole calculations, the BP86 [133, 134] density functional is chosen with Ahlrichs' def2-SV(P), def2-SVP and def2-TZVP basis sets [129].

Furthermore, we performed constrained optimizations by employing the Gaussian [191] program (version 09, Revision C.1) to obtain transition-state guess structures. Essentially, one internal coordinate was kept fixed at defined values and for all other degrees of freedom a structure optimization was carried out. Furthermore, intrinsic reaction coordinates were calculated with Gaussian. In these calculations we employed BP86 with the def2-SVP basis set [192, 193]. We have chosen the default convergence criteria (scfconv=tight, which means that the energy difference between two SCF iterations was less than 10^{-8} Hartree, and that the structure optimizations were considered converged when the root-mean-square force acting on all atoms was below 3×10^{-4} a.u.).

We should note that we provide data for the eigenvalues of the Hessian as 'frequencies' (reported in units of wave numbers). I.e., we take the square root of the eigenvalues, which corresponds to a harmonic vibrational frequency for a stationary structure, even for *non*-stationary structures and denote it a 'frequency' for the sake of convenience (eventually, these data become harmonic frequencies upon convergence of the stationary-structure optimization). Moreover, to highlight imaginary frequencies, we add a minus sign in front of them (this is possible as the square of such a frequency still yields the correct eigenvalue of the Hessian matrix). Note also that we use the term ’mode’ to denote an eigenvector of the Hessian matrix.

5.5 Test systems for MTsearch

To study the capabilities of MTsearch, we have chosen four intramolecular reactions involving molecules of different sizes (4 atoms, 8 atoms, 41 atoms, and 284 atoms; shown in Figure 5.2).

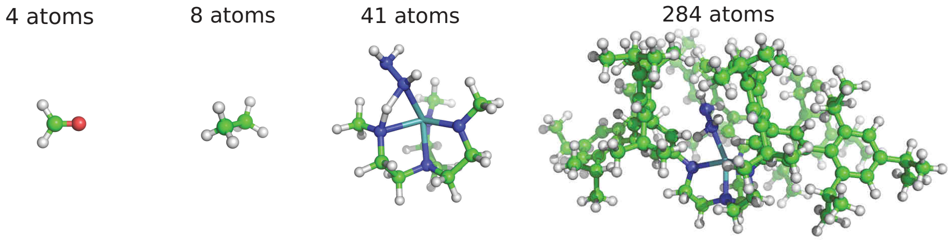


Figure 5.2: Molecular models considered in this work: H_2CO (left), ethane (second from left), a small (third from left) and the full Schrock hydrazine tris(amido)amine Mo complex (right). Element color code: green, C; red, O; blue, N; cyan, Mo; white, H.

We start with the investigation of the rotational barrier in the C_2H_6 molecule, because the transition-state structure is well defined and the system is small, which allows us to investigate the suitable settings and thresholds for MTsearch. Next, we analyze the possibility of MTsearch to optimize several transition-state structures starting from one minimum-energy structure using hydroxymethylene as an example.

The last two reactions considered are possible side reactions of the Chatt–Schrock cycle of N_2 fixation at a molybdenum containing catalyst [172, 173], in which N_2 is reduced to ammonia under acidic and reductive conditions. Under these reaction conditions, it is possible that several unwanted intermediates are formed. Exemplarily, we have chosen one possible side-reaction pathway, where one proton of N_2H_4 coordinated to molybdenum shifts to one of the amido nitrogens.

The Schrock catalyst is ligated by a tetradentate hiptN$_3$N ligand (hipt = hexa-*iso*-propyl terphenyl). It has been intensively studied both experimentally [173, 174, 194] and theoretically [160, 194–198]. Because of the relatively large system size of the hiptN_3N ligated hydrazine molybdenum complex (278 atoms), several smaller generic model system of the catalyst, in which the aryl substituents have been substituted, e.g., by H atoms or CH_3 groups, have been studied [194, 195, 197, 199–202]. In the following, the small and large Schrock catalyst refer to the $MeNCH_2CH_3N$ or hipt ligated molybdenum catalyst, respectively, with a hydrazine ligand as shown in Figure 5.2. Our focus is the optimization of transition-state structures for the proton-transfer reaction in the small and large Schrock catalysts.

5.5.1 Benchmark Example: C_2H_6 rotation

We first calculated the transition-state structure of ethane rotation from staggered conformation to eclipsed conformation and back to staggered conformation. To obtain starting structures and initial guesses for the Mode-Tracking scheme, we performed a linear synchronous transit in internal coordinates with six nodes on the path including reactant and product structures (both staggered). Since only one dihedral angle is changed during the transition from one minimum structure to the other, the choice of internal coordinates is very useful in this example. Due to the symmetry in the LST path, only three of the six structures are different. The first (minimum), second, and third structures of the LST pathway were chosen as starting structures for EVF procedures performed with MTsearch and, for comparison, Turbomole.

Although the second structure of the LST pathway does not feature negative eigenvalues of the Hessian, the EVF algorithm optimization started from these structures converged towards the transition-state structure with both Turbomole and MTsearch (see Appendix E for details).

If we start from the energy minimum structure, EVF relying on one negative eigenvalue is not able to find the TS, because the structure is far away from the quadratic region around the TS. Therefore, one would usually start from a guess structure closer to the TS. By contrast, MTsearch locates the TS starting from the energy-minimum structure with a mode specified by the LST and by a manually chosen mode corresponding to the rotation of one CH_3 group around the C–C axis. The initial-guess structures and converged TSs can be found in Appendix E.

We analyzed the effect of various parameters on the convergence of MTsearch. First, we investigated the optimal length of the first Newton–

Raphson step. If the starting structure is close to the energy minimum structure, the optimizer has to accomplish a larger step out of the minimum. We observed that a first Newton–Raphson step size of 1.0–1.5 $Å/\sqrt{\text{amu}}$ is appropriate (see Appendix E for details).

Next, we adapted the number of orthogonal optimization steps performed until the next mode is optimized by Mode-Tracking to values between 2 and 10. To generalize the algorithm, we defined a protocol which stops the orthogonal optimization if the norm of the gradient for the optimization orthogonal to the transition path drops below $1{\times}10^{-3}$ a.u., which means that the maximum number of orthogonal-optimization steps needs not to be reached. Then, the next Mode-Tracking calculation and Newton–Raphson step in the direction of the converged transition vector is performed.

5.5.2 Explorative Example: Isomerization of H_2CO

In this section, we study the possibility to find several TSs with MTsearch starting from one minimum-energy structure only. We have chosen the isomerization reaction of formaldehyde to hydroxymethylene and a subsequent trans-/cis-isomerization of hydroxymethylene as an example (see Figure 5.3). The transition-state structures are well known [168–170, 203]. Since two reaction pathways are possible from trans-hydroxymethylene, a selective way of choosing the eigenvector of interest is important.

The trans-hydroxymethylene structure, from which we start the explorative TS search, can either undergo an internal hydrogen transfer from the oxygen atom to the carbon atom (over **TS-1**) that leads to formaldehyde or a rotation around the C-O axis that leads to cis-hydroxymethylene (over **TS-2**). MTsearch is able to locate both transition-state structures by following the modes which are shown in Figure 5.3 next to the arrows indicating the reaction direction. The three lowest modes of the starting structure (obtained by a full Hessian calculation) have the following frequencies: 1100 cm^{-1}, 1188 cm^{-1}, and 1317 cm^{-1}. The first one leads to **TS-2** and the second and third ones to **TS-1**. MTsearch can find **TS-1** and **TS-2** also by starting from guess modes which are based on chemical intuition (see Appendix E).

The standard EVF optimization from hydroxymethylene following the lowest vibrational frequency mode does not converge to a TS, but falls back to the minimum-energy structure. Already a minor distortion

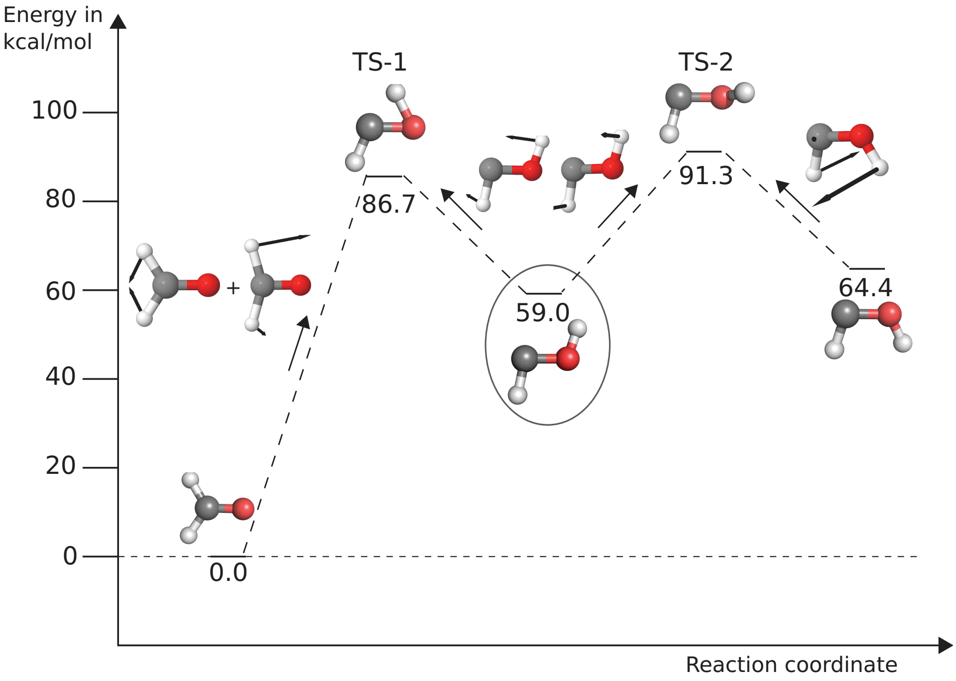


Figure 5.3: BP86/RI/def2-SV(P) reaction path from formaldehyde to trans-hydroxymethylene and cis-hydroxymethylene with transition-state structures optimized with MTsearch. The modes taken from a full vibrational analysis that lead to the TSs, **TS-1** and **TS-2**, are also depicted. A maximum number of three orthogonal optimization steps has been chosen in MTsearch. Element color code: gray, C; red, O; white, H.

of the minimum-energy structure towards the TS can already lead to a successful location of **TS-2** (see Appendix E for details).

Besides the TS optimization starting from trans-hydroxymethylene, we have also carried out TS localizations from formaldehyde and cis-hydroxymethylene. For cis-hydroxymethylene, **TS-2** was found by following the LST guess mode. For formaldehyde, neither a LST guess mode nor a guess mode based on chemical intuition led to convergence to **TS-1**.

5.5.3 Intramolecular proton-transfer reaction in a hydrazine Mo complex

In this section, we study the hydrazine intermediate of Schrock's nitrogen-reducing catalyst and a generic model complex with aryl groups substituted by methyl groups. For the Schrock hydrazine complex, the lowest-energy spin state is a doublet. All spin states with higher multiplicity are at least 60 kJ/mol higher in energy [197]. We investigated the transition-state structure for a proton shift reaction from the nitrogen atom of N_2H_4 that ligates to molybdenum to one of the amido nitrogen atoms.

For the generic model complex, we carried out a constrained optimization scan along the N_{amido}-H distance that changes from reactant to product in 12 steps including reactants and products of 0.2 Å step size to obtain a guess structure close to the TS. From this constrained optimization scan, we selected the highest-energy structure (**Scan12**, see Figure 5.4) and an initial guess mode based on the structures **Scan11** and the product. This mode has been simplified by retaining only those entries that refer to the transferring hydrogen atom. This restriction to the "moving" part in the system improves convergence as other motions of parts of the system are discarded. Moreover, it produces a guess mode that is transferable between homologous species (see the large complex below).

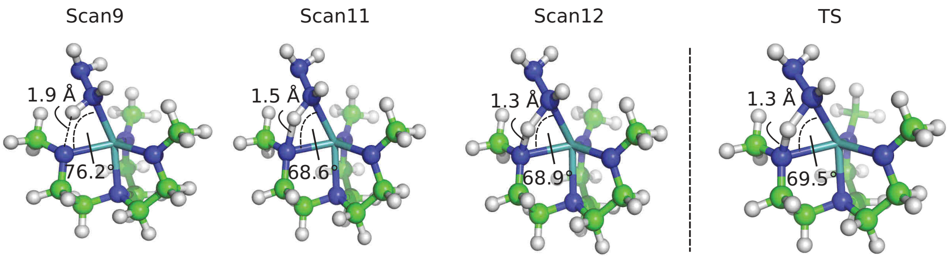


Figure 5.4: Initial guess structures chosen from a constrained optimization scan along one H(N_2H_4)-N_{amido} distance. **Scan9**, **Scan11** and **Scan12** are the 9th, 11th and 12th structure from a 12-step constrained optimization scan with the program GAUSSIAN (0.2 Å increase of the N_{amido}-H distance in each step) along the N_{amido}-H distance starting from the hydrazine bound complex.

The frequency analysis of structures **Scan11** and **Scan9** revealed that the lowest eigenvalue modes do not correspond to the desired transition vector. We observed that the standard EVF algorithm often fails to find the TS in this situation (see Appendix E for details). Only for structure **Scan12**, which is already very close to the TS, the standard EVF optimization converges towards the TS. By contrast, MTsearch was able to find the TS also from **Scan11** and **Scan9** (see Appendix E).

The root-mean-square deviation between the MTsearch-optimized transition-state structures and the one calculated with Turbomole's EVF is only 0.04 Å, which means that the two algorithms converged to the same structure. The vibrational analysis revealed exactly one imaginary frequency of -i1244 cm^{-1}, and the intrinsic reaction coordinates (IRC) connect the reactant and product structures, which confirms that we found the desired TS. The stationary points calculated by MTsearch are shown in Figure 5.5.

It is noteworthy that the initial guess modes cannot only be obtained from a constrained scan, but also from a LST or NEB pathway or based on chemical intuition. For isomerization reactions, in which only one atom re-positions, as in our example, it is straightforward to manually choose an approximate transition pathway (cf. Appendix E). However, the manual set-up of a proper molecular distortion that is likely to resemble a reaction pathway is possible and potentially useful also for other types of reactions.

In a Mode-Tracking-based TS search, one should confirm whether the very first mode converged with Mode-Tracking corresponds to the desired reaction pathway, since all following optimization steps are based on the direction of this initial mode. However, this can be done automatically by calculating the overlap of the initial guess mode and the converged one. If the initial mode is not close to the transition vector, the optimization may lead to a different TS than the desired one (in our case, the TS for a rotation of the terminal NH_2 moiety of the N_2H_4 ligand was often found when the initial guess mode was not clearly dominated by the shifting proton).

For the optimization of the analogous transition-state structure in the significantly larger hiptN_3N-ligated Schrock Mo catalyst, we rotated the coordinating N_2H_4 ligand in the minimum-energy structure such that an initial guess structure comparable to the TS of the generic model complex is obtained. Then, we performed a constrained optimization (fixed atoms are: molybdenum, the proton that moves and the two nitrogen atoms to which the proton binds in the reactant structures). We choose as an initial guess mode the converged mode of the TS in the generic model complex (after alignment of the large and small homologous complexes, and choosing only those entries for atoms that occur in both complexes; all other entries are set to zero). Due to the significantly larger system size, the maximum number of orthogonal optimization steps performed within MTsearch is increased to 10. The stationary points with the TSs calculated by MTsearch are displayed in Figure 5.5.

With Mode-Tracking we obtain one imaginary frequency of -i1323 cm^{-1} for the TS, which is similar to the one of the TS in the small model system. The mode is located on the proton that shifts. To study the performance of our algorithm, we also calculated the full Hessian and obtained one negative frequency mode of -i1338 cm^{-1}.

Since the complete Hessian calculation and diagonalization for this molecule consisting of 284 atoms takes significantly longer than the Mode-Tracking calculation (about a week vs. 2 hours on 12 cores on a blade system featuring two six-core AMD Opteron 2435 processors (i.e., a total of 12 cores)), the computational time needed for the TS search is considerably reduced.

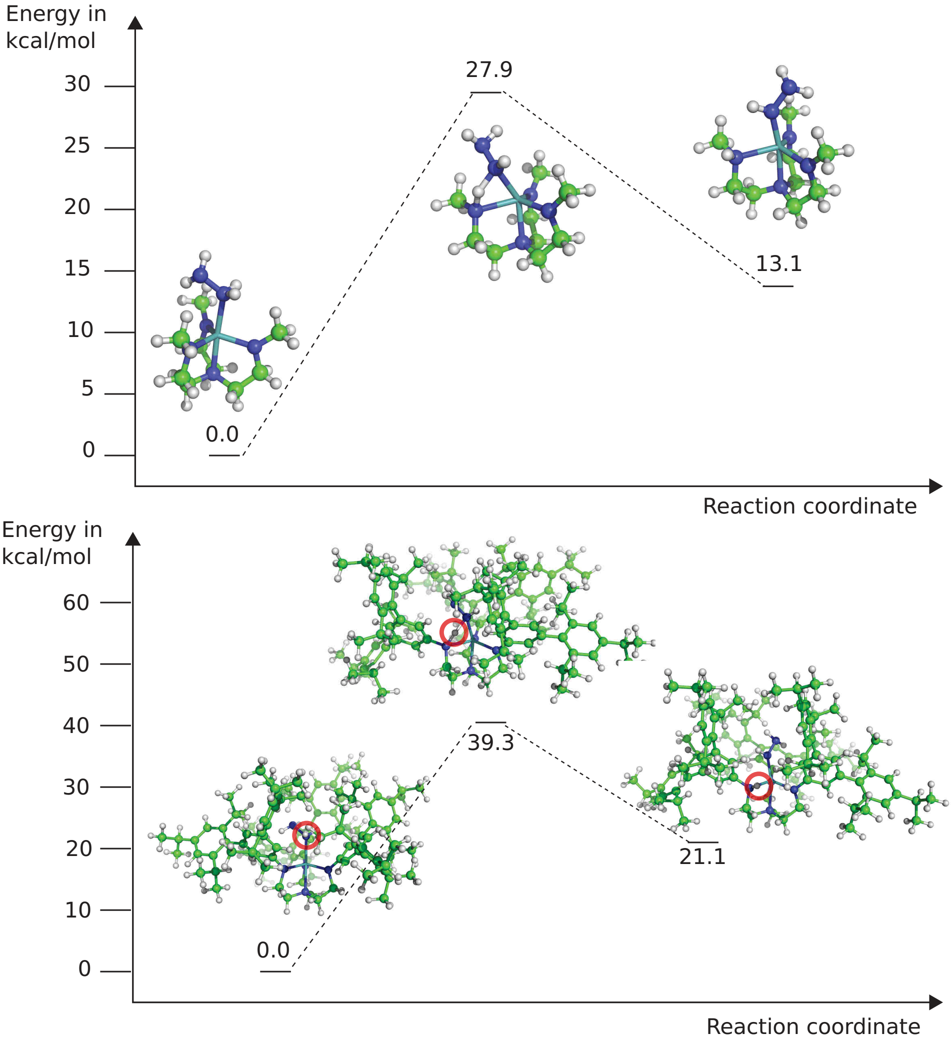


Figure 5.5: Internal proton shift reaction pathway of the generic Schrock model system (top) and the full Schrock hydrazine Mo complex (bottom) with BP86/RI/def2-SV(P) transition-state structures optimized with MTSEARCH. Element color code: green, C; blue, N; cyan, Mo; white, H. The red circles highlight the proton that shifts during the reaction.

6

Heuristic Reaction Network Explorations

Several types of chemical reactions are known, in which a variety of side products can occur. In these cases, either a reactive species (such as a radical, a valence-unsaturated species, a charged particle, a strong acid, or a strong base) is involved or the energy deposited into the system is high (e.g. due to a high reaction temperature). Highly reactive species lead to a large variety of products, because such reactions may proceed in an unselective manner. High thermal energy increases the number of side products, because parts of the reaction network that are usually not reachable (e.g. at room temperature) become accessible.

Reaction networks map out sequences of elementary reaction steps, in which two valleys on the potential energy surface (PES) are connected by one transition-state structure (TS). For a detailed analysis of the kinetics and thermodynamics of a net reaction energy resulting from many elementary steps, a complete network of potentially accessible intermediates is required. Only then can the global lowest-energy paths be identified and the probabilities to reach certain minimum-energy structures be compared with each other. The number of stationary points in such reaction networks can be estimated along the lines of the following example: For protonation reactions, we assume that the number of protonated species can be determined by identifying all potential protonation sites within the initial structure to which protons can be added. For a reaction network in which all these possible protonated intermediates are considered, the number of possible minimum-energy structures is given by

$$N = \sum_{p=1}^{N_{\mathrm{PS}}} \binom{N_{\mathrm{PS}}}{p}, \tag{6.1}$$

where N_{PS} is the number of possible protonation sites and p the number of protons. In principle, one can and should also add more than one type of reactive species to a substrate. The total number of protons added to the molecule varies between one and the number of possible acceptor sites. For example, for 11 protonation sites, $N = 2047$. The number of generated structures exceeds 1000 for substrates with more than 8 protonation sites. Clearly, the transfer of so many protons to a substrate is not very sensible from a chemical point of view as the charge increases dramatically and the $\mathrm{p}K_a$ value of the protonated intermediates might be unfavorable with respect to further protonation. However, for strong acids accompanied by strong reductants many of these multiply protonated species will become accessible, if also reduction steps are considered.

The number of possible reaction paths between N structures is bounded by the total number $N(N-1)/2$. Since most of the structure pairs are not connected by an elementary step, the number of reaction paths can be significantly reduced by introducing empirical rules for the identification of elementary steps. Nevertheless, the number of elementary steps can still be too large for a manual investigation and an automated optimization of the relevant minimum-energy structures and TSs is highly desirable.

A reaction network exploration algorithm should be able to identify all relevant intermediates of a chemical reaction without any additional information apart from basic input structures (i.e., reactant, product, or intermediate) and to generate the reaction network with all relative potential energies and activation barriers. Naturally, no reaction network exploration algorithm can guarantee the localization of all minimum-energy structures and TSs. Nevertheless, since the manual calculation of all elementary steps and the evaluation of the reaction paths is very time-consuming, algorithms that systematically screen the PES represent an important leap forward.

Automated reaction network exploration algorithms can be divided into two main categories. The first one contains algorithms that require only one starting structure (i.e., the minimum-energy structure) and evolve along the PES by local curvature information obtained at the starting structure. Algorithms that first search for a diverse range of minimum-energy structures and then optimize the TSs based on previously assigned minimum-energy structure pairs belong to the second category.

Algorithms of the first category usually employ information of the curvature of the PES to walk uphill and identify new minimum-energy structures, ideally after passing by TSs or their vicinities. One example is the global reaction route mapping (GRRM) algorithm based on the anharmonic downward distortion following method and the artificial force induced reaction method [171, 204, 205]. For large molecules, GRRM becomes computationally unfeasible, because of the computational cost for the PES exploration around a local minimum. Although one can apply a micro-iteration method within a mixed quantum mechanics and molecular mechanics framework [206] to partially overcome this problem, methods starting from only one minimum-energy structure remain time-consuming.

In contrast to the first category of automated reaction network generators, the second category utilizes experience-based transformation rules for the generation of new input structures and can subsequently make use of the robustness and computational efficiency of double-ended TS-search algorithms. This type of network exploration algorithms is referred to as a heuristic approach. Although one cannot ensure that all minimum-energy structures are obtained, heuristic approaches are very useful for chemical reactions for which transformation rules for the generation of possible new structures are known. Rappoport et al. [207] and also Zimmerman [155] employed such a heuristic approach for the generation of new structures according to transformation rules that are applicable to the main group elements of the periodic table. Atom connectivities are the basis for the generation of new bonds or breaking of existing ones. However, for intermediates and transition metal atoms, it may be difficult to define these atom connectivities.

Other approaches are based on molecular dynamics, Monte-Carlo methods, or the reaction path Hamiltonian method [208–211]. They mainly aim at the localization of the global minimum of a system, instead of all possible reaction pathways.

Here, we describe a reaction network generation algorithm of the second category combined with automated TS-search algorithms. Guess structures for potential intermediates are generated by formalizing heuristic principles. We implement an automated set-up of reactive complexes by placing a reactant in the vicinity of a substrate target. Specifically, we focus on two reactive species which are often involved in chemical reactions, namely protons and electrons. To identify all side reactions (if possible), all sites to which protons might be added are considered in a combinatorial way. These structures are then optimized. A subsequent assignment of neighboring structure pairs is carried out and between them a TS is optimized.

We chose to apply the algorithm to an important and still unsolved problem in chemistry: catalytic nitrogen fixation under ambient conditions at the example of the molybdenum complex developed by Schrock and coworkers [172, 173]. We map out the reaction network for the methyl-substituted model system with one- to four-fold protonation and reduction steps for the first half of the catalytic cycle, which is based on a N_2 moiety bound to the Mo center.

The content of this Chapter has been published in Ref. [25].

6.1 Structure of HeurEx

HeurEx is a parallelized meta-program written in the programming language Python [212] that accesses standard quantum-chemical programs for the calculation of electronic energies and for the localization of minimum-energy and TS structures.

An overview on the Heuristic Explorer (HeurEx) reaction network exploration program for the investigation of unspecific protonations and reductions of a substrate is given in Figure 6.1.

As an initialization, HeurEx aims at locating all relevant minimum-energy structures starting from reactive-complex structures. In our example, all proton-acceptor atoms in the substrate and subsequently the number of protonation sites at each proton-acceptor atom are identified first. Protonation sites are filled according to the maximum number of protons to be added (by default, this is equal to the number of protonation sites) by the “structure generator”.

In a next step, these structures are optimized by the “structure optimizer”, which drives standard quantum-chemical programs for this purpose (currently Turbomole and Gaussian are available). All optimized structures are then structurally aligned by the “elementary-step finder“. Redundant and decomposed structures are removed.

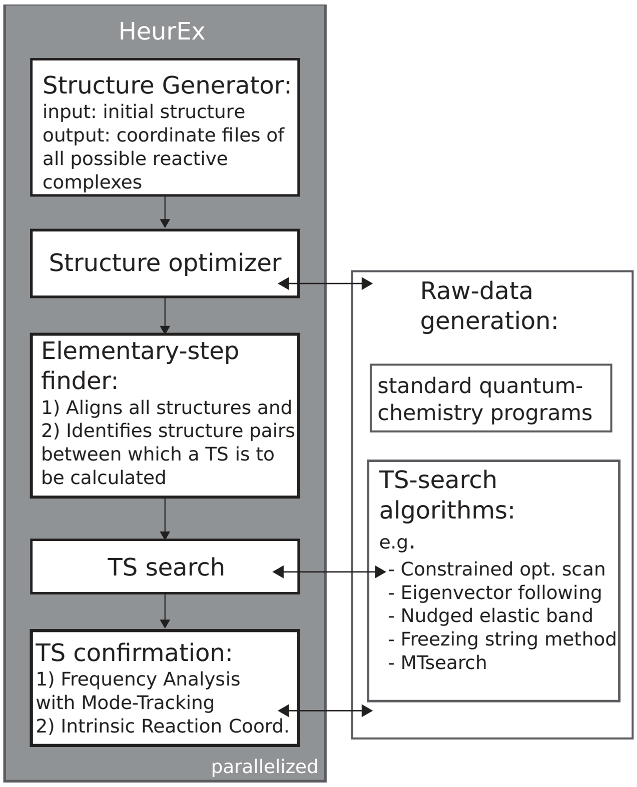

HeurEx
Structure Generator:
input: initial structure
output: coordinate files of all possible reactive complexes
Structure optimizer
Elementary-step finder:
1) Aligns all structures and
2) Identifies structure pairs between which a TS is to be calculated
TS search
TS confirmation:
1) Frequency Analysis with Mode-Tracking
2) Intrinsic Reaction Coord.
parallelized
Raw-data generation:
standard quantum-chemistry programs
TS-search algorithms:
e.g.
- Constrained opt. scan
- Eigenvector following
- Nudged elastic band
- Freezing string method
- MTsearch

Figure 6.1: Flow chart for the steps performed in the HeurEx reaction network exploration algorithm.

Noteworthy, we are only interested in structures that are accessible under the given reaction conditions. If a minimum-energy structure is too high in energy, it shall be discarded. The number of TS optimizations can hereby be significantly reduced. The probability distribution of an ensemble of structures over all possible minimum-energy structure states at a given temperature may be described by the Boltzmann distribution. Either an energy cutoff for minimum-energy structures, a temperature threshold, or a Boltzmann probability threshold can be chosen to restrict the reaction network to the relevant minimum-energy structures. In this work, we have chosen a temperature threshold of 298 K, which is approximately equal to an energy cutoff of 25 kcal/mol compared to the lowest-energy structure, because structures with an energy above the cutoff are not observed at room temperature (according to Ref. [213]):

For a unimolecular reaction assuming the (simplified) Eyring equation [6] the rate constant k reads,

$$k = \frac{k_B T}{h} \exp^{-\Delta G^{\ddagger}/RT}, \tag{6.2}$$

with Boltzmann constant k_B, Planck constant h, the activation (free) energy of the reaction $\Delta G^{\ddagger}$ and the temperature T. For the half life of a species ($t_{1/2} = \ln 2/k$) we then obtain the following equation,

$$t_{1/2} = \left(\frac{\ln 2h}{k_B T}\right) \exp^{\Delta G^{\ddagger}/RT} . \tag{6.3}$$

If $\Delta G^{\ddagger}$=100 kJ/mol (23.9 kcal/mol) and T = 298 K, then $t_{1/2} = 10^{4.5}$ s = 31622.8 s. We assume that $\Delta G^{\ddagger}$ is approximately equal to $\Delta E^{\ddagger}_{\mathrm{el}}$. Since the half life for a $\Delta G^{\ddagger}$ of 23.9 kcal/mol is already very long, we assume that a an energy cutoff of 25-30 kcal/mol will be sufficient. The activation energy $\Delta G^{\ddagger}$ from the energetically lower minimum-energy structure to the energetically-higher local-minimum structure must be at least as large as the energy difference between the two minimum-energy structures. Therefore, one may consider the threshold for the activation energy (i.e., the TS-structure energy) also for differences between minimum-energy structures. The structures that remain after this analysis step build the basis of the reaction-network, the nodes.

Subsequently, structure pairs connected by an elementary reaction are determined for TS optimizations. After the TS localization performed by the module ”TS search“, the TSs are verified by a Mode-Tracking based lowest-frequency analysis. Mode-Tracking is a protocol which employs the Davidson subspace-iteration technique to selectively localize only the mode of interest instead of calculating the full Hessian matrix [156–164]. Especially for large molecules, Mode-Tracking is significantly faster compared to a full Hessian analysis. Furthermore, intrinsic reaction coordinates (IRC) are calculated to confirm that the TS found is connecting the minimum-energy structures under consideration. From the TS energies, we can conclude whether a certain reaction takes place under the given conditions or not. The energy differences between the TS and the minimum-energy structures which are connected by the TS correspond to the activation energies for the reactions in forward and backward direction. The rates of these forward and backward reactions under the given reaction conditions can be obtained by Eq. (6.2). We can exclude all edges from the reaction network for which neither the forward nor the backward reaction can be accomplished within a reasonable time-scale.

The final results are plotted in a reaction network, where all minimum-energy structures, their energies relative to the lowest-energy structure, and the respective transition-state energies relative to the lowest-energy structure are reported in one graph. The global lowest-energy pathways can thus easily be identified.

All single steps are discussed in detail in the following sections.

6.2 Structure generator and optimizer

The substrate to be protonated and a list of atom connectivities in the substrate is the only required input. To define proton-acceptor atoms within the substrate, HeurEx counts the bonds of each atom with adjacent atoms and compares this value to the maximum number of possible bonds this atom can form (according to its valence, up to 4 bonds are possible for main group elements). Although this approach easily works out for the main group elements, the determination of protonation sites at transition metals is more elaborate. In those cases additional rules are applied as described in the listing below (key point 5). According to the actual number of bonds formed with atom i, $N_{b,i}$, of a complex consisting of M atoms, and the maximum possible number of bonds to that atom, $N_{\max,i}$ ($N_{\max,i} = 4$ for main group elements), the number of protonation sites is defined as

$$N_{\mathrm{p}} = \sum_{i}^{M} N_{\max,i} - N_{b,i}. \tag{6.4}$$

The total number of protons added to the cluster varies between one and the number of possible protonation sites. To define protonation sites, we first determine the geometry of each acceptor atom with its neighbors and then add (a) new proton(s) according to the following rules:

1. The default value for the distance between the acceptor atom and the proton is 1.0 Å. The molybdenum-proton distance is increased to 1.7 Å. A more general setup can be achieved by reading in the optimized atom–proton distances from a reference table.

2. If the acceptor atom has only one neighbor atom, up to 3 protons can be added in an approximately tetrahedral geometry.

3. If the acceptor atom has two neighbor atoms and forms an approximately linear arrangement with them, up to two protons are added in the plane perpendicular to the 2 adjacent atoms (with the constraint of a 180° angle between the two protons, but a random position of the first proton within this plane); if the neighbor–acceptor–neighbor angle is bent, the new atom positions are located at a position where the angle above 180° is halved;

4. If the acceptor atom has three neighbor atoms and forms an approximately planar geometry with them, two protonation sites are created perpendicular to this plane (additionally, a constraint of only occupying one of these protonation sites is set); if the four atoms are not planar, the new protonation site is added in an approximately tetrahedral geometry;

5. If the acceptor atom has more than three neighbor atoms and the acceptor atom does not belong to the main group elements, e.g., transition metals, the sterically least demanding positions are chosen for the protonation sites. Otherwise, no proton is added. The geometry of the ligand sphere of the transition metal is determined according to the seven ideal symmetries employed in the crystal field theory, which are: octahedral, pentagonal bipyramidal, pentagonal pyramidal, square planar, square pyramidal, tetrahedral and trigonal bipyramidal. For example, if the transition metal has a trigonal bipyramidal ligand sphere, up to 3 protons are added in the layer spanned by the trigonal atoms such that they divide the 120° ligand-transition metal-ligand angles. Analogously, for the other types of ligand spheres the sterically least demanding 3 positions are chosen as protonation sites.

Noteworthy, we do not create new input structures for the differently charged clusters, but assume that the same input structures can be employed for the structure optimizations at the different charges. We assume that for each proton one electron can be added such that the number of structure optimizations is given by

$$N_{\text{opt}} = \sum_{p=1}^{N_{\text{PS}}} \binom{N_{\text{PS}}}{p} (p+1), \tag{6.5}$$

where p is the number of protons and N_{PS} the number of protonation sites.

We have to ensure that the heuristic input generation does not lead to energetically unfavorable starting geometries. For example, a chosen guess structure may be convenient for a certain number of protons added, but inappropriate for another number of protons added. Especially for the linear arrangement of case 3, the initial-guess structure has to be carefully analyzed, because a structural rearrangement of one of the two neighbor atoms or groups, respectively, might be necessary to obtain the desired minimum-energy structure for a certain number of protons added. One option may be to implement a function that searches for the sterically and electronically best arrangement of a proton in a given environment, or to relax part of the molecule with the constraint that the proton is bound to the desired acceptor atom instead of applying the geometrical rules described above.

Moreover, the number of protonation sites might be larger than the actual number of protons that can be added to the acceptor atom (compare rule 4). In those cases, we define a constraint that allows to only occupy as many protonation sites as the atom can accept. The number of input structures can hereby be significantly reduced.

If a structure optimization fails, because a self-consistent field calculation did not converge, HEUR𝖤x changes the damping parameters automatically (see Section 6.7) and restarts the calculation. Structures for which the optimization does not converge within 1200 optimization steps, are saved and we have to investigate those structures manually to decide whether they shall be covered in the reaction network or not.

6.3 Elementary-step finder

All optimized structures in which the molecule decomposes are automatically removed. This is achieved by calculating the distance between the added proton(s) and the center of mass of the molecule and the spatial root-mean square deviation (RMSD) between the initial and the optimized structures (without the added proton(s)). If a proton–center-of-mass distance or the RMSD exceed a predefined threshold, the structure is discarded.

Within a collection of optimized structures with the same charge and number of atoms, multiple structure optimizations may lead to the same minimum-energy structures. To identify these structures, the RMSD for every pair of structures within each collection is calculated after their alignment. Structures within a certain RMSD threshold (here we choose 0.5 Å) are collected in so called clusters.

The next step is the identification of all structure pairs between which a TS search shall be carried out. For the studied network of protonated intermediates, this means that we have to identify all structure pairs which are related by a shift of exactly one proton. If more than one possible proton shift pathway exists for a given pair, the one with the smallest replacement of the transferred proton is chosen. Noteworthy, this procedure might not always be the best choice, e.g., in cases where roaming reactions are favored. For these cases, one may calculate both reaction pathways.

6.4 Automated transition-state searches between two minimum-energy structures

The basis of the proceeding TS searches is the collection of structure pairs with the same charge, the same chemical formula and a difference in exactly one proton position. For an automated transition-state search between two minimum-energy structures, one can in principle utilize several TS-search methods (e.g., interpolation methods [139], eigenvector-following (EVF) [140–144], string methods [145], the scaled hypersphere search method [146], a constrained optimization technique [214], or the recently published Davidson subspace-iteration based TS-search algorithm [166]).

Here, we choose a combination of constrained structure-optimization scans with a subsequent EVF run. According to the structural differences between the two minimum-energy structures, the direction (distance, angle, etc.) and step size (default: 8 steps) for the constrained optimization scan has been chosen. Out of the constrained optimization scan, the highest-energy structure is selected for a subsequent Mode-Tracking calculation of the lowest-frequency mode. If the initial-guess structure has one imaginary frequency, an EVF procedure is carried out. If not, the first procedure failed. In this case, the freezing-string method [215] (FSM) is chosen in combination with EVF. If both TS-search algorithms fail to converge the desired TS, we assume that the respective TS is very high in energy or there is no TS in between the respective structure pair.

Clearly, none of the current TS-search methods can ensure that the lowest-energy barrier will be found. However, one can easily interface or include other TS-search algorithms in HeurEx and thus increase the reliability.

6.5 Verification of the TSs

For the verification of the TSs, HeurEx first analyzes whether the TS calculation has converged or not. Second, HeurEx starts a vibrational analysis with Mode-Tracking on every presumable TS. Structures with only positive frequencies and with frequencies that are all larger than e.g., $-i100$ cm^{-1} (by default) are removed, because they are presumably minimum-energy structures. If exactly one imaginary frequency is obtained, we visualized the corresponding transition mode to decide whether the correct TS has been found. For structures with a transition mode not obviously leading to the minimum-energy structures the TS shall connect, the IRCs are calculated to make sure, that the TS connects the correct two minimum-energy structures. One can easily generalize this procedure by calculating the IRC for all TS structures and performing an automated alignment between the minimum-energy structures of the IRC and the starting structures to figure out whether the TS connects the desired minimum-energy structures or not.

6.6 Visualization of the results in reaction networks

In general, a reaction network consists of nodes and edges connecting the nodes. In our example, the nodes of the reaction network represent the minimum-energy structures and edges represent TS paths between the respective minimum-energy structures. If the TS search for a structure pair is successful and the activation energy is below the chosen threshold, the energies of the TSs compared to the lowest-energy structure are written on the edge lines. To obtain the activation energies for the forward and backward reactions, one must thus calculate the differences between the TS energy and the energies of the two minimum-energy structures. The optimized minimum-energy structures are graphically displayed and the energy differences compared to the lowest-energy structure are reported. If the number of nodes is large, we do not display the optimized structures in the nodes, but label the respective node with a structure identification number. The optimized structures are then visualized in a separate list. The colors of the nodes correspond to the energy differences with respect to the lowest energy. Dark blue represents the lowest-energy structures and dark-red the highest-energy structures. The calculated TS energies are presented above each edge. For a clearer visualization of energetically low and high barriers, we scale the line width of the edges and the gray scale according to the TS energy such that a thick line and black color represents a low-energy barrier and a thin line in light gray represents a high-energy barrier. The intermediates of the catalytic cycle are highlighted by the gray background to distinguish them from side products.

6.7 Generation of raw data

Density-functional calculations are performed with the program package Turbomole [190] (versions 6.3.1 and 6.4.0) using Ahlrich's def2-SV(P) basis set [129] with polarization functions on all atoms. Unrestricted and restricted all-electron Kohn–Sham calculations are carried out. The density functional BP86 [133, 134] in combination with the resolution-of-the-identity (RI) technique is chosen. The self-consistent-field (SCF) single-point calculations are considered to be converged when the total electronic-energy difference between two iteration steps is less then

10^{-7} hartree. The SCF damping parameters are adapted from the default values to start = 5.000, step = 0.050, min = 0.50 and scforbitalshift automatic = 1.0, if the energies oscillated for the given structure. Structure optimizations are considered converged when the length of the geometry gradient dropped below 10^{-4} atomic units.

In the constrained optimizations with GAUSSIAN [191] (version 09, Revision C.1) we employ BP86 with the def2-SVP basis set [192, 193]. One internal coordinate is kept fixed at defined values (obtained from the reactant- and product-structure differences), and along all other degrees of freedom a structure optimization has been carried out. For most cases, this procedure led to a reasonable TS guess structure. These guess structures are then refined with TURBOMOLE's trust-radius image-based EVF optimization (with a trust-radius of 0.2 Å). The mode to follow is obtained from a Mode-Tracking calculation. If the EVF optimization converged, a second Hessian calculation has been carried out to figure out whether the stationary structure has exactly one imaginary frequency. In these cases, intrinsic reaction coordinates are calculated with GAUSSIAN (BP86/def2-SVP) to figure out whether the correct TS has been found. We employ the default convergence criteria for all GAUSSIAN calculations (scfconv=tight, which means that the energy difference between two cycles is less than 10^{-8} hartree, and optimizations are considered converged when the root-mean-square force is below $3{\times}10^{-4}$ a.u.).

If the constrained optimization scan with proceeding EVF did not converge to the desired TS, an alternative algorithm is employed. Here, we choose the freezing-string method (FSM) with 12 nodes (B3LYP/3-21G) from Q-CHEM (version 4.0.1) [216].

The structures are visualized with PYMOL [188] and the python modules PYGRAPHVIZ and NETWORKX are employed for the pictorial representation of the reaction networks.

6.8 Reaction Networks for Intermediates of the Chatt–Schrock cycle

The main catalytic reaction steps of the full nitrogen-fixation cycle (i.e., the Chatt–Schrock cycle) are shown in Figure 6.2, on the left. In this Section, we investigate the first half of the Chatt–Schrock cycle, which is highlighted in gray. All intermediates are based on an N_2 fragment bound

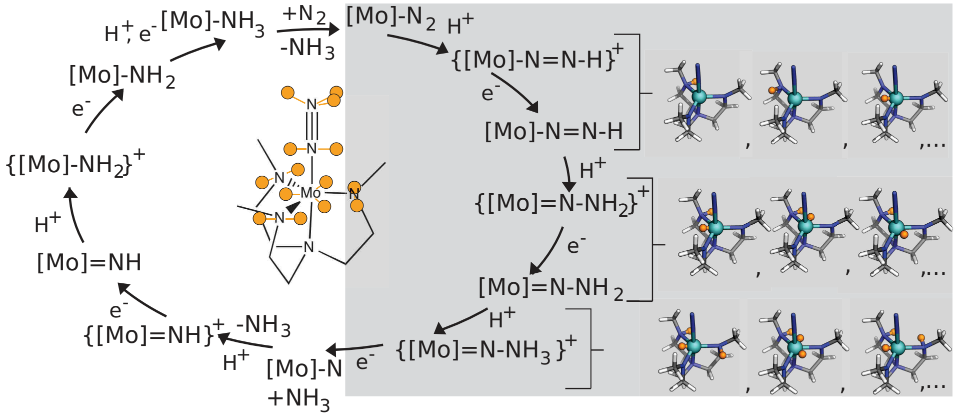


Figure 6.2: Left: The Chatt–Schrock nitrogen-fixation cycle is shown. [Mo] refers to the HIPTN$_3$N(HIPT = hexa-*iso*-propyl terphenyl)-ligated molybdenum complex. The part of the catalytic cycle shaded in gray is studied in this work. In the middle of the reaction cycle a simplified Lewis structure (without the H atoms) of the model system is shown and in orange all possible protonation sites are highlighted. Right: Singly, doubly and triply protonated model systems are shown to indicate that each intermediate shown in the cycle represents a full reaction network. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, protonation sites.

to the Mo atom. The second part of the catalytic cycle contains a [Mo]–N fragment, since one nitrogen atom is already released from the complex in form of NH_3. [Mo] refers to the HIPTN$_3$N(HIPT = hexa-*iso*-propyl terphenyl)-ligated molybdenum complex.

Our reaction networks are based on the Schrock model complex with aryl rests substituted by methyl groups (see Figure 6.2, in the middle of the catalytic cycle). The lowest-energy spin state of the neutral N_2-bound Schrock model complex is a doublet. The spin states with higher multiplicity are at least 60 kJ/mol higher in energy. Since this energy difference is quite large, we can assume that the ground-state spin state is adopted. Proton-acceptor atoms are the amido nitrogen atoms, the N_2 group atoms and molybdenum atoms. All protonation sites that are considered in this study are highlighted in orange. According to the valency of the nitrogen atoms, the amido nitrogen atoms can only accept one proton. However, due to the fixed geometry of the HIPT ligand, one should take into account two different positions for these protons (as shown in the Lewis structure representation in Figure 6.2). Thus, a constraint of only occupying one of the two protonation sites has been set. To prevent the generation of structures with an artificially increased energy, we compare the protonation input structures of the nitrogen atom of N_2 that is directly bound to Mo in a linear manner versus a bent one. We did not observe any differences in the structure-optimization results and thus assume that the guess structures based on the defined proton positions are decent. Optimized structures in which H_2 is formed are removed from the reaction networks and only reported (on the bottom of the reaction network), if the energy of the H_2 formation is preferred compared to the other intermediates within the reaction network.

The BP86/RI/def-SV(P) reaction networks of all one- to four-fold protonated intermediates at the different charge are calculated. Each reaction network is based on the same number of protons and electrons. To obtain the energy differences between the reaction networks and thus to connect all relative energies with each other, the energy differences between the Chatt-Schrock cycle intermediate structures (see Figure 6.2, main intermediates in the catalytic cycle) are calculated in each reaction network. These structures are highlighted by a gray background in each reaction network. The calculated energy differences between the reaction networks correspond to the one-fold protonation and one-e^- reduction energies and they are reported in Figure 6.3. On the right-hand side of Figure 6.3, we present the energies of the Chatt–Schrock cycle intermediates (marked by a *) and the lowest energies of each reaction network relative to the reference structure, which is the N_2-coordinated Schrock complex.

Noteworthy, for the calculation of protonation reactions in the gas phase, protons do not contribute to the reaction energy since they have no electronic energy (assuming the Born–Oppenheimer approximation (zero kinetic energy) and an infinite distance between the two reactants (zero potential energy)). Therefore, an artificially large exothermic protonation reaction is observed. One common solution is to take the solvation energy of a proton in a certain medium as reference. In bioinorganic reactions, the energy of a solvated proton is often assumed to be −262.4 kcal/mol [217, 218] (solvation in water). A range between −252.6 and −262.4 kcal/mol for the solvated-proton energy is given in Ref. [218]. If we assume lutidine-H^+ (lutidine=2,6-dimethylpyridine) as proton source, the BP86/RI/def2-SV(P) reference energy for the proton is −237.7 kcal/mol. For the electron oxidation reaction from $Cp^*{}_2Cr$ (Cp^* = 1,2,3,4,5-Pentamethylcyclopentadienyl) to $Cp^*{}_2Cr^+$ a BP86/RI/def2-SV(P) reaction energy of +103.7 kcal/mol is obtained. An advantage of HEUREX is that one can

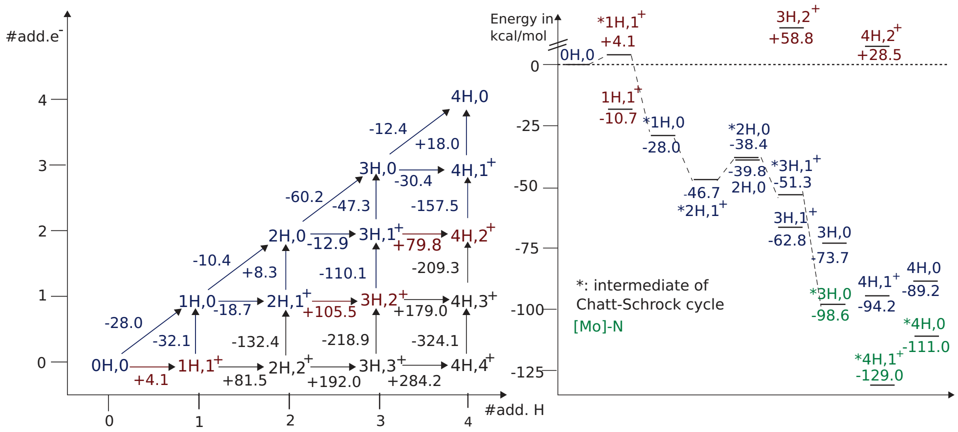

#add.e$^-$
4
3
2
1
0
4H,0
-12.4
+18.0
3H,0
-30.4
4H,1$^+$
-60.2
-47.3
-157.5
2H,0
-12.9
3H,1$^+$
+79.8
4H,2$^+$
-10.4
+8.3
-110.1
-209.3
1H,0
-18.7
2H,1$^+$
+105.5
3H,2$^+$
+179.0
4H,3$^+$
-28.0
-32.1
-132.4
-218.9
-324.1
0H,0
+4.1
1H,1$^+$
+81.5
2H,2$^+$
+192.0
3H,3$^+$
+284.2
4H,4$^+$
0
1
2
3
4
#add. H
Energy in kcal/mol
0
-25
-50
-75
-100
-125
*1H,1$^+$
+4.1
0H,0
3H,2$^+$
+58.8
4H,2$^+$
+28.5
1H,1$^+$
-10.7
*1H,0
-28.0
*2H,0
-38.4
-39.8
2H,0
*3H,1$^+$
-51.3
-46.7
*2H,1$^+$
3H,1$^+$
-62.8
3H,0
-73.7
*: intermediate of Chatt-Schrock cycle
[Mo]-N
*3H,0
-98.6
4H,1$^+$
-94.2
4H,0
-89.2
*4H,0
-111.0
*4H,1$^+$
-129.0

Figure 6.3: BP86/RI/def2-SV(P) electronic-energy differences (in kcal/mol) between the Chatt-Schrock cycle intermediates in each reaction network (characterized by a certain number of protons and electrons added). The energy of a solvated proton is assumed to be −237.7 kcal/mol. For the electron reduction, Cp$^*{}_2$Cr was chosen as electron source, which corresponds to a BP86/RI/def2-SV(P) energy difference of +103.7 kcal/mol. The dark-blue intermediates are lower in energy than the red ones and the black ones are not accessible at the given conditions ($T = 298$K).

easily change the proton and electron reference values depending on the specific reaction conditions. The program can automatically highlight those reaction energies that depend on the solvated-proton reference energy and might thus be interpreted carefully in order to not accidentally exclude a part of the reaction network that is accessible.

First of all, our results are in accordance with the calculated and experimental reaction energies between intermediates of the Chatt–Schrock cycle [202, 219]. From the energy differences between the reaction networks (with a fixed number of protons and electrons) we can estimate which types of intermediates occur at the given conditions and which can be neglected. We label the reaction networks according to the number of protons added to the N_2 bound Schrock complex reference structure (xH, with x being the number of protons) and the total charge of the system. For better readability, the two values that determine each reaction network are given in parentheses separated by a comma (i.e., (xH,charge)). We observe that the intermediates with one to three protons added, which are neutral and single-positively charged are easily accessible at room temperature. The reaction networks of these intermediates are shown in Figures 6.4, 6.5, 6.6 and 6.9. Reaction networks (3H,2^+) and (4H,2^+)

may be accessible, but they are significantly higher in energy than most of the 0 and +1 charged species. All reaction networks with higher charges ((3H,3$^+$), (4H,3$^+$) and (4H,4$^+$)) and reaction network (2H,2$^+$) are energetically disfavored and thus neglected in the main analysis (these reaction networks can be found in Appendix F).

From the reaction networks, one can extract the number and type of minimum-energy structures accessible at the given reaction conditions. For the singly-protonated intermediates, we obtain four intermediates with an energy below 25 kcal/mol. Interestingly, in the one-fold protonated and 1-e$^-$ reduced species, the intermediate of the catalytic cycle is by 14.8 kcal/mol higher in energy than the lowest-energy structure. Most of the activation energies are larger than 30 kcal/mol. The TS with an energy of +29.9 and +38.6 kcal/mol, may be reached from the energetically higher minimum-energy structures which are connected by the TS.

For the two-fold protonated intermediates with charges 1+ and 0, ten and thirteen side products next to the intermediate of the Chatt–Schrock cycle are accessible, respectively. Noteworthy, the H_2 dissociation product is by −11.6 kcal/mol lower in energy than all structures in the (2H,0) reaction network. The first species of the catalytic cycle is obtained again. This may indicate a hindrance for the catalytic formation of NH_3, because the H_2 formation appears to be an energetically preferred side-reaction. Noteworthy, a detailed investigation of possible H_2-formation pathways at the model complex has to be carried out.

In the three-fold protonated species, we obtain several low-energy structures, in which the ligand is destroyed. Since we have chosen a model complex with the aryl rests substituted by methyl groups and thus, the original complex is sterically significantly more demanding, these destroyed ligand sphere structures may be artificial. We reported the structures and their relative energies compared to the lowest-energy structure in the above reaction network on the bottom of the reaction network, separated by a dashed line. In the (3H,0) species the H_2-formation energy is by −5.7 kcal/mol more exothermic than the lowest-energy structures in the reaction network. However, the formation of NH_3 is by −24.9 kcal/mol more exothermic than the lowest-energy structures in the reaction network. Thus, the NH_3 formation is clearly favored over the other possible side reactions.

In total, the reaction networks for the one- to three-fold protonated species with charges 0, and +1, demonstrate that a large number of possible side products next to the desired intermediate of the Chatt–Schrock cycle is formed. Hence, by investigating only the reaction energies between these intermediates, one may miss an energetically preferable side reaction. A manual study of the minimum-energy structures and reaction pathways would be very time-consuming and thus, the automated program HeurEx can contribute significantly to the exploration of the full reaction network.

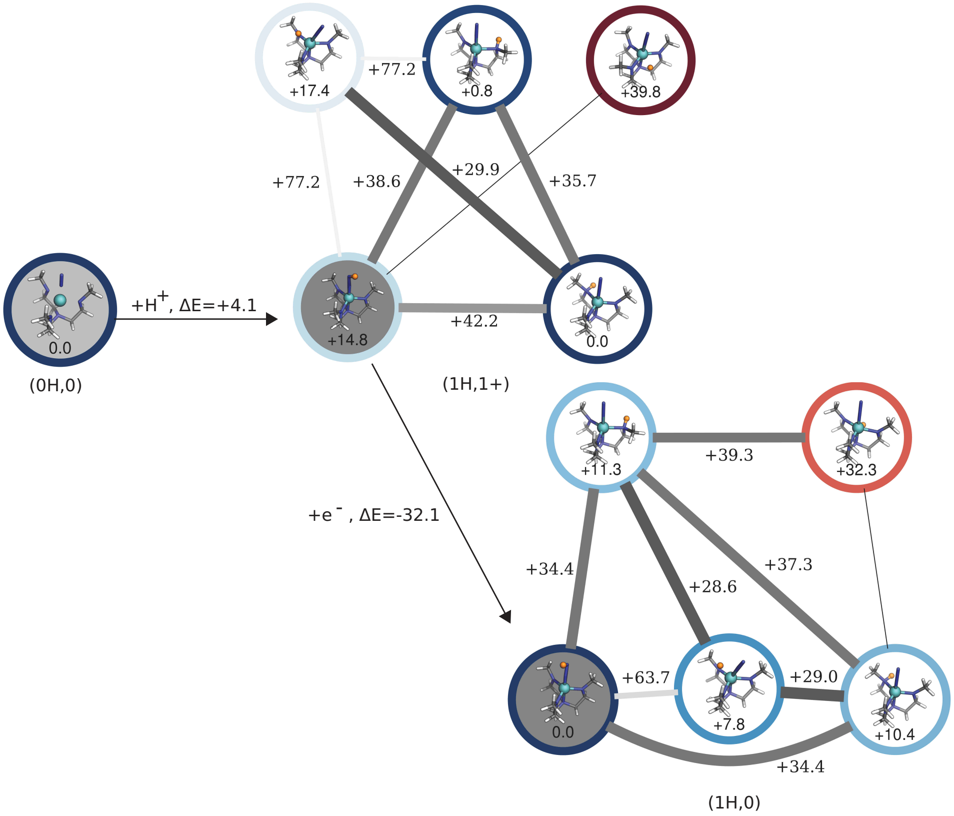

+17.4
+77.2
+0.8
+39.8
+77.2
+38.6
+29.9
+35.7
0.0
+H$^+$, ΔE=+4.1
+14.8
+42.2
0.0
(0H,0)
(1H,1+)
+11.3
+39.3
+32.3
+e$^-$, ΔE=-32.1
+34.4
+28.6
+37.3
+63.7
+29.0
0.0
+7.8
+10.4
+34.4
(1H,0)

Figure 6.4: BP86/RI/def2-SV(P) reaction network for the first protonation and reduction step of the N_2-bound generic model complex of the Schrock catalyst. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol for the energy-differences of minimum-energy structures between which an elementary step is assumed. A dark-blue circle around the minimum-energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high energy structure (40 kcal/mol higher energy than lowest energy of each sub-reaction network). The intermediates of the Chatt–Schrock cycle are highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

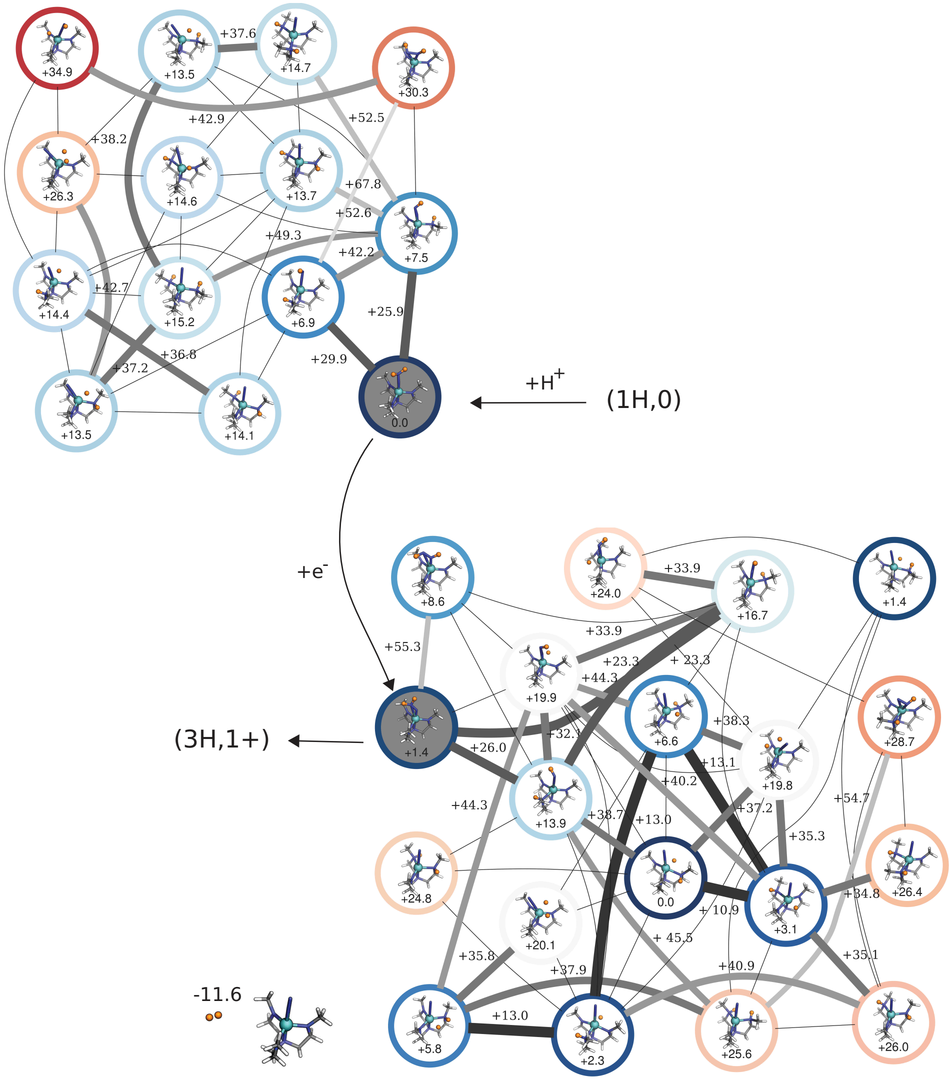

+34.9
+13.5
+37.6
+14.7
+30.3
+42.9
+52.5
+38.2
+26.3
+14.6
+13.7
+67.8
+52.6
+49.3
+7.5
+42.2
+14.4
+42.7
+15.2
+6.9
+25.9
+36.8
+37.2
+29.9
+13.5
+14.1
0.0
+H+
(1H,0)
+e-
+8.6
+24.0
+33.9
+16.7
+1.4
+55.3
+33.9
+23.3
+23.3
+44.3
+19.9
(3H,1+)
+1.4
+26.0
+32.1
+6.6
+38.3
+13.1
+19.8
+28.7
+40.2
+44.3
+13.9
+38.7
+13.0
+37.2
+54.7
+35.3
+24.8
0.0
+34.8
+26.4
+10.9
+3.1
+20.1
+45.5
+35.8
+35.1
+37.9
+40.9
-11.6
+13.0
+5.8
+2.3
+25.6
+26.0

Figure 6.5: BP86/RI/def2-SV(P) reaction network for the two-fold protonated and 1 e^--reduced (top) and the two-fold protonated and 2 e^--reduced (bottom) intermediates of the N_2-bound generic model complex of the Schrock catalyst. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol for the energy-differences of minimum-energy structures between which an elementary step is assumed. A dark-blue circle around the minimum-energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high energy structure (40 kcal/mol higher energy than lowest energy of each sub-reaction network). The intermediates of the Chatt–Schrock cycle are highlighted by a gray background. For those edges which are not labeled and only a thin line is drawn, the TS is either very high in energy or no TS exists. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

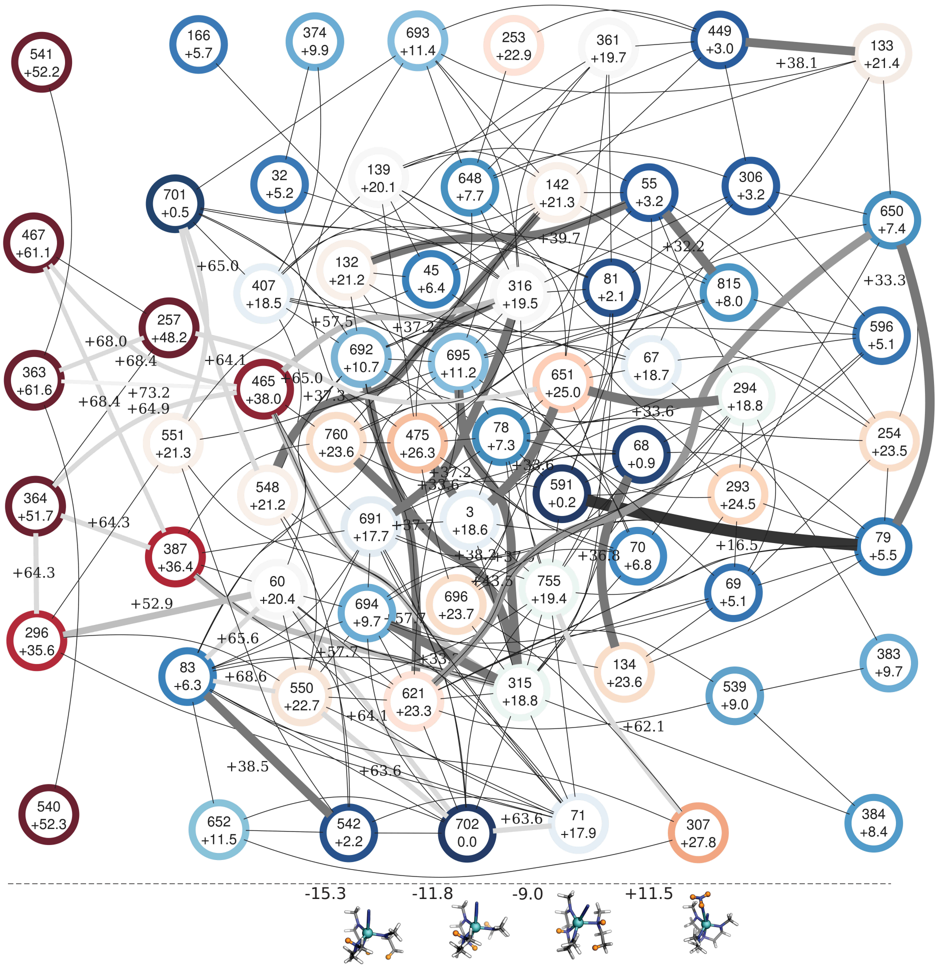

541 +52.2
166 +5.7
374 +9.9
693 +11.4
253 +22.9
361 +19.7
449 +3.0
133 +21.4
+38.1
701 +0.5
32 +5.2
139 +20.1
648 +7.7
142 +21.3
55 +3.2
306 +3.2
650 +7.4
467 +61.1
+65.0
407 +18.5
132 +21.2
45 +6.4
316 +19.5
+39.7
81 +2.1
+32.2
815 +8.0
+33.3
257 +48.2
+57.5
+37.2
692 +10.7
695 +11.2
67 +18.7
596 +5.1
+68.0
+68.4
+64.1
465 +38.0
+65.0
+37.3
651 +25.0
294 +18.8
363 +61.6
+68.4
+73.2
+64.9
551 +21.3
760 +23.6
475 +26.3
78 +7.3
+33.6
68 +0.9
254 +23.5
548 +21.2
591 +0.2
293 +24.5
364 +51.7
+64.3
691 +17.7
3 +18.6
387 +36.4
70 +6.8
+16.5
79 +5.5
+64.3
60 +20.4
694 +9.7
696 +23.7
755 +19.4
69 +5.1
296 +35.6
+52.9
+65.6
+57.7
134 +23.6
383 +9.7
83 +6.3
+68.6
550 +22.7
621 +23.3
315 +18.8
539 +9.0
+64.1
+62.1
+38.5
+63.6
540 +52.3
652 +11.5
542 +2.2
702 0.0
+63.6
71 +17.9
307 +27.8
384 +8.4
-15.3
-11.8
-9.0
+11.5

Figure 6.6: BP86/RI/def2-SV(P) reaction network for the three-fold protonated and 2 e$^-$-reduced intermediates (total charge of 1+ elementary unit). Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol for the energy-differences of minimum-energy structures between which an elementary step is assumed. A dark-blue circle around the minimum-energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high energy structure (40 kcal/mol higher energy than lowest energy of sub-reaction network). The intermediate of the Chatt–Schrock cycle is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's. For those edges which are not labeled and only a thin line is drawn, the TS is either very high in energy or no TS exists. On the bottom structures in which part of the Schrock complex fell apart and their energies relative to the lowest-energy structure of the network are shown.

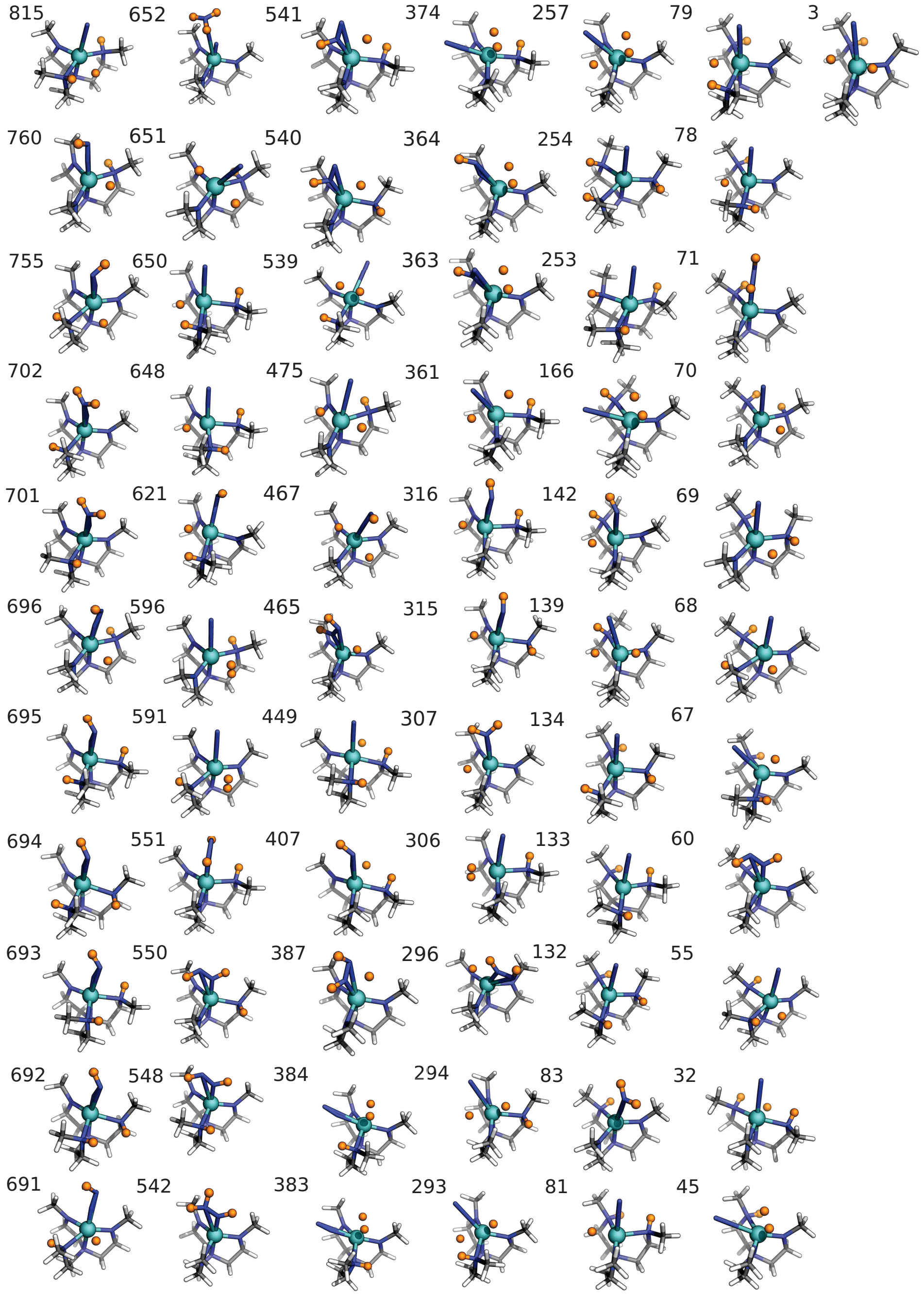

815
652
541
374
257
79
3
760
651
540
364
254
78
755
650
539
363
253
71
702
648
475
361
166
70
701
621
467
316
142
69
696
596
465
315
139
68
695
591
449
307
134
67
694
551
407
306
133
60
693
550
387
296
132
55
692
548
384
294
83
32
691
542
383
293
81
45

Figure 6.7: BP86/RI/def2-SV(P) optimized structures of the three-fold protonated and 2 e$^-$-reduced intermediates (total charge of 1+ elementary unit) and their identification number.

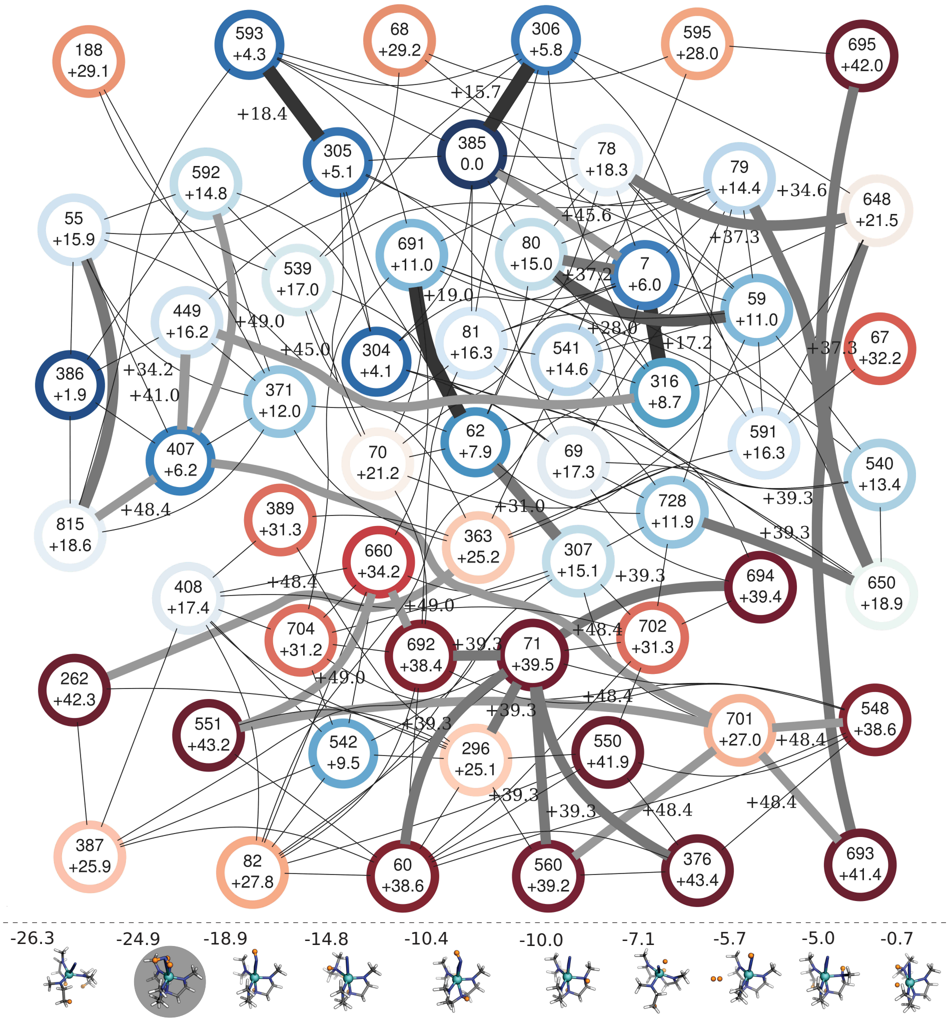


Figure 6.8: BP86/RI/def2-SV(P) reaction network for the three-fold protonated and 3 e^{-}-reduced. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol for the energy-differences of minimum-energy structures between which an elementary step is assumed. A dark-blue circle around the minimum-energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high energy structure (40 kcal/mol higher energy than lowest energy of sub-reaction network). The intermediate of the Chatt–Schrock cycle is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's. For those edges which are not labeled and only a thin line is drawn, the TS is either very high in energy or no TS exists. On the bottom structures in which part of the N_2 bound Schrock complex fell apart and their energies relative to the lowest-energy structure of the network are shown.

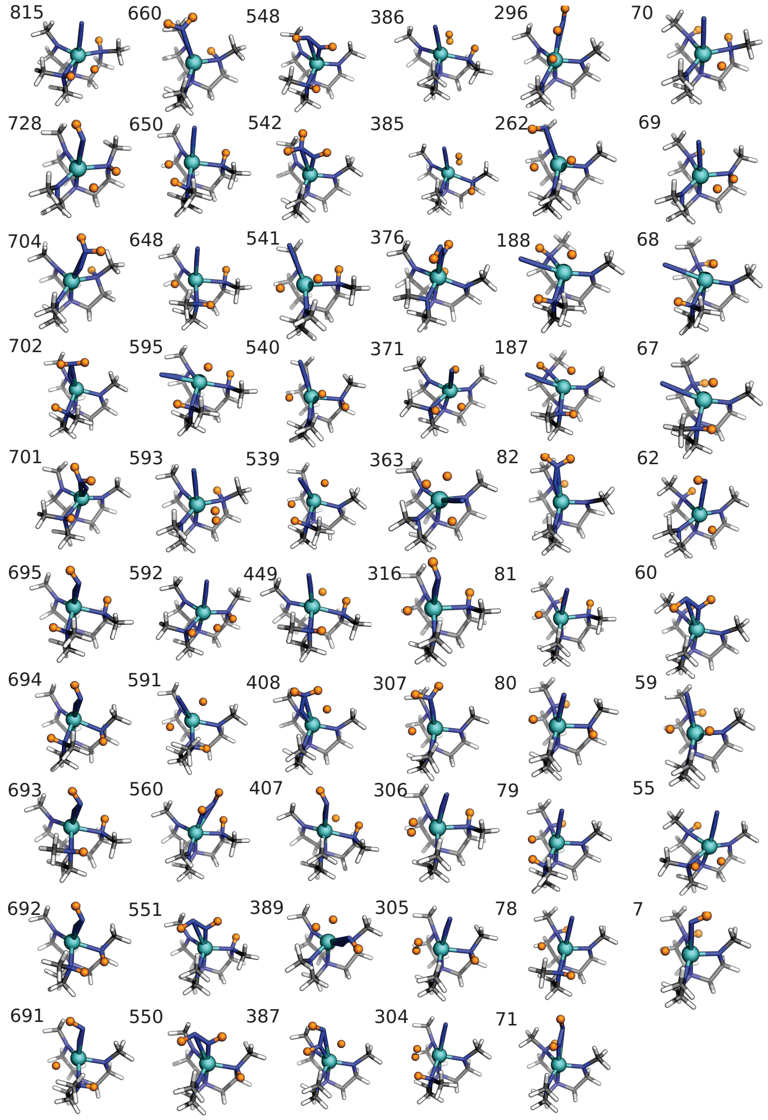

815
660
548
386
296
70
728
650
542
385
262
69
704
648
541
376
188
68
702
595
540
371
187
67
701
593
539
363
82
62
695
592
449
316
81
60
694
591
408
307
80
59
693
560
407
306
79
55
692
551
389
305
78
7
691
550
387
304
71

Figure 6.9: BP86/RI/def2-SV(P) optimized structures of the three-fold protonated and 3 e^--reduced intermediates of the N_2-bound generic model complex of the Schrock catalyst and their identification numbers.

7

Conclusions and Outlook

In two case studies, we observed that one of the biggest challenges for mechanistic explorations is the multitude of structures that has to be considered. If many reaction pathways are accessible, the manual calculation of TSs is often tedious or completely unfeasible and an automated calculation of the reaction pathways is needed. Such an automated program requires two main ingredients, i.e., an algorithm to obtain new structures and an efficient and computationally feasible TS optimization algorithm. Along these lines, we have developed two new algorithms, one which allows for computationally achievable TS optimization of large molecules (MTSEARCH) and one, which explores reaction networks based on chemical intuition in a heuristic manner (HEUREX).

The search for multiple reaction pathways starting from one minimum structure is still a main obstacle of current transition-state optimization programs. In general, chemical intuition is needed to chose a suitable starting mode, which connects the educts with the products. In Chapter 5, we have described the MTSEARCH program that efficiently combines the calculation of selected normal modes by the Mode-Tracking scheme [156, 183, 220] and the eigenvector-following procedure to eventually optimize transition-state structures. Since Mode-Tracking avoids the time-consuming calculation of the complete Hessian matrix and instead only optimizes the modes of interest, MTSEARCH is particularly suitable for optimizing transition-state structures of large reactive molecular systems. The search for several transition-state structures is feasible and the starting structures for a TS search may be outside the quadratic region of the transition-state structure.

We investigated the performance of MTsearch on four intramolecular reaction pathways: the rotational barrier of C_2H_6, the isomerization of H_2CO, a proton shift in the hydrazine bound intermediate of the [Mo(hiptN$_3$N)] catalyst by Schrock as well as in a model system with methyl substituents instead of the aryl substituents. We found that the initial guess modes for the Mode-Tracking procedure can be either extracted from the LST or NEB pathway, or based on chemical intuition. A TS optimization can be started either from two or just from one minimum-energy structure only. The potential energy surface can be explored in a customized way along the desired directions. Even for a large molecule such as the hydrazine Mo complex of Schrock and coworkers with more than 200 atoms, we were able to efficiently locate a transition-state structure.

By choosing different initial guess modes and/or branching off at certain structures during the optimization, one may scan the potential energy surfaces along different directions.

In Chapter 6, we presented the second algorithm, HeurEx, which efficiently constructs protonation- and electron-reduction-based reaction networks. Guess structures for reactive complexes were generated by exploiting chemical heuristics and they were subjected to structure minimization procedures. In particular, protonation sites were identified and protonated. However, the whole algorithm can be easily generalized to arbitrary reactive species to be placed into the vicinity of a target. HeurEx compares all optimized structures with the same number of protons and electrons with each other, searches for elementary-step reactions and starts TS searches between the chosen structure pairs. An intuitive visualization of the results is achieved by labeling all color-coded nodes (minimum-energy structures) and edges (TSs) with the respective relative energies. This simplifies the analysis of the global lowest energy paths.

At the example of the nitrogen-reducing molybdenum catalyst synthesized by Yandulov and Schrock, we studied all possible protonated and reduced clusters (but no negatively-charged clusters) as well as chemically relevant reaction pathways. The reaction networks clearly demonstrate that the number of accessible minimum-energy structures of several intermediates of the Chatt–Schrock cycle is very large. Thus, for the investigation of all reaction pathways, an automated program such as HeurEx is required. From the reaction networks of the one- to four-fold protonated intermediates, the lowest energy structures for a specific number of protonation and electron reduction reactions were identified. Furthermore, it became clear which internal proton transitions may occur at room temperature. We observed that for the two-fold protonated and 2-e^- reduced species, the H_2 formation product is by −11.6 kcal/mol favored compared to the lowest-energy structure of the reaction network.

With all relative energies at hand, the next step is to implement the automated kinetic modeling for this (as well as the second) half of the reaction cycle, which is current work in progress in our group. Then, we will be able to estimate the amount of each individual species formed within a certain time frame. A further goal is to connect the implemented algorithms with a haptic device to explore the chemical reaction network interactively. Since ultra-fast real-time calculations allow for an immediate force feed-back [185–187], the user can determine the development of the reaction network by steering the algorithm towards the chemical regions of interest. This is also work in progress in our group and may lead to significantly facilitated chemical-mechanism and kinetic-modeling studies.

Appendix

A [FeFe] hydrogenase calculations

Out of the protein structure with manually inserted $[2Fe]_H$ subcluster, we first select a large model system, with the complete amino acids Glu373, Met375, Cys129 (protonated), Cys377, Cys381, Cys130, Cys185, Lys188, Glu191, Ser153, Gln155, Pro154 and Ala54 taken into account. We cut our large model system such that the complete peptide bond with the next amino acid is taken into account and the C_α atom becomes a methyl group. The backbone atoms (peptide C, N and O atoms) and the methyl groups of the former C_α atoms are kept fixed in the structure optimizations. It should be noted, that three of the amino acids in the protein environment of the large model system are charged Glu373: −, Lys188: + and Glu191: − and the overall charge of the cubane and H-cluster is either 2− or 3− elementary charges. In the WT model system water molecules with labels 2, 6, 18, 34, 40, 58, 81, 107, which are closest to the active center, are taken into account. The large model systems of the WT and its mutant are displayed in the main text (Figure 2.1).

In a second step, we choose a small model system. Here, we take into account only those parts of certain side chains, which directly interact with the $[2Fe]_H$ subcluster. The amino acids which are considered herein, are Lys188, Glu191, Ser153, Gln155, Pro154, Cys377, Cys381, Cys130, Cys185 and Cys129 (protonated). The C atoms of the side chains which are closest to the backbone are fixed during the structure optimizations. By comparing the calculated properties (such as coordination energies and HOMO-LUMO gaps) of the small and the large model system with each other, we are able to estimate the effect of the more distant protein environment, which is only taken into account explicitly by the large model system. The continuum solvation method (COSMO, with $\epsilon = 4$) is employed in the small and the large model system to account for electrostatic screening effects of the protein environment further away from the catalytic center.

All density functional calculations are performed with the program package TURBOMOLE [190] (version 6.3.1). Unrestricted, all-electron Kohn–Sham calculations are carried out using Ahlrichs' valence double-ζ def2-SVP and triple-ζ def2-TZVP basis set [129] with polarization functions on all atoms. The TPSS [136] density functional is chosen in combination with the resolution-of-the-identity (RI) technique. In order to test the dependence of the HOMO-LUMO gaps on the exact Hartree–Fock exchange employ in the density functional, we apply TPSSh [136] (containing 10% HF exchange) and B3LYP [221–223] (containing 20% HF exchange). The self-consistent-field single-point calculation are considered to be converged when the total electronic energy difference between two iteration steps is less then 10^{-6} hartree for the large model system and 10^{-7} hartree for the small model system. The constrained structure optimizations are considered converged when the length of the geometry gradient is below 10^{-3} atomic units for the large model system and below 10^{-4} atomic units for the small model system.

The H-cluster models are optimized within the broken symmetry (BS) approach [12, 15, 19]. The different spin-coupling schemes are optimized by swapping coordinates of the iron atoms and employing converged orbitals defined with respect to the original coordinate listing [224]. Local spin expectation values are obtained by a local version of TURBOMOLE's MOLOCH module employing Löwdin's projection operators [20–22, 225, 226].

The low-spin singlet and doublet wave functions are lowest in energy. In principle, several broken-symmetry (BS) spin coupling schemes between the six iron atoms of the hydrogenase H-cluster are possible. In the H^{ox} clusters we restrict our study to the spin-coupling scheme where the two iron centers of the $[2Fe]_{H^{ox}}$ subcluster are antiferromagnetically coupled to the two closest iron atoms in the Fe_4S_4 cubane and those are again antiferromagnetically coupled to the other two iron atoms in the cubane. By contrast, the $[2Fe]_{H^{red}}$ subcluster of the H^{red} clusters are ferromagnetically coupled to the two closest iron atoms in the Fe_4S_4 cubane. Within the cubane there is still an antiferromagnetic coupling between the two iron atoms closer to the $[2Fe]_{H^{red}}$ subcluster and the other two iron atoms. Especially for the calculation of reaction energies, we ensure that the same coupling scheme between educts and products is achieved, although this BS spin-coupling scheme is not always the lowest-energy one.

The mutation of amino acid 187 from arginine to aspartate changes the total charge of the system by two elementary units. Interestingly, the charged amino acid is very close to the Fe_4S_4 cubane (5–6 Å to the cubane center), which suggests that the effect of the charge on the H cluster is not significantly damped by the protein environment or water molecules which might be located between the H-cluster and the charged amino acid. Due to the different overall charge of the wild type and its mutant, we can not directly compare the electronic energy differences between the two systems.

The structures and molecular orbitals are visualized with PyMOL [188] and Jmol [189].

B Supplementary material for Fe_4S_4 calculations

All density functional calculations are performed with the program package Turbomole [190] (version 6.3.1) using Ahlrichs' valence triple-ζ def2-TZVP basis set [129] with polarization functions on all atoms. Unrestricted, all-electron Kohn–Sham calculations are carried out. As density functionals BP86 [133, 134] in combination with the resolution of the identity (RI) technique and the hybrid functional B3LYP [221–223] are chosen. We mainly employ BP86, because in iron-sulfur clusters from [FeFe] hydrogenase it is found that BP86 yields reliable structures and reaction energetics [34, 36, 40, 227]. The self-consistent-field single-point calculation are considered to be converged when the total electronic energy difference between two iteration steps is less then 10^{-7} hartree.

The Fe_4S_4 clusters are optimized within the broken symmetry (BS) approach [12, 15, 19]. The different spin-coupling schemes are optimized by swapping coordinates of the iron atoms and employing converged orbitals defined with respect to the original coordinate listing [224]. Local spin expectation values are obtained by a local version of Turbomole's moloch module employing Löwdin's projection operators [20–22, 225, 226]. The spin couplings we obtain for the singly negatively charged clusters correspond to OS1 and OS2 according to Noodleman et al. [73]. Noodleman and coworkers define coupling schemes OS1, OS2 and OS3, which only differ in the spin populations of the cubane sulfur atoms. These are either both negative, of mixed sign or both positive, respectively. The iron-atom spin populations, which are the largest compared to the spin populations of all atoms in the cluster, are about the same in the three coupling schemes. For doubly negatively charged clusters Noodleman et al. obtained local spin populations on the cubane sulfur atoms which are of different sign compared to the spin populations on the iron atoms within the same ferromagnetically coupled Fe_2S_2 subcluster. By contrast, we obtain a spin-coupling pattern with spin populations of the same sign on these iron and sulfur atoms. For the clusters with elementary charge 3− Noodleman et al. defined two spin couplings, OC1 and OC2, which differ in the spin population of the ligands bound to the reduced or oxidized Fe_2S_2 sub-clusters. We obtain mainly OC2 spin coupling, in which the spin populations of the iron and sulfur atoms of the oxidized subcluster have the same sign. Also, the spin population of the iron and sulfur atoms of the reduced subcluster have the same sign, but different from the oxidized subcluster. It has to be noted that in the oxidized subcluster sometimes the two sulfur atoms possess different signs in the spin populations.

During structure optimizations either the four hydrogen atoms or the four sulfur atoms of the HS^- ligands are kept fixed in order to mimic the anchoring by the protein backbone. Structure optimizations are considered converged when the length of the geometry gradient is below 10^{-4} atomic units in such constrained optimizations.

As starting point for the calculation of the correlation diagrams the structure with the lowest *a* and *b* distances is selected. To obtain smooth correlation diagrams, converged α and β orbitals are provided for the next grid point with one step larger *b* distance. For the structure with one step larger *a* distance at lowest *b* distance, the molecular orbitals of the structure with one step smaller *a* distance are employed. In order to make sure that the lowest-energy solution is found, α and β orbitals from structures with larger *a* and *b* distances are provided in additional optimizations, if we observe kinks in the potential energy surfaces.

In order to choose a reasonable range of cluster distortions, the energetically lowest cluster geometry has to be found first. Distortions are then applied to this starting structure. An ideal structure is obtained by structure optimization (in C_1) of a D_{2d}-symmetric guess structure while no atoms are kept fixed. The three possible BS spin-coupling solutions A, B, and C are calculated. We observe that for this ideal ligand arrangement one BS spin-coupling scheme is about 2 kcal mol^{-1} higher in energy than the other two. In this spin-coupling scheme the cluster adopts the D_{2d}-1 form. In the lowest-energy spin-coupling scheme the cluster structure changes to an almost D_{2d}-2 form.

The optimized C_1-symmetric structures without constraints are then used to set up the distorted structures in the two D_{2d}-symmetric forms. The starting structures are symmetrized in order to obtain perfectly symmetric structures and the hydrogen atoms or the sulfur atoms of the HS^- ligands are fixed to yield the distance constraints a and b.

For the calculations of protein-embedded clusters from the PDB, the nitrogen and carbonyl carbon atoms of the cysteine ligands are changed to hydrogen atoms and these two hydrogen atoms of the ethyl thiolate rest are kept fixed at a C-H distance of 1.1 Å.

The continuum solvation model COSMO [228] is applied (with $\epsilon = 4$) to mimic electrostatic effects and to minimize charge artifacts. In nature, the protein scaffold compensates for the high charges of the active center.

All energies presented in this Section are calculated for 0 K without zero-point vibrational and temperature corrections. The structure–property plots are created with Mathematica 7.0 and 6.0 by Wolfram Research [229]. Molecular structures are visualized with PyMOL [188] and ChemBioDraw [230].

B.1 Analysis of S(Cys)-S(Cys)/C_β-C_β distances in selected crystal structures

In Tables B.1 and B.2 the S(Cys)-S(Cys) and C_β-C_β distances of cysteine-ligated Fe_4S_4 clusters of selected protein data bank entries with resolutions below 2.1 Å are listed. The clusters are arranged as shown in Figure B.1. In all crystal structures analyzed, one C_β-C_β distance is significantly shorter than all other C_β-C_β distances. This distance and the one on the opposite site of the cubane are labeled a. The remaining four C_β-C_β distances are labeled b. The S(Cys)-S(Cys) distances listed in Tables B.1 and B.2 refer to the corresponding C_β-C_β distances of the same cysteines.

The C_β-C_β distances in the chosen crystal structures vary between 4.5 Å and 9.1 Å, of which the shortest C_β-C_β distances range from 4.5 Å to 6.2 Å, four intermediate ones between 6.6–8.8 Å and the largest C_β-C_β distances between 8.5–9.1 Å. In our model cluster, $[Fe_4S_4(SH)_4]^{1-/2-/3-}$, the H-H distances of the HS^- ligands correspond to the C_β-C_β distances of the cysteine-ligated clusters. The chosen region of structural distortions

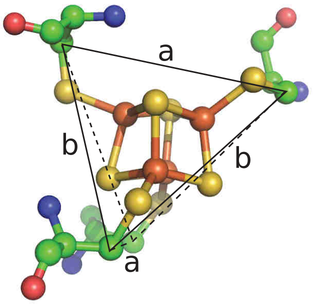


Figure B.1: Graphical representation of the *a* and *b* distances for a cysteine-ligated cubane cluster exemplarily shown for crystal structure 2FDN. Element color code: yellow, S; orange, Fe; green, C; blue, N; red, O.

in our correlation plots depends on the chosen D_{2d}-symmetric structure. For D_{2d}-1, a varies between 7.4 and 8.8 Å and b between 6.0 Å and 7.0 Å. For D_{2d}-2, a range of $a = 4.1–5.9$ Å and $b = 7.4–8.8$ Å is chosen, which approximately covers the natural distances. It should be noted, that the cysteine-ligated clusters from the crystal structures, especially in the HiPIPs, correspond rather to a D_{2d}-2 form than to D_{2d}-1. Therefore, the a and b distances in the correlation diagrams of structures in D_{2d}-2 form are more similar to the corresponding distances measured in the crystal structures.

The simplification of cysteine-ligands to HS^- ligands leads to small differences in the cluster-ligand geometry compared to cysteine-ligated clusters. The preferred FeSH angle differs from the favored $FeSC_\beta$ angle. In BP86 optimized unconstrained $[Fe_4S_4(SEt)_4]^{2-}$ clusters, the $FeSC(H_2)$ angle turns out to be 106°. This angle approximately resembles the $FeSC_\beta$ angle in cysteine-ligated clusters. For the HS^- ligands, the corresponding angle is considerably smaller, about 95° for D_{2d}-1 and 97° for D_{2d}-2. In the selected proteins the $FeSC_\beta$ angles range from 98–125°. In comparison in our calculations with fixed hydrogen atoms according to D_{2d}-2 form the distortions applied lead to variations in the FeSH angle ranging from 87° to 118° (BP86) and from 81° to 110° (B3LYP). In the S-fixed clusters the FeSH angle adopts values of 95–101° (BP86) and 94–102° (B3LYP). Although the FeSH angles in the H-fixed clusters are considerably smaller than the

$FeSC_\beta$ angles, the deviations from the angle of the lowest-energy structure are about −10° to +20°, which is in the region of the naturally occurring deviations. Furthermore, the Fe-Fe distances in the BP86 optimized H-fixed clusters range from 2.55 to 2.85 Å. The Fe-S cubanes in the protein data bank considered here have Fe-Fe distances of 2.54–2.84 Å, which is in good agreement with our model system. The Fe-S(Cys) distances of the experimental structures vary between 2.20 Å and 2.39 Å (one exception of 2.13 Å). The BP86 optimized structures in our correlation plots feature Fe-S(H) distances between 2.20 Å and 2.40 Å, which is within the range of the distances observed in proteins.

The S-S distances of the crystal structures vary between 5.9 Å and 6.9 Å (in the HiPIPs one S-S distance is shorter than the others, 5.5–6.0 Å). Our corresponding correlation plots are based on distances of 6.0–7.0 Å in *a* and 6.0–7.1 Å in *b* and hence cover the range of naturally occurring clusters. The fixation of sulfur atoms mainly affects the Fe-S and Fe-Fe distances and cluster-internal angles. The FeSH angles are not significantly affected by these constraints, since the hydrogen atoms are allowed to freely relax and can adopt any suitable position in space. The BP86 optimized $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ clusters in our correlation plots feature Fe-S(H) distances between 2.14 Å and 2.53 Å, which exceed the range of distances observed in proteins (2.20–2.39 Å). The Fe-Fe distances range from 2.49–3.04 Å. Compared with the Fe-Fe distances measured in the crystal structures (2.54–2.84 Å), the S-fixation obviously constrains the cluster slightly too strong for the largest distortions.

B.2 Comparison of the two D_{2d}-symmetric structures

In this section we compare the two D_{2d}-symmetric starting structures, D_{2d}-1 and D_{2d}-2, to understand whether the effect of distortions of distances *a* and *b* on cluster properties is similar in these systems. The main difference between the two configurations is the arrangement of the hydrogen atoms of the HS^- ligands. Hence, if the sulfur atoms are constrained, only small differences between the correlation plots are expected, because the sulfur positions are nearly identical in both configurations. The situation changes if the hydrogen atoms are fixed. The two directions along which *b* is measured in both models are not parallel. For example, we define a *z*-direction along the main rotation axis as shown in Figure B.2. For the same b the projection onto z is much larger in the D_{2d}-2 symmetric system than in the D_{2d}-1 symmetric system.

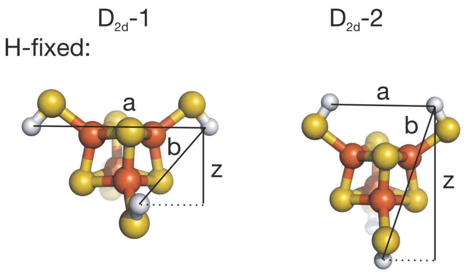


Figure B.2: Graphical representation of the two cluster-internal distance constraints a and b as well as an external z distance, which is introduced here for comparison of the two D_{2d}-symmetric structures.

In Figure B.3 the BP86 potential energy surface of $[Fe_4S_4(SH)_4]^{2-}$ in D_{2d}-1 form is shown. For a few structures with small a and large b the structure optimization does not converge. These values are missing in the correlation diagram.

In Figure B.4 the differences in the electronic energy between coupling schemes A and B (s. Figure 5 in the article), ΔE_{el}, are presented for both D_{2d}-symmetric structures. In order to visualize the dependence of ΔE_{el} on a structural constraint z instead of b, the axes in the correlation diagrams shown in Figure B.4 (upper panel) are transformed to Δa and Δz, where Δa and Δz describe the difference in a and z ”direction“, respectively, compared to the distances in the lowest-energy structure (see Figure B.4, lower panel). The changes in ΔE_{el} are qualitatively very similar for both definitions of distortions. We do not calculate the energies for structures, which we expected to be larger than 12 kcal mol^{-1}. Therefore, in the ΔE_{el} correlation diagram these values are missing.

Furthermore, the coordination energies of one $HO_2^\bullet$ to $[Fe_4S_4(SH)_4]^{2-}$ in D_{2d}-1 form are shown in Figure B.5. We obtain a similar range of coordination energies as for the D_{2d}-2 form.

Qualitatively the changes in the cluster properties are very similar for the two configurations, D_{2d}-1 and D_{2d}-2. Therefore, we can reduce our analysis to only one of the two D_{2d}-symmetric model systems, the D_{2d}-2 form.

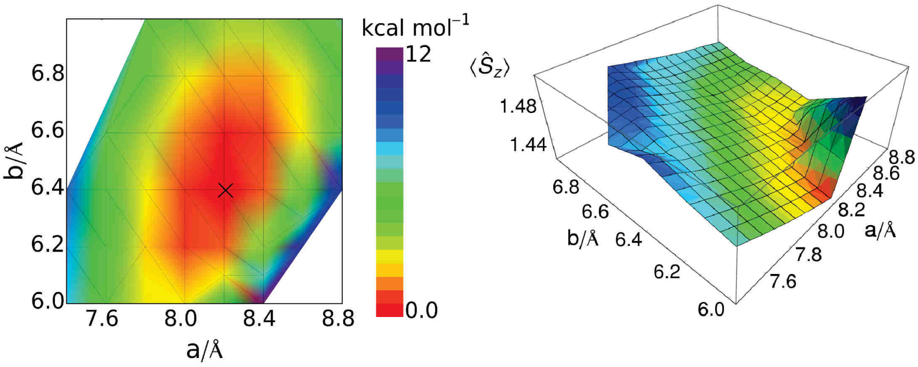


Figure B.3: **D_{2d}-1/H-fixed:** BP86/RI/def2-TZVP/COSMO potential energy surface (left) and plot of local spin $\langle \hat{S}_z \rangle$ expectation values (right, exemplarily of one iron atom with excess of α spin density) of structurally constrained $[Fe_4S_4(SH)_4]^{2-}$ clusters originally in D_{2d}-1 form with fixed hydrogen atoms. The cross indicates the minimum energy structure.

B.3 Results for the S-fixed model system

In Figure B.6 and Figure B.7 the BP86 and B3LYP potential energy surfaces of the S-fixed $[Fe_4S_4(SH)_4]^{1-/2-/3-}$ clusters are presented in Figure B.6 and Figure B.7. Similarly to the H-fixed model system, the potential energy surface is slightly affected by the charge of the cluster, whereby the ideal ligand arrangement of $[Fe_4S_4(SH)_4]^{1-}$ is shifted to about 0.5 Å smaller S-S distances (in a) compared to $[Fe_4S_4(SH)_4]^{3-}$. The potential energies of two structures with fixations of 6.0 Å in a and 6.3 Å and 6.4 Å in b are lower than we expected. In these structures, the HS^- ligands adopt an unusual SFeSH dihedral angle (about 65° instead of 0°).

The BP86 and B3LYP differences in electronic energy between spin-coupling scheme A and B are presented in Figure B.8 for $[Fe_4S_4(SH)_4]^{2-}$ and $[Fe_4S_4(SH)_4]^{3-}$ clusters with fixed S atoms. The ΔE_{el}'s of the two structures with the unusual dihedral angle are left out here. The correlation diagrams of ΔE_{el} for the S-fixed system differ from the results obtained for the H-fixed system. In the S-fixed $[Fe_4S_4(SH)_4]^{3-}$ clusters coupling scheme B is favored for almost all structural constraints and in the $[Fe_4S_4(SH)_4]^{2-}$ cluster coupling scheme B also gets favored for large b distances, which is not observed for the H-fixed model system. Still, the qualitative trend in the ΔE_{el}'s is the same.

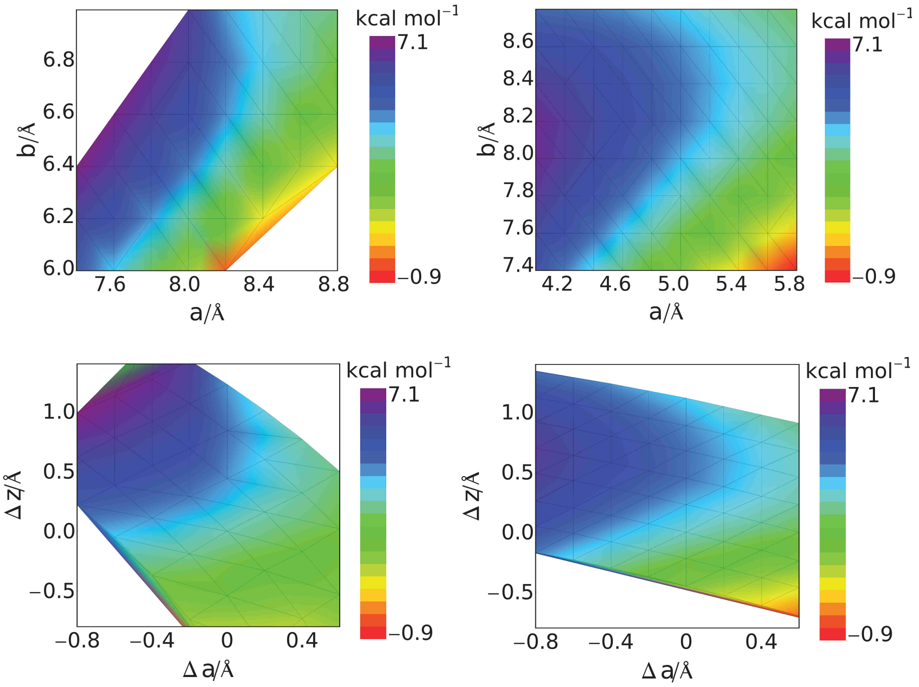


Figure B.4: **D$_{2d}$-1 and D$_{2d}$-2/H-fixed:** BP86/RI/def2-TZVP/COSMO electronic energy differences ΔE_{el} for the coupling schemes A and B of structurally constrained $[Fe_4S_4(SH)_4]^{2-}$ clusters in D$_{2d}$-1 form (left) and D$_{2d}$-2 form (right) (in singlet spin state) with respect to changes in a and b (upper panel) and changes in Δa and Δz (lower panel). The hydrogen atoms of the ligands are kept fixed.

The adiabatic detachment energies, ADE(2−/1−) and ADE(3−/2−), are presented in Figure B.9. The ADE correlation diagrams are similar to the ones obtained for the H-fixed model system, but the overall ADE change for the chosen structural distortion is slightly larger.

The energies for the coordination of one $HO_2^\bullet$ radical to $[Fe_4S_4(SH)_4]^{2-}$ are shown in Figure B.10. Here, for several structures the structure optimization of the $HO_2^\bullet$ radical adduct failed to converge with B3LYP. Therefore, the B3LYP correlation diagram is incomplete.

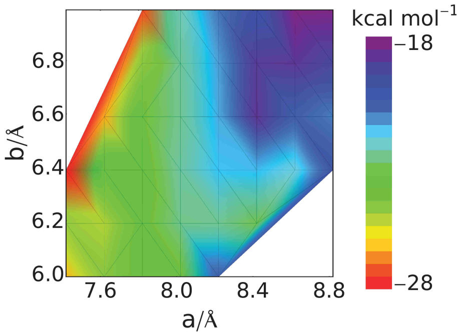


Figure B.5: **D_{2d}-1/H-fixed:** Plot of the coordination energies of distorted, originally D_{2d}-1 symmetric $[Fe_4S_4(SH)_4]^{2-}$ clusters with $HO_2^{\bullet}$. The hydrogen atoms of the ligands are kept fixed at distances *a* and *b* during BP86/RI/def2-TZVP/COSMO structure optimization.

Table B.1: Overview of S(Cys)-S(Cys) and C_β-C_β distances in cysteine-ligated Fe_4S_4 clusters from the PDB. Abbreviations: Res. = resolution, atom = the atoms in the crystal structures between the distances are measured.

PDB entry (Res./Å)	atom	*a*/Å	*b*/Å
Ferredoxins			
1A6L (2.10)	S(Cys)	6.59, 6.59	6.39, 6.14, 6.10, 6.73
	C_β	6.06, 8.11	9.08, 8.14, 8.05, 8.33
2FDN (0.94), cluster 61	S(Cys)	6.53, 6.71	6.38, 6.01, 6.19, 6.44
	C_β	5.74, 8.16	8.90, 8.16, 8.15, 7.80
6FD1 (1.35)	S(Cys)	6.55, 6.65	6.41, 6.04, 6.12, 6.82
	C_β	5.85, 8.12	9.03, 7.99, 7.95, 8.46
2FGO (1.32), cluster 201	S(Cys)	6.55, 6.64	6.37, 6.14, 6.08, 6.86
	C_β	5.68, 8.06	8.91, 8.06, 8.02, 8.37
2Z8Q (1.70)	S(Cys)	6.18, 6.71	6.26, 5.90, 6.16, 6.59
	C_β	6.18, 8.19	8.79, 8.06, 8.13, 8.06
2ZVS (1.65), cluster 101	S(Cys)	6.54, 6.55	6.28, 6.06, 6.11, 6.93
	C_β	6.06, 8.20	6.63, 8.33, 7.81, 8.50
Other Fe-S clusters			
1CP2 (1.93)	S(Cys)	6.55, 6.56	6.21, 6.14, 6.34, 6.44
	C_β	5.65, 5.94	8.79, 8.83, 7.72, 8.69
2C42 (1.78), cluster B2235	S(Cys)	6.18, 5.98	6.45, 6.78, 6.40, 6.59
	C_β	4.87, 7.96	8.83, 7.39, 7.97, 7.81
3C8Y (1.39), cluster 577	S(Cys)	6.45, 6.59	6.30, 6.48, 6.18, 6.76
	C_β	5.25, 8.20	8.99, 8.71, 8.19, 8.39
2E7Z (1.26)	S(Cys)	6.31, 6.14	6.24, 6.37, 6.57, 6.85
	C_β	4.78, 8.09	8.56, 8.66, 7.61, 7.93
2YIV (1.28)	S(Cys)	6.18, 6.27	6.60, 6.40, 6.82, 6.44
	C_β	4.54, 7.26	8.80, 8.80, 8.23, 8.02

Table B.2: Overview of S(Cys)-S(Cys)/C_β-C_β distances in cysteine-ligated Fe_4S_4 clusters of HiPIPs from the PDB. Abbreviations: Res. = resolution, atom = the atoms in the crystal structures between the distances are measured.

PDB entry (Res./Å)	atom	*a*/Å	*b*/Å
HiPIPs			
2AMS (1.40)	S(Cys)	6.38, 5.87	6.52, 6.53, 6.20, 6.55
	C_β	5.14, 8.07	8.55, 6.83, 7.82, 7.95
1B0Y (0.93)	S(Cys)	6.41, 6.01	6.48, 6.52, 6.24, 6.61
	C_β	5.10, 8.25	8.49, 6.71, 7.69, 7.99
1HLQ (1.45)	S(Cys)	6.48, 6.02	6.83, 6.66, 6.38, 6.59
	C_β	5.14, 8.05	8.61, 6.66, 7.71, 7.99
1HPI (1.80)	S(Cys)	6.38, 5.94	6.54, 6.48, 6.11, 6.48
	C_β	5.04, 8.32	8.54, 6.64, 7.62, 7.98
3H31 (1.00)	S(Cys)	6.38, 5.94	6.61, 6.61, 6.17, 6.60
	C_β	5.07, 7.92	8.53, 6.80, 7.90, 7.94
1ISU (1.50)	S(Cys)	6.39, 5.45	6.69, 6.58, 6.13, 6.45
	C_β	5.11, 7.75	8.54, 6.87, 7.85, 7.82

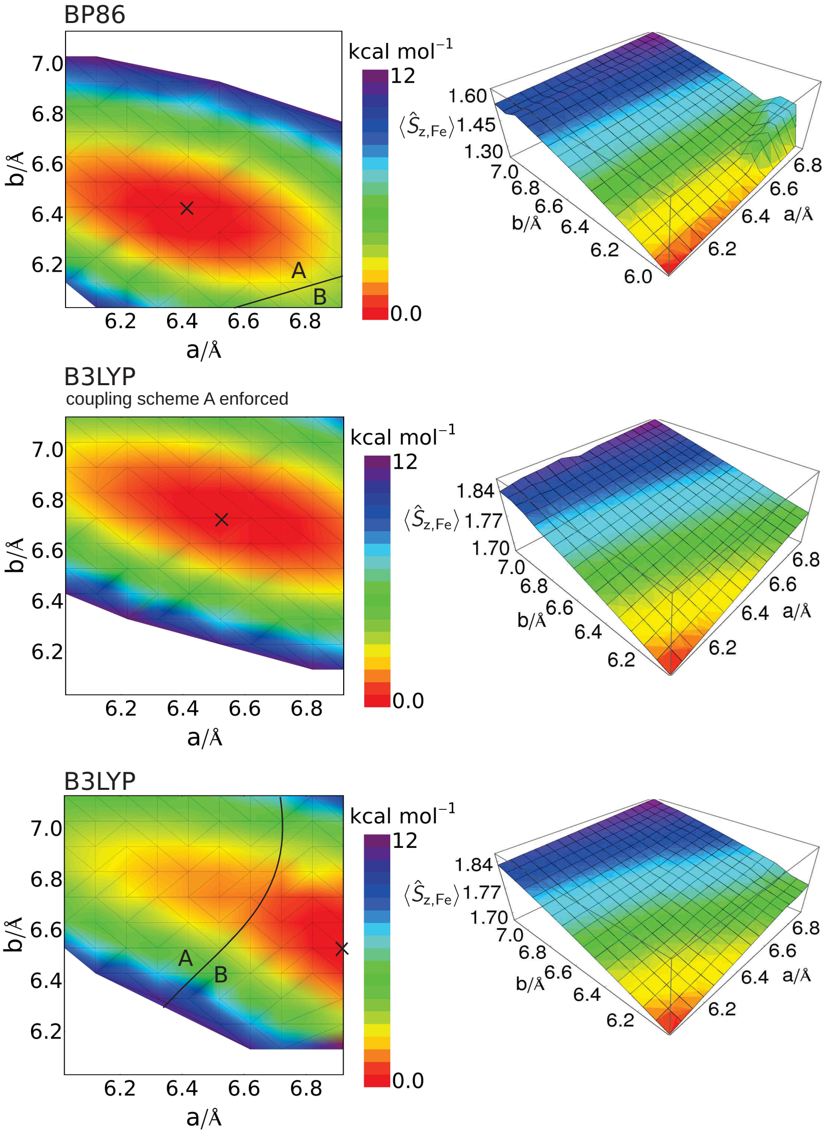


Figure B.6: **D$_{2d}$-2/S-fixed:** Potential energy surfaces (PES's, left) and local spin $\langle \hat{S}_{z,\mathrm{Fe}} \rangle$ expectation values (right, exemplarily for one iron atom with excess of α-spin density) of constrained optimized $[\mathrm{Fe}_4\mathrm{S}_4(\mathrm{SH})_4]^{2-}$ clusters (in open-shell singlet BS spin state) with fixed S(H) atoms. On the top, the PES of BP86/RI/def2-TZVP/COSMO optimized structures with the lowest-energy spin-coupling scheme is plotted. In the middle, the B3LYP/def2-TZVP/COSMO results are depicted for clusters in coupling scheme A, and at the bottom the B3LYP PES for the lowest-energy spin-coupling scheme is plotted. The crosses indicate minimum energy structures.

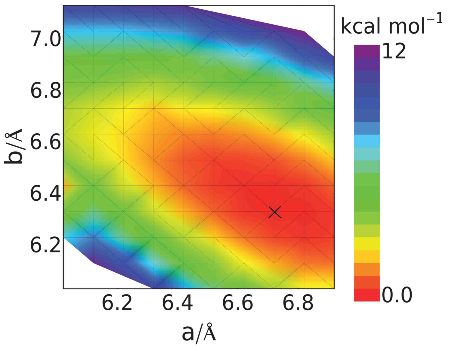


Figure B.7: **D_{2d}-2/S-fixed:** BP86/RI/def2-TZVP/COSMO potential energy surface plot of structurally constrained, originally D_{2d}-2 symmetric $[Fe_4S_4(SH)_4]^{3-}$ (left) and $[Fe_4S_4(SH)_4]^{1-}$ clusters (right) in doublet spin state. The sulfur atoms of the ligands are kept fixed at distances a and b. The crosses indicate minimum energy structures.

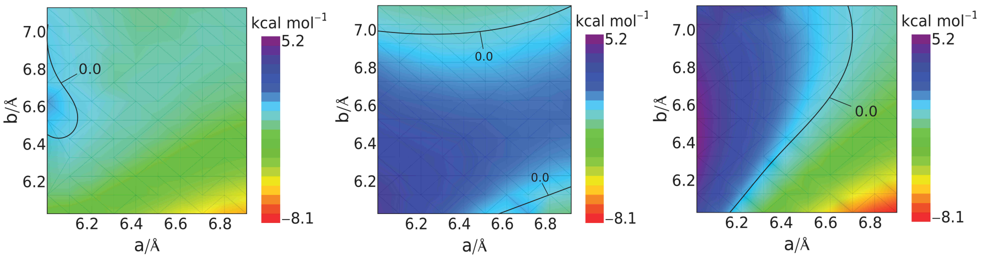


Figure B.8: **D_{2d}-2/S-fixed/different charges:** BP86/RI/def2-TZVP/COSMO electronic energy differences between spin-coupling schemes A and B (ΔE_{el}) of structurally constrained $[Fe_4S_4(SH)_4]^{3-}$ (in doublet spin state, left)/$[Fe_4S_4(SH)_4]^{2-}$ (open-shell singlet, middle) clusters and B3LYP/def2-TZVP/COSMO ΔE_{el}'s of $[Fe_4S_4(SH)_4]^{2-}$ (right). The sulfur atoms of the ligands are kept fixed at distances a and b.

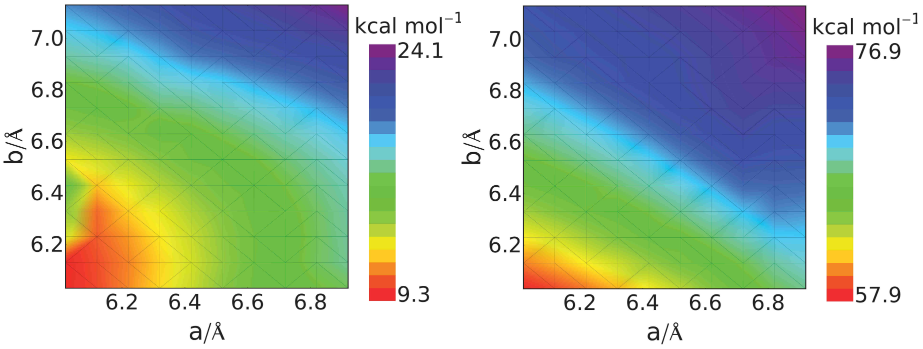


Figure B.9: **D_{2d}-2/S-fixed:** BP86/RI/def2-TZVP/COSMO ADE(3−/2−) (left) and ADE(2−/1−) (right) correlation diagrams for structurally constrained $[Fe_4S_4(SH)_4]^{1-/3-}$ clusters in doublet spin state and $[Fe_4S_4(SH)_4]^{2-}$ clusters in singlet spin state. The sulfur atoms of the ligands are kept fixed at distances *a* and *b*.

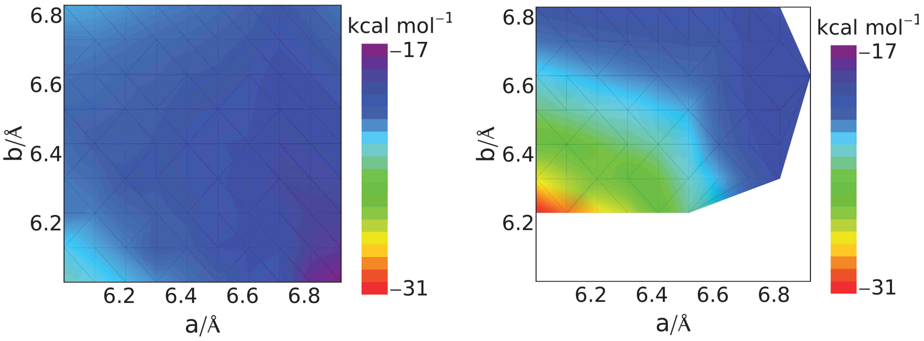


Figure B.10: **D_{2d}-2/S-fixed:** Plot of the coordination energies of structurally constrained $[Fe_4S_4(SH)_4]^{2-}$ clusters with respect to $HO_2^\bullet$ (doublet spin state for $[Fe_4S_4(SH)_4(HO)]^{2-}$) calculated with BP86/RI/def2-TZVP/COSMO (left) and B3LYP/def2-TZVP/COSMO (right). The sulfur atoms of the ligands are kept fixed at distances*a* and *b*.

C Conformer analysis of double-decker-type rotamers

We perform density functional calculations with the program package TURBOMOLE [190] (versions 6.3.1) using Ahlrichs' valence double-ζ SVP [192] and triple-ζ def2-TZVP basis set [129, 193] with polarization functions on all atoms. Unless otherwise noted in the text, we employ the def2-TZVP basis set [129, 193]. BP86 [133, 134], B97-D [127] and TPSS [131, 132, 135, 136] density functionals in combination with the resolution of the identity (RI) technique are chosen. Besides the dispersion correction in B97-D, the DFT-D3 [128] dispersion corrections are employed. The self-consistent-field single-point calculation are considered to be converged when the energy difference between two cycles drops below 10^{-7} hartree. Structure optimizations are considered converged when the length of the geometry gradient is below 10^{-4} a.u. On the TPSS-D3/SVP optimized structures, we perform single-point calculations using the COSMO solvation model [231] with a dielectric constant of $\epsilon = 46.7$ for DMSO.

Second-order Møller-Plesset (MP2) energies of structures **1A**, **1B** and **1C** are obtained with TURBOMOLE (versions 6.4.0) with def2-TZVP [129, 193], def2-TZVPP [129] and def2-QZVPP [130] basis sets.

Gibbs free energies are calculated with FREEH as implemented in TURBOMOLE for a temperature of T = 298.15 K and p = 0.1 MPa. Here, the translational, rotational and vibrational degrees of freedom are approximated by the ideal gas, the quantum rigid rotor and the quantum harmonic oscillator. For the calculation of the BP86 free energies, the wavenumbers are scaled by a factor of 0.9914 [232]. The mode with the imaginary frequency is not included in the calculation of the free energies. It should be noted that the free energy differences (ΔG's) are very similar to the electronic energy differences (ΔE_{el}'s). For this reason, we only report the electronic energy differences in the main text.

In the GAUSSIAN [191] (version 09, Revision C.1) calculations we employ B97-D [127] with the TZVP basis set [192, 193]. We mainly employ the default convergence criteria (scfconv = tight, which means that the energy difference between two cycles is less than 10^{-8} hartree, and optimizations are considered converged when the root mean square force is below 3×10^{-4} a.u.), but investigate the energy difference between these results and the results we obtain with tighter convergence criteria of opt = verytight (root mean square force below 1×10^{-6} a.u.) and grid = UltraFine for some structures.

A vibrational analysis (with Turbomole and Gaussian) is performed on every structure with vanishing gradient, in order to verify that the structure is a minimum or transition state. The intrinsic reaction coordinate plots are created with Mathematica 9.0 by Wolfram Research [229]. Molecular structures are visualized with PyMOL [188].

For the exploratory Born-Oppenheimer molecular dynamics (BOMD) simulations we apply the program CP2K (version 2.4.0/12993) [233]. We solvate the optimized structures **1A** and **1B** by 32 DMSO molecules in a cubic box of cell length 19 Å using packmol [234]. The BOMD is performed for an NVT ensemble at T = 300K under periodic boundary condition. The Gaussian and augmented plane wave method is employed with a double-ζ DZVP basis set in the optimized short-range form and with Goedecker-Teter-Hutter (GTH) pseudo potentials [235, 236] for the core electrons. The BLYP [133, 223] density functional is chosen.

Out of the trajectories over approximately 15 ps, 34 snapshots are chosen, which contain the solute and a reduced number of 19 DMSO molecules. The snapshots are optimized with BP86 under structurally constrained DMSO sulfur and oxygen atom positions and two fixated nitrogen atoms, which are on top of each other in the double-layered structure **1**. The def2-SVP basis set is chosen for the DMSO molecules and the def2-TZVP basis set is chosen for the solute. To confirm that the constraint of one fixated pair of N-N does not lead to an artificial energy increase, we once release the constraint and only fixate all DMSO molecules. The potential energy differences do not change significantly and hence, the chosen constraints are appropriate.

D Reaction pathways of an organic phosphaketene

We perform density functional calculations with the Turbomole [190] (versions 6.3.1 and 6.5) and Gaussian [191] (version 09, Revision C.1) program packages. In the Turbomole DFT calculations we employ Ahlrichs' def2-SVP and def2-TZVP basis set [129, 193]. BP86 [133, 134] and PBE [237, 238] (in combination with the resolution of the identity

(RI) technique) as well as the hybrid functionals PBE0 [239] and B3LYP [221–223] are chosen. Furthermore, second-order Møller-Plesset (MP2) calculations are performed with the def2-TZVPP [129] basis set. Since similar relative energies are obtained for the different quantum-chemical methods, we only discuss the BP86 energies explicitly. The convergence criterion for the self-consistent-field single-point calculations is set to an energy difference of 10^{-7} hartree between the two last energy iterations and structure optimizations are considered converged when the length of the geometry gradient drops below 10^{-4} hartree/bohr.

In the GAUSSIAN calculations we employ BP86 [133, 134] with the TZVP basis set [193]. We choose the default convergence criteria (scfconv = tight, which means that the energy difference between two iteration steps is less than 10^{-8} hartree, and optimizations are considered converged when the root mean square force is below 3×10^{-4} hartree/radians).

We locate the transition state structures by performing a constrained optimization scan with GAUSSIAN, selecting the highest energy structure, calculating the vibrational frequencies (with AOFORCE) and following the lowest eigenvalue mode with the trust-radius image based eigenvector-following procedure implemented in TURBOMOLE. A preoptimization is carried out with def2-SVP basis set and on the converged structure we perform a TS calculation with def2-TZVP basis set.

We perform a vibrational analysis on every structure with vanishing gradient, in order to verify that the structure is a minimum or transition state. The intrinsic reaction coordinates are calculated with GAUSSIAN to verify that the transition state structure between the correct local minima is found. Molecular structures are visualized with PYMOL [188].

E Mode-Tracking based Transition-State Optimizations

E.1 Choice of the Mode-Tracking guess modes

In this section, we investigate whether an initial-guess mode obtained from an LST pathway is sufficient to find the transition-state structure.

The isomerization reaction from H_2CO to trans-HOCH serves as an example. We create an internal LST path of 11 points including the reactant and product structures and always calculate the full Hessian (with Turbomole) and the LST-based mode obtained from Mode-Tracking. Then, we compare the mode obtained by Mode-Tracking with the lowest frequency ones calculated by Turbomole's aoforce tool to probe whether MTsearch chooses the mode which is closest to the transition mode. In Table E.3 the results are reported. We highlight the modes that are closest to the transition mode at each initial-guess structure. The modes based on the LST guess mode correspond to the TS mode for all structures. Of course, the performance of the TS search strongly depends on the LST path, which is not necessarily close to the intrinsic reaction coordinates.

After the orthogonal optimization steps (i.e., before the next Mode-Tracking calculation) the next guess mode is by default the converged mode of the previous Mode-Tracking calculation. For some model systems, it might be useful to steer the TS localization not only by providing the first eigenvector, but several eigenvectors defined by a sequence of structures. Then, one may circumvent the problem that EVF follows the wrong mode and the actual TS mode is lost during the TS optimization. Therefore, an alternative TS-optimization procedure is implemented: in addition to the initial-guess mode we aligned the following structures during the optimization with the LST structures and always selected the guess mode of the structure with lowest root mean square deviation (RMSD) compared to the actual structure for the subsequent Mode-Tracking calculation.

An approximate transition pathway can also be set up manually according to chemical insights. One may also employ a haptic device for a fast and simple creation of an approximate transition pathway [185–187].

It is important to note that for large reactive systems, which may feature many similar modes, one might easily choose a mode that does not lead to the TS, by LST or by other methods. Since no general criterion exists to assess whether a mode is still pointing towards the desired path, for manually constructed guess modes, we recommend to ensure that the first converged mode resembles the mode of interest.

The manually chosen guess modes for locating both TSs from the trans-hydroxamethylene structure are summarized in Table E.4.

Table E.3: Mode-Tracking (MT) and TURBOMOLE calculated BP86/RI/def2-SV(P) frequencies in cm^{-1} of the first 8 structures out of an 11-point LST pathway between H_2CO and trans-HOCH. Here, a convergence criterion of 1×10^{-8} a.u. is chosen for the orbital optimization. The highlighted frequencies are the ones of those modes that are closest to the reaction pathway. Note that 'frequencies' simply denotes the square root of the Hessian eigenvalues and must not be mistaken for a vibrational frequency in case of non-stationary structures. Moreover, we have added a minus sign to the imaginary square root of a negative Hessian eigenvalue in order to highlight such imaginary frequencies. Initial guess modes were obtained from an LST path. Here (and only here), a single-point electronic structure calculation was carried out for each point on the LST path before starting MTSEARCH. We found that this can affect the number of steps needed in the initial eigenvector following run, but does not affect the final results.

Initial-guess structure	MT Freq.	TURBOMOLE, full Hessian calculation Freq. 1	Freq. 2	Freq. 3	Freq. 4
LST1=Reactant	+1814	+1144	$\boxed{+1232}$	+1500	$\boxed{+1814}$
LST2	+1473	+1147	$\boxed{+1287}$	$\boxed{+1473}$	+1733
LST3	+1342	+1088	+1282	$\boxed{+1343}$	+1606
LST4	+1407	+825	+988	$\boxed{+1252}$	$\boxed{+1402}$
LST5	+1058	−i580	+851	$\boxed{+1058}$	+1228
LST6	−i754	−i1133	$\boxed{-i757}$	+696	+1252
LST7	−i1793	$\boxed{-i1793}$	−i328	+693	+1355
LST8	−i1779	$\boxed{-i1778}$	+846	+901	+1401

E.2 Newton–Raphson step length and maximum number of orthogonal optimization steps

In general, in the EVF procedure a smaller step size and a smaller number of orthogonal optimization steps leads to a structure more similar to the previous one. In other words, MTSEARCH will follow the initially chosen eigenvector as closely as possible if the Newton–Raphson step length and the number of orthogonal optimization steps is small. However, if the first step size is too small, the optimization might take longer to leave the

Table E.4: Initial starting distortions of trans-hydroxamethylene constructed manually for application in MTsearch.

	x	y	z	x	y	z
C	0	0	0	0	0	0
O	0	0	0	0	0	0
H	0.1	0	−0.6	0	0.1	0
H	0	0	0	0	0.1	0

energy minimum. Moreover, for exploring the potential energy surface, a larger step may be useful to locate transition-state structures. The first Newton–Raphson step size should not be increased too much, because then the structure can not adapt to this distortion during the orthogonal optimization and in the worst case the self-consistent-field calculation does not converge. We investigate the effect of the first Newton–Raphson step length on the convergence of the TS search in the C_2H_6 and H_2CO model systems. For this, we apply (first) Newton–Raphson steps in the range from 0.5 to 2.5 $Å/\sqrt{\text{amu}}$ within the MTsearch algorithm starting from minimum-energy structures. We report the computing times (wall times) for locating the TS in Table E.5.

The first Newton–Raphson step influences the computing times significantly. For the calculation of the rotational barrier of C_2H_6 starting from the minimum-energy structure, one can apply relatively large first Newton–Raphson steps up to about 2.5 $Å/\sqrt{\text{amu}}$ and thereby decrease the computing time significantly. However, in the trans-HOCH↔H_2CO isomerization a distortion of 2.0 to 2.5 $Å/\sqrt{\text{amu}}$ is already too large, because MTsearch then overlooks the TS and converges to the opposite minimum-energy structure. We conclude that first Newton–Raphson step lengths between 1.0 and 1.5 $Å/\sqrt{\text{amu}}$ is a reasonable choice for these types of reactions.

Table E.5: Computing times in min:sec required for the TS optimization of the C_2H_6 rotational transition-state barrier and the trans-HOCH↔H_2CO isomerization with MTsearch and Turbomole eigenvector following for a given first Newton–Raphson (NR) step size (in Å/$\sqrt{\text{amu}}$). All calculations have been performed on a blade system featuring two six-core AMD Opteron 2435 processors (i.e., a total of 12 cores). The computing times for optimizations with a maximum number of orthogonal optimization steps of 10 are given first and after the slash the computing times for a maximum number of orthogonal optimization steps of 3 are given. BP86/RI/def2-SVP (C_2H_6) and def2-SV(P) (H_2CO) calculations have been performed with 12 core processors, MPI-parallelized, and an SCF convergence criterion of $1{\times}10^{-7}$ a.u. The calculations were considered to be converged when the gradient norm was below $1{\times}10^{-3}$ a.u. 'No' denotes that with the given setting the optimization does not converge to the TS.

	Length of 1st NR step in Å/$\sqrt{\text{amu}}$				
Model system	0.5	1.0	1.5	2.0	2.5
C_2H_6	10:04/14:11	8:35/9:58	6:31/8:13	4:05/4:15	2:22/3:54
trans-HOCH↔H_2CO	No	No/7:16	No/7:14	No	No/9:23

E.3 Convergence criteria for MTsearch

In Table E.6 we compare the structures and energies of MTsearch optimized TSs applying convergence criteria for the gradient norm of 10^{-4} a.u. and 10^{-3} a.u.

Since the RMSD and the energy differences between the TSs converged according to the different criteria are small, we conclude that a convergence criterion of 10^{-3} a.u. for the gradient norm is sufficiently accurate, which is in accordance with the default threshold for a Turbomole structure optimization.

Table E.6: Structural (RMSD) and energy (ΔE) differences between TSs optimized with MTSEARCH applying convergence criteria for the gradient norm of 10^{-4} a.u. and 10^{-3} a.u. BP86/RI/def2-SVP (C_2H_6) and def2-SV(P) (H_2CO) calculations have been performed with an SCF convergence criterion of 10^{-7} a.u.

Model system	RMSD in Å	ΔE in kcal/mol
H_2CO, TS-1 (from trans-HOCH minimum)	0.001	0.00
C_2H_6 (from LST2)	0.056	0.09

E.4 TS-search examples

E.4.1 C_2H_6

For the second structure of the 6-node LST we obtain a lowest frequency of 245.5 cm^{-1} (TURBOMOLE) and 251.3 cm^{-1} (MTSEARCH), respectively. The overlap between this lowest-frequency eigenvector of the second LST structure and the optimized TS mode still amounts to 93.6%. The third structure of the LST is very close to the TS. An eigenvector with an overlap of 99.5% with the TS mode and an imaginary frequency of −i195.1 cm^{-1} (TURBOMOLE) and −i185.4 cm^{-1} (MTSEARCH), respectively, is obtained. The imaginary mode of the TS is −301.1 cm^{-1}.

E.4.2 H_2CO

The lowest eigenvalue of the Hessian of the optimized trans-hydroxymethylene structure corresponds to a wavenumber of 1100.71 cm^{-1}. An EVF procedure along this eigenvector does not lead to the TS, but the optimization stays at the minimum-energy structure. We introduce small distortions in the minimum-energy structure by constructing LST pathways with 100, 50, 25, 10, 5, or 4 nodes, i.e., introduce distortions to the minimum-energy structure of $1/100^{th}$, $1/50^{th}$, $1/10^{th}$, $1/5^{th}$ and $1/4^{th}$ of the total structural change from cis-hydroxymethylene to trans-hydroxymethylene. We choose the second node (i.e., the first distorted structure, since node 1 is equal to the educt structure) as starting structure for a TS localization and calculate its lowest-frequency eigenvector.

Only for the second structure of the 4-node LST we obtain an imaginary frequency, but the EVF procedure localizes the TS (with a frequency of $-i1559.28$ cm^{-1}) also for the structures obtained from the LST based on 5 to 100 nodes. The second structure in the 100-node LST is very similar to the trans-hydroxymethylene structure, their RMSD is only 0.008 Å.

E.4.3 [Mo(MeNCH$_2$CH$_2$)$_3$N]+N$_2$H$_4$

- Turbomole calculations:

  In structure **Scan9**, the lowest frequency of the Hessian amounts to $-i33.4$ cm^{-1} and the eigenvector is only to a very small amount localized on the proton that is transfered. Therefore, the contribution of the eigenvector on the proton diminishes during the EVF procedure and finally the optimization leads to a minimum-energy structure (the hydrazine-bound Schrock complex). Also in structure **Scan11**, the part of the eigenvector which is located on the proton is very small compared to the rest of the eigenvector contributions and the optimization leads to the hydrazine-bound Schrock complex.

- MTSearch calculations:

  For **Scan9**, an overall number of 471 steps was required for convergence, and for **Scan11**, it was 75 steps.

F Additional Reaction Networks

In Figures F.11, F.12, F.13, F.20, F.21 and F.22 the reaction networks, in which the energies of all structures are larger than the energy of the N$_2$ bound Schrock model complex are reported. We assume lutidine-H$^+$ (lutidine = 2,6-dimethylpyridine) as proton source (the reference energy for the proton is −237.7 kcal/mol) and Cp*$_2$Cr (Cp* is 1,2,3,4,5-Pentamethylcyclopentadienyl) as electron donor (the reference energy for the electron amounts to −103.7 kcal/mol). In detail, these high-energy reaction networks are the two-fold protonated and 2-e$^-$ reduced species, the three-fold protonated species with a total charge of 2+ and 3+ and the four-fold protonated species with total charges of 2+, 3+ and 4+. Although the four-fold protonates species with a neutral or singly positive charge have lower energies than the energy of the N_2 bound Schrock model complex, these reaction networks are reported here (Figures F.14, F.17), because the [Mo]-N intermediates with NH_3 dissociated from the complex are significantly lower in energy and thus, we assume that the four-fold protonated clusters in which the [Mo]-N_2 fragment is still present (i.e., the N-N bond is not broken) are not relevant. Moreover, in contrast to the cutoff criterium chosen for the reaction networks in the main Chapter, we have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures, which reduces the number of nodes significantly.

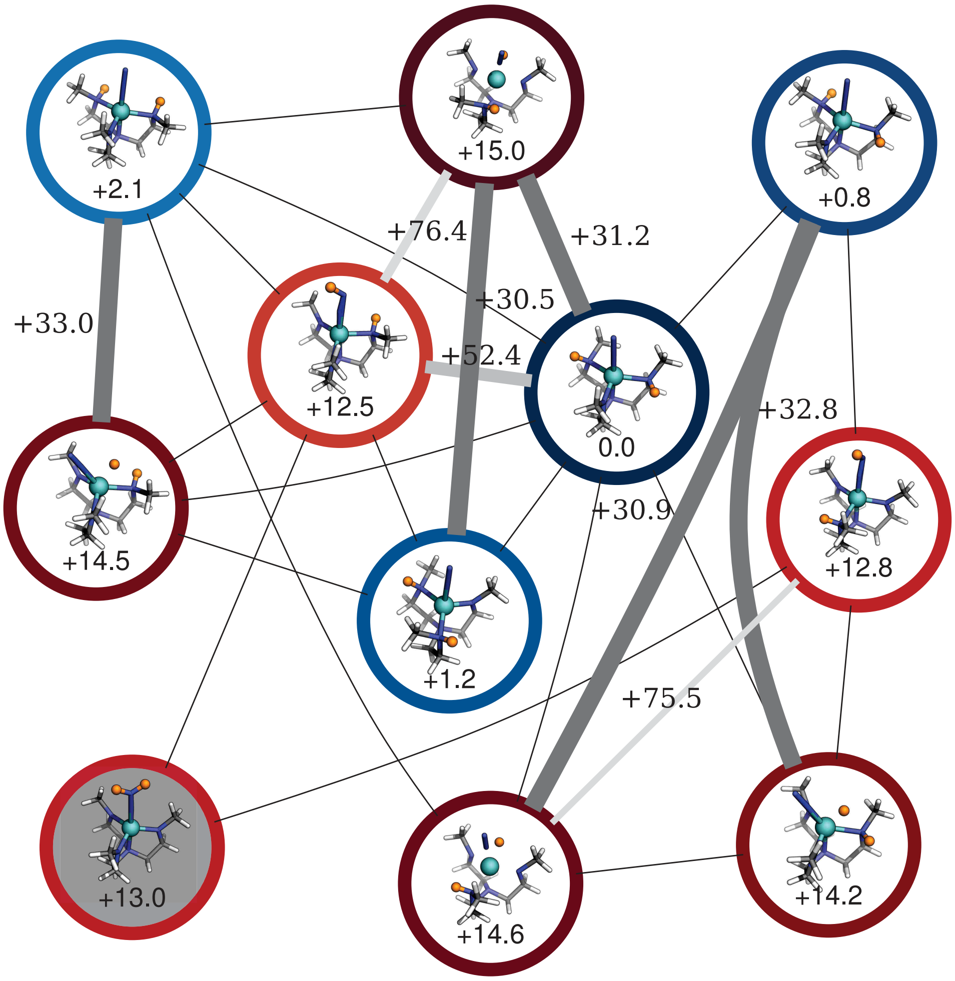


Figure F.11: BP86/RI/def2-SV(P) reaction network for the two-fold protonated intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of 2+ elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

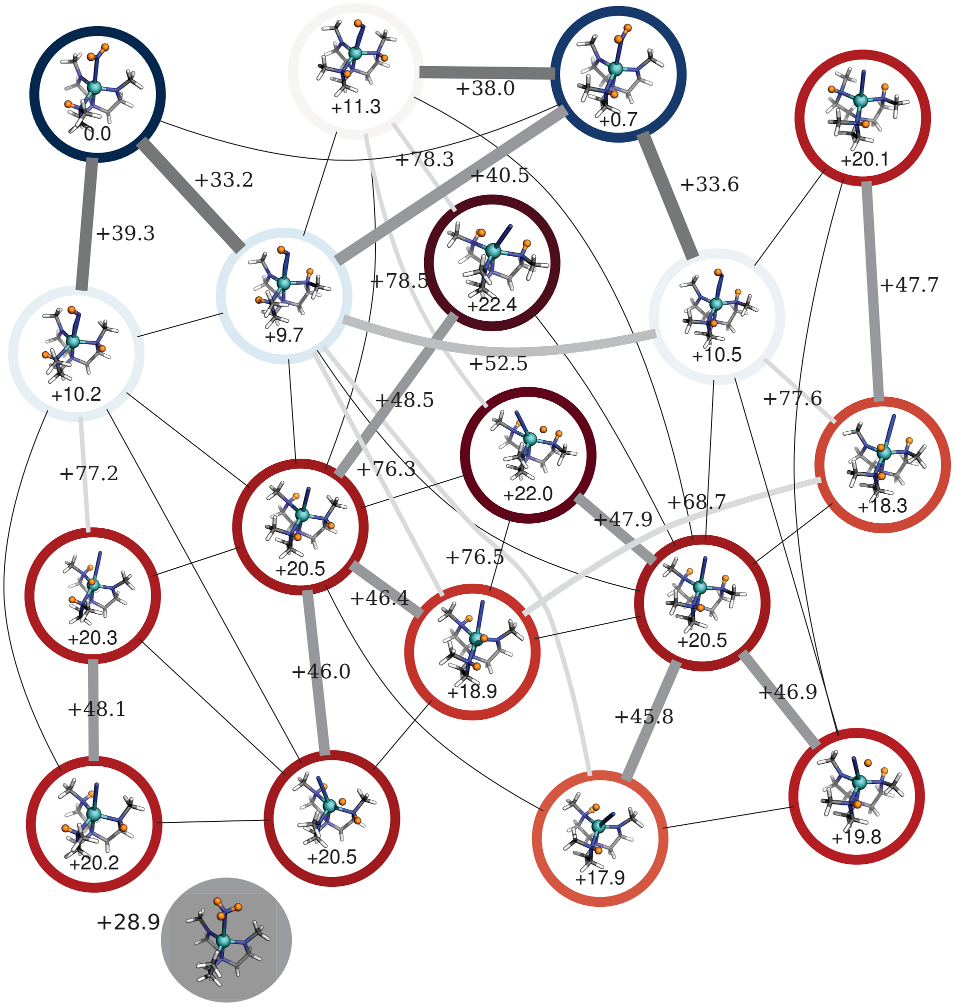

0.0
+11.3
+38.0
+0.7
+20.1
+33.2
+78.3
+40.5
+33.6
+39.3
+78.5
+22.4
+47.7
+9.7
+10.5
+10.2
+52.5
+48.5
+77.6
+76.3
+22.0
+18.3
+77.2
+68.7
+47.9
+20.5
+76.5
+46.4
+20.3
+20.5
+18.9
+46.0
+48.1
+46.9
+45.8
+19.8
+20.2
+20.5
+17.9
+28.9

Figure F.12: BP86/RI/def2-SV(P) reaction network for the three-fold protonated intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of 2+ elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

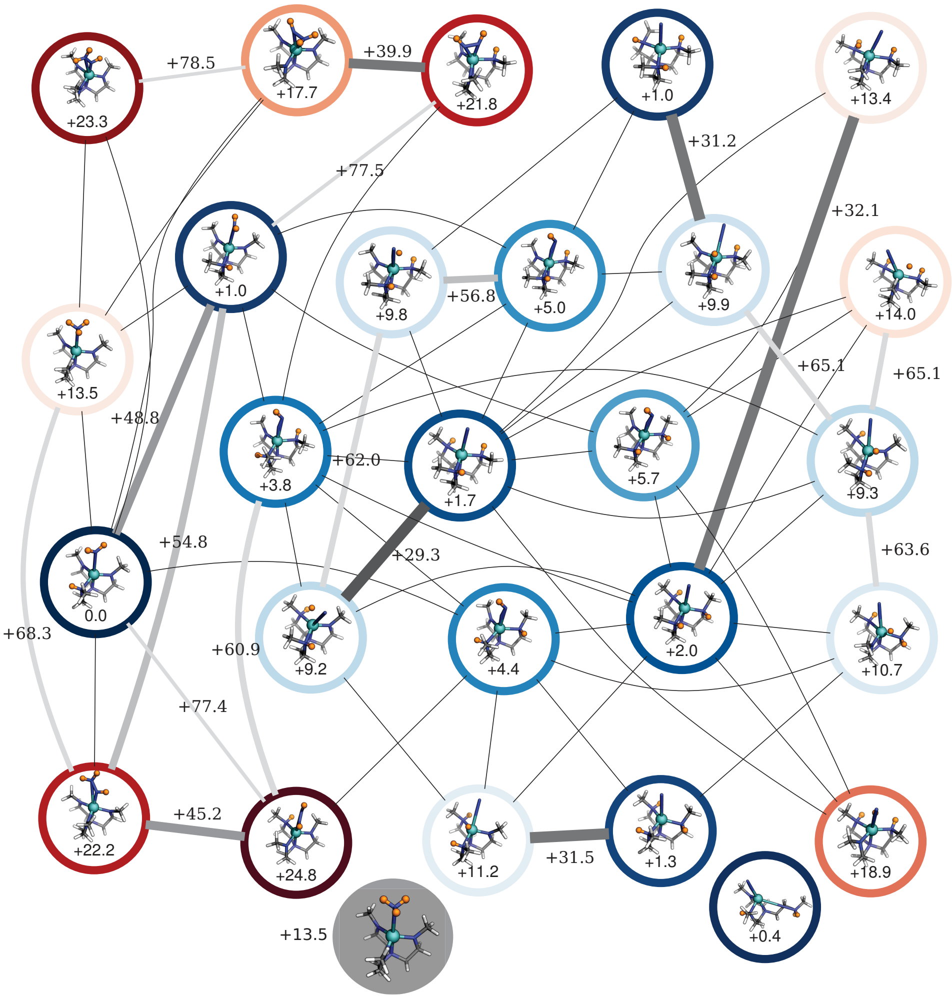

+78.5
+39.9
+23.3
+17.7
+21.8
+1.0
+13.4
+31.2
+77.5
+32.1
+1.0
+9.8
+56.8
+5.0
+9.9
+14.0
+13.5
+48.8
+65.1
+65.1
+3.8
+62.0
+1.7
+5.7
+9.3
+54.8
+29.3
+63.6
0.0
+68.3
+60.9
+9.2
+4.4
+2.0
+10.7
+77.4
+45.2
+22.2
+24.8
+11.2
+31.5
+1.3
+18.9
+13.5
+0.4

Figure F.13: Reaction network for the three-fold protonated intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of 3+ elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

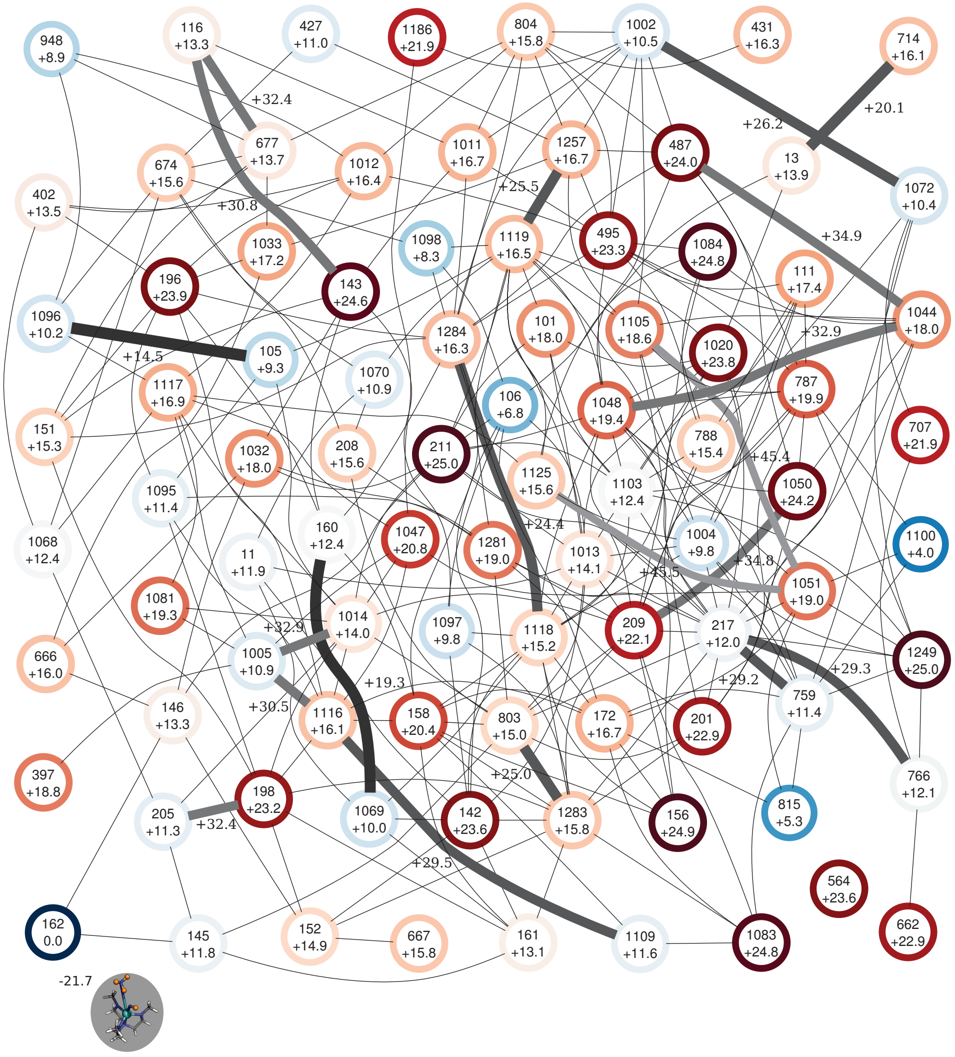

948 +8.9
116 +13.3
427 +11.0
1186 +21.9
804 +15.8
1002 +10.5
431 +16.3
714 +16.1
+32.4
+26.2
+20.1
674 +15.6
677 +13.7
1012 +16.4
1011 +16.7
1257 +16.7
487 +24.0
13 +13.9
1072 +10.4
402 +13.5
+30.8
+25.5
1033 +17.2
1098 +8.3
1119 +16.5
495 +23.3
1084 +24.8
+34.9
196 +23.9
143 +24.6
111 +17.4
1096 +10.2
101 +18.0
1105 +18.6
1284 +16.3
1044 +18.0
+32.9
1020 +23.8
+14.5
105 +9.3
1070 +10.9
1117 +16.9
106 +6.8
1048 +19.4
787 +19.9
151 +15.3
788 +15.4
707 +21.9
1032 +18.0
208 +15.6
211 +25.0
+45.4
1125 +15.6
1103 +12.4
1050 +24.2
1095 +11.4
+24.4
160 +12.4
1047 +20.8
1281 +19.0
1013 +14.1
1004 +9.8
1100 +4.0
1068 +12.4
11 +11.9
+45.5
+34.8
1051 +19.0
1081 +19.3
+32.9
1014 +14.0
1097 +9.8
1118 +15.2
209 +22.1
217 +12.0
666 +16.0
1005 +10.9
1249 +25.0
+29.3
+29.2
+19.3
759 +11.4
146 +13.3
+30.5
1116 +16.1
158 +20.4
803 +15.0
172 +16.7
201 +22.9
397 +18.8
766 +12.1
+25.0
205 +11.3
+32.4
198 +23.2
1069 +10.0
142 +23.6
1283 +15.8
156 +24.9
815 +5.3
+29.5
564 +23.6
162 0.0
145 +11.8
152 +14.9
667 +15.8
161 +13.1
1109 +11.6
1083 +24.8
662 +22.9
-21.7

Figure F.14: BP86/RI/def2-SV(P) reaction network for the four-fold protonated and 4 e$^-$-reduced intermediates of the N_2-bound generic model complex of the Schrock catalyst. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures and 30 kcal/mol for the activation energies between a minimum-energy structure and the respective TS. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

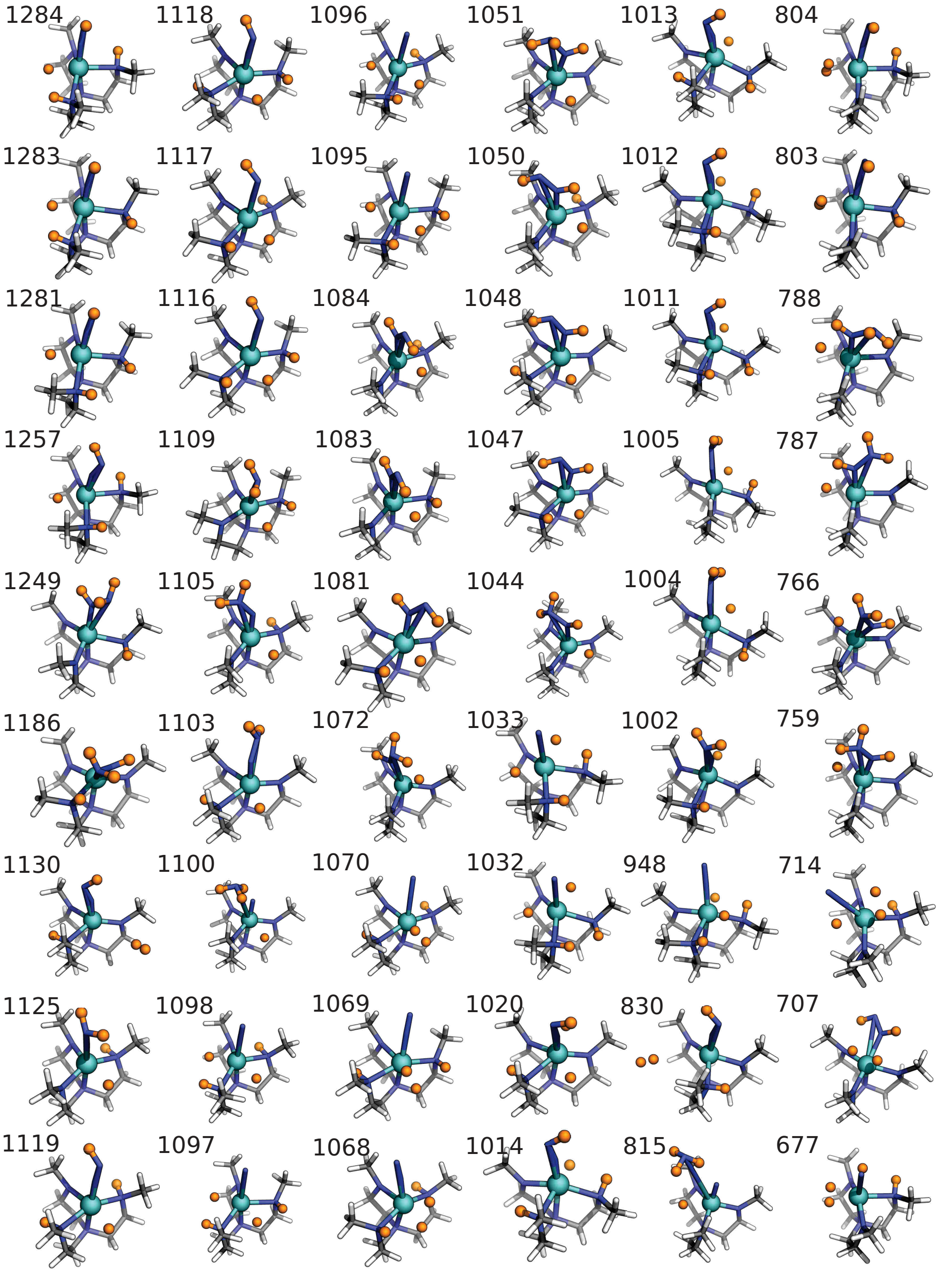

1284
1118
1096
1051
1013
804
1283
1117
1095
1050
1012
803
1281
1116
1084
1048
1011
788
1257
1109
1083
1047
1005
787
1249
1105
1081
1044
1004
766
1186
1103
1072
1033
1002
759
1130
1100
1070
1032
948
714
1125
1098
1069
1020
830
707
1119
1097
1068
1014
815
677

Figure F.15: **Part I**: BP86/RI/def2-SV(P) optimized structures of the four-fold protonated and 4 e$^-$-reduced intermediates and their identification number.

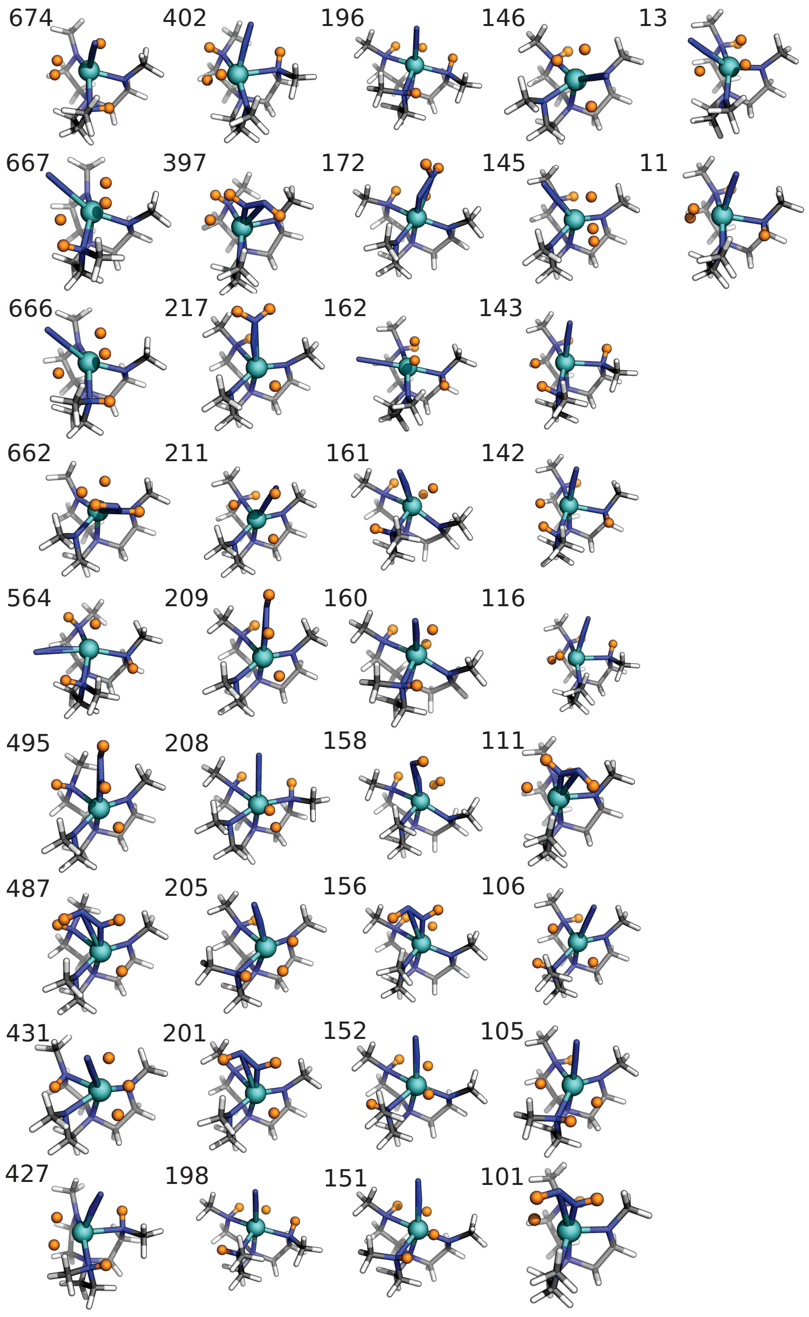

674
402
196
146
13
667
397
172
145
11
666
217
162
143
662
211
161
142
564
209
160
116
495
208
158
111
487
205
156
106
431
201
152
105
427
198
151
101

Figure F.16: **Part II**: BP86/RI/def2-SV(P) optimized structures of the four-fold protonated and 4 e$^-$-reduced intermediates and their identification number.

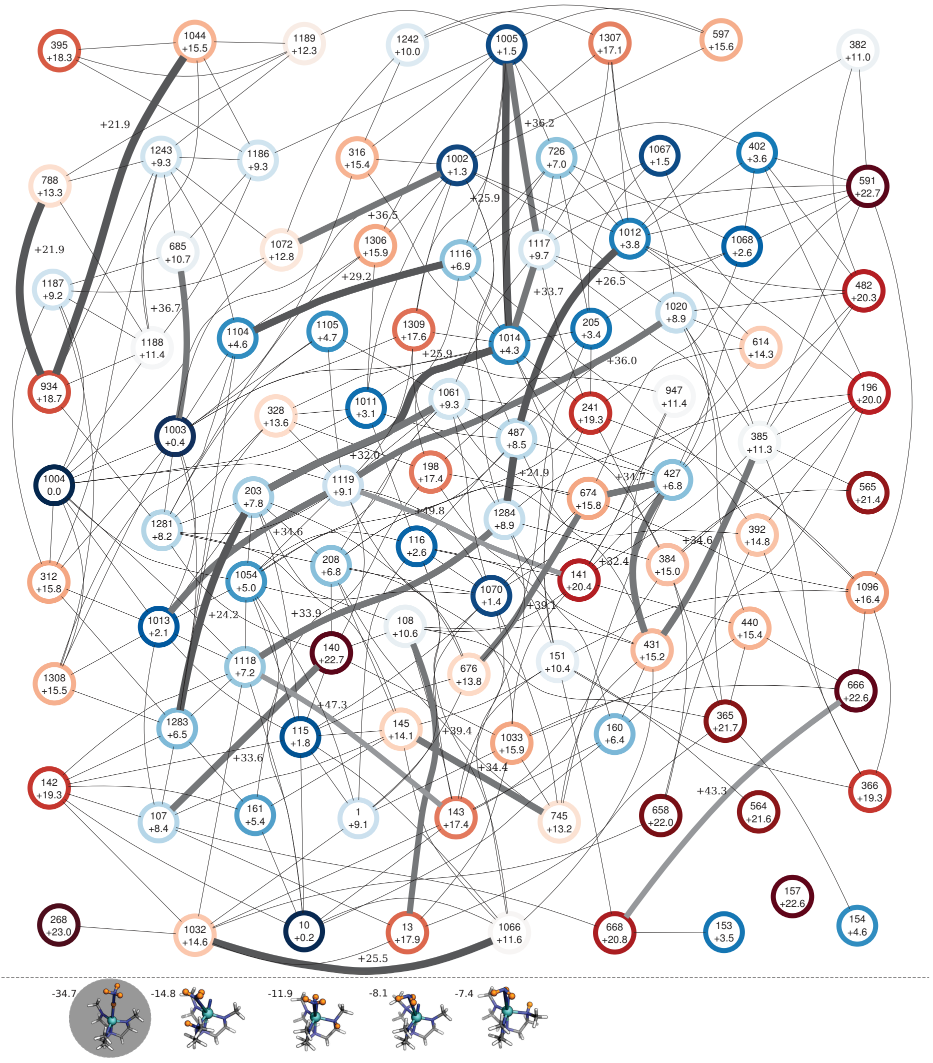

395 +18.3
1044 +15.5
1189 +12.3
1242 +10.0
1005 +1.5
1307 +17.1
597 +15.6
382 +11.0
+21.9
1243 +9.3
1186 +9.3
316 +15.4
1002 +1.3
+36.2
726 +7.0
1067 +1.5
402 +3.6
591 +22.7
788 +13.3
+25.9
+36.5
685 +10.7
1072 +12.8
1306 +15.9
1116 +6.9
1117 +9.7
1012 +3.8
1068 +2.6
+21.9
1187 +9.2
+29.2
+33.7
+26.5
482 +20.3
+36.7
1188 +11.4
1104 +4.6
1105 +4.7
1309 +17.6
1014 +4.3
205 +3.4
1020 +8.9
614 +14.3
+25.9
+36.0
934 +18.7
947 +11.4
196 +20.0
1011 +3.1
1061 +9.3
328 +13.6
241 +19.3
1003 +0.4
487 +8.5
385 +11.3
+32.0
198 +17.4
+24.9
+34.7
427 +6.8
1004 0.0
203 +7.8
1119 +9.1
674 +15.8
565 +21.4
1284 +8.9
+49.8
1281 +8.2
+34.6
392 +14.8
116 +2.6
+34.6
208 +6.8
+32.4
384 +15.0
141 +20.4
312 +15.8
1054 +5.0
1070 +1.4
1096 +16.4
+39.1
+24.2
+33.9
1013 +2.1
108 +10.6
440 +15.4
140 +22.7
151 +10.4
431 +15.2
1118 +7.2
676 +13.8
1308 +15.5
666 +22.6
+47.3
365 +21.7
1283 +6.5
145 +14.1
115 +1.8
+39.4
1033 +15.9
160 +6.4
+33.6
+34.4
142 +19.3
+43.3
366 +19.3
564 +21.6
107 +8.4
161 +5.4
1 +9.1
143 +17.4
745 +13.2
658 +22.0
157 +22.6
268 +23.0
1032 +14.6
10 +0.2
13 +17.9
1066 +11.6
668 +20.8
153 +3.5
154 +4.6
+25.5
-34.7
-14.8
-11.9
-8.1
-7.4

Figure F.17: BP86/RI/def2-SV(P) reaction network for the four-fold protonated and 3 e$^-$-reduced intermediates of the N$_2$-bound generic model complex of the Schrock catalyst at a total charge of +1 elementary unit. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures and 30 kcal/mol for the activation energies between a minimum-energy structure and the respective TS. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

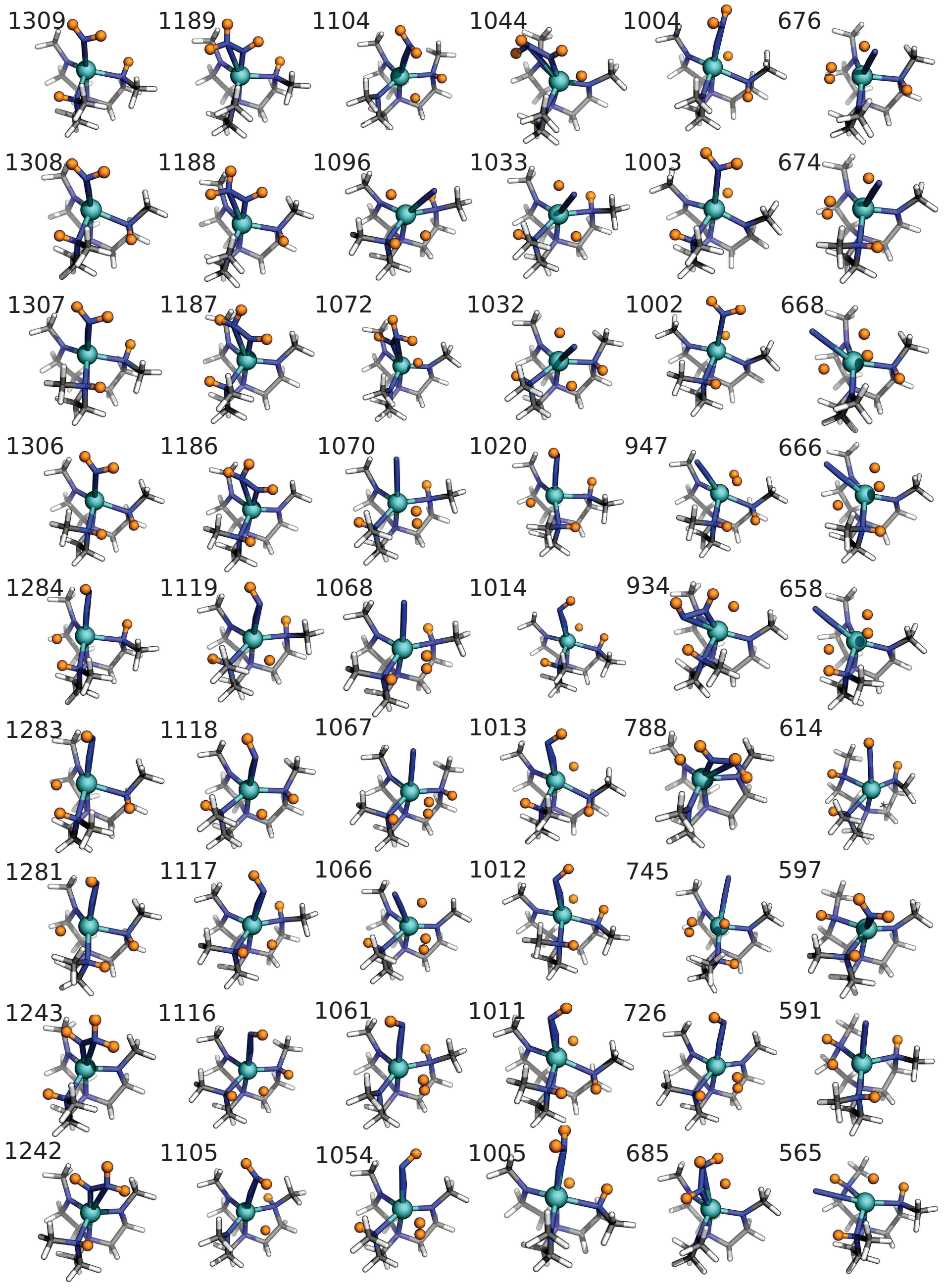

1309
1189
1104
1044
1004
676
1308
1188
1096
1033
1003
674
1307
1187
1072
1032
1002
668
1306
1186
1070
1020
947
666
1284
1119
1068
1014
934
658
1283
1118
1067
1013
788
614
1281
1117
1066
1012
745
597
1243
1116
1061
1011
726
591
1242
1105
1054
1005
685
565

Figure F.18: **Part I**: BP86/RI/def2-SV(P) optimized structures of the four-fold protonated and 3 e^--reduced intermediates (total charge of +1 elementary unit) and their identification number.

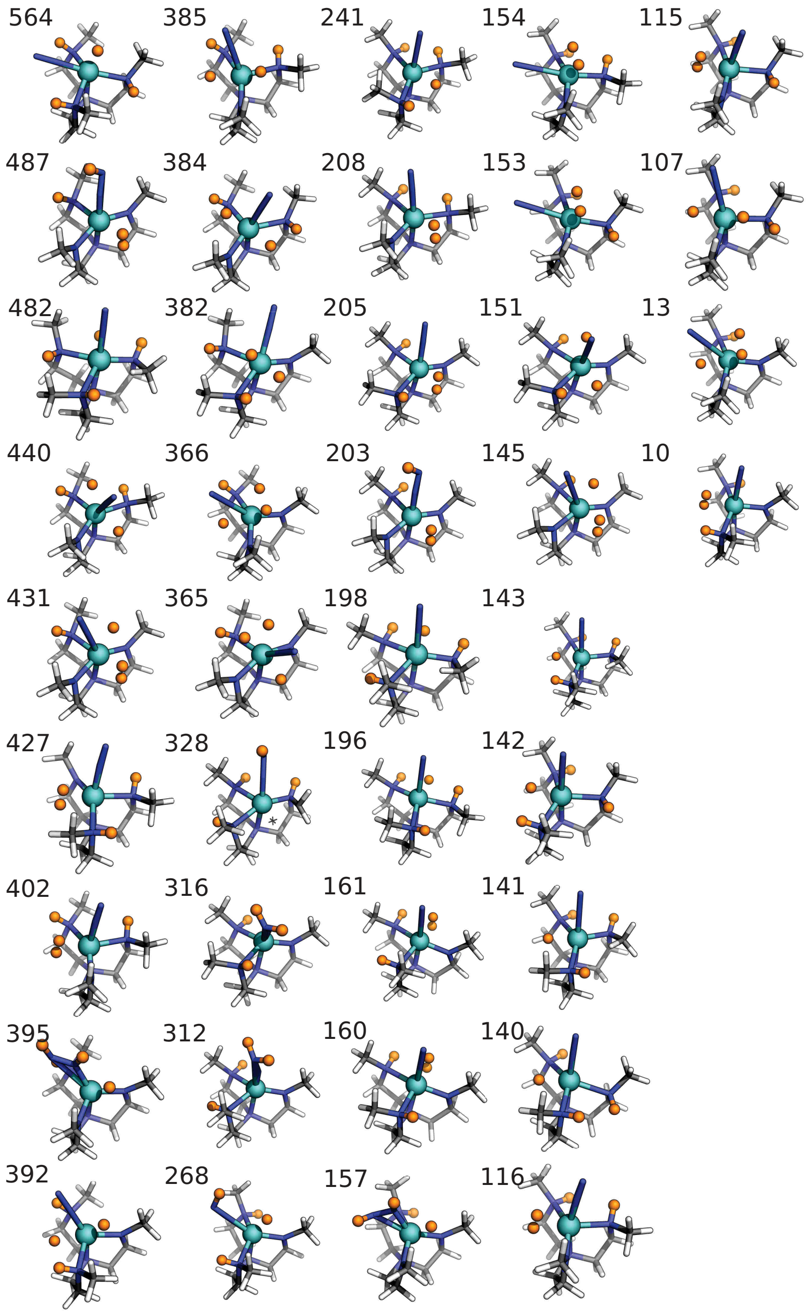

564
385
241
154
115
487
384
208
153
107
482
382
205
151
13
440
366
203
145
10
431
365
198
143
427
328
196
142
402
316
161
141
395
312
160
140
392
268
157
116

Figure F.19: **Part II**: BP86/RI/def2-SV(P) optimized structures of the four-fold protonated and 3 e$^-$-reduced intermediates (total charge of +1 elementary unit) and their identification number.

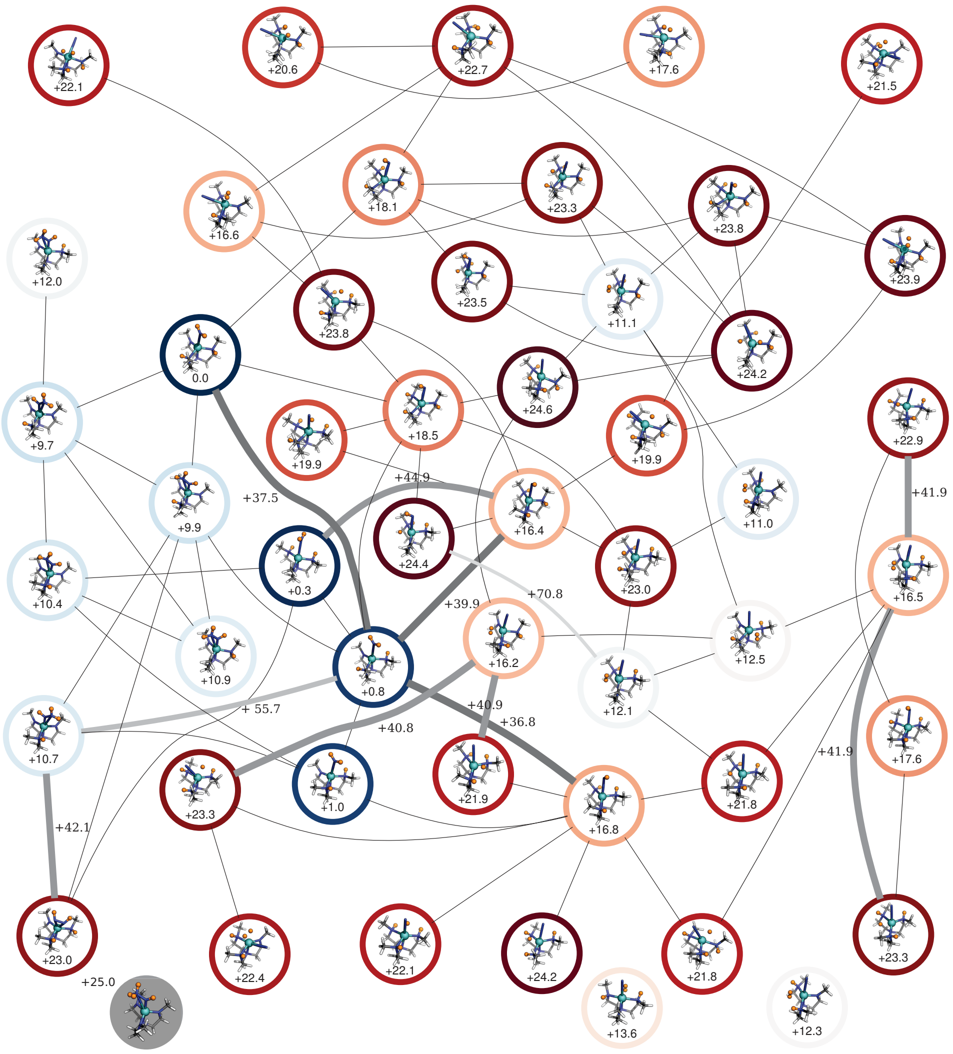

+22.1
+20.6
+22.7
+17.6
+21.5
+18.1
+23.3
+16.6
+23.8
+12.0
+23.9
+23.5
+11.1
+23.8
0.0
+24.2
+24.6
+9.7
+18.5
+19.9
+19.9
+22.9
+44.9
+37.5
+41.9
+9.9
+11.0
+16.4
+24.4
+0.3
+23.0
+16.5
+70.8
+39.9
+10.4
+16.2
+12.5
+10.9
+0.8
+12.1
+ 55.7
+40.9
+36.8
+40.8
+10.7
+41.9
+17.6
+23.3
+1.0
+21.9
+21.8
+16.8
+42.1
+23.0
+22.4
+22.1
+24.2
+21.8
+23.3
+25.0
+13.6
+12.3

Figure F.20: BP86/RI/def2-SV(P) reaction network for the four-fold protonated and $2\ e^-$-reduced intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of $+2$ elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

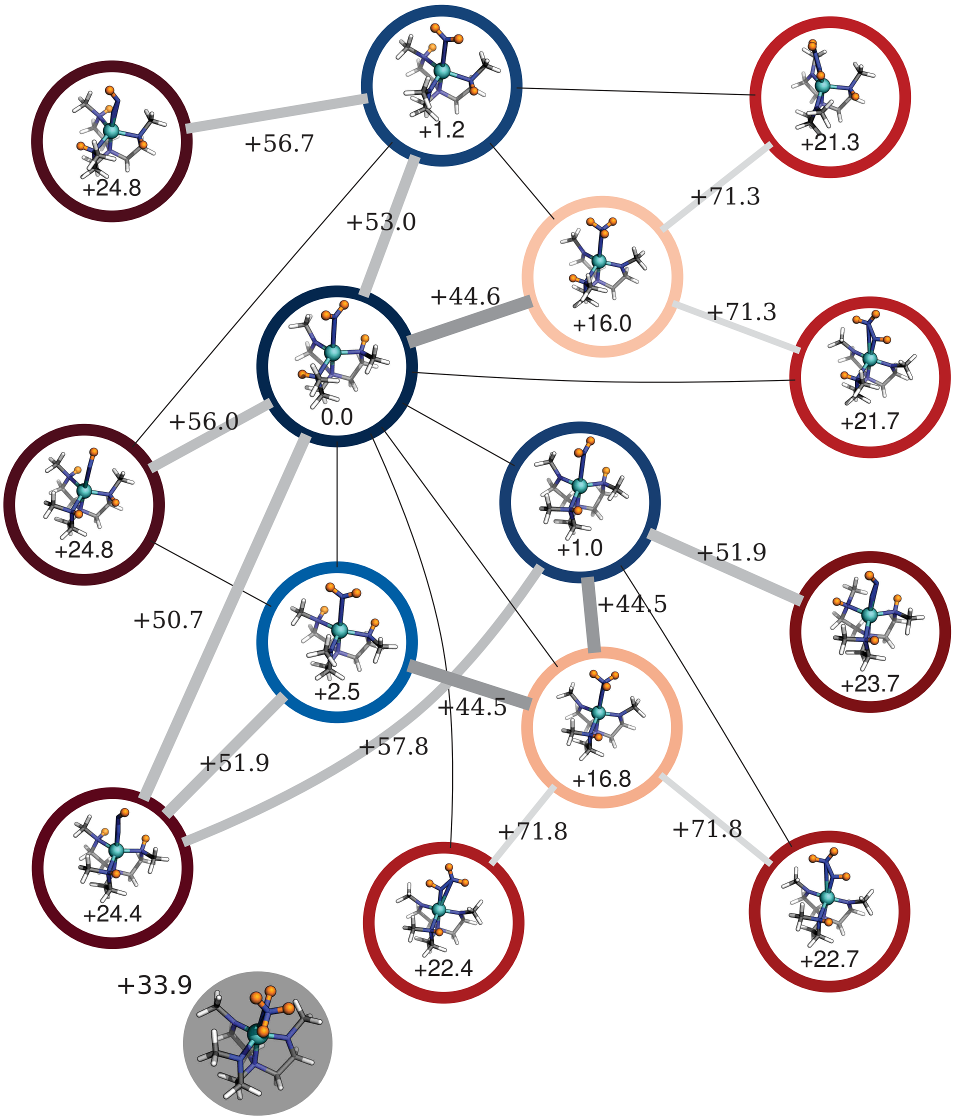

+24.8
+56.7
+1.2
+21.3
+53.0
+71.3
+44.6
+16.0
+71.3
0.0
+21.7
+56.0
+24.8
+1.0
+51.9
+44.5
+50.7
+2.5
+23.7
+44.5
+57.8
+51.9
+16.8
+71.8
+71.8
+24.4
+22.4
+22.7
+33.9

Figure F.21: BP86/RI/def2-SV(P) reaction network for the four-fold protonated and 1 e^--reduced intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of +3 elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

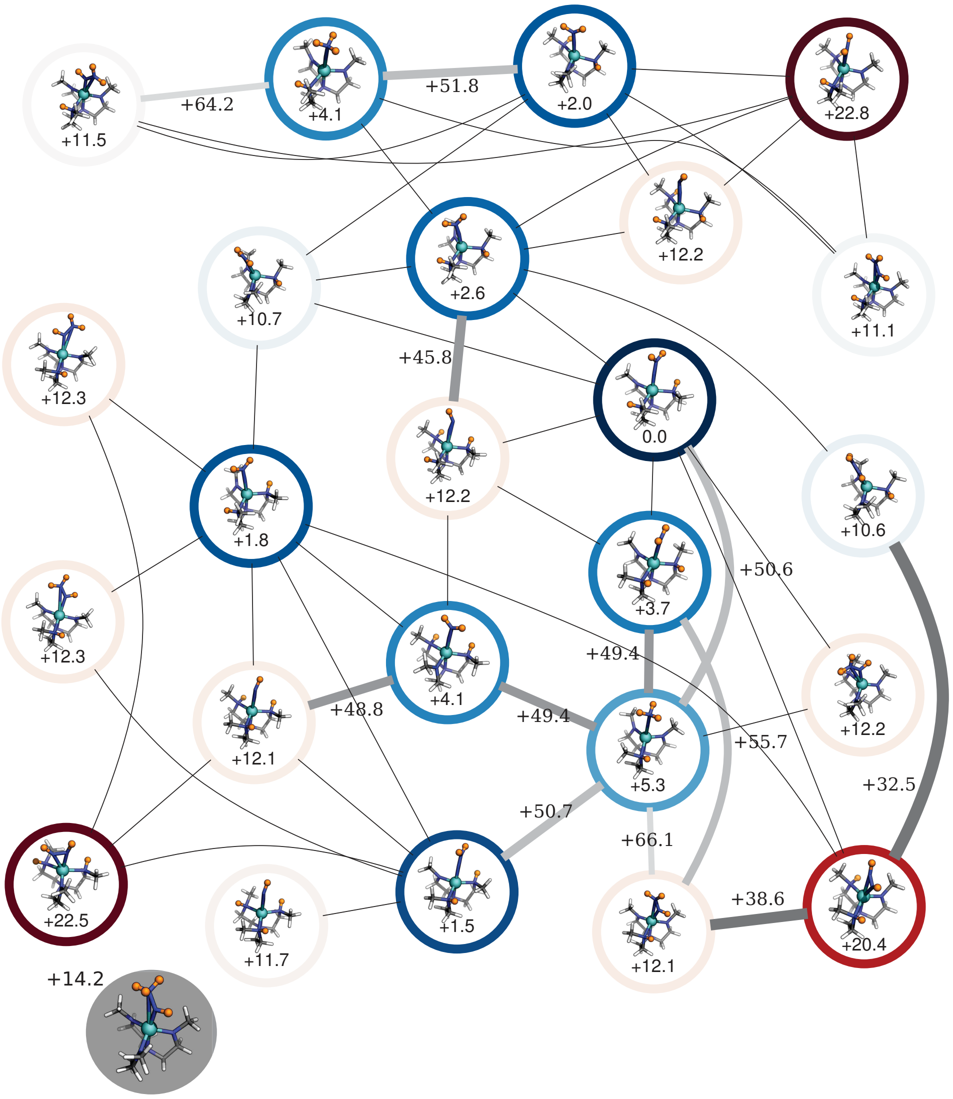

+64.2
+51.8
+11.5
+4.1
+2.0
+22.8
+12.2
+2.6
+10.7
+11.1
+45.8
+12.3
0.0
+12.2
+1.8
+10.6
+50.6
+3.7
+12.3
+49.4
+4.1
+48.8
+49.4
+12.1
+12.2
+55.7
+5.3
+32.5
+50.7
+66.1
+38.6
+22.5
+1.5
+11.7
+12.1
+20.4
+14.2

Figure F.22: BP86/RI/def2-SV(P) reaction network for the four-fold protonated intermediates of the N_2-bound generic model complex of the Schrock catalyst at a total charge of +4 elementary units. Energies are given in kcal/mol. We have chosen an energy cutoff of 25 kcal/mol with respect to the lowest-energy structure for all minimum-energy structures. A dark-blue circle around the minimum energy structure highlights the lowest-energy structure of a given composition and a dark-red circle a high-energy structure (closest to 25 kcal/mol above lowest-energy structure). The intermediate which is chosen as reference for the protonation and electron reduction step is highlighted by a gray background. Element color code: gray, C; blue, N; turquoise, Mo; white, H of substrate; orange, added H's.

Bibliography

[1] H. Eyring and M. Polanyi. In: *Z. Phys. Chem.* B12 (1931), pp. 279–311.

[2] M. Born and J. R. Oppenheimer. In: *Ann. Phys.* 84 (1927), pp. 457–484.

[3] K. Fukui. In: *Acc. Chem. Res.* 14 (1981), pp. 363–368.

[4] H. P. Hratchian and H. B. Schlegel. *Theory and Applications of Computational Chemistry: The First 40 Years*. Amsterdam: Elsevier, 2005.

[5] S. Arrhenius. In: *Z. Phys. Chem.* 4 (1889), pp. 226–248.

[6] H. Eyring. In: *J. Chem. Phys.* 3 (1935), pp. 107–115.

[7] F. Jensen. *Introduction to Computational Chemistry*. Chichester: 2nd ed. Wiley & Sons, 1989.

[8] M. Reiher. In: *Unterlagen zur Quantenchemie-Vorlesung* (2010).

[9] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press, 1981.

[10] P. Hohenberg and W. Kohn. In: *Phys. Rev.* 136 (1964), pp. 864–871.

[11] W. Kohn and L. J. Sham. In: *Phys. Rev.* 140 (1965), A1133–A1138.

[12] L. Noodleman. In: *J. Chem. Phys.* 74 (1981), pp. 5737–5743.

[13] L. Noodleman, D. Post, and E. Baerends. In: *Chem. Phys.* 64 (1982), pp. 159–166.

[14] L. Noodleman, D. A. Case, and A. Aizman. In: *J. Am. Chem. Soc.* 110 (1988), pp. 1001–1005.

[15] L. Noodleman and J. G. Norman. In: *J. Chem. Phys.* 70 (1979), pp. 4903–4906.

[16] C. Herrmann, L. Yu, and M. Reiher. In: *J. Comput. Chem.* 27 (2006), pp. 1223–1239.

[17] M. Reiher. In: *Faraday Disc.* 135 (2007), pp. 97–124.

[18] M. Podewitz and M. Reiher. In: *Adv. Inorg. Chem.* 62 (2010), pp. 177–230.

[19] C. R. Jacob and M. Reiher. In: *Int. J. Quantum Chem.* 112 (2012), pp. 3661–3684.

[20] A. E. Clark and E. R. Davidson. In: *J. Chem. Phys.* 115 (2001), pp. 7382–7392.

[21] C. Herrmann, M. Reiher, and B. A. Hess. In: *J. Chem. Phys.* 122 (2005), p. 034102.

[22] M. Podewitz, C. Herrmann, A. Malassa, M. Westerhausen, and M. Reiher. In: *Chem. Phys. Lett.* 451 (2008), pp. 301–308.

[23] M. Bergeler, M. T. Stiebritz, and M. Reiher. In: *Chem. Plus. Chem.* 78 (2013), pp. 1082–1098.

[24] M. Bergeler, C. Herrmann, and M. Reiher. In: *submitted* (2015).

[25] M. Bergeler, G. Simm, J. Proppe, and M. Reiher. In: *to be submitted* (2015).

[26] M. W. Adams. In: *Biochim. Biophys. Acta* 1020 (1990), pp. 115–145.

[27] M. Frey. In: *ChemBioChem* 3 (2002), pp. 153–160.

[28] K. A. Vincent, A. Parkin, and F. A. Armstrong. In: *Chem. Rev.* 107 (2007), pp. 4366–4413.

[29] J. C. Fontecilla-Camps, A. Volbeda, C. Cavazza, and Y. Nicolet. In: *Chem. Rev.* 107 (2007), pp. 4273–4303.

[30] W. Lubitz, E. Reijerse, and van Gastel. In: *Chem. Rev.* 107 (2007), pp. 4331–4365.

[31] P. M. Vignais and B. Billoud. In: *Chem. Rev.* 107 (2007), pp. 4206–4272.

[32] A. Adamska-Venkatesh, D. Krawietz, J. Siebel, K. Weber, T. Happe, E. Reijerse, and W. Lubitz. In: *J. Am. Chem. Soc.* 136 (2014), pp. 11339–11346.

[33] P. Chernev, C. Lambertz, A. Brünje, N. Leidel, K. G. V. Sigfridsson, R. Kositzki, C.-H. Hsieh, S. Yao, R. Schiwon, M. Driess, C. Limberg, T. Happe, and M. Haumann. In: *Inorg. Chem.* 53 (2014), pp. 12164–12177.

[34] M. T. Stiebritz and M. Reiher. In: *Inorg. Chem.* 48 (2009), pp. 7127–7140.

[35] M. K. Bruska, M. T. Stiebritz, and M. Reiher. In: *J. Am. Chem. Soc.* 133 (2011), pp. 20588–20603.

[36] M. T. Stiebritz, A. R. Finkelmann, and M. Reiher. In: *Eur. J. Inorg. Chem.* (2011), pp. 1163–1171.

[37] M. T. Stiebritz and M. Reiher. In: *Chem. Sci.* 3 (2012), pp. 1739–1751.

[38] M. K. Bruska, M. T. Stiebritz, and M. Reiher. In: *Dalton Trans.* 42 (2013), pp. 8729–8735.

[39] M. Bruschi, C. Greco, P. Fantucci, and L. De Gioia. In: *Inorganic Chemistry* 47 (2008), pp. 6056–6071.

[40] M. Bruschi, C. Greco, L. Bertini, P. Fantucci, U. Ryde, and L. De Gioia. In: *J. Am. Chem. Soc.* 132 (2010), pp. 4992–4993.

[41] A. R. Finkelmann, M. T. Stiebritz, and M. Reiher. In: *Chem. Sci.* 5 (2014), pp. 215–221.

[42] W. Kohn, A. D. Becke, and R. G. Parr. In: *J. Phys. Chem.* 100 (1996), pp. 12974–12980.

[43] Q. Zhao and R. G. Parr. In: *Phys. Rev. A* 46A (1992), pp. 2337–2343.

[44] Q. Zhao and R. G. Parr. In: *J. Chem. Phys.* 98 (1993), pp. 543–548.

[45] E. J. Baerends and O. V. Gritsenko. In: *J. Phys. Chem.* 101 (1997), pp. 5383–5403.

[46] R. Stowasser and R. Hoffmann. In: *J. Am. Chem. Soc.* 121 (1999), pp. 3414–3420.

[47] Z. Cao and M. B. Hall. In: *J. Am. Chem. Soc.* 123 (2001), pp. 3734–3742.

[48] H.-J. Fan and M. B. Hall. In: *J. Am. Chem. Soc.* 123 (2001), pp. 3828–3829.

[49] A. R. Finkelmann, M. T. Stiebritz, and M. Reiher. In: *J. Phys. Chem. B* 117.17 (2013), pp. 4806–4817.

[50] A. R. Finkelmann, M. T. Stiebritz, and M. Reiher. In: *Chem. Comm.* 49 (2013), pp. 8099–8101.

[51] A. R. Finkelmann, H. M. Senn, and M. Reiher. In: *Chem. Sci.* 5 (2014), pp. 4474–4482.

[52] J. A. Imlay. In: *Mol. Microbiol.* 59 (2006), pp. 1073–1082.

[53] I. R. Tsaneva and B. Weiss. In: *J. Bacteriol.* 172 (1990), pp. 4197–4205.

[54] K. A. Vincent, A. Parkin, O. Lenz, S. P. J. Albracht, J. C. Fontecilla-Camps, R. Cammack, B. Friedrich, and F. A. Armstrong. In: *J. Am. Chem. Soc.* 127 (2005), pp. 18179–18189.

[55] H. Beinert. In: *J. Biol. Inorg. Chem.* 5 (2000), pp. 2–15.

[56] M. T. Stiebritz and M. Reiher. In: *Chem. Sci.* 3 (2012), pp. 1739–1751.

[57] B. A. Averill, T. Herskovitz, R. H. Holm, and J. A. Ibers. In: *J. Am. Chem. Soc.* 95 (1973), pp. 3523–3534.

[58] Y. Ohki, K. Tanifuji, N. Yamada, M. Imada, T. Tajima, and K. Tatsumi. In: *Proc. Natl. Acad. Sci.* 108 (2011), pp. 12635–12640.

[59] C. W. Carter Jr., J. Kraut, S. T. Freer, R. A. Alden, L. C. Sieker, E. Adman, and L. H. Jensen. In: *Proc. Nat. Acad. Sci.* 69 (1972), pp. 3526–3529.

[60] R. Cammack. In: *Adv. Inorg. Chem.* 38 (1992), pp. 281–322.

[61] A. Dey, F. E. Jenney Jr., M. W. W. Adams, E. Babini, Y. Takahashi, K. Fukuyama, K. O. Hodgson, B. Hedman, and E. I. Solomon. In: *Science* 318 (2007), pp. 1464–1468.

[62] M. Kodaka, T. Tomohiro, and H. Okuno. In: *J. Phys. Chem.* 95 (1991), pp. 6741–6744.

[63] R. Langen, G. M. Jensen, U. Jacob, P. J. Stephens, and A. Warshel. In: *J. Biol. Chem.* 267 (1992), pp. 25625–25627.

[64] P. D. Swartz, B. W. Beck, and T. Ichiye. In: *Biophys. J.* 71 (1996), pp. 2958–2969.

[65] I. Bertini, G Gori-Savellini, and C. Luchinat. In: *J. Biol. Inorg. Chem.* 2 (1997), pp. 114–118.

[66] B. S. Perrin Jr. and T. Ichiye. In: *Proteins* 78 (2010), pp. 2798–2808.

[67] X. Yang, S. Niu, T. Ichiye, and L.-S. Wang. In: *J. Am. Chem. Soc.* 126 (2004), pp. 15790–15794.

[68] S. Niu and T. Ichiye. In: *J. Am. Chem. Soc.* 131 (2009), pp. 5724–5725.

[69] X.-B. Wang, S. Niu, X. Yang, S. K. Ibrahim, C. J. Pickett, T. Ichiye, and L.-S. Wang. In: *J. Am. Chem. Soc.* 125 (2003), pp. 14072–14081.

[70] K. Chen, G. J. Tilley, V. Sridhar, G. S. Prasad, C. D. Stout, F. A. Armstrong, and B. K. Burgess. In: *J. Biol. Chem.* 274 (1999), pp. 36479–36487.

[71] C. Zhou, J. W. Raebiger, B. M. Segal, and R. H. Holm. In: *Inorg. Chim. Acta* 300-302 (2000), pp. 892–902.

[72] L. Noodleman and D. A. Case. In: *Adv. Inorg. Chem.* 38 (1992), pp. 423–470.

[73] L. Noodleman, C. Y. Peng, D. A. Case, and J.-M. Mouesca. In: *Coord. Chem. Rev.* 144 (1995), pp. 199–244.

[74] T. Glaser, B. Hedman, K. O. Hodgson, and E. I. Solomon. In: *Acc. Chem. Res.* 33 (2000), pp. 859–868.

[75] E. Sigfridsson, M. H. M. Olsson, and U. Ryde. In: *Inorg. Chem.* 40 (2001), pp. 2509–2519.

[76] J. A. Fee, J. M. Castagnetto, D. A. Case, L. Noodleman, C. D. Stout, and R. A. Torres. In: *J. Biol. Inorg. Chem.* 8 (2003), pp. 519–526.

[77] R. A. Torres, T. Lovell, L. Noodleman, and D. A. Case. In: *J. Am. Chem. Soc.* 125 (2003), pp. 1923–1936.

[78] A. Dey, C. L. Roche, M. A. Walters, K. O. Hodgson, B. Hedman, and E. I. Solomon. In: *Inorg. Chem.* 44 (2005), pp. 8349–8354.

[79] E. I. Solomon, B. Hedman, K. O. Hodgson, A. Dey, and R. K. Szilagyi. In: *Coord. Chem. Rev.* 249 (2005), pp. 97–129.

[80] L. Noodleman and W.-G. Han. In: *J. Biol. Inorg. Chem.* 11 (2006), pp. 674–694.

[81] S. Niu and T. Ichiye. In: *Theor. Chem. Acc.* 117 (2007), pp. 275–281.

[82] A. A. Voityuk. In: *Chem. Phys. Lett.* 495 (2010), pp. 131–134.

[83] M. Chakrabarti, E. Münck, and E. L. Bominaar. In: *Inorg. Chem.* 50 (2011), pp. 4322–4326.

[84] D. P. Bhave, W.-G. Han, S. Pazicni, J. E. Penner-Hahn, K. S. Carroll, and L. Noodleman. In: *Inorg. Chem.* 50 (2011), pp. 6610–6625.

[85] S. Niu and T. Ichiye. In: *Mol. Simulat.* 37 (2011), pp. 572–590.

[86] A. D. Walsh. In: *J. Chem. Soc.* 466 (1953), pp. 2260–2331.

[87] S. S. Shaik and A. Pross. In: *J. Am. Chem. Soc.* 104 (1982), pp. 2708–2719.

[88] A. Pross. *Theoretical & Physical Principles of Organic Reactivity*. New York: John Wiley & Sons, Inc., 1995.

[89] Y. Tanabe and S. Sugano. In: *J. Phys. Soc. Jpn.* 9 (1954), pp. 753–779.

[90] Y. Tanabe and S. Sugano. In: *J. Phys. Soc. Jpn.* 11 (1954), pp. 864–877.

[91] M. Podewitz, M. T. Stiebritz, and M. Reiher. In: *Faraday Discuss.* 148 (2011), pp. 119–135.

[92] T. Herskovitz, B. A. Averill, R. H. Holm, J. A. Ibers, W. D. Phillips, and J. F. Weiher. In: *Proc. Natl. Acad. Sci.* 69 (1972), pp. 2437–2441.

[93] P. Giastas, N. Pinotsis, G. Efthymiou, M. Wilmanns, P. Kyritsis, J.-M. Moulis, and I. M. Mavridis. In: *J. Biol. Inorg. Chem.* 11 (2006), pp. 445–458.

[94] M. N. Johannessen, M. S. Nielsen, B. L. Ooi, H. E. M. Christensen, and P Harris. In: *To be published.* ().

[95] C. D. Stout, E. A. Stura, and D. E. McRee. In: *J. Mol. Biol.* 278 (1998), pp. 629–639.

[96] H. S. Gao-Sheridan, M. A. Kemper, R. Khayat, G. J. Tilley, F. A. Armstrong, V. Sridhar, G. S. Prasad, C. D. Stout, and B. K. Burgess. In: *J. Biol. Chem.* 273 (1998), pp. 33692–33701.

[97] E. Saridakis, P. Giastas, G. Efthymiou, V. Thoma, J.-M. Moulis, P. Kyritsis, and I. M. Mavridis. In: *J. Biol. Inorg. Chem.* 14 (2009), pp. 783–799.

[98] I. Rayment, G. Wesenberg, T. E. Meyer, M. A. Cusanovich, and H. M. Holden. In: *J. Mol. Biol.* 228 (1992), pp. 672–686.

[99] Z. Dauter, K. S. Wilson, L. C. Sieker, J. Meyer, and J.-M. Moulis. In: *Biochemistry* 36 (1997), pp. 16065–16073.

[100] E. Parisini, F. Capozzi, P. Lubini, V. Lamzin, C. Luchinat, and G. M. Sheldrick. In: *Acta Crystallogr. D* 55 (1999), pp. 1773–1784.

[101] A. González, S. Benini, and S. Ciurli. In: *Acta Crystallogr. D* 59 (2003), pp. 1582–1588.

[102] M. M. Benning, T. E. Meyer, I. Rayment, and H. M. Holden. In: *Biochemistry* 33 (1994), pp. 2476–2483.

[103] M. Stelter, A. M. P. Melo, G. O. Hreggvidsson, S. Hjorleifsdottir, L. M. Saraiva, M. Teixeira, and M. Archer. In: *J. Biol. Inorg. Chem.* 15 (2010), pp. 303–313.

[104] L. M. Hunsicker-Wang, W. Han, C. D. Stout, L. Noodleman, and J. A. Fee. *To be published.*

[105] J. L. Schlessman, D. Woo, L. Joshua-Tor, J. B. Howard, and D. C. Rees. In: *J. Mol. Biol.* 280 (1998), pp. 669–685.

[106] A. S. Pandey, T. V. Harris, L. J. Giles, J. W. Peters, and R. K. Szilagyi. In: *J. Am. Chem. Soc.* 130 (2008), pp. 4533–4540.

[107] G. B. Seiffert, G. M. Ullmann, A. Messerschmidt, B. Schink, P. M. H. Kroneck, and O. Einsle. In: *Proc. Natl. Acad. Sci.* 104 (2007), pp. 3073–3077.

[108] C. Cavazza, C. Contreras-Martel, L. Pieulle, E. Chabrière, E. C. Hatchikian, and J. C. Fontecilla-Camps. In: *Structure* 14 (2006), pp. 217–224.

[109] J.-H. Jeoung and H. Dobbek. In: *J. Biol. Inorg. Chem.* 17 (2012), pp. 167–173.

[110] *http://metallo.scripps.edu/ClusterGeom (accessed on 10.04.13).*

[111] H. Beinert, R. H. Holm, and E. Münck. In: *Science* 277 (1997), pp. 653–659.

[112] A. Aizman and D. A. Case. In: *J. Am. Chem. Soc.* 104 (1982), pp. 3269–3279.

[113] M. Reiher, D. Sellmann, and B. A. Hess. In: *Theor. Chem. Acc.* 106 (2001), pp. 379–392.

[114] J. Thar and B. Kirchner. In: *J. Phys. Chem. A* 110 (2006), pp. 4229–4237.

[115] L. Noodleman, J. G. Norman Jr., J. H. Osborne, A. Aizman, and D. A. Case. In: *J. Am. Chem. Soc.* 107 (1985), pp. 3418–3426.

[116] J.-M. Mouesca, J. L. Chen, L. Noodleman, D. Bashford, and D. A. Case. In: *J. Am. Chem. Soc.* 116 (1994), pp. 11898–11914.

[117] P. J. Stephens, D. R. Jollie, and A. Warshel. In: *Chem. Rev.* 96 (1996), pp. 2491–2514.

[118] P. D. Swartz and T. Ichiye. In: *Biophys. J.* 73 (1997), pp. 2733–2741.

[119] I. Bertini, S. Ciurli, A. Dikiy, and C. Luchinat. In: *J. Am. Chem. Soc.* 115 (1993), pp. 12020–12028.

[120] E. D. Duée, E. Fanchon, J. Vicat, L. C. Sieker, J. Meyer, and J.-M. Moulis. In: *J. Mol. Biol.* 243 (1994), pp. 683–695.

[121] A. Dey, T. Okamura, N. Ueyama, B. Hedman, K. O. Hodgson, and E. I. Solomon. In: *J. Am. Chem. Soc.* 127 (2005), pp. 12046–12053.

[122] L. Noodleman, T. Lovell, T. Liu, F. Himo, and R. A. Torres. In: *Curr. Opin. Chem. Biol.* 6 (2002), pp. 259–273.

[123] R. Kümmerle, J. Gaillard, P. Kyritsis, and J.-M. Moulis. In: *J. Biol. Inorg. Chem.* 6 (2001), pp. 446–451.

[124] E. Babini, I. Bertini, M. Borsari, F. Capozzi, C. Luchinat, X. Zhang, G. L. C. Moura, I. V. Kurnikov, D. N. Beratan, A. Ponce, A. J. Di Bilio, J. R. Winkler, and H. B. Gray. In: *J. Am. Chem. Soc.* 122 (2000), pp. 4532–4533.

[125] M. J. Kory, M. Bergeler, M. Reiher, and A. D. Schlüter. In: *Chem. Eur. J.* 20 (2014), pp. 6934–6938.

[126] M. J. Kory, M. Wörle, T. Weber, P. Payamyar, J. Dshemuchadse, N. Trapp, and A. D. Schlüter. In: *Nat. Chem.* 6 (2014), pp. 779–784.

[127] S. Grimme. In: *J. Comp. Chem.* 27 (2006), pp. 1787–1799.

[128] S. Grimme, J. Antony, S. Ehrlich, and H. A. Krieg. In: *J. Chem. Phys.* 132 (2010), p. 154104.

[129] F. Weigend and R. Ahlrichs. In: *Phys. Chem. Chem. Phys.* 7 (2005), pp. 3297–3305.

[130] F. Weigend, F. Furche, and R. Ahlrichs. In: *J. Chem. Phys.* 119 (2003), pp. 12753–12762.

[131] P. A. M. Dirac. In: *Proc. Royal. Soc. (London)* 123 (1929), pp. 714–733.

[132] J. C. Slater. In: *Phys. Rev.* 81 (1951), pp. 385–390.

[133] A. D. Becke. In: *Phys. Rev. A* 38 (1988), pp. 3098–3100.

[134] J. P. Perdew. In: *Phys. Rev. B* 33 (1986), pp. 8822–8824.

[135] J. P. Perdew and Y. Wang. In: *Phys. Rev. B* 45 (1992), pp. 13244–13249.

[136] J. M. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. In: *Phys. Rev. Lett.* 91 (2003), p. 146401.

[137] Z. Li, X. Chen, M. Bergeler, M. Reiher, C. Su, and H.-J. Grützmacher. In: *Dalton Transactions* (2015), DOI: 10.1039/C4DT04012K.

[138] R. Sustmann and R. Schubert. In: *Angew. Chem. Int. Ed.* 11 (1972), p. 840.

[139] T. A. Halgren and W. N. Lipscomb. In: *Chem. Phys. Lett.* 49 (1977), pp. 225–232.

[140] C. J. Cerjan and W. H. Miller. In: *J. Chem. Phys.* 75 (1981), pp. 2800–2806.

[141] J. Simons, P. Jorgensen, H. Taylor, and J. Ozment. In: *J. Phys. Chem.* 87 (1983), pp. 2745–2753.

[142] D. J. Wales. In: *J. Chem. Soc., Faraday Trans.* 88 (1992), pp. 653–657.

[143] D. J. Wales. In: *J. Chem. Soc., Faraday Trans.* 89 (1993), pp. 1305–1313.

[144] F. Jensen. In: *J. Chem. Phys.* 102 (1995), pp. 6706–6718.

[145] G. Henkelman, B. P. Uberuaga, and H. Jónsson. In: *J. Chem. Phys.* 113 (2000), pp. 9901–9904.

[146] K. Ohno and S. Maeda. In: *Chem. Phys. Lett.* 384 (2004), pp. 277–282.

[147] C. Peng and H. P. Schlegel. In: *Isr. J. Chem.* 33 (1993), pp. 449–454.

[148] G. Henkelman and H. Jónsson. In: *J. Chem. Phys.* 113 (2000), pp. 9978–9985.

[149] C. G. Broyden. In: *Math. Comput.* 21 (1967), pp. 368–381.

[150] R. Fletcher. *Practical Methods of Optimization. Unconstrained Optimization*. New York: Band 1. Wiley, 1980.

[151] J. M. Bofill. In: *J. Comput. Chem.* 15 (1994), pp. 1–11.

[152] M. J. D. Powell. In: *Math. Prog.* (1971), pp. 26–57.

[153] S. M. Sharada, P. M. Zimmermann, A. T. Bell, and M. Head-Gordon. In: *J. Chem. Theory Comput.* 8 (2012), pp. 5166–5174.

[154] P. M. Zimmerman. In: *J. Chem. Phys.* 138 (2013), p. 184102.

[155] P. M. Zimmerman. In: *J. Comput. Chem.* 34 (2013), pp. 1385–1392.

[156] M. Reiher and J. Neugebauer. In: *J. Chem. Phys.* 118 (2003), pp. 1634–1641.

[157] J. Neugebauer and M. Reiher. In: *J. Phys. Chem. A* 108 (2004), pp. 2053–2061.

[158] J. Neugebauer and M. Reiher. In: *J. Comput. Chem.* 25 (2004), pp. 587–597.

[159] M. Reiher and J. Neugebauer. In: *Phys. Chem. Chem. Phys.* 6 (2004), pp. 4621–4629.

[160] M. Reiher, B. Le Guennic, and B. Kirchner. In: *Inorg. Chem.* 44 (2005), pp. 9640–9642.

[161] T. B. Adler, N. Bohro, M. Reiher, and M. A. Suhm. In: *Angew. Chem.* 118 (2006), pp. 3518–3523.

[162] C. Herrmann and M. Reiher. In: *Surf. Science* 600 (2006), pp. 1891–1900.

[163] C. Herrmann, J. Neugebauer, and M. Reiher. In: *New. J. Chem.* 31 (2007), pp. 818–831.

[164] C. Herrmann, J. Neugebauer, and M. Reiher. In: *J. Comput. Chem.* 29 (2008), pp. 2460–2470.

[165] P. Deglmann and F. Furche. In: *J. Chem. Phys.* 117 (2002), pp. 9535–9538.

[166] S. M. Sharada, A. T. Bell, and M. Head-Gordon. In: *J. Chem. Phys.* 140 (2014), p. 164115.

[167] S. M. Sharada, A. T. Bell, and M. Head-Gordon. In: *J. Chem. Phys.* 140 (2014), p. 229902.

[168] H. L. Davis, D. J. Wales, and R. S. Berry. In: *J. Chem. Phys.* 92 (1990), pp. 4308–4319.

[169] K. Bondensgård and F. Jensen. In: *J. Chem. Phys.* 104 (1996), pp. 8025–8031.

[170] W. Quapp, M. Hirsch, and D. Heidrich. In: *Theor. Chem. Acc.* 100 (1998), pp. 285–299.

[171] S. Maeda, K. Ohno, and K. Morokuma. In: *Phys. Chem. Chem. Phys.* 15 (2013), pp. 3683–3701.

[172] D. V. Yandulov and R. R. Schrock. In: *J. Am. Chem. Soc.* 124 (2002), pp. 6252–6253.

[173] D. V. Yandulov and R. R. Schrock. In: *Science* 301 (2003), pp. 76–78.

[174] R. R. Schrock. In: *Acc. Chem. Res.* 38 (2005), pp. 955–962.

[175] J. Baker and F. Chan. In: *J. Comput. Chem.* 17 (1996), pp. 888–904.

[176] Y. Zhao, N González-García, and D. G. Truhlar. In: *J. Phys. Chem. A* 109 (2005), pp. 2012–2018.

[177] V. Guner, K. S. Khuong, A. G. Leach, P. S. Lee, M. D. Bartberger, and K. N. Houk. In: *J. Phys. Chem. A* 107 (2003), pp. 11445–11459.

[178] P. M. Zimmerman. In: *J. Chem. Theory Comput.* 9 (2013), pp. 3043–3050.

[179] D. J. Wales. In: *J. Chem. Phys.* 101 (1994), pp. 3750–3762.

[180] D. O. Neal, H. Taylor, and J. Simons. In: *J. Phys. Chem.* 88 (1984), pp. 1510–1513.

[181] A. Banerjee, N. Adams, J. Simons, and R. Shepard. In: *J. Phys. Chem.* 89 (1985), pp. 52–57.

[182] D. J. Wales. In: *Mol. Phys.* 74 (1991), pp. 1–25.

[183] *www.reiher.ethz.ch/software/akira (accessed on 30.11.2014).*

[184] T. Weymuth, M. P. Haag, K. Kiewisch, S. Luber, S. Schenk, C. R. Jacob, C. Herrmann, J. Neugebauer, and M. Reiher. In: *J. Comput. Chem.* 33 (2012), pp. 2186–2198.

[185] K. H. Marti and M. Reiher. In: *J. Comput. Chem.* 30 (2009), pp. 2010–2020.

[186] M. P. Haag and M. Reiher. In: *Faraday Discuss.* 169 (2014), pp. 89–118.

[187] M. P. Haag, A. C. Vaucher, M. Bosson, S. Redon, and M. Reiher. In: *ChemPhysChem* 15 (2014), pp. 3301–3319.

[188] *DeLano,W. L. DeLano Scientific: San Carlos, CA*. 2002.

[189] *http://www.jmol.org/ (accessed on 30.11.14).*

[190] R. Ahlrichs, M Bär, M Häser, H. Horn, and C Kölmel. In: *Chem. Phys. Lett.* 162 (1989), pp. 165–169.

[191] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. *Gaussian 09*. Wallingford CT: Gaussian, Inc., 2009.

[192] A. Schäfer, H. Horn, and R. Ahlrichs. In: *J. Chem. Phys.* 97 (1992), pp. 2571–2577.

[193] A. Schäfer, C. Huber, and R. Ahlrichs. In: *J. Chem. Phys.* 100 (1994), pp. 5829–5835.

[194] B. LeGuennic, B. Kirchner, and M. Reiher. In: *Chem. Eur. J.* 11 (2005), pp. 7448–7460.

[195] F. Studt and F. Tuczek. In: *Angew. Chem. Int. Ed.* 44 (2005), pp. 5639–5642.

[196] S. Schenk, B. L. Guennic, B. Kirchner, and M. Reiher. In: *Inorg. Chem.* 47 (2008), pp. 3634–3650.

[197] S. Schenk and M. Reiher. In: *Inorg. Chem.* 48 (2009), pp. 1638–1648.

[198] S. Schenk, B. Kirchner, and M. Reiher. In: *Chem. Eur. J.* 15 (2009), pp. 5073–5082.

[199] D. V. Khoroshun, D. G. Musaev, and K. Morokuma. In: *Mol. Phys.* 100 (2002), pp. 523–532.

[200] Z. Cao, Z. Zhou, H. Wan, and Q. Zhang. In: *Int. J. Quantum Chem.* 103 (2005), pp. 344–353.

[201] F. Studt and F. Tuczek. In: *Angew. Chem.* 117 (2005), pp. 5783–5787.

[202] A. Magistrato, A. Robertazzi, and P. Carloni. In: *J. Chem. Theory Comput.* 3 (2007), pp. 1708–1720.

[203] L. Deng, T. Ziegler, and L. Fan. In: *J. Chem. Phys.* 99 (1993), pp. 3823–3835.

[204] S. Maeda, E. Abe, M. Hatanaka, T. Taketsugu, and K. Morokuma. In: *J. Chem. Theory Comput.* 8 (2012), pp. 5058–5063.

[205] S. Maeda, T. Taketsugu, and K. Morokuma. In: *Z. Phys. Chem.* 227 (2013), pp. 1421–1433.

[206] S. Maeda, K. Ohno, and K. Morokuma. In: *J. Chem. Theory Comput.* 5 (2009), pp. 2734–2743.

[207] D. Rappoport, C. J. Galvin, D. Y. Zubarev, and A. Aspuru-Guzik. In: *J. Chem. Theory Comput.* 10 (2013), pp. 897–907.

[208] S. Goedecker. In: *J. Chem. Phys.* 120 (2004), pp. 9911–9917.

[209] S. Schweiger, B. Hartke, and G. Rauhut. In: *Phys. Chem. Chem. Phys.* 6 (2004), pp. 3341–3349.

[210] S. Roy, S. Goedecker, M. J. Field, and E. Penev. In: *J. Phys. Chem. B* 113 (2009), pp. 7315–7321.

[211] J. M. Dietrich and B. Hartke. In: *J. Comput. Chem.* 32 (2011), pp. 1377–1385.

[212] G. van Rossum. *http://www.python.org/ (accessed on 27.11.14).*

[213] E. G. Lewars. *Computational Chemistry — Introduction to the Theory and Applications of Molecular and Quantum Mechanics.* 2nd ed. Springer, 2011.

[214] C. Gonzalez and H. B. Schlegel. In: *J. Phys. Chem.* 94 (1990), pp. 5523–5527.

[215] A. Behn, P. M. Zimmerman, A. T. Bell, and M. Head-Gordon. In: *J. Chem. Phys.* 135 (2011), p. 224108.

[216] Y. Shao, L. Fusti-Molnar, Y. Jung, J Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio Jr., R. C. Lochan, T. Wang, G. J. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. V. Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Daschel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khaliullin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. Gill, and M. Head-Gordon. In: *Phys. Chem. Chem. Phys.* 8 (2006), pp. 3172–3191.

[217] C. E. Klots. In: *J. Phys. Chem.* 85 (1981), pp. 3585–3588.

[218] I. A. Topol, G. J. Tawa, S. K. Burt, and A. A. Rashin. In: *J. Phys. Chem.* 101 (1997), pp. 10075–10081.

[219] R. R. Schrock. In: *Angew. Chem. Int. Ed.* 47 (2008), pp. 5512–5522.

[220] J. Neugebauer, M. Reiher, C. Kind, and B. A. Hess. In: *J. Comput. Chem.* 23 (2002), pp. 895–910.

[221] A. D. Becke. In: *J. Chem. Phys.* 98 (1993), pp. 5648–5652.

[222] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. In: *J. Phys. Chem.* 98 (1994), pp. 11623–11627.

[223] C. Lee, W. Yang, and R. G. Parr. In: *Phys. Rev. B* 37 (1988), pp. 785–789.

[224] C. Greco, P. Fantucci, U. Ryde, and L. De Gioia. In: *Int. J. Quantum Chem.* 111 (2011), pp. 3949–3960.

[225] I. Mayer. In: *Chem. Phys. Lett.* 440 (2007), pp. 357–359.

[226] P.-O. Löwdin. In: *J. Chem. Phys.* 18 (1950), pp. 365–375.

[227] M. Bruschi, C. Greco, M. Kaukonen, P. Fantucci, U. Ryde, and L. De Gioia. In: *Angew. Chem. Int. Ed.* 48 (2009), pp. 3503–3506.

[228] A. Klamt and V. Jonas. In: *J. Chem. Phys.* 105 (1996), pp. 9972–9981.

[229] *http://www.wolfram.com/ (accessed on 28.08.13).*

[230] „ChemBioDraw Ultra 12.0, Cambridge Soft: Cambridge, MA.“ In: (2010).

[231] A. Klamt and G. Schüürmann. In: *J. Chem. Soc. Perk. T. 2* (1993), pp. 799–805.

[232] A. P. Scott and L. Radom. In: *J. Phys. Chem.* 100 (1996), pp. 16502–16513.

[233] *http://www.cp2k.org/ (accessed on 4.12.13).*

[234] L. Martínez, R. Andrade, E. G. Brigin, and J. M. Martínez. In: *J. Comput. Chem.* 30 (2009), pp. 2157–2164.

[235] S. Goedecker, M. Teter, and J. Hutter. In: *Phys. Rev. B* 54 (1996), pp. 1703–1710.

[236] C. Hartwigsen, S. Goedecker, and J. Hutter. In: *Phys. Rev. B* 58 (1998), pp. 3641–3662.

[237] J. P. Perdew, K. Burke, and M. Ernzerhof. In: *Phys. Rev. Lett.* 77 (1996), pp. 3865–3868.

[238] J. P. Perdew, K. Burke, and M. Ernzerhof. In: *Phys. Rev. Lett.* 78 (1997), p. 1396.

[239] C. Adamo and V. Barone. In: *J Chem Phys* 110 (1999), pp. 6158–6169.

Abbreviations

Part I: The following abbreviations are used in this thesis.

abbreviation	meaning
ADE	Adiabatic detachment energy
BOMD	Born–Oppenheimer molecular dynamics
BS	Broken–symmetry
Cys	Cysteine
DMB	Dimethyl butadiene
DMSO	Dimethyl sulfoxide
DFT	Density functional theory
EVF	Eigenvector following
Fdn	Ferredoxin
FSM	Freezing-string method
GRRM	Global reaction route mapping
HF	Hartree–Fock
HiPIP	High-potential iron protein
HIPT	Hexa-*iso*-propyl terphenyl
HOMO	Highest-occupied molecular orbital
IR	Infrared
IRC	Intrinsic reaction coordinates
KS	Kohn–Sham
LST	Linear synchroneous transit
LUMO	Lowest-unoccupied molecular orbital
MP2	Møller-Plesset second-order perturbation
NEB	Nudged elastic band
PES	Potential energy surface
	(continued on the next page)...

Part II: The following abbreviations are used in this thesis.

abbreviation	meaning
	... *(continued from the previous page)*
QST	Quadratic synchroneous transit
PDB	Protein data bank
RI	Resolution-of-the-identity
RMSD	Root-mean square deviation
SCF	Self-consistent-field
TS	Transition-state structure
WT	Wild type
ZPE	Zero-point energy

Publications

The following publications are included in parts or an extended version in this thesis:

- M. Bergeler, M. T. Stiebritz, M. Reiher, *Structure-property relationships of Fe_4S_4 clusters*, *ChemPlusChem*, **2013**, *78*, 1082-1098. (Section 2.2 and Chapter 4)
- M. Bergeler, C. Herrmann, M. Reiher, *Mode-Tracking Based Stationary-Point Optimization*, *J. Comput. Chem.*, **2015**, submitted. (Chapters 1, 5 and 7)
- M. Bergeler, G. Simm, J. Proppe, M. Reiher, *Heuristic Reaction Network Explorations*, **2015**, to be submitted. (Chapters 1, 6 and 7)
- M. Kory, M. Bergeler, M. Reiher, A. D. Schlüter *Facile synthesis and theoretical conformation analysis of a triazine-based, double decker rotor molecule with three anthracene blades*, *Chem. Eur. J.*, **2014**, *20*, 6934-6938. (Section 3.1 and Chapter 4)
- Z. Li, X. Chen, M. Bergeler, M. Reiher, C.-Y. Su, H.-J. Grützmacher *A Stable Organic Phosphaketene and Its Reactivity*, *Dalton Transactions*, **2015**, DOI: 10.1039/C4DT04012K. (Section 3.2)

Additional publication:

- J. R. Allison, M. Bergeler, N. Hansen, W. F. van Gunsteren, *Current Computer Modeling Cannot Explain Why Two Highly Similar Sequences Fold into Different Structures*, *Biochemistry*, **2011**, *50*, 10965-10973.

Acknowledgement

First of all, I would like to thank Prof. Dr. Markus Reiher for giving me the chance to work in his research group and for being an excellent supervisor. I very much enjoyed the discussions about our research projects as well as discussions about research in general. I will always be fascinated about his passion about chemistry in general and theoretical chemistry in specific. And I am grateful for all I have learned from him.

I would like to thank Prof. Dr. Sereina Riniker and Prof. Dr. Carmen Herrmann for accepting to be co-examiners of this thesis. Prof. Dr. Carmen Herrmann I also want to thank for her support on the implementation of the MTsearch algorithm and for giving me the chance to visit her at the University of Hamburg.

Furthermore, I would like to thank all members of the group, which I have met during the last three years (in alphabetical order): Dr. Katharina Boguslawski, Marta Bruska, Dr. Steven Donald, Arndt Finkelmann, Dr. Samuel Fux, Moritz Haag, Erik Hedegård, Sebastian Keller, Dr. Stefan Knecht, Dr. Arseny Kovyrshin, Florian Krausbeck, PD Dr. Hans Peter Lüthi, Dr. Yingjin Ma, Halua Pinto de Magalhaes, Dr. Edit Matyus, Andrea Muolo, Dr. Daoling Peng, Jonny Proppe, Oliver Sala, Benjamin Simmen, Christopher Stein, Dr. Martin Stiebritz, Dr. Pawel Tecmer, Dr. Elizabeth Chirackal Varkey, and Alain Vaucher.

Especially I want to thank Jonny, Stefan, Oliver and Halua for the nice time, the lunch breaks at WOKA and all discussions we had.

Many thanks also go to our secretary, Romy Isenegger, for being always very supportive and contributing significantly to the friendly and productive atmosphere within the group.

I would like to thank our collaboration partners, Danuta Krawietz, Martin Winkler and Prof. Happe from Ruhr University Bochum, Max Kory and Prof. Schlüter from the D-MATL at ETH Zurich as well as Zhongshu Li and Prof. Grützmacher from the D-CHAB at ETH Zurich. It was a pleasure for me to work together with you on the projects.

Moreover, I had the chance to supervise a very motivated master student, Gregor Simm, during his semester project. I enjoyed the time and wish you all the best for your future.

Besides work, there is one person I owe infinitely many thanks to, which is Stephan Bachmann. He has always been there for me and supported me wherever I needed it. Thank you for the great time, for being always available and giving the best advices in every sense.

Furthermore, I am very grateful to my friends in Zurich and from several other places in the world. I want to thank Noah Bieler, Alice Lonardi, Pavel Oborsky, Lenka Cienikova, Pascal Merz, Jagna Witek, Annick Rennevey, Francesco Ravotti, Andrea Barna, Anja Billeter, Alena Kuhn, Lisa Eisert, Ignaz Höhlein and Julia Weiss. Many thanks also go to the Female Scientists in Structural Biology group, to which Alice has introduced me. Thanks for the very nice atmosphere and all the personal and academic discussions we had. I enjoyed the time a lot.

During my studies and PhD I have been part of a synchronized skating team in Zurich. I appreciated the great atmosphere a lot and I would like to thank all of them. For me it was always the best imaginable compensation besides work I could have. I will always remember the 7 years in our team and hope to come back at one point.

At the end I want to thank the persons I am most thankful at all, which are my parents and my twin-sister. They always support me in every aspect of my live.

Maike Bergeler

Date of birth	October 31, 1989
Place of birth	Mainz, Germany
Nationality	German

Education

10/2011–present	PhD in Theoretical Chemistry, ETH Zurich, Switzerland Thesis: Towards Automated Quantum Chemical Explorations of Chemical Mechanisms Supervisor: Prof. Dr. Markus Reiher
09/2010–09/2011	MSc Chemistry, ETH Zurich, Switzerland Thesis: Free energy calculations using DFTMD: A study of H_2O_2 and iron complexes in aqueous solution Supervisors: Prof. M. Sprik (University of Cambridge) and Prof. Dr. Markus Reiher
09/2007–09/2010	BSc Chemistry, ETH Zurich, Switzerland
08/2004–06/2007	Internatsschule Schloss Hansenberg, Geisenheim, Germany High school diploma (Abitur)

Scholarships

2012–2014	PhD Scholarship of the Fond der chemischen Industrie
2011	Erasmus Exchange Scholarship