Origins of the Solvent Chain-Length Dependence of Gibbs Free Energies of Transfer

Marcus G. Martin,[†] Nikolaj D. Zhuravlev,[†] Bin Chen,[†] Peter W. Carr,[†] and J. Ilja Siepmann*[,†,‡]

Department of Chemistry and Department of Chemical Engineering and Materials Science, University of Minnesota, 207 Pleasant Street Southeast, Minneapolis, Minnesota 55455-0431

Received: December 2, 1998; In Final Form: February 2, 1999

Experimentally measured partition coefficients show that the solubilities of small solutes in normal alkanes depend on the solvent chain length (n_C). The causes for this n_C dependence have not yet been unambiguously determined, and there is considerable controversy as to whether different interactions with methyl and methylene groups or entropic Flory−Huggins-like effects might play the major role. We have performed Gibbs-ensemble Monte Carlo simulations to study the vapor−liquid partitioning of methane in normal alkanes (with 6−12 carbon atoms) and related model solvents. The simulations show that the increase in solvent density with increasing n_C is the main origin of the n_C dependence for normal alkanes; that is, the solute molecule feels a different environment depending on the alkane chain length.

1. Introduction

Oil−water, oil−vapor, and water−vapor partitioning experiments are used extensively to study the hydrophobic effect,[1−4] and the measured transfer free energies have been very instrumental in modeling biomolecular processes, particularly protein folding and ligand binding.[5,6,7] However, the accuracy of predictions obtained from biomolecular modeling exercises depends on the assumed similarity of the oil phase and the biomolecular phase, e.g., the protein interior. Following this rationale, we may conjecture that a solute molecule should feel essentially the same environment whether the oil phase is hexane or hexadecane.[8] In contrast to this expectation, partitioning experiments have established a chain-length (n_C) dependence of the solubilities of small solutes in alkanes.[9,10]

There has been considerable controversy about the origins of this n_C dependence. Ben-Naim[11,12] has suggested that this trend is caused by different interactions of the solute with methyl and methylene groups. In contrast, Sharp et al.[13,14] and Dill and co-workers[8,15] have argued that the n_C dependence can be rationalized by using Flory−Huggins type arguments to account for the complexities of the oligomeric solvent molecules. While Sharp et al. advocate that free volume effects (volume entropy) upon solvation cause the n_C dependence, Dill and co-workers argue for steric interferences among chain fluids to be responsible.

The thermodynamic definition of the standard Gibbs free energy of transfer from phase 1 to phase 2 is[12,16]

$$\Delta G^* = -RT \ln(\rho_{a,2}/\rho_{a,1}) \quad (1)$$

where the partition coefficient is equal to the ratio of the equilibrium number densities (molarities) of the solute *a* in phase 2 and 1. The free energy of transfer can be viewed as a measure of the difference in affinities of the solute to the two phases. Following Flory−Huggins theory,[13,15] an additional term has been suggested for the right-hand side of eq 1 that is supposed to account for free energies resulting from changes in the entropy of the chain solvent upon solute insertion. In the special case of a solute *a* partitioning between its own pure vapor in phase 1 and chain solvent *b* in phase 2, the modified Gibbs free energy of transfer (environment swap energy, ESE) becomes[15]

Sharp et al.

$$\Delta G^*_{ESE} = -RT\left[\ln\left(\frac{\rho_{a,2}}{\rho_{a,1}}\right) - \frac{N_a + N_b}{N_a + (V_b/V_a)N_b}\right] \quad (2.a)$$

Chan and Dill

$$\Delta G^*_{ESE} = -RT\left[\ln\left(\frac{\rho_{a,2}}{\rho_{a,1}}\right) - \left(\frac{N_a + N_b}{N_a + (m_b/m_a)N_b} - m_a\right)\right] \quad (2.b)$$

where *N*, *V*, and *m* denote the number of particles, the molar volumes, and the length of the particles in a Flory−Huggins lattice approximation. Chan and Dill have already pointed out that the correction term obtained from the Flory-Huggins treatment might not be quantitatively accurate and that cyclohexane might be a better solvent to determine ESEs.[8] Ben-Naim and Mazo[17] and Holtzer[18] have argued against these correction terms on thermodynamic grounds; that is, ΔG^*_{ESE} cannot be used to describe phase or chemical equilibria. Sharp and co-workers have in turn criticized the thermodynamic arguments.[14] The motivation for our work is to assess the applicability of the Flory−Huggins correction terms and to determine the origins for the n_C dependence of the solubilities of small solutes in alkanes from molecular simulations of model systems.

2. Simulation Details

We have used four different models for the oligomeric solvent: (i) model AA, the OPLS all-atom alkane force field;[19] (ii) model UA, the TraPPE united-atom alkane force field;[20] (iii) model UM, united-atom chains consisting solely of TraPPE methylene groups; and (iv) model FJ, chains consisting of freely jointed Lennard-Jones beads. Comparison of models AA and UA allows us to test whether the more articulated structure of the all-atom representation leads to a different n_C dependence for the solubility. Model UM allows us to investigate whether the n_C dependence is caused by the different polarizabilities of methyl and methylene groups. Finally, model FJ has been designed to be an oligomeric solvent for which the segment

* Corresponding author. E-mail: siepmann@chem.umn.edu.
† Department of Chemistry.
‡ Department of Chemical Engineering and Materials Science.

10.1021/jp984583z CCC: $18.00
Published on Web 03/26/1999

density does not depend strongly on n_C; that is, while model FJ retains the conformational characteristics of a chain solvent, it resembles closely a sea of monomers.

For the methane solutes, either the OPLS-AA force field (for model AA) or the TraPPE force field (for models UA, UM, and FJ) is used. Accurate bond-bending and torsional potentials are employed for models AA, UA, and UM. Such bonded potentials were not used for model FJ, resulting in more flexibility for the FJ chains. The Lennard-Jones parameters for model FJ were derived by fitting the vapor−liquid phase diagram of an eight-segment FJ chain to that of *n*-octane. Using the constraint that the FJ bond length is equal to the minimum-energy separation of Lennard-Jonesium ($2^{1/6}\sigma$), we obtain $\epsilon/k_B = 236$ K and $\sigma = 2.91$ Å. Lorentz−Berthelot combining rules are used for all unlike interactions.

Simulations in the Gibbs ensemble[21] offer a very elegant and accurate approach to the direct determination of Gibbs free energies of transfer via the corresponding number densities.[22,23] The Gibbs ensemble utilizes two separate simulation boxes which are in thermodynamic contact but do not have an explicit interface.[21] In the isobaric version of the Gibbs ensemble, mechanical equilibrium for each phase is established by volume exchanges with an external pressure bath. Configurational-bias Monte Carlo swaps[24,25] of molecules between the two phases are used to equalize the chemical potential of each species. Thermal equilibrium in each phase is reached via translational, rotational, and configurational-bias conformational moves. As a result, for a given state point the properties of the coexisting phases, such as the partitioning of solute molecules, can be determined directly from a single simulation.

Simulations were performed for systems of 256 alkanes (*n*-hexane, *n*-octane, *n*-decane, or n-dodecane) and 64 methane molecules (128 and 64 molecules for model AA). The resulting dimensions of the liquid-phase simulation boxes range from 37 Å (*n*-hexane) to 45 Å (*n*-dodecane)[26] and the dimension of all vapor-phase boxes is close to 140 Å. The stable phase of the six-mer for model UM is the vapor phase, and simulations were therefore not carried out for this case. A spherical potential truncation at 14 Å (9 Å for model AA) was employed for the Lennard-Jones interactions. Coulombic interactions for model AA were truncated by using a carbon-based cutoff at 9 Å for entire methyl and methylene groups.[27] The systems were equilibrated for at least 5×10^4 Monte Carlo cycles (one Monte Carlo cycle consists of N trial moves, where N is the total number of molecules), and production periods consisted of (1 to 2) $\times 10^5$ Monte Carlo cycles. 10^4 Monte Carlo cycles for the decane system take approximately 20 h of CPU time on an Intel Pentium II.

3. Results and Discussion

The calculated Ostwald solubities and Gibbs free energies of transfer for methane ($T = 298.15$ K, $P = 1$ atm) are compared in Figures 1 and 2 to their experimental counterparts.[10] The numerical values of the Ostwald solubilities are also listed in Table 1. It should be noted here that neither model AA nor UA is capable of yielding perfect predictions for the free energies of transfer, but the mean errors between simulation and experiments are small (+0.5 and −0.4 kJ/mol for AA and UA, respectively), and the errors for a given model are similar in magnitude over the entire range of n_C. In agreement with experiment, models AA, UA, and UM all show a strong solvent n_C dependence of the methane solubility. By constructing a thermodynamic cycle (alkane A to vapor to alkane B) we can estimate the transfer free energies among liquid alkanes (see Figure 3). However, it should be emphasized that the four alkane solvents (hexane to dodecane) are fully miscible and form a one-phase system at standard conditions. The differences in the transfer free energies for the solvents *n*-dodecane and *n*-hexane,

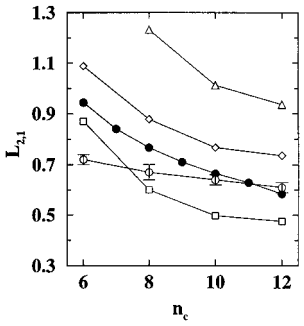

Figure 1. Ostwald solubilities for methane in alkane-like chain solvents as a function of chain length. The experimental data[10] are shown as filled circles. Open squares, diamonds, triangles, and circles denote the calculated solubilities for models AA, UA, UM, and FJ. Standard errors of the mean[29] are shown only for model FJ. The symbols are connected by thin lines for clarity.

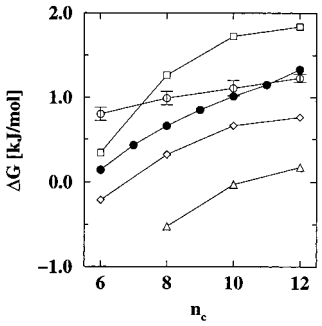

Figure 2. Gibbs free energies of transfer for methane in alkane-like chain solvents as a function of chain length. Symbols as in Figure 1.

TABLE 1: Ostwald Solubilities (Partition Coefficients) and Carbon Number Densities (in nm^{-3}) for Models AA, UA, UM, and FJ and Experiment[10]

solvent	*n*-hexane	*n*-octane	*n*-decane	*n*-dodecane
$L_{2,1}$ (exp)	0.944	0.767	0.665	0.585
$L_{2,1}$ (AA)	0.87_3[a]	0.60_4	0.50_3	0.48_3
$L_{2,1}$ (UA)	1.09_2	0.88_3	0.77_2	0.73_2
$L_{2,1}$ (UM)		1.23_4	1.01_2	0.94_2
$L_{2,1}$ (FJ)	0.72_3	0.68_3	0.65_3	0.61_2
ρ_C (exp)	27.459	29.461	30.746	31.615
ρ_C (AA)	27.66_6	30.00_8	31.40_3	32.25_5
ρ_C (UA)	27.55_2	29.65_2	30.99_2	31.93_2
ρ_C (UM)		26.18_6	28.32_7	29.73_9
ρ_C (FJ)	29.43_3	29.63_2	29.78_4	29.91_2

[a] The subscript gives the statistical uncertainty in the last digit(s).[29]

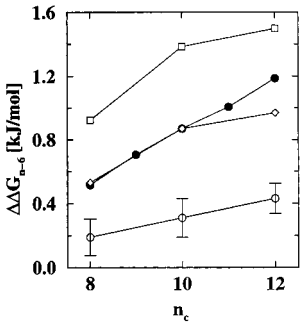

Figure 3. Differences in Gibbs free energies of transfer for methane between solvent hexane and the other *n*-alkane solvents. Symbols as in Figure 1.

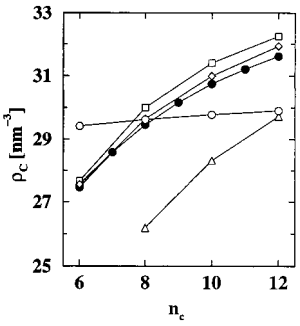

Figure 4. Saturated carbon (segmental) number densities of the alkane solvents as a function of chain length. Standard errors of the mean are much smaller than the symbol sizes. Symbols as in Figure 1.

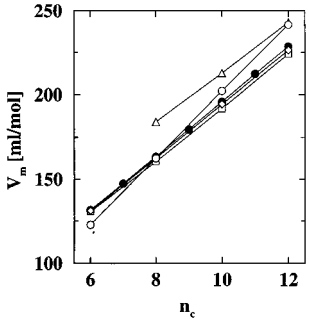

Figure 5. Molar volumes of the alkane solvents as a function of chain length. Standard errors of the mean are much smaller than the symbol sizes. Symbols as in Figure 1.

$\Delta\Delta G^*_{12-6}$, are 1.2, 1.5, and 1.0 kJ/mol for the experiment, AA, and UA, respectively. The fact that the n_C dependence for model UM is similar in magnitude as observed experimentally and for model UA is unambiguous evidence that differences in the interactions of the solute with methyl and methylene segments are not responsible for the n_C dependence. Here it should be noted that model UA employs significantly different well-depths for methylene ($\epsilon_{CH_2}/k_B = 46$ K) and methyl groups ($\epsilon_{CH_3}/k_B = 98$ K).[20] The most striking result, however, is that in contrast to the experimental data and those for the other three models, the solubility in the freely jointed chain solvent (model FJ) shows a much smaller n_C dependence with $\Delta\Delta G^*_{12-6} = 0.4$ kJ/mol.

The saturated carbon number densities and molar volumes of the solvents are shown in Figures 4 and 5 (see also Table 1). While experiment and models AA, UA, and UM are characterized by marked increases in the carbon number densities with increasing n_C (e.g., the experimental carbon number density increases by 14% from hexane to dodecane), the increase for model FJ is approximately 1 order of magnitude smaller (1.6% from hexane to dodecane). In contrast, as should be expected for all oligomeric liquids, the molar volumes increase almost linearly with n_C. Note that model FJ exhibits similar incremental molar volumes for both end and middle segments ($\Delta V_{end} = 22.0$ mL/mol and $\Delta V_{mid} = 19.8$ mL/mol), whereas models AA, UA, and experiment give values around $\Delta V_{end} = 33.6$ mL/mol and $\Delta V_{mid} = 15.9$ mL/mol. Nevertheless the total molar volumes are to within a few percent equal for all models. Thus if the change in molar volume (or length) of the solvent (V_b or m_b in eq 2) would be responsible for the n_C dependence as inferred from Flory−Huggins theory, then all four model solvents should show very similar n_C dependencies. This is clearly not the case.

For a more quantitative comparison, eq 2.a has been used to calculate ΔG^*_{ESE} for every value of n_C as function of V_a for models UA, UM, and FJ (all other variables required for eq 2.a

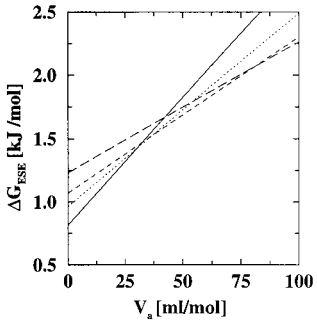

Figure 6. Envrionment swap energies for model FJ as a function of V_a (see eq 2.a). The solid, dotted, dashed, and long-dashed lines show the results for the 6-mer, 8-mer, 10-mer, and 12-mer solvents, respectively.

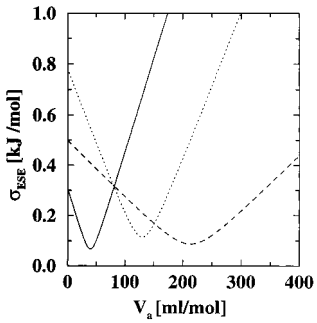

Figure 7. Standard deviations in the environment swap energies (ΔG^*_{ESE}) as a function of V_a (see eq 2.a). The solid, dotted, and dashed lines show the results for model FJ, UA, and UM, respectively.

are obtained directly from the simulation averages). Figure 6 shows the results for model FJ. A chain-length independent ΔG^*_{ESE} would require the curves for the four chain lengths to intersect in a common point. However, curve intersections are found over a V_a range from 30 to 80 mL/mol. Similar spreads of the intersection points are also observed for models UA and UM. To determine the V^*_a values that are most consistent with a constant ΔG^*_{ESE}, the standard deviations in ΔG^*_{ESE} are plotted as functions of V_a in Figure 7. The values for V^*_a range from 40 over 130 to 210 mL/mol for models FJ, UA, and UM, respectively. Considering that the same methane parameters are used in the simulations for these three models, it should be expected that the three models would yield a common V^*_a, too. However, one might argue that V^*_a should not only depend on the solute parameters but also on the solute−solvent interactions. Nevertheless, values of 130 and 210 mL/mol that are similar to the molar volumes of *n*-hexane and *n*-undecane appear to be rather large for a methane solute. Chan and Dill[8] have also reported a too large V^*_a value for the xenon solubility in *n*-alkanes. On the other hand, a V^*_a value of 40 mL/mol as found for model FJ is consistent with the incremental molar volume of a methyl group (≈34 mL/mol, see above), i.e., one should expect a methane solute to be slightly larger than the methyl group increment. Thus, it can be argued that the n_C dependence of the transfer free energies for model FJ, which exhibits only small density variations with increasing n_C, can be described by Flory−Huggins entropy corrections. Whereas comparison of Figures 2 and 4 point to the segmental number density (or specific density) of the alkane solvent as the most important contribution to the n_C dependence for models AA, UA, UM, and experiment, an increase in solvent density with increasing n_C leads to an increase in the free energy of transfer. The seemingly small differences in the alkane densities are able to cause a pronounced effect on the chemical affinity between

solute and solvent. Thus, we would like to argue that an environmental swap energy[8] that is independent of the solvent chain length *does not* exist for medium-length alkanes because for a given absolute temperature and pressure the solvent density *does* depend on n_C.

Finally, it is worthwhile to address the relative importance of attractive and repulsive forces on the solvent n_C dependence of the solubilities. An increase in the solvent density will lead to increases in both the attractive forces (cohesive energy density) and repulsive forces (cost of cavity formation). While an increase of the attractive forces will favor partitioning into the liquid phase (increase solubility), an increase of the repulsive forces will favor partitioning into the vapor phase (decrease solubility). As is evident from Figures 1 and 4, an increase in solvent density leads to a decrease of the methane solubility for both experiment and simulation. Thus, we may conclude that the contributions from the repulsive forces are the dominant factor for the observed solvent n_C dependence. This is also in good agreement with results obtained from scaled particle theory that point to the high cost of creating a cavity in a more dense (or less compressible) solvent.[10,28]

4. Conclusions

In conclusion, using Gibbs ensemble Monte Carlo simulations to calculate the partitioning of a methane solute into medium-length alkane model solvents (with 6−12 carbons), it is demonstrated that the increase in the segmental densities (or specific densities) of the solvents is the dominating factor for the observed chain-length dependence of Ostwald solubilities and transfer free energies. Different interactions with methyl and methylene segments appear to be of far lesser importance. Flory−Huggins entropy corrections fail to account for the qualitative differences observed for the solubilities in alkane-like solvents (models AA, UA, and UM) on one side and a freely jointed-bead solvent (model FJ) on the other side. However, if viewed alone, then the chain-length dependence found for model FJ is consistent with the predictions of Flory−Huggins entropy corrections.

Acknowledgment. We would like to thank Peter Rossky and Hue Sun Chan for a critical reading of this manuscript and for helpful discussions. Financial support from the National Science Foundation (Grant CTS-9813601), the Petroleum Research Fund, administered by the American Chemical Society, a Camille and Henry Dreyfus New Faculty Award, a McKnight Land-Grant Assistant Professorship, an Alfred P. Sloan Research Fellowship, a DOE Computational Science Graduate Fellowship (M.G.M.), and a NSF-REU Summer Research Fellowship (N.D.Z.) is gratefully acknowledged. Part of the computer resources were provided by the Minnesota Supercomputing Institute.

References and Notes

(1) Tanford, C. *The Hydrophobic Effect*, 2nd ed.; Wiley: New York, 1980.

(2) Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Lee, R. H.; Zehfus, M. H. *Science* **1985**, *229*, 834.

(3) Dill, K. A. *Science* **1990**, *250*, 297.

(4) Carr, P. W.; Li, J.; Dallas, A. J.; Eikens, D. I.; Tan, L. C. *J. Chromatogr. A* **1993**, *656*, 113.

(5) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199.

(6) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.

(7) Honig, B.; Yang, A.-S. *Adv. Protein Chem.* **1995**, *46*, 27.

(8) Chan, H. S.; Dill, K. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 425.

(9) Pollack, G. L.; Himm, J. F. *J. Chem. Phys.* **1982**, *77*, 3221.

(10) Hesse, P. J.; Battino, R.; Scharlin, P.; Wilhelm, E. *J. Chem. Eng. Data* **1996**, *41*, 195.

(11) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *80*, 4438.

(12) Ben-Naim, A. *Solvation Thermodynamics*; Plenum: New York, 1987.

(13) Sharp, K. A.; Nicholls, A.; Fine, R. F.; Honig, B. *Science* **1991**, *252*, 106. Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. *Biochemistry* **1991**, *30*, 9686.

(14) Kumar, S. K.; Szleifer, I.; Sharp, K. A.; Rossky, P. J.; Friedman, R.; Honig, B. *J. Phys. Chem.* **1995**, *99*, 8382. Sharp, K. A.; Kumar, S. K.; Rossky, P. J.; Friedman, R.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 14166.

(15) De Young, L. R.; Dill, K. A. *J. Phys. Chem.* **1990**, *94*, 801. Chan, H. S.; Dill, K. A. *J. Chem. Phys.* **1994**, *101*, 7007. Krukowski, A. E.; Chan, H. S.; Dill, K. A. *J. Chem. Phys.* **1995**, *103*, 10675.

(16) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792.

(17) Ben-Naim, A.; Mazo, R. *J. Phys. Chem.* **1993**, *97*, 10829; *J. Phys. Chem. B* **1997**, *101*, 11221.

(18) Holtzer, A. *Biopolymers* **1992**, *32*, 711; *Biopolymers* **1994**, *34*, 315.

(19) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J, *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(20) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569.

(21) Panagiotopoulos, A. Z. *Mol. Phys.* **1987**, *61*, 813. Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. *J. Mol. Phys.* **1988**, *63*, 527.

(22) Martin, M. G.; Siepmann, J. I. *J. Am. Chem. Soc.* **1997**, *119*, 8921.

(23) Martin, M. G.; Siepmann, J. I. *Theor. Chem. Acc.* **1998**, *99*, 347.

(24) Siepmann, J. I. *Mol. Phys.* **1990**, *70*, 1145. Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59.

(25) Mooij, G. C. A. M.; Frenkel, D.; Smit, B. *J. Phys.: Condens. Matt.* **1992**, *4*, L255.

(26) The dimensions of the liquid-phase simulation boxes for the OPLS-AA model are 20% ($1-0.5^{1/3}$) smaller. These system sizes are sufficient to obtain accurate solubilities and Gibbs free energies from Gibbs ensemble simulations.[23]

(27) Coulombic interactions do not play an important role for the fluid phases of alkanes (as is evident from the low dielectric constant of alkanes). In particular, we have recently reported that the vapor−liquid coexistence curves of *n*-pentane using the OPLS-AA model with and without partial charges are indistinguishable [Chen, B.; Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2578]. Thus group-based spherical truncations are adequate for simulations of the OPLS-AA model.

(28) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. *J. Chem. Phys.* **1959**, *31*, 369.

(29) Standard errors of the mean were calculated from blocks of 2.5×10^4 MC cycles. A systematic analysis of the statistical errors showed that this block length is sufficient to yield statistically independent samples.