

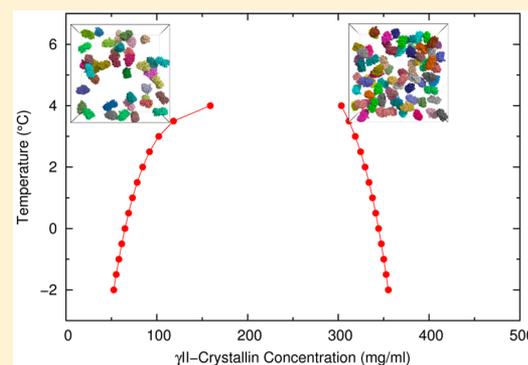
Fast Method for Computing Chemical Potentials and Liquid–Liquid Phase Equilibria of Macromolecular Solutions

Sanbo Qin and Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, United States

Supporting Information

ABSTRACT: Chemical potential is a fundamental property for determining thermodynamic equilibria involving exchange of molecules, such as between two phases of molecular systems. Previously, we developed the fast Fourier transform (FFT)-based method for Modeling Atomistic Protein-crowder interactions (FMAP) to calculate excess chemical potentials according to the Widom insertion. Intermolecular interaction energies were expressed as correlation functions and evaluated via FFT. Here, we extend this method to calculate liquid–liquid phase equilibria of macromolecular solutions. Chemical potentials are calculated by FMAP over a wide range of molecular densities, and the condition for coexistence of low- and high-density phases is determined by the Maxwell equal-area rule. When benchmarked on Lennard-Jones fluids, our method produces an accurate phase diagram at 18% of the computational cost of the current best method. Importantly, the gain in computational speed increases dramatically as the molecules become more complex, leading to many orders of magnitude in speed up for atomistically represented proteins. We demonstrate the power of FMAP by reporting the first results for the liquid–liquid coexistence curve of γ II-crystallin represented at the all-atom level. Our method may thus open the door to accurate determination of phase equilibria for macromolecular mixtures such as protein–protein mixtures and protein–RNA mixtures, that are known to undergo liquid–liquid phase separation, both in vitro and in vivo.



INTRODUCTION

Simple molecular fluids have long served as models for studying phase equilibria.^{1,2} The gas–liquid phase coexistence of these fluids is a prototype for the separation of two liquid phases in colloidal and biomacromolecular solutions, one poor in solutes and one rich in solutes.³ The latter phase, widely found in cells as membraneless microcompartments and appearing as droplets under light and other microscopes, is implicated in many normal cellular functions as well as in diseases and has received great attention lately.^{4–8} For example, droplet formation of BuGZ, an evolutionarily conserved low-complexity protein, promotes microtubule polymerization and the assembly of the spindle apparatus during cell division to ensure equal partitioning of chromosomes into daughter cells.⁹ The low-complexity region of another protein, FUS, drives the reversible transitions between the dissolved, droplet, and hydrogel-like phases, but mutations increase the propensity of FUS to form irreversible fibrillar hydrogels and may ultimately lead to neurodegenerative diseases.¹⁰ Many proteins that form intracellular droplets bind RNA, and RNA molecules can either be active components of the droplets or be present as cargo.^{11,12} Liquid–liquid phase separation is also an important factor in protein crystallization and in the formulation of protein-based therapeutics. Whereas the presence of a metastable fluid–fluid critical point can enhance crystal nucleation and decrease the propensity of forming amorphous

aggregates,^{13,14} liquid–liquid phase separation is a serious concern in the formulation of therapeutics based on highly concentrated proteins such as monoclonal antibodies.^{15,16} A practical method that allows for computation of liquid–liquid phase equilibria under a realistic representation of intermolecular interactions will lead to the much needed physical understanding of the phase behaviors of the wide range of biomacromolecular systems and the numerous functional implications.

Because of the vast number of atoms in a single protein molecule and the large number of protein molecules to be treated in phase-diagram calculations, such calculations with proteins represented at the atomistic level have not been feasible. Nearly all previous calculations were limited to a spherical representation of protein molecules, such as isotropic, patchy, and hybrid models,^{17–21} aeolotropic models,²² and embedded charge modes.²³ The limitations of spherical models and importance of anisotropic electrostatic interactions have been noted.²⁴ In particular, a spherical model overestimated the electrostatic repulsion of proteins when compared to that by a structure-based model in which individual amino-acid residues

Special Issue: J. Andrew McCammon Festschrift

Received: February 16, 2016

Revised: May 16, 2016

Published: June 21, 2016

were coarse-grained into interaction sites, and the discrepancy worsens with increasing anisotropy in molecular shape and charge.²⁵ Arguably the most detailed model in protein-phase calculations has all neutral residues collapsing into a sphere but all charged residues represented as interaction sites.²⁶ The system size in this study was relatively small, with only 40 protein molecules in the simulations box. An atomistic representation has been used in Brownian dynamics simulations of concentrated protein solutions,^{27,28} but these simulations have not been used for calculating phase equilibria.

We have developed a fast Fourier transform (FFT)-based method for Modeling Atomistic Protein-crowder interactions (FMAP) to calculate the excess chemical potential (μ^{ex}) for inserting a protein molecule into a crowded (i.e., concentrated) solution of macromolecules.^{29,30} FMAP implements the Widom insertion³¹ but with enormous gain in computational speed arising from the key idea of expressing intermolecular interactions as correlation functions and evaluating these functions via FFT. Compared to that in a brute-force implementation of the Widom insertion, FMAP produced over 4 orders of magnitude in speed up for calculating excess chemical potentials of atomistically represented proteins.²⁹ We further found that numerical errors arising from using relatively large grid spacing in discretizing for FFT can be compensated by modest adjustment of the interaction energy parameters; therefore, accurate results can be achieved at low computational cost.³⁰ These features position FMAP well for calculating the chemical potentials and phase equilibria of atomistically represented protein solutions.

For simple molecular fluids, Monte Carlo simulation is the method of choice in phase calculations, which has resulted in a detailed understanding of the phase behaviors of these systems. A state of the art method, namely, grand canonical transition matrix Monte Carlo (GCTMMC), produces a histogram of particle numbers in a single simulation run at a fixed chemical potential (μ), volume (V), and temperature (T).¹ With appropriate reweighting, histograms of particle numbers at other chemical potentials, in particular, the one corresponding to phase coexistence, can be obtained.³² In this way, the phase diagram can be constructed by repeating the simulations at different temperatures. In simulations of relatively large systems of Lennard-Jones fluids, metastable configurations such as droplets near the gas phase, slabs near the critical density, and bubbles near the liquid phase, have been observed.²

At coexistence, the chemical potentials of the two phases are equal. Hence, calculating the chemical potentials constitutes a conceptually simple route to calculating phase equilibria. The chemical potential can be separated into an ideal part and excess part

$$\mu = \mu^{\text{id}} + \mu^{\text{ex}} \quad (1)$$

The ideal part results from dispersion of the molecules in a given volume

$$\mu^{\text{id}} = k_{\text{B}}T \ln(\rho/\rho_0) \quad (2)$$

where k_{B} is Boltzmann's constant, ρ is the molecule number density, and ρ_0 is an unimportant constant. The excess part accounts for the intermolecular interactions and can be obtained by free energy perturbation methods such as the Widom insertion,³¹ particle deletion, and combinations thereof. According to Widom's original formulation,

$$\exp(-\beta\mu^{\text{ex}}) = \langle \exp(-\beta U_{\text{int}}) \rangle_N \quad (3)$$

where $\beta = (k_{\text{B}}T)^{-1}$, $\langle \dots \rangle_N$ denotes averaging over a simulation at fixed particle number N (thus mimicking the canonical ensemble) and U_{int} is the interaction energy of a particle that is fictitiously inserted into the N -particle system. Hereafter, this formulation of μ^{ex} is referred to as EXP. Combinations of particle insertion and deletion steps can improve the convergence and robustness of μ^{ex} .^{33–35} In the overlap sampling (OS) method, one obtains μ^{ex} as the free energy difference between a half insertion step and half deletion step,³⁴

$$\exp(-\beta\mu^{\text{ex}}) = \frac{\langle \exp(-\beta U_{\text{int}}/2) \rangle_N}{\langle \exp(\beta U_{\text{int}}/2) \rangle_{N+1}} \quad (4)$$

where $U_{\text{int}}/2$ in the numerator represents the energy for inserting a "half" particle into the N -particle system and $-U_{\text{int}}/2$ in the denominator represents the energy for deleting a "half" particle from the $(N + 1)$ -particle system. In the Bennett acceptance ratio (BAR) method,^{33–35} to minimize the variance of the μ^{ex} estimate, weights are introduced to the two Boltzmann factors,

$$\exp(-\beta\mu^{\text{ex}}) = \frac{\langle w(U_{\text{int}}) \exp(-\beta U_{\text{int}}/2) \rangle_N}{\langle w(U_{\text{int}}) \exp(\beta U_{\text{int}}/2) \rangle_{N+1}} \quad (5)$$

where

$$w(U_{\text{int}}) = 1/\cosh[\beta(U_{\text{int}} - C)/2] \quad (6)$$

and the parameter C therein is in turn related to μ^{ex} , as well as the number of insertion attempts, n_0 , and the number of deletion attempts, n_1

$$\beta C = \beta\mu^{\text{ex}} + \ln(n_1/n_0) \quad (7)$$

Hence, eqs 5–7 have to be solved iteratively to find μ^{ex} .

Lomakin et al.¹⁷ and Vörtler et al.³⁶ implemented the Widom insertion (the EXP version) in Monte Carlo simulations of square-well fluids to obtain the chemical potentials at various densities. The subcritical chemical potential versus density isotherms were used to calculate the corresponding gas–liquid coexistence densities and chemical potentials by means of the Maxwell equal-area rule. This approach, although conceptually straightforward, is not widely used in practice because of the significant computational expenses. In a brute-force implementation of the Widom insertion, most insertion attempts encounter strong particle–particle repulsion and the chance of inserting into the rare cavities large enough to accommodate the inserted particle in a dense fluid is extremely small, and yet it is the latter insertions that make the dominant contribution to the excess chemical potential.³⁷ Recently, an effort to speed up the calculation by graphics processing units was reported.³⁵ In FMAP, our FFT-based implementation of the Widom insertion, we obtain U_{int} values for inserting a particle at all points on a uniform grid covering the simulation box of the N -particle system, thereby allowing for accurate calculation of μ^{ex} . Potentially, FMAP may hold a breakthrough for phase-equilibrium calculations.

In this paper, we demonstrate this potential of FMAP in a proof-of-concept study of Lennard-Jones fluids. In line with our previous study,²⁹ we find that FMAP produces significant speed up for treating particle insertion, especially for large system sizes. We combine the FMAP-enabled insertion with particle deletion to calculate μ^{ex} according to the OS and BAR methods. We then use the Maxwell equal-area rule to locate the gas–

liquid coexistence condition. The resulting phase diagram agrees very well with previous GCTMMC results³⁸ but at only 18% of the computational cost.

We emphasize that, according to the McMillan and Mayer theory of solutions,³⁹ a macromolecular solution is equivalent to a fluid of the macromolecules alone when the interactions of the latter molecules represent potentials of the mean force with the solvent degrees of freedom averaged out. For example, the expansion for osmotic pressure in powers of the solute concentration is completely analogous to the virial expansion of the pressure of a molecular fluid. Similarly, the liquid–liquid phase separation of a macromolecular solution is analogous to the gas–liquid coexistence of a molecular fluid.³ Therefore, in exactly the same way as that for calculating gas–liquid phase equilibria of molecular fluids, the FMAP method can be applied for calculating liquid–liquid phase equilibria of macromolecular solutions, as long as interactions of solute molecules account for the solvent implicitly.

Importantly, the FMAP-enabled gain in computational speed in treating particle insertion increases drastically as the molecules become more complex, leading to many orders of magnitude in speed up for atomistically represented proteins.^{29,30} With the computational cost of particle insertion reduced to a feasible range, even for atomistically represented proteins, it is anticipated that configurational sampling of the N -particle and $(N + 1)$ -particle systems for a protein solution may present a remaining obstacle. Although Brownian dynamics simulations of such systems appear promising,^{27,28} it may well be that cruder treatments are necessary for more exhaustive configurational sampling (e.g., to break up transient clusters). To explore how configurations that cover the configurational space fully but with inexact population impact the accuracy of phase diagram calculations, we use simulations run at a temperature above the critical temperature to generate the whole phase diagram and find that the EXP method, involving only particle insertion, is more tolerant of imperfection in the configurational population than the OS and BAR methods. We apply the forgoing ideas to γ II-crystallin solutions, with the protein molecules represented at the all-atom level, and report the first such results for the liquid–liquid coexistence curve. Together, these results suggest that FMAP may open the door to accurate determination of phase equilibria for macromolecular mixtures, such as protein–protein mixtures and protein–RNA mixtures that are known to undergo a liquid–liquid phase separation, both in vitro and in vivo.

METHODS

Monte Carlo Simulations. We ran conventional Metropolis Monte Carlo simulations of Lennard-Jones fluids in the canonical ensemble (i.e., constant NVT) to generate particle configurations for insertion and deletion. Each system started with a prescribed number of particles on the grid points of a cubic lattice. Rather than the full Lennard-Jones potential

$$u_{\text{LJ}}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (8)$$

where r , σ , and ϵ denote the interparticle distance, particle diameter, and energy well depth, respectively, we used a shifted-force form

$$u(r) = u_{\text{LJ}}(r) - u_{\text{LJ}}(r_c) - (r - r_c)u'_{\text{LJ}}(r_c) \quad \text{for } r \leq r_c \\ = 0 \quad \text{for } r > r_c \quad (9)$$

where r_c denotes the cutoff distance and a prime denotes a derivative. The periodic boundary condition was applied and treated by the minimum image convention. For simplicity, we use reduced units, that is, $\sigma = 1$, $\epsilon = 1$, and $k_B = 1$. As a result, all temperatures listed below are scaled by ϵ/k_B , and all number densities scaled by σ^{-3} . The cutoff distance chosen was 2.5; for this choice, the reported critical temperature, T_c was 0.937.⁴⁰

The side length of the cubic simulation box was 8 in all of the simulations. Simulations were run at six subcritical temperatures ranging from 0.65 to 0.90, at an increment of 0.05. At each temperature, a total of 15 N -particle systems were simulated for insertion, with N ranging from 30 to 450 (at an increment of 30) and, correspondingly, the number density ranging from 0.059 to 0.879. In parallel, the same number of $(N + 1)$ -particle systems were simulated for deletion.

At each Monte Carlo simulation step, a particle was randomly chosen for a trial move. The displacement of the move in each direction was random, with a uniform distribution limited to a maximum displacement of 1.2. The trial move was accepted with the probability $\min[1, \exp(-\beta\Delta U_{\text{int}})]$, where ΔU_{int} is the change, upon the trial move, in the interaction energy of the chosen particle with the other particles in the simulation box. The actual acceptance rates for the trial moves ranged from 0.89 to 0.21 in the simulations at low and high densities. For each system, 2 million steps were used for equilibration; the average interaction energy per particle reached equilibrium over a much smaller number of steps. Afterwards, 20 million steps were used for collecting configurations for particle insertion and deletion.

Brute-Force Insertion and Deletion. Configurations were collected at intervals of 10 000 steps from the simulations of the N -particle and $(N + 1)$ -particle systems for insertion and deletion, respectively. For brute-force insertion (referred to as Ins hereafter), a particle was randomly placed into any location in the simulation box and its interaction energy with the N particles was calculated. In each of the 2000 configurations collected, a total of 512 000 (see below) fictitious insertions were made. For deletion, a particle was randomly selected and its interaction energy with the other N particles was calculated. In each of the 2000 configurations, 10 000 fictitious deletions (equivalent to one deletion per Monte Carlo simulation step) were made.

FMAP-Enabled Insertion. The same 2000 snapshots collected for each system were used for FMAP-enabled insertion. In FMAP, the interaction energy of the inserted particle with the N -particle system is expressed as a correlation function.^{29,30} For the present case, if the inserted particle is located at \mathbf{r} and the N particles are located at \mathbf{r}_i , $i = 1-N$, then the interaction energy is

$$U_{\text{int}}(\mathbf{r}) = \sum_{i=1}^N u(|\mathbf{r} - \mathbf{r}_i|) \quad (10a)$$

$$= \int d^3\mathbf{s} u(|\mathbf{s} - \mathbf{r}|) \sum_{i=1}^N \delta(\mathbf{s} - \mathbf{r}_i) \quad (10b)$$

$$\equiv u(r) \times q(\mathbf{r}) \quad (10c)$$

$$(10)$$

which is akin to the interaction of an “electric” potential field, $u(r)$, with an “electric” charge distribution, $q(\mathbf{r})$. Equation 10 highlights the fact that the interaction energy is expressed as a cross correlation of $u(r)$ and $q(\mathbf{r})$. Both of the latter functions are discretized on a cubic grid, and their correlation is evaluated via FFT, according to the well-known theorem that the correlation of two functions in direct space corresponds to the product of the two functions in Fourier space.

The discretization of $u(r)$ is straightforward, whereas the discretization of $q(\mathbf{r})$ entails “distributing” the Lennard-Jones particles (represented by the delta functions in eq 10) to the grid points. Following our previous work,^{29,30} each Lennard-Jones particle was distributed to the eight grid points forming the smallest enclosing cube, with weights according to trilinear interpolation. The discretization incurs numerical errors. Again, following earlier work,³⁰ these errors were corrected by scaling the parameters of the energy function. Specifically, ϵ was scaled by a factor λ , and simultaneously, σ was scaled by the factor $6/(\lambda + 5)$. For the grid spacing of 0.1 used, benchmarking against results from brute-force insertion allowed us to determine the appropriate λ value to be 1.06.

The discretization of $u(r)$ was done once and its Fourier transform used for all 2000 configurations of each system. For each configuration, the discretized version of $q(\mathbf{r})$ was Fourier transformed, and the product of the two Fourier transforms was then transformed back to direct space. The results are the values of $U_{\text{int}}(\mathbf{r})$ at all $80^3 = 512\,000$ grid points. In other words, each FMAP-enabled insertion calculation is equivalent to brute-force insertions into 512 000 locations.

Excess Chemical Potential Calculations. The interaction energies from insertion (either Ins or FMAP-enabled) and deletion were used to calculate the excess chemical potential according to three methods: EXP, OS, and BAR. The EXP and OS calculations were performed on the fly, but for the BAR calculations, we saved the interaction energies in the form of histograms (to avoid having to save the large number of actual interaction energies). The histogram bin width was 0.001, and the lower and upper bounds were -100 and 100 , respectively, for the insertion energy and negative of the deletion energy. The iteration of eqs 5–7 was started with the μ^{ex} value of the OS method and terminated when the change in parameter C was less than 0.001.

Errors in μ^{ex} were estimated using a Python code, written by J. Fung, R. W. Perry, and T. G. Dimiduk (<https://github.com/manoharan-lab/flyvbjerg-std-err/>), that implements Flyvbjerg and Petersen’s block decorrelation technique⁴¹ and also reports convergence of the error estimate. For a given method (e.g., FMAP/BAR) for calculating μ^{ex} , the input to the error estimate consisted of μ^{ex} values calculated by that method on blocks of snapshots with a given block size. For example, blocks with size 10 were composed of snapshots 1–10, 11–20, etc.

Polynomial Fitting of Excess Chemical Potentials. The excess chemical potential as a function of number density was fit to a fifth-order polynomial

$$\beta\mu^{\text{ex}} = \sum_{l=1}^5 b_l \rho^l \quad (11)$$

which effectively is a truncated virial expansion.⁴² (Similar forms of fitting were used by Lomakin et al.¹⁷ and Vörtler et al.³⁶) The first coefficient, b_1 , in theory is equal to twice the more commonly known second-order coefficient, B_2 , of the

virial expansion of pressure (P). The latter coefficient in turn is given by a Mayer cluster integral

$$B_2 = -2\pi \int dr r^2 [\exp(-\beta u(r)) - 1] \quad (12)$$

The full chemical potential is the sum of the excess part and the ideal part (see eqs 1 and 2). The resulting μ – ρ relation constitutes isotherms in the μ – ρ plane.

RESULTS AND DISCUSSION

Isotherms in the μ – ρ Plane. In Figure 1, we display the values of the excess chemical potential at 15 number densities

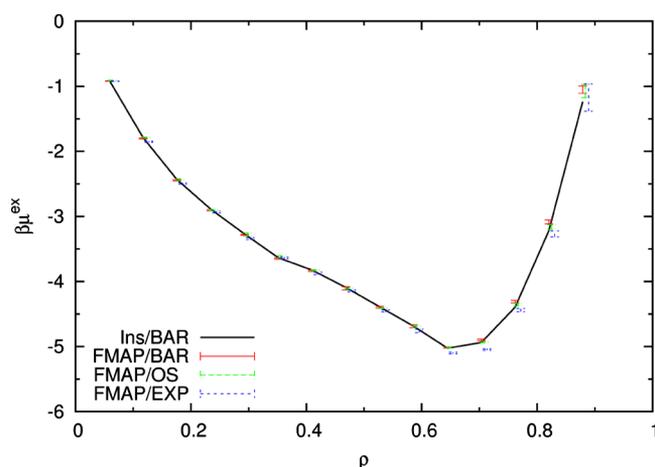


Figure 1. Dependence of the excess chemical potential on the number density for Lennard-Jones fluids at $T = 0.65$. Particle insertion is either by brute-force (Ins) or using FMAP. The excess chemical potential is calculated using only insertion (EXP) or combinations with deletion, either by the OS method or by the BAR method. For clarity, the FMAP results by the three methods at a given density are slightly shifted horizontally from each other.

for $T = 0.65$. The results of brute-force insertion (Ins) are reproduced well by FMAP-enabled insertion. From here on, we will focus on the latter results. Among the three methods of combining insertion and deletion, as expected, the standard errors are the lowest for BAR, intermediate for OS, and highest for EXP. Comparing OS and EXP, it is apparent that, at high ρ in particular, including deletion significantly reduces the error in μ^{ex} . It also significantly speeds up the convergence of calculated μ^{ex} . The value of μ^{ex} becomes very stable after 6 million Monte Carlo simulation steps (of a total 20 million steps; Figure S1). Similar results were obtained at five higher temperatures.

Figure 2A displays the polynomial fit of the FMAP/BAR results for the μ^{ex} dependence on ρ at $T = 0.65$. The data agree with the fit very closely. We also checked the value of the first-order coefficient, b_1 , against the theoretical value, $2B_2$ (with B_2 given by eq 12; Table S1). The agreement between b_1 and $2B_2$ is moderate at the lowest temperature studied, with a 20% discrepancy. This discrepancy at lower temperatures can partly be attributed to stronger effects of interparticle interactions, upon scaling by $k_B T$. As a result, the higher-order terms in the expansion of μ^{ex} in terms of ρ can be important, and truncation at the fifth order can lead to deviation of b_1 from $2B_2$. The agreement improves at higher temperatures, with the discrepancy decreasing to 8% at $T = 0.90$.

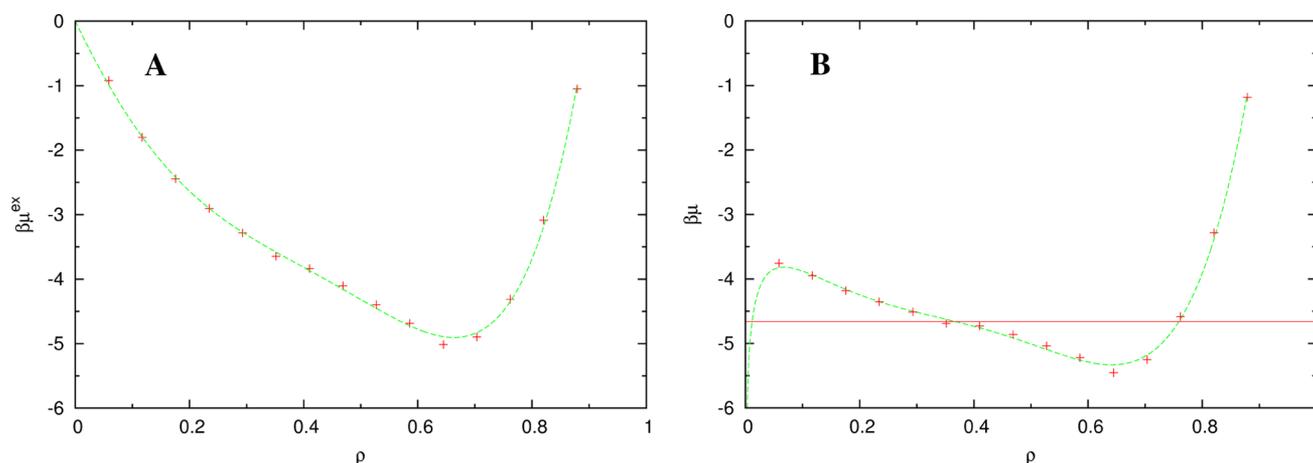


Figure 2. (A) Polynomial fit (green curve) of the density dependence of the excess chemical potential (calculated by FMAP/BAR, shown as red crosses) at $T = 0.65$. (B) The full chemical potential, after adding the ideal part. A van der Waals loop, characterized by a decrease in μ with increasing ρ in the mid-range ρ , is present. The red horizontal line crosses the van der Waals loop with equal areas enclosed above and below. The height of this line corresponds to the chemical potential at phase coexistence, whereas the low and high ρ values where the red line intersects the green curve are the densities of the gas and liquid phases, respectively.

Table 1. Chemical Potential and Densities at Gas–Liquid Phase Coexistence

T	Maxwell			pseudo-GCMC ^a			GCTMMC		
	$\beta\mu_{\text{co}}$	$10\rho_{\text{g}}$	ρ_{l}	$\beta\mu_{\text{co}}$	$10\rho_{\text{g}}$	ρ_{l}	$\beta\mu_{\text{co}}$	$10\rho_{\text{g}}$	ρ_{l}
0.65	-4.659	0.117	0.759	-4.662	0.123	0.757	-4.655	0.113	0.762
0.70	-4.208	0.204	0.721	-4.210	0.214	0.719	-4.195	0.195	0.729
0.75	-3.829	0.333	0.680	-3.830	0.349	0.677	-3.812	0.319	0.693
0.80	-3.511	0.528	0.636	-3.512	0.556	0.632	-3.491	0.504	0.652
0.85	-3.234	0.818	0.592	-3.234	0.869	0.586	-3.218	0.795	0.601
0.90	-2.997	1.23	0.531	2.997	1.39	0.514	-2.986	1.35	0.524

^aGCMC, grand canonical Monte Carlo.

Figure 2B shows the isotherm in the μ – ρ plane, featuring a “van der Waals loop” (see next), after adding the ideal part.

Chemical Potential and Densities at Gas–Liquid Coexistence. In a certain temperature range, isotherms for Lennard-Jones fluids on the PV diagram are theoretically expected to be nonmonotonic and contain a van der Waals loop. The pressure approaches zero at very large volumes. When the volume first decreases, the pressure increases, as expected from the ideal gas law. When the volume decreases more, such that the average interparticle distance is around σ and the particles experience appreciable attraction toward each other, the pressure decreases. When the volume decreases even further, the particles start to repel each other, and finally, the pressure increases again, ultimately to infinity.

The van der Waals loop can be observed on a small system of Lennard-Jones particles, which maintains homogeneity in particle dispersion when the volume varies over a wide range. However, for a macroscopic system, in the density region where the van der Waals loop is expected, the particles separate into two phases, one with low density (hence the gas phase) and the other with high density (hence the liquid phase). Maxwell⁴³ recognized that these two densities correspond to the end points of a horizontal (i.e., constant V) line segment that crosses (the portion of the isotherm that we now call) the van der Waals loop in such a way as to enclose equal areas above and below. In effect, as would be required for the coexistence of the gas and liquid phases, the difference in Gibbs free energy (G) between the two end points of the line segment is zero, because this difference equals the sum of the two areas (with a

positive sign for the one above and a negative sign for the one below). The vertical height of this line segment is the pressure at coexistence. This approach is known as the Maxwell equal-area rule.

As illustrated by Figure 2B, isotherms in the μ – ρ plane can also contain a van der Waals loop. An analogous Maxwell equal-area rule can also be used to locate the phase coexistence condition. On the basis of the following relation, valid at constant temperature,

$$dP = \rho d\mu \quad (13a)$$

we find the difference in pressure between the end points of a horizontal line segment crossing the van der Waals loop in the μ – ρ plane to be

$$P_2 - P_1 = \int_{\text{loop}} \rho d\mu \quad (13b)$$

If the height of the horizontal line segment is chosen such that the areas enclosed above and below are equal, then the integral on the right-hand side of eq 13b is zero and hence $P_2 = P_1$; then, this height is the chemical potential, μ_{co} , at gas–liquid coexistence, and the two end points correspond to the number densities of the two phases, ρ_{g} and ρ_{l} .

Using the foregoing Maxwell equal-area rule (Figure 2B), the value of $\beta\mu_{\text{co}}$ is found to be -4.659 , and the number densities of the gas and liquid phases are 0.0117 and 0.759 , respectively, at $T = 0.65$. These agree very well with the corresponding GCTMMC results,³⁸ with $\beta\mu_{\text{co}} = -4.655$, $\rho_{\text{g}} = 0.0113$, and $\rho_{\text{l}} = 0.762$.

Phase Diagram. The chemical potentials and number densities at phase coexistence were similarly determined for the five higher temperatures. The results are listed in Table 1 (under the heading “Maxwell”), and the phase diagram in the temperature-density plane is displayed in Figure 3. Overall

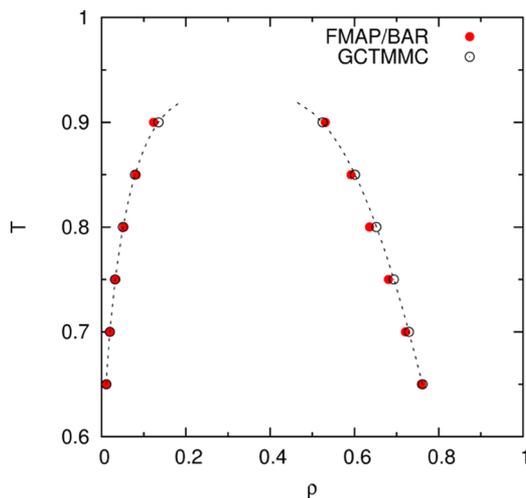


Figure 3. Gas–liquid coexistence densities at different subcritical temperatures, calculated by FMAP/BAR and previous GCTMMC results. The dotted curve is a fit to the GCTMMC results for guiding the eye.

there is very good agreement with the GCTMMC results,³⁸ although a mild discrepancy appears at higher temperatures. This discrepancy is not due to any insufficient elimination of the discretization error of FMAP-enabled insertion; results calculated by brute-force insertion also show a growing, but mild, trend of deviation from the GCTMMC results at higher temperatures (Figure S2).

Instead, we attribute a major part of the discrepancy to differing manifestations of system size effects in our calculations based on canonical ensemble (i.e., constant NVT) simulations and in those based on grand canonical Monte Carlo (GCMC) (i.e., constant μVT) simulations. Recall that in our calculations phase coexistence is identified by choosing the height of a line, $\mu = \mu_{co}$, that intersects the isotherm in the μ – ρ plane with equal areas enclosed, so as to ensure pressure equality between the two phases; the two outer intersection points give the gas and liquid number densities. In GCMC, one builds a histogram of particle numbers at some chemical potential and reweights it to obtain the histogram at $\mu = \mu_{co}$, which has equal populations under the peaks representing the gas and liquid phases.^{1,44} This equal-population rule is meant to ensure pressure equality between the two phases. Furthermore, the gas and liquid number densities are calculated as population averages in the two phases. Below, we show that the isotherm in the μ – ρ plane can be used to construct a grand canonical histogram of particle numbers. In doing so, we reveal that both the enforcement of pressure equality and the determination of gas and liquid number densities are done in subtly different ways under the Maxwell equal-area rule and GCMC equal-population rule.

According to statistical mechanics, the grand canonical histogram of particle numbers is given by

$$\Pi(N; \mu, V, T) = \frac{Q(N, V, T) \exp(\beta\mu N)}{\Xi(\mu, V, T)} \quad (14)$$

where $Q(N, V, T)$ is the canonical partition function and $\Xi(\mu, V, T)$ is the grand canonical partition function. Taking the logarithm of both sides and noting $\ln Q(N, V, T) = -\beta F$, where F denotes Helmholtz free energy, we have

$$\ln \Pi(N; \mu, V, T) = -\beta F + \beta\mu N - \ln \Xi(\mu, V, T) \quad (15)$$

As in the canonical ensemble $\mu(N, V, T) = \left(\frac{\partial F}{\partial N}\right)_{V, T}$, under constant V and T we can express F as an integral of $\mu(N, V, T)$ over N (or, equivalently, over ρ). Consequently

$$\begin{aligned} \ln \Pi(N; \mu, V, T) &= -\beta \int_0^N dN \mu(N, V, T) + \beta\mu N + \ln \Pi(0; \mu, V, T) \end{aligned} \quad (16)$$

which constitutes our recipe for constructing the grand canonical histogram of particle numbers and involves integrating our canonical ensemble chemical potential over N (the first term on the right-hand side). The second term is a reweighting factor that allows the histograms to be obtained for different chemical potentials,³² whereas the third term corresponds to a normalization constant. We call this approach pseudo-GCMC.

The pseudo-GCMC histograms of particle numbers at $T = 0.90$ are displayed in Figure 4 for two choices of μ in the second

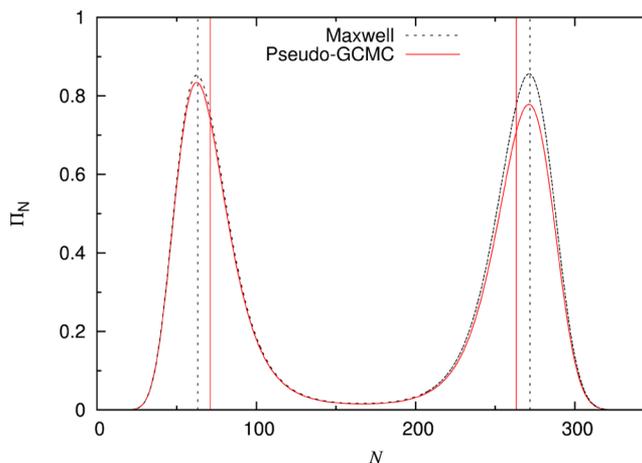


Figure 4. Histograms of particle numbers constructed by the pseudo-GCMC approach at $T = 0.90$, illustrating subtle differences between the two methods of determining phase coexistence properties. Application of the Maxwell equal-area rule leads to the black curve, with equal peak heights, for the histogram and peak positions (indicated by black vertical dashes) for the gas and liquid number densities. Application of the GCMC equal-population rule leads to the red curve, with equal areas under the peaks, for the histogram and population averages (indicated by red vertical lines) for the gas and liquid number densities.

term on the right-hand side of eq 16. One (black curve; $\mu = -2.99704$) follows the Maxwell equal-area rule, which results in equal peak heights. The other (red curve; $\mu = -2.99747$) follows the GCMC equal-population rule, which means equal areas under the two peaks of the histogram. Moreover, the gas and liquid number densities determined by the Maxwell equal-area rule correspond to the two N values (indicated by black vertical dashes) at which the peaks occur, whereas these densities determined by the GCMC equal-population rule correspond to the population averages (indicated by red

Table 2. Computational Costs of Three Approaches to Phase Equilibrium Calculations

	number of systems	total number of steps (10 ⁶)	computational cost (10 ¹⁰ pair interactions)			
			move	insertion	deletion	total
GCTMMC ^a	400	8000	128	47.9	47.9	223
Ins/BAR ^b	30	600	27.9	369	7.2	404
FMAP/BAR ^b	30	600	27.9	5.4	7.2	40.5

^aSimulations spanned a range of N from 1 to 400, with 20 million steps at each N .³⁸ Each step was one of three choices: move, insertion, or deletion, with probabilities of 0.4, 0.3, and 0.3, respectively. ^bParallel simulations were run for N -particle and $(N + 1)$ -particle systems for 15 N values ranging from 30 to 450 (at an increment of 30). For each system, 20 million steps were simulated. At intervals of 10 000 steps, 512 000 insertions (or the equivalent amount by FMAP) and 10 000 deletions were made.

vertical lines). For a macroscopic system, the two peaks of the histogram are infinitely sharp and symmetric so that the phase coexistence results determined by these two methods would be indistinguishable. However, for the finite-size systems studied here, the peaks become broad and asymmetric at high temperatures, giving rise to mild differences in the phase coexistence properties determined by the two methods. The pseudo-GCMC results are also listed in Table 1. Overall, the mild discrepancy arising from finite-size effects (and other factors) notwithstanding, the good agreement with the GCTMMC results provides strong validation of our FMAP-based approach to phase equilibrium calculations.

One may suggest further repressing the mild finite-size effects observed here by increasing the system size. Unfortunately, as the system size increases, the van der Waals loop becomes less prominent.² Hence, there is a limit to how much the system size can be increased. Similarly, the histograms of particle numbers become harder to converge in GCMC simulations when the system size is increased.

Gain in Computational Speed. To simplify the comparison of computational costs, we use the calculation of a pair interaction as the cost unit. Monte Carlo simulation moves, brute-force insertion, and deletion can all be measured in this unit. However, FMAP-enabled insertion does not involve directly calculating pair interactions. Whereas the cost of brute-force insertion increases in proportion to the number of particles in the system, the cost of FMAP-enabled insertion remains the same for all system sizes. FMAP-enabled insertion into each of the 2000 selected configurations produces interaction energies at 512 000 positions (i.e., equivalent to that many brute-force insertions), but the cost is equivalent to just 10 000 brute-force insertions into a system with 180 particles. Hence, there is a 51.2-fold gain in speed over that of brute-force insertion for that system size, and the gain further increases as the system size increases.

We list in Table 2 the costs of calculating the phase equilibrium at a single temperature by three methods. With brute-force insertion, the total cost is equivalent to 4.04×10^{12} pair interaction calculations, of which the bulk, that is, 3.69×10^{12} pair interaction calculations, is due to the insertion part. With FMAP, this cost item is reduced by 68.3-fold, whereas the other two cost items, deletion and Monte Carlo move, remain the same, and the latter item now rises to be the major contributor to computational cost. The effect of FMAP is to reduce the total computational cost by 10-fold.

As explained above, our approach to phase equilibrium calculation, i.e., applying the Maxwell equal-area rule on canonical simulations, is, in practice, very similar to the common GCMC approach based on the equal-population rule. In this connection, we expect that the computational cost of our approach, when brute-force insertion is used, is comparable

to the computational cost of the GCMC approach. The expectation is borne out by the cost for a typical GCTMMC setup,³⁸ with a total of 2.23×10^{12} pair interaction calculations required per temperature (Table 2). By comparison, our approach, with insertion accelerated by FMAP, incurs only 18% of the computational cost of GCTMMC.

Importantly, the gain in computational speed afforded by FMAP in treating insertion increases dramatically as the particles become more complex.^{29,30} This is because the cost of FMAP calculations, with the particles mapped to a grid, remains the same when the level of details in representing the particles is increased. That is, whether the particles are represented by a single atom, as for Lennard-Jones fluids, or thousands of atoms, as for protein molecules, FMAP costs the same. This feature is manifested in a small way by the aforementioned constancy in the cost of FMAP-enabled insertion for all system sizes. Previously, we have demonstrated that FMAP can produce over 4 orders of magnitude in speed up over brute-force insertion for atomistically represented proteins.²⁹ We anticipate that FMAP will reduce the computational cost of particle insertion to a feasible range, even for atomistically represented proteins, in phase equilibrium calculations (see below).

Tolerance of Imperfection in Configurational Population. As noted above and shown in Table 2, with FMAP substantially cutting the computational cost of insertions, configurational sampling by Monte Carlo moves becomes the computationally most expensive item. We anticipate that this situation will also hold true in studies to calculate phase diagrams of atomistically represented protein solutions (see below). Configurational sampling of such systems has been done by Brownian dynamics simulations,^{27,28} but full coverage of the configurational space may require cruder treatments, for example, to break up transient clusters. The resulting configurations may cover the configurational space fully but with an inexact population. Below, we report some preliminary results on how an inexact configurational population impacts the phase diagram for the systems studied here.

To mimic fast sampling that results in full coverage of the configurational space but with an inexact population, we ran Monte Carlo simulations to generate configurations at $T = 1.0$, which is above the reported critical temperature of 0.937.⁴⁰ We used these same configurations for insertion and deletion for all six subcritical temperatures to calculate excess chemical potentials and the whole phase diagram. Specifically, the interaction energies for insertion and deletion were calculated once and stored as histograms. Different β values for the corresponding temperatures were then used in eqs 3–5 to obtain μ^{ex} according to the EXP, OS, and BAR methods, respectively. We note in passing that the particular imperfection studied here in configurational population can in theory be

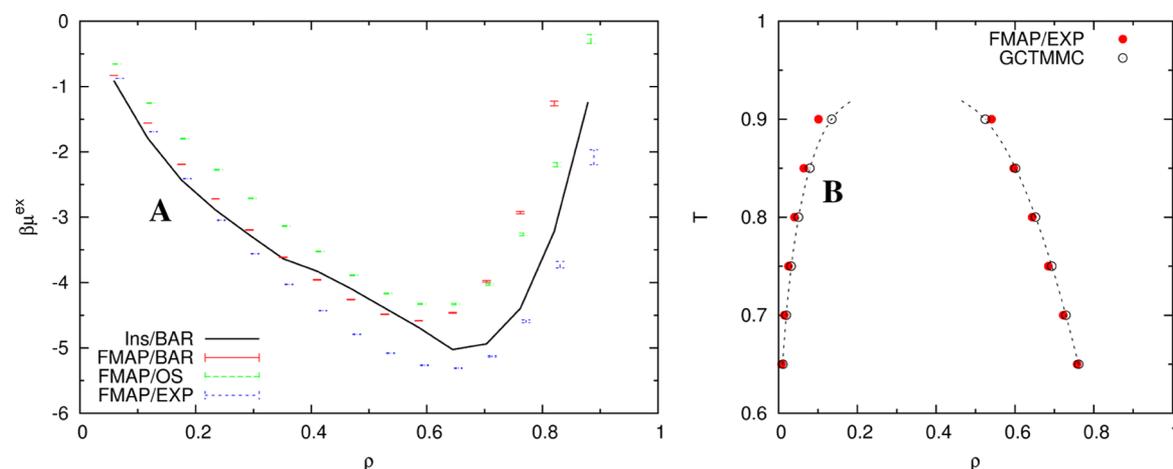


Figure 5. Impact of an inexact configurational population on excess chemical potential and the phase diagram. Configurations from simulations at $T = 1.0$ were used for calculations at all subcritical temperatures, except for the benchmarks. (A) Density dependence of the excess chemical potential at $T = 0.65$. The Ins/BAR result is the same as that in Figure 1 and is used here as a benchmark. (B) Gas–liquid coexistence densities calculated by FMAP/EXP. The GCTMMC results are the same as those in Figure 3 and are used here as a benchmark.

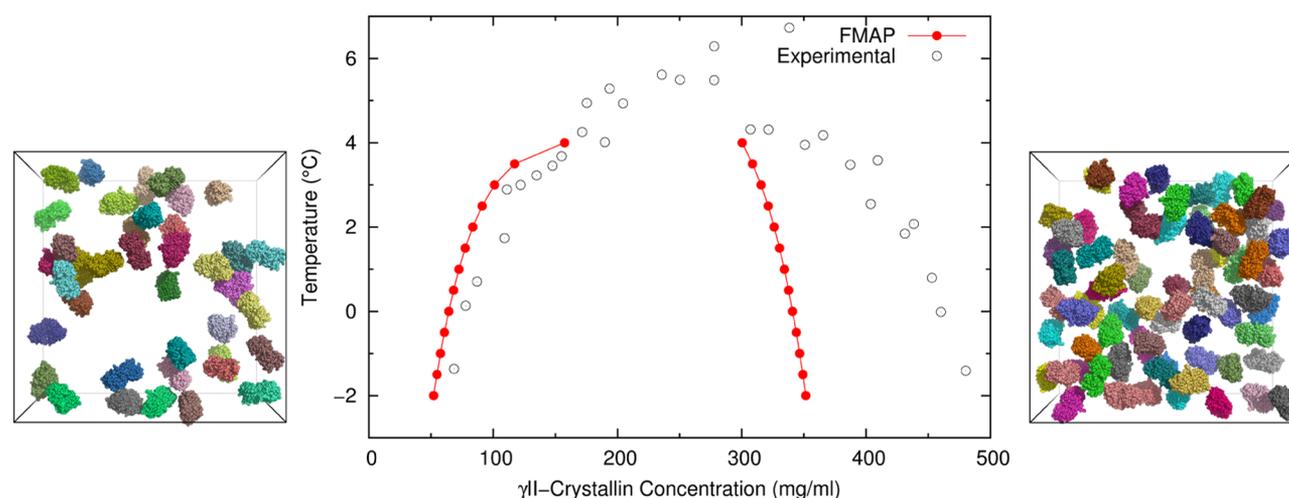


Figure 6. Comparison of our calculated liquid–liquid coexistence curve for γ II-crystallin with the experimental result (ref 46). In the calculation, γ II-crystallin molecules are represented at the all-atom level, and their interactions are modeled by Lennard-Jones potentials and Debye–Hückel potentials. Snapshots of protein configurations at 123 and 307 mg/mL in an 81 Å slice are shown to the left and right, respectively.

corrected by properly reweighting the configurations.^{17,45} This correction was not pursued here because our goal was to examine the impact of the imperfect configurational population.

As Figure 5A shows, the EXP method, involving only particle insertion, is the best at tolerating imperfection in configurational population, followed by BAR and then OS. The deletion contribution to μ^{ex} is dominated by deletions of particles that form strong favorable interactions with neighboring particles; such configurations are undersampled in the high-temperature simulations, leading to underestimation of the magnitude of μ^{ex} . These erroneously calculated deletion contributions are heavily weighted in OS but are tempered in BAR. Thus, we conclude that in cases in which the configurational population is inexact EXP should be the preferred method for calculating μ^{ex} . Figure 5B shows that the phase diagram determined by FMAP/EXP using the configurations from the high-temperature simulations agree reasonable well with the GCTMMC benchmark.

Liquid–Liquid Coexistence Curve of γ II-Crystallin Represented at the All-Atom Level. We applied the preceding idea of using FMAP/EXP on one set of configurations to calculate the liquid–liquid coexistence curve

of the γ II-crystallin solution. The phase diagram of this system has been determined experimentally⁴⁶ and has been the subject of computational and theoretical studies using a spherical representation of protein molecules.^{17,21} Here, we represented γ II-crystallin at the all-atom level. The protein structure was from Protein Data Bank entry 1AMM,⁴⁷ with hydrogens added and AMBER charges assigned using the PDB2PQR server⁴⁸ (protonation states according to PROPKA⁴⁹ for pH 7). Our potential function, as described in previous work,³⁰ consisted of Lennard-Jones potentials for modeling steric, van der Waals, and hydrophobic interactions and Debye–Hückel potentials for modeling electrostatic interactions between protein molecules (at ionic strength 0.24 M). The energy parameters were the same as those in the previous work, except that the overall scaling factor of the Lennard-Jones potentials was slightly adjusted (to 0.156) and the overall scaling factor of the Debye–Hückel potentials was dependent on protein concentration ρ (increasing by 2-fold as ρ increased from zero to infinity), to model the decrease in the dielectric constant of the solution with increasing ρ .^{50,51}

To finesse the difficulty in configurational sampling, we used the 2000 configurations collected for Lennard-Jones fluids at $T = 1.0$ and a given ρ as the basis for generating protein configurations in a γ II-crystallin solution. Each Lennard-Jones was replaced by a protein molecule with a random orientation; the side length of the simulation box was equated to 324 Å, such that 450 molecules in the box translated to a concentration of 461 mg/mL. For FMAP calculations, the simulation box was discretized at a grid spacing of 0.6 Å, resulting in 1.57×10^8 voxels. The FMAP/EXP calculations for 2000 protein configurations took approximately 7 h on 16 cores of two Intel Xeon E5-2650 2.6 GHz CPUs. FMAP-enabled insertion was repeated 10 times, each with a different orientation for the test protein; the histograms of interaction energies from the 10 repeats were merged. For a series of temperatures, μ^{ex} was calculated from these histograms, and its dependence on ρ was fit to eq 11, and the liquid–liquid coexistence densities were determined according to the Maxwell equal-area rule.

In Figure 6, we compare the calculated liquid–liquid coexistence curve with the experimental result. The calculation shows good agreement with the experiment for the low-concentration branch but underestimates the protein concentrations of the upper branch. The potential function used here has not been carefully parameterized for modeling interprotein interactions in protein solutions. So, it is possible that fine-tuning the parameters may lead to broadening of the liquid–liquid coexistence curve. However, our preliminary study suggests another likely source for the insufficient broadness. The underestimation of protein concentrations in the upper branch apparently can be traced to a rapid rise in the contribution of the steric part of the potential function to μ^{ex} . Our preliminary explicit-solvent molecular dynamics simulations of concentrated protein solutions show that protein molecules tend to form clusters. Cluster formation would result in more free space than that in a more uniform distribution of protein molecules, such as those generated on the basis of configurations of Lennard-Jones fluids, and thereby slows down the rise in the contribution of the steric part of the potential function to μ^{ex} at high protein concentrations.

As a possible extension of the idea of using simulations at one temperature to calculate phase coexistence properties at other temperatures, simulations for one protein system may be used, after homology modeling, for calculations on mutants or homologous proteins. In particular, homologous proteins can have very different critical temperatures for liquid–liquid phase separation. Such examples include antibodies⁵² and γ -cristallins.^{46,53} Although some of the differences in phase behaviors can be qualitatively rationalized by differences in protein surface properties,⁵³ it will be interesting to quantitatively analyze these potential correlates by methods and ideas presented in this study.

CONCLUSIONS

We have presented a proof-of-concept study to demonstrate the potential of FMAP as a fast method for calculating phase equilibria. We have shown that the common GCMC approach for such calculations is equivalent, both operationally and in computational cost, to applying the Maxwell equal-area rule to isotherms in the μ – ρ plane, with μ determined by brute-force particle insertion and deletion. FMAP drastically reduces the computational cost of particle insertion, especially for complex molecules including proteins, such that the burden shifts to

configurational sampling. We have also explored ideas to maximize the use of configuration samples, for example, for calculating phase-coexistence properties at multiple temperatures, and presented the first results for the liquid–liquid coexistence curve of γ II-crystallin represented at the all-atom level. It now appears that the door to accurate modeling of liquid–liquid phase separation for protein–protein and protein–RNA mixtures is opening.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b01607.

Two additional figures (Figures S1 and S2) and one additional table (Table S1) (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: hzhou4@fsu.edu. Phone: (850) 645-1336. Fax: (850) 644-7244.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by NIH Grants GM88187 and GM118091. We thank Prof. Cameron Abrams for distributing his Monte Carlo simulation and analysis codes online at <http://www.pages.drexel.edu/~cfa22/msim/>, which were instructive for writing our own codes. Computation reported in this work used resources at the Florida State University Research Computing Center and the Extreme Science and Engineering Discovery Environment (XSEDE),⁵⁴ which is supported by National Science Foundation grant number ACI-1053575.

ABBREVIATIONS

FFT, fast Fourier transform; FMAP, FFT-based method for modeling atomistic protein–crowder interactions; GCMC, grand canonical Monte Carlo; GCTMMC, grand canonical transition matrix Monte Carlo; OS, overlap sampling; BAR, Bennett acceptance ratio

REFERENCES

- (1) Errington, J. R. Direct Calculation of Liquid-Vapor Phase Equilibria from Transition Matrix Monte Carlo Simulation. *J. Chem. Phys.* **2003**, *118*, 9915–9925.
- (2) Binder, K.; Block, B. J.; Virnau, P.; Troster, A. Beyond the Van Der Waals Loop: What Can Be Learned from Simulating Lennard-Jones Fluids inside the Region of Phase Coexistence. *Am. J. Phys.* **2012**, *80*, 1099–1109.
- (3) Asherie, N.; Lomakin, A.; Benedek, G. B. Phase Diagram of Colloidal Solutions. *Phys. Rev. Lett.* **1996**, *77*, 4832–4835.
- (4) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoegge, C.; Gharakhani, J.; Jülicher, F.; Hyman, A. A.; Germline, P. Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **2009**, *324*, 1729–1732.
- (5) Brangwynne, C. P.; Mitchison, T. J.; Hyman, A. A. Active Liquid-Like Behavior of Nucleoli Determines Their Size and Shape in *Xenopus* Oocytes. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 4334–4339.
- (6) Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58.
- (7) Garber, K. Cell Biology. Protein ‘Drops’ May Seed Brain Disease. *Science* **2015**, *350*, 366–367.

- (8) Strzyz, P. Molecular Networks: Protein Droplets in the Spotlight. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 639.
- (9) Jiang, H.; Wang, S.; Huang, Y.; He, X.; Cui, H.; Zhu, X.; Zheng, Y. Phase Transition of Spindle-Associated Protein Regulate Spindle Apparatus Assembly. *Cell* **2015**, *163*, 108–122.
- (10) Murakami, T.; Qamar, S.; Lin, J. Q.; Schierle, G. S. K.; Rees, E.; Miyashita, A.; Costa, A. R.; Dodd, R. B.; Chan, F. T. S.; Michel, C. H.; et al. ALS/PTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015**, *88*, 678–690.
- (11) Anderson, P.; Kedersha, N. RNA Granules: Post-Transcriptional and Epigenetic Modulators of Gene Expression. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 430–436.
- (12) Zhang, H.; Elbaum-Garfinkle, S.; Langdon, E. M.; Taylor, N.; Occhipinti, P.; Bridges, A. A.; Brangwynne, C. P.; Gladfelter, A. S. RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell* **2015**, *60*, 220–230.
- (13) ten Wolde, P. R.; Frenkel, D. Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science* **1997**, *277*, 1975–1978.
- (14) Wedekind, J.; Xu, L.; Buldyrev, S. V.; Stanley, H. E.; Reguera, D.; Franzese, G. Optimization of Crystal Nucleation Close to a Metastable Fluid-Fluid Phase Transition. *Sci. Rep.* **2015**, *5*, No. 11260.
- (15) Wang, Y.; Lomakin, A.; Latypov, R. F.; Benedek, G. B. Phase Separation in Solutions of Monoclonal Antibodies and the Effect of Human Serum Albumin. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16606–16611.
- (16) Raut, A. S.; Kalonia, D. S. Liquid-Liquid Phase Separation in a Dual Variable Domain Immunoglobulin Protein Solution: Effect of Formulation Factors and Protein-Protein Interactions. *Mol. Pharmaceutics* **2015**, *12*, 3261–3271.
- (17) Lomakin, A.; Asherie, N.; Benedek, G. B. Monte Carlo Study of Phase Separation in Aqueous Protein Solutions. *J. Chem. Phys.* **1996**, *104*, 1646–1656.
- (18) Kern, N.; Frenkel, D. Fluid-Fluid Coexistence in Colloidal Systems with Short-Ranged Strongly Directional Attraction. *J. Chem. Phys.* **2003**, *118*, 9882–9889.
- (19) Gogelein, C.; Nagele, G.; Tuinier, R.; Gibaud, T.; Stradner, A.; Schurtenberger, P. A Simple Patchy Colloid Model for the Phase Behavior of Lysozyme Dispersions. *J. Chem. Phys.* **2008**, *129*, No. 085102.
- (20) Bianchi, E.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Theoretical and Numerical Study of the Phase Diagram of Patchy Colloids: Ordered and Disordered Patch Arrangements. *J. Chem. Phys.* **2008**, *128*, 144504.
- (21) Kastelic, M.; Kalyuzhnyi, Y. V.; Hribar-Lee, B.; Dill, K. A.; Vlachy, V. Protein Aggregation in Salt Solutions. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6766–6770.
- (22) Lomakin, A.; Asherie, N.; Benedek, G. B. Aeolotopic Interactions of Globular Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9465–9468.
- (23) Rosch, T. W.; Errington, J. R. Investigation of the Phase Behavior of an Embedded Charge Protein Model through Molecular Simulation. *J. Phys. Chem. B* **2007**, *111*, 12591–12598.
- (24) Lund, M. Anisotropic Protein-Protein Interactions Due to Ion Binding. *Colloid Surf., B* **2016**, *137*, 17–21.
- (25) Kaieda, S.; Lund, M.; Plivelic, T. S.; Halle, B. Weak Self-Interactions of Globular Proteins Studied by Small-Angle X-Ray Scattering and Structure-Based Modeling. *J. Phys. Chem. B* **2014**, *118*, 10111–10119.
- (26) Kurut, A.; Åkesson, T.; Forsman, J.; Lund, M. Anisotropic Interactions in Protein Mixtures: Self Assembly and Phase Behavior in Aqueous Solution. *J. Phys. Chem. Lett.* **2012**, *3*, 731–734.
- (27) McGuffee, S. R.; Elcock, A. H. Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm. *PLoS Comput. Biol.* **2010**, *6*, No. e1000694.
- (28) Mereghetti, P.; Gabdouliline, R. R.; Wade, R. C. Brownian Dynamics Simulation of Protein Solutions: Structural and Dynamical Properties. *Biophys. J.* **2010**, *99*, 3782–3791.
- (29) Qin, S.; Zhou, H. X. FFT-Based Method for Modeling Protein Folding and Binding under Crowding: Benchmarking on Ellipsoidal and All-Atom Crowders. *J. Chem. Theory Comput.* **2013**, *9*, 4633–4643.
- (30) Qin, S.; Zhou, H. X. Further Development of the FFT-Based Method for Atomistic Modeling of Protein Folding and Binding under Crowding: Optimization of Accuracy and Speed. *J. Chem. Theory Comput.* **2014**, *10*, 2824–2835.
- (31) Widom, B. Some Topics in the Theory of Fluids. *J. Chem. Phys.* **1963**, *39*, 2808–2812.
- (32) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo Technique for Studying Phase Transitions. *Phys. Rev. Lett.* **1988**, *61*, 2635–2638.
- (33) Bennett, C. H. Efficient Estimation of Free-Energy Differences from Monte-Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (34) Lu, N.; Singh, J. K.; Kofke, D. A. Appropriate Methods to Combine Forward and Reverse Free-Energy Perturbation Averages. *J. Chem. Phys.* **2003**, *118*, 2977–2984.
- (35) Daly, K. B.; Benziger, J. B.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Massively Parallel Chemical Potential Calculation on Graphics Processing Units. *Comput. Phys. Commun.* **2012**, *183*, 2054–2062.
- (36) Vörtler, H. L.; Schäfer, K.; Smith, W. R. Simulation of Chemical Potentials and Phase Equilibria in Two- and Three-Dimensional Square-Well Fluids: Finite Size Effects. *J. Phys. Chem. B* **2008**, *112*, 4656–4661.
- (37) Hong, Q.-J.; van de Walle, A. Direct First-Principles Chemical Potential Calculations of Liquids. *J. Chem. Phys.* **2012**, *137*, No. 094114.
- (38) Shen, V. K.; Siderius, D. W.; Krekelberg, W. P. NIST Standard Reference Simulation Website, NIST Standard Reference Database Number 173. http://www.nist.gov/mml/csd/informatics_research/sat_tmnc_2d5.cfm (accessed Nov 24, 2015).
- (39) McMillan, W. G.; Mayer, J. E. The Statistical Thermodynamics of Multicomponent Systems. *J. Chem. Phys.* **1945**, *13*, 276–305.
- (40) Errington, J. R.; Debenedetti, P. G.; Torquato, S. Quantification of Order in the Lennard-Jones System. *J. Chem. Phys.* **2003**, *118*, 2256–2263.
- (41) Flyvbjerg, H.; Petersen, H. G. Error Estimates on Averages of Correlated Data. *J. Chem. Phys.* **1989**, *91*, 461–466.
- (42) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Dover Publications, Inc.: New York, 1986.
- (43) Clerk-Maxwell, J. Xxii.—On the Dynamical Evidence of the Molecular Constitution of Bodies. *J. Chem. Soc.* **1875**, *28*, 493–508.
- (44) Binder, K.; Landau, D. P. Finite-Size Scaling at First-Order Phase Transitions. *Phys. Rev. B* **1984**, *30*, 1477–1485.
- (45) Ferrenberg, A. M.; Landau, D. P.; Swendsen, R. H. Statistical Errors in Histogram Reweighting. *Phys. Rev. E* **1995**, *51*, 5092–5100.
- (46) Broide, M. L.; Berland, C. R.; Pande, J.; Ogun, O. O.; Benedek, G. B. Binary-Liquid Phase Separation of Lens Protein Solutions. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5660–5664.
- (47) Kumaraswamy, V. S.; Lindley, P. F.; Slingsby, C.; Glover, I. D. An Eye Lens Protein–Water Structure: 1.2 Å Resolution Structure of γ B-Crystallin at 150 K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1996**, *52*, 611–622.
- (48) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PBD2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32*, W665–W667.
- (49) Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (50) Tjong, H.; Zhou, H. X. Prediction of Protein Solubility from Calculation of Transfer Free Energy. *Biophys. J.* **2008**, *95*, 2601–2609.
- (51) Predeus, A. V.; Gul, S.; Gopal, S. M.; Feig, M. Conformational Sampling of Peptides in the Presence of Protein Crowders from AA/CG-Multiscale Simulations. *J. Phys. Chem. B* **2012**, *116*, 8610–8620.
- (52) Wang, Y.; Lomakin, A.; Latypov, R. F.; Laubach, J. P.; Hideshima, T.; Richardson, P. G.; Munshi, N. C.; Anderson, K. C.;

Benedek, G. B. Phase Transitions in Human Igg Solutions. *J. Chem. Phys.* **2013**, *139*, 121904.

(53) Norledge, B. V.; Hay, R. E.; Bateman, O. A.; Slingsby, C.; Driessen, H. P. Towards a Molecular Understanding of Phase Separation in the Lens: A Comparison of the X-Ray Structures of Two High Tc γ -Crystallins, γ E and γ F, with Two Low Tc γ -Crystallins, γ B and γ D. *Exp. Eye Res.* **1997**, *65*, 609–630.

(54) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I. T.; Gaither, K. P.; Grimshaw, A. S.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; et al. XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16*, 62–74.