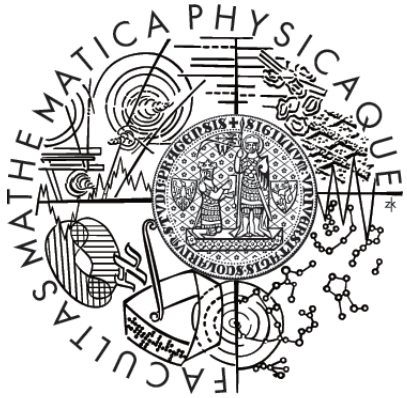
Faculty of Mathematics and Physics
Charles University in Prague



Ph.D. Thesis

Interaction of Metal Cations in Bioorganic Environment

Computational Study Using Quantum Mechanics and Molecular Mechanics Tools

Zdeněk Futera

Supervisor: Prof. RNDr. Ing. Jaroslav Burda, DrSc.

Prague 2012

Foremost I would like to thank to my supervisor, Prof. Jaroslav Burda, for his guidance of my work. I am also very thankful to Dr. Yoshitaka Tateyama and Dr. Keitaro Sodeyama for their encouragement and many fruitful discussions. Last but not least, I am grateful to Ola Macáková for moral support necessary for successful completion of the thesis.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work related to the rigths and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague on June 6, 2012 Zdeněk Futera

Název práce: Interakce kovových iontů v bioorganickém prostředí; kvantově–chemická a molekulově–mechanická výpočetní studie

Autor: RNDr. Zdeněk Futera

Katedra / Ústav: Katedra chemické fyziky a optiky

Vedoucí doktorské práce: Prof. RNDr. Ing. Jaroslav Burda, DrSc.

Abstrakt: Biologicky významné interakce piano–stool Ru(II) komplexů s DNA jsou v této práci studovány QM/MM výpočetní metodou. Reakční mechanismus je rozdělen do tří částí — hydratace $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ komplexu, následné navázání na DNA a vytvoření můstku mezi dvěma guaniny na stejném vlákně. Profily volné energie všech studovaných reakcí jsou spočteny metodou umbrella sampling z QM/MM MD simulací během nichž byl Ru(II) komplex popsán na úrovni DFT. Pro tento účel byl vytvořen QM/MM software, který umožňuje provázat programy Gaussian a Amber. Spočtené energetické bariéry hydratační reakce a procesu navázání Ru(II) komplexu na DNA jsou v souladu s experimentálně změřenými rychlostními konstantami těchto reakcí. Vytvoření můstků bylo předpovězeno jako uskutečnitelné z termodynamického i kinetického hlediska.

Klíčová slova: rutheniové komplexy, piano–stool struktura, hybridní QM/MM metoda, DFT, umbrella sampling, WHAM

Title: Interactions of Metal Cations in Bioorganic Environment; Computational Study Using Quantum Mechanics and Molecular Mechanics Tools

Author: RNDr. Zdeněk Futera

Department / Institute: Department of Chemical Physics and Optics

Supervisor of the doctoral thesis: Prof. RNDr. Ing. Jaroslav Burda, DrSc.

Abstract: Biologically relevant interactions of piano–stool ruthenium(II) complexes with *ds*-DNA are studied by QM/MM computational technique. The whole reaction mechanism is divided into three phases — hydration of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$, consequent binding DNA and final intra–strand cross–link formation between two adjacent guanines. Free energy profiles of all reactions are explored by QM/MM MD umbrella sampling approach where the Ru(II) complex is described by DFT. For that purpose, special QM/MM software was developed to couple Gaussian and Amber programs. Calculated free energy barriers of Ru(II) hydration as well as DNA binding process are in good agreement with experimentally determined rate constants. Reaction pathway for cross–link formation was predicted that is feasible from both thermodynamical and kinetical point of view.

Keywords: ruthenium(II) complexes, piano–stool structure, hybrid QM/MM method, DFT, umbrella sampling, WHAM

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

acac	Acetylacetonate
ACM	Adiabatic Connection Method
Aq	Aqua ligand
B	Becke (1988) exchange functional
B3LYP	B exchange, LYP correlation hybrid functional
B86	Becke (1986) exchange functional
B88	Becke (1988) correlation functional
B95	Becke (1995) correlation functional
BFGS	Broyden–Fletcher–Goldfarb–Shanno minimization algorithm
Bz	Benzene ligand
CC	Coupled Clusters method
CCSD	CC with single and double excitations
CCSD(T)	CCSD with perturbative estimation for triple excitations
CCSDT	CC with single, double and triple excitations
CCSDTQ	CC with single, double, triple and quadruple excitations
CG	Conjugate Gradients algorithm
CI	Configuration Interaction method
CISD	CI with single and double excitations
CISDT	CI with single, double and triple excitations
CISDTQ	CI with single, double, triple and quadruple excitations
C-PCM	Conductor Polarized Continuum Model
COSMO	Conductor like Screening Model
CPU	Central Processing Unit (computer processor)
DC	Divide and Conquer method
DFT	Density Functional Theory
DMSO	Dimethylsulfoxide
en	Ethylenediamine
ECP	Effective Core Potential
ESP	Electrostatic Potential
FCI	Full Configuration Interaction method
FEP	Free Energy Perturbation
FF	Force Field
FFT	Fast Fourier Transformation
FIFO	First In, First Out program container

FMO	Fragment Molecular Orbital method
FMO–LCMO	FMO enhanced by Linear Combination of Molecular Orbitals
GGA	Generalized Gradient Approximation
GTO	Gaussian Type Orbital
Im	Imidazole
Ind	Indasole
HF	Hartree–Fock method/approximation
KCIS	Krieger, Chen, Iafrate, Savin meta–GGA functional
KP1019	*trans*-[$Ru^{III}Cl_4(Ind)_2$] complex
L-BFGS	Limited memory BFGS
LCAO	Linear Combination of Atomic Orbitals
LDA	Local Density Approximation
LSDA	Local Spin Density Approximation
LYP	Lee, Yang, Parr correlation functional
MD	Molecular Dynamics
MM	Molecular Mechanics
MPn	Moller–Pleset Perturbation theory of n-th order
mPBE	Modified PBE exchange functional
mPW	Modified PW exchange functional
NAMI-A	*trans*-[$Ru^{III}Cl_4$(DMSO)(Im)] complex
NPA	Natural Population Analysis
O	OptX exchange functional
P	Perdew exchange functional
P86	Perdew correlation functional
PBC	Periodic Boundary Conditions
PBE	Perdew, Burke, Enzerhof exchange functional
PCM	Polarized Continuum Model
PES	Potential Energy Surface
PME	Particle Mesh Ewald method/algorithm
PMF	Potential of Mean Force
PT	Perturbation Theory
PTA	1,3,5-triaza-7-phosphaadamantane
PTSS	Perdew, Tao, Staroverov, Scuseria meta–GGA functional
PW	Perdew–Wang (1991) exchange functional
PW91	Perdew–Wang (1991) correlation functional
QCISD	Quadratic CI with single and double excitations

QCISD(T)	QCISD with perturbative estimate for triple excitations
QM	Quantum Mechanics
QM/MM	Hybrid Quantum Mechanics – Molecular Mechanics
RAPTA	Ruthenium–Arene–PTA complex
SCF	Self–Consistent Field
SD	Steepest Descent algorithm
STO	Slater Type Orbital
TI	Thermodynamic Integration
UI	Umbrella Integration method
US	Umbrella Sampling method
VdW	Van der Waals
VSXC	Van Voorhis, Scuseria meta–GGA functional
VWN	Vosko, Wilk, Nusair correlation functional
WHAM	Weighted Histogram Analysis Method
X	Exchange functional from Xu and Goddard

R0	Hydration reaction of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex
$\mathbf{R_{R0}}$	Reactant of **R0** ($[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex)
$\mathbf{TS_{R0}}$	Transition state of **R0**
$\mathbf{P_{R0}}$	Product of **R0** ($[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex)
R1	Binding of $[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex to DNA
$\mathbf{R_{R1}}$	Reactant of **R1** ($[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex)
$\mathbf{I_{R1}}$	Intermediate state of **R1** (Ru(II)-O6(G) adduct)
$\mathbf{TS^D_{R1}}$	Transition state of direct **R1** path
$\mathbf{TS^1_{R1}}$	Transition state between $\mathbf{R_{R1}}$ and $\mathbf{I_{R1}}$
$\mathbf{TS^2_{R1}}$	Transition state between $\mathbf{I_{R1}}$ and $\mathbf{P_{R1}}$
$\mathbf{P_{R1}}$	Product of **R1** (Ru(II)-N7(G) adduct)
R2	Ru(II) guanine–guanine cross–link formation
$\mathbf{P_{R1}}$	Reactant of **R2** (Ru(II)-N7(G) adduct)
$\mathbf{I^1_{R2}}$	Cross–link with coordinated benzene and aqua ligand
$\mathbf{TS^{11}_{R2}}$	Transition state between $\mathbf{P_{R1}}$ and $\mathbf{I^1_{R2}}$
$\mathbf{TS^{12}_{R2}}$	Transition state between $\mathbf{I^1_{R2}}$ and $\mathbf{P_{R2}}$
$\mathbf{I^2_{R2}}$	Cross–link with coordinated benzene ligand
$\mathbf{TS^{21}_{R2}}$	Transition state between $\mathbf{P_{R1}}$ and $\mathbf{I^2_{R2}}$
$\mathbf{TS^{22}_{R2}}$	Transition state between $\mathbf{I^2_{R2}}$ and $\mathbf{P_{R2}}$
$\mathbf{P_{R2}}$	Cross–link with coordinated aqua ligand (product of **R2**)

1 Introduction

Living cell, basic unit of all organisms, is one of the most complex and fascinating things in the world. Great effort was done during last decades in molecular biology to understand how the living cell is organized and especially how it works. [1–3]. Although today the structure and functions of main cellular parts are at least partially understood including fundamental concept of expression of genes into the proteins, there are still many secrets.

Experimentally measured structures of nucleic acids and many proteins are available, obtained in most cases by X-ray crystallographic methods and NMR spectroscopy. Proteins are very complex macromolecules, often consisting of more than one cooperating units, and functionality differs from type to type. Some of them serve only as building material, their function is exclusively structural. However, large group of proteins have very specific functions, like enzymes catalyzing specific chemical reactions or transfer proteins carrying ions and small molecules through cellular membrane.

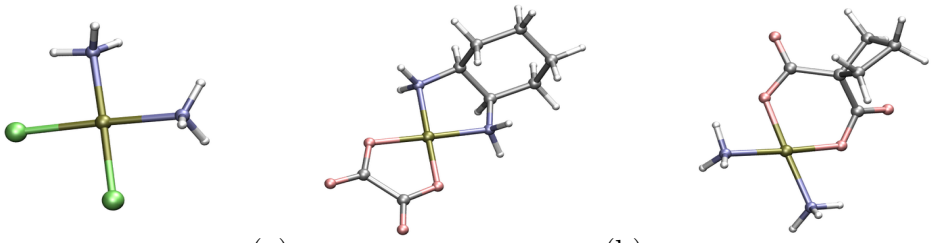
Interestingly, a lot of proteins contain one or more metal cations in their structure. These metals can be found in active sites of enzymes as well as deep inside membrane proteins. While there are only a few biogenic elements used for formation of all biological material (C, H, O, N, S, P), range of metals that can be found in biomacromolecules is surprisingly wide (Fe, Co, Mg, Zn, Cu, Ni, Cr, Mo, Se,...) [3, 4].

Nature exploits the high coordination numbers of transition metals and different spin states, which can influence geometry of their nearest ligands. Thanks to these properties and strong interaction with organic material, metals are often used in organic synthesis [5, 6] and in pharmacology [7, 8]. Several transition metal complexes are successfully used as drugs in medical treatment of various diseases such as arthritis (Au), gastric ulcers (Bi) or various skin diseases (Zn) [9]. Special group of drugs are metal-based chemotherapeutic agents used against cancer tumours [10].

1.1 Cisplatin

First metallodrug which was applied in anti–cancer treatment is *cis*-diamminedichloroplatinum(II) complex (Fig. 1a), usually called cisplatin. It was discovered accidentally in 1950s by Prof. Rosenberg [11] and approved as chemotherapeutic drug in 1978 after series of medical trials [12]. Cisplatin is up today used in 70% cases of chemotherapy administration [13].

Intensive experimental research [14–18] complemented by several computational studies [19, 20] followed and finally mechanism of cisplatin anti-tumour activity was elucidated. Cisplatin enters a cancer cell through cellular membrane by either passive or ac-



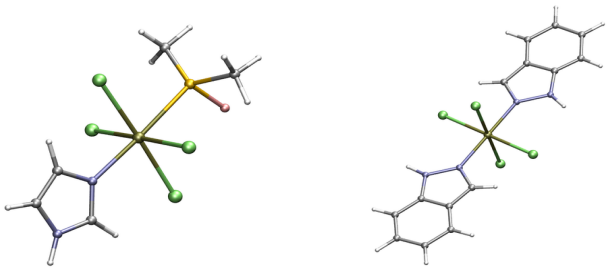
(a) (b) (c)

Figure 1: Platinum complexes: (a) cisplatin, (b) oxaliplatin, (c) carboplatin

tive transport mediated by Ctr1 copper transporter [21–23]. Inside the cell, in cytoplasm, is significantly lower concentration of Cl^- anions than in extracellular environment (110 mmol/dm^3 vs. 10 mmol/dm^3). This causes replacement of one or both cisplatin chloride ligands by water molecule [14, 24]. This hydration reaction is known as activation process because the resulting platinum aqua complex is much more reactive than cisplatin itself. Afterwards, the aqua complex can interact with various biological material, however, the most important for anticancer activity is its interaction with DNA in a cellular nucleus.

It is well known that cisplatin prefers binding to guanine base in DNA where it interacts with N7 nitrogen [12]. This binding site is well accessible from major groove of DNA double–helix. The aqua ligand of platinum(II) complex is released during the binding. In guanine rich sequences of DNA, like telomeres, cisplatin can form a bridge between two nucleic bases. Although inter–strand cross–linked structures were also observed, the most often type is intrastrand bridge between N7 nitrogens of two adjacent guanines. When this structure is formed, DNA is locally unwounded and deformed as a result of strong interaction with cisplatin. Such a deformation leads to blocking of replication or transcription of DNA and consequently to apoptosis.

As for cytostatic activity, cisplatin is very effective against several types of tumours, especially ovarian cancer or cancer of neck and head [25], however, many other kinds are resistant to cisplatin intrinsically or the resistance can be induced during the treatment [26]. Moreover, there are many severe side effects, such as kidney damage, hair loss or nausea, that can complicate the medical treatment. Therefore the research in this area is still very active. New platinum complexes, which succeeded medical trial, and are used in medical practice now, were discovered in 1980s or later. Among most successful and important ones belong oxaliplatin (Fig. 1b, [27]) and carboplatin (Fig. 1c, [28]). Other promising metallodrugs still undergo clinical trials [25].



(a) (b)

Figure 2: Ru(III) complexes: (a) NAMI-A, (b) KP1019

1.2 Ruthenium complexes

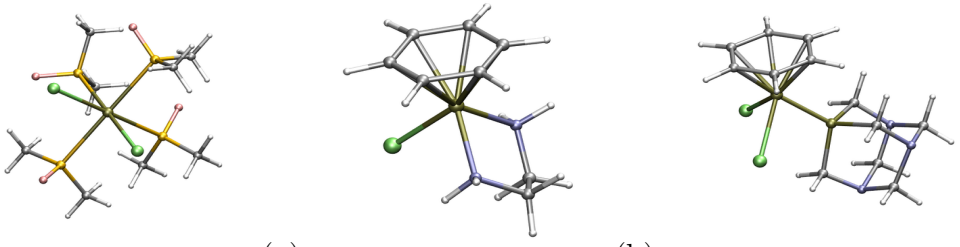
Beside platinum, metals chemically similar to iron are used for synthesis of anti-cancer active complexes in most cases. Research in recent years concerned among others on ruthenium and rhodium compounds and several promising complexes were already discovered in this area [29, 30].

First, Ru(III) complexes and their properties were explored extensively and two successful drugs with very good anti-tumour activity were discovered. These are known by abbreviations NAMI-A (*trans*-[$Ru^{III}Cl_4$(DMSO)(Im)], DMSO = dimethylsulfoxide, Im = imidazole; Fig. 2a) and KP1019 (*trans*-[$Ru^{III}Cl_4(Ind)_2$], Ind = indasole; Fig. 2b). Both of them are still undergoing clinical testing.

It is assumed that oxidation state of Ru(III) in NAMI-A and KP1019 is reduced to Ru(II) in biological environment and so the research moved to Ru(II) complexes [31]. First, cisplatin analogue compound, *cis*-[$Ru^{II}(DMSO)_4Cl_2$], was discovered and recently interesting half–sandwich complexes with η^6–coordinated arene ligand were prepared. These compounds have [$Ru^{II}(\eta^6$-arene)(chelate)X]$^+$(chelate = en or acac, X = Cl or Br) structure and are also known as Ru(II) "piano–stool" complexes, named after their typical shape [32–35]. Similar biologically active complex including bulky PTA ligand is known as RAPTA complex ([$Ru^{II}(\eta^6$-arene)(PTA)X_2]) [36, 37]. Structure of these complexes is shown in Fig. 3.

1.3 Piano stool Ru(II) complexes

Behaviour of piano–stool Ru(II) complexes in biological environment is very similar to cisplatin. First of all, before any interaction with biological material, the hydration reaction occurs in the cell and chloride ligand in [$Ru^{II}(\eta^6$-benzene)(en)Cl]$^+$ is substituted by water molecule. The reaction, called activation process, proceeds partly in cytoplasm but mainly



(a) (b) (c)

Figure 3: Ru(II) complexes: (a) *cis*-[$Ru^{II}(DMSO)_4Cl_2$], (b) [$Ru^{II}(\eta^6$-benzene)(en)Cl]$^+$, (c) [$Ru^{II}(\eta^6$-benzene)(PTA)Cl_2]

in nucleus of the cell where is very low concentration of Cl^- anions (4 mmol/dm^3) [34]. Hydration of Ru(II) complexes is ten times faster than activation of cisplatin, however, in both cases the reaction is endothermic [34, 38]. It is known that type of chelate ligand has strong influence on kinetics of the hydration [39, 40], however, size of arene ligand has negligible effect.

Resulting aqua complex [$Ru^{II}(\eta^6$-benzene)(en)$(H_2O)]^{2+}$ is more reactive than its chloride–type predecessor and can interact with various nucleophile centers in oligopeptides, proteins, free aminoacids or nucleic acids. As in the case of cisplatin, the most interesting is the interaction with DNA. From experimental data [32, 34] as well as computation studies [41–43] is known that also Ru(II) piano–stool complexes prefers N7 guanine position for binding. However, this substitution of aqua ligand can proceed by two reaction pathways as is schematically shown in Fig. 4. Besides direct binding to N7 nitrogen, two–step reaction mechanism going through Ru(II)–O6(G) intermediate state is possible [42].

Since the behaviour of Ru(II) complexes, as fas as is known, is analogous to cisplatin,

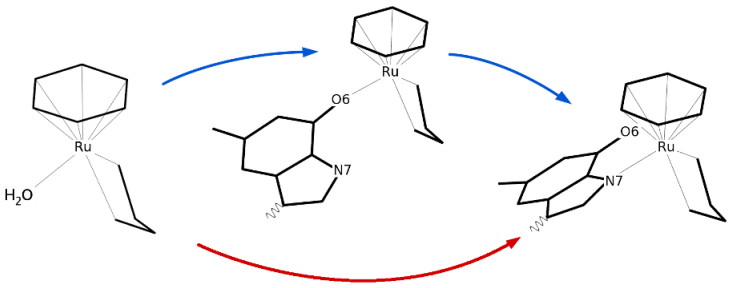


Figure 4: Reaction scheme of DNA ruthenization: [$Ru^{II}(\eta^6$-benzene)(en)$(H_2O)]^{2+}$ reacts with DNA directly and binds to N7 guanine position (red path) or undergoes two–step mechanism through O6 gunine position (blue path)

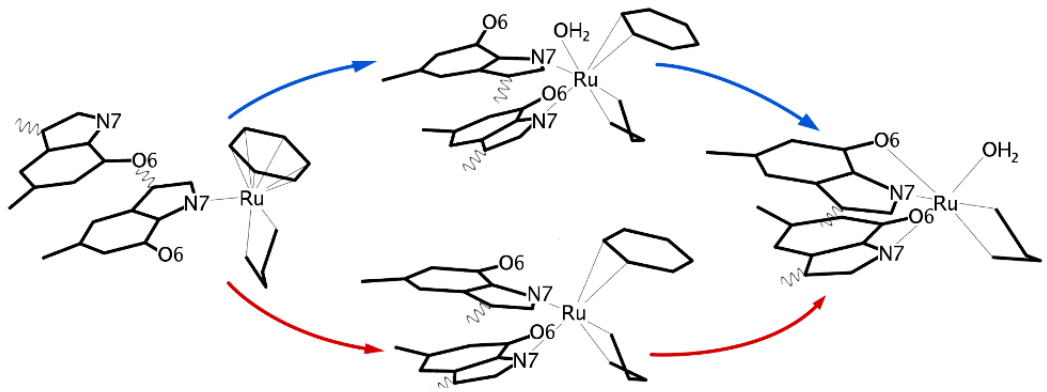


Figure 5: Suggested reaction scheme of Ru(II) cross–link formation. Two reaction pathways are considered: (1) transition from Ru(II)-N7(G1) mono–adduct to cross–link with η^2–coordinated benzene ligand and Ru(II)-O6(G2) interaction followed by benzene–water substitution (red path), (2) formation of cross–link with both aqua and benzene ligand and consequent benzene release (blue path).

there is a possibility that they can also create a similar cross–link between two adjacent guanines on the same DNA strand. This kind of binding was observed in case of dinuclear Ru(II) complexes but it was not assumed in monofunctional $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ type discussed here [34]. However, when we consider that benzene ligand could change its coordination from η^6 to η^2 then two valences on Ru(II) cation are available for interaction with N7 nitrogen on adjacent guanine and one more nucleophile site. Reaction scheme for such process is suggested in Fig. 5. Saturation of remaining free valence on Ru(II) by either O6 oxygen on guanine or aqua ligand is assumed together with complete release of benzene ring from the complex. Validity of this reaction scheme is studied in this work by methods of computational chemistry.

The thesis is organized as follows. In the next section is reviewed theoretical background of computational chemistry methods, that is molecular mechanics, quantum chemical approach and their mutual combination – hybrid QM/MM technique. This technique was implemented as part of this work in special software and details are described in Sec. 3. Main part of the thesis, that is study of Ru(II) reaction mechanism by these computational methods, is discussed in Sec. 4. Hydration reaction, binding to DNA and consequential creation of cross–linked structures is systematically explored. In the last section Sec. 5 are briefly summarized all obtained results of this work.

2 Theory

Brief overview of basic approximations and methods used for calculation geometries and electronic properties of molecular systems is presented in this section. First, simple molecular mechanical approach is described here followed by more sophisticated quantum mechanical methods. Afterwards, their combination, that is a hybrid quantum mechanic – molecular mechanic technique (QM/MM) and their specific problems are discussed.

2.1 Molecular mechanics

Simple but straightforward and intuitive approach how to describe potential function of molecular system is to use rough molecular mechanical (MM) approximation and treat molecules as system of "solid balls" connected by "strings" as is schematically shown in Fig. 6. Although such description is far from being accurate, it is widely exploited for geometry optimization and time evolution of large molecular systems because of relatively low computational cost.

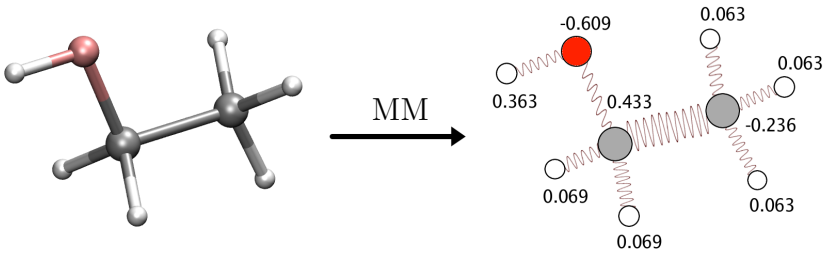


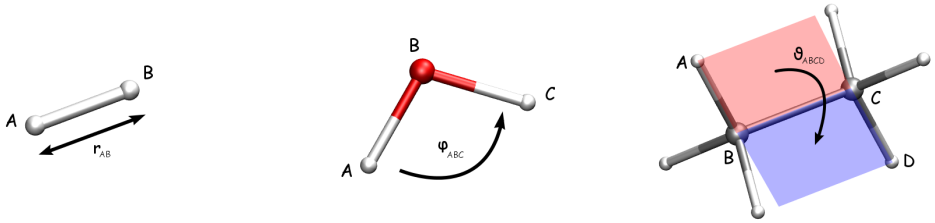
Figure 6: Schematic illustration of MM approximation on ethanol molecule. Interaction between bonded atoms is simulated by mechanical strings. Each atom has assigned partial charge.

2.1.1 Potential energy function

Potential energy function of molecular system, E_{MM}, is constructed as a sum of several independent contributions

$$E_{\mathrm{MM}} = E_{\mathrm{B}} + E_{\mathrm{A}} + E_{\mathrm{D}} + E_{\mathrm{VdW}} + E_{\mathrm{Cl}}, \tag{1}$$

where first three terms represent bonding interactions (stretching E_{B}, bending E_{A} and torsion energy E_{D}) and last two terms are non-bonding energy contributions (Van der Waals and Coulomb energy).



(a) (b) (c)

Figure 7: Schematic vizualization of MM bonding terms: (a) bond distance r_{AB}, (b) angle φ_{ABC}, (c) dihedral angle θ_{ABCD}

Bonding energy

For the sake of simplicity and low computational demands the bonding terms of potential energy should have analytical form, be continuously differentiable and dissociation energy of bond should be positive [44]. Required analytical form is obtained using Taylor expansion of potential energy in equilibrium geometry $\mathbf{r_{eq}}$

$$U(\mathbf{r}) = \sum_{n=0}^{\infty} \frac{1}{n!} (\mathbf{r} - \mathbf{r_{eq}})^n \frac{d^n U(\mathbf{r})}{d\mathbf{r}^n}\bigg|_{\mathbf{r}=\mathbf{r_{eq}}} \tag{2}$$

In harmonic approximation, all terms in sum (2) higher than second order are neglected. The first term is constant that can be set to zero by definition and first order is not contributing since forces in minimum of PES are zero. Therefore the entire sum (2) is reduced to second, quadratic term representing parabolic potential.

Potential energy $U(\mathbf{r})$ is not known and its derivation in (2) has to be substituted by empirical constant k^{B}. The energy of bonds, E_{B}, is then evaluated as a sum of contributions from all N_B bonds

$$E_{\mathrm{B}} = \frac{1}{2} \sum_{i=1}^{N_B} k_i^B (r_i - r_{\mathrm{eq}})^2 \tag{3}$$

Harmonic potential is satisfyingly good in vicinity of equilibrium distance but dissociation than should occur for longer distances can not be described in this approximation. Such a behaviour can be at MM level of theory described for example by Morse potential [45] with two empirical constants: dissociation energy D and fitting constant α. Energy of bonds is then

$$E_{\mathrm{B}}^{\mathrm{Morse}} = \sum_{i=1}^{N_B} D_i \left[1 - e^{-\alpha_i(\mathbf{r}-\mathbf{r_{eq}})}\right]^2 \tag{4}$$

Potential for angle bending in harmonic approximation is constructed analogously. If an angle between two bonds with common center atom (see Fig. 7b) is marked as φ and k^{A} represent force constant modulating width of potential around equilibrium value of bending angle φ_{eq}, then energy of angles is sum of N_A contributions

$$E_{\mathrm{A}} = \frac{1}{2} \sum_{i=1}^{N_A} k_i^A (\varphi_i - \varphi_{\mathrm{eq}})^2 \tag{5}$$

Torsions can be in general described by Fourier series because they are periodic. The Fourier series used in MM has a general form

$$U(\vartheta) = \frac{1}{2} \sum_{i=0}^{\infty} V_i \left[1 + (-1)^{i+1} \cos\left(i\vartheta + \delta\right)\right], \tag{6}$$

where ϑ is torsional dihedral angle, δ is phase shift and V_i are Fourier coefficients of the series. Usually, the first term in (6) is set to zero by definition and all terms greater than first order are neglected. Then the energy of torsions is a sum of N_D contributions from each dihedral angle:

$$E_{\mathrm{D}} = \frac{1}{2} \sum_{i=1}^{N_D} k_i^D [1 + \cos(n\vartheta_i - \delta_i)] \tag{7}$$

Fourier coefficient in Eq. 7 is replaced by force constant k^D and periodicity n is introduced in the formula.

Non-bonding energy

Non-bonding interactions are pair interactions between atoms that are not connected by chemical bonds. They are of two kinds: short–range Van der Waals interactions and long–range Coulombic electrostatic interactions.

Van der Waals interactions are usually in MM described by Lennard–Jones 12–6 potentials [46] with two empirical parameters for each atomic type: potential depth ϵ and interaction distance r^0. Parameters for pair interaction between two different types of atoms are obtained as geometric resp. arithmetic average of these constants:

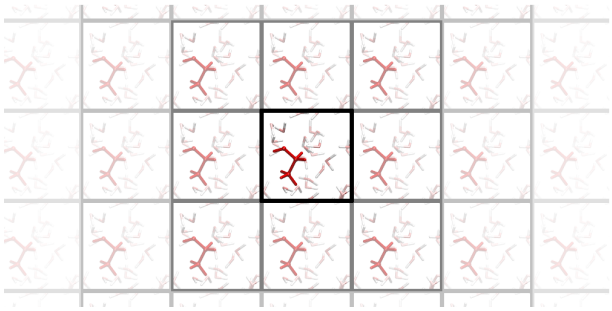


Figure 8: Schematic vizualization of periodic boundary conditions in 2D: simulation box is in the center surrounded by its virtual images

$$E_{\mathrm{VdW}} = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \epsilon_{ij} \left[\left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{6} \right], \quad r_{ij}^0 = \frac{1}{2}(r_i + r_j), \quad \epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \tag{8}$$

Electrostatic interaction is described by classical Coulomb law, so atomic charges Q for each atom have to be defined

$$E_{\mathrm{Cl}} = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{Q_i Q_j}{r_{ij}} \tag{9}$$

From (8) and (9) is obvious that non-bonding interactions are evaluated for each pair of atoms. This can be very computational demanding task especially in huge systems with several thousands of atoms. Therefore some cut-offs for Van der Waals as well as Coulombic interaction are usually introduced. Then, interaction of all atoms with mutual distance greater than chosen cut-off are set to be zero.

2.1.2 Periodic boundary conditions

Although MM approximation allow us to simulate large systems with several thousands of atoms, sometimes much bigger models are needed for proper description of a studied phenomenon. This is the case mainly when the studied system is periodic (crystals) or if infinite solution environment should be simulated. In such cases the system can be enlarged artificially using periodic boundary conditions (PBC).

Concept of PBC is shown on Fig. 8. The basic simulation box, which may be relatively small, is surrounded by its images in all directions. Pair interaction between two atoms

in given cut-off region is then calculated according to expression [47]:

$$\mathbf{F}_{ij}^{\mathrm{PBC}} = \sum_{\mathbf{n}} \mathbf{F} \left(\left| \mathbf{r}_i - \mathbf{r}_j + \sum_{\mu=1}^{3} a_\mu n_\mu \right| \right) \tag{10}$$

where a_μ is box size in μ direction and $\mathbf{n}$ is a vector pointing to one of the virtual images of the basic cell.

This approach is memory saving because only information about basic cell atoms is kept in computer memory. There are no unwanted surface effects and is computationally less demanding as using huge simulation box. However, basic cell size has to be set carefully to avoid self-interaction of large molecules.

Ewald summation

With PBC the simulated system has effectively infinite size and so it is not possible to evaluate non-bonding interaction energies (8) and (9) directly as in finite box. Van der Waals interactions are short–range, decaying as r^{-6}, therefore relatively small cut-off can be applied to evaluate sum (8). However, this is not possible in case of long–range Coulombic interaction decaying as r^{-1}.

However, there is a way how to calculate exact electrostatic interaction in periodic system by summing all contribution in reciprocal space [48]. Derivation was first done by P. P. Ewald and today can be found in standard textbooks of computational physics, e.g. [47, 49]

$$\begin{aligned} E_{\mathrm{El}}^{\mathrm{Ewald}} = & \frac{1}{4\pi\epsilon_0} \sum_{\mathbf{n}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{Q_i Q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}L|} \mathrm{erfc}\left(\frac{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}L|}{\sqrt{2}\sigma} \right) \\ & + \frac{1}{2V\epsilon_0} \sum_{\mathbf{k}\neq 0} \frac{e^{-\sigma^2 k^2/2}}{k^2} |S(\mathbf{k})|^2 \\ & - \frac{1}{4\pi\epsilon_0} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{N} Q_i^2 \end{aligned} \tag{11}$$

Ewald sum (11) consists of three terms. First is short-ranged in real space and is truncated by error function, second is short-ranged in reciprocal space and is truncated by quadratic exponential function and the last one is charge self–interaction energy.

Ewald sum algorithm scales as $\mathcal{O}(N^{3/2})$ and so it can be very demanding for large molecular systems. To improve that, particle mesh Ewald algorithm (PME) was derived in 1990s [50] which uses Fast Fourier transformation (FFT) on a grid and scales only as

$\mathcal{O}(N \log N)$. PME is now implemented in almost all MM software.

2.1.3 Force field

Formulas (3), (5), (7), (8) and (9) together with set of needed empirical parameters (force constants k^{B}, k^{A} and k^{D}, equilibrium distances r_{eq}, angles φ_{eq}, torsional periodicities n and phase shifts δ, Lennard–Jones atomic parameters r, ϵ and charges Q) define MM force field (FF). Because force fields are often used for molecular dynamics (MD, see Sec. 2.1.4) they also include atomic masses m.

FF parameters are fitted to reproduce experimental or calculated values of various quantities like bond lengths, interatomic angles, dipole moments, enthalpies of vaporization or sublimation etc. Parametrization is done on specific set of molecules, e.g. hydrocarbons, proteins, nucleic acids, inorganic crystals, glasses.

Among most popular force fields belong AMBER [51] (proteins, nucleic acids), GAFF [52] (small organic molecules), CHARMM [53] (small molecules as well as macromolecules), GROMACS [54], GROMOS [55] (proteins, nucleotides, sugars), OPLS [56], UFF [57] (parameters for all elements), CFF [58], MM2/MM3/MM4 [59–63] (hydrocarbons, small organic molecules), SIBFA [64] (small molecules, flexible proteins, metalloproteins) or AMOEBA [65] (polarizable force field for organic molecules).

2.1.4 Molecular dynamics

Because MM description of molecular systems has relatively low computational cost, it is possible calculate also the time evolution of the systems [66]. This is called molecular dynamics (MD) and it is based on integration of classical equations of motions, i.e. basically Newton's equations.

State of a particle in classical mechanics is completely set by knowledge of its position $\mathbf{r}$ and momentum $\mathbf{p}$. Time evolution of these two quantities can be in general written as

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \int_{t_0}^{t} \frac{\mathbf{p}(t')}{m} dt' \tag{12}$$

$$\mathbf{p}(t) = \mathbf{p}(t_0) + m \int_{t_0}^{t} \mathbf{a}(t') dt' \tag{13}$$

where $\mathbf{a}(t) = \mathbf{F}(t)/m$ is acceleration of an atom at given time t. There are several algorithms for evaluation of integrals in (12) and (13) differing by complexity and accuracy. One of the most often used algorithm for MD of molecular systems is known as "Velocity Verlet" and is described below.

Velocity Verlet

Original Verlet algorithm is derived from forward and backward Taylor expansion of position in time [67]. This leads to formula for position updating

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)(\Delta t)^2 \tag{14}$$

with error scaling as $\mathcal{O}((\Delta t)^4)$. But no explicit formula for velocities is given and so velocities have to be calculated by differencing of two subsequent positions and that leads to error $\mathcal{O}((\Delta t)^2)$.

To avoid this problem Leap–Frog type algorithm was invented later [68] and by its combination with (14) so the called velocity Verlet algorithm can be derived [69]. Positions and velocities in this integration method are updated according to following formulas

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}\Delta t + \frac{1}{2}\mathbf{a}(t)(\Delta t)^2 \tag{15}$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t+\Delta t) + \mathbf{a}(t)]\Delta t \tag{16}$$

Velocity Verlet method is numerically more stable then original Verlet algorithm while the error scaling is the same [47].

Equilibrium thermodynamic properties

Statistical physics and its methods provide a connection between atomic description of the system and macroscopic parameters such as temperature or pressure [70]. Time averaging of variable A during the MD productive trajectory and its relating to statistical averages $\langle A \rangle$ is justified by ergodic hypothesis [71]

$$\bar{A} \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T A(t)dt = \frac{\iint A(\mathbf{r},\mathbf{p}) \exp\left(-\frac{H(\mathbf{r},\mathbf{p})}{k_B T}\right) d\mathbf{r}d\mathbf{p}}{\iint \exp\left(-\frac{H(\mathbf{r},\mathbf{p})}{k_B T}\right) d\mathbf{r}d\mathbf{p}} \equiv \langle A \rangle \tag{17}$$

In computer simulations it is not possible to calculate infinitely long trajectory to get exact time average $\bar{A}$ but finite set averaging from N samples is used to approximate this value

$$\bar{A} \approx \frac{1}{N}\sum_{i=1}^{N} A_i \tag{18}$$

Because samples gathered from MD trajectory are not exactly independent, standard statistical formula for variance can not be used to get deviation from the average [47, 71]. The simplest approach to overcome this problem is to block averaging method [72].

Variation is then calculated as

$$\sigma^2 = \frac{1}{N_b - 1} \sum_{\beta=1}^{N_b} \left(A_\beta^2 - \langle A \rangle_b^2 \right) \tag{19}$$

where N_b is number of used blocks, A_β is a block average and $\langle A \rangle_\beta$ is overall average [71].

Probably the most important quantities besides total energy of the system are its temperature and pressure. Both are usually kept fixed during the productive simulations of molecular systems. Temperature is calculated from Boltzmann's equipartition theorem

$$\frac{3}{2} N k_\mathrm{B} T = \left\langle \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2 \right\rangle \tag{20}$$

where N is total number of particles in the system.

Virial theorem can be then used for pressure evaluation. If the simulation box has volume V, temperature T and contains N particles then it holds

$$pV = N k_\mathrm{B} T - \frac{1}{3} \left\langle \sum_{i=1}^{N} \mathbf{r}_i \mathbf{F}_i \right\rangle \tag{21}$$

2.2 Quantum mechanics

Molecular mechanic approach described in Sec. 2.1 is computationally very fast and so suitable for MD studies of large molecular systems. However, the great disadvantage of MM is that it can not describe chemical changes, i.e. making and/or breaking of chemical bonds. Adequate theory for proper description of molecules is quantum mechanics (QM).

Quantum mechanics is based on the concept of wave function $\Psi(\mathbf{x}, t)$ which includes information about state of given system. Time evolution of wave function is then described by Schrödinger equation

$$i\hbar \frac{d\Psi(\mathbf{x}, t)}{dt} = \hat{H}(\mathbf{x}, t) \Psi(\mathbf{x}, t) \tag{22}$$

where $\hbar$ is reduced Plank's constant and $\hat{H}$ is generally time-dependent Hamiltonian of the system. If the Hamiltonian is time-independent, Eq. 22 can be separated for spacial and time part of wave function $\Psi(\mathbf{x}, t) = \psi(\mathbf{x}) \varphi(t)$. Time independent Schrödinger equation for stationary states has then form

$$\hat{H}(\mathbf{x}) \psi(\mathbf{x}) = E \psi(\mathbf{x}) \tag{23}$$

Hamiltonian of molecular system in non-relativistic approximation consists of kinetic

energy of nuclei, kinetic energy of electrons, electron–nuclei interaction, electron–electron interaction and nucleus–nucleus interaction term

$$\hat{H} = -\frac{1}{2}\sum_{A=1}^{N_n}\frac{1}{m_A}\Delta_A - \frac{1}{2}\sum_{i=1}^{N_e}\Delta_i - \sum_{i=1}^{N_e}\sum_{A=1}^{N_n}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N_e-1}\sum_{j=i+1}^{N_e}\frac{1}{r_{ij}} + \sum_{A=1}^{N_n-1}\sum_{B=A+1}^{N_n}\frac{Z_A Z_B}{r_{AB}} \quad (24)$$

Atomic units are used in (24) as is usual in QM texts, i.e. electron mass m_e, electron charge e and Coulombic prefactor $1/4\pi\epsilon_0$ are equal to one. Number of electrons and nuclei are denoted as N_e and N_n, m_A represent mass of nucleus A and Z_A is its charge.

2.2.1 Born–Oppenheimer approximation

Because nuclei of atoms are much more heavier than electrons it is possible to separate motion of these two kinds of particles and solve only motion of electrons in a electrostatic field of fixed nuclei. This is called Born–Oppenheimer approximation and it simplifies the Hamiltonian to form

$$\hat{H}_{\mathrm{BO}} = -\frac{1}{2}\sum_{i=1}^{N_e}\Delta_i - \sum_{i=1}^{N_e}\sum_{A=1}^{N_n}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N_e-1}\sum_{j=i+1}^{N_e}\frac{1}{r_{ij}} + \sum_{A=1}^{N_n-1}\sum_{B=A+1}^{N_n}\frac{Z_A Z_B}{r_{AB}} \quad (25)$$

The last term in (25), the electrostatic interaction of nuclei, is now constant because of assumed fixed atom nuclei. This term is standardly marked as V_{nn}. The Schrödinger equation (23) is solved without V_{nn} and the nuclei contribution to total energy $E_{nn} = \langle\psi|V_{nn}|\psi\rangle$ is added afterwards.

2.2.2 LCAO and Basis set

For solving (23), it is convenient represent the wave function $\psi(\mathbf{x})$ in some practically chosen set of basis functions. For molecular systems is natural to use atomic orbitals $\phi(\mathbf{x})$, i.e. mathematical function representing distribution of individual electrons in hydrogen atom as a basis for projection of $\psi(\mathbf{x})$

$$\psi(\mathbf{x}) = \sum_{i=1}^{K} d_i\phi_i(\mathbf{x}) \quad (26)$$

Expansion (26) is called LCAO (Linear Combination of Atomic Orbitals) and it is an approximative expression because the sum is truncated to finite number of terms. This is necessary for practical numerical solving on computers.

Next approximation that is usually done is a simplification of mathematical form of

$\phi(\mathbf{x})$ because of complexity of hydrogen functions. These can be approximated by Slater type orbitals (STO). However, STO are not so convenient, because atomic integrals with them are not analytical. Therefore they are often expanded into series of Gaussian type orbitals (GTO) with general form [44]

$$\phi_{ijk}^{\mathrm{GTO}}(\mathbf{r}) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \left[\frac{(8\alpha)^{i+j+k} i!j!k!}{(2i)!(2j)!(2k)!}\right]^{1/2} x^i y^j z^k e^{-\alpha r^2} \tag{27}$$

where indices i,j and k specifies type of the orbital and α is exponential factor different for each type. More details about different basis sets and their specifications can be found in computational chemistry textbooks [44, 73–75] or manuals of quantum chemical software [76].

Effective core potentials

Heavy atoms like transition metals contain large number of electrons, but chemically important are mainly those included in valence shells. Moreover, inner electrons are relatively close to nucleus and there are non-negligible relativistic effects. To avoid demanding relativistic description and reduce number of electrons, core electrons can be substitute by analytical function W, so called effective core potential (ECP). Schrödinger equation (23) is then solved only for valence electrons

$$\left[\hat{H}(\mathbf{x}) + \hat{W}(\mathbf{x})\right] \psi_v(\mathbf{x}) = E_{\mathrm{ECP}} \psi_v(\mathbf{x}) \tag{28}$$

Operator $\hat{W}$ is generalized Phillips–Kleinmann pseudopotential, non-local function depending on $\psi_v(\mathbf{x})$ [77]. To avoid computationally demanding iterative solving, $\hat{W}$ is substituted by local pseudopotential with similar shape. For example Stuttgart–Dresden pseudopotentials [78] are defined as

$$\hat{W}_{\mathrm{SD}} = -\frac{Z-n}{r} + \sum_l \hat{W}_l(r)\hat{P}_l \tag{29}$$

where $\hat{P}_l$ is projection operator to angular momentum functions and $\hat{W}_l$ is radial part of the pseudopotential specified by parameters c and α

$$\hat{P}_l = \sum_{m=-l}^{l} |Y_{lm}\rangle\langle Y_{lm}|, \qquad \hat{W}_l(r) = \sum_k c_{lk} \frac{e^{\alpha_{lk} r^2}}{r^{n_{lk}}} \tag{30}$$

2.2.3 Hartree–Fock method

Because electrons are fermions, the total wave function of the system $\psi(\mathbf{x})$ has to be antisymmetric in electron variables. The easiest way how to set up this condition is to use one–electron approximation and represent $\psi(\mathbf{x})$ by Slater determinant

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{N_e!}} \sum_{\mathcal{P}} (-1)^{\mathrm{sgn}(\mathcal{P})} \mathcal{P}[\chi_1(\mathbf{x}_1), \ldots, \chi_n(\mathbf{x}_n)] \tag{31}$$

$\mathcal{P}$ represents permutation of electrons $\mathbf{x}_i$ in set of electron functions χ_j. These function are called molecular orbitals. LCAO is in this case applied on each molecular orbital instead of total wave function.

Using Slater determinant (31) in (23) and applying variational principle leads after several mathematical operations to set of canonical Hartree-Fock equations [73, 75]

$$\left[\hat{H}_1^{\mathrm{core}} + \sum_{j=1}^{N_e} \hat{J}_j - \hat{K}_j \right] \chi_i(\mathbf{x}_1) = \epsilon_i \chi_i(\mathbf{x}_1) \tag{32}$$

where $\hat{H}_1$ is one–electron Hamiltonian and $\hat{J}_i$ and $\hat{K}_i$ are Coulombic and exchange operators acting on i-th electron. ϵ_i has the meaning of one–electron energy. Operator in square brackets is called Fock operator.

Solution of Hartree–Fock equations (32) is set one–electron of molecular orbitals $\chi(\mathbf{x})$ and total energy of the system given by expression

$$E_{\mathrm{el}}^{\mathrm{HF}} = \sum_{i}^{\mathrm{occ}} \epsilon_i - \sum_{i<j}^{\mathrm{occ}} (J_{ij} - K_{ij}) = \frac{1}{2} \sum_{i}^{\mathrm{occ}} (\epsilon_i + H_{ii}) \tag{33}$$

Because electron–electron interaction is in (32) counted twice, the electronic energy (33) is not simple summation of one–electron contributions. H_{ii} in (33) represent one electron integral

$$H_{ii} = \int \chi_i^*(\mathbf{x}_1) \hat{H}_1^{\mathrm{core}} \chi_i(\mathbf{x}_1) d\mathbf{x}_1 \tag{34}$$

Two electron integrals are of two kinds, symmetric Coulombic J_{ij} and antisymmetric exchange K_{ij} integrals defined by formulas [73]

$$J_{ij} = \int \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \chi_j(\mathbf{x}_2) \chi_i(\mathbf{x}_1) d\mathbf{x}_1 d\mathbf{x}_2 \tag{35}$$

$$K_{ij} = \int \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \chi_j(\mathbf{x}_1) \chi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \tag{36}$$

Closed shells

Before applying LCAO expansion on molecular orbitals in HF equations (32) it is convenient to distinguish between systems with closed and open shells, i.e. systems with odd or even number of electrons. In so called restricted approximation it is assumed that space parts of orbitals are the same for pairs of electron differing only by spin states.

When the system is closed shell type, application of LCAO to HF equations leads to Roothaan–Hall equations for expansion coefficients $c_{\nu i}$ of molecular orbitals $\psi_i(\mathbf{r})$ into the set of atomic orbitals $\phi_\nu(\mathbf{r})$ [79]

$$\sum_{\nu=1}^{K}\left[H_{\mu\nu}^{\text{core}}+\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K}P_{\lambda\sigma}\left((\mu\nu|\lambda\sigma)-\frac{1}{2}(\mu\sigma|\lambda\nu)\right)\right]c_{\nu i}=\epsilon_i\sum_{\nu=1}^{K}S_{\mu\nu}c_{\nu i} \tag{37}$$

$H_{\mu\nu}^{\text{core}}$ and $(\mu\nu|\lambda\sigma)$ are one and two–electron integrals (34), (35) and (36) transformed into the set of atomic orbitals (26) [73]. Chemical notation is used for the two–electron integrals. Matrix $\mathbb{S}$ is an overlap matrix of atomic orbitals and $\mathbb{P}$ is a density matrix constructed as

$$P_{\lambda\sigma}=2\sum_{i}^{\text{occ}}c_{\sigma i}^{*}c_{\lambda i} \tag{38}$$

Total electronic energy calculated from Roothaan-Hall equations (37) is then given by expression [73]

$$E_{\text{el}}^{\text{RH}}=\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}P_{\mu\nu}H_{\mu\nu}+\frac{1}{2}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K}P_{\mu\nu}P_{\lambda\sigma}\left[(\mu\nu|\lambda\sigma)-\frac{1}{2}(\mu\sigma|\lambda\nu)\right] \tag{39}$$

Open shells

In case of open shell system is necessary to distinguish between two spin states of electron, marked as α and β. This leads to coupled Pople–Nesbet's equations for two set of molecular orbitals [73, 80]

$$\sum_{\nu=1}^{K}\left[H_{\mu\nu}^{\text{core}}+\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K}P_{\lambda\sigma}^{T}(\mu\nu|\lambda\sigma)-P_{\lambda\sigma}^{\kappa}(\mu\sigma|\lambda\nu)\right]c_{\nu i}^{\kappa}=\epsilon_i^{\kappa}\sum_{\nu=1}^{K}S_{\mu\nu}c_{\nu i}^{\kappa}\quad\text{for}\quad\kappa=\alpha,\beta \tag{40}$$

where $P_{\lambda\sigma}^{\kappa}=\sum_{i}^{\text{occ}}c_{\sigma i}^{\kappa *}c_{\lambda i}^{\kappa}$ is κ-spin density matrix and $P^{T}=P^{\alpha}+P^{\beta}$ is total density matrix. Finally, the total electronic energy for open shell system in HF approximation is

calculated as

$$E_{\mathrm{el}}^{\mathrm{PN}} = \sum_{\mu=1}^{K}\sum_{\nu=1}^{K} P_{\mu\nu}^{T} H_{\mu\nu}^{\mathrm{core}} + \frac{1}{2}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K}\left(P_{\mu\nu}^{T}P_{\lambda\sigma}^{T} - P_{\mu\sigma}^{\alpha}P_{\lambda\nu}^{\alpha} - P_{\mu\sigma}^{\beta}P_{\lambda\nu}^{\beta}\right)(\mu\nu|\lambda\sigma) \quad (41)$$

Post Hartree–Fock methods

In HF approximation each electron feels the average field of other electrons but there is no dynamical correlation with remaining electrons. That causes difference between HF energy and exact non-relativistic energy of the system, defined as a correlation energy

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}} \quad (42)$$

Lack of dynamical correlation of electrons in HF method is caused by using single determinant representation of system wave function (31). Naturally, the improvement can be achieved by inclusion of more Slater determinants representing electronic excitations. This is base of configuration interaction method (CI) [44, 73, 75]. E_{exact} can be calculated by taking into account all possible Slater determinants as it is done in computationally very demanding full-CI (FCI) method.

Other approach is to calculate correlation energy using perturbation theory (PT). It can be shown that HF energy is exact to first order of PT and so usually second order is used to get an improvement. Computational algorithm for such calculations is known by abbreviation MPn (Moller–Pleset perturbation theory of n-th order [81]).

Most accurate, besides FCI, although also very computationally demanding is Coupled Clusters method (CC). Similar to CI it takes into account Slater determinants representing electron excitations but it uses different approximations [82, 83].

2.2.4 Density functional theory

Hartree–Fock method and its improvements such as CI, CC or MP2 are all based on many–particle wave function. Conceptually different approach is to evaluate total energy of the system and other properties directly from electron density. By this approach it is possible to get relatively precise method, which are less demanding than conventional post–HF methods (see Tab. 2).

Hohenberg–Kohn theorems

Rigorous density functional theory is based on two Hohenberg–Kohn theorems [85]. The first one states that ground state density of a system determines its external potential.

Method	Scaling
DFT	N^3
HF	N^4
MP2	N^5
MP3, CISD, CCSD, QCISD	N^6
MP4, CCSD(T), QCISD(T)	N^7
MP5, CISDT, CCSDT	N^8
MP6	N^9
MP7, CISDTQ, CCSDTQ	N^{10}

Table 2: Scaling behaviour of MO based ab initio methods and DFT as a function of number of electrons N_e [84]

Electron density includes the information about number of electrons and position and charge of atom nuclei

$$N = \int \rho(\mathbf{r}) d\mathbf{r}, \qquad \left.\frac{\partial \bar{\rho}(r_A)}{\partial r_A}\right|_{r_A=0} = -2Z_A \rho(\mathbf{r}_A) \tag{43}$$

where $\bar{\rho}(r_A)$ is spherically averaged density around nucleus A with charge Z_A. So the Hamiltonian of the whole system is determined by the density

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N_e} \Delta_i + \sum_{i=1}^{N_e} v(\mathbf{r}_i) + \sum_{i=1}^{N_e-1}\sum_{j=i+1}^{N_e} \frac{1}{r_{ij}} \equiv \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} \tag{44}$$

The second theorem concerns variational principle. For each density $\rho(\mathbf{r})$ that integrates to proper number of electrons N holds

$$E_0 \leq E[\rho(\mathbf{r})] \tag{45}$$

As result, total energy of the system is unique functional of ground state electron density. However, Hohenberg–Kohn theorems give no information about the form of this functional.

Kohn–Sham method

Kohn–Sham method [86] for finding ground state electron density is based on treating with non-interacting reference system of electron representing the real system. According to Hohenberg–Kohn theorems, density of reference and real system have to be the same. Density can be represented by one-electron orbitals and in analogy to HF method,

canonical Kohn–Sham orbital equations can be derived

$$\left[-\frac{1}{2}\Delta + v_{\text{eff}}(\mathbf{r})\right]\psi_i(\mathbf{x}) = \epsilon_i\psi_i(\mathbf{x}) \tag{46}$$

Density is constructed from one-electron functions $\psi_i(\mathbf{x})$ as a sum of their squares

$$\rho(\mathbf{r}) = \sum_{\kappa=\alpha,\beta}\sum_{i=1}^{N_e}|\psi_i(\mathbf{x})|^2 \tag{47}$$

and effective potential in Kohn–Sham equations (46) has a form

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{\text{xc}}(\mathbf{r}) \tag{48}$$

where $v(\mathbf{r})$ is potential of nuclei and $v_{\text{xc}}(\mathbf{r})$ is so called exchange–correlation potential. It is defined as functional variation of exchange–correlation energy E_{xc}, which includes differences between system of non-interacting electrons and the real system. Finally, total energy of the system is given by expression

$$E_{\text{el}}^{\text{KS}} = \sum_{i=1}^{N_e}\epsilon_i - \frac{1}{2}\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{\text{xc}}[\rho] - \int v_{\text{xc}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{49}$$

Kohn–Sham methods works, in electron gas approximation, with exact functional of kinetic energy. However, exact form of the exchange–correlation functional $E_{\text{xc}}[\rho]$ is not known and approximations are used.

Exchange–correlation functionals

Theoretically, the exchange–correlation functional depends not only on electron density but also on all its derivatives [87]. Simple approximation, popular in solid–state physics, is to neglect dependence on derivatives. This is called local density approximation (LDA)

$$E_{\text{xc}}^{\text{LDA}}[\rho] = \int\epsilon_{\text{xc}}[\rho]\rho(\mathbf{r})d\mathbf{r} \tag{50}$$

Functional $\epsilon_{\text{xc}}[\rho]$ can be divided on exchange and correlation part. Exchange part $\epsilon_{\text{x}}[\rho]$ can be written in analytical form

$$\epsilon_{\text{x}}[\rho] = -\frac{9\alpha}{8}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3}(\mathbf{r}) \tag{51}$$

where α is parameter specific for different derivations. It equals 1 for Slater's model, $\frac{2}{3}$ for Dirac's derivation [88] or $\frac{3}{4}$ for Xα model [89]. Analytical form for correlation part $\epsilon_{\mathrm{c}}[\rho]$ was derived as interpolation of quantum Monte–Carlo results by Vosko, Wilk and Nusair (VWN) [90].

LDA method can be extended to spin–polarized regime when two different densities ρ^{α}, ρ^{β} for two spin states are used instead of single one. This is known as Local Spin Density Approximation (LSDA) [87].

$$E_{\mathrm{xc}}^{\mathrm{LSDA}}[\rho^{\alpha},\rho^{\beta}] = \int \epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta})\rho(\mathbf{r})d\mathbf{r} \tag{52}$$

Exchange part of functional $\epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta})$ for spin κ has in this case following form

$$\epsilon_{\mathrm{x}}[\rho^{\kappa}] = \epsilon_{\mathrm{x}}^{0}[\rho] + \left[\epsilon_{\mathrm{x}}^{1}[\rho]0 - \epsilon_{\mathrm{x}}^{0}[\rho]\right]\left[\frac{(1+\kappa)^{4/3}+(1-\kappa)^{4/3}-2}{2(2^{1/3}-1)}\right] \tag{53}$$

where $\epsilon_{\mathrm{x}}^{0}[\rho]$ is exchange functional (51) for standard electron gas while $\epsilon_{\mathrm{x}}^{1}[\rho]$ refers to formula derived for "ferromagnetic" case.

Because LDA was derived from model of uniform electron gas it works well for extended solid state problems but not suitable for molecules. A way how to improve it is to construct exchange–correlation functional explicitly dependent not only on density but also on its first derivations. This approach is known as Generalized Gradient Approximation (GGA) [91–94]

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[\rho^{\alpha},\rho^{\beta},\nabla\rho^{\alpha},\nabla\rho^{\beta}] = \int \epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta},\nabla\rho^{\alpha},\nabla\rho^{\beta})\rho(\mathbf{r})d\mathbf{r} \tag{54}$$

Exchange–correlation function inside integral (54) is usually constructed from LSDA $\epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta})$ by adding a gradient–dependent correction.

$$\epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta},\nabla\rho^{\alpha},\nabla\rho^{\beta}) = \epsilon_{\mathrm{xc}}(\rho^{\alpha},\rho^{\beta}) + \Delta\epsilon_{\mathrm{xc}}\left[\frac{|\nabla\rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}\right] \tag{55}$$

Several different exchange functional forms for GGA ϵ_{x} was constructed now known by abbreviations like B [95], PW [96], mPW [97], O [98], X [99], B86 [100], P [101], PBE [102] or mPBE [103]. Popular correlation functionals are for example B88 [104], P86 [101], PW91 [105] or LYP [106]. Complete exchange–correlation functional is specified by concatenation of abbreviation for its exchange and correlation part. Recently, new types of functionals depending on higher order density gradients or on kinetic energy density were developed. These are called Meta-GGA functionals and among the most famous belong B95 [107], KCIS [108], PTSS [109] or VSXC [110].

While exchange functional is only approximative in DFT, it is computed exactly in HF method. So it is possible to include this exact HF exchange into DFT functionals. This approach is based on Adiabatic Connection Method (ACM) [111] and leads to hybrid DFT functionals. Probably the most popular one in calculations of organic molecules is B3LYP [95, 106, 107] functional with 3 empirical parameters $a = 0.20$, $b = 0.72$, $c = 0.81$

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = (1-a)E_{\mathrm{x}}^{\mathrm{LSDA}} + aE_{\mathrm{x}}^{\mathrm{HF}} + b\Delta E_{\mathrm{x}}^{\mathrm{B}} + (1-c)E_{\mathrm{c}}^{\mathrm{LSDA}} + cE_{\mathrm{c}}^{\mathrm{LYP}} \tag{56}$$

2.3 Hybrid QM/MM method

With great development of computational sources during last two decades and improvement of MM force fields it is possible to model and simulate structures of large molecular systems, especially biological macromolecules. However, MM approach can not describe chemical reactions and demanding of QM methods scale rapidly with size of the system (Tab. 2). Therefore several new methods based on division of the system smaller parts were developed recently.

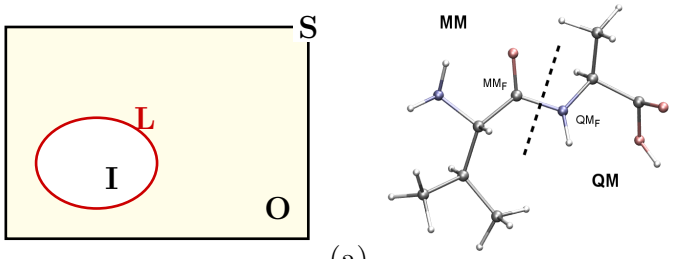
One group of methods, such as Divide and Conquer (DC) [112–116], Fragment Molecular Orbitals (FMO) [117] or its enhancement FMO–LCMO [118], divide a molecular system into several small fragments, calculate their energy by QM and then sum up all contributions to get total energy of the whole system. This approach works well for example for proteins which can be naturally divided into small fragments, i.e. aminoacids.

On the other hand, if there is only small part in large system where some interesting chemical changes occur, like catalysis of chemical reactions in an active sites of proteins, it is better to describe only this part by accurate QM method and rest of the system can be described by MM force field. This is basic idea of hybrid QM/MM method [44, 119–130]. Division of the system, so called QM/MM partitioning, is schematically shown in Fig. 9a where the labeling is also defined. The QM part is referred as inner part while the rest of the system is called outer part.

There are basically two approaches how to construct the total QM/MM energy [129, 130]. The first possibility is straightforward additive scheme where the MM contribution from outer part is added to QM energy of inner part

$$E_{\mathrm{QM/MM}}(\mathrm{S}) = E_{\mathrm{MM}}(\mathrm{O}) + E_{\mathrm{QM}}(\mathrm{I}+\mathrm{L}) + E_{\mathrm{QM-MM}}(\mathrm{S}) \tag{57}$$

The last term in (57) is mutual interaction of inner and outer part. This interaction is



(a) (b)

Figure 9: QM/MM partitioning: (a) System S divided into inner part I with boundary layer L and outer part O, (b) Boundary cutting covalent bond between QM frontier (QM_F) and MM frontier (MM_F)

sometimes hard to evaluate and so subtractive scheme for QM/MM energy can be used.

$$E_{\mathrm{QM/MM}}(\mathrm{S}) = E_{\mathrm{MM}}(\mathrm{S}) + E_{\mathrm{QM}}(\mathrm{I}+\mathrm{L}) - E_{\mathrm{MM}}(\mathrm{I}+\mathrm{L}) \tag{58}$$

Interaction of two system parts is included implicitly in subtraction of MM contribution of inner part from MM energy of the total system. Disadvantage of this approach is that MM parametrization has to be done also for inner part of the system that is described by QM.

Forces, needed for geometry optimization and molecular dynamics, are calculated in standard way as first derivation of total QM/MM energy according to position. However, forces from inner part (and outer part in additive scheme case) needs to be transformed to coordinates of the whole system [131, 132].

$$\mathbf{F}_{\mathrm{QM/MM}}(\mathrm{S}) = \nabla E_{\mathrm{MM}}(\mathrm{S}) + [\nabla E_{\mathrm{QM}}(\mathrm{I}+\mathrm{L}) - \nabla E_{\mathrm{MM}}(\mathrm{I}+\mathrm{L})] \cdot \mathbb{J} \tag{59}$$

Jacobian matrix $\mathbb{J} = \partial\mathbf{R}(\mathrm{S})/\partial\mathbf{R}(\mathrm{I}+\mathrm{L})$ is non–trivial only if there is a chemical bond between between inner and outer part of the system. This is known as covalent embedding.

2.3.1 Covalent embedding

If the boundary between inner and outer part cross a covalent bond as in Fig. 9b, a special treatment has to be applied to saturate that bond in QM calculation. Atoms participating in cut bond are usually names MM frontier (MM_F) and QM frontier (QM_F). There are several more or less complex approaches how to treat covalent embedding. The most popular are link atom method and localized frozen orbital approach.

Figure 10: Link atom approach for treating covalent embedding

Link atom method

In link atom approach, the cut QM_F–MM_F is saturated by monovalent atom or group of atoms [120, 121]. In most cases, hydrogen atom is used for that purpose. However, artificial degrees of freedom are introduced into the system by adding link atom and they have to be transformed out from QM/MM force vector.

To avoid this problem, position of link atom is usually constrained to QM_F–MM_F line as is illustrated in Fig. 10. The position is then unambiguously determined from coordinates of MM and QM frontiers and so no independent degrees of freedom are added into the system.

$$\mathbf{R}(\mathrm{L}) = \mathbf{R}(\mathrm{QM_F}) + \alpha_\mathrm{L}(\mathbf{R}(\mathrm{MM_F} - \mathbf{R}(\mathrm{QM_F}) \tag{60}$$

Constant α_L is ratio between QM_F–MM_F bond length in the whole system and QM_F–L bond length in QM part determined before starting QM/MM calculation.

Localized frozen orbitals approach

Drawback of link atom method described above is that cut QM_F–MM_F bond is saturated by chemically different atom or atomic group. To keep chemical character of the bond, localized frozen orbitals can be used [133–135].

In that method, localized orbitals optimized for given system are put on MM_F and kept frozen. Only orbital directly participating in the QM_F–MM_F bond is used freely in SCF procedure during QM calculation. The same approach is used for covalent embedding in FMO method [117]. Although this technique seems to be more accurate than link atom method, it was shown that there are no big differences between these to approaches [136].

2.3.2 Charge embedding

Non-bonding interactions, i.e. Coulombic electrostatic interaction and short range Van der Waals interactions, are main contributions to interaction term $E_{\mathrm{QM-MM}}(\mathrm{S})$ in (57). While the VdW contribution is evaluated from FF, there are two different ways how to

treat Coulombic interaction.

Mechanical embedding

The simplest approach is to calculate electrostatic interaction in classical way, that means define atomic charges of all atoms of the whole system and evaluate electrostatic contribution to total energy using Coulomb law

$$E_{\text{m.embd}} = \sum_{A=1}^{N_{MM}} \sum_{B=1}^{N_{QM}} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{61}$$

Atomic charges in QM part are set before the QM/MM calculation. If optimization or MD is performed these charges can remain unchanged or be updated in each step by their calculation from QM electron density. In that case, various population analysis methods can be used, such as Mulliken's method [137], NPA [138] or ESP fitting [139,140], depending on required accuracy, consistency with used force field and computation cost. Nevertheless, the mutual QM–MM interaction remains on MM level and electron density of QM part does not feel surrounding MM atoms.

Electronic embedding

More accurate description is reached by putting atomic charges from MM atoms into QM Hamiltonian (25) as a background charges. Electrostatic interaction is then evaluated self–consistently during SCF procedure through contributing one–electron integrals

$$E_{\text{e.embd}} = \sum_{A=1}^{N_{\text{MM}}} \int_V \frac{Z_A \rho(\mathbf{r})}{|\mathbf{R}_A - \mathbf{r}|} d\mathbf{r} \tag{62}$$

This leads to response of QM electron density to its electrostatic environment resulting in its polarization. If link atoms are used in the system, atomic charge on MM frontier has to be screened or set to zero. This is necessary because of nonphysically short distance between MM frontier and link atom (see Fig. 10) that would cause artificial polarization of the electron density in this area.

Polarized embedding

Although electronic embedding approach is more accurate than mechanical embedding, description of the system is unbalanced because inner part is polarized while outer part has fixed point charges. To improve that, it is possible to include polarization terms into MM part, too. Basically, there are three approaches, which are mainly used: fluctuating charge method, Drude oscillator description or induce dipole models [141].

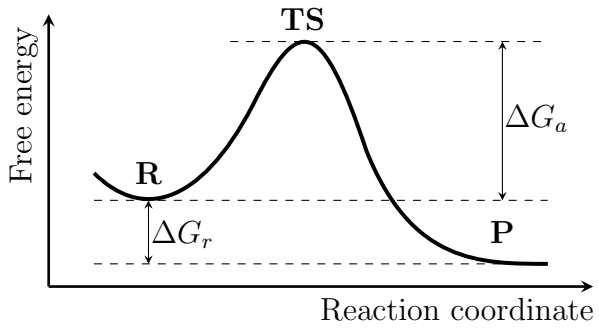


Figure 11: Schematic illustration of reaction profile. Reactant **R** overcomes activation barrier ΔG_a, transforms to transition state **TS** and subsequently change to product **P**.

In fluctuating charge method, atomic charges are varied iteratively according to atom electronegativities to get optimized electrostatic energy of the system [142–145].

Drude oscilators, also known as shell model, uses an additional point charges restrained to atomic centers by harmonic potential. Charge magnitudes and harmonic force constants are optimized to reproduce experimental data, i.e. atomic and molecular polarizabilities and energies [146–148].

The most often used is induce dipole method. Point polarizabilities are assigned to each atom and dipoles are induced by electrostatic field from permanent atomic charges, $\mu_i^{\text{ind}} = \alpha_i(\mathbf{E}_i^0 + \mathbf{E}_i^{\text{pol}})$. Final electrostatic field is get consistently by iterative approach and contribution of the polarization energy to the total non-bonded energy is [149]

$$E_{\text{pol}} = -\frac{1}{2}\sum_{i=1}^{N} \mu_i \cdot \mathbf{E}_i^0 \tag{63}$$

2.4 Free energy calculation

Using computational methods described above, that is MM, QM or their combination – hybrid QM/MM, total energy of different molecular configurations can be directly evaluated and compared. However, for discussion of chemical stability and kinetics, differences of free energies are important [150, 151]. In chemistry temperature T and pressure p of the system are often controlled. Using Legendre transformation, it can be derived that thermodynamic potential with these natural variables is Gibbs free energy [152, 153]

$$G = U + pV - TS \tag{64}$$

Keeping T and p constant, G decreases for spontaneous processes and seeks a min-

imum. Difference of thermodynamical potential between two states is independent on transition path between them. Therefore chemical reactions can be characterized by reaction energy $\Delta G_r = \sum_i G_i^P - \sum_i G_i^R$, that is the difference of product P and reactant R free energies. While the ΔG_r provide information about spontaneity, kinetics of the reaction is governed by its barrier ΔG_a (see Fig. 11).

From Eyring's transition state theory follows that rate of a reaction is exponential function of the reaction barrier [154–156]. There is a relation for temperature dependent rate constant

$$k(T) = \frac{k_B T}{h} \frac{Z^{\mathrm{TS}}}{Z^{\mathrm{R}}} \exp\left[-\frac{\Delta E_a}{k_B T}\right] = \frac{k_B T}{h} \exp\left[-\frac{\Delta G_a}{RT}\right] \tag{65}$$

where k_B is Boltzmann constant, R is universal gas constant and Z are partition functions of reactant and transition state of the reaction. Transition state has one imaginary frequency related with an antisymmetric stretching movement between reactant and product structure.

From computational point of view the evaluation of Gibbs free energy is difficult task because entropy included in (64) is related to accessible volume of configuration space. Therefore, sufficient sampling of configuration space is required and that is computationally very demanding.

2.4.1 Corrections calculated on optimized structures

In demanding QM calculations, corrections to total energies of optimized structures are computed using harmonic frequency analysis [157]. These are then used for desired free energy differences, for example reaction energy

$$\Delta G_r = \sum_{i=1}^{N_P} (E_i^P + G_{i,\mathrm{corr}}^P) - \sum_{i=1}^{N_R} (E_i^R + G_{i,\mathrm{corr}}^R) \tag{66}$$

The free energy correction, G_{corr}, is calculated in ideal gas approximated where non–interacting particles are assumed. When the partition function $Q(V,T)$ of the molecule is constructed from its translational, rotational and vibrational contributions, internal energy U and entropy S can be evaluated according to following relations [157]

$$S = Nk_B + Nk_B \ln\left(\frac{Q}{N}\right) + Nk_B T \left(\frac{\partial \ln Q}{\partial T}\right)_V \tag{67}$$

$$U = Nk_B T^2 \left(\frac{\partial \ln Q}{\partial T}\right) \tag{68}$$

Finally, using (64) the desired free energy correction is obtained from these functions

$$G_{\text{corr}} = U + k_B T - TS \tag{69}$$

2.4.2 Corrections calculated by MD sampling

As was already explained, correct way how to calculate free energy is to sample configuration space. Molecular dynamics (MD) or Monte Carlo (MC) algorithms can be applied. However, this is computationally very demanding and therefore it was for a long time domain only of MM. Nowadays, semiempirical QM or hybrid QM/MM methods are also used.

The problem is that different configurations of molecular system are separated by energy barriers that are poorly sampled or even inaccessible during direct MD. Several different techniques and algorithms are available to overcome this [49, 158, 159]. Most often used are thermodynamic integration (TI), free energy perturbation (FEP), umbrella sampling (US) or metadynamics. Umbrella sampling technique which is used in this work is described below.

Because computer simulations are mainly performed in canonical NVT ensemble where volume V is kept fixed rather than pressure, computational methods work with Helmholtz free energy $A = U - TS$ instead of G.

Umbrella sampling

Umbrella sampling (US) is free energy calculation approach similar to thermodynamic integration (TI) [49, 159–163]. Reaction coordinate ξ is divided into N_w windows and these are sampled individually. To keep system in i-th window, a bias potential w_i is set to restrain reaction coordinate. Typically, a harmonic form of bias is used:

$$w_i(\xi) = \frac{1}{2} K_i (\xi - \xi_i^0)^2. \tag{70}$$

Force constant K_i has to be chosen carefully to get efficient sampling around pre-set value of reaction coordinate ξ_i^0. When the force constant K increase ad infinitum, restrain becomes to constrain and umbrella sampling change to thermodynamic integration method [162].

Biased distribution for each window, $\mathcal{P}_i^b$, are constructed as normalized histograms of reaction coordinate sampling. From these distributions, unbiased free energies can be calculated [163]

$$A_i(\xi) = -\frac{1}{\beta} \ln \mathcal{P}_i^b(\xi) - w_i(\xi) + F_i \tag{71}$$

where $F_i = -\frac{1}{\beta} \ln \langle \exp [-\beta w_i(\xi)] \rangle$ are free energies related to bias potentials. F_i are calculated from overlaps of distributions from neighbouring windows. Free energy profiles can be constructed in several different ways. Among the most popular belong Weighted Histogram Analysis Method (WHAM, [164, 165]) and its modification, Umbrella Integration (UI, [166, 167]).

Weighted histogram analysis method

Free energy profile is connected with total unbiased distribution $\mathcal{P}^u$ by relation $A(\xi) = -\frac{1}{\beta} \ln \mathcal{P}^u(\xi)$. In umbrella sampling method, $\mathcal{P}^u$ is not available directly but it can be calculated as weighted average of biased distribution from individual windows

$$\mathcal{P}^u(\xi) = \sum_{i=1}^{N_w} p_i(\xi) \mathcal{P}_i^b(\xi) \tag{72}$$

Weights p_i fulfill normalization condition $\sum_{i=1}^{N_w} p_i = 1$ and are chosen in a way to minimize statistical error of $\mathcal{P}^u(\xi)$ [164]

$$p_i(\xi) = \frac{N_i \exp [-\beta(w_i(\xi) - F_i)]}{\sum_{j=1}^{N_w} N_j \exp [-\beta(w_i(\xi) - F_i)]} \tag{73}$$

N_i in (73) represents number of samples in i-th window used for construction of histogram $\mathcal{P}_i^b$. Free energies F_i are then calculated by integration of total unbiased distribution weighted by Boltzmann factor of bias w_i

$$\exp(-\beta F_i) = \int \mathcal{P}^u(\xi) \exp [-\beta w_i(\xi)] d\xi \tag{74}$$

Equations (72) and (74) are coupled together through (73) and therefore they have to be solved iteratively until sufficient convergence in F_i is reached. This procedure is called Weighted Histogram Analysis Method (WHAM, [164, 165]).

Umbrella integration

Umbrella Integration (UI) is approximation to WHAM derived by J. Kästner and W. Thiel [166, 167] where harmonic form of bias (70) is assumed. In contrast to WHAM, mean forces are taken into account as is typical for thermodynamic integration

$$\frac{\partial A}{\partial \xi} = \sum_{i=1}^{N_w} p_i(\xi) \frac{\partial A_i^u}{\partial \xi} \tag{75}$$

The weights p_i in sum (75) are for each histogram bin of reaction coordinate determined directly from biased distributions $\mathcal{P}_i^b$

$$p_i(\xi) = \frac{N_i \mathcal{P}_i^b(\xi)}{\sum_{j=1}^{N_w} N_j \mathcal{P}_j^b(\xi)} \tag{76}$$

Advantage of this approach is that unbiased mean forces from individual windows are independent on F_i and can be calculated directly from biased distributions $\mathcal{P}_i^b$

$$\frac{\partial A_i^u}{\partial \xi} = -\frac{1}{\beta} \frac{\partial \ln \mathcal{P}_i^b(\xi)}{\partial \xi} - \frac{dw_i}{d\xi} \tag{77}$$

When the windows are small enough and appropriate bias force constants K_i are chosen, histogram $\mathcal{P}_i^b$ has only one peak and can be well approximated by Gaussian normal distribution

$$\mathcal{P}_i^b(\xi) = \frac{1}{\sigma_i^b \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{\xi - \bar{\xi}_i^b}{\sigma_i^b} \right)^2 \right] \tag{78}$$

Because harmonic shape of bias was assumed, equation (75) can be rewritten using (70) and (78) into form

$$\frac{\partial A_i^u}{\partial \xi} = \frac{\xi - \bar{\xi}_i^b}{\beta (\sigma_i^b)^2} - K(\xi - \xi_i^0) \tag{79}$$

After construction of total mean force (75), the free energy profile resp. potential of mean force (PMF) is obtained by simple integration $A(\xi) = \int_0^\xi \frac{\partial A(\xi')}{\partial \xi'} d\xi'$. Comparing to WHAM, umbrella integration method does not require overlapping distributions, however, they are desirable for better accuracy. As a result of approximating $\mathcal{P}_i^b$ by normal distribution, the $A(\xi)$ is a smooth curve.

3 QM/MM implementation

QM/MM methodology was during last decade implemented in many software packages. In many cases it was added as an extension to already well established QM and MM programs. However, such approach has its drawbacks in accuracy and sophistication of these extensions. In MM programs (e.g. Amber [168], Gromacs [169]) usually only empirical or semi-empirical QM methods were implemented because of algorithm complexity of advanced methods. On the contrary, only basic, mainly universal, force fields are supported by QM software (e.g. ADF [170], Gamess [171], Gaussian [76], MolCas [172], NWChem [173], Q-Chem [174], Turbomole [175]).

Because of complexity of MM and especially QM methods and their implementation, there are not so many standalone QM/MM programs (e.g. QSite [176, 177]). However, it is possible to create software layer operating with existing QM and MM software (e.g. ChemShell [178], QMMM [179], ComQum [180–182]). Variety of QM methods can be then combined with different force fields because energy and forces are calculated for each part of the system by specific QM or MM program. Disadvantage of this approach is necessity to implement, besides QM/MM coupling, all algorithms needed for geometry optimization or molecular dynamics.

This program layer or QM/MM interface approach is employed in this work. QM/MM coupling was implemented in our own software called *QMS*–Uni (Universal Program Layer for Coupling QM and MM Software) in order to get flexible, easy controlled program suitable for working with organometallic systems in bioorganic environment. External QM and MM programs are called for energy and force evaluation as described below. More details can be found in a *QMS*–Uni manual [183].

3.1 Program structure

Program *QMS*–Uni consists of three main binary modules called *QMS*–Prep, *QMS*–Step and *QMS*–Traj supported by several small scripts handling with external QM and MM software. While the *QMS*–Prep and *QMS*–Traj are standalone utilities for data preparation and analysis of trajectory file, communication with *QMS*–Step is interfaced by *QMS*–Run script. This structure is schematically shown in Fig. 12 where user interface is marked by red frame while blue frame indicated internal communication.

Binary modules are coded in C program language and all scripts are written in BASH. The entire program is design to run on UNIX–like systems and support MM program Amber [168, 184] and QM programs Gamess [171], Gaussian [76, 185], MolPro [186] and

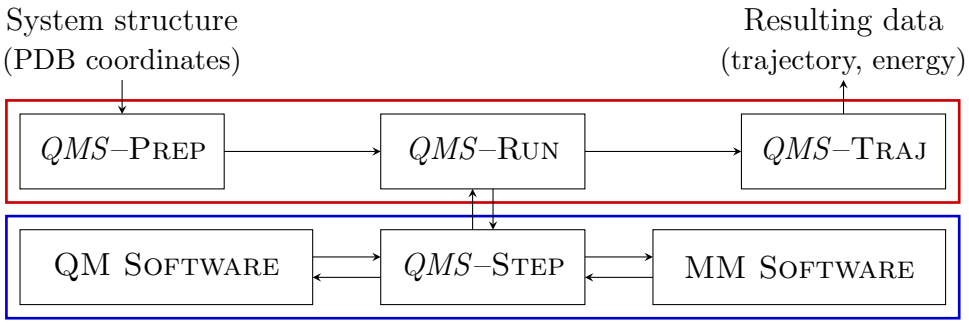


Figure 12: Schematic structure of *QMS*–Uni program layer. Red frame indicated user interface while the the blue frame is internal comunication. The main computational module is *QMS*–Step.

Turbomole [175]. Details about implemented QM/MM coupling and computational methods are described below.

3.1.1 Module functions

Module *QMS*–Prep is interactive program that helps user to prepare all necessary input files for QM/MM calculation. At the beginning, structure of the whole system in PDB format is expected and QM/MM partitioning is done by selecting QM atoms. Because *QMS*–Uni support multi–QM/MM computations, more than one QM part can be defined if desired. Link atom approach (60) is used for covalent embedding and correct scaling factors has to be set here. Both mechanical and electrostatic embedding are supported (see Sec. 2.3.2). In electronic embedding it is possible to select layer of MM atoms around QM part(s) that will be included in QM Hamiltonian. At last, type of calculation is set. *QMS*–Uni can perform geometry optimization, molecular dynamics or harmonic frequency analysis on one structure.

Subtractive energy scheme (58) is used for QM/MM coupling in *QMS*–Uni. Therefore, input files for all system parts including MM parametrization of QM parts have to be prepared before staring QM/MM calculation. The calculation is started by *QMS*–Run and performed by *QMS*–Step module.

Various information such as structure coordinates, energies or values of defined internal coordinate are saved into special files during geometry optimization or MD. For manipulation with these files is available *QMS*–Traj module. It can extract trajectory of specified QM part only, structures from given time or with minimal energy. Other useful functions can be found in the manual [183].

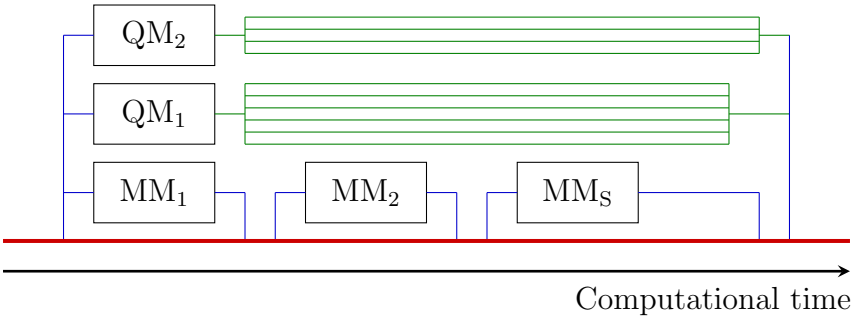


Figure 13: Schematic illustration of *QMS*–Step parallelization. Master thread (red) calls QM and MM programs, gathers data from all parts and compile them. QM calculation of each inner part can be further parallelized by external software (green, 4 CPUs and 6 CPUs per 1 QM part in this example).

3.1.2 Parallelization

In order to increase efficiency of the QM/MM calculation, evaluation of QM electron densities is done parallel to MM calculations as is illustrated in Fig. 13. POSIX Thread library [187] is employed to create new thread for each part of the system when external programs are called. While the demanding QM parts are being computed much faster MM calculations for all parts are done in series.

Standard QM software can usually perform calculation on more than one CPU to speed up the SCF iterative cycle. This level of parallelization can be also employed in *QMS*–Uni by specifying required number of CPUs in relevant QM input file. Different QM parts can be evaluated using different number of CPUs as indicated by green part in Fig. 13. Other speed up of SCF can be reached by using wave function from previous step as an initial guess for calculation during geometry optimization or MD.

In QM/MM MD simulations, it is possible to do sampling at different level of theory of QM part(s) of the system. This is often needed in case of large QM parts where compromise between accuracy and speed has to be done in choice of method and basis set. In *QMS*–Uni trajectory is sampled automatically during the simulation every given number of steps. Structures are saved to FIFO container [188] and their energies are evaluated one by one in thread parallel to MD faster run.

3.2 Computational methods

Subtractive energy scheme (58) is implemented in *QMS*–Uni. Relevant forces are calculated according to (59) and used in energy minimization algorithms or in molecular

dynamics. These methods are coded in *QMS*–Step module.

3.2.1 Geometry optimization

In geometry optimization methods, structure of the system is changed to reach minimum of the potential energy function. In each step of the optimization cycle, a search direction **d** is chosen in which system coordinates are subsequently moved to get new structure with lower energy [189]

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{d}_n \tag{80}$$

Step length α in (80) should be set in each geometry optimization step n to minimize structure energy in direction **d**. However, this would be computationally very demanding and so an approximative value of α is set by line–search algorithms.

Search direction **d**, chosen by geometry optimization procedure, has general form

$$\mathbf{d} = -\mathbb{B}^{-1}\nabla E \tag{81}$$

where matrix $\mathbb{B}$ is symmetric and nonsingular matrix. In Newton's quadratic method $\mathbb{B}$ is exact Hessian of energy function, however, less computational demanding algorithms use simplified matrices [190].

There are three different optimization methods implemented in *QMS*–Uni: Steepest descent (SD) algorithm, Conjugate Gradients (CG) and Limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method. Because QM/MM calculations are computational demanding, all these methods do not use higher order derivation of total energy than gradient.

Line–search algorithms

Ideal line–search algorithm should find a global minimum of energy function in given direction, i.e. minimize the function

$$f(\alpha) = E(\mathbf{x} + \alpha\mathbf{d}) \tag{82}$$

To avoid large computational demand it is not possible to accept mere decrease of $f(\alpha)$ because that can not assure convergence of geometry optimization algorithm to real minimum of energy. Therefore, a sufficient decrease condition (83) and curvature condition (84), together known as Wolfe condition [191], have to be fulfilled

$$E(\mathbf{x}_n + \alpha_n \mathbf{d}_n) \leq E(\mathbf{x}_n) + c_1 \alpha_n \nabla E(\mathbf{x}_n)^T \cdot \mathbf{d}_n \tag{83}$$

$$\nabla E(\mathbf{x}_n + \alpha_n \mathbf{d}_n)^T \cdot \mathbf{d}_n \geq c_2 \nabla E(\mathbf{x}_n)^T \cdot \mathbf{d}_n \tag{84}$$

Constants c_1, c_2 have to be set to hold $0 < c_1 < c_2 < 1$ relation and they are usually chosen as $c_1 = 10^{-4}$ and $c_2 = 0.9$ for Newton and quasi–Newton methods or $c_2 = 0.1$ for nonlinear conjugate gradient method.

Wolfe condition line–search is implemented in *QMS*–Unitogether with backtracking algorithm and extrapolation schemes. Although these approaches do not control sufficient decrease condition, they are faster to estimate value of α.

Steepest descent

SD is a simplest optimization algorithm for finding local minimum of potential energy function. Matrix $\mathbb{B}$ in (81) is simply unit matrix $\mathbb{I}$ and so the energy function $E(\mathbf{x})$ is minimized in direction opposite to its gradient in each step

$$\mathbf{x}_{n+1}^{\mathrm{SD}} = \mathbf{x}_n - \alpha_n \nabla(\mathbf{x}_n) \tag{85}$$

Steepest descent always gets structure with lower energy in subsequent optimization step, however, convergence is very slow near the minimum region of energy where gradients are small. That is why SD is recommended for pre–optimization of the system only, before more sophisticated method is used to reach minimum.

Nonlinear conjugate gradient

CG, as well as SD, is an optimization method based on first derivations of energy function. First optimization step is the same as in SD method (85) but direction in other steps is corrected by direction from the previous step

$$\mathbf{x}_{n+1}^{\mathrm{CG}} = \mathbf{x}_n - \alpha_n [\nabla E(\mathbf{x}_n) - \beta_n \mathbf{d}_{n-1}] \tag{86}$$

There are three different ways how to set a conjugation factor β, all of them are implemented in *QMS*–Uni code: Fletcher–Reeves (FR) [192], Polak–Ribière (PR) [193] and Hestenes–Stiefel (HS) [194] derivation

$$\beta_n^{\mathrm{FR}} = \frac{\mathbf{F}_n^T \cdot \mathbf{F}_n}{\mathbf{F}_{n-1}^T \cdot \mathbf{F}_{n-1}}, \quad \beta_n^{\mathrm{PR}} = \frac{\mathbf{F}_n^T \cdot (\mathbf{F}_n - \mathbf{F}_{n-1})}{\mathbf{F}_{n-1}^T \cdot \mathbf{F}_{n-1}}, \quad \beta_n^{\mathrm{HS}} = \frac{\mathbf{F}_n^T \cdot (\mathbf{F}_n - \mathbf{F}_{n-1})}{\mathbf{d}_n \cdot (\mathbf{F}_n - \mathbf{F}_{n-1})} \tag{87}$$

For the sake of simplicity, negative gradient was marked as $\mathbf{F} = -\nabla E(\mathbf{x})$ in (87).

If the minimizing function is an exact quadratic form of N variables, CG method reaches minimum in N steps [189, 190]. Although energy function of molecular system has general, often complicated, non–quadratic shape, CG method has still much better

convergence behaviour than SD in region close to the minimum.

Limited memory BFGS

L-BFGS is a memory saving variant of original Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi–Newton optimization method [195]. In BFGS method, inversion Hessian matrix $\mathbb{B}^{-1}$ from (81) is approximated by matrix $\mathbb{C}$ constructed from energy gradients

$$\mathbf{x}_{n+1}^{\mathrm{L-BFGS}} = \mathbf{x}_n - \alpha_n \mathbb{C}_n \cdot \nabla E(\mathbf{x}_n) \tag{88}$$

However, full matrix $\mathbb{C}$ can be kept in computer memory only for system with limited size. Therefore, in L-BFGS modification only two vectors $\mathbf{s}$ and $\mathbf{y}$ from last N_s steps are saved

$$\mathbf{s}_n = \mathbf{x}_{n+1} - \mathbf{x}_n \tag{89}$$

$$\mathbf{y}_n = \nabla E(\mathbf{x}_{n+1}) - \nabla E(\mathbf{x}_n) \tag{90}$$

Matrix $\mathbb{C}$ is then constructed from these vectors

$$\begin{aligned}\mathbb{C}_{n+1} =& \rho_n \mathbf{s}_n \cdot \mathbf{s}_n^T + \\ &+ \sum_{i=1}^{N_s} \rho_{n-i} \left(\prod_{j=0}^{i-1} \mathbb{V}_{n-j} \right) \mathbf{s}_{n-i} \cdot \mathbf{s}_{n-i}^T \left(\prod_{j=i-1}^{N_s} \mathbb{V}_{n-N_s+j} \right) + \\ &+ \left(\prod_{j=0}^{N_s} \mathbb{V}_{n-j} \right) \mathbb{C}_0 \left(\prod_{n-N_s+j}^{N_s} \mathbb{V}_{n-N_s+j} \right)\end{aligned} \tag{91}$$

where $\rho_n = (\mathbf{y}_n^T \cdot \mathbf{s}_n)^{-1}$, $\mathbb{C}_0$ is symmetric and positive definite starting matrix (identity matrix $\mathbb{I}$ as a simplest choice is used in *QMS*–Uni code) and matrices $\mathbb{V}_n$ are defined as $\mathbb{V}_n = \mathbb{I} - \rho_n \mathbf{y}_n \cdot \mathbf{s}_n^T$.

Microiterations

Optimization algorithms described above minimize the total QM/MM energy function. The system is, however, divided into two or more parts that are differently flexible and whose energy evaluation is differently computational demanding. Therefore, in *QMS*–Uni code, individual parts can be optimized with fixed rest of the system or microiteration optimization technique can be used [131].

Microiteration–based optimization exploits the fact that MM part of the system has much more degrees of freedom than QM part, is more flexible, but on the other hand its energy evaluation is much less demanding. Each optimization step is divided into two parts: first MM part is optimized to nearest minimum in MM coordinates only ($\mathbf{x}^{\mathrm{MM}}$)

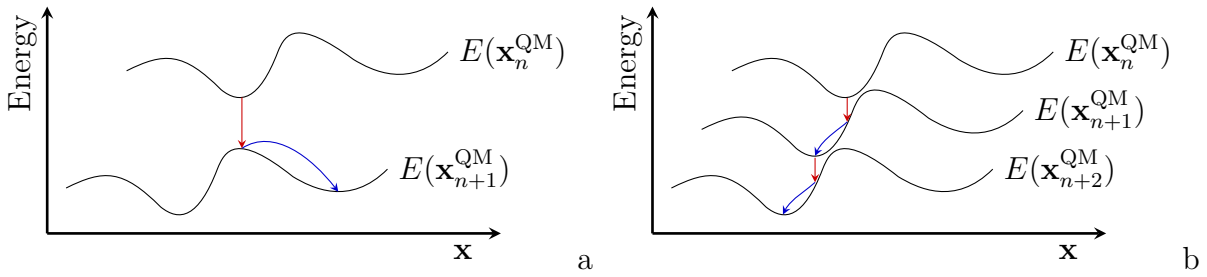


Figure 14: QM/MM microiteration–based optimization procedure. (a) large step of MM optimizer (red arrow) can mislead the system to different minimum, (b) correct behaviour with properly chosen setting.

and then one optimization step is performed in QM coordinates, $\mathbf{x}^{\mathrm{QM}}$ but using gradient of total QM/MM energy. Series of MM steps is called microiterations.

Microiterations are well defined for mechanical embedding where energy of QM part does not depend on MM coordinates and vice versa. This is no longer true for electronic embedding where point–charges from MM atoms are included in QM Hamiltonian (see Sec. 2.3.2). Moreover, optimization step has to be chosen carefully to avoid misleading of optimization procedure to different energy minimum by microiterations [131] as illustrated in Fig. 14. Proper line–search algorithm as well as maximum step length has to be chosen to maximize efficiency of microiteration method.

3.2.2 Molecular dynamics

Born–Oppenheimer molecular dynamics is implemented in *QMS*–Uni code. That means that total QM/MM energy (57) and force (59) are used for integration of classical equation of motions as it is described in Sec. 2.1.4. Velocity Verlet algorithm, that is coded in *QMS*–Step module, can be directly used for MD in microcanonical (NVE) ensemble.

For simulation of real molecular systems is more important MD in canonical (NVT) ensemble. For that purpose, three simple thermostat algorithms are implemented to keep temperature of studied systems at specified value: Andersen thermostat, velocity scaling algorithm and Berendsen thermostat.

Andersen thermostat

Andersen thermostat [196] is a stochastic algorithm that correctly sample NVT ensemble. At each step of MD several particles in the system are chosen randomly and their tem-

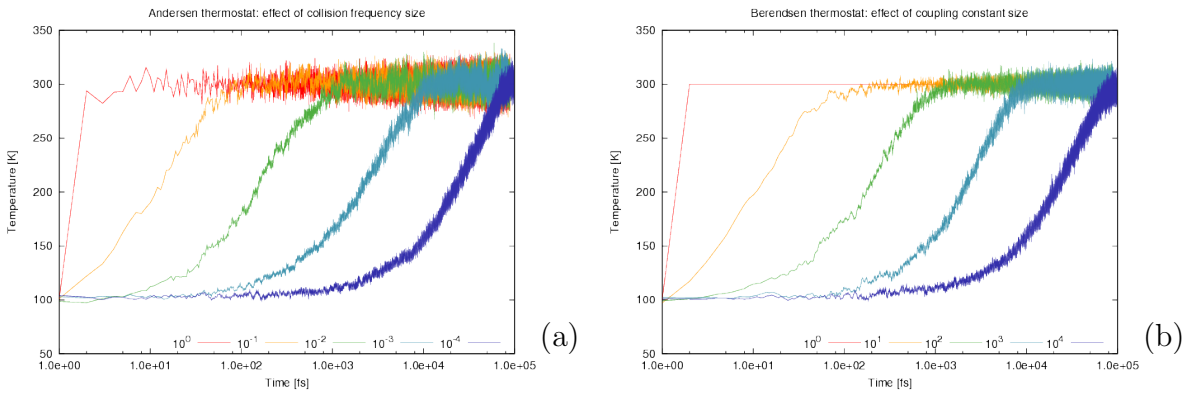


Figure 15: Influence of collision frequency parameter resp. coupling parameter on (a) Andersen thermostat and (b) Berendsen thermostat. Simulated system consisted of 369 water molecules in rectangular box in PBC tempered to 100 K. Time evolution of temperature is shown in logarithmic scale from the time when the pre–set temperature T_0 changed from 100 K to 300 K.

perature is reassign according to Maxwell–Boltzmann distribution for given temperature

$$\mathcal{P}(v_i) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{m v_i^2}{2 k_B T}\right) \tag{92}$$

Strength of system coupling to the heat bath, i.e. number of particles chosen for velocity reassigning, is specified by collision frequency ν. Velocity of a particle is reset if $\nu \Delta t$ is less than random number from [0,1] interval.

Velocity scaling

The simplest algorithm for keeping temperature is to rescale velocities of all particles in the system by factor

$$\lambda = \sqrt{\frac{T_0}{T(t)}} \tag{93}$$

where T_0 is desired temperature and $T(t)$ is actual temperature at time t [197]. However, velocity scaling algorithm is too strict, it does not allow any temperature fluctuation, is not time–reversible and do not follow canonical ensemble. That is why this algorithm is scarcely used, mainly for pre–equilibration of the system.

Berendsen thermostat

More sophisticated version of velocity scaling is Berendsen thermostat [198]. Coupling to

the heat bath is described by differential equation for temperature

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_0 - T(t)) \tag{94}$$

where τ is a coupling constant. Solution is an exponential decay to pre–set temperature T_0 and scaling factor for velocities is given by expression

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau}\left(\frac{T}{T_0} - 1\right)} \tag{95}$$

Strength of coupling between the system and heat bath is specified by value of τ. The bigger value, the weaker coupling and for $\tau \to \infty$ system converts to microcanonical ensemble. Although better then simple velocity scaling, Berendsen thermostat does not conform canonical ensemble neither.

Comparison of all these thermostats is shown on Fig. 15 where time response of temperature after sudden change of T_0 from 100 K to 300 K is plotted for different collision frequencies resp. coupling constants. All thermostats heat the system correctly to 300 K. There are larger fluctuations in temperature in case of Anderesen thermostat than Berendsen one. Velocity scaling is limit case of Berendsen thermostat for coupling constant equals 1 fs^{-1} where no fluctuations are allowed.

4 Reaction mechanism of Ru(II) complexes

In this section, the theory and computational approaches introduced in previous sections are used to study reaction mechanism of piano–stool Ru(II) complexes. The $[Ru^{II}(\eta^6$-benzene)(en)Cl$]^+$ is chosen as a representant of given group of complexes and its interaction with DNA is studied. As was already explained in Sec. 1.3, hydration reaction, during which the chloride anion is replaced by aqua ligand, proceeds at first. This reaction is in further text marked as **R0** and experimental value of its rate constant is used for checking accuracy of chosen computational method. Resulting aqua complex, $[Ru^{II}(\eta^6$-benzene)(en)$(H_2O)]^{2+}$ then interacts with DNA.

The binding to DNA (denoted as **R1**) can proceed by two different reaction pathways resulting in Ru(II)-N7 guanine monoadduct. One possibility is direct bonding to N7 guanine position going through single energy barrier. The other is two step mechanism when first the Ru(II) complex coordinates to O6 guanine position and then move from this intermediate state to final N7 position. Both pathways were explored and their thermodynamic as well as kinetic preference is discussed in Sec. 4.3.

Next step is formation of intrastrand cross–linked structure where Ru(II) connects N7 nitrogens from two adjacent guanines. This last step of reaction mechanism, which is studied here, is marked **R2**. η^6 coordination of benzene ligand is transformed to η^2 in order to create vacant position for N7 of the second guanine and water molecule released in **R1** is coordinated back. Three types of cross–linked structures can be distinguished where only benzene ligand, water molecule or both of them are coordinated to Ru(II).

4.1 Computational model

Ru(II) complexes are at QM level described by DFT method with hybrid B3LYP functional. Choice of this method is based on comparison of several computational levels for hydration reaction (**R0**) rate constant calculation where experimental value is available (see Sec. 4.2 for details). The 6-31G(d) basis set is used for geometry optimization and larger 6-31++G(2df,2pd) for electron analysis. Stuttgart–Dresden pseudopotential [78, 199, 200] are used for 28 core electrons on Ru atom and 10 electrons on Cl to include partially relativistic corrections and speed up the calculation. Original set of pseudo–orbitals is augmented by polarization functions with exponents $\alpha_f(\mathrm{Ru}) = 1.29$ and $\alpha_d(\mathrm{Cl}) = 0.618$ for optimization. In single point calculations (SP), diffuse functions are added consistently to this augmented basis set: $\alpha_s(\mathrm{Ru}) = 0.008$, $\alpha_p(\mathrm{Ru}) = 0.011$, $\alpha_d(\mathrm{Ru}) = 0.025$, $\alpha_s(\mathrm{Cl}) = 0.09$ and $\alpha_p(\mathrm{Cl}) = 0.0075$.

Supermolecular approach is used to study all considered reactions. That means, all reactant, transition state and product are treated as single molecules. This allow us directly compare energies of these structures.

4.1.1 Hydration reaction

Hydration reaction of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex is treated in special way in this study because the complex is small enough to apply high precision QM methods. The hydration occurs in aqueous environment of a cell, independently on other molecular environment like proteins and nucleic acids. Therefore, simple QM computational model can be used where water solution is simulated by C-PCM model [201,202] with solvent dielectric constant $\epsilon = 78.39$, solvent radius $R_{\text{solv}} = 1.385$ Å and Klamt radii as implemented in program Gaussian 03 [185].

QM/MM model with explicit water molecules for the hydration reaction was built by program LEaP [203]. QM part of the system consists of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ together with one water molecule participating in the reaction. This structure was parametrized by GAFF [52] and compatible RESP charges [204–207] fitted from HF/6-31G(d) electron density were assigned to all atoms. The complex was then surrounded by 1111 TIP3P water molecules [208] shaped in rectangular box together with one Cl^- anion to compensate positive charge of the complex. Temperature and density of the water solution was equilibrated by NVT ($T = 298.15$ K) and subsequent NPT ($p = 1$ atm) MD simulation using PBC by program Amber 8 [184]. Resulting box parameters $a = 33.633$ Å, $b = 31.547$ Å, $c = 30.918$ Å were used in subsequent calculations.

QM/MM MD umbrella sampling simulations were performed in canonical ensemble with $T = 298.15$ K kept fixed by Berendsen thermostat [198] using *QMS*–Uni. Reaction coordinate was defined as $\xi = \Delta R_{\text{Aq}}/(\Delta R_{\text{Cl}} + \Delta R_{\text{Aq}})$ where $\Delta R_{\text{x}} = R_{\text{x}} - R_{\text{x}}^{\text{ref}}$ is displacement of Ru(II)–X distance (R_{x}) from its reference value $R_{\text{x}}^{\text{ref}}$. These reference values were taken from QM/MM optimized structure at the same QM level, that is DFT(B97D) with LANL2DZ or SDD/6-31G(d) basis set (see Sec. 4.2).

4.1.2 QM/MM model with DNA

On the contrary to the hydration reaction, studying binding to DNA and subsequent cross–link formation requires more extended computational model including oligonucleotid of DNA. Two different QM/MM models are employed in this work. At first, 6 Watson–Crick base–pair *ds*-DNA with sequence 5'-GCG*G*GC-3' containing two adjacent guanines in the middle was prepared by NucGen utility of Amber 8 [184]. Structural pa-

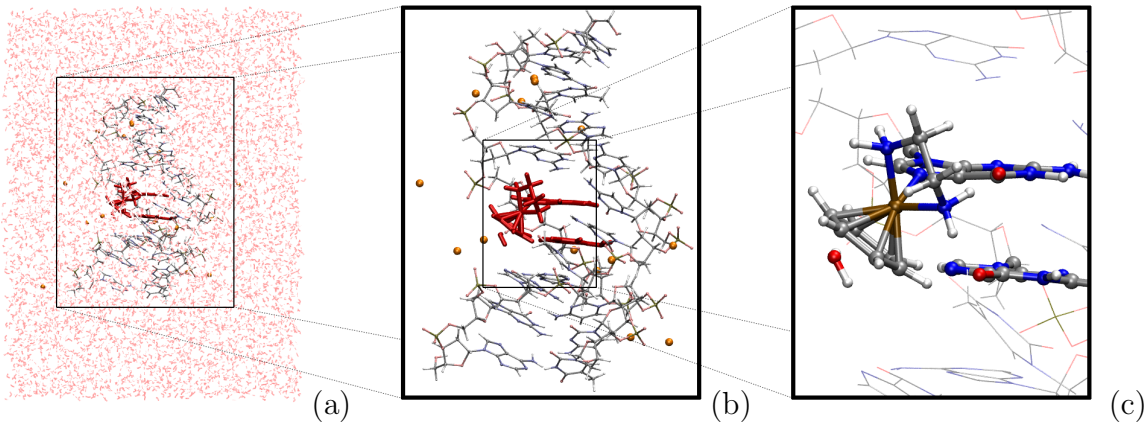


Figure 16: QM/MM computational model with periodic TIP3P water box (a), extracted DNA structure with coordinated Ru(II) complex and Na^+ cations (b) and detail of QM part of the system (c).

rameters for standard B–DNA geometry were applied. 10 sodium cations were added to compensate negative charge of phosphate groups. The DNA structure was relaxed by 100 ps MD simulations in explicit water box using Amber FF96 force field [51, 209] and TIP3P water model [208]. After that, total energy of the model was minimized by conjugate gradient algorithm coded in Amber 8 [184] and DNA structure was extracted from water box. $[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex was added to the vicinity of central guanine pair and two of Na^+ cations were removed to keep zero net charge of the model. Such prepared structure was used for ONIOM calculation [125, 126] as implemented in Gaussian 03 [185] where the Ru(II) complex and two adjacent guanines defined the QM part of the system. Link atom approach was used to treat N9(G)–C1'(S) glycosidic bond between guanine (G) and deoxyribose (S) crossing QM–MM boundary, N9(G) nitrogen was capped by hydrogen with scaling factor $\alpha_L = 1.464$. This model, in further text called simply as ONIOM model, was described at DF-DFT(BLYP)/SDD/6-31G(d):FF96 level and its purpose is to check feasibility of suggested reaction mechanism and prepare structures for more complex QM/MM model described below.

Next, 10 base–paired *ds*-DNA structure with sequence 5'–AATGG*G*ACCT–3' was prepared analogically. In this case FF03 [210] force field in Amber 11 [168] was used for DNA relaxation. After adding Ru(II) complex and 16 Na^+ cations, the structure was surrounded by 4000 TIP3P water box and equilibrated in the same way as described in case of hydration reaction. Resulting box parameters $a = 45.641$ Å, $b = 59.046$ Å, $c = 47.345$ Å were used in other calculations. QM/MM partitioning was done in the same way

Type	Translation [Å]			Rotation [deg]		
	Slide	Shift	Rise	Twist	Roll	Tilt
FF03	-2.00	-0.43	3.66	31.15	5.03	6.37
BLYP	-1.57	-0.37	3.65	31.64	-1.11	0.93
B3LYP	-1.59	-0.56	3.65	32.47	3.10	4.05
B97D	-1.47	-0.17	3.35	31.01	1.75	2.75
M06	-1.95	-0.41	3.60	29.41	2.91	6.23
Exp [211]	0.00	0.00	3.38	36.00	0.00	0.00

Table 3: Comparisson of calculated structure parameters of adjacent guanine pair in B-DNA with experimental values [211]. Computed numbers are averaged values from 10 ps QM/MM MD calculation of B-DNA in explicit water box where guanine pair was included into QM part of the system.

as in ONIOM case and the complete system, further referred as QM/MM model, is shown in Fig. 16. This model was treated by our *QMS*–Uni program and its main purpose is to construct free energy profiles of studied reactions by umbrella sampling method.

Because of intended QM/MM MD simulations, DFT functional with dispersion correction has to be used to describe correctly stacking interaction between to adjacent guanines in the QM part and thus keep proper structure of DNA. B97D functional was chosen [212] based on QM/MM MD simulations of DNA in water box as shown in Tab. 3 and because of reasonable speed of the calculation. On the other hand, compromise in basis set had to be done and LANL2DZ [213–215] pseudopotentials were used for all atoms in QM part instead of demanding SDD/6-31G(d) description used in previous models. However, sampling was done during QM/MM MD each 10 fs and free energy profiles were corrected to DFT(B3LYP)/SDD/6-31++G(d,p) by FEP [49], that is

$$\Delta G^{\mathrm{SDD}}_{\mathrm{LAN}} = G_{\mathrm{SDD}} - G_{\mathrm{LAN}} = -k_B T \ln \left\langle \exp\left(-\frac{E_{\mathrm{SDD}} - E_{\mathrm{LAN}}}{k_B T}\right)\right\rangle_{\mathrm{LAN}} \tag{96}$$

Electronic embedding was used to polarize electron density of QM part by surrounding MM atomic charges.

4.2 Hydration reaction

Hydration reaction of $[\mathrm{Ru^{II}}(\eta^6\text{-benzene})(\mathrm{en})\mathrm{Cl}]^+$ complex (**R0**, Fig. 17) proceeds when the complex passes cellular membrane and enters the cell. Driving force of this reaction is low concentration of Cl^- anions inside the cell in comparison to intercellular environment [12, 216]. From UV-VIS spectrometry measurements is known that the Ru(II) hydration

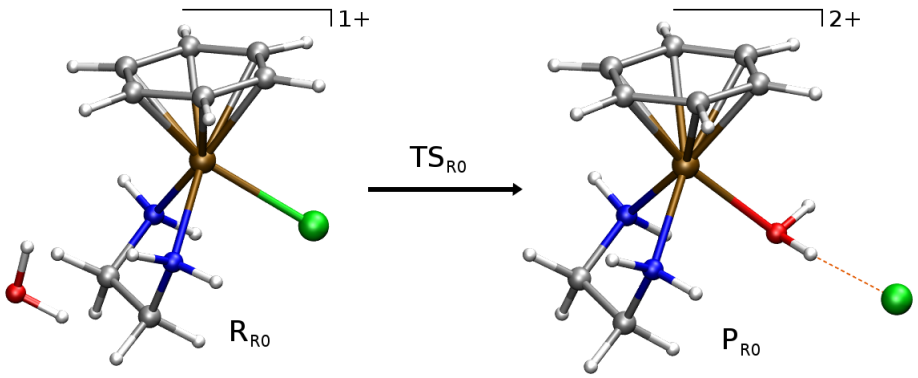


Figure 17: QM/MM optimized complexes of hydration reaction

is endotermic process which proceeds with rate constant $(1.98 \pm 0.02) \cdot 10^{-3}$ s^{-1} [217]. This value corresponds, using Eyring's transition state theory, with activation barrier 21.1 kcal/mol and it can be used for accuracy check of our calculations.

4.2.1 Accuracy check

Using DFT method with hybrid B3LYP functional was judged by calculating **R0** reaction barrier and corresponding rate constant. All stationary points of the reaction, i.e. reactant, transition state and product, were optimized at DFT(B3LYP)/6-31G(d) level in gas phase as well as in implicit solvent (C-PCM, $\epsilon = 78.39$). Energies of optimized structures were then recalculated by DFT(B3LYP), MP2 and CCSD using five different basis sets as shown in Tab. 4.

Using accurate CCSD/6-31++G(d,p) method with C-PCM, the reaction barrier 21.7 kcal/mol was determined. That is very close to experimental value if we take into consideration accuracy of C-PCM model, ECPs, basis set and B3LYP functional used for optimization. MP2 perturbation theory overestimates activation barrier as well as reaction energies by ca 2 kcal/mol. On the other hand results from DFT(B3LYP) are underestimated and converge to 20.0 kcal/mol barrier in CBS limit. However, this is still in very good agreement with experimental value 21.1 kcal/mol. Obtained results indicate that DFT(B3LYP) method provide sufficiently accurate description of Ru(II) piano–stool complexes. Based on that, all structures discussed in this work are optimized by this approach in 6-31G(d) basis set and electronic analysis are obtained in more extensive 6-31++G(d,p) if not specified otherwise.

Meth. \ En.	QM					QM(C-PCM)				
	ΔE_a	ΔE_r	ΔG_a	ΔG_r	k	ΔE_a	ΔE_r	ΔG_a	ΔG_r	k
DFT/B1	27.3	20.7	28.1	20.7	1.6E-08	17.2	-1.2	18.1	-1.2	3.6E-01
DFT/B2	27.9	22.1	28.8	22.0	5.0E-09	18.5	1.9	19.4	1.8	4.2E-02
DFT/B3	28.8	23.6	29.7	23.6	1.1E-09	19.2	2.8	20.1	2.8	1.4E-02
DFT/B4	28.8	23.4	29.6	23.4	1.2E-09	19.1	2.9	20.0	2.9	1.5E-02
DFT/B5	28.6	23.2	29.5	23.1	1.6E-09	19.1	2.9	20.0	2.9	1.5E-02
MP2/B1	32.5	21.8	33.4	21.8	2.3E-12	22.8	3.2	23.6	3.2	3.2E-05
MP2/B2	33.1	23.1	34.0	23.1	7.7E-13	22.9	6.1	23.7	6.0	2.7E-05
MP2/B3	36.3	29.2	37.2	22.1	3.7E-15	23.9	6.8	24.8	6.7	4.5E-06
CCSD/B1	32.6	24.6	33.5	24.6	1.9E-12	20.8	1.7	21.6	1.6	9.3E-04
CCSD/B2	32.0	26.6	32.9	26.5	5.3E-12	20.9	4.5	21.7	4.4	7.6E-04

Table 4: Reaction barriers and energies [kcal/mol] of hydration reaction and TST rate constants k [s^{-1}] calculated in different methods and basis sets (B1 = 6-31G(d), B2 = 6-31++G(d,p), B3 = 6-31++G(2df,2pd), B4 = Aug-CC-PVTZ, B5 = Aug-CC-PVQZ)

4.2.2 Geometry parameters

Geometry parameters of optimized stationary points of **R0** (Fig. 17) are shown in Tab. 6. Distances between central Ru(II) cation and its ligands obtained by geometry optimization using three different approaches (QM calculation in gas phase, C-PCM and QM/MM with electronic embedding) are compared. In second half of Tab. 6 are shown average values of these distances from MD sampling. Structure of transition state $\mathbf{TS_{R0}}$ is shown in separate picture Fig. 18 where used atom labels are defined.

It can be seen that significant changes in distance between different computational method occur only in case of Ru–Cl and Ru–OH_2 bonds. This is effect of electrostatic forces that are not screened, and thus overestimated, in pure QM calculations without any solvent. In QM/MM model with explicit water solution participates the Cl resp. H_2O ligand in two hydrogen bonds and that causes slightly longer distance than in implicit solvent (C-PCM). This effect is enhanced in QM/MM MD simulation where lenghts and strenghts of H–bonds are changing because of thermal motion.

4.2.3 Energy profile

Reaction energies and barriers computed by static QM calculation in gas phase as well as implicit solvent are shown in Tab. 4 and are discussed in Sec. 4.2.1. Free energy corrections to total energy differences were calculated from frequency analyses of optimized stationary points. By this standard approach for small molecular systems we got close to

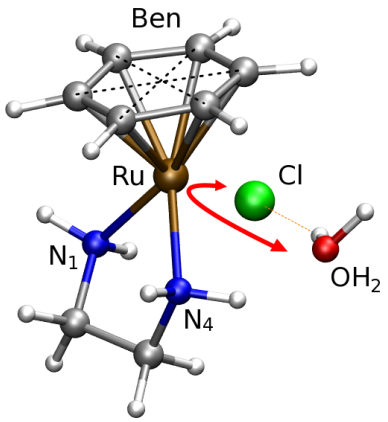


Figure 18: Transition state of hydration reaction ($\mathbf{TS_{R0}}$). Antisymmetric stretching movement of Cl—H_2O replacement is marked by red arrow.

At. \ Meth.	QM	QM(C-PCM)
Ru-Cl	3.10	3.24
Ru-OH_2	2.58	2.88
Ru-Ben	1.70	1.73
Ru-N1(en)	2.16	2.14
Ru-N4(en)	2.18	2.14

Table 5: Structural parameters of $\mathbf{TS_{R0}}$ calculated by geometry optimization at DFT(B3LYP)/6-31G(d) level. Atom labels are defined in Fig. 18.

experimental value of free energy barrier 21.1 kcal/mol.

Because we want to study interaction of Ru(II) complexes with DNA (see Sec. 4.3 and Sec. 4.4) we need to use QM/MM technique with explicit water box. Therefore accuracy of that approach was also checked on known hydration reaction energy profile. Complex $\mathbf{TS_{R0}}$ optimized in C-PCM was put into the box with 1111 water molecules and equilibrated as described in Sec. 4.1. Then, free energy profile shown in Fig. 19 was computed by umbrella sampling method.

First, DFT(B3LYP)/6-31G(d) with SDD pseudopotential for Ru and Cl was used for description of QM part (blue curve in Fig. 19). Free energy barrier 17.0 kcal/mol (WHAM) resp. 16.6 kcal/mol (UI) was obtained by this approach, which is consistent with static C-PCM result 18.1 kcal/mol (see DFT/B1 in Tab. 4). Inconsistency can be seen in reaction energy which is 4.3 kcal/mol (WHAM) resp. 4.6 kcal/mol (UI) in explicit water while -1.2 kcal/mol in C-PCM. This can be explained by different nature of the model. In C-PCM, solvent is simulated implicitly by charged cavity constructed around the complex. Chloride anion released from $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ can interact only with this complex and in $\mathbf{P_{R0}}$ stabilized by two hydrogen bonds. On the contrary, in explicit solvent Cl^- interacts mainly with water molecules and there is only one H–bond between Cl^- and Ru(II) complex, in average. More accurate behaviour in C-PCM has to be reached by better description of the model, i.e. larger basis set and more accurate computational method as can be seen from Tab. 4 where $\Delta G_r = 4.4$ kcal/mol in CCSD/6-31++G(d,p). The free energy profile was also corrected by FEP method from

Complex \ Method		Geometry optimization			Molecular dynamics		
		QM	QM(C-PCM)	QM/MM	QM	QM(C-PCM)	QM/MM
$\mathbf{R_{R0}}$	Ru-Cl	2.42	2.48	2.50	2.42	2.48	2.55
	Ru-OH_2	4.14	4.22	5.26	4.38	4.59	4.97
	Ru-Ben	1.72	1.71	1.71	1.73	1.73	1.75
	Ru-N1(en)	2.15	2.15	2.13	2.17	2.16	2.15
	Ru-N4(en)	2.17	2.15	2.15	2.19	2.17	2.15
$\mathbf{P_{R0}}$	Ru-Cl	3.91	4.27	4.73	3.85	4.37	4.40
	Ru-OH_2	2.14	2.17	2.18	2.14	2.19	2.21
	Ru-Ben	1.72	1.71	1.70	1.73	1.73	1.78
	Ru-N1(en)	2.14	2.14	2.14	2.16	2.16	2.15
	Ru-N4(en)	2.16	2.14	2.15	2.18	2.16	2.20

Table 6: Geometry parameters of **R0** complexes [Å] obtained by geometry optimization and molecular dynamics. MD values are averages of 100 samples separated by 20 fs at temperature 298.15 K. QM calculation was done at DFT(B3LYP)/6-31G(d) level.

MD sampling with DFT(B3LYP)/SDD/6-31++G(d,p) description of the complex with resulting free energy barrier 19.5 kcal/mol and reaction energy 2.0 kcal/mol. The barrier height is in good agreement with experimental value 21.1 kcal/mol.

Although B3LYP functional with 6-31G(d) basis set gives satisfying results for hydration reaction it is not possible to use it for exploring of reaction mechanism including DNA. The reason is lack of dispersion which is needed for proper description of guanine–guanine stacking interaction and too big computational demandness for QM/MM MD. For that purpose it is better to use DFT(B97D)/LANL2DZ description of QM part as discussed in Sec. 4.1. Therefore, the free energy profile of hydration reaction was constructed also at this level (red curve in Fig. 19). As you can see, poor LANL2DZ basis set where in all atoms but hydrogens are core electrons described by ECP leads to underestimation of both activation barrier $\Delta G_a = 9.1$ kcal/mol and reaction energy $\Delta G_r = 1.0$ kcal/mol. Discrepancy between WHAM and UI in transition state area is caused by shape of histograms that slightly differ from gaussian distribution here.

4.2.4 Electron density analyses

Basic analyses of electron density, that is calculation of ligand binding energies and charge population analyses, were done on optimized complexes. Changes proceeding during the reaction as well as differences between individual computational models are discussed below.

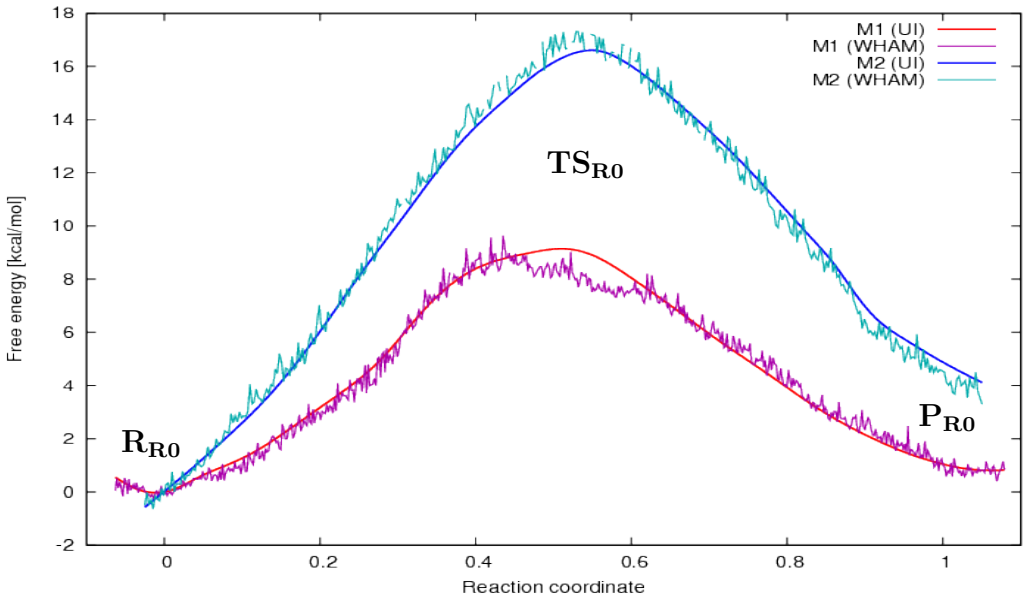


Figure 19: Free energy profile of hydration reaction (**R0**) computed by umbrella sampling QM/MM MD method. DFT(B97D)/LANL2DZ method (red curve) is compared with DFT(B3LYP)/SDD/6-31G(d) approach (blue curve).

Binding energies

Binding energies, E_{BE}, of relevant ligands are calculated according to following formula

$$E_{\mathrm{BE}} = E_{\mathrm{c-l}} - E_{\mathrm{l}} - E_{\mathrm{c}} \tag{97}$$

where three terms on the right hand side represent total energy of the whole complex, energy of given ligand and energy of the complex without the ligand. BSSE corrections are taken into account.

In case of hydration reaction, binding energies were calculated at DFT(B3LYP)/6-31++G(2df,2pd) level of theory and their values are shown in Tab. 7. First of all, one can notice that values calculated in gas phase are overestimated because of unscreened electrostatic interaction. The most stable part of the complex is ethylenediamine ligand (en) with binding energy between 100 and 110 kcal/mol. On the other hand, η^6–coordinated benzene ligand interacts with central Ru(II) cation relatively weakly, between 73 and 88 kcal/mol.

Charge population analyses

To illustrate electron density distribution in the complex, partial atomic charges were

Lig. \ Mol.	QM			QM(PCM)		
	$\mathbf{R_{R0}}$	$\mathbf{TS_{R0}}$	$\mathbf{P_{R0}}$	$\mathbf{R_{R0}}$	$\mathbf{TS_{R0}}$	$\mathbf{P_{R0}}$
Ben	-72.5	-102.4	-82.7	-72.6	-87.7	-81.7
En	-102.6	-105.3	-112.2	-99.3	-104.9	-110.2
Cl	-224.9	-191.8	-181.1	-46.7	-21.3	-19.2
Wat	-9.5	-13.0	-46.3	-4.4	-7.0	-34.1

Table 7: Ru–ligand interaction energies [kcal/mol] in complexes hydration reaction (QM, C-PCM, DFT(B3LYP)/6-31++G(2df,2pd))

Atom \ Mol.		QM			QM(C-PCM)			QM/MM		
		$\mathbf{R_{R0}}$	$\mathbf{TS_{R0}}$	$\mathbf{P_{R0}}$	$\mathbf{R_{R0}}$	$\mathbf{TS_{R0}}$	$\mathbf{P_{R0}}$	$\mathbf{R_{R0}}$	$\mathbf{TS_{R0}}$	$\mathbf{P_{R0}}$
NPA	Ru	-0.117	0.092	0.095	-0.078	0.156	0.123	-0.049	0.143	0.161
	Cl	-0.356	-0.634	-0.770	-0.459	-0.782	-0.872	-0.527	-0.812	-0.988
	O(Aq)	-0.989	-0.983	-0.878	-1.026	-0.997	-0.882	-1.086	-1.004	-0.882
CHELPG	Ru	0.007	0.116	0.080	0.184	0.413	0.278	0.353	0.470	0.376
	Cl	-0.395	-0.542	-0.609	-0.580	-0.823	-0.856	-0.696	-0.915	-0.988
	O(Aq)	-0.718	-0.740	-0.518	-0.862	-0.870	-0.669	-1.001	-0.939	-0.624

Table 8: Atomic charges [e] in Ru(II)-Cl hydration reaction calculated at DFT(B3LYP)/6-31++G(d,p) level. Charges from NPA are compared with CHELPG method.

calculated by population analysis. Because partial charges are not well defined there is no unambiguous method for this purpose. We compared results of Natural Population Analysis (NPA, [138]) and CHELPG method (CHarges from Electrostatic Potentials using a Grid based method, [218]) because these two methods do not depent strongly on basis set and give meaningful results in most cases [44] and because we have good experience with these methods from previous studies [42, 219].

In Tab. 8 atomic partial charges obtained on hydration reaction complexes by three different computational models are compared: QM with and without implicit solvent (C-PCM) and QM/MM with explicit solvent. In all these cases there is the same trend through the reaction – electron density donated to Ru(II) by Cl^- moves out as this ligand is withdrawn from the complex and then water oxygen begins to donate to the central metal. The NPA charges differs from CHELPG ones mainly on Ru atom and in $\mathbf{R_{R0}}$ complex differs even in sign. This discrepancy is caused by different nature of these methods. NPA is orbital approach and so it reflects strong Cl^- donation in reactant structure while CHELPG construct charges from overall electrostatic potential.

Changes of atomic charges of Cl^- and O(Aq) correspond with differences in binding

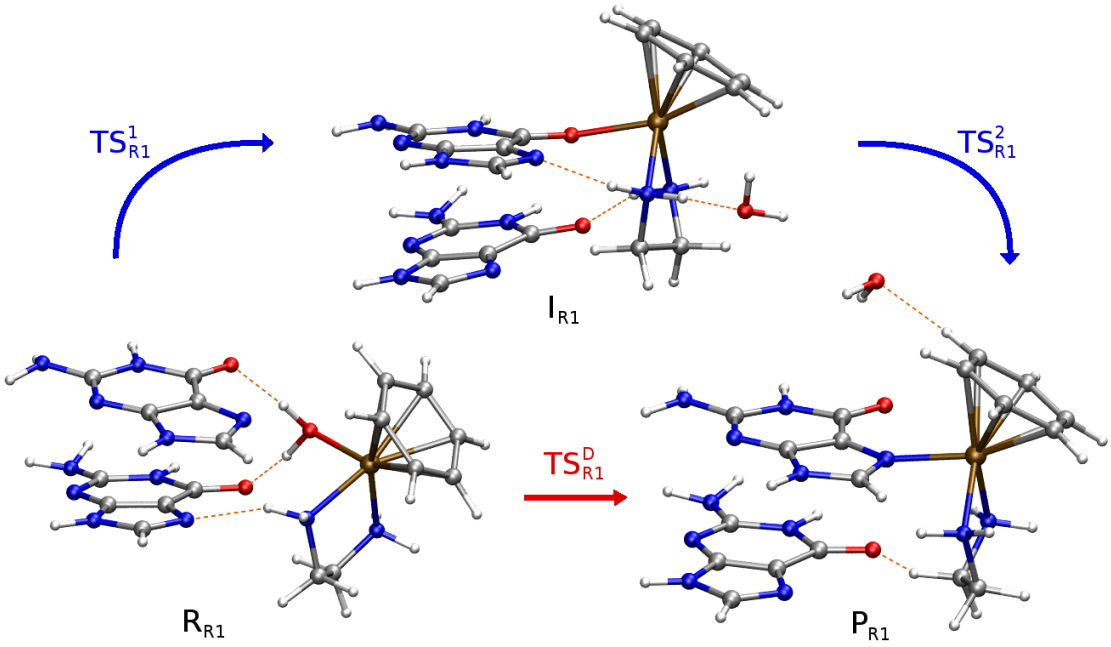


Figure 20: QM/MM optimized complexes of Ru(II)-DNA binding process

energies of these ligands. From all these analyses, it is obvious that hydration reaction of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ proceeds via dissociative mechanism. That means, the Cl^- ligands is released at the beginning of the reaction and is practically free in $\mathbf{TS_{R0}}$ structure. Similarly, the aqua ligand interacts in $\mathbf{TS_{R0}}$ very weakly and is not coordinated to Ru(II) up to last stage of the reaction.

4.3 Binding to DNA

Aqua complex $[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$, i.e. product of hydration reaction, is more reactive than $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$and interacts easily with nucleophilic binding sites of bio–molecules. As was explained in Sec. 1.3, binding to DNA, preferably to N7 guanine position, is important for cytostatic activity of this complex. This process can proceed by two reaction pathes that are both shown in Fig. 20. One possibility is direct binding mechanism with single transition state $\mathbf{TS^D_{R1}}$. This is the way how also cisplatin binds to N7 guanine. Alternative path is going through intermediate state $\mathbf{I_{R1}}$ where the Ru(II) complex interacts with oxygen O6 on guanine and then relocate to N7 position.

Complex \ Method		Geometry optimization QM(C-PCM)	ONIOM	QM/MM	Molecular dynamics QM/MM
$\mathbf{R_{R1}}$	Ru-OH_2	2.13	2.16	2.15	2.13
	Ru-N7(G)	4.68	3.98	4.45	4.78
	Ru-O6(G)	4.26	4.71	4.14	5.50
$\mathbf{I_{R1}}$	Ru-OH_2	4.17	4.48	4.52	10.50
	Ru-N7(G)	3.82	3.83	3.81	3.66
	Ru-O6(G)	2.14	2.28	2.29	2.25
$\mathbf{P_{R1}}$	Ru-OH_2	4.85	6.01	5.72	18.54
	Ru-N7(G)	2.17	2.20	2.20	2.18
	Ru-O6(G)	3.73	3.96	4.00	3.87

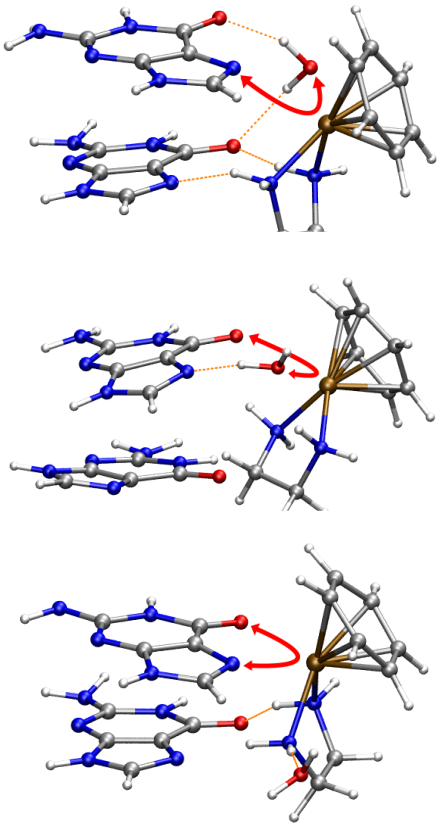
Table 9: Geometry parameters of **R1** complexes [Å]. MD values are averages of 100 samples separated by 20 fs at temperature 298.15 K.

4.3.1 Geometry parameters

To study interaction of the Ru(II) complex with DNA, larger computational model than in case of simple hydration reaction is required. DNA structure together with Ru(II) complex was build and used as QM/MM model the Ru(II) complex and two adjacent guanines from the middle of double helix are defined as QM part. From now on we are distinguishing two different QM/MM models. First, marked as ONIOM, with 6 base-paired DNA structure without any solvent and second, marked as QM/MM, with 10 base-paired DNA double helix with explicit water box (see Sec. 4.1 for more details).

Structure parameters of PES stable minima are shown in Tab. 9. Although the QM optimization with C-PCM was started from ONIOM minimized structure, final QM geometry differs from ONIOM resp. QM/MM QM part. This is caused by lack of steric influence of DNA structure in this small model. Results from ONIOM and QM/MM model differs only slightly which is caused by different solvent environment. For comparison, average values of selected distances from QM/MM MD simulation are also shown in Tab. 9. Longer Ru-O6(G) distance than Ru-N7(G), obvious from the data, suggests better stability of $\mathbf{P_{R1}}$ than $\mathbf{I_{R1}}$. Benzene ligand does not undergo any changes in **R1** reaction, its distance to Ru(II) cation is the same as in isolated $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex, i.e. 1.71 Å (see Tab. 6).

Structures of transition states are shown in Fig. 21 where the antisymmetric stretching movement related to imaginary frequency is indicated by red arrow in each geometry. The $\mathbf{TS^D_{R1}}$ structure has a compact geometry where the Ru(II) complex and aqua replacing ligand are held in suitable positions by hydrogen bonds to the both guanines. This is



At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	3.03	3.02
Ru-N7(G1)	2.86	2.83
Ru-O6(G1)	4.47	4.45
Ru-N7(G2)	4.37	4.34
Ru-O6(G2)	4.29	4.21

(a)

At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	2.72	2.71
Ru-N7(G1)	4.10	4.08
Ru-O6(G1)	3.11	3.10
Ru-N7(G2)	6.75	6.74
Ru-O6(G2)	4.92	4.90

(b)

At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	4.28	3.96
Ru-N7(G1)	2.64	2.79
Ru-O6(G1)	2.63	2.49
Ru-N7(G2)	4.51	4.62
Ru-O6(G2)	4.48	4.14

(c)

Figure 21: Transition states of **R1** reaction mechanism: (a) $\mathbf{TS^{D}_{R1}}$, (b) $\mathbf{TS^{1}_{R1}}$, (c) $\mathbf{TS^{2}_{R1}}$. Red arrows indicate antisymmetric stretching movements corresponding with imaginary frequency in each structure.

not the case of $\mathbf{TS^{1}_{R1}}$ and $\mathbf{TS^{2}_{R1}}$ where the Ru(II) complex interacts practically with one guanine only.

4.3.2 Energy profile

Free energy characteristics of Ru(II) binding to DNA (**R1**), i.e. activation barriers ΔG_a and reaction energies ΔG_r, were calculated by both static and dynamical approach and their values are shown in Tab. 10. In QM and ONIOM model where no explicit solvent is used, free energy correactions were obtained on optimized structures by static approach at DFT(B3LYP)/6-31++G(d,p) level.

In QM/MM model with explicit water box and periodic boundary conditions, the umbrella sampling method was applied to construct free energy profile of the reactions. This technique is computationally very time consuming and so description of the QM part has

Reac. \ Meth.	QM(C-PCM)		ONIOM		QM/MM_{MD}		QM/MM_{corr}	
	ΔG_a	ΔG_r	ΔG_a	ΔG_r	ΔG_a	ΔG_r	ΔG_a	ΔG_r
$\mathbf{R_{R1}} \rightarrow \mathbf{P_{R1}}$	29.1	-1.3	26.5	-12.2	15.3	-1.8	22.5	-4.5
$\mathbf{R_{R1}} \rightarrow \mathbf{I_{R1}}$	28.5	1.4	27.8	-1.1	15.7	9.2	21.1	7.8
$\mathbf{I_{R1}} \rightarrow \mathbf{P_{R1}}$	34.7	-2.7	29.3	-11.1	4.8	-11.3	10.8	-11.2

Table 10: Reaction barrier and energies [kcal/mol] in reaction mechanism of binding to DNA. Free energies for QM and ONIOM are calculated by static corrections, QM/MM_{MD} is umbrella sampling result for DFT(B97D)/LANL2DZ and QM/MM_{corr} are values corrected to DFT(B3LYP)/6-31++G(d,p) by FEP.

to be performed at the less demanding DFT(B97D)/LANL2DZ level to reach reasonable speed of QM/MM MD simulations. Resulting biased distributions were transformed to unbiased free energy profiles by WHAM as well as UI method (see Sec. 2.4.2 for details) and are shown in Fig. 22. Because ΔG_a and ΔG_r calculated at this level of theory can differ considerably from more accurate description, as discussed in Sec. 4.2.1, corrections to DFT(B3LYP)/6-31++G(d,p) were calculated by FEP. Final values are shown in last part of Tab. 10, marked as QM/MM_{corr}.

From the comparison of ONIOM and QM/MM model in Tab. 10 it is obvious that there are big differences between static and dynamical approach of free energy calculation. In direct path reaction through $\mathbf{TS^D_{R1}}$ both methods give the same barrier but differs in reaction energy. While the ONIOM predicts exothermicity -12.2 kcal/mol, in QM/MM it is only -4.5 kcal/mol. Comparison with QM result calculated in C-PCM suggest that this is mainly because of solvation effect. In the two step reaction mechanism via intermediate state $\mathbf{I_{R1}}$ the situation is a little more complicated. All models agree in preference of $\mathbf{P_{R1}}$ rather than $\mathbf{I_{R1}}$ as it was already suggested from optimized geometries. If the system gets into the $\mathbf{I_{R1}}$state, which can be possible because $\mathbf{TS^1_{R1}}$ and $\mathbf{TS^D_{R1}}$ are very close in energy, the probability that system comes back to $\mathbf{R_{R1}}$ is lower than for transition to $\mathbf{P_{R1}}$ (barriers differ by 1 kcal/mol in ONIOM and 3 kcal/mol in QM/MM model). From that we can conclude that both direct path and two–step mechanism through O6 to final N7 binding site are kinetically feasible and the whole process is spontaneous from thermodynamical point of view.

4.3.3 Electron density analyses

As in case of hydration reaction, electron density is analyzed here by several methods to better understand interaction with DNA. Binding energies of individual ligands of Ru(II) cation were calculated as well as atomic partial charges and Bader's AIM analysis. All

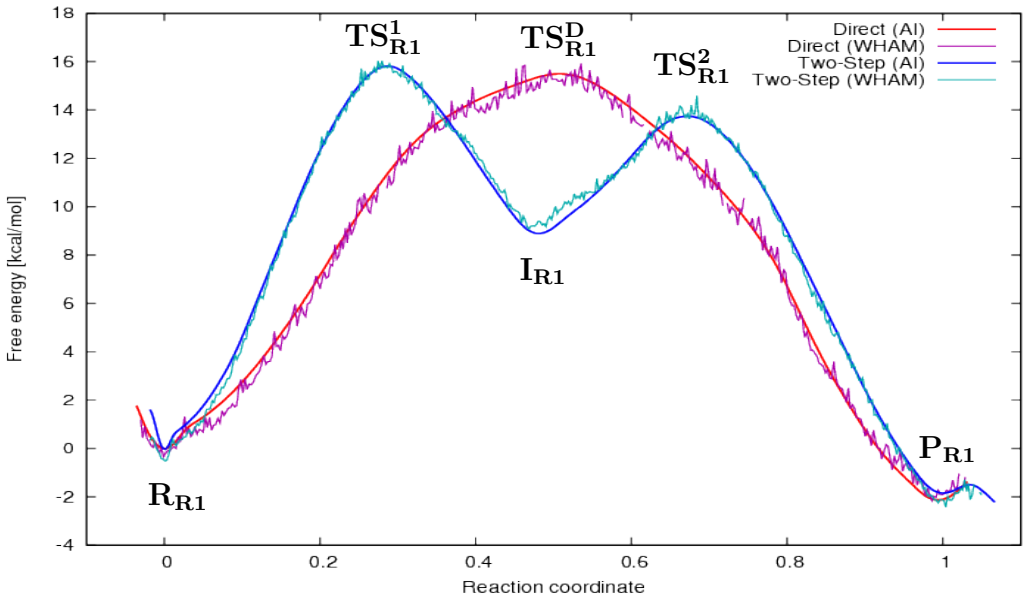


Figure 22: Free energy profile of Ru(II) binding to DNA (**R1**) computed by umbrella sampling QM/MM MD method. Direct binding to N7 guanine (red curve) as well as two–step mechanism going through O6 position (blue curve) is shown.

the calculations were done on QM part from QM/MM model reoptimized in C-PCM by DFT(B3LYP)/6-31++G(d,p) with constrained position of guanines.

Binding energies

Binding energies (BE) of all ligands interacting with central Ru(II) cation were computated using Eq. 97 corrected to BSSE and are shown in Tab. 11. For the sake of better insight, BE of both guanines is shown although only G1 is coordinating to Ru(II) while G2 participates in hydrogen bonding to ethylenediamine as apparent from Fig. 20.

Because water molecule is being replaced by G1 in **R1** reaction, marginal changes

Lig. \ Mol.	$\mathbf{R_{R1}}$	$\mathbf{I_{R1}}$	$\mathbf{P_{R1}}$	$\mathbf{TS^1_{R1}}$	$\mathbf{TS^2_{R1}}$	$\mathbf{TS^D_{R1}}$
Ben	-67.8	-66.2	-67.9	-80.5	-68.1	-77.9
En	-102.9	-109.0	-101.1	-100.5	-95.2	-103.6
Wat	-37.8	-3.4	-6.4	-16.7	-4.1	-5.7
G1	-17.1	-31.4	-40.2	-12.5	-7.9	-8.5
G2	-9.3	-3.9	-5.4	-1.1	-10.2	-16.0

Table 11: Ru–ligand interaction energies [kcal/mol] in complexes of Ru-N7(G) adduct formation reaction (QM, C-PCM, DFT(B3LYP)/6-31++(d,p))

Atom \ Mol.	$\mathbf{R_{R1}}$	$\mathbf{I_{R1}}$	$\mathbf{P_{R1}}$	$\mathbf{TS^1_{R1}}$	$\mathbf{TS^2_{R1}}$	$\mathbf{TS^D_{R1}}$
Ru	0.123	0.120	0.040	0.238	0.154	0.233
O(Aq)	-0.880	-0.999	-1.027	-1.007	-1.006	-1.009
N7(G1)	-0.528	-0.521	-0.415	-0.545	-0.525	-0.471
O6(G1)	-0.718	-0.592	-0.670	-0.674	-0.696	-0.698

Table 12: NPA charges [e] in complexes of Ru-N7(G) adduct formation reaction (QM, C-PCM, DFT(B3LYP)/6-31++G(d,p))

in BE occur in these two species. Comparing G1 interaction in $\mathbf{I_{R1}}$ and $\mathbf{P_{R1}}$show that coordination of Ru(II) to O6 oxygen is weaker by ca 9 kcal/mol than coordination to N7 position. This is in agreement with long Ru–O6 distance than Ru–N7 as discussed in Sec. 4.3.1 and it help to explain N7 preference in free energy profile.

From energy decomposition of transition state structures, it may seem that direct path through $\mathbf{TS^D_{R1}}$ proceeds by dissociative reaction mechanism as was the case of hydration reaction **R0**. Aqua ligand is already released but G1 is not coordinated yet in $\mathbf{TS^D_{R1}}$ structure. However, antisymmetric stretching movement related to imaginary frequency of $\mathbf{TS^D_{R1}}$ indicates that the reaction mechanism is rather associative. Different situation is in two–step reaction mechanims where both the ligands are coordinated in $\mathbf{TS^1_{R1}}$and water molecule is thus not fully released as far as the barrier is overcome.

Charge population analyses

Partial atomic charges were calculated by NPA method on atoms directly participating in reaction mechanism **R1** and their values are shown in Tab. 12. As was discussed in the case of hydration reaction, NPA reflects better arrangement of molecular orbitals than other methods and so only charges from this population analysis are presented here.

Charge on Ru atom is almost the same in reactant $\mathbf{R_{R1}}$ and intermediate state $\mathbf{I_{R1}}$ because in both structures Ru(II) complex interacts with oxygen. On the contrary, in Ru(II)-N7 adduct $\mathbf{P_{R1}}$ is the charge on Ru significantly reduced by electron donation from nitrogen. The Ru(II) interaction with O6 resp. N7 atom is also apparent from the charge changes of these nucleophilic binding sites in relevant structures, i.e. $\mathbf{I_{R1}}$ resp. $\mathbf{P_{R1}}$.

AIM analysis

Bader's AIM analysis of electron density allows to indentify binding interaction between atom pairs thanks to formation of a bond critical point (BCP). Values of electron density in BCPs then indicates the strenght of interaction between these two atoms. Density in

Bond \ Struct.		$\mathbf{R_{R1}}$	$\mathbf{I_{R1}}$	$\mathbf{P_{R1}}$	$\mathbf{TS^1_{R1}}$	$\mathbf{TS^2_{R1}}$	$\mathbf{TS^D_{R1}}$
Ben	Ru-C1	8.23	7.46	7.72	6.97	—	—
	Ru-C2	8.02	—	7.86	—	—	8.96
	Ru-C3	—	8.38	—	8.99	8.89	8.77
	Ru-C4	—	8.19	7.31	9.20	9.29	—
	Ru-C5	7.74	—	7.41	—	—	—
	Ru-C6	—	7.36	7.74	7.36	7.77	7.47
En	Ru-N1	8.41	9.01	8.88	9.19	8.83	8.27
	Ru-N4	8.90	9.38	8.44	8.49	7.96	8.27
Wat	Ru-O	7.30	—	—	2.19	—	1.57
G1	Ru-N7	—	—	7.61	—	3.17	—
	Ru-O6	—	6.74	—	6.74	2.84	—

Table 13: Electron density [10^{-2} $e/\text{Å}^3$] in important BCPs of **R1** complexes.

all BCPs found between Ru(II) cation and complex ligands are shown in Tab. 13.

Interaction of ethylenediamine with Ru(II) cation is in all structures except $\mathbf{TS^D_{R1}}$ asymmetric. This is caused by participating of hydrogens from amine groups in H–bonding to O6(G2) or N7(G2), which induces change of electron density on relevant (en) nitrogen atom and so influences the coordination strenght to Ru(II). Between formally η^6–coordinated benzene ligand and Ru(II) were found 3, 4 or 5 BCPs.

4.4 Cross–linked structures

Computational results presented above are supported by existing experimental data. Rate constant of the hydration reaction is known and was used for accuracy check of chosen method. Similarly, the preference of N7 binding site of guanine was observed. This agreement allow us to use the same computational setting for exploration of intrastrand cross–linked structure. As was discussed in Sec. 1.1, this metallic bridge between two adjacent guanines is crutial for cytostatic activity of cisplatin. However, it was not observed in case of Ru(II) piano–stool complexes yet.

The suggested reaction mechanism for Ru(II) cross–link formation is in Fig. 23 where the Ru(II)-N7 mono–adduct $\mathbf{P_{R1}}$, product of **R1** reaction, is the initial structure. During the reaction the coordination number of Ru(II) was 6. It means that benzene ligand was first partially and then fully released when the bridge between N7 nitrogen was created and water molecule saturated the free valences on the metal. This gives us three possible cross–linked structures: first, where η^6 coordination of benzene ligand is transformed to η^2 and O6 oxygen interacts with Ru(II) (this structure is marked as $\mathbf{I^2_{R2}}$); second, where

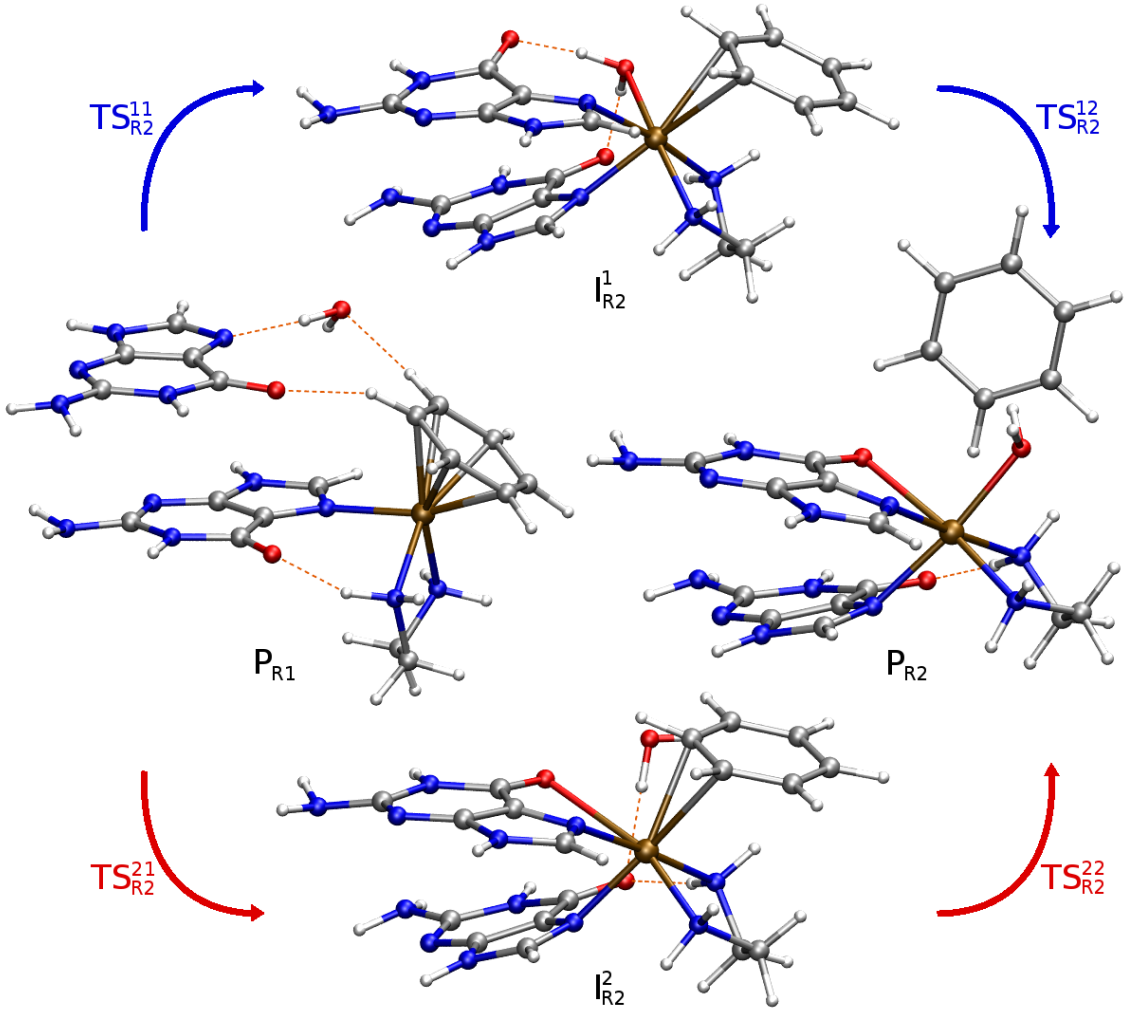


Figure 23: QM/MM optimized complexes of cross–link formation process

the water molecule binds to Ru(II) instead of O6 oxygen on G2 ($\mathbf{I^1_{R2}}$) and third, where the benzene ligand is fully released and both water molecule and O6 oxygen on G2 are coordinated to Ru(II) ($\mathbf{P_{R2}}$). The last structure, $\mathbf{P_{R2}}$, is assumed to be the most stable complex in this reaction mechanism.

4.4.1 Geometry parameters

Geometries of all the complexes from **R2** reaction mechanism were optimized in the same way as in previous **R1** reaction. Again, three models were used: ONIOM model without solvent, QM/MM model with explicit water box and small QM model with implicit solvent simulated by C-PCM. Structure parameters of stable minima of PES are shown in Tab. 14 where also their avarage values calculated from QM/MM MD sampling are present. QM

Complex \ Method		Geometry optimization QM(PCM)	ONIOM	QM/MM	Molecular dynamics QM/MM
$\mathbf{P_{R1}}$	Ru-OH_2	4.85	6.01	5.72	18.54
	Ru-$C(Bz)_{min}$	2.22	2.23	2.20	2.27
	Ru-N7(G1)	2.17	2.20	2.20	2.18
	Ru-O6(G1)	3.73	3.96	4.00	3.87
	Ru-N7(G2)	5.46	4.61	4.53	5.96
	Ru-O6(G2)	4.62	3.85	3.95	7.97
$\mathbf{I^1_{R2}}$	Ru-OH_2	2.17	2.17	2.16	2.11
	Ru-$C(Bz)_{min}$	2.58	2.54	2.62	2.44
	Ru-N7(G1)	2.10	2.26	2.14	2.13
	Ru-O6(G1)	3.75	3.60	3.64	3.58
	Ru-N7(G2)	2.28	2.18	2.22	2.04
	Ru-O6(G2)	3.78	3.73	3.71	3.51
$\mathbf{I^2_{R2}}$	Ru-OH_2	4.72	4.79	4.10	3.16
	Ru-$C(Bz)_{min}$	2.42	2.36	2.45	2.42
	Ru-N7(G1)	2.14	2.11	2.12	2.12
	Ru-O6(G1)	2.73	2.57	2.64	2.53
	Ru-N7(G2)	2.15	2.23	2.25	2.08
	Ru-O6(G2)	3.69	3.72	3.70	3.62
$\mathbf{P_{R2}}$	Ru-OH_2	2.17	2.22	2.18	2.13
	Ru-$C(Bz)_{min}$	4.88	8.58	8.57	9.97
	Ru-N7(G1)	2.19	2.09	2.11	2.08
	Ru-O6(G1)	3.70	2.40	2.40	2.21
	Ru-N7(G2)	2.11	2.19	2.23	2.08
	Ru-O6(G2)	2.48	3.75	3.70	3.59

Table 14: Geometry parameters of **R2** complexes [Å]. MD values are averages of 100 samples separated by 20 fs at temperature 298.15 K.

parts from optimized QM/MM structures are shown in Fig. 23. Geometries of transition states and relevant distances in these structures are shown in Fig. 25.

Distances from Ru(II) to O6 and N7 position on both G1 and G2 guanine as well as to aqua and benzene ligand are shown in Tab. 14. In the case of benzene, the distance from Ru(II) is not to the center of ring but to the nearest carbon of the benzene ring. The ligand arrangement described above is apparent from Fig. 23. Data from the ONIOM and QM/MM model differ only slightly, both models predict more or less the same structures. Larger changes are in C-PCM reoptimized QM part where the complex can have guanines in different position because of lacking steric effects of DNA structure. However, this mainly occur only in relatively flexible mono adduct $\mathbf{P_{R1}}$ than in cross–linked structure.

Geometries of transition states in Fig. 25 indicate that cross–link formation is quite

Reac. \ Meth.	QM(C-PCM)		ONIOM		QM/MM_{MD}		QM/MM_{corr}	
	ΔG_a	ΔG_r	ΔG_a	ΔG_r	ΔG_a	ΔG_r	ΔG_a	ΔG_r
$\mathbf{P_{R1}} \rightarrow \mathbf{I^1_{R2}}$	57.9	2.4	30.1	-12.8	23.8	1.1	37.0	15.3
$\mathbf{P_{R1}} \rightarrow \mathbf{I^2_{R2}}$	38.2	13.2	23.7	-9.7	10.1	-9.8	16.9	1.5
$\mathbf{I^1_{R2}} \rightarrow \mathbf{P_{R2}}$	20.7	4.2	10.4	-8.6	1.2	-15.4	3.2	-28.8
$\mathbf{I^2_{R2}} \rightarrow \mathbf{P_{R2}}$	15.2	-6.5	2.3	-11.7	8.3	-6.8	9.3	-11.2

Table 15: Reaction barrier and energies [kcal/mol] in reaction mechanism of cross–link formation. Free energies for QM and ONIOM are calculated by static corrections, QM/MM_{MD} is umbrella sampling result for DFT(B97D)/LANL2DZ and QM/MM_{corr} are values corrected to DFT(B3LYP)/6-31++G(d,p) by FEP.

complex process where significant structural changes occur. This suggests that these structures, i.e. $\mathbf{TS^{11}_{R2}}$resp. $\mathbf{TS^{21}_{R2}}$, can be energetically high above the mono adduct $\mathbf{P_{R1}}$. On the other hand, transition states between individual cross–linked structures ($\mathbf{I^1_{R2}}$, $\mathbf{I^2_{R2}}$ and $\mathbf{P_{R2}}$), where the benzene ligand is being released ($\mathbf{TS^{12}_{R2}}$) or replaced by water molecule ($\mathbf{TS^{22}_{R2}}$) have simple vibrational mode related to imaginary frequency.

4.4.2 Energy profile

Reaction mechanism **R2** has two possible pathways leading to final $\mathbf{P_{R2}}$ product structure. Both pathways can be divided into two consequent reactions with activation barriers ΔG_a. Free energy values of these barriers as well as reaction energies ΔG_r were calculated by static or dynamical approach, depending on the model, and are shown in Tab. 15.

For dynamical approach, QM/MM MD umbrella sampling technique was used with DFT(B97D)/LANL2DZ description of QM part as mentioned in Sec. 4.3. Resulting free energy profile constructed by WHAM as well as umbrella integration method are shown in Fig. 24. In order to get values for better description of QM part, FEP corrections to DFT(B3LYP)/6-31++G(d,p) were calculated and final barriers and reaction energies are shown in last column of Tab. 15.

From the energy profile in Fig. 24, it is obvious that creation of $\mathbf{I^1_{R2}}$ cross–link with η^2–coordinated benzene ligand and recoordinated water molecule is kinetically inhibited by the high energy barrier. This was already suggested by relatively complex geometry of relevant transition state $\mathbf{TS^{11}_{R2}}$. Also the steric repulsion of ligands in $\mathbf{I^1_{R2}}$ is significant, resulting in low stability of this structure where benzene ligand can be very easily released. This is apparent from Fig. 24 where the barrier between $\mathbf{I^1_{R2}}$ and $\mathbf{P_{R2}}$ is only 1 kcal/mol (3.2 kcal/mol after FEP corrections).

In the second reaction path, the $\mathbf{I^2_{R2}}$ cross–linked structure is formed first and then

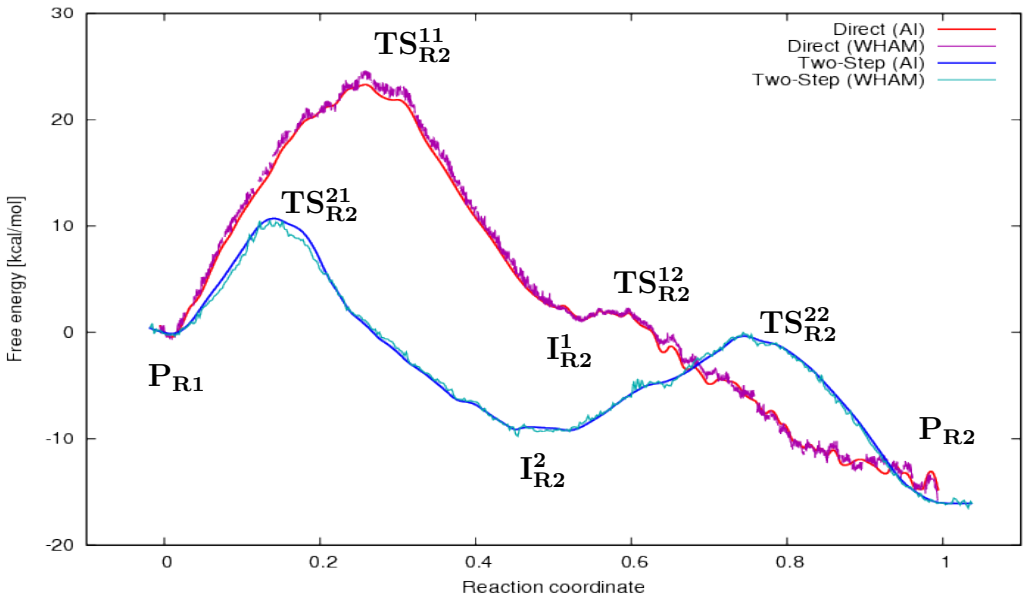
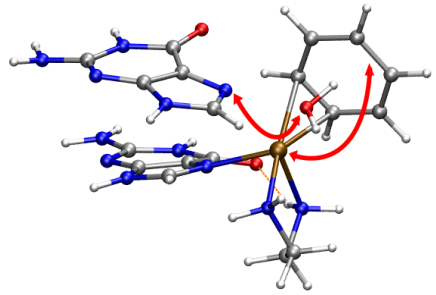


Figure 24: Free energy profile of Ru(II) cross–link formation (**R2**) computed by umbrella sampling QM/MM MD method. Two reaction pathways going through $\mathbf{I^1_{R2}}$ (red curve) or $\mathbf{I^2_{R2}}$ (blue curve) to final $\mathbf{P_{R2}}$ are shown.

the benzene ligand is substituted by water molecule. This reaction mechanism has significantly lower energy barrier. Energy of the $\mathbf{TS^{21}_{R2}}$ structure is lower than $\mathbf{TS^{11}_{R2}}$ energy in all used computational models and considering that ca. 2 kcal/mol correlate roughly with one order in rate constant, the second reaction pathway is apparently kinetically preferred. Height of the $\mathbf{TS^{21}_{R2}}$ barrier is 16.9 kcal/mol in corrected QM/MM model, which corresponds with rate constant of 2.5 $mol^{-1}s^{-1}$. Therefore, cross–link formation should be even faster process than binding to DNA.

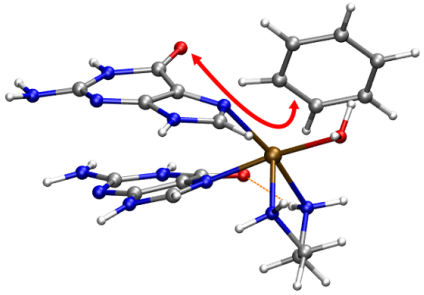
4.4.3 Electron density analyses

Binding energies of all interacting ligands and partial charges on relevant atoms together with Bader's AIM topological analysis were performed in order to get deeper insight to the electronic properties of the individual complexes. As in the part where Ru(II) binding to DNA was studied, all calculations in this section were done on DFT(B3LYP)/6-31++G(d,p) electron density evaluated in C-PCM.



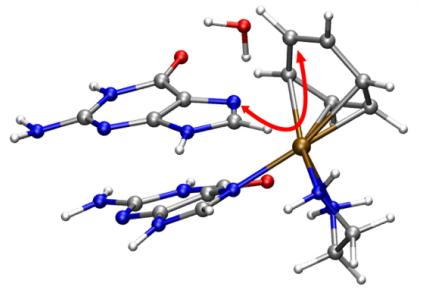
At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	2.54	2.50
Ru-$C(Bz)_{min}$	2.49	2.51
Ru-$C(Bz)_{max}$	3.81	3.83
Ru-N7(G1)	2.18	2.08
Ru-O6(G1)	3.67	3.52
Ru-N7(G2)	2.64	2.51
Ru-O6(G2)	4.10	3.82

(a)



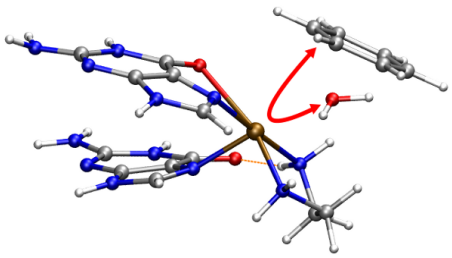
At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	2.94	2.18
Ru-$C(Bz)_{min}$	2.49	2.59
Ru-$C(Bz)_{max}$	4.19	4.05
Ru-N7(G1)	2.19	2.06
Ru-O6(G1)	3.11	3.74
Ru-N7(G2)	2.21	2.25
Ru-O6(G2)	3.81	3.86

(b)



At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	5.30	4.74
Ru-$C(Bz)_{min}$	2.26	2.24
Ru-$C(Bz)_{max}$	2.97	2.97
Ru-N7(G1)	2.75	2.15
Ru-O6(G1)	4.09	3.22
Ru-N7(G2)	2.20	2.72
Ru-O6(G2)	3.59	4.06

(b)



At. \ Meth.	ONIOM	QM/MM
Ru-OH_2	2.95	2.80
Ru-$C(Bz)_{min}$	3.43	3.40
Ru-$C(Bz)_{max}$	5.39	4.71
Ru-N7(G1)	2.10	2.03
Ru-O6(G1)	2.63	3.70
Ru-N7(G2)	2.10	2.11
Ru-O6(G2)	3.76	2.61

(c)

Figure 25: Transition states of **R2** reaction mechanism: (a) $\mathbf{TS^{11}_{R2}}$, (b) $\mathbf{TS^{12}_{R2}}$, (c) $\mathbf{TS^{21}_{R2}}$, (d) $\mathbf{TS^{22}_{R2}}$. Red arrows indicate antisymmetric stretching movements corresponding with imaginary frequency in each structure.

Lig. \ Mol.	$\mathbf{P_{R1}}$	$\mathbf{I^1_{R2}}$	$\mathbf{I^2_{R2}}$	$\mathbf{P_{R2}}$	$\mathbf{TS^{11}_{R2}}$	$\mathbf{TS^{12}_{R2}}$	$\mathbf{TS^{21}_{R2}}$	$\mathbf{TS^{22}_{R2}}$
Ben	-67.9	-11.6	-17.5	0.3	-13.0	-14.3	-36.1	0.8
En	-101.1	-95.5	-106.8	-101.3	-102.3	-109.9	-105.4	-105.2
Wat	-6.4	-28.2	-3.1	-23.3	-9.6	-2.3	-3.3	-0.5
G1	-40.2	-47.0	-41.5	-35.4	-49.6	-38.6	-42.2	-49.1
G2	-5.4	-35.9	-42.4	-46.5	-7.1	-45.4	-1.5	-39.0

Table 16: Ru–ligand binding energies [kcal/mol] in complexes of cross–linked structures (CL) and single guanine interaction complex (Ru-G) calculated in QM model with PCM.

Binding energies

BSSE corrected binding energies of ligands in all PES stationary points of **R2** reaction mechanism were calculated using Eq. 97 and their values are shown in Tab. 16. Looking at $\mathbf{P_{R1}} \rightarrow \mathbf{I^1_{R2}}$ transition, it is apparent that transformation of η^6 benzene coordination to η^2 proceeds before $\mathbf{TS^{11}_{R2}}$ is reached. Binding energy of this arene ligand is reduced from -68 to -13 kcal/mol during this process. On the contrary, water molecule and second guanine are coordinated to Ru(II) after the activation barrier is overcomed. In $\mathbf{P_{R1}} \rightarrow \mathbf{I^2_{R2}}$ reaction, the benzene ligand changes its coordination consequently from η^6 in $\mathbf{P_{R1}}$ to η^4 in $\mathbf{TS^{21}_{R2}}$ (with 32 kcal/mol reduction of BE) and later to final η^2 in $\mathbf{I^2_{R2}}$ (18.5 kcal/mol reduction to -17.5 kcal/mol). The Ru–bridge to N7 nitrogen of G2 is not created until the $\mathbf{I^2_{R2}}$geometry is reached.

Population analyses

Partial atomic charges were calculated by NPA method and their values are shown in Tab. 17. Redistribution of electron density reflects to the ligand coordination and geometry arrangement. Atomic charges in Ru(II)–N7(G) mono–adduct $\mathbf{P_{R1}}$ are also shown in Tab. 17 for the sake of comparison. It can be seen that charge of N7 nitrogen is less negative if the relevant guanine is coordinated to Ru(II) cation, which is caused by electron density donation of the nitrogen to ruthenium. This effect is also apparent in case of O6 oxygen when the $\mathbf{I^1_{R2}}$ and $\mathbf{I^2_{R2}}$structures are compared. However, electronegativity of oxygen is greater than of nitrogen and so the atomic charge of O6 is significantly larger in absolute value than charge of N7.

Because benzene ligand is being released from the Ru(II) complex in this reaction scheme, atomic charge of the nearest carbon of the benzene ring to Ru(II) is shown in Tab. 17. This charge differs significantly between Ru(II) mono–aduct and cross–linked structures where is more negative. On the other hand, there is no big change in charge

Atom \ Mol.	$\mathbf{P_{R1}}$	$\mathbf{I^1_{R2}}$	$\mathbf{I^2_{R2}}$	$\mathbf{P_{R2}}$	$\mathbf{TS^{11}_{R2}}$	$\mathbf{TS^{12}_{R2}}$	$\mathbf{TS^{21}_{R2}}$	$\mathbf{TS^{22}_{R2}}$
Ru	0.120	0.240	0.298	0.322	0.479	0.500	0.220	0.515
O(Aq)	-0.999	-0.899	-1.021	-0.867	-0.968	-0.887	-1.025	-0.978
$C(Bz)_{min}$	-0.188	-0.256	-0.266	-0.249	-0.262	-0.235	-0.226	-0.247
N7(G1)	-0.521	-0.406	-0.446	-0.437	-0.424	-0.445	-0.437	-0.427
O6(G1)	-0.592	-0.697	-0.614	-0.680	-0.674	-0.672	-0.701	-0.673
N7(G2)	-0.507	-0.449	-0.433	-0.437	-0.447	-0.472	-0.459	-0.467
O6(G2)	-0.701	-0.681	-0.701	-0.609	-0.662	-0.649	-0.658	-0.624

Table 17: Atomic charges [e] in complexes of cross–linked structures (CL) and single guanine interaction complex (Ru-G) calculated in QM model with PCM.

Bond \ Struct.		$\mathbf{P_{R1}}$	$\mathbf{I^1_{R2}}$	$\mathbf{I^2_{R2}}$	$\mathbf{P_{R2}}$	$\mathbf{TS^{11}_{R2}}$	$\mathbf{TS^{12}_{R2}}$	$\mathbf{TS^{21}_{R2}}$	$\mathbf{TS^{22}_{R2}}$
Ben	Ru-C1	7.72	—	5.22	—	—	—	—	0.78
	Ru-C2	7.86	3.44	4.02	—	—	—	—	—
	Ru-C3	—	2.94	—	—	3.97	1.05	—	—
	Ru-C4	7.31	—	—	—	4.76	—	7.40	—
	Ru-C5	7.41	—	—	—	—	—	7.35	—
	Ru-C6	7.74	—	—	—	—	—	—	—
En	Ru-N1	8.88	8.95	8.77	8.54	8.99	8.79	9.51	9.09
	Ru-N4	8.44	8.10	9.45	9.02	8.54	9.47	7.99	8.44
Wat	Ru-O	—	6.29	—	5.86	3.34	5.63	—	1.38
G1	Ru-N7	7.61	8.66	7.77	6.95	7.19	7.91	6.98	8.59
	Ru-O6	—	—	—	—	—	—	—	—
G2	Ru-N7	—	5.51	7.64	8.11	2.68	6.24	2.38	8.24
	Ru-O6	—	—	2.19	3.30	—	—	—	2.38

Table 18: Electron density [10^{-2} $e/$Å^3] in important BCPs of **R2** complexes.

between geometries with η^2–coordinated benzene ligand and final $\mathbf{P_{R2}}$ structure. This indicates weak binding interaction of this ligand in $\mathbf{I^1_{R2}}$ and $\mathbf{I^2_{R2}}$ complexes in agreement with binding energy values shown in Tab. 16.

AIM analysis

Bader's AIM analysis, which was done on DFT(B3LYP)/6-31++G(d,p) C-PCM electron density of QM part, provides insight into topological changes of **R2** complexes during the cross–link formation process. Electron densities in bond critical points (BCP) between Ru(II) and its ligands are shown in Tab. 18.

First of all, change of η^6 coordination of benzene ligand to η^2 type is clearly visible. Ru–(Bz) BCPs in $\mathbf{I^1_{R2}}$ has lower values of electron density than in $\mathbf{I^2_{R2}}$ which correlates

with weaker binding interaction of benzene in crowded $\mathbf{I^1_{R2}}$ cross–linked structure. In $\mathbf{P_{R2}}$, the benzene ligand is fully released and substituted by water molecule. Since the Ru(II) cation has coordination number 6, there can be found BCP between Ru(II) and O6(G2) oxygen in $\mathbf{I^2_{R2}}$ and $\mathbf{P_{R2}}$ structures to saturate the remaining free coordination. All PES minima have structures with pseudo–octahedral symmetry.

5 Conclusions

In this work the reaction mechanism of Ru(II) piano–stool complexes interaction with DNA was investigated by QM and QM/MM computational methods. The reaction mechanism of Ru(II) complexes is divided into three phases that are studied separately. First phase is hydration reaction of $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex, which was chosen as a representative compound of the group of Ru(II) piano–stool complexes. Its hydration results in $[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex that consequently binds to N7 guanine position in DNA, this is the second phase of reaction mechanism. The last phase, suggested in analogy to cisplatin, is the transformation of Ru(II)-N7 mono–adduct to intrastrand cross–linked structure where Ru(II) complex forms a bridge between two adjacent guanines.

First of all, the hydration reaction was explored extensively using various QM methods. Since the $[Ru^{II}(\eta^6\text{-benzene})(en)Cl]^+$ complex is relatively small the accurate QM methods with sufficiently large basis sets could be used for calculation of reaction energy and activation barrier of this process. Not surprisingly, the best agreement with experimental value of rate constant provides the CCSD/6-31++G(d,p) method in C-PCM with difference ca. 0.6 kcal/mol to the experimentally determined reaction barrier. While MP2 perturbation method overestimates the barrier by more than 3.5 kcal/mol relatively faster DFT(B3LYP) method underestimates this value only by 1.1 kcal/mol in CBS limit. Therefore, this method was chosen for all other calculations. Similar accuracy check was done also for QM/MM MD umbrella sampling method used for free energy profile construction from models with explicit water solvent. The computationally demanding MD simulations were performed with LANL2DZ pseudopotentials and corresponding double–ζ basis set. Resulting free energy profile was corrected by FEP from MD sampling at DFT(B3LYP)/SDD/6-31G(d) level that yields free energy barrier height 19.5 kcal/mol which is 0.6 kcal/mol below experimental value.

Interactions of Ru(II) piano–stool complexes with DNA were studied by QM/MM method with model containing fragment of double–stranded DNA. Two different models were taken into account. First, Ru(II) complex with 6–bp DNA without any solvent where geometry optimization were performed using ONIOM method implemented in program Gaussian 03. The latter, more realistic model with 10–bp DNA and explicit water solvent was used for QM/MM MD simulations performed by the *QMS*–Uni software developed in our laboratory. Based on previous QM calculations, two reaction pathways were considered for binding of $[Ru^{II}(\eta^6\text{-benzene})(en)(H_2O)]^{2+}$ complex to N7 nitrogen of guanine in DNA: direct binding and two–step mechanism with the Ru(II)–O6(G) intermediate state.

This intermediate state is lower in Gibbs free energy than the reactant in ONIOM model, however, it is about 8 kcal/mol above the reactant when the explicit water is included in QM/MM model. The entire binding process is exothermic with reaction free energy -4.5 kcal/mol and barrier 22.5 kcal/mol for direct reaction path that is 2.4 kcal/mol above experimentally determined activation energy. Because barrier for Ru(II)–O6(G) is comparable (21.1 kcal/mol) and backward reaction is connected with barrier about 3 kcal/mol higher barrier than consequent transition to final N7 binding site, the two–step reaction pathway can proceed as well.

Intrastrand cross–link formation is the last phase of studied reaction mechanism and it was investigated using the same models and computational approaches as in the case of Ru–binding to DNA. Reaction scheme for Ru(II)–N7(G) mono–adduct transformation into guanine–guanine cross–linked structure was suggested. The benzene coordination is reduced from η^6 to η^2–type and the ligand is subsequently completely released from the complex. This process can theoretically proceeds in two different ways. First, complex cross–linked structure is created where besides two guanines and ethylenediamine also benzene and aqua ligand are simultaneously coordinated to Ru(II) cation. However, formation of this structure is blocked by high energy barrier, 37.0 kcal/mol. The other way is to create guanine–guanine bridge first and consequently replace the benzene ligand by water molecule. This two–step mechanism leads to the most stable geometry, by -14.3 kcal/mol lower in free energy than initial Ru(II) mono–adduct. Barriers of Ru–bridge creation process and benzene substitution have reasonable heights 16.9 kcal/mol resp. 9.3 kcal/mol.

For performing QM/MM calculations special software called *QMS*–Uni was developed that can couple QM programs (Gaussian, GAMESS, Turbomole or MolPro) with Amber MM package. Methods for geometry optimization and Born–Oppenheimer molecular dynamics were implemented in the program. Because harmonic restrains can be applied on Cartesian coordinates of atoms or interatomic distance between them, this code can be easily used for QM/MM MD umbrella sampling calculations as it is demonstrated in this work. For constructing free energy profile from umbrella sampling histograms, WHAM and umbrella integration methods were implemented in standalone program. Source code of all these software written in C program language is available on attached CD together with brief manuals how to use it.

It can be concluded that reaction mechanism of Ru(II) piano–stool complexes interaction with DNA was explored systematically by QM and QM/MM computational techniques. Calculated results are in good agreement with available experimental data that justifies chosen methods.

References

[1] Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; and Walter, P. *Molecular Biology of the Cell.* Garland Science, 2002.

[2] Karp, G. *Cell and Molecular Biology: Concepts and Experiments.* Wiley, 2007.

[3] Murray, Robert, K.; Granner, D.K.; Mayes, P.A.; and Rodwell, V.W. *Harper's Illustrated Biochemistry.* McGraw-Hill Companies, Inc., 2003.

[4] Zalups, R.K. and Koropatnick, J., editors. *Cellular and Molecular Biology of Metals.* CRC Press, 2010.

[5] Beller, M. and Bolm, C., editors. *Transition Metals for Organic Synthesis: Building Blocks and Fine Chemicals.* Wiley-VCH, 2004.

[6] Davies, P.W. *Annu. Rep. Prog. Chem., Sect. B: Org. Chem*, **106**:98–119, 2010.

[7] Reedijk, J. *Curr. Opin. Chem. Biol.*, **3**(2):236 – 240, 1999.

[8] Ronconi, L. and Sadler, P.J. *Coord. Chem. Rev.*, **251**(13-14):1633 – 1648, 2007.

[9] Dabrowiak, J.C. *Metals in Medicine.* Wiley, 2009.

[10] Desoize, B. *Anticancer Res.*, **24**:1529–1544, 2004.

[11] Rosenberg, B.; van Camp, L.; and Krigas, T. *Nature*, **205**:698–6999, 1965.

[12] Lippert, B., editor. *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug.* Wiley-VCH Verlag GmbH, Wienheim, Germany, 1999.

[13] National cancer institute, 2011.

[14] Miller, S.E.; Gerard, K.J.; and House, D.A. *Inorg. Chim. Acta*, **190**(1):135 – 144, 1991.

[15] Arpalahti, J.; Mikola, M.; and Mauristo, S. *Inorg. Chem.*, **32**(15):3327–3332, 1993.

[16] Mikola, M. and Arpalahti, J. *Inorg. Chem.*, **33**:4439–4445, 1994.

[17] Hindmarsch, K.; House, D.A.; and Turnbull, M.M. *Inorg. Chim. Acta*, **257**(1):11 – 18, 1997.

[18] Marzilli, L.G.; Ano, S.O.; Intini, F.P.; and Natile, G. *J. Am. Chem. Soc.*, **121**:9133–9142, 1999.

[19] Burda, J.V.; Zeizinger, M.; and Leszczynski, J. *J. Chem. Phys.*, **120**(3):1253–1262, 2004.

[20] Burda, J.V.; Zeizinger, M.; and Leszczynski, J. *J. Comput. Chem.*, **26**:907–914, 2005.

[21] Ishida, S.; Lee, J.; Thiele, D.J.; and Herskowitz, I. *Proc Natl Acad Sci USA*, **99**(22):14298–14302, 2002.

[22] Sinani, D.; Adle, D.J.; Kim, H.; and Lee, J. *J. Biol. Chem.*, **37**:26775–26785, 2007.

[23] Pabla, N.; Murphy, R.F.; Liu, K.; and Dong, Z. *Am. J. Physiol. Renal Physiol.*, **296**:F505–F511, 2009.

[24] Miller, Sian, E. and House, D.A. *Inorg. Chim. Acta*, **187**:125–132, 1991.

[25] Wheate, N.J.; Walker, S.; Craig, G.E.; and Oun, R. *Dalton Trans.*, **39**:8113–8127, 2010.

[26] Perez, R.P. *Eur. J. Cancer*, **34**:1535–1542, 1998.

[27] Graham, J.; Mushin, M.; and Kirkpatrick, P. *Nat. Rev. Drug. Discov.*, **3**:11–12, 2004.

[28] Canetta, R.; Rozencweig, M.; and Carter, S.K. *Cancer Treat Rev.*, **12**:125–136, 1985.

[29] Gielen, M. and Tiekink, E.R.T., editors. *Metallotherapeutic Drugs and Metal–Based Di-*

agnostic Agents: The Use of Metals in Medicine. John Wiley & Sons, Ltd., 2005.

[30] Ang, W.H. and Dyson, P.J. *Eur. J. Inorg. Chem.*, pages 4003–4018, 2006.

[31] Allardyce, S.C. and Dyson, J.P. *Platinum Metals Review*, **45**(2):62–69, 2001.

[32] Morris, R.E.; Aird, R.E.; Murdoch, P.d.S.; Chen, H.; Cummings, J.; Hughes, N.D.; Parsons, S.; Parkin, A.; Boyd, G.; Jodrell, D.I.; and Sadler, P.J. *J. Med. Chem.*, **44**:3616–3621, 2001.

[33] Aird, R.; Cummings, J.; Ritchie, A.A.; Muir, M.; Morris, R.E.; Chen, H.; Sadler, P.J.; and Jodrell, D.I. *Br. J. Cancer*, **86**:1652–1657, 2002.

[34] Yan, Y.K.; Melchart, M.; Habtemariam, A.; and Sadler, P.J. *Chem. Comm.*, pages 4764–4776, 2005.

[35] Guichard, S.; Else, R.; Reid, E.; Zeitlin, B.; Aird, R.; Muir, M.; Dodds, M.; Fiebig, H.; Sadler, P.; and Jodrell, D. *Biochem. Pharm.*, **71**(4):408 – 415, 2006.

[36] Allardyce, C.S.; Dyson, P.J.; Ellis, D.J.; and Heath, S.L. *Chem. Comm.*, page 1396, 2001.

[37] Allardyce, C.S.; Dyson, P.J.; Ellis, D.J.; Salter, P.A.; and Scopelliti, R. *J. Organomet. Chem.*, **668**:35–42, 2003.

[38] Jennerwein, M. and Andrews, P.A. *Drug Metab. Dispos.*, **23**:178–184, 1995.

[39] Stebler-Röthlisberger, M.; Hummel, W.; Pittet, P.A.; Bürgi, H.B.; Ludi, A.; and Merbach, A.E. *Inorg. Chem.*, **27**:1358–1363, 1988.

[40] Peacock, A.F.A.; Habtemariam, A.; Fernandez, R.; Walland, V.; Fabbiani, F.P.A.; Parsons, S.; Aird, R.E.; Jodrell, D.I.; and Sadler, P.J. *J. Am. Chem. Soc.*, **128**:1739–1748, 2006.

[41] Deubel, D.V. and Kai-Chi Lau, J. *Chem. Comm.*, **23**:2451–2453, 2006.

[42] Futera, Z.; Klenko, J.; Sponer, J.E.; Sponer, J.; and Burda, J.V. *J. Comput. Chem.*, 2009.

[43] Gossens, C.; Tavernelli, I.; and Rothlisberger, U. *J. Phys. Chem. A*, **113**:11888–11897, 2009.

[44] Cramer, C.J. *Essentials of Computational Chemistry - Theories and Models*. John Wiley & Sons, Ltd., 2004.

[45] Morse, P.M. *Phys. Rev.*, **34**:57–64, 1929.

[46] Lennard-Jones, J. *Proc. R. Soc. Lond. A*, **106**(738):463–477, 1924.

[47] Thijssen, J. *Computational Physics*. Cambridge University Press, 2007.

[48] Ewald, P.P. *Ann. Physik*, **369**(64):253–287, 1921.

[49] Frenkel, D. and Smit, B. *Understanding Molecular Simulation - From Algorithms to Applications*. Elsevier, 1996.

[50] Daren, T.; York, D.; and Pedersen, L. *J. Chem. Phys.*, **98**:10089–10092, 1993.

[51] Cornell, W.D.; Cieplak, P.; Bayly, C.; Gould, I.R.; Merz, Kenneth M., J.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; and Kollman, P.A. *J. Am. Chem. Soc.*, **117**:5179–5197, 1995.

[52] Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; and Case, D.A. *J. Comput. Chem.*, **25**(9):1157–1174, 2004.

[53] Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; and Karplus, M. *J. Comp. Chem.*, **4**(2):187–217, 1983.

[54] Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; and Berendsen, H.J.

J. Comp. Chem., **26**(16):1701–1718, 2005.

[55] van Gunsteren, W.F. and Berendsen, H.J.C. *Groningen Molecular Simulation (GROMOS) Library Manual*. BIOMOS b.v., Groningen, 1987.

[56] Jorgensen, W.L.; Maxwell, D.S.; and Tirado-Rives, J. *J. Am. Chem. Soc.*, **118**:11225–11236, 1996.

[57] Rappe, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A.I.; and Skiff, W.M. *J. Am. Chem. Soc.*, **114**:10024–10035, 1992.

[58] Maple, J.R.; Hwang, M.J.; Stockfisch, T.P.; Dinur, U.; Waldman, M.; Ewig, C.S.; and Hagler, A.T. *J. Comp. Chem.*, **15**:162–182, 1994.

[59] Allinger, N.L. *J. Am. Chem. Soc.*, **99**:8127–8134, 1977.

[60] Allinger, N.L.; Yuh, Y.H.; and Lii, J.H. *J. Am. Chem. Soc.*, **111**:8551–8565, 1989.

[61] Lii, J.H. and Allinger, N.L. *J. Am. Chem. Soc.*, **111**:8566–8575, 1989.

[62] Lii, J.H. and Allinger, N.L. *J. Am. Chem. Soc.*, **111**:8576–8582, 1989.

[63] Allinger, N.L.; Chen, K.; and Lii, J.H. *J. Comp. Chem.*, **17**:642–668, 1996.

[64] Gresh, N.; Cisneros, G.A.; Darden, T.A.; and Piquemal, J.P. *J. Chem. Theory Comput.*, **3**:1960, 2007.

[65] Ponder, J.W. *TINKER - Software Tools for Molecular Design. Users's Guide*. Washington University School of Medicine, 1990.

[66] Alder, B.J. and Wainwright, T.E. *J. Chem. Phys.*, **31**(2):459–466, 1959.

[67] Verlet, L. *Phys. Rev.*, **159**(1):98–103, 1967.

[68] Beeman, D. *J. Comput. Phys.*, **20**:130–139, 1976.

[69] Swope, W.C. and Andersen, H.C. *J. Chem. Phys.*, **76**(1):637–649, 1982.

[70] Kvasnica, J. *Statistická fyzika*. Academia, 1998.

[71] Rapaport, D.C. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.

[72] Flyvbjerg, H. and Petersen, H.G. *J. Chem. Phys.*, **91**(1):461–466, 1989.

[73] Szabo, A. and Ostlund, N.S. *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory*. McGraw-Hill Companies, Inc., 1982.

[74] Helgaker, T.; Jorgensen, P.; and Olsen, J. *Molecular Electronic Structure Theory*. John Wiley & Sons, Ltd., 2000.

[75] Jensen, F. *Introduction to Computational Chemistry*. John Wiley & Sons, Ltd., 2007.

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

Ortiz, J.V.; Cioslowski, J.; and Fox, D.J. Gaussian 09, Revision A.1, 2009. Gaussian, Inc., Wallingford, CT.

[77] Phillips, J.C. and Kleinman, L. *Phys. Rev.*, **116**:287–294, 1959.

[78] Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; and Preuß, H. *Theor. Chim. Acta*, **77**(2):123–14, 1990.

[79] Roothaan, C.C.J. *Rev. Mod. Phys.*, **23**:69–89, 1951.

[80] Pople, J.A. and Nesbet, R.K. *J. Chem. Phys.*, **22**:571–572, 1954.

[81] Møller, C. and Plesset, M.S. *Phys. Rev.*, **46**:618–622, 1934.

[82] Čížek, J. *J. Chem. Phys.*, **45**:4256–4266, 1966.

[83] Lipkowitz, K.B.; Boyd, D.B.; Crawford, T.D.; and Schaefer, H.F. *Rev. Comp. Chem.*, **14**, 2007.

[84] Sousa, S.F.; Fernandes, P.A.; and Ramos, M.J. *J. Phys. Chem. A*, **111**:10439–10452, 2007.

[85] Hohenberg, P. and Kohn., W. *Phys. Rev.*, **136**:B864–B871, 1964.

[86] Kohn, W. and Sham, L.J. *Phys. Rev.*, **140**:A1133–A1138, 1965.

[87] Parr, R.G. and Yang, W. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, 1994.

[88] Dirac, P.A.M. *Math. Proc. Cambridge Philos. Soc.*, **26**:376–385, 1930.

[89] Slater, J.C. *Phys. Rev.*, **81**:385–390, 1951.

[90] Vosko, S.H.; Wilk, L.; and Nusair, M. *Can. J. Phys.*, **58**:1200–1211, 1980.

[91] Langreth, D.C. and Perdew, J.P. *Phys. Rev. B*, **21**:5469–5493, 1980.

[92] Langreth, D.C. and Mehl, M.J. *Phys. Rev. B*, **28**:1809–1834, 1983.

[93] Perdew, J.P. and Yue, W. *Phys. Rev. B*, **33**:8800–8802, 1986.

[94] Perdew, J.P. *Phys. Rev. B*, **33**:8822–8824, 1986.

[95] Becke, A.D. *Phys. Rev. A*, **38**(6):3098–3100, 1988.

[96] Perdew, J.P. and Wang, Y. *Phys. Rev. B*, **33**:8800–8802, 1986.

[97] Adamo, C. and Barone, V. *J. Chem. Phys.*, **108**:664–675, 1998.

[98] Handy, N.C. and Cohen, A.J. *Mol. Phys.*, **99**:403–412, 2001.

[99] Xu, X. and Goddard, William A., I. *Proc. Natl. Acad. Sci, U.S.A.*, **101**:2673–2677, 2004.

[100] Becke, A.D. *J. Chem. Phys.*, **84**:4524–4529, 1986.

[101] Perdew, J.P. *Phys. Rev. B*, **33**:8822–8824, 1986.

[102] Perdew, J.P.; Burke, K.; and Ernzerhof, M. *Phys. Rev. Lett.*, **77**:3865–3868, 1996.

[103] Adamo, C. and Barone, V. *J. Chem. Phys.*, **116**:5933–5940, 2002.

[104] Becke, A.D. *J. Chem. Phys.*, **88**:1053–1062, 1988.

[105] Perdew, J.P. In P. Ziesche and H. Eschig, editors, *Electronic Structure of Solids '91*, pages 11–20. Akademie Verlag, Berlin, Germany, 1991.

[106] Lee, C.; Yang, W.; and Parr, R.G. *Phys. Rev. B*, **37**:785–789, 1988.

[107] Becke, A.D. *J. Chem. Phys.*, **104**:1040–1046, 1996.

[108] Krieger, J.B.; Chen, J.; Iafrate, G.J.; and Savin, A. *Electron Correl. Mater. Prop.*, **3**:463–477, 1999.

[109] Tao, J.; Perdew, J.P.; Staroverov, V.N.; and Scuseria, G.E. *Phys. Rev. Lett.*, **91**:146401, 2003.

[110] Voorhis, T.V. and Scuseria, G.E. *J. Chem. Phys.*, **109**:400–410, 1998.

[111] Becke, A.D. *J. Chem. Phys.*, **98**:5648–5652, 1993.

[112] Yang, W. *Phys. Rev. Lett*, **66**:1438–1441, 1991.

[113] Kobayashi, M.; Akama, T.; and Nakai, H. *J. Chem. Phys.*, **125**:204106, 2006.

[114] Kobayashi, M.; Imamura, Y.; and Nakai, H. *J. Chem. Phys.*, **127**:074103, 2007.

[115] Kobayashi, M. and Nakai, H. *J. Chem. Phys.*, **129**:044103, 2008.

[116] Kobayashi, M. and Nakai, H. *J. Chem. Phys.*, **131**:114108, 2009.

[117] Kitaura, K.; Ikeo, F.; Asada, T.; Nakano, T.; and Uebayashi, M. *Chem. Phys. Lett.*, **313**(3-4):701–706, 1999.

[118] Tsuneyuki, S.; Kobori, T.; Akagi, K.; Sodeyama, K.; Terakura, K.; and Fukuyama, H. *Chem. Phys. Lett.*, **476**(1-3):104–108, 2009.

[119] Warshel, A. and Levit, M. *J. Mol. Biol.*, **103**(2):227–249, 1976.

[120] Singh, U.C. and Kollman, P.A. *J. Comput. Chem.*, **7**:718–730, 1986.

[121] Field, M.J.; Bash, P.A.; and Karplus, M. *J. Comput. Chem.*, **11**(6):700–733, 1990.

[122] Bakowies, D. and Thiel, W. *J. Phys. Chem.*, **100**:10580–10594, 1996.

[123] Gao, J. *Acc. Chem. Res.*, **29**:298–305, 1996.

[124] Maseras, F. and Morokuma, K. *J. Comput. Chem.*, **16**(9):1170–1179, 1995.

[125] Dapprich, S.; Komaromi, I.; Byun, K.S.; Morokuma, K.; and Frisch, M. *J. Mol. Struct. Theochem*, **461-462**:1–21, 1999.

[126] Svensson, M.; Humbel, S.; Froese, R.D.J.; Matsubara, T.; Sieber, S.; and Morokuma, K. *J. Phys. Chem.*, **100**:19357–19363, 1996.

[127] Sherwood, P. *Modern Methods and Algorithms of Quantum Chemistry*, **3**:285–305, 2000.

[128] Vreven, T.; Byun, K.S.; Komaromi, I.; Dapprich, S.; Montgomery, John A., J.; Morokuma, K.; and Frish, M.J. *J. Chem. Theory Comput.*, **2**:815–826, 2006.

[129] Senn, H.M. and Thiel, W. *Top. Curr. Chem.*, **268**:173–290, 2007.

[130] Thiel, W. *Multiscale Sim. Meth. Mol. Sci.*, **42**:203–214, 2009.

[131] Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H.B.; and Frisch, M.J. *J. Comput. Chem.*, **24**:760 –769, 2003.

[132] Vreven, T.; Frisch, M.J.; Kudin, K.N.; Schlegel, H.B.; and Morokuma, K. *Mol. Phys.*, **104**:701–714, 2006.

[133] Théry, V.; Rinaldi, D.; Rivail, J.L.; Maigret, B.; and Ferencz, G.G. *J. Comp. Chem.*, **15**(3):269–282, 1994.

[134] Monard, G.; Loos, M.; Théry, V.; Baka, K.; and Rivai, J.L. *Int. J. Quant. Chem.*, **58**:153–159, 1996.

[135] Assfeld, X. and Rivail, J.L. *Chem. Phys. Lett.*, **263**(1-2):100–106, 1996.

[136] Reuter, N.; Dejaegere, A.; Maigret, B.; and Karplus, M. *J. Phys. Chem. A*, **104**(8):1720–1735, 2000.

[137] Mulliken, R.S. *J. Chem. Phys*, **23**:1833, 1955.

[138] Reed, A.E.; Weinstock, R.B.; and Weinhold, F. *J. Chem. Phys.*, **83**:735–746, 1985.

[139] Singh, U.C. and Kollman, P.A. *J. Comp. Chem.*, **5**(2):129–145, 1984.

[140] Besler, B.H.; Merz, K.M.; and Kollman, P.A. *J. Comp. Chem.*, **11**(4):431–439, 1990.

[141] Ponder, J.W. and Case, D.A. *Adv. Protein Chem.*, **66**:27–85, 2003.

[142] Rappe, A.K. and Goddard, W.A.I. *J. Phys. Chem.*, **95**:3358–3363, 1991.

[143] Rick, S.W.; Stuart, S.J.; and Berne, B.J. *J. Chem. Phys.*, **101**:6141–6156, 1994.

[144] Kitao, O. and Ogawa, T. *Mol. Phys.*, **1-2**:3–17, 2003.

[145] Patel, S.; MacKerell, Alexander D., J.; and Brooks, Charles L., I. *J. Comput. Chem.*, **25**:1504–1514, 2004.

[146] Mitchell, P.J. and Fincham, D. *J. Phys.: Condens. Matter*, **5**:1031–1038, 1993.

[147] Anisimov, V.M.; Lamoureux, G.; Vorobyov, I.V.; Huang, N.; Roux, B.; and Mackerell, A. D., J. *J. Chem. Theory Comput.*, **1**:153–168, 2005.

[148] Vorobyov, I.V.; Anisimov, V.M.; and MacKerell, A. D., J. *J. Phys. Chem. B*, **109**:18988–18999, 2005.

[149] Halgren, T.A. and Damm, W. *Curr. Opin. Chem. Biol.*, **11**:236–242, 2001.

[150] Pauling, L. *General Chemistry*. Dover Publications, Inc., 1988.

[151] Olander, D.R. *General Thermodynamics*. CRC Press, 2008.

[152] Alberty, R.A. and Oppenheim, I. *J. Chem. Phys.*, **89**(6):3689–3693, 1988.

[153] Alberty, R.A. *J. Chem. Thermodynamics*, **29**:501–516, 1997.

[154] Eyring, H. *J. Chem. Phys*, **3**:107–115, 1935.

[155] Laidler, K.J. and King, C.M. *J. Phys. Chem.*, **87**:2657–2664, 1983.

[156] Truhlar, D.G.; Garrett, B.C.; and Klippenstein, S.J. *J. Phys. Chem.*, **100**(31):12771, 1996.

[157] Ochterski, J.W. Thermochemisty in gaussian, 2000.

[158] van Gunsteren, W.F.; Daura, X.; and Mark, A.E. *Helv. Chim. Acta*, **85**:3113–3128, 2002.

[159] Chipot, C. and Pohorille, A., editors. *Free Energy Calculations*. Springer Verlag, 2007.

[160] Torrie, G.M. and Valleau, J.P. *Chem. Phys. Lett.*, **28**(4):578–581, 1974.

[161] Torrie, G.M. and Valleau, J.P. *J. Comput. Chem.*, **23**:187–199, 1977.

[162] Kästner, J.; Senn, H.M.; Thiel, S.; Otte, N.; and Thiel, W. *J. Chem. Theory Comput.*, **2**(2):452–461, 2006.

[163] Kästner, J. *Comput. Mol. Sci.*, **1**:932, 2011.

[164] Kumar, S.; Rosenberg, J.M.; Bouzida, D.; Swendsen, R.H.; and Kollman, P.A. *J. Comp. Chem.*, **13**(8):1011–1021, 1992.

[165] Souaille, M. and Roux, B. *Comput. Phys. Commun.*, **135**:40–57, 2001.

[166] Kästner, J. and Thiel, W. *J. Chem. Phys.*, **123**:144101, 2005.

[167] Kästner, J. and Thiel, W. *J. Chem. Phys.*, **124**:234106, 2006.

[168] Case, D.A.; Darden, T.A.; Cheatham, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Walker, R.C.; Zhang, W.; Merz, K.M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K.F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.J.; Cui, G.; Roe, D.R.; Mathews, D.H.; Seetin, M.G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; and Kollman, P.A. AMBER 11, 2010. University of California, San

Francisco.

[169] Spoel, D.v.d.; Lindahl, E.; Hess, B.; Buuren, A.R.v.; Apol, E.; Meulenhoff, P.J.; Tieleman, D.P.; Sijbers, A.L.T.M.; Feenstra, K.A.; Drunen, R.v.; Tieleman, D.A.; Sijbers, A.L.T.M.; Feenstra, K.A.; Drunen, R.v.; and Berendsen, H.J.C. *Gromacs User Manual version 4.5.4.* University of Groningen, 2010.

[170] Baerends, E.J.; Ziegler, T.; Autschbach, J.; Bashford, D.; Berces, A.; Bickelhaupt, F.M.; Bo, C.; Boerrigter, P.M.; Cavallo, L.; Chong, D.P.; Deng, L. Dickson, R.M.; Ellis, D.E.; Faassen, M.v.; Fan, L.; Fischer, T.H.; Fonseca Guerra, C.; Ghysels, A.; Giammona, A.; Gisbergen, S.J.A.v.; Gotz, A.W.; Groeneveld, J.A.; Gritsenko, O.V.; Gruning, M.; Gusarov, S.; Harris, F.E.; Hoek, P.v.d.; Jacob, C.R.; Jacobsen, H.; Jensen, L.; Kaminski, J.W.; Kessel, G.v.; Kootstra, F.; Kovalenko, A.; Krykunov, M.V.; Lenthe, E.v.; McCormack, D.A.; Michalak, A.; Mitoraj, M.; Neugebauer, J.; Nicu, V.P.; Noodleman, L.; Osinge, V.P.; Patchkovskii, S.; Philipsen, P.H.T.; Post, D.; Pye, C.C.; Ravenek, W.; Rodriguez, J.I.; Ros, P.; Schipper, P.R.T.; Schreckenbach, G.; Seldenthuis, J.S.; Seth, M.; Snijders, J.G.; Sola, M.; Swart, M.; Swerhome, D.; Velde, G.t.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T.A.; Wezenbeek, E.M.v.; Wiesenekker, G.; Wolff, S.K.; Woo, T.K.; and Yakovlev, A.L. *ADF Program System, Release 2010, SCM.* Scientific Computing & Modelling NV, Vrije Universiteit, 2010.

[171] Schmidt, M.W.; Baldridge, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.; Windus, T.L.; Dupuis, M.; and A. M.J. *J. Comput. Chem.*, **14**:1347–1363, 1993.

[172] Aquilante, F.; Vico, L.D.; Ferre, N.; Ghigo, G.; Malmqvist, P.a.; Neogrady, P.; Pedersen, T.B.; Pitonak, M.; Reiher, M.; Roos, B.O.; Serrano-Andres, L.; Urban, M.; Veryazov, V.; and Lindh, R. *J. Comput. Chem.*, **31**:224–247, 2010.

[173] Valiev, M.; Bylaska, E.J.; Govind, N.; Kowalski, K.; Straatsma, T.P.; Dam, H.J.J.V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.L.; and Jong, W.A. *Comp. Phys. Comm.*, **181**:1477–1489, 2010.

[174] Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S.T.; Gilbert, A.T.B.; Slipchenko, L.V.; Levchenko, S.V.; O'Neill, D.P.; DiStasio, Robert A., J.; Lochan, R.C.; Wang, T.; Beran, G.J.O.; Besley, N.A.; Herbert, J.M.; Lin, C.Y.; Voorhis, T.V.; Chien, S.H.; Sodt, A.; Steele, R.P.; Rassolov, Vitaly A. Maslen, P.E.; Korambath, P.P.; Adamson, R.D.; Austin, B.; Baker, J.; Byrd, E.F.C.; Daschel, H.; Doerksen, R.J.; Dreuw, A.; Dunietz, B.D.; Dutoi, A.D.; Furlani, T.R.; Gwaltney, S.R.; Heyden, A.; Hirata, S.; Hsu, C.P.; Kedziora, G.; Khaliullin, R.Z.; Klunzinger, P.; Lee, A.M.; Lee, M.S.; Liang, W.Z.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E.I.; Pieniazek, P.A.; Rhee, Y.M.; Ritchie, J.; Rosta, E.; Sherrill, D.; Simmonett, A.C.; Subotnik, J.E.; Woodcock, H. Lee, I.; Zhang, W.; Bell, A.T.; Chakraborty, A.K.; Chipman, D.M.; Keil, F.J.; Warshel, A.; Hehre, W.J.; Schaefer, Henry F., I.; Kong, J.; Krylov, A.I.; Gill, P.M.W.; and Head-Gordon, M. *Phys. Chem. Chem. Phys.*, **8**:3172, 2006.

[175] Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; and Kölmel, C. *Chem. Phys. Letters*, **162**:165, 1989.

[176] Philipp, D.M. and Friesner, R.A. *J. Comp. Chem.*, **20**(14):1468–1494, 1999.

[177] Murphy, R.B.; Philipp, D.M.; and Friesne, R.A. *J. Comp. Chem.*, **21**(16):1442–1457, 2000.

[178] Sherwood, P.; Vries, A.H.d.; Guest, M.F.; Schreckenbach, G.; Catlow, C.R.A.; French,

S.A.; Sokol, A.A.; Bromley, S.T.; Thiel, W.; Turner, A.J.; Billeterc, S.; Terstegenc, F.; Thiel, S.; Kendrick, J.; Rogers, S.C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schaferg, A.; and Lennartz, C. *J. Mol. Struct. Theochem*, **632**:1–28, 2003.

[179] Lin, H.; Y., Z.; Pezeshki, S.; and Truhlar, D.G. Qmmm - version 1.3.8, 2009. University of Minnesota, Minneapolis.

[180] Ryde, U. *J. Comput.-Aided Mol. Des.*, **10**:153–164, 1996.

[181] Ryde, U. and Olsson, M.H.M. *Int. J. Quantum Chem.*, **81**:335–347, 2001.

[182] Ryde, U.; Olsen, L.; and Nilsson, K. *J. Comput. Chem.*, **23**:1058–1070, 2002.

[183] Futera, Z. *Universal Program Layer for Coupling QM and MM Software: QMS-Uni. User Manual.* Charles University in Prague, 2011.

[184] Case, D.A.; Darden, T.A.; Cheatham, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Merz, K.M.; Wang, B.; Pearlman, D.A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J.W.; Ross, W.S.; and Kollman, P. AMBER 8, 2004. University of California, San Francisco.

[185] Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Montgomery, Jr., J.A.; Vreven, T.; Kudin, K.N.; Burant, J.C.; Millam, J.M.; Iyengar, S.S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J.E.; Hratchian, H.P.; Cross, J.B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Ayala, P.Y.; Morokuma, K.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Zakrzewski, V.G.; Dapprich, S.; Daniels, A.D.; Strain, M.C.; Farkas, O.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Ortiz, J.V.; Cui, Q.; Baboul, A.G.; Clifford, S.; Cioslowski, J.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Challacombe, M.; Gill, P.M.W.; Johnson, B.; Chen, W.; Wong, M.W.; Gonzalez, C.; and Pople, J.A. Gaussian 03, Revision C.02, 2004. Gaussian, Inc., Wallingford, CT.

[186] Werner, H.J.; Knowles, P.J.; Lindh, R.; Manby, F.R.; Schütz, M.; Celani, P.; Korona, T.; Mitrushenkov, A.; Rauhut, G.; Adler, T.B.; Amos, R.D.; Bernhardsson, A.; Berning, A.; Cooper, D.L.; Deegan, M.J.O.; Dobbyn, A.J.; Eckert, F.; Goll, E.; Hampel, C.; Hetzer, G.; Hrenar, T.; Knizia, G.; Köppl, C.; Liu, Y.; Lloyd, A.W.; Mata, R.A.; May, A.J.; McNicholas, S.J.; Meyer, W.; Mura, M.E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Schumann, U.; Stoll, H.; Stone, A.J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; and Wolf, A. MOLPRO, version 2006.1, a package of ab initio programs, 2008.

[187] POSIX.1c, Threads extensions (IEEE Std 1003.1c-1995).

[188] Kruse, R.L. *Data Structures Program Design.* Prentice-Hall, 1994.

[189] Nocedal, J. and Wright, S.J. *Numerical Optimization.* Springer Verlag, 1999.

[190] Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; and Vetterling, W.T. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 1992.

[191] Wolfe, P. *Society for Industrial and Applied Mathematics*, **11**:226–235, 1967.

[192] Fletcher, R. and Reeves, C.M. *Comp. J.*, **7**:149–154, 1964.

[193] Polak, B. and Ribiere, G. *Rev. Fr. Imform. Rech. Oper.*, **16**:35–43, 1969.

[194] Hestenes, M.R. and Stiefel, E. *Journal of Research of the National Bureau of Standards*,

49:409–436, 1952.

[195] Liu, D.C. and Nocedal, J. *Mathematical Programming*, **45**:503–528, 1989.

[196] Andersen, H.C. *J. Chem. Phys.*, **72**:2384–2393, 1980.

[197] Hünenberger, P.H. *Adv. Polymer. Sci.*, **173**:105–149, 2005.

[198] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; and Haak, J.R. *J. Chem. Phys.*, **81**:3684–3690, 1984.

[199] Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; and Preuß, H. *Theor. Chim. Acta*, **78**:247–266, 1991.

[200] Bergner, A.; Dolg, M.; Kuechle, W.; Stoll, H.; and Preuss, H. *Mol. Phys.*, **80**:1431, 1993.

[201] Klamt, A. and Schüürmann, G. *J. Chem. Soc. Perkin Trans.*, **2**:799, 1993.

[202] Barone, V. and Cossi, M. *J. Phys. Chem. A*, **102**:1998, 1995.

[203] Zhang, W.; Hou, T.; Schafmeister, C.; Ross, W.S.; and Case, D.A. Leap, 2010. In AmberTools v1.4.

[204] Bayly, C.I.; Cieplak, P.; Cornell, W.; and Kollman, P.A. *J. Phys. Chem.*, **97**(40):10269–10280, 1993.

[205] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; and Kollman, P.A. *J. Am. Chem. Soc.*, **115**(21):9620–9631, 1993.

[206] Cieplak, P.; Cornell, W.D.; Bayly, C.; and Kollman, P.A. *J. Comput. Chem.*, **16**:1357–1377, 1995.

[207] Fox, T. and Kollman, P.A. *J. Phys. Chem. B*, **102**(41):8070–8079, 1998.

[208] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; and Klein, M.L. *J. Chem. Phys.*, **79**(2):926–935, 1983.

[209] Kollman, P.A. *Acc. Chem. Res.*, **29**:461–469, 1996.

[210] Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M.C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Taisung, L.; Claldwell, J.; Wang, J.; and Kollman, P. *J. Comp. Chem.*, **24**(16):1999–2012, 2003.

[211] Arnott, S. and Hukins, D.W.L. *J. Mol. Biol.*, **81**(2):93–105, 1973.

[212] Grimme, S. *J. Comp. Chem.*, **27**(15):1787–1799, 2006.

[213] Hay, P.J. and Wadt, W.R. *J. Chem. Phys*, **82**:270–283, 1985.

[214] Wadt, W.R. and Hay, P.J. *J. Chem. Phys*, **82**:284–298, 1985.

[215] Hay, P.J. and Wadt, W.R. *J. Chem. Phys*, **82**:299–310, 1985.

[216] Zimmermann, T.; Leszczynski, J.; and Burda, J.V. *J. Mol. Mod.*, **17**(9):2385–2393, 2011.

[217] Wang, F.Y.; Habtemariam, A.; van der Geer, E.P.L.; Fernandez, R.; Melchart, M.; Deeth, R.J.; Aird, R.; Guichard, S.; Fabbiani, F.P.A.; Lozano-Casal, P.; Oswald, I.D.H.; Jodrell, D.I.; Parsons, S.; and Sadler, P.J. *Proc Natl Acad Sci USA*, **102**:18269, 2005.

[218] Breneman, C.M. and Wiberg, K.B. *J. Comp. Chem.*, **11**(3):361–373, 1990.

[219] Zimmermann, T.; Zeizinger, M.; and Burda, J.V. *J. Inorg. Biochem.*, **99**:2184–2196, 2005.

List of Publications

1. Futera, Zdeněk; Klenko, Julia; Šponer, Judit E.; Šponer, Jiří; Burda, Jaroslav V.: Interaction of the "Piano–stool" [Ruthenium(II)(η^6–arene)(en)Cl]$^+$ Complexes With Water and Nucleobases; Ab initio and DFT Study, *J.Comput.Chem.* **30**, 2009, p1758-1770

2. Chval, Zdeněk; Futera, Zdeněk; Burda, Jaroslav V.: Comparison of hydration reactions for "piano–stool" RAPTA-B and [Ru(η^6–arene)(en)Cl]$^+$ complexes: Density functional theory computational study, *J.Chem.Phys.* **134**, 2011, 024520

3. Futera, Zdeněk; Koval, Tomáš; Leszczynski, Jerzy; Gu, Jiande; Mitoraj, Mariusz; Srebro, Monika; Burda, Jaroslav V.: Exploring a Reaction Mechanism for Acetato Ligand Replacement in Paddlewheel Tetrakisacetatodirhodium(II,II) Complex by Ammonia: Computational Density Functional Theory Study, *J.Phys.Chem. A* **115**, 2011, p784-794

4. Li, Jinghua; Futera, Zdeněk; Li, Hongfang; Tateyama, Yoshitaka; Higuchi, Masayoshi: Conjugation of organic–metallic hybrid polymers and calf–thymus DNA, *Phys.Chem.Chem.Phys.* **13**, 2011, p4839-4841

5. Futera, Zdeněk; Platts, James A.; Burda, Jaroslav V.: Binding of Piano–Stool Ru(II) Complexes to DNA; QM/MM Study, *J.Comput.Chem.*, 2012, in print

6. Li, Jinghua; Futera, Zdeněk; Li, Hongfang; Tateyama, Yoshitaka; Murakami, Tatsuya; Higuchi, Masayoshi: DNA Binding Property and Cytotoxicity of Metallo-Supramolecular Polymers, *Chem.Mat.*, submitted

7. Futera, Zdeněk and Burda, Jaroslav V.: Reaction Mechanism of Ru(II) Piano–Stool Complexes – QM/MM MD Study, in preparation

Attachments