

# **Theoretical Studies of Ru- and Re-based Catalysts for Artificial Photosynthesis**

Dissertation  
for the award of the degree  
“Doctor rerum naturalium” (Dr.rer.nat)  
of the Georg-August-Universität Göttingen

within the doctoral program of Chemistry  
of the Georg-August University School of Science (GAUSS)

submitted by  
**Thorsten Stolper**  
from Sulz am Neckar

Göttingen, 2017



## Thesis Committee

**Prof. Dr. Ricardo Mata**

Institute of Physical Chemistry  
Georg-August University Göttingen

**Prof. Dr. Franc Meyer**

Institute of Inorganic Chemistry  
Georg-August University Göttingen

**Prof. Dr. Peter Blöchl**

Institute of Theoretical Physics  
Clausthal University of Technology

## Members of the Examination Board

Reviewer:

**Prof. Dr. Ricardo Mata**

Institute of Physical Chemistry  
Georg-August University Göttingen

Second Reviewer:

**Prof. Dr. Peter Blöchl**

Institute of Theoretical Physics  
Clausthal University of Technology

## Further members of the Examination Board

**Prof. Dr. Inke Siewert**

Institute of Inorganic Chemistry  
Georg-August University Göttingen

**Prof. Dr. Jörg Behler**

Institute of Physical Chemistry  
Georg-August University Göttingen

**Prof. Dr. Martin Suhm**

Institute of Physical Chemistry  
Georg-August University Göttingen

Date of the oral examination: 8.12.2017



Newton, quasi-Newton, and GDIIS methods  
will converge [...] if the surface is quadratic.  
However, life is not quadratic.

---

Bernhard-”Berny”-Schlegel



# Acknowledgements

First and foremost I would like to thank my supervisor, Prof. Dr. Ricardo Mata, for the possibility to work on this very current and interesting topic. He guided me through several years of research, always had time for discussions and enabled me to further my personal development with skill building seminars and visits to a number of conferences. Many thanks also go to the other members of the Computational Chemistry and Biochemistry group and especially to those who offered to read and correct parts of this thesis.

I would like to thank the other members of my thesis committee, Prof. Dr. Franc Meyer and Prof. Dr. Peter Blöchl, for all the interesting discussions and comments during the course of my PhD and of course for accepting the extra work that comes with a position in the thesis committee. On the same note I want to thank all mentioned and the additional members of my examination board, Prof. Dr. Inke Siewert, Prof. Dr. Jörg Behler and Prof. Dr. Martin Suhm, for agreeing to take part in the examination and the uncomplicated way of choosing a date for the examination.

Of course many additional scientists contributed to invigorating conversations that helped to broaden my horizon. In particular many thanks go to Marian Bongers and Mohsen Sotoudeh, who were always as interested as I in understanding the work of the other and with whom I had many exciting discussions.

Last but not least I want to thank all my relatives and friends, who always helped to motivate me and understood the time constraints that writing a thesis includes. It goes without saying that Christine deserves a large part of my gratitude, as she always had to put up with my mood swings and late-night working sessions. She supported me through all this time and I cannot imagine having done this without her.





# Abstract

The conversion of light into energy stored in a sustainable and economic way is one of the main challenges of the 21st century. Artificial photosynthesis is therefore an active research field in many parts of science. Molecular catalysts have come a long way in the reduction of carbon dioxide and the oxidation of water to dioxygen. Although a lot has been learned about the reactivity and stability of these catalysts, many unanswered questions remain. This thesis hopefully takes a further step towards answering some of these questions.

Since computational studies rely heavily on the investigation of relevant points on the potential hypersurface, the optimisation procedure can become a major bottleneck. Several approximations have been suggested for the step prediction, second derivatives and coordinates. The latter have been shown to exhibit a great influence on the efficiency of the minimisation, with approximate decoupling for carefully designed internal coordinates. Nevertheless, the implementation and thus performance varies widely for different software suites. A new geometry optimisation software is therefore developed that allows a detailed investigation of each employed algorithm. The focus is set on the impact of the primitive coordinates, as used in the redundant internals scheme, and tested on the convergence of several drug-like organic molecules. Some corner-cases of molecular structures and a CO<sub>2</sub> reduction catalyst are optimised with several approaches to corroborate the superior performance of the developed procedure.

Investigations into CO<sub>2</sub> reduction catalysts have been conducted for some time, with a large number of results reported for rhenium-based complexes. A recently published dinuclear rhenium catalyst with a proton responsive ligand combines the knowledge gained in those studies and reports a detailed analysis of the occurring reduction intermediates by infrared spectro-electrochemistry. To achieve a greater insight on the molecular structure of the catalyst in solution, a computational basis for the calculation of infrared frequency shifts is devised. The calibration of theoretical values to the crystal structure enables further comparison of calculated to the experimental spectra. An extensive search for possible side-reaction products during reduction is then undertaken to assign each of the experimentally encountered intermediates, which underlines the usefulness of the approach in catalysis investigations.

A novel dinuclear ruthenium catalyst with an activity in the oxidation of water to dioxygen is inspected in the fifth chapter. It has been proven experimentally to react by a water nucleophilic attack mechanism, contrary to other dinuclear complexes. The computational examination of this catalyst and its comparison to other molecular systems with a similar active site may thus reveal the subtle structural details that determine the mechanism. In a first step, the full range of possible reaction pathways is scanned, including all probable multiplicities in two oxidation states. Fuelled by the failure of some density functionals to reproduce the correct mechanistic preference, a range of different methods and the influence of exact exchange is tested. With the determined computational protocol the catalyst is compared to two similar water oxidation complexes on the same level of theory, which allows to unearth the difference in their structure and reactivity. The gained insight suggests further modifications of the ligand structure that could steer the mechanism into one direction or the other.



# Contents

<b>Acknowledgements</b>	<b>I</b>
<b>Abstract</b>	<b>III</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theory</b>	<b>5</b>
2.1 Hartree-Fock Theory . . . . .	6
2.2 Density Functional Theory . . . . .	8
2.3 Implicit Solvation Models . . . . .	12
2.4 Geometry Optimisation . . . . .	14
<b>3 Development of an Improved Optimisation Procedure</b>	<b>23</b>
3.1 The StOpt-Software . . . . .	23
3.2 The Best Combination of Approximations . . . . .	26
3.2.1 Step Prediction . . . . .	28
3.2.2 The Approximation of the Hessian . . . . .	30
3.2.3 The Coordinate Set . . . . .	32
3.3 Comparison to Other Software . . . . .	35
3.4 A Real-World Test . . . . .	40
3.5 Conclusion . . . . .	41
<b>4 Molecular Reduction of CO<sub>2</sub></b>	<b>45</b>
4.1 Computational Determination of Reduction Intermediates . . . . .	47
4.1.1 Computational Details . . . . .	48
4.1.2 Results . . . . .	49
4.2 Conclusion . . . . .	55
<b>5 Artificial Water Oxidation</b>	<b>57</b>
5.1 Computational Details . . . . .	60
5.2 Computational Investigations on a novel dinuclear Ruthenium Catalyst . . . . .	62
5.2.1 Experimental Basis <sup>[13]</sup> . . . . .	62
5.2.2 A Reasonable Model System for [Ru <sub>2</sub> (μ-Mebbp)(py-SO <sub>3</sub> ) <sub>4</sub> (H <sub>2</sub> O) <sub>2</sub> ] . . . . .	63
5.2.3 Assessing the Validity of the Model Catalyst . . . . .	63
5.2.4 Comparison of Mechanisms . . . . .	65
5.3 Comparison of Water Oxidation Catalysts . . . . .	69
5.3.1 Ru-cbim: Similar to Ru-Hbpb or to Ru-Mebbp? . . . . .	73
5.3.2 Indices to Predict Water Oxidation Mechanisms . . . . .	74

5.4 Conclusion . . . . .	79
<b>6 Summary</b>	<b>81</b>
<b>Bibliography</b>	<b>85</b>
<b>Appendix</b>	<b>95</b>
Additional Benchmarks of StOpt's Methods . . . . .	95
Functional Tests on Reduction Intermediates of <b>1</b> . . . . .	95
Cartesian Structures of Intermediates and Transition States . . . . .	96
<b>Curriculum Vitae</b>	<b>105</b>

# 1 Introduction

Global warming, greenhouse gases and their effects on the environment are dominating topics in research, media and politics. It has been proposed that a rise in global mean temperature change relative to pre-industrial levels should be kept below 2 °C,<sup>[1]</sup> in order to avoid irreversible changes to the planet. As most of the emissions created by humans result from the always growing industrialisation, one might wonder whether current standards of living can be kept or even improved, while striving for the goal of an environmentally friendly industry. Sustainable energy production and storage goes a long way and is therefore one of the main areas of focus in reaching that goal. Many great inventions of humanity were made by taking nature as an example and thus it comes to no surprise that plants seem to have perfected the whole cycle of converting sunlight into chemical energy in form of carbohydrates and oxygen. Consequently there is an active field of research on artificial photosynthesis to develop procedures for an improved light-harvesting, energy transfer, CO<sub>2</sub> reduction and water oxidation.<sup>[2]</sup> Ultimately it is hoped that this activity uncovers part of the intricate path towards an ecologic economy by devising routes to catalytically active systems with a performance that is superior to nature's.

The successful development of efficient and stable catalysts requires a strategy to control the effect of different catalyst structures, solvent interactions and side-reactions. As this is often a difficult task in reactions on surfaces, homogeneous molecular catalysts are employed as model systems which can ideally be tuned by several modifications to the ligand structure. Even then the most interesting states and transformations occur during the catalysis, rendering the experimental analysis a rather difficult endeavor. Theoretical methods are therefore often applied to complement the experimental assignments and in some cases stimulate further investigation. Usual multi-nuclear transition-metal-based catalysts of medium to large size present difficult challenges to the computational chemist. Transition metals often take on several different oxidation and spin states, which can also be coupled either directly or through the ligand backbone. Essentially this means an extensive correlated multireference approach would have to be used that generally leads to a steep scaling with growing system size. The ligands may be non-innocent and lead to charge transfer processes or exhibit a large contribution of dispersion interactions, which requires a well-balanced correlation treatment. Taking the solvent into account opens up a range of additional possibilities and approximations that generally extend the computation time significantly. To achieve a balance between the cost of electronic structure method and the possible size of the system under investigation, *density functional theory* (DFT) is usually the method of choice. These have been shown to be less sensitive to problems stemming from multi-reference situations than wavefunction methods and, depending on the actual functional involved, yield an average accuracy that is comparable to more demanding approaches.

Although routine calculations can be carried out with DFT functionals, any method becomes expensive if the optimisation, be it in the self-consistent field or geometry relaxation procedure, does not converge. If memory permits and analytic gradients are available, Quasi-Newton algorithms have been found to accelerate the convergence to a minimum most efficiently and thus has been implemented in

several electronic structure codes as a means to quickly traverse the *potential energy surface* (PES).<sup>[3]</sup> Due to the subjectively varying importance of geometry optimisation algorithms, “quickly” may differ between several codes. It has been found that a *redundant internal coordinate* set is more robust than the efficient but tedious to define natural internals, while still accelerating the convergence of the relaxation. Influences of other approximations such as model Hessians, update algorithms and trust radius update have also been investigated already.<sup>[3,4]</sup> However, the accurate determination of the step and the influence of the primitive coordinates on the overall performance has gained less attention. A software to optimise molecular geometries has therefore been developed in the course of this work and applied to molecules with a range of structural motifs. Several influential approximations are tested, with a particular focus on the redundant internal coordinate set and the step restriction.

CO<sub>2</sub> reduction catalysts promise a route to limit the amount of greenhouse gases that reside in the atmosphere and can additionally serve as a source for chemicals. Due to the high potentials associated with the reduction of CO<sub>2</sub>, the main product released from these catalysts has been found to be CO.<sup>[5]</sup> A lot of progress has been made in unraveling the mechanistic details of the catalysis, with a large percentage based on metal complexes containing rhenium. The work of several groups lead to the suggestion of an improved catalysis by protic groups near the active site and multi-nuclear assemblies.<sup>[6–8]</sup> These effects were consequently incorporated into a novel rhenium catalyst that was extensively characterised by cyclovoltammetry (CV) and infrared-spectroelectrochemistry (IR-SEC), which allowed to assign several intermediates occurring during multi-electron reductions. With the successful application of relatively simple DFT functionals, the experimental deductions are tested against theoretical harmonic frequency calculations. Since the crystal structure has been determined, experimentally well-characterised structures are used to reference the computational protocol and acquire correction parameters. Intermediates with less experimental insight can then be subject to a thorough investigation of the structure, giving a comprehensive picture of the reduction process.

Water oxidation forms the complementary part to the reduction of CO<sub>2</sub> for sustainable energy storage. The *oxygen evolving complex* (OEC) in Photosystem II of plant leaves is known to catalyse this reaction via a tetramanganese cluster CaMn<sub>4</sub>O<sub>5</sub> efficiently by successive charge accumulation and proton-coupled electron transfer processes. Due to the multitude of intermediate states as summarised by the Kok cycle, the overpotential has been found to stay low.<sup>[9]</sup> Many approaches to mimic the effects of the OEC in small molecular catalysts have therefore been developed, with successful applications from bioinspired molecular clusters to mono- and dinuclear catalysts.<sup>[10–12]</sup> In order to advance beyond the simple replication of systems in nature a matter of tuning catalyst activity is needed, which has been tackled in a combination of experimental and theoretical investigations. However, consistent computations on water oxidation catalysts with similar structures to unravel the ability of influence on the mechanism has been scarce. A novel dinuclear ruthenium-based complex was developed that exhibited unprecedented activity and stability, as well as shifting the reactivity to a different mechanistic pathway. The thorough characterisation of the latter is therefore undertaken with DFT functionals to arrive at a rational explanation of the experimental findings. As it is a fitting model system to compare with other recent catalysts, calculations on the same level of theory are envisioned to reveal the subtle differences that determine the underlying mechanism.

This thesis is partitioned in the following way:

Chapter 2 gives a short introduction to the theoretical foundation of the methods mentioned through-

out the following chapters. It starts with the reasoning that lead to the development of Hartree-Fock theory which is the basis of modern wavefunction theory and is employed in popular density functional approximations. The latter are then derived in general and several levels of sophistication are discussed, taking Perdew's suggested "Jacob's ladder" as an example. As the reactions investigated were all conducted in solution, an introduction to implicit solvation models follows the electronic structure methods. The basic formulae and approximations needed in the development of a geometry optimisation software concludes this chapter.

A description of the procedure used in the optimisation software is then presented in chapter 3, where differences between cartesian and internal schemes, as well as peculiarities of primitive internal coordinates are highlighted. The influence of several approaches in the minimisation is then tested on a medium-sized set of molecular structures with crude starting geometries and demanding structural motifs. Since the actual performance has to be referenced against known codes, the best methods are then applied to benchmark sets proposed in the literature. As a final assessment of efficiency and robustness the optimisation scheme is applied to a transition metal complex.

Chapter 4 examines the possibility of characterising the reduction intermediates in solution of a recently proposed CO<sub>2</sub> reduction catalyst by density functional approximations. With a functional of balanced accuracy for vibrational frequencies and computational cost, the basis for the investigation is set by the calculation of experiment-theory shifts from the crystal structure of the complex. The established procedure is then applied to a comprehensive characterisation of its three reduction intermediates by correlating a large number of relative vibrational shifts.

In chapter 5 the oxidation of water to dioxygen with homogeneous ruthenium-based dinuclear catalysts is investigated. Particularly density functional methods are examined for describing mechanisms of the newly characterised complex Ru-Mebbp.<sup>[13]</sup> Several effects such as functional sophistication, exact exchange percentage, oxidation state of the catalyst and influence of the model system are studied. With these results it is compared to other literature known catalysts, Ru-Hbpp<sup>[14]</sup> and Ru-cbim<sup>[15]</sup>, on the basis of the mechanistic barrier heights. Because of the structurally similar Ru-Hbpp and Ru-Mebbp, both are subjected to an inspection of the minimum energy path for the mechanism determining state to unearth the catalytically important differences between both catalysts.





## 2 Theory

At the beginning of the twentieth century it became apparent that any description of particles which would still be appropriate at atomistic scales needed to take wave-like behaviour into account. This required a new paradigm, culminating in the development of quantum mechanics. The ruling physics are described by the time-dependent Schrödinger equation

$$\hat{H}\Psi(x, t) = i\frac{h}{2\pi} \frac{\partial\Psi(x, t)}{\partial t},$$

where  $\Psi(x, t)$  is the wavefunction containing all of the information about the system,  $h$  is Planck's constant and  $\hat{H}$  the Hamiltonian operator. For an isolated system without any external perturbations the Hamiltonian is independent of the time, which allows the separation of the wavefunction into a spatial and time-dependent part. These can thus be solved separately, leading to a simple phase factor for the time-dependent part and the usually applied time-independent Schrödinger equation

$$\hat{H}\Psi(x) = E\Psi(x). \quad (2.1)$$

Here the eigenvalue  $E$  corresponds to the total energy of the system and the Hamiltonian is easily expressed in Hartree atomic units

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee} = -\frac{1}{2} \sum_a^M \frac{1}{m_a} \nabla_a^2 - \frac{1}{2} \sum_i^N \nabla_i^2 + \sum_{a<b}^{M,M} \frac{Z_a Z_b}{r_{ab}} - \sum_{i,a}^{N,M} \frac{Z_a}{r_{ia}} + \sum_{i<j}^{N,N} \frac{1}{r_{ij}}, \quad (2.2)$$

where  $m_a, Z_a$  are the mass and charge of nucleus  $a$ ,  $\mathbf{r}_i, \mathbf{r}_a$  the position vector of electron  $i$  and nucleus  $a$ ,  $N$  the total number of electrons and  $M$  the total number of nuclei. As such the Hamiltonian in equation (2.2) simply contains all the kinetic and potential energy contributions between the nuclei and electrons. Notwithstanding the apparent simplicity of equation (2.1), not much is known about its solutions  $\Psi$  except that it should be single-valued, square integrable and that its square norm  $|\Psi|^2$  can be interpreted as a probability density. This conundrum is partly solved by invoking the *variational principle*, which is based on the assumption of a complete orthonormal set of functions  $\Psi_i$  in which a trial wavefunction can be expanded with  $\tilde{\Psi} = \sum_i^\infty c_i \Psi_i$ . Taking  $E_0$  to be the exact energy eigenvalue to  $\Psi$ , it can then easily be shown that  $E_0$  is always lower than the eigenvalue of any approximate expansion of  $\Psi$

$$E_0 \leq \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}, \quad (2.3)$$

where the bracket notation has been used to signify integration over all variables and space. Thus the variational principle states that any model for the wavefunction can be used to approximate the true energy and the best solution within this model is found by minimisation.

Nevertheless, solving equation (2.1) is additionally complicated by the terms  $\hat{V}_{NN}$ ,  $\hat{V}_{eN}$  and  $\hat{V}_{ee}$ .

which depend explicitly on the positions of two particles and implicitly on all of them. Consequently any system with more than one nucleus and electron cannot be solved analytically as all movements are coupled. In the *Born-Oppenheimer approximation*<sup>[16]</sup> advantage is taken of the huge difference in mass between the two types of particles which results in a generally much higher velocity for the electrons than the nuclei. As such the nuclei can be assumed as fixed within the timescale of electronic rearrangements, resulting in an electronic wavefunction  $\Psi(\mathbf{r}_i; \mathbf{r}_a)$  that only depends parametrically on the positions of the nuclei. The only problematic term of equation (2.2) that remains is then the electron-electron interaction  $\hat{V}_{ee}$ , which is elegantly treated in the Hartree-Fock method by applying a model theory.

## 2.1 Hartree-Fock Theory

In the Hartree-Fock method the electrons are approximated as quasi-independent particles which only interact with each other through an average potential. This would allow the corresponding wavefunction  $\Psi$  to be written as a Hartree product  $\Psi = \prod_i^N \psi_i$  of one-electron functions, the molecular orbitals. However, the Hartree product does not comply with the Pauli principle,<sup>[17]</sup> which states that the wavefunction has to change sign if two electrons are exchanged. In the Hartree-Fock method a single Slater determinant is therefore used, which automatically ensures the antisymmetry requirement

$$\Psi_{\text{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_1(2) & \dots & \psi_1(N) \\ \psi_2(1) & \psi_2(2) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(1) & \dots & \dots & \psi_N(N) \end{vmatrix}, \quad (2.4)$$

where  $N$  is the number of electrons and  $\psi_i$  the spin-orbital  $i$  with the coordinates of the electron enumerated by the columns of the determinant. The spin-orbitals thus contain the spatial orbital  $\phi_i$  and the spin function  $\sigma$  which is limited to the two values  $\alpha$  and  $\beta$  that correspond to majority and minority spin, respectively. The spin contributions can be integrated separately and the working equations derived with the spatial parts only. In the special closed-shell case every two electrons occupy the same orbital and thus only  $\frac{N}{2}$  are needed.

Since the Hamiltonian in equation (2.2) contains only one- and two-electron operators, applying the Slater-Condon rules<sup>[17]</sup> to the Slater determinant (2.4) gives the energy expression

$$E = \langle \Psi_{\text{SD}} | \hat{H} | \Psi_{\text{SD}} \rangle = 2 \sum_i^{N/2} \langle i | \hat{h} | i \rangle + \sum_i^{N/2} \sum_j^{N/2} [2 \langle ii | jj \rangle - \langle ij | ji \rangle] \quad (2.5)$$

$$\langle i | \hat{h} | j \rangle = \int \phi_i^*(\mathbf{r}_1) \hat{h} \phi_j(\mathbf{r}_1) d\mathbf{r}_1 \quad (2.6)$$

$$\langle ij | kl \rangle = \int \phi_i^*(\mathbf{r}_1) \phi_j(\mathbf{r}_1) \frac{1}{r_{12}} \phi_k^*(\mathbf{r}_2) \phi_l(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (2.7)$$

where equation (2.6) employs the one-electron operator  $\hat{h}_i = -\frac{1}{2}\nabla_i^2 + \sum_a^M \frac{Z_a}{r_{ai}}$ . This energy expression then has to be minimised with respect to the orbitals for the optimal solution according to the variational principle and under the constraint of keeping the orbitals  $\phi_i$  orthonormal. The method of Lagrange multipliers is used to acquire a functional with this constraint, which is then minimised and results in

the eigenvalue equations

$$\hat{f}_i \phi_i = \sum_j^{N/2} \epsilon_{ij} \phi_j. \quad (2.8)$$

Because the total wavefunction  $\Psi$  is not changed by a unitary transformation, the orbitals can be chosen so that they diagonalise the Fock matrix, which leads to a number of pseudo-eigenvalue equations

$$\hat{f}_i \phi'_i = \epsilon_i \phi'_i. \quad (2.9)$$

The molecular orbitals with this property are called *canonical orbitals*. The above Fock operator  $\hat{f}_i$  only contains effective one-electron operators

$$\hat{f}_i = \hat{h}_i + \hat{g}_i = \hat{h}_i + \sum_j^{N/2} [2\hat{j}_{ij} - \hat{k}_{ij}], \quad (2.10)$$

where  $\hat{g}_i$  is the operator accounting for the electron-electron interaction in an average way and depends on the Coulomb and exchange operators  $\hat{j}_{ij}$  and  $\hat{k}_{ij}$ . The Coulomb operator describes the classical repulsion of two charge distributions, whereas the exchange operator is a completely quantum mechanical phenomenon and results from the anti-symmetry requirement of the wavefunction  $\Psi$ . This can directly be seen by their respective integrals

$$J_{ij} = \langle i | \hat{j}_{ij} | j \rangle = (ii|jj) = \int \frac{|\phi_i(\mathbf{r}_1)|^2 |\phi_j(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \quad (2.11)$$

$$K_{ij} = \langle i | \hat{k}_{ij} | j \rangle = (ij|ji) = \int \frac{\phi_i^*(\mathbf{r}_1) \phi_j(\mathbf{r}_1) \phi_j^*(\mathbf{r}_2) \phi_i(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.12)$$

From equation (2.10) it is obvious that terms with  $i = j$  occur, which are unreasonable in the physical sense that an electron cannot interact with itself. The Coulomb term  $J_{ii}$  would therefore be responsible for the effect of *self-interaction*, but these contributions are exactly canceled by the exchange term  $K_{ii}$ .

Nevertheless, the particular structure of the molecular orbitals is still not known so that an approximation is necessary. For this an expansion of the molecular orbitals  $\phi_i$  in a set of basis functions  $\chi_i$  is applied

$$\phi_i = \sum_{\mu} c_{\mu i} \chi_{\mu}. \quad (2.13)$$

Inserting equation (2.13) into the eigenvalue equations (2.9) and integrating leads to the Roothaan-Hall formulation in matrix representation

$$\mathbf{FC} = \mathbf{SCE} \quad (2.14)$$

$$F_{\mu\nu} = \langle \chi_{\mu} | \hat{f} | \chi_{\nu} \rangle \quad (2.15)$$

$$S_{\mu\nu} = \langle \chi_{\mu} | \chi_{\nu} \rangle, \quad (2.16)$$

where  $\mathbf{F}$  is the Fock matrix with integrals over all basis functions,  $\mathbf{C}$  is the expansion coefficient matrix,  $\mathbf{S}$  contains the overlap between each function and  $\mathbf{e}$  is the corresponding eigenvalue vector. The coefficients  $c_{\mu i}$  are then optimised in an iterative solution of equation (2.14) until a minimum of the total energy has been found.

In the above derivations it was exemplified that the ansatz of a single determinant for the total wavefunction leads to a mean-field treatment of the electron-electron interaction. A consequence is the neglect of electron correlation  $E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$  which is important in the description of reactions with a strong change in bonding and weak interactions. Nevertheless the Hartree-Fock method usually retrieves about 99 % of the total energy of a system and thus forms a reasonable basis for more involved methods that employ multiple Slater determinants. The single determinant approach has the additional benefit of being consistent with the chemists view of an independent electron in an orbital which determines its reactivity, as expressed in Koopman's theorem for the calculation of ionisation potentials.

## 2.2 Density Functional Theory

The Hartree-Fock method leads to an approximate solution of the Schrödinger equation that retrieves a large portion of the total energy. Nevertheless, the wavefunction depends on  $4N$  variables, that is three spatial and one spin coordinate per electron. A route to the total energy with a reduced number of variables and thus computational demand would consequently be desired.

In general the electronic density  $\rho$  is ideal in that regard as it only depends on three spatial coordinates, defines the number of electrons  $N = \int \rho(\mathbf{r})d\mathbf{r}$  and the position of the nuclei by its maxima and cusps. The proof that the density uniquely defines the ground-state properties of an  $N$ -electron system was developed by Hohenberg and Kohn.<sup>[18]</sup> In the first theorem they showed that the external potential  $v_{\text{ext}}$ , which is exerted by the nuclear framework, determines the electronic density. There may be a direct correspondence, or over the wavefunction

$$v_{\text{ext}} \rightarrow \Psi_0 \rightarrow \rho$$

as shown in the previous section. If the mapping is unique so that each  $v_{\text{ext}}$  determines a certain density, the process should be invertible and the ground-state wavefunction can be determined from the electronic density. Without a variational principle similar to wavefunction methods this correspondence would not be practical. However, Hohenberg and Kohn proved in their second theorem that this principle also applies to a variation of the density.

Calculation of the total electronic energy is then simply a matter of determining the contributions resulting from the attraction of electrons and nuclei  $V_{\text{eN}}$ , the nucleic repulsion  $V_{\text{NN}}$ , as well as the potential and kinetic terms  $V_{\text{ee}}$  and  $T$  that arise due to interelectronic interactions. These terms should be universal, but  $V_{\text{ee}}$  and  $T$  are unknown. Kohn and Sham<sup>[19]</sup> thus suggested the use of orbitals, since they are directly connected to the density

$$\rho = \sum_i |\psi_i|^2. \quad (2.17)$$

The orbitals are introduced by assuming a system of non-interacting electrons with the same density as the real system. Similarly to Hartree-Fock theory the system can then be described by a single Slater determinant, giving the same contributions to the coulomb  $J[\rho]$  and kinetic energy  $T_s[\rho]$

$$E[\rho] = V_{\text{NN}} + V_{\text{eN}}[\rho] + T_s[\rho] + J[\rho] + E_{\text{xc}}[\rho] \quad (2.18)$$

$$T_s[\rho] = -\frac{1}{2} \sum_i \langle \psi_i | \nabla^2 | \psi_i \rangle \quad (2.19)$$

$$J[\rho] = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \quad (2.20)$$

$$E_{xc}[\rho] = \Delta T[\rho] + \Delta V_{ee}[\rho] = T - T_s[\rho] + V_{ee} - J[\rho]. \quad (2.21)$$

The dilemma of finding the exact terms for the kinetic and electron repulsion energy has thus been transformed into the problem of determining small corrections to the kinetic energy  $\Delta T[\rho]$  and electron repulsion  $\Delta V_{ee}[\rho]$ , assembled in the exchange correlation functional  $E_{xc}[\rho]$ .

The simplest approximation to the exchange-correlation functional was derived from the uniform electron gas (UEG),<sup>[20,21]</sup> a fictitious infinite system of electrons with a positive background charge and constant density. For systems with a slowly varying density even this crude approximation is then assumed to be reasonably accurate if applied *locally*. Dirac<sup>[21]</sup> and Slater<sup>[22]</sup> derived the exchange contribution to this *Local Density Approximation* (LDA) already before the rigorous development of DFT to

$$E_x[\rho] = -\frac{3}{4} \left( \frac{6}{\pi} \right)^{1/3} \int \rho_\alpha^{4/3}(\mathbf{r}) + \rho_\beta^{4/3}(\mathbf{r}) d\mathbf{r}, \quad (2.22)$$

where the integral explicitly depends on the densities of both spin-functions  $\alpha$  and  $\beta$  since only electrons of like spin contribute to the exchange part. The solutions to the correlation part cannot be solved analytically, but highly accurate numerical values have been obtained by Quantum Monte Carlo methods.<sup>[23]</sup> Parameterised functionals can then be designed by fitting to these results as exercised by Vosko, Wilk and Nusair for example.<sup>[24]</sup> The surprising result that finite systems like molecules can be described by a local adaption of formulae derived for a constant density is offset by the only fair accuracy of the LDA. It retrieves 90 % of the exact exchange energy as returned by Hartree-Fock, but leads to a large overbinding and therefore too low barrier heights. On top of that the self-interaction is not exactly cancelled as in HF theory which results in an excessive delocalisation of the electron density.

A range of corrections were therefore suggested, which have been assembled into a hierarchical *Jacob's Ladder* by Perdew,<sup>[25]</sup> that is ordered by functional complexity. LDA forms the basis of this ladder as the simplest approximation to the exact exchange-correlation functional. The second rung therefore also includes the gradient of the density, resulting in the development of the *generalised gradient approximation* (GGA).<sup>[26]</sup> A correction  $\Delta E_{xc}$  to the LDA has to be developed that explicitly depends on the reduced density gradient  $\chi_\sigma = \frac{|\nabla \rho_\sigma(\mathbf{r})|}{\rho_\sigma^{4/3}(\mathbf{r})}$ , which is either done by invoking constraints or by parameters that are fit to experimental values. One of the first successful GGAs that is still in use today is the BP86 functional, constructed from the B88 exchange<sup>[27]</sup> and P86<sup>[28]</sup> correlation functional because it is known to give structural parameters in good agreement with experiment.

Further improvement can consequently be envisioned by adding second derivatives or the kinetic energy density  $\tau(\mathbf{r}) = \frac{1}{2} \sum_i |\nabla \psi_i(\mathbf{r})|^2$  which makes the third rung: *meta-GGAs*. With the development of (meta-)GGAs the inclusion of empirically fitted parameters was initiated to achieve better agreement with experiment or a broader applicability. Famous congeners of recent meta-GGA functionals are the non-empirical TPSS<sup>[29]</sup> which includes the laplacian of the density and the kinetic energy density and has been shown to perform well for transition metal compounds.<sup>[30]</sup> Local meta-GGAs of the Minnesota range were also constructed to be used on transition metal complexes, but the M06-L functional<sup>[31]</sup> already contains 34 parameters fitted to experimental values. Several reasons for both approaches can

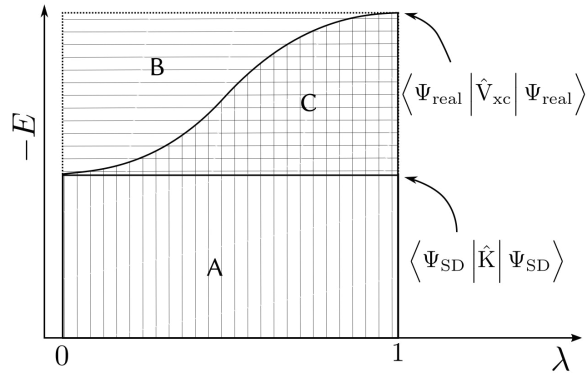


Figure 2.1: The adiabatic connection. The exact result is at  $(1, \langle \Psi_{\text{real}} | \hat{V}_{\text{xc}} | \Psi_{\text{real}} \rangle)$  which corresponds to the area  $A + B$ . With a non-interacting system  $\Psi_{\text{SD}}$  only exchange remains and the integral is thus equal to the Hartree-Fock exchange  $\langle \Psi_{\text{SD}} | \hat{K} | \Psi_{\text{SD}} \rangle$ . The unknown contribution to the exact exchange-correlation functional of the fully interacting system is  $C$ . Scheme adapted from [33].

be found,<sup>[32]</sup> but neither has so far resulted in huge advantages over the other.

Despite the improved accuracy to experiment in (meta-)GGAs they still suffer from the same problems as LDA: self-interaction and insufficient description of non-local effects like charge-transfer, static correlation and dispersion interaction. The problem of self-interaction derives from the assumption of a non-interacting system of electrons in approximate density functionals. In Hartree-Fock theory this artificial contribution is cancelled exactly by the exchange term, which is only included in DFT within the model of the exchange-correlation functional. The value of the exact exchange-correlation energy can be expressed as a function of a parameter  $\lambda$  that, varying from zero to one, slowly turns the interaction between the electrons on. Integration then yields

$$E_{\text{xc}} = \int_0^1 \langle \Psi^{\text{real}}(\lambda) | \hat{V}_{\text{xc}}(\lambda) | \Psi^{\text{real}}(\lambda) \rangle d\lambda, \quad (2.23)$$

where  $\Psi^{\text{real}}$  is the wavefunction of the fully interacting system and  $\hat{V}_{\text{xc}}$  the exact interaction. Equation (2.23) is called the *adiabatic connection* and suggests the use of exact exchange in functional development, which is schematically visualised in Figure 2.1.<sup>[33]</sup> As it is the integral from zero to one below the curve, it can be separated into several contributions. The area  $A$  corresponds to the exact exchange energy of a non-interacting system as determined by the operator from Hartree-Fock theory.  $B$  is the area within the rectangle on top of  $A$ , which is the difference between the integrals of the exact exchange-correlation  $E_{\text{xc}}^{\text{real}}$  and the HF exchange  $E_{\text{x}}^{\text{HF}}$ . If the fraction  $z = \frac{C}{B}$  and  $\hat{V}_{\text{xc}}$  were known, the exact result  $E_{\text{xc}} = E_{\text{x}}^{\text{HF}} + z(E_{\text{xc}}^{\text{real}} - E_{\text{x}}^{\text{HF}})$  could be calculated from the non-interacting system. Since the exact exchange-correlation is not known either  $E_{\text{xc}}^{\text{real}}$  is approximated with an arbitrary functional and  $z$  is used in a parameter  $a = 1 - z$ <sup>[33]</sup>

$$E_{\text{xc}} = (1 - a)E_{\text{xc}}^{\text{DFT}} + aE_{\text{x}}^{\text{HF}}. \quad (2.24)$$

Thus the incorporation of exact exchange from the Hartree-Fock theory is expected to be a promising route to better density functionals, albeit at the necessary introduction of the arbitrary parameter  $a$ . One of the most often utilised GGA *hybrid functionals* is B3LYP,<sup>[34]</sup> which contains B88<sup>[27]</sup> exchange

and LYP<sup>[35]</sup> for the correlation correction

$$E_{xc}^{B3LYP} = aE_x^{HF} + (1 - a)E_x^{LDA} + bE_x^{B88} + E_c^{LDA} + c(E_c^{LYP} - E_c^{LDA}). \quad (2.25)$$

The three parameters  $a = 0.20$ ,  $b = 0.72$  and  $c = 0.82$  have been determined by fitting to thermodynamic data of small organic molecules, for which it achieves impressive accuracy. In addition it was used successively to describe reactions with transition metal compounds, although it was suggested that the amount of exact exchange be lowered to acquire better spin-state energetics.<sup>[36]</sup> The ideal amount of exact exchange is thus not constant, but varies with the system under study.

Some deficiencies of the mentioned DFT functionals were already introduced due to the assumptions made. By employing a single Slater determinant the calculation of systems with multiple low-lying electronic states is problematic, as spin-polarised calculations can lead to a large contamination of the result with states of higher multiplicity. In multinuclear transition metal complexes just two unpaired electrons can exhibit a certain degree of coupling between the two spin centres which already requires several determinants for an appropriate wavefunction. The *broken-symmetry* ansatz<sup>[37-39]</sup> solves this problem by deliberately employing a single determinant guess  $\Psi_{BS}^{guess} = |(\text{core})\eta_a\bar{\eta}_b|$  with a wrong spin symmetry.  $\eta_a$  and  $\eta_b$  are essentially localised molecular orbitals obtained from a linear combination of the singly occupied MOs as returned from the SCF procedure. A variational procedure on the guess yields the broken-symmetry solution  $\Psi_{BS} = |(\text{core})\tau_a\bar{\tau}_b|$ . However, the relaxation of  $\eta_a$  and  $\eta_b$  is not constrained in any way and they may therefore return to a closed-shell solution. Otherwise a spin-polarised result is obtained that reduces the total energy due to additional flexibility in the variational procedure of relaxing  $\Psi_{BS}^{guess}$ . The same determinant is then often used to describe the spin-state energetics by a model Hamiltonian<sup>[40]</sup> approach or to correct errors due to spin-contamination.<sup>[41]</sup>

Because of the local nature of most DFT functionals the description of dispersion interactions is evidently rudimentary, as the asymptotic interaction declines exponentially.<sup>[32]</sup> As many systems' stability depends on the inclusion of dispersion effects, several corrections have been suggested. One of the simplest, but still accurate is the empirical D3 method by Grimme and co-workers<sup>[42]</sup> which employs reference data to determine functional-dependent corrections. With the efficient Becke-Johnson damping<sup>[43,44]</sup> the term reads

$$E_{\text{disp}}^{\text{D3(BJ)}} = -\frac{1}{2} \sum_{A \neq B} \left[ s_6 \frac{C_6^{\text{AB}}}{R_{\text{AB}}^6 + [f(R_{\text{AB}}^0)]^6} + s_8 \frac{C_8^{\text{AB}}}{R_{\text{AB}}^8 + [f(R_{\text{AB}}^0)]^8} \right] \quad (2.26)$$

$$f(R_{\text{AB}}^0) = a_1 R_{\text{AB}}^0 + a_2 \quad (2.27)$$

$$R_{\text{AB}}^0 = \sqrt{\frac{C_8^{\text{AB}}}{C_6^{\text{AB}}}} \quad (2.28)$$

where  $s_6 = 1$  for GGA and hybrid functionals, and  $s_8$ ,  $a_1$ ,  $a_2$  are free fit parameters.

Further developments on density functionals has led to range-separated methods<sup>[45-47]</sup> that try to capture more of the non-local effects which are important in charge-transfer processes, double-hybrid GGAs<sup>[48-50]</sup> that include correlated wavefunction methods to capture dispersion and generally achieve better thermochemistry and the random-phase approximation<sup>[51]</sup> which adds correlation from an inclusion of the virtual orbitals and is therefore seen as the fifth and last rung of Jacob's ladder.

### 2.3 Implicit Solvation Models

The electronic energy determined by wavefunction or density functional methods, with the added interaction of the nuclei, corresponds to the inner energy  $U$  of a system in gas phase. Calculating the same observable for a molecule in solution requires the application of a model for the solvent. This can be accomplished by *explicit* inclusion of a large number of solvent molecules and the successive averaging over all significant conformations, or by an *implicit solvation* method which models the bulk effects with a continuum approach. The simplification of the molecular system due to the neglect of solvent molecules is traded for an increase in the number of approximations needed to describe the important interactions. In accordance with many textbooks on this topic, the following equations are expressed in the esu unit system.

The change in free enthalpy associated with the transfer from gas phase to solution can be expressed by

$$\Delta G_s^\circ = \Delta G_{\text{ENP}} + G_{\text{CDS}} + \Delta G_{\text{conc}}^\circ, \quad (2.29)$$

where  $\Delta G_{\text{ENP}}$  contains the terms due to electronic (E), nuclear (N) and polarisation (P) contributions<sup>[33,52]</sup>

$$\Delta G_{\text{ENP}} = \langle \Psi^{\text{sol}} | \hat{H} | \Psi^{\text{sol}} \rangle - \langle \Psi^{\text{gas}} | \hat{H} | \Psi^{\text{gas}} \rangle + \langle \Psi^{\text{sol}} | \hat{G}_{\text{P}} | \Psi^{\text{sol}} \rangle = \Delta E_{\text{EN}} + G_{\text{P}}. \quad (2.30)$$

If the molecular geometry is assumed to be the same in condensed and gas phase the nuclear term vanishes. The remaining parts are then the difference in electronic structure  $\Delta E_{\text{E}}$  and the polarisation of the solute  $G_{\text{P}}$ .  $G_{\text{CDS}}$  is the effect of cavitation (C), change in dispersion energy (D) and solvent structure (S), whereas  $\Delta G_{\text{conc}}^\circ$  is the contribution due to a change in standard states between both phases and amounts to a value of 7.9 kJ/mol for a change from 1 atm to 1 mol/l.

By moving the molecule from the gas phase into solution the charge distribution of the solute interacts with the solvent and is consequently polarised, adding a favourable contribution  $G_{\text{P}}$  to the free enthalpy. However, the solvent is generally also polarisable, which modulates the capability to influence the solute's electronic structure, resulting in the positive increase  $\Delta E_{\text{E}}$  of the solute's electronic energy. The stabilisation due to mutual interaction thus has to be determined iteratively until a stationary point is found. In *continuum solvent models* (CSMs) the solute which is in interaction with an electric field is taken to reside within a cavity.

The electrostatic interaction of the solute's charge distribution with the solvent continuum is then calculated classically by the Poisson equation

$$\nabla [\epsilon(\mathbf{r}) \cdot \nabla \varphi(\mathbf{r})] = -4\pi\rho(\mathbf{r}), \quad (2.31)$$

with  $\epsilon(\mathbf{r})$  being the position-dependent dielectric constant,  $\varphi(\mathbf{r})$  the electric field potential and  $\rho(\mathbf{r})$  the charge distribution. An explicit dependence of the dielectric constant on the position within the solution allows to take into account varying regions of polarity as within large solute proteins ( $\epsilon \approx 2-5$ ) or the surrounding water environment ( $\epsilon \approx 80$ ). However, equation (2.31) is only valid if no ions are taken into account, that is the ionic strength is assumed to be zero. To appropriately take screening exerted by these into account the Poisson-Boltzmann (PB) equation is employed

$$\nabla [\epsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] - \epsilon(\mathbf{r}) \lambda(\mathbf{r}) \kappa^2 \frac{k_{\text{B}} T}{q} \sinh \left[ \frac{q\varphi(\mathbf{r})}{k_{\text{B}} T} \right] = -4\pi\rho(\mathbf{r}), \quad (2.32)$$



where  $q$  is the charge of the ions,  $\lambda$  is zero or one, depending on whether  $\mathbf{r}$  can be accessed by the ions and  $\kappa$  is the inverse of the Debye length which measures the distance the field can permeate into the solution. Since equation (2.32) is difficult to solve, it is usually linearised by assuming low ionic strength, which exchanges the hyperbolic sine with a linear term of  $\varphi(\mathbf{r})$ .

The PB equation can be solved analytically for some symmetric solute structures like spheres and ellipsoids, but these are accurate for relatively few molecular structures. Attempts have been made to change the radius of the sphere to an effective value to compensate this effect. Since the radius is then very system dependent comparison to other values is difficult and the influence on the results can be very strong. As such cavities in the shape of the molecular surface are preferred, which can be acquired in two different ways. The simplest approach is creating a cavity by overlapping spheres with a scaled van-der-Waals radius that are centred at the atomic positions. However, this may lead to enclosed spaces that are not within the cavity, but separated from the outer continuum and thus can introduce errors. A more involved method therefore depends on determining the solvent-accessible surface (SASE) which is modeled by moving a sphere over the molecular van-der-Waals surface. The resulting SASE thus only contains areas that can be in direct contact with solvent or electrolyte molecules.

Since the Poisson equation has to be solved numerically, the space occupied by the solute molecule is often discretised with a number of grid points. The charge distribution is then mapped onto the grid points by dividing the charge according to some rationale and for each of the points the Poisson equation can then be solved numerically. An alternative possibility is to project the potential onto the surface of the cavity in the form of a charge density.<sup>[53]</sup> The molecular surface is then divided into spherical triangles and the enclosed charge is collected in the midpoint which simplifies the integral over the charge density to a sum over interactions between point charges. This approach can be problematic in cases where midpoints of triangles corresponding to different atoms are very close, which leads to a large interaction contribution that can destabilise the procedure. In this case triangles in regions of overlap between spheres of different atoms are ignored, thus reducing the accuracy.

As the charge density on the molecular surface has a direct influence on the electronic structure of the solute, the reaction field  $V$  resulting from the charge has to be included in the Hamiltonian<sup>[33]</sup>

$$(\hat{H} - \frac{1}{2}V)\Psi = E\Psi. \quad (2.33)$$

In the case of a single determinant approximation to the total wavefunction, the reaction field is directly included in the self-consistent solution

$$(\hat{f}_i - V)\psi_i = e_i\psi_i, \quad (2.34)$$

therefore effecting the relaxation of the mutual polarisation.

The described procedure is for example utilised in the *C*Ontinuum Solvent *M*Odel (COSMO) by Klamt and co-workers,<sup>[54,55]</sup> or in the *P*olarisable *C*ontinuum *M*odel (PCM) by Miertuš, Scrocco and Tomasi.<sup>[56]</sup> COSMO additionally assumes that the dielectric model of the solvent gives similar results as a conductor with  $\epsilon \rightarrow \infty$ . As such no potential exists within the solvent, but an image charge is created on the surface which largely simplifies the calculation of the free enthalpy of polarisation  $G_p$ . Of course the difference in actual dielectric constant and the conductor has to be accounted for, so that the determined free enthalpy of formation is corrected with the Onsager factor  $\frac{2(\epsilon-1)}{(2\epsilon+1)}$ . Still, it could be expected that

such a treatment is only correct for solvents with a high dielectric constant, but it was found that COSMO also works well for non-polar solvents.<sup>[57]</sup>

Up to now only the electrostatic contributions as described by the Poisson equation were taken into account. However, the insertion of a molecule into a solution can also effect changes in the geometry of the solute and first solvation shell, resulting in a different interaction moiety. As such these effects can play an important role in the accurate description of non-polar solvents. The *Solvent Model using the full electron Density* (SMD) by Marenich, Cramer and Truhlar<sup>[52]</sup> incorporates such non-bulk effects in a parameterised manner

$$G_{\text{CDS}} = \sum_k^{\text{atoms}} \sigma_k A_k + \sigma^{[\text{M}]} \sum_k^{\text{atoms}} A_k, \quad (2.35)$$

where  $\sigma_k$  is the atomic surface tension of atom  $k$ ,  $\sigma^{[\text{M}]}$  the respective molecular surface tension and  $A_k$  the solvent-accessible surface area. Other methods to calculate the non-bulk effects exist, such as the cavity-dispersion-repulsion terms of PCM methods.<sup>[58]</sup> Nevertheless, it was found that approaches which do not take the arbitrary definition of the cavity into account cannot give quantitative results.<sup>[59,60]</sup> The SMD method solves this by a parameterisation that is consistent for a certain set of solute radii and thus the solute-solvent boundary.<sup>[60]</sup>

Notwithstanding the great variety of different approaches to implicitly account for the effect of molecule solvation, all of the mentioned methods have been used to study reactions and equilibria in solution. Various measurable quantities such as redox potentials,<sup>[61,62]</sup>  $\text{pK}_a$  values<sup>[63-65]</sup> and even infrared<sup>[66]</sup> or UV/Vis spectra<sup>[67]</sup> have been investigated with continuum models, which have an important role in the efficiency, stability and characterisability of molecular catalysts.

## 2.4 Geometry Optimisation

The Born-Oppenheimer approximation to the electronic structure problem allows to decouple the atomic movement from that of the electrons creating the concept of a *potential energy surface* (PES). As such the dynamics of the nuclei can be calculated classically on the surface resulting from a particular energy function. However, in most ab initio methods the analytical energy function is not known and only calculations for a few points on that surface can be carried out. Consequently the full PES is only accessible in the simplest cases, as the dimensionality and cost rise exponentially with the system size and methods of higher accuracy.

For many cases only some points on the PES with sufficiently low energy are needed to describe systems in equilibrium or even reactions. The latter are represented by paths on the PES of lowest energy by going from one minimum to another. These minimum energy pathways (MEPs) correspond to the most likely trajectory for a certain reaction of a system. As shown in Figure 2.2 the minima are separated by a saddle point of first order, the transition state (TS). Since the TS is the point of highest energy on the reaction pathway it is often a good indicator for the reactivity of a system and is used in several models for the prediction of reaction rates.<sup>[68,69]</sup>

Computational investigations of molecular reactions therefore depend on the determination of these stationary points, which is accomplished by numerical *optimisation algorithms*. Several procedures have been developed to specifically minimise the molecular energy or maximise it in a single direction or coordinate. As the accuracy and cost of the electronic structure method is increased the requirements

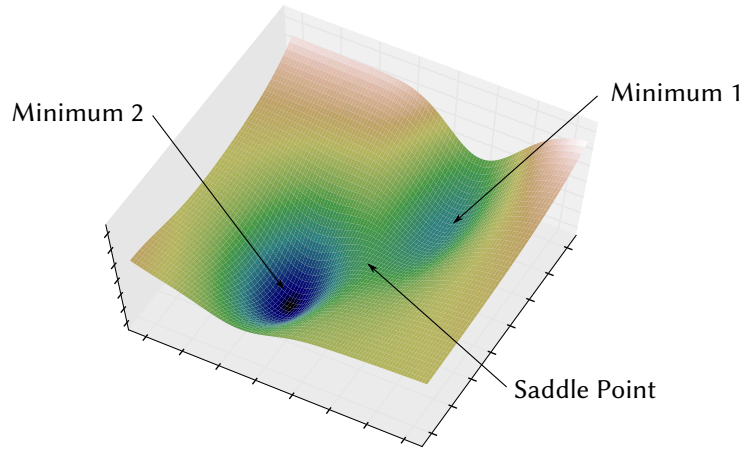


Figure 2.2: Schematic representation of a potential energy surface.

on the efficiency of the optimisation algorithms are stronger. Most of the modern electronic structure methods allow the calculation of the energy gradient at very little additional cost, but higher derivatives are often prohibitively expensive in a general application. However, second order optimisation algorithms require significantly fewer steps until convergence is reached, so that a balance between the number of steps and the computational demand has to be devised. Since the calculation of energetic derivatives becomes more and more prohibitive with the highest order involved, the Taylor expansion of the potential energy surface (PES) is often stopped at second order

$$E(\mathbf{x}) = E(\mathbf{x}_0) + \mathbf{g}_0^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H}_0 \Delta \mathbf{x}, \quad (2.36)$$

with the electronic energy  $E$ , gradient  $\mathbf{g}$  and Hessian  $\mathbf{H}$  at the geometry  $\mathbf{x}$  in the Born-Oppenheimer approximation. Since stationary points on the PES are sought, the first derivative of the energy has to vanish

$$\frac{dE(\mathbf{x})}{d\mathbf{x}} = \mathbf{g}_0 + \mathbf{H}_0 \Delta \mathbf{x} \stackrel{!}{=} \mathbf{0}. \quad (2.37)$$

The required step  $\Delta \mathbf{x}$  to the minimum in the quadratic approximation can be readily derived according to

$$\Delta \mathbf{x} = -\mathbf{H}_0^{-1} \mathbf{g}_0 \quad (2.38)$$

and is equivalent to a Newton-Raphson step for the gradient function  $\mathbf{g}(\mathbf{x})$ .<sup>[3]</sup> Notwithstanding, for PESs that strongly deviate from the quadratic approximation, this step does not lead to the exact solution and an iterative procedure is required. Additionally the step may become very small and thus converge slowly, or very large and overshoot the minimum. For very difficult cases the latter effect can result in oscillative behaviour, leading to a number of unnecessary steps and therefore single point calculations. A simple way to control the step size is to restrict the length of the predicted step to the radius  $\tau$  of a hypersphere in which the quadratic approximation is assumed to still be valid. This can be trivially achieved by choosing a maximum value for the step length and cutting the vector to this value if it is exceeded. More appropriately the step can be determined by applying a Lagrange multiplier  $\lambda$  that minimises  $E(\mathbf{x}) = \frac{1}{2} \lambda ((\Delta \mathbf{x})^2 - \tau^2)$  which leads to the *trust radius method* (TRM)<sup>[3]</sup>

$$\Delta \mathbf{x} = -(\mathbf{H} - \lambda \mathbf{I})^{-1} \mathbf{g}. \quad (2.39)$$

A benefit of the trust radius method is that  $\lambda$  can be used to force the optimisation into the direction of a minimum even if the Hessian matrix is not positive definite. The multiplier  $\lambda$  therefore has to be lower than the lowest eigenvalue of  $\mathbf{H}$ .

In a related approach, the rational function (RF) algorithm, the quadratic approximation to the potential energy surface is changed to a rational function approximation

$$\Delta E(\mathbf{x}) = \frac{\mathbf{g}^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H} \Delta \mathbf{x}}{1 + \Delta \mathbf{x}^T \mathbf{S} \Delta \mathbf{x}}, \quad (2.40)$$

where  $\mathbf{S}$  is an arbitrary matrix. The solution of the above equation, that is  $d\Delta E/d\Delta \mathbf{x} = 0$ , results in an eigenvalue equation in which the eigenvalues are proportional to the predicted energy change  $\Delta E$  and the eigenvectors to the step  $\Delta \mathbf{x}$ .

$$\begin{bmatrix} \mathbf{H} & \mathbf{g} \\ \mathbf{g}^T & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ 1 \end{bmatrix} = 2\Delta E \begin{bmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ 1 \end{bmatrix} \quad (2.41)$$

In the leftmost part of equation (2.41) the Hessian is assembled with the gradient into a larger matrix which is therefore also called the *augmented Hessian*. Since  $\mathbf{S}$  can be any matrix many varieties of the RF procedure are possible. Usually it is set to a constant, say  $\lambda$ , times the identity matrix which reduces the first row in equation (2.41) to the trust region method in equation (2.39). Similarly to TRM the eigenvectors of solving equation (2.41) corresponding to the lowest eigenvalues can be used to optimise stationary points of different order. For a more robust saddle point search the rational function approach is usually partitioned into the eigenvector to be maximised and the minimisation space.<sup>[70]</sup>

During the optimisation to the stationary point the quadratic or rational function approximation to the local potential energy surface should become increasingly better. Consequently the arbitrarily defined threshold for the trust radius  $\tau$  may not be appropriate over the whole range of steps. A common update to the trust radius thus compares the last step and energy to the values of an ideal quadratic surface and the threshold is then adjusted according to the deviation. One typical approach is to estimate the accuracy of the quadratic approximation as the ratio  $\rho$  of the real and predicted change in energy<sup>[3]</sup>

$$\rho = \frac{\Delta E}{\mathbf{g}^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H} \Delta \mathbf{x}}. \quad (2.42)$$

The trust radius is doubled if the approximation fits reasonably well, e.g.  $\rho > 0.75$  and  $\tau^{\text{old}} < \frac{5}{4} |\Delta \mathbf{x}|$ , whereas it is reduced to  $\tau^{\text{new}} = \frac{1}{4} |\Delta \mathbf{x}|$  if the surface deviates strongly from a quadratic description, that is  $\rho < 0.25$ . In between, the trust radius is not changed at all.

As a different approach to determine a step length the previous steps can be inter- or extrapolated to achieve a better geometry. The *direct inversion in the iterative subspace* (DIIS) method<sup>[71]</sup> minimises a set of error vectors  $\mathbf{e}_i$  of previous steps  $i$  and has been used successfully in electronic structure methods to accelerate convergence. Since the error vector can be arbitrarily defined, choosing the Newton-Raphson step  $\mathbf{e}_i = -\mathbf{H}\mathbf{g}_i$  or the change in energy  $\mathbf{e}_i \mathbf{e}_j = (\mathbf{g}_i - \mathbf{g}_j)(\mathbf{x}_i - \mathbf{x}_j)$  enables the geometry optimisation with this method (GDIIS).<sup>[3,72,73]</sup> Several different approaches using DIIS or Newton-Raphson algorithms have been found to be advantageous in separate parts of the optimisation which led to the suggestion of a hybrid method.<sup>[73]</sup> Nevertheless, the GDIIS method is known to be drawn to stationary points of

higher order if the Hessian is not properly conditioned.<sup>[3]</sup> Furthermore, one has to ensure that the set of steps used for the interpolation does not become linearly dependent and a robust DIIS algorithm thus employs a range of checks to ensure non-redundancy and a reasonable step direction.<sup>[74]</sup>

Another option to determine the step length is to employ a *line search* along the predicted direction of the step. This reduces the dimensionality of the problem to one variable by fitting cubic or quartic polynomials to the energy and gradient of the last steps.<sup>[75,76]</sup> However, as the line search is not exact it can also lead to large steps and then additional energy and gradient calculations may be required.

The calculation of supplementary points to determine the step length can be very costly and may only be justified if the improved convergence more than compensates for the increase in the number of gradient and Hessian evaluations. The latter is seldom so fast to calculate that an exact evaluation at every step is computationally competitive. As a first approximation one may assume the identity matrix for the Hessian, which effectively results in a step along the direction of steepest descent, that is the gradient. In the course of the optimisation the second derivatives can be updated numerically with the history of previous steps and gradients. An unlimited number of these *update methods* can be constructed with each exhibiting different benefits and disadvantages.

One of the most widely used algorithm for this task is the one by Broyden, Fletcher, Goldfarb and Shanno (BFGS)<sup>[77-80]</sup>

$$\mathbf{H}_{n+1} = \mathbf{H}_n + \frac{\Delta \mathbf{g}_n \Delta \mathbf{g}_n^T}{\Delta \mathbf{g}_n^T \Delta \mathbf{x}_n} - \frac{\mathbf{H}_n \Delta \mathbf{x} \Delta \mathbf{x}^T \mathbf{H}_n^T}{\Delta \mathbf{x}^T \mathbf{H}_n \Delta \mathbf{x}}, \quad (2.43)$$

which has the important property of keeping the second derivative matrix positive definite and thus enforces a step towards a minimum. The denominator in equation (2.43) contains the difference of the gradient  $\Delta \mathbf{g}_n$  between the last two steps and the corresponding coordinate change  $\Delta \mathbf{x}_n$ . It is therefore evident that the BFGS method can become unstable if the product of both approaches zero.<sup>[4]</sup> Other formulations like SR1<sup>[81]</sup> and PSB<sup>[82]</sup> do not exhibit this instability and also relax the positive definiteness of the correction, which renders them useful in the optimisation of transition states. Due to their specific disadvantages some combinations to improve upon the existing algorithms have also been suggested.<sup>[4,83]</sup>

Using the identity as the second derivative matrix is a very crude approximation. Although it is known that diagonal matrices in combination with well constructed internal coordinates can be sufficient,<sup>[84,85]</sup> a better model Hessian can greatly increase the speed of convergence. General force-field-based analytic functions lend themselves to the development of efficient approximations which was taken up by Schlegel,<sup>[86]</sup> Fischer and Almlöf<sup>[87]</sup> as well as in the group of Malmqvist.<sup>[88]</sup> In the latter (Lindh's model) a super-redundant set of internal coordinates consisting of distances, angles and dihedrals for every combination of atoms is set up and the contribution to the energy is taken to be a simple harmonic term

$$\begin{aligned} E_{\text{stretch}} &= \frac{1}{2} \sum k_{ij}(r_{ij} - r_{0,ij})^2 \\ E_{\text{bend}} &= \frac{1}{2} \sum k_{ijk}(\theta_{ijk} - \theta_{0,ijk})^2 \\ E_{\text{torsion}} &= \frac{1}{2} \sum k_{ijkl}(\tau_{ijkl} - \tau_{0,ijkl})^2 \\ E_{\text{ges}} &= E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} \end{aligned}$$

where the size of the force constant is damped according to the distance between the atoms

$$\begin{aligned}\rho_{ij} &= \exp(\alpha_{ij}(r_{\text{ref},ij}^2 - r_{ij}^2)) \\ k_{ij} &= k_r \rho_{ij} \\ k_{ijk} &= k_\theta \rho_{ij} \rho_{jk} \\ k_{ijkl} &= k_\tau \rho_{ij} \rho_{jk} \rho_{kl}.\end{aligned}$$

The parameters  $\alpha_{ij}$  and  $r_{\text{ref},ij}$  depend on the elements contributing to the bond, but the values are tabulated for the first three rows.<sup>[88]</sup> In the original paper the Hessian in super-redundant internal coordinates is transformed back to cartesian before further use, which renders the model coordinates independent of the actual optimisation space. As a consequence the model is evaluated in a global coordinate system for any molecular structure which justifies the use of the Hessian's eigenvectors as a non-redundant set of coordinates.<sup>[88,89]</sup> A drawback is the huge size of the transformation matrix for systems with more than 50–100 atoms, which are routinely investigated nowadays. To reduce this dependence the calculation of the force constants is often done in the internal coordinate space used in the optimisation.<sup>[85]</sup> Similarly to the quadratic approximation in the methods for the step prediction, the use of a parameterised Hessian corresponds to a different surface than the exact solution. Consequently the step size and direction may differ and thus the convergence to the same stationary point is not ensured.

The question whether a stationary point is reached cannot be answered easily due to the numerical nature of the optimisation procedure. By approaching such a point the gradient is reduced in magnitude and vanishes at the minimum or saddle point. However, the method used to evaluate the energy and gradient only allows a certain accuracy so that several thresholds have to be defined to identify a converged solution. Many different schemes have been suggested and are in use by several program suites.

An often cited scheme was invented by Baker<sup>[84]</sup> to be able to properly compare different optimisation protocols. In that scheme convergence is signalled if the maximum component of the gradient is below  $3 \cdot 10^{-4}$  a.u. and either the maximum displacement ( $3 \cdot 10^{-4}$  a.u.) or the energy change of the last step ( $10^{-6}$  a.u.) is small. Thus, the gradient is the main criterion since it is the necessary condition for a stationary point, whereas the additional checks serve to signal convergence in special cases. The test for a small change in energy helps to converge floppy molecules which can change strongly with a chemically insignificant change in energy.<sup>[84]</sup>

Slightly different schemes are employed by software suites like Molpro<sup>[89]</sup> and Gaussian.<sup>[90]</sup> The former just moves the point at which convergence is checked before the gradient evaluation at the current step and thus saves unnecessary computation time as well as steps.<sup>[89]</sup> The software package Gaussian uses somewhat looser criteria which require several values to be within the limits at the same time.<sup>[91]</sup> In addition to the maximum components of the gradient and the cartesian step being lower than  $4.5 \cdot 10^{-4}$  a.u. and  $1.8 \cdot 10^{-3}$  a.u. respectively, the root-mean-square (RMS) values are required to be within  $3 \cdot 10^{-4}$  a.u. and  $1.2 \cdot 10^{-3}$  a.u. for both quantities. As a further check, if the RMS and maximum value of the gradient are less than a hundredth than the thresholds, convergence is also signalled in that scheme.

Even though the above mentioned methods are all important in an efficient optimisation and partic-

ularly for pathologic cases, the coordinates in which a molecular geometry is described often have an even stronger effect on the rate of convergence. On top of this the coordinates are more sensitive to chemical properties of the system and an efficient set should accordingly take the different interactions into account.

Hence many different coordinate systems have been suggested and used, with the cartesian coordinates being one of the simplest and directly accessible to general optimisation algorithms. Many molecular structures are already supplied or can easily be converted into cartesian coordinates so that they are simple to set up. Large biomolecules like proteins are often optimised in cartesian coordinates since an efficient internal set would be difficult to define and would use up too much memory and CPU cycles in the transformation between the internal and cartesian space.<sup>[92]</sup> Nevertheless, they are by no means ideal because cartesian coordinates also describe the translational and rotational motion of the molecule. The latter is represented by the eigenvectors of the Hessian corresponding to eigenvalues with a value of zero at stationary points. In general cartesian coordinates exhibit a high degree of coupling with regard to the internal degrees of freedom which slows down any optimisation procedure.

One option to reduce the coupling and get rid of the non-internal degrees of freedom is to describe the geometry in the eigenvector space of the Hessian, which also reduces the harmonic coupling of the coordinates. These so-called *local normal coordinates*<sup>[89]</sup> (LNC) therefore work best with an exact Hessian or a good approximation to it. Since the potential energy surface away from the minimum is seldom quadratic, the coordinates change at every step in the optimisation and have to be recalculated. Due to the diagonalisation of the Hessian, translational and rotational contributions can be singled out by examining the eigenvalues, which results in a reduced coordinate set with the necessary  $3N - 6(5)$  degrees of freedom. Although this is true for translation motion, eigenvalues for rotational eigenvectors are only exactly zero at a minimum.<sup>[93]</sup> Consequently care has to be taken to project out rotations from the Hessian matrix before diagonalisation, so that all external degrees of freedom give eigenvalues that are exactly zero.

The latter caveat can be circumvented by using internal coordinates. One of the simplest and best known version of these is the *z-matrix* in which one bond, angle and dihedral is assigned to every atom successively to fully describe the geometry. For this, only  $3N - 6(5)$  coordinates are required in total which is an advantage over the  $3N$  coordinates of optimisations in plain cartesians. Similar to the Hessian eigenvectors, internal coordinates also reduce the harmonic coupling which should enhance the rate of convergence in an optimisation procedure. Notwithstanding, z-matrices for a geometry can be built in multiple ways of which every single construction may show a different convergence behavior in the optimisation.

Another variation are the *natural internal coordinates* which take coordinates similar to the ones used in spectroscopy as an example to build a non-redundant set.<sup>[94]</sup> In addition to the usual bonds, angles and dihedrals coupled coordinates like ring deformations are taken into account, which not only reduces harmonic couplings, but also anharmonic contributions. These have been shown to achieve superior convergence for a variety of molecular structures. However, the problem of setting the coordinates up remains and algorithms defining natural internals can be tedious to implement. It is therefore evident that some cases may not be handled correctly and the corresponding optimisation can fail without taking any step.

Pulay and co-workers thus suggested an easy to build set of coordinates which can and even is

expected to be redundant and still lead to satisfactory convergence.<sup>[95]</sup> These *redundant internal coordinates* (RIC) are created by a set of *primitive coordinates* (PICs) made up of a sensible definition of bonds, angles and dihedrals, but other primitives which are useful for certain geometrical features can be added. Calculating the values and derivatives of these PICs is generally straightforward, only during the actual optimisation when geometry differences are needed some corner cases need to be treated. For one, the sign of any dihedral has to be conserved during the course of the relaxation which can easily be taken care of by comparing to the last value. Another more sensible problem is the transition between normal angles and those which are almost linear. The latter require the definition of two orthogonal vectors which are also perpendicular to the bonding axis of the three linearly arranged atoms. This can either be accomplished by searching for useful reference vectors like Bakken and Helgaker<sup>[85]</sup> suggested, or by simply using the cartesian basis vectors which are nearest to orthogonality.<sup>[96]</sup> In any case, if during an optimisation an angle becomes linear, a suitable transition between the standard definition of a bending coordinate and the corresponding linear bends must be ensured.

With the creation of a suitable primitive internal coordinate set, the determination of current values  $\mathbf{q}$  of this set for a particular optimisation step is trivial. Relating the differences of two cartesian geometries to a change in internal degrees of freedom on the contrary is not directly accessible because of the curvilinear nature of the primitive coordinates. However, infinitesimally small deviations can be approximately described by a linear relationship and thus the transformation of small changes in  $\delta\mathbf{x}$  into internal displacements  $\delta\mathbf{q}$  are realised by Wilson's B-matrix of first derivatives  $\partial q_i/\partial x_j$ .<sup>[97]</sup>

$$\mathbf{B}\delta\mathbf{x} = \delta\mathbf{q}. \quad (2.44)$$

Due to the high redundancy of the internal coordinate system,  $\mathbf{B}$  is rectangular and therefore not directly invertible. A pseudo-inverse matrix  $\mathbf{B}^{-1}$  can be defined which has the property of reducing the norm  $\|\mathbf{B}\delta\mathbf{x} - \delta\mathbf{q}\|$ , or  $\|\delta\mathbf{x}\|$  in the case of over- or underdetermined systems of linear equations, respectively.<sup>[85]</sup> For a non-singular symmetric matrix this results in the standard inverse. Although most linear algebra libraries provide functions to directly calculate the pseudo-inverse, for example by singular value decomposition, it can also be determined by standard methods. This is accomplished by projecting the  $\mathbf{B}$  matrix onto the non-redundant subset, which can be accomplished by building and diagonalising the  $\mathbf{G}$  matrix

$$\mathbf{G} = \mathbf{B}\mathbf{B}^T = (\mathbf{U}\mathbf{R}) \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{U} \\ \mathbf{R} \end{pmatrix}. \quad (2.45)$$

Since  $\mathbf{G}$  is a real symmetric matrix, it can readily be inverted by diagonalisation and inversion of the non-zero eigenvalues. The actual pseudo-inverse is then acquired by

$$\mathbf{B}^{-1} = \mathbf{G}^{-1}\mathbf{B}. \quad (2.46)$$

For the prediction of an optimisation step the gradient and Hessian have to be known in internal coordinates, which can be readily calculated with the pseudo-inverse of the transposed  $\mathbf{B}$  matrix

$$\mathbf{g}_q = (\mathbf{B}^T)^{-1}\mathbf{g}_x \quad (2.47)$$

$$\mathbf{H}_q = (\mathbf{B}^T)^{-1}(\mathbf{H}_x - \mathbf{K})\mathbf{B}^{-1} \quad (2.48)$$



with  $K_{jk} = \sum_i \mathbf{g}_{q,i} \frac{\partial^2 q_i}{\partial x_j \partial x_k}$  being the second derivatives of the internal by the cartesian coordinates. Nevertheless, due to the redundancy in the coordinate system the acquired internal gradients and Hessian might not correspond to a physically meaningful change in the coordinates. A projector  $\mathbf{P} = \mathbf{B}\mathbf{B}^{-1}$  is therefore employed that removes first order redundancies and reduces the influence of numerical noise

$$\tilde{\mathbf{g}}_q = \mathbf{P}\mathbf{g}_q \quad (2.49)$$

$$\tilde{\mathbf{H}}_q = \mathbf{P}\mathbf{H}_q\mathbf{P}. \quad (2.50)$$

After the determination of a new geometrical change  $\mathbf{s} = \mathbf{q}_{n+1} - \mathbf{q}_n$  in internal coordinates, the step has to be transformed back from curvilinear to rectilinear coordinates which necessitates an iterative procedure

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{B}^{-1}\Delta\mathbf{q}_k \quad (2.51)$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{B}^{-1}\mathbf{s} \quad (2.52)$$

$$\Delta\mathbf{q}_k = \mathbf{s} - (\mathbf{q}_k - \mathbf{q}_n). \quad (2.53)$$

The above procedure would require a complete calculation and inversion of the  $\mathbf{B}$  matrix which scales as  $\mathcal{O}(N^3)$  and thus would add a significant amount of computational time for larger molecules. Often  $\mathbf{B}$  is therefore taken to be constant which leads to an increase in the number of iteration steps, but effectively reduces the time needed for the whole back-transformation.<sup>[98]</sup> Since this step is a major bottleneck in the procedure, several attempts have been made to reduce the cost by lowering the overall scaling due to evaluation of the  $\mathbf{B}^{-1}$  matrix<sup>[98,99]</sup> or by utilising polynomial fits in conjunction with automatic differentiation techniques.<sup>[100]</sup>

As seen from equation (2.45) the diagonalisation of  $\mathbf{G}$  reveals the redundancy of the primitive internal coordinate set. This information can therefore be used to create a set of non-redundant internal coordinates with the correct size of  $3N - 6(5)$  degrees of freedom and may be a better representation of the molecular structure. The new set of coordinates are simply the eigenvectors of the non-redundant space  $\mathbf{U}$  as defined in equation (2.45).<sup>[98,101]</sup> Thus they consist of a linear combination of PICs and are consequently often delocalised over the whole molecule. However,  $\mathbf{G}$  might possess several very small nonzero eigenvalues which necessitates a proper definition of a threshold for the selection of the eigenvectors. While the utilisation of a threshold can lead to a larger set of coordinates than necessary, just choosing the  $3N - 6(5)$  vectors corresponding to the highest eigenvalues may neglect important combinations. The main advantage of these *delocalised internal coordinates* (DIC) is the much smaller memory footprint of the method if the intermediate matrices are not kept, since the size of the coordinate set is reduced. On the other hand additional transformations to and from the redundant set have to be carried out and if the coordinates are not determined at every step, a possible mixture of the non-redundant  $\mathbf{U}$  and the redundant space  $\mathbf{R}$  has to be taken into account. The latter effectively leads to slower geometry convergence and thus the delocalised internal coordinates are expected to mostly perform worse than the corresponding redundant set. Nevertheless, in the limit of a perfect set of delocalised coordinates this reduction in efficiency would be vanishing.

If all the mentioned coordinate systems are taken into account, it becomes obvious that all of them suffer from one or more deficiencies and have to be well-designed to be used in routine applications. Cartesian based coordinates are easy to set up and handle, but can lead to slow convergence or even to

saddle points of higher order when a good Hessian approximation is not available. Internal sets require more transformations which can make up the major part of the computation time and memory usage if large molecules are investigated, but offer superior performance. Nevertheless the creation of these coordinates can be error prone (natural internals) or not appropriate for the optimisation of a certain geometry (z-matrix). With the suggestion of the redundant internals many problems influencing the building of the coordinate set and the efficiency of the optimisation have been resolved. The delocalised internals additionally address the memory requirement of the redundant approach, but are not expected to improve the convergence. Since the redundant internals are made up of several simpler coordinates their influence on the optimisation efficiency should be large. This of course depends on the step prediction and Hessian approximation as well. A concise investigation of the influences of every part in the optimisation routine and the performance of several approximations is therefore necessary. To accomplish this and allow the modification of the routine in detail, the software program StOpt was developed and interfaced to the Molpro suite<sup>[102]</sup> for the electronic structure calculations. As Molpro offers several optimisation procedures itself it was additionally used as the reference to benchmark the performance of StOpt.

# 3 Development of an Improved Optimisation Procedure

## 3.1 The StOpt-Software

The general program flow for cartesian- and internal-based optimisation procedures is shown schematically in Figure 3.1. In general both procedures require the same successive steps of determining the energy and gradient, predicting a new step and finally checking the convergence criteria before continuing the cycle. Several additional steps could be added in-between to add constraints, external potentials or to carry out a line search. These modifications would be similar in the flow-chart of both coordinate spaces, however. The change from cartesian to internal coordinates instead requires a more extensive modification of the procedure.

Although the local normal coordinates as used in Molpro<sup>[102]</sup> show similarities to internal coordinates their treatment is similar to a cartesian optimisation as presented in the left-hand side of Figure 3.1. With the translation and rotation projected out of the Hessian matrix their use results in  $3N - 6(5)$  coordinates to optimise, which are determined by diagonalisation and successive transformation with the eigenvectors. Because the transformation is linear it is easily carried out using matrix multiplication. On the contrary internal coordinate schemes require basic chemical information about bonds and the type of their interaction to build an efficient set of primitives. With a robust definition of the PICs the construction of transformation matrices is straightforward, although they require more matrix operations than local normal coordinates. As the step prediction method is only applied in the coordinate system of the optimisation, the last big difference between both flow-charts is the back-transformation. This is a major part of the whole procedure and as it is not exact the iterative algorithm might diverge, which can lead to crashes or unreasonable cartesian geometries.

By default Molpro 2012<sup>[102]</sup> minimises geometries supplied as cartesian coordinates in local normal coordinates with Lindh's global model Hessian. The latter was designed with LNCs in mind and thus is an appropriate match for these coordinates. If no other options are given to the program, its optimiser also uses the rational function ansatz to determine the step and limits it to a length of 0.3 a.u. by scaling the step back if it exceeds that margin. Molpro's optimiser consequently follows the approach of Figure 3.1 with the small but significant difference that the gradient is only calculated after the convergence check, as mentioned in section 2.4.

As StOpt was developed with the goal of an efficient redundant-internal-based optimisation method, represented by the flow-chart on the right hand side of Figure 3.1, a procedure to determine a set of primitives had to be chosen. A super-redundant set of PICs or a total connection scheme in which the distance between all possible combination of atoms are used would be a simple solution, but quickly become unfeasible or even lead to slower convergence. The developed software thus employed a procedure described in the literature<sup>[85]</sup> which defines stretching coordinates in a chemical sense, that is

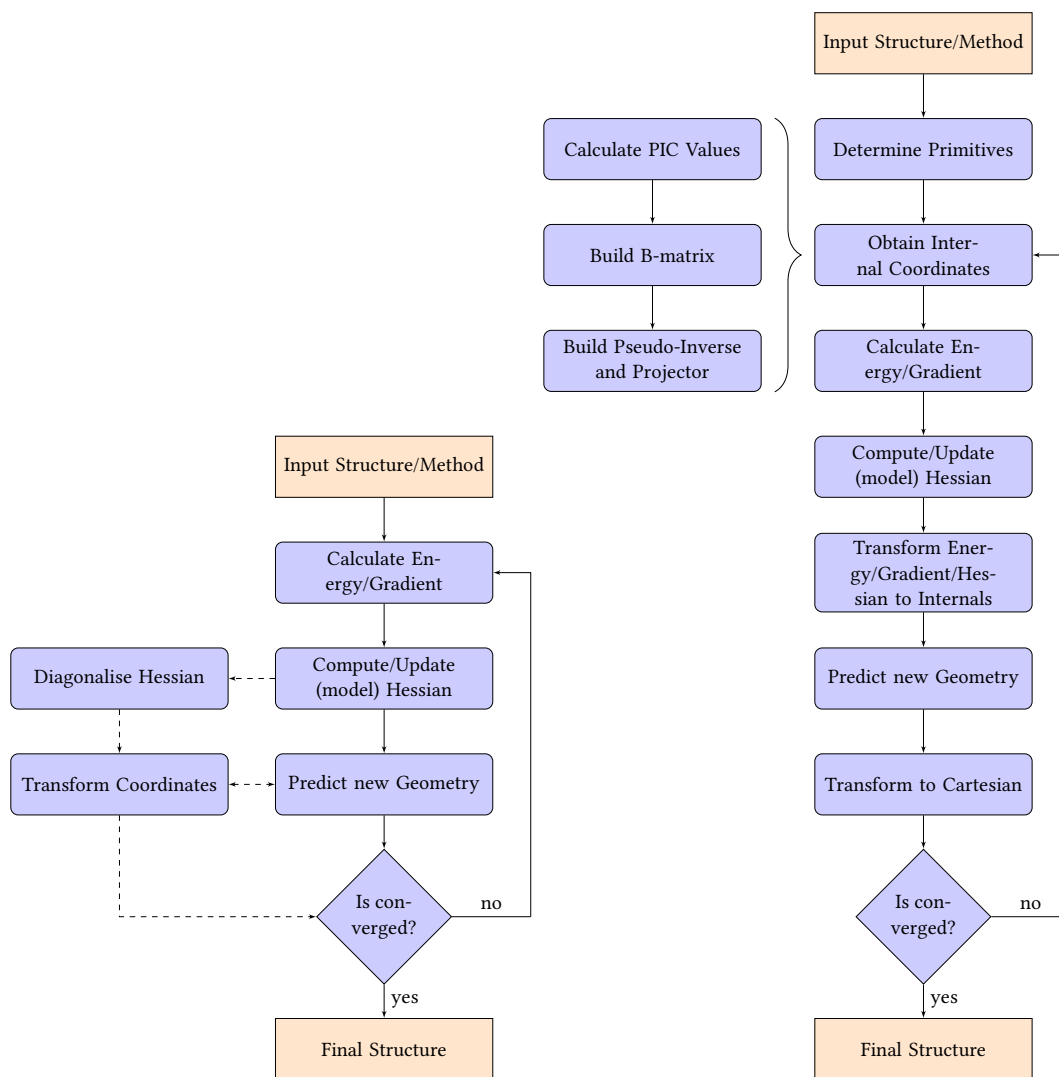


Figure 3.1: General optimising flowcharts in cartesian (left) and redundant internal (right) coordinate space. The dashed lines represent the alternative flow in case of local normal coordinates as used in Molpro's optimiser. As convergence is generally tested on cartesian values, the transformation of coordinates works in both directions for the LNC optimisation. The flowchart on the right corresponds to the approach implemented in StOpt.

by comparing the distance to the sum of covalent radii multiplied by the factor 1.3. An extra-redundant set of stretches can be defined by the same check, but with a larger scaling factor of 2.5, which mostly results in bonds between nearest neighbor atoms. Similarly hydrogen bonds are assigned to any distance between a heteroatom and a hydrogen bound to a heteroatom for which the distance is greater than the sum of the covalent, but lower than the sum of the van-der-Waals radii and with an X-H-X angle larger than  $90^\circ$ . X in this case are the heteroatoms which currently have to be one of the range N, O, F, P, S and Cl. If multiple fragments are identified that are not connected by any coordinates, a regular interfragment (IF) bond is added by searching for the atoms with the smallest distance between the fragments. The extra-redundant set defines auxiliary interfragment bonds for distances between the fragments which are lower than 2.0, but larger than 1.3 times the standard IF bond. Values for the covalent and van-der-Waals radii were taken from tabulated results.<sup>[103,104]</sup>

With the bonds set up the bending and torsional coordinates are determined for every unique combination of connected stretches which are not auxiliary. If an angle between three atoms is larger than  $168^\circ$  and the central atom is only connected to the other two atoms, the bending coordinate is exchanged with two linear bends. In the current implementation the linear bends are taken into account by the two cartesian position vectors of the central atom which are as close to orthogonality with the bond vectors as possible. This is determined by the smallest scalar product of the cartesian basis vectors and the distance vector between the outer atoms. The cartesian direction with the smallest scalar product might change during the optimisation or the back-transformation which could lead to an inconsistent set of coordinates. To avoid problems of this kind the indices of the chosen directions are only assigned at the beginning of the optimisation.

Torsional coordinates are similarly defined by three non-auxiliary bonds that share one atom with the next bond and do not exert any angles larger than  $168^\circ$ . If the molecule contains any three connected collinear atoms all torsions including these atoms are added by traversing the molecular chain until a well-defined dihedral is found. Since a dihedral angle is usually defined for the range  $-\pi \leq \tau \leq \pi$ , whereas the inverse of the cosine only returns values from zero to  $\pi$ , the phase of the angle is determined by comparison. The sign of the angle is set by the scalar product of the first bond with the cross product of the other two. This phase has to be kept consistent for geometry differences so that a change in angle from  $\tau = -178^\circ$  to  $\tau = 178^\circ$  does not result in  $\Delta\tau = 356^\circ$ , but  $4^\circ$ . As such the state of the dihedral is saved at the beginning of every back-transformation and a value is assigned to indicate whether the angle is within the quadrant  $-90^\circ \geq \tau \geq -180^\circ$ ,  $90^\circ \leq \tau \leq 180^\circ$  or in between. The angle is then shifted so that the right difference is acquired by subtraction and therefore the actual value may be larger than  $\pm\pi$ . The same approach is also employed in the Psi4 package.<sup>[105]</sup>

Due to the dependence of the PICs on the definition of sensible bonds, angles and dihedrals the actual coordinate set used depends on the geometry and thus may change during the optimisation. To allow the use of an appropriate set of primitive coordinates in every step an option to recalculate the PICs after a number of steps was also included.

The further steps in an optimisation scheme with redundant internal coordinates require the calculation of the first derivatives regarding the cartesian coordinates and the derived matrices, that is the (pseudo-)inverse and projector, as detailed in section 2.4. Energy, gradient and Hessian are similar to an optimisation in cartesians, although parameterised second derivatives are often evaluated directly in the primitive internal coordinate space. Multiple approximate Hessian guesses have therefore been

implemented in the software. As the model Hessians published in the literature are often used in a modified definition in several software suites, the only comparable Hessian is the scaled diagonal one. It includes values of  $0.5 E_h/\text{Bohr}^2$  for stretches,  $0.2 E_h/\text{rad}^2$  for bends and  $0.1 E_h/\text{rad}^2$  for torsional primitives. Since cartesian coordinates are used to represent linear bends, a value of  $1.0 E_h/\text{Bohr}^2$  is used for those instead. For a more elaborate Hessian Almlöf's<sup>[87]</sup> and Lindh's<sup>[88]</sup> approximation have been added to the library of second derivative guesses. However, the former is determined in a separate PIC set as detailed above, converted to cartesian space and transformed to the internal coordinate set of the optimisation. In contrast to the original formulation by Fischer and Almlöf, StOpt does not include out-of-plane bending coordinates. The implementation of Lindh's model Hessian follows the approach in<sup>[85]</sup>, so that the force constants are determined on the primitive internal coordinates used in the optimisation procedure and consequently no additional transformation is necessary. Since only minimisation algorithms have been implemented, the main updating method is BFGS and can be applied for any number of steps. The model Hessian is then calculated for the geometry  $n_{\text{steps}}\%n_{\text{upd}}$  steps before the current one, where  $n_{\text{steps}}$  is the number of optimisation steps,  $n_{\text{upd}}$  the number of steps used for the update and % the modulo operator.

With the projected gradient and Hessian the new step is then predicted by the rational function approach and scaled if the length of the step exceeds 0.3. The value is not checked in the cartesian space, but in the coordinate system used. As it is more appropriate to determine the step under the constraint of a certain step length the restricted-step RF (RS-RF) method was implemented as well, which requires an iterative solution of the rational function ansatz. In the last steps of the optimisation cycle a new geometry in cartesians has to be determined which requires the iterative back-transformation from curvilinear to rectilinear coordinates. The standard procedure is used to accomplish this without the recalculation of the B-matrix in every step to save matrix operations. As noted above the phases of the dihedral angles are fixed at the beginning of the back-transformation so that a smooth convergence is achieved.

Optimisation convergence is checked at the end of the loop and signalled if the maximum gradient and either the cartesian step or the change in energy is below its threshold. The approach was designed to reproduce the one used in Molpro and thus the gradient used in the convergence check is taken from the last but one structure, although energy and gradient calculations are done in one single point calculation.

Designing StOpt with parameters similar to those used in Molpro ensures a consistent picture and allows the comparison of their optimisation efficiency. Since an exact correspondence is almost impossible not much insight is gained by comparing similar methods of different software, but the best or standard methods of each should be compared. Deficiencies in the developed procedure are better unearthed by benchmarking several settings and approximations of the optimiser.

## 3.2 The Best Combination of Approximations

Several benchmark sets can be envisioned to test an optimisation scheme. Ideally those sets should contain a balanced number of many different binding motifs so that the robustness and general applicability of the scheme can be ensured. However, a comprehensive assembly of molecular geometries that takes into account every possible situation grows quickly too large for routine tests. A focus of this study is the development and assessment of a redundant-internal-coordinate-based optimisation

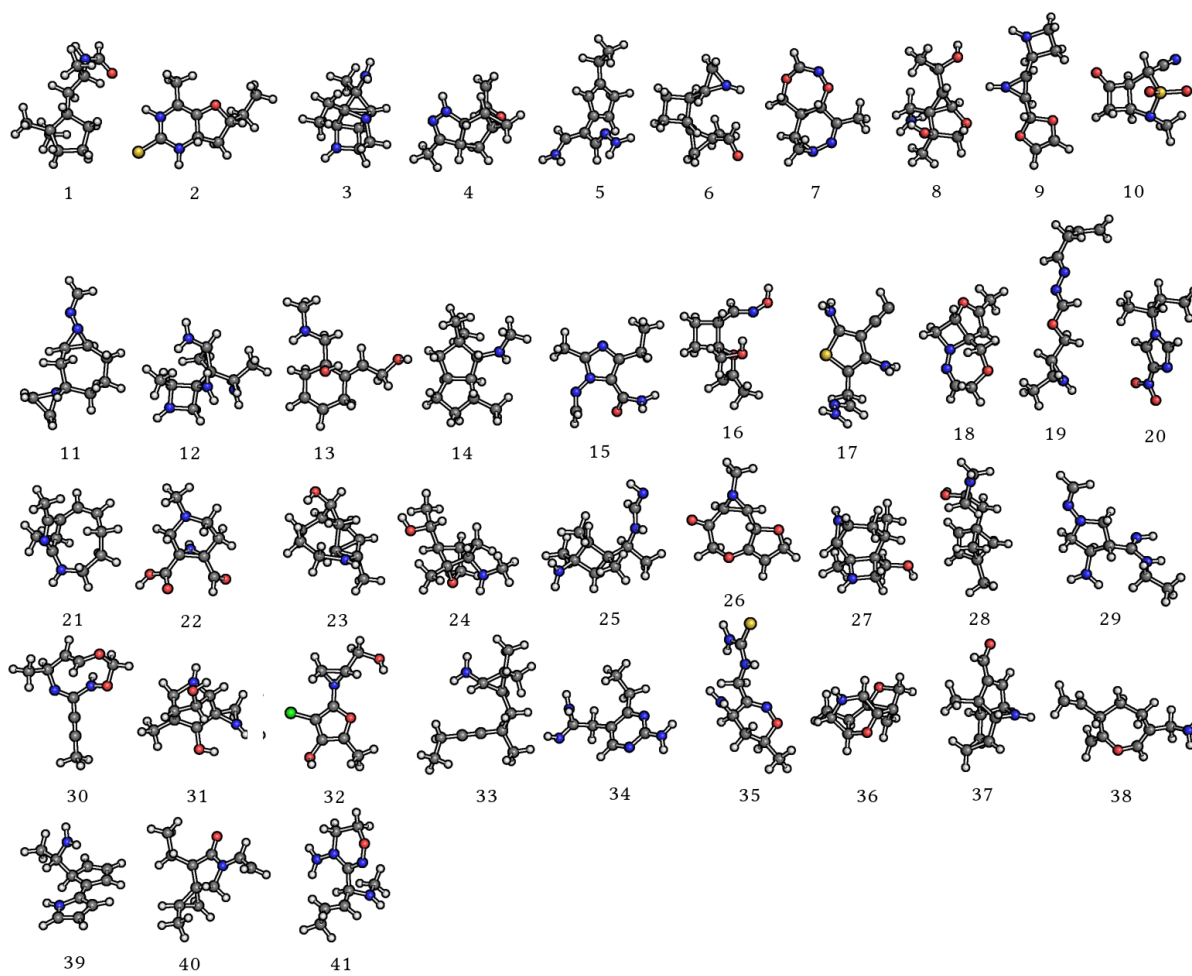


Figure 3.2: All of the starting structures of OrgSet1 as used in the optimisation tests shown in the following plots.

procedure, which depends heavily on the quality of the primitive coordinates. Therefore, a set of small and diverse geometries that exhibits difficult cases for these coordinates is a more appropriate choice for a benchmark set.

As such a set of 41 structures was obtained from the GDB-13 database<sup>[106]</sup> for small organic druglike molecules which are shown in Figure 3.2. They were preoptimised with the simple DREIDING force field,<sup>[107]</sup> leading to rather crude starting structures. In one case (structure 23) the geometry even exhibits two seemingly interlocked ring structures which is unfavourable and almost impossible to optimise as a closed-shell molecule. Nevertheless, it is retained in the set as a measure of robustness for the optimisation procedure. Others contain one or several rings and linear domains which can also be problematic for internal coordinates and thus serve to unravel problems in the PIC set.

This benchmark set was therefore optimised with several different options in order to find handles for an improvement of the optimiser and assert the best combination of approximations. All of the single point calculations used Molpro<sup>[102]</sup> in the version 2012.1 with the DFT functional PBE<sup>[108]</sup> and the def2-SVP<sup>[109]</sup> basis set. Density fitting was used in combination with the automatically assigned auxiliary basis set def2-SVP/JFIT.<sup>[110]</sup> For converged energies and gradients the grid size parameter was set to achieve an energy accuracy of  $1.0 \times 10^{-9} E_h$  and the threshold for energy convergence to  $1.0 \times 10^{-8} E_h$ . All molecules were taken to be neutral and of singlet multiplicity which allowed the use

of the restricted Kohn-Sham method.

The comparison of approximations on OrgSet1 is mostly presented in bar graphs which show the number of iterations needed for convergence, with a maximum of 100. A converged geometry may not correspond to an actual minimum and thus all the stationary points were checked with a successive frequency calculation on the same level of theory. Due to numerical inaccuracies in the optimisation and frequency run, only results with imaginary frequencies larger than  $25\text{ cm}^{-1}$  were assumed to not be sufficiently near to a minimum. If an optimisation converged to such a point it is marked with an asterisk.

#### 3.2.1 Step Prediction

The first task in setting up a numerical optimisation procedure is choosing the right step prediction algorithm. Many possibilities exist to minimise an arbitrary function, but if second derivatives are available the (Quasi-)Newton (QN) types have been shown to give the best results.<sup>[3]</sup> Consequently the QN methods have already been investigated up to the point at which further improvement is likely to be small.<sup>[4]</sup> One of the most popular choices is the rational function algorithm, which is known to give reasonable steps and have competitive convergence properties. The level-shifted Quasi-Newton method behaves similarly in the limit of the exact Hessian, but is inferior for more crude approximations.<sup>[85]</sup> As such a further benchmark of step prediction methods is unlikely to help in improving the efficiency of the optimisations.

On the contrary, the restriction of the step has gained much less attention and thus a possible improvement could be made on that end. In most optimisation routines the predicted step length is simply scaled in the coordinate system of the step prediction if it exceeds the trust radius. As this generally does not correspond to the actual solution of the RF equations, the direction of the true step may be different than the scaled one. To solve this problem an iterative algorithm is needed, as suggested in the literature.<sup>[111,112]</sup> The *restricted step rational function optimisation*<sup>[112]</sup> (RS-RF) is such an approach and was thus tested on OrgSet1 against the standard RF algorithm with a simple maximum step size.

Figure 3.3 shows the number of iterations for each structure obtained with Lindh's model Hessian determined in the PIC set and either with RF or RS-RF as the step predicting method. The second derivative model was recalculated at the geometry five steps before or at the starting structure if less steps were taken. It was then sequentially updated by BFGS using the five last steps. In both runs the step restriction was set to a length of 0.3 in the coordinate system of choice and thus the only difference should result from the enforcement of the step length.

In the case of a diagonal Hessian, as shown in the appendix, the restricted step variant of the rational function method seems inferior to a simple scaling of the step length. However, from Figure 3.3 it becomes evident that this deficiency decreases with a better approximation to the exact Hessian as Lindh's model reduces the number of iterations largely, except for structures 17 and 33. Even those are much improved by an increase of BFGS update steps as deduced from the graphs in the appendix. This additionally hints at the importance of a good parameterisation of the second derivatives for an efficient optimisation using RS-RF. Notwithstanding, the character of the stationary points is also affected by the Hessian approximation. In the case of the best performance of RS-RF the difference in efficiency and non-minima points in comparison to RF is found to be almost insignificant. On the other side of the spectrum, that is by using a diagonal approximation to the Hessian, the number of higher-order



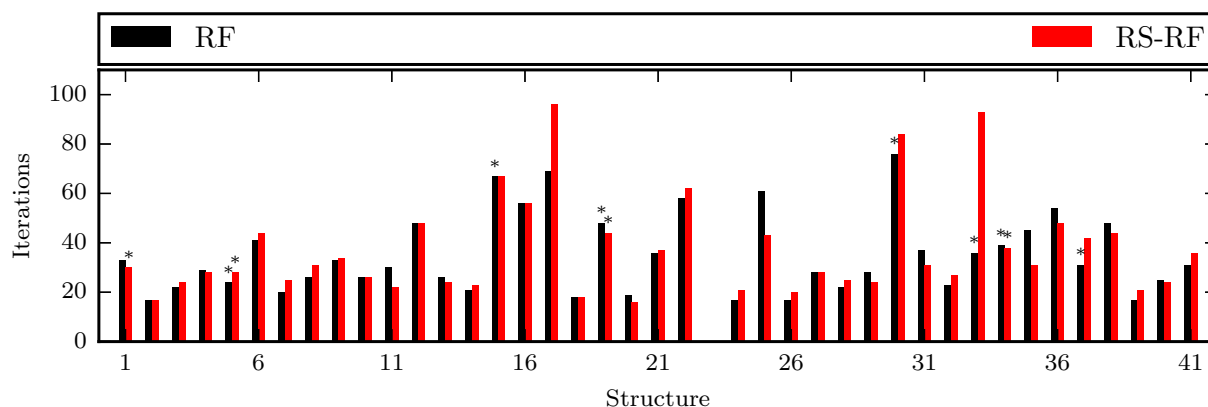


Figure 3.3: Number of iterations needed in for each structure in OrgSet1 using rational function (black) or the restricted-step variety (red) for the step prediction. Lindh's model Hessian with BFGS updates of five previous steps was employed with a maximum step length of 0.3 in the coordinate system used for the prediction.

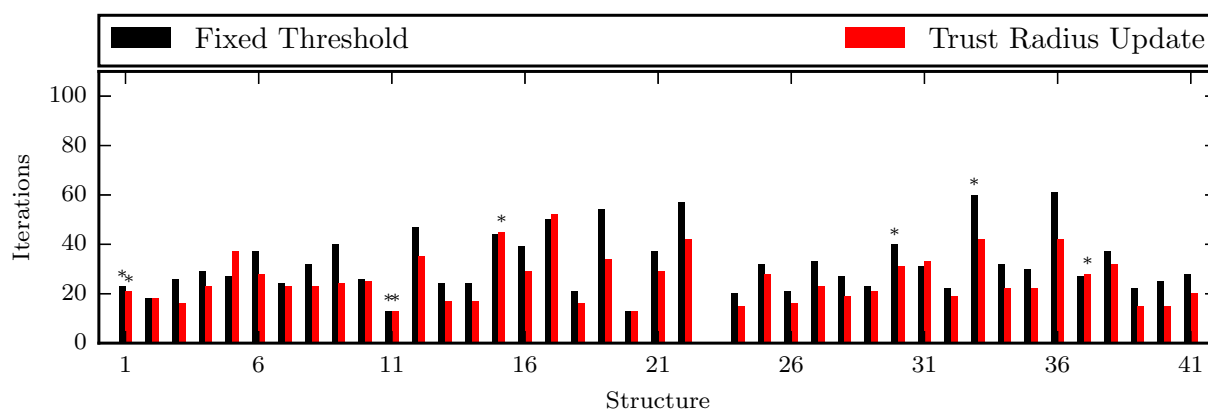


Figure 3.4: Convergence of OrgSet1 with (red) and without (black) an update of the trust radius. The initial value was 0.3 in the internal coordinate set in both cases. A rational function algorithm with Lindh's model Hessian as initial guess and successive BFGS updates were otherwise employed.

stationary points is much lower for RS-RF than for RF. As such it seems that RS-RF is in general a more robust method than simply scaling the RF step, which is likely even more pronounced in transition state optimisations. Only if a good analytical approximation cannot be carried out the simple ansatz is going to be more economic.

Independent of the approach used for the step restriction, the magnitude of the trust radius itself has an obvious influence on the number of iterations. A fixed step size generally leads to sufficient performance although a strong dependence on the actual value is often reported. On the other hand the optimum trust radius has been determined to be between 0.3 and 0.4, seemingly independent of the particular internal coordinate system used.<sup>[85,89]</sup> The latter is rather surprising since largely different coordinate systems should lead to a variety of step sizes and directions, which is expressed by the improved performance of internal over cartesian coordinates. However, this finding is in line with the competitive convergence of cartesians with good Hessian approximations, although this competitiveness likely diminishes for more difficult potential energy surfaces. The usefulness of scaling the length dynamically is therefore also dependent on the structure and thus the local potential. Figure 3.4 clearly shows that despite the relatively small difference, a trust radius update is generally advantageous. Most optimisations are significantly accelerated by allowing a larger step size which is especially pronounced

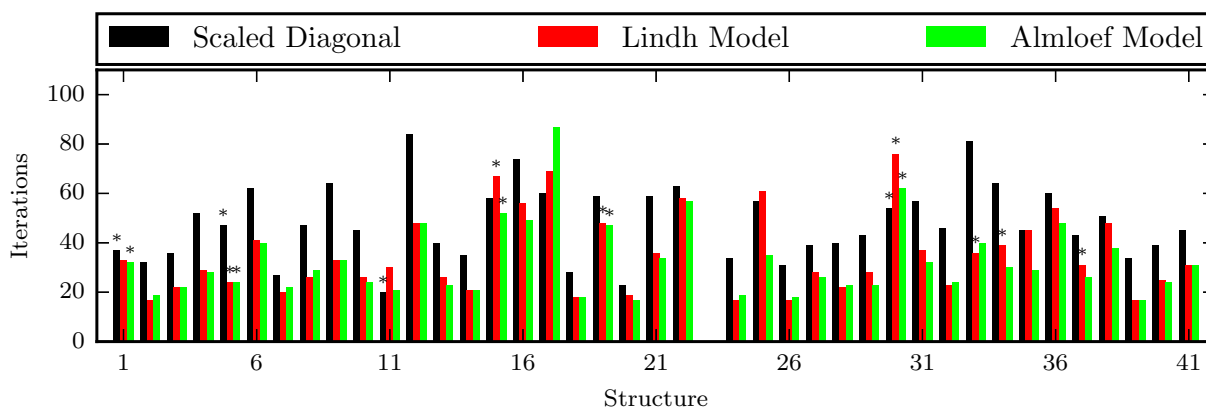


Figure 3.5: Number of iterations needed for convergence due to different Hessian approximations. The RF method was used with the standard redundant internal coordinate set and a step length scaling if it exceeded 0.3. A diagonal (black) approximation, Lindh’s model (black) and Almlöf’s model (green) were used. BFGS was used to update Diag at every point, whereas only the five most recent steps were used for the other model Hessians.

in the first steps that exhibit large gradients. Although the use of an update of the trust radius can thus be recommended for routine applications it has to be noted that it might lead to worse convergence for more difficult potential energy surfaces, as the update only takes the last two steps into account.

As expected from the modifications investigated in this section the reduction of the iterations needed for convergence is very small for this benchmark set. Nevertheless the correct prediction and scaling of the step results in an overall performance gain of the optimiser. More importantly, the desired stationary point is found more reliably than with standard procedures that simply scale the rational function step to a certain cutoff. Nevertheless, the direction and length of the step is closely linked to the second derivatives used in the prediction of the step, which warrants a separate investigation.

### 3.2.2 The Approximation of the Hessian

The improved performance of internal coordinates is a result of the approximate decoupling of the coordinate space. As a consequence optimisation in an internal framework is said to already be efficient with a simple diagonal Hessian. A common practice is thus to scale each type of coordinate in an identity matrix with a different factor, which also allows to artificially prefer the optimisation of specific coordinates. With a better description of the true Hessian even strongly coupled bases like cartesians show competitive convergence for many systems. The question thus arises whether internal coordinates exert the same improvement of convergence due to more sophisticated approximations.

There are many ways to obtain approximate second derivatives, however. For all except the smallest systems an exact Hessian matrix is prohibitively expensive in the course of an optimisation. In these cases using an update formula is necessary to achieve a decent efficiency. To also circumvent the exact Hessian calculation at the beginning, many different model Hessians have been suggested that are sufficiently accurate for the relaxation of molecules. Therefore the most commonly used Hessian approximations in optimisation schemes for molecules were tested: a scaled diagonal Hessian (Diag), Lindh’s model (Lindh) and Almlöf’s model (Almlöf).

The results in Figure 3.5 clearly show an improvement of the molecular mechanics Hessians over the diagonal approximation. In general the convergence with the diagonal Hessian is very smooth in contrast to model-based versions, which is a direct result of the smaller step size inflicted by the diagonal

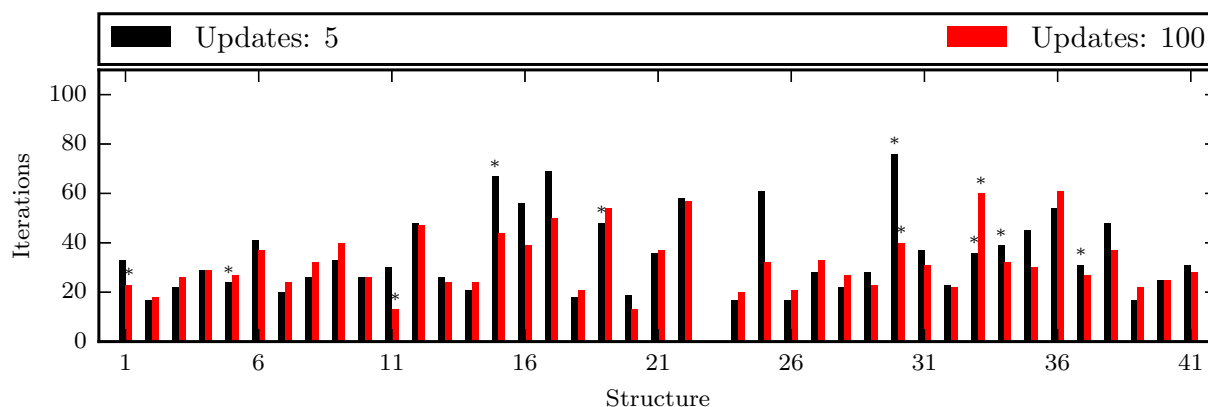


Figure 3.6: Effect of the change in the number of steps used for the Hessian update. RF was employed for the step prediction, which was scaled back to a length of 0.3 if that threshold was exceeded. The standard redundant internal set served as the basis for optimisation and evaluation of Lindh’s model Hessian which was updated with either the five (black) most recent or all (red) steps.

approximation. This is not surprising as using an unscaled diagonal, that is the identity, matrix reduces the related Quasi-Newton method to a gradient descent. One could assume that due to the approximate character of model Hessians, a step might become unreasonably large or lead to an unfavourable direction. That is certainly often the case, but at the start of an optimisation the geometry is sufficiently distorted that even a step into a slightly different direction leads to an overall reduction in potential energy, whereas it may hinder convergence near the minimum. Generally the overall reduction in the number of cycles more than compensates for the latter effect.

Nevertheless, plain use of any model would lead to bad optimisation behaviour because the curvature of the true PES may deviate strongly from the approximate derivatives. To compensate for this effect update methods need to be applied for a number of steps. Often the model Hessian is just used as a guess at the beginning of the optimisation run and updated successively, but it was also found that using a limited amount of update steps can enhance convergence further.<sup>[89]</sup> Figure 3.6 thus compares the influence of using the Hessian update with a different number of steps, which should decrease the influence of the model with increasing number of updates.

The difference in the results is most pronounced for structures with already slow convergence behaviour, as expressed by structures 15–17, 25, 30 and 33. For the average converging structures like 1–10 the difference is small, but favouring the smaller number of steps used for the update as suggested in the literature. However, for those difficult minimisations the use of a model Hessian only at the beginning reduces the number of needed iterations significantly and may compensate for the additional few iterations lost in the other cases. Additionally the large number of updates seems to reduce the convergence to saddle points as using the full step history only resulted in four geometries that do not qualify for a minimum, whereas the restriction to five steps leads to seven non-minimum structures.

As such the use of model Hessians as an initial guess with successive BFGS corrections in every step not only saves CPU time by avoiding expensive recalculation, but also improves the overall efficiency of the optimiser. From the analysis in this section it is not clear which model may be the best in average. Both Lindh’s and Almlöf’s approximations showed similar performance. With a different benchmark set that includes a wider range of interactions the discrepancy may be larger. However, both model Hessians were used differently from the definition in the original papers so that with the results obtained a similar behaviour can at least be expected in their current implementation. On the

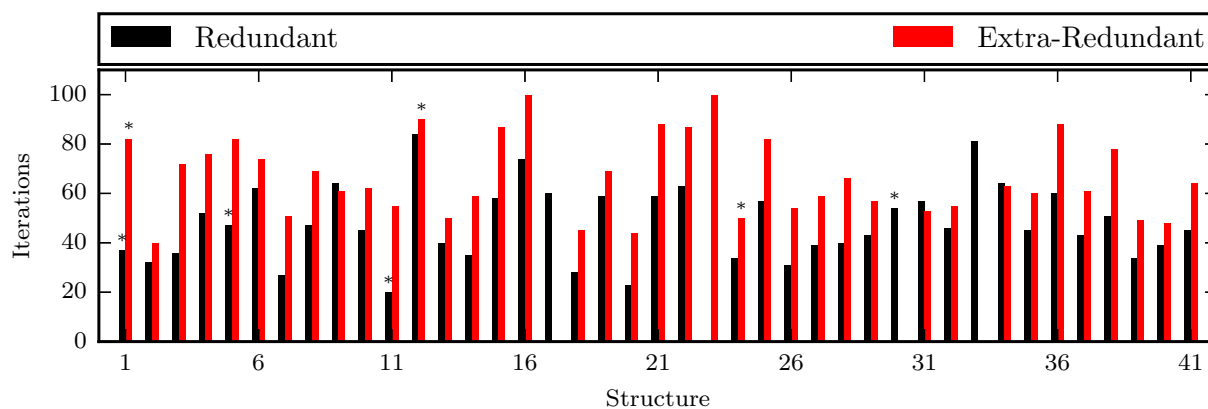


Figure 3.7: Convergence dependence on the primitive internal coordinate set if a scaled diagonal Hessian is used as a guess at the beginning and updated by BFGS in successive steps, determined by the RF method with a maximum step size of 0.3. The standard primitive set (black) is compared against an extended set with a number of auxiliary bonds (red).

other hand both are based on the PIC set as defined internally, so that the coordinates could have a significant influence, which is therefore investigated in further detail.

### 3.2.3 The Coordinate Set

The internal coordinate set has been found to enhance the convergence of molecular geometry optimisations due to the reduction of couplings between movements that occur throughout the search for a stationary point. With a careful combination of these primitive coordinates both harmonic and anharmonic couplings can be significantly reduced.<sup>[84,113]</sup> In exchange for the acceleration of convergence compared to cartesian-based spaces and the ease of construction in contrast to natural internals, a redundant set of coordinates becomes sensitive to the setup and handling of the primitives. The redundancies are usually removed to first order by the projector  $\mathbf{P}$  as in equation (2.50), which ensures a coordinate step that is actually representable in the set of primitives.<sup>[95,113]</sup> However, this cannot redeem the dependence on the actual definition and setup routine. As is the case with the choice of a particular z-matrix a certain set of primitives may be a bad representation of a molecule's equilibrium geometry and thus a very small, fixed coordinate set might consequently lead to oscillations and even prevent convergence. On the other hand plainly adding all possible stretches, bends and torsions would result in a highly redundant basis which mostly slows down the convergence by reducing the step size. As such the construction of the internal basis for optimisation has to yield a well-balanced set of coordinates to be efficient.

Bakken and Helgaker<sup>[85]</sup> suggested an *extra-redundant* coordinate set in addition to the standard definition, that consisted of auxiliary bonds which mostly connect next-neighbour atoms. The difference in convergence due to the additional redundancy can be seen in Figure 3.7 and Figure 3.8. From these results it is evident that the higher redundancy does not improve the overall convergence in the case of a diagonal starting guess, which supports the notion of generally smaller step sizes for a larger internal coordinate set. With a better approximation for the Hessian the picture changes somewhat as seen in Figure 3.8, but not sufficiently to prefer the inclusion of the auxiliary bonds by default. An advantage due to a model Hessian guess was already reported<sup>[85]</sup> and attributed to a better description of the approximate second derivatives by the larger coordinate set in which the Hessian guess was built. Here

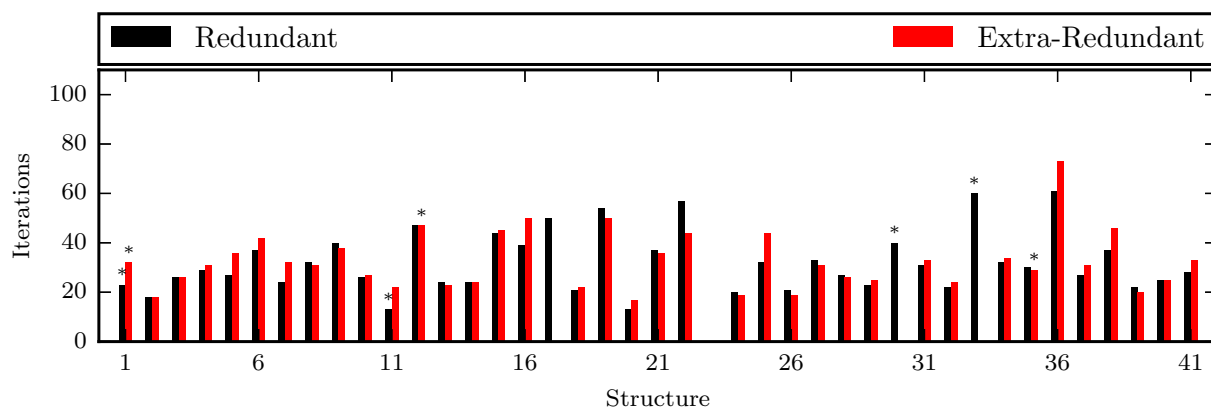


Figure 3.8: Influence of different primitive internal coordinate systems on the convergence with RF, a step scaling to 0.3 and Lindh’s model Hessian used as a guess at the beginning with successive BFGS updates. The black bars correspond to optimisations with the standard primitive internal set, whereas the red ones were determined with the extra-redundant coordinates.

the extended PIC set does not lead to a significant improvement of the average convergence though. On the contrary, some structures fail to optimise in the extended coordinates, specifically 17, 30 and 33 which all exhibit linear carbon chains in the initial geometry. All three optimisations stop because of a diverging back-transformation which was not further investigated. Still, as bonds are added in between nearest neighbours, chain-like structures lead to a highly redundant set of bonds at any point on the PES. It is thus possible that the extra-redundant coordinates are unsuitable to optimise such a molecule.

As shown above, even when projecting out the first-order redundancies, the separate primitive coordinate sets can behave very differently in traversing the potential energy surface. This is not only expressed in the efficiency of the optimiser, but also directly in the iterative back-transformation after the step-prediction. An inflexible or rigid coordinate set would make it difficult to represent the predicted step in cartesian space so that a large internal residual  $\Delta r = |\mathbf{q}' - \tilde{\mathbf{q}}'|$  remains.<sup>[113]</sup> Consequently a metric for the rigidity of an internal coordinate set can be defined as

$$\chi_r = \Delta r / |\mathbf{s}|, \quad (3.1)$$

where  $\mathbf{s}$  is the predicted step,  $\mathbf{q}'$  the corresponding predicted and  $\tilde{\mathbf{q}}'$  the actual internal coordinates after convergence of the back-transformation. This measure could consequently be used to find improved internal coordinate sets and can lead to insight about its usefulness for the current optimisation.

As such, structure 36 was investigated, since it exhibits one of the slowest convergence and a pronounced discrepancy between the standard and extra-redundant coordinates. To compare with a well-behaved optimisation structure 3 serves as a reference. Looking at the graphs in Figure 3.9 one clearly notices the similarity in the rigidity behaviour for both coordinate sets if the structure is well described by the primitive coordinates. The contrary can be said about the minimisation of structure 36, which exhibits a slower convergence and reduction in rigidity for the extra-redundant than the standard PIC set. Interestingly the maximum  $\chi_r$  is higher and the reduction more erratic for the standard set which is similar to the change in energy. Comparing the behaviour of the coordinate rigidity during the optimisation of structure 3 with that of structure 36, one may be led to assume that the difference is due to an unsatisfactory choice of coordinates. Selected molecular geometries during the minimisation of

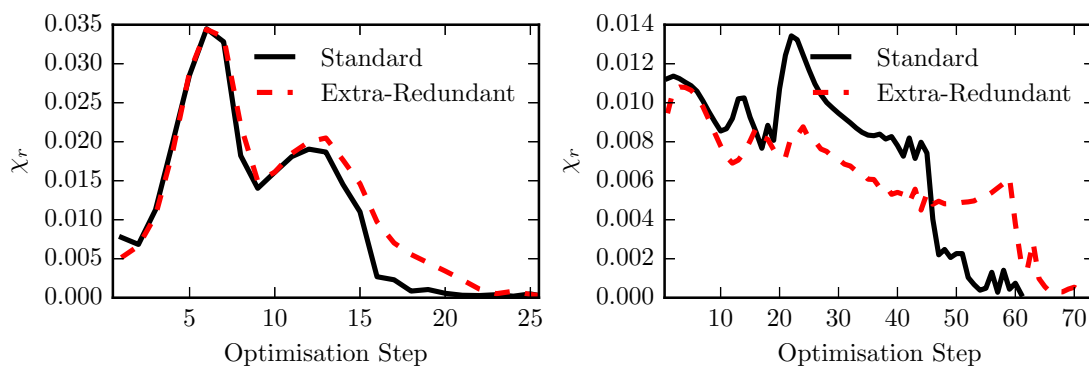


Figure 3.9: Rigidity of the standard (black) or the extra-redundant (red, dashed) PIC set throughout the optimisation of structures 3 (left) and 36 (right). The RF method was used to predict a step that was scaled to 0.3 if it exceeded that value. Lindh's model in the used PIC set provided the initial Hessian guess which was successively updated by the BFGS method.

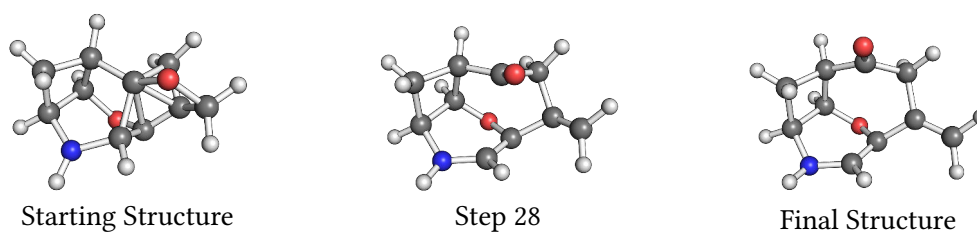


Figure 3.10: Geometries of structure 36 as occurring during the optimisation using the standard set of primitive internals, rational function optimisation, Lindh's model Hessian and BFGS updates in all successive steps.

structure 36 are depicted in Figure 3.10 and clearly corroborate this assumption.

The rigidity does not give clear guidance on how to improve the coordinate set for the optimisation and thus a specific procedure would have to be developed. A simple first approach to this problem can be achieved by recalculating the primitive internal coordinates at several steps during the optimisation. Therefore structure 36 was optimised with the standard coordinates and a reset at several optimisation steps, resulting in the convergence plot Figure 3.11. From the graphs it is evident that an appropriate coordinate set improves the convergence considerably. The recalculation of the coordinates is not necessary in every step as the influence near the minimum is small and is mostly not changed. As a globally usable, but still efficient collection of primitives is impossible to create for all structures of interest, the recalculation seems a necessary improvement for a robust optimiser.

Although some optimisations of molecular structures could possibly be improved by adding other primitive coordinates as out-of-plane bends,<sup>[114,115]</sup> and special intermolecular bonds,<sup>[116]</sup> the currently used standard combination of stretching, bending, linear bending and twisting coordinates exhibits satisfactory convergence properties. The extra-redundant set may be advantageous for structures in which some degrees of freedom are insufficiently characterised by the standard set, but is not recommended for general optimisations. However, due to the change in geometry of crude starting structures the recalculation of the internal coordinate set is advisable.

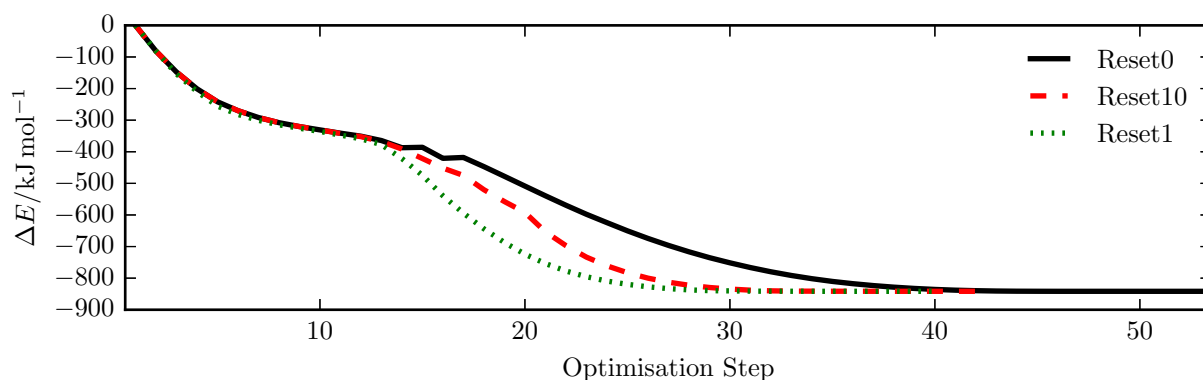


Figure 3.11: Optimisation of structure 36 with and without recalculation of the standard primitive internal coordinate set. Reset0: no recalculation of the PICs, Reset10: recalculation after every ten iterations, Reset1: recalculation at every step. The optimisation otherwise employed RF for the step prediction, Lindh’s model for the initial Hessian and BFGS for successive updates.

Table 3.1: Algorithms used in the optimisation methods to compare Molpro’s internal routine against methods available in StOpt.

Method	Step Prediction	Step Length	Coordinates	Model Hessian	Update	Update Steps
Molpro1	RF	0.3	LNC	Lindh	BFGS	5
Molpro2	RF	0.3	NIC	Diag	BFGS	5
StOpt1	RF	0.3	RIC	Diag	BFGS	all
StOpt2	RF	0.3	RIC	Lindh	BFGS	all

### 3.3 Comparison to Other Software

Without a common basis for optimisation tests the comparison of literature results would not be meaningful. Baker introduced a set of 30 small organic molecules, shown in Figure 3.12, near to a stationary point to benchmark the efficiency of optimisation schemes.<sup>[84]</sup> Owing to the computational resources at that time the method of choice was restricted Hartree-Fock theory with the STO-3G<sup>[117]</sup> basis set, as only single and uncharged closed-shell molecules were calculated. Due to several benchmarks against Baker’s results this combination grew to be a de-facto standard with only few additional investigations of the basis set influence.<sup>[85]</sup> A less often used set of molecular structures, was introduced as a test for delocalised internal coordinates.<sup>[101]</sup> Some of these structures, shown in Figure 3.13, were chosen as difficult cases for optimisation procedures in any coordinate space. Cubane and the other molecules with multiple annulated rings can be problematic for the setup of internal coordinates, whereas the string of three-membered rings in jawsamycin makes it difficult to converge, especially in cartesian coordinates.<sup>[101]</sup> This additional collection of molecules is thus a good measure for the general applicability of the implemented internals and the performance of the optimisation scheme.

To appropriately set the efficiency of StOpt into perspective Baker30 was optimised with two promising procedures in StOpt and two of Molpro’s own methods, which is summarised in Table 3.1. The first scheme, StOpt1, is based on the standard primitive internal coordinate set as detailed in section 3.1, the rational function algorithm with a simple scaling of the step length to 0.3 and the scaled diagonal Hessian as an initial guess that is successively updated using BFGS. In internal coordinates diagonal Hessians are said to be competitive with the optimisation in cartesians and better approximations to the second derivatives, which is thus tested with StOpt1. StOpt2 expresses the performance of a gen-

### 3 Development of an Improved Optimisation Procedure

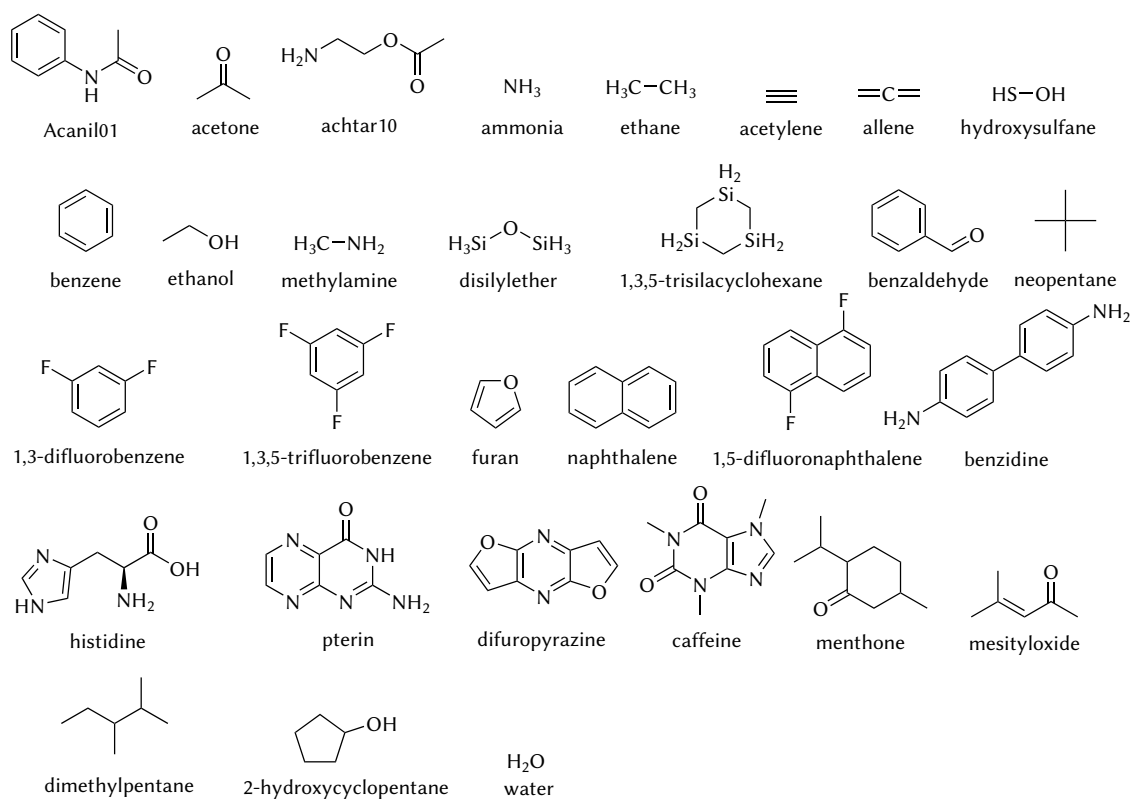


Figure 3.12: Baker's test set (Baker30) of 30 molecules used to compare the efficiency of natural internal and cartesian coordinates in a general optimisation procedure.

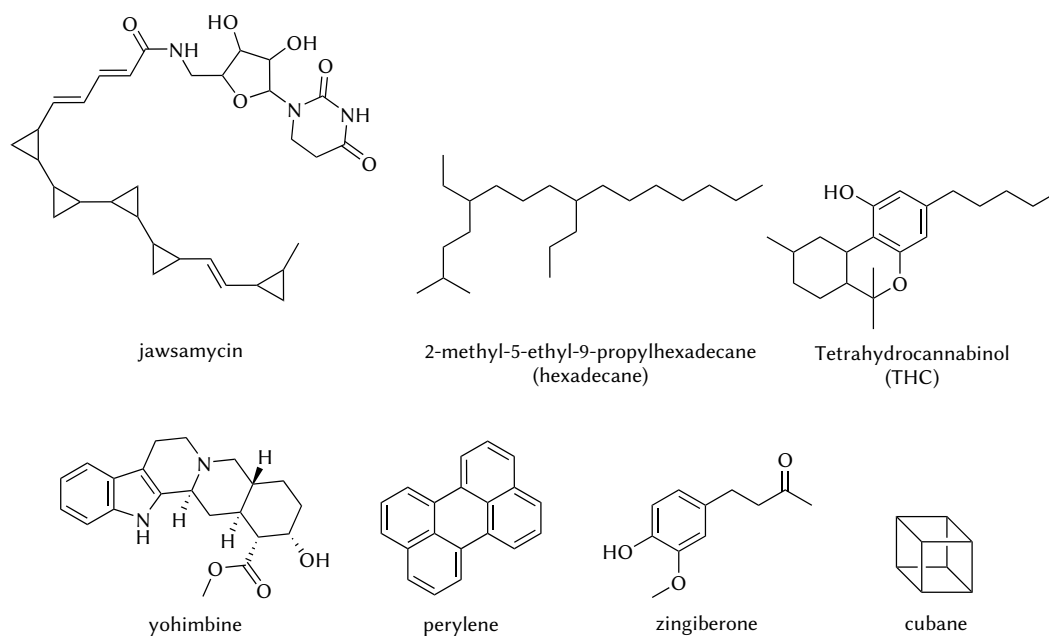


Figure 3.13: The seven molecules used to benchmark StOpt's minimisation performance against other optimisation routines. The collection is inspired by Baker's test set for delocalised internal coordinates, as fused ring structures and rare structural motifs are problematic cases for internal coordinates. The names in parantheses are abbreviations that are used alternatively.



erally applicable optimisation procedure in StOpt with an improved convergence due to the use of a molecular mechanics Hessian. As such Lindh's model was used as a guess for the second derivatives and successively updated at every step.

Molpro1 is the standard optimisation approach in Molpro 2012.1, which employed the rational function step prediction with a simple scaling of the step length to 0.3 and local normal coordinates determined from Lindh's Hessian model which is updated with five previous steps. Molpro2 resembles the best performing procedure as it is based on natural internal coordinates, but used a diagonal matrix for the second derivatives. The latter is more robust with the natural internals than molecular mechanics Hessians as used in Molpro1, which otherwise utilises the same approximations.

As for the literature values, results obtained by Baker<sup>[84]</sup> employed natural internal coordinates, the EF method for step prediction and a molecular mechanics Hessian guess with BFGS update. The EF algorithm is similar to the rational function optimisation.<sup>[85]</sup> Values in the column for Eckert<sup>[89]</sup> were determined with a similar procedure as Molpro1. Bakken<sup>[85]</sup> represents the reference method for the approach developed in this study, as a similar PIC set is used. However, the listed results were acquired in the extra-redundant internal coordinate set. Furthermore the iterations in the columns Bakken and StOpt2 were obtained with Lindh's model Hessian in the space of the primitive coordinates, contrary to the other approaches. The initial guess was then updated with BFGS in every new step like in StOpt1. All of the calculations used similar convergence criteria as defined by Baker.<sup>[84]</sup>

The average of needed iterations in the last row of Table 3.2 show a striking variation in the efficiency of optimisation schemes. Baker retrieved the worst results of the literature values even though he used natural internal coordinates. The particular definition of those coordinates may therefore be behind the inferior convergence. Eckert obtained very good performance using simple local normal coordinates, which could not be reproduced with Molpro1 despite the assumed similarity of the optimisation procedures. Bakken is also very efficient and was expected to be comparable to StOpt2. Notwithstanding, there is still a significant difference in the results between procedures in StOpt and comparable literature methods.

This discrepancy might be caused by a difference in the updating procedure or even just the step restriction, since a trust radius update was not employed in StOpt1 or StOpt2. Additionally subtle differences in the internal coordinate set can have a visible effect even in molecules as small as those in Baker30. The latter may be responsible for the difference between Bakken and StOpt2 as a slightly different approach was used in both cases. Irrespective of this even model Hessians like Lindh's are used or constructed differently in separate codes, which can lead to deviations in the behaviour of the minimisation. As Bakken employed the extra-redundant coordinate set, not only the flexibility of the coordinates might have been better, but also the initial Hessian guess which was evaluated in the same coordinate space.

Column Molpro2 prominently shows the advantage of natural internals in optimisations as it exhibits the lowest number of iterations for any of the procedures, even with a diagonal Hessian. Due to the simple Hessian approximation this advantage is going to be less for more difficult structures. Additionally the optimisation of acetylene did not even commence, since the linear structure resulted in a program failure. StOpt1 exhibits only fair performance for an optimisation in internal coordinates which can probably be attributed to the simple diagonal Hessian used at the start of the optimisation, as corroborated by the much better results of StOpt2. Additionally StOpt1 requires a similar total number

Table 3.2: Number of geometry optimisation steps required for convergence of molecules in Baker’s set of 30 molecules. Program crashes are marked by a dash. The average iterations until convergence is achieved is listed in the last row. In the case of Molpro2 only 29 structures were used for the average. For details on the methods used see the text.

Molecule	Baker	Eckert	Bakken	Molpro1	Molpro2	StOpt1	StOpt2
water	6	4	4	4	4	5	4
ammonia	6	5	5	5	6	4	4
ethane	5	4	3	4	3	4	4
acetylene	6	5	4	5	-	5	4
allene	5	5	4	3	4	5	4
hydroxysulphane	8	7	7	8	7	9	7
benzene	4	3	3	3	3	3	3
methylamine	6	5	4	4	4	4	5
ethanol	6	5	4	6	5	6	5
acetone	6	5	4	5	5	5	5
disilylether	8	10	8	11	8	9	8
1.3.5-trisilacyclohexane	8	8	9	10	5	12	8
benzaldehyde	6	5	4	5	4	6	4
1.3-difluorobenzene	5	5	4	4	4	4	4
1.3.5-trifluorobenzene	5	4	4	4	4	4	4
neopentane	5	4	4	4	4	4	4
furan	8	6	5	5	6	6	6
naphtalene	5	6	5	5	6	5	5
1.5-difluoronaphtalene	6	6	5	5	6	5	5
2-hydroxybicyclopentane	15	9	9	13	10	17	12
achtar10	12	8	8	9	7	12	8
acanil01	8	8	7	8	8	8	7
benzidine	9	8	9	12	6	10	8
pterin	10	9	8	9	9	8	9
difuropyrazine	9	7	6	7	7	6	7
mesityloxide	7	5	5	5	6	7	5
histidine	19	23	16	27	9	23	15
dimethylpentane	12	10	9	11	7	11	10
caffeine	12	7	6	6	7	7	7
menthone	13	13	12	21	11	23	13
Average	8.0	7.0	6.2	7.6	6.0	7.9	6.5

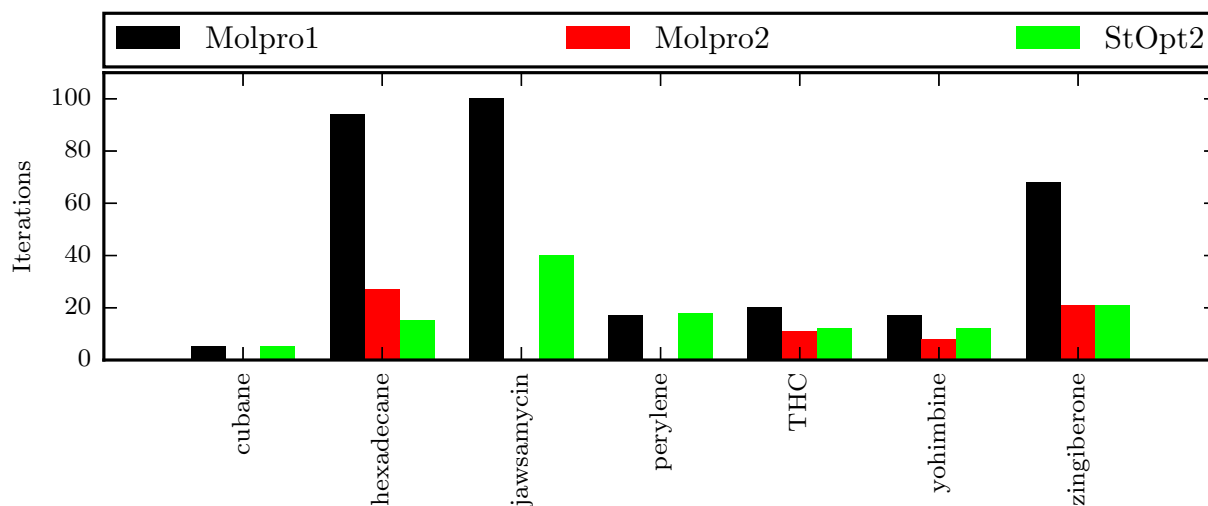


Figure 3.14: Minimisation of the OrgLarge set with the approaches Molpro1 (RF, LNC, Lindh’s model with five steps for the BFGS update), Molpro2 (RF, natural internals, scaled diagonal Hessian with five steps for the BFGS update) and StOpt2 (RF, standard redundant PIC set, Lindh’s model as initial guess with successive BFGS updates). The setup of the natural internals in Molpro2 failed for three of the structures due to the demanding molecular structure.

of iterations as Molpro1 and thus supports the adequate convergence of redundant internal coordinates with simple diagonal approximations, analogous to the natural internals. Although StOpt2 is not the most efficient procedure for this set of structures it is not very far off from the other approaches and thus proves to be a decent approach. Furthermore the advantage of redundant internal coordinate systems may become more pronounced for larger and structurally involved systems.

As such the OrgLarge set can be used to test this hypothesis, as its molecules exhibit a much larger number of variables to optimise and more difficult starting structures. StOpt2, Molpro1 and Molpro2 were thus used to minimise the energy of these structures on the same level of theory as the Baker30 set. Figure 3.14 clearly corroborates the assumption that StOpt2 exhibits the best performance mainly for larger and more difficult systems. The standard optimisation in Molpro 2012.1, Molpro1, converges every structure except jawsamycin, which however is already near the minimum and would only require a number of additional steps. This observation is in line with the one made by Baker,<sup>[101]</sup> who stated that cartesian optimisation procedures converged very slowly since they take very small steps already from the start. For the easier to converge systems like cubane and perylene Molpro1 is very competitive or even a few steps faster. Notwithstanding, those structures exhibit the smallest number of iterations anyway and therefore only contribute little to the overall performance. In the case of more demanding structures like hexadecane and jawsamycin the approach exerts a much slower convergence than Molpro2 or StOpt2. These are almost similarly efficient, although one has to admit the difference in the Hessian employed in both cases. Molpro2 used a diagonal Hessian with five steps in the BFGS update, whereas StOpt2 started with Lindh’s model, which was successively updated at every step. Thus the efficiency of Molpro2 could probably still be improved with a better approximation to the Hessian. Nevertheless, due to the failure of determining the natural internal coordinates, the optimisation did not even start for three of the structures, whereas StOpt2 did not exhibit any problems. The performance of StOpt is thus competitive to other optimisation approaches used in current software suites and seems to be robust enough to be used in routine applications.

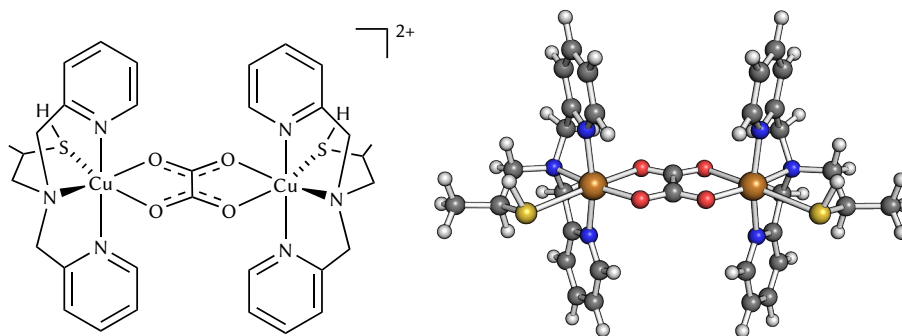


Figure 3.15: The model system for the product of  $\text{CO}_2$  reduction to oxalate with a dinuclear copper based catalyst.<sup>[118]</sup> The original tetranuclear intermediate was capped with hydrogen atoms and cut at the sulfide bond, halving the systems size. Left: Lewis structure of the model system. Right: ball-and-sticks representation of the pre-optimised structure which was used as a starting point.

### 3.4 A Real-World Test

Up to now most benchmarks were done on small organic molecules with the exception of OrgLarge, which unearthed the superiority of redundant internal over local normal coordinates, as used in the standard optimisation routine of the Molpro software suite. In large structures with several different interactions like molecular clusters with hydrogen- or van-der-Waals-bound fragments or even transition metal complexes and proteins, the setup of internal coordinates may become difficult and result in an insufficient convergence in comparison to cartesians. Consequently an assessment of the performance on molecules that are of relevance in current research is necessary.

To redeem this lack of tests, a recently proposed copper-based transition-metal complex, which was found to catalyse the electrochemical reduction of  $\text{CO}_2$  to oxalate,<sup>[118]</sup> was chosen as a final test of performance. In Figure 3.15 the model system used in the optimisation is depicted by its Lewis structure on the left. The actual resting state of the catalyst after the reduction contains the dinuclear complex bridged by two oxalate ligands to another catalyst molecule. To acquire a smaller structure more accessible to state-of-the-art methods the tetranuclear complex was cut at the sulfide bond and capped with hydrogen atoms, thus almost reducing the number of atoms by half. Since the structure cannot be expected to converge within the chosen maximum number of iterations, it was then preoptimised until convergence was signalled. The starting geometry of this optimisation was then used as the reference for the benchmark and is shown in Figure 3.15 on the right-hand side.

Preoptimisation and the final minimisation run were then carried out with Molpro 2012.1 for the electronic structure calculation, whereas the PBE<sup>[108]</sup> functional was chosen in combination with the def2-SVP<sup>[109]</sup> basis set. The accuracy of the energy and first derivatives had to be high for a smooth optimisation with StOpt and thus the integration grid was set to a size for an energy convergence of  $1.0 \times 10^{-8} E_h$  and the SCF thresholds for the energy and density to  $1.0 \times 10^{-7}$  and  $3.16 \times 10^{-8} E_h$ , respectively. Since the model complex is dinuclear with a total charge of +2 the unrestricted Kohn-Sham method was used with the density fitting approach and a triplet multiplicity.

The final optimisation in Molpro again employed the rational function step prediction with a simple scaling of the step, local normal coordinates and Lindh's model Hessian which is updated with a maximum of five steps using BFGS. Minimisation with StOpt was done with the same step prediction, but the extra-redundant coordinate set and a diagonal Hessian at the beginning, which is successively

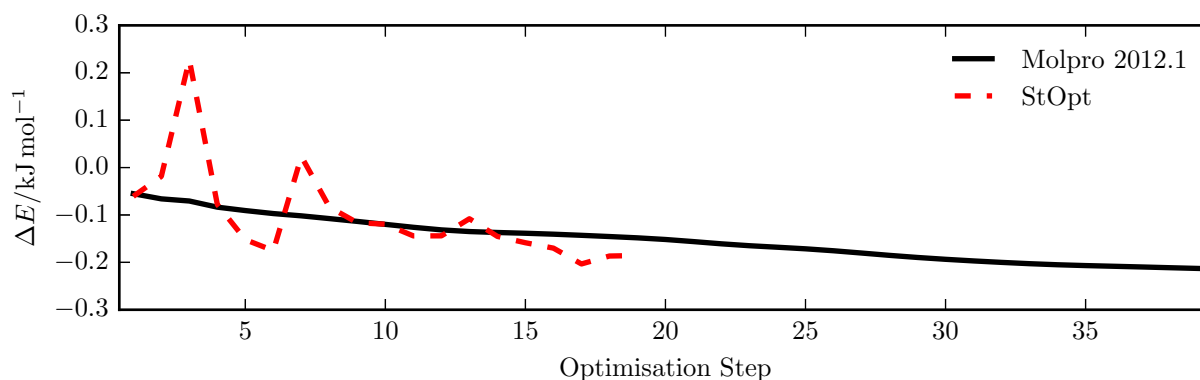


Figure 3.16: Convergence for the last steps of the catalyst’s minimisation using PBE/SVP with either Molpro’s internal or the self-written optimiser (StOpt).

updated with BFGS.

From Figure 3.16 one can see that the structure is already nearly converged at the start of the minimisation run and thus the difference in energy of each step is very small. However, as the thresholds for geometry convergence are not satisfied at that point the comparison is still valid and even more so, because Quasi-Newton methods are known to converge slowly at the end of the optimisation.<sup>[73]</sup> Although GDIIS-based algorithms could accelerate the minimisation at that stage, a faster converging approach with the same step prediction is preferable.

The optimisation with local normal coordinates in Molpro decreases the energy very slowly and smoothly, despite the extensive use of a model Hessian. In contrast StOpt exhibits an oscillative, but decaying behaviour which still converges twice as fast. A reduction of these oscillations may further lessen the number of iterations needed, as the positive steps are likely small and unnecessary. This argument is supported by the failure to converge the optimisation with the standard primitive internal coordinate set in StOpt. Since the extra-redundant coordinates only add bonds to atoms with a greater distance, some important angles or dihedrals may not be described properly, which can lead to the observed oscillations. Utilisation of Lindh’s model Hessian resulted in large steps that were unable to converge in the back-transformation and made a diagonal approximation necessary. As the model was calculated in the PIC space, this additionally supports the necessity to improve the internal coordinates for transition metals compounds. Nevertheless, this graph clearly corroborates the advantage of using redundant internal coordinates in general optimisation work.

### 3.5 Conclusion

In this chapter the development of a molecular geometry minimiser based on redundant internal coordinates was outlined. It was shown that even in the framework of Quasi-Newton methods some parts of the optimisation procedure can still be improved.

The step prediction with Quasi-Newton methods is already a mature field, so that the investigation was restricted to the rational function algorithm. Even with the best method the approximation to the potential energy surface cannot be always accurate and therefore the step is often restricted to a maximum step size to avoid overstepping of the minimum. As the correct approach to restrict the step to a certain size requires the iterative solution of the RF equations, the RS-RF method was tested against the usually applied simple scaling. RS-RF was found to require a larger number of iterations in comparison

to the simpler procedure if used with diagonal starting Hessians, but resulted in a better convergence to the minimum. With an improved approximation to the second derivatives, the difference in the step length treatment became negligible. Dynamically adjusting the maximum step size, that is the trust radius, reduced the number of iterations in the investigated benchmark set significantly.

A variety of model Hessians exist to calculate second derivatives from simple analytical potentials. The suggested efficiency of internal coordinates with scaled diagonal matrices was therefore tested against two model Hessians. Although the usability of the diagonal approximation is evident, more elaborate parameterisations still lead to a significant acceleration of convergence. A pronounced difference between the definition of the model potentials within the current implementation could not be observed. Similarly the Hessian update plays an important role to correct for deficiencies in the approximations. Contrary to suggestions in the literature, the optimisation resulted in more minima of the OrgSet1 when the model Hessian was only calculated at the beginning and successively updated, whereas the number of optimisation steps changed only marginally.

The primitive coordinates play an important role in the performance of a redundant internal set. As such the investigation of OrgSet1 using a PIC set determined by a standard routine and an extended set with additional nearest-neighbour bonds were compared with different Hessian approximations. In all cases the smaller set of coordinates was superior to the extended set, which exposed problems in the case of structures with linear chains. Nevertheless, for some cases the extra-redundant coordinates can be advantageous as a larger internal set generally reduces the step size and thus leads to smoother convergence. This effect was necessary to converge a transition metal catalyst taken from the literature. Even with the extra bonds added the number of iterations was halved in comparison to a local normal coordinate based procedure.

A comparison on a small set of 30 molecules to other optimisation schemes exhibited just mediocre performance of the standard procedure in StOpt. Due to the limited size and bonding patterns of the test set, the reason for this could not be concluded. Notwithstanding, a lot of differences between each of the methods persisted like trust radius updates, varying definitions of model Hessians and changes in the internal coordinates. It was thus suggested that the advantage of redundant internal coordinates can only be fully revealed with a set of larger molecules that exhibit more difficult bonding patterns. Calculations using StOpt and Molpro on the OrgLarge set corroborated this hypothesis. As such there is still room for improvement in StOpt, but the procedure described and tested has been found to be already quite robust and superior to the standard used in Molpro 2012.1.

Several ways to enhance the current implementations could still be envisioned. A major influence on the efficiency is exerted by the PIC set and thus adding or exchanging several coordinates for a better description of other motions like out-of-plane bends could go a long way toward a better optimisation. This argument is in the same line as using natural internals and even a combination of those and redundant internal coordinates may be advantageous.<sup>[119]</sup> The recalculation of the coordinates in every step was shown to improve convergence for some structures, therefore it is obvious that a method to control a change in coordinates will be superior. Furthermore, several step prediction algorithms are known to excel at different parts during the optimisation, suggesting the implementation of hybrid optimisation schemes.<sup>[73]</sup> Similarly, although BFGS is generally the most efficient Hessian update for minimisations, a hybrid scheme for the update methods was suggested and found to marginally speed up convergence.<sup>[4,83]</sup> In summary a flowchart-type optimisation routine that chooses the most appro-

appropriate approach in each phase of the optimisation and takes care to properly scale the step is the most promising direction for further development.





## 4 Molecular Reduction of CO<sub>2</sub>

The reduction of carbon dioxide holds the potential for great developments in renewable energies. CO<sub>2</sub>, used by nature itself to convert sunlight into carbohydrates,<sup>[120–122]</sup> is ubiquitously available. At the same time it is touted as one of the main greenhouse gases produced by humans.<sup>[123]</sup> In light of the impending consequences of global warming there is a great incentive to artificially bind this molecule and also to utilise it in the formation of chemicals for energy storage or other industrial large-scale applications. Fixations of CO<sub>2</sub> at electrodes is possible, but the energy conversion efficiency is relatively low.<sup>[124]</sup> The simple one-electron reduction is also thermodynamically very demanding, with a potential of  $-1.90\text{ V}$  against the standard hydrogen electrode (SHE).<sup>[125,126]</sup> In protic environments a range of more favourable reactions are available with a very narrow potential distribution.<sup>[126]</sup> The

Table 4.1: CO<sub>2</sub> reduction reactions and their corresponding potentials (V vs SCE).<sup>[127]</sup>

Reaction	$E^\circ/\text{V}$
$\text{CO}_2 + 2\text{H}^+ + 2\text{e}^- \longrightarrow \text{HCO}_2\text{H}$	$-0.85$
$\text{CO}_2 + 2\text{H}^+ + 2\text{e}^- \longrightarrow \text{CO} + \text{H}_2\text{O}$	$-0.77$
$\text{CO}_2 + 4\text{H}^+ + 4\text{e}^- \longrightarrow \text{C} + 2\text{H}_2\text{O}$	$-0.44$
$\text{CO}_2 + 4\text{H}^+ + 4\text{e}^- \longrightarrow \text{HCHO} + \text{H}_2\text{O}$	$-0.72$
$\text{CO}_2 + 6\text{H}^+ + 6\text{e}^- \longrightarrow \text{CH}_3\text{OH} + \text{H}_2\text{O}$	$-0.62$
$\text{CO}_2 + 8\text{H}^+ + 8\text{e}^- \longrightarrow \text{CH}_4 + 2\text{H}_2\text{O}$	$-0.48$

selectivity for a particular product is consequently low if no additional measure of control is added to the process.<sup>[5]</sup> Homogeneous catalysts are particularly advantageous for this task due to the often well-defined structure and tunable reactivity. In particular ruthenium-based complexes with  $\text{ReL}(\text{CO})_3$  moieties have been found to catalyse the CO<sub>2</sub> reduction selectively with a very high preference over dihydrogen evolution.

Several groups therefore investigated the  $[\text{Re}(\text{L})(\text{CO})_3\text{Cl}]$  range of catalysts (L = substituted bipyridine) which lead to the suggestion of two different mechanisms that differ in the number of electrons inserted into a single complex (Figure 4.1).<sup>[128–131]</sup>

It was found that the first reduction is ligand centred and the complex bound chloride is successively dissociated. Further reaction can advance in two different directions, depending on the reaction conditions.<sup>[130]</sup> In the one-electron pathway the formed  $[\text{Re}(\text{bpy})(\text{CO})_3]^\cdot$  radical reacts directly with CO<sub>2</sub> to form an intermediate,  $[\text{Re}(\text{bpy})(\text{CO}_3)(\text{CO}_2)]^\cdot$ , which can then undergo reductive disproportionation with another CO<sub>2</sub> molecule to CO and  $\text{CO}_3^{2-}$ . If the radical is long-living enough a dimer may form which reduces the activity, but still catalyses CO<sub>2</sub> reduction. On the other hand, with strongly coordinating solvents the vacant coordination site can be occupied and the radical is stabilised. This enables a further reduction of the complex before reacting with another CO<sub>2</sub> molecule, leading to the two-electron pathway. In contrast to the one-electron path the resulting intermediate transfers the spare oxygen onto an acceptor, for example  $\text{H}^+$  which leads to carbon monoxide and water as products.

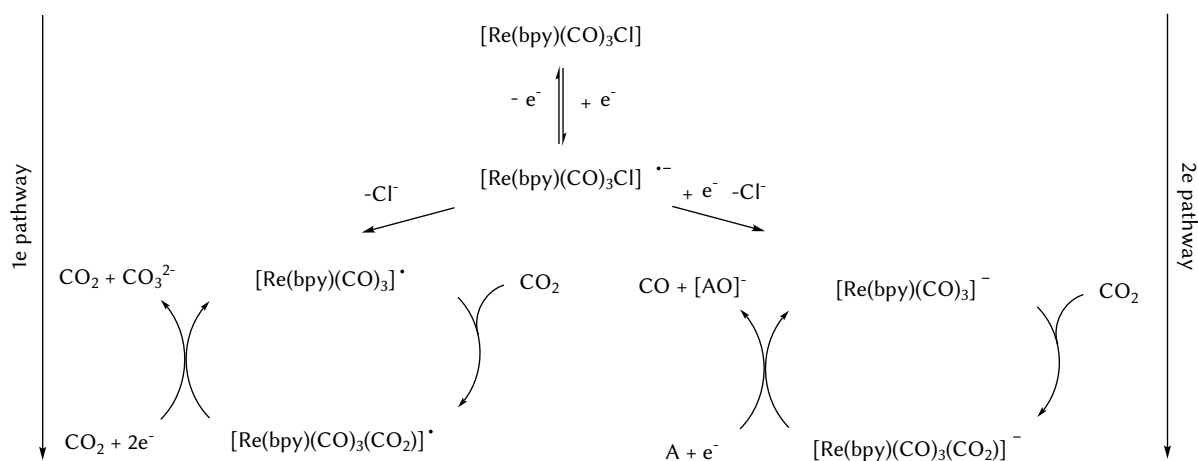


Figure 4.1: One- and two-electron pathways suggested by the group of T. J. Meyer for the mechanism of CO<sub>2</sub> reduction with *fac*-Re(bpy)(CO)<sub>3</sub>Cl.<sup>[128]</sup>

The intermediates occurring during reduction can be followed by *infrared spectro-electrochemistry* (IR-SEC), as already exemplified in the course of the investigations on *fac*-[Re(CO)<sub>3</sub>]-based catalysts with different ligands.<sup>[130–133]</sup> The symmetric and asymmetric vibrational modes of the facial carbon monoxide arrangement show characteristic frequencies in the infrared spectrum with a strong splitting of the symmetric and antisymmetric CO-bands.<sup>[130,133,134]</sup> A variation in the electronic structure of the complex, most prominently in the vicinity of the transition metal, has a direct influence on those frequencies.<sup>[131]</sup> Thus the reduced species during cyclovoltammetry or catalysis can be distinguished and with additional information they can ideally also be assigned to certain structures.

The aromatic ligand framework already proved to be non-innocent during reduction. A further side-reaction in addition to Cl<sup>-</sup> dissociation might thus be possible. Indeed it has been found for [Re(CO)<sub>3</sub>]-type complexes with protic ligands that dihydrogen may be released. This was shown to occur for hydroxy-substituted ligands,<sup>[135]</sup> leading to a small IR shift of about 10 cm<sup>-1</sup> per proton due to the slight increase of the ligand's donor strength. A similar observation was made with mononuclear rhenium complexes containing imidazoles.<sup>[136]</sup> Due to the inserted electrons and effective loss of protons this process has also been termed *reductive deprotonation*.<sup>[136]</sup>

The actual ligand-based reduction of [Re(bpy)(CO)<sub>3</sub>Cl] was found to lead to a shift of approximately 20 cm<sup>-1</sup> and would thus be distinguishable from the mentioned deprotonation.<sup>[130,132]</sup> Kubiak and co-workers studied a methyl-substituted catalyst, [Re(Mebpy)(CO)<sub>3</sub>Cl], and showed that the slow chloride loss results in a charge-transfer from the ligand to the rhenium.<sup>[133,134]</sup> This ligand to metal charge transfer (LMCT) adds to the shifts of the CO bands which are then around 40 cm<sup>-1</sup> lower compared to the unreduced complex.<sup>[133]</sup> They therefore corroborated the first steps in the proposed mechanism, but also showed the utility of the (CO)<sub>3</sub> moiety as a probe for the change in electronic structure.

Mechanistic investigations were also conducted on a supramolecular arrangement of [ReL(CO)<sub>3</sub>] fragments which self-assembled due to hydrogen bridges in the ligand.<sup>[7]</sup> Cyclovoltammetric measurements on this arrangement showed a higher catalytic activity for the reduction of CO<sub>2</sub> in comparison to the mononuclear parent complex at the potential of the first reduction. This increase in activity suggests that a multinuclear arrangement could be beneficial for the transformation of CO<sub>2</sub> by the one electron pathway. However, at the second reduction potential no change in activity was found and thus the 2 e<sup>-</sup>-mechanism is most likely not influenced.<sup>[7]</sup> In the case of the 2 e<sup>-</sup>-mechanism one of the

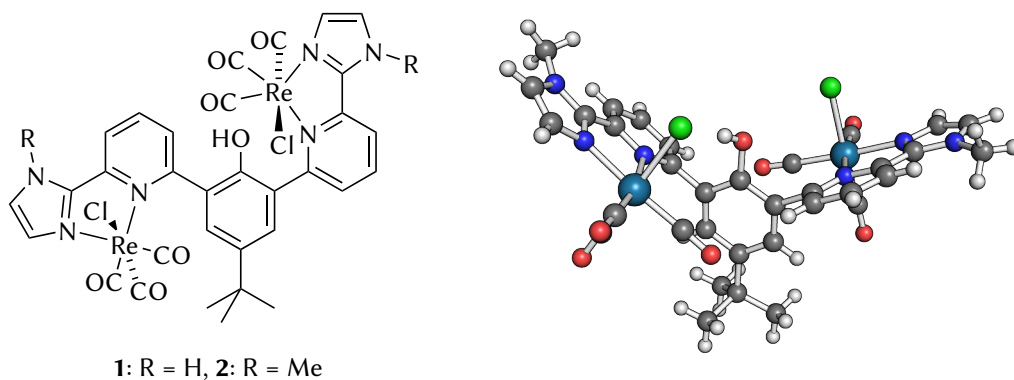


Figure 4.2: Left: Lewis structure of 1, as reported in reference [137]. Right: ball and stick representation of the experimentally determined crystal structure of the methylated backbone ligand.

products stems from an oxygen transferred to acceptor molecules or atoms, likely  $\text{H}^+$  in protic media. Consequently it seems likely that acidic or other proton delivering groups could enhance the activity of  $\text{CO}_2$  reduction catalysts. The validity of this hypothesis was tested by incorporating phenol groups to an iron porphyrin catalyst.<sup>[126]</sup> Catalytic activity of the latter was shown to be a billion times higher than for the methoxy-substituted congener.

Stimulated by the aforementioned results an even more active catalyst could be envisioned that incorporates all these improvements: a dinuclear arrangement with a  $\text{Re}(\text{CO})_3$  active centre and a proton donating group. Such a catalyst has recently been designed in the group of I. Siewert.<sup>[137]</sup>

Incorporating the above features is enabled by the backbone ligand L (4-tert-butyl-2,6-bis(6-(1H-imidazol-2-yl)-pyridin-2-yl)phenol) as depicted in Figure 4.2, which was also synthesised with 1Me-imidazolyl substituents. The unmethylated complex was not crystallised and thus Figure 4.2 only shows the ball and sticks representation of the catalyst with the methylated imidazolyl moieties. However, the influence was expected to be marginal. As for complexes with similar  $\text{Re}(\text{CO})_3$  moieties 1 also exhibits facial coordination by the carbon monoxide ligands and the rhenium ions are enclosed in a mostly undistorted octahedral environment. This enables a hydrogen bond of the phenolic OH-group to one of the chloride ions. The Lewis picture of the ligand suggests a conjugated  $\pi$ -system. As seen in the representation of Figure 4.2, the planes of the sidearms are almost orthogonal to each other. An electronic coupling of both catalytic centres is therefore not intrinsic to the structure. Part of the results presented in the following sections has already been published elsewhere.<sup>[137]</sup>

## 4.1 Computational Determination of Reduction Intermediates

The complex was thoroughly investigated with UV/Vis-SEC and IR-SEC so that each of the reduction potentials of 1 were assigned infrared and UV/Vis spectra. With the CO group stretching bands unobstructed in the  $1700\text{--}2100\text{ cm}^{-1}$  range and sensitive to any change in the catalyst's environment the corresponding intermediate structures should be identifiable. Nonetheless, without further information a direct assignment or a differentiation between stereochemical configurations is challenging. Additional clues can however be supplied by theoretical methods with the carbon monoxide vibrations serving as the interface to experiment.

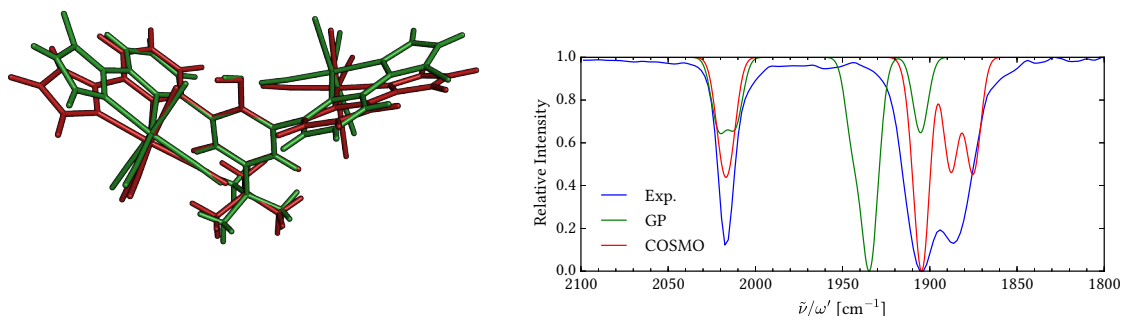


Figure 4.3: Left: Overlay of the experimental crystal structure (green) of **2** with the methyl groups replaced by a hydrogen atom and the optimised minimum on the BP86-D3(BJ)/TZVP/COSMO level (red). Right: Comparison of calculated IR spectra with a solvent model (COSMO) and in gas phase (GP) to the experimental bands (Exp). Theoretical spectra were shifted to coincide at the band of highest frequency.

#### 4.1.1 Computational Details

Starting with the unreduced complex, the infrared spectra were computed for a number of viable molecular geometries on the basis of density functional theory. For the optimisation and frequency calculation the GGA functional BP86<sup>[27,28]</sup> was used with the def2-TZVP basis set<sup>[109]</sup> and D3 dispersion corrections<sup>[42]</sup> with Becke-Johnson damping.<sup>[44]</sup> The integral calculation was accelerated using the resolution of the identity approximation<sup>[138]</sup> for which the associated def2 fitting basis<sup>[110]</sup> set was applied. To further reduce the computation time the first 60 core electrons of rhenium were substituted by the effective core potential ECP60MWB.<sup>[139]</sup> COSMO was used with the standard parameters for DMF ( $\epsilon = 38.3$ ) to model bulk solvent effects.<sup>[54]</sup> The frequencies and intensities were determined in the usual double harmonic approximation.

In DFT the parametric nature of most functionals precludes the prediction of a functional's accuracy based on the underlying theoretical sophistication. Therefore it is common practice to apply scaling factors to vibrational frequencies which were determined using DFT methods. As the transferability is debatable and the mentioned combination of functional and basis set was already used successfully for vibrational fingerprints,<sup>[66,140]</sup> a predefined factor was not utilised in this study. Instead the accuracy was evaluated by direct comparison of computed values to the experimental spectrum. To accomplish this the structure corresponding to the experimental spectrum has to be known so as to be able to exclude any deviation introduced by the neglect of structural sampling. Only the crystal structure of **2** was available, but **1** is expected to behave similarly and the former was accordingly used with the methyl groups replaced by hydrogen atoms. As shown in Figure 4.3 on the left, after optimisation of the crystal structure the geometry changes only slightly, which suggests a similar minimum for the complex in solution. The acquired structure should thus be a reasonable approximation to the one in the experiment. To assess the influence of the solvation environment the vibrational frequencies of **1** were additionally calculated without a solvent model, but on the same level of theory. By aligning the spectra at the frequency of the symmetric stretching mode of the CO ligands as in Figure 4.3, the difference of the computed to the experimental values are clearly visible. Unshifted results would largely diminish the agreement, but keeping the original values is not necessary in this case since only relative vibrational shifts are needed. This is made possible by the high certainty of the complex's structure in

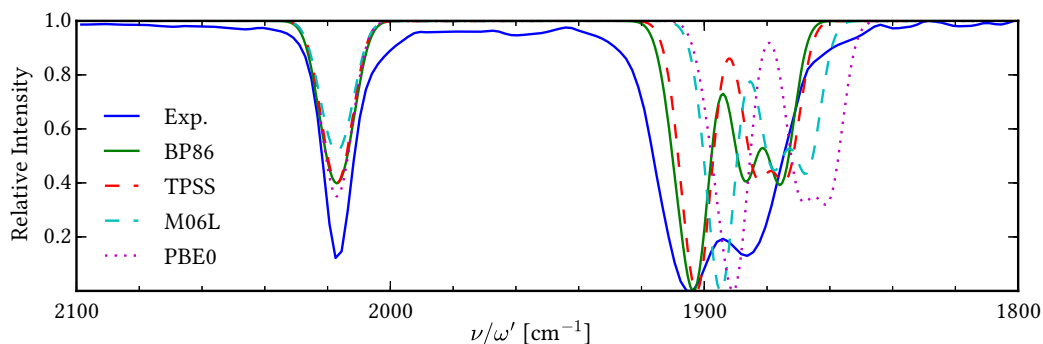


Figure 4.4: Comparison of the functional influence on the accuracy to experiment. Vibrational spectra were obtained on the X/TZVP/COSMO level of theory, where X is one of BP86, M06-L, TPSS or PBE0. The experimental spectrum and BP86 use solid lines, whereas meta-GGAs are dashed and hybrids employ a dotted linestyle. Intensities are presented after normalisation to the highest band in the range  $1800 \leq \omega' \leq 2100$ . The spectra are aligned to coincide at the experimental band for the symmetric stretching modes at  $2017 \text{ cm}^{-1}$ .

solution and using it as the reference for all the other intermediates. From the comparison it is evident that the solvent effect is not negligible and a good agreement for the separation of the symmetric to the asymmetric bands is only achieved by including the continuum model in the computational protocol. A constant shift of  $+31 \text{ cm}^{-1}$  was determined as the deviation of the calculated to the experimental spectrum, which further serves as the correction to all vibrations that are compared to experiment.

It could be argued that BP86 is too simple in construction and newer meta-GGA or hybrid functionals might be more appropriate for the calculation of vibrational frequencies. To evaluate this, several functionals were subsequently used to optimise the crystal structure and calculate the infrared spectra with the solvent model. Dispersion corrections were not included to acquire only the effect of the functional's construction. The results for BP86, M06-L,<sup>[31]</sup> TPSS<sup>[29]</sup> and PBE0<sup>[141]</sup> are collected in Figure 4.4. Comparing the deviation of BP86 to the experimental bands for the asymmetric stretching modes, a slightly reduced agreement is found than if dispersion corrections are included as shown in Figure 4.3. This corroborates the use of the D3 method, but also suggests that BP86 is already appropriate without corrections for the calculation of infrared spectra. The TPSS functional gives very similar results with only marginally larger deviations in the separation of bands within the region of asymmetric stretching modes. On the other hand M06-L and PBE0 agree significantly less with experimental values and would thus not be recommended. This discrepancy between the functionals may be somewhat mitigated by employing dispersion corrections. In the course of this study only relative shifts of frequencies are calculated and this is expected to be very similar for all methods, justifying the use of other functionals within this investigation. Another graph to support this hypothesis can be found in the appendix. In any case, BP86 overall reveals the highest accuracy which reinforces confidence in its further use.

#### 4.1.2 Results

The shift as well as the corresponding level of theory set the basis for a larger, comprehensive investigation of the reduction intermediates. In the experiment two irreversible reduction waves were observed at about  $-2.0 \text{ V}$  and  $-2.6 \text{ V}$  versus the ferrocene couple respectively, with two electrons added in the first step. For the second reduction the number of transferred electrons was not as clear and two to three

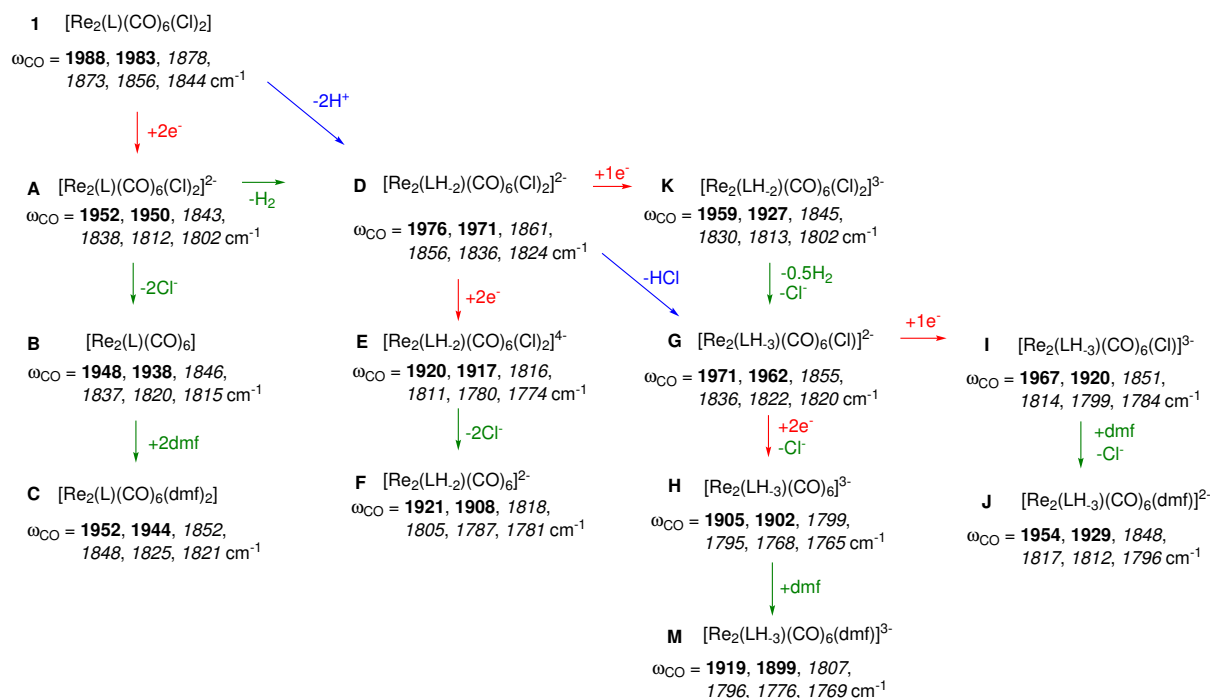


Figure 4.5: Intermediates considered during the reduction. Reduction pathways are marked by red arrows, deprotonation in blue and ligand dissociation or exchange in green. Computed harmonic CO stretching frequencies are listed below the chemical formulae in  $\text{cm}^{-1}$ . Symmetric stretching modes are typed in boldface. Scheme adapted from reference [137].

could be assigned. The optimised complex derived from the crystal structure sets the starting point for the theoretical investigation of the intermediates. Thus only a few meaningful chemical processes are possible of which the most likely are shown in Figure 4.5.

### The first reduction

Following the description of reactivity during the reduction of  $\text{Re}(\text{L})(\text{CO})_3$ -type complexes in the literature, the first reduction was expected to be ligand based and possibly lead to a subsequent chloride dissociation. An IR shift of approximately  $-18$  to  $-30 \text{ cm}^{-1}$  should consequently be observed. Since it was not known whether a successive reaction took place at all the structures **A**, **B** and **C** qualify as possible intermediates. Additionally the doubly deprotonated intermediate **D** was taken into account. Although the latter reaction was not expected to occur, similar examples with dihydrogen production during reduction exist for imidazolyl groups as in the case of investigations by Hartl and co-workers.<sup>[136]</sup> Likewise the experimental shift of  $-11 \text{ cm}^{-1}$  for the symmetric and  $-18 \text{ cm}^{-1}$  for the asymmetric modes are much lower than what is expected in the case of ligand-based reduction.<sup>[131–133]</sup> A superposition of the computed values of the aforementioned structures with the experimental spectrum is shown in Figure 4.6 and Figure 4.7. By simply comparing the computed shifts of **1** and **A**, **B** or **C** the values agree well with the ones from the literature, although already for the plain  $2e^-$ -reduction shifts larger than  $30 \text{ cm}^{-1}$  are achieved. The dissociation of chloride ions indeed leads to a further increase of the band shift (**A**  $\rightarrow$  **B**) similarly to the reported effects in  $[\text{Re}(\text{bpy})(\text{CO})_3\text{Cl}]$ , which however is partly reversed by recoordination with solvent molecules (**B**  $\rightarrow$  **C**). Notwithstanding, the shifts of **A**, **B** and **C** are far off the experimental values as depicted in Figure 4.6 and Figure 4.7. In contrast, the shifts of the doubly deprotonated structure **D** are in very good agreement with the experimental

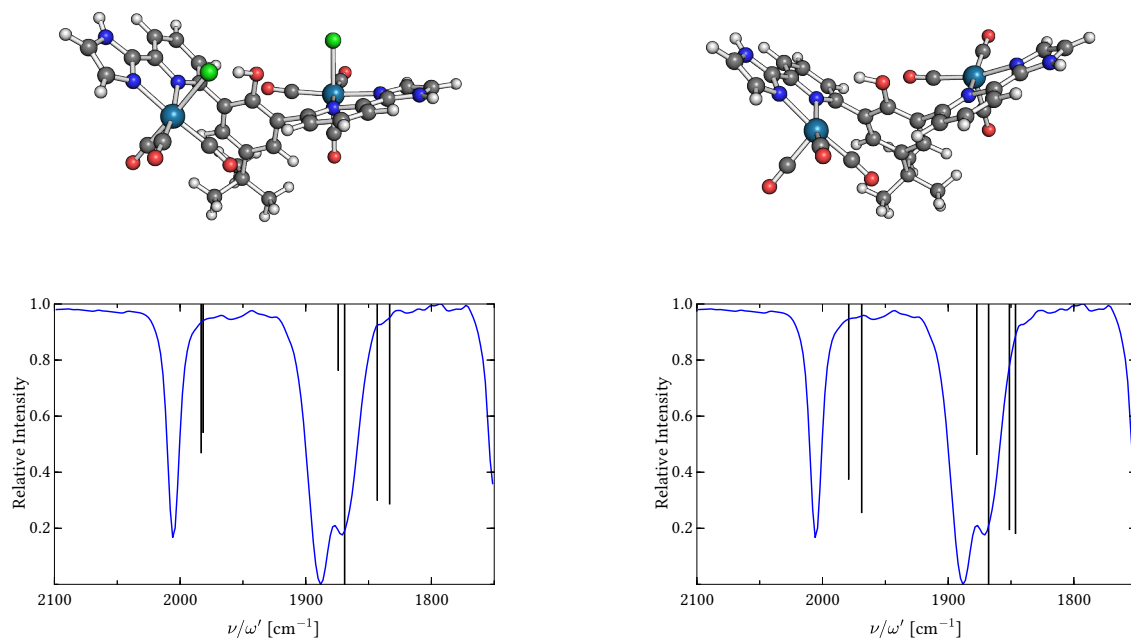


Figure 4.6: Experimental (blue curve) and calculated (black bars) vibrational spectra with the corresponding geometries (ball-and-stick pictures) of  $1^{\text{red1}}$ , A and B. All calculated values were shifted by  $+31 \text{ cm}^{-1}$ .

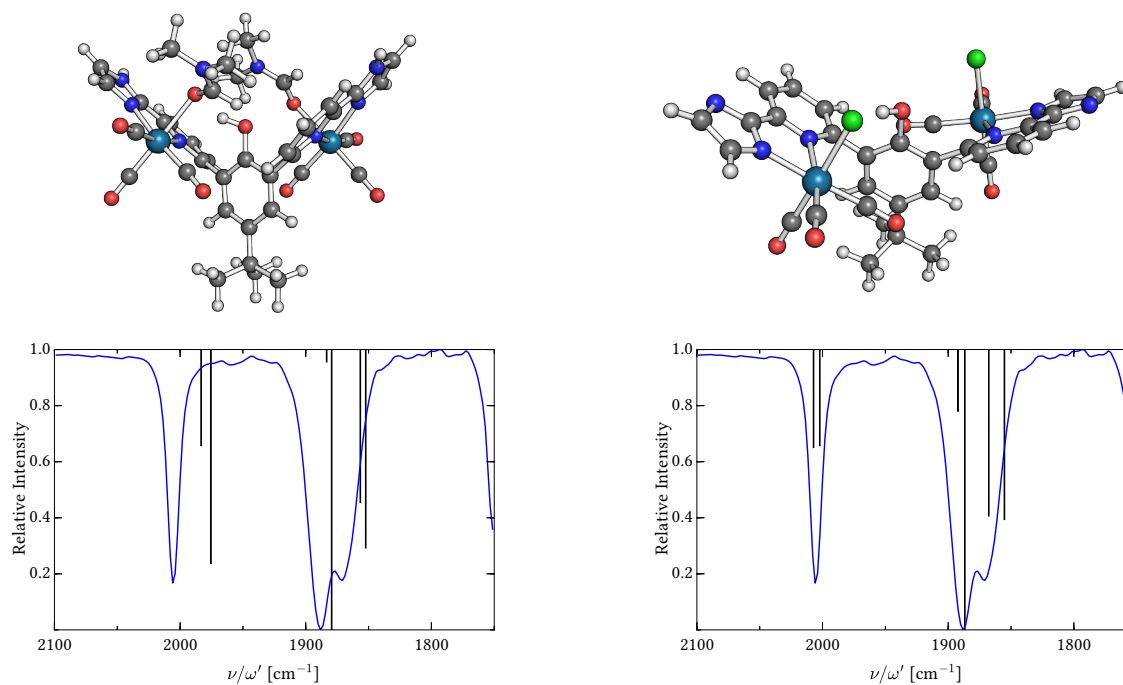


Figure 4.7: Experimental (blue curve) and calculated (black bars) vibrational spectra with the corresponding geometries (ball-and-stick pictures) of  $1^{\text{red1}}$ , C and D. All calculated values were shifted by  $+31 \text{ cm}^{-1}$ .

spectrum. Theoretically this is a very strong hint that **D** should be the correct intermediate. However, due to the uncertainty in the accuracy of the method this finding has to be corroborated by experiment. Because  $1 \rightarrow \mathbf{D}$  is effectively a deprotonation step, Wilting and Siewert added 2 eqLDA to a solution of **1**.<sup>[137]</sup> The reduction and deprotonation lead to the same spectrum which in combination with theoretical results suggests that **D** is actually the correct structure for  $1^{\text{red1}}$ . This is also conceptually in line with the small IR shift since the local electronic structure of the transition metal is barely changed by deprotonation of the imidazoles. The only difference is a slight increase in donor strength of the backbone ligand.

### The second reduction

Up to  $1^{\text{red1}}$  the experimental evidence would have been sufficient to identify the intermediate. However, during further reduction the experimental assignment of structures becomes more difficult. For one, at the potential of the second reduction, two different intermediates were postulated based on the changes in the IR spectrum. For another the dissociation and coordination of ligands is becoming more likely with additional negative charge inserted into the complex. A direct consequence of ligand loss, which would also be visible in experiment, would likely be the splitting of the bands and thus an asymmetric IR spectrum. The first structure seen in experiment at the second reduction, **Int**, can again correspond to several different reactions, starting from **D**. Either the electrons are just inserted into the complex (**E**) and the successive dissociation of chloride ligands (**F**) leads to the second and last reaction product  $1^{\text{red2}}$ , or **Int** is reached by only a single electron reduction (**K**) with possible successive ligand exchanges giving  $1^{\text{red2}}$ . Since the shift in the experimental vibrations between  $1^{\text{red1}}$  and **Int** is again quite small, an additional dihydrogen formation may be suggested. However, not many polar bonds with hydrogen atoms remain and as such **G** is deprotonated at the phenolic group of the backbone ligand. Some structures can already be ruled out without doing any time consuming calculations and actually have already been omitted from Figure 4.5. Looking at the experimental spectrum in Figure 4.8 a slight asymmetry in the bands can be recognised. The corresponding structure would therefore also have to feature a difference in coordination around the transition metal. The spectra of the resting state **1** and the product of the first reduction  $1^{\text{red1}}$  do not exert such an asymmetry. Secondary differences in coordination like the hydrogen bond of the phenol group seem to be too small to induce such a change. That being said the difference in the newly coordinating ligand should not be too strong, otherwise a much larger splitting of the single bands would be expected. Consequently structures **E** and **F** can already be ruled out. In the case of **E** the symmetry of the calculated spectrum does not fit to experiment whereas it would depend on the actual configuration for **F**. Even though the computed frequencies show a significant splitting for the latter, both spectra exhibit redshifts of too large magnitude to be consistent with experiment. The single reduction product **K** features a pronounced splitting due to the geometrically near orthogonal orientation of the ligand sidearms, which reduces electronic coupling between the metal-sites.<sup>[142,143]</sup>

Without any further information one could argue that the uncertainty of the method prohibits any further distinction and **K** could well be the correct intermediate even though the agreement is unsatisfactory. However, due to the confidence gained by identifying the preceding intermediates such a discrepancy seems unlikely. Furthermore the differences in the frequencies compared to experiment is still quite large for **K** and the experimental shifts are reminiscent of the first reduction which hints at another dihydrogen evolution step. A complex with a free phenolate group in a relatively unpolar and



already negatively charged moiety is deemed unfavourable. Thus the dissociation of a chloride ligand can be expected as a result of deprotonation. Combined with the dihydrogen evolution this transformation would effectively be a HCl dissociation ( $\mathbf{D} \rightarrow \mathbf{G}$ ). An open coordination site as well as the association of a solvent DMF molecule do not seem likely in light of the spectral features. Intermediate  $\mathbf{G}$  therefore could involve the coordination of the phenolate oxygen to one of the rhenium sites which is in line with the notion that the chloride would have to be substituted for a similar ligand to arrive at a spectrum close to the experimental one observed for  $\mathbf{Int}$ .

Reviewing the reactions necessary to reach  $\mathbf{G}$  from  $\mathbf{1}$ , one may notice that no electrons are actually transferred to the complex. In total only three protons were abstracted and a chloride ion dissociated. Similarly to the process of assigning  $\mathbf{1}^{\text{red1}}$ , deprotonation with three equivalents of a base should lead to the structure and spectrum of  $\mathbf{Int}$ . This only holds as long as the chloride loss is induced by the accumulation of negative charge due to the deprotonation. The deprotonation was checked by Wilting and Siewert by adding LDA and corroborated the hypothesis of further  $\text{H}_2$  development.<sup>[137]</sup> Loss of chloride was independently checked in experiment by reduction in the presence of either  $n\text{-Bu}_4\text{NPF}_6$  or a 100-fold excess of  $n\text{-Bu}_4\text{NCl}$ .<sup>[137]</sup> They acquired a spectrum that resembled  $\mathbf{1}^{\text{red1}}$  with using  $n\text{-Bu}_4\text{NCl}$  which increases the chloride concentration in solution. This supports the assumption of an effective HCl removal, but still leaves the fate of the empty coordination site unanswered. An extensive computational investigation of all configurations and ligands could provide an explanation to the possible structures for  $\mathbf{Int}$ . Nevertheless, an educated guess can lead to a huge saving in resources invested into one structure. Although a chloride definitely dissociates in the course of the reduction, the environment of the affected rhenium atom cannot change significantly. Binding of a solvent molecule is thus already ruled out since it would break the symmetry of the vibrations as seen already for  $\mathbf{A}$  and  $\mathbf{C}$ . Additionally DMF is not charged and a similar ligand as chloride should also carry a negative charge. One of the most likely candidates that could serve as a replacement for chloride is therefore the phenolic oxygen which was also used for the structure of  $\mathbf{G}$ . As seen in Figure 4.8 the spectra of the deprotonation coincide and agree well with the theoretical results for  $\mathbf{G}$ .

Experimentally the IR-SEC results suggest an additional intermediate  $\mathbf{1}^{\text{red2}}$  that directly follows the occurrence of  $\mathbf{Int}$  at the second reduction potential.<sup>[137]</sup> As such the final structure of  $\mathbf{1}^{\text{red2}}$  should be somewhat similar to  $\mathbf{G}$  with possible reactions including additional chloride loss and DMF coordination. In experiment, with the same test as for  $\mathbf{Int}$ , a dependence on the chloride concentration was shown.<sup>[137]</sup> The total number of transferred electrons at the second reduction potential was determined to be slightly higher than at the first reduction, which includes the reaction to  $\mathbf{Int}$ .<sup>[137]</sup> With the currently suggested pathway to  $\mathbf{G}$  this leaves one electron for the further transformation  $\mathbf{Int} \rightarrow \mathbf{1}^{\text{red2}}$ . Nevertheless, the error margin in the calculation of the transferred electrons did not rule out a  $3e^-$  reduction.<sup>[137]</sup> Thus the two electron pathways to  $\mathbf{H}$  and  $\mathbf{M}$  were also calculated, but the strong shifts invalidate this assumption and support the experimental assignment of one additionally transferred electron. In light of this, two structures qualify to represent  $\mathbf{1}^{\text{red2}}$ .  $\mathbf{I}$  is simply  $\mathbf{Int}$  with an additional electron. The calculated spectrum roughly resembles the shape of the experimental one for  $\mathbf{1}^{\text{red2}}$ , but the splitting of the symmetric vibrations is with  $\approx 47\text{ cm}^{-1}$  larger than in the experiment. By exchanging the chloride ligand for a solvent molecule structure  $\mathbf{J}$  is acquired. The relative intensities and separation of the asymmetric vibration bands coincide more with experiment and the symmetric bands reveal a separation which fits better than for  $\mathbf{I}$ . All in all the agreement is only fair in comparison to previous as-

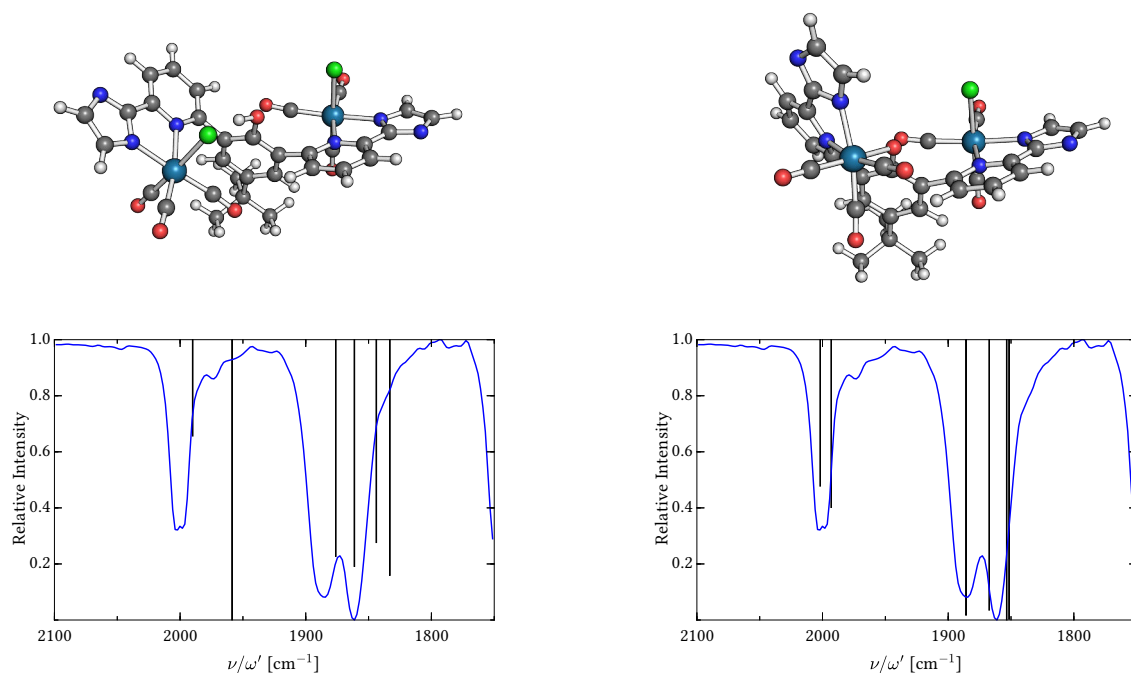


Figure 4.8: Overlays of the experimental (line) and calculated (bars) vibrational spectra for the intermediate at the second reduction potential, **Int**. Left: calculated values for **K**. Right: calculated values for **G**. All theoretical values were shifted by  $+31 \text{ cm}^{-1}$ .

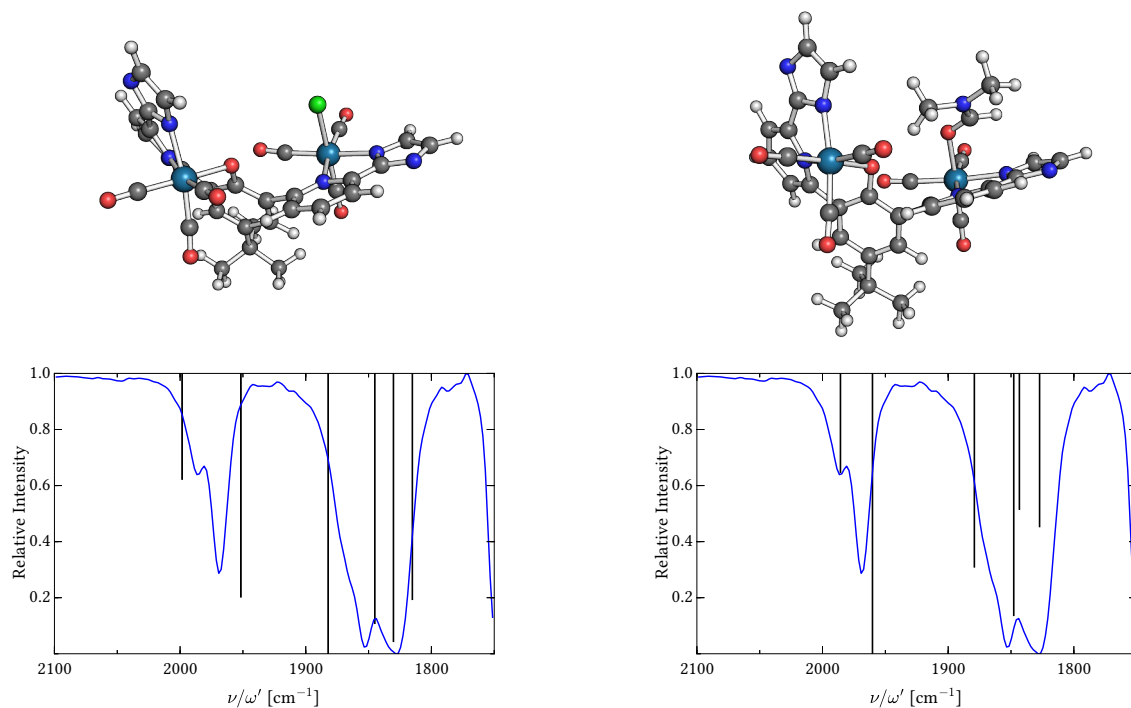


Figure 4.9: Experimental (blue curve) and calculated (black bars) vibrational spectra for **1<sup>red2</sup>**, **I** (left) and **J** (right) respectively. Theoretical frequencies were shifted by  $+31 \text{ cm}^{-1}$ . Structures on top of the spectra show the corresponding minimum geometries.

signments and due to the lack of additional experimental evidence for  $\mathbf{1}^{\text{red}2}$  a small uncertainty remains. However, none of the other calculated geometries fits nearly as well as **J** and even the asymmetries in the low frequency modes of the experimental spectrum are quite well reproduced. The remaining discrepancy may thus be due to a larger number of conformations that play a role in solution due to the freedom of rotation for the DMF ligand. A sampling of the the conformational space would thus be necessary and may lead to a better agreement, but is prohibitive in this case.

## 4.2 Conclusion

This study revealed the possibility of characterising reaction intermediates which would otherwise be hard to define if not completely inaccessible. Even in situations where experimental evidence is available, as was the case for deprotonation and chloride loss of **Int**, some parts of the reaction can more easily be verified by theoretical methods. In the example of **Int** this was the coordination of the phenolate group to rhenium, which is the reason for the small experimental IR shift instead of a larger one. With the comprehensive calculation of several other possible intermediate structures these results also serve as a reference for similar investigations and compounds. Based on the work described here, a better understanding of  $\text{CO}_2$  reduction intermediates by combined experimental and computational approaches should be possible. One of the spectacular results is the overall good agreement between experiment and theory by using a rather simple and computationally economical functional. Therefore the approach presented here should be applicable to a wider range of systems and a larger audience than state-of-the-art methods that lead to the same conclusions.

As a summary of the understanding gained about the system under study the experimental and theoretical spectra are depicted in Figure 4.10. Three separate intermediates were observed experimentally by IR-SEC,<sup>[137]</sup> with hints of chloride dissociation and hydrogen evolution. For each of the intermediates a likely structure could be assigned through the synergy of experimental and computational results. In total the processes during reduction were determined to consist of a deprotonation with  $\text{H}_2$  formation at the first potential ( $\mathbf{1}^{\text{red}1}$ ,  $[\text{Re}_2(\text{LH}_{-2})(\text{CO})_6(\text{Cl})_2]^{2-}$ ) after transferring two electrons into the system. At the second reduction an intermediate (**Int**,  $[\text{Re}_2(\text{LH}_{-3})(\text{CO})_6\text{Cl}]^{2-}$ ) is generated by one electron and the successive dissociation of a chloride ligand, accompanied by  $\text{H}_2$  evolution and coordination of the phenolate to the vacant site. **Int** is readily reduced by another electron, leading to a further dissociation of the second chloride and subsequent solvent (DMF) coordination ( $\mathbf{1}^{\text{red}2}$ ,  $[\text{Re}_2(\text{LH}_{-3})(\text{CO})_6(\text{DMF})]^{2-}$ ).

Although the task and aim of this study were fulfilled some questions about **1** still remain to be answered. A major part is the description of the important structures created during catalysis since the intermediates identified previously were all based on measurements in exclusion of  $\text{CO}_2$ . There is also evidence that different intermediates are reached during catalysis, as a faster scan rate during the reduction leads to the possible occurrence of an intermediate with a slightly different potential.<sup>[137]</sup> This effect was attributed to the direct reduction of the complex without dissociation of the chloride ligand, for example structure **H**. The same is seen during catalysis which hints at a different series of intermediates for the actual reaction with  $\text{CO}_2$ . Further insight in this direction would be needed to describe the activity of the catalyst, and the current work can serve as a basis for this endeavor.

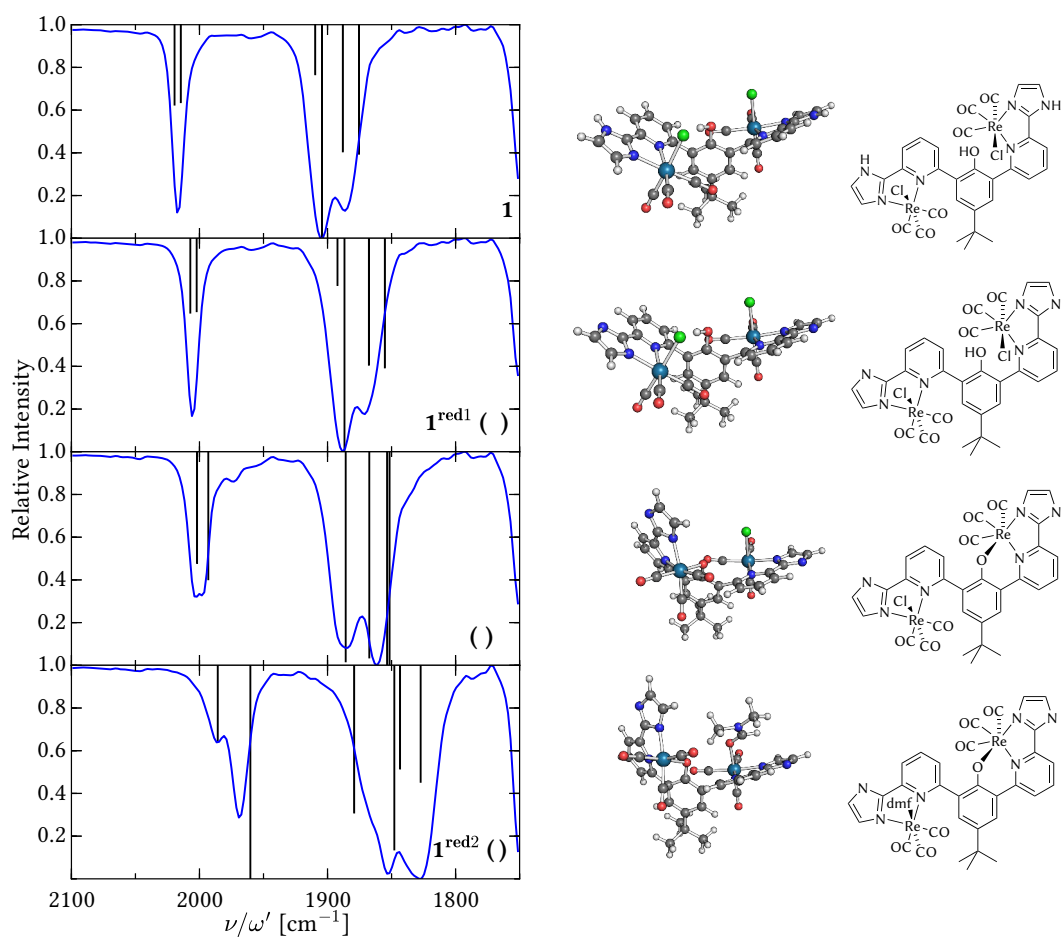
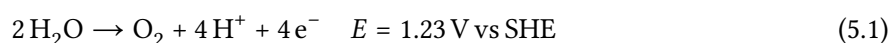


Figure 4.10: Left: experimental spectra for the separate intermediates that occur during electrochemical reduction (blue curves), superposed with calculated harmonic vibrational frequencies of the CO stretching modes (black bars). The theoretical spectra were uniformly shifted by +31 cm<sup>-1</sup>. Right: ball and sticks representation of the corresponding minimum structures responsible for the theoretical spectra on the left. Graph of IR spectra and Lewis structures adapted from reference [137].

## 5 Artificial Water Oxidation

Splitting water into its components hydrogen and oxygen is one very promising method in the storage of energy using clean fuels. However, the oxidation half-reaction to create dioxygen (OER) is thermodynamically demanding as expressed by a high standard potential (5.1).



Additionally the accumulation or separation of four charge equivalents in one confined space raises structural challenges in the design of catalysts for this reaction. In nature this is accomplished by Photosystem II (PSII, Figure 5.1) which channels energy gained from electronic excitations into the oxygen evolving complex (OEC) that oxidises water molecules with a high efficiency.<sup>[9]</sup> The OEC is made of a

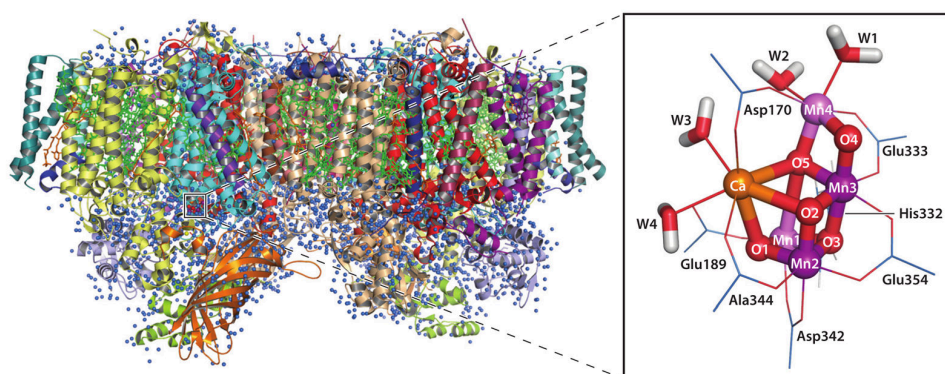


Figure 5.1: Photosystem II (left) and the oxygen evolving complex (OEC, right). Picture created by Vinyard and Brudvig.<sup>[9]</sup>

manganese oxide cluster with a supporting calcium ion, as shown in Figure 5.1. Manganese can reach a range of different oxidation states,<sup>[144]</sup> enabling the cluster to accumulate positive charge. Due to the protein environment the oxidation steps in PSII may additionally be eased by concerted removal of a proton.<sup>[145]</sup> This redox-leveling<sup>[146,147]</sup> necessitates a multi-step process to create dioxygen.

Mimicking and eventually surpassing the efficiency of PSII would go a long way towards creating a viable alternative to current energy storage. To develop successful artificial water oxidation catalysts one has to gain a profound understanding of the reactivity and influences on the mechanism. Beyond the bioinspired OEC mimics,<sup>[10,147,148]</sup> oxygen evolution has also been investigated in molecular organic frameworks,<sup>[149,150]</sup> heterogeneous<sup>[151–153]</sup> and homogeneous catalysts based on various transition metals,<sup>[12,154–159]</sup> as well as molecular clusters.<sup>[160–162]</sup> One of the most intensely studied and well characterised catalysts are ruthenium based, however. The proposed mechanisms for the reaction of these catalysts can be divided into two pathways: the interaction of two metal-oxyl radicals (I2M) and the water-nucleophilic attack (WNA). I2M can occur between separate molecules, or intramolecularly (i-I2M) as depicted in Figure 5.2 for a general dinuclear ruthenium complex.<sup>[14,163]</sup> In the former path-

way two oxygen radicals directly form a bond, resulting in a peroxide intermediate (AC  $\rightarrow$  I1). One of those Ru–O bonds is successively broken, accompanied by the binding of a water molecule (I1  $\rightarrow$  I2). The same intermediate can be reached by the water-nucleophilic attack pathway (W1) which only differs from I2 formally by the position of a proton. Consequently the actual oxygen evolution step in the reaction is very similar in I2M and WNA. In this step the last Ru–O bond has to be cleaved while the peroxide is simultaneously oxidised by the complex.

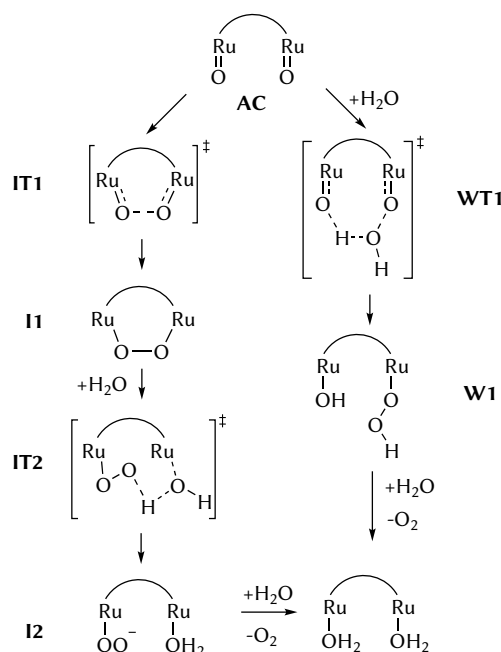


Figure 5.2: I2M (left) and WNA (right) mechanism as often applied to computational studies of dinuclear ruthenium water oxidation catalysts.<sup>[14]</sup>

Notwithstanding the extensive attention that ruthenium based water oxidation catalysts have received, the reason why a catalyst reacts through a certain mechanism is still not fully understood. Several attempts have been made to rationalise the reactivity based on several qualitative measures like the oxidation state of the metal site or the radical character of the metal-bound oxygen.<sup>[164]</sup> For mononuclear complexes most catalysts have been found to exhibit WNA reactivity, which was attributed to the high entropic penalty that a bimolecular I2M mechanism entails. Nevertheless the WNA mechanism is assumed to be generally slower than I2M for similar complexes because the Ru=O group has to be more electrophilic in the former and therefore requires higher oxidation potentials.<sup>[164]</sup> Thus the reactivity can be steered to I2M by incorporating negatively charged ligands as has been shown with the development of the Ru-bda catalyst by Sun and co-workers.<sup>[165,166]</sup> Additionally the negative charge is expected to increase the radical character of the oxyl groups which is also beneficial for I2M catalysts. As such the computed spin-density on the ruthenium bound oxyl groups was already used as an index to predict the mechanism of catalysis.<sup>[167]</sup>

In the case of dinuclear catalysts the multiple transition metals can enhance the reactivity by redox leveling. Compared to their mononuclear counterpart the intermolecular radical coupling pathway is often disfavoured by shielding of the active site due to steric constraints. As shown with the design of the Ru-Hbpp (3) catalyst the intramolecular I2M can be enforced by careful design of the ligands to create a pre-arranged active site beneficial for i-I2M.<sup>[14,169,170]</sup> Notwithstanding, a change in the axial

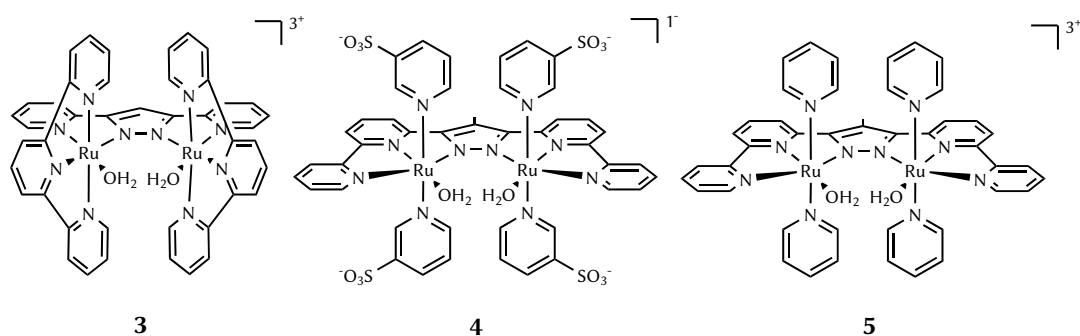


Figure 5.3: Three dinuclear water oxidation catalysts with a similar ligand framework and characterised experimentally,<sup>[13,14]</sup> Ru-Hbpp (3), Ru-Mebbp (5) and Ru-Mebbp-SO<sub>3</sub> (4) with sulfonated axial ligands.

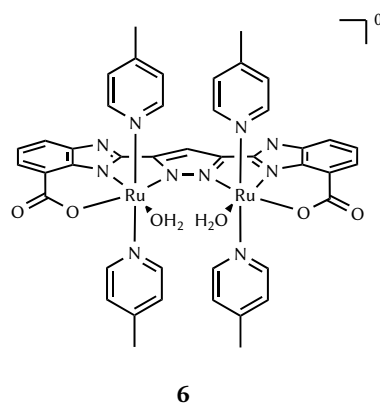


Figure 5.4: The  $[(\text{H}_2\text{L})\text{Ru}_2^{\text{II,III}}(\text{pic})_4]^{2+}$  complex Ru-cbim.<sup>[168]</sup> Equatorial picoline ligands were exchanged with water molecules for consistency with the other Lewis diagrams.

ligands to direct the Ru=O groups away from each other did not result in WNA reactivity, but intermolecular I2M.<sup>[171]</sup> Water nucleophilic attack could be achieved by a change in the backbone ligand, leading to the Ru-Mebbp-SO<sub>3</sub> (4) catalyst.<sup>[13]</sup> This complex exhibited catalytic activity with different substituents on the axial pyridines<sup>[172,173]</sup> with some structurally similar to Ru-Hbpp, for example Ru-Mebbp (5). The same reasoning as for mononuclear complexes led to the incorporation of negatively charged ligands in the Ru-cbim (6) catalyst.<sup>[168]</sup> Calculations on 6 suggested an i-I2M mechanism under the premise of an additional oxidation.<sup>[15]</sup> Although several efficient dinuclear water oxidation catalysts were mechanistically investigated in recent years, the steering effects are still elusive.

Computational investigations can help to reach a better understanding of the influences on the reactivity of catalysts. In addition to the blue dimer<sup>[174,175]</sup> a couple of newer and more active mono-<sup>[167,176–180]</sup> and dinuclear<sup>[14,15,163]</sup> catalysts were investigated theoretically to explain their reactivity. Notwithstanding, to unravel ligand influences on the mechanism similar catalysts that take different mechanistic pathways need to be compared on the same footing. The mechanisms of Ru-Hbpp and Ru-Mebbp have been experimentally determined and exhibit quite similar active sites. However, Ru-Mebbp has not been investigated computationally and thus a detailed theoretical basis has to be established. Additionally Ru-cbim, although structurally quite different, can serve as a probe for the validity of the predictions on Ru-Mebbp as it is also an efficient water oxidation catalyst with detailed computational results.<sup>[15]</sup> As such this study tries to redeem the lack of theoretical results for 4 and compare to the other catalysts on the same footing.

## 5.1 Computational Details

The calculation of reactions occurring in catalytic cycles requires an accurate description of several energetic contributions. Due to the size of the systems and constantly changing electronic structure in catalytic processes, DFT is often used to probe the energetics. However, reaction barriers can vary greatly with different functionals employed.<sup>[181–183]</sup> To arrive at a result with high confidence, functionals are often benchmarked to a reference value, either obtained from more accurate wavefunction methods or experiment. Consequently several different functionals have been used in this study for optimisation and electronic energy calculations. Minimum structures were determined using the functionals M06-L<sup>[31]</sup>, BP86<sup>[27,28]</sup> or B3LYP\*<sup>[36]</sup> combined with the def2-SVP<sup>[109]</sup> basis set and its associated effective core potential ECP28MWB.<sup>[139]</sup> B3LYP\* is a modified version of the famous B3LYP hybrid functional<sup>[34,35,184]</sup> with the exact exchange factor reduced to 15%. The electronic energies were refined using the larger def2-TZVP<sup>[109]</sup> basis and a variety of GGA (BP86,<sup>[27,28]</sup> PBE<sup>[108]</sup>), meta-GGA (M06-L,<sup>[31]</sup> TPSS<sup>[29]</sup>), hybrid (B3LYP, B3LYP\*) and meta-hybrid (M06,<sup>[185]</sup> TPSSH<sup>[186]</sup>) functionals. Integration grids were set to Grid4 for optimisations and larger ones for frequency and single point calculations, as implemented in the ORCA suite of programs.<sup>[187]</sup> Radial grids were increased to 7.0 on the ruthenium atoms to account for the stronger varying density around the metal sites. To speed up the calculations of the Coulomb integrals the resolution of the identity (RI) approximation<sup>[138]</sup> was used with the associated fitting basis,<sup>[110]</sup> whereas the exchange part was approximated by the chain-of-spheres approach.<sup>[188]</sup> Dispersion effects were corrected for using the D3 method<sup>[42]</sup> with Becke-Johnson damping<sup>[44]</sup> developed by Grimme and co-workers. The influence of the bulk solvent was taken into account with the continuum solvent model COSMO<sup>[54]</sup> and the corresponding standard values in ORCA, that is a dielectric constant of 80.4 was assigned for water as the solvent.

Thermodynamic corrections were determined within the rigid-rotor harmonic-oscillator (RRHO) approximation and added to the inner energy as retrieved with the above methods.<sup>[33]</sup> In RRHO the molecular geometry is assumed to be fixed during rotation and the vibrations are treated in a harmonic framework which enables the analytic calculation of the corrections as shown in equations (5.2)–(5.7).

$$H_{\text{trans}} = \frac{5}{2}RT \quad (5.2)$$

$$H_{\text{rot}} = \frac{3}{2}RT \quad (5.3)$$

$$H_{\text{vib}} = R \sum_{i=1}^{3N-6} \left( \frac{h\nu_i}{2k_B} + \frac{h\nu_i}{k_B(e^{h\nu_i/k_B T} - 1)} \right) \quad (5.4)$$

$$S_{\text{trans}} = R \left( \frac{5}{2} + \ln \left( \frac{V(2\pi M k_B T)^{3/2}}{N_A h^3} \right) \right) \quad (5.5)$$

$$S_{\text{rot}} = R \left( \frac{3}{2} + \ln \left( \frac{\sqrt{\pi}}{\sigma} \left( \frac{8\pi^2 k_B T}{h^2} \right)^{3/2} \sqrt{I_1 I_2 I_3} \right) \right) \quad (5.6)$$

$$S_{\text{vib}} = R \sum_{i=1}^{3N-6} \left( \frac{h\nu_i}{k_B T(e^{h\nu_i/k_B T} - 1)} - \ln(1 - e^{-h\nu_i/k_B T}) \right) \quad (5.7)$$

Here  $R$  is the universal gas constant,  $T$  the temperature,  $k_B$  the Boltzmann constant,  $N_A$  is Avogadro's constant,  $V$  the volume of a virtual container,  $M$  the mass of the molecule,  $\sigma$  the symmetry constant,  $I_1$ ,



$I_2, I_3$  the moments of inertia in the principle axis system,  $N$  the number of atoms,  $h$  the Planck constant and  $\nu_i$  the harmonic frequency of the  $i$ -th normal mode. The vibrational contributions are determined from second derivatives of the electronic energy on the PES, as the diagonalisation of the mass-weighted Hessian leads to frequencies and normal-modes as eigenvalues and eigenvectors, respectively. Low frequency modes can have a great impact on  $S_{\text{vib}}$  and thus lead to significant errors in relative barriers. An ubiquitous example is the internal rotation of methyl groups which generally results in a low frequency mode. Due to the rotational nature of the mode, other models have been suggested to increase the accuracy of the description of motion and the resulting partition functions.<sup>[189,190]</sup> Other, simpler solutions to circumvent the influence of these vibrations are substituting lower values with a certain threshold or neglecting low frequencies altogether. In the ORCA suite of programs vibrational frequencies of  $35 \text{ cm}^{-1}$  or lower are neglected in the calculation of the thermodynamic corrections. This approach is therefore also used in the determination of the free enthalpies reported in this work.

The above procedure only results in a free enthalpy  $G_{i,(g)}^\circ$  for one structure  $i$  in gas phase, whereas reaction energies in solution are needed. A comprehensive treatment would take all important conformations in solution into account in a Boltzmann weighted fashion.<sup>[60]</sup>

$$G_{(\text{solv})}^\circ = -RT \ln \left( \sum_i e^{G_{i,(\text{solv})}^\circ/RT} \right) \quad (5.8)$$

Since only low energy conformers play a role in the weight, well defined reaction pathways in a rigid environment can often be described by a single conformation and thus the minimum energy path. The free enthalpy in solution  $G_{(\text{solv})}^\circ$  can be decomposed into the state in gas phase  $G_{i,(g)}^\circ$  and a correction  $\Delta_{\text{solv}} G_i^\circ$  for transferring 1 atm of substance to a 1 M state in solution.

$$G_{i,(\text{solv})}^\circ = G_{i,(g)}^\circ + \Delta_{\text{solv}} G_i^\circ \quad (5.9)$$

This separation of contributions enables the use of a thermodynamic cycle to calculate reaction energies as depicted in Figure 5.5. The free enthalpy of reaction in solution  $\Delta_{\text{r}} G_{298,\text{aq}}^\circ$  is thus linked to the

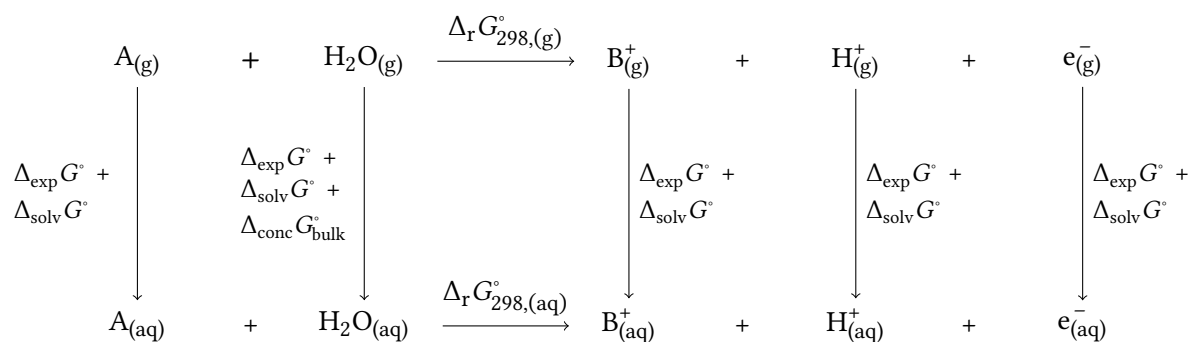


Figure 5.5: Thermodynamic cycle for an arbitrary proton coupled electron transfer reaction. The thermodynamic contributions of the electron are often neglected since they cancel by using appropriate references.

one in gas-phase. Due to the difference in standard state between gas-phase and solution the change of expanding an ideal gas from 1 atm to a concentration of 1 M,  $\Delta_{\text{exp}} G^\circ \approx 8 \text{ kJ mol}^{-1}$ , has to be added in addition to the solvation correction. In the case of bulk water the standard concentration is 55.6 M

which requires a further correction  $\Delta_{\text{conc}}G_{\text{bulk}}^{\circ} \approx 18 \text{ kJ mol}^{-1}$ . These thermodynamic cycles are often used to calculate  $\text{pK}_{\text{a}}$ s and oxidation or reduction potentials.<sup>[60,191]</sup> However, accurate results can only be achieved by a careful referencing, as mixing different standards can lead to several kJ/mol deviation.<sup>[60]</sup> This is due to different conventions in use for the proton and electron which influence the calculated result and reference potentials. Additionally, calculated free enthalpies are referenced relative to the free particles, whereas tabulated values refer to the elements in their standard state. In unbalanced reactions this change in reference state leads to an error in the calculated values compared to experiment. These variabilities in conventions can be made irrelevant if sensible reference reactions are included in the computational protocol. An effective reduction in the average error was thus shown for redox potential calculations using this internal referencing.<sup>[61]</sup>

## 5.2 Computational Investigations on a novel dinuclear Ruthenium Catalyst

### 5.2.1 Experimental Basis<sup>[13]</sup>

$[\text{Ru}_2(\mu\text{-Mebbp})(\text{py-SO}_3)_4(\text{H}_2\text{O})_2]$  (4) exhibits a Lewis-structure similar to 3, with an extended ligand backbone and consequently more flexible  $\text{py-SO}_3$  as axial ligands. It was already found experimentally by crystallisation of a mimic of 4, but without the sulfonates (5), that the structural change leads to a significantly different geometry in the active site.<sup>[13]</sup> Representations of the crystal structures are provided in Figure 5.6 which show the almost planar arrangement of 5 and the large Ru-N-N-Ru torsion in 3. The latter leads to an enlargement of the Ru-Ru distance by about 0.2 Å compared to 5.

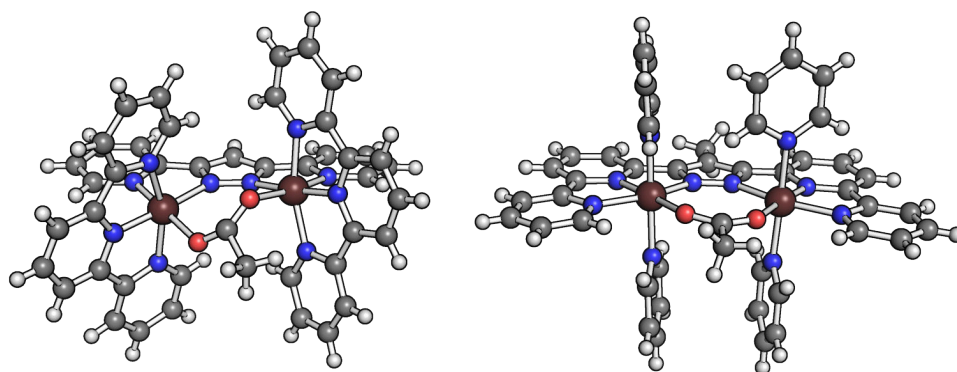


Figure 5.6: Crystal structures with bound acetate for 3 (left) and 5 (right), taken from the respective publications.<sup>[13,14]</sup> The tilting of the acetate ligand reflects the Ru-N-N-Ru torsion which varies between the two complexes.

This difference in the active site geometry is considered to be responsible for the variation in lability of the coordinating acetate and the choice of reaction pathway during catalysis.<sup>[13]</sup> Modification of the axial pyridines was found to have a significant effect on the activity<sup>[173]</sup> and the potentials<sup>[172]</sup> of the catalyst depending on whether the functional groups exert electron withdrawing or donating effects. Similarly to studies<sup>[192]</sup> on 3 the possible occurrence of an additional oxidation,  $[4,4] \rightarrow [4,5]$ , was proposed.<sup>[13,172]</sup>

### 5.2.2 A Reasonable Model System for $[\text{Ru}_2(\mu\text{-Mebbp})(\text{py-SO}_3)_4(\text{H}_2\text{O})_2]$

The rather similar performance<sup>[173]</sup> of  $[\text{Ru}_2(\mu\text{-Mebbp})(\text{py-SO}_3)_4]$  (4) and  $[\text{Ru}_2(\mu\text{-Mebbp})(\text{py})_4]$  (5) suggests that a mechanistic investigation of both catalysts should give the same results. Due to the high number of electrons and the charge introduced by the sulfonate groups a smaller model complex would be desired for the computational investigation. Since 5 is expected to lead to the same conclusions as 4 it can be directly compared to the  $[\text{Ru}_2(\text{Hbpb})(\text{tpy})_2]$  catalyst owing to the similar local structures around the transition metals.

Due to the high activity of 4 and other water oxidation catalysts, few is known about the complex in its active state despite the prevalent mechanism. Even the charge at which catalysis commences is difficult to assess, because the minimum formal oxidation state of the two metal centres needed for oxygen evolution would be [4,4], but a further oxidation might be hidden within the catalytic wave.<sup>[13]</sup> Additionally  $[\text{Ru}^{\text{IV}}]$  retains  $4e^-$  in its d-manifold, leaving at least two spin-states (singlet and triplet) for one ion in low-spin configuration of an octahedral ligand-field splitting. Since 4 is dinuclear this number is increased to at least three states: singlet, triplet and quintet. Additionally antiferromagnetic coupling has to be taken into account, which is not well described by a single determinant. Thus broken-symmetry approaches should be used to circumvent this problem,<sup>[39,40,193]</sup> which adds another two spin degrees of freedom to consider. On top of that the preferred spin-states depend on the functional<sup>[194–196]</sup> and an extensive calibration by CASSCF, as in the case of 3,<sup>[14]</sup> was not feasible.

One option to short-cut the calibration step with wavefunction approaches is basing this study on the methods as calibrated and tested by Cramer and co-workers on 3.<sup>[14]</sup> In the case of 3 the mechanism was scanned at the M06-L/6-31G(d) level<sup>[31,197]</sup> with the Stuttgart-Dresden pseudopotential and valence basis set SDD (ECP28MWB) for ruthenium<sup>[139]</sup> and refined by single point calculations using M06-L/6-311+G(2df,p)/SDD. Solvent effects were treated in the framework of the SMD model,<sup>[52]</sup> although it was also noted that the inclusion of explicit water molecules was necessary.<sup>[14]</sup>

### 5.2.3 Assessing the Validity of the Model Catalyst

With the clear experimental assignments of the mechanistic pathways for the Ru-Hbpb (3) and Ru-Mebbp-SO<sub>3</sub> (4) catalysts and the similarity of 3 and 5 the usefulness of the model system can be asserted. Calculations on 5 were thus carried out using M06-L and def2-SVP with the associated effective core potential. The size of def2-SVP is similar to 6-31G(d), but of a balanced and economical construction so that a better ratio of accuracy to basis function and a reliable convergence to the basis set limit can be expected.<sup>[198,199]</sup>

Probing the stationary points along the first reaction steps of 5, the active catalyst state was found to reside in a quintet state. The electronic coupling could be described as weak since the broken-symmetry state was within a few kJ/mol of the ferromagnetically coupled solution. In effect any of the two spin-states could have been chosen for the investigation, because the energies of reaction were found to differ marginally. Using the antiferromagnetically coupled state would have the added benefit of avoiding the problem with switching between different spin-state surfaces. Nevertheless, the ferromagnetically coupled multiplicities are easier to converge and were consequently used instead, where the difference between the two configurations varied only marginally. In the further reaction of WNA and I2M a covalent oxygen-oxygen bond is formed which suggests a redistribution of electrons in the system which are formally assigned to the metal centres. As such the multiplicity for WT1, W1, IT1 and I1

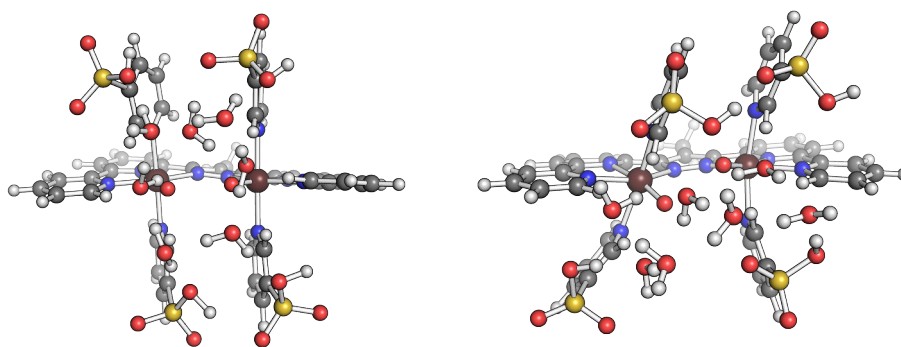


Figure 5.7: 4 with seven water molecules to model the non-bulk effects during catalysis. Left: starting structure of the optimisation with M06-L/SVP. Right: converged structure.

are expected to be lower than for the active catalyst. Indeed a similar situation is found for the other states with a triplet-derived multiplicity.

Notwithstanding, on the level of theory employed any attempt to converge the WNA transition state, **WT1**, was unsuccessful. Several effects could play a role in this observation, from unsuitable parameters of the optimisation to a non-stationary character of the target geometry within this electronic structure method. **WT1** requires a water molecule in the reaction coordinate which is likely involved in several hydrogen bonds before the reaction. The accurate modeling of the first solvation shell may therefore be an important factor for this transition state. As such 4 with seven explicit solvent molecules was successively investigated to probe the effect of the pyridine-bound sulfonate groups and the emerging cage-like water clusters. Calculations on this larger model system for [4,4] exhibited further problems, with the optimisation of **AC** changing the structure from a chemically sensible to a significantly deformed geometry as depicted in Figure 5.7.

An answer to this surprising behaviour may be found in the literature. M06-L was designed to be used for thermochemistry and noncovalent interactions in main group and transition metal compounds.<sup>[31]</sup> However, like many modern functionals it contains a large number of empirically fitted parameters—in the case of M06-L it is 34. Although one cannot determine the accuracy or reliability of a functional based on the number of its parameters, the approach of fitting to certain properties of data sets is an ongoing debate.<sup>[200–202]</sup> Since the basis of DFT for determining the energy of a system is its density, a better approximation to the true functional is expected to also reproduce the latter correctly. It has been recently argued that this may not hold in some newer functionals due to massive parameterisation.<sup>[203,204]</sup> Whether these deductions actually transfer to chemical problems like geometries and reaction barriers of multi-atom systems is still being debated at the time of writing.<sup>[205–207]</sup> On the same note the M06 range of functionals also exhibits an additional problem due to their specific parametric form as Wheeler and Houk showed.<sup>[208]</sup> For dispersion controlled dimers it was furthermore found that some meta-GGAs in combination with generally used grids can lead to optimisation problems and even additional minima on the PES.<sup>[200,208,209]</sup>

In these cases it is often better to stay with older, but thoroughly tested functionals. One of those is BP86 which has been found in similar applications to give geometries in good agreement with experiment.<sup>[210]</sup> Not surprisingly, changing to this functional led to more reasonable relaxations and converged structures that resemble the crystal data. To arrive at a converged description of the system and to set up a reference against which other model calculations can be compared, the larger model

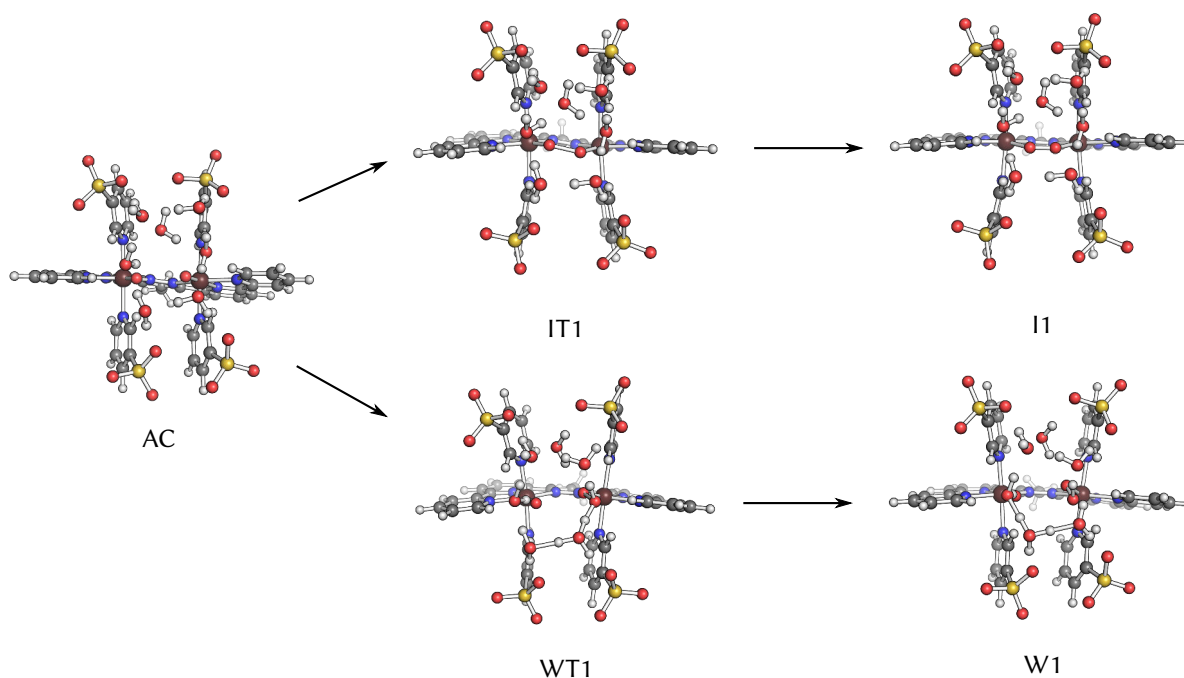


Figure 5.8: Stationary state geometries for the first steps of WNA and I2M as determined for 4 with 7 H<sub>2</sub>O on the formal oxidation state [4,4].

was again evaluated with BP86/def2-SVP used in the optimisation. Despite the change in functional the spin-state ordering did not change significantly. The geometries for the first transition states as determined are depicted in Figure 5.8.

#### 5.2.4 Comparison of Mechanisms

Because of the erratic behaviour in the larger model system the results of the reference functional M06-L seemed inappropriate for 4, whereas BP86 was shown to give reasonable structures even for the sulfonated catalyst. However, the size of the system prohibits an extensive analysis and the inclusion of several explicit water molecules could require the sampling over a multitude of geometries, thus raising the computational demand even further. With an increased confidence in the functional used to determine the geometries, a more thorough look into the smaller model system 5 is reasonable. Due to the effect of directed interactions of water molecules with the complex and the necessity of including it for the WNA pathway, a single H<sub>2</sub>O was added to all stationary points. Using just one water molecule in the transition states may also simplify the optimisation, although some could need more explicit solvent molecules to stabilise the structure energetically.<sup>[14]</sup> An additional benefit of retaining the same amount of molecules in every step of the reaction is a less biased view of the barrier heights. Often solvent molecules to stabilise a stationary state are only added when needed and the reaction is then treated as separated educts with the corresponding thermodynamic corrections.<sup>[14,192]</sup> Since the latter are mostly determined in the gas-phase a large penalty due to the loss of translational degrees of freedom is obtained in multimolecular reactions, which may reach a sizable margin. Calculations on this system were therefore expected to result in better accuracy.

The functional BP86 is known to not perform well for reaction barriers and hybrid functionals should

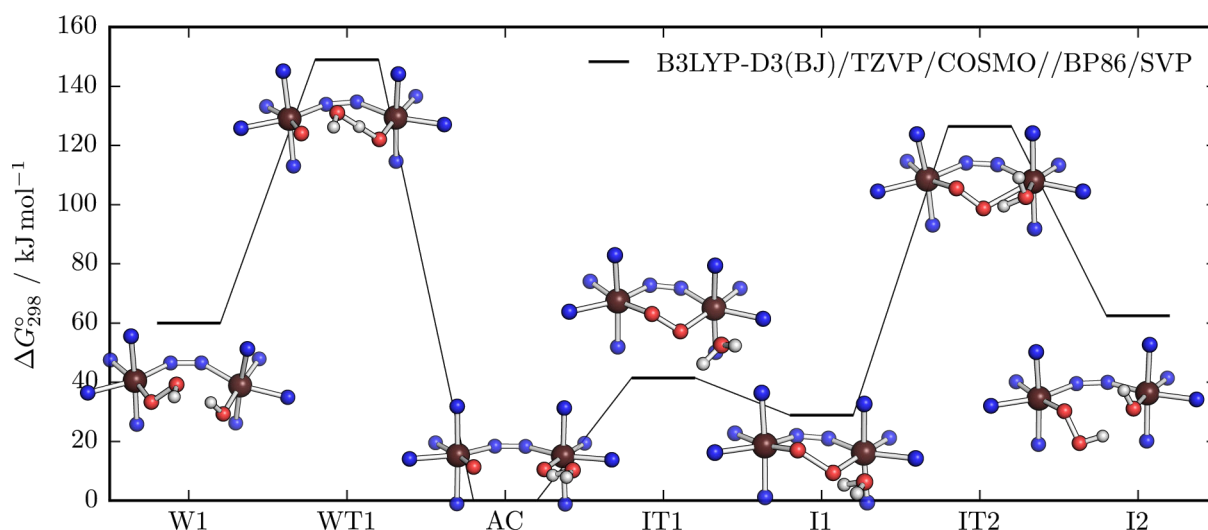


Figure 5.9: Full reaction path up to the hydroperoxide intermediate in oxidation state [4,4]. Geometries and thermodynamic corrections were determined using BP86/def2-SVP with single-point corrections for the electronic energy (B3LYP), dispersion (D3 with Becke-Johnson damping) and bulk solvent effects (COSMO for water).

be used to refine the electronic energies if possible.<sup>[183,211]</sup> B3LYP has been applied extensively to calculate reactions of organic molecules and was also successfully applied with transition metal compounds involved, but dispersion contributions should be corrected for.<sup>[212]</sup> As such the energetic profile has been calculated at the B3LYP-D3/TZVP/COSMO//BP86/SVP level of theory for the oxidation state [4,4]. The corresponding energy profile and the ball and stick representation of the stationary structures is shown in Figure 5.9. It is clearly visible that the IT2 has a significantly larger barrier than IT1 and is therefore likely the mechanism determining state for the i-I2M pathway. This may have been expected since the reaction intermediates W1 and I2 exhibit a very similar structure and thus energy. Both steps create and break a bond and are likely dependent on the solvent environment to stabilise the transition state. The process AC  $\rightarrow$  I2 could therefore be regarded as a two-step equivalent of AC  $\rightarrow$  W1. Nevertheless the bonds broken and created are quite different, as in IT2 a dative bond would be expected to emerge due to the coordination of a water molecule to the ruthenium site, but a covalent bond is created in WT1. A strong influence of the corrections to the inner energy would therefore be expected.

Since thermodynamic corrections are determined at the optimisation level of theory, the effect of the functional may be evaluated by comparing the single point results at a higher level of theory. Four different functionals were thus used on the converged structures with the TZVP basis set: BP86, M06-L, B3LYP and B3LYP\*. The first combination gives a hint on the basis set convergence, since the geometries were optimised using BP86/SVP. M06-L failed previously, but employing it for single point calculations may still give reasonable results for less CPU time than hybrid functionals. The latter has been added with B3LYP, essentially showing the dependence on the exact exchange in comparison to B3LYP\* which contains 5% less Hartree-Fock exchange. These values are assembled in Figure 5.10. Although many more functionals could be tested the selection in the figure should be representative of the main rungs used. Double hybrids with corrections based on perturbation theory could have been added for setting up a reference, but the increase in cost is mostly unjustified for this application.<sup>[213,214]</sup> And even with this small set of methods one can see that both transition states are varying in relative energy, although the more elaborate methods give a similarly inconclusive picture. Interestingly the barrier

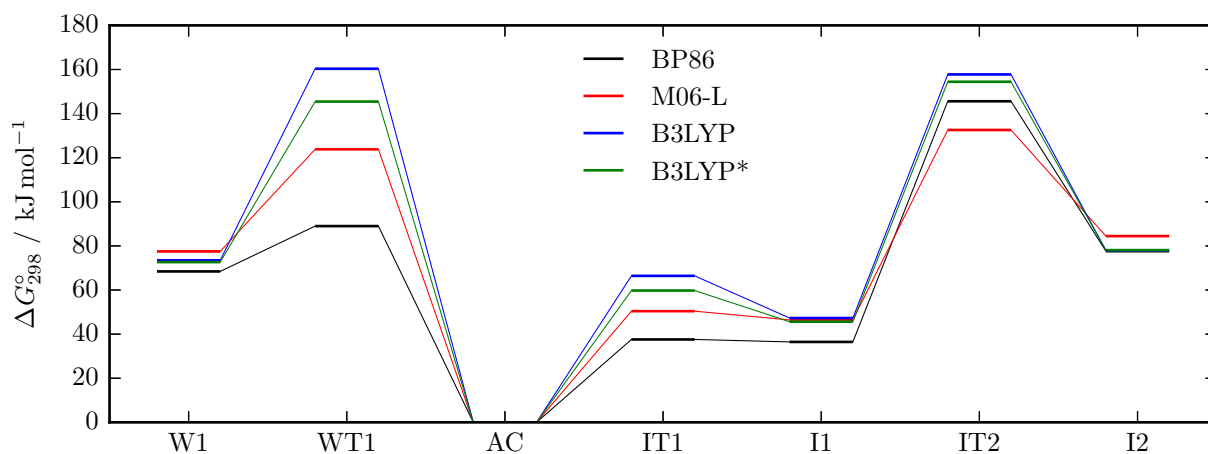


Figure 5.10: Comparison of barrier heights due to a change in the functional in oxidation state [4,4]. The same level of theory as in Figure 5.9 was used otherwise, without the dispersion corrections to keep the focus on the nature of the functionals.

height as retrieved by BP86 agrees well with the experimental selectivity. Even though it might be correct, this result is more likely to be accidental. All the other functionals result in a relatively similar barrier height for both pathways and are generally more accurate for reaction energies.<sup>[183,215,216]</sup> The apparent accuracy of BP86 in this case should therefore not be trusted without any further comparison to experimental values. Effectively the [4,4] oxidation state does not give a conclusive picture about the reactivity of 4. It is thus quite likely that the reason for the catalyst's selectivity lies somewhere else.

Despite the failure to describe the catalyst's reactivity in the [4,4] state an additional oxidation right before the catalysis commences can be envisioned and was deemed likely by experiment.<sup>[13,172]</sup> In the case of 5 the preferred stationary structure in the [4,5] oxidation state is expected to be equivalent to a simple one electron removal, without any further atom transfer. Thus the only additional degree of freedom is the choice of multiplicity. In contrast to the multitude of spin states for the preceding oxidation state the loss of an electron reduces the number of sensible multiplicities to a doublet, quartet and possibly a broken-symmetry doublet. As in the case of  $[\text{Ru}^{\text{IV}}\text{Ru}^{\text{IV}}]$  the latter takes into account the possibility of constructive interaction between the spins of the two metal sites. Nevertheless, the most stable PES was found to be the quartet throughout with similarly narrow broken-symmetry states as in the lower oxidation state.

If the additional oxidation of 5 was the reason for its WNA selectivity, at least two computed results would support this notion. One is the energy needed for oxidation of the catalyst which has to be lower than the actual potential of the oxidant, in this case  $\text{Ce}^{\text{IV}}$ . Further support could be gained if the majority of tested functionals lower the WT1 in relation to the IT2 barrier, since similar results were also obtained for the reduced state. The mechanism determining barriers were thus calculated for different popular functionals at the [4,5] state as shown in Figure 5.11. From these results several effects can be extracted: The variation of barrier heights with the functional is larger for IT2 than WT1 at [4,5]. This may be due to the different bonding situations and charge separation in those transition states. Another effect represented in Figure 5.11 is the rung of the functionals. More precisely, BP86 and PBE as two famous examples of GGA functionals give very similar barriers as do the (meta-)hybrid functionals B3LYP and TPSSh. In contrast the results of the M06-type functionals are neither consistent



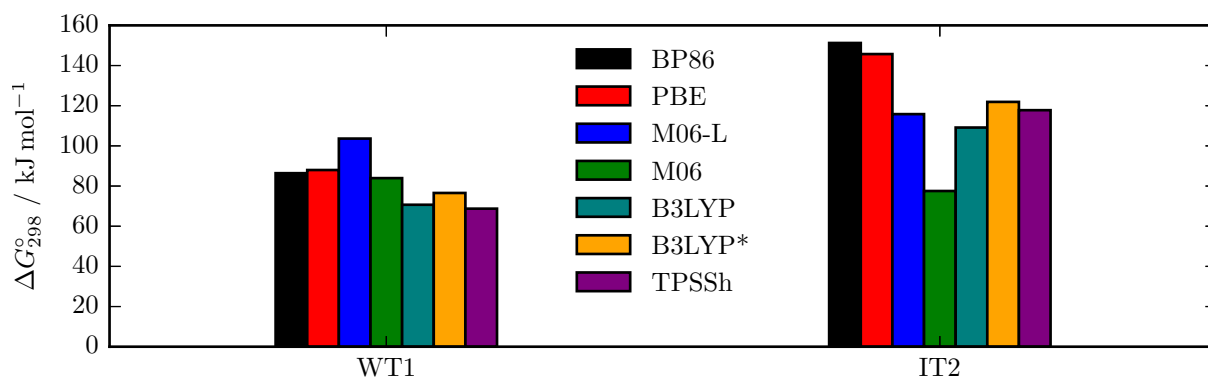


Figure 5.11: Barrier heights of IT2 and WT1 in the [4,5] oxidation state for different approximations X/TZVP/COSMO//BP86/SVP where X is the functional listed in the plot.

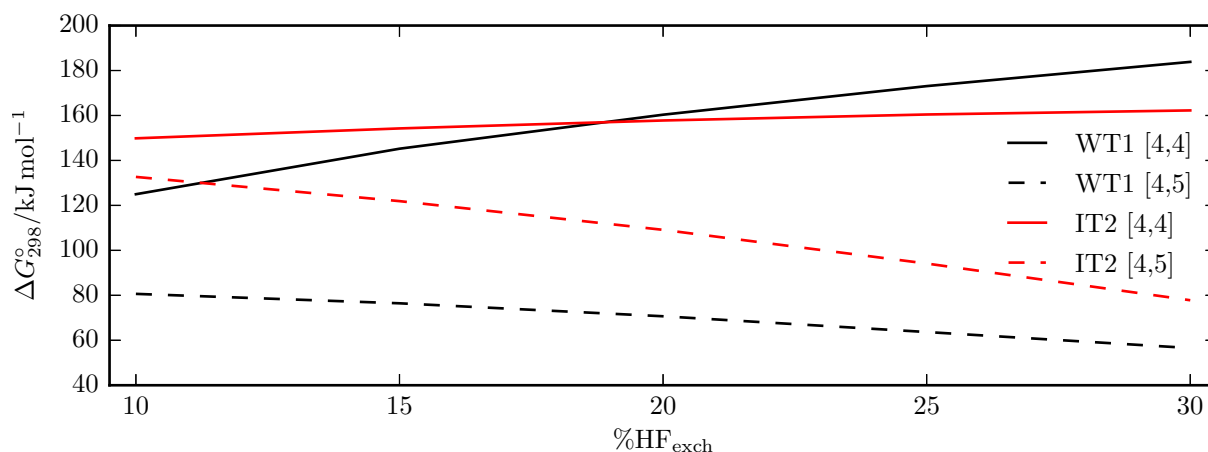


Figure 5.12: Comparison of IT2 and WT1 barrier heights for both oxidation states [4,4] and [4,5]. Energies are reported using B3LYP/TZVP/COSMO//BP86/SVP with differing amounts of exact exchange (%HF<sub>exch</sub>).

to each other, nor to their respective rung. Additionally the results returned by the hybrid functionals depend on the amount of exact exchange which is most notably seen by comparing B3LYP and B3LYP\*. The last effect is only of significance if it also changes the relative difference of the barriers. To validate this notion the exact exchange parameter of the B3LYP functional was thus modified in a small range and used to calculate both energy differences again. As shown in Figure 5.12 the exchange parameter significantly modulates the barriers and can even invert the preference for the mechanism. Reducing %HF<sub>exch</sub> would change the preference for one pathway over the other in the case of oxidation state [4,5]. For [4,4] the ordering does not swap in the range investigated, but the difference between the barriers still changes strongly. Intriguingly the change is in the opposite direction between both oxidation states and consequently a lower amount of exact exchange supports the current reasoning in both states. This by itself does not justify an adjustment of the parameters would it not be for several other investigations with similar preference for lower percentage of exact exchange.<sup>[217,218]</sup> Combined with the extensive analysis of the reduced state and the comparison to experiment the necessity of an oxidation to [4,5] to explain the reported selectivity is very likely.

Notwithstanding this significant insight into the active state of the catalyst several questions which are important to the development of more efficient complexes remain: what is the difference between 4 and 3 that leads to such a significant change in mechanism? How can it be rationalised to predict the



Catalyst	WNA				i-I2M			
	Int1	TS1	Start	TS1	Int1	ITS2	Int2	TS3
Ru-Hbpp <sup>[14]</sup>	W1	WT1	Min1	TS1	Min2	TS2	Min3	TS3
Ru-Mebbp	W1	WT1	AC	IT1	I1	IT2	I2	IT3
Ru-cbim <sup>[15]</sup>	int2dp	ts2dp	4dp	ts1dp	int1dp/int3	ts3	int4/int4dp	ts4

Table 5.1: Translation table for several names of stationary points used by different groups.

mechanism? Is it possible to tune the activity by modulating the preference for one mechanism? How does 4 fare against other dinuclear ruthenium catalysts which show I2M reactivity? A complete answer to all of these question cannot be given in a single study due to the dimensionality of each problem. As a first step toward an explanation of the reactivity, 5 was compared to the reference 3 and another recently computationally investigated catalyst (6).

### 5.3 Comparison of Water Oxidation Catalysts

Proper comparison of barrier heights and properties in DFT can only be reliably made for a similar combination of approximations—be it functionals, basis sets or further corrections like for solvent or dispersion. The mechanism of Ru-cbim had been evaluated<sup>[15]</sup> at the B3LYP\*-D2/6-311+G(2df,2p)/SMD//B3LYP/6-31G(d,p) level of theory, which is quite similar to the method used in this study. However, this is only true as long as the geometry is only slightly influenced by the functional. In the case of Ru-Hbpp the difference in methodology is more striking as the calculations were based on the meta-GGA functional M06-L. More precisely the results reported employed a computational scheme using M06-L/6-311+G(2df,p)/SMD//M06-L/6-31G(d).<sup>[14]</sup> To unite these different approaches the stationary points were recalculated using the same combination of methods throughout. Due to the experience with and literature-known deficiencies of the M06 type functionals only BP86 and B3LYP or B3LYP\* were a sensible choice for the reoptimisation. The latter two are expected to provide similar geometries and resemble the method used on Ru-cbim.<sup>[15]</sup> In the same line the basis set, solvent model and dispersion correction approach from the previously determined scheme were used, as differences should be marginal.

For a better comparison of the methods' influence the reported free energies of Ru-Hbpp,<sup>[192]</sup> Ru-Mebbp and Ru-cbim<sup>[15]</sup> were then overlaid with the results of the new computational approach as shown in Figure 5.13. In the figure the original labels of each stationary point have been changed to the same nomenclature as used for Ru-Mebbp to achieve a more consistent comparison. As the additional oxidation of Ru-cbim results in different formal oxidation states than for the other catalysts, these differences are noted by an additional **ox** if needed for the discussion. The labels of each group for structures of similar character are assembled in Table 5.1. Figure 5.13c shows the deprotonated (dp) pathway of Ru-cbim, which was reported to be less stable in the active state than the singly protonated one.<sup>[15]</sup> However, the deprotonation was possible with an energetic penalty of just 1.5 kcal/mol and even became the preferred pathway in the later course of the mechanism.<sup>[15]</sup> The calculations were thus restricted to this pathway. All of the catalysts have been assumed to require the [4,5] oxidation state in the active state for their reactivity, as Ru-Hbpp and Ru-cbim have been suggested to be active in that state.

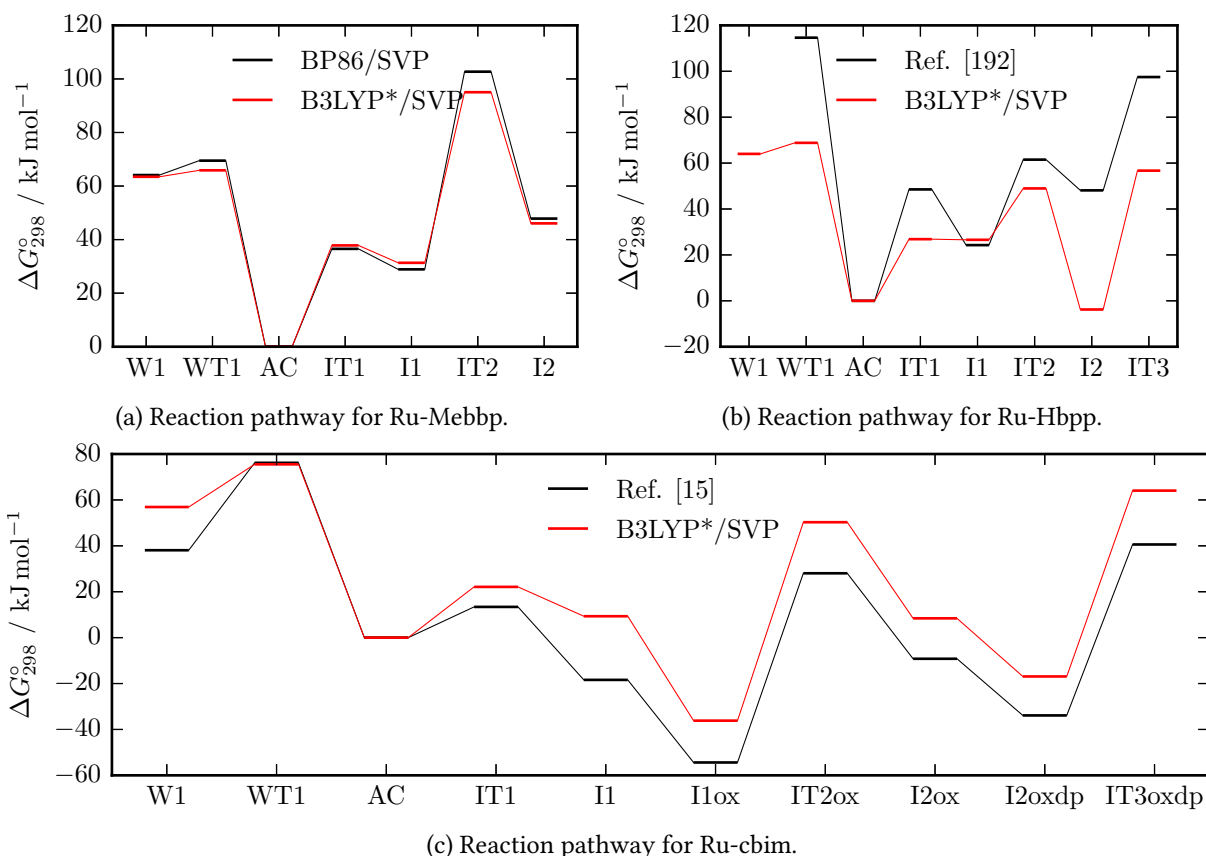


Figure 5.13: Comparison of reaction pathways for the [4,5] oxidation state calculated using B3LYP\*/SVP geometries with thermodynamic contributions and B3LYP\*-D3(BJ)/TZVP/COSMO corrections. a) Recalculated barriers for Ru-Hbpb compared to the results obtained in the group of Cramer and co-workers.<sup>[192]</sup> b) Results on the new level for Ru-Mebbp compared against values obtained by optimisation with BP86/SVP. c) The reaction profiles for the deprotonated states of Ru-cbim.<sup>[15]</sup>

The influence of using a hybrid functional like B3LYP\* instead of the GGA functional BP86 for the geometry relaxation is demonstrated by the results in Figure 5.13a, as the single point calculation was kept the same. BP86 was previously chosen for the geometry optimisation because it is known to return useful structures and thus the change in barrier height was expected to be small between structures determined by B3LYP\* or BP86. The largest deviation is found in IT2 of Figure 5.13a in which a water molecule reacts with the catalyst's active site. Nevertheless the barrier to WT1 is increased as well, which reduces the variation between the two transition states due to the optimisation method. The variation induced by the change in functional is thus insignificant for the current analysis.

A similar calculation for Ru-Hbpb, as shown in Figure 5.13a, and compared to the values reported in the literature<sup>[14]</sup> gives a much larger discrepancy. Except for the first i-I2M minimum I1 the differences in the relative energy vary between approximately 20–50 kJ mol<sup>-1</sup>. The disagreement between the two profiles is obvious, confirming a large impact from the used functional. In section 5.2.3 it was shown that the M06 range of functionals can exhibit erratic behaviour in optimisations. Additionally the barrier heights obtained with M06 and M06-L varied most strongly in comparison with other functionals, as shown in Figure 5.11. The literature values in Figure 5.13b were determined using M06-L for the optimisation and single point energies and consequently very different results to B3LYP\* can be expected. Since Ru-Hbpb is structurally similar to Ru-Mebbp B3LYP\* should work equally well to

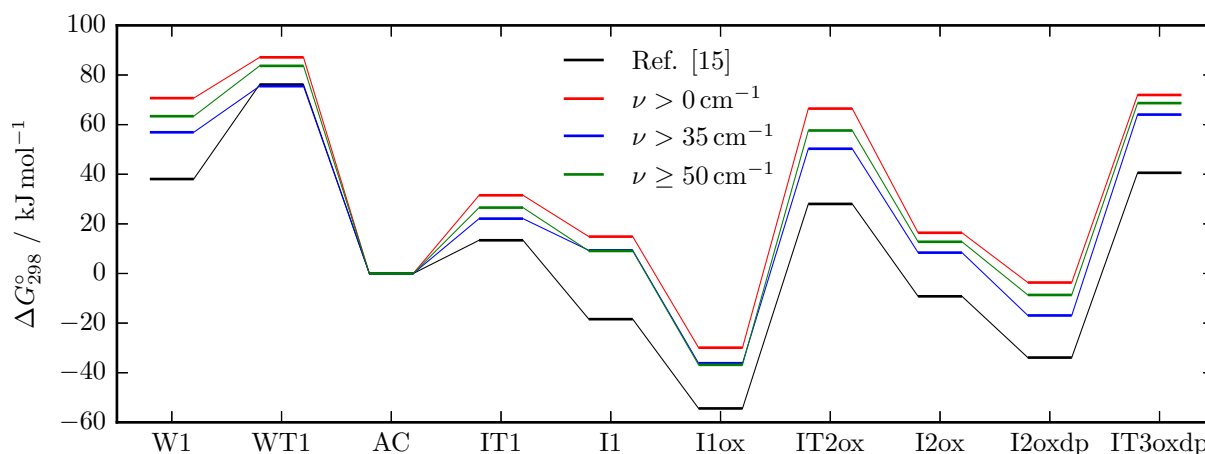


Figure 5.14: The influence of different approximations to low frequency vibrations on the barrier heights. Red: all frequencies are used, Blue: only frequencies above  $35.0 \text{ cm}^{-1}$ , Green: substituting low frequencies with  $50 \text{ cm}^{-1}$ , Black: reference values from the literature using methods as implemented in Gaussian.<sup>[15]</sup>

investigate the former.

Less deviation is expected for energy pathways of Ru-cbim in Figure 5.13c because the same functionals were used for the single point calculations. However, except for the **WT1** barrier the agreement of the two calculation schemes is only fair. From **I1ox** on the two i-I2M pathways exhibit a constant difference which hints at a subtle variation in the computational procedure. The choice of dispersion correction and continuum solvent model employed should not result in an error of that size as similar types were used in both cases. On the contrary thermodynamic contributions of multimolecular reactions in solution can be sensitive to the approximations used. For one, small frequencies of molecular vibration can have a large influence on the thermodynamic functions defined by the RRHO approach, which results in an error if the number of those frequencies changes during the reaction. To reduce this error small values are often neglected or substituted which may break the comparability to other calculations depending on the computational scheme. For another, continuum models include corrections to take the loss of translation into account which occurs due to the change from gas to solvent phase. This correction may overcompensate the effect in the case of multimolecular reactions and including the reacting species in every stationary point of the mechanism may therefore give more accurate results. The influence of these approximations on the barrier heights therefore needs to be assessed for a more consistent picture. To emphasise the influence of the low frequency treatment, several approaches were tested on Ru-cbim: all frequencies are included ( $\nu > 0 \text{ cm}^{-1}$ ), only those above  $35.0 \text{ cm}^{-1}$  ( $\nu > 35 \text{ cm}^{-1}$ ) as used in the ORCA suite<sup>[187]</sup> and frequencies below  $50 \text{ cm}^{-1}$  replaced with the latter ( $\nu \geq 50 \text{ cm}^{-1}$ ). As directly visible from Figure 5.14 the deviation can be large with up to  $20 \text{ kJ/mol}$ , depending on the state. The largest barriers thus result if all frequencies are included as calculated ( $\nu > 0 \text{ cm}^{-1}$ ), except for the transition mode. Neglecting the small frequencies ( $\nu > 35 \text{ cm}^{-1}$ ), which may also be more subject to noise in numerical frequency calculations, therefore leads to lower relative energies. In fact most states and barriers calculated using this approximation are more similar to the reference profile than the other approximations.  $\nu \geq 50 \text{ cm}^{-1}$  leads to values in between without a clear similarity to either of the other approaches. Neither of the three treatments leads to changes that would invalidate the finding of i-I2M reactivity, although the difference between **WT1** and **IT2ox** gets smaller than predicted by

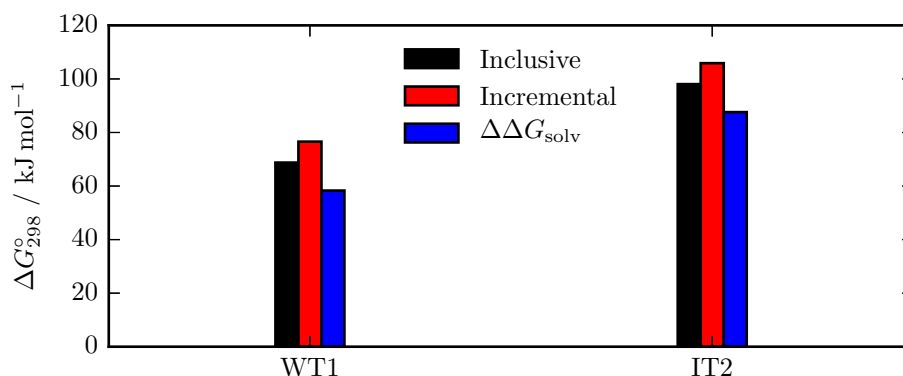


Figure 5.15: The mechanism determining barriers WT1/AC and IT2/AC as calculated using B3LYP\*-D3/TZVP/COSMO//B3LYP\*/SVP and three different approaches. Inclusive: the same number of atoms is retained in every state, Incremental: the required number of water molecules is added at the transition state when needed,  $\Delta\Delta G_{\text{solv}}$ : incremental version with an approximate correction for the bimolecular reaction in solution.

the reference calculations. Thus the qualitative agreement is retained irrespective of the approximate treatment of low frequency vibrations. Additionally, since the energy profile for  $\nu > 35 \text{ cm}^{-1}$  gives similar barriers for each step except IT1 and I1 the currently used computational scheme is corroborated by this investigation.

As for the different treatment of solvent effects in Figure 5.13, Ru-Hbpb and Ru-cbim have been calculated with a scheme that only includes explicit water molecules in a structure if required for the transition state or intermediate.<sup>[14,15,192]</sup> Specifically, Figure 5.13b includes two H<sub>2</sub>O molecules from IT2 on and three in IT3. Figure 5.13c on the other hand was determined using only one H<sub>2</sub>O starting from IT2ox and two at IT3oxdp. In contrast the values for Ru-Mebpb were calculated with the same number of water molecules and thus atoms in every stationary point, which was expected to reduce the uncertainty in the barriers due to the difference of water-complex, water-solvent interactions and thermodynamic corrections. Although the correctness of this assumption could not be verified during the development of the computational scheme, the difference in including or neglecting water molecules in certain states influences the obtained results. Indeed it was suggested that continuum models be released of the requirement to correct for the change in entropy encompassed by moving a molecule from gas-phase to solution.<sup>[219]</sup> This is mostly a problem of multimolecular reactions, for which an approximate correction  $\Delta\Delta G_{\text{solv}}$  was derived.<sup>[219]</sup>

To assess the magnitude of this influence, the energy profile for Ru-Mebpb was additionally evaluated without H<sub>2</sub>O at the active state. The Gibbs free enthalpy can then be calculated by the scheme used for Ru-Hbpb and Ru-cbim, or additionally corrected using  $\Delta\Delta G_{\text{solv}}$ . Figure 5.15 shows the results for these approximations, where the values *incremental* were acquired by adding H<sub>2</sub>O only to the transition state. In contrast *inclusive* also contained a water molecule in intermediate AC. As expected the choice of approach for bimolecular reactions in solution can have a pronounced effect on the barrier height of about 20 kJ mol<sup>-1</sup>, whereas the addition of  $\Delta\Delta G_{\text{solv}}$  to the incremental scheme reduces the activation energies and thus is in the same direction as the inclusive scheme. Since only AC was changed for this test the relative energy between the transition states is not modulated by the different approaches. In a more extended microsolvation environment the effect on the transition states should be different, in which case a consistent description with the same atoms in every calculation is of advantage. This

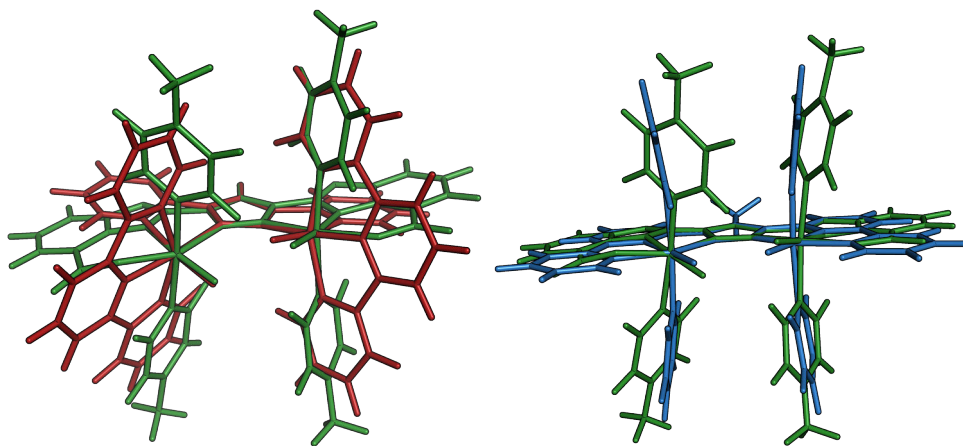


Figure 5.16: Overlays of active catalyst state geometries in the [4,5] oxidation state for Ru-cbim (green) with Ru-Hbpbp (red) and with Ru-Mebbp (blue).

result suggests that the inclusive scheme is a more consistent model for the reactivity of catalysts like Ru-Mebbp.

### 5.3.1 Ru-cbim: Similar to Ru-Hbpbp or to Ru-Mebbp?

In the previous section it was shown that the current approach to calculating the free enthalpies is consistent with the proposed mechanisms for all three catalysts and despite remaining differences in barrier height, the results are still comparable to calculations of other groups. The Ru-Hbpbp and Ru-Mebbp complexes are already very similar so that the change in reactivity is likely subtle and a precise investigation is needed. However, in the case of Ru-cbim the carboxylate substituents lead to a lower effective charge in the active state and throughout the catalytic cycle. This lowered charge contributes to the reason for the suggested additional oxidation by  $[\text{Ru}(\text{bpy})_3]_3^+$  in the i-I2M pathway.<sup>[15]</sup> In the active state AC the Ru=O groups exhibit a large Ru–O–O–Ru torsion similar to Ru-Hbpbp as seen in Figure 5.16. The latter may be expected since it was suggested that this pre-arrangement is the cause of i-I2M reactivity.<sup>[14]</sup> Notwithstanding, the coordination sphere around the rhenium atoms looks rather similar to Ru-Mebbp otherwise.

Turning again to the energy profiles in Figure 5.13 one can easily see that the WNA barriers of both complexes have very similar values if calculated at the same level. This means the i-I2M reactivity of Ru-Hbpbp and Ru-cbim effectively results from a relative stabilisation of the complete I2M compared to the WNA pathway. For Ru-cbim this is supported by the additional oxidation after the O–O bond formation, so the question arises whether this still holds without removing the electron at I1. Indeed, comparing the barriers for the steps I1  $\rightarrow$  IT2 of Ru-Mebbp and Ru-cbim, fairly similar values are obtained despite the difference in oxidation state and composition.

Since the computational scheme used on Ru-Mebbp was found to qualitatively give the same results as the literature procedure for Ru-cbim, the second i-I2M transition state was recalculated without the additional oxidation to compare the effect. For insight on the influence of the oxidation on the mechanism only the relative energies of the transition state IT2 to the preceding intermediate of the respective oxidation state (I1/I1ox) or the active catalyst (AC) is important. The latter can directly be compared to the WNA transition state WT1 and determines the preferred mechanism. The effect of

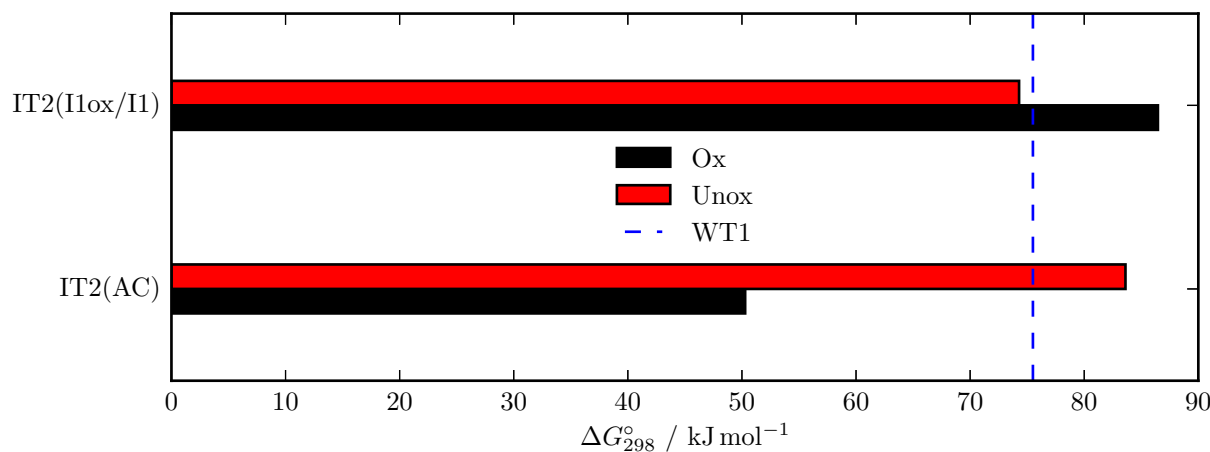


Figure 5.17: Selected calculated barriers for the deprotonated pathways of Ru-cbim. Black bars correspond to the procedure with further oxidation  $I1 \rightarrow I1ox$ , the red bars are retrieved if this is neglected. Top: comparison of energy differences between IT2/IT2ox and the preceding intermediate (I1 or I1ox). Bottom: energy difference between IT2/IT2ox and active catalyst state (AC).

the oxidation on the forward barrier of the step allows to deduce whether the higher oxidation state actually enhances the catalysis or if it is only the reason for the mechanistic pathway.

As seen in Figure 5.17 the relative energy of IT2 to AC strongly increases if the catalyst is taken to stay on the [4,5] oxidation state and actually surpasses WT1. Thus an i-I2M reactivity would not be assigned on that level and the relative energies between IT2 and WT1 actually resemble the results on Ru-Mebbp in the [4,4] state. Nevertheless, the barrier for the step to IT2 is increased by the oxidation and could well become the rate determining step of the catalytic cycle. Summing these results up the Ru-cbim catalyst seems to exhibit i-I2M reactivity only due to the possibility of another oxidation at the I1 intermediate which traps the complex in this pathway. Otherwise a mixed or even WNA reactivity would be expected which then would be more similar to Ru-Mebbp than Ru-Hbpb. This notion is supported by the more similar active site coordination geometry for Ru-cbim and Ru-Mebbp. The results suggest that the reactivity could be tuned by careful adjustment of the oxidation potential, which would make an interesting handle for the reactivity tuning of water oxidation catalysts, but is out of the scope of this study.

### 5.3.2 Indices to Predict Water Oxidation Mechanisms

In section 5.2.4 it was shown that a correct computational representation of the mechanistic insight gained through experiment was mostly dependent on the oxidation state. Nevertheless, this was only possible combined with the notion that the highest state in the mechanistic branch is actually defining the barrier for the pathway. The question then arises whether a specific state or state difference can be made out which explains the change in mechanism from Ru-Hbpb to Ru-Mebbp.

Comparing the calculated profiles in Figure 5.13b and Figure 5.13a, the relative energies for the first I2M and WNA states are very similar. A pronounced difference thus only starts with the second transition state IT2, but the reason for this could not yet be concluded. In all transition states except WT1 and IT1 water molecules react directly within the active site, that is they form bonds with the transition metals and consequently the catalyst has to rearrange, which is more strongly influenced by the solvent environment than reactions between oxygen atoms within a confined space. Notwithstanding,

WT1 also includes the reaction of a water molecule to form new bonds. This change in bonding and the difference to the interactions within the solvent environment could mean that the relative energies vary even more than for IT2. It is therefore not obvious why IT2 should be the mechanism determining transition state, since it does not seem to be mainly the interaction with the water molecule.

In the literature the integrated spin density on the Ru=O oxygen has been suggested to be an indicator of I2M reactivity.<sup>[167]</sup> If that was true also for dinuclear ruthenium catalysts the mechanism could be derived from a simple single-point calculation of the active state intermediate. The search for new catalysts with differing preferred pathways would then be largely simplified. Notwithstanding, the latter relation only holds if IT1 would be the mechanism determining state as the radical character is said to influence the creation of an oxygen–oxygen bond. In the following I2M states the peroxide is only displaced and Ru–O bonds are broken so that the radical character should not have much of an effect. Since IT1 is not the important stationary point for the mechanism of Ru-Hbpp and Ru-Mebbp a correlation seems unlikely. Still, the use of this reactivity index in the literature warrants an investigation.

The lowest unoccupied molecular orbital (LUMO) was used as an additional criterion to determine the likely mechanism, as an energy similar to the highest occupied molecular orbital (HOMO) of water would suggest a lower barrier for the WNA pathway.<sup>[167]</sup> Thus an indication for I2M reactivity would be expected from Ru=O oxygens with high spin density and a LUMO of high energy. The same work compares two mononuclear water oxidation catalysts that react by different mechanisms which is induced by a slight modification of the ligand. This modification changed the calculated integrated spin densities on the oxygen from 0.67 to 0.63 and the LUMO from  $-4.37$  to  $-4.52$  eV. It was proposed that this change led to the switch in mechanism from I2M to WNA. In the case of Ru-Hbpp and Ru-Mebbp the Löwdin integrated spin densities in the active state AC were found to be 0.54 and 0.57 respectively. This is in line with similar active sites of mononuclear complexes that were investigated, as it was suggested that a seven-coordinated metal would increase the integrated spin density strongly.<sup>[167]</sup> However, Ru-Hbpp and Ru-Mebbp exhibit the same change in mechanistic pathway with a similar coordination in the active state but only marginally different spin densities. Even the LUMOs differ only by about 0.05 eV with Ru-Mebbp having the higher energy. The conclusion for these two complexes using the spin density and LUMO as an index would thus be contrary to the actual reactivity. It is therefore evident that the approach suggested is not applicable to dinuclear ruthenium catalysts in general. Also it does not seem appropriate to derive any conclusion about reactivity by examining only one stationary state.

Nevertheless a kind of barrier–property or barrier–structure correlation is needed to further devise procedures to improve homogeneous catalysts. A successful pursuit of these structural indices requires a major influence of the pathways by that index. Since the most strongly influenced transition state is IT2 a structural modulation of the mechanism can likely be found by a detailed investigation of that state. Several primitive coordinates that could change significantly during the reaction step were thus compared between I1, IT2 and I2 for Ru-Hbpp, Ru-Mebbp and Ru-cbim. The employed coordinate nomenclature is schematically depicted in Figure 5.18 and the results assembled in Table 5.2. The difference in geometries at the transition state is clearly visible in the ball-and-sticks representations in Figure 5.19. From the table it is obvious that many of the coordinates are still similar in I1 for Ru-Hbpp and Ru-Mebbp which corroborates the assumption that both catalysts are structurally very similar. On



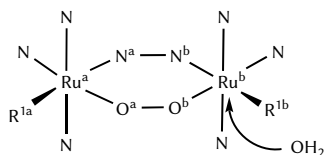


Figure 5.18: First coordination sphere of a general dinuclear ruthenium catalyst ( $R=N_O$ ) with the naming convention used for declaring the internal coordinates in Table 5.2.

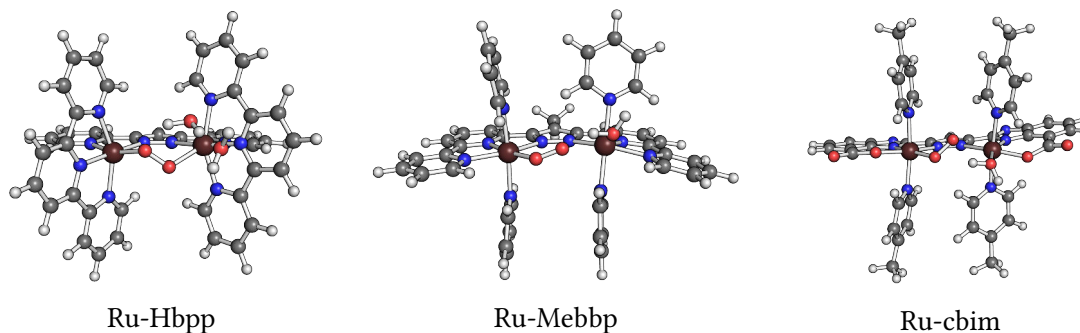


Figure 5.19: Ball and stick representation of IT2 determined using B3LYP\*/SVP for all three catalysts.

the other hand a significant deviation is notable in IT2, at which Ru-Mebbp exhibits a stronger deformation than Ru-Hbpbp as seen from the change in Ru–Ru distance and torsion of the two ruthenium sites. The values for  $r(\text{Ru}^a\text{--Ru}^b)$  further hint at an early transition state for Ru-Hbpbp and a late one for Ru-Mebbp. According to the Hammond postulate the transition state and following intermediate are expected to be higher for a late transition state compared to an earlier one. This is indeed the case for those two catalysts as the IT2 and I2 of Ru-Hbpbp are lower in relative energy than those of Ru-Mebbp.

Staying in the framework of the Hammond postulate an early transition state would be represented in a steeper deformation curve of the main coordinate that is changed. Since a water molecule has to be incorporated in the active site in IT2 the largest change in the geometry is expected to occur for the Ru–Ru distance and Ru–N–N–Ru dihedral as already deduced from Table 5.2. Both coordinates were therefore subjected to a relaxed surface scan in I1 to unravel the energy dependence of these coordinates. Because I1 is a minimum, few unwanted deformations in the range of investigation were expected. The resulting potentials are depicted in Figure 5.20 and clearly support the previous assumption that Ru-Hbpbp exhibits steeper potentials than Ru-Mebbp. Nevertheless these curves were determined at minimum structures and thus do not include the influence of the reacting water molecule.

Table 5.2: Significant coordinates that highlight structural differences between 3 and 5. All distances are given in Å, all angles in degrees. For the definition of the coordinates see Figure 5.18.

Coordinate	Ru-Mebbp (5)			Ru-Hbpbp (3)			Ru-cbim (6)		
	I1	IT2	I2	I1	IT2	I2	I1ox	IT2ox	I2ox
$r(\text{Ru}^a\text{--O}^a)$	2.11	1.95	1.95	1.98	1.98	1.95	1.91	1.96	1.93
$r(\text{Ru}^b\text{--O}^b)$	2.00	2.61	4.27	1.98	2.06	4.84	1.91	2.55	3.72
$r(\text{O}^a\text{--O}^b)$	1.31	1.28	1.27	1.34	1.33	1.30	1.41	1.32	1.32
$r(\text{Ru}^a\text{--Ru}^b)$	4.08	4.53	4.66	4.10	4.20	4.71	4.02	4.50	4.68
$\tau(\text{R}^{1a}\text{--Ru}^a\text{--Ru}^b\text{--R}^{1b})$	-6.1	35.9	14.5	0.7	-14.6	-7.8	53.0	33.9	17.6



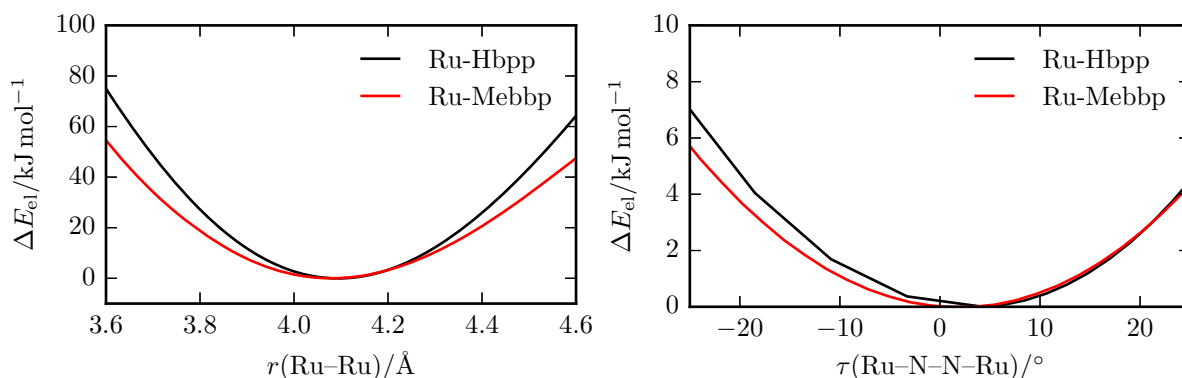


Figure 5.20: Electronic energy potentials retrieved by deforming a single internal coordinate of **3** and **5** at the stationary point **I1**. Left: contraction and stretching of the Ru–Ru-distance. Right: twist of the Ru–N–N–Ru dihedral out of equilibrium. The energies were determined with B3LYP\*/def2-SVP.

For a clearer insight into the structure dependence of the reaction step a more extensive investigation is needed.

A possible approach is to characterise the important contributions to the energy along the reaction coordinate  $\zeta$  from **I1** to **I2**. As a simple first approximation the total energy  $\Delta E(\zeta)$  can be decomposed into the energy of interaction,  $\Delta E_{\text{int}}(\zeta)$ , and the deformation energy  $\Delta E_{\text{strain}}(\zeta)$ . The interaction energy can be further separated into several contributions if energy decomposition schemes are applied. Comparing the separate contributions to the energy difference along the reaction pathway was already suggested in the *activation strain model* (ASM)<sup>[220]</sup> which allows to characterise the lateness of the transition states and to unravel the energetic reason for a trend in reactivity. As such ASM extends the previous analysis of the minimum potentials and may lead to a better understanding of the ligand's influence. Judging from the structures of the transition states (Figure 5.19) and the difference in the coordinates' changes (Table 5.2), a more pronounced influence of the deformation energy than the interaction contribution is expected. If this assumption was right it would be directly visible in the results from the ASM.

The minimum energy path was therefore determined by steepest descent and selected points were improved by single point calculations on the B3LYP\*-D3/TZVP/COSMO level as for the mechanistic investigations. Figure 5.21 supports the assignment of the geometrical position of the transition state for Ru-Hbpp and Ru-Mebbp. The curve for Ru-Mebbp additionally exhibits a more pronounced variation in the progression from the transition state towards the minima. At the beginning of the reaction path this is caused by the rotation of a pyridine ligand which couples to the path of the approaching water molecule. This required movement does not take place in the case of Ru-Hbpp as the axially coordinating ligands are fixed to the equatorial plane due to the terpyridine ligands. Consequently the total absolute path length of Ru-Hbpp is much shorter than the one of Ru-Mebbp in mass-weighted coordinates.

From the other charts in Figure 5.21 some hints as to the influence of the different ligand framework may be extracted. Interestingly the change in interaction energy diverges between Ru-Hbpp and Ru-Mebbp around half of the reaction coordinate value whereas the deformation energy difference does not increase monotonically. This leads to a larger influence of the interaction contribution over the whole pathway and one may argue that a small increase in the deformation in Ru-Hbpp creates the possibility of a better interaction with the water molecule. The same effect may be less pronounced for

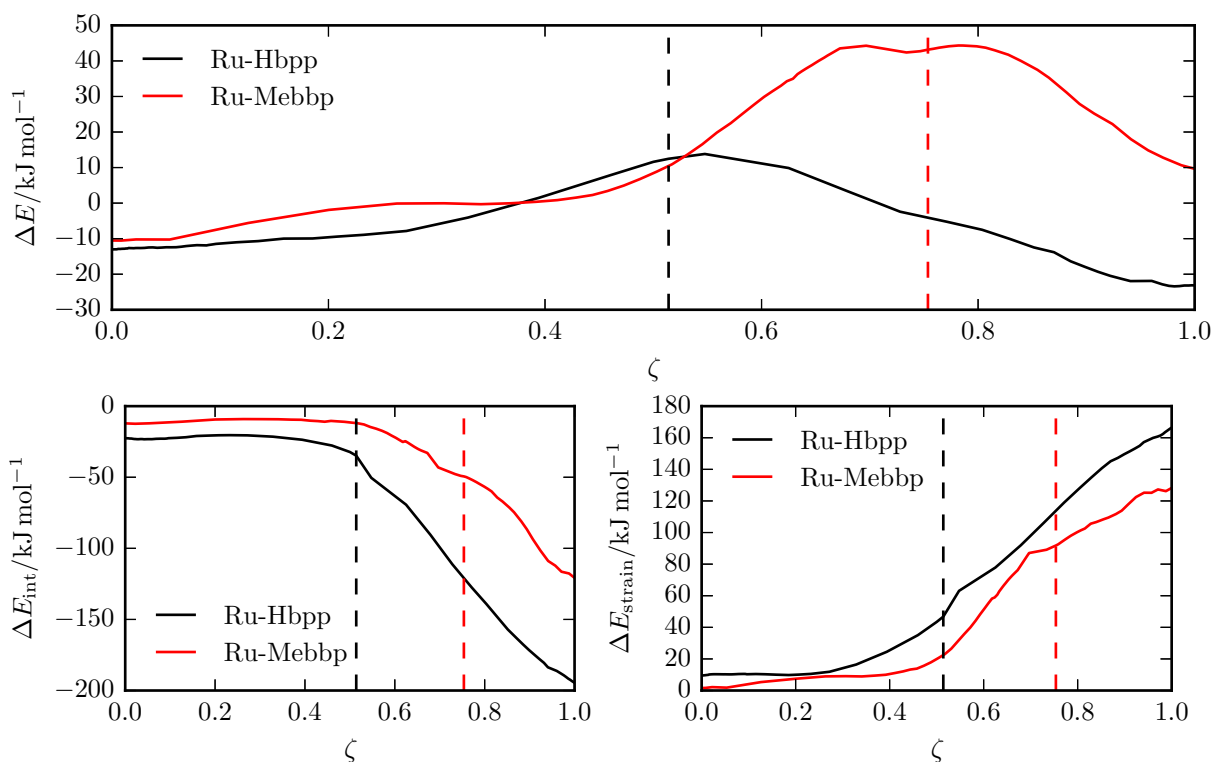


Figure 5.21: The ASM pathway along the reaction coordinate  $\zeta$  on the minimum energy pathway for Ru-Hbpp and Ru-Mebbp on the oxidation state [4,5]. The listed energies were calculated using B3LYP\*/D3/TZVP/COSMO//B3LYP\*/SVP. The position of the transition states as determined on the optimisation level are marked by dashed vertical lines.

Ru-Mebbp due to the independence of the axial ligands on the equatorial distortion. Comparing the values at the transition state coordinate the similarity in interaction energy is directly visible, whereas the deformation contribution differs greatly. Thus the interaction energy between water and the catalyst may actually not be the determining factor for the barrier height and thereupon selectivity, but the internal constraints induced by the ligands.

It was shown with the scan of the Ru–N–N–Ru torsion and Ru–Ru stretch in Figure 5.20 that the change in dihedral angle only has a negligible influence on the deformation energy as compared to the distance between the metal sites. A coordinate with a large contribution to the total barrier of the step is therefore expected to correlate with the Ru–Ru distance. This is also true for the torsion of course and since its change is much larger for Ru-Hbpp than for Ru-Mebbp the actual geometrical cause for the difference in barrier is not obvious from the investigated coordinates. Comparing the transition state structures in Figure 5.19 one might ask whether the axial ligands' structure is the actual reason for the mechanism. As the pyridyl ligand in IT2 of Ru-Mebbp has to rotate approximately  $90^\circ$  to arrive at the geometry of the transition state, the water molecule displacing the peroxide from the metal site approaches mostly from the front of the active site. The same can be said about Ru-cbim, even though the picoline ligands are not rotated as much. Nevertheless the structural parameters of Ru-cbim are very similar to Ru-Mebbp at the transition state as listed in Table 5.2. This suggests similar behaviour and although i-I2M reactivity was suggested for the mechanism it was shown above that this assignment relies on an additional oxidation step. Contrary to both catalysts Ru-Hbpp exhibits fixed orientations of the axial ligands and thus displays an angle between the tripyridines of approximately  $66^\circ$  in **I1** as

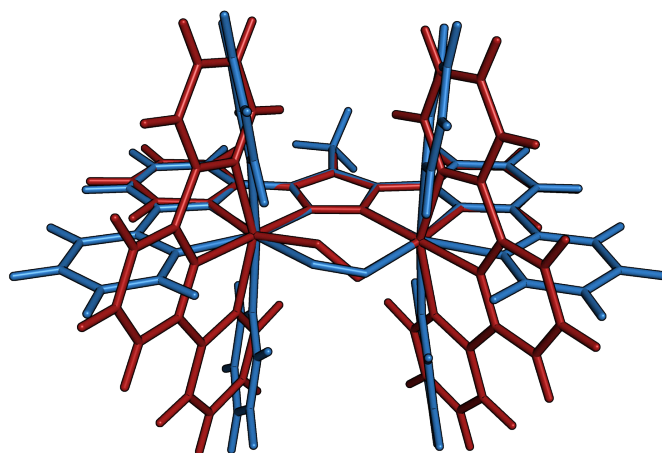


Figure 5.22: Overlay of Ru-Mebbp (blue) and Ru-Hbbp (red) in the intermediate state I1, as determined using B3LYP\*/SVP.

shown with the superposition in Figure 5.22. The water attack therefore should occur slightly diagonal to the active site which supports the torsional deformation of the complex to make space for the water molecule. Although this assumption cannot be supported quantitatively with the current calculations the acquired minimum energy paths support the reasoning qualitatively. An argument against this hypothesis may be that the pyridines of Ru-Mebbp are free to rotate and the water approach should therefore not be hindered by it. However, the argument does not hold in that case as the space opened by a 90° rotated pyridine does not seem to be sufficient to allow a diagonal attack and thus the catalyst has to deform in the equatorial plane to increase the Ru–Ru distance and to create space for the H<sub>2</sub>O. On top of that the rotational freedom is somewhat constrained in solution and even more so when the sulfonated groups of Ru-Mebbp-SO<sub>3</sub> are in interaction with water.

## 5.4 Conclusion

A novel dinuclear water oxidation catalyst was investigated computationally and its mechanistic pathway thoroughly characterised. The theoretical basis to treat the water nucleophilic attack and the interaction of two metal sites on equal footing and with a consistent method was developed. This enabled a concise analysis of the Ru-Mebbp (5) catalyst which was used as a model system for the active species in a water environment. It was shown that a WNA type reactivity of the catalyst can only occur in the [4,5] oxidation state as the reduced species exerts high barriers and a negligible selectivity in that state which was corroborated by application of a variety of different DFT functionals. Nevertheless the use of M06 range of functionals is discouraged as erratic behaviour already described in the literature,<sup>[200,208]</sup> was found for the larger sulfonate-containing catalyst. The proposed reduction of exact exchange in the B3LYP functional suggested in the literature was also in line with the predictions of reactivity for Ru-Mebbp, which further supported the robustness of the approach employed in this study. A comparison of computational schemes to literature values of different catalysts signified the importance of a consistent treatment of low vibrational frequencies in the thermodynamic corrections and the inclusion of solvent molecules in each stationary point.

To unravel the effects that determine the reactivity of water oxidation catalysts and thus also the mechanistic pathway, Ru-Mebbp was compared to the similar dinuclear catalysts Ru-Hbpb and Ru-cbim. The latter displays a comparable octahedral coordination environment with a much higher absolute charge, whereas Ru-Hbpb only exhibits a change in geometry so that both make excellent reference complexes.

Detailed examination of Ru-cbim showed a pronounced dependence of the mechanism on the oxidation state. Without a further electron abstraction at **I1**, WNA reactivity was found to be prominent, whereas the suggested i-I2M pathway seems to derive from the low oxidation potential at this intermediate, which prevents the back-reaction and eliminates the possibility of a water nucleophilic attack. The oxidation state dependence of the barrier heights actually resembled that of Ru-Mebbp in the oxidation states [4,4] and [4,5] and similarly the structural parameters in the most influenced transition state **IT2ox** corroborated the analogy between Ru-cbim and Ru-Mebbp.

Ru-Hbpb on the other hand is electronically close, but modifies the mechanistic pathway by a difference in ligand geometry which is expressed in an earlier transition state at **IT2** than for Ru-Mebbp. The difference in reactivity was attributed in the literature to the pre-arrangement of the Ru=O groups due to the ligand backbone, which was partly supported by the determination of the minimum energy path along the **I1** → **IT2** → **I2** coordinate. A focus on the relative oxyl orientation suggests that the bond creating reaction step is mechanism determining which was not found to be true for any of the three catalysts investigated. Additionally the geometry of Ru-Hbpb and Ru-Mebbp at **I1** exhibited similar values for significant primitive coordinates of the motion during reaction and thus the mechanism determining state **IT2** should not be influenced much. However, the deformation in **AC** of Ru-Hbpb hints at the repulsion of the terpyridyl ligands, which also change the accessibility of the active site for a water molecule in comparison to single pyridines. This effect was made responsible for the higher **IT2** barrier and relative position of the transition state. For further insight on the ligands' influence a stronger focus should thus be set on the axially coordinating ligands.

## 6 Summary

In this thesis several topics of importance for the larger field of artificial photosynthesis have been addressed. As the systems investigated usually consist of 50–100 atoms or more with at least one transition metal ion, density functional theory was the method of choice and the investigations all employed one or the other functional. Systems of this size, especially with possible explicit solvent molecules, require the sampling of a large number of oxidation states, intermediates and conformations. With such a large number of structures to calculate, a sound strategy is required to keep the computations accurate and the cost low.

In a first step a range of different approximations for geometry relaxation procedures were thus implemented into a newly created software, StOpt, and applied to several molecular test sets. The main optimisation scheme employed the rational function algorithm for the step prediction, which is known to excel if first derivatives from the electronic structure calculation are available. Since a correct step restriction would require an iterative solution of the RF equations, the RS-RF algorithm was implemented as an alternative.<sup>[3,112]</sup> A standard set of redundant internals was used as the primary coordinate system for the optimisation procedure, as they proved to be robust and lead to fast convergence in comparison to cartesian coordinates.<sup>[85]</sup> Optionally an extended set that includes auxiliary bonds between next neighbours could be chosen. Approximate second derivatives covered a diagonal parameterisation and several model Hessians defined in the literature, which were updated with previous steps in the course of the minimisation.

Several of these options have then been tested on 41 structures of small drug-like organic molecules which were preoptimised with a simple force field, resulting in crude starting geometries of structures with many ring and linear-chain motifs. The investigation focussed on the influence of the step restriction, Hessian approximation and coordinate system. It was shown that the convergence depends strongly on the sophistication of the parameterised second derivatives, despite the use of internal coordinates. Nevertheless, although the RF and RF-RS methods performed comparably with a model Hessian, a correct step restriction helped to converge the desired stationary point in the case of simple diagonal approximations. Employing an update of the trust radius reduced the overall number of iterations. By increasing the number of optimisation steps used in the update of the Hessian matrix it was found that a successive updating scheme is superior to using only five steps if combined with this coordinate set. Since the primitive coordinates have a significant impact on the convergence of the minimisation, the standard definition was compared against the extra-redundant internal coordinates. This revealed an advantage of the smaller coordinate set which resulted in faster convergence, but may be too restricted for difficult systems. The flexibility of the PIC set was studied with several program runs that recalculated the internal coordinates after a number of steps and tracked the change of the rigidity index during convergence. For geometries far from the minimum, a pronounced difference between the rigidity behaviour of the standard and extra-redundant coordinates was found and thus a recalculation of the primitives is recommended.

Two efficient and robust methods derived from this benchmarking were then compared to literature-known results which employed a set of 30 small molecules introduced by Baker. The obtained results were encouraging, with a similar performance as the procedure on which the coordinates were based, although other optimisation schemes reported less iterations for this benchmark set. Nevertheless, the literature values demonstrated large discrepancies in the average number of iterations between several codes. In additional minimisation runs on larger molecules it was shown that the advantage of redundant internal coordinates is stronger in systems which are either hard to converge or exhibit many uncommon structural motifs, as in the case of jawsamycin and cubane. For a final test a recently developed CO<sub>2</sub> reduction catalyst was optimised with StOpt and Molpro's internal procedure. This hinted at some deficiencies in the internal coordinate set for transition metal complexes, but still halved the iterations needed for convergence.

Despite the overall competitive performance and robustness of StOpt, some parts in the optimisation can still be improved. The Hessian model has a significant influence on the number of steps until convergence is achieved, but a lot of different parameterisations exist. In the redundant internals primitive coordinates like bonds, angles and dihedrals can be directly addressed in the second derivatives, such that a weighting according to the primary coordinates of interest can be beneficial. On top of that the Hessian update was shown to achieve slightly better convergence if a combination of several algorithms is employed, which was not tested within this study.<sup>[4]</sup> However, a large part of the performance stems from the primitive coordinate set, and thus a more specific definition of primitives or a combination with natural internals could further reduce the number of optimisation steps.<sup>[116,119,221,222]</sup> Most of the problems in optimisation procedures arise with difficult or uncommon geometries so that a flow-chart approach, which chooses the appropriate method and coordinates in each situation, should be the first step for improvement.<sup>[4]</sup> For larger molecular structures the diagonalisation or inversion of matrices used in the optimisation procedure may become the actual bottleneck and require additional algorithms to reduce the overall cost.<sup>[223]</sup> Similarly important, the robustness of the procedure has to be tested on a larger variety of bond types and strengths. As such a more diverse benchmark set should be devised.

In chapter 4 an approach to acquire computational fingerprints for intermediates in the field of CO<sub>2</sub> reduction catalysis was reported. As an exemplary system a recently published rhenium-based dinuclear complex with a protic ligand and catalytic CO<sub>2</sub> reactivity was chosen, which had been comprehensively characterised by cyclovoltammetry and infrared-spectroelectrochemistry.<sup>[137]</sup> The reported crystal structure served as a reference to test the functional and solvent influence on the predicted vibrational frequencies of the tricarbonyl group-stretching modes in comparison with experiment. A difference in the density functional construction exhibited a negligible influence, but the inclusion of solvent contributions in the form of a continuum model was found to be necessary. A total shift of the vibrational frequencies between the experimental and theoretical results persisted, but a good agreement could be obtained for the relative difference of the symmetric and antisymmetric stretching modes. The consequently determined difference in frequencies of the symmetric stretching modes was then used as a constant correction to the computed values, which led to similarly accurate spectra for the first two reduction intermediates.

A more ambitious investigation was needed for the structure of the final reduction. Lack of additional experimental evidence and a multitude of possible side-reactions required the calculation of a large amount of molecular structures and charge states. Reduced accuracy in comparison with the first

two assignments had to be expected, since the corresponding first transformations did not change the number of electrons in the system. In spite of this, the collection of investigated structures only returned a few spectra with reasonable agreement, of which just one had an asymmetry in the intensities similar to the experimental results. As such the developed computational protocol also proved reliable without additional evidence other than the infrared spectrum taken from experiment. The large number of relative shifts for the vibrational bands of the CO stretching modes can thus serve as a reference for future developments of CO<sub>2</sub> reduction catalysts.

The oxygen evolution reaction is known to be the bottleneck of the processes occurring in artificial photosynthesis and thus is actively studied experimentally and computationally. A novel dinuclear ruthenium complex, Ru-Mebbp, that has recently been developed and thoroughly characterised, was thus examined in chapter 5, as it exhibits complementary reactivity to the Ru-Hbpb catalyst with only subtle changes in the ligand framework<sup>[13,14]</sup> and therefore lends itself to an inspection of possible reaction barrier tuning. The interaction of two metal-oxo sites and water nucleophilic attack pathways were consequently calculated for two possible oxidation and several spin states, giving a concise picture of the barrier heights for both mechanisms. Due to this investigation it was revealed that the computational approach used on the Ru-Hbpb complex fails to describe the experimental reactivity of Ru-Mebbp, as the activation free enthalpies with M06-L always suggest a preference for the other pathway or no significant difference at all. In addition the assigned mechanism could only be reproduced in the [4,5] state of the complex, suggesting that a further oxidation is necessary for the observed reactivity. This is in line with experimental suggestions for Ru-Mebbp and other complexes, that state an additional oxidation should occur.<sup>[13,192]</sup> Notwithstanding, the relative barrier height of the active intermediate of the complex to the mechanism determining state of both pathways were found to be dependent on the amount of exact exchange in the density functional. Since it was proposed that an amount of 15 % Hartree-Fock exchange improves the spin-state energies in first-row transition metal complexes,<sup>[36]</sup> the B3LYP\* functional was then employed as the method of choice throughout the further investigation.

With a high confidence in the computational protocol gained by the mechanistic studies, a search for the differences in structure that define the reactivity of Ru-Mebbp and other catalysts was conducted. For this Ru-Mebbp, Ru-Hbpb and the similar Ru-cbim<sup>[15,168]</sup> were evaluated on the same level of theory, making a direct comparison possible. Even though a discrepancy between the determined reaction barriers and the literature values remained, the qualitative conclusion stayed the same and a connection between the catalysts is justified. It was consequently shown that Ru-cbim, despite the proposed difference in its mechanism, is similar to Ru-Mebbp, which is attributed to the comparable coordination environment. As such the difference in reactivity stems from the additional electron removal after the O–O bond formation and similar selectivity as the one found for Ru-Mebbp in the [4,4] oxidation state are acquired without it. Thus the lower oxidation potentials for Ru-cbim resulting from a more negative charge of the ligand backbone and a higher number of possible proton-coupled electron transfer steps are expected to be responsible for its mechanism. The disparity between Ru-Hbpb and Ru-Mebbp is more subtle, as the charge and oxidation states seem to be very similar, which is expressed in the barrier height of the water nucleophilic attack pathway. The corresponding barrier for the intramolecular oxygen–oxygen coupling mechanism on the other hand exhibited a large variation in height, and thus the minimum energy path along its reaction coordinate was scanned for Ru-Hbpb and Ru-Mebbp. Due

to this analysis it could be suggested that the lowered barriers exhibited by Ru-Hbpb result from the structure of the axial ligands, as they form a rather open active site. Contrarily the ligands of Ru-Mebbp hinder the breaking of a peroxide bond. Consequently the second transition state in the intramolecular pathway of Ru-Hbpb was found to be early, whereas the axial pyridines in Ru-Mebbp lead to a late transition state geometry.

The latter investigation has only scratched at the surface of a structural indicator for the mechanistic pathway. As such a more detailed study of the minimum energy pathway may give additional insight. However, this also depends on the chosen electronic structure method, or the modeling of the solvent. Since Ru-Mebbp was used as a model system for the actual catalyst Ru-Mebbp-SO<sub>3</sub>, the first solvation shell might play an increased role if the larger structure is employed. A sampling of several geometries is then required for accurate reaction energies, which may only be computationally feasible by QM/MM or lower-cost methods.<sup>[224]</sup> The extension of the comparison to a third catalyst of similar coordination environment<sup>[171]</sup> could also lead to the final hint about the structural difference. Ideally the calculations should give a complete picture on the variation in barrier height by the modification of ligands, enabling the theoretician to suggest routes for development to experimentators. For this an extensive probing of substituents could be undertaken and correlated to the mechanism determining states.<sup>[192]</sup> Even employing largely different axial ligands could be a viable route for further calculations that might corroborate the proposed ligand dependence. Nevertheless, a complete impression is only obtained if all aspects of the catalysis are taken into account, which includes the oxidation potentials and the rate-determining step of the whole reaction cycle. The latter is generally found to be the final dioxygen displacement by a second water molecule in dinuclear catalysts,<sup>[15,192]</sup> which was not determined within this study, as this transition state does not influence the preferred mechanism of the complex. On the other hand an additional oxidation at an intermediate could change the successive barriers, or lead to mechanistic trapping as in the case of Ru-cbim, and thus may be a promising target for further investigations of mechanism tuning.



# Bibliography

- [1] C.-F. Schlessner, J. Rogelj, M. Schaeffer, T. Lissner, R. Licker, E. M. Fischer, R. Knutti, A. Levermann, K. Frieler, W. Hare, *Nature Climate Change* **July 25, 2016**, 6, nclimate3096.
- [2] T. Hisatomi, K. Domen, *Faraday Discussions* **2017**, 198, 11–35.
- [3] H. B. Schlegel, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **Sept. 2011**, 1, 790–809.
- [4] A. B. Birkholz, H. B. Schlegel, *Theoretical Chemistry Accounts* **Apr. 2016**, 135, DOI 10.1007/s00214-016-1847-3.
- [5] H. Takeda, C. Cometto, O. Ishitani, M. Robert, *ACS Catalysis* **Jan. 6, 2017**, 7, 70–88.
- [6] C. Costentin, S. Drouet, M. Robert, J.-M. Saveant, *Science* **Oct. 5, 2012**, 338, 90–94.
- [7] C. W. Machan, S. A. Chabolla, J. Yin, M. K. Gilson, F. A. Tezcan, C. P. Kubiak, *Journal of the American Chemical Society* **Oct. 15, 2014**, 136, 14598–14607.
- [8] J. Agarwal, E. Fujita, H. F. Schaefer, J. T. Muckerman, *Journal of the American Chemical Society* **Mar. 21, 2012**, 134, 5180–5186.
- [9] D. J. Vinyard, G. W. Brudvig, *Annual Review of Physical Chemistry* **May 5, 2017**, 68, 101–116.
- [10] G. C. Dismukes, R. Brimblecombe, G. A. N. Felton, R. S. Pryadun, J. E. Sheats, L. Spiccia, G. F. Swiegers, *Accounts of Chemical Research* **Dec. 21, 2009**, 42, 1935–1943.
- [11] R.-Z. Liao, M. D. Kärkäs, B.-L. Lee, B. Åkermark, P. E. M. Siegbahn, *Inorganic Chemistry* **Jan. 5, 2015**, 54, 342–351.
- [12] J. D. Blakemore, R. H. Crabtree, G. W. Brudvig, *Chemical Reviews* **Dec. 9, 2015**, 115, 12974–13005.
- [13] S. Neudeck, S. Maji, I. López, S. Meyer, F. Meyer, A. Llobet, *Journal of the American Chemical Society* **Jan. 8, 2014**, 136, 24–27.
- [14] F. Bozoglian, S. Romain, M. Z. Ertem, T. K. Todorova, C. Sens, J. Mola, M. Rodríguez, I. Romero, J. Benet-Buchholz, X. Fontrodona, C. J. Cramer, L. Gagliardi, A. Llobet, *Journal of the American Chemical Society* **Oct. 28, 2009**, 131, 15176–15187.
- [15] R.-Z. Liao, M. D. Kärkäs, T. M. Laine, B. Åkermark, P. E. M. Siegbahn, *Catal. Sci. Technol.* **2016**, 6, 5031–5041.
- [16] M. Born, R. Oppenheimer, *Annalen der Physik* **1927**, 389, 457–484.
- [17] A. Szabo, N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*, Dover Publications, Mineola, N.Y., **1996**.
- [18] P. Hohenberg, W. Kohn, *Physical review* **1964**, 136, B864.
- [19] W. Kohn, L. J. Sham, *Physical Review* **Nov. 15, 1965**, 140, A1133–A1138.

- [20] F. Bloch, *Zeitschrift für Physik* **July 1929**, 57, 545–555.
- [21] P. a. M. Dirac, *Mathematical Proceedings of the Cambridge Philosophical Society* **July 1930**, 26, 376–385.
- [22] J. C. Slater, *Physical Review* **1951**, 81, 385.
- [23] D. M. Ceperley, B. J. Alder, *Physical Review Letters* **Aug. 18, 1980**, 45, 566–569.
- [24] S. H. Vosko, L. Wilk, M. Nusair, *Canadian Journal of Physics* **Aug. 1980**, 58, 1200–1211.
- [25] J. P. Perdew, K. Schmidt, *AIP Conference Proceedings* **July 6, 2001**, 577, 1–20.
- [26] F. Neese, *Coordination Chemistry Reviews* **Mar. 2009**, 253, 526–563.
- [27] A. D. Becke, *Physical Review A* **Sept. 1, 1988**, 38, 3098–3100.
- [28] J. P. Perdew, *Physical Review B* **June 15, 1986**, 33, 8822–8824.
- [29] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Physical Review Letters* **Sept. 30, 2003**, 91, 146401.
- [30] F. Furche, J. P. Perdew, *The Journal of Chemical Physics* **Jan. 24, 2006**, 124, 044103.
- [31] Y. Zhao, D. G. Truhlar, *The Journal of Chemical Physics* **Nov. 15, 2006**, 125, 194101.
- [32] A. D. Becke, *The Journal of Chemical Physics* **May 14, 2014**, 140, 18A301.
- [33] C. J. Cramer, *Essentials of computational chemistry: theories and models*, 2nd ed, Wiley, Chichester, West Sussex, England ; Hoboken, NJ, **2004**, 596 pp.
- [34] P. J. Stephens, F. J. Devlin, C. Chabalowski, M. J. Frisch, *The Journal of Physical Chemistry* **1994**, 98, 11623–11627.
- [35] C. Lee, W. Yang, R. G. Parr, *Physical Review B* **Jan. 15, 1988**, 37, 785–789.
- [36] M. Reiher, O. Salomon, B. Artur Hess, *Theoretical Chemistry Accounts: Theory Computation and Modeling (Theoretica Chimica Acta)* **Dec. 1, 2001**, 107, 48–55.
- [37] L. Noodleman, *The Journal of Chemical Physics* **May 15, 1981**, 74, 5737–5743.
- [38] L. Noodleman, E. R. Davidson, *Chemical Physics* **Nov. 1, 1986**, 109, 131–143.
- [39] F. Neese, *Journal of Physics and Chemistry of Solids* **Apr. 2004**, 65, 781–785.
- [40] D. A. Pantazis, V. Krewald, M. Orio, F. Neese, *Dalton Transactions* **May 11, 2010**, 39, 4959–4967.
- [41] K. Yamaguchi, F. Jensen, A. Dorigo, K. Houk, *Chemical Physics Letters* **Sept. 2, 1988**, 149, 537–542.
- [42] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *The Journal of Chemical Physics* **2010**, 132, 154104.
- [43] E. R. Johnson, A. D. Becke, *The Journal of Chemical Physics* **May 5, 2006**, 124, 174104.
- [44] S. Grimme, S. Ehrlich, L. Goerigk, *Journal of Computational Chemistry* **May 2011**, 32, 1456–1465.
- [45] A. Dreuw, J. L. Weisman, M. Head-Gordon, *The Journal of Chemical Physics* **Aug. 8, 2003**, 119, 2943–2946.
- [46] T. Yanai, D. P. Tew, N. C. Handy, *Chemical Physics Letters* **July 21, 2004**, 393, 51–57.
- [47] R. Baer, E. Livshits, U. Salzner, *Annual Review of Physical Chemistry* **Mar. 2010**, 61, 85–109.

- [48] Y. Zhao, B. J. Lynch, D. G. Truhlar, *The Journal of Physical Chemistry A* **May 2004**, *108*, 4786–4791.
- [49] Y. Zhao, B. J. Lynch, D. G. Truhlar, *Physical Chemistry Chemical Physics* **2005**, *7*, 43.
- [50] L. Goerigk, S. Grimme, *Journal of Chemical Theory and Computation* **Feb. 8, 2011**, *7*, 291–309.
- [51] G. E. Scuseria, T. M. Henderson, D. C. Sorensen, *The Journal of Chemical Physics* **Dec. 21, 2008**, *129*, 231101.
- [52] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *The Journal of Physical Chemistry B* **May 7, 2009**, *113*, 6378–6396.
- [53] R. J. Zauhar, A. Varnek, *Journal of Computational Chemistry* **May 1996**, *17*, 864–877.
- [54] A. Klamt, G. Schüürmann, *Journal of the Chemical Society Perkin Transactions 2* **1993**, 799–805.
- [55] A. Klamt, V. Jonas, T. Bürger, J. C. W. Lohrenz, *The Journal of Physical Chemistry A* **June 1998**, *102*, 5074–5085.
- [56] S. Miertuš, E. Scrocco, J. Tomasi, *Chemical Physics* **Feb. 1981**, *55*, 117–129.
- [57] D. M. Dolney, G. D. Hawkins, P. Winget, D. A. Liotard, C. J. Cramer, D. G. Truhlar, *Journal of Computational Chemistry* **Apr. 15, 2000**, *21*, 340–366.
- [58] J. Tomasi, B. Mennucci, R. Cammi, *Chemical Reviews* **Aug. 2005**, *105*, 2999–3094.
- [59] C. J. Cramer, D. G. Truhlar, *Accounts of Chemical Research* **Apr. 21, 2009**, *42*, 493–497.
- [60] A. V. Marenich, J. Ho, M. L. Coote, C. J. Cramer, D. G. Truhlar, *Physical Chemistry Chemical Physics* **July 2, 2014**, *16*, 15068–15106.
- [61] S. J. Konezny, M. D. Doherty, O. R. Luca, R. H. Crabtree, G. L. Soloveichik, V. S. Batista, *The Journal of Physical Chemistry C* **2012**, *116*, 6349–6356.
- [62] R. C. Johnston, J. Zhou, J. C. Smith, J. M. Parks, *The Journal of Physical Chemistry B* **July 8, 2016**, DOI 10.1021/acs.jpcc.6b02701.
- [63] M. G. Ullmann, L. Noodleman, D. A. Case, *JBIC Journal of Biological Inorganic Chemistry* **June 2002**, *7*, 632–639.
- [64] R. Casanovas, J. Ortega-Castro, J. Frau, J. Donoso, F. Muñoz, *International Journal of Quantum Chemistry* **Oct. 15, 2014**, *114*, 1350–1363.
- [65] B. Thapa, H. B. Schlegel, *The Journal of Physical Chemistry A* **June 22, 2017**, *121*, 4698–4706.
- [66] M. Reiher, G. Brehm, S. Schneider, *The Journal of Physical Chemistry A* **Feb. 1, 2004**, *108*, 734–742.
- [67] V. Barone, A. Polimeno, *Chemical Society Reviews* **2007**, *36*, 1724.
- [68] P. Pechukas, *Annual Review of Physical Chemistry* **1981**, *32*, 159–177.
- [69] D. Chandler, D. E. Manolopoulos, *Faraday Discuss.* **2016**, *195*, 699–710.
- [70] A. Banerjee, N. Adams, J. Simons, R. Shepard, *The Journal of Physical Chemistry* **1985**, *89*, 52–57.
- [71] P. Pulay, *Journal of Computational Chemistry* **1982**, *3*, 556–560.
- [72] P. Császár, P. Pulay, *Journal of Molecular Structure* **Jan. 1984**, *114*, 31–34.

- [73] X. Li, M. J. Frisch, *Journal of Chemical Theory and Computation* **May 2006**, 2, 835–839.
- [74] Ö. Farkas, H. B. Schlegel, *Physical Chemistry Chemical Physics* **Jan. 7, 2002**, 4, 11–15.
- [75] H. B. Schlegel, *Journal of Computational Chemistry* **1982**, 3, 214–218.
- [76] H. B. Schlegel, *Journal of computational chemistry* **2003**, 24, 1514–1527.
- [77] C. G. Broyden, *Mathematics of Computation* **1970**, 24, 365–382.
- [78] R. Fletcher, *The Computer Journal* **Mar. 1, 1970**, 13, 317–322.
- [79] D. Goldfarb, *Mathematics of computation* **1970**, 24, 23–26.
- [80] D. F. Shanno, *Mathematics of computation* **1970**, 24, 647–656.
- [81] B. A. Murtagh, *The Computer Journal* **Feb. 1, 1970**, 13, 185–194.
- [82] M. J. D. Powell, *IMA Journal of Applied Mathematics* **1971**, 7, 21–36.
- [83] Ö. Farkas, H. B. Schlegel, *The Journal of Chemical Physics* **1999**, 111, 10806.
- [84] J. Baker, *Journal of computational chemistry* **1993**, 14, 1085–1100.
- [85] V. Bakken, T. Helgaker, *The Journal of Chemical Physics* **2002**, 117, 9160.
- [86] H. B. Schlegel, *Theoretica chimica acta* **1984**, 66, 333–340.
- [87] T. H. Fischer, J. Almlof, *The Journal of Physical Chemistry* **1992**, 96, 9768–9774.
- [88] R. Lindh, A. Bernhardsson, G. Karlström, P.-Å. Malmqvist, *Chemical physics letters* **1995**, 241, 423–428.
- [89] F. Eckert, P. Pulay, H.-J. Werner, *Journal of Computational Chemistry* **Sept. 1997**, 18, 1473–1483.
- [90] 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, D. J. Fox, Gaussian~09 Revision E.01, Gaussian Inc. Wallingford CT 2009.
- [91] C. Peng, P. Y. Ayala, H. B. Schlegel, M. J. Frisch, *Journal of Computational Chemistry* **1996**, 17, 49–56.
- [92] E. F. Koslover, D. J. Wales, *The Journal of Chemical Physics* **2007**, 127, 234105.
- [93] I. Kolossváry, C. McMartin, *Journal of mathematical chemistry* **1992**, 9, 359–367.
- [94] G. Fogarasi, X. Zhou, P. W. Taylor, P. Pulay, *Journal of the American Chemical Society* **Oct. 1, 1992**, 114, 8191–8201.
- [95] P. Pulay, G. Fogarasi, *The Journal of Chemical Physics* **Feb. 15, 1992**, 96, 2856–2860.
- [96] S. R. Billeter, A. J. Turner, W. Thiel, *Physical Chemistry Chemical Physics* **2000**, 2, 2177–2186.

- [97] E. B. Wilson, J. C. Decius, P. C. Cross, *Molecular vibrations: the theory of infrared and Raman vibrational spectra*, Dover Publications, New York, **1980**, 388 pp.
- [98] J. Baker, D. Kinghorn, P. Pulay, *The Journal of chemical physics* **1999**, *110*, 4986.
- [99] B. Paizs, J. Baker, S. Suhai, P. Pulay, *The Journal of Chemical Physics* **2000**, *113*, 6566.
- [100] V. V. Rybkin, U. Ekström, T. Helgaker, *Journal of Computational Chemistry* **Aug. 5, 2013**, *34*, 1842–1849.
- [101] J. Baker, A. Kessi, B. Delley, *The Journal of chemical physics* **1996**, *105*, 192.
- [102] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **Mar. 1, 2012**, *2*, 242–253.
- [103] P. Pyykkö, M. Atsumi, *Chemistry - A European Journal* **Jan. 2009**, *15*, 186–197.
- [104] S. Alvarez, *Dalton Transactions* **2013**, *42*, 8617.
- [105] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, T. D. Crawford, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **July 2012**, *2*, 556–565.
- [106] L. C. Blum, J.-L. Reymond, *Journal of the American Chemical Society* **July 2009**, *131*, 8732–8733.
- [107] S. L. Mayo, B. D. Olafson, W. A. Goddard, *Journal of Physical chemistry* **1990**, *94*, 8897–8909.
- [108] J. P. Perdew, K. Burke, M. Ernzerhof, *Physical review letters* **1996**, *77*, 3865.
- [109] F. Weigend, R. Ahlrichs, *Physical Chemistry Chemical Physics* **2005**, *7*, 3297.
- [110] F. Weigend, *Physical Chemistry Chemical Physics* **2006**, *8*, 1057.
- [111] J. M. Anglada, J. M. Bofill, *International Journal of Quantum Chemistry* **1997**, *62*, 153–165.
- [112] E. Besalú, J. M. Bofill, *Theoretical Chemistry Accounts* **1998**, *100*, 265–274.
- [113] K. Németh, M. Challacombe, M. Van Veenendaal, *Journal of Computational Chemistry* **2010**, NA–NA.
- [114] K. J. Miller, *Journal of Computational Chemistry* **1990**, *11*, 336–345.
- [115] S.-H. Lee, K. Palmo, S. Krimm, **1999**.
- [116] P. E. Maslen, *The Journal of Chemical Physics* **2005**, *122*, 014104.
- [117] W. J. Hehre, R. F. Stewart, J. A. Pople, *The Journal of Chemical Physics* **Sept. 15, 1969**, *51*, 2657–2664.
- [118] R. Angamuthu, P. Byers, M. Lutz, A. L. Spek, E. Bouwman, *Science* **Jan. 14, 2010**, *327*, 313–315.
- [119] M. von Arnim, R. Ahlrichs, *The Journal of Chemical Physics* **1999**, *111*, 9183.
- [120] I. E. Woodrow, J. A. Berry, *Annual Review of Plant Physiology and Plant Molecular Biology* **1988**, *39*, 533–594.
- [121] A. R. Portis, M. A. J. Parry, *Photosynthesis Research* **Aug. 20, 2007**, *94*, 121–143.

- [122] A. M. Appel, J. E. Bercaw, A. B. Bocarsly, H. Dobbek, D. L. DuBois, M. Dupuis, J. G. Ferry, E. Fujita, R. Hille, P. J. A. Kenis, C. A. Kerfeld, R. H. Morris, C. H. F. Peden, A. R. Portis, S. W. Ragsdale, T. B. Rauchfuss, J. N. H. Reek, L. C. Seefeldt, R. K. Thauer, G. L. Waldrop, *Chemical Reviews* **Aug. 14, 2013**, *113*, 6621–6658.
- [123] T. Herzog, *World Resources Institute* **2009**.
- [124] *Modern Aspects of Electrochemistry*, (Eds.: C. G. Vayenas, R. E. White, M. E. Gamboa-Aldeco), DOI: 10.1007/978-0-387-49489-0, Springer New York, New York, NY, **2008**.
- [125] H. A. Schwarz, R. W. Dodson, *The Journal of Physical Chemistry* **Jan. 1989**, *93*, 409–414.
- [126] C. Costentin, M. Robert, J.-M. Savéant, *Chemical Society Reviews* **2013**, *42*, 2423.
- [127] A. J. Morris, G. J. Meyer, E. Fujita, *Accounts of Chemical Research* **Dec. 21, 2009**, *42*, 1983–1994.
- [128] B. P. Sullivan, C. M. Bolinger, D. Conrad, W. J. Vining, T. J. Meyer, *Journal of the Chemical Society Chemical Communications* **1985**, 1414–1416.
- [129] T. R. O’Toole, B. Sullivan, M. R.-M. Bruce, L. D. Margerum, R. W. Murray, T. J. Meyer, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* **Feb. 1989**, *259*, 217–239.
- [130] P. Christensen, A. Hamnett, A. V. G. Muir, J. A. Timney, *Journal of the Chemical Society Dalton Transactions* **1992**, 1455–1463.
- [131] F. P. Johnson, M. W. George, F. Hartl, J. J. Turner, *Organometallics* **1996**, *15*, 3374–3387.
- [132] G. J. Stor, F. Hartl, J. W. M. van Outersterp, D. J. Stufkens, *Organometallics* **Mar. 1, 1995**, *14*, 1115–1131.
- [133] J. M. Smieja, C. P. Kubiak, *Inorganic Chemistry* **Oct. 18, 2010**, *49*, 9283–9289.
- [134] E. E. Benson, C. P. Kubiak, *Chemical Communications* **2012**, *48*, 7374.
- [135] G. F. Manbeck, J. T. Muckerman, D. J. Szalda, Y. Himeda, E. Fujita, *The Journal of Physical Chemistry B* **June 18, 2015**, *119*, 7457–7466.
- [136] Q. Zeng, M. Messaoudani, A. Vlček, F. Hartl, *European Journal of Inorganic Chemistry* **Jan. 2012**, *2012*, 471–474.
- [137] A. Wilting, T. Stolper, R. A. Mata, I. Siewert, *Inorganic Chemistry* **Apr. 3, 2017**, *56*, 4176–4185.
- [138] F. Neese, *Journal of Computational Chemistry* **Nov. 15, 2003**, *24*, 1740–1747.
- [139] D. Andrae, U. Häußermann, M. Dolg, H. Stoll, H. Preuß, *Theoretica chimica acta* **Mar. 1, 1990**, *77*, 123–141.
- [140] L. Yu, C. Greco, M. Bruschi, U. Ryde, L. De Gioia, M. Reiher, *Inorganic Chemistry* **May 2, 2011**, *50*, 3888–3900.
- [141] C. Adamo, V. Barone, *The Journal of Chemical Physics* **Mar. 23, 1999**, *110*, 6158–6170.
- [142] A. C. Benniston, A. Harriman, P. Li, C. A. Sams, M. D. Ward, *Journal of the American Chemical Society* **Oct. 2004**, *126*, 13630–13631.
- [143] A. C. Benniston, A. Harriman, *Chem. Soc. Rev.* **2006**, *35*, 169–179.
- [144] D. R. J. Kolling, N. Cox, G. M. Ananyev, R. J. Pace, G. C. Dismukes, *Biophysical Journal* **July 18, 2012**, *103*, 313–322.

- [145] T. J. Meyer, M. H. V. Huynh, H. H. Thorp, *Angewandte Chemie International Edition* **July 9, 2007**, *46*, 5284–5304.
- [146] R. J. Pace, R. Stranger, S. Petrie, *Dalton Transactions* **2012**, *41*, 7179.
- [147] C. Zhang, C. Chen, H. Dong, J.-R. Shen, H. Dau, J. Zhao, *Science* **May 8, 2015**, *348*, 690–693.
- [148] M. M. Najafpour, G. Renger, M. Hołyńska, A. N. Moghaddam, E.-M. Aro, R. Carpentier, H. Nishihara, J. J. Eaton-Rye, J.-R. Shen, S. I. Allakhverdiev, *Chemical Reviews* **Mar. 9, 2016**, *116*, 2886–2936.
- [149] T. Zhang, W. Lin, *Chemical Society Reviews* **2014**, *43*, 5982–5993.
- [150] B. A. Johnson, A. Bhunia, S. Ott, *Dalton Trans.* **2017**, *46*, 1382–1388.
- [151] B. Klahr, S. Gimenez, F. Fabregat-Santiago, J. Bisquert, T. W. Hamann, *Energy & Environmental Science* **May 23, 2012**, *5*, 7626–7636.
- [152] X. Chen, D. Aschaffenburg, T. Cuk, *Journal of Materials Chemistry A* **June 13, 2017**, *5*, 11410–11417.
- [153] A. Grimaud, O. Diaz-Morales, B. Han, W. T. Hong, Y.-L. Lee, L. Giordano, K. A. Stoerzinger, M. T. M. Koper, Y. Shao-Horn, *Nature Chemistry* **Jan. 9, 2017**, DOI 10.1038/nchem.2695.
- [154] R. Cao, W. Lai, P. Du, *Energy & Environmental Science* **July 18, 2012**, *5*, 8134–8157.
- [155] M. D. Kärkäs, O. Verho, E. V. Johnston, B. Åkermark, *Chemical Reviews* **Dec. 24, 2014**, *114*, 11863–12001.
- [156] A. R. Parent, K. Sakai, *ChemSusChem* **Aug. 1, 2014**, *7*, 2070–2080.
- [157] I. Corbucci, K. Ellingwood, L. Fagiolari, C. Zuccaccia, F. Elisei, P. L. Gentili, A. Macchioni, *Catalysis Today* **July 2017**, *290*, 10–18.
- [158] S. J. Koepke, K. M. Light, P. E. VanNatta, K. M. Wiley, M. T. Kieber-Emmons, *Journal of the American Chemical Society* **June 28, 2017**, *139*, 8586–8600.
- [159] J. Li, R. Güttinger, R. Moré, F. Song, W. Wan, G. R. Patzke, *Chem. Soc. Rev.* **2017**, DOI 10.1039/C7CS00306D.
- [160] L.-P. Wang, T. Van Voorhis, *The Journal of Physical Chemistry Letters* **Sept. 2011**, *2*, 2200–2204.
- [161] M. Muuronen, S. M. Parker, E. Berardo, A. Le, M. A. Zwijnenburg, F. Furche, *Chem. Sci.* **2017**, *8*, 2179–2183.
- [162] J. Soriano-López, D. G. Musaev, C. L. Hill, J. R. Galán-Mascarós, J. J. Carbó, J. M. Poblet, *Journal of Catalysis* **June 2017**, *350*, 56–63.
- [163] X. Yang, M.-H. Baik, *Journal of the American Chemical Society* **Dec. 3, 2008**, *130*, 16231–16240.
- [164] D. W. Shaffer, Y. Xie, J. J. Concepcion, *Chemical Society Reviews* **Sept. 1, 2017**, DOI 10.1039/C7CS00542C.
- [165] L. Duan, F. Bozoglian, S. Mandal, B. Stewart, T. Privalov, A. Llobet, L. Sun, *Nature Chemistry* **Mar. 25, 2012**, *4*, 418–423.
- [166] L. Wang, L. Duan, Y. Wang, M. S. G. Ahlquist, L. Sun, *Chem. Commun.* **Sept. 1, 2014**, *50*, 12947–12950.

- [167] T. Fan, L. Duan, P. Huang, H. Chen, Q. Daniel, M. S. G. Ahlquist, L. Sun, *ACS Catalysis* **Apr. 7, 2017**, *7*, 2956–2966.
- [168] T. M. Laine, M. D. Kärkäs, R.-Z. Liao, T. Åkermark, B.-L. Lee, E. A. Karlsson, P. E. M. Siegbahn, B. Åkermark, *Chem. Commun.* **2015**, *51*, 1862–1865.
- [169] C. Sens, I. Romero, M. Rodríguez, A. Llobet, T. Parella, J. Benet-Buchholz, *Journal of the American Chemical Society* **June 2004**, *126*, 7798–7799.
- [170] S. Romain, F. Bozoglian, X. Sala, A. Llobet, *Journal of the American Chemical Society* **Mar. 4, 2009**, *131*, 2768–2769.
- [171] S. Maji, L. Vigara, F. Cottone, F. Bozoglian, J. Benet-Buchholz, A. Llobet, *Angewandte Chemie International Edition* **June 11, 2012**, *51*, 5967–5970.
- [172] S. Berardi, L. Francàs, S. Neudeck, S. Maji, J. Benet-Buchholz, F. Meyer, A. Llobet, *ChemSusChem* **Nov. 2015**, *8*, 3688–3696.
- [173] S. Neudeck, S. Maji, I. López, S. Dechert, J. Benet-Buchholz, A. Llobet, F. Meyer, *Inorganic Chemistry* **Mar. 7, 2016**, *55*, 2508–2521.
- [174] F. Liu, J. J. Concepcion, J. W. Jurss, T. Cardolaccia, J. L. Templeton, T. J. Meyer, *Inorganic Chemistry* **Mar. 1, 2008**, *47*, 1727–1752.
- [175] R. Bianco, P. J. Hay, J. T. Hynes, *The Journal of Physical Chemistry A* **July 14, 2011**, *115*, 8003–8016.
- [176] L. Tong, L. Duan, Y. Xu, T. Privalov, L. Sun, *Angewandte Chemie International Edition* **Jan. 10, 2011**, *50*, 445–449.
- [177] Y. Wang, M. S. G. Ahlquist, *Dalton Transactions* **Aug. 19, 2014**, *43*, 13776–13782.
- [178] R. Kang, K. Chen, J. Yao, S. Shaik, H. Chen, *Inorganic Chemistry* **June 26, 2014**, 7130–7136.
- [179] A. F. Abdel-Magied, A. Shatskiy, R.-Z. Liao, T. M. Laine, W. A. A. Arafa, P. E. M. Siegbahn, M. D. Kärkäs, B. Åkermark, E. V. Johnston, *ChemSusChem* **Dec. 20, 2016**, *9*, 3448–3456.
- [180] R. Matheu, M. Z. Ertem, J. Benet-Buchholz, E. Coronado, V. S. Batista, X. Sala, A. Llobet, *Journal of the American Chemical Society* **Aug. 12, 2015**, 150812101708002.
- [181] J. N. Harvey, *Annual Reports Section 'C' (Physical Chemistry)* **Aug. 3, 2006**, *102*, 203–226.
- [182] G. F. Mangiatordi, E. Brémond, C. Adamo, *Journal of Chemical Theory and Computation* **Sept. 11, 2012**, *8*, 3082–3088.
- [183] M. Steinmetz, S. Grimme, *ChemistryOpen* **June 2013**, *2*, 115–124.
- [184] A. D. Becke, *The Journal of Chemical Physics* **Apr. 1993**, *98*, 5648–5652.
- [185] Y. Zhao, D. G. Truhlar, *Theoretical Chemistry Accounts* **Apr. 2008**, *119*, 525–525.
- [186] V. N. Staroverov, G. E. Scuseria, J. Tao, J. P. Perdew, *The Journal of Chemical Physics* **Dec. 15, 2003**, *119*, 12129–12137.
- [187] F. Neese, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **Jan. 1, 2012**, *2*, 73–78.
- [188] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chemical Physics* **Feb. 2009**, *356*, 98–109.
- [189] D. G. Truhlar, *Journal of computational chemistry* **1991**, *12*, 266–270.



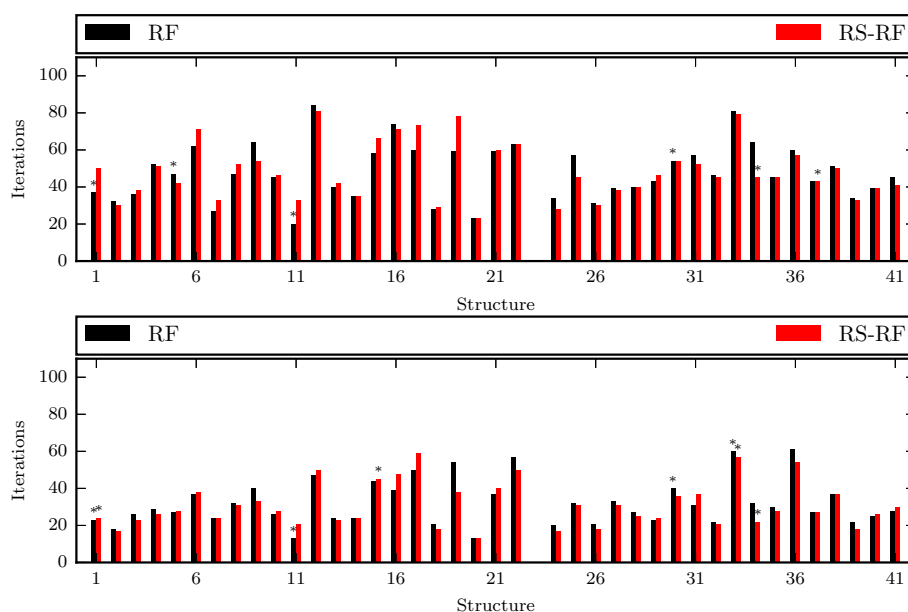
- [190] P. Y. Ayala, H. B. Schlegel, *The Journal of Chemical Physics* **Feb. 8, 1998**, *108*, 2314–2325.
- [191] A. V. Marenich, A. Majumdar, M. Lenz, C. J. Cramer, D. G. Truhlar, *Angewandte Chemie International Edition* **Dec. 14, 2012**, *51*, 12810–12814.
- [192] S. Roeser, F. Bozoglian, C. J. Richmond, A. B. League, M. Z. Ertem, L. Francàs, P. Miró, J. Benet-Buchholz, C. J. Cramer, A. Llobet, *Catal. Sci. Technol.* **2016**, *6*, 5088–5101.
- [193] E. R. Batista, R. L. Martin, *Journal of the American Chemical Society* **June 1, 2007**, *129*, 7224–7225.
- [194] M. Radon, *Physical Chemistry Chemical Physics* **2014**, *16*, WOS:000338442900015, 14479–14488.
- [195] E. I. Ioannidis, H. J. Kulik, *The Journal of Physical Chemistry A* **Feb. 2, 2017**, *121*, 874–884.
- [196] P. Verma, Z. Varga, J. E. M. N. Klein, C. J. Cramer, L. Que, D. G. Truhlar, *Physical Chemistry Chemical Physics* **May 24, 2017**, *19*, 13049–13069.
- [197] *Ab initio molecular orbital theory*, (Ed.: W. J. Hehre), Wiley, New York, **1986**, 548 pp.
- [198] X. Xu, D. G. Truhlar, *Journal of Chemical Theory and Computation* **Sept. 13, 2011**, *7*, 2766–2779.
- [199] J. Zheng, X. Xu, D. G. Truhlar, *Theoretical Chemistry Accounts* **Feb. 1, 2011**, *128*, 295–305.
- [200] D. Roy, M. Marianski, N. T. Maitra, J. J. Dannenberg, *The Journal of Chemical Physics* **Oct. 4, 2012**, *137*, 134109.
- [201] D. Bousquet, E. Brémond, J. C. Sancho-García, I. Ciofini, C. Adamo, *Journal of Chemical Theory and Computation* **Aug. 13, 2013**, *9*, 3444–3452.
- [202] E. Brémond, I. Ciofini, J. C. Sancho-García, C. Adamo, *Accounts of Chemical Research* **Aug. 16, 2016**, *49*, 1503–1513.
- [203] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, K. A. Lyssenko, *Science* **Jan. 6, 2017**, *355*, 49–52.
- [204] D. Hait, M. Head-Gordon, **Sept. 15, 2017**.
- [205] K. P. Kepp, **Feb. 2, 2017**.
- [206] K. R. Brorsen, Y. Yang, M. V. Pak, S. Hammes-Schiffer, *The Journal of Physical Chemistry Letters* **May 4, 2017**, *8*, 2076–2081.
- [207] M. Korth, *Angewandte Chemie International Edition* **May 8, 2017**, *56*, 5396–5398.
- [208] S. E. Wheeler, K. N. Houk, *Journal of Chemical Theory and Computation* **Feb. 9, 2010**, *6*, 395–404.
- [209] E. R. Johnson, R. A. Wolkow, G. A. DiLabio, *Chemical Physics Letters* **Aug. 2004**, *394*, 334–338.
- [210] F. Neese, *JBIC Journal of Biological Inorganic Chemistry* **Sept. 2006**, *11*, 702–711.
- [211] P. E. M. Siegbahn, *JBIC Journal of Biological Inorganic Chemistry* **Sept. 2006**, *11*, 695–701.
- [212] A. Hansen, C. Bannwarth, S. Grimme, P. Petrović, C. Werlé, J.-P. Djukic, *ChemistryOpen* **Oct. 1, 2014**, *3*, 177–189.
- [213] W. Jiang, M. L. Laury, M. Powell, A. K. Wilson, *Journal of Chemical Theory and Computation* **Nov. 13, 2012**, *8*, 4102–4111.
- [214] P. A. Denis, *Theoretical Chemistry Accounts* **May 1, 2011**, *129*, 219–227.
- [215] R. Kang, J. Yao, H. Chen, *Journal of Chemical Theory and Computation* **Apr. 9, 2013**, *9*, 1872–1879.

- [216] L. Hu, H. Chen, *Journal of Chemical Theory and Computation* **Oct. 13, 2015**, *11*, 4601–4614.
- [217] E. I. Ioannidis, H. J. Kulik, *The Journal of Chemical Physics* **July 21, 2015**, *143*, 034104.
- [218] P. Milko, M. A. Iron, *Journal of Chemical Theory and Computation* **Jan. 14, 2014**, *10*, 220–235.
- [219] D. Ardura, R. López, T. L. Sordo, *The Journal of Physical Chemistry B* **Dec. 2005**, *109*, 23618–23623.
- [220] L. P. Wolters, F. M. Bickelhaupt, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **July 2015**, *5*, 324–343.
- [221] M. Swart, F. Matthias Bickelhaupt, *International Journal of Quantum Chemistry* **2006**, *106*, 2536–2544.
- [222] L.-P. Wang, C. Song, *The Journal of Chemical Physics* **June 7, 2016**, *144*, 214108.
- [223] W. Liang, H. Wang, J. Hung, X. Li, M. J. Frisch, *Journal of Chemical Theory and Computation* **July 13, 2010**, *6*, 2034–2039.
- [224] V. Kunz, J. O. Lindner, M. Schulze, M. I. S. Röhr, D. Schmidt, R. Mitrić, F. Würthner, *Energy & Environmental Science* **Sept. 4, 2017**, DOI 10.1039/C7EE01557G.

# Appendix

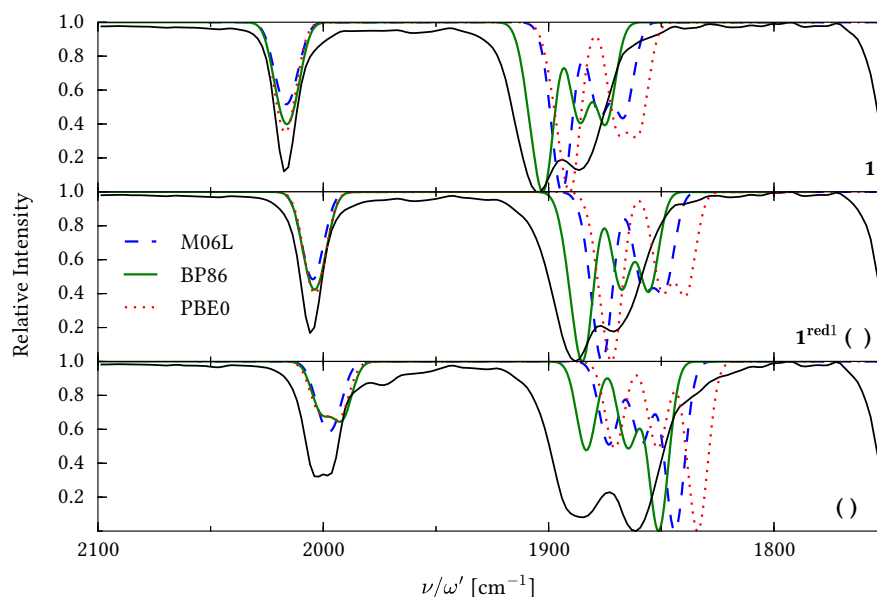
## Additional Benchmarks of StOpt's Methods

Below are the plots referred to in section 3.2.1. Shown are the required iterations to convergence of each structure in the benchmark set OrgSet1. The two rational function methods are compared, either with a simple step scaling (RF), or an iterative solution to the step length (RS-RF). The trust radius was set to 0.3 in the standard redundant internal coordinate set. Hessian guesses were only calculated at the start of the optimisation and then successively updated with BFGS. The first figure shows the results for an initial diagonal parameterisation to the second derivatives and the second figure used Lindh's model Hessian.



## Functional Tests on Reduction Intermediates of 1

The functional test on the splitting between the symmetric and antisymmetric CO stretching bands in Figure 4.4 showed a large variation with the method employed. The relative shifts between the intermediates may be less influenced, which is shown for the first three states and the functionals BP86, M06-L and PBE0 in the following figure. The calculations were carried out with def2-TZVP and COSMO. Computed spectra were determined by broadening the harmonic frequencies with gaussians and  $10\text{ cm}^{-1}$  for the full width at half maximum.



## Cartesian Structures of Intermediates and Transition States

In the following, structures in XYZ format corresponding to the stationary points of the I2M and WNA mechanisms are listed for 5. As the geometries should be relatively similar between oxidation and spin states, values for the quartet surface and a total charge of +4 are reported. Geometry optimisations employed B3LYP\*/def2-SVP with the effective core potential ECP28MWB.

AC

Ru	1.686805	0.073576	0.571797	C	1.561153	-4.272763	0.047180
O	1.303579	-0.174348	-1.104280	H	1.750048	-2.434674	-1.092016
Ru	-2.326832	0.488124	-0.777725	H	1.018803	-1.858777	2.976484
O	-1.112333	-0.045853	-1.863055	C	1.130733	-3.933325	2.386231
H	-0.618824	-0.062163	-5.078857	C	3.894304	0.379778	4.286872
O	-1.373283	0.541452	-4.977382	C	1.578095	0.891874	4.839820
H	-1.549684	0.846937	-5.884213	C	2.543986	4.422115	0.785505
N	1.405054	-2.036334	0.927187	C	3.073093	3.704487	-1.445284
N	2.212344	0.384130	2.599906	C	-1.366590	1.226537	3.321065
N	2.152164	2.136505	0.133935	C	-2.174644	1.101814	2.168518
N	-0.144100	0.573717	1.517152	H	3.938246	-0.734915	-1.426817
N	3.763882	-0.295684	0.603106	C	5.857460	-0.884617	-0.414767
C	1.585742	-2.888178	-0.110909	C	5.757500	-0.397967	1.942135
C	1.170526	-2.559379	2.153189	C	1.339162	-4.812598	1.318518
C	3.505444	0.207921	2.951731	H	1.725476	-4.915186	-0.823333
C	1.242195	0.726491	3.487976	H	4.932742	0.246478	4.599696
C	2.108644	3.125650	1.056074	C	2.914472	0.720120	5.227846
C	2.619336	2.432026	-1.102937	H	0.821080	1.141603	5.584658
C	-0.080665	0.860808	2.858287	C	3.046350	4.721544	-0.485121
N	-1.393530	0.711672	1.108832	C	-1.769542	1.662411	4.698256
C	4.482858	-0.645052	-0.482279	N	-4.049282	0.598935	-1.999384
C	4.383288	-0.170111	1.820213	N	-3.887100	1.087592	0.558974

N	-1.923560	2.525396	-1.319478	H	-3.338585	-1.331190	1.613520
N	-2.837575	-1.579559	-0.402673	C	-3.166989	-3.772006	-1.341348
C	-3.599533	1.287074	1.873022	H	-2.373196	-2.047174	-2.385948
H	6.401248	-1.165949	-1.321782	H	-4.433191	1.754328	3.828870
C	6.506497	-0.758315	0.815733	C	-5.936569	1.736748	2.271654
H	6.247219	-0.295574	2.913935	C	-6.402767	0.729100	-3.490782
H	1.330653	-5.896820	1.474811	H	-7.354934	1.254670	-1.618731
H	3.194114	0.849677	6.278840	H	-5.124373	0.208056	-5.173409
H	-2.053006	0.804926	5.336322	H	-7.225184	1.625525	0.534566
H	-2.624190	2.354457	4.673443	C	-1.588127	5.165006	-2.191535
H	-0.955799	2.200409	5.207347	C	-3.659900	-4.242961	-0.119535
C	-5.230502	0.923006	-1.382700	H	-3.096574	-4.424427	-2.217152
C	-4.022858	0.359100	-3.327814	H	-6.753675	1.992806	2.954420
C	-5.141460	1.193175	0.066196	H	-7.329154	0.779148	-4.072968
C	-2.137322	3.564804	-0.478972	H	-3.994982	-5.280296	-0.012379
C	-1.527915	2.783821	-2.590645	H	0.948190	-4.302310	3.400078
C	-3.298214	-2.038988	0.783189	H	-4.091218	-3.675547	1.938927
C	-2.766624	-2.441292	-1.445293	H	2.620562	1.618414	-1.833293
C	-4.631363	1.610759	2.765656	H	3.452036	3.885304	-2.455794
H	7.583309	-0.938292	0.903089	H	3.411773	5.726807	-0.721278
C	-6.417819	0.992548	-2.116215	H	2.496689	5.181185	1.572319
H	-3.051923	0.153486	-3.799868	H	1.720763	2.872717	2.044413
C	-5.187908	0.412277	-4.100130	H	-1.472986	6.197874	-2.538029
C	-6.204808	1.530155	0.913404	H	-1.040566	4.247796	-4.088919
C	-1.981915	4.890746	-0.877102	H	-1.357909	1.933370	-3.262744
C	-1.351709	4.088797	-3.052054	H	-2.450855	3.329413	0.539841
C	-3.716761	-3.356734	0.961454	H	-2.180751	5.694746	-0.162169

IT1

Ru	1.498160	0.157088	0.571334	C	1.675397	0.997034	4.788605
N	1.198091	-1.933922	0.826674	C	2.380855	4.510545	0.870762
N	2.152803	0.420382	2.522817	C	3.002340	3.795295	-1.337685
N	2.015251	2.225835	0.200839	C	-1.416653	1.291692	3.372861
N	-0.196822	0.605248	1.589731	Ru	-2.285224	0.342238	-0.634720
N	3.595196	-0.258387	0.425439	C	-2.217295	1.171464	2.206405
O	0.706623	0.198233	-1.095341	H	3.665035	-0.690993	-1.612468
C	1.367144	-2.732710	-0.254830	C	5.629520	-0.863366	-0.702289
C	0.961867	-2.513347	2.026972	C	5.655581	-0.386047	1.659494
C	3.469822	0.246006	2.789009	C	1.106861	-4.722294	1.077556
C	1.238436	0.788978	3.472607	H	1.484436	-4.723884	-1.068387
C	1.934039	3.213574	1.121089	H	4.998067	0.304091	4.339634
C	2.532762	2.523868	-1.014392	C	3.031461	0.816690	5.089153
C	-0.122458	0.916226	2.924754	H	0.976340	1.295321	5.572059
N	-1.429527	0.744046	1.170554	C	2.935654	4.811174	-0.377504
C	4.255174	-0.609819	-0.693837	C	-1.865143	1.721210	4.737397
C	4.279290	-0.142760	1.608275	N	-3.937729	0.330294	-2.006559
C	1.329087	-4.122782	-0.167336	N	-3.862560	1.056129	0.520747
H	1.533070	-2.233042	-1.212743	O	-0.771519	-0.467644	-1.310928
H	0.822607	-1.849593	2.883144	N	-1.789681	2.253419	-1.400421
C	0.911247	-3.896704	2.190285	N	-2.910641	-1.659201	-0.094517
C	3.942213	0.439010	4.092719	C	-3.620057	1.374609	1.829227

Appendix

H	6.123827	-1.145722	-1.636844	H	-4.503640	2.111529	3.673406
C	6.341485	-0.749701	0.494194	C	-5.947125	1.962412	2.060667
H	6.197734	-0.294366	2.604091	C	-6.239679	0.376116	-3.585247
H	1.090129	-5.812948	1.179528	H	-7.235722	1.115568	-1.811474
H	3.384744	0.974476	6.113766	H	-4.930296	-0.355322	-5.160598
H	-2.757391	1.159260	5.062228	H	-7.166325	1.715127	0.288071
H	-2.124187	2.795465	4.757711	C	-1.346013	4.732343	-2.618908
H	-1.094241	1.559371	5.502810	C	-3.833575	-4.265323	0.366702
C	-5.129343	0.748331	-1.468235	H	-3.241086	-4.621642	-1.700209
C	-3.881879	-0.045659	-3.298710	H	-6.780844	2.327307	2.669871
C	-5.088297	1.157919	-0.045538	H	-7.145153	0.393806	-4.201260
C	-1.309033	2.319092	-2.667574	H	-4.208458	-5.279427	0.542723
C	-2.030888	3.405819	-0.730907	H	0.731785	-4.315851	3.184997
C	-3.409683	-2.013937	1.110905	H	-4.281339	-3.532435	2.368247
C	-2.850983	-2.599223	-1.067148	H	2.564978	1.715037	-1.749585
C	-4.672091	1.841941	2.628592	H	3.423007	3.977752	-2.331202
H	7.419308	-0.941927	0.525553	H	3.309995	5.816886	-0.597400
C	-6.290595	0.778733	-2.244783	H	2.299854	5.269466	1.654965
H	-2.902589	-0.334067	-3.707176	H	1.502927	2.961135	2.091705
C	-5.018087	-0.038992	-4.116829	H	-1.191477	5.702738	-3.103162
C	-6.167063	1.619637	0.719624	H	-2.053565	5.559855	-0.730446
C	-1.076180	3.540690	-3.299474	H	-2.412400	3.320108	0.288965
C	-1.825848	4.657572	-1.305808	H	-1.115087	1.376330	-3.194265
C	-3.877998	-3.299944	1.378035	H	-0.698138	3.546236	-4.326084
H	-3.441181	-1.245590	1.886471	H	-0.439078	-0.933907	-4.857752
C	-3.301680	-3.903817	-0.876395	O	-1.075557	-0.206729	-4.759670
H	-2.426977	-2.290714	-2.026720	H	-1.228878	0.085582	-5.675054

I1

Ru	1.492115	0.096357	0.700267	C	1.655618	1.047271	4.805941
N	1.263440	-1.994109	1.052703	C	2.572448	4.378448	0.633114
N	2.141531	0.413432	2.562233	C	2.821989	3.544121	-1.608363
N	2.001740	2.091676	0.130515	C	-1.420179	1.288717	3.364624
N	-0.203345	0.602628	1.570923	Ru	-2.340483	0.407455	-0.662998
N	3.572830	-0.325279	0.471986	C	-2.223537	1.156169	2.185188
O	0.661539	-0.149932	-1.222323	H	3.646857	-0.804090	-1.552071
C	1.400104	-2.838473	-0.000355	C	5.613545	-0.945168	-0.640212
C	1.106486	-2.529250	2.288352	C	5.639508	-0.409010	1.711283
C	3.456937	0.242160	2.828999	C	1.256821	-4.773403	1.426829
C	1.220707	0.808088	3.494609	H	1.536551	-4.859671	-0.732355
C	2.112417	3.118178	1.008621	H	4.980372	0.339758	4.386186
C	2.343451	2.313915	-1.162801	C	3.012703	0.871807	5.114940
C	-0.134841	0.925705	2.925733	H	0.957366	1.367014	5.581683
N	-1.446243	0.738606	1.152136	C	2.946604	4.599254	-0.697373
C	4.238714	-0.698835	-0.636572	C	-1.887040	1.717498	4.721447
C	4.261368	-0.175865	1.654687	N	-3.965325	0.416999	-2.033854
C	1.406350	-4.223547	0.148410	N	-3.870062	1.056601	0.505323
H	1.511163	-2.379093	-0.985973	O	-0.624182	-0.288287	-1.408585
H	0.993453	-1.834848	3.123605	N	-1.879912	2.371673	-1.355725
C	1.093464	-3.904279	2.511819	N	-2.861817	-1.628970	-0.229644
C	3.925301	0.469261	4.132542	C	-3.627322	1.353822	1.821286

H	6.111591	-1.247078	-1.566562	H	-4.500849	2.036183	3.687511
C	6.325397	-0.797612	0.554304	C	-5.960444	1.910319	2.086648
H	6.183690	-0.289297	2.651638	C	-6.281039	0.474917	-3.595104
H	1.270347	-5.858754	1.575663	H	-7.275881	1.141905	-1.792324
H	3.362085	1.054502	6.136551	H	-4.971757	-0.186395	-5.200444
H	-2.728121	1.093690	5.071236	H	-7.182037	1.691437	0.315088
H	-2.233718	2.766847	4.714549	C	-1.501259	4.909844	-2.484671
H	-1.097609	1.643084	5.480896	C	-3.698784	-4.286392	0.096077
C	-5.159738	0.797489	-1.477289	H	-3.055258	-4.527015	-1.971941
C	-3.913940	0.082776	-3.336558	H	-6.793324	2.251082	2.710409
C	-5.103482	1.162313	-0.046819	H	-7.190662	0.497693	-4.204606
C	-1.419091	2.501234	-2.624416	H	-4.040190	-5.319879	0.218787
C	-2.133667	3.493801	-0.643638	H	0.968434	-4.283911	3.530500
C	-3.371446	-2.056799	0.947386	H	-4.213857	-3.660348	2.117261
C	-2.750640	-2.522470	-1.240826	H	2.225521	1.474794	-1.853733
C	-4.677312	1.790938	2.638197	H	3.099792	3.665158	-2.659774
H	7.404532	-0.982115	0.589079	H	3.330566	5.574240	-1.016891
C	-6.328582	0.834439	-2.242373	H	2.646278	5.171368	1.383728
H	-2.932809	-0.178033	-3.758816	H	1.827193	2.924306	2.044616
C	-5.056715	0.095948	-4.146720	H	-1.372477	5.901162	-2.932528
C	-6.180826	1.596205	0.742392	H	-2.194875	5.646582	-0.555158
C	-1.220042	3.750186	-3.214271	H	-2.500430	3.364542	0.377255
C	-1.959045	4.772757	-1.169875	H	-1.214811	1.584124	-3.191457
C	-3.798252	-3.369505	1.147731	H	-0.859184	3.802709	-4.245769
H	-3.447610	-1.324263	1.754534	H	-0.447962	-0.719918	-4.952423
C	-3.156714	-3.850158	-1.118073	O	-1.091997	-0.010061	-4.794073
H	-2.325077	-2.154960	-2.179254	H	-1.273883	0.331485	-5.686878

IT2

Ru	2.209567	-0.015154	-0.170369	N	-2.348230	2.170705	0.338178
Ru	-2.314710	0.081938	-0.135191	C	-1.178613	-0.661343	2.451367
N	3.163250	-0.089634	1.588461	N	-4.308382	0.315513	-0.796037
O	0.166336	0.019364	-1.786822	C	-2.633466	-0.685520	2.656868
O	-1.026596	0.410674	-1.562223	N	-2.436099	-1.959305	-0.798165
C	4.498485	0.169145	1.638163	C	-3.335944	-0.960986	3.840320
N	0.593104	-0.292212	1.189842	H	-2.807055	-1.201724	4.763184
C	5.197869	0.024595	2.843279	C	-4.735345	-0.911417	3.829923
H	6.270027	0.228547	2.892327	H	-5.292792	-1.126814	4.747498
N	-0.745332	-0.306883	1.198084	C	-5.426760	-0.584117	2.657362
C	4.505487	-0.381854	3.986262	H	-6.518983	-0.546316	2.655883
H	5.035930	-0.499601	4.936939	C	-4.693741	-0.301347	1.498350
N	-3.337599	-0.361683	1.535185	H	-2.498710	-1.317776	-2.787336
C	3.130324	-0.624930	3.917412	C	1.467349	-2.954443	-0.436050
H	2.580724	-0.911779	4.814094	H	0.487237	-2.475720	-0.507821
N	2.526292	-2.133871	-0.240779	C	1.605039	-4.336200	-0.549664
C	2.461294	-0.477381	2.693662	H	0.718067	-4.955782	-0.709088
N	4.203005	0.588723	-0.697134	C	2.879833	-4.908137	-0.462351
C	1.028507	-0.612070	2.457610	H	3.017985	-5.991211	-0.550985
N	1.893926	2.090239	-0.227693	C	3.975318	-4.061700	-0.264476
C	-0.066027	-0.889759	3.297565	H	4.994164	-4.454907	-0.194617
C	-0.056437	-1.335840	4.729115	C	3.759059	-2.688157	-0.158785

Appendix

C	4.634951	1.029708	-1.896700	C	-6.609310	0.190258	-0.097630
C	5.947507	1.458649	-2.111010	H	-7.349246	-0.002020	0.683506
H	6.245949	1.795050	-3.108806	C	-7.028946	0.551199	-1.383560
C	6.853157	1.441980	-1.047454	H	-8.096458	0.642558	-1.610850
H	7.889710	1.768129	-1.183843	C	-6.065587	0.790285	-2.366631
C	6.405997	1.007364	0.203505	H	-6.347180	1.074182	-3.385274
H	7.089587	1.001736	1.056165	C	-4.714457	0.662512	-2.030945
C	5.076712	0.596250	0.358337	C	-2.570340	-2.187064	-2.127443
C	1.744233	2.694919	-1.429280	C	-2.795007	-3.457066	-2.656302
C	1.677988	4.080801	-1.564023	H	-2.904465	-3.579494	-3.738361
H	1.571782	4.520939	-2.560269	C	-2.889737	-4.549460	-1.786865
C	1.766966	4.881459	-0.419607	H	-3.084223	-5.556798	-2.170895
H	1.738974	5.973897	-0.496715	C	-2.735771	-4.321165	-0.414855
C	1.902587	4.255240	0.824047	H	-2.803941	-5.139662	0.308594
H	1.978652	4.835708	1.748447	C	-2.509825	-3.021017	0.038122
C	1.962994	2.863023	0.880114	H	-2.395578	-2.823711	1.106431
H	2.084330	2.350460	1.838080	O	2.390340	-0.529529	-2.654985
C	-2.379430	2.658412	1.599385	H	1.521775	-0.595704	-3.091265
H	-2.276145	1.940163	2.415850	H	2.928209	-1.265662	-2.994649
C	-2.548164	4.015018	1.876547	H	-2.420525	2.633848	-1.698663
H	-2.578464	4.352945	2.916920	H	4.603971	-2.010882	-0.016269
C	-2.691675	4.915239	0.815916	H	3.910969	1.009180	-2.711721
H	-2.845369	5.983504	1.003218	H	1.681846	2.040245	-2.302724
C	-2.638694	4.416428	-0.490620	H	-3.934032	0.844847	-2.777314
H	-2.742952	5.076814	-1.356961	H	-0.937692	-1.947672	4.969329
C	-2.464582	3.047721	-0.687076	H	-0.040410	-0.483443	5.433682
C	-5.243025	0.075785	0.177795	H	0.816263	-1.968678	4.949191

12

Ru	2.371636	-0.032431	-0.269344	N	-2.323285	2.146720	0.208410
Ru	-2.288732	0.081737	-0.388700	C	-1.083719	-0.654671	2.187932
N	3.263353	0.055711	1.496639	N	-4.284743	0.319502	-0.987021
O	-0.827787	0.108351	-3.088012	C	-2.530047	-0.845263	2.363911
O	-1.361623	0.568363	-2.033281	N	-2.400455	-1.934548	-1.119067
C	4.599306	0.295736	1.529412	C	-3.197370	-1.342667	3.493754
N	0.709394	-0.248283	0.976775	H	-2.642993	-1.677322	4.371288
C	5.275563	0.271158	2.756828	C	-4.596326	-1.423982	3.477437
H	6.348936	0.469385	2.813431	H	-5.129024	-1.811722	4.352068
N	-0.630218	-0.369351	0.925226	C	-5.315463	-1.012679	2.350779
C	4.553947	-0.039448	3.915665	H	-6.406923	-1.067506	2.345296
H	5.068207	-0.084866	4.881321	C	-4.611543	-0.543457	1.232552
N	-3.256588	-0.488128	1.269655	H	-2.980056	-1.272141	-3.018764
C	3.180998	-0.304256	3.847228	C	1.703909	-2.941805	-0.778280
H	2.633670	-0.568791	4.752259	H	0.763933	-2.447017	-1.037664
N	2.705925	-2.143623	-0.329461	C	1.846879	-4.320709	-0.906925
C	2.530461	-0.230568	2.603454	H	1.008994	-4.917186	-1.278848
N	4.341931	0.348861	-0.866664	C	3.065761	-4.918778	-0.563252
C	1.106304	-0.404072	2.306937	H	3.209880	-6.000120	-0.662927
N	2.124677	2.082176	-0.404643	C	4.096661	-4.100417	-0.091006
C	-0.008765	-0.659222	3.115261	H	5.068604	-4.515906	0.192317
C	-0.069168	-0.836687	4.603111	C	3.879971	-2.727436	0.013202



C	4.804526	0.504628	-2.123834	C	-6.569983	0.014445	-0.286483
C	6.135643	0.823900	-2.394250	H	-7.289132	-0.295522	0.476005
H	6.469126	0.933119	-3.430964	C	-7.023469	0.500682	-1.518441
C	7.018342	1.000437	-1.322426	H	-8.096929	0.570174	-1.724074
H	8.067973	1.258760	-1.499982	C	-6.084477	0.895004	-2.474329
C	6.543257	0.842873	-0.015091	H	-6.391963	1.282520	-3.450487
H	7.220131	0.980928	0.832271	C	-4.723938	0.787815	-2.169906
C	5.200144	0.514557	0.196994	C	-2.811972	-2.150597	-2.392359
C	2.062418	2.672394	-1.624235	C	-3.024320	-3.427535	-2.911697
C	2.024723	4.056017	-1.784536	H	-3.362192	-3.537585	-3.946615
H	1.989370	4.479703	-2.792932	C	-2.815074	-4.540067	-2.091316
C	2.047991	4.875524	-0.650502	H	-2.993689	-5.554389	-2.464511
H	2.037291	5.966840	-0.746825	C	-2.377874	-4.322781	-0.779052
C	2.093124	4.267842	0.609617	H	-2.204472	-5.157701	-0.093092
H	2.113360	4.862869	1.527681	C	-2.177645	-3.016316	-0.334897
C	2.132514	2.877603	0.692672	H	-1.836961	-2.827955	0.685569
H	2.179551	2.383756	1.665587	O	1.672355	-0.254056	-2.341591
C	-2.354403	2.549145	1.499777	H	0.741918	-0.099364	-2.673500
H	-2.252224	1.777166	2.265241	H	2.000779	-1.031798	-2.824268
C	-2.521896	3.882826	1.871644	H	-2.403128	2.754668	-1.789292
H	-2.552474	4.146593	2.933224	H	4.678432	-2.076088	0.372806
C	-2.664200	4.855767	0.877270	H	4.079531	0.369288	-2.931784
H	-2.815719	5.908626	1.138594	H	2.052403	2.007071	-2.490286
C	-2.614041	4.447455	-0.459889	H	-3.963052	1.083778	-2.898841
H	-2.718802	5.165508	-1.279126	H	-1.043728	-0.531197	5.010107
C	-2.443215	3.095266	-0.751649	H	0.680555	-0.216895	5.116462
C	-5.196089	-0.067221	-0.038244	H	0.104054	-1.886563	4.904123

WT1

Ru	2.292652	-0.349396	0.299638	C	2.049490	0.451197	4.508300
N	2.093413	-2.445338	0.753880	C	3.020033	4.018400	0.371729
N	2.758524	-0.002652	2.282481	C	3.472920	3.262735	-1.863562
N	2.668471	1.708253	-0.209805	C	-0.820820	0.873642	2.926027
N	0.387350	0.124399	1.142949	Ru	-2.020548	0.084060	-1.100681
N	4.385035	-0.615988	0.325908	C	-1.628065	0.730927	1.776394
O	1.892312	-0.714196	-1.423237	H	4.623648	-0.957175	-1.714183
C	2.649334	-3.338085	-0.100343	C	6.526743	-1.059120	-0.662329
C	1.507850	-2.913999	1.879887	C	6.349505	-0.668705	1.709728
C	4.047397	-0.173437	2.671147	C	2.068992	-5.190508	1.324113
C	1.754035	0.315593	3.142376	H	3.123138	-5.385259	-0.573355
C	2.635024	2.717173	0.691327	H	5.417716	-0.153608	4.362522
C	3.073370	1.986343	-1.472781	C	3.372584	0.281594	4.935603
C	0.450387	0.453591	2.479389	H	1.267274	0.671862	5.235350
N	-0.867879	0.284583	0.717393	C	3.457620	4.301375	-0.926262
C	5.143592	-0.882509	-0.754133	C	-1.210906	1.393763	4.278077
C	4.969616	-0.499188	1.560141	N	-3.750800	0.248776	-2.300393
C	2.655044	-4.708910	0.148394	N	-3.424743	0.779901	0.238679
H	3.101139	-2.935448	-1.009993	O	-1.016319	-0.460681	-2.528913
H	1.051070	-2.181186	2.547907	N	-1.567557	2.110763	-1.703695
C	1.480267	-4.270824	2.198580	N	-2.597433	-1.926204	-0.612869
C	4.388136	-0.025792	4.019501	C	-3.059817	0.962442	1.534879

Appendix

H	7.105630	-1.275571	-1.565472	H	-3.785955	1.451491	3.528140
C	7.139326	-0.953618	0.589061	C	-5.356398	1.513810	2.044312
H	6.813067	-0.575239	2.695217	C	-6.143558	0.596269	-3.699646
H	2.072043	-6.261562	1.554090	H	-6.967413	1.208936	-1.795366
H	3.615794	0.386670	5.998400	H	-4.992619	-0.058867	-5.429427
H	-1.627885	0.600854	4.926436	H	-6.729942	1.473678	0.370649
H	-1.966483	2.190368	4.200703	C	-1.241644	4.739003	-2.638077
H	-0.355541	1.838478	4.806540	C	-3.531582	-4.526032	-0.107694
C	-4.871282	0.682881	-1.642489	H	-3.547337	-4.718784	-2.277557
C	-3.813731	-0.001608	-3.621674	H	-6.127661	1.798970	2.767524
C	-4.695345	0.947786	-0.199617	H	-7.083476	0.733088	-4.244853
C	-1.249241	2.357255	-2.996486	H	-3.908538	-5.535301	0.089149
C	-1.702426	3.168315	-0.870159	H	1.006800	-4.596098	3.129808
C	-2.773559	-2.369053	0.653663	H	-3.376190	-3.950802	1.986476
C	-2.873524	-2.779343	-1.627908	H	3.065785	1.153761	-2.180797
C	-4.036207	1.325961	2.474493	H	3.800947	3.430224	-2.894072
H	8.221160	-1.087439	0.696179	H	3.782428	5.310745	-1.201265
C	-6.078387	0.860386	-2.327053	H	2.985338	4.794418	1.142518
H	-2.889950	-0.344363	-4.097210	H	2.299265	2.476393	1.701692
C	-4.990745	0.159309	-4.357033	H	-1.133269	5.764810	-3.006666
C	-5.698001	1.331294	0.700023	H	-1.683184	5.303357	-0.581192
C	-1.080680	3.647168	-3.497193	H	-1.949646	2.955884	0.171664
C	-1.548654	4.486725	-1.296969	H	-1.122759	1.486634	-3.644835
C	-3.238147	-3.650922	0.943324	H	-0.835967	3.785306	-4.554844
H	-2.540296	-1.673409	1.461540	O	-0.016992	-1.790473	-2.396962
C	-3.335796	-4.076647	-1.417299	H	0.972141	-1.258340	-1.949794
H	-2.718310	-2.406979	-2.641892	H	0.126352	-1.965771	-3.350661

W1

Ru	2.310695	-0.311218	0.308380	C	2.040325	0.288634	4.532512
N	2.143492	-2.420319	0.661990	C	2.944794	4.062500	0.538177
N	2.765215	-0.050545	2.286376	C	3.379764	3.403643	-1.731757
N	2.637548	1.769489	-0.128207	C	-0.807669	0.837689	2.939607
N	0.399848	0.093461	1.153679	Ru	-2.042291	0.051296	-1.076361
N	4.411482	-0.516372	0.319037	C	-1.603961	0.736183	1.778957
O	1.900391	-0.612137	-1.465332	H	4.660774	-0.699384	-1.742185
C	2.793139	-3.265341	-0.176186	C	6.562434	-0.839925	-0.692037
C	1.472613	-2.946597	1.713531	C	6.372279	-0.629699	1.703136
C	4.055534	-0.233743	2.679452	C	2.134665	-5.185835	1.117615
C	1.751989	0.223414	3.157775	H	3.353100	-5.281192	-0.687529
C	2.594996	2.741694	0.812940	H	5.411024	-0.292660	4.377290
C	3.015504	2.104017	-1.385911	C	3.357315	0.097330	4.960265
C	0.458452	0.394386	2.498996	H	1.250616	0.470912	5.262407
N	-0.846652	0.286330	0.716333	C	3.355271	4.404851	-0.754551
C	5.176541	-0.684232	-0.776969	C	-1.201202	1.341519	4.296304
C	4.989898	-0.476085	1.560043	N	-3.746003	0.240488	-2.283089
C	2.809179	-4.644550	0.017196	N	-3.402690	0.797767	0.246057
H	3.312959	-2.818278	-1.027072	O	-1.018064	-0.693880	-2.529749
H	0.940688	-2.249929	2.364380	N	-1.508818	2.003757	-1.795450
C	1.450707	-4.315452	1.973664	N	-2.687655	-1.926561	-0.517178
C	4.382423	-0.156366	4.034156	C	-3.031884	0.996747	1.537360

*Cartesian Structures of Intermediates and Transition States*

H	7.148201	-0.977389	-1.606294	H	-3.736888	1.552748	3.520897
C	7.169340	-0.816735	0.567032	C	-5.312752	1.612539	2.043492
H	6.831827	-0.598109	2.694519	C	-6.125632	0.644985	-3.689974
H	2.142216	-6.265530	1.303420	H	-6.938480	1.289391	-1.791597
H	3.593175	0.145164	6.028936	H	-4.988268	-0.047831	-5.413789
H	-1.712118	0.566279	4.897029	H	-6.691502	1.568415	0.375370
H	-1.881187	2.204356	4.221072	C	-1.094461	4.552426	-2.899090
H	-0.332993	1.688758	4.873701	C	-3.709089	-4.476775	0.082614
C	-4.856661	0.709352	-1.630607	H	-3.857047	-4.697118	-2.079435
C	-3.811863	-0.012666	-3.603527	H	-6.074329	1.930309	2.763100
C	-4.671185	0.987581	-0.191821	H	-7.060248	0.802893	-4.238386
C	-1.104098	2.149890	-3.079061	H	-4.119162	-5.464861	0.316796
C	-1.686007	3.121743	-1.053460	H	0.905572	-4.689573	2.845459
C	-2.815683	-2.344440	0.762888	H	-3.413813	-3.877724	2.154457
C	-3.058479	-2.782523	-1.499095	H	3.010418	1.301355	-2.127710
C	-3.994448	1.405049	2.472137	H	3.686361	3.618851	-2.760050
H	8.253470	-0.937438	0.667872	H	3.651745	5.431650	-0.995089
C	-6.057511	0.914304	-2.318613	H	2.903707	4.807580	1.338642
H	-2.895988	-0.385177	-4.072147	H	2.281742	2.454543	1.818352
C	-4.982542	0.175857	-4.342615	H	-0.953396	5.547048	-3.335599
C	-5.662103	1.409607	0.705276	H	-1.664220	5.272599	-0.923882
C	-0.889694	3.398344	-3.661930	H	-2.007171	2.987044	-0.018411
C	-1.490331	4.403683	-1.565626	H	-0.948168	1.229926	-3.647985
C	-3.319293	-3.600097	1.100260	H	-0.577601	3.455872	-4.709158
H	-2.510945	-1.647702	1.545895	O	-0.198373	-1.871124	-2.258189
C	-3.567069	-4.054406	-1.242780	H	1.129354	-1.178357	-1.819237
H	-2.948757	-2.430399	-2.526680	H	-0.129789	-2.272804	-3.150222



# Curriculum Vitae

## Personal Data

Name Thorsten Stolper  
Nationality German  
Date of birth 04.11.1986  
Place of birth Sulz am Neckar  
Address Föhrenweg 4, 37077 Göttingen  
E-mail tstolpe@gwdg.de

## Education

2013–2017 Doctoral studies in the Computational Chemistry and Biochemistry group of Prof. Dr. Ricardo A. Mata at the Georg-August University of Göttingen, Germany.  
2010–2013 Master's studies in Chemistry at the Georg-August University of Göttingen, Germany.  
Thesis: *Hybrid QM/QM approaches using local explicitly correlated methods*  
2006–2010 Bachelor's studies in Chemistry at the Georg-August University of Göttingen, Germany.  
Thesis: *Optimierungsverfahren für die Untersuchung organischer Dimere*  
1999–2006 High School "Gymnasium Bad Zwischenahn–Edeweicht", Abitur (A-levels).

## Publications

1 A. Wilting, T. Stolper, R. A. Mata, I. Siewert, *Inorganic Chemistry* **2017**, *56*, 4176–4185.