

Insight into the density-dependence of pair potentials for predictive coarse-grained models

Maria C. Lesniewski and W.G. Noid*

*Department of Chemistry, The Pennsylvania State University, University Park,
Pennsylvania 16802*

E-mail: wnoid@chem.psu.edu

Abstract

We investigate the temperature- and density-dependence of effective pair potentials for 1-site coarse-grained (CG) models of two industrial solvents, 1,4-dioxane and tetrahydrofuran. We observe that the calculated pair potentials are much more sensitive to density than to temperature. The generalized-Yvon-Born-Green framework reveals that this striking density-dependence reflects corresponding variations in the many-body correlations that determine the environment-mediated indirect contribution to the pair mean force. Moreover, we demonstrate, perhaps surprisingly, that this density-dependence is not important for accurately modeling intermolecular structure. Accordingly, we adopt a density-independent interaction potential and transfer the density-dependence of the calculated pair potentials into a configuration-independent volume potential. Furthermore, we develop a single global potential that accurately models the intermolecular structure and pressure-volume equation of state across a very wide range of liquid state points. Consequently, this work provides fundamental insight into the density-dependence of effective pair potentials and also provides a significant step towards developing predictive CG models for efficiently modeling industrial solvents.

1 Introduction

Many computational studies adopt coarse-grained (CG) models to simulate length- and time-scales that can not be effectively addressed with traditional all-atom (AA) models.^{1–7} By representing systems in reduced detail, CG models provide both exceptional computational efficiency^{8–11} and also a conceptually simpler perspective for studying complex materials.^{12–14} Unfortunately, the computational and conceptual advantages of averaging over atomic details come at the cost of effective potentials that vary with thermodynamic conditions.^{10,12,15} Thus, effective potentials that accurately model one set of thermodynamic conditions may provide a poor description of other conditions. Clearly, the computational advantages of CG models will be severely undermined if they must be reparameterized for each condition of interest. Moreover, the temperature-, composition-, and density-dependence of effective CG models also complicates the treatment of thermodynamic properties.^{16–21} For example, the energetic and entropic contributions to effective potentials must be properly treated when modeling energetic and entropic properties.^{12,22,23} Accordingly, a large literature has investigated the dependence of effective potentials upon thermodynamic state point.^{24,25}

For instance, many studies have investigated the temperature-range over which bottom-up pair potentials provide reasonable structural fidelity.^{26–29} Conversely, other studies have explicitly calculated pair potentials for a range of temperatures and, in some cases, attempted to model these potentials, e.g., as a linear function of temperature.^{30,30–41} These studies suggest that bottom-up pair potentials often become more repulsive with increasing temperature at constant density. Interestingly, Voth and coworkers have demonstrated that this observed temperature-dependence can be used to model entropic quantities that would seem beyond the scope of CG models.⁴² Moreover, several recent studies have predicted the temperature-dependence of bottom-up pair potentials based upon simulation data at a single state point via the dual^{43–45} or microcanonical⁴⁶ framework. In particular, the dual approach parameterizes an operator for modeling atomic energetics and then uses this operator to infer the entropic contribution and, thus, the temperature-dependence of CG pair

potentials.

Similarly, many studies have investigated the sensitivity of bottom-up pair potentials to variations in composition or density.^{17,33,36,47–67} Since these potentials tend to be overly repulsive, they are often modified to reproduce the average internal pressure of the AA model,⁶⁸ although this modification often results in discrepancies in the compressibility of the CG model.⁶⁹ Several studies suggest that bottom-up pair potentials tend to become increasingly attractive with decreasing density at constant temperature.^{33,41,51} In fact, bottom-up pair potentials appear more sensitive to density than to temperature, since they become increasingly attractive with increasing temperature as the volume increases at constant external pressure.^{33,41,51} This density-dependence is typically treated “passively” and simply ignored when calculating the internal pressure of bottom-up CG models,^{20,69–72} which may partially account for their tendency to poorly model the AA pressure equation of state. Conversely, Guenza and coworkers have recently demonstrated an “active” approach that accounted for this density-dependence when computing the internal pressure.⁷³ This suggests the possibility of developing a dual approach for predicting the density-dependence of pair potentials based upon the internal pressures sampled by the AA model at a single state point.

Recent studies suggest that local density (LD) potentials provide another promising approach for treating the composition- and density-dependence of bottom-up potentials.^{48,74–78} By employing one-body LD potentials of the density around each particle,⁷⁴ CG models can accurately describe the AA pressure equation of state,^{78–80} the interfacial profiles of inhomogeneous systems,⁸¹ and also the liquid-liquid phase separation of immiscible solvents.^{82,83} Unfortunately, in comparison to conventional pair potentials, LD potentials are less computationally efficient and also more challenging to parameterize.^{76,84} Moreover, they appear to demonstrate rather complex temperature-dependence.^{85,86}

Alternatively, Das and Andersen pioneered a simpler framework for treating the density-dependence of CG potentials for homogeneous systems.⁸⁷ Specifically, Das and Andersen introduced a configuration-independent volume potential that can be readily parameterized

to ensure that the CG model accurately reproduces the volume-dependence of the AA free energy and, as a consequence, the AA pressure equation of state.⁸⁷⁻⁸⁹ In particular, CG models with volume potentials can not only reproduce the average density and internal pressure of AA models, but also their compressibility. In contrast to LD potentials, the volume potential is easy to parameterize and introduces essentially no additional computational cost to simulations of the CG model. Few studies have investigated the temperature-dependence of volume potentials, though.^{41,90} The interesting study of Rosenberger and van der Vegt determined a single fixed pair potential and then parameterized a series of volume potentials for constant pressure simulations at 1 bar external pressure across a rather wide range of temperatures.⁹⁰ They observed that this volume potential varied linearly with temperature. Remarkably, the resulting models quite accurately reproduced the pair structure, bulk density, coefficient of thermal expansion, and isothermal compressibility of hexane and perfluorohexane across a wide temperature range at 1 bar external pressure. However, they did not consider the performance of these models at elevated pressures.

In this work, we further investigate the temperature- and density-dependence of pair and volume potentials for two homogeneous molecular liquids. Specifically, we consider 1-site CG models for 1,4-dioxane ($\text{C}_4\text{H}_8\text{O}_2$) and for tetrahydrofuran ($\text{C}_4\text{H}_8\text{O}$), which are two widely used industrial solvents. Although they have very similar molecular structures, dioxane and tetrahydrofuran have rather different physical properties. While dioxane has historically been employed in cosmetics, shampoos, and detergents, current research into dioxane focuses on its potential role as a carcinogen and the difficulty of separating it from aqueous solutions in the environment.^{91,92} Conversely, tetrahydrofuran remains widely used in chromatography, organic synthesis, and polymer chemistry, but demonstrates limited miscibility in water.^{93,94}

We first parameterize a series of CG potentials for modeling dioxane at a rather wide range of densities and temperatures. As in prior studies, we observe that the calculated pair potentials are much more sensitive to density than to temperature. We employ the generalized-Yvon-Born-Green framework⁹⁵⁻⁹⁸ to elucidate the molecular origin of this strik-

ing density-dependence. However, we find that the density-dependence of the pair potentials has relatively little impact upon the structural properties of the CG model. Consequently, this density-dependence can be directly transferred into the volume potential without reducing the model’s structural fidelity. We then develop a global model for the volume potential that allows us to accurately describe both the intermolecular structure and the pressure equation of state across the entire range of liquid state points for dioxane. Having established this procedure for dioxane, we conclude by validating it for tetrahydrofuran.

The remainder of this manuscript is organized as follows. Section 2 briefly reviews the relevant theory for bottom-up coarse-graining. Sections 3 and 4 present the details and results of our computational studies. Finally, Section 5 summarizes our findings and provides concluding comments.

2 Theory

2.1 Exact Coarse Graining

Bottom-up approaches develop CG models based upon information from a high resolution all-atom (AA) model of the given system.⁹⁹ We denote the AA configuration for n atoms by $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_n\}$ and the potential $u(\mathbf{r}; V)$, which may depend upon the volume V . The instantaneous internal pressure of the AA model is $p_{\text{int}}(\mathbf{r}; V) = nk_{\text{B}}T/V + p_{\text{xs}}(\mathbf{r}; V)$ and the excess pressure is $p_{\text{xs}}(\mathbf{r}; V) = -(\partial u/\partial V)_{\hat{\mathbf{r}}} = \frac{1}{3V} \sum_{i=1}^n \mathbf{f}_i \cdot \mathbf{r}_i - (\partial u/\partial V)_{\mathbf{r}}$, where $\mathbf{f}_i(\mathbf{r}; V) = -(\partial u/\partial \mathbf{r}_i)_V$ is the force on atom i , $\hat{\mathbf{r}}_i = V^{-1/3}\mathbf{r}_i$ is the scaled coordinates for atom i , and $\hat{\mathbf{r}} = \{\hat{\mathbf{r}}_i\}$.¹⁰⁰ For simplicity, we have assumed that the system is spatially homogeneous. The canonical configuration distribution for the AA model is $p_{\text{r}}(\mathbf{r}; V, T) = z^{-1}(V, T) \exp[-\beta u(\mathbf{r}; V)]$, where $z(V, T) = \int_{V^n} d\mathbf{r} \exp[-\beta u(\mathbf{r}; V)]$ and $\beta = 1/k_{\text{B}}T$.

We denote the CG configuration for N sites by $\mathbf{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_N\}$. We define a mapping, $\mathbf{M} : \mathbf{r} \rightarrow \mathbf{R} = \mathbf{M}(\mathbf{r})$, that maps each AA configuration, $\mathbf{r}$, to a particular CG configuration,

$\mathbf{R} = \mathbf{M}(\mathbf{r})$. This determines the canonical mapped distribution,

$$p_{\mathbf{R}}(\mathbf{R}; V, T) = \int_{V^n} d\mathbf{r} p_{\mathbf{r}}(\mathbf{r}; V, T) \delta(\mathbf{R} - \mathbf{M}(\mathbf{r})), \quad (1)$$

which gives the probability (density) that the AA model samples a configuration that maps to $\mathbf{R}$. We seek to approximate the many-body potential of mean force (PMF), W , which may be defined

$$\exp[-\beta W(\mathbf{R}; V, T)] \equiv V^{-(n-N)} z_{\mathbf{R}}(\mathbf{R}; V, T), \quad (2)$$

where

$$z_{\mathbf{R}}(\mathbf{R}; V, T) = \int_{V^n} d\mathbf{r} \exp[-\beta u(\mathbf{r}; V)] \delta(\mathbf{R} - \mathbf{M}(\mathbf{r})) \quad (3)$$

is the total Boltzmann weight associated with the CG configuration $\mathbf{R}$ at the given V and T . Thus, the many-body PMF is defined as the excess Helmholtz potential of the AA model as a function of the CG coordinates. Consequently, if the CG model employs the PMF as an interaction potential, then it will perfectly reproduce the mapped distribution of CG configurations that is implied by the AA model and the CG mapping. Moreover, if the PMF is known as a function of configuration and thermodynamic state point, then the CG model will be able to perfectly reproduce the thermodynamic properties of the AA model.

Because the PMF is an excess Helmholtz potential, the variations of the PMF may be expressed^{20,23}

$$dW = - \sum_{I=1}^N \bar{\mathbf{f}}_I \cdot (d\mathbf{R}_I)_V - \bar{p}_{\text{xs}} dV - S_W dT, \quad (4)$$

where $(d\mathbf{R}_I)_V = V^{1/3} d\hat{\mathbf{R}}_I$ denotes variations in configuration at constant volume and $\hat{\mathbf{R}}_I = V^{-1/3} \mathbf{R}_I$ denotes the scaled CG coordinates. Each term in Eq. (4) relates variation in the PMF to a conditioned average of a conjugate variable over the subensemble of AA

configurations that map to the given CG microstate. In particular,

$$\bar{\mathbf{f}}_I \equiv \bar{\mathbf{f}}_I(\mathbf{R}; V, T) \equiv \langle \mathbf{f}_I(\mathbf{r}; V) \rangle_{\mathbf{R}; VT} \quad (5)$$

$$\bar{p}_{\text{xs}} \equiv \bar{p}_{\text{xs}}(\mathbf{R}; V, T) \equiv \langle p_{\text{xs}}(\mathbf{r}; V) \rangle_{\mathbf{R}; VT}, \quad (6)$$

where (assuming that the mapping associates sites with disjoint atomic groups) $\mathbf{f}_I$ denotes the net atomic force^{101–103} on CG site I and the subscripted brackets indicate an average according to the conditioned distribution:

$$p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}; V, T) \equiv p_{\mathbf{r}}(\mathbf{r}; V, T) \delta(\mathbf{R} - \mathbf{M}(\mathbf{r})) / p_{\mathbf{R}}(\mathbf{R}; V, T). \quad (7)$$

The temperature-dependence of the PMF is determined by

$$S_W \equiv S_W(\mathbf{R}; V, T) \equiv -k_B \left\langle \ln \left[\frac{p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}; V, T)}{q_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}; V, T)} \right] \right\rangle_{\mathbf{R}; VT}, \quad (8)$$

where $q_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}; V, T) = V^{-(n-N)} \delta(\mathbf{R} - \mathbf{M}(\mathbf{r}))$ is the uniform conditioned distribution. Thus, $S_W(\mathbf{R}; V, T)$ is the excess configurational entropy associated with the subensemble of AA configurations that map to $\mathbf{R}$.²²

Note that according to Eqs. (4) – (6)

$$\langle p_{\text{xs}}(\mathbf{r}; V) \rangle_{\mathbf{R}; VT} = \frac{1}{3V} \sum_{I=1}^N \bar{\mathbf{f}}_I \cdot \mathbf{R}_I - \left(\frac{\partial W}{\partial V} \right)_{\mathbf{R}, T}. \quad (9)$$

In general one expects that the conditioned mean of the AA excess pressure can not be reproduced by mean forces. Consequently, one generally expects that the PMF will depend explicitly upon volume, i.e., $(\partial W / \partial V)_{\mathbf{R}, T} \neq 0$. Moreover, by equating mixed partial derivatives of the PMF, we see that

$$\left(\frac{\partial \bar{\mathbf{f}}_I}{\partial V} \right)_{\hat{\mathbf{R}}, T} = - \left(\frac{\partial}{\partial V} \left(\frac{\partial W}{\partial \mathbf{R}_I} \right)_{\mathbf{R}_I^*, V, T} \right)_{\hat{\mathbf{R}}, T} = \left(\frac{\partial \bar{p}_{\text{xs}}}{\partial \mathbf{R}_I} \right)_{\mathbf{R}_I^*, V, T}, \quad (10)$$

where $\hat{\mathbf{R}} = \{\hat{\mathbf{R}}_I\}$ and $\mathbf{R}_I^*$ denotes the set of $N-1$ Cartesian coordinates for sites $J \neq I$. Thus, the volume-dependence of the mean forces corresponds to the configuration-dependence of the mean excess pressure. Finally, we note that in order to reproduce the mapped distribution in the isothermal-isobaric ensemble at a constant external pressure, P_{ext} , it is necessary to account for the missing ideal contributions from the $n-N$ particles that have been integrated out of the AA model. Consequently, the proper interaction potential for modeling the isothermal-isobaric ensemble is

$$W_P(\mathbf{R}, V; T) \equiv W(\mathbf{R}; V, T) - (n - N)k_B T \ln(V/\Lambda_0^3) \quad (11)$$

where Λ_0 is an arbitrary constant with dimensions of length that may depend upon temperature. By accounting for the ideal contribution to the internal pressure, the interaction potential, W_P , ensures that the CG model will properly sample the isothermal-isobaric ensemble, i.e.,

$$\exp[-\beta(W_P(\mathbf{R}, V; T) + P_{\text{ext}}V)] = \Lambda_0^{-3(n-N)} z_R(\mathbf{R}; V, T) \exp[-\beta P_{\text{ext}}V] \quad (12)$$

$$\propto p_{RV}(\mathbf{R}, V; P_{\text{ext}}, T), \quad (13)$$

where $p_{RV}(\mathbf{R}, V; P_{\text{ext}}, T)$ is the joint configuration-volume distribution that is implied by the AA model and the CG mapping.

2.2 Approximate Coarse Graining

We consider CG models for simulating 1-component systems of organic molecular liquids in the isothermal-isobaric ensemble at temperature, T , and external pressure, P_{ext} . Following Das and Andersen,⁸⁷ we approximate the many-body PMF, $W_P(\mathbf{R}, V)$, according to

$$U(\mathbf{R}, V) = U_R(\mathbf{R}) + U_V(V). \quad (14)$$

Here U_R is an interaction potential that determines the forces on each site, $\mathbf{F}_I = -(\partial U_R / \partial \mathbf{R}_I)_{\mathbf{R}_I^*, V}$, but does not explicitly depend upon the volume. Since we consider neutral organic molecules, we represent each molecule with a single CG site and describe inter-molecular interactions with short-ranged central pair potentials, U_2 , between each distinct pair, (I, J) , of molecules:

$$U_R(\mathbf{R}) = \sum_{(I,J)} U_2(R_{IJ}), \quad (15)$$

where R_{IJ} is the distance between molecules I and J . Conversely, $U_V(V)$ is a “volume potential” that does not generate forces, but directly contributes a “volume force,” $F_V(V) = -dU_V(V)/dV$, to the excess pressure

$$P_{\text{xs}}(\mathbf{R}, V) = \frac{1}{3V} \sum_{(I,J)} F_2(R_{IJ}) R_{IJ} + F_V(V) \quad (16)$$

where $F_2(R) = -dU_2(R)/dR$. We model the volume potential with the simple power series

$$U_V(V) = N [\psi_1(V/V_{\text{ref}} - 1) + \psi_2(V/V_{\text{ref}} - 1)^2], \quad (17)$$

where ψ_1 and ψ_2 are parameters, while V_{ref} is the average volume of the AA model at the reference state point (T, P_{ext}) .

While there exist many bottom-up methods for parameterizing CG models,^{49,68,104–106} in this work we adopt the MS-CG force-matching method.^{107,108} The MS-CG method determines the pair potential, U_2 , by minimizing

$$\chi^2[U_2; V_{\text{ref}}, T] = \left\langle \frac{1}{3N} \sum_{I=1}^N \left| \mathbf{f}_I(\mathbf{r}) - \mathbf{F}_I(\mathbf{M}(\mathbf{r})) \right|^2 \right\rangle_{V_{\text{ref}}, T}, \quad (18)$$

where the angular brackets denote a canonical average according to $p_{\text{r}}(\mathbf{r}; V_{\text{ref}}, T)$. The MS-CG variational principle^{102,109,110} ensures that the resulting interaction potential provides an optimal approximation for the configuration-dependence of the PMF at the given state point.

Consequently, we expect that the resulting MS-CG pair potential will explicitly depend upon the volume and temperature of the reference AA ensemble, i.e., $U_2(R) = U_2(R; V_{\text{ref}}, T)$. While Eq. (18) defines the MS-CG variational principle for the canonical ensemble, we also employ the force-matching method for simulations at constant pressure. In practice one obtains the same potentials from minimizing χ^2 over the canonical ensemble and over the corresponding isothermal-isobaric ensemble.¹¹¹

Das and Andersen proposed optimizing the volume potential, $U_V(V)$, by minimizing a pressure-matching functional⁸⁷

$$\chi_V^2[U_V|U_R; P_{\text{ext}}, T] \equiv \langle |\delta P_0(\mathbf{r}, V) - F_V(V)|^2 \rangle_{P_{\text{ext}}T} \quad (19)$$

where $\delta P_0(\mathbf{r}, V) = (n - N)k_B T/V + p_{\text{xs}}(\mathbf{r}, V) - P_{\text{xs}}(\mathbf{M}(\mathbf{r}), V)|_{F_V=0}$ is the difference between the instantaneous pressure of the AA and CG models in the absence of the volume force, while the angular brackets denote an average over the isothermal-isobaric distribution, $p_{rV}(\mathbf{r}, V; P_{\text{ext}}, T) \propto p_r(\mathbf{r}; V, T) \exp[-\beta P_{\text{ext}}V]$. Note that the optimized volume potential, U_V , will depend upon the reference AA ensemble that was used in the parameterization, i.e., $\psi = \psi(V_{\text{ref}}, T)$ where V_{ref} and T are the mean volume and temperature of the constant NPT simulation. Moreover, U_V will also depend upon the interaction potential, U_R , and, thus, reflect the state-point dependence of U_R . In the following, we first estimate U_V via a related pressure-matching functional and then systematically refine U_V until it quantitatively reproduces the AA pressure-volume equation of state.^{88,111} This self-consistent pressure-matching method minimizes the relative entropy^{105,112–115} between the AA and CG constant pressure ensembles with respect to the volume potential.⁸⁹

3 Computational Methods

3.1 Atomistic Simulations

We simulated two homogeneous liquid systems with periodic boundary conditions in all three dimensions. The dioxane system consisted of 685 molecules, while the tetrahydrofuran (THF) system consisted of 768 molecules. We performed AA simulations of both liquids with Gromacs version 2019.6,^{116,117} while propagating dynamics with the Gromacs ‘md’ leapfrog integrator. These simulations employed a 1 fs time step and did not rigidly constrain any bonds. We employed the Bussi stochastic thermostat¹¹⁸ with a relaxation time of 0.5 ps to sample energy fluctuations at constant temperature. We employed an isotropic Parrinello-Rahman barostat¹¹⁹ with a relaxation time of 5 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ to sample volume fluctuations at constant external pressure.

We employed the optimized potentials for liquid simulations (OPLS-AA) force field to model interactions in the AA simulations.¹²⁰ We generated the topology files for both dioxane and tetrahydrofuran with the LigParGen browser tool.^{121–123} We modeled long-ranged electrostatic interactions with the particle-mesh Ewald method,¹²⁴ while using a grid spacing of 0.08 nm and a short-ranged cutoff of 1.3 nm. We truncated short-ranged non-bonded interactions at $r_c = 1.3 \text{ nm}$, while switching the corresponding forces to zero between 1.0 and 1.3 nm. We employed dispersion corrections to account for truncating these interactions. We employed a neighbor list with a cut-off of 1.4 nm that was updated every step.

We performed AA simulations of both systems at the 15 state points indicated in Fig. 1. We constructed this simulation matrix by first performing constant NPT simulations at external pressure $P_{\text{ext}} = P^0 \equiv 1 \text{ bar}$ for 5 evenly spaced temperatures that span the liquid phase range at atmospheric pressure. These simulations correspond to the diagonal of the matrix. We determined the equilibrium density, $\bar{\rho}_T^0 \equiv \langle \rho \rangle_{TP^0}$, at each temperature, T , from these constant NPT simulations. We then performed a constant NVT simulation with the fixed density, $\bar{\rho}_T^0$, at each temperature, $T' > T$, to characterize the 10 state points above the

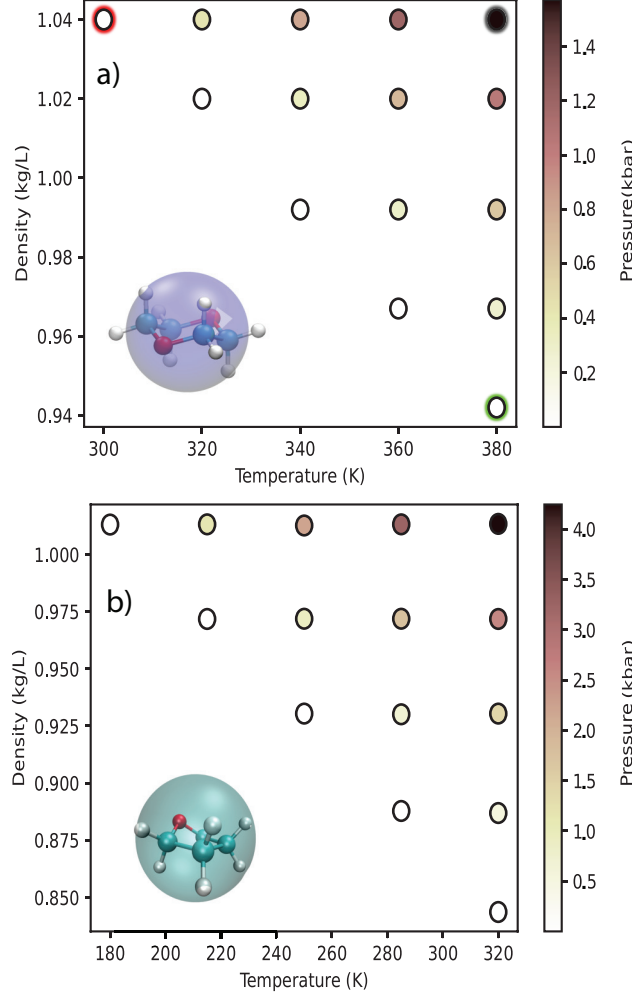


Figure 1: Simulated state points, (T, ρ) , for dioxane (a) and THF (b). The color of each circle indicates the pressure of the state point according to the color bar at right. In particular, the circles along the diagonal correspond to ambient pressure, $P = 1 \text{ bar} = 0.001 \text{ kbar}$. The reference state, $(T_{\text{ref}}, \rho_{\text{ref}})$, corresponds to the top right circle in each matrix. The top left (red circle) and bottom right (green circle) in panel a correspond to the extreme state points considered later. The inset of each panel indicates the 1-site CG representation of the corresponding molecule.

diagonal in Fig. 1. We determined the equilibrium internal pressure, $P_{\text{int}}(\bar{\rho}_T^0, T') = \langle P_{\text{int}} \rangle_{\bar{\rho}_T^0, T'}$, for each of these state points from these constant NVT simulations. Finally, we performed a corresponding constant NPT simulation at each of these 10 state points with the temperature, T' , and the external pressure $P'_{\text{ext}} = P_{\text{int}}(\bar{\rho}_T^0, T')$. Tables 1 and 2 report the pressure and density of each simulated state point for dioxane and THF, respectively.

The AA simulations are quite consistent with experimental measurements for the bulk

density of these liquids at $T_{\text{rm}} = 298$ K and ambient pressure, $P_{\text{ext}} = 1$ bar.¹²⁵ Specifically, the simulated equilibrium density for dioxane is 1.04 kg/L under these conditions, while experimental measurements report a density of 1.03 kg/L. Similarly, experimental measurements for THF report a density of 0.88 kg/L under these conditions, while our simulations at the closest state point, $T = 285$ K and $P_{\text{ext}} = 1$ bar, predict a density of 0.89 kg/L.

Table 1: Simulated dioxane state points in Fig. 1a. The top row indicates the simulated temperatures, while the entries in the corresponding column report the equilibrium density and pressure, $(\bar{\rho}, P)$, of each simulation in units of kg/L and bar, respectively.

300 K	320 K	340 K	360 K	380 K
1.041, 1	1.043, 447	1.042, 811	1.042, 1193	1.041, 1569
X	1.016, 1	1.016, 346	1.017, 714	1.017, 1053
X	X	0.992, 1	0.992, 307	0.992, 623
X	X	X	0.967, 1	0.967, 278
X	X	X	X	0.942, 1

Table 2: Simulated THF state points in Fig. 1b. The top row indicates the simulated temperatures, while the entries in the corresponding column report the equilibrium density and pressure, $(\bar{\rho}, P)$, of each simulation in units of kg/L and bar, respectively.

180 K	215 K	250 K	285 K	320 K
1.013, 1	1.013, 1152	1.013, 2222	1.013, 3257	1.013, 4245
X	0.972, 1	0.972, 911	0.972, 1770	0.972, 2590
X	X	0.930, 1	0.930, 714	0.930, 1413
X	X	X	0.888, 1	0.887, 555
X	X	X	X	0.844, 1

We constructed the configurations for our initial NPT simulations by employing the `gromacs gmx_solvate` command to insert molecules into a cubic box with a volume of 96.6 nm³. We then performed an energy minimization with the resulting configuration. We performed each constant NPT simulation for 32 ns and analyzed the last 30 ns after discarding the first 2 ns as equilibration.

Given a constant NPT simulation at temperature, T , and ambient pressure, $P = 1$ bar, we determined the initial condition for a subsequent constant NVT simulation as the sampled

configuration with an instantaneous density, ρ , that was closest to the mean, $\bar{\rho}_T^0$. After performing an energy minimization with this configuration, we then simulated the resulting configuration in the constant NVT ensemble at the target temperature, $T' > T$, for 7 ns and discarded the first 2 ns for equilibration. After the equilibration period, we sampled our simulations every 1 ps.

3.2 CG Models

We developed several different potentials for 1-site CG models of dioxane and THF. We initially employed the bottom-up open-source coarse-graining software (BOCS) package v. 5.0 to parameterize a potential, $U(T, \rho) = U_R(T, \rho) + U_V(T, \rho)$, for each state point (T, ρ) in Fig. 1. Here and in the following, we often suppress the explicit arguments of potentials in order to focus on their parametric dependence upon the thermodynamic state point for which they were parameterized, e.g., $U_R(T, \rho) \equiv U_R(\mathbf{R}; T, \rho) = \sum_{(I,J)} U_2(R_{IJ}; T, \rho) \equiv \sum_{(I,J)} U_2(T, \rho)$.

We first mapped the corresponding AA constant NPT simulation to the CG resolution by representing each molecule with its mass center and determining the net force on each molecule. We determined the optimal MS-CG pair potential, $U_2(T, \rho)$, for each state point by minimizing Eq. (18) via the default singular value decomposition solver^{126,127} in BOCS. In this force-matching calculation we represented the pair potential with cubic spline basis functions on a uniformly spaced grid over the interval $0.0 \text{ nm} \leq R \leq 1.4 \text{ nm}$ with $\Delta R = 0.02 \text{ nm}$. We represented the volume potential, $U_V(T, \rho)$, with the simple form in Eq. (17). We determined the corresponding parameters by performing self-consistent pressure-matching, as described in Ref. 88. In particular, we employed the `mdrun -rerun` feature from `gromacs/2016.6` to obtain pressure information from the mapped atomistic ensemble for the initial step of pressure-matching. We employed 6 iterations at each state point in order to adequately converge the optimal ψ parameters.

Given the potentials, $U(T, \rho)$, that were optimized for each dioxane state point, we de-

veloped a CG potential, $U_{\text{lin}}(T, \rho)$, by linearly extrapolating about a fixed reference state point. We identify the top right state point in Fig. 1a, i.e., ($T_{\text{ref}} = 380$ K, $\rho_{\text{ref}} = 1.04$ kg/L), as a reference state for dioxane and defined $U_{2;\text{ref}} = U_2(T_{\text{ref}}, \rho_{\text{ref}})$. In order to quantify the density- and temperature-dependence of the MS-CG pair potentials, $U_2(T, \rho)$, we calculated the finite differences

$$m_T(T) \equiv \Delta U_2(T, \rho_{\text{ref}})/\Delta T \equiv \left[U_2(T, \rho_{\text{ref}}) - U_2(T_{\text{ref}}, \rho_{\text{ref}}) \right] / \left[T - T_{\text{ref}} \right] \quad (20)$$

$$m_\rho(\rho) \equiv \Delta U_2(T_{\text{ref}}, \rho)/\Delta \rho \equiv \left[U_2(T_{\text{ref}}, \rho) - U_2(T_{\text{ref}}, \rho_{\text{ref}}) \right] / \left[\rho - \rho_{\text{ref}} \right] \quad (21)$$

for each $T < T_{\text{ref}}$ and $\rho < \rho_{\text{ref}}$. We defined $\bar{m}_T$ and $\bar{m}_\rho$ by simply averaging the resulting 4 finite differences $m_T(T)$ and $m_\rho(\rho)$. We then modeled linear variations in the MS-CG pair potentials according to

$$U_{2;\text{lin}}(T, \rho) = U_{2;\text{ref}} + \bar{m}_T(T - T_{\text{ref}}) + \bar{m}_\rho(\rho - \rho_{\text{ref}}), \quad (22)$$

which defines a corresponding interaction potential, $U_{\text{R};\text{lin}}(T, \rho)$. We combined $U_{\text{R};\text{lin}}(T, \rho)$ with the volume potential, $U_V(T, \rho)$, that was optimized for the original MS-CG pair potential, $U_2(T, \rho)$, to define the extrapolated total potential: $U_{\text{lin}}(T, \rho) = U_{\text{R};\text{lin}}(T, \rho) + U_V(T, \rho)$. In the following, we consider the radial distribution functions and pressure-volume equations of state obtained from CG simulations with this potential, $U_{\text{lin}}(T, \rho)$.

We also developed a density-decoupled (DD) potential, U_{DD} , for each liquid by neglecting the density-dependence of the interaction potentials. In this case, we defined the density-independent interaction potential for each temperature, T , by the MS-CG interaction potential, $U_{\text{R}}(T, \rho_{\text{ref}})$, determined for the reference density, ρ_{ref} , i.e., $U_{\text{R};\text{DD}}(T) \equiv U_{\text{R}}(T, \rho_{\text{ref}})$. Given this fixed density-independent interaction potential, $U_{\text{R};\text{DD}}(T)$, we repeated our self-consistent pressure-matching calculations to determine a new volume potential, $U_{\text{V};\text{DD}}(T, \rho)$,

for each state point (T, ρ) in Fig. 1. We define the DD potential

$$U_{\text{DD}}(T, \rho) \equiv U_{\text{R;DD}}(T) + U_{\text{V;DD}}(T, \rho). \quad (23)$$

In contrast to the original MS-CG models, the volume forces for the DD model, $F_{\text{V;DD}}(V; T, \rho) = -dU_{\text{V;DD}}(V; T, \rho)/dV$, are well described by a continuous function of V that is independent of the density, ρ , of the state point (T, ρ) . Consequently, we approximated $F_{\text{V;DD}}(V; T, \rho)$ with a global function of V and T :

$$F_{\text{V;glob}}(V; T) \equiv (a_0 + a_1 \Delta T) V^2 + (b_0 + b_1 \Delta T) V + (c_0 + c_1 \Delta T), \quad (24)$$

where $\{a_0, a_1, b_0, b_1, c_0, c_1\}$ are six free parameters and $\Delta T = T - T_{\text{ref}}$. We employed the Python curve fit function in the `scipy.optimize` package¹²⁸ to determine the six free parameters that best approximated the volume forces $F_{\text{V;DD}}(V; T, \rho)$ for the 15 state points in Fig. 1. This determines a single global potential, U_{glob} , for modeling the liquid phase that depends explicitly upon $\mathbf{R}$ and V , while depending parametrically upon T :

$$U_{\text{glob}}(\mathbf{R}, V; T) = U_{\text{R;DD}}(\mathbf{R}; T) + U_{\text{V;glob}}(V; T), \quad (25)$$

where $U_{\text{V;glob}}(V; T)$ is the global volume potential obtained by integrating Eq. (24). We constructed this global model for dioxane about the single reference state point, $(T_{\text{ref}} = 380 \text{ K}, \rho_{\text{ref}} = 1.04 \text{ kg/L})$. In the case of THF, we constructed two different global models based upon two different reference state points, $(T_{\text{ref};1} = 320 \text{ K}, \rho_{\text{ref};1} = 1.01 \text{ kg/L})$ and $(T_{\text{ref};2} = 320 \text{ K}, \rho_{\text{ref};2} = 0.93 \text{ kg/L})$. Table 3 reports the parameters for the corresponding global volume potentials.

Table 3: Parameters for the global volume potentials, $U_{V;\text{glob}}$, defined by Eq. (24).

DD Model	Dioxane	THF	THF
T_{ref} (K)	380	320	320
ρ_{ref} (kg/L)	1.04	1.01	0.93
a_0 (bar/nm ⁶)	2.72	3.59	5.84
b_0 (bar/nm ³)	-471.51	-550.68	-1,152.21
c_0 (bar)	11,930.03	4,610.14	49,411.44
a_1 (bar/nm ⁶ K)	-0.053	-0.047	-0.051
b_1 (bar/nm ³ K)	10.795	9.362	9.796
c_1 (bar / K)	-560.448	-486.056	-487.652

3.3 CG Simulations

We simulated CG models in the constant NVT ensemble with Gromacs 2019.6. We employed a 1 fs timestep for propagating dynamics with the leapfrog integrator, while employing the Bussi stochastic thermostat¹¹⁸ with a relaxation time of 0.5 ps. We truncated interactions at 1.4 nm and employed a neighbor list of 1.5 nm that was updated every step. We performed each constant NVT simulation for 7 ns, discarded the first 2 ns for equilibration, and sampled the remainder of the simulation after every 1 ps.

We simulated CG models in the constant NPT ensemble with Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS),^{129,130} while employing the BOCS package that implements volume potentials.^{88,111} We propagated dynamics by integrating the Martyna-Tuckerman-Tobias-Klein^{131,132} equations of motion with a 1 fs time step. We employed time scales of 100 fs and 1 ps for the thermostat and barostat, respectively, and employed the default chain length ($n = 3$) for the Nosé-Hoover chain.^{133,134} We employed the analytic functional form of Eq. (17) to represent the volume potential in all cases, except when using the global volume potential, $U_{V;\text{glob}}(V; T)$. In this case, we employed a tabulated representation of Eq. (24) with a grid space $\Delta V = 0.001 \text{ nm}^3$. We employed a neighbor-list with a cut-off of 1.4 nm that was updated every time step. We performed each constant NPT simulation for 5 ns, while sampling the simulation after every 1 ps.

The results section examines the relationship between the density-dependence of pair

and volume potentials. We performed this analysis by first creating a scatter plot of the sampled volumes and internal pressures, $\{(V, P_{\text{int}})\}$, where the internal pressure, $P_{\text{int}} = Nk_{\text{B}}T/V + P_{\text{xs}}$, includes an ideal term, $Nk_{\text{B}}T/V$, and an excess contribution, $P_{\text{xs}} = \mathcal{W}_{\text{R}} + F_{\text{V}}$, that reflects both the virial, $\mathcal{W}_{\text{R}} = \frac{1}{3V} \sum_{(I,J)} F_2(R_{IJ})R_{IJ}$, and also the volume force, F_{V} . We subtracted the volume force, $F_{\text{V}} = F_{\text{V}}(V)$, from each sampled internal pressure to determine $P_0 = P_{\text{int}} - F_{\text{V}} = Nk_{\text{B}}T/V + \mathcal{W}_{\text{R}}$. We employed the Python `scipy.optimize` package to fit the resulting scatter plot $\{(V, P_0)\}$ to determine $\bar{P}_0(V) = \langle P_0 \rangle_V$ as a linear function of V . Given two CG potentials, $U_1 = U_{\text{R1}} + U_{\text{V1}}$ and $U_2 = U_{\text{R2}} + U_{\text{V2}}$, that are defined for the same state point, we determined the difference in the corresponding average virials, $\Delta\bar{\mathcal{W}}_{\text{R}}(V) = \bar{P}_{0;2}(V) - \bar{P}_{0;1}(V)$. Finally, we correlated $\Delta\bar{\mathcal{W}}_{\text{R}}(V)$ with the difference in the corresponding volume forces, $\Delta F_{\text{V}}(V) = F_{\text{V};2}(V) - F_{\text{V};1}(V)$.

4 Results and Discussion

4.1 Analysis of dioxane reference state

In this work, we develop 1-site CG models for simulating liquid 1,4-dioxane across the range of state points indicated in Fig. 1a. In particular, we consider the temperature range $300 \text{ K} \leq T \leq 380 \text{ K}$ for which dioxane is a liquid at ambient pressure. Additionally, we consider a rather wide range of pressures: $1 \text{ bar} \leq P \leq 1569 \text{ bar}$. We select the state point, $T_{\text{ref}} = 380 \text{ K}$ and $P_{\text{ref}} = 1569 \text{ bar}$, as a reference state because it allows us study the variation of CG potentials over a wide temperature and density range.

Figure 2 presents results from AA constant NPT simulations of dioxane at this reference state point, $T_{\text{ref}} = 380 \text{ K}$ and $P_{\text{ref}} = 1569 \text{ bar}$. The solid black curve in Fig. 2a presents the AA radial distribution function (rdf) for the dioxane mass center. (Except when otherwise indicated, we report rdf's from constant NPT simulations.) The rdf vanishes for $r \leq 0.38 \text{ nm}$ and features a prominent first peak at $r_0 \approx 0.57 \text{ nm}$, as well as a pronounced minimum at $r_1 \approx 0.76 \text{ nm}$. The AA rdf demonstrates rather long-ranged structure at this high external

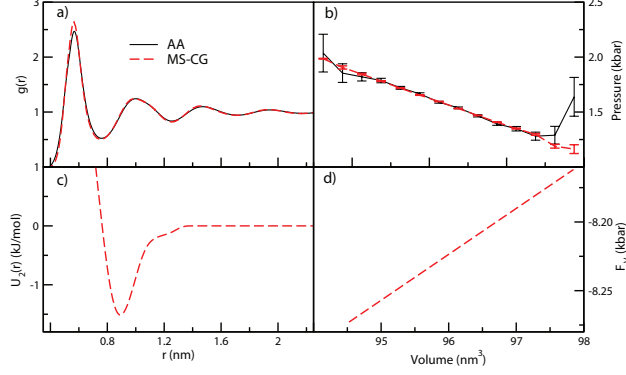


Figure 2: Analysis of AA and 1-site CG models for dioxane at the reference state point. Solid black curves indicate results for the AA model, while dashed red curves indicate results for the CG model that was optimized for this reference state point. Panels a and b present the mass center rdf and pressure-volume equation of state. Panels c and d present the MS-CG pair potential and volume force.

pressure with discernible second, third, and fourth peaks at $r \approx 1.0$ nm, 1.5 nm, and 1.9 nm, respectively. The solid black curve in Fig. 2b presents the simulated pressure-volume equation of state for the AA model. At the elevated pressure $P_{\text{ref}} = 1569$ bar, the average volume of the AA model is 96.2 nm^3 , which corresponds to an equilibrium density, $\rho_{\text{ref}} = 1.04 \text{ kg/L}$.

We employed the MS-CG force-matching variational principle to determine an optimal pair potential for this reference state point. Figure 2c presents the calculated MS-CG pair potential, U_2 . This pair potential features a single shallow well with a depth of approximately $0.48 k_B T_{\text{ref}}$, a minimum near $r_{\text{min}} \approx 0.9$ nm, and a slight shoulder over the distance range $1.0 \text{ nm} \leq r \leq 1.3$ nm. The pair potential vanishes for $r \geq 1.4$ nm by construction. It is rather striking that MS-CG pair potential is purely repulsive in the first solvation shell and that the attractive well coincides with the second solvation shell.

As is often observed for bottom-up CG models,^{20,68,87} constant NPT simulations with the pair-additive MS-CG potential dramatically over-estimate the internal pressure of the AA model and rapidly vaporize. Consequently, we employed self-consistent pressure-matching to determine a volume potential, $U_V(V)$, of the form given by Eq. (17). Figure 2d presents the resulting volume force $F_V(V) = -dU_V(V)/dV$. The volume force reduces the internal pressure by more than 8000 bar over the simulated volume range. The (absolute) magnitude

of the volume force slightly decreases as the system expands. Note that here and in the following, we plot $F_V(V)$ only over the volume range sampled by the corresponding AA simulation.

The dashed red curves in Figs. 2a and 2b present the rdf and pressure-volume equation of state obtained from constant NPT simulations of the resulting CG model. The MS-CG model slightly overestimates the height of the first peak in the AA rdf. Aside from this slight discrepancy, the MS-CG model quantitatively reproduces the AA rdf. The MS-CG model quantitatively reproduces the AA pressure-volume equation of state by construction.

4.2 Density- and temperature-variations in effective potentials

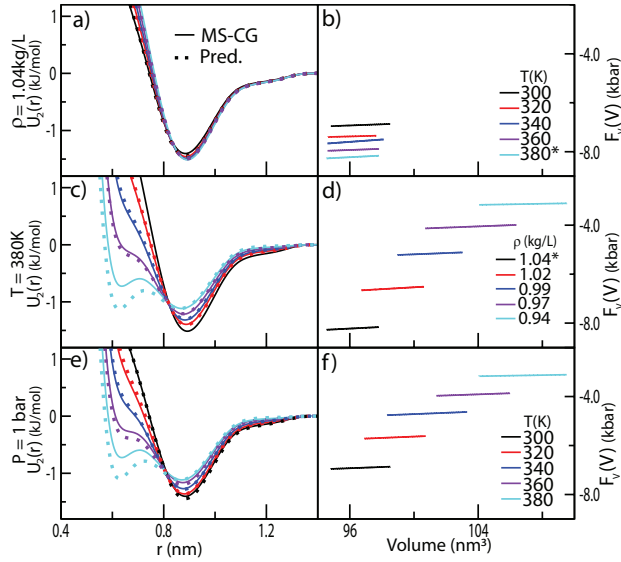


Figure 3: State-point dependent potentials for 1-site CG model of dioxane. The solid curves in the left and right columns present MS-CG pair potentials, U_2 , and corresponding volume forces, F_V , that were optimized for each state point (T, ρ) . The top row analyzes temperature-variations at fixed density for the state points (T, ρ_{ref}) . The middle row analyzes density-variations at fixed temperature for the state points (T_{ref}, ρ) . The bottom row analyzes simultaneous temperature- and density variations (T, ρ) at fixed external pressure $P_{\text{ext}} = 1$ bar. The legends are specific to each row, while ‘*’ identifies the reference state $(T_{\text{ref}}, \rho_{\text{ref}})$ that is analyzed in Fig. 2. The dotted curves in the left column present pair potentials, $U_{2;\text{lin}}$, that were predicted for each state point via linear extrapolation according to Eq. (27).

We employed force-matching and self-consistent pressure-matching to determine an op-

timal pair potential, $U_2(T, \rho)$, and volume potential, $U_V(T, \rho)$, for a 1-site CG model of dioxane at each state point, (T, ρ) , in Fig. 1a. The solid curves in the left and right columns of Fig. 3 present the resulting pair potentials and volume forces, respectively. The top row of Fig. 3 considers temperature-variations at constant density, while the middle row considers density-variations at constant temperature. The bottom row considers simultaneous temperature- and density-variations at constant external pressure.

Figure 3a investigates the temperature-dependence of the MS-CG pair potentials, $U_2(T, \rho_{\text{ref}})$, at the fixed reference density, $\rho_{\text{ref}} = 1.04 \text{ kg/L}$, which corresponds to the top row of Fig. 1a. The cyan curve corresponds to the reference state with $T_{\text{ref}} = 380 \text{ K}$ that was analyzed in Fig. 2. As the temperature increases at constant density, the minimum of the potential very slightly deepens, while the pair potentials also become noticeably more repulsive at short distances, $r < r_{\text{min}}$. However, the minimum remains at $r_{\text{min}} \approx 0.9 \text{ nm}$ and the shape of the potential varies relatively little.

Figure 3b presents the corresponding volume forces, $F_V(T, \rho_{\text{ref}})$, that were optimized to reproduce the AA pressure equation of state at the state points (T, ρ_{ref}) . As the temperature increases from 300 K to 380 K at the fixed density, the internal pressure of the AA model increases from 1 bar to 1569 bar. Conversely, the volume forces become increasingly attractive as the temperature increases. The absolute magnitude of the volume force increases from 6910 bar to 8217 bar over this temperature range in order to compensate for the increasingly repulsive pair potentials.

Figure 3c investigates the density-dependence of the MS-CG pair potentials, $U_2(T_{\text{ref}}, \rho)$, at the fixed reference temperature, $T_{\text{ref}} = 380 \text{ K}$, which corresponds to the right column of Fig. 1a. The black curve corresponds to the reference state point, $\rho_{\text{ref}} = 1.04 \text{ kg/L}$, that was analyzed in Fig. 2. As previously observed, the MS-CG potentials vary dramatically with density.⁴¹ The global minimum of the pair potential becomes increasingly shallow and slightly shifts to smaller distances as the density decreases at constant temperature, but remains in the second solvation shell. Even more strikingly, as the density decreases the

potential becomes much softer and a second contact minimum appears in the first solvation shell. This second contact minimum occurs at a distance, $r \approx 0.64$ nm, that is slightly greater than the first peak of the AA rdf. This second minimum is almost as deep as the global minimum at the lowest density, $\rho = 0.94$ kg/L, which corresponds to ambient pressure.

Figure 3d presents the corresponding volume forces, $F_V(T_{\text{ref}}, \rho)$, that were parameterized for the state points (T_{ref}, ρ) . These volume forces become much more attractive as the density increases. Each volume force has been parameterized to reproduce the same AA pressure-volume equation of state, $\bar{p}_{\text{int}}(V, T_{\text{ref}})$, for the fixed reference temperature, $T_{\text{ref}} = 380$ K. Because they have been calculated from constant NPT simulations with different equilibrium densities, ρ , the volume forces are determined over different volume ranges. Moreover, the calculated volume forces differ in overlapping volume ranges because they correspond to different pair potentials. For instance, the red and blue curves correspond to models that were parameterized for $\rho = 1.02$ kg/L and 0.99 kg/L, respectively. Both models reproduce the internal pressure of the AA model at the volume $V = 100$ nm³. Because the model parameterized at $\rho = 1.02$ kg/L employs a more repulsive pair potential (i.e., the red curve in Fig. 3c), this model requires a more attractive volume force to match the AA internal pressure (i.e., the red curve in Fig. 3d). Therefore, the marked variation in the volume forces of Fig. 3d corresponds to the dramatic density-dependence of the pair potentials in Fig. 3c.

Figures 3e and 3f present the pair potentials and volume forces that were calculated from AA simulations at the temperature range $300 \text{ K} \leq T \leq 380 \text{ K}$ and constant external pressure $P_{\text{ext}} = 1$ bar. As previously observed,⁴¹ the pair potentials in Fig. 3e become increasingly attractive as the temperature increases at constant external pressure. Similarly, the variation in the volume forces in Fig. 3f appears more similar to the density-variation in Fig. 3d than the temperature-variation in Fig. 3b. Thus, the calculated potentials appear much more sensitive to density than temperature.

4.3 g-YBG analysis of effective potentials

The generalized-Yvon-Born-Green (YBG) formalism provides physical insight into the MS-CG pair potentials, $U_2(r)$, and their sensitivity to temperature and density.^{95,96,98,110} Given the approximate potential in Eq. (15), the MS-CG method determines the pair force, $F_2(r) = -dU_2(r)/dr$, according to a force-balance equation^{95,135} that is analogous to the YBG equation from liquid state theory¹³⁶

$$\overline{F}_2(r) = F_2(r) + \int dr' K(r, r') F_2(r'). \quad (26)$$

Here $\overline{F}_2(r) = -dw_{AA}(r)/dr = -k_B T g'_{AA}(r)/g_{AA}(r)$ is the AA pair mean force, i.e., the magnitude of the average force on each molecule of a pair when they are separated by a distance, r .¹³⁷ The kernel, $K(r, r')$, of the integral term is a many-body correlation function that describes the density of surrounding molecules. Importantly, both $\overline{F}_2$ and K are completely specified by the underlying AA model and the CG mapping. Thus, the MS-CG method determines the pair force by decomposing the AA pair mean force into two contributions:⁹⁸ (1) a direct contribution, $F_2(r)$, that describes the interaction between the pair; and (2) an indirect contribution, $F_{ind}(r) = \int dr' K(r, r') F_2(r')$, that describes the interactions of the pair with the surrounding many-body environment.^{97,98}

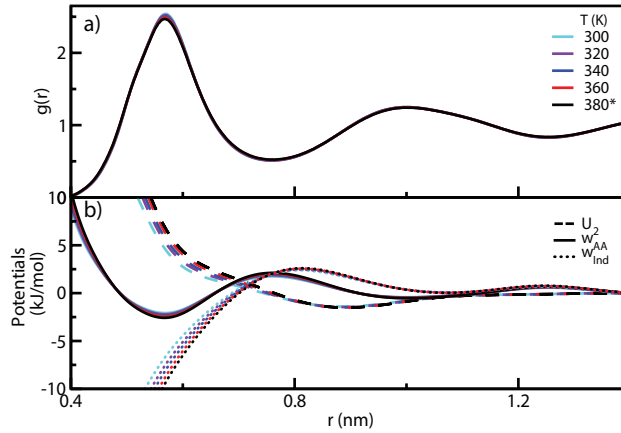


Figure 4: Analysis of temperature-dependent MS-CG potentials at constant density. (a) Temperature-dependence of AA rdfs. (b) Decomposition of pair potentials of mean force, w_{AA} , (solid) into direct contributions, U_2 , (dashed) and indirect contributions, w_{ind} (dotted).

Figure 4 employs this generalized-YBG (g-YBG) equation to interpret the variation in the MS-CG pair potentials with decreasing temperature at constant density. Figure 4a presents the mapped AA rdf's, $g_{AA}(r)$, for dioxane at the fixed reference density and each simulated temperature. As the temperature decreases at constant density, the first peak of the AA rdf at $r_0 \approx 0.57$ nm very slightly increases, while the minimum at $r_1 \approx 0.76$ nm almost imperceptibly decreases. Otherwise the AA rdf appears independent of temperature at the fixed density. Consequently, the solid curves in Fig. 4b demonstrate that the AA pair potential of mean force (ppmf), $w_{AA}(r) = -k_B T \ln g_{AA}(r)$, varies almost linearly with temperature. In particular, as the temperature decreases, $w_{AA}(r)$ becomes slightly less repulsive for short distances $r < r_0$ near contact, but becomes slightly less attractive over the rest of the first solvation shell, $r_0 < r < r_1$.

The dashed and dotted curves in Fig. 4b present the direct, U_2 , and indirect, w_{ind} , contributions to the ppmf, w_{AA} , respectively. The indirect contribution, w_{ind} , is determined by simply integrating (the negative of) F_{ind} . Within the first solvation shell, $r \leq r_1$, the indirect contribution to the ppmf is purely attractive (i.e., $F_{\text{ind}}(r) = -dw_{\text{ind}}(r)/dr < 0$) due to crowding by the surrounding liquid. This solvent-induced effective attraction between the pair is analogous to the Asakura-Oosawa depletion-induced attraction between colloids due to a repulsive solvent^{138–140} and becomes quite large as $r \rightarrow 0$.

At large distances, $w_{AA}(r) \approx w_{\text{ind}}(r)$ because $U_2(r) \approx 0$ for $r \gtrsim 1.0$ nm. At shorter distances, the direct and indirect contributions to the ppmf tend to be anti-correlated and both make significant contributions. For $r > r_0$, the indirect contribution appears larger than the direct contribution. However, as molecules approach contact at very short distances, $r < r_0$, the direct contribution begins to dominate.

The Supporting Information (SI) demonstrates that the AA many-body correlation function, $K(r, r')$, varies little with temperature at constant density. Consequently, the temperature-dependence of the MS-CG pair potentials stems from the temperature-dependence of the ppmf. In particular, as the temperature decreases, $U_2(r)$ becomes slightly more attractive

for $r < r_0$ because the ppmf also becomes more attractive at such short distances. As the direct repulsion between the pair decreases, the solvent-induced attraction, w_{ind} , also decreases from reduced repulsion with the surrounding environment. Because w_{ind} dominates U_2 at slightly larger distances, $r_0 < r < r_1$, the ppmf becomes more repulsive over the remainder of the first solvation shell. (More precisely, the dominance of w_{ind} and the increasing repulsion in the ppmf for $r_0 < r < r_1$ cause U_2 to be slightly less repulsive over this distance range.) The MS-CG pair potential demonstrates somewhat greater sensitivity to temperature than the ppmf because the variations in U_2 and w_{ind} tend to cancel each other.

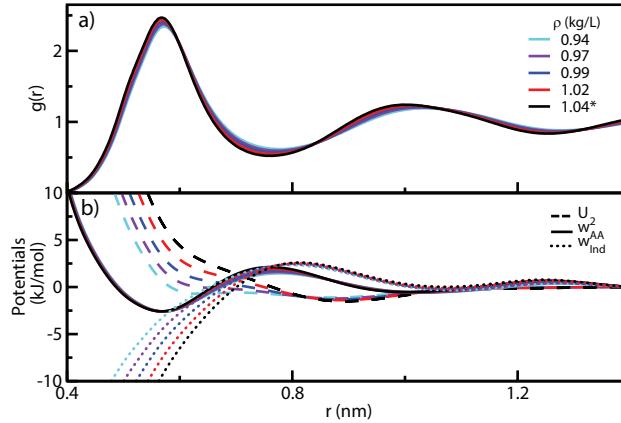


Figure 5: Analysis of density-dependent MS-CG potentials at constant temperature. (a) Density-dependence of AA rdfs. (b) Decomposition of pair potentials of mean force, w_{AA} , (solid) into direct contributions, U_2 , (dashed) and indirect contributions, w_{ind} (dotted).

Figure 5 similarly employs the g-YBG formalism to interpret the variations in the MS-CG pair potentials with decreasing density at constant temperature. As the density decreases, the AA rdf systematically shifts to larger distances, while the maxima and minima both shift towards 1. Accordingly, the ppmf becomes very slightly more repulsive over the entire first solvation shell, $r < r_1$. This suggests that the MS-CG pair potential should also become more repulsive at short distances, which is in direct contrast to the density-variation that is observed in Fig. 3. The SI demonstrates that, if K were density-independent, then the MS-CG pair potential would indeed become more repulsive at short distances as the density decreases. However, the SI also demonstrates that the many-body correlation func-

tion, $K(r, r')$, varies significantly with density. This density-variation significantly reduces the strength of the solvent-induced effective attraction, w_{ind} , as the liquid expands. Consequently, the MS-CG pair potential must become much more attractive at short distances in order to compensate for the reduced solvent-induced effective attraction. In particular, this increasing direct attraction gives rise to the contact minimum that emerges near $r \approx r_0$ as the density decreases at constant temperature. Interestingly, the SI demonstrates that this contact minimum corresponds to a feature in the many-body correlation matrix, $K(r, r')$. Thus, the g-YBG formalism traces the dramatic density-variation in the MS-CG pair potentials and, in particular, the contact minimum to the density-dependence of the many-body liquid structure.

4.4 Structural and thermodynamic fidelity

Figure 6 assesses the structural fidelity and thermodynamic accuracy of the 1-site MS-CG models in constant NPT simulations at 1 bar external pressure. The solid curves present the results of AA simulations, while the dashed curves present the results of CG simulations with the pair and volume potentials that were optimized for each state point. In particular, Figs. 6a and 6b demonstrate that these MS-CG models very accurately reproduce the AA rdf's at $T = 300$ K and 380 K, respectively. As in Fig. 2a, the largest errors occur in the first peak. The MS-CG models overestimate the first peak of the AA rdf by only 10% at 300 K and reproduce the first peak even more accurately at 380 K. The dashed curves in Fig. 6c, which are obscured by the solid AA curves, demonstrate that the MS-CG models also quantitatively reproduce the AA pressure-volume equations of state by construction.

Given the observed temperature- and density-variations in the calculated pair potentials from Figs. 3a and 3c, we calculated empirical estimates for the temperature- and density-derivatives of the MS-CG pair potentials, $\overline{m}_T \approx (\partial U_2 / \partial T)_\rho$ and $\overline{m}_\rho \approx (\partial U_2 / \partial \rho)_T$, according to Eqs. (20) and (21). We modeled the MS-CG pair potentials as a linear function of T and

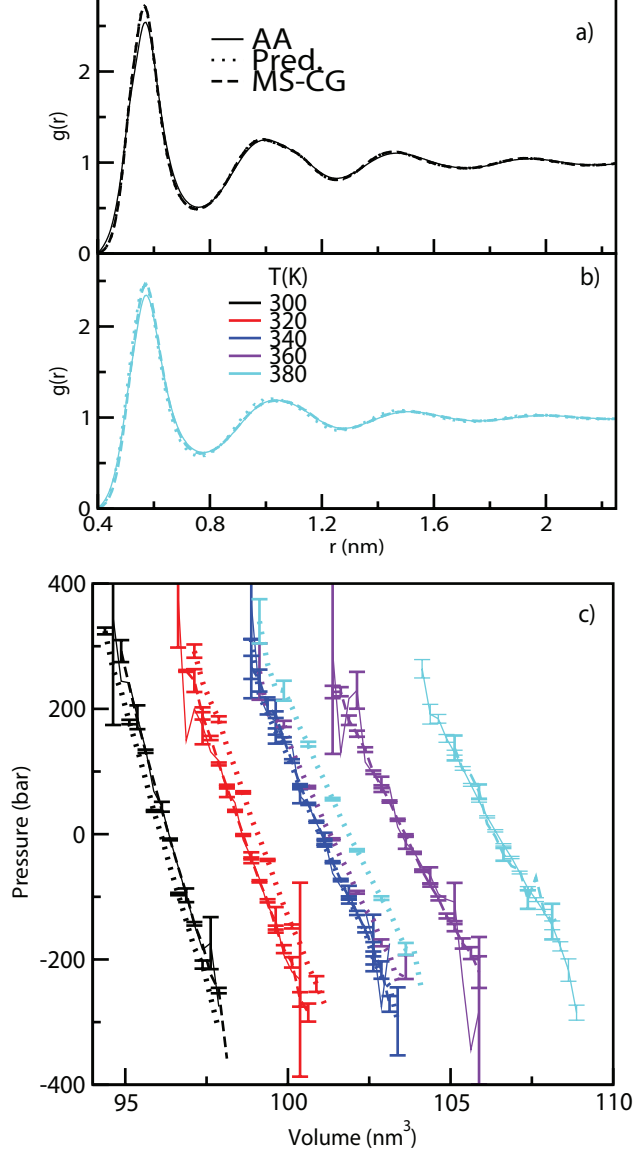


Figure 6: Accuracy of 1-site dioxane models at 1 bar external pressure. Panels a and b present rdf's for the molecular mass center from simulations at $T = 300$ K and 380 K, respectively. Panel c presents simulated pressure-volume equations of state. The solid curves present results for the AA model, while the dashed curves present results for the MS-CG model with the potential, $U(T, \rho) = U_R(T, \rho) + U_V(T, \rho)$, that was optimized for the specified state point (T, ρ) . The dotted curves present results for CG models with the predicted interaction potential, $U_{\text{lin}}(T, \rho) = U_{R;\text{lin}}(T, \rho) + U_V(T, \rho)$.

ρ according to

$$U_{2;\text{lin}}(T, \rho) = U_{2;\text{ref}} + \bar{m}_T (T - T_{\text{ref}}) + \bar{m}_\rho (\rho - \rho_{\text{ref}}), \quad (27)$$

where $U_{2;\text{ref}} \equiv U_2(T_{\text{ref}}, \rho_{\text{ref}})$ is the pair potential in Fig. 2. The dotted curves in Figs. 3a, 3c,

3e present the resulting pair potentials. The linear predictions very accurately reproduce the modest temperature-variations observed in Fig. 3a, as well as the density-variations in Fig. 3c for $r > r_{\min}$. Moreover, the linear predictions also accurately reproduce the density-variations for $r < r_{\min}$ until the second minimum begins to emerge at lower densities. Interestingly, the linear prediction over-estimates the depth of this minimum. The SI indicates that m_T is approximately independent of T , while m_ρ varies approximately linearly with ρ over this range. Consequently, expanding the Taylor series in Eq. (27) to second order in $\Delta\rho$ provides a slightly improved model for the calculated pair potentials.

The dotted curves in Fig. 6 present the results of CG simulations with these predicted pair potentials, while still employing the volume potentials that were optimized for each state point. Despite the noticeable differences between the calculated MS-CG and predicted pair potentials for $T = 380$ K in Fig. 3e, Figs. 6a and 6b demonstrate that the predicted potentials very accurately reproduce the pair structure of the AA model. Similarly, Fig. 6c demonstrates that the predicted potentials quite accurately reproduce the AA equation of state for $T \leq 340$ K. However, for $T \geq 360$ K the predicted potentials are significantly more attractive than the corresponding MS-CG potentials. Accordingly, simulations that employ the predicted potentials with the original volume potentials significantly overestimate the AA density for $T \geq 360$ K at ambient pressure.

4.5 Transferability of effective potentials

While Fig. 6 assesses the accuracy of the MS-CG potentials at the state point for which they were parameterized, Fig. 7 assesses their transferability to other state points. In particular, the top row of Fig. 7 analyzes the importance of temperature-variations in the MS-CG potentials for simulations at the fixed reference density, ρ_{ref} . The inset of Fig. 7a compares the pair potential, $U_2(T_{\text{ref}}) \equiv U_2(T_{\text{ref}}, \rho_{\text{ref}})$, that was optimized for the reference temperature $T_{\text{ref}} = 380$ K with the pair potential, $U_2(T_{\text{lo}}) \equiv U_2(T_{\text{lo}}, \rho_{\text{ref}})$, that was optimized for the lowest temperature, $T_{\text{lo}} = 300$ K. As already noted in Fig. 3a, as T decreases at constant

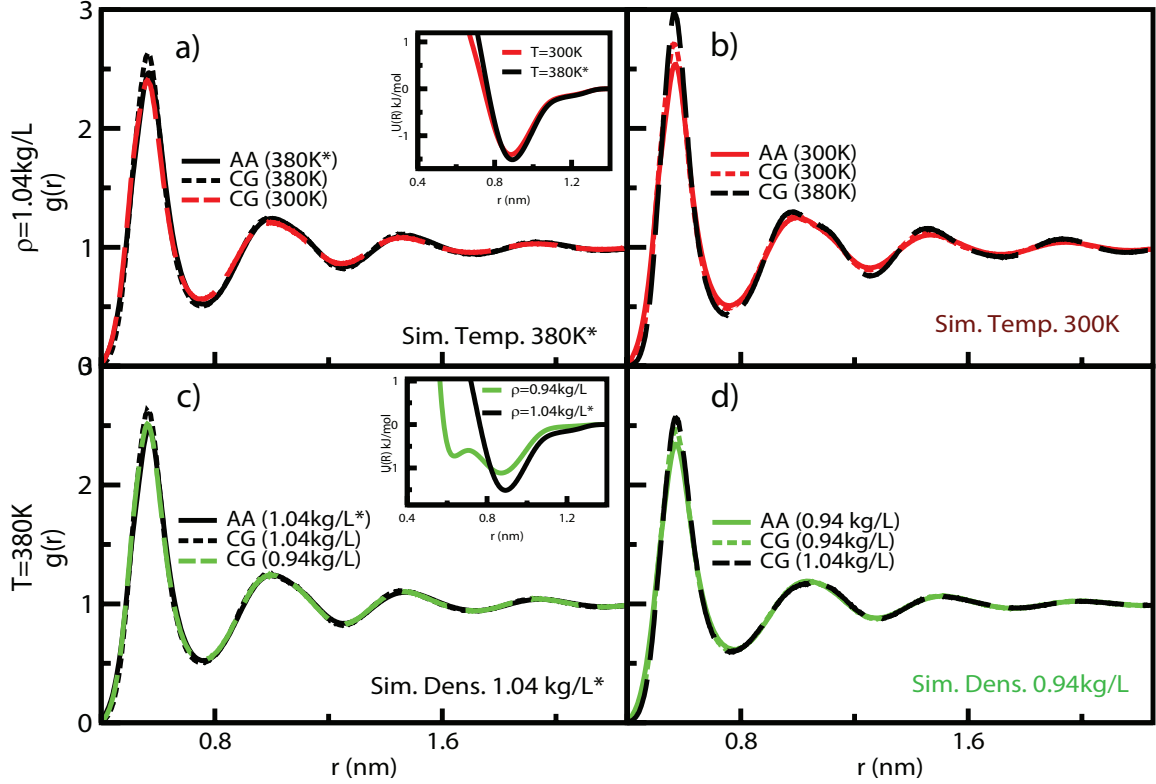


Figure 7: Transferability of 1-site MS-CG models for dioxane. The top row presents rdf's from simulations at the reference density, $\rho_{\text{ref}} = 1.04 \text{ kg/L}$: panels a and b correspond to simulations at the reference temperature, $T_{\text{ref}} = 380 \text{ K}$, and at the reduced temperature, $T_{\text{lo}} = 300 \text{ K}$, respectively. The inset of panel a compares the pair potentials that were calculated for the two state points. The bottom row presents rdf's from simulations at the reference temperature, $T_{\text{ref}} = 380 \text{ K}$: panels c and d correspond to simulations at the reference density, $\rho_{\text{ref}} = 1.04 \text{ kg/L}$, and at the reduced density, $\rho_{\text{lo}} = 0.94 \text{ kg/L}$, respectively. The inset of panel c compares the pair potentials that were calculated for the two state points. The black curves correspond to the reference state point, $(T_{\text{ref}}, \rho_{\text{ref}})$; the red curves correspond to the low temperature state, $(T_{\text{lo}}, \rho_{\text{ref}})$, which corresponds to the red circle in Fig. 1a; and the green curves correspond to the low density state, $(T_{\text{ref}}, \rho_{\text{lo}})$, which corresponds to the green circle in Fig. 1a. Solid curves present results from constant NPT AA simulations, while dashed curves present results from constant NVT CG simulations at the corresponding state point.

ρ , the MS-CG pair potential becomes slightly less repulsive at short range and slightly less attractive at longer range.

Figure 7a presents rdf's obtained from constant NVT simulations of two different CG models at the reference state point $(T_{\text{ref}}, \rho_{\text{ref}})$. As already observed in Fig. 2a, the potential, $U_2(T_{\text{ref}})$, that was optimized for this state point very accurately reproduces the AA rdf, although it does slightly over-estimate the height of the first peak. Surprisingly, constant NVT

simulations at T_{ref} with the low-temperature MS-CG potential, $U_2(T_{\text{lo}})$, also very accurately reproduce the AA rdf. In fact, the low-temperature MS-CG potential, $U_2(T_{\text{lo}})$, reproduces the first peak of the AA rdf more accurately than the reference MS-CG potential, $U_2(T_{\text{ref}})$, that was optimized for this state point. However, the low-temperature potential does appear to slightly under-estimate the subsequent minima and maxima of the AA rdf.

Figure 7b presents rdf's obtained from constant NVT simulations at the reduced temperature, T_{lo} , and fixed reference density, ρ_{ref} . As before, the MS-CG potential, $U_2(T_{\text{lo}})$, that was optimized for this state point very accurately reproduces the AA rdf, although it very slightly over-estimates the height of the first peak. In contrast, the MS-CG potential, $U_2(T_{\text{ref}})$, that was optimized for the higher temperature reference state provides a significantly less accurate description of the AA rdf. Because it is notably more repulsive at short distances, the reference potential, $U_2(T_{\text{ref}})$, significantly over-estimates the height of the first peak in the AA rdf and, more generally, over-estimates the pair structure of the AA model.

Thus, Figs. 7a and 7b demonstrate that the low-temperature MS-CG potential, $U_2(T_{\text{lo}})$, models intermolecular packing quite accurately at both low- and high-temperatures. Conversely, the high-temperature reference MS-CG potential, $U_2(T_{\text{ref}})$, accurately describes the high-temperature reference state, but provides a significantly less accurate description of intermolecular packing at the reduced temperature, T_{lo} . Thus, it appears that the potential that was optimized for lower temperatures provides improved transferability. In fact, several previous studies have also reported similar observations.^{27,28,90} Moreover, these results suggest that the relatively subtle temperature variations in the MS-CG potentials are important for accurately modeling the intermolecular structure of the AA model.

The bottom row of Fig. 7 then analyzes the importance of density-variations in the MS-CG potentials for simulations at the fixed reference temperature, $T_{\text{ref}} = 380$ K. The inset of Fig. 7c compares the pair potential, $U_2(\rho_{\text{ref}}) \equiv U_2(T_{\text{ref}}, \rho_{\text{ref}})$, that was optimized for the reference density, $\rho_{\text{ref}} = 1.04$ kg/L, with the pair potential, $U_2(\rho_{\text{lo}}) \equiv U_2(T_{\text{ref}}, \rho_{\text{lo}})$, that was optimized for the reduced density, $\rho_{\text{lo}} = 0.94$ kg/L. As already noted in Fig. 3c, as ρ

decreases at constant T_{ref} , the calculated MS-CG pair potential becomes significantly softer and develops a second minimum in the first solvation shell.

Figure 7c presents rdf's obtained from constant NVT simulations at the reference state point $(T_{\text{ref}}, \rho_{\text{ref}})$. While it is well known that rdf's are quite insensitive to many aspects of pair potentials,^{141,142} we find it somewhat surprising that such dramatically different pair potentials generate such similar rdf's at the reference temperature. In fact, the low-density MS-CG potential, $U_2(\rho_{\text{lo}})$, reproduces the AA rdf with slightly better accuracy than the reference MS-CG potential, $U_2(\rho_{\text{ref}})$, that was calculated for this state point.

Figure 7d presents the rdf's obtained from constant NVT simulations at the reduced density, ρ_{lo} , and fixed reference temperature, T_{ref} . In this case, the two MS-CG potentials generate almost identical rdf's.

Figures 7c and 7d demonstrate that the low-density MS-CG potential, $U_2(\rho_{\text{lo}})$, provides slightly better accuracy and transferability than the MS-CG potential, $U_2(\rho_{\text{ref}})$, that was optimized for the higher reference density. Perhaps more surprisingly, despite the dramatic differences in the inset of Fig. 7c, both MS-CG potentials reproduce the AA rdf quite accurately across the entire density range. Thus, it appears that the dramatic density-dependence in the MS-CG pair potentials is not important for modeling the pair structure of the AA model.

4.6 A globally transferable model

Figure 7 indicates that the density-variation in the MS-CG pair potentials, $U_2(T, \rho)$, is not very important for accurately modeling the configuration-dependence of the PMF. This suggests that this density-dependence may be accurately modeled by a configuration-independent volume potential. Consequently, we decided to decouple the density-dependence of the pair and volume potentials. We adopted the MS-CG interaction potential, $U_{\text{R}}(T_{\text{ref}}, \rho_{\text{ref}})$, that was calculated for the reference state point, $(T_{\text{ref}}, \rho_{\text{ref}})$, as a fixed density-independent potential for modeling all state points at the reference temperature, T_{ref} . Given this density-decoupled

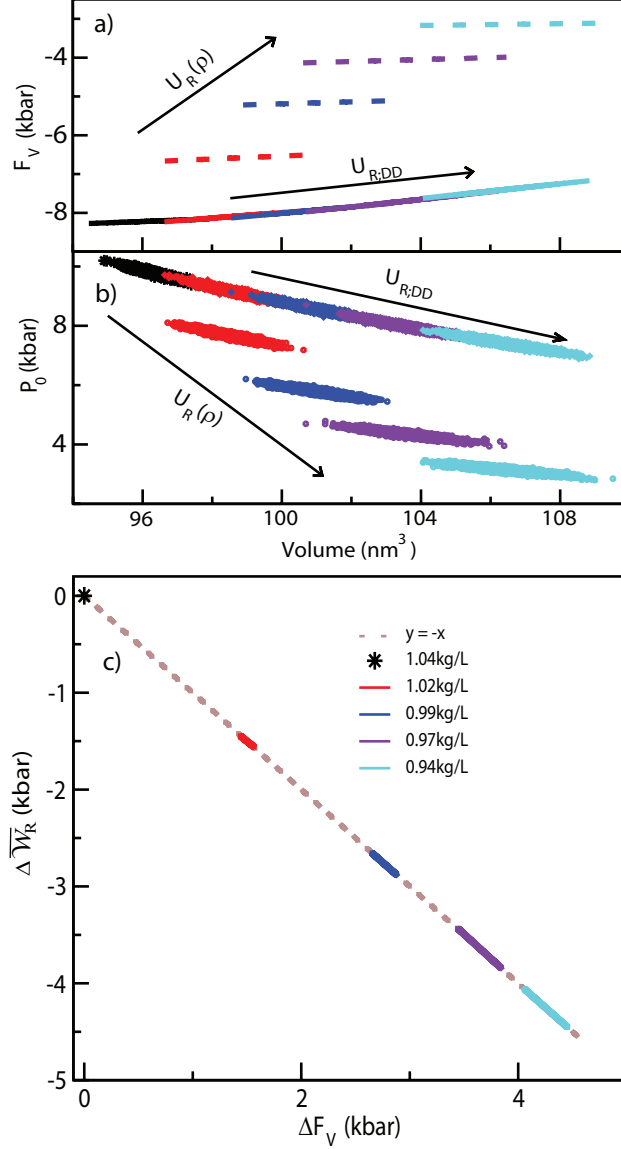


Figure 8: Analysis of calculated volume forces at $T_{\text{ref}} = 380$ K. Panel a: the dashed and solid curves present the volume forces, $F_V(V) = -dU_V(V)/dV$, that were calculated for the density-dependent interaction potentials, $\{U_R(\rho)\}$, and for the density-independent DD interaction potential, $U_{R;DD}$. Panel b: Scatter plot of the volume, V , with the naïve internal pressure, $P_0 = Nk_B T/V + \mathcal{W}_R$, that includes the ideal, $Nk_B T/V$, and configurational virial, $\mathcal{W}_R = \frac{1}{3V} \sum_{(I,J)} F_2(R_{IJ})R_{IJ}$, contributions evaluated for configurations $(\mathbf{R}, V)$ sampled by constant NPT CG simulations with the density-independent DD interaction potential, $U_{R;DD}$, and with the density-dependent MS-CG interaction potential, $U_R(\rho)$. Panel c: Correlation between the difference in the mean configurational virial, $\Delta \overline{\mathcal{W}}_R \equiv \langle \overline{\mathcal{W}}_R[U_R(\rho)] \rangle_V - \langle \overline{\mathcal{W}}_R[U_{R;DD}] \rangle_V$, calculated for $U_R(\rho)$ and for $U_{R;DD}$ with the difference in the corresponding volume forces, $\Delta F_V \equiv F_V[U_R(\rho)] - F_V[U_{R;DD}]$.

(DD) interaction potential, $U_{\text{R};\text{DD}} \equiv U_{\text{R}}(T_{\text{ref}}, \rho_{\text{ref}})$, we repeated our self-consistent pressure-matching calculations to determine a new volume potential, $U_{\text{V};\text{DD}}(\rho)$, for modeling each state point (T_{ref}, ρ) .

Figure 8 analyzes the resulting volume forces that were calculated for modeling the constant NPT ensemble at the fixed reference temperature, T_{ref} , and various external pressures, P_{ext} . We define $P_0(\mathbf{R}, V) = Nk_{\text{B}}T/V + \mathcal{W}_{\text{R}}(\mathbf{R}, V)$ as the naïve CG pressure, where $\mathcal{W}_{\text{R}}(\mathbf{R}, V) \equiv \mathcal{W}_{\text{R}}(\mathbf{R}, V|U_{\text{R}}) \equiv \frac{1}{3V} \sum_{(I,J)} F_2(R_{IJ})R_{IJ}$ is the configurational virial that is defined by the interaction potential, U_{R} .

The dashed lines in Fig. 8a reproduce from Fig. 3d the set of volume forces, $\{F_{\text{V}}(\rho) \equiv F_{\text{V}}(V; T_{\text{ref}}, \rho)\}$, that were calculated for the original density-dependent MS-CG potentials, $\{U_{\text{R}}(\rho) \equiv U_{\text{R}}(\mathbf{R}; T_{\text{ref}}, \rho)\}$. As already noted in Fig. 3d, this set of volume forces, $\{F_{\text{V}}(\rho)\}$, cannot be described by a single continuous function of V . The solid lines in Fig. 8a present the set of volume forces, $\{F_{\text{V};\text{DD}}(\rho) \equiv -dU_{\text{V};\text{DD}}(\mathbf{R}; \rho)/dV\}$, that were determined for each state point, (T_{ref}, ρ) , when employing the fixed density-independent interaction potential, $U_{\text{R};\text{DD}}$. Importantly, the set of DD volume forces, $\{F_{\text{V};\text{DD}}(\rho)\}$, can be represented by a single continuous function of V , $F_{\text{V};\text{glob}} \equiv F_{\text{V};\text{glob}}(V; T_{\text{ref}})$. Specifically, this global volume force, $F_{\text{V};\text{glob}}$, is independent of the density, ρ , of the simulated state point, (T_{ref}, ρ) .

The discontinuities in the original volume forces, $\{F_{\text{V}}(\rho)\}$, reflect the changes in the virial due to the density-dependence of the original MS-CG pair potentials, $\{U_{\text{R}}(\rho)\}$, in Fig. 3c. Figure 8b presents a scatter plot of (V, P_0) that was sampled by CG simulations with the density-dependent MS-CG interaction potential, $U_{\text{R}}(\rho)$, and with the density-independent DD interaction potential, $U_{\text{R};\text{DD}}$. If one employs two different interaction potentials, $U_{\text{R}}(\rho_1)$ and $U_{\text{R}}(\rho_2)$, to simulate the same configuration $(\mathbf{R}, V)$, then one obtains two different values for configurational virial, $\mathcal{W}_{\text{R}1}(\mathbf{R}, V)$ and $\mathcal{W}_{\text{R}2}(\mathbf{R}, V)$. Consequently, two different volume potentials, $U_{\text{V}}(\rho_1)$ and $U_{\text{V}}(\rho_2)$, are necessary to model the same configuration in simulations with the two interaction potentials, $U_{\text{R}}(\rho_1)$ and $U_{\text{R}}(\rho_2)$. Conversely, if one assumes that the interaction potential is density-independent, then constant NPT simulations at different

densities, ρ_1 and ρ_2 , obtain the same configurational virial for a given configuration $(\mathbf{R}, V)$. In this case $\overline{\mathcal{W}}_{\mathbf{R}}(V) \equiv \langle \mathcal{W}_{\mathbf{R}} \rangle_V$ varies continuously with volume and is independent of the density, ρ , of the simulated state point (T_{ref}, ρ) .

Figure 8c further supports this conclusion by presenting a scatter plot of the difference in the configurational virial for the density-dependent MS-CG interaction potential and for the DD interaction potential, $\Delta \overline{\mathcal{W}}_{\mathbf{R}} \equiv \overline{\mathcal{W}}_{\mathbf{R}}[U_{\mathbf{R}}(\rho)] - \overline{\mathcal{W}}_{\mathbf{R}}[U_{\mathbf{R};\text{DD}}]$, and the difference in the corresponding volume forces, $\Delta F_V \equiv F_V[U_{\mathbf{R}}(\rho)] - F_V[U_{\mathbf{R};\text{DD}}]$. Clearly, the variation in the configurational virial due to density-dependent potentials is exactly compensated for by the change in the volume force. Moreover, this suggests that the density-dependence of the original MS-CG pair potentials primarily stems from the density-dependence of the AA pressure and, consequently, can be quite accurately captured by a single global volume potential, $U_{\text{V;glob}}(V; T_{\text{ref}})$.

We next repeated this process for each temperature $300 \text{ K} \leq T < T_{\text{ref}}$. At each temperature, T , we defined the fixed density-independent DD interaction potential by the MS-CG interaction potential that had been optimized for the equilibrium state, (T, ρ_{ref}) , i.e., $U_{\mathbf{R};\text{DD}}(T) \equiv U_{\mathbf{R};\text{DD}}(\mathbf{R}; T) = U_{\mathbf{R}}(\mathbf{R}; T, \rho_{\text{ref}})$. We then repeated self-consistent pressure-matching to determine a corresponding volume potential, $U_{\text{V;DD}}(T, \rho) = U_{\text{V;DD}}(V; T, \rho)$, for performing constant NPT simulations at each state point, (T, ρ) , in Fig. 1a.

The solid lines in Fig. 9a present the resulting set of volume forces, $\{F_{\text{V;DD}}(T, \rho)\}$. As seen in Fig. 8a, the DD volume forces for each temperature, T , can be accurately modeled by a single continuous function of V , $F_{\text{V;glob}}(T) \equiv F_{\text{V;glob}}(V; T)$ that is independent of the density, ρ , of the simulated state point, (T, ρ) . Moreover, this global volume force, $F_{\text{V;glob}}(T)$, can be accurately modeled as a quadratic function of V for each temperature. Accordingly, we fit the entire set of DD volume forces, $\{F_{\text{V;DD}}(V; T, \rho)\}$, to a single global function of V and T :

$$F_{\text{V;glob}}(V; T) \equiv a(T)V^2 + b(T)V + c(T), \quad (28)$$

where $a(T)$, $b(T)$, and $c(T)$ are linear functions of temperature, e.g., $a(T) \equiv a_0 + a_1 (T - T_{\text{ref}})$.

The dotted orange curves in Fig. 9a present the resulting global volume force. The fit is extremely good at elevated temperatures, but slightly deteriorates at the lowest temperatures. We define a corresponding global volume potential, $U_{V;\text{glob}}(V;T)$, by simply integrating $F_{V;\text{glob}}(V;T) = -dU_{V;\text{glob}}(V;T)/dV$.

Given $U_{V;\text{glob}}(V;T)$, we define a global potential:

$$U_{\text{glob}}(\mathbf{R}, V; T) \equiv U_{R;\text{DD}}(\mathbf{R}; T) + U_{V;\text{glob}}(V; T). \quad (29)$$

Figures 9b - 9d assess the accuracy of this global potential for modeling the liquid phase of dioxane. In particular, Fig. 9b demonstrates that this global potential quite accurately reproduces the AA pressure-volume equation of state across the entire range of liquid state points. Somewhat surprisingly, U_{glob} accurately describes the AA pressure equation of state even at state points where the global fit to the optimized volume forces deteriorates. Figures 9c and 9d present rdf's from constant NPT simulations at ambient pressure for the two state points that are farthest from the reference state point in Fig. 1a. These figures demonstrate that the global potential also accurately reproduces the pair structure of the AA model at each state point. While the global model slightly over-estimates the height of the first peak of the AA rdf, it reproduces the AA rdf as accurately as the MS-CG model that was optimized for the state point.

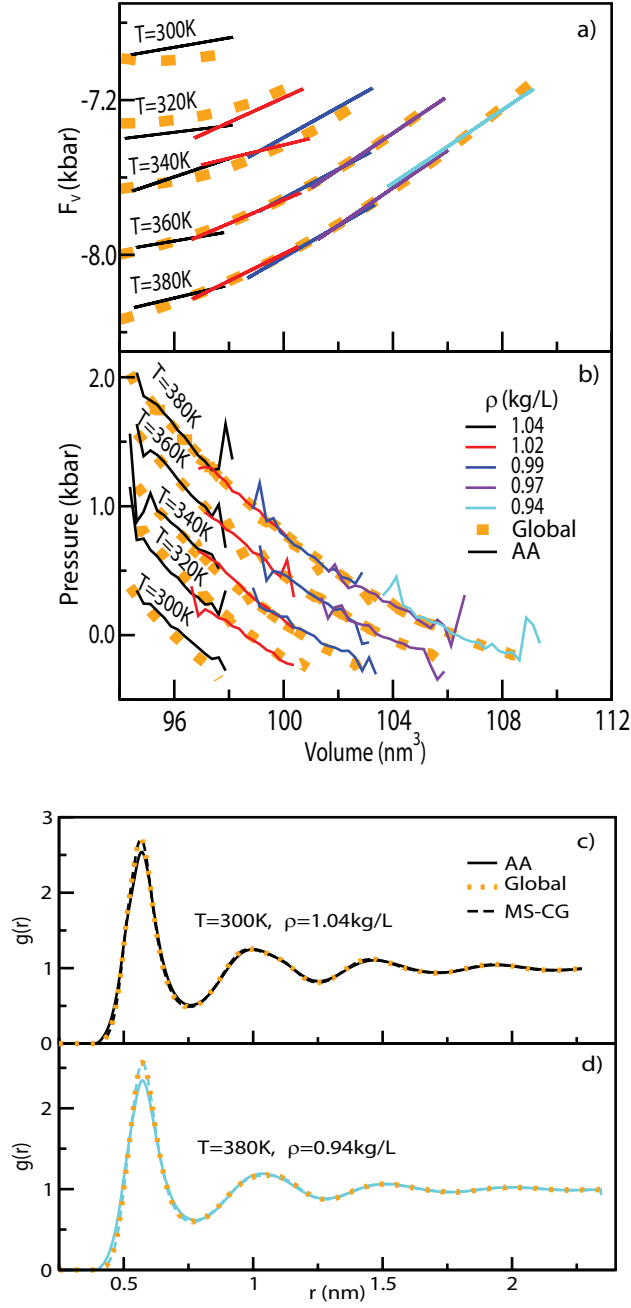


Figure 9: Global liquid dioxane model defined for $\rho_{\text{ref}} = 1.04 \text{ kg/L}$. Panel a compares the DD volume forces, $F_{V;\text{DD}}(V; T, \rho)$, that were calculated for each state point (solid lines) with the global volume force, $F_{V;\text{glob}}(V; T)$, defined by Eq. (28) (dotted orange curves). Panel b presents simulated pressure-volume equations of state for the AA model (solid curves) and for the global model defined by Eq. (29) (dotted orange curves). Panels c and d present rdf's from constant NPT simulations at the two state points in Fig. 1a that are farthest from the reference state point. In panels c and d the solid curves correspond to the AA model, the dashed curves correspond to the MS-CG models parameterized for the corresponding state point, and the dotted orange curves correspond to the global model.

4.7 Assessment for THF

In the preceding subsections we developed a global potential for modeling liquid dioxane across a rather wide range of thermodynamic state points. In this final subsection, we briefly consider the generality of this approach for a 1-site CG model of tetrahydrofuran (THF), which is a structurally similar organic liquid. Specifically, we consider the range of liquid state points $180 \text{ K} \leq T \leq 320 \text{ K}$ and $1 \text{ bar} \leq P \leq 4245 \text{ bar}$. In this subsection we present results for two global THF models that differ in their choice of reference state. The SI presents detailed results that are analogous to Figs. 2 - 9 for THF.

Figure 10 presents results for a global THF model that employs the top right state point in Fig. 1b as the reference state, i.e., $T_{\text{ref}} = 320 \text{ K}$, $P_{\text{ref}} = 4245 \text{ bar}$, $\rho_{\text{ref}} = 1.01 \text{ kg/L}$. Figure 10a demonstrates that the global volume force, $F_{\text{V;glob}}$, quite accurately matches the volume forces that were independently calculated for each state point in Fig. 1b, although the agreement slightly deteriorates at the lowest temperatures. Figure 10c and 10d demonstrate that the structural fidelity of this global model slightly deteriorates at the extremes of the state point matrix and, in particular, at the state point with highest temperature and lowest density. Nevertheless, the global model quite accurately describes the intermolecular structure of the AA model across the entire liquid phase. More impressively, Fig. 10b demonstrates that the global model very accurately describes the AA pressure-volume equation of state across the entire range of liquid state points.

Finally, Fig. 11 presents the results of employing an intermediate density, $\rho_* = 0.93 \text{ kg/L}$, as a reference for parameterizing a second global potential, U_{glob}^* . In this case, we only determine the DD interaction potential, $U_{\text{R;DD}}^*(T) \equiv U_{\text{R}}(T, \rho_*)$, for $T \geq T_* \equiv 250 \text{ K}$. State points $(T < T_*, \rho_*)$ fall below the diagonal of Fig. 1b and, consequently, correspond to negative internal pressures. We determine a new DD volume force, $F_{\text{V;DD}}^*(V; T, \rho)$, for each state point (T, ρ) and fit these to determine a new global volume force, $F_{\text{V;glob}}^*(V; T)$, of the form given by Eq. (28). Figure 11a demonstrates that the new global volume force accurately describes the DD volume forces that were independently optimized for each state

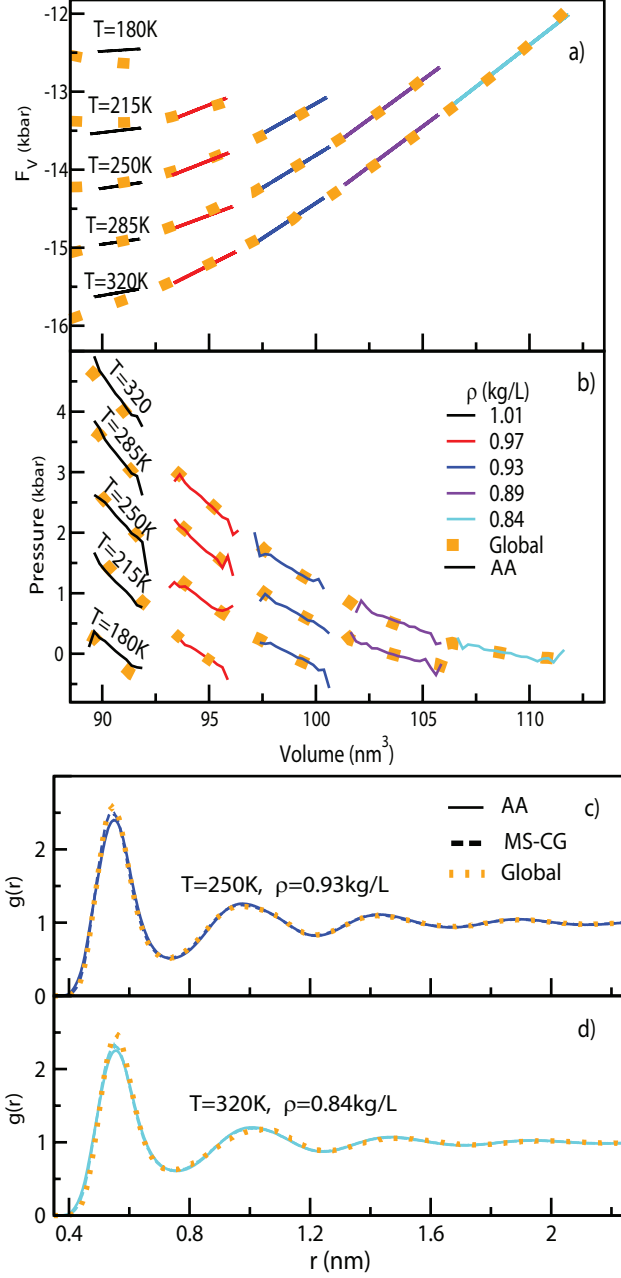


Figure 10: Global model for liquid THF with $\rho_{\text{ref}} = 1.01 \text{ kg/L}$. Panel a compares the DD volume forces, $F_{V;\text{DD}}(V; T, \rho)$, that were calculated for each state point (solid lines) with the global volume force, $F_{V;\text{glob}}(V; T)$, defined by Eq. (28) (dotted orange curves). Panel b presents the pressure-volume equations of state for the AA model (solid curves) and for the global model defined by Eq. (29) (dotted orange curves). Panels c and d present rdf's from constant NPT simulations at two representative two state points in Fig. 1b. In panels c and d the solid curves correspond to the AA model, the dashed curves correspond to the MS-CG models parameterized for the corresponding state point, and the dotted orange curves correspond to the global model.

point. Interestingly, in both Fig. 10a and 11a, the volume forces are most attractive at the state points used to define the density-independent DD interaction potential, i.e., (T, ρ_{ref}) and (T, ρ_*) , respectively. Figure 11b demonstrates this second global model also accurately reproduces the AA pressure-volume equations of state at each temperature $T \geq T_*$. Similarly, Figs. 11c and 11d demonstrate that the second global model accurately reproduces the intermolecular structure of the AA model at each state point. In this case, the structural fidelity of the global model appears slightly improved with respect to Fig. 10 because we consider smaller variations from the central reference state point.

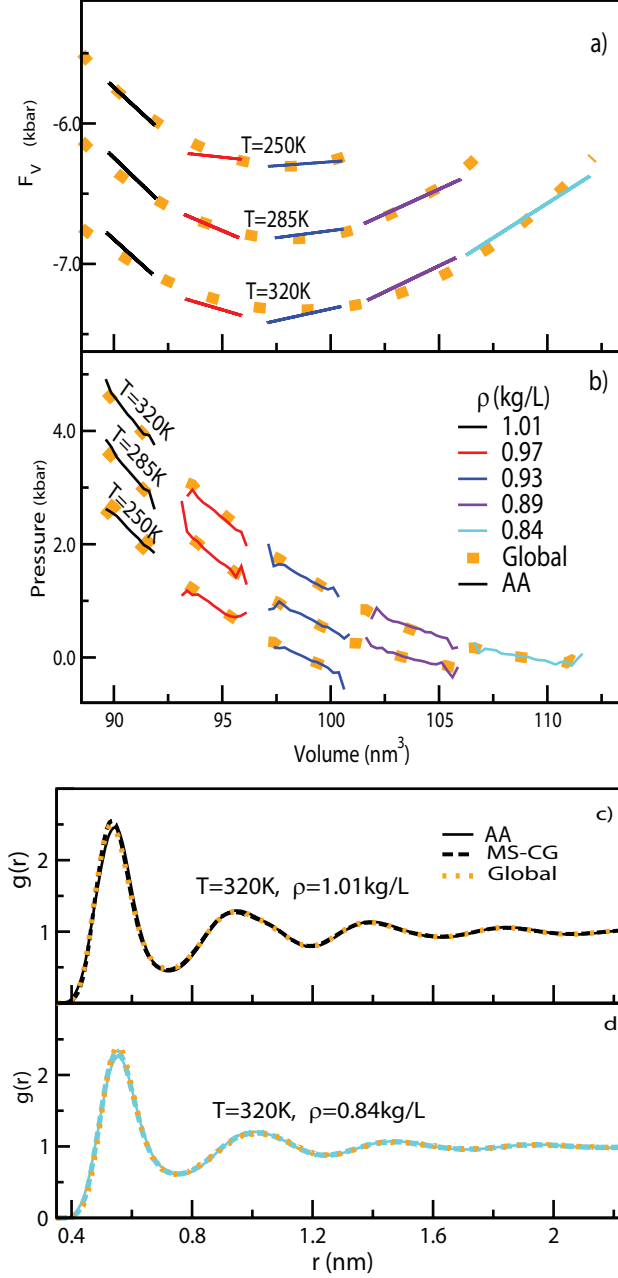


Figure 11: Global model for liquid THF defined for the intermediate density $\rho_* = 0.93 \text{ kg/L}$. Panel a compares the DD volume forces, $F_{V;\text{DD}}^*(V; T, \rho)$, that were calculated for each state point (solid lines) with the global volume force, $F_{V;\text{glob}}^*(V; T)$, defined by Eq. (28) (dotted orange curves). Panel b presents simulated pressure-volume equations of state for the AA model (solid curves) and for the global model defined by Eq. (29) (dotted orange curves). Panels c and d present rdf's from constant NPT simulations at two representative state points in Fig. 1b. In panels c and d the solid curves correspond to the AA model, the dashed curves correspond to the MS-CG models parameterized for the corresponding state point, and the dotted orange curves correspond to the global model.

5 Conclusions

In this study, we have investigated the temperature- and density-dependence of effective potentials for 1-site MS-CG models that accurately describe the intermolecular structure and internal pressure of AA models for dioxane and THF. The calculated pair potentials are highly repulsive with very shallow minima that correspond to the second solvation shell. The pair potentials become slightly more repulsive as the temperature increases at constant density, while the volume potentials simultaneously become somewhat more attractive. As previously reported, we observe that the pair and volume potentials depend much more sensitively upon density than temperature. Specifically, as the density decreases at constant temperature, the short-ranged component of the pair potential becomes much more attractive, while the corresponding volume forces become much less attractive. At the lowest densities, a second contact minimum begins to compete with the global minimum in the pair potential.

We employed the g-YBG framework to elucidate the physical origin of this striking density-dependence.^{95–98} The g-YBG framework decomposes the mean force between each pair of molecules into a direct contribution, which corresponds to the MS-CG pair potential, and an indirect contribution, which results from correlated interactions with the environment. While the direct contribution between neighboring molecules is highly repulsive, the indirect contribution to this pair mean force is highly attractive due to repulsive interactions with the surrounding molecules.^{97,135} Our calculations demonstrate that the pair mean force between contacting molecules is almost density-independent for dioxane and THF. However, the attractive indirect contribution to this pair mean force significantly decreases as the density of the surrounding environment decreases. Consequently, the MS-CG pair potential between contacting molecules must become more attractive to compensate.

This analysis relied upon special properties of the MS-CG method^{102,107–109} and, in particular, its relationship with the g-YBG framework.^{95,98,135} In contrast to other structure-based methods, such as Iterative Boltzmann Inversion (IBI)⁶⁸ or Inverse Monte Carlo (IMC),¹⁰⁴

the MS-CG method determines interaction potentials directly from the AA mapped ensemble without requiring successive simulations of CG models with trial potentials. Moreover, the g-YBG framework interprets the resulting MS-CG potentials in terms of direct and indirect contributions to mean forces that can be determined from the mapped ensemble.^{97,98} This provides a simple physical basis for understanding the physical origin of MS-CG potentials, as well as their sensitivity to temperature- and density-variations. In contrast, the temperature- and density-dependence of IBI or IMC potentials cannot be so directly traced to properties of the AA model because they also reflect the temperature- and density-dependence of the simulated CG ensemble. In this case, one must consider two sets of g-YBG equations: one for the mapped AA ensemble and one for the simulated CG ensemble.^{8,135,143} Nevertheless, we anticipate that the findings of our g-YBG analysis will likely qualitatively hold also for potentials obtained via other structure-based bottom-up approaches. Specifically, we speculate that bottom-up potentials for small molecular liquids will often demonstrate relatively weak temperature-dependence, which will be primarily determined by the temperature-dependence of the pair mean force or, equivalently, the mapped rdf. Conversely, we speculate that these potentials will demonstrate greater density-dependence and that this density-dependence will be more sensitive to density-dependent many-body correlations. It may be instructive for future studies to assess this hypothesis.

We had expected that the striking density-dependence of the MS-CG pair potentials would be important for accurately modeling the liquid structure. Surprisingly, though, we find that it has relatively little impact upon the structural fidelity of the MS-CG models. Consequently, we decoupled the density-dependence of the pair and volume potentials. Specifically, for each temperature, T , we defined a fixed density-independent interaction potential, $U_{\text{R;DD}}(\mathbf{R}; T)$, and then determined an independent volume potential, $U_{\text{V;DD}}(V; T, \rho)$, for modeling each state point, (T, ρ) . We were able to fit this set of 15 independent volume potentials to a single global volume potential, $U_{\text{V;glob}}(V; T)$, that depends cubically upon V and linearly upon T . The resulting global potential, $U_{\text{glob}}(\mathbf{R}, V; T) = U_{\text{R;DD}}(\mathbf{R}; T) + U_{\text{V;glob}}(V; T)$,

described the AA intermolecular structure and pressure-volume equation of state with remarkable accuracy across the range of liquid phase state points for both systems. These global potentials, U_{glob} , appear very promising for developing highly efficient CG models that accurately model the structure and thermodynamic properties of homogeneous industrial solvents across their entire liquid phase.

In the present work, we employed AA simulations at 15 different state points to determine this global potential. Specifically, we fixed a reference density, ρ_{ref} , and then employed AA simulations at a range of temperatures, T , in order to determine a density-independent interaction potential, $U_{\text{R;DD}}(T)$, for each T . In the future, it may be possible to employ the dual⁴³ or microcanonical⁴⁶ framework to predict this temperature-dependence based upon simulating a single state point. Additionally, we calculated a volume potential for all 15 state points in order to determine the global volume potential, $U_{\text{V;glob}}(V;T)$. This global volume potential varied quadratically with V and linearly with T , while depending upon 6 parameters. Given the accuracy of the fit, it is likely that these six parameters could be accurately determined from a small fraction of these 15 state points. We have also previously demonstrated that it is sometimes possible to predict the temperature-dependence of the volume potential via the dual approach.^{43,44} Moreover, it may be possible to determine the cubic volume dependence of $U_{\text{V;glob}}$ from simulations at a single state point by including an additional cubic basis function into Eq. (17) when performing self-consistent pressure-matching. However, accurately resolving the curvature of the volume force from a single state point may be challenging in practice.

We anticipate that it may often be possible to extend the present approach to mixtures of miscible solvents. We have previously combined⁸⁹ self-consistent pressure-matching⁸⁸ with the extended ensemble framework⁶⁶ in order to develop transferable pair and volume potentials for modeling heptane-toluene mixtures at ambient pressure and a single fixed temperature. In this case, we found that the volume potentials could be easily predicted as a function of solution composition based upon regular solution considerations. Rather

remarkably, the resulting transferable potentials described these solutions more accurately than MS-CG models that had been specifically parameterized for each composition.⁸⁹ Quite recently, Kanekal et al. demonstrated that this extended ensemble framework can improve not only the transferability, but also the accuracy of CG models by averaging over system-specific details that cannot be accurately described by MS-CG potentials.⁶⁷ Conversely, Shen et al. demonstrated that the extended ensemble approach does not always succeed.¹⁴⁴

In order to extend our prior work with heptane-toluene mixtures,⁸⁹ we anticipate that it would be necessary to first perform AA simulations at a range of temperatures and external pressures for a representative set of solution compositions. We would then employ the extended ensemble force-matching variational principle⁶⁶ to determine a set of composition-independent, density-independent (but possibly temperature-dependent) pair potentials that optimally reproduced the configuration-dependence of the many-body PMF for each representative composition at each simulated state point. Given these extended ensemble interaction potentials, we would employ self-consistent pressure-matching to determine a six parameter global volume force of the form given by Eq. (24) for each solution composition. We anticipate that it may often be possible to model the global volume forces for miscible liquids as a simple function of composition, as in our work with heptane-toluene mixtures at a single temperature and pressure.⁸⁹ However, we anticipate this approach will likely fail to describe liquid-liquid phase separation.

It may be more challenging to extend this approach to model inhomogeneous and interfacial systems. Volume potentials provide an extremely efficient method for accurately reproducing pressure-volume equations of state for homogeneous fluids. However, because they apply a force directly to the system volume, volume potentials cannot be employed to simulate inhomogeneous systems, such as liquid-vapor interfaces. In these cases, local density potentials can be used to accurately model both the internal pressure and also the interfacial profile, while providing outstanding transferability between bulk and interfacial environments. In some cases it is possible to directly map volume potentials to local density

potentials.⁷⁸ This may require defining the local density over a relatively long length-scale, though, which can generate significant artifacts when modeling liquid interfaces.^{81,84}

Similarly, in this work we only considered equilibrium properties of small molecular solvents with short-ranged interactions. Future studies should certainly investigate this approach for modeling more complex systems, such as ionic and polymeric liquids. Furthermore, future studies should extend this work to consider dynamical quantities^{145–148} and non-equilibrium phenomena.^{149–151} Nevertheless, we hope that this study provides fundamental insight into the density-dependence of effective pair potentials. Moreover, we hope that this approach may represent a significant practical advance towards developing predictive CG models for modeling industrial solvents across wide ranges of their phase diagram.

Supporting Information Available

The SI further investigates the predicted interaction potentials, U_{lin} , and employs the g-YBG formalism to analyze the temperature- and density-dependence of MS-CG potentials. The SI also presents rdf's that explicitly assess the structural fidelity of the MS-CG and global potentials for dioxane and THF at each state point.

Acknowledgement

WGN gratefully acknowledges Prof. G. A. Voth for the opportunity to study systematic coarse-graining in the stimulating environment of his group, as well as for the inspiration of his many pioneering contributions to the field. The authors also gratefully acknowledge financial support from the National Science Foundation (Grant Nos. CHE-1856337 and CHE-2154433). Portions of this research were conducted with Advanced CyberInfrastructure computational resources provided by The Institute for Computational and Data Sciences at The Pennsylvania State University (<http://icds.psu.edu>). Additionally, parts of this research used the Expanse resource at the San Diego Supercomputer Center through allocation TG-

CHE170062 from the Extreme Science and Engineering Discovery Environment (XSEDE)¹⁵², which was supported by National Science Foundation grant number TG-CHE170062. This work also used Rockfish at Johns Hopkins University through allocation CHE170062 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296. Figure 1 employed VMD.¹⁵³ VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of 2510 Illinois at Urbana-Champaign.

References

- (1) Klein, M. L.; Shinoda, W. Large-scale molecular dynamics simulations of self-assembling systems. Science **2008**, 321, 798–800.
- (2) Saunders, M. G.; Voth, G. A. Coarse-graining methods for computational biology. Annu. Rev. Biophys. **2013**, 42, 73–93.
- (3) Li, Y.; Abberton, B. C.; Kröger, M.; Liu, W. K. Challenges in Multiscale Modeling of Polymer Dynamics. Polymers **2013**, 5, 751–832.
- (4) Brini, E.; Algaer, E. A.; Ganguly, P.; Li, C.; Rodríguez-Ropero, F.; van der Vegt, N. F. A. Systematic coarse-graining methods for soft matter simulations - a review. Soft Matter **2013**, 9, 2108–2119.
- (5) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. Wiley Interdiscip. Rev. Comput. Mol. Sci. **2014**, 4, 225–248.
- (6) Gartner, T. E.; Jayaraman, A. Modeling and Simulations of Polymers: A Roadmap. Macromolecules **2019**, 52, 755–786.
- (7) Schmid, F. Understanding and Modeling Polymers: The Challenge of Multiple Scales. ACS Polymers Au **2023**, 3, 28–58.
- (8) 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, 1869–92.
- (9) Peter, C.; Kremer, K. Multiscale simulation of soft matter systems. Faraday Discuss. **2010**, 144, 9–24.

- (10) 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, 12423–12430.
- (11) Guenza, M.; Dinpajooh, M.; McCarty, J.; Lyubimov, I. Accuracy, transferability, and efficiency of coarse-grained models of molecular liquids. J. Phys. Chem. B **2018**, 122, 10257–10278.
- (12) Muller, M.; Katsov, K.; Schick, M. Biological and synthetic membranes: What can be learned from a coarse-grained description? Phys. Rep. **2006**, 434, 113–176.
- (13) Noid, W. G. Perspective: coarse-grained models for biomolecular systems. J. Chem. Phys. **2013**, 139, 090901.
- (14) Giuliani, M.; Rigoli, M.; Mattiotti, G.; Menichetti, R.; Tarenzi, T.; Fiorentini, R.; Potestio, R. From System Modeling to System Analysis: The Impact of Resolution Level and Resolution Distribution in the Computer-Aided Investigation of Biomolecules. Front. Mol. Biosci. **2021**, 8, 676976.
- (15) Rowlinson, J. Intermolecular potentials that are functions of thermodynamic variables. Mol. Phys. **1984**, 52, 567–572.
- (16) van der Hoef, M. A.; Madden, P. A. Three-body dispersion contributions to the thermodynamic properties and effective pair interactions in liquid argon. J. Chem. Phys. **1999**, 111, 1520–1526.
- (17) Louis, A. A.; Bolhuis, P. G.; Hansen, J. P.; Meijer, E. J. Can polymer coils be modeled as “soft colloids”? Phys. Rev. Lett. **2000**, 85, 2522–2525.
- (18) D’Adamo, G.; Pelissetto, A.; Pierleoni, C. Predicting the thermodynamics by using state-dependent interactions. J. Chem. Phys. **2013**, 138, 234107.

- (19) McCarty, J.; Clark, A. J.; Copperman, J.; Guenza, M. G. An analytical coarse-graining method which preserves the free energy, structural correlations, and thermodynamic state of polymer melts from the atomistic to the mesoscale. J. Chem. Phys. **2014**, 140, 204913.
- (20) Dunn, N. J. H.; Foley, T. T.; Noid, W. G. van der Waals perspective on coarse-graining: progress toward solving representability and transferability problems. Acc. Chem. Res. **2016**, 49, 2832–2840.
- (21) 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, 044108.
- (22) Foley, T. T.; Shell, M. S.; Noid, W. G. The impact of resolution upon entropy and information in coarse-grained models. J. Chem. Phys. **2015**, 143, 243104.
- (23) Kidder, K. M.; Szukalo, R. J.; Noid, W. G. Energetic and entropic considerations for coarse-graining. Eur. Phys. J. B **2021**, 94, 153.
- (24) Jin, J.; Pak, A. J.; Durumeric, A. E. P.; Loose, T. D.; Voth, G. A. Bottom-up Coarse-Graining: Principles and Perspectives. Journal of Chemical Theory and Computation **2022**, 18, 5759–5791.
- (25) Noid, W. G. Perspective: Advances, Challenges, and Insight for Predictive Coarse-Grained Models. J. Phys. Chem. B **2023**, 127, 4174–4207.
- (26) Vettorel, T.; Meyer, H. Coarse graining of short polyethylene chains for studying polymer crystallization. J. Chem. Theory Comput. **2006**, 2, 616–629.
- (27) Ghosh, J.; Faller, R. State point dependence of systematically coarse-grained potentials. Mol. Simu. **2007**, 33, 759–767.

- (28) Szukalo, R. J.; Noid, W. Investigation of coarse-grained models across a glass transition. Soft Mater. **2020**, 18, 1–15.
- (29) Wilson, M. R.; Yu, G. Computer Simulations of a Twist Bend Nematic (NTB): A Coarse-Grained Simulation of the Phase Behaviour of the Liquid Crystal Dimer CB7CB. Crystals **2023**, 13.
- (30) 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, 024103.
- (31) Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Ołdziej, S.; Wachucik, K.; Scheraga, H. A. Modification and optimization of the united-residue (UNRES) potential energy function for canonical simulations. I. Temperature dependence of the effective energy function and tests of the optimization method with single training proteins. J. Phys. Chem. B **2007**, 111, 260–85.
- (32) Qian, H.-J.; Carbone, P.; Chen, X.; Karimi-Varzaneh, H. A.; Liew, C. C.; Müller-Plathe, F. Temperature-Transferable Coarse-grained potentials for Ethylbenzene, Polystyrene and their mixtures. Macromolecules **2008**, 41, 9919–29.
- (33) Chaimovich, A.; Shell, M. S. Anomalous waterlike behavior in spherically-symmetric water models optimized with the relative entropy. Phys. Chem. Chem. Phys. **2009**, 11, 1901–1915.
- (34) Farah, K.; Fogarty, A. C.; Böhm, M. C.; Müller-Plathe, F. Temperature dependence of coarse-grained potentials for liquid hexane. Phys. Chem. Chem. Phys. **2011**, 13, 2894–902.
- (35) Lu, L.; Voth, G. A. The multiscale coarse-graining method. VII. Free energy decomposition of coarse-grained effective potentials. J. Chem. Phys. **2011**, 134, 224107.

- (36) Izvekov, S. Towards an understanding of many-particle effects in hydrophobic association in methane solutions. J. Chem. Phys. **2011**, 134, 034104.
- (37) Mirzoev, A.; Lyubartsev, A. P. Effective solvent mediated potentials of Na⁺ and Cl⁻ ions in aqueous solution: temperature dependence. Phys. Chem. Chem. Phys. **2011**, 13, 5722–5727.
- (38) Cao, Z.; Dama, J. F.; Lu, L.; Voth, G. A. Solvent Free Ionic Solution Models from Multiscale Coarse-Graining. J. Chem. Theory Comput. **2013**, 9, 172–178.
- (39) Solano Canchaya, J. G.; Dequidt, A.; Goujon, F.; Malfreyt, P. Development of DPD coarse-grained models: From bulk to interfacial properties. J. Chem. Phys. **2016**, 145, 054107.
- (40) Rudzinski, J. F.; Lu, K.; Milner, S. T.; Maranas, J. K.; Noid, W. G. Extended ensemble approach to transferable potentials for low-resolution coarse-grained models of ionomers. J. Chem. Theory Comput. **2017**, 13, 2185–2201.
- (41) Lebold, K. M.; Noid, W. G. Systematic study of temperature and density variations in effective potentials for coarse-grained models of molecular liquids. J. Chem. Phys. **2019**, 150, 014104.
- (42) 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, 4549–4557.
- (43) Lebold, K. M.; Noid, W. G. Dual approach for effective potentials that accurately model structure and energetics. J. Chem. Phys. **2019**, 150, 234107.
- (44) Lebold, K. M.; Noid, W. G. Dual-potential approach for coarse-grained implicit solvent models with accurate, internally consistent energetics and predictive transferability. J. Chem. Phys. **2019**, 151, 164113.

- (45) Szukalo, R. J.; Noid, W. G. Investigating the energetic and entropic components of effective potentials across a glass transition. J. Phys.: Condens. Matter **2021**, 33, 154004.
- (46) Pretti, E.; Shell, M. S. A microcanonical approach to temperature-transferable coarse-grained models using the relative entropy. J. Chem. Phys. **2021**, 155, 094102.
- (47) Lyubartsev, A. P.; Laaksonen, A. Osmotic and activity coefficients from effective potentials for hydrated ions. Phys. Rev. E **1997**, 55, 5689–5696.
- (48) Allen, E. C.; Rutledge, G. C. A novel algorithm for creating coarse-grained, density dependent implicit solvent models. J. Chem. Phys. **2008**, 128, 154115.
- (49) Savelyev, A.; Papoian, G. A. Molecular renormalization group coarse-graining of electrolyte solutions: Applications to aqueous NaCl and KCl. J. Phys. Chem. B **2009**, 113, 7785–93.
- (50) Shen, J.-W.; Li, C.; van der Vegt, N. F.; Peter, C. Transferability of Coarse Grained Potentials: Implicit Solvent Models for Hydrated Ions. J. Chem. Theory Comput. **2011**, 7, 1916–1927.
- (51) Johnson, M. E.; Head-Gordon, T.; Louis, A. A. Representability problems for coarse-grained water potentials. J. Chem. Phys. **2007**, 126, 144509.
- (52) Villa, A.; Peter, C.; van der Vegt, N. F. A. Transferability of Nonbonded Interaction Potentials for Coarse-Grained Simulations: Benzene in Water. J. Chem. Theory Comput. **2010**, 6, 2434–2444.
- (53) Li, C.; Shen, J.; Peter, C.; van der Vegt, N. F. A. A Chemically Accurate Implicit-Solvent Coarse-Grained Model for Polystyrenesulfonate Solutions. Macromolecules **2012**, 45, 2551–2561.

- (54) Ganguly, P.; Mukherji, D.; Junghans, C.; van der Vegt, N. F. A. Kirkwood–Buff Coarse-Grained Force Fields for Aqueous Solutions. J. Chem. Theory Comput. **2012**, 8, 1802–1807.
- (55) Ganguly, P.; van der Vegt, N. F. A. Representability and Transferability of Kirkwood–Buff Iterative Boltzmann Inversion Models for Multicomponent Aqueous Systems. J. Chem. Theory Comput. **2013**, 9, 5247–5256.
- (56) Tsourtis, A.; Harmandaris, V.; Tsagkarogiannis, D. Parameterization of coarse-grained molecular interactions through potential of mean force calculations and cluster expansion techniques. Entropy **2017**, 19, 395.
- (57) Dinpajoo, M.; Guenza, M. G. Coarse-graining simulation approaches for polymer melts: the effect of potential range on computational efficiency. Soft Matter **2018**, 14, 7126–7144.
- (58) Izvekov, S.; Chung, P. W.; Rice, B. M. Particle-based multiscale coarse graining with density-dependent potentials: Application to molecular crystals (hexahydro-1,3,5-trinitro-s-triazine). J. Chem. Phys. **2011**, 135, 044112.
- (59) 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, 2180–2197.
- (60) Dijkstra, M.; van Roij, R.; Evans, R. Direct Simulation of the Phase Behavior of Binary Hard-Sphere Mixtures: Test of the Depletion Potential Description. Phys. Rev. Lett. **1999**, 82, 117–120.
- (61) Silbermann, J.; Klapp, S. H. L.; Shoen, M.; Channamsetty, N.; Block, H.; Gubbins, K. E. Mesoscale modeling of complex binary fluid mixtures: Towards an atomistic foundation of effective potentials. J. Chem. Phys. **2006**, 124, 074105.

- (62) Fischer, J.; Paschek, D.; Gieger, A.; Sadowski, G. Modeling of aqueous poly(oxyethylene) solutions. 2. Mesoscale simulations. J. Phys. Chem. B **2008**, 112, 13561–71.
- (63) Montes-Saralegui, M.; Kahl, G.; Nikoubashman, A. On the applicability of density dependent effective interactions in cluster-forming systems. J. Chem. Phys. **2017**, 146, 054904.
- (64) Kempfer, K.; Devémy, J.; Dequidt, A.; Couty, M.; Malfreyt, P. Realistic Coarse-Grain Model of *cis* -1,4-Polybutadiene: From Chemistry to Rheology. Macromolecules **2019**, 52, 2736–2747.
- (65) Hess, B.; Holm, C.; van der Vegt, N. Modeling Multibody Effects in Ionic Solutions with a Concentration Dependent Dielectric Permittivity. Phys. Rev. Lett. **2006**, 96, 147801.
- (66) Mullinax, J. W.; Noid, W. G. Extended ensemble approach for deriving transferable coarse-grained potentials. J. Chem. Phys. **2009**, 131, 104110.
- (67) Kanekal, K. H.; Rudzinski, J. F.; Bereau, T. Broad chemical transferability in structure-based coarse-graining. The Journal of Chemical Physics **2022**, 157, 104102.
- (68) Müller-Plathe, F. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. ChemPhysChem **2002**, 3, 754 – 769.
- (69) Wang, H.; Junghans, C.; Kremer, K. Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining? Eur. Phys. J. E: Soft Matter Biol. Phys. **2009**, 28, 221–229.
- (70) Stillinger, F. H.; Sakai, H.; Torquato, S. Statistical mechanical models with effective potentials: Definitions, applications, and thermodynamic consequences. J. Chem. Phys. **2002**, 117, 288–296.

- (71) McCarty, J.; Clark, A. J.; Lyubimov, I. Y.; Guenza, M. G. Thermodynamic Consistency between Analytic Integral Equation Theory and Coarse-Grained Molecular Dynamics Simulations of Homopolymer Melts. Macromolecules **2012**, 45, 8482–8493.
- (72) Guenza, M. Thermodynamic consistency and other challenges in coarse-graining models. Eur. Phys. J.: Spec. Top. **2015**, 224, 2177–2191.
- (73) Dinpajoo, M.; Guenza, M. G. On the Density Dependence of the Integral Equation Coarse-Graining Effective Potential. J. Phys. Chem. B **2018**, 122, 3426–3440.
- (74) Pagonabarraga, I.; Frenkel, D. Dissipative particle dynamics for interacting systems. J. Chem. Phys. **2001**, 115, 5015–5026.
- (75) Papoian, G. A.; Ulander, J.; Eastwood, M. P.; Luthey-Schulten, Z.; Wolynes, P. G. Water in protein structure prediction. Proc. Natl. Acad. Sci. U.S.A. **2004**, 101, 3352–3357.
- (76) Sanyal, T.; Shell, M. S. Coarse-grained models using local-density potentials optimized with the relative entropy: Application to implicit solvation. J. Chem. Phys. **2016**, 145, 034109.
- (77) 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, 044113.
- (78) DeLyser, M. R.; Noid, W. G. Extending pressure-matching to inhomogeneous systems via local-density potentials. J. Chem. Phys. **2017**, 147, 134111.
- (79) 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.

- (80) Moore, J. D.; Barnes, B. C.; Izvekov, S.; Lísal, M.; Sellers, M. S.; Taylor, D. E.; Brennan, J. K. A coarse-grain force field for RDX: Density dependent and energy conserving. J. Chem. Phys. **2016**, 144, 104501.
- (81) DeLyser, M.; Noid, W. Bottom-up coarse-grained models for external fields and interfaces. J. Chem. Phys. **2020**, 153, 224103.
- (82) Sanyal, T.; Shell, M. S. 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.
- (83) Rosenberger, D.; Sanyal, T.; Shell, M. S.; van der Vegt, N. F. A. Transferability of Local Density-Assisted Implicit Solvation Models for Homogeneous Fluid Mixtures. J. Chem. Theory Comput. **2019**, 15, 2881–2895.
- (84) DeLyser, M. R.; Noid, W. Analysis of local density potentials. J. Chem. Phys. **2019**, 151, 224106.
- (85) Shahidi, N.; Chazirakis, A.; Harmandaris, V.; Doxastakis, M. Coarse-graining of polyisoprene melts using inverse Monte Carlo and local density potentials. J. Chem. Phys. **2020**, 152, 124902.
- (86) Szukalo, R. J.; Noid, W. A temperature-dependent length-scale for transferable local density potentials. J. Chem. Phys. **2023**, 159.
- (87) Das, A.; Andersen, H. C. The multiscale coarse-graining method. V. Isothermal-isobaric ensemble. J. Chem. Phys. **2010**, 132, 164106.
- (88) Dunn, N. J. H.; Noid, W. G. Bottom-up coarse-grained models that accurately describe the structure, pressure, and compressibility of molecular liquids. J. Chem. Phys. **2015**, 143, 243148.

- (89) Dunn, N. J. H.; Noid, W. G. 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.
- (90) Rosenberger, D.; van der Vegt, N. F. A. Addressing the temperature transferability of structure based coarse graining models. Phys. Chem. Chem. Phys. **2018**, 20, 6617–6628.
- (91) Stohr, H.; Vaidya, R.; Wilson, C.; Pruden, A.; Salazar-Benites, G.; Bott, C. Cometabolic Treatment of 1, 4-Dioxane in Biologically Active Carbon Filtration with Tetrahydrofuran and Propane at Relevant Concentrations for Potable Reuse. ACS ES&T Water **2023**, 3, 2948–2954.
- (92) Zenker, M. J.; Borden, R. C.; Barlaz, M. A. Occurrence and Treatment of 1,4-Dioxane in Aqueous Environments. Environmental Engineering Science **2003**, 20, 423–432.
- (93) Long, J.; Zhang, Q.; Wang, T.; Zhang, X.; Xu, Y.; Ma, L. An efficient and economical process for lignin depolymerization in biomass-derived solvent tetrahydrofuran. Bioresource Technology **2014**, 154, 10–17.
- (94) Smith, M. D.; Mostofian, B.; Petridis, L.; Cheng, X.; Smith, J. C. Molecular driving forces behind the tetrahydrofuran–Water miscibility gap. The Journal of Physical Chemistry B **2016**, 120, 740–747.
- (95) Mullinax, J. W.; Noid, W. G. A generalized Yvon-Born-Green theory for molecular systems. Phys. Rev. Lett. **2009**, 103, 198104.
- (96) Mullinax, J. W.; Noid, W. G. A generalized Yvon-Born-Green theory for determining coarse-grained interaction potentials. J. Phys. Chem. C **2010**, 114, 5661–5674.
- (97) Rudzinski, J. F.; Noid, W. G. The role of many-body correlations in determining

- potentials for coarse-grained models of equilibrium structure. J. Phys. Chem. B **2012**, 116, 8621–8635.
- (98) Rudzinski, J. F.; Noid, W. G. A generalized-Yvon-Born-Green method for coarse-grained modeling. Eur. Phys. J.: Spec. Top. **2015**, 224, 2193–2216.
- (99) Tuckerman, M. E. Statistical Mechanics: Theory and Molecular Simulation; Oxford University Press: Oxford, Great Britain, 2013.
- (100) Thompson, A. P.; Plimpton, S. J.; Mattson, W. General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions. J. Chem. Phys. **2009**, 131, 154107.
- (101) Ciccotti, G.; Kapral, R.; Vanden-Eijnden, E. Blue Moon Sampling, Vectorial Reaction Coordinates, and Unbiased Constrained Dynamics. ChemPhysChem **2005**, 6, 1809–14.
- (102) 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, 244114.
- (103) Krämer, A.; Durumeric, A. E. P.; Charron, N. E.; Chen, Y.; Clementi, C.; Noé, F. Statistically Optimal Force Aggregation for Coarse-Graining Molecular Dynamics. The Journal of Physical Chemistry Letters **2023**, 14, 3970–3979.
- (104) Lyubartsev, A. P.; 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. Topics **1995**, 52, 3730–3737.
- (105) Shell, M. S. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. J. Chem. Phys. **2008**, 129, 144108.
- (106) Dequidt, A.; Solano Canchaya, J. G. Bayesian parametrization of coarse-grain dissipative dynamics models. J. Chem. Phys. **2015**, 143, 084122.

- (107) Izvekov, S.; Voth, G. A. A multiscale coarse-graining method for biomolecular systems. J. Phys. Chem. B **2005**, 109, 2469 – 2473.
- (108) Izvekov, S.; Voth, G. A. Multiscale coarse graining of liquid-state systems. J. Chem. Phys. **2005**, 123, 134105.
- (109) Noid, W. G.; Liu, P.; Wang, Y. T.; Chu, J.-W.; Ayton, G. S.; Izvekov, S.; Andersen, H. C.; Voth, G. A. The multiscale coarse-graining method. II. Numerical implementation for molecular coarse-grained models. J. Chem. Phys. **2008**, 128, 244115.
- (110) Noid, W. G. Systematic methods for structurally consistent coarse-grained models. Methods Mol Biol **2013**, 924, 487–531.
- (111) Dunn, N. J. H.; Lebold, K. M.; DeLyser, M. R.; Rudzinski, J. F.; Noid, W. G. BOCS: Bottom-up Open-source Coarse-graining Software. J. Phys. Chem. B **2018**, 122, 3363–3377.
- (112) Kullback, S.; Leibler, R. A. On Information and Sufficiency. Ann. Math. Stat. **1951**, 22, 79–86.
- (113) Cover, T. M.; Thomas, J. A. Elements of Information Theory, 2nd ed.; Wiley Interscience, 2006.
- (114) Chaimovich, A.; Shell, M. S. Coarse-graining errors and numerical optimization using a relative entropy framework. J. Chem. Phys. **2011**, 134, 094112.
- (115) Rudzinski, J. F.; Noid, W. G. Coarse-graining entropy, forces, and structures. J. Chem. Phys. **2011**, 135, 214101.
- (116) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. J. Comp. Chem. **2005**, 26, 1701–18.

- (117) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX **2015**, 1-2, 19 – 25.
- (118) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. J. Chem. Phys. **2007**, 126, 014101.
- (119) Parrinello, M.; Rahman, A. Strain fluctuations and elastic constants. J. Chem. Phys. **1982**, 76, 2662–6.
- (120) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS All-Atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. **1996**, 118, 11225–11236.
- (121) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. Proceedings of the National Academy of Sciences **2005**, 102, 6665–6670.
- (122) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. J. Phys. Chem. B **2017**, 121, 3864–3870.
- (123) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. Nucleic Acids Research **2017**, 45, W331–W336.
- (124) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. J. Chem. Phys. **1993**, 98, 10089–10092.
- (125) Lide, D. R., Ed. CRC Handbook of Chemistry and Physics, 75th Edition, 75th ed.; CRC Press: Ann Arbor, MI USA, 2009.

- (126) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. Numerical Recipes in FORTRAN: The Art of Scientific Computing; Cambridge University Press: New York, NY USA, 1992.
- (127) Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Croz, J. D.; Greenbaum, A.; Hammarling, S.; McKenney, A. et al. LAPACK Users' Guide; SIAM: Philadelphia, PA USA, 1999.
- (128) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods **2020**, 17, 261–272.
- (129) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. J. Comp. Phys. **1995**, 117, 1 – 19.
- (130) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Computer Physics Communications **2022**, 271, 108171.
- (131) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. J. Chem. Phys. **1992**, 97, 2635–2643.
- (132) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. Mol. Phys. **1996**, 87, 1117–1157.
- (133) Nose, S. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys. **1984**, 52, 255–68.

- (134) Evans, D. J.; Holian, B. L. The Nose-Hoover thermostat. J. Chem. Phys. **1985**, 83, 4069–4074.
- (135) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Voth, G. A. Multiscale coarse-graining and structural correlations: Connections to liquid state theory. J. Phys. Chem. B **2007**, 111, 4116–4127.
- (136) Hansen, J.-P.; McDonald, I. R. Theory of Simple Liquids, 2nd ed.; Academic Press: San Diego, CA USA, 1990.
- (137) Hill, T. L. An Introduction to Statistical Thermodynamics; Addison Wesley Publishing Company, 1987.
- (138) Asakura, S.; Oosawa, F. On interaction between two bodies immersed in a solution of macromolecules. J. Chem. Phys. **1954**, 22, 1255.
- (139) Asakura, S.; Oosawa, F. Interaction between particles suspended in solutions of macromolecules. J. Polym. Sci. **1958**, 33, 183–192.
- (140) Binder, K.; Virnau, P.; Statt, A. Perspective: The Asakura Oosawa model: A colloid prototype for bulk and interfacial phase behavior. J. Chem. Phys. **2014**, 141, 140901.
- (141) Potestio, R. Effective interactions in soft condensed matter physics. J. Unsolved Quest. **2013**, 3, 13.
- (142) Wang, H.; Stillinger, F. H.; Torquato, S. Sensitivity of pair statistics on pair potentials in many-body systems. J. Chem. Phys. **2020**, 153, 124106.
- (143) Rudzinski, J. F.; Noid, W. G. Investigation of coarse-grained mappings via an iterative generalized Yvon-Born-Green method. J. Phys. Chem. B **2014**, 118, 8295–8312.
- (144) Shen, K.; Sherck, N.; Nguyen, M.; Yoo, B.; Köhler, S.; Speros, J.; Delaney, K. T.; Fredrickson, G. H.; Shell, M. S. Learning composition-transferable coarse-grained mod-

- els: Designing external potential ensembles to maximize thermodynamic information. J. Chem. Phys. **2020**, 153, 154116.
- (145) Rondina, G. G.; Böhm, M. C.; Müller-Plathe, F. Predicting the Mobility Increase of Coarse-Grained Polymer Models from Excess Entropy Differences. J. Chem. Theory Comput. **2020**, 16, 1431–1447.
- (146) Meinel, M. K.; Müller-Plathe, F. Loss of molecular roughness upon coarse-graining predicts the artificially accelerated mobility of coarse-grained molecular simulation models. J. Chem. Theory Comput. **2020**, 16, 1411–1419.
- (147) Jin, J.; Schweizer, K. S.; Voth, G. A. Understanding dynamics in coarse-grained models. I. Universal excess entropy scaling relationship. The Journal of Chemical Physics **2023**, 158, 034103.
- (148) Jin, J.; Schweizer, K. S.; Voth, G. A. Understanding dynamics in coarse-grained models. II. Coarse-grained diffusion modeled using hard sphere theory. The Journal of Chemical Physics **2023**, 158, 034104.
- (149) Rudzinski, J. F. Recent progress towards chemically-specific coarse-grained simulation models with consistent dynamical properties. Comput. **2019**, 7, 42.
- (150) Klippenstein, V.; Tripathy, M.; Jung, G.; Schmid, F.; van der Vegt, N. F. A. Introducing Memory in Coarse-Grained Molecular Simulations. The Journal of Physical Chemistry B **2021**, 125, 4931–4954.
- (151) Schilling, T. Coarse-grained modelling out of equilibrium. Physics Reports **2022**, 972, 1–45.
- (152) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazelwood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D. et al. XSEDE: Accelerating Scientific Discovery. Computing in Science & Engineering **2014**, 16, 62–74.

- (153) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. J. Mol. Graph. **1996**, 14, 33–38.

TOC Graphic

