Article

Construction of Multiscale Dissipative Particle Dynamics (DPD) Models from Other Coarse-Grained Models

Yinhan Wang and Rigoberto Hernandez*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We present a general scheme for converting coarse-grained models into Dissipative Particle Dynamics (DPD) models. We build the corresponding DPD models by analogy with the de novo DPD coarse-graining scheme suggested by Groot and Warren (*J. Chem. Phys.*, 1997). Electrostatic interactions between charged DPD particles are represented though the addition of a long-range Slater Coulomb potential as suggested by González-Melchor et al. (*J. Chem. Phys.*, 2006). The construction is illustrated by converting MARTINI models for various proteins into a DPD representation, but it not restricted to the usual potential form in the MARTINI model—viz., Lennard-Jones potentials. We further extended the DPD scheme away from the typical use of homogeneous particle sizes, therefore faithfully representing the variations in the particle sizes seen in the underlying MARTINI model. The accuracy of the resulting construction of our generalized DPD models with respect to several structural observables has been benchmarked favorably against all-atom and MARTINI models for a selected set of peptides and proteins, and variations in the scales of the coarse-graining of the water solvent.
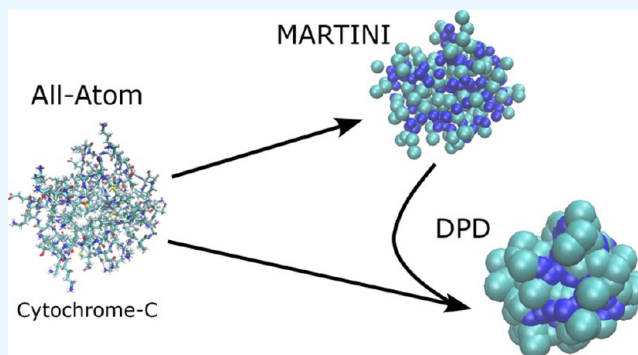
1. INTRODUCTION

Force fields for all-atom models (e.g., CHARMM36[1]) and for coarse-grained models (e.g., MARTINI[2,3]) have been widely used to reveal the structure and dynamics of molecular and biomolecular systems.[4−6] While the all-atom force fields provide highly accurate descriptions of protein structures, their use requires large amounts of computational resources limiting the size of systems that are accessible.[4,7,8] On the other hand, coarse-grained models reduce computational costs, while hopefully retaining acceptable accuracies.[9] The MARTINI force field[2] is among the most widely used coarse-graining model,[10] but it lacks a general description for non-biomolecular systems and omits, for example, interactions with and between arbitrary engineered nanoparticles. While coarse-graining schemes in Dissipative Particle Dynamics (DPD)[11] have accounted for such nanoparticles,[12] they have generally lacked an accurate representation of protein interactions.

In this paper, a generalized scheme for building a DPD model from existing finer-grained—but still sufficiently coarse-grained—force fields is shown to be useful for incorporating biomolecules. Our aim is not to make a better model for proteins but rather a model for proteins that we can use within a larger material framework. This is needed because, at the so-called nano-bio interface, we see interactions between engineered nanoparticles (including metals and other non-biological components), lipids (of varying complexity), proteins, and other large biomolecules.[13−16] These complex

mixtures become sufficiently large that they are often out of reach for all-atom models.[14−16] A challenge to the implementation of a coarse-grained model at the DPD scale(s) in describing the interaction of a nanoparticle with a bilayer in the presence of a membrane binding protein,[17,18] for example, is that we do not have a clear path for including proteins using a representation comparable in accuracy to the MARTINI representation. This is the challenge that we address through the scheme presented here. Moreover, we illustrate the scheme by coarse-graining several proteins at the DPD level through a construction based on an underlying MARTINI representation.[3]

Coarse-graining in DPD projects atoms and molecules into soft particles that interact through quadratic potentials at short-range and not at all after a specified cutoff distance.[11,19] There are at least three DPD protein models now available in the literature[20−22] making such an identification in ways relevant to this work: they will be referred to as the Pivkin,[20] Neimark,[21] and Gao[22] models throughout. Among these three models, the Neimark model does not have a complete

representation for all naturally occurring amino acids in proteins. The Pivkin model uses a severe approximation in which the pressure is assumed to be linear in the number density [refer to Figure 3]. While this works reasonably well for DPD liquids at very small number densities, it turns out not to be generally appropriate because many cases of interest are outside the circle of convergence of the underlying linear approximation. Instead, we can include a quadratic equation to fully characterize the inverse compressibility of DPD water particles while suffering only a modest computational cost in its implementation.

The present scheme is inspired by the Gao model[22]—which was developed to describe lipids—for the DPD parametrization of MARTINI particles.[22] As in the Gao construction, we first employ the coarse-graining approach suggested by Groot[23] to construct an initial working DPD model. To overcome the absence of charges in their model, we then introduce a Slater Coulomb potential.[24] We also introduce variations in the size of the particles in the revised DPD potential (commensurate with such variations in the higher fidelity MARTINI model).

In developing the model, we also need to account for the coarse-graining of the water solvent. As per the usual DPD construction,[23] we have coarse-grained water into beads (or particles)—referred to as W_4 throughout—corresponding to 4 water molecules at a length scale that is consistent with the coarse-graining of the protein. In so doing, we also found that further coarse-graining of the water solvent into ever larger particles can lead to modest errors while reducing the cost of the simulations. The limit of such an inconsistent coarse-graining would be to introduce a uniform or implicit solvent which is not an uncommon practice.[22] Coarse-graining 20 water molecules into a W_{20} particle, for example, allows for some heterogeneity in the solvent that would not be available from the implicit solvent. Herein, we also demonstrate the use of the W_{20} water solvent, not as a better water model (which it is not) but rather as a tool for solvating the target system which is immersed within it. It is accessible within the proposed scheme because the construction of the coarse-grained (CG) particles from one model to another also allows the connection of subsets of CG particles to be coarse-grained at a different scale.

The present scheme provides a direct construction of the DPD model based on arbitrary forms of the underlying potential. Notably, it is not limited to the Lennard-Jones forms of the potentials available in MARTINI. One part of the challenge in building a DPD model is the need for obtaining the complete interaction matrix. Flory−Huggins theory addresses this challenge well through a mean-field description of polymer solvation. We find that a corresponding approach can be used to formally convert any finer-grained coarse-grained model to a coarser-grained coarse-grained model, as represented in eq 17. We tested our DPD parametrization scheme for several small peptides and proteins building on the underlying MARTINI force field. The results on several benchmarking observables—viz., figures of merit related to the protein structure—suggest that the DPD models constructed in this way have a comparable accuracy to that resulting from MARTINI. We further tested the efficacy of this scheme to generate a DPD model for solvated proteins in a water solvent—e.g., W_{20} water particles—that is coarse-grained on a larger scale. The structure of the proteins in this W_{20} solvent differs from that in W_4 solvent, but the differences are not significant for features larger than the length scale of the coarse-graining of W_{20}.

2. METHODS

In the following section, we illustrate the process of constructing a coarse-grained model and specifically the DPD model as the chosen paradigm. This DPD model is designed to be equivalent to another coarse-grained model, namely, MARTINI, up to the extent of information loss in the second round of coarse-graining. It is notable that the parameters in the DPD model depend on the MARTINI parameters, and the accuracy of the former is subject to that of the latter model.

In addition, we report the further coarse-graining from 4 to 20 waters in the CG particles, aiming to reduce the computational costs associated with the solvent when solvent effects are indirect relative to the structure and motion of the solutes. A justification of the accuracy of this model is provided in the following section where DPD models are benchmarked against the all-atom and MARTINI models for several peptides and proteins.

2.1. Introduction to DPD Forces. DPD relies on the coarse-grained force field suggested by Groot and Warren.[11] It is widely used in the simulation of fluids.[25] The forces between DPD particles are short-ranged and linearly repulsive. The DPD interaction is soft, allowing particles to have partial overlap within their effective radii. The net forces include Langevin terms, accounting for the apparent fluctuations resulting from the fine-grained variables that are omitted. That is, for each particle, the force is a sum

$$f_i = \sum_{j \neq i} (F_{ij}^C + F_{ij}^D + F_{ij}^R) \tag{1}$$

of the constant force term F_{ij}^C, the dissipation term F_{ij}^D, and the random force term F_{ij}^R

$$F_{ij}^C = \begin{cases} a_{ij} w(r_{ij}) \widehat{r_{ij}} & r_{ij} < R_c \\ 0 & r_{ij} \geq R_c \end{cases} \tag{2}$$

$$F_{ij}^D = -\gamma w^2(r_{ij})(\widehat{r_{ij}} \cdot v_{ij}) \widehat{r_{ij}} \tag{3}$$

$$F_{ij}^R = \sigma w(r_{ij}) \theta_{ij} \widehat{r_{ij}} \tag{4}$$

where

$$w(r_{ij}) = 1 - r_{ij}/R_c \tag{5}$$

is the weighting function. R_c in the weighting function represents the cutoff distance. It is useful to define different R_c to accommodate differences in the sizes of the particles represented in the model.

The dissipation and random forces are correlated through the fluctuation−dissipation theorem ($\sigma^2 = 2\gamma k_B T$). Consequently, the only parameters needed to specify the force field are the repulsion parameter a_{ij} and the friction γ in units consistent with the rescaling. Notably, this dissipation parameter in the DPD force field is not the total dissipation γ_t that the system experiences. Instead, assuming linear response, it represents the contribution from the coarse-graining of the fine-grained variables, and it adds to the friction γ_0 from the solvent. The total effective friction γ_t is thus the sum, $\gamma_0 + \gamma$, as shown in ref 26, for the case of two additive random forces acting on a particle through a Langevin

equation. For example, we will see in Figure 4(a) that, while the plot exhibiting the ratio of terms in the fluctuation dissipation relation is linear in γ, it has a non-zero intercept due to γ_0.

2.2. Coarse-Graining All-Atom Models to DPD Particles.

The particle mapping of DPD proteins implemented here is based on the previous work of Gao and co-workers.[22] They represent particles similarly to the coarse-grained particles in the MARTINI force field, except for additional contractions of different sets of particles represented by a corresponding common representative particle.

In both force fields, particles are categorized according to 5 main types: water (W), polar (P), non-polar (N), a-polar (C), and charged (Q). The degree of coarse-graining of water into a DPD particle is denoted throughout as W_N particles where N denotes the number of water molecules in the effective DPD particle. The polar type is divided into 5 subtypes with different polarities: P1 with weakest polarity and P5 with highest polarity. It should be noted that a water particle is the same as a P4 particle. Type N and Q can have various hydrogen bonding designations, including donor (d), acceptor (a), both (da), and none (0). There are thus 8 possible combinations: Q_0, Q_d, Q_a, Q_{da}, N_0, N_d, N_a, and N_{da}. The mass of each of these DPD particle types is 1.0 in reduced units. A mesh of a coarse-grained MARTINI Cytochrome-C (CytC) equipotential is overlaid on the fine-grained all-atom structure in Figure 1.
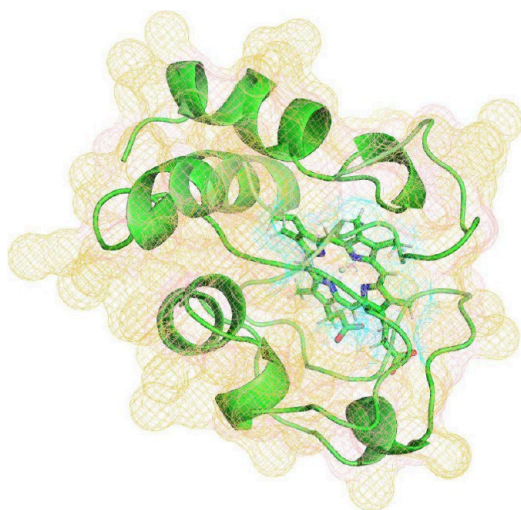

Figure 1. The structure of CytC (in yellow) is the same in both the coarse-grained MARTINI and DPD models and shown here relative to its cartoon representation (in green).

The MARTINI 2.2 force field has two types of particle sizes: large particles that are the size of 4 water molecules ($N_m = 4$) and small particles on aromatic rings that are the size of approximately 3 water molecules ($N_m = 3$), which are the particles denoted with a starting letter "S" in the MARTINI force field. Though Gao's work neglected the size difference between the particles, our simulations show that the stability of the DPD protein structure is improved by restoring this hybrid particle size in the DPD protein model. As a result, in the following sections, the regular sized $N_m = 4$ particles with mass 1.0 are labeled as L particles and the smaller $N_m = 3$ particles with mass 0.625 are labeled as S particles. The latter mass is not 0.75, as one might expect from the ratio 3:4 of particle

sizes S:L, because the S coarse-grained particles modeled here represent fine-grained particles on aromatic rings that have smaller densities. We still use a mass of 0.75 in the process of calculating parameters for S particles because these particles represent water during the parameter calibration process. The detailed calibration of the underlying parameters is discussed in Section 2.3 and Section 2.5.

In addition to the L and S particles, there are also charged Q and C1/C2 particles in MARTINI which are characterized by "IX" interactions. The latter are characterized by $\sigma = 0.62$ nm in the Lennard-Jones potential, while L and S particles have $\sigma = 0.47$ nm and $\sigma = 0.43$ nm, respectively. This is a correction in distance, instead of size, as the ϵ is unchanged in the L particle and IX interactions. The IX interaction allows for the insertion of a particle between the interacting particles as is seen in the fined-grained models.

Such insertions are possible for particles experiencing Lennard-Jones interactions within the MARTINI force field because all the forces between three particles are mildly attractive or repulsive. Unfortunately, they are rarely accommodated in DPD force field models because the force between any two particles is typically a large repulsion. Likewise, in the present work, we ignore the IX interaction in the DPD model by treating all IX interactions as those from regular L particles.

The MARTINI force-field representation is converted to DPD through a mapping from most of the all-atom protein structure to a coarse-grained structure performed by the MARTINI-maker module in CHARMM-GUI.[27] The exception is the heme in CytC which is not available from CHARMM-GUI but can be built by hand following a scheme analogous to that employed in CHARMM-GUI. The interactions between particles are characterized by two types of pair interactions: the Slater Coulomb potential between charged particles and the DPD potential between all particles. These potentials are implemented within the hybrid/overlay package in Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) and retained in input files.

Though only two particle sizes are included in the MARTINI model, larger particles can also be introduced in the converted DPD model to support larger-scale heterogeneous simulations. The possibility of such a multiscale DPD model by further coarse-graining water from W_4 particles into W_{20} particles is reported in Section 3.3.

2.3. DPD Unit Length.

In the present model, the majority of the particles represent a volume of 4 water molecules ($N_m = 4$). The mass of each such DPD water particle is set to 1, and all other masses are ratioed to this effective mass—that is, 72 g/mol for four water molecules or 1.20×10^{-23} g per particle. The DPD unit length is determined by matching the volume of the water molecules and the volume of a DPD particle. We can estimate the volume of a water molecule as 29.9 Å³ because the molar density of water at SATP is 18.0151 cc/mol. Initially, we take the number density of the DPD liquid as $\rho = 3$ in keeping with the earlier work of Groot.[11,23] As each unit DPD cube thereby contains 3 DPD water particles this amounts to setting the effective DPD length scale as[11]

$$R_c^* = \sqrt[3]{\rho \times N_m \times 30 \text{ Å}^3} = \sqrt[3]{3 \times 4 \times 30 \text{ Å}^3} = 7.11 \text{ Å}$$

(6)

Recall that the diameter of water is around 2.75 Å. Thus, 7.11 Å is a length equal to that spanned by 2.59 water molecules in

line. Consequently, this corresponds to about 3 DPD L particles. The natural unit length, R_c^*, is distinguished from the DPD cutoff parameter R_c because the latter is in reduced units; i.e., distances are converted to DPD units by dividing them by R_c^*. For a DPD model with a single particle size, the cutoff parameter R_c can be set to 1 for all particles, while for the multiscale DPD model with different particle sizes, we can represent particles with different sizes by setting a different cutoff parameter for each particle, all ratioed to a common R_c^*. In this work, we choose the coarse-grained water particle with $N_m = 4$ as the reference DPD particle with a cutoff distance $R_c = 1$.

In what follows, we take a default non-water DPD particle to be of the same size at the water DPD particle and call them L particles. Smaller S particles correspond to $N_m = 3$, and their cutoff parameter R_c decreases to 0.909. Instead of matching different particle sizes by their compressibility, Kacar et al.[28] suggested that the pressure across a multiscale model would be a more appropriate choice for matching. That is, particles with different sizes should exhibit the same pressures rather than the same compressibilities. Indeed, we found that this choice leads to better agreement with the continuum behavior in the limit of coexistence of particles with large size differences. Meanwhile, the cross-term interactions between L and S DPD particles are treated in the same way as the interactions between two L DPD particles. This follows the approach used in the MARTINI force field in treating the cross term interactions.[3] Our attempts to treat the L−S cross-term interactions by representing them as originating from two middle-sized particles led to an unacceptably large radius of gyration and a Root Mean Square Deviation (RMSD) that was too broad.

The effective equations of motion are those shown above in eq 1. The effective length scale leads to a way of mapping the coordinates of atoms in an all-atom model to the DPD model as described in Section 2.2. It is also used in converting the bonded interaction parameters in Section 2.8. The integration of this model leads to a scaling of time, though the resulting trajectories sample the space correctly.

2.4. Electrostatic Interactions. In modeling the interactions between lipids, Gao and co-workers accounted for the simulated electrostatic interactions by assigning larger repulsion parameters to the DPD particles containing repulsive charges.[22] Here, we avoid making this reassignment of the DPD parameters by adding instead Slater−Coulomb potentials to the bare DPD potential between all charged beads.

As the DPD potential allows particles to overlap with each other, the electrostatic interactions between charged particles should also be modified to allow overlap. Groot[29] suggested a linearly smeared charge to allow overlap, while González-Melchor et al.[24] simplified the expression by changing it to Slater potentials—viz., exponential smeared charges

$$\rho(r) = \frac{q\beta^2}{\pi r} \exp(-2\beta r) \tag{7}$$

where the potential and force acting on the particles are (see also Figure 2)[24]

$$U_{ij} = \frac{\Gamma q_i q_j}{4\pi r_{ij}}[1 - (1 + \beta r_{ij}) \exp(-2\beta r_{ij})]$$

$$\vec{F}_{ij}^{\mathrm{E}} = \frac{\Gamma q_i q_j}{4\pi r_{ij}}[1 - \exp(-2\beta r_{ij})(1 + 2\beta r_{ij}(1 + \beta r_{ij}))]\frac{\vec{r}_{ij}}{r_{ij}} \tag{8}$$
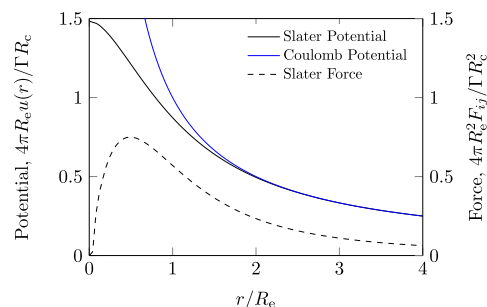

Figure 2. Slater potential (black) of eq 8 and the Coulomb potential (blue) $1/r$ using the parameters implemented in this work shown as solid curves corresponding to the left axis labels. The underlying Slater potential is shown as a dashed curve corresponding to the right axis labels. All three are in agreement with the corresponding curves in Figure 1 of ref 24.

The parameter β is an auxiliary number chosen such that the smeared potential generates neither ion pairing nor dissociated electrolytes.[29] For the case of a model with $N_m = 3$ and $R_c = 6.46$ Å, González-Melchor et al.[24] set $\beta_{N_m} = \beta_3 = 0.929$ in units of inverse DPD lengths. Assuming linearity in the coarse-grained scales, the auxiliary parameter for the present model is

$$\beta_4 = \frac{\beta_3 \times R_{c,N_m=3}}{R_{c,N_m=4}} = \frac{0.929 \times 6.46}{7.11} = 0.844 \tag{9}$$

The unit of charge is also converted to DPD units through the relation[30]

$$q_{\mathrm{DPD}} = \frac{q}{\sqrt{4\pi \epsilon_0 R_c^* k_B T}} = 8.86 \tag{10}$$

for each unit of electric charge. The dielectric constant of the Slater potential is set to $\epsilon_r = 78.3$, which is consistent with typical continuum water models and was also employed in DPD by González-Melchor et al.[24] and Groot.[29] The potential is available using the "coul/slater/long" pair style command in versions of LAMMPS which include bug fixes as of July 2022.

2.5. Determination of the DPD Water Repulsion Parameter. The repulsion parameter a_{ww} between two DPD water particles is determined by matching the inverse compressibility of the DPD water with real water. The inverse compressibility of DPD water is measured by the slope of pressure vs number density ρ at a given ρ, from the equation suggested by Groot[23]

$$\frac{1}{k_B T}\left(\frac{\partial P}{\partial \rho}\right)_{\mathrm{simulation}} = \frac{N_m}{k_B T}\left(\frac{\partial P}{\partial \rho}\right)_{\mathrm{experiment}} = N_m \times \kappa^{-1} \tag{11}$$

where $\kappa^{-1} = 15.9835$ is the experimental value of the dimensionless inverse compressibility of water. The resulting inverse compressibilities and parameters are shown in Figure 3.

2.6. Determination of the Dissipative Parameter γ. The dissipative parameter of an L particle is $\gamma = 4.5$, as
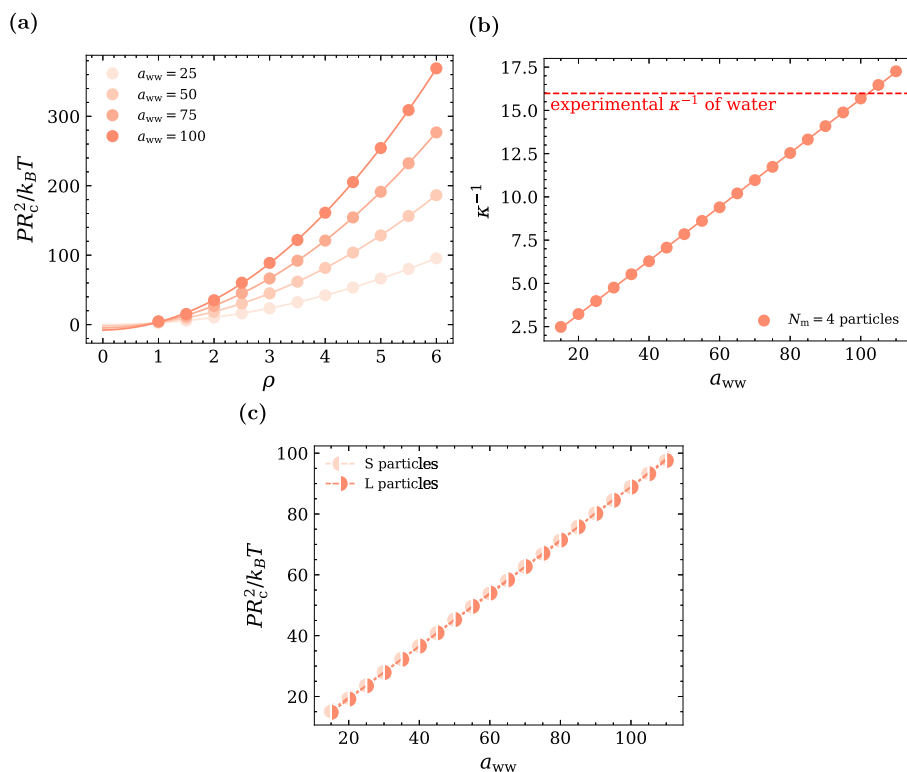
Figure 3. (a) The density-dependent pressure at different ρ and a_{ww}, at fixed $R_c = 1$, exhibits second order behavior for several values of the repulsion parameter a_{ww}. (b) The inverse compressibility κ^{-1} as a function of the repulsion parameter a_{ww} for large (L) particles. The inverse compressibility of water at room temperature is highlighted as a dashed red horizontal line. (c) The normalized pressure as a function of a_{ww} for small (S) and large (L) particles.

suggested by Groot.[29] Meanwhile, the parameter γ for S particles needs to be determined. We assume the diffusion coefficient as a function of γ obeys a Stokes−Einstein relation

$$D = \frac{k_B T}{\gamma'} = \frac{k_B T}{\alpha(\gamma + \gamma_0)} \tag{12}$$

where γ' $[\equiv \alpha(\gamma + \gamma_0) = 6\pi R\mu v]$ is the effective dissipation constant. This dissipation includes a sum of the naive (or internal) dissipation—viz., γ_0—arising from the coarse-graining of the degrees of freedom within the dissipative particle, and the dissipation—viz., γ—arising from the solvent and the other dissipative particles. The sum is rescaled by α to account for the corresponding renormalization of time due to coarse-graining. The observed dynamics in the simulations follows the Stokes−Einstein relation of eq 12, as shown in Figure 4(b). See also the discussion in Figure 4(a) regarding the role of γ' in Stokes law.

It is notable that MARTINI water has been seen to exhibit an erroneous freezing point and gives rise to incorrect diffusion.[31] On the contrary, by inducing a dissipative term, the DPD potential prevents freezing during the simulation. The time scale is also carefully calibrated in DPD by benchmarking the diffusion rate to the appropriate limits, as shown in Figure 4(a). This is a possible advantage for the use of DPD over MARTINI.

2.7. Interaction Parameters for Non-Water Particles. The parameter matrix a_{AB} represents the repulsion between arbitrary types of beads, A and B, and extends the character beyond the interaction between water beads, a_{ww}. It can be written relative to the water−water interaction as[11]

$$a_{AB} \equiv a_{ww} + \Delta a_{AB} \tag{13}$$

The relative interaction Δa_{AB} can be obtained from the Flory−Huggins parameter using the following linear relation[11]

$$\chi_{AB} = \lambda \Delta a_{AB} \tag{14}$$

where χ_{AB} is the Flory−Huggins parameter between beads A and B and λ is a fitting parameter independent of the nature of the interacting beads. In the present case, we further constrain the self-interactions of the coarse-grained beads A such that $a_{AA} = a_{ww}$—viz., $\Delta a_{AA} = 0$—in correspondence with the analogous choice in the MARTINI model in which the effective Lennard-Jones radius σ_{AA} is set to σ_{ww} for all such A.

According to Groot and Rabone, χ is linearly dependent on particle size.[23] This suggests that for different coarse-grained sizes—viz., volumes which are linear in the number of particles—the values in χ are in correspondence according to

$$\chi_{AB} = \overline{\chi}_{AB} N_m \tag{15}$$

where $\overline{\chi}$ is independent of particle size.

The parameter λ in eq 14 is determined by simulating the volume fraction of two types of DPD particles in a box, as reported in Figure 5(a). That is, we determine the values of χ given an initial Δa and then infer λ from the linear fit of eq 14 between them. We need only consider the special case in which there are only two types particles, say A_1 and A_2, that initially occupy the upper and lower halves of the simulation box, respectively, because that is enough to fix λ. The self-interacting parameters are the same within particle A_1 and particle A_2—that is, $a_{11} = a_{22}$—yet the interacting parameter between them differs by Δa which we sample to obtain a series of resulting data pairs with χ. After several thousand time steps
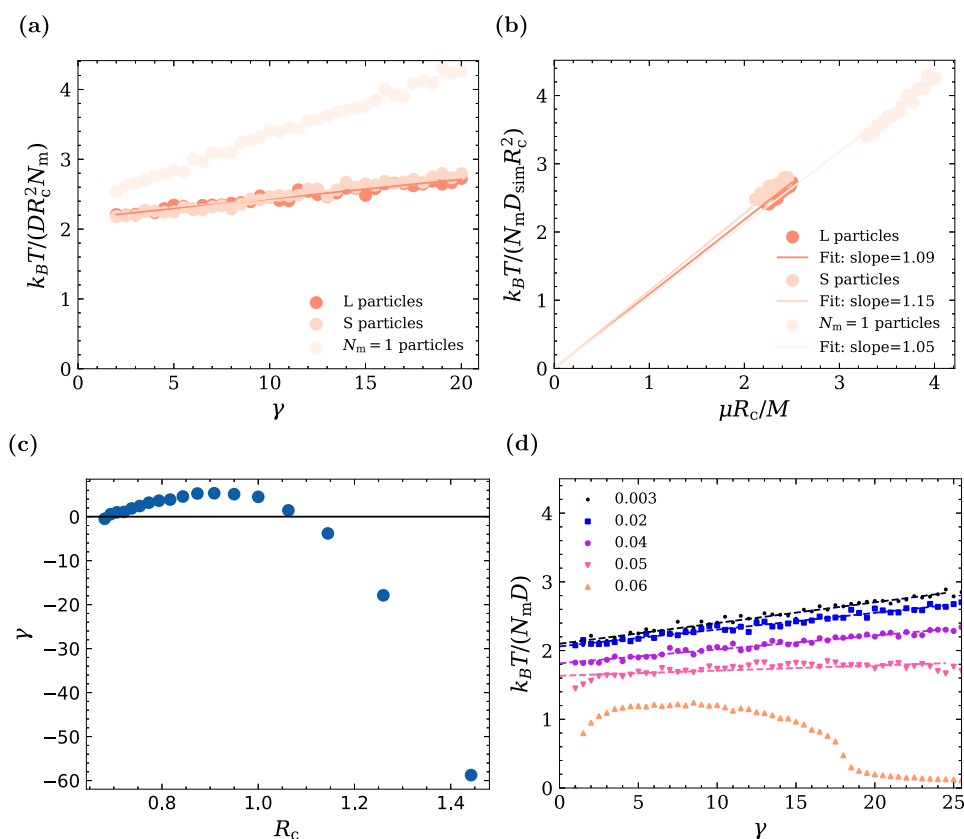
Figure 4. (a) Diffusion coefficient of coarse-grained water particles with varying size—small (S) and large (L) as defined in the text—when modeled within DPD for varying values of γ. (b) Diffusion coefficient as a function of viscosity—which depends on γ—for L particles, S particles, and $N_m = 1$ particles. (c) Corrected γ for various R_c. (d) Diffusion coefficient vs γ parameter, simulated with different time steps as labeled in the legend. The slopes are 0.0302 for DPD step 0.003, 0.0244 for DPD step 0.02, 0.0202 for DPD step 0.04, and 0.00746 for DPD step 0.05.

in the simulation for each Δa, both particle type A_1 and particle type A_2 diffuse and occupy a portion of the other side of the box, as shown in Figure 5(a). The majority volume fractions ϕ in each of the two phases can then be measured. The Flory–Huggins χ parameter is related to the degree of mixing through the expression

$$\chi = \frac{\ln[(1 - \phi)/\phi]}{1 - 2\phi} \tag{16}$$

where ϕ is the maximum volume fraction of either particle A_1 or A_2. This numerical experiment is repeated several times with different Δa; then λ is inferred from the linear fit calculated χ results across different Δa, as reported in Figure 5(b).

When obtaining the matrix of χ_{AB} between beads, Groot and Rabone[23] summarized several experimental χ that can be used to calculate a_{ij} from previous experimental reports.[32−39] Gao et al.[22] used this consensus set of parameters with small adjustments. In this work, we have found that, given the interaction matrix of another coarse-grained model (e.g., MARTINI), the χ_{AB} matrix can be calculated using the mean field equation

$$\bar{\chi}_{AB} N_m = \chi_{AB} = \frac{z}{k_B T}\left[\varepsilon_{AB} - \frac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB})\right] \tag{17}$$

where z is the effective number of nearest neighbors in the liquid. It is assigned to 8.85, such that $\bar{\chi} = 2.0$ for water and C_1, which is consistent with Groot and Gao et al.[22,23] The values of

ε are taken from the MARTINI model (Table 1 of Marrink et al.[2]).

After λ and the matrix of $\bar{\chi}_{AB}$ are calculated, the finalized parameter matrix a_{AB} can be calculated by a modified eq 13:

$$a_{ij} = a_{ww} + \bar{\chi}_{ij} N_m/\lambda_m \tag{18}$$

2.8. Determination of DPD Bonded Parameters. There are four types of bonded interactions in this model: bond, angle, dihedral, and improper. The bonded parameters in Gao's paper are simplified MARTINI bonded interactions, where the equilibrium distances/angles are preserved but different force constants are simplified into one single force constant. In this work, the force constants are directly converted from MARTINI for better accuracies. The constraint interactions in MARTINI are converted to bond interactions with a large force constant of 30,000 kJ/(nm^2· mol). All force constants are converted into DPD units.

The bond, angle, dihedral, and improper interactions are simulated with bond style set to "harmonic", angle style set to "cosine/squared", dihedral style set to "charmm", and improper style set to "harmonic" in LAMMPS, respectively.

2.9. HEME in CytC. Heme is a component of the CytC protein, but its interaction potential is not readily available in all coarse-graining representations. In particular, the original MARTINI force field does not contain heme, so the coarse-graining scheme and bonded interactions need to be defined somehow. One possibility lies in borrowing the structure for heme from the MARTINI model built by de Jong et al.[40] As
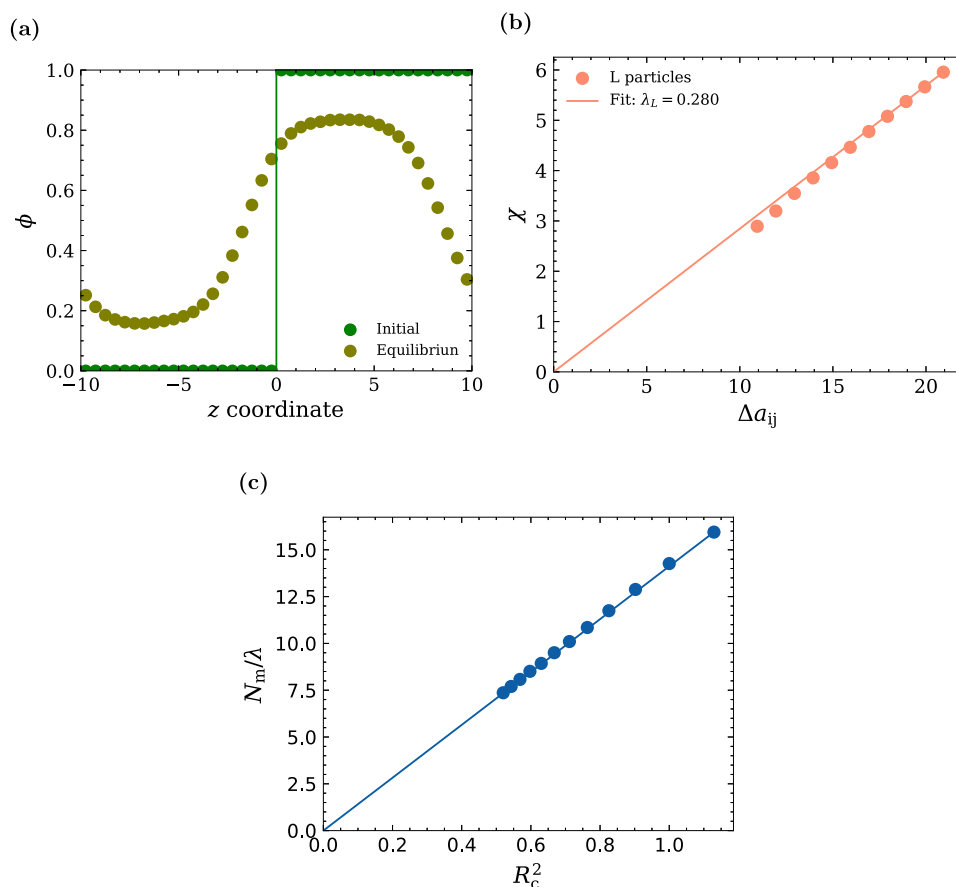
Figure 5. (a) The observed volume fraction ϕ vs z coordinate in the initial and equilibrium ensembles. (b) χ vs Δa_{ij} for L particles. The fits correspond to eq 14, resulting in the observed values of λ which are the slopes of the fitted lines. (c) N_m/λ plotted vs R_c^2. The slope of the fitted line is 0.204, and the intercept is 0.

the latter excludes particle–particle interactions within a heme, the resulting DPD potential will likewise exclude them.

The parameters of a covalent bond between heme c and MET80 of CytC, as well as angles that contain this bond, are not provided by the current CHARMM36m force field. For the all-atom simulations in this paper, these bonds and angles are consequently ignored. In the MARTINI model, the inclusion of the two bonds between the heme and MET80 leads to a merging of the heme and backbone into a single coarse-grained molecule. Such a merge is required by Gromacs, for example, if the bonded interaction is to be accounted for, but would result in a force field that could not distinguish these important components. Fortunately, in LAMMPS, we can stipulate an effective interaction—viz., bond—between DPD beads of different molecules. These bonds are modeled with the same force parameter used for the bonds in the protein backbone. The angle parameters resulting from the bonds are modeled as 500 kJ/mol and perpendicular to the heme plane in the DPD model.

2.10. Coarser-Grained Water Models. In order to access even larger systems (and longer time scales) using coarse-grained representations, the scaling factor denoting the ratio of fine-grained to coarse-grained particles must necessarily be larger than that in existing DPD models. Here, we introduce a scheme for generating multiscale coarse-grained models at higher scaling factors. We illustrate the scheme by increasing the number density to $\rho = 20$, and thereby accommodate larger

coarse-grained water particles. We keep the baseline cut off parameter at 1 for $N_m = 4$ particles for simplicity.

We denote coarse-grained water particles with $N_m = n$ water molecules as W_n; e.g., W_{20} and W_4 correspond to water particles with 20 and 4 water molecules, respectively. This notation should not be confused with L particles and S particles which refer to different coarse-grained groupings of atoms within a protein. In introducing coarse-grained water particles with larger N_m, we must necessarily rescale its associated unit length R_c^* to retain the correct number density. According to eq 6, the resulting unit length is

$$R_c^* = \sqrt[3]{20 \times 4 \times 30 \, \text{Å}^3} = 13.4 \, \text{Å} \tag{19}$$

The unit charge q_{DPD} (= 6.46) can be calculated from eq 10. Consequently, the parameter β of the Slater potential determined earlier in eq 9 now takes on the value

$$\beta_{20} = \frac{\beta_3 \times R_{c,N_m=3}}{R_{c,N_m=20}} = \frac{0.929 \times 6.46}{13.4} = 0.448 \tag{20}$$

The determination of the repulsion parameters a_{ij} and dissipation parameter γ is discussed in Section 3.2.

Modeling the interactions between particles with different volumes—that is, different scaling factors—requires specification through some kind of mixing rule. In both W_4 and W_{20} systems, all S particles with $N_m = 3$ are treated as L particles ($N_m = 4$) when they are not interacting with another S particle. This is the simplest choice as it entails no additional

parameters, and it is consistent with the treatment of the MARTINI2.2 force field. For the interaction between water particles ($N_m = 20$) and other protein particles ($N_m = 4$) in the system, we use the customary Lorentz−Berthelot rules. That is, we use the same LJ-form of the potential and set the effective interaction radius R_c for the mixed interaction as the geometric mean of the radii in the pure interactions. The repulsion parameter can be calculated using eq 25. The dissipation parameter γ between particles with different N_m is determined by the assumption that the densities are preserved through the equation

$$N_{m,\text{int}} = \frac{R_{c,\text{int}}^3}{R_{c,4}^3} N_{m,4} \tag{21}$$

For interactions between W_4 and W_{20}, the resulting effective number water molecules in the coarse-grained particle interaction is $N_m = 9.95$.

The W_{20} model is effectively fit to the pressure of the particles and the diffusion coefficient, as shown in Section 3.2. Like most coarse-grained water models, its accuracy is limited by the loss of information from the loss of the fine-grained degrees of freedom. This limitation is exacerbated in molecular models as the broader scale of the coarse-grained water model necessarily also obscures quantum effects. That is, in capturing the intricate behavior of fine-grained water, which encompasses quantum effects and more, with a coarse-grained classical framework like the MARTINI model, these effects are approximated at best up to a mean-field level. In the present work, we confirmed the use of W_{20} by comparing the structure of solvated proteins resulting from simulations using it and W_4.

2.11. Software Tools. The all-atom protein models and MARTINI protein models are built using Charmm-GUI.[27,41−43] All-atom simulations are performed under the CHARMM36m force field[44] using the NAMD 2.13[45] molecular dynamics engine. MARTINI simulations are performed under the MARTINI2.2 force field[3] using the Gromacs 5[46] particle dynamics engine. DPD particle mapping is the same as for the MARTINI mapping. The DPD simulations are integrated using LAMMPS.[47]

3. RESULTS AND DISCUSSION

3.1. Determined Parameters. The pressure versus different ρ is assumed to be a second order curve by Groot.[11] The fitted second order curves are shown in Figure 3(a), which are $P/k_BT = 2.70\rho^2 - 0.25\rho + 0.09$ at $a_{ww} = 25$, $P/k_BT = 5.27\rho^2 - 0.27\rho - 1.83$ at $a_{ww} = 50$, $P/k_BT = 7.73\rho^2 + 0.55\rho - 4.79$ at $a_{ww} = 75$, and $P/k_BT = 10.25\rho^2 + 1.27\rho - 7.88$ at $a_{ww} = 100$. The curves become steeper as a_{ww} increases because larger repulsion generates larger pressure in the simulation box. The fact that the curves do not intersect at the origin suggests that this approximation works for a limited range of ρ in agreement with earlier reports by Groot on the efficacy of the approach in the regime when $\rho > 2$.[11,48]

The simulated inverse compressibility

$$\kappa^{-1} = \frac{1}{N_m k_B T} \left(\frac{\partial P}{\partial \rho} \right)_{\text{simulation}} \tag{22}$$

is the slope of the pressure in ρ shown in Figure 3. The fitted linear equation for L particles at $\rho = 3$ and various a_{ww} is $\kappa^{-1} = 0.156 a_{ww} + 0.0769$.

From experiment, the dimensionless inverse compressibility of water is $\kappa^{-1} = 15.9835$.[11] The resulting values in the corresponding parameters for DPD water are $a_{ww} = 102$ for L particles. This result is close to Gao et al.[22] who set the parameter at $a_{ww} = 100$.

As noted above, the unit length of the L particles is taken to be the reference length in DPD units, and hence $R_c = 1$ and $a_{ww} = 100$ for them. To determine the parameter a_{ww} for other sized particles, we accounted for pressure as being proportional to the cross section of the coarse-grained particles. As shown in Figure 3(c), the pressures of the L particles and the S particles are identical after this correction. This leads to a corresponding rescaling of the water−water repulsion parameter for the particles with $R_c \neq 1$ that goes as

$$a_{ww}' = a_{ww} R_c^2 \tag{23}$$

The diffusion coefficient D and the viscosity μ are measured under different γ as reported in Figure 4. The behavior in the diffusion coefficient versus γ appears to violate the Stokes−Einstein relation, as the fitted line does not start from the origin in Figure 4(a). This limiting behavior arises because the DPD particles are intrinsically dissipated particles, and the non-zero intercept reflects the dissipation they experience when the mean-field solvent—represented by γ—disappears. Nevertheless, the viscosities of the DPD particles, shown in Figure 4(b), do follow the linearity of the Stokes−Einstein relation. In addition to L particles and S particles, we also measured the viscosity and diffusion rate for $N_m = 1$ particles to represent particles with the volume identical to water molecules. We find from Figure 4(b) that $k_B T/(N_m D_{\text{sim}} R_c^2) = 1.09\mu R_c/M$ for L particles, $k_B T/(N_m D_{\text{sim}} R_c^2) = 1.15\mu R_c/M$ for S particles, and $k_B T/(N_m D_{\text{sim}} R_c^2) = 1.05\mu R_c/M$ for $N_m = 1$ particles. These slopes are nearly identical and imply that probes diffusing in different-sized coarse-grained solvent particles have the same diffusion rate at a given viscosity. This confirms that such solvents are representing the same liquid at least with respect to this observable.

One should not be tempted to impose the same γ for different-sized particles because of the apparent similarity in the diffusion constants D for a given γ. As shown in Figure 4(a), the regression lines for L particles and S particles are indistinguishable, but the regression lines for $N_m = 1$ particles are different from the other two. Specifically, the dependence of the diffusion coefficient on the parameter γ seen in Figure 4(a) is $\frac{k_B T}{N_m D R_c^2} = 0.0280\gamma + 2.15$ for L particles, $\frac{k_B T}{N_m D R_c^2} = 0.0350\gamma + 2.09$ for S particles, and $\frac{k_B T}{N_m D R_c^2} = 0.0959\gamma + 2.37$ for $N_m = 1$ particles. One can select different γ on different lines to keep the diffusion rate D constant across different particle sizes. Such corrected γ parameters are shown in Figure 4(c). This correction also defines the upper (viz., ca. 1.1) and lower limit (viz., ca. 0.4) of particle sizes accessible to the model as γ cannot be less than zero. Particles with R_c greater than 1.2 can still be included by varying ρ accordingly.

The time scale of the simulation is determined by matching the diffusion constant of DPD water particles and the experimental water diffusion constant. The resulting effective time for a DPD step is[23]
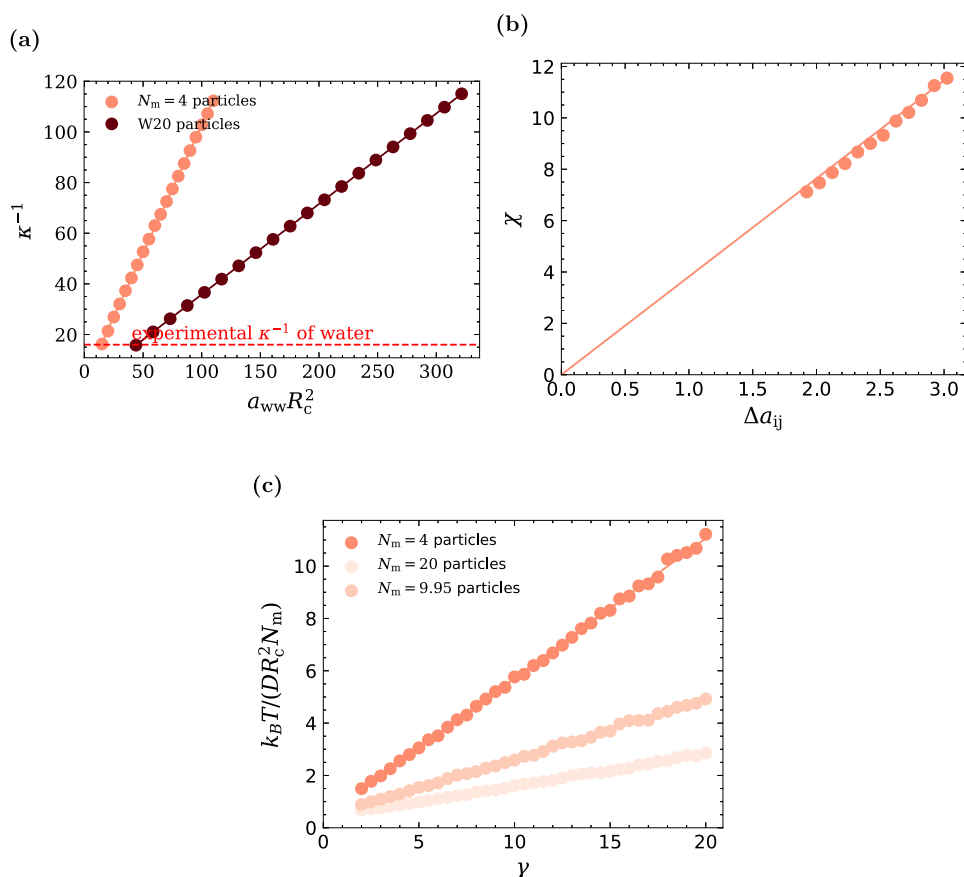
Figure 6. (a) The inverse compressibility of W_4 and W_{20}. (b) The calculated χ for $N_m = 4$ particles in W_{20} systems. (c) The calculated γ parameters for various sized particles in the $\rho = 20$ system.

$$\tau = \frac{N_m D_{sim} R_c^2}{D_{water}} = \frac{\frac{1}{0.0280 \times 4.5 + 2.15} \times (7.11 \times 10^{-8} \text{ cm})^2}{2.43 \times 10^{-5} \text{ cm}^2/\text{s}}$$

$$= 91.4 \text{ ps} \tag{24}$$

While the size of the integration step is formally arbitrary, in practice it must be set according to the accuracy of the integrator. As shown in Figure 4(d), integration of the DPD system with a DPD step of 0.06—as used by Groot in his simulations[11]—led the parameter γ to lose linearity. This indicates that the internal dissipation of the system would also become non-linear with γ for even larger integration DPD steps. In addition, we observed that, as the DPD integration step goes above 0.02, the regression lines of the diffusion coefficient versus γ begin to differ substantially. The time step is also limited by the stiffness of the spring constants for the bonds. In converging the integrator, we therefore found a need to use reduced time steps below even those earlier implemented by Groot.[11] Specifically, the Verlet integrator was implemented with a DPD step of 0.003 in the present simulations that is 20 times smaller. According to eq 24, each such step of the DPD simulation represents a time step of 0.003×91.4 ps = 0.274 ps. Although this value is relatively small, it still represents a speed up of more than a factor of 100 over typical implementations of all-atom simulations which use 2 fs time steps at best.[49]

The need for the smaller DPD step reported here does not imply that the results reported by Groot using a larger step are incorrect. In fact, the reason that Groot was able to use a much larger DPD step is due to the modified Verlet integrator he

employed.[11] Unfortunately, this integrator cannot be incorporated into our model easily because the current implementation in LAMMPS is not compatible with charged particles. As an updated integrator allowing for charges could significantly improve the performance, there is promise for the use of larger DPD steps in the future within an updated LAMMPS package.

The distributions of volume fractions ϕ observed for the initial and equilibrated distributions are shown in Figure 5(a). The fitting parameters λ for the slopes of the lines, $\chi = \lambda \Delta a$, in Figure 5(b) were found to be $\lambda = 0.280$ for L particles. This process for determining the fitting parameter λ was justified in Section 2.7. In turn, λ is used to calculate the non-self-interacting DPD parameters a_{ij}, as shown in eq 14. The experimental Flory–Huggins parameters χ are calculated using the mean field eq 17, based on the parameters in Table 1 of Marrink et al.[2] The parameter a_{ww} is determined using the method of Section 2.5. These parameters can then be used to complete pair-interaction parameter tables such as the ones reported in Table S1 of the Supporting Information.

We repeated the numerical experiment on different particle sizes, and the result is shown in Figure 5(c). Therein, N_m/λ changes linearly with R_c^2, and the intercept is zero. This leads to $\Delta a_{ij}' = \Delta a_{ij} R_c^2$, which, combined with $a_{ww}' = a_{ww} R_c^2$, gives

$$a_{ij}' = a_{ij} R_c^2 \tag{25}$$

This simple and elegant relationship implies that one only needs to use R_c^2 to correct for the repulsion parameter a_{ij} for different particle sizes. The process of obtaining repulsion parameters is consequently simplified to only the determi-
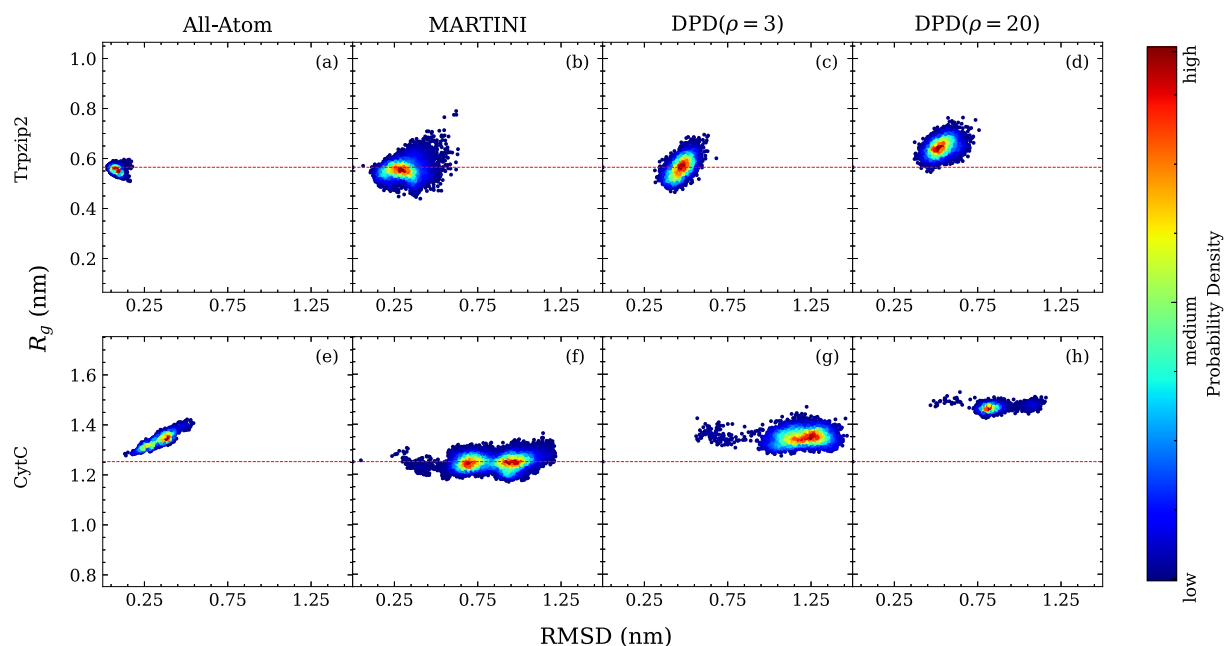
Figure 7. Probability density functions for Trpzip2 and CytC in the all-atom model, MARTINI model, and DPD model. The RMSD is on the x axis, the radius of gyration is on the y axis, and the radius of gyration of the PDB structure is represented by the red dashed line. (a) Trpzip2 all-atom model. (b) Trpzip2MARTINI model. (c) Trpzip2 DPD model with W_4. (d) Trpzip2 DPD model with W_{20}. (e) CytC all-atom model. (f) CytC MARTINI model. (g) CytC DPD model with W_4. (h) CytC DPD model with W_{20}.

nation of a_{ww} and λ. Though the process of calculating corrected a_{ij} is simple enough, the γ parameter still needs to be corrected to account for the correct diffusion rate. This is because plots of γ versus the reciprocal of diffusion rate typically do not exhibit a zero intercept; thus a simple linear correction does not arise. The process described above should be repeated to calculate γ for each particle size in the simulation. An example for generating a multiscale W_{20} model of coarse-grained particles with $N_m = 4$ particles for protein subunits and $\rho = 20$ representing coarse-grained 20 waters molecules is shown in Section 3.2.

3.2. Building the $\rho = 20$ Model. The water−water repulsion parameter of a W_{20} model is inferred from the reverse compressibility curves shown in Figure 6(a). As a result of the fit, the inverse compressibility is $\kappa^{-1} = 1.01a_{ww} + 1.77 = 15.9835$ using the water−water repulsion parameter of $a_{ww} = 14.1$. For W_{20}, a_{ww} must be corrected using eq 25. Using the fit for W_{20} in Figure 6(a), the resulting inverse compressibility is $\kappa^{-1} = 0.357a_{ww}R_c^2 + 0.113 = 14.8$, which is close to the experimental value of 15.9835. The χ parameter is then determined from the same numerical procedure as in Figure 5(a), and the slope of the regression line in Figure 6(b) is 3.75. With both a_{ww} and χ in hand, we can then calculate the repulsion parameter for all the non-water particles using the same energy table. This also leads to the a_{ij} parameters for arbitrarily sized particles, as noted in the previous section.

The γ parameter is calibrated to make the diffusion coefficient consistent across different particle sizes. The resulting regression lines shown in Figure 6(c) are $0.535\gamma + 0.366 = k_B T/(N_m D_{sim} R_c^2)$ for W_4, $0.122\gamma + 0.385 = k_B T/(N_m D_{sim} R_c^2)$ for W_{20}, and $0.225\gamma + 0.390 = k_B T/(N_m D_{sim} R_c^2)$ for $W_{9.94}$.

The γ parameter is then calculated from the regression lines with $\gamma = 4.5$ for W_4, which leads to $\gamma = 19.6$ for W_{20} and $\gamma = 10.6$ for $W_{9.94}$.

The time scale of the simulation can be determined using eq 24:[23]

$$\tau = \frac{\frac{1}{0.535} \times 4.5 + 0.366 \times (13.4 \times 10^{-8}\ \text{cm})^2}{2.43 \times 10^{-5}\ \text{cm}^2/\text{s}} = 264\ \text{ps} \tag{26}$$

The time scale of the W_{20} model is about 3 times faster than that for the $\rho = 3$ model. Combined with a reduced number of particles by further coarse-grained water, the W_{20} model consequently reduces the required simulation time significantly.

3.3. Model Validation on Proteins and Peptides. We validated the DPD protein model relative to the more fine-grained MARTINI and DPD models across a representative set of small peptides and proteins. The probability density functions for Trpzip2 (1LE1) and CytC (1AKK) over the domain of the radius of gyration (R_g), and the RMSD shown in Figure 7 support the main findings. The remaining density functions for Trpcage (1L2Y), Villin Headpiece (1VII), WW-Dgmain (1E0L), peripheral-binding-subunit (2PDD), Domain B Protein A (1BDC), GB1 (2J52), Trp Inhibitor (1H34), and Ubiquitin (2LD9) are included as Figures S1−S10 in the Supporting Information.

To compare across different models, we use metrics that depend on variables accessible to all three, namely, the MARTINI representation. Consequently, the structures in each frame of the all-atom trajectories are coarse-grained to the MARTINI model before calculating values. The MARTINI structures are ipso facto in the MARTINI representation, and the DPD particles are the same as the MARTINI representation. The radius of gyration of the native structures, obtained from the MARTINI projection of the PDB structure, is indicated with a red dashed line. For all the probability density functions reported in Figure 7, red represents a high probability and blue represents a low probability. Values below

that threshold are indicated in white. In this case, negligible values (shown in white) are defined by the lack of any captured frame in the simulated trajectories. These densities are not normalized, as they represent an arbitrary number of steps in the simulation. The all-atom simulation, MARTINI simulation, and DPD simulation are all simulated for different lengths of times as needed to reach equilibrium. The probability density functions in each of the all-atom and coarse-grained representations of two proteins are shown in Figure 7, and the others are available in the Supporting Information. The Root Mean Square Fluctuations (RMSFs) were also obtained and are shown in Figure 9. The underlying structures for the proteins whose RMSD and R_g match those of the corresponding averages of these distributions also serve to provide a comparison across the representations. Figure 8

(a) Trpcage (1.0x) **(b)** Trpzip2 (0.9x)

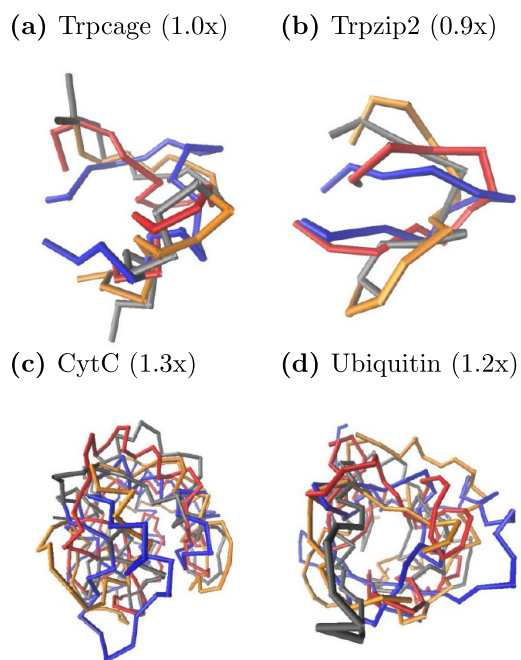**(c)** CytC (1.3x) **(d)** Ubiquitin (1.2x)

Figure 8. Comparison of selected structures of the proteins obtained from analysis at each of the scales of Figure 7, and all shown using a consistent coarse-grained representation. Structures are selected as the one whose RMSD and R_g matched the corresponding averages in the distributions shown in Figure 7. The coarse-grained (MARTINI) representations shown here illustrate the backbone for the selected structures from the all-atom model (blue), the MARTINI model (red), the W_4 DPD model (gray), and the W_{20} DPD model (yellow). The all-atom model is the reference for the alignment, and the displayed orientation is selected to best illustrate the clamped or helical structure of the protein. Projected 3D structures are zoomed according to the factor listed in parentheses, but exact distances across images are difficult to compare because aspect ratios also change the effective distance of bonds as they move in and out of the plane.

displays the alignment of these structures for the four proteins reported here—Trpcage, Trpzip2, CytC, and ubiquitin—with the coarse-grained structure of the all-atom model. The small peptides Trpcage and Trpzip2 show good agreement between all models, while larger proteins CytC and ubiquitin show more differences.

3.4. Accuracy of the DPD Protein Model. For all DPD proteins, as shown in Figure 7 and Figures S1−S10, the near-equilibrium dynamics of the $\rho = 3$ DPD protein model can be

seen to exhibit close agreement with the MARTINI results and good agreement with the all-atom results. Specifically, the shape and position of their probability density functions are similar. Not surprisingly, the probability density functions of DPD are closer to those of MARTINI than to those of the all-atom model. The W_{20} DPD protein models are less accurate than the W_4 DPD protein models. Their RMSDs are larger, and the radii of gyration are further away from the default positions. This is the expected behavior as the W_{20} DPD model has further coarse-grained the water solvent, so the result should be less accurate.

Nevertheless, the differences resulting from the two DPD models relative to the MARTINI model are smaller than the lengths of the structural differences between the DPD models. The length scale of the DPD models—that is, the diameter of the CG water particle—can be calculated from the relation, $30 \text{ Å}^3 N_m = 4\pi r^3/3$, where N_m is the number of water molecules per water particle and r is the radius of the water particle. The resulting diameters for the W_4 and W_{20} are 0.612 and 1.05 nm, respectively. The differences between the MARTINI and DPD models as seen in Figures S1−S10 lie well within the error bars of these length scales.

The RMSF results, as shown in Figure 9 and Figures S1−S10, also show similar behavior across all three models, with W_4 DPD and MARTINI more similar to each other. Surprisingly, the W_{20} DPD protein models in general have a RMSF closer to the all-atom result than the MARTINI model and the W_4 DPD model. This is probably due to the fact that
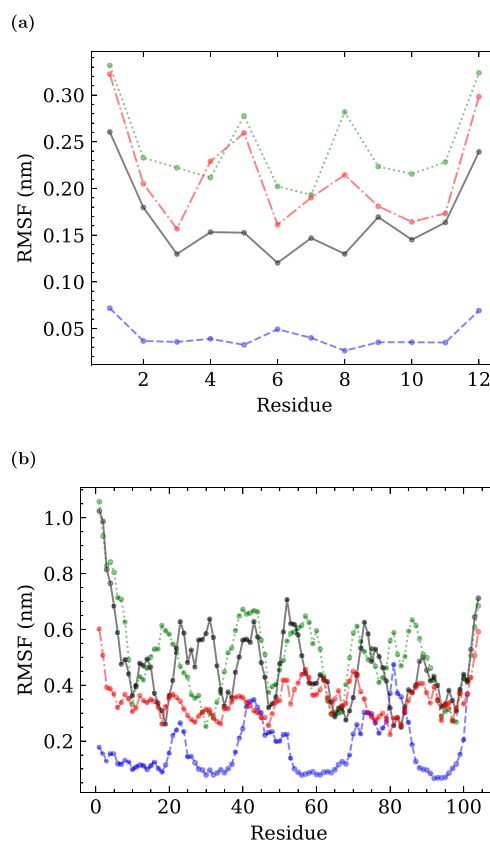

Figure 9. RMSF for Trpzip2 in panel (a) and CytC in panel (b) obtained from the all-atom (blue dashed), MARTINI (green dotted), W_4 DPD (black real), and W_{20} DPD (red dash-dot) models.

the increased number density also increased the accuracy of the protein model.

The all-atom model has lower RMSD and RMSF in comparison to the coarse-grained models, especially on smaller peptides. The larger deviations in the RMSFs seen for the MARTINI and DPD models relative to the all-atom models should not be a cause for alarm. The need for MARTINI to better stabilize protein structure was observed earlier, and is one of the issues that the program ElNeDyn was constructed to correct.[50] Nevertheless, addressing this discrepancy within our DPD model is an aim for future work. We also observe that the RMSF results for CytC are less consistent across all-atom, MARTINI, and DPD. We expect that improved parameters for the heme could increase the accuracy of all methods and resolve the relatively larger shifting in the radius of gyration in all-atom and W_{20} DPD in CytC.

4. CONCLUDING REMARKS

In this paper, we have consolidated an approach for building a DPD protein model based on the scheme suggested by Groot[11] in combination with the MARTINI[2] model. We benchmarked the accuracy of this scheme on several small peptides and larger proteins. We found that the resulting DPD proteins solvated in a similarly coarse-grained DPD water solvent have a comparable accuracy to the MARTINI protein models.

When a DPD protein was dissolved in a solvent of coarser-grained water—viz., a DPD model with particle radii increased by a factor of 5—the resulting protein structure was similar and tended to showed similar accuracy, thus providing additional computational savings. However, this coarser-grained solvent—referred to as the W_{20} model—in the text did lead to some anomalies. For example, salts which are known to be soluble were often seen to aggregate in the solution during the simulations. A possible resolution could involve the implementation of a polarizable water model[51] and a scheme to coarse-grain it within the framework of DPD.

Although DPD implementations are often empirical—that is, the parameters are usually fitted to match simulations with benchmarking data—we show here that DPD can also be built *bottom up* from finer-grained or bottom-up models (such as MARTINI). In so doing, we retain high accuracies and clear physical meaning for the nature of the system at the coarser DPD scale.

The scheme implemented here for building a DPD protein model from the MARTINI force field can also be applied to other reference coarse-grained potentials. That is, the mean field equation (eq 17) does not limit the type of potentials used as reference in obtaining the effective parameters at the DPD scale.

ASSOCIATED CONTENT

Data Availability Statement

The primary data that support the findings of this study are available in the Supporting Information. All data reported can be generated readily using the scripts provided online via Github at https://github.com/rxhernandez/MultiScaleDPD.

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c01868.

The complete set of parameters (Tables S1 and S2) and the complete results for small peptides and proteins (Figures S1−S10) (PDF)

AUTHOR INFORMATION

Corresponding Author

Rigoberto Hernandez − *Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218, United States;* ⓞ orcid.org/0000-0001-8526-7414; Email: r.hernandez@jhu.edu

Author

Yinhan Wang − *Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218, United States;* ⓞ orcid.org/0000-0002-5124-1942

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c01868

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Huang, J.; MacKerell, A. D., Jr. CHARMM36 All-Atom Additive Protein Force Field: Validation Based Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.

(2) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(3) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.

(4) Hadley, K. R.; McCabe, C. Coarse-Grained Molecular Models of Water: a Review. *Mol. Sim.* **2012**, *38*, 671−681.

(5) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801−6822.

(6) Nerenberg, P. S.; Head-Gordon, T. New Developments in Force Fields for Biomolecular Simulations. *Curr. Opin. Struct. Biol.* **2018**, *49*, 129−138.

(7) Zwier, M. C.; Chong, L. T. Reaching Biological Timescales with All-Atom Molecular Dynamics Simulations. *Curr. Opin. Pharmacol.* **2010**, *10*, 745−752.

(8) Ozboyaci, M.; Kokh, D. B.; Corni, S.; Wade, R. C. Modeling and Simulation of Protein−Surface Interactions: Achievements and Challenges. *Q. Rev. Biophys.* **2016**, *49*, e4.

(9) Saunders, M. G.; Voth, G. A. Coarse-Graining Methods for Computational Biology. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.

(10) de Jong, D. H.; Baoukina, S.; Ingolfsson, H. I.; Marrink, S. J. Martini Straight: Boosting Performance Using a Shorter Cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1−7.

(11) Groot, R. D.; Warren, P. B. Dissipative Particle Dynamics: Bridging the Gap Between Atomistic and Mesoscopic Simulation. *J. Chem. Phys.* **1997**, *107*, 4423−4435.

(12) Chong, G.; Laudadio, E. D.; Wu, M.; Murphy, C. J.; Hamers, R. J.; Hernandez, R. Density, Structure, and Stability of citrate^{3-}and H_2citrate$^-$on Bare and Coated Gold Nanoparticles. *J. Phys. Chem. C* **2018**, *122*, 28393−28404.

(13) Murphy, C. J.; Vartanian, A. M.; Geiger, F. M.; Hamers, R. J.; Pedersen, J.; Cui, Q.; Haynes, C. L.; Carlson, E. E.; Hernandez, R.; Klaper, R. D.; et al. Biological Responses to Engineered Nanomaterials: Needs for the Next Decade. *ACS Cent. Sci.* **2015**, *1*, 117.

(14) Cui, Q.; Hernandez, R.; Mason, S. E.; Frauenheim, T.; Pedersen, J. A.; Geiger, F. Sustainable Nanotechnology: Opportunities and Challenges for Theoretical/Computational Studies. *J. Phys. Chem. B* **2016**, *120*, 7297−7306.

(15) Gupta, C.; Sarkar, D.; Tieleman, D. P.; Singharoy, A. The ugly, bad, and good stories of large-scale biomolecular simulations. *Curr. Opin. Struct. Biol.* **2022**, *73*, No. 102338.

(16) Casalini, T.; Limongelli, V.; Schmutz, M.; Som, C.; Jordan, O.; Wick, P.; Borchard, G.; Perale, G. Molecular modeling for nanomaterial−biology interactions: Opportunities, challenges, and perspectives. *Front. bioeng. biotechnol.* **2019**, *7*, 268.

(17) Melby, E. S.; Allen, C.; Foreman-Ortiz, I.; Caudill, E.; Kuech, T. R.; Vartanian, A. M.; Zhang, X.; Murphy, C. J.; Hernandez, R.; Pedersen, J. A. Peripheral Membrane Proteins Facilitate Nanoparticle Binding at Lipid Bilayer Interfaces. *Langmuir* **2018**, *34*, 10793.

(18) Tollefson, E. J.; Allen, C. R.; Chong, G.; Zhang, X.; Rozanov, N. D.; Bautista, A.; Cerda, J. J.; Pedersen, J. A.; Murphy, C. J.; Carlson, E. E.; et al. Preferential Binding of Cytochrome C to Anionic Ligand-Coated Gold Nanoparticles: A Complementary Computational and Experimental Approach. *ACS Nano* **2019**, *13*, 6856−6866.

(19) Hoogerbrugge, P.; Koelman, J. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. *Europhys. Lett.* **1992**, *19*, 155−160.

(20) Peter, E. K.; Lykov, K.; Pivkin, I. V. A Polarizable Coarse-Grained Protein Model for Dissipative Particle Dynamics. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24452−24461.

(21) Vishnyakov, A.; Talaga, D. S.; Neimark, A. V. DPD Simulation of Protein Conformations: From α-Helices to β-Structures. *J. Phys. Chem. Lett.* **2012**, *3*, 3081−3087.

(22) Wan, M.; Gao, L.; Fang, W. Implicit-Solvent Dissipative Particle Dynamics Force Field Based on a Four-to-One Coarse-Grained Mapping Scheme. *PLoS One* **2018**, *13*, No. e0198049.

(23) Groot, R. D.; Rabone, K. L. Mesoscopic Simulation of Cell Membrane Damage, Morphology Change and Rupture by Nonionic Surfactants. *Biophys. J.* **2001**, *81*, 725−736.

(24) González-Melchor, M.; Mayoral, E.; Velázquez, M. E.; Alejandre, J. Electrostatic Interactions in Dissipative Particle Dynamics Using the Ewald Sums. *J. Chem. Phys.* **2006**, *125*, 224107.

(25) Liu, M. B.; Liu, G. R.; Zhou, L. W.; Chang, J. Z. Dissipative Particle Dynamics (DPD): An Overview and Recent Developments. *Arch. Comput. Methods Eng.* **2015**, *22*, 529−556.

(26) Shepherd, T. D.; Hernandez, R. Chemical Reaction Dynamics with Stochastic Potentials Beyond the High-Friction Limit. *J. Chem. Phys.* **2001**, *115*, 2430−2438.

(27) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 4486−4494.

(28) Kacar, G.; Peters, E. A. J. F.; de With, G. A Generalized Method for Parameterization of Dissipative Particle Dynamics for Variable Bead Volumes. *Europhys. Lett.* **2013**, *102*, 40009.

(29) Groot, R. D. Electrostatic Interactions in Dissipative Particle Dynamics−Simulation of Polyelectrolytes and Anionic Surfactants. *J. Chem. Phys.* **2003**, *118*, 11265−11277.

(30) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: New York, 1987.

(31) Jarin, Z.; Newhouse, J.; Voth, G. A. Coarse-Grained Force Fields From the Perspective of Statistical Mechanics: Better Understanding of the Origins of a MARTINI Hangover. *J. Chem. Theory Comput.* **2021**, *17*, 1170−1180.

(32) Shaw, D. G. *Hydrocarbons with water and seawater, part I and II*; Solubility Data Series; Pergamon Press: Oxford, UK, 1989; Vol. 37, pp 317−319.

(33) Barneveld, P. A.; Scheutjens, J. M. H. M.; Lyklema, J. Bending moduli and spontaneous curvature. 1. Bilayers and monolayers of pure and mixed nonionic surfactants. *Langmuir* **1992**, *8*, 3122−3130.

(34) Leermakers, F. A. M.; Scheutjens, J. M. H. M. Statistical thermodynamics of association colloids. I. Lipid bilayer membranes. *J. Chem. Phys.* **1988**, *89*, 3264−3274.

(35) Wijmans, C. M.; Smit, B.; Groot, R. D. Phase behavior of monomeric mixtures and polymer solutions with soft interaction potentials. *J. Chem. Phys.* **2001**, *114*, 7644−7654.

(36) Cohen Stuart, M. A.; Waajen, F. H. W. H.; Cosgrove, T.; Vincent, B.; Crowley, T. L. Hydrodynamic thickness of adsorbed polymer layers. *Macromolecules* **1984**, *17*, 1825−1830.

(37) Barneveld, P. A. In *The Bending Elasticity of Surfactant Monolayers and Bilayers and Its Effect on the Swelling of Free Liquid Films*; Landbouwuniversiteit te Wageningen, 1991, https://edepot. wur.nl/206447 (accessed March 27, 2024).

(38) Saeki, S.; Kuwahara, N.; Nakata, M.; Kaneko, M. Upper and lower critical solution temperatures in poly (ethylene glycol) solutions. *Polymer* **1976**, *17*, 685−689.

(39) Lu, J. R.; Li, Z. X.; Thomas, R. K.; Staples, E. J.; Tucker, I.; Penfold, J. Neutron reflection from a layer of monododecyl hexaethylene glycol adsorbed at the air-liquid interface: the configuration of the ethylene glycol chain. *J. Phys. Chem.* **1993**, *97*, 8012−8020.

(40) de Jong, D. H.; Liguori, N.; van den Berg, T.; Arnarez, C.; Periole, X.; Marrink, S. J. Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex. *J. Phys. Chem. B* **2015**, *119*, 7791−7803.

(41) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.

(42) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405−413.

(43) Hsu, P.-C.; Bruininks, B. M. H.; Jefferies, D.; Cesar Telles de Souza, P.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y.; Khalid, S.; Im, W. CHARMM-GUI Martini Maker for Modeling and Simulation of Complex Bacterial Membranes with Lipopolysaccharides. *J. Comput. Chem.* **2017**, *38*, 2354−2363.

(44) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71−73.

(45) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, No. 044130.

(46) 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*, 19−25.

(47) Plimpton, S. J. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(48) Kumar, A.; Asako, Y.; Abu-Nada, E.; Krafczyk, M.; Faghri, M. From Dissipative Particle Dynamics Scales to Physical Scales: A Coarse-Graining Study for Water Flow in Microchannel. *Microfluid Nanofluidics* **2009**, *7*, 467.

(49) Kim, S. Issues on the Choice of a Proper Time Step in Molecular Dynamics. *Phys. Procedia* **2014**, *53*, 60−62. 26th Annual CSP Workshop on "Recent Developments in Computer Simulation Studies in Condensed Matter Physics", CSP 2013.

(50) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531−2543.

(51) Wu, Z.; Cui, Q.; Yethiraj, A. A new coarse-grained model for water: the importance of electrostatic interactions. *J. Phys. Chem. B* **2010**, *114*, 10524−10529.