Published in final edited form as:

Faraday Discuss. 2010 ; 146: 247–401.

Hydrophobic interactions in model enclosures from small to large length scales: non-additivity in explicit and implicit solvent models

Lingle Wang, Richard A. Friesner, and B. J. Berne

Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027, USA

Abstract

The binding affinities between a united-atom methane and various model hydrophobic enclosures were studied through high accuracy free energy perturbation methods (FEP). We investigated the non-additivity of the hydrophobic interaction in these systems, measured by the deviation of its binding affinity from that predicted by the pairwise additivity approximation. While only small non-additivity effects were previously reported in the interactions in methane trimers, we found large cooperative effects (as large as $-1.14 \text{ kcal mol}^{-1}$ or approximately a 25% increase in the binding affinity) and anti-cooperative effects (as large as $0.45 \text{ kcal mol}^{-1}$) for these model enclosed systems. Decomposition of the total potential of mean force (PMF) into increasing orders of multi-body interactions indicates that the contributions of the higher order multi-body interactions can be either positive or negative in different systems, and increasing the order of multi-body interactions considered did not necessarily improve the accuracy. A general correlation between the sign of the non-additivity effect and the curvature of the solute molecular surface was observed. We found that implicit solvent models based on the molecular surface area (MSA) performed much better, not only in predicting binding affinities, but also in predicting the non-additivity effects, compared with models based on the solvent accessible surface area (SASA), suggesting that MSA is a better descriptor of the curvature of the solutes. We also show how the non-additivity contribution changes as the hydrophobicity of the plate is decreased from the dewetting regime to the wetting regime.

1 Introduction

Hydrophobic interactions (HI) play a very important role in the formation and stability of many self-assembled aggregates and biological structures,¹ and are considered to be the driving force for protein folding.^{2,3} A fundamental understanding of HI is crucial to the study of many important biological phenomena, such as protein folding, micelle formation and protein–ligand binding.

An important question concerning the protein folding problem is whether hydrophobic associations are pairwise additive, cooperative, or anti-cooperative.^{4,5} In other words, is the potential of the mean force (PMF) holding a cluster of n hydrophobic particles together equal to the sum of pairwise interaction free energies (or pair pmfs), or is it more negative (cooperative) or more positive (anti-cooperative) than what is predicted by the pairwise additivity approximation.

Nemethy and Scheraga⁶ found that the hydrophobic interactions between more than two solute particles can't be expressed as a sum of pairwise solute–solute interactions. Palma also observed the non-additivity of solvent induced potential of mean force for hydrophobic particle solutions by molecular dynamics simulation.⁷ In 1997, Rank and Baker⁸ studied three methane molecules in an isosceles triangle geometry with two methane molecules at

contact distance forming a fixed base, and they found that the three-body PMF was anti-cooperative for distances up to 6.5 Å. The conclusions of this work were not reliable, however, because the error associated with the baseline of the PMF was of the same order as the non-additive term itself. After that, Chan^{4,9} and Scheraga^{10–12} performed studies of methane trimers using the weighted histogram analysis method (WHAM) and the test particle insertion method respectively, and found contradictory results. After exchange of comments between these two groups,^{13–16} they found that the disagreement came from the assignment of the baseline for the PMFs. Comparing these two methods Scheraga¹⁷ stated that the methane dimer seemed to be the largest system that can be treated by the test particle insertion method. The effect of size, pressure, temperature, and salts on the non-additivity of the three particle system was also investigated.^{18–21} Recently, cooperative effects on the association of four methane molecules were also investigated by Scheraga's group.²²

In this paper, we used the FEP method to study the binding affinities between a united-atom methane and various model hydrophobic enclosures. The binding affinities were compared with the predictions of the pairwise additivity approximation to assess the non-additive contributions to the hydrophobic interactions in these systems. Two different empirical models were also used to predict the binding affinities and non-additivity effects, and comparisons were made with the FEP reference data. We found that the molecular surface area (MSA, or Connolly surface area) model performed much better than the solvent accessible surface area (SASA) model for the predictions of both the binding affinities and the non-additivity effects, which is consistent with previous findings.^{4,8,10,11} Detailed analysis of these two models indicates that the SASA model can't predict the cooperative effects, but the MSA model can. Decomposition of the total PMF into increasing orders of multi-body interactions indicated that the higher order multi-body interactions can be either positive or negative, and increasing the order of the multi-body interactions considered did not necessarily improve the accuracy.

2 Definition of cooperative and anti-cooperative effects

In general, for a solution with n solute particles forming a cluster, the potential of mean force (PMF), $W(1, 2, \dots, n)$, is related to the n -particle correlation function $g^{(n)}(1, 2, \dots, n)$ by the following definition:²³

$$g^{(n)}(1, 2, \dots, n) = e^{-\beta W(1, 2, \dots, n)} \quad (1)$$

where $\beta^{-1} = k_B T$, k_B is Boltzmann's constant, and T is the temperature. $W(1, 2, \dots, n)$ is a short-hand notation for $W(r_1, r_2, \dots, r_n)$, where r_i denotes the position of the i -th particle, and corresponds to the free energy required to bring n particles from infinitely far apart to the current configuration.

The n -particle PMF can be decomposed into single-body, pairwise and multi-body contributions:

$$\begin{aligned} W(1, 2, \dots, n) &= F(1, 2, \dots, n) - \sum_{i=1}^n F(i) \\ &= \sum_{i < j} \delta F(i, j) + \sum_{i < j < k} \delta F(i, j, k) + \dots + \delta F(1, 2, \dots, n) \end{aligned} \quad (2)$$

$$W(1, 2, \dots, n) = W_2 + W_3 + \dots + W_n \quad (3)$$

Where $F(1, 2, \dots, n)$ is the hydration free energy for the specified configuration of the solute particles, $F(i)$ is the hydration free energy of solute article i in an infinitely dilute solution, $\delta F(i, j)$ is the same as the normalized two-body PMF $W(i, j)$, and $\delta F(i, j, \dots)$ corresponds to subsequent higher order multi-body interactions. W_2 is the sum of pairwise contributions, and W_m is the sum of m -body interactions. Truncating the series to an n -body term leads to the Generalized Kirkwood Superposition Approximation (GKSA) to the n -th order.^{24,25} Specifically, the hypothesized pairwise additivity of PMF is obtained by truncating the series at the pairwise term. In clusters this approximation neglects shielding effects where a third particle could shield a pair from other particles and from the solvent.

For the three-particle case, pairwise additivity is equivalent to the Kirkwood super-position approximation,^{26–28} and $\delta F(i, j, k)$ measures the non-additive part of the three body interactions. Cooperativity is defined when $\delta F(i, j, k)$ is negative, meaning the free energy between the third particle and the remaining two particles is more negative and the configuration is more favorable than pairwise additivity predicts, and similarly, anti-cooperativity is defined when $\delta F(i, j, k)$ is positive, meaning the free energy between the third particle and the remaining two particles is more positive and the configuration is less favorable than pairwise additivity predicts. For the n -particle case, if the interaction between a specific particle and the remaining $n - 1$ particles is equal to the sum of the pairwise interactions between that specific particle with each of the other $n - 1$ particles, and if this condition holds for each of the n particles, the total n -body PMF will be equal to the sum of the pairwise interactions. So we further generalize the cooperative and anti-cooperative concepts to the n -particle case: when the interaction energy between one specific particle and the remaining $n - 1$ particles is more negative than the sum of pairwise free energies between the specific particle and each particle in the remaining $n - 1$ cluster, we term it “cooperative”; “anti-cooperative” is similarly defined. In other words, if we label the specific particle as n , and introduce the notation $\delta W(1, 2, \dots, n - 1; n)$, which is the sum of all higher than two body interactions involving particle n ,

$$\delta W(1, 2, \dots, n - 1; n) = F(1, 2, \dots, n) - F(1, 2, \dots, n - 1) - F(n) - \sum_{i=1}^{n-1} \delta F(i, n) \quad (4)$$

$$\delta W(1, 2, \dots, n - 1; n) = \sum_{i < j \leq n-1} \delta F(i, j, n) + \sum_{i < j < k \leq n-1} \delta F(i, j, k, n) + \dots + \delta F(1, 2, \dots, n - 1, n) \quad (5)$$

then, cooperativity or anti-cooperativity is defined when $\delta W(1, 2, \dots, n - 1; n)$ is negative or positive, respectively. In eqn (4), the term, $F(1, 2, \dots, n) - F(1, 2, \dots, n - 1) - F(n)$, gives the interaction energy between particle n and the remaining $n - 1$ particles, and the term,

$\sum_{i=1}^{n-1} \delta F(i, n)$, gives the sum of pairwise interactions between particle n and each of the other $n - 1$ particles. The difference between these two terms provides a measure of the non-additive contribution to the interaction.

3 Simulation details

In this paper, molecular dynamics simulations were performed using the DES-MOND program²⁹ to study the binding affinities between a united-atom methane and 13 model hydrophobic enclosures depicted in Fig. 1. The geometry of the model hydrophobic plate in these systems is displayed in the bottom right of Fig. 1. It consists of 19 single-layer atoms arranged in a triangular lattice with a bond length of 3.2 Å. For systems consisting of two plates, the two plates were parallel and in-registry with a separation distance of $D = 7.46$ Å. The LJ atoms forming the enclosures were uniformly represented with Lennard Jones parameters $\sigma = 3.73$ Å and $\varepsilon = 0.294$ kcal mol⁻¹, which are the same as the united-atom methane parameters used in these simulations.³⁰ The inserted methane particle displayed in black (online version green) in Fig. 1 was placed at the contact distance ($d = 3.73$ Å) with the other atoms and plate(s) forming the enclosures. In order to study the non-additivity of hydrophobic interactions between the insertion methane and the enclosures, another 10 systems corresponding to the subsystems of the 13 model enclosures (depicted in Fig. 2, and named after the corresponding system in Fig. 1 by adding a “prime”) were also studied. The binding affinities for the other four systems in Fig. 2 [G', H', I', K'] were calculated by combining the binding affinities for related systems in a thermodynamic cycle. For example, the binding affinity for system G' was calculated by combination of binding affinities calculated for systems G', G and E.

The free energy perturbation (FEP) method was used to determine the binding affinities between the inserted methane and each of the enclosures. The Maestro System Builder utility³¹ was used to insert each enclosure into a cubic water box with a 10 Å buffer. The SPC water model³² was used to describe the solvent. The atoms of the enclosures were constrained to their initial positions throughout the dynamics, and only the solvent degrees of freedom were sampled. The united-atom methane was “turned on” inside the model enclosures over 9 lambda windows with $\lambda = [0, 0.125, 0.25, 0.375, 0.50, 0.625, 0.75, 0.875, 1]$, where λ is the coupling parameter to turn on/off the LJ interaction between the methane and the rest of the system with initial and final states corresponding to $\lambda = 0$ and $\lambda = 1$ respectively. In these simulations, the core of the LJ potential is made softer³³ as $\lambda \rightarrow 0$ to avoid singularities and numerical instabilities. For each of the λ windows, molecular dynamics simulations were performed. The energy of the system was minimized, and then equilibrated to 298 K and 1 atm with Nose–Hoover^{34,35} temperature and Martyna–Tobias–Klein³⁶ pressure controls over 100 ps of molecular dynamics. A cut-off distance of 9 Å was used to model the Lennard-Jones interactions, and the particle-mesh Ewald method³⁷ was used to model the electrostatic interactions. Following the equilibration, a 20 ns production molecular dynamics simulation was performed and configurations of the system were collected every 1.002 ps. The energy difference between neighboring λ windows for each configuration saved was calculated and the Bennett acceptance ratio method³⁸ was used to calculate the free energy difference between neighboring states. The sum of the free energy difference between neighboring states gave the solvation free energy of methane in the enclosures. The same procedure was followed to calculate the solvation free energy of methane in bulk water. The difference between the two solvation free energies gave the binding free energy to bring a methane from infinitely far to inside the hydrophobic enclosure, which is the potential of mean force (PMF) between the methane and the enclosure. The error associated with these binding affinities was of the order of ± 0.02 kcal mol⁻¹.

Implicit solvent model calculations

Due to the large computational cost of running explicit solvent model simulations, in protein folding or protein–ligand binding problems, one is often forced to use implicit solvent models to reduce the computational cost. Many of the implicit solvent models attempt to

account for hydrophobicity in terms of nonpolar surface exposure to water,³⁹ among which solvent accessible surface area (SASA)⁴⁰ and molecular-surface-area (or Connolly surface area, MSA)⁴¹ models are the most popular. The solvent accessible surface (SAS) is traced out by the probe sphere center as it rolls over the solute, and the molecular surface (MS) is the surface traced by the inward-facing surface of the probe sphere.

In this paper, the SASA and MSA of each enclosure, both with and without the bound methane, were computed with the Connolly molecular surface package,⁴² as was the SASA and MSA of the methane particle by itself. From this data the buried surface area upon methane-enclosure complexation was determined. The direct Lennard-Jones interaction energy upon the binding of methane to each enclosure was similarly computed. The buried surface area multiplied by a surface tension like quality is often used to approximate to the solvent induced potential of the mean force. Together with the direct Lennard-Jones interaction energy, the total binding affinity between the methane and each enclosure can then be calculated, as has been done in many empirical methods for calculating binding affinities.

In addition, to investigate whether these two implicit solvent models can predict the non-additivity of the hydrophobic effect, the predicted pairwise additive buried surface area upon methane binding to each enclosure was also calculated. The deviation of the actual buried surface area from the pairwise additivity predicted allowed us to estimate the non-additive part of hydrophobic interactions based on these models.

4 Results and discussion

The binding free energies between methane and the model hydrophobic enclosures, as measured by FEP, are reported in Table 1. It is found that the range of binding free energies of the methane for the model enclosures is nearly 5 kcal mol⁻¹. Also reported in Table 1 are the pairwise additivity predicted binding affinity upon complexation, the buried surface area upon complexation, (both SASA and MSA), the pairwise additivity predicted buried surface area upon complexation, and the deviation between corresponding terms which gives the non-additive contributions.

Comparison for different implicit solvent models in predicting the binding affinity

From the data presented in Table 1, we can determine how well the buried surface area/molecular mechanics model predicts the binding affinity. Tuning the surface tension coefficient to minimize the mean-average-error (MAE) of fit with FEP reference data, we obtained an optimal surface tension coefficient of $\gamma = 0.00763$ kcal mol⁻¹ Å⁻² for the SASA and $\gamma = 0.03767$ kcal mol⁻¹ Å⁻² for the MSA models of these enclosures. For comparison, the surface tension for SASA determined by fitting to experimental solvation free energy for linear or branched alkanes by Honig⁴³ was 0.005 kcal mol⁻¹ Å⁻², whereas the macroscopic water-alkanes surface tension was 0.070 kcal mol⁻¹ Å⁻². Since there are no overlaps between the inserted methane and the enclosures, the buried van der Waals area is zero for all these systems. So the van der Waals surface model would just predict the binding affinity to be the direct LJ interaction energy. The predicted binding affinities *versus* the FEP reference data are reported in Fig. 3. From this figure we see that, for the most part, the MSA model performed better than the SASA model. This is indicated by a higher R^2 value (0.89 vs. 0.76) and smaller MAE (0.40 vs. 0.57), which is consistent with previous findings.^{4,8,10,11} Both of these models performed much better than the van der Waals surface area model, which has $R^2 = 0.70$ and MAE = 0.94. The SASA based model cannot differentiate the hydrophobicity between systems [J, K, K', K'', L, L', M, and M'] (predicting similar binding affinities of about -4.1 kcal mol⁻¹) nor between systems [D, E, F, F', G, G', G'', H, H', H'', I, I', and I''] (predicting similar binding affinities of about -2.2 kcal mol⁻¹), while

the MSA model performed much better for these systems and predicted the right order of hydrophobicity among these systems to some extent.

The SASA model found enclosures J and K' in which a methane molecule is bound between two hydrophobic plates to be the most hydrophobic of all the systems [J, K, K', K'', L, L', M, M']. The buried SASAs for these two systems, upon methane complexation, were the largest because large swaths of formerly accessible surface area on the faces of the plates are buried by the presence of the binding methane. For the other enclosures, several methane molecules already lie between the plates in the absence of the binding methane so that part of the surface area buried by the binding methane was already buried by the other particles. For the extreme cases of systems L and M, which are most hydrophobic, the SASA model strongly underestimates the binding affinity, because the buried SASAs are smaller for these two systems than in systems J and K'. On the other hand, the MSA model to some extent predicts the right order of binding affinity among these systems. A similar analysis of systems [D, E, F, F', G, G', G'', H, H', H'', I, I', I''], yields similar conclusions.

Comparison of MSA and SASA model predictions of the non-additivity effect

The non-additive contributions to the binding affinities of methane to all of the enclosures are listed in Table 1. We see from the table that while only small non-additive effects (± 0.2 kcal mol⁻¹) were observed for the methane trimers by the Chan and Scheraga groups,^{4,9,12,22} large cooperative effects (as large as -1.14 kcal mol⁻¹ for system L') and anti-cooperative effects (as large as 0.45 kcal mol⁻¹ for system J) were observed for these systems.

The data shown in Table 1 allows us to phenomenologically identify a connection between the sign of the non-additivity effect and the curvature of the molecular surface. To wit, for all the enclosures that exhibit an anti-cooperative effect, the molecular surfaces of the enclosures were convex without any concave or saddle parts, whereas for all enclosures that exhibit a cooperative effect, the molecular surfaces had concave or saddle parts, and the methane binds close to the saddle or concave part of the surface.

Consider systems E and G for example. The molecular surface of enclosure E has no concave part, and the binding affinity of methane is anti-cooperative, (less favored than pairwise additivity predicted). In the presence of another methane in enclosure G, the molecular surface has a saddle part, and when the methane binds to this saddle part, it shows a cooperative effect. Similar analyses can be done on systems J and K, K' and M', K' and K''. Systems I', H', and F have the same enclosure whose molecular surface exhibits a saddle region, but the inserted methane binds at different locations of the enclosure. Comparing the methane binding affinities for these systems, we found that as the methane moves closer and closer to the saddle part (from systems I' to H' to F), the binding affinity gets more and more favorable (from -1.74 kcal mol⁻¹ to -1.97 kcal mol⁻¹ to -2.63 kcal mol⁻¹), and the cooperative effect changes from relatively weak (-0.05 kcal mol⁻¹ for system I') to strong (-0.51 kcal mol⁻¹ and -0.37 kcal mol⁻¹ for systems H' and F, respectively). Similar analysis can be done for systems M', L', and K.

The cooperative effects in systems H' and L' were found to be stronger than for systems F and K respectively. This might seem to be at odds with intuition because systems F and K are more compact than systems H' and L' respectively. This is not surprising though: the distance between the two methane molecules in systems H' and L' corresponds to the de-solvation barrier on the PMF curve of two methanes molecules in bulk water. In the presence of the plate(s), the de-solvation barrier between the two methanes may not exist, so that the binding affinities might be much more favorable than would be predicted by the pairwise additivity approximation.

In addition to the data for the non-additive contribution to the changes of SASA and MSA, we can investigate whether either of these implicit solvent models can predict the non-additivity in the hydrophobic interactions. Multiplying the non-additive part of the changes in MSA and SASA in methane complexation by the surface tension coefficient obtained in the previous section, we can determine the MSA and SASA predicted non-additive contributions to the hydrophobic interactions. Fig. 4 depicts the relationship between the SASA/MSA predicted non-additive contributions and the FEP results of the non-additive hydrophobic effects. It can be seen that the MSA predicted results have a strong correlation with the FEP results while SASA predicted results anticorrelate with the FEP results.

At first glance at Fig. 4, it might seem that MSA anticorrelates with SASA, in contradiction to our common understanding of these models. To better understand this “strange” behavior, the non-additive contribution of SASA and MSA for a model methane trimer system was investigated. With the two methanes kept at their contact distance, the position of the third methane can be specified by two coordinates (θ , d). (See the top right corner of Fig. 5). Fig. 5 shows the predictions of the non-additive part of hydrophobic interactions given by calculating the changes in the MSA and the SASA each multiplied respectively by their corresponding surface tension coefficients (as in the previous section) as a function of the distance d when $\theta = 0$. Actually, this figure is representative of what we found for all of the angles θ . (See Fig. 9 of ref. ⁴) We see from this figure that for $d = 3.23$ Å, which corresponds to the configuration where the third methane is in contact with the other two methanes, the predictions of MSA and SASA have opposite signs. For all of the systems studied in this paper, the binding methane is in contact with the other methane(s) and/or plate(s) in the enclosures, thus it is not surprising to find the anticorrelation between MSA and SASA predictions observed in Fig. 4.

In addition, it can be seen from Fig. 5 that while the MSA model can predict additive, anti-cooperative, and cooperative effects, the SASA model only predicts additive and anti-cooperative effects. This conclusion seems to be generally valid. In the SASA model, the surface considered is formed from overlapping spheres, each of whose radii is the sum of the corresponding atomic van der Waals radii plus the radius of water.⁴⁰ The buried surface area between a sphere and a cluster of spheres, when the cluster of spheres has no overlaps, is always equal to the sum of the buried surface areas between the sphere and each individual sphere in the cluster, and smaller than this when the cluster of spheres has overlaps. This means that the SASA model can never predict the cooperative effects. However, in the MSA model, the surface is traced by the inward-facing surface of the probe sphere,⁴¹ and it has the potential to predict all of the non-additivity effects. This partially explains why the MSA based model performs better than the SASA based model both for the binding affinity and for the non-additivity effects.

It should be noted from Fig. 4 that while the MSA model successfully predicts how the non-additivity effects correlate with the FEP results, there were a few outliers. We will interpret these outliers case by case.

In system M', the region between the plates is enclosed on four different sides by hydrophobic moieties. This causes the density of water in the enclosed region to be much smaller than in bulk (1/4 of the bulk value), close to a hydrophobic dewetting condition. Part of the time there was one water molecule in the enclosed region and part of time it was empty, indicating an interface different from the normal nonpolar solute–liquid water interface. It is well known that in hydrophobic dewetting there is a strong driving force to bring the hydrophobic particles together,¹ which explains the strong cooperative effect observed in system M'.

For systems H' and L', as previously mentioned, the distance between the two methane molecules corresponds to the de-solvation barrier on the PMF curve of two methane molecules in bulk water, and in the presence of the plate(s), the de-solvation barrier may not exist at all. Because of the barriers in the pair potential of mean force, the pairwise additivity approximation predicts that this configuration will be very unfavorable, but the full calculation shows that in fact they are not unfavorable. In fact, the binding affinities in these particular configurations are much more favorable than would be predicted by assuming pairwise additivity. This may be the reason for the strong cooperative effects observed for these systems. In addition, the dewetting argument discussed in connection with systems M' also applies to system L', which further validates the strong cooperative effect observed for system L'.

To investigate the reason for the strong anti-cooperative effects observed in systems J and K', we analyzed the structure of water between the two plates and found that, at the surface of one hydrophobic plate (system D), water breaks one hydrogen bond on average, which caused the average interaction energy between water at the surface and the rest of the system to be higher than that in bulk water by $1.12 \text{ kcal mol}^{-1}$. However, between two hydrophobic plates (system J), water breaks less than two hydrogen bonds on average because of its flexibility in making hydrogen bonds. This is supported by the fact that the average interaction energy between water molecules located between the two plates and the rest of the system is higher than in bulk water by only $2.05 \text{ kcal mol}^{-1}$, less than twice $1.12 \text{ kcal mol}^{-1}$ ($2 \times 1.12 = 2.24$) expected from doubling the effect of one plate. It is well known that the hydrophobic effect for large scale systems is enthalpy driven,^{1,44} because of breaking hydrogen bonds at the surface of the hydrophobic plates. A large contribution of the methane-enclosure binding affinity comes from de-solvation of solvent between the plates.⁴⁵ The anti-cooperative effect observed for systems J and K' may be due to the fact that the excess of the interaction energy of water located between two hydrophobic plates (system J) over that of bulk water is found to be smaller than twice the value of the water at one hydrophobic plate (system D).

Non-additivity effect at wetting-dewetting transition

In the previous sections, we have shown how the non-additivity effects of methane binding affinities in enclosures with different topologies correlate with the MSA measurements. It is well known that the hydrophobicity of the enclosures depends not only on the topologies but also on the LJ parameters for atoms making up the enclosures.¹ So it will be interesting to study how the non-additivity effects depend on the LJ parameters for particles making up the enclosure. In this section, we will explore this effect by changing the LJ ϵ parameter for particles making up the plates for one representative enclosure, enclosure J.

Fig. 6 depicts how the binding affinities of methane in enclosure J (ΔG_J) and in enclosure D (ΔG_D) changes as a function of the LJ ϵ parameter for particles making up the plate(s) from FEP simulations. The pairwise additivity predicted binding affinities for enclosure J, which is two times that for enclosure D, are also depicted in the figure. As can be seen from this figure, while the binding affinities for enclosure D decreases (or alternatively the free energy becomes more positive) monotonically with increasing values of ϵ , the binding affinities for enclosure J first increases (the free energy becomes more negative) and then decreases, and the non-additivity effect goes from slightly anti-cooperative at very low ϵ regions (smaller than $0.06 \text{ kcal mol}^{-1}$), to cooperative at intermediate ϵ regions (between 0.06 and $0.23 \text{ kcal mol}^{-1}$), back to anti-cooperative at high ϵ regions (higher than $0.23 \text{ kcal mol}^{-1}$).

The solvation free energy of methane in the enclosure can be decomposed into two components: the free energy to create a cavity with the size of methane in the enclosure and the free energy to turn on the attractive part of interactions between methane and the rest of

the system. For enclosure D, when increasing the ϵ parameter, the free energy to create the cavity will become more unfavorable because the solvent will become denser at the surface of the plate; however, the free energy to turn on the attractive part of interactions between methane and the plate will become more favorable with increasing values of ϵ . These two factors have opposite effects, but the first component dominates, so the overall binding affinity will decrease slightly. For enclosure J, at low values of ϵ , the region between the plates dewets, so the free energy to create the cavity is almost zero and changes slightly with increasing values of ϵ , but the free energy to turn on the attractive interactions between the methane and the plates becomes more negative, so the binding affinity will increase in this dewetting region. The critical value of ϵ corresponding to the wetting–dewetting transition is $\epsilon \approx 0.15 \text{ kcal mol}^{-1}$. At this point, the probability for observing a dry inter-plate region is 50%. For ϵ larger than this, the free energy to create the cavity grows rapidly with increasing values of ϵ , becoming the dominant effect, so that the overall binding affinity decreases rapidly with increasing values of ϵ . At the critical value of ϵ , that is at the wetting–dewetting transition, there is a large cooperative non-additive effect on the binding of the methane between the plates. With increasing values of ϵ , the two plates affect the density fluctuation of solvent by more than twice what one plate does, and the slope of the binding affinity *versus* ϵ is therefore much larger for enclosure J than that for enclosure D. For sufficiently large ϵ there is a large anti-cooperative deviation from additivity. At $\epsilon = 0.23$, these two effects balance each other, so the free energy is additive at this point. The ϵ value for system J studied in the previous section is $\epsilon = 0.294 \text{ kcal mol}^{-1}$ which is higher than 0.23, so we observed an anti-cooperative effect there. As ϵ is decreased below the critical value of ϵ , the binding affinity increases for enclosure D and decreases for enclosure J and becomes anti-cooperative in this very low ϵ region.

Higher order multi-body interactions

In the previous section, we investigated the non-additive effects manifested when methane is inserted into different model enclosures. For systems consisting of three components, the non-additive effect corresponds to three-body interactions; for systems consisting of more than three components, the non-additive effect is the summation of all higher-than two-body interactions involving the insertion methane, corresponding to $\delta W(1, 2, \dots, n-1; n)$, defined in eqn(5). We now investigate the contributions beyond three body interactions defined recursively in eqn(3).

Table 2 lists the total PMF, the two-body contribution to the PMF, W_2 , the three-body contribution to the PMF, W_3 , and the subsequent higher order contributions for all the systems consisting of more than three components. The deviations found by truncating the series up to order n , ΔW_n , equivalent to the GKSA approximation to the n -th order, is also listed in the table. (For example, $\Delta W_2 = W(1, 2, \dots, n) - W_2$.) From this table, we see that the error arising from truncation of the total PMF at the pairwise term can be as large as $1.43 \text{ kcal mol}^{-1}$ (system L), indicating the importance of non-additivity effects in these systems. In addition, the higher order multi-body contributions can be either positive or negative, suggesting that hydrophobic interactions can be quite complex.

Fig. 7 and 8 gives the deviations of the PMF predicted by the GKSA from the total PMF, when the latter is truncated to order n , as a function of n . For most of the cases, the higher the order of multi-body interactions considered, the smaller the magnitude of the error (systems G, H, I, L', L, M). However, this is not generally true since for systems K and M', inclusion of the three-body interactions increases the error.

5 Conclusion

In this paper, the binding affinities between a united-atom methane and various model hydrophobic enclosures were determined using high accuracy FEP molecular dynamics simulations. Comparisons were made between the binding affinities from FEP and predictions based on assuming pairwise additivity, and through this it was possible to investigate the non-additive contributions of the hydrophobic interactions. Small non-additivity effects were found in the methane trimer systems by the Chan and Scheraga groups,^{4,9,12,22} but we find large cooperative effects (as large as $-1.14 \text{ kcal mol}^{-1}$) and anti-cooperative effects (as large as $0.45 \text{ kcal mol}^{-1}$) for our relatively larger systems. Although approximations based on the pairwise additivity of the PMF have been used to study the transition state of protein folding⁴⁶ and the force-extension behavior of protein,⁴⁷ simulations done in this paper indicate that higher order correlations may be very important in real biological systems such as protein folding, protein–ligand binding and other relevant fields. This should not be surprising since the Kirkwood superposition approximation fails in dense simple fluids.

Phenomenologically, the sign of the non-additive contributions to the binding affinities of methane in the enclosures was found to be correlated with the curvature of the enclosures. To be specific, anti-cooperative effects were observed only in enclosures whose molecular surfaces were convex without any concave or saddle part, and cooperative effects were observed in enclosures whose molecular surfaces had concave or saddle parts, in which case the methane was found to bind close to the saddle or concave part of the surface. Such observations might be useful for further development of models to incorporate the non-additivity effect.

We also investigated whether two kinds of implicit solvent models are consistent with the observed binding affinities and non-additivity effects. In these models the solvent induced free energy for particle insertion was computed from the product of a “surface tension” and the area of the buried surface upon methane complexation. The area of the buried surface was computed by using the solvent accessible surface area (SASA) and by using the molecular surface area (MSA) as described in the text. We found that the MSA based model performed much better than the SASA based model in predicting the binding affinities, an observation consistent with previous findings.^{4,8,10,11} In addition, the SASA based model always predicts non-additive effects which anticorrelate with the FEP reference data, whereas the MSA based model performed reasonably well in predicting the non-additive effects, except for a few outliers which were explained in the text. Further analysis indicated that the MSA based model predicts all cooperative and anti-cooperative non-additivity effects, and the SASA based model exhibits an intrinsic problem in that it can never predict cooperative effects. Furthermore, because of the correlation we observed between the non-additive contributions and the curvature of the molecular surface in the MSA based model, we believe that the MSA based model is a far better descriptor of the curvature of simple solutes than the SASA based model.

The non-additivity effect depends not only on the topology but also on the LJ parameters for atoms making up the enclosure. By changing the LJ ϵ parameter for atoms making up the plates for enclosure J, we observed a wetting–dewetting transition in the inter-plate region, and the non-additivity effect changes from anti-cooperative in the low ϵ region, to cooperative in the intermediate ϵ region, and then to anti-cooperative again in the higher ϵ region. This complicated reentrant behavior of the non-additivity effect results from the competition between the two factors contributing to the solvation free energy: the free energy to create the cavity and the free energy to turn on the attractive interactions between

the solute and the rest of the system. While the first factor dominates in the higher ε region (wetting region), the second factor dominates in the lower ε region (dewetting region).

The decomposition of the PMF for cluster formation can be expressed as a sum of increasing orders of multi-body interactions. We found that the multi-body correlations can be either negative or positive, implying that the hydrophobic interaction will depend on the topology of the surfaces enclosing the particle. In addition, increasing the order of multi-body interactions included can usually improve the accuracy, but this was not generally true.

Acknowledgments

We thank Dr Robert Abel and Prof. Tom Young for many helpful discussions. BJB acknowledges support by NIH GM 43340. RAF acknowledges that this work was supported in part by the National Science Foundation through TeraGrid resources provided by NCSA and ABE, (MCA08X002), and NIH GM 52018.

References

1. Berne BJ, Weeks JD, Zhou R. *Annu Rev Phys Chem* 2009;60:85–103. [PubMed: 18928403]
2. Kauzmann W. *Adv Protein Chem* 1959;14:1–63. [PubMed: 14404936]
3. Dill KA. *Biochemistry* 1990;29:7133–7155. [PubMed: 2207096]
4. Shimizu S, Chan HS. *Proteins: Struct, Funct, Genet* 2002;48:15–30. [PubMed: 12012334]
5. Li X, Liang J. *Proteins: Struct, Funct, Bioinf* 2005;60:46–65.
6. Némethy G, Scheraga HA. *J Phys Chem* 1962;66:1773–1789.
7. Bruggè F, Fornili SL, Malenkov GG, Palma-Vittorelli MB, Palma MU. *Chem Phys Lett* 1996;254:283–291.
8. Rank JA, Baker D. *Protein Sci* 1997;6:347–354. [PubMed: 9041636]
9. Shimizu S, Chan HS. *J Chem Phys* 2001;115:1414–1421.
10. Czaplewski C, Rodziewicz-Motowidlo S, Liwo A, Ripoll DR, Wawak RJ, Scheraga HA. *Protein Sci* 2000;9:1235–1245. [PubMed: 10892816]
11. Czaplewski C, Ripoll DR, Liwo A, Rodziewicz-Motowidlo S, Wawak RJ, Scheraga HA. *Int J Quantum Chem* 2002;88:41–55.
12. Czaplewski C, Liwo A, Ripoll DR, Scheraga HA. *J Phys Chem B* 2005;109:8108–8119. [PubMed: 16851948]
13. Czaplewski C, Rodziewicz-Motowidlo S, Liwo A, Ripoll DR, Wawak RJ, Scheraga HA. *J Chem Phys* 2002;116:2665–2667.
14. Shimizu S, Chan HS. *J Chem Phys* 2002;116:2668–2669.
15. Shimizu S, Moghaddam MS, Chan HS. *J Phys Chem B* 2005;109:21220–21221. [PubMed: 16853749]
16. Czaplewski C, Kalinowski S, Liwo A, Ripoll DR, Scheraga HA. *J Phys Chem B* 2005;109:21222–21224.
17. Czaplewski C, Kalinowski S, Liwo A, Scheraga HA. *Mol Phys* 2005;103:3153–3167.
18. Shimizu S, Chan HS. *J Chem Phys* 2000;113:4683–4699.
19. Moghaddam MS, Shimizu S, Chan HS. *J Am Chem Soc* 2005;127:303–316. [PubMed: 15631480]
20. Ghosh T, Garcia AE, Garde S. *J Phys Chem B* 2003;107:612–617.
21. Ghosh T, Kalra A, Garde S. *J Phys Chem B* 2005;109:642–651. [PubMed: 16851057]
22. Czaplewski C, Rodziewicz-Motowidlo S, Dabal M, Liwo A, Ripoll DR, Scheraga HA. *Biophys Chem* 2003;105:339–359. [PubMed: 14499903]
23. Mcquarrie, DA. *Statistical Mechanics*. Sausalito, CA: University Science Books; 2000.
24. Fisher IZ, Kopeliovich BL. *Dokl Akad Nauk SSSR* 1960;133:81–83.
25. Reiss H. *J Stat Phys* 1972;6:39–47.
26. Kirkwood JG. *J Chem Phys* 1935;3:300–313.
27. Kirkwood JG. *J Chem Phys* 1939;7:919–925.

28. Kirkwood JG, Boggs EM. J Chem Phys 1942;10:394–402.
29. Bowers, KJ.; Chow, E.; Xu, H.; Dror, RO.; Eastwood, MP.; Gregersen, BA.; Klepeis, JL.; Kolossvary, I.; Moraes, MA.; Sacerdoti, FD.; Salmon, JK.; Shan, Y.; Shaw, DE. SC '06: Proceedings of the 2006 ACM/IEEE conference on Supercomputing; New York, NY, USA. 2006. p. 84
30. Jorgensen WL, Madura JD, Swenson CJ. J Am Chem Soc 1984;106:6638–6646.
31. Maestro, version 8.5. Schrodinger, LLC; New York, NY: 2008.
32. Berendsen, H.; Postma, J.; van Gunsteren, W.; Hermans, J. Intermolecular Forces. Pullman, B., editor. Reidel: Dordrecht; 1981. p. 331-342.
33. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF. Chem Phys Lett 1994;222:529–539.
34. Nosé S. J Chem Phys 1984;81:511–519.
35. Hoover WG. Phys Rev A: At, Mol, Opt Phys 1985;31:1695–1697.
36. Martyna GJ, Tobias DJ, Klein ML. J Chem Phys 1994;101:4177–4189.
37. Darden T, York D, Pedersen L. J Chem Phys 1993;98:10089–10092.
38. Bennett CH. J Comput Phys 1976;22:245–268.
39. Roux B, Simonson T. Biophys Chem 1999;78:1–20. [PubMed: 17030302]
40. Lee B, Richards FM. J Mol Biol 1971;55:379–380. [PubMed: 5551392]
41. Connolly ML. Science 1983;221:709–713. [PubMed: 6879170]
42. Connolly, ML. The molecular surface package, version 3.9.3. 2006.
43. Tannor DJ, Marten B, Murphy R, Friesner RA, Sitkoff D, Nicholls A, Honig B, Ringnalda M, Goddard WA. J Am Chem Soc 1994;116:11875–11882.
44. Lum K, Chandler D, Weeks JD. J Phys Chem B 1999;103:4570–4577.
45. Abel R, Wang L, Friesner RA, Berne BJ. J Chem Theory Comput. 2010 submitted.
46. Clementi C, Nymeyer H, Onuchic JN. J Mol Biol 2000;298:937–953. [PubMed: 10801360]
47. Grater F, Heider P, Zangi R, Berne BJ. J Am Chem Soc 2008;130:11578–11579. [PubMed: 18693720]

Fig. 1.
The 13 model systems studied. The hydrophobic enclosures are depicted in gray. The location of the methane molecule when bound to the respective hydrophobic enclosure is depicted in black. The geometry of the hydrophobic plate is depicted in the right bottom of the figure.

Fig. 2.

The 14 model systems corresponding to subsystems depicted in Fig. 1. The black particle in each system denotes the methane that will bind to the enclosure which is depicted in gray, and the small white particle denotes a pseudo-particle that specified the position of the binding methane for the corresponding system in Fig. 1. (The binding affinity for system G'', H'', I'' and K'' were calculated through thermodynamic cycles by combination of the binding affinities calculated for related systems).

Fig. 3.

Buried surface area/molecular mechanics prediction of methane-enclosure binding affinities. The surface tension coefficients were chosen to minimize the MAE (mean average error) between the predicted and FEP results for the binding affinities. Predictions based on the MSA (molecular surface area) model performed better than those based on the SASA (solvent accessible surface area) model, indicated by a higher R^2 value (0.89 vs. 0.76) and smaller MAE (0.40 vs. 0.57). Both of these models performed much better than predictions based on the van der Waals surface model which takes into account the direct LJ interaction (giving $R^2 = 0.70$, and MEA = 0.94 kcal mol⁻¹). SASA can not differentiate the hydrophobicity among systems [J, K, K', K'', L, L', M, M'] (predicting a similar binding affinity of about -4.1 kcal mol⁻¹) nor among systems [D, E, F, F', G, G', G'', H, H', H'', I, I', I''] (predicting a similar binding affinity of about -2.2 kcal mol⁻¹), while the MSA based model performed much better for these systems and predicted the right order of hydrophobicity among these systems to some extent.

Fig. 4. Relationship between the surface area models predicted non-additivity of hydrophobic effects and the corresponding FEP results. There is a strong correlation between MSA (molecular surface area) predictions and FEP results, but the SASA (solvent accessible surface area) predicted results anticorrelate with the FEP results.

Fig. 5.

The MSA/SASA (molecular surface area/solvent accessible surface area) predicted non-additivity effects of hydrophobic interactions (non-additive part of the change in MSA and SASA multiplied by the corresponding surface tension coefficient) for the methane trimer system as a function of the distance d for the configuration depicted in the top right corner when $\theta = 0$. The distance corresponding to the configuration where the third methane is in contact with the remaining two were indicated by the dotted line labeled “cn”.

Fig. 6.

The methane enclosure binding affinities for enclosure D and J as a function of the LJ ϵ parameter for atoms making up the plate(s). The binding free energy increases monotonically for enclosure D with increasing values of ϵ , while it decreases at the lower ϵ region and increases at the higher ϵ region for enclosure J. The non-additivity effect for enclosure J goes from anti-cooperative in the lower ϵ region, (smaller than 0.06 kcal mol⁻¹), to cooperative in the intermediate ϵ region, (between 0.06 and 0.23 kcal mol⁻¹), and then to anti-cooperative again in the higher ϵ region (larger than 0.23 kcal mol⁻¹).

Fig. 7.

The deviation of GKSA (generalized Kirkwood superposition approximation) predicted PMF (potential of mean force) from the total PMF by truncating the total PMF to the second-, third-, and fourth- order as a function of the order.

Fig. 8.

The deviation of GKSA (generalized Kirkwood superposition approximation) predicted PMF (potential of mean force) from the total PMF by truncating the total PMF to the second-, third-, fourth- and fifth- order as a function of the order.

Table 1

The binding thermodynamics of methane for the various model hydrophobic enclosures^a

	ΔG_{FEP}	ΔG_2	$\Delta \Delta G$	ΔS_{ASA}	ΔS_{ASA_2}	$\Delta \Delta S_{\text{ASA}}$	$\Delta \Delta S_{\text{ASA}_2}$	$\Delta \Delta S_{\text{ASA}}$	$\Delta \Delta S_{\text{ASA}_2}$	$\Delta \Delta S_{\text{ASA}}$	$\Delta \Delta S_{\text{ASA}_2}$	Sign
A	-0.60	-	-	-57.44	-	-	-4.15	-	-	-	-	-
B'	-0.06	-	-	0.00	-	-	0.00	-	-	-	-	-
C'	0.18	-	-	-25.75	-	-	3.54	-	-	-	-	-
D	-1.66	-	-	-89.09	-	-	-14.44	-	-	-	-	-
E'	-0.06	-	-	0.00	-	-	0.00	-	-	-	-	-
F'	-1.64	-	-	-88.54	-	-	-14.16	-	-	-	-	-
B	-1.15	-1.21	0.06	-114.88	-114.88	0.00	-8.29	-8.29	0.00	0.00	0.00	anti
E	-2.17	-2.26	0.09	-146.53	-146.53	0.00	-18.58	-18.58	0.00	0.00	0.00	anti
G'	-1.44	-1.46	0.02	-114.29	-114.29	0.00	-10.61	-10.61	0.00	0.00	0.00	anti
J	-2.86	-3.31	0.45	-178.17	-178.17	0.00	-28.87	-28.87	0.00	0.00	0.00	anti
K'	-2.83	-3.27	0.44	-177.08	-177.08	0.00	-28.31	-28.31	0.00	0.00	0.00	anti
C	-1.41	-1.21	-0.20	-95.61	-114.88	19.27	-11.46	-8.29	-3.16	-3.16	-3.16	coop
F	-2.63	-2.26	-0.37	-115.03	-146.53	31.50	-25.14	-18.58	-6.56	-6.56	-6.56	coop
G	-3.41	-2.86	-0.54	-153.20	-203.97	50.77	-24.48	-14.76	-9.72	-9.72	-9.72	coop
G''	-2.68	-2.06	-0.62	-120.96	-171.73	50.77	-24.48	-14.76	-9.72	-9.72	-9.72	coop
H	-3.44	-2.86	-0.57	-129.81	-203.97	74.15	-38.00	-22.73	-15.27	-15.27	-15.27	coop
H'	-1.97	-1.46	-0.51	-101.36	-114.29	14.16	-13.01	-10.61	-2.40	-2.40	-2.40	coop
H''	-2.77	-2.06	-0.71	-116.14	-171.73	55.59	-25.88	-14.76	-11.12	-11.12	-11.12	coop
I	-3.47	-2.86	-0.60	-140.97	-203.97	63.00	-35.84	-22.73	-13.11	-13.11	-13.11	coop
I'	-1.74	-1.69	-0.05	-88.54	-88.54	0.00	-14.16	-14.16	0.00	0.00	0.00	coop
I''	-2.57	-2.29	-0.28	-114.49	-145.98	31.49	-24.86	-18.31	-6.55	-6.55	-6.55	coop
K	-4.59	-3.92	-0.67	-172.62	-235.61	62.99	-46.13	-33.46	-13.11	-13.11	-13.11	coop
K''	-4.56	-3.77	-0.68	-171.53	-234.52	62.99	-45.57	-32.46	-13.11	-13.11	-13.11	coop
L	-5.24	-4.52	-0.72	-164.01	-293.05	129.04	-64.10	-37.17	-26.93	-26.93	-26.93	coop
L'	-4.23	-3.09	-1.14	-176.97	-202.83	25.86	-29.57	-24.77	-4.80	-4.80	-4.80	coop
M	-5.45	-4.52	-0.94	-167.06	-293.05	125.99	-63.39	-37.17	-26.22	-26.22	-26.22	coop
M'	-3.80	-3.33	-0.48	-177.09	-177.09	0.00	-28.32	-28.32	0.00	0.00	0.00	coop

ΔG_{FEP} denotes the binding free energies from FEP, (free energy perturbation) ΔG_2 denotes the predicted binding affinities by assuming pairwise additivity, and $\Delta\Delta G$ denotes the deviation of the binding affinity results from the pairwise additivity predicted binding affinities (which corresponds to non-additivity of hydrophobic interactions defined by formula 5). ΔSASA denotes the change of SASA (solvent accessible surface area) upon methane binding, the $\Delta\Delta\text{SASA}_2$ denotes the pairwise-additive contribution to the change of SASA. $\Delta\Delta\text{SASA}$ denotes the deviation of ΔSASA_2 from ΔSASA . The notations for MSA (molecular surface area) are similarly defined. “anti” is a short-hand notation for “anti-cooperative” and “coop” stands for “cooperative”. Free energies in kcal mol^{-1} , surface area in $\text{\AA}^2$.

Table 2

Multi-body PMF (potential of mean force) calculated from FEP, (free energy perturbation) and the contribution from two-, three-, four-, five-body interactions^a

Systems	W	W ₂	ΔW ₂	W ₃	ΔW ₃	W ₄	ΔW ₄	W ₅	ΔW ₅
G	-4.90	-4.37	-0.52	-0.47	-0.06	-0.06	0.00	-	-
H	-7.04	-5.95	-1.08	-1.46	0.38	0.38	0.00	-	-
I	-6.85	-6.19	-0.65	-0.74	0.09	0.09	0.00	-	-
K	-7.42	-7.19	-0.24	0.14	-0.37	-0.37	0.00	-	-
L'	-7.07	-6.36	-0.71	-0.14	-0.56	-0.56	0.00	-	-
M'	-6.64	-6.60	-0.04	0.78	-0.82	-0.82	0.00	-	-
L	-12.31	-10.88	-1.43	-1.40	-0.03	-0.53	0.51	0.51	0.00
M	-12.09	-11.12	-0.98	-0.21	-0.73	-1.39	0.63	0.63	0.00

^aThe PMF between two plates was set to zero when $D = 7.46 \text{ \AA}$, which corresponded to the configuration for all the systems including two plates. The choice of base line for the PMF between the two plates does not affect the multi-body contributions. Free energies in kcal mol⁻¹.