Supporting Information

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

Huong T.L. Nguyen and David M. Huang

Contents

S1 Derivation of Force-Matching Equation .............................................. S2
S2 Derivation of Expressions for CG Mass and Inertia Tensor ....................... S4
S3 Force Basis Functions and Least-Squares Solver .................................. S9
  S3.1 Scalar Variables of Basis Functions ........................................... S9
  S3.2 Multivariate Potential and Force Basis Functions ................................. S9
  S3.3 Basis Functions for Benzene and Perylene Parametrizations ................. S11
  S3.4 Least-Squares Solver .................................................................. S11
S4 Modified S-function Fit to Anisotropic CG Potential .............................. S12
  S4.1 Modified S-function Force and Torque Calculations ............................ S13
S5 Benzene Simulations ........................................................................... S15
  S5.1 Distributions of Principal Moments of Inertia ................................. S15
  S5.2 AFM-CG Model Parameters ....................................................... S16
  S5.3 NVT Simulations ..................................................................... S19
    S5.3.1 Radial Distribution Function (RDF) ............................................ S19
    S5.3.2 Angular–Radial Distribution Function (ARDF) .............................. S20
  S5.4 NPT Simulations ..................................................................... S21
    S5.4.1 Radial Distribution Function (RDF) ............................................ S21
    S5.4.2 Angular–Radial Distribution Function (ARDF) .............................. S22
    S5.4.3 Mean Squared Displacement (MSD) ........................................... S23
    S5.4.4 Orientational Time Correlation Function (OTCF) ......................... S24
S6 Perylene Simulations ........................................................................... S25
  S6.1 Distributions of Principal Moments of Inertia ................................. S25
  S6.2 AFM-CG Model Parameters ....................................................... S26
  S6.3 NVT Simulations ..................................................................... S29
    S6.3.1 Radial Distribution Function (RDF) ............................................ S29
    S6.3.2 Angular–Radial Distribution Function (ARDF) .............................. S30
  S6.4 NPT Simulations ..................................................................... S32
    S6.4.1 Angular–Radial Distribution Function (ARDF) .............................. S32
S1 Derivation of Force-Matching Equation

In the main paper, the coarse-grained (CG) potential energy function \( U \) was derived from matching the fine-grained (FG) and CG equilibrium configurational probability distributions. An expression for the force on CG particle \( I \) was subsequently derived using the expression for the CG potential and the definitions of the CG configurational mapping from the FG coordinates. Here, a full derivation of the CG force equation is provided.

Enforcing consistency between FG and CG equilibrium configurational distributions requires that

\[
P_{CG,c}(R^N, \Omega^N) = \int \, dr^n \, P_{FG,c}(r^n) \delta \left( M_R^N (r^n) - R^N \right) \delta \left( M_{\Omega}^N (r^n) - \Omega^N \right),
\]

(S1)

which can be written in terms of the FG and CG potential energy functions, \( u(r^n) \) and \( U(R^N, \Omega^N) \), respectively, using eqs (3) and (5) of the main paper as

\[
\exp \left( -\frac{U(R^N, \Omega^N)}{k_B T} \right) / Z_{CG} = \frac{1}{Z_{FG}} \int \, dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta \left( M_R^N (r^n) - R^N \right) \delta \left( M_{\Omega}^N (r^n) - \Omega^N \right),
\]

(S2)

where \( Z_{FG} = \int \, dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \) and \( Z_{CG} = \int \, d\Omega^N \exp \left( -\frac{U(R^N, \Omega^N)}{k_B T} \right) \) are the configurational partition functions of the FG and CG systems, respectively. Equation (S2) can be rearranged to give

\[
U(R^N, \Omega^N) = -k_B T \ln Z(R^N, \Omega^N) - k_B T \ln \left( \frac{Z_{CG}}{Z_{FG}} \right),
\]

(S3)

and the second term in eq (S3) is independent of FG or CG coordinates. Using eqs (S3) and (S4), the force on CG particle \( I \) is

\[
F_I (R^N, \Omega^N) = -\frac{\partial U(R^N, \Omega^N)}{\partial R_I} = \frac{k_B T}{Z(R^N, \Omega^N)} \frac{\partial Z(R^N, \Omega^N)}{\partial R_I}
\]

\[
= \frac{k_B T}{Z(R^N, \Omega^N)} \frac{\partial}{\partial R_I} \left[ \int \, dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta \left( M_R^N (r^n) - R^N \right) \delta \left( M_{\Omega}^N (r^n) - \Omega^N \right) \right]
\]

\[
= \frac{k_B T}{Z(R^N, \Omega^N)} \int \, dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta \left( M_R^N (r^n) - \Omega^N \right)
\]

\[
\times \prod_{j \neq I} \delta (M_{IR} (r^n) - R_j) \frac{\partial}{\partial R_I} \delta (M_{IR} (r^n) - R_I)
\]

(S5)

Using the definition of the linear mapping from FG to CG positions of the \( n_I \) FG particles in the set \( \zeta_I \) that are mapped onto CG particle \( I \),

\[
R_I = M_{RI} (r^n) = \frac{\sum_{i \in \zeta_I} m_i r_i}{\sum_{i \in \zeta_I} m_i},
\]

(S6)

and applying the chain rule, the delta function derivative in eq (S5) can be recast in terms of FG coordinates as

\[
\frac{\partial}{\partial R_I} \delta (M_{RI} (r^n) - R_I) = -\sum_{i \in \zeta_I} \frac{\partial}{\partial r_i} \delta (M_{RI} (r^n) - R_I).
\]

(S7)

Inserting eq (S7) into eq (S5) and integrating by parts, gives

\[
F_I (R^N, \Omega^N) = \frac{k_B T}{Z(R^N, \Omega^N)} \int \, dr^n \sum_{i \in \zeta_I} \frac{\partial}{\partial r_i} \left[ \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta \left( M_{RI}^N (r^n) - \Omega^N \right) \prod_{j \neq I} \delta (M_{RI} (r^n) - R_j) \right]
\]

\[
\times \delta (M_{RI}^N (r^n) - R_I)
\]

\[
= \frac{k_B T}{Z(R^N, \Omega^N)} \int \, dr^n \sum_{i \in \zeta_I} \frac{\partial}{\partial r_i} \left[ \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta \left( M_{RI}^N (r^n) - \Omega_I \right) \prod_{j \neq I} \delta (M_{RI} (r^n) - R_j) \right]
\]

\[
\times \delta (M_{RI}^N (r^n) - R^N).
\]

(S8)
The second equality in eq (S8) follows from the fact that derivatives with respect to positions of FG particles mapped on to CG particle \( I \) vanish for terms involving CG particles \( J \neq I \), because no FG particle contributes to more than one CG particle. The factor operated on by the derivatives in eq (S8) includes the delta function for the mapping to CG orientation coordinates for particle \( I \). As the orientation coordinates are defined by the inertia tensor with respect to the center-of-mass of the FG particles mapped onto CG particle \( I \), by the chain rule the derivatives of this delta function are proportional to the derivatives applied to the inertia tensor. Using the formula for the inertia tensor of the FG particles, mapping to CG orientational coordinates in eq (S8) is zero. This means that the only factor affected by \( \delta \) \( \sum_{j \in \zeta_I} m_i (\Delta r_i^T \Delta r_i E - \Delta r_i \Delta r_i^T) \) \( \sum_{j \in \zeta_I} \delta \frac{\partial \rho_{fg}}{\partial r_j} \sum_{j \in \zeta_I} m_i (\Delta r_i^T \Delta r_i E - \Delta r_i \Delta r_i^T) \)

\[
\sum_{j \in \zeta_I} \frac{\partial \Delta r_i}{\partial r_j} = \sum_{j \in \zeta_I} \frac{\partial}{\partial r_j} \left( r_i - \sum_{k \in \zeta_I} \frac{m_k r_k}{\sum_{k \in \zeta_I} m_k} \right) = \frac{\partial r_i}{\partial r_i} - \sum_{j \in \zeta_I} \frac{m_j \partial r_j}{\sum_{k \in \zeta_I} m_k} = E - \sum_{j \in \zeta_I} \frac{m_j}{\sum_{k \in \zeta_I} m_k} E = 0.
\]

The derivation in eq (S11) uses the fact that the derivative \( \frac{\partial \rho_{fg}}{\partial r_j} \) is an identity matrix when \( i = j \) and a zero matrix otherwise, and likewise for \( \frac{\partial \rho_{fg}}{\partial r_j} \). From eq (S11), the right-hand side of eq (S10) is zero, which implies that the sum of derivatives with respect to FG positions of the delta function for the mapping to CG orientational coordinates in eq (S8) is zero. This means that the only factor affected by the derivatives with respect to the FG positions that map on to CG particle \( I \) in eq (S8) is the Boltzmann factor of the FG potential energy function. Thus, eq (S8) becomes

\[
F_i(R^N, \Omega^N) = \frac{1}{Z(R^N, \Omega^N)} \int dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta(M_{t_j}^N (r^n) - \Omega^N) \delta(M_{t_j}^N (r^n) - R^N) \left[ \sum_{i \in \zeta_I} \frac{\partial}{\partial r_i} (-u(r^n)) \right] = \frac{1}{Z(R^N, \Omega^N)} \int dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta(M_{t_j}^N (r^n) - R^N) \delta(M_{t_j}^N (r^n) - \Omega^N) \left[ \sum_{i \in \zeta_I} f_i(r^n) \right].
\]

where \( f_i = -\frac{\partial u(r^n)}{\partial r_i} \) is the force on particle \( i \) in the FG system. Inserting \( Z(R^N, \Omega^N) \) from eq (S4) gives

\[
F_{CG,i}(R^N, \Omega^N) = \frac{\int dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta(M_{t_j}^N (r^n) - R^N) \delta(M_{t_j}^N (r^n) - \Omega^N) \left( \sum_{i \in \zeta_I} f_i(r^n) \right)}{\int dr^n \exp \left( -\frac{u(r^n)}{k_B T} \right) \delta(M_{t_j}^N (r^n) - R^N) \delta(M_{t_j}^N (r^n) - \Omega^N)} = \left( \sum_{i \in \zeta_I} f_i(r^n) \right)_{R^N, \Omega^N} = \left( F_{FG,i}(r^n) \right)_{R^N, \Omega^N}
\]

where \( \langle \cdots \rangle_{R^N, \Omega^N} \) denotes an equilibrium ensemble average over FG configurations that are mapped to CG configuration \( (R^N, \Omega^N) \) and \( F_{FG,i}(r^n) \equiv \sum_{i \in \zeta_I} f_i(r^n) \) is the total force acting on the FG particles mapped onto CG particle \( I \). Equation (S13) generalizes the force-matching condition of the MS-CG method for isotropic CG particles \(^1\) to anisotropic particles.
S2 Derivation of Expressions for CG Mass and Inertia Tensor

The mapping from FG to CG linear and angular momenta are derived from Hamilton’s equations and the definitions of the CG configuration as described in the main paper. The expressions for the CG linear and angular momenta for CG particle \( I \) in terms of FG variables are

\[
P_I = M_{FG}(p^n) = M_I \dot{R}_I = \frac{M_I}{\sum m_i} \sum p_i \tag{S14}
\]

and

\[
L_I = M_{FG}(p^n, r^n) = \mathbf{I}_I \Omega_I = \mathbf{I}_I \mathbf{I}_{FG,I}^{-1} \sum \Delta r_i \times p_i = \mathbf{I}_I \mathbf{I}_{FG,I}^{-1} \sum [\Delta r_i] p_i, \tag{S15}
\]

where \( \mathbf{I}_I \) is the inertia tensor with respect to the center-of-mass of CG particle \( I \), \( \mathbf{I}_{FG,I} \) is the inertia tensor relative to the center-of-mass of the set of FG particles mapped onto CG particle \( I \), and \( \Delta r_i \) is the position vector of the FG particles relative to their center-of-mass. The mapping \( M_{FG}(p^n, r^n) \) for the angular momentum depends not only on the momenta but also on the positions of the FG particles due to the dependence of the angular momentum on FG particle positions with respect to the center-of-mass. In the last step in eq (S15), the cross product of the FG position relative to the center-of-mass and the FG angular momentum depends not only on the momenta but also on the positions of the FG particles due to the dependence of the angular momentum on FG particle positions with respect to the center-of-mass. The mapping from FG to CG linear and angular momenta are derived from Hamilton’s equations and the definitions of the CG configuration as described in the main paper. The expressions for the CG linear and angular momenta for CG particle \( I \) are derived from Hamilton’s equations and the definitions of the CG configuration as described in the main paper.

The FG inertia tensor can be expressed in terms of this matrix as

\[
\mathbf{I}_{FG,I} = - \sum_{i \in \mathcal{G}_I} m_i [\Delta r_i]^2. \tag{S17}
\]

Enforcing consistency of the FG and CG equilibrium phase-space distributions requires that

\[
\mathbb{P}_{CG,m}(\mathbf{P}^N, \mathbf{L}^N) \mathbb{P}_{CG,c}(\mathbf{R}^N, \Omega^N) = \int d\mathbf{P}^n \int d\mathbf{p}^n \mathbb{P}_{FG,m}(\mathbf{p}^n) \mathbb{P}_{FG,c}(\mathbf{r}^n) \delta (\mathbf{M}_I^N(p^n) - \mathbf{P}^N) \delta (\mathbf{M}_{I,F}^N(p^n, r^n) - \mathbf{L}^N) \\
\times \delta (\mathbf{M}_R^N(r^n) - \mathbf{R}^N) \delta (\mathbf{M}_I^N(r^n) - \Omega^N), \tag{S18}
\]

where the normalized CG and FG momentum distributions are

\[
\mathbb{P}_{CG,m}(\mathbf{P}^N, \mathbf{L}^N) = \int d\mathbf{P}^n \int d\mathbf{L}^N \exp \left( -\frac{1}{2k_B T} \sum_{I=1}^N \left( \frac{p_i^2}{m_i} + \mathbf{L}_I^T \mathbf{I}_I^{-1} \mathbf{L}_I \right) \right) \\
= \frac{1}{(2\pi k_B T)^{3N}} \prod_{I=1}^N \left\{ \exp \left( -\frac{1}{2k_B T} \left( \frac{p_i^2}{m_i} + \mathbf{L}_I^T \mathbf{I}_I^{-1} \mathbf{L}_I \right) \right) \right\} \tag{S19}
\]

and

\[
\mathbb{P}_{FG,m}(\mathbf{p}^n) = \int d\mathbf{p}^n \exp \left( -\frac{1}{2k_B T} \sum_{i=1}^n \frac{p_i^2}{m_i} \right) \\
= \frac{1}{(2\pi k_B T)^{3n/2}} \prod_{i=1}^n \left[ \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right], \tag{S20}
\]

respectively. To derive eqs (S19) and (S20), we have used the identity

\[
\int dx \exp \left( -\frac{1}{2} \mathbf{A}^T \mathbf{x} + \mathbf{B}^T \mathbf{x} \right) = \sqrt{\frac{(2\pi)^n}{\det(\mathbf{A})}} \exp \left( \frac{1}{2} \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B} \right), \tag{S21}
\]

where \( \mathbf{A} \) is an \( m \times m \) positive-definite matrix, \( \mathbf{B} \) and \( \mathbf{x} \) are column vectors of length \( m \), and \( \det(\mathbf{A}) \) denotes the determinant of \( \mathbf{A} \), with \( \mathbf{B} = \mathbf{0} \) in the cases of eqs (S19) and (S20).
Substituting eqs (S19) and (S20) into eq (S18) and re-arranging gives

\[
(2\pi k_B T)^{(3n/2-3N)} \left\{ \prod_{i=1}^{N} \exp \left[ -\frac{1}{2m_i k_B T} \left( \frac{p_i^2}{m_i} + L_i^T M_i^{-1} L_i \right) \right] \right\} \rho_{CG,c}(R^N, \Omega^N) \\
= \int dr^n \int dp^n \left[ \prod_{i=1}^{n} \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] P_{FG,c}(r^n) \delta (M_{\Pi}^{\scriptscriptstyle N}(p^n) - P^N) \delta (M_{\Omega}^{\scriptscriptstyle N}(p^n, r^n) - L^N) \\
 \times \delta (M_{\Pi}^{\scriptscriptstyle N}(r^n) - R^N) \delta (M_{\Omega}^{\scriptscriptstyle N}(r^n) - \Omega^N) \\
= \int dr^n P_{FG,c}(r^n) \delta (M_{\Pi}^{\scriptscriptstyle N}(r^n) - R^N) \delta (M_{\Omega}^{\scriptscriptstyle N}(r^n) - \Omega^N) \\
 \times \prod_{i=1}^{N} \int \left[ \prod_{i \in \zeta_i} dp_i \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] \delta (M_{P}^{\scriptscriptstyle N}(p^n) - P_I) \delta (M_{L}^{\scriptscriptstyle N}(p^n, r^n) - L_I) \\
\tag{S22}
\]

where in the last step we have used the fact that no FG particle is mapped to more than one CG particle to factor the right-hand side of eq (S22) into a product of factors that each involves only the FG particles that are mapped to a specific CG particle. Integration of the factor that involves momenta of the $n_I$ FG particles mapped onto CG particle $I$ can be carried out analytically using the mapping operators in eqs (S14) and (S15) and the Fourier representation of the Dirac delta function, which in $m$ dimensions is

\[
\delta(x - a) = \frac{1}{(2\pi)^m} \int dk \exp \left[ i k^T (x - a) \right], \tag{S23}
\]

where $x$, $a$, and $k$ are vectors of length $m$. Using the auxiliary variables $\Psi$ and $\Phi$ in the Fourier representations of the two delta functions gives

\[
\int \left[ \prod_{i \in \zeta_i} dp_i \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] \delta (M_{P}^{\scriptscriptstyle N}(p^n) - P_I) \delta (M_{L}^{\scriptscriptstyle N}(p^n, r^n) - L_I) \\
= \int \left[ \prod_{i \in \zeta_i} dp_i \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] \delta (c_p \sum_{i \in \zeta_i} p_i - P_I) \delta (C_L \sum_{i \in \zeta_i} |\Delta r_i| p_i - L_I) \\
= \frac{1}{(2\pi)^6} \int d\Phi \int d\Psi \int \left[ \prod_{i \in \zeta_i} dp_i \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] \times \exp \left[ i\Psi^T \left( c_p \sum_{i \in \zeta_i} p_i - P_I \right) \right] \exp \left[ i\Phi^T \left( C_L \sum_{i \in \zeta_i} |\Delta r_i| p_i - L_I \right) \right] \\
= \frac{1}{(2\pi)^6} \int d\Phi \exp \left( -i\Phi^T L_I \right) \int d\Psi \exp \left( -i\Psi^T P_I \right) \prod_{i \in \zeta_i} \int dp_i \exp \left[ -\frac{p_i^2}{2m_i k_B T} + i \left( c_p \Psi^T + \Phi^T C_L |\Delta r_i| p_i \right) \right] \tag{S24}
\]

where

\[
c_p \equiv \frac{M_I}{\sum_{i \in \zeta_i} m_i} \tag{S25}
\]

and

\[
C_L \equiv I_{FG,I}^{-1}. \tag{S26}
\]

Evaluating the integrals over the FG particle momenta $p_i$ in eq (S24) using the Gaussian identity in
eq (S21) gives

\[ \prod_{i \in \zeta_i} \int dp_i \exp \left[ -\frac{p_i^2}{2m_i k_B T} + i \left( c_p \Psi^T + \Phi^T C_L [\Delta r_i] \right) p_i \right] \]

\[ = \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \exp \left[ -\frac{1}{2} m_i k_B T \left( c_p \Psi^T + \Phi^T C_L [\Delta r_i] \right)^T \left( c_p \Psi^T + \Phi^T C_L [\Delta r_i] \right) \right] \]

\[ = \left[ \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \right] \times \exp \left[ -\frac{1}{2} m_i k_B T \sum_{i \in \zeta_i} m_i \left( c_p^2 \| \Psi \|^2 - \Phi^T C_L [\Delta r_i]^2 C_L^T \Phi - c_p \Psi^T [\Delta r_i]^T C_L^T \Phi + c_p \Phi^T C_L [\Delta r_i] \Psi \right) \right] \]

\[ = \left[ \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \right] \exp \left\{ -\frac{1}{2} m_i k_B T \left[ \sum_{i \in \zeta_i} m_i \left( \sum_{i \in \zeta_i} m_i \| \Psi \|^2 - \sum_{i \in \zeta_i} [\Delta r_i]^2 C_L^T \Phi \right) \right] \right\} \]

\[ = \left[ \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \right] \exp \left[ -\frac{1}{2} m_i k_B T \left( \sum_{i \in \zeta_i} m_i \left( \sum_{i \in \zeta_i} \frac{M_i^2}{m_i} \| \Psi \|^2 + \Phi^T I_{\text{FG},i} I_{\text{FG},i}^{-1} I_i \Phi \right) \right] \right\} \]  

(S27)

where in the second step we have used \([\Delta r_i]^T = -[\Delta r_i]\), in the third step we have used \(\sum_{i \in \zeta_i} m_i [\Delta r_i] = 0\) (from definition of the position of the CG particle in terms of the center-of-mass of the constituent FG particles) to eliminate the last two terms in the exponent, and in the last step we have used the definitions of \(I_{\text{FG},i}\), \(c_p\), and \(C_L\) in eqs (S17), (S25), and (S26), respectively. Inserting eq (S27) into eq (S24) and rearranging gives

\[ \int \left[ \prod_{i \in \zeta_i} dp_i \exp \left( -\frac{p_i^2}{2m_i k_B T} \right) \right] \delta \left( M_{P,F} (p^n) - P_f \right) \delta \left( M_{L,F} (p^n, r^n) - L_f \right) \]

\[ = \frac{1}{(2\pi)^6} \left[ \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \right] \int d\Psi \exp \left[ -i \Psi^T P_f - \frac{1}{2} k_B T \sum_{i \in \zeta_i} m_i \| \Psi \|^2 \right] \]

\[ \int d\Phi \exp \left[ -i \Phi^T L_f - \frac{1}{2} k_B T \Phi^T I_{\text{FG},i}^{-1} I_{\text{FG},i} \Phi \right] \]

\[ = \frac{1}{(2\pi k_B T)^3} \left[ \prod_{i \in \zeta_i} (2\pi m_i k_B T)^{3/2} \right] \left( \sum_{i \in \zeta_i} m_i \right) \frac{3/2}{M_i^2} \det \left( I_{\text{FG},i}^{-1} I_f \right)^{-1/2} \]

\[ \times \exp \left[ -\frac{1}{2} k_B T \left( \frac{\sum_{i \in \zeta_i} m_i P_f^2}{m_i} + L_f^T I_{\text{FG},i}^{-1} I_{\text{FG},i}^{-1} L_f \right) \right] \]  

(S28)

where we have applied the Gaussian identity in eq (S21). Inserting eq (S28) into eq (S22) and using \(\det \left( I_{\text{FG},i}^{-1} I_f \right)^{-1/2} = \det(I_{\text{FG},i})^{1/2} / \det(I_f)\) gives

\[ (2\pi k_B T)^{(3n/2-3N)} \left\{ \prod_{I=1}^{N} \exp \left[ -\frac{1}{2k_B T} \left( \frac{P_f^2}{M_f^2} + L_f^T I_{\text{FG},i}^{-1} I_f \right) \right] \right\} \]

\[ = \int dr^n p_{\text{FG},i} (r^n) \delta \left( M_{P,F}^N (r^n) - R^N \right) \delta \left( M_{L,F}^N (r^n) - \Omega^N \right) \]

\[ \times \prod_{I=1}^{N} \left( \frac{1}{2\pi k_B T} \right)^{3} \left[ \prod_{i \in \zeta_i} m_i^{-3/2} (2\pi m_i k_B T)^{3/2} \right] \left( \frac{\sum_{i \in \zeta_i} m_i}{M_f^2} \right)^{3/2} \frac{\det(I_{\text{FG},i})^{1/2}}{\det(I_f)} \]

\[ \times \exp \left[ -\frac{1}{2k_B T} \left( \frac{\sum_{i \in \zeta_i} m_i P_f^2}{m_i} + L_f^T I_{\text{FG},i}^{-1} I_{\text{FG},i}^{-1} L_f \right) \right] \].  

(S29)
Noting that
\[ \prod_{i=1}^{N} \frac{1}{(2\pi k_B T)^3} \left[ \prod_{i \in \zeta} m_i^{-3/2} (2\pi m_i k_B T)^{3/2} \right] = (2\pi k_B T)^{(3n/2-3N)}, \] (S30)
canceling common terms on either side of eq (S29), and pulling the factor involving the CG linear momentum \( P_I \) out of the integral, gives
\[ \left[ \prod_{i=1}^{N} \exp \left( -\frac{P_i^2}{2k_B T M_i} \right) \right] \left[ \prod_{i=1}^{N} \exp \left( -\frac{L_i^T I^{-1} L_i}{2k_B T} \right) \right] \mathbb{P}_{CG,c}(R^N, \Omega^N) \\
= \left[ \prod_{i=1}^{N} \left( \frac{\sum_{i \in \zeta} m_i}{M_i} \right)^{3/2} \exp \left( -\frac{\sum_{i \in \zeta} m_i P_i^2}{2M_i^2 k_B T} \right) \right] \int \mathrm{d}r^N \mathbb{P}_{FG,c}(r^n) \delta \left( \mathbf{M}_R^N (r^n) - R^N \right) \delta \left( \mathbf{M}_C^N (r^n) - \Omega^N \right) \\
\times \left[ \prod_{i=1}^{N} \left( \frac{\det(\mathbf{I}_{FG,I})}{\det(\mathbf{I}_I)} \right)^{1/2} \exp \left( -\frac{L_i^T I^{-1} \mathbf{I}_{FG,I} I^{-1} L_i}{2k_B T} \right) \right]. \] (S31)

For this equation to hold for all values of the CG linear momentum for all CG particles, each factor involving the linear momentum \( P_I \) of a CG particle \( I \) must be equal on the two sides of the equation, i.e.
\[ \exp \left( -\frac{P_I^2}{2k_B T M_I} \right) = \left( \frac{\sum_{i \in \zeta} m_i}{M_I} \right)^{3/2} \exp \left( -\frac{\sum_{i \in \zeta} m_i P_i^2}{2M_i^2 k_B T} \right), \text{ for } I = 1, 2, \ldots, N, \] (S32)
which requires the masses of the CG particles to satisfy
\[ M_I = \sum_{i \in \zeta} m_i \text{ for } I = 1, 2, \ldots, N, \] (S33)
i.e. the mass of each CG particle must equal the total mass of the FG particles that are mapped to it. This consistency condition for the CG linear momenta had previously been derived, in a more general form for more general linear mappings of FG to CG positions, in the development of the MS-CG method for isotropic CG particles.\(^1\)

Inserting eq (S32) into eq (S31) and rearranging gives
\[ \prod_{I=1}^{N} \det(\mathbf{I}_I)^{1/2} \exp \left( -\frac{L_i^T I^{-1} L_i}{2k_B T} \right) = \int \mathrm{d}r^N \mathbb{P}_{FG,c}(r^n) \delta \left( \mathbf{M}_R^N (r^n) - R^N \right) \delta \left( \mathbf{M}_C^N (r^n) - \Omega^N \right) \]
\[ \times \left[ \prod_{I=1}^{N} \det(\mathbf{I}_{FG,I})^{1/2} \exp \left( -\frac{L_i^T I^{-1} \mathbf{I}_{FG,I} I^{-1} L_i}{2k_B T} \right) \right] \mathbb{P}_{CG,c}(R^N, \Omega^N) \]
\[ = \left[ \prod_{I=1}^{N} \det(\mathbf{I}_{FG,I})^{1/2} \exp \left( -\frac{L_i^T I^{-1} \mathbf{I}_{FG,I} I^{-1} L_i}{2k_B T} \right) \right]_{R^N, \Omega^N} \] (S34)
where we have used the consistency condition between the FG and CG configurational probability distributions, \( \mathbb{P}_{FG,c}(r^n) \) and \( \mathbb{P}_{CG,c}(R^N, \Omega^N) \), respectively, given in eq (S1) to write the integral as an equilibrium average \( \langle \cdots \rangle_{R^N, \Omega^N} \) over FG configurations that are mapped to CG configuration \( (R^N, \Omega^N) \).

Assuming that the inertia tensor \( \mathbf{I}_{FG,I} \) of the group of FG particles mapped onto a particular CG particle does not depend on the configuration of the other particles in the FG system, the average on the right-hand side of eq (S34) can be factored into independent averages for each CG particle, each of which must equal the corresponding factor on the left-hand side, which gives
\[ \det(\mathbf{I}_I)^{1/2} \exp \left( -\frac{L_i^T I^{-1} L_i}{2k_B T} \right) = \left\langle \det(\mathbf{I}_{FG,I})^{1/2} \exp \left( -\frac{L_i^T I^{-1} \mathbf{I}_{FG,I} I^{-1} L_i}{2k_B T} \right) \right\rangle_{R^N, \Omega^N}, \] (S35)
for \( I = 1, 2, \ldots, N \), where \( \langle \cdots \rangle_{R^N, \Omega^N} \) denotes an equilibrium average over FG configurations consistent with the coordinate mapping of CG particle \( I \).

Writing eq (S35) in terms of moments of inertia, \( I_{FG,I,q} \) and \( I_{I,q} \), and angular momenta, \( L_{I,q} \), about the principal axes, \( q = a, b, c \), in the body-fixed frame, assuming that the factors for each principal axis are independent of one another, gives
\[ I_{I,q}^{1/2} \exp \left( -\frac{L_{I,q}^2}{2I_{I,q} k_B T} \right) = \left\langle I_{FG,I,q}^{1/2} \exp \left( -\frac{I_{FG,I,q} L_{I,q}^2}{2I_{I,q}^2 k_B T} \right) \right\rangle_{R^N, \Omega_N}. \] (S36)
for $I = 1, 2, \ldots, N$ and $q = a, b, c$, where $I_{FG,I,q}$, $I_{I,q}$, and $L_{I,q}$ are the FG moment of inertia, CG moment of inertia, and angular momentum about the $q$ axis for CG particle $I$.

This consistency condition for the CG angular momenta can be simplified to eliminate the CG moment of inertia $I_{I,q}$ from the right-hand side by using eq (S15) to write it in terms of CG angular velocities as

$$I_{I,q}^{1/2} \exp \left( \frac{I_{I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) = \left( I_{I,q}^{1/2} \exp \left( \frac{I_{FG,I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) \right)_{R_I, \Omega_I} \tag{S37}$$

for $I = 1, 2, \ldots, N$ and $q = a, b, c$, where $\dot{\Omega}_{I,q}$ is the angular velocity about the $q$ axis.

Defining $\delta \dot{I} = I_{FG,I,q} - \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}$, writing for simplicity $\langle I_{FG,I,q} \rangle_{R_I, \Omega_I} \equiv \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}$, and writing the square root and exponential on the right-hand side as Taylor series, eq (S37) can be written as

$$I_{I,q}^{1/2} \exp \left( \frac{I_{I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) = \left( I_{I,q}^{1/2} \exp \left( \frac{I_{I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) \right)_{R_I, \Omega_I}$$

$$= \left( \langle I_{FG,I,q} \rangle_{R_I, \Omega_I} + \delta \dot{I} \right)^{1/2} \exp \left[ - \frac{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I} \dot{\Omega}_{I,q}^2}{2k_B T} \right]_{R_I, \Omega_I}$$

$$= \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}^{1/2} \exp \left( - \frac{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I} \dot{\Omega}_{I,q}^2}{2k_B T} \right)$$

$$\times \left[ 1 + \frac{1}{2} \frac{\delta \dot{I}}{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I}} - \frac{1}{8} \left( \frac{\delta \dot{I}}{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I}} \right)^2 + \cdots \right] \left[ 1 - \frac{\delta \dot{I}}{2k_B T} + \frac{\delta \dot{I}^2}{8} \left( \frac{\dot{\Omega}_{I,q}}{k_B T} \right)^2 - \cdots \right]_{R_I, \Omega_I}$$

$$= \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}^{1/2} \exp \left( - \frac{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I} \dot{\Omega}_{I,q}^2}{2k_B T} \right)$$

$$\times \left[ 1 + \frac{1}{4} \frac{\langle \delta \dot{I} \rangle^2}{\langle I_{FG,I,q} \rangle_{R_I, \Omega_I}} + \frac{1}{16} \langle \delta \dot{I} \rangle^3 \right] + \cdots \right] \langle \dot{\Omega}_{I,q}^2 \rangle_{k_B T} + \left[ \frac{1}{8} \langle \delta \dot{I} \rangle^2 + \frac{1}{16} \langle \delta \dot{I} \rangle^3 \right] - \cdots \right] \langle \dot{\Omega}_{I,q}^2 \rangle_{k_B T} + \cdots \right] \tag{S38}$$

where $\langle (\delta \dot{I})^m \rangle_R \equiv \langle (\delta \dot{I})^m \rangle_{R_I, \Omega_I}$, where $m$ is an integer, and we have used the fact that $\langle \delta \dot{I} \rangle_{R_I, \Omega_I} = 0$.

For the left- and right-hand sides of eq (S38) to be consistent, they must have the same functional dependence on the angular velocity $\dot{\Omega}_{I,q}$. This is the case if the terms on the right-hand side inside the curly braces that depend on $\dot{\Omega}_{I,q}$ are negligible for values of $\dot{\Omega}_{I,q}$ that occur with significant probability. An order-of-magnitude estimate of such values can be obtained by assuming a Gaussian distribution of $\dot{\Omega}_{I,q}$ for a rigid body with principal moment of inertia $I_{FG,I,q}$ about the $q$ axis, for which the largest values of $\dot{\Omega}_{I,q}$ that occur with significant probability are on the order of $k_B T/\langle I_{FG,I,q} \rangle$. Inserting this value into eq (S38), it can be seen that consistency between the left- and right-hand sides of eq (S38) is achieved if $\langle (\delta \dot{I})^m \rangle_R \ll \langle I_{FG,I,q} \rangle^m$ for integers $m \geq 2$. Using the definition of $\delta \dot{I}$, this condition is

$$\langle (I_{FG,I,q} - \langle I_{FG,I,q} \rangle_{R_I, \Omega_I})^m \rangle_{R_I, \Omega_I} \ll \langle I_{FG,I,q} \rangle^m_{R_I, \Omega_I} \tag{S39}$$

i.e. the $m$-th order cumulant of the distribution of $I_{FG,I,q}$ is much smaller than the $m$-th power of its mean value. For a Gaussian distribution of $I_{FG,I,q}$, for which all cumulants can be written in terms of the second moment (variance), this condition reduces to the standard deviation of $I_{FG,I,q}$ being smaller than its mean, i.e. $\langle (I_{FG,I,q} - \langle I_{FG,I,q} \rangle_{R_I, \Omega_I})^2 \rangle_{R_I, \Omega_I}^{1/2} \ll \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}$. With the condition in eq (S39) satisfied, eq (S38) reduces to

$$I_{I,q}^{1/2} \exp \left( \frac{I_{I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) \approx \langle I_{FG,I,q} \rangle_{R_I, \Omega_I}^{1/2} \exp \left( \frac{I_{FG,I,q} \dot{\Omega}_{I,q}^2}{2k_B T} \right) \tag{S40}$$

for $I = 1, 2, \ldots, N$ and $q = a, b, c$, from which it can be identified that

$$I_{I,q} \approx \langle I_{FG,I,q} \rangle_{R_I, \Omega_I} \tag{S41}$$

for $I = 1, 2, \ldots, N$ and $q = a, b, c$. This means that the angular-momentum distributions of the FG and CG systems will be consistent if the fluctuations of the principal moments of inertia of the group of FG particles that is mapped onto a CG particle are small compared with their mean values, in which case the principal moments of inertia of the CG particle are approximately equal to the corresponding averages of the FG principal moments of inertia.
S3  Force Basis Functions and Least-Squares Solver

S3.1  Scalar Variables of Basis Functions

In the applications of the anisotropic force-matching coarse-graining (AFM-CG) algorithm presented in this paper, the basis functions for the CG potential were assumed to be functions of a set of scalar variables that are in turn functions of the relative positions and orientations of particle pairs. For two anisotropic CG particles \(A\) and \(B\) with position vectors \(R_A\) and \(R_B\), respectively, and orientation defined by the matrices \(\mathbf{A} \equiv (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)\) and \(\mathbf{B} \equiv (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)\), respectively, whose columns are unit vectors aligned with the particle’s principal axes in the space-fixed frame, the set of scalar variables is defined by the matrices

\[
\mathbf{R} = \mathbf{B}^T \mathbf{A} = \begin{bmatrix}
\cos \alpha \cos \gamma & -\sin \alpha & \cos \beta \sin \gamma & -\cos \alpha \sin \gamma & \sin \alpha \cos \beta & \sin \alpha \sin \beta \\
\sin \alpha \cos \gamma & \cos \alpha & \cos \beta \cos \gamma & \sin \alpha \cos \gamma & -\cos \alpha \sin \beta & \cos \beta \\
\sin \beta \sin \gamma & 0 & \sin \beta \cos \gamma & 0 & 0 & 0
\end{bmatrix}.
\]

(S42)

S3.2  Multivariate Potential and Force Basis Functions

A CG pair potential between CG particles \(A\) and \(B\) can be written as a linear combination of \(N_b\) basis functions that are multivariate functions of the scalar variables in section S3.1, as described in eq (35) in the main paper. In this work, each basis function \(\mathcal{B}_i\) for \(i = 1, 2, \ldots, N_b\) was taken to have the form

\[
\mathcal{B}_i(R, \alpha, \beta, \gamma, \phi, \theta) \equiv \mu_n(R) \cos(k_\alpha \alpha) \cos(k_\beta \beta) \cos(k_\gamma \gamma) \cos(k_\phi \phi) \cos(k_\theta \theta),
\]

(S43)

where \(\mu_n(R)\) is a cubic spline basis function of the inter-particle distance variable, and \(k_\alpha, k_\beta, k_\gamma, k_\phi, k_\theta\) are parameters defining the cosine basis functions of the angular variables. A cubic spline was chosen for the radial function as it ensures the smoothness and continuity of the basis force vectors. For the univariate case, a cubic spline is a piece-wise continuous function of cubic polynomials with continuous derivatives up to order 2 that interpolates the data between sub-intervals of the variable space. The sub-intervals are connected through predetermined points called knots between which the cubic polynomials are constructed, and the boundary knots are the boundaries of the variable interval. In addition to the continuity conditions, a cubic spline can have additional constraints, namely that the function is linear beyond the boundary knots, for which the function is called a natural cubic spline. In this work, natural cubic splines were used to define \(N_R\) basis functions of the inter-particle distance \(R\) in terms of \(N_R\) predetermined knots \((R_0, R_1, \ldots, R_{N_R})\) as

\[
\mu_1(R) = 1, \quad \mu_2(R) = R, \quad \mu_n(R) = d_{n-2}(R) - d_{N_R-1}(R) \quad \text{for} \quad n = 3, 4, \ldots, N_R,
\]

(S44)

where

\[
d_m(R) = \frac{(R - R_m)^3}{R_{N_R} - R_m} - \frac{(R - R_m)^3}{R_{N_R} - R_m},
\]

(S45)

with \(f^+\) denoting the positive part of \(f\) (negative values are set to zero).

The force basis functions are derived as minus the gradient for the basis functions with respect to the CG position. A basis function to construct the force that particle \(B\) exerts on particle \(A\) is derived using

S9
the chain rule as

\[-\sum_{j=1}^{6} \frac{\partial B_i(\xi)}{\partial \xi_j} \frac{\partial \xi_j}{\partial R_A} = - \left( \frac{\partial B_i(\xi)}{\partial R} \frac{\partial R}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \alpha} \frac{\partial \alpha}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \beta} \frac{\partial \beta}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \gamma} \frac{\partial \gamma}{\partial R_A} \right) \]

\[+ \left( \frac{\partial B_i(\xi)}{\partial \phi} \frac{\partial \phi}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \theta} \frac{\partial \theta}{\partial R_A} \right) \]

\[= - \left( \frac{\partial B_i(\xi)}{\partial R} \frac{\partial R}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \cos \phi} \frac{\partial \cos \phi}{\partial R_A} + \frac{\partial B_i(\xi)}{\partial \cos \theta} \frac{\partial \cos \theta}{\partial R_A} \right). \quad (S46)\]

Here the terms involving partial derivatives of the Euler angles with respect to \( R_A \) vanish because these angles do not depend on the position coordinates of the particles, and derivatives have been expressed with respect to \( \cos \phi \) and \( \cos \theta \) instead of \( \phi \) and \( \theta \), respectively, to facilitate the calculations. The partial derivatives of the basis function \( B_i(\xi) \) with respect to each variables in eq (S46) can be easily derived, using

\[
\frac{d\mu_n(R)}{dR} = \begin{cases} 
0 & \text{if } n = 1, \\
1 & \text{if } n = 2, \\
\frac{d}{dR} d_{n-2}(R) - \frac{d}{dR} d_{N_2}(R) & \text{for } n = 3, 4, ..., N_R,
\end{cases} \quad (S47)
\]

and

\[
\frac{\partial \cos k \phi}{\partial \cos \phi} = k \phi \sin k \phi, \quad (S48)
\]

\[
\frac{\partial \cos k \theta}{\partial \cos \theta} = k \theta \sin k \theta, \quad (S49)
\]

with \( \frac{d}{dR} d_m(R) = 3 \left[ \frac{(R-R_m)^2}{R_{N2} - R_m} \right] \) in eq (S47). The remaining partial derivatives with respect to \( R_A \) in eq (S46) are

\[
\frac{\partial R}{\partial R_A} = \frac{\partial R}{\partial R_{AB}} = \hat{R}_{AB}, \quad (S50)
\]

\[
\frac{\partial \cos \phi}{\partial R_A} = \frac{\partial \left( \hat{a}_3 \cdot R_{AB}/R \right)}{\partial R_{AB}} = \hat{a}_3 - \left( \hat{a}_3 \cdot \hat{R}_{AB} \right) \hat{R}_{AB}, \quad (S51)
\]

and

\[
\frac{\partial \cos \theta}{\partial R_A} = \frac{\partial \left( \hat{b}_3 \cdot R_{AB}/R \right)}{\partial R_{AB}} = \hat{b}_3 - \left( \hat{b}_3 \cdot \hat{R}_{AB} \right) \hat{R}_{AB}. \quad (S52)
\]
S3.3 Basis Functions for Benzene and Perylene Parametrizations

Because benzene and perylene are coarse-grained into uniaxial particles in this work, only terms with the variables \( \{R, \beta, \phi, \theta\} \) appear in the expression of the basis functions (eq (S43)) (\( \beta \) is the angle between the unique molecular axes of the two particles if these axes are defined as \( \hat{a}_3 \) and \( \hat{b}_3 \), respectively). Details of the basis functions used in the parametrization of the benzene and perylene models are as follows:

- **Benzene**
  - \( N_R = 23 \) cubic splines were used for the basis functions the inter-particle distance. The knots for these cubic splines were between 3–10 Å and chosen so that the knot density was higher at small distances and lower at large distances. Specifically, the knot \( R_n \) was taken to be 
    \[ R_n = R_{N_R} - (R_{N_R} - R_0) \cos \left( \frac{n\pi}{2(N_R - 1)} \right) \text{ for } n = 0, \ldots, N_R. \]
  - The parameters \( k_\beta, k_{\phi} \) and \( k_\theta \) in the cosine terms were elements of the set \{0, 2, 4, 6, 8\}. Only even integers were used because the CG potential has a period of \( \pi \) with respect to \( \beta, \phi, \) and \( \theta \) due to the symmetry of the benzene molecule.
  - In total, \( 23 \times 5 \times 5 \times 5 = 2875 \) basis functions were used for the CG potential.

- **Perylene**
  - \( N_R = 39 \) cubic splines were used for the basis functions of the inter-particle distance. The knots for these cubic splines are between 2.5–19 Å and chosen so that the knot density is higher at small distances and lower at large distances. The same function to determine the knots was used as for benzene: 
    \[ R_n = R_{N_R} - (R_{N_R} - R_0) \cos \left( \frac{n\pi}{2(N_R - 1)} \right) \text{ for } n = 0, \ldots, N_R. \]
  - The parameters \( k_\beta, k_{\phi}, \) and \( k_\theta \) in the cosine terms were elements of the set \{0, 2, 4, 6, 8, 10\}. Only even integers were used because the CG potential has a period of \( \pi \) with respect to \( \beta, \phi, \) and \( \theta \) due to the symmetry of the perylene molecule.
  - In total, \( 39 \times 6 \times 6 \times 6 = 8424 \) basis functions were used for the CG potential.

S3.4 Least-Squares Solver

After evaluating the force matrix as described in the main paper, the basis coefficients that optimize the CG force field can be found using a least-squares method. Various strategies for solving the force matrix have been addressed and implemented in the MS-CG method for isotropic particles.\(^6\) We chose to apply a strategy that involves the sequential householder transformation\(^7\) of the force matrix to an equivalent matrix of smaller size, and solving the resulting matrix equation using singular value decomposition.\(^8\) The key idea and advantages of this strategy have been discussed in detail previously by Lu et al.\(^6\). The method is especially useful in the applications of the AFM-CG method in this work, in which the number of atomistic configurations (which correspond to the number of matrix rows) is large compared to the number of basis functions (number of matrix columns).
S4 Modified S-function Fit to Anisotropic CG Potential

The general expression for the modified S-function pair potential \( U_{S}^{\text{mod}} \) for uniaxial particles used in the CG MD simulations of benzene and perylene is described in the main paper. Here the expressions for the various terms in the potential and the expressions for the resulting forces and torques are provided. Including the pressure correction \( \Delta U \), the total pair potential is

\[
U_{S}^{\text{mod}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = U_{S}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) + U_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) + U_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) + \Delta U(\mathbf{R}_{i,j}),
\]

where \( \mathbf{R}_{i,j} = \mathbf{R}_{i} - \mathbf{R}_{j} \) is the vector that connects the particle centers, \( \mathbf{u}_{j} \) and \( \mathbf{u}_{j} \) are the unit vectors along the unique axis of each particle, and \( \mathbf{R}_{i,j} = ||\mathbf{R}_{i,j}|| \). The form of the S-function potential used in this work is

\[
U_{S}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = 4\varepsilon \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) \left[ A \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) \left( \frac{\sigma_{0}}{d(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})} \right)^{p(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})} - \left( \frac{\sigma_{0}}{d(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})} \right)^{6} \right],
\]

(S54)

which differs from the more conventional Lennard-Jones-like form used in previous studies\(^{9,10}\) in including an orientation-dependent prefactor \( A(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) \) and an orientation-dependent exponent \( p(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) \), where \( \mathbf{R}_{i,j} = \mathbf{R}_{i}/\mathbf{R}_{j} \), in the repulsive term instead of constant values of 1 and 12, respectively. These functions are defined as

\[
p(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = c_{p1} \exp \left[ -c_{p2}\sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) \right] + c_{p3}
\]

(S55)

and

\[
A(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = c_{A1} \left[ \sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) - c_{A2} \right]^{3} + c_{A3}\sigma + c_{A5}.
\]

(S56)

The exponential term has the form

\[
U_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = \epsilon_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) \exp \left[ -\kappa_{1} \frac{d_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})}{\sigma_{0}} \right].
\]

(S57)

The oscillatory term has the form

\[
U_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = \epsilon_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) \left[ \cos \left( \frac{\kappa_{2} d_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})}{\sigma(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})} \right) + C_{2} \right] \exp \left[ -\kappa_{2} \frac{d_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})}{\sigma(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j})} \right].
\]

(S58)

The pressure-correction term for the potential cut-off distance \( R_{c} \) has the form

\[
\Delta U(\mathbf{R}_{i,j}) = b \left( 1 - \frac{\mathbf{R}_{i,j}}{R_{c}} \right).
\]

(S59)

The various anisotropic functions in eqs (S54)–(S58) are defined as

\[
\sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) = \sigma_{0} \left[ \sigma_{000} S_{000} + \sigma_{cc2} (S_{202} + S_{220}) + \sigma_{c22} S_{220} + \sigma_{c22} S_{222} + \sigma_{c24} S_{224} + \sigma_{c24} (S_{404} + S_{044}) \right],
\]

(S60)

\[
\epsilon \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) = \epsilon_{0} \left[ \epsilon_{00} S_{000} + \epsilon_{cc2} (S_{202} + S_{220}) + \epsilon_{c22} S_{220} + \epsilon_{c22} S_{222} + \epsilon_{c24} S_{224} + \epsilon_{c24} (S_{404} + S_{044}) \right],
\]

(S61)

\[
\epsilon_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = \epsilon_{\exp0} \left[ \epsilon_{exp}^{000} S_{000} + \epsilon_{exp}^{cc2} (S_{202} + S_{220}) + \epsilon_{exp}^{c22} S_{220} + \epsilon_{exp}^{c22} S_{222} + \epsilon_{exp}^{c24} S_{224} + \epsilon_{exp}^{c24} (S_{404} + S_{044}) \right],
\]

(S62)

\[
\epsilon_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = \epsilon_{\text{osc0}} \left( f_{1} f_{2} \right)^{6},
\]

(S63)

\[
d(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = R_{ij} - \sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) + \sigma_{0},
\]

(S64)

\[
d_{\exp}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = R_{ij} - \sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) + \sigma_{\exp},
\]

(S65)

\[
d_{\text{osc}}(\mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j}) = R_{ij} - \sigma \left( \mathbf{R}_{i,j}, \mathbf{u}_{j}, \mathbf{u}_{j} \right) + \sigma_{\text{osc}},
\]

(S66)

with

\[
f_{0} = \mathbf{u}_{j} \cdot \mathbf{u}_{j}, \quad f_{1} = \mathbf{u}_{j} \cdot \mathbf{R}_{ij}, \quad f_{2} = \mathbf{u}_{j} \cdot \mathbf{R}_{ij}.
\]

(S67)
Implementation of the modified S-function potential for MD simulations require calculations of the forces presented below in Tables S2 and S4, respectively. Note that not all of the terms in $U_S^{\text{mod}}$ are used in both models, with unused parameters set to zero.

### S4.1 Modified S-function Force and Torque Calculations

Implementation of the modified S-function potential for MD simulations require calculations of the forces and torques on the CG particles. The expressions for the forces and torques for a pair of uniaxial particles are

$$F_{IJ} = -\frac{\partial U_S^{\text{mod}}}{\partial R_{IJ}} = \left( \frac{\partial U_S^{\text{mod}}}{\partial f_1} \right) \left( \frac{\hat{R}_{IJ}}{R_{IJ}} \right) \left( \frac{\hat{u}_J - f_1 \hat{R}_{IJ}}{R_{IJ}} \right) \left( \frac{\hat{u}_J - f_2 \hat{R}_{IJ}}{R_{IJ}} \right)$$

and

$$\tau_{IJ} = -\hat{u}_I \times \left( \frac{\partial U_S^{\text{mod}}}{\partial f_1} \hat{R}_{IJ} \right) \left( \frac{\partial U_S^{\text{mod}}}{\partial f_2} \hat{R}_{IJ} \right).$$

The derivative with respect to the inter-particle separation $R_{IJ}$ is

$$\frac{\partial U_S^{\text{mod}}}{\partial R_{IJ}} = \frac{\partial U_S}{\partial R_{IJ}} + \frac{\partial U_{\exp}}{\partial R_{IJ}} + \frac{\partial U_{\osc}}{\partial R_{IJ}} - \frac{b}{R_c},$$

where

$$\frac{\partial U_S}{\partial R_{IJ}} = -4\epsilon \left[ A_p \left( \frac{\sigma_0}{a} \right)^p - 6 \left( \frac{\sigma_0}{a} \right)^6 \right],$$

$$\frac{\partial U_{\exp}}{\partial R_{IJ}} = -\frac{\kappa exp}{\sigma_0} \exp \left( -\frac{d exp}{\sigma_0} \right),$$

and

$$\frac{\partial U_{\osc}}{\partial R_{IJ}} = \frac{\epsilon_{osc}}{\sigma} \exp \left( -\frac{d_{osc}}{\sigma} \right) \left\{ -\frac{1}{2} \sin \left( \frac{d_{osc}}{\sigma} \right) - \frac{1}{2} \left[ \cos \left( \frac{d_{osc}}{\sigma} \right) + C_2 \right] \right\}. $$

The derivatives with respect to the orientational variables $f_i$, where $i = 0, 1, 2$, are

$$\frac{\partial U}{\partial f_i} = \frac{\partial U_S}{\partial f_i} + \frac{\partial U_{\exp}}{\partial f_i} + \frac{\partial U_{\osc}}{\partial f_i} + \frac{\partial U_{\exp}}{\partial \sigma} \frac{\partial \sigma}{\partial f_i} + \frac{\partial U_{\osc}}{\partial \sigma} \frac{\partial \sigma}{\partial f_i}.$$

where

$$\frac{\partial U_S}{\partial \sigma} = 4 \left[ A_p \left( \frac{\sigma_0}{a} \right)^p - 6 \left( \frac{\sigma_0}{a} \right)^6 \right],$$

$$\frac{\partial U_{\exp}}{\partial \sigma} = 4 \epsilon \left[ \frac{3 c_{A1}^2 - c_{A2}^2}{A_p c_{A1} c_{A2}} \ln \left( \frac{\sigma_0}{a} \right) + A_p \frac{\sigma_0}{a} \right] - 6 \frac{\sigma_0^6}{a^6},$$

$$\frac{\partial U_{\osc}}{\partial \sigma} = 4 \epsilon c_{A3} \left( \frac{\sigma_0}{a} \right)^p,$$

and

$$\frac{\partial \sigma}{\partial f_i} = \epsilon_0 \left[ c_{000} \frac{\partial S_{000}}{\partial f_i} + c_{cc2} \left( \frac{\partial S_{022}}{\partial f_i} + \frac{\partial S_{202}}{\partial f_i} \right) + \frac{\partial S_{222}}{\partial f_i} \right],$$

where $c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, s_0, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4}, c_0, c_{000}, c_{cc2}, c_{cc4,
\[
\frac{\partial \sigma}{\partial f_i} = \sigma_0 \left[ \sigma_{000} \frac{\partial S_{000}}{\partial f_i} + \sigma_{0c2} \left( \frac{\partial S_{020}}{\partial f_i} + \frac{\partial S_{022}}{\partial f_i} \right) + \sigma_{220} \frac{\partial S_{220}}{\partial f_i} \\
+ \sigma_{222} \frac{\partial S_{222}}{\partial f_i} + \sigma_{2c2} \frac{\partial S_{024}}{\partial f_i} + \sigma_{c4} \left( \frac{\partial S_{044}}{\partial f_i} \right) \right], \\
\frac{\partial U_{\exp}}{\partial \epsilon} = \exp \left( -\kappa_1 \frac{d_{\exp}}{\sigma_0} \right), \\
\frac{\partial U_{\exp}}{\partial \sigma} = \kappa_1 \frac{\epsilon_{\exp}}{\sigma_0} \exp \left( -\kappa_1 \frac{d_{\exp}}{\sigma_0} \right), \\
\frac{\partial \epsilon_{\exp}}{\partial f_i} = \epsilon_{\exp} \left[ \sigma_{000} \frac{\partial S_{000}}{\partial f_i} + \sigma_{0c2} \left( \frac{\partial S_{020}}{\partial f_i} + \frac{\partial S_{022}}{\partial f_i} \right) + \epsilon_{220} \frac{\partial S_{220}}{\partial f_i} + \epsilon_{2c2} \frac{\partial S_{024}}{\partial f_i} + \epsilon_{c4} \left( \frac{\partial S_{044}}{\partial f_i} \right) \right], \\
\frac{\partial U_{\osc}}{\partial \epsilon_{\osc}} = \left[ \cos \left( k_2 \frac{d_{\osc}}{\sigma} \right) + C_2 \right] \exp \left( -\kappa_2 \frac{d_{\osc}}{\sigma} \right), \\
\frac{\partial U_{\osc}}{\partial \sigma} = \left( \frac{\epsilon_{\osc}}{\sigma} \right) \left[ k_2 \sin \left( k_2 \frac{d_{\osc}}{\sigma} \right) + k_2 \left( \cos \left( k_2 \frac{d_{\osc}}{\sigma} \right) + C_2 \right) \right] \exp \left( -\kappa_2 \frac{d_{\osc}}{\sigma} \right), \\
\frac{\partial \epsilon_{\osc}}{\partial f_0} = 0, \quad \frac{\partial \epsilon_{\osc}}{\partial f_1} = 6 \epsilon_{\osc} f_1^3 f_2^6, \quad \frac{\partial \epsilon_{\osc}}{\partial f_2} = 6 \epsilon_{\osc} f_1^3 f_2^6, \\
\frac{\partial S_{000}}{\partial f_0} = \frac{\partial S_{000}}{\partial f_1} = \frac{\partial S_{000}}{\partial f_2} = 0, \\
\frac{\partial S_{022}}{\partial f_0} = 0, \quad \frac{\partial S_{022}}{\partial f_1} = 0, \quad \frac{\partial S_{022}}{\partial f_2} = 0, \\
\frac{\partial S_{222}}{\partial f_0} = 0, \quad \frac{\partial S_{222}}{\partial f_1} = 3 f_2 / \sqrt{5}, \quad \frac{\partial S_{222}}{\partial f_2} = 0, \\
\frac{\partial S_{222}}{\partial f_0} = 3 f_2 / \sqrt{5}, \quad \frac{\partial S_{222}}{\partial f_1} = 0, \quad \frac{\partial S_{222}}{\partial f_2} = 0, \\
\frac{\partial S_{222}}{\partial f_0} = 0, \quad \frac{\partial S_{222}}{\partial f_1} = \frac{9 f_1 f_2 - 6 f_0}{\sqrt{70}}, \quad \frac{\partial S_{222}}{\partial f_2} = \frac{(9 f_0 f_2 - 6 f_1)}{\sqrt{70}}, \quad \frac{\partial S_{222}}{\partial f_0} = \frac{(9 f_0 f_1 - 6 f_2)}{\sqrt{70}}, \\
\frac{\partial S_{224}}{\partial f_0} = (4 f_0 - 20 f_1 f_2) / (4 \sqrt{70}), \quad \frac{\partial S_{224}}{\partial f_1} = (-10 f_1 - 20 f_0 f_2 + 70 f_1 f_2^2) / (4 \sqrt{70}), \quad \frac{\partial S_{224}}{\partial f_2} = (-10 f_2 - 20 f_0 f_1 + 70 f_2 f_1^2) / (4 \sqrt{70}), \\
\frac{\partial S_{044}}{\partial f_0} = 0, \quad \frac{\partial S_{044}}{\partial f_1} = \frac{-60 f_1 + 140 f_1^3}{24}, \quad \frac{\partial S_{044}}{\partial f_2} = \frac{-60 f_2 + 140 f_2^3}{24}, \quad \frac{\partial S_{244}}{\partial f_0} = 0, \quad \frac{\partial S_{244}}{\partial f_1} = \frac{-60 f_1 + 140 f_2^3}{24}, \quad \frac{\partial S_{244}}{\partial f_2} = \frac{-60 f_2 + 140 f_2^3}{24}, \\
\frac{\partial S_A}{\partial f_0} = \left( f_0 - f_1 f_2^2 \right) (1 - f_0 f_1 f_2) \left[ -c_{4f} f_1 f_2 (f_0^2 - f_1^2 f_2^2) + 4 f_0 (1 - f_0 f_1 f_2) \right], \\
\frac{\partial S_A}{\partial f_1} = -f_2 (f_0 - f_1 f_2) (1 - f_0 f_1 f_2) \left[ c_{4f} f_0 (f_0^2 - f_1^2 f_2^2) + f_1 f_2 (1 - f_0 f_1 f_2) \right], \\
\frac{\partial S_A}{\partial f_2} = -f_1 (f_0 - f_1 f_2) (1 - f_0 f_1 f_2) \left[ c_{4f} f_0 (f_0^2 - f_1^2 f_2^2) + f_1 f_2 (1 - f_0 f_1 f_2) \right].
S5 Benzene Simulations

S5.1 Distributions of Principal Moments of Inertia

The distributions of the principal moments of inertia from the fine-grained (FG) atomistic MD simulation of benzene at 300 K and 1 atm are shown in Figure S1 and the mean values are given in Table S1. The fluctuations of the principal moments are small compared to the mean values, with the standard deviation less than 2% of the mean value of each principal moment.

![Fig. S1: Distributions of the principal moments of inertia of FG benzene at 300 K and 1 atm.](image)

**Table S1**: Mean values of principal moments of inertia of FG benzene at 300 K and 1 atm.

| Property | Value  | Unit         |
|----------|--------|--------------|
| $I_a$    | 88.04  | g mol$^{-1}$ Å$^2$ |
| $I_b$    | 92.36  | g mol$^{-1}$ Å$^2$ |
| $I_c$    | 180.11 | g mol$^{-1}$ Å$^2$ |
S5.2 AFM-CG Model Parameters

The parameters used in the AFM-CG MD simulations of benzene are presented in Table S2.

| Parameter | Value  | Unit   |
|-----------|--------|--------|
| $\sigma_0$ | 3.4976 | Å      |
| $\sigma_{000}$ | 1.5828 | –      |
| $\sigma_{cc2}$ | -0.5860 | –   |
| $\sigma_{220}$ | -0.1522 | –   |
| $\sigma_{222}$ | -0.0617 | –    |
| $\sigma_{224}$ | 0.6244 | –   |
| $\sigma_{cc4}$ | 0      | –    |
| $\epsilon_0$ | 0.5911 | kcal mol$^{-1}$ |
| $\epsilon_{000}$ | 1.7456 | –     |
| $\epsilon_{cc2}$ | 0.7297 | –    |
| $\epsilon_{220}$ | 0.7904 | –   |
| $\epsilon_{222}$ | -1.9927 | –   |
| $\epsilon_{224}$ | -6.5476 | –   |
| $\epsilon_{cc4}$ | 0      | –    |
| $\epsilon_{p1}$ | -4.4531 | –   |
| $\epsilon_{p2}$ | 0.0564 | Å$^{-1}$ |
| $\epsilon_{p3}$ | 12.2418 | –   |
| $\epsilon_{A1}$ | -0.0052 | Å$^{-3}$ |
| $\epsilon_{A2}$ | 4.3459 | Å    |
| $\epsilon_{A3}$ | 0.2087 | –    |
| $\epsilon_{A4}$ | 10.3869 | –   |
| $\epsilon_{A5}$ | 0.7536 | –    |

(a) S-function parameters

| Parameter | Value | Unit |
|-----------|-------|------|
| $\epsilon_{exp0}$ | 0 | kcal mol$^{-1}$ |
| $\epsilon_{000}^{exp}$ | 0 | –   |
| $\epsilon_{cc2}^{exp}$ | 0 | –   |
| $\epsilon_{220}^{exp}$ | 0 | –   |
| $\epsilon_{222}^{exp}$ | 0 | –   |
| $\epsilon_{224}^{exp}$ | 0 | –   |
| $\sigma_{exp}$ | 0 | Å   |
| $\kappa_1$ | 0 | –   |
| $\epsilon_{osc0}$ | -0.0542 | kcal mol$^{-1}$ |
| $k_2$ | 5.6156 | – |
| $C_2$ | -0.6455 | – |
| $\sigma_{osc}$ | -4.4082 | Å |
| $\kappa_2$ | 2.9079 | – |
| $b$ | -0.5290 | kcal mol$^{-1}$ |

(b) Additional + pressure correction parameters

| Parameter | Value | Unit |
|-----------|-------|------|
| $I_a$ | 90.20 | g mol$^{-1}$ Å$^2$ |
| $I_b$ | 90.20 | g mol$^{-1}$ Å$^2$ |
| $I_c$ | 180.11 | g mol$^{-1}$ Å$^2$ |

(c) Principal moments of inertia
To obtain the optimized parameters for the modified S-function potential (eq (S53)) in Table S2 used in the AFM-CG MD simulations, a least-squares fit of the S-function term (eq (S54)) in this potential to the basis-expansion potential was first carried out for 300,000 dimer configurations with a separation distance up to 9 Å randomly sampled from the equilibrated FG simulation used for parametrization. This gave a reasonable fit, but the depth of the potential well for the edge-to-edge configuration was overestimated, which led to an overestimate of the probability of this configuration in CG simulations. To overcome this issue, the parameters in the orientation-dependent prefactor $A(R_{ij}, \hat{u}_i, \hat{u}_j)$ and orientation-dependent exponent $p(R_{IJ}, \hat{u}_I, \hat{u}_J)$ were optimized again by a least-squares fit with all other parameters in eq (S54) fixed using an augmented data set that included an additional 1000 edge-to-edge configurations with separation distances uniformly distributed between 6–9 Å. Then, to fit the significant positive hump in the potential for configurations close to face-to-face, the oscillatory term (eq (S58)) in eq (S53) was added to the fit potential, the set of dimer configurations was augmented with 2000 additional face-to-face and parallel-displaced (PD) configurations (PD-10, PD-30, PD-50 and PD-70 as denoted in Figure S2) with separation distances uniformly distributed up to 9 Å, and a least-squares fit was used to optimize the parameters in the oscillatory term with the previously optimized S-function parameters kept fixed. The lower limit of the separation distance of the dimer configurations added in this step was chosen such that the potential was less than 5 kcal mol$^{-1}$. Basis expansion and fitted modified S-function potential curves for various benzene–benzene dimer relative orientations are shown in Figure S2.

**Fig. S2:** Benzene AFM-CG pair potential (without pressure correction) for various dimer relative orientations using the basis set expansion (solid lines) and fitted modified S-function potential (dashed lines). The scalar variables $\beta, \phi, \theta$ (defined in section S3) that describe the relative orientation for theses configurations are: face-to-face ($0^\circ,0^\circ,0^\circ$), edge-to-face ($90^\circ,0^\circ,90^\circ$), edge-to-edge ($0^\circ,90^\circ,90^\circ$), PD-10 (parallel-displaced) ($0^\circ, x^\circ,x^\circ$), Cross-x ($x^\circ,90^\circ,90^\circ$), SP-x ($\pi$-stacking) ($x^\circ,0^\circ,x^\circ$).
The pressure-corrected potential does not differ substantially the modified S-function potential fit to the uncorrected AFM-CG potential, as shown in Figure S3.

**Fig. S3:** Benzene AFM-CG pair potential for various dimer relative orientations using the basis set expansion (solid lines) and fitted to the modified S-function form with (dotted lines) and without (dashed lines) the pressure correction.
S5.3 NVT Simulations

Properties of the FG benzene model from NPT simulations at 1 atm are compared in this section with those from NVT simulations of CG benzene models in which the density was set to the average density of the FG simulation at the corresponding temperature.

S5.3.1 Radial Distribution Function (RDF)

RDFs of the FG, AFM-CG, and Bowen CG\textsuperscript{9} models are compared at two different temperatures in Figure S4.

![RDF Comparison](image)

**Fig. S4:** RDFs from NPT simulations of FG benzene model at 1 atm and (left) 300 K and (right) 330 K, and from NVT simulations of the AFM-CG and Bowen CG\textsuperscript{9} benzene models with density set to the average density of the FG simulation at the corresponding temperature.
S5.3.2 Angular–Radial Distribution Function (ARDF)

ARDFs of the FG, AFM-CG, and Bowen CG benzene models at 300 K and 330 K are compared in Figure S5 and Figure S6, respectively.

**Fig. S5:** ARDFs from (top) NPT simulation of FG benzene model at 300 K and 1 atm, and from NVT simulations of the (middle) AFM-CG and (bottom) Bowen CG benzene models with density set to the average density of the FG simulation at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. An angle of 0° between out-of-plane axes indicates parallel alignment (e.g. face-to-face, parallel-displaced or edge-to-edge configuration) and angle of 90° indicates perpendicular alignment (e.g. face-to-edge or cross configuration).
Fig. S6: ARDFs from (top) NPT simulation of FG benzene model at 330 K and 1 atm, and from NVT simulations of the (middle) AFM-CG and (bottom) Bowen CG benzene models with density set to the average density of the FG simulation at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. An angle of 0° between out-of-plane axes indicates parallel alignment (e.g. face-to-face, parallel-displaced or edge-to-edge configuration) and angle of 90° indicates perpendicular alignment (e.g. face-to-edge or cross configuration).

S5.4 NPT Simulations

S5.4.1 Radial Distribution Function (RDF)

RDFs from NPT simulations at 1 atm of the FG, AFM-CG, and Bowen CG benzene models are compared at two different simulation temperatures in Figure S7.

Fig. S7: RDFs from NPT simulations of FG, AFM-CG, and Bowen CG benzene models at 1 atm and (left) 300 K and (right) 330 K.
S5.4.2 Angular–Radial Distribution Function (ARDF)

ARDFs from NPT simulations of the FG, AFM-CG, and Bowen CG benzene models at 1 atm and 330 K are compared in Figure S8, while those at 300 K and the same pressure are compared in the main paper.

**Fig. S8:** ARDFs from NPT simulations of (top) FG, (middle) AFM-CG, and (bottom) Bowen CG benzene models at 330 K and 1 atm, shown as (left) 2D color maps and (right) 1D slices at fixed angles on the right. An angle of $0^\circ$ between out-of-plane axes indicates parallel alignment (e.g. face-to-face, parallel-displaced or edge-to-edge configuration) and angle of $90^\circ$ indicates perpendicular alignment (e.g. face-to-edge or cross configuration).
S5.4.3 Mean Squared Displacement (MSD)

The mean squared displacement (MSD) was calculated as described in the main paper and used to calculate the translational diffusion coefficients of the FG and AFM-CG benzene models.

Fig. S9: MSD versus time from NPT simulations of FG and AFM-CG benzene models at 1 atm and various temperatures.
S5.4.4 Orientational Time Correlation Function (OTCF)

The orientational time correlation function (OTCF) was calculated as described in the main paper and used to calculate the rotational diffusion coefficients with respect to the out-of-plane and in-plane molecular axes of the FG and AFM-CG benzene models.

Fig. S10: OTCF of out-of-plane molecular axis versus time from NPT simulations of FG and AFM-CG benzene models at 1 atm and various temperatures.

Fig. S11: OTCF of in-plane molecular axis versus time of FG and AFM-CG benzene models from NPT simulations at 1 atm and various temperatures.
S6  Perylene Simulations

S6.1 Distributions of Principal Moments of Inertia

Distributions of the moments of inertia about three principal axes for FG atomistic perylene are shown in Figure S12. The fluctuations of the principal moments are small compared to the mean values, with the standard deviation less than 2% of the mean value of each principal moment.

![Figure S12: Distributions of principal moments of inertia of FG perylene at 570 K and 1 atm.](image)

**Table S3:** Mean values of principal moments of inertia of atomistic perylene at 570 K and 1 atm.

| Property | Value   | Unit          |
|----------|---------|---------------|
| $I_a$    | 820.23  | g mol$^{-1}$ Å$^2$ |
| $I_b$    | 1541.91 | g mol$^{-1}$ Å$^2$ |
| $I_c$    | 2353.61 | g mol$^{-1}$ Å$^2$ |
S6.2 AFM-CG Model Parameters

The parameters used in the AFM-CG MD simulations of perylene are presented in Table S4.

| Parameter | Value | Unit |
|-----------|-------|------|
| $\sigma_0$ | 6.8732 | Å |
| $\sigma_{000}$ | 0.7127 | – |
| $\sigma_{cc2}$ | -0.4305 | – |
| $\sigma_{220}$ | 0.0209 | – |
| $\sigma_{222}$ | -0.0205 | – |
| $\sigma_{224}$ | 0.6406 | – |
| $\sigma_{cc4}$ | -0.0296 | – |
| $\epsilon_0$ | 0.3156 | kcal mol$^{-1}$ |
| $\epsilon_{000}$ | 20.3054 | – |
| $\epsilon_{cc2}$ | 0.0686 | – |
| $\epsilon_{220}$ | -3.8036 | – |
| $\epsilon_{222}$ | -33.7787 | – |
| $\epsilon_{224}$ | -26.2274 | – |
| $\epsilon_{cc4}$ | 2.9238 | – |
| $\epsilon_{000}$ | 6.8376 $\times 10^2$ | – |
| $\epsilon_{cc2}$ | -4.3553 $\times 10^2$ | – |
| $\epsilon_{220}$ | -1.4224 $\times 10^3$ | – |
| $\epsilon_{222}$ | 1.1941 $\times 10^3$ | – |
| $\epsilon_{224}$ | -1.1000 $\times 10^2$ | – |
| $\sigma_{exp}$ | 14.2260 | Å |
| $\kappa_1$ | 8.4527 | – |
| $\epsilon_{osc0}$ | -20.1197 | kcal mol$^{-1}$ |
| $\kappa_2$ | 0.4506 | – |
| $C_2$ | 0.3268 | – |
| $\sigma_{osc}$ | 12.0749 | Å |
| $b$ | -0.4000 | kcal mol$^{-1}$ |

(a) S-function parameters

(b) Additional + pressure correction parameters

(c) Principal moments of inertia

To obtain the optimized parameters for the modified S-function potential (eq (S53)) used in the AFM-CG MD simulations, a least-squares fit of the sum of the S-function term (eq (S54)) and exponential term (eq (S57)) to the basis-expansion potential was first carried out for 500 000 dimer configurations up to a separation distance of 14 Å randomly sampled from the equilibrated FG simulation used for parametrization. Then, similarly to benzene, to fit the significant positive hump in the potential for configurations close to face-to-face, the oscillatory term (eq (S58)) in eq (S53) was added to the fit potential, the set of dimer configurations was augmented with 4200 additional face-to-face, various parallel-displaced (PD) (PD-10, PD-30, PD-50 and PD-70 as denoted in Figure S13), and pi-stacking (SP) (SP-10 and SP-30 as denoted in Figure S13) configurations with separation distances uniformly distributed up to 19 Å, and a least-squares fit was used to optimize the parameters in the oscillatory term with the previously
optimized parameters kept fixed. The lower limit of the separation distance of the dimer configurations added in this step was chosen such that the potential was less than $5 \text{kcal mol}^{-1}$. Basis expansion and fitted modified S-function potential curves for various perylene–perylene dimer relative orientations are shown in Figure S13.

![Fig. S13: Perylene AFM-CG pair potential (without pressure correction) for various dimer relative orientations using the basis expansion (solid lines) and fitted modified S-function (dashed lines). The scalar variables $\beta, \phi, \theta$ (defined in section S3) that describe the relative orientation for these configurations are face-to-face ($0^\circ,0^\circ,0^\circ$), edge-to-face ($90^\circ,0^\circ,90^\circ$), edge-to-edge ($0^\circ,90^\circ,90^\circ$), PD-x (parallel-displaced) ($0^\circ,x^\circ,x^\circ$), Cross-x ($x^\circ,90^\circ,90^\circ$), SP-x ($\pi$-stacking) ($x^\circ,0^\circ,x^\circ$), and ST-x ($x^\circ,x^\circ,90^\circ$).](image-url)
The pressure-corrected potential does not differ substantially to the modified S-function potential fit to the uncorrected AFM-CG potential, as shown in Figure S14.

Fig. S14: Perylene AFM-CG pair potential for various dimer relative orientations using the basis expansion (solid lines) and fitted modified S-function with (dotted lines) and without (dashed lines) the pressure correction.
S6.3 NVT Simulations

Properties of the FG perylene model from NPT simulations at 1 atm are compared in this section with those from NVT simulations of CG perylene models in which the density was set to the average density of the FG simulation at the corresponding temperature.

S6.3.1 Radial Distribution Function (RDF)

RDFs of the FG, AFM-CG, Babadi CG, and Berardi CG perylene models are compared at three different temperatures in Figure S15.

**Fig. S15:** RDFs from NPT simulations of FG benzene model at 1 atm and (top) 570 K, (middle) 610 K, and (bottom) 670 K, and from NVT simulations of AFM-CG, Babadi CG, and Berardi CG perylene models with density set to the average density of the FG simulation at the corresponding temperature.
S6.3.2 Angular–Radial Distribution Function (ARDF)

ARDFs from the NVT simulations at 570 K of the FG and AFM-CG perylene models are plotted in Figure S16, while those for the Babadi CG\textsuperscript{11} and Berardi CG\textsuperscript{12} perylene models are plotted in Figure S17. Analogous plots at 670 K are shown in Figures S18 and S19, respectively.

**Fig. S16:** ARDFs from (top) NPT simulation of FG perylene model at 570 K and 1 atm and (bottom) NVT simulation of AFM-CG perylene model with density set to the average FG simulation density at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. The distribution is only visible for 0–20° as it is close to 0 at other angles.

**Fig. S17:** ARDFs from NVT simulations of (top) Babadi CG\textsuperscript{11} and (bottom) Berardi CG\textsuperscript{12} models at 570 K with density set to the average FG simulation density at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. The distribution is only visible for 0–20° as it is close to 0 at other angles.
Fig. S18: ARDF from (top) NPT simulation of FG perylene model at 670 K and 1 atm and (bottom) NVT simulation of AFM-CG perylene model with density set to the average FG simulation density at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. The distribution is only visible for 0–20° as it is close to 0 at other angles.

Fig. S19: ARDFs from NVT simulations of (top) Babadi CG\textsuperscript{11} and (bottom) Berardi CG\textsuperscript{12} models at 670 K with density set to the average FG simulation density at the same temperature, shown as (left) 2D color maps and (right) 1D slices at fixed angle. The distribution is only visible for 0–20° as it is close to 0 at other angles.
S6.4 NPT Simulations

S6.4.1 Angular–Radial Distribution Function (ARDF)

ARDFs from NPT simulations at 570 K and 1 atm of the FG and AFM-CG models are plotted in Figure S20, while those of the Babadi CG\textsuperscript{11} and Berardi CG\textsuperscript{12} perylene models at the same conditions are plotted in Figure S21. Analogous plots at 670 K are shown in Figures S22 and S23, respectively.

**Fig. S20:** ARDFs from NPT simulations of (top) FG and (bottom) AFM-CG perylene models at 570 K and 1 atm, shown as (left) 2D color maps and (right) 1D slices at fixed angle.

**Fig. S21:** ARDFs from NPT simulations of (top) Babadi CG\textsuperscript{11} and (bottom) Berardi CG\textsuperscript{12} perylene models at 570 K and 1 atm, shown as (left) 2D color maps and (right) 1D slices at fixed angle.
Fig. S22: ARDFs from NPT simulations of (top) FG and (bottom) AFM-CG perylene models at 670 K and 1 atm, shown as (left) 2D color maps and (right) 1D slices at fixed angle.

Fig. S23: ARDFs from NPT simulations of (top) Babadi CG and (bottom) Berardi CG perylene models at 670 K and 1 atm, shown as (left) 2D color maps and (right) 1D slices at fixed angle.

References

[1] 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.

[2] Zee, A. Quantum Field Theory in a Nutshell, 2nd ed.; Princeton University Press: Princeton, 2010.

[3] Zhang, L. Dirac Delta Function of Matrix Argument. Int. J. Theor. Phys. 2021, 60, 2445–2472.

[4] Goldstein, H. Classical Mechanics, 3rd ed.; Addison-Wesley: San Francisco, 2002.
[5] Friedman, J.; Hastie, T.; Tibshirani, R., et al. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, 2009.

[6] 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.

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

[8] Van Loan, C. F.; Golub, G. *Matrix Computations*, 4th ed.; The Johns Hopkins University Press: Baltimore, 2013.

[9] 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.

[10] 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.

[11] Babadi, M.; Everaers, R.; Ejtehadi, M. R. Coarse-grained Interaction Potentials for Anisotropic Molecules. *J. Chem. Phys.* 2006, 124, 174708.

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