Systematic Bottom-Up Molecular Coarse-Graining via Force Matching using Anisotropic Particles

Huong T.L. Nguyen and David M. Huang*

Department of Chemistry, School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

E-mail: david.huang@adelaide.edu.au

Abstract

We derive a systematic and general method for parametrizing coarse-grained molecular models consisting of anisotropic particles from fine-grained (e.g. all-atom) models for condensed-phase molecular dynamics simulations. The method, which we call anisotropic force-matching coarse-graining (AFM-CG), is based on rigorous statistical mechanical principles, enforcing consistency between the coarse-grained and fine-grained phase-space distributions to derive equations for the coarse-grained forces, masses, and moments of inertia in terms of properties of a condensed-phase fine-grained system. We verify the accuracy and efficiency of the method by coarse-graining liquid-state systems of two different anisotropic organic molecules, benzene and perylene, and show that the parametrized coarse-grained models more accurately describe properties of these systems than previous anisotropic coarse-grained models parametrized using other methods that do not account for finite-temperature and many-body effects on the condensed-phase coarse-grained interactions. The AFM-CG method will be useful for developing accurate and efficient dynamical simulation models of condensed-phase

systems of molecules consisting of large, rigid, anisotropic fragments, such as nucleic acids, liquid crystals, and organic semiconductors.

1 Introduction

Molecular dynamics (MD) simulations enable the study of molecular-scale behavior that is often challenging or impossible to observe experimentally. While all-atom simulations can accurately describe the microscopic features of molecular interactions, they are often unable to access the time and length scales needed to model many phenomena of fundamental and technological interest, particularly for macromolecular systems commonly encountered in biology and materials science. A solution to this problem is to use coarse-grained (CG) molecular simulations,[1,2] in which groups of atoms are mapped onto a smaller set of CG particles, which can speed up simulations by orders of magnitude. CG simulations using isotropic (spherical) particles have been used to study large biomolecular[3,4] and synthetic polymer[5–7] systems with reasonable accuracy. Isotropic-particle CG models predominate due to their simplicity and widely developed CG parametrization methods.[2,8–12]

However, isotropic CG models may not always provide the most accurate, efficient, or physically transparent CG description of system properties. For molecules consisting of large, rigid, anisotropic fragments, which includes important classes of molecules such as nucleic acids,[13,14] liquid crystals,[15] and organic semiconductors[16–20] (due to their extended π-conjugation), anisotropic-particle CG models can more accurately represent the shape and interaction anisotropy using a smaller number of CG particles. This potentially increases simulation speed compared with the use of isotropic CG particles for a given level of accuracy. The smaller number of CG variables for an anisotropic CG model compared with an isotropic CG model of equivalent accuracy or fidelity can also enable more straightforward interpretation or prediction of system properties, particularly those related to the orientation of anisotropic molecules. This is the case for organic semiconductors, for which molecu-

lar ordering and alignment at the material interfaces play crucial roles in determining the performance of organic electronic devices such as solar cells, light-emitting diodes and transistors.[17,21–23] Implementation of MD simulations of anisotropic CG models is facilitated by well-established analytical non-bonded pair potentials, with the most commonly used being the Gay-Berne[24] and RE-squared[25] potentials for ellipsoidal particles. Advantages of using anisotropic CG models have been demonstrated in recent approaches to simulate organic material systems,[26–29] in which both large-scale conformational properties and small-scale anisotropic molecular arrangements were obtained.

Despite the aforementioned advantages, the use of anisotropic CG models to accurately simulate behavior of condensed-phased systems is limited by a lack of systematic parametrization methods with a rigorous theoretical basis. Some parametrizations of anisotropic CG models have taken a "top-down" approach, in which interactions are tuned to match relevant macroscopic thermodynamic and mechanical properties such as melting temperatures and elastic moduli.[14] For those anisotropic CG models derived from a fine-grained (FG) model such as an all-atom model using a "bottom-up" approach, most have been parametrized by matching the interaction potential between CG particles to the FG potential energy for specific configurations of pairs of molecules or molecular fragments,[13,25–27,30,31] which neglects finite-temperature and many-body effects that can significantly influence the effective interactions between particles in the condensed phase under thermodynamically relevant conditions.

CG interactions have also been fit to the orientation-dependent potential-of-mean-force (PMF) between molecule pairs in condensed-phase FG simulations,[32,33] which is equivalent to the iterative Boltzmann inversion (IBI) approach, in which CG interactions are iteratively tuned to match structural distribution functions in a FG simulation, in this case applied to pair correlation functions.[8] Although this approach has a basis in statistical mechanics theory, it uses reduced information about the condensed-phase system rather than the full many-body CG PMF,[2] and so is not readily systematically improved to account for higher

order correlations.

Tripathy et al.[34] used the total instantaneous force between pairs of molecules extracted from condensed-phase all-atom simulations to fit the parameters in an anisotropic CG pair potential for several different polyaromatic hydrocarbons. Tanis et al.[28] used a similar force-matching approach, but instead used average pair forces and torques from condensed-phase all-atom simulations as a function of the pair configuration to fit analytical CG pair potentials. Although demonstrated to accurately reproduce properties of the FG all-atom models for specific systems, these methods do not rigorously account for many-body effects on the CG PMF, even though the molecule pairs used in the parametrization were extracted from condensed-phase simulations. They also rely on CG potentials that can be decomposed into pair interactions, limiting their ability to be generalized to more complex effective interactions.

In this work, we derive a systematic and general CG parametrization method for anisotropic particles with a strong theoretical basis in rigorous statistical mechanical principles and develop a computer algorithm for implementing it. The method, which we have called anisotropic force-matching coarse-graining (AFM-CG), is a generalization to anisotropic CG particles of the multi-scale coarse-graining (MS-CG) method for isotropic CG particles, which uses force matching to relate interactions in the underlying FG model to those in the CG model. Unlike existing anisotropic force-matching methods,[28,34] the AFM-CG method is developed to approximate the FG PMF in a finite-temperature, many-body condensed-phase system without assuming pair interactions between CG particles and is not limited to specific functional forms of the CG potentials, forces, or torques. As proof-of-principle, the method is applied to systems of two anisotropic molecules, benzene and perylene. A number of structural, thermodynamic and dynamical properties are compared between the CG and FG models to verify the accuracy of the CG models and AFM-CG method.

2 Theory

The theoretical background and computational implementation of the multi-scale coarse-graining (MS-CG) method that uses force matching to parametrize interactions between isotropic CG particles has been introduced previously.[9,35,36] The method provides a rigorous statistical-mechanical framework for determining the interactions between CG particles so that the equilibrium phase-space distribution of the CG system is consistent with that of the atomically detailed FG system from which it is derived. Here we generalize the MS-CG method to CG models consisting of anisotropic particles, in which the system configuration is defined not only by particle positions, as in the case of isotropic (spherical) particles, but also by particle orientations.

Consider a FG system comprising n spherical particles (atoms), with particle positions and momenta $\boldsymbol{r}^n \equiv (\boldsymbol{r}_1, ..., \boldsymbol{r}_n)$ and $\boldsymbol{p}^n \equiv (\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, respectively, for which the potential energy as a function of FG particle coordinates is $u(\boldsymbol{r}^n)$. This FG system is to be mapped onto a CG representation comprising N anisotropic sites, where $N < n$, whose configuration is defined by both the CG particle positions $\boldsymbol{R}^N \equiv (\boldsymbol{R}_1, ..., \boldsymbol{R}_N)$ and their orientations $\boldsymbol{\Omega}^N \equiv (\boldsymbol{\Omega}_1, ..., \boldsymbol{\Omega}_N)$. The conjugate momenta to these CG position and angular coordinates are the linear momenta $\boldsymbol{P}^N \equiv (\boldsymbol{P}_1, ..., \boldsymbol{P}_N)$ and angular momenta $\boldsymbol{L}^N \equiv (\boldsymbol{L}_1, ..., \boldsymbol{L}_N)$, respectively. (We use lower-case symbols for FG variables and the corresponding upper-case symbols for the corresponding CG variables.) The Hamiltonians of the FG and CG systems are

$$H_{\mathrm{FG}}\left(\boldsymbol{r}^n, \boldsymbol{p}^n\right) = \sum_{i=1}^{n} \frac{\boldsymbol{p}_i^2}{2m_i} + u\left(\boldsymbol{r}^n\right) \tag{1}$$

and

$$H_{\mathrm{CG}}\left(\boldsymbol{R}^N, \boldsymbol{\Omega}^N, \boldsymbol{P}^N, \boldsymbol{L}^N\right) = \sum_{I=1}^{N} \left(\frac{\boldsymbol{P}_I^2}{2M_I} + \frac{\boldsymbol{L}_I^{\mathrm{T}} \mathbb{I}_I^{-1} \boldsymbol{L}_I}{2} \right) + U\left(\boldsymbol{R}^N, \boldsymbol{\Omega}^N\right) \tag{2}$$

respectively, where m_i is the mass of FG particle i, M_I and $\mathbb{I}_I$ are the mass and inertia tensor, respectively, of CG particle I, $U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$ is the CG potential energy as a function of CG coordinates, and the superscript "T" denotes a matrix transpose. The first and second

terms in the sum in eq (2) correspond to the translational and rotational kinetic energies, respectively, of the CG particle.[37] The aims of the coarse-graining procedure are (1) to define a physically appropriate mapping from FG to CG coordinates and (2) to derive expressions for the CG masses, inertia tensors, and potential energy function that best represent the FG system.

In the MS-CG method for a dynamical system of isotropic CG particles,[35] the CG mapping and potential energy function are chosen so that the equilibrium phase-space distribution of the CG system is consistent with that of the FG system from which it is derived, since this choice guarantees that any equilibrium average property of the CG system will be consistent with that of the FG system. We impose the same condition here for anisotropic CG particles and assume that the equilibrium phase-space probability distribution for both the FG and CG systems can be factored into the product of a configurational probability distribution and a momentum probability distribution. From the FG Hamiltonian in eq (1), the normalized equilibrium configurational distribution of the FG system in the canonical ensemble is

$$\mathbb{P}_{\mathrm{FG,c}}\left(\boldsymbol{r}^n\right) = \frac{1}{Z_{\mathrm{FG}}} \exp\left[-\frac{u\left(\boldsymbol{r}^n\right)}{k_{\mathrm{B}}T}\right] \tag{3}$$

where $Z_{\mathrm{FG}} \equiv \int \mathrm{d}\boldsymbol{r}^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right)$ is the configurational partition function of the FG system, T is the temperature, and k_{B} is the Boltzmann constant. The normalized equilibrium momentum distribution of the FG system is

$$\mathbb{P}_{\mathrm{FG,m}}\left(\boldsymbol{p}^n\right) = \frac{1}{(2\pi k_{\mathrm{B}}T)^{3n/2}} \prod_{i=1}^{n} \frac{1}{m_i^{3/2}} \exp\left(-\frac{\boldsymbol{p}_i^2}{2m_i k_{\mathrm{B}}T}\right) \tag{4}$$

Similarly, from the CG Hamiltonian in eq (2), the normalized equilibrium configurational distribution of the CG system is

$$\mathbb{P}_{\mathrm{CG,c}}\left(\boldsymbol{R}^N, \boldsymbol{\Omega}^N\right) = \frac{1}{Z_{\mathrm{CG}}} \exp\left[-\frac{U\left(\boldsymbol{R}^N, \boldsymbol{\Omega}^N\right)}{k_{\mathrm{B}}T}\right] \tag{5}$$

where $Z_{\mathrm{CG}} \equiv \int \mathrm{d}\boldsymbol{R}^N \int \mathrm{d}\boldsymbol{\Omega}^N \exp\left(-\frac{U(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)}{k_{\mathrm{B}}T}\right)$ is the configurational partition function of the CG system, and the normalized equilibrium momentum distribution of the CG system is

$$\mathbb{P}_{\mathrm{CG,m}}(\boldsymbol{P}^N,\boldsymbol{L}^N) = \frac{1}{(2\pi k_{\mathrm{B}}T)^{3N}} \prod_{I=1}^{N} \frac{\exp\left[-\frac{1}{2k_{\mathrm{B}}T}\left(\frac{\boldsymbol{P}_I^2}{M_I} + \boldsymbol{L}_I^{\mathrm{T}}\mathbb{I}_I^{-1}\boldsymbol{L}_I\right)\right]}{M_I^{3/2}\det\left(\mathbb{I}_I\right)^{1/2}} \tag{6}$$

where $\det(\mathbb{I}_I)$ is the determinant of $\mathbb{I}_I$. The normalization factors in eqs (4) and (6) are derived in the Supporting Information (SI), section S2.

2.1 Configuration-Space Consistency Conditions and Coarse-Grained Forces

The FG particle positions $\boldsymbol{r}^n$ are mapped onto anisotropic CG particle positions $\boldsymbol{R}^N$ and orientations $\boldsymbol{\Omega}^N$ via the mapping operators

$$\mathbb{M}_{\mathrm{R}}^N\left(\boldsymbol{r}^n\right) \equiv \left(\mathbb{M}_{\mathrm{R}1}(\boldsymbol{r}^n), \mathbb{M}_{\mathrm{R}2}(\boldsymbol{r}^n), \ldots, \mathbb{M}_{\mathrm{R}N}(\boldsymbol{r}^n)\right) = \boldsymbol{R}^N \tag{7}$$

and

$$\mathbb{M}_{\Omega}^N\left(\boldsymbol{r}^n\right) \equiv \left(\mathbb{M}_{\Omega 1}(\boldsymbol{r}^n), \mathbb{M}_{\Omega 2}(\boldsymbol{r}^n), \ldots, \mathbb{M}_{\Omega N}(\boldsymbol{r}^n)\right) = \boldsymbol{\Omega}^N \tag{8}$$

respectively. These mapping operators are defined so that the CG coordinates have well-defined physical definitions in terms of the FG coordinates. In the derivation of the MS-CG method for isotropic CG particles,[35] it was shown that consistency between FG and CG models in momentum space requires that no FG particle contribute to more than one CG particle. Thus, we impose the same condition and, more specifically, represent the position and orientation of each CG particle as the center-of mass and principal axes of inertia of the group of FG particles that are mapped onto it. We define N non-intersecting subsets, $\zeta_1, \zeta_2, \ldots, \zeta_N$, of the FG particle indices such that ζ_I contains the indices of FG particles that

are mapped onto CG site I. The position of CG particle I is defined as

$$\boldsymbol{R}_I = \mathbb{M}_{\mathrm{R}I}(\boldsymbol{r}^n) = \frac{\sum_{i\in\zeta_I} m_i \boldsymbol{r}_i}{\sum_{i\in\zeta_I} m_i} \tag{9}$$

The orientation of the CG particle is represented by its three principal axes of inertia, determined from the inertia tensor relative to the center-of-mass of the group of FG particles mapped onto it,

$$\mathbb{I}_{\mathrm{FG},I} = \sum_{i\in\zeta_I} m_i \left(\|\Delta\boldsymbol{r}_i\|^2 \boldsymbol{E} - \Delta\boldsymbol{r}_i \Delta\boldsymbol{r}_i^{\mathrm{T}} \right) \tag{10}$$

where $\Delta\boldsymbol{r}_i = \boldsymbol{r}_i - \boldsymbol{R}_I$ is the position of FG particle i relative to the center-of-mass (CG particle position) and $\boldsymbol{E}$ is the 3×3 identity matrix. The principal axes are the column vectors of the invertible matrix that diagonalizes the inertia tensor. Figure 1 illustrates the mapping from FG to CG coordinates for the example of two benzene molecules.

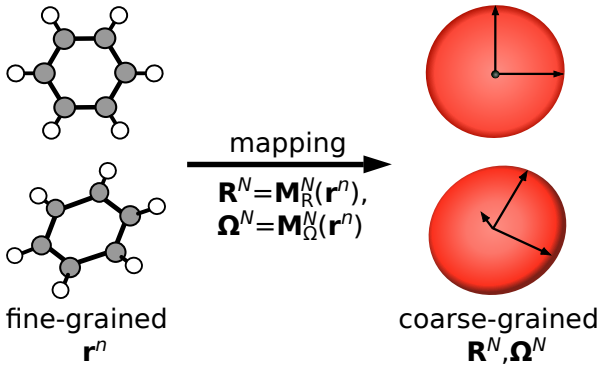

Figure 1: Mapping of atomistic FG positions to CG position and orientation coordinates for two benzene molecules coarse-grained to disk-shaped particles. The black arrows on the CG particles indicate the principal axes of inertia that define the CG orientations.

Given this mapping from FG to CG configurations, enforcing consistency between FG and CG equilibrium configurational distributions requires the probability of configuration $(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$ in the CG system to match the probability of the FG system being in a configuration that maps on to this CG configuration. This condition is

$$\mathbb{P}_{\mathrm{CG,c}}(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) = \int \mathrm{d}\boldsymbol{r}^n \mathbb{P}_{\mathrm{FG,c}}(\boldsymbol{r}^n) \delta\left(\mathbb{M}_{\mathrm{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N\right) \delta\left(\mathbb{M}_{\Omega}^N(\boldsymbol{r}^n) - \boldsymbol{\Omega}^N\right) \tag{11}$$

where the Dirac delta functions constrain the mapped FG configuration to match the CG configuration. Inserting eqs (3) and (5) into eq (11) gives the CG potential energy function as

$$U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) = -k_{\mathrm{B}}T \ln Z(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) - k_{\mathrm{B}}T \ln \left(\frac{Z_{\mathrm{CG}}}{Z_{\mathrm{FG}}} \right) \tag{12}$$

where

$$Z(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) = \int \mathrm{d}\boldsymbol{r}^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T} \right) \delta\left(\mathbb{M}_{\mathrm{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N \right) \delta\left(\mathbb{M}_{\Omega}^N(\boldsymbol{r}^n) - \boldsymbol{\Omega}^N \right) \tag{13}$$

Using eqs (12) and (13), the force on CG site I is

$$\begin{aligned} \boldsymbol{F}_I(\boldsymbol{R}^N, \boldsymbol{\Omega}^{\boldsymbol{N}}) &= -\frac{\partial U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)}{\partial \boldsymbol{R}_I} \\ &= \frac{k_{\mathrm{B}}T}{Z(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)} \int \mathrm{d}\boldsymbol{r}^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T} \right) \delta\left(\mathbb{M}_{\Omega}^N(\boldsymbol{r}^n) - \boldsymbol{\Omega}^N \right) \\ &\quad \times \prod_{J \neq I}^{N} \delta\left(\mathbb{M}_{\mathrm{R}J}(\boldsymbol{r}^n) - \boldsymbol{R}_J \right) \frac{\partial}{\partial \boldsymbol{R}_I} \delta\left(\mathbb{M}_{\mathrm{R}I}(\boldsymbol{r}^n) - \boldsymbol{R}_I \right) \end{aligned} \tag{14}$$

Using the definition of the linear mapping $\mathbb{M}_{\mathrm{R}I}(\boldsymbol{r}^n)$ of FG to CG positions (eq (9)) and applying the chain rule, the delta function derivative in eq (14) can be recast in terms of FG coordinates as[35]

$$\frac{\partial}{\partial \boldsymbol{R}_I} \delta\left(\mathbb{M}_{\mathrm{R}I}(\boldsymbol{r}^n) - \boldsymbol{R}_I \right) = -\sum_{i \in \zeta_I} \frac{\partial}{\partial \boldsymbol{r}_i} \delta\left(\mathbb{M}_{\mathrm{R}I}(\boldsymbol{r}^n) - \boldsymbol{R}_I \right) \tag{15}$$

Inserting eq (15) into eq (14) and integrating by parts, using the fact that derivatives with respect to positions of FG particles mapped on to CG particle I vanish for the term involving the orientation of CG particle I or for any term involving CG particles $J \neq I$ because no

FG particle contributes to more than one CG particle, gives

$$\begin{aligned}\boldsymbol{F}_I(\boldsymbol{R}^N,\boldsymbol{\Omega}^N) = \frac{1}{Z(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)}\int \mathrm{d}\boldsymbol{r}^n \exp\left(-u(\boldsymbol{r}^n)/k_\mathrm{B}T\right)\left(\sum_{i\in\zeta_I}\boldsymbol{f}_i(\boldsymbol{r}^n)\right) \\ \times\,\delta\left(\mathbb{M}^N_\mathrm{R}\left(\boldsymbol{r}^n\right)-\boldsymbol{R}^N\right)\delta\left(\mathbb{M}^N_\Omega\left(\boldsymbol{r}^n\right)-\boldsymbol{\Omega}^N\right) \\ = \left\langle\sum_{i\in\zeta_I}\boldsymbol{f}_i(\boldsymbol{r}^n)\right\rangle_{\boldsymbol{R}^N,\boldsymbol{\Omega}^N} \end{aligned} \tag{16}$$

where $\boldsymbol{f}_i = -\frac{\partial u(\boldsymbol{r}^n)}{\partial \boldsymbol{r}_i}$ is the force on particle i in the FG system and $\langle\cdots\rangle_{\boldsymbol{R}^N,\boldsymbol{\Omega}^N}$ denotes an average over FG configurations that are mapped to CG configuration $(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)$. Equation (16) generalizes the force-matching condition of the MS-CG method for isotropic CG particles[35] to anisotropic particles. It shows that consistency between the FG and CG models in configuration space requires the force on a CG particle in a particular CG system configuration to equal the average total force acting on its constituent FG particles in FG configurations mapped onto the CG configuration. A full derivation of the expression for CG forces can be found in the SI, section S1.

2.2 Momentum-Space Consistency Condition

The FG particle momenta $\boldsymbol{p}^n$ are mapped onto anisotropic CG particle linear momenta $\boldsymbol{P}^N$ and angular momenta $\boldsymbol{L}^N$ via the mapping operators

$$\mathbb{M}^N_\mathrm{P}(\boldsymbol{p}^n) \equiv (\mathbb{M}_{\mathrm{P}1}(\boldsymbol{p}^n),\mathbb{M}_{\mathrm{P}2}(\boldsymbol{p}^n),...,\mathbb{M}_{\mathrm{P}N}(\boldsymbol{p}^n)) = \boldsymbol{P}^N \tag{17}$$

and

$$\mathbb{M}^N_\mathrm{L}(\boldsymbol{p}^n,\boldsymbol{r}^n) \equiv (\mathbb{M}_{\mathrm{L}1}(\boldsymbol{p}^n,\boldsymbol{r}^n),\mathbb{M}_{\mathrm{L}2}(\boldsymbol{p}^n,\boldsymbol{r}^n),...,\mathbb{M}_{\mathrm{L}N}(\boldsymbol{p}^n,\boldsymbol{r}^n)) = \boldsymbol{L}^N \tag{18}$$

respectively. The definitions of these mappings follow from the relationship between generalized coordinates and momenta from Hamilton's equations and the definitions of the mappings from FG positions to CG positions and orientations given by eqs (9) and (10). The mapping

to the linear momentum for CG particle I is

$$\mathbb{M}_{\mathrm{P}I}(\boldsymbol{p}^n) = \boldsymbol{P}_I = M_I \dot{\boldsymbol{R}}_I = M_I \frac{\sum_{i\in\zeta_I} m_i \dot{\boldsymbol{r}}_i}{\sum_{i\in\zeta_I} m_i} = \frac{M_I}{\sum_{i\in\zeta_I} m_i} \sum_{i\in\zeta_I} \boldsymbol{p}_i \tag{19}$$

where we have used the mapping of particle positions in eq (9) and Hamilton's equations, $\dot{\boldsymbol{r}}_i = \frac{\partial H_{\mathrm{FG}}}{\partial \boldsymbol{p}_i}$ and $\dot{\boldsymbol{R}}_I = \frac{\partial H_{\mathrm{CG}}}{\partial \boldsymbol{P}_I}$, applied to the FG and CG Hamiltonians in eqs (1) and (2), respectively, to relate the FG and CG linear momenta and velocities. The mapping to the angular momentum of CG particle I is

$$\mathbb{M}_{\mathrm{L}I}(\boldsymbol{p}^n, \boldsymbol{r}^n) = \boldsymbol{L}_I = \mathbb{I}_I \dot{\boldsymbol{\Omega}}_I = \mathbb{I}_I \mathbb{I}_{\mathrm{FG},I}^{-1} \mathbb{I}_{\mathrm{FG},I} \dot{\boldsymbol{\Omega}}_I = \mathbb{I}_I \mathbb{I}_{\mathrm{FG},I}^{-1} \sum_{i\in\zeta_I} \Delta\boldsymbol{r}_i \times \boldsymbol{p}_i \tag{20}$$

where we have applied Hamilton's equation, $\dot{\boldsymbol{\Omega}}_I = \frac{\partial H_{\mathrm{CG}}}{\partial \boldsymbol{L}_I}$, for the CG angular velocity to eq (2) and used the fact that the mapping of particle orientations defined by eq (10) requires the mapped FG angular velocity to equal the CG angular velocity. This identifies $\mathbb{I}_{\mathrm{FG},I}\dot{\boldsymbol{\Omega}}_I$ as the total angular momentum of the FG particles with respect to their center-of-mass, $\sum_{i\in\zeta_I} \Delta\boldsymbol{r}_i \times \boldsymbol{p}_i$, given the definition of the CG orientation $\boldsymbol{\Omega}_I$ in terms of the principal axes of inertia of the group of FG particles that are mapped onto the CG particle. The mapping $\mathbb{M}_{\mathrm{L}I}(\boldsymbol{p}^n, \boldsymbol{r}^n)$ for the angular momentum depends on the FG positions due to the dependence of the angular momentum on FG particle positions with respect to the center-of-mass.

Consistency between the FG and CG systems in momentum space requires the linear and angular momentum probability distributions in the CG system to match the corresponding probability distributions of the mapped FG system. As the CG angular momentum mapping depends on both FG coordinates and momenta, derivation of momentum-space consistency conditions must use the overall matching of the equilibrium phase-space distributions of the

FG and CG systems,

$$\mathbb{P}_{\mathrm{CG,m}}(\boldsymbol{P}^N,\boldsymbol{L}^N)\mathbb{P}_{\mathrm{CG,c}}(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)=\int\int \mathrm{d}\boldsymbol{r}^n\mathrm{d}\boldsymbol{p}^n\mathbb{P}_{\mathrm{FG,m}}(\boldsymbol{p}^n)\mathbb{P}_{\mathrm{FG,c}}(\boldsymbol{r}^n)\delta\left(\mathbb{M}_{\mathrm{P}}^N(\boldsymbol{p}^n)-\boldsymbol{P}^N\right)$$
$$\times\,\delta\left(\mathbb{M}_{\mathrm{L}}^N(\boldsymbol{p}^n,\boldsymbol{r}^n)-\boldsymbol{L}^N\right)\delta\left(\mathbb{M}_{\mathrm{R}}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{R}^N\right)\delta\left(\mathbb{M}_{\Omega}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{\Omega}^N\right) \tag{21}$$

Inserting eqs (4) and (6) for the FG and CG momentum probability distributions and eqs (17)–(20) for the linear and angular momentum mappings gives

$$(2\pi k_{\mathrm{B}}T)^{(3n/2-3N)}\left\{\prod_{I=1}^{N}\frac{\exp\left[-\frac{1}{2k_{\mathrm{B}}T}\left(\frac{\boldsymbol{P}_I^2}{M_I}+\boldsymbol{L}_I^{\mathrm{T}}\mathbb{I}_I^{-1}\boldsymbol{L}_I\right)\right]}{M_I^{3/2}\det\left(\mathbb{I}_I\right)^{1/2}}\right\}\mathbb{P}_{\mathrm{CG,c}}(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)$$
$$=\int \mathrm{d}\boldsymbol{r}^n\mathbb{P}_{\mathrm{FG,c}}(\boldsymbol{r}^n)\delta\left(\mathbb{M}_{\mathrm{R}}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{R}^N\right)\delta\left(\mathbb{M}_{\Omega}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{\Omega}^N\right)$$
$$\times\prod_{I=1}^{N}\int\left[\prod_{i\in\zeta_I}\mathrm{d}\boldsymbol{p}_i m_i^{-3/2}\exp\left(-\frac{\boldsymbol{p}_i^2}{2m_ik_{\mathrm{B}}T}\right)\delta(c_{\mathrm{p}}\sum_{i\in\zeta_I}\boldsymbol{p}_i-\boldsymbol{P}_I)\delta(\mathbb{C}_{\mathrm{L}}\sum_{i\in\zeta_I}\Delta\boldsymbol{r}_i\times\boldsymbol{p}_i-\boldsymbol{L}_I)\right] \tag{22}$$

where $c_{\mathrm{p}}\equiv\frac{M_I}{\sum_{i\in\zeta_I}m_i}$ and $\mathbb{C}_{\mathrm{L}}\equiv\mathbb{I}_I\mathbb{I}_{\mathrm{FG},I}^{-1}$. The integrals in eq (22) over the FG momenta $\boldsymbol{p}_i$ can be carried out analytically by using the Fourier representation of the Dirac delta functions and Gaussian identities to give

$$\left[\prod_{I=1}^{N}\exp\left(-\frac{\boldsymbol{P}_I^2}{2k_{\mathrm{B}}TM_I}\right)\right]\left[\prod_{I=1}^{N}\exp\left(-\frac{\boldsymbol{L}_I^{\mathrm{T}}\mathbb{I}_I^{-1}\boldsymbol{L}_I}{2k_{\mathrm{B}}T}\right)\right]\mathbb{P}_{\mathrm{CG,c}}(\boldsymbol{R}^N,\boldsymbol{\Omega}^N)$$
$$=\left[\prod_{I=1}^{N}\left(\frac{\sum_{i\in\zeta_I}m_i}{M_I}\right)^{3/2}\exp\left(-\frac{\sum\limits_{i\in\zeta_I}m_i\boldsymbol{P}_I^2}{2M_I^2k_{\mathrm{B}}T}\right)\right]$$
$$\times\int\mathrm{d}\boldsymbol{r}^n\mathbb{P}_{\mathrm{FG,c}}(\boldsymbol{r}^n)\delta\left(\mathbb{M}_{\mathrm{R}}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{R}^N\right)\delta\left(\mathbb{M}_{\Omega}^N\left(\boldsymbol{r}^n\right)-\boldsymbol{\Omega}^N\right)$$
$$\times\prod_{I=1}^{N}\left(\frac{\det(\mathbb{I}_{\mathrm{FG},I})}{\det(\mathbb{I}_I)}\right)^{1/2}\exp\left(-\frac{\boldsymbol{L}_I^{\mathrm{T}}\mathbb{I}_I^{-1}\mathbb{I}_{\mathrm{FG},I}\mathbb{I}_I^{-1}\boldsymbol{L}_I}{2k_{\mathrm{B}}T}\right) \tag{23}$$

For this equation to hold for all values of the CG linear momentum for all CG particles, each

factor involving the linear momentum $\boldsymbol{P}_I$ of a CG particle I must be equal on the two sides of the equation, which requires the CG particle mass to satisfy

$$M_I = \sum_{i \in \zeta_I} m_i \tag{24}$$

This consistency condition for the CG linear momenta had previously been derived (in a more general form for a more general linear mapping from FG to CG positions) in the development of the MS-CG method for isotropic CG particles.[35]

Inserting eq (24) into eq (23) and matching the FG and CG configurational distributions using eq (11) gives

$$\begin{aligned} &\prod_{I=1}^{N} \det(\mathbb{I}_I)^{1/2} \exp\left(-\frac{\boldsymbol{L}_I^{\mathrm{T}} \mathbb{I}_I^{-1} \boldsymbol{L}_I}{2k_{\mathrm{B}}T}\right) \\ &\quad = \left\langle \prod_{I=1}^{N} \det(\mathbb{I}_{\mathrm{FG},I})^{1/2} \exp\left(-\frac{\boldsymbol{L}_I^{\mathrm{T}} \mathbb{I}_I^{-1} \mathbb{I}_{\mathrm{FG},I} \mathbb{I}_I^{-1} \boldsymbol{L}_I}{2k_{\mathrm{B}}T}\right) \right\rangle_{\boldsymbol{R}^N, \boldsymbol{\Omega}^N} \end{aligned} \tag{25}$$

where $\langle \cdots \rangle_{\boldsymbol{R}^N, \boldsymbol{\Omega}^N}$ denotes an average over FG configurations that are mapped to CG configuration $(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$. The left-hand side of eq (25) consists of a product of separate factors for each CG particle that do not depend on the configuration of the CG system, whereas the right-hand side in general is not separable into independent factors and does depend on the CG configuration. Consistency between FG and CG angular momenta requires the right-hand side of eq (25) to be separable into independent factors for each principal axis of each each CG particle, which should be achievable (at least approximately) via a judicious choice of the CG mapping, e.g. by mapping groups of FG particles that form approximately rigid bodies into CG particles such that the inertia tensor $\mathbb{I}_{\mathrm{FG},I}$ of each group of FG particles is largely unperturbed by the surrounding particles. In this case, from eq (25) the consistency

condition for the angular momentum about the principal axis q of CG particle I is

$$I_{I,q}^{1/2} \exp\left(-\frac{L_{I,q}^2}{2I_{I,q}k_{\mathrm{B}}T}\right) = \left\langle I_{\mathrm{FG},I,q}^{1/2} \exp\left(-\frac{I_{\mathrm{FG},I,q}L_{I,q}^2}{2I_{I,q}^2k_{\mathrm{B}}T}\right)\right\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I} \tag{26}$$

where $I_{\mathrm{FG},I,q}$, $I_{I,q}$, and $L_{I,q}$ are the FG moment of inertia, CG moment of inertia, and angular momentum about the q axis for CG particle I, and $\langle\cdots\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}$ denotes an equilibrium average over FG configurations consistent with the coordinate mapping of CG particle I. This condition can be simplified to eliminate the CG moment of inertia $I_{I,q}$ from the right-hand side by using eq (20) to write it in terms of CG angular velocities as

$$I_{I,q}^{1/2} \exp\left(-\frac{I_{I,q}\dot{\Omega}_{I,q}^2}{2k_{\mathrm{B}}T}\right) = \left\langle I_{\mathrm{FG},I,q}^{1/2} \exp\left(-\frac{I_{\mathrm{FG},I,q}\dot{\Omega}_{I,q}^2}{2k_{\mathrm{B}}T}\right)\right\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I} \tag{27}$$

where $\dot{\Omega}_{I,q}$ is the angular velocity component corresponding to the q axis. For the left- and right-hand sides of eq (27) to be consistent, they must have the same functional dependence on the angular velocity $\dot{\Omega}_{I,q}$, which is shown in section S2 of the SI to be the case if the fluctuations of $I_{\mathrm{FG},I,q}$ are small compared with its mean. (For a Gaussian distribution of $I_{\mathrm{FG},I,q}$, this condition corresponds to the standard deviation of $I_{\mathrm{FG},I,q}$ being much smaller than its mean, i.e. $\left\langle \left(I_{\mathrm{FG},I,q} - \langle I_{\mathrm{FG},I,q}\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}\right)^2\right\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}^{1/2} \ll \langle I_{\mathrm{FG},I,q}\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}$.) In this case, eq (27) reduces to

$$I_{I,q}^{1/2} \exp\left(-\frac{I_{I,q}\dot{\Omega}_{I,q}^2}{2k_{\mathrm{B}}T}\right) \approx \langle I_{\mathrm{FG},I,q}\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}^{1/2} \exp\left(-\frac{\langle I_{\mathrm{FG},I,q}\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I}\dot{\Omega}_{I,q}^2}{2k_{\mathrm{B}}T}\right) \tag{28}$$

from which comparing the left- and right-hand sides gives

$$I_{I,q} \approx \langle I_{\mathrm{FG},I,q}\rangle_{\boldsymbol{R}_I,\boldsymbol{\Omega}_I} \tag{29}$$

for each principal axis q. This expression implies that consistency between the FG and CG angular-momentum distributions is achieved if the CG principal moments of inertia are equal

to the corresponding averages of the principal moments of inertia of the FG particles mapped onto the CG particle, provided their fluctuations are small compared to the average values. Note that if this condition is satisfied, the assumption used to derive eq (25) from eq (26) of independent FG inertia tensors for each CG particle will be approximately satisfied. The full derivation for the momentum-space consistency conditions for anisotropic particles is given in the SI in section S2.

2.3 Anisotropic force-matching coarse-graining (AFM-CG) consistency conditions

In summary, the following conditions are sufficient to yield a CG model consisting of anisotropic particles whose equilibrium canonical phase-space distribution is consistent with a particular FG model:

1. Each CG particle represents a group of one or more atoms in the FG system, and each atom is only involved in the definition of one CG particle. The CG position and orientation coordinates correspond to the center-of-mass and principal axes of inertia, respectively, of the FG particles that constitute the CG particle.

2. The CG potential is defined such that the force on each CG particle in a particular CG configuration is equal to the average over equivalent FG configurations of the total force on the FG particles that constitute the CG particle, as expressed in eq (16).

3. Given condition 1, the mass of each CG particle is defined as the sum of the constituent FG particles.

4. Given condition 1, each CG particle is defined such that the fluctuations of each principal moment of inertia of the group of constituent FG particles is much smaller than its mean, and the CG principal moments of inertia are defined to be equal to the mean values of the corresponding FG principal moments of inertia.

3 Coarse-Graining and Simulation Methods

3.1 Anisotropic force-matching coarse-graining (AFM-CG) algorithm

The AFM-CG method produces anisotropic CG force fields from the FG many-body PMF by means of force-matching. In practice, identifying the CG potential $U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$ that yields CG forces that satisfy the force-matching condition in eq (16) exactly for any CG configuration $(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$ is generally non-trivial. However, by approximating the CG potential $U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda})$ as a set of functions that depends on a set of parameters $\boldsymbol{\lambda}$, the optimal CG potential of this form subject to the force-matching condition can be found as the one whose parameters $\boldsymbol{\lambda}$ minimize the average least-squares residual between the CG and FG forces in corresponding system configurations,[35] given by

$$\chi^2[\boldsymbol{\lambda}] = \frac{1}{3N} \left\langle \sum_{I=1}^{N} \left| \sum_{i \in \zeta_I} \boldsymbol{f}_i(\boldsymbol{r}^n) - \boldsymbol{F}_I(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda}) \right|^2 \right\rangle \tag{30}$$

where $\boldsymbol{F}_I(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda}) = -\partial U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda})/\partial \boldsymbol{R}_I$ is the force on CG particle I for the specified force field, $\left(\boldsymbol{R}^N, \boldsymbol{\Omega}^N\right) = \left(\mathbb{M}_{\mathrm{R}}^N(\boldsymbol{r}^n), \mathbb{M}_{\Omega}^N(\boldsymbol{r}^n)\right)$ are the mapped CG coordinates corresponding to the FG configuration $\boldsymbol{r}^n$, and the angle brackets denote an average over the equilibrium ensemble of the FG model. Optimization of the CG force field requires only that a FG simulation be conducted to give the FG forces as a function of FG positions. For each equilibrium FG simulation configuration used for CG parametrization, the FG positions are mapped onto CG positions and orientations to calculate the CG forces used to evaluate the force residual in eq (30).

In general, the CG potential is a sum of intramolecular and intermolecular interactions, including non-bonded electrostatic and van de Waals dispersion interactions, and bond-stretching, angle-bending, and dihedral-torsion interactions. While the AFM-CG method can be applied to construct all of these CG interaction types, in this work we focus on anisotropic non-bonded interactions, as the other interactions are readily approximated using

existing coarse-graining methods, such as the MS-CG method for isotropic CG particles[35] or the iterative Boltzmann inversion method.[38] In the general case, a hybrid approach using the AFM-CG method for non-bonded interactions and another method for bonded interactions could be used. Several previous studies of isotropic CG models have obtained the CG potential using hybrid approaches employing more than one coarse-graining method and have demonstrated good structural and thermodynamic accuracy.[39,40]

Similarly to the implementation of the MS-CG method,[35] we have used a linear combination of scalar basis functions to approximate the non-bonded CG potential.[36] In general, this potential is a many-body interaction that depends on the positions and orientations of all CG particles in the system. In the implementation presented here, we have assumed that this potential is the sum of pair-wise interactions between CG particles, which significantly simplifies the CG potential into a function only of the relative positions and orientations of pairs of CG particles. Extension of the method beyond two-body interactions is possible. With this approximation, in general each basis function in the CG potential can be written as a function of a set of 6 scalar variables, $\boldsymbol{\xi} \equiv (\xi_1, \xi_2, \ldots, \xi_6)$, that are functions of the relative position $\boldsymbol{R}_{IJ}$ and orientation $\boldsymbol{\Omega}_{IJ}$ of a pair of CG particles I and J. For the simplest case in which the pair interaction is the same for all CG particles (generalization to different interaction types is straightforward), the CG potential can be written as a sum over N_b scalar basis functions $\mathcal{B}_b$ per pair of particles as

$$U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda}) = \sum_{b=1}^{N_\mathrm{b}} \lambda_b \sum_{I=1}^{N} \sum_{J \neq I}^{N} \mathcal{B}_b \left(\boldsymbol{\xi} \left(\boldsymbol{R}_{IJ}, \boldsymbol{\Omega}_{IJ} \right) \right) \tag{31}$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_{N_\mathrm{b}})$ is the corresponding basis coefficient vector. The force on CG particle I due to this potential is thus

$$\boldsymbol{F}_I(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda}) = -\sum_{b=1}^{N_\mathrm{b}} \lambda_b \sum_{J \neq I}^{N} \sum_{j=1}^{6} \frac{\partial \mathcal{B}_b \left(\boldsymbol{\xi} \right)}{\partial \xi_j} \frac{\partial \xi_j}{\partial \boldsymbol{R}_I} = \sum_{b=1}^{N_\mathrm{b}} \lambda_b \boldsymbol{G}_{b,I}(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) \tag{32}$$

where $\boldsymbol{G}_{b,I}(\boldsymbol{R}^N, \boldsymbol{\Omega}^N) = -\sum_{J\neq I}^{N}\sum_{j=1}^{6}\frac{\partial \mathcal{B}_b(\boldsymbol{\xi})}{\partial \xi_j}\frac{\partial \xi_j}{\partial \boldsymbol{R}_I}$ is a force basis vector function for CG particle I. Inserting eq (32) into eq (30) results in

$$\chi^2\left[\boldsymbol{\lambda}\right] = \frac{1}{3N}\left\langle \sum_{I=1}^{N}\left|\sum_{i\in\zeta_I}\boldsymbol{f}_i(\boldsymbol{r}^n) - \sum_{b=1}^{N_\mathrm{b}}\lambda_b \boldsymbol{G}_{b,I}(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)\right|^2\right\rangle \tag{33}$$

The minimization of the force residual in eq (33) is equivalent to obtaining the least-squares solution $\boldsymbol{\lambda}$ of the matrix equation[36]

$$\mathbb{G}\boldsymbol{\lambda} \approx \mathbb{F} \tag{34}$$

where $\mathbb{G}$ is a $3Nt \times N_b$ matrix and $\boldsymbol{\lambda}$ and $\mathbb{F}$ are column vectors with N_b and $3Nt$ elements, respectively, with t being the number of FG configurations used for CG parametrization. Each of the N_b columns of the matrix $\mathbb{G}$ contains the three components of the basis force vectors $\boldsymbol{G}_{b,I}(\boldsymbol{R}^N, \boldsymbol{\Omega}^N)$ for each of the N CG particles in each of the t FG configurations. The vector $\mathbb{F}$ contains the three components of the corresponding total FG force, $\sum_{i\in\zeta_I}\boldsymbol{f}_i(\boldsymbol{r}^n)$, on the FG particles comprising each of the N CG particles in each of the same t FG configurations. As the matrix generally contains many rows compared with the number of columns, it is beneficial to apply prior sequential QR factorizations to transform it into square form[41] to reduce the computational cost of producing the least-squares solution. After finding the optimal parameter set $\boldsymbol{\lambda}$, the CG potential $U(\boldsymbol{R}^N, \boldsymbol{\Omega}^N; \boldsymbol{\lambda})$ can be reconstructed using eq (31), from which forces and torques can be calculated for use in dynamical simulations.

3.2 Basis Set Expansion of Anisotropic CG Potential

In the implementation of the MS-CG method for isotropic CG particles, several different types of basis functions of a single variable (the inter-particle separation) were tested, including discrete delta functions, linear spline basis functions, and piecewise continuous cubic polynomials.[36] For anisotropic CG particles in the AFM-CG method, multivariate basis functions must be used as the potential depends on both the relative position and orientation of

CG particles.

For two anisotropic CG particles I and J, the basis functions in the pair-wise expansion of the CG potential can be written in general as a function of the set of variables $\boldsymbol{\xi} = \{R, \alpha, \beta, \gamma, \phi, \theta\}$, where R is the inter-particle distance, α, β, and γ are Euler angles that describe the rotation of particle I into the frame of particle J, and θ and ϕ are angles between the inter-particle vector and a specific principal axis on each particle. Details of the calculations of the relative positions and orientations of pairs of CG particles and the formulas for the scalar variables are given in section S3 of the SI.

The form of these basis functions plays an important role in defining the shape of the CG potential and the accuracy of the method. Here we have assumed that each multivariate basis function is a product of a cubic spline function of the inter-particle distance R, which ensures smoothness and continuity of the forces, and cosine functions of each the other orientational variables in the case where the CG non-bonded potentials are even periodic functions of angular variables. Hence, we have represented the linear combination of basis functions in eq (31) as

$$\sum_{i=1}^{N_\mathrm{b}} \lambda_i \mathcal{B}_i(R, \alpha, \beta, \gamma, \phi, \theta) \equiv \sum_{n=1}^{N_\mathrm{R}} \sum_{k_\alpha=0}^{N_\alpha} \sum_{k_\beta=0}^{N_\beta} \sum_{k_\gamma=0}^{N_\gamma} \sum_{k_\phi=0}^{N_\phi} \sum_{k_\theta=0}^{N_\theta} \lambda_{n,k_\alpha,k_\beta,k_\gamma,k_\phi,k_\theta} \mu_n(R) \cos(k_\alpha \alpha) \cos(k_\beta \beta) \times \cos(k_\gamma \gamma) \cos(k_\phi \phi) \cos(k_\theta \theta) \qquad (35)$$

where there are $N_\mathrm{b} = N_\mathrm{R} N_\alpha N_\beta N_\gamma N_\phi N_\theta$ basis coefficients $\lambda_{n,k_1,k_2,k_3,k_4,k_5}$ in total to be optimized, and $\mu_n(R)$ is one of N_R cubic spline basis functions of the inter-particle distance variable. For CG models that have rotational symmetry around one or more axes, the number of angular variables required to specify the pair configuration is reduced. For example, for uniaxial CG particles with one axis of rotational symmetry, only one Euler angle is needed to describe rotation between the molecular coordinate frames, reducing the number of variables in eq (35) by two. Explicit formulas for the cubic spline basis functions and expressions for the force basis vectors are given in section S3 of the SI.

3.3 Modified S-Function Fit to Anisotropic CG Potential

Evaluation of CG forces and torques from the expression for the CG potential in eq (35) can be computationally costly if the number of basis functions is large. To expedite force and torque evaluations so as to enable more efficient CG MD simulations, we have fit the optimized CG potential obtained by AFM-CG to a simpler analytical form containing a smaller number of parameters. For this simpler potential, we have used a modified form of the S-function potential U_{S} for disk-shaped particles,[42–44] to which two additional terms, U_{exp} and U_{osc}, have been added to enable a better fit to the CG potential: U_{exp} enables a better fit to the repulsive part of the potential, while U_{osc} captures oscillations in the potential. Thus, the pair-wise potential between uniaxial CG particles I and J, which is a function of the vector $\boldsymbol{R}_{IJ} = \boldsymbol{R}_I - \boldsymbol{R}_J$ that connects the particle centers and the unit vectors $\hat{\boldsymbol{u}}_I$ and $\hat{\boldsymbol{u}}_J$ along the unique axis of each particle, was fitted to the form

$$U_{\mathrm{S}}^{\mathrm{mod}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = U_{\mathrm{S}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) + U_{\mathrm{exp}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) + U_{\mathrm{osc}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) \tag{36}$$

where

$$U_{\mathrm{S}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = 4\epsilon\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)\left[A\left(\frac{\sigma_0}{d(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J)}\right)^p - \left(\frac{\sigma_0}{d(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J)}\right)^6\right] \tag{37}$$

$$U_{\mathrm{exp}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = \epsilon_{\mathrm{exp}}\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)\exp\left[-\kappa_1 \frac{d_{\mathrm{exp}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)}{\sigma_0}\right] \tag{38}$$

and

$$\begin{aligned} U_{\mathrm{osc}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = \epsilon_{\mathrm{osc}}\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)\left\{\cos\left[k_2 \frac{d_{\mathrm{osc}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)}{\sigma}\right] + C_2\right\} \\ \times \exp\left[-\kappa_2 \frac{d_{\mathrm{osc}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)}{\sigma}\right] \end{aligned} \tag{39}$$

with

$$d\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = R_{IJ} - \sigma\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) + \sigma_0 \tag{40}$$

$$d_{\text{exp}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = R_{IJ} - \sigma\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) + \sigma_{\text{exp}} \tag{41}$$

and

$$d_{\text{osc}}\left(\boldsymbol{R}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) = R_{IJ} - \sigma\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right) + \sigma_{\text{osc}} \tag{42}$$

Here $R_{IJ} \equiv \|\boldsymbol{R}_{IJ}\|$ and $\hat{\boldsymbol{R}}_{IJ} \equiv \boldsymbol{R}_{IJ}/R_{IJ}$, while σ_0, σ_{exp}, κ_1, σ_{osc}, k_2, κ_2, and C_2 are fit parameters. $\epsilon\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$, $\epsilon_{\text{exp}}\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$, and $\epsilon_{\text{osc}}\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$ are anisotropic energy functions, and $\sigma\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$ is an anisotropic distance function, which contain a number of fit parameters. The standard form of the S-function potential[43] was modified to include the prefactor $A\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$ and exponent $p\left(\hat{\boldsymbol{R}}_{IJ}, \hat{\boldsymbol{u}}_I, \hat{\boldsymbol{u}}_J\right)$, which are functions of the pair orientation. The forms of these anisotropic fitted functions and derivation of the forces and torques are defined in section S4 of the SI.

To enable CG simulations at constant pressure and temperature (NPT ensemble) that accurately capture the density of the FG system, an isotropic linear correction was added to the fitted CG potentials of the form[8]

$$\Delta U(R_{ij}) = b\left(1 - \frac{R_{ij}}{R_{\text{c}}}\right) \tag{43}$$

where R_{c} is the cutoff distance for the non-bonded CG interactions and b is a parameter that was chosen so that the pressure in the CG simulation matched that of the corresponding FG simulation under the parametrization conditions. This pressure correction was relatively small compared with the fitted modified S-function potential for the systems studied, as shown in the SI in Figures S3 and S14. More rigorous methods of pressure-matching for CG parametrization in the NPT ensemble involving thermodynamic state-dependent corrections are possible, along the lines of those previously applied to isotropic CG models,[12,45–47] and will be the focus of future study.

3.4 Molecular Dynamics (MD) Simulations

All MD simulations were conducted with a modified version of the LAMMPS software package[48–50] (version 3Mar20) incorporating the modified S-function pair potential (eq (36)). All-atom FG simulations used parameters from the OPLS-AA force field[51–55] without modification, using a cut-off distance for short-ranged non-bonded interactions of 10 Å and calculating the long-ranged component of the electrostatic interactions with the particle–particle particle–mesh (PPPM) method.[50,56] These simulations were carried out at constant temperature and pressure (NPT ensemble), with the pressure set to 1 atm. Although the derivation of the AFM-CG method given above is for the constant-volume conditions of the NVT ensemble rather than constant pressure, volume fluctuations (as measured by their standard deviation) were less than 1% of the average system volume in the FG simulations used for CG parametrization for all systems studied, and so the use of NPT simulations is expected to have minimal impact on coarse-graining. Temperature and pressure were controlled with a Nosé-Hoover thermostat and barostat.[57,58]

CG simulations were carried out with the fitted modified S-function potential (eq (36)) with interactions truncated at a system-dependent cut-off distance R_c (given in the sections below). After optimizing the pressure-correction potential (eq (43)) in the NVT ensemble at the same temperature used for the AFM-CG parametrization, CG simulations were undertaken in the NPT ensemble. CG simulations in the NVT ensemble with densities set to the FG average values were also conducted to compare with the FG simulations without the confounding effect of differing densities.

4 Results and Discussion

As a proof-of principle, we have applied the AFM-CG method to coarse grain pure liquid systems of benzene and perylene, with both molecules modeled as uniaxial anisotropic particles with a single non-equivalent principal axis. In both cases, the principal moments of

inertia of the CG particles were defined in terms of the average principal moments of inertia of the molecules in the FG simulation used for parametrization. The fluctuations in the principal moments of inertia in the FG simulations were verified to be small compared with their mean values in both cases (SI Figures S1 and S12, respectively), so approximating the CG principal moments by the corresponding average FG values should accurately satisfy the momentum-space consistency condition described in section 2.3. While benzene is approximately uniaxial, perylene is more accurately described as biaxial, although the two in-plane axes are relatively similar in length compared with the out-of-plane axis. Thus, we have used a uniaxial CG perylene model here for simplicity as it significantly reduces the number of basis functions needed to represent the anisotropic CG potential. For both molecules, the moments of inertia about the two in-plane principal axes in the CG model were set to be equal to the average of the FG principal moments of inertia about the two in-plane axes. Extension of the method to biaxial CG models is possible. Simulation speedups of 40–140 times—due to both the reduction in the computational time per simulation time step and the larger time step that was feasible—were achieved using the CG models compared with the FG models for simulations carried out on eight Intel Core i7 6700K CPUs. Further speedups could be achieved by porting the simulation code to graphics processing units (GPUs).

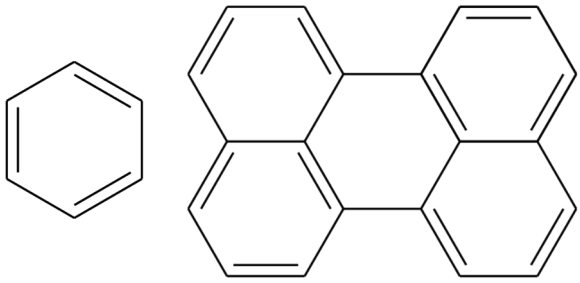

Figure 2: Chemical structures of (a) benzene and (b) perylene

4.1 Benzene

An all-atom simulation of 500 benzene molecules at 300 K and 1 atm was conducted to parametrize the CG model, in which each benzene molecule was mapped to a single CG particle. Simulations of the same system were also carried out at 280, 320, 330 and 350 K to verify transferability of the CG model across thermodynamic conditions. The simulation time step was 2 fs. The total duration of each simulation was 25 ns, with data from the last 20 ns used for analysis. Atom coordinates and forces from simulation at 300 K were output every 1000 time steps (2 ps) to produce a total of 10 000 simulation snapshots as input data for CG parametrization.

The cut-off for CG pair-wise interactions was set to 10 Å. As the CG model had uniaxial symmetry, only the position and out-of-plane principal axis of each molecule were needed to define the pair interaction. Following optimization of the CG potential, the root-mean-square error (RMSE) between the CG and FG forces (the square root of the force residual in eq (30)) was 2.57 kcal mol^{-1} Å^{-1}. This value is significantly lower than RMSEs obtained for isotropic CG models using the MS-CG method with a similar coarse-graining level[36] and is comparable to the RMSE for a multi-site CG benzene model with many-body interactions.[59] The optimized CG pair potential is shown as a function of intermolecular distance for several relative orientations of benzene dimers in Figure 3 and for more relative orientations in the SI (section S5.2). The minimum of the potential curve for the edge-to-face configuration is at about 5 Å, which agrees with previous experimental results[60,61] and quantum calculations[61,62] for benzene dimers. The energy barrier for the face-to-face configuration at 5–7 Å can be attributed to the quadrupolar electrostatic repulsion that dominates dispersion at this distance, which has been observed previously in a quantum-chemical study of benzene dimers.[62]

The fitted modified S-function potential curves for the face-to-face, edge-to-face, and edge-to-edge configurations are compared with the AFM-CG curves in Figure 3. The additional oscillatory term in the fitted potential improves the fit considerably compared with

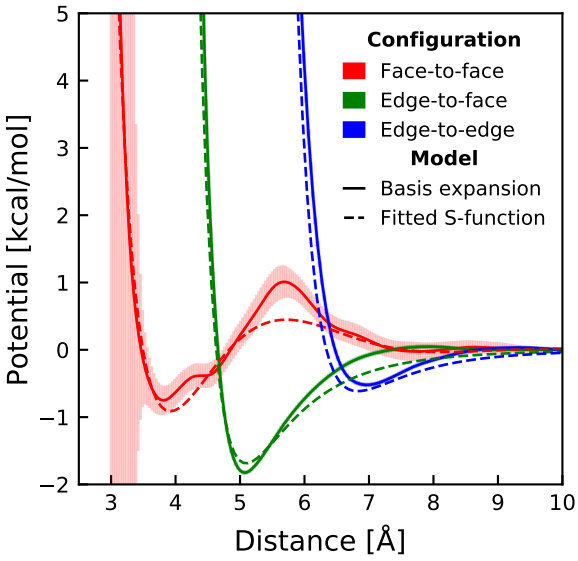

Figure 3: Benzene AFM-CG pair potential (without pressure correction) for various dimer configurations calculated using basis set expansion (solid lines with error bars) and fitted modified S-function potential (dashed lines)

using the S-function form on its own. The fit is generally good with a RMSE between the final fitted model and the FG forces of 2.65 kcal mol^{-1} Å^{-1}, which is 2.8% higher than for the basis-set expansion. The CG potential curves with and without the pressure correction are similar, as shown in SI Figure S3). The parameters used for the CG simulation of benzene are given in Table S2 in the SI.

NPT CG simulations of 500 CG benzene molecules were carried out for 15 ns using a time step of 10 fs, with data from the last 10 ns used for analysis. The average density of liquid benzene from NPT simulations of the pressure-corrected CG model at 1 atm pressure and for temperatures between 280 and 350 K are compared with those from corresponding FG simulations and experiment in Figure 4. The average densities of the FG model are slightly lower but agree reasonably well with experimental values,[63] deviating by less than 4% over the temperature range studied. Given that the CG model was parametrized based on the FG model rather than experiment, closer agreement with the former is to be expected; despite the simplicity of the pressure correction applied, the density of the CG model is close to that of the FG model over the temperature range studied, although the density of the CG

model decreases more rapidly with temperature, with a maximum deviation of less than 2% difference at the highest temperature of 350 K.

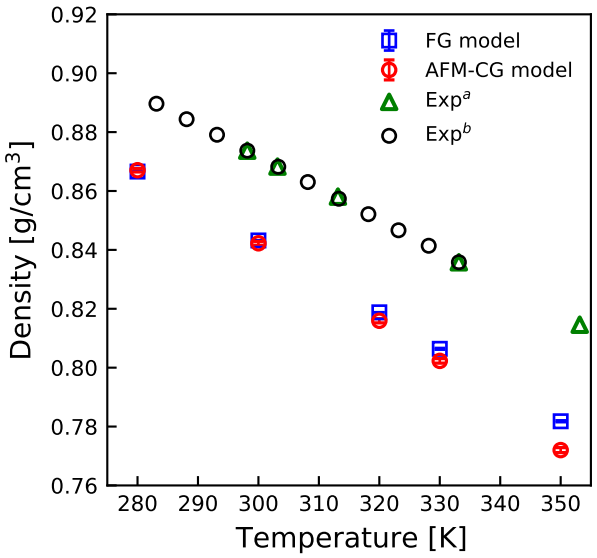

Figure 4: Benzene density versus temperature at 1 atm for the FG and AFM-CG models compared with experimental values (Exp^a: ref 63, Exp^b:ref 64). Error bars for simulated densities are smaller than the symbols.

The radial distribution functions (RDFs) of the FG and CG systems in NPT simulations at 1 atm and for temperatures between 280 and 350 K are compared in Figure 5, which shows close agreement between the FG and CG models, indicating that the CG model captures the intermolecular structural correlations of the FG model well over this temperature range.

To further test the accuracy of the CG model in describing the structural correlations of the FG model, we have also compared the angular–radial distribution function (ARDF) of the two models, which quantities the local orientational structure of the system. If $\langle n(r,\beta)\rangle$ is the average number of molecules whose centers-of-mass are within the distance range r to $r+\Delta r$ and whose symmetry axes are within an angular range β to $\beta+\Delta\beta$ of a tagged molecule centered at the origin, the ARDF is

$$g(r,\beta) = \frac{\langle n(r,\beta)\rangle}{\frac{4}{3}\pi\rho\left[(r+\Delta r)^3 - r^3\right]\sin\beta\Delta\beta} \tag{44}$$

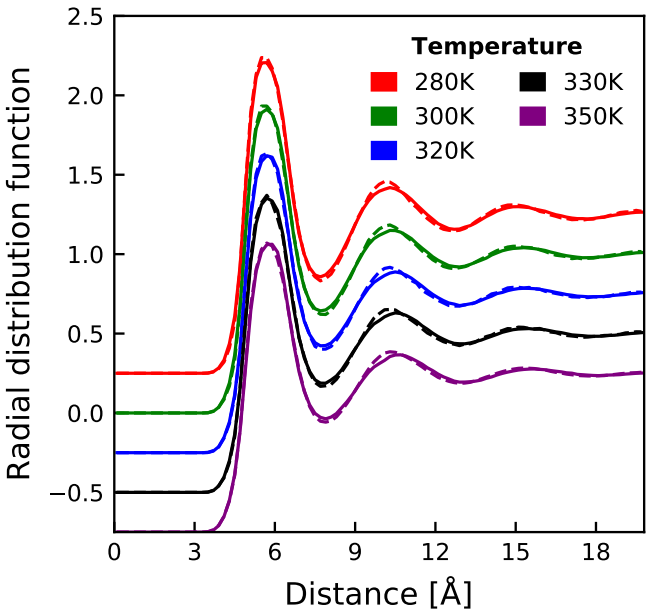

Figure 5: RDFs of FG and AFM-CG benzene models at 1 atm and various temperatures, for AFM-CG model parametrized at 300 K. The RDFs for 280, 300, 320 and 330 K have been shifted vertically for ease of viewing.

where ρ is the number density of molecules. The ARDFs at 300 K of the FG and AFM-CG benzene models are compared in Figure 6, along with that of a recent single-site benzene CG model by Bowen et al.[44], which we will refer to as the "Bowen CG model". The Bowen CG model was parametrized from the same FG OPLS-AA force field but by matching the interaction energy between pairs of rigid benzene molecules in a variety of configurations, and so its parametrization does not account for finite-temperature many-body effects that are taken into account in the AFM-CG method. Figure 6 shows that the ARDF for the Bowen CG model is relatively isotropic, except at around 5 Å where the edge-to-face configuration is most favorable, whereas the distributions vary anisotropically for the FG and AFM-CG models over a wider range of intermolecular separations. The ARDFs for the FG and AFM-CG models are in good agreement, except for some slight differences at a distance of around 6 Å. This can be explained by minor differences between the fitted S-function potential compared with the AFM-CG basis expansion potential shown in Figure 3 and in Figure S2 in the SI.

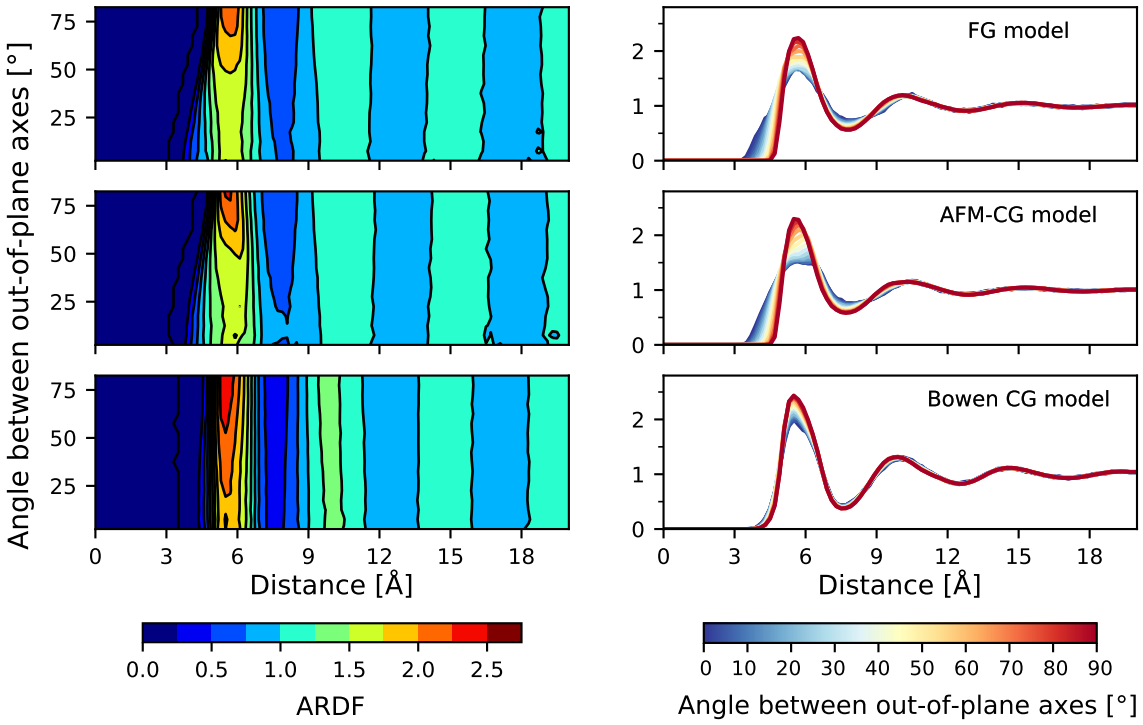

Figure 6: ARDFs for FG and CG benzene models at 300 K (AFM-CG model: this work; Bowen CG model: ref 44), shown as 2D color maps on the left and 1D slices at fixed angle on the right. An angle of 0° between out-of-plane axes indicates parallel alignment (e.g. face-to-face or edge-to-edge) and 90° indicates perpendicular alignment (e.g. edge-to-face configuration).

The poorer agreement between the ARDFs FG and Bowen CG models can be partly explained by the higher density (0.935 g cm^{-3}) of the Bowen CG model under the conditions simulated. To eliminate the effect of density in the comparison, RDFs and ARDFs from FG NPT simulations and CG NVT simulations with densities set to those of the FG simulations at 300 and 330 K are compared in the SI in section S5.3. In addition, ARDFs from FG and CG NPT simulations at 330 K are compared in Figure S8. The RDFs and ARDFs of the AFM-CG model agree better with the FG model than the Bowen CG model in all cases, although the agreement of the Bowen CG model with the FG model is significantly improved when its density matches that of the FG model. Overall, the AFM-CG model reproduces the FG liquid structure of benzene very well.

Although the AFM-CG method does not optimize the CG model for dynamics, we have

compared the dynamics of the parametrized CG model with the FG model by measuring translational and rotational time correlation functions in these systems. The translational diffusion coefficient D_{T} was obtained as 1/6 of the slope of the linear region at long times of a plot of the mean squared displacement (MSD) of the molecular center-of-mass versus time t, defined as $\langle \Delta R^2(t) \rangle = \left\langle \left| \boldsymbol{R}(t+t') - \boldsymbol{R}(t') \right|^2 \right\rangle$, where $\boldsymbol{R}(t')$ is the position of a molecule's center-of-mass at time t' and the angle brackets denote an ensemble average. The MSD is plotted versus time at various temperatures for the AFM-CG and FG models in Figure S9 in the SI. The diffusion coefficients calculated from the fitted slopes of these plots are plotted as a function of temperature along with experimental data in Figure 7. By linear interpolation and extrapolation of the line-of-best-fit to the experimental points, the FG diffusion coefficients are found to overestimate the experimental values by approximately 31%, 21%, 7% and 10% at 280, 300, 320 and 330 K, respectively. Finite-size effects lower the translational diffusion coefficient in simulated systems with periodic boundaries relative to the infinite system-size limit,[65] and so cannot explain the discrepancy between the FG model and experiment. The most likely explanation for the overestimated FG diffusion coefficients is the underestimated densities shown in Figure 4. The AFM-CG model shows faster translational diffusion than the FG model at all temperatures studied, with the highest relative error of less than 40% at 280 K. Faster dynamics compared with the underlying FG model is a general feature of CG models with only conservative forces[66,67] and can be explained by the coarse-graining process integrating out degrees of freedom of the FG system that act as sources of dissipation.

Rotational diffusion coefficients were calculated from the decay of orientational time correlation functions (OTCFs) of the molecular axes. The OTCF for a molecular axis i is $C_i(t) = \langle \hat{\boldsymbol{u}}_i(t') \cdot \hat{\boldsymbol{u}}_i(t+t') \rangle$, where $\hat{\boldsymbol{u}}_i(t')$ is the unit vector aligned with axis i at time t'. For diffusive motion, the OTCF decays exponentially with time t with a decay constant $2D_{\mathrm{R},i}$, where $D_{\mathrm{R},i}$ is the rotational diffusion coefficient of this axis ($C_i(t) \sim \exp(-2D_{\mathrm{R},i}t)$).[70] For benzene, the rotational diffusion coefficients of the out-of-plane axis and an in-plane molecu-

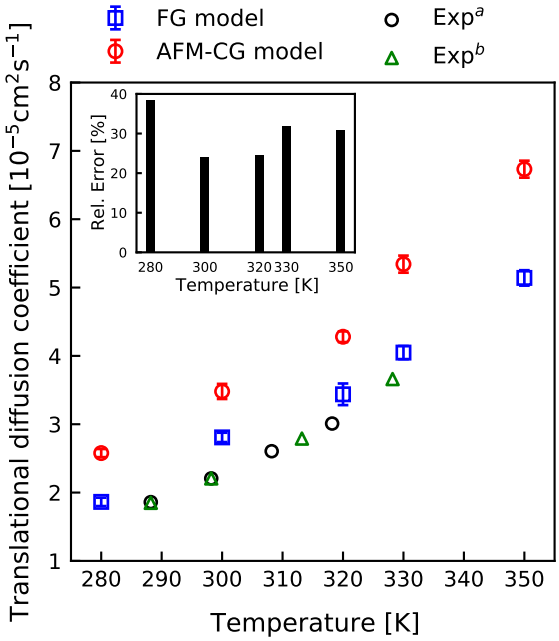

Figure 7: Translational diffusion coefficient at 1 atm as a function of temperature for the FG and AFM-CG benzene models and from experiments (Exp^a: ref 68; Exp^b: ref 69). The relative error between the AFM-CG and FG values is plotted in the inset.

lar axis were calculated from an exponential fit to the OTCFs (given in Figures S10 and S11 in the SI) and are compared at various temperatures for the FG and AFM-CG models in Figure 8. The faster rotational dynamics of the in-plane axis compared with the out-of-plane axis for the FG model at all temperatures is consistent with results from several previous experimental[71–73] and atomistic simulation studies[74–76] of the dynamics of benzene. However, the opposite behavior is observed for the CG model, with the diffusion coefficients for the out-of-plane and in-plane axes respectively overestimated and underestimated compared with the FG model. The explanation for this discrepancy is unclear and is the subject of ongoing work.

4.2 Perylene

An all-atom FG simulation of 1000 perylene molecules was conducted at 570 K and 1 atm to parametrize the CG model, in which each perylene molecule was mapped to a single CG

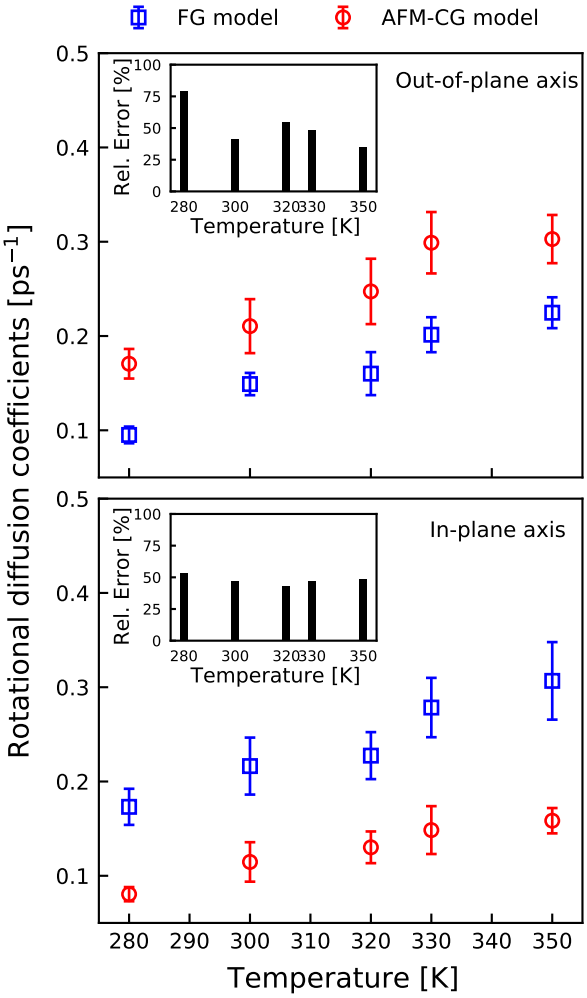

Figure 8: Rotational diffusion coefficient at 1 atm versus temperature for out-of-plane and in-plane axes for FG and AFM-CG benzene models. The relative error between the AFM-CG and FG values is plotted in the insets.

particle. Simulations of the same system were also carried out at 20 K temperature increments from 550 to 670 K to test the transferability of the CG model to other thermodynamic conditions. The simulation timestep was 1.5 fs. The total duration of each simulation was 25 ns, with data from the last 10 ns used for analysis. FG coordinates and forces from every 5000 timesteps (7.5 ps) of the simulation at 570 K were used as input for CG parametrization. The cut-off for non-bonded CG interactions was 19 Å. The optimized basis expansion and modified S-function fit of the AFM-CG pair potential for perylene are plotted as a function of intermolecular distance for the face-to-face, edge-to-face, and edge-to-edge configurations

in Figure 9. The potential well for the face-to-face configuration is significantly deeper than the other two, indicating a preference for this configuration due to strong molecular π–π stacking.

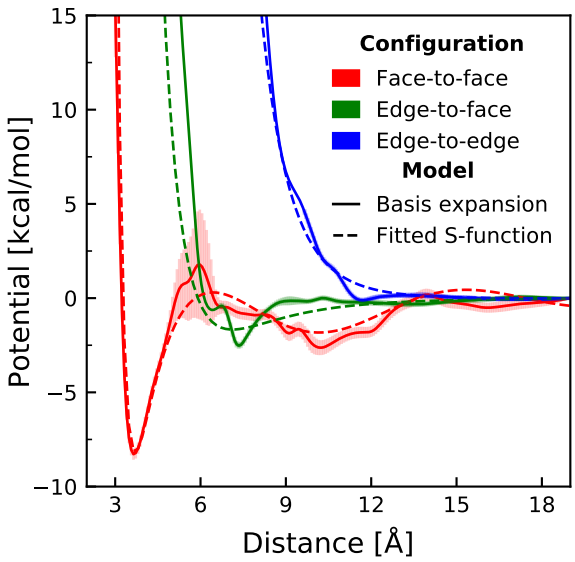

Figure 9: Perylene AFM-CG pair potential for various dimer configurations calculated using basis set expansion (solid lines with error bars) and fitted modified S-function potential (dashed lines).

The force-matching RMSE for perylene was 6.63 kcal mol^{-1} Å^{-1}, which is more than twice that obtained for benzene. The higher RMSE is in part due to the larger size of the perylene molecule compared with benzene resulting in larger intermolecular forces and in part due to coarse-graining perylene molecules into uniaxial particles, which neglects their biaxial shape. The modified S-function potential for perylene was obtained by fitting the basis expansion of the AFM-CG potential for 500 000 dimer configurations sampled from the equilibrium distribution of the FG simulation. The additional exponential potential in the fit potential (see eqs (36) and (38)) allows the modified S-function to fit both the deep well for the face-to-face configuration, and the relatively repulsive interactions for the edge-to-face and edge-to-edge configurations, as shown in Figure 9. The fit is generally good, with a RMSE between the final fitted model and the FG forces of 7.10 kcal mol^{-1} Å^{-1} which is 7% higher than for the basis expansion. CG potential curves for other configurations not shown in

Figure 9 can be found in Figure S13 in the SI. The AFM-CG potentials with and without the pressure correction are almost identical, as shown in Figure S14 in the SI. The optimized simulation parameters are given in Table S4 in the SI.

NPT CG simulations of 1000 CG perylene molecules were carried out with the optimized modified S-function potential at 1 atm pressure and the same temperatures as the FG simulations using a time step of 10 fs. CG simulations were also carried out of two existing biaxial ellipsoidal models of perylene developed by Babadi et al.[25] and by Berardi et al.[77] under the same conditions and compared to the FG and AFM-CG models (we will refer to these models as the "Babadi CG model" and the "Berardi CG model"). These CG models of perylene were parametrized from atomistic FG models by matching the interaction energy between pairs of rigid perylene molecules in a variety of configurations, which neglects finite-temperature many-body effects that are taken into account in the AFM-CG method. It should be noted that the Babadi and Berardi CG models were parametrized using different FG models (MM3[78] and the Universal Force Field (UFF),[79] respectively) from the OPLS-AA FG model used here. So the comparison with the FG and AFM-CG models in this work is not completely equitable and for this reason we also compare with experiment below. For temperatures 550–670 K, which are approximately at or above the experimental melting point of perylene of 554 K,[80] NPT simulations of the models showed that both the FG and AFM-CG models were in the liquid phase, whereas the Babadi and Berardi CG models rapidly crystallized and remained solid at all temperatures, as illustrated in the simulation snapshots of the models at $T = 570$ K in Figure 10, together with the RDFs at 570, 610, and 670 K. The RDFs of the AFM-CG model show a stronger preference for parallel packing than the FG model at separations less than 5 Å, especially at lower temperatures, but the peak and trough positions agree well with those for the FG model. The stronger alignment in the AFM-CG model simulations can also be seen in the simulation snapshots and could be due to the more symmetric uniaxial shape and reduced molecular flexibility of the CG model compared with the FG model enhancing short-range molecular packing. On

the other hand, the RDFs of the Babadi and Berardi CG models differ drastically from those for the FG model due to these CG systems being in the solid phase across this temperature range and thus these models do not accurately describe the phase behavior of perylene.

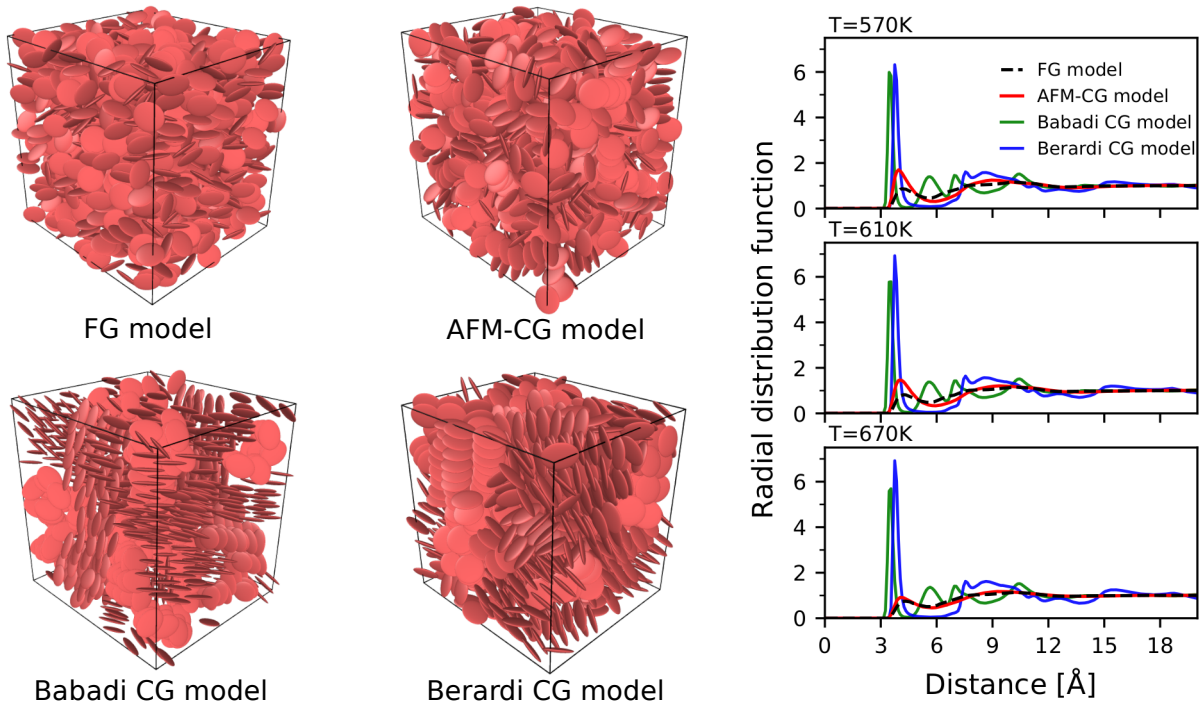

Figure 10: Simulation snapshots for different perylene models at 570 K and 1 atm, and RDFs of perylene at 1 atm and temperatures of 570, 610 and 670 K for the FG, AFM-CG, Babadi CG and Berardi CG models.

The average density of the AFM-CG model agrees significantly better with the FG model than the Babadi or Berardi CG models at 1 atm over the 550–670 K temperature range, as shown in Figure 11. The average density of the AFM-CG model is within 4% of the FG value for 550–630 K and within 9% for the higher temperatures, which are close to the experimental boiling point of 673 K.[81] On the other hand, the average densities for the Berardi and Babadi CG models are, respectively, more than one-and-a-half and two times those of the FG system over the same temperature range. The average densities of the Berardi and Babadi CG models are also unrealistic compared with experimental densities of perylene in the solid state, which range from 1.29 to 1.40 g cm^{-3} over the temperature of 440–140 K.[82] (Experimental densities in its liquid state are not available for direct comparison

to the values obtained for the FG and CG models.) NVT simulations of the CG models with density set to the average density of the NPT FG simulations were also carried out at temperatures of 570, 610 and 670 K, with similar phase behavior for each system observed, as shown in section S6.3 of the SI, indicating that the significant ordering in the Babadi and Berardi CG models is not simply due to their higher density.

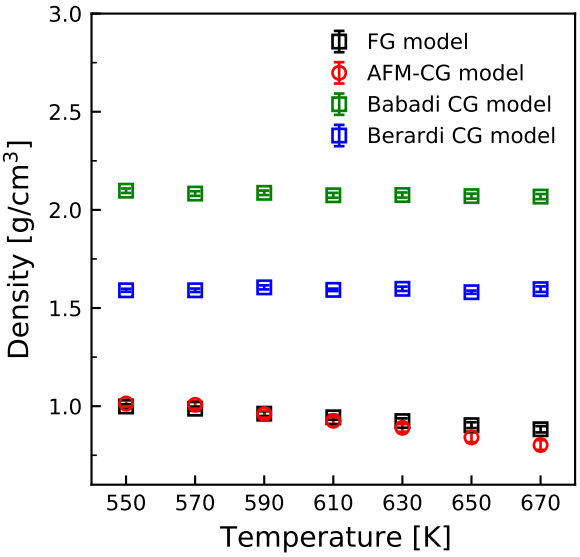

Figure 11: Perylene density at 1 atm for temperature from 550 to 670 K for the FG model and the AFM-CG, Babadi and Berardi CG models.

The ARDFs for the different models at 570 and 670 K in NPT simulations at 1 atm and NVT simulations with average density matching those of the FG systems are compared and presented in Figures S16–S23 in the SI. Because the AFM-CG model shows stronger molecular parallel packing than the FG model, the ARDFs for the AFM-CG model also have larger peaks at short range around 4 Å. Figure 12 compares slices of the ARDFs of the different models at several specific angles between the out-of-plane principal axes of perylene pairs, in which the height of the most intense peak (for the parallel configuration at 0°) for the AFM-CG, Babadi and Berardi CG models have been scaled to match that of the FG model. Only distributions at 0, 10 and 20° are shown as the probability of other angles is not significant due to the strong parallel alignment of perylene molecules in all models.

The scaled ARDF of the AFM-CG model agrees well with the FG model, indicating that the relative probability of orientations of molecular pairs at fixed distance are similar in the two models, even if the absolute probabilities do not match quantitatively. On the other hand, only the parallel (0°) configuration occurs with significant probability in the Babadi CG model. The Berardi CG model shows non-negligible probability for molecular pairs at 10°, but the relative heights and positions of the peaks do not match those for the FG model very well.

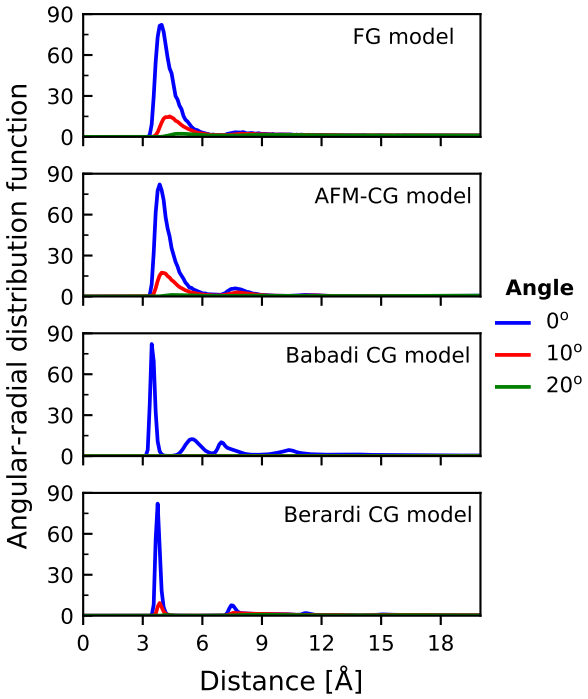

Figure 12: 1D slices of scaled ARDFs at angles of 0, 10 and 20° between out-of-plane principal axes for FG, AFM-CG, Babadi, and Berardi CG perylene models at 570 K and 1 atm. The scaling factors for the AFM-CG, Babadi and Berardi CG models are 1.70, 31.55 and 9.94, respectively.

5 Conclusion

In summary, we have developed a systematic bottom-up method for coarse-graining molecular simulations models using anisotropic coarse-grained (CG) particles, called the anisotropic

force-matching coarse-graining (AFM-CG) method, which generalizes the multiscale coarse-graining (MS-CG) method for parametrizing CG models with isotropic (spherical) particles via force matching. We have derived rigorous conditions for the CG masses, moments of inertia, and forces that must satisfy to ensure thermodynamic consistency in the canonical equilibrium ensemble between the aniostropic CG model and the fine-grained (FG) model from which it is derived, given the interaction potential between FG particles and a definition CG position and orientation coordinates in terms of FG coordinates. As in the MS-CG method for isotropic CG particles, the force on an anisotropic CG particle in the AFM-CG method must equal the average total force acting on the FG particles mapped onto the CG particle for equivalent FG and CG configurations. We have also presented an algorithm for optimizing the CG potential subject to this force-matching condition using configuration and force data from a simulation of the FG model, for the case of a pair-additive CG potential expressed as a linear combination of position- and orientation-dependent basis functions.

We have applied the AFM-CG method to parametrize uniaxial CG models of benzene and perylene in the liquid phase. Force-matching of both models gave reasonable force errors, with the error for the AFM-CG benzene model comparable with that obtained previously for a multi-site CG benzene model that considered many-body interactions.[59] As the AFM-CG method uses the pair-additive approximation for the non-bonded potential, the AFM-CG potential was fitted to an analytical form of pair-wise potential to enhance simulation speed. A linear correction was added to the fitted CG potential to match the FG NPT simulation pressure for CG simulations in the NPT ensemble. The fitted AFM-CG benzene model reproduced the FG liquid structures well, especially compared with another recent single-site model of benzene.[44] Reasonable agreement with the FG benzene model was also obtained for system density and dynamical properties. Although the CG model was not optimized for dynamics, the translational and rotational diffusion coefficients were within a factor-of-two of the FG values; however, the relative rate of rotational diffusion of the out-of-plane and in-plane molecular axes differed qualitatively from the FG model. The perylene CG

model showed significant improvement over previous single-site CG models,[25,77] in terms of both structural and thermodynamic properties. Even though there was stronger molecular alignment in the AFM-CG model compared with the FG model, which is in part due to the more symmetric CG uniaxial shape, agreement between the models for the relative probability of orientations of molecular pairs was good. Improved accuracy of the AFM-CG perylene model could be obtained using biaxial CG particles.

Acknowledgement

HTLN thanks the University of Adelaide for a PhD. scholarship. The research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government, and from The University of Adelaide's Phoenix High Performance Computing service. We thank Alec Bowen for providing us with the LAMMPS code for the S-function pair potential used in ref 44, which was modified for this work.

Supporting Information Available

Full derivation of force-matching equation, full derivation of expressions for CG mass and inertia tensor, description of scalar variables and basis functions used to define CG potential, description of least-squares solver used to optimize basis function coefficients in CG potential, description of modified S-function potential and derivatives used to calculate forces and torques in CG MD simulations, CG model parameters for benzene and perylene, and additional results from MD simulations of FG and CG models of benzene and perylene. Python code used to parametrize the CG potential using particle coordinates and forces from a FG MD simulation trajectory is available at https://github.com/dmhuanglab/afmcg.

References

(1) Noid, W. G. Perspective: Coarse-grained Models for Biomolecular Systems. *J. Chem. Phys.* **2013**, *139*, 090901.

(2) Saunders, M. G.; Voth, G. A. Coarse-graining Methods for Computational Biology. *Ann. Rev. Biophys.* **2013**, *42*, 73–93.

(3) Kmiecik, S.; Kouza, M.; Badaczewska-Dawid, A. E.; Kloczkowski, A.; Kolinski, A. Modeling of Protein Structural Flexibility and Large-scale Dynamics: Coarse-grained Simulations and Elastic Network Models. *Int. J. Mol. Sci.* **2018**, *19*, 3496.

(4) Sadeghi, M.; Noé, F. Large-scale Simulation of Biomembranes Incorporating Realistic Kinetics into Coarse-grained Models. *Nat. Commun.* **2020**, *11*, 1–13.

(5) Lee, C.-K.; Pao, C.-W. Nanomorphology Evolution of P3HT/PCBM Blends during Solution-processing from Coarse-grained Molecular Simulations. *J. Phys. Chem. C* **2014**, *118*, 11224–11233.

(6) Morita, H.; Tanaka, K.; Kajiyama, T.; Nishi, T.; Doi, M. Study of the Glass Transition Temperature of Polymer Surface by Coarse-grained Molecular Dynamics Simulation. *Macromolecules* **2006**, *39*, 6233–6237.

(7) Huang, D. M.; Faller, R.; Do, K.; Moulé, A. J. Coarse-grained Computer Simulations of Polymer/fullerene Bulk Heterojunctions for Organic Photovoltaic Applications. *J. Chem. Theory Comput.* **2010**, *6*, 526–537.

(8) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving Effective Mesoscale Potentials from Atomistic Simulations. *J. Comput. Chem.* **2003**, *24*, 1624–1636.

(9) Lu, L.; Izvekov, S.; Das, A.; Andersen, H. C.; Voth, G. A. Efficient, Regularized, and Scalable Algorithms for Multiscale Coarse-graining. *J. Chem. Theory Comput.* **2010**, *6*, 954–965.

(10) Chaimovich, A.; Shell, M. S. Coarse-graining Errors and Numerical Optimization Using a Relative Entropy Framework. *J. Chem. Phys.* **2011**, *134*, 094112.

(11) Brini, E.; Algaer, E. A.; Ganguly, P.; Li, C.; Rodríguez-Ropero, F.; van der Vegt, N. F. Systematic Coarse-graining Methods for Soft Matter Simulations – A Review. *Soft Matter* **2013**, *9*, 2108–2119.

(12) Dunn, N. J. H.; Noid, W. G. Bottom-up Coarse-grained Models That Accurately Describe the Structure, Pressure, and Compressibility of Molecular Liquids. *J. Chem. Phys.* **2015**, *143*, 243148.

(13) Morriss-Andrews, A.; Rottler, J.; Plotkin, S. S. A Systematically Coarse-grained Model for DNA and Its Predictions for Persistence Length, Stacking, Twist, and Chirality. *J. Chem. Phys.* **2010**, *132*, 035105.

(14) Doye, J. P. K.; Ouldridge, T. E.; Louis, A. A.; Romano, F.; Sulc, P.; Matek, C.; Snodin, B. E. K.; Rovigatti, L.; Schreck, J. S.; Harrison, R. M.; Smith, W. P. J. Coarse-graining DNA for Simulations of DNA Nanotechnology. *Phys. Chem. Chem. Phys.* **2013**, *15*, 20395–20414.

(15) Berardi, R.; Muccioli, L.; Orlandi, S.; Ricci, M.; Zannoni, C. Computer Simulations of Biaxial Nematics. *J. Phys.: Condens. Matter* **2008**, *20*, 463101.

(16) Quinn, J. T.; Zhu, J.; Li, X.; Wang, J.; Li, Y. Recent Progress in the Development of n-Type Organic Semiconductors for Organic Field Effect Transistors. *J. Mater. Chem. C* **2017**, *5*, 8654–8681.

(17) Boehm, B. J.; Nguyen, H. T. L.; Huang, D. M. The Interplay of Interfaces, Supramolecular Assembly, and Electronics in Organic Semiconductors. *J. Phys. Condens. Matter* **2019**, *31*, 423001.

(18) Jackson, N. E. Coarse-Graining Organic Semiconductors: The Path to Multiscale Design. *The Journal of Physical Chemistry B* **2021**, *125*, 485–496.

(19) Bronstein, H.; Nielsen, C. B.; Schroeder, B. C.; McCulloch, I. The Role of Chemical Design in the Performance of Organic Semiconductors. *Nat. Rev. Chem.* **2020**, *4*, 66–77.

(20) Huang, Y.; Elder, D. L.; Kwiram, A. L.; Jenekhe, S. A.; Jen, A. K.; Dalton, L. R.; Luscombe, C. K. Organic Semiconductors at the University of Washington: Advancements in Materials Design and Synthesis and toward Industrial Scale Production. *Adv. Mater.* **2021**, *33*, 1904239.

(21) Vandewal, K.; Tvingstedt, K.; Gadisa, A.; Inganäs, O.; Manca, J. V. Relating the Open-circuit Voltage to Interface Molecular Properties of Donor:Acceptor Bulk Heterojunction Solar Cells. *Phys. Rev. B* **2010**, *81*, 125204.

(22) Tumbleston, J. R.; Collins, B. A.; Yang, L.; Stuart, A. C.; Gann, E.; Ma, W.; You, W.; Ade, H. The Influence of Molecular Orientation on Organic Bulk Heterojunction Solar Cells. *Nat. Photon.* **2014**, *8*, 385–391.

(23) Min, J.; Jiao, X.; Ata, I.; Osvet, A.; Ameri, T.; Bäuerle, P.; Ade, H.; Brabec, C. J. Time-Dependent Morphology Evolution of Solution-Processed Small Molecule Solar Cells during Solvent Vapor Annealing. *Adv. Energy Mater.* **2016**, *6*, 1502579.

(24) Gay, J. G.; Berne, B. J. Modification of the Overlap Potential to Mimic a Linear Site–site Potential. *J. Chem. Phys.* **1981**, *74*, 3316–3319.

(25) Babadi, M.; Everaers, R.; Ejtehadi, M. R. Coarse-grained Interaction Potentials for Anisotropic Molecules. *J. Chem. Phys.* **2006**, *124*, 174708.

(26) Haxton, T. K.; Mannige, R. V.; Zuckermann, R. N.; Whitelam, S. Modeling Sequence-specific Polymers Using Anisotropic Coarse-grained Sites Allows Quantitative Comparison with Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 303–315.

(27) Ricci, M.; Roscioni, O. M.; Querciagrossa, L.; Zannoni, C. MOLC. A Reversible Coarse Grained Approach Using Anisotropic Beads for the Modelling of Organic Functional Materials. *Phys. Chem. Chem. Phys.* **2019**, *21*, 26195–26211.

(28) Tanis, I.; Rousseau, B.; Soulard, L.; Lemarchand, C. A. Assessment of an Anisotropic Coarse-grained Model for cis-1, 4-Polybutadiene: A Bottom-up Approach. *Soft Matter* **2021**, *17*, 621–636.

(29) Cohen, A. E.; Jackson, N. E.; De Pablo, J. J. Anisotropic Coarse-Grained Model for Conjugated Polymers: Investigations into Solution Morphologies. *Macromolecules* **2021**, *54*, 3780–3789.

(30) Li, G.; Shen, H.; Zhang, D.; Li, Y.; Wang, H. Coarse-Grained modeling of nucleic acids using anisotropic Gay–Berne and electric multipole potentials. *J. Chem. Theory Comput.* **2016**, *12*, 676–693.

(31) Tillack, A. F.; Johnson, L. E.; Eichinger, B. E.; Robinson, B. H. Systematic Generation of Anisotropic Coarse-grained Lennard-Jones Potentials and Their Application to Ordered Soft Matter. *J. Chem. Theory Comput.* **2016**, *12*, 4362–4374.

(32) Lee, C. K.; Hua, C. C.; Chen, S. A. An Ellipsoid-chain Model for Conjugated Polymer Solutions. *J. Chem. Phys.* **2012**, *136*, 084901–11.

(33) Lee, C.-K.; Pao, C.-W. Multiscale Molecular Simulation of Solution Processing of SMDPPEH:PCBM Small-Molecule Organic Solar Cells. *ACS Appl. Mater. Interfaces* **2016**, *8*, 20691–20700.

(34) Tripathy, M.; Agarwal, U.; Kumar, P. S. Toward Transferable Coarse-Grained Potentials for Poly-Aromatic Hydrocarbons: A Force Matching Approach. *Macromol. Theory. Simul.* **2019**, *28*, 1800040.

(35) Noid, W. G.; Chu, J.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. The Multiscale Coarse-graining Method. I. A Rigorous Bridge between Atomistic and Coarse-grained Models. *J. Chem. Phys.* **2008**, *128*, 244114.

(36) Noid, W. G.; Liu, P.; Wang, Y.; Chu, J. W.; Ayton, G. S.; Izvekov, S.; Andersen, H. C.; Voth, G. A. The Multiscale Coarse-graining Method. II. Numerical Implementation for Coarse-grained Molecular Models. *J. Chem. Phys.* **2008**, *128*, 244115.

(37) Gray, C. G.; Gubbins, K. E. *Theory of Molecular Fluids: Vol. 1: Fundamentals*; Clarendon Press: Oxford, 1984.

(38) Tschöp, W.; Kremer, K.; Batoulis, J.; Bürger, T.; Hahn, O. Simulation of Polymer Melts. I. Coarse-graining Procedure for Polycarbonates. *Acta Polym.* **1998**, *49*, 61–74.

(39) Rühle, V.; Junghans, C. Hybrid Approaches to Coarse-Graining using the VOTCA Package: Liquid Hexane. *Macromol. Theory Simul.* **2011**, *20*, 472–477.

(40) Potter, T. D.; Tasche, J.; Wilson, M. R. Assessing the Transferability of Common Top-down and Bottom-up Coarse-grained Molecular Models for Molecular Mixtures. *Phys. Chem. Chem. Phys.* **2019**, *21*, 1912–1927.

(41) Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*, 1st ed.; Prentice-Hall: Englewood Cliffs, 1974.

(42) Zewdie, H. Computer Simulation Studies of Liquid Crystals: A New Corner Potential for Cylindrically Symmetric Particles. *J. Chem. Phys.* **1998**, *108*, 2117–2133.

(43) Cinacchi, G.; Tani, A. Computer Simulations of Pure and Mixed Systems of Disklike Particles Interacting with the S-function Corner Potential. *J. Chem. Phys.* **2002**, *117*, 11388–11395.

(44) Bowen, A. S.; Jackson, N. E.; Reid, D. R.; de Pablo, J. J. Structural Correlations and

Percolation in Twisted Perylene Diimides Using a Simple Anisotropic Coarse-Grained Model. *J. Chem. Theory Comput.* **2018**, *14*, 6495–6504.

(45) Das, A.; Andersen, H. C. The Multiscale Coarse-graining Method. V. Isothermal-isobaric Ensemble. *J. Phys. Chem. Phys.* **2010**, *132*, 164106.

(46) Sanyal, T.; Shell, M. S. Transferable Coarse-grained Models of Liquid–liquid Equilibrium Using Local Density Potentials Optimized with the Relative Entropy. *J. Phys. Chem. B* **2018**, *122*, 5678–5693.

(47) Jin, J.; Yu, A.; Voth, G. A. Temperature and Phase Transferable Bottom-up Coarse-grained Models. *J. Chem. Theory Comput.* **2020**, *16*, 6823–6842.

(48) Plimpton, S. Fast Parallel Algorithms for Short-range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(49) Brown, W. M.; Wang, P.; Plimpton, S. J.; Tharrington, A. N. Implementing Molecular Dynamics on Hybrid High Performance Computers – Short Range Forces. *Comput. Phys. Commun.* **2011**, *182*, 898–911.

(50) Brown, W. M.; Kohlmeyer, A.; Plimpton, S. J.; Tharrington, A. N. Implementing Molecular Dynamics on Hybrid High Performance Computers – Particle–Particle Particle–Mesh. *Comput. Phys. Commun.* **2012**, *183*, 449–459.

(51) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(52) McDonald, N. A.; Jorgensen, W. L. Development of an All-atom Force Field for Heterocycles. Properties of Liquid Pyrrole, Furan, Diazoles, and Oxazoles. *J. Phys. Chem. B* **1998**, *102*, 8049–8059.

(53) Rizzo, R. C.; Jorgensen, W. L. OPLS All-atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.

(54) Price, M. L.; Ostrovsky, D.; Jorgensen, W. L. Gas-phase and Liquid-state Properties of Esters, Nitriles, and Nitro Compounds with the OPLS-AA Force Field. *J. Comput. Chem.* **2001**, *22*, 1340–1352.

(55) Jorgensen, W. OPLS All-Atom Parameters for Organic Molecules, Ions, Peptides & Nucleic Acids. *Yale University* **2009**,

(56) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles*, 1st ed.; CRC Press: Boca Raton, 1988.

(57) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(58) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-space Distributions. *Phys. Rev. A* **1985**, *31*, 1695.

(59) John, S. T.; Csányi, G. Many-body Coarse-grained Interactions Using Gaussian Approximation Potentials. *J. Phys. Chem. B* **2017**, *121*, 10934–10949.

(60) Arunan, E.; Gutowsky, H. S. The Rotational Spectrum, Structure and Dynamics of a Benzene Dimer. *J. Chem. Phys.* **1993**, *98*, 4294–4296.

(61) Headen, T. F.; Howard, C. A.; Skipper, N. T.; Wilkinson, M. A.; Bowron, D. T.; Soper, A. K. Structure of π–π Interactions in Aromatic Liquids. *J. Am. Chem. Soc.* **2010**, *132*, 5735–5742.

(62) Sinnokrot, M. O.; Sherrill, C. D. Highly Accurate Coupled Cluster Potential Energy Curves for the Benzene Dimer: Sandwich, T-shaped, and Parallel-displaced Configurations. *J. Phys. Chem. A* **2004**, *108*, 10200–10207.

(63) Beg, S. A.; Tukur, N. M.; Al-Harbi, D. K.; Hamad, E. Z. Saturated Liquid Densities of Benzene, Cyclohexane, and Hexane from 298.15 to 473.15 K. *J. Chem. Eng. Data* **1993**, *38*, 461–464.

(64) Sun, T. F.; Schouten, J. A.; Trappeniers, N. J.; Biswas, S. N. Measurements of the Densities of Liquid Benzene, Cyclohexane, Methanol, and Ethanol As Functions of Temperature at 0.1 MPa. *J. Chem. Thermodyn.* **1988**, *20*, 1089–1096.

(65) Yeh, I.-C.; Hummer, G. System-size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

(66) Li, Y.; Abberton, B. C.; Kröger, M.; Liu, W. K. Challenges in Multiscale Modeling of Polymer Dynamics. *Polymers* **2013**, *5*, 751–832.

(67) Rondina, G. G.; Böhm, M. C.; Müller-Plathe, F. Predicting the Mobility Increase of Coarse-Grained Polymer Models from Excess Entropy Differences. *J. Chem. Theory Comput.* **2020**, *16*, 1431–1447.

(68) Collings, A. F.; Mills, R. Temperature-dependence of Self-diffusion for Benzene and Carbon Tetrachloride. *Trans. Faraday Soc.* **1970**, *66*, 2761–2766.

(69) McCool, M. A.; Collings, A. F.; Woolf, L. A. Pressure and Temperature Dependence of the Self-diffusion of Benzene. *J. Chem. Soc., Faraday Trans. 1* **1972**, *68*, 1489–1497.

(70) Beckmann, P. A. Spectral Densities and Nuclear Spin Relaxation in Solids. *Phys. Rep.* **1988**, *171*, 85–128.

(71) Tanabe, K. Determination of Rotational Diffusion Constants of Liquid Benzene from Measurements of Infrared and Raman Line Widths. *Chem. Phys. Lett.* **1979**, *63*, 43–46.

(72) Tanabe, K. Raman Study of Reorientational Motion of Liquid Benzene. *Chem. Phys.* **1978**, *31*, 319–325.

(73) Bauer, D. R.; Alms, G. R.; Brauman, J. I.; Pecora, R. Depolarized Rayleigh Scattering And 13C NMR Studies of Anisotropic Molecular Reorientation of Aromatic Compounds in Solution. *J. Chem. Phys.* **1974**, *61*, 2255–2261.

(74) Witt, R.; Sturz, L.; Dölle, A.; Müller-Plathe, F. Molecular dynamics of benzene in neat liquid and a solution containing polystyrene. 13C nuclear magnetic relaxation and molecular dynamics simulation results. *J. Phys. Chem. A* **2000**, *104*, 5716–5725.

(75) Linse, P.; Engström, S.; Jönsson, B. Molecular Dynamics Simulation of Liquid and Solid Benzene. *Chem. Phys. Lett.* **1985**, *115*, 95–100.

(76) Steinhauser, O. On the Structure and Dynamics of Liquid Benzene. *Chemi. Phys.* **1982**, *73*, 155–167.

(77) Berardi, R.; Fava, C.; Zannoni, C. A Gay–Berne Potential for Dissimilar Biaxial Particles. *Chem. Phys. Lett.* **1998**, *297*, 8–14.

(78) Lii, J. H.; Allinger, N. L. Molecular Mechanics. The MM3 Force Field For Hydrocarbons .3. The van der Waals Potentials And Crystal Data For Aliphatic And Aromatic Hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8576–8582.

(79) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

(80) Casellato, F.; Vecchi, C.; Girelli, A.; Casu, B. Differential Calorimetric Study of Polycyclic Aromatic Hydrocarbons. *Thermochim. Acta* **1973**, *6*, 361–368.

(81) Karaffa, L. S. *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*; RSC Publishing: Cambridge, 2013.

(82) Botoshansky, M.; Herbstein, F. H.; Kapon, M. Towards a Complete Description of a

Polymorphic Crystal: The Example of Perylene: Redetermination of the Structures of the (Z = 2 and 4) Polymorphs. *Helv. Chim. Acta* **2003**, *86*, 1113–1128.

Graphical TOC Entry