Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics Letters 381 (2003) 362–367
CHEMICAL PHYSICS LETTERS
www.elsevier.com/locate/cplett

A coupled RISM/MD or MC simulation methodology for solvation free energies

Holly Freedman, Thanh N. Truong *

Department of Chemistry, Henry Eyring Center for Theoretical Chemistry, University of Utah, 315 South 1400 East, Rm. 2020, Salt Lake City, UT 84112, USA

Received 25 March 2003; in final form 5 September 2003
Published online: 24 October 2003

Abstract

We propose a new computational methodology that couples reference interaction site model (RISM) and molecular dynamics (MD) or Monte Carlo (MC) simulation methods for determination of solvation free energies. We employ the RISM formulation of solvation free energy. The correlation functions entering this expression are derived from radial distribution functions supplied by MD or MC simulations, instead of coming from simultaneous solution of the RISM and hyper-netted chain equations. We apply this approach to determining free energies of solvation for several small molecules. Our results are in good agreement with experimental values, demonstrating the potential of this method for applications to larger systems.

1. Introduction

One major obstacle to the computational study of solvated chemical systems has been the difficulty associated with calculating free energies of solvation. This is especially true for studies of large systems such as aqueous biosystems. However the solvation free energy of a biosystem is a significant factor in its total thermochemistry, often contributing to free energy differences by an amount comparable to that due to interactions between the biomolecules themselves. Therefore it is equally important that these solvation free energy differences be calculated accurately when applying computational methods to problems such as rational drug design.

The simplest method for calculating a free energy of solvation is by using a dielectric continuum model, which represents the solvent implicitly. However, its accuracy is often questionable and its application is limited by the availability of the necessary parameters such as atomic radii [1]. A more accurate approach that has been applied to calculating solvation free energies is the use of either molecular dynamics (MD) or Monte Carlo (MC) simulations, for which the solute is surrounded with explicit solvent molecules. Using this explicit solvent approach, solvation free

* Corresponding author. Fax: +1-801-581-4353.
E-mail address: truong@chem.utah.edu (T.N. Truong).

0009-2614/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.cplett.2003.09.108

energy differences between similar compounds can be obtained. The displacement of water molecules surrounding the solute can be accounted for by integrating over a coupling parameter that varies as one solute molecule is gradually 'mutated' into the other using the thermodynamic integration (TI) or free energy perturbation (FEP) technique. Although simulations have been demonstrated to produce acceptable solvation free energy differences for small organic molecules and peptides, these simulation methods are tedious and time consuming. It is difficult to set up the path transforming one compound into the other, in which the two sets of parameters of the two solutes must be linked. Also a large amount of time is required to converge all of the necessary simulations [1,2].

An alternative method for calculating solvation free energies is by using the reference interaction site model (RISM) [3–8]. In the RISM theory, absolute solvation free energies can be expressed as a function of the radial distribution functions and the direct correlation functions. Again, absolute solvation free energies are expressed as integrals over a coupling parameter, λ, in order to create the solute molecule within the solvent. However, within the RISM formalism an exact differential corresponding to $d\Delta\mu^S/d\lambda$ can be derived, and thus integration over the coupling parameter can be done analytically [9,10]. Thus once radial distribution functions and direct correlation functions for a solute have been calculated from RISM theory, its solvation free energy can easily be determined by using a simple expression [11–13]. Despite the simplicity of this formulation, it has been known that RISM can only provide qualitatively correct solvent structure [14–16]. Consequently, our preliminary results indicate that RISM is often not accurate for calculating absolute solvation free energies. However there have been successful applications of RISM to the study of relative solvation free energies [17,18].

The convenience of the expression used to obtain solvation free energies in RISM and the greater accuracy in the radial distribution functions obtained from MD or MC simulations suggest the possibility of coupling these two methodologies. In this study we show that it is possible to extract radial distribution functions from MD or MC simulations and use them in the RISM formulation for free energy of solvation. We then compare our results to those of a RISM/HNC stand-alone calculation. Although by our comparison we cannot definitively conclude that the use of more accurate distribution functions from simulations leads to more accurate solvation free energies, the coupled RISM/simulation methodology provides a viable alternate means to calculate solvation free energy, particularly when stand-alone RISM/HNC encounters convergence difficulties such as for biological systems; at the same time it has provided us with insights into the improvement of stand-alone RISM/HNC absolute solvation free energy results.

RISM has been utilized in combination with MD and MC simulations in the past, but in a different framework and for a different purpose. Free energy profiles of peptides have been found by simulations using RISM calculations to represent the solvent. A stochastic dynamics (SD) simulation was carried out by Pettitt and coworkers, who used look-up tables of RISM-calculated solute/solute correlation functions to calculate solvent contributions to potentials of mean force (PMFs) in the superposition approximation, which could be added on to molecular mechanics force field calculations [19–22]. Additionally, in an MD simulation, Go and coworkers summed free energies of solvation calculated in the RISM formalism with conformational energies to obtain total free energy profiles [23]. Likewise Hirata and coworkers have performed combined MC and RISM calculations for a similar purpose [24,25]. These studies employed the RISM method to provide the solvent structure rather than explicitly sampling the solvent configuration space, thus significantly reducing the computational demand in MD, MC, or SD simulations. However we emphasize that our goal is not to use RISM to expedite simulations by providing approximate solvent distribution functions, but rather to use simulation results to aid in the calculation of solvation free energies as formulated within RISM theory.

Below we describe our implementation of the coupled RISM/MD or MC method and illustrate its application to a few small molecules. We also

present some comparison with results of the standalone RISM/HNC method.

2. Computational methodology

Within the RISM formalism, the radial distribution functions, $g = h + 1$, and the direct correlation function, c, satisfy the equation:

$$\hat{\mathbf{h}}(k) = \hat{\boldsymbol{\omega}}(k)\hat{\mathbf{c}}(k)\hat{\boldsymbol{\omega}}(k) + \rho\hat{\boldsymbol{\omega}}(k)\hat{\mathbf{c}}(k)\hat{\mathbf{h}}(k), \tag{1}$$

where the carets indicate spatial Fourier transforms, the bold-face denotes matrices, and $\hat{\boldsymbol{\omega}}$ is the intramolecular correlation function [3,4,8]. The RISM equation, (1) simplifies to one-dimensional the positional dependence of h. This averaging over angular coordinates of the exact molecular Ornstein–Zernike equation corresponds to the first of two approximations made in RISM theory. The second approximation made in RISM theory comes from the closure relations, equations derived to be solved in conjunction with the RISM equation. These closure relations were found by expressing h as a diagrammatic expansion, grouping terms corresponding to c, and throwing out the so-called 'bridge diagrams' [8]. For a polar solvent, the hyper-netted chain (HNC) closure relation is often used and is given by

$$g_{\alpha\gamma}(r) = \exp\left(-\beta u_{\alpha\gamma}(r) + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r)\right), \tag{2}$$

where u is the intermolecular pair potential. When the RISM equation is combined with the HNC relation, the two equations can be simultaneously solved for h and c, using an iterative technique [5–7].

Given solvent distribution functions from an MD or MC simulation, iterative solution of the RISM and HNC equations, (1) and (2), is no longer needed in order to determine the direct correlation function, c. Instead, c can be solved for from one or the other of these two equations. However, neither Eq. (1) nor (2), can be solved for c over its entire range, so that both equations are, in fact, necessary. In particular, the HNC equation, Eq. (2), cannot be solved for c at small distances where the radial distribution function vanishes. On the other hand, the necessary matrix inversions become ill-conditioned when k is near 0 when the RISM equation, Eq. (1), is used to solve for $\hat{c}$. Fortunately, c can be decomposed into two components that can be calculated separately from Eqs. (1) and (2), as described below.

Consider the expression for solvation free energy [10]:

$$\Delta\mu^{(\mathrm{HNC})} = -\frac{\rho}{2\beta}\sum_{\alpha\gamma}\int 4\pi r^2 \times \left[2c_{\alpha\gamma}(r) + h_{\alpha\gamma}(r)c_{\alpha\gamma}(r) - h^2_{\alpha\gamma}(r)\right]\mathrm{d}r. \tag{3}$$

Note that the integral of the first term is proportional to $\hat{c}(k)$ evaluated at $k = 0$, which, as we just mentioned, cannot be found from the RISM approach; nor can this term, or the third one, be found from the HNC approach, where c is unknown at small r. Fortunately, if we rewrite the free energy as

$$\Delta\mu^{(\mathrm{HNC})} = -\frac{\rho}{2\beta}\sum_{\alpha\gamma}\int 4\pi r^2\left[2c'_{\alpha\gamma}(r)\left(1 + h_{\alpha\gamma}(r)\right) - h_{\alpha\gamma}(r)c_{\alpha\gamma}(r) - h^2_{\alpha\gamma}(r)\right]\mathrm{d}r, \tag{4}$$

where c is calculated from the RISM equation and c' from the HNC equation, then the data available for these two function sets complement each other. This is because the range of r for which c' is unknown coincides with $1 + h = 0$ and makes no contribution to the first term in this integral. Also, since the last term is equal to a multiple of the integral of the product of the Fourier transforms, $\hat{h}$ and $\hat{c}$, over $\mathbf{k}$, the error in this term coming from our uncertainty in $\hat{c}$ at low k is reduced in this range by being multiplied by a factor of k^2.

It is important to point out that the free energy expression (Eq. (4)) inherits the invariant property of the solution to the RISM and HNC equations pointed out by Singer and Chandler [10]. We will discuss this feature and its effects on the error analysis of the coupled RISM/MD or MC method in the forthcoming full paper.

In addition to the HNC expression for the free energy of solvation, there are two other expressions that can also be employed. One is derived from Gaussian field (GF) theory [26]:

$$\Delta\mu^{(\mathrm{GF})} = \frac{-\rho}{2\beta}\sum_{\alpha\gamma}\int 4\pi r^2\left[2c_{\alpha\gamma}(r) + h_{\alpha\gamma}(r)c_{\alpha\gamma}(r)\right]\mathrm{d}r. \tag{5}$$

The other is based on 'partial wave' (PW) expansions of correlation functions [27,28]:

$$\Delta\mu^{(\mathrm{PW})} = -\frac{\rho}{2\beta}\sum_{\alpha\gamma}\int 4\pi r^2\Big[2c_{\alpha\gamma}(r) + h_{\alpha\gamma}(r)c_{\alpha\gamma}(r) - h_{\alpha\gamma}h^{(\mathrm{PW})}_{\alpha\gamma}(r)\Big]\mathrm{d}r. \quad (6)$$

It was shown that this expression can improve the calculation of solvation free energies from correlation functions solved for using RISM/HNC, when compared to the original 'HNC' equation (3), and, what is more, its derivation is more rigorous [28]. We will examine the accuracy of these three expressions for solvation free energy.

3. Computational details

In this study, distribution functions were calculated by Monte Carlo simulations at 298 K using the BOSS package [29]. The dimensions of solvent boxes used were large enough to contain solvent water molecules at distances of up to ~11 Å from the center of any solute atom.

Originally, OPLS all-atom parameters were used for solutes and the TIP3P model was used for water in all MC simulations. For consistency with MC simulations, we used the same force fields to evaluate the pair potential when finding direct correlation functions from the RISM and HNC equations in the coupled RISM/MC approach. In applying the RISM stand-alone method, the OPLS parameter set requires some small modifications. In particular, the OPLS parameter set assigns values of zero to the Van der Waals parameters of certain hydrogen atoms; this yields high artificial peaks in the calculated distribution functions. We found that if these parameters are assigned non-zero values ($\epsilon = 0.046$, $\sigma = 0.40$) [13], reasonable distributions can be obtained. Coupled RISM/MC simulations were also repeated, using the modified OPLS parameters consistently in all calculations for comparison with the stand-alone RISM/HNC results.

The Fourier transformations for the RISM calculations were carried out on a linear grid of 512 points, with zero-aliasing to twice this range [5]. We have employed the renormalization technique to obtain the contribution of the long-range Coulomb asymptotic behavior of c to h [5,6,30]. Using the RISM equation, the short-range part of c is then solved for in terms of the difference when this contribution is subtracted from h [31]. This involves the inversion of the intramolecular correlation matrix, $\hat{\boldsymbol{\omega}}$, which has dimensions equal to the number of solute sites. As we alluded to above, this matrix becomes ill-conditioned when k is close to 0; we used a linear extrapolation at low k. For consistency, this same linear extrapolation was used when evaluating the solvation free energy in the RISM/HNC stand-alone approach.

For simultaneous solution of the RISM and HNC equations we used the MDIIS convergence method [32].

Finally, in carrying out the numerical integration for the solvation free energy, radial distribution functions were cutoff at 11 Å. An adjustment was made to the solvation free energy expression based upon differences between coordination numbers of solvent O and H atoms, respectively, with respect to each solute site.

4. Results

We tested the coupled RISM/MC method on water and methanol. Figs. 1(a) and 2(a) illustrate some solvent/solvent and solute/solvent radial distribution functions, respectively, which were used in our coupled RISM/MC calculations, obtained from MC simulations with the OPLS parameter set. Shown in parts (b) of the same figures are the same distribution functions, this time calculated from MC simulations using the modified OPLS parameter set. The third graph in each figure shows the radial distribution functions determined by RISM/HNC, using the modified parameter set. Note that, especially in Fig. 1, the graphs shown in part (a) compare very closely to those in part (b), that is our modification of Van Der Waals parameters had little effect upon solvent distributions determined by MC.

Solvation free energies calculated by HNC, GF, and PW methods are listed in Table 1 along with experimental results. We found that the partial

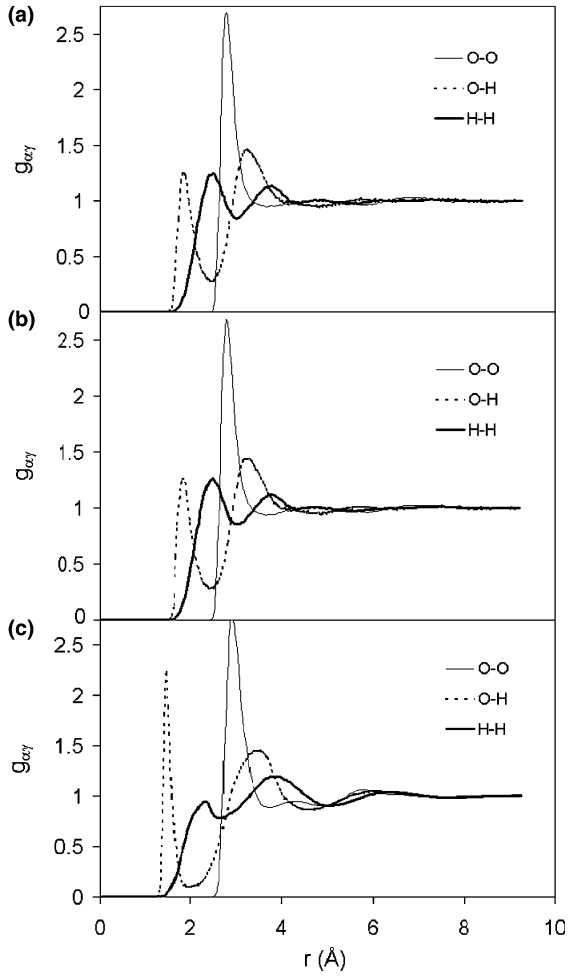

Fig. 1. Calculated solvent/solvent radial distribution functions for TIP3P water: (a) MC, (b) MC with modified OPLS parameters, (c) RISM/HNC with modified OPLS parameters.

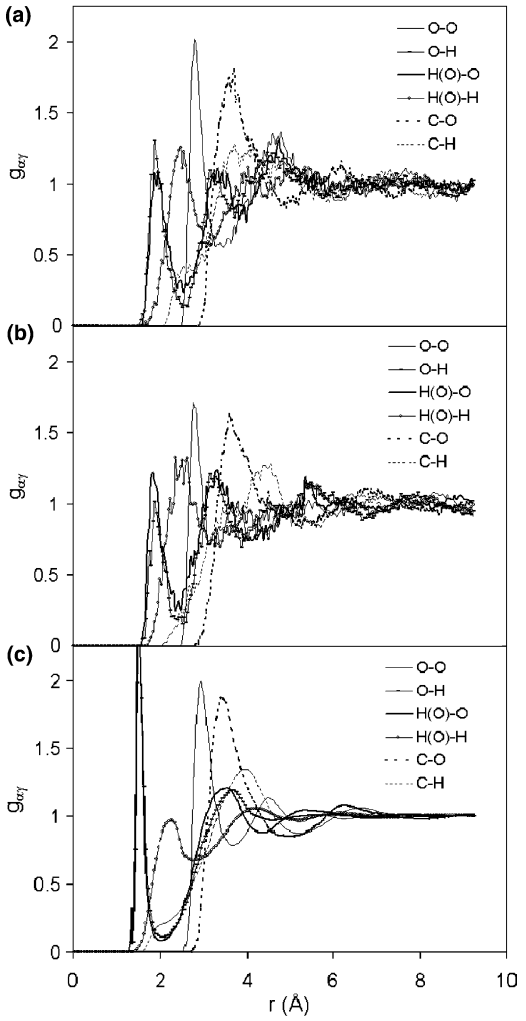

Fig. 2. Calculated solute/solvent radial distribution functions for methanol in TIP3P water: (a) MC, (b) MC with modified OPLS parameters, (c) RISM/HNC with modified OPLS parameters.

wave free energy expression gives the highest accuracy. The results obtained from the coupled RISM/MC method compared well with experimental solvation free energies.

In Table 2, we compare results of coupled RISM/MC calculations to those produced by the stand-alone RISM/HNC approach. The modified OPLS parameter set and the PW free energy expression were used in each calculation. The third

Table 1
Free energies of solvation (kcal/mol)

Solute	Experiment	$\Delta\mu^{HNC}$	$\Delta\mu^{GF}$	$\Delta\mu^{PW}$
TIP3P water	−6.3	−5.39	−10.06	−6.47
Methanol	−5.1	3.00	−11.16	−5.92

Table 2
Free energies of solvation (kcal/mol)

Solute	Experiment	RISM(PW)/HNC	RISM(PW)/MC
TIP3P water	−6.3	−8.31	−5.90
Methanol	−5.1	−6.19	−5.18

column shows the results of stand-alone RISM/HNC, while the fourth column gives the results of coupled RISM/MC calculations. Coupled RISM/MC results are somewhat better-quality, although a larger test set is needed to draw a more definitive conclusion.

5. Conclusions

In this study, we have proposed a new computational methodology for determining solvation free energies. The new method takes advantage of the RISM integral equation formalism for its simple expression for solvation free energy and the MD or MC simulation method for its accurate description of radial distribution functions. Note that, in contrast to the case with free energy simulations, we need only one MD or MC simulation to calculate radial distribution functions for the coupled RISM/MD or MC method.

We have shown that the coupled RISM/simulation approach gives satisfactory results for a few small molecules. A careful analysis of the accuracy of the coupled RISM/MD or MC method is needed and will be presented in the forthcoming full paper. Applications of this method to larger test-cases such as polypeptides and proteins are being considered.

Acknowledgements

We are grateful to Prof. Shao-Wen Zhang for assistance in the implementation of the RISM method and to Prof. Monte Pettitt, Prof. John Perkyns, and Prof. Frank Stenger for helpful discussion.

References

[1] M. Orozco, F.J. Luque, Chem. Rev. 100 (2000) 4187.
[2] A. Agarwal, F. Brown, M.R. Reddy, in: M.R. Reddy, M.D. Erion (Eds.), Free Energy Calculations in Rational Drug Design, Kluwer Academic/Plenum Publishers, New York, 2001.
[3] H.C. Andersen, D. Chandler, J. Chem. Phys. 57 (1972) 1918.
[4] D. Chandler, H.C. Andersen, J. Chem. Phys. 57 (1972) 1930.
[5] P.J. Rossky, H.L. Friedman, J. Chem. Phys. 72 (1980) 5694.
[6] F. Hirata, P.J. Rossky, Chem. Phys. Lett. 83 (1981) 329.
[7] F. Hirata, B.M. Pettitt, P.J. Rossky, J. Chem. Phys. 77 (1982) 509.
[8] J.P. Hansen, I.R. McDonald, Theory of Simple Liquids, Academic Press, London, 1986.
[9] T. Morita, K. Hiroike, Prog. Theoret. Phys. 23 (1960) 1003.
[10] S.J. Singer, D. Chandler, Mol. Phys. 55 (1985) 621.
[11] D.A. Zichi, P.A. Rossky, J. Chem. Phys. 84 (1985) 1712.
[12] H.A. Yu, M. Karplus, J. Chem. Phys. 89 (1988) 2366.
[13] H.A. Yu, B. Roux, M. Karplus, J. Chem. Phys. 92 (1990) 5020.
[14] T. Lazaridis, M. Karplus, in: E. Di Cera (Ed.), Thermodynamics in Biology, Oxford University Press, New York, 2000, p. 3.
[15] T. Ichiye, D. Chandler, J. Phys. Chem. 92 (1988) 5257.
[16] H.-A. Yu, B.M. Pettitt, M. Karplus, J. Am. Chem. Soc. 113 (1991) 2425.
[17] M. Kinoshita, Y. Okamoto, F. Hirata, J. Chem. Phys. 107 (1997) 1586.
[18] M. Kinoshita, Y. Okamoto, F. Hirata, J. Comput. Chem. 19 (1998) 1724.
[19] N.V. Prabhu, J.S. Perkyns, H.D. Blatt, P.E. Smith, B.M. Pettitt, Biophys. Chem. 78 (1999) 113.
[20] B.M. Pettitt, M. Karplus, P.J. Rossky, J. Phys. Chem. 90 (1986) 6335.
[21] B.M. Pettitt, M. Karplus, Chem. Phys. Lett. 121 (1985) 194.
[22] P.E. Smith, B.M. Pettitt, M. Karplus, J. Phys. Chem. 97 (1993) 6907.
[23] A. Kitao, F. Hirata, N. Go, J. Phys. Chem. 97 (1993) 10231.
[24] M. Kinoshita, Y. Okamoto, F. Hirata, J. Am. Chem. Soc. 120 (1998) 1855.
[25] M. Kinoshita, Y. Okamoto, F. Hirata, J. Chem. Phys. 110 (1999) 4090.
[26] D. Chandler, Y. Singh, D. Richardson, J. Chem. Phys. 81 (1984) 1975.
[27] S. Ten-No, S. Iwata, J. Chem. Phys. 111 (1999) 4865.
[28] S. Ten-No, J. Chem. Phys. 115 (2001) 3724.
[29] W.L. Jorgensen, Yale University, 2000.
[30] A. Kovalenko, F. Hirata, J. Chem. Phys. 112 (2000) 10391.
[31] F. Hirata, P.J. Rossky, B.M. Pettitt, J. Chem. Phys. 78 (1983) 4133.
[32] A. Kovalenko, S. Ten-No, F. Hirata, J. Comput. Chem. 20 (1999) 928.