THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

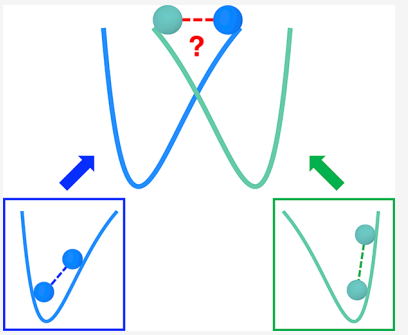
pubs.acs.org/JPCL
Letter

Statistical Mechanical Design Principles for Coarse-Grained Interactions across Different Conformational Free Energy Surfaces

Jaehyeok Jin and Gregory A. Voth*

Cite This: J. Phys. Chem. Lett. 2023, 14, 1354−1362
Read Online

ABSTRACT: Systematic bottom-up coarse-graining (CG) of molecular systems provides a means to explore different coupled length and time scales while treating the molecular-scale physics at a reduced level. However, the configuration dependence of CG interactions often results in CG models with limited applicability for exploring the parametrized configurations. We propose a statistical mechanical theory to design CG interactions across different configurations and conditions. In order to span wide ranges of conformational space, distinct classical CG free energy surfaces for characteristic configurations are identified using molecular collective variables. The coupling interaction between different CG free energy surfaces can then be systematically determined by analogy to quantum mechanical approaches describing coupled states. The present theory can accurately capture the underlying many-body potentials of mean force in the CG variables for various order parameters applied to liquids, interfaces, and in principle proteins, uncovering the complex nature underlying the coupling interaction and imparting a new protocol for the design of predictive multiscale models.

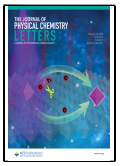


Molecular dynamics (MD) computer simulation has become an important tool for probing the properties of condensed matter since its origination in the 1950s and 1960s.[1,2] Due to a wide range of length and time scales in soft matter systems, multiscale modeling is required to deliver a consistent description across these scales.[3−12] Bottom-up coarse-grained (CG) models provide an attractive yet powerful solution by averaging over unnecessary details from the fine-grained (FG) level.[7,13−24] While maintaining the essential microscopic physics,[10,25−27] this reductionist perspective allows for exploring many different scales from quantum[28] to mesoscopic levels[29,30] and beyond.[31] Despite its efficiency and ability to access large spatiotemporal scales, the reduced and approximate nature of CG models may not be sufficient to correctly capture the underlying free energy surface, i.e., the many-body potential of mean force (PMF) for the CG variables. Since conventional CG models generally approximate the many-body CG PMF as simple interactions based on the molecular mechanics, this problem is exacerbated when the underlying atomistic systems of interest exhibit multiple distinct conformations, e.g., as in protein conformational change or condensed phase system heterogeneity. Using a relatively simple set of CG variables, distinct characteristic interactions from different conformations are often degraded, and the CG models are incapable of distinguishing conformational states at the reference level.

In order to surmount this challenge, recent contributions in the field aim to increase the expressiveness of CG models by introducing multiple conformations or configurations.[32−35] The so-called multi-configurational CG models are designed to account for changes in the CG force field from different conformational free energy surfaces using additional collective variables (CVs)[36] for the same set of CG variables. The very essence of multi-configurational CG modeling can be thought of as an extension of anisotropic network models that have been widely employed in physics and biology, where the desired model can switch between multiple equilibrium states.[37−41] However, for systematic CG models, defining these "conformational states" is a nontrivial task. A well-defined state and the corresponding free energy surfaces should be defined along some CVs that can correctly distinguish distinctive conformations and free energies of the reference system.[42−44]

More importantly, introducing multiple free energy surfaces naturally requires a *coupling* between distinct CG free energy surfaces, so addressing this coupling is crucial for describing the underlying energy surface. At the molecular level, contributions by Warshel and others have established the empirical valence bond (EVB) approach by introducing product and reactant diabatic states to describe chemical reactions.[45,46] In the multi-configurational EVB formalism, the coupling between two diabatic states reveals information about

Received: December 19, 2022
Accepted: January 27, 2023
Published: February 2, 2023



https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

the transition state and barrier.[47−50] However, at the classical CG level, this coupling should be related to how CG particles across different conformational free energy surfaces interact with each other (known as the "cross-interaction"). Understanding the CG cross-interaction is not only important in terms of the accuracy of CG models but also can enable one to design more *transferable* CG models outside of specific parametrized conditions. Since many-body CG PMFs are configuration-dependent,[15−17] bottom-up CG models are dependent on other thermodynamic variables (even as simple as temperature[51]), and their applicability to systems outside of the parametrized conditions may be limited.[10,22,52−56] Therefore, it is desirable to design more transferable CG models if one knows how to correctly couple different CG interactions together. In this Letter, we derive a unique theoretical framework motivated by analogy to quantum mechanical (multi-configurational) state space that can elucidate complex couplings between distinct CG free energy surfaces, thereby imparting a statistical mechanical design principle for CG interactions across diverse conformations.

Generally speaking, interaction potentials ("force fields") governing the CG equation of motion $U_{\text{tot}}(\mathbf{R}^N)$ are commonly designed based on molecular mechanics functional forms of

$$U_{\text{tot}}(\mathbf{R}^N) = U(\mathbf{R}^N) + \sum_{\text{bonds}} U_b(d_{IJ}) + \sum_{\text{angles}} U_\theta(\theta_{IJK}) + \sum_{\text{dihedrals}} U_\psi(\psi_{IJKL}) \tag{1}$$

consisting of nonbonded, bonded, angle, and dihedral potentials, respectively.[57−62] To clearly demonstrate our idea of multi-configurational CG models, we limit our discussions to the nonbonded interaction here, even though an extension to bonded interactions and beyond can be readily done (see refs 32−35). Due to computational cost, low-dimensional basis sets are commonly used to simplify the nonbonded interactions $U(\mathbf{R}^N)$, i.e., the pairwise approximation,

$$U(\mathbf{R}^N) = \sum_{I,J} U(R_{IJ}) \tag{2}$$

Nevertheless, eq 2 may not be able to correctly represent many-body free energies along different CV values (Figure 1). Since changes in CV values are due to the changes in molecular conformations at a given time in CG MD trajectories, a multi-configurational CG model can be constructed by introducing internal states into a CG model under the quasi-equilibrium condition. In other words, CG force fields can be expressed as a weighted average of multi-configurational interactions, where weights correspond to each CG particle's probability of being at each state s_I

$$U_{\text{multi}}(\mathbf{R}^N) = \sum_{I,J} \sum_{s_I,s_J} p_I(s_I)\, p_J(s_J)\, U_{s_I,s_J}(R_{IJ}) \tag{3}$$

where the relevant CV enters through the state probability $p_I(s_I) = p_I(s_I|\mathbf{R}^N)$ to reflect a correct configurational basin conditioned to $\mathbf{R}^N$ that satisfies $\sum_{s_I} p_I(s_I) = 1$ Equation 3, known as the ultra-coarse-grained (UCG) interaction,[63] has been applied to various chemical systems with different physical characteristics by imposing internal states, e.g., self-assembly,[63] interfaces,[64] hydrogen bonding,[65] and phases.[56] To note, despite its complexity, the interaction parameters in eq 3 can be variationally parametrized from the atomistic

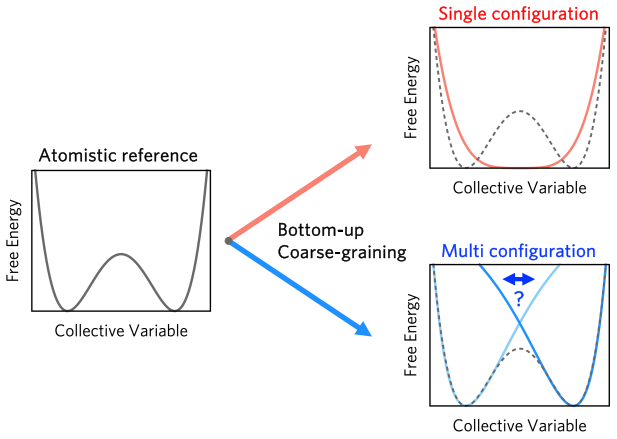


Figure 1. Schematic diagram illustrates a limitation of conventional bottom-up CG approaches. The usual practice (red) based on the single conformation may not be able to capture the underlying free energy surface having different conformations (black); this description can be improved if more than one conformational state is introduced (blue).

information from a bottom-up perspective using multiscale coarse-graining force-matching (MSCGFM).[17,18,66,67]

For two-state models denoted herein as $\{s_I\} = \{1, 2\}$, eq 3 becomes

$$U_{\text{multi}}(\mathbf{R}^N) = \sum_{I,J} [p_{1,I}p_{1,J}U_{11}(R_{IJ}) + (p_{1,I}p_{2,J} + p_{2,I}p_{1,J}) U_{12}(R_{IJ}) + p_{2,I}p_{2,J}U_{22}(R_{IJ})] \tag{4}$$

Here, U_{11} and U_{22} denote the self-interactions at each of the "diabatic" configurational free energy surfaces, but U_{multi} also contains a cross-interaction U_{12}, which is essentially a coupling interaction between the U_{11} and U_{22} energy surfaces. For the sake of simplicity and clarity, we elaborate our argument for two-state models hereafter, but it should be emphasized that our approach is general enough to go beyond just two coupled states. While several bottom-up CG methodologies allow for variational parametrization of these statewise interactions for a given system, a systematic and physical understanding of the coupling interaction in terms of the coupled free energy surfaces has remained largely unknown. Therefore, the current state-of-the-art bottom-up CG methodologies lack a statistical mechanical principle for designing CG interactions coupling different CG free energy surfaces at a predictive level.

The central idea here to understand the cross-interactions is inspired by the quantum mechanical description of a two-state coupling scenario, e.g., as in electron transfer.[68] A continuous generalization of the quantum-like picture[45,46] to the atomistic[69−73] and CG[33,34] regimes further motivates this approach. While the proposed idea here shares a similar description with the existing multi-configurational CG approaches[32−35] by decomposing a complex CG free energy surface onto separate free energy surfaces, the physical definition of each conformational surface and surface dynamics is fundamentally different. Previous approaches such as the surface hopping-like method[32,35] and two-state coupling scheme[33,34] assumed that the time scale of the transitions between the conformational surfaces is of the same order of magnitude as the CG molecular motion. This assumption allows for defining the individual free energy basins according to intramolecular degrees of freedom, i.e., bonds, angles, and dihedrals, but does not impart adequate

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

information about cross-interactions. However, our decomposition of the CG free energy surface is based on determining local CVs that can characterize the heterogeneous, multi-configurational nature. Then, the CG (nonbonded) interactions should immediately adjust when a local chemical environment varies.[63] By attributing the multiple configurations to the intrinsic differences in CV spaces, eq 3 provides a natural approach for understanding the cross-interactions across different configurations in CG models.

Following the multi-configurational approach, a CG Hamiltonian $\mathbf{H}_{CG}(\mathbf{R}^N)$ for the two configurations of CG systems can be written as a matrix form containing the diabatic CG free energy surfaces that was introduced in eq 3,

$$\mathbf{H}_{CG}(\mathbf{R}^N) = \begin{pmatrix} U_{11}(\mathbf{R}^N) & U_{12}(\mathbf{R}^N) \\ U_{12}(\mathbf{R}^N) & U_{22}(\mathbf{R}^N) \end{pmatrix} \tag{5}$$

These two diabatic states 1 and 2 are coupled with some nontrivial coupling free energy $U_{12}(\mathbf{R}^N)$. By analogy to the quantum mechanical treatment, the overall CG energy surface can be obtained by solving the characteristic equation $\mathbf{H}_{CG}c = E_0 c$, where $E_0(\mathbf{R}^N)$ denotes the lowest eigensolution of $E(\mathbf{R}^N)$. Then, if there exists a nontrivial solution for this equation, the lowest-energy eigenvalue can be solved via the following secular determinant

$$\begin{vmatrix} U_{11}(\mathbf{R}^N) - E_0(\mathbf{R}^N) & U_{12}(\mathbf{R}^N) \\ U_{12}(\mathbf{R}^N) & U_{22}(\mathbf{R}^N) - E_0(\mathbf{R}^N) \end{vmatrix} = 0 \tag{6}$$

It is then straightforward that the coupling $U_{12}(\mathbf{R}^N)$ should satisfy[33]

$$U_{12}(\mathbf{R}^N)^2 = [U_{11}(\mathbf{R}^N) - E_0(\mathbf{R}^N)][U_{22}(\mathbf{R}^N) - E_0(\mathbf{R}^N)] \tag{7}$$

Even though eq 7 poses the necessary condition for $E_0(\mathbf{R}^N)$, eq 7 itself does not clearly show how $U_{12}(\mathbf{R}^N)$ can be designed from two diabatic states, i.e., $U_{11}(\mathbf{R}^N)$ and $U_{22}(\mathbf{R}^N)$. Therefore, currently available CG methodologies, e.g., refs 33 and 34, rely only on eq 7 and are limited to approximating $U_{12}(\mathbf{R}^N)$ by utilizing the free energy values from the atomistic reference to eq 7.

Nevertheless, if two diabatic free energy surfaces are well-separated by certain CVs with locality, then we immediately notice that $E_0(\mathbf{R}^N)$ obtained by eq 7 should be consistent with the multi-configurational CG free energy $U_{multi}(\mathbf{R}^N)$ from eq 3. We note that solving eq 7 gives

$$E_0(\mathbf{R}^N) = \frac{1}{2}[U_{11}(\mathbf{R}^N) + U_{22}(\mathbf{R}^N)] - \left[\left(\frac{U_{11}(\mathbf{R}^N) - U_{22}(\mathbf{R}^N)}{2}\right)^2 + U_{12}(\mathbf{R}^N)^2\right]^{1/2} \tag{8}$$

To further explore this possible correspondence, we introduce the pairwise decomposability for the CG free energies in eq 8 in line with a previous CG model for studying protein conformations[40]

$$E_0(\mathbf{R}^N) = \sum_{IJ}\left[\frac{1}{2}[U_{11}(R_{IJ}) + U_{22}(R_{IJ})] - \left[\left(\frac{U_{11}(R_{IJ}) - U_{22}(R_{IJ})}{2}\right)^2 + U_{12}(R_{IJ})^2\right]^{1/2}\right] \tag{9}$$

We immediately notice that despite their different derivations, both eqs 3 and 9 describe a general form of a multi-configurational free energy for the same system in a pairwise form. Within the pairwise approximation, $E_0(\mathbf{R}^N)$ and $U_{multi}(\mathbf{R}^N)$ must be equivalent for any molecular systems exhibiting a heterogeneous nature characterized by specific CVs. In principle, solving $E_0(\mathbf{R}^N) = U_{multi}(\mathbf{R}^N)$ involves a system of equations for every particle pair, and there is more than one solution. As such, it is unlikely that the pairwise coupling interaction $U_{12}(R_{IJ})$ can be analytically expressed in terms of the other variables involving state probabilities $p_I(s_I)$ and $p_J(s_J)$ due to the diversity and complexity of molecular systems. In turn, one can define only the following analytical expression

$$\sum_{s_I,s_J} p_I(s_I)p_J(s_J)U_{s_I,s_J}(R_{IJ}) = \frac{1}{2}[U_{11}(R_{IJ}) + U_{22}(R_{IJ})] - \left[\left(\frac{U_{11}(R_{IJ}) - U_{22}(R_{IJ})}{2}\right)^2 + U_{12}(R_{IJ})^2\right]^{1/2} \tag{10}$$

for a CG particle pair R_{IJ} in a dilute-solution limit, where it is assumed that introducing additional particles to the system does not affect the local state assignment. In the later sections, we will explore if eq 10 is sufficient to describe the pairwise cross-interaction for practical CG systems that are no longer under dilute conditions. In turn, eq 10 can provide a means to enrich our understanding of the interactions between distinct CG free energy surfaces.

Mean-Field Limit. Since a full consideration of eq 3 involves tracking four state probability terms from two pairwise interacting CG particles, we first demonstrate our theory by starting from the simplest *mean-field* scenario. When the choice of CV is not too localized such that all CG particles follow an identical CV distribution exerted by other particles in terms of ensemble averages, we can remove the particle dependency on the ensemble average of state probability distributions, i.e., $\langle p_{I,1}\rangle = \langle p_{J,1}\rangle = p_1$ and $\langle p_{I,2}\rangle = \langle p_{J,2}\rangle = p_2$. A practical realization of this CV will be introduced later. In this case, it is convenient to introduce an ensemble average of the multi-configurational CG Hamiltonian (many-body CG PMFs) under the pairwise approximation:

$$\bar{U}_{UCG}(\mathbf{R}^N) = \sum_{I,J} p_1^2 U_{11}(R_{IJ}) + 2p_1p_2U_{12}(R_{IJ}) + p_2^2U_{22}(R_{IJ}) \tag{11}$$

Then, solving eq 10 is equivalent to

$$p_1^2 U_{11}(R_{IJ}) + 2p_1p_2U_{12}(R_{IJ}) + p_2^2U_{22}(R_{IJ}) = \frac{1}{2}[U_{11}(R_{IJ}) + U_{22}(R_{IJ})] - \left[\left(\frac{U_{11}(R_{IJ}) - U_{22}(R_{IJ})}{2}\right)^2 + U_{12}(R_{IJ})^2\right]^{1/2} \tag{12}$$

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

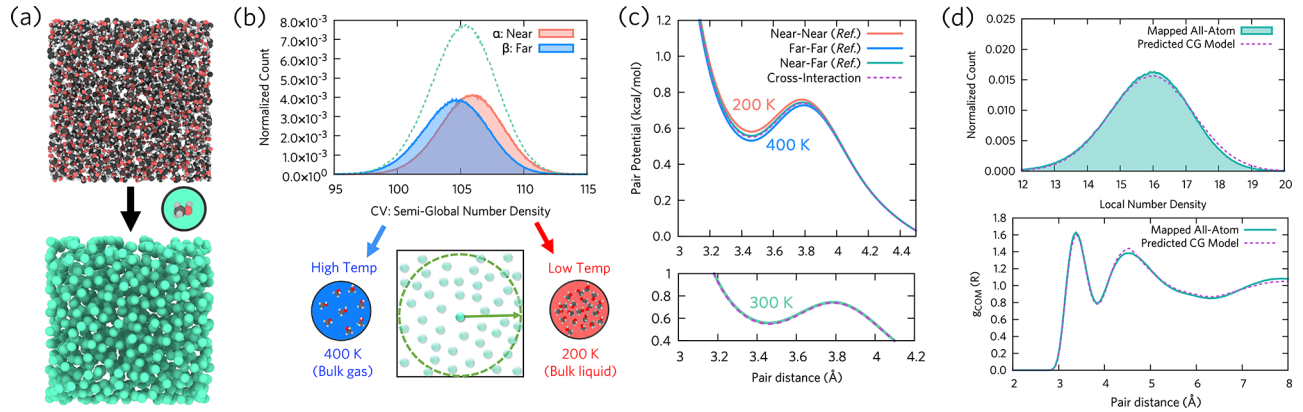


Figure 2. Coupling interaction between CG free energy surfaces of methanol liquids using the semi-global number density CV (*mean-field limit*): (a) Representation of the atomistic (top) and the CG (bottom) models of liquid methanol. (b) Configurational states using the semi-global number density (green dashed line) as α (red) corresponds to a denser state and β (blue) corresponds to a less dense state. (c) Self- and cross-interactions between α and β states. The reference statewise interactions are obtained from the UCG parametrization. Both top and bottom panels compare the estimated coupling interaction from this work using eq 13 (purple dashed line) and the parametrized reference (green). (d) Performance of the CG model with the estimated cross-interaction (purple dashed line) by examining the local density fluctuation within the second coordination shell (top) and the intermolecular pair correlation function $g_{COM}(R)$ (bottom) in comparison with the mapped all-atom results (green).

for each pair distance R_{IJ}. We note that eq 12 can be further simplified due to the fact that $p_1 + p_2 = 1$. After some algebraic manipulation (Supporting Information), we arrive at the coupling interaction

$$U_{12}(R_{IJ}) = \frac{p_1 p_2 U_{11}(1 - 2p_1^2) + p_1 p_2 U_{22}(1 - 2p_2^2) + \sqrt{2p_1 p_2 (U_{11}U_{22} - p_1^2 U_{11}^2 - p_2^2 U_{22}^2)}}{1 - 4p_1^2 p_2^2} \tag{13}$$

where the pairwise distance (R_{IJ}) dependency on the right-hand side is omitted here for clarity.

We first validate the consistency of eq 13 for limiting cases: U_{12} is zero when $p_1 = 0$ or $p_2 = 0$, corresponding to the single configurational basin case. Interestingly, eq 13 suggests that the coupling interaction is not just a simple average between state interactions. Instead, U_{12} contains information about the averaged state probability of each conformational basin. For the simplest case, where $p_1 = p_2 = 1/2$, we arrive at the mixed pair interaction

$$U_{12}(R_{IJ}) = \frac{1}{3}\left(\frac{U_{11}}{2} + \frac{U_{22}}{2} + \sqrt{8U_{11}U_{22} - 2U_{11}^2 - 2U_{22}^2}\right) \tag{14}$$

which contains a nonlinear term between U_{11} and U_{22}, and the cross-interaction at the CG level does not follow simple arithmetic or geometric averages expected from the combining rule in molecular mechanics.[74,75] In turn, eqs 13 and 14 suggest that the CG cross-interactions are inherently complex, but their coupling interaction can be analytically derived in terms of their underlying free energy surfaces. This observation is rather different from the conventional quantum mechanical treatment (where approximations to some extent are required), indicating that it is theoretically possible to design a transferable CG interaction across various configurational spaces.

Next, we applied the derived mean-field description to some realistic molecular systems. All computer simulations were carried out using the LAMMPS MD engine,[76−78] where simulation details are described in the Supporting Information. We note that readers may refer to refs 56, 64, and 65 regarding the construction of UCG models for the molecules reported in this work. In practice, we consider a bulk methanol liquid at ambient temperature (300 K) coarse-grained to its center of mass [Figure 2(a)]. Despite being a homogeneous liquid, a recent CG study showed that homogeneous bulk CG models could be viewed as a mixture of a low-temperature liquid-like state at 200 K and a high-temperature gas-like state at 400 K.[56] This argument was further corroborated by examining the number density of CG particles using a fairly large cutoff distance, referred to as the *semi-global number density*. At this limit, the semi-global number density distribution follows a normal distribution function (substantiated by the Gaussian fluctuation model[79−81] and information theory approach[82]), and its mean value is directly related to the volume of the system. This analysis allows us to describe the CG methanol particle as a mixture of low-temperature and high-temperature states [Figure 2(b)]. Even though we note that the relative populations in low- and high-temperature states are quite similar for the 300 K conditions, the parametrized statewise interaction from force-matching distinguishes a difference in relative interaction strengths at the two temperature limits with their coupling interaction [Figure 2(c)].

From the high-temperature and low-temperature CG free energy surfaces, their cross-interaction can be described by applying eq 13, since this semi-global CV also satisfies $\langle p_I \rangle = \langle p_J \rangle = \langle p \rangle$ and $\langle p_I p_J \rangle = \langle p \rangle^2$. Figure 2(c) compares the predicted cross-interaction with the actual cross-interactions from the UCG parametrization of the atomistic reference.[63,67] Given the observation that the two potentials are identical to each other, we further substantiate the accuracy of the predicted cross-interaction by performing the CG MD simulation using the predicted interaction shown in Figure 2(c). The nearly identical agreement between the predicted

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

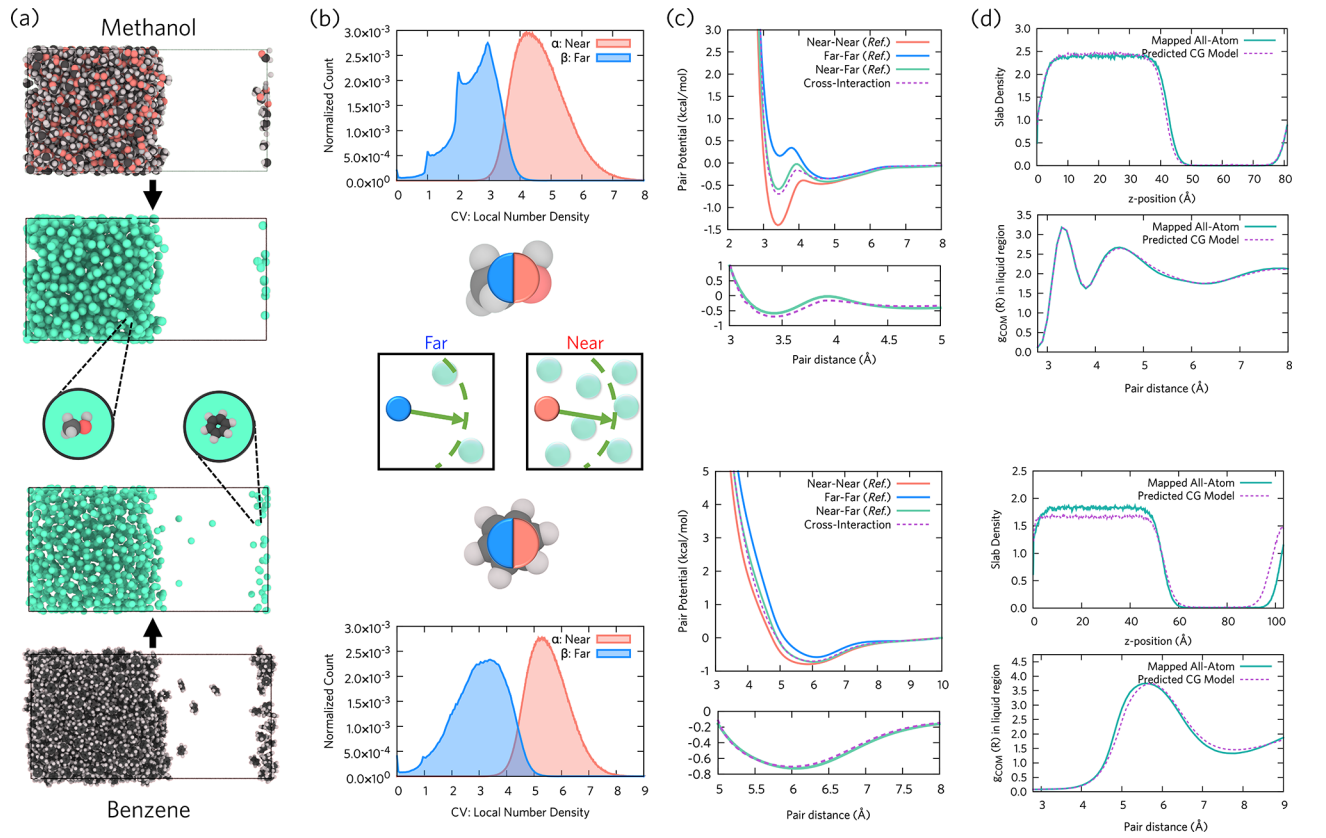


Figure 3. Coupling interaction between CG free energy surfaces of liquid−vapor interfaces of methanol (top panels) and benzene (bottom panels) using the local number density CV. (a) Representation of the atomistic and CG models of liquid−vapor interfaces. (b) Definition of configurational states using the local number density, where α (red) corresponds to a denser state and β (blue) corresponds to a less dense state. (c) Self- and cross-interactions between α and β states. The reference statewise interactions are obtained from the UCG parametrization. Both panels compare the estimated coupling interaction from this work using eq 16 (purple dashed line) and the parametrized reference (green line). (d) Performance of the transferred CG model with the estimated cross-interaction (purple dashed line) by examining the slab density profile along the slab z axis and the intermolecular pair correlation function $g_{\mathrm{COM}}(R)$ in the liquid-like inner region compared with the mapped all-atom MD results (green line).

CG model and the atomistic reference for the two- and N-body static correlations [Figure 2(d)] indicates that eq 13 can accurately derive the interaction between CG particles located on different many-body free energy surfaces.

Beyond the Mean-Field Limit. While the mean-field approximation adopted in eq 11 relates the large fluctuations in density to bulk properties of a system, the semi-global density might not be an optimal CV to correctly delineate changes in the local environment. Therefore, we now establish a link between diabatic energy surfaces and coupling interactions beyond the mean-field limit. From eq 10, a generalization of eq 13 requires including per-particle state probabilities in the CG Hamiltonian and can be written as

$$\frac{1}{2}[U_{11}(R_{IJ}) + U_{22}(R_{IJ})] - \left[\left(\frac{U_{11}(R_{IJ}) - U_{22}(R_{IJ})}{2}\right)^2 + U_{12}(R_{IJ})^2\right]^{1/2} = [p_{1,I}p_{1,J}U_{11}(R_{IJ}) + (p_{1,I}p_{2,J} + p_{2,I}p_{1,J})U_{12}(R_{IJ}) + p_{2,I}p_{2,J}U_{22}(R_{IJ})] \tag{15}$$

where the additional constraints are now imposed by the UCG design principles in which $p_{1,I} + p_{2,I} = 1$ and $p_{1,J} + p_{2,J} = 1$ from eq 5. In spite of its complicated form, eq 15 is still quadratic, and thus using similar algebra, the coupling interaction can be derived as

$$U_{12}(R_{IJ}) = \frac{[\Omega_1] + \sqrt{[\Omega_2]}}{2(1 - p_{1,I}^2p_{2,J}^2 - 2p_{1,I}p_{1,J}p_{2,I}p_{2,J} - p_{1,J}^2p_{2,I}^2)} \tag{16}$$

where

$$[\Omega_1] = 2p_{1,I}p_{2,I}p_{1,J}^2U_{11} + 2p_{1,I}p_{2,I}p_{2,J}^2U_{22} + 2p_{1,I}^2p_{1,J}p_{2,J}U_{11} + 2p_{2,I}^2p_{1,J}p_{2,J}U_{22} - p_{1,I}p_{2,J}U_{11} - p_{2,I}p_{1,J}U_{11} - p_{1,I}p_{2,J}U_{22} - p_{2,I}p_{1,J}U_{22} \tag{17a}$$

$$[\Omega_2] = 4p_{1,I}^2p_{1,J}^2U_{11}^2 + p_{1,I}^2p_{2,J}^2U_{11}^2 - 2p_{1,I}^2p_{2,J}^2U_{11}U_{22} + p_{1,I}^2p_{2,J}^2U_{22}^2 + 2p_{1,I}p_{2,I}p_{1,J}p_{2,J}U_{11}^2 + p_{1,I}p_{2,I}p_{1,J}p_{2,J}U_{11}U_{22} + 2p_{1,I}p_{2,I}p_{1,J}p_{2,J}U_{22}^2 - 4p_{1,I}p_{1,J}U_{11}^2 - 4p_{1,I}p_{1,J}U_{11}U_{22} + p_{1,J}^2p_{2,I}^2U_{11}^2 - 2p_{1,J}^2p_{2,I}^2U_{11}U_{22} + p_{1,J}^2p_{2,I}^2U_{22}^2 + 4p_{2,I}^2p_{2,J}^2U_{22}^2 - 4p_{2,I}p_{2,J}U_{11}U_{22} - 4p_{2,I}p_{2,J}U_{22}^2 + 4U_{11}U_{22} \tag{17b}$$

As expected, the Ω_1 and Ω_2 terms in eqs 16, 17a, and 17b contain an ensemble average of up to four-point probability correlations involving $p_{1,I}$, $p_{1,J}$, $p_{2,I}$, and $p_{2,J}$. These state

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

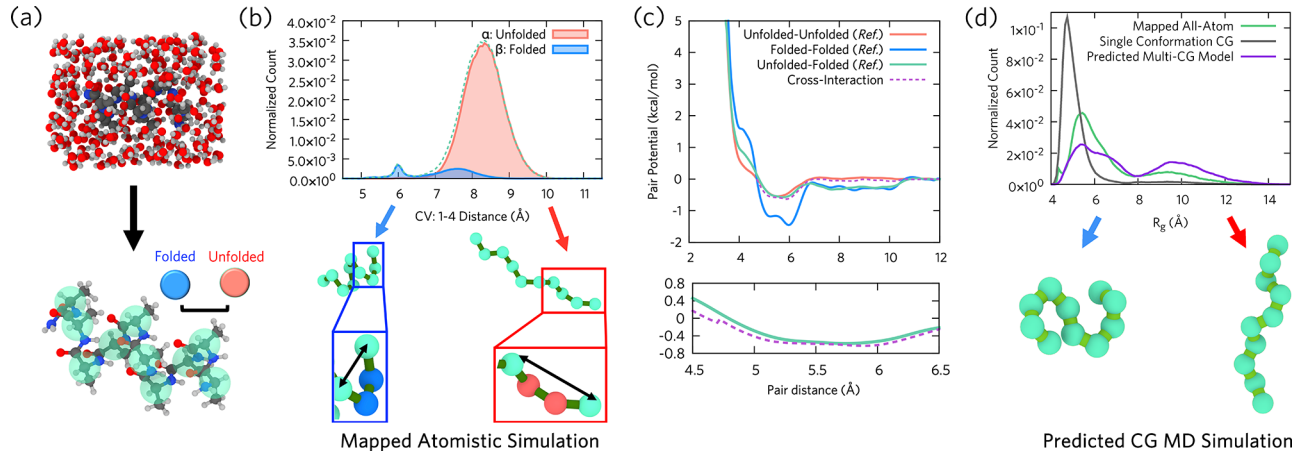


Figure 4. Coupling interaction between CG free energy surfaces of deca-alanine solvated in water using the 1−4 distance CV. (a) Representation of the atomistic (top) and the minimalist CG (bottom) models of deca-alanine. (b) Definition of configurational states using the 1−4 distance, where two α (red) sites per dihedral correspond to an unfolded state and two β (blue) sites correspond to a folded state. (c) Self- and cross-interactions between α and β states. The reference statewise interactions are obtained from the UCG parametrization. Both panels compare the estimated coupling interaction from this work using eq 16 (purple dashed line) and the parametrized reference (green line). (d) Performance of the transferred multi-conformational CG model (purple line) and comparison to the atomistic reference (green line) as well as single conformation CG model (gray line) by computing the radius of gyration, R_g (eq 18). Snapshots from the predicted CG simulation readily capture the folded and unfolded conformations (bottom).

probability products can be efficiently evaluated during a computer simulation (Supporting Information).

In order to assess the multi-configurational cross-interaction in eq 16, we considered the heterogeneous liquid−vapor interface system for molecular liquids. Despite their apparent simplicity,[83,84] liquid−vapor interfaces inherently pose representability and transferability problems because a single CG interface model cannot easily distinguish liquid-like behavior from gas-like behavior.[64] We therefore first approached this problem by constructing the single-site CG models for liquid−vapor interfaces of methanol and benzene [Figure 3(a)]. At the single-site center-of-mass resolution, the local density of each CG particle evaluated up to the first coordination shell shows a clear bimodal distribution corresponding to liquid-like (near) and gas-like (far) states [Figure 3(b)]. Having two internal states per CG particle, the parametrized statewise interaction from the atomistic trajectories confirms the molecular intuition underlying the internal states: the liquid-like states are sticky to each other, whereas the gas-like states are more repulsive than liquid. This trend is consistent with our recent paper showing that these self-interactions are transferable to bulk liquid and gas phases [Figure 3(c)].[64] However, it has been completely unknown how to understand the cross-interactions between liquid-like and gas-like internal states and establish their transferability, which "lives" between the two self-interactions. Here, our theory predicts the cross-interaction remarkably accurately for both methanol and benzene by utilizing only information about two self-interactions, e.g., the minimum at the first coordination shell, as depicted in Figure 3(c).

To demonstrate the fidelity and transferability of our approach beyond the mean-field limit, we constructed *transferable* CG interface models with the self-interactions directly transferred from bulk liquid and gas systems (see ref 56 for the detailed setup) and the estimated cross-interaction shown in Figure 3(c). As depicted in Figure 3(d), the transferable CG models can recapitulate the pair structures within the liquid region of interfaces as well as correctly describe liquid−vapor phase coexistence. Notably, the finely reproduced slab and interface widths with respect to the atomistic reference confirm the accuracy of our theory. This finding indicates that modeling complex many-body CG interactions could be achieved by transferring the CG interactions from relatively simpler bulk systems and coupling them.

To further demonstrate the applicability and fidelity of the present approach to more challenging complex systems and CVs, we consider the minimalist polypeptide system: deca-alanine solvated in water.[85−87] As depicted in Figure 4(a), the CG model was constructed by mapping 1 CG site per amino acid residue and with implicit solvent. Unlike the example of molecular liquids, conventional bottom-up CG models of polypeptides and proteins can have clear limitations in terms of reproducing secondary structure, conformational states, and self-assembly.[88] In particular, due to the complex CG free energy surface originating from the helix−coil transition, the density-based CVs introduced earlier are often not sufficient to correctly capture important structural features of the polypeptide, e.g., its secondary structural content.[89] Hence, we chose the 1−4 distance as a CV to characterize the local heterogeneity and to assign internal states for nonbonded interactions. The detailed CG model design is described in the Supporting Information. Based on the 1−4 distance distribution, we assigned folded or unfolded states to Cα and Cβ sites per dihedral with their state probabilities [Figure 4(b)].

The parametrized CG interactions between the folded−folded and unfolded−unfolded states exhibit a notable difference at 4−6 Å; the more stable folded interaction shows a more attractive profile than the unfolded state interaction. Interestingly, in this system, the parametrized cross-interaction is closer to the folded interaction instead of falling at the midpoint of the two self-interactions, as seen in previous cases. This trend can be explained by several metastable basins, and the helix-like global minimum is rarely visited throughout the MD simulation, resulting in the cross-interaction being closer to the unfolded−unfolded interaction.

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

Remarkably, our theory can still recapitulate this trend correctly, indicating that the present theoretical framework does not treat the cross-interaction as a simple arithmetic average but can correctly encode the molecular nature underlying complex molecules as described by CVs. Finally, we compared the performance of our approach by running the CG simulation and calculating the radius of gyration at the CG level

$$R_g^2(\mathbf{R}^N) = \frac{1}{N}\sum_I (\mathbf{R}_I - \mathbf{R}_{\mathrm{COM}})^2 \tag{18}$$

where $\mathbf{R}_{\mathrm{COM}}$ is the center-of-mass of the deca-alanine molecule. As expected, the conventional CG model exhibits a single conformation near $R_g \approx 5$ Å, whereas the transferable CG model can recapitulate the two distinct conformations for the biosystem as well, consistent with the atomistic reference [Figure 4(d)]. The small difference between the transferable CG model and the reference can be understood from the multistate nature of intramolecular interactions in polyalanine motifs, as reported in ref 32 for tetra-alanine, that are missing in our model. Hence, we expect that extending the proposed multi-configurational theory to bonded interactions shown in eq 1 can provide a more accurate description for higher-dimensional systems. By combining these approaches to leverage the correct configuration-dependent nature of CG interactions, the successful demonstration of our theory here for this minimal polypeptide suggests promising directions for future work to capture large conformational changes in more complex proteins such as lipid bilayers[90] and the endophilin H0 helix.[91]

In this work, we have presented a new theoretical multi-configurational UCG (MC-UCG) approach that can characterize the complex nature of cross-interactions between the CG free energy surfaces in different free energy basins. The molecular cross-interaction was derived by equating the internal state UCG Hamiltonian (akin to Ehrenfest dynamics at the quantum level) to the multi-configurational CG Hamiltonian (akin to multi-configurational EVB at a quantum level). Albeit with various approximations, our finding demonstrates that a quantum mechanical-like description of multi-configurations can be extended to a more "mesoscopic" molecular level. It may be that the physical principle underlying such quantum-meso isomorphism might be understood from the relationship between quantum delocalization and mesoscopic metastable states. A rigorous theory thus might be based on a connection between the quantum transition state theory[92] and CG dynamics[93,94] and will be pursued as future work.

Our theory further reveals how to systematically design CG interactions that can transfer across different conformational basins. The development of design principles for bottom-up and increasingly transferable CG models is of utmost importance in multiscale modeling, yet this effort is still in its early stages due to the inability of simple and approximate CG models to correctly describing the underlying free energy surfaces having multiple basins. Nevertheless, akin to the combining rules for nonbonded interactions[74,75] which are commonplace in the design of atomistic interaction potentials and some top-down CG models,[95] the findings in this work are general enough to encompass various possible CVs and order parmaters. We therefore expect our framework to further facilitate the development of increasingly accurate and general bottom-up CG models.[24]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.2c03844.

Detailed simulation settings for atomistic and coarse-grained simulations; parametrization protocols for coarse-grained models; and derivation of the coupling interaction and practical implementation described in this work (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Gregory A. Voth − *Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and James Franck Institute, The University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0002-3267-6748; Email: gavoth@uchicago.edu

Author

Jaehyeok Jin − *Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and James Franck Institute, The University of Chicago, Chicago, Illinois 60637, United States;* Present Address: Department of Chemistry, Columbia University, New York, New York 10027, United States; orcid.org/0000-0003-3859-3545

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.2c03844

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation (NSF grant CHE-2102677). Simulations were performed using computing resources provided by the University of Chicago Research Computing Center (RCC). J.J. thanks Boksu Jin for insightful motivation and acknowledges extensive discussion and feedback during the Les Houches-TSRC Protein Dynamics Workshop in 2022. J.J. acknowledges the Harper Dissertation Fellowship from the University of Chicago and an Arnold O. Beckman Postdoctoral Fellowship from the Arnold and Mabel Beckman Foundation.

■ REFERENCES

(1) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. I. General method. *J. Chem. Phys.* **1959**, *31* (2), 459−466.
(2) Rahman, A. Correlations in the motion of atoms in liquid argon. *Phys. Rev.* **1964**, *136* (2A), A405.
(3) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586−3616.
(4) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Elsevier, 2001.
(5) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9* (9), 646−652.
(6) Adcock, S. A.; McCammon, J. A. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006**, *106* (5), 1589−1615.

(7) Voth, G. A. *Coarse-Graining of Condensed Phase and Biomolecular Systems*; CRC Press, 2008.
(8) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. Atomic-level characterization of the structural dynamics of proteins. *Science* **2010**, *330* (6002), 341−346.
(9) Schlick, T.; Collepardo-Guevara, R.; Halvorsen, L. A.; Jung, S.; Xiao, X. Biomolecular modeling and simulation: A field coming of age. *Q. Rev. Biophys.* **2011**, *44* (2), 191−228.
(10) Saunders, M. G.; Voth, G. A. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.
(11) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 2017.
(12) Pak, A. J.; Voth, G. A. Advances in coarse-grained modeling of macromolecular complexes. *Curr. Opin. Struct. Biol.* **2018**, *52*, 119−126.
(13) Müller-Plathe, F. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *ChemPhysChem* **2002**, *3* (9), 754−769.
(14) Scheraga, H. A.; Khalili, M.; Liwo, A. Protein-folding dynamics: Overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **2007**, *58*, 57−83.
(15) Izvekov, S.; Voth, G. A. A multiscale coarse-graining method for biomolecular systems. *J. Phys. Chem. B* **2005**, *109* (7), 2469−2473.
(16) Izvekov, S.; Voth, G. A. Multiscale coarse graining of liquid-state systems. *J. Chem. Phys.* **2005**, *123* (13), 134105.
(17) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* **2008**, *128* (24), 244114.
(18) Noid, W. G.; Liu, P.; Wang, Y.; Chu, J.-W.; Ayton, G. S.; Izvekov, S.; Andersen, H. C.; Voth, G. A. The multiscale coarse-graining method. Ii. Numerical implementation for coarse-grained molecular models. *J. Chem. Phys.* **2008**, *128* (24), 244115.
(19) Peter, C.; Kremer, K. Multiscale simulation of soft matter systems−from the atomistic to the coarse-grained level and back. *Soft Matter* **2009**, *5* (22), 4357−4366.
(20) Murtola, T.; Bunker, A.; Vattulainen, I.; Deserno, M.; Karttunen, M. Multiscale modeling of emergent materials: Biological and soft matter. *Phys. Chem. Chem. Phys.* **2009**, *11* (12), 1869−1892.
(21) Riniker, S.; Allison, J. R.; van Gunsteren, W. F. On developing coarse-grained models for biomolecular simulation: A review. *Phys. Chem. Chem. Phys.* **2012**, *14* (36), 12423−12430.
(22) Noid, W. G. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **2013**, *139* (9), 090901.
(23) Shell, M. S. Coarse-graining with the relative entropy. *Adv. Chem. Phys.* **2016**, *161*, 395−441.
(24) Jin, J.; Pak, A. J.; Durumeric, A. E.; Loose, T. D.; Voth, G. A. Bottom-up coarse-graining: Principles and perspectives. *J. Chem. Theory Comput.* **2022**, *18* (10), 5759−5791.
(25) Chaimovich, A.; Shell, M. S. Relative entropy as a universal metric for multiscale errors. *Phys. Rev. E* **2010**, *81* (6), 060104.
(26) Chaimovich, A.; Shell, M. S. Coarse-graining errors and numerical optimization using a relative entropy framework. *J. Chem. Phys.* **2011**, *134* (9), 094112.
(27) Noid, W. G. Systematic methods for structurally consistent coarse-grained models. *Methods Mol. Biol.* **2013**, *924*, 487−531.
(28) Han, Y.; Jin, J.; Wagner, J. W.; Voth, G. A. Quantum theory of multiscale coarse-graining. *J. Chem. Phys.* **2018**, *148* (10), 102335.
(29) Han, Y.; Dama, J. F.; Voth, G. A. Mesoscopic coarse-grained representations of fluids rigorously derived from atomistic models. *J. Chem. Phys.* **2018**, *149* (4), 044104.
(30) Han, Y.; Jin, J.; Voth, G. A. Constructing many-body dissipative particle dynamics models of fluids from bottom-up coarse-graining. *J. Chem. Phys.* **2021**, *154* (8), 084122.
(31) Sherck, N.; Shen, K.; Nguyen, M.; Yoo, B.; Köhler, S.; Speros, J. C.; Delaney, K. T.; Shell, M. S.; Fredrickson, G. H. Molecularly informed field theories from bottom-up coarse-graining. *ACS Macro Lett.* **2021**, *10* (5), 576−583.
(32) Bereau, T.; Rudzinski, J. F. Accurate structure-based coarse graining leads to consistent barrier-crossing dynamics. *Phys. Rev. Lett.* **2018**, *121* (25), 256002.
(33) Sharp, M. E.; Vázquez, F. X.; Wagner, J. W.; Dannenhoffer-Lafage, T.; Voth, G. A. Multiconfigurational coarse-grained molecular dynamics. *J. Chem. Theory Comput.* **2019**, *15* (5), 3306−3315.
(34) Dannenhoffer-Lafage, T.; Voth, G. A. Reactive coarse-grained molecular dynamics. *J. Chem. Theory Comput.* **2020**, *16* (4), 2541−2549.
(35) Rudzinski, J. F.; Bereau, T. Coarse-grained conformational surface hopping: Methodology and transferability. *J. Chem. Phys.* **2020**, *153* (21), 214110.
(36) Wagner, J. W.; Dannenhoffer-Lafage, T.; Jin, J.; Voth, G. A. Extending the range and physical accuracy of coarse-grained models: Order parameter dependent interactions. *J. Chem. Phys.* **2017**, *147* (4), 044113.
(37) Bahar, I.; Atilgan, A. R.; Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding Des.* **1997**, *2* (3), 173−181.
(38) Doruker, P.; Atilgan, A. R.; Bahar, I. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins: Struct., Funct., Bioinf.* **2000**, *40* (3), 512−524.
(39) Atilgan, A. R.; Durell, S.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80* (1), 505−515.
(40) Chu, J.-W.; Voth, G. A. Coarse-grained free energy functions for studying protein conformational changes: A double-well network model. *Biophys. J.* **2007**, *93* (11), 3860−3871.
(41) Zheng, W.; Brooks, B. R.; Hummer, G. Protein conformational transitions explored by mixed elastic network models. *Proteins: Struct., Funct., Bioinf.* **2007**, *69* (1), 43−57.
(42) Hayward, S.; Go, N. Collective variable description of native protein dynamics. *Annu. Rev. Phys. Chem.* **1995**, *46* (1), 223−250.
(43) Rohrdanz, M. A.; Zheng, W.; Clementi, C. Discovering mountain passes via torchlight: Methods for the definition of reaction coordinates and pathways in complex macromolecular reactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 295−316.
(44) Noé, F.; Clementi, C. Collective variables for the study of long-time kinetics from molecular trajectories: Theory and methods. *Curr. Opin. Struct. Biol.* **2017**, *43*, 141−147.
(45) Warshel, A.; Weiss, R. M. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J. Am. Chem. Soc.* **1980**, *102* (20), 6218−6226.
(46) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley and Sons, 1991.
(47) Chang, Y. T.; Miller, W. H. An empirical valence bond model for constructing global potential energy surfaces for chemical reactions of polyatomic molecular systems. *J. Phys. Chem.* **1990**, *94* (15), 5884−5888.
(48) Jensen, F. Transition structure modeling by intersecting potential energy surfaces. *J. Comput. Chem.* **1994**, *15* (11), 1199−1216.
(49) Kim, Y.; Corchado, J. C.; Villa, J.; Xing, J.; Truhlar, D. G. Multiconfiguration molecular mechanics algorithm for potential energy surfaces of chemical reactions. *J. Chem. Phys.* **2000**, *112* (6), 2718−2735.
(50) Schlegel, H. B.; Sonnenberg, J. L. Empirical valence-bond models for reactive potential energy surfaces using distributed gaussians. *J. Chem. Theory Comput.* **2006**, *2* (4), 905−911.
(51) Krishna, V.; Noid, W. G.; Voth, G. A. The multiscale coarse-graining method. Iv. Transferring coarse-grained potentials between temperatures. *J. Chem. Phys.* **2009**, *131* (2), 024103.
(52) Wagner, J. W.; Dama, J. F.; Durumeric, A. E. P.; Voth, G. A. On the representability problem and the physical meaning of coarse-grained models. *J. Chem. Phys.* **2016**, *145* (4), 044108.
(53) Dunn, N. J. H.; Foley, T. T.; Noid, W. G. Van der waals perspective on coarse-graining: Progress toward solving represent-

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362

ability and transferability problems. *Acc. Chem. Res.* **2016**, *49* (12), 2832−2840.
(54) Potter, T. D.; Tasche, J.; Wilson, M. R. Assessing the transferability of common top-down and bottom-up coarse-grained molecular models for molecular mixtures. *Phys. Chem. Chem. Phys.* **2019**, *21* (4), 1912−1927.
(55) Jin, J.; Pak, A. J.; Voth, G. A. Understanding missing entropy in coarse-grained systems: Addressing issues of representability and transferability. *J. Phys. Chem. Lett.* **2019**, *10* (16), 4549−4557.
(56) Jin, J.; Yu, A.; Voth, G. A. Temperature and phase transferable bottom-up coarse-grained models. *J. Chem. Theory Comput.* **2020**, *16* (11), 6823−6842.
(57) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The mm3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111* (23), 8551−8566.
(58) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. Dreiding: A generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94* (26), 8897−8909.
(59) Rappé, A. K.; Casewit, C. J.; Colwell, K.; Goddard, W. A., III; Skiff, W. M. Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024−10035.
(60) Scott, W. R.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The gromos biomolecular simulation program package. *J. Phys. Chem. A* **1999**, *103* (19), 3596−3607.
(61) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668−1688.
(62) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. Charmm: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(63) Dama, J. F.; Jin, J.; Voth, G. A. The theory of ultra-coarse-graining. 3. Coarse-grained sites with rapid local equilibrium of internal states. *J. Chem. Theory Comput.* **2017**, *13* (3), 1010−1022.
(64) Jin, J.; Voth, G. A. Ultra-coarse-grained models allow for an accurate and transferable treatment of interfacial systems. *J. Chem. Theory Comput.* **2018**, *14* (4), 2180−2197.
(65) Jin, J.; Han, Y.; Voth, G. A. Ultra-coarse-grained liquid state models with implicit hydrogen bonding. *J. Chem. Theory Comput.* **2018**, *14* (12), 6159−6174.
(66) Lu, L.; Izvekov, S.; Das, A.; Andersen, H. C.; Voth, G. A. Efficient, regularized, and scalable algorithms for multiscale coarse-graining. *J. Chem. Theory Comput.* **2010**, *6* (3), 954−965.
(67) Dama, J. F.; Jin, J.; Voth, G. A. 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* (4), 2288−2288.
(68) Marcus, R. A. Chemical and electrochemical electron-transfer theory. *Annu. Rev. Phys. Chem.* **1964**, *15* (1), 155−196.
(69) Schmitt, U. W.; Voth, G. A. Multistate empirical valence bond model for proton transport in water. *J. Phys. Chem. B* **1998**, *102* (29), 5547−5551.
(70) Day, T. J.; Soudackov, A. V.; Čuma, M.; Schmitt, U. W.; Voth, G. A. A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.* **2002**, *117* (12), 5839−5849.
(71) Wu, Y.; Chen, H.; Wang, F.; Paesani, F.; Voth, G. A. An improved multistate empirical valence bond model for aqueous proton solvation and transport. *J. Phys. Chem. B* **2008**, *112* (2), 467−482.
(72) Park, K.; Lin, W.; Paesani, F. A refined ms-evb model for proton transport in aqueous environments. *J. Phys. Chem. B* **2012**, *116* (1), 343−352.
(73) Biswas, R.; Tse, Y.-L. S.; Tokmakoff, A.; Voth, G. A. Role of presolvation and anharmonicity in aqueous phase hydrated proton solvation and transport. *J. Phys. Chem. B* **2016**, *120* (8), 1793−1804.
(74) Lorentz, H. Ueber die anwendung des satzes vom virial in der kinetischen theorie der gase. *Ann. Phys.* **1881**, *248*, 127−136.
(75) Berthelot, D. Sur le mélange des gaz. *C. R. Hebd. Séances Acad. Sci.* **1898**, *126*, 1703.
(76) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1−19.
(77) Brown, W. M.; Wang, P.; Plimpton, S. J.; Tharrington, A. N. Implementing molecular dynamics on hybrid high performance computers−short range forces. *Comput. Phys. Commun.* **2011**, *182* (4), 898−911.
(78) Brown, W. M.; Kohlmeyer, A.; Plimpton, S. J.; Tharrington, A. N. Implementing molecular dynamics on hybrid high performance computers−particle−particle particle-mesh. *Comput. Phys. Commun.* **2012**, *183* (3), 449−459.
(79) Chandler, D. Gaussian field model of fluids with an application to polymeric fluids. *Phys. Rev. E* **1993**, *48* (4), 2898.
(80) Crooks, G. E.; Chandler, D. Gaussian statistics of the hard-sphere fluid. *Phys. Rev. E* **1997**, *56* (4), 4217.
(81) Truskett, T. M.; Torquato, S.; Debenedetti, P. G. Density fluctuations in many-body systems. *Phys. Rev. E* **1998**, *58* (6), 7369.
(82) Hummer, G.; Garde, S.; Garcia, A.; Paulaitis, M. E.; Pratt, L. R. Hydrophobic effects on a molecular scale. *J. Phys. Chem. B* **1998**, *102* (51), 10469−10482.
(83) Benjamin, I. Molecular structure and dynamics at liquid-liquid interfaces. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 407−451.
(84) Benjamin, I. Reaction dynamics at liquid interfaces. *Annu. Rev. Phys. Chem.* **2015**, *66*, 165−188.
(85) Hazel, A.; Chipot, C.; Gumbart, J. C. Thermodynamics of deca-alanine folding in water. *J. Chem. Theory Comput.* **2014**, *10* (7), 2836−2844.
(86) Rudzinski, J. F.; Noid, W. G. Bottom-up coarse-graining of peptide ensembles and helix−coil transitions. *J. Chem. Theory Comput.* **2015**, *11* (3), 1278−1291.
(87) Tomar, D. S.; Weber, V.; Pettitt, B. M.; Asthagiri, D. Importance of hydrophilic hydration and intramolecular interactions in the thermodynamics of helix−coil transition and helix−helix assembly in a deca-alanine peptide. *J. Phys. Chem. B* **2016**, *120* (1), 69−76.
(88) Thorpe, I. F.; Zhou, J.; Voth, G. A. Peptide folding using multiscale coarse-grained models. *J. Phys. Chem. B* **2008**, *112* (41), 13079−13090.
(89) Marrink, S. J.; Tieleman, D. P. Perspective on the martini model. *Chem. Soc. Rev.* **2013**, *42* (16), 6801−6822.
(90) Katira, S.; Mandadapu, K. K.; Vaikuntanathan, S.; Smit, B.; Chandler, D. The order-disorder transition in model lipid bilayers is a first-order hexatic to liquid phase transition. *arXiv (Condensed Matter)* **2015**, 1506.04310v3, DOI: 10.48550/arXiv.1506.04310.
(91) Cui, H.; Mim, C.; Vázquez, F. X.; Lyman, E.; Unger, V. M.; Voth, G. A. Understanding the role of amphipathic helices in n-bar domain driven membrane remodeling. *Biophys. J.* **2013**, *104* (2), 404−411.
(92) Voth, G. A.; Chandler, D.; Miller, W. H. Rigorous formulation of quantum transition state theory and its dynamical corrections. *J. Chem. Phys.* **1989**, *91* (12), 7749−7760.
(93) Jin, J.; Schweizer, K. S.; Voth, G. A. Understanding dynamics in coarse-grained models: I. Universal excess entropy scaling relationship. *J. Chem. Phys.* **2023**, *158* (3), 034103.
(94) Jin, J.; Schweizer, K. S.; Voth, G. A. Understanding dynamics in coarse-grained models: Ii. Coarse-grained diffusion modeled using hard sphere theory. *J. Chem. Phys.* **2023**, *158* (3), 034104.
(95) Halgren, T. A. The representation of van der waals (vdw) interactions in molecular mechanics force fields: Potential form, combination rules, and vdw parameters. *J. Am. Chem. Soc.* **1992**, *114* (20), 7827−7843.

https://doi.org/10.1021/acs.jpclett.2c03844
J. Phys. Chem. Lett. 2023, 14, 1354−1362