Three-dimensional reference interaction site model solvent combined with a quantum mechanical treatment of the solute

Bo Li [a], Alexei V. Matveev [a], Notker Rösch [b,c,]*

[a] Department Chemie, Technische Universität München, 85747 Garching, Germany
[b] Department Chemie and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany
[c] Institute of High Performance Computing, Agency for Science, Technology and Research, 1 Fusionopolis Way, Connexis #16-16, Singapore 138632, Singapore

ABSTRACT

We implemented and applied a hybrid approach for modeling the solvation of molecules where we combined a three-dimensional reference interaction site model (3D RISM) with quantum mechanical (QM) calculations of the solute. The electrostatic potential induced by the solute is derived directly from its electron density. For neutral solutes, we analyzed the accuracy of calculated solvation free energies which is mainly determined by the cavity formation energy. In an aqueous medium the solute electronic structure relaxation also has a noticeable influence on the results of SCF-RISM calculations. We apply a known partial molar volume correction for which we give an alternative interpretation. The results of this hybrid model agree well with experiment and results of a polarizable continuum model (PCM). As a model with an atomic solvent representation, RISM accounts for the effect of discrete hydrogen bonds which in PCM models are included on average only.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The solvent environment of a molecule can considerably affect its properties and thus also change the products as well as the rate of its chemical reactions and even the mechanism of such transformations. Therefore, modeling solvation effects is of central importance for understanding many chemical reactions of organic and inorganic species as well as for industrial applications [1]. Examples of pertinent application fields, also of industrial importance, are catalytic reactions for the production of hydrogen [2], the conversion of CO_2 [3], transformations of biomass [4], and the complex chemistry of actinides in aqueous solution, which is highly relevant for the nuclear fuel cycle and radioactive waste disposal [5].

In support of current experimental efforts in various areas of chemistry and materials science in aqueous media, computational studies are needed to examine or predict species and processes at the atomic level. One of the two most popular strategies in computational chemistry for studying solution systems is the combination of quantum mechanics (QM) calculations with implicit solvation models, e.g., the polarizable continuum model (PCM)

[6] and its variants based the conductor-like solvation model (COSMO) [7]. Yet, despite the importance of solvation models, these implicit methods do not provide insight into the structure and the statistical fluctuations of the solvent. Another approach to solvation is the so-called explicit solvation model, where one treats solute and solvent equivalently by carrying out molecular dynamics (MD) simulations from first principles (*ab initio* MD – AIMD) or MD simulations based on a force-field description (molecular mechanics – MM) of both the solute and the solvent medium (MM–MD). However, the statistical sampling of large numbers of solvent configurations entails significant computational cost. Therefore QM dynamic simulations are restricted to models with a comparably small numbers of solvent molecules [1]. The MM solvent description of MM–MD or hybrid QM/MM methods [8,9] reduce the computational cost. Yet, the results of statistical sampling are only as good as the effective force field used to parameterize the solvent–solvent and solute–solvent interactions. Also, to compete with continuum models in reproducing electrostatic interactions, one has to account for the polarizability of the solvent medium [10].

Given the difficulties mentioned above, strategies which combine QM methods with integral equation theory of molecular liquids [1,11] are regarded as a promising direction for studying ions and complexes and their chemical reactions in aqueous solution with sufficient accuracy and at affordable cost. One successful approach of this type is the reference interaction site model (RISM)

* Corresponding author at: Department Chemie and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany.
E-mail addresses: alexei.matveev@gmail.com (A.V. Matveev), roesch@mytum.de (N. Rösch).

[12] which yields solvation (free) energies and also averaged solvent structures via site–site radial distribution functions (RDF). The hybrid approach of RISM together with a self-consistent field (SCF) QM solute description was first proposed at the level of Hartree–Fock theory [13]. Later-on Kohn–Sham density functional theory (KS-DFT) was combined with 3D RISM to examine the electronic structure of a metal–water interface [14]. Thus far, several implementations [15–21] combine RISM theory with a state-of-the-art QM method to provide an alternative approach for classical implicit and explicit solvation models. Hybrid SCF-RISM calculations have been applied for calculating the structure of molecular systems in solution, their acidity and basicity, their excited states, as well as chemical equilibria and chemical reactions [1,22].

Mainly two factors determine the solvent effect in the SCF-RISM method, the contributions of the solute electronic relaxation and the excess chemical potential [1]. The excess chemical potential, which measures the free energy change of coupling a solute molecule to the solvent, serves as the approximation of the solvation free energy in MM-RISM methods. The excess chemical potential calculated from 3D RISM is well known [1] 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 [23]. To overcome this weakness, several correction schemes have been proposed for the solvation free energy, including the repulsive bridge correction (RBC) [1,23], the cavity formation energy correction [24], and the partial molar volume (PMV) correction [25–27]. These correction methods are mainly applied in MM-RISM calculations, in which the solute electronic structure is assumed not to be relaxed. Yet, as in the PCM method [28], the energy due to the electronic relaxation of the solute is not negligible in the SCF-RISM method. The change of the solvation free energy due to solute polarization is almost proportional to this term [22,29], but the influence of the solute polarization on the accuracy of the solvation free energy from SCF-RISM calculations has not been studied in detail yet.

Therefore, we will examine in the following the accuracy of the solvation free energy in hybrid QM RISM calculations for a test set of 43 neutral organic solute molecules in aqueous solution. This test set comprises hydrocarbon chains and aromatic rings, and various organic molecules with functional groups. A list of all species is given as Supplementary Material (SM). We explore the effect of solute electronic relaxation in stepwise fashion. Starting from MM-RISM as a reference, we introduce the polarization of the solute in the solvent medium, first by applying first-order perturbation theory (PT1-RISM), going on to a second-order perturbation theory treatment (PT2-RISM), and finally moving to the full SCF-RISM relaxation. For all method variants, we compare the performance of the recently reported PMV correction term [26] to the solvation free energy. To analyze the accuracy the RISM approach, we compare these results to those of PCM calculations and to experiment.

2. Methodology

2.1. 3D-RISM theory

As a recent development of the liquid state theory, the 3D RISM provides a discrete description of the solvent around the solute, using solute–solvent correlation functions as fundamental variables. The relevant theory has thoroughly been discussed [1,14,22,30]. Here we sketch only the fundamental formulae of the solute–solvent 3D RISM integral equation, in the form of a convolution [1]:

$$\mathbf{h}^{uv} = \mathbf{c}^{uv} * \boldsymbol{\chi}^{v} \tag{1}$$

$\mathbf{h}^{uv}$ and $\mathbf{c}^{uv}$ are (rectangular) matrices of the 3D total correlation functions $h_{i\alpha}^{uv}(r)$ and 3D direct correlation functions $c_{i\alpha}^{uv}(r)$, both unknowns of the equation. The superscripts u and v label the solute and the solvent, respectively. The subscript α enumerates the solvent site interacting with the whole solute. In Eq. (1), the solvent is represented by its (square symmetric) susceptibility matrix $\boldsymbol{\chi}^v = \boldsymbol{\omega}^v + \rho \mathbf{h}^{vv}$ which can be obtained self-consistently from the pure solvent RISM equation [1,31]; in practice one operates with the dimensionless Fourier transform $\tilde{\chi}^v(k)$ of the solvent susceptibility. The low-k limit $\tilde{\chi}_{\alpha\gamma}^v(0) = \rho\kappa/\beta$ encodes the solvent compressibility κ; the next leading $o(k^2)$ term is calibrated by the dielectric response of the solvent [32]. $\beta = 1/T$ is the inverse of the temperature T.

Similar to the 1D RISM, a closure relation is needed to render Eq. (1) solvable. In all RISM calculations we use the closure proposed by Kovalenko and Hirata (KH) [14]: $h_\alpha = f(-\beta u_\alpha + h_\alpha - c_\alpha)$ with $f(x) = \exp(x) - 1$ for $x \leqslant 0$ and $f(x) = x$ otherwise. Here $u_\alpha(\vec{r})$ is the interaction potential between solvent site α and the solute.

The excess chemical potential of solvation is given by the distribution functions. For the KH closure, it has the analytical form [1,14]:

$$\mu = \frac{\rho}{\beta} \int \sum_{\alpha=1}^{N_v} [\Theta(-h_\alpha)h_\alpha^2/2 - c_\alpha - c_\alpha h_\alpha/2]d^3r \tag{2}$$

Here $\Theta(x)$ is the Heaviside step function which ensures that the terms h^2 contribute only in the depletion regions, $h_\alpha < 0$ [14]. This expression, together with the linear convolution relation, Eq. (1), between the total and direct correlation functions h and c, leads to seriously overestimated cavity formation energies [25,33]. Indeed, by assuming a simple liquid exposed to a potential sufficiently repulsive within an extended volume V_0 and approximating $h \approx -1$ and $c \approx -z^{-1}$ within the volume V_0, one obtains:

$$\beta\mu \approx \rho V_0(z^{-1} + 1)/2. \tag{3}$$

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, $z \approx 1/15.8$, at normal conditions [34]. With the KH closure and a modified TIP3P water model [35,36], RISM theories yield a comparable value, $1/15.2$. Thus the asymptotic RISM expression for the formation energy of the solute cavity suggests that, when forming the cavity at normal conditions, the energy $a = \beta^{-1}(z^{-1} + 1)/2 \approx 5$ kcal/mol is required to displace each of ρV_0 water solvent molecules. 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 cal/mol per displaced water molecule. Because of this low prefactor, for most chemical applications the water surface tension of 104 cal/mol/Å^2 is significantly more important for medium-sized cavities of nanosized solutes [37]. Nevertheless, the large error of the RISM expression for the chemical potential cannot be ignored in practical calculations.

The erroneous behavior of Eq. (3) likely is due to the fact that the method relies on the solvent susceptibility χ for the (dense) liquid water also inside the cavity where the solvent density is vanishingly small. When one substitutes in Eq. (3) the compressibility factor by its low-density (ideal gas) limit $z = 1$, the RISM asymptotic expression acquires the correct form $\mu \approx \rho V_0/\beta = pV_0$.

Palmer et al. noted that the partial molar volume (PMV) correlates with the difference between experimental and RISM calculated solvation free energy values. Therefore, they suggested the PMV correction [25],

$$\mu^{PMV} = \mu + (a\rho V_u + b), \tag{4}$$