

Toward Quantitative Coarse-Grained Models of Lipids with Fluids Density Functional Theory

Laura J. Douglas Frink,^{*,†} Amalie L. Frischknecht,^{*,‡} Michael A. Heroux,[‡] Michael L. Parks,[‡] and Andrew G. Salinger[‡]

[†]Colder Insights Corp., Shoreview, Minnesota, 55126, United States

[‡]Sandia National Laboratories, Albuquerque, New Mexico 87185, United States

ABSTRACT: We describe methods to determine optimal coarse-grained models of lipid bilayers for use in fluids density functional theory (fluids-DFT) calculations. Both coarse-grained lipid architecture and optimal parametrizations of the models based on experimental measures are discussed in the context of dipalmitoylphosphatidylcholine (DPPC) lipid bilayers in water. The calculations are based on a combination of the modified-iSAFT theory for bonded systems and an accurate fundamental measures theory (FMT) for hard sphere reference fluids. We furthermore discuss a novel approach for pressure control in the fluids-DFT calculations that facilitates both partitioning studies and zero tension control for the bilayer studies. A detailed discussion of the numerical implementations for both solvers and pressure control capabilities are provided. We show that it is possible to develop a coarse-grained lipid bilayer model that is consistent with experimental properties (thickness and area per lipid) of DPPC provided that the coarse-graining is not too extreme. As a final test of the model, we find that the predicted area compressibility moduli and lateral pressure profiles of the optimized models are in reasonable agreement with prior results.

1. INTRODUCTION

Coarse-grained modeling of biological systems has become more widely used in recent years. The interest in coarse-grained (CG) models stems from the terrific cost associated with treating these complex fluid systems with classical all-atom molecular dynamics (MD) simulations. Lipid bilayers in particular have been studied with a diverse range of models and methods (theory and simulation). Early MD simulations employed fairly generic CG lipid models;^{1,2} more recent efforts have included extensive fitting of CG models for lipids based on experiment and/or atomistic MD simulations in order to obtain realistic bilayer properties from CG simulations.^{3–12}

Statistical mechanical theories have also been used to study coarse-grained models of lipids. Self-consistent field theory (SCFT) has been widely applied to polymeric fluids¹³ and more recently to lipid bilayers.^{14,15} For biological membranes where the length scale of atoms is likely important, a complementary statistical mechanical approach is fluids density functional theory (DFT), which differs from SCFT in that it retains information on the length scale of the CG atoms. The advantages of using statistical theories to study lipid bilayers include less computational expense than MD simulations and direct access to the equilibrium free energy of the system as a natural output of the theory.¹⁶ We have previously applied a fluids-DFT to model single component lipid bilayers,^{17,18} the effects of alcohols on lipid bilayers,¹⁹ and the formation of pores in assemblies of simple α -helix antimicrobial peptides embedded in bilayers.²⁰

These previous fluids-DFT studies of biological membranes applied the Chandler–McCoy–Singer (CMS) DFT for polymer fluids^{16,21,22} to a simple, generic CG lipid model. This theory is based on a second order expansion of the free energy. One limitation of the CMS theory is that it cannot predict liquid–vapor transitions. Without access to a basic fluid

phase diagram, it is impossible to clearly define the thermodynamic state of the system, and therefore it is difficult to rationally set coarse-grained parameters.

More recent fluids-DFTs for bonded systems have been developed that are based on a perturbation to an atomistic reference fluid.^{23,24} As a result, the bonding constraints can be built on an accurate fundamental measures theory (FMT) model for the hard sphere reference fluid.²⁵ When FMT-based perturbative approaches are used to study attractive atomistic fluids, both bulk phase transitions and many novel phase transitions in confined media have been predicted.²⁶ The extension of FMTs to bonded fluids provides a new opportunity for careful parametrizations of complex fluid systems such as bilayer-forming lipids in water.

In this paper, we use a particular version of the newer bonded fluids-DFTs, namely, the “modified-iSAFT” DFT,²⁴ to investigate a variety of coarse-grained models of dipalmitoylphosphatidylcholine (DPPC) bilayers. The goal is to determine if it is possible to find parametrizations of the underlying molecular parameters that (1) obey qualitative expectations and key quantitative measures of the underlying chemical groups in the system and (2) are consistent with available experimental measures of DPPC bilayers and their constituent groups. Our coarse-graining approach is similar to that used in the development of the MARTINI force field.^{4,6} However, it is currently not possible to use a CG force field developed for use in MD simulations directly in the fluids-DFT because MD force fields often include terms not currently available to the theory (such as bond angle constraints). Furthermore, the mean-field nature of the theory leads to discrepancies in phase diagrams when compared to MD simulations. Thus, it is necessary to

Received: October 4, 2011

Published: March 9, 2012

develop CG models specifically for use with the fluids-DFT. Here, we demonstrate the feasibility of constructing a sensible CG model of DPPC for use with the modified iSAFT fluids-DFT.

We give a brief overview of modified-iSAFT and our computational methods in sections 2 and 3 and then proceed to describe coarse-grained models and their parametrization in sections 4 and 5. Novel numerical approaches that allow direct comparison between experimental results and fluids-DFT calculations are discussed along with the results. Detailed appendices that include background on theoretical and computational details are also included.

2. THEORY

This paper is based on a particular fluids-DFT approach known as the modified-iSAFT method. A description of the modified-iSAFT equations is provided in Appendix A, and a complete derivation of the theory can be found elsewhere.²⁴ In modified i-SAFT, the reference fluid consists of a simple fluid of hard spheres, which are the sites that make up the molecules of interest. The bonding constraints are treated as a perturbation to this reference hard sphere fluid. Attractions between the sites are also included as a perturbation.

This fluids-DFT approach starts with a definition of the grand free energy Ω , which is a functional of the densities of each fluid component $\rho_\alpha(\mathbf{r})$:

$$\Omega[\rho_\alpha(\mathbf{r})] = F[\rho_\alpha(\mathbf{r})] - \int d\mathbf{r}(\mu_\alpha - V_\alpha(\mathbf{r})) \rho_\alpha(\mathbf{r}) \quad (1)$$

where $F[\rho_\alpha(\mathbf{r})]$ is the intrinsic Helmholtz free energy, μ_α is the chemical potential of species α , and $V_\alpha(\mathbf{r})$ is an external field. This functional obeys a variational principle. The equilibrium density profile $\rho_\alpha^0(\mathbf{r})$ minimizes the functional $\Omega[\rho_\alpha(\mathbf{r})]$:

$$\left. \frac{\delta \Omega[\rho_\alpha(\mathbf{r})]}{\delta \rho_\alpha(\mathbf{r})} \right|_{\rho_\alpha^0} = 0 \quad (2)$$

and results in the grand potential free energy of the system, $\Omega = \Omega[\rho_\alpha^0(\mathbf{r})]$.²⁷ Equation 2 leads to a set of nonlinear integral equations that are solved for the density profiles.

The intrinsic Helmholtz free energy in the modified i-SAFT theory consists of a series of terms:

$$F[\rho_\alpha(\mathbf{r})] = F_{id}[\rho_\alpha(\mathbf{r})] + F_{hs}[\rho_\alpha(\mathbf{r})] + F_{att}[\rho_\alpha(\mathbf{r})] + F_{ch}[\rho_\alpha(\mathbf{r})] \quad (3)$$

where the terms on the right-hand side represent the intrinsic Helmholtz free energies for the ideal gas, hard sphere, attractive interactions, and chain constraints, respectively. More specifically, the chain term $F_{ch}[\rho_\alpha(\mathbf{r})]$ is based on an extension of Wertheim's first order thermodynamic perturbation theory,^{24,28} and the segment density profiles $\rho_\alpha(\mathbf{r})$ are the densities of each segment type, α , in the model fluid. Note that the Helmholtz free energy is constructed as a sum of independent terms related to various physical interactions. As a result, various hard sphere and attractive functionals could be used in conjunction with the modified iSAFT theory for bonded systems. This separation of physical effects gives the method considerable flexibility to combine existing fluids-DFT methods.

3. COMPUTATIONAL METHODS

The system of nonlinear integral equations that make up the modified i-SAFT theory are solved using the open source

software, Tramonto (see <http://software.sandia.gov/tramonto>) developed at Sandia National Laboratories. The set of residual equations associated with modified-iSAFT are detailed in eqs 13–17 of Appendix A and are solved in real space on a Cartesian grid. An iterative Newton's method is used to solve the nonlinear problem, and the embedded linear matrix problem is solved with a segregated Schur complement approach. These algorithms work well across several classes of fluids-DFTs on computational platforms ranging from desktop workstation platforms to distributed memory massively parallel computers.²⁹ The details of the Schur complement method in the context of the modified i-SAFT method are included in Appendix B.

The studies presented here rely heavily on arc-length continuation (ALC) and binodal constraint algorithms developed previously in the context of studies on porous materials.^{30,31} Binodal constraint calculations allow the simultaneous solution of density profiles in two different computational domains in which the chemical potentials of all species as well as the free energy are constrained to be equal across both domains. These are precisely the constraints of phase equilibria, and the method is reminiscent of Gibbs ensemble molecular simulations.^{32,33} When either or both of the computational domains contains a uniform bulk fluid, the free energy is directly related to the pressure:

$$\Omega_b = -pV \quad (4)$$

and so the binodal method can be used as a means to directly control pressure in the fluids-DFT calculation. The algorithms used for ALC and binodal constraints are discussed in detail in Appendix C.

4. COARSE-GRAINED MODELS OF DPPC

The key property of lipids that gives rise to their ability to self-assemble in water is their amphiphilic nature. Typical phospholipids consist of a polar segment (the *head* group), connected to two hydrocarbon chains which are nonpolar in character (the *tail* group). When this molecule is submerged in a polar solvent (water), the head groups form an interface between a nonpolar region rich in tail groups and the polar solvent. Depending on the precise geometry of the molecules, lamellar phases, hexagonal phases, or other complex phases such as the bicontinuous phase may form. In biological systems, the lamellar phase is the most important since lipid bilayers form the membranes of cells. This work will focus on a common model lipid, DPPC. Its structure is shown in Figure 1.

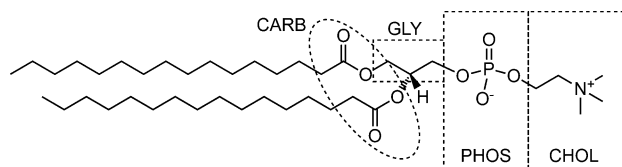


Figure 1. Chemical structure of DPPC, along with a definition of various parts of the head groups.

In virtually all CG models of lipids, the molecule will be split into hydrophobic and hydrophilic parts. Those parts will often be simplified into connected beads with uniform properties. Any CG model of a lipid bilayer also requires some treatment of the surrounding solvent (usually an electrolyte). The simplest CG solvent is a continuum model (implicit solvent) where the solvent has no discrete size and may or may not have a defined dielectric constant.¹¹

In this study, DPPC is treated as either a two-bead or a four-bead headgroup connected to a tail that is composed of two strands of CG hydrocarbon beads. Each tail strand is four, eight, or 16 beads long. The solvent is modeled as a single site that may represent one, two, or four water molecules. Note that the 16-mer tails are similar to a united atom type of approach, the 8-mer tails are similar to our previous work on CG lipids,¹⁷ and the 4-mer tails most closely correspond to work related to the MARTINI force field.^{4,6} In this work, all bonds have a fixed length, and the beads in the lipid are tangent to each other and freely jointed, with no angle or dihedral interactions. This freely jointed chain model for the lipids typically precludes the presence of a gel transition but is adequate for describing lipid bilayers in the fluid phase.^{17,18}

Here, we consider both the optimization of the CG lipid architecture and specific parametrizations of interaction potential parameters needed to fully define a CG model. Experimental measures for water, water–hexadecane interfaces, and DPPC bilayers are used as benchmarks for the study.

The nonbonded CG potentials consist of the hard sphere term, with attractions of a Lennard-Jones form:

$$u_{ij}(r) = 4\epsilon_{ij} \left\{ \left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right\} \quad (5)$$

For all interactions, a cut and shift of the potential is applied, so the interaction is hard sphere for distances less than the particle interaction radii and zero beyond a cutoff distance, r_c (see Appendix A for details). A fixed cutoff distance of $r_c = 3.5\sigma_{ss}$ is used throughout the paper. Electrostatics are not explicitly included. This approach assumes that electrostatic effects are highly screened and that in a coarse-grained sense, potentials are not dominated by the $1/r$ character of electrostatic interactions.

Assuming that $\epsilon_{ij} = \epsilon_{ji}$, $\sigma_{ij} = \sigma_{ji}$, and $\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2$, there are nine independent ϵ_{ij} and σ_{ii} parameters as well as two state parameters (temperature, T , and pressure, p) that must be rationally set to fully parametrize this rather simple three-component (head/tail/solvent) CG model of a lipid/solvent system. Note that the solvent–solvent parameters will be written without indices ($\sigma = \sigma_{ss}$ and $\epsilon = \epsilon_{ss}$). Section 5 presents calculations used to set the parameters and test different CG lipid architectures.

5. RESULTS

The overall strategy for developing a good CG model is to start with the simplest component of the system and then build complexity. So, the solvent model is considered first, followed by the hydrocarbon tails and the water–hydrocarbon interface. Finally, we consider known properties of DPPC bilayers to parametrize the CG head groups in the model.

5.1. State Point and Solvent Models. The thermodynamic state of water under physiological conditions ($T = 310$ K, $p = 1$ atm) is used as a guide for defining the state point and water parameters of the CG model. Most importantly, we require that the pure CG solvent will be in the liquid state near the triple point. This constraint reflects the state of water relative to the water phase diagram, which has a critical point at $T_c = 647$ K and a triple point at $T_t = 273$ K.

A single site LJ model is used for the solvent regardless of whether it represents one, two, or four water molecules. For the atomic (or 1-mer) LJ model, there is only one possible bulk

phase diagram in reduced units. The solution to the mean field fluids-DFT equations of Appendix A for this phase diagram is shown in Figure 2.

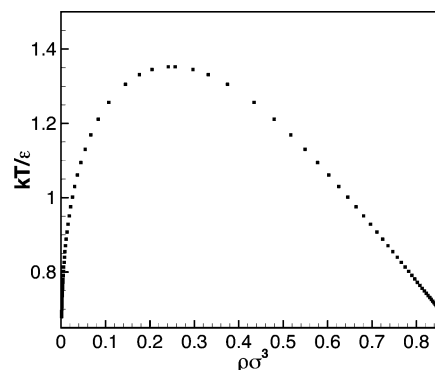


Figure 2. Liquid–vapor coexistence curve for a single segment solvent with the fluids-DFT discussed in Appendix A and the cutoff distance of $r_c = 3.5\sigma$.

The triple point of the LJ fluid is $kT_t/\epsilon \approx 0.65$ while the critical point is $kT_c/\epsilon = 1.35$, where k is the Boltzmann constant. Respecting the water phase diagram requires a temperature close to the triple point of the model system. So, we have used $kT/\epsilon = 0.68$ in all calculations. Energetics will be treated only in reduced units, so both T and ϵ are set by this choice.

Setting the density of the CG solvent is more ambiguous than the temperature both because the liquid-state densities of the LJ model fluid are systematically lower than water and because the density, $\rho\sigma^3$, is very sensitive to the size of the solvent σ . We have used the relative pressure of water as a guide. At 310 K, water has a relative pressure (compared to the saturation pressure) of $p/p_{\text{sat}} = 1 \text{ atm}/0.0728 \text{ atm} \approx 14$. We set the CG fluid density to $\rho\sigma^3 = 0.8656$ where the relative pressure of the model is also $p(\sigma^3/kT)/p_{\text{sat}} = 0.0175/0.001248 \approx 14$. All subsequent calculations are performed at constant pressure, with $p\sigma^3/kT = 0.0175$.

The specific value of σ in the CG solvent model depends on the number of water molecules that are represented by the CG site. A simple volumetric scaling is used to determine σ for three different models. Table 1 provides the values of σ for

Table 1. Summary of Solvent Parameters (σ and ϵ) for Various Coarse-Grained Models^a

model	# waters per bead	CG solvent volume	σ (nm)	ϵ/kT
M1	1	V_1	0.315	1.47
M2	2	$2V_1$	0.397	1.47
M4	4	$4V_1$	0.500	1.47

^aThe size of the solvent in M1 is set to a typical van der Waals diameter for oxygen.

cases M1, M2, and M4 that correspond to one, two, or four water molecules per CG site.

5.2. Hydrocarbon Models. The second constituent group of the lipid system that we consider is the hydrocarbon tail group. Hexadecane is most similar to the tail groups of DPPC and is used as the benchmark for determining tail–tail and tail–solvent parameters.

5.2.1. CG-Hydrocarbon Size. The CG hydrocarbon models are coarse-grained consistently with the solvent models so the M1, M2, and M4 models discussed above correspond to cases where there are one, two, and four CH_2 groups per CG segment, respectively. Using a QSAR based estimate of the volume of hexadecane ($V_{\text{C16}} \approx 283 \text{ \AA}^3$),³⁴ we find that $\sigma_{\text{tt}} = 1.03\sigma$ for all models. Neglecting the 3% difference, we set $\sigma_{\text{tt}} = \sigma$. A modern united atom approach will differ from the M1 case (e.g., $\sigma_{\text{tt}} = 0.396 \text{ nm}$ for CH_2 groups in the TraPPE potential³⁵) because those models are based on overlapping bonded spheres while the models here are based on tangent spheres.

5.2.2. CG Hydrocarbon–Hydrocarbon Interactions. The bulk liquid density of hexadecane is expected to be slightly lower than the bulk liquid density of water. We assume a reduced hexadecane density of $\rho\sigma^3 = 0.85$ and then find the isotherm, $kT/\varepsilon_{\text{tt}}$, that contains the state point $\{\rho\sigma^3 = 0.85, p\sigma^3/kT = 0.0175\}$. The desired state point is found with straightforward fluids-DFT calculations of pure bulk CG hexadecane to be located on the isotherms $\varepsilon_{\text{tt}}/kT = 0.96$ (M1), $\varepsilon_{\text{tt}}/kT = 1.0$ (M2), and $\varepsilon_{\text{tt}}/kT = 1.06$ (M4).

5.2.3. CG Hydrocarbon–Solvent Interactions. Two experimental measures of the interaction between water and hexadecane were used to parametrize CG solvent–hydrocarbon interactions. Water and hexadecane do not mix but rather form two nearly pure phases with a liquid–liquid interface. This demixing is the primary driving force for lipid bilayer formation, so it is particularly important to find parameters that reflect experimental measures. The two experimental measures used here are the interfacial tension, γ , and the partition coefficient, Θ . Known values for the hexadecane–water interface are $\gamma = 53.5 \text{ mN/m}$ ³⁶ and $\Theta = \Delta G/kT = -9.69$.³⁷

Calculation of Θ requires bulk (uniform) solutions to the fluids-DFT equations for coexisting alkane-rich and solvent-rich phases. The binodal constraint approach discussed in section 3 and Appendix C is used to find the coexisting states. It involves a simultaneous solution of the fluids-DFT equations in two computational domains along with the constraints of constant free energy (or bulk pressure) and constant chemical potentials for all species across the two domains.

Given the two coexisting solutions, it is straightforward to calculate the partition coefficient as

$$\Theta = \Delta G/kT = -\ln(\rho_{\text{w}}^{\text{C16}}/\rho_{\text{w}}^{\text{s}}) \quad (6)$$

where $\rho_{\text{w}}^{\text{C16}}$ is the density of water in the hexadecane-rich phase and $\rho_{\text{w}}^{\text{s}}$ is the density of water in the water-rich phase. ΔG can be interpreted as a measure of the free energy needed to move a solvent particle from a solvent-rich environment into a hydrocarbon-rich environment. Figure 3 shows how ΔG varies with the molecular parameter ε_{ts} . We observe that all model systems can achieve the experimental value for the partition coefficient.

Calculating γ requires the solution to the fluids-DFT equations in a single computational domain, but with inhomogeneous boundary conditions. On one side of the domain, there is the bulk water-rich phase, and on the other, the bulk hydrocarbon-rich phase. Those phases are guaranteed to have equal pressure, Ω_{b} , and chemical potentials because they are found from the bulk liquid–liquid equilibria calculation needed to compute partition coefficients above. In the calculation of the interface between those bulk phases, the fluids-DFT calculation yields the structure and free energy per

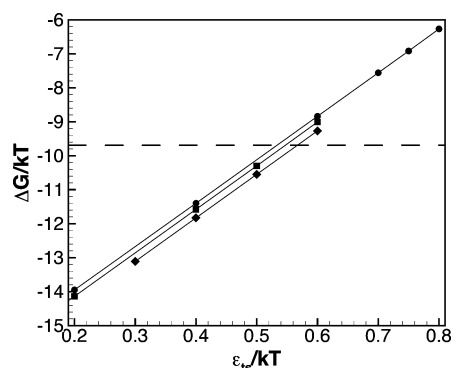


Figure 3. Partition coefficient as a function of ε_{ts} for the three coarse grained models at a pressure of $p\sigma^3/kT = 0.0175$. Curves are M1 (circles), M2 (squares), M4 (diamonds), and the experimental value (dashed line).

unit area (Ω/A from eq 27) of the water–hexadecane interface. The interfacial tension is then defined as

$$\gamma = (\Omega/A - \Omega_{\text{b}}/A) = (\Omega/A + pL) \quad (7)$$

where $\Omega_{\text{b}} = -pV$ is the free energy of (either) bulk homogeneous fluid.

To obtain an accurate γ , the computational domain must be large enough to fully resolve the interfacial structure while the solutions reach the bulk state at the boundaries. The calculations here were based on a domain size of $L = 40\sigma$. To ensure a stable interface, integration quadratures must be accurate enough in the bulk that the two pressures on the edges of the domain are equal. Appendix D discusses the computation of quadratures in Tramonto and the normalization of those quadratures that ensures accurate bulk calculations.

Figure 4 shows how the interfacial tension varies with ε_{ts} for the three CG models of the hexadecane–water interface. All of

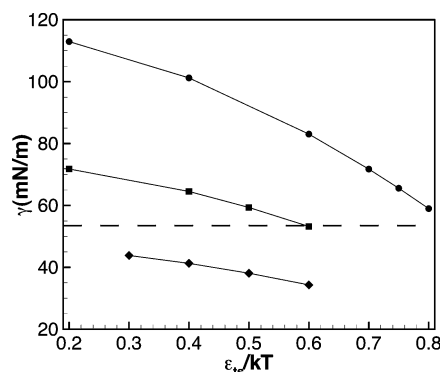


Figure 4. Interfacial tension as a function of ε_{ts} for the three coarse-grained models at constant pressure. Curves are identical to Figure 3.

the models can approach the desired value of γ , but at very different values of ε_{ts} . Figure 5 combines the data in Figures 3 and 4 to show the errors from expectation values of γ and Θ simultaneously. While a very good fit (<10% error) to both Θ and γ can be achieved for the 8-mer CG model (M2), the 4-mer (M4) and 16-mer (M1) models can only minimize the errors to within about 25% of the experimental values. Taking the best possible parametrizations, the 16-mer model has γ and Θ values that are too large while for the 4-mer model, they are too small. These results suggest that the 8-mer tail might be the

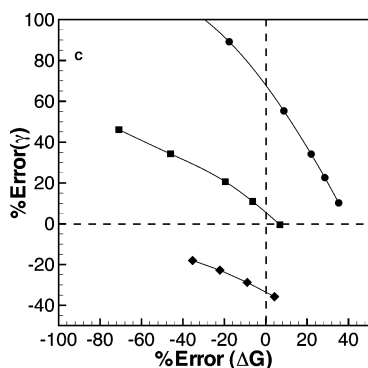


Figure 5. Simultaneous errors from experimental expectation values for interfacial tension, γ , and partition coefficient ΔG with variation in ϵ_{ts} . Curve identification is provided in Figure 3.

optimal model for bilayers since they best represent the properties of hexadecane.

Table 2 summarizes the best parameters for tail–tail and tail–solvent interactions of the CG models based on

Table 2. Coarse-Grained Molecular Parameters That Yield the Best Partition Coefficients, $\Theta = \Delta G/kT$, and Interfacial Tensions, γ , for Hexadecane/Water Systems^a

model	ϵ_{tt}/kT	ϵ_{ts}/kT	$p\sigma^3/kT$	$\Delta G/kT$	γ (mN/m)
M1	0.96	0.75	0.0175	−6.92	65.57
M2	1.0	0.6	0.0175	−9.01	53.23
M4	1.06	0.4	0.0175	−11.83	41.29

^aExperimental values are $\Theta = \Delta G/kT = -9.69$ and $\gamma = 53.5$ mN/m. $\sigma_{tt} = \sigma$ for all models.

experimental measures of the hexadecane–water interface. The table also includes computed values for the partition coefficient and surface tension. The parameters, ϵ_{tt} and ϵ_{ts} , shown in Table 2 are used in all of the studies of CG DPPC bilayers below.

Finally, Figure 6 shows hexadecane–solvent interfaces associated with the parameter sets in Table 2 for the three

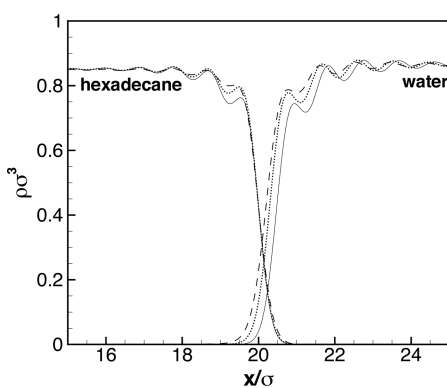


Figure 6. Density profiles at a hexadecane/water interface for all three coarse-grained models using the best parameters found in this study. The curves are solid lines (M4), dotted lines (M2), dashed lines (M1).

CG models. The three interface profiles are remarkably similar with a small increase in structure and broadening of the interface as the models become more coarse-grained.

5.3. Head Group Models. In this section, we discuss the construction of headgroup models for DPPC. The DPPC headgroup is composed of several chemical groups as shown in

Figure 1. Here, we consider three possible coarse-grained models for this system, as sketched in Figure 7. The first model

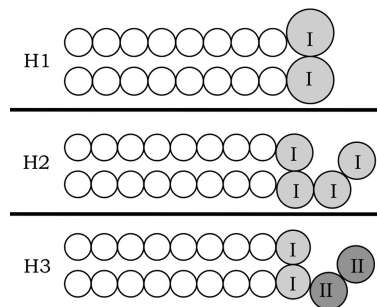


Figure 7. Sketches of three model coarse-grained headgroups using the same 8-mer (model M2) tail for each case. Both the H1 and H2 models have only one type of headgroup bead, while the H3 model has two types of headgroup beads.

(H1) is composed of two rather large beads. This model results in a linear lipid model.¹⁷ The second model (H2) is composed of four smaller identical beads.⁴ The final model (H3) has four CG beads of two different types. The larger outer beads would represent the ionic groups PHOS and CHOL, while the smaller inner beads (type I) correspond to CARB and GLY groups. This allows type II beads to be more hydrophilic than the type I beads.

The headgroup interaction parameters, ϵ_{hs} , ϵ_{hh} , and ϵ_{ht} (H1 and H2 models), are determined using the area per lipid, A_L , and the bilayer thickness, t , as benchmarks. Expected values for DPPC from the literature are $A_L \approx 61 \text{ \AA}^2$ and $t \approx 38.3 \text{ \AA}$.^{38,39}

5.3.1. CG Head-Group Size. The size of the CG beads in the various models in Figure 7 are computed on the basis of volumetric considerations, using the group volume data developed by Armen et al.⁴⁰ (Table 2 of that paper). Specifically, the total headgroup is taken to occupy 320 \AA^3 . The mean volumes for various chemical groups in the Armen et al. study were found to be 53.7 \AA^3 (phosphate group, PHOS), 108.60 \AA^3 (choline group, CHOL), 39.0 \AA^3 (carbonyl group, CARB), and 63.59 \AA^3 (glycerol group, GLY). The bead sizes in each of the CG models is given in Table 3. Note that the size of the CG headgroup beads is

Table 3. Particle Sizes for Three Head Group Models

model	σ_h^I (nm)	σ_h^{II} (nm)
H1	0.675	N/A
H2	0.536	N/A
H3	0.516	0.556

independent of the solvent and tail group models (M1, M2, or M4). So, any of the headgroup models could be coupled to any of the solvent/tail models.

5.3.2. Calculating CG Bilayers. The calculation of fluid bilayer structures with fluids-DFT has been discussed in detail elsewhere.^{16–18} The calculations presented here were performed in computational domains of size $L/\sigma = 20$. Reflecting boundaries were placed on each side of the domain with the center of the lipid bilayer at the right boundary. The full bilayer (including a reflected domain) is shown in the figures in this section.

Only zero tension bilayers are considered in the calculations presented here. The zero tension constraint can be expressed as

$$\gamma = 0 = \Omega/A - \Omega_{ps}/A \quad (8)$$

where $\Omega_{\text{ps}} = -pV$ is the free energy of a bulk pure (or nearly pure) solvent phase. The zero tension constraint is achieved in practice using the same binodal constraint algorithm used for calculation of the partition coefficients in the previous section. One computational domain contains the bilayer, while the other contains a nearly pure solvent bath (chemical equilibria demands a trace amount of lipid). The binodal algorithm enforces constant Ω between the two domains. Any of the headgroup parameters (ϵ_{hh} , ϵ_{hs} , or ϵ_{ht}) can be varied, and the chemical potential of the lipids is allowed to adjust to maintain equilibrium constraints. Since the lipid concentration in the bath is near zero, this has virtually no impact on solvent density in the solvent bath domain. This approach maintains a constant pressure and achieves a zero tension constraint for the bilayers.

5.3.3. Sensitivity Studies. Since there are two experimental benchmarks and three headgroup interactions of interest in the H1 and H2 models, a sensitivity study was done to determine if the bilayer properties were equally sensitive to all of the parameters. Both the H1 and H2 CG headgroup models were combined with the 8-mer chain tail (model M2) for these studies.

Figure 8 shows the sensitivity of A_L and t to the interaction parameters. Note that these figures provide a sample of the sensitivity studies performed (varying one molecular interaction parameter while the others are held constant). However, the studies did not employ formal optimization algorithms.

The experimental values for DPPC require the model to predict a thick, dense bilayer as indicated by the dashed lines in Figure 8. Thicker and more dense bilayers are achieved when ϵ_{hh} is increased or ϵ_{hs} is decreased. The area per lipid is relatively insensitive to ϵ_{ht} although the bilayer thickness is sensitive to all three parameters. To simplify the headgroup parametrizations, the head–tail interactions were set to $\epsilon_{\text{ht}}/kT \approx 0.3$ (or $\epsilon_{\text{ht}}/kT \approx 0.3$ for the H3 model) for all cases below.

Most of the data from the sensitivity studies in Figure 8 fall below the target for thickness and above the target for area per lipid, indicating that the CG models tend to predict bilayers that are too thin and not dense enough to reflect DPPC. However, the four-bead headgroup model (H2) is a significant improvement over the two-bead headgroup (model H1). The H2/M2 model can predict both a bilayer thickness and area per lipid that differs from experimental results by <20%. In contrast, the H1/M2 model never approaches the correct area per lipid and only achieves the correct thickness for unphysical parameter values where head–solvent interactions are becoming hydrophobic.

Therefore, we conclude that the H1 model is suitable only for qualitative studies of bilayers but would not likely produce good quantitative results. The large size of the CG headgroup beads in the H1 model relative to the solvent and tail group beads, $\sigma_{\text{h}}/\sigma = 1.7$, simply prevents efficient packing of the chains in the bilayer for reasonable parameter sets.

Select sensitivity studies were also performed for the H3 model that involved varying all seven of the headgroup interactions ϵ_{h1s} , ϵ_{h1t} , ϵ_{h1h1} , ϵ_{h1h2} , ϵ_{h2s} , ϵ_{h2h2} , and ϵ_{h2t} . The range of observed bilayer properties fell within the same general range as the H2 model results in Figure 8.

5.3.4. Bilayer Structure. In this section, we present a variety of bilayer structures predicted for several different combinations of head and tail/solvent models. We present both the optimal CG bilayer structures and demonstrate the effects of less optimal choices for CG lipid architectures.

To present density profiles for bilayers in a way that is most easily compared among the data sets, it is necessary to adjust the results for the coarse graining. The following adjustments must be made to density profiles computed for the various CG models:

- M2: multiply solvent and tail densities by 2
- M4: multiply solvent and tail densities by 4
- H1: multiply headgroup densities by 2
- M1, H2, H3: no adjustments for coarse-graining

The first two items above allow all of the solvent/tail profiles to reflect the density of individual water molecules and CH_2 groups rather than the CG groups that contain two or four of these chemical constituents. The third item allows all of the head groups densities to reflect four chemical groups within the CG model. This is clearly more ambiguous than the solvent/tail models because of the diversity of chemical groups that make up the headgroup. But, it does provide a common scale for comparison of the results.

For all of the calculations shown below, the parameters $\sigma_{\text{ss}} = \sigma_{\text{tt}}$, ϵ_{ss}/kT , ϵ_{tt}/kT , ϵ_{ts}/kT , σ_{hh} , and pressure are set to the values discussed previously and shown in Tables 1, 2, and 3. Table 4 shows all of the headgroup interactions that were used to prepare the figures in this section. The resulting bilayer properties are also included in Table 4.

Figure 9 shows two results for bilayers using the combined H1/M2 models. As was discussed above, the large H1 headgroup precludes obtaining the proper bilayer density regardless of molecular parameters. This chain architecture only approaches the proper thickness when the solvent–headgroup interactions are decreased dramatically. Figure 9B shows what

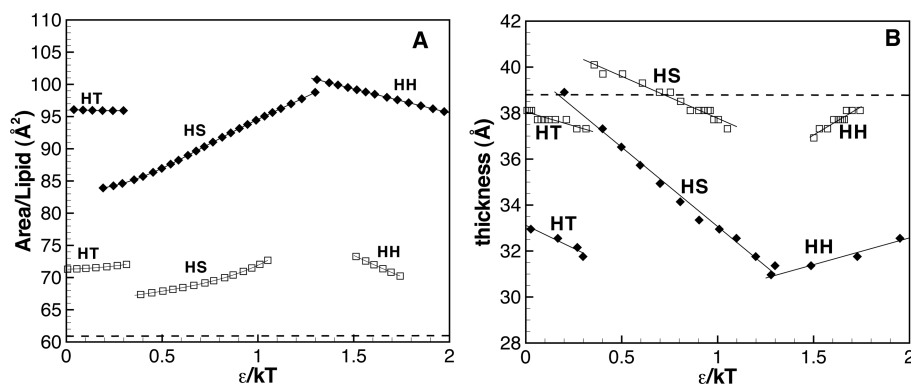


Figure 8. Sensitivity of the area per lipid (A) and bilayer thickness (B) to changes in the molecular parameters, ϵ_{hh} , ϵ_{ht} , and ϵ_{hs} (as indicated) for headgroup models H1 (diamonds) and H2 (squares). The experimental values are indicated by dashed lines.

Table 4. Head Group Interaction Parameters and Observed Bilayer Properties (Thickness t and Area Per Lipid A_L) Corresponding to Various Lipid Architectures Shown in Figures 9–11^a

figure	model	$\beta\epsilon_{hh}$	$\beta\epsilon_{hs}$	$\beta\epsilon_{ht}$	t (Å)	A_L (Å ²)
9A	M2/H1	2.5	1.3	0.3	32.6	94.7
9B	M2/H1	2.5	0.53	0.3	38.7	82.5
10A	M2/H2	2.0	1.1	0.3	38.3	69.2
11A	M4/H2	2.03	1.07	0.29	33.0	77.9
11B	M1/H2	1.97	1.13	0.31	41.0	66.0
10B	M2/H3	1.3	0.8	0.6	37.3	70.4
10B		$\beta\epsilon_{h_2h_2}$	$\beta\epsilon_{h_2s}$	$\beta\epsilon_{h_2t}$	$\beta\epsilon_{hh_2}$	
10B		2.5	1.3	0.3	1.3	

^aThe notation $\beta = 1/kT$ is used here, and in the case of the H3 head group, the notation h indicates type I interactions while h_2 indicates type II head bead interactions. The type II interactions are given in the last line of the table.

happens at low ϵ_{hs} . The head groups become significantly desolvated, and the chains pack more tightly together. A shoulder forms on the headgroup peak, suggesting that the system would prefer to have two headgroup peaks, but this is not possible with the linear lipid model.

This significant problem with the H1 model is mostly corrected in both the H2 and H3 models where four smaller beads make up the headgroup of the lipid. Figure 10 compares bilayers from the H2 and H3 models where the H2

parametrizations for ϵ_{hs} and ϵ_{hh} are rough averages of the two types in the H3 model as shown in Table 4. The parameters of the H3 model were set so the interactions of the type II headgroup beads (composed of the ionic CHOL-PHOS constituents) would be strong and hydrophilic while the type I groups (composed of CARB-GLY constituents) have a more hydrophobic character.

Both the H2 and H3 model bilayers in Figure 10 have about a 15% error in A_L . This could be improved by making the head groups more hydrophobic as was done for the H1 model in Figure 9. However, it is not physically rational to make this adjustment, and we conclude that the CG models in Figure 10 have been optimized as much as is reasonable for these simple models.

Figure 10 demonstrates that averaging properties among the chemical groups in the headgroup (as in model H2) can yield reasonable lipid bilayers with respect to the measures t and A_L . However, to study bilayer interactions with peptides or small molecules, it may be crucial to introduce different chemical groups as in the H3 model.

Finally, we compare lipid bilayer structures for a four-bead headgroup model (H2) coupled to the 4-mer chain tail (M4) in Figure 11A and the 16-mer chain tail (model M1) in Figure 11B. The interaction parameters used for these studies have only insignificant differences with those used for the 8-mer (M2) chain in Figure 10A. Comparing bilayer properties among these three cases, we find that the 16-mer chain model produces the best bilayer with less than 10% error in both

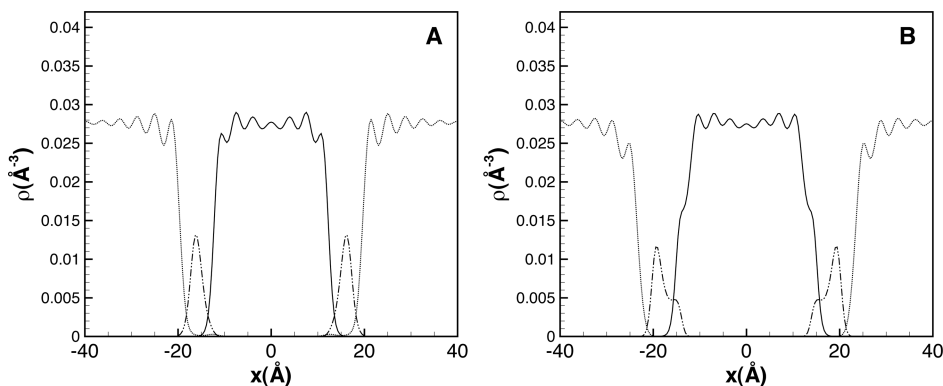


Figure 9. A comparison of two bilayer structures predicted with the H1 model headgroup. One case (A) has strong head–solvent interactions ($\beta\epsilon_{hs} = 1.3$). The other (B) has weak head–solvent interactions ($\beta\epsilon_{hs} = 0.53$). Dotted, solid, and dashed-dot-dot lines are the density profiles for solvent, tail, and head groups respectively.

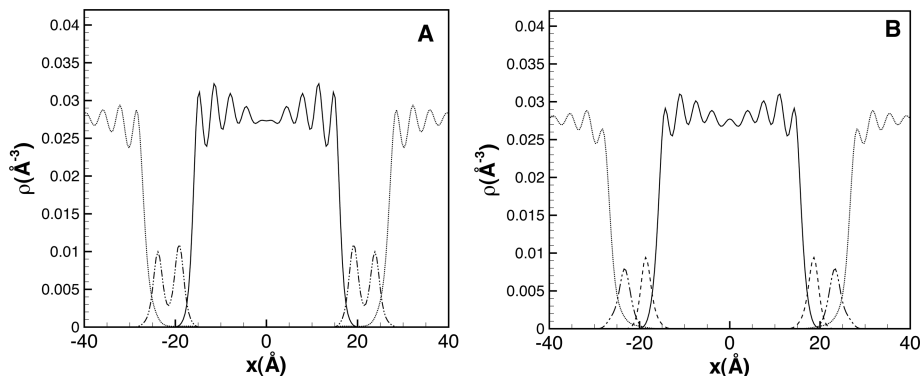


Figure 10. A comparison of H2 (A) and H3 (B) model bilayers using the optimized M2 solvent/tail model. The headgroup interactions of the uniform head model (H2) are set to be approximate averages of the nonuniform headgroup model (H3). The line types are the same as in Figure 9 except that the type I head groups are shown with a dashed line while the type II head groups are the dash-dot-dot line in B.

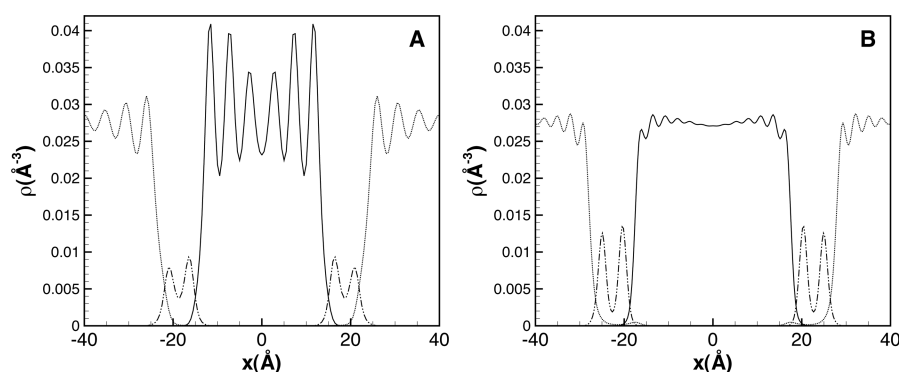


Figure 11. A comparison of 4-mer tails (M4) (A) and 16-mer tails (M1) (B) model bilayers using an H2 headgroup model with interactions nearly identical to those in Figure 10A. The line types are defined in the caption of Figure 9.

thickness and area per lipid. As the tail and solvent models are coarsened (from M1 to M2 to M4), additional structure is observed in the tail region of the bilayer. This structure is significant in the case of the 4-mer (M1) tails. Coarsening of the model also results in a thinning of the predicted bilayer along with a decrease in the lipid density.

All of the bilayer density profiles show multiple peaks in the solvent and tail group regions. These peaks are typically absent in bilayer density profiles obtained from MD simulations of CG lipid bilayers because of thermal fluctuations in the bilayer which lead to these peaks being averaged out. Being an equilibrium, statistical mechanical theory, the fluids-DFT does not include these thermal fluctuations, and hence the theory resolves the packing of the molecules. The packing effects are stronger here than in our previous fluids-DFT studies of lipid bilayers^{17,18} because the optimized models are at higher liquid densities and lower temperatures, where water is a relatively dense fluid.

5.3.5. Mechanical Properties. As a further test of the CG models discussed here, we calculate the area compressibility modulus K_A and the lateral pressure profiles of the bilayers. The experimental value of the area compressibility modulus for DPPC is $K_A = 231 \pm 20$ mN/m;³⁸ however, atomistic simulations on small bilayers typically report values of K_A that are about a factor of 2 larger than the experimental value.⁴¹ The coarse-grained MARTINI model gives $K_A = 400 \pm 30$ mN/m for a small system of 256 lipids, but $K_A = 260 \pm 40$ mN/m for a larger system with 6400 lipids.^{4,6} Because the fluids-DFT calculations do not include thermal fluctuations, we would expect the calculations to be consistent with the small system MD measures.

The area compressibility modulus is obtained from the dependence of the surface tension on the area per lipid near the point of zero tension:

$$\gamma \approx K_A(A_L - A_{L0})/A_{L0} \quad (9)$$

where A_{L0} is the area per lipid when $\gamma = 0$. For small deviations of A_L from its zero tension value, we find that γ is indeed linear in $(A_L - A_{L0})/A_{L0}$.^{1,18} Values of K_A for the model bilayers studied here are shown in Table 5. The area compressibility modulus is clearly quite sensitive to the degree of coarse-graining applied to DPPC. However, the fluids-DFT predictions are consistent with the observations of small MD simulations (K_A about a factor of 2 larger than experiment) as expected.

Table 5. Area Compressibility Modulus for Various Model Bilayers

figure	model	K_A (mN/m)
9A	M2/H1	251.8
10A	M2/H2	398.1
11A	M4/H2	325.5
11B	M1/H2	578.1

The lateral pressure profiles $\pi(x)$ are in general not accessible experimentally, but in the DFT they can be calculated directly from the excess grand free energy density $\omega(x)$, where^{17–19}

$$\pi(x) = -(\omega(x)/V - \Omega_b/V), \quad \Omega = \int \omega(x) dx \quad (10)$$

and V is the total volume of the system. The pressure profiles for two of the 8-mer models, models M2/H1 and M2/H2, are shown in Figure 12, while the lateral pressure profile for the 4mer model (M4/H2) is shown in Figure 13A and that for the 16mer model (M1/H2) is shown in Figure 13B. The pressure profiles show significant structure corresponding to the dense packing of the lipid chains. The broad features of the pressure profiles are similar to those calculated previously with simpler CG models.¹⁸ The pressure profiles are zero in the solvent region and display a negative peak near the headgroup/solvent interface corresponding to the headgroup/solvent interfacial tension, followed by a positive peak in the headgroup region corresponding to positive pressure from headgroup attractions, and then show a second negative peak at the headgroup/tail interface due to the head/tail interfacial tension. In all cases, the pressure is positive in the tail region, with a dip in the middle of the bilayer due to a lessening of the pressure near the tail ends. These general features are easiest to see in the least ordered bilayer, the M2/H1 model of Figure 12A.

The pressure profiles found here are similar to those calculated from MD simulations of CG lipid bilayers.^{1,8,18} Atomistic simulations also show qualitatively similar behavior, although there are known differences between atomistic and CG models with respect to lateral pressure profiles.⁸ In particular, the CG models generally have a dip in pressure in the middle of the tail region consistent with our calculations, while atomistic simulations generally show a peak in the middle of the tail region. Atomistic models yield peak pressures for the headgroup maxima in the range of 200–600 bar and head/tail interfacial troughs in the range of –500 to –1000 bar.^{8,42–44} DFT pressure profiles should have larger peak magnitudes than pressures calculated from simulation due to the lack of thermal

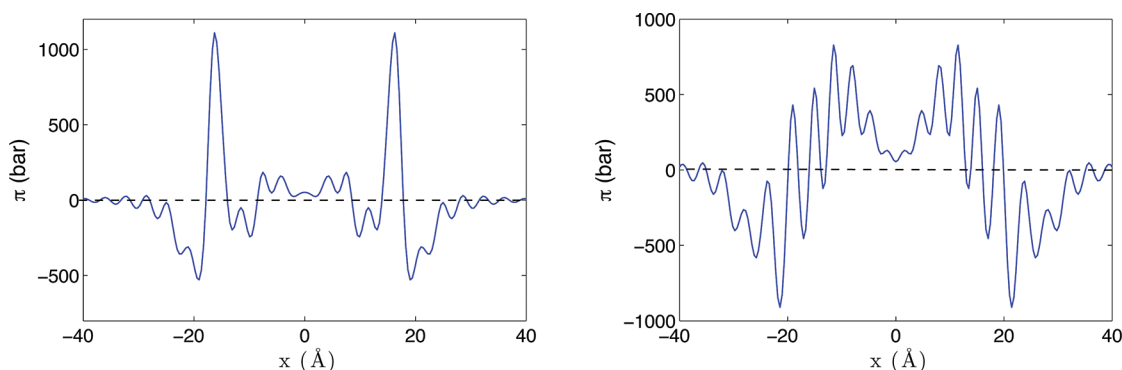


Figure 12. Lateral pressure profiles for 8-mer models M2/H1(left) and M2/H2(right), corresponding to the density profiles in Figures 9A and 10A, respectively.

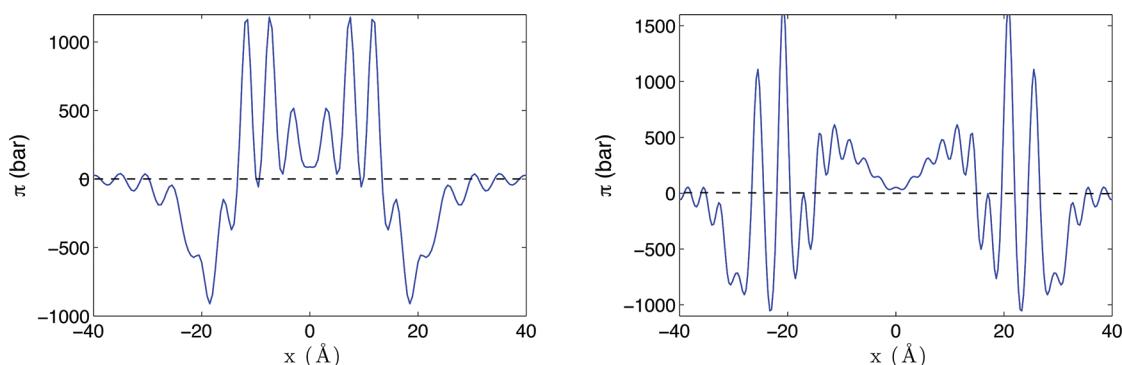


Figure 13. Lateral pressure profiles for models M4/H2(left) and M1/H2(right), corresponding to the density profiles in Figures 11A and B, respectively.

fluctuations in the fluids-DFT. We conclude that our calculated pressure profiles are in accordance with expectations for the type of CG model developed here.

6. CONCLUSIONS

We have applied a recently developed fluids density functional theory, the modified-iSAFT theory, to calculating the structure and thermodynamic properties of coarse-grained models of lipid bilayers. The goals of the study were to determine if the coarse-grained models could be systematically parametrized from experimental properties to predict realistic bilayers that will be useful for quantitative investigations.

In order to perform the parametrizations, a diverse set of numerical techniques was utilized including high performance solvers, bifurcation algorithms, and inhomogeneous boundaries. A novel binodal approach for performing pressure controlled fluids-DFT calculations was specifically developed for these studies and was discussed here. A further numerical extension that would be useful in future parametrization studies would be to couple automated optimization algorithms⁴⁵ to the Tramonto code.

The results presented here for coarse-grained lipid models demonstrate that it is possible to obtain lipid bilayers with structural properties (thickness and lipid density) within 10–20% of expected experimental values at approximately physiological conditions (temperature and pressure). These reasonable quantitative results were obtained with rational molecular parameters consistent with the physical expectations for various chemical groups of the lipid–solvent system.

However, the results also caution against extreme coarse-graining of lipid systems. When the head-groups are reduced to just two coarse-grained beads, it is no longer possible to achieve quantitative bilayer properties with any molecular interaction

potentials. Likewise when the hexadecane-like tails are reduced to just four beads, the errors in bilayer properties increase, and significant (most likely unrealistic) structure is introduced into the observed bilayers. In the former case, qualitative trends in bilayer behavior are likely reasonable; in the latter case, it is difficult to predict how strongly the artificial structure will affect other qualitative or quantitative studies.

The results presented in this paper provide a benchmark for improving coarse-grained models of lipid systems using fluids-DFT approaches. They also provide the numerical framework necessary for performing a variety of calculations that are needed for careful parametrization of lipid systems. Finally, given the bilayer models developed here, it is possible to begin to study more complex lipid bilayer based phenomena (bilayer interactions, small-molecule, and peptide insertion etc.) with the possibility of obtaining quantitative results.

■ APPENDIX A: THEORETICAL APPROACH

Here, we present the system of equations that form the basis for our calculations. We refer the reader elsewhere for a full derivation of the modified-iSAFT equations for linear chains.^{16,24,46,47} As with all fluids-DFT approaches, applying the functional derivative in eq 2 to the grand free energy in eq 1 results in a coupled set of integral (Euler–Lagrange) equations. The integral equations that are solved in the studies in this paper are presented in eqs 13–17 in the form of residuals. Solving eqs 13–17 by setting $R_{\text{type},\alpha}(\mathbf{r}) = 0$ yields the equilibrium density profiles, $\rho_a^0(\mathbf{r})$.

We have generalized the original modified-iSAFT equations to apply to branched chains.⁴⁸ A similar generalization for the case of associating fluids was recently done by Bymaster and Chapman.⁴⁹ The extension to branched chains is relatively

straightforward and is based on an analysis of the density distribution of segments on linear chains in a mean field $U(r)$. This mean field variable $U(r)$ is equivalent to the variable $\Xi(r)$ defined in the context of modified-iSAFT below. From the development of Donley et al. in the context of the CMS theory,²¹ we can write the density of segment i on a chain of length N as

$$\rho_i(r) = \rho_b \int dr_1 \dots \int dr_N e^{-U_N} \left(\prod_{j=1}^{N-1} w_{j,j+1} e^{-U_j} \right) \delta(r_i - r) \quad (11)$$

where ρ_b is the bulk monomer density and $w_{ij} = w_{ij}(r_i, r_j)$ is a delta function that constrains bond lengths to be fixed (see Table 6). For branched chains, which can have segments with more than two bonds, this expression generalizes to

Table 6. Weight Functions Associated with Various Parts of the Fluids-DFT Approach Used Here^a

variable	definition	purpose
w_{ij}	$\delta(\sigma_{ij} - r)/(4\pi\sigma_{ij}^2)$	define bond lengths
$w_{\alpha,HS}^{r=0}$	$\delta(R_\alpha - r)/(4\pi R_\alpha^2)$	HS reference fluid
$w_{\alpha,HS}^{r=1}$	$\delta(R_\alpha - r)/(4\pi R_\alpha)$	HS reference fluid
$w_{\alpha,HS}^{r=2}$	$\delta(R_\alpha - r)$	HS reference fluid
$w_{\alpha,HS}^{r=3}$	$\Theta(R_\alpha - r)$	HS reference fluid
$w_{\alpha,HS}^{r=v1}$	$(\vec{r}/ r) \delta(R_\alpha - r)/(4\pi R_\alpha)$	HS reference fluid
$w_{\alpha,HS}^{r=v2}$	$(\vec{r}/ r) \delta(R_\alpha - r)$	HS reference fluid
$w_{\alpha,CAV}^{r=2}$	$(\pi/6)\sigma_\alpha^2 \Theta(\sigma_\alpha - r - r')$	cavity correlations
$w_{\alpha,CAV}^{r=3}$	$(\pi/6)\sigma_\alpha^3 \Theta(\sigma_\alpha - r - r')$	cavity correlations
$w_{\alpha,\beta}^u$	$u_{\alpha\beta}^{\text{att}}(r)\Theta(r_c - r)$	dispersion interaction

^aIn all cases, $|r| = |r - r'|$. The δ -function terms are nonzero only when the pairwise distance r matches the defined set point: either the particle radius, R_α or particle diameter, σ_α . Θ is the Heaviside step function and is nonzero for all $|r|$ less than or equal to a distance given by R_α , σ_α , or a cut-off distance, r_c .

$$\rho_i(r) = \rho_b \int \left\{ \prod_{j=1}^N dr_j e^{-U_j} \right\} \left(\prod_{j=1}^{N-1} \prod_{b=1}^{N_b(j)} w_{j,k(b)} \right) \delta(r_i - r) \quad (12)$$

where $N_b(j)$ is the number of bonds of site j and the weight function $w_{j,k(b)}$ is for sites j bonded to sites k , indexed by the bond number b . In both the CMS-DFT and modified iSAFT, the multiple integrals in eq 11 or 12 are simplified by introducing a set of “propagator functions” $G_i^j(r)$, that propagate the probability distribution of sites along the chain from the end points. Through the definition of these functions as given below in eq 15, one can rewrite eq 12 as eq 13. This is then the generalized expression for the site densities for arbitrarily branched chains. Finally, we note that the modified-iSAFT equations provided here are based on segment type densities $\rho_{\alpha(i)}$ rather than on individual segment densities ρ_i as in the original modified-iSAFT derivation.²⁴ There is no loss of generality with this modification.

The system of equations that we solve has five types of residual equations. The R_1 and R_2 equations define the density and mean field variables, ρ_α and Ξ_α . The R_3 , R_4 , and R_5 equations define the propagators and two types of nonlocal density variables, ξ and n . We simplify the notation found in the

modified-iSAFT derivation by introducing a new field variable $\Xi_\alpha(\mathbf{r}) = \Xi_i(\mathbf{r}) = \exp(D_\alpha(\mathbf{r}) - \beta V_\alpha^{\text{ext}}(\mathbf{r}))$, where $V_\alpha^{\text{ext}}(\mathbf{r})$ is a one-body external field and $D_\alpha(\mathbf{r})$ (defined precisely below) is a measure of the intrinsic (fluid induced) field acting on segments of type α . The propagators defined here are given by $G_i(\mathbf{r}) = I_i(\mathbf{r}) \Xi_{\alpha(i)}(\mathbf{r})$, where $I_i(\mathbf{r})$ is the propagator used in the original modified i-SAFT derivation.²⁴ The five residual equation types are

$$R_{1,\alpha} = \rho_\alpha(\mathbf{r}) - \sum_{i \in \{\alpha\}} e^{\beta \mu_{M(i)}} \frac{\prod_{b=1}^{N_b(i)} G_i^{j(b)}(\mathbf{r})}{\Xi_\alpha^{(N_b(i)-1)}(\mathbf{r})} \quad (13)$$

$$R_{2,\alpha} = \ln \Xi_\alpha(\mathbf{r}) - D_\alpha(\mathbf{r}) + \beta V_\alpha^{\text{ext}}(\mathbf{r}) \quad (14)$$

$$R_{3,ij} = \frac{G_i^j(\mathbf{r})}{\Xi_\alpha} - \int \left\{ y_{ij}(\mathbf{r}, \mathbf{r}') w_{ij}(\mathbf{r}, \mathbf{r}') \times \left(\prod_{b=1}^{N_b(j); k(b) \neq i} G_j^{k \neq i}(\mathbf{r}') \right) \Xi_j^{(N_b(j)-2)}(\mathbf{r}') \right\} d\mathbf{r}' \quad (15)$$

$$R_{4,\varepsilon} = \xi_\varepsilon(\mathbf{r}) - \sum_\alpha \int \rho_\alpha(\mathbf{r}') w_{\alpha,CAV}^{(\varepsilon)}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \quad (16)$$

$$R_{5,\gamma} = n_\gamma(\mathbf{r}) - \sum_\alpha \int \rho_\alpha(\mathbf{r}') w_{\alpha,HS}^{(\gamma)}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \quad (17)$$

The indices i , j , and k refer to individual segments that are bonded together to form chains, α refers to unique components (or segment types, i.e. tail, head, and solvent groups), $N_b(j)$ is the number of chemical bonds (indexed with b) associated with the j th segment, and $\mu_{M(i)}$ is the chemical potential of a chain M on which segment i resides. The indices γ and ε indicate different individual nonlocal density variables. These nonlocal density variables are simple linear functionals (or spherical averages or convolutions) of the density profiles and a geometric weight function. The weight functions w_{HS} and w_{CAV}^ε are detailed in Table 6.

The propagators $G_i^j(\mathbf{r})$ defined in eq 15 enforce stoichiometry of the bonded system. The weight functions associated with bonding constraints, w_{ij} , are also defined in Table 6. Equation 15 is a recursion where the subscript i (or j) refers to a particular segment on the chain, and the superscript j (or k) refers to a neighboring bonded segment on the chain. Every segment will have N_b propagator equations associated with it except for the end points of the chain. Propagators G_i^{end} that represent an end segment i (and not any of the segments $j(b)$ bonded to i) obey the initial conditions

$$R_{3,i,\text{end}} = \frac{G_i^{\text{end}}(\mathbf{r})}{\Xi_{\alpha(i)}(\mathbf{r})} - 1 \quad (18)$$

The recursion relation eq 15 is biased by both the cavity correlation function $y_{ij}(\mathbf{r}, \mathbf{r}')$ and by w_{ij} . Since the latter is a δ function, only the contact value of $y_{ij}(\mathbf{r}, \mathbf{r}')$ is needed. The cavity correlation function in the inhomogeneous fluid is unknown, so following previous work we make the approximation

$$y_{ij}(\mathbf{r}, \mathbf{r}') = \sqrt{y_{ij}^{\text{contact}}[\xi(\mathbf{r})] y_{ij}^{\text{contact}}[\xi(\mathbf{r}')] } \quad (19)$$

where the contact value is calculated in the bulk reference fluid

$$\gamma_{\alpha\alpha'}^{\text{contact}}(\{\xi_i\}) = \frac{1}{1 - \xi_3} + \frac{3\sigma_\alpha\sigma_{\alpha'}}{\sigma_\alpha + \sigma_{\alpha'}} \frac{\xi_2}{(1 - \xi_3)^2} + 2 \left(\frac{\sigma_\alpha\sigma_{\alpha'}}{\sigma_\alpha + \sigma_{\alpha'}} \right)^2 \frac{\xi_2^2}{(1 - \xi_3)^3} \quad (20)$$

and where the ξ_i are defined by eq 16.^{46,47}

The most complex of the residual equations is R_2 because the intrinsic fluid contribution has several terms with different physical origins as in a typical perturbation scheme. Specifically,

$$D_\alpha(\mathbf{r}) = \frac{1}{2} \int \sum_j \rho_j(\mathbf{r}') \left\{ \sum_{b=1}^{N_b(j)} \frac{1}{y^{jk(b)}(\mathbf{r}')} \times \sum_{\epsilon=2,3} \frac{\partial y^{jk(b)}}{\partial \xi_\epsilon(\mathbf{r}')} w_{\alpha, \text{CAV}}^{(\epsilon)}(\mathbf{r}, \mathbf{r}') \right\} d\mathbf{r}' - \int \sum_\gamma \frac{\partial \Phi}{\partial \mathbf{n}_\gamma(\mathbf{r}')} w_{\alpha, \text{HS}}^{(\gamma)}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' - \sum_\beta \int \rho_\beta(\mathbf{r}') w_{\alpha\beta}^u(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \quad (21)$$

where the first term contains cavity correlations and comes from the chain bonding energy, the second term comes from the volume exclusion effects, and the last term is a strict mean field perturbation for attractions. The last two of these terms are identical to terms found in FMT-DFTs of nonbonded systems. This connection to those well studied theories is one of the key advantages of the modified iSAFT approach to bonded systems.

In the calculations presented here, the hard sphere term comes from the FMT known as the White Bear theory.⁵⁰ In this case, there are four scalar and two vector nonlocal densities $n_r(\mathbf{r})$. The excess Helmholtz free energy of this hard sphere model system is

$$\beta F_{\text{hs}}[\rho_\alpha(\mathbf{r})] = \int d\mathbf{r} \Phi[n_\alpha(\mathbf{r})] \quad (22)$$

where the energy density for the hard sphere system, Φ , is given by

$$\Phi = -n_0 \ln(1 - n_3) + \frac{n_1 n_2 - \vec{n}_1 \cdot \vec{n}_2}{1 - n_3} + (n_2^3 - 3n_2 \vec{n}_2 \cdot \vec{n}_2) \frac{n_3 + (1 - n_3)^2 \ln(1 - n_3)}{36\pi n_3^2 (1 - n_3)^2} \quad (23)$$

The various nonlocal densities $n_r(\mathbf{r})$ are defined by eq 17 along with the weight functions w_{HS} in Table 6.

The last term in eq 21 comes from a strict mean field treatment of attractions using the approximation

$$\beta F_{\text{att}}[\rho_\alpha(\mathbf{r})] = \frac{1}{2} \sum_{\alpha\beta} \int d\mathbf{r} d\mathbf{r}' u_{\alpha\beta}^{\text{att}}(|\mathbf{r} - \mathbf{r}'|) \rho_\alpha(\mathbf{r}) \rho_\beta(\mathbf{r}') \quad (24)$$

The attractive part of the potential u_{att} is based on the Weeks, Chandler, and Anderson (WCA) approach:⁵¹

$$u_{\alpha\beta}^{\text{att}}(r) = u_{\alpha\beta}(r_{\text{min}}) - u_{\alpha\beta}(r_c), \quad r \leq r_{\text{min}} \\ u_{\alpha\beta}^{\text{att}}(r) = u_{\alpha\beta}(r) - u_{\alpha\beta}(r_c), \quad r_{\text{min}} < r < r_c \\ u_{\alpha\beta}^{\text{att}}(r) = 0, \quad r > r_c \quad (25)$$

where the interaction potential $u_{\alpha\beta}(r)$ is cut and shifted at r_c . In this paper, $u_{\alpha\beta}(r)$ is the Lennard-Jones potential as given in eq 5.

The first term in eq 21 modifies the mean field for bonded systems. Notice that the density in this term, ρ_j , is a segment density. Since our solution is formulated for segment type densities ρ_ω this term is replaced by its definition

$$\rho_j(\mathbf{r}') = e^{\beta \mu_{M(j)}} \frac{\prod_{b=1}^{N_b(j)} G_j^{k(b)}(\mathbf{r}')}{\Xi_{\alpha(j)}^{(N_b(j)-1)}(\mathbf{r}')} \quad (26)$$

The solution to the system of equations provided here results in the equilibrium density profiles $\rho_\alpha^0(\mathbf{r})$, which are then used to compute the equilibrium grand free energy from²⁴

$$\beta \Omega = \sum_\alpha \int d\mathbf{r} \rho_\alpha^0(\mathbf{r}) [D_\alpha(\mathbf{r}) + \sum_{i,i \in \alpha} (N_b(i)/2 - 1)] + \beta F_{\text{hs}}[\rho_\alpha^0(\mathbf{r})] + \beta F_{\text{att}}[\rho_\alpha^0(\mathbf{r})] \quad (27)$$

■ APPENDIX B: NUMERICAL METHODS

The open source Tramonto software (see <http://software.sandia.gov/tramonto>) solves the nonlinear integral equations of eqs 13–17 in real space on a Cartesian grid. The nonlinear problem is treated with an iterative method based on Newton's method. The solution of the nonlinear problem is found through an iterative process where a new solution x_{i+1} is obtained from an old solution x_i and some update vector Δx . The update vector is found by solving the linear problem defined below. To maintain the physicality of solutions, the Newton step is scaled back if a full step would lead to unphysical values (negative densities for example). Using this numerical approach, solutions are obtained in 10–20 nonlinear iterations when a good initial guess is supplied to the solver. Subsequent solutions at nearby parameter values will typically converge in 3–6 Newton iterations without the need to curtail the Newton step.

Linear Solver

The discretized linear system can be written in the usual fashion as

$$A \Delta x = b(x_i) \quad (28)$$

where x_i is the vector of unknowns, $A_{ij} = \delta R_i / \delta x_j$ is the matrix of functional derivatives of the various residual equations, and $b(x_i)$ is the residual vector at the i th step. The structure of fluids-DFT is amenable to a segregated Schur complement approach for the linear solves.²⁹

The first step of the segregated Schur complement approach is to rewrite eq 28 in a 2×2 block formulation, as

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (29)$$

such that the inverse of A_{11} exists, and the action of A_{11}^{-1} can be easily applied. The Schur complement of A is

$$S = A_{22} - A_{21}A_{11}^{-1}A_{12} \quad (30)$$

Note that one can compute a matrix-vector multiplication with S without ever forming S directly. The solution for Δx_2 can be found by solving

$$S\Delta x_2 = b_2 - A_{21}A_{11}^{-1}b_1 \quad (31)$$

with a Krylov method, such as GMRES, which requires only matrix-vector multiplications with S . After x_2 is determined, Δx_1 is easily computed as

$$\Delta x_1 = A_{11}^{-1}(b_1 - A_{12}\Delta x_2) \quad (32)$$

Matrix Structure of Modified-iSAFT Equations

To construct the 2×2 block matrix described above, Tramonto includes all of the variables, ρ_α , Ξ_α , G_β , n_γ , and ξ_ϵ , explicitly in the matrix solve. The unknowns are segregated to an appropriate position in the matrix given the constraints on A_{11} . Specifically, A_{11} is composed of the equations that define G_β , n_γ , and ξ_ϵ (eqs 15–17), while the other variables (eqs 13 and 14) are located in A_{22} . If properly assembled, A_{11} can be formed as a lower triangular block.

By including G_β , n_γ , and ξ_ϵ as independent variables in the matrix problem, the fill complexity of matrix entries associated with eqs 13 and 14 is considerably reduced. This approach is akin to the FOSLS approach of reducing the order of higher order partial differential equations by introduction of variables. As written in eqs 13–17, the structure of the Schur matrix is

	n	ξ	G	Ξ	ρ
R_5	I	0	0	0	$B(\sigma/2)/C$
R_4	0	I	0	0	$B(\sigma)/C$
R_3	0	$B(\sigma)$	LT	D	0
R_2	$B(\sigma/2)$	$B(\sigma)$	$B(\sigma)$	$B(\sigma)$	$B(r_c)/C$
R_1	0	0	$D2$	D	D

(33)

where the residual type is included on the left and the variable is included on the top to clarify the segregation of variables. The structure of each physics based block as well as the overall 2×2 structure of the matrix are shown in eq 33. The notation in the blocks is as follows: I indicates an identity matrix with 1 on the main diagonal of the block, D indicates a diagonal (one entry per row) not necessarily unity, $D2$ indicates two entries per row, C indicates a constant block, and B indicates a banded structure where the width of the band is indicated as σ , $\sigma/2$, or r_c . The width of the banded blocks is controlled by the definition of the weight functions in Table 6.

Improving Numerical Performance

Extending the FOSLS-Inspired Approach for Attractions. To improve numerical performance, we make three modifications to the development above. First, we use the same FOSLS-inspired approach to introduce a new residual equation

$$R_6 = a_\alpha(\mathbf{r}) - \sum_\beta \int \rho_\beta(\mathbf{r}') w_{\alpha\beta}^\mu(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \quad (34)$$

This additional variable a can be located in the upper half of the 2×2 matrix because it will introduce only a diagonal in A_{11} . Thus, the dense block associated with attractions is

removed from A_{22} . This reduces the impact of attractions on memory requirements and generally decreases the solve time by a factor of 2 to 4. The structure of the matrix with this modification is

	n	ξ	a	G	Ξ	ρ
R_5	I	0	0	0	0	$B(\sigma/2)/C$
R_4	0	I	0	0	0	$B(\sigma)/C$
R_6	0	0	I	0	0	$B(r_c)/C$
R_3	0	$B(\sigma)$	0	LT	D	0
R_2	$B(\sigma/2)$	$B(\sigma)$	I	$B(\sigma)$	$B(\sigma)$	0
R_1	0	0	0	$D2$	D	D

(35)

Matrix Approximations. The second modification involves simplification of the R_2 equation. Rather than making the substitution for the segment density as indicated by eq 26, we can form an approximate matrix by setting

$$\rho_i \approx \rho_\alpha / N_\alpha \quad (36)$$

where N_α is the number of segments of type α on the chain. This approximation allows us to apply preconditioners that were previously developed for the CMS-DFT equations using this segregated Schur approach.²⁹ With this modification, the upper left sub-block of the A_{22} block becomes diagonal, and the structure of the R_2 equation is simplified. The revised matrix is then

	n	ξ	a	G	Ξ	ρ
R_5	I	0	0	0	0	$B(\sigma/2)/C$
R_4	0	I	0	0	0	$B(\sigma)/C$
R_6	0	0	I	0	0	$B(r_c)/C$
R_3	0	$B(\sigma)$	0	LT	D	0
R_2	$B(\sigma/2)$	$B(\sigma)$	I	0	D	$B(\sigma)$
R_1	0	0	0	$D2$	D	D

(37)

While there is no guarantee that the approximate matrix will still allow robust convergence to the solution, in practice we have found that this approximation worked well for 4-mer and 8-mer polymers. In the case of the 16-mer polymers, an exact Jacobian with new preconditioners was needed.

Physics Scaling. The R_1 equation (eq 13) contains one inconvenient term e^{μ_M} in the density residual equation. Since μ_M can be large, we have introduced a scaling factor, S , to reduce its effect on convergence. Specifically, we replace the density residual equation with

$$R_{1,\alpha} = \rho_\alpha(\mathbf{r}) \Xi_\alpha^* - \sum_{i;\{\alpha(i)\}} e^{(\beta\mu_M(i) + S_\alpha)} G_i^{i-1} G_i^{i+1} \quad (38)$$

where the sum is taken over all segments i of type α and $\Xi_\alpha^* = \Xi_\alpha^S$. The scaling factor S_α can be set to any convenient value. One option⁵² is to set

$$S_\alpha = N_{i,M}^* \mu_i \quad (39)$$

where $N_{i,M}$ is the number of segments of the segment type i on polymer type M and μ_i is the chemical potential of segment type i . However, any convenient scaling factor can be applied.

■ APPENDIX C: BORDERING ALGORITHMS AND BINODAL CONSTRAINTS

In order to study two phase equilibria, apply a constant pressure constraint, and obtain the large numbers of solutions needed for optimization studies, the solution of the DFT residual equations is coupled to bifurcation algorithms found in the open source library of continuation algorithms (LOCA). The two methods applied here were arc-length continuation and binodal constraints.

Arc-Length Continuation

The pseudo arclength continuation algorithm of Keller⁵³ allows continuation around turning points in systems with multiple solutions and phase transitions. To perform arclength continuation, the linear system defined by eq 28 is augmented with an arclength parameter s and an arclength equation g . The augmented system is then described by

$$\mathbf{f}(\mathbf{x}(s), \chi(s)) = 0 \quad (40)$$

$$g(\mathbf{x}(s), \chi(s), s) = 0 \quad (41)$$

where \mathbf{f} is the set of Newton residuals ($\{\mathbf{R}\}$ in eqs 13–17, χ is a parameter that will be varied, and g is an arclength equation defined below. Note that χ could be a state parameter such as μ or a molecular parameter such as ϵ . Both \mathbf{x} and χ are parameterized as functions of s . Following a solution curve requires the computation of \mathbf{x} and χ given the previous state \mathbf{x}_p and χ_p and an increment in the arclength parameter, Δs . In this work, we use the pseudo arclength equation:^{53,54}

$$g = \left(\frac{\partial \mathbf{x}}{\partial s} \right)^T (\mathbf{x} - \mathbf{x}_p) + \frac{\partial \chi}{\partial s} (\chi - \chi_p) - \Delta s = 0 \quad (42)$$

where the superscript T is the transpose operation (which could alternatively be written as a dot product). Equation 42 is just a linearization of the usual Euclidean distance formula in $N + 1$ dimensional space: $\sum (\mathbf{x} - \mathbf{x}_p)^2 + (\chi - \chi_p)^2 = (\Delta s)^2$. The augmentation of the original system with eq 42 frees the χ parameter as an additional unknown and introduces the step size along the curve Δs as a parameter, allowing for continuation around turning points (numerical spinodals).

The Newton iteration for the augmented system can be written as

$$\begin{bmatrix} \mathbf{J} & \frac{\partial \mathbf{f}}{\partial \chi} \\ \frac{\partial \mathbf{x}}{\partial s}^T & \frac{\partial \chi}{\partial s} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \Delta \chi \end{bmatrix} = - \begin{bmatrix} \mathbf{f} \\ g \end{bmatrix} \quad (43)$$

where \mathbf{J} , \mathbf{f} , and $(\partial \mathbf{f})/(\partial \chi)$ are recomputed each Newton iteration using the current iterate of χ . To solve the augmented system (eq 43), it is necessary to update both χ and the solution update \mathbf{x} at each iteration. While it is possible to construct a full $(N + 1) \times (N + 1)$ system and solve it once, the resulting matrix would be more dense than the original Jacobian. A bordering algorithm⁵⁴ exploits the typically sparse nature of \mathbf{J} and requires no new matrix formulation, by performing two solves of \mathbf{J} per iteration:

$$\mathbf{J}\mathbf{a} = -\mathbf{f} \quad (44)$$

$$\mathbf{J}\mathbf{b} = -\frac{\partial \mathbf{f}}{\partial \chi} \quad (45)$$

where \mathbf{a} and \mathbf{b} are temporary vectors. The new updates are then found by

$$\Delta \chi = - \left(g + \frac{\partial \mathbf{x}}{\partial s} \cdot \mathbf{a} \right) / \left(\frac{\partial \chi}{\partial s} + \frac{\partial \mathbf{x}}{\partial s} \cdot \mathbf{b} \right) \quad (46)$$

$$\Delta \mathbf{x} = \mathbf{a} + \Delta \chi \mathbf{b} \quad (47)$$

The arclength continuation algorithm finds all solution branches that are simply connected. Note that the solutions to eq 2 are not limited to energy minima but can also be saddle points. These are encountered after continuing through a spinodal point and do not lead to physically realizable equilibria yet typically reconnect to branches of minima. While there is no guarantee that all solution branches will be located, the approach is of great practical utility in defining solution complexity and identifying coexisting solutions.

Binodal Constraints

Two phase coexistence occurs when two solutions with different density profiles have the same free energy at the same thermodynamic state point. A phase transition can be located graphically by plotting the free energy Ω versus some parameter χ and seeing where two branches with different morphologies cross. However, this approach is very time consuming particularly when many solutions with the binodal constraint are required. Phase diagram mapping is not the only situation where this tedious process is encountered. In the current work, the pressure constraints needed to enforce zero tension for the bilayers while varying molecular parameters (section 5.3) and the calculation of coexisting water rich and hydrocarbon rich phases for partitioning calculations (section 5.2) both required multiple connected solutions with binodal constraints.

The binodal constraint algorithm used in the Tramonto/LOCA software uses Newton's method to converge simultaneously to a parameter value and two solution vectors which have equal free energies. Simple zero order continuation is then used to track the coexistence point as a function of a second parameter. The coexistence point is found by solving the following $2N + N_\chi$ equations:

$$\mathbf{f}(\mathbf{x}_1, \chi) = 0 \quad (48)$$

$$\mathbf{f}(\mathbf{x}_2, \chi) = 0 \quad (49)$$

$$G = \Omega(\mathbf{x}_1, \chi) - \Omega(\mathbf{x}_2, \chi) = 0 \quad (50)$$

where the $2N + 1$ unknowns are two distinct solution vectors, \mathbf{x}_1 and \mathbf{x}_2 , and the parameter χ which is identical in the two domains. The final equation $G = 0$ requires that the two profiles have equal free energies, Ω . In the studies presented here, a chemical potential acts as the *bifurcation* parameter ($\chi = \mu_M$), and a molecular interaction parameter, ϵ , acts as the *continuation* parameter.

A full Newton method for this system has the form

$$\begin{bmatrix} \mathbf{J}_1 & 0 & \frac{\partial \mathbf{f}_1}{\partial \chi} \\ 0 & \mathbf{J}_2 & \frac{\partial \mathbf{f}_2}{\partial \chi} \\ \frac{\partial \Omega_1}{\partial x_1} & -\frac{\partial \Omega_2}{\partial x_2} & \frac{\partial G}{\partial \chi} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_1 \\ \Delta \mathbf{x}_2 \\ \Delta \chi \end{bmatrix} = - \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ G \end{bmatrix} \quad (51)$$

Here, the subscript i on the variables \mathbf{f} , \mathbf{J} , and Ω represents evaluation with solution vector \mathbf{x}_i ($i = 1$ or 2).

As with the pseudo arclength continuation algorithm, a bordering algorithm is used to solve the system of equations in eq 51. The linear equations in the Newton iteration for the phase transition algorithm eq 51 can be equivalently formulated with two linear solves of the matrix \mathbf{J}_1 , two of the matrix \mathbf{J}_2 , and some simple algebra:

$$\mathbf{J}_1 \mathbf{a} = -\mathbf{f}_1 \quad (52)$$

$$\mathbf{J}_1 \mathbf{b} = -\frac{\partial \mathbf{f}_1}{\partial \chi} \quad (53)$$

$$\mathbf{J}_2 \mathbf{c} = -\mathbf{f}_2 \quad (54)$$

$$\mathbf{J}_2 \mathbf{d} = -\frac{\partial \mathbf{f}_2}{\partial \chi} \quad (55)$$

$$\Delta \chi = -\frac{G + \left(\frac{\partial \Omega_1}{\partial \mathbf{x}_1} \cdot \mathbf{a} - \frac{\partial \Omega_2}{\partial \mathbf{x}_2} \cdot \mathbf{c} \right)}{\frac{\partial G}{\partial \chi} + \left(\frac{\partial \Omega_1}{\partial \mathbf{x}_1} \cdot \mathbf{b} - \frac{\partial \Omega_2}{\partial \mathbf{x}_2} \cdot \mathbf{d} \right)} \quad (56)$$

$$\Delta \mathbf{x}_1 = \mathbf{a} + \Delta \chi \mathbf{b} \quad (57)$$

$$\Delta \mathbf{x}_2 = \mathbf{c} + \Delta \chi \mathbf{d} \quad (58)$$

The variables \mathbf{a} , \mathbf{b} , \mathbf{c} , and \mathbf{d} are temporary vectors of length N . The second solves of matrices \mathbf{J}_1 and \mathbf{J}_2 can be accelerated by reusing LU factorizations if using a direct solver and reusing preconditioners if using an iterative solver, as in this work. This algorithm is memory efficient since \mathbf{J}_1 and \mathbf{J}_2 are not needed at the same point in the algorithm and can use the same storage space.

The quantities in eq 56 with the $(\partial \Omega_i)/(\partial \mathbf{x}_i)$ terms can be quickly approximated with directional derivatives, such as

$$\frac{\partial \Omega_1}{\partial \mathbf{x}_1} \mathbf{a} - \frac{\partial \Omega_2}{\partial \mathbf{x}_2} \mathbf{c} \approx \frac{G(\mathbf{x}_1 + \varepsilon \mathbf{a}, \mathbf{x}_2 + \varepsilon \mathbf{c}) - G(\mathbf{x}_1, \mathbf{x}_2)}{\varepsilon} \quad (59)$$

where the perturbation is chosen as

$$\varepsilon = \delta \sqrt{\frac{\mathbf{x}_1 \cdot \mathbf{x}_1}{\mathbf{a} \cdot \mathbf{a} + \delta} + \frac{\mathbf{x}_2 \cdot \mathbf{x}_2}{\mathbf{c} \cdot \mathbf{c} + \delta}} \quad (60)$$

with $\delta = 10^{-6}$. This saves on the need to develop code for computing $(\partial \Omega_i)/(\partial \mathbf{x}_i)$.

The algorithm requires initial guesses for \mathbf{x}_1 , \mathbf{x}_2 , and χ , which usually come from picking two solutions from near a phase transition from an initial pseudo arclength continuation run.

■ APPENDIX D: QUADRATURE NORMALIZATION

When computing systems with inhomogeneous boundaries (as in Figure 6), it is particularly important that the bulk properties (density, chemical potential, and pressure) are computed exactly as even small errors in integration quadratures can lead to inhomogeneous pressures on the two boundaries resulting in an unstable interface. To eliminate this possibility, the numerical quadratures used in Tramonto are normalized to guarantee that the correct bulk properties for a uniform density profile will be obtained.

For all of the weight functions in Table 6, the value of the spherically integrated weight function V_s ,

$$V_s = \int w_s(\mathbf{r}) \, d\mathbf{r} \quad (61)$$

is known analytically. A standard numerical quadrature has points and weights, w_i , on the computational grid such that $V_s \approx \sum w_i$. To ensure correct bulk behavior Tramonto normalizes all numerical quadratures (formed as described below) using

$$w_i^* = w_i \left(\frac{V_s}{\sum_i w_i} \right) \quad (62)$$

so that V_s is exactly obtained as $V_s = \sum w_i^*$. This ensures exact bulk fluid results and therefore constant pressure for cases with inhomogeneous coexisting boundaries.

Analytical Quadratures

The numerical weights w_i for the 1-dimensional (1D) numerical problems (all cases in this paper) are assigned to mesh nodes i offset from an arbitrary (translatable) origin, using a linear projection of many gaussian quadrature points that sample each element in 1D. Each of those gauss point weights includes a geometric contribution from a circular slice of the full 3D weight function. The geometric weight at a given gauss point in 1D will be

$$w_{1D}^g(|x - x_0|) = \int_0^{2\pi} d\theta \int_0^h w_{3D}(|r - r_0|, R) r \, dr \quad (63)$$

where R is the radius of the 3D stencil defined in Table 6; x is the position of the 1D quadrature point relative to the origin of the stencil function, $r_0 = \{x_0, y_0, z_0\}$, and $h = (R^2 - |x - x_0|^2)^{1/2}$ is the radius of the circle within the 3D stencil function that has its origin at x .

If the integrations can be performed analytically, the geometric weights at any gauss point are simple to calculate. For example, if $w_{3D}(x_0, R) = \theta(R - |r|)$ (uniform weight of unity within the volume of a sphere of radius R), the geometric factor associated with out-of-plane integrations is

$$w_{G1D}^\theta(|x - x_0|) = \pi h^2 = \pi(R^2 - |x - x_0|^2) \quad (64)$$

If $w_{3D}(x_0, R) = \delta(R - |r|)$ (uniform weight of 1 on the surface of a sphere of radius R), the 1D weight function at any gauss point is

$$w_{G1D}^\delta(|x - x_0|, R) = 2\pi R \quad (65)$$

Notice that eq 65 will have nonzero contributions for all $|x - x_0| \leq R$ so the 3D δ function becomes a volumetric weight function when it is reduced for a 1D or 2D problem.

The analytic geometric factors in eqs 64 and 65 of course depend on the dimensionality, d , of the problem. For 2D problems, the analytical gauss weights for the θ and δ function stencils are

$$w_{G2D}^\theta(x, y) = 2h \quad (66)$$

and

$$w_{G2D}^\delta(x, y) = 2R/h \quad (67)$$

where $h^2 = R^2 - \{|x - x_0|^2 + |y - y_0|^2\}$ is the out-of-plane distance to the integration cutoff from any point (x, y) .

For 3D calculations, there are no out-of-plane geometric factors and these weights are simply

$$w_{G_{3D}}^{\theta}(x, y, z) = w_{G_{3D}}^{\delta}(x, y, z) = 1 \quad (68)$$

However, since the δ function is truly a surface integration in this case, it will require a unique treatment.

Numerical Quadratures

Not all of the 3D weight functions in Table 6 are simple enough to allow analytical out-of-plane integrations. One case where *numerical* out-of-plane calculations are required is the attractive pair potential stencil $w^u(|r - r'|)$, which is a convolution of a pair interaction u^{att} and a θ function with a range of r_c . In this case, gaussian quadrature is used for out-of-plane (in the r direction in eq 63) calculations with 12–40 gauss points depending on the distance of an element from the origin of the stencil. Given the long range nature of the attractions ($r_c = 3.5\sigma$ in this paper), the cost of the numerical out-of-plane calculations is significant so only one out-of-plane calculation is done per element.

Constructing the Nodal Integration Stencils

As discussed above, the integration stencils used in Tramonto to calculate the various terms in eqs 13–17 are based on the normalized nodal weight functions w_i^* . Before the normalization is applied, the nodal weight functions, w_i , are constructed using weights calculated on a subgrid of points within each element where the analytical or numerical out-of-plane contributions are determined. These subgrid weights are then split among all the nodes at the corners of the element. The number of subgrid points depends on the type of out-of-plane integration (analytical or numerical) and whether a particular element is fully within the integration radius. The default rules in Tramonto are

- *Analytical - element within integration radius:* Apply six subgrid points per dimension per element. Use gauss weights to construct w_i .
- *Analytical - element crosses integration radius:* Apply 40 subgrid points per dimension per element. Use a dense midpoint integration scheme to construct w_i .
- *Numerical integration:* Apply one subgrid point per element located at the center of the element. Split out-of-plane weight equally among all nodes at the corners of the element to obtain w_i .
- *3D surface integral:* Apply $40 \times 40 \times 40$ subgrid with spacing τ to all elements that cross integration radius. For all subgrid points that fall within $\tau/2$ of the integration surface, add a contribution to w_i at the corners of the element.

Once the nodal weights, w_i , are constructed as defined above, the normalization of eq 62 is applied. In all cases except the 3D surface integrals, the normalization factor $V_s/\sum w_i$ will be smaller than a few percent. In the case of the 3D surface integrals, the procedure outlined above is not expected to yield a surface integral directly but rather a scaled weighting that captures the surface of the sphere. Since V_s is known exactly, the normalization procedure is used to generate w_i^* values that are representative of the surface integral.

It has been pointed out by others⁵⁵ that there are no straightforward quadratures for accurately computing spherical surface integrals, and it has been suggested that the Tramonto 3D δ -function quadrature may not be adequate. It is quite simple to test this assertion by considering a problem with

symmetry (e.g., a hard sphere fluid near a planar wall) and comparing computations using 1D stencils to a full 3D calculation. We have performed this test and have found very good agreement. Specifically, a factor of two decrease in the mesh spacing decreases the discrepancies between 1D and 3D calculations by a factor of 8. Simple calculations also demonstrate that quadrature errors are typically much smaller than discretization errors. We conclude that the 3D δ function quadratures in Tramonto are adequate for 3D-fluids-DFT computations.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ljfrink@colderinsights.com; alfrisc@sandia.gov.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. This work was supported in part by the Sandia National Laboratories LDRD program and in part by the DOE Office of Science Advanced Scientific Computing Research (ASCR) Applied Mathematics program.

REFERENCES

- (1) Goetz, R.; Lipowsky, R. *J. Chem. Phys.* **1998**, *108*, 7397–7409.
- (2) Stevens, M. J. *J. Chem. Phys.* **2004**, *121*, 11942–11948.
- (3) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464–4470.
- (4) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (5) Elezgaray, J.; Laguerre, M. *Comput. Phys. Commun.* **2006**, *175*, 264–268.
- (6) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (7) Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. *J. Struct. Biol.* **2007**, *157*, 593–605.
- (8) Orsi, M.; Haubertin, D. Y.; Sanderson, W. E.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 802–815.
- (9) Lyubartsev, A.; Mirzoev, A.; Chen, L.; Laaksonen, A. *Faraday Discuss.* **2010**, *144*, 43–56.
- (10) Shinoda, W.; DeVane, R.; Klein, M. L. *J. Phys. Chem. B* **2010**, *114*, 6836–6849.
- (11) Wang, Z.-J.; Deserno, M. *J. Phys. Chem. B* **2010**, *114*, 11207–11220.
- (12) Lyubartsev, A. P.; Rabinovich, A. L. *Soft Matter* **2011**, *7*, 25–39.
- (13) Fredrickson, G. H. *The Equilibrium Theory of Inhomogeneous Polymers*; Oxford University Press: New York, 2006.
- (14) Mueller, M.; Katsov, K.; Schick, M. *Phys. Rep.* **2006**, *434*, 113–176.
- (15) Pogodin, S.; Baulin, V. A. *Soft Matter* **2010**, *6*, 2216–2226.
- (16) Frischknecht, A. L.; Frink, L. J. D. *Molecular Theory Applied to Lipid Bilayers and Lipid-Protein Interactions*. In *Biomembrane Frontiers: Nanostructures, Models, and the Design of Life*; Faller, R., Jue, T., Longo, M. L., Risbud, S. H., Eds.; Humana Press: New York, 2009; Vol. 2, Chapter 1, pp 1–39.
- (17) Frink, L. J. D.; Frischknecht, A. L. *Phys. Rev. E* **2005**, *72*, 041923.
- (18) Frischknecht, A. L.; Frink, L. J. D. *Phys. Rev. E* **2005**, *72*, 041924.
- (19) Frischknecht, A. L.; Frink, L. J. D. *Biophys. J.* **2006**, *91*, 4081–4090.
- (20) Frink, L. J. D.; Frischknecht, A. L. *Phys. Rev. Lett.* **2006**, *97*, 208701.

- (21) Donley, J.; Rajasekaran, J.; McCoy, J. D.; Curro, J. G. *J. Chem. Phys.* **1995**, *103*, 5061–5069.
- (22) Hooper, J.; McCoy, J.; Curro, J. *J. Chem. Phys.* **2000**, *112*, 3090–3093.
- (23) Yu, Y. X.; Wu, J. Z. *J. Chem. Phys.* **2002**, *117*, 2368–2376.
- (24) Jain, S.; Dominik, A.; Chapman, W. G. *J. Chem. Phys.* **2007**, *127*, 244904.
- (25) Roth, R. *J. Phys.-Condens. Mater.* **2010**, *22*, 063102.
- (26) *Fundamentals of Inhomogeneous Fluids*; Henderson, D., Ed.; Dekker: New York, 1992.
- (27) Evans, R. *Adv. Phys.* **1979**, *28*, 143–200.
- (28) Wertheim, M. J. *Stat. Phys.* **1984**, *35*, 35–47.
- (29) Heroux, M. A.; Salinger, A. G.; Frink, L. J. D. *Siam. J. Sci. Comput.* **2007**, *29*, 2059–2077.
- (30) Salinger, A. G.; Frink, L. J. D. *J. Chem. Phys.* **2003**, *118*, 7457–7465.
- (31) Frink, L. J. D.; Salinger, A. G. *J. Chem. Phys.* **2003**, *118*, 7466–7476.
- (32) Panagiotopoulos, A. Z. *Mol. Phys.* **1987**, *61*, 813.
- (33) Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. J. *Mol. Phys.* **1988**, *63*, 527.
- (34) Zhao, Y.; Abraham, M.; Zissimos, A. *J. Org. Chem.* **2003**, *68*, 7368–7373.
- (35) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.
- (36) Wu, D. M.; Hornof, V. *Chem. Eng. Commun.* **1999**, *172*, 85–106.
- (37) Schatzberg, P. J. *J. Phys. Chem.* **1963**, *67*, 776–779.
- (38) Nagle, J. F.; Tristram-Nagle, S. *Bba-Rev. Biomembranes* **2000**, *1469*, 159–195.
- (39) Kucerka, N.; Nagle, J. F.; Sachs, J. N.; Feller, S. E.; Pencer, J.; Jackson, A.; Katsaras, J. *Biophys. J.* **2008**, *95*, 2356–2367.
- (40) Armen, R. S.; Uitto, O. D.; Feller, S. E. *Biophys. J.* **1998**, *75*, 734–744.
- (41) Anézo, C.; de Vries, A. H.; Hölte, H. D.; Tieleman, D. P.; Marrink, S. J. *J. Phys. Chem. B* **2003**, *107*, 9424.
- (42) Lindahl, E.; Edholm, O. *J. Chem. Phys.* **2000**, *113*, 3882–3893.
- (43) Patra, M. *Eur. Biophys. J.* **2005**, *35*, 79–88.
- (44) Ollila, S.; Hyvonen, M. T.; Vattulainen, I. *J. Phys. Chem. B* **2007**, *111*, 3139–3150.
- (45) Design Analysis Kit for Optimization and Terascale Applications. <http://dakota.sandia.gov> (accessed March 5, 2012).
- (46) Tripathi, S.; Chapman, W. G. *Phys. Rev. Lett.* **2005**, *94*, 4.
- (47) Tripathi, S.; Chapman, W. G. *J. Chem. Phys.* **2005**, *122*, 094506.
- (48) Note that earlier work employing iSAFT included results on branched chains, although the theoretical derivation was only written explicitly for linear chains.^{46,47}
- (49) Bymaster, A.; Chapman, W. G. *J. Phys. Chem. B* **2010**, *114*, 12298–12307.
- (50) Roth, R.; Evans, R.; Lang, A.; Kahl, G. *J. Phys., Condens. Mater.* **2002**, *14*, 12063–12078.
- (51) Weeks, J. D.; Chandler, D.; Anderson, H. C. *J. Chem. Phys.* **1971**, *54*, 5237.
- (52) Chapman, W. Personal communication.
- (53) Keller, H. B. In *Applications of Bifurcation Theory*; Rabinowitz, P. H., Ed.; Academic: New York, 1977; Chapter: Numerical solution of bifurcation and nonlinear eigenvalue problems, pp 159–384.
- (54) Shadid, J. N. Ph.D. thesis, University of Minnesota: Minneapolis, MN, 1987.
- (55) Knepley, M.; Karpeev, D.; Davidovits, S.; Eisenberg, R.; Gillespie, D. *J. Chem. Phys.* **2010**, *132*, 124101.