

DEVELOPMENT OF HYBRID QM/QM LOCAL CORRELATION METHODS FOR THE STUDY OF METAL SITES IN BIOMOLECULAR CATALYSIS

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

within the doctoral program of IRTG1422
of the Georg-August University School of Science (GAUSS)

submitted by
Milica Andrejić
from Niš, Serbia

Göttingen, 2015

Thesis Committee

Jun.-Prof. Dr. Ricardo Mata

Institute for Physical Chemistry, Georg-August-Universität Göttingen

Prof. Dr. Ulf Ryde

Department of Theoretical Chemistry, Lund University

Prof. Dr. Guido Clever

Institute for Inorganic Chemistry, Georg-August-Universität Göttingen

Members of the Examination Board

Reviewer:

Jun.-Prof. Dr. Ricardo Mata

Institute for Physical Chemistry, Georg-August-Universität Göttingen

Second Reviewer:

Prof. Dr. Ulf Ryde

Department of Theoretical Chemistry, Lund University

Further members of the Examination Board

Prof. Dr. Guido Clever

Institute for Inorganic Chemistry, Georg-August-Universität Göttingen

Prof. Dr. Franc Meyer

Institute for Inorganic Chemistry, Georg-August-Universität Göttingen

Prof. Dr. Burkhard Geil

Institute for Physical Chemistry, Georg-August-Universität Göttingen

Priv.-Doz. Dr. Thomas Zeuch

Institute for Physical Chemistry, Georg-August-Universität Göttingen

Date of the oral examination: 27.03.2015

Declaration

I hereby declare that I have created this thesis without the help of third parties and by my own efforts. No other than the referenced sources have been employed and knowledge or ideas taken directly or indirectly from prior existing works were referenced as such.

Milica Andrejić

Contents

Acknowledgements	XI
Abstract	XV
Citations to Published Work	XIX
Abbreviations	XXIII
1 Introduction	XXVII
2 Theoretical Background	3
2.1 Second Quantization	5
2.2 Møller-Plesset Perturbation Theory	10
2.3 Coupled Cluster Theory	13
2.3.1 Coupled Cluster Theory with Singles and Doubles	14
2.3.2 Coupled Cluster Theory in Diagrammatic Representation	20
2.3.3 Perturbative Triple Excitations	35
2.4 Local Correlation Methods	35
2.4.1 Selection of domains and orbital pairs	37
2.4.2 Local Møller-Plesset Method	39
2.4.3 Local Coupled Cluster Method	41
2.5 Quantum Mechanics/Quantum Mechanics Method	42
2.6 Density Functional Theory	43
3 Auophilic Complexes	47
3.1 Ligand Effects in Auophilic Interactions	50
3.1.1 Computational Details	52
3.1.2 Results and Discussion	54
3.1.3 Final Remarks	69
4 Molybdenum Enzymes	71
4.1 Reaction pathways for SO	74
4.1.1 Computational details	78
4.1.2 Results and Discussion	81
4.1.3 Final Remarks	92

4.2	LMOMO Calculations	93
4.2.1	Computational details	94
4.2.2	LMOMO method for SO	95
4.2.3	LMOMO method for DMSOR	102
4.2.4	Final Remarks	107
5	Open Shell Systems	109
5.1	Open-Shell Test Systems	111
5.1.1	Computational Details	112
5.1.2	Benchmark Results	113
5.2	Nitrite Reductase	117
5.2.1	Computational Details	120
5.2.2	Results and Discussion	122
5.3	Final Remarks	129
6	Conclusions	131
	Bibliography	137
	Curriculum Vitae	149

Acknowledgements

First of all, I would like to thank Prof. Dr. Mata for welcoming me and for introducing me to the for me new world of theoretical chemistry. I am also very thankful for his encouragement, guidance and patience.

My sincere thanks to Prof. Dr. Ulf Ryde for the nice welcoming in his group in Lund University. For giving me the opportunity to work also on topics which were not closely related to my thesis and I would like to thank Dr. Pär Söderhjelm for working with me on these topics.

I acknowledge the financial support from the International Research Training Group 1422 - Metal Sites in Biomolecules: Structures, Regulation and Mechanisms.

My thanks to the past and current members of the Computational Chemistry and Biochemistry Group as well as to all members of the theoretical department of Lund University. For their help in numerous situations and for creating a pleasant and enthusiastic environment for everyday work. With special thanks to Thorsten Stolper, since he always positively replied on my question "Can I ask you something?" even when he knew that something nagging is coming up and for sharing his expertise with me. Thanks go also to Martin Werner, since he always was in the mood for board games which I enjoy very much. The biggest thanks go to Jonas Feldt, for solving all my L^AT_EX problems and more importantly for keeping me motivated during the writing of my thesis. And also to Axel Wuttke for working with me on the aurophilic complexes.

I would also like to thank to Diego, Joao and Fernanda for all the laughs and the great moments that we had together and for understanding how hard it can be sometimes to be abroad.

Last but not least, I would like to thank my parents ("Hvala vam, dragi roditelji, što ste me rodili ovako lepu i pametnu!") and especially to my brother and his sentence "Uozbilji se, deportovaće te". For their unconditional support.

Abstract

Metal-containing systems present a challenge for both theoretical and experimental chemists. Joint efforts are required to fully understand their complex electronic structure and their role in catalysis. On the side of theory, accuracy is essential but the cost of wave function methods is deterrent to this goal.

The emphasis of this work is on the development of a QM/QM method for biomolecules which can reach wavefunction accuracy. To properly describe the electron correlation in transition metals inclusion of higher-order excitations is required. However, some of these effects are relatively local in nature (regularly connected with the metal center) and hybrid or embedding approaches offer a cost-effective alternative for their computation. In the proposed QM/QM scheme, localized orbitals are used to split the system into different groups. This allows for high accuracy in regions where bond breaking/formation takes place, while the remaining environment is described at a low level. Coupled cluster and MP2 approaches can be combined in a single calculation, without resource to model systems.

In the present thesis, this QM/QM method will be applied for the first time in the treatment of closed-shell metal systems and for open-shell systems in general. The application on different complexes containing a wide variety of metals will be presented. It will be shown that this method provides valuable insights into the aurophilic interaction. On the example of molybdenum enzymes, it will be shown that high accuracy for the calculation of reaction pathways can be obtained with significantly reduced computational costs. On the example of nitrite reductase, it will also be shown that properties, such as electron affinities, can be reliably calculated.

Citations to Published Work

Most of the Chapter "Auophilic Complexes" has been published or it will be published in the following papers:

"Study of ligand effects in auophilic interactions using local correlation methods",
M. Andrejić and R. A. Mata, Phys. Chem. Chem. Phys. **15**, 18115-18122 (2013)

"Competing effects in crystal formation - auophilic vs π - π interactions",
A. Wuttke, M. Andrejić and R. A. Mata, to be submitted

The Chapter "Molybdenum Enzymes" has been published in:

"A quantum-mechanical study of the reaction mechanism of sulfite oxidase",
M. van Severen, M. Andrejić, J. Li, K. Starke, R. A. Mata, E. Nordlander,
U. Ryde, Journal of Biological Inorganic Chemistry **19**, 1165-1179 (2014)

"Local hybrid QM/QM calculations of reaction pathways in metollobiosites",
M. Andrejić and R. A. Mata, J. Chem. Theory Comput. **10**, 5397-5404 (2014)

The Chapter "Open Shell Systems" will be published in:

"Local hybrid QM/QM calculations for open shell systems",
M. Andrejić and R. A. Mata, to be submitted

Abbreviations

BSSE	Basis Set Superposition Error
CBS	Complete Basis-Set Limit
CC	Coupled Cluster
CCSD	Coupled Cluster with Singles and Doubles
CCSDT	Coupled Cluster with Singles, Doubles and Triples
CCSD(T)	Coupled Cluster with Singles, Doubles and Perturbative Triples
COSMO	Conductor-like Screening Model
CP	Counterpoise
DF	Density-fitting
DFT	Density Functional Theory
DFT-NL	DFT non-local density dependent dispersion corrections
DMDT	1,2-dimethyldithiolene
DMS	Dimethyl Sulfide
DMSO	Dimethyl Sulfoxide
DMSOR	Dimethyl Sulfoxide Reductase
ECP	Effective Core Potential
ee	edge-to-edge
ff	face-to-face
GGA	Generalized Gradient Approximation
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
IM	Intermediate
IMOMO	Integrated Molecular Orbital Molecular Orbital
KS	Kohn-Sham
LCCSD(T)	Local Coupled Cluster with Singles, Doubles and Perturbative Triples
LCCSD(T0)	Local Coupled Cluster with Singles, Doubles and non-iterative Perturbative Triples
LDA	Local Density Approximation
LI	Linear
LMO	Localized Molecular Orbital
LMOMO	Localized Molecular Orbital Molecular Orbital
LMP2	Local Møller-Plesset Perturbation Theory of Second Order

LRMP2	Local Restricted Møller-Plesset Perturbation Theory of Second Order
LUCCSD(T0)	Local Unrestricted Coupled Cluster with Singles, Doubles and non-iterative Perturbative Triples
LUMO	Lowest Unoccupied Molecular Orbital
Me ₂ bimy	1,3-bismethylbenzimidazole-2-ylidene
MM	Molecular Mechanics
MNT	maleonitrile-1,2-dithiolate
Moco	Molybdopterin cofactor
MP	Møller-Plesset Perturbation Theory
MP2	Møller-Plesset Perturbation Theory of Second Order
NHC-H	imidazol-2-ylidene
NHC-B	1-(benzyl)-3-(N- <i>tert</i> -butylacetamido)imidazole-2-ylidene
NAO	Natural Atomic Orbital
NiR	Nitrate Reductase
NLMO	Natural Localized Molecular Orbitals
NPA	Natural Population Analysis
OAT	Oxygen-Atom Transfer
OSV	Orbital Specific Virtuals
PAO	Projected Atomic Orbitals
PCM	Polarizable Continuum Model
PES	Potential Energy Surface
PNO	Pair Natural Orbital
PS	Product State
QCISD	Quadratic Configuration Interaction with Singles and Doubles
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
QM/QM	Quantum Mechanics/Quantum Mechanics
RI	Resolution-of-Identity
RMP2	Restricted Møller-Plesset Perturbation Theory of Second Order
RMSD	Root Mean Square Deviation
RCCSD(T)	Restricted Coupled Cluster with Singles, Doubles and Perturbative Triples

RS	Reactant State
RSPT	Rayleigh-Schrödinger Perturbation Theory
SCF	Self-consistent Field Method
SCS	Spin Component Scaled
SD	Slater Determinant
SO	Sulfite Oxidase
SP	Square Pyramidal
SQ	Square Planar
TDDFT	Time Dependent DFT
TH	Tetrahedral
TS	Transition State
UCCSD(T)	Unrestricted Coupled Cluster with Singles, Doubles and Perturbative Triples
XO	Xanthine Oxidase
ZPE	Zero Point Energy

CHAPTER 1

Introduction

The role of metals in biological processes is receiving ever-increasing recognition. Over one third of all proteins contain at least one metal ion as an essential prosthetic group. It is generally known that these metallobiomolecules are of key importance for a large number of biological reactions and phenomena, e.g. electron transfer and enzymatic catalysis in oxidation. A comprehensive understanding of the regulation and molecular mechanism of metallobiomolecules function is not only of fundamental interest, but is of great importance for the various industries, e.g. medical, biotechnological, environmental etc. Metalloenzymes are capable of carrying out energetically difficult chemical transformations at ambient conditions and the understanding of these processes can lead to their application in various technological contexts.

One way to gain a detailed understanding of metallobiomolecules is through theoretical calculations. The method which is widely used in the computational research of such species is density functional theory (DFT), since it is cheaper than wavefunction alternatives and it can compete in accuracy. However, this method is less suited for the description of the organic substrates, their weak interactions with metal centers and reactivity, due to the approximate correlation functionals used. Therefore, further development of wavefunction methods is necessary, since electronic correlation in transition metals is particularly difficult to describe. However, some of these effects are relatively local in nature (regularly connected with the metal centers) and hybrid or embedding approaches offer a cost-effective alternative for their computation.

This thesis is structured as follows. In Chapter 2 theoretical methods which were used throughout the thesis are shortly reviewed. Since the derivation of coupled cluster equations can be cumbersome, diagrammatic representation is presented and applied in the context of coupled cluster theory.

In Chapter 3, theoretical studies on the aurophilic effect are presented. The latter have been based on local correlation methods, which have been used to partition the correlation energy and to inspect the contribution of the Au...Au interaction to the correlation energy in a variety of dimers.

In Chapter 4 molybdenum enzymes were investigated. This chapter is divided into two parts. In the first part three possible reaction mechanisms for sulfite oxidase were investigated using the LCCSD(T0) method. In the second part the LMOMO method was evaluated on the example of sulfite oxidase as well as dimethyl sulfoxide reductase.

In Chapter 5 LMOMO calculations on open-shell systems are presented. The chapter starts with the application of the LMOMO method on small benchmark systems. In

the second part of the chapter the LMOMO method was applied on the copper nitrite reductase and calculation of electron and proton affinities in the system.

CHAPTER 2

Theoretical Background

In the Hartree-Fock (HF) theory the motion of electrons is described in the average field of the other electrons. The instantaneous interaction between electrons due to their mutual repulsion is not included. The purpose of all post-Hartree-Fock methods is to describe this omitted interaction. The latter is called electron correlation and represents the difference between the HF description and the exact solution. The correction to the energy is called correlation energy and is strictly defined as

$$\Delta E_{corr} = E_{exact} - E_{HF}. \quad (2.1)$$

The correction to the wave function which is present in some of post-HF methods determines the electron density and all other properties of molecules:

$$\Psi_{exact} = \Phi_{HF} + \chi_{corr}. \quad (2.2)$$

However, this type of the correlation is not the only one which is not properly described in the HF theory. Two types of the correlation which are lacking in the HF theory exist: dynamic and static correlation. Dynamic correlation reflects the instantaneous correlation in the motion of electrons. Static correlation can be present in the case of excited states and other open-shell systems, and sometimes even in closed-shell systems where bonds are stretched near breaking. These systems cannot be properly described using a single-configuration description. To properly deal with these two types of electron correlation two different treatments can be applied. In the case of dynamic correlation one can use perturbation methods and coupled cluster theory. In the case of static correlation one reaches for multireference methods. In the following text perturbation methods and coupled cluster theory will be described.

2.1 Second Quantization

Before we derive the equations for post-Hartree-Fock methods we will introduce the notation which will be used throughout. In the standard formulation of quantum mechanics (QM) observables are represented by operators and states by functions. In the formalism of second quantization [1–3], states are also represented by operators. Every state can be expressed as a combination of creation and annihilation operators working on the vacuum state. In this way a compact and convenient representation is introduced

and important relationships can be developed in an elegant manner.

A creation operator \hat{a}_i^\dagger can be associated with a spin orbital χ_i and it is defined as:

$$\hat{a}_i^\dagger |\chi_k \cdots \chi_l\rangle = |\chi_i \chi_k \cdots \chi_l\rangle \quad (2.3)$$

where $|\chi_k \cdots \chi_l\rangle$ represents an arbitrary Slater determinant (SD). Thus \hat{a}_i^\dagger creates an electron in the spin orbital χ_i . On the other hand, an annihilation operator \hat{a}_i destroys an electron from the spin orbital χ_i from an arbitrary SD:

$$\hat{a}_i |\chi_i \chi_k \cdots \chi_l\rangle = |\chi_k \cdots \chi_l\rangle. \quad (2.4)$$

It is important to notice that an annihilation operator can only act on a determinant if the spin orbital, which will disappear, is immediately on the left. In the case where this is not fulfilled the spin orbital should be properly reordered by interchanging the columns of the determinant:

$$\hat{a}_i |\chi_k \chi_l \chi_i\rangle = -\hat{a}_i |\chi_i \chi_l \chi_k\rangle = -|\chi_l \chi_k\rangle = |\chi_k \chi_l\rangle. \quad (2.5)$$

The other important thing to notice is that in the case of the creation operator the order in which two operators are applied to a determinant is crucial. For example:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |\chi_k \cdots \chi_l\rangle = \hat{a}_i^\dagger |\chi_j \chi_k \cdots \chi_l\rangle = |\chi_i \chi_j \chi_k \cdots \chi_l\rangle, \quad (2.6)$$

on the other hand, if we change the order we obtain:

$$\hat{a}_j^\dagger \hat{a}_i^\dagger |\chi_k \cdots \chi_l\rangle = \hat{a}_j^\dagger |\chi_i \chi_k \cdots \chi_l\rangle = |\chi_j \chi_i \chi_k \cdots \chi_l\rangle = -|\chi_i \chi_j \chi_k \cdots \chi_l\rangle. \quad (2.7)$$

Since $|\chi_k \cdots \chi_l\rangle$ is an arbitrary determinant, we obtain the anticommutator relation for the creation operator:

$$[\hat{a}_r^\dagger, \hat{a}_s^\dagger]_+ = \hat{0}. \quad (2.8)$$

Since the annihilation operator is the adjoint of the creation operator we also obtain the anticommutator relation for the annihilation operator:

$$[\hat{a}_r, \hat{a}_s]_+ = \hat{0}. \quad (2.9)$$

One can also obtain a general anticommutator relation between a creation and an annihilation operator. We first consider how the operator $\hat{a}_i\hat{a}_i^\dagger + \hat{a}_i^\dagger\hat{a}_i$ acts on an arbitrary determinant $|\chi_k \cdots \chi_l\rangle$. If the spin orbital χ_i is occupied we have:

$$\begin{aligned}
(\hat{a}_i\hat{a}_i^\dagger + \hat{a}_i^\dagger\hat{a}_i) |\chi_k \cdots \chi_i \cdots \chi_l\rangle &= \hat{a}_i^\dagger\hat{a}_i |\chi_k \cdots \chi_i \cdots \chi_l\rangle \\
&= -\hat{a}_i^\dagger\hat{a}_i |\chi_i \cdots \chi_k \cdots \chi_l\rangle \\
&= -\hat{a}_i^\dagger |\cdots \chi_k \cdots \chi_l\rangle \\
&= -|\chi_i \cdots \chi_k \cdots \chi_l\rangle \\
&= |\chi_k \cdots \chi_i \cdots \chi_l\rangle.
\end{aligned} \tag{2.10}$$

In the case when the spin orbital χ_i is not occupied we have:

$$\begin{aligned}
(\hat{a}_i\hat{a}_i^\dagger + \hat{a}_i^\dagger\hat{a}_i) |\chi_k \cdots \chi_l\rangle &= \hat{a}_i\hat{a}_i^\dagger |\chi_k \cdots \chi_l\rangle \\
&= \hat{a}_i |\chi_i \chi_k \cdots \chi_l\rangle \\
&= |\chi_k \cdots \chi_l\rangle.
\end{aligned} \tag{2.11}$$

Since the same determinant was recovered in both cases we obtain:

$$\hat{a}_i\hat{a}_i^\dagger + \hat{a}_i^\dagger\hat{a}_i = \hat{1} = [\hat{a}_i, \hat{a}_i^\dagger]_+. \tag{2.12}$$

Finally, one should consider $(\hat{a}_j^\dagger\hat{a}_i + \hat{a}_i\hat{a}_j^\dagger) |\chi_k \cdots \chi_l\rangle$ when $i \neq j$. For this term it should be shown that it is equal to zero, since the only case when this expression can be nonzero is when the spin orbital χ_i appears and the spin orbital χ_j is not present in $|\chi_k \cdots \chi_l\rangle$. However, even in that case we obtain zero as a result of the antisymmetry property of determinants. Thus, for the anticommutation relation between a creation and an annihilation operator we obtain:

$$[\hat{a}_i^\dagger, \hat{a}_j]_+ = [\hat{a}_i, \hat{a}_j^\dagger]_+ = \hat{\delta}_{ij}, \tag{2.13}$$

where $\hat{\delta}_{ij}$ is the Kronecker-delta operator.

The anticommutation relations for the creation and annihilation operators can then

be summarized as

$$\begin{aligned} [\hat{a}_i, \hat{a}_j]_+ &= \hat{0}, \\ [\hat{a}_i^\dagger, \hat{a}_j^\dagger]_+ &= \hat{0}, \\ [\hat{a}_i^\dagger, \hat{a}_j]_+ &= [\hat{a}_i, \hat{a}_j^\dagger]_+ = \hat{\delta}_{ij}. \end{aligned} \quad (2.14)$$

They determine all other algebraic properties of the second-quantization formalism.

Since the electronic Hamiltonian consists of one- and two-electron operators we want to know how to represent the latter in second quantization. It can be shown that the one-electron operator within the Fock space can be written as:

$$\hat{F} = \sum_{\mu=1}^N \hat{h}_\mu = \sum_{kl} \langle k | \hat{h} | l \rangle \hat{a}_k^\dagger \hat{a}_l. \quad (2.15)$$

Similarly, for a symmetric two-electron operator we have:

$$\hat{G} = \sum_{\mu < \nu}^N \hat{g}_{\mu\nu} = \frac{1}{2} \sum_{\mu \neq \nu}^N \hat{g}_{\mu\nu} = \frac{1}{2} \sum_{ijkl} \langle i(1)j(2) | \hat{g}_{12} | k(1)l(2) \rangle \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k. \quad (2.16)$$

With the definitions from (2.15) and (2.16), we can rewrite the Hamiltonian as

$$\hat{H} = \sum_{ij} \langle i | \hat{h} | j \rangle \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{ijkl} \langle ij | \hat{v} | kl \rangle \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k, \quad (2.17)$$

where $\langle ij | \hat{v} | kl \rangle \equiv \langle i(1)j(2) | \hat{v}_{12} | k(1)l(2) \rangle$.

The antisymmetric two-electron operator can be abbreviated as:

$$\langle ij | \hat{v} | kl \rangle - \langle ij | \hat{v} | lk \rangle = \langle ij | \hat{v} | kl \rangle_A = \langle ij | | kl \rangle. \quad (2.18)$$

If we also know that:

$$\langle ij | | kl \rangle = -\langle ij | | lk \rangle \quad \text{and} \quad \hat{a}_l \hat{a}_k = -\hat{a}_k \hat{a}_l, \quad (2.19)$$

we obtain the final expression for the second-quantized Hamiltonian:

$$\hat{H} = \sum_{ij} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \frac{1}{4} \sum_{ijkl} \langle ij|kl \rangle \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k. \quad (2.20)$$

In order to obtain the spin-free formulation of the Hamiltonian in second quantization we introduce the spin-free operators. The spin-free one-electron operator is given by:

$$\hat{f} = \sum_{ij}^{N/2} h_{ij} \hat{E}_{ij}, \quad (2.21)$$

where singlet excitation operators are introduced in the form:

$$\hat{E}_{ij} = a_{i\alpha}^\dagger a_{j\alpha} + a_{i\beta}^\dagger a_{j\beta}. \quad (2.22)$$

In the same way, the spin-free two-electron operator is:

$$\hat{g} = \frac{1}{2} \sum_{ijkl}^{N/2} \langle ij|kl \rangle \hat{E}_{ij,kl} \quad (2.23)$$

where two-electron excitation operators are introduced:

$$\hat{E}_{ij,kl} = \sum_{\sigma\tau} a_{i\sigma}^\dagger a_{k\tau}^\dagger a_{j\tau} a_{l\sigma}. \quad (2.24)$$

One-electron integrals display the usual Hermitian permutational symmetry:

$$h_{ij} = h_{ji}^*, \quad (2.25)$$

and for two-electron integrals different symmetries are present for real and complex orbitals. For complex orbitals, Hermitian symmetry is satisfied:

$$\langle ij|kl \rangle = \langle ji|lk \rangle^*, \quad (2.26)$$

and in the case of real orbitals, permutational symmetry is also present

$$\langle ij|kl \rangle = \langle ji|kl \rangle = \langle ij|lk \rangle = \langle ji|lk \rangle. \quad (2.27)$$

Before we write the spin-free Hamiltonian, we will write some helpful operator relations:

$$\hat{E}_{rs}\hat{E}_{ai} = \hat{E}_{ai,rs} + \delta_{as}\hat{E}_{ri}, \quad (2.28)$$

$$\hat{E}_{rs,tu}\hat{E}_{ai} = \hat{E}_{ai,rs,tu} + \delta_{as}\hat{E}_{ri,tu} + \delta_{au}\hat{E}_{rs,ti}. \quad (2.29)$$

The nonrelativistic and spin-free molecular electronic Hamiltonian in the second-quantization representation is given by:

$$\hat{H} = \sum_{rs} h_{rs}\hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu)\hat{E}_{rs,tu}. \quad (2.30)$$

The one- and two-electron integrals are the same as before, except that the integration runs over the spatial coordinates only.

Finally, the following density matrix relations which are valid for the reference function will also be used

$$\langle \Phi_0 | \hat{E}_{ij} | \Phi_0 \rangle = 2\hat{\delta}_{ij}, \quad (2.31)$$

$$\langle \Phi_0 | \hat{E}_{ik,jl} | \Phi_0 \rangle = 4\hat{\delta}_{ik}\hat{\delta}_{jl} - 2\hat{\delta}_{il}\hat{\delta}_{jk}. \quad (2.32)$$

2.2 Møller-Plesset Perturbation Theory

One way to include the correlation interaction into wave function calculations is by the use of Møller-Plesset (MP) perturbation theory [4]. In this theory one assumes that the HF solution is a sufficiently good approximation to the total energy of the system and the missing contributions are obtained through a perturbation expansion. Thus, the Hamiltonian can be divided into a reference $\hat{H}^{(0)}$ part and a perturbation $\hat{H}^{(1)}$ part:

$$(\hat{H}^{(0)} + \lambda\hat{H}^{(1)})|\Psi\rangle = E|\Psi\rangle \quad (2.33)$$

with

$$\hat{H}^{(0)} = \sum_i^N \hat{f}(i) = \sum_i^N [\hat{h}(i) + \hat{g}(i)] \quad (2.34)$$

$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)} \quad (2.35)$$

where the unperturbed part of the Hamiltonian represents the sum of the Fock operators. The next step is to expand the energy and the wavefunction in a Taylor series:

$$E = \sum_{k=0}^N \lambda^k E^{(k)} = E^{(0)} + \lambda^1 E^{(1)} + \lambda^2 E^{(2)} + \dots \quad (2.36)$$

$$\Psi = \sum_{k=0}^N \lambda^k \Psi^{(k)} = \Psi^{(0)} + \lambda^1 \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \dots \quad (2.37)$$

These expansions are now inserted into the Equation (2.33) and since this holds for any λ , an equation for each power of λ is obtained:

$$\begin{aligned} \lambda^{(0)} : \hat{H}^{(0)} |\Psi^{(0)}\rangle &= E^{(0)} |\Psi^{(0)}\rangle \\ \lambda^{(1)} : \hat{H}^{(0)} |\Psi^{(1)}\rangle + \hat{H}^{(1)} |\Psi^{(0)}\rangle &= E^{(0)} |\Psi^{(1)}\rangle + E^{(1)} |\Psi^{(0)}\rangle \\ \lambda^{(2)} : \hat{H}^{(0)} |\Psi^{(2)}\rangle + \hat{H}^{(1)} |\Psi^{(1)}\rangle &= E^{(0)} |\Psi^{(2)}\rangle + E^{(1)} |\Psi^{(1)}\rangle + E^{(2)} |\Psi^{(0)}\rangle \\ &\dots \\ \lambda^{(k)} : \hat{H}^{(0)} |\Psi^{(k)}\rangle + \hat{H}^{(1)} |\Psi^{(k-1)}\rangle &= \sum_{i=0}^n E^{(i)} |\Psi^{(k-i)}\rangle \end{aligned} \quad (2.38)$$

This expansion can be truncated at any point. The zeroth order energy presents the sum of the orbital energies and the first order recovers the HF energy. Therefore, the second order is the first correction to the HF energy. These energies are obtained by multiplying the first three equations with $\langle \Psi^{(0)} |$, where $\Psi^{(0)}$ is the HF wave function:

$$\begin{aligned} \lambda^{(0)} : E^{(0)} &= \langle \Psi^{(0)} | \hat{H}^{(0)} | \Psi^{(0)} \rangle = \sum_i^N \epsilon_i \\ \lambda^{(1)} : E^{(1)} &= \langle \Psi^{(0)} | \hat{H}^{(1)} | \Psi^{(0)} \rangle = -\frac{1}{2} \sum_{ij} [2(ii|jj) - (ij|ji)] \\ \lambda^{(2)} : E^{(2)} &= \langle \Psi^{(0)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle \end{aligned} \quad (2.39)$$

As one can see, to obtain the second order correction one needs the first order wavefunction. According to the Brillouins theorem, single excited configurations do not interact with the reference state. Therefore, the first order wavefunction depends only on the

double excited configurations and can be presented as:

$$|\Psi^{(1)}\rangle = \frac{1}{2} \sum_{ij} \sum_{ab} T_{ab}^{ij} |\Phi_{ij}^{ab}\rangle, \quad (2.40)$$

where $|\Phi_{ij}^{ab}\rangle$ is:

$$|\Phi_{ij}^{ab}\rangle = \hat{E}_{ai} \hat{E}_{bj} |\Psi^0\rangle. \quad (2.41)$$

Since the double excited configurations are neither orthogonal nor normalized, we will use contravariant configurations and amplitudes:

$$\tilde{\Phi}_{ij}^{ab} = \frac{1}{6} (2\Phi_{ij}^{ab} + \Phi_{ji}^{ab}), \quad (2.42)$$

$$\tilde{T}_{ab}^{ij} = 2T_{ab}^{ij} - T_{ab}^{ji}. \quad (2.43)$$

In this way, the final equations are greatly simplified. Relationships between contravariant and covariant forms of configurations and amplitudes are as presented here:

$$\langle \tilde{\Phi}_{ij}^{ab} | \Phi_{kl}^{cd} \rangle = \delta_{ac} \delta_{bd} \delta_{ik} \delta_{jl} + \delta_{ad} \delta_{bc} \delta_{il} \delta_{jk}, \quad (2.44)$$

$$\langle \tilde{\Phi}_{ij}^{ab} | \Psi^{(1)} \rangle = T_{ab}^{ij}, \quad (2.45)$$

$$\langle \tilde{\Phi}_{ij}^{ab} | \hat{H} | \Psi^{(0)} \rangle = (ai|bj) = K_{ab}^{ij}, \quad (2.46)$$

where K_{ab}^{ij} denotes the exchange integral. Using these relations one can obtain the second order correction to the energy:

$$\begin{aligned} \Delta E_{MP2} = E^{(2)} &= \langle \Psi^{(0)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle = \langle \Psi^{(0)} | \hat{H} - \hat{H}^{(1)} | \Psi^{(1)} \rangle \\ &= \sum_{ijab} \langle \Psi^{(0)} | \hat{H} | \tilde{\Phi}_{ij}^{ab} \rangle \tilde{T}_{ab}^{ij} = \sum_{ijab} K_{ab}^{ij} \tilde{T}_{ab}^{ij}. \end{aligned} \quad (2.47)$$

If the second term of Equation (2.39) is multiplied from the left with $\langle \tilde{\Phi}_{ab}^{ij} |$ one obtains the double residuals equation

$$R_{ij}^{ab} = \langle \tilde{\Phi}_{ab}^{ij} | \hat{H}^{(0)} - E^{(0)} | \Psi^{(1)} \rangle + \langle \tilde{\Phi}_{ab}^{ij} | \hat{H} | \Psi^{(0)} \rangle = 0. \quad (2.48)$$

This equation should be equal to zero for the converged solution. Using second quantization the double residuals can be written as:

$$R_{ij}^{ab} = K_{ij}^{ab} + \sum_c (f_{ac} T_{cd}^{ij} + T_{ac}^{ij} f_{cb}) - \sum_k (f_{ik} T_{ab}^{kj} + T_{ab}^{ik} f_{kj}) = 0 \quad (2.49)$$

These double residuals are orbital independent and can be used with any type of orbitals. Solving this equation for double amplitudes and considering canonical orbitals one obtains:

$$T_{ab}^{ij} = -\frac{K_{ab}^{ij}}{(\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)}, \quad (2.50)$$

and the final energy correction is:

$$\Delta E_{MP2} = \sum_{ij} \sum_{ab} \frac{K_{ab}^{ij} (2K_{ab}^{ij} - K_{ab}^{ji})}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \quad (2.51)$$

In the same way, higher order perturbations can be included leading to the MPn series. However, their inclusion does not guarantee the convergence of the series. Nevertheless, the MP2 method is widely used since the improvement over HF theory is significant and it is the cheapest correlation method (with a formal scaling of N^5 with the system size). On the other hand, the error can be large in some cases, e.g. for molecules with a small HOMO-LUMO gap and when the HF reference is a poor description. Empirical corrections to MP2 [5, 6] have been introduced to correct for the well-known overestimation of the correlation energy.

2.3 Coupled Cluster Theory

Coupled cluster (CC) [1, 3] takes the basic Hartree-Fock molecular orbital method and constructs multi electron wavefunctions using the exponential cluster operator to account for electron correlation. The cluster operator is given as:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots, \quad (2.52)$$

a sum over all excitation operators, including singles (\hat{T}_1), doubles (\hat{T}_2) and all other remaining higher excitations. These operators are represented as:

$$\hat{T}_1 = \sum_{ia} t_a^i \hat{E}_{ai}, \quad (2.53)$$

$$\hat{T}_2 = \frac{1}{(2!)^2} \sum_{ijab} t_{ab}^{ij} \hat{E}_{ai} \hat{E}_{bj}, \quad (2.54)$$

$$\hat{T}_3 = \frac{1}{(3!)^2} \sum_{ijkabc} t_{abc}^{ijk} \hat{E}_{ai} \hat{E}_{bj} \hat{E}_{ck}, \quad (2.55)$$

where $t_{ab\dots}^{ij\dots}$ are the amplitudes for the corresponding operators. In general, an excitation operator is given by:

$$\hat{T}_m = \frac{1}{(m!)^2} \sum_{ij\dots ab\dots} t_{ab\dots}^{ij\dots} \hat{E}_{ai} \hat{E}_{bj} \dots \quad (2.56)$$

The indices i, j, k, l refer to occupied orbitals (hole indices) and a, b, c, d to unoccupied orbitals (particle indices). Indices r, s, t, u will refer to general orbitals. The $\frac{1}{(m!)^2}$ factor accounts for the redundancy created by the unrestricted summations, since any permutation of the m hole indices or of the m particle indices does not produce a distinct contribution.

The wavefunction of the coupled cluster theory is written as an exponential ansatz:

$$|\Psi\rangle = e^{\hat{T}} |\Phi_0\rangle \quad (2.57)$$

The exponential operator $e^{\hat{T}}$ can be expanded in a Taylor series, resulting in:

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2!} \hat{T}^2 + \dots \quad (2.58)$$

2.3.1 Coupled Cluster Theory with Singles and Doubles

In case of coupled cluster with singles and doubles (CCSD), the cluster operator will be restricted to

$$\hat{T} = \hat{T}_1 + \hat{T}_2. \quad (2.59)$$

Taking \hat{T} in Equation (2.59), counting only up to double excitations, one obtains:

$$\begin{aligned} e^{\hat{T}} &= 1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2)^2 + \dots \\ &= 1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2}(\hat{T}_1^2 + 2\hat{T}_1\hat{T}_2 + \hat{T}_2^2) + \dots \end{aligned} \quad (2.60)$$

If we now insert the coupled cluster wavefunction into the Schrödinger equation and use the exponential operator in form which we got in Equation (2.60), we will obtain

$$\hat{H}(1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2)^2 + \dots) |\Phi_0\rangle = E(1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2)^2 + \dots) |\Phi_0\rangle \quad (2.61)$$

where $E = E_0 + E_{corr}$. If we now multiply Equation (2.61) from the left with $\langle\Phi_0|$ we obtain an expression for the energy:

$$\langle\Phi_0| \hat{H}(1 + \hat{T}_1 + (\frac{1}{2}\hat{T}_1^2 + \hat{T}_2)) |\Phi_0\rangle = \langle\Phi_0| E(1 + \hat{T}_1 + (\frac{1}{2}\hat{T}_1^2 + \hat{T}_2)) |\Phi_0\rangle, \quad (2.62)$$

where only the terms which contribute to the energy are kept.

If we want to obtain the single amplitudes for CCSD we need to keep all terms from Equation (2.58) which contain up to triple excitations, as they will couple to the single excited space through the Hamiltonian. In the case of double residuals, quadruple terms will be present. Therefore, if we include the exponential operator $e^{\hat{T}}$ as described and multiply Schrödinger equation with $\langle\Phi_i^a|$ we obtain the connected amplitude equation for single amplitudes:

$$\begin{aligned} \langle\Phi_i^a| \hat{H}(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3) |\Phi_0\rangle &= \\ = \langle\Phi_i^a| E(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3) |\Phi_0\rangle. \end{aligned} \quad (2.63)$$

We will start by solving for the energy, by reforming Equation (2.62):

$$\begin{aligned} \langle\Phi_0| \hat{H} |\Phi_0\rangle + \langle\Phi_0| \hat{H}\hat{T}_1 |\Phi_0\rangle + \langle\Phi_0| \hat{H}(\frac{1}{2}\hat{T}_1^2 + \hat{T}_2) |\Phi_0\rangle &= \\ \langle\Phi_0| E |\Phi_0\rangle + \langle\Phi_0| E\hat{T}_1 |\Phi_0\rangle + \langle\Phi_0| E(\frac{1}{2}\hat{T}_1^2 + \hat{T}_2) |\Phi_0\rangle, \end{aligned} \quad (2.64)$$

$$E_0 + \langle\Phi_0| \hat{H}\hat{T}_1 |\Phi_0\rangle + \langle\Phi_0| \hat{H}(\frac{1}{2}\hat{T}_1^2 + \hat{T}_2) |\Phi_0\rangle = E. \quad (2.65)$$

From the last equation the correlation energy can be computed as:

$$E_{corr} = \langle \Phi_0 | \hat{H} \hat{T}_1 | \Phi_0 \rangle + \langle \Phi_0 | \hat{H} (\frac{1}{2} \hat{T}_1^2 + \hat{T}_2) | \Phi_0 \rangle. \quad (2.66)$$

In order to obtain working equations, we will make use of the Hamiltonian in second quantization as given in Equation (2.30). This results in

$$\begin{aligned} E_{corr} = & \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] \hat{T}_1 | \Phi_0 \rangle + \\ & + \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] (\frac{1}{2} \hat{T}_1^2 + \hat{T}_2) | \Phi_0 \rangle \end{aligned} \quad (2.67)$$

We will look in these terms separately, the first term is:

$$\begin{aligned} A = & \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] \hat{T}_1 | \Phi_0 \rangle \\ = & \sum_{ai} \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] \hat{E}_{ai} | \Phi_0 \rangle t_a^i \\ = & \sum_{ai} \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} \hat{E}_{ai} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu} \hat{E}_{ai}] | \Phi_0 \rangle t_a^i \end{aligned} \quad (2.68)$$

Using relations (2.28) and (2.29) one obtains:

$$A = \sum_{ai} \langle \Phi_0 | [\sum_{rs} h_{rs} (\hat{E}_{ai,rs} + \delta_{as} \hat{E}_{ri}) + \frac{1}{2} \sum_{rstu} (rs|tu) (\hat{E}_{ai,rs,tu} + \delta_{as} \hat{E}_{ri,tu} + \delta_{au} \hat{E}_{rs,ti})] | \Phi_0 \rangle t_a^i \quad (2.69)$$

The first term in the above equation with $\hat{E}_{ai,rs}$ is equal to zero since a cannot be equal to i , and the third term with $\hat{E}_{ai,rs,tu}$ is equal to zero since $\langle \Phi_0 | \Phi_{ijk}^{abc} \rangle$ will give zero. The equation can then be rewritten as

$$A = \sum_{ai} [\langle \Phi_0 | \sum_{rs} h_{rs} \delta_{as} \hat{E}_{ri} | \Phi_0 \rangle + \frac{1}{2} \sum_{rstu} (rs|tu) [\langle \Phi_0 | \delta_{as} \hat{E}_{ri,tu} | \Phi_0 \rangle + \langle \Phi_0 | \delta_{au} \hat{E}_{rs,ti} | \Phi_0 \rangle]] t_a^i \quad (2.70)$$

Combining Equations (2.31) and (2.32) one obtains:

$$\begin{aligned}
A &= \sum_{ai} [2 \sum_{rs} h_{rs} \delta_{as} \delta_{ri} + \frac{1}{2} \sum_{rstu} (rs|tu) [\delta_{as} (4\delta_{ri} \delta_{tu} - 2\delta_{ru} \delta_{it}) + \delta_{au} (4\delta_{rs} \delta_{ti} - 2\delta_{ri} \delta_{st})]] t_a^i \\
&= \sum_{ai} [2h_{ia} + \frac{1}{2} \sum_{rstu} (rs|tu) [4\delta_{as} \delta_{ri} \delta_{tu} - 2\delta_{as} \delta_{ru} \delta_{it} + 4\delta_{au} \delta_{rs} \delta_{ti} - 2\delta_{au} \delta_{ri} \delta_{st}]] t_a^i \\
&= \sum_{ai} [2h_{ia} + 2 \sum_u (ia|uu) - \sum_u (ua|iu) + 2 \sum_r (rr|ia) - \sum_t (it|ta)] t_a^i \\
&= 2 \sum_{ai} [h_{ia} + \sum_r [2(ia|rr) - (ra|ir)]] t_a^i \tag{2.71}
\end{aligned}$$

If we know that

$$(\mathbf{F})_{rs} = (\mathbf{F}^r)_s = h_{rs} + \sum_k [2(rs|kk) - (rk|ks)] \tag{2.72}$$

for the first term we will obtain

$$A = 2F_{ia}^\dagger t_a^i, \quad \text{where} \quad F_{ia}^\dagger = F_{ia}. \tag{2.73}$$

Now we will consider the second term from Equation (2.67):

$$\begin{aligned}
B &= \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] (\frac{1}{2} \hat{T}_1^2 + \hat{T}_2) | \Phi_0 \rangle \\
&= \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] \sum_{abij} (\frac{1}{2} t_a^i t_b^j + \frac{1}{4} t_{ab}^{ij}) | \Phi_{ij}^{ab} \rangle \tag{2.74}
\end{aligned}$$

If we introduce $C_{ab}^{ij} = \frac{1}{2} t_a^i t_b^j + \frac{1}{4} t_{ab}^{ij}$ one obtains:

$$\begin{aligned}
B &= \sum_{abij} C_{ab}^{ij} \langle \Phi_0 | [\sum_{rs} h_{rs} \hat{E}_{rs} + \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu}] | \Phi_{ij}^{ab} \rangle \\
&= \sum_{abij} C_{ab}^{ij} [\langle \Phi_0 | \sum_{rs} h_{rs} \hat{E}_{rs} | \Phi_{ij}^{ab} \rangle + \langle \Phi_0 | \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu} | \Phi_{ij}^{ab} \rangle] \\
&= \sum_{abij} C_{ab}^{ij} [\langle \Phi_0 | \sum_{rs} h_{rs} \hat{E}_{rs} \hat{E}_{ai} \hat{E}_{bj} | \Phi_0 \rangle + \langle \Phi_0 | \frac{1}{2} \sum_{rstu} (rs|tu) \hat{E}_{rs,tu} \hat{E}_{ai} \hat{E}_{bj} | \Phi_0 \rangle]. \tag{2.75}
\end{aligned}$$

Combining Equations (2.28) and (2.29) the operator relation

$$\begin{aligned}
\hat{E}_{rs}\hat{E}_{ai}\hat{E}_{bj} &= [\hat{E}_{ai,rs} + \delta_{as}\hat{E}_{ai}]\hat{E}_{bj} \\
&= \hat{E}_{bj,ai,rs} + \delta_{bi}\hat{E}_{aj,rs} + \delta_{bs}\hat{E}_{ai,rj} + \delta_{as}[\hat{E}_{bj,ri} + \delta_{bi}\hat{E}_{rj}] \\
&= \hat{E}_{bj,ai,rs} + \delta_{bi}\hat{E}_{aj,rs} + \delta_{bs}\hat{E}_{ai,rj} + \delta_{as}\hat{E}_{bj,ri} + \delta_{as}\delta_{bi}\hat{E}_{rj}
\end{aligned} \tag{2.76}$$

is obtained, as well as

$$\begin{aligned}
\hat{E}_{rs,tu}\hat{E}_{ai}\hat{E}_{bj} &= [\hat{E}_{ai,rs,tu} + \delta_{as}\hat{E}_{ri,tu} + \delta_{au}\hat{E}_{rs,ti}]\hat{E}_{bj} \\
&= \hat{E}_{ai,rs,tu}\hat{E}_{bj} + \delta_{as}\hat{E}_{ri,tu}\hat{E}_{bj} + \delta_{au}\hat{E}_{rs,ti}\hat{E}_{bj} \\
&= \hat{E}_{ai,rs,tu}\hat{E}_{bj} + \delta_{as}(\hat{E}_{bj,ri,tu} + \delta_{bi}\hat{E}_{rj,tu} + \delta_{bu}\hat{E}_{ri,tj}) \\
&\quad + \delta_{au}(\hat{E}_{bj,rs,ti} + \delta_{bs}\hat{E}_{rj,ti} + \delta_{bi}\hat{E}_{rs,tj})
\end{aligned} \tag{2.77}$$

Since $a \neq i, j, \dots$ and $b \neq i, j, \dots$, some of the terms in Equation (2.76) are equal zero, the rest of the terms is equal zero due to $\langle \Phi_0 | \Phi_{ijk}^{abc} \rangle$ or $\langle \Phi_0 | \Phi_{ijkl}^{abcd} \rangle$ being zero. Therefore, for Equation (2.76) we obtain

$$\hat{E}_{rs}\hat{E}_{ai}\hat{E}_{bj} = 0 \tag{2.78}$$

and due to the same reason we obtain for Equation (2.77)

$$\hat{E}_{rs,tu}\hat{E}_{ai}\hat{E}_{bj} = \delta_{as}\delta_{bu}\hat{E}_{ri,tj} + \delta_{au}\delta_{bs}\hat{E}_{rj,ti} \tag{2.79}$$

Inserting the relations above into Equation (2.75) leads to

$$\begin{aligned}
B &= \sum_{abij} C_{ab}^{ij} [\langle \Phi_0 | \frac{1}{2} \sum_{rstu} (rs|tu) [\delta_{as}\delta_{bu}\hat{E}_{ri,tj} + \delta_{au}\delta_{bs}\hat{E}_{rj,ti}] | \Phi_0 \rangle] \\
&= \frac{1}{2} \sum_{rstu} (rs|tu) \sum_{abij} C_{ab}^{ij} [\delta_{as}\delta_{bu} \langle \Phi_0 | \hat{E}_{ri,tj} | \Phi_0 \rangle + \delta_{au}\delta_{bs} \langle \Phi_0 | \hat{E}_{rj,ti} | \Phi_0 \rangle]
\end{aligned} \tag{2.80}$$

And again, applying Equations (2.31) and (2.32):

$$\begin{aligned}
B &= \frac{1}{2} \sum_{rstu} (rs|tu) \sum_{abij} C_{ab}^{ij} [\delta_{as}\delta_{bu}(4\delta_{rj}\delta_{ti} - 2\delta_{jt}\delta_{ri}) + \delta_{au}\delta_{bs}(4\delta_{ri}\delta_{tj} - 2\delta_{it}\delta_{rj})] \\
&= \frac{1}{2} \sum_{abij} C_{ab}^{ij} [4(jb|ia) - 2(ib|ja) + 4(ia|jb) - 2(ja|ib)] \\
&= \sum_{abij} C_{ab}^{ij} [4(jb|ia) - 2(ib|ja)] \tag{2.81}
\end{aligned}$$

The correlation energy is then given as:

$$\begin{aligned}
E_{corr} &= 2 \sum_{ai} F_{ia}^\dagger t_a^i + \sum_{abij} [(\frac{1}{2}t_a^i t_b^j + \frac{1}{4}t_{ab}^{ij}) [4(jb|ia) - 2(ib|ja)]] \\
&= 2 \sum_{ai} F_{ia}^\dagger t_a^i + \sum_{abij} [\frac{1}{2}t_a^i t_b^j [4(jb|ia) - 2(ib|ja)] + \frac{1}{4}t_{ab}^{ij} [4(jb|ia) - 2(ib|ja)]] \\
&= 2 \sum_{ai} F_{ia}^\dagger t_a^i + \sum_{abij} [t_a^i t_b^j [2(jb|ia) - (ib|ja)] + \frac{1}{2}t_{ab}^{ij} [2(jb|ia) - (ib|ja)]]. \tag{2.82}
\end{aligned}$$

if we instead use:

$$(\mathbf{L}^{ij})_{ab} = [2\mathbf{K}^{ij} - \mathbf{K}^{ji}]_{ab} = 2(ai|jb) - (aj|ib) \tag{2.83}$$

we have:

$$B = 2 \sum_{abij} C_{ab}^{ij} L_{ab}^{ij} \tag{2.84}$$

This leads to a relatively compact expression for the correlation energy

$$\begin{aligned}
E_{corr} &= \sum_{ai} \langle \Phi_0 | \hat{H} | \Phi_i^a \rangle t_a^i + \sum_{abij} \langle \Phi_0 | \hat{H} | \Phi_{ij}^{ab} \rangle C_{ab}^{ij} \\
&= 2 \sum_{ai} F_{ia}^\dagger t_a^i + 2 \sum_{abij} L_{ab}^{ij} C_{ab}^{ij} \\
&= 2 \sum_i f^{i\dagger} t^i + 2 \sum_{ij} tr[L^{ij} C^{ij}] \\
&= 2 \sum_i f^{i\dagger} t^i + 2 \sum_{i \geq j} (2 - \delta_{ij}) tr[L^{ij} C^{ij}] \tag{2.85}
\end{aligned}$$

2.3.2 Coupled Cluster Theory in Diagrammatic Representation

An alternative to derive the CC equations is to use diagrammatic representation. The purpose of diagrammatic representation is:

- to ease the task of finding all non-vanishing distinct terms in the CC equations,
- to bring out certain cancellations in these sums,
- to provide systematic manipulation of the various surviving terms,
- to construct the CC energy and amplitude equations far more quickly than by direct application of Wick's theorem or any other algebraic way.

Diagrammatic representation has its origin in quantum field theory, in the form of Feynman diagrams, in an explicit time-dependent format. Initially, in the Rayleigh-Schrödinger perturbation theory (RSPT) diagrammatic representation was also in time-dependent form. This was indicated in the diagrams by a time axis. This time axis can be vertical or horizontal. In the case where the time axis is vertical, if we want to represent $\hat{U} |\Phi_{ij}^{ab}\rangle$, we start with a representation of $|\Phi_{ij}^{ab}\rangle$ at the bottom, and the operator \hat{U} is given above, leading to a representation of the result at the top. The second case is when the time axis is horizontal, then one writes terms going from right to left in the same way as for the vertical axis. In the following text we will make use of the vertical representation. First we will go through basic rules and representations. Later we will proceed to obtain an equation for the correlation energy of CCSD, showcasing the power of diagrammatic representations.

Slater determinants

In the diagrammatic representation the Slater determinant (SD) for the reference state (Fermi vacuum) is represented by an empty space, i.e. by a position on the time axis at which there are no lines or other symbols. All other SDs are represented by vertical or diagonal directed lines, pointing upward for particles and downwards for holes, with labels identifying the spin orbitals.

$$\Phi_i^a = \begin{array}{c} | \\ | \\ \downarrow i \\ \uparrow a \\ | \\ | \end{array} \qquad \Phi_{ij}^{ab} = \begin{array}{c} | \\ | \\ \downarrow i \\ \uparrow a \\ | \\ | \end{array} \quad \begin{array}{c} | \\ | \\ \downarrow j \\ \uparrow b \\ | \\ | \end{array}$$

In the case where a *ket* or *bra* should be indicated different notations can be used:

$$\begin{aligned}
 |\Phi^a\rangle = \hat{a}^\dagger |0\rangle &= \begin{array}{c} | \\ \uparrow a \\ \hline \hline \end{array} & |\Phi_i\rangle = \hat{i} |0\rangle &= \begin{array}{c} | \\ \downarrow i \\ \hline \hline \end{array} \\
 |\Phi_i^a\rangle = \{\hat{a}^\dagger \hat{i}\} |0\rangle &= \begin{array}{c} | \\ \downarrow i \quad \uparrow a \\ \hline \hline \end{array} & \langle \Phi^a | = \langle 0 | \{\hat{i}^\dagger \hat{a}\} &= \begin{array}{c} \hline \hline \\ \downarrow i \quad \uparrow a \end{array} \\
 |\Phi_{ij}^{ab}\rangle = \{\hat{a}^\dagger \hat{b}^\dagger \hat{j} \hat{i}\} |0\rangle = \{(\hat{a}^\dagger \hat{i})(\hat{b}^\dagger \hat{j})\} |0\rangle &= \begin{array}{c} | \\ \downarrow i \quad \uparrow a \quad \downarrow j \quad \uparrow b \\ \hline \hline \end{array}
 \end{aligned}$$

The horizontal double line represents the point of operation of the normal-product operator, and below or above it we have the Fermi vacuum. In this case a new notation was also used for the creation and annihilation operators, so instead of \hat{a}_i^\dagger for the creation operator we used only \hat{i}^\dagger and for the annihilation operator \hat{i} was used instead of \hat{a}_i .

One-particle operators

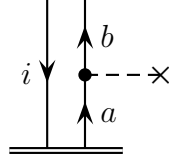
The representation of operators will be considered next. We will start with the one-electron operator $\hat{U}_N = \sum_{pq} \langle p | \hat{u} | q \rangle \{\hat{p}^\dagger \hat{q}\}$ which acts on a singly excited SD $|\Phi_i^a\rangle = \{\hat{a}^\dagger \hat{i}\} |0\rangle$. Since the representation will depend on whether p and q are particle or hole indices we will start with a particle-particle term. Using the generalized Wick's theorem we obtain:

$$\begin{aligned}
 \langle b | \hat{u} | c \rangle \{\hat{b}^\dagger \hat{c}\} \{\hat{a}^\dagger \hat{i}\} |0\rangle &= \langle b | \hat{u} | c \rangle \{\hat{b}^\dagger \hat{c} \hat{a}^\dagger \hat{i}\} |0\rangle + \langle b | \hat{u} | c \rangle \{\hat{b}^\dagger \hat{c} \hat{a}^\dagger \hat{i}\} |0\rangle \\
 &= \langle b | \hat{u} | c \rangle \hat{b}^\dagger \hat{a}^\dagger \hat{i} \hat{c} |0\rangle + \langle b | \hat{u} | c \rangle \delta_{ac} \{\hat{b}^\dagger \hat{i}\} |0\rangle \\
 &= 0 + \langle b | \hat{u} | c \rangle \delta_{ac} |\Phi_i^b\rangle.
 \end{aligned} \tag{2.86}$$

One can observe that the only non-zero term for a particle-particle interaction is

$$\langle b | \hat{u} | a \rangle \{ \hat{b}^\dagger \hat{a} \} | \Phi_i^a \rangle = \langle b | \hat{u} | a \rangle | \Phi_i^b \rangle \quad (2.87)$$

which is given in diagram form as



At the bottom $|\Phi_i^a\rangle$ is represented and at the top $|\Phi_i^b\rangle$. The point of the operator's action is marked by the *interaction line (vertex)*. X is a marker for the operator \hat{u} . In case we need to represent more than one one-electron operator, different markers have to be used ($\#$ or \triangle or \square). Also, it is important to note that the *bra* corresponds to the line leaving the vertex while the *ket* corresponds to the entering line.

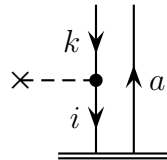
Now we will consider a hole-hole term. Again using Wick's theorem we obtain:

$$\begin{aligned} \langle j | \hat{u} | k \rangle \{ \hat{j}^\dagger \hat{k} \} \{ \hat{a}^\dagger \hat{i} \} | 0 \rangle &= \langle j | \hat{u} | k \rangle \{ \hat{j}^\dagger \hat{k} \hat{a}^\dagger \hat{i} \} | 0 \rangle + \langle j | \hat{u} | k \rangle \overline{\{ \hat{j}^\dagger \hat{k} \hat{a}^\dagger \hat{i} \}} | 0 \rangle \\ &= - \langle j | \hat{u} | k \rangle \hat{k} \hat{a}^\dagger \hat{i} \hat{j}^\dagger | 0 \rangle + \delta_{ij} \langle j | \hat{u} | k \rangle \{ \hat{k} \hat{a}^\dagger \} | 0 \rangle \\ &= 0 - \delta_{ij} \langle j | \hat{u} | k \rangle \{ \hat{a}^\dagger \hat{k} \} | 0 \rangle \\ &= -\delta_{ij} \langle j | \hat{u} | k \rangle | \Phi_k^a \rangle \end{aligned} \quad (2.88)$$

The only non-vanishing term is:

$$\langle i | \hat{u} | k \rangle \{ \hat{i}^\dagger \hat{k} \} | \Phi_i^a \rangle = - \langle i | \hat{u} | k \rangle | \Phi_k^a \rangle \quad (2.89)$$

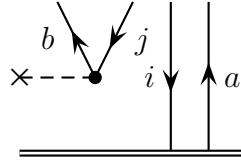
which in diagrammatic representation looks like:



In the case of the particle-hole term one obtains:

$$\begin{aligned}
 \langle b | \hat{u} | j \rangle \{ \hat{b}^\dagger \hat{j} \} \{ \hat{a}^\dagger \hat{i} \} | 0 \rangle &= \langle b | \hat{u} | j \rangle \{ \hat{b}^\dagger \hat{j} \hat{a}^\dagger \hat{i} \} | 0 \rangle \\
 &= \langle b | \hat{u} | j \rangle \hat{a}^\dagger \hat{b}^\dagger \hat{j} \hat{i} | 0 \rangle \\
 &= \langle b | \hat{u} | j \rangle | \Phi_{ij}^{ab} \rangle
 \end{aligned} \tag{2.90}$$

which is represented as:



Finally, the hole-particle term is given by:

$$\begin{aligned}
 \langle j | \hat{u} | b \rangle \{ \hat{j}^\dagger \hat{b} \} \{ \hat{a}^\dagger \hat{i} \} | 0 \rangle &= \langle j | \hat{u} | b \rangle \{ \hat{j}^\dagger \hat{b} \hat{a}^\dagger \hat{i} \} | 0 \rangle + \langle j | \hat{u} | b \rangle \overline{\{ \hat{j}^\dagger \hat{b} \hat{a}^\dagger \hat{i} \}} | 0 \rangle \\
 &+ \langle j | \hat{u} | b \rangle \overline{\{ \hat{j}^\dagger \hat{b} \hat{a}^\dagger \hat{i} \}} | 0 \rangle + \langle j | \hat{u} | b \rangle \overline{\overline{\{ \hat{j}^\dagger \hat{b} \hat{a}^\dagger \hat{i} \}}} | 0 \rangle \\
 &= \langle j | \hat{u} | b \rangle \{ 0 + \delta_{ij} \times 0 + \delta_{ab} \times 0 + \delta_{ij} \delta_{ab} | 0 \rangle \\
 &= \langle j | \hat{u} | b \rangle \delta_{ij} \delta_{ab} | 0 \rangle
 \end{aligned} \tag{2.91}$$

with the only non-zero term being:

$$\langle i | \hat{u} | a \rangle \{ \hat{i}^\dagger \hat{a} \} | \Phi_i^a \rangle = \langle i | \hat{u} | a \rangle | 0 \rangle \tag{2.92}$$

represented by:



Some simple rules can be derived for the interpretation of the one-particle vertices. One can see that the *bra* in the integral corresponds to an outgoing line, while the *ket* corresponds to an incoming line. Also, a creation operator corresponds to an outgoing line and an annihilation operator to an incoming line.

bra index	\leftrightarrow	creation operator	\leftrightarrow	outgoing line
ket index	\leftrightarrow	annihilation operator	\leftrightarrow	incoming line

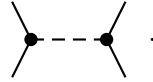
We also saw that in the case of the hole-hole term a phase factor of -1 is present. The phase factor in this case is associated with one hole being replaced by another. It can be shown that when we generalize this we obtain a phase factor of $(-1)^{h-l}$, where h is the number of hole lines in the loop and l is the number of loops. A path is a sequence of connected lines, including any connection that indicates particle-hole pairs in the initial and final state. A loop is a connected path.

Two-particle operators

Now we will consider a two-particle operator in normal-product form,

$$\hat{W} = \frac{1}{2} \sum_{pqrs} \langle pq | rs \rangle \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\} = \frac{1}{4} \sum_{pqrs} \langle pq | rs \rangle \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\}. \quad (2.93)$$

A two-particle operator is represented with an interaction line which connects two half-vertices at the same level:



In the case of two-particle operators a single vertex is composed of two half-vertices and the interaction line. Line labels are associated in the same way as for the one-particle operators, but with the added feature for the two-body case that

electron 1	\leftrightarrow	left half-vertex
electron 2	\leftrightarrow	right half-vertex

The integral indices associated with a two-body vertex are assigned according to the scheme

\langle left-out right-out left-in right-in \rangle ,

while the corresponding operator product can be described by

$\{(\text{left-out})^\dagger (\text{right-out})^\dagger (\text{right-in}) (\text{left-in})\}.$
--

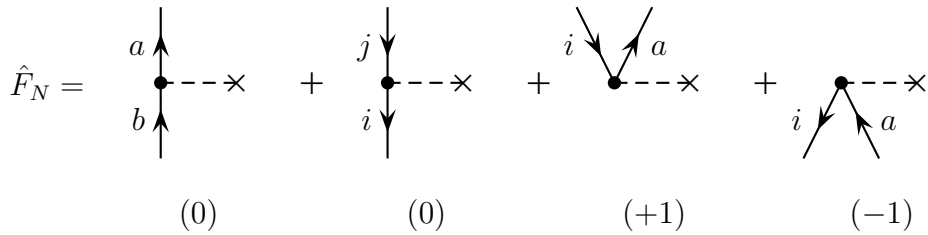
Diagrams which employ representations based on non-antisymmetrized integrals are called Goldstone diagrams.

Basics for coupled cluster

Based on the diagrammatic representations of one-electron and two-electron operators introduced in the previous section, the electronic Hamiltonian

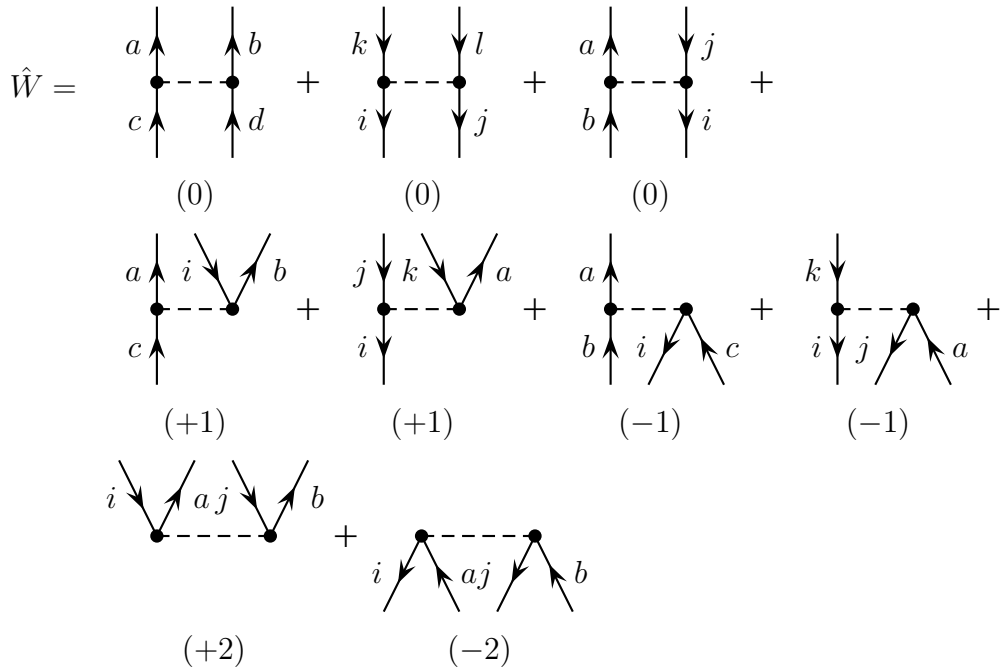
$$\hat{H} = \hat{F} + \hat{W} = \sum_{pq} f_{pq} \{\hat{p}^\dagger \hat{q}\} + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \{\hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r}\} \quad (2.94)$$

can be represented by



for the one electron operator. The numbers below the diagrams denote the excitation level represented by each diagram. They are calculated as the number of particle-hole open-line pairs at the top over the number of such pairs at the bottom.

For the case of two-electron operators we have:



For the coupled cluster methods, the diagrammatic representation of cluster operators is also necessary. Using the Goldstone diagrams they are represented as shown below.

$$\begin{array}{ccc}
 \hat{T}_1 = & \begin{array}{c} \text{Diagram 1: A vertex with two incoming lines labeled } i \text{ and } j \text{ from the left, and two outgoing lines labeled } a \text{ and } b \text{ to the right.} \\ (+1) \end{array} & \hat{T}_2 = & \begin{array}{c} \text{Diagram 2: A vertex with two incoming lines labeled } i \text{ and } j \text{ from the left, and two outgoing lines labeled } a \text{ and } b \text{ to the right.} \\ (+2) \end{array} & \hat{T}_3 = & \begin{array}{c} \text{Diagram 3: A vertex with three incoming lines labeled } i, j, \text{ and } k \text{ from the left, and three outgoing lines labeled } a, b, \text{ and } c \text{ to the right.} \\ (+3) \end{array}
 \end{array}$$

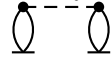


To obtain the diagrams which represent the coupled cluster energy, or some of the amplitudes equations, we need to combine diagrams which represent the Hamiltonian with diagrams which represent the cluster operator on all possible ways so that we can obtain the appropriate excitation level. When we have the needed diagrams, and using certain rules for interpretation we can obtain equations for both energies and amplitudes.

Interpretation rules for coupled-cluster diagrams

Here we will summarize the interpretation rules needed for obtaining coupled cluster diagrams. First, we will write some general rules for interpretation of different lines and vertexes in diagrams:

1. In one diagram all lines should have different labels. First the external (open) lines should be labeled. Downgoing lines are labeled with the hole target indices (i, j, \dots). Upgoing lines are labeled with the particle target indices (a, b, \dots). (The target indices are those that occur in the bra part of the amplitude equation.) After that, internal lines should be labeled with hole and particle indices that are different from the target indices.
2. Each one-particle interaction vertex $\text{---}\times\text{---}$ is associated with a factor $f_{out,in}$.
3. Each two-particle interaction vertex $\text{---}\times\text{---}\times\text{---}$ is associated with an antisymmetric two-electron integral $\langle \text{left-out right-in} | | \text{left-in right-out} \rangle$.
4. With every \hat{T}_m vertex $\text{---}\bigvee\text{---}\bigvee\text{---}$ an amplitude $t_{ij\dots}^{ab\dots}$ should be associated.
5. At the end, a summation over all indices should be performed.

We already saw that a phase factor of -1 can appear in some terms. Other factors also need to be accounted for and can be obtained following the rules:

1. A factor $\frac{1}{2}$ should be associated with each pair of equivalent internal lines. (Two internal lines are considered equivalent if they connect the same two vertices, going in the same direction.)
2. A factor $\frac{1}{2}$ should also be associated with each pair of equivalent \hat{T}_m vertices. (Two \hat{T} vertices are considered equivalent if they have the same number of line pairs and are connected in equivalent ways to the interaction vertex; examples are provided by the two \hat{T}_1 vertices in  and also in  but not those in .)
3. With each term a sign $-1^{(h-l)}$ should be associated, where h is the number of hole lines and l is the number of loops. For the purpose of counting loops, paired external lines (lines with labels such as (i,a) or (j,b) that are paired vertically in the bra part of the amplitude equation) are considered to be externally connected through imaginary extensions, forming quasiloops.
4. A summation over all distinct permutations \hat{P} of labels of inequivalent external particle lines and of inequivalent external hole lines, including a parity factor $(-1)^{\sigma(\hat{P})}$ should be performed. Lines that are quasi-equivalent (i.e. that would be equivalent if they became internal by the addition of an interaction vertex) are not considered inequivalent for this purpose. These sums over permutations are represented by operators of the form $\hat{P}(ij...|ab...)$.
5. In open diagrams with equivalent vertices, cancel each factor $\frac{1}{2}$ arising from rule 2 above with a permutation of the labels of a pair of external lines connected to the equivalent vertices (going in the same direction).

Correlation energy and amplitudes for CCSD

Now, having discussed basic diagrams and with the knowledge on how to combine them and the rules for reading them we can obtain the correlation energy for CCSD. Starting from the equation:

$$\langle \Phi_0 | \hat{H}(1 + \hat{T}_1 + (\frac{1}{2}\hat{T}_1^2 + \hat{T}_2)) | \Phi_0 \rangle = \langle \Phi_0 | E(1 + \hat{T}_1 + (\frac{1}{2}\hat{T}_1^2 + \hat{T}_2)) | \Phi_0 \rangle \quad (2.95)$$

we see that the only non-vanishing terms will correspond to combination of excitation operators with a resulting null excitation. Three possible combinations exist:

The diagram shows the equation $\Delta E =$ followed by three terms separated by plus signs. The first term is a two-loop diagram with two vertices (black dots) connected by a horizontal dashed line. Each vertex has two curved arrows forming a loop. The second term is a one-loop diagram with one vertex and two curved arrows forming a loop, with a dashed line extending from the top of the vertex to a cross symbol. The third term is another two-loop diagram with two vertices connected by a horizontal dashed line, and each vertex has two curved arrows forming a loop.

Translating our result into working equations we obtain an expression for the CCSD correlation energy:

$$\Delta E = \frac{1}{4} \sum_{ijab} \langle ij | ab \rangle t_{ij}^{ab} + \sum_{ia} f_{ia} t_i^a + \frac{1}{2} \sum_{ijab} \langle ij | ab \rangle t_i^a t_j^b \quad (2.96)$$

One should notice that this equation is for spin orbitals and to be able to compare the result with Equation (2.82), we need to transform these one- and two-electron integrals from spin orbitals into the spatial orbitals [2]. Therefore, we integrate over the spin functions α and β . For the one-electron integral, and after integration, we obtain:

$$\sum_i^N \langle \chi_i | \hat{h} | \chi_i \rangle = 2 \sum_i^{N/2} (\psi_i | \hat{h} | \psi_i) \quad (2.97)$$

where $\chi_i(\mathbf{x})$ is a spin orbital and ψ_i is the spatial orbital. The two-electron integral term is:

$$\begin{aligned} \frac{1}{2} \sum_{ab}^N \langle \chi_i \chi_j | \chi_i \chi_j \rangle &= \frac{1}{2} \sum_{ab}^N [\langle \chi_i \chi_j | \chi_i \chi_j \rangle - \langle \chi_i \chi_j | \chi_j \chi_i \rangle] \\ &= \sum_{ab}^{N/2} [2(\psi_a \psi_b | \psi_a \psi_b) - (\psi_a \psi_b | \psi_b \psi_a)]. \end{aligned} \quad (2.98)$$

Applying these equations to the expression for the energy we obtain:

$$\Delta E = \frac{1}{2} \sum_{ijab} [2(ij|ab) - (ij|ba)] t_{ij}^{ab} + 2 \sum_{ia} f_{ia} t_i^a + \sum_{ijab} [2(ij|ab) - (ij|ba)] t_i^a t_j^b. \quad (2.99)$$

We see that this equation is equivalent to the equation for the correlation energy obtained algebraically in Equation (2.82).

In order to calculate the singles amplitudes, we start from the equation:

$$\langle \Phi_i^a | \hat{H}_N (1 + \hat{T}_2 + \hat{T}_1 + \hat{T}_1 \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 + \frac{1}{3!} \hat{T}_1^3) | \Phi_0 \rangle = 0 \quad (2.100)$$

To represent this in a diagrammatic form we need to combine diagrams which represent one- and two-electron operators from the Hamiltonian with diagrams of cluster-operators resulting in a single excitation. The diagrams are shown in Figure 2.1.

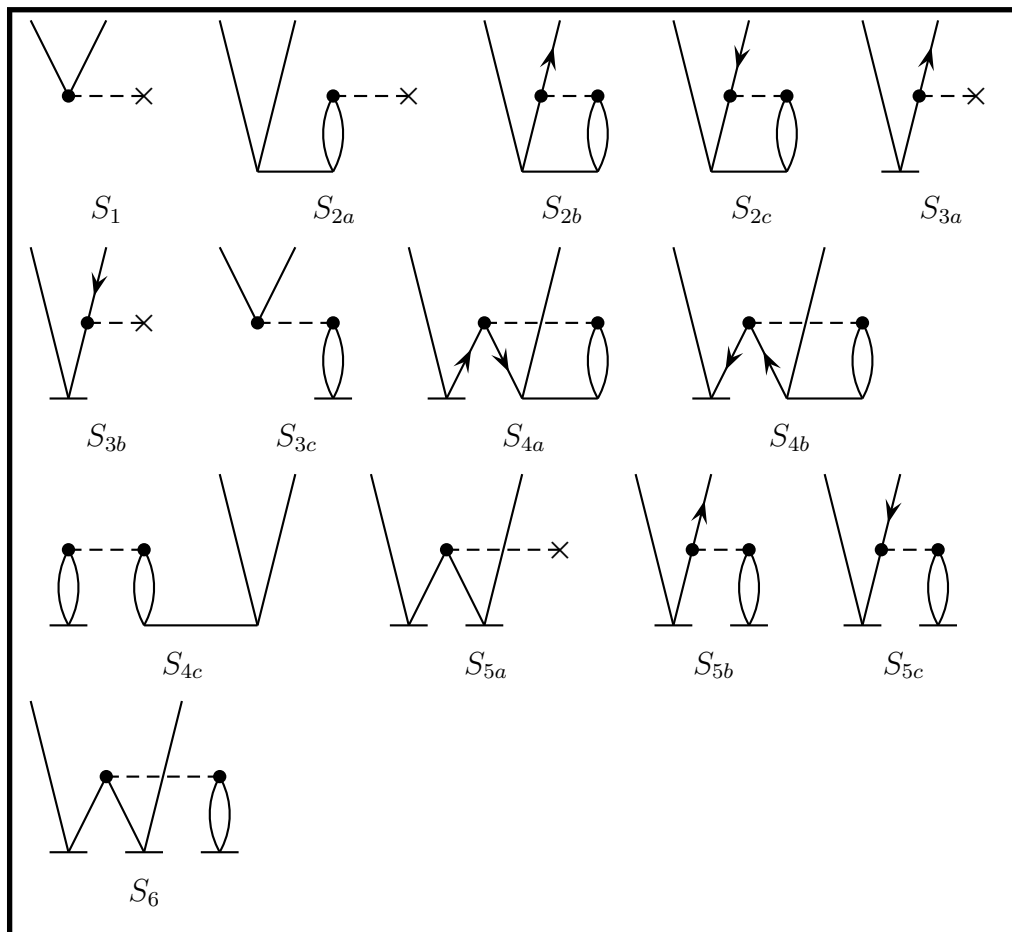


Figure 2.1: Antisymmetrized Goldstone diagrams representing the CCSD \hat{T}_1 equations.

The contributions obtained from the diagrams in Figure 2.1 are given in Table 2.1. The summations are over all the common indices k, l, c, d that occur in each term. It is also important to know that the algebraic interpretation of the diagrams assumes that labels are assigned in alphabetical order, from left to right, within each class of lines, under the condition that a and i are on the same continuous path, as well as b and j and so on.

Table 2.1: Interpretation of the diagrams for CCSD single residuals.

Diagram	Interpretation
1 (no \hat{T} vertex, requires +1 interaction vertex):	
S_1	f_{ai}
\hat{T}_2 (requires -1 interaction vertex):	
S_{2a}	$\sum f_{kc} t_{ik}^{ac}$
S_{2b}	$\frac{1}{2} \sum \langle ak cd \rangle t_{ik}^{cd}$
S_{2c}	$-\frac{1}{2} \sum \langle kl ic \rangle t_{kl}^{ac}$
\hat{T}_1 (requires 0 interaction vertex):	
S_{3a}	$\sum f_{ac} t_i^c$
S_{3b}	$-\sum f_{ki} t_k^a$
S_{3c}	$\sum \langle ak ic \rangle t_k^c$
$\hat{T}_1 \hat{T}_2$ (requires -2 interaction vertex):	
S_{4a}	$-\frac{1}{2} \sum \langle kl cd \rangle t_i^c t_{kl}^{ad}$
S_{4b}	$-\frac{1}{2} \sum \langle kl cd \rangle t_k^a t_{il}^{cd}$
S_{4c}	$\sum \langle kl cd \rangle t_k^c t_{li}^{da}$
$\frac{1}{2} \hat{T}_1^2$ (requires -1 interaction vertex):	
S_{5a}	$-\sum f_{kc} t_i^c t_k^a$
S_{5b}	$\sum \langle ak cd \rangle t_i^c t_k^d$
S_{5c}	$-\sum \langle kl ic \rangle t_k^a t_l^{ac}$
$\frac{1}{3!} \hat{T}_1^3$ (requires -2 interaction vertex):	
S_6	$-\sum \langle kl cd \rangle t_i^c t_k^a t_l^d$

In the case of the double residuals we have the equation:

$$\langle \Phi_{ij}^{ab} | \hat{H}_N (1 + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2 + \hat{T}_1 + \hat{T}_1 \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 + \frac{1}{2} \hat{T}_1^2 \hat{T}_2 + \frac{1}{3!} \hat{T}_1^3 + \frac{1}{4!} \hat{T}_1^4) | \Phi_0 \rangle = 0 \quad (2.101)$$

Here, we need to combine diagrams for the Hamiltonian and cluster operators to obtain double excitations. Diagrams for the \hat{T}_1 contributions to the double residuals are shown in Figure 2.3 and for the \hat{T}_2 contributions in Figure 2.2. The contributions to the double amplitudes obtained from the diagrams from Figures 2.3 and 2.2 are presented in Tables 2.2 and 2.3, respectively. The summations are over all the internal labels k, l, c, d

that occur in each term. More information about the diagrammatic representations of perturbation and coupled cluster theory can be found in the book on this topic written by Shavitt and Bartlett [3].

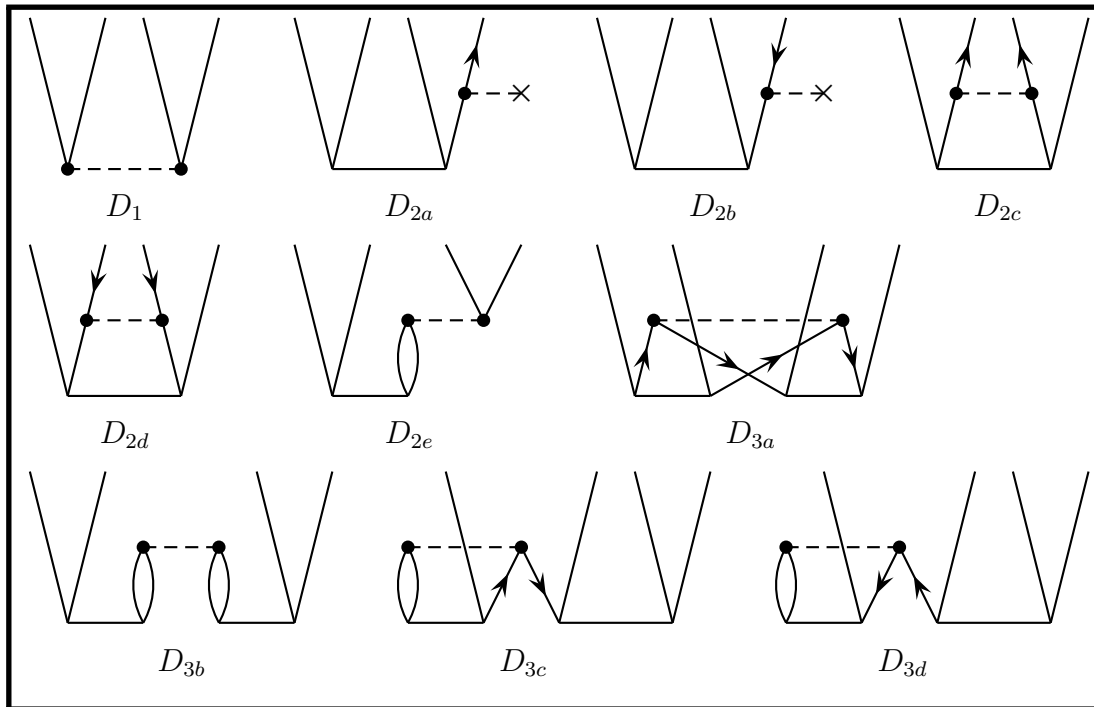


Figure 2.2: Antisymmetrized Goldstone diagrams representing \hat{T}_2 contributions to the CCSD \hat{T}_2 equations.

Table 2.2: Interpretation of diagrams for \hat{T}_2 contributions to the CCSD double residuals.

Diagram	Interpretation
1 (no \hat{T} vertex, requires +2 interaction vertex):	
D_1	$\langle ab ij \rangle$
\hat{T}_2 (requires 0 interaction vertex):	
D_{2a}	$\hat{P}(ab) \sum f_{bc} t_{ij}^{ac}$
D_{2b}	$-\hat{P}(ij) \sum f_{kj} t_{ik}^{ab}$
D_{2c}	$\frac{1}{2} \sum \langle ab cd \rangle t_{ij}^{cd}$
D_{2d}	$\frac{1}{2} \sum \langle kl ij \rangle t_{kl}^{ab}$
D_{2e}	$\hat{P}(ij ab) \sum \langle kb cj \rangle t_{ik}^{ac}$
$\frac{1}{2} \hat{T}_2^2$ (requires -2 interaction vertex):	
D_{3a}	$\frac{1}{4} \sum \langle kl cd \rangle t_{ij}^{cd} t_{kl}^{ab}$
D_{3b}	$\hat{P}(ij) \sum \langle kl cd \rangle t_{ik}^{ac} t_{jl}^{bd}$
D_{3c}	$-\frac{1}{2} \hat{P}(ij) \sum \langle kl cd \rangle t_{ik}^{dc} t_{lj}^{ab}$
D_{3d}	$-\frac{1}{2} \hat{P}(ab) \sum \langle kl cd \rangle t_{lk}^{ac} t_{ij}^{db}$

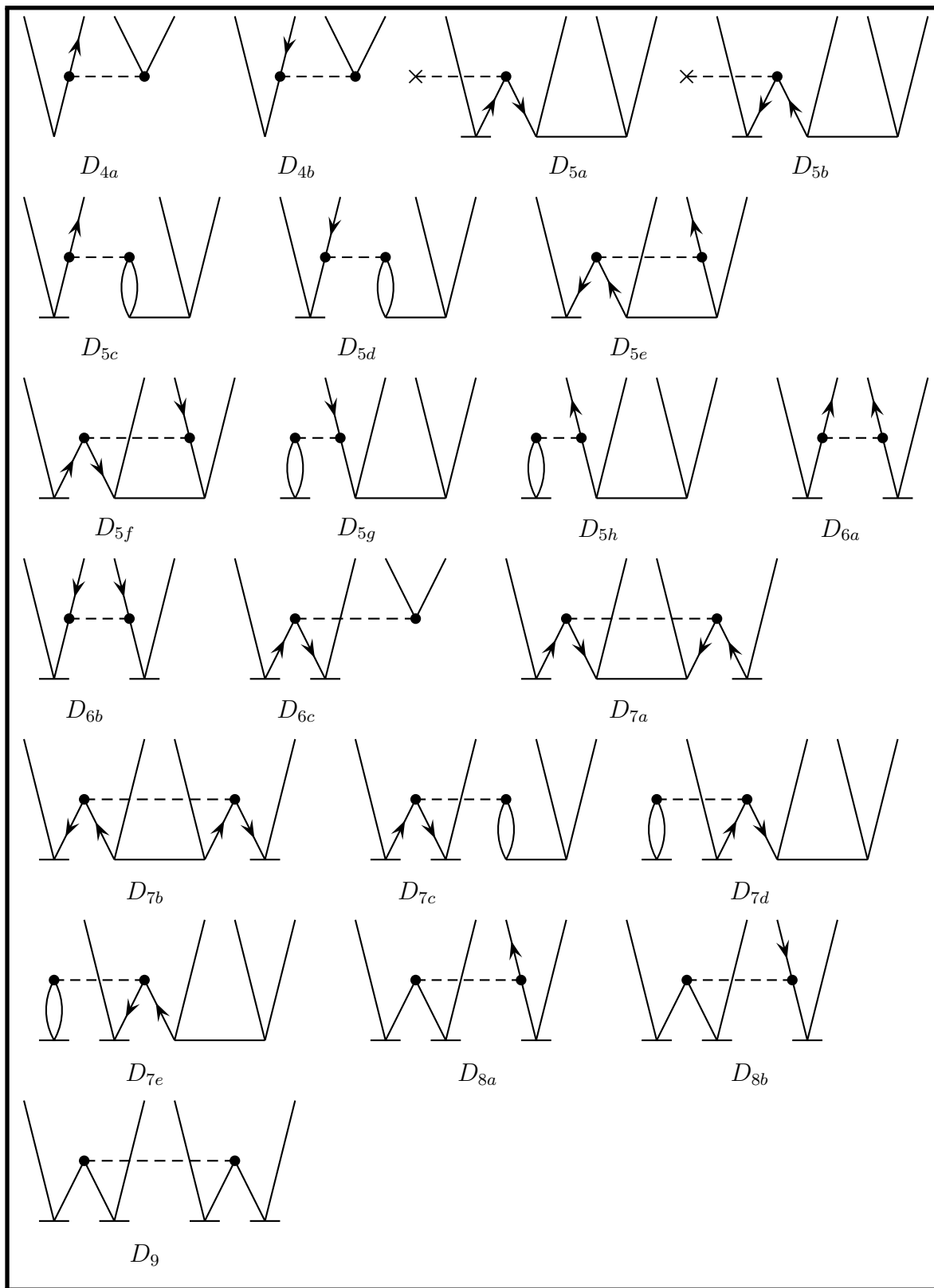


Figure 2.3: Antisymmetrized Goldstone diagrams representing \hat{T}_1 contributions to the CCSD \hat{T}_2 equations.

Table 2.3: Interpretation of diagrams for \hat{T}_1 contributions to the CCSD double residuals.

Diagram	Interpretation
\hat{T}_1 (requires +1 interaction vertex):	
D_{4a}	$\hat{P}(ij) \sum \langle ab cj \rangle t_i^c$
D_{4b}	$-\hat{P}(ab) \sum \langle kb ij \rangle t_k^a$
$\hat{T}_1 \hat{T}_2$ (requires -1 interaction vertex):	
D_{5a}	$-\hat{P}(ij) \sum f_{kc} t_i^c t_{kj}^{ab}$
D_{5b}	$-\hat{P}(ab) \sum f_{kc} t_k^a t_{ij}^{cb}$
D_{5c}	$\hat{P}(ij ab) \sum \langle ak cd \rangle t_i^c t_{kj}^{db}$
D_{5e}	$-\frac{1}{2} \hat{P}(ab) \sum \langle kb cd \rangle t_k^a t_{ij}^{cd}$
D_{5g}	$\hat{P}(ab) \sum \langle ka cd \rangle t_k^c t_{ij}^{db}$
D_{5d}	$-\hat{P}(ij ab) \sum \langle kl ic \rangle t_k^a t_{lj}^{cb}$
D_{5f}	$\frac{1}{2} \hat{P}(ij) \sum \langle kl cj \rangle t_i^c t_{kl}^{ab}$
D_{5h}	$-\hat{P}(ij) \sum \langle kl ci \rangle t_k^c t_{lj}^{ab}$
$\frac{1}{2} \hat{T}_1^2$ (requires 0 interaction vertex):	
D_{6a}	$\frac{1}{2} \hat{P}(ij) \sum \langle ab cd \rangle t_i^c t_j^d = \sum \langle ab cd \rangle t_i^c t_j^d$
D_{6b}	$\frac{1}{2} \hat{P}(ab) \sum \langle kl ij \rangle t_k^a t_l^b = \sum \langle kl ij \rangle t_k^a t_l^b$
D_{6c}	$-\hat{P}(ij ab) \sum \langle kb cj \rangle t_i^c t_k^a$
$\frac{1}{2} \hat{T}_1^2 \hat{T}_2$ (requires -2 interaction vertex):	
D_{7a}	$\frac{1}{4} \hat{P}(ij) \sum \langle kl cd \rangle t_i^c t_{kl}^{ab} t_j^d = \frac{1}{2} \sum \langle kl cd \rangle t_i^c t_{kl}^{ab} t_j^d$
D_{7b}	$\frac{1}{4} \hat{P}(ab) \sum \langle kl cd \rangle t_k^a t_{ij}^{cd} t_l^b = \frac{1}{2} \sum \langle kl cd \rangle t_k^a t_{ij}^{cd} t_l^b$
D_{7c}	$-\hat{P}(ij ab) \sum \langle kl cd \rangle t_i^c t_k^a t_{lj}^{db}$
D_{7d}	$-\hat{P}(ij) \sum \langle kl cd \rangle t_k^c t_i^d t_{lj}^{ab}$
D_{7e}	$-\hat{P}(ab) \sum \langle kl cd \rangle t_k^c t_l^a t_{ij}^{db}$
$\frac{1}{3!} \hat{T}_1^3$ (requires -1 interaction vertex):	
D_{8a}	$\frac{1}{2} \hat{P}(ij ab) \sum \langle kb cd \rangle t_i^c t_k^a t_j^d = \hat{P}(ab) \sum \langle kb cd \rangle t_i^c t_k^a t_j^d$
D_{8b}	$\frac{1}{2} \hat{P}(ij ab) \sum \langle kl cj \rangle t_i^c t_k^a t_l^b = \hat{P}(ij) \sum \langle kl cj \rangle t_i^c t_k^a t_l^b$
$\frac{1}{4!} \hat{T}_1^4$ (requires -2 interaction vertex):	
D_9	$\frac{1}{4} \hat{P}(ij ab) \sum \langle kl cd \rangle t_i^c t_j^d t_k^a t_l^b = \sum \langle kl cd \rangle t_i^c t_j^d t_k^a t_l^b$

2.3.3 Perturbative Triple Excitations

It is known that CCSD does not treat electron correlation properly and for example in the case of atomization energies, the errors are of the same magnitude as in the case of MP2 but of opposite sign [1]. To improve these results one needs to include higher-order excitations. In most cases the inclusion of triple excitations is sufficient. However, this leads to a computationally demanding method which scales as N^8 . Nonetheless, the most important terms are those linear in triple excitations as it was shown by Rayleigh-Schrödinger perturbation theory (RSPT) for the treatment of higher excitations. Therefore, the correlation energy of CCSD (Equation (2.82)) can be improved by the inclusion of a perturbative correction of the form:

$$\Delta E^{(T)} = \langle \Psi^{\text{HF}} | (\hat{T}_1 + \hat{T}_2)^\dagger \hat{V} \hat{T}_3 | \Psi^{\text{HF}} \rangle, \quad (2.102)$$

where \hat{V} represents the perturbation operator. This method is generally denoted as CCSD(T) [7–9]. In this case, the CCSD results for atomization energies are improved by an order of magnitude [1]. A coupling between the individual amplitudes is not present in this method and the correction can be obtained non-iteratively. Despite that this method scales as N^7 , since the CCSD iterations are done independently, a more cost effective approach is obtained than in the case of the full triples. The CCSD(T) method is the "gold" standard in most of the fields of computational chemistry and it is found that it is generally more accurate than CCSDT.

2.4 Local Correlation Methods

Dynamic electron correlation is the instantaneous interaction between electrons in movement and it is a short range effect which decreases as r^{-6} . In the previously presented methods this property of dynamic correlation was not taken into account, since canonical orbitals were used. Canonical orbitals are highly delocalized over the whole system. This causes the unfortunate scaling with the number of electrons in these methods. To overcome this problem a mathematically equivalent representation of the canonical orbitals has been suggested. Different methods which are using the locality of electron correlation have been introduced [10–15]. The most successful one was introduced by Pulay [16] and first implemented by Pulay and Saebø [10, 17–20]. The occupied orbital

space is transformed using an unitary transformation to the local orbital basis:

$$|\phi_i^{loc}\rangle = \sum_{\mu} |\chi_{\mu}\rangle L_{\mu i} = \sum_k |\phi_i^{can}\rangle U_{ki} \quad (2.103)$$

where $\mathbf{L}=\mathbf{C}\mathbf{U}$. Different localization schemes have been proposed, the most popular ones from Boys [21], Pipek-Mezey [22] and Edmiston-Ruedenberg [23]. The Boys localization procedure is a very inexpensive procedure where the distance between orbital centroids is maximized. On the other hand, this procedure has problems with radial localization as well as with multiple bonds, which are usually represented as "banana bonds". In the case of the Pipek-Mezey localization scheme this is not the case and the σ and π orbitals can be separated by maximizing the orbital partial charges. The Edmiston-Ruedenberg localization scheme has the same advantages as the Pipek-Mezey scheme, but due to the transformation of the two-electron integrals it is quite expensive. In the last years, an alternative localization scheme was proposed which made use of natural localized molecular orbitals (NLMO) [24].

The virtual orbital space can be obtained from projected atomic orbitals (PAOs) [16], pair natural orbitals (PNOs) [25] or orbital specific virtuals (OSVs) [26]. In the case of PAOs, localized molecular orbitals (LMOs) are projected out of the atomic orbital basis:

$$|\tilde{\chi}_r\rangle = (1 - \sum_i |\phi_i^{loc}\rangle \langle \phi_i^{loc}|) |\chi_r\rangle = \sum_{\mu} |\chi_{\mu}\rangle P_{\mu r}, \quad (2.104)$$

with the projection matrix \mathbf{P} given by:

$$\mathbf{P} = \mathbf{1} - \mathbf{L}\mathbf{L}^{\dagger}\mathbf{S}. \quad (2.105)$$

This combination of occupied and virtual orbitals has some specific properties:

- both occupied and virtual space are localized.
- the occupied orbitals are orthogonal among themselves and also they are orthogonal to the virtual space.
- the virtual orbitals are not orthogonal among themselves and have the overlap defined as:

$$\langle \tilde{\chi}_r | \tilde{\chi}_s \rangle = (\mathbf{P}^{\dagger}\mathbf{S}\mathbf{P})_{rs} = \tilde{\mathbf{S}}_{rs}. \quad (2.106)$$

- linear dependencies in the virtual space are present.

2.4.1 Selection of domains and orbital pairs

Two approximations are present in local correlation methods. The first one is the domain approximation. In the domain approximation the excitations are restricted to subspaces of PAOs, so called domains $[i]$. The second approximation is the pair approximation. This is coming from the fact that the correlation energy decreases quickly with the distance between two correlated localized orbitals. Therefore, the orbital pairs very far from each other can be neglected or treated at a lower level of theory.

The construction of the domains is carried out using one of the two available criteria for the domain selection. The first one was proposed by Boughton and Pulay [27]. In this case one starts by ordering the atoms according to decreasing Löwdin charges:

$$l_A^i = 2 \sum_{\mu \in A} [\mathbf{S}^{\frac{1}{2}} \mathbf{L}]_{\mu i}. \quad (2.107)$$

Atoms are added to the domain list according to a threshold and the overlap criteria. The Boughton-Pulay criteria can be represented as:

$$B^i = 1 - \int |(\phi_i - \hat{\phi}_i)^2| d\tau > T_{BP}, \quad (2.108)$$

where T_{BP} is the threshold. This threshold should vary with the basis set, and the recommended values are $T_{BP} = 0.980$, $T_{BP} = 0.985$ and $T_{BP} = 0.990$ for double, triple and quadruple-zeta basis set quality, respectively. Pair domains are the union of single domains $[ij] = [i] \cup [j]$.

The second one is Natural Population Analysis-based domain criteria (NPA-based domain criteria) [24]. In this case the atoms are added to the domain if they exceed a certain threshold T_{NPA} for the NPA charges. For T_{NPA} recommended values are between 0.03 and 0.01. It was found that the NPA domain criteria is less sensitive to the change of the basis set than the Boughton-Pulay domain criteria [24].

In the case where aromatic molecules with high symmetry are present, it can happen that the localization is not unique. An example for this is benzene, where the localized π -orbitals can be freely rotated. In these cases, the merging of domains is the recommended procedure. In the case of benzene one would merge the three π -orbital domains into one

domain, and this one domain is then used for all three orbitals. This procedure is also useful in the case of reaction pathways, where it can happen that domains change from one to another stationary point. In that case the domains are merged to correspond to the largest domain definition along the pathway.

The orbital pairs (ij) can be classified using a distance, connectivity or even mixed distance-connectivity criteria. In the distance criteria the orbital pairs are classified according to the closest distance $R^{(ij)}$ between atoms in the primary domains $[i]$ and $[j]$. The connectivity criteria is on the other hand dependent on the minimum number of bonds between the two orbital domains. In the case of the mixed criteria one of the conditions should be fulfilled, distance or connectivity, to assign orbital pairs into the appropriate group. Five different pair types are present: strong pairs, close pairs, weak pairs, distant pairs and very distant pairs (Table 2.4).

Table 2.4: Pair types used in the local coupled cluster calculations.

Strong	$r_p < R_c$	Treated at the CCSD level
Close	$R_c \leq r_p < R_w$	Treated at the MP2 level and included in the triples calculation
Weak	$R_w \leq r_p < R_d$	Treated at the MP2 level
Distant	$R_d \leq r_p < R_{vd}$	Treated at the MP2 level
Very distant	$r_p \geq R_{vd}$	Neglected

The strong pairs include the largest contributions to the correlation energy and they are treated at the highest level of theory, for example at LCCSD(T0). Close, weak and distant pairs are normally treated at the LMP2 level, and very distant pairs are neglected. Coupling between strong and close pairs is possible through the use of LMP2 amplitudes of the close pairs in the calculation of the LCCSD residuals of strong pairs. It is also important to know that close pairs may not have any influence on the LCCSD energy, but affect the treatment of the triple excitations. An important feature of this classification is that only the number of very distant pairs scales quadratically with the molecular size, and all other classes scale linearly, independent of the choice of the distance criteria.

The default values for the distance criteria are $R_c=1$, $R_w=3$, $R_d=8$ and $R_{vd}=15$ Bohr. In the case of the connectivity criteria $I_c=1$, $I_w=2$, $I_d=5$ and $I_{vd}=8$ are default values. This means that strong pairs are separated by no bonds, close pairs by 1 bond, weak

pairs by 2-4 bonds, distant pairs by 5-7 bonds and very distant pairs by at least 8 bonds. The advantage of the connectivity criteria is the independence of the bond lengths and the advantage of the distance criteria is that it is also effective in non-bonding situations.

Triple excitations are also restricted to domains. The triples list (ijk) contains the pairs (ij) , (ik) and (jk) and they must be either strong or close pairs. Additionally, at least one of these pairs must be strong. In the case of the calculation of the triple contributions, amplitudes of close pairs from LMP2 calculation are included. This is important to obtain accurate triples corrections.

2.4.2 Local Møller-Plesset Method

When we know how to localize the occupied and virtual space we can rewrite the MP2 method in this scope. We start with defining the first order wave function in the LMO and PAO basis as:

$$|\Psi^{(1)}\rangle = \frac{1}{2} \sum_{ij \in P} \sum_{rs \in [ij]} \tilde{T}_{rs}^{ij} |\Phi_{ij}^{rs}\rangle \quad \text{with} \quad \tilde{T}_{rs}^{ij} = \tilde{T}_{sr}^{ji}, \quad (2.109)$$

where P represents the pair list and $[ij]$ are pair domains. \tilde{T}_{rs}^{ij} are amplitudes in the PAO basis [28]. One should notice that the number of projected functions $r, s \in [ij]$ for a given pair (ij) is independent of molecular size. The next step is to transform the equation for the residuals from the canonical into the local basis. The detailed derivation can be found in Reference [28] and here is shown only the final result:

$$\begin{aligned} \tilde{R}_{rs}^{ij} = & \tilde{K}_{rs}^{ij} + \sum_{tu \in [ij]} \tilde{f}_{rt} \tilde{T}_{tu}^{ij} \tilde{S}_{us} + \sum_{tu \in [ij]} \tilde{S}_{rt} \tilde{T}_{tu}^{ij} \tilde{f}_{us} \\ & - \sum_k \left[\sum_{tu \in [kj]} \tilde{S}_{rt} \tilde{f}_{ik} \tilde{T}_{tu}^{kj} \tilde{S}_{us} + \sum_{tu \in [ki]} \tilde{S}_{rt} \tilde{f}_{kj} \tilde{T}_{tu}^{ik} \tilde{S}_{us} \right] = 0 \end{aligned} \quad (2.110)$$

where \tilde{S}_{rs} is the overlap matrix. If one compares above equation with Equation (2.49) one can see that the only difference is that in the above equation the overlap matrix is present. The occupied-occupied and virtual-virtual blocks of the Fock matrices f are not diagonal in the case of the local basis. Therefore, the equation for the residuals has to be solved iteratively for the amplitudes \tilde{T}_{rs}^{ij} . In the case of local methods the excitations are restricted to the domains $[ij]$ of PAOs which implies that only the elements \tilde{T}_{rs}^{ij} with

$r, s \in [ij]$ are non-zero. Thus, only the corresponding elements of the residuals \tilde{R}_{rs}^{ij} must vanish at the convergence. In the end, the second order energy can be calculated as:

$$E^{\text{LMP2}} = \sum_{ij \in P} \sum_{rs \in [ij]} (2\tilde{T}_{rs}^{ij} - \tilde{T}_{sr}^{ij}) \tilde{K}_{rs}^{ij}, \quad (2.111)$$

where \tilde{K}_{rs}^{ij} are exchange integrals. Only a small subset of these exchange integrals is needed to compute the residuals [29]. The exchange integrals which are needed are those of the form:

$$\tilde{K}_{rs}^{ij} = (ri|sj) = \sum_{\nu\mu} P_{\mu r} P_{\nu s} \left[\sum_{\rho\sigma} L_{\rho i} L_{\sigma j} (\mu\rho|\nu\sigma) \right], \quad (2.112)$$

where all r, s are close either to i or j . The calculation of these exchange integrals is the most demanding part in the LMP2 calculations. However, due to the domain and pair approximations, the LMP2 method runs significantly faster in comparison with MP2. Analytical gradients for LMP2 have been also developed by Werner and coworkers [30, 31]. Density fitting approximations can be used in the combination with LMP2. In this way the speed of the calculation can be improved with negligible errors [32].

Due to the local character of the occupied and virtual spaces, one can decompose the intermolecular interaction energy into individual contributions of different excitation classes [33]. The excitation classes are shown in Figure 2.4 and explained in detail in the following list:

- (a) Intramolecular correlation effects are described by double substitutions $i \rightarrow [i, i']$ and $i' \rightarrow [i, i']$, where LMOs i, i' and their domains are located on the same monomer.
- (b) Dispersive coupling between the monomers is described by the simultaneous single excitations on two different monomers $i \rightarrow [i], j \rightarrow [j]$.
- (c) Dispersion-exchange effects are described by cross excitations, where an electron from an occupied LMOs of monomer A is transferred to the correlating space of the monomer B ($i \rightarrow [j], j \rightarrow [i]$).
- (d) Ionic effects represent the single excitation on one monomer coupled with the transfer of an electron from an occupied LMOs from monomer B to the virtual

space of monomer A ($i \rightarrow [i], j \rightarrow [i]$).

- (e) Double ionic substitutions from the type $j \rightarrow [i]$ and $j' \rightarrow [i]$ are responsible for basis set superposition errors (BSSE) and they are excluded from local methods by construction.
- (f) A second class of ionic substitutions from the type $i \rightarrow [i]$ and $i' \rightarrow [j]$ are also excluded from local methods by construction, yet they should not contribute significantly to the correlation energy.

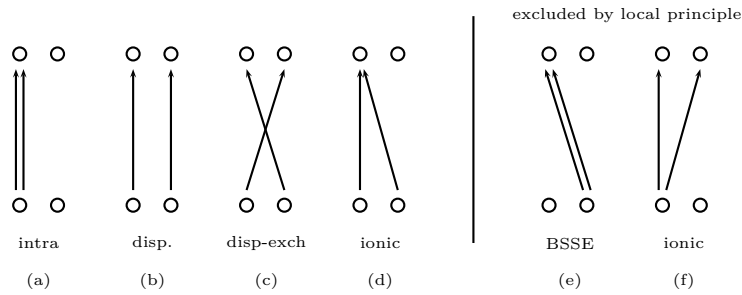


Figure 2.4: Schematic representation of the different double excitation classes of local correlation methods in the context of intermolecular interactions. The lower and upper circles represent different monomers in the ground and excited states, respectively. The arrows symbolize the corresponding excitations. Adapted from Reference [33].

2.4.3 Local Coupled Cluster Method

Just as in the LMP2 case one can obtain the LCCSD equations by transforming the residuals from the molecular orbital to the LMO/PAO basis. Equations obtained in this way differ from canonical ones only in additional matrix multiplications with the PAO overlap matrix and couplings through the non-diagonal Fock-matrix. It was also shown that in the case of double residuals the most efficient explicit formulation is obtained using the contravariant functions [34, 35]. As in the case of LMP2, the coefficient matrices \tilde{T}_{rs}^{ij} and the vectors \tilde{t}_r^i are nonzero if $r, s \in [ij]$ and $r \in [ii]$, respectively. Therefore, it is sufficient to compute only the corresponding residuals \tilde{R}^{ij} and \tilde{R}^i :

$$\tilde{\mathbf{R}}^{ij} = \tilde{\mathbf{K}}^{ij} + \tilde{\mathbf{K}}(\tilde{\mathbf{D}}^{ij}) + \tilde{\mathbf{G}}^{ij} + \tilde{\mathbf{G}}^{ji\dagger} + \tilde{\mathbf{S}} \left[\sum_{kl} (\alpha_{ij,kl} - \delta_{jl}\beta_{ki} - \delta_{ki}\beta_{lj}) \tilde{\mathbf{C}}^{kl} \right] \tilde{\mathbf{S}}, \quad (2.113)$$

$$\tilde{\mathbf{R}}^i = \tilde{\mathbf{s}}^i + \tilde{\mathbf{S}} \left[\sum_k (2\tilde{\mathbf{T}}^{ik} - \tilde{\mathbf{T}}^{ki}) \tilde{\mathbf{r}}^k - \beta_{ki} \tilde{\mathbf{t}}^k \right]. \quad (2.114)$$

The detailed derivation can be found in Reference [36]. At the end, for the correlation energy one obtains:

$$E_{corr} = \sum_{i \geq j} (2 - \delta_{ij}) \text{tr}[\tilde{\mathbf{L}}^{ij} \tilde{\mathbf{C}}^{ij}] + 2 \sum_i \tilde{\mathbf{f}}^{i\dagger} \tilde{\mathbf{t}}^i. \quad (2.115)$$

For one to compute the local perturbative triples correction (T) [37, 38] the triple excitations $\hat{E}_i^r \hat{E}_j^s \hat{E}_k^t$ should be restricted to triples domains $[ijk] = [i] \cup [j] \cup [k]$. Furthermore, the list of the orbital triples ijk is also restricted on the way that related pairs ij , ki and kj are close or strong and at least one of them is strong. In the triples equations strong and close pair amplitudes are included as well as the single amplitudes.

In the local basis the Fock matrix is not diagonal. Therefore, the perturbative triples equations have to be solved iteratively. Despite that this is possible, this is not commonly used since it is very expensive. All triple amplitudes would have to be saved. To overcome this problem, the off-diagonal elements of the Fock matrix f_{ij} of the occupied-occupied block can be neglected. This is called the T0 approximation [37, 38]. In this way triples can be computed very efficiently and about 97% of the full local (T) energy are recovered independently of the size of the molecular system as well as the settings of R_s and R_c .

2.5 Quantum Mechanics/Quantum Mechanics Method

It is the common case that one deals with system sizes too large to be investigated with conventional wavefunction QM methods. However, one is often interested in a chemical event, so the most important part of the molecule is relatively localized around a point of interest. In inorganic catalysis, for example, this would correspond to the metal center or the part where the bond formation/breaking occurs. Therefore, one can restrict the use of accurate correlation methods to an even smaller part. In the scope of local correlation methods this can be done by assigning groups of LMOs and related domains to the specific region and then apply different methods to the different regions. This approach is called Local Molecular Orbital : Molecular Orbital (LMOMO) method [39].

The LMOs can be assigned to the region with the following procedure:

1. A list of atoms for the high level region as well as an assignment of methods for the specific region are given as input.
2. The LMOs are assigned to the high level region if in their domain list is at least one atom which is in the high level region.
3. If there is a second region the first two points are repeated.
4. The rest of the orbitals is treated at the low level region which is per default HF, if it is not specified in the first point.

On this way only one calculation is necessary to obtain the result. Also, the definition of a model system which would be treated at the high level is not needed and cutting of bonds is avoided. Since the high level region is computed in the presence of the low level region the coupling between the regions is included in the calculation. All these points are an improvement over the methods which are more in use, e.g. IMOMO [40].

In the case when LCCSD(T0) and LMP2 methods are used for the high and low level regions, respectively, the correlation energy of the LMOMO method is calculated as:

$$E^{\text{corr}}[\text{LCCSD(T0)} : \text{LMP2}] = \sum_{ijk \in \text{HL}} \varepsilon_{ijk, [ijk]}^{\text{CC}} + \sum_{ij \in \text{HL}} \varepsilon_{ij, [ij]}^{\text{CC}} + \sum_{ij \notin \text{HL}} \varepsilon_{ij, [ij]}^{\text{MP2}}, \quad (2.116)$$

where HL and LL are high level and low level, respectively. One can observe that triples are only computed within the HL region.

2.6 Density Functional Theory

The theoretical footing for density functional theory (DFT) was introduced through the seminal work of Hohenberg and Kohn, consisting of two theorems [41]. The first theorem states that the non-degenerated density determinant of the ground state uniquely determines the external potential. The second theorem states that the energy functional reaches its minimum for the ground state density corresponding to the external potential. From these two theorems one can conclude that the density determines the external potential, determining the Hamiltonian which ultimately determines the wave function.

Thomas and Fermi introduced the first form of DFT, where all energy components were represented as density functionals. In this way one has only three coordinates

instead of $3N$ coordinates (after spin integration) which are necessary in wave function methods. The energy functional can be divided into specific components which are easier to analyze:

$$E[\rho(\mathbf{r})] = T_{ni}[\rho(\mathbf{r})] + V_{ne}[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + \Delta T[\rho(\mathbf{r})] + \Delta V_{ee}[\rho(\mathbf{r})], \quad (2.117)$$

where T_{ni} is the kinetic energy of non-interacting electrons, the following two terms are the nuclear-electron and the classical electron-electron repulsion and the last two terms are a correction to the kinetic energy, derived from the interacting nature of the electrons and all non-classical corrections to the electron-electron repulsion. If we look only at the last two terms their sum represents the exchange-correlation energy, E_{xc} . Since Thomas-Fermi DFT as well as an improved version of it, the Thomas-Fermi-Dirac model, predict that all molecules are unstable relative to the dissociation into their constituents, DFT did not have any impact on chemistry at the time.

The real improvement of DFT and its application in quantum chemistry started when Kohn and Sham proposed that one can use orbitals to represent the electron density when the kinetic energy is calculated [42]. By self-consistently solving the equation:

$$\left[-\frac{1}{2}\nabla^2 + V(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right] |\phi_i\rangle = \epsilon_i |\phi_i\rangle \quad (2.118)$$

for each orbital and defining the electron density as:

$$\rho(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2, \quad (2.119)$$

one obtains the so called Kohn-Sham (KS) orbitals. Their meaning as well as the meaning of respective energies is not clear. A one-to-one correspondence between the density and the orbitals is not proven.

The energy functional can be now written in the orbital formalism:

$$\begin{aligned} E[\rho(\mathbf{r})] = & -\frac{1}{2} \sum_i^N \langle \phi_i | \nabla_i^2 | \phi_i \rangle - \sum_i^N \sum_k^M \langle \phi_i | \frac{Z_k}{|\mathbf{r}_i - \mathbf{r}_k|} | \phi_i \rangle + \\ & + \frac{1}{2} \sum_i^N \langle \phi_i | \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' | \phi_i \rangle + E_{xc}[\rho(\mathbf{r})] \quad , \end{aligned} \quad (2.120)$$

where N is the number of electrons, M is the number of nuclei and $\rho(\mathbf{r}')$ is the density for the non-interacting system.

By examining Equation (2.120) we see that DFT is an exact theory if E_{xc} is known. Unfortunately, this is not the case. Therefore, different approximations to E_{xc} have been proposed over the years which have resulted in the endless list of functionals in the literature. Some groups suggested that E_{xc} in a given position can be computed exclusively from the value of the density in that position. This is known as the local density approximation (LDA). Density functionals which fit this description are derived from the uniform electron gas. The most used LDA functionals are the functionals derived by Vosko *et al.* [43] and Perdew [44].

LDA functionals result in large errors when applied to molecules since the uniform electron gas approximation is inadequate in such cases. To correct for this behavior the gradient of the density has been introduced. The corrections can be added to the exchange or correlation part of E_{xc} . This is known as the generalized gradient approximation (GGA). Some examples for this family are the B86 [45, 46] and PBE [47] functionals for exchange and the B88 [45] and LYP [48] for the correlation.

It is often found that the inclusion of exact HF exchange improves the results. Therefore, a new class of functionals is introduced, where the exact HF exchange is combined with the exchange and correlation from other sources. This class of functionals is called hybrid functionals and the most famous among these methods is the B3LYP [49] functional. In this case, E_{xc} is defined as:

$$E_{xc}^{\text{B3LYP}} = 0.2E_x^{\text{HF}} + 0.72E_x^{\text{B88}} + 0.08E_x^{\text{S}} + 0.81E_c^{\text{LYP}} + 0.19E_c^{\text{VWN80}} \quad (2.121)$$

where the exchange is composed of the exact Hartree-Fock exchange (E_x^{HF}), the exchange from the gradient corrected B88 functional (E_x^{B88}) and the Slater-Dirac exchange (E_x^{S}) [50]. The correlation part is built from the gradient corrected and local approximations, E_c^{LYP} and E_c^{VWN80} [43], respectively. Present numerical parameters were fitted to reproduce atomization energies, proton affinities and ionization potentials of the G1 molecule test.

Despite the large application field of DFT, this theory has some disadvantages. The first one is that although the theory promises an universally valid functional, this has still not been found. Another point is that most of the available functionals will work sensible only in the systems for which they were parametrized. Further, among the

proposed density functionals, a well defined hierarchy does not exist, thus there is no systematic way to improve on a DFT estimate. At the end one can point out that the current functionals are unable to describe weak interactions like dispersion forces. In the last years significant work was dedicated to overcome this problem. As a result different empirically based corrections to the DFT functionals have been proposed with significant success [51–53].

CHAPTER 3

Aurophilic Complexes

In the late 1970s a rather unusual interaction was observed in the field of structural gold chemistry. The attractive interaction between linearly two-coordinated gold complexes in the +1 oxidation state, a d^{10} closed shell system was verified. In the absence of steric hindrance, the equilibrium distance between two gold atoms in these complexes was in the range from 2.50–3.50 Å, which is below the sum of two van der Waals radii (3.80 Å). Therefore, it was proposed that these complexes were stabilized by the attraction of the gold kations, which was not possible to predict theoretically. In the late eighties, Schmidbaur, Jansen and respective coworkers [54–56] found empirical evidence for an $\text{Au} \cdots \text{Au}$ attraction of about -30 kJ/mol. This attraction was named aurophilicity. The effect was later observed for many other metals with closed d-shell, giving rise to the concept of metallophilicity.

Theoretical work in this area was pioneered by Pyykkö [57]. His group carried out HF and MP2 calculations to quantify the interaction between gold atoms in $[\text{X-Au-PH}_3]_2$, where $\text{X} = \text{H}$ or Cl . It was discovered that the HF level interaction energy curves were repulsive, but when electron correlation was included at the MP2 level attractive curves were obtained. This finding was particularly relevant since till then aurophilicity was believed to be an hybridization effect. This theory faded out since in such a case the interaction would be already visible at the HF level. Through the contributions from several research groups, Pyykkö and coworkers foremost, we know today that aurophilicity is linked to electron correlation and is significantly strengthened by relativity.

Three categories of aurophilic bonding are known: semi-supported, fully-supported and unsupported interactions (Figure 3.1). Semi-supported and fully-supported interactions are intramolecular interactions where the gold atoms are connected through two ligands. In rare cases the coordination number of gold can increase to three. In the case of intermolecular aurophilic bonding any interaction is unsupported.

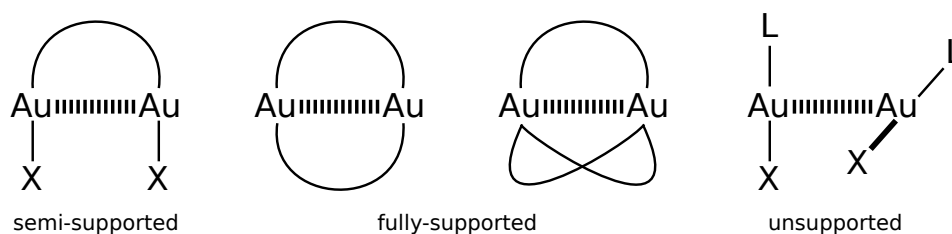


Figure 3.1: Three types of interaction in aurophilic systems.

3.1 Ligand Effects in Auophilic Interactions

In this work we focused on the effect of electron correlation, as well as the interplay between Au...Au contacts and other weak interactions involving the ligands in unsupported systems. We have investigated bicoordinated complexes with the general formula Cl-Au-L. All ligands (L) used in our study are shown in Figure 3.2. First, we benchmarked electron correlation methods by calculating the potential energy curve of the Cl-Au-PH₃ dimer.

In later sections, we looked into larger ligands and also evaluated the influence of the different dimer orientations on the results. For this purpose Au(I) complexes with an imidazol-2-ylidene (NHC-H) and a phosphonium ylide (C(PH₃)₂) were used. Different orientations of the complexes are shown in Figure 3.3. In a previous work, Muniz *et al.* [58] already investigated the effect of varying ligands. Their study included a variety of N-heterocyclic carbenes, phosphanes and other ligands. They came to the conclusion that various N-heterocyclic carbenes yield the strongest auophilic interaction. Another aspect of their research was the influence of different orientation of monomers on the interaction energy. They found that large calculated difference between the different orientation of monomers was present in these systems. However, it was still unclear how the weight of the d¹⁰-d¹⁰ interaction changes between the different complexes and/or their orientations.

In a final set of calculations two more realistic systems have been considered. The first one is ClAu[1-(benzyl)-3-(N-*tert*-butylacetamido)imidazole-2-ylidene] which is referred to as Cl-Au-(NHC-B) and the second one is ClAu[1,3-bismethylbenzimidazole-2-ylidene], referenced as Cl-Au-(Me₂bimy).

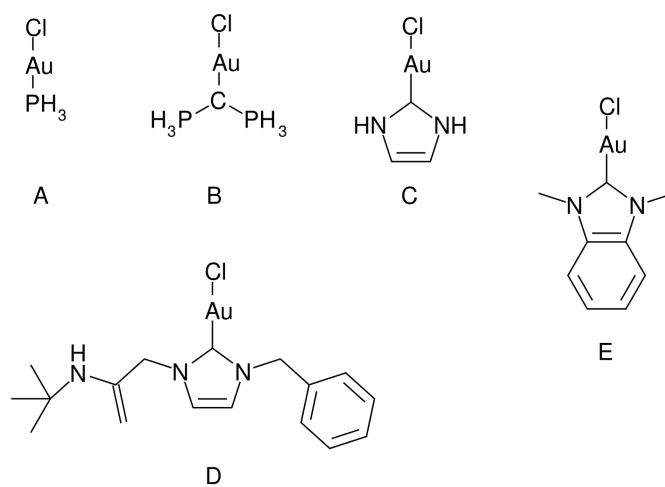


Figure 3.2: Lewis representation of the monomers considered in this study. (A) $\text{L} = \text{PH}_3$, (B) $\text{L} = \text{C}(\text{PH}_3)_2$, (C) $\text{L} = \text{NHC-H}$, (D) $\text{L} = \text{NHC-B}$ and (E) $\text{L} = \text{Me}_2\text{bimy}$.

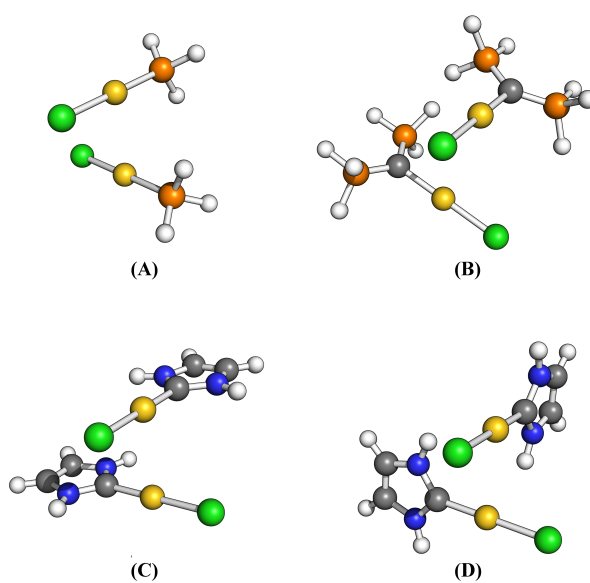


Figure 3.3: Orientation of Cl-Au-PH_3 (A), $\text{Cl-Au-C}(\text{PH}_3)_2$ (B) and both Cl-Au-NHC-H dimers: (C) face-to-face (ff) and (D) edge-to-edge (ee).

3.1.1 Computational Details

The correlation energy can be decomposed into strictly localized pair contributions within local correlation methodologies (e.g. LMP2 or LCCSD). These pairs are localized in a given region of the monomer. In the case of the dimers $(\text{Cl}_A\text{-Au}_A\text{-L}_A)(\text{Cl}_B\text{-Au}_B\text{-L}_B)$, the energy can be divided into six classes, each computed as:

$$\Delta E(X \cdots Y) = \Delta E(X_A \cdots Y_B) + \Delta E(X_B \cdots Y_A), \quad (3.1)$$

with $X, Y = \text{Au}, \text{Cl}$ or L . Each ΔE refers to the correlation component of the interaction energy between two fragments $X \cdots Y$. Since gold shares a covalent bond with chloride and the carbene ligand, the terms arising from these two covalent bonding orbitals have to be split between the respective fragments. The Natural Population Analysis (NPA) [24, 59] orbital charges allow one to split the interaction. The charge in a given atom A is determined according to

$$P_{A,i} = \sum_{r \in \{A\}} \left(\frac{V_{ri}^2}{\sum_j V_{rj}^2} \right) \tilde{D}_{rr}, \quad (3.2)$$

where the matrix \mathbf{V} is the transformation matrix from Natural Atomic Orbitals (NAOs) to local occupied orbitals and \tilde{D}_{rr} is the occupation number for a given NAO r [24]. The sum only runs over occupied orbitals indices. The interaction energy obtained from each covalent bonding orbital at the correlated level can then be split. The fraction belonging to the Au atom is determined as

$$\Delta E_i^{\text{Au}} = \frac{P_{\text{Au},i}}{P_{\text{Au},i} + \sum_{B \in \{\text{L}\}} P_{B,i}} \Delta E_i. \quad (3.3)$$

The remaining fraction is attributed to the other fragments $\{\text{L}\}$. The interaction energy does not have to be partitioned in the case where an orbital has a strictly located domain within the fragment.

In the following we used density fitted local second order Moller-Plesset perturbation theory (DF-LMP2) [32] as well as local coupled cluster singles and doubles with perturbative non-interactive triples (DF-LCCSD(T0)) [60] calculations. The spin component scaled LMP2 (DF-SCS-LMP2) method [51] was also used allowing for a straightforward correction to the LMP2 energies. This was carried out at the pair energy level which we

used for our decomposition. Since we used density fitting for all methods, the prefix 'DF-' is left out. In all local correlation calculations, Pipek-Mezey localization scheme [22] for occupied orbitals was used. The orbital domains were assigned according to the NPA criteria [24], with $T_{\text{NPA}}=0.03$. All intermolecular pairs were treated as strong pairs [36] thus they were fully included in the coupled cluster treatment in the local coupled cluster calculations. All calculations were carried out with a development version of Molpro 2010.2 [61].

The Dunning aug-cc-pVnZ ($n = \text{D, T}$) orbital basis set [62, 63] was used for all atoms except gold where the aug-cc-pVnZ-PP basis set was used in combination with the effective core potential ECP60MDF [64, 65]. These basis sets will be referred to as AVDZ and AVTZ for $n = \text{D}$ and $n = \text{T}$, respectively. An alternative basis set without diffuse functions was also applied (cc-pVTZ and cc-pVTZ-PP for Au) and will be referred to as VTZ. The density fitting approximation was used throughout with corresponding default basis sets (for the aug-cc-pVTZ and aug-cc-pVTZ-PP basis) [66, 67] except for SCF calculations with Au where the JKFIT def2-QZVPP [68] basis set was used. In cases where the VTZ basis set was used the corresponding non-augmented JKFIT and MP2FIT basis sets were applied and for gold def2-QZVPP/JKFIT and cc-pVTZ-PP/MP2FIT.

In the case of explicitly correlated calculations, the cc-pVTZ-F12 [69, 70] explicitly correlated basis set was used for all atoms except gold where the aug-cc-pVTZ-PP basis set was used as proposed by Peterson *et al.* [71]. In the case of explicitly correlated methods, a third auxiliary basis set was used for the resolution-of-identity (RI) approximation. The RI approximation was used to reduce the computationally expensive three- and four-electron integrals arising from the explicitly correlated terms to products of at most two-electron integrals. The basis set for the RI approximation was aug-cc-pVTZ-PP/OptRI [71] for gold and cc-pVTZ-F12/OptRI [69] for all other atoms. The other two auxiliary basis sets used were the same as in the AVTZ basis set. This basis set will be referred to as VTZ-F12.

DFT calculations were also carried out with the BP86 [45, 46], B3LYP [49] and PBE [47] functionals. In the DFT energy calculations the def2-TZVPP basis set was used in combination with the ECP60MDF pseudopotential for Au [72, 73].

The monomer geometries were optimized at the BP86/def2-TZVP level and then frozen monomers were placed in the arrangement of interest, except if stated otherwise. To ensure a smooth potential profile of the interaction curves, the orbital domains

were computed at a $r(\text{Au-Au})$ distance of 100 Å and kept fixed in all calculation steps.

In the case of Cl-Au-(Me₂bimy), the calculations were carried out as part of the master thesis of Axel Wuttke [74].

3.1.2 Results and Discussion

Cl-Au-PH₃ dimer

We start with the Cl-Au-PH₃ dimer which is a well known benchmark system for the investigation of aurophilic interactions. The monomer internal geometries were taken from Reference [75] and then placed in a non-planar orientation with a 90° dihedral angle to minimize the leading dipole-dipole term between the two monomers. This made it possible to focus on the aurophilic interaction itself. The energy profile of the dimer was computed by calculating the energy with fixed monomer geometries varying the distance between the gold atoms.

Most of the previous theoretical work was carried out using MP2 or LMP2 methods, except in the case of O’Grady *et al.* [76] where CCSD(T) and quadratic configuration interaction with singles and doubles (QCISD) methods were also applied. Here, we want to make use of LCCSD(T0) and the AVTZ basis set. If we compare this method with LMP2 and SCS-LMP2 using the same basis set (Figure 3.4 (left)), one can see that LMP2 strongly overestimates the results. On the other hand, the SCS-LMP2 curve is almost indistinguishable from the LCCSD(T0) curve.

It is well known that a significant basis set dependence is present in gold complexes [77]. Therefore, we wanted to evaluate the basis set dependence of LCCSD(T0) for this system. Two basis sets were used, AVDZ and AVTZ. Since calculations with larger basis set were prohibitive we instead carried out LCCSD(T0)-F12 calculations with VTZ-F12 basis set. It was already shown before that explicitly correlated methods with triple-zeta basis set obtain results of quadruple-zeta quality or even above [78]. Therefore, the LCCSD(T0)-F12 method was used as a reference. All three potential curves are shown in Figure 3.4 (right) and one can see that when the basis set is decreased LCCSD(T0) underestimates the results. In the case of LCCSD(T0)/AVTZ results are slightly underestimated, but the found minimum distance (3.2 Å) agrees well with the F12 results. The difference in energy at the minimum is ~ 2 kJ/mol. Therefore, we can conclude that our results are almost converged with the basis set and that AVTZ is a good choice for these

calculations.

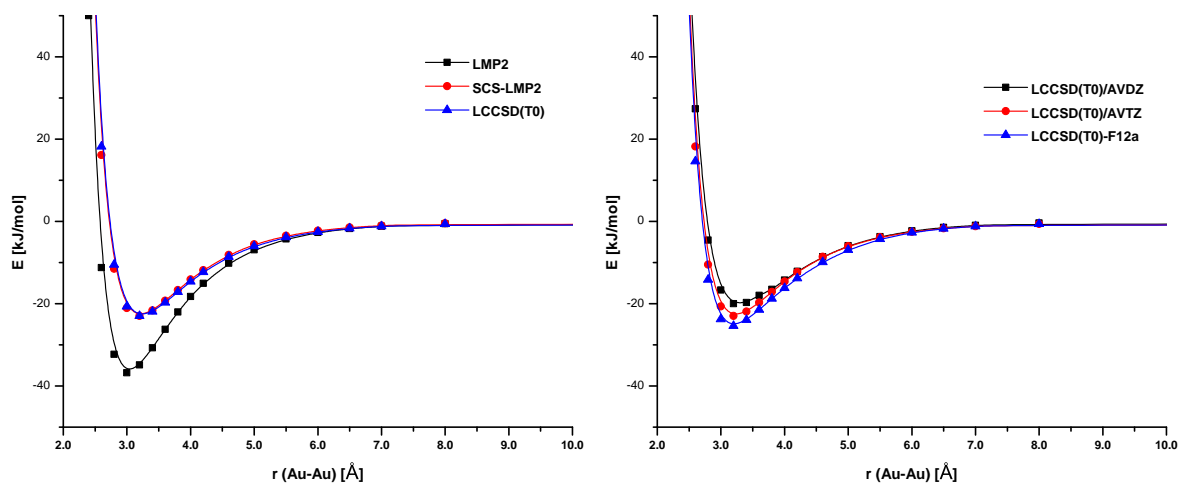


Figure 3.4: (left) Potential energy curves of the Cl-Au-(PH₃) dimer at different levels of theory with AVTZ basis sets. (right) Potential energy curves of the Cl-Au-(PH₃) dimer calculated with explicitly correlated CC with AVTZ-F12 basis set and the LCCSD(T0) results with AVTZ and AVDZ basis sets.

[ClAuPH₃]₂ potential energy curves obtained with several different methods are shown in Figure 3.6. All calculations were carried out with the AVTZ basis set. Since in the most previous theoretical studies the MP2 method was used and it was shown that LMP2 method overestimates the results, we wanted to evaluate the behaviour of MP2. Here, one can see that MP2 overestimates the results even more than LMP2. The reason for this is not only the lack of the higher order excitations but also the basis set superposition effect (BSSE). The LMP2 method is less prone to BSSE by construction. Excitations which are responsible for BSSE, the double cross-excitations from one monomer to the other monomer's virtual space, are excluded from local correlation methods (Figure 3.5). Therefore, CP-MP2 is found to be in a much better agreement to LMP2. The difference at the minimum is only 2 kJ/mol in comparison to 9 kJ/mol without counterpoise correction (CP). This indicates that the local approximation does not have almost any effect on the potential. The same behaviour is also observed in the work of Runeberg *et. al* [79]. In their study, they also carried out the CP-LMP2 calculation and obtained a curve which is quite close to the LMP2 curve.

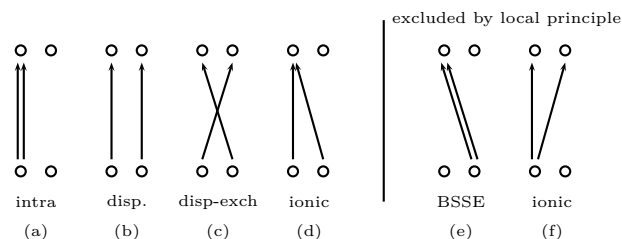


Figure 3.5: Schematic representation of the different double excitation classes of local correlation methods in the context of intermolecular interactions. The lower and upper circles represent different monomers in the ground and excited states, respectively. The arrows symbolize the corresponding excitations.

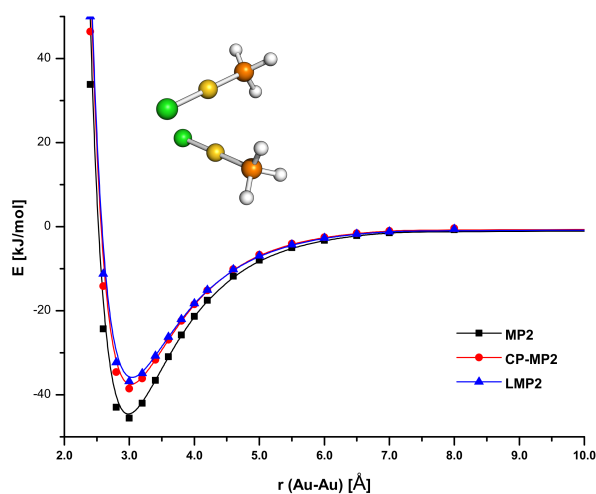


Figure 3.6: Potential energy curve of the Cl-Au-(PH₃) dimer at different levels of theory. All results have been computed with the AVTZ basis set.

Since aurophilicity is based on the Au \cdots Au interaction solely we can use the LMOMO method to focus this interaction. The LMP2 method fails to describe the aurophilic interaction since it overestimates dispersion forces. For this reason we treat the Au \cdots Au interaction at the LCCSD(T0) level and the rest at the LMP2 level. This should be enough to obtain the potential energy curve which would be in the close agreement with the potential energy curve obtained with LCCSD(T0). As one can see in Figure 3.7, despite the treatment of the gold atoms at the high level, the difference between LCCSD(T0) and LCCSD(T0):LMP2 results is still present. The difference between these two methods at the minimum is about 5 kJ/mol. Hence, we can conclude that these two monomers do not interact solely through the Au \cdots Au interaction. Despite the small system size, the ligands play an important role and the dispersion interactions between

gold atoms and the ligands or even between the ligands are not correctly described at the LMP2 level.

After partitioning the energy, Magnko *et al.* [75] and Runeberg *et al.* [79] arrived to the same conclusion, that this system does not interact only through the Au \cdots Au interaction. In the study by Runeberg *et al.*, the correlation energy was first partitioned according to the excitation classes shown in Figure 3.5. They observed that the dispersive and the ionic attractions near the equilibrium are of comparable size. Afterwards, dispersive and ionic attractions were partitioned even further and individual pair energies were obtained. From those pair energies it was seen that the largest contributions to the dispersion and ionic terms are from the pairs where at least one excitation was from an Au $5d$ orbital. Since the ligands were small, they could not fully assess the influence of the ligands.

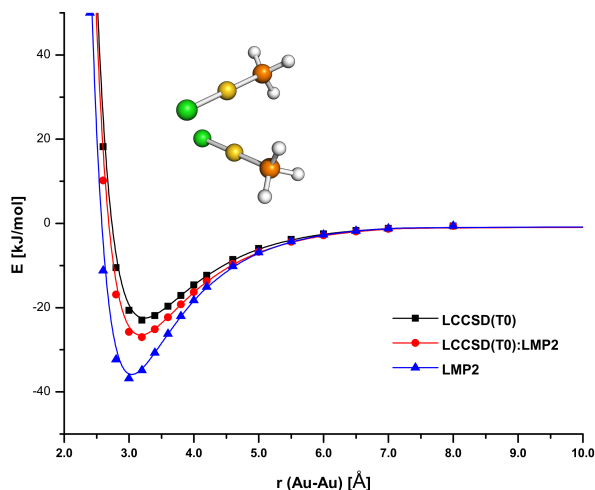


Figure 3.7: LCCSD(T0), LMP2 and LCCSD(T0):LMP2 results for the Cl-Au-(PH₃) dimer. In the latter calculation, only the Au orbitals are treated at the CC level. All results have been computed with the AVTZ basis.

To understand the LMOMO results better we analysed the orbital interactions according to the interaction groups already mentioned. Since we obtained a good agreement between the SCS-LMP2 and LCCSD(T0) results, we used the orbital contributions as computed at the SCS-LMP2 level to explore the role of different groups. The interaction groups for this dimer are $\Delta E(\text{PH}_3 \cdots \text{PH}_3)$, $\Delta E(\text{PH}_3 \cdots \text{Au})$, $\Delta E(\text{PH}_3 \cdots \text{Cl})$, $\Delta E(\text{Au} \cdots \text{Au})$, $\Delta E(\text{Au} \cdots \text{Cl})$ and $\Delta E(\text{Cl} \cdots \text{Cl})$. These results are shown in Figure 3.8. One can observe that the correlation energy contribution from the Au \cdots Au interaction

is the largest term around the minimum distance, but contributions from $\text{Cl}\cdots\text{Au}$ and $\text{L}\cdots\text{Au}$ are also important and summed together are as significant as the direct $\text{Au}\cdots\text{Au}$ contributions. Even for such a small ligand we obtained that these two monomers do not interact only through the $\text{Au}\cdots\text{Au}$ centres, but that the terms between gold and ligands are also significant.

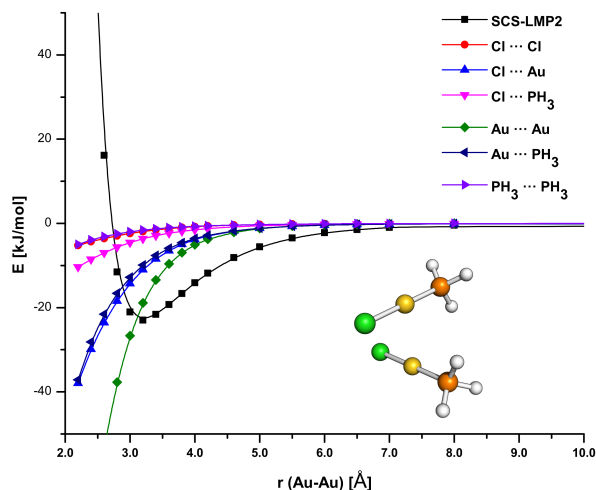


Figure 3.8: Correlation energy decomposition of the $\text{Cl-Au-(PH}_3\text{)}$ dimer potential curve (SCS-LMP2/AVTZ) together with the full result.

As previously noted by Pyykkö and O’Grady [76, 80], DFT methods cannot reliably describe van der Waals-like interactions, since the R^{-6} term which results from the dipole-dipole Hamiltonian is not explicitly included in DFT. On the other hand, the dispersion correction (D3) includes such terms through an empirical correction and improves the DFT results. Hence, we decided to compare DFT and DFT-D3 results to the LCCSD(T0)/AVTZ values. The B3LYP and PBE functionals were used. For the D3 correction Becke–Johnson damping [52] as well as non-local density dependent dispersion corrections (DFT-NL) were used as implemented in the Orca program package [81, 82]. The D3 corrections were computed with the `dftd3` program [51], using fitting coefficients for Au kindly provided by the authors [83]. The latter were obtained by TDDFT calculations on the AuH_2^+ system following the standard procedure of Grimme and co-workers [51].

The results can be found in Figure 3.9. As it was expected, bare DFT values strongly underestimate the interaction energy. An almost purely repulsive curve was obtained

when the B3LYP functional was used. This is naturally linked to the lack of dispersion forces. When D3 corrections were added, the DFT values come in very close agreement with the coupled cluster curve. D3 corrections also brought two different functionals in good agreement. The remaining discrepancies between DFT-D3 and the wave function values can be linked to the difference in the basis set convergence of the two methods as we have not included any corrections for basis set incompleteness.

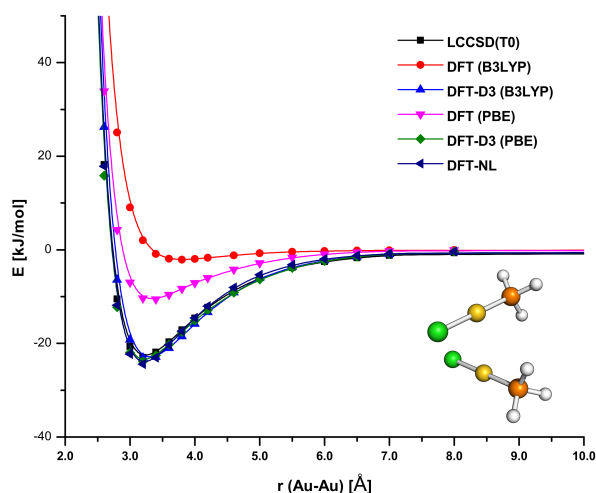


Figure 3.9: Potential energy curve of the Cl-Au-(PH₃) dimer at the LCCSD(T0)/AVTZ level and DFT results. All DFT calculations have been carried out with the def2-TZVP basis set.

Cl-Au-(C(PH₃)₂) and Cl-Au-(NHC-H) dimers

We now consider two different Au(I) complexes. On the example of the Cl-Au-(C(PH₃)₂) dimer we have checked the influence of the ligands on the LCCSD(T0):LMP2 energies as well as their contribution to the total correlation interaction. Then, also using the Cl-Au-(C(PH₃)₂) dimer, we have tested how the change of ligands influences the agreement between SCS-LMP2 and LCCSD(T0) results. For the energy decomposition we have again made use of SCS-LMP2. On the example of the Cl-Au-(NHC-H) dimer, we have investigated how the orientation of the ligands influence the interaction between two monomers as well as the auophilic interaction.

First, we have shortly compared the LCCSD(T0) and LCCSD(T0):LMP2 results for the Cl-Au-(C(PH₃)₂) dimer. From Figure 3.10 (left) it is noticeable that the difference between the full and the region calculation is even larger than in the previous case, suggesting that the ligands have a larger effect than before. This is what one would

expect, since these ligands are larger than in the previous case. On the other hand, the SCS-LMP2 results are still in good agreement with LCCSD(T0). It is also important to notice that in this dimer, the minimum is found at $r > 4 \text{ \AA}$ and the energy at the minimum is $\sim 12 \text{ kJ/mol}$. Therefore, this dimer is an example of a complex in the weak interaction regime.

In the decomposition profile in Figure 3.10 (right) one can see that the major correlation contribution is not from $\text{Au} \cdots \text{Au}$ terms, but rather from the $\text{C}(\text{PH}_3)_2$ ligand. It can also be seen that around a distance of 4 \AA contributions from all terms are relatively equal. At shorter distances the $\text{Au} \cdots \text{Au}$ interaction becomes more important, but still not the largest one. If we think about a simple distance criteria which suggests that systems with Au-Au distances shorter than 3.5 \AA are characterized as auropophilic this system would not be in such a group.

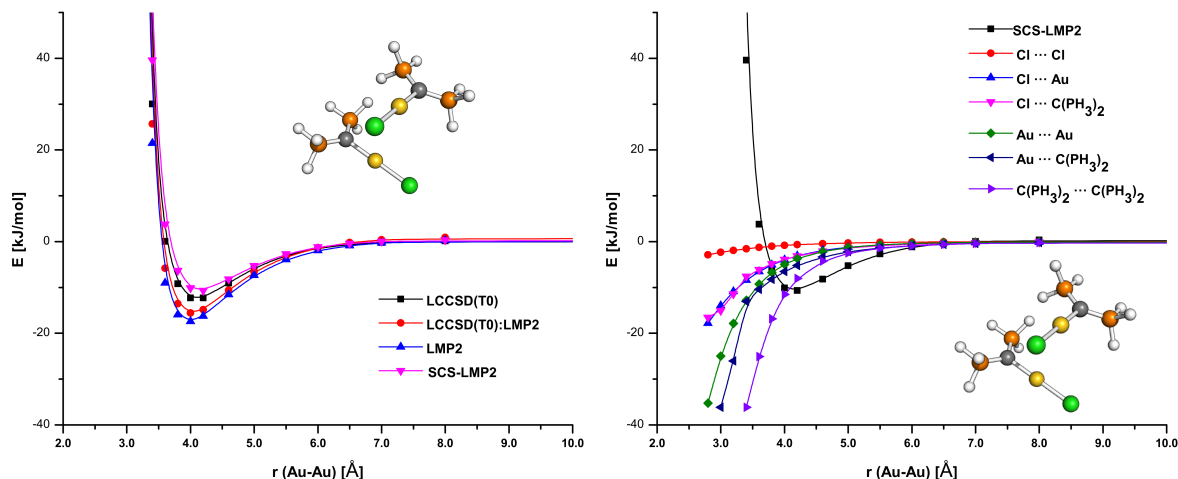


Figure 3.10: (left) LCCSD(T0), SCS-LMP2, LMP2 and LCCSD(T0):LMP2 results for the Cl-Au-(C(PH₃)₂) dimer. In the latter calculation, only the Au orbitals are treated at the CC level. (right) Correlation energy decomposition of the Cl-Au-(C(PH₃)₂) dimer potential curve (SCS-LMP2/AVTZ) together with the full result.

Furthermore, we have investigated systems with a N-heterocyclic carbene an imidazol-2-ylidene (NHC-H) ligand. Two different orientations of the system were considered, face-to-face (ff) and edge-to-edge (ee), which are shown in Figure 3.3. In both systems minima were observed at similar distances, 3.2 \AA for the ff dimer and 3.3 \AA for the ee dimer. On the other side, the well depth of the systems is considerably different. It is more than doubled in the ee dimer (-18.0 kJ/mol in the ff dimer and -42.4 kJ/mol in

the ee dimer). Similar results for these systems were obtained in previous theoretical studies [58]. The ff dimer shows an energy profile quite close to the Cl-Au-(PH₃) dimer (Figures 3.8 and 3.11 (left)). The minimum distances and total interaction energies are very similar. Also, the most significant contribution is linked to the Au···Au interaction, followed by Au···Cl and Au···L.

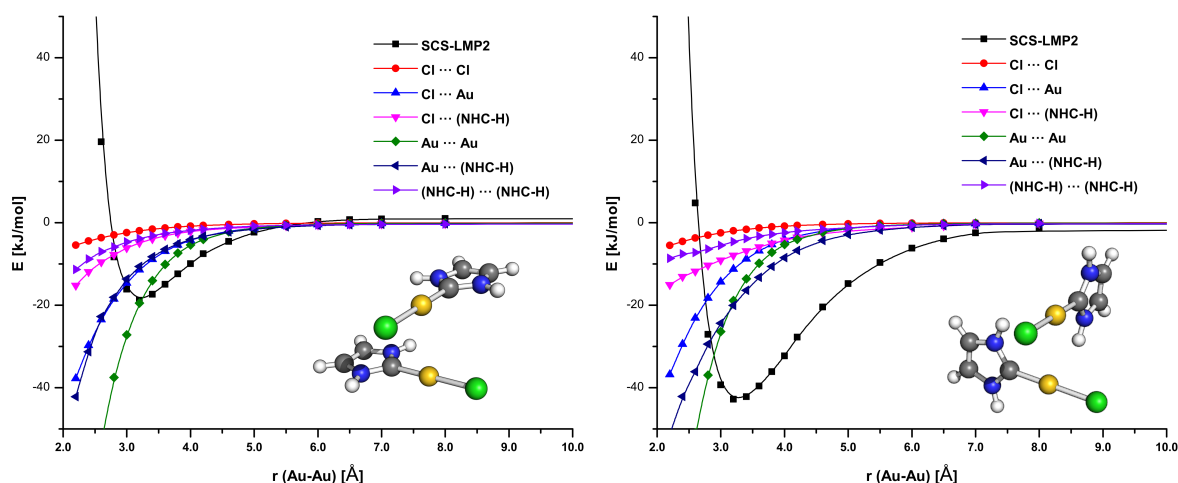


Figure 3.11: Correlation energy decomposition of the Cl-Au-(NHC-H) face-to-face (left) and edge-to-edge (right) dimers potential curves (SCS-LMP2/AVTZ) together with the full result.

In the case of the ee dimer (Figure 3.11 (right)), the energy is significantly stabilized in comparison with the ff dimer. Looking at the decomposition of the correlation energy, one can see that this stability is not due to the Au···Au interaction, since this remains mostly unchanged. The largest difference is in the Au-(NHC-H) contact energy. The latter is even slightly larger than the Au···Au contribution near the minimum. The interaction between the two monomers is increased since the N-H groups are now closer to the Au centres.

After observing the influence of the ligands on the interaction curves, we wanted to see if the ligands could impact the Au···Au interaction itself. For this purpose we plotted only the $\Delta E(\text{Au-Au})$ curves for all four dimers in Figure 3.12. One can see that different ligands do not influence the Au···Au interaction. In three systems, the same gold coordination was present (Cl-Au-C), but there are still reasonable differences between the studied ligands.

In previous work metallophilicity along the group 11 was investigated [76]. It was found that at the MP2 level of theory metallophilicity would increase down the group,

in agreement with other works using the same method. On the other hand, when metal-philicity was investigated using the QCISD or coupled cluster approaches the strongest interaction was obtained for the $[\text{Cl-Ag-(PH}_3)_2]$ system. Stronger argentophilic than aurophilic interaction is also observed experimentally by Ray *et al.* [84] and in the other complexes with N-heterocyclic carbenes [85–87]. To cast a new light over such findings, we have also computed the Cl-Ag-PH_3 and Cl-Cu-PH_3 dimers. In Figure 3.12 the metal-metal interactions also for silver and copper are shown. The curves demonstrate that the direct metal-metal interaction increases when going down the group. However, the differences are surprisingly large. As one can see in Figure 3.12 the $\text{Cu}\cdots\text{Cu}$ interaction becomes relevant only for distance smaller than 3 Å. Thus, in the case of Cu and Ag, the ligands have an even bigger influence on the total interaction energy.

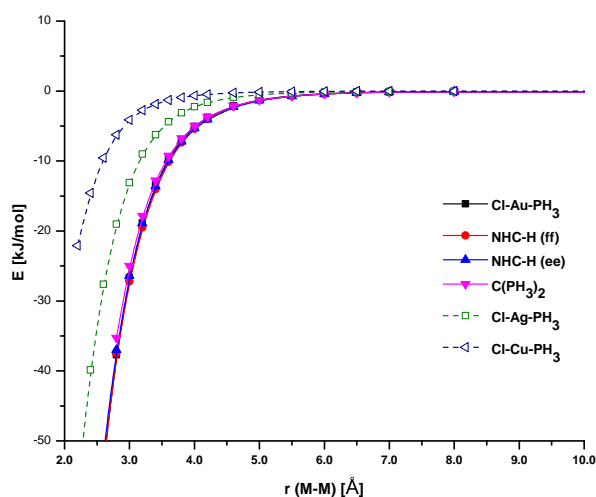


Figure 3.12: Au-Au SCS-LMP2 correlation energy decomposition for all complexes, including a comparison of the Ag-Ag and Cu-Cu interactions in Cl-M-PH_3 .

Our results are in agreement with the study of Magnko and coworkers [75]. They compared intermolecular, ionic and dispersive contributions of the Cl-M-PH_3 , where $M = \text{Cu, Ag or Au}$ and obtained that in all three cases the aurophilic interaction is the strongest one followed by argentophilic interaction and the weakest one cuprophilic. After they partitioned the correlation energy into orbital pair-energies, they observed that in practically all cases, the main part of the attraction involves pair correlation between one $M(d^{10})$ entity and non- $M(d^{10})$ localized orbitals of the partner monomer. But for the Cu they found that pair correlations not involving any of $M(d^{10})$ centres

become the most significant ones around the minimum. This again comes in support of our observations.

Cl-Au-(NHC-B) dimer

Observing that the Au \cdots Au interaction is stronger than Ag \cdots Ag interaction in a wide variety of complexes, an investigation of the Cl-Au-(NHC-B) dimer was warranted. The full Cl-Au-(NHC-B) system amounts to 86 atoms. To avoid changes in the conformation of the NHC-B ligand we decided to optimize the dimer and not monomer as in previous cases. The dimer was optimized at the BP86/def2-SVP level of theory [72, 88]. The starting structure for the optimization was the crystal structure of the system from Reference [84]. In the previous systems dimers were orientated to diminish dipole-dipole interaction. However, in the crystal other relative conformations are adopted. The C1-Au1-Cl1 angle amounts to 172.38° in the crystal and the C1-Au1-Au2 angle 96.8°. After the optimization these two angles are 175.17° and 96.4°, respectively. Therefore, dipole-dipole interactions do not amount to zero as in the previous systems. The obtained geometry is shown in Figure 3.13. To obtain the interaction energy curve, the Au-Au distance was varied. Because of the size of the system, a smaller basis set was used (VTZ).

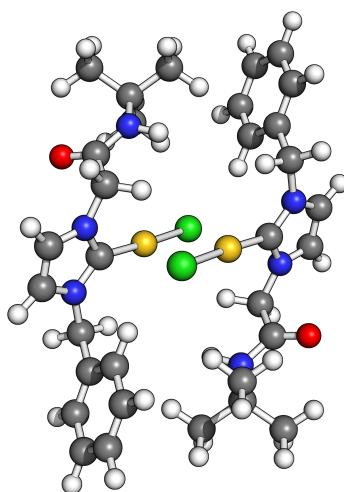


Figure 3.13: Structure of the Cl-Au-(NHC-B) dimer.

In Figure 3.14 one can see that already at the HF level the potential curve has a

minimum indicating that electrostatic interactions (dipole and higher moments) contribute to stabilize the structure. By including correlation effects, one is obtaining a much deeper minimum (SCS-LMP2 curve in Figure 3.14). This means that the most significant interaction terms arise due to the correlation effects. Also one can see that correlating the gold atoms at the LCCSD(T0) level does not improve the LMP2 results. Thus, we can conclude that Au \cdots Au interactions have an almost negligible effect.

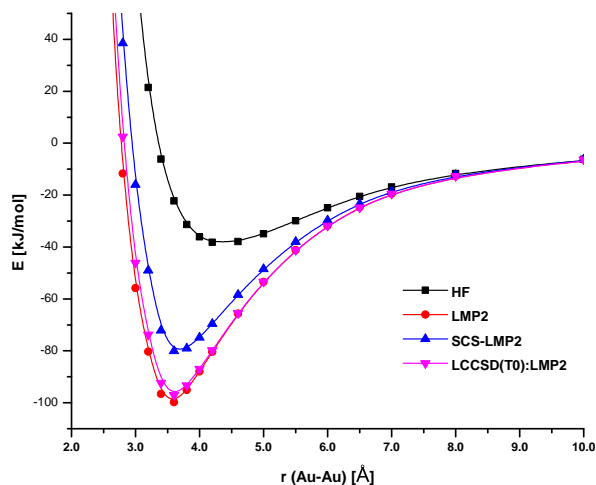


Figure 3.14: Potential curves calculated at different levels of theory using the VTZ basis set.

Observing the energy decomposition (Figure 3.15) one can see that the correlation contributions are dominated by the $\Delta E(\text{Au}\cdots\text{L})$ and $\Delta E(\text{L}\cdots\text{L})$ terms, where L is the NHC-B ligand. The Au \cdots Au interaction contributes only $\sim 10\%$ to the correlation energy. In such a case one would be inclined to consider this system only slightly auophilic. Even the Au \cdots L interaction outweighs the Au \cdots Au interaction.

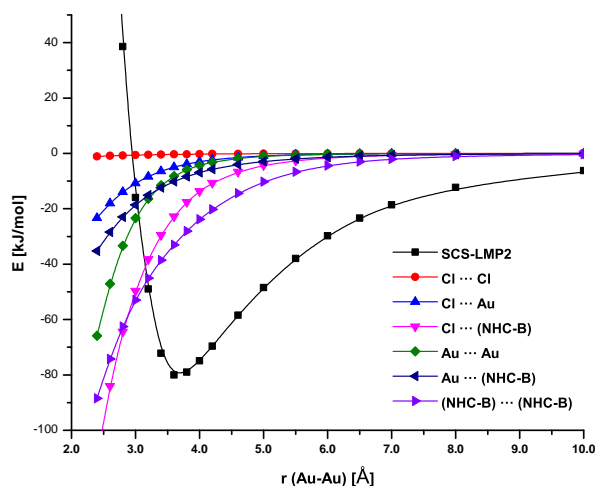


Figure 3.15: Energy decomposition analysis of the Cl-Au-(NHC-B) dimer potential curve (SCS-LMP2/VTZ), together with the full results.

In the case of the Ag complex the two monomers can come closer to each other than in the case of the gold complex. Therefore, the dispersion interaction between the ligands will be more important, so it is reasonable to obtain a stronger interaction between the complexes which contain Ag than the complexes which contain Au. This, however, does not mean that the $\text{Ag}\cdots\text{Ag}$ interaction is stronger than the $\text{Au}\cdots\text{Au}$. It is solely a ligand effect.

Cl-Au-(Me₂bimy) dimer

At the end we consider one more system which is experimentally well characterized, the Cl-Au-(Me₂bimy) dimer. The Cl-Au-(Me₂bimy) dimer is one of the units which build the crystal. This crystal was characterized by Wang *et al.* [89]. Despite the similarity with the Cl-Au-(NHC-H) dimer, this dimer is present in the crystal in the ff orientation. Therefore, one can assume that in this case the weak interactions between ligands can add up and become the determining factor in the dimer formation. Thus, we have compared two model dimers in ff and ee orientations and then we have also compared the ff dimer and the dimer in the same orientation as in the crystal.

In case of the Cl-Au-(Me₂bimy) dimer calculations, the monomers were optimized with BP86 using the def2-TZVP basis set for all atoms except Au where the ECP60MDF core potential was used in combination with the dhf-TZVP basis functions [90]. Using the frozen monomer geometries, different arrangements were built.

As one can see from Figure 3.16 in the case of the Cl-Au-Me₂bimy ff dimer a minimum is found at 3.6 Å in comparison with 6 Å for the ee orientation. Also, the ff orientation is more stable and the energy at the minimum amounts to -26.2 kJ/mol, opposite to the shallow minimum of -9.0 kJ/mol for the ee case.

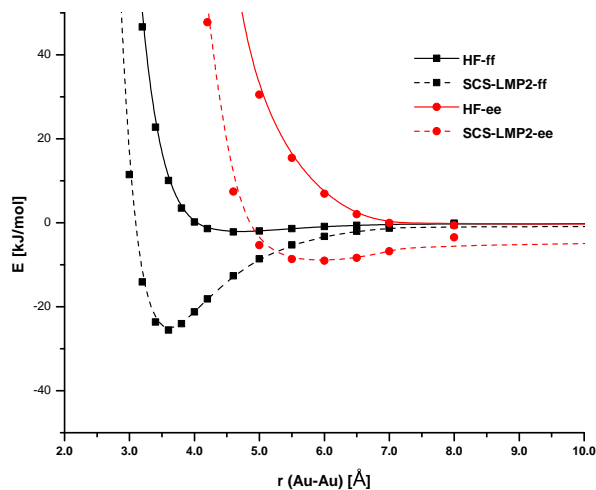


Figure 3.16: Potential energy curves of the Cl-Au-(Me₂bimy) dimer in face-to-face and edge-to-edge orientation at different levels of theory. All results have been computed with the AVTZ basis set.

Looking at the orientation of this dimer (Figure 3.17) one can see that the methyl groups have a high steric influence in the case of the ee orientation and can be responsible for the early energy minimum. If we now compare the results from Cl-Au-(NHC-H) and Cl-Au-(Me₂bimy) dimers, both in the ff orientation, one can see that in the case of the latter the well depth has been increased by about ~8 kJ/mol and the bonding distance lengthened by 0.4 Å. This can be also attributed to steric repulsion effects.

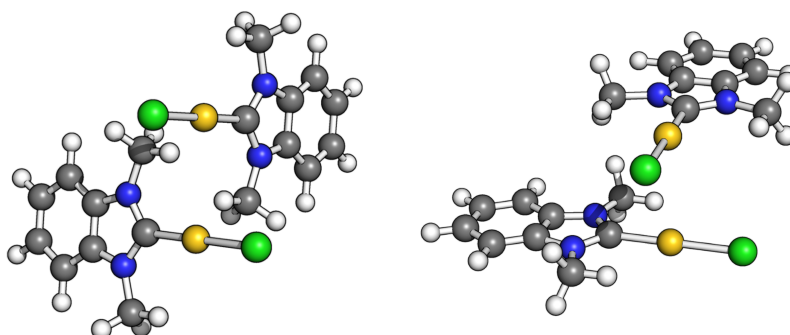


Figure 3.17: Cl-Au-(Me₂bimy) dimer in the face-to-face (left) and edge-to-edge (right) orientation.

Now we take a closer look at the energy decomposition for both orientations shown in Figure 3.18. In both cases the dominant component at the minimum was the L···L component. In the ff conformation this is due to the fact that methyl groups are above each other and can interact through van der Waals interactions. The largest difference is in the Au···L interaction. In the ee orientation the methyl groups are closer to the gold atom, so this interaction is much stronger. Stronger interactions in ee orientation were also obtained for L···L and Cl···L components. Due to the orbital overlap at smaller distances, the values for an ee orientation are unreliable for distances below 4 Å.

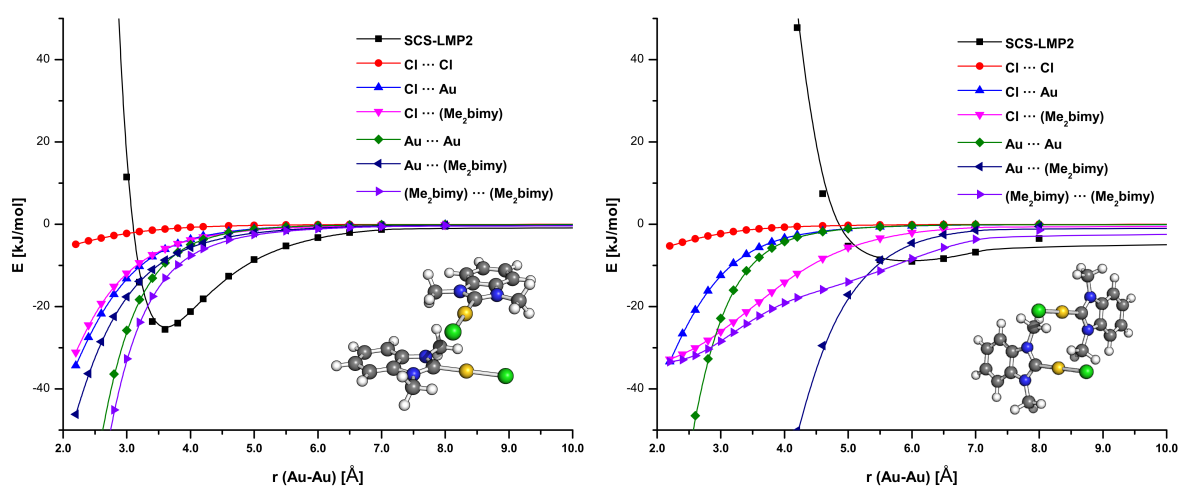


Figure 3.18: Correlation energy decomposition of the Cl-Au-(Me₂bimy) face-to-face (left) and edge-to-edge (right) dimers potential curves (SCS-LMP2/AVTZ) together with the full result.

We also compared the results obtained for the ff model structure (model) and the dimer in the same orientation as in the crystal (crystal). The difference between these

two structures is that in the model system the dihedral angle C1-Au1-Au2-C2 is 90° and in the crystal structure 94.7° . The other difference is that Au1-Au2-C2 are forming an angle of 90° in the model (84.4° in the crystal).

Potential energy curves for both systems were calculated at the HF and SCS-LMP2 levels of theory and they are shown in Figure 3.19 (left). As one can see in the case of the crystal dimer, an attractive curve is obtained already at the HF level. This means that a dipole-dipole interaction is present in the crystal. It should also be noted that the difference between the HF results for the different systems relative to SCS-LMP2 is mostly kept. Therefore, one can conclude that the increase in the well depth for the crystal system is only due to dipole-dipole interactions and other higher order electrostatic contributions captured by HF.

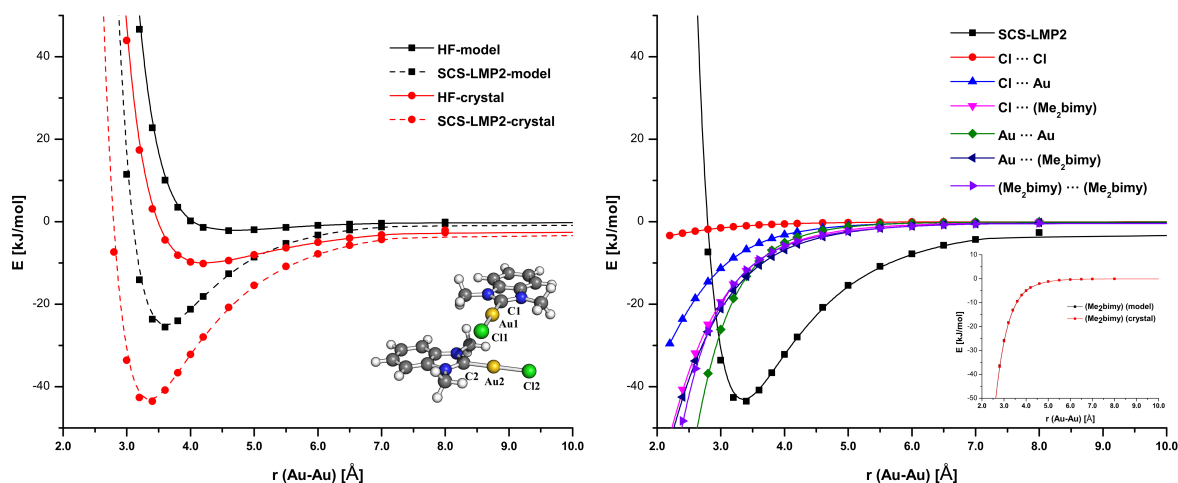


Figure 3.19: Potential energy curves of the Cl-Au-(Me₂bimy) dimer in model system and crystal at different levels of theory. All results have been computed with the AVTZ basis set (left). Correlation energy decomposition of the Cl-Au-(Me₂bimy) crystal dimer potential curve (SCS-LMP2/AVTZ) together with the full result (right).

The energy decomposition provides further insights when comparing the model to the crystal dimer. One can see that Cl... (Me₂bimy), Au... (Me₂bimy) and (Me₂bimy)... (Me₂bimy) interactions contribute almost equally to the correlation energy in the case of the crystal dimer. The same does not happen in the case of the model dimer. The largest change is obtained for the (Me₂bimy)... (Me₂bimy) interaction, but Cl... (Me₂bimy) and Au... (Me₂bimy) interactions are also noticeably altered. This is due to the shorter distance between the ligands and also between gold atoms and

the ligands. On the other hand, the Au \cdots Au interaction stayed unchanged as it can be seen in the inset in Figure 3.18 (right).

3.1.3 Final Remarks

In this chapter we have investigated the auophilic interaction. The first conclusions about the character of this interaction were drawn from the comparison of LCCSD(T0) and LCCSD(T0):LMP2 results. Since the difference between these two methods was significant, we concluded that the Au \cdots Au interaction was not as dominant as previously thought. The ligands have a significant contribution to the interaction.

To obtain correct results using the LMOMO scheme, one can increase the region which is correlated at the LCCSD(T0) level. In our studied systems, that would not be so much of interest, since then just a few orbitals would stay treated at the LMP2 level, except for the last two systems. The other possibility is that one uses the SCS-LMP2 method in the LMOMO calculations instead of LMP2. This was not carried out, since our goal was to understand the auophilic interaction rather than to obtain a good agreement between the results from LMOMO and the full calculation.

Given the good agreement between SCS-LMP2 and LCCSD(T0), potential energy curves were calculated with the former method. The use of this method enabled us to decompose energies into the interaction groups. In this way we could see how much each group contributes to the electronic correlation energy. We saw that already in the case of the smallest ligands, Au \cdots L and L \cdots L interactions summed together contribute almost as much as the Au \cdots Au interaction. When ligands are increased in size their influence on the correlation energy becomes even larger. In the case of the experimentally identified complexes, the interaction between ligands is even the dominant one. In these systems the energy is also stabilized by electrostatic interactions, so Au \cdots Au interactions have an even smaller influence, on the total interaction energy.

Since we had the possibility to separate the Au \cdots Au interaction from all other contributions, we also investigated how the different ligands, orientation of the ligands and orientation of the monomers influence the Au \cdots Au interaction. We found that the Au \cdots Au interaction itself is independent of all these factors. Auophilicity has the same strength in systems with ee and ff orientation meaning that the dispersion interaction between the ligands cannot influence the Au \cdots Au interaction. A slight reorientation of the system as in the case of the Cl-Au-(Me₂bimy) crystal dimer also did not influence the

Au...Au interaction despite the presence of dipole-dipole and higher order electrostatic interactions. Therefore, we can conclude that interaction between the closed d-shells in gold atoms is unchangeable. The difference in the interaction observed in these systems is due to ligands contacts and the different orientation of the ligands and/or monomers.

CHAPTER 4

Molybdenum Enzymes

Molybdenum is one of the most common metals in biological systems. It plays a prominent role in naturally bio-catalysts as found in several molybdoenzymes in nature. The coordination chemistry of molybdenum is diverse occurring in several different oxidation states. Depending on the latter it will show affinity for both hard oxo donor and soft sulfido donor ligands. Coordination numbers ranging from 4 to 8 have been identified.

A common feature of molybdenum enzymes is the molybdopterin ligand. It should be noted that despite the name, molybdopterin does not contain molybdenum. Molybdopterin consists of a pterin moiety, a dithiolene group as a side chain and a phosphoester group (Figure 4.1). Molybdopterin coordinates to Mo through a dithiolene moiety. The entire ligand with molybdenum is called the molybdopterin cofactor (Moco).

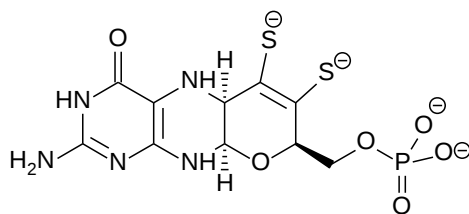


Figure 4.1: The structure of the pyranopterin cofactor.

Molybdenum enzymes are classified into two main groups. The first group are nitrogenases, which convert N_2 into NH_3 through a $Fe_7S_9C Mo$ cofactor. The second group contains the mononuclear molybdenum enzymes, which perform oxygen-transfer and hydroxylation reactions. Hille [91] has suggested that this second group can be divided into three enzyme families, depending on the ligands of the central Mo atom. These families are named after a typical example, viz. the xanthine oxidase (XO), dimethyl sulfoxide reductase (DMSOR) and sulfite oxidase (SO) families (Figure 4.2).

In the case of the XO family, the molybdenum ion in its oxidized state is coordinated to the two thiolate groups (Mo-S) from molybdopterin, one terminal oxygen ligand (Mo=O), one terminal sulfido ligand (Mo=S) and one hydroxy ligand (Mo-OH). These enzymes are responsible for the catalysis of oxidative hydroxylation reactions with a variety of aldehyde and purine substrates.

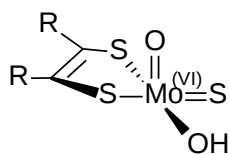
The DMSOR family is largest and most diverse one. It is responsible for the catalysis of the oxygen-atom transfer (OAT) from dimethyl sulfoxide (DMSO) to the Mo^{IV} active

site, yielding dimethyl sulfide (DMS) and Mo^{VI} . In this family, molybdenum is coordinated to two molybdopterin molecules and one serine residue, which in some enzymes is replaced by cysteine or selenocysteine. In the oxidized state, Mo also binds to an additional oxo group.

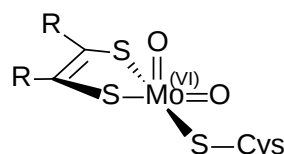
The last is the SO family and it catalyzes the oxidation of sulfite to sulfate. In this case, the Mo center is coordinated to one molybdopterin molecule through the thiolate ligands, one cysteine residue and two oxo ligands (in the oxidized state).

All these enzymes can be found in a variety of species from bacteria to plants and animals, including humans. The value of these enzymes lies not only in the importance to their host organisms for metabolism and energy generation, but also in their major impact in the nitrogen, sulfur, carbon and arsenic cycles.

xanthine oxidase family



sulfite oxidase family



DMSO reductase family

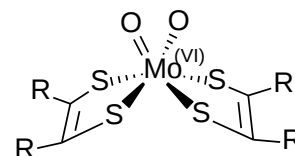


Figure 4.2: Active site structures for the three families of mononuclear molybdenum enzymes in their oxidized states. The structures shown are, from left to right, for xanthine oxidase, sulfite oxidase, and DMSO reductase.

In the following chapter we will investigate two of these families, SO and DMSOR. In the first part we will investigate possible mechanisms for sulfur oxidation in the scope of the SO family. Further, we will evaluate the accuracy of the LMOMO method in the case of sulfur oxidation. This method will be used to calculate the activation barrier for the reduction of DMSO.

4.1 Reaction pathways for SO

The oxidation of sulfite to sulfate [92] is the terminal step in the biological sulfur cycle of a variety of organisms. The reaction is catalyzed by sulfite oxidase [93], which contains a molybdenum cofactor with a Mo^{VI} ion in its oxidized state. Two oxo ligands and a cysteine residue from the protein are connected to the Mo ion. The molybdenum cofactor

is in a square-planar geometry with one oxo ligand in the axial position. Approximately 30 Å away from the Mo ion, a heme group is located, which is also important for the sulfur cycle. At the beginning of the catalytic cycle, the Mo ion is in the oxidized +VI state, and the iron in the heme group is in the +III state. Then sulfite binds to the Mo ion, being oxidized to sulfate while Mo is reduced to the +IV state. In the next step of the catalytic cycle, the Mo ion binds water. This is followed by two coupled one-electron/proton-transfer steps, which bring the Mo ion to the active Mo^{VI}=O form of the cofactor. The electrons are transferred via reduction of the heme group. The heme group is reoxidized by cytochrome c. The catalytic cycle is shown in Figure 4.3.

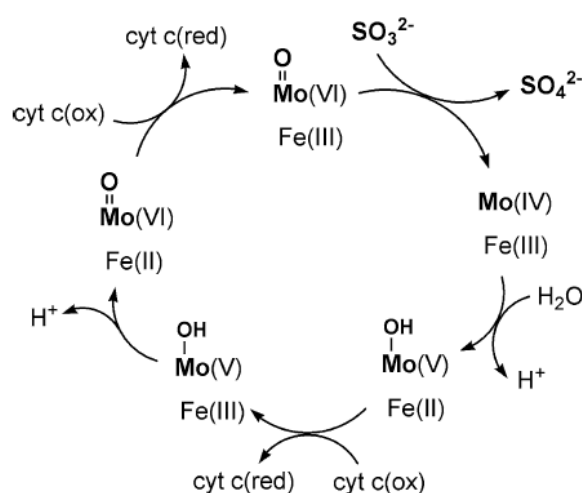


Figure 4.3: The general catalytic cycle for sulfite oxidase, Mo refers to the molybdenum cofactor and Fe refers to heme. Reproduced from Reference [94] with permission of copyright holders.

We will now focus on the first step of the catalytic cycle for sulfite oxidase – the oxidation of SO_3^{2-} to SO_4^{2-} . A mechanism consistent with Michaelis-Menten kinetics have been established [95–99]. Three different reaction mechanisms for this step have been proposed in previous studies. The first is a mechanism where the lone pair of the sulfur atom of SO_3^{2-} attacks the equatorial oxygen of Mo (S→OMo mechanism). In the second mechanism, the oxygen of SO_3^{2-} first attacks Mo, forming Mo-O bond and then the sulfur from SO_3^{2-} builds a bond with the equatorial O of Moco (O→Mo mechanism). In the last mechanism the sulfur from SO_3^{2-} first forms a bond with Mo and then with the equatorial O (S→Mo mechanism). All three mechanisms are presented in Figure 4.4.

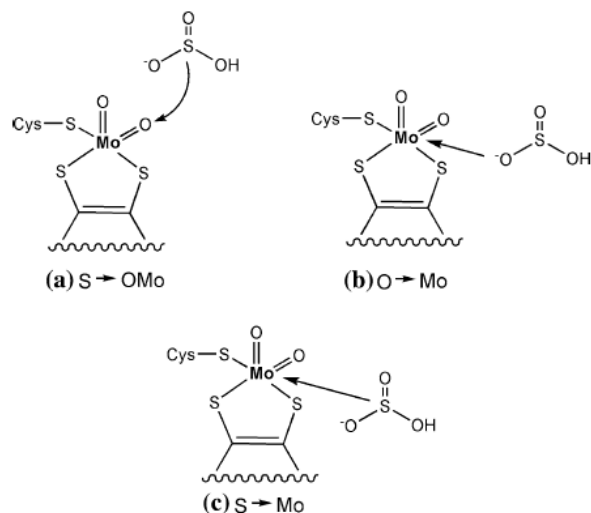


Figure 4.4: The three considered reaction mechanisms: (a) S→OMo mechanism, (b) O→Mo mechanism and (c) S→Mo mechanism. Reproduced from Reference [94] with permission of copyright holders.

The S→OMo mechanism was first proposed by Hille and coworkers [100, 101]. The latter was based on the observation that anions are able to bind to the molybdenum center in its resting state. It was therefore suggested that catalysis is initiated by the direct binding of the substrate to the molybdenum center through one of its oxo groups. This led to the conclusion that the lone pair of substrate is essential for reactivity. This mechanism was later theoretically investigated by different theoretical groups [102–104]. In Figure 4.5 is shown as proposed by Marin-Hernandez and Ziegler [102].

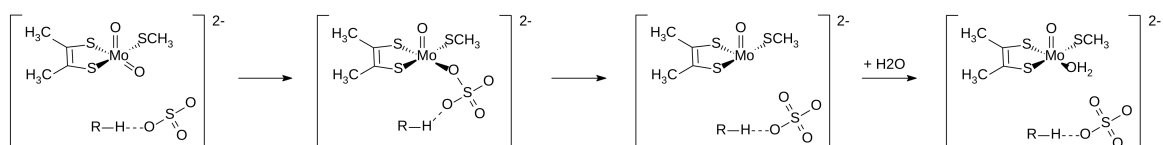


Figure 4.5: S→OMo mechanism as proposed in reference [102].

The O→Mo mechanism was proposed by Sarkar and coworkers [104]. In this theoretical study two different model systems were used. The first model system ($[\text{MoO}_2(\text{S}_2\text{C}_2(\text{CN})_2)_2]^{2-}$) was used to represent a model complex mimicking the active-site function of SO. The second one ($[\text{MoO}_2(\text{S}_2\text{C}_2\text{Me}_2)\text{SMe}]^-$) was a model of the active-

site of native SO. They proposed the formation of a Michaelis complex through oxoanionic binding of HSO_3^- at the Mo center (Figure 4.6). Experimental studies of Rajagopalan [105] showed that the active site of the native enzyme can adopt both five and six coordinated geometries, suggesting the possibility of oxoanionic binding of HSO_3^- . In this study the $\text{S} \rightarrow \text{OMo}$ mechanism was also investigated, but since no Michaelis complex was obtained, it was dismissed. The $\text{S} \rightarrow \text{Mo}$ mechanism was not considered in their investigation.

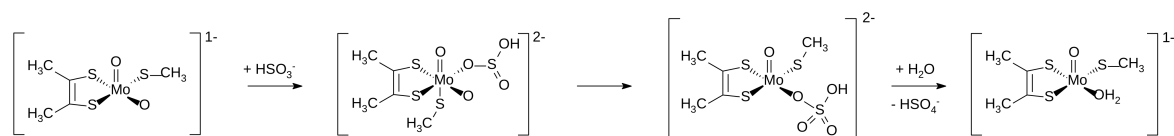


Figure 4.6: O \rightarrow Mo mechanism as proposed in reference [104].

The $\text{S} \rightarrow \text{Mo}$ mechanism was investigated by Thapper *et al.* [103] using DFT calculations. The system was modelled by $[\text{MoO}_2(\text{MNT})_2]^{2-}$ (MNT = maleonitrile-1,2-dithiolate). The first step of the $\text{S} \rightarrow \text{Mo}$ mechanism is shown in Figure 4.7. In the same study the $\text{S} \rightarrow \text{OMo}$ mechanism was also investigated. However, the first transition state for the $\text{S} \rightarrow \text{Mo}$ mechanism was found to be 50 kJ/mol higher than the rate-limiting step in the $\text{S} \rightarrow \text{OMo}$ case. Therefore, they decided that further calculations were unnecessary. In this study the O \rightarrow Mo mechanism was not considered.

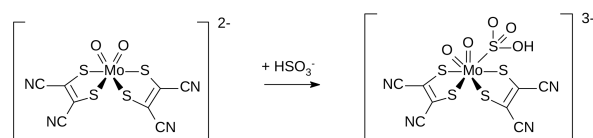


Figure 4.7: S \rightarrow Mo mechanism as proposed in reference [103].

As the previous summary shows, all three mechanisms have been investigated in different studies, but never on the same footing. This resulted in different suggestions for the most likely reaction. To solve this dilemma, we investigated all three mechanisms using the same methods, in an attempt to find the balanced description of the different pathways.

4.1.1 Computational details

In our study, the active site of sulfite oxidase was modelled by a Mo ion coordinated to one or two oxo ligands, and three sulfur atoms. One of the sulfur atoms belongs to the cysteine ligand which was modeled by SCH_3^- and the other two sulfur atoms come from the molybdopterin which was modelled by 1,2-dimethyldithiolene (DMDT).

In order to calculate the final electronic energies, we applied the local methods DF-LMP2 [106], DF-SCS-LMP2 [5] and DF-LCCSD(T0) [36]. Pipek-Mezey localized orbitals [22] were used. The orbital domains were determined according to a natural population analysis occupation with a threshold of $T_{NPA} = 0.03$ [24]. Mixed distance and connectivity criteria was used for the classification of orbital pairs [107]. The values which were used for the distance criterion were $R_c=3$, $R_w=5$, $R_d=8$ and $R_{vd}=15$. In the case of the connectivity criterion $I_c=2$, $I_w=3$, $I_d=5$ and $I_{vd}=8$ were used. All local correlation methods were carried out with a development version of Molpro 2012.2 [61].

The aug-cc-pVTZ orbital basis [63] set was used for all atoms except for H (cc-pVTZ [62]) and Mo (aug-cc-pVTZ-PP with the Stuttgart-Dresden ECP28MDF pseudopotential [108]). We will refer to this basis set as AVTZ. Also, we used a quadruple zeta basis set where all atoms were described with aug-cc-pVQZ except H(cc-pVQZ) and Mo (aug-cc-pVQZ-PP and corresponding pseudopotential) (AVQZ). DF approximations were used throughout for both the Hartree-Fock as well as the correlation part. The auxiliary basis sets were aug-cc-pVnZ/JKFIT [66] and aug-cc-pVnZ/MP2FIT [67] for all elements except for Mo where def2-nZVPP/JKFIT [68] and def2-nZVPP/MP2FIT [109] basis sets were applied ($n = T, Q$).

The geometries were optimized at the DFT level with the TPSS [110] functional and def2-SV(P) basis set [72]. For Mo a relativistic 28-electron core potential was used. To speed up calculations the RI approximation was used with the corresponding def2-SV(P) [111, 112] auxiliary basis set [94].

By comparing LMP2 and MP2 relative energies for the same basis set, one can estimate the effect of the domain approximation

$$\Delta E_{\text{domain}} = E_{\text{LMP2/XZ}} - E_{\text{MP2/XZ}}. \quad (4.1)$$

Therefore, MP2 and LMP2 calculations were carried out using the AVTZ and AVQZ basis sets. ΔE_{domain} is shown in Figure 4.8 for the $\text{S} \rightarrow \text{OMo}$ mechanism. One can see

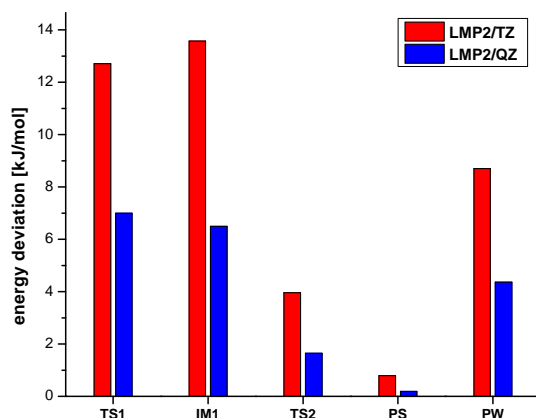


Figure 4.8: Difference between LMP2 and MP2 relative energies of the stationary points of the S→OMo mechanism for the TZ and QZ basis sets.

that the difference between these two methods using the AVTZ basis set reaches up to 13 kJ/mol. Using a larger basis set this difference is reduced to less than 2 kJ/mol. This trend was also observed in a previous study [113] with a molybdenum complex. The large difference for AVTZ is not only due to the domain approximation, but also due to BSSE. BSSE is significant for MP2/AVTZ, but it is reduced by increasing the number of functions. On the other hand, in LMP2 BSSE is significantly reduced by construction [33, 114].

Also to account for the basis-set incompleteness effects, the DF-LCCSD(T0)/AVTZ energies were corrected by estimating the canonical MP2 complete basis-set limit (CBS). For this, we used a n^{-3} extrapolation of the correlation energy from the two points (MP2/CBS[3:4]) [115]. Correlation energies were calculated at the DF-MP2/AVTZ and DF-MP2/AVQZ levels of theory. This extrapolated correlation energy was then added to the HF/AVQZ reference energy. The difference in the relative energies between MP2/CBS[3:4] and MP2, as well as LMP2, is shown in Figure 4.9 for the S→OMo mechanism. A significant difference was observed for the AVTZ basis set. Increasing the basis set to AVQZ this difference was reduced. Despite the reduction, this effect is still important and should be corrected for.

Assuming that this effect is similar at the MP2 and CCSD(T) levels of theory one can add a correction to the LCCSD(T0) values in the form:

$$E_{\text{EC}} = E(\text{MP2/CBS}[3:4]) - E(\text{LMP2/AVTZ}) \quad (4.2)$$

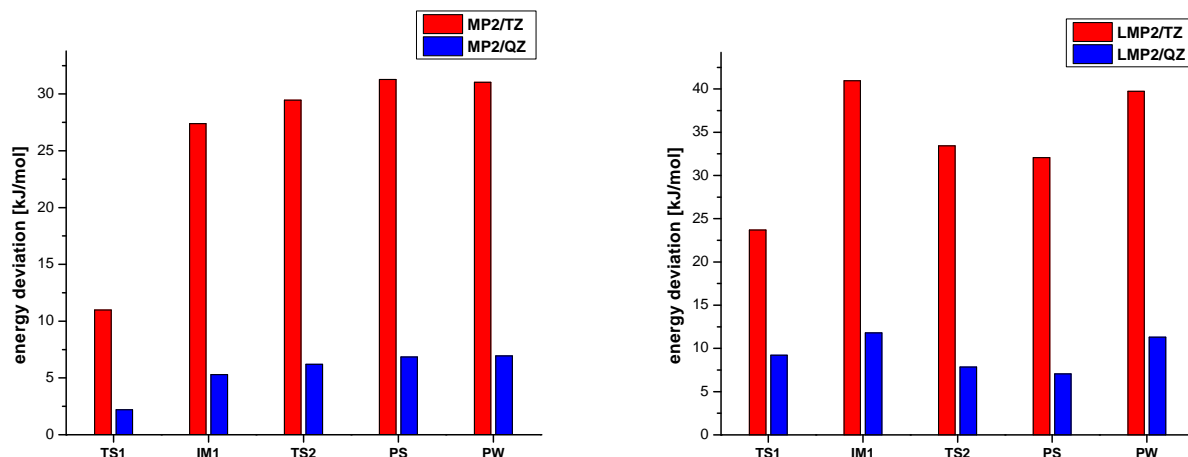


Figure 4.9: Differences from the MP2/CBS[3:4] relative energies of the stationary points of the S→OMo mechanism: MP2 (left) and LMP2 (right).

resulting in:

$$E(\text{CC/CBS}) = E(\text{LCCSD(T0)/AVTZ}) - E(\text{LMP2/AVTZ}) + E(\text{MP2/CBS[3:4]}) \quad (4.3)$$

as a final expression for the gas phase electronic energy. The total free energy was obtained as:

$$\Delta G_{tot} = \Delta E(\text{CC/CBS}) + \Delta G(\text{PCM}) + \Delta G(\text{therm}) + \Delta G(\text{solv}, \epsilon), \quad (4.4)$$

where $\Delta G(\text{PCM})$ is the sum of the Polarizable Continuum Model (PCM) cavitation, dispersion and repulsion energies, $\Delta G(\text{therm})$ is the thermal correction to the Gibbs free energy and $\Delta G(\text{solv}, \epsilon)$ is the solvation energy (the energy difference between the Conductor-like Screening Model (COSMO) calculations with dielectric constant of ϵ and 1). Four different energies will be discussed, starting with the ΔE which is the vacuum electronic energy without any thermal corrections (including only the first two terms from Eq. 4.4), then ΔG_1 which is the free energy without any solvent corrections and finally the two free energies with continuum solvent corrections. The dielectric constants used were 4 and 80, leading to ΔG_4 and ΔG_{80} , respectively [94].

4.1.2 Results and Discussion

S→OMo mechanism

The first investigated pathway is the S→OMo mechanism. In this case the reaction starts with a nucleophilic attack from the sulfur in HSO_3^- or SO_3^{2-} at the oxo ligand of molybdenum. Two oxo ligands are present, but it was previously shown that only the equatorial oxo ligand was reactive in this mechanism [116]. All results have been computed for this specific conformation.

The mechanism starts from the reactant state (RS) with the oxidized active site without any substrate. After sulfite attacks RS the first intermediate (IM1) is formed, via a transition state (TS1). In this IM1 state sulfate is coordinated to Mo(IV), so the oxygen abstraction is already under way. After this, sulfate dissociates from the molybdenum complex via a second transition state, TS2, which brings us to the product state, PS, which is four-coordinated. In agreement with crystal structures of the product, we investigated one more state where a water ligand is bound to the Mo complex instead of the product (PW). This allows us to keep the Mo complex five-coordinated. All structures are depicted in Figure 4.10.

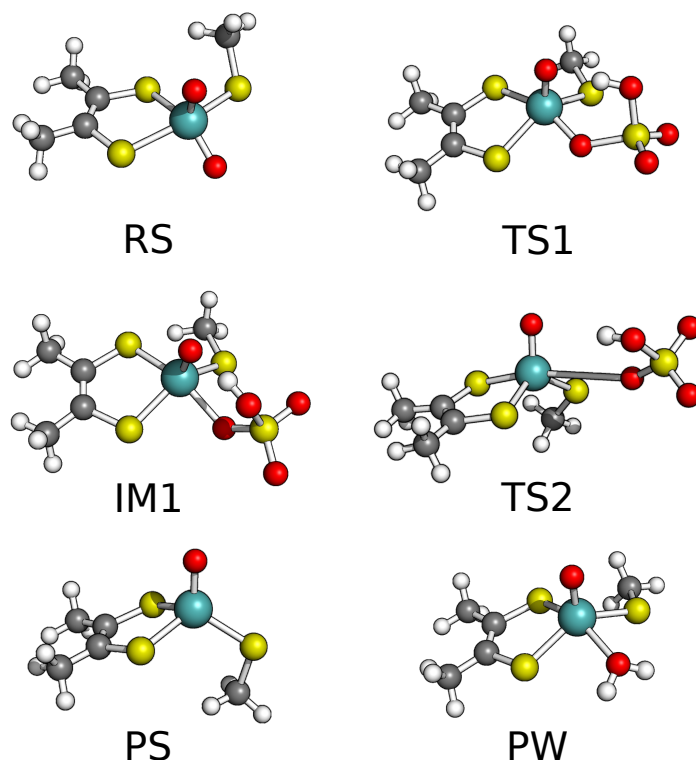


Figure 4.10: Optimized structures for S→OMo mechanism using the model with DMDT and HSO_3^- .

In Figure 4.11 the relative energies for different methods are presented. Comparing LMP2 and LCCSD(T0) results it is visible that LMP2 strongly overestimates relative energies by about 90 – 115 kJ/mol. The reason for this strong discrepancy is the known deficiency in MP2 in the description of oxidation states changes (molybdenum is changing from +VI to +IV state) [113, 117, 118]. In the case of LCCSD, the difference is much smaller, as much as 55 kJ/mol but in the opposite direction, bringing us to the conclusion that the triples and singles contributions are important for a proper description of the system.

Four different energy profiles are shown in Figure 4.12 for HSO_3^- as a substrate and in Figure 4.13 for SO_3^{2-} . The presented energies are the electronic energies in vacuum (ΔE), and the free energies in vacuum and in continuum solvents with dielectric constants of 4 and 80 (ΔG_1 , ΔG_4 and ΔG_{80}). From the figures it can be seen that ΔE and ΔG_1 results are similar for both substrates. The reason for this is that the thermal effects are small, besides an entropic effect of binding a free ligand of ~ 58 kJ/mol for HSO_3^- and

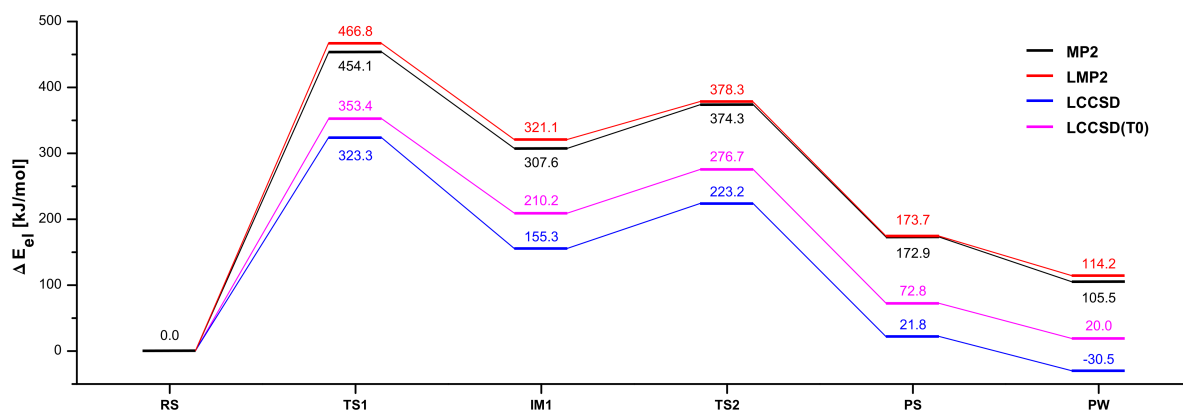


Figure 4.11: Reaction pathway for S→OMo mechanism with HSO_3^- calculated at the different levels of theory.

~ 40 kJ/mol for SO_3^{2-} . On the other hand, if we compare ΔG_1 with ΔG_4 or ΔG_{80} we can see that solvation effects are large. Solvation effects strongly reduce the activation barriers as well as the energy of the intermediate. The reason for this is that both reactant state and substrate are charged species, so the reaction energies are dominated by Coulomb repulsion, which is lowered in solution. The lowest energies are observed in aqueous solution. However, the highest barrier (TS1) is still considerable, 159 kJ/mol for HSO_3^- and 107 kJ/mol for SO_3^{2-} . This means that the reaction in aqueous solution is possible, but very slow.

Taking a closer look at the values for free energies ΔG_1 and ΔG_{80} one can see that the energies are stabilized by ~ 190 kJ/mol. Looking into the ΔG_{80} energies along the path, one can see that the intermediate (IM1) is 7 kJ/mol higher in energy than the isolated reactants. Furthermore, the second transition state (TS2) is significantly lower in energy than TS1 (89 kJ/mol). The product state (PS) is 48 kJ/mol higher in energy than the reactant state and the binding of water should decrease this difference by 15 kJ/mol.

In the case of the SO_3^{2-} substrate (Figure 4.13) the general conclusion is the same, but the ΔG_1 energies are somewhat higher and the ΔG_{80} are somewhat lower. In this case the difference between ΔG_1 and ΔG_{80} energies is ~ 300 kJ/mol. One would expect this trend. In the case of ΔG_1 energies, the activation barriers are larger for the SO_3^{2-} substrate since the total charge is increased in comparison with the HSO_3^- substrate. Since solvent effects have a stronger influence on higher charged species, we obtained lower ΔG_{80} energies for SO_3^{2-} . The highest barrier is still the first transition state with

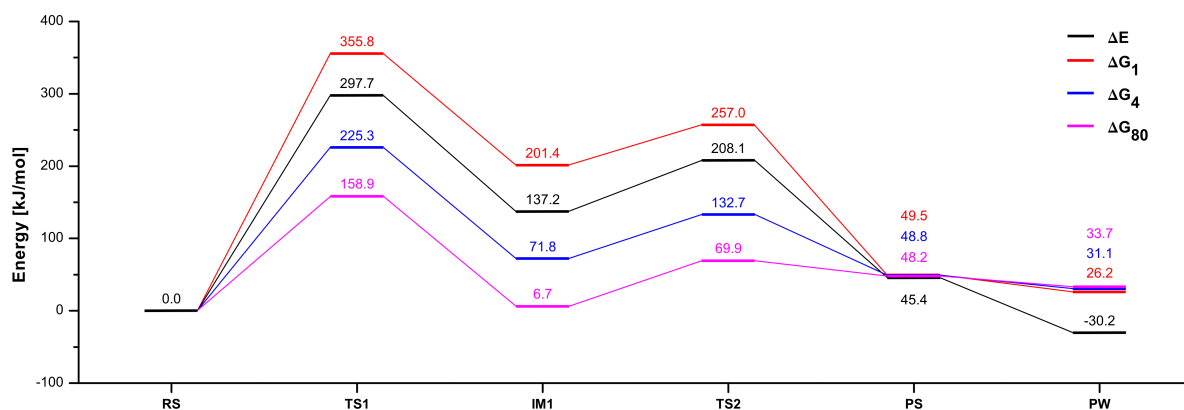


Figure 4.12: Reaction pathway for S→OMo mechanism using the model with DMDT and HSO₃⁻.

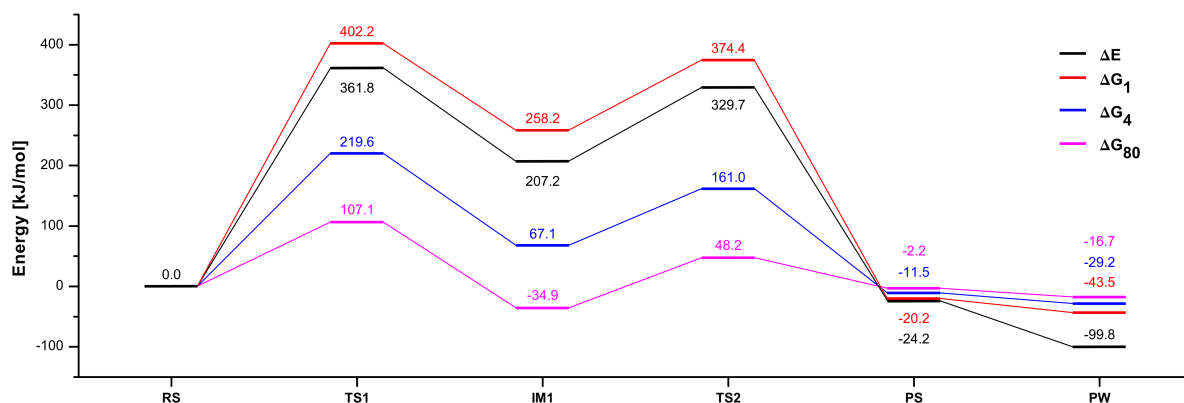


Figure 4.13: Reaction pathway for S→OMo mechanism using the model with DMDT and SO₃²⁻.

a relative energy of 107 kJ/mol (ΔG_{80} energy), which is almost 50 kJ/mol lower than in the case of HSO₃⁻. All other energies are stabilized as well, so IM1 is now 35 kJ/mol lower in energy than RS. PS and PW states are also lower than RS by 2 and 17 kJ/mol, respectively. At the end TS2 is also stabilized by 22 kJ/mol.

Multireference effects

Given the large discrepancy between Møller-Plesset and coupled cluster results, we set out to investigate multireference effects and their weight in our system. First we computed the T1 diagnostic along the reaction path. Observed values were in a range between 0.020 – 0.027, depending on the chosen structure. These values indicate some multireference character, but they are still below the suggested thresholds for the appli-

cation of CCSD(T) (0.04 [119] or even 0.05 [120]). We also compared canonical and local double amplitudes. For this purpose, we needed to reduce our system. Smaller model systems were built mimicking the reactant $[\text{Mo}^{\text{VI}}(\text{O})_2(\text{SH})_3]^-$ and product $[\text{Mo}^{\text{IV}}(\text{O})(\text{SH})_3]^-$ states. Structures were optimized at the B3LYP/def2-TZVP level of theory, including the respective ECP on Mo. These two model systems are shown in Figure 4.14. In both complexes, double amplitudes did not exceed the value 0.05 which is taken as a threshold to identify strong multireference character.

We also wanted to compare the T1 diagnostic obtained from full canonical CCSD(T) and LCCSD(T0) calculations. In the $[\text{Mo}^{\text{VI}}(\text{O})_2(\text{SH})_3]^-$ complex, T1 diagnostics were 0.026 and 0.025 from CCSD(T) and LCCSD(T0) calculations, respectively. In the case of $[\text{Mo}^{\text{IV}}(\text{O})(\text{SH})_3]^-$, we obtained 0.033 and 0.028. One can conclude that the truncation of the virtual space did not influence the value of the T1 diagnostic, despite the significant reduction of accessible configurations. At the end, taking into account the two different diagnostics for the verification of the multireference character, we see that our system has only weak multireference character and it is suitable for an investigation with the LCCSD(T0) method.

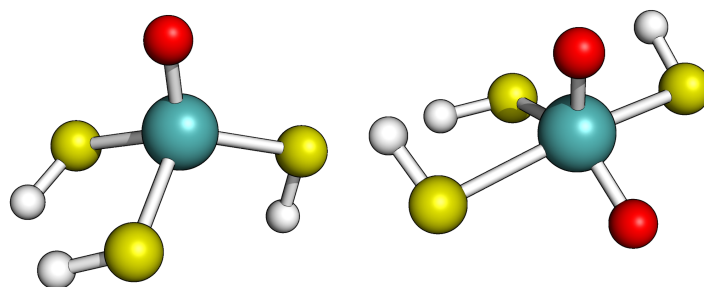


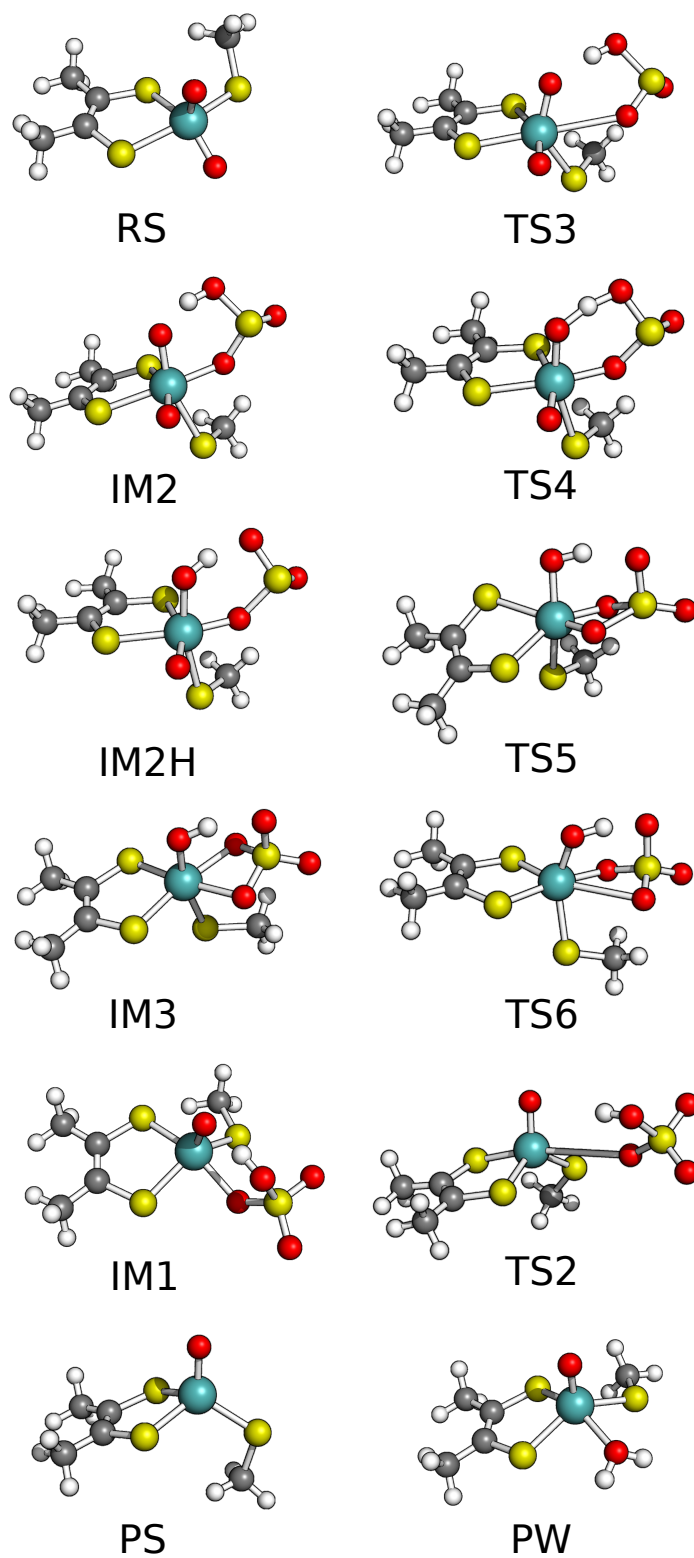
Figure 4.14: Model systems with Mo(IV) (left) and Mo(VI) (right).

O→Mo mechanism

The second investigated pathway was the O→Mo mechanism. This mechanism was studied by Sarkar and coworkers [104], who proposed a three step mechanism. However, we could not find such a simple pathway which would be properly connected. A more intricate mechanism was instead found and is shown in Figure 4.15.

In this case the reaction starts with the same reactant state (RS) as in the S→OMo mechanism. In the next step HSO_3^- approaches the Mo complex and through the tran-

sition state TS3 forms IM2. IM2 is a six coordinated $\text{Mo}^{\text{VI}}\text{-HSO}_3^-$ complex where Mo is bonded to the O from the substrate. In the next step a proton transfers from HSO_3^- to the axial oxo ligand resulting in a $\text{Mo}(\text{OH})(\text{SO}_3)$ complex (IM2H) via the transition state TS4. After that, the sulfur atom of HSO_3^- can attack the equatorial oxo ligand and form IM3 which is a six coordinated Mo complex with a bidentate sulfate ligand. The IM3 state is formed through TS5 which is the chemical step when sulfate is formed and Mo reduced to Mo^{IV} . In this step the Mo-O_{sub} bond is intact. It is first broken in the next step when one of the ligating atoms in HSO_4^- dissociates from the Mo ion. At the same time, the proton is transferred back to the sulfate, so it arrives at the same intermediate IM1 as in $\text{S}\rightarrow\text{OMo}$ mechanism (via TS6). From this point onward the reaction occurs as in the $\text{S}\rightarrow\text{OMo}$ mechanism. If we use SO_3^{2-} as a substrate the mechanism is similar except for the proton transfer, so IM2H and TS4 are omitted.

Figure 4.15: Optimized structures for the O→Mo mechanism using the model with DMDT and HSO_3^- .

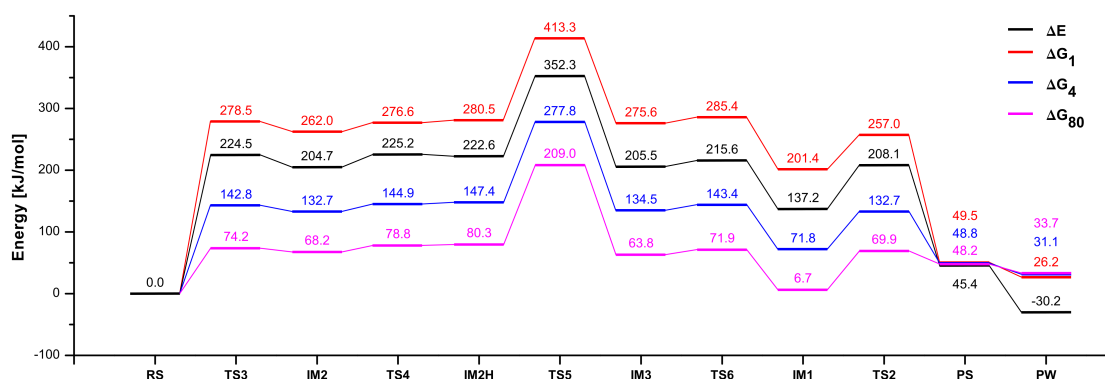


Figure 4.16: Reaction pathway for O→Mo mechanism using the model with DMDT and HSO_3^- .

Figures 4.16 and 4.17 show the relative energies for the O→Mo mechanism with two different substrates. Here it is again observed that the solvation effects decrease the relative energies for all transition states and intermediates. The entropic and thermal effects show the opposite effect.

In this mechanism a large difference between the two substrates is observed. In the case of the HSO_3^- substrate the transition state TS5 holds the obviously highest energy ($\Delta G_{80} = 209$ kJ/mol), so it is the rate-limiting step. In this step the formation of the S-O bond and the binding of the substrate takes place. All other states have similar energies between 63 and 78 kJ/mol, in aqueous solution, except the IM1 which is much lower in energy (7 kJ/mol).

For the SO_3^{2-} substrate the situation is somewhat different and the energies show larger variation. Despite the evidence that TS5 has such a high energy ($\Delta G_{80} = 163$ kJ/mol), this state is not the rate-limiting step, because TS3 state has an energy of 168 kJ/mol. In the case of the HSO_3^- substrate the hydrogen atom in HSO_3^- strongly facilitates the formation of the first Mo-O bond by forming a hydrogen bond with the O_{ax} atom. The TS3 state is then lowered by 94 kJ/mol.

In both cases the rate-limiting steps show higher activation energies than in the previous mechanism. In the case of HSO_3^- substrate TS5 from O→Mo mechanism is 50 kJ/mol higher in energy than TS1, the rate-limiting step of S→OMo mechanism. For the SO_3^{2-} substrate TS3 is 60 kJ/mol higher than TS1. At the end we can conclude that this mechanism is less favorable than the S→OMo mechanism.

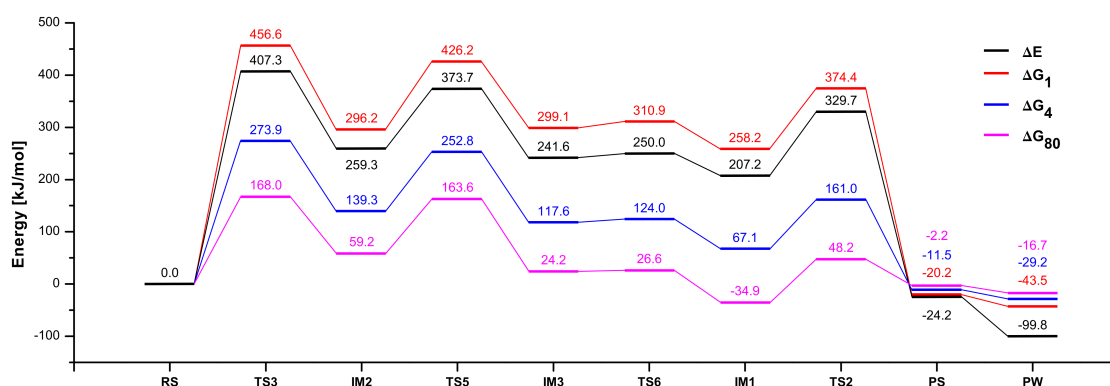


Figure 4.17: Reaction pathway for O→Mo mechanism using the model with DMDT and SO_3^{2-} .

S→Mo mechanism

The last investigated mechanism was the S→Mo mechanism. In this case, the lone pair of the S atom from the substrate attacks Mo in the first step. The IM4 intermediate is formed via transition state TS7. This intermediate is a six-coordinated complex with the substrate bound to Mo by the S atom. In the following, an approximate transition state TS8 was obtained with an intact Mo-S_{sub} bond and the shortest possible O_{eq}-S_{sub} bond. This transition state leads to the same IM1 as in the previous mechanisms and from here onward the reaction continues through TS2 and PS as before. This mechanism is shown in the Figure 4.18.

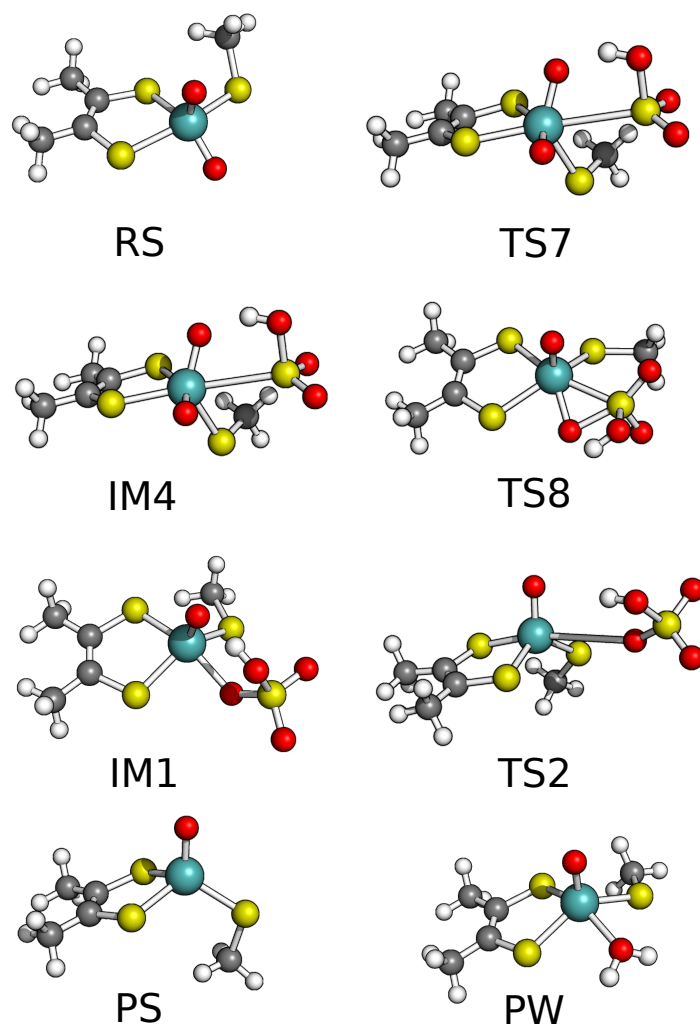


Figure 4.18: Optimized structures for S→Mo mechanism using the model with DMDT and HSO_3^- .

Energies for the last investigated mechanism are shown in Figures 4.19 and 4.20. In this mechanism, high barriers are obtained for both protonated and deprotonated substrates. These energies are higher than in the S→OMo mechanism. The activation barrier for the HSO_3^- is 256 kJ/mol, 98 kJ/mol higher than in the S→OMo mechanism. For SO_3^{2-} , the activation barrier is significantly lowered to 161 kJ/mol, but it is still 54 kJ/mol higher than in the case of the S→OMo mechanism. Therefore, we can also rule out this mechanism.

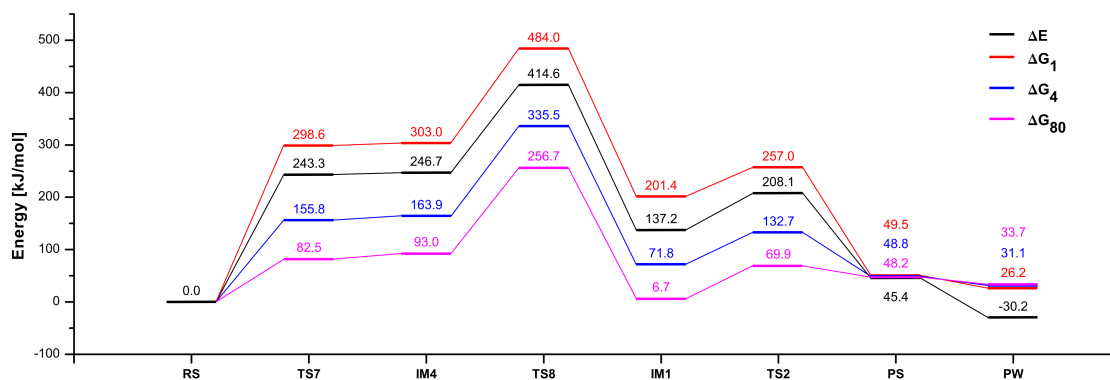


Figure 4.19: Reaction pathway for S→Mo mechanism using the model with DMDT and HSO₃⁻.

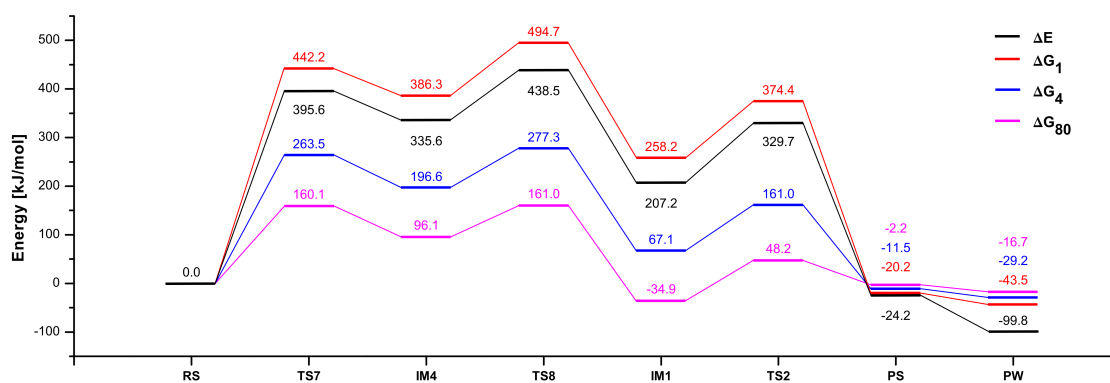


Figure 4.20: Reaction pathway for S→Mo mechanism using the model with DMDT and SO₃²⁻.

4.1.3 Final Remarks

In this study our goal was to investigate all SO mechanisms under the same method and the same model system. Our model system followed similar conventions as in previous theoretical studies. For most cases, the conclusions were similar. The only study where lower activation energies were obtained was the study by Hernandez-Marin and Ziegler [102]. In this case a slightly different model system was used. An arginine residue was included in the model system and on that way the approach of two negatively charged species was avoided. The conclusion was nevertheless the same. The S→OMo mechanism is favoured.

In the end we can conclude that in both cases, HSO_3^- and SO_3^{2-} substrates, the pathway with the lowest barrier is found in the S→OMo mechanism. One intermediate state was identified, a Mo-sulfate complex, and two transition states. In the first transition state the formation of the $\text{S}_{sub}\text{-O}_{eq}$ bond occurs and in the second transition state the Mo-O_{eq} bond is cleaved. The first transition state involves the chemical step and the reduction of the Mo ion. This step is rate-limiting.

It was also observed that solvation had a large effect on the activation barriers. In the previous studies for realistic protein calculations dielectric constants of 2 – 20 were used [121–124]. In the case of charged groups it was also observed that larger values for dielectric constants are necessary [125]. In our study the smallest barriers were obtained in solution with a dielectric constant of 80. Nevertheless, those barriers were still relatively high, meaning that the reaction would be very slow.

In Sarkar's study [104] they discarded the S→OMo mechanism because in their intermediate state Mo-O_{eq} and $\text{O}_{eq}\text{-S}_{sub}$ the bond lengths were found to be 2.4 and 1.5 Å, respectively. This suggested that the oxygen atom transfer was completed. This would imply an one-step mechanism without the formation of a Michaelis complex. In our case the bond lengths are different, so for Mo-O_{eq} we obtained 2.23 and 2.07 Å, for HSO_3^- and SO_3^{2-} substrates, respectively, so O_{eq} is still weakly bound to Mo. On the other hand, the $\text{O}_{eq}\text{-S}_{sub}$ bond length was found to be larger than in Sarkar's study, 1.54 and 1.65 Å for HSO_3^- and SO_3^{2-} substrates, respectively, confirming that O_{eq} is not fully transferred to sulfite. Observing only the bond distances we can say that we obtain a Michaelis complex for our intermediate state IM1.

Another strong argument in support of the S→OMo mechanism is found in the experimental work from Brody and Hille [126]. In their study they investigated the interaction

of the enzyme center with HSO_3^- and dimethylsulfite. They found that both substrates react with the same rate constant. Therefore, they concluded that the reaction should occur through a nucleophilic attack of the sulfur lone pair at the equatorial oxygen. The only difference between these two substrates is that in the case of HSO_3^- the intermediate state is more stabilized. They explained that this stabilization appears from the substrate oxyanion groups which can be directly coordinated to the molybdenum or participate in electrostatic or hydrogen interactions with other functional groups of the active site. Since we did not have any other functional groups included in our model we could not test this effect. On the other hand, in the study of Marin-Hernandez and Ziegler [102] they included in their model an arginine group and in this way included hydrogen bonds between this residue and the substrate. The same mechanism as in our case was verified.

In the end we can conclude that there is also evidence from other theoretical as well as experimental works that the reaction occurs through the $\text{S} \rightarrow \text{OMo}$ mechanism. The only problem remaining is that the obtained energies are too high. Since, we observed that solvation effects are large and inclusion of an extra residue as in the case of Marin-Hernandez and Ziegler [102] study significantly stabilizes the barriers one can envision a QM/MM approach for this system. In this way other residues of the active site could be included in the calculation and unbiased reaction energies obtained.

4.2 LMOMO Calculations

In the following we will continue investigating the Mo-systems, but this time using the hybrid LMOMO approach. Two different reaction mechanisms will be under investigation. The first one will be again the enzyme sulfite oxidase (SO) which was previously studied. In this case a slightly different model system will be used. The second system is the dimethyl sulfoxide reductase (DMSOR). For DMSOR two different model systems will be used as it will be described later.

These two systems were chosen because of their similarities, but also because of some specific properties. In both systems oxygen atom transfer reaction occurs. The main difference is that they proceed in different directions (to and from the metal). Another important feature is that the most stable state for this species is the singlet state, so it is possible to make use of closed-shell methods. Another issue is the multireference

character of the systems under the investigation. As we have already shown, only a small multireference character was observed, so that single-reference CC methods are applicable.

Based on these systems we will test the applicability of the LMOMO approach to reaction mechanisms in metallobiosites. Also, we will test the convergence of our method by increasing the high-level region in the systematic fashion. Our goal is to show that the application of accurate methods such as coupled cluster is not necessary on the whole investigated system, but rather on the small part where reaction occurs.

4.2.1 Computational details

Wave function calculations were carried out with the cc-pVTZ basis set [62] in combination with the aug-cc-pVTZ-PP orbital basis and the Stuttgart/Dresden ECP28MDF pseudopotential [108] for the description of Mo (AVTZ). The Pipek-Mezey localization scheme [22] was used for the occupied orbitals. Orbital domains were determined according to the NPA criteria, with $T_{\text{NPA}}=0.03$ [24]. In the case of reaction pathways calculations, it is important that orbital domains are of the same size in every structure. For some of the structures along the DMSOR reaction path that was not the case, so domain merging was applied [114]. In this way, the same number of orbitals was used in the high level region for all stationary points.

DF approximations were applied throughout [31, 32, 127]. The auxiliary basis sets for Coulomb and exchange were def2-TZVPP/JKFIT for Mo and cc-pVTZ/JKFIT for all other atoms [66, 68]. In the case of the correlation part the cc-pVTZ/MP2FIT and the def2-TZVPP/MP2FIT basis sets were used for Mo [67, 128]. Orbital pairs were classified according the distance criteria, strong pairs were defined within a distance of 3 bohr and close pairs as within 5 bohr [107]. In the coupled cluster treatment strong orbital pairs were fully included. Close pairs were treated at the MP2 level, but partly included in the triples calculation [39].

When free energies are presented, all of the corrections to the electronic energies were taken from the previous studies. In the case of SO, the added corrections were obtained as a difference between the values of Hernandez-Marin and Ziegler [102] and our DFT energies computed with the BP86 functional and the cc-pVTZ basis set, which is similar to the basis set applied in the original calculations, a triple- ζ -quality STO basis set. In the case of DMSOR the corrections were taken from Reference [129].

4.2.2 LMOMO method for SO

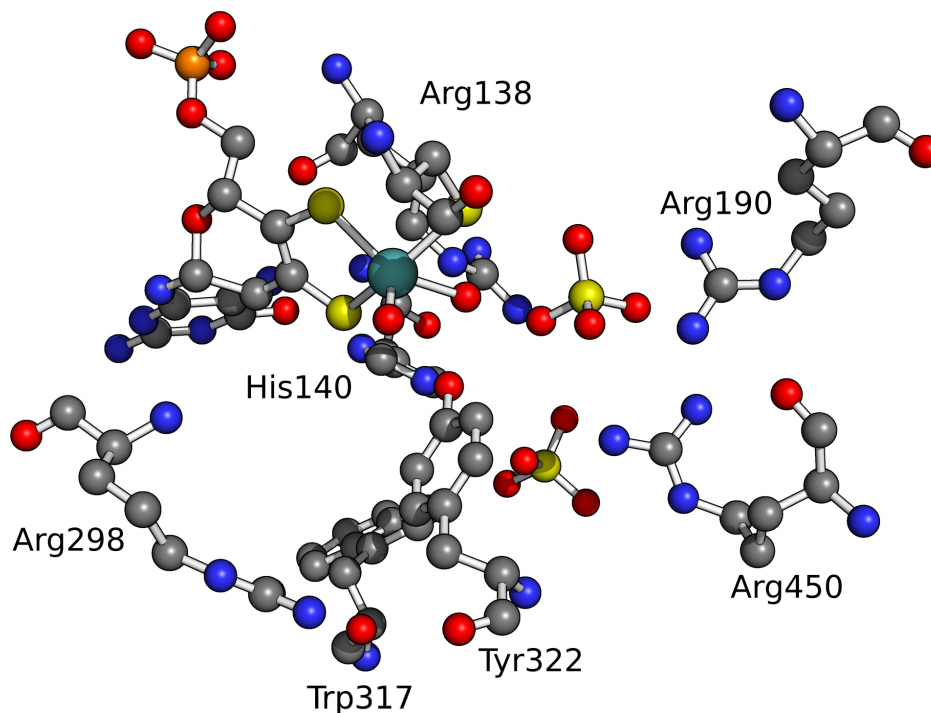


Figure 4.21: Active center of sulfite oxidase.

In the previous study by Hernandez-Marin *et al.* [102] they used a slightly modified model system than what we described in the previous section. They included the residue Arg-138 from the chicken sulfite oxidase X-ray structure, PDB code 1SOX [130] (Figure 4.21), modelled by the methyl guanidinium group. In this way the approach of two negatively charged species is avoided and reaction barriers are stabilized. The other important factor is the possibility of forming the hydrogen bonds between this residue and the substrate, which was proposed as an important effect. On the other hand it was also shown in the previous clinical studies that this residue was important for the catalytic process. In the case when Arg-138 was mutated to glutamine the protein exhibited only 2% of its native activity [131]. Thus, it was proposed that this amino acid has an important role in the attraction of the anionic substrate sulfite to the binding site near molybdenum. In the following, we will use the same model system as in Reference [102] with a small modification. Two methyl groups were added to change the ligand to

DMDT, which was not the case in the original study. Here we will use only SO_3^{2-} as a substrate.

Despite the experimental evidence which suggests the importance of Arg-138 one should notice that around the active site other charged residues are present and should be also included in the calculation. However, to include all of these, a QM/MM approach would be necessary. The problem with such a method is that one usually uses DFT in the QM region. As we saw in our previous study, the difference between DFT and LCCSD(T0) results was 2 – 62 kJ/mol depending on the structure [94]. Therefore, one is in need to find new methods to approach the CC limit. Thus, we want to test the hybrid LMOMO method which can later be included in a QM/MM scheme at the QM level.

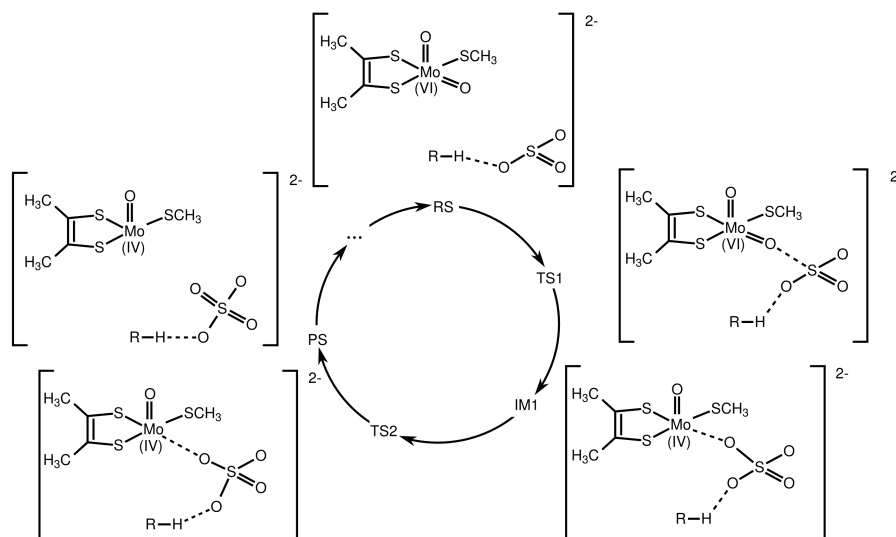


Figure 4.22: Reaction scheme for the oxidation of sulfite in SO following the S→OMo mechanism.

As observed before, the oxidation of sulfite in SO follows S→OMo mechanism. This mechanism is a two step mechanism, where in the TS2 the O atom transfer is completed with full dissociation of the bond to Mo. Also, the arginine residue is hydrogen bonded to the substrate, but it is not directly involved in the reaction.

In Figure 4.23 the LMP2, LCCSD and LCCSD(T0) results are plotted. It can be seen that the same difference between the methods is observed as in the chapter before, suggesting that the Arg residue does not affect the convergence of the QM results. The

only difference between these and previously presented results is that the activation barriers are significantly lower. For example, in the case of the TS1 state, which is a rate-limiting step, the energy barrier is only 37 kJ/mol in comparison with 107 kJ/mol in the previous calculations.

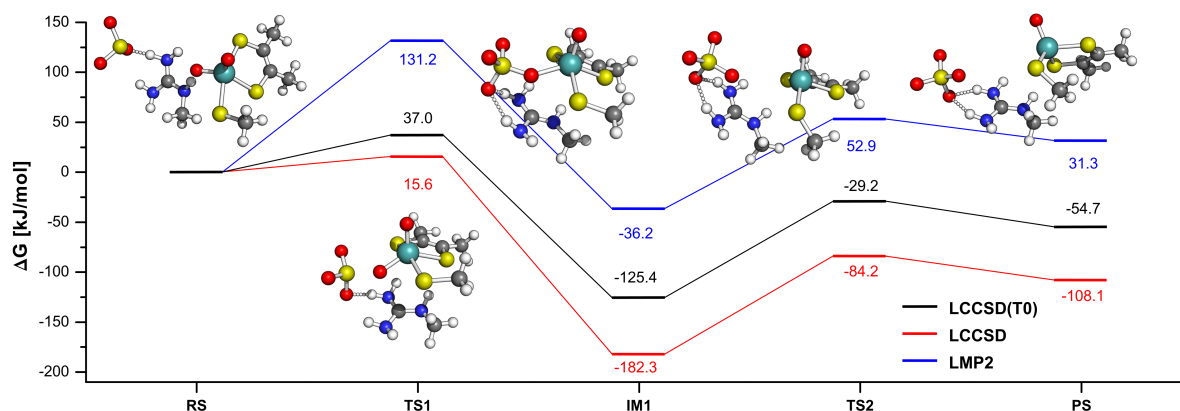


Figure 4.23: Free energies (in kJ/mol) along the reaction path of SO.

Before we start evaluating the LMOMO method, one more difference between the present system and the one used in the first part of this chapter should be noted. In the first model system we saw that the energy changes drastically when one goes from the free energy in vacuum to the free energy in solution and also when one changes from $\epsilon = 4$ to $\epsilon = 80$. It was our intention to reevaluate this trend with the added arginine residue. One can see in Figure 4.24 that in the second case (model+Arg-138) the correction is significantly smaller and amounts to a maximum of 20 kJ/mol in comparison with 110 kJ/mol for the first case (model). As stated before the large difference observed in the first system are due to the interaction of two charged species. On the other hand, with the inclusion of arginine in the model, a charge is neutralized, reducing the effect.

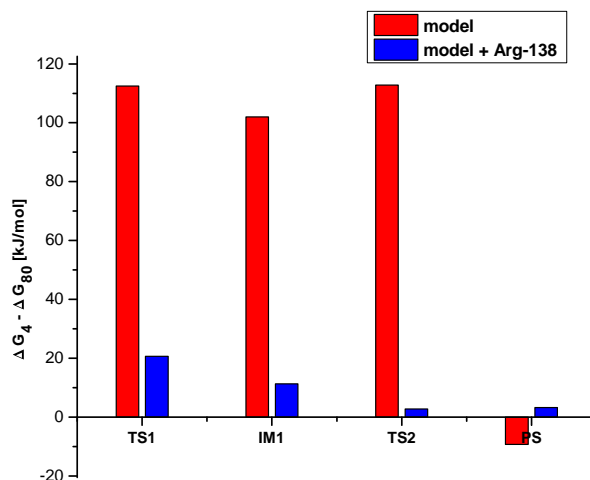


Figure 4.24: Difference between the free energies with $\epsilon = 4$ and $\epsilon = 80$ for the two different model systems along the S→OMo reaction path.

For the LMOMO calculations, three different high level regions were selected. They are presented in Figure 4.25 in red. The TS2 structure was taken as an example. In the smallest region (R1) only orbitals which had Mo in their domain list were correlated at the coupled cluster level. In the next selection, atoms directly included in the bond breaking/formation were also included in the high level. This selection will be denoted as R2. Finally, we include not only molybdenum, oxygen and sulfur but also the first neighbours of the molybdenum. This will correspond to the R3 selection. In this way we are increasing the number of orbitals in the high level region, so in the case of the R1 selection we have 15 orbitals correlated at the high level, in R2 27 orbitals and in the R3 selection the total number of orbitals at the high level is 33. The total number of valence orbitals in LCCSD(T0) calculations for this system was 61. In this way the number of orbitals correlated at the LCCSD(T0) level was significantly reduced in all cases.

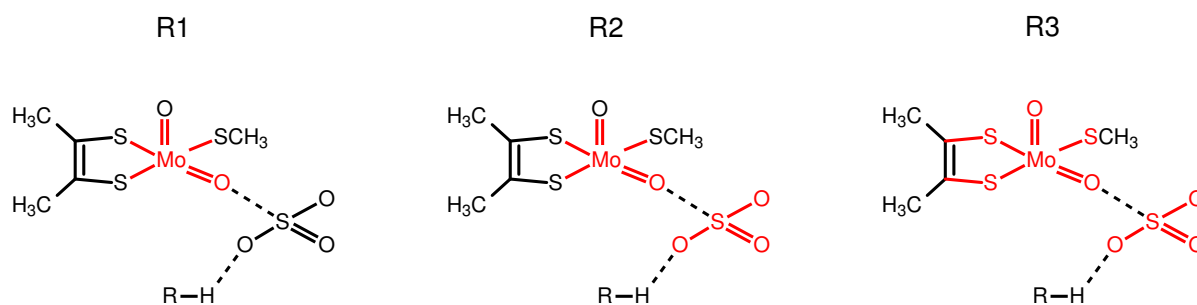


Figure 4.25: Three different regions for the SO active site.

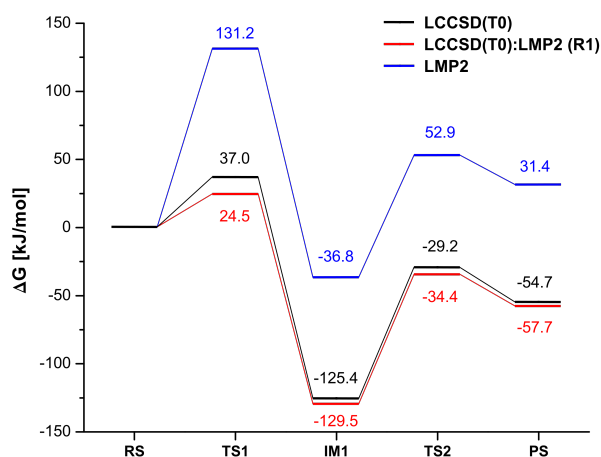


Figure 4.26: Free energies (in kJ/mol) along the reaction path of SO. The LMOMO results are shown for selection R1.

In Figure 4.26 LCCSD(T0):LMP2 results for the R1 selection, as well as LMP2 and LCCSD(T0) results are presented. One can see that despite the large difference between LMP2 and LCCSD(T0) results, even the smallest region was enough to obtain reasonable results. The results for all three region selection in comparison with LMP2 and LCCSD(T0) results are shown in Table 4.1. If we now look into the results for the R2 selection, deviations of 2 – 3 kJ/mol are present. For this type of systems, this would be already a reasonable compromise. In the R3 selection, the results are close to converged and only in the case of TS1 the error was a little bit above 1 kJ/mol.

Table 4.1: Relative electronic energies (in kJ/mol) computed for the SO reaction at different levels of theory. Three different selections for the high-level region have been considered, as depicted in Figure 4.25.

	LMP2	LCCSD(T0):LMP2			LCCSD(T0)
		R1	R2	R3	
RS	0.0	0.0	0.0	0.0	0.0
TS1	90.1	-16.7	-6.2	-5.2	-4.1
IM1	-68.3	-160.1	-158.2	-156.4	-156.9
TS2	65.7	-21.7	-19.2	-16.9	-16.5
PS	55.6	-33.4	-31.5	-29.3	-30.4

As stated earlier, one important aspect is the multireference character of the system. The easiest way to check this is to obtain the T1 diagnostic. The T1 diagnostic is defined as the norm of the singles amplitude vector divided by the square root of the number of electrons [132]:

$$T_1 = \frac{1}{\sqrt{N_{elec}}} |\mathbf{t}_1|. \quad (4.5)$$

In this way the quality of a CCSD wave function is evaluated and if $T_1 < 0.04$ [119] or even $T_1 < 0.05$ [120] the coupled cluster methods should still be applicable, as long as triple excitations are included. Therefore, we decided to check the T1 diagnostic obtained from LMOMO calculations in comparison to the one from a full LCCSD(T0) run. As one can see in Table 4.2 a good agreement was achieved independent of the region size. One observes only a slight underestimation. Thus, we can conclude that this method is also suited to obtain such a valuable piece of information, the multireference character of the system, at a relatively low cost.

We showed that the energies as well as the T1 diagnostic obtained with the LMOMO method were in good agreement with the full QM calculation. This places it as an alternative to the custom DFT applications. On the other hand, it is well known that coupled cluster methods are much more time consuming in comparison with DFT. For this reason, we wanted to test the timings for the hybrid LMOMO approach. To test the timings we performed calculations on the model system, where we included the full molybdopterin group (Figure 4.27). The selection of the orbitals for the high level region was the same as in the previous case. The number of orbitals in the high level region was 15, 27 and 33 for the R1, R2 and R3 selections, respectively. In the full LCCSD(T0)

Table 4.2: T1 diagnostics for all states in the SO reaction pathway calculated at the different levels of theory. Three different selections for the high-level region have been considered, as depicted in Figure 4.25.

	LCCSD(T0):LMP2			LCCSD(T0)
	R1	R2	R3	
RS	0.016	0.019	0.019	0.021
TS1	0.020	0.023	0.024	0.025
IM1	0.016	0.018	0.019	0.020
TS2	0.016	0.018	0.019	0.021
PS	0.017	0.018	0.019	0.021

calculation the number of correlated orbitals was 97. In Figure 5 the real times for the LCCSD(T0) and LCCSD(T0):LMP2 calculations are shown, excluding the HF part. As one can see, even for the R3 selection we obtained significant savings in comparison to the full LCCSD(T0) calculation (270 min for R3 selection and 903 min for full LCCSD(T0)). Therefore, we introduced a method which is less expensive than full LCCSD(T0), but with comparable accuracy. These two important features of the method qualify it as a good replacement for DFT in the investigation of reaction mechanisms.

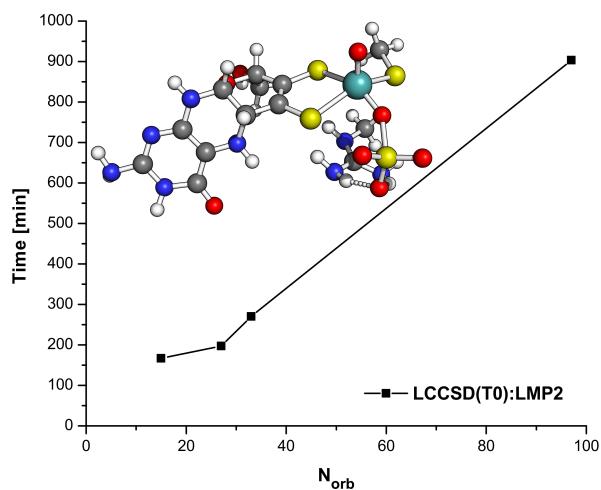


Figure 4.27: Time needed for LCCSD(T0):LMP2 calculations dependent on the number of orbitals in high level region.

4.2.3 LMOMO method for DMSOR

DMSOR catalyses the oxygen-atom transfer from dimethyl sulfoxide (DMSO) to the Mo^{IV} active site yielding dimethyl sulfide (DMS) and Mo^{VI}. The active site for DMSOR was modelled by [Mo(DMDT)₂(CH₃O)]⁻, where CH₃O⁻ is a model of the serine ligand. All structures used in these calculations were taken from the Reference [129] and they are shown in Figure 4.31.

In the case of DMSOR a relatively broad consensus about the mechanism exists. This reaction follows a two-step mechanism and starts when DMSO enters the active site. Via the first transition state the intermediate state with weakly bound DMSO to the Mo^{IV} was built. After that, in the second step the S-O bond is cleaved in an oxygen atom transfer reaction. This reaction mechanism is depicted in Figure 4.28. All previous studies agreed that the second transition state is the rate-limiting step, but the activation barriers were different. Because of that we wanted to employ the hybrid LMOMO approach to see if we can obtain better results.

Webster and Hall [133] performing B3LYP calculations on the model [Mo(DMDT)₂(OCH₃)]⁻ obtained an activation energy of 37 kJ/mol. On the other hand, Thapper *et al.* [134] studied a slightly different model system and obtained an activation energy of 76 kJ/mol. For this system, activation energies in the range from 37 kJ/mol to 80 kJ/mol were obtained in different studies [133–144]. In the recent paper from Li *et al.* [129] they improved their results using the LCCSD(T0) method and obtained excellent agreement with the experiment. In our studies we wanted to show that one can restrict the CC treatment to a relatively small region, just as in the case of SO. Thus, we again applied the hybrid LMOMO approach in the calculation of the pathway.

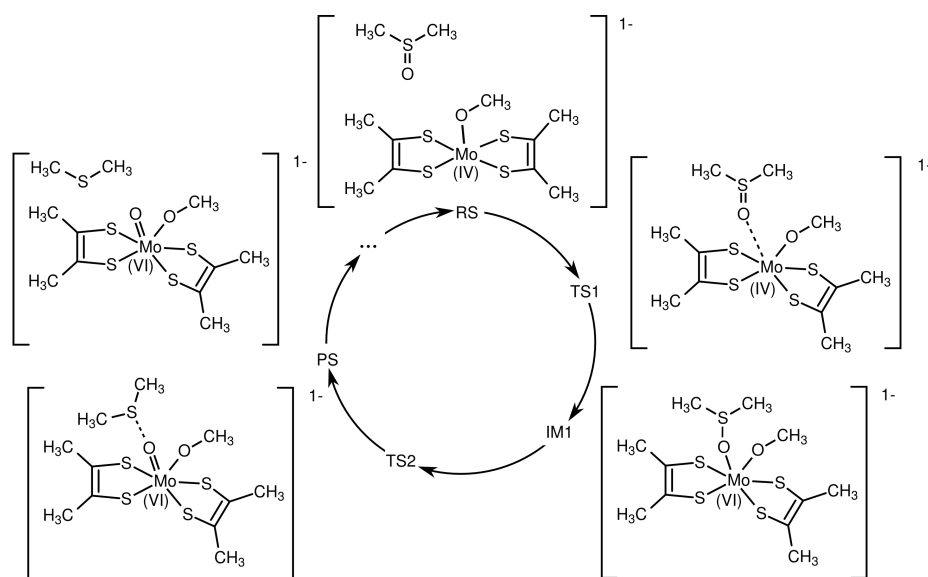


Figure 4.28: Reaction scheme for the reduction of dimethyl sulfide catalysed by DMSOR.

In the previous investigation [129], a biomimetic model was also considered. The difference between this biomimetic model and the one which we already introduced is one of the ligands, PhO^- , instead of CH_3O^- . TS2 geometries for both complexes are shown in Figure 4.29. In both cases the same mechanism is followed and similar stationary points were identified.

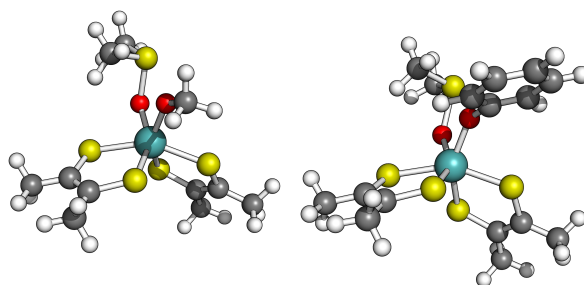


Figure 4.29: TS2 structures for the DMSOR cluster model (left) and a biomimetic complex (right) of the same enzyme.

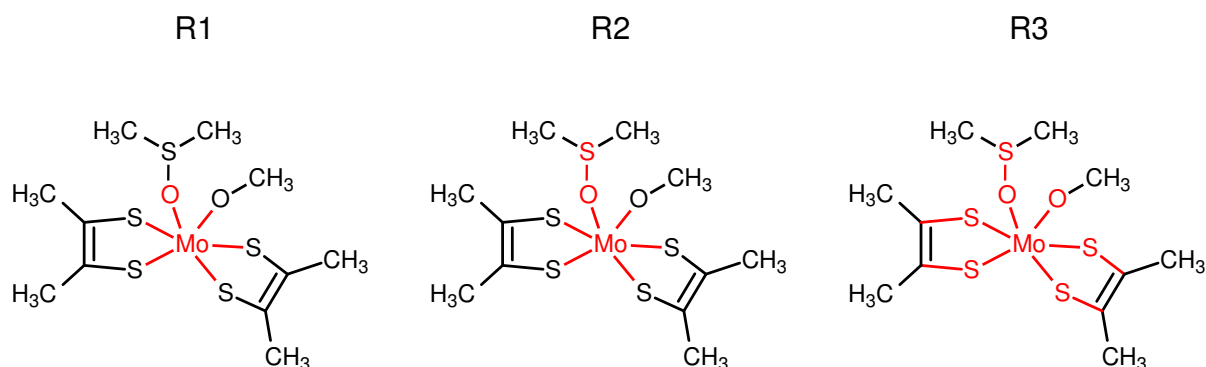


Figure 4.30: Three different high-level region selections for the DMSOR reaction. The number of orbitals included in the high-level are 17 (R1), 20 (R2) and 28 (R3) from a total of 57 valence orbitals.

Three different selections for LMOMO calculations were again used. In the first selection, the region consisted of only orbitals which belong to Mo. This region will be noted as R1. In the second selection, the region was extended to the sulfur and oxygen which took part in the reaction. This will be region R2. In the last case, we included also the first neighbours of Mo to the R2. All three region selections are depicted in Figure 4.30.

Before we look into the LMOMO results we want to compare results obtained from different local methods, LMP2, LCCSD and LCCSD(T0) (Figure 4.31). In the TS1 and IM1 states the difference between LMP2 and LCCSD(T0) as well as LCCSD and LCCSD(T0) is less than 5 kJ/mol. On the other hand, LMP2 and LCCSD results show larger errors in the TS2 and PS states. In the PS state LMP2 even shows an error of about 100 kJ/mol and triple contributions for this state are also important and amount to 57 kJ/mol. In the TS2 state the situation is somewhat better and the LMP2 error amounts to only 22 kJ/mol.

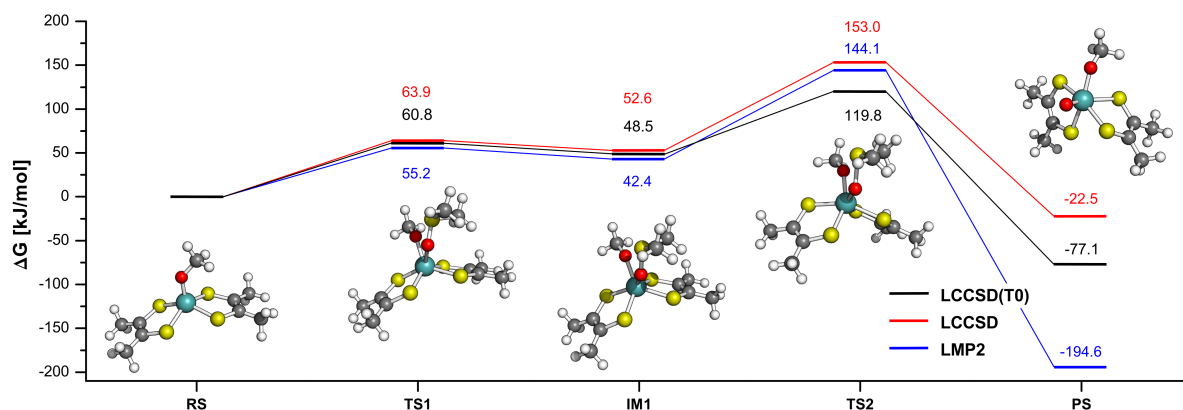


Figure 4.31: Free energies (in kJ/mol) along the reaction path of DMSOR.

If we now compare the results for the different ligands (Table 4.3), one observes that the LCCSD(T0) results are quite close for each stationary point. On the other hand, looking at the LMP2 values, one observes much larger errors with PhO^- . This effect should be connected to the aromatic character of the ligand.

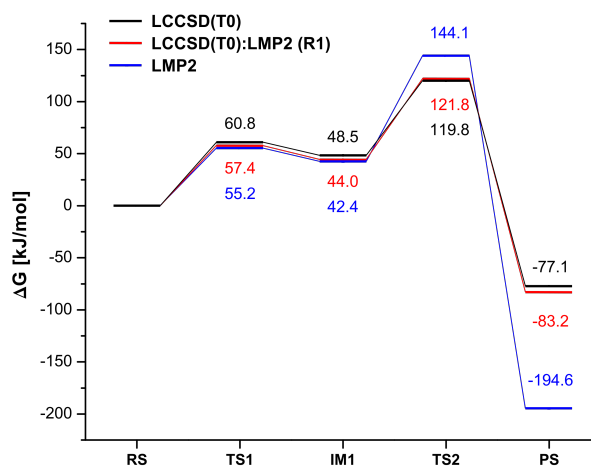


Figure 4.32: Free energies (in kJ/mol) along the reaction path of DMSOR. The LMOMO results are shown for selection R1.

In Figure 4.32 LMP2, LCCSD(T0) and LCCSD(T0):LMP2 (R1) results are shown. One can see that the LCCSD(T0):LMP2 results are properly describing the reaction pathway even in the points where LMP2 fails (TS2, PS). Looking closely into the results (Table 4.3), smaller differences than in the SO reaction for the R1 region are observed.

Table 4.3: Relative electronic energies (in kJ/mol) computed for the DMSOR reaction path at different levels of theory. Three different selections for the high-level region have been considered. Values for the cluster (CH_3O^-) are shown, corresponding to the DMSOR active site model, together with the energies for the biomimetic complex (PhO^-).

	LMP2	LCCSD(T0):LMP2			LCCSD(T0)
		R1	R2	R3	
<hr/>					
CH_3O					
RS	0.0	0.0	0.0	0.0	0.0
TS1	15.5	17.7	19.1	21.4	21.2
IM1	10.5	12.0	14.1	17.0	16.6
TS2	93.6	71.3	69.5	69.1	69.4
PS	-252.2	-140.9	-139.1	-138.0	-134.8
<hr/>					
PhO					
RS	0.0	0.0	0.0	0.0	0.0
TS1	28.8	26.4	26.5	25.6	23.1
IM1	8.2	6.7	9.4	12.1	12.0
TS2	87.4	58.4	56.5	55.5	57.1
PS	-257.4	-143.0	-141.2	-139.3	-134.9

On the other hand, that was not the case for R2 and R3 selections. In the R3 selection, the LCCSD(T0):LMP2 deviations are within the 1 kJ/mol except for the product state, for the CH_3O^- ligand. The product state for this ligand shows deviation of about 3 kJ/mol. In the case of PhO^- ligand, all deviations are somewhat larger, but not more than 5 kJ/mol.

At the end, we also looked into the T1 diagnostic for the DMSOR reaction and the same conclusion as in the SO case can be drawn. A weak multireference character is observed, with a good agreement between LCCSD(T0):LMP2 and LCCSD(T0) T1 values. Results for the T1 diagnostic are shown in Table 4.4.

Table 4.4: T1 diagnostics for all states in the DMSOR reaction pathway calculated at the different levels of theory. Three different selections for the high-level region have been considered, as depicted in Figure 4.30.

	LCCSD(T0):LMP2			LCCSD(T0)
	R1	R2	R3	
RS	0.016	0.016	0.017	0.019
TS1	0.017	0.017	0.019	0.020
IM1	0.017	0.017	0.019	0.020
TS2	0.023	0.026	0.028	0.029
PS	0.018	0.016	0.018	0.020

4.2.4 Final Remarks

In this chapter, reactivity studies of molybdenum enzymes highlighted the importance of highly accurate methods for the description of reaction pathways. Results obtained with LCCSD(T0), LMP2 and LCCSD(T0):LMP2 were compared. We concluded that LMP2 is not a suitable approach for the calculation of such reaction pathways, since it fails to properly describe the change in the oxidation state of the metal center. Since the change of the oxidation state is localized on the metal center and only slightly affects the rest of the system, the effect can be handled locally. Therefore, in the hybrid LCCSD(T0):LMP2 scheme we treated the metal center at the LCCSD(T0) level and the rest of the system at the LMP2 level. This was applied for both the S→OMo mechanism for SO and also the main step in DMSOR.

In the SO mechanism the difference between the LMP2 and LCCSD(T0) results was significant. However, we showed that by including only the orbitals which belong to the metal center this difference was drastically reduced. By adding the direct coordination shell to the metal center, the relative energies converged to the LCCSD(T0) results and the biggest observed difference for this reaction mechanism was 1.1 kJ/mol.

In the DMSOR case the difference between LMP2 and LCCSD(T0) was observed only in the second transition state and in the product state. In this case we also obtained a good agreement between the LCCSD(T0):LMP2 and LCCSD(T0) results. The deviations are somewhat larger than in the SO case, but not more than 5 kJ/mol.

In the end we also tested the values for the T1 diagnostic obtained from full canonical, local and hybrid LMOMO calculations. It was shown that T1 diagnostics obtained from

all three methods were in good agreement in both cases, SO and DMSOR. Therefore, one can use the LMOMO scheme as a quick check of the multireference character of the system.

Another important use of the LMOMO method is to test the accuracy of LMP2 since we saw that already the smallest region can correct for the LMP2 behaviour. In the case when it is observed that LMP2 might lead to artifacts, the LMOMO scheme should be used, by correlating the metal center and the first neighbours of the metal at the high level and the rest at the lower level. In this way energies which are in good agreement with LCCSD(T0) results can be obtained.

CHAPTER 5

Open Shell Systems

In the previous chapters the LMOMO method was applied to closed-shell systems. We will make one step further by expanding this treatment to open shell metal centres. This is of utmost importance since open shell metal centres commonly occur in enzymes and model coordination compounds. At the same time, it is generally known that the electronic configuration of open-shell systems is more difficult to describe than the one of closed-shell systems. In this chapter we will start by evaluating the LMOMO accuracy on the example of small benchmark systems where binding energies will be calculated. After assessing the accuracy for open-shell systems we will make use of this method for the investigation of the mechanism of the nitrite reductase and more demanding properties such as electron affinities.

5.1 Open-Shell Test Systems

One of the fundamental questions in bioinorganic chemistry is metal ion selectivity. To address this question accurate computational protocols of the species involved are needed. In the previous study of Gutten *et al.* [145], they tested the influence of the optimization procedure, the solvation effects and the accuracy of the calculated electronic energy on the results. In this study we focus only on the calculation of the electronic energy. In their study, they found out that RI-MP2 results were generally in good agreement with UCCSD(T). On the other hand, the comparison between DFT and UCCSD(T) was not so satisfactory. DFT performed reasonably well for uncharged ligands, but in the case of charged ligands, the results deviated by more than 10 kcal/mol. Therefore, we decided to use the same complexes to study the accuracy of our newly developed LMOMO method for open shell systems.

In this study binding energies were calculated using different levels of theory, e.g. UCCSD(T), LUCCSD(T0) and LMOMO. Four complexes $[\text{M}^{\text{II}}(\text{CH}_3\text{S})(\text{H}_2\text{O})]^+$ (LI), $[\text{M}^{\text{II}}(\text{H}_2\text{O})_2(\text{H}_2\text{S})(\text{NH}_3)]^{2+}$ (TH), $[\text{M}^{\text{II}}(\text{CH}_3\text{S})(\text{NH}_3)(\text{H}_2\text{O})(\text{CH}_3\text{COO})]$ (SQ), $[\text{M}^{\text{II}}(\text{H}_2\text{O})_2(\text{SH})(\text{CH}_3\text{COO})(\text{Im})]\cdot\text{H}_2\text{O}$ (SP) with three different metal ions (Fe^{2+} , Cu^{2+} and Mn^{2+}) (Figure 5.1) were used as test systems. These three metals were taken in their high-spin state ($S = \frac{1}{2}$ for Cu^{2+} , $S = 2$ for Fe^{2+} and $S = \frac{5}{2}$ for Mn^{2+}) which is assumed to be their ground electronic state. The ligands were chosen to represent all three metal-binding atoms which commonly occur in proteins (O, S, N) and to be of varied size and charge.

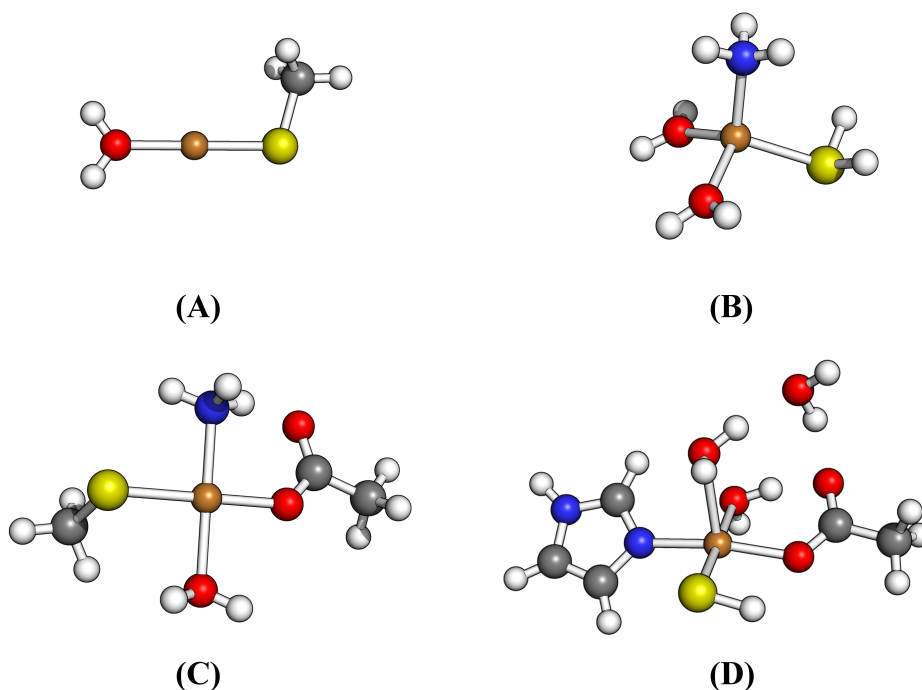


Figure 5.1: Model complexes:

- (A) $[\text{M}^{\text{II}}(\text{CH}_3\text{S})(\text{H}_2\text{O})]^+$ in linear coordination geometry;
- (B) $[\text{M}^{\text{II}}(\text{H}_2\text{O})_2(\text{H}_2\text{S})(\text{NH}_3)]^{2+}$ in tetrahedral coordination geometry;
- (C) $[\text{M}^{\text{II}}(\text{CH}_3\text{S})(\text{NH}_3)(\text{H}_2\text{O})(\text{CH}_3\text{COO})]$ in square planar coordination geometry;
- (D) $[\text{M}^{\text{II}}(\text{H}_2\text{O})_2(\text{SH})(\text{CH}_3\text{COO})(\text{Im})] \cdot \text{H}_2\text{O}$ in square pyramidal coordination geometry with one water in the second coordination shell.

5.1.1 Computational Details

Energies were calculated using the aug-cc-pVTZ basis set [63] for all atoms except hydrogen and the metal. For hydrogen the cc-pVTZ basis set [62] was used. In the case of the metals, aug-cc-pVTZ-PP [65] and the effective core potential ECP10MDF [64] were used for Cu and aug-cc-pVTZ [146] was used for Fe and Mn. In all local calculations density fitting was used with corresponding basis sets, except in the case of the metals and the JKFIT basis set, where def2-QZVPP/JKFIT was used [68]. All calculations were carried out with a development version of Molpro 2012.2 [61].

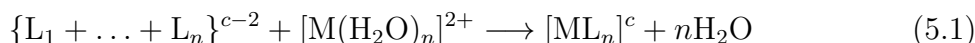
In local calculations orbitals were localized using the Pipek-Mezey localization [22]. Orbital domains were determined according to the NPA criteria [24], $T_{\text{NPA}} = 0.03$. A distance criteria was used for the classification of the orbital pairs. Strong pairs were

defined within a distance of 3 bohr and close pairs within a distance of 5 bohr.

The geometry optimizations were performed using B3LYP [48, 49] with def2-TZVP [72, 109] starting from the structures from Reference [145]. The L-M-L angles were constrained in all geometry optimizations to preserve the desired coordination geometry, except in the case of the SP complex. In the SP complex the starting structure was in an octahedral coordination geometry and the optimization did not converge when the L-M-L angles were constrained. In the case of the relaxed optimization, one water significantly moved away from the copper and at the end of the optimization was located in the second coordination shell. The rest of the complex slipped into a square pyramidal geometry.

5.1.2 Benchmark Results

The binding energies were calculated corresponding to the process:



where the $\{L_i\}^{c-2}$ notation is used to denote the overall charge of the (non-interacting) ligands. In almost all cases the number of ligated water molecules before the reaction was equal to the total number of ligands bound to the metal center after the reaction. It should be pointed out that in the case of the SP complex the starting complex was $[\text{M}^{\text{II}}(\text{H}_2\text{O})_6]^{2+}$ in octahedral coordination geometry and the SP complex was in a square pyramidal conformation with one water molecule in the second shell. One can define the binding energy ΔE_{bind} as:

$$\Delta E_{\text{bind}}(L_n) = E([\text{ML}_n]^c) + nE(\text{H}_2\text{O}) - E([\text{M}(\text{H}_2\text{O})_n]^{2+}) - \left(\sum_i E(L_i)\right)^{c-2} \quad (5.2)$$

Single point calculations were carried out on all reactant and product states using different levels of theory. Both RCCSD(T) and UCCSD(T) were used, but the results were similar so we will only show UCCSD(T) results, which simplifies the comparison to the results featured in Reference [145]. Since we could not obtain the UCCSD(T) results for the SP system we carried out LUCCSD(T0) calculations and evaluated the accuracy of this method. Furthermore, the LUCCSD(T0) method was used as a reference to evaluate the hybrid LMOMO method. In the LMOMO calculations only orbitals containing the metal center were treated at the coupled cluster level, and the rest of

the system at the LRMP2 level. In Figure 5.2 on the example of the copper complexes the interactions for all four systems are shown, including the representation of the used regions. The high level region is represented in red.

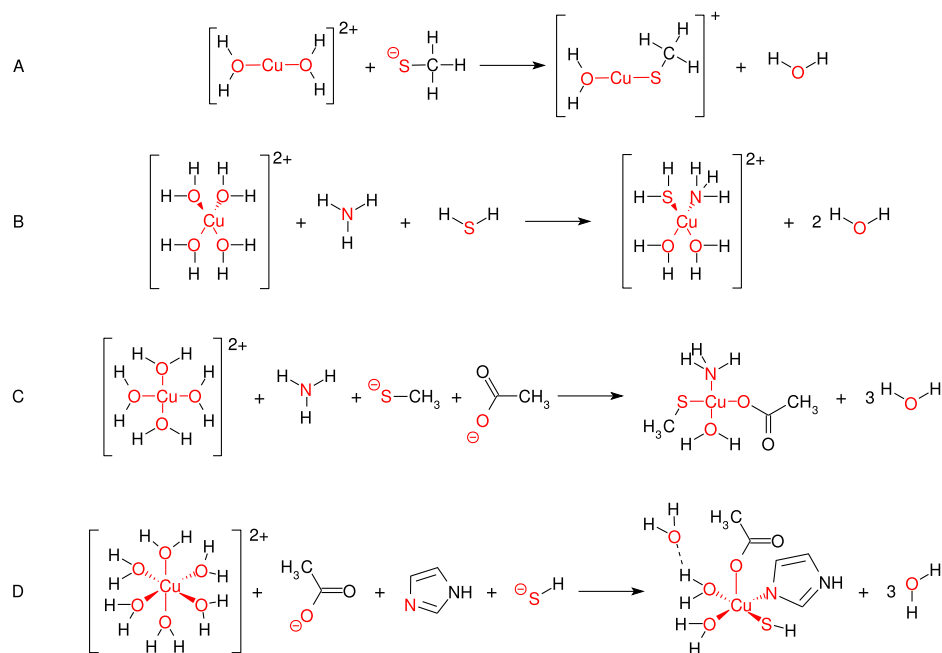


Figure 5.2: Reaction mechanism of the Cu complex in (A) linear, (B) tetrahedral, (C) square planar and (D) square pyramidal coordination geometry with one water molecule in the second coordination shell. Orbitals treated at the high level in LUCCSD(T0):LRMP2 calculations are depicted in red.

In the study of Gutten *et al.* [145] it was already shown that UCCSD(T) results in combination with the aug-cc-pVTZ basis set are reasonably converged, therefore we did not test the convergence of this method. On the other hand, we wanted to investigate the accuracy of LUCCSD(T0) using the same basis set. In the original study of the LUCCSD(T0) method, it was shown that the error of this method is mostly less than 1 kcal/mol, but can go up to 4 kcal/mol [147, 148]. Since the investigated systems in their study did not contain any metal centers our first step was to evaluate the accuracy of the LUCCSD(T0) method in comparison to UCCSD(T).

In Table 5.1 the results obtained using different wavefunction methods including canonical, local and hybrid LMOMO methods are shown. As one can see, the difference between canonical and local MP2 results is in most cases negligible. Only in the case of

systems in square planar coordination geometry the difference is slightly larger, but not more than 1.5 kcal/mol. Therefore, we can conclude that the domain approximation almost does not effect the results. On the other hand, the difference between local and canonical coupled cluster results is larger, although not in excess of 3 kcal/mol. This is in agreement with the results obtained in the original study [147, 148]. This difference can be due to the pair approximation. Since the domain approximation does not produce almost any error, there is no error cancellation in this case. It was shown in the previous work that error cancellation between the domain and pair approximation is important and can lead the local coupled cluster results to be in good agreement with canonical ones. Furthermore, one can notice that the difference is present in the systems with anionic ligands, since in those systems the charge is not as well localized as in systems with uncharged ligands. Therefore, the local correlation approximation should be taken as a good approximation if we take into the account that we are dealing with open shell species which are more demanding for the calculation and the degree of uncertainty is higher for these systems.

Before we look into the LMOMO results, we want to compare the LRMP2 and LUCCSD(T0) results. One can see that in almost all cases the difference between these two methods is less than a few kcal/mol, except for the linear complex of $[\text{Cu}^{\text{II}}(\text{CH}_3\text{S})(\text{H}_2\text{O})]^+$, where the difference is formidable, reaching 20 kcal/mol. The same difference was observed also in the case of the canonical calculations. Concerning the performance of the LMOMO method one can see that in all systems investigated in this study the difference between the LMOMO and LUCCSD(T0) results is more or less constant and does not amount to more than 1 kcal/mol. It was already seen in the previous study [145] that in these complexes DFT methods deviate more from the UCCSD(T) results dependent of the charge of the ligand. Therefore, we can conclude that the proposed LMOMO method performs significantly better than the DFT methods independent of the charge of the ligands.

Table 5.1: Interaction Energies of the Studied Metal Ions with Model Binding Sites, $\Delta E_{int}(L_n)$, (for the Reaction $(L_1 + \dots + L_n)^{c-2} + [M(H_2O)_n]^{2+} \rightarrow [ML_n]^c + nH_2O$), Calculated Using Various ab Initio Methods and $[M(H_2O)_n]^{2+}$ as References.

coord	complex	UCCSD(T)	LUCCSD(T0)	LMOMO	LRPM2	RMP2
LI	$[CuX_2]^{+(a)}$	-295.4	-296.3	-295.9	-318.3	-318.3
	$[FeX_2]^+$	-265.8	-263.3	-263.8	-259.7	-259.6
	$[MnX_2]^+$	-256.8	-255.2	-255.6	-252.9	-252.9
TH	$[CuX_4]^{2+(b)}$	-17.4	-17.5	-17.8	-15.4	-15.4
	$[FeX_4]^{2+}$	-6.6	-6.9	-7.5	-6.4	-6.4
	$[MnX_4]^{2+}$	-5.2	-5.5	-5.9	-5.0	-5.0
SQ	$[CuY_4]^{(c)}$	-377.9	-374.9	-375.5	-370.3	-371.7
	$[FeY_4]$	-367.9	-364.9	-365.2	-363.6	-364.8
	$[MnY_4]$	-354.0	-351.3	-351.6	-350.2	-351.5
SP	$[CuX_6]^{(d)}$	-	-348.7	-348.6	-344.1	-
	$[FeX_6]$	-	-325.0	-325.	-324.6	-
	$[MnX_6]$	-	-322.2	-322.7	-321.7	-

(a) $[MX_2]^+$ stands for the $[M^{II}(CH_3S)(H_2O)]^+$ complex.

(b) $[MX_4]$ stands for the $[M^{II}(CH_3S)(NH_3)(H_2O)(CH_3COO)]$ complex.

(c) $[MY_4]^{2+}$ stands for the $[M^{II}(H_2O)_2(H_2S)(NH_3)]^{2+}$ complex.

(d) $[MX_6]^+$ stands for the $[M^{II}(H_2O)_2(SH)(CH_3COO)(Im)] \cdot H_2O$ complex.

Since we observe that LRMP2 gives reasonable results for the investigated systems we wanted to compare the overall performance of the LRMP2 and LMOMO method. Therefore, we plotted normalized Gaussians in Figure 5.3 for both methods. The center of the Gaussian represents the average difference between LRMP2 or LUCCSD(T0):LRMP2 and LUCCSD(T0). The width represents the root mean square deviation (RMSD). As one can see both methods have an average deviation below 1 kcal/mol. However, the main disadvantage of LRMP2 is that the results can deviate much stronger as it is seen from the width of the Gaussian. This is not the case for the hybrid LMOMO method. These results show that our hybrid LMOMO method gives results which are in good agreement with the full calculations independent of the investigated system.

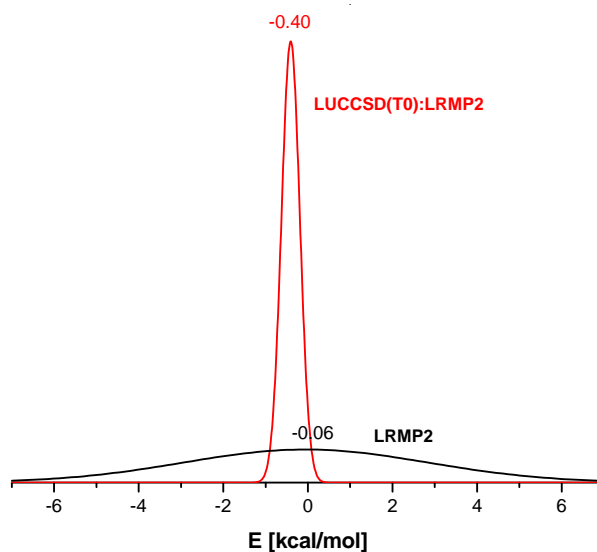
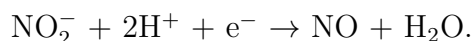


Figure 5.3: Normalized Gaussians centered at the average difference, the width represents the RMSD. LUCCSD(T0) was taken as a reference.

5.2 Nitrite Reductase

Nitrite reductases (NiRs) catalyze the one-electron reduction of nitrite (NO_2^-) into gaseous nitric oxide (NO) [149–151], which is the first step in microbial denitrification:



Two types of nitrite reductases (NiR) are known to exist. The first one contains a prosthetic cd_1 -heme iron in the active site. The second type makes use of copper atoms, copper nitrite reductase (CuNiR). Two groups of CuNiRs can be distinguished on the basis of their absorption spectra. They are blue and green CuNiRs. These two groups have a high degree of homology and structural similarity. Both enzymes have been determined to be 3-fold symmetrical homotrimers [152–154]. Also, in both enzymes two copper sites are present, one of type 1 (T1) and one of type 2 (T2) separated by ~ 13 Å [151]. In the T1 copper site, copper is coordinated to two histidine ligands, one cysteine and one methionine ligand. The T2 copper has three histidine ligands plus a number of water molecules (Figure 5.4) [155, 156]. The T1 copper site is the redox active site and the T2 copper site is the catalytic one.

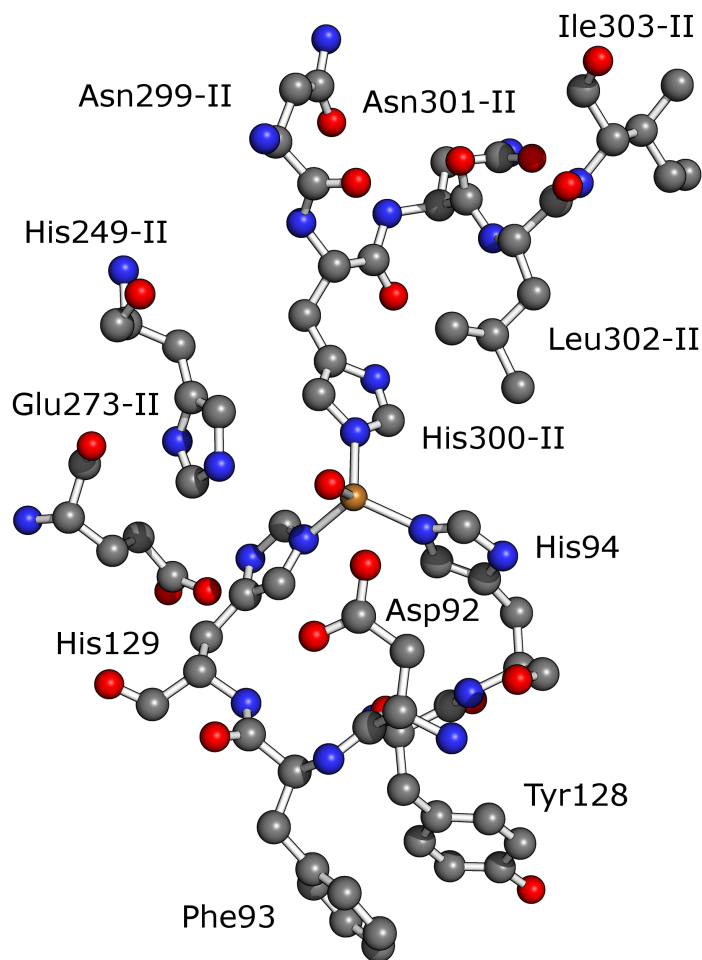


Figure 5.4: X-ray structure of the catalytically active T2-copper site in NiR from *Alcaligenes xylosoxidans* [157].

Based on the experimental evidence different groups proposed several mechanisms for the nitrite reduction [155, 156, 158–165]. Averill proposed the first mechanism for CuNiR based on the postulated mechanism of cd_1 -NiR [163]. In this mechanism nitrite binds to the reduced copper site. After that two consecutive proton transfers occur from nearby protein residues. In this way, the $Cu(I)-NO^+$ complex is formed. This complex is in resonance with the $Cu(II)-NO(\text{radical})$ form. After the loss of NO from the $Cu(II)-NO(\text{radical})$ the catalytic cycle is completed.

On the other hand, Suzuki and coworkers proposed an alternative mechanism based on the crystal structures of the NO_2^- and NO bound forms of different CuNiR enzymes [164]. In this case nitrite binds to the oxidized active site. After binding occurs the copper site

is reduced and a Cu(I)-NO₂⁻ complex is formed. A fast protonation of NO₂⁻ occurs and NO is quickly released and the oxidized Cu(II) active site with bound water is formed. This proposed mechanism was further changed by Hasnain and coworkers [165], since a different binding mode of NO₂⁻ was discovered.

Since all these mechanisms were lacking details, Marothy *et al.* [166] performed theoretical calculations on several possible mechanisms. The investigated mechanisms are shown in Figure 5.4. The initial state is the O state in which the T2 site is oxidized and Asp-92 is not protonated. In the first step, the nitrite binds to the copper center instead of the water (S state). After that the protonation of Asp-92 occurs (P1 state) which triggers the electron transfer from the T1 site resulting in a reduced, substrate-bound state with a single proton (R state). After the R state two pathways were investigated. They found that most likely a second proton enters the active site and binds to one of the nitrite group oxygen atoms, forming a nitrous acid group (P2 state). After this, the reaction occurs and the nitrite N-O bond is broken, forming nitric oxide. At the same time the T2-bound hydroxyl part is protonated from Asp-92 producing a water molecule. At the end nitric oxide is removed and one water is bound to the copper and another one is in the second shell and can be rearranged to return the enzyme to the initial state. Their study supported the mechanism proposed by Suzuki and coworkers [164].

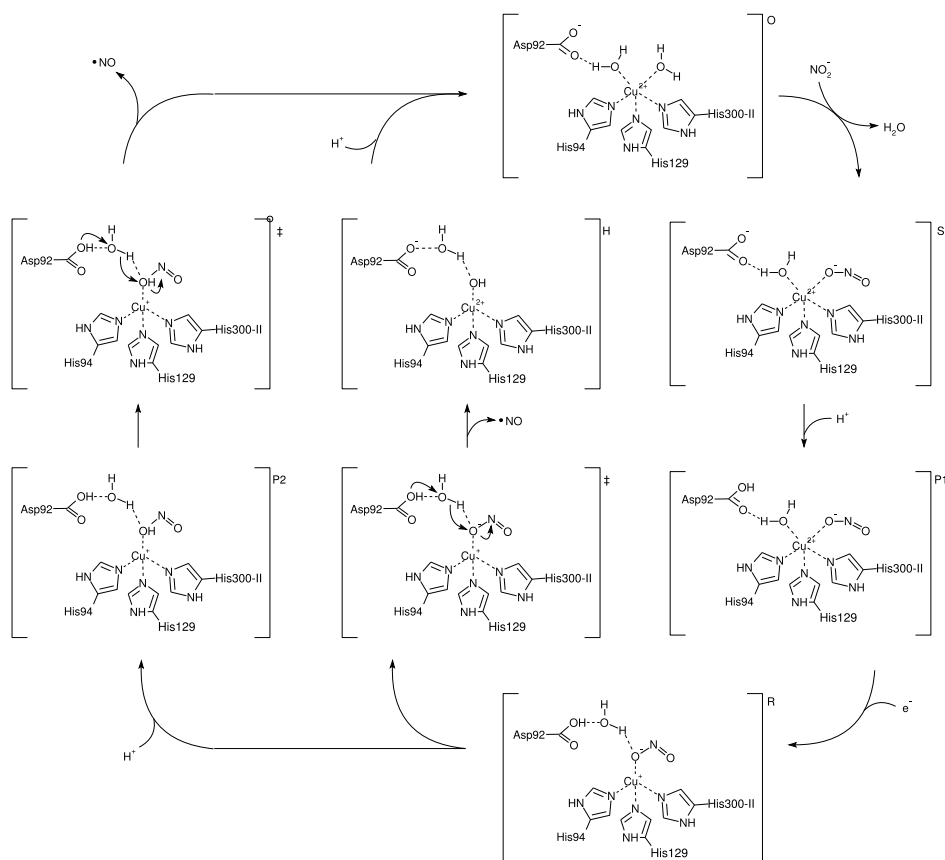


Figure 5.5: Putative reaction cycle of copper nitrite reductase.

In their study the order of the first three steps in the mechanism was based on the computed electron and proton affinities. The problem was that the calculated electron affinity of the P1 state was smaller than of the nitrite-free O state, while it should be larger. Therefore, we deemed this to be a good system to evaluate the accuracy of the LMOMO method. Since our focus is on open-shell systems we restricted our study to the first three steps in the proposed mechanism.

5.2.1 Computational Details

The active site was modelled starting from the crystal structure of the blue *Alcaligenes xylooxidans* NiR [157]. The model includes the T2 copper, which was coordinated to three histidines (His-94, His-129 and His-300-II) as well as Asp-92 in the second coordination shell. In some of the crystal structures several water molecules around the

Asp-92 are present [157, 167], so we included two of them in our model. All the ligands have been truncated at the peptide chain backbone.

The geometry optimization on this model complexes were done using the B3LYP-D3 method with the def2-SVP basis set [72, 112]. To speed up the calculations the RIJCOSX approximation was used in combination with the def2-SVP/JK auxiliary basis set [68]. In all optimizations the α -carbons were constrained to emulate the strain of the backbone, yet allowing a certain amount of flexibility. All geometry optimizations were carried out with the ORCA program package [168].

The Dunning cc-pVTZ basis set [62] was used for all atoms except Cu where the aug-cc-pVTZ-PP basis set [146] in combination with the ECP10MDF effective core potential [64] was used. The Pipek-Mezey localization scheme was used for orbitals [22]. Orbital domains were determined according to the NPA criteria [24] with $T_{\text{NPA}} = 0.03$. Orbital pairs were classified according to the distance criteria, strong pairs were defined within a distance of 3 bohr and close pairs within 5 bohr. Strong orbital pairs were fully included in the coupled cluster part and close pairs were treated exclusively at the MP2 level. The density fitting approximation was used throughout. Auxiliary basis sets were used for the Coulomb and exchange part as well as for the correlation part. In the first case def2-QZVPP/JKFIT [68] for Cu and for all other atoms cc-pVTZ/JKFIT [66] auxiliary basis sets were used. For the second part cc-pVTZ/MP2FIT [67] for all atoms except for Cu where aug-cc-pVTZ-PP/MP2FIT [169] basis sets were used. All single point energy calculations were carried out with a development version of Molpro 2012.2 [61].

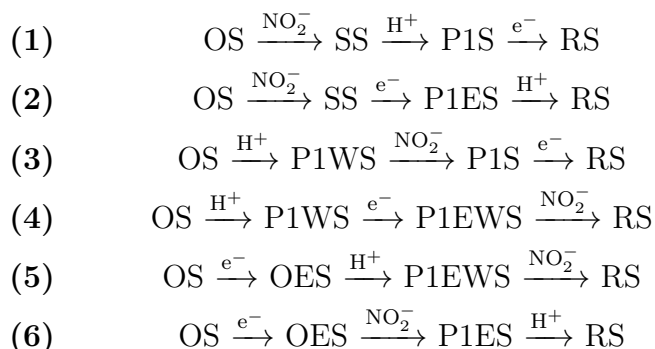
All shown energies present the free energies in the enzyme environment and were calculated as:

$$\Delta G_{tot} = \Delta E(\text{elec}) + \Delta G(\text{solv}, \epsilon) + \Delta G(\text{ZPE}) + \Delta G(\text{therm}), \quad (5.3)$$

where $\Delta E(\text{elec})$ is the electronic energy calculated with LUCCSD(T0) or LRMP2 or LUCCSD(T0):LRMP2, $\Delta G(\text{solv}, \epsilon)$ is the solvation energy (the difference between the COSMO calculations with dielectric constant of $\epsilon = 4$ and 1 obtained with the LRMP2 method), $\Delta G(\text{therm})$ is the thermal correction to the Gibbs free energy and $\Delta G(\text{ZPE})$ is the zero point energy (ZPE). The last two corrections were obtained from the frequency calculations using the B3LYP-D3 method with def2-SVP basis set.

5.2.2 Results and Discussion

Looking into the first three steps of the proposed mechanism one can see that three different processes occur, protonation, reduction and substrate binding. To thoroughly investigate the first three steps in the reaction mechanism all six possible combinations of these three processes will be considered:



The initial state is the same for all possibilities, the T2 copper is in its oxidized state and Asp-92 is deprotonated (OS). In **(1)** in the first step NO_2^- replaces the water molecule and binds to copper (SS). In the second step the protonation of Asp-92 takes place (P1S) and at the end the reduction occurs (RS). This is the same as proposed by Marothy *et al.* [166]. In **(2)** the first step is the same as in **(1)**, but protonation and reduction change their place leading to the P1ES and RS states, respectively. In **(3)** the protonation of Asp-92 occurs first (P1WS) followed by nitrite binding (P1S) and reduction (RS). In **(4)** the mechanism again starts with the protonation of Asp-92 as in the previous case, but the order of nitrite binding and reduction is changed which leads to formation of P1EWS and RS states respectively. In **(5)** and **(6)** the mechanism starts with the reduction of the active site (OES) which is followed by protonation (P1ES) and nitrite binding (RS) in **(5)** and nitrite binding (P1S) and protonation (RS) in **(6)**. All these states are shown in Figure 5.6 with the active site in the oxidized state and in Figure 5.7 with the active site in the reduced state.

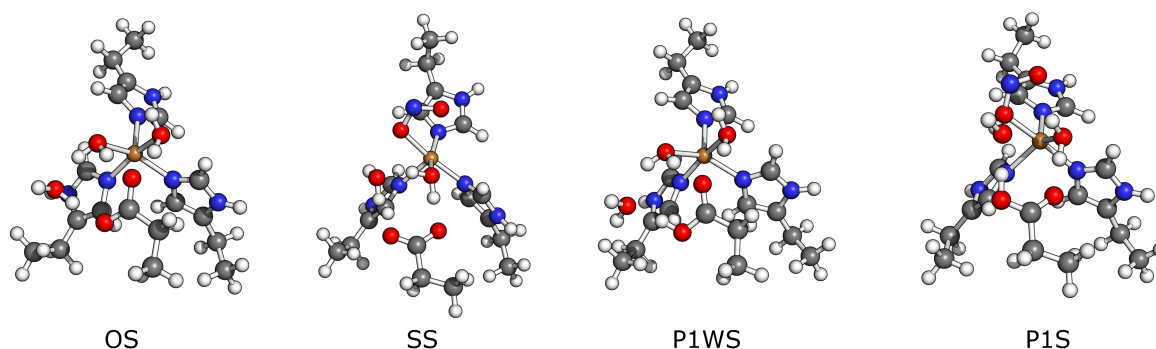


Figure 5.6: Optimized structures of the T2 site with copper in the oxidized state.

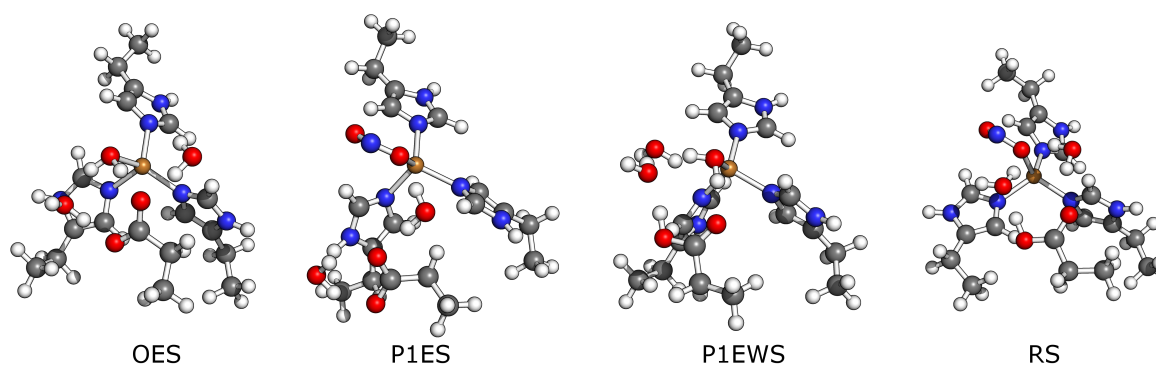


Figure 5.7: Optimized structures of the T2 site with copper in the reduced state.

Proton and Electron Affinities

For one to determine the protonation state of different intermediates, pK_a values should be calculated and compared with the proton donor. The pK_a is obtained from the relation:

$$pK_a = \frac{\Delta G_{solv}^{o,pK_a}}{RT \ln 10}, \quad (5.4)$$

where by $\Delta G_{solv}^{o,pK_a}$ is the free energy of deprotonation and $RT \ln 10 = 5.743$ kJ/mol (R is the gas constant and T is the temperature). The free energy of deprotonation of molecule HA is defined as the free energy change of the reaction:



Therefore, for the $\Delta G_{solv}^{o,pK_a}$ we have:

$$\Delta G_{solv}^{o,pK_a} = \Delta G_{solv}^o(A^+) + G_{solv}^o(H^+) - G_{solv}^o(HA) \quad (5.6)$$

where by $G_{solv}^o(A^+)$ and $G_{solv}^o(AH)$ are the energies in solution including zero-point energies and thermal corrections of deprotonated and protonated species. The $G_{solv}^o(H^+)$ energy is the solvation energy of H^+ including the translational free energy of H^+ and a correction for the change in reference state from 1 atm to 1 M. It amounts to -1119.4 kJ/mol ([170]).

Similarly, for one to determine at which point the reduction occurs the redox potential of different intermediates should be calculated. The redox potential is calculated using the equation:

$$E^o = -\frac{\Delta G_{solv}^{o,red}}{nF} \quad (5.7)$$

where by $\Delta G_{solv}^{o,red}$ is free energy of reduction, F is the Faraday constant and n is the number of electrons being transferred ($n = 1$ in our case). The free energy of reduction is defined as the free energy change of the reaction:



where Ox_{solv} and Red_{solv} are the oxidized and reduced species in solution. To calculate experimentally relevant redox potentials a reference electrode must be considered. One way to incorporate a reference electrode in the calculation of redox potentials is to use previously published experimental or theoretical values for the absolute reduction potential of the reference electrode. In this study we will use the value of 4.281 V obtained by Isse and coworkers [170]. Thus, we will now look into the pK_a values and redox potential to try to determine the order of the steps in which the reaction occurs.

In the OS state the pK_a has a negative value (-12.8) meaning that the protonation can not occur in this state. The pK_a value of the SS state is 5.1 and is slightly smaller than the one obtained for OES state (9.7). The largest pK_a value is obtained for the P1ES state and it amounts to 27.6. All these values are above the pK_a value of 3.9 for the β -carboxy group of aspartic acid [171]. The pK_a value for the β -carboxy group was taken as a reference since the protonation in our calculations occurs at this position. Therefore, four of the proposed mechanisms are possible, mechanisms (1), (2), (5) and

(6). Thus, if we look only into the proton affinities we cannot conclude through which of the proposed mechanisms the reaction proceeds.

If we now take a look in the redox potentials we can see that in the case of the SS state we have a negative value (-1.1 V) meaning that the reduction does not occur in this step. The values for the three other states are 0.06 V, 0.26 V and 1.40 V for OS, P1S and P1WS states, respectively. Since we know that electron transfer should occur from the T1 copper site we should compare these values with the redox potential of the T1 copper site which amounts to 247 mV [172]. Therefore we can conclude that only the mechanisms in which the reduction of P1S or P1WS states occurs are possible. This leave us with three possible mechanisms: (1), (3), (4).

Combing the results for pK_a values and redox potentials we see that the reaction could happen only through mechanism (1). The other mechanisms are rejected due to the too low proton or electron affinities. This is in the agreement with the experimental results for this system. It is known that the electron is transferred to the T2 copper site from the T1 copper site. It was observed by Solomon and coworkers [173] that there is no such transfer before the protonation of the T2 site. In an other experimental study it was indicated that the nitrite does not bind to the reduced Cu(I) T2 site [174]. Hence, we have that in the first step nitrite binds to the T2 copper, after that the protonation of the Asp-92 occurs and at the end the electron is transferred from the T1 copper. This is in agreement with the proposed mechanism of Marothy and coworkers [166].

In a recent theoretical study the mechanism of NiR was investigated [175]. Despite the use of a different active site the same conclusion was obtained. The electron transfer occurs after the nitrite binds. The difference between our study and the study from Li *et. al* is that in their case nitrite binds to the copper in a bidentate fashion, which is not the case in our study. The reason for this as they found out is that nitrite binds in a bidentate way only when both residues Asp-92 and His-249 are protonated. Since we did not include in our model the His-249 residue we could not obtain such a binding mode.

LMOMO calculations for NiR

After we knew through which mechanism the reaction occurs we used the LMOMO approach to calculate the energetic of the reaction pathway. Three different region selections were used. In the first one only the metal center was correlated at the high

level and it will be denoted as R1. In the second one the substrate was treated at the high level too (R2 selection) and in the third one the position where the proton binds was also treated at the high level (R3). All three selections are shown in Figure 5.8. The total number of valence orbitals in the LUCCSD(T0) calculation is reduced to almost one third, from 98 to 34 and 36 for R2 and R3, respectively. In the case of the R1 region only 14 orbitals are correlated at the coupled cluster level.

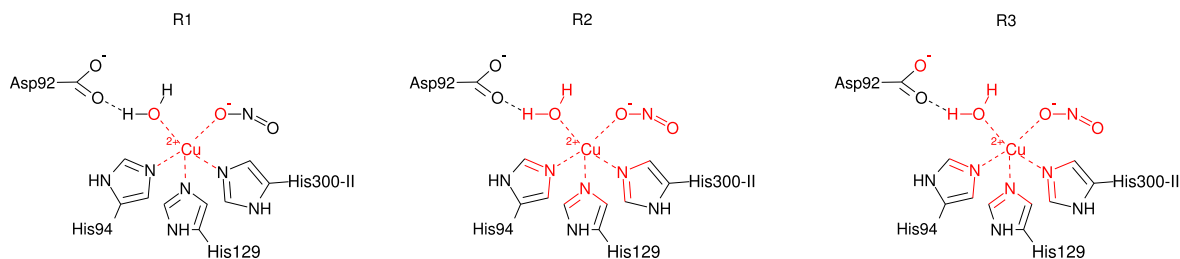


Figure 5.8: Selection of high level regions (represented in red) on the example of SS structure.

Table 5.2: Comparison of the LCCSD(T0):LRMP2 method with LRMP2 and LCCSD(T0) methods for the mechanism **(1)**.

	LRMP2	LUCCSD(T0):LRMP2			LUCCSD(T0)
		R1	R2	R3	
OS	0.00	0.00	0.00	0.00	0.00
SS	-37.1	-36.9	-38.9	-38.9	-38.7
P1S	-309.6	-308.6	-311.2	-311.7	-313.2
RS	-440.6	-415.5	-417.3	-417.4	-417.9

In Table 5.2 results for all three region selections as well as for LRMP2 and LUCCSD(T0) calculations are presented. One can see that in the SS and P1S states the difference between the LRMP2 and LUCCSD(T0) is significantly smaller than in the case of the RS state. This is what one would expect, since only in the last step the reduction occurs and the electronic configuration is drastically changed. Møller-Plesset then fails to describe this change. Therefore, this state is our best choice to test the LUCCSD(T0):LRMP2 results. As one can see the treatment of the copper center at the LUCCSD(T0) level is already enough to recover most of the energy obtained with the LUCCSD(T0) method. Increasing the size of the region one is improving even further the results and for the

largest region the difference between the LUCCSD(T0) and LUCCSD(T0):LMP2 results is below 1 kcal/mol in comparison with 22.7 kcal/mol in the case of LRMP2. In the case of other structures the difference is slightly larger, but not more than 1.5 kcal/mol and in all cases it is smaller than the difference in the case of the LRMP2 results.

Proton affinities were calculated using the LCCSD(T0) and LMP2 methods and also using the LMOMO scheme. The results are shown in Table 5.3. One can see that the difference between R1 and the full calculation is a few kcal/mol and it is slightly decreasing when the region is increased. At the end the largest error is less than 3 kcal/mol and it is comparable with the error obtained from LRMP2 results. Therefore, we can conclude that for the calculation of proton affinities we did not obtain any significant improvements treating only the metal center at the coupled cluster level since the LRMP2 results were already sufficient.

Table 5.3: Proton affinities calculated with LCCSD(T0):LMP2 in comparison with the LCCSD(T0) and LMP2 methods. All values are in kcal/mol.

	LRMP2	LUCCSD(T0):LRMP2			LUCCSD(T0)
		R1	R2	R3	
OS	248.0	247.1	248.0	248.6	250.0
SS	272.5	271.8	272.3	272.8	274.5
OES	279.4	278.6	279.6	280.3	280.9
P1ES	302.9	302.2	302.5	302.6	305.5

On the other hand, in the case of electron affinities the LRMP2 results deviate strongly from the CC results, up to 25 kcal/mol. DFT methods were also found unsuitable to calculate electron affinities accurately [166]. Therefore, we wanted to test the hybrid LMOMO method for the calculation of this property. As one can see in Table 5.4 the treatment of the metal at the LUCCSD(T0) level is already enough to obtain more accurate results. Increasing the size of the high level region one can reach an accuracy of about 3 kcal/mol. This was not the case neither for LRMP2, which strongly overestimated the results, nor with DFT which even did not obtain the correct order of the electron affinities [166]. Therefore, we can conclude that the hybrid LMOMO approach is capable of obtaining accurate results even in cases where most of the other available methods fail.

Table 5.4: Electron affinities calculated with LUCCSD(T0):LRMP2 in comparison with LUCCSD(T0) and LRMP2 methods. All values are in kcal/mol.

	LRMP2	LUCCSD(T0):LRMP2			LUCCSD(T0)
		R1	R2	R3	
OS	124.9	101.5	100.7	100.3	100.1
SS	100.6	76.5	75.9	76.0	73.8
P1WS	156.3	133.0	132.4	132.1	131.0
P1S	130.9	106.9	106.1	105.7	104.7

Since we proved one more time that the LMOMO method gives results which are in good agreement with LUCCSD(T0) results using the same basis set, we want to show that a significant saving of time is also present. In Figure 5.9 the timings for LUCCSD(T0):LRMP2 calculations are presented on the example of the OS structure. The time needed for the HF part is not included. As one can see, savings by a factor of 5 are obtained for the smallest region selection and 2.5 times for the two larger selections (R2 and R3). It should be also noticed that the difference in time needed for the R2 and R3 selections is ≈ 5 minutes, since the number of orbitals going from R2 to R3 changes only by 2.

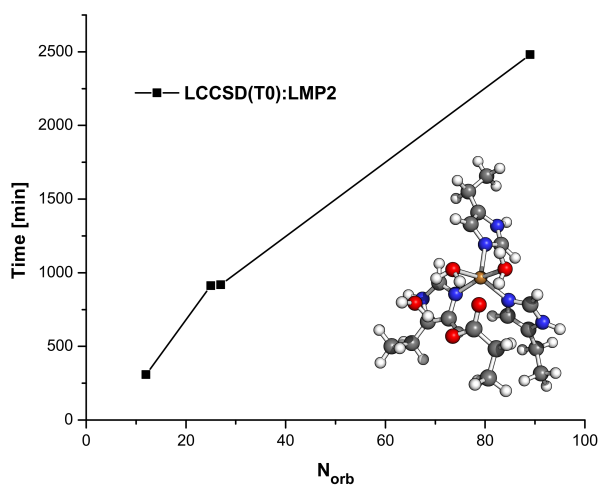


Figure 5.9: Time needed for LCCSD(T0):LMP2 calculations dependent on the number of orbitals in high level region.

5.3 Final Remarks

In this chapter the LMOMO method was applied to open-shell systems for the first time. In the first part of the chapter the accuracy of the LUCCSD(T0) method as well as the LMOMO method was evaluated. The UCCSD(T) method was used as a reference. It was observed that the LUCCSD(T0) results deviate from the UCCSD(T) results in the case of anionic ligands, but not more than 3 kcal/mol. On the other hand, the LMOMO results deviated only up to 1 kcal/mol from LUCCSD(T0) results independent of the calculated system.

After satisfactory accuracy of the investigated methods was obtained, we applied the LUCCSD(T0) and LMOMO methods for the investigation of the first three steps in the reaction mechanism of NiR. Starting from the six possible mechanisms we calculated proton and electron affinities for all states using the LUCCSD(T0) method. We obtained that the most favourable mechanism is the first one. In that mechanism, copper is in its oxidized state and Asp-92 is not protonated in the initial state (OS). Already in the first step, the nitrite binds to the copper center (SS). After nitrite binds, the protonation of Asp-92 occurs (P1S). This process is the trigger for the electron transfer from the T1 copper site to the T2 copper site (RS). This mechanism is in agreement with previous theoretical [166, 175] and experimental work [164, 173].

Afterwards, the LMOMO method was used to calculate the energies for these states (OS, SS, P1S and RS). A good agreement between the LUCCSD(T0) and LMOMO results was obtained even for the smallest region selection. The LMOMO method was also used to calculate proton and electron affinities. In the case of proton affinities already at the LRMP2 level accurate results are obtained, therefore in that case there is no real need for the hybrid LMOMO scheme. In any case, the LMOMO results were also of the same accuracy as the LRMP2 results. On the other hand, in the case of electron affinities, the situation is drastically different. The LRMP2 method significantly overestimates the results. In all states the difference was around 25 kcal/mol. On the contrary, the LMOMO results were in good agreement with the LUCCSD(T0) results. For the largest region selection the errors were below 3 kcal/mol.

CHAPTER 6

Conclusions

The purpose of this thesis was to expand the application range of the LMOMO method to metal-containing systems and open-shell molecules in general. This opens the door to studies in catalysis, particularly metallobiosites, where the accuracy of QM methods is an essential prerequisite. The local correlation approach allows one to combine in a single calculation different levels of theory. In this way, one can correlate at a high level of theory parts of the system which are chemically relevant and at the low level the remaining system.

The first investigated systems were gold dimer complexes. It has long been postulated that such aggregates are primarily defined through the Au \cdots Au interaction, the aurophilic effect. The treatment of only the gold atoms at the high level of theory showed that this is not the case. This led us to the conclusion that the ligands play an important role in these complexes and that the gold interaction alone has a less significant effect in systems previously labeled as aurophilic than what was thought. It was found that the ligands dominate the interaction in bulky complexes. By observing systems which were experimentally characterized it was found that dipole-dipole and higher order electrostatic contributions were also responsible for the formation of crystals. On the other hand, when Au \cdots Au interactions from different complexes were compared, an essentially unchanged interaction was observed along a vast range of ligands. By comparing the Au \cdots Au and Ag \cdots Ag interaction it was observed that the former is stronger. Since in the latter ligands play an even larger effect, stronger interactions and shorter distances were observed in these complexes and therefore it was wrongly believed that would be more metallophilic. In the end, the results warrant a reevaluation of the concept of aurophilicity. Aurophilicity should be taken as the direct Au \cdots Au interaction only which may contribute to the stabilization of said complexes, but it is not necessarily the driving force for their formation.

Continuing our research we turned to the study of molybdenum enzymes and their reaction mechanisms. A debate about the preferred mechanism persists in the community. In that regard, three different mechanisms were investigated, the S \rightarrow OMo, the O \rightarrow Mo and S \rightarrow Mo mechanisms. At the end of our investigation, the S \rightarrow OMo mechanism was found as the most likely one. Our conclusions were drawn from the comparison of barrier heights computed with consistent model systems and levels of theory. Although the latter values were too high, this effect is clearly due to the reduced model size used, whereby the active site is charged and the reaction slowed due to the Coulomb repulsion to the substrate. One way to overcome this problem is to include in our model system

one of the surrounding positively charged residues. In the second part of this chapter the LMOMO scheme was applied for the calculation of the reaction pathways for two molybdenum enzymes, SO and DMSOR. It was found that the processes which occur along the reaction path are relatively local and therefore suitable to be treated with this hybrid scheme. The results obtained in this part were in excellent agreement with the results obtained from our reference values. This is highly relevant, since other available methods failed to obtain results of such accuracy.

After series of studies on closed-shell systems we expanded the applicability of LMOMO to open-shell systems. This is of particular interest, since open-shell systems are commonly found in enzymes and proteins. It was shown that the localization leads to slightly larger errors than in the case of closed-shell systems, since the electron density is not so well localized. However, the obtained accuracy was still satisfactory compared to other available methods. In the case of the proposed LMOMO scheme, the results did not deviate strongly from the high level local method.

At the end of this thesis, the reaction mechanism of the copper nitrite reductase was investigated. The order of the first three steps was studied by calculating the proton and electron affinities. It was found that the reaction starts with the nitrite binding, followed by protonation which then triggers the electron transfer, leading to the reduction of the copper center. Probably, the most important finding in this part of the thesis is that the proposed hybrid scheme is capable of calculating the electron affinities with the required accuracy. This was a particularly demanding test since the electron correlation picture is changed in this process and most of the proposed methods are not capable to describe this properly.

At the end, I would like to summarize a few important points which can be taken as a take-home message. The first and the most important one is that it was shown that LMOMO can obtain the results of the same accuracy as LCCSD(T0) for a wide range of both closed- and open-shell systems. The other important information is how one can choose the high level region. It was shown that the treatment of the metal center and the closest neighbours is enough to properly describe the systems investigated in this thesis. Last but not least, we were able to confirm the computational savings brought forth by the hybrid scheme. I hope that this method will find its way into the chemistry community for the calculation of reaction mechanisms and other properties where high accuracy is necessary beyond commonly used DFT.

Bibliography

-
- [1] Helgaker, T.; Jorgensen, P.; Olsen, J. , 1st ed.; Wiley: Chichester ; New York, 2013.
- [2] Szabo, A.; Ostlund, N. S. , new edition ed.; Dover Publications: Mineola, N.Y, 1996.
- [3] Shavitt, I.; Bartlett, R. J. , 1st ed.; Cambridge University Press: Cambridge ; New York, 2009.
- [4] Jensen, F. , 2nd ed.; Wiley: Chichester, England ; Hoboken, NJ, 2007.
- [5] Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- [6] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–9802.
- [7] Urban, M.; Noga, J.; Cole, S. J.; Bartlett, R. J. *J. Chem. Phys.* **1985**, *83*, 4041–4046.
- [8] Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.
- [9] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- [10] Saebø, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, *113*, 13–18.
- [11] Ayala, P. Y.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 3660–3671.
- [12] Scuseria, G. E.; Ayala, P. Y. *J. Chem. Phys.* **1999**, *111*, 8330–8343.
- [13] Maslen, P. E.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 7093–7099.
- [14] Maslen, P. E.; Head-Gordon, M. *Chem. Phys. Lett.* **1998**, *283*, 102–108.
- [15] Flocke, N.; Bartlett, R. J. *J. Chem. Phys.* **2004**, *121*, 10935–10944.
- [16] Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151–154.
- [17] Pulay, P.; Saebø, S. *Theor. chim. acta* **1986**, *69*, 357–368.
- [18] Saebø, S.; Pulay, P. *J. Chem. Phys.* **1987**, *86*, 914–922.

- [19] Saebo, S.; Pulay, P. *J. Chem. Phys.* **1988**, *88*, 1884–1890.
- [20] Saebø, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.
- [21] Boys, S. F. In *Quantum Theory of Atoms, Molecules, and the Solid State, A Tribute to John C. Slater*; Academic Press, 1966; p 253.
- [22] Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- [23] Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457–464.
- [24] Mata, R. A.; Werner, H.-J. *Mol. Phys.* **2007**, *105*, 2753–2761.
- [25] Neese, F.; Hansen, A.; Liakos, D. G. *J. Chem. Phys.* **2009**, *131*, 064103.
- [26] Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2012**, *136*, 144105.
- [27] Boughton, J. W.; Pulay, P. *J. Comput. Chem.* **1993**, *14*, 736–740.
- [28] Knowles, P.; Schütz, M.; Werner, H.-J. In ; Grotendorst, J., Ed.; John von Neumann Institute for Computing: Jülich, 2000; pp 97–179.
- [29] Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H.-J. *J. Chem. Phys.* **2000**, *113*, 9443–9455.
- [30] Azhary, A. E.; Rauhut, G.; Pulay, P.; Werner, H.-J. *J. Chem. Phys.* **1998**, *108*, 5185–5193.
- [31] Schütz, M.; Werner, H.-J.; Lindh, R.; Manby, F. R. *J. Chem. Phys.* **2004**, *121*, 737–750.
- [32] Werner, H.-J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149–8160.
- [33] Schütz, M.; Rauhut, G.; Werner, H.-J. *J. Phys. Chem. A* **1998**, *102*, 5997–6003.
- [34] Pulay, P.; Saebo, S.; Meyer, W. *J. Chem. Phys.* **1984**, *81*, 1901–1905.
- [35] Hampel, C.; Peterson, K. A.; Werner, H.-J. *Chem. Phys. Lett.* **1992**, *190*, 1–12.
- [36] Hampel, C.; Werner, H.-J. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

- [37] Schütz, M. *J. Chem. Phys.* **2000**, *113*, 9986–10001.
- [38] Schütz, M.; Werner, H.-J. *Chem. Phys. Lett.* **2000**, *318*, 370–378.
- [39] Mata, R. A.; Werner, H.-J.; Schütz, M. *J. Chem. Phys.* **2008**, *128*, 144106.
- [40] Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959–1967.
- [41] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- [42] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [43] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [44] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.
- [45] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [46] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- [47] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [48] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- [49] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [50] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.
- [51] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- [52] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [53] Grimme, S. *Chem. – Eur. J.* **2012**, *18*, 9955–9964.
- [54] Jansen, M. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 1098–1110.
- [55] Schmidbaur, H.; Graf, W.; Müller, G. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 417–419.
- [56] Schmidbaur, H. *Gold Bulletin* **1990**, *23*, 11–21.
- [57] Pyykkö, P.; Zhao, Y. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 604–605.

- [58] Muñiz, J.; Wang, C.; Pyykkö, P. *Chem. Eur. J.* **2011**, *17*, 368–377.
- [59] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
- [60] Werner, H.-J.; Schütz, M. *J. Chem. Phys.* **2011**, *135*, 144116.
- [61] Werner, H.-J.; et al.; *MOLPRO, version 2012.1, a package of ab initio programs*; 2012; see www.molpro.net.
- [62] Jr, T. H. D. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [63] Kendall, R. A.; Jr, T. H. D.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- [64] Figen, D.; Rauhut, G.; Dolg, M.; Stoll, H. *Chem. Phys.* **2005**, *311*, 227–244.
- [65] Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283–296.
- [66] Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
- [67] Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.
- [68] Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.
- [69] Yousaf, K. E.; Peterson, K. A. *J. Chem. Phys.* **2008**, *129*, 184108.
- [70] Hill, J. G.; Mazumder, S.; Peterson, K. A. *J. Chem. Phys.* **2010**, *132*, 054108.
- [71] Hill, J. G.; Peterson, K. A. *J. Chem. Theory Comput.* **2012**, *8*, 518–526.
- [72] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [73] Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuß, H. *Theor. Chim. Acta* **1990**, *77*, 123–141.
- [74] Wuttke, A.; Master's thesis; Georg-August-Universität Göttingen; 2014.
- [75] Magnko, L.; Schweizer, M.; Rauhut, G.; Schütz, M.; Stoll, H.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1006–1013.
- [76] O'Grady, E.; Kaltsoyannis, N. *Phys. Chem. Chem. Phys.* **2004**, *6*, 680–687.
- [77] Pyykkö, P.; Zaleski-Ejgierd, P. *J. Chem. Phys.* **2008**, *128*, 124309.

- [78] Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2011**, *135*, 144117.
- [79] Runeberg, N.; Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **1999**, *110*, 7210–7215.
- [80] Pyykkö, P. *Chem. Rev.* **1997**, *97*, 597–636.
- [81] Lee, K.; Murray, D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.
- [82] Vydrov, O. A.; Voorhis, T. V. *J. Chem. Phys.* **2010**, *133*, 244103.
- [83] Grimme, S.; Ehrlich, S. *personal communication* **2013**.
- [84] Ray, L.; Shaikh, M. M.; Ghosh, P. *Inorg. Chem.* **2008**, *47*, 230–240.
- [85] Baker, M. V.; Barnard, P. J.; Berners-Price, S. J.; Brayshaw, S. K.; Hickey, J. L.; Skelton, B. W.; White, A. H. *J. Organomet. Chem.* **2005**, *690*, 5625–5635.
- [86] de Frémont, P.; Scott, N. M.; Stevens, E. D.; Nolan, S. P. *Organometallics* **2005**, *24*, 2411–2418.
- [87] de Frémont, P.; Scott, N. M.; Stevens, E. D.; Ramnial, T.; Lightbody, O. C.; Macdonald, C. L. B.; Clyburne, J. A. C.; Abernethy, C. D.; Nolan, S. P. *Organometallics* **2005**, *24*, 6301–6309.
- [88] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- [89] Wang, H. M. J.; Chen, C. Y. L.; Lin, I. J. B. *Organometallics* **1999**, *18*, 1216–1223.
- [90] Weigend, F.; Baldes, A. *J. Chem. Phys.* **2010**, *133*, 174102.
- [91] Hille, R. *Chem. Rev.* **1996**, *96*, 2757–2816.
- [92] Mendel, R. R. *J. Exp. Bot.* **2007**, *58*, 2289–2296.
- [93] Feng, C.; Tollin, G.; Enemark, J. H. *Biochim. Biophys. Acta (BBA)* **2007**, *1774*, 527–539.
- [94] Severen, M.-C. v.; Andrejić, M.; Li, J.; Starke, K.; Mata, R. A.; Nordlander, E.; Ryde, U. *J. Biol. Inorg. Chem.* **2014**, 1–15.
- [95] Wilson, H. L.; Rajagopalan, K. V. *J. Biol. Chem.* **2004**, *279*, 15105–15113.

- [96] Thapper, A.; Lorber, C.; Fryxelius, J.; Behrens, A.; Nordlander, E. *J. Inorg. Biochem.* **2000**, *79*, 67–74.
- [97] Chaudhury, P. K.; Das, S. K.; Sarkar, S. *Biochem. J.* **1996**, *319*, 953–959.
- [98] Johnson, J. L.; Rajagopalan, K. V. *J. Clin. Invest.* **1976**, *58*, 543–550.
- [99] Brody, M. S.; Hille, R. *Biochem.* **1999**, *38*, 6668–6677.
- [100] Hille, R. *J. Biol. Inorg. Chem.* **1997**, *2*, 804–809.
- [101] Hille, R. *Biochim. Biophys. Acta (BBA) - Bioenergetics* **1994**, *1184*, 143–169.
- [102] Hernandez-Marin, E.; Ziegler, T. *Inorg. Chem.* **2009**, *48*, 1323–1333.
- [103] Thapper, A.; Deeth, R. J.; Nordlander, E. *Inorg. Chem.* **1999**, *38*, 1015–1018.
- [104] Pal, K.; Chaudhury, P.; Sarkar, S. *Chem. Asian J.* **2007**, *2*, 956–964.
- [105] Doonan, C. J.; Wilson, H. L.; Bennett, B.; Prince, R. C.; Rajagopalan, K. V.; George, G. N. *Inorg. Chem.* **2008**, *47*, 2033–2038.
- [106] Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691–5705.
- [107] Werner, H.-J.; Pflüger, K. In *Annu. Rep. Comput. Chem.*; Spellmeyer, D. C., Ed.; Elsevier, 2006; Vol. 2, pp 53–80.
- [108] Peterson, K. A.; Figgen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.
- [109] Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143–152.
- [110] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- [111] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–290.
- [112] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

- [113] Dieterich, J. M.; Werner, H.-J.; Mata, R. A.; Metz, S.; Thiel, W. *J. Chem. Phys.* **2010**, *132*, 035101.
- [114] Mata, R. A.; Werner, H.-J. *J. Chem. Phys.* **2006**, *125*, 184110.
- [115] Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- [116] Peariso, K.; McNaughton, R. L.; Kirk, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9006–9007.
- [117] Ilich, P.; Hille, R. *J. Phys. Chem. B* **1999**, *103*, 5406–5412.
- [118] Ilich, P.; Hille, R. *J. Am. Chem. Soc.* **2002**, *124*, 6796–6797.
- [119] Martin, J. M. L.; Taylor, P. R. *J. Chem. Phys.* **1996**, *100*, 6047–6056.
- [120] Jiang, W.; DeYonker, N. J.; Wilson, A. K. *J. Chem. Theory Comput.* **2012**, *8*, 460–468.
- [121] Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729–738.
- [122] Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651.
- [123] Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301–332.
- [124] Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144–1149.
- [125] Schutz, C. N.; Warshel, A. *Proteins: Struct., Funct., and Bioinf.* **2001**, *44*, 400–417.
- [126] Brody, M. S.; Hille, R. *Biochim. Biophys. Acta (BBA) - Protein Structure and Molecular Enzymology* **1995**, *1253*, 133–135.
- [127] Polly, R.; Werner, H.-J.; Manby, F. R.; Knowles, P. J. *Mol. Phys.* **2004**, *102*, 2311–2321.
- [128] Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. *Theor. Chem. Acc.* **2007**, *117*, 587–597.
- [129] Li, J.-L.; Mata, R. A.; Ryde, U. *J. Chem. Theory Comput.* **2013**, *9*, 1799–1807.

- [130] Kisker, C.; Schindelin, H.; Pacheco, A.; Wehbi, W. A.; Garrett, R. M.; Rajagopalan, K. V.; Enemark, J. H.; Rees, D. C. *Cell* **1997**, *91*, 973–983.
- [131] Garrett, R. M.; Johnson, J. L.; Graf, T. N.; Feigenbaum, A.; Rajagopalan, K. V. *Proc. Natl. Acad. Sci.* **1998**, *95*, 6394–6398.
- [132] Lee, T. J.; Taylor, P. R. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.
- [133] Webster, C. E.; Hall, M. B. *J. Am. Chem. Soc.* **2001**, *123*, 5820–5821.
- [134] Thapper, A.; Deeth, R. J.; Nordlander, E. *Inorg. Chem.* **2002**, *41*, 6695–6702.
- [135] Lim, B. S.; Sung, K.-M.; Holm, R. H. *J. Am. Chem. Soc.* **2000**, *122*, 7410–7411.
- [136] Mohr, M.; McNamara, J. P.; Wang, H.; Rajeev, S. A.; Ge, J.; Morgado, C. A.; Hillier, I. H. *Faraday Discuss.* **2003**, *124*, 413–428.
- [137] McNamara, J. P.; Hillier, I. H.; Bhachu, T. S.; Garner, C. D. *Dalton Trans.* **2005**, 3572–3579.
- [138] McNamara, J. P.; Joule, J. A.; Hillier, I. H.; Garner, C. D. *Chem. Commun.* **2005**, 177–179.
- [139] Hernandez-Marin, E.; Ziegler, T. *Can. J. Chem.* **2010**, *88*, 683–693.
- [140] Tenderholt, A. L.; Wang, J.-J.; Szilagyi, R. K.; Holm, R. H.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J. Am. Chem. Soc.* **2010**, *132*, 8359–8371.
- [141] Tenderholt, A. L.; Hodgson, K. O.; Hedman, B.; Holm, R. H.; Solomon, E. I. *Inorg. Chem.* **2012**, *51*, 3436–3442.
- [142] Sung, K.-M.; Holm, R. H. *J. Am. Chem. Soc.* **2002**, *124*, 4312–4320.
- [143] Hofmann, M. *J. Mol. Struct. (THEOCHEM)* **2006**, *773*, 59–70.
- [144] Hofmann, M. *Inorg. Chem.* **2008**, *47*, 5546–5548.
- [145] Gutten, O.; Beššeová, I.; Rulíšek, L. *J. Phys. Chem. A* **2011**, *115*, 11394–11402.
- [146] Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.

- [147] Liu, Y.; Ph.D. thesis; Institut für Theoretische Chemie der Universität Stuttgart; 2011.
- [148] Liu, Y.; Werner, H.-J. *unpublished*.
- [149] Averill, B. A. *Chem. Rev.* **1996**, *96*, 2951–2964.
- [150] Moura, I.; Moura, J. J. G. *Curr. Opin. Chem. Biol.* **2001**, *5*, 168–175.
- [151] Wasser, I. M.; de Vries, S.; Moënne-Loccoz, P.; Schröder, I.; Karlin, K. D. *Chem. Rev.* **2002**, *102*, 1201–1234.
- [152] Libby, E.; Averill, B. A. *Biochem. Biophys. Res. Commun.* **1992**, *187*, 1529–1535.
- [153] Abraham, Z. H.; Lowe, D. J.; Smith, B. E. *Biochem. J.* **1993**, *295*, 587–593.
- [154] Kukimoto, M.; Nishiyama, M.; Murphy, M. E. P.; Turley, S.; Adman, E. T.; Horinouchi, S.; Beppu, T. *Biochemistry* **1994**, *33*, 5246–5252.
- [155] Murphy, M. E. P.; Turley, S.; Adman, E. T. *J. Biol. Chem.* **1997**, *272*, 28455–28460.
- [156] Boulanger, M. J.; Murphy, M. E. P. *Biochemistry* **2001**, *40*, 9132–9141.
- [157] Ellis, M. J.; Dodd, F. E.; Sawers, G.; Eady, R. R.; Hasnain, S. S. *J. Mol. Biol.* **2003**, *328*, 429–438.
- [158] Suzuki, S.; Kataoka, K.; Yamaguchi, K. *Acc. Chem. Res.* **2000**, *33*, 728–735.
- [159] Boulanger, M. J.; Kukimoto, M.; Nishiyama, M.; Horinouchi, S.; Murphy, M. E. P. *J. Biol. Chem.* **2000**, *275*, 23957–23964.
- [160] Adman, E. T.; Godden, J. W.; Turley, S. *J. Biol. Chem.* **1995**, *270*, 27458–27474.
- [161] Pinho, D.; Besson, S.; Brondino, C. D.; de Castro, B.; Moura, I. *Eur. J. Biochem.* **2004**, *271*, 2361–2369.
- [162] Fittipaldi, M.; Wijma, H. J.; Verbeet, M. P.; Canters, G. W.; Groenen, E. J. J.; Huber, M. *Biochemistry* **2005**, *44*, 15193–15202.
- [163] Averill, B. A. *Angew Chem. Int. Ed.* **1994**, *33*, 2057–2058.

-
- [164] Kataoka, K.; Furusawa, H.; Takagi, K.; Yamaguchi, K.; Suzuki, S. *J. Biochem.* **2000**, *127*, 345–350.
- [165] Antonyuk, S. V.; Strange, R. W.; Sawers, G.; Eady, R. R.; Hasnain, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12041–12046.
- [166] De Marothy, S. A.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Comput. Chem.* **2007**, *28*, 528–539.
- [167] Boulanger, M. J.; Murphy, M. E. *Protein Sci.* **2003**, *12*, 248–256.
- [168] Neese, F.; et al.; *ORCA, version 3.0.2, An Ab Initio, DFT and Semiempirical electronic structure package*; 2012.
- [169] Hill, J. G.; Platts, J. A. *J. Chem. Phys.* **2008**, *128*, 044104.
- [170] Isse, A. A.; Gennaro, A. *The Journal of Physical Chemistry B* **2010**, *114*, 7894–7899.
- [171] Lide, D. R. , 85th ed.; CRC Press: Boca Raton, 2004.
- [172] Olesen, K.; Veselov, A.; Zhao, Y.; Wang, Y.; Danner, B.; Scholes, C. P.; Shapleigh, J. P. *Biochemistry* **1998**, *37*, 6086–6094.
- [173] Ghosh, S.; Dey, A.; Sun, Y.; Scholes, C. P.; Solomon, E. I. *J. Am. Chem. Soc.* **2008**, *131*, 277–288.
- [174] Strange, R. W.; Murphy, L. M.; Dodd, F. E.; Abraham, Z. H. L.; Eady, R. R.; Smith, B. E.; Hasnain, S. S. *J. Mol. Biol.* **1999**, *287*, 1001–1009.
- [175] Li, Y.; Hodak, M.; Bernholc, J. *Biochemistry* **2015**.

Curriculum Vitae

Name:	Milica Andrejić
Address:	Ulmenweg 2b 37077 Göttingen Germany
E-Mail:	mandrej@gwdg.de
Place of birth:	Niš, Serbia
Date of birth:	27.03.1987
Nationality:	Serbian

Education

Mar 2012 - today	PhD student in Chemistry under the supervision of Jun.-Prof. Dr. Ricardo Mata, Institute of Physical Chemistry, Georg-August-Universität Göttingen, Germany.
Oct 2011 - Jan 2012	Master of Science in Physics at Faculty of Physics, University of Belgrade, Serbia. Graduated with 9.67/10.
Oct 2006 - Oct 2011	Bachelor of Science in Physics at Faculty of Physics, University of Belgrade, Serbia. Graduated with 8.88/10.

Research Stays

May 2014	Research stay in the group of Prof. Dr. Ulf Ryde, Lund University.
May 2013 - June 2013	Research stay in the group of Prof. Dr. Ulf Ryde, Lund University.

Scholarships

Mar 2012 - Apr 2015	Doctoral Fellowship in IRTG1422 - Metal Sites in Biomolecules: Structures, Regulation and Mechanisms
Oct 2007 - Sept 2010	Student Scholarship from the Ministry of Science and Education, Republic of Serbia

Teaching Experience

Mar 2012 - today	Assistant for "PC Experimentieren: Spektroskopie" (WS14/15) Assistant for "Advanced Molecular Modeling" (SS14) Co-supervised 1 Master student (WS13/14) Student Assistant for "Mathematik für Chemiker I" (WS12/13, WS13/14)
-----------------------------	---

Oral Presentations

June 2014	<i>Application of Local Coupled-Cluster Methods to Enzymatic Reactivity</i> Contributed talk at the FemEx-Oslo, Oslo, Norway
May 2014	<i>Hybrid QM/QM Studies of Molybdenum Enzymes - Reaction Pathways and Redox Potentials</i> Talk at the IRTG workshop, Lund, Sweden
Dec 2013	<i>Accurate Calculations of Reaction Mechanisms in Bioinorganic Complexes</i> Talk at the "Weihnachtskolloquium" of the Faculty of Chemistry, Göttingen, Germany
May 2013	<i>Hybrid QM/QM Open-Shell Local Correlation Methods for the Study of Metal Ions with Biomolecular Ligands</i> Talk at the IRTG workshop, Homberg (Hülsa), Germany

Poster Presentations

- | | |
|------------------|---|
| Oct 2014 | <i>Hybrid QM/QM Calculations on Metal Complexes - Approaching the Coupled Cluster Limit</i>
Triennial Congress of the World Association of Theoretical and Computational Chemists (WATOC), Santiago, Chile |
| Sept 2014 | <i>Hybrid QM/QM Studies of Molybdenum Enzymes - Reaction Pathways and Redox Potentials</i>
ORCA summer school, Gelsenkirchen, Germany |
| Sept 2013 | <i>Hybrid QM/QM local correlation methods for the study of metal ions with biomolecular ligands</i>
49th Symposium on Theoretical Chemistry, Erlangen, Germany |
| July 2013 | <i>Hybrid QM/QM open-shell local correlation methods for the study of metal sites in biomolecular catalysis</i>
32nd International Symposium on Free Radicals, Potsdam, Germany |
| Sept 2012 | <i>Theoretical studies of metallophilic interactions through the use of local correlation methods</i>
4th EuCheMS Chemistry Congress, Prague, Czech Republic |

Personal Skills and Competences

Native Language	Serbian
Other Languages	English: TOEFL 82/120 German: Level A1.2 following GER/CEFR
Programming	basic knowledge of C++, Fortran and Python
Computational Chemistry	Molpro, ORCA, Turbomole

Publication List

- **M. Andrejić** and R. A. Mata, *Local hybrid QM/QM calculations of reaction pathways in metallobiosites*, J. Chem. Theory Comput. **10**, 5397-5404 (2014).
- **M. Andrejić**, U. Ryde, R. A. Mata, P. Söderhjelm, *Coupled-Cluster Interaction Energies for 200-Atom Host-Guest Systems*, Chem. Phys. Chem. **15**, 3270-3281 (2014).
- M. van Severen, **M. Andrejić**, J. Li, K. Starke, R. A. Mata, E. Nordlander, U. Ryde, *A quantum-mechanical study of the reaction mechanism of sulfite oxidase*, J. Biol. Inorg. Chem. **7**, 1165-1179 (2014).
- P. Mikulskis, D. Cioloboc, **M. Andrejić**, S. Khare, J. Brorsson, S. Genheden, R. A. Mata, P. Söderhjelm, U. Ryde, *Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host-guest binding energies*, J. Comput. Aided. Mol. Des. **28**, 375-400 (2014).
- **M. Andrejić** and R. A. Mata, *Study of ligand effects in aurophilic interactions using local correlation methods*, Phys. Chem. Chem. Phys. **15**, 18115-18122 (2013).