Published in final edited form as:

J Chem Theory Comput. 2020 November 10; 16(11): 6823–6842. doi:10.1021/acs.jctc.0c00832.

Temperature and Phase Transferable Bottom-up Coarse-Grained Models

Jaehyeok Jin,

Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States

Alvin Yu,

Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States

Gregory A. Voth

Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States

Abstract

Despite the high fidelity of bottom-up coarse-grained (CG) approaches to recapitulate the structural correlations in atomistic simulations, the general use of bottom-up CG methods is limited because of the nontransferable nature of these CG models under different thermodynamic conditions. Because bottom-up CG potentials usually correspond to configuration-dependent free energies of the system, recent studies have focused on adjusting enthalpic or entropic contributions to account for issues with transferability. However, these approaches can require a manual adjustment of the CG interaction a priori and are usually limited to constant volume ensembles. To overcome these limitations, we construct temperature and phase transferable CG models under constant pressure by developing the ultra-coarse-graining (UCG) methodology in the mean-field limit. In the mean-field ansatz, an embedded semi-global order parameter recapitulates global changes to the system by automatically adjusting the effective CG interactions, thus bridging free energy decompositions with UCG theory. The method presented is designed to faithfully capture structural correlations under different thermodynamic conditions, using a single UCG model. Specifically, we test the applicability of the developed theory in three distinct cases: (1) different temperatures at constant pressure in liquids, (2) different temperatures across thermodynamic phases, and (3) liquid/vapor interfaces. We demonstrate that the systematic construction of both

Corresponding Author: Gregory A. Voth - gavoth@uchicago.edu .

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00832

Supporting Information

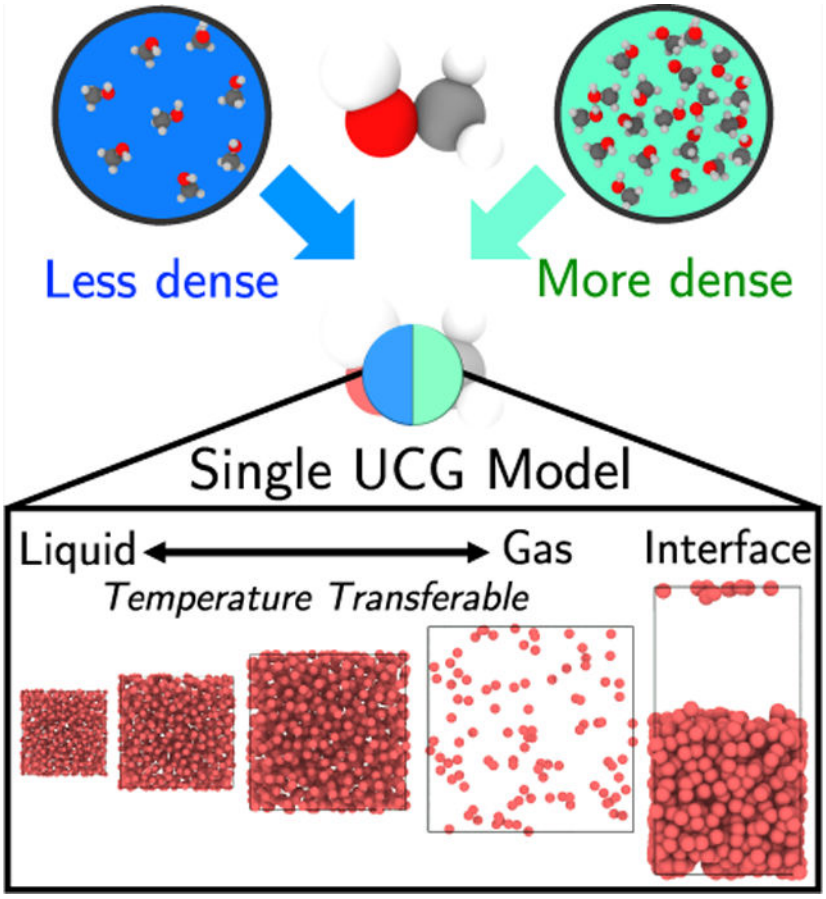
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00832.

Computation details of both FG and CG simulations; additional analyses including the semi-global order parameter of acetone; full error metric data for bulk CG liquids; additive density coupling; and pressure transferability (PDF)

The authors declare no competing financial interest.

temperature and phase transferable bottom-up CG models is possible using this generalized UCG theory. Based on our findings, this approach significantly extends the transferability and applicability of the bottom-up CG theory and method.

Graphical Abstract



1. INTRODUCTION

Bottom-up coarse-grained (CG) approaches are typically constructed to preserve fine-grained (FG) static correlations (e.g., full all-atom resolution) by fitting CG interaction parameters to all-atom reference trajectories.[1–3] Iterative Boltzmann inversion,[4] inverse Monte Carlo,[5] multiscale coarse-graining (MS-CG),[6–10] and relative entropy minimization[11–13] are a few such approaches. It has been shown that these bottom-up CG approaches can accurately recapitulate pairwise correlations[14] or three-body correlations[15] in the reference FG system. In comparison with top-down approaches where CG parameters are designed to reproduce experimental thermodynamics or phenomenology,[16–19] a physically accurate description of these static correlations is the main advantage of bottom-up methods. However, transferability remains a central issue in bottom-up CG models.[20–24]

The transferability problem stems from the fact that bottom-up CG models are often designed to have probability distribution functions, $p_R(\mathbf{R}^N)$, which are consistent with that of the FG system, $p_I(\mathbf{r}^n)$, in configuration space. These spaces are linked by the mapping operator $M_{\mathrm{R}}^{N}:\mathrm{r}^{n}\rightarrow\mathrm{R}^{N}$, where $\mathbf{r}^n$ and $\mathbf{R}^N$ refer to the configurations of the FG and CG systems, respectively.[8] The consistency condition can be written as

$$\exp\left[-\beta U_{\mathrm{CG}}\left(\mathbf{R}^{N}\right)\right]=C\int\mathrm{d}\mathbf{r}^{n}\delta\left(M_{\mathbf{R}}^{N}(\mathbf{r}^{n})-\mathbf{R}^{N}\right)\times\exp\left[-\beta u(\mathbf{r}^{n})\right] \quad (1)$$

where $u(\mathbf{r}^n)$ is the FG (e.g., all-atom) potential energy function and $\beta = 1/k_{\mathrm{B}}T$. The condition in eq 1 establishes that the CG effective potential $U_{\mathrm{CG}}(\mathbf{R}^N)$ is a many-body potential of mean force (PMF), i.e., a configuration-dependent free energy function. The CG potential is thus substantially different from conventional definitions of potential energy functions. Because the position-dependent CG free energy function can change for different thermodynamic conditions, $U_{\mathrm{CG}}(\mathbf{R}^N)$ is arguably not inherently transferable between those conditions.

Some fundamental thermodynamic relationships should, however, still hold in a many-body functional form as in eq 1. Inspired by a free energy decomposition (FED) scheme, several efforts have been made to partition the CG free energy functional into energetic and entropic components.[25,26] This approach allows for the construction of temperature transferrable CG models with interpolation.[27] Recently, we also showed that one can design combining rules for the CG PMF of mixed-composition systems by decomposing PMFs of individual components and mixing the terms separately.[24] However, these approaches require that the "interpolation" must be carried out under constant volume conditions.

Given the diverse length scales for CG simulations in chemical or biological systems, it is not only practical but also necessary to develop a method for the construction of temperature transferable CG models at constant pressure.[28,29] However, evaluating pressure in the CG system is challenging because of the missing degrees of freedom lost in the CG procedure. Without a properly designed barostat that accounts for representability issues,[30] both CG model parametrization and simulations at constant pressure are not feasible. In addition, FED schemes are intrinsic interpolations that require manual adjustments to entropic or volumic terms under different temperature and pressure conditions after performing atomic simulations of the target system in a range of temperature or pressure, reducing the predictiveness of the CG simulations.

One holy grail for CG modeling is to develop a single CG model that can be applied to any system without *a posteriori* adjustments.[31] A possible scheme to achieve this goal would be to construct a CG model that automatically detects the system condition and alters its entropic or volumic contributions *on the fly*. In practice, one would need to determine an appropriate order parameter that reflects the changes in the system. By implicitly coupling the order parameter to the CG interaction, the resultant CG model can be expected to more faithfully adjust its interaction under different system conditions. While conceptually sound, this requires a new CG theory to account for these changes and the design of order parameters that can determine global system properties *on the fly*.

We recently developed the ultra-coarse-graining (UCG) theory where CG sites can adopt internal states to represent the physical or chemical transitions occurring beneath the resolution of each CG site.[32–35] In contrast to conventional CG force fields, the UCG force field is expressed as a mixed interaction between each substate of the CG site. In principle, by carefully designing the internal states to be dependent on the order parameter, the single UCG model can represent more than one characteristic state of the system to help capture the many-body correlations using only pairwise interaction forms. Practically, by imposing a separation of time scales, this theory was further elaborated in two opposing limits of internal state dynamics: a discrete slow switching limit[33] and a rapid local equilibrium (RLE) limit.[34,35] The RLE limit allows UCG models to continuously change state contributions to the overall effective CG interaction based on the local environment *on the fly*[34,35] and can be applied to construct more transferable CG models.[36] In this limit, we have demonstrated that UCG models are able to capture various structural correlations that are often missing and not described in conventional CG models. To note a few, the missing hydrogen bonding interactions are faithfully described by introducing hydrogen bonding donor and acceptor states,[37] and the hydrophobic association of molecules is captured by constructing internal states based on the local associated density.[34] Despite its expressiveness, the physical connection between the UCG theory and the transferability of CG models over different state points is less clear.

In this work, we develop temperature transferable CG models under constant NPT dynamics by developing UCG methodology in the mean-field limit. From the mean-field ansatz, we aim to rigorously bridge general FED schemes with UCG models that are governed by a semi-global order parameter. We demonstrate the applicability of this approach in CG simulations of liquids over different temperatures under constant pressure conditions. We further develop methods to extend our temperature transferable CG model from a single phase to multiple phases and to generate “phase combinational” CG models that can capture structural correlations at liquid/vapor interfaces. Extending from the previous work that focused on phase transferability for either homogeneous or heterogeneous systems, we argue that the UCG methodology can be utilized to its fullest extent to encompass both conditions for the first time.

2. THEORY

2.1. Transferability Challenges and the FED of CG PMFs.

The direct use of CG interaction potentials at nonparametrized state points may be problematic because the CG interaction potential is essentially a configuration $\mathbf{R}^N$ dependent free energy, i.e., the CG PMF[8,9]

$$U(\mathbf{R}^N) = -k_{\mathrm{B}}T\ln\int \mathrm{d}\mathbf{r}^n \exp\left(-\frac{u(\mathbf{r}^n)}{k_{\mathrm{B}}T}\right)\delta\left(M^N_{\mathbf{R}}(\mathbf{r}^n)-\mathbf{R}^N\right) + (\text{ const. }) \tag{2}$$

As the CG potential unambiguously changes at different state points, several bottom-up CG approaches have attempted to determine this change.[24,29,34,36,38–48] Among those, our particular focus is to treat the CG interaction as a *free energy functional*.[49] FED-based

interpolation schemes have been developed to generate CG models by altering the entropic component of the free energy.[24–26] However, these methods were limited to constant volume conditions (*NVT* ensembles), in which the CG PMF corresponds to the Helmholtz free energy: $A = E - TS$.[24–26] Because various chemical reactions take place under the constant pressure condition (e.g., 1 atm), we will further extend our approach to constant pressure conditions.

Nevertheless, a simple extension of constant *NVT* to *NPT* ensembles is not possible because of pressure representability issues.[20–22,24] Because of the missing degrees of freedom in the CG models, calculating the CG pressure naïvely from the FG virial will result in significant deviations from the FG pressure.[50] In this respect, MS-CG parametrizations for different FG pressure conditions are still generally conducted for constant *NVT* dynamics after performing constant *NPT* dynamics with different pressure conditions.[30] Figure 1 highlights the temperature transferability issues in CG modeling as a result of difficulties with the representability of pressure.

It is worth noting an alternative approach called “pressure-matching”,[30,46,51–53] inspired from the work by Das and Andersen (DA), which attempts to resolve this inconsistency.[28] Instead of designing CG interactions as a function of configurational variables, the DA approach introduces a volume term U_V to the CG potential. As a consequence, variational optimization for both configurational and volumic interactions is feasible, enabling simulations in isothermal-isobaric CG ensembles. Although pressure-matching is appealing because of its extensibility and accuracy, in this work, we use a more conventional definition from the MS-CG methodology where the CG interaction is only an explicit function of configurational variables. Details of the MS-CG models with force-matching parametrization are discussed in the Supporting Information Section S1.

To account for changes in different pressure conditions, we introduce an enthalpy function to the Helmholtz free energy under constant pressure conditions, given as[49]

$$\Delta A = \Delta H - P\Delta V - T\Delta S \tag{3}$$

Equation 3 can account for changes in pressure conditions in constant *NVT* ensembles and further suggests that temperature transferability under constant pressure may be attained by manual adjustment of the $-P\Delta V - T\Delta S$ terms, as in our prior interpolation scheme for constant volume conditions. Ultimately, however, the goal here is to construct a temperature transferable CG model with a force field that determines the effective temperature and volume and is not explicitly or manually fitted on the basis of FED. We will further discuss how this can be implemented. In the current work, we aim to develop an automated *single* CG force field that is transferable to various temperatures by adjusting contributions to the free energy on the fly.

2.2. Theory of UCG.

Detailed illustrations of the UCG theory[32–35] and implementation[36,37,54–57] are discussed extensively in prior papers, and we recommend readers to refer ref 32 for general formalism, ref 33 for the MC-like switching limit, and refs 34 and 35 for the quasi-equilibrium (or RLE)

limit. Under state-free dynamics by adopting a separation of time scales, here, we use the RLE limit to explicitly couple order parameters to the CG Hamiltonian.

In spirit, the UCG model is defined by two additional variables: internal state s and the substate probability given the CG configuration, $p(s|\mathbf{R}^N)$. In $p(s|\mathbf{R}^N)$, the conditionality to CG configurations $\mathbf{R}^N$ derives from a bottom-up manner. In practical terms, we simplify particle interactions to be pairwise, and the pairwise approximation also applies to the internal states. Altogether, the mixed UCG interaction can thus be written as a pair summation of internal state interactions

$$U_{\text{mix}}^{\text{UCG}}\left(\mathbf{R}^N\right) = \sum_{I,J}\sum_{s_I,s_J} p\left(s_I|\mathbf{R}^N\right)p\left(s_J|\mathbf{R}^N\right)U_{s_I,s_J}^{(2)}\left(R_{IJ}\right) \tag{4}$$

In eq 4, the superscript *(2)* denotes that $U_{s_I,s_J}^{(2)}\left(R_{IJ}\right)$ has a pairwise form, and we will omit this superscript in subsequent equations for simplicity. We design the substate probabilities $p(s|\mathbf{R}^N)$ as differentiable so that the effective forces, $-\nabla U_{\text{mix}}^{\text{UCG}}\left(\mathbf{R}^N\right)$, can be written as the gradient of the mixed UCG interaction energy

$$\begin{aligned} -\nabla U_{\text{mix}}^{\text{UCG}}\left(\mathbf{R}^N\right) = & -\sum_{I,J}\sum_{s_I,s_J}\Big[p\left(s_I|\mathbf{R}^N\right)p\left(s_J|\mathbf{R}^N\right)\nabla U_{s_I,s_j}\left(R_{IJ}\right) \\ & + \left(p\left(s_J|\mathbf{R}^N\right)U_{s_I,s_J}\left(R_{IJ}\right)\nabla p\left(s_I|\mathbf{R}^N\right) + p\left(s_I|\mathbf{R}^N\right)U_{s_I,s_J}\left(R_{IJ}\right)\nabla p\left(s_J|\mathbf{R}^N\right)\right)\Big] \end{aligned} \tag{5}$$

The first term in eq 5 is the force between CG sites I and J, weighted by each state probability. Unlike other conventional CG methodologies, the UCG force expression accounts for contributions from the changes to the substate probabilities ∇p (the second term), which is related to the driving force for switching between internal states. This additional probability force facilitates the modeling of rapidly changing substates.[35] Because the force expression in eq 5 is still pairwise decomposable, MS-CG force-matching approaches can be readily applied to parametrize eq 5. Prior UCG implementations used this approach to minimize the following force residual[58]

$$\begin{aligned} & \chi^2\left[U\left(\sigma,\mathbf{R}^N\right)\right] \\ & = \lim_{T_s\to\infty}\frac{1}{T_s}\sum_t\frac{1}{N_s}\sum_n \left\|\nabla\left(U\left(\sigma_n,M\left(r_t^n\right)\right)+k_{\text{B}}T\ln p_{\Sigma}\left(\sigma_n|M\left(r_t^n\right)\right)-M^{+}\left(\nabla U\left(r_t^n\right)\right)\right)\right\|^2 \end{aligned} \tag{6}$$

Equation 6 suggests that UCG substate interactions can be directly obtained from FG simulations using force-matching.[34,36,37] However, this parametrization scheme limits the transferability of substate interactions, as it is unclear that substate interactions derived in this fashion can be fully transferable to bulk systems. In the previous UCG work, we witnessed that substate interactions derived from interfacial systems had a limited

transferability to pure CG interactions under bulk conditions because of mismatches of forces in long-range regimes.[37] In turn, one way to enhance the transferability of the CG model would be to *directly* utilize the CG interactions from bulk conditions as substate interactions. The primary remaining problem is then how to choose such CG interactions as UCG substate interactions and couple them to the UCG framework with an appropriate order parameter.

2.3. Mean-Field UCG Ansatz.

In the prior subsections, we introduced FEDs of the CG PMF and UCG methodology. Nevertheless, an actual and physical link between these two theories remains ambiguous. Here, we now aim to develop an ansatz to connect the two separate theories explicitly.

In order to parametrize UCG force fields, an appropriate order parameter, or collective variable, must first be determined to distinguish underlying internal UCG states. As an example, we recently demonstrated that the local (number) density of a particle is a suitable order parameter for differentiating phases in interfacial systems and also imparting transferrable interactions from bulk liquids to the UCG models of interfacial systems.[37] However, in the systems of interest here where the temperature is an explicitly controlled variable, there are no straightforward collective variables that are directly coupled to the temperature. This is because the UCG force field, eq 4, in the RLE limit is built upon the assumption of the locality in each UCG substate, and extensive variables such as temperature are not well-defined order parameters for the UCG parametrization.[34,35] Thus, we investigate under which conditions the local density, or other such order parameters, can represent the global temperature of the system.

A conventional choice for local density from previous UCG studies is a differentiable ($\in \mathbf{C}^1$) switching function $w(R_{IJ})$ that uses a sigmoidal hyperbolic tangent function

$$\rho_I(R_I) = \sum_{IJ} \frac{1}{2}\left(1 - \tanh\left(\frac{R_{IJ} - R_0}{\bar{R}}\right)\right) \tag{7}$$

Despite other possible forms of switching functions, we decide to utilize the form in eq 7 as it can recapitulate structural correlations emergent in various systems.[59–62] Using the local density order parameter $\rho_I(R_I)$, the substate probabilities for more dense (α) and less dense (β) states are given as

$$p_{s_I = \{\alpha,\beta\},I}(R_I) = \frac{1}{2}\left(1 \pm \tanh\left(\frac{\rho_I(R_I) - \rho_0}{\bar{\rho}}\right)\right) \tag{8}$$

The probability and order parameter defined in eqs 7 and 8 are local quantities by definition, which depend on the CG site *I*, and vary for different particles in the system. However, an averaged quantity over the CG ensemble may provide a feasible way to remove locality. The idea here is thus to introduce the mean-field description

$$\left\langle \rho_I(R_I) \right\rangle = \left\langle \sum_{IJ} \frac{1}{2}\left(1 - \tanh\left(\frac{R_{IJ} - R_0}{\bar{R}}\right)\right)\right\rangle \tag{9a}$$

$$\left\langle p_{I,s_I} \right\rangle = \frac{1}{2}\left(1 \pm \tanh\left(\frac{\langle \rho_I(R_I)\rangle - \rho_0}{\bar{\rho}}\right)\right) \tag{9b}$$

An average scheme introduced in eq 9a can be alternatively constructed by defi ning a counting function $n_I(R) = 1/2\left(1 - \tanh\left(R - R_0/\bar{R}\right)\right)$ and then calculating the structural average over the system using a radial distribution function (RDF), $g(r)$, such that[63]

$$\left\langle \rho_I \right\rangle = \int_{r=0}^{r=r_{\text{cut}}} 4\pi r^2 \rho_{\text{b}} g(r) n_I(r) \mathrm{d}r \tag{10}$$

Because of nontrivial correlations in liquids, an exact analytical form is impossible to determine. However, certain approximations can be still applied. Our key assumption is that if the cutoff distance R_0 is much larger than the correlation (or characteristic) length of the system, the structural average may give a local density value ρ_I that is no longer strongly dependent on the particle environment. Mathematically, this assumption can be reformulated as "there exists R_0 such that $\int_{r=0}^{r=R_0} 4\pi r^2 \rho_{\text{b}} \cdot g(r)\mathrm{d}r \approx \int_{r=0}^{r=R_0} 4\pi r^2 \rho_{\text{b}} \cdot 1\mathrm{d}r$". In this limit, we can drop the dependence on the Ith particle so the average local density ρ_I becomes ρ and depends only on the bulk density (global density), ρ_{b}. The local density order parameter ρ_I, then, is a quantity that depends on global properties as well. If the system is a homogeneous liquid, we expect that the densities follow a Gaussian distribution by the law of large numbers.[64,65] We will later validate these assumptions in simulations of real molecular liquids. In a practical manner, the computational cost of the CG simulation should not be affected by the choice of a large cutoff distance R_0 as long as R_0 is shorter than the cutoff used here to evaluate the pairwise interactions. On the other hand, it is expected that this semi-global feature allows for much larger integration time steps, which will be discussed later.

Imposing a global order parameter condition, average substate probabilities can be similarly treated as global properties

$$\left\langle p_{I,s_I} \right\rangle = \frac{1}{2}\left(1 \pm \tanh\left(\frac{\langle \rho \rangle - \rho_0}{\bar{\rho}}\right)\right) = \left\langle p_{s_I} \right\rangle \tag{11}$$

Because the averaged substate probability $\left\langle p_{I,s_I} \right\rangle$ no longer depends on a particular CG particle, this ansatz allows us to assess the average many-body UCG PMF. Recall that the full expression of the UCG free energy $U_{\text{mix}}^{\text{UCG}}\left(\mathbf{R}^N\right)$ is

$$U_{\mathrm{mix}}^{\mathrm{UCG}}(\mathbf{R}^N) = \sum_{I,J}\Big[\langle p_{I,\alpha}p_{J,\alpha}\rangle U_{\alpha\alpha}(R_{IJ}) + \langle p_{I,\beta}p_{J,\beta}\rangle U_{\beta\beta}(R_{IJ}) + \big(\langle p_{I,\alpha}p_{J,\beta}\rangle + \langle p_{I,\beta}p_{J,\alpha}\rangle\big)U_{\alpha\beta}(R_{IJ})\Big] \tag{12}$$

Equation 12 is the UCG Hamiltonian for the CG simulations. However, it should be noted that an actual assessment of eq 12 is not feasible because p_{I,s_I} varies over different particles, and thus, we are not able to project it on its pairwise basis sets to visualize the interaction. To rigorously bridge UCG with FED, we define α as a lower temperature T_{low} (thus more dense) state and β as a higher temperature T_{high} (less dense) state at constant pressure. In turn, we obtain a mean-field approximation of the UCG model at the averaged density. The collective behavior of substate probability in a given ensemble results in the averaged UCG PMF $\overline{U_{\mathrm{mix}}^{\mathrm{UCG}}}(\mathbf{R}^N)$ as

$$\overline{U_{\mathrm{mix}}^{\mathrm{UCG}}}(\mathbf{R}^N) = \sum_{I,J}\Big[\langle p_{I,\alpha}p_{J,\alpha}\rangle U_{\alpha\alpha}(R_{IJ}) + \langle p_{I,\beta}p_{J,\beta}\rangle U_{\beta\beta}(R_{IJ}) + (\langle p_{I,\alpha}p_{J,\beta}\rangle + \langle p_{I,\beta}p_{J,\alpha})\rangle U_{\alpha\beta}(R_{IJ})\Big] \tag{13}$$

In eq 13, $\langle p_{I,s_I}p_{J,s_J}\rangle$ is the averaged correlation of substate probabilities. However, in the mean-field ansatz, p_{I,s_I} and p_{J,s_J} become uncorrelated

$$\langle p_{I,s_I}p_{J,s_J}\rangle = \langle p_{I,s_I}\rangle\langle p_{J,s_J}\rangle \tag{14}$$

Equation 14 can be thought of as a natural property of the Gaussian distribution function. Finally, the mean-field UCG free energy is given as

$$\overline{U_{\mathrm{mix}}^{\mathrm{UCG}}}(\mathbf{R}^N) = \sum_{I,J}\Big[\langle p_\alpha\rangle^2 U_{\alpha\alpha}(R_{IJ}) + \langle p_\beta\rangle^2 U_{\beta\beta}(R_{IJ}) + 2\langle p_\alpha\rangle\langle p_\beta\rangle U_{\alpha\beta}(R_{IJ})\Big] \tag{15}$$

Here, we note that ρ_I from the mean-field ansatz is still computed *locally* for each CG site *I*. This also satisfies the locality assumption made in the RLE theory as p_I and p_J are both governed by the Gaussian distribution, as shown in eq 14.[34] However, in order to distinguish it from the conventional local density that has been utilized in the previous UCG studies,[34,36,37] we denote the specific local density that accounts for the mean-field limit as the "semi-local" density hereafter. Figure 2 highlights two major advantages of this semi-global density order parameter in contrast to conventional local or global density features. Because the semi-global density is evaluated from pair contributions at each configuration, it still contains the density fluctuation information during the CG simulation, unlike the global density with a single value. In the previous UCG studies, we have noticed that density

fluctuations are essential to model the driving force of rapidly changing substates from the $-\nabla p$ contribution.[34,36,37] By correctly recapitulating the density fluctuations, the UCG Hamiltonian in eq 12 can be faithfully applied to inhomogeneous systems, where the changes in density are expected to be captured by modulating the internal UCG states. We will revisit this point by the end of the next section.

Despite density fluctuations, the semi-global density is still related to the single global density under the large R_0 value as long as eqs 10 and 11 remain valid. This connection between two densities allows for resolving the transferability issue; this will be examined in the next subsection. In turn, even though we adopt the mean-field approach, we emphasize that p_I can still account for fluctuations at the CG site I during the CG simulation because of changes in the density ρ_I around the mean-field density $\langle\rho\rangle$, which is directly related to the global property of the system.

2.4. Connection between UCG and FED.

To ensure a more complete transferability of the UCG model, we constructed the effective UCG interaction in a *hybrid* manner: bulk interactions that are parametrized at different temperatures using MS-CG are utilized as substate interactions in eq 12. We aim to obtain CG interactions at different systems in equilibrium and couple these CG interactions using an order parameter (semi-global density in this case). However, it should be noted that one can also rescale the FG canonical configuration integrals $\int \mathrm{d}\mathbf{r}^n \exp[-\beta U(\mathbf{r}^n)]$ to numerically transfer a single CG interaction to different temperatures.[27]

In eq 12, we define α as the bulk state at T_{low} and β as the bulk state at T_{high}. In other words, the $\alpha\alpha$ substate interaction corresponds to interactions in bulk at T_{low}, and the $\beta\beta$ substate interaction corresponds to interactions in bulk at T_{high}. Then, naturally, there exists an intermediate temperature T_{int} that satisfies $T_{\mathrm{low}} < T_{\mathrm{int}} < T_{\mathrm{high}}$ and which corresponds to the $\alpha\beta$ cross-interaction. Taken together, the mean-field UCG free energy can be cast as

$$\overline{U_{\mathrm{mix}}^{\mathrm{UCG}}}(\mathbf{R}^N) = \sum_{I,J}\Big[\langle p_\alpha\rangle^2 U_{T=T_{\mathrm{low}}}(R_{IJ}) + \langle p_\beta\rangle^2 U_{T=T_{\mathrm{high}}}(R_{IJ}) + 2\langle p_\alpha\rangle\langle p_\beta\rangle U_{T=T_{\mathrm{int}}}(R_{IJ})\Big] \tag{16}$$

Introducing the FED scheme from eq 3 to the right-hand side of eq 6, one finds

$$\begin{aligned}\overline{U_{\text{mix}}^{\text{UCG}}}(\mathbf{R}^N) &= \sum_{I,J}\Big[\langle p_\alpha\rangle^2(\Delta H(R_{IJ}) - P_{\text{low}}\Delta V(R_{IJ}) - T_{\text{low}}\Delta S(R_{IJ}))\\ &+ \langle p_\beta\rangle^2(\Delta H(R_{IJ}) - P_{\text{high}}\,\Delta V(R_{IJ}) - T_{\text{high}}\,\Delta S(R_{IJ}))\\ &+ 2\langle p_a\rangle\langle p_\beta\rangle(\Delta H(R_{IJ}) - P_{\text{int}}\,\Delta V(R_{IJ}) - T_{\text{int}}\,\Delta S(R_{IJ}))\Big]\\ &= \sum_{I,J}\Big[\Delta H(R_{IJ})\big(\langle p_\alpha\rangle^2 + \langle p_\beta\rangle^2 + 2\langle p_\alpha\rangle\langle p_\beta\rangle\big)\\ &- \Delta S(R_{IJ})\big(T_{\text{low}}\langle p_\alpha\rangle^2 + T_{\text{high}}\langle p_\beta\rangle^2 + 2T_{\text{int}}\langle p_\alpha\rangle\langle p_\beta\rangle\big) - P\Delta V(R_{IJ})\Big]\end{aligned} \tag{17}$$

In eq 17, we used a constant pressure condition $P_{\text{low}} = P_{\text{high}} = P_{\text{int}} = 1$ atm. Because $T_{\text{low}} < T_{\text{int}} < T_{\text{high}}$, there exists $\vartheta \in (0, 1)$ such that $T_{\text{int}} = \vartheta T_{\text{low}} + (1 - \vartheta)T_{\text{high}}$, and

$$\begin{aligned}&\big(T_{\text{low}}\langle p_\alpha\rangle^2 + T_{\text{high}}\langle p_\beta\rangle^2 + 2T_{\text{int}}\langle p_\alpha\rangle\langle p_\beta\rangle\big)\\ &= \big(T_{\text{low}}\big(\langle p_\alpha\rangle + 2\big(1 - \vartheta\langle p_\beta\rangle\big)\langle p_\beta\rangle\big)\langle p_\alpha\rangle + T_{\text{high}}\big(\langle p_\beta\rangle + 2\vartheta\langle p_\alpha\rangle\big)\langle p_\beta\rangle\end{aligned} \tag{18}$$

The summation of the coefficients from eq 18, $(\langle p_\alpha\rangle + 2(1 - \vartheta)\langle p_\beta\rangle)\langle p_\alpha\rangle$ and $(\langle p_\beta\rangle + 2\vartheta\langle p_\alpha\rangle)\langle p_\beta\rangle$, further satisfies

$$\begin{aligned}&\big(\langle p_\alpha\rangle^2 + 2\langle p_\alpha\rangle\langle p_\beta\rangle - 2\vartheta\langle p_\alpha\rangle\langle p_\beta\rangle + \langle p_\beta\rangle^2 + 2\vartheta\langle p_\alpha\rangle\langle p_\beta\rangle\big)\\ &= \big(\langle p_\alpha\rangle^2 + 2\langle p_\alpha\rangle\langle p_\beta\rangle + \langle p_\beta\rangle^2\big)\\ &= 1\end{aligned} \tag{19}$$

Therefore, the UCG model based on different temperatures can be interpreted as a single CG model at an intermediate temperature given below, which is determined by an ensemble average of the semi-global density and substate probability $\langle p_\alpha\rangle$ in the mean-field limit

$$T_{\text{UCG}} = \big(T_{\text{high}} - T_{\text{low}}\big) \times \left(\left\langle p_\alpha \right\rangle + 2\left(\frac{T_{\text{high}} - T_{\text{int}}}{T_{\text{high}} - T_{\text{low}}}\right) \cdot \big(1 - \langle p_\alpha\rangle\big)\right) \tag{20}$$

It should be noted that formulating temperature transferable CG models using the UCG framework is possible because of the linearity in (i) coupling enthalpic and entropic terms in the free energy and (ii) coupling between the UCG statewise interactions.

2.5. Alternative CG Model: Empirical Design Principle for Temperature Transferability.

An alternative approach to derive temperature transferable CG potentials has been suggested for CG polymers, in which the CG potentials at different temperatures are empirically rescaled on the basis of temperatures under constant pressure conditions[66,67]

$$U^{\mathrm{emp}}(R,T) = U\left(R,T_0\right) \times \sqrt{\frac{T}{T_0}} \tag{21}$$

Here, we use $U^{\mathrm{emp}}(R,T)$ to denote the rescaled PMF with respect to the reference temperature T_0, which differs from the true PMF $U(R,T)$. We note that the proposed scaling rule in eq 21 is valid only if the entropic contribution from $U(R,T)$ obeys

$$\Delta S(R) \approx U\left(R,T_0\right) \cdot \frac{5T_0 - T}{8T_0^2} \tag{22}$$

A detailed derivation for eq 22 is given in the Supporting Information Section S2. Although this empirical scaling scheme has shown to be useful for polymer systems and hexane, it is unclear if eq 21 would provide accurate CG interactions for liquids because of strong approximations made in eq 22. Thus, we assess the applicability of the empirical design principle in liquid systems and compare performances to that of the systematic approach in Section 3.1.

3. RESULTS

3.1. Temperature Transferable CG Model for Bulk Liquids.

3.1.1. Changes in CG Interaction over Different Temperatures.—Methanol is one of the most extensively studied liquid systems in bottom-up CG modeling.[7,24,25,30,68] Acetone is also a polar liquid, but it does not have any hydrogen bonding, resulting in different energetics and pair structures compared to methanol.[69] We considered single-site CG models for both liquid methanol and acetone to develop transferable CG models for simple liquids. Thus, we could distinguish CG interactions that originate from different FG energetics: hydrogen bonding (methanol) and van der Waals-type (acetone).

Methanol-methanol and acetone-acetone pair CG interactions were computed from different temperature ranges based on the ensemble conditions. For methanol, in Figure 3a and 3b, we selected a temperature range of 225–400 K for the constant *NVT* ensemble and 200–400 K for the constant *NPT* ensemble. For acetone, because of different melting and boiling points, we adopted 225–375 K under the constant *NVT* and 150–400 K under constant *NPT* dynamics, as shown in Figure 3c and 3d. For both constant *NVT* ensembles, we used the fixed volume from the $T = 300$ K condition under $P = 1$ atm. Pressures for constant pressure simulations were maintained as $P = 1$ atm. Full details for FG and CG simulations are available in the Supporting Information Section S3. For these temperature ranges, we note that all simulated CG models have standard liquid structures.

In Figure 3, different ensemble conditions result in distinct changes in the CG interaction as temperature varies. A notable trend is that $U_{\mathrm{CG}}(R)$ seems increasing with temperature (i.e., $(\Delta U_{\mathrm{CG}}(R)/\Delta T)|_{NVT} > 0$) under constant *NVT* conditions, whereas an opposite trend is observed for the constant *NPT* systems (and then the *NVT* runs were performed for analysis) in both liquids: $(\Delta U_{\mathrm{CG}}(R)/\Delta T)|_{NPT} < 0$. Whereas the temperature derivative of the CG PMF

under fixed volume conditions is expected to be positive,[24–26] the CG PMF with a negative temperature derivative at constant pressure is surprising at first glance. The positive $(\Delta U_{\mathrm{CG}}(R)/\Delta T)|_{NVT}$ value can be understood from the definition of the mapping entropy during the CG process where missing modes beneath the CG resolution are integrated into the entropic component of the CG PMF: $-\Delta S(R) \coloneqq S_V(R)$.[24–26] The subscript V denotes the mapping entropy carried out under the constant volume conditions. However, the Helmholtz free energy term obtained from the constant pressure simulations behaves differently because the ΔV term in eq 5 will no longer be zero. This gives rise to $[(\Delta H - P\Delta V)/(\Delta T)]_{\mathrm{MS-CG}} = -P\cdot(\Delta V/\Delta T) \neq 0$. Therefore, a natural extension of eq 5 suggests that $(\Delta U_{\mathrm{CG}}(R)/\Delta T)|_{NPT}$ can be cast as

$$\left.\frac{\Delta U_{\mathrm{CG}}(R)}{\Delta T}\right|_{NPT} = S_V(R) - P\frac{\Delta V}{\Delta T} \tag{23}$$

From eq 23, the additional $-P\cdot(\Delta V/\Delta T)$ term has a negative sign, resulting in the opposite sign of the temperature derivative of the CG PMF. This is consistent with another study reported last year.[30] From Figure 3b, we find that the $P\cdot(\Delta V/\Delta T)$ does not depend on the temperature or the $-\Delta S(R)$ term as the CG PMF is equally spaced. In this sense, we can further associate $P\cdot(\Delta V/\Delta T)$ from inspection of eq 23 as a mapping entropic term that originates from volume differences

$$S_{\Delta \mathrm{V}}(R) \coloneqq P\frac{\Delta V}{\Delta T} \tag{24}$$

In turn, the contributions from $S_V(R)$ and $S_{\Delta V}(R)$ complicate the development of transferable CG models among various state points.

3.1.2. Semi-global Density Distributions at Different Temperatures.—

To construct UCG models that provide temperature transferability in the liquid phase, we computed density histograms from different temperatures at constant pressures. According to eqs 17–21, we used a fairly large cutoff $R_0 = 12$ Å with a smoothly decaying function: $\rho_I(R_I) = \sum_{IJ} 1/2\left(1 - \tanh\left((R_{IJ} - 12)/1.2\right)\right)$.

Figure 4 illustrates the changes in density distribution at various temperatures for both liquids. As expected, a general trend is that as temperature increases, the volume of the system also increases, resulting in a lower overall density.

There are two other notable features in the semi-global density histograms that are witnessed in Figure 4. First, the densities are normally distributed, i.e., Gaussian distributions. Although other density distributions with a shorter distance cutoff often exhibit bimodal or more complicated distributions, the normal distributions, here, suggest that the proposed order parameter can be utilized to quantify the global properties of the system.[64,65] Changes to the average density from each histogram across temperatures are gradual and monotonic, as noted in Table 1. We now validate whether the order parameter is suitable as a global order parameter in the UCG models.

3.1.3. Semi-global Order Parameter: Validation.—The main assumption that links UCG theory to FED requires the locally computed order parameter to directly respond to the global properties of the system. In such a case, a "semi-global" number density variable $\langle\rho\rangle$ can be further simplified as follows

$$\left\langle \rho \right\rangle \approx \rho_{\mathrm{b}} \cdot V_{\mathrm{density}} = \frac{N_{\mathrm{bulk}}}{V_{\mathrm{bulk}}} \cdot V_{\mathrm{density}} \tag{25}$$

The equality holds if and only if the system is homogeneous and densities are evenly distributed. Therefore, if we equate $\langle\rho\rangle$ to $(N_{\mathrm{bulk}}/V_{\mathrm{bulk}})\cdot V_{\mathrm{density}}$, we obtain V_{density} as an idealized volume that gives the number density $\langle\rho\rangle$ of the system. This approximation is valid for methanol, as $\langle\rho\rangle$ and $1/V_{\mathrm{bulk}}$ vary linearly (Table 1, Figure 5a). Thus, a density radius from the idealized volume is represented as

$$R_0^{\rho} = \left(\frac{3\langle\rho\rangle}{4\pi} \cdot \frac{V_{\mathrm{bulk}}}{N_{\mathrm{bulk}}} \right)^{1/3} \tag{26}$$

Figure 5b depicts the trend of the density radius R_0^{ρ} over different temperatures. Different from the semi-global density where $\langle\rho\rangle$ decreases as temperature increases, we find that the R_0^{ρ} value is nearly constant as the temperature changes and is almost equal to the density cutoff R_0, indicating that our locally computed order parameter also acts as a global property. We evaluate these assumptions for acetone in the Supporting Information Section S4. In spite of small changes in R_0^{ρ} for acetone, we conclude that this semi-global approximation holds for acetone as well.

To summarize, we conclude that the proposed mean-field ansatz can be readily applied for the chosen order parameter, indicating temperature transferability due to the connection to the FED scheme. This connection is possible because the UCG theory allows for one to couple the semi-global density and pair interaction term multiplicatively. We also tested if transferability can be achieved additively using a recently developed theory.[70] As expected, although an additive coupling in CG models may capture structural correlations correctly, it does not produce transferable CG interactions over different temperatures (see the Supporting Information Section S6).

3.1.4. Temperature Transferable UCG Model: Results.—Based on Figure 4, we constructed the UCG model of liquids using bulk interactions at different temperatures. To encompass all possible liquid temperature ranges, we chose more dense-more dense interaction as the bulk interaction at the lowest temperature and less dense-less dense interaction as the highest temperature bulk interaction. Then, the intermediate interaction was chosen as the CG interaction where its temperature gives a median density histogram between the highest and lowest temperature. For the case of methanol, this gives T_{low} = 200 K and T_{high} = 400 K with the intermediate temperature of T_{int} = 300 K (see Figure 4a). Then, based on Figure 4a, we used ρ_0 = 105.0 and $\bar{\rho} = 0.1 \cdot \rho_0$ = 10.5 to impose internal UCG

states, followed by eq 8. Final UCG trajectories were generated using the following free energy functional

$$U_{\text{mix}}^{\text{UCG}}(\mathbf{R}^N) = \sum_{I,J}\Big[p_{I,\alpha}p_{J,\alpha}U_{T_{\text{low}}}(R_{IJ}) + p_{I,\beta}p_{J,\beta}U_{T_{\text{high}}}(R_{IJ}) + \left(p_{I,\alpha}p_{J,\beta} + p_{I,\beta}p_{J,\alpha}\right)U_{T_{\text{int}}}(R_{IJ})\Big] \tag{27}$$

Figure 6 presents the RDFs obtained by the UCG models from 200 to 400 K under constant pressure. To evaluate the performance of other CG models, we introduce four CG models: three naïve MS-CG models from T_{low}, T_{int}, and T_{high}, and an empirically scaled CG model described earlier. It is clear from Figure 6 that the naïve model from T_{low} approximates the low-temperature structure relatively well, and the high-temperature naïve model can provide reasonable pair correlations at higher temperatures. However, these naïve models are clearly not transferable over different temperature ranges, highlighting the transferability issue. We also generated the scaled CG interaction from T_0 = 300 K based on the empirical relationship. Although scaled to different temperatures, we discover that the scaled CG interaction still gives similar pair structures as directly employing T_{int} = 300 K interaction, indicating that eq 21 does not impart transferable interactions. It is apparent, however, that the single UCG model enhances the pair correlation, capturing the first minimum near 4 Å with two peaks at 3.5 and 4.5 Å. The UCG theory designed in this work is able to precisely reproduce the pair structures across a temperature range of 200 K.

Similarly, we generated the temperature transferable UCG model for acetone using the following substate interactions T_{low} = 150 K, T_{high} = 400 K, and T_{int} = 275 K (see Figure 4b). The cutoff distance was fixed to R_0 = 12 Å, but the density cutoff was changed to ρ_0 = 49.0 and $\bar{\rho} = 0.1 \cdot \rho_0 = 4.9$ because of the changes in the system size, as can be seen in Figure 4b. RDFs computed from these CG trajectories are depicted in Figure 7. As expected, all of the naïve CG models cannot fully encompass the whole temperature range. Interestingly, the empirically scaled CG models give a fair agreement to the reference correlations, whereas they had failed for methanol. Nevertheless, several deviations are noted near the cusp after the first maximum peak at 6.5 Å. Significant enhancement in pair correlations can be seen in the UCG models at all different temperatures as delineated by red lines in Figure 7. Among different temperatures ranging from 150 K to 400 K, the single UCG model can faithfully recapitulate the pair correlations, even in the long-range regime.

In order to quantitatively examine the performance of various CG models, we introduce here a statistical measure for probing the accuracy of pair correlations: $\Delta(g\{\mathscr{L}\})$, where the error of model $\mathscr{L}$ via $g(\mathbf{r})$ compared to the target RDF $g^{\text{target}}(\mathbf{r})$ is defined as

$$\Delta(g\{\mathscr{L}\}) = \frac{\int_0^{r_{\text{cut}}} |g^{\mathscr{L}}(\mathbf{r}) - g^{\text{target}}(\mathbf{r})| d\mathbf{r}}{\int_0^{r_{\text{cut}}} g^{\text{target}}(\mathbf{r}) d\mathbf{r}} \tag{28}$$

This measure manifests how close is the model $\mathscr{L}$ to the target under the MS-CG framework, bounded from 0 to 1. This error metric is first introduced as a convergence metric to match the RDF[71] and later utilized as a quantified measure of the RDF.[72]

In Figure 8, the relative error, $100 \times \Delta(g\{\mathscr{L}\})$, is visualized, and the associated heat map further confirms the high fidelity of the UCG model for all temperature ranges. For methanol, the UCG model provides the closest RDF values compared to the reference data for all temperature ranges. A similar situation is encountered in acetone where the UCG model improves the accuracy of the existing CG models, except for some temperatures (150, 175, and 250 K) where we witness that the UCG model is less accurate than other UCG models. Nevertheless, it is apparent that the UCG models do a superb job compared to other naïve CG models. Deviations observed at these temperatures can be understood as the imperfection in pairwise basis sets because these deviations also occur in other pairwise CG models.

Finally, we also computed the density histograms from the UCG simulations to check the many-body structural correlations. Because locally computed density is a many-body quantity, density histograms can be regarded as alternative measures to assess *N*-body properties. Figure 9 depicts the semi-global density histogram of the UCG models compared to the reference data, indicating that the UCG models can generally capture the profile of semi-global density with correct widths and tails for both liquids. We notice that normalized counts at each peak differ from the reference model, but the relative error is within 0–20% for both methanol and acetone. Moreover, the averaged density values $\langle\rho\rangle$ are almost identical to the reference value, validating the semi-global hypothesis.

3.1.5. Computational Efficiency.—Computational performance in CG modeling involves two different procedures: parametrization and simulation of the designed CG model. Even though introducing internal states to the CG model might slow down the overall performance compared to the conventional CG model, the present model does not have to be adjusted under any desired conditions of the system once it is built. Namely, without manually adjusting the entropic contribution to the system, the present UCG model instead automatically finds the effective temperature of the system. This feature emphasizes the effectiveness of the model.

Table 2 lists the effective speed-up factors from UCG simulations compared to FG resolutions. The speed-up factor was directly calculated by evaluating the ratio of time needed to complete the fixed amount of time compared to the UCG model. Because the energy landscape of the CG model is smoother than the one at the FG resolution, we adopted a time step of 50 fs for the CG simulation. It is known that the UCG model is designed to calculate the order parameter *on the fly* to adjust its interaction based on the order parameter value, and thus, the integration time step should not be large to perform a stable CG simulation (usually about 10 fs); however, in this work, this large integration step is feasible because we use the semi-global order parameter, which is not highly affected by the local environment.

Table 2 shows that the speed-up factor is ~30–40 for the methanol system and ~90 for acetone systems under the liquid condition.

3.1.6. Extension to Pressure Transferability.—Even though the current model is developed for constant pressure conditions, it would be of great interest to extend it to different pressure conditions in terms of transferability. We further argue that such an extension is also possible because the equilibrium volume of the system will decrease as the system pressure increases. Hence, changes in CG interactions at different pressures can be viewed as if the temperature decreases under the constant pressure conditions. Alternatively, this can be shown from eq 17 using a similar analogy under constant temperature with varying pressure. We tested the pressure transferability by extrapolating the current methanol UCG model to high pressure conditions (see the Supporting Information Section S7) where methanol is still in the liquid phase.[73,74] As shown in Figure S3, we confirmed that the current methodology is also applicable to different pressures that are not previously parametrized without any modifications.

Beyond temperature and pressure, a perfectly transferable CG model should be, in principle, applicable to other conditions. As a next step toward this direction, we aim to develop a phase transferable UCG model in the next subsection.

3.2. Phase Transferable UCG Model.

3.2.1. Gas Phase: Setup and Interaction.—A gas-liquid phase transferable UCG model can be designed if one can adequately differentiate two phases with a proper order parameter. In principle, the semi-global density order parameter can be utilized in this work because there are large differences in the global densities of liquid and gas phases. A UCG design principle would be to model any interaction that appeared in either phase as a linear combination of liquid and gas phase interactions. Without loss of generality, we choose methanol to construct phase transferable UCG models in this section.

One apparent observation from the gas phase simulation is that the simulation volume increases significantly compared to that of the liquid phase. For example, at 500 K, in a system composed of $N = 1000$ molecules, a system box length becomes 181.80 Å, which is 4.5 times larger than the system size at 300 K. From the identical initial configurations, the constant NPT simulation for 1 ns of the gas phase (500 K) is 24.6 times slower than that of the liquid phase (300 K). For the final constant NVT simulation for 5 ns, the gas phase simulation is 11.1 times slower than the liquid phase simulation, indicating that the simulation of 1000 gas molecules would be about 100 times slower than liquid. Hence, we decided to simulate a smaller system to perform more inexpensive gas phase FG simulations. However, it must be checked if the CG interaction in the reduced system retains the same interaction profile. In principle, CG interactions should not change in reduced system sizes at the same density because of periodic boundary conditions. To validate this assumption, we constructed the gas phase of methanol composed of differing numbers of methanol molecules ranging from 100 to 750, as depicted in Figure 10a. We set the target temperature at 500 K, which is certainly beyond the boiling point of methanol, as shown in Figure S4.

As expected, the CG pair potentials remain nearly identical across system sizes, and we chose the smallest system $N = 100$ that reproduces the CG interaction accurately at inexpensive computational costs. To understand how gas phase interactions varied at different temperatures, we also performed simulation at $T = 500$, 550, and 575 K. Furthermore, we obtained interactions at $T = 450$ K, but this phase is near to phase transition limit, as shown in Figure S4. Figure 10b compares the trend from the liquid phase interactions (350 and 400 K) to that of gas phase interactions.

The most striking feature from Figure 10b is that there is an abrupt change in the CG interactions near the phase transition. The repulsive short-ranged interaction profile present in liquid disappears and becomes highly attractive. Moreover, the way in which the CG potentials generally change in response to temperature appears reversed in the gas phase compared to that in the liquid phase. Under constant pressure conditions, it is seen that CG interaction becomes more attractive as temperature increases. However, in the gas phase, the CG interaction becomes more repulsive as temperature increases.

Because of the two peculiar changes during the liquid-gas phase transition, a CG interaction between liquid and gas phases can be expressed by a linear combination between cold liquid interactions (e.g., the repulsive potential from Figure 10b) and warm gas-like interactions (the most attractive potential). In other words, the UCG framework is applied for treating these two interactions as $\alpha\alpha$ and $\beta\beta$ substate interactions and interpolating the intermediate $\alpha\beta$ interaction.

3.2.2. UCG Model Design.—Based on Figure 10, we extended the two-phase UCG model to achieve phase transferability. Our primary goal of the UCG model is to (1) accurately reproduce the changes in the liquid phase under different temperatures and (2) reasonably capture the structural correlations of the gas phase. This is a challenging goal because liquid and gas have noticeably different interaction profiles and the corresponding structural correlations.

The final phase transferable UCG model was constructed using the following substate definitions

$$\rho_I(R_I) = \sum_{IJ} \frac{1}{2}\left(1 - \tanh\left(\frac{R_{IJ} - 12}{1.2}\right)\right) \tag{29a}$$

$$p_{s_I = \{\alpha,\beta\}, I}(R_I) = \frac{1}{2}\left(1 \pm \tanh\left(\frac{\rho_I(R_I) - 65}{42.5}\right)\right) \tag{29b}$$

As in the previous section, the choice of $R_0 = 12$ Å and $\bar{R} = 0.1R_0 = 1.2$ Å is invariant. For the new definition of internal states $p_{s_I = \{\alpha,\beta\}, I}(R_I)$, we chose $\rho_0 = 65$ and $\bar{\rho} = 42.5$ to effectively differentiate liquid and gas phases.

Figure 11 provides a general landscape of the designed phase transferable UCG model where there are two substate probabilities that represent the liquid state (as density increases) and gas state (as density decreases). As discussed above, to fully enhance the transferability of the UCG model, we directly insert CG potentials obtained from the bulk system into the UCG substate interactions. The choice of CG potentials was based on Figure 11: more dense-more dense interaction is modeled by bulk CG interaction at T_{low} = 250 K and less dense-less dense interaction is borrowed from T_{high} = 525 K. Finally, the intermediate interaction (more dense-less dense) is chosen from T_{int} = 450 K, where T_{int} is located between the liquid and gas phases.

3.2.3. Results.—From the UCG trajectories simulated at different temperatures, we calculated the RDF to evaluate the performance of the UCG model. To emphasize the different pair correlations emergent in gas, a naïve liquid CG model is prepared from 300 K and directly applied to the gas phase.

From Figure 12, it is evident that the phase transferable model can successfully bridge between liquid and gas phases. The designed UCG model can clearly distinguish two different structures for the liquid (Figure 12a–d) and gas states (Figure 12e–h). As expected, pair correlations in gas are less structured, whereas naïve models of liquids are not capable of capturing either peak intensity or overall correlation profile. Furthermore, we highlight that the phase transferable UCG model can recapitulate the changes in liquid pair correlations for different temperatures, where the peak becomes less structured and separated as temperature increases. Altogether, the phase transferable UCG model distinguishes structural correlations that originate from differences in phases as well as the temperature dependent changes in the structure.

3.2.4. Mean-Field Explanation.—Given the performance demonstrated in Figure 12, phase and temperature transferability of the present UCG model is not an accident. The high fidelity of the UCG model can be attributed to the mean-field description introduced in the underlying ansatz. Because the density distributions follow the Gaussian form because of semi-globality, one can effectively extract the ensemble-averaged UCG PMF for each system. Recall that the mean-field expression of the UCG PMF for this system is given by

$$\overline{U_{\text{mix}}^{\text{UCG}}}\left(\mathbf{R}^{\text{N}}, T\right) = \sum_{I,J}\Big[\langle p_\alpha\rangle^2 U_{T=250\text{K}}(R_{IJ}) + \langle p_\beta\rangle^2 U_{T=525\text{K}}(R_{IJ}) + 2\langle p_\alpha\rangle\langle p_\beta\rangle U_{T=450\text{K}}(R_{IJ})\Big] \tag{30}$$

With an appropriately projected UCG PMF, we compare $\overline{U_{\text{mix}}^{\text{UCG}}}\left(\mathbf{R}^{\text{N}}, T\right)$ and $U_{\text{CG}}(\mathbf{R}^{\text{N}}, T)$ at different temperatures and phases, as shown in Figure 13.

Although the CG interaction potential for liquid and gas phases are distinct, the average UCG interaction can adjust the PMF in response from 300 to 575 K. A slight deviation in the gas phase is understood from having only two states for the UCG model, but the differences are relatively small. In the liquid phase, we found that the phase transferrable

UCG model behaves similarly as our liquid temperature transferable model. Again, this performance is a result of the bimodal characteristics observed in Figure 10b, enabling the UCG model to span intermediate interactions in a finely detailed manner.

3.3. UCG Phase Combinational Model.

3.3.1. FG System Design.—In the aforementioned sections, we constructed UCG models that can be transferred across different temperatures (1) in liquids (2) in different phases (liquid and gas phases). In spite of its good performance, there was a clear limitation that these models were developed for homogeneous systems. Hence, an end goal would be extending the same methodology to a heterogeneous system, i.e., a liquid/vapor interface. Given the prevalence and importance of interfaces,[75–78] a high fidelity CG model should encompass the liquid/liquid and liquid/vapor interfaces, including those present in complex biomolecular systems, e.g., lipids and membranes.[79,80] Following our previous work on interfacial systems,[36] we adopted the same protocol to generate the liquid/vapor interface of methanol. A detailed setup to build the atomistic configurations of the interface is elaborated in Figure S1.

3.3.2. UCG Model Design.—We generalized the present UCG theory to treat inhomogeneous interfacial systems comprising both liquid and gas phases. A structural profile of the interfacial system can be identified by the slab density profile along the axis normal to the surface.[81] Figure 14 highlights this density profile, characterizing two major regimes underlying the interface: inner and outer regions. The inner region corresponds to a liquid-like phase with high number density, whereas the regime with low number density (nearly zero) corresponds to the outer region. The phase boundary is an intermediate region, located between the inner and outer regions with a monotonically decaying number density. The link between the number density and the corresponding regimes suggests that a UCG model using the semi-global density order parameter is possible. In particular, the nonzero probability force term $-\nabla p$ in the UCG force originating from the semi-global density fluctuation is expected to modulate the phase coexistence in an inhomogeneous system with the present UCG theory without further modification.

In prior work, we demonstrated the capability of the local density-based UCG model for the various interfacial systems, although transferability was somehow limited to only CG force terms because of non-negligible artifacts introduced in the long-range interactions because these interactions were affected by the phase boundary of the interface.[36] In this work, we resolve the limited transferability issue using the theory developed in the current approach. As mentioned in prior sections, we do not explicitly force-match the UCG interaction from interfacial systems. Instead, we directly employ the bulk MS-CG interaction under different temperature (volume) conditions as the UCG substate interaction. In this way, the UCG statewise interactions do not contain the long-range artifact from the phase boundary and no longer suffer from the long-range force contributions to the CG potentials. The goal of this section is to provide a complete description of liquid/vapor interfaces modeled by the mean-field UCG methodology constructed from other homogeneous systems: a phase combinational model.

Because the liquid/vapor interfaces are a heterogeneous mixture of liquid and vapor systems, we argue that the design principle of the presented phase transferable UCG model can be readily utilized without further modifications. The only difference between them is that more dense-more dense statewise interactions are borrowed from bulk liquid at 300 K in this design because the atomistic system was prepared at 300 K. The choice of other statewise interactions is invariant. The finalized scheme is depicted in Figure 14 where the resultant UCG model is governed by the following interaction form

$$U_{\mathrm{mix}}^{\mathrm{UCG}}\left(\mathbf{R}^{N}\right)=\sum_{I,J}\Big[p_{I,\alpha}p_{J,\alpha}U_{T=300\mathrm{K}}\left(R_{IJ}\right)+p_{I,\beta}p_{J,\beta}U_{T=450\mathrm{K}}\left(R_{IJ}\right)+\left(p_{I,\alpha}p_{J,\beta}+p_{I,\beta}p_{J,\alpha}\right)U_{T=425\mathrm{K}}\left(R_{IJ}\right)\Big] \tag{31}$$

After defining the interaction form, the last step to design the UCG model is the substate probability $p_{s_I=\{\alpha,\beta\},I}\left(R_I\right)$ and the corresponding density functions $\rho_I(R_I)$. As opposed to the bulk systems, designing the density threshold and distance cutoff may not be straightforward because of the inhomogeneous nature of the system and should be different from the bulk case. In contrast to the bulk density cutoff $R_0^{\mathrm{bulk}}=12\,\mathring{A}$, we used a short cutoff value to avoid counting the particles across the inner and outer regions at the phase boundary. We chose $R_0^{\mathrm{interface}}=6\,\mathring{A}$ where it does not exceed half the length of the phase boundary ranging from 38 < z < 50 Å, as shown in Figure 14. From the $R_0^{\mathrm{interface}}$ value, the variable ρ_0 is determined by the density histograms along the slab axis to impose bimodality. It is expected that the interfacial system exhibits three different density profiles corresponding to the inner region, phase boundary, and outer region. Because the system is no longer periodic along the slab axis, we follow the previous work[36] and evaluate the density histogram by constructing a quasi-periodic subsystem where the subsystem is composed of CG particles located in the specific z value range. Within a quasi-periodic condition, the density histograms are then computed, as seen in Figure 15.

In Figure 15, two curves computed from 10 < z < 20 Å and 20 < z < 30 Å, corresponding to the inner region, are almost identical. A bimodal shape from the inner region (peaks near ρ_1 = 8 and ρ_2 = 12) suggests that this might be further decomposed into liquid-like phase density and phase boundary density. For the outer region, in an interval of both 50 < z < 60 Å and 60 < z < 70 Å, we see almost zero density as a gas phase. The intermediate region 32 < z < 42 Å fills this gap where the intermediate density locates exactly between the inner and outer region curves. Interestingly, the average density in the phase boundary is near ρ_1 = 8, confirming our initial hypothesis about bimodality. Because the outer region has mostly zero density during the simulation, we selected $\rho_0^{\mathrm{interface}}$ as 11.5 to effectively distinguish the inner region and phase boundary.[36]

Finally, we arrive at the UCG phase combinational defined by the following definitions

$$\rho_I(R_I) = \sum_{IJ} \frac{1}{2}\left(1 - \tanh\left(\frac{R_{IJ} - 6}{0.6}\right)\right) \tag{32a}$$

$$p_{s_I = \{\alpha,\beta\},I}(R_I) = \frac{1}{2}\left(1 \pm \tanh\left(\frac{\rho_I(R_I) - 11.5}{1.15}\right)\right) \tag{32b}$$

3.3.3. Results.—We tested the fidelity of the phase combinational UCG model by evaluating two major structural characteristics: the slab density profile and pair correlation function. The calculated average distribution along the *z*-axis is shown in Figure 16a. For the conventional MS-CG model, the density profile follows a bell-like Gaussian shape rather than the trapezoidal profile observed in the all-atom reference system.[36,53,70] Furthermore, the MS-CG model cannot fully capture the vapor phase where the number density along 50–70 Å is higher than zero. The UCG model can ameliorate these problems by improving both the inner and outer regions. It can be discerned that there are two shoulders near the phase boundary (near 0 and 40 Å) where the liquids are strongly structured compared to the reference system. Although these structured shoulders were previously reported by other CG modeling studies,[36,53] a clear physical explanation underlying this anomaly has not been suggested in the literature and thus should be addressed.

We propose here that this deviation can be explained because CG interactions in the gas phase are much more attractive than CG interactions in the liquid phase.[82,83] As depicted in Figure 10, a gas phase contains fewer molecules but has stronger interactions. This tends to keep the gas phase stabilized in an enlarged system compared to the liquid system. However, at the phase boundary, the CG particle becomes more attractive as it enters a gas-like state, whereas the number density is not as low as in the gas phase. Hence, the CG particles experiencing attractive interactions at the phase boundary become more strongly structured, resulting in a shoulder-like anomaly. However, even with this shouldered profile, the UCG model correctly reproduces the liquid/vapor phase separation.

To demonstrate that interactions at the interface are more complex than the interactions obtained in bulk, we performed the MS-CG simulations at three different temperatures that correspond to the temperature of statewise interactions: 300, 425, and 450 K. The slab density profiles from these naïve MS-CG models are shown in Figure 16b. As expected, we find that all naïve models produce incorrect interfacial structures. To be specific, naïve models parametrized at both 300 and 425 K fail to capture phase separation and generate a homogeneous fluid instead. Another naïve model from 425 K seems to reproduce a signature of the interface, but the resultant profile greatly deviates from both the reference and UCG models. This performance is quantified by fitting the density profiles to the following function based on the previous studies[84,85]

$$\rho(z) = \rho_{\mathrm{v}} + (\rho_1 - \rho_{\mathrm{v}})\frac{1 - \tanh\frac{2z+\omega}{\delta}\tanh\frac{2z-\omega}{\delta}}{1 + \tanh^2\frac{\omega}{\delta}} \tag{33}$$

where the widths of the slab and interface are represented as ω and δ (units in Å), and the phase densities at the liquid and vapor phases are represented as ρ_l and ρ_v (units in $Å^{-1}$), respectively. The fitted parameters for the slab density profile using two sigmoid functions are listed in Table 3 given below.

From Table 3, we observe that the UCG model notably enhances the description of δ, ρ_l, and ρ_v values from the all-atom values, as well as an incremental improvement seen in slab width ω. In conjunction with Figure 16b, all naïve models completely fail to reproduce any of the parameters that are required to describe the slab structure.

In addition to the slab density profile, we calculated $g(r)$ functions for the models we studied here. Because of the inhomogeneous nature of interfacial systems, a direct calculation of $g(r)$ results in an incorrect normalization in the large distance limit. To mitigate this problem, we followed the same approach to construct the quasi-periodic system used in Figure 15 that gives a correct asymptotic behavior. The adjusted $g(r)$ from the inner region of the interface is depicted in Figure 17.

The introduced scheme gives a correctly converged $g(r)$ value of unity for all simulations. Of particular note, a general trend demonstrated by the inner region of the atomistic simulation is equivalent to $g(r)$ from bulk methanol, confirming the validity of our quasi-periodic scheme. As expected, the naïve CG models generally have poor performance in which the inner region is too much or less structured. Interestingly, the correlation obtained from 425 K appears to fairly close to the atomistic reference data from Figure 17, but Figure 16b shows that this model cannot reproduce the slab structure correctly. The UCG model, however, is the only model where local pair correlations containing two peaks are both quantitatively and qualitatively captured while maintaining the slab structure.

This analysis suggests that the modeling of CG interfaces may be quite sensitive. At the phase boundary, the number density starts to decay, whereas the CG interaction becomes attractive. An accurate high fidelity CG model for an interface must account for such a trade-off between number density and interaction strength. One possible way of modeling this behavior is to introduce a third state for the UCG model. In this way, the third state can bridge between liquid-like and vapor-like states, balancing the right trade-off. This idea is also supported by Figure 16, which shows the existence of the intermediate state (blue line corresponding to the 32–42 Å regime) between the gas-like and liquid-like densities. Nevertheless, the overall performance exhibited by the UCG model clearly suggests that having only two internal states can faithfully reproduce the structural properties that originate from the bimodal nature of liquid and vapor states.

3.3.4. Computational Efficiency.—Finally, we evaluate the computational performance of the phase combinational UCG model by calculating the effective speed-up factor. Benchmark calculations were performed with the same protocol used in the earlier section. Effective speed-up values are directly obtained by the ratio of relative time per step between the all-atom and UCG simulations, as listed in Table 4. Here, we used a time step of 30 fs to propagate the UCG model. We see that the effective speed-up factor is about 100 times, and this factor is larger than the previous UCG models shown in Table 2. We attribute

this enhanced performance to the following factors: (1) the interface system is not as compact as the homogeneous system and (2) the density cutoff value is half the bulk case. To conclude, the UCG model presented in this work can substantially enhance the physical range of a CG model, encompassing not only the single-phase but also their inter-phases while reproducing structural correlations both qualitatively and quantitatively.

4. CONCLUSIONS

In this paper, we employed our recently developed UCG framework to construct a single CG model which is temperature and phase transferable. Bottom-up CG models suffer from transferability issues when the interaction potential should change in response to the thermodynamic variables of the system. Even though the UCG methodology can greatly enhance the many-body correlations lacking in conventional CG models, prior UCG applications have been limited in their ability to impart transferability to CG models. Here, we developed a mean-field UCG ansatz expressed by a semi-global order parameter. Furthermore, we corroborated that the UCG theory provides an FED of the many-body CG PMF, confirming that the UCG framework naturally imparts temperature transferable CG models. In particular, we utilized the bulk CG interactions at different temperatures as the UCG statewise interaction, maintaining the theoretical rigor of the UCG theory while extending its transferability. In a nutshell, Figure 18 summarizes the overall protocols that are introduced in this work.

By utilizing the effectiveness of the proposed theory to the maximum extent, we designed three distinct UCG models: liquid, gas, and interface. The first demonstration of the methodology was to design temperature transferrable liquid state CG models under constant pressure conditions. We found that the UCG temperature transferable model is not only predictive of the effective CG interactions at different temperatures but it also quantitatively captures the structural correlations of the underlying FG models. Notably, no manual adjustment of the CG potential was needed to simulate the model at different temperatures. By going one step further, we constructed a phase transferable UCG model encompassing both liquid and gas phases. Even with just two internal UCG states, the UCG phase transferable model reproduces structural correlations over a wide range of temperatures (ΔT = 300 K). The accuracy of the model was attributed to the correct recapitulation of the CG PMF using the mean-field UCG ansatz. Finally, we expanded our UCG theory to describe a heterogeneous liquid/vapor interface. Although two distinct phases coexist in the same system, our UCG method can be applied without any alterations. Unlike conventional CG models, the correct phase coexistence feature with accurate pair correlations is observed for the UCG model. We emphasize that there are currently no other unified bottom-up CG models that are transferrable to different thermodynamic states and can maintain a similar degree of accuracy and computational performance.

This paper opens several interesting avenues for future research. The current model can be improved by including an additional internal state that bridges the liquid and vapor states. By including the transitional behavior of liquid to vapor phases, we would expect this three-state UCG model to increase the accuracy and transferability of our CG modeling. Furthermore, the ultimate goal along this direction is to construct a single UCG model that is

completely transferrable across the solid, liquid, and vapor phases. Naturally, this requires a number of internal UCG states, i.e., a minimum of three (for the solid, liquid, and vapor states) and a maximum of five (including two phase transitional regions). Another important direction is to design transferable bottom-up CG models of more complex biomolecules and material systems. Given that the direct bottom-up parametrization of the large-scale systems is computationally demanding, an effective approach could be to divide such large molecules into several building blocks, e.g., functional groups or amino acids. One could combine these higher resolution UCG models into a single, transferable multistate UCG model that can span larger spatiotemporal scales. The optimization of computationally efficient multistate UCG models will be pursued in future studies. Altogether, these research directions may lead to a single, unified approach for transferable “bottom-up” CG modeling.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work is based on work supported by the National Science Foundation (NSF Grant CHE-1465248). Simulations were performed using computing resources provided by the University of Chicago Research Computing Center (RCC). We gratefully acknowledge Dr. Yining Han for stimulating conversations on this project.

REFERENCES

(1). Saunders MG; Voth GA Coarse-Graining Methods for Computational Biology. Annu. Rev. Biophys 2013, 42, 73–93. [PubMed: 23451897]

(2). Noid WG Systematic Methods for Structurally Consistent Coarse-Grained Models. Biomolecular Simulations; Springer, 2013; pp 487–531.

(3). Shell MS Coarse-Graining with the Relative Entropy. Adv. Chem. Phys 2016, 161, 395–441.

(4). Lyubartsev AP; Laaksonen A Calculation of Effective Interaction Potentials from Radial Distribution Functions: A Reverse Monte Carlo Approach. Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top 1995, 52, 3730.

(5). McGreevy RL; Pusztai L Reverse Monte Carlo Simulation: A New Technique for the Determination of Disordered Structures. Mol. Simul 1988, 1, 359–367.

(6). Izvekov S; Voth GA A Multiscale Coarse-Graining Method for Biomolecular Systems. J. Phys. Chem. B 2005, 109, 2469–2473. [PubMed: 16851243]

(7). Izvekov S; Voth GA Multiscale Coarse Graining of Liquid-State Systems. J. Chem. Phys 2005, 123, 134105. [PubMed: 16223273]

(8). Noid WG; Chu J-W; Ayton GS; Krishna V; Izvekov S; Voth GA; Das A; Andersen HC The Multiscale Coarse-Graining Method. I. A Rigorous Bridge between Atomistic and Coarse-Grained Models. J. Chem. Phys 2008, 128, 244114. [PubMed: 18601324]

(9). Noid WG; Liu P; Wang Y; Chu J-W; Ayton GS; Izvekov S; Andersen HC; Voth GA The Multiscale Coarse-Graining Method. II. Numerical Implementation for Coarse-Grained Molecular Models. J. Chem. Phys 2008, 128, 244115. [PubMed: 18601325]

(10). Lu L; Izvekov S; Das A; Andersen HC; Voth GA Efficient, Regularized, and Scalable Algorithms for Multiscale Coarse-Graining. J. Chem. Theory Comput 2010, 6, 954–965. [PubMed: 26613319]

(11). Shell MS The Relative Entropy Is Fundamental to Multiscale and Inverse Thermodynamic Problems. J. Chem. Phys 2008, 129, 144108. [PubMed: 19045135]

(12). Chaimovich A; Shell MS Relative Entropy as a Universal Metric for Multiscale Errors. Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top 2010, 81, 060104.

(13). Chaimovich A; Shell MS Coarse-Graining Errors and Numerical Optimization Using a Relative Entropy Framework. J. Chem. Phys 2011, 134, 094112. [PubMed: 21384955]

(14). Reith D; Pütz M; Müller-Plathe F Deriving Effective Mesoscale Potentials from Atomistic Simulations. J. Comput. Chem 2003, 24, 1624–1636. [PubMed: 12926006]

(15). Noid WG; Chu J-W; Ayton GS; Voth GA Multiscale Coarse-Graining and Structural Correlations: Connections to Liquid-State Theory. J. Phys. Chem. B 2007, 111, 4116–4127. [PubMed: 17394308]

(16). Thirumalai D; Klimov DK Deciphering the Timescales and Mechanisms of Protein Folding Using Minimal Off-Lattice Models. Curr. Opin. Struct. Biol 1999, 9, 197–207. [PubMed: 10322218]

(17). Sorenson JM; Head-Gordon T Toward Minimalist Models of Larger Proteins: A Ubiquitin-Like Protein. Proteins: Struct., Funct., Bioinf 2002, 46, 368–379.

(18). Friedel M; Sheeler DJ; Shea J-E Effects of Confinement and Crowding on the Thermodynamics and Kinetics of Folding of a Minimalist B-Barrel Protein. J. Chem. Phys 2003, 118, 8106.

(19). Marrink SJ; Risselada HJ; Yefimov S; Tieleman DP; De Vries AH The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. J. Phys. Chem. B 2007, 111, 7812–7824. [PubMed: 17569554]

(20). Noid WG Perspective: Coarse-Grained Models for Biomolecular Systems. J. Chem. Phys 2013, 139, 090901. [PubMed: 24028092]

(21). Wagner JW; Dama JF; Durumeric AEP; Voth GA On the Representability Problem and the Physical Meaning of Coarse-Grained Models. J. Chem. Phys 2016, 145, 044108. [PubMed: 27475349]

(22). Dunn NJH; Foley TT; Noid WG Van Der Waals Perspective on Coarse-Graining: Progress toward Solving Representability and Transferability Problems. Acc. Chem. Res 2016, 49, 2832–2840. [PubMed: 27993007]

(23). Potter TD; Tasche J; Wilson MR Assessing the Transferability of Common Top-down and Bottom-up Coarse-Grained Molecular Models for Molecular Mixtures. Phys. Chem. Chem. Phys 2019, 21, 1912–1927. [PubMed: 30632568]

(24). Jin J; Pak AJ; Voth GA Understanding Missing Entropy in Coarse-Grained Systems: Addressing Issues of Representability and Transferability. J. Phys. Chem. Lett 2019, 10, 4549–4557. [PubMed: 31319036]

(25). Lu L; Voth GA The Multiscale Coarse-Graining Method. VII. Free Energy Decomposition of Coarse-Grained Effective Potentials. J. Chem. Phys 2011, 134, 224107. [PubMed: 21682507]

(26). Foley TT; Shell MS; Noid WG The Impact of Resolution Upon Entropy and Information in Coarse-Grained Models. J. Chem. Phys 2015, 143, 243104. [PubMed: 26723589]

(27). Krishna V; Noid WG; Voth GA The Multiscale Coarse-Graining Method. IV. Transferring Coarse-Grained Potentials between Temperatures. J. Chem. Phys 2009, 131, 024103. [PubMed: 19603966]

(28). Das A; Andersen HC The Multiscale Coarse-Graining Method. V. Isothermal-Isobaric Ensemble. J. Chem. Phys 2010, 132, 164106. [PubMed: 20441257]

(29). Agrawal V; Peralta P; Li Y; Oswald J A Pressure-Transferable Coarse-Grained Potential for Modeling the Shock Hugoniot of Polyethylene. J. Chem. Phys 2016, 145, 104903. [PubMed: 27634275]

(30). Lebold KM; Noid WG Systematic Study of Temperature and Density Variations in Effective Potentials for Coarse-Grained Models of Molecular Liquids. J. Chem. Phys 2019, 150, 014104. [PubMed: 30621403]

(31). Voth GA A Multiscale Description of Biomolecular Active Matter: The Chemistry Underlying Many Life Processes. Acc. Chem. Res 2017, 50, 594–598. [PubMed: 28945406]

(32). Dama JF; Sinitskiy AV; McCullagh M; Weare J; Roux B; Dinner AR; Voth GA The Theory of Ultra-Coarse-Graining. 1. General Principles. J. Chem. Theory Comput 2013, 9, 2466–2480. [PubMed: 26583735]

(33). Davtyan A; Dama JF; Sinitskiy AV; Voth GA The Theory of Ultra-Coarse-Graining. 2. Numerical Implementation. J. Chem. Theory Comput 2014, 10, 5265–5275. [PubMed: 26583210]

(34). Dama JF; Jin J; Voth GA The Theory of Ultra-Coarse-Graining. 3. Coarse-Grained Sites with Rapid Local Equilibrium of Internal States. J. Chem. Theory Comput 2017, 13, 1010–1022. [PubMed: 28112956]

(35). Dama JF; Jin J; Voth GA Correction to the Theory of Ultra-Coarse-Graining. 3. Coarse-Grained Sites with Rapid Local Equilibrium of Internal States. J. Chem. Theory Comput 2018, 14, 2288. [PubMed: 29509418]

(36). Jin J; Voth GA Ultra-Coarse-Grained Models Allow for an Accurate and Transferable Treatment of Interfacial Systems. J. Chem. Theory Comput 2018, 14, 2180–2197. [PubMed: 29481754]

(37). Jin J; Han Y; Voth GA Ultra-Coarse-Grained Liquid State Models with Implicit Hydrogen Bonding. J. Chem. Theory Comput 2018, 14, 6159–6174. [PubMed: 30354110]

(38). Golubkov PA; Wu JC; Ren P A Transferable Coarse-Grained Model for Hydrogen-Bonding Liquids. Phys. Chem. Chem. Phys 2008, 10, 2050–2057. [PubMed: 18688358]

(39). Carbone P; Varzaneh HAK; Chen X; Müller-Plathe F Transferability of Coarse-Grained Force Fields: The Polymer Case. J. Chem. Phys 2008, 128, 064904. [PubMed: 18282071]

(40). Mullinax JW; Noid WG Extended Ensemble Approach for Deriving Transferable Coarse-Grained Potentials. J. Chem. Phys 2009, 131, 104110.

(41). Brini E; Van der Vegt NFA Chemically Transferable Coarse-Grained Potentials from Conditional Reversible Work Calculations. J. Chem. Phys 2012, 137, 154113. [PubMed: 23083154]

(42). Dalgicdir C; Sensoy O; Peter C; Sayar M A Transferable Coarse-Grained Model for Diphenylalanine: How to Represent an Environment Driven Conformational Transition. J. Chem. Phys 2013, 139, 234115. [PubMed: 24359360]

(43). Kar P; Gopal SM; Cheng Y-M; Predeus A; Feig M PRIMO: A Transferable Coarse-Grained Force Field for Proteins. J. Chem. Theory Comput 2013, 9, 3769–3788. [PubMed: 23997693]

(44). Kar P; Feig M Recent Advances in Transferable Coarse-Grained Modeling of Proteins. Advances in Protein Chemistry and Structural Biology; Elsevier, 2014; Vol. 96, pp 143–180. [PubMed: 25443957]

(45). Hsu DD; Xia W; Arturo SG; Keten S Thermomechanically Consistent and Temperature Transferable Coarse-Graining of Atactic Polystyrene. Macromolecules 2015, 48, 3057–3068.

(46). Dunn NJH; Noid WG Bottom-up Coarse-Grained Models with Predictive Accuracy and Transferability for Both Structural and Thermodynamic Properties of Heptane-Toluene Mixtures. J. Chem. Phys 2016, 144, 204124. [PubMed: 27250296]

(47). Sanyal T; Shell MS Transferable Coarse-Grained Models of Liquid-Liquid Equilibrium Using Local Density Potentials Optimized with the Relative Entropy. J. Phys. Chem. B 2018, 122, 5678–5693. [PubMed: 29466859]

(48). Rosenberger D; Sanyal T; Shell MS; van der Vegt NF A. Transferability of Local Density Assisted Implicit Solvation Models for Homogeneous Fluid Mixtures. J. Chem. Theory Comput 2019, 15, 2881–2895. [PubMed: 30995034]

(49). McQuarrie DA; Simon JD Physical Chemistry: A Molecular Approach; University Science Books Sausalito, CA, 1997; Vol. 1.

(50). Dannenhoffer-Lafage T; Wagner JW; Durumeric AEP; Voth GA Compatible Observable Decompositions for Coarse-Grained Representations of Real Molecular Systems. J. Chem. Phys 2019, 151, 134115. [PubMed: 31594316]

(51). Dunn NJH; Noid WG Bottom-up Coarse-Grained Models That Accurately Describe the Structure, Pressure, and Compressibility of Molecular Liquids. J. Chem. Phys 2015, 143, 243148. [PubMed: 26723633]

(52). Dunn NJH; Lebold KM; DeLyser MR; Rudzinski JF; Noid WG BOCS: Bottom-up Open-Source Coarse-Graining Software. J. Phys. Chem. B 2017, 122, 3363–3377.

(53). DeLyser MR; Noid WG Extending Pressure-Matching to Inhomogeneous Systems Via Local-Density Potentials. J. Chem. Phys 2017, 147, 134111. [PubMed: 28987122]

(54). Katkar HH; Davtyan A; Durumeric AEP; Hocky GM; Schramm AC; De La Cruz EM; Voth GA Insights into the Cooperative Nature of ATP Hydrolysis in Actin Filaments. Biophys. J 2018, 115, 1589–1602. [PubMed: 30249402]

(55). Grime JMA; Voth GA Highly Scalable and Memory Efficient Ultra-Coarse-Grained Molecular Dynamics Simulations. J. Chem. Theory Comput 2014, 10, 423–431. [PubMed: 26579921]

(56). Grime JMA; Dama JF; Ganser-Pornillos BK; Woodward CL; Jensen GJ; Yeager M; Voth GA Coarse-Grained Simulation Reveals Key Features of HIV-1 Capsid Self-Assembly. Nat. Commun 2016, 7, 11568. [PubMed: 27174390]

(57). Pak AJ; Grime JMA; Yu A; Voth GA Off-Pathway Assembly: A Broad-Spectrum Mechanism of Action for Drugs That Undermine Controlled HIV-1 Viral Capsid Formation. J. Am. Chem. Soc 2019, 141, 10214–10224. [PubMed: 31244184]

(58). Multiscale Coarse-Graining Force-Matching (MSCGFM) v. 1.7 https://github.com/uchicago-voth/MSCG-release (accessed 2020-09-07).

(59). Monaghan JJ Smoothed Particle Hydrodynamics. Annu. Rev. Astron. Astrophys 1992, 30, 543–574.

(60). Monaghan JJ Smoothed Particle Hydrodynamics. Rep. Prog. Phys 2005, 68, 1703.

(61). Monaghan JJ Smoothed Particle Hydrodynamics and Its Diverse Applications. Annu. Rev. Fluid. Mech 2012, 44, 323–346.

(62). Sanyal T; Shell MS Coarse-Grained Models Using Local-Density Potentials Optimized with the Relative Entropy: Application to Implicit Solvation. J. Chem. Phys 2016, 145, 034109. [PubMed: 27448876]

(63). Chandler D Introduction to Modern Statistical Mechanics; Oxford University Press: Oxford, U.K., 1987.

(64). Hsu PL; Robbins H Complete Convergence and the Law of Large Numbers. Proc. Natl. Acad. Sci. U.S.A 1947, 33, 25. [PubMed: 16578237]

(65). Dekking FM; Kraaikamp C; Lopuhaä HP; Meester LE A Modern Introduction to Probability and Statistics: Understanding Why and How; Springer: Heidelberg, Berlin, 2005.

(66). Qian H-J; Carbone P; Chen X; Karimi-Varzaneh HA; Liew CC; Müller-Plathe F Temperature-Transferable Coarse-Grained Potentials for Ethylbenzene, Polystyrene, and Their Mixtures. Macromolecules 2008, 41, 9919–9929.

(67). Farah K; Fogarty AC; Böhm MC; Müller-Plathe F Temperature Dependence of Coarse-Grained Potentials for Liquid Hexane. Phys. Chem. Chem. Phys 2011, 13, 2894–2902. [PubMed: 21157607]

(68). Lebold KM; Noid WG Dual-Potential Approach for Coarse-Grained Implicit Solvent Models with Accurate, Internally Consistent Energetics and Predictive Transferability. J. Chem. Phys 2019, 151, 164113. [PubMed: 31675902]

(69). Sega M; Fábián B; Horvai G; Jedlovszky P How Is the Surface Tension of Various Liquids Distributed Along the Interface Normal? J. Phys. Chem. C 2016, 120, 27468–27477.

(70). Wagner JW; Dannenhoffer-Lafage T; Jin J; Voth GA Extending the Range and Physical Accuracy of Coarse-Grained Models: Order Parameter Dependent Interactions. J. Chem. Phys 2017, 147, 044113. [PubMed: 28764380]

(71). Merabia S; Pagonabarraga I Density Dependent Potentials: Structure and Thermodynamics. J. Chem. Phys 2007, 127, 054903. [PubMed: 17688360]

(72). Fu C-C; Kulkarni PM; Shell MS; Gary Leal L A Test of Systematic Coarse-Graining of Molecular Dynamics Simulations: Thermodynamic Properties. J. Chem. Phys 2012, 137, 164106. [PubMed: 23126694]

(73). Sindzingre P; Klein ML A Molecular Dynamics Study of Methanol near the Liquid–Glass Transition. J. Chem. Phys 1992, 96, 4681.

(74). Gonzalez Salgado D; Vega C Melting Point and Phase Diagram of Methanol as Obtained from Computer Simulations of the OPLS Model. J. Chem. Phys 2010, 132, 094505. [PubMed: 20210403]

(75). Robinson GW; Zhu SB; Singh S; Evans M Water in Biology, Chemistry, and Physics: Experimental Overviews and Computational Methodologies; World Scientific: Singapore, 1996; Vol. 9.

(76). Volkov AG Interfacial Catalysis; CRC Press: Boca Raton, FL, 2002.

(77). Kazarinov VE The Interface Structure and Electrochemical Processes at the Boundary between Two Immiscible Liquids; Springer: Berlin, 1987.

(78). Starks CM; Liotta CL; Halpern ME Phase-Transfer Catalysis: Fundamentals, Applications, and Industrial Perspectives; Chapman & Hall: New York, 1994.

(79). Patel AJ; Varilly P; Jamadagni SN; Hagan MF; Chandler D; Garde S Sitting at the Edge: How Biomolecules Use Hydrophobicity to Tune Their Interactions and Function. J. Phys. Chem. B 2012, 116, 2498–2503. [PubMed: 22235927]

(80). Gennis RB Biomembranes: Molecular Structure and Function; Springer-Verlag: New York, 1989.

(81). Benjamin I Chemical Reactions and Solvation at Liquid Interfaces: A Microscopic Perspective. Chem. Rev 1996, 96, 1449–1476. [PubMed: 11848798]

(82). Fowkes FM Additivity of Intermolecular Forces at Interfaces. I. Determination of the Contribution to Surface and Interfacial Tensions of Dispersion Forces in Various Liquids. J. Phys. Chem 1963, 67, 2538–2541.

(83). Fowkes FM Attractive Forces at Interfaces. Ind. Eng. Chem 1964, 56, 40–52.

(84). Chapela GA; Saville G; Thompson SM; Rowlinson JS Computer Simulation of a Gas–Liquid Surface. Part 1. J. Chem. Soc., Faraday Trans 2 1977, 1133–1144.

(85). Jochum M; Andrienko D; Kremer K; Peter C Structure-Based Coarse-Graining in Liquid Slabs. J. Chem. Phys 2012, 137, 064102. [PubMed: 22897250]

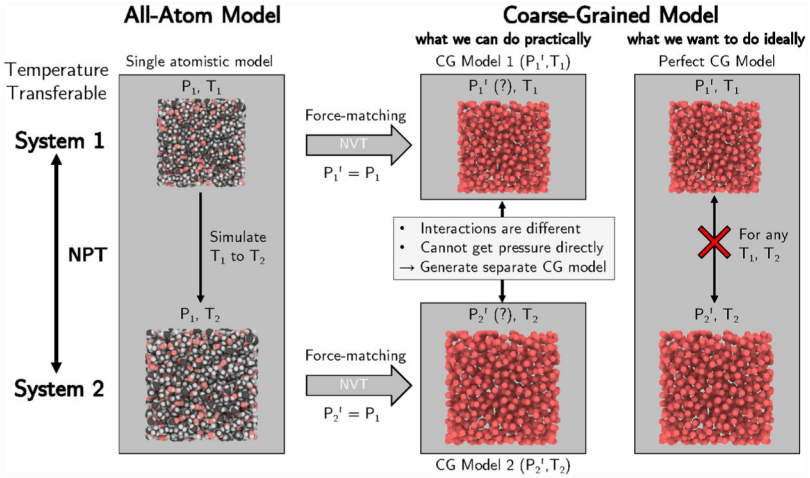


Figure 1.

Schematic diagram illustrates challenges in temperature transferability for bottom-up CG models at *constant pressure*. Ideally, one could transfer a single CG model to different temperatures by maintaining pressure in CG resolution (rightmost column) as we perform simulations at FG resolution (leftmost column). However, in bottom-up CG models, this is limited because of (1) transferability and (2) representability problems (middle column).

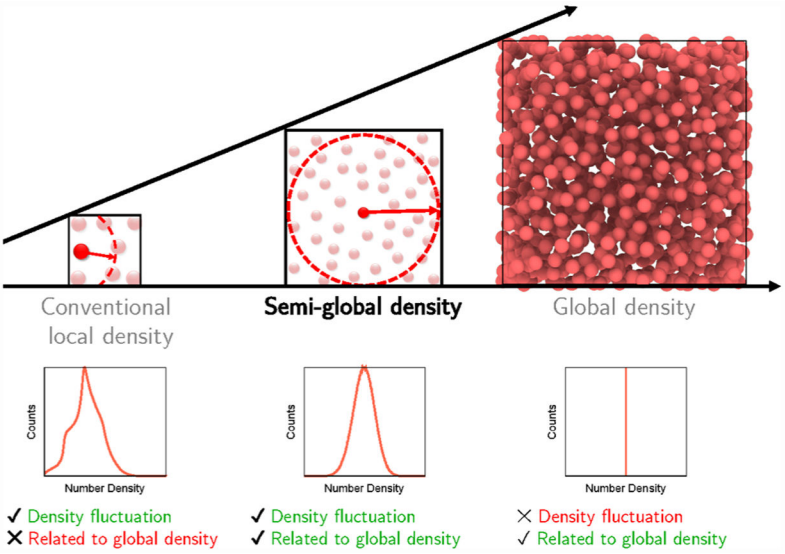


Figure 2.

Illustration of the semi-global density order parameter introduced in this work. The semi-global density (middle) demonstrates the Gaussian distribution with a mean value that is related to the global density. Unlike the semi-global density, a global density of the system (right) is a single value and, thus, does not account for changes to the local environment. For the conventional local density with a smaller cutoff (left), the density fluctuation is still observed throughout the CG simulation, but the density profile becomes too local to accurately reflect the global density, and it is no longer normally distributed.

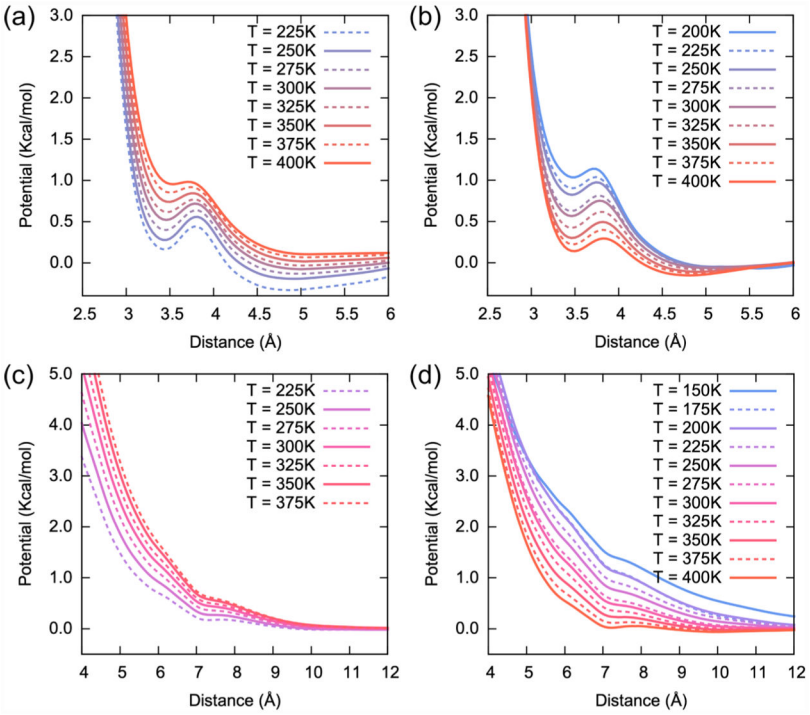


Figure 3.

Effective CG interactions of the methanol and acetone systems for various temperatures under different ensemble conditions. (a) Methanol-methanol CG PMFs under constant *NVT* conditions from 225 to 400 K (blue to red). (b) Methanol-methanol CG PMFs under constant *NPT* conditions from 200 to 400 K (blue to red). (c) Acetone-acetone CG PMFs under constant *NVT* conditions from 225 to 375 K (purple to red). (d) Acetone-acetone CG PMFs under constant *NPT* conditions from 150 to 400 K (blue to red). FG calculation and CG parametrization protocols are detailed in the Supporting Information Sections S1 and S3.

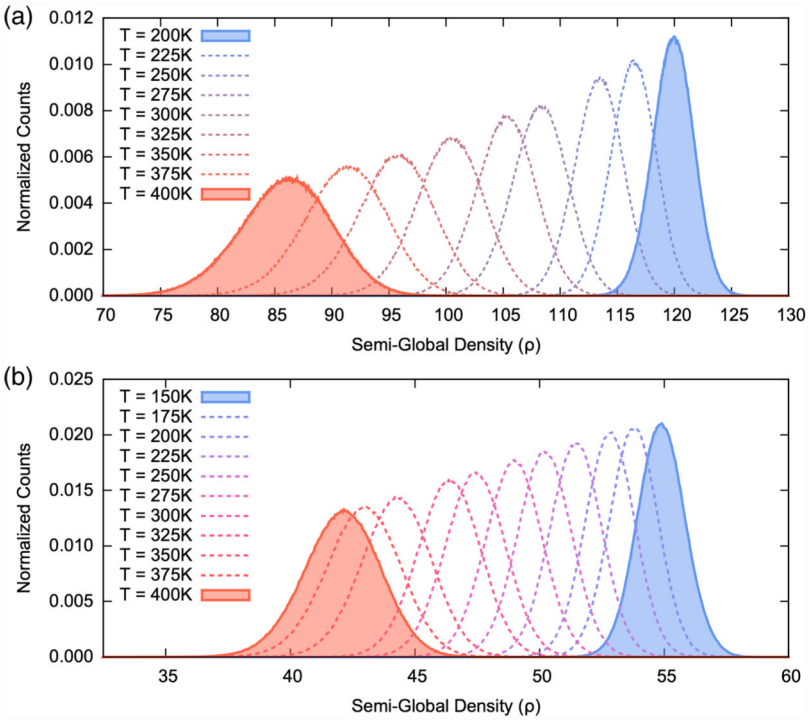


Figure 4.

Density histogram of mapped all-atom (FG) systems at various temperatures. (a) Methanol, from 200 K (blue, filled) to 400 K (red, filled). Intermediate temperatures are shown in dashed lines from right to left for every 25 K. (b) Acetone, from 150 K (blue, filled) to 400 K (red, filled). Intermediate temperatures are shown in dashed lines from right to left for every 25 K.

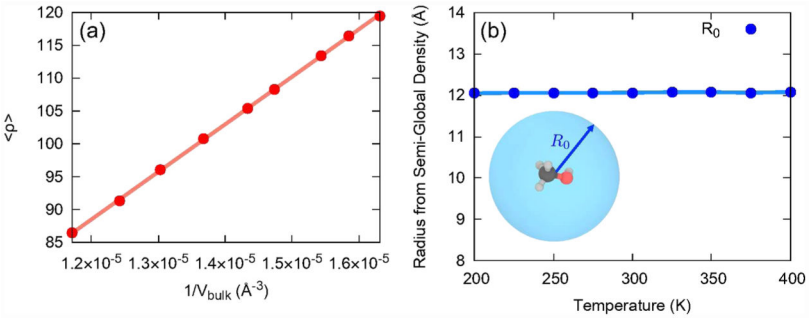


Figure 5.

Analyzing the semi-global nature of the density order parameter used in this work. (a) Relationships between the average semi-global densities $\langle\rho\rangle$ and the reciprocal of overall volumes $1/V_{\text{bulk}}$, confirming the semi-global nature of the utilized density. (b) Density radius R_0^{ρ} calculated from the average densities $\langle\rho\rangle$ over different temperatures using eq 26.

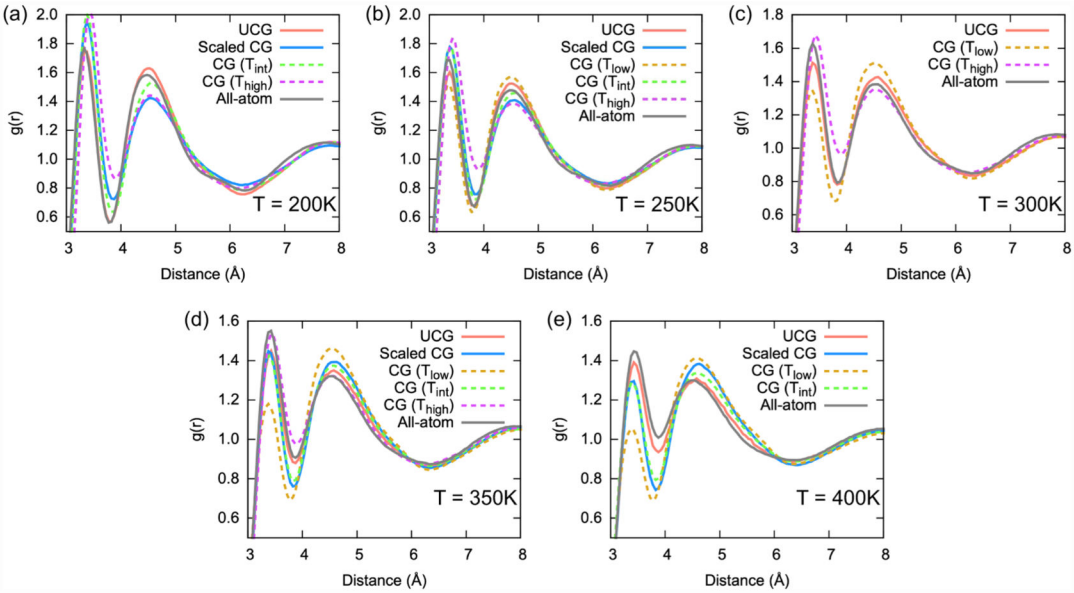


Figure 6.

Methanol-methanol intermolecular pair correlation functions $g(r)$ for atomistic (black line) and various CG models at different temperatures under *constant pressure*: (a) 200, (b) 250, (c) 300, (d) 350, and (e) 400 K. Here, we tested four different CG models: UCG (red line), empirically scaled CG (blue line), naïve CG from T_{low} = 200 K (gold dashed), naïve CG from T_{int} = 300 K (green dashed), and naïve CG from T_{high} = 400 K (purple dashed).

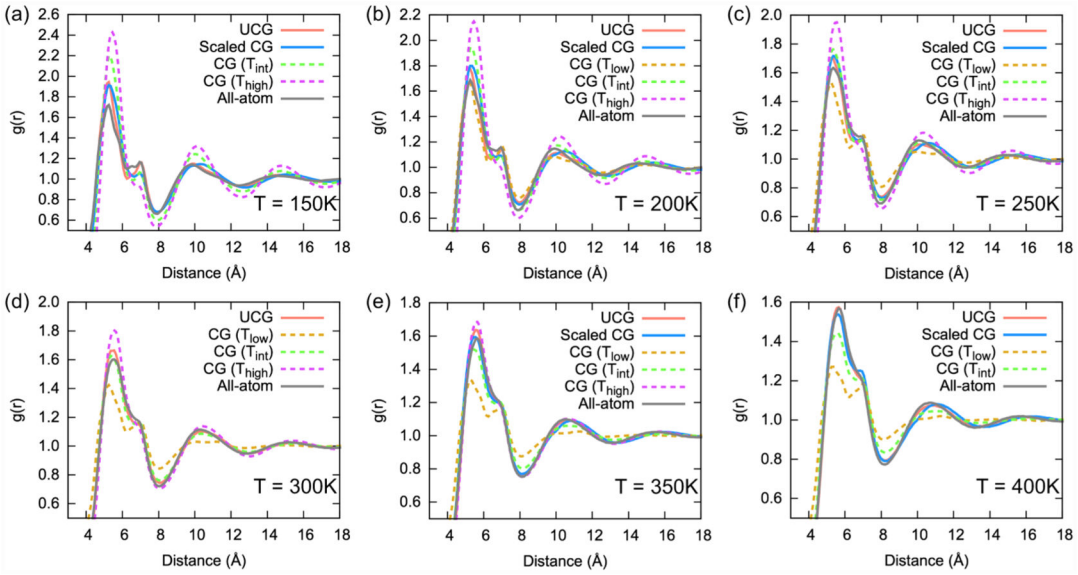


Figure 7.

Acetone-acetone intermolecular pair correlation functions $g(r)$ for atomistic (black line) and various CG models at different temperatures under *constant pressure*: (a) 150, (b) 200, (c) 250, (d) 300, (e) 350, and (f) 400 K. Here, we tested four different CG models: UCG (red line), empirically scaled CG (blue line), naïve CG from T_{low} = 150 K (gold dashed), naïve CG from T_{int} = 275 K (green dashed), and naïve CG from T_{high} = 400 K (purple dashed).

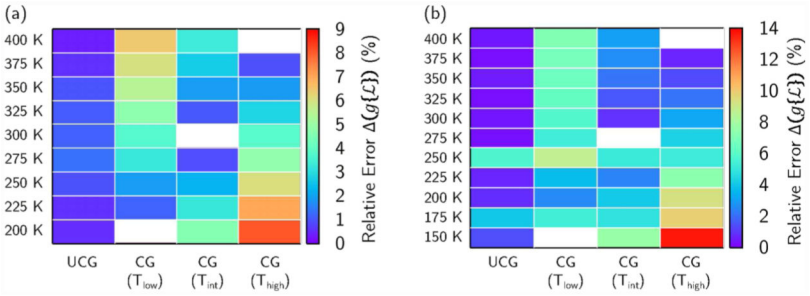


Figure 8.

Heat map of error metric percentage $100 \times \Delta(g\{\mathscr{L}\})$ to reproduce the RDF at given temperatures for various CG models of liquids: (a) methanol and (b) acetone. For each liquid, we calculated the error metric percentage by comparing it to the actual MS-CG model (target). Four CG models are introduced in this figure: UCG (first column), naïve CG at T_{low} (second column), naïve CG at T_{int} (third column), and naïve CG at T_{high} (fourth column). For naïve CG models, we did not calculate errors at the parametrized temperatures. These values are depicted as white blank. Detailed percentage values are listed in Table S2.

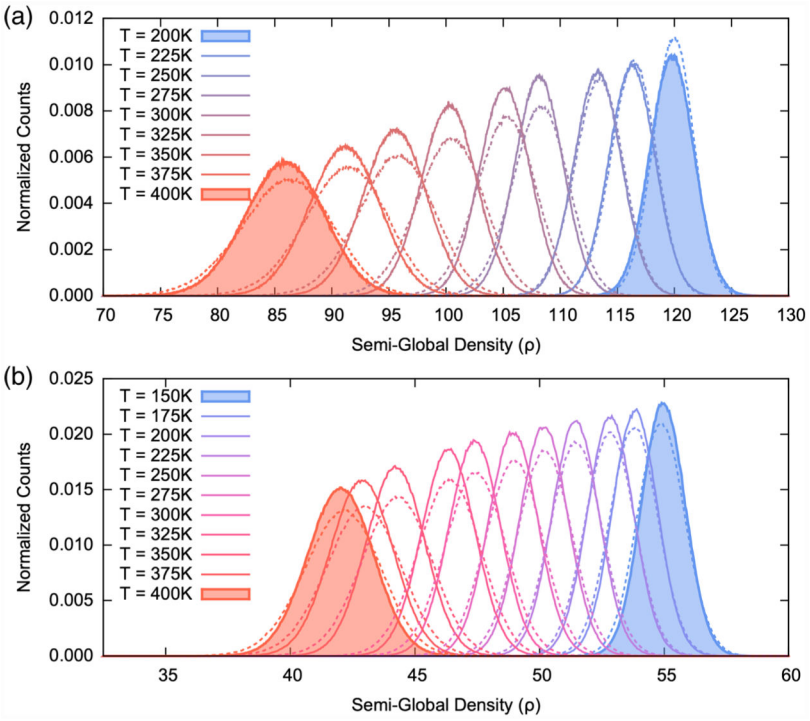


Figure 9.

Assessment of the many-body correlation in the UCG model by calculating density histogram (solid lines) in comparison with the mapped all-atom (dashed lines) from Figure 4. (a) For methanol, density histograms are calculated in the temperature range from 200 K (blue, filled) to 400 K (red, filled). (b) For acetone, density histograms are calculated at various temperatures ranging from 150 K (blue, filled) to 400 K (red, filled). For both insets, intermediate temperatures are located in between two filled curves for every 25 K.

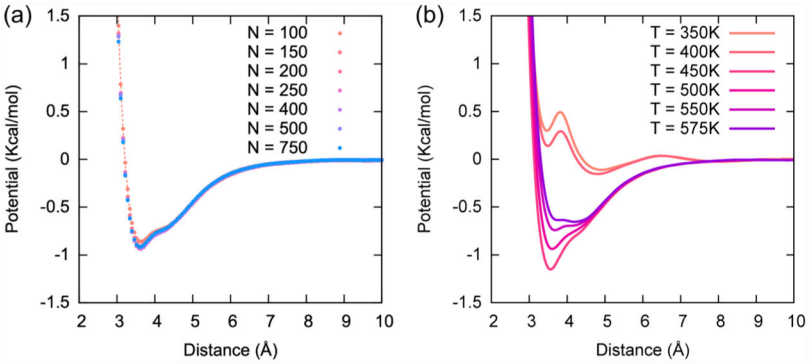


Figure 10.

Methanol-methanol pair CG interaction profiles for different conditions. (a) Gas phase CG interaction at $T = 500$ K with different numbers of methanol molecules from $N = 100$ (orange) to $N = 750$ (blue). Within the hard-core region $R < 3$ Å, the CG interaction from $N = 100$ molecules is shown in dashed orange lines. (b) Pair CG interaction profiles for liquid and gas phases while varying temperature. As temperature increases, liquid state interaction tends to decrease from 350 K (top, salmon) to 400 K (orange), whereas gas state interaction becomes less attractive from 450 K (bottom, red), 500 K (pink), 550 K (magenta), and 575 K (violet).

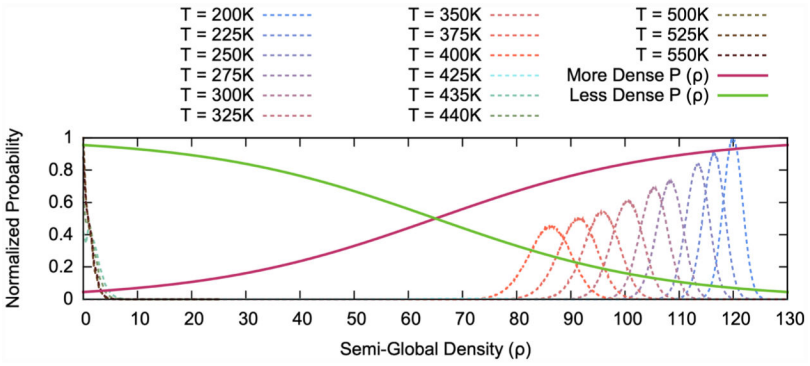


Figure 11.

Normalized semi-global density distributions for different phases (gas: left corner and liquid: right corner) at different temperatures with the corresponding substate probabilities (more dense state: magenta line and less dense state: green line) for the phase transferable UCG model.

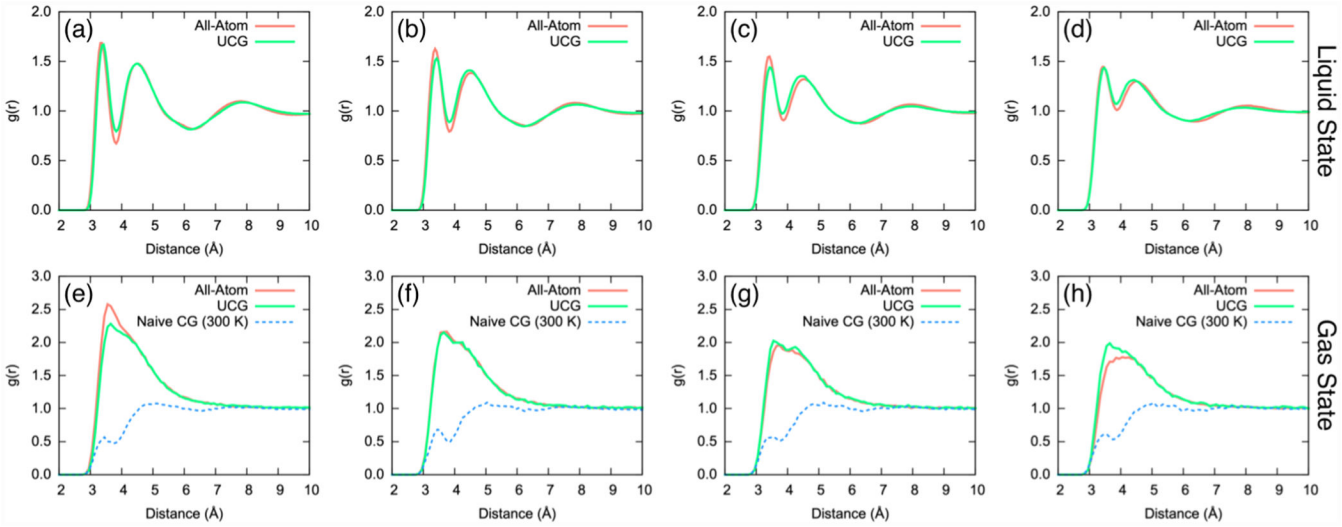


Figure 12.

RDFs $g(r)$ from the mapped all-atom (red line) and UCG simulations (green line) for liquid, panel (a–d), and gas states, panel (e–h), with different temperatures: (a) 250, (b) 300, (c) 350, (d) 400, (e) 500, (f) 525, (g) 550, and (h) 575 K. For the gas state, the naïve CG model where the interaction parameter is directly borrowed from bulk liquid at 300 K was also employed, and the corresponding $g(r)$ is shown in blue dashed lines.

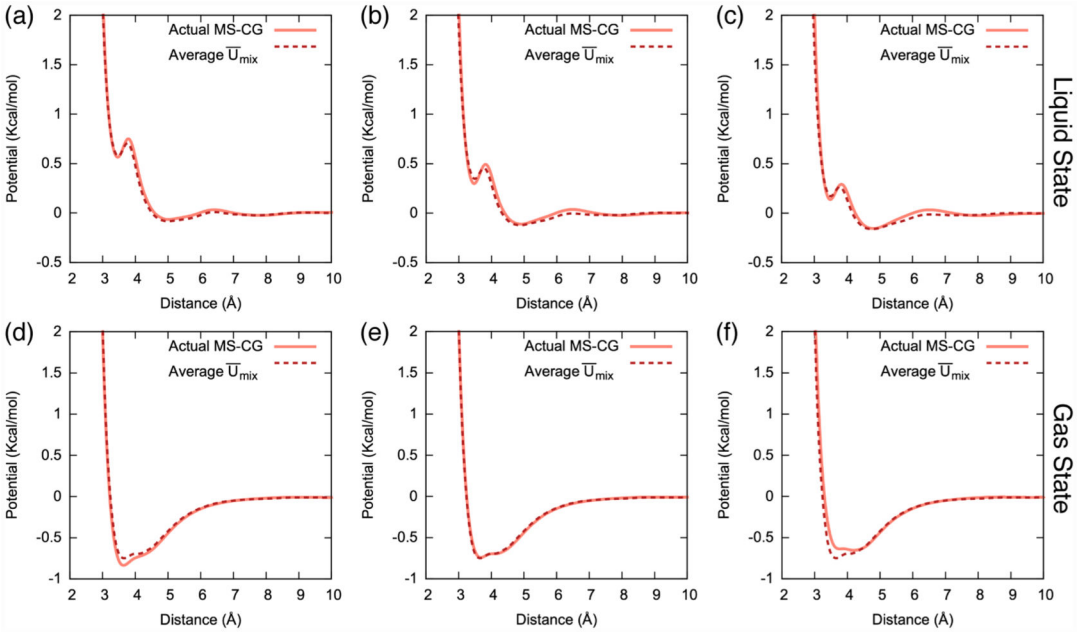


Figure 13.

Confirming a high fidelity of the phase transferable UCG model by comparing the actual MS-CG interactions obtained from each temperature (solid red line) and the mean-field UCG interactions averaged from the CG ensembles at the certain temperature (dashed red line). MS-CG/UCG interactions from the liquid state are shown in panel (a–c) and the interactions from the gas state are shown in panel (d–f), accordingly: (a) T = 300, (b) 350, (c) 400, (d) 525, (e) 550, and (f) 575 K.

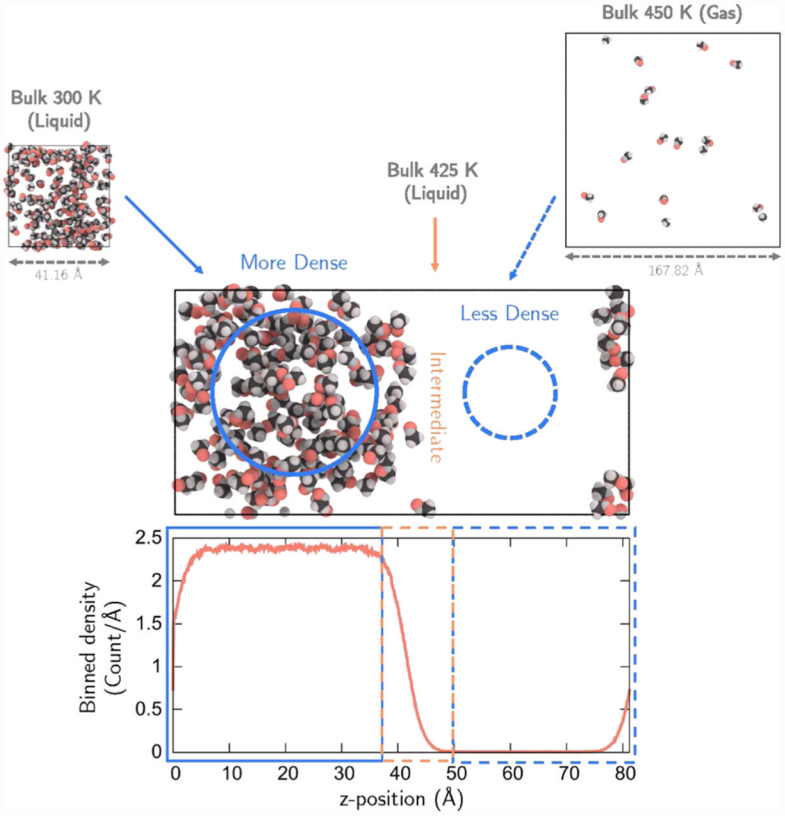


Figure 14.

Schematic diagram illustrating the liquid/vapor interface and construction scheme of the corresponding UCG model (top). Based on the number density along the slab axis (*z*-axis), the binned number density is calculated (bottom). From the number density profile, three distinct regions are identified: inner region (bottom, blue line), phase boundary (bottom, red dashed), and outer region (bottom, blue dashed). The UCG model is then constructed to faithfully represent each region by utilizing the bulk CG interaction from 300, 425, and 450 K, accordingly (top).

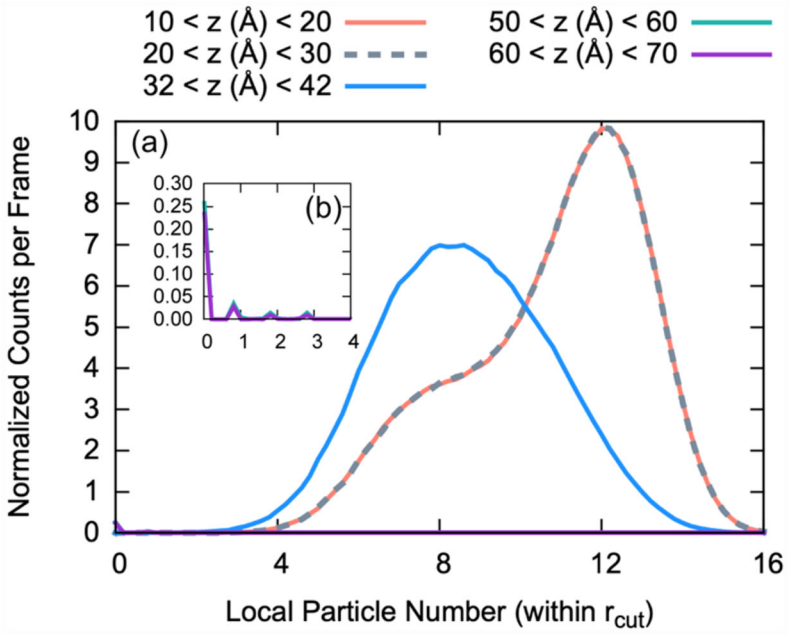


Figure 15.

Histograms of CG methanol particle's density in the liquid/vapor interface at different distances along the slab axis. (a) Each histogram is sampled with 10 Å interval along the *z*-axis: 10–20 Å (red line), 20–30 Å (dashed gray line), 32–42 Å (blue line), 50–60 Å (green line), and 60–70 Å (purple line). (b) Magnified histograms corresponding to the outer region area (50–70 Å). The plotted densities are normalized counts per frame from the reference atomistic trajectory.

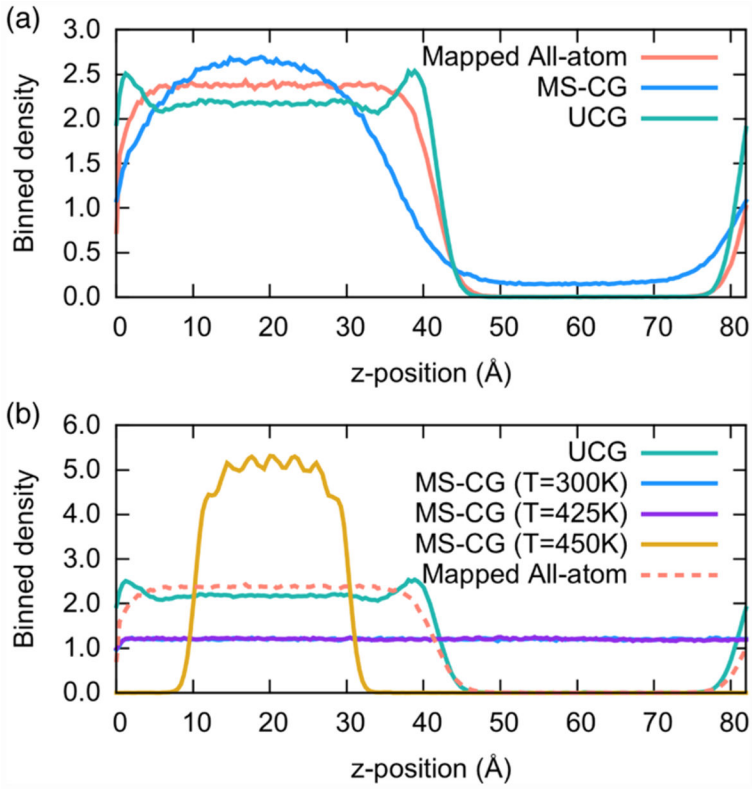


Figure 16.

Slab density profiles $\rho(z)$ of the methanol liquid/vapor interface system using various models: (a) Direct comparison of the all-atom (red line) to the MS-CG (blue line) and UCG (green line). (b) Evaluation of naïve MS-CG models where interaction parameters are directly obtained from bulk systems at 300 K (blue line), 425 K (purple line), and 450 K (gold line). Slab density profiles from the all-atom (red line) and UCG simulations (green line) are also shown here.

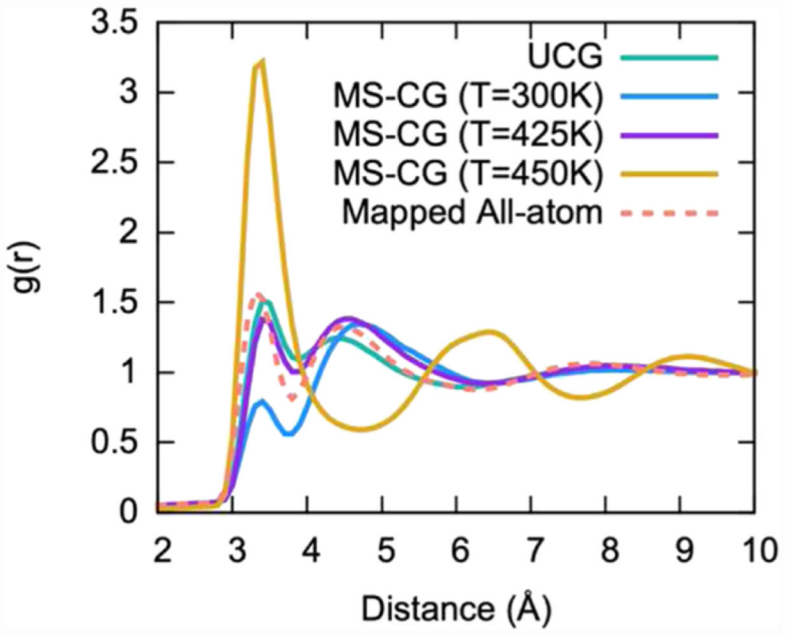


Figure 17.

RDFs $g(r)$ at the inner region of the liquid/vapor interface from the all-atom (red dashed line), naïve MS-CG models at different temperatures (300 K: blue line, 425 K: purple line, and 450 K: gold line), and UCG simulations (green line).

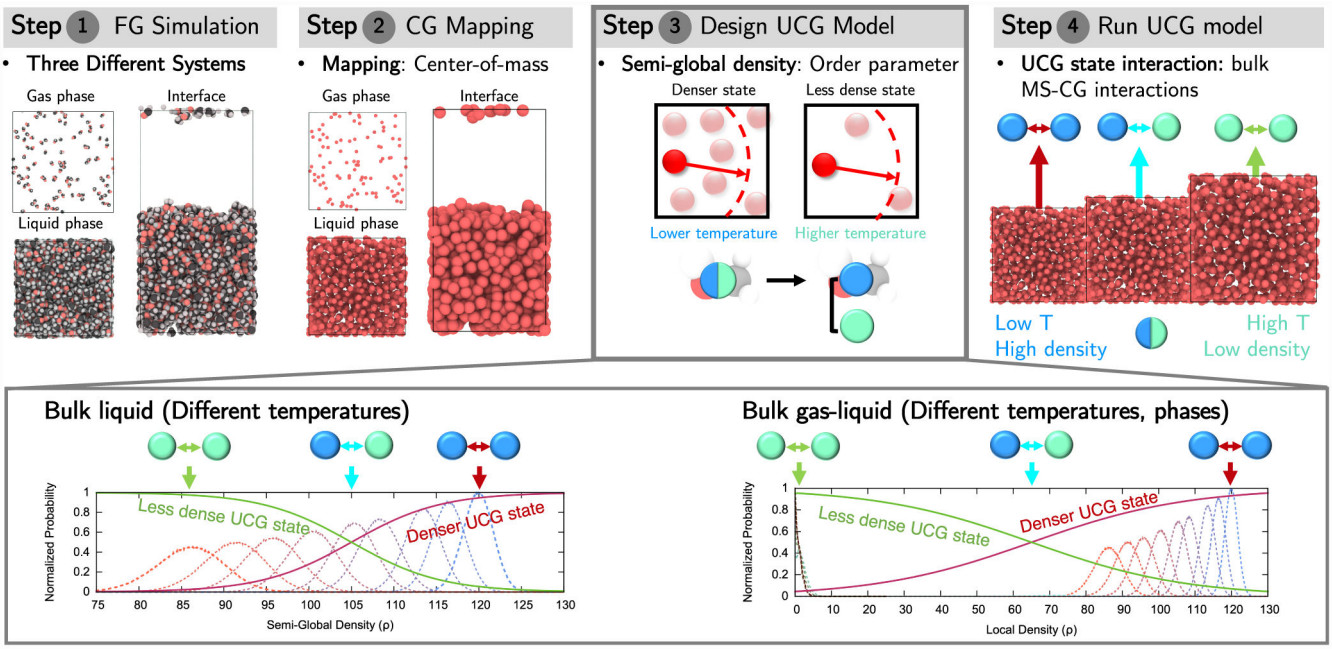


Figure 18.

Schematic procedure of the temperature and phase transferable UCG models in four steps summarizing the current work. Step 1: Running the FG simulation. Step 2: Performing CG mapping (COM mapping is used). Step 3: Designing the UCG model based on the semi-global density. From the system of interest, we calculate the density histogram with R_0 to determine ρ_0 that distinguishes the denser and less dense states. Step 4: Constructing the UCG free energy function from the bulk MS-CG interactions. Here, the bulk systems are chosen based on the density histogram from Step 3. Then, with the UCG Hamiltonian in place, the final UCG trajectories are generated. For clarity, we provide the two density histograms from Sections 3.2 and 3.3 in Step 3.

Table 1.

Average Semi-global Density Values Calculated from the Mapped All-Atom Simulations at Different Temperatures Ranging from 200 to 400 K for (a) Methanol and (b) Acetone[a]

(a) methanol		(b) acetone	
temperature (K)	**average semi-global density $\langle\rho\rangle$**	**temperature (K)**	**average semi-global density $\langle\rho\rangle$**
200	119.50	200	54.90
225	116.40	225	53.85
250	113.45	250	52.90
275	108.30	275	51.40
300	105.45	300	50.10
325	100.85	325	49.00
350	96.10	350	47.45
375	91.35	375	46.45
400	86.50	400	44.20

[a]As seen in Figure 4, all of the density histograms follow almost the Gaussian distribution.

Table 2.

Effective Speed-up Factor of the UCG Simulations with Respect to the Fully Atomistic Simulations at Different Temperatures from 250 to 400 K: (a) Methanol and (b) Acetone

(a) methanol		(b) acetone	
temperature (K)	**speed-up factor**	**temperature (K)**	**speed-up factor**
250	26.44	250	89.10
275	27.77	275	80.49
300	36.80	300	92.68
325	38.40	325	91.86
350	39.36	350	96.72
375	39.73	375	86.60
400	41.25	400	92.90

Table 3.

Fitted Parameters for the Slab Density Profiles from the Mapped All-Atom, MS-CG, UCG, and Naïve MS-CG Models of Various Molecules That Are Delineated in Figure 16[a]

model		ω (Å)	δ (Å)	ρ_l (Å^{-1})	ρ_v (Å^{-1})
all-atom		41.84	5.797	2.381	0.0
MS-CG		35.50	16.61	2.833	0.154
UCG		45.73	2.98	2.207	0.0
	300 K		1.203	1.203	1.203
naïve MS-CG	425 K		1.203	1.203	1.203
	450 K	19.80	5.078	5.078	0.0

[a]Models with no ω values indicate that the interfacial structure is not observed in such cases.

Table 4.

Effective Speed-up Analysis of the Phase Combinational UCG Model by Calculating the Relative Time Per Step

model	relative time per step (step·fs^{-1})	effective speed-up factor
all-atom	67.75	1 (reference)
UCG[a]	0.67	101.11

[a]UCG simulation was performed with a larger time step.