

Hydrophobic Force Field as a Molecular Alternative to Surface-Area Models

G. Hummer

Contribution from the Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received December 28, 1998. Revised Manuscript Received April 29, 1999

Abstract: An effective force field for hydrophobic interactions is developed based on a modified potential-of-mean-force (PMF) expansion of the effective many-body interactions between nonpolar molecules in water. For the simplest nonpolar solutes in water, hard particles, the modified PMF expansion is exact in both limiting cases of infinite separation and perfect overlap. The hydrophobic interactions are parametrized by using the information-theory model of hydrophobic hydration. The interactions between nonpolar solutes are short-ranged and can be evaluated efficiently on a computer. The force field is compared with simulation data for alkane conformational equilibria in water as well as a model for the formation of a hydrophobic core of a protein. The modified PMF expansion can be extended to solutes with attractive interactions. The observed accuracy, computational efficiency, and atomic detail of the model suggest that this simple hydrophobic force field can lead to a molecular alternative for phenomenological surface-area models with applications in ligand-binding and protein-folding studies.

Introduction

Protein folding,^{1–3} membrane formation,⁴ drug binding to proteins,⁵ and the formation of amyloid plaques in Alzheimer's disease tissues⁶ are largely driven by hydrophobic interactions. In a simplified view, water squeezes out nonpolar molecules or groups,⁷ resulting in attractive interactions between hydrophobic solutes. From this simple picture one already expects many-body effects to be important. The most commonly used *phenomenological* model of hydrophobic interactions assumes a proportionality of hydration free energies to surface areas, corresponding to an effective many-body potential. Various definitions of surface areas, corrections for curvature on a molecular scale, and proportionality constants are in use,^{8–16} with the latter differing for dissolution and conformational changes of nonpolar molecules.¹⁷ Scaled-particle theory (SPT)¹⁸

motivates the various surface-area models of hydrophobic hydration and interactions. In SPT adapted to hydrophobic hydration,^{19,20} the exact results in the limit of small solutes are interpolated with the expected limit for macroscopically large solutes. However, molecular processes of drug binding, protein folding, and conformational equilibria fall in the gap well above the small-solute limit and far from the macroscopic limit. The resulting lack of atomic detail of surface-area models in this molecular regime may thus be seen as a reason for their ambiguities¹⁷ and shortcomings.²¹

Hydrophobic effects have been studied with numerous theoretical models and computer simulations.^{8–17,19–57} A thor-

- (1) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.
- (2) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.
- (3) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545.
- (4) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; John Wiley & Sons: New York, 1973.
- (5) Levy, R. M.; Gallicchio, E. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531.
- (6) Fink, A. L. *Folding Des.* **1998**, *3*, R9.
- (7) Richards, F. M. *Sci. Am.* **1991**, *264*, 54.
- (8) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754.
- (9) Chothia, C. *Nature* **1974**, *248*, 338.
- (10) Oobatake, M.; Ooi, T. *Prog. Biophys. Mol. Biol.* **1993**, *59*, 237.
- (11) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.
- (12) Sharp, K. A.; Nicholls, A.; Fine, R. F.; Honig, B. *Science* **1991**, *252*, 106.
- (13) Simonson, T.; Brünger, A. T. *J. Phys. Chem.* **1994**, *98*, 4683.
- (14) Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1994**, *98*, 4141.
- (15) Tuñón, I.; Silla, E.; Pascual-Ahuir, J. L. *J. Phys. Chem.* **1994**, *98*, 377.
- (16) Fukunishi, Y.; Suzuki, M. *J. Phys. Chem.* **1996**, *100*, 5634.
- (17) Ashbaugh, H. S.; Kaler, E. W.; Paulaitis, M. E. *Biophys. J.* **1998**, *75*, 755.
- (18) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. *J. Chem. Phys.* **1959**, *31*, 369.

- (19) Pierotti, R. A. *J. Phys. Chem.* **1963**, *67*, 1840.
- (20) Stillinger, F. H. *J. Solution. Chem.* **1973**, *2*, 141.
- (21) Rank, J. A.; Baker, D. *Protein Sci.* **1997**, *6*, 347.
- (22) Lee, B. *Biopolymers* **1985**, *24*, 813.
- (23) Lee, B. *Biopolymers* **1991**, *31*, 993.
- (24) Gibson, K. D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 480.
- (25) Hopfinger, A. J. *Macromolecules* **1971**, *4*, 731.
- (26) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1987**, *91*, 4105.
- (27) Blokzijl, W.; Engberts, J. B. F. *N. Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545.
- (28) Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13*, 507.
- (29) Némethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1962**, *36*, 3382.
- (30) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683.
- (31) Ben-Naim, A. *Hydrophobic Interactions*; Plenum: New York, 1980.
- (32) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 1712.
- (33) Lazaridis, T.; Paulaitis, M. E. *J. Phys. Chem.* **1992**, *96*, 3847.
- (34) Chandler, D. *Phys. Rev. E* **1993**, *48*, 2898.
- (35) Besseling, N. A. M.; Lyklema, J. *J. Phys. Chem. B* **1997**, *101*, 7604.
- (36) (a) Madan, B.; Sharp, K. *J. Phys. Chem. B* **1997**, *101*, 11237. (b) Matubayasi, N.; Reed, L. H.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 10640.
- (37) Kinoshita, M.; Hirata, F. *J. Chem. Phys.* **1997**, *106*, 5202.
- (38) Hummer, G.; Garde, S. *Phys. Rev. Lett.* **1998**, *80*, 4193.
- (39) Pangali, C.; Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, *71*, 2975.
- (40) Lee, C. Y.; McCammon, J. A.; Rossky, P. J. *J. Chem. Phys.* **1984**, *80*, 4448.
- (41) Watanabe, K.; Andersen, H. C. *J. Phys. Chem.* **1986**, *90*, 795.
- (42) Smith, D. E.; Haymet, A. D. J. *J. Chem. Phys.* **1993**, *98*, 6445.
- (43) van Belle, D.; Wodak, S. J. *J. Am. Chem. Soc.* **1993**, *115*, 647.

oughly molecular model of hydrophobic hydration and interactions has been developed⁵⁸ that uses information-theory (IT) concepts⁵⁹ to calculate approximately the free energy of cavity formation^{49–52} in water. In its simplest form, the IT model is related to Pratt–Chandler integral equation theory³⁰ and Gaussian field models³⁴ (for a discussion, see refs 58, 60–62). IT calculations led to a molecular model for the temperature dependence of hydrophobic hydration relevant for hydrophobic contributions to protein-unfolding entropies.⁶³ In addition, IT calculations also led to a novel description of the pressure-induced denaturation of proteins,⁶⁴ addressing a longstanding puzzle in protein stability.⁶⁵ However, the simplest IT model requires the calculation of absolute solvation free energies.^{58,61} This limits applications requiring high computational efficiency such as drug binding to proteins. Extensions of the IT model to inhomogeneous systems have been discussed.⁶¹ However, those are more complicated, involving solute–water correlation functions.

Here, we follow a different route to developing a simple, effective force-field model of hydrophobic interactions. We start from an expansion of the n -body interaction free energy into two-body, three-body, and higher-order contributions. That modified potential-of-mean-force (PMF) expansion is motivated by our earlier work revealing the shortcomings of the conventional PMF expansion in the limit of close contact or overlap.⁶⁶ For the simplest nonpolar solutes in water, hard particles, the modified PMF expansion is exact in the two limiting cases of infinite separation and perfect overlap, even when solute–solute interactions are not included. For soft particles with attractive interactions, the expansion is exact in those two limits within first-order perturbation theory.

We then use the IT model to parametrize the resulting force field for hydrocarbon solutes. In addition to two-body interactions, we derive a simple expression for the many-body terms involving volumes of intersection between water-exclusion

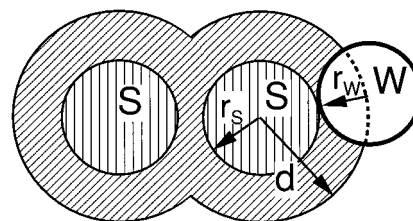


Figure 1. Schematic representation of the excluded volume of two solute spheres S , each of radius r_s . Also shown is a water molecule W of radius r_w . The joint excluded volume of the two solute spheres is the total shaded area, defined as the union of the two spheres of radius $d = r_s + r_w$ inaccessible to water. For methane and CH_n groups, an exclusion radius of $d = 0.33$ nm is used, corresponding to the distance at which the solute–water radial distribution function between methane and water for typical models reaches a value of one.⁵⁸

spheres. This results in a short-ranged and computationally efficient expression for the many-body contributions to the force field.

We compare the predictions of the hydrophobic force field with various simulation data for free energies of alkane conformational equilibria in water.⁶⁷ In addition, we also compare it with a model for the formation of the hydrophobic core of a protein involving a cluster of fourteen methane molecules in water.²¹

The resulting force-field description of hydrophobic interactions provides a computationally efficient, molecular alternative to phenomenological surface-area models. Potential applications include drug binding to proteins and protein-folding phenomena.

Modified Potential-of-Mean-Force Expansion

In the following, we consider the solvation free energy $\mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n)$ of n identical hard-sphere solutes in water, with straightforward generalizations to nonidentical, nonspherical particles. The solute particles $1, \dots, n$ can, e.g., represent the carbon groups of an alkane, or of a nonpolar amino acid side chain such as leucine. The size of the hard spheres corresponds to the water-exclusion volume. For methane and CH_n groups, a typical distance of closest approach of a water–oxygen atom to the center of the sphere is 0.33 nm.⁵⁸ The excluded volume of two solute spheres is illustrated in Figure 1. The free energy of hydration of n hard-sphere solutes is that of forming a cavity corresponding to the union of the exclusion volumes v_i of the particles $1, \dots, n$,

$$\mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) = \mu\left(\bigcup_{i=1}^n v_i\right) \quad (1)$$

We know from previous work on the structural hydration of nonpolar solutes in water⁶⁶ that a conventional PMF expansion for the free energy,⁶⁸

$$\mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) = \sum_{i=1}^n \delta\mu^{(1)}(\mathbf{r}_i) + \sum_{\substack{i,j=1 \\ i < j}}^n \delta\mu^{(2)}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{\substack{i,j,k=1 \\ i < j < k}}^n \delta\mu^{(3)}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (2)$$

is inadequate for chemically bonded carbon groups. $\delta\mu^{(1)}(\mathbf{r}_i) = \mu^{(1)}(\mathbf{r}_i)$ is the free energy of hydration of a single particle,

(67) Ashbaugh, H. S.; Garde, S.; Hummer, G.; Kaler, E. W.; Paulaitis, M. E. *Biophys. J.* **1999**, in press.

(68) Münster, A. *Statistical Thermodynamics*; Springer: Berlin, 1969, Chapter 5.4, p 338.

(44) Wallqvist, A.; Berne, B. J. *J. Phys. Chem.* **1995**, *99*, 2893.

(45) Head-Gordon, T. *J. Am. Chem. Soc.* **1995**, *117*, 501.

(46) Garde, S.; Hummer, G.; Paulaitis, M. E. *Faraday Discuss.* **1996**, *103*, 125.

(47) Lüdemann, S.; Abseher, R.; Schreiber, H.; Steinhauser, O. *J. Am. Chem. Soc.* **1997**, *119*, 4206.

(48) Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55.

(49) Pratt, L. R.; Pohorille, A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2995.

(50) Guillot, B.; Guissani, Y. *J. Chem. Phys.* **1993**, *99*, 8075.

(51) Forsman, J.; Jönsson, B. *J. Chem. Phys.* **1994**, *101*, 5116.

(52) Beutler, T. C.; Béguelin, D. R.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 3787.

(53) Beglov, D.; Roux, B. *J. Chem. Phys.* **1994**, *100*, 9050.

(54) Silverstein, K. A. T.; Haymet, A. D. J.; Dill, K. A. *J. Am. Chem. Soc.* **1998**, *120*, 3166.

(55) Arthur, J. W.; Haymet, A. D. J. *J. Chem. Phys.* **1998**, *109*, 7991.

(56) Cheng, Y. K.; Rosky, P. *J. Nature* **1998**, *392*, 696.

(57) Lum, K.; Luzar, A. *Phys. Rev. E* **1997**, *56*, R6283.

(58) Hummer, G.; Garde, S.; García, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951.

(59) Jaynes, E. T. In *E. T. Jaynes: Papers on Probability, Statistics, and Statistical Physics*; Rosenkrantz, R. D., Ed.; Reidel: Dordrecht, Holland, 1983.

(60) Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8800.

(61) Hummer, G.; Garde, S.; García, A. E.; Paulaitis, M. E.; Pratt, L. R. *J. Phys. Chem. B* **1998**, *102*.

(62) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*, 4570.

(63) Garde, S.; Hummer, G.; García, A. E.; Paulaitis, M. E.; Pratt, L. R. *Phys. Rev. Lett.* **1996**, *77*, 4966.

(64) Hummer, G.; Garde, S.; García, A. E.; Paulaitis, M. E.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1552.

(65) Kauzmann, W. *Nature (London)* **1987**, *325*, 763.

(66) Garde, S.; Hummer, G.; García, A. E.; Pratt, L. R.; Paulaitis, M. E. *Phys. Rev. E* **1996**, *53*, R4310.

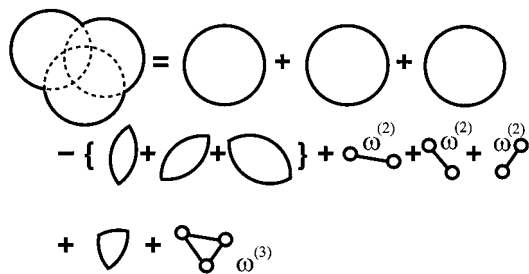


Figure 2. Schematic representation of the modified PMF expansion for three spherical solutes. Volume terms $\mu^{(1)}(\mathbf{r}_i)$, $\mu(v_i \cap v_j)$, and $\mu(v_1 \cap v_2 \cap v_3)$, are shown as outlined areas; pair and triplet interactions are shown as pair-bonds $\omega^{(2)}$ and a triangle $\omega^{(3)}$, respectively.

$\delta\mu^{(2)}(\mathbf{r}_i, \mathbf{r}_j)$ is a two-body PMF, etc. This is caused by an over-counting of solvent-excluded volumes, such that in the limit of perfect overlap of the n particles, one would have to include all orders $\delta\mu^{(1)}$ to $\delta\mu^{(n)}$ for the expansion to become exact.

We can avoid the problem of overlapping volumes by modifying the PMF expansion eq 2

$$\begin{aligned} \mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \sum_{i=1}^n \delta\mu^{(1)}(\mathbf{r}_i) \\ &+ \sum_{\substack{i,j=1 \\ i < j}}^n [\omega^{(2)}(\mathbf{r}_i, \mathbf{r}_j) - \mu(v_i \cap v_j)] \\ &+ \sum_{\substack{i,j,k=1 \\ i < j < k}}^n [\omega^{(3)}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \mu(v_i \cap v_j \cap v_k)] + \dots \\ &+ \omega^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) - (-1)^n \mu\left(\bigcap_{i=1}^n v_i\right) \end{aligned} \quad (3)$$

where $\mu(v_i \cap v_j)$ and $\mu(v_i \cap v_j \cap v_k)$ are the free energies of forming cavities with shape and size corresponding to the volumes of intersection of two (i, j) and three particles (i, j, k), respectively. Note that the contributions from intersection volumes will be zero beyond some order l of intersections. Figure 2 schematically represents the modified PMF expansion for three spheres.

Equation 3 can easily be inverted⁶⁸ to give the effective pair, triplet, and higher-order interactions $\omega^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$, $\omega^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, and $\omega^{(k)}(\mathbf{r}_1, \dots, \mathbf{r}_k)$. For the pair interaction, we find

$$\begin{aligned} \omega^{(2)}(\mathbf{r}_1, \mathbf{r}_2) &= \mu^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - \delta\mu^{(1)}(\mathbf{r}_1) - \delta\mu^{(1)}(\mathbf{r}_2) \\ &+ \mu(v_1 \cap v_2) \end{aligned} \quad (4)$$

For distances at which the cavity volumes of the two particles do not overlap ($v_1 \cap v_2 = 0$), $\omega^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ is identical to the conventional PMF $\delta\mu^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ defined by eq 2. For shorter distances, a correction $\mu(v_1 \cap v_2)$ is added that guarantees that $\omega^{(2)}$ vanishes at perfect overlap, $\omega^{(2)}(\mathbf{r}, \mathbf{r}) = 0$.

Equation 3 is exact with $\omega^{(2)} = \omega^{(3)} = \dots = \omega^{(n)} = 0$ in both limits of infinite separation and of complete overlap of the particles $1, \dots, n$. In the infinite-separation limit ($|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$ for all i, j), we simply have $\mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) = \sum_{i=1}^n \mu^{(1)}(\mathbf{r}_i)$. In the limit of perfect overlap of n identical particles, we obtain the correct result $\mu^{(n)}(\mathbf{r}, \dots, \mathbf{r}) = \mu^{(1)}(\mathbf{r})$ because each of the volumes of intersection $v_i \cap v_j$, $v_i \cap v_j \cap v_k$, etc., is identical to that of a single sphere, such that

$$\mu^{(n)}(\mathbf{r}, \dots, \mathbf{r}) = -\mu^{(1)}(\mathbf{r}) \sum_{k=1}^n (-1)^k \binom{n}{k} = \mu^{(1)}(\mathbf{r}) \quad (5)$$

where we used that the sum of binomial coefficients is $(1 - 1)^n - 1$. Equation 5 is a consequence of the inclusion-exclusion principle. By definition, we therefore have

$$\omega^{(2)}(\mathbf{r}, \mathbf{r}) = \omega^{(3)}(\mathbf{r}, \mathbf{r}, \mathbf{r}) = \dots = \omega^{(n)}(\mathbf{r}, \dots, \mathbf{r}) = 0 \quad (6)$$

that is, the effective interactions $\omega^{(k)}$ are zero both in the limit of infinite separation and perfect overlap. We expect that truncation of the ω expansion at the pair level introduces only a small error, with many-body effects contained largely in the intersection-volume terms, $\mu(\cap_i v_i)$.

Additional Many-Body Effects

Many-body interactions have not been eliminated completely in the modified PMF expansion. For a single sphere interacting with $n - 1$ perfectly overlapping spheres, the many-body interactions $\omega^{(n)}$ can be calculated exactly

$$\omega^{(n)}(\mathbf{r}, \mathbf{r}, \dots, \mathbf{r}, \mathbf{s}) = (-1)^n \omega^{(2)}(\mathbf{r}, \mathbf{s}) \quad (7)$$

This means that the higher-order correlations for this particular case are of identical magnitude and alternating sign. The error introduced by truncating the modified PMF expansion eq 3 at the pair level is

$$\mu^{(n)}(\mathbf{r}, \mathbf{r}, \dots, \mathbf{r}, \mathbf{s}) - \mu_{\text{pair-trunc}}^{(n)}(\mathbf{r}, \mathbf{r}, \dots, \mathbf{r}, \mathbf{s}) = (2 - n) \omega^{(2)}(\mathbf{r}, \mathbf{s}) \quad (8)$$

This is the expected result correcting for the over-counting of $n - 2$ pair interactions when $n - 1$ particles perfectly overlap.

From this analysis, we can estimate the effect of neglected many-body interactions between a single methane probe and the chemically bonded carbon groups in a linear alkane molecule. Truncation at the pair level overestimates the total interaction free energy $\sum \omega^{(k)}$ ($k \geq 2$) by approximately a factor of 2 since every pair of bonded carbons is close to the overlap limit. This suggests two corrections: (1) by using an effectively scaled $\omega^{(2)}$, or (2) by replacing the two bonded groups with an interaction site on the bond⁶⁹ in the calculation of the pair interactions, $\sum \omega^{(2)}(\mathbf{r}_i, \mathbf{r}_j)$. Here, we will scale the theoretical $\omega^{(2)}$ by a factor 0.5 to account for many-body effects neglected in a pair-level truncation of the modified PMF expansion.

Alternatively, we could include the higher-order interactions $\omega^{(k)}$ for $k \geq 3$ calculated from IT. Equation 7 also suggests approximate expressions for $\omega^{(k)}$ for $k \geq 3$:

$$\begin{aligned} \omega^{(k)}(\mathbf{r}_1, \dots, \mathbf{r}_k) &\approx (-1)^k \omega^{(2)}\left(\max_{1 \leq i, j \leq k} |\mathbf{r}_j - \mathbf{r}_i|\right) \times \\ &f_k\left(\min_{1 \leq i, j \leq k} |\mathbf{r}_j - \mathbf{r}_i|\right) \end{aligned} \quad (9)$$

where $f_k(r)$ is an appropriately chosen scaling function.

In applications involving a large number of interaction sites, such as drug binding to proteins, both uncertainties in the potential parametrization and additional many-body effects contained in $\omega^{(k)}$ for $k \geq 3$ are likely to become relevant. In those cases, proximity approximations involving only interactions with the nearest sites^{46,66,67,70} could provide suitable alternatives to including $\omega^{(k)}$ with $k \geq 3$.

Soft Repulsive and Attractive Solute-Water Interactions

The modified PMF expansion can be generalized to solutes with soft repulsive and attractive solute-solvent interactions.

(69) Pratt, L. R.; Chandler, D. *Methods Enzymol.* **1986**, *127*, 48.

(70) Ashbaugh, H. S.; Paulaitis, M. E. *J. Phys. Chem.* **1996**, *100*, 1900.

We first consider a collection of hard sphere solutes of equal size in water with an additional weak attractive solute–water interaction (e.g., $1/r^6$ or Yukawa potential). The goal is again to find a modified PMF expansion that is exact in both limits of perfect overlap and infinite separation with $\omega^{(k)} = 0$. Perturbation theory⁷¹ approximates the hydration free energy of an individual solute as

$$\mu^{(1)} \approx \mu_{\text{HS}} + \langle u_{\text{att}} \rangle \quad (10)$$

where μ_{HS} is the hydration free energy of the hard-sphere solute, and $\langle u_{\text{att}} \rangle$ is the attractive solute–water interaction energy, averaged over the hydration structure of the hard solute. This first-order perturbation treatment was found to give accurate results for a wide range of solute sizes and solute–water interaction strengths.⁷¹

In the limit of n overlapping hard-sphere solutes with attractive solute–water interactions, the corresponding perturbation expression for the hydration free energy is

$$\mu^{(n)} \approx \mu_{\text{HS}} + n \langle u_{\text{att}} \rangle \quad (11)$$

The following modified PMF expansion will reproduce this first-order perturbation result

$$\begin{aligned} \mu^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \sum_{i=1}^n \delta \mu^{(1)}(\mathbf{r}_i) \\ &+ \sum_{\substack{i,j=1 \\ i < j}}^n [\omega^{(2)}(\mathbf{r}_i, \mathbf{r}_j) - \mu_{\text{hard}}(v_i \cap v_j)] \\ &+ \sum_{\substack{i,j,k=1 \\ i < j < k}}^n [\omega^{(3)}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \mu_{\text{hard}}(v_i \cap v_j \cap v_k)] + \dots \\ &+ \omega^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) - (-1)^n \mu_{\text{hard}}\left(\bigcap_{i=1}^n v_i\right) \end{aligned} \quad (12)$$

In eq 12, many-body effects due to overlapping volumes of water-exclusion spheres are taken into account by including the cavity-formation free energies of hard particles μ_{hard} .

For soft repulsive and attractive solute–water interactions, we can either define effective hard-sphere radii and use eq 12 to define the corresponding effective interactions $\omega^{(k)}$, or we can separate the solute–water interactions into repulsive and attractive contributions, by using for instance a Weeks–Chandler–Andersen separation.⁷² In the latter case, a suitable definition for the solute–water interaction potential of the “intersection” of n solutes with a distance-dependent repulsive solute–water interaction $u_{\text{rep}}(r)$ is obtained by using the interaction of a water molecule at position \mathbf{r} with the most distant solute site

$$u_{\text{rep}}^{(n)}(\mathbf{r}; \mathbf{r}_1, \dots, \mathbf{r}_n) = u_{\text{rep}}(\max_{1 \leq k \leq n} |\mathbf{r} - \mathbf{r}_k|) \quad (13)$$

For hard solute–water interactions, this definition coincides with that of intersecting cavity volumes. A suitable PMF expansion

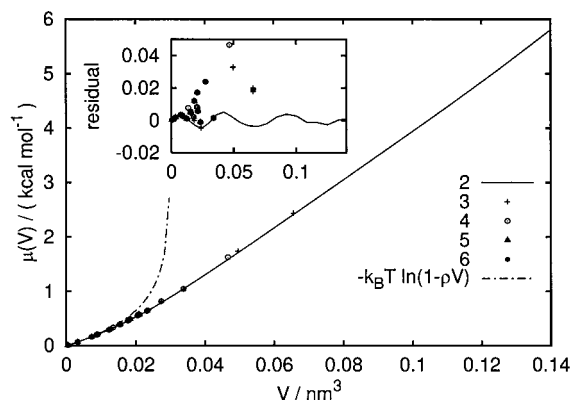


Figure 3. Hydration free energy in units of kcal/mol for particles corresponding to the volumes of intersection of two (solid line), three (plus symbols), four (open circles), five (triangles), and six spheres (filled circles) as a function of their volume V . The free energies are calculated from IT for intersection volumes corresponding to the six hexane trajectories of Table 2, with dihedral angle steps of 10° . Also shown is the limiting curve $-k_B T \ln(1 - \rho V)$ from SPT which is exact for $V \rightarrow 0$ (dot–dashed line). The inset shows the deviations from the curve fitted to the two-sphere data, as described in the Appendix.

would then define the effective interactions $\omega^{(k)}$ by replacing μ_{hard} in eq 12 with μ_{rep} , corresponding to solute–water interactions defined by eq 13.

Note that even for strong solute–water interactions such as hydrogen bonds or ionic interactions, we can use the modified PMF expansion to define effective interactions $\omega^{(k)}$ and expect that many-body contributions due to volume exclusion are approximately included. Additional many-body effects due to direct electrostatic interactions have been studied by Ashbaugh and Paulaitis.⁷³ However, further analysis and comparisons with explicit computer simulations are required to establish the usefulness of an $\omega^{(k)}$ -expansion for hydrogen-bond or ionic interactions.

Parametrization of the Hydrophobic Force Field

We simplify and parametrize the modified PMF expansion eq 3 to derive a many-body force-field representation of hydrophobic interactions. The IT model of hydrophobic hydration^{58,61} allows us to calculate approximate free energies of forming molecular-size cavities in water. Here, we use it to determine the chemical potential of hard particles defined by the volume of intersection of k spheres. In the following numerical calculations, the spheres are assumed to be of methane size (i.e., 0.33 nm distance of closest approach of a water oxygen to the center of the sphere).⁵⁸

First, we observe that the volumes of intersection of k spheres of equal radii are convex-shaped objects with a volume less than that of a single sphere. This lets us assume that the chemical potential of all these intersection volumes is approximately a monotonic function of their volume, irrespective of the specific shapes. This is indeed supported by the IT calculations. Shown in Figure 3 is the chemical potential of particles corresponding to the volume of intersection of two and more spheres as a function of their volume V . We find that free energies for representative volumes of intersection of more than two spheres fall almost exactly onto the curve of the two-sphere overlap. We also find the correct limiting behavior $-k_B T \ln(1 - \rho V)$ for $V \rightarrow 0$, with k_B the Boltzmann constant, T the temperature, and ρ the bulk density of the solvent water.^{18–20}

(71) Garde, S.; García, A. E.; Pratt, L. R.; Hummer, G. *Biophys. Chem.* **1999**, *78*, 21.

(72) Weeks, J. D.; Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 5237.

(73) Ashbaugh, H. S.; Paulaitis, M. E. *J. Phys. Chem. B* **1998**, *102*, 5029.

Table 1. Nodes r_i (in nm), Function Values y_i (in kcal/mol), and Second Derivatives y_i'' [in kcal/(mol nm²)] of the Pair Interaction $\omega^{(2)}(r)$, as Represented by a Cubic Spline Function

r_i	y_i	y_i''
0.000000000	0.000000000	0.000000000
0.241081824	0.355013511	-10.160976371
0.440195386	0.351795405	0.066022732
0.643794334	0.097754778	-26.618586944
0.668034896	0.023946352	55.015275692
0.827075306	0.001397748	-13.222689211
0.983437774	0.002274700	3.041789473
1.200000000	0.000000000	-1.375389056

This allows us to approximate the chemical potential of excluded-volume particles formed by the intersection of k spheres as a function of their volume irrespective of shape

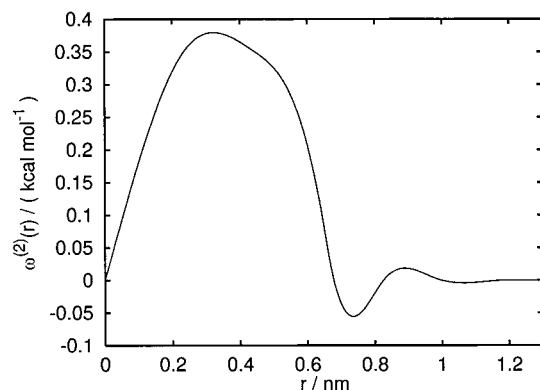
$$\mu(\nu_1 \cap \nu_2 \cap \dots \cap \nu_k) \approx \bar{\mu}(\nu_1 \cap \nu_2 \cap \dots \cap \nu_k) \quad (14)$$

where $\bar{\mu}(V)$ is a function of the magnitude of the volume V irrespective of its shape. $\bar{\mu}(V)$ has been fitted to the IT results for the overlap of two spheres of identical radii (see Appendix). This simplifies the practical calculation of the many-body interactions $\mu(\nu_i \cap \nu_j)$, $\mu(\nu_i \cap \nu_j \cap \nu_k)$, etc., to the calculation of the volume of intersection of spheres.^{74–85}

We can use IT also to calculate the effective interactions $\omega^{(n)}$. Here, we truncate the ω expansion at the pair level, assuming $\omega^{(k)} = 0$ for $k \geq 3$. The pair interaction $\omega^{(2)}$ is defined in eq 4. All terms on the right-hand side can be calculated by using the IT model. However, we scale $\omega^{(2)}$ by a factor of 0.5 to account approximately for neglected higher-order interactions, as discussed above. Table 1 compiles the positions of the nodes, amplitudes, and second derivatives of the resulting scaled $\omega^{(2)}(r)$ ($r = |\mathbf{r}_2 - \mathbf{r}_1|$) in a cubic-spline representation⁸⁶ with a cutoff of $r = 1.2$ nm [$\omega^{(2)}(r) = 0$ for $r \geq 1.2$ nm]. Figure 4 shows the pair interaction $\omega^{(2)}(r)$. We find a maximum for $\omega^{(2)}(r)$ of about 0.4 kcal/mol at $r \approx 0.3$ nm, correcting for the overestimated free-energy gains of bringing cavities to overlap when only the volume contributions are considered in eq 3. A minimum near $r \approx 0.7$ nm stabilizes solvent-separated configurations.

Comparison to Simulation Data

In the following, we compare the hydrophobic force field to explicit computer simulation data by Ashbaugh et al.⁶⁷ for alkane conformations in water, and by Rank and Baker²¹ for a methane cluster in water. Ashbaugh et al.⁶⁷ calculated the solvent contribution to the free energy (or PMF) along six trajectories of hexane, two of pentane, and one of butane, each corresponding to changing one of the dihedral angles and keeping the others fixed (see Table 2). Rank and Baker²¹ studied the free energy of changing the length of the edge of a cubic cluster of 14

**Figure 4.** Pair interaction $\omega^{(2)}(r)$ in units of kcal/mol as a function of the distance r in units of nm (solid line). The $\omega^{(2)}(r)$ shown was calculated from IT and scaled by a factor 0.5. A quadratic function was added between $r = 0$ and 0.325 nm to correct for the non-additivity in the IT free energy for two spheres.⁶¹ A cubic-spline representation of $\omega^{(2)}(r)$ is given in Table 1.**Table 2.** Simulation Trajectories⁶⁷ for Free-Energy Profiles of Alkanes along Torsional Degrees of Freedom^a

trajectory	φ_1	φ_2	φ_3
Butane			
x	ϕ		
Pentane			
gx	60°	ϕ	
tx	180°	ϕ	
Hexane			
$g'xg$	60°	ϕ	60°
txt	180°	ϕ	180°
$gg'x$	60°	60°	ϕ
gtx	60°	180°	ϕ
$xg't$	ϕ	-60°	180°
$x'tt$	ϕ	180°	180°

^a ϕ denotes the variable torsional angle. g , g' , and t stand for gauche, gauche-prime and trans dihedral angles (60° , -60° , and 180°).

methane molecules in a face-centered cubic lattice. Bringing the methane molecules into contact corresponds to the formation of a hydrophobic cluster comparable in size to those observed in protein cores.

Figure 5 compares the simulation⁶⁷ and force-field results for the solvent-contributions to the conformational free energies of the nine alkane trajectories. We find excellent agreement for all simulation PMFs. The modified PMF expansion reproduces not only the positions of the minima but also the amplitudes of the curves. The main difference is that the simulation curves tend to be somewhat more structured for dihedral angles φ , $120^\circ < |\varphi| \leq 180^\circ$. Note, however, that the estimated statistical errors in the simulation data are about 0.2–0.4 kcal/mol and thus larger than the observed differences.

Rank and Baker²¹ studied the hydration free energy of a cluster of 14 methane molecules in a face-centered cubic arrangement. A strong hydrophobic attraction was observed for the formation of a mini-hydrophobic core, together with a high barrier for cluster formation. Surface-area models failed qualitatively and quantitatively to reproduce the observed hydration free energy for cluster formation.

Figure 6 compares the force-field and simulation curves for the hydration free energy as a function of the length d of the cubic cluster. Again, we observe good agreement between the optimized force field and the simulation data. The main differences are a shift by about 0.03 nm towards larger r values, and a less pronounced second maximum. The latter might be caused in part by cutoff effects in the simulations. Alternatively,

(74) Powell, M. J. D. *Mol. Phys.* **1964**, *7*, 591.
 (75) Gibson, K. D.; Scheraga, H. A. *J. Phys. Chem.* **1987**, *91*, 4121.
 (76) Gibson, K. D.; Scheraga, H. A. *Mol. Phys.* **1987**, *62*, 1247.
 (77) Dodd, L. R.; Theodorou, D. N. *Mol. Phys.* **1991**, *72*, 1313.
 (78) Sastry, S.; Corti, D. S.; Debenedetti, P. G.; Stillinger, F. H. *Phys. Rev. E* **1997**, *56*, 5524.
 (79) Liang, J.; Edelsbrunner, H.; Fu, P.; Sudhakar, P. V.; Subramaniam, S. *Proteins: Struct., Funct., Genet.* **1998**, *33*, 1.
 (80) Richmond, T. J. *J. Mol. Biol.* **1984**, *178*, 63.
 (81) Guerrero-Ruiz, G.; Ocádiz-Ramírez, A.; Garduño-Juárez, R. *Comput. Chem.* **1991**, *15*, 351.
 (82) Connolly, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 1118.
 (83) Gogonea, V.; Osawa, E. *J. Comput. Chem.* **1995**, *16*, 817.
 (84) Petitjean, M. *J. Comput. Chem.* **1994**, *15*, 507.
 (85) Lustig, R. *Mol. Phys.* **1985**, *55*, 305.
 (86) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran*, 2nd ed.; Cambridge University Press: New York, 1992, Chapter 3.3, pp 107–110.

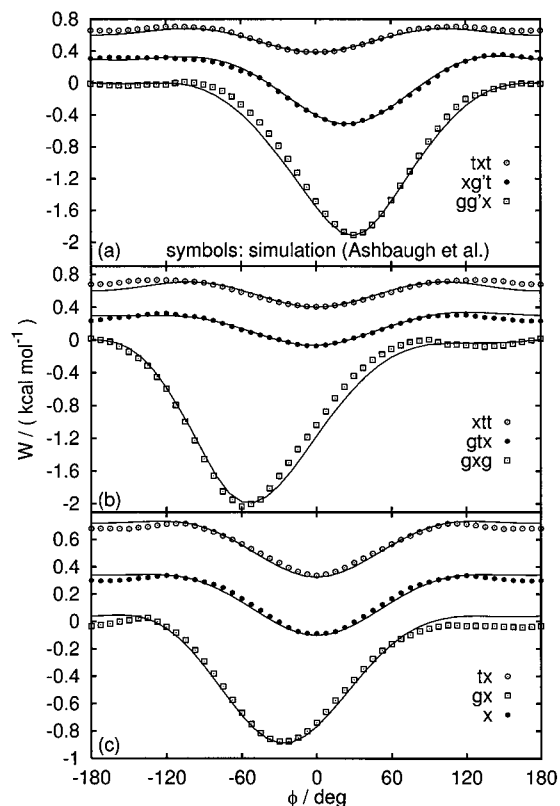


Figure 5. Comparison of the hydration free energies for alkane trajectories obtained from simulations (symbols)⁶⁷ and the hydrophobic force field (lines). The conformational free energies (or PMFs) are shown in units of kcal/mol as a function of a dihedral angle in degrees (see Table 2 for details). Panels a and b show results for the six hexane trajectories. Panel c shows the results for the butane and the two pentane trajectories. The total free-energy profiles can be obtained by adding the internal energy contributions to the hydration free energies shown here.

the scaling of $\omega^{(2)}(r)$ by a factor of 0.5 might underestimate the pair interactions between non-overlapping methane molecules. To investigate this possibility, we have scaled $\omega^{(2)}(r)$ from IT only for distances r smaller than $r = 0.6778$ nm, where it crosses the zero axis just beyond the overlap limit of $r = 0.66$ nm. The resulting free-energy profile is also shown in Figure 6 and indeed has a more pronounced second maximum and a solvent-separated minimum at $d \approx 1$ nm considerably lower than the limit of infinite separation. We note that the alkane curves of Figure 5 are left unchanged by this partial scaling, as the largest pair distance even for the fully extended hexane is below the overlap limit of $r = 0.66$ nm. Inclusion of $\omega^{(3)}$ is not likely to improve the agreement further. The three-body interactions $\omega^{(3)}$ are expected to be of short range with only small contributions at the distance of the second maximum in the cluster formation PMF of Figure 6.

The positional shift of the the free energy curves may be due in part to a different water model (TIP4P vs SPC water; whereas SPC water was used consistently in the alkane studies discussed above). In addition, large statistical uncertainties were reported for the simulation data caused by insufficient statistical convergence.²¹ Figure 8 of ref 21 illustrates these uncertainties in the simulation curves, resulting in considerable variations in the positions of the extrema as well as the amplitudes. For instance, the position of the desolvation barrier in the five reported curves varies by about 0.1 nm between $r = 0.75$ and 0.85 nm. The observed agreement between the free-energy profile calculated from the modified PMF expansion and the

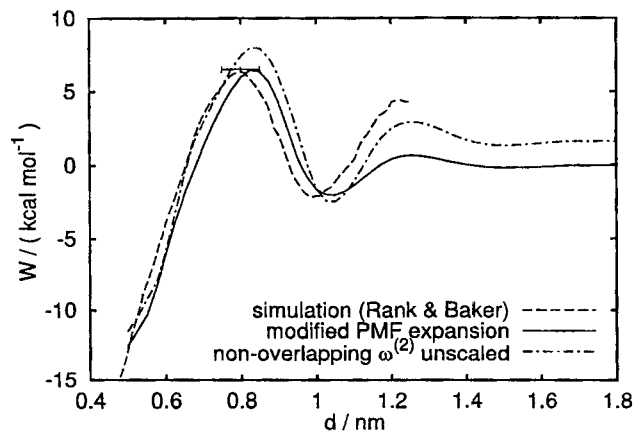


Figure 6. Comparison of the free energies of hydration of a cubic 14-methane cluster in a face-centered cubic arrangement obtained from simulation (dashed line)²¹ and the hydrophobic force field (solid line). Also shown is the free-energy profile obtained by scaling $\omega^{(2)}(r)$ from IT by a factor of 0.5 only for distances $r < 0.6778$ nm (dot-dashed line). At $r = 0.6778$ nm, $\omega^{(2)}(r)$ crosses the zero-axis just beyond the overlap limit of $r = 0.66$ nm. This partial scaling introduces a kink into $\omega^{(2)}(r)$. The conformational free energies (or PMFs) are shown in units of kcal/mol as a function of the length d (in nm) of the edge of the cubic cluster. The total free-energy profiles can be obtained by adding the direct methane–methane interaction energy to the hydration free-energy profiles shown here. The error bar indicates the variations in the barrier position in the simulation data between different runs.²¹

simulation data can thus be considered excellent when these uncertainties in the reference simulation data are taken into account. In particular, the height of the kinetically important desolvation barrier at about $r = 0.8$ nm agrees exactly with the simulation result by Rank and Baker.²¹

Concluding Remarks

We have developed a hydrophobic force field as a molecular alternative to phenomenological surface-area models. This force field takes into account the many-body contributions arising from overlapping cavity volumes. The importance of a careful treatment of overlapping solvent-exclusion volumes has been realized before, leading to the “hydration-shell model” of solvation thermodynamics^{24–26} and related approaches.⁸⁰ The theoretical basis of the hydrophobic force field is a modified PMF expansion which is exact for hard solutes in both limits of perfect overlap and infinite separation even in the absence of interaction terms $\omega^{(k)}$. This modified PMF expansion thus reduces the non-additivity problems identified for hydrophobic clusters which led to the development of so-called proximity approximations for hydrophobic solvation.^{46,66,67,70} Alternatively, Pellegrini et al.^{87–89} estimate conformational free energies by using a truncated PMF expansion^{90,91} of the inhomogeneous solvent density. However, a pair-level truncation (i.e., Kirkwood superposition approximation) in the solvent-density expansion leads to small errors even in methane-pair PMFs. These errors are expected to be amplified for larger hydrophobic assemblies for which solvent densities at contact have been shown to be over-estimated by the Kirkwood superposition approximation.⁶⁶

(87) Pellegrini, M.; Doniach, S. *J. Chem. Phys.* **1995**, *103*, 2696.

(88) Pellegrini, M.; Grønbech-Jensen, N.; Doniach, S. *J. Chem. Phys.* **1996**, *104*, 8639.

(89) Pellegrini, M.; Grønbech-Jensen, N.; Doniach, S. *Physica A* **1997**, *239*, 244.

(90) Hummer, G.; García, A. E.; Soumpasis, D. M. *Biophys. J.* **1995**, *68*, 1639.

(91) (a) García, A. E.; Hummer, G.; Soumpasis, D. M. *Proteins: Struct., Funct., Genet.* **1997**, *27*, 471. (b) Soumpasis, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5116.

Two approximations make the hydrophobic force field developed here a computationally efficient method for free energy calculations: (1) We find that the hydration free energies of cavity volumes corresponding to intersections of spheres of equal radii can be approximated accurately as a function of the volume alone, irrespective of the shape; (2) the expansion in many-body interactions $\omega^{(k)}$ can be truncated at the pair level. The parameters of the hydrophobic force field can be calculated directly from the IT model of hydrophobic interactions.⁵⁸ The pair interaction $\omega^{(2)}(r)$ between methane-like solutes calculated from IT is scaled by a factor 0.5 to account for additional many-body contributions, but more refined approaches have been discussed.

With these approximations, we find that the hydrophobic force field reproduces simulation data for alkane conformational equilibria.⁶⁷ A more challenging test is the formation of a mini-hydrophobic core by bringing 14 methane molecules together in a close-packed configuration.²¹ The free-energy profile of cluster formation exhibits barriers of almost 10 kcal/mol and free energy gains from the desolvation barrier to the contact minimum (near $d = 0.55$ nm) of about 15 kcal/mol. The hydrophobic force field also reproduces this highly collective aggregate formation, where a conventional two-body PMF expansion and surface-area models fail dramatically.²¹ We believe that the hydrophobic force field developed here provides a molecular alternative to commonly used surface-area models based on its accuracy, computational efficiency, and formal simplicity.

The IT model can be used to parametrize the hydrophobic force field for different temperatures⁶³ and pressures with potential applications to the phase behavior of hydrophobically-stabilized aggregates.⁶⁴ Moreover, the statistical-mechanical foundation of the modified PMF expansion makes this approach well-suited to include also attractive interactions. We have briefly discussed a perturbative approach which will be explored in future studies. In principle, a complete force-field representation of hydration free energies can be derived on the basis of the modified PMF expansion. At least for solutes with weak attractions to water, the inclusion of the many-body cavity effects should result in a considerable increase in accuracy.

Applications include ligand binding and macromolecular complex formation, as well as protein folding where many-body interactions are expected to be important.⁹² However, further analysis and possibly optimization is necessary to achieve the accuracy required for applications such as drug binding.

Acknowledgment. The author is indebted to Dr. Henry S. Ashbaugh for providing the simulation data for alkane conformational equilibria in water prior to publication. The author also wants to thank Dr. Lawrence R. Pratt, Dr. Angel E. García, Dr. Shekhar Garde, Professor Michael E. Paulaitis, and Dr. Henry S. Ashbaugh for numerous discussions on hydrophobic effects and their helpful comments on the manuscript. This work was supported by the U.S. Department of Energy through a Los Alamos National Laboratory LDRD-ER grant.

Appendix

All IT calculations⁵⁸ reported here are performed for the SPC model of water⁹³ at a temperature of 298 K and a number density of 33.33 nm^{-3} . A flat default model was used in the IT calculations.^{58,61} For the dependence of the hydration free energy on the volume of intersection, eq 14, we fitted the two-sphere intersection data to a function

$$\frac{\bar{\mu}(V)}{k_B T} = \frac{a_1 V + a_2 V^2 + a_3 V^n}{1 + b_1 V + b_2 V^2} \quad \text{for } V \leq 4\pi d^3/3 \quad (15)$$

for $d = 0.33$ nm. With $\bar{\mu}$ in units of kcal/mol, V in units of nm^3 , and $k_B T \approx 0.59219$ kcal/mol, we find $a_1 = 32.0966142769287$, $a_2 = 467.831641477464$, $a_3 = 27068.1637856813$, $b_1 = -6.31717234308783$, $b_2 = 166.339344803134$, $n = 3.68238239111433$. This fit of $\bar{\mu}(V)$ should be accurate for cavity spheres also of different sizes than those studied here (exclusion radius 0.33 nm).

JA984414S

(92) Sorenson, J. M.; Head-Gordon, T. *Folding Des.* **1998**, *3*, 523.

(93) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; pp 331–342.