Technische Universität München

Department Chemie

Fachgebiet Theoretische Chemie

Computational Modeling of Chemical Systems in Solution:

A Microscopic Approach for Efficient Density Functional Calculations with ParaGauss

Bo Li

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Ville R. I. Kaila
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Dr. h.c. Notker Rösch (i.R.)
	2.	Univ.-Prof. Dr. Klaus Köhler

Die Dissertation wurde am 22.09.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 02.11.2015 angenommen.

Acknowledgment

It is my great honor to be the last PhD student supervised by Professor Notker Rösch at Technische Universität München (TUM). To him I would like to express my gratitude for giving me the opportunity to study this interesting topic in his group and his guidance of my scientific work.

My special thanks further go to my doctoral project team leader, Dr. Alexei V. Matveev, for the various discussions of scientific topics and mathematical problems, as well as his generous help of improving my programming skills. Another person whom I am greatly indebted to is Dr. Sven Krüger, for his continuous support of both scientific and administrative issues during the past years. I also would like to thank Dr. Alexander Genest for his help and various discussions during my overseas research stay at the Institute of High Performance Computing (IHPC) in Singapore.

I want thank my second supervisor of my doctoral project, Prof. Hans-Joachim Bungartz, for his collaboration to make this project possible and various discussions during the method implementation phase.

I especially want to thank my colleagues and friends Dr. Cheng-chau Chiu and Dr. Thomas Soini, for their help both in work and life.

Furthermore, I would like to thank all present and past colleagues at the TUM and at the Catalysis Modelling Group at IHPC, Benjamin Chen, Dr. Duygu Başaran, George Beridze, Dr. Ion Chiorescu, Dr. Konstantina Damianos, Dr. Shrabani Dinda, Dr. Wilhelm Eger, Dr. Agalya Govindasamy, Stefan Kienzle, Dr. Alena Kremleva, Dr. Xiufang Ma, Dr. Remi Marchal, Dr. Astrid Nikodem, Dr. Suwit Suthirakun, Dr. Yin Wu, Dr. Lili Zhao, Dr. Zhijian Zhao, for the good working atmosphere.

I also thank the International Graduate School of Science and Engineering at the TUM for the generous scholarship to support my doctoral project.

Last but not least I thank my family for their love, support, understanding, and encouraging, which enabled me to complete this work.

Contents

List of Abbreviations

ABS	Auxiliary Basis Set
AIMD	Ab initio Molecular Dynamics
ASC	Apparent Surface Charges
BEM	Boundary Element Method
BGY	Born-Green-Yvon
BO	Born-Oppenheimer
COSMO	Conductor-like Screening Model
DFT	Density Functional Theory
FF	Force Field
HF	Hartree-Fock
IET	Integral Equation Theory
KS	Kohn-Sham
LCGTO-FF	the Linear Combination of Gaussian-Type Orbitals Fitting-Functions
MC	Monte Carlo
MD	Molecular Dynamics
MM	Molecular Mechanics
MO	Molecular Orbital
MSA	Mean-spherical Approximation
OPLS	Optimized Potential for Liquid Simulations
OPLS-AA	Optimized Potential for Liquid Simulations All-Atom
OZ	Ornstein-Zernike
PCM	Polarizable Continuum Model
PES	Potential Energy Surface
PMF	Potential of Mean Force
PMV	Partial Molar Volume
PSE*n*	Partial Series Expansion of order *n*
PT1	the first-order Perturbation Theory
PT2	the second-order Perturbation Theory
PY	Percus-Yevick
QM	Quantum Mechanics
RDF	Radial Distribution Function
RISM	Reference Interaction Site Model
SAS	Solvent-Accessible Surface

SCF	Self-Consistent Field
SES	Solvent-Excluded Surface
SPC	Simple Point Charge
SPC/E	Extended Simple Point Charge
SSOZ	Site-site Ornstein Zernike
TI	Thermodynamic Integration
TIP3P	Transferable-Intermolecular-Potential-3-Point-charge
TIP4P	Transferable-Intermolecular-Potential-4-Point-charge
UFF	Universal Force Field
vdW	van der Waals
VWS	van der Waals Surface

1. Introduction

1.1. Solvent Effects in Computational Chemistry

Understanding the thermodynamics and kinetics of chemical reactions is one of the most important issues in the field of chemistry.[1] While chemical thermodynamics can be described by the free energy changes of the species involved in chemical reactions, reaction kinetics is characterized by the rates of the chemical reactions. Key quantities are the energy barriers between reactants and products; the related barrier heights determine the branching ratios of competing products.[1,2] The development of computational chemistry nowadays allows one to determine the electronic structure of molecules in the gas phase with a high level accuracy at affordable cost.[3,4] The electronic structure description of quantum chemical applications in turn permits one to explore the stationary points on the potential energy surface (PES) of the system under study.[3,4] To this end free energy differences between minima on the PES can be calculated and the transition state structures related to individual reaction paths and the activation barriers of chemical reactions can be localized and characterized.[2-4] For isolated molecules, the electronic energy changes of the species involved are solely determining the chemical reaction.[2,4] However, for many chemical reactions occurring in the solution phase, the solvation energies of the reactants due to the presence of the solvent are often comparable to the change of the electronic energy of the solute; in short, the PES can be affected notably due to a solvent.[3-5] In addition, the solute electronic structure and thus also the physical properties of the solute itself can be significantly influenced by the solvent medium.[5-7] As a result the solvation environment of a molecule can considerably affect its properties and thus also change the products, the rate of its chemical reactions, and even the reaction mechanism.

One striking example for such solvent effects is the Menshutkin type S_N2 reaction[8] where the nucleophilic displacement of a halide in chloromethane (CH_3Cl) by an amine (NH_3) generates a halide anion (Cl^-) and an alkylammonium cation ($H_3NCH_3^+$). In the gas phase this reaction is endergonic by 110 kcal/mol.[9] On the other hand, the solvation energies of the Cl^- and $H_3NCH_3^+$ ions in water are large enough that the reaction becomes exergonic by –34 kcal/mol with the activation energy barrier estimated as about 24 kcal/mol.[4,5,9] With the proper solvation models, quantum mechanics (QM) calculations successfully interpret the solvent effect on the reaction mechanism and are in good agreement with the experiment.[9-12]

This example demonstrates that modeling solvation effects in computational chemistry is of central importance for understanding chemical reactions of organic and inorganic species in solution for scientific and industrial applications.[5] Examples of pertinent application fields,

which also possess industrial importance, are catalytic reactions for the production of hydrogen from water,[13] the conversion of CO_2 into hydrocarbon fuel,[14] and the transformations of biomass into fuels such as ethanol.[15] Another important application relates to the nuclear fuel cycle and radioactive waste disposal, where the thermodynamic and structural properties of actinides in aqueous solution are crucial for understanding the complex chemistry and chemical processes in this field.[16]

1.2. Modeling the Solvation Effect

Categorized by the interactions between solute and solvent molecules the solvation effects can fall into two broad categories, the *long-range* solvation effects including the polarization and dipole orientation and the *short-range* solvation effects including hydrogen bonding, van der Waals (vdW) interaction, solvent shell structure, solvent-solute dynamics, charge transfer effects, and hydrophobic effects.[3] The electrostatic nature of the long-range solvation effects causes a screening of charge interactions which leads to a macroscopic dielectric constant measuring the polarity. The short-range solvation effects depend on the microscopic structures of the solvent molecules such as their shapes and their ability to form hydrogen bonds. In support of current experimental efforts in various areas of chemistry and materials science in solution, computational methods modeling solvation effects can therefore be partitioned into three types, treating the solvent *implicitly* as a continuum medium, taking account of the solvent molecules *explicitly*, and establishing a hybrid approach that combines aspects of implicit and explicit methods.[5]

The combination of QM calculations with implicit solvation models, e.g. the polarizable continuum model (PCM)[17] and its variants based on the conductor-like solvation model (COSMO),[18] is one of the most popular strategies in computational chemistry for studying systems in solution. This type of implicit solvation models, so-called continuum solvation models, do not provide much insight into the structure and the statistical fluctuations of the solvent. Instead, continuum solvation model solvents are generally characterized by the use of the dielectric constant to account for the long-range electrostatic interactions.[19] With this treatment, however, solvent species with very different physical properties that only share the similar dielectric constants, e.g. ethanol and acetonitrile, can be nearly indistinguishable.[5] To discriminate such solvents, continuum solvation models have to rely on the empirical parameters such as effective radius to define the interface between solute and solvent by constructing the van der Waals surface (VWS), the solvent-accessible surface (SAS) and the solvent-excluded surface (SES).[20] Note, however, that this simple approximation still does not properly describe important short-range interaction between solute and solvent, such as

hydrogen bonding.[4,5]

Another approach to solvation is the so-called explicit solvation model, where one treats solute and solvent of the same level of theory by carrying out molecular dynamics (MD) simulations from first principles (*ab initio* MD – AIMD, or, more appropriately first-principles MD – FPMD) or MD simulations based on a force-field description (molecular mechanics – MM) of both the solute and the solvent medium (MM-MD). Such methods explore the phase-space or configuration space of the solution system to obtain physical observables in terms of statistical averages of thermal quantities.[3,4] However, the statistical sampling of large numbers of solvent configurations also entails significant computational costs. Explicit QM dynamic simulations are therefore restricted to models with a comparably small number of solvent molecules.[5] The MM solvent description provided by MM-MD or hybrid QM/MM methods[21,22] can reduce the computational cost. Yet, the results of statistical sampling are only as good as the effective force field used to describe the solvent-solvent and solute-solvent interactions. Furthermore, to compete with continuum models in reproducing electrostatic interactions also one has to account for the polarizability of the solvent medium.[23]

Given the difficulties mentioned above for both continuum and explicit solvation models, strategies which combine QM methods with integral equation theory (IET) of molecular liquids[5,24] are regarded as promising for studying ions and complexes and their chemical reactions in solution with sufficient accuracy and at an affordable cost. A very intuitive approach of IET is based on the Yvon-Born-Green (YBG) hierarchy equations,[24] which has successfully been applied to simple fluids[24] and molecular systems[25,26] to determine the particle pair correlation functions.

Another more popular IET approach is the reference interaction site model (RISM) also known as site-site Ornstein Zernike (SSOZ) equation,[27] which provides solvation (free) energies and also averaged solvent structures via site-site radial distribution functions (RDF). In its original version of 1D RISM, the site-site RDFs are treated based on their spherically symmetric averages.[5,27] With its extension (XRISM)[28] to treat properly the electrostatic interaction in the molecular fluid system,[29] 1D RISM has successfully been applied to study aqueous and ionic solutions.[5,6,30] However, the radial averaging treatment of the site-site RDF is unable to describe properly the spatial information of the solvent structure around a molecular solute of arbitrary shape.[5,30] Therefore, a three-dimensional representation of RISM, 3D RISM[31] was developed to provide a more detailed picture of the solvation structure, which is particularly important in the case of large complex solute molecules.[5,30] Realizations of 3D RISM not only successfully describe the solvation of simple ions[32-34] as well as of polar and non-polar solute molecules,[35] but have proven useful also when applied to

supramolecules,[36-38] proteins,[36,39] drug design,[40,41] and interfaces.[42,43]

Inspired by these promising applications of force-field based RISM, the combination of RISM with high-level QM methods attracts more and more interest and has been explored extensively during the past decade.[6,30] The hybrid approach of RISM together with a self-consistent field (SCF) QM description of the solute system was first proposed at the level of Hartree-Fock theory[44] while later on also Kohn-Sham density functional theory (KS-DFT) was combined with 3D RISM.[45] Thus far, several implementations[46-52] combined the RISM approach with state-of-the-art QM descriptions of the solute system and provide an appealing alternative to classical implicit and explicit solvation models. Hybrid RISM-SCF methods have successfully been applied to calculate the acidity of hydrogen halides[53] and the basicity of substituted methylamines,[54] the conformational equilibrium of different isomers,[50,55,56] the molecular structures of urea,[57] triiodide ion,[58] carbonate and nitrate cations[59] altered by solvation, important chemical reactions such as the S_N2 reactions[11,60] and excited states in solution systems.[5,6,30]

Compared to the traditional continuum and explicit solvation models, RISM approaches account for the microscopic structural information of the solvent at an average level and calculate the thermodynamic properties of solution systems at an affordable cost. However, for being a rather novel tool for modeling solvent effects, there still remain many open issues yet to be clarified in this context.[5,6,30] From the conceptual aspect, RISM methods have mostly been considered as qualitative tools because the solvation free energy is overestimated.[5,6,30] Recent developments in combining both MM[61-64] and QM[65] methods with RISM are exploring the determining factors for the predictive accuracy of this approach, aiming at making high-accuracy calculations possible in the near future.[30] Also the popularity of RISM regarding applications nowadays is steadily increasing. However, RISM methods have not been extensively used to take solvation effects into account in catalytic reactions with transition metal complexes and in actinyl chemistry.[6]

1.3. Thesis Outline

This thesis focuses on the development, implementation, and application of a hybrid approach that combines RISM and KS-DFT. Thereby the electrostatic potential induced by the solute is directly derived from its electron density. This combination, which will be referred to as QM+RISM in this thesis, was added to the parallel density functional program package ParaGauss.[66] The factors affording high-accuracy QM+RISM solvation energy calculations have been explored by suggesting a conceptual interpretation for the correction term of the solvation energy. In addition, QM+RISM is applied for the first time to studying the solvation of actinyl ions and their aqua complexes in an aqueous medium as well as to the overall catalytic cycle for CO_2 conversion in acetonitrile solution with a Ru-bipyridine complex. The remaining part of this introduction is dedicated to a brief description of the structure of this thesis.

Chapter 2 provides the theoretical background of the currently popular solvation models by briefly introducing the state-of-the-art continuum solvation models, explicit solvation models, and integral equation theory (IET) including RISM methods.

Chapter 3 presents details of the implementation and numerical aspects of the coupling of the RISM method with the DFT program ParaGauss. After the basic theory of the combination of RISM with QM is introduced, the second section addresses the interfacing of the RISM solvent environment and the solute, with the latter being treated at the QM level. This requires implementing and evaluating the electrostatic interaction functional for the charge density of the QM region and that of the solvent medium. The third section deals with the construction of free energy functionals in the hybrid approach and presents the expressions approximated by the unperturbed electronic structure of the solute as well as a fully SCF treatment of the solute-solvent interaction. The last section discusses numerical aspects of the implementation including its validation, based on the electrostatic interaction between solute and solvent, as well as its convergence during the corresponding iterative solution of the working equations.

Chapter 4 presents a methodology interpretation and several applications of the hybrid approach. First a benchmark for the RISM implementation is presented: the solvation energies of simple alkali and halide ions in water solution are calculated with 1D and 3D RISM. The values are then compared with those from other RISM solvers, MD calculations, and experiments. Second, the combination of KS-DFT with 3D RISM is systematically studied for a test set of 43 neutral organic solute molecules by applying the known partial molar volume correction for which an alternative interpretation is given. Also the influence of the

relaxation term on the solvation energy is examined based on the numerical results from both approximate and fully self-consistent approaches, as well as based on a simple linear model. Furthermore, the solvation of uranyl in water is thoroughly investigated using the hybrid method of QM with both 1D and 3D RISM. In this application the solvation energies and geometries of uranyl complexes with 4–6 explicit aqua ligands (i.e. the first solvation shell) in the presence of an implicit bulk solvent modeled by 1D and 3D RISM are compared to PCM as well as experiments. In addition, the activation energy for the exchange of water ligands is examined with a hybrid QM+RISM model. As a last application presented in this thesis, the hybrid method is used to explore the CO_2 conversion to $HCOO^-$ in acetonitrile solution with a Ru-bipyridine complex.

2. Theory of Solvation Models

2.1. Continuum Solvation Model

2.1.1. *Polarizable Solvation Model*

The continuum solvation model relies on the basic assumption of using a continuous medium with the macroscopic dielectric constant, that equals the value of the solvent bulk, to represent the environment created by the latter around the solute.[17] To this end, the solute molecule is located in the dielectric continuum in a cavity of arbitrary shape. In this cavity one has the solute electron density $\rho_{\mathrm{M}}(\mathbf{r})$. Then the electrostatic interaction potential $V_{\mathrm{el}}(\mathbf{r})$ between solute and the solvent medium can be determined from the Poisson equation for the space inside the cavity and from the Laplace equation outside the cavity:[67]

$$\begin{aligned} &\Delta V_{\mathrm{el}}(\mathbf{r}) = -4\pi\rho_{\mathrm{M}}(\mathbf{r}), \mathbf{r} \in \textit{inside cavity};\\ &\Delta V_{\mathrm{el}}(\mathbf{r}) = 0, \mathbf{r} \in \textit{outside cavity}. \end{aligned} \tag{2.1.1}$$

In the case of an irregular shape of the cavity of the solute, it is more practical to use the apparent surface charges (ASCs) method.[17] One utilizes the apparent surface charges $\sigma(\mathbf{s})$ spread on the surface Γ of the cavity to obtain the electrostatic potential[17,67]

$$V_{\sigma}(\mathbf{r}) = \int_{\Gamma} \frac{\sigma(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} d^2 s\,. \tag{2.1.2}$$

Here $V_{\sigma}(\mathbf{r})$ corresponds to the interaction potential in Eq. (2.1.1). In general one achieves the solution of ASC method with the help of the boundary element method (BEM) where the cavity surface Γ is approximated by a set of finite elements (tesserae), to which individual apparent surface charges $\sigma(\mathbf{s})$ are assigned.[17]

At the QM level, the problem associated with the continuum solvation model can be expressed by the following Schrödinger equation (SE)[17]

$$[H^0 + V_{\mathrm{R}}]\Psi = E\Psi\,. \tag{2.1.3}$$

Here H^0 is the Hamiltonian of the solute in vacuum, Ψ is the wave function of the solute, E is the energy of the solute, and V_{R} is the interaction potential between the solute and the solvent medium. In the standard polarizable continuum model (PCM),[17] V_{R} can be decomposed as follows

$$V_{\mathrm{R}} = V_{\mathrm{el}} + V_{\mathrm{dis}} + V_{\mathrm{rep}}, \tag{2.1.4}$$

whereas V_{el}, V_{dis} and V_{rep} denote the electrostatic, dispersion and repulsion contributions respectively. Correspondingly, the free energy functional of the solution system is given by[17,19]

$$G_{\mathrm{sol}} = G_{\mathrm{el}} + G_{\mathrm{dis}} + G_{\mathrm{rep}} + G_{\mathrm{cav}} + G_{\mathrm{tm}}, \tag{2.1.5}$$

where the contributions G_{el}, G_{dis} and G_{rep} correspond to the individual terms on the r.h.s of Eq. (2.1.4), respectively. These three terms involve the interaction with the solvent charge density and are included in the Hamiltonian in Eq. (2.1.3) as discretized contributions on the cavity surface.[19] In addition, the electric polarization of the solvent is included in G_{el} as a polarization function $\mathbf{P}$ in terms of the electric field generated by the electrostatic potential, either directly from Eq. (2.1.1) or via the apparent charges as Eq. (2.1.2),[19]

$$\mathbf{P} = -\frac{\varepsilon - 1}{4\pi}\boldsymbol{\Delta} V . \tag{2.1.6}$$

Regarding the last two terms in Eq. (2.1.5), G_{cav} represents the energy required for creating the cavity with the appropriate shape inside the solute. Note that G_{cav} only depends on the geometry of the solute. The last term G_{tm} describes the contributions due to thermal motions of the nuclei.

2.1.2. *Conductor-like Screening Model*

In the original PCM method, the determination of ASCs with the aid of BEM requires special care in order to have a good representation of electrostatic potential everywhere and to save computational cost during the iterative procedure.[67] As an improvement, Klamt and Schüürmann[68] suggested a new approach using a screening conductor theory to describe the reaction field of the continuous solvent. This so-called conductor-like solvation model (COSMO) describes the apparent polarization charge distribution on the cavity surface with the condition of a vanishing total electrostatic potential, Eq. (2.1.1), at the cavity boundary.

With this boundary condition, the dielectric constant of the solvent medium is set to $\varepsilon = \infty$, hence the value of an electric conductor. Afterwards the ASCs of Eq. (2.1.2) are determined by the local value of the electrostatic potential. To recover the dielectric constant of the solvent medium, the cavity surface charge density is defined as the product of the unscreened charge density σ^* with a scaling function $f(\varepsilon)$,[19]

$$\sigma(\mathbf{s}) = f(\varepsilon)\sigma^*(\mathbf{s}) . \tag{2.1.7}$$

The scaling function $f(\varepsilon)$ is then empirically determined by comparing the energies that result from the unscaled COSMO and the correct electrostatic solute-solvent interaction. The

corresponding expression read as[19]

$$f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon - k}, \tag{2.1.8}$$

with the parameter k set to $k = 0.5$ in the original COSMO approach[68] while later on $k = 0$ became more popular.[19] With these modifications the point charges on the cavity surface can be directly calculated instead of using the iterative procedure of the original PCM. Therefore COSMO allows one to include a calculation of the solvent reaction field directly in the SCF procedure. As a result the solute electron density and the solvent reaction field can simultaneously reach convergence, which significantly accelerates the QM calculation.[19]

2.1.3. *Summary*

Continuum solvation models constitute the most widely used approaches for treating solution systems due to their simple formalism and their convenient implementation in QM software packages.[6,17,69] Besides COSMO PCM, other models have been also developed and applied, e.g. integral equation formalism polarizable continuum model (IEFPCM)[70] and self-consistent isodensity polarizable continuum model (SCIPCM).[71]

However, as seen from the conceptual aspects presented in Sections 2.1.1–2.1.2, no continuum model provides insight into the microscopic structure of the solvent. As mean-field theories they attempt to reproduce the average effect of the solute-solvent interactions on the electric structure of the solute by the effects of (correlated) thermal fluctuations of the microscopic structure on thermodynamics properties of the system. The long-range electrostatic interaction can reasonably be described with a macroscopic dielectric constant that represents the solvent medium. However, the treatment of short-range interactions such as dispersion forces or hydrogen bonding relies on empirical quantities defining the molecular cavities and shapes.[4,5,19] The proper choice of the solute cavity size and shape to build the van der Waals surface (VWS), the solvent-accessible surface (SAS) and the solvent-excluded surface (SES) still remains a subject of debate.[17,19,20]

2.2. Explicit Solvation Models

2.2.1. *Molecular Mechanics Models*

As mentioned in Section 2.1.3, continuum solvation models do not include a microscopic description of the solvent structure. A straightforward alternative is to treat molecular solvents explicitly at an atomic level in the energy expression.[4] Molecular mechanics (MM) methods, also referred to as force field (FF) methods, provide simple and empirical descriptions of the solute/solvent molecules. The MM potential energy functional U_{MM} is the sum of various terms:[3,4]

$$U_{\mathrm{MM}} = U_{\mathrm{str}} + U_{\mathrm{bend}} + U_{\mathrm{tors}} + U_{\mathrm{cross}} + U_{\mathrm{vdw}} + U_{\mathrm{el}}\,, \tag{2.2.1}$$

Here U_{str} is the potential energy for bond stretching, U_{bend} represents the potential energy for bending an bond angle, U_{tors} is the torsional energy for rotation around a bond, and U_{cross} describes the coupling between the first three terms. U_{vdw} and U_{el} are the non-bonding interaction potential energies, with the van der Waals (vdW) term U_{vdw} representing the combination of dispersion and repulsion energies and U_{el} the electrostatic potential energy in terms of atomic point charges mimicking the (re)distribution of electrons.

With the appropriate form of the interaction potential energy functional for the solution system, molecular simulation techniques such as molecular dynamics (MD)[3,4] and Monte Carlo (MC) methods[3,4] can be employed to explore the phase space or the configuration space of the solution system to acquire observables as an average of mechanical quantities.[3,4,72]

However, it is worth emphasizing that the MM model is basically a collection of functional forms that depend on the associated constants.[3,4] The parameterization of the force field (FF) is usually done by reproducing experimental measurements as well as by fitting the MM molecular potential energy surface (PES) to the results of high-level QM calculations.[4] For non-aqueous molecules, the MM parameters can be picked from the common dataset, such as the universal force field (UFF)[73] or the optimized potentials for liquid simulations (OPLS) and the OPLS all-atom (OPLS-AA) force field.[74,75] For water as solvent, the most widely applied models are the transferable-intermolecular-potential-3- and -4-point-charge (TIP3P, TIP4P),[76] the simple point charge (SPC),[77] and the extended simple point charge (SPC/E) models.[78] To account properly the solvent polarity, also polarizable force fields have been suggested.[79] However, such polarizable force fields entail higher computational cost and adequate care is required when determining such a parameterization.[80] Therefore, polarizable force fields are not yet widely employed.[3,4]

2.2.2. *Quantum Mechanics Models*

MM models have a number of limitations. The most striking one is that re-parameterization of the specific solute-solvent system is necessary to describe appropriately the bond-breaking/forming events, which reduces the transferability of the FF model.[3,4] To overcome this drawback, the so-called *ab initio* molecular dynamics (AIMD) method was proposed.[81]

AIMD method combines finite temperature dynamics with forces obtained from electronic structure calculations carried out 'on the fly' as the MD simulation proceeds. In the AIMD method, each MD step involves taking a phase point, calculating the electronic energy and the gradients for that point with respect to changes of the positions of the nuclei, then propagating these positions for a short time step to enter the next iteration cycle.[81] Compared to MM MD methods, this process requires a significantly higher computational cost. Therefore even with very powerful computer resources, present-day applications of AIMD remain somewhat limited.[5]

2.2.3. *Hybrid QM/MM Models*

As mentioned in Sections 2.1.3 and 2.2.1, in continuum models the short-range solute-solvent interactions are not properly treated, while in MM models the appropriate description for bond-breaking/forming is missing. To overcome these drawbacks, the hybrid QM/MM method was proposed in which the electronic structure of the solute is fully treated at the QM level while the MM model is introduced to describe the solvent molecules. The effective Hamiltonian of the hybrid system reads as

$$H_{\mathrm{hyb}} = H_{\mathrm{QM}} + H_{\mathrm{MM}} + H_{\mathrm{QM/MM}}\,, \tag{2.2.2}$$

where H_{QM} represents the Hamiltonian of QM solute, H_{MM} accounts for the contribution of all MM solvent molecules, and $H_{\mathrm{QM/MM}}$ represent the interaction between QM solute and MM solvent. The latter contribution usually results from the non-bonding interactions of the MM model, which include the electrostatic interaction and the van der Waals interaction.

Nevertheless, similar to other explicit models (Sections 2.2.1 and 2.2.2), a statistical sampling of the phase space or trajectory space of the MM solvent molecules is required in hybrid QM/MM methods.[3,21] Such a statistical averaging entails significant additional cost, because a reasonable sampling involves a very large number of QM calculations for the solute part.[4] Furthermore the results of the statistical sampling are only as good as the effective force field used to parameterize the solvent-solvent and solvent-solute interactions.

2.2.4. *Summary*

Compared to the continuum solvation models in Section 2.1, explicit solvation models are capable of recovering appropriately the short-range interactions between solvent and solute molecules. However, the simple MM description of both solute and solvents will result in loosing important information such as bond-breaking and forming, while the full QM treatment will significantly increase the cost and, thus, become less applicable. As already highlighted by the Nobel Prize in Chemistry in 2013,[82] the hybrid QM/MM approach seems promising as it combines the advantages of both MM and QM methods. Yet, the system properties of the solution system modelled by a QM/MM method still require the statistical sampling which entails an enormous expense of computational resources.[4] Regarding this difficulty, methods based on integral equation theories (IET)[5] are able to provide the average solvation structures by solving a set of integral equations instead of carrying out the statistical sampling. IET approaches are in some sense intermediate between continuum solvation models and explicit solvation models and, as such, they intend to overcome the drawbacks in the latter two types of solvation models.[5] As IET methods show great potential for modeling solution systems with adequate accuracy and at an affordable cost, they are discussed in the following sections in more detail.

2.3. Liquid State Integral Equation Theory

2.3.1. *The Yvon-Born-Green Hierarchy*

Instead of an explicit treatment of solvent molecules in explicit solvation models, the liquid state integral equation theory (LS-IET) is based on information of the solvent density around the solute in terms of correlation functions. Taking a subsystem of *n* particles from a liquid in equilibrium of *N* particles in total ($n < N$), the distribution of the *n*-particle can be represented as the correlation function $g^{(n)}$ of order *n*. Two successive correlation functions $g^{(n)}$ and $g^{(n+1)}$ are related via the Yvon-Born-Green (YBG-) hierarchy[24]

$$\nabla_{\mathbf{x}_1} g^{(n)}(\mathbf{x}_{(n)}) = \beta \sum_{i=2}^{n} \mathbf{F}(\mathbf{x}_1, \mathbf{x}_i) g^{(n)}(\mathbf{x}_{(n)}) + \beta\rho \int \mathbf{F}(\mathbf{x}_1, \mathbf{x}_{n+1}) g^{(n+1)}(\mathbf{x}_{(n+1)}) d\mathbf{x}_{n+1} \qquad (2.3.1)$$

Here the short-hand notation $\mathbf{x}_{(n)}$ represents the coordinates of *n* particles $(\mathbf{x}_1, ..., \mathbf{x}_n)$, while the pair-wise force contributions $\mathbf{F}_{ij}$ are the derivatives of the pair interaction potential $v(\mathbf{x}_i, \mathbf{x}_j)$

$$\mathbf{F}(\mathbf{x}_i, \mathbf{x}_j) = -\nabla_i v(\mathbf{x}_i, \mathbf{x}_j) . \qquad (2.3.2)$$

Combining the case $n = 2$ for pair correlation functions with the Kirkwood superposition approximation[83]

$$g^{(3)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = g^{(2)}(\mathbf{x}_1, \mathbf{x}_2) g^{(2)}(\mathbf{x}_1, \mathbf{x}_3) g^{(2)}(\mathbf{x}_2, \mathbf{x}_3) , \qquad (2.3.3)$$

one can write Eq. (2.3.1) as the Born-Green equation[24]

$$\nabla_{\mathbf{x}_1} [\ln g^{(2)}(\mathbf{x}_1, \mathbf{x}_2) + \beta v(\mathbf{x}_1, \mathbf{x}_2)] = \beta\rho \int \nabla_{\mathbf{x}_1} v(\mathbf{x}_1, \mathbf{x}_3) g(\mathbf{x}_1, \mathbf{x}_3) [g(\mathbf{x}_2, \mathbf{x}_3) - 1] d\mathbf{x}_3 . \qquad (2.3.4)$$

The Born-Green equation Eq. (2.3.4) can be solved for a given pair potential $v(\mathbf{x}_1, \mathbf{x}_2)$, to yield the pair correlation function $g^{(2)}$.

2.3.2. *The Three-dimensional Born-Green-Yvon Equation*

Based on the YBG-hierarchy and the Born-Green equations as in Eqs. (2.3.1) and (2.3.4), a three-dimensional Born-Green-Yvon equation, denoted as the BGY3dM model,[26,84] was suggested for an arbitrary solute immersed into a molecular solvent.[26,84] The essence of this model, the BGY3d equation, describes the solvent distribution $g(x)$ around a fixed solute with *m*-tuple coordinates $\mathbf{x}_{(m)}$ as follows:

$$\nabla \ln g(\mathbf{x}) = \beta \mathbf{F}(\mathbf{x}, \mathbf{x}_{(m)}) + \beta \int \mathbf{F}(\mathbf{x}, \mathbf{x}') g^{(2)}(\mathbf{x}, \mathbf{x}') g(\mathbf{x}') d\mathbf{x}' , \qquad (2.3.5)$$

Here the pair distribution function $g^{(2)}$ appears as an input and can be computed by the Born-

Green equation, Eq. (2.3.4), which in turn can be regarded as a special case of the BGY3d equation, Eq. (2.3.5). The BGY3d equation can be generalized further for *s* molecular solvent sites, and the generalized equation is called the BGY3dM equation.[26,84] For every solvent site $\alpha = 1,\ldots,s$, the BGY3dM equation reads as

$$\begin{aligned}\nabla_{\mathbf{x}_1^\alpha} \ln g_\alpha(\mathbf{x}_1^\alpha) = \beta\mathbf{F}(\mathbf{x}_1^\alpha) + \beta\rho\sum_{\eta=1}^{s}\int \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_{\alpha\eta}^{(2)}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_\eta(\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta \\ +\beta\sum_{\eta\neq\alpha}^{s}\nabla_{\mathbf{x}_1^\alpha}\int \omega(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g_\eta(\mathbf{x}_1^\eta) d\mathbf{x}_1^\eta\end{aligned}, \qquad (2.3.6)$$

with $\omega_{a\eta}(x_1^\alpha, x_1^\eta) = \delta(r_1^{\alpha\eta} - r_0^{\alpha\eta}) / 4\pi(r_0^{\alpha\eta})^2$ denoting the intramolecular pair distribution function, $r_1^{\alpha\eta} = \left|\mathbf{x}_1^\alpha - \mathbf{x}_1^\eta\right|$ being the distance between particles α and η, and $r_0^{\alpha\eta}$ representing their equilibrium intramolecular distance.

The BGY3dM model yields reasonable solvent distribution functions for solvent models such as hydrogen chloride (HCl) and carbon disulfide (CS_2).[26,84] However, it fails to provide even qualitatively correct results for water, which plays a crucial role as solvent as well as reactant in chemical and biological systems. Alternatively, another model based on IET is known as the reference interaction site model (RISM).[5,24] It does not only provides accurate information on the structure of a molecular solvent such as water, but also approaches chemical accuracy in terms of predictions of solvation thermodynamics.[5] In the next section, a brief outline of the RISM theory will be presented.

2.4. Reference Interaction Site Model

2.4.1. *Ornstein-Zernike Equation*

The theory framework of RISM is based on the seminal work of Ornstein and Zernike from 1914,[85] in which they used an equation to define the total correlation function $h(r)$ between particles 1 and 2 in a uniform fluid

$$h(r_{12}) = c(r_{12}) + \rho\int c(r_{13})c(r_{32})dr_3 + \rho^2\int\int c(r_{13})c(r_{34})c(r_{42})dr_3dr_4 + \ldots, \qquad (2.4.1)$$

where ρ is the density of the liquid and r represents the distance between two particles. $c(r)$ is the so-called direct correlation function and the integral terms in Eq. (2.4.1) represent the indirect correlation functions mediated through other particles 3, 4 and all the rest. By recursively referring to the definition of $h(r)$, Eq. (2.4.1) is rewritten as the famous Ornstein-Zernike (OZ) equation[5,24]

$$h(r_{12}) = c(r_{12}) + \rho\int c(r_{13})h(r_{32})dr_3\,. \qquad (2.4.2)$$

Both functions in Eq. (2.4.2), $h(r)$ and $c(r)$, are unknown, so that another equation is required to get a solution. A general relation which couples $h(r)$ and $c(r)$ with the interaction potential $u(r)$ of fluid particles is the so-called "closure equation"[5,24]

$$h(r) = \exp[-\beta u(r) + h(r) - c(r) + b(r)] - 1 \qquad (2.4.3)$$

where $\beta = 1 / T$ is the inverse of the temperature T and $b(r)$, the so-called "bridge" function, is a functional of $h(r)$. By selecting the simplest approximations of $u(r)$ and $b(r)$, i.e. using the Lennard-Jones (LJ) potential and setting $b(r) = 0$ one recovers the hypernetted-chain (HNC) closure.[5,86] Eqs. (2.4.2) and (2.4.3) indeed can be solved simultaneously. However, this original form is only applicable to atomic fluids. The OZ equation must be extended to obtain the correlation functions of a molecular liquid.

2.4.2. *1D Reference Interaction Site Model*

Chandler and Anderson first generalized the OZ equation to polyatomic molecules by introducing the intramolecular correlation function $\omega(r)$ to represent the mutual orientations of molecules in liquids.[27] To this end, the solvent liquid is considered to consist of atomic or rigid polyatomic species around the solute whose interaction with the solute is approximated by a two-particle interatomic potential (site-site potential). This site-site Ornstein–Zernike (SSOZ) equation, which later became known as 1D reference interaction site model (1D RISM), averages the correlation functions between solute and solvent molecules over orientations at fixed distances r between their interaction sites. For infinitely dilute solutions,

the ansatz can be written in the following matrix form [5,28,29,87]

$$\mathbf{h}^{\mathrm{uv}} = \mathbf{w}^{\mathrm{u}} * \mathbf{c}^{\mathrm{uv}} * \boldsymbol{\chi}^{\mathrm{v}} . \tag{2.4.4}$$

Here $\mathbf{h}^{\mathrm{uv}}$ and $\mathbf{c}^{\mathrm{uv}}$ are (rectangular) matrices of the total correlation functions $h_{i\alpha}^{\mathrm{uv}}(r)$ and direct correlation functions $c_{i\alpha}^{\mathrm{uv}}(r)$,

$$\mathbf{h}^{\mathrm{uv}} = \begin{pmatrix} h_{11}^{\mathrm{uv}} & \cdots & h_{i1}^{\mathrm{uv}} \\ \vdots & \ddots & \vdots \\ h_{1\alpha}^{\mathrm{uv}} & \cdots & h_{i\alpha}^{\mathrm{uv}} \end{pmatrix}, \mathbf{c}^{\mathrm{uv}} = \begin{pmatrix} c_{11}^{\mathrm{uv}} & \cdots & c_{i1}^{\mathrm{uv}} \\ \vdots & \ddots & \vdots \\ c_{1\alpha}^{\mathrm{uv}} & \cdots & c_{i\alpha}^{\mathrm{uv}} \end{pmatrix}. \tag{2.4.5}$$

The superscripts u and v label the solute and the solvent, respectively, while the subscripts *i* and α enumerate solute and solvent interaction sites, respectively. In Eq. (2.4.4) the star $*$ denotes the convolution integral between two matrices **a** and **b**,

$$(\mathbf{a} * \mathbf{b})_{ik}(r) = \sum_j \int a_{ij}(r - r') b_{jk}(r') dr' \tag{2.4.6}$$

with the summation over the corresponding matrix indices. With this convention Eq. (2.4.4) is calculated as

$$h_{i\alpha}^{\mathrm{uv}}(r) = \sum_{j=1}^{N_{\mathrm{u}}} \sum_{\gamma=1}^{N_{\mathrm{v}}} \iint \omega_{ij}^{\mathrm{u}}(r - r') c_{j\gamma}^{\mathrm{uv}}(r' - r'') \chi_{\gamma\alpha}^{\mathrm{v}}(r'') dr' dr'' \tag{2.4.7}$$

by summing over all the solute and solvent sites, $j = 1, \ldots, N_{\mathrm{u}}$ and $\gamma = 1, \ldots, N_{\mathrm{v}}$, respectively. The solvent is represented by its (square symmetric) susceptibility matrix $\chi^{\mathrm{v}} = \omega^{\mathrm{v}} + \rho h^{\mathrm{vv}}$, which is obtained self-consistently from the pure solvent RISM equation[5,87]

$$\mathbf{h}^{\mathrm{vv}} = \mathbf{w}^{\mathrm{v}} * \mathbf{c}^{\mathrm{vv}} * \boldsymbol{\chi}^{\mathrm{v}}, \tag{2.4.8}$$

at a specific solvent number density ρ. Finally, $\mathbf{w}^{\mathrm{u}}$ and $\mathbf{w}^{\mathrm{v}}$ are the (square symmetric) matrices built of intra-molecular correlation functions of solute and solvent, respectively, which are radial δ-functions, best specified by their Fourier representations, e.g. $\tilde{\omega}_{\alpha\gamma}^{\mathrm{v}}(k) = j_0(k\ell_{\alpha\gamma})$. Here, j_0 is the zeroth-order spherical Bessel function,[5] k the wave number, and $\ell_{\alpha\gamma}$ the distance between two interaction sites α and γ of a rigid solvent species.[5,27]

Similar to the original OZ equation in Section 2.4.1, a closure relation is necessary to admit solving Eqs. (2.4.4) and (2.4.8). In the framework of 1D RISM theory, the HNC closure reads[86]

$$h_{i\alpha} = \exp[-\beta u_{i\alpha} + h_{i\alpha} - c_{i\alpha}] - 1 \tag{2.4.9}$$

Here $u_{i\alpha}$ is the pair-wise potential between interaction sites *i* and α which accounts for short-range repulsion and dispersion interactions. It commonly is represented by the sum of the

short-range LJ term $u_{i\alpha}^{\mathrm{LJ}}$ and the long-range Coulomb term $u_{i\alpha}^{\mathrm{C}}$,

$$\begin{aligned} &u_{i\alpha} = u_{i\alpha}^{\mathrm{LJ}} + u_{i\alpha}^{\mathrm{C}}, \\ &u_{i\alpha}^{LJ}(r) = 4\varepsilon_{i\alpha}[(\sigma_{i\alpha}/r)^{12} - (\sigma_{i\alpha}/r)^{6}]; u_{i\alpha}^{C}(r) = q_i q_\alpha / r. \end{aligned} \tag{2.4.10}$$

The parameters of the pair interactions $\varepsilon_{i\alpha}$ and $\sigma_{i\alpha}$ were derived from site parameters using the Lorentz-Berthelot rules,[5] which are $\sigma_{i\alpha} = (\sigma_i + \sigma_\alpha)/2$ and $\varepsilon_{i\alpha} = \sqrt{\varepsilon_i \varepsilon_\alpha}$. q_i and q_α are the charges carried by the interaction sites.

In the original 1D RISM theory, the solution of Eqs. (2.4.4) and (2.4.8) yields far too low dielectric constants of polar liquids according to the expression[30,88]

$$\varepsilon_{\mathrm{RISM}} = 1 + 4\pi\beta\rho\mu^2/3. \tag{2.4.11}$$

Here μ is the dipole moment of the solvent. A popular solution to this problem has been suggested,[89] namely the dielectrically consistent RISM (DRISM) with an additional term introduced in Eq. (2.4.8):[32,88,89]

$$\mathbf{h}' = \mathbf{w}' * \hat{\mathbf{c}} * (\mathbf{w}' + \boldsymbol{\rho}\mathbf{h}') \tag{2.4.12}$$

Eq. (2.4.12) involves $\mathbf{h}' = \mathbf{h} - \boldsymbol{\zeta}$, $\mathbf{w}' = \mathbf{w} + \boldsymbol{\rho}\boldsymbol{\zeta}$ and $\hat{\mathbf{c}} = \mathbf{c} - \mathbf{b}$, whereas $\mathbf{b}$ is the bridge correction term coming into Eq. (2.4.12) from the general closure relation Eq. (2.4.3), and $\boldsymbol{\zeta}$ is the renormalized correction satisfying $\boldsymbol{\zeta} = \mathbf{w} * \mathbf{b} * (\mathbf{w} + \boldsymbol{\rho}\boldsymbol{\zeta})$, which can be obtained in reciprocal space using experimental results for the dielectric constant.[32,88]

2.4.3. *3D Reference Interaction Site Model*

Due to the radial average of correlation functions, the site-site treatment in 1D RISM theory does not represent the full tree-dimensional structure of the solvation.[5,30] An intuitive approach to overcome this issue may consist in taking the molecular orientation into account in the OZ equation, Eq. (2.4.2), which leads to the molecular OZ (MOZ) equation[5,24]

$$h(r_{12}, \Omega_1, \Omega_2) = c(r_{12}, \Omega_1, \Omega_2) + \frac{\rho}{\bar{\Omega}} \int c(r_{13}, \Omega_1, \Omega_3) h(r_{32}, \Omega_3, \Omega_2) dr_3 d\Omega_3, \tag{2.4.13}$$

in terms of the molecular orientation Ω; $\bar{\Omega} = \int d\Omega$ is the normalization factor of the orientational integral for different species. Although several trials have been made in this direction,[90-93] the difficulty of solving the six-dimensional (6D) MOZ equation hampered its broad application.[30] Instead, Beglov and Roux[94] pioneered solving the RISM equation by using a 3D grid space and employing the density functional theory of non-uniform polyatomic liquids.[95] Kovalenko and Hirata derived the 3D RISM integral equations later from the 6D MOZ equation by partial integration over the orientational coordinates.[96] The 3D RISM

equation reads as[5]

$$\mathbf{h}^{uv} = \mathbf{c}^{uv} * \boldsymbol{\chi}^{v}. \tag{2.4.14}$$

Here $\mathbf{h}^{uv}$ and $\mathbf{c}^{uv}$ are matrices of the three-dimensional total correlation functions $h_\alpha^{uv}(\vec{r})$ and the 3D direct correlation functions $c_\alpha^{uv}(\vec{r})$ of solvent site α around a solute of arbitrary shape. The star symbol $*$ has the same meaning as in Eq. (2.4.4). As in the 1D RISM theory in Section 2.4.2, $\chi_{\alpha\gamma}^{v}(r)$ is the bulk solvent susceptibility represented by a (symmetric square) matrix. It can be derived from the radial site-site correlation functions of the pure solvent in 1D RISM equations, Eqs. (2.4.8) and (2.4.9) for a specific temperature and solvent number density ρ.[5,32]

By comparing the 1D and 3D RISM equations, hence Eq. (2.4.4) with Eq. (2.4.14), one notes that the latter does not require the intramolecular correlations to represent the molecular orientation. Instead, the orientation dependence is retained via the partial integration of the total and direct correlation vectors in a 3D uniform space. In analogy to 1D RISM, the closure relation is again required for solving Eq. (2.4.14). The HNC closure in 3D RISM reads[5]

$$h_\alpha = \exp[-\beta u_\alpha + h_\alpha - c_\alpha] - 1. \tag{2.4.15}$$

Here u_α is the interaction potential between solvent site α and the solute. In actual applications, this quantity is estimated as a superposition of the site-site interaction potentials $u_{i\alpha}$, Eq. (2.4.10), between the solute site i and the solvent site α

$$u_\alpha(\vec{r}) = \sum_i u_{i\alpha}(|\vec{r} - \vec{r}_i|). \tag{2.4.16}$$

With Eqs. (2.4.14) and (2.4.15), the spatial correlation functions of a solvent around the solute molecule with an arbitrary shape can be solved.

2.4.4. *Closure Relations*

As discussed in Sections 2.4.1–2.4.3 the closure relation, which couples the total and direct correlations functions with the interaction potential, is necessary for solving the OZ equation as well as 1D and 3D RISM equations. Setting the bridge term b(r) to zero in the general form Eq. (2.4.3) yields the HNC closures Eq. (2.4.9) and Eq. (2.4.15). However, in this case the exponential form in the HNC closure leads to a slow convergence rate or may even cause numerical instabilities.[5] For a numerically more stable closure expression, Kovalenko and Hirata partially linearized the exponent term of the HNC closure.[97] The new closure, which is referred to as partial-linearized HNC (PLHNC) or KH closure, reads as $h_{i\alpha} = f(-\beta u_{i\alpha} + h_{i\alpha} - c_{i\alpha})$ in 1D RISM and $h_\alpha = f(-\beta u_\alpha + h_\alpha - c_a)$ in 3D RISM, with

$$f(x) = \begin{cases} e^x - 1 & x \le 0 \\ x & x > 0 \end{cases}. \tag{2.4.17}$$

The KH closure, Eq. (2.4.17) with the relevant expressions in 1D and 3D RISM, combines the HNC approximation in the solvent density depleted regions where $h < 1$, and the linearization of the entire exponent in the enriched regions where $h > 1$. Being more numerically stable, the KH closure is widely applied in the studies performed with RISM method.[5,30] Recently, Kast and Kloss systematically investigated the convergence problem of the HNC closure with a partial series expansion of order n (PSEn), which yields $f(x)$ in terms of the following expression[98]

$$f(x) = \begin{cases} e^x - 1 & x \le 0 \\ \sum_{k=1}^{n} x^k / k! & x > 0 \end{cases}. \tag{2.4.18}$$

The formulation of PSEn closure of Eq. (2.4.18) provides a good balance between the limiting cases of HNC ($n \to \infty$) and KH ($n = 1$) closures. The partial series expansion for $f(x)$ became thus popular in the recent studies of polar and charged system as the convergence of solving RISM equations can be improved altering n.[30] Other closures which were proposed for early RISM variants, e.g. the Percus-Yevick (PY) approximation and the mean-spherical approximation (MSA), provide different linearization of the exponent term in the HNC closure.[5] However, these latter formulations were found to behave erratically for Coulombic systems and are therefore rarely used in applications.[5]

2.4.5. *Excess Chemical Potential*

The excess chemical potential of a solute at infinite dilution, also known as the "solvation free energy", corresponds to the free energy change associated with a process in which a solute molecule is coupled to the solvent. The coupling process can be expressed in terms of the Kirkwood coupling parameter ξ [5]

$$U(\xi) = U_{vv} + \xi U_{uv} \tag{2.4.19}$$

where U, U_{vv}, and U_{uv} represent the interaction potential of the total system, the solvent-solvent contribution, and the solute-solvent part. The coupling parameter ξ thereby varies from 0 to 1, representing all states, from no solute coupling to full coupling. The free energy change related to the coupling process can be written in terms of the configuration integral of the system which is a functional of the coupling state at a given temperature T [5]

$$\mu = -kT \ln Z(\xi = 1) - [-kT \ln Z(\xi = 0)] \tag{2.4.20}$$

In the context of 1D RISM theory, Eq. (2.4.20) can be expressed in terms of the pair

correlation functions as thermodynamic integration (TI) over ξ [5]

$$\mu = \frac{4\pi\rho}{\beta}\sum_{i=1}^{N_u}\sum_{\alpha=1}^{N_v}\int_0^1 d\xi\int_0^\infty u_{i\alpha}(r)g_{i\alpha}(r,\xi)r^2dr\,. \tag{2.4.21}$$

Here $g_{i\alpha} = h_{i\alpha} + 1$ is the site-site radial distribution function (RDF). The numeric integration in Eq. (2.4.21) requires a numeric solution of the RISM equation for each finite increment of ξ, which can be expensive for large systems. In practice, one of the analytical expressions for the closure relations discussed in Section 2.4.4 is used to calculate the excess chemical potential. When using the HNC closure, the excess chemical potential reads[86]

$$\mu_{1D}^{HNC} = \frac{2\pi\rho}{\beta}\sum_{i=1}^{N_u}\sum_{\alpha=1}^{N_v}\int_0^\infty [|h_{i\alpha}|^2 - 2c_{i\alpha} - c_{i\alpha}h_{i\alpha}]r^2dr\,. \tag{2.4.22}$$

For the KH closure and the generalized form of the PSE*n* closure $h_{i\alpha} = f(-\beta u_{i\alpha} + h_{i\alpha} - c_{i\alpha})$, the excess chemical potential is[98]

$$\mu_{1D}^{PSEn} = \mu_{1D}^{HNC} + \frac{2\pi\rho}{\beta}\sum_{i=1}^{N_u}\sum_{\alpha=1}^{N_v}\int_0^\infty \phi(x_{i\alpha})r^2dr\,, \tag{2.4.23}$$

with $\phi(x) = f(x) - x - \int_0^x f(y)dy$. Eq. (2.4.23) has a particularly simple form with $f(x)$ defined in Eq. (2.4.18): $\phi_n(x) = -x^{n+1}/(n+1)!$ for $x > 0$ and zero otherwise. In the case of the KH (PSE1) closure, the additional term $-x^2/2$ cancels the $h^2/2$ term of the HNC integrand, Eq. (2.4.22), in the regions where $x = h > 0$.

Another popular expression was proposed by Chandler, Singh and Richardson assuming Gaussian fluctuations (GF) of the solvent molecules,[99]

$$\mu_{1D}^{GF} = \mu_{1D}^{HNC} - \frac{2\pi\rho}{\beta}\sum_{i=1}^{N_u}\sum_{\alpha=1}^{N_v}\int_0^\infty |h_{i\alpha}|^2 r^2dr\,. \tag{2.4.24}$$

The GF expression in Eq. (2.4.24) omits contribution of order h^2 in the HNC expression in Eq. (2.4.22) and thus reduces the overestimated energies calculated by the HNC-like closures. However, it was found problematic for polar solutes.[5]

Analogously, the excess chemical potential can be expressed in 3D RISM theory. To this end, the HNC, PSE*n*, and GF functionals read as

$$\mu_{3D}^{HNC} = \frac{\rho}{\beta}\sum_{\alpha=1}^{N_v}\int [|h_\alpha|^2 - 2c_\alpha - c_\alpha h_\alpha]d^3r\,, \tag{2.4.25}$$

$$\mu_{3D}^{PSEn} = \mu_{3D}^{HNC} + \frac{\rho}{\beta}\sum_{\alpha=1}^{N_v}\int \phi(x_\alpha)d^3r\,, \tag{2.4.26}$$

$$\mu_{3D}^{GF} = \mu_{3D}^{HNC} - \frac{\rho}{\beta}\sum_{\alpha=1}^{N_v}\int |h_\alpha|^2 \, d^3r \text{ , respectively.} \tag{2.4.27}$$

Here $\phi(x)$ has the same definition as in 1D RISM in Eq. (2.4.23).

2.4.6. *Cavity Formation Correction*

Due to the explicit treatment of the spatial orientation of the molecular solute, 3D RISM was expected to yield energetic results of higher accuracy than 1D RISM. However, when only applying the various functionals for the closure relations and the excess chemical potential as in Section 2.4.4 and 2.4.5, 3D RISM fails to provide satisfactory results.[5,30] Several studies showed that for organic solutes the calculated solvation free energy can deviate from experiment by an order magnitude.[5,30,62]

Recently, several systematical studies covering a large test set of molecular solutes reported a correlation between the error of the calculated solvation free energy and the partial molar volume (PMV) of the solute.[64,100-102] Followed by both semi-empirical functionals with parameter fitted against the error[64] as well as a first principle theoretical argument,[61,63,64] PMV corrections provide significantly improved solvation free energy results for a wide range of molecular solutes and are becoming increasingly popular.[30]

Within the 3D RISM approach the dimensionless PMV of the solute[5]

$$\rho V_u = \frac{\rho\kappa_T}{\beta}\left(1 - \rho\sum_{\alpha=1}^{N_v}\tilde{c}_\alpha(0)\right), \tag{2.4.28}$$

involves the Fourier representation of the solvent-solute direct correlation functions $\tilde{c}_\alpha(k)$ at $k = 0$. κ_T denotes the compressibility of the pure solvent[5]

$$\kappa_T = \frac{\beta}{\rho\left(1 - \rho\sum_{\alpha}^{N_v}\sum_{\gamma}^{N_v}\tilde{c}_{\alpha\gamma}^{v}(0)\right)}, \tag{2.4.29}$$

with $\tilde{c}_{\alpha\gamma}^{v}(k)$ being the Fourier representation of the direct correlation function of the pure solvent. Based on the strong linear correlation of errors in the solvation free energy and the PMV, Palmer et al. suggested a semi-empirical correction expression[64]

$$\mu^{PMV} = \mu + (a\rho V_u + b), \tag{2.4.30}$$

to compensate for the asymptotic deficiency of the original 3D RISM expression, by fitting the parameters a and b to a data set of 189 simple organic molecules. Following this fitting strategy, several research works reported refinements and further extensions of the semi-empirical functional.[30,100,101,103] Yet the physical meaning of the correction coefficients

remained unclear.

Sergiievskyi et al.[63] proposed a first conceptual interpretation of the correction coefficients in an isothermal-isobaric ensemble by molecular density functional theory (MDFT) which they further extended to 3D RISM. They derived the correction coefficients as $\beta\alpha = (\rho/2)\sum_{\alpha\gamma}\tilde{c}^{\mathrm{v}}_{\alpha\gamma}(0)$ and b = 0. However, their thermodynamic arguments for the derivation of this particular PMV correction form were questioned later[104] and the recent work of Minsin et al.[61] favors a different correction coefficient, $\beta\alpha = (\rho/2)\sum_{\alpha\gamma}\tilde{c}^{\mathrm{v}}_{\alpha\gamma}(0) - 1$. Both correction functionals have been justified for a data set of small organic molecules. As will be presented in Section 4.3, those two choices of correction coefficient can be interpreted in an asymptotic analysis of the excess chemical potential functional.

3. Implementation of a Hybrid Approach of RISM and QM

3.1. Basics of the Hybrid Method

3.1.1. *Hybrid Method of RISM with Hartree-Fock Theory*

With the wide range of applications and the rapid development of RISM theory, the combination of RISM with QM is regarded as a promising alternative to the QM/MM method and the PCM method for studying the electronic structure of molecular species in the presence of a solvent. The hybrid approach of 1D RISM together with a self-consistent field (SCF) QM solute description was first proposed by Ten-no et al. at the level of Hartree-Fock (HF) theory.[44] The implementation of this hybrid approach combining HF with RISM is not covered in this thesis, but an overview of fundamental aspects is briefly presented in the following.

The combination of HF theory with RISM can be derived by considering the reaction field of the solvent molecules. The variational equation subject to the orthonormalization constraint reads as[44]

$$\sum_i \langle \delta\psi_i \,|\, F_i - f_i \sum_\lambda V_\lambda b_\lambda \,|\, \psi_i \rangle = 0 \tag{3.1.1}$$

where F_i is the Fock operator of an isolated molecule and the term $F_i - f_i \sum_\lambda V_\lambda b_\lambda$ denotes the Fock operator F_i^{solv} of the solvated system. f_i are the occupation numbers of the orbitals $\{\psi_i\}$, b_λ is the population operator for a solute atomic site λ, and V_λ is the electrostatic potential induced by the solvent medium on the solute site

$$V_\lambda = \rho \sum_\alpha \int 4\pi r^2 \frac{q_\alpha}{r} h_{\lambda\alpha}(r) dr \,. \tag{3.1.2}$$

Here q_α represents the partial charge of a site α of the solvent molecule with number density ρ and $h_{\lambda\alpha}(r)$ is the total correlation function which results from RISM theory as described in Section 2.4.2. The RISM equations (2.4.4-2.4.9) are solved with the appropriate partial charge q_λ assigned to the solute site λ to calculate the site-site Coulomb interaction $u^C_{\lambda\alpha}(r) = q_\lambda q_\alpha / r$. The charge q_λ can be obtained by summing over all the orbitals $\{\psi_i\}$ and including the contribution due to nuclei

$$q_\lambda = q_\lambda^{(N)} - \sum_i f_i \langle \psi_i | b_\lambda | \psi_i \rangle \,, \tag{3.1.3}$$

where $q_\lambda^{(N)}$ is the charge contributed by nucleus *N*. The correlation functions are updated in

the RISM equations, 2.4.4 to 2.4.9, with partial charges of solute sites determined from Eq. (3.1.3). Then the new Fock operator can be obtained via Eq. (3.1.1). One repeats this procedure until convergence, so that the electronic structure of a solute and the surrounding solvent environment are determined in a self-consistent manner.

This first generation of the hybrid method combining RISM with HF theory was developed further. The analytical expression of the energy gradient can be derived with respect to the nuclear coordinates.[56] It was suggested to treat explicitly the spatial electron density distribution (SEDD) when calculating the electrostatic interaction between solute and solvent instead of using the fitted atomic charges of solute sites.[105] Discussing these aspects in detail would lead away from the main topic of this thesis which aims at combining the RISM method with electronic density functional theory; the details of that approach can be found in the original papers[56,105] as well as recent reviews.[6,30]

3.1.2. *Hybrid Method of 3D RISM with Density Functional Theory*

The Kohn-Sham density functional theory (KS-DFT) can be considered as an improvement of the HF theory, since the many-body effect of electron correlation is approximated by a functional of the electron density.[3] As the KS-DFT approach requires the spatial density profiles but not the radial site-site distribution functions, the 3D RISM method which yields the 3D correlation profiles of solvents is more suited for combining with the KS-DFT approach in a self-consistent description of the electronic structure of solvated molecules.[45,48] In the following, the fundamental aspects of combining 3D RISM with KS-DFT will be presented.

The KS equation of the KS-DFT approach, is used to determine the spin-polarized KS molecular orbitals $\{\psi_a^\sigma(\mathbf{r})\}$, represented in terms of the spatial electronic coordinates $\mathbf{r}$ [45,48]

$$[\hat{h}_a^{KS,\sigma} + \phi_{\mathrm{v}}(\mathbf{r}_a)]\psi_a^\sigma(\mathbf{r}) = \varepsilon_a^\sigma \psi_a^\sigma(\mathbf{r}) \tag{3.1.4}$$

where $\hat{h}_a^{KS,\sigma}$ is the single-particle KS Hamiltonian of the solute molecule, $\{\varepsilon_a^\sigma\}$ denote the eigenvalues of the KS orbitals, and $\phi_{\mathrm{v}}(\mathbf{r})$ is the solvent electrostatic potential acting on the KS orbitals. Solving the KS-DFT equation Eq. (3.1.4) yields the solute charge density $n_{\mathrm{u}}(\mathbf{r})$ which includes both the contribution from electrons and nuclei.

At the level of 3D RISM theory, the interaction energy between the QM solute and the solvent molecule needs to be calculated as in Eq. (2.4.16). It consists of two parts, the short-range interaction modeled by the LJ potential $\underline{u}^{\mathrm{LJ}}$ between the solute nuclei and the solvent site as the expression in Eq. (2.4.10), and the long-range electrostatic interaction u^{C} evaluated by solute potential $\phi_{\mathrm{u}}(\mathbf{r})$ acting on the solvent charge density $n_{\mathrm{v}}(\mathbf{r}) = \rho\sum_\alpha q_\alpha g_\alpha(\mathbf{r})$. Here

$\phi_u(\mathbf{r})$ can be derived from the charge density of the QM solute $n_u(\mathbf{r})$ obtained from KS-DFT. To this end, the general form of the electrostatic interaction between the QM solute and the solvent molecules can be written in terms of the charge densities n_u and n_v

$$u^C = \int n_u(\mathbf{r}) f^C(|\mathbf{r}-\mathbf{r}'|) n_v(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \tag{3.1.5}$$

where $f^C = 1/r$ is the Coulomb kernel. From Eq. (3.1.5), the solute and solvent electrostatic potentials ϕ_u and ϕ_v can be defined as:

$$\begin{aligned} \phi_u &= f^C * n_u \\ \phi_v &= f^C * n_v \end{aligned} \tag{3.1.6}$$

Eq. (3.1.5) connects the charge density n_u of the QM solute obtained from KS-DFT and the solvent charge density n_v evaluated by 3D RISM, so that their determination can be carried out by combining KS-DFT with 3D RISM in a self-consistent manner.

Figure 3.1.1 illustrates the self-consistency cycle of the hybrid method of 3D RISM with KS-DFT. Starting with the solutions of the KS equation of the molecular solute in the gas phase without the solvent potential term in Eq. (3.1.4), the initial solute charge density n_u is determined and transferred to the 3D RISM solver. The converged 3D RISM calculation yields the solvent distribution g as well as the solvent potential ϕ_v, Eq. (3.1.6). With the solvent potential ϕ_v present, the KS equation, Eq. (3.1.4), is solved for the electronic structure of the solute as well as the charge density n_u of the solvated solute. Repeating this process until the DFT calculation converges to yield the solute electronic structure, perturbed by the solvent medium, and the solvent structure in the presence of a QM solute.

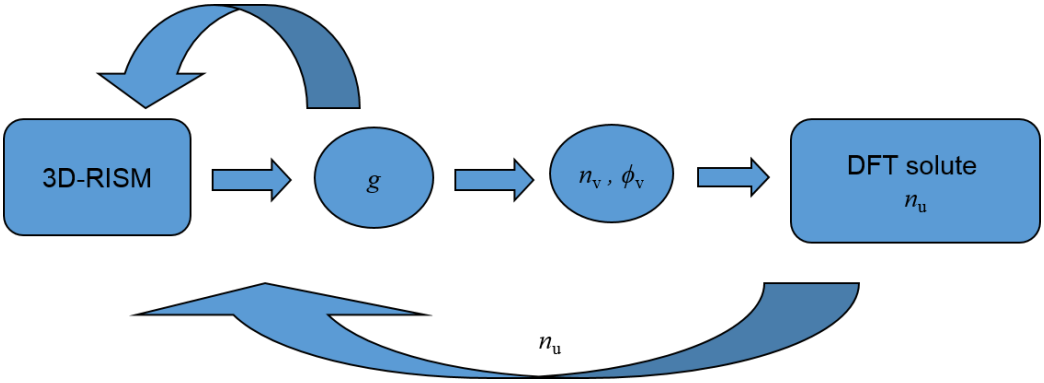


Figure 3.1.1: Flow chart of the self-consistency cycle of the hybrid 3D RISM DFT method.

3.2. Interfacing a 3D RISM Solvent and a DFT Solute

The self-consistent combination of the 3D RISM with KS-DFT involves a two-phase implementation that implies a) updating the solvent distribution in the electrostatic field of the solute charge density and b) a KS-DFT calculation in the mean field of the solvent distribution that determines the electronic structure and the electron charge distribution of the solute. This section describes details of the implementation of these two aspects.

3.2.1. *Numerical Solution of the 3D RISM Equations*

The non-linear equations of 3D RISM requires an iterative solution. The solving procedure of 3D RISM with a QM description for the solute is conducted analogously to the case of a MM representation.[45,48] The QM solute affects the solvent medium via an electrostatic potential acting on the charged solvent sites (Section 3.1.2). The following discussion therefore briefly introduces the numerical solution of the 3D RISM equation of which the original implementation was contributed by A.V. Matveev based on the earlier studies and the refinements were conducted in the context of this thesis.[65,106] The modification to the solvent-solute interaction due to the presence of a QM solute will then be presented in the subsequent section.

The elements of the correction function matrices in the non-linear equations (2.4.14-2.4.15) are central for this. For the sake of brevity the subscripts are omitted and one sets $\beta = 1$ in the closure relation. Then the non-linear system of equations reads

$$\begin{cases} h = \chi * c \\ c = f(-u + h - c). \end{cases} \tag{3.2.1}$$

The solute-solvent interaction potential u and the solvent susceptibility χ are calculated externally and thus act as "constants" in Eq. (3.2.1). As the two unknown sets of variables h and c appear both in the non-linear functional f of the closure relation in Eq. (3.2.1), it is numerically advantageous to operate with a single set of unknowns defined as

$$t = h - c \tag{3.2.2}$$

t is called the indirect correlation function.[107] The closure relation in Eq. (3.2.1) is then reformulated as a functional $C(t)$

$$C(t) = c = f(-u + t) - t \tag{3.2.3}$$

With Eq. (3.2.3) the non-linear equations system of Eq. (3.2.1) is transformed into a non-linear functional $T(t)$

$$T(t) = t = (\chi - 1) * C(t) \tag{3.2.4}$$

Solving Eq. (3.2.4) yields *t* and, together with Eqs. (3.2.2-3.2.3), *h* and *c* can be obtained. A naive solution of Eq. (3.2.4) can be achieved by searching for a fix point of the simple iterative scheme $t_{n+1} = T(t_n)$. However, this is not the most efficient approach as it tends to diverge for the highly non-linear closure relations defined in Eq. (2.4.17). Instead, a better choice is to solve the reduced non-linear problem

$$F(t) = T(t) - t = 0 \tag{3.2.5}$$

using a Newton-Krylov non-linear solver included in the *Portable, Extensible Toolkit for Scientific Computation* (PETSc) library.[108] With that solver the solution of Eq. (3.2.5) is achieved by iteratively searching *t* via

$$\begin{aligned} &J(t_n)\Delta t_n = -F(t_n); \\ &t_{n+1} = t_n + \Delta t_n, \end{aligned} \tag{3.2.6}$$

where $J(t)$ denotes the corresponding Jacobian of $F(t)$. Also note that in practice the indirect correlation function *t* is represented as a vector on a 3D uniform grid so that it is enormously expensive to compute and store all elements of the Jacobian matrix. It is therefore advantageous to provide the Jacobian as an analytical matrix-free linear operator $J = \delta F / \delta t$ defined by its action on an arbitrary vector δt

$$\begin{aligned} J * \delta t &= \delta F \\ &= \delta[T(t) - t] \\ &= \delta[(\chi - 1) * C(t) - t] \\ &= (\chi - 1) * \delta C - \delta t \end{aligned}. \tag{3.2.7}$$

Indeed, the expression in Eq. (3.2.7) containing the response $\delta C = (\delta C / \delta t) * \delta t$ as an intermediate provides a more intuitive and straightforward way for computing the operator result $J * \delta t$. The convolution operations in Eqs. (3.2.4) and (3.2.7) are carried out using the fast Fourier transformation[109] according to the convolution theorem. To this end, the calculation of $T(t)$ in Eq. (3.2.4) is transformed as

$$\begin{aligned} &\mathcal{F}_3[T(t)] = \mathcal{F}_3[(\chi - 1) * C(t)] = \mathcal{F}_3(\chi - 1)\mathcal{F}_3[C(t)]; \\ &T(t) = \mathcal{F}_3^{-1}\{\mathcal{F}_3[T(t)]\}, \end{aligned} \tag{3.2.8}$$

where $\mathcal{F}_3$ denotes the three-dimensional Fourier transformation operator and $\mathcal{F}_3^{-1}$ its inverse.

The solvent susceptibility χ, one of the two "constants" χ and *u* in the non-linear equations system, Eq. (3.2.1), can be obtained by solving the 1D RISM equations (2.4.8-2.4.9) with a similar approach as described above. The interaction *u* between solute and solvent defined by the pair-wise potential as in Eq. (2.4.10) requires a special treatment due to the long-range nature of the electrostatic potential. According the scheme proposed by Ng,[110] the long-range

asymptote u_{L} of the interaction potential is separated from the indirect correlation *t* as well as $-u_{\mathrm{L}}$ from the direct correlation *c*. The operator *T*(*t*) in Eq. (3.2.4) is then adapted to act primarily with the short-range counterparts $u_{\mathrm{S}}, c_{\mathrm{S}}$, and t_{S}

$$\begin{aligned} T_{\mathrm{S}}(t_{\mathrm{S}}) &= (\chi - 1) * (c_{\mathrm{S}} - u_{\mathrm{L}}) - u_{\mathrm{L}} \\ &= (\chi - 1) * \mathrm{c}_{\mathrm{S}} - \tau \end{aligned}, \tag{3.2.9}$$

with $c_{\mathrm{S}} = C_{\mathrm{S}}(t_{\mathrm{S}}) = f(-u_{\mathrm{S}} + t_{\mathrm{S}}) - t_{\mathrm{S}}$ and the constant "renormalization" term $\tau = \chi * u_{\mathrm{L}}$. The corresponding non-linear functional in Eq. (3.2.5) is transformed to

$$\begin{aligned} F_S(t_{\mathrm{S}}) &= T_{\mathrm{S}}(t_{\mathrm{S}}) - t_{\mathrm{S}} \\ &= [(\chi - 1) * c_{\mathrm{S}} - t_{\mathrm{S}}] - \tau \end{aligned} \tag{3.2.10}$$

Eqs. (3.2.10) and (3.2.11) have essentially the same functional form as Eqs. (3.2.4) and (3.2.5), albeit with a "constant" addition $-\tau$ in the non-linear functional. While this constant term does not affect the form of Jacobian, it depends on the form of u_{L}, which also determines the direct correlation $c_{\mathrm{L}} = -\beta u_{\mathrm{L}}$ and the indirect correlation $t_{\mathrm{L}} = \beta u_{\mathrm{L}}$ at long distances. The asymptotic form of u_{L} must therefore be chosen sufficiently smooth to be represented by a truncated Fourier expansion with known coefficients, thus avoiding a numerically problematic fast Fourier transforms of these terms.

In the atomic charge representation of the solute, the long-range behavior of the Coulomb interaction u^{C} in Eq. (2.4.10) is separated into two components with the Ewald summation scheme[111]

$$u^{\mathrm{C}}(r) = u_{\mathrm{S}}^{\mathrm{C}}(r) + u_{\mathrm{L}}^{\mathrm{C}}(r) = q_i q_\alpha a f_{\mathrm{S}}^{\mathrm{C}}(ar) + q_i q_\alpha a f_{\mathrm{L}}^{\mathrm{C}}(ar), \tag{3.2.11}$$

using the Ewald range parameter a = 1.2 Å^{-1}. The functions $f_{\mathrm{S}}^{\mathrm{C}}$ and $f_{\mathrm{L}}^{\mathrm{C}}$ are defined via the error function

$$\begin{aligned} f_{\mathrm{L}}^{\mathrm{C}}(r) &= \frac{\mathrm{erf}(r)}{r} \\ f_{\mathrm{S}}^{\mathrm{C}}(r) &= \frac{1 - \mathrm{erf}(r)}{r} = \frac{\mathrm{erfc}(r)}{r} \end{aligned} \tag{3.2.12}$$

The potential kernel $f_{\mathrm{S}}^{\mathrm{C}}$ is singular short-range and $f_{\mathrm{L}}^{\mathrm{C}}$ is regular long-range, so that the separation of the long-range asymptote u_{L} from the interaction u can be written as

$$u = u_{\mathrm{S}} + u_{\mathrm{L}} = (u^{\mathrm{LJ}} + u_{\mathrm{S}}^{\mathrm{C}}) + u_{\mathrm{L}}^{\mathrm{C}} \tag{3.2.13}$$

3.2.2. *Solution of the 3D RISM Equations in the Presence of a QM Solute*

The Coulomb kernel f^{C} in Eq. (3.1.5) needs to be adapted due to the presence of the QM

solute. Compared to the site atomic charge representation of molecular solute in MM/3D RISM, the solute charge density n_{u} of a QM solute is the sum of the distributed electron density n_{e} and a localized core density n_{n}

$$n_{\mathrm{u}}(r) = n_{\mathrm{e}}(r) + n_{\mathrm{n}}(r)\,. \tag{3.2.14}$$

Here n_{n} represents the atomic nuclei or effective core charges and can be approximated by Dirac-δ functions

$$n_{\mathrm{n}}(r) = \sum_i z_i \delta(r - r_i) \tag{3.2.15}$$

Here z_i is the positive (effective) core charge of atomic nucleus *i*. In practice, when evaluating the Coulomb potential, a more stable procedure results by "broadening" the core charges in the form of a Gaussian kernel *p* of scale parameter η

$$p(r) = \left(\frac{\eta}{\sqrt{\pi}}\right)^3 \exp(-\eta^2 r^2) \tag{3.2.16}$$

This yields the broadened core density n'_{n}

$$\begin{aligned} n'_{\mathrm{n}}(r) &= p(r) * n_{\mathrm{n}}(r) \\ &= \sum_i z_i p(r) * \delta(r - r_i) \\ &= \sum_i z_i p(r - r_i). \end{aligned} \tag{3.2.17}$$

Then the solute electrostatic potential ϕ_{u}, Eq. (3.1.6) can be expressed as

$$\phi_{\mathrm{u}} = f^{\mathrm{C}} * n_{\mathrm{u}}(r) = f^{\mathrm{C}} * \{n_{\mathrm{e}}(r) + n'_{\mathrm{n}}(r) + [n_{\mathrm{n}}(r) - n'_{\mathrm{n}}(r)]\}, \tag{3.2.18}$$

where the effects caused by broadening the singular core charges needs to be evaluated with $f^{\mathrm{C}} * (n_{\mathrm{n}} - n'_{\mathrm{n}})$. The contributions of the electrons $f^{\mathrm{C}} * n_{\mathrm{e}}$ and the broadened core charges $f^{\mathrm{C}} * n'_{\mathrm{n}}$ are both long-range. With the latter term

$$\begin{aligned} f^{\mathrm{C}} * n'_{\mathrm{n}} &= f^{\mathrm{C}} * p * n_{\mathrm{n}} \\ &= \frac{\mathrm{erf}\,(\eta r)}{r} * n_{\mathrm{n}} = f^{\mathrm{C}}_{\mathrm{L}} * n_{\mathrm{n}}, \end{aligned} \tag{3.2.19}$$

the contribution of $f^{\mathrm{C}} * (n_{\mathrm{n}} - n'_{\mathrm{n}})$ yields a short-range expression

$$\begin{aligned} f^{\mathrm{C}} * (n_{\mathrm{n}} - n'_{\mathrm{n}}) &= f^{\mathrm{C}} * n_{\mathrm{n}} - f^{\mathrm{C}} * n'_{\mathrm{n}} = f^{\mathrm{C}} * n_{\mathrm{n}} - f^{\mathrm{C}}_{\mathrm{L}} * n_{\mathrm{n}} \\ &= (f^{\mathrm{C}} - f^{\mathrm{C}}_{\mathrm{L}}) * n_{\mathrm{n}} = f^{\mathrm{C}}_{\mathrm{S}} * n_{\mathrm{n}} \\ &= \sum_\iota z_i f^{\mathrm{C}}_{\mathrm{S}}(r - r_i) = \sum_i z_i \frac{\mathrm{erfc}[\eta(r - r_i)]}{r - r_i}. \end{aligned} \tag{3.2.20}$$

When inserting Eqs. (3.2.19) and (3.2.20) into Eq. (3.2.18), the solute electrostatic potential ϕ_{u} can be written as

$$\phi_{\mathrm{u}} = f^{\mathrm{C}} * n_{\mathrm{e}} + f_{\mathrm{L}}^{\mathrm{C}} * n_{\mathrm{n}} + f_{\mathrm{S}}^{\mathrm{C}} * n_{\mathrm{n}} . \tag{3.2.21}$$

In this way, the separation of long-range asymptote u_{L} required to solve the 3D RISM equation as described in Section 3.2.1 also applies in the presence of a QM solute. u_{L} is calculated by applying the long-range part of ϕ_{u}, which is $f^{\mathrm{C}} * n_{\mathrm{e}} + f_{\mathrm{L}}^{\mathrm{C}} * n_{\mathrm{n}}$, with the solvent charge density $n_{\mathrm{v}} = \rho \sum_{\alpha} q_{\alpha} g_{g}$. The short-range part u_{S} is computed in similar fashion as in Eq. (3.2.13), hence by evaluating the explicit expression at each point with the short-range potential $f_{\mathrm{S}}^{\mathrm{C}} * n_{\mathrm{n}}$ defined in Eq. (3.2.20) and the LJ potential. Here the LJ field takes the coordinates of the solute nuclei as the superposition of solute site in Eq. (2.4.10).

In this way, the 3D RISM equations can be solved in the presence of a QM description of the solute following strategies discussed in Section 3.2.1. The corresponding solvent electrostatic potential ϕ_{v} can then be obtained and transferred to the DFT program ParaGauss, in which it is included in the Hamiltonian when solving the KS equation, Eq. (3.1.4). In the implementation of this thesis, the solvent electrostatic potential ϕ_{v} and that of the solute ϕ_{u} are calculated by the Poisson equation using fast Fourier transform techniques with periodic boundary conditions; for details, see Section 3.4.1.

Figure 3.2.1 summarizes the procedure of solving the 3D RISM equations with the effective core charges and electron density representation of the solute, as described in Sections 3.2.1-3.2.2. Each step depicted in Fig. 3.2.1 represents the name of the relevant implementation module. The short-range potential $\boldsymbol{u}^{\mathrm{LJ}}$, $\boldsymbol{u}_{\mathrm{S}}^{\mathrm{C}}$ and the long-range part $\boldsymbol{u}_{\mathrm{L}}^{\mathrm{C}}$ are evaluated by the module `solute_field`, the solvent susceptibility χ is pre-processed by the module `solvent_kernel`. Then the non-linear functional $T(t)$ is constructed within the module `iterator_T` and passed to the non-linear equation solver `snes_solver` for solving Eq. (3.2.6). Once the solution of the non-linear equations is converged, the solvent electrostatic potential ϕ_{v} is evaluated, ready to be passed on to ParaGauss.

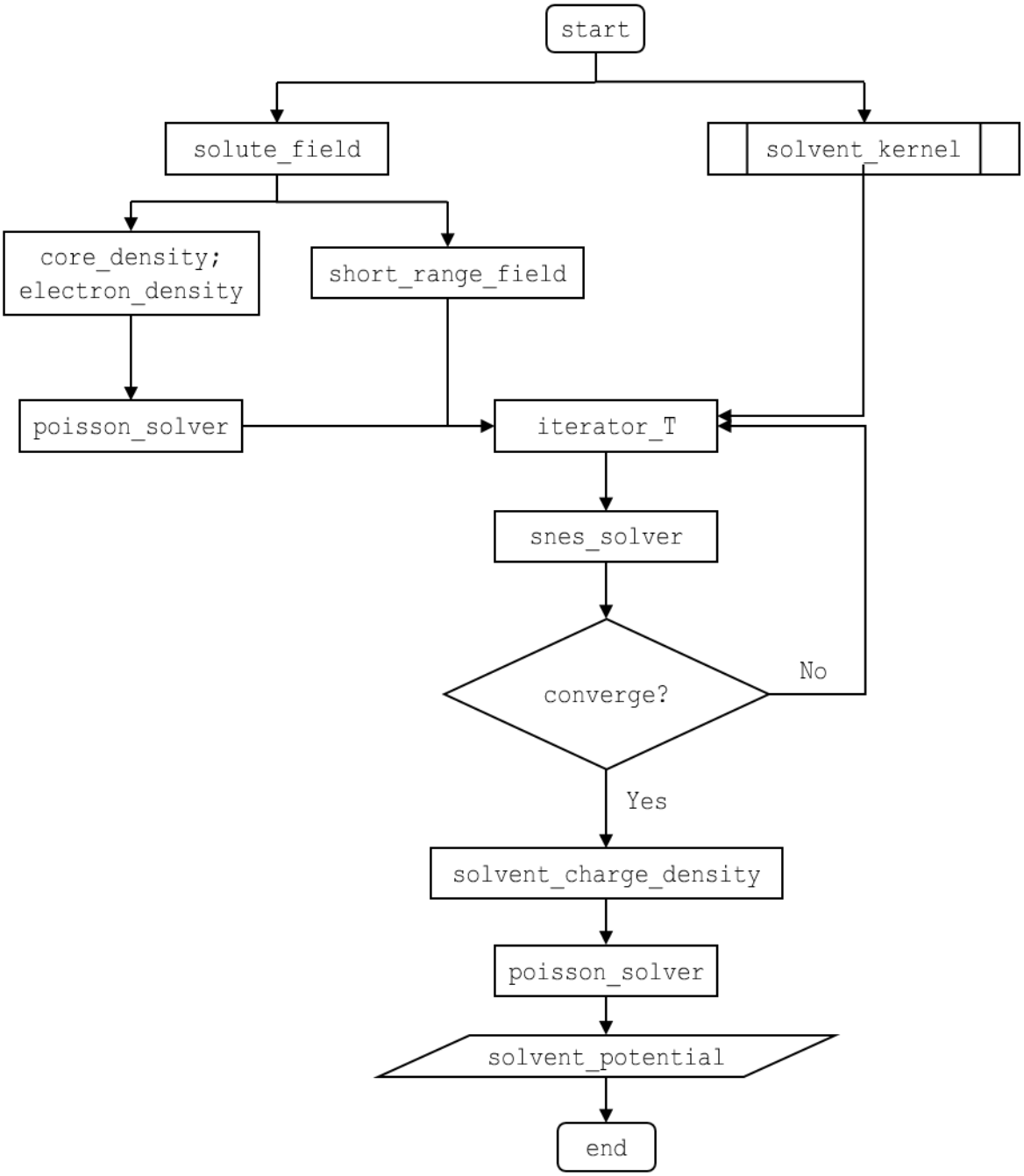


Figure 3.2.1. The flow chart of solving 3D RISM with a QM solute. The names of the various implementation modules are also shown.

3.2.3. *Implementation in ParaGauss with 3D RISM Solvent*

The non-linear equation solver in 3D RISM uses vector data types which are of the form `Vec[1:Ni][1:Nj][1:Nk]` and distributed over the available computing nodes in parallel program runs to represent the 3D vector on a uniform Cartesian grid. Here N_i, N_j and N_k are the numbers of the grid points in each dimension labelled by the indices `i`, `j` and `k`. To pass the value of the electron density n_e from ParaGauss to the 3D RISM solver in a convenient

fashion, and to return the value of the solvent potential filed ϕ_v from the 3D RISM solver to ParaGauss, a grid mapping scheme is used for transforming the coordinate vector in 3D RISM to a two-dimensional array in ParaGauss and vice versa

$$r_{RISM}[1:N_i][1:N_j][1:N_k] \longleftrightarrow r_{PG}[1:3,1:N_{ijk}] \tag{3.2.22}$$

where $N_{ijk}=N_i \times N_j \times N_k$ is the total number of grid points. The operation with the grid points is performed in serial batches. Each batch has the size no larger than the memory cache pre-defined in ParaGauss, thus the additional memory consumption for direct storing the large-size 3D RISM vectors is avoided when running ParaGauss.

With the transformed coordinates, the electron density of the solute molecule can be calculated by the following summation over the KS molecular orbitals $\{\psi_a^\sigma(\mathbf{r})\}$ [4,112]

$$n_e(\mathbf{r}) = \sum_\sigma \sum_a \left|\psi_a^\sigma(\mathbf{r})\right|^2 \tag{3.2.23}$$

The procedure for calculating the electron density is implemented as the function `density` in the RISM solvent module of ParaGauss, and the relevant pseudo code is presented in Scheme 3.2.1.

```
do while (get_gridpoints)
   do i = 1, nportion
      gridpoints unit conversion
   end do
   call orbital_calculate
   call density_calculate
   sum the total density
   move to next batch
end do
```

Scheme 3.2.1. Pseudo code for calculating the electron density.

As for the solvent potential ϕ_v transferred from the 3D RISM solver to ParaGauss, the contribution to the KS Hamiltonian is evaluated in the latter code by the integral over the MOs. Its matrix form read as[3,4]

$$\mathbf{V}_{solv} = \sum_{a\sigma} < \psi_a^\sigma \left|\phi_v\right| \psi_a^\sigma > \tag{3.2.24}$$

The integration is implemented in the subroutine `integrate`, and its main loop structure is shown in Scheme 3.2.2. With the solvent contribution included, the SCF calculation is carried

out to obtain an updated electronic structure of the solute system. Thus, the new electron density n_e is updated and transferred to the 3D RISM solver.

```
do while (get r and V from batch)
  do i = 1, n
    r(i, :) and V(i) units conversion
  end do
  call orbital_calculate
  do i = 1, n_irrep ! loop over irreps
    do j = 1, partners(i) ! loop partners
      do k = 1, dims(i)
        work array: |V|j>
      end do
      call dgemm to calculate <i|V|j>
    end do ! end loop over partners
  end do ! end loop over irreps
end do ! end loop to get values
```

Scheme 3.2.2. Pseudo code for calculating the integral of the solvent potential over the MOs.

3.3. Free Energy Functional

3.3.1. *A Simple Linear Model*

Before introducing the functionals which are used when calculating the free energy with the hybrid QM+RISM approach, this section analyzes the solvation energy contribution to the total free energy with a simple linear model originally suggested by A.V. Matveev.[65] Based on the following discussions, approximative expressions for the free energy functional will be derived in the next section.

For simplicity, assume a model that is characterized by a single degree of freedom *d*, e.g. the dipole moment. Minimization of the energy expression, for $a > 0$,

$$e[d] = a(d - d_0)^2 / 2 \tag{3.3.1}$$

leads to a "permanent" dipole moment d_0. The response to an additional electric field, $\delta e = -fd$, results in $d - d_0 = a^{-1} f$ and suggests identifying a^{-1} as the "polarizability" of the system. The simple expression for the solvation energy of a rigid dipole,

$$\mu[d] = -b(d - d_*)^2 / 2, \tag{3.3.2}$$

is sufficiently flexible for parametrizing the electric cavity reaction field $f_*[d] = -\partial\mu / \partial d$ to have a linear form with some generally non-vanishing intercept $f_*[0] = -bd_*$. Embedding a frozen system (with d_0 in the "gas phase") into a solvent leads to the solvation energy

$$\mu[d_0] = -b(d_0 - d_*)^2 / 2. \tag{3.3.3}$$

Minimizing the sum $G[d] = e[d] + \mu[d]$ results in to a new minimum at

$$d_1 = \frac{ad_0 - bd_*}{a - b} \tag{3.3.4}$$

assuming $0 < b < a$. Finally, one obtains the energy

$$e[d_1] = \frac{a}{2}\left[\frac{b}{a-b}\right]^2 (d_0 - d_*)^2, \tag{3.3.5}$$

the "solvation" energy

$$\mu[d_1] = -\frac{b}{2}\left[\frac{a}{a-b}\right]^2 (d_0 - d_*)^2, \tag{3.3.6}$$

and the total free energy

$$G[d_1] = -\frac{ab}{2(a-b)}(d_0 - d_*)^2 \tag{3.3.7}$$

Assuming $0 < b < a$, the ratio of the changes in the energy defined in Eq. (3.3.5) and the "solvation" energy defined in Eq. (3.3.6)

$$\frac{\mu[d_1]-\mu[d_0]}{e[d_1]-e[d_0]} = -(2-b/a) \tag{3.3.8}$$

has a value between −2 and −1, independent of the system parameters d_0 and d_*. Naturally only the ratio of b / a, defining the scale of the two contributions, matters for this result. The relaxation effect of the free energy,

$$G[d_1]-G[d_0] = -\frac{b^2}{2(a-b)}(d_0-d_*)^2 \tag{3.3.9}$$

is $\mathcal{O}(b^2)$, compared to $G[d_0]$ which is $\mathcal{O}(b)$.

Previously, Sato et al.[113] demonstrated that the ratio in Eq. (3.3.8) is equal to –2 in the linear-response electrostatic limit. This result is confirmed by the present linear model in the limit $b \to 0$. This limit has to be carried out with care, e.g. by setting $b = \varepsilon$ and $d_* = -f_*/\varepsilon$ with $\varepsilon \to 0$, so that $bd_* = -f_* = const$. By discarding from $\mu[d]$ and $G[d]$ divergent terms independent of d, one obtains

$$G[d_1]-G[d_0] \to \frac{f_*^2}{2a} \tag{3.3.10}$$

Thus, the relaxation effect of the free energy is constant when the ratio, Eq. (3.3.8), of the two contributions to this form of the relaxation energy is exactly –2.

3.3.2. *Expressions for the Free Energy Functional*

In electronic DFT, the ground state of the electronic structure of a solute in the gas phase is determined by minimizing the KS energy functional with respect to the KS orbitals in agreement with the Hohenberg-Kohn theorem,[114] which states that the energy functional is uniquely determined by the ground state electron density.[2-4] In the following, for brevity, the energy is expressed as a functional $E[n_u]$ of the solute charge density n_u, whereas $\boldsymbol{n}_u^{(o)}$ denotes the solute charge density in the ground state. The free energy expression in the self-consistent field (SCF) QM+RISM model is assumed to be obtained from the free energy functional

$$G_{SCF} = G[n_u] = E[n_u] + \mu[n_u], \tag{3.3.11}$$

with the minimization of the self-consistent solution given for n_u. This self-consistent model will denoted as "SCF". With the unperturbed charge density $n_u^{(0)}$ for the solute in the gas-phase, Eq. (3.3.11) leads to the first-order perturbation theory (PT1) expression

$$G_{PT1} = G[n_u^{(0)}] = E[n_u^{(0)}] + \mu[n_u^{(0)}], \tag{3.3.12}$$

It will be denoted as "PT1" model. With the force-field treatment, the charge distribution of the solute sites evaluates the ground state charge density $n_u^{(0)}$ with a coarse approximation $\bar{n}_u^{(0)}$. Accordingly, RISM calculations of the excess chemical potential with a discrete force-field will give the approximation $\mu[\bar{n}_u^{(0)}]$ for the solvation free energy of the solute in the solvent medium. The corresponding free energy

$$G_{MM} = E[n_u^{(0)}] + \mu[\bar{n}_u^{(0)}], \tag{3.3.13}$$

will be referred as "MM" model. In the following, it will turn out to be more convenient to shift the free energy values by the energy $-E[n_u^{(0)}]$ of the solute in the gas phase. The corresponding deviations $\Delta G = G - E[n_u^{(0)}]$ of Eqs. (3.3.11-3.3.13) then are

$$\Delta G_{MM} = \mu[\bar{n}_u^{(0)}], \tag{3.3.14}$$

$$\Delta G_{PT1} = \mu[n_u^{(0)}], \tag{3.3.15}$$

and

$$\Delta G_{SCF} = \left(E[n_u] - E[n_u^{(0)}]\right) + \mu[n_u] \tag{3.3.16}$$

By construction, the solvent reaction field forms an electrostatic potential $\delta\mu / \delta n_u = \phi_u$ to be added to the effective Kohn-Sham potential. The corresponding electric field leads to a redistribution of the electron charge density $\Delta n_u^{SCF} = n_u - n_u^{(0)}$. The difference $\Delta E_{SCF} = E[n_u] - E[n_u^{(0)}]$ in Eq. (3.3.16) thereby arises due to the polarization of the solute by the solvent and is always positive for any given n_u, because n_u represents the charge density of the solute electronic structure distorted from the minimization in gas-phase.

In contrast, one always has $G[n_u] \le G[n_u^{(0)}]$ so that the SCF relaxation of the excess chemical potential $\Delta\mu_{SCF} = \mu[n_u] - \mu[n_u^{(0)}]$ is necessarily a non-positive value with a magnitude larger or equal to that of the polarization energy: $\Delta\mu_{SCF} \le -\Delta E_{SCF}$. From the simplified linear model presented in Section 3.3.1, the upper and lower bounds for the ratio of the two relaxation terms for the energy functionals E and μ can be predicted as $-2 \le \Delta\mu_{SCF} / \Delta E_{SCF} \le -1$. The lower bound is realized with the limit of a linear μ corresponding to a static reaction field. From the application of the organic solutes solvated by water as will

be presented in Section 4.3, that ratio is demonstrated quite close to –2. The magnitude of the polarization term $E[n_u^{(1)}] - E[n_u^{(0)}]$ is evaluated by applying a static solvent reaction field to the QM part. This solvent reaction field reads as $\delta\mu[\bar{n}_u^{(0)}] / \delta n_u$ and can be estimated from the force-field solute model. With the approximation $n_u^{(1)} \approx n_u$ and the assumption that the two relaxation terms for the energy functionals E and μ have a fixed ratio of –2,[21,22,113] the solvation free energy functional including an approximate term of second order in the solvent reaction field is

$$\Delta G_{PT2} = \Delta G_{PT1} - \left(E[n_u^{(1)}] - E[n_u^{(0)}] \right) \tag{3.3.17}$$

It is referred to as "PT2" model. The result $\phi_u = \delta\mu / \delta n_u$ for the reaction field contributes to the KS potential as a mean field of the solvent electrostatic potential. This treatment also depends on the observation that the expressions of the excess chemical potential, as discussed in Section 2.4.5, are path-independent integrals of work,[86,98] and on the assumption that the electrons interact with the solvent charged sites according to an unscreened Coulomb law.[48] Therefore the PMV correction $\mu^{PMV} - \mu$, as previously suggested in Eq. (2.4.30) in Section 2.4.6, can be added as an empirical post-SCF term instead of being part of the self-consistent expression, Eq. (3.3.11). This post-SCF treatment is also based on the observation that the molar volume V_u of the solute is almost independent of the various options to approximate the solute charge density n_u ; this will be demonstrated in Section 4.3.

3.4. Numerical Aspects

3.4.1. *Solving the Poisson Equation with FFT*

In the following we denote the charge density $n(\mathbf{r})$ and the electrostatic potential $\phi(\mathbf{r})$ represented on a 3D Cartesian grid with a grid spacing of $h = L / N$ as

$$\begin{aligned} n(\mathbf{r}) &\rightarrow n(x, y, z) \rightarrow n(i, j, k) \\ \phi(\mathbf{r}) &\rightarrow \phi(x, y, z) \rightarrow \phi(i, j, k) \end{aligned} \tag{3.4.1}$$

To this end, $i = x / h$, $j = y / h$, and $k = z / h$ denote the grid indices along the x, y, and z-axis. The Poisson equation can be expressed as

$$\frac{n^2}{L^2} \Delta\phi(i, j, k) = -4\pi n(i, j, k) \tag{3.4.2}$$

When applying the Fourier transformation to n and ϕ, one obtains

$$\frac{4\pi^2 \mathbf{k}^2}{L^2} \mathcal{F}_3(\phi) = \mathcal{F}_3(n) \tag{3.4.3}$$

with $\mathbf{k}^2 = \mathrm{k}_x^{\,2} + \mathrm{k}_y^{\,2} + \mathrm{k}_z^{\,2}$ being the square summation of the wavenumbers in k-space. After applying the inverse Fourier transformation to $\mathcal{F}_3(n)$ and dividing by the normalization factor N^3, ϕ is conveniently calculated from Eq. (3.4.3) by the FFTW3 implementation:

$$\phi = \frac{1}{N^3} \mathcal{F}_3^{-1}[\mathcal{F}_3(\phi)] = \frac{1}{N^3} \mathcal{F}_3^{-1}[\frac{L^2 \mathcal{F}_3(n)}{4\pi^2 \mathbf{k}^2}] \tag{3.4.4}$$

Table 3.4.1. Excess chemical potentials calculated by 3D RISM, with the long-range Coulomb potential expressed analytically and obtained by solving the Poisson equation with FFT.[a]

Solvent model	Analytical	FFT
Two-point charge[b]	16.29	16.29
PR-SPC/E water[c]	–6.32	–6.31

[a] Energies in kcal/mol. [b] Pseudo solvent model with point charge 0.1 *e* and –0.1 *e* on two sites. [c] Ref. 115.

To validate the electrostatic potential obtained from Eq. (3.4.4) with the 3D RISM formalism, the excess chemical potentials calculated from this latter intermediate are compared to those obtained from the analytical expression of the Coulomb potential. Table 3.4.1 shows the results of a pseudo solvent model with point charges 0.1 *e* and –0.1 *e* on the two sites and those from the PR-SPC/E water model.[115] As can be seen, for the pseudo two-point charge model, the analytical and FFT treatments yield nearly identical results. When the water model is applied, the excess chemical potentials resulting from these two treatments

only differs by 0.01 kcal/mol. This comparison shows that the electrostatic potential from solving the Poisson equation with FFT yields the same results as when directly applying the analytical expression of the Coulomb potential. The FFT treatments is comparatively efficient as it allows for the direct usage of the solute charge density obtained from the DFT calculation. In contrast to that, the analytical expression needs to be tabulated on a real-space grid using the effective solute site charges. For a solute with a QM description its effective site charges are usually fitted from the population analysis,[3,4] with which additional evaluations of the charge fitting procedure may be required.[6,105]

3.4.2. *Evaluating the Electrostatic Interaction in QM+RISM*

From Eqs. (3.1.5) and (3.1.6), one obtains

$$u^{\mathrm{C}} = \int n_{\mathrm{u}}(\mathbf{r})\phi_{\mathrm{v}}(\mathbf{r})d\mathbf{r} \tag{3.4.5}$$

for the electrostatic interaction between the solute in a QM description and the solvent determined by 3D RISM. Calculating the integral in Eq. (3.4.5) yields the electrostatic interaction between the solute in a DFT description and the solvent environment determined by 3D RISM. On the one hand, this calculation can be done within the DFT program package ParaGauss, by forming a trace of the product of the density matrix with the matrix representation of the solvent potential and including the contribution arising from the nuclei. This result is labeled as $u^{\mathrm{C}}_{\mathrm{QM}}$. Alternatively, the results can also be obtained by integrating the solvent electrostatic potential over the diffuse charge density of the solute, which, as described in Section 3.2.2, includes the electron density and the broadened core charge density. This latter approach yields the long-range contribution $u^{\mathrm{CL}}_{\mathrm{RISM}}$. The short-range part $u^{\mathrm{CS}}_{\mathrm{RISM}}$ in turn needs to be estimated with the analytical expression of the electrostatic kernel in Eq. (3.2.20) to compensate the broadened core charges. Given the different implementation details in ParaGauss and the 3D RISM solver, it is necessary to compare the value of u^{C} which can be calculated either way. The main goal of this comparison consists in validating the transfer of the solute charge density and the solvent electrostatic potential between the ParaGauss program and the 3D RISM solver via the grid mapping scheme (Section 3.2.3)

In Table. 3.4.2 the electrostatic interactions calculated between water as solvent described by the TIP3P model[76,115] with different solute molecules are presented. The values in the second column exactly match those in the first column. Thus the electrostatic interaction between solute and solvent is described essentially identical by ParaGauss and 3D RISM. Therefore, the transfer of the electron density and the solvent potential through the interface between 3D RISM and ParaGauss in the hybrid approach does not result in any loss of

accuracy. Besides, when comparing analogous values in the first and third columns one notices that by integrating only the "diffuse" part of charge density with the solvent potential, the electrostatic interactions are overestimated by nearly 20%. Thus, the compensation of the broadened core charges is indeed necessary.

Table 3.4.2. Electrostatic interaction calculated in ParaGauss and in the 3D RISM solver.[a]

Solute	$u_{\mathrm{QM}}^{\mathrm{C}}$	$u_{\mathrm{RISM}}^{\mathrm{CL}} + u_{\mathrm{RISM}}^{\mathrm{CS}}$	$u_{\mathrm{RISM}}^{\mathrm{CL}}$	$u_{\mathrm{RISM}}^{\mathrm{CS}}$
Methanol	–13.65	–13.65	–15.95	2.30
Formic acid	–14.31	–14.31	–17.03	2.72
Formaldehyde	–9.04	–9.04	–10.91	1.87

[a] Energies in kcal/mol. Computational details and parameters for the solute can be found in Section 4.1

3.4.3. *Convergence of 3D RISM with Grid Settings*

Table 3.4.3. Convergence of the excess chemical potential with respect to the grid settings, obtained in 3D RISM calculations of a pure solvent. **N** is the number of grid points in each of the three dimensions, **L** the length of the cubic cell.[a]

	L			
N	20	25	30	35
128	–5.79	–5.81	–5.84	–5.83
160	–5.78	–5.82	–5.84	–5.83
192	–5.78	–5.82	–5.84	–5.85
224	–5.79	–5.82	–5.84	–5.85

[a] Energies in kcal/mol. Length in Å. TIP3P water was used.[76,115]

Before using the hybrid approach in actual applications the convergence of the excess chemical potential was also examined with regard to the grid settings. The results were obtained with 3D RISM pure solvent calculation using the TIP3P water model[76,115] and are shown in Table 3.4.3. Apparently, for a given size of the cubic unit cell, the reduction of the grid step, hence increasing the number of grid points from N = 128 to N = 224 in each dimension, yields only a slight change of at most 0.02 kcal/mol. The results are more sensitive to an increase of the cubic cell size from L = 20 Å to L = 35 Å, which may be due to the long-range interactions. However, even the difference between the smallest cubic cell (L = 20 Å) and the largest one (L = 35 Å) only amount to 0.07 kcal/mol. Therefore, it can be claimed that the most efficient grid setting (N = 128, L = 20 Å) in the 3D RISM calculation yields a sufficiently high accuracy. This setting is therefore used in the applications discussed in Chapter 4.

4. Applications

4.1. General Computational Details

4.1.1. *RISM Calculations*

All the RISM calculations were carried out for the temperature T = 298 K. For the benchmark calculations on monatomic alkali and halide ions, the coincident SPC/E water model (cSPC/E)[116] with σ_H = 1.1658 Å and ε_H = 0.01553 kcal/mol was used. The dielectric constant of water solvent was set to 78.4 with the corresponding water solvent density ρ = 0.0333295 $Å^{-3}$ to be consistent with the reference RISM calculation.[88] The force field parameters for alkali and halide ions were adopted from reference RISM and MD calculations.[88,117] The three-dimensional grid for representing the correlation functions was chosen with 4096 grid points in each Cartesian direction and a uniform spacing of 0.0390625 Å; this amounts to a radial range of 160 Å. In the 3D RISM test calculations, a moderate setting for the grid resolution was used, with 64^3 points in total in a cubic volume of 20^3 $Å^3$.

The geometries of 43 neutral organic molecules chosen as solutes in Section 4.3 were taken from the database of Mobley et al.[118] and the NIST database.[119] The LJ parameters and the solute site charges were taken from the OPLS-AA force field.[120] The TIP3P water model,[76] was used with the modified LJ parameters for hydrogen atoms, σ_H = 0.4 Å and ε_H = 0.046 kcal/mol, as suggested by Pettitt and Rosky.[115] The corresponding water solvent particle density was set to ρ = 0.033427745 $Å^{-3}$, equivalent to 1.0 g/cm^3 which slightly differs from the parameter chosen for the initial benchmark calculations but is more commonly used in other 3D RISM studies.[35,47,121,122] A uniform grid with a spacing of 0.15625 Å and 128 grid points in each of the three dimensions was chosen to represent the correlation functions. This amounts to a cubic grid of (20 Å)^3, consistent with other RISM applications.[46,62,123]

Various options of the uranyl force-field models were examined in the application of uranyl solvation. First force-field parameters as suggested by Guilbaud and Wipff (GW)[124] were chosen and the results were compared to those of models developed by Maginn et al. (RM12 and PM13),[125,126] see Table 4.1.1. Recently, Kerisit and Liu modified the parameters of the GW model to strengthen the uranyl–water interaction.[127] For their models KL1 and KL2 they reported improved values for the hydration free energy and the water exchange rates of the uranyl aqua complex. Therefore, in this thesis the results of KL1 and KL2 models were also compared with those obtained with the GW, RM12, and PM13 variants. For the water solvent, the atomic charges and LJ potential parameters of the extended simple point charge

model SPC/E[78] were adopted. While the original SPC/E water model does not contain a repulsive potential for hydrogen atoms, a LJ-type potential for hydrogen atoms is necessary in RISM calculations to avoid numerical singularities. In their early study, Pettitt and Rosky[115] introduced a parameter set with σ = 0.4 Å and ε = 0.046 kcal/mol (PR-SPC/E). Sato and Hirata[128] adopted the parameters for the hydrogen atom as σ = 1.0 Å and ε = 0.05455 kcal/mol (SH-SPC/E). Table 4.1.2 summarizes the details of these three popular modified water models. The PR-SPC/E and KL2 models for water and uranyl, respectively, were applied in the 3D RISM study of uranyl solvation reported later on.

In the RISM applications of uranyl solvation, the same dielectric constant and number density of water solvent were employed as in the benchmark calculations of monatomic akali and halide ions. The grid setting in 1D RISM was kept the same and the solvent susceptibility was determined from the pure water site-site RISM equation by Newton iterations with a convergence tolerance of 10^{-14} for the L_2 norm of the indirect correlation matrix. The discrete sine transform was applied for the Fourier transformation when calculating convolution integrals in 1D RISM.[109] The grid setting was slightly refined for the 3D RISM applications, in which the 3D domain was represented by a cubic cell of 20^3 Å^3, a Cartesian grid with 96^3 points, and a resolution of 0.208 Å. The radial grid used to precompute and interpolate the solvent susceptibility and its contractions with the Coulomb kernel was chosen with a radial range of 40 Å and 1536 points corresponding to a resolution of about 0.026 Å.

Table 4.1.1. Lennard–Jones interaction parameters of uranyl: σ (Å), ε (kcal/mol), and atomic charge q (e).

Uranyl model	σ_U	ε_U	q_U	σ_O	ε_O	q_O
GW	2.8152	0.4	2.5	3.1181	0.2	−0.25
RM12	3.3334	0.0268	2.5	2.8344	0.4290	−0.25
PM13	2.95	0.1266	2.5	3.83	0.0136	−0.25
KL1	2.8509	0.12	3.25	3.1181	0.2	−0.625
KL2	2.8152	0.3	3.5	3.1181	0.2	−0.75

The geometry optimizations for the applications on uranyl solvation have been carried out using the BFGS method as implemented in the utility package ParaTools.[129,130] During the QM calculations, the U–O bonds were relaxed under C_{8h} symmetry constraints. In the MM and MM+RISM calculations uranyl was kept rigid, however. Motivated by earlier results, the U–O distances were fixed to 176 pm in the RISM calculations of the bare uranyl ion.[125,126] In MM+RISM geometry optimizations of uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$, the uranyl

bonds were fixed to 178 pm for n = 4 and n = 5, and set to 177 pm for n = 6. These choices for the fixed uranyl bond lengths are consistent with results of previous studies.[131-133] However, for a study of the water exchange with the MM+RISM approach, a flexible uranyl model without symmetry constraints was adopted and compared to two sets of intramolecular parameters[127,134] from the results of unconstrained QM+RISM calculations. The two uranyl force field models employed differ only by the uranyl equilibrium bond length and the force constants for stretching and bending modes. The SPC/E geometry of explicit water ligands in the aqua complexes $[UO_2(H_2O)_n]^{2+}$ with n = 4–6, features an O–H distance of 100 pm and an H–O–H angle of 109.47 °.

Table 4.1.2. Lennard–Jones interaction parameters of water: σ (Å), ε (kcal/mol), and atomic charge q (e).

Water model	σ_O	ε_O	q_O	σ_H	ε_H	q_H
SH-SPC/E	3.1656	0.1553	−0.8476	1.0	0.0545	0.4238
PR-SPC/E	3.1656	0.1553	−0.8476	0.4	0.046	0.4238
cSPC/E	3.1656	0.1553	−0.8476	1.1658	0.01553	0.4238

In the application with acetonitrile as solvent (Section 4.5), the LJ parameters and the solvent site charges were taken from a new six-site acetonitrile model.[135] The dielectric constant of acetonitrile solvent was set to its experimental value ε = 36,[136] with a corresponding solvent number density ρ = 0.033427745 Å^{-3}.[35,47,120,121] The LJ parameter for the H,C,N and O sites in the Ru-pyridine complex were selected from the OPLS-AA force field,[120] by referring to the pyridine group and formic acid group. For the ruthenium atom, we applied the parameters of a recent molecular dynamic study of Ru-pyridine in water.[137] To represent the correlation functions in this case a uniform grid with a spacing of 0.15625 Å and 96 grid points in each of the three Cartesian directions was chosen. This amounts to a cubic grid extending over a spatial volume of (32 Å)^3, which, due to the large size of the Ru-pyridine complex catalysts, is more than 50% larger than the previous 3D RISM applications.

4.1.2. *QM Calculations*

For the uranyl solvation applications, all-electron calculations were carried out with the linear combination of Gaussian-type orbitals fitting-functions (LCGTO-FF) density functional method[138] as implemented in the parallel code ParaGauss.[66] The second-order Douglas–Kroll–Hess all-electron scalar-relativistic approach[139] was used to account for relativistic effects in the electronic structure of the actinide compound.[140] The gradient-corrected exchange-correlation functional (generalized gradient approximation, GGA)[2,141] was chosen as suggested by Becke and Perdew (BP).[142,143] The Kohn–Sham orbitals were represented by flexible Gaussian-type basis sets, which were contracted in a generalized fashion using atomic eigenvectors. Basis sets and further computational details are the same as in earlier studies on uranyl(VI) species.[131,144-146] For U, a basis set of the size (24s, 19p, 16d, 11f) contracted to [10s, 7p, 7d, 4f] was used, while O and H atoms were described by standard basis sets featuring the contractions (9s, 5p, 1d) → [5s, 4p, 1d] and (6s, 1p) → [4s, 1p], respectively. Atomic eigenvectors were employed to construct the general contractions for the calculation of the Hartree potential in ParaGauss.[147] The electronic density is approximately represented by a set of fitting functions.[138,148] The *s*- and r^2 exponents of this set of fitting functions were determined from a subset of the *s*- and *p*-orbital exponents scaled by a factor of 2; the fit basis was augmented by sets of five *p*- and *d*-type "polarization" exponents,[148], chosen as geometric series with factors of 2.5, starting at 0.1 and 0.2 au, respectively.

For the application addressing 43 neutral solutes in Section 4.3, the GGA exchange-correlation functional suggested by Perdew, Burke and Ernzerhof (PBE) was used.[143,149] The CRENBL effective core potential (ECP),[150-153] which entails an orbital basis with a large-core ECP was employed, together with the Coulomb fitting basis set developed by Ahlrichs and coworkers.[154,155]

For the QM calculations carried out with ParaGauss in Sections 4.3 and 4.4, the COSMO PCM solvation model was used for comparison,[69] with the cavity defined via an effective solvent radius of 1.4 Å and van der Waals radii of solute atoms as tabulated by Bondi[156] and scaled by 1.125 (except for hydrogen). The cavity tessellation with the FIXPVA approach[146,157] (fixed points with variable areas) was employed to obtain numerically stable energies and forces.[146]

For the application of CO_2 conversion in acetonitrile solvent, the geometries and electronic energies for the species in the previous DFT study[158] were used as calculated with the conductor-like polarizable continuum model (CPCM).[159] For the underlying electronic structure calculations the Gaussian 03 program package[160] and the B3LYP[142,161] hybrid functional in the unrestricted KS formalism were applied on geometries obtained in the

preceding study of Damianos et al.[158] For the Ru atom, the Stuttgart-Dresden effective core potential MWB28 ECP[162] was employed and for the lighter atoms H,C,N and O the 6-31G(d,p) basis set was used. Additional charge analyses for all species were carried out by the natural population analysis (NPA); these atomic charges were later used later on in the 3D RISM calculations.

4.2. Monoatomic Ions Test

To evaluate the RISM approach implemented in the framework of this thesis, the recently reported 1D RISM results[88] for the solvation of alkali and halide ions were reproduced by applying the cSPC/E water model. As shown in Table 4.2.1, the results for the excess chemical potential of simple ions as obtained with 1D RISM in the context of this thesis agree well with the reported values and are also comparable to the results from MD calculations with SPC/E water as well as those from experiments.[163] The most significant technical difference to Ref. 88 is the choice of a smaller radial range, 160 Å (819 Å in Ref. 88), and a somewhat coarser grid spacing, 0.0390625 Å (0.025 Å in Ref. 88). With the moderate grid setting described in Section 4.1.1 the 3D RISM results match the 1D cases within 0.2 kcal/mol at most (Table 4.2.1).

Table 4.2.1. Solvation energies (excess chemical potentials) for alkali metal and halide ions in water by 1D and 3D RISM approaches with the cSPC/E model.[a]

	Li^+	Na^+	K^+	Cl^-	Br^-	I^-
1D KH[b]	−108.5	−85.0	−68.5	−78.5	−72.6	−63.1
3D KH[b]	–108.5	–85.0	–68.5	–78.4	–72.5	–62.9
1D KH[c]	−108.3	−85.0	−68.6	−78.1	−72.2	−62.7
1D HNC[b]	−111.7	−86.3	−69.3	−79.9	−74.2	−64.9
3D HNC[b]	–111.7	–86.3	–69.3	–79.8	–74.1	–64.7
1D HNC[c]	−111.4	−86.2	−69.4	−79.5	−73.7	−64.5
1D KH+PMV[d]	−108.5	−87.2	−73.0	−85.1	−80.5	−73.4
1D HNC+PMV[d]	−110.3	−87.4	−72.7	−84.8	−80.0	−72.8
MD[e]	−113.3	−88.4	−71.0	−89.3	−82.7	−74.4
Exp[f]	−113.8	−88.7	−71.2	−89.1	−82.7	−74.3

[a] MD results and experimental values are also shown. Energies in kcal/mol. [b] This work. [c] Ref. 88. [d] This work, with partial molar volume correction as in Ref. 164; see Section 2.4.6. [e] MD with SPC/E water from Ref. 117. [f] Ref. 163.

The excess chemical potential is insensitive to the specific value of the radial range. With the radial range set to 320 Å and 80 Å while keeping the same grid spacing, the excess chemical potential changes by 0.01 kcal at most. In contrast, explicit solvent models may require a size-dependent correction.[165,166] The RISM model predicts a double charge layer around a neutral ion leading to a finite electrostatic cavity potential of 10–14 kcal/(mol·*e*) depending on the size of the ion and the dispersion parameters. This is comparable to estimates

from MD for the jump of the vacuum–liquid interfacial potential[167] and about half of the value for an orientationally disordered liquid,[166] which in explicit methods are highly sensitive to the Coulomb summation scheme.

Overall the HNC closure relation yields slightly better results than the KH closure as compared to MD and experimental reference data. Compared to experiments, the RISM model underestimates the magnitude of the solvation free energies of cations by a few percent only and the results for anions are slightly worse. The difference results in part directly from the RISM approximation and in part from the difference in LJ parameters of the hydrogen centers of water. Another intrinsic problem of the RISM approach is that it overestimates the cavity formation energy which may be approximated as the solvation energy of the uncharged solute. For larger anions the cavity formation energy amounts to 14–19 kcal/mol whereas the MD suggests smaller values of ~7 kcal/mol.[88,168] Although this error is notable for neutral solutes, it is less important for ions where the electrostatic contribution dominates. This holds even more so for the uranyl di-cation. To some extent this overestimation may be compensated by a correction proportional to the molar volume of the solute;[164] see the calculations with the PMV corrections in Table 4.2.1.The more pronounced compensation effects of the PMV corrections for neutral solutes will be discussed in Section 4.3.

4.3. 3D RISM Combined with a QM Description of the Solute

4.3.1. *Introduction*

Following the discussions in Section 3.3 it can be anticipated from the free energy functionals that two main factors determine the solvent effect in the free energy calculations when combining QM calculation with RISM. These are the contributions of the solute electronic relaxation and the excess chemical potential.[5] As discussed in Section 2.4.6, the excess chemical potential calculated from 3D RISM is well known[5] to overestimate the energy required for the formation of the solute cavity. Thus 3D RISM poorly predicts the thermodynamics of hydrophobic hydration of neutral solutes.[169] To overcome this weakness, several correction schemes have been proposed for the solvation free energy, including the repulsive bridge correction (RBC),[5,169] the cavity formation energy correction,[6,62] and the partial molar volume (PMV) correction;[61,63,64] see Section 2.4.6. Thus far, these correction methods have mainly been applied in MM-RISM calculations, in which a static electronic structure of the solute is implied. Yet, the energy due to the electronic relaxation of the solute is always found significant in the case of SCF-RISM and PCM calculations.[6,170] The change of the solvation free energy due to solute polarization is known to be almost proportional to this term,[6,113] but so far the influence of the solute polarization on the accuracy of the solvation free energy from SCF-RISM calculations has not been studied in detail.

Therefore, in the following, the accuracy of the solvation free energy in hybrid QM RISM calculations will be examined for a test set of 43 neutral organic solute molecules in aqueous solution. This test set comprises hydrocarbon chains and aromatic rings, as well as various organic molecules with functional groups. A list of all species can be found in Figure 4.3.1. The effect of solute electronic relaxation is explored in stepwise fashion following the free energy functionals in Section 3.3. Starting from MM-RISM as a reference, the polarization of the solute in the solvent medium is then introduced by applying first- and second-order perturbation theory, labeled as PT1-RISM and PT2-RISM, respectively, and finally in terms of the full SCF-RISM relaxation. For all method variants, the performance of a recently reported PMV correction[63] is examined with its absence and presence in the solvation free energy. These results are compared to those from PCM calculations and to experimental measurements, to analyze the accuracy the RISM approach.

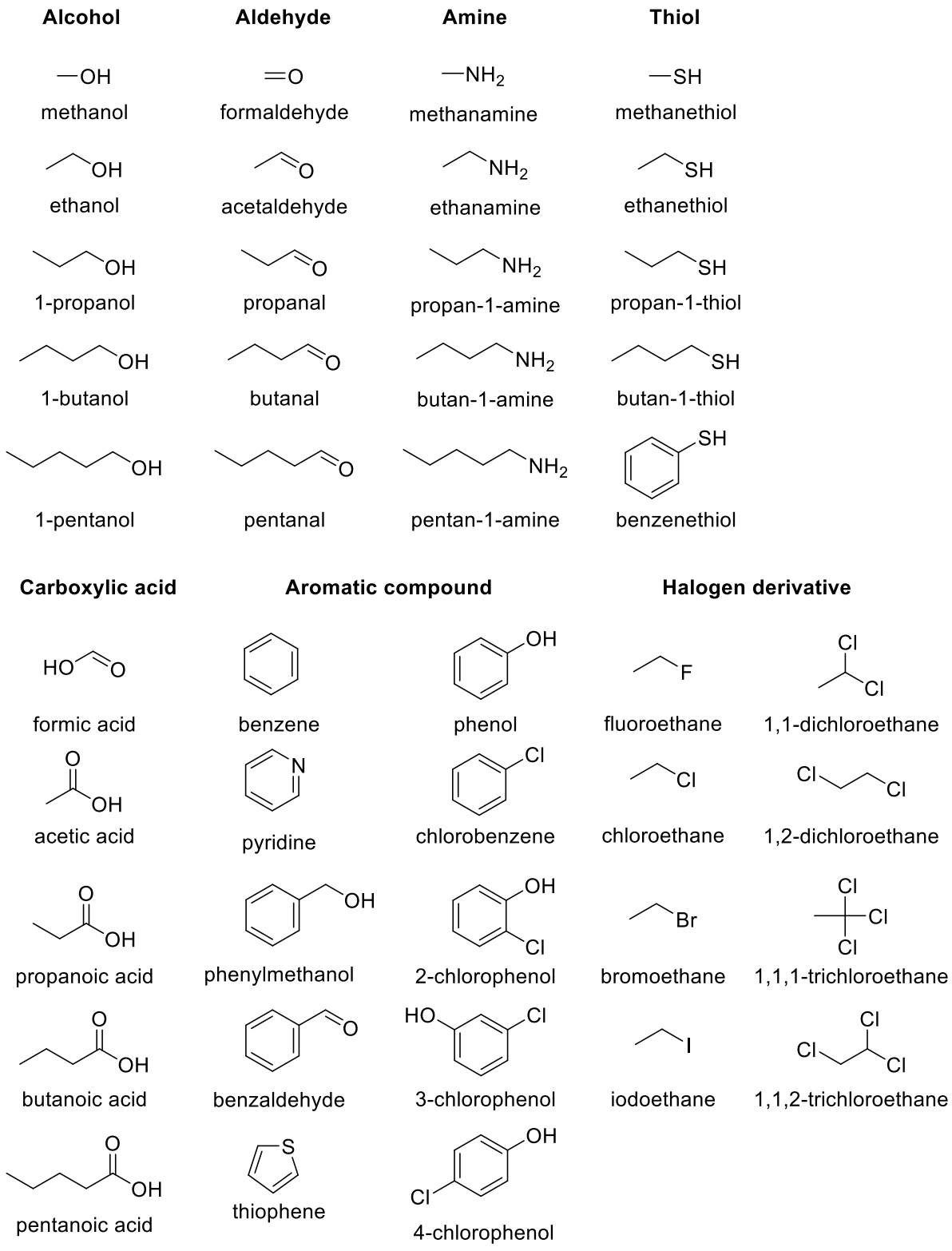


Figure 4.3.1. 43 compounds, categorized in seven groups, that were examined as solutes.

Table 4.3.1. Experimental solvation free energies (kcal/mol) and the corresponding values calculated with PCM and 3D RISM using various methods. Results without and with PMV corrections are shown.[a]

Solute	ΔG_{MM}	ΔG_{MM}^{PMV}	ΔG_{PT1}	ΔG_{PT1}^{PMV}	ΔE_{PT2}	ΔG_{PT2}	ΔG_{PT2}^{PMV}	ΔE_{SCF}	ΔG_{SCF}	ΔG_{SCF}^{PMV}	ΔG_{PCM}	ΔG_{Exp}
methanol	2.77	-5.89	2.66	-5.90	1.72	0.94	-7.62	2.46	0.70	-7.77	-5.42	-5.11
ethanol	6.17	-6.10	7.12	-5.02	2.14	4.98	-7.16	2.67	5.07	-7.00	-4.26	-5.00
1-propanol	11.36	-4.35	11.95	-3.71	1.72	10.22	-5.43	2.32	10.15	-5.44	-2.61	-4.83
1-butanol	14.79	-4.25	15.72	-3.25	1.92	13.79	-5.17	2.47	13.82	-5.07	-1.47	-4.72
1-pentanol	19.32	-3.44	20.10	-2.60	1.89	18.22	-4.48	2.52	18.17	-4.45	-0.28	-4.52
formaldehyde	4.17	-3.71	4.97	-2.87	1.58	3.4	-4.45	2.31	3.34	-4.46	-3.76	-2.76
acetaldehyde	8.19	-3.12	8.40	-2.90	1.60	6.8	-4.5	2.92	6.43	-4.84	-3.34	-3.51
propanal	12.45	-2.30	12.53	-2.23	1.49	11.04	-3.72	2.77	10.64	-4.10	-2.17	-3.44
butanal	16.51	-1.67	16.64	-1.55	1.48	15.16	-3.03	2.74	14.77	-3.40	-1.00	-3.18
pentanal	20.69	-1.19	21.19	-0.73	1.62	19.58	-2.35	2.93	19.24	-2.68	0.14	-3.03
formic acid	3.51	-4.98	0.87	-7.46	1.17	-0.29	-8.62	2.19	-0.87	-9.07	-7.41	-6.99
acetic acid	7.29	-4.60	4.14	-7.63	1.20	2.94	-8.83	2.67	2.09	-9.57	-6.61	-6.70
propanoic acid	10.99	-4.46	7.61	-7.11	1.24	6.37	-8.34	2.87	5.44	-9.19	-6.02	-6.48
butanoic acid	15.86	-3.04	12.73	-6.10	1.14	11.59	-7.25	2.72	10.66	-8.09	-4.39	-6.38
pentanoic acid	20.30	-2.37	17.26	-5.34	1.19	16.07	-6.53	2.70	15.19	-7.30	-2.91	-6.20
methylamine	3.87	-5.34	2.13	-7.00	1.92	0.2	-8.93	3.74	-0.64	-9.64	-4.76	-4.55
ethylamine	8.19	-4.47	7.00	-5.60	1.94	5.06	-7.55	3.82	4.24	-8.24	-3.47	-4.50
propan-1-amine	12.21	-3.92	11.09	-5.00	1.95	9.14	-6.95	3.88	8.30	-7.68	-2.37	-4.39

[a] ΔG represents the solvation free energy, in which the subscripts MM, PT1, PT2, and SCF denotes the results calculated with MM-, PT1-, PT2- and SCF-RISM models, respectively, and the superscript PMV denotes the results corrected by the PMV correction. PCM and experimental results are labeled as ΔG_{PCM} and ΔG_{Exp}. For PT2- and SCF-RISM models the solute relaxation energies due to the solvent medium are labeled as ΔE_{PT2} and ΔE_{SCF}, respectively.

Table 4.3.1 (continued).

Solute	ΔG_{MM}	ΔG_{MM}^{PMV}	ΔG_{PT1}	ΔG_{PT1}^{PMV}	ΔE_{PT2}	ΔG_{PT2}	ΔG_{PT2}^{PMV}	ΔE_{SCF}	ΔG_{SCF}	ΔG_{SCF}^{PMV}	ΔG_{PCM}	ΔG_{Exp}
butan-1-amine	16.15	-3.48	15.24	-4.35	2.04	13.2	-6.4	4.05	12.36	-7.11	-1.17	-4.24
pentan-1-amine	20.47	-2.74	19.54	-3.66	2.02	17.52	-5.67	4.02	16.67	-6.40	-0.02	-4.09
methanethiol	10.23	-0.26	8.84	-1.54	0.69	8.16	-2.23	1.53	7.72	-2.54	-3.30	-1.24
ethanethiol	14.27	0.28	13.23	-0.68	0.76	12.47	-1.43	1.63	12.06	-1.72	-2.06	-1.14
propan-1-thiol	18.55	1.10	17.26	-0.14	0.67	16.59	-0.81	1.56	16.13	-1.14	-0.89	-1.06
butan-1-thiol	22.36	1.45	21.71	0.83	0.78	20.93	0.05	1.61	20.58	-0.19	0.57	-0.99
benzenethiol	21.24	0.53	20.03	-0.47	0.61	19.42	-1.07	1.31	19.03	-1.31	-2.70	-2.55
benzene	17.89	0.04	18.97	2.22	0.53	18.44	0.73	0.31	18.71	1.02	0.36	-0.86
pyridine	13.86	-2.94	13.86	-3.85	1.74	12.12	-4.63	4.00	11.39	-5.35	-3.11	-4.69
phenylmethanol	17.62	-4.86	18.32	-3.98	1.80	16.52	-5.79	2.48	16.39	-5.82	-3.00	-6.62
benzaldehyde	18.77	-2.68	19.29	-2.09	1.41	17.88	-3.5	3.37	17.17	-4.20	-2.57	-4.02
thiophene	15.77	-0.03	16.28	0.58	0.42	15.86	0.16	0.26	16.06	0.37	-1.70	-1.42
phenol	13.82	-4.91	14.62	-3.90	1.76	12.87	-5.65	1.95	13.08	-5.32	-4.60	-6.61
chlorobenzene	21.51	1.29	21.68	1.52	0.28	21.4	1.24	0.29	21.44	1.30	0.11	-1.12
2-chlorophenol	19.01	-2.18	19.38	-1.68	1.09	18.29	-2.77	1.20	18.45	-2.56	-2.88	-4.55
3-chlorophenol	17.05	-4.09	17.47	-3.49	1.56	15.91	-5.04	1.72	16.08	-4.75	-4.68	-6.62
4-chlorophenol	16.89	-4.24	17.25	-3.70	1.62	15.63	-5.32	1.85	15.77	-5.05	-5.02	-7.03
fluoroethane	12.23	0.23	11.89	-0.11	0.48	11.41	-0.59	0.86	11.21	-0.78	-0.82	-0.40
chloroethane	14.20	0.70	14.30	0.81	0.39	13.91	0.42	0.50	13.91	0.43	-0.39	-0.63
bromoethane	14.35	0.46	13.29	-0.46	0.48	12.81	-0.94	0.82	12.67	-1.05	-0.41	-0.74
iodoethane	16.29	1.32	15.82	0.87	0.23	15.6	0.64	0.31	15.58	0.64	-1.57	-0.74
1,1-dichloroethane	17.87	1.87	16.83	0.89	0.16	16.67	0.73	0.43	16.48	0.58	-1.01	-0.84
1,2-dichloroethane	14.61	-1.34	15.45	-0.45	0.95	14.49	-1.41	1.00	14.66	-1.17	-3.09	-1.79
1,1,1-trichloroethane	20.15	1.67	20.27	1.79	0.23	20.04	1.56	0.22	20.08	1.62	0.17	-0.19
1,1,2-trichloroethane	18.28	-0.18	18.30	-0.09	0.73	17.57	-0.82	0.89	17.59	-0.71	-3.27	-1.99

4.3.2. *Asymptotic Analysis of the Cavity Formation Energy*

The expressions of Eqs. (2.4.25) to (2.4.26) for the excess chemical potential in combination with the linear convolution relation $h=c*\chi$ between the total and direct correlation functions h and c, lead to seriously overestimated cavity formation energies.[64,103] The possible reason can be attributed to the contribution of the $|h|^2$ term to the integral in the low-solvent-density region where $h \approx -1$, while natively omitting this term as in the GF functional, Eq. (2.4.26), can yield problematic results for polar solute.[5] Indeed, by assuming a simple liquid exposed to a potential which is sufficiently repulsive within an extended volume V_0 and approximating $h \approx -1$ and $c \approx -z^{-1}$ within that volume, one obtains the excess chemical potential as

$$\beta\mu \approx \rho V_0 (z^{-1} + 1) / 2 \,. \tag{4.3.1}$$

To this end, one uses the convolution relation and the short-hand notation $z = \rho\kappa / \beta$ for the compressibility factor.

Water is nearly incompressible with the compressibility factor $z \approx 1 / 15.8$ at normal conditions.[171] With the KH closure and a modified TIP3P water model,[76,115] both 1D and 3D RISM theories yield a comparable value of about $z = 1/15$. The asymptotic RISM expression for the formation energy of the solute cavity therefore suggests that the energy $a = \beta^{-1}(z^{-1}+1)/2 \approx 5$ kcal/mol is required to displace each of ρV_0 water solvent molecules when forming the cavity at normal conditions. This asymptotic behavior is far too hydrophobic in comparison to the correct asymptote $\mu \approx pV_0$ and the corresponding rate $a = p / \rho$ which amounts to 0.4×10^{-3} kcal/mol per displaced water molecule. Because of this low pre-factor, for many chemical applications the water surface tension of 104×10^{-3} kcal/mol/Å^2 is significantly more important for medium-sized cavities of nanosized solutes.[172] Nevertheless, the large error of the RISM expression for the chemical potential cannot be ignored in practice.

The erroneous behavior of Eq. (4.3.1) likely arises from the fact that this method also relies on the solvent susceptibility χ for the (dense) liquid water inside the cavity where the solvent density is in fact vanishingly small. When one substitutes the compressibility factor in Eq. (4.3.1) by its low-density (ideal gas) limit $z = 1$, the RISM asymptotic expression acquires the correct form $\mu \approx \rho V_0 / \beta = pV_0$.

With regard to the PMV correction expression, Eq. (2.4.30), the correction coefficient a proposed by Sergiievskyi et al.[63] may alternatively be expressed via the solvent compressibility factor $\beta a = -(z^{-1} - 1)/2$. An extension of the asymptotic analysis for simple

liquids to the PMV correction shows that it cancels the main term $(\beta z)^{-1}\rho V_0/2$ of the asymptotic 3D RISM error in the excess chemical potential for nearly incompressible liquids with $z \ll 1$ when $V_u \approx V_0 \to \infty$. Therefore, another value for the correction coefficient ⍺ from a recent work,[61] $\beta a = -(z^{-1}+1)/2$, can also be derived from the asymptotic analysis for simple liquids via Eq. (4.3.1), by assuming $V_u \approx V_0$ and neglecting the pressure term.

4.3.3. *Force Field Solute Model: MM-RISM*

The theory of 3D RISM was originally based on pair-wise interactions between solute and solvent sites.[5] Thus the analysis of this application starts with a MM-RISM model using the LJ parameters and the site charges from a well-tested standard force-field for organic solutes in water.[76,115,120] Similar to other studies,[47,63,64,169] the solvation free energies for 43 neutral organic solutes calculated with MM-RISM are significantly overestimating experiment (Table 4.3.2). Without the PMV correction, the correlation coefficient *r* between MM-RISM and experimental values is only 0.30. The dimensionless PMV factor ρV_u ranges from about 1.9 to 5.5 for the 43 solutes examined. The error grows with the (dimensionless) molar volume at a rate of $a \approx 5$ kcal/mol as estimated in Section 4.3.2. As a consequence the product of *a* and ρV_u dominates the energy of the relatively weak solvation of these organic molecules, ranging from −7 kcal/mol to ~0 kcal/mol.[118,173] The correlation of the calculated solvation energies with experiment is poor. The calculated values are overestimated by 18 kcal/mol on average; see the mean absolute deviation (MAD) between MM-RISM and experimental values in Table 4.4.2. With the PMV correction applied, the quality of the linear correlation between MM-RISM and experiments improves significantly with a correlation coefficient of $r = 0.88$* (Table 4.3.2) and MAD = 1.57 kcal/mol (Table 4.3.3).

The effect of the PMV correction is confirmed by Figure 4.3.2a, which depicts an obvious linear correlation of the deviation from the experiment and the solute molar volume. Table 4.3.4 provides the parameters of the linear regression between the calculated (dimensionless) partial molar volume ρV_u and the difference $\Delta^2 G$ between the calculated and the experimentally measured solvation free energies. According to this regression, the coefficient of determination of the linear fit of the uncorrected results is $R^2 = 0.97$. The linear correlation between ρV_u and the remaining error $\Delta^2 G$ is significantly weakened to $R^2 = 0.37$

* In the correlation analysis both quantities are independent random variables (e.g. computed energy vs. experimental energy). Their relation is investigated in terms of the correlation coefficient *r*. In regression analysis one quantity is an independent random variable while the other random variable depends on it, e.g. $\Delta^2 G$ vs. ρV_u as will be shown in the following. In this case the coefficient of determination R^2 is used to evaluate the linear fit; see Ref. 174.

after applying the PMV correction (Table 4.3.4). This reduction indicates that not the cavity formation error but other factors affect the overall accuracy of the solvation free energy in 3D RISM most. The remaining correlation of the corrected results with PMV manifests itself in the slope $A' = 0.66$ kcal/mol (Table 4.3.4), which may be rationalized with the accuracy of the universal correction coefficient, Eq. (2.4.12). This is can be confirmed by the fitted slope of the uncorrected error $A = 4.88$ kcal/mol (Table 4.3.4) which is somewhat higher than the magnitude of the universal correction coefficient $a = -4.22$ kcal/mol. As suggested in Ref. 63, the latter is derived from the solvent properties.

Table 4.3.2. Matrix of correlation coefficients *r* of solvation free energies, calculated[a] with variants of 3D RISM and PCM as well as from experiment.[b] Values are given both without (uncorr.) and with (corr.) a PMV correction.

	Method	MM	PT1	PT2	SCF	PCM	Exp.
uncorr.	MM	1.00	0.98	0.98	0.97	0.67	0.30
	PT1		1.00	1.00	0.99	0.71	0.33
	PT2			1.00	1.00	0.72	0.40
	SCF				1.00	0.72	0.41
	PCM					1.00	0.73
corr.	MM	1.00	0.90	0.93	0.90	0.70	0.88
	PT1		1.00	0.99	0.99	0.78	0.88
	PT2			1.00	1.00	0.73	0.89
	SCF				1.00	0.71	0.86

[a] This work. [b] Refs. 118,173. The corresponding energy values are listed in Table 4.3.1

Table 4.3.3. Matrix of mean absolute deviations MAD of solvation free energies (kcal/mol), calculated[a] with variants of 3D RISM and PCM as well as from experiment.[b] Values are given both without (uncorr.) and with (corr.) a PMV correction.

	uncorr.				corr.	
Model	*A*	*B*	R^2	MAD	*A´*	R^2
MM	4.88	-1.22	0.97	0.73	0.66	0.37
PT1	5.12	-2.49	0.97	0.72	0.91	0.51
PT2	5.09	-3.59	0.94	1.03	0.87	0.33
SCF	5.13	-3.97	0.92	1.24	0.92	0.27

[a] This work. [b] Refs. 118,173. The corresponding energy values are listed in Table 4.3.1

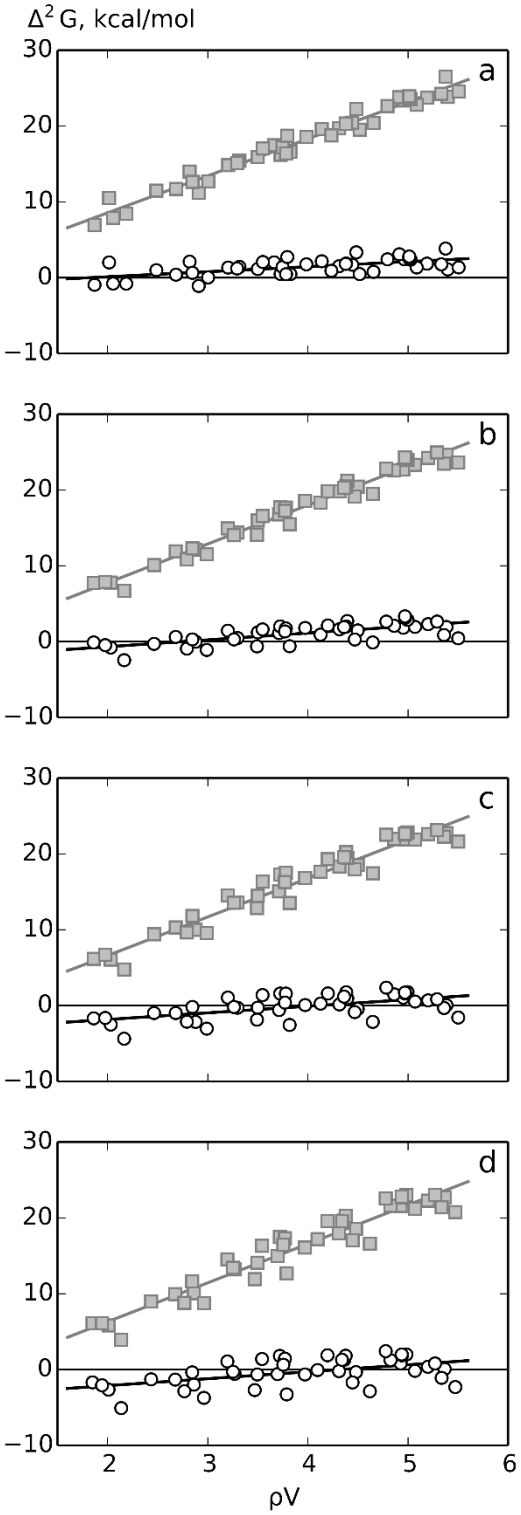


Figure 4.3.2. Correlation between the calculated solute partial molar volume ρV_u and the difference $\Delta^2 G$ of the computed and the experimental solvation free energies. Results are shown for four approaches to the solute-solvent interactions: a) MM, b) PT1, c) PT2, and d) SCF. Gray squares represent uncorrected results, white circles designate PMV-corrected results. For both datasets a linear regression line is shown.

Table 4.3.4. Partial molar volume V scaled by the solvent number density ρ and the dipole moment of the solute calculated with various methods.[a]

	ρV			Dipole				
Solute	MM	PT1/2	SCF	MM	PT1	PT2	SCF	PCM
methanol	2.05	2.03	2.01	0.48	0.42	0.54	0.56	0.51
ethanol	2.91	2.88	2.86	0.50	0.39	0.54	0.55	0.50
1-propanol	3.73	3.71	3.70	0.49	0.43	0.57	0.59	0.53
1-butanol	4.52	4.50	4.48	0.49	0.39	0.56	0.58	0.52
1-pentanol	5.40	5.38	5.37	0.49	0.40	0.57	0.59	0.52
formaldehyde	1.87	1.86	1.85	0.56	0.52	0.67	0.71	0.66
acetaldehyde	2.68	2.68	2.67	0.59	0.58	0.77	0.84	0.77
propanal	3.50	3.50	3.50	0.59	0.60	0.78	0.85	0.78
butanal	4.31	4.32	4.31	0.59	0.61	0.80	0.87	0.79
pentanal	5.19	5.20	5.20	0.59	0.57	0.77	0.85	0.78
formic acid	2.01	1.98	1.94	0.30	0.27	0.36	0.38	0.35
acetic acid	2.82	2.79	2.77	0.28	0.34	0.43	0.49	0.45
propanoic acid	3.66	3.49	3.47	0.28	0.34	0.42	0.49	0.43
butanoic acid	4.48	4.47	4.45	0.27	0.39	0.47	0.55	0.48
pentanoic acid	5.38	5.36	5.34	0.28	0.38	0.47	0.55	0.48
methylamine	2.19	2.17	2.14	0.38	0.37	0.49	0.54	0.46
ethylamine	3.00	2.99	2.96	0.38	0.36	0.50	0.55	0.46
propan-1-amine	3.83	3.82	3.79	0.38	0.37	0.52	0.57	0.48
butan-1-amine	4.66	4.65	4.62	0.38	0.36	0.51	0.56	0.47
pentan-1-amine	5.51	5.50	5.47	0.38	0.37	0.52	0.57	0.48
methanethiol	2.49	2.46	2.43	0.42	0.38	0.50	0.54	0.50
ethanethiol	3.32	3.30	3.27	0.43	0.38	0.51	0.56	0.52
propan-1-thiol	4.14	4.13	4.93	0.43	0.39	0.52	0.57	0.52
butan-1-thiol	4.96	4.95	4.10	0.43	0.36	0.51	0.55	0.51
benzenethiol	4.91	4.86	4.83	0.30	0.27	0.36	0.41	0.38
benzene	4.23	4.20	4.20	0.00	0.00	0.00	0.00	0.00
pyridine	3.99	3.97	3.97	0.50	0.52	0.74	0.85	0.74
phenylmethanol	5.33	5.29	5.27	0.48	0.49	0.63	0.68	0.59
benzaldehyde	5.09	5.07	5.07	0.53	0.74	0.91	1.14	1.02
thiophene	3.75	3.72	3.72	0.30	0.09	0.19	0.11	0.12
phenol	4.44	4.39	4.36	0.45	0.33	0.48	0.47	0.45
chloroethane	3.20	3.20	3.20	0.45	0.42	0.53	0.54	0.55
bromoethane	3.29	3.26	3.25	0.48	0.39	0.52	0.56	0.53
iodoethane	3.55	3.55	3.54	0.38	0.34	0.44	0.45	0.46
1,1-dichloroethane	3.79	3.78	3.77	0.31	0.44	0.51	0.56	0.58
1,2-dichloroethane	3.78	3.77	3.75	0.75	0.60	0.79	0.79	0.82
1,1,1-trichloroethane	4.38	4.38	4.38	0.47	0.40	0.49	0.49	0.51
1,1,2-trichloroethane	4.38	4.36	4.34	0.70	0.61	0.77	0.79	0.84

[a] The product of ρV is dimensionless, the dipole moment of the solute is in unit of eÅ.

4.3.4. *Unperturbed QM Solute Model: PT1-RISM*

In a second step towards self-consistency, the point charge representation of the solute charge density is replaced by the QM charge density of the solute in the gas phase. The corresponding solvent distribution functions are solutions of the 3D RISM equations with the electrostatic field of the unperturbed solute in the gas phase. This new electrostatic field replaces that of point charges when defining the interaction potential $u_\alpha(\vec{r})$ of the "external" field in the closure relations; see the discussions in Section 3.1.2. From the matrix of correlation coefficients (Table 4.3.2) one notices that the PT1-RISM and MM-RISM solvation energies are strongly correlated with $r = 0.98$. The corresponding moderate MAD = 0.95 kcal/mol (Table 4.3.3) is another indication for the similarity between these computational approaches. Therefore, it is also not surprising to find again a weak correlation between uncorrected PT1-RISM and experimental results which yields $r = 0.33$ and a MAD of 17.6 kcal/mol (Tables 4.3.2, 4.3.3). When the PMV correction is introduced, the correlation again significantly improves and reaches to $r = 0.88$ and MAD = 1.40 kcal/mol. The PMV values differ by no more than 5% between PT1 and MM. Therefore, the cross correlation of the PMV-corrected MM and PT1 solvation energies is nearly as strong as that of the uncorrected variants with $r = 0.90$ and MAD = 0.94 kcal/mol (Tables 4.3.2, 4.3.3).

Table 4.3.5. Parameters of linear regression, $y \approx Ax + B$, between the calculated solute partial molar volume measure $x = \rho V_u$ and the solvation free energy difference $y = \Delta^2 G$ between computed and experimental results for four variants of 3D RISM solvation models, without and with PMV correction.[a]

	uncorr.				corr.	
Model	A	B	R^2	MAD	A'	R^2
MM	4.88	-1.22	0.97	0.73	0.66	0.37
PT1	5.12	-2.49	0.97	0.72	0.91	0.51
PT2	5.09	-3.59	0.94	1.03	0.87	0.33
SCF	5.13	-3.97	0.92	1.24	0.92	0.27

[a] Slope A, $A'=A-a$, intercept B, and mean average deviation (MAD) in kcal/mol. Intercept B and MAD are the same for the uncorrected and the corrected regression models.

Inspection of the errors in the solvation energy from PT1-RISM calculations in Figure 4.3.2b suggests a very similar correlation between PMV and the MM-RISM approach. The error of the uncorrected solvation free energy correlates linearly with the dimensionless PMV with a slope of $A = 5.12$ kcal/mol, intercept $B = -2.49$ kcal/mol, and $R^2 = 0.97$ (Table 4.3.4). By applying the PMV correction with $a = -4.22$ kcal/mol, most of the linear correlation is

removed and the corresponding measures are reduced to A' = 0.91 kcal/mol and R^2 = 0.51 (Table 4.3.4). A slight growth of the remaining error with the system size (hence measured by PMV) can be detected from Figure 4.3.2b despite the scattering due to other factors affecting the error. Nevertheless, the error in the formation energy of the solute cavity appears to constitute the largest part of the deviation of the solvation free energy from experiment for the neutral solutes examined here.

Some difference between the MM and PT1 approaches is nevertheless to be expected. As the electrostatic field derived from the atomic charges of the force-field representation represents a relatively crude approximation to the field of the charge distribution of the more sophisticated QM model. Figure 4.3.3a provides evidence for this suggestion as it shows the relation of the solute dipole moments of the MM and PT1 solute models. The correlation is far from ideal with the slope of the linear fit equal to 0.84 and R^2 = 0.68. Apparently, there is no clear trend regarding the difference between the MM and PT1 approaches. Despite the differences of basic electrostatic solute properties in the MM and PT1 approaches, it can be concluded that the solvation free energies calculated using these two methods are at the similar levels of accuracy.

4.3.5. *Perturbed QM Solute Model: PT2-RISM*

In both the MM-RISM and PT1-RISM approaches the charge distribution of the solute does not adapt to the solvent medium. In either case, the solute is embedded as static species in the solvent environment. Then the perturbation of the solute electronic structure due to the reaction field of the solvent medium is introduced. However, the solvent field taken from a MM-RISM calculation is fixed during the electronic structure calculations. The polarization energy $\delta E \geq 0$ of the QM subsystem is then used to estimate the relaxation of the excess chemical potential $\delta\mu \approx -2\delta E$ and to compute the second-order free energy contribution as in Eq. (3.3.17).

For the systems under study, the polarization due to the reaction field affects the solute electronic energies only moderately with $\delta E \leq 2.1$ kcal/mol. Because of the small scale of the second-order correction, solvation free energies from PT2-RISM calculations are strongly correlated with the MM-RISM and PT1-RISM results, with r = 0.98 and MAD= 1.68 kcal/mol as well as r > 0.99 and MAD = 1.22 kcal/mol, respectively (Tables 4.3.2, 4.3.3; without PMV correction). Similar as the findings obtained with MM and PT1 methods, the PT2 results without PMV correction do not compare well with the experimental results with r = 0.4 and a corresponding MAD of 16.4 kcal/mol (Table 4.3.3). The correlation with the experimental measurements improves to r = 0.89 and MAD = 1.27 kcal/mol (Tables 4.3.2, 4.3.3) when the

PMV correction is applied. These estimates are comparable to those of the MM and PT1 approaches and better than in the case of the PCM results. Although the magnitude of the second-order effect is comparable to the accuracy of the MM and PT1 results, its addition does not improve the overall accuracy. This supports the interpretation that other error sources dominate the PMV-corrected RISM results.

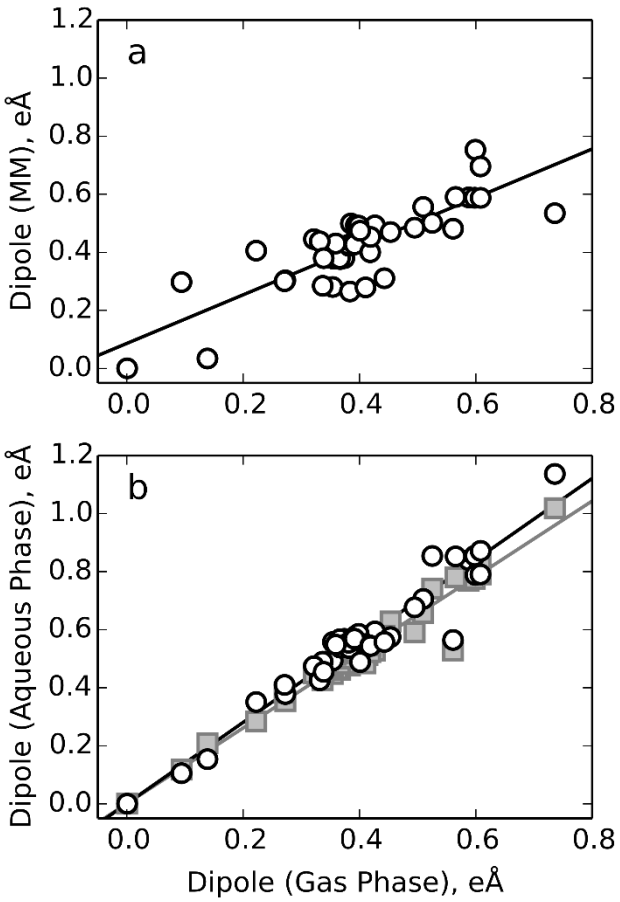


Figure 4.3.3. Comparison of dipole moment of solutes: a) MM vs. gas phase QM (with a linear regression line $y = 0.84x + 0.086$), b) aqueous phase vs. gas phase for SCF-RISM and PCM solvent models. In b), gray square represent PCM results (with a linear regression line $y = 1.30x + 0.001$), and white solid cycles represent SCF-RISM results (with a linear regression line $y = 1.40x + 0.001$).

The effect of the second-order correction can be examined in Figure 4.3.3c. The scattering around the linear regression line slightly increases compared to MM and PT1, which is confirmed by a slightly larger MAD from Table 4.3.3. However, on average the PMV-corrected results are closer to the base line in Figure 4.3.3c because of the net negative shift due to the second-order relaxation term. As discussed above this base line reflects a marginally better comparison to experiment. The slope of the regression line for the corrected PT2 results is very similar to that of the PT1 results, while the intercept of $B = –3.59$ kcal/mol

is lower by ~2.1 kcal/mol because of the overall negative shift.

4.3.6. *Self-consistent QM Solute Model: SCF-RISM*

Finally, the relaxation of the solute electron distribution upon self-consistent treatment of the solvent distribution is discussed. In contrast to the variants elaborated above, the solvent rearrangement is now taken into account. The error in the energy for forming the solute cavity leads to a poor correlation with the experiment yielding r = 0.41 and MAD = 16.1 kcal/mol (Tables 4.3.2, 4.3.3) for the SCF-RISM solvation energies without PMV corrections. The approximations considered in the previous sections strongly correlate with the self-consistent approach. The correlation coefficients are r = 0.97, 0.99, and 1.00 for the MM, PT1, and PT2 methods, respectively (Table 4.3.2). The mean absolute deviations of self-consistent and approximate results differ notably and amount to 2.01 kcal/mol, 1.52 kcal/mol, and 0.37 kcal/mol for MM, PT1, and PT2, respectively (Table 4.3.3). Note that the rough estimate of the second-order correction brings the PT2 data closer to the SCF results than the corrections due to the two first-order approximations.

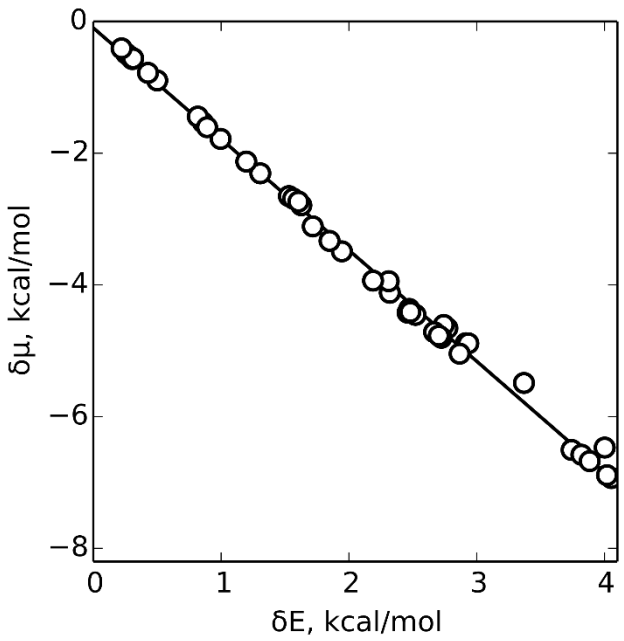


Figure 4.3.4. Correlation of two contributions to the relaxation energy $\delta G = \delta E + \delta\mu$ from the self-consistent treatment of the electronic charge distribution and the solvent charge density, with a linear regression line $\delta\mu = -1.69\ \delta E - 0.10$.

With the SCF results the assumptions of the previous section can be verified now. For nearly all the solutes shown in Figure 4.3.1, the self-consistent approach yields larger solute polarization energies δE than the PT2 estimates, which is reflected by a wider range of

$0 \leq \delta E \leq 4.1$ kcal/mol (Figure 4.3.4). The relaxation of the solvent chemical potential $\delta\mu$ is indeed strongly correlated with the polarization energy δE. The ratio of the two contributions to the relaxation energy estimated from a linear regression is about −1.69, see Figure 4.3.4. This value thus falls between the limits of −1 and −2 derived from the simple linear model (Section 3.3.1). The mean signed deviation (MSD) of PMV-corrected data shows that the on average PT2 model underestimates experiment by −0.16 kcal/mol while the average underestimation from SCF results is −0.39 kcal/mol. This indicates that the net effect of the self-consistent relaxation is slightly more negative than the second-order correction.

The PMV correction is hardly affected by a redistribution of the solute charge density in response to the solvent field. The effect is below 2% for the 43 solutes studied in this section. Therefore, the PMV correction of solvation energies has almost the same effect as for the other three approximations. In consequence, the PMV-corrected SCF-RISM results agree reasonably well with the experiment; $r = 0.86$ and MAD = 1.51 kcal/mol (Tables 4.3.2, 4.3.3).

Judged by these metrics, the SCF-RISM method together with the PMV correction does not improve upon PT2 or even PT1 (Tables 4.3.2, 4.3.3). A possible reason may be the overestimation of the solute polarization due to the volume charge of the solvent medium, because the repulsive exchange interaction between solute and solvent electrons is neglected in the SCF-RISM model. An unphysical distortion of the electron cloud by a nearby positive charge is known to cause artifacts in hybrid models[175] while omission of the exchange repulsion in negatively charged model atoms is less critical.[176] The solvation energies calculated with PT2 and SCF, hence method variants that account for electron polarization are on average slightly too negative than the experimental results. This is especially pronounced for the self-consistent variant. Figure 4.3.3b compares the dipole moments of the solute molecules in an aqueous environment and in the gas phase. It can be seen that the dipole moment increases by about 40% on average in SCF-RISM and by only 30% in PCM calculations. Note that the latter method also does not explicitly account for exchange repulsion.

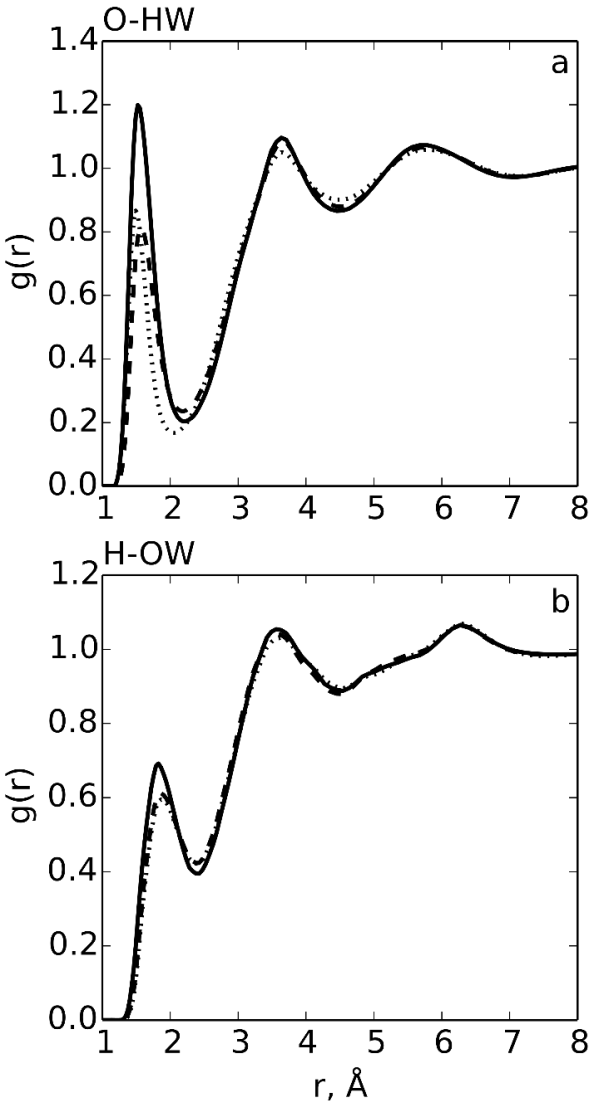


Figure 4.3.5. Radial distribution functions of solvated acetic acid: a) O–H_W (carbonyl oxygen – water hydrogen) and b) H–O_W (hydroxyl hydrogen – water oxygen). Results obtained with the MM, PT1 and SCF methods are shown as dotted line, dashed line and solid line, respectively.

This discrepancy of ~10% in the dipole moment between SCF-RISM and PCM calculations may be another indication of an overestimated solute polarization in SCF-RISM. Related evidence can also be obtained by inspecting radial distribution functions (RDFs). Figure 4.3.5 shows the RDFs of the pairs O–HW (carbonyl oxygen – water hydrogen) and H–OW (hydroxyl hydrogen – water oxygen) of solvated acetic acid. Comparing the results obtained from the MM, PT1, and SCF methods, respectively, one notices that the peak of positively charged H_W around O is significantly increased, by ~50%, when the electronic structure of the solute is relaxed self-consistently in the solvent field, while that of the negatively charged O_W around H increases by less than 10%. These findings agree with results

of Du and Wei[47] who observed an increase of water hydrogen and oxygen distributions around trans-*N*-methylamine from cycle 0 to cycle 5 during QM/3D RISM-HNC SCF iterations. Similarly, an earlier SCF-RISM study by Kawata et al.[54] also confirmed an increasing density of water hydrogen covalently bonded to the nitrogen center of the amine moiety when the electrostatic potential is switched on in a self-consistent treatment.

The plot of RISM solvation energies in Figure 4.3.2d is similar to that of PT2-RISM in Figure 4.4.1c. The magnitude of the residual error, which is unrelated to the PMV factor, is slightly larger for SCF-RISM than for PT2-RISM. This is confirmed by the larger MAD from the regression line (Table 4.3.4). The correlation of the SCF-RISM error with PMV is still obvious with $R^2 = 0.92$ and the fitted slope of $A = 5.13$ kcal/mol is close to the PT2 and PT1 results (Table 4.3.4). The PMV correction reduces these measures to $R^2 = 0.27$ and $A' = 0.92$ kcal/mol. The net negative shift due to the self-consistent solute density relaxation is slightly larger than for PT2 and the intercept $B = -3.97$ kcal/mol is the most negative among the regression lines of all considered methods.

4.3.7. *Concluding Remarks*

This application combined a QM description for the solute with a 3D RISM treatment of the solvent medium by a series of approximations as well as a fully self-consistent hybrid approach. The results of these treatments reveal that the accuracy of the calculated solvation free energies is mainly determined by the asymptotic error of the cavity formation energies for neutral solutes. Significantly better agreement with experimental results is obtained in previous non-hybrid studies[61,63,64] by applying the "partial molar volume" (PMV) correction. On empirical grounds the certainly useful PMV correction was justified,[64] although the thermodynamic arguments were disputed later.[63,104] Recently, a different pre-factor for PMV correction term was advocated.[61] This application presents an asymptotic analysis that offers an alternative interpretation of the origin of the error. The analysis leads to the same pre-factor as used in Ref. 61 for simple liquids, but does not rely on the thermodynamic arguments of Ref. 63.

When comparing approximate methods it is found that the representation of the solute charge density by point charges of a force field or by the gas-phase QM electron density leads to quite comparable solvation energies. The mean absolute deviations of these variants from each other amount to ~1 kcal/mol while these results differ by 1.4–1.6 kcal/mol from the experimental measurements. An approximate way to estimate the electron relaxation caused by the solvent reaction field and the resulting solvent relaxation energy are also examined. On average this approximate approach and the fully self-consistent 3D RISM QM method agree

to better than 0.4 kcal/mol. The ratio of the solute polarization and the solvent relaxation energies is nearly constant for all solutes studied, but it differs notably from the linear-response electrostatic limit of −2 (Section 3.3.1).

Relaxation terms have been suggested to be important for neutral molecules with overall rather small solvation energies.[6,113] The magnitude of the relaxation terms determined in the present context seems to support this argument. Nevertheless, using the approximate second-order (PT2) method to address the relaxation terms reduces the deviation from the experimental reference only slightly, while a self-consistent approach does not have any further essential effect. This latter finding may be due to an artificial polarization of the electron density by nearby localized positive volume charges of the force-field representation of the solvent medium. That effect likely is overestimated in the absence of repulsive exchange interactions between solute and solvent electrons. The slight overestimation of the induced solute dipole moments in comparison to the PCM approach is indicative for this artifact. An increased density of hydrogen and oxygen centers of water molecules around the solute in the SCF-RISM also supports this hypothesis. Alternatively, an asymptotic correction limited to a term which is strictly proportional to the solute volume may not suffice for reducing the intrinsic error of 3D RISM for neutral solutes in water to less than 1–2 kcal/mol. A careful analysis of the asymptotic behavior of 3D RISM may offer a more accurate error estimate and resolve some of the remaining issues.

4.4. Uranyl Solvation by 1D and 3D RISM

4.4.1. *Introduction*

Understanding the thermodynamic and structural properties of uranium in aqueous solution is crucial for modeling processes during geological radioactive waste disposal and recycling of used nuclear fuel.[177,178] Thus, both experimental and theoretical studies have been carried out to elucidate structural and energetic information of the uranyl(VI) cation UO_2^{2+} which is the most common stable oxidation state of this species in aqueous solution.[179,180] While in experimental studies of uranyl cations researchers have to face the radiotoxicity of uranium, theoretical methods, including QM and classical MM simulations, provide alternative ways for studying the solvation of such systems with considerable accuracy. Given the difficulties of the popular PCM and MD methods discussed in Sections 2.1 and 2.2, RISM approaches are regarded as promising direction for studying ions and complexes as well as their chemical reactions in aqueous solution with sufficient accuracy at affordable costs. Therefore, in this section RISM approaches including both 1D and 3D RISM techniques were applied to the solvation of uranyl(VI) in aqueous solution. While the 1D RISM calculations were conducted in the context of this thesis, the discussion related to 3D RISM is based on the results reported by A.V. Matveev et al..[106]

4.4.2. *Calculation of Uranyl Solvation Energy*

Due to the complex models used in this application, the free energy functional in terms of G and E described in Section 3.3 needs further refinement to yield the correct solvation energy of uranyl in water. Here the solvation energy of uranyl is estimated as

$$\Delta G_{\mathrm{sol}} = G[\mathrm{UO_2(H_2O)}_n{}^{2+}] - nG[\mathrm{H_2O}] - E[\mathrm{UO_2}^{2+}] \tag{4.4.1}$$

for $n \geq 0$ explicit aqua ligands. Without proper statistical sampling over the solute degrees of freedom this approximation of the solvation energy, while being justified to some extent for sufficiently rigid solutes, does not hold for systems with a large number $n \to \infty$ of explicit water molecules outside of the tightly bound first solvation shell. Indeed, in a hypothetical experiment the difference $G[(\mathrm{H_2O})_n] - nG[\mathrm{H_2O}]$ for the above definition of G applied to water droplets asymptotically approaches $n(\bar{E}[\mathrm{H_2O}] - G[\mathrm{H_2O}])$, where $\bar{E}$ is the average configurational energy of a water molecule. The value of $G[(\mathrm{H_2O})_n] - nG[\mathrm{H_2O}]$ is about −9.9 kcal/mol.[78,181] This value is to be compared to the solvation energy of −6.3 kcal/mol

derived from the ratio of the equilibrium vapor/liquid concentrations.[182] The intramolecular energy ΔE_{intra} of the solute model complex is obtained by substituting *E* for *G* in Eq. (4.4.1).

4.4.3. *Comparison of Water Models*

As a first test, well-established[124] and newly developed[125-127] empirical force field parameterizations for uranyl as well as a suitable water model, solvent structures around the *bare* uranyl ion were inspected with the 1D RISM method and compared with results of MD simulations. In this section, the PM13 uranyl model as shown in Table 4.1.1 was chosen to explore the performance of the three water models that differ only by the LJ parameters of the hydrogen sites, Table 4.1.2. Based on the parameters used in the classical MD calculation of Maginn et al.,[125] the radial distribution functions (RDFs) of the SPC/E water was reproduced for comparison using Gromacs 3.[183] The RDFs of the oxygen and hydrogen sites of water molecules, O_W and H_W, around the uranium center and the oxygen sites of the uranyl ion, U and O, are shown in Figure 4.4.1.

When comparing the 1D RISM results with those from MD calculations one notices that the RDFs for site pairs carrying charges of opposite sign, i.e. for the pairs U–O_W and O–H_W, show quite similar shapes although the double peak feature of the latter is not resolved in the RISM RDF. For the pairs U–H_W and O–O_W which carry charges of the same sign, the differences between MD and 1D RISM models are more pronounced and shared by all three water models. From the analysis of the RDF plots it can be concluded that the three variations of the SPC/E water model differ slightly in the intensity of the peaks, while the positions do not vary much. The positions of the first peaks of the U–O_W radial distribution function are between 241 pm and 243 pm for the various water models (Table 4.4.1) while the corresponding coordination numbers vary from 4.9 to 5.8. The MD reference results provide a peak at a slightly longer distance, 246 pm, that integrates to a coordination number of 5.0. The excess chemical potential, calculated between −309 kcal/mol and −293 kcal/mol, is somewhat higher than the solvation free energy of −332 kcal/mol determined with the MD approach (Table 4.4.1). The observation that the 1D RISM model underestimates the solvation free energy of uranyl ion by 9–13% is not surprising, given the results and discussions for simple ions, Section 4.2.

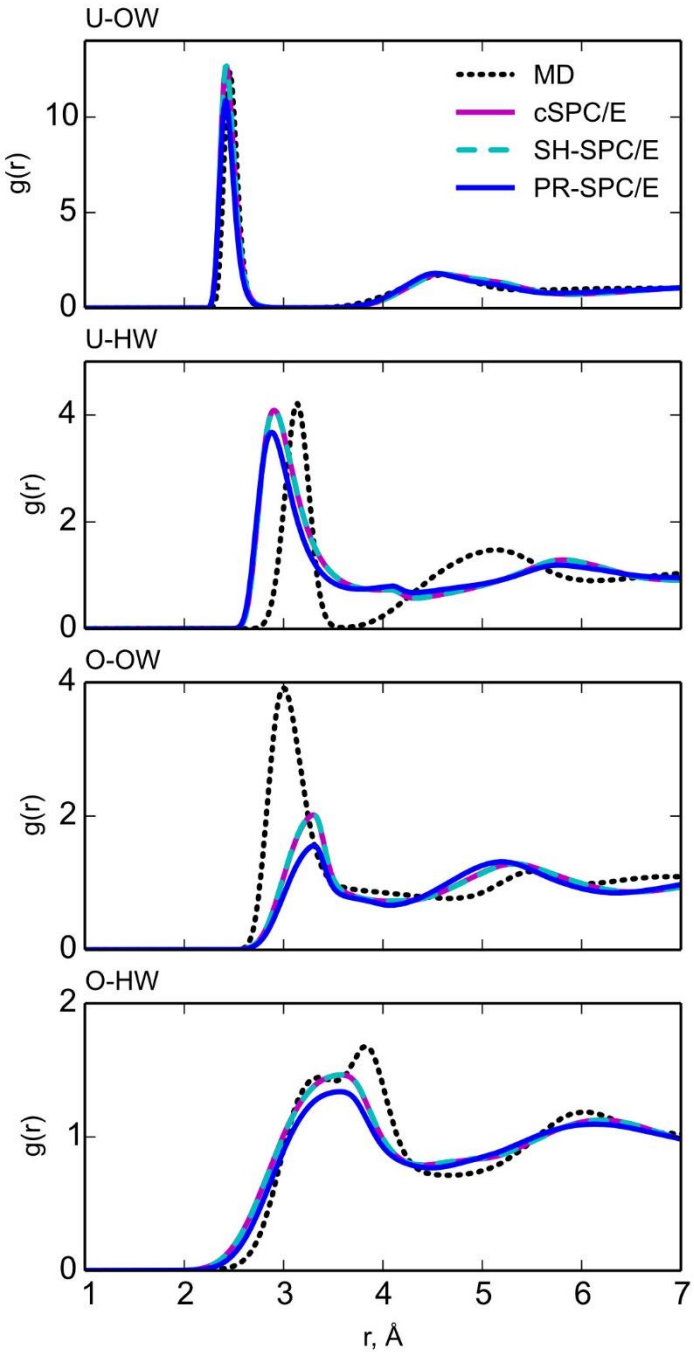


Figure 4.4.1. Radial distribution function of solvated uranyl, calculated with three modifications of the SPC/E water model by using 1D RISM method. SPC/E MD results are shown for comparison.

Table 4.4.1. RISM results for solvated uranyl described by the PM13 force field and various water models: coordination number N of the first solvation shell, position of the first peak d_1(U–O$_W$) (pm), and solvation free energy ΔG_{sol} (kcal/mol).

Water model	N	d_1(U–O$_W$)	ΔG_{sol}
SH-SPC/E	5.8	243	–309
PR-SPC/E	4.9	241	–293
cSPC/E	5.8	243	–308
SPC/E MD[a]	5.0	246	–332

[a] Ref. 125.

4.4.4. *Comparison of Uranyl Models in 1D RISM*

To compare the performance of 1D RISM for the five uranyl models listed in Table 4.1.1, the PR-SPC/E water model (Table 4.1.2) was chosen as it yields a coordination number of 4.9 which is closest to 5 among all water models studied (Table 4.4.1). As can be seen from Figure 4.4.2 and Table 4.4.2, the coordination numbers 4.9–5.2 as determined by GW and Maginn's models, RM12 and PM13, deviate from the results of *ab-initio* MD (AIMD)[184,185] quantum mechanical charge field MD (QMCF-MD)[186] and MM-MD.[126,187] The latter three methods predict coordination number 5, while the range 4–5 is spanned by experiments of X-ray scattering,[188] extended X-ray Absorption Fine Structure (EXAFS),[189] and high-energy X-ray scattering (HEXS).[190] The positions of the first U–O$_W$ peaks calculated with RM12 and RM13 vary from 241 pm to 242 pm. These results are close to those of MM-MD calculations[126,187] and experimental values,[188-190] while the position of 248 pm obtained with the GW model falls in the range of 246–249 pm of AIMD and QMCF-MD calculations.[184-186] The variations between different uranyl models are not surprising as they were fitted to different experimental solvation energies (see below). However, RISM calculations applying these models generally show the solvent structure around uranyl ion to be comparable with the results of experiments and explicit solvation model simulations.

The calculated solvation free energies are slightly below the experimental range determined with Fourier transform ion cyclotron resonance mass spectrometer (FTICR–MS).[191] On the other hand, Kerisit and Liu intentionally increased the charge separation in the uranyl force field to yield solvation free energies of −375 kcal/mol and −380 kcal/mol, which fall right into the experimental range.[127] A larger charge separation in the models KL1 and KL2 leads to stronger uranyl-water interactions which become manifest in shorter U–OW bonds and higher RDF peaks (Figure 4.4.2) as well as shell coordination numbers close to 6

(Table 4.4.2). The original MD simulation employing the models KL1 and KL2 for uranyl predicted the first U–O_W peak with a height of about 16 and a coordination number of 5.[127] The U–O_W distance of the first RDF peak in the 1D RISM calculations is somewhat shorter than in the MD simulation: 232 pm for KL1 RISM and 238 pm for KL2 RISM (Table 4.4.2) as compared to 236 pm and 240 pm, respectively, in the MD calculations.[127] Despite these geometric differences of RISM and MD, the free energy of solvation is well reproduced. MD predicted −304 kcal/mol, −371 kcal/mol, and −377 kcal/mol for the models GW, KL1, and KL2, respectively,[127] the corresponding results of the present RISM approach are −296 kcal/mol, −375 kcal/mol, and −380 kcal/mol, Table 4.4.2.

Table 4.4.2. Calculated (1D RISM, MD) and experimental results for uranyl (VI) UO_2^{2+}**,** solvated in water: coordination number *N* of the first solvation shell, positions (pm) of the first, d_1 (U–O_W), and second, d_2 (U–O_W), peaks of the radial distribution function, solvation free energy ΔG_{sol} (kcal/mol).

	Method		*N*	d_1 (U–O_W)	d_2 (U–O_W)	ΔG_{sol}
Calc.	**RISM**	**GW**	5.2	248	455	−296
		RM12	5.1	241	448	−305
		PM13	4.9	241	452	−293
		KL1	5.8	232	451	−375
		KL2	6.1	238	452	−380
	AIMD[a]		5.0	246	459	
	AIMD[b]		5.0	248	460	
	QMCF–MD[c]		5.0	249	430–500	
	MM–MD[d]		5.0	242	450	
	MM–MD[e]		5.0	241–245	460	
Exp.	**X-ray scat.**[f]		4.9	242	446	
	EXAFS[g]		4.5 ± 0.5	241		
	HEXS[h]		4.61	242	450	
	FTICR–MS[i]					−293 ± 5
	FTICR–MS[k]					−369 ± 15
	FTICR–MS[l]					−437 ÷ −318

[a] Ref. 185. [b] Ref. 184. [c] Ref. 186. [d] Ref. 187. [e] Ref. 126. [f] Ref. 188. [g] Ref. 189. [h] Ref. 190. [i] Calculated from $\Delta_{hyd}H^{\circ}_{abs}$ = −1345 ± 20 kJ/mol and $\Delta_{hyd}S^{\circ}_{abs}$ = −399 ± 5 J/(mol K) as in Ref. 192. [k] Calculated from ΔH_{hyd} = −1665 ± 64 kJ/mol and ΔS_{hyd} = −400 J/(mol K) as in Ref. 193. [l] Calculated from ΔH°_{f} = −371 ± 60 kcal/mol for uranyl in the gas phase, from Ref. 191, using $\Delta_h S^{\circ}$ = −78.6 cal/(mol K) following Ref. 131.

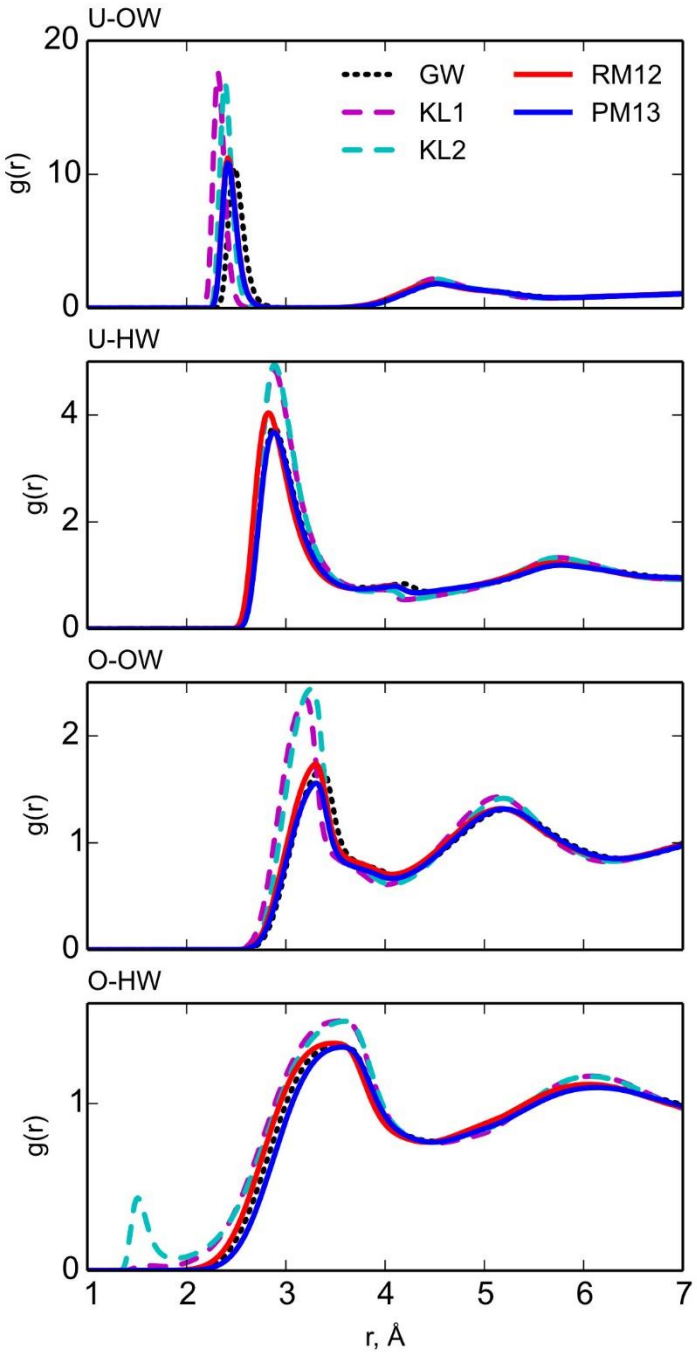


Figure 4.4.2. Radial distribution function from 1D RISM calculations applying various force field models for uranyl.

In an AIMD calculation[185] and in the classical MD simulations of Keresit und Liu with KL1 and KL2[127] there is a small peak at 196 pm in the RDF of the O–H_W pair that corresponds to water molecules which form a hydrogen bond to the terminal oxygen centers of uranyl. The results of the RISM calculation with the KL2 uranyl model also show the same small peak, albeit at a shorter O–H_W separation of 160 pm (Figure 4.4.2). This feature in the RDF was already observed in the early work[187] when the site charge on the uranium atom was increased from 3 *e* to 6 *e*. Hagberg et al.[194] also stated that in some configurations during their MD

simulations, there exists a weak hydrogen bond of 200–250 pm in length to the uranyl oxygen. In the AIMD simulations, it was observed that such close contacts between uranyl oxygen and water with O–H_W distances below 200 pm are rarely maintained for longer than a few hundred femtoseconds.[184] Thus, a corresponding peak is absent in the AIMD RDF.

Based on the values of solvation free energies of a bare uranyl solute model, which does not include explicit aqua ligands in the definition of the solute complex, the uranyl models can be roughly divided into two groups regarding to the strength of the uranyl-water interactions. The weaker FF group (GW, RM12, PM13) and the stronger FF group (KL1, KL2). The latter two uranyl models yield solvation free energies which lie close to the center of the interval of the experimental results, Table 4.4.2. However, this improvement of the solvation free energy comes at the expense of higher coordination numbers which exceed the experimental references, Table 4.4.2.

One possible reason for the existence of these two classes of models is that the earlier force fields were parameterized on older and less negative experimental estimates, $\Delta G_{sol} = -293 \pm 5$ kcal/mol.[192] The largest uncertainty of the uranyl solvation energy, -437 kcal/mol $\leq \Delta G_{sol} \leq -318$ kcal/mol,[131] as well as of the other ΔG_{sol} entries in Table 4.3.2 is due to the experimental value of the heat of formation of uranyl in the gas phase, $\Delta H_f^\circ = -371 \pm 60$ kcal/mol.[191] A more recent work provides a narrower interval for that value, $\Delta H_f^\circ = -364 \pm 15$ kcal/mol,[193] which is close to the estimate, $\Delta H_f^\circ = -370 \pm 12$ kcal/mol, obtained by comparing relativistic DFT results with experimental enthalpies of model reactions.[144,145]

4.4.5. *Free Energy Minimization for Uranyl Aqua Complexes Using 1D RISM*

More insight into the structure of the first solvation shell and details of the uranyl-water interactions can be gained by adding explicit water ligands to the solute model embedded in a solvent model described by 1D RISM. Hence the application turned to the *standard model* of uranyl solvation in water where the solute is described as aqua complexes $[UO_2(H_2O)_n]^{2+}$ with n = 4–6, i.e. the first solvation shell of uranyl is represented by explicitly treated water molecules in equatorial positions.[69,131,145,184,195-197] For this *standard model* free energy surfaces of uranyl-water complexes in the presence of a 1D RISM water solvent environment outside the first solvation shell were computed both with MM and QM methods. The uranyl geometry is either fixed (MM, MM+1D RISM) or not (QM, QM+1D RISM, see below) while the positions and orientations of the water ligands are always optimized. The results of the MM calculations will be discussed first followed by those of QM treatments of the solute, embedded in water treated at the 1D RISM level.

MM+1D RISM

Table 4.4.3 lists the intra-molecular energy ΔE_{intra} of the solute and the solvation free energy ΔG_{sol} of uranyl, as defined in Section 4.4.2. Both of them were determined for the solvated aqua complex $[UO_2(H_2O)_5]^{2+}$, using MM+RISM with various uranyl force fields. These energies confirm the classification of the uranyl force fields into two groups. The three weaker force fields (GW, RM12, PM13) suggest the intra-molecular energy between uranyl and its five first shell aqua ligands to lie in the range of −216 kcal/mol ≤ ΔE_{intra} ≤ −210 kcal/mol, while the two stronger force fields suggest notably larger ligand binding energies, −268 kcal/mol (KL1) and −271 kcal/mol (KL2). The solvation free energies correlate with the strength of the uranyl-water interactions, Table 4.4.4. A comparison of the corresponding rows in Tables 4.4.3 and 4.4.4 also shows that the free energies based on the model with five explicit water ligands are by 10–15 kcal/mol lower than those of a bare uranyl model. A part of this difference is attributed to the missing statistical average over various accessible arrangements of the ligands.

Table 4.4.3. Properties of solvated uranyl from various force-field models: average uranyl–water distance, d(U–O$_W$), intra-molecular interaction energy, ΔE_{intra}, between a uranyl ion and its water ligands, and solvation free energy, ΔG_{sol}, for $[UO_2(H_2O)_5]^{2+}$ calculated with MM+1D RISM method.[a]

Uranyl model	d(U–O$_W$)	ΔE_{intra}	ΔG_{sol}
GW	251	−210	−311
RM12	246	−216	−321
PM13	247	−212	−309
KL1	238	−268	−384
KL2	243	−271	−390

[a] Distance in pm, energies in kcal/mol.

Based on the results in Table 4.4.3 and the discussion in Section 4.4.4, the force field KL2 gave reasonable agreement for geometry parameters with experimental measurements and favorable solvation energies for the bare uranyl model and the aqua complex model $[UO_2(H_2O)_5]^{2+}$. Therefore, the further free energy minimizations of the uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$ (n = 4–6) was limited to the force field KL2.

Geometry optimizations of these complexes were carried out in the gas phase at the MM level and in solution at the MM+RISM level, both without symmetry constraints. The force field optimization in the gas phase led to nearly symmetric structures for n = 4 and 5, see Figure 4.4.3. In the case of n = 6 a single water ligand is moved to a position outside the first

solvation shell: $d(U–O_W)$ = 411 pm. The RISM solvent environment was also found to compensate the fifth water molecule in the undercoordinated uranyl aqua complex $[UO_2(H_2O)_4]^{2+}$, which results in the asymmetric structure depicted in Figure 4.4.3a. This compensation is also reflected by the average $U–O_W$ distance in the first solvation shell. This quantity, 239 pm in the gas phase for $[UO_2(H_2O)_4]^{2+}$, is shorter than that of 244 pm for $[UO_2(H_2O)_5]^{2+}$, while in solution it increases to 245 pm and compares well with the result for $[UO_2(H_2O)_5]^{2+}$ which amounts to 243 pm (Figure 4.4.3a). For $[UO_2(H_2O)_6]^{2+}$ a similar structure as in the gas phase is obtained, albeit with a considerably longer $U–O_W$ distance, 453 pm, to the second-shell aqua ligand. Free energy values in Table 4.4.4 show that embedding into the RISM solvent yields a solvation free energy that is essentially constant and thus independent of the number of explicit water ligands. This quantity is again more negative by about 10 kcal/mol than the solvation energy of the bare uranyl solvent model, Table 4.4.2.

Table 4.4.4. Calculated properties of solvated uranyl from molecular mechanics calculations using *n* explicit water ligands: average uranyl–water distance, $d(U–O_W)$, to the first solvation shell, intra-molecular energy, ΔE_{intra},[a] with and without 1D RISM solvent, and solvation free energy, ΔG_{sol}.[b]

	N	0	4[c]	5[c]	6[c]
MM	$d(U–O_W)$	–	239	244	244
	ΔE_{intra}	0	–232	–271	–294
MM+RISM	$d(U–W)$	–	245	243	244
	ΔE_{intra}	0	–217	–271	–275
	ΔG_{sol}	–380	–388	–390	–391

[a] Value corresponds to the binding energy of the water ligands in $[UO_2(H_2O)_n]^{2+}$, n = 4–6, but is identically zero for rigid bare uranyl, n = 0. [b] Distances in pm, energies in kcal/mol. [c] For the structure, see Figure 4.4.3.

QM+1D RISM

The results from calculations with gas phase, PCM, and 1D RISM models for bare uranyl as well as systems with 4–6 explicit aqua ligands are compared in Table 4.4.5. The uranyl-water interactions described by DFT in the QM, QM+PCM, and QM+1D RISM calculations in general are comparable to those obtained from the KL1 and KL2 force field models. The typical intra-molecular energy corresponding to the binding of the five ligands of the penta-aqua uranyl complex, including uranyl relaxation, amounts to about −270 kcal/mol, see Table 4.4.5. On the other hand, inspection of the attractive branch of the solute-solvent interaction

potential as a function of $d(U–O_W)$ reveals that it is less steep when calculated with a QM method than with the KL2 force field. This inconsistency may partially rationalize why during some free energy minimizations with the QM+1D RISM approach without symmetry constraints explicit water molecules were found to be replaced by the 1D RISM medium. Hence, when comparing QM, QM+PCM, and QM+1D RISM models, symmetric conformations of the solute complexes were applied to avoid the substitution effect described above. The point group symmetries were assigned as D_{4h} for $n = 4$, D_{5h} for $n = 5$, and D_{3d} for $n = 6$.

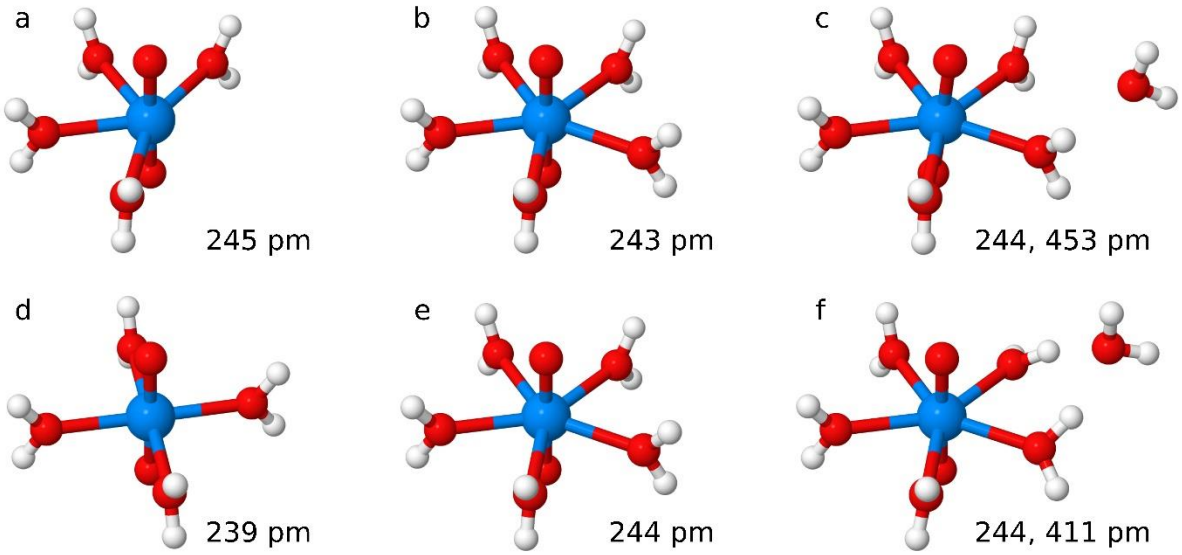


Figure 4.4.3. Optimized structure of selected local minima of the uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$ $n = 4–6$, with (first row) and without (second row) embedding in an aqueous solvent medium described by 1D RISM method. $U–O_W$ distances (pm) are averaged over the solvation shell.

Table 4.4.5 compares the results of QM, QM+PCM, and QM+1D RISM calculations for such symmetric uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$ ($n = 4–6$) with those obtained for a bare uranyl ion as solute ($n = 0$). In all cases the intra-molecular energies ΔE_{intra} with explicit water molecules were found almost independent of the method used to optimize the geometry of the solvated complex, as indicated by their deviations $\left|\Delta^2 E_{\text{intra}}\right| \leq 2.2$ kcal/mol to the analogous quantities for the gas phase. In the case of a bare uranyl solute, ΔE_{intra} corresponds to the relaxation energy of uranyl itself. Its change of 6.7 kcal/mol upon solvation as described by the 1D RISM approach is larger than for the other complexes. In fact, the U–O bond increases from 172 pm in the gas phase to 180 pm in solution, see Table 4.4.5. The explicit modeling of the first solvation shell explicitly hardly affects this distance. The terminal distance U–O of 181 pm is therefore only 1 pm longer for all $n = 4–6$ than for the bare uranyl

model. The results of the QM+PCM approach are rather different. Here the relaxation of bare uranyl due to the medium is much smaller. The energy change is 0.4 kcal/mol with a corresponding bond elongation of only ~2 pm. The U–O bond reaches values in the range of 179–180 pm only when explicit water molecules are introduced in the first solvation shell.

Another qualitative difference between the QM+PCM and QM+1D RISM results is the consistent shortening of the U–O_W bonds by 4–6 pm with the QM+PCM approach in comparison to the gas phase QM results. This solvent effect is not confirmed by the QM+1D RISM results. In this latter case the changes of d(U–O_W) due to solvation are found between −1 pm and +4 pm, see Table 4.4.5. The relatively long U–O_W bonds of the 1D RISM models may tentatively be rationalized by an increased first shell coordination of uranyl due to a 1D RISM solvent contribution in addition to the explicit aqua ligands. This effect manifests itself as a peak in the U–O_W radial distribution function.

While most experimental studies reported the number of water ligands bound to UO_2^{2+} cation in aqueous solution to vary between 4 and 5,[188-190] most computational results suggest five-coordinated uranyl to be the dominating species.[132,133,195,197]

Neglecting the small differences due to relaxation, the binding energy ΔE_{intra} varies by about −26 kcal/mol when going from n = 4 to n = 5 and by another −19 kcal/mol when going to n = 6 (QM results). Therefore it is not useful to compare the stability of these conformations in solution. However, the solvation free energy of uranyl for various conformations of the first solvation shell shows quite a different trend when being estimated with complexes $[UO_2(H_2O)_n]^{2+}$ with n = 4–6. According to QM+PCM and QM+1D RISM calculations, the five-coordinated uranyl complex is the most stable conformation with ΔG_{sol} = −426 kcal/mol (QM+PCM) and ΔG_{sol} = −390 kcal/mol (QM+1D RISM). The latter value is close to the center of the experimental range of results, from −437 kcal/mol to −318 kcal/mol, while the former is closer to its lower bound. Apart from the difference in magnitude, there is also a qualitative difference between QM+PCM and QM+1D RISM results for the relative stability of the four-, five-, and six-coordinated conformations. The QM+PCM model predicts the four-coordinated and six-coordinated complexes to be less stable than the five-coordinated one by 10 kcal/mol and 1 kcal/mol, respectively (Table 4.4.5). The relative order is opposite at the QM+1D RISM level, the four-coordinated complex is only by about 1 kcal/mol less stable than the five-coordinated one while the six-coordinated complex lies by about 6 kcal/mol higher in energy, see Table 4.4.5. It is not surprising that the free solvation energies at the QM+1D RISM and the MM+1D RISM levels agree very well for comparable structures (Table 4.4.4 and 4.4.5), because the KL2 force field was parameterized to yield nearly the same intra-molecular energies as the results from QM calculations for the uranyl aqua

complex models.[127]

Table 4.4.5. Calculated properties of solvated uranyl from QM calculations, either pure or with 1D RISM solvent models: uranyl bond length, d(U–O), uranyl–water distance, d(U–O_W), intra-molecular energy, ΔE_{intra},[a] its changes due to the solvent environment, $\Delta^2 E_{intra}$, and the solvation free energy, ΔG_{sol}, of uranyl.[b]

	n	0	4[c]	5[c]	6[c]
QM	d(U–O)	172	177	177	178
	d(U–O_W)	–	242	249	253
	ΔE_{intra}	0	–244	–270	–289
QM+PCM	d(U–O)	174	179	179	180
	d(U–O_W)	–	237	243	249
	$\Delta^2 E_{intra}$	0.4	1.1	1.3	2.0
	ΔG_{sol}	–322	–416	–426	–425
QM+RISM	d(U–O)	180	181	181	181
	d(U–O_W)	–	246	248	254
	$\Delta^2 E_{intra}$	6.7	2.2	1.1	2.2
	ΔG_{sol}	–379	–388	–390	–384

[a] Value corresponds to the binding energy of the water ligands in $[UO_2(H_2O)_n]^{2+}$, n = 4–6, but is identically zero for rigid bare uranyl, n = 0. [b] Bond lengths in pm, energies in kcal/mol. Uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$ (n = 4–6), and a bare uranyl (n = 0) treated by QM while interactions with the PR-SPC/E water solvent are described by the KL2 force field. [c] Symmetry constraints (see text), D_{4h} for n = 4, D_{5h} for n = 5, and D_{3d} for n = 6.

While most quantities in QM+1D RISM calculations are rather independent of the number of aqua ligands, the U–O_W distances to the explicit aqua ligands increase from 246 pm (n = 4) to 254 pm (n = 6); see Table 4.4.5. These values of d(U–O_W) are larger than the results of QM+PCM calculations and the experimental results of 241–242 pm (Table 4.4.5). This overestimation is due to the effectively higher coordination of uranyl in the QM+1D RISM model. Besides the explicit aqua ligands, also 1D RISM water contributes to the first shell coordination of uranyl and these additional contributions lead to effective coordination numbers of 5.8 (n = 4), 6.2 (n = 5), and 6.9 (n = 6).

Table 4.4.6. Properties[a] of solvated uranyl, modeled as $[UO_2(H_2O)_5]^{2+}$ (with two exceptions[b,c]) calculated by combinations of QM and 1D RISM as well as various solvation models: uranyl bond length d(U– O), uranyl–water distance d(U–O$_W$), solvation energy ΔG_{sol} of uranyl, reaction energy ΔG_R for adding an explicit water ligand to the tetra-coordinated complex, Eq. (4.4.2). Experimental values shown for comparison.

Method	d(U–O)	d(U–O$_W$)	ΔG_{sol}	ΔG_R
BP+COSMO[d]	179	243	–426	–10.3
BP+RISM[d]	181	248	–390	–1.6
MP2+IEFPCM[e]	178	244–248	–464	–24.7
B3LYP+COSMO[f]	176	249		–11.1
PBE+COSMO[g]	178	247	–384[h]	–4.9
BLYP+CPMD[b]	178	250		–8.7
MP2+SCIPCM[i]	175	250	–412	–7.8
MP2+SCIPCM[c]	176	244–255	–411	–2.5
HEXS[j]	177	242		–1.2

[a] Distances in pm, energies in kcal/mol. [b] Car-Parinello molecular dynamics (CPMD) calculation on uranyl, UO_2^{2+}, in a water solution, Ref. 132. [c] Self-consistent isodensity polarizable continuum model (SCIPCM), Ref. 196; structure optimization at the B3LYP level on a model in the gas phase that includes the second solvation shell. [d] This work. [e] Integral equation formalism polarizable continuum model (IEFPCM), Ref. 133. [f] Small-core effective core potential, Ref. 195. [g] All-electron four-component scalar relativistic calculation, Ref. 195, except the value of ΔG_{sol}. [h] Single-point calculation at the all-electron zeroth order regular approximation (ZORA) level, Ref. 195. [i] Self-consistent isodensity polarizable continuum model (SCIPCM), Ref. 196; structure optimization at the B3LYP level on a model in the gas phase. [j] Experimental results, Ref. 190.

A useful quantity for determining the preferred coordination number is the reaction energy ΔG_R for the water exchange process between four- and five-coordinated uranyl:

$$[UO_2(H_2O)_4]^{2+} + H_2O \leftrightarrow [UO_2(H_2O)_5]^{2+}. \qquad (4.4.2)$$

Table 4.4.6 compares the results of various QM calculations of uranyl solvation to those of a high-energy x-ray scattering (HEXS) experiment.[190] A small free energy difference of −1.2 kcal/mol between four- and five-coordinated uranyl in aqueous solution has been determined in the HEXS experiment[190] by estimating the relative weights of four- and five-coordinated species, based on an average coordination number. Various computational results reveal that ΔG_R is commonly overestimated when determined by combinations of QM and continuum solvation models as well as by first-principles MD calculations (Table 4.4.6). Density functional and MP2 calculations combined with COSMO or SCIPCM

variants[131,133,195] yield ΔG_R values between −11.1 kcal/mol and –4.9 kcal/mol (Table 4.4.6), which are more negative than the experimental estimate. The strongest overestimation, ΔG_R = –24.7 kcal/mol, of the experimental value is determined with the MP2+IEFPCM approach due to the overestimation of hydrogen bonds in MP2 method.[133] A noteworthy better result, ΔG_R = −2.5 kcal/mol, has been achieved with the MP2+SCIPCM approach when explicitly accounting for the second solvation shell.[196] A Car−Parinello MD simulation also somewhat overestimated the preference for the five-coordinated species, ΔG_R = −8.7 kcal/mol.[132] Without significantly increasing the computational cost compared to those of QM calculations, the QM+1D RISM approach leads to a favorable result of ΔG_R = −1.6 kcal/mol. This value agrees best with the experimental result of ΔG_R = −1.2 kcal/mol. However, the quality of the present result should not be over-interpreted as it likely arises from a favorable method combination, i.e. the force fields, the exchange-correlation functional, and the 1D RISM approach.

Finally, the geometric parameters shown in Table 4.4.6 will be briefly discussed. The uranyl bond 181 pm of 1D RISM calculations is by 2.3% larger than the experimental value 177−178 pm[188-190] and also longer than the results obtained from the PCM solvation models, 178−179 pm (see above). The uranyl–water distance $d(U–O_W)$ = 248 pm as calculated with RISM is also longer than the experimental results of 241–242 pm.[188-190] However, this value agrees with the results calculated with other methods like BLYP and MP2 combined with PCM models (Table 4.4.6). In summary, when studying chemical processes in the solvent environment, the 1D RISM approach seems to provide an accuracy that is equal or superior to that of traditional continuum solvation models.

4.4.6. *Combination of 3D RISM with MM and QM*

As discussed in Section 4.4.5 there remain deficiencies when applying the 1D RISM technique to study the uranyl solvation problem. The uranyl-water coordination number is overestimated and the U–O and $U–O_w$ bonds are excessively elongated. In this section, 3D RISM is applied to study the uranyl solvation with the same models while aiming to resolve the aforementioned issues.

MM+3D RISM

The combination of 3D RISM for bulk water and a force field potential for the interactions of uranyl and *n* explicit water ligands was first examined for *n* = 4−6. For the optimized geometries there is a qualitative difference for the structures obtained by the MM and MM+3D RISM methods, of which the latter features a missing equatorial ligand for the complex with

n = 4. This finding is consistent with the results obtained by MM+1D RISM (Figure 4.4.3). The location of this missing water ligand corresponds to the vacancy of the solvent distribution in the equatorial plane of the uranyl. While the information of this vacancy is hard to be interpreted from RDFs averaged over the radial grid in 1D RISM, 3D RISM allows its direct observation from the spatial distribution of the solvent, see the water oxygen distribution plot in Ref. 106. A straightforward MM calculation yields a minimum at a D_{4h}-like structure with four water ligands evenly distributed in the equatorial plane, while structures with a vacancy in the equatorial plane are by 13−15 kcal/mol higher in energy. The average U−O_w bond of 239 pm in the MM-optimized structure is slightly shorter than the corresponding bonds of 240−242 pm for the structures optimized at the MM+3D RISM level.[106] The shorter distances indicate a weaker steric ligand−ligand repulsion in the *undercoordinated* complex in the gas phase. Note however, this effect is opposite in the five-coordinated uranyl complexes. U−O_w distances are shortened from 243−244 to 240 pm after embedding a complex with a structure optimized for the gas phase into the RISM medium.[106] This result may indicate a partial screening of the steric repulsion of the equatorial ligands of a *saturated* complex by the second solvation shell. This effect seems to be present also in PCM calculations (see below).

The mechanism by which the MM+3D RISM model compensates for the missing ligand is described as a local increase of the water density at the location of the missing explicit ligand, see Ref. 106. Yet, the 3D RISM representation of the missing ligand has to be considered as an approximate substitute for an explicit water ligand. This can be interpreted from the notable width of the local distribution both in equatorial and in apical directions but less so when being averaged over the volume in radial direction.[106] Indeed, the average distances U−O_w between the complexes with 4 and 5 explicit aqua ligands in the first solvation shell differ, 241−242 pm for n = 4 and 240 pm for n = 5.[106]

QM+3D RISM

Symmetric models of uranyl aqua complexes were used in the QM+3D RISM calculations to facilitate a comparison with previous results obtained with PCM and 1D RISM representations of the solvent. In these symmetric models the equatorial aqua ligands are evenly distributed. The symmetric structures of the five- and six-coordinated uranyl aqua complexes are qualitatively consistent with the structures obtained from an unconstrained optimization (see next section). In contrast, the unconstrained four-coordinated species features a vacancy in the first solvation shell which is filled by the 3D RISM solvent. These QM models feature the uranyl bond length as an additional intramolecular degree of freedom, unlike the MM models discussed in the previous section where this distance was kept

constant.

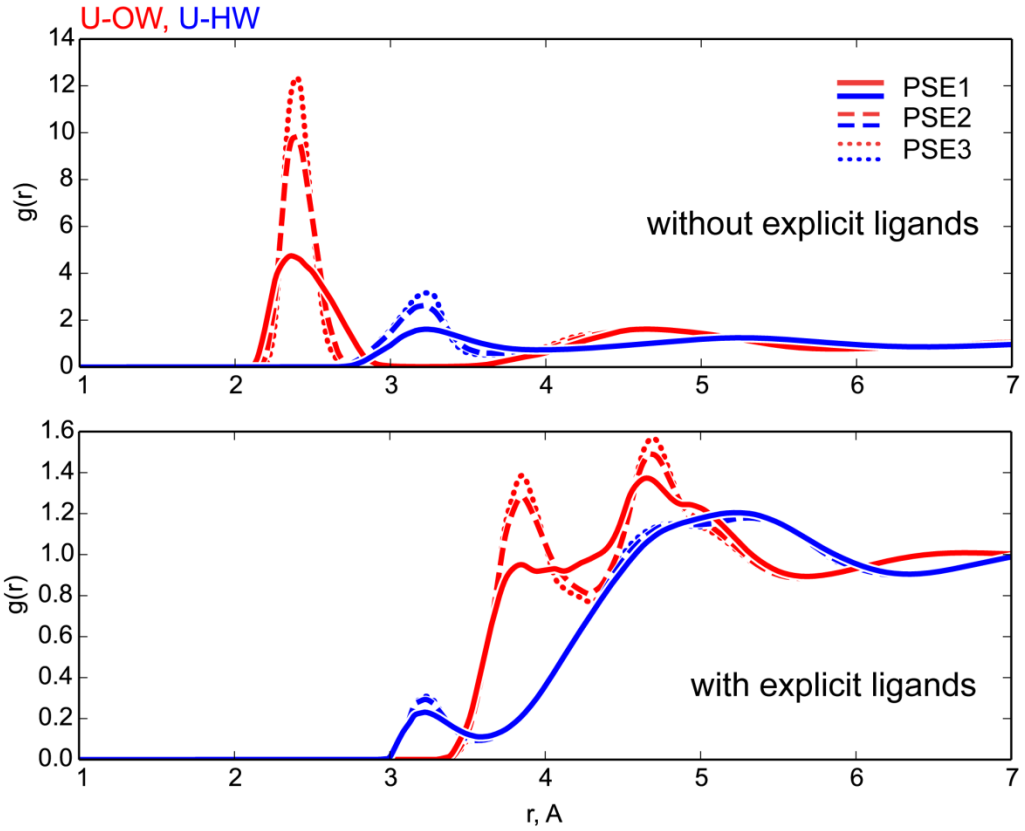


Figure 4.4.4. Radial distribution functions of solvated uranyl without (top panel) and with (bottom panel) explicit aqua ligands from QM+3D RISM calculations: U−O_W (red) and U−H_W (blue). Results obtained with the PSE1 closure are shown as solid lines; results obtained with the closure relations PSE2 and PSE3 are displayed using dashed lines and dotted lines, respectively.

Consistently longer U−O and U−O_w bond lengths were determined in QM+1D RISM calculations compared to QM+PCM and experiment when treating these symmetric models (Section 4.4.5). This trend may be attributed to the excess coordination of uranyl as manifested by the presence of a peak in the U−O_w RDF at the distance that corresponds to the first solvation shell for models with explicit ligands. From these 3D RISM calculations, it is clear that this artifact is specific to 1D RISM. Comparing the U−O_w RDFs from the models without and with five water ligands (Figure 4.4.4), such a feature is absent in the RDF plot. However, the number integrals 5.3, 6.2, and 6.4 for the first peak of the U−O_w RDFs with PSE closures of orders 1 to 3, respectively, are overestimated in the model without explicit aqua ligands (Figure 4.4.4).

Table 4.4.7. Uranyl bond length, d(U−O), uranyl−water distance, d(U−O$_W$), and estimate, ΔG_{sol}, of the uranyl solvation energy, as obtained using a QM model for a system in the gas phase (GP) as well as QM+PCM and QM+3D RISM models.[a]

n		0	4[b]	5[c]	6[d]
d(U–O)	GP	172	177	177	178
	PCM	174	179	179	180
	PSE1	179	179	179	180
	PSE2	179	179	179	180
	PSE3	179	179	179	180
d(U–O$_W$)	GP		242	249	253
	PCM		237	243	249
	PSE1		237	243	248
	PSE2		237	243	248
	PSE1		237	243	248
ΔG_{sol}	GP[e]		–244	–270	–289
	PCM	–322	–416	–426	–425
	PSE1	–357	–377	–388	–393
	PSE2	–373	–379	–389	–394
	PSE3	–378	–379	–389	–393

[a] Bond lengths in pm, energies in kcal/mol. Symmetric uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$, $4 \le n \le 6$, and the bare uranyl (n = 0) are treated by QM, and interactions with the PR-SPC/E water solvent are treated by the KL2 force field. [b] D_{4h} structure. [c] D_{5h} structure. [d] D_{3d} structure. [e] Binding energies of n aqua ligands.[106]

The QM+3D RISM and QM+PCM results agree for uranyl bond lengths of 179−180 pm with the QM+PCM result of 174 pm, with the bare uranyl remaining the only less satisfactory exception (Table 4.4.7). On the other hand, the QM+3D RISM results for uranyl bond lengths are around 179 pm, hence equally long (PSE1) or even marginally longer by 0.3 and 0.7 pm for PSE2 and PSE3, respectively, than for four-coordinated uranyl. The effect of the RISM medium on the bond length of bare uranyl is thus comparable to that of the explicit water ligands; the elongation of the uranyl bond by 7−8 pm corresponds to 5−6 kcal/mol change in the intra-molecular energy.

Another notable discrepancy between the PCM and 1D RISM results mentioned in Section 4.4.5 is also resolved by the 3D RISM approach. Namely the uranyl−water distances U−O$_W$ are consistently shorter in solvated complexes than in the gas phase by a measurable

amount of 4−6 pm. The 1D RISM approach did not show this effect in a convincing manner. Recall that this effect of the medium can be interpreted as a screening of repulsive ligand−ligand interactions in the first coordination shell due to the solvent.

The results for the solvation energy ΔG_{sol} of uranyl estimated with QM+3D RISM models are comparable to those obtained in the previous work using 1D RISM. The estimates of −388 kcal/mol and −389 kcal/mol using the standard model of hydrated uranyl with five explicit water ligands vary only little with the order of the closure and agree very well with the result −390 kcal/mol from the previous work in Section 4.4.5. However, the relative order of the estimates based on the four- and six-coordinated complexes is not reproduced. The estimate of ΔG_{sol} from the four-coordinated complex is by ∼10 kcal/mol less negative than the result of the standard model with five aqua ligands, see Table 4.4.7. That difference of ΔG_{sol} values is the same as obtained from the PCM calculations, although the PCM values are by ∼40 kcal/mol more negative than the 3D RISM results. With 1D RISM the ΔG_{sol} values differ by 1.6 kcal/mol only, see Table 4.4.5.

Water Exchange

Besides the overestimation of the bond lengths of U–O and U–O_W exposed when using 1D RISM method was resolved when applying 3D RISM, the spatial representation of the solvent structure in the latter method also allows one to investigate the process of exchanging an aqua ligand of uranyl and a water molecule from the surrounding solvent. The activation energy for such exchanging process is estimated at about 9 kcal/mol when one applies the Eyring equation to the experimentally available water exchange rate of 1.3×10^6 s^{-1}.[198] Many theoretical studies[134,185,186,194,199,200] predict a much faster water exchange than that the available experimental data suggests. As the experimentally observed exchange rate is many orders of magnitude lower than a diffusion-controlled reaction rate $4\pi\rho DR$ for a solvent molecule characterized by self-diffusion coefficient D and a solute approximated as spherical particle of radius R (uranyl, R = ∼3 Å), the rate limiting step of the reaction must be an exchange of a uranyl water ligand and a solvent molecule from the second solvation shell.[134] Such an exchange involves breaking a uranyl−ligand bond for the molecule leaving the first solvation shell and dissolving a network of hydrogen bonds for the incoming water molecule. The section below will focus on the water exchange between the first and second solvation shells, without assuming any specific mechanism of this reaction. However, given that a six-coordinated uranyl species was determined to be significantly more stable in the gas phase than a four coordinated species (Table 4.4.7), it can be anticipated that an associative mechanism via the six-coordinated intermediate is preferred over a dissociative mechanism via the four-coordinated species. Furthermore, the equivalence of the solvent molecules

participating in the exchange allows one to focus on the first half-reaction with the intermediate as final state.

The analysis of the potential of mean force (PMF) in Ref. 106 predicted the activation energy barriers ranging from 3.1 kcal/mol to 5.2 kcal/mol,[106] which are significantly lower than the experiment estimate, 9.1 kcal/mol.[198] The possible reason can be attributed to the fact that the radical reduction of the dimensionality of the free energy surface will likely obscure mechanistic details. Hence, the (one-dimensional) PMF profile does not indicate any metastable intermediate. It is not possible to draw any conclusion regarding the collective behavior of the incoming and leaving ligands from one-particle PMF profiles. To investigate the collective behavior of the ligands in the water exchange process, one needs an explicit treatment of the first solvation shell and the incoming water molecule. Besides, a flexible uranyl moiety has to be used to evaluate accurately the exchange process.[134] To describe a flexible uranyl species, two intramolecular force fields that primarily differ by the force constant of the uranyl bending modes and the equilibrium bond length d_0(U−O) were employed. The force field referred to as "soft" yields d_0(U−O) = 176 pm while the harmonic frequency of the bending mode is 159 cm^{-1}.[134] The alternative "hard" force field features a somewhat longer uranyl bond d_0(U−O) = 180 pm.[127] The corresponding frequency of the uranyl bending mode (393 cm^{-1}) is also significantly higher than the experimental references, which have been characterized in the range 180−270 cm^{-1} from the bending vibrations in solution and in crystals.[201-204] The DFT models at the BP level employed in the present and in earlier works yield the bending frequency in this experimental range when one uses 5−6 explicit water ligands.[205] For uranyl species in the gas phase, such standard DFT calculations with the BP exchange−correlation functionals predict a too soft bending mode. The three free energy profiles are shown along the reaction coordinate $q = d(U{-}O_{w1}) - d(U{-}O_{w2})$ where the labels "1" and "2" refer to the incoming and leaving water molecules, respectively. Small values of q correspond to a metastable six-coordinated complex and the large absolute values of q correspond to a five-coordinated uranyl complex with one ($q > 0$) or the other ($q < 0$) explicit water molecule in the second shell. The full profile is symmetric around $q = 0$. Due to the apparent symmetry of the metastable state and the exchange process as the whole, only the non-redundant part of the actual reaction profile is displayed in Figure 4.4.5.

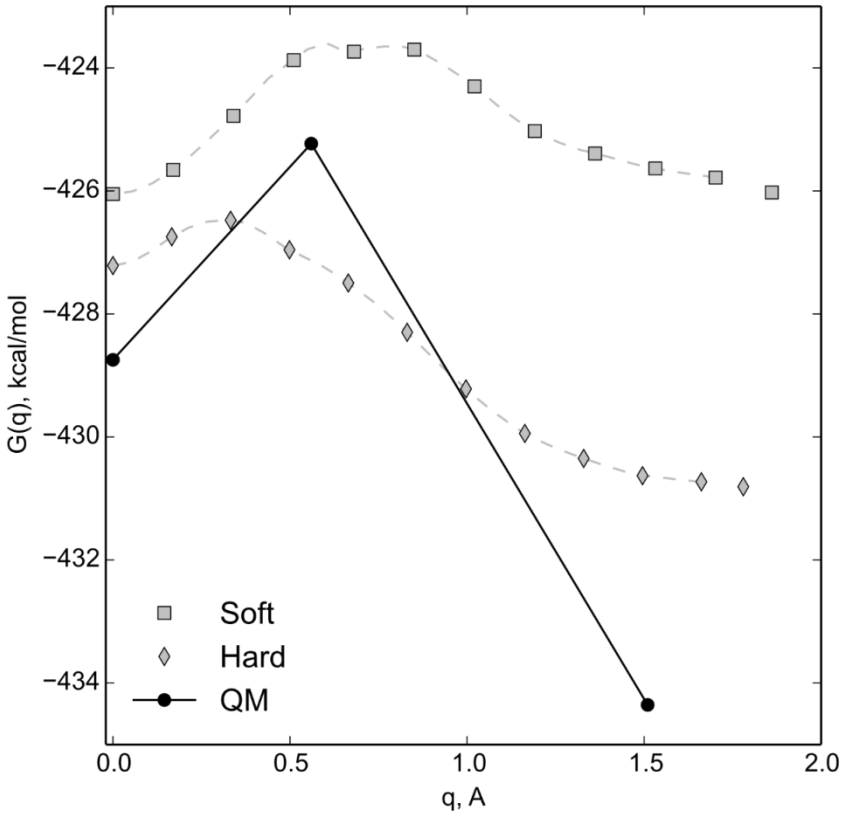


Figure 4.4.5. Free energy profiles, $G(q)$, along the water exchange pathway using uranyl force fields with soft (squares) and hard (diamonds) uranyl bending modes. The QM+3D RISM profile is represented by three stationary states (circles) connected by strait lines to guide the eye. The reaction coordinate q is the difference of U−O_W distances of two selected water molecules.[106]

The QM+RISM profile is approximated by just three stationary points, two local minima and a saddle point obtained as root of the Lagrange multiplier in a series of constrained optimizations. The reaction path was sampled in a preparation step using lower quality settings for the QM and 3D RISM energy terms. All three profiles feature an intermediate metastable state at around $q = 0$, roughly corresponding to the six-coordinated D_{3d}-like uranyl aqua complex. With the "soft" force field, the uranyl species in this metastable state is bent with an O−U−O angle of 155 ° and the distances of uranium atom to the two water ligands participating in the exchange are about 15 pm longer than the distances to the other four water ligands. The deviations from the D_{3d} geometry are less pronounced for the other two models, QM and "hard" force field.

The transition state approximated by the structure with the highest free energy is characterized by $q > 0.6$ Å for both the QM and the "soft" uranyl models, with the longest U−O_W distances at 303 pm and 310 pm, respectively. These distances are comparable to or

slightly below the estimates derived from the 1D and 3D PMFs, see Ref. 106. On the other hand, the "hard" uranyl model suggests a transition state with a much shorter distance, ~280 pm, to the outermost water ligand and a shallow metastable state. This is likely a consequence of the significantly longer U−O bond of the solvated uranyl species, which is ~186 pm and more than 6 pm longer than the corresponding distances of the other models employed here and in other works.[127,134] This particular elongation should be considered as a deficiency of the parametrization. In turn, the "soft" model does not convincingly show a preference for five-coordinated uranyl with the sixth water in the second solvation shell over the six-coordinated broken symmetry metastable structure. The QM model and the "hard" force field model predict the six-coordinated D_{3d}-like state to be less stable by 5.6 and 3.6 kcal/mol, respectively, than the five-coordinated structure with an extra water in the second shell (Table 4.4.8). The presence of such a six-coordinated intermediate may be taken to suggest an associative two-step water exchange mechanism. This latter finding is consistent with the results of other theoretical studies,[134,199,206-208] which report an associative mechanism more favorable than the dissociative mechanism via the four-coordinated intermediate structure.

Table 4.4.8. Free energy values $G(q)$ at local minima $q = q^*$, Corresponding to a five-coordinated uranyl complex with an explicit water molecule in the second shell, and $q = 0$, representing a six-coordinated metastable intermediate, as well as at the corresponding transition state, $q = q^{\#}$, obtained by MM+3D RISM and QM+3D RISM approaches, Figure 4.4.5.[a]

	MM+RISM		
	soft	hard	QM+RISM
$G(q^*)$	–426.0	–430.8	–434.4
$G(q^{\#})$	–423.7	–426.5	–425.2
$G(0)$	–426.1	–427.2	–428.7
$\Delta G^{\#}$ [b]	2.3	4.3	9.1

[a] Free energies in kcal/mol. The exchange reaction corresponds to a transition from q^* to $-q^*$, but due to symmetry $G(q^*) = G(-q^*)$ and $G(q^{\#}) = G(-q^{\#})$. [b] The activation free energy is estimated as $\Delta G^{\#} = G(q^{\#}) - G(q^*)$.

The water exchange barrier of 2.3 kcal/mol and 4.3 kcal/mol calculated with the "soft" and the "hard" force field models, respectively (Table 4.4.8), are rather close to the estimates obtained from the PMF analysis ranging from 3.4 kcal/mol to 5.2 kcal/mol.[106] Yet, these barrier values are significantly lower than the result determined with the QM+3D RISM model (Table 4.4.8) and the experimental estimate.[198] These latter two values of ~9 kcal/mol

are in excellent agreement. A comparison of the three models regarding the location and the orientation of the second-shell explicit water molecule suggests that the force field models slightly underestimate the strength and alter the directional nature of the hydrogen bonds between the incoming explicit water molecule and two of the explicit aqua ligands of uranyl. This may explain the low activation energy for water exchange and the stability of the six-coordinated metastable state relative to the structure with an explicit water molecule in the second shell when predicted by the force field models (Figure 4.4.5).

Finally, the present results for the activation barrier of exchanging two water molecules of the first and the second solvation shell of uranyl are compared to the results of some state-of-the-art QM calculations. The computed values are sensitive to the choice of the reaction path, the level of theory, and the solvation model.[134] Tsushima[206] reported activation barriers of 10.1 kcal/mol and 4.8 kcal/mol at the B3LYP level with the CPCM approach by considering final states with one and two hydrogen bonds, respectively. The latter value is close to the prediction of 4.5 kcal/mol of an earlier study,[199] yet both values are lower than the experimental result. Rotzinger applied a complete active space self-consistent field method in combination with two solvation models, a PCM approach and a spherical cavity self-consistent reaction field (SCRF) solvation model.[208] For the associative pathway he obtained barriers of 9.4 kcal/mol and 6.5 kcal/mol, respectively. Bühl and Kabrede calcualted the free energy profile for the ligand exchange of uranyl from the thermodynamic sampling of a constrained Car−Parrinello molecular dynamic calculation for which they combined the BLYP functional with norm-conserving pseudopotentials.[207] They observed a “shallow ragged plateau” around the six-coordinated intermediate and estimated the activation barrier by 6.7 kcal/mol, which is only slightly lower than the experiment and the present QM+3D RISM value.

The good agreement between the QM+3D RISM result and experiment[198] as well as with the values of other accurate theoretical models indicates that the 3D RISM model provides an attractive approach for solvation in particular as a hybrid solvation strategy applied to uranyl aqua complexes. Using a semi-local DFT method in combination with the 3D RISM strategy offers an adequate accuracy at an affordable cost.

4.4.7. *Conclusions*

The atomistic nature of the 1D RISM model offers a more detailed and consistent view on solvation effects of uranyl aqua complexes in comparison to the state-of-the-art PCM models. Combined QM+1D RISM models predict essentially the same uranyl bond lengths and similar solvation free energies, independent of the number of explicit water molecules introduced into

the first solvation shell. The results of the PCM solvation model depend more strongly on the structure of an explicitly treated first solvation shell. Exploration of the free energy surface with a MM+1D RISM method suggests a qualitative difference to the polarizable continuum models. The solvation energy and the structure of the first solvation shell of an under-coordinated complex of uranyl with four water molecules match those of a five-coordinated complex. Thus, explicit water molecules in the first solvation shell and their net statistical effect as modeled with the 1D RISM approach are largely interchangeable.

Without the need for statistical sampling required for explicit solvation models, the 1D RISM method predicts a water-solvent structure around a uranyl ion in good qualitative agreement with experimental measurements. The QM+1D RISM approach yields a small free energy difference of −1.6 kcal/mol between four- and five-coordinated uranyl complexes in aqueous solution, which is in excellent agreement with the experimental result of −1.2 kcal/mol. This success stands in contrast to the results of polarizable continuum models and AIMD calculations which both notably overestimate this difference. However, one qualitative difference between the PCM and 1D RISM models remains arguable. The former method predicts uranyl-water distances shorter in solution than in the gas phase while the latter yields only a small change for this quantity upon solvation.

3D RISM is not susceptible to this deficiency. When applying the 1D RISM technique, an artificial superposition of explicit aqua ligands and the RISM medium leads to an overestimation of the effective uranyl-water coordination, which results in an excessive elongation of the U−O and U−O_W bonds. With the present 3D RISM method, semi-local DFT calculations on uranyl aqua complexes with 4−6 explicit ligands essentially predict the same structure of the first solvation shell as the conventional PCM strategy. The net effect of a missing explicit aqua ligand in the first solvation shell is correctly represented in a qualitative fashion by an excess solvent density of the RISM medium, leading to comparable results for uranyl with 5 and 4 explicit water ligands.

Any hybrid approach that combines two models of ion−water interactions, “first-principles” and force field, relies heavily on the consistency of the two interaction models. For uranyl hydration, the force field KL2 features large charges of interacting sites as selected in the 1D RISM study for predicting the largest uranyl solvation energy in absolute terms. This model is far from being practical for routine calculations because it strongly favors substitution of explicit QM aqua ligands by the 3D RISM medium. Such artifacts affect the 3D RISM approach, suggesting that this particular QM+RISM combination may require some revision.

Nevertheless, by employing the 3D RISM model for the second solvation shell and the

bulk water, the hybrid QM+3D RISM model produces meaningful structures of the first solvation shell and results in solvation energies that are competitive to the state-of-the-art PCM procedures.

By examining the activation energy for the exchange of aqua ligands, this application convincingly demonstrated the benefits of the hybrid QM+3D RISM model for addressing properties that are not easily reproduced by pure force field models. This type of activation energy is predicted too low with the current force field models, likely due to a misrepresentation of hydrogen bonding. In contrast, the present hybrid QM+3D RISM model yields an activation barrier for ligand exchange that agrees remarkably well with the available experimental results.[198]

4.5. CO_2 Conversion to $HCOO^-$ in Solution by 3D RISM

4.5.1. *Introduction*

The conversion of CO_2 into hydrocarbon fuels is a crucial process which would benefit our environment and the energy industry.[14] However, due to the high kinetic and thermodynamic stability of the CO_2 molecule, this conversion process consumes a large amount of energy. CO_2 conversion is therefore only reasonable in combination with renewable energy sources such as solar energy as well as with the aid of an appropriate catalyst to overcome the otherwise very high reaction barriers.[14,209,210] The photochemical reduction process of CO_2 seems a promising way to solve both of these difficulties. However, understanding the way in which photons provide the energy for the reaction represents a major concern in both, experimental and theoretical studies.[14,209,211] For this purpose the use of homogeneous catalysts such as molecular catalysts that possess pyridine-based ligands is greatly desired, as it is well known since early eighties that those ligands in combination with, for instance, a Ru metal center can photochemically promote the reduction of CO_2.[211,212] Using the catalyst $Ru(bpy)_2(CO)H^+$ (bpy: 2,2′ bipyridine) suggested in an earlier electrochemical study,[213] Damianos et al.[158] carried out a computational study by utilizing the PCM method to investigate the mechanism of the overall catalytic cycle for the reduction of CO_2 into formate in acetonitrile solution.

However, as noted in Section 2.1, PCM method misses the microscopic picture of the solvent structure. This information can be conveniently obtained from the 3D RISM approach, and it has been shown in Sections 4.3 and 4.4 that the hybrid method of 3D RISM and QM yields solvation energy results comparable to state-of-the-art continuum solvation model. To this end, the combination of 3D RISM with QM calculation is applied to this problem to examine the overall catalytic reaction cycle of the CO_2 conversion investigated in the aforementioned PCM study. The free energies of all the reaction species are calculated with the first-order perturbation theory (PT1) QM+RISM model as described in Section 3.3. The PT1 model approximates the solvation energy results at an adequate level of accuracy and its simplicity allows its application to model the complicate molecular solutes in this section. To compare with the results obtained in the preceding PCM study, in the free energy functional $G = E + \Delta G_{\mathrm{sol}}$ the solvation free energy ΔG_{sol} calculated with the PCM method is replaced the excess chemical potential term $\mu[n_{\mathrm{u}}^{(0)}]$ calculated with 3D RISM, see Eq. (3.3.12).

For brevity we used **A** and $\mathbf{A}^+$ to represent the reduced and non-reduced catalyst complex, respectively. The molecular structures of the catalyst $\mathbf{A}^+$ and the acetonitrile solvent are shown in Figure. 4.5.1. Note, that the preceding PCM study revealed only minor structural

differences (e.g. Ru-ligand distances varying less than 2 pm) between the systems **A** and $\mathbf{A}^+$. For brevity the two bpy ligands in the catalyst are labeled as **L1** and **L2**, respectively. The CO ligand is nearly perpendicular to the pyridine rings of **L2** while being almost coplanar to **L1**.

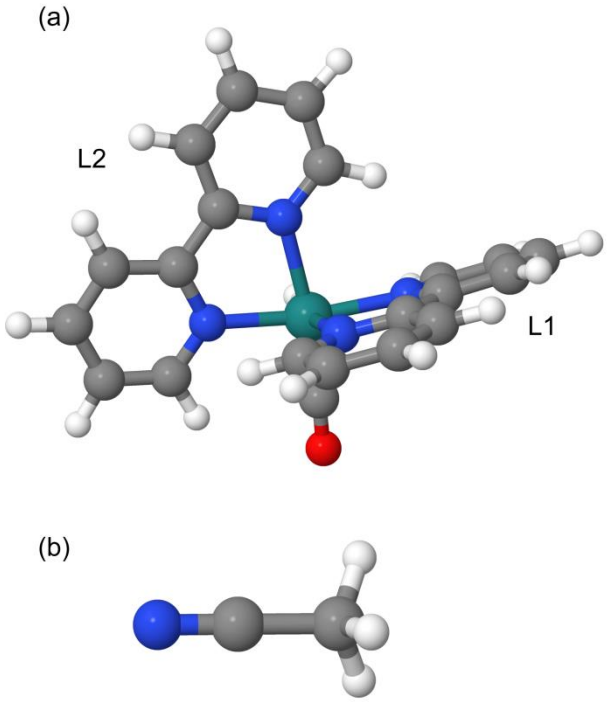
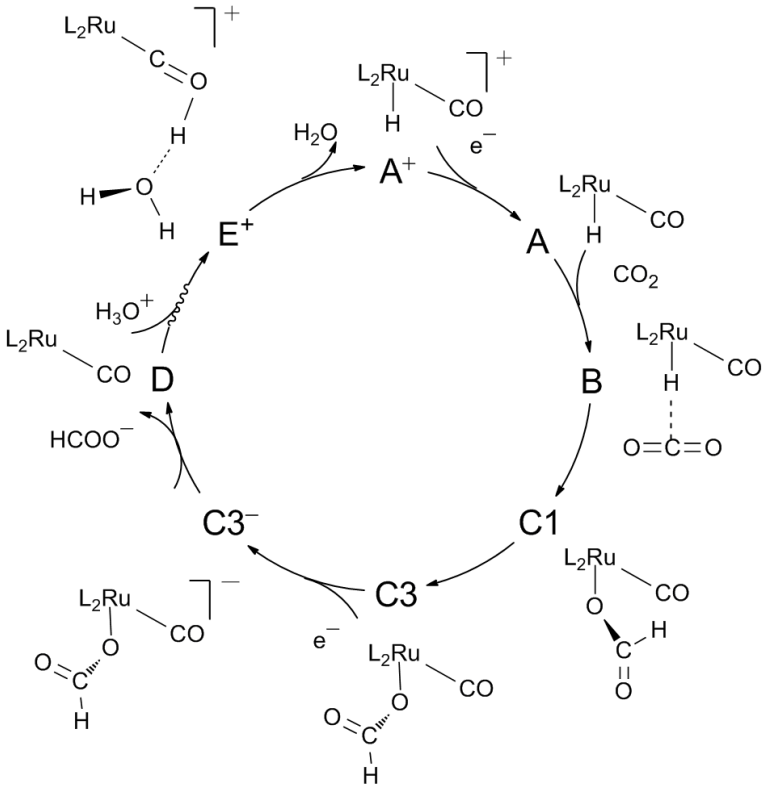


Figure 4.5.1. Molecular structures of (a) the catalyst $\mathbf{A}^+$: $Ru(bpy)_2(CO)H^+$ and (b) the acetonitrile solvent. The two bpy ligands in the catalyst molecule are labeled as **L1** and **L2**, respectively.

The discussion of the results in this section starts by examining the solvent structure of acetonitrile around the catalyst, for both reduced and non-reduced species. Then following the catalytic cycle in Scheme 4.5.1 the reactions happened in the first and second one-electron reduction steps, including the CO_2 insertion (**A^+→A→B→C1**) and isomerization (**C1→C3**), $HCOO^-$ formation (**C3→$C3^-$→D**) and catalyst regeneration (**D→A^+**) are discussed by comparing the 3D RISM and PCM calculations for all those reaction steps. Finally the overall catalytic cycle obtained by the 3D RISM approach is presented and also compared with that from PCM calculations.[158]



Scheme 4.5.1. Main catalytic cycle that has been suggested in Ref. 158 for the formation of $HCOO^-$ starting from the catalyst complex $\mathbf{A}^+$ as well as the two one-electron transfer steps are included. L_2 represents the bpy ligands.

4.5.2. *Solvation Structure of the Catalyst*

The spatial information of acetonitrile solvent around the positively charged catalyst $\mathbf{A}^+$ and the neutral species **A** in the presence of reduction can be found in Fig. 4.5.2. As shown in Fig.4.5.2a, there are three types of regions where the carbon atom of methyl group is concentrated in the first solvation shell: the outer sphere around the plane of the pyridine rings of the ligand **L1**, the regions above and below the pyridine ring, and the region near the oxygen center below one pyridine ring. Note that there is a narrow gap between the outer sphere of the methyl carbon distribution and the cavity formed by the catalyst solute. By comparing the distribution of the nitrogen center (Fig. 4.5.1b), the relevant outer sphere is closely adjoined to the cavity; the distribution of the nitrogen center does not show any localization above and below the pyridine rings of the ligand **L2** or around the oxygen site of the solute. From these observations one can conclude that in the first solvation shell, the acetonitrile solvents are partly distributed around the plane of the pyridine ligand with the nitrogen atom pointing into the solute. Above and below the two pyridine rings of the ligand **L2** as well as around the oxygen center of the CO ligand, acetonitrile solvent molecules are oriented with their methyl group pointing towards the solute. Another region of noticeable concentration of the solvent

is located in the second solvation shell with the methyl group pointing to the oxygen atom of the CO ligand, which is also visible in Figures 4.5.2a and 4.5.2b.

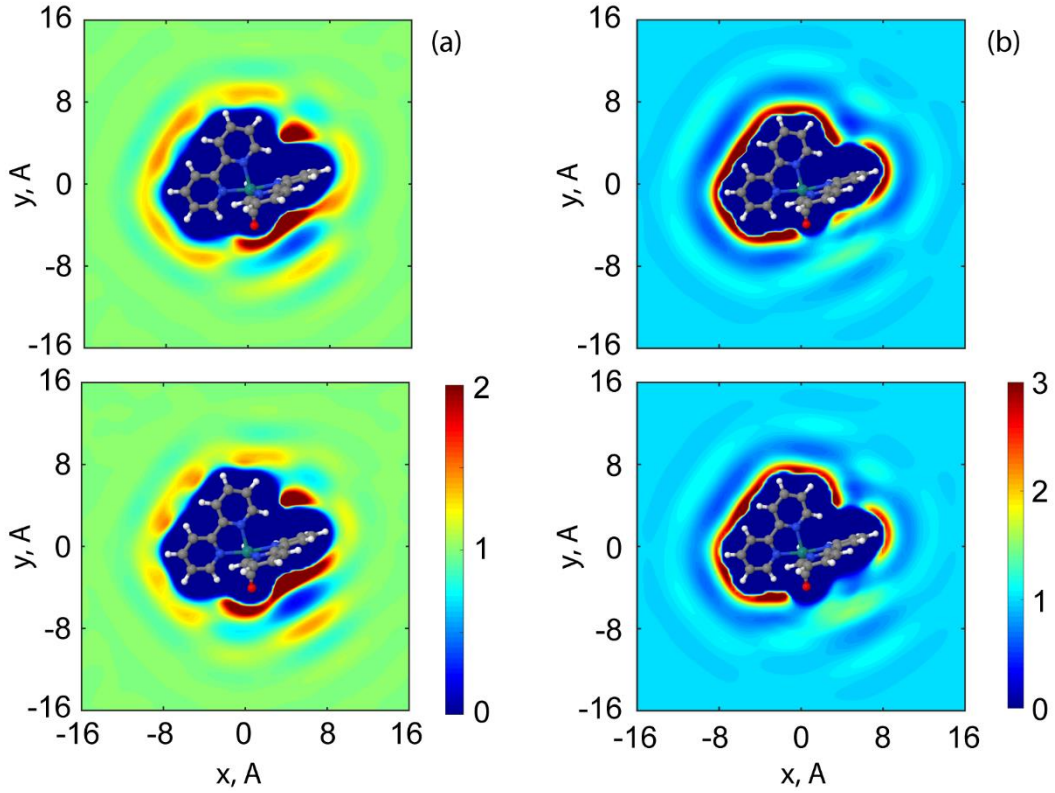


Figure 4.5.2. Acetonitrile carbon atom of methyl group (a) and nitrogen atom (b) distribution around the positively charge catalyst $\mathbf{A}^+$ (top panel) and neutral species **A** (bottom panel). The distribution profiles are cut at the plane $z = 0$ of the cubic cell, with which the CO ligand and the **L1** ligand (Figure 4.5.1) approach.

The differences for the non-reduced and reduce catalysts can be observed when comparing the top panels (solute is positive charged, $\mathbf{A}^+$) with the bottom panels (solute is neutral, **A**) in Figure 4.5.2a and 4.5.2b. When going from solute $\mathbf{A}^+$ to **A**, the surrounding solvent distribution decreases in the outer sphere around the two pyridine rings of the ligand **L1** while increasing in the regions parallel to the two pyridine rings of the ligand **L2**, those around the oxygen atom, and in the second solvation shell with acetonitrile methyl group pointing to the oxygen. This observation can be rationalized by analyzing the charge differences[158] for each atomic site on the positively charged catalyst $\mathbf{A}^+$ and the neutral specie **A** upon one-electron reduction. When $\mathbf{A}^+$ is reduced to **A**, a positive charge of 0.03 *e* on average is shifted to each atomic site on the hydride ligands of all the pyridine rings, while a negative charge of up to 0.06 *e* is transferred to every atomic site on each bi-pyridine ring and the Ru atom as well as to the oxygen atom of the CO ligand. When considering that in the acetonitrile solvent the nitrogen sites possess partially negative charges while the methyl

group is positively charged, the increase of the solvent distribution with methyl groups pointing to the negatively charged part of the solute can be understood in terms of electrostatic interaction. With more electron charge density being transferred to the negative charged sites of the solute upon reduction, the electrostatic interactions between those negatively charged sites of the solute and the positively charged sites of the solvent are enhanced. In consequence the solvent distribution shows more concentrations in the relevant regions.

4.5.3. *CO_2 Insertion and Isomerization Reaction*

Following the reaction path suggested in the preceding (PCM) study (Scheme 4.5.1, **A^+→A→B→C1**),[158] the influence of reduction on the reaction of inserting CO_2 to break the metal-hydride bond will be addressed in this section. As shown in Figure 4.5.3a (without reduction), the free energy differences of the intermediate **B^+** against the initial reactants **A^+** + CO_2 obtained from the PCM and QM+3D RISM results are quite comparable, which yield 9.8 kcal/mol and 9.7 kcal/mol, respectively. With reduction the relevant quantities increase to 12.0 kcal/mol and 11.8 kcal/mol for the latter two variants, respectively, see Figure 4.5.3b. This trend due to the presence of the reduction is reversed when it comes to the transition states **B-C1** and **B-$C1^+$**. With QM+3D RISM the energy barrier for forming the Ru-O species **C1** is lowered by 7.2 kcal/mol from 35.7 kcal/mol for the non-reduction case to 28.5 kcal/mol with reduction. Both of these values are smaller than the barriers of 42.5 kcal/mol and 31.8 kcal/mol obtained with PCM where the energy barrier decreases by about 11 kcal/mol upon reduction.[158] However, even the lowered energy barrier of 28.5 kcal/mol obtained with QM+3D RISM is still significantly higher than the commonly used rule-of-thumb limit of 21 kcal/mol for a reaction to take place at room temperature. It is also larger than the energy barrier of 15 kcal/mol for CO_2 insertion, using a Ru-complex catalyst with a larger ligand as reported by Creutz et al.[214].

The dashed lines in Figure 4.5.3 represent the subsequent isomerization reactions following the CO_2 insertion. A detailed demonstration of the reaction path of the formato-complex from **C1** to **C3** (Scheme 4.5.1, **C1→C3**) can be found in Figures 4.5.4a and 4.5.4b, for cases without and with reduction, respectively. The results obtained from QM+3D RISM are close to those from PCM. As shown in Figure 4.5.4a, QM+3D RISM yields the reaction barriers of formation of the isomers **$C2^+$** and **$C3^+$** as 6.1 kcal/mol and 5.4 kcal/mol, respectively, which are quite comparable with the values of 5.4 kcal/mol and 5.2 kcal/mol calculated by PCM.[158] The trend shown in Figure 4.5.4b is similar to that in Figure 4.5.4a, when using the 3D RISM method the reaction barriers of forming **C2** and **C3** are 6.7 kcal/mol and 5.3 kcal/mol, respectively, marginally higher than the corresponding values of

5.8 kcal/mol and 4.7 kcal/mol from PCM.[158] Both, QM+3D RISM and PCM, results show that the one-electron reduction does not have significant effects on the isomerization reactions. Although QM+3D RISM yields somewhat higher energy barriers for the CO_2 insertion reaction than PCM, the difference of the energy barrier for isomerization reaction is less than 1 kcal/mol only.

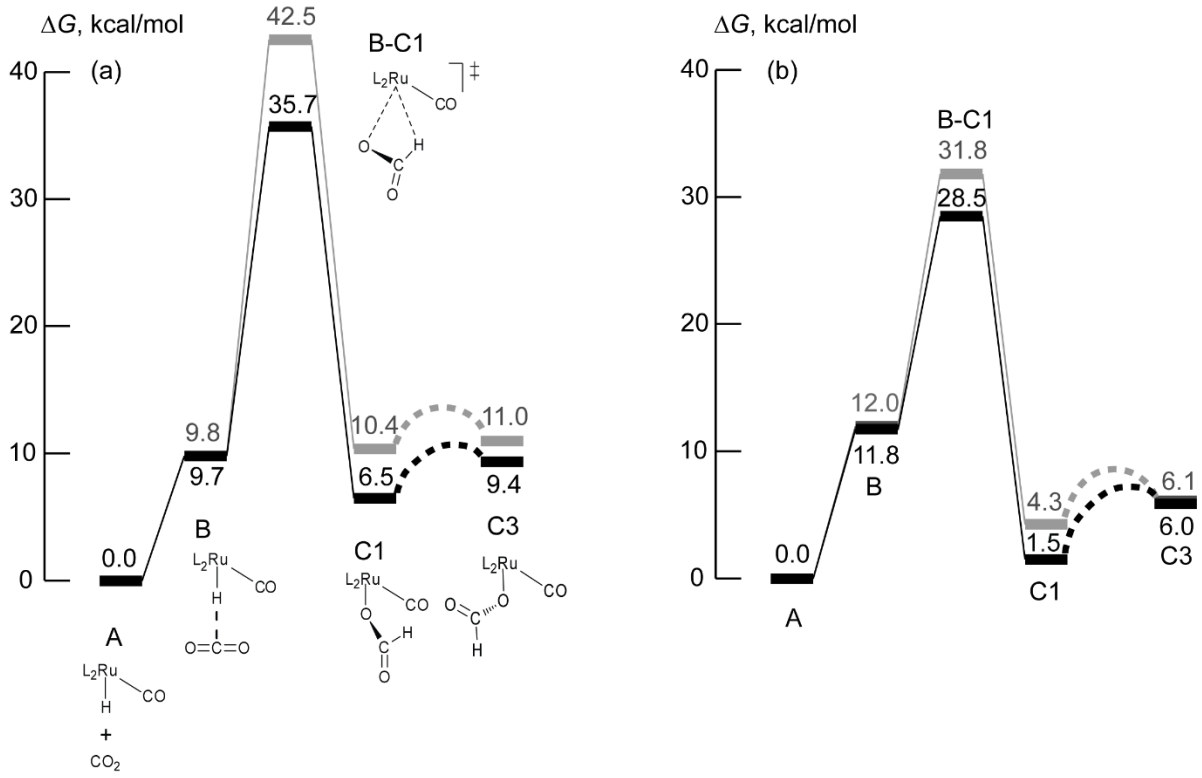


Figure 4.5.3. (a) Free Energy profiles of CO_2 insertion in the (a) absence and (b) presence of reduction leading to the Ru-O bound species **C1** through transition state **B-C1**. Species are positively charged in (a) and neutral in (b). As reference energy the sum of the total energies of the catalyst complex and CO_2 are considered. Free energies are given in kcal/mol. QM+3D RISM results are in black and PCM results are in gray.

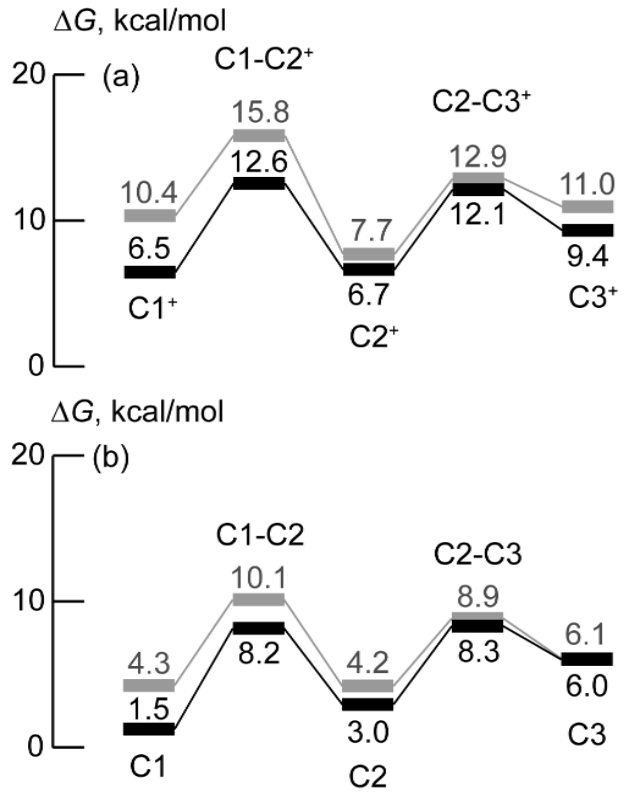


Figure 4.5.4. Isomerization profile of product **C1** to **C3** (a) without and (b) with the presence of reduction. Free energies are given in kcal/mol. QM+3D RISM results are in black, PCM in gray.

4.5.4. *Second Reduction of $HCOO^-$ Formation*

The dissociative path was proposed in the PCM study[158] to start from the format–species **C3**, go through the transition state **C3-D,** and finally form the 5-coordinated species **D** with the generation of $HCOO^-$ (Scheme 4.5.1, **C3→C3$^-$→D**). As shown in Fig. 4.5.5a for the case without reduction, QM+3D RISM predicts the reaction barrier for this dissociation to be 19.4 kcal/mol, hence slightly higher than the PCM value of 17.3 kcal/mol.[158] Upon the second one-electron reduction a relatively large decrease of the energy barrier amounting to 6.6 kcal/mol is obtained from the PCM calculation[158] while that from QM+3D RISM is only 3.3 kcal/mol. As a result, with reduction, the QM+3D RISM method yields a notably higher reaction barrier of 16.1 kcal/mol compared to the PCM barrier of 9.7 kcal/mol.[158]

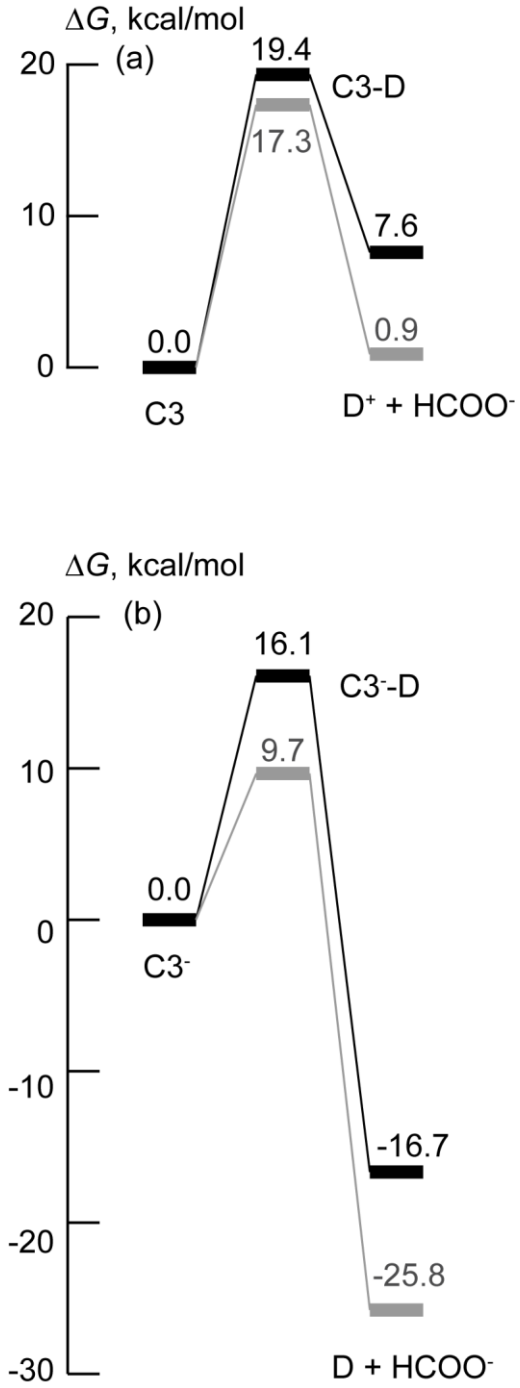


Figure 4.5.5. Free energy profile of the Dissociative mechanism for the formate liberation and the formation of the five-coordinated Ru-complex (a) without and (b) with the presence of the second reduction. QM+3D RISM results are in black, PCM in gray.

This result is in contrast to the CO_2 insertion reaction upon the first one-electron reduction where the reaction barrier is lower in QM+3D RISM than in PCM. It can be rationalized with the solvation free energy term ΔG_{sol}, whose changes are different in these two methods when going from reactants **C3**/**C3**$^{-}$ to transition states **C3-D**/**C3**$^{-}$**-D**. Without reduction ΔG_{sol} decreases by 8.5 kcal/mol from 10.7 kcal/mol for **C3** to 2.2 kcal/mol for **C3-D** in PCM,[158]

while in QM+3D RISM it is reduced to go down from –11.6 kcal/mol to –18.1 kcal/mol with the decrease of 7.5 kcal/mol. However, when the solutes are negatively charged, namely upon the reduction ΔG_{sol} decreases by 8.9 kcal/mol from –18.8 kcal/mol to –27.7 kcal/mol in PCM[158] while only a smaller decrease of 2.5 kcal/mol from –47.2 to –49.7 kcal/mol is obtained in QM+3D RISM. Given that the electronic energy is the same for species in comparison in both calculations, the much higher activation energy barrier of forming $HCOO^-$ can be attributed to the different ways of evaluating of the changes of ΔG_{sol} from the reactant to the transition state obtained from QM+3D RISM and PCM. Yet, despite this difference, 3D RISM results still support the conclusions of the PCM study;[158] without reduction the formation of $HCOO^-$ is endergonic while the one-electron reduction lowers the reaction barrier.

4.5.5. *Catalyst Regeneration*

Following the reaction path in the PCM study,[158] the catalyst $\mathbf{A}^+$ is regenerated from the 5-coordinated species **D** (Scheme 4.5.1, $\mathbf{D}\rightarrow\mathbf{A}^+$). The catalyst regeneration profiles of the two-step pathway modeled by PCM and QM+3D RISM methods are shown in Figure 4.5.6. At the start of the catalyst regeneration process, the five-coordinated species **D** is protontated with the combination of H_3O^+ and H_2O to form the species $\mathbf{G}^+$. While PCM calculates the free energy difference between the pronated **D** and $\mathbf{G}^+$ to –3.9 kcal/mol, the QM+3D RISM calculation in this thesis yields a more energetically favored result with a free energy difference of –9.8 kcal/mol.

This noticeable difference can be observed during the intermediate process $\mathbf{G}^+\rightarrow\mathbf{E}^+$ with the loss of the water molecule loss from the pronated species. As can be seen from Figure 4.5.6, the PCM results show that this process has two steps overcoming first an energy barrier of 1.7 kcal/mol to complete the proton transfer from H_3O^+ and H_2O to species $\mathbf{H}^+$ (path $\mathbf{G}^+\rightarrow\mathbf{H}^+$) and then another barrier of 1.4 kcal/mol to attach a water molecule to the CO ligand (path $\mathbf{H}^+\rightarrow\mathbf{E}^+$). In contrast, QM+3D RISM shows that this process is single-step with an energy barrier of about 2.8 kcal/mol (considering $\mathbf{H}^+$ with two free water molecules as the intermediate). No additional search for a transition state was attempted with QM+3D RISM were performed and the geometries and electronic structures of the relevant species are identical in both, PCM and QM+3D RISM calculations. Therefore the argument of "two-step" vs. "one-step" in these two models must be taken carefully into account. It was also mentioned in the PCM study[158] that the reaction path $\mathbf{D}\rightarrow\mathbf{G}^+\rightarrow\mathbf{H}^+$ lacks sufficient optimization for a species that comes from the forward path of the intrinsic reaction coordinate. The discrepancy between PCM and QM+3D RISM results presented above demonstrates that not only the

quantitative differences of the free energies, but also qualitatively different reaction mechanisms can be obtained when using those two solvation models.

Regardless of the difference in the intermediate process, the final step from $\mathbf{E}^+$ to $\mathbf{A}^+$ is again similar in PCM and QM+3D RISM, with a reaction barrier of 1.6 kcal/mol obtained from both methods.

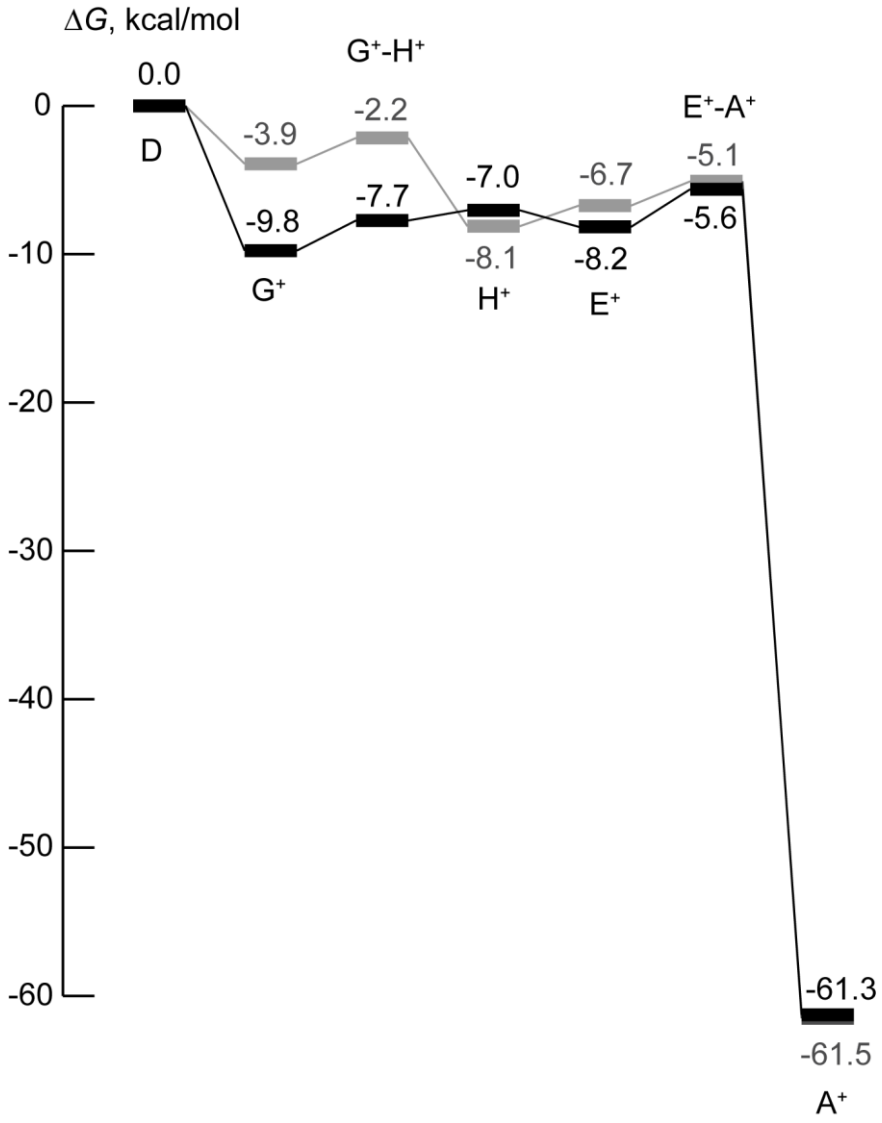


Figure 4.5.6. Catalyst regeneration profiles that starts from the neutral, 5-coordinated species complex **D**. Free energy levels are given with respect to the summation of species **D**, H_3O^+ and H_2O. The step from $\mathbf{E}^+$ to $\mathbf{A}^+$ benefits from an explicit H_2O ligand. QM+3D RISM results are in black, PCM in gray.

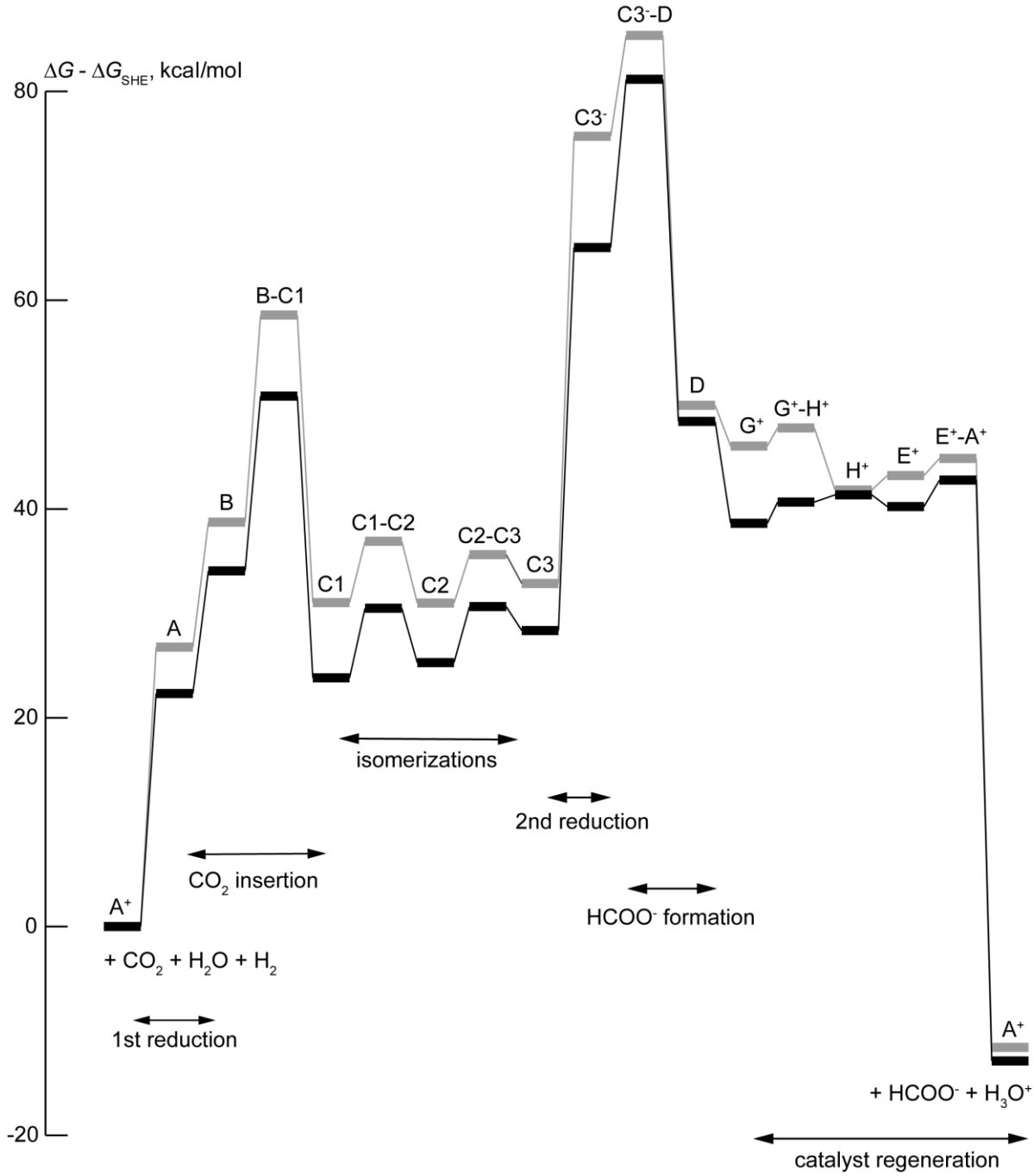


Figure 4.5.7. The overall reduction profile leading to the formation of $HCOO^-$. Free energies are given in kcal/mol and drawn against the standard hydrogen electrode reaction (SHE), Eq. (4.5.1), as in the preceding PCM study.[158] The reference energy is the sum of the energies of CO_2, H_2O and H_2 in which the two electrons of each half cell reaction cancel out. The two, one-electron reductions are indicated by the paths $\mathbf{A^+ \rightarrow A}$ and $\mathbf{C3 \rightarrow C3^-}$. QM+3D RISM results are in black, PCM in gray.

4.5.6. *The Overall Catalytic Cycle*

Finally, the energy profile of the overall reaction cycles is shown in Figure 4.5.7. Analogously to the reference PCM study[158], each step of the first and second reductions calculated by QM+3D RISM is considered against the half-cell oxidation reaction of the standard hydrogen electrode (SHE)[215]

$$\tfrac{1}{2}\mathrm{H_2} + \mathrm{H_2O} \rightarrow \mathrm{H_3O^+} + 1e^- \qquad (4.5.1)$$

whose reaction energy is about 98.7 kcal/mol. With this treatment on one hand the largely negative absolute values of the free energies can be avoided when expressing the overall reaction path, on the other hand the half-cell oxidation reaction, Eq. (4.5.1), can be considered as the source of consumed electrons for reduction.[158] To this end, the overall absolute reaction

$$\mathrm{CO_2} + \mathrm{H_3O^+} + 2e^- \rightarrow \mathrm{HCOO^-} + \mathrm{H_2O} \qquad (4.5.2)$$

can be transformed to the relative reduction reaction

$$\mathrm{CO_2} + \mathrm{H_2} + \mathrm{H_2O} \rightarrow \mathrm{HCOO^-} + \mathrm{H_3O^+}\,. \qquad (4.5.3)$$

For the reduction processes $\mathbf{A^+ \rightarrow A}$ and $\mathbf{C3 \rightarrow C3^-}$, QM+3D RISM yields free energy differences of 22.3 kcal/mol and 36.7 kcal/mol, respectively, which are slightly lower than the PCM results of 26.8 kcal/mol and 42.8 kcal/mol, respectively. The CO_2 insertion exhibits the highest reaction barrier which results to about 30 kcal/mol in both, PCM and 3D RISM calculations. With H_2 and H_2O as educts, the reaction of Eq. (4.5.2) is exergonic by –12.9 kcal/mol in QM+3D RISM, slightly higher than the PCM result of –11.6 kcal/mol.

4.5.7. *Conclusions*

In this application the conversion of CO_2 to $HCOO^-$ in acetonitrile solution with the aid of a Ru-based homogeneous catalyst was examined with the QM+3D RISM model. The solvation structure obtained from 3D RISM shows that the acetonitrile solvents distribute along the outer sphere around the pyridine rings of the ligand **L1** (Figure 4.5.1) with the nitrogen atom pointing to the solute, while above and below the pyridine rings of the ligand **L2** (Figure 4.5.1) as well as around the oxygen atom of CO ligand the solvent distributes itself with its methyl group pointing to the solute. In the case of the reduced catalyst complex more negative electronic charge is located at the negative charge sites of the solute. As a result of this electron transfer the solvent distribution increases in the regions corresponding to those sites.

The conversion free energy profile was also calculated with the solvation free energies obtained from QM+3D RISM in comparison with PCM results.[158] The QM+3D RISM model yields results generally in good agreement with PCM results. The highest activation free energy barrier for the CO_2 insertion is 31.8 kcal/mol in PCM and only slightly lowered to

28.5 kcal/mol in QM+3D RISM. The catalyst regeneration process shows some differences when using these two methods. PCM predicts this process as a a two-step reaction, while from the same geometry and electronic structure a one-step reaction is suggested by the QM+3D RISM calculations. Both methods yield quite low barriers, ~ 2 kcal/mol, for the catalyst regeneration process. By taking the overall catalytic cycle against the standard hydrogen electrode, Eq. (4.5.1), QM+3D RISM and PCM predict both that the $HCOO^-$ formation is an exergonic process of around –10 kcal/mol, with only 1.3 kcal/mol difference for the free energies of the final products.

This application demonstrates that for complicated catalytic reactions, QM+3D RISM is capable of produce free energy results comparable to those of state-of-the-art continuum solvation models. In addition, QM+3D RISM calculations produce information about the atomistic structure of the solvent from which the changes due to the presence of reduction can be clearly observed. Therefore, when comparing to the continuum solvation model, QM+3D RISM method seems more advantageous for investigating chemical reactions which are sensible to the solvation environment.

5. Summary

This thesis presented a novel implementation of a hybrid approach to modeling the microscopic solvent effects that occur in chemical systems and reactions in solution. This hybrid approach combines the reference interaction site model (RISM) representation of the solvent with a quantum mechanical (QM) electronic structure description of the solute. This QM+RISM method has been implemented in the parallel density functional program ParaGauss. The resulting implementation was then verified and applied in several case studies where solvation effects play an important role.

State-of-the-art solvation models either implicitly describe the solvent in terms of a continuum medium or by taking the explicit solvent molecules into account. Both strategies are widely applied despite of several drawbacks associated with them. Continuum solvation models are by design unable to provide any microscopic information about the solvent which in turn prevents an appropriate description of the short-range, non-covalent interactions between solvent and solute such as hydrogen bonding. Explicit solvation models account for both, short- and long-range interactions. However, their applicability is limited either by the accuracy loss originating from a molecular mechanics (MM) description of the whole solution system or by the enormous computational costs of the very many full QM or hybrid MM/QM calculations required for an accurate statistical sampling. On the other hand, a hybrid RISM approach allows one to overcome these drawbacks of traditional solvation models. In contrast to continuum solvation models, RISM methods recover the microscopic description of the solvent medium, while in comparison with explicit solvation models they are free of statistical sampling which significantly reduces the computational costs of their applications.

Based on integral equation theories (IET), RISM methods describe the solvent microscopic structure in terms of site-site distribution functions. RISM theories also provide an analytical expression for conveniently calculating the solvation energy from these site-site distribution functions. Two main variants of this approach exist, 1D RISM and 3D RISM, depending on the specific treatment of the distribution functions. In the 1D version of RISM the site-site distribution functions are spherically averaged, while in the 3D version they are represented as the spatial distribution functions and thus provide a more realistic, spatially resolved description of the solvent structure.

Both approaches can be combined with QM methods. However, for the applications examined with RISM in the context of this thesis, the 3D variant is more favorable as it is more feasible for modeling complicated molecular solutes. Therefore a combination of 3D RISM with Kohn-Sham density functional theory (KS-DFT) was implemented in the parallel

density functional program ParaGauss in the context of this thesis. One of the most important aspects when interfacing 3D RISM with KS-DFT is an adequate evaluation of the electrostatic interaction between the QM solute and the solvent molecules. For this purpose, a robust and efficient scheme is suggested in this thesis. This strategy allows one to determine the perturbation of the solute electronic structure by the solvent environment and vice versa within QM+3D RISM in a self-consistent manner. A comparison with the electrostatic interactions computed with ParaGauss thereby shows no significant loss in accuracy. Furthermore, the solvation energy of the hybrid method can be evaluated either in a fully self-consistent fashion or by approximate expressions that correspond to different orders of perturbation theory. The latter cases were thereby implemented based on a simple linear model that describes the relation between of the solute polarization and the solvent relaxation.

Subsequently the RISM solver implemented in the context of this thesis was carefully validated. In a first step its convergence behavior with respect to different grid resolutions was examined. Furthermore, the solvation energies of monoatomic ions in water were computed and compared with experiments as well as with the results of MD simulations and other RISM solvers. In addition, the PMV corrections were found to compensate different percentages of errors for anions and cations. As a satisfying level of accuracy could be confirmed in those tests, the hybrid approach of QM+RISM was subsequently applied in several case studies.

In a systematical study involving a test set of 43 neutral organic solute molecules in aqueous medium the accuracy of the solvation free energy delivered by the QM+3D RISM method was analyzed. The solvation free energy is mainly determined by two factors, the contribution from excess chemical potential upon solvation and that from the electronic relaxation of the solute. The excess chemical potential is known to be calculated too large due to the overestimated cavity formation error in MM+3D RISM studies. The results obtained from the hybrid QM+3D RISM method in this thesis essentially confirmed this source of error. A significantly better agreement with experimental results can be achieved when the partial molecular volume (PMV) is suitably corrected. This study also presented an asymptotic scaling analysis and thereby provided an alternative interpretation of the origin of the aforementioned error. Furthermore, this analysis also showed an alternative route for deriving the pre-factors required for the individual terms of the PMV correction. In addition, the importance of the relaxation terms, hence the electron relaxation caused by the solvent reaction field and the resulting solvent relaxation energy, was investigated. The ratio of the solute polarization and the solvent relaxation energies was calculated to be nearly constant for all solutes studied. Using the second-order (PT2) approximation or a self-consistent approach

for the relaxation term reduced the deviation of the calculated solvation energy from the experimental reference only slightly. This latter finding likely arises due to the artificial polarization of the electron density by nearby localized positive charge accumulations in the solvent medium described by MM force field. In the absence of repulsive exchange interactions between solute and solvent electrons this effect is likely overestimated, which is also suggested by the slight overestimation of the induced solute dipole moments in comparison to the PCM approach. The above interpretation is further supported by an increased water hydrogen and oxygen density around the solute in the SCF-RISM description. In summary, this systematical studies reveals that when the main error source is removed in the hybrid QM+3D RISM method, the approximations of the free energy functional yield results quite comparable with the fully SCF treatment as well as the PCM method. Although the SCF-RISM method is able to show the solute polarization in solvent medium, this effect is overestimated possibly due to the absence of the solute-solvent exchange repulsion.

Based on the findings in the first application, a second application study addressed the solvation of uranyl in aqueous medium by means of 1D and 3D RISM calculations with the approximative free energy expression. The 1D RISM method was applied first to bare uranyl ions in order to compare the various force field parameterizations for uranyl and water. The models determined most accurate thereby were then used in calculations of uranyl aqua complexes $[UO_2(H_2O)_n]^{2+}$ with n = 4–6. The net effect of a missing explicit aqua ligand in the first solvation shell is represented in a qualitative fashion by an excess solvent density described by 1D RISM and 3D RISM, leading to comparable result for the four- and five-coordinated uranyl complexes. QM+1D RISM and 3D RISM both predict the uranyl solvation free energy quite close to the experimental value. In agreement with experiment, the QM+1D RISM approach also yields a small free energy difference between four- and five-coordinated uranyl complexes in aqueous solution. However, an artificial superposition of explicit aqua ligands and the 1D RISM medium leads to an overestimation of the effective uranyl-water coordination, which results in an excessive elongation of the U−O and U−O_W bonds. 3D RISM is not susceptible to this deficiency and the hybrid QM+3D RISM model thus produces meaningful structures of the first solvation shell. The corresponding results for the solvation energies are comparable to those of state-of-the-art PCM procedures. In addition, the qualitative analyses in 1D and 3D RISM both indicate the interchangeability of the explicit ligands with their RISM representation while this microscopic feature is missing in continuum solvation models like PCM. This advantage is also confirmed by the quantitative aspect, that QM+3D RISM method yields an activation barrier for ligand exchange which agrees remarkably well with the available experimental result compared to the traditional continuum

and explicit solvation models.

To further extend the applicability of this novel hybrid method, the final study examined the solvation effects during the conversion of CO_2 to $HCOO^-$ in acetonitrile solution with the aid of a Ru-based homogeneous catalyst by comparing the results obtained from QM+3D RISM and PCM solvation models. First the solvent structure and especially the orientation of the individual acetonitrile molecules around the catalyst was determined with QM+3D RISM. The resulting figures nicely depict the accumulation of solvent molecules in the case of a cationic, reduced Ru complex. Upon a stepwise comparison of the reaction energy barriers of the catalytic reaction cycle the QM+3D RISM model was found to agree in general with PCM. Consistency between these two variants was also found by taking the overall catalytic cycle against the standard hydrogen electrode, where both methods predicted that the $HCOO^-$ formation to be an exergonic process. However, PCM predicts the catalyst regeneration process is a two-step reaction, while a one-step reaction is suggested by QM+3D RISM calculations from the same geometry and electronic structure.

Overall, this thesis demonstrated that the hybrid method of QM+3D RISM is competitive or even superior to traditional solvation models for the description of chemical systems and reactions in solution. However, the applications presented in this thesis also suggested cases in which the accuracy of RISM results can still be further improved. Also the analysis of the cavity formation error and the electron relaxation term in the solvation energy can be extended to ionic solute molecules. As suggested from the MM+RISM calculations of monoatomic ions, the cavity formation error does not play a major role in QM+RISM calculations of charged solutes. Indeed, in the case of ionic solutes the electrostatic interaction terms are much larger and thus, the corresponding errors in these terms may become important instead. In addition, the current implementation of SCF-RISM in this thesis is dependent on employing the effective core potentials in the QM calculations. For more general purpose one may consider the extension of utilizing larger basis sets with the numerical convergence issues being resolved. Also the accuracy can be improved by taking the repulsive exchange interactions between solute and solvent electrons into account.

Until now, theoretical developments of the general accurate energy functionals are still pending in the hybrid QM+RISM approaches. We hope that the work conducted in this thesis in combination with ParaGauss will help to speed up the popularity and development of this novel hybrid method, which ultimately will improve our understanding of solution science.

Appendix A. Flowchart of the SCF-RISM Cycle of ParaGauss

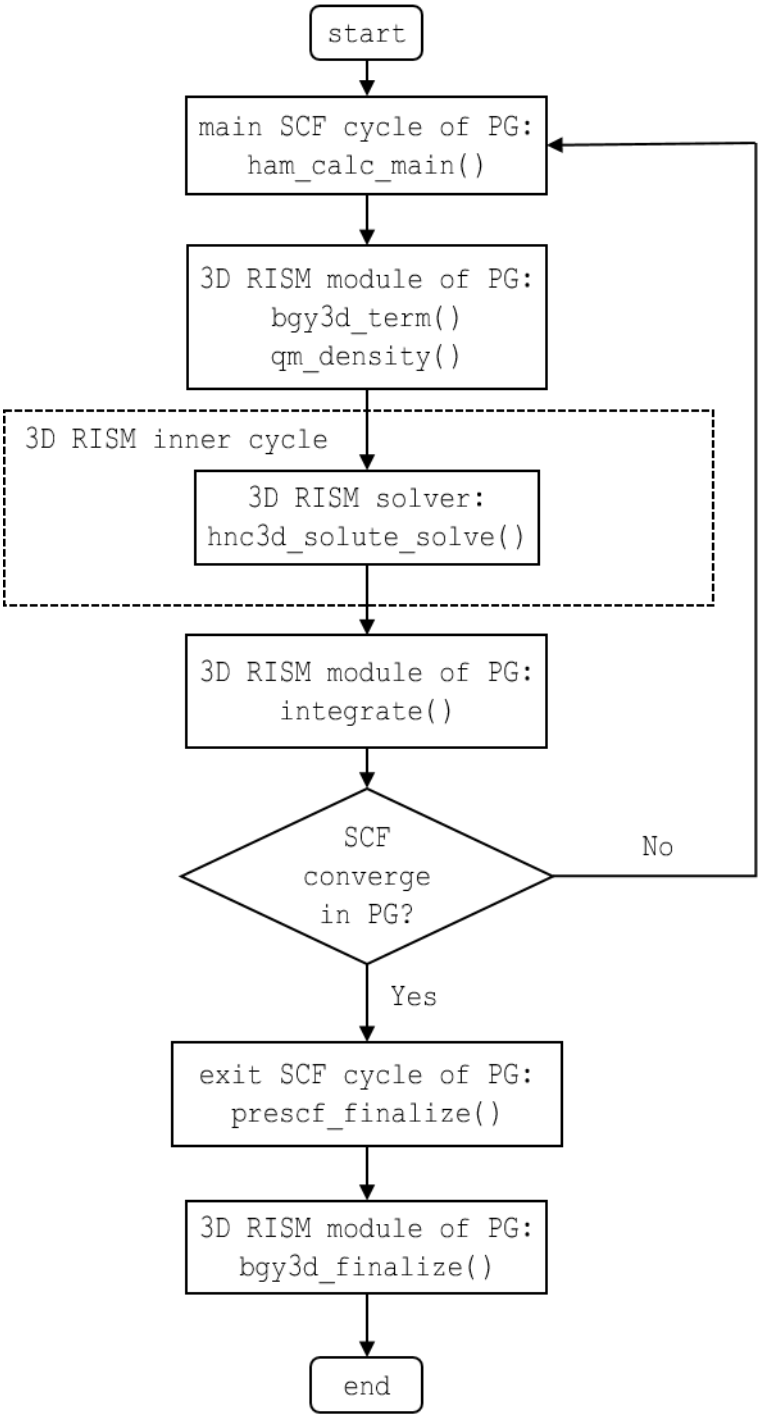


Figure A1. Flowchart for the calling sequence of the modules in the SCF-RISM implementation of ParaGauss. Each process in the rectangle is labeled with the real name of the relevant subroutines. The details for the 3D RISM inner cycle can be found in Section 3.2.2. PG represents the density functional program ParaGauss.

Appendix B. Usage of the RISM Implementation of ParaGauss

Overview

The stable version of reference interaction site model (RISM) implementation in combination with the density functional program ParaGauss, has the following features:

- Site-site radial distribution function (RDF) in 1D RISM,
- Spatial distribution function in 3D RISM,
- Closure relations: HNC, KH, and GF expressions.
- Excess chemical potential both in 1D and 3D RISM,
- Partial molar volume (PMV) corrections available in 1D and 3D RISM,
- Repulsive bridge correction (RBC) evaluated with the thermodynamic perturbation theory (TPT) available in 1D RISM,
- Charged solutes described by molecular mechanics in 1D RISM,
- Full SCF-RISM calculation in combination with ParaGauss, available in 3D RISM for neutral solutes.

The developer version also allows:

- Charged solutes described by molecular mechanics in 3D RISM,
- PSE*n* closure relations.
- Import electron density data from other QM programs into 3D RISM solver,
- Geometry optimization for solute molecules using MM-RISM and PT1-RISM free energy functionals in combination with ParaTools.

Compiling the RISM solver

Those prerequisites of libraries and software packages are needed to compile and run the RISM solver:

- GNU Compiler Collection (GCC)
- GNU Ubiquitous Intelligent Language for Extensions (Guile)
- Open MPI
- parallel FFTW library (libfftw3-mpi3)
- Portable, Extensible Toolkit for Scientific Computation (PETSc) library
- Python

The compilation and installation of the PETSc library also require a lot of prerequisite libraries of which more information can be found from the PETSc manual.[108] The following versions

of the packages have been tested and used for production in this thesis; `GCC 4.7.2, Guile-2.0, Open MPI 1.4.5, libfftw3-mpi3 3.3.2-3.1, libpetsc3.2-dev_3.2.dfgs-6,` and `python 2.7.3`.

The environment variables for PETSc library (shown for the bash-console here) need to be correctly set before compiling the RISM solver. In bash, one can execute the commands:

```
export PETSC_DIR=<path-to-petsc-library>
export PETSC_ARCH=linux-gnu-c-opt
```

or write them into the user's "`.bashrc`" file. After that one can start the compilation from the source code directory of the RISM

```
cd <path-to-RISM-folder>
make
```

to generate the binary executable `bgy3d` there. The command "`make distclean`" can be used prior to a full re-compilation.

The executable `bgy3d` is a guile interpreter. To run the RISM solver, one needs to modify the script `<path-to-RISM-folder>/guile/runbgy.scm` by setting the path in the header as

```
#!<path-to-RISM-folder>/bgy3d \s
!#
```

and adapting the line

```
(set! %load-path (cons "<path-to-RISM-folder>" %load-path)))
```

It is strongly suggested to run the regression test after compilation by typing "`make test-all`" to check the results for a few test cases in comparison with those obtained in this thesis. This regression test depends on using the Python binding for PETSc which called `petsc4py`.

Compiling the RISM module in ParaGauss

The RISM module in ParaGauss is to be compiled by setting the tags in the Makefile of ParaGauss:

```
WITH_GUILE = 1
WITH_BGY3D = 1
```

It is also required to set a link to the directory of the RISM source code where the libraries `libbgy3d.a` and `libbgy3d.so.1` can be located. This can be done by typing:

```
cd <path-to-paragauss-folder>
ln –sf <path-to-RISM-folder> ./bgy3d
```

Then type "`make`" to compile the ParaGauss program and later "`make link`" to link the ParaGauss executable `guile-qm` to the RISM solver library.

Running the SCF-RISM Calculation

The script `<path-to-RISM-folder>/test-qm/runqm` is used to run the SCF-RISM calculation, in which one needs to set the correct interpreter:

```
#!<path-to-paragauss-folder>/guile-qm
!#
```

and the paths:

```
(set! %load-path (cons "<path-to-RISM-folder>" %load-path))
(set! %load-path (cons "<path-to-paragauss-folder>" %load-path)))
```

An example of the input files "`hcl.scm`" for SCF-RISM calculation is:

```
;;
;; HCl, ECP, RKS.
;;
((operations
  (operations_symm #t)
  (operations_integral #t)
  (operations_scf #t)
  (operations_dipole #t)
  (operations_properties #t))
 (main_options
  (integrals_on_file #f)                ; This is faster
  (relativistic "false")                ; This is an ECP calculation
  (spin_restricted #t))                 ; HCl is closed shell
 (geo
  (units angstrom)
  ("H" (0.6285 0.0 0.0))
  ("Cl" (-0.6285 0.0 0.0)))
 (mixing (chmix 0.5) (start_after_cycle 5))
 (grid (sym_reduce #t) (weight_grads #t))
 (rep 2 (gridatom (nrad 30) (nang 131)))
 (xc_control (xc "pbe"))
 (bas "nwchem" "H" "crenbl_ecp" "ahlrichs_coulomb_fitting")
 (ecp "nwchem" "Cl" "crenbl_ecp" "ahlrichs_coulomb_fitting")
 (properties
  (plot_orbitals #t))
 (orbital_plot
  (n_input_lines 0)
  (density_plot #t)
  (density_tot #t)))

;;;
;;; RISM input  follows. The parameter solute specifies  the table to
;;; take the  site-specific force field  parameters. The names  of the
;;; sites are those specified in PG input above:
;;;
((solute "hydrogen chloride")
 (solvent "water")
 (L 10.0)                               ; box size, A
 (N 96)                                 ; grid dimension
 (rho 0.033427745)                    ; solvent number density, A^-3
 (beta 1.6889)                        ; inverse temperature, kcal^-1
 (closure KH))
```

The first half of the input file is for running the QM calculation and the second half is for 3D

RISM. The force field parameters for the solute and solvent molecules are listed in the database files "<path-to-RISM-folder>/guile/solutes.scm" and "<path-to-RISM-folder>/guile/force-fields.scm". It is also possible to update and add more molecular systems by following the syntax in these files. The input file only contains the most important parameters for running 3D RISM calculations, e.g. solute and solvent names, grid dimension, solvent number density, temperature and closure relation option. More settings can be given with the list pairs:

```
    (norm-tol 1.0e-7)                ; convergence threshold
    (max-iter 1500)                  ; max number of iterations
    (damp-start 1.0)                 ; scaling factor
    (lambda 0.02)                   ; Mixing parameter for Picard iteration
scheme
    (bond-length-thresh 1.0)   ; scale for covalent bond autodetection
    (closure HNC)                    ; HNC, KH or PY
    (derivatives #f)                 ; #t for calculating the derivatives
or #f for not
    (snes-solver "trial")           ; "jager", "newton", "picard", "trial"
```

The SCF-RISM calculation for the example input file "hcl.scm" can be run serially by typing:

```
<path-to-RISM-folder>/test-qm/runqm hcl.scm
```

or in parallel with N processors by using Open MPI:

```
mpirun –np N <path-to-RISM-folder>/test-qm/runqm hcl.scm
```

With the successful ending of the calculation, the normal exit information of both ParaGauss and the RISM solver can be found in the output. Information of the excess chemical potential results is printed with the PMV correction values:

```
Nonlinear solve converged due to CONVERGED_PNORM_RELATIVE
 # Chemical potentials, default is marked with *:
 # mu = 4.833630 kcal (HNC)
 # mu = 2.737230 kcal (KH)*
 # mu = -0.746960 kcal (GF)
 # Calculated partial molar volume (PMV), need to be scaled by kappa:
 # PMV = 18.543616
 # PMV * rho = 0.619871
```

The spatial distribution functions will be stored as binary files as "g*uv*.bin", where *u* and *v* are the indices for the solute and solvent sites starting from 0. The binary files can be dumped to the output by running:

```
<path-to-RISM-folder>/guile/runbgy.scm dump guv.bin
```

References

1. Reichardt, C., *Solvents and Solvent Effects in Organic Chemistry*. Wiley: 2006.

2. Koch, W.; Holthausen, M. C., *A Chemist's Guide to Density Functional Theory*. Wiley-VCH Verlag GmbH: Weinheim, Germany, 2001.

3. Jensen, F., *Introduction to Computational Chemistry*. Wiley: 2007.

4. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*. Wiley: 2005.

5. Hirata, F., *Molecular Theory of Solvation*. Kluwer Academic Publisher: Dordrecht, 2003.

6. Sato, H., A modern solvation theory: quantum chemistry and statistical chemistry. *Phys. Chem. Chem. Phys.* **2013,** *15* (20), 7450-7465.

7. Sato, H., A Combination of Electronic Structure and Liquid-state Theory: RISM–SCF/MCSCF Method. In *Continuum Solvation Models in Chemical Physics: From Theory to Applications*, Mennucci, B.; Cammi, R., Eds. John Wiley & Sons Ltd: Chichester, 2007; pp 593-605.

8. Menschutkin, N., Beiträgen zur Kenntnis der Affinitätskoeffizienten der Alkylhaloide und der organischen Amine. *Z. Physik. Chem.* **1890,** *5*, 589.

9. Gao, J., A priori computation of a solvent-enhanced SN2 reaction profile in water: the Menshutkin reaction. *J. Am. Chem. Soc.* **1991,** *113* (20), 7796-7797.

10. Fradera, X.; Amat, L.; Torrent, M.; Mestres, J.; Constans, P.; Besalú, E.; Martí, J.; Simon, S.; Lobato, M.; Oliva, J. M.; Luis, J. M.; Andrés, J. L.; Solà, M.; Carbó, R.; Duran, M., Analysis of the changes on the potential energy surface of Menshutkin reactions induced by external perturbations. *J. Mol. Struct. Theochem* **1996,** *371*, 171-183.

11. Naka, K.; Sato, H.; Morita, A.; Hirata, F.; Kato, S., RISM-SCF study of the free-energy profile of the Menshutkin-type reaction $NH_3+CH_3Cl \rightarrow NH_3CH_3^+ + Cl^-$ in aqueous solution. *Theor. Chem. Acc.* **1999,** *102* (1-6), 165-169.

12. Gao, J.; Xia, X., A two-dimensional energy surface for a type II SN2 reaction in aqueous solution. *J. Am. Chem. Soc.* **1993,** *115* (21), 9667-9675.

13. Lewis, N. S.; Nocera, D. G., Powering the planet: Chemical challenges in solar energy utilization. *Proc. Natl. Acad. Sci.* **2006,** *103* (43), 15729-15735.

14. Roy, S. C.; Varghese, O. K.; Paulose, M.; Grimes, C. A., Toward Solar Fuels: Photocatalytic Conversion of Carbon Dioxide to Hydrocarbons. *ACS Nano* **2010,** *4* (3), 1259-1278.

15. Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T., The Path Forward for Biofuels and Biomaterials. *Science* **2006,** *311* (5760), 484-489.

16. Choppin, G. R.; Jensen, M. P., Actinides in Solution: Complexation and Kinetics. In *The Chemistry of the Actinide and Transactinide Elements*, Morss, L.; Edelstein, N.; Fuger, J., Eds. Springer Netherlands: 2006; pp 2524-2621.

17. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005,** *105* (8), 2999-3094.

18. Klamt, A., Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995,** *99* (7), 2224-2235.

19. Tomasi, J.; Cancès, E.; Pomelli, C. S.; Caricato, M.; Scalmani, G.; Frisch, M. J.; Cammi, R.; Basilevsky, M. V.; Chuev, G. N.; Mennucci, B., Modern Theories of Continuum Models. In *Continuum Solvation Models in Chemical Physics*, John Wiley & Sons, Ltd: 2007; pp 1-123.

20. Cossi, M.; Rega, N.; Soteras, I.; Blanco, D.; Huertas, O.; Bidon-Chanal, A.; Luque, F. J.; Truhlar, D. G.; Pliego, J. R.; Ladanyi, B. M.; Newton, M. D.; Domcke, W.; Sobolewski, A. L.; Laage, D.; Burghardt, I.; Hynes, J. T.; Persico, M.; Granucci, G.; Huxter, V. M.; Scholes, G. D.; Curutchet, C., Chemical Reactivity in the Ground and the Excited State. In *Continuum Solvation Models in Chemical Physics*, John Wiley & Sons, Ltd: 2007; pp 313-497.

21. Gao, J., Hybrid Quantum and Molecular Mechanical Simulations: An Alternative Avenue to Solvent Effects in Organic Chemistry. *Acc. Chem. Res.* **1996,** *29* (6), 298-305.

22. Gao, J.; Xia, X., A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* **1992,** *258* (5082), 631-635.

23. Leontyev, I.; Stuchebrukhov, A., Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* **2011,** *13* (7), 2613-2626.

24. Hansen, J.-P.; McDonald, I. R., *Theory of Simple Liquids*. 3rd ed.; Academic Press: Burlington, 2006.

25. Mullinax, J. W.; Noid, W. G., Generalized Yvon-Born-Green Theory for Molecular Systems. *Phys. Rev. Lett.* **2009,** *103* (19), 198104.

26. Griebel, M.; Jager, L., The BGY3dM model for the approximation of solvent densities. *J. Chem. Phys.* **2008,** *129* (17), 174511-14.

27. Chandler, D.; Andersen, H. C., Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids. *J. Chem. Phys.* **1972,** *57* (5), 1930-1937.

28. Hirata, F.; Rossky, P. J., An extended rism equation for molecular polar fluids. *Chem. Phys. Lett.* **1981,** *83* (2), 329-334.

29. Hirata, F.; Pettitt, B. M.; Rossky, P. J., Application of an extended RISM equation to dipolar and quadrupolar fluids. *J. Chem. Phys.* **1982,** *77* (1), 509-520.

30. Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V., Solvation Thermodynamics of Organic Molecules by the Molecular Integral Equation Theory: Approaching Chemical Accuracy. *Chem. Rev.* **2015,** *115* (13), 6312-56.

31. Beglov, D.; Roux, B., Numerical solution of the hypernetted chain equation for a solute of arbitrary geometry in three dimensions. *J. Chem. Phys.* **1995,** *103* (1), 360-364.

32. Kovalenko, A.; Hirata, F., Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach. *J. Chem. Phys.* **2000,** *112* (23), 10391-10402.

33. Kovalenko, A.; Hirata, F., Potentials of mean force of simple ions in ambient aqueous solution. II. Solvation structure from the three-dimensional reference interaction site model approach, and comparison with simulations. *J. Chem. Phys.* **2000,** *112* (23), 10403-10417.

34. Howard, J.; Pettitt, B. M., Integral Equations in the Study of Polar and Ionic Interaction Site Fluids. *Journal of Statistical Physics* **2011,** *145* (2), 441-466.

35. Du, Q.; Beglov, D.; Roux, B., Solvation Free Energy of Polar and Nonpolar Molecules in Water: An Extended Interaction Site Integral Equation Theory in Three Dimensions. *J. Phys. Chem. B* **2000,** *104* (4), 796-805.

36. Kovalenko, A.; Kobryn, A. E.; Gusarov, S.; Lyubimova, O.; Liu, X.; Blinov, N.; Yoshida, M., Molecular theory of solvation for supramolecules and soft matter structures: application to ligand binding, ion channels, and oligomeric polyelectrolyte gelators. *Soft Matter* **2012,** *8* (5), 1508-1520.

37. Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A., Plant Biomass Recalcitrance: Effect of Hemicellulose Composition on Nanoscale Forces that Control Cell Wall Strength. *J. Am. Chem. Soc.* **2013,** *135* (51), 19048-19051.

38. Silveira, R. L.; Stoyanov, S. R.; Gusarov, S.; Skaf, M. S.; Kovalenko, A., Supramolecular Interactions in Secondary Plant Cell Walls: Effect of Lignin Chemical Composition Revealed with the Molecular Theory of Solvation. *J. Phys. Chem. Lett.* **2015,** *6* (1), 206-211.

39. Imai, T.; Ohyama, S.; Kovalenko, A.; Hirata, F., Theoretical study of the partial molar volume change associated with the pressure-induced structural transition of ubiquitin. *Protein Sci.* **2007,** *16* (9), 1927-1933.

40. Imai, T.; Oda, K.; Kovalenko, A.; Hirata, F.; Kidera, A., Ligand Mapping on Protein Surfaces by the 3D-RISM Theory: Toward Computational Fragment-Based Drug Design. *J. Am. Chem. Soc.* **2009,** *131* (34), 12430-12440.

41. Nikolić, D.; Blinov, N.; Wishart, D.; Kovalenko, A., 3D-RISM-Dock: A New Fragment-Based Drug Design Protocol. *J. Chem. Theory Comput.* **2012,** *8* (9), 3356-3372.

42. Shapovalov, V.; Truong, T. N.; Kovalenko, A.; Hirata, F., Liquid structure at metal oxide–water interface: accuracy of a three-dimensional RISM methodology. *Chem. Phys. Lett.* **2000,** *320* (1–2), 186-193.

43. Howard, J. J.; Perkyns, J. S.; Choudhury, N.; Pettitt, B. M., Integral Equation Study of the Hydrophobic Interaction between Graphene Plates. *J. Chem. Theory Comput.* **2008,** *4* (11), 1928-1939.

44. Ten-no, S.; Hirata, F.; Kato, S., A hybrid approach for the solvent effect on the electronic structure of a solute based on the RISM and Hartree-Fock equations. *Chem. Phys. Lett.* **1993,** *214* (3–4), 391-396.

45. Kovalenko, A.; Hirata, F., Self-consistent description of a metal--water interface by the Kohn--Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.* **1999,** *110* (20), 10095-10112.

46. Sato, H.; Kovalenko, A.; Hirata, F., Self-consistent field, ab initio molecular orbital and three-dimensional reference interaction site model study for solvation effect on carbon monoxide in aqueous solution. *J. Chem. Phys.* **2000,** *112* (21), 9463-9468.

47. Du, Q.; Wei, D., Solvation and Polarization of the N-Methyl Amine Molecule in Aqueous Solution: A Combined Study of Quantum Mechanics and Integral Equation Theory in Three Dimensions. *J. Phys. Chem. B* **2003,** *107* (48), 13463-13470.

48. Gusarov, S.; Ziegler, T.; Kovalenko, A., Self-Consistent Combination of the Three-Dimensional RISM Theory of Molecular Solvation with Analytical Gradients and the Amsterdam Density Functional Package. *J. Phys. Chem. A* **2006,** *110* (18), 6083-6090.

49. Casanova, D.; Gusarov, S.; Kovalenko, A.; Ziegler, T., Evaluation of the SCF Combination of KS-DFT and 3D-RISM-KH; Solvation Effect on Conformational Equilibria, Tautomerization Energies, and Activation Barriers. *J. Chem. Theory Comput.* **2007,** *3* (2), 458-476.

50. Kloss, T.; Heil, J.; Kast, S. M., Quantum Chemistry in Solution by Combining 3D Integral Equation Theory with a Cluster Embedding Approach. *J. Phys. Chem. B* **2008,** *112* (14), 4337-4343.

51. Yoshida, N., Efficient implementation of the three-dimensional reference interaction site model method in the fragment molecular orbital method. *J. Chem. Phys.* **2014,** *140* (21), 214118.

52. Yoshida, N.; Kiyota, Y.; Hirata, F., The electronic-structure theory of a large-molecular system in solution: Application to the intercalation of proflavine with solvated DNA. *J. Mol. Liq.* **2011,** *159* (1), 83-92.

53. Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F., Solvent Effect on Acidity: A Hybrid Approach Based on the RISM and the Hartree−Fock Equations. *J. Phys. Chem.* **1996,** *100* (4), 1111-1117.

54. Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F., Theoretical study for the basicities of methylamines in aqueous solution: A RISM-SCF calculation of solvation thermodynamics. *Chem. Phys.* **1996,** *203* (1), 53-67.

55. Sato, H.; Hirata, F., The syn-/anti-conformational equilibrium of acetic acid in water studied by the RISM-SCF/MCSCF method1. *J. Mol. Struct. Theochem* **1999,** *461–462*, 113-120.

56. Sato, H.; Hirata, F.; Kato, S., Analytical energy gradient for the reference interaction site model multiconfigurational self-consistent-field method: Application to 1,2-difluoroethylene in aqueous solution. *J. Chem. Phys.* **1996,** *105* (4), 1546-1551.

57. Ishida, T.; Rossky, P. J.; Castner, E. W., A Theoretical Investigation of the Shape and Hydration Properties of Aqueous Urea: Evidence for Nonplanar Urea Geometry. *J. Phys. Chem. B* **2004,** *108* (45), 17583-17590.

58. Sato, H.; Hirata, F.; Myers, A. B., Theoretical Study of the Solvent Effect on Triiodide Ion in Solutions. *J. Phys. Chem. A* **1998,** *102* (11), 2065-2071.

59. Vchirawongkwin, V.; Sato, H.; Sakaki, S., RISM-SCF-SEDD Study on the Symmetry Breaking of Carbonate and Nitrate Anions in Aqueous Solution. *J. Phys. Chem. B* **2010,** *114* (32), 10513-10519.

60. Sato, H.; Sakaki, S., Comparison of Electronic Structure Theories for Solvated Molecules: RISM-SCF versus PCM. *J. Phys. Chem. A* **2004,** *108* (9), 1629-1634.

61. Misin, M.; Fedorov, M. V.; Palmer, D. S., Communication: Accurate hydration free energies at a wide range of temperatures from 3D-RISM. *J. Chem. Phys.* **2015,** *142* (9), 091105.

62. Truchon, J.-F.; Pettitt, B. M.; Labute, P., A Cavity Corrected 3D-RISM Functional for Accurate Solvation Free Energies. *J. Chem. Theory Comput.* **2014,** *10* (3), 934-941.

63. Sergiievskyi, V. P.; Jeanmairet, G.; Leyesque, M.; Borgis, D., Fast Computation of Solvation Free Energies with Molecular Density Functional Theory: Thermodynamic-Ensemble Partial Molar Volume Corrections. *J. Phys. Chem. Lett.* **2014,** *5* (11), 1935-1942.

64. Palmer, D. S.; Frolov, A. I.; Ratkova, E. L.; Fedorov, M. V., Towards a universal method for calculating hydration free energies: a 3D reference interaction site model with partial molar volume correction. *J. Phys.: Condens. Matter* **2010,** *22* (49), 492101.

65. Li, B.; Matveev, A. V.; Rösch, N., Three-dimensional reference interaction site model solvent combined with a quantum mechanical treatment of the solute. *Comput. Theor. Chem.* **2015,** *1070*, 143-151.

66. Rösch, N.; Belling, T.; Grauschopf, T.; Krüger, S.; Nörtemann, F.; Staufer, M.; Mayer, M.; Nasluzov, V. A.; Birkenheuer, U.; Hu, A.; Matveev, A. V.; Shor, A. M.; Fuchs-Rohr, M. S.; Neymann, K. M.; Ganyushin, D. I.; Kerdcharoen, T.; Woiterski, A.; Gordienko, A. B.; Majumder, S.; Huix-Rotllant, M.; Ramakrishnan, R.; Dixit, G.; Nikodem, A.; Soini, T.; Roderus, M. *PARAGAUSS*, Version 4.0; Technische Universität München: Munich, 2012.

67. Tomasi, J.; Persico, M., Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994,** *94* (7), 2027-2094.

68. Klamt, A.; Schuurmann, G., COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2* **1993,** (5), 799-805.

69. Fuchs, M. S. K.; Shor, A. M.; Rösch, N., The hydration of the uranyl dication: Incorporation of solvent effects in parallel density functional calculations with the program PARAGAUSS. *Int. J. Quantum Chem* **2002,** *86* (5), 487-501.

70. Cancès, E.; Mennucci, B.; Tomasi, J., A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997,** *107* (8), 3032-3041.

71. Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J., Solvent Effects. 5. Influence of Cavity Shape, Truncation of Electrostatics, and Electron Correlation on ab Initio Reaction Field Calculations. *J. Phys. Chem.* **1996,** *100* (40), 16098-16104.

72. Frenkel, D.; Smit, B., *Understanding Molecular Simulation*. 2nd ed.; Academic Press: San Diego, 2002.

73. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M., UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992,** *114* (25), 10024-10035.

74. Sambasivarao, S. V.; Acevedo, O., Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. *J. Chem. Theory Comput.* **2009,** *5* (4), 1038-1050.

75. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

76. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

77. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*, Pullman, B., Ed. Springer Netherlands: 1981; Vol. 14, pp 331-342.

78. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. *J. Phys. Chem.* **1987,** *91* (24), 6269-6271.

79. Ren, P.; Ponder, J. W., Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **2002,** *23* (16), 1497-1506.

80. Rizzo, R. C.; Jorgensen, W. L., OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999,** *121* (20), 4827-4836.

81. Marx, D.; Hutter, J., Ab-initio Molecular Dynamics: Theory and Implementation. In *Modern Methods and Algorithms of Quantum Chemistry*, Grotendorst, J., Ed. John von Neumann Institute for Computing: Jülich, 2000; Vol. 1, pp 301-449.

82. Nobelprize.org *The Nobel Prize in Chemistry 2013 - Advanced Information*. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/advanced.html (accessed 6 Sep).

83. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935,** *3* (5), 300-313.

84. Jager, L. Fluid Density Approximation for an Implicit Solvent Model. Dissertation, Rheinischen Friedrich-Wilhelms-Universität Bonn, Bonn, 2007.

85. Ornstein, L. S.; Zernike, F., Acculental deviations of density and opalescence at the critical point of a simple substance. *Proc. K. Ned. Akad. Wet Amsterdam* **1914,** *17*, 793-806.

86. Singer, S. J.; Chandler, D., Free energy functions in the extended RISM approximation. *Mol. Phys.* **1985,** *55* (3), 621-625.

87. Hirata, F.; Rossky, P. J.; Pettitt, B. M., The interionic potential of mean force in a molecular polar solvent from an extended RISM equation. *J. Chem. Phys.* **1983,** *78* (6), 4133-4144.

88. Joung, I. S.; Luchko, T.; Case, D. A., Simple electrolyte solutions: Comparison of DRISM and molecular dynamics results for alkali halide solutions. *J. Chem. Phys.* **2013,** *138* (4), 044103.

89. Perkyns, J. S.; Pettitt, B. M., A dielectrically consistent interaction site theory for solvent—electrolyte mixtures. *Chem. Phys. Lett.* **1992,** *190* (6), 626-630.

90. Gendre, L.; Ramirez, R.; Borgis, D., Classical density functional theory of solvation in molecular solvents: Angular grid implementation. *Chem. Phys. Lett.* **2009,** *474* (4–6), 366-370.

91. Zhao, S.; Ramirez, R.; Vuilleumier, R.; Borgis, D., Molecular density functional theory of solvation: From polar solvents to water. *J. Chem. Phys.* **2011,** *134* (19), 194102.

92. Ishizuka, R.; Yoshida, N., Application of efficient algorithm for solving six-dimensional molecular Ornstein-Zernike equation. *J. Chem. Phys.* **2012,** *136* (11), 114106.

93. Ishizuka, R.; Yoshida, N., Extended molecular Ornstein-Zernike integral equation for fully anisotropic solute molecules: Formulation in a rectangular coordinate system. *J. Chem. Phys.* **2013,** *139* (8), 084119.

94. Beglov, D.; Roux, B., An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* **1997,** *101* (39), 7821-7826.

95. Chandler, D.; McCoy, J. D.; Singer, S. J., Density functional theory of nonuniform polyatomic systems. I. General formulation. *J. Chem. Phys.* **1986,** *85* (10), 5971-5976.

96. Kovalenko, A.; Hirata, F., Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach. *Chem. Phys. Lett.* **1998,** *290* (1–3), 237-244.

97. Kovalenko, A.; Hirata, F., Potential of Mean Force between Two Molecular Ions in a Polar Molecular Solvent: A Study by the Three-Dimensional Reference Interaction Site Model. *J. Phys. Chem. B* **1999,** *103* (37), 7942-7957.

98. Kast, S. M.; Kloss, T., Closed-form expressions of the chemical potential for integral equation closures with certain bridge functions. *J. Chem. Phys.* **2008,** *129* (23), 236101.

99. Chandler, D.; Singh, Y.; Richardson, D. M., Excess electrons in simple fluids. I. General equilibrium theory for classical hard sphere solvents. *J. Chem. Phys.* **1984,** *81* (4), 1975-1982.

100. Ratkova, E. L.; Fedorov, M. V., On a relationship between molecular polarizability and partial molar volume in water. *J. Chem. Phys.* **2011,** *135* (24), -.

101. Ratkova, E. L.; Chuev, G. N.; Sergiievskyi, V. P.; Fedorov, M. V., An Accurate Prediction of Hydration Free Energies by Combination of Molecular Integral Equations Theory with Structural Descriptors. *J. Phys. Chem. B* **2010,** *114* (37), 12068-12079.

102. Palmer, D. S.; Sergiievskyi, V. P.; Jensen, F.; Fedorov, M. V., Accurate calculations of the hydration free energies of druglike molecules using the reference interaction site model. *J. Chem. Phys.* **2010,** *133* (4), 044104.

103. Frolov, A. I.; Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V., Hydration Thermodynamics Using the Reference Interaction Site Model: Speed or Accuracy? *J. Phys. Chem. B* **2011,** *115* (19), 6011-6022.

104. Chong, S.-H.; Ham, S., Thermodynamic-Ensemble Independence of Solvation Free Energy. *J. Chem. Theory Comput.* **2015**, 378-380.

105. Yokogawa, D.; Sato, H.; Sakaki, S., New generation of the reference interaction site model self-consistent field method: Introduction of spatial electron density distribution to the solvation theory. *J. Chem. Phys.* **2007,** *126* (24), 244504.

106. Matveev, A.; Li, B.; Rösch, N., Uranyl Solvation by a Three-Dimensional Reference Interaction Site Model. *J. Phys. Chem. A* **2015,** *119* (32), 8702-13.

107. Lee, L. L., Hard sphere properties obtained from a consistent closure. *J. Chem. Phys.* **1999,** *110* (15), 7589-7590.

108. Balay, S.; Adams, M. F.; Brown, J.; Brune, P.; Buschelman, K.; Eijkhout, V.; Gropp, W. D.; Kaushik, D.; Knepley, M. G.; McInnes, L. C.; Rupp, K.; Smith, B. F.; Zhang, H. *PETSc Users Manual*; ANL-95/11 - Revision 3.4; Argonne National Laboratory: 2013.

109. Frigo, M.; Johnson, S. G., The Design and Implementation of FFTW3. *Proc. IEEE* **2005,** *93* (2), 216--231.

110. Ng, K.-C., Hypernetted chain solutions for the classical one-component plasma up to Gamma = 7000. *J. Chem. Phys.* **1974,** *61* (7), 2680-2689.

111. Ewald, P. P., Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik* **1921,** *369* (3), 253-287.

112. Kohn, W.; Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965,** *140* (4A), A1133-A1138.

113. H.Sato; D.Yokogawa; S.Sakaki, Polyatomic molecules in condensed phase: bond order index and solvation energy studied by RISM-SCF theory. *Condens. Matter Phys.* **2007**, 373-380.

114. Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Physical Review* **1964,** *136* (3B), B864-B871.

115. Pettitt, B. M.; Rossky, P. J., Integral equation predictions of liquid state structure for waterlike intermolecular potentials. *J. Chem. Phys.* **1982,** *77* (3), 1451-1457.

116. Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J.; Kovalenko, A., Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2010,** *6* (3), 607-624.

117. Joung, I. S.; Cheatham, T. E., Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008,** *112* (30), 9020-9041.

118. Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A., Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2009,** *5* (2), 350-358.

119. Linstrom, P. J.; Mallard, W. G., *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. National Institute of Standards and Technology: Gaithersburg MD, 2015.

120. Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J., Free Energies of Hydration from a Generalized Born Model and an All-Atom Force Field. *J. Phys. Chem. B* **2004,** *108* (41), 16264-16270.

121. Hoffgaard, F.; Heil, J.; Kast, S. M., Three-Dimensional RISM Integral Equation Theory for Polarizable Solute Models. *J. Chem. Theory Comput.* **2013,** *9* (11), 4718-4726.

122. Kast, S.; Heil, J.; Güssregen, S.; Schmidt, K. F., Prediction of tautomer ratios by embedded-cluster integral equation theory. *J. Comput. Aided Mol. Des.* **2010,** *24* (4), 343-353.

123. Gusarov, S.; Pujari, B. S.; Kovalenko, A., Efficient treatment of solvation shells in 3D molecular theory of solvation. *J. Comput. Chem.* **2012,** *33* (17), 1478-1494.

124. Guilbaud, P.; Wipff, G., Force field representation of the UO_2^{2+} cation from free energy MD simulations in water. Tests on its 18-crown-6 and NO_3^- adducts, and on its calix[6]arene^{6-} and CMPO complexes. *J. Mol. Struct. Theochem* **1996,** *366* (1–2), 55-63.

125. Pomogaev, V.; Tiwari, S. P.; Rai, N.; Goff, G. S.; Runde, W.; Schneider, W. F.; Maginn, E. J., Development and application of effective pairwise potentials for UO_2^{n+}, NpO_2^{n+}, PuO_2^{n+}, and AmO_2^{n+} (n = 1, 2) ions with water. *Phys. Chem. Chem. Phys.* **2013,** *15* (38), 15954-15963.

126. Rai, N.; Tiwari, S. P.; Maginn, E. J., Force Field Development for Actinyl Ions via Quantum Mechanical Calculations: An Approach to Account for Many Body Solvation Effects. *J. Phys. Chem. B* **2012,** *116* (35), 10885-10897.

127. Kerisit, S.; Liu, C., Structure, Kinetics, and Thermodynamics of the Aqueous Uranyl(VI) Cation. *J. Phys. Chem. A* **2013,** *117* (30), 6421-6432.

128. Sato, H.; Hirata, F., Ab initio study of water. II. Liquid structure, electronic and thermodynamic properties over a wide range of temperature and density. *J. Chem. Phys.* **1999,** *111* (18), 8545-8555.

129. Chaffey-Millar, H.; Nikodem, A.; Matveev, A. V.; Krüger, S.; Rösch, N., Improving Upon String Methods for Transition State Discovery. *J. Chem. Theory Comput.* **2012,** *8* (2), 777-786.

130. Nikodem, A.; Matveev, A. V.; Zheng, B.-X.; Rösch, N., Efficient Two-Step Procedures for Locating Transition States of Surface Reactions. *J. Chem. Theory Comput.* **2013,** *9* (1), 588-599.

131. Moskaleva, L. V.; Krüger, S.; Spörl, A.; Rösch, N., Role of Solvation in the Reduction of the Uranyl Dication by Water: A Density Functional Study. *Inorg. Chem.* **2004,** *43* (13), 4080-4090.

132. Bühl, M.; Diss, R.; Wipff, G., Coordination Environment of Aqueous Uranyl(VI) Ion. *J. Am. Chem. Soc.* **2005,** *127* (39), 13506-13507.

133. Cao, Z.; Balasubramanian, K., Theoretical studies of $UO_2(H_2O)_n^{2+}$,$NpO_2(H_2O)_n^{+}$, and $PuO_2(H_2O)_n^{2+}$ complexes (n=4–6) in aqueous solution and gas phase. *J. Chem. Phys.* **2005,** *123* (11), 114309.

134. Tiwari, S. P.; Rai, N.; Maginn, E. J., Dynamics of actinyl ions in water: a molecular dynamics simulation study. *Phys. Chem. Chem. Phys.* **2014,** *16* (17), 8060-8069.

135. Nikitin, A. M.; Lyubartsev, A. P., New six-site acetonitrile model for simulations of liquid acetonitrile and its aqueous mixtures. *J. Comput. Chem.* **2007,** *28* (12), 2020-2026.

136. Venables, D. S.; Schmuttenmaer, C. A., Far-infrared spectra and associated dynamics in acetonitrile–water mixtures measured with femtosecond THz pulse spectroscopy. *J. Chem. Phys.* **1998,** *108* (12), 4935-4944.

137. Moret, M.-E.; Tavernelli, I.; Rothlisberger, U., Combined QM/MM and Classical Molecular Dynamics Study of $[Ru(bpy)_3]^{2+}$ in Water. *J. Phys. Chem. B* **2009,** *113* (22), 7737-7744.

138. Dunlap, B. I.; Rösch, N., The Gaussian-Type Orbitals Density-Functional Approach to Finite Systems. *Adv. Quantum Chem.* **1990,** *21*, 317-339.

139. Rösch, N.; Krüger, S.; Mayer, M.; Nasluzov, V. A., The douglas-kroll-hess approach to relativistic density functional theory: Methodological aspects and applications to metal

complexes and clusters. In *Recent Developments and Applications of Modern Density Functional Theory*, Seminario, J. M., Ed. Elsevier: 1996; Vol. Volume 4, pp 497-566.

140. Kaltsoyannis, N.; Hay, P. J.; Li, J.; Blaudeau, J.-P.; Bursten, B., Theoretical Studies of the Electronic Structure of Compounds of the Actinide Elements. In *The Chemistry of the Actinide and Transactinide Elements*, Morss, L.; Edelstein, N.; Fuger, J., Eds. Springer Netherlands: 2011; pp 1893-2012.

141. Görling, A.; Trickey, S. B.; Gisdakis, P.; Rösch, N., A Critical Assessment of Density Functional Theory with Regard to Applications in Organometallic Chemistry. In *Organometallic Bonding and Reactivity*, Brown, J.; Hofmann, P., Eds. Springer Berlin Heidelberg: 1999; Vol. 4, pp 109-163.

142. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988,** *38* (6), 3098-3100.

143. Perdew, J. P., Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986,** *33* (12), 8822-8824.

144. Moskaleva, L. V.; Matveev, A. V.; Dengler, J.; Rösch, N., The heat of formation of gaseous PuO_2^{2+} from relativistic density functional calculations. *Phys. Chem. Chem. Phys.* **2006,** *8* (32), 3767-3773.

145. Moskaleva, L. V.; Matveev, A. V.; Krüger, S.; Rösch, N., The Heat of Formation of the Uranyl Dication: Theoretical Evaluation Based on Relativistic Density Functional Calculations. *Chem. Eur. J.* **2006,** *12* (2), 629-634.

146. Kremleva, A.; Zhang, Y.; Shor, A. M.; Krüger, S.; Joseph, C.; Raditzky, B.; Schmeide, K.; Sachs, S.; Bernhard, G.; Rösch, N., Uranyl(VI) Complexation by Sulfonate Ligands: A Relativistic Density Functional and Time-Resolved Laser-Induced Fluorescence Spectroscopy Study. *Eur. J. Inorg. Chem.* **2012,** *2012* (22), 3636-3644.

147. Koitz, R.; Soini, T. M.; Genest, A.; Trickey, S. B.; Rösch, N., Scalable properties of metal clusters: A comparative study of modern exchange-correlation functionals. *J. Chem. Phys.* **2012,** *137* (3), 034102.

148. Dunlap, B. I.; Rösch, N.; Trickey, S. B., Variational fitting methods for electronic structure calculations. *Mol. Phys.* **2010,** *108* (21-23), 3167-3180.

149. Perdew, J. P., Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986,** *34* (10), 7406-7406.

150. Dunning, T. H., Jr.; Hay, P. J., Gaussian Basis Sets for Molecular Calculations. In *Methods of Electronic Structure Theory*, Schaefer, H., III, Ed. Springer US: 1977; Vol. 3, pp 1-27.

151. Fernandez Pacios, L.; Christiansen, P. A., Ab initio relativistic effective potentials with spin-orbit operators. I. Li through Ar. *J. Chem. Phys.* **1985,** *82* (6), 2664-2671.

152. Hurley, M. M.; Pacios, L. F.; Christiansen, P. A.; Ross, R. B.; Ermler, W. C., Abinitio relativistic effective potentials with spin-orbit operators. II. K through Kr. *J. Chem. Phys.* **1986,** *84* (12), 6840-6853.

153. LaJohn, L. A.; Christiansen, P. A.; Ross, R. B.; Atashroo, T.; Ermler, W. C., Abinitio relativistic effective potentials with spin–orbit operators. III. Rb through Xe. *J. Chem. Phys.* **1987,** *87* (5), 2812-2824.

154. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R., Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995,** *240* (4), 283-290.

155. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R., Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997,** *97* (1-4), 119-124.

156. Bondi, A., van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68* (3), 441-451.

157. Su, P.; Li, H., Continuous and smooth potential energy surface for conductorlike screening solvation model using fixed points with variable areas. *J. Chem. Phys.* **2009,** *130* (7), 074109.

158. Damianos, K.; Genest, A.; Rösch, N., On the overall catalytic cycle of CO_2 conversion to $HCOO^-$ in solution: A DFT study. *in preparation*.

159. Barone, V.; Cossi, M., Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998,** *102* (11), 1995-2001.

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

Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford CT, 2004.

161. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988,** *37* (2), 785-789.

162. Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H., Energy-adjusted *ab initio* pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta* **1990,** *77* (2), 123-141.

163. Schmid, R.; Miah, A. M.; Sapunov, V. N., A new table of the thermodynamic quantities of ionic hydration: values and some applications (enthalpy-entropy compensation and Born radii). *Phys. Chem. Chem. Phys.* **2000,** *2* (1), 97-102.

164. Sergiievskyi, V. P.; Jeanmairet, G.; Leyesque, M.; Borgis, D., Fast Computation of Solvation Free Energies with Molecular Density Functional Theory: Thermodynamic-Ensemble Partial Molar Volume Corrections. *J. Phys. Chem. Lett.* **2014,** *5* (11), 1935-1942.

165. Kastenholz, M. A.; Hünenberger, P. H., Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J. Chem. Phys.* **2006,** *124* (22), 224501.

166. Kastenholz, M. A.; Hünenberger, P. H., Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J. Chem. Phys.* **2006,** *124* (12), 124106.

167. Lee Warren, G.; Patel, S., Hydration free energies of monovalent ions in transferable intermolecular potential four point fluctuating charge water: An assessment of simulation methodology and force field performance and transferability. *J. Chem. Phys.* **2007,** *127* (6), 064509.

168. Reif, M. M.; Hünenberger, P. H., Computation of methodology-independent single-ion solvation properties from molecular simulations. IV. Optimized Lennard-Jones interaction parameter sets for the alkali and halide ions in water. *J. Chem. Phys.* **2011,** *134* (14), 144104.

169. Kovalenko, A.; Hirata, F., Hydration free energy of hydrophobic solutes studied by a reference interaction site model with a repulsive bridge correction and a thermodynamic perturbation method. *J. Chem. Phys.* **2000,** *113* (7), 2793-2805.

170. Orozco, M.; Marchán, I.; Soteras, I.; Vreven, T.; Morokuma, K.; Mikkelsen, K. V.; Milani, A.; Tommasini, M.; Zoppo, M. D.; Castiglioni, C.; Aguilar, M. A.; Sánchez, M. L.; Martín, M. E.; Galván, I. F.; Sato, H., Beyond the Continuum Approach. In *Continuum Solvation Models in Chemical Physics*, John Wiley & Sons, Ltd: 2007; pp 499-605.

171. Rodnikova, M. N., A new approach to the mechanism of solvophobic interactions. *J. Mol. Liq.* **2007,** *136* (3), 211-213.

172. Dzubiella, J.; Hansen, J.-P., Competition of hydrophobic and Coulombic interactions between nanosized solutes. *J. Chem. Phys.* **2004,** *121* (11), 5514-5530.

173. Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L., Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.* **1981,** *10* (8), 563-595.

174. Kreyszig, E., *Advanced Engineering Mathematics: Maple Computer Guide*. John Wiley & Sons, Inc.: New York, 2000; p 312.

175. Yudanov, I. V.; Nasluzov, V. A.; Neyman, K. M.; Rösch, N., Density functional cluster description of ionic materials: Improved boundary conditions for MgO clusters with the help of cation model potentials. *Int. J. Quantum Chem* **1997,** *65* (5), 975-986.

176. Matveev, A. V.; Neyman, K. M.; Yudanov, I. V.; Rösch, N., Adsorption of transition metal atoms on oxygen vacancies and regular sites of the MgO(001) surface. *Surf. Sci.* **1999,** *426* (1), 123-139.

177. Silva, R.; Nitsche, H., Actinide Environmental Chemistry. *Radiochim. Acta* **1995,** *70*, 377-396.

178. Grenthe, I.; Fuger, J.; Konings, R.; Lemire, R.; Muller, A.; Nguyen-Trung, C.; Wanner, H., *Chemical Thermodynamics of Uranium*. North-Holland: Amsterdam, 1992.

179. Arnold, P. L.; Love, J. B.; Patel, D., Pentavalent uranyl complexes. *Coord. Chem. Rev.* **2009,** *253* (15–16), 1973-1978.

180. Katz, J. J.; Seaborg, G. T.; Morss, L. R., *The Chemistry of the Actinide Elements*. 2 ed.; Chapman and Hall: New York, 1986; Vol. 2.

181. Mahoney, M. W.; Jorgensen, W. L., A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000,** *112* (20), 8910-8922.

182. Bridgeman, O. C.; Aldrich, E. W., Vapor Pressure Tables for Water. *J. Heat Transfer* **1964,** *86* (2), 279-286.

183. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.

184. Bühl, M.; Kabrede, H.; Diss, R.; Wipff, G., Effect of Hydration on Coordination Properties of Uranyl(VI) Complexes. A First-Principles Molecular Dynamics Study. *J. Am. Chem. Soc.* **2006,** *128* (19), 6357-6368.

185. Nichols, P.; Bylaska, E. J.; Schenter, G. K.; de Jong, W., Equatorial and apical solvent shells of the UO_2^{2+} ion. *J. Chem. Phys.* **2008,** *128* (12), 124507.

186. Frick, R. J.; Hofer, T. S.; Pribil, A. B.; Randolf, B. R.; Rode, B. M., Structure and Dynamics of the UO_2^{2+} Ion in Aqueous Solution: An Ab Initio QMCF MD Study†. *J. Phys. Chem. A* **2009,** *113* (45), 12496-12503.

187. Guilbaud, P.; Wipff, G., Hydration of uranyl (UO_2^{2+}) cation and its nitrate ion and 18-crown-6 adducts studied by molecular dynamics simulations. *J. Phys. Chem.* **1993,** *97* (21), 5685-5692.

188. Neuefeind, J.; Soderholm, L.; Skanthakumar, S., Experimental Coordination Environment of Uranyl(VI) in Aqueous Solution. *J. Phys. Chem. A* **2004,** *108* (14), 2733-2739.

189. Wahlgren, U.; Moll, H.; Grenthe, I.; Schimmelpfennig, B.; Maron, L.; Vallet, V.; Gropen, O., Structure of Uranium(VI) in Strong Alkaline Solutions. A Combined Theoretical and Experimental Investigation. *J. Phys. Chem. A* **1999,** *103* (41), 8257-8264.

190. Soderholm, L.; Skanthakumar, S.; Neuefeind, J., Determination of actinide speciation in solution using high-energy X-ray scattering. *Anal. Bioanal. Chem.* **2005,** *383* (1), 48-55.

191. Cornehl, H. H.; Heinemann, C.; Marçalo, J.; de Matos, A. P.; Schwarz, H., The "Bare" Uranyl(2+) Ion, UO_2^{2+}. *Angew. Chem. Int. Ed. Engl.* **1996,** *35* (8), 891-894.

192. Marcus, Y.; Loewenschuss, A., Standard thermodynamic functions of the gaseous actinyl ions MO and for their hydration. *J. Chem. Soc., Faraday Trans. 1* **1986,** *82* (9), 2873-2886.

193. Gibson, J. K.; Haire, R. G.; Santos, M.; Marçalo, J.; Pires de Matos, A., Oxidation Studies of Dipositive Actinide Ions, An^{2+} (An = Th, U, Np, Pu, Am) in the Gas Phase: Synthesis and Characterization of the Isolated Uranyl, Neptunyl, and Plutonyl Ions UO_2^{2+}(g), NpO_2^{2+}(g), and PuO_2^{2+}(g). *J. Phys. Chem. A* **2005,** *109* (12), 2768-2781.

194. Hagberg, D.; Karlström, G.; Roos, B. O.; Gagliardi, L., The Coordination of Uranyl in Water: A Combined Quantum Chemical and Molecular Simulation Study. *J. Am. Chem. Soc.* **2005,** *127* (41), 14250-14256.

195. Shamov, G. A.; Schreckenbach, G., Density Functional Studies of Actinyl Aquo Complexes Studied Using Small-Core Effective Core Potentials and a Scalar Four-Component Relativistic Method. *J. Phys. Chem. A* **2005,** *109* (48), 10961-10974.

196. Gutowski, K. E.; Dixon, D. A., Predicting the Energy of the Water Exchange Reaction and Free Energy of Solvation for the Uranyl Ion in Aqueous Solution. *J. Phys. Chem. A* **2006,** *110* (28), 8840-8856.

197. Bühl, M.; Wipff, G., Insights into Uranyl Chemistry from Molecular Dynamics Simulations. *ChemPhysChem* **2011,** *12* (17), 3095-3105.

198. Farkas, I.; Bányai, I.; Szabó, Z.; Wahlgren, U.; Grenthe, I., Rates and Mechanisms of Water Exchange of UO_2^{2+}(aq) and UO_2(oxalate)F$(H_2O)_2^-$: A Variable-Temperature ^{17}O and ^{19}F NMR Study. *Inorg. Chem.* **2000,** *39* (4), 799-805.

199. Vallet, V.; Wahlgren, U.; Schimmelpfennig, B.; Szabó, Z.; Grenthe, I., The Mechanism for Water Exchange in $[UO_2(H_2O)_5]^{2+}$ and $[UO_2(oxalate)_2(H_2O)]_2^-$, as Studied by Quantum Chemical Methods. *J. Am. Chem. Soc.* **2001,** *123* (48), 11999-12008.

200. Kerisit, S.; Liu, C., Molecular simulation of the diffusion of uranyl carbonate species in aqueous solution. *Geochim. Cosmochim. Acta* **2010,** *74* (17), 4937-4952.

201. Gál, M.; Goggin, P. L.; Mink, J., Vibrational spectroscopic studies of uranyl complexes in aqueous and non-aqueous solutions. *Spectrochimica Acta Part A: Molecular Spectroscopy* **1992,** *48* (1), 121-132.

202. Toth, L. M.; Begun, G. M., Raman spectra of uranyl ion and its hydrolysis products in aqueous nitric acid. *J. Phys. Chem.* **1981,** *85* (5), 547-549.

203. Denning, R. G.; Foster, D. N. P.; Snellgrove, T. R.; Woodwark, D. R., The electronic structure of the uranyl ion. *Mol. Phys.* **1979,** *37* (4), 1089-1107.

204. Denning, R. G.; Snellgrove, T. R.; Woodwark, D. R., The electronic structure of the uranyl ion. *Mol. Phys.* **1976,** *32* (2), 419-442.

205. Matveev, A. V.; Rösch, N., Atomic approximation to the projection on electronic states in the Douglas-Kroll-Hess approach to the relativistic Kohn-Sham method. *J. Chem. Phys.* **2008,** *128* (24), 244102.

206. Tsushima, S., Hydration and Water-Exchange Mechanism of the UO_2^{2+} Ion Revisited: The Validity of the "n + 1" Model. *J. Phys. Chem. A* **2007,** *111* (18), 3613-3617.

207. Bühl, M.; Kabrede, H., Mechanism of Water Exchange in Aqueous Uranyl(VI) Ion. A Density Functional Molecular Dynamics Study. *Inorg. Chem.* **2006,** *45* (10), 3834-3836.

208. Rotzinger, F. P., The Water-Exchange Mechanism of the $[UO_2(OH_2)_5]^{2+}$ Ion Revisited: The Importance of a Proper Treatment of Electron Correlation. *Chem. Eur. J.* **2007,** *13* (3), 800-811.

209. Kumar, B.; Llorente, M.; Froehlich, J.; Dang, T.; Sathrum, A.; Kubiak, C. P., Photochemical and Photoelectrochemical Reduction of CO_2. *Annu. Rev. Phys. Chem.* **2012,** *63* (1), 541-569.

210. Duan, L.; Tong, L.; Xu, Y.; Sun, L., Visible light-driven water oxidation-from molecular catalysts to photoelectrochemical cells. *Energy & Environmental Science* **2011,** *4* (9), 3296-3313.

211. Lehn, J.-M.; Ziessel, R., Photochemical generation of carbon monoxide and hydrogen by reduction of carbon dioxide and water under visible light irradiation. *Proc. Natl. Acad. Sci.* **1982,** *79* (2), 701-704.

212. Kelly, J. M.; Vos, J. G., *cis*-$[Ru(bpy)_2(CO)H]^+$—A Possible Intermediate in the Photochemical Production of H_2 from Water Catalyzed by $[Ru(bpy)_3]^{2+}$? *Angew. Chem. Int. Ed. Engl.* **1982,** *21* (8), 628-629.

213. Pugh, J. R.; Bruce, M. R. M.; Sullivan, B. P.; Meyer, T. J., Formation of a metal-hydride bond and the insertion of carbon dioxide. Key steps in the electrocatalytic reduction of carbon dioxide to formate anion. *Inorg. Chem.* **1991,** *30* (1), 86-91.

214. Creutz, C.; Chou, M. H.; Hou, H.; Muckerman, J. T., Hydride Ion Transfer from Ruthenium(II) Complexes in Water: Kinetics and Mechanism. *Inorg. Chem.* **2010,** *49* (21), 9809-9822.

215. Truhlar, D. G.; Cramer, C. J.; Lewis, A.; Bumpus, J. A., Molecular Modeling of Environmentally Important Processes: Reduction Potentials. *J. Chem. Educ.* **2004,** *81* (4), 596.