Rapid Prediction of Solvation Free Energy and Vapor Pressure of Liquid and Solid from Molecular Dynamics Simulation

Li Yang
Dept. of Chemical Engineering, National Taiwan University, Taipei 10617, Taiwan

Key Laboratory of Green Chemical Process of Ministry of Education, Key Laboratory of Novel Reactor and
Green Chemical Technology of Hubei Province, School of Chemical Engineering and Pharmacy, Wuhan Institute
of Technology, Wuhan, Hubei 430073, China

Shiang-Tai Lin
Dept. of Chemical Engineering, National Taiwan University, Taipei 10617, Taiwan

DOI 10.1002/aic.14859

We show that the solvation free energy and vapor pressure, important thermodynamic properties of pure substances in liquid or solid states, can be obtained from short, about 20 ps, molecular dynamics simulations. The method combines the determination of free energy of a chemical in vacuum using the normal-mode analysis (NMA, energy minimization), and in the condensed phase using the two-phase thermodynamic (2PT) model. We have examined the calculation results for common liquids and solids, including water, alcohol, acid, aromatics, and alkanes. The results, referred to as 2PT-NMA, is comparable to those calculated from thermodynamic integration (TI) for liquids, and is readily applicable to solids, where simple TI is not applicable. Furthermore, the free energy from 2PT-NMA converges (20 ps) much faster than that from TI (1 ns). The new method could be a very useful tool for fast screening of condensed phase pressure from the trajectory of MD simulations. © 2015 American Institute of Chemical Engineers *AIChE J*, 61: 2298–2306, 2015

Keywords: solvation free energy, vapor pressure, two-phase thermodynamic model, thermodynamic integration, free energy of liquid, free energy of solid

Introduction

The solvation free energy, describing the interactions between a solute and its surrounding solvent, is one of the important properties for research in different areas, such as phase equilibrium,[1] drug design,[2] protein folding,[3] molecule blending,[4] and membrane design[5] etc. There have been many methods developed to determine free energy from molecular simulations, such as the Widom insertion method,[6] free energy perturbation method,[7,8] umbrella sampling method,[9] histogram analysis method,[10] expend ensemble method,[11] potential of mean force method,[12] and thermodynamic integration (TI) method etc.[1,13–17] The solvation free energy is often determined by varying the state of the solute from none interaction to full interaction via tuning a coupling parameter λ base on Krikwood thermodynamic integration equation.[3,18–20] However, in order to remain reversibility, only small incremental changes of λ are allowed. Therefore, the simulation times required in such TI process are often very long, on the order of 1 ns.[17]

In this work, we proposed a new approach to calculate solvation free energy and vapor pressure base on separate determination of Gibbs energy in the vapor and liquid phases, respectively. In the vapor phase, the system is ideal and the free energy can be easily determined from the translation, rotation, and intramolecular vibrations (normal mode analysis [NMA]). In the liquid phase, the free energy is obtained from a short, about 20 ps, trajectory of molecular dynamic simulation based on the two-phase thermodynamic (2PT) model. The 2PT determines the free energy of a fluid based on the vibrational density of states (DOS), obtained from the Fourier transfer of the velocity auto-correlation function. In particular, the anharmonicity of diffusive modes and quantum corrections for the high frequency modes are readily included in the 2PT model. The 2PT model has been successful in the prediction of many physical chemistry properties such as diffusion,[21] heat capacity,[22] absolute entropy,[23] and free energy.[24] The challenge is that the small solvation free energy is obtained from taking the difference of two (often) large Gibbs energies in the vapor and liquid states. We show that such systematic differences between the NMA and 2PT calculations can be minimized by scaling of the vibrational frequencies in the gas phase. Therefore, the resulting approach, referred to as 2PT-NMA(svib), can be used for rapid determination of solvation free energy and vapor pressure from MD simulations.

Theory

The solvation energy and vapor pressure

The solvation free energy is defined as the work needed to transfer a molecule from a fixed position in the ideal gas

---

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

Correspondence concerning this article should be addressed to S.T. Lin at e-mail: stlin@ntu.edu.tw.

phase to another fixed position in solution under constant temperature and pressure

$$\Delta\underline{G}_i^{*\mathrm{sol}}(T,P)=\underline{G}_i^{*,L}(T,P)-\underline{G}_i^{*,\mathrm{IG}}(T,P) \tag{1}$$

where $\underline{G}_i^{*,L}$ and $\underline{G}_i^{*,\mathrm{IG}}$ are the "pseudo" chemical potential of solute i in the solution and ideal gas phase, respectively,[25] with the asterisk in the superscripts denoting the fact that the solute molecule is not allowed to move. It can be shown that the same work can be attained from transferring the molecule from the ideal gas phase having the same temperature and molar volume as the solution, that is[25]

$$\Delta\underline{G}_i^{*\mathrm{sol}}(T,P)=\underline{G}_i^{L}(T,\underline{V}_i^L)-\underline{G}_i^{\mathrm{IG}}(T,\underline{V}_i^L) \tag{2}$$

where $\underline{V}_i^L$ is the molar volume of the solution at temperature T and pressure P. The advantage of this process is that the position of solute molecule is not fixed and the formal chemical potential (rather than the pseudo chemical potential) can be used to determine the solvation free energy.

The vapor pressure of a liquid can be obtained by solving the pressure of the system at vapor-liquid equilibrium[25]

$$\underline{G}_i^{L}(T,P_i^{\mathrm{vap}})=\underline{G}_i^{V}(T,P_i^{\mathrm{vap}}) \tag{3}$$

If the vapor pressure is sufficiently low such that the vapor phase can be assumed ideal, the solvation free energy is related to the vapor pressure as[25]

$$\begin{aligned}\Delta\underline{G}_i^{*\mathrm{sol}}(T,P_i^{\mathrm{vap}})&=\underline{G}_i^{L}(T,\underline{V}_i^L(T,P_i^{\mathrm{vap}}))-\underline{G}_i^{\mathrm{IG}}\left(T,P=\frac{RT}{\underline{V}_i^L}\right)\\&=\underline{G}_i^{L}(T,P_i^{\mathrm{vap}})-\underline{G}_i^{\mathrm{IG}}(T,P_i^{\mathrm{vap}})+\underline{G}_i^{\mathrm{IG}}(T,P_i^{\mathrm{vap}})-\underline{G}_i^{\mathrm{IG}}\left(T,\frac{RT}{\underline{V}_i^L}\right)\\&=RT\ln\frac{P_i^{\mathrm{vap}}\underline{V}_i^L}{RT}\end{aligned} \tag{4}$$

where the first two terms of the second identity cancel because of Eq. 3. As the solvation free energy is not sensitive to small pressure variations, the vapor pressure can be evaluated based on solvation free energy evaluated at any pressure of convenience, that is[25]

$$\ln P_i^{\mathrm{vap}}=\frac{\Delta\underline{G}_i^{*\mathrm{sol}}(T,P)}{RT}+\ln\frac{RT}{\underline{V}_i^L} \tag{5}$$

The two-phase thermodynamic model for Gibbs free energy

The Gibbs free energy can be calculated from the trajectory of molecular dynamic simulation using the 2PT model.[21,22,24,26,27] The vibrational DOS function $S(\upsilon)$ is defined as the distribution of vibrational normal modes in the system. The number of modes of a system at some frequency υ is calculated as the sum of contributions from all atoms in the system[21]

$$S(\upsilon)=\frac{2}{kT}\sum_{j=1}^{N}\sum_{k=1}^{3}m_j s_j^k(\upsilon) \tag{6}$$

where m_j is the mass of atom j, N is the total number of atoms. The velocity spectral density $s^k_j(\upsilon)$ of atom j in the kth coordinate ($k = x$, y, and z in the Cartesian coordinate) is determined from the square of the Fourier transform of the velocities as[21]

$$s_j^k(\upsilon)=\lim_{\tau\to\infty}\frac{1}{2\tau}\left|\int_{-\tau}^{\tau}v_j^k(t)e^{-i2\pi\upsilon t}\mathrm{d}t\right|^2 \tag{7}$$

where $v^k_j(t)$ is the kth velocity component of atom j at time t. The DOS function can also be obtained from the Fourier transform of the velocity autocorrelation function.[28] Several important features of DOS should be noted. First of all, the integration of S(υ) over the whole frequency domain gives the total degree of freedom (3N) of the system. In other words, S(υ)dυ is the number of normal modes with frequencies between υ and $\upsilon + \mathrm{d}\upsilon$. Furthermore, the zero frequency intensity, S(0), is related to the diffusivity (motion of zero frequency) of the system. For system of polyatomic molecules the $S(\upsilon)$ is decomposed into three components, translation (trn), rotation (rot), and intramolecular vibration motions (vib)

$$S(\upsilon)=S_{\mathrm{trn}}(\upsilon)+S_{\mathrm{rot}}(\upsilon)+S_{\mathrm{vib}}(\upsilon) \tag{8}$$

Each DOS component is determined from the center of mass velocity, angular velocity, and vibrational velocities.[24]

The essence of the 2PT model is the decomposition of the DOS of the system with 3N degrees of freedom to a gas-like and a solid-like portion[21]

$$S_{\mathrm{m}}(\upsilon)=S_{\mathrm{m}}^g(\upsilon)+S_{\mathrm{m}}^s(\upsilon) \tag{9}$$

where the subscript m can be translation (*trn*), rotation (*rot*), or vibration (*vib*). For internal vibration, the gas-like component is zero. For translation and rotation, the hard sphere model is used to describe the gas-like diffusive component,[29] aiming to take up all the diffusive and anharmonic modes

$$S_{\mathrm{m}}^g(\upsilon)=\frac{S_{\mathrm{m}}(0)}{1+\left(\frac{\pi S_{\mathrm{m}}(0)\upsilon}{6fN}\right)^2} \tag{10}$$

The fluidicity factor f, whose value falls between 0 and 1, represents the fraction of gas-like component, and is determined from the dimensionless diffusivity constant Δ_{m} as

$$2\Delta_{\mathrm{m}}^{-9/2}f^{15/2}-6\Delta_{\mathrm{m}}^{-3}f^5-\Delta_{\mathrm{m}}^{-3/2}f^{7/2}+6\Delta_{\mathrm{m}}^{-3/2}f^{5/2}+2f-2=0 \tag{11}$$

with Δ_{m} determined as a function of material properties[21]

$$\Delta_{\mathrm{m}}(T,\rho,m,s_0)=\frac{2S_{\mathrm{m}}(0)}{9N}\left(\frac{\pi kT}{m}\right)^{1/2}\rho^{1/3}\left(\frac{6}{\pi}\right)^{2/3} \tag{12}$$

Once the gas-like component is determined, the solid-like component can be obtained from the difference of the total DoS and the gas-like DoS.

The thermodynamic properties P_{m} (energy E or entropy S) associated with molecular motion m is determined from the weighting the individual DOS component with proper functions[21]

$$P_{\mathrm{m}}=\int_0^\infty d\upsilon\ S_{\mathrm{m}}^S(\upsilon)W_{\mathrm{P,m}}^S(\upsilon)+\int_0^\infty d\upsilon S_{\mathrm{m}}^g(\upsilon)W_{\mathrm{P,m}}^g(\upsilon) \tag{13}$$

where $W_{\mathrm{P,m}}^S(\upsilon)$ and $W_{\mathrm{P,m}}^g(\upsilon)$ are the weighting functions of a harmonic oscillator and the corresponding gas part of component, respectively. For the absolute entropy and energy, they are

$$W_{\mathrm{E}}^{s}(\upsilon)=\frac{h\upsilon}{2}+\frac{h\upsilon}{\exp(\beta h\upsilon)-1} \tag{14}$$

$$W_{\mathrm{S}}^{\mathrm{QHO}}(\upsilon)=\frac{k_{\mathrm{b}}\beta h\upsilon}{\exp(\beta h\upsilon)-1}-k_{\mathrm{b}}\ln[1-\exp(-\beta h\upsilon)] \tag{15}$$

$$W_{\mathrm{E,trn}}^{g}(\upsilon)=W^{g}{}_{\mathrm{E,rot}}(\upsilon)=0.5\beta^{-1} \tag{16}$$

$$W_{\mathrm{S,trn}}^{g}(\upsilon)=\frac{S^{\mathrm{HS}}}{3} \tag{17}$$

$$W_{\mathrm{S,trn}}^{g}(\upsilon)=\frac{S^{\mathrm{FR}}}{3} \tag{18}$$

where $\beta=1/k_bT$, S^{HS} and S^{FR} are the hard-sphere entropy and rotational entropy of molecules at the ideal gas state, respectively

$$\frac{S^{\mathrm{HS}}}{R}=\frac{5}{2}+\ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\frac{V}{f_{\mathrm{trn}}N}z(y)\right]+\frac{y(3y-4)}{(1-y)^2} \tag{19}$$

$$\begin{cases}\dfrac{S^{\mathrm{FR}}}{R}=\ln\left(\dfrac{T}{\sigma\Theta_{\mathrm{r}}}\right)+1 & \text{(linear molecule)}\\[2ex] \dfrac{S^{\mathrm{FR}}}{R}=\ln\left(\dfrac{\pi^{1/2}T^{3/2}}{\sigma_{\mathrm{r}}\left(\Theta_{\mathrm{r},x}\Theta_{\mathrm{r},y}\Theta_{\mathrm{r},z}\right)^{1/2}}\right)+\dfrac{3}{2} & \text{(nonlinear molecule)}\end{cases} \tag{20}$$

where $y=f_{\mathrm{trn}}{}^{5/2}/\Delta_{\mathrm{trm}}{}^{3}$ and $z(y)$ is the compressibility factor of hard sphere gases from the Carnahan-Starling equation of state,[30] Θ is the rotation temperature, and σ is the rotational symmetry number. The thermodynamic properties are determined from the sum of translation, rotation, and intramolecular vibration contributions

$$E=E_0+E_{\mathrm{trn}}+E_{\mathrm{rot}}+E_{\mathrm{vib}} \tag{21}$$

$$S=S_{\mathrm{trn}}+S_{\mathrm{rot}}+S_{\mathrm{vib}} \tag{22}$$

where the reference energy form E_0 is described as

$$E_0=E^{\mathrm{MD}}-\beta^{-1}3N\left(1-0.5f_{\mathrm{trn}}-0.5f_{\mathrm{rot}}\right) \tag{23}$$

The reference energy is obtained by comparing the MD potential energy to the 2PT energy based on classical harmonic oscillators.[21] The Helmholtz energy and Gibbs energy can be calculated through thermodynamic relationship equation as following

$$A=E-ST \tag{24}$$

$$G=A+PV \tag{25}$$

where P and V are total pressure and volume of system, respectively.

The normal mode analysis for Gibbs free energy of an ideal gas

Although the 2PT method is applicable to all types of state of aggregation of a chemical, the DOS converges much slowly in the vapor phase as enough collisions would be necessary.[24,26] A more efficient approach to obtain the ideal gas free energy is from the NMA,[29] where the property Q of an ideal gas is calculated with contributions from molecular translation, rotation, and vibration as in Eqs. 21 and 22. For energy, they are[29]

$$E_{\mathrm{trn}}=\frac{3RT}{2} \tag{26}$$

$$E_{\mathrm{rot}}=\frac{\delta RT}{2}\ \text{with}\ \delta=2\ \text{for linear molecules and 3 for nonlinear molecules} \tag{27}$$

$$E_{\mathrm{vib}}=RT\sum_{j=1}^{3n-3-\delta}\left[\left(\frac{\Theta_{\upsilon,j}}{2T}\right)+\frac{\Theta_{\upsilon,j}/T}{e^{\Theta_{\upsilon,j}/T}-1}\right] \tag{28}$$

And for entropy

$$S_{\mathrm{trn}}=R\left\{\ln\left[\left(\frac{2\pi mk_{\mathrm{B}}T}{h^2}\right)^{3/2}V\right]+1+\frac{3}{2}\right\} \tag{29}$$

$$S_{\mathrm{vib}}=R\sum_{j=1}^{3n-3-\delta}\left[\frac{\Theta_{\upsilon,j}/T}{e^{\Theta_{\upsilon,j}/T}-1}-\ln\left(1-e^{-\Theta_{\upsilon,j}/T}\right)\right] \tag{30}$$

S_{rot} is the same as Eq. 20.

The reference value for entropy is zero. For energy, E_0 is potential energy is calculated based on force field at reference temperature during energy minimization (0 K). In order to match the reference energy E_0 from 2PT (Eq. 21), the following equation is used

$$E_0=E_{\mathrm{opt}}(0\ \mathrm{K})+E_{\mathrm{trn}}+E_{\mathrm{rot}}=E_{\mathrm{opt}}(0\ \mathrm{K})+\frac{3}{2}RT+\frac{3}{2}RT \tag{31}$$

The Helmholtz and Gibbs energy are calculated the same manner as in Eq. 24 and 25.

Computational Details

The solvation free energies of 18 molecules are used to validate the proposed method. Molecular dynamic simulations are performed using the GROMACS 4.5.7 package.[23,31] All organic molecules are described by the optimized potentials for liquid simulations for all atoms (OPLS-AA) force field[32,33] which has been widely used for pure liquid simulation of common chemicals. The TraPPE united-atom force field[34] was also employed for alkane molecules. The non-bonded interaction or van der Waals interaction of the atoms was carried in Lennard-Jones (LJ) potential and electrostatic potential

$$V_{ij}=4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\frac{q_iq_j}{4\pi\varepsilon_0r_{ij}} \tag{32}$$

where the V_{ij} is the sum of potential energy, the first term on the right of Eq. 36 is LJ formula, the second term on the right of Eq. 36 is Columbic interaction. A spherical cutoff of 12 Å was both used to evaluate van der Waals and Columbic interactions. The calculation of LJ potential is corrected using Switch algorithm. The long range Columbic interaction was calculated by PME (fast Particle-Mesh Ewald) method.[23] Three types of intramolecular interactions were employed for calculation of intramolecular potential including bond, angle and dihedral torsions, respectively.[32–34] The water molecule was described by three rigid water models, SPC, TIP3P, and TIP4P. The bonds and angles are fixed in special values using the SETTLE algorithm.[22]

The liquid phase simulations were conducted using cubic boxes containing 500 molecules, with the initial density set to that of pure solvent. Energy minimization was then performed with the box volume allowed to vary in order to relax extra-stresses in the system. Then a 10-ns simulation under constant temperature and pressure (NPT) were conducted to equilibrate the system. The temperature was controlled at desired temperature using the Nose-Hoover extended ensemble,[31] and the Berendsen barostat method[31] was employed to stabilize

Table 1. The Solvation Free Energy Determined from 2PT and TI Methods at 298.15 K and 1 bar

Compounds	ΔG^{*sol} (2PT-NMA) (kJ/mol)	ΔG^{*sol} (2PT-NMA(novib)) (kJ/mol)	ΔG^{*sol} (2PT-NMA(svib)) (kJ/mol)	SD (kJ/mol)	ΔG^{*sol} (TI) (kJ/mol)
OPLS-AA (all-atoms)					
Water(spc)	−28.58	−28.58	−28.58	±0.03	−27.01 ± 0.11
Water(tip3p)	−28.42	−28.42	−28.42	±0.03	−26.28 ± 0.03
Water(tip4p)	−27.51	−27.51	−27.51	±0.04	−25.72 ± 0.09
Methane[a]	−5.42	−4.84	−7.35	±0.64	−5.36 ± 0.01
Methanol	−22.65	−24.94	−25.36	±0.83	−18.48 ± 0.26
Acetone	−18.05	−19.27	−22.61	±0.79	−16.05 ± 0.09
Acetic acid	−29.64	−32.62	−32.25	±0.76	−29.07 ± 0.88
Acetonitrile	−20.84	−20.75	−22.93	±0.63	−17.54 ± 0.09
Benzene	−18.45	−18.40	−24.21	±0.32	−17.87 ± 0.09
Chloroform	−12.80	−14.89	−14.03	±0.71	−16.61 ± 0.20
Ethanol	−18.02	−21.45	−22.33	±0.31	−19.39 ± 0.33
1,4-Dioxane	−12.42	−15.52	−18.53	±0.94	−18.06 ± 0.23
Furan	−15.25	−14.25	−19.05	±0.30	−15.60 ± 0.09
THF	−14.23	−13.86	−19.96	±0.57	−16.51 ± 0.37
Toluene	−10.33	−18.54	−17.79	±0.79	−21.10 ± 0.35
Butane	−7.03	−10.66	−14.17	±0.50	−10.62 ± 0.10
Pentane	−4.76	−10.36	−13.63	±0.76	−13.62 ± 0.19
Hexane	−4.21	−11.88	−14.97	±0.24	−15.96 ± 0.19
Heptane	−4.33	−14.11	−17.25	±0.40	−19.13 ± 0.46
Octane	−3.74	−16.26	−18.86	±0.12	−21.73 ± 0.50
TraPPE (united-atoms)					
Butane	−16.52	−15.49	−16.54	±0.64	−10.84 ± 0.10
Pentane	−16.93	−15.51	−17.10	±0.35	−13.16 ± 0.18
Hexane	−17.47	−16.06	−17.52	±0.55	−16.04 ± 0.29
Heptane	−18.47	−17.11	−18.65	±0.36	−18.54 ± 0.10
Octane	−18.84	−17.95	−19.26	±0.62	−20.84 ± 0.15
R^2	0.41	0.77	0.73		—
RMSD	6.44	3.06	3.47		—

[a] Methane simulations were conducted at 111.65 K and 1 bar.
2PT-NMA $= G_L(\text{2PT}) - G_V(\text{NMA})$.
2PT-NMA(novib): without including contributions from the intramolecular vibration.
2PT-NMA(svib): the vibrational frequencies from NMA are scaled as given in Eq. 35.
SD: standard deviation from 2PT.
R^2: correlation coefficient against results from TI.
RMSD: root mean square deviation against results from TI.

pressure at 1 atm. The equilibrated system serves as the initial state of the subsequent simulations for 2PT or TI analysis. The solid phase simulations were conducted using triclinic or orthorhombic boxes based upon the crystallographic data. All crystal structures of solid phase were downloaded from Cambridge Structural Database.[35] The process of simulation was employed the same way as for the liquid phase simulations.

For 2PT analysis additional 20 ps NVT simulation was performed with the coordinates and velocities saved every 2 fs. Five independent simulations were performed in order to obtain the statistical uncertainty.

For the TI analysis, 10 ns NPT simulation was used to slowly turn on the LJ interaction (with the coupling parameter λ increased from 0 [start point] to 1 [end point]). Additional 10 ns simulation followed to incorporate the Columbic interactions (with the coupling parameter λ increased from 0 [start point] to 1 [end point]). The temperature was controlled at desired value using the Langevin stochastic dynamic method and the Berendsen barostat method was used to stabilize pressures at 1 atm. The detailed simulation results from TI are provided in the supporting information.

Results and Discussion

Solvation free energy and vapor pressure from 2PT model for fluids

Table 1 summarizes the calculated solvation free energy from 2PT and TI. Although the solvation free energies from 2PT-NMA for most chemicals are consistent with those from the TI method, the results for alkanes are obviously erroneous both in the absolute values and the change of solvation free energy with the number of carbon atoms: the solvation free energy becomes more negative in TI as the alkane chain length increases. This results in a low overall correlation coefficient R^2 between 2PT-NMA and TI of 0.41.

An analysis regarding the free energy components helps to understand the source of problem. Table 2 shows the five solvation free energy components from reference energy (ΔE_0), translation (ΔA^{*sol} (trn)), rotation (ΔA^{*sol} (rot)), vibrational (ΔA^{*sol} (vib)) motions, and the pressure-volume term (ΔPV). Note that ΔG^{*sol} (2PT-NMA) $= \Delta E_0 + \Delta A^{*sol}$ (trn) $+ \Delta A^{*sol}$ (rot) $+ \Delta A^{*sol}$ (vib) $+ \Delta$PV. It can be seen that the solvation energy is dominated by the large negative contributions from ΔE_0, with positive corrections from ΔA^{*sol} (trn) and ΔA^{*sol} (rot). The internal vibration ΔA^{*sol} (vib) is usually small, except for a few cases, such as tetrahydrofuran (THF) (−62.68 kJ/mol) and acetic acid (−22.43 kJ/mol). In the case of THF, the contribution from ΔE_0 becomes positive. Figure 1 illustrates the vibrational DOS from 2PT and NMA for THF. It is quite clear that there is a significant discrepancy in the two methods in the high frequency modes (>2500 cm^{-1}). The large difference in the vibration frequencies in the gas phase and in liquid implies that a very different valence potential energy and zero point energy in these two states, such a large different results in a positive value of ΔE_0 and a large negative of ΔA^{*sol} (vib) for THF. One

Table 2. The Solvation Free Energy Components from the 2PT-NMA Method

Compounds	ΔE_0 (kJ/mol)	ΔE_0 (novib) (kJ/mol)	$\Delta A^{*\text{sol}}$ (trn) (kJ/mol)	$\Delta A^{*\text{sol}}$ (rot) (kJ/mol)	$\Delta A^{*\text{sol}}$ (vib) (kJ/mol)	$\Delta A^{*\text{sol}}$ (svib) (kJ/mol)	ΔPV (kJ/mol)
OPLS-AA (all-atom)							
Water(spc)	−62.75	−62.75	16.14	20.51	0.00	0.00	−2.48
Water(tip3p)	−60.90	−60.90	15.21	19.74	0.00	0.00	−2.48
Water(tip4p)	−62.56	−62.56	16.43	21.09	0.00	0.00	−2.48
Methane	−16.10	−15.30	7.46	3.92	0.23	−1.71	−0.93
Methanol	−55.44	−55.47	17.15	15.86	2.26	−0.45	−2.47
Acetone	−38.64	−49.19	16.30	16.10	−9.34	−13.89	−2.47
Acetic acid	−48.05	−73.47	20.89	22.42	−22.43	−25.04	−2.47
Acetonitrile	−47.85	−47.90	15.58	14.04	−0.15	−2.24	−2.47
Benzene	−45.92	−52.00	18.66	17.41	−6.15	−11.90	−2.47
Chloroform	−47.89	−47.24	18.25	16.56	2.74	1.52	−2.47
Ethanol	−38.08	−59.97	19.56	21.43	−18.46	−22.77	−2.47
1,4-Dioxane	−42.30	−51.72	18.86	19.80	−6.32	−12.42	−2.47
Furan	−32.55	−47.47	16.53	19.16	−15.92	−19.72	−2.47
Tetrahydrofuran	14.75	−47.55	17.87	18.29	−62.68	−68.41	−2.47
Toluene	−49.73	−59.14	19.51	23.55	−1.18	−8.65	−2.47
Butane	−31.29	−37.74	14.83	14.72	−2.82	−9.95	−2.47
Pentane	−33.21	−42.31	16.63	17.78	−3.50	−12.37	−2.47
Hexane	−36.53	−47.23	18.01	19.81	−3.04	−13.81	−2.47
Heptane	−39.87	−52.32	19.12	21.55	−2.67	−15.59	−2.46
Octane	−43.13	−57.16	20.24	23.12	−1.51	−16.63	−2.46
TraPPE (united-atom)							
Butane	−35.12	−35.11	12.74	9.35	−1.02	−1.04	−2.47
Pentane	−39.33	−39.12	13.83	12.25	−1.20	−1.37	−2.47
Hexane	−43.51	−43.03	15.13	14.30	−0.92	−0.97	−2.47
Heptane	−47.49	−46.77	16.14	15.98	−0.63	−0.81	−2.46
Octane	−51.71	−50.56	17.38	17.70	−0.25	−0.17	−2.46

$\Delta G^{*\text{sol}}$ (2PT-NMA) = $\Delta E_0 + \Delta A^{*\text{sol}}$ (trn) $+\Delta A^{*\text{sol}}$ (rot) $+\Delta A^{*\text{sol}}$ (vib) $+\Delta PV$
$\Delta G^{*\text{sol}}$ (2PT-NMA(novib)) = ΔE_0(novib) $+ \Delta A^{*\text{sol}}$ (trn) $+\Delta A^{*\text{sol}}$ (rot) $+\Delta PV$
$\Delta G^{*\text{sol}}$ (2PT-NMA(svib)) = $\Delta E_0 + \Delta A^{*\text{sol}}$ (trn) $+\Delta A^{*\text{sol}}$ (rot) $+\Delta A^{*\text{sol}}$ (svib) $+\Delta PV$

possible cause of the shift of vibrational frequencies is the use of 0 K molecular structure in NMA, whereas the liquid phase is equilibrated at 298 K. Figure 2 shows the DOS from 2PT and NMA for octane. It can be seen that the shifts of the high frequency modes are similar as in the case of THF. We thus performed additional simulations for alkanes using the united atom force field, TraPPE, where the high frequency bond stretching modes do not appear. As can be seen in Figure 2b, the high frequency modes (>2500 cm^{-1}) associated with CH stretching are not present in the DOS. Furthermore, correct dependence of solvation free energy on carbon number can be retained and good agreement between 2PT and TI can be achieved, as seen in Table 1.

Note that in TI the contributions from internal vibrations are completely ignored. In the case of rigid waters, the agreements between 2PT-NMA and TI are pretty good. It is thus intriguing to see how the calculation changes without consideration of the internal vibration modes. The $\Delta G^{*\text{sol}}$ (2PT-NMA(novib)) in Table 1 (third column) are results without considering internal vibrations in both 2PT and NMA analysis. In other words, the solvation free energy is calculated as $\Delta G^{*\text{sol}}$ (2PT-NMA(novib)) = ΔE_0(novib) $+ \Delta A^{*\text{sol}}$ (trn) $+\Delta A^{*\text{sol}}$ (rot) $+\Delta$PV given in Table 2. The reference energy needs to be modified (adding back MD energy of intramolecular vibration) in order to match the reference state in 2PT and NMA. The agreement between 2PT-NMA and TI improves significantly when the internal vibrations are ignored, with the correlation coefficient R^2 improves to 0.77 (root mean square deviation [RMSD] error improves from about 6.44 kJ/mol to 3.06 kJ/mol), as shown in Figure 3.

While in practice the internal vibrations should vary with the chemical environment, the complete neglect of such contributions in TI and 2PT-NMA(novib) is not strictly correct. Since the high frequency modes should be less affected by the environment where a solute molecule is situated, we suggest a different treatment that the vibrational frequencies from the NMA can be rescaled by a constant factor κ determined from the highest frequency from both calculations as

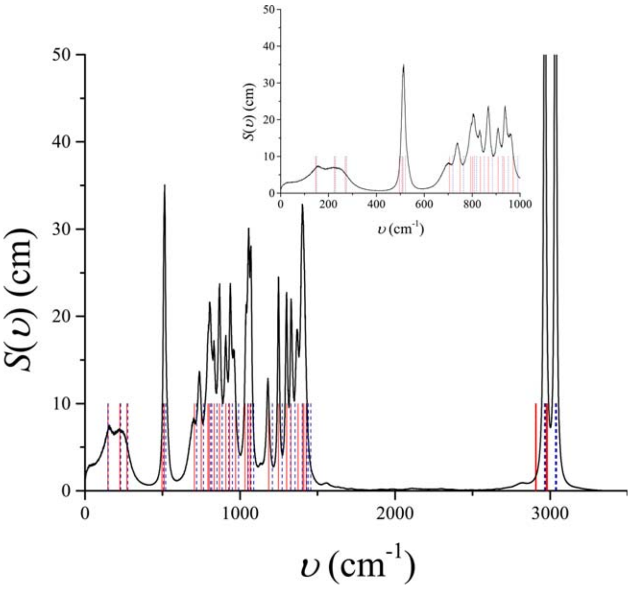

Figure 1. The vibrational DOSs of tetrahydrofuran from 2PT (black line), NMA (red line), and NMA with the frequencies rescaled (dotted blue line).

The inset figure emphasizes the data in the low frequency region. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

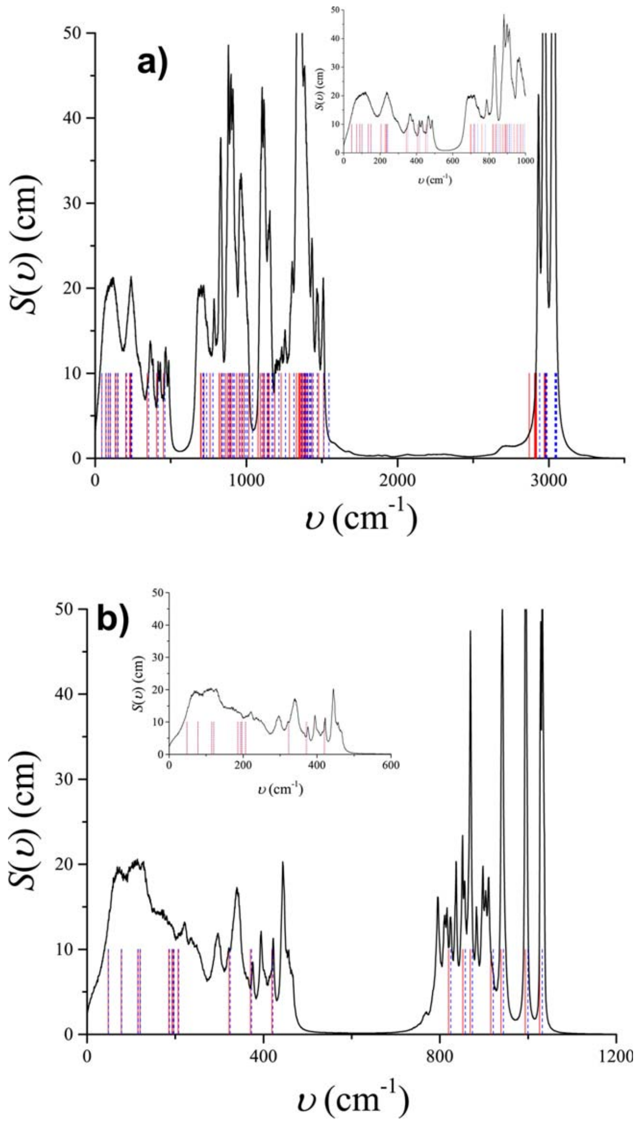

Figure 2. The vibrational DOSs of octane, using all-atom model (a) and united-atom model (b), from 2PT (black line), NMA (red line), and NMA with the frequencies rescaled (dotted blue line).

The inset figure emphasizes the data in the low frequency region. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\kappa = \frac{f^{\mathrm{2PT}}_{\mathrm{highest}}}{f^{\mathrm{NMA}}_{\mathrm{old,highest}}} \tag{33}$$

The value of κ is found to range from 1.01 to 1.04 for the 18 molecules studied here. The NMA frequencies are then scaled as

$$f^{\mathrm{NMA}}_{\mathrm{new},i} = \kappa f^{\mathrm{NMA}}_{\mathrm{old},i} \tag{34}$$

The modified vibrational frequencies of NMA model have a better matching with 2PT model as the blue dash line shown on Figures 1 and 2. The solvation free energy obtained from this method (denoted as 2PT-NMA(svib)) still gives much improved agreement with the TI results (RMSD error reduces to 3.47 kJ/mol, and correlation coefficient $R^2 = 0.73$), as shown in Figure 3. The rescaling of internal vibration frequency also provides a correct solvation free energy dependency on alkane chain length for both all atom and united atom force fields. In other words, 2PT-NMA(svib) may minimize the systematic error caused by the used of different simulation methods in the liquid and vapor phases and produces satisfactory accuracy for the solvation free energy for a variety of species. It is noteworthy that an exact agreement between TI and 2PT-NMA should not be expected as the latter approach includes quantum corrections by applying quantum harmonic statistics (Eq. 14, 15, 29, 31), whereas TI produces purely classical results.

The vapor pressure determined from the calculated solvation free energy (Eq. 5) are summarized in Table 3 and compared in Figure 4. The RMSD of calculated vapor pressure from 2PT-NMA, 2PT-NMA(novib), 2PT-NMA(svib), and TI are 2.75, 1.43, 1.38, and 0.61 (unit is ln(pa)), respectively. The results from TI best agree with experiment[36] perhaps because that the classical force fields were developed with inclusion of phase transition data.[27] The larger errors from 2PT-NMA methods can be expected since the force fields were developed without consideration of quantum effects, which is important for proper description of other properties such as enthalpy and heat capacities.[37]

Solvation free energy and vapor pressure from 2PT model for solid crystals

One prevailing advantage of 2PT-NMA method over TI is the determination of properties of solid crystals, where quantum effects are significant. To determine the free energy of a solid with TI, one often need to choose a solid model (e.g., Einstein crystal) to be the reference state, and then device a TI path to transform the ideal crystal to the system of interest.[38,39] Recently, Aragones et al.[40] demonstrated the evaluation of free energy of solids with TI without consideration of intramolecular vibrations (i.e., rigid molecules). To evaluate the solvation free energy or sublimation pressure, one would also need to change the reference state the ideal gas phase. In contrast, the 2PT-NMA method, which inherently includes the quantum corrections to the calculated properties, can

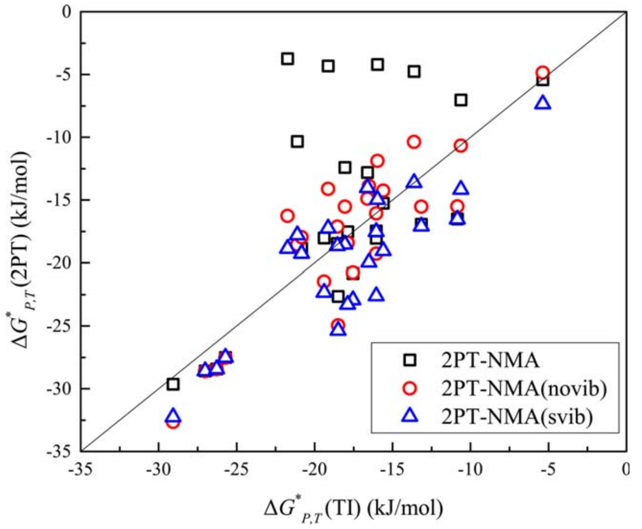

Figure 3. The comparison of solvation free energies determined from 2PT-NMA and TI.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 3. The Vapor Pressure Determined from 2PT and TI Methods at 298.15 K

Compounds	$\ln P^{cal}$(Pa)				
	2PT-NMA	2PT-NMA (novib)	2PT-NMA (svib)	TI	Exp[36]
OPLS-AA (all-atoms)					
Water(spc)	7.2	7.2	7.2	7.8	8.1
Water(tip3p)	7.3	7.3	7.3	8.1	8.1
Water(tip4p)	7.6	7.6	7.6	8.3	8.1
Methane[a]	11.4	12.0	9.3	11.3	11.5
Methanol	8.8	7.9	7.7	10.6	9.7
Acetone	10.1	9.6	8.2	10.8	10.3
Acetic acid	5.6	4.4	4.6	5.9	7.6
Acetonitrile	9.2	9.3	8.4	10.5	9.4
Benzene	10.1	9.7	7.8	9.9	9.4
Chloroform	12.1	11.3	11.6	10.5	10.2
Ethanol	10.3	8.9	8.6	9.7	9.0
1,4-Dioxane	12.1	10.9	9.7	9.8	8.5
Furan	11.2	11.6	9.7	11.1	11.3
THF	11.5	11.6	9.2	10.5	10.0
Toluene	12.8	9.5	9.8	8.5	8.2
Butane	14.2	12.7	11.3	12.7	12.4
Pentane	15.0	12.7	11.4	11.4	11.1
Hexane	15.1	12.0	10.7	10.3	9.9
Heptane	14.9	11.0	9.7	8.9	8.7
Octane	15.0	10.0	8.9	7.8	7.5
TraPPE (united-atoms)					
Butane	10.3	10.8	10.3	12.6	12.4
Pentane	10.1	10.6	10.0	11.5	11.1
Hexane	9.7	10.3	9.7	10.3	9.9
Heptane	9.2	9.7	9.1	9.1	8.7
Octane	8.9	9.3	8.8	8.1	7.5
RMSD	2.75	1.43	1.38	0.61	

[a]Vapor pressure of methane is determined at 111.65 K.

readily provide the properties of solids without devising any additional simulations.

In this work, several typical compounds were chosen to calculate their solid solvation free energy below their melting temperatures by about 50 K. The use of a lower temperature is to ensure a stable solid phase using the OPLS force field. Every crystal configuration contained about 250 molecules to 640 molecules base on the crystalline unit cell obtained from the CCDC database.[35] The calculated solvation free energy and vapor pressure of solid were shown in Table 4. The solvation energy is dominated by the large negative contributions from ΔE_0 with a small negative contributions form ΔA^{*sol} (svib) and ΔPV and positive corrections from ΔA^{*sol} (trn) and ΔA^{*sol} (rot). The predicted sublimation pressure of solid agrees reasonably with experiments. Note that the some of the experimental values[36] (denoted by symbol "†" in Table 4) were extrapolated from data at higher temperatures. The results show that the 2PT-NMA method is useful for estimating the sublimation pressure of solids, whose values are often very small and difficult to be measured accurately by experiment.

Comparison of merits between 2PT and TI methods

One important merit of the proposed 2PT-NMA method is its efficiency. Figure 5 illustrates the solvation free energies determined from different lengths of a MD trajectory for octane. It can be seen that the 2PT-NMA results are converged within 20ps. In contrast, TI requires at least 200 ps to achieve convergence. TI usually requires a long simulation time to allow slow changes in the coupling parameter λ. The simulation time would increase significantly especially for large sized molecules. The convergence of 2PT is not much affected by the size of molecules. On other hand, while TI is convenient for obtaining free energy in the liquid phase, additional modifications (as discussed previously) are necessary for the solid phase. The 2PT-NMA method is readily applicable to all condensed phases because the 2PT method includes both diffusive effects (important for liquids) and quantum corrections (important for solids). Therefore, there is a great potential for its application in different scientific areas.

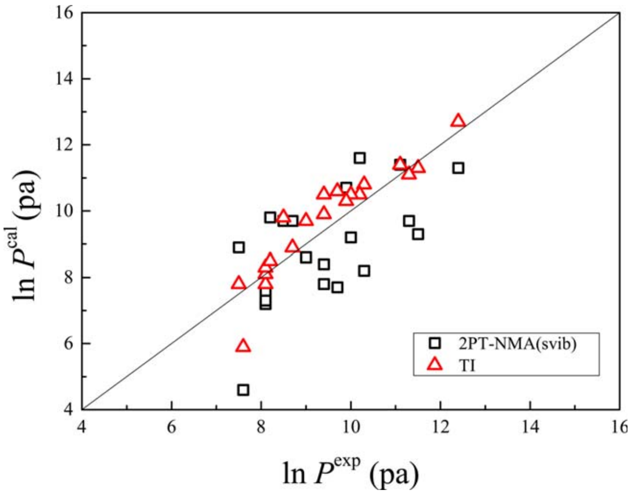

Figure 4. The comparison of vapor pressure predictions from 2PT-NMA(svib) and TI against experimental data.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 4. The Solvation Free Energy and Vapor Pressure Determined from 2PT-NMA(svib) for Solid Crystals

Compounds	T (K)	ΔE_0 (kJ/mol)	$\Delta A^{*\text{sol}}$ (trn) (kJ/mol)	$\Delta A^{*\text{sol}}$ (rot) (kJ/mol)	$\Delta A^{*\text{sol}}$ (svib) (kJ/mol)	ΔPV (kJ/mol)	$\Delta G^{*\text{sol}}$ (2PT-NMA(svib)) (kJ/mol)	SD (kJ/mol)	P^{vap}(2PT) (Pa)	P^{vap}(exp) (Pa)
Acetic acid	240.15	−50.02	17.97	18.10	−18.80	−1.99	−34.75	±0.96	1.14141	4.26550
Acetonitrile	178	−48.93	12.45	11.47	−2.70	−1.48	−29.19	±0.55	0.10271	0.17867
Benzene	228.64	−48.98	16.77	13.52	−9.52	−1.89	−30.10	±1.56	3.38859	61.93800
Ethanol	108.8	−36.69	9.34	12.48	−15.02	−0.90	−30.79	±0.99	3.28730E-08	8.7985E-12[a]
Butane	86	−35.90	7.76	9.10	−3.09	−0.71	−22.83	±0.94	1.46688E-07	3.1715E-08[a]
1,4-Dioxane	210	−45.02	16.26	17.94	−11.01	−1.74	−23.56	±1.96	32.08876	1.08206[a]

SD: standard deviation from 2PT methods

[a]The experiments extrapolated from data at higher temperatures

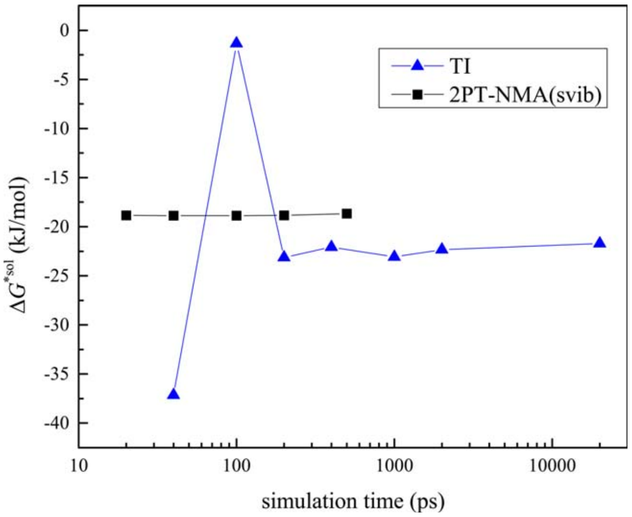

Figure 5. The convergence of solvation free energy from 2PT-NMA(svib) and TI based on different lengths of simulation trajectory.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Conclusion

In this research, we propose a new method to determine the solvation free energy from the results of MD simulations. This new method combines the Gibbs free energy determined from NMA in ideal gas phase and 2PT analysis in the condensed phase. The systematic error from taking the free energy difference between two methods are minimized by rescaling the internal vibration frequencies based on the highest vibration frequency in the real phase. The resulting method, 2PT-NMA(svib) produces solvation free energies that correlates reasonably with those from TI in fluid phase. The advantages of the proposed approach are to include quantum effects (zero point energy, harmonic vibrations) in the solvation free energies. Because of the inclusion of quantum corrections, this method can be readily employed for the calculation of solvation free energy in solid phase. Therefore, the proposed method has the potential applications for chemicals in all kinds of states, liquid, solid, glassy or rubbery states. The method produces converged free energy within about 20 ps simulation trajectory, which is orders of magnitude faster than that of TI. This opens up opportunities of studying solvation driving forces using ab initio MD or systems involving macromolecules.

Acknowledgments

This research was partially supported by the National Science Council of Taiwan (101–2628-E-002–014-MY3 and 102–3113-P-002-01) and Ministry of Education of Taiwan (102R4000 and 103R4000). Dr. Li Yang would also like to acknowledge the support from the National Natural Science Foundation of China (No. 21406172). The computation resources from the National Center for High-Performance Computing of Taiwan and the Computing and Information Networking Center of the National Taiwan University are acknowledged.

Literature Cited

1. 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(6):1929–1938.

2. Reddy MR, Erion MD. *Free Energy Calculations in Rational Drug Design*. New York: Kluwer Academic/Plenum Publishers, 2001.
3. Beveridge DL, Dicapua FM. Free energy via molecular simulation: Application to chemical and biomolecular systems. *Annu Rev Biophys Biophys Chem*. 1989;18:431–492.
4. Johansson ACV, Lindahl E. Position-resolved free energy of solvation for amino acids in lipid membranes from molecular dynamics simulations. *Proteins: Struct Funct Bioinf*. 2008;70(4):1332–1344.
5. Tieleman DP, MacCallum JL, Ash WL, Kandt C, Xu ZT, Monticelli L. Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid-protein interactions, side chain transfer free energies and model proteins. *J Phys Condens Matter*. 2006; 18(28):S1221–S1234.
6. Widom B. Some topics in the theory of fluids. *J Chem Phys*. 1963; 39(11):2808–2812.
7. Singh UC, Brown FK, Bash PA, Kollman PA. An approach to the application of free energy perturbation methods using molecular dynamics: applications to the transformations of methanol –> ethane, oxonium –> ammonium, glycine –> alanine, and alanine –> phenylalanine in aqueous solution and to H3O+(H2O)3 –> NH4+(H2O)3 in the gas phase. *J Am Chem Soc*. 1987;109(6):1607–1614.
8. Bash PA, Field MJ, Karplus M. Free energy perturbation method for chemical reactions in the condensed phase: a dynamic approach based on a combined quantum and molecular mechanics potential. *J Am Chem Soc*. 1987;109(26):8092–8094.
9. Souaille M, Roux Bt. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun*. 2001;135(1):40–57.
10. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The weighted histogram analysis method for free-energy calculations on biomolecules. *I. The method*. *J Comput Chem*. 1992;13(8):1011–1021.
11. 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(3):1776–1783.
12. Roux B. The calculation of the potential of mean force using computer simulations. *Comput Phys Commun*. 1995;91(1–3):275–282.
13. Allen MP, Tildesley DJ. *Computer Simulation of Liquids*. New York: Oxford University Press, 1987.
14. Hockney R, Eastwood J. *Computer-Simulation using Particles*. New York: Adam Hilger, 1989.
15. Frenkel D, Smit B. *Understanding Molecular Simulation from Algorithms to Applications*. San Diego: Academic Press, 2002.
16. Chang J, Lenhoff AM, Sandler SI. Solvation free energy of amino acids and side-chain analogues. *J Phys Chem B*. 2007;111(8):2098–2106.
17. Yang L, Ahmed A, Sandler SI. Comparison of two simulation methods to compute solvation free energies and partition coefficients. *J Comput Chem*. 2013;34(4):284–293.
18. Kirkwood JG. Statistical mechanics of fluid mixtures. *J Chem Phys*. 1935;3(5):300–313.
19. Ben-Nairn A, Marcus Y. Solvation thermodynamics of non ionic solutes. *J Chem Phys*. 1984;81(4):2016–2027.
20. Ben-Nairn A. *Molecular Theory of Solutions*. Oxford: Oxford University Press, 2006.
21. Lin ST, Blanco M, Goddard WA. The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: validation for the phase diagram of Lennard-Jones fluids. *J Chem Phys*. 2003;119(22):11792–11805.
22. Lai P-K, Hsieh C-M, Lin S-T. Rapid determination of entropy and free energy of mixtures from molecular dynamics simulations with the two-phase thermodynamic model. *Phys Chem Chem Phys*. 2012; 14(43):15206–15213.
23. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701–1718.
24. Lin S-T, Maiti PK, Goddard WA III. Two-phase thermodynamic model for efficient and accurate absolute entropy of water from molecular dynamics simulations. *J Phys Chem B*. 2010;114(24): 8191–8198.
25. Lin S-T, Hsieh C-M, Lee M-T. Solvation and chemical engineering thermodynamics. *J Chin Inst Chem Eng*. 2007;38(5-6):467–476.
26. Huang S-N, Pascal TA, Goddard WA III, Maiti PK, Lin S-T. Absolute entropy and energy of carbon dioxide using the two-phase thermodynamic model. *J Chem Theory Comput*. 2011;7(6):1893–1901.
27. Pascal TA, Lin S-T, Goddard WA III. Thermodynamics of liquids: standard molar entropies and heat capacities of common solvents from 2PT molecular dynamics. *Phys Chem Chem Phys*. 2011;13(1): 169–181.
28. Berens PH, Mackay DHJ, White GM, Wilson KR. Thermodynamics and quantum corrections from molecular dynamics for liquid water. *J Chem Phys* 1983;79(5):2375–2389.
29. McQuarrie DA. *Statistical Mechanics*. New York: Harper & Row Publishers, 1976.
30. Carnahan NF, Starling KE. Thermodynamic properties of a rigid-sphere fluid. *J Chem Phys*. 1970;53(2):600–603.
31. 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(3):435–447.
32. Jorgensen WL, Severance DL. Aromatic-aromatic interactions: free energy profiles for the benzene dimer in water, chloroform, and liquid benzene. *J Am Chem Soc*. 1990;112(12):4768–4774.
33. Jorgensen WL, Maxwell DS, TiradoRives 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(45): 11225–11236.
34. Martin MG, Siepmann JI. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J Phys Chem B*. 1998; 102(14):2569–2577.
35. Allen FH. The Cambridge structural database: a quarter of a million crystal structures and rising. *Acta Crystallogr B*. 2002;58:380–388.
36. Wilding WV, Rowley RL, Oscarson JL. DIPPR (R) Project 801 evaluated process design data. *Fluid Phase Equilib*. 1998;150:413–420.
37. Caleman C, van Maaren PJ, Hong M, Hub JS, Costa LT, van der Spoel D. Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J Chem Theory Comput*. 2012;8(1):61–74.
38. Broughton JQ, Gilmer GH. Molecular dynamics investigation of the crystal–fluid interface. I. Bulk properties. *J Chem Phys* 1983;79(10): 5095–5104.
39. Frenkel D, Ladd AJC. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *J Chem Phys*. 1984;81(7):3188–3193.
40. Aragones JL, Noya EG, Valeriani C, Vega C. Free energy calculations for molecular solids using GROMACS. *J Chem Phys*. 2013; 139(3):034104.

Manuscript received Mar. 7, 2015;