ELSEVIER

Biophysical Chemistry 71 (1998) 199–204

Biophysical Chemistry

Contributions of solvent–solvent hydrogen bonding and van der Waals interactions to the attraction between methane molecules in water

Jeffrey A. Rank, David Baker *

Department of Biochemistry, University of Washington, Seattle, WA, USA

Received 7 November 1997; revised 7 January 1998; accepted 7 January 1998

Abstract

The contribution of solvent–solvent hydrogen bonding and van der Waals interactions to the attraction between methane molecules in water was investigated by comparing the potential of mean force (PMF) between two methane molecules in TIP4P water to those in a series of related liquids in which the solvent–solvent interactions were progressively turned off while keeping the solvent–solute interactions unchanged. The magnitude of the attraction between methanes was not significantly changed when the hydrogen bonding interaction between solvent molecules was eliminated and the solvent was maintained in the liquid state by increasing either the pressure or the magnitude of the solvent–solvent van der Waals interaction. However, when solvent–solvent excluded volume interactions were eliminated, the methane molecules interacted no more strongly than in the gas phase. The results are consistent with the idea that the primary contribution of hydrogen bonding to the hydrophobic interaction is to keep water molecules in a liquid state; at constant density, packing interactions rather than hydrogen bonding appear to be critical as suggested by scaled particle theories of solvation. The overall shape of the PMF was, however, changed in the absence of hydrogen bonding, pointing to an influence of hydrogen bonding on the detailed form of the interactions between nonpolar solutes in water. The effects of correlations between the configurations sampled during the Monte Carlo procedure used in the free energy calculations on the estimation of errors was also characterized.

Keywords: Hydrophobic effect; Methane dimer; Potential of mean force

---

1. Introduction

Although it is universally acknowledged that the hydrophobic effect plays a critical role in the stabilization of biological macromolecules, there is surprisingly little consensus on its physical origins; recent discussions have traced the hydrophobic effect to very different features of water [1–3]. A better understanding of hydrophobic interactions is essential for understanding processes such as protein folding which are largely driven by the clustering of nonpolar atoms in protein cores [4,5].

As the power of computers has increased, simulations of biological macromolecules have become an increasingly powerful research tool. Such simulations require a computationally tractable model of

---

* Corresponding author. E-mail: dabaker@u.washington.edu

0301-4622/98/$19.00
PII S0301-4622(98)00103-3

water, and a number of such models have been proposed. Here we use one of the most commonly used of such models, TIP4P [6] in conjunction with standard free energy perturbation methods [7], to investigate a very simple system: the interaction between two methane particles in water. Previous studies ([8,9] and references therein) have shown that there is a pronounced attraction between methane molecules in water. Our goal here is to characterize the features of water responsible for this attraction.

2. Results

2.1. Free energy calculations

The perturbation method was used to compute the free energy change as a function of the separation between two methane molecules in a box of water (see methods):

$$\Delta G(d \rightarrow d') = -kT \ln < \exp\{-\Delta E_i(d \rightarrow d')/kT\} >_b \quad (1)$$

where $\Delta G(d \rightarrow d')$ is the change in free energy and $\Delta E_i(d \rightarrow d')$ the change in energy associated with a perturbation in the solute separation from d to d' [7]. $\langle\rangle_b$ denotes a Boltzman weighted average over all states of the unperturbed system. The Boltzman average is approximated by averaging over a set of configurations generated in a long Monte Carlo simulation in which successive configurations are generated by rotating and translating a single water molecule. Although configurations separated by a small number of moves are highly correlated, it is generally assumed that the sets of configurations generated over the course of a long simulation approach the broad, representative cross section of phase space required for the accurate application of Eq. (1).

A standard procedure for calculating errors in free energy perturbation calculations is to calculate averages of quantities of interest over blocks of typically one to 300 000 Monte Carlo moves. The data from each of these 'block averages' are then used to compute the average and standard deviation for the measured quantity over the simulation. The implicit assumption with such a method is that the different blocks are independent samples. To test this assumption, we calculated the correlation in the measured change in solute–solvent interaction energy for a small displacement of the solutes over the course of a long Monte Carlo run. The correlations persist over hundreds of thousands of Monte Carlo moves which corresponds to tens of thousands of moves per water molecule (Fig. 1a).

To explore the effect of these long correlations on the estimation of errors in the free energy calculations, mean free energy changes and standard deviations for a representative perturbation were calculated from five independent Monte Carlo runs of 5

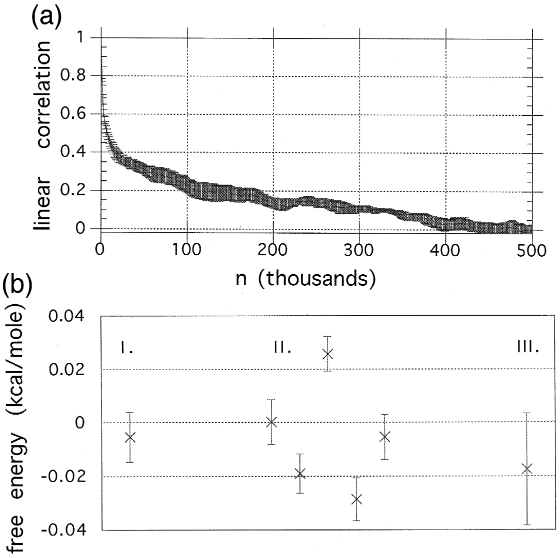

Fig. 1. Errors in free energy calculations from persistent correlations in Monte Carlo simulations. (a) Linear correlation between $\Delta E_i(d \rightarrow d')$ for configurations separated by n monte carlo moves. The curve shown was calculated from five independent runs of five million configurations each; the perturbation was a change in the methane–methane separation from 5.95 to 6.00 Å. Error bars plotted are one sigma. (b) Analysis of errors in the computation of the change in free energy for a change in the methane–methane separation from 5.95 to 6.00 Å. Averages and standard errors were computed from five independent Monte Carlo runs of 5 million configurations each. The leftmost point (I) is the average of the values obtained from each of the 5 runs with the associated one sigma error bar. The 5 center points (II) are the averages for each of the 5 runs with error bars calculated using 50 block averages each consisting of 100 000 configurations. The rightmost point (III) was calculated from the 5 independent averages corresponding to the first 1 million configurations of each run.

million configurations each using I) all 25 million configurations, II) data from the 5 million configurations of a single Monte Carlo run with a block average size of 100 000 configurations, and III) data from the first 1 million configurations of each of the five independent Monte Carlo runs. Although the value calculated using method III is based on at least as much statistically independent data, the calculated standard error is much larger than that obtained using method II. Furthermore, several of the values obtained using block averaging are nearly two sigma from the best estimate (I) obtained from averaging all of the available data. Thus block averaging can result in a considerable underestimate of the real uncertainty in the quantities being calculated.

To avoid inaccuracies associated with block averaging all subsequent data and standard errors were calculated from a minimum of ten independently equilibrated runs for each point. Since correlations persist over lengths on the order of that required for equilibration, performing independent runs does not significantly increase simulation time, but does allow rigorous statistical uncertainties to be computed.

2.2. Methane–methane PMF in TIP4p and related liquids

Fig. 2 (dashed line) shows the free energy of association or potential of mean force (PMF) for two methane molecules in TIP4P water at 25°C. There is a pronounced minimum at contact, and a second minimum for a solvent separated pair. As described previously, the results are largely independent of the choice of water model and the size of the water box [9,10]. The average error per Å computed from the ten independent runs, 0.044 kcal/(mol Å), was small compared to the magnitude of the free energy changes.

In order to investigate the origin of the attraction between methane molecules in water, a series of simplifications were made to the solvent. First, the extent to which hydrogen bonding contributes to the PMF was investigated by turning off all coulombic interactions, leaving a Leonard Jones fluid with the TIP4P LJ parameters. In order to keep the solvent in a liquid state a constant volume ensemble was used, with the density fixed to the density of TIP4P water at 25°C (see methods; a similar strategy was used by Madan and Lee [11] to explore the contribution of hydrogen bonding to the cavity distribution function). The PMF so obtained (Fig. 2, triangles) resembles the PMF obtained in hydrogen bonded water (Fig. 2, dashed line). The attraction is somewhat reduced, with the free energy difference between solvent separated and contact mimina change from 0.59 ± 0.06 kcal/mol to 0.30 ± 0.06 kcal/mol.

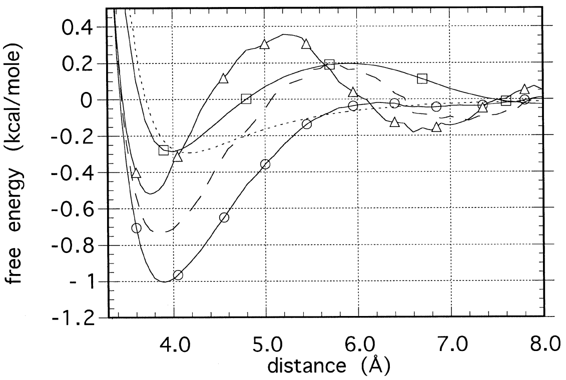

Fig. 2. Methane–methane PMF in TIP4P water and simplified model liquids. Dashed line, TIP4P water; triangles, TIP4P without coulombic interactions with the density maintained at that of water using high pressure; circles, model liquid used by Pratt and Pohorille [3] with increased LJ attraction; squares, noninteracting TIP4P like fluid; dotted line, methane–methane LJ interaction. Data for all curves were collected at .05 Å intervals; for clarity only a small subset of the datapoints are indicated with symbols.

A potential complication in the interpretation of this result is the large difference in the pressure in the two simulations. The liquid state can be preserved in the absence of hydrogen bonding at atmospheric pressure if the strength of the Leonard Jones (LJ) interactions between solvent molecules are substantially increased. Pratt and Pohorille [3] used a model liquid with LJ parameters chosen to reproduce the radial distribution function and density of water at 1 atm to assess the influence of hydrogen bonding on the cavity distribution function. To assess the contribution of hydrogen bonding without large increases in the pressure, we computed the methane–methane PMF in this LJ liquid (Fig. 2, circles). The free energy loss upon bringing the two methanes into contact in this non hydrogen bonding liquid is greater than in TIP4P water. The decrease in the size of the barrier may be related to the increased orientational freedom of the non hydrogen bonded liquid: there may be more low energy configurations of solvent molecules around the highly curved cavity formed by the two methanes at intermediate separations. The comparison between TIP4P and this LJ liquid is also interesting because the integral equation theory of Pratt and Chandler [12] related the methane–methane PMF to the solvent–solvent radial distribution function. Because the two liquids have very similar radial distribution functions, the similarities and differences between the corresponding methane–methane PMFs indicate the extent to which pair correlations, which the integral equation theory is based on, account for the observed interactions. In particular, the barrier at intermediate solute separations appears to reflect higher order correlations between molecules.

The fact that in both calculations with TIP4P like liquids lacking hydrogen bonds the free energy decrease in contact formation is greater than that due solely to the LJ interaction between the two methanes (Fig. 2, dotted line) suggests that solvent hydrogen bonding is not critical to the added attraction between methanes in TIP4P water. To determine whether the solvent contribution to the interaction between the methanes depends on LJ interactions between solvent molecules, the model was further simplified by turning off all solvent–solvent interactions leaving two methane particles interacting with an ideal gas-like fluid at the density of water. In this model, calculation of the PMF requires only the computation of two 3 dimensional configuration integrals (see methods). The attraction between methanes in such a liquid (Fig. 2, squares) is no greater than the LJ attraction between two methanes in the gas phase (Fig. 2, dotted line). Thus, solvent–solvent excluded volume interactions contribute to the attraction between methanes in water, as expected given the large differences between the cavity distribution function of a hard sphere liquid and an ideal gas.

To obtain further insights into the differences between the liquids, and to enable comparison with experimental data, the solvent contribution to the methane–methane PMF was determined for solute separations ranging from 1.5 to 8.5 Å. Ben-Naim [13] pointed out that the solvent contribution to the free energy change for bringing two methanes together to a separation distance of 1.5 Å (the C–C separation in ethane) should correspond to the difference in the free energy of solvation of two methanes compared to a single ethane. The solvent contribution for this two methane to ethane transformation was -3.8, -3.2 and -6.1 kcal/mol for TIP4P water, the LJ liquid of Pratt and Pohorile, and the TIP4P like liquid without coulombic interactions held together by high pressure. The large increase in the free energy change for the solvent under high pressure is due to the reduction in the volume excluded to the solvent in the transformation; the $P\,\Delta V$ term is only significant for the high pressure simulation. The free energy changes obtained for the first two liquids, while similar, are considerably larger than the estimate of -2.15 kcal/mol obtained from the solvation energies of methane and ethane [13]. The origins of the discrepancy between the estimates from simulation and experiment are not clear; one possibility is the neglect of polarizability implicit in the TIP4P model.

3. Discussion

Our results suggest that excluded volume interactions between solvent molecules are critical to the attraction between nonpolar solutes in water, while solvent–solvent hydrogen bonding plays only a secondary role. Hydrogen bonding is critical to keeping water a liquid, but when the solvent density is maintained in the absence of hydrogen bonding by increasing either the van der Waals interactions between solvent molecules or the pressure, the attraction between nonpolar solutes is largely preserved. The results are consistent with the observation of collapse of hexadecane in non-hydrogen bonded model liquids [14] and with the argument that the unusual feature of water which gives rise to the hydrophobic effect is not the extensive hydrogen bonding but the small size of the water molecule [2]. While perhaps surprising to the biochemist, the assumption that hydrogen bonding plays a secondary role is implicit in scaled particle theory based treatments of solvation [15,16]. However, it is clear from Fig. 2 that solvent hydrogen bonding does influence the methane–methane PMF; this is consistent with previous observations that there are subtle differences between the cavity distribution functions of

TIP4P water and the two non hydrogen bonding model liquids [3,11]. The bilobed cavities relevant to the methane–methane PMF for solute separations less than ~5 Å differ considerably from the spherical cavities studied previously [3,11], but the same general trends are likely to hold.

Despite the lack of sensitivity of the overall free energy change in solvation or solute association processes to the details of the solvent–solvent interactions, many of the anomalous features of the hydrophobic interaction, including the large change in heat capacity, may be directly related to hydrogen bonding between solvent molecules. Solvent reconfiguration in the vicinity of nonpolar solutes contributes to both the enthalpy and entropy changes in nonpolar solvation and solute association processes, but the two contributions largely cancel in the overall free energy change [17–20]. Because of this enthalpy–entropy compensation, estimation of the contribution of hydrophobic interactions to the free energy changes in complex reactions such as protein folding should be more straightforward than estimation of the contributions to the enthalpy and entropy changes; our results suggest that very simple solvent models may suffice for the former.

An interesting recent advance in the modeling of hydrophobic interactions employed maximum entropy methods to estimate the cost of cavity formation from only the density and radial distribution function of the solvent [22]. This simple approach quantitatively reproduced the TIP4P cavity distribution function along with solvation free energies and PMFs for hard sphere solutes computed using explicit solvent models. The model also predicted the entropy convergence characteristic of hydrophobic solvation [23], suggesting that the details of hydrogen bonding may not even be required for understanding the entropy changes in hydrophobic hydration. Since the density and radial distribution function of each of the model solvents used in this study are readily obtainable, our results provide an opportunity to further test the remarkable success of this simple approach. Accurate prediction of the differences between the methane–methane PMF in the different solvents from the (relatively small) differences in the solvent radial distribution functions would be an impressive confirmation of the power of the method.

4. Methods

4.1. MC simulations

Monte Carlo calculations were performed using a $20 \times 20 \times 30$ Å box containing 396 water molecules with periodic boundary conditions. The TIP4P model [6] was used for water, and a united atom approximation for the solute molecules. The methane–methane and methane–solvent interactions were modeled by a LJ potential

$$U_{\mathrm{LJ}}(r) = 4\varepsilon\left((\sigma/r)^{12} - (\sigma/r)^{6}\right) \tag{2a}$$

with parameters $\sigma_{\mathrm{m-m}} = 3.71$ Å, $\varepsilon_{\mathrm{m-m}} = 0.294$ kcal/mol [8]. The parameters used for solute–solvent interactions were determined using the Lorentz–Berthelot mixing rules [21]. All solvent interactions were quadratically damped to zero between 8.0 and 8.5 Ås.

Simulations were performed at 25°C and 1 atm in an NPT ensemble for all calculations except for those for the TIP4P liquid without electrostatic interactions (Fig. 2, triangles). The latter simulation was performed at 25 C° and 12 000 Å^3 in an NVT ensemble. The difference in ensembles is to avoid large volume expansions when the hydrogen bonding (electrostatic) interactions between the water molecules are turned off. This is equivalent to applying a very large pressure (about 35 000 atm) to the same system in an NPT ensemble. The potential of mean force is independent of the choice of ensemble since for this process $(\Delta F)_{V,T} = (\Delta G)_{p,T}$.

The perturbation method was used to compute the free energy change as a function of the distance between the solutes. The reaction coordinate was the solute–solute separation distance and a perturbation step size of 0.05 Å was used. Double wide sampling and preferential solute sampling were used to increase sampling efficiency [21]. The solute pair was oriented along the long axis of the box with a fixed separation distance and placed at a random position in one of 25 pre-equilibrated water boxes. High energy waters were removed and the system was equilibrated for 1 million MC moves. Sampling was then carried out over an additional 1.5 million MC moves. Errors were estimated by summing the standard deviations of the mean free energy changes over the ten independent runs for each point in quadrature and then dividing by the total length of the PMF.

The PMF for the non-interacting solvent model (Fig. 2, squares) was calculated by evaluating the 3 dimensional integral for the configurational part of the partition function for a single water molecule in the presence of the solutes, and normalizing the result to the density of water:

$$\begin{aligned}\Delta F(d) &= -kT \ln[Z(d)/Z(\infty)] \\ &= kT\phi_s \lim_{v \to \infty}\{V/V_w[1 - I(d)/(2I_1 - 1)]\}\end{aligned} \qquad (2b)$$

where $\Delta F(d)$ is the solvent contribution to the PMF, ϕ_s is the solvent volume packing fraction $= 0.54$ for TIP4P water, and V_w is the volume of a single solvent molecule. $I(d)$ is the configurational integral for a single water molecule in a box containing two methane molecules separated by distance d, and I_1, the configurational integral for a single water molecule in a box containing a single methane molecule.

Acknowledgements

We would like to thank Drs. D. Shortle, R. Baldwin, K. Plaxco, B. Lee, and W. Parson for helpful comments on the manuscript. JR was supported by an NIH molecular biophysics predoctoral fellowship. This work was supported by young investigator awards to DB from the NSF and the Packard foundation.

References

[1] N. Muller, Acc. Chem. Res. 23 (1990) 23–28.
[2] B. Lee, Biopolymers 24 (1985) 813–823.
[3] L.R. Pratt, A. Pohorille, PNAS 89 (1992) 2995–2999.
[4] W. Kauzmann, Adv. Protein Chem. 14 (1959) 1–63.
[5] K. Dill, Biochemistry 29 (1990) 7133–7155.
[6] W.L. Jorgenson, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926–935.
[7] D.L. Beveridge, F.M. DiCapua, Annu. Rev. Biophys. Biophys. Chem. 18 (1989) 431–492.
[8] W.L. Jorgenson, J.K. Buckner, S. Boudon, J. Tirado-Rives, J. Chem. Phys. 89 (1988) 3742.
[9] D.E. Smith, A.D.J. Haymet, J. Chem. Phys. 98 (1993) 6445–6454.
[10] J.A. Rank, D. Baker, Protein Sci. 6 (1997) 347–354.
[11] B. Madan, B. Lee, Biophys. Chem. 51 (1994) 279–289.
[12] L.R. Pratt, D. Chandler, J. Chem. Phys. 67 (1977) 3683–3704.
[13] A. Ben-Naim, J. Chem. Phys. 54 (1971) 1387–1404.
[14] A. Wallqvist, D.G. Covell, Biophys. J. 71 (1988) 600–608.
[15] H. Reiss, Adv. Chem. Phys. 9 (1965) 1–84.
[16] F.H. Stillinger, J. Solution Chem. 2 (1973) 141–158.
[17] H. Yu, M. Karplus, J. Chem. Phys. 89 (1988) 2366–2379.
[18] A. Ben-Naim, Solvation thermodynamics, Plenum, New York, 1987.
[19] B. Lee, Biophys. Chem. 51 (1994) 271–278.
[20] B. Lee, Graziano, JACS 118 (1996) 5163–5168.
[21] M.P. Allen, D.J. Tildesly, Computer simulations in liquids, Clarendon Press, Oxford, 1986.
[22] G. Hummer, S. Garde, A. Garcia, A. Pohorille, L. Pratt, PNAS 93 (1996) 8951–8955.
[23] S. Garde, G. Hummer, A. Garcia, M. Paulaitis, L. Pratt, Phys. Rev. Lett. 77 (1996) 4966–4968.