Soft Matter

PAPER

View Article Online
View Journal | View Issue

Cite this: *Soft Matter*, 2026, 22, 3370

Molecular dynamics investigation of the impact of methylation on the nematic phase of phenyl benzoate mesogens and dimers

Jacob S. Votava, Emily C. Davidson * and Michael A. Webb *

Liquid crystals (LCs) possess anisotropic mechanical and optical properties with applications ranging from soft robotics to display technology. Despite advances in the precise synthesis of liquid crystalline materials, the microscopic origins of substituent effects, which impact functional performance, are not always well understood. Here, we use molecular dynamics (MD) simulations to investigate how methyl substitution affects the nematic phase behavior of liquid crystal monomers and dimers composed of phenyl benzoate cores flanked by aliphatic tails. Methylation induces a decrease in the nematic–isotropic transition temperature. Using data-driven analysis, we find that for monomers this decrease is associated with an increase in flexibility of core-adjacent aliphatic torsions that influence overall mesogen conformation. For dimers, this manifests as a shift from a continuum of accessible conformations in the isotropic phase to occupying more distinct hairpin and extended states in the nematic phase. The latter exhibits a bend angle consistent with experimental signatures of a modulated nematic phase. Together, these results show how minor changes in chemical structure can impact the conformational ensemble of liquid crystals, trading local conformational entropy for global nematic order, in turn influencing their macroscopic transition temperatures.

Received 7th March 2026,
Accepted 13th April 2026

DOI: 10.1039/d6sm00196c

rsc.li/soft-matter-journal

1 Introduction

Liquid crystalline materials possess phases between a disordered isotropic liquid and a crystalline solid. As small molecules, their hierarchical ordering has been exploited in display technology[1,2] and dielectric materials.[3] Polymerizing LCs to form liquid crystal polymers (LCPs) and elastomers can further tailor their mechanical properties and phase behavior, with applications in soft robotics,[4] thermal camouflage,[5] and elastocaloric devices.[6] The nematic phase, one of the most common LC phases, possesses orientational order yet is translationally disordered. Its stability ultimately depends on both entropic and enthalpic interactions, and thus it is highly dependent on molecular shape, rendering it sensitive to small structural changes.[7–9] However, anticipating how any chemical modifications affect phase behavior is challenging, hindering design of novel liquid crystalline materials.

Minor chemical modifications can often yield substantial impacts on phase behavior. For instance, varying terminal chain length or adding lateral methyl/fluoro groups can cause large changes in the transition temperatures.[10–12] In bent-core mesogens, relocating a chiral/branched center in the terminal side chains has been shown to shift the core-chain bend-angle distribution and the torsion statistics near the core-chain junction, which impacts the curvature of B4-type assemblies.[13,14] Yet even for comparatively simple systems, the microscopic origins of these substituent effects remain poorly understood beyond heuristics. For example, lateral methyl substitution is suspected to suppress nematic stability by widening the molecular cross-section,[15] but whether this steric effect alone or more subtle changes to the conformational ensemble are responsible remains unclear. In polymers and oligomers the problem is amplified, as dispersity and lack of sequence control obscure causal relationships.

Using iterative growth polymerization,[16] a platform was recently developed to obtain nearly monodisperse, sequence-defined oligomers of mesogens with phenyl benzoate cores.[17] This has enabled experimental study[18] of the monomers and dimers shown in Fig. 1, revealing pronounced substituent effects. In particular, introducing a single methyl substituent ($P_1 \rightarrow M_1$) suppresses crystallization, stabilizes the nematic phase, and lowers the nematic–isotropic transition temperature, T_{NI}, by 56 K. In addition, dimerization increases T_{NI}, while methylation continues to depress it relative to unmethylated species. The asymmetric heterodimers (PM and MP in Fig. 1B) exhibit sequence-dependent crystallization and differences in

Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08540, USA. E-mail: edavidson@princeton.edu, mawebb@princeton.edu

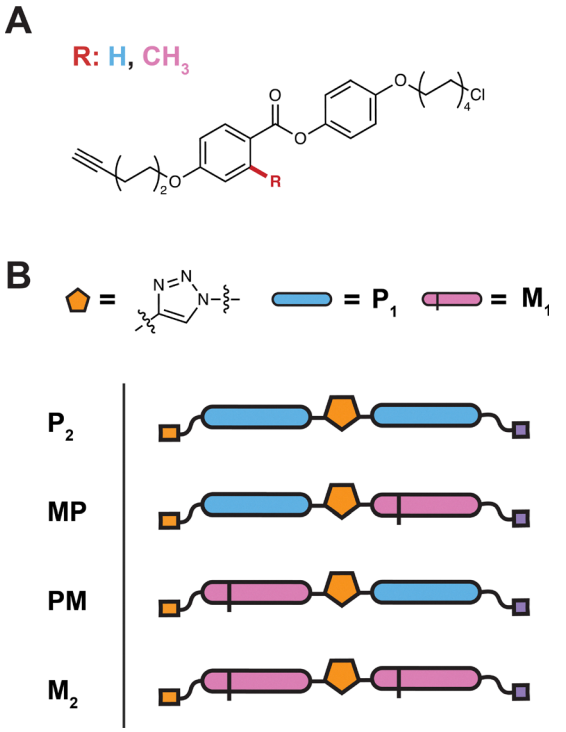

Fig. 1 System overview. (A) Phenyl-benzoate monomer family with R = H,CH_3 defining P_1 and M_1. (B) Triazole-linked dimer sequences P_2, MP, PM, and M_2. Vertical black lines indicate the lateral methyl position on M_1; orange and purple rectangles indicate the alkyne and chlorine terminal groups, respectively.

hierarchical organization, despite minimal differences in transition temperatures. These observations raise questions as to how a single methyl group precisely impacts the phase behavior.

Atomistic molecular dynamics (MD) simulations can help elucidate the origins of LC phase behavior.[19,20] In practice, accurately reproducing LC behavior has historically been approached by refining existing, general-purpose force fields that were originally parameterized for broader chemical spaces.[21,22] For example, the united-atom Transferable Potentials for Phase Equilibria (TraPPE) force field was refined to reproduce the T_{NI} of 4-cyano-4′-pentylbiphenyl by reparameterizing selected torsional terms,[20] and re-optimizing the alkane and biphenyl torsions in the general Amber force field (GAFF), together with scaling select nonbonded parameters, improved agreement with experiments. In particular, the revised model better reproduced densities and heats of vaporization for fragments containing 2,5-diphenyl-1,3,4-oxadiazole and phenyl benzoate cores.[21] However, because many properties, including phase behavior, are sensitive to the specific torsions and observables chosen for fitting, it can be challenging to generalize this targeted re-parameteriztion strategy across various mesogens and architectures. Recent developments in the context of soft materials highlight machine-learned force fields[23] as an attractive, alternative route that could in principle simplify parameterization while capturing more complex many-body interactions that are generally neglected in aforementioned general purpose force fields. Such an approach has been recently exemplified on the canonical model system of 4-cyano-4-pentylbiphenyl (5CB).[24] Nevertheless, accessing system sizes and timescales that are relevant for phase behavior remains a formidable challenge for machine-learned force fields, such that re-parameterizing conventional force fields with high-quality DFT may be the most pragmatic, effective option in the near-term.

In this work, we use atomistic MD to interrogate substituent effects in phenyl-benzoate IEG monomers (P_1/M_1) and their triazole-linked dimers, linking methylation-driven conformational changes to nematic behavior. This investigation is inspired by the experimental systems developed in ref. 18. We model these systems using the general Sage force field[25] but with re-optimization of all rotatable torsions using density functional theory (DFT). The resulting models capture relative shifts in T_{NI} across monomer and dimer systems, in good agreement with experimental trends, while detailed conformational analysis elucidates their origins. In particular, methylation induces a shift in torsion populations and biases larger-scale conformations with a tendency to destabilize nematic phases. Ultimately, this study provides molecular-level interpretation of substituent effects in these phenyl benzoate mesogens, illustrating how minor chemical modifications can alter conformational ensembles and thus phase stability. Additionally, the systematic force-field parameterization and machine-learning analysis approaches demonstrated here are readily transferable to other LC systems.

The remainder of this paper is organized as follows. Section 2 presents our simulation and analysis methods, with particular attention to the force-field optimization procedure. Section 3 discusses our results: we first establish that the re-parameterized models yield good agreement with experimental trends, then analyze temperature-dependent conformational behavior in the monomer systems to identify key differences driven by methylation. Machine-learning models facilitate an unbiased assessment of which molecular features govern these differences. We subsequently examine analogous trends in the dimer systems. Section 4 summarizes the key findings and offers perspective for future work.

2 Methods

2.1 Simulation details

2.1.1 General procedures. All molecular dynamics (MD) simulations were performed using GROMACS 2024.3 with periodic boundary conditions in all dimensions.[26] Equations of motion were integrated with the leapfrog scheme, and bonds involving hydrogen were constrained using the LINCS algorithm, which applies linear constraint solvers to maintain bond lengths, allowing for a 2 fs timestep.[27] The stochastic cell-rescaling barostat and velocity-rescaling thermostat were used to maintain pressure and temperature of each system.[28,29]

Interactions were described using the Sage force field with reoptimization of torsional potentials (see Section 2.2). Charges were determined using the espaloma charge graph neural

network.[30] Real-space nonbonded interactions were smoothly switched to zero between 0.8 and 0.9 nm for all simulations with the force-switch specified for the Sage force field. Long-range electrostatics were handled with particle-mesh Ewald (PME) with a 10^{-5} convergence accuracy.

Following initialization, all systems first underwent energy minimization by steepest-descent with a force tolerance of 1000 kJ mol^{-1} nm^{-1} and a minimization step size of 0.001 nm. Subsequently, systems were equilibrated for 5 ns at a target temperature and 1 atm. Target temperatures ranged from 260–440 K for monomers and 375–475 K for dimers, both in increments of 2.5 K. Configurational sampling was facilitated by replica-exchange molecular dynamics (REMD) across the range of temperatures with exchange attempts every 500 steps. Monte Carlo exchange acceptance rates are provided in SI, Fig. S1.

2.1.2 Monomer systems. For systems of monomers, 256 molecules were randomly placed in a $9 \times 9 \times 9$ nm^3 cubic simulation cell. For P_1, REMD simulations were run for 1.2 s. For M_1, REMD simulations were run for 7.4 s. For both, the final 200 ns were used for analysis. Fig. S2 and S3 establish this is sufficient to obtain reliable statistics on the nematic order parameter. Reported errors correspond to standard errors estimated using block averaging with 10 blocks of 20 ns.

2.1.3 Dimer systems. For dimer systems, 1024 dimers were packed into a $15 \times 15 \times 15$ nm^3 cubic cell by first placing one dimer at the box center and then inserting the remaining dimers at randomly sampled center-of-mass positions with a prescribed, common orientation. This aligned initialization has been shown to reduce equilibration times in similar LC systems.[31,32] Scaling effects are unlikely to impact these simulations, as prior work on nematic transitions in LC systems has shown little difference in phase behavior between system sizes ranging from 256 to 2048 mesogens.[22] Despite this, it does lead to a decrease in the Monte Carlo acceptance rate, reported in Fig. S1. Across all systems, REMD simulations were run for 480 ns, and the final 40 ns were used for analysis. Fig. S4–S7 show the equilibration of the nematic order parameter over this time period. Reported errors correspond to standard errors estimated using block averaging with 10 blocks of 4 ns.

2.2 Force-field optimization

Torsional parameters were re-optimized using the Open Force Field (OpenFF) initiative's BespokeFit software[33] and the protocols described in Section 2.2.1. These protocols were applied separately to the P_1, M_1, and P_2 systems. The resulting parameters were then combined, with redundant terms pruned automatically by the software. All fragments and corresponding torsional scans are provided in SI, Section S2.

2.2.1 Optimization procedure. Torsional parameters were fit using a truncated Fourier series,[33]

$$U_t(\phi) = \sum_{m=1}^{N_t} k_{t,m}\left[1 + \cos\left(n_{t,m}\phi - \psi_{t,m}\right)\right]. \tag{1}$$

Here, $U_t(\phi)$ is the torsional contribution to the potential energy for torsion t as a function of torsion angle ϕ. The sum runs over N_t Fourier terms (typically $N_t \leq 4$). For term m of torsion t, $k_{t,m}$ is the force constant, $n_{t,m}$ is the periodicity (integer), and $\psi_{t,m}$ is the phase offset.

The fit was performed with ForceBalance,[34] using the weighted least-squares torsion objective employed in OpenFF BespokeFit:[33]

$$\mathscr{L}_t(\Phi) = \frac{1}{S_f^2} \frac{\sum_{i=1}^{N_{\mathrm{grid},t}} w\left(E_{\mathrm{QM},t}(x_{t,i})\right)\left[E_{\mathrm{QM},t}(x_{t,i}) - E_{\mathrm{MM},t}(x_{t,i};\Phi)\right]^2}{\sum_{i=1}^{N_{\mathrm{grid},t}} w\left(E_{\mathrm{QM},t}(x_{t,i})\right)}. \tag{2}$$

In eqn (2), $\mathscr{L}_t$ is the objective for a single torsion t, Φ denotes the vector of optimizable force-field parameters, $x_{t,i}$ is the i-th molecular configuration (grid point) along the torsion scan for torsion t, and $N_{\mathrm{grid},t}$ is the number of grid points for that scan. $E_{\mathrm{QM},t}(x_{t,i})$ and $E_{\mathrm{MM},t}(x_{t,i};\Phi)$ are the relative QM and molecular mechanics (MM) energies of configuration $x_{t,i}$, respectively, each referenced to its corresponding minimum. QM energies are computed at the B3LYP-D3BJ/DZVP level of theory, consistent with the transferrable Sage force field. S_f is an energy scaling factor (set to 1.0 kcal mol^{-1}).[33] The weights were assigned as a function of the QM relative energy,[33]

$$w(E) = \begin{cases} 1, & E < E_{\mathrm{low}}, \\ \left[1 + (E - E_{\mathrm{low}})^2\right]^{-1/2}, & E_{\mathrm{low}} \leq E < E_{\mathrm{high}}, \\ 0, & E \geq E_{\mathrm{high}}. \end{cases} \tag{3}$$

In eqn (3), $w(E)$ is the weight assigned to a grid point as a function of its QM relative energy E. The cutoffs were set to E_{low} = 1.0 kcal mol^{-1} and E_{high} = 10.0 kcal mol^{-1}.[33]

Using contributions from all torsions, an aggregate objective was minimized with regularization,[33]

$$\mathscr{L}_{\mathrm{total}}(\Phi) = \sum_{t=1}^{N_{\mathrm{tors}}} w_t \mathscr{L}_t(\Phi) + w_{\mathrm{reg}} \sum_{p=1}^{N_p} \frac{|\Delta\Phi_p|}{\sigma_p^2}. \tag{4}$$

In eqn (4), N_{tors} is the number of torsions (targets) included in the fit, and w_t is an optional user-chosen weight for torsion t. w_{reg} is the regularization prefactor, N_p is the number of optimizable parameters, $\Delta\Phi_p = \Phi_p - \Phi_p^{(0)}$ is the change in parameter p relative to its initial value $\Phi_p^{(0)}$, and σ_p is the prior width for parameter p (σ_p = 6.0).[33]

2.3 Structural analysis

2.3.1 Nematic order parameter. The nematic order parameter S is used to track the nematic–isotropic phase transition.[35] For any given configuration, it is computed as the largest eigenvalue of the tensor

$$\mathbf{Q}(t) = \frac{1}{2N}\sum_{i=1}^{N}\left[3\hat{\mathbf{u}}_i(t) \otimes \hat{\mathbf{u}}_i(t) - \mathbf{I}\right] \tag{5}$$

where $\hat{\mathbf{u}}_i$ is the unit director vector, which points along the axis connecting the centers of mass of the two phenyl rings in the LC core.

2.3.2 Nematic–isotropic transition temperature. To determine the T_{NI}, the empirical Haller equation is used:[32,36]

$$S = (1 - S_0)\left(1 - \frac{T}{T_{NI}}\right)^{\beta} + S_0 \quad (6)$$

where S_0 is the order parameter of the isotropic phase and β is an empirical fitting parameter. Least-squares regression of the Haller fit was performed over temperature ranges of 340–415 K for P_1, 300–340 K for M_1, and 375–472.5 K for all dimer systems, allowing for assignment of T_{NI}. This empirical form is used solely to provide a consistent estimate of T_{NI} across systems for comparison with experimental trends, rather than for mechanistic interpretation.

2.3.3 Radius of gyration. The radius of gyration R_g was computed as a general conformational characterization of monomers and dimers. This was obtained by first calculating the gyration tensor as

$$\mathbf{G} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{r}_i - \mathbf{R}^{COM}) \otimes (\mathbf{r}_i - \mathbf{R}^{COM}) \quad (7)$$

where $\mathbf{r}_i$ is the position of atom i, and $\mathbf{R}^{COM}$ is the center of mass of atoms $1 - N$. Diagonalizing $\mathbf{G}$ yielded associated eigenvalues $\lambda_x{}^2 \leq \lambda_y{}^2 \leq \lambda_z{}^2$, which represent the principal moments of the molecule. The radius of gyration was then calculated with

$$R_g = \sqrt{\lambda_x{}^2 + \lambda_y{}^2 + \lambda_z{}^2}. \quad (8)$$

2.3.4 Bend angle of dimers. For dimers, molecular conformation was additionally characterized by the bend angle between the two mesogenic cores. As described earlier, each LC core possesses a unit director vector $\hat{\mathbf{u}}_i$ between the two aryl rings. Using these, the bend angle θ between the two cores was computed as

$$\theta = \arccos[\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2] \quad (9)$$

where $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ correspond to the two mesogenic cores within the dimer.

2.4 Machine learning details

2.4.1 R_g regressor. To facilitate attribution of specific molecular characteristics to mesogen conformation, machine learning models were constructed to predict radius of gyration from molecular features. In particular, input features based on torsion angle distributions were supplied to gradient-boosted decision tree models optimized with the XGBoost package with mean-squared-error loss.[37]

Input feature vectors were obtained as follows. First, torsion angles ϕ_i were computed for each mesogen in every simulation frame. For each torsion type i, circular statistics were computed across all 256 mesogens within a frame. The use of circular statistics is to properly account for the periodicity of torsional angles. The circular resultant length was calculated as

$$\bar{r}_i = \sqrt{\langle \sin \phi_i \rangle^2 + \langle \cos \phi_i \rangle^2}, \quad (10)$$

from which the circular standard deviation was obtained as

$$\sigma_i = \sqrt{-2 \ln \bar{r}_i}. \quad (11)$$

Each frame was then represented by a feature vector consisting of the mean circular standard deviations $\{\bar{\sigma}_i\}$ from each frame, with the frame-averaged R_g as the target value.

Hyperparameters were selected *via* grid search over tree depth (5–7), learning rate (0.03–0.05), number of estimators (400–800), subsample fraction (0.8–0.9), and column subsample fraction (0.8–0.9), using three-fold cross-validation with shuffled splits and R^2 scoring as implemented in scikit-learn.[38] The final model achieved R^2 = 0.988 and root-mean-squared error of 0.0218 Å, indicating strong predictive performance. The model parity plot (SI Fig. S11) and the corresponding feature distributions (SI Fig. S12) are provided in the SI to assess fit quality and support feature attribution.

2.4.2 Shapley additive explanations (SHAP) analysis. To rank the importance of various torsions on the predictions of the R_g regressor and system classifier we calculated the SHAP value of each model. SHAP assigns each feature a fair share of a model's prediction by averaging its marginal contribution over all possible feature coalitions (Shapley values). Explanations were computed with a random subsample (up to 4000 examples) and a 1000-sample background using the SHAP package.[39] For each torsion feature ϕ_i^{std}, importance was reported as the mean absolute SHAP value,

$$I(\phi_i) = \text{mean}\left(\left|\text{SHAP}_{\phi_i^{std}}\right|\right). \quad (12)$$

where $I(\phi_i)$ represents the importance of torsion i. For visualization, the smallest contributions were combined into another category by summing their mean absolute SHAP values.

3 Results and discussion

3.1 Systematic reparameterization of torsion potentials

Given the absence of prior computational studies on the LC systems in Fig. 1, we first sought to establish a robust force field for simulation. Fig. 2A compares single-molecule torsional energies produced by the general Sage force field with those obtained from DFT at the B3LYP-D3BJ/DZVP level of theory; the substantial discrepancies (R^2 = 0.742) establishes the need for force-field refinement. By re-parameterizing all rotatable torsional modes as described in Section 2.2, we markedly improved correspondence with the quantum-chemical reference data (R^2 = 0.986; Fig. 2B). Notably, this agreement is achieved through systematic optimization of all rotatable bonds, rather than selective reparameterization of specific torsions chosen by ansatz. This highlights some potential for the approach to be readily extended to other LC systems leveraging the OpenFF software ecosystem.[33]

Fig. 2C–E presents three representative torsional profiles that illustrate the quality of refinement achieved through re-optimization; the full set is provided in the SI. First, the *trans–gauche* energy barrier in aliphatic torsions decreases from ~7 kJ mol^{-1} to ~2.5 kJ mol^{-1} for the hexane fragment

(Fig. 2C), consistent with modifications made to GAFF in prior work.[21,22] This reduction increases the population of *gauche* conformers, yielding more flexible chains that can influence mesophase stability. Second, re-optimization of the aryl-ester torsion in the phenyl benzoate core (Fig. 2D) removes two spurious energy minima and lowers the barrier between the remaining states, allowing greater flexibility that may help relieve packing frustrations in the nematic phase. Third, the torsion controlling core planarity (Fig. 2E) acquires a new local minimum upon optimization, influencing resistance to deformation. These examples illustrate subtle but important corrections that systematic optimization captures but *ad hoc* reparameterization of select torsions could miss.

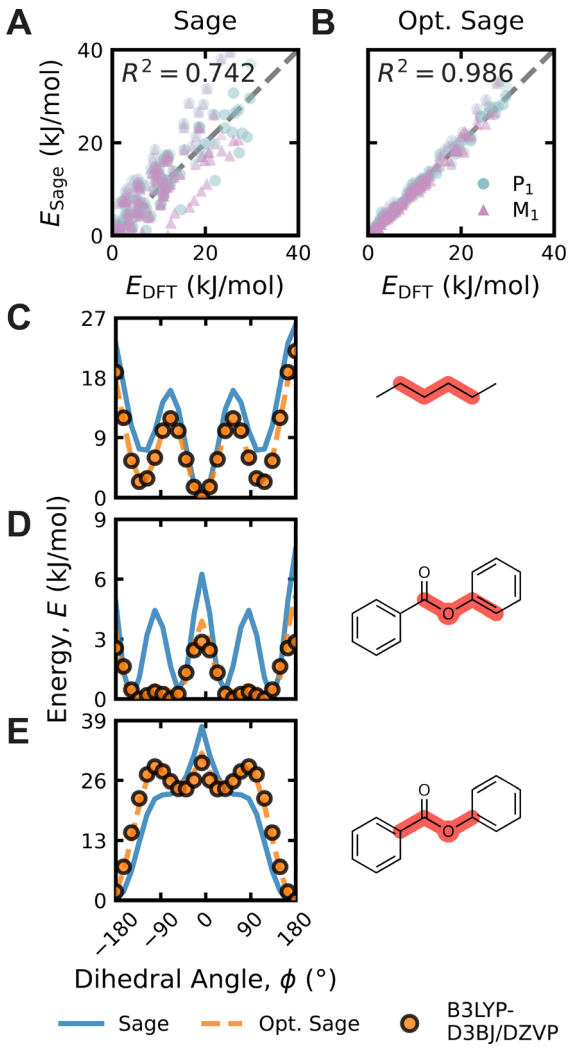

Fig. 2 Force-field optimization and representative mode scans. Parity plots comparing molecular mechanics (MM) and density functional theory (DFT) total energies of Sage (A) before (labeled as 'Sage') and (B) after force-field optimization (labeled as 'Opt. Sage'). Coefficients of determination R^2 are reported in each panel. (C)–(E) Representative torsional energy scans of selected fragments, comparing total energy with initial Sage parameters, refitted parameters, and DFT reference data at the B3LYP-D3BJ/DZVP level of theory. The legend below panel E also applies to panels C and D. For each panel, the shaded red region highlights the torsional mode being scanned, with ϕ denoting the dihedral angle defined by the four highlighted atoms.

3.2 Simulations capture trends in methylation on nematic transition temperatures

Having established agreement between our force field and quantum chemical calculations, we next assessed whether the simulated systems recapitulate experimental trends. Experimentally, a 56 K difference in T_{NI} is reported between P_1 (337 K) and M_1 (281 K).[18] To probe this, we performed replica-exchange MD across a wide temperature range and calculated the nematic order parameter as a function of temperature. The simulation snapshots in Fig. 3A illustrate that the models transition from a nematic phase at low temperatures to an isotropic phase at high temperatures. Fig. 3B tracks this behavior quantitatively, comparing S across all simulated temperatures for the P_1 and M_1 systems. It is apparent that the simulated P_1 forms a more ordered nematic phase relative to M_1 ($S \approx 0.7$ compared to $S \approx 0.4$). Importantly, both systems are amenable to extracting T_{NI} by eqn (6). Although the transition temperatures ($\approx$400 K for P_1, $\approx$350 K for M_1) are systematically high relative to experiment, which is a common artifact, their relative difference aligns with experiment in terms of both the magnitude and direction. This validates that the underlying models are sensitive to methylation as a substituent effect.

The phase transition observed in both systems is also consistent with general expectations for a nematic–isotropic transition. Namely, the sigmoidal decrease in S in Fig. 3B is the hallmark of a weakly first-order transition, while Fig. 3C displays that the system density is continuous through the transition. All together, this agreement suggests that the simulations reliably capture the influence of methylation on an isotropic-nematic transition, motivating detailed conformational analysis to identify the molecular origins of these effects.

3.3 Methylation induces subtle conformational changes in the monomer systems

As a first step towards reconciling the differences in observed phase behavior between M_1 and P_1, we computed the radius of gyration R_g as a convenient proxy for the spatial extent of the mesogens. The distribution of alternative shape descriptors can be found in the SI, Section S5, and reflect similar trends. Fig. 4A compares the temperature-dependent R_g distributions for M_1 and P_1. In the nematic phase (low temperatures), the breadth of the distributions is notably different, and methylation seemingly induces slightly more compact (smaller R_g) mesogens. However, the distributions appear to converge at higher temperatures, for which both M_1 and P_1 are in the isotropic phase. These trends are quantitatively confirmed in Fig. 4B, which tracks the variance of the R_g distributions across temperatures. The variance of M_1 is initially higher than that of P_1 at low temperatures and through the former's T_{NI}, but the values effectively converge at and above the T_{NI} of P_1. This data establishes that P_1 and M_1 are characterized by different conformational ensembles in their respective ordered phases. This

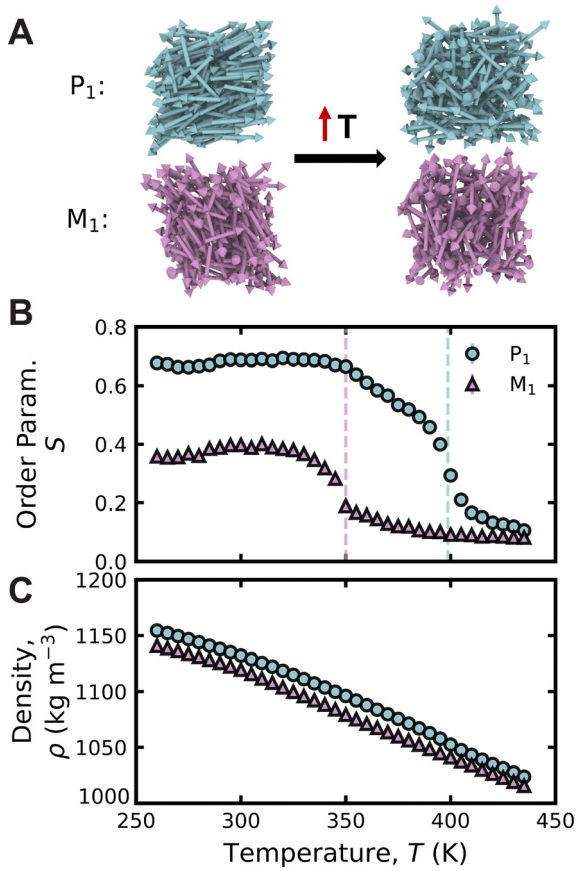

Fig. 3 Analysis of macroscopic phase behavior in the monomers. (A) Simulation snapshots showing the M_1 (bottom) and P_1 orientations at 300 K (left) and 435 K (right). Mesogens are represented by arrows which represent their director vector. (B) Comparison of the nematic order parameter S between P_1 and M_1. Dashed lines represent the results of the Haller fit, eqn (6), yielding $T_{NI} \approx 400$ K for P_1 and ≈ 350 K for M_1. (C) Comparison of mass density ρ *versus* T for the same systems. For B, C, error bars represent the standard error from block averaging and are not visible under the markers. Data is shown in 5 K increments for visual clarity.

data establishes that P_1 and M_1 are characterized by different conformational ensembles in their respective ordered phases.

For chemically similar mesogens, broader configurational distributions would be expected to correspond to higher-entropy systems. This is consistent with the observed increase in R_g variance as both systems approach the isotropic phase (Fig. 4B). Notably, P_1 exhibits consistently narrower R_g distributions than M_1, indicating more uniform molecular conformations that correlate with its higher nematic order parameter at a given temperature. This conformational uniformity suggests a more rod-like ensemble that effectively supports nematic alignment and packing in P_1. By contrast, methylation appears to broaden the conformational distribution, disrupting the uniformity which supports stable nematic ordering.

To identify which specific molecular features underlie these conformational differences, we trained a gradient-boosted regression model to predict R_g from characteristics of underlying torsional modes. In particular, the model was trained on torsion angle distributions, with the mean circular standard deviation of each rotatable bond serving as the input features. Our rationale was that knowing the flexibility of each torsion would enable effective prediction of R_g, and then identifying the most important modes for this prediction would provide insight into salient differences between the P_1 and M_1 systems. While only the mean circular standard deviations are used for this analysis, the full underlying torsional probability distributions are reported in the SI, Section S4.

Fig. 5 examines the relative importance of the various torsional modes in terms of their influence for predicting the mean R_g. This is specifically conveyed by rank-ordering modes by their mean absolute SHAP values. In the context of explainable machine learning, SHAP is an attribution analysis method that quantifies each feature's contribution to model predictions.[39] Here, the analysis reveals that aliphatic torsions dominate the model output, with those closest to the LC core (*e.g.*, ϕ_1 and ϕ_6) contributing most strongly to predictions of R_g. Although statistical correlations between successive torsional angles are expected[40] and SHAP analysis does not explicitly account for such correlations, the observed decay in importance with distance from the core supports the physical interpretation that torsions nearest the core most strongly influence mesogen conformation. This ultimately informs a physical hypothesis that methylation influences nematic stability by altering torsional populations in the aliphatic fragments adjacent to the mesogen core.

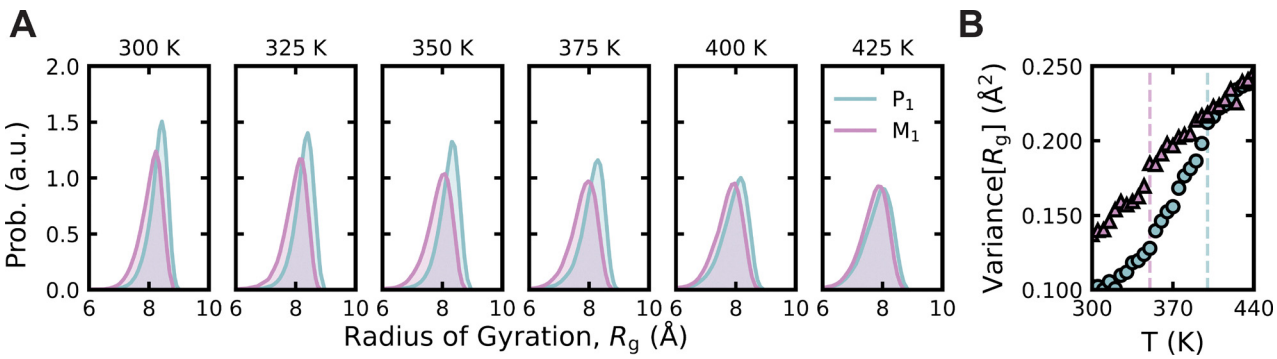

Fig. 4 Analysis of temperature-dependent conformational distributions. (A) Comparison of R_g distributions for M_1 and P_1 (300–425 K in 25 K increments). (B) Comparison of the variance of the R_g distributions between M_1 (pink triangles) and P_1 (blue circles) across temperatures.

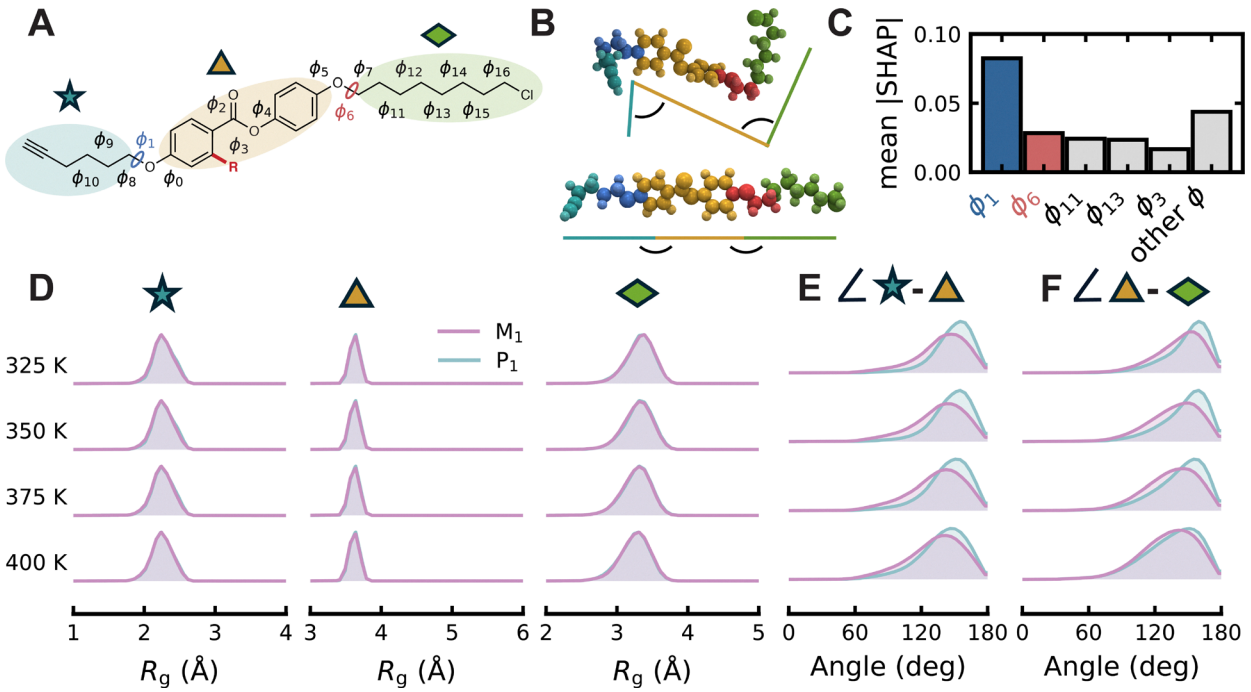

Fig. 5 Machine learning analysis of torsion importance in monomers. (A) Rotatable bond labels for P_1 and M_1; fragments for gyration tensor analysis are highlighted in blue, gold, and green (star: alkyne tail; triangle: biphenyl; diamond: chlorine tail). Torsion distributions can be found in SI Section S4. (B) Molecular render of P_1 illustrating inter-fragment angles, with ϕ_6 atoms in red and ϕ_1 atoms in blue. P_1 system is shown in two different states. (C) SHAP feature importance for the R_g regressor, highlighting ϕ_1 (blue) and ϕ_6 (red). (D) Fragment R_g distributions for the alkyne, biphenyl, and chlorine tail fragments in P_1 (blue) and M_1 (pink), showing negligible conformational differences. (E) Principal-axis angle distribution between the alkyne and biphenyl fragments. (F) Principal-axis angle distribution between the biphenyl and chlorine tail fragments.

To characterize the structural implications of the key torsional modes, we partitioned each mesogen into three fragments defined by the ϕ_1 and ϕ_6 torsions identified as most important by the regressor. The probability distributions of these dihedrals can be found in Fig. S13 and S14. Fig. 5A illustrates the partitioning: two aliphatic tails and the biphenyl core, with Fig. 5B showing a render of these fragments and the angles between them. Fig. 5D shows that the spatial extent of each fragment is similar between P_1 and M_1, indicating that conformational differences do not arise within these individual fragments. However, Fig. 5E and F show that the angles formed between adjacent fragments differ markedly between the two systems. In P_1, the distribution of angles peaks around 150°, whereas in M_1 the distributions peak at somewhat higher values while also being broader. This implies that the methyl substituent promotes more bent mesogen geometries. Because the individual fragments have similar spatial extents, this geometric bending explains the enhanced compaction (lower R_g) observed in M_1 relative to P_1 in Fig. 4A.

Based on this analysis, we hypothesize that methylation alters phase behavior by increasing the flexibility of torsional modes adjacent to the mesogen core. This added torsional freedom leads to an increase in the conformational entropy, which ultimately disrupts global orientational order, lowering the T_{NI}.

3.4 Increasing methylation reduces the nematic transition temperature in dimeric systems

Having understood what drives phase behavior in the monomers, we now turn to understanding the nematic phase of the dimers. Towards this, we apply the systematically optimized force field to the dimer systems. Experimentally, P_2, MP, PM, and M_2 each possess a nematic–isotropic phase transition at 358 K, 323 K, 326 K and 296 K respectively, representing an increase in nematic stability relative to the monomers in all dimers.[18] Similar to the monomeric system, increasing methylation leads to a relative decrease in T_{NI}. This is reflected in Fig. 6A, where each system appears to undergo a weakly first-order transition in the nematic order parameter, while the observed phase transitions are continuous in density (Fig. 6B). We calculate the biaxial order parameter (SI Section S6) to demonstrate that the nematic is uniaxial. Notably, the dimer systems exhibit more consistent uniaxial nematic order parameter values across a narrower range compared to their monomer counterparts in the nematic phase. Among the dimers, P_2 displays a slightly reduced order parameter relative to P_1, whereas M_2 shows a modest increase over M_1. The density again decreases with methylation, reflecting less efficient packing.

The dimers share many chemical features with the monomers, but the linker introduces added complexity. Similar to the monomer system, we characterize the structural differences

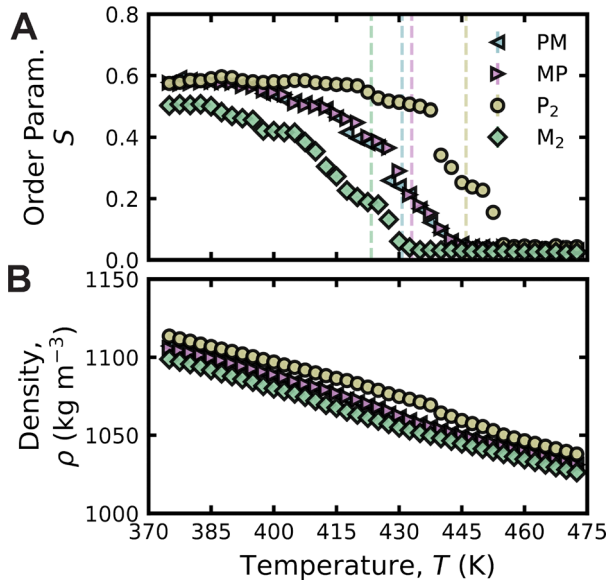

Fig. 6 Analysis of macroscopic phase behavior in the dimers. (A) Comparison of the first-order transition in nematic order parameter *S versus* temperature *T* for dimer systems. Dashed lines are the results of the Haller fits, giving $T_{NI} \approx 435$ K for PM and MP, $\approx$423 K for M_2, and $\approx$447 K for P_2. (B) Smooth transition of mass density ρ *versus* *T* for the same systems. PM and MP show nearly overlapping densities across the entire temperature range. For A, B, error bars represent the standard error from block averaging and are not visible under the markers.

in each phase through R_g in Fig. 7A. Again, we see a tradeoff in conformational entropy and thermal stability of global orientational order. Unlike the monomeric systems where there is a difference in the variance of a unimodal distribution, the R_g distributions of the dimers are bimodal. To further understand these phases, we turn to the distribution of the angle between mesogenic cores in Fig. 7B and look at the joint probability distributions in Fig. 7E and F to elucidate the differences in the conformational ensemble of the dimers.

Across the T_{NI}, the R_g distributions reveal a clear conformational shift from a bimodal distribution to a more uniform one. This is coincident with similar changes in the bend-angle distribution. Examining the joint probabilities below the T_{NI} clearly conveys the bimodal character of both order parameters. This is reflective of a extended conformation (rendered in Fig. 7C) with a bend angle of $\sim$140–150° and an R_g of $\sim$15–16 Å and a hairpin conformer (rendered in Fig. 7D) with a bend angle of $\sim$20–30° and $\sim$9–10 Å. These conformers have been reported previously in the nematic phases of other dimeric, bent-core mesogen simulations.[31,42] The relative populations of extended and hairpin states evolve across their T_{NI}, eventually blending into the continuous distribution seen in Fig. 7F as the systems enter the isotropic state. This is reminiscent of the monomer systems, where an entropic tradeoff increases the available states of the aliphatic tails while disrupting nematic stability. The aliphatic linkers allow for even more flexibility, increasing the accessible states in the dimeric conformational ensembles.

Taken together, these results support the claim that methylation increases conformational entropy, destabilizing global orientational order. Similar to the monomers, the number of accessible states increases with increasing methylation. (Fig. 7A and B) In this picture, methylation shifts the balance between

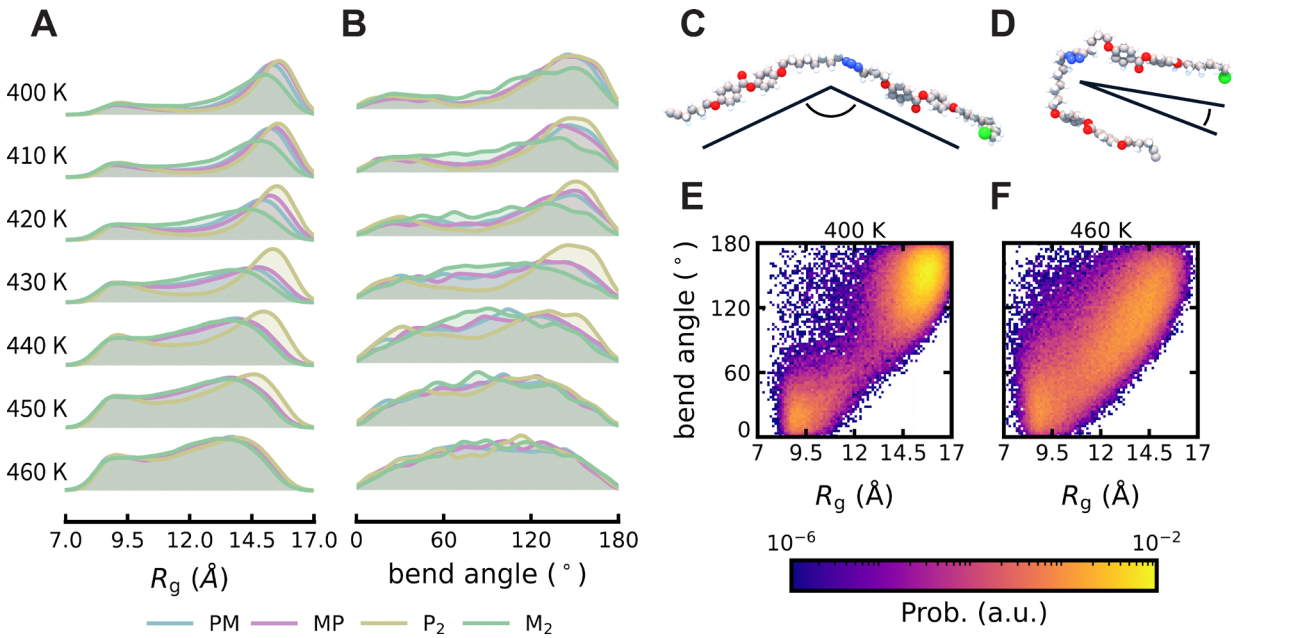

Fig. 7 Conformational analysis of the dimers. (A) and (B) Normalized probability distributions for the radius of gyration R_g and molecular bend angle respectively, for PM, MP, P_2, and M_2 across temperatures from 400–460 K in 10 K increments. (C) and (D) Representative molecular conformations illustrating extended/bent and hairpin-like states. Molecular renderings were generated in OVITO with atoms colored by element (O = red, Cl = green, C = gray, H = white).[41] The bend angle is visually represented by black lines. (E) and (F) Comparison of joint probability distributions of R_g and bend angle for P_2 at 400 K (nematic) and 460 K (isotropic).

extended and hairpin-like states and smooths the bimodal distribution, consistent with increased conformational entropy and the observed reduction in T_{NI}.

We note that oligomerization can give rise to emergent behavior, including twist-bend nematic phases associated with flexible, bent molecular shapes.[43,44] In the homotrimer of the P system, cooling-induced changes in viscosity and optical birefringence indicate a nematic–nematic transition, likely involving a lower-symmetry nematic phase such as the twist-bend nematic.[17] Although capturing such behavior is beyond the scope of the present study, the results here show that nonlinear extended dimer conformations are common. It is conceivable that conformational biases induced by substituents could propagate in oligomers and thereby influence hierarchical ordering and complex phase behavior.

4 Conclusions

We investigated the microscopic origins of substituent effects in phenyl benzoate liquid-crystalline monomers and dimers, focusing on how small chemical changes influence packing (*e.g.*, density and ordering) and conformational ensembles, in turn changing the T_{NI}. Using simulations with a refined force field, we showed that the monomer shift in T_{NI} stems from methylation driving a broader distribution in core-adjacent aliphatic torsions, trading local conformational entropy for reduced nematic order. These simulations not only reproduced experimental shifts between monomers but also revealed temperature-dependent conformational changes that connect molecular shape to phase behavior in dimers providing a more complete microscopic picture of these effects. When applied to dimers, the optimized force field reproduced the experimentally observed trend that increasing methylation lowers the nematic transition temperature. MD simulations revealed two dominant conformations a bent state at $\approx 150^\circ$ and a hairpin state at $\approx 30^\circ$. These findings align with experimental indications of a possible nematic–nematic phase transition, with a bend angle previously supported only by gas-phase DFT calculations.[17] Together, these results demonstrate that while lateral methylation suppresses nematic stability, consistent with heuristics about molecular cross-section, it also leads to changes in the conformational ensemble, increasing conformational entropy and thus lowering the T_{NI} While this study focuses on a particular set of LCs, the insights have broader relevance to the LC field. Because these materials can be synthesized with precise sequence control and narrow dispersity, they provide a platform for probing sequence effects in LC polymers. We also demonstrate that the Sage force field, when re-optimized with DFT but without specific reference to experiments, can reproduce experimental trends in transition temperatures for these systems. This provides a foundation for future studies of how the triazole linker influences dimer geometry in the uniaxial nematic phase and for establishing general structure–shape relationships in this class of materials. In the future, it may be informative to to investigate the difference induced by the triazole linker *versus* more commonly seen aliphatic linkers.[43] Additionally, the conformational changes identified here are expected to produce measurable changes in the Frank elastic moduli (K_1, K_2, K_3), connecting single-molecule shape to macroscopic nematic elasticity and representing a natural extension of this work. These results may also have implications for bottom-up coarse-grained model development, as the inter-fragment angles identified here as key to the conformational differences suggest that interaction terms between fragments must be carefully tuned to reproduce mesophase behavior. In recent work, the orientation-dependent interactions between anisotropic sites have been shown to be critical in coarse-grained models of related systems.[45] It may therefore be of interest to probe similar phenomena in the context of coarse-grained models that are also amenable to analysis of assembly and ordering at much larger scales.

Author contributions

Conceptualization: E. C. D., J. S. V., M. A. W. Methodology: E. C. D., J. S. V., M. A. W. Investigation: J. S. V. Supervision: E. C. D. and M. A. W. Writing, review and editing: E. C. D., J. S. V., M. A. W.

Conflicts of interest

There are no conflicts to declare.

Data availability

Relevant code and simulation files associated with this study are available at https://github.com/webbtheosim/md-simulation-files/tree/main/2026-pblcs-mon-dim.

Supplementary information (SI): torsiondrive mode scans, supplemental simulation results, machine learning metrics and feature distributions, and torsional distributions from bulk MD simulations. See DOI: https://doi.org/10.1039/d6sm00196c.

Acknowledgements

J. S. V. acknowledges Dr Ryan Szukalo, Dr Zachary Lipel, Dr Chun Lam Clement Chan, Dr Hang Zhang, and Emily C. Ostermann for their helpful discussions and mentorship. Simulations and analyses were performed using resources from Princeton Research Computing at Princeton University, which is a consortium led by the Princeton Institute for Computational Science and Engineering (PICSciE) and Office of Information Technology's Research Computing. This work was funded by the US Department of Energy grant DE-SC0023023 (J. S. V., E. C. D.) and the National Science Foundation under grant no. 2237470 (J. S. V., M. A. W).

Notes and references

1 D. W. Berreman, *J. Opt. Soc. Am.*, 1973, **63**, 1374.
2 M. Schadt and W. Helfrich, *Appl. Phys. Lett.*, 1971, **18**, 127–128.
3 M. Schadt, *J. Chem. Phys.*, 1972, **56**, 1494–1497.
4 A. Kotikian, R. L. Truby, J. W. Boley, T. J. White and J. A. Lewis, *Adv. Mater.*, 2018, **30**, 1706164.
5 Y. Liu, J. Song, W. Zhao, X. Ren, Q. Cheng, X. Luo, N. X. Fang and R. Hu, *Nanophotonics*, 2020, **9**, 855–863.
6 J. A. Koch, J. A. Herman and T. J. White, *Phys. Rev. Mater.*, 2021, **5**, L062401.
7 Y. Lim, S. Lee and S. C. Glotzer, *ACS Nano*, 2023, **17**, 4287–4295.
8 L. Onsager, *Ann. N. Y. Acad. Sci.*, 1949, **51**, 627–659.
9 P. F. Damasceno, M. Engel and S. C. Glotzer, *Science*, 2012, **337**, 453–457.
10 S. Dixit and R. A. Vora, *Mol. Cryst. Liq. Cryst.*, 2014, **592**, 133–140.
11 X. Chai, Z. Che, J. Li, M. Hu, J. Zhang, D. Wan, L. Mo, H. Xia and J. Li, *Liq. Cryst.*, 2024, **51**, 2094–2103.
12 M. M. Naoum, A. A. Fahmi, G. R. Saad and M. H. Ali, *Liq. Cryst.*, 2017, **44**, 1664–1677.
13 J. Liu, S. Shadpour, M. E. Prévôt, M. Chirgwin, A. Nemati, E. Hegmann, R. P. Lemieux and T. Hegmann, *ACS Nano*, 2021, **15**, 7249–7270.
14 A. Gowda, G. Acharjee, S. K. Pathak, G. A. R. Rohaley, A. Shah, R. P. Lemieux, M. E. Prévôt and T. Hegmann, *Mater. Horiz.*, 2024, **11**, 5550–5563.
15 M. M. Naoum, A. A. Fahmi, N. H. Ahmed and G. R. Saad, *Liq. Cryst.*, 2015, 1–11.
16 Y. Jiang, M. R. Golder, H. V.-T. Nguyen, Y. Wang, M. Zhong, J. C. Barnes, D. J. C. Ehrlich and J. A. Johnson, *J. Am. Chem. Soc.*, 2016, **138**, 9369–9372.
17 C. L. C. Chan, E. C. Ostermann, S. M. Maguire, Z. Schmidt, J. S. Votava, P. Wasik, M. A. Webb and E. C. Davidson, *Sci. Adv.*, 2025, **11**(34), eadw5327.
18 E. Ostermann, C. L. C. Chan, E. Reed, S. Maguire and E. Davidson, *Chemrxiv*, 2026, preprint.
19 M. R. Wilson, G. Yu, T. D. Potter, M. Walker, S. J. Gray, J. Li and N. J. Boyd, *Crystals*, 2022, **12**, 685.
20 J. Zhang, J. Su and H. Guo, *J. Phys. Chem. B*, 2011, **115**, 2214–2227.
21 N. J. Boyd and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2015, **17**, 24851–24865.
22 N. J. Boyd and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1485–1496.
23 N. E. Jackson, M. A. Webb and J. J. de Pablo, *Curr. Opin. Chem. Eng.*, 2019, **23**, 106–114.
24 Y. Jin, G. R. Perez-Lemus, P. F. Zubieta Rico and J. J. De Pablo, *J. Phys. Chem. A*, 2024, **128**, 7257–7268.
25 S. Boothroyd, P. K. Behara, O. C. Madin, D. F. Hahn, H. Jang, V. Gapsys, J. R. Wagner, J. T. Horton, D. L. Dotson, M. W. Thompson, J. Maat, T. Gokey, L.-P. Wang, D. J. Cole, M. K. Gilson, J. D. Chodera, C. I. Bayly, M. R. Shirts and D. L. Mobley, *J. Chem. Theory Comput.*, 2023, **19**, 3251–3275.
26 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
27 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
28 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
29 M. Bernetti and G. Bussi, *J. Chem. Phys.*, 2020, **153**, 114107.
30 Y. Wang, I. Pulido, K. Takaba, B. Kaminow, J. Scheen, L. Wang and J. D. Chodera, *J. Phys. Chem. A*, 2024, **128**, 4160–4167.
31 G. Yu and M. R. Wilson, *Soft Matter*, 2022, **18**, 3087–3096.
32 R. Berardi, L. Muccioli and C. Zannoni, *ChemPhysChem*, 2004, **5**, 104–111.
33 J. T. Horton, S. Boothroyd, J. Wagner, J. A. Mitchell, T. Gokey, D. L. Dotson, P. K. Behara, V. K. Ramaswamy, M. Mackey, J. D. Chodera, J. Anwar, D. L. Mobley and D. J. Cole, *J. Chem. Inf. Model.*, 2022, **62**, 5622–5633.
34 L.-P. Wang, T. J. Martinez and V. S. Pande, *J. Phys. Chem. Lett.*, 2014, **5**, 1885–1891.
35 N. J. Mottram and C. J. P. Newton, Introduction to Q-tensor theory, *arXiv*, 2014, preprint, arXiv:1409.3542[cond-mat], DOI: 10.48550/arXiv.1409.3542, https://arxiv.org/abs/1409.3542.
36 M. R. Wilson, *Int. Rev. Phys. Chem.*, 2005, **24**, 421–455.
37 T. Chen and C. Guestrin, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
38 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
39 S. Lundberg and S.-I. Lee, A Unified Approach to Interpreting Model Predictions, *arXiv*, 2017, preprint, arXiv:1705.07874 [cs], DOI: 10.48550/arXiv.1705.07874, https://arxiv.org/abs/1705.07874.
40 P. Flory, *Statistical Mechanics of Chain Molecules*, Wiley, New York, NY, 1st edn, 1969.
41 A. Stukowski, *Modell. Simul. Mater. Sci. Eng.*, 2010, **18**, 015012.
42 J. L. Hobbs, C. J. Gibb, E. Cruickshank, R. Walker and R. J. Mandle, *Liq. Cryst.*, 2024, **51**, 1022–1034.
43 R. J. Mandle, *Chemical Record*, 2018, **18**, 1341–1349.
44 C. T. Imrie, R. Walker, J. M. D. Storey, E. Gorecka and D. Pociecha, *Crystals*, 2022, **12**, 1245.
45 A. E. Cohen, N. E. Jackson and J. J. De Pablo, *Macromolecules*, 2021, **54**, 3780–3789.