Predicting the Solubility of Solid Phenanthrene: A Combined Molecular Simulation and Group Contribution Approach

Andrew S. Paluch and Edward J. Maginn
Dept. of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556

DOI 10.1002/aic.14020
Published online March 4, 2013 in Wiley Online Library (wileyonlinelibrary.com)

A simple correction to the infinite dilution activity coefficient computed via molecular simulation for a nonelectrolyte solid solute in solution is proposed. The methodology adopts the concept that the activity coefficient may be fundamentally interpreted as a product of a residual and combinatorial term. The residual contribution is assumed to be insensitive to concentration, and the combinatorial term is modeled using the athermal Flory–Huggins theory. The proposed method uses only properties for the solute computed at infinite dilution to estimate solution-phase properties at finite concentrations. Properties of the pure solid solute are estimated using the group contribution method of Gani and coworkers, allowing for efficient blind solubility predictions to be made. The method is applied to predict the solubility of solid phenanthrene in 17 different solvents. For all cases, the combinatorial correction lowers the predicted solubility relative to the infinite dilution approximation, and in general, improves agreement with experiment. © 2013 American Institute of Chemical Engineers *AIChE J*, 59: 2647–2661, 2013
Keywords: thermodynamics/statistical, thermodynamics/classical, phase equilibrium, computer simulations (MC and MD)

Introduction and Motivation

Knowledge of the solubility behavior of crystalline solids in different solvents is of great practical interest. Of particular interest in this study is the scenario of finding the most suitable solvent to dissolve a particular crystalline solute. This is important, for instance, in the formulation of pharmaceutical compounds for which novel drug candidates often have poor aqueous solubility.[1,2] Chemists and process engineers are confronted with the challenge of determining a suitable solvent or cosolvent system for the production and formulation of the drug. Another important industrial application is the design of extraction processes in which it is desirable to extract only specific components from a heavy petroleum stream.[3,4] In recent years, a tremendous amount of experimental effort has been devoted toward developing state-of-the-art ionic liquid based extraction processes for use with specific systems.[5]

As a result of the vast possibility of solvents and the chemical composition of the solute, along with the time and expense associated with experimental solubility measurements, it is desirable to have methods that can predict the ability of a given solvent to dissolve a particular solute. Because solvation is a complex phenomenon in which many different competing forces interact to determine the behavior of a given solute–solvent system, the development of such methods is extremely challenging. An impressive amount of research has been done in this area, but it is still generally the case that the prediction of solvation behavior relies on empirical and semiempirical correlations, and group contribution (GC) methods for which the underlying molecular level driving forces are not always evident.[1] Many efforts have also been aimed at developing theoretically based models.[1,6–11] Although many of these models hold great promise, they typically require experimental data to allow for their "training." For cases in which the solute is a novel molecule, the necessary amount of experimental data that these approaches require is likely unavailable.

Given this situation, it is desirable to have atomistic-based molecular dynamics (MD) and Monte Carlo (MC) methods that can predict solubility with no experimental data as input. Not only would such methods provide guidance in solvent selection but atomistic models provide insight into the molecular level details governing solubility behavior. Only a limited number of molecular modeling approaches have been used to predict the solubility of solids.[12–18] In some cases, atomistic simulations have been performed to compute parameters for empirical correlations,[12] whereas in others, simulations have been used to directly compute the solubility of solutes in solvents.[13–18] In the latter case, the phase equilibria condition requires equality of the temperature, pressure, and chemical potential of each species between the phases in equilibrium. The solubility may be computed in the complete absence of experimental solubility or thermodynamic data, so long as the solute solid crystal structure is known.[13–15,17,18] Computing the chemical potential in such systems requires precise simulation results and accurate molecular models.[15,17,19] Alternatively, the chemical potential of the solute in the solid-phase may be computed from experimental data.[16]

Additional Supporting Information may be found in the online version of this article.

Correspondence concerning this article should be addressed to E. J. Maginn at ed@nd.edu.

Frequently, the solubility of nonelectrolyte solids of industrial interest is low and is commonly assumed to be infinitely dilute. For this case, our previous work[20] attempted to introduce simplifying assumptions to greatly reduce the computational cost of making solubility predictions. In this manner, solubility predictions may be made in a range of solvents in a rapid and efficient fashion. Our previous work used expressions for the infinite dilution solute activity coefficient from the literature.[21,22] In this study, the working expressions used to make solubility predictions are rigorously derived with a firm thermodynamic foundation, with resulting expressions that differ from our previous work.[20] Moreover, the working expressions naturally allow for the correction of the infinite dilution approximation through use of established solution theories. The proposed methodology is aimed at efficiently predicting the solution-phase behavior of the solute at finite concentrations using only properties computed at infinite dilution. With knowledge of the fugacity of the pure solid solute, the equilibrium solubility may be predicted. The pure solid solute fugacity could be computed directly via molecular simulation.[14,15,17–19] Likewise, this could be computed using experimental data.[16] However, in future studies, it will be desirable to study novel, complex molecular solutes for which experimental data may not be available, and molecular simulations of the solid state would be difficult. While the development of predictive GC methods to predict the behavior of an arbitrary solid solute in a wide range of arbitrary solvents is a challenge, the GC method of Gani and coworkers[23,24] has been shown to accurately predict pure component properties. Therefore, in this study, the pure solid solute fugacity will be estimated using the GC method of Gani and coworkers, allowing for blind solubility predictions. Moreover, for the simulator who is interested in computing partition coefficients using infinite dilution properties, this study introduces a combinatorial correction term to account for the composition dependence of the partition coefficient resulting from the size dissimilarity of the solute and solvents of interest. In this study, the proposed methodology is additionally used to make solubility predictions using the 2005 revised MOSCED model developed by Eckert and coworkers,[11,25] which is used to predict infinite dilution activity coefficients.

Methodology

Thermodynamic approach to solid–liquid equilibrium

The solubility of a solid, nonelectrolyte solute (Component 2) in a single component solvent (Component 1) is described by the classical equations of solid–liquid equilibrium. Assuming that the solvent does not penetrate the crystalline solid solute in equilibrium with the solution-phase, the equilibrium expression is[26]

$$\ln\left[x_2\gamma_2(T,p,x_2)\right]=\ln\frac{f_2^{\mathrm{L}}(T,p,x_2)}{f_2^{0}(T,p)}=\ln\frac{f_2^{\mathrm{S}}(T,p)}{f_2^{0}(T,p)} \tag{1}$$

where x_2 and γ_2 are the equilibrium mole fraction and activity coefficient of Component 2 in solution, respectively, T is the temperature, p is the pressure, and f_2^{L}, f_2^{S}, and f_2^{0} are the fugacity of Component 2 in solution, in a pure solid, and in an arbitrary reference state, respectively. In this study, we adopt a conventional Raoult's Law standard state, which is typically used to describe the phase equilibria of nonelectrolyte solids in solution using classical thermodynamics.[26] The use of a Raoult's Law standard state will allow for the marriage of molecular simulation with solution theory, ultimately simplifying the task at hand. The adoption of a Raoult's Law standard state is analogous to adopting a hypothetical, pure, subcooled liquid solute reference state.

Next, it is desirable to substitute into Eq. 1 quantities directly obtainable via molecular simulation. Using conventional free energy calculations in the isothermal–isobaric (NpT) ensemble, one may compute the residual chemical potential of the solute in solution, μ_2^{res}, defined as[15,27,28]

$$\mu_2^{\mathrm{res}}(T,p,N_1,N_2)=\mu_2(T,p,N_1,N_2)-\mu_2^{\mathrm{ig}}\left(T,\langle V\rangle_{T,p,N_1,N_2-1},N_2\right) \tag{2}$$

where N_1 and N_2 are the number of molecules of Components 1 and 2, respectively, μ_2 is the chemical potential of Component 2 in solution at the same conditions as μ_2^{res}, and μ_2^{ig} is the chemical potential of Component 2 in a hypothetical, noninteracting, ideal gas state at the same T, but at a fixed density. The relevant volume is given by the ensemble average volume of the system in the absence of the solute molecule that is being coupled/decoupled to the system, $\langle V\rangle_{T,p,N_1,N_2-1}$.[15,27,28] In what follows, it is important to emphasize that μ_2^{ig} is at a fixed density and not at p. As an aside, the terms residual chemical potential, excess chemical potential, and Gibbs free energy of solvation are often used interchangeably to describe the free energy change of coupling a single solute molecule to a system in an NpT ensemble.

It follows directly from Eq. 2 and the definition of fugacity[26] that

$$\beta\mu_2^{\mathrm{res}}(T,p,N_1,N_2)=\ln\frac{f_2^{\mathrm{L}}(T,p,x_2)}{f_2^{\mathrm{ig}}\left(T,\langle V\rangle_{T,p,N_1,N_2-1},N_2\right)} \tag{3}$$

where $\beta^{-1}=k_{\mathrm{B}}T$, k_{B} is Boltzmann's constant, and f_2^{ig} is the ideal gas fugacity of Component 2 at the same conditions as μ_2^{ig} in Eq. 2. The ideal gas fugacity of Component 2 is equivalent to the partial pressure of Component 2 in a hypothetical ideal gas state,[26] $f_2^{\mathrm{ig}}=p_2^{\mathrm{ig}}$. Use of the ideal gas equation of state leads to

$$f_2^{\mathrm{ig}}\left(T,\langle V\rangle_{T,p,N_1,N_2-1},N_2\right)=p_2^{\mathrm{ig}}\left(T,\langle V\rangle_{T,p,N_1,N_2-1},N_2\right)=\frac{N_2k_{\mathrm{B}}T}{\langle V\rangle_{T,p,N_1,N_2-1}}=x_2\frac{(N_1+N_2)k_{\mathrm{B}}T}{\langle V\rangle_{T,p,N_1,N_2-1}} \tag{4}$$

where the definition of the mole fraction of Component 2, $x_2=N_2/(N_1+N_2)$ is used in the last term.

Substituting Eq. 4 for f_2^{ig} and Eq. 1 for f_2^{L} into Eq. 3 and rearranging gives the following expression for the activity coefficient of Component 2

$$\ln\gamma_2(T,p,x_2)=\beta\mu_2^{\mathrm{res}}(T,p,N_1,N_2)+\ln\left[\frac{(N_1+N_2)k_{\mathrm{B}}T}{\langle V\rangle_{T,p,N_1,N_2-1}}\right]-\ln f_2^{0}(T,p) \tag{5}$$

where the adoption of a Raoult's Law standard state requires that $\gamma_i\rightarrow 1$ as $x_i\rightarrow 1$, where $i=\{1,2\}$. Equation 5 may be used to predict the solubility of Component 2 through use of Eq. 1. Note that the fugacity terms appearing in Eqs. 1 and 5 are solute-dependent constants, and the other terms are composition-dependent properties of the solution that can be computed from a liquid-phase simulation.

Computing the activity coefficient

Infinite Dilution Approximation. Assume that solute–solute interactions are negligible, and the mole fraction of the solute in solution is sufficiently small so as to be considered infinitely dilute. That is, assume $x_2 \to 0$. This is realized in a simulation by adding a single-solute molecule ($N_2=1$) to a system of pure solvent for which $N_1 \gg N_2$. It follows that Eq. 5 becomes

$$\ln \gamma_2(T,p,x_2) \approx \ln \gamma_2^{\infty}(T,p,0)=\beta\mu_2^{\mathrm{res},\infty}(T,p,N_1,N_2=1) + \ln\left[\frac{(N_1+1)k_BT}{\langle V\rangle_{T,p,N_1}}\right]-\ln f_2^0(T,p) \quad (6)$$

where the superscript ∞ corresponds to properties for an infinitely dilute solute ($x_1 \to 1$ and $x_2 \to 0$). For an infinitely dilute solute, $N_1 \gg N_2$ such that

$$\frac{(N_1+1)k_BT}{\langle V\rangle_{T,p,N_1}} \approx \frac{N_1k_BT}{\langle V\rangle_{T,p,N_1}}=\frac{RT}{v_1(T,p)} \quad (7)$$

where v_1 is the intensive molar volume of the solvent, $R=N_{\mathrm{avo}}k_B$ is the molar gas constant, and N_{avo} is Avogadro's constant. Equation 6 may be rewritten as

$$\ln \gamma_2(T,p,x_2) \approx \ln \gamma_2^{\infty}(T,p,0)=\beta\mu_2^{\mathrm{res},\infty}(T,p,N_1,N_2=1) + \ln\frac{RT}{v_1(T,p)}-\ln f_2^0(T,p) \quad (8)$$

From Eq. 1, the final expression assuming an infinitely dilute solute is

$$\ln x_2=-\beta\mu_2^{\mathrm{res},\infty}(T,p,N_1,N_2=1)-\ln\frac{RT}{v_1(T,p)}+\ln f_2^{\mathrm{S}}(T,p) \quad (9)$$

Note that the last two terms in Eqs. 8 and 9 form a dimensionless group, requiring that f_2^0 and f_2^{S} have the same units as RT/v_1. Throughout this study, f_2^0 and f_2^{S} will have units of Pa. As mentioned previously, f_2^{S} is a solute-dependent constant, allowing relative solubilities (or partition coefficients) to be computed in the complete absence of knowledge of the solid-phase fugacity. Converting the units of concentration from mole fraction to molar density, the traditional expression to compute a partition coefficient via molecular simulation may be obtained.[29,30]

Combinatorial Correction. Following the work of Abrams and Prausnitz in the development of the universal quasichemical (UNIQUAC) excess Gibbs free energy model,[4,26,31] assume that $\ln \gamma_2$ may be fundamentally interpreted as a sum of a residual and combinatorial term. The residual term is the result of intermolecular interactions, whereas the combinatorial term stems from the size dissimilarity of the solute and solvent. The outcome of this subsection is the derivation of an expression that neglects solute–solute interactions involved in the residual term (as in Eq. 8) but accounts for the composition dependence of the combinatorial term using only properties computed at infinite dilution.

In the spirit of Debenedetti and Kumar,[32] $\ln \gamma_2$ may be expressed as a Maclaurin series expansion (MSE)

$$\ln \gamma_2(T,p,x_2) \approx \ln \gamma_2^{\infty}(T,p,0)+\left(\frac{\partial \ln \gamma_2}{\partial x_2}\right)_{T,p}\Bigg|_{x_2\to 0} x_2+\mathcal{O}\left(x_2^2\right) \quad (10)$$

Maintaining only the first term in the expansion, $\ln \gamma_2^{\infty}$, would be analogous to the infinite dilution approximation. In this light, the higher-order terms in the expansion may be conceptually thought of as corrections to the infinite dilution approximation. For the purpose of this study, Eq. 10 introduces a formalism by which Eq. 8 may be improved.

In this study, $\ln \gamma_2^{\infty}$ will be taken from Eq. 8, which includes the residual and combinatorial contributions at infinite dilution. For many of the problems of interest, a large-size dissimilarity exists between the solute and solvent. Therefore, it is reasonable to assume solute–solute interactions, and, hence, the correction from the residual term is negligible compared to the combinatorial correction. The outcome of this assumption is that $\ln \gamma_2^{\infty}$ will be corrected via Eq. 10 using only the combinatorial term.

As a result of the ability to estimate the infinite dilution partial molar volume of the solute without additional simulation effort when computing μ_2^{res} (as will be discussed momentarily) and its simplicity, the athermal Flory–Huggins (FH) theory[26,33,34] will be used to model the combinatorial (comb) contribution of $\ln \gamma_2$

$$\ln \gamma_2^{\mathrm{comb}}(T,p,x_2)=\ln\left(\frac{\phi_2}{x_2}\right)-(r-1)\phi_1 \quad (11)$$

where ϕ_1 and ϕ_2 are the volume fraction of Components 1 and 2, respectively, and r is the ratio of the excluded volume of Component 2 relative to Component 1. In this study, the ratio of the excluded volumes will be estimated as the ratio of the infinite dilution partial molar volume of Component 2 ($\bar{v}_2^{\infty}$) relative to the molar volume of Component 1 (v_1^0), $r=\bar{v}_2^{\infty}/v_1^0$. Furthermore, when computing the volume fractions, the molar volume of Component 2 (v_2^0) will be taken as $\bar{v}_2^{\infty}$. In this fashion, $\phi_1=x_1/(x_1+rx_2)$ and $\phi_2=rx_2/(x_1+rx_2)$. Noting that $x_1=1-x_2$, the corresponding MSE of Eq. 11 is

$$\ln \gamma_2^{\mathrm{comb}}(T,p,x_2) \approx \ln \gamma_2^{\mathrm{comb},\infty}(T,p,0) + \sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}(nr-1)(r-1)^n x_2^n+\mathcal{O}\left(x_2^{N+1}\right) \quad (12)$$

where N is the order of the expansion, and the zeroth-order term is $\ln \gamma_2^{\mathrm{comb},\infty}=\ln(r)-(r-1)$. Taking $\ln \gamma_2^{\infty}$ from Eq. 8 and the correction terms in Eq. 10 from the combinatorial contribution in Eq. 12 gives

$$\ln \gamma_2(T,p,x_2)=\beta\mu_2^{\mathrm{res},\infty}(T,p,N_1,N_2=1)+\ln\frac{RT}{v_1(T,p)} -\ln f_2^0(T,p)+\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}(nr-1)(r-1)^n x_2^n \quad (13)$$

This leads to the final expression for the equilibrium solubility of Component 2

$$\ln x_2=-\beta\mu_2^{\mathrm{res},\infty}(T,p,N_1,N_2=1)-\ln\frac{RT}{v_1(T,p)}+\ln f_2^{\mathrm{S}}(T,p) -\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}(nr-1)(r-1)^n x_2^n \quad (14)$$

where the summation term may be seen as a combinatorial correction to the infinite dilution solubility prediction (and activity coefficient) given by Eq. 9 (and Eq. 8). However, as compared to the infinite dilution expression, Eq. 14 will require a trivial numerical solution to find x_2. Note that alternative combinatorial expressions exist and may be applied to correct the infinite dilution activity coefficient (see, e.g., the

work of Kikic et al.[35]). Also, as mentioned previously, f_2^S is a solute-dependent constant, allowing relative solubilities (or partition coefficients) to be computed in the complete absence of knowledge of the solid-phase fugacity. However, with Eq. 14, the partition coefficient is no longer a constant.

Computing the solid fugacity

Equations 9 and 14 each contain an (unknown) term involving the pure solid solute fugacity. We now discuss how this may be determined so as to obtain predictions of the equilibrium solubility. First, we will consider computing $\ln f_2^S/f_2^0$. This is advantageous, as it has previously been considered numerous times in the literature.[4,6,26,36] Next, we will consider computing $\ln f_2^0$. In general, the vapor pressure of the pure subcooled liquid below the normal melting point is small.[37] Therefore, it is reasonable to assume that the vapor-phase in equilibrium with the pure subcooled liquid may be approximated as an ideal gas. For this case, $\ln f_2^0 = \ln p_2^{0,\mathrm{sat}}$, where $p_2^{0,\mathrm{sat}}$ is the vapor pressure of the pure subcooled liquid. This is advantageous as many efforts have been devoted to correlating and estimating the vapor pressure of pure liquids.[37] Once $\ln f_2^S/f_2^0$ and $\ln f_2^0$ are known, $\ln f_2^S$ is known.

Computing $\ln f_2^S/f_2^0$.
By definition of the fugacity[26]

$$\ln\frac{f_2^{\mathrm{S}}(T,p)}{f_2^0(T,p)} = \frac{1}{RT}\mu_2^{\mathrm{S}}(T,p) - \frac{1}{RT}\mu_2^0(T,p) = \frac{1}{RT}g_2^{\mathrm{S}}(T,p) - \frac{1}{RT}g_2^0(T,p) = -\frac{1}{RT}\Delta g_2^{\mathrm{fus}}(T,p) \quad (15)$$

where μ_2^{S} and μ_2^0 are the chemical potential of Component 2 in the pure solid and subcooled liquid states, respectively, which for a pure species are equivalent to the corresponding molar Gibbs free energies, given by g_2^{S} and g_2^0, respectively, and $\Delta g_2^{\mathrm{fus}}$ is the molar Gibbs free energy change of fusion. The Gibbs–Helmholtz relation may next be applied to give[26]

$$\left(\frac{\partial \Delta g_2^{\mathrm{fus}}(T,p)/RT}{\partial T}\right)_p = -\frac{\Delta h_2^{\mathrm{fus}}(T,p)}{RT^2} \quad (16)$$

where $\Delta h_2^{\mathrm{fus}}$ is the molar enthalpy change of fusion. As commonly done in the literature,[4,6,26,36] assume that the difference in heat capacity between the subcooled liquid and solid ($\Delta c_{\mathrm{p}}^{\mathrm{fus}} = c_{\mathrm{p}}^0 - c_{\mathrm{p}}^{\mathrm{S}}$) is insensitive to temperature between the temperature of interest (T) and the normal melting point temperature (T_{m}). Integrating Eq. 16 from T_{m} to T gives the following expression[4,6,26,36]

$$\ln\frac{f_2^{\mathrm{S}}(T,p)}{f_2^0(T,p)} = \frac{\Delta h_2^{\mathrm{fus}}(T_{\mathrm{m}},p) - \Delta c_{\mathrm{p}}^{\mathrm{fus}} T_{\mathrm{m}}}{R}\left[\frac{T-T_{\mathrm{m}}}{T_{\mathrm{m}}T}\right] + \frac{\Delta c_{\mathrm{p}}^{\mathrm{fus}}}{R}\ln\frac{T}{T_{\mathrm{m}}} \quad (17)$$

Via Eq. 17, $\ln f_2^S/f_2^0$ may be computed with knowledge of T_{m}, $\Delta h_2^{\mathrm{fus}}$ at T_{m}, and $\Delta c_{\mathrm{p}}^{\mathrm{fus}}$. Note that when integrating Eq. 16 from T_{m} to T, we have ignored any enthalpic contributions resulting from transformations of the crystalline structure; this information is likely unavailable for novel molecules of interest. T_{m} and $\Delta h_2^{\mathrm{fus}}$ at T_{m} are properties that may readily be computed by molecular simulation.[38] However, presently, such simulations are challenging as one must first know the crystal structure either from experiment or by some other predictive means, and accurately modeling the solid-phase is challenging using conventional force fields.[15,17,19] In the absence of experimental data, the values may alternatively be estimated using GC methods. In this study, values computed using the GC method of Gani and coworkers[24] will be used and compared to values taken from experiment. Equation 17 offers the flexibility of using any number of approaches to estimate the necessary parameters.

Several methods have been proposed to estimate $\Delta c_{\mathrm{p}}^{\mathrm{fus}}$.[4,6,26,36,39,40] In this study, we will consider five different methods. For the first three cases, the notation of Nordstrom and Rasmuson[6] is adopted

$$\Delta c_{\mathrm{p}}^{\mathrm{fus}} = \zeta\frac{\Delta h_2^{\mathrm{fus}}(T_{\mathrm{m}},p)}{T_{\mathrm{m}}} \quad (18)$$

where ζ is a known constant. For the case of $\zeta=0$, $\Delta c_{\mathrm{p}}^{\mathrm{fus}}=0$, and $\Delta h_2^{\mathrm{fus}}$ is, therefore, assumed to be constant over the temperature range of interest. This would produce the result that $\ln f_2^S/f_2^0$ changes linearly with T^{-1} and is analogous to the Clausius/Clapeyron equation often used to correlate the vapor pressure of liquids over narrow temperature ranges.[37,41] The case of $\zeta=1$ was proposed by Hildebrand et al.[36] using simple thermodynamic relations assuming that $\Delta c_{\mathrm{p}}^{\mathrm{fus}}$ is a constant. The third case considered was proposed by Nordstrom and Rasmuson,[6] where ζ was treated as a universal, empirical constant. At 298.15 K, they report $\zeta=1.897$, which was found by regressing to 115 temperature-dependent solubility curves for a wide range of industrial and pharmaceutical solids.

The last two methods considered are based on the GC method of Chickos and coworkers.[39,40] First, it is assumed that

$$\Delta c_{\mathrm{p}}^{\mathrm{fus}} = c_{\mathrm{p}}^0(T_0,p) - c_{\mathrm{p}}^{\mathrm{S}}(T_0,p) \quad (19)$$

where $\Delta c_{\mathrm{p}}^{\mathrm{fus}}$ is equal to the corresponding heat capacity difference at $T_0=298.15$ K. Chickos and coworkers[39,40] have also proposed an empirical expression to use to compute $\Delta h_2^{\mathrm{fus}}$ at $T_0=298.15$ K when $\Delta h_2^{\mathrm{fus}}$ at T_{m} is known. Assuming that the heat capacity is constant, an effective heat capacity difference may be extracted from their expression

$$\Delta c_{\mathrm{p}}^{\mathrm{fus}} = \Delta\hat{c}_{\mathrm{p}}^{\mathrm{fus}} = 983 + 026c_{\mathrm{p}}^0(T_0,p) - 0.15c_{\mathrm{p}}^{\mathrm{S}}(T_0,p) \quad (20)$$

where 9.83 has units of J/mol K, and the two prefactors are dimensionless. The values of c_{p}^0 and $c_{\mathrm{p}}^{\mathrm{S}}$ are calculated using the GC method of Chickos and coworkers.[39,40]

Computing $\ln f_2^0$. To compute $\ln f_2^{\mathrm{S}}$ via Eq. 17, $\ln f_2^0$ must be known. In general, the vapor pressure of the pure subcooled liquid below the normal melting point is small.[37] Therefore, it is reasonable to assume that the vapor-phase in equilibrium with the pure subcooled liquid may be approximated as an ideal gas. For this case, $\ln f_2^0 = \ln p_2^{0,\mathrm{sat}}$, where $p_2^{0,\mathrm{sat}}$ is the vapor pressure of the pure subcooled liquid. Many efforts have been devoted to correlating and estimating the vapor pressure of pure liquids, and the interested reader is directed to chapter seven of the text "The Properties of Gases and Liquids," by Poling et al.[37] Vapor pressure data obtained from experiment or from molecular simulation above the normal melting point may be extrapolated to the subcooled liquid conditions.[37,42] Although molecular simulations to compute solid–liquid and solid–vapor phase equilibria properties are challenging, simulation capabilities have advanced to efficiently compute vapor–liquid phase equilibria properties of highly complex fluids.[43–46] Alternatively, parameters computed using GC methods may be used. In

this study, we will adopt the GC method of Gani and coworkers[24] to estimate $\ln f_2^S/f_2^0$ as a result of the previously mentioned challenges of modeling the solid-phase. Therefore, we will additionally use the GC method of Gani and coworkers to estimate $\ln f_2^0$. To accomplish this, we will use the Clapeyron equation[41] in conjunction with the Watson equation,[37] which is particularly well suited for use with the GC method of Gani and coworkers to model the vapor pressure between the normal boiling point temperature (T_b) and the temperature of interest (here $T=298.15$ K). Although accurate corresponding state theories exist, their application is commonly limited to a particular class of compounds.[37] Alternatively, the proposed method is free of this constraint.

We begin with the rigorous Clapeyron equation[41]

$$\frac{d\ln p_2^{sat}}{dT}=\frac{\Delta h_2^{vap}}{RT^2\Delta Z_2^{vap}} \tag{21}$$

where p_2^{sat} is the vapor pressure and Δh_2^{vap} is enthalpy of vaporization of Component 2 at saturation, and ΔZ_2^{vap} is the compressibility-factor change of vaporization of Component 2. If the temperature and pressure dependence of Δh_2^{vap} and ΔZ_2^{vap} are known in addition to a single saturation point (i.e., p_2^{sat} at a given T), Eq. 21 may be used to estimate $p_2^{0,sat}$. In this study, Δh_2^{vap} will be estimated using the Watson equation[37,42]

$$\Delta h_2^{vap}=\Delta h_{2,b}^{vap}\left(\frac{T_c-T}{T_c-T_b}\right)^n \tag{22}$$

where T_c is the critical temperature and $\Delta h_{2,b}^{vap}$ is the enthalpy of vaporization of Component 2 at T_b. n is a constant, and a common choice is 0.375 or 0.38.[37,42] In the work of Miller,[42] the Clapeyron equation was integrated numerically using the Watson equation with $n=0.38$ and with ΔZ_2^{vap} modeled with a temperature- and pressure-dependent expression due to Haggenmacher. Miller found that over the (low) pressure range from approximately 0.1 to 2 MPa, the predictions were in excellent agreement with experiment for a wide range of compounds.

For the low pressures of interest in this study (less than 0.1 MPa), a reasonable approximation to simplify the Clapeyron equation is to assume that the vapor-phase is an ideal gas and that the molar volume of the liquid is negligible compared to the molar volume of the vapor. This is equivalent to assuming that $\Delta Z_2^{vap}=1$.[41] Although n in Eq. 22 is commonly taken to be 0.375 or 0.38, with knowledge of Δh_2^{vap} at any T other than T_b, a unique value of n may be computed.

The final expression to estimate $p_2^{0,sat}$ at the temperature of interest (T) is found by integrating Eq. 21 from the normal boiling point to the saturation point at T to give

$$\ln\left(\frac{p_2^{0,sat}}{p_{2,b}^{sat}}\right)=\int_{T_b}^{T}\frac{\Delta h_2^{vap}}{RT^2}dT \tag{23}$$

where by definition, the vapor pressure at T_b is $p_{2,b}^{sat}=0.1\text{MPa}$. Note that with use of Eq. 22, the integral in Eq. 23 may not be evaluated analytically.

Computing solution-phase thermodynamic quantities

An expanded-ensemble (EE)[47–49] procedure was used to compute the residual chemical potential of the solute. The adopted methodology is described in detail in our previous work,[50] so only a brief summary of the essential concepts and modifications from our previous work are provided here. The basic idea behind the EE method is to construct an augmented ensemble as a sum of a reference and M_{Total} subensembles. This series of $M_{Total}+1$ subensembles connects two systems of interest by gradually performing transitions between the two systems. In this work, the systems of interest are a noninteracting solute molecule (Component 2) in an ideal gas reference state at a fixed density at T (governed by the ensemble average volume of the pure solvent at T and p),[15,27,28] and a single fully interacting solute molecule in the solvent at T and p. The free energy difference between these two systems gives the residual chemical potential of the solute. The intermediate subensembles between the noninteracting and fully interacting solute subensembles serve to scale only the intermolecular interaction potential of the solute. A specific subensemble is designated by index m, where the solute intermolecular Lennard-Jones (LJ) and electrostatic interactions are regulated by the subensemble-dependent coupling parameters λ_m^{LJ} and λ_m^{elec}, respectively. These coupling parameters vary such that $0\le\lambda_m^{LJ}\le 1$ and $0\le\lambda_m^{elec}\le 1$.

Within a given subensemble, configurational-phase space is sampled using MD in an NpT ensemble. Periodically, a stochastic transition to an adjacent subensemble is attempted. The transitions are accepted based on an appropriate MC acceptance rule.[50] However, as the free energy difference between subensembles increases, the probability of accepting a transition decreases exponentially. This difficulty is overcome using a biasing scheme[51] that utilizes a combined Wang–Landau (WL)[52–54] and Bennett's acceptance ratio (BAR)[55–58] method.[50]

The EE procedure has advantages over conventional MD methods in that it decreases the configurational correlation time of the system, enhances the sampling of the important regions of configurational-phase space, and requires only a single simulation.[50] If necessary, EE may also be used to enhance conformational sampling of complex solutes.[59] The free energy difference between subensembles was computed using BAR.[50] Moreover, in this study, perturbation theory was used to estimate the difference in volume between subensembles as[60]

$$V_n(T,p,N_1,N_2)-V_o(T,p,N_1,N_2)=\frac{\langle V_n(\mathbf{x})\exp\left(-\beta[U_n(\mathbf{x})-U_o(\mathbf{x})]\right)\rangle_o}{\langle\exp\left(-\beta[U_n(\mathbf{x})-U_o(\mathbf{x})]\right)\rangle_o}-\langle V_o(\mathbf{x})\rangle_o \tag{24}$$

where V_n-V_o is the difference of the ensemble average volume between subensemble n and o, respectively, $V_n(\mathbf{x})$ and $V_o(\mathbf{x})$ correspond to the instantaneous configurational value of the volume in subensemble n and o, respectively, $U_n(\mathbf{x})$ and $U_o(\mathbf{x})$ correspond to the instantaneous configurational value of the potential energy in subensemble n and o, respectively, and the brackets correspond to an ensemble average taken with respect to the probability distribution of subensemble o, as indicated by the subscript. While the difference in volume between subensembles could be computed by taking ensemble averages within each subensemble, use of Eq. 24 decreases the bias of the calculation by computing the difference from the same probability distribution.[61] The difference in Gibbs free energy and volume between the target ($m=M_{Total}$) and reference ($m=0$) subensemble is computed as the sum of the computed (M_{Total}) differences of the property between each subensemble. By finite difference

arguments and standard thermodynamic definitions,[15,27,28] the computed difference between the target and reference subensemble gives the residual chemical potential (μ_2^{res}) and the partial molar volume ($\bar{v}_2$) of the solute.

Summary

In summary, the two expressions used to estimate the equilibrium solubility of a solid, nonelectrolyte solute are Eqs. 9 and 14, which correspond to the infinite dilution approximation and the combinatorial correction, respectively. For both of these cases, f_2^S is required and is calculated via Eqs. 17, 22, and 23, which require the following pure component properties: Δh_2^{fus} at T_m, T_m, Δc_p^{fus}, T_b, $\Delta h_{2,b}^{vap}$, and T_c. Also, rather than assuming n in Eq. 22 is a universal constant, a unique value may be estimated with knowledge of Δh_2^{vap} at a reference temperature T_0 ($\Delta h_{2,0}^{vap}$). In addition, for both cases, $\mu_2^{res,\infty}$ and v_1 are computed via molecular simulation. Lastly, for the combinatorial correction, $\bar{v}_2^{\infty}$ is also computed via molecular simulation.

Simulation Details

All of the EE calculations were performed with a modified version of the MD simulation package M.DynaMix 5.2.[62,63] For simulations involving water, LJ interactions were truncated at a distance of $r_{cut}=1.45$ nm, and for all other cases, LJ interactions were truncated at a distance of $r_{cut}=1.60$ nm. Standard uniform fluid tail corrections were applied to both the energy and the pressure, assuming $g(r)=1$ beyond the cutoff.[64] For interactions between unlike LJ sites, Lorentz-Berthelot[64] combining rules were used. Electrostatic interactions were evaluated with an Ewald summation with tin foil boundary conditions,[64] with real space interactions truncated at r_{cut}. A damping parameter of $\alpha=\pi/r_{cut}$ was used, and the maximum number of reciprocal space lattice vectors was set by $K_{max}=\alpha L$, where L is the box length of the cubic simulation cell.[65] Integration of the equations of motion was performed with the multiple-time step method of Tuckerman et al.[66] in Cartesian coordinates. A short time step of 0.2 fs was used for all intramolecular degrees of freedom and nonbonded interactions within a cutoff of $r_{short}=0.8$ nm, and a time step of 4 fs was used for all other interactions. An Andersen thermostat[67] was used with the collision time for the thermostat set to 0.5 ps for water and 1.0 ps for all other cases. An Andersen–Hoover barostat[67–69] was used with the time constant for the barostat set to 1.5 ps. Modifications to M.DynaMix include implementation of the Andersen thermostat, the "soft-core" potential,[70–72] separate decoupling of LJ and electrostatic solute intermolecular interactions for EE calculations, WL-BAR, modification of the Ewald summation with EE fractional particles,[50] and other minor additions.

With the exception of water, all of the force fields were optimized to reproduce vapor–liquid phase equilibria properties, making their use appropriate for this study. All of the alcohol solvents were modeled with the United-Atom Transferable Potential for Phase Equilibria (TraPPE-UA) force field of Siepmann and coworkers.[73–75] The nine alcohols studied were: methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1-heptanol, 1-octanol, and 2-propanol. The solvents composed of benzene, pyridine, and thiophene were modeled with the TraPPE Explicit-Hydrogen (TraPPE-EH) force field.[45,76,77] The phenanthrene solute was also modeled with TraPPE-EH. Acetone and 2-butanone were modeled with both TraPPE-UA[78] and the model of Kamath, Georgiev, and Potoff (KGP).[79] KGP modified the TraPPE-UA ketone model by optimizing the carbonyl group parameters to reproduce the vapor–liquid phase equilibria properties of a binary acetone–chloroform system. A chloroform solvent was also studied and modeled using the KGP force field. 1,4-Dioxane was modeled with the force field of Yazaydin and Thompson.[80]

Save for 1,4-dioxane, all of the mentioned models contain rigid bond lengths and for TraPPE-EH, also contain rigid bond angles. To avoid the use of constraints during the MD simulations, missing harmonic bond and angle parameters were introduced. For the alcohol and ketone models, the alcohol H—O stretch was taken from AMBER Parm99,[81] and all other harmonic bond parameters were taken from the united-atom Gromacs (ffgmx) force field from Gromacs 4.05.[82,83] For chloroform, benzene, and pyridine, all missing bond stretching, angle bending, and dihedral parameters were taken from the OPLS-AA[84] (ffoplsaa) force field within Gromacs 4.05. For thiophene, all missing bond stretching, angle bending, and dihedral parameters were taken from the general AMBER force field (GAFF).[85,86] For phenanthrene, all intramolecular parameters were taken from GAFF.

Water was modeled with the simple point charge flexible water (SPC/Fw) model.[87] In this model, the nonbonded parameters are taken from the rigid SPC model, and the intramolecular parameters are optimized to reproduce the experimental properties of bulk water at ambient conditions.[87] As compared to the conventional, rigid SPC water model, it has been shown that the introduction of flexibility has a similar effect as the increased partial charges used by the rigid extended SPC (SPC/E) model.[87–89] As an outcome, the static dielectric constant of the SPC/Fw model and the predicted vapor–liquid coexistence properties are in better agreement with experiment as compared to both the SPC and the SPC/E models.[87–89] In addition, the flexibility of the model allows for the model to alter its conformation (and ultimately dipole moment) in response to perturbations to its environment.[89] Overall, the model should be well suited for this study.

To increase the MD time step of the EE simulations, the mass of the studied molecules was redistributed.[90,91] The procedure involved redistributing a small amount of the mass of carbon and oxygen atoms bonded to a hydrogen to slow the corresponding vibrational frequency and to increase the simulation stability. For chloroform, the hydrogen atom mass was increased by 2 Da, and the masses of the three chlorine atoms were decreased to compensate. Moreover, the mass of the heavy sulfur (in thiophene) and chlorine (in chloroform) atoms were scaled by a factor of 1/2. Note that while manipulating the mass of the molecules alters the dynamics of the system, it has no effect on the salvation-free energy.[59,90,91] All of the force field files (which include atomic masses) used in this study are provided in the Supporting Information.

For each solute (phenanthrene)/solvent system studied, five independent simulations were performed. The systems were set up by randomly inserting a gas phase minimized phenanthrene molecule into five previously equilibrated pure solvent boxes for each of the systems studied. The number of solvent molecules was chosen to give a cubic box length of ~3.2 nm for water and 4.0 nm for all other solvents.[92] The velocities of each of the systems were initialized from a Maxwell–Boltzmann distribution with a unique seed to the

random number generator. Production runs were carried out in an EE-*NpT* ensemble at 298.15 K (25°C) and 0.1 MPa. All production runs were 24 ns in duration. Each of the five independent simulations for each system was initialized with a unique random seed for the random number generator used by the thermostat and for the MC random walk in subensembles.

All of the systems began in the reference subensemble. In the reference subensemble (i.e., ideal gas state), $m=0$, the solute is noninteracting with the rest of the system ($\lambda_0^{\mathrm{LJ}}=0$ and $\lambda_0^{\mathrm{elec}}=0$). The solute was taken from the reference subensemble to the target subensemble, $m=\mathrm{M_{Total}}$, with a fully interacting solute molecule $\left(\lambda_{M_{\mathrm{Total}}}^{\mathrm{LJ}}=1,\lambda_{M_{\mathrm{Total}}}^{\mathrm{elec}}=1\right)$ by first bringing the intermolecular LJ interactions to full strength and then adding in the intermolecular electrostatic interactions. The addition of the LJ and electrostatic interactions were performed separately in $M_{\mathrm{LJ}}=15$ and $M_{\mathrm{elec}}=4$ steps, respectively, for a total of $M_{\mathrm{Total}}=19$ steps. For the M_{LJ} steps, the intermolecular electrostatic interactions were turned off, and the intermolecular LJ interactions were strengthened as $\lambda_m^{\mathrm{LJ}}=\{0.05,0.10,0.20,0.30,0.40,0.50,0.60,0.65,0.70,0.75,0.80,0.85,0.90,0.95,\text{and }1.0\}$ over the range $1\leq m\leq 15$ using a modified "soft-core" potential.[70–72] Next, the intermolecular electrostatic interactions were strengthened as $\lambda_m^{\mathrm{elec}}=\{0.25,0.50,0.75,\text{and }1.0\}$ over the range $16\leq m\leq 19$. The intermolecular subensembles were chosen to agree with the previous work of Mobley et al.[93] Details regarding the coupling/decoupling of intermolecular interactions with both Ewald summation and the "soft-core" potential may be found in our previous work.[50]

Attempts to change subensembles were made every 24 fs. Over the first 0.5 ns, the MC random walk was carried out with WL biasing, in which the WL weight factor was initially taken to be $\upsilon_{\mathrm{WL}}=0.5$ and reduced as $\upsilon_{\mathrm{WL}}^{\mathrm{new}}=0.25\upsilon_{\mathrm{WL}}^{\mathrm{old}}$ every 0.1 ns. During the entire course of the simulation, transition energies (in both directions) were computed each time a transition between subensembles was attempted/proposed, and new subensemble weights were computed from BAR every 0.5 ns.[50] Additionally, each time a transition was attempted/proposed, the potential energies and volumes necessary for Eq. 24 were saved for postsimulation analysis. All of the reported thermodynamic quantities from the EE simulations are taken as the average value computed in the five independent simulations, and the uncertainty is taken as one standard deviation.

Results and Discussion

Accuracy of the combinatorial correction

It is informative to begin by evaluating the performance of the proposed combinatorial correction to the infinite dilution activity coefficient. First, in Figure 1, the analytic FH theory (Eq. 11) is compared to the corresponding MSE (Eq. 12). Note that for a given value of r, the infinite dilution term in the MSE is a constant and simply shifts vertically the resulting curve of ln γ_2 vs. x_2. Therefore, the results will shed insight into the range of applicability of the combinatorial correction when used with molecular simulation. In the top pane of Figure 1, the case of $r=2$ is considered. Using an 11th-order expansion, the MSE is in excellent agreement with the analytic FH theory up to $x_2=0.6$. As compared to the infinite dilution approximation at $x_2=0.6$, the 11th-order

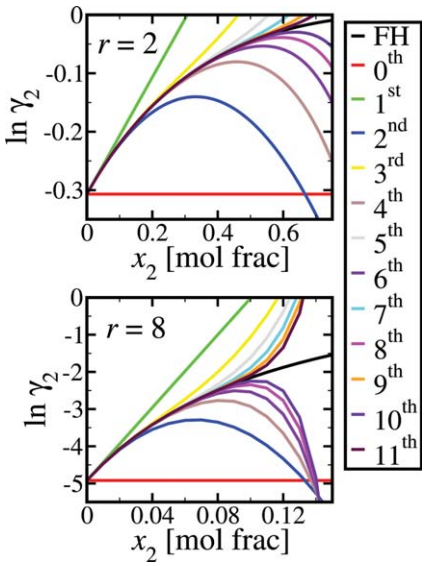

Figure 1. A graphical comparison of ln γ_2 given by the analytic athermal FH theory (Eq. 11) and the corresponding MSE to *N*th order (Eq. 12).

Note that the infinite dilution approximation ($\ln\gamma_2^{\infty}$) corresponds to a zeroth-order expansion. The top pane considers the case of $r=2$, and the bottom pane considers the case of $r=8$. [Color figure can be viewed in the online issue, which is available at wileyonline library.com.]

expansion predicts that ln γ_2 is greater by ~0.3 ln units. In the bottom pane, the case of $r=8$ is considered, and it is shown that an 11th-order MSE is in excellent agreement with the analytic FH theory up to $x_2=0.1$. For this case, at $x_2=0.1$, the infinite dilution approximation and the 11th-order expansion differ by ~2.5 ln units, an appreciable deviation. It is impressive that the MSE is applicable to such a degree. Additionally, from Figure 1, it is advisable to use an odd order expansion. For even order expansions, ln γ_2 predicted using a MSE goes through a maximum, which would cause uniqueness issues when solving for ln γ_2 and would be undesirable for this study.

Next, in Figure 2 the ability of the combinatorial correction (Eq. 13) to predict the activity coefficient of phenanthrene in chloroform is compared to simulation results for phenanthrene at finite concentrations (Eq. 5). All of the simulation data for all of the systems at infinite dilution are contained in Table 1, and the simulation data for phenanthrene at finite concentrations in chloroform are contained in Table 2. Note that for convenience in Table 2, the relevant molar volume of the system, v^*, is defined as

$$v^*(T,p,N_1,N_2)=\frac{\langle V\rangle_{T,p,N_1,N_2-1}N_{\mathrm{avo}}}{N_1+N_2} \tag{25}$$

Within Figure 2, the circles correspond to the finite concentration simulations (Eq. 5), and the dotted (red) horizontal line corresponds to the infinite dilution approximation (Eq. 8). As expected, the discrepancy between the finite concentration simulation results and the infinite dilution approximation becomes more pronounced as the concentration of phenanthrene is increased. For a phenanthrene concentration of $x_2=0.2505$, the finite concentration and infinite dilution results differ by ~0.6 ln units. Also within Figure 2, results are obtained using the combinatorial correction (Eq. 13) with

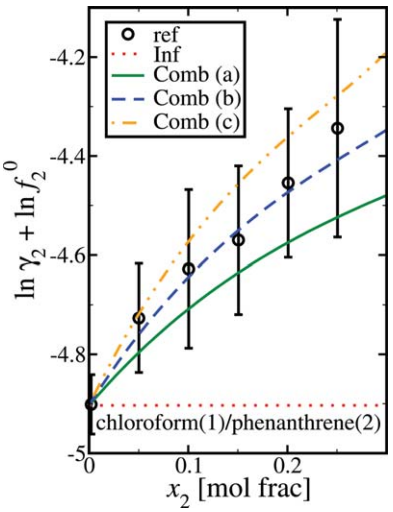

Figure 2. A comparison of the composition dependence of $\ln\gamma_2+\ln f_2^0$ for phenanthrene(2) in chloroform(1) computed by performing reference (ref) simulations at finite concentrations of phenanthrene (Eq. 5), using the infinite dilution approximation (Inf, Eq. 8), and using the combinatorial correction (Comb, Eq. 13) with a seventh-order expansion.

Cases (a)–(c) for the combinatorial correction correspond to use of r=2.5, 2.8, and 3.0, respectively, as described in the text. f_2^0 has units of Pa. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

a seventh-order expansion. A seventh-order expansion was used as higher-order terms are negligible for the value of r and composition range of interest. Three cases are plotted for the combinatorial correction wherein each case is unique in the value of $\bar{v}_2^\infty$ used. For (a), $\bar{v}_2^\infty=200\,\mathrm{cm}^3/\mathrm{mol}$ and r=2.5, which is computed from the simulation of phenanthrene in chloroform at infinite dilution. For (b), $\bar{v}_2^\infty=\hat{\bar{v}}_2^\infty=220\,\mathrm{cm}^3/\mathrm{mol}$ and r=2.8, which is the average value of $\bar{v}_2^\infty$ computed in all of the systems considered at infinite dilution. For (c), $\bar{v}_2^\infty=\hat{\bar{v}}_2=240\,\mathrm{cm}^3/\mathrm{mol}$ and r=3.0, which is the average value of $\bar{v}_2$ computed in all of the simulations of phenanthrene in chloroform. For all three cases, the results obtained are in excellent qualitative and quantitative agreement with the finite concentration simulation results. However, it is apparent that the results are sensitive to the value of $\bar{v}_2^\infty$ used. Be that as it may, within the uncertainty of the finite concentration simulation results, all three cases are in quantitative agreement with the simulation results. Clearly, the proposed simple combinatorial correction to the infinite dilution activity coefficient greatly improves the range of applicability of the simulation results. Also, note that the computed value of $\bar{v}_2^\infty$ differs from the experimentally measured value for the same system of $\bar{v}_2^\infty=155.2\mathrm{cm}^3/\mathrm{mol}$.[94] As a result of the sensitivity of Eq. 13 on $\bar{v}_2^\infty$ and the uncertainty of the computed value of $\bar{v}_2^\infty$ in this study, a future study will be necessary to find the preferred method of calculation. It is also interesting to note that the estimate of the subcooled liquid molar volume of phenanthrene at 298 K estimated using the GC method of Gani and coworkers[24] is v_2=152 cm^3/mol, in good agreement with the experimental value of $\bar{v}_2^\infty$.

Table 1. A Summary of the Dimensionless Residual Chemical Potential ($\beta\mu_2^{\mathrm{res},\infty}$) and Partial Molar Volume of Phenanthrene at Infinite Dilution ($\bar{v}_2^\infty$), and the Molar Volume of the Solvent (v_1) Computed Via Molecular Simulation

Solvent	$\beta\mu_2^{\mathrm{res},\infty}$	$\bar{v}_2^\infty(\mathrm{cm}^3/\mathrm{mol})$	$v_1(\mathrm{cm}^3/\mathrm{mol})$
Methanol	−17.4±0.1	190±20	40.88±0.03
Ethanol	−17.8±0.1	150±10	58.77±0.05
1-Propanol	−18.07±0.08	180±10	75.57±0.06
1-Butanol	−18.43±0.09	190±20	92.36±0.07
1-Pentanol	−18.6±0.1	180±30	108.80±0.04
1-Hexanol	−18.9±0.2	170±30	125.13±0.07
1-Heptanol	−19.1±0.2	180±40	141.2±0.1
1-Octanol	−19.4±0.3	160±40	157.0±0.2
2-Propanol	−17.3±0.2	190±30	77.2±0.1
Water	−5.8±0.2	210±60	17.96±0.01
Chloroform	−22.17±0.06	200±30	78.89±0.02
Acetone[a]	−20.0±0.1	210±10	74.21±0.08
Acetone[b]	−20.82±0.08	200±20	74.8±0.1
2-Butanone[a]	−20.26±0.08	220±20	91.09±0.03
2-Butanone[b]	−20.7±0.1	220±20	91.50±0.07
Benzene	−20.1±0.1	410±70	89.56±0.09
Pyridine	−20.6±0.1	400±100	82.84±0.02
Thiophene	−20.7±0.2	280±20	80.07±0.03
1,4-Dioxane	−20.3±0.2	280±30	85.91±0.05

[a]Using the KGP ketone model.
[b]Using the TraPPE-UA ketone model.

Computing the solid fugacity

Computing $\ln f_2^S/f_2^0$. While solution-phase properties of the solute may readily be computed using molecular simulation, computing the solubility of the solute also requires the determination of $\ln f_2^S$, a property of the pure solid solute. First, focus will be given to the calculation of $\ln f_2^S/f_2^0$ at 298.15 K. Although T_m and Δh_2^{fus} at T_m may in principle be computed by molecular simulation,[38] such simulations are challenging as one must first know the solid solute crystal structure. Early in the design process, the experimental crystal structure is likely unknown, and methods to predict the crystal structure *a priori* have not yet advanced to make accurate predictions.[95] As an alternative, in this study the GC method of Gani and coworkers[24] was used to estimate T_m and Δh_2^{fus} at T_m using the parameters found using the simultaneous fitting procedure.

In Table 3, comparison is made of T_m and Δh_2^{fus} at T_m for phenanthrene from experiment and estimated using the GC method of Gani and coworkers. Throughout the manuscript, agreement with experiment is quantified by computing the mean percent deviation (MPD), defined as

Table 2. A Summary of the Simulation Conditions, the Dimensionless Residual Chemical Potential ($\beta\mu_2^{\mathrm{res}}$) and Partial Molar Volume ($\bar{v}_2$) of Phenanthrene, and the Relevant Molar Volume of the System (v^*, Eq. 25) Computed Via Molecular Simulation for Phenanthrene(2) in Chloroform(1) at Finite Concentrations

x_2 [mol frac]	N_1	N_2	$\beta\mu_2^{\mathrm{res}}$	$\bar{v}_2(\mathrm{cm}^3/\mathrm{mol})$	$v^*(\mathrm{cm}^3/\mathrm{mol})$
0.0021	477	1	−22.17±0.06	200±30	78.72±0.02
0.0502	454	24	−21.9±0.1	190±30	82.26±0.05
0.1004	430	48	−21.8±0.2	230±40	86.03±0.05
0.1506	406	72	−21.7±0.1	240±30	89.88±0.03
0.2008	382	96	−21.5±0.1	320±30	93.73±0.05
0.2505	359	120	−21.4±0.2	260±50	97.56±0.07

Table 3. A Comparison of the Thermophysical Properties of Phenanthrene from Experiment (Exp) and Estimated Using GC Methods

Property	Exp	GC	MPD (%)
T_b (K)	613.45[a]	609.021	0.7
T_c (K)	869.25	867.781	0.2
T_m (K)	372.38	373.087	0.2
h^{fus} (kJ/mol)	16.46	15.881	3.5
$h^{v,0}$ (kJ/mol)	70.40[b]	75.495	
$h^{v,b}$ (kJ/mol)	54.04	56.521	4.6
v^0 (cm^3/mol)	167[b]	152	
c_p^0(J/mol K)	351.8[c]	279.2	
c_p^S(J/mol K)	190.3[d]	208.96	

All of the experimental data are from Yaws' Handbook.[96] The GC estimates for c_p^S and c_p^0 are at a temperature of 298.15 K and were computed using the GC method of Chickos and coworkers.[39,40] All of the other GC properties were computed using the GC method of Gani and coworkers.[24]
[a]Using the Cox equation of Ruuzicka et al.[97] and finding where p^{sat}=0.1 MPa, T_b=607.54 K.
[b]Experimental data at 372 K.
[c]Experimental data at 382 K.
[d]Experimental data at 270 K.

$$\text{MPD} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i^{pred}-y_i^{exp}}{y_i^{exp}}\right| \quad (26)$$

where y^{pred} and y^{exp} are the value of the property of interest predicted (pred) and from experiment (exp), and N is the number of data for which comparison is made. The agreement between the GC method and experiment is very satisfying. The MPD for T_m was 0.2%. With knowledge of the experimental crystal structure, T_m predicted via molecular simulation using conventional force fields can deviate from experiment by up to 20% for small aromatic compounds.[38] The MPD for Δh_2^{fus} at T_m was found to be 3.5%. Via molecular simulation, deviations from experiment can be greater than 10% for small aromatic compounds.[38] Therefore, in the complete absence of experimental data, the GC method appears to perform at least as well as molecular simulation. However, note that phenanthrene is a common industrial compound and is likely included in the parameterization of the GC method. Nonetheless, the method has been validated and tested on a wide range of compounds with good agreement with experiment,[24] suggesting its appropriateness in this study.

Table 3 also includes values of the heat capacity of the solid (c_p^S) and subcooled liquid (c_p^0) at 298.15 K estimated using the GC method of Chickos and coworkers.[39,40] Comparison to experimental values is challenging as a result of the difficulty of the corresponding experiments.[4,6,26,36] An experimental value of c_p^S is available from Yaws' Handbook[96] at 270 K. The deviation from experiment of the estimated value at 298.15 K is 9.8%. The experimental value reported for c_p^0 from Yaws' Handbook[96] is at 382 K. Given the nearly 100 K difference between the experimental value and the GC estimate, a quantitative comparison is not warranted. However, the predicted value appears to be reasonable. A summary of the calculation of all of the GC properties used in this study are contained in the Supporting Information.

A summary of $\ln f_2^S/f_2^0$ estimated at 298.15 K via Eq. 17 and the various approximations for Δc_p^{fus} are summarized in Table 4. For each of the five methods to estimate Δc_p^{fus}, a reference (ref) estimate of $\ln f_2^S/f_2^0$ was computed using the experimental values of T_m and Δh_2^{fus} at T_m, and a GC estimate was made using estimated parameters. A comparison of both estimates is made to the experimental (exp) result computed using the accurate Cox equations of Ruuzicka et al.[97] and assuming $\ln f_2^S/f_2^0 = \ln p_2^{S,sat}/p_2^{0,sat}$. For all cases considered, the MPD was found to range from approximately 3 to 25%. As one might expect, for all of the methods to estimate Δc_p^{fus}, the MPD of the ref estimate is lower than the GC estimate. However, the GC estimates are quite good and are likely suitable for design purposes. Considering just the GC estimates, the MPD was found to range from 5.7 to 25.0%. Having shown that the MPD of the GC estimate for Δh_2^{fus} at T_m was 3.5%, the GC estimates are quite impressive. GC appears promising for design purposes for which molecular simulations are challenging, and experimental data are absent.

Computing $\ln f_2^0$. We next consider the prediction of $\ln f_2^0$. Once $\ln f_2^S/f_2^0$ and $\ln f_2^0$ are known, $\ln f_2^S$ is known. In Table 3, comparison is made of T_b, T_c, $\Delta h_{2,b}^{vap}$, and Δh_2^{vap} at T_0=298 K ($\Delta h_{2,0}^{vap}$) from experiment and estimated using the GC method of Gani and coworkers.[24] For both T_b and T_c, the GC estimates are excellent, with MPDs of 0.7 and 0.2%, respectively. For $\Delta h_{2,b}^{vap}$, the GC estimate deviates from experiment by 4.6%, which corresponds to ~2.5 kJ/mol. The agreement with experiment for these properties is comparable to the agreement found by Rai and Siepmann[76,77] in the parameterization of the TraPPE-EH force field for aromatic compounds, again suggesting that the GC method appears to perform at least as well as molecular simulation for pure component property prediction.

Next, $\ln f_2^0$ may be estimated via Eqs. 22 and 23. As before, a reference (ref) estimate was computed using the experimental values of T_b, T_c, $\Delta h_{2,b}^{vap}$, and $\Delta h_{2,0}^{vap}$, and a GC estimate was made using estimated parameters. A comparison of both estimates is made to the experimental (exp) result computed using the accurate Cox equation of Ruuzicka et al.[97] for liquid phenanthrene. All of the results are presented in Table 5. First, in the top of Table 5, comparison is made at 372 K. This is informative, as it corresponds to the experimental T_m and is the lowest temperature for which the liquid experimental vapor pressure and enthalpy of vaporization are available. Note that the lowest experimental temperature considered by Ruuzicka et al.[97] was 373 K, and the results are, therefore, extrapolated 1 K using the accurately parameterized Cox equation. Rather than assuming n is a universal constant, a unique value was computed for our ref and GC estimates. When computing n for the ref estimate,

Table 4. A Comparison of $\ln f_2^S/f_2^0$ for Phenanthrene at 298.15 K Computed Directly from Experimental (Exp) Vapor Pressure Data[97] and Computed from Both a Reference (Ref) and a GC Estimate Using Eq. 17

	$-\ln f_2^S/f_2^0$				
	Exp	Ref	MPD (%)	GC	MPD (%)
ζ=0	1.3649	1.3236	3.0	1.2868	5.7
ζ=1	1.3649	1.1819	13.4	1.1479	15.9
ζ=1.897	1.3649	1.0548	22.7	1.0233	25.0
Δc_p^{fus} GC	1.3649	1.0984	19.5	1.0576	22.5
$\Delta \hat{c}_p^{fus}$ GC	1.3649	1.1599	15.0	1.1201	17.9

The reference approach uses experimental values of T_m and Δh_2^{fus} at T_m, whereas the GC method uses values estimated using the GC method of Gani and coworkers.[24] Additionally, use of Eq. 17 with five different approaches to estimate Δc_p^{fus} was used, as described in the text. The ζ estimates correspond to Eq. 18, and Δc_p^{fus} GC and $\Delta \hat{c}_p^{fus}$ GC correspond to Eqs. 19 and 20, respectively.

Table 5. A Comparison of $\ln f_2^0$ for Phenanthrene at 372 K (T_m) and for $\ln f_2^0$ and $\ln f_2^S$ at 298.15 K Computed Directly from Experimental (Exp) Vapor Pressure Data[97] and Computed Using Both the GC Properties of Gani and Coworkers[24] in Addition to Using the Reference (Ref) Thermophysical Properties from Yaws' Handbook[96] (see Table 3)

Exp	Ref	MDP (%)	GC	MDP (%)
$\ln f_2^0$ at $T=372$K				
3.3347[a]	3.3571	0.7	3.2244	3.3
$\ln f_2^0$ at $T=298.15$K				
−2.6091	−2.4544	5.9	−2.6845	2.9
$\ln f_2^S$ at $T=298.15$K				
−3.9740	−3.7780	4.9	−3.9713	0.1

For the low pressures of interest, we assume that $f_2^0=p_2^{0,sat}$ and $f_2^S=p_2^{S,sat}$. The experimental data are from the accurate Cox equations of Ruuzicka et al.,[97] parameterized over the range of 371–512 K for the liquid-phase and 200–363 K for the solid-phase. For all cases, f_2^0 and f_2^S have units of Pa.
[a]This agrees well with the value of $\ln f_2^0=3.3460$ from the vapor pressure data in Yaws' Handbook.[96]

$\Delta h_{2,0}^{vap}$ was at 372 K and was at 298 K for the GC estimate. We find that both the ref and the GC estimate are in excellent agreement with experiment. The deviation with experiment is 0.7 and 3.3% for the ref and GC estimate, respectively. The excellent agreement gives confidence in the proposed equations used to estimate $\ln f_2^0$.

In the middle of Table 5, comparison is made at 298.15 K, where phenanthrene is a subcooled liquid. For this case, the experimental value is an extrapolation from the experimental temperature range of 373–512 K using the accurately parameterized Cox equation.[97] Once again, both the ref and the GC estimates are in excellent agreement with experiment. The deviation with experiment is 5.9 and 2.9% for the ref and GC estimate, respectively. The MPD of the GC estimate at 298.15 and 372 K are comparable, with the vapor pressure underestimated in both cases. The level of agreement is quite satisfying considering the experimental vapor pressure at 298.15 K is 7.3×10^{-2} Pa, and the GC estimate is 6.8×10^{-2} Pa; note that the vapor pressure is seven orders of magnitude smaller than atmospheric pressure (1.0×10^5 Pa).

Computing $\ln f_2^S$. Having calculated $\ln f_2^0$ and $\ln f_2^S/f_2^0$ at 298.15 K, in the bottom of Table 5, we compare our ref and GC estimate of $\ln f_2^S$ to experiment. For both the ref and the GC estimates, $\ln f_2^S/f_2^0$ calculated with $\zeta=0$ is used, as it yielded the closest agreement with experiment for both cases. Once again, both the ref and the GC estimates are in excellent agreement with experiment. The deviation with experiment is 4.9 and 0.1% for the ref and GC estimate, respectively. For the ref estimate, $\ln f_2^0$ and $\ln f_2^S/f_2^0$ at 298.15 K were both overestimated with MPDs of 5.9 and 3.0%, respectively, and $\ln f_2^S$ is in turn also overestimated with an MPD of 4.9%. On the other hand, for the GC estimate, $\ln f_2^0$ is underestimated with an MPD of 2.9% and $\ln f_2^S/f_2^0$ is overestimated with an MPD of 5.7%. Therefore, the low MPD of 0.1% for $\ln f_2^S$ results from a fortunate cancellation of errors. Future testing will be necessary to determine the accuracy with which GC estimates may be made for other solutes, but the results of this study are promising.

Solubility predictions

Molecular Simulation. All of the necessary properties computed at infinite dilution via molecular simulation for phenanthrene in each solvent of interest are summarized in Table 1, and the computed solubility data are summarized in Table 6. The value of $\ln f_2^S$ was computed using GC properties, as discussed in the previous subsection. When predicting the solubility using the combinatorial corrected infinite dilution activity coefficient (Eq. 14), two cases are again considered. For (a), $\bar{v}_2^\infty$ is computed from the simulation of phenanthrene in the solvent of interest at infinite dilution. For (b), $\bar{v}_2^\infty=\hat{\bar{v}}_2^\infty=220\,\mathrm{cm}^3/\mathrm{mol}$, which is the average value of $\bar{v}_2^\infty$ computed in all of the solvents at infinite dilution. For both cases, a seventh-order expansion was used.

Of the solvents studied, the experimental solubility of phenanthrene in chloroform[98] of $x_2=0.26778$ is greater than in any of the other studied solvents. This also corresponds to the case for which the combinatorial correction offers the greatest improvement over the infinite dilution approximation (Eq. 9). In fact, the infinite dilution approximation predicts an unphysical value of $x_2=2.5$ for the solubility of phenanthrene in chloroform. This is clearly in error as $x_2\leq1$ by definition. As demonstrated in Figure 2, for phenanthrene in chloroform, the combinatorial correction accurately captures the composition dependence of the activity coefficient as compared to finite concentration simulation results. If the combinatorial correction is used to estimate the solubility of phenanthrene in chloroform, the predicted solubility is (a) $x_2=0.62$ and (b) $x_2=0.53$ for each case. That is, using the simple combinatorial correction, the prediction is drastically improved. This result is also extremely informative in relation to Figure 2. At a phenanthrene concentration in chloroform of $x_2=0.2505$, it was shown that $\ln\gamma_2$ from the finite concentration simulation results and the infinite dilution approximation differed by $\sim$0.6 ln units. Although this may seem minute, it has a large implication when predicting phase equilibria due to the exponential relation between x_2 and $\ln\gamma_2$.

In Table 6, the experimental solubility of phenanthrene in 17 solvents is compared to results using both the infinite dilution approximation and the combinatorial correction. For all cases, when computing the MPD, the solubility of phenanthrene in chloroform for the infinite dilution approximation was taken to be $x_2=1.0$. Although this has the effect of lowering the corresponding MPD, this was done as the predicted value is unphysical. Overall, the MPD computed using the combinatorial correction is about 50% lower than the infinite dilution approximation.

If only the nine alcohol solvents and water are considered, the computed MPD found using each method are in reasonably close agreement. For these cases, either the solubility or r is small, such that the combinatorial correction is minor. Additionally, for all of these cases, the MPD with experiment is quite large, between 224 and 250%. For the studied alcohols, the solubility is always overpredicted, and the solubility is underpredicted for water using the employed force fields. Although the employed force fields have been optimized to reproduce pure component properties, this would suggest that an improvement in the parameterization of the solute–solvent interactions could ultimately improve the solubility predictions. A detailed study by Potoff et al.[99] found that when predicting vapor–liquid coexistence properties of binary mixtures via molecular simulation, the choice of mixing rules may play an important rule. Likewise, describing binary interactions using pure component properties with mixing rules has long been a challenge of equation of state modeling.[26] Also, the molecular simulation work of Mobley

Table 6. A Summary of the Predicted Solubility of Phenanthrene(2) in Each Solvent as Compared to Experiment (Exp) Using Both the Infinite Dilution Approximation (Inf, Eq. 9) and the Combinatorial Correction (Comb, Eq. 14) Using a Seventh-Order Expansion, along with the Computed MPD

	x_2 (mol frac)			
Solvent	Exp	Inf	Comb (a)	Comb (b)
Methanol	0.00589[101]	0.011±0.001	0.010±0.001	0.010±0.001
Ethanol	0.01114[101]	0.024±0.003	0.022±0.003	0.020±0.003
1-Propanol	0.01355[101]	0.041±0.003	0.038±0.003	0.036±0.003
1-Butanol	0.01771[101]	0.071±0.007	0.066±0.006	0.064±0.006
1-Pentanol	0.02491[101]	0.10±0.01	0.09±0.01	0.09±0.01
1-Hexanol	0.03028[101]	0.15±0.02	0.15±0.02	0.14±0.02
1-Heptanol	0.03937[101]	0.21±0.04	0.21±0.04	0.20±0.03
1-Octanol	0.05418[101]	0.3±0.1	0.3±0.1	0.3±0.1
2-Propanol	0.00977[101]	0.019±0.003	0.018±0.003	0.018±0.003
Water	1.170×10^{-7}[c]	$4.4\pm0.8 \times 10^{-8}$	$4.4\pm0.8 \times 10^{-8}$	$4.4\pm0.8 \times 10^{-8}$
Chloroform	0.26778[98]	2.5±0.2[e]	0.62±0.04	0.53±0.03
Acetone[a]	0.17145[98]	0.28±0.03	0.19±0.02	0.18±0.02
Acetone[b]	0.17145[98]	0.63±0.05	0.36±0.03	0.32±0.02
2-Butanone[a]	0.20900[101]	0.43±0.04	0.30±0.03	0.30±0.03
2-Butanone[b]	0.20900[101]	0.71±0.08	0.45±0.05	0.45±0.05
Benzene	0.20909[98]	0.37±0.04	0.13±0.01	0.26±0.03
Pyridine	0.24125[d]	0.53±0.08	0.14±0.02	0.32±0.05
Thiophene	0.23401[d]	0.6±0.1	0.25±0.04	0.34±0.06
1,4-Dioxane	0.21650[101]	0.42±0.08	0.21±0.04	0.28±0.05
Total MPD (%)[a]		199[f]	156	148
Total MPD (%)[b]		219[f]	166	157
R—OH MPD (%)[g]		250	238	224
Non-R—OH MPD (%)[a]		128[f]	38	39

For Comb (a), $\bar{v}_2^{\infty}$ is computed from the simulation of phenanthrene in the solvent of interest at infinite dilution (see Table 1). For Comb (b), $\bar{v}_2^{\infty}=\bar{v}_2^{\infty}=220\,\mathrm{cm}^3/\mathrm{mol}$, which is the average value of $\bar{v}_2^{\infty}$ computed in all of the solvents at infinite dilution.
[a]Using the KGP ketone model.
[b]Using the TraPPE-UA ketone model.
[c]DECHEMA.[104]
[d]Estimated by extrapolating the data of Choi and McLaughlin[105] to 298.15 K using a van't Hoff plot.
[e]The infinite dilution approximation predicts an unphysical result for this case.
[f]Using $x_2=1$ for phenanthrene in chloroform.
[g]R—OH corresponds to the alcohols and water.

et al.[93] demonstrated the sensitivity of the computed value of $\mu_2^{\mathrm{res},\infty}$ for nonelectrolyte solutes in water on the method with which the solute partial charges were parameterized. A similar artifact may be the culprit in this study.

On the other hand, examining the solubility in only chloroform, acetone, 2-butanone, benzene, pyridine, thiophene, and 1,4-dioxane, there is a significant difference in the MPD found using the infinite dilution approximation and using the combinatorial correction. Here, we will consider the results using the KGP ketone model, as it was found to yield improved predictions over TraPPE-UA, and both the KGP and the TraPPE-UA models ultimately lead to similar observations. Although the infinite dilution approximation yields an MPD of 128%, it drops to 38 and 39% when using the combinatorial correction with (a) and (b), respectively. For all cases, we find that the combinatorial correction lowers the predicted solubility relative to the infinite dilution approximation and improves the agreement with experiment. This result is very satisfying and is comparable to the level of agreement found with the latest 2005 revision of MOSCED[11] (as will be described in detail momentarily). For the same test set minus thiophene, for which MOSCED parameters are not available, the computed MPD using the combinatorial correction is 30%. Also, this is comparable to the level of agreement with experiment found by Acree and Abraham[100] when predicting the solubility of phenanthrene in a range of solvents using the semitheoretical Abraham general solvation model. Using two different forms of the Abraham general solvation model to predict the solubility of phenanthrene in a range of 23 solvents, Acree and Abraham[100] report MPDs of 26.5 and 42.5% (note that the reported solubilities were given in units of molarity). Additionally, note that the experimental data for chloroform, acetone, and benzene were all taken from Henstock,[98] which was published in 1922; the purity and accuracy of the experiments may be imperfect by modern standards.[1] The solubility of phenanthrene in methanol and ethanol at 298.15 K is 26 and 27%, respectively, larger in the work of Henstock[98] from 1922 as compared to the modern work of Hernandez et al.[101] from 1999.

For all cases, the combinatorial correction lowers the predicted solubility relative to the infinite dilution approximation. This is expected as the FH theory predicts only negative deviations from Raoult's Law for $r>1$.[26] This fact additionally limits the systems for which the combinatorial correction is anticipated to be valid. For the case of a solute containing a carboxylic acid functional group that may associate with other solute molecules in a nonpolar solvent,[26] use of only the combinatorial correction would not be suitable. Within the context of this study, this shortcoming may potentially be overcome by accounting for the residual activity coefficient using solubility parameter-based methods (i.e., Hildebrand–Scatchard or Hansen), which predict only positive deviations from Raoult's Law.[4,26,36] The Hildebrand–Scatchard solubility parameter of the solvent may readily be computed via molecular simulation,[102] and the Hildebrand–Scatchard and Hansen solubility parameters of the solvent and solid solute may readily be estimated using the GC

Table 7. A Summary of the Solvent Molar Volume (v_1) at 293.15 K and the Infinite Dilution Activity Coefficient of Phenanthrene ($\ln \gamma_2^\infty$) at 298.15 K Calculated Using the 2005 Revised MOSCED Model[11]

			x_2 (mol frac)	
Solvent	v_1 (cm^3/mol)	$\ln \gamma_2^\infty$	Inf	Comb
Methanol	40.6	4.2667	0.0039	0.0037
Ethanol	58.6	3.7175	0.0067	0.0066
1-Propanol	75.1	3.2541	0.0107	0.0105
1-Butanol	92.0	2.9146	0.0150	0.0148
1-Pentanol	108.5	2.7023	0.0185	0.0184
1-Hexanol	125.2	2.8365	0.0162	0.0162
1-Octanol	158.2	2.5521	0.0215	0.0215
2-Propanol	76.8	3.4775	0.0085	0.0084
Water	18.0	15.2562	6.538×10^{-8}	6.538×10^{-8}
Chloroform	80.5	−0.2277	0.3468	0.2792
Acetone	73.8	1.4544	0.0645	0.0592
2-Butanone	90.2	0.8221	0.1214	0.1131
Benzene	89.5	0.1167	0.2457	0.2172
Pyridine	80.9	0.8145	0.1223	0.1100
1,4-Dioxane	85.7	0.0069	0.2743	0.2350
Total MPD (%)			35	33
R—OH MPD (%)			33	34
Non-R—OH MPD (%)			38	30

These data are used to predict the solubility of phenanthrene with the infinite dilution approximation (Inf) and the proposed combinatorial correction (Comb) using a seventh-order expansion and v_2=167.1 cm^3/mol, as used in the MOSCED model.[11] The experimental solubility data may be found in Table 6 and were used to compute the corresponding MPD of the estimates. The MPD is further broken down as in Table 6, where R—OH corresponds to the alcohols and water. Note that when computing $\ln \gamma_2^\infty$ using MOSCED, v_1 of water is artificially doubled.

methods of Gani and coworkers.[24,103] The investigation of such an approach will be left for a future study.

MOSCED. In addition to the use of molecular simulation, here we consider the use of the 2005 revised MOSCED model developed by Eckert and coworkers.[11,25] MOSCED is a semitheoretical model to predict infinite dilution activity coefficients and describes each molecule using five adjustable parameters. The interested reader is directed to the original publication for a description of its formulation and basis.[25] In the 2005 revision, the molecular descriptors of 132 organic solvents, water, and five permanent gases were fit to a set of 6441 experimental infinite dilution activity coefficient data points.[11] In addition, parameters for 26 organic solids were regressed. For the case of phenanthrene, the parameters were obtained by regressing to solubility data in 37 organic solvents. This was accomplished by calculating $\ln f_2^S/f_2^0$ using experimental data and using infinite dilution activity coefficients computed using MOSCED to obtain the parameters for either the two parameter Wilson or UNIQUAC excess Gibbs free energy model, which were used to estimate the solubility.[11] For the 37 organic solvents to which MOSCED was fit, MOSCED obtained an MPD of 23.8%, whereas predictions using the state-of-the-art modified UNIFAC model obtained an MPD of 36.7%.

In this study, we consider the use of MOSCED to predict the solubility of phenanthrene in a set of 15 solvents. The solvent set is identical to that considered using molecular simulation save for thiophene and 1-heptanol, for which MOSCED parameters do not exist. In Table 7, we list the value of $\ln \gamma_2^\infty$ for phenanthrene predicted using MOSCED, along with the molar volume of the solvent at 293.15 K, as reported by Eckert and coworkers.[11] Note that for water, when predicting $\ln \gamma_2^\infty$, the molar volume of water is artificially doubled to 36 cm^3/mol.[11] In Table 7, we report the experimental value, as it will be used in application of the combinatorial correction. Using $\ln f_2^S/f_2^0$ computed in this study with ζ=0, we predicted the solubility of phenanthrene using an analogous procedure as with our molecular simulation data, with results reported in Table 7. For the combinatorial correction prediction, only a single case is considered, using the value of v_2=167.1 cm^3/mol reported by Eckert and coworkers.[11]

If only the eight alcohol solvents and water are considered, the computed MPD found using each method are in close agreement. This is in agreement with our molecular simulation results. For these cases, either the solubility or r is small, such that the combinatorial correction is minor. The MPDs of 33% using the infinite dilution approximation and 34% using the combinatorial correction are both superior to the MPDs found using molecular simulation, again suggesting shortcomings in the employed mixing rules and/or the employed molecular models.

Next, considering the solubility in only chloroform, acetone, 2-butanone, benzene, pyridine, and 1,4-dioxane, the difference in the MPD found using the infinite dilution approximation and using the combinatorial correction of 8% is small compared to the large difference observed with our molecular simulation results. However, first let us take a closer look at acetone, 2-butanone, and pyridine. For these cases, the infinite dilution approximation underestimates the experimental solubility by ~0.1 mol frac. Further, for these cases, the infinite dilution approximation estimates range from ~0.06 to 0.12 mol frac. As a result, the combinatorial correction is small and serves to slightly decrease the estimated solubility. Next, let us look closer at chloroform, benzene, and 1,4-dioxane. For these cases, the infinite dilution approximation overestimates the experimental solubility by 0.08, 0.04, and 0.06 mol frac for chloroform, benzene, and 1,4-dioxane, respectively, with an overall MPD of 24.6%. For these cases, the infinite dilution approximation estimates range from 0.25 to 0.35 mol frac, and the combinatorial correction is, therefore, appreciable. For these cases, the combinatorial correction lowers the predicted solubility as compared to the infinite dilution approximation, yielding excellent agreement with experiment with an MPD of 5.6%.

For the 15 solvents considered, the overall MPD was 35 and 33% using the infinite dilution approximation and the

combinatorial correction, respectively. This is larger than the MPD reported by Eckert and coworkers[11] of 23.8% for the 37 solvents to which the model was parameterized. However, this is largely the result of the value of $\ln f_2^S/f_2^0$ used in this study, which was obtained using GC parameters from Gani and coworkers[24] with $\zeta=0$. The corresponding MPD of the GC estimate of $\ln f_2^S/f_2^0$ is 5.7% as compared to experiment. While the MPD of $\ln f_2^S/f_2^0$ is only 5.7%, x_2 is exponentially related to $\ln f_2^S/f_2^0$. Ultimately, we chose to estimate $\ln f_2^S/f_2^0$ using GC parameters rather than use the experimental value, as it allows for purely predictive solubility estimates, devoid of experimental data.

Summary and Conclusions

A simple combinatorial correction to the infinite dilution activity coefficient for a solute in solution has been proposed. In the spirit of the UNIQUAC excess Gibbs free energy model, $\ln \gamma_2$ is fundamentally interpreted as a sum of a residual and a combinatorial term. The proposed method assumes that the residual term is relatively insensitive to composition and that the composition dependence of the combinatorial term may be accounted for using the athermal Flory-Huggins (FH) theory. While any suitable expression for the combinatorial term may be used, the necessary terms required by the FH theory may readily be computed via molecular simulation. The proposed method uses only properties for the solute computed at infinite dilution using conventional free energy simulation methodologies to estimate the solution-phase behavior of the solute at finite concentrations. The necessary methodology is implemented in most open-source simulation software and is, therefore, readily adoptable by other simulators.

Although the properties of the solution-phase were calculated using molecular simulation, calculating the pure solid solute fugacity using molecular simulation is challenging. On the other hand, although predicting properties of an arbitrary solute in an arbitrary solvent may prove challenging, GC methods may accurately predict pure component properties. As a result, in this study, the pure solid solute fugacity was accurately computed using the GC method of Gani and coworkers. In this fashion, solubility predictions may be made in the complete absence of experimental data.

The proposed methodology was used to predict the solubility of phenanthrene in a range of 17 unique solvents at ambient conditions. Overall, predictions made using the combinatorial correction are in better agreement with experiment as compared to the infinite dilution approximation, especially for systems where the molar solubility is greater than 5%. When the molar solubility is less than 5%, the combinatorial correction term was found to be negligible. For the specific case of phenanthrene in chloroform using molecular simulation, the proposed methodology results in a substantial improvement over the infinite dilution approximation, which yields an unphysical result.

The molecular simulation results predicted for phenanthrene in water and the studied alcohols are in poor quantitative agreement with experiment. On the other hand, for the other seven solvents considered, the results of this study are in good agreement with experiment, with deviations comparable to the latest 2005 revision of MOSCED. The results in water and the studied alcohols may likely be improved through improvements of the employed mixing rules and/or molecular models. While the method was used in this study to predict the solubility of phenanthrene in solution, the proposed method is additionally suited to estimate partition coefficients. The promising results of this study warrant future studies to probe the application to a wider range of solvents and solutes and application to predict partition coefficients.

Acknowledgments

The authors thank Neeraj Rai (Notre Dame) for assistance with the TraPPE-EH force field and Wilfried Cordes and Bastian Schmid (DDBST GmbH) for assistance with MOSCED. A.S.P. gratefully acknowledges a fellowship from the Arthur J. Schmitt Foundation and additional funding provided by the Notre Dame Sustainable Energy Initiative. Computing support was provided by Notre Dame's Center for Research Computing.

Literature Cited

1. Jouyban A. *Handbook of Solubility Data for Pharmaceuticals*. Boca Raton, FL: CRC Press/Taylor and Francis Group, 2010.
2. Liu R, editor. *Water-Insoluble Drug Formulation*, *2nd ed*. Boca Raton, FL: CRC Press, 2008.
3. Boduszynski MM. Composition of heavy petroleums. 2. Molecular characterization. *Energy Fuels*. 1988;2:597–613.
4. Walas SM. *Phase Equilibria in Chemical Engineering*. Stoneham, MA: Butterworth Publishers, 1985.
5. Meindersma GW, Hansmeier AR, de Haan AB. Ionic liquids for aromatic extraction. Present status and future outlook. *Ind Eng Chem Res*. 2010;49:7530–7540.
6. Nordstrom F, Rasmuson AC. Predicting the solubility curves and melting properties of organic and pharmaceutical compounds. *Eur J Pharm Sci*. 2009;36:330–344.
7. Abildskov J, O'Connell JP. Predicting the solubilities of complex chemicals. I. Solutes in different solvents. *Ind Eng Chem Res*. 2003;42:5622–5634.
8. Cassens J, Ruether F, Leonhard K, Sadowski G. Solubility calculation of pharmaceutical compounds—a priori parameter estimation using quantum-chemistry. *Fluid Phase Equilib*. 2010;299:161–170.
9. Klamt A, Eckert F, Hornig M, Beck ME, Burger T. Prediction of aqueous solubility of drugs and pesticides with COSMO-RS. *J Comput Chem*. 2002;23:275–281.
10. Hsieh CM, Wang S, Lin ST, Sandler SI. A predictive model for the solubility and octanol–water partition coefficient of pharmaceuticals. *J Chem Eng Data*. 2011;56:936–945.
11. Lazzaroni MJ, Bush D, Eckert CA, Frank TC, Gupta S, Olson JD. Revision of MOSCED parameters and extension to solid solubility calculations. *Ind Eng Chem Res*. 2005;44:4075–4083.
12. Jorgensen WL, Duffy EM. Prediction of drug solubility from Monte Carlo simulations. *Bioorg Med Chem Lett*. 2000;10:1155–1158.
13. Ferrario M, Ciccotti G, Spohr E, Cartailler T, Turq P. Solubility of KF in water by molecular dynamics using the Kirkwood integration method. *J Chem Phys*. 2002;117:4947–4953.
14. Sanz E, Vega C. Solubility of KF and NaCl in water by molecular simulation. *J Chem Phys*. 2007;126:014507.
15. Paluch AS, Jayaraman S, Shah JK, Maginn EJ. A method for computing the solubility limit of solids: application to sodium chloride in water and alcohols. *J Chem Phys*. 2010;133:124504.
16. Moucka F, Lisal M, Skvor J, Jirsak J, Nezbeda I, Smith WR. Molecular simulation of aqueous electrolyte solubility. 2. Osmotic ensemble Monte Carlo methodology for free energy and solubility calculations and application to NaCl. *J Phys Chem B*. 2011;115:7849–7861.
17. Schnieders MJ, Baltrusaitis J, Shi Y, Chattree G, Zheng LQ, Yang W, Ren PY. The structure, thermodynamics, and solubility of organic crystals from simulation with a polarizable force field. *J Chem Theory Comput*. 2012;8:1721–1736.
18. Aragones JL, Sanz E, Vega C. Solubility of NaCl in water by molecular simulation revisited. *J Chem Phys*. 2012;136:244508.
19. Jayaraman S, Thompson AP, von Lilienfeld A, Maginn EJ. Molecular simulation of the thermal and transport properties of three alkali nitrate salts. *Ind Eng Chem Res*. 2010;49:559–571.

20. Paluch AS, Cryan DD, Maginn EJ. Predicting the solubility of the sparingly soluble solids 1,2,4,5-tetramethylbenzene, phenanthrene, and fluorene in various organic solvents by molecular simulation. *J Chem Eng Data*. 2011;56:1587–1595.
21. Haile JM. On the use of computer simulation to determine the excess free energy in fluid mixtures. *Fluid Phase Equilib.* 1986;26: 103–127.
22. Liu Y, Li X, Wang L, Sun H. Prediction of partition coefficients and infinite dilution activity coefficients of 1-ethylpropylamine and 3-methyl-1-pentanol using force field methods. *Fluid Phase Equilib.* 2009;285:19–23.
23. Marrero J, Gani R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib*. 2001;183–184:183–208.
24. Hukkerikar AS, Sarup B, Kate AT, Abildskov J, Sin G, Gani R. Grop-contribution+ (GC+) based estimation of properties of pure components: improved property estimation and uncertainty analysis. *Fluid Phase Equilib*. 2012;321:25–43.
25. Thomas ER, Eckert CA. Prediction of limiting activity coefficients by a modified separation of cohesive energy density model and UNIFAC. *Ind Eng Chem Process Des Dev*. 1984;23:194–209.
26. Prausnitz JM, Lichtenhaler RN, de Azevedo EG. Molecular Thermodynamics of Fluid-Phase Equilibria, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1986.
27. Shing KS, Chung ST. Computer simulation methods for the calculation of solubility in supercritical extraction systems. *J Phys Chem.* 1987;91:1674–1681.
28. Shirts MR, Pitera JW, Swope WC, Pande VS. Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. *J Chem Phys*. 2003;119:5740–5761.
29. Jorgensen WL, Briggs JM, Contreras ML. Relative partition coefficients for organic solutes from fluid simulations. *J Phys Chem.* 1990;94:1683–1686.
30. Garrido NM, Economou IG, Queimada AJ, Jorge M, Macedo EA. Prediction of the n-hexane/water and 1-octanol/water partition coefficients for environmentally relevant compounds using molecular simulation. *AIChE J*. 2012;58:1929–1938.
31. Abrams DS, Prausnitz JM. Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE J*. 1975;21:116–128.
32. Debenedetti PG, Kumar SK. Infinite dilution fugacity coefficients and the general behavior of dilute binary systems. *AIChE J*. 1986;32:1253–1262.
33. Sandler SI. The generalized van der Waals partition function as a basis for excess free energy models. *J Supercrit Fluids*. 2010;55: 496–502.
34. Hildebrand JH. The entropy of solution of molecules of different size. *J Chem Phys*. 1947;15:225–228.
35. Kikic I, Alessi P, Rasmussen P, Fredenslund A. On the combinatorial part of the UNIFAC and UNIQUAC models. *Can J Chem Eng*. 1980;58:253–258.
36. Hildebrand JH, Prausnitz JM, Scott RL. Regular and Related Solutions. New York, NY: Van Nostrand Reinhold Company, 1970.
37. Poling BE, Prausnitz JM, O'Connell JP. The Properties of Gases and Liquids, 5th ed. New York, NY: McGraw-Hill Companies, Inc. 2001.
38. Eike DM, Maginn EJ. Atomistic simulation of solid–liquid coexistence for molecular systems: application to triazole and benzene. *J Chem Phys*. 2006;124:164503.
39. Chickos JS, Hesse DG, Liebman JF. A group additivity approach for the estimation of heat capacities of organic liquids and solids at 298 K. *Struct Chem*. 1993;4:261–269.
40. Acree WE, Chickos JS. Phase transition enthalpy measurements of organic and organometallic compounds. Sublimation, vaporization and fusion enthalpies from 1880 to 2010. *J Phys Chem Ref Data*. 2010;39:043101.
41. Smith JM, Ness HCV, Abbott MM. Introduction to Chemical Engineering Thermodynamics, 7th ed. New York, NY: McGraw-Hill, 2005.
42. Miller DG. Estimating vapor pressures—a comparison of equations. *Ind Eng Chem*. 1964;56:46–57.
43. Siepmann JI, Karaborni S, Smit B. Simulating the critical properties of complex fluids. *Nature*. 1993;365:330–332.
44. Paluch AS, Shen VK, Errington JR. Comparing the use of Gibbs ensemble and grand-canonical transition-matrix Monte Carlo methods to determine phase equilibria. *Ind Eng Chem Res*. 2008;47:4533–4541.
45. Rai N, Bhatt D, Siepmann JI, Fried LE. Monte Carlo simulations of 1,3,5-triamino-2,4,6-trinitrobenzene (TATB): pressure and temperature effects for the solid phase and vapor–liquid phase equilibria. *J Chem Phys*. 2008;129:194510.
46. Rai N, Maginn EJ. Critical behaviour and vapour–liquid coexistence of 1-alkyl-3-methylimidazolium bis-(trifluoromethylsulfonyl)amide ionic liquids via Monte Carlo simulations. *Faraday Discuss*. 2012;154:53–69.
47. Lyubartsev AP, Martsinovski AA, Shevkunov SV, Vorontsov-Velyaminov PN. New approach to Monte Carlo calculation of the free energy: method of expanded ensembles. *J Chem Phys*. 1992;96: 1776–1783.
48. Lyubartsev AP, Laaksonen A, Vorontsov-Velyaminov PN. Free energy calculations for Lennard-Jones systems and water using the expanded ensemble method: a Monte Carlo and molecular dynamics simulation study. *Mol Phys*. 1994;82:455–471.
49. Wilding NB, Muller M. Accurate measurements of the chemical potential of polymeric systems by Monte Carlo simulation. *J Chem Phys*. 1994;101:4324–4330.
50. Paluch AS, Shah JK, Maginn EJ. Efficient solvation free energy calculations of amino acid analogs by expanded ensemble molecular simulation. *J Chem Theory Comput*. 2011;7:1394–1403.
51. Torrie GM, Valleau JP. Monte Carlo free energy estimates using non-Boltzmann sampling: application to the sub-critical Lennard-Jones fluid. *Chem Phys Lett*. 1974;28:578–581.
52. Wang F, Landau DP. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett*. 2001;86: 2050–2053.
53. Yan Q, Faller R, de Pablo JJ. Density-of-states Monte Carlo method for simulation of fluids. *J Chem Phys*. 2002;116:8745–8749.
54. Shell MS, Debenedetti PG, Panagiotopoulos AZ. Generalization of the Wang–Landau method for off-lattice simulations. *Phys Rev E*. 2002;66:056703.
55. Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys*. 1976;22:245–268.
56. Lu N, Singh JK, Kofke DA. Appropriate methods to combine forward and reverse free-energy perturbation averages. *J Chem Phys*. 2003;118:2977–2984.
57. Shirts MR, Bair E, Hooker G, Pande VS. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys Rev Lett*. 2003;91:140601.
58. Fenwick MK, Escobedo FA. On the use of Bennett's acceptance ratio method in multi-canonical-type simulations. *J Chem Phys*. 2004;120:3066–3074.
59. Paluch AS, Mobley DL, Maginn EJ. Small molecule solvation free energy: enhanced conformational sampling using expanded ensemble molecular dynamics simulations. *J Chem Theory Comput*. 2011;7: 2910–2918.
60. Chipot C, Pohorille A, editors. Free Energy Calculations: Theory and Applications in Chemistry and Biology, vol. 86 of Springer Series in Chemical Physics. New York, NY: Springer, 2007.
61. Lu N, Kofke DA, Woolf TB. Staging is more important than perturbation method for computation of enthalpy and entropy changes in complex systems. *J Phys Chem B*. 2003;107:5598–5611.
62. Lyubartsev AP, Laaksonen A. M.DynaMix—a scalable portable parallel MD simulation package for arbitrary molecular mixtures. *Comp Phys Commun*. 2000;128:565–589.
63. Lyubartsev AP, Laaksonen A. MDynaMix: a molecular dynamics program. Available at: http://www.fos.su.se/sasha/mdynamix/ (accessed February 1, 2010).
64. Allen MP, Tildesley DJ. Computer Simulation of Liquids. New York, NY: Oxford University Press, Inc., 1987.
65. Finchman D. Optimisation of the Ewald Sum for Large Systems. *Mol Simul*. 1994;13:1–9.
66. Tuckerman M, Berne BJ, Martyna GJ. Reversible multiple time scale molecular dynamics. *J Chem Phys*. 1992;97:1990–2001.
67. Andersen HC. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys*. 1980;72:2384–2393.
68. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys*. 1994;101:4177–4189.
69. Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML. Explicit reversible integrators for extended system dynamics. *Mol Phys*. 1996;87: 1117–1157.
70. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett*. 1994;222:529–539.

71. Shirts MR, Pande VS. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J Chem Phys*. 2005;122:134508.
72. Steinbrecher T, Mobley DL, Case DA. Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J Chem Phys*. 2007;127:214108.
73. Martin MG, Siepmann JI. Transferable potentials for phase equilibria. 1. United-atom description of *n*-alkanes. *J Phys Chem B*. 1998;102:2569–2577.
74. Martin MG, Siepmann JI. Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes. *J Phys Chem B*. 1999;103:4508–4517.
75. Chen B, Potoff JJ, Siepmann JI. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *J Phys Chem B*. 2001;105:3093–3104.
76. Rai N, Siepmann JI. Transferable potentials for phase equilibria. 9. explicit hydrogen description of benzene and five-membered and six-membered heterocyclic aromatic compounds. *J Phys Chem B*. 2007;111:10790–10799.
77. Rai N, Siepmann JI. Transferable potentials for phase equilibria. 10. Explicity-hydrogen description of substituted benzenes and polycyclic aromatic compounds. *J Phys Chem B*. 2013;117:273–288.
78. Stubbs JM, Potoff JJ, Siepmann JI. Transferable potentials for phase equilibria. 6. United-atom description for ethers, glycols, ketones, and aldehydes. *J Phys Chem B*. 2004;108:17596–17605.
79. Kamath G, Georgiev G, Potoff JJ. Molecular modeling of phase behavior and microstructure of acetone–chloroform–methanol binary mixtures. *J Phys Chem B*. 2005;109:19463–19473.
80. Yazaydin AO, Thompson RW. Simulating the vapour–liquid equilibria of 1,4-dioxane. *Mol Simul*. 2006;32:657–662.
81. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformation energies of organic and biological molecules? *J Comput Chem*. 2000;21:1049–1074.
82. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4:435–447.
83. GROMACS: Fast, flexible, and free. Available at: http://www.gromacs.org/ (accessed July 1, 2009).
84. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc*. 1996;118:11225–11236.
85. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. 2004;25:1157–1174.
86. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Modell*. 2006;25:247–260.
87. Wu Y, Tepper HL, Voth GA. Flexible simple point-charge water model with improved liquid-state properties. *J Chem Phys*. 2006;124:024503.
88. Raabe G, Sadus RJ. Influence of bond flexibility on the vapor–liquid phase equilibria of water. *J Chem Phys*. 2007;126:044701.
89. Raabe G, Sadus RJ. Molecular dynamics simulation of the dielectric constant of water: the effect of bond flexibility. *J Chem Phys*. 2011;134:234501.
90. Bennett CH. Mass tensor molecular dynamics. *J Comput Phys*. 1975;19:267–279.
91. Feenstra KA, Hess B, Berendsen HJC. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J Comput Chem*. 1999;20:786–798.
92. Haynes WM, editor. *CRC Handbook of Chemistry and Physics*, *92nd ed*. Boca Raton, FL: CRC Press/Taylor and Francis, 2011–2012; Internet Version 2012.
93. Mobley DL, Dumont E, Chodera JD, Dill KA. comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. *J Phys Chem B*. 2007;111:2242–2254.
94. Ohba M, Takigawa T, Ogawa H, Murakami S, Nomura H. Thermodynamic properties of rigid polycyclic molecules. 2. Partial molar volumes of polycyclic aromatics compared with the RISM integral equation theory. *Fluid Phase Equilib*. 1997;136:289–297.
95. Day GM, Cooper TG, Cruz-Cabeza AJ, Hejczyk KE, Ammon HL, Boerrigter SXM, Tan JS, Della Valle RG, Venuti E, Jose J, Gadre SR, Desiraju GR, Thakur TS, van Eijck BP, Facelli JC, Bazterra VE, Ferraro MB, Hofmann DWM, Neumann MA, Leusen FJJ, Kendrick J, Price SL, Misquitta AJ, Karamertzanis PG, Welch GWA, Scheraga HA, Arnautova YA, Schmidt MU, van de Streek J, Wolf AK, Schweizer B. Significant progress in predicting the crystal structures of small organic molecules—a report on the fourth blind test. *Acta Crystallogr Sect B*. 2009;65:107–125.
96. Yaws CL. *Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds*. Knovel, 2003. Available at: http://www.knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=667&VerticalID=0, Electronic ISBN: 978-1-59124-444-8 (accessed August 1, 2012).
97. Ruuzicka K, Mokbel I, Majer V, Ruuzicka V, Jose J, Zabransky M. Description of vapour–liquid and vapour–solid equilibria for a group of polycondensed compounds of petroleum interest. *Fluid Phase Equilib*. 1998;148:107–137.
98. Henstock H. The solubility of phenanthrene in various organic solvents. *J Chem Soc Trans*. 1922;121:2124–2128. DOI: 10.1039/CT9222102124.
99. Potoff JJ, Errington JR, Panagiotopoulos AZ. Molecular simulation of phase equilibria for mixtures of polar and non-polar components. *Mol Phys*. 1999;97:1073–1083.
100. Acree WE, Abraham MH. Solubility predictions for crystalline nonelectrolyte solutes dissolved in organic solvents based upon the Abraham general solvation model. *Can J Chem*. 2001;79:1466–1476.
101. Hernandez CE, Fina KMD, Roy LE, Sharp TL, Acree WE. Solubility of phenanthrene in organic nonelectrolyte solvents. Comparison of observed versus predicted values based upon Mobile Order theory. *Can J Chem*. 1999;77:1465–1470.
102. Rai N, Wagner AJ, Ross RB, Siepmann JI. Application of the TraPPE force field for predicting the Hildebrand solubility parameters of organic solvents and monomer units. *J Chem Theory Comput*. 2008;4:136–144.
103. Modarresi H, Conte E, Abildskov J, Gani R, Crafts P. Model-based calculation of solid solubility for solvent selection—a review. *Ind Eng Chem Res*. 2008;47:5234–5242.
104. Abildskov J, editor. Solubility and Related Properties of Large Complex Chemicals. Part 2: Organic Solutes ranging from C2 to C41. Frankfurt am Main, Germany: DECHEMA, 2005.
105. Choi PB, McLaughlin E. Solubility of aromatic hydrocarbon solids in pyridine and thiophene. *Ind Eng Chem Res*. 1983;22:46–51.

Manuscript received Aug. 15, 2012, and revision received Dec. 5, 2012.