Estimation of Hydration Free Energy for Polyethylene Glycol Oligomers and the Relative Contributions of Hydrophobic and Short-Ranged Chemical Interactions

S. Utiramerur, M.E. Paulaitis*
Department of Chemical & Biomolecular Engineering,
The Ohio State University, Columbus, OH

September 22, 2008

1 Introduction

Polyethylene glycol (PEG) is a non-ionic water soluble polymer that is used as an additive in protein solutions to induce protein crystallization. PEG in dilute solution concentrations and within a molecular weight range, induces an attractive protein-protein interaction. However the strength of this attractive protein-protein interaction weakens with increasing PEG concentration and molecular weights. Understanding the solvation thermodynamics of PEG is crucial to understanding the molecular mechanisms for the induced protein-protein interaction and its dependence on PEG molecular weights. We calculate the hydration free energy of 1,2-Dimethoxyethane - the shortest oligomer of Polyethylene glycol, using an approach similar to Quasi-chemical theory[1, 2]. This approach enables us to calculate the hydration free energy directly from the simulation data and overcome the limitation in van der Waals approach of identifying a suitable reference interaction model. Using this approach we rearrange the estimation of excess chemical potential into a packing contribution of DME molecule in solvent and an interaction contribution due to explicit interaction of DME with solvent. Thus we are able to estimate the relative contributions of hydrophobic, short-ranged chemical interactions identified by quasi-chemical theory and long-ranged interactions to the overall hydration free energy of DME. The hydration free energy of DME was estimated to be 5.14 *kcal/mole*, which is in good agreement with experimental result[3]. Influence of ether conformations on solvation of DME was studied by estimating the excess chemical potential for some of the most probable conformers in solution.

2 Theory

The partial molar Gibbs free energy or chemical potential for a solute with no internal structure (*i.e.*, no intramolecular degrees of freedom) is given by,

$$\beta\mu_i = \ln \rho_i \Lambda_i^3 + \beta\mu_i^{ex} \ , \tag{1}$$

---

*Corresponding author. Address: Department of Chemical & Biomolecular Engineering, The Ohio State University, 125 Koffolt Labs. W.19th Ave, Columbus, OH 43210, U.S.A., Tel.: (614)247-8847

where $\beta^{-1} = kT$ is the thermal energy, ρ_i is the solute number density, and Λ_i is the thermal deBroglie wave length of the solute. The thermodynamic quantity of interest for solvation is the solute excess chemical potential, μ_i^{ex}, which can be derived from simulation data using the inverse form of the potential distribution theorem,

$$e^{\beta\mu_i^{ex}} = \left\langle e^{\beta\varepsilon} \right\rangle \equiv \int p(\varepsilon) e^{\beta\varepsilon} d\varepsilon \ , \tag{2}$$

by calculating $p(\varepsilon)$, the probability density function of solute-solvent interaction or binding energies, ε, for a distinguished solute molecule. In the inverse PDT, $p(\varepsilon)$ is generated with the solute fully coupled to the solvent in the simulation. Calculation of the integral in Eq. 2 is sensitive to the high-energy tail of $p(\varepsilon)$, which is least likely to be well-sampled in a simulation. We account for inadequate sampling by rewriting Eq. 2 in terms of a reference potential for high-energy solute-solvent intermolecular interactions, which we assume is well-characterized. Specifically, we define a hard-core model of the distinguished solute molecule, and write

$$e^{\beta(\mu_i^{ex} - \mu_{HS}^{ex})} = p(n_\lambda = 0) \int p(\varepsilon | n_\lambda = 0) e^{\beta\varepsilon} d\varepsilon \ , \tag{3}$$

where the marginal probability, $p(n_\lambda = 0)$, is the fraction of the total number of configurations for which there are no solvent molecules within a radius λ of the solute, and $p(\varepsilon | n_\lambda = 0)$ is the corresponding conditional probability of solute-solvent interaction energies for those configurations. The excess chemical potential, μ_{HS}^{ex}, is also derived from the simulation data,

$$\mu_{HS}^{ex}(\lambda) = -kT \ln p_0(\lambda) \ , \tag{4}$$

by computing $p_0(\lambda)$, the probability of finding a cavity in the solvent of radius λ. For simple solutes, $p_0(\lambda)$ is obtained straightforwardly from cavity statistics [4, 5], or occupancies probabilities in the context of information theory [6, 1]. An alternative derivation of Eq. 3 defines a reference potential that takes into explicit consideration those solvent molecules chemically associated with the solute within the framework of quasichemical theory. The marginal probability, $p(n_\lambda = 0)$, in this case is defined in terms of chemical equilibrium constants for solute-solvent association in the first hydration shell of the solute [2, 7]. Values of the conditioning radius, λ, that define $\mu_{HS}^{ex}(\lambda)$ are on the order of the length scale of the distinguished solute molecule. Our interest here is in those values that return a gaussian distribution for $p(\varepsilon | n_\lambda = 0)$. If that distribution has a mean $\langle\varepsilon\rangle$ and variance σ^2, then the integral in Eq. 3 can be evaluated to give,

$$\mu_i^{ex} = \mu_{HS}^{ex}(\lambda) + kT \ln p(n_\lambda = 0) + \langle\varepsilon\rangle + /2kT. \tag{5}$$

For a molecular solute with intramolecular conformational degrees of freedom, μ_i^{ex} is obtained by averaging over all possible conformations, denoted as $\mathcal{R}^n \equiv (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ for a solute molecule consisting of n atoms,

$$e^{\beta\mu_i^{ex}} = \int e^{\beta\mu_i^{ex}(\mathcal{R}^n)} p\left(\mathcal{R}^n\right) d\mathcal{R}^n \ , \tag{6}$$

where the probability distribution of solute conformations, $p\left(\mathcal{R}^n\right)$, is generated with the solute and solvent fully coupled, and $\mu_i^{ex}(\mathcal{R}^n)$ for a specific solute conformation, $\mathcal{R}^n$, is given by Eq. 5,

$$\begin{aligned} \mu_i^{ex}(\mathcal{R}^n) = {} & \mu_{HS}^{ex}(\lambda_1, \lambda_2, \ldots | \mathcal{R}^n) + kT \ln p(n_\lambda = 0 | \mathcal{R}^n) \\ & + \langle\varepsilon | \mathcal{R}^n\rangle + \sigma^2\left(\mathcal{R}^n\right)/2kT \ . \end{aligned} \tag{7}$$

Eq. 7 also recognizes that multiple conditioning radii, $\lambda_1, \lambda_2, \ldots$, may be required to define the hard-core model for a molecular solute.

The challenge to evaluating $\mu_i^{ex}(\mathcal{R}^n)$ from simulation data using Eq. 7 is adequate sampling to compute $\mu_{HS}^{ex}(\lambda_1, \lambda_2, \ldots|\mathcal{R}^n)$, or equivalently, the corresponding cavity probability,

$$\mu_{HS}^{ex}(\lambda_1, \ldots, \lambda_n|\mathcal{R}^n) = -kT \ln p_0(\lambda_1, \ldots, \lambda_n|\mathcal{R}^n) \ . \tag{8}$$

For 1,2-dimethoxyethane, we compute this cavity probability as the product of the cavity probability for dimethyl ether, $p_0(\lambda_1, \lambda_2|\tilde{\mathcal{R}}^n)$, and the conditional probability of finding a cavity in water corresponding to 1,2-dimethoxyethane in conformation $\mathcal{R}^n$, given the cavity for dimethyl ether in conformation $\tilde{\mathcal{R}}^n$,

$$p_0(\lambda_1, \lambda_2|\mathcal{R}^n) = p_0(\lambda_1, \lambda_2|\tilde{\mathcal{R}}^n) p_0(\lambda_1, \lambda_2, |\mathcal{R}^n, \tilde{\mathcal{R}}^n) \ , \tag{9}$$

where $p_0(\lambda_1, \lambda_2, |\mathcal{R}^n, \tilde{\mathcal{R}}^n)$ is the conditional probability of finding a cavity corresponding to 1,2-dimethoxyethane with conformation $\mathcal{R}^n$ in water given the cavity corresponding to dimethyl ether with conformation $\tilde{\mathcal{R}}^n$.

3 Results and Discussion

The dependence of μ_i^{ex} for dimethyl ether on the conditioning radii, λ_1 and λ_2, is shown in Figure 1. Individual contributions to μ_i^{ex} (Eq.5) for each λ_1-λ_2 pair are given in Table 1. These results show that μ_i^{ex} becomes more positive as both λ_1 and λ_2 increases and becomes insensitive to conditioning at large values of λ_1 and λ_2. The effect of conditioning on μ_i^{ex} is much more pronounced for λ_2, the conditioning radius for the methyl group compared to λ_1 for the ether oxygen. The effect of conditioning on the net hydration free energy can be better understood by looking at its effect on the individual contributions to hydration free energy like the packing contribution and the mean binding-energy.

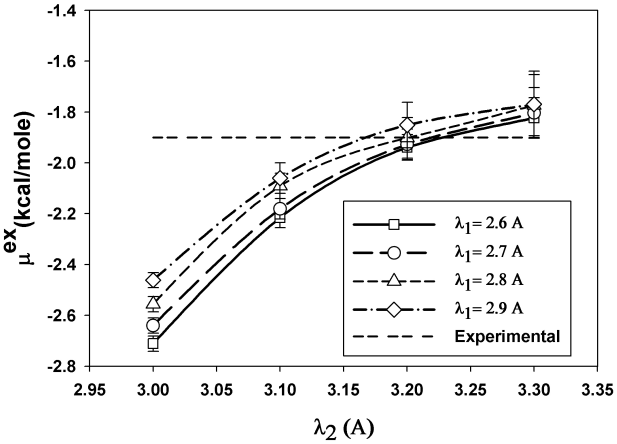

Figure 1: μ_i^{ex} for dimethyl ether (Eq 5) as function of λ_1, the conditioning radius for the ether oxygen, and λ_2, the conditioning radius for the methyl groups. The experimental value of the excess chemical potential [8] is shown as the horizontal line $\cdots$

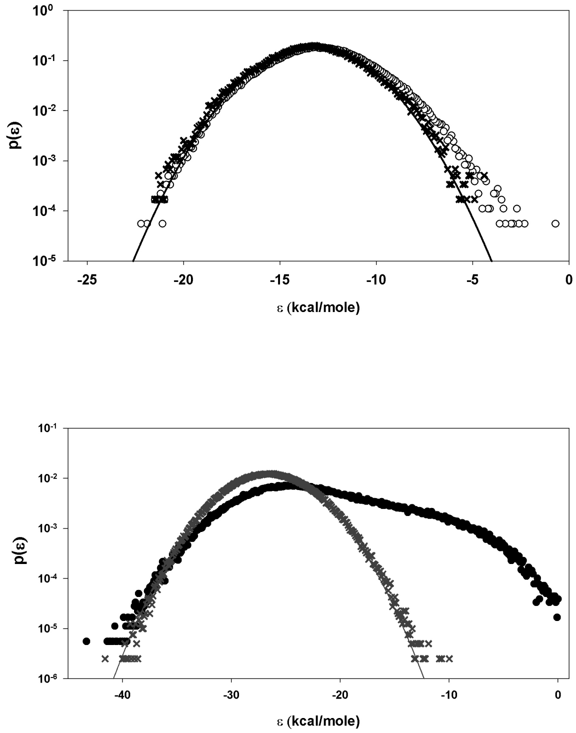

Figure 2: Binding energy distributions for united atom models of dimethyl ether and 1,2-dimethoxyethane. Top panel (binding energy distributions for dimethyl ether): open circles represents the unconditioned binding energy distribution for dimethyl ether; cross symbol represents the conditioned binding energy distribution with $\lambda_1 = 2.7$ Å , and $\lambda_2 = 3.2$ Å ; solid line represents the gaussian fit of conditioned binding energies; Bottom panel (binding energy distributions averaged over all conformations of 1,2-dimethoxyethane): Filled circles represents the unconditioned binding energy distribution for 1,2-dimethoxyethane; crosses represents the conditioned binding energy distribution with $\lambda_1 = 2.7$ Å and $\lambda_2 = 3.2$ Å ; solid line represents the gaussian fit of the conditioned binding energies

The conditional mean binding energies $\langle\varepsilon\rangle$ listed in Table 1, becomes more positive (unfavorable) with increasing values of λ_1. This indicates that increasing conditioning about the ether oxygen eliminates low energy (favorable) conformations more than the high energy conformations. This observation is similar to the one made by Shah *et al.* [9] on applying conditioning to water oxygen's. However conditioning on the methyl carbons (λ_2) has much more pronounced effect on the mean binding energies and is in contrast to the effect of conditioning on ether oxygen's (λ_1). The mean binding energies become more negative (favorable) with increasing values of λ_2 and this shows that the conditioning eliminates the unfavorable (high) energy conformations, more than the well-bound (low) energy conformations. This is consistent with the observation made by Asthagiri *et al.* [10] on applying conditioning to hydrophobic solutes like CF_4. This shows that the methyl carbons in dimethyl ether are hydrated like individual hydrophobic methane molecules.

The net effect of conditioning is evident from observing the conditioned and unconditioned binding-energy distributions for dimethyl ether and 1,2-dimethoxyethane shown in Figure 2. The unconditioned binding-

Table 1: Free energy contributions (Eq 5) for dimethyl ether as a function of conditioning radius for ether oxygen - λ_1 and the methyl groups - λ_2. All the free energy contributions are in the units of kcal/mole; the conditioning radius are in units of Å

λ_1	λ_2	$-kT\ln p_0$	$+kT\ln p(n_\lambda = 0)$	$+\langle\varepsilon\rangle$	$+\sigma^2/2kT$	$=\mu^{ex}$
2.6	3.0	7.45	-0.02	-12.92	+2.77	-2.71 ±0.03
2.6	3.1	8.17	-0.07	-13.02	+2.70	-2.21 ±0.04
2.6	3.2	8.98	-0.25	-13.18	+2.51	-1.94 ±0.05
2.6	3.3	9.92	-0.63	-13.34	+2.22	-1.82 ±0.08
2.7	3.0	7.57	-0.08	-12.84	+2.71	-2.64 ±0.03
2.7	3.1	8.24	-0.14	-12.94	+2.65	-2.18 ±0.04
2.7	**3.2**	**9.03**	**-0.32**	**-13.10**	**+2.46**	**-1.92** ±0.06
2.7	3.3	9.95	-0.69	-13.21	+2.13	-1.80 ±0.10
2.8	3.0	7.64	-0.25	-12.58	+2.63	-2.55 ±0.03
2.8	3.1	8.27	-0.30	-12.67	+2.61	-2.09 ±0.05
2.8	3.2	9.25	-0.48	-12.83	+2.15	-1.90 ±0.08
2.8	3.3	10.02	-0.85	-12.99	+2.04	-1.77 ±0.12
2.9	3.0	7.74	-0.51	-12.13	+2.43	-2.46 ±0.03
2.9	3.1	8.46	-0.56	-12.21	+2.25	-2.06 ±0.06
2.9	3.2	9.18	-0.73	-12.37	+2.06	-1.85 ±0.09
2.9	3.3	10.10	-1.09	-12.53	+1.75	-1.77 ±0.13

Table 2: Excess chemical potential (*kcal/mole*) for 1,2-dimethoxyethane in the tgt and tgg conformations in water, calculated with $\lambda_1 = 2.7$Å and $\lambda_2 = 3.2$Å . Values of $\mu^{ex}(\mathcal{R}^n)$ in parenthesis were computed by numerical evaluation of the integral in Eq 3, rather than assuming a gaussian distribution for $p(\varepsilon|n_\lambda = 0)$. Values of $p_{gas}(\mathcal{R}^n)$ were obtained from [11]

Conformation, $\mathcal{R}^n$	$-kT\ln p_0$	$+kT\ln p(n_\lambda = 0)$	$+\langle\varepsilon\rangle$	$+\sigma^2/2kT$	$=\mu^{ex}(\mathcal{R}^n)$	$p_{aq}(\mathcal{R}^n)$	$p_{gas}(\mathcal{R}^n)$
tgt	14.95	-0.51	-28.56	+8.78	-5.34 (-4.65)	0.63	0.25
tgg	13.64	-0.48	-27.76	+8.67	-5.93 (-5.63)	0.28	0.05
other	14.13	-0.46	-26.21	+8.35	-4.19 (-4.08)	0.09	0.70
Boltzmann average of all conformers **-5.14 kcal/mole** (-4.68)							

energy distributions have a non-gaussian high-ϵ tail, which can be attributed to the highly repulsive solute-solvent interactions. The unconditioned binding-energy distributions for 1,2-dimethoxyethane shows a much larger high-ϵ tail compared to that of dimethyl ether and is attributed to the highly repulsive interactions of tg$^+$g$^-$ conformer with water. Conditioning is able to eliminate these unfavorable near-neighbor interactions and as a result the conditioned binding-energy distributions are more gaussian in nature.

The packing contribution (μ^{ex}_{HS}) evaluated using equation 4 increases with conditioning radius pair λ_2-λ_1 and is affected by values of λ_2 more than by the values of λ_1. This shows that the ether oxygen in dimethyl ether is highly shielded by the larger methyl groups on either side. The $kT\ln p(n_\lambda = 0)$ term makes a small but favorable contribution, which would otherwise been neglected in traditional van der Waals approaches.

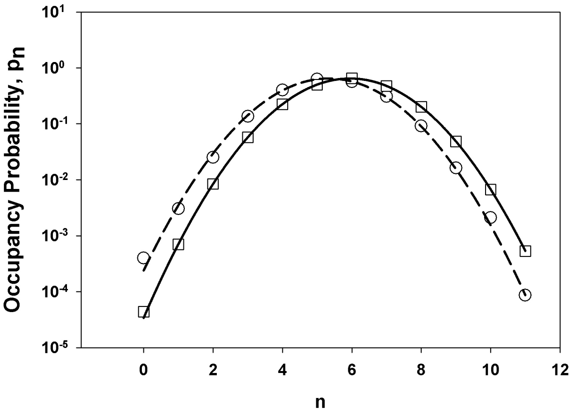

Figure 3: Conditional occupancy probabilities $p_n(\lambda_1, \lambda_2|\tilde{\mathcal{R}}^n, \mathcal{R}^n)$ for the tgg (open circles) and the tgt (open squares) conformations of 1,2-dimethoxyethane in water. The solid line and dashed lines are gaussian fits to the probability distributions for the *tgt* and *tgg* conformations, respectively.

The fluctuation term makes a relatively larger contribution and it decreases with increasing conditioning radius. The net hydration free energy of 1.93 *kcal/mole* evaluated using $\lambda_1 = 2.7$ Å and $\lambda_2 = 3.2$ Å , matches with experimental value[8] of 1.9 *kcal/mole*.

The individual terms in the net hydration free energy of 1,2-dimethoxyethane conformers evaluated using equation 7 are listed in Table 2. The packing contribution $\mu^{ex}_{HS}(\lambda_1,\lambda_2|\mathcal{R}^n)$ was obtained from cavity probabilities using eq. 8. The conditional occupancy probability $p_n(\lambda_1, \lambda_2|\tilde{\mathcal{R}}^n, \mathcal{R}^n)$ for *tgt* and *tgg* conformers of 1,2-dimethoxyethane are shown in Figure 3. The cavity probability for the *tgg* conformer is higher than that of *tgt* conformer by a factor of 10, and this result is expected given the fact that *tgt* is a more extended conformer compared to *tgg*. Overall cavity probability $p_0(\lambda_1, \lambda_2|\mathcal{R}^n)$, for 1,2-dimethoxyethane conformers were calculated using equation 9. The $\mu^{ex}_{HS}(\mathcal{R}^n)$ makes a large and significant contribution to the net free energy of solvation. The μ^{ex}_{HS} contribution for the *tgt* conformer is larger than that for the *tgg* conformer.

The conditioned binding-energy distributions $p(\varepsilon|n_\lambda = 0)$ for the *tgt*, *tgg* and 'other' conformers of 1,2-dimethoxyethane are shown in Figure 4. The individual binding-energy distributions have been scaled down by a factor of 0.1 for purpose of clarity. The strength of ether-water interactions as a function of conformers can be determined from the mean binding energy $\langle\varepsilon\rangle$ for the individual conformers. The results show that the *tgt* conformer has the strongest interaction with water compared to other conformers ($\langle\varepsilon\rangle$ is nearly 1 *kcal/mole* more favorable). The conditioned binding-energy distributions display a small non-gaussian high-ε tail, which could lead to over estimation of hydration free energy by using the gaussian approximation for $p(\varepsilon|n_\lambda = 0)$. The net hydration free energies evaluated using the gaussian approximation (Eq 7) are listed in Table 2 and the net hydration free energies calculated by numerical evaluation of the integral $\int p(\varepsilon|n_\lambda = 0)e^{\beta\varepsilon}d\varepsilon$ are also listed within parenthesis. It is evident from the results, that using the gaussian approximation slightly over-estimates the hydration free energies. The net hydration free energy for the *tgg* conformer (-5.93 *kcal/mole*) is more favorable than that for the *tgt* conformer (-5.34 *kcal/mole*).

The conformer population of 1,2-dimethoxyethane in ideal gas phase [11] and in aqueous solution are listed in Table 2. The *tgt* and *tgg* conformers - characterized by the torsional angles about the C-O-C-C, O-C-C-O and C-C-O-C bonds, are the two most probable conformers in water. The mean binding-energy

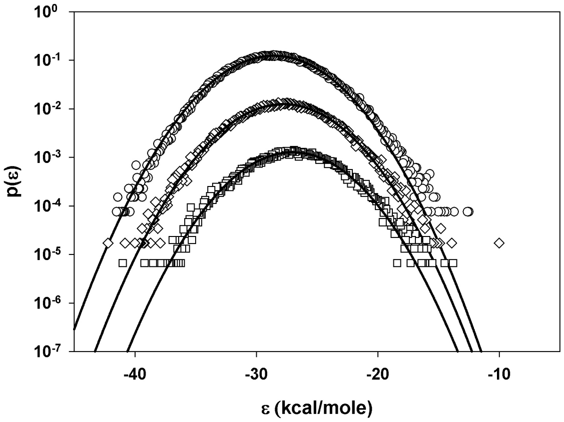

Figure 4: Binding energy distributions for the tgt, tgg and 'other' conformations of 1,2-dimethoxyethane in water. The distributions for tgg and 'other' conformations have been scaled down by factors of 0.1 and 0.01 respectively, for the purpose of clarity. The symbols represent the following, open circles – tgt conformer binding energies, open diamonds – tgg conformer binding energies, open squares – all other conformers binding energy distribution. The solid lines represent Gaussian fit of individual conformers binding energy distribution.

$\langle \varepsilon \rangle$ for *tgt* conformer is nearly 1 $kcal/mole$ more favorable than that for the tgg conformer and it explains the predominance of *tgt* conformer in water (63%). However the net hydration free energy for *tgt* conformer is less favorable than that for the *tgg* conformer. At first the results of net hydration free energies for 1,2-dimethoxyethane conformers and their solution probabilities seem to contradict each other. This apparent contradiction can be explained by looking at the relative increase in conformer population in going from ideal gas phase to water. The *tgt* conformer shows nearly a two-fold increase in population relative to ideal gas phase, whereas the *tgg* conformer shows over a five-fold increase.

The overall hydration free energy of 1,2-dimethoxyethane, calculated as the Boltzmann average (Eq 6) of individual conformer hydration free energies is -5.14 $kcal/mole$ and is in good agreement with the available experimental result [3] of -4.8 $kcal/mole$. Thus the approach of using the Gaussian statistical thermodynamic model provides a reliable method for estimating the hydration free energy for large solutes with conformational degrees of freedom like 1,2-dimethoxyethane.

References

[1] G. Hummer, S. Garde, A.E Garcia, and L.R. Pratt. New perspectives on hydrophobic effects. *Chem Phys*, 258:349–370, 2000.

[2] M.E. Paulaitis and L.R. Pratt. Hydration theory for molecular biophysics. *Adv. Protein Chem*, 62:283–310, 2002.

[3] S. Cabani, P. Gianni, V. Mollica, and L. Lepori. Group contributions to the Thermodynamic properties of non-ionic organic solutes in dilute aqueous-solution. *J Solution Chem*, 10(8):563–595, 1981.

[4] A. Pohorille and L. R. Pratt. Cavities in molecular liquids and the the theory of hydrophobic solubilities. *Journal of Amercian Chemical Society*, 112:5066–5074, 1990.

[5] L. R. Pratt and A. Pohorille. Theory of hydrophobicity - transient cavities in molecular liquids. *PNAS*, 89:2995–2999, 1992.

[6] G. Hummer, S. Garde, A. Garcia, M.E. Paulaitis, and L.R. Pratt. Hydrophobic Effects on a Molecular Scale. *J. Phys. Chem. B*, 102(51):10469–10482, 1998.

[7] T.L. Beck, M.E. Paulaitis, and L.R. Pratt. *The potential distribution theorem and models of molecular solutions*. Cambridge University Press, New York, 2006.

[8] J. Hine and P.K. Mookerjee. Intrinsic hydrophilic character of organic compounds. Correlations in terms of structural contributions. *J. Org. Chem*, 40(3):292–298, 1975.

[9] J.K. Shah, D. Asthagiri, L.R. Pratt, and M.E. Paulaitis. Balancing local order and long-ranged interactions in the molecular theory of liquid water. *J. Chem Phys*, 127(14):1445081–1445087, 2007.

[10] D. Asthagiri, H.S. Ashbaugh, A. Priyatinski, M.E. Paulaitis, and L.R. Pratt. Non-van der Waals Treatment of the Hydrophobic solubilities of CF_4. *J. Am. Chem. Soc*, 129(33):10133–10140, 2007.

[11] H.S. Ashbaugh. *The hydration of Amphiphilic solutes: A theoretical and modeling perspective*. PhD thesis, University of Delaware, 1998.