Multiscale Modelling of Self-assembly in Soft Matter

GARY YU

Multiscale Modelling of Self-assembly in Soft Matter

Gary Yu

A Thesis submitted for the degree of Doctor of Philosophy

Department of Chemistry
Durham University
2022

Abstract

This thesis presents all-atom molecular dynamics simulations and the development of coarse-grained models for various classes of liquid crystals. The overall aim was to parametrise chemically specific models, propagating information between different resolutions through multiscale modelling approaches, to investigate hierarchical self-assembly in soft matter systems. Common coarse-graining methods were assessed in terms of their representability and transferability for applications involving thermotropic calamitic and discotic mesogens, and lyotropic chromonic liquid crystals.

Extensive all-atom simulations were performed on: bent liquid crystal dimers, such as CB7CB; ionic cyanine dyes in aqueous solution (PIC, PCYN, TTBC and BIC); a chromonic perylene bisimide dye (PER); and its thermotropic discotic analogue (PER_{OEG}). These serve as references to parametrise/validate lower resolution models and to provide insights into these systems at the molecular level. For CB7CB, the twist-bend nematic (N_{TB}) phase is observed and characterised. The self-assembly of cyanine dyes and chromonic mesogens was studied by calculating ΔG_{assoc}, ΔH_{assoc} and ΔS_{assoc} for the association of n-mers (where n = 2, 3 or 4). Structures of H-aggregate stacks, with shift and Y junction defects, and J-aggregates with a brickwork arrangement were detected.

Coarse-graining approaches including iterative Boltzmann inversion (IBI), multiscale coarse-graining (MS-CG) in the form of hybrid force matching (FM) and the Martini 3 force field were utilised for the aforementioned systems. A FM model of CB7CB demonstrates high representability and transferability; the N_{TB} phase is captured and the full phase diagram can be explored via heating or cooling. An optimised Martini model correctly exhibits the chromonic nematic and hexagonal phases for PER at the expected concentrations. For PER_{OEG}, an IBI model was found to be superior in modelling the columnar-hexagonal phase. This thesis discusses, in detail, the successes and failures of the various coarse-graining strategies. While successful coarse-graining of liquid crystals remains a challenge, this thesis demonstrates that, with the right choice of method, high-quality coarse-grained models can be developed for both thermotropic and lyotropic systems.

Declaration

The material contained within this thesis has not previously been submitted for a degree at Durham University or any other university. The research reported within this thesis has been conducted by the author unless indicated otherwise.

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Acknowledgements

Principally I would like to thank my supervisor, Prof. Mark Wilson. His enthusiasm and support over the years have been instrumental to my studies, and it has been my privilege to have worked with him. I would also like to thank Dr Martin Walker for his guidance and insightful discussions.

I am grateful to everyone from CG200X for creating a fun, interesting and unique work environment. My time there was truly enjoyable and full of entertaining moments.

I would like to thank the Advanced Research Computing services at Durham University for providing the resources necessary to complete my research.

Finally, I want to express my appreciation to my family and friends for their support during my PhD.

List of Acronyms

N Nematic
Sm Smectic
N_{TB} Twist-bend nematic
Col_h Columnar-hexagonal
M Chromonic hexagonal
CB7CB 1,7-bis-4-(4′-cyanobiphenyl) heptane
PIC Pseudoisocyanine chloride
PCYN Pinacyanol chloride
TTBC 5,5′,6,6′-tetrachloro-1,1′,3,3′-tetraethylbenzimidazolylcarbocyanine chloride
BIC 1,1′-disulfopropyl-3,3′-diethyl- 5,5′,6,6′-tetrachlorobenzimidazolylcarbocyanine sodium salt
PER bis-(N,N-diethylaminoethyl)perylene-3,4,9,10-tetracarboxylic diimide dihydrochloride
AA All-atom
CG Coarse-grained
PMF Potential of Mean Force
RB Ryckaert–Bellemans
LINCS Linear Constraint Solver
PME Particle Mesh Ewald
LJ Lennard-Jones
MD Molecular Dynamics
SD Stochastic Dynamics
COM Centre of Mass
RDF Radial Distribution Function
IBI Iterative Boltzmann Inversion
MS-CG Multiscale Coarse-graining
FM (hybrid) Force Matching
CB Cyanobiphenyl
GAFF General Amber Force Field
FWHM Full Width at Half Maximum
DFT Density Functional Theory
DPD Dissipative Particle Dynamics
REMD Replica Exchange Molecular Dynamics
GB Gay–Berne

Contents

Chapter 1

Introduction

1.1 Self-assembly in soft matter

1.1.1 Liquid crystals

Liquid crystals are a class of matter with properties that are intermediate between those of a liquid and those of a crystal.[1] The molecules within a liquid crystal sample can exhibit various degrees of order which manifests in a vast assortment of possible mesophases, each with distinct optical, physical and mechanical properties. The first liquid crystals to be discovered were derivatives of cholesterol, initially examined by botanist Friedrich Reinitzer in 1888.[2] Reinitzer corresponded with physicist Otto Lehmann stating in a letter "The substance exhibits two melting points, if one may say so. At 145.5 °C it melts to a turbid but absolutely fluid liquid which becomes suddenly clear not until 178.8 °C. On cooling, violet and blue colours appear which quickly vanish with the sample leaving lactescently turbid but fluid. On further cooling the violet and blue colours reappear but very soon the sample solidifies forming a white crystalline mass." After systematically studying these substances, Lehmann presented the paper '*On Flowing Crystals*', laying the foundation for research into liquid crystals.[3] Subsequently, advances in the field were made by numerous scientists, with de Gennes receiving the 1991 Nobel prize in Physics for discovering that "methods developed for studying order phenomena in simple systems can be generalised to more complex forms of matter, in particular to liquid crystals and polymers."[4]

The molecule or component of a liquid crystal responsible for its structure is known

as a mesogen. Mesogenic molecules are generally composed of a rigid core and flexible parts. The rigid component of a mesogen promotes alignment of molecules resulting in the formation of liquid crystal phases, whereas the flexible part provides mobility between molecules and maintains fluidity in the phase. The shape of the mesogen is an important factor in determining the structure of the phase with examples such as rod, disc and bent-core based mesogens.

There are two classes of liquid crystal phases, thermotropic and lyotropic, which are defined by how their phase transitions are controlled. Thermotropic liquid crystals exist in the absence of a solvent and display changes in phase behaviour depending on temperature and pressure. At high temperatures, liquid crystals will adopt an isotropic liquid phase (a phase that appears the same when viewed from any direction) and at low temperatures, will adopt a conventional crystalline or amorphous solid phase.[5] The typical thermotropic phases formed by calamitic (rod-like) mesogens are the nematic and smectic phases (see Figure 1.1). In a nematic (N) phase, molecules have no positional ordering but there is long-range orientational order. These mesophases are highly dependent on the molecular geometry of the mesogens. Smectic (Sm) phases display more order than nematic phases by forming well-defined layers perpendicular to the director, in the smectic A (SmA) phase, or at a tilted angle relative to the director, known as the smectic C (SmC) phase. Lyotropic liquid crystals exist in solvents and their phase behaviour is dependent on both concentration of the components and temperature. Molecules which exhibit lyotropic behaviour tend to be amphiphilic, possessing both hydrophilic and hydrophobic components within the molecule. Lyotropic liquid crystals exhibit structures such as micelles, hexagonal columnar phases, lamellar phases and inverse micellar phases depending on the concentration of the amphiphile.

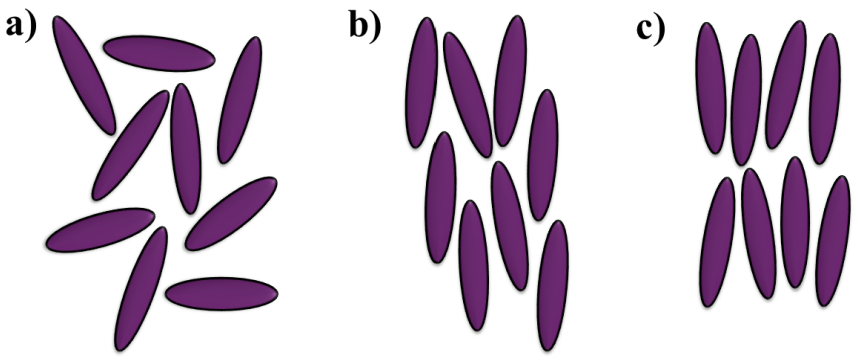


Figure 1.1: Schematic of thermotropic liquid crystal phases formed by calamitic mesogens: (a) isotropic, (b) nematic and (c) smectic.

The simplest model for the isotropic–nematic transition was introduced by Onsager.[6] This theory models mesogens as infinitely long hard rods and explores the effect of repulsion/excluded volume on the free energy of the two phases. As the centres of mass for two idealised cylinders approach, their orientations will determine the amount of excluded volume between them. Thus, an increase in orientational entropy results in a decrease of positional entropy. In the low density regime, maximising orientational entropy results in orientational freedom and an isotropic phase. At a higher density, the balance of positional and orientational entropy leads to parallel arrangements of cylinders being preferred. Therefore, it is predicted that at a sufficiently high density there will be a phase transition from a disordered state to an ordered arrangement due to a change in the dominant component of the entropy at that density.

The nematic phase is characterised by the orientational alignment of molecules in a preferred direction and by low positional order. Beyond the uniaxial classification, nematic phases can also be biaxial, where in addition to the primary director, there is also orientational order along a secondary axis.[7] The chiral nematic (N*) phase features a helical structure where there is a periodic variation of the direction of the nematic director axis through space, and the period in which a full 360 ° rotation is completed is known as the pitch (see Figure 1.2).[1] The helices acquire a handedness depending on the enantiomer responsible for the direction of twist. Principally, the N* phase is achieved in two ways: including chirality in the mesogens or induced from a nematic phase through the addition of a chiral dopant.[8] Similar to nematic phases, chiral smectic phases exist. In the chiral smectic C phase, the helix is formed by a precession of the tilt director of each successive layer around the normal director axis of the phase.[9] Other chiral liquid crystal phases reported are the blue phases and twist grain boundary phases.[10–12]

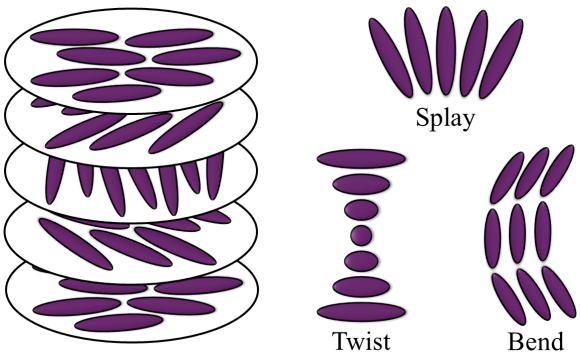


Figure 1.2: Schematic of a chiral nematic phase (left) and splay, twist and bend deformations (right).

The application of an external field to a liquid crystal can be used to align molecules into a phase. When the strength of the electric or magnetic field reaches a threshold, an undistorted nematic liquid crystal will undergo deformation in splay, twist or bend geometries, where this effect is labelled the Fréedericksz transition.[13,14] The three curvature strains of a nematic liquid crystal can be described by the Frank–Oseen free energy density

$$\mathcal{F}_{\text{elastic}} = \frac{1}{2}\left[k_{11}(\nabla \cdot \boldsymbol{n})^2 + k_{22}(\boldsymbol{n} \cdot \nabla \times \boldsymbol{n})^2 + k_{33}(\boldsymbol{n} \times \nabla \times \boldsymbol{n})^2\right], \tag{1.1}$$

where k_{11}, k_{22} and k_{33} are the splay, twist and bend elastic constants, respectively, and $\boldsymbol{n}$ is the director.[15,16] The high sensitivity of nematics to electric fields was first exploited in liquid crystal display devices using the twisted nematic field effect.[17] A pixel in its off-state consists of a twisted nematic sandwiched between two polarising filters, with its molecules aligned parallel to the substrates. Light which passes through the first polariser is rotated by the helical configuration such that it properly passes through the second polariser, resulting in a transparent image. In the on-state, an electric field is applied which reorients the molecules such that they are perpendicular to the substrates. When light is passed through the liquid crystal, it is no longer correctly oriented to pass through the second polariser and the image appears opaque. Nowadays, displays based on vertical alignment[18] and in-plane switching[19] modes are more common. Chiral nematic liquid crystals whose pitch varies with temperature can be used as rudimentary thermometers, as the colour of the reflected light corresponds to the wavelength of the pitch.[20]

The formation of chiral liquid crystal phases from achiral mesogens was predicted by Dozov, who proposed that bent-shaped molecules could cause symmetry breaking in the nematic phase through splay-bend or twist-bend deformations.[21] This was later supported by simulations of achiral banana-shaped molecules in which a helical superstructure was observed.[22] Subsequently, the twist-bend nematic (N_{TB}) phase was reported in cyanobiphenyl dimers with a hydrocarbon spacer of odd parity.[23–26] The N_{TB} phase exhibits periodic twist and bend deformations forming a helical structure with no layering, and occurs at temperatures below the N phase. Bent-core molecules, which have a rigid V-shaped core with flexible chains, can also form unique liquid crystal phases.[27] One such phase, the helical nanofilament phase (HNF), is comprised of twisted bundles of filaments that are constructed from twisted layers.[28] Each layer exhibits strong saddle-splay deformation and curvature which drive the formation of thin, twisted filaments. The rolling of

layers results in heliconical-layered nanocylinders which can further organise into a hierarchy of complex structures.[29,30] The emergent chirality exhibited by achiral mesogens can be rationalised through simulation studies, where distributions of a chirality order parameter demonstrate statistical achirality of bent-core molecules.[31–33] The presence of instantaneous highly chiral conformers was found, with large helical twisting powers, which are able to template regions of chiral order within a bulk phase.

Discotic mesogens are composed of a flat aromatic core with peripherally-attached flexible chains. The least ordered liquid crystal phase formed by these disc-like molecules is the discotic-nematic (N_D) phase, in which molecules have low translational and rotational order around the disc normal but are oriented along a director.[34] The strong π–π interactions between the cores can result in periodic stacking of molecules into columns, which can act as rod-like units for further mesophase organisation (see Figure 1.3). The simplest phase of this category is the columnar-nematic (N_{col}) phase, where the long axis of the stacks are oriented along a director. Higher order columnar phases can be distinguished by the arrangement of columns on a 2D lattice.[35] The columnar-hexagonal (Col_h) phase is defined as having untilted columns packed on a hexagonal lattice with six-fold symmetry. Further columnar phases of a different symmetry are the columnar-rectangular (Col_r) phases or a hexagonal phase with tilted columns, termed the columnar-oblique (Col_{ob}) phase.[35]

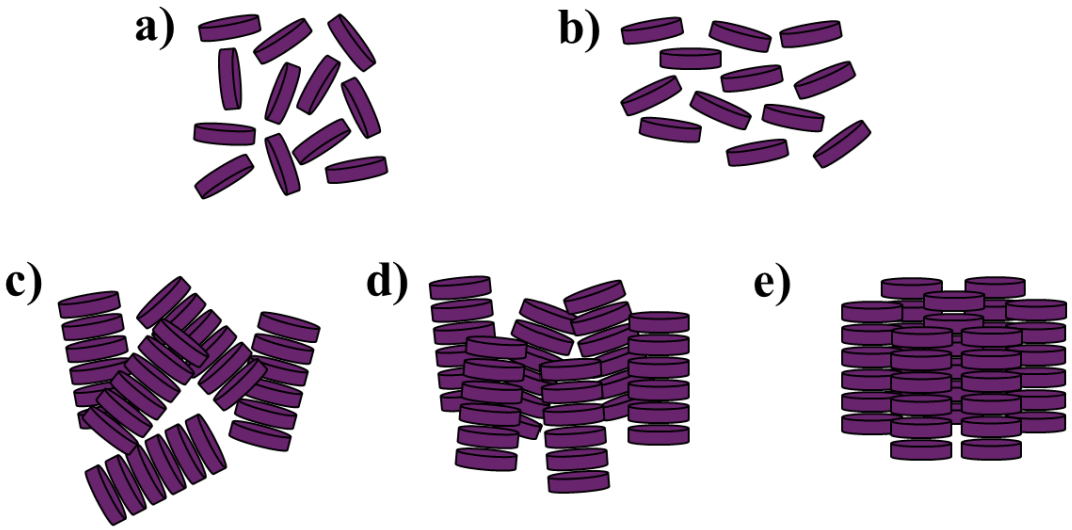


Figure 1.3: Schematic of thermotropic liquid crystal phases formed by discotic mesogens: (a) discotic-isotropic, (b) discotic-nematic, (c) columnar-isotropic, (d) columnar-nematic and (e) columnar-hexagonal.

1.1.2 Chromonic liquid crystals

Amphiphilic compounds can form lyotropic phases upon aggregation of the molecules in a solvent. In water, this is attributed to the hydrophobic effect. It is argued that this effect is largely entropic in origin.[36] Hydrophobic residues disrupt the hydrogen bonding network in water resulting in an unfavourable structuring of water in solvation shells. Aggregation of hydrophobic residues results in reduced surface area exposed to water allowing for minimisation of disruption to the hydrogen bonding network and an increase in entropy from the liberation of water. Amphiphilic compounds forming lyotropic liquid crystals are more commonly referred to as surfactants which, upon sufficient concentration, aggregate into spherical micelles in aqueous media; these are characterised by the favourable solvation of hydrophilic head groups and the shielding of hydrophobic tails from water (see Figure 1.4). The concentration at which surfactant molecules favour aggregation into micelles over existing as monomers in solution is known as the critical micelle concentration (CMC). Further increase in concentration of surfactant results in ordering of self-assembled structures such as bilayers and lamellar phases.

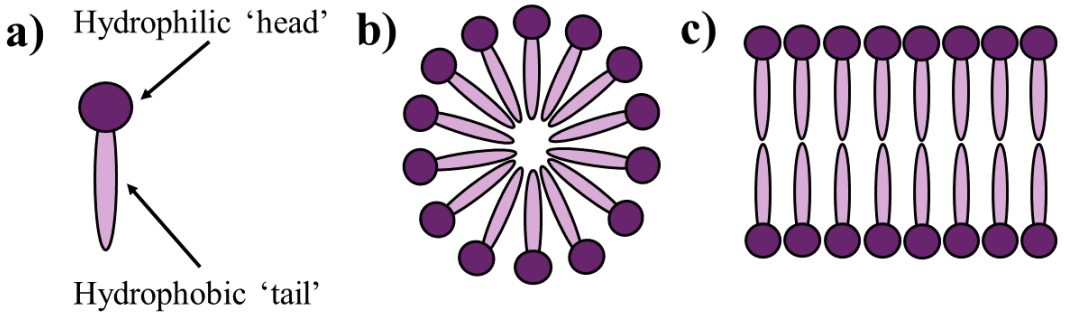


Figure 1.4: Schematic of lyotropic self-assembly: (a) an amphiphile, (b) micelle and (c) bilayer.

An unusual class of lyotropic systems are chromonic liquid crystals. Chromonic mesogens generally consist of a hydrophobic aromatic core with flexible, hydrophilic groups on the periphery for solubility. These molecules aggregate into stacks through face-to-face interactions between the π systems where these stacks can further organise into nematic (N) or hexagonal (M) phases, akin to thermotropic discotic phases.[37] Unlike typical lyotropics, self-assembly into stacks can occur at very low concentrations (without a specific concentration for aggregation like a CMC) and aggregation is unconstrained; aggregation occurs over all concentrations and the aggregation number can greatly vary from low to very

high values.[38] Chromonic self-assembly is considered isodesmic; the addition of a new molecule to a stack provides more or less the same free energy increment, independent of the size of the aggregate.[39–41] The aggregation of molecules into stacks is largely enthalpic in origin rather than entropic, and the association is driven by attractive interactions. Typical association free energies extracted from experimental studies of chromonic systems are in the range of 7 to 14 $k_{\mathrm{B}}T$.[42–46] These free energies of association are also consistent with those calculated via atomistic simulation (Sunset Yellow:[47] 7 $k_{\mathrm{B}}T$, TP6EO2M:[48] 12 $k_{\mathrm{B}}T$, a thiacyanine dye:[49] 14 $k_{\mathrm{B}}T$ and a series of cyanine dyes:[50] 8 to 15 $k_{\mathrm{B}}T$).

The association of aromatic cores in chromonic mesogens gives rise to distinct optical properties which are determined by the angle between adjacent molecular planes.[51–55] Direct face-to-face stacking results in a hypsochromic shift whereas offset staggered stacking is attributed to a bathochromic shift, where these are termed H- and J-aggregates, respectively (see Figure 1.5(a)). J-aggregates exhibit a narrow red-shifted absorption band, compared to the monomer band, which was first observed in a pseudoisocyanine dye.[56,57] The corresponding blue-shift, in H-aggregates, has been observed for similar dyes.[58] From these two association motifs, further complexity can be found for long ranged structures, with suggestions such as layered phases, double-width columns, hollow chimneys and brickwork models.[59,60]

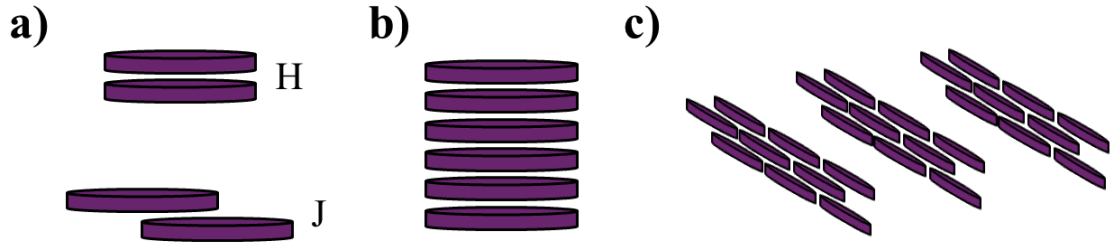


Figure 1.5: Schematic of chromonic self-assembly: (a) arrangement of molecules in H- and J-aggregates, (b) a H-aggregate stack and (c) a brickwork structure in a layered arrangement.

The properties of chromonic mesophases can be exploited for many uses, where these can be from the direct involvement of phases or the use of chromonic phases in preparing materials.[37] In principle, aligned films of chromonic phases can be produced by drying. When an isotropic solution of chromonic mesogens is deposited on a substrate with microfeatures and evaporated, highly ordered ribbons can be obtained.[61] A layer of a photoresponsive polymer on a substrate can be used to align a chromonic N phase. When irra-

diated, the polymer layer reorients which can direct the alignment of the chromonic phase and maintain this orientation during the drying process.[62] Thin films of dried chromonic phases can be used as polarising filters, where limitations in the range of absorption wavelengths can be improved by mixtures of two dyes.[63] A N phase with a chiral dopant can act as a compensator in a twisted nematic display cell, addressing the common greyscale inversion and viewing angle problems.[64] Supramolecular assemblies of extended aromatic dyes, in general, have desirable electronic transport properties which make them useful in organic electronics.[65] A photovoltaic device can be constructed by depositing chromonic columns on a layer of fullerenes, where the electron-accepting properties and high conductivity of the stacks can be useful for light-harvesting purposes.[66] Other applications of chromonics include their use in controllable assembly of gold nanorods[67] and in biosensors[68,69].

1.2 Molecular simulation

Computer simulation methods are powerful tools for the study of problems in soft matter chemistry. Molecular simulation can reinforce or even offer new understanding into experimental studies by providing detailed molecular level insights into a chemical system and allowing for exploration of these systems with a level of freedom unattainable to experiment. For a given phenomenon to be investigated by simulation, a balance must be struck between the level of theory appropriate and its associated computational cost (see Figure 1.6). The use of quantum chemical methods, such as density functional theory[70], can be used to study the electronic structure of single molecules. A classical mechanics approach with atomistic resolution models can allow for study of relatively sizeable systems using molecular dynamics.[71] However, conventional atomistic simulation becomes intractable for the study of phenomena requiring many hundreds or thousands of molecules. In order to access larger length and longer time scales, a coarse-grained simulation approach may be used, where multiple atomistic sites can be combined into a single interaction site.[72] The aim of this approach is to develop models of lower resolution with a reduced number of degrees of freedom, whilst retaining the key chemistry and physics required to study the relevant phenomena. At even longer length and time scales, mesoscale methods can be employed[73], with techniques such as dissipative particle dynamics[74] and lattice Boltzmann

methods[75].

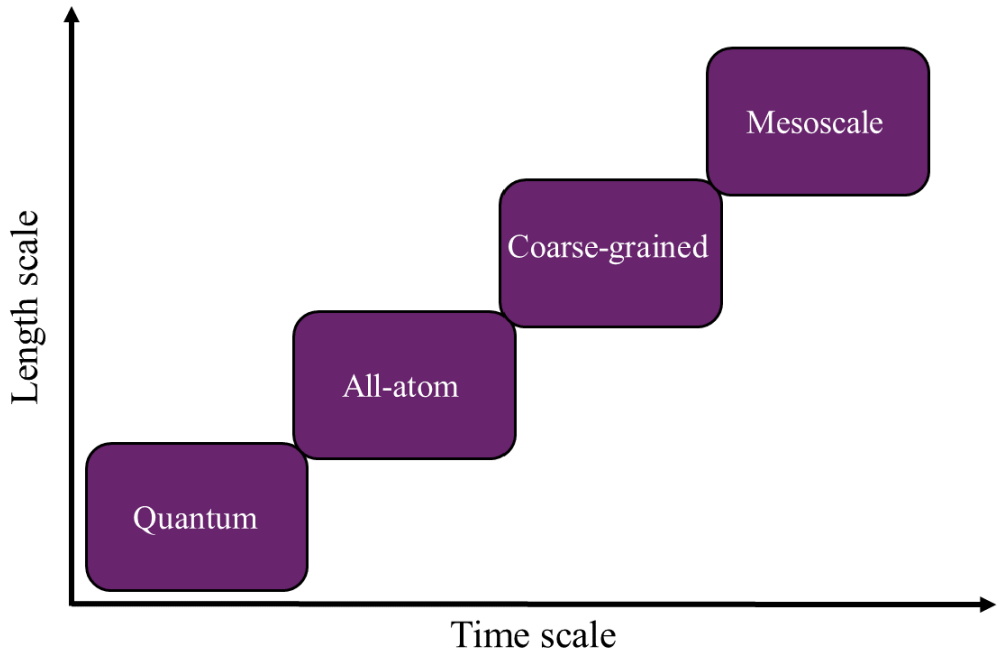


Figure 1.6: Different resolutions for molecular simulation, as a function of accessible length and time scale.

1.2.1 Atomistic simulation and molecular dynamics

Bill Bryson wrote "Protons give an atom its identity, electrons its personality." in his book '*A Short History of Nearly Everything*'. This sentiment forms the basis for quantum chemical methods which model nuclei and electrons independently and explicitly, where the aim is to attempt to solve the Schrödinger equation with the molecular Hamiltonian. The determination of the electronic structure of a system allows for various molecular properties to be calculated with high accuracy. However, this comes at a large computational cost and is limited to small systems and short timescales. Alternatively, many molecular systems can be modelled using a classical mechanics approach which avoids the complexity of computing electronic motions, and allows for significantly larger systems to be simulated.

In molecular mechanics at an all-atom resolution, the nucleus and electrons for an atom are unified into a single particle. Interactions between atoms, arising from electronic effects, are represented by potentials. Generally, an atom will have an associated partial charge, to represent its electron density, which is invariant in simulation. Instantaneous effects and interactions originating from shifts in electron density, termed van der Waals interactions, are typically modelled by a Lennard-Jones 12:6 potential. The molecular ge-

ometry (bond stretching, angle bending and torsional profiles) can be treated as springs with a harmonic oscillator. Overall, these components are combined into a set of functions/parameters to calculate the potential energy of a system and is colloquially known as a molecular *force field*. Further details can be found in section 2.1.

The parameters for the defined functions within a force field determine the behaviour of an atom with typical values required being partial charges and Lennard-Jones parameters. For groups of atoms, equilibrium bond lengths, angles and dihedrals and their associated force constants will need to be defined. Parameter sets can be obtained from experimental data or quantum chemical calculations. Equilibrium values for intramolecular terms can be extracted from X-ray diffraction or crystallography data, and their force constants from vibrational spectra. The parameters for intermolecular terms are more difficult to determine and can be gained through an iterative protocol that aims to reproduce the target data. Furthermore, force field parameters as a whole may need to be optimised to reproduce the target data. The purpose of the force field introduces several considerations into its parametrisation: i) is the goal maximum accuracy for a specific type of system or to have a wide coverage of chemical space?, ii) what are the target properties the force field aims to reproduce and how do they pertain to its intended use? and iii) how transferable is the force field across thermodynamic state points? To this effect, many force fields exist which differ in the set of data used for parametrisation, functional forms of its terms and compounds for which it specialises in. Examples of popular force fields are AMBER[76,77], CHARMM[78], GROMOS[79–81] and OPLS[82,83] whose parametrisations commonly target properties such as heats of vaporization and densities in the condensed phase. A modification can be made to all-atom force fields where hydrogen atoms on hydrocarbons, for example, are treated implicitly in an united-atom approach, increasing the computational efficiency of the force field.[80,84,85]

The treatment of electrostatics in force fields by fixed point charges has limitations in capturing the response of charge distribution to the environment. Polarisable force fields have been developed to address this issue.[86] The Drude oscillator model uses a massless virtual particle that carries a partial charge which is attached to an atom with a harmonic spring.[87] The change in position and orientation of this site in response to a change in electrostatic environment mimics an induced dipole. Polarisability can also be modelled by a multipole representation, as presented in the AMOEBA force field.[88] In this framework,

each atom has a permanent multipole moment consisting of a monopole/partial charge, a dipole vector and a quadrupole tensor. The interaction of permanent multipoles produces induced dipoles at each site, which can further induce dipoles themselves.

For a given configuration of a system, the potential energy can be calculated with its force field. The aim of simulation is to sample an ensemble of configurations which can be analysed. By concatenating the positions of all the sites in a system to a $3N$ dimensional vector $\boldsymbol{R}$, we can define a point in configuration space which describes the instantaneous configuration of the system, from which the potential energy $V(\boldsymbol{R})$ can be determined. Similarly, the collective momenta of each particle in the system can be represented by a point in momentum space. Combined together, this is referred to as phase space and simulation methods allow for the exploration of this. Molecular dynamics simulations involve the calculation of forces through $\boldsymbol{F} = -\nabla V(\boldsymbol{R})$. Newton's second law, $\boldsymbol{F} = m\boldsymbol{a}$, then gives the acceleration of a particle which is the second derivative of the position with respect to time, and velocity is the time derivative of position. Hence, solving this equation provides a method to advance the positions of atoms as a function of time (described in greater detail in section 2.2). Alternatively, Monte Carlo methods can be used to sample configuration space through the Metropolis–Hastings algorithm.[89,90] In this method, a trial move involving a change in conformation, orientation or position of the molecules is made. If the new configuration has a lower potential energy than the previous one, the move is accepted. If the potential energy of the new configuration is higher, the move is accepted or rejected by comparing a random number to the acceptance probability (according to its Boltzmann factor).

1.2.2 Coarse-grained modelling

A coarse-grained modelling approach, where groups of atoms are represented as single interaction sites, facilitates molecular dynamics investigations into soft matter systems well beyond the time and length scales available for all-atom simulations.[72] There are a number of ways by which coarse-grained models allow for a speed-up in simulation time or an increase in system size: i) the number of particles representing the system is reduced, ii) the potential energy surface of the system is smoother, iii) the interaction potentials are often simplified and iv) a larger time step can be employed. The first reason is straightforward in its effect; a reduced number of particles per molecule allows for more molecules to be

present in the system for the same total number of particles in an all-atom system. The second reason arises from the mapping of an all-atom model to a lower resolution coarse-grained model. The coarse-grained model will be defined by the slow, relevant degrees of freedom of the atomistic model whereas the fast degrees of freedom are removed or averaged into the coarse-grained model. The loss of degrees of freedom removes their explicit entropic contributions which are compensated for in the enthalpic contributions. Overall, this results in a flattening of the all-atom potential energy surface into a smoother landscape which improves the efficiency at which configuration space is sampled. The third reason arises from a neglect of contributions to the interaction potentials which often simplifies their form. This can simply be a neglect of electrostatic forces in the resulting model, where their contributions are usually implicit in the coarse-grained potentials. A more fundamental approximation to the interaction potentials is their implementation in coarse-grained simulations as effective pair potentials in order to limit computational expense. The effective nature of a potential refers to the amalgamation of multiple effects into a single potential which allows a given problem to be decomposed to fewer dimensions. The fourth reason is related to a combination of the removal of fast degrees of freedom and the use of more simplified potentials, which enables a larger time step to be used in molecular dynamics.

The partition function for an all-atom (AA) system can be expressed as

$$Q^{\mathrm{AA}} = c \int \mathrm{d}\boldsymbol{R} \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-U^{\mathrm{AA}}(\boldsymbol{R},\boldsymbol{r})/k_{\mathrm{B}}T}, \tag{1.2}$$

where c is a constant, U^{AA} is the all-atom interaction potential, k_{B} is the Boltzmann constant and T is the temperature.[72] $\boldsymbol{R}$ and $\boldsymbol{r}$ denote 'slow' and 'fast' degrees of freedom, respectively. A multibody potential of mean force can be obtained by integration and is defined by

$$U_{\mathrm{eff}}(\boldsymbol{R}) = -k_{\mathrm{B}}T \ln \left[c \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-U^{\mathrm{AA}}(\boldsymbol{R},\boldsymbol{r})/k_{\mathrm{B}}T} \right]. \tag{1.3}$$

This can be regarded as the free energy for the removed degrees of freedom of a coarse-grained (CG) configuration and it follows that

$$Q^{\mathrm{AA}} = Q^{\mathrm{CG}} = \int \mathrm{d}\boldsymbol{R}\, \mathrm{e}^{-U_{\mathrm{eff}}(\boldsymbol{R})/k_{\mathrm{B}}T}. \tag{1.4}$$

While this basic formulation is thermodynamically consistent, the multibody nature of this potential makes it difficult to evaluate and implement practically in computer simulations. Thus, most coarse-graining procedures will employ the use of pair potentials to simplify the nature of the interactions; however, this comes at the cost of thermodynamic consistency and the loss or averaging of multibody contributions to the potential energy of a coarse-grained configuration.

The aforementioned origins for the acceleration of molecular dynamics simulations in a coarse-grained framework may seem practical and beneficial, but they often come at a cost of accuracy. The development of coarse-grained potentials is a key challenge in computational chemistry and there is considerable interest in the systematic generation of potentials, given a defined mapping scheme, which are representative in terms of structure prediction and sufficiently transferable over a range of thermodynamic conditions. There are two classes of systematic coarse-graining procedures which differ in the type of data used in the parametrisation process. A *bottom-up* approach aims to construct potentials which reproduce microscopic properties obtained from an atomistic level simulation. On the other hand, a *top-down* approach attempts to parametrise potentials based on macroscopic quantities such as thermodynamic data from an experiment.

A bottom-up coarse-graining approach constructs lower resolution models on the basis of a more detailed (fine-grained) system. In principle, these methods derive coarse-grained potentials that reproduce target quantities, commonly obtained from an all-atom simulation, that describe the effective interactions between coarse-grained sites. The many-body potential of mean force (PMF) can be completely specified as[91]

$$W(\boldsymbol{R}) = -k_{\mathrm{B}}T \ln \left[\int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-U(\boldsymbol{r})/k_{\mathrm{B}}T} \delta(\boldsymbol{M}(\boldsymbol{r}) - \boldsymbol{R}) \right], \tag{1.5}$$

where $W(\boldsymbol{R})$ is the many-body PMF, $U(\boldsymbol{r})$ is the potential for the atomistic model and $\boldsymbol{M}(\boldsymbol{r})$ is a mapping operator. Thus, the many-body PMF is entirely defined by the atomistic model and the mapping operator. Here, the Dirac delta function $\delta(\boldsymbol{M}(\boldsymbol{r}) - \boldsymbol{R})$ selects all the atomistic configurations, $\boldsymbol{r}$, which map to a given coarse-grained configuration, $\boldsymbol{R}$. $W(\boldsymbol{R})$ is a measure of the weight of each coarse-grained configuration according to the Boltzmann weight of the atomistic configurations mapped to it. Therefore, this function can be used to evaluate a distribution of coarse-grained configurations that is equivalent to the atomistic model for a given a mapping scheme. The many-body PMF is considered a configuration-

dependent free energy, where its gradient corresponds to the conditional average of the atomistic forces and, as a function of $\boldsymbol{R}$, V and T, can be used to reconstruct structural and thermodynamic properties of the all-atom model.[91] However, the dependence of this many-body PMF on the coordinates of all coarse-grained configurations results in a problem too difficult to solve or too complex for use in simulation. It is the goal of bottom-up coarse-graining strategies to approximate $W(\boldsymbol{R})$ and determine potentials that are tractable for use yet accurate in recreating atomistic behaviour in further simulations.

As defined by Brini et al., bottom-up coarse-graining strategies can be grouped into two categories.[72] In *parametrised* methods, the aim is to construct potentials that will reproduce target properties or distributions calculated by atomistic simulation. These can be structure based with methods such as iterative Boltzmann inversion[92] (see section 2.5), inverse Monte Carlo (IMC)[93] and relative entropy (RE)[94,95], where these methods target reproduction of radial distribution functions through different implementations. For IMC, cross correlations within and between interactions are computed for the update of potentials, with the update itself being based in statistical mechanics. The RE method seeks to minimise the relative entropy, a quantity which describes the extent of phase space overlap between a coarse-grained and an atomistic system in terms of probability density distributions. Another approach is to match force distributions from the underlying atoms to the forces acting between their coarse-grained sites, known as force matching or multiscale coarse-graining (see section 2.6).[96–98] The second category of methods are termed *derived* methods, where the potentials are directly calculated from the atomistic model. This includes methods such as using the pair potential of mean force[99], effective force coarse-graining[100] and the conditional reversible work method[101].

A top-down coarse-grained model is, generally, developed from experimental thermodynamic data that is observed at or above the length scale of the coarse-grained model. The natural question that arises is the chemical specificity of top-down models, in disparity from bottom-up models for which chemical specificity is innate. In principle, a top-down model can reproduce a given experimental property without any relation to the fine-grained system, but this can be addressed by constructing models with the aim of capturing particular properties of the system of interest. The Martini force field (see section 2.7) is a popular top-down coarse-graining framework which provides a library of bead types to construct models with.[102–106] The interactions between these beads are

parametrised against partitioning behaviour between aqueous–hydrophobic environments with the goals of high transferability and wide coverage of chemical space. Using statistical associating fluid theory (SAFT), coarse-grained potentials can be obtained through an algebraic link between thermodynamic data and a molecular bead model.[107–109] In SAFT-γ Mie theory, parameters for Mie potentials for a tangentially-connected chain of beads are fitted to reproduce thermodynamic data over a series of experimental reference points.

Machine learning techniques have emerged as powerful tools to study large datasets, and have recently been exploited in analysing data from molecular simulation and in parametrising coarse-grained force fields.[110,111] One use of coarse-grained models is to generate equilibrated configurations of complex or large systems, and then backmap to a fine-grained atomistic representation for further analysis. The challenge is backmapping to an accurate physical structure, where a machine learning approach has been proposed.[112] Neural networks and unsupervised learning methods have been used as analysis tools to distinguish local structures in single-particle model colloidal or glass former systems.[113–115] From a bottom-up perspective, neural networks can be trained on all-atom simulations to learn the many-body PMF and extract the free energy surface at a coarse-grained resolution.[116] By formulating force matching as a supervised machine learning problem, in a deep learning[117] or graph neural network[118] approach, coarse-grained models have been developed which inherently capture multibody effects. The application of machine learning has also been used to parametrise coarse-grained water models using an extensive training set of properties derived from atomistic simulations and supplemented by experimental data.[119]

In coarse-grained modelling, there are two big issues that can arise: the representability and transferability problems. Representability is concerned with the ability of the coarse-grained model to reproduce and represent properties at the thermodynamic state point at which it is parametrised.[120] Transferability is concerned with the ability of the model to be predictive at different state points that it was not parametrised at.[72] The origins of these problems can be traced to the state point dependence of the effective pair potentials.[121] Coarse-graining changes the balance of enthalpic and entropic contributions to the free energy, which can lead to drastically different behaviour in the coarse-grained system. However, the correct phase behaviour may be observed if free energy changes are represented well. Another problem is the unrealistic pressures obtained in coarse-grained

simulations, where structure based methods can greatly overestimate the pressure of water at ambient conditions.[122] Moreover, a set of potentials that provide an exact reproduction of structure may not result in correct thermodynamic properties being reproduced, as a coarse-grained potential parametrised to one observable may not be able to reproduce other observables at the same thermodynamic state point.[123]

1.3 Scope of Thesis

The aim of this thesis is to investigate hierarchical self-assembly in soft matter systems by molecular dynamics simulations using models at various resolutions, with a focus on propagating information between different resolutions through multiscale modelling approaches. There are two classes of systems of interest which both self-assemble into ordered structures or phases, and constitute a comprehensive range of liquid crystals to study the effectiveness of coarse-graining here. This includes thermotropic liquid crystals of calamitic and discotic mesogens such as the cyanobiphenyl dimer CB7CB, which forms the twist-bend nematic phase, and a discotic perylene dye with oligoether substituents. Progress in the molecular simulation of liquid crystals has been reviewed.[124,125] The main focus in this work is on lyotropic systems of ionic dyes, containing an extended aromatic core, in aqueous solution, which aggregate into columns or complex stacked structures and can further organise into liquid crystal phases. This classification encompasses compounds such as chromonic liquid crystals which have been extensively studied using all-atom molecular dynamics simulations.[47–50,126–135] Using atomistic models, insights into the thermodynamics of self-assembly, local molecular ordering and interplay between structures have been gained. Since the study of mesophases and complex large-scale aggregates of these compounds requires simulation on time and length scales inaccessible to all-atom models, coarse-grained simulation approaches are indispensable in facilitating further investigation. However, the coarse-graining of chromonic systems is highly challenging as a subtle balance between hydrophilic and hydrophobic interactions within the molecule must be achieved. Numerous studies on the coarse-graining of chromonics were focused on the non-ionic mesogen TP6EO2M.[136–139] This thesis continues the pursuit in developing coarse-grained methodologies to study hierarchical self-assembly across a range of liquid crystals through the application, assessment and improvement of

systematic approaches to induce a paradigm shift in coarse-graining soft matter systems.

The organisation of this thesis is as follows. This first and current chapter presents a comprehensive background into liquid crystals and the unusual class of lyotropic systems known as chromonic liquid crystals. The field of molecular simulation is then generally introduced, covering the various levels of computational chemistry.

Chapter 2 provides a more in-depth background into the theory and methods of computer simulation used in this work. The theoretical details, accompanied with a review of the literature, for the coarse-graining methods utilised are presented.

Chapter 3 presents coarse-graining results on a thermotropic nematic liquid crystal based on the cyanobiphenyl moiety, CB7CB. The aim of this work is to study the effect of coarse-graining on this nematogen and venture toward simulating the twist-bend nematic phase. Molecular liquids for fragments of this mesogen, benzonitrile and heptane, and their miscible mixture are investigated. This serves as a prelude to further coarse-graining of more complex compounds in aqueous solution.

Chapter 4 focuses on all-atom molecular dynamics simulations of a series of four ionic cyanine dyes: PIC, PCYN, TTBC and BIC. The atomistic models were used to study the thermodynamics of self-assembly of the dyes and demonstrate their association into two classes of structures, where these observations allow for increased insight into these systems and validation of subsequent coarse-grained models.

Chapter 5 looks into the performance of different coarse-grained mappings and methods on models for cyanine dyes. Firstly, the aim is to assess and compare unmodified approaches for a single system. Secondly, approaches to improve the representability of coarse-grained models for cyanine dyes are tested.

Chapter 6 details all-atom simulations and the development of coarse-grained models for a chromonic perylene dye. While thematically similar to the cyanine dyes, this mesogen features a number of structural and behavioural differences that should render it more conducive to coarse-graining. Furthermore, a thermotropic analogue is studied under the hypothesis that the absence of an aqueous nature will diminish the difficulty of coarse-graining compared to a chromonic mesogen.

Finally, chapter 7 will summarise the research presented in this thesis.

Chapter 2

Computer Simulation and Coarse-Graining Methods

2.1 Force fields

In classical molecular simulations, a force field is used to determine the potential energy of the system and is comprised of a set of potential energy functions which represent the various contributions to the total energy. These can be categorised into two sets of contributions given by

$$U_{\mathrm{tot}} = U_{\mathrm{bonded}} + U_{\mathrm{nonbonded}}, \tag{2.1}$$

where each term is composed of a number of subterms which describe more detailed components of the system. The general formulation used is

$$U_{\mathrm{bonded}} = \sum U_{\mathrm{bond}} + \sum U_{\mathrm{angle}} + \sum U_{\mathrm{dihedral}} \tag{2.2}$$

$$U_{\mathrm{nonbonded}} = \sum U_{\mathrm{electrostatic}} + \sum U_{\mathrm{van\ der\ Waals}}, \tag{2.3}$$

where each component is a summation of the interactions of that class present in the system and can assume a variety of functional forms.

Bonded interactions

The intramolecular bonded terms of the force field determine the molecular geometry and structure of the molecules. The first term represents bond stretching and is defined between pairs of connected sites. This is generally modelled using harmonic potentials of the form

$$U_{\text{bond}} = \frac{1}{2}k_r(r - r_0)^2, \tag{2.4}$$

where k_r is the force constant, r is the bond length and r_0 is the equilibrium bond length. The force constant of the bond is proportional to the vibrational frequency of the motion. According to Nyquist's theorem[140], a caveat is introduced such that the time step for simulation must be less than half the period of the fastest vibration. The use of a harmonic potential to approximate bond stretching is reasonably accurate close to the equilibrium bond length, but may require the use of an asymmetric Morse potential[141] to better represent bond stretching for a wider range of distances.

The angle bending term is defined between three connected sites and is commonly described by a harmonic potential

$$U_{\text{angle}} = \frac{1}{2}k_\theta(\theta - \theta_0)^2, \tag{2.5}$$

where k_θ is the force constant, θ is the angle and θ_0 is the equilibrium angle. Compared to bond stretching, a deviation from the equilibrium value causes an increase in energy but on a much smaller scale.

For a group of four sites i, j, k and l, a dihedral angle can be defined between the planes formed by sites ijk and jkl. This may be simply represented by the potential

$$U_{\text{dihedral}} = k_\phi(1 + \cos(n\phi - \phi_0)), \tag{2.6}$$

where k_ϕ is the force constant, n is the periodicity, ϕ is the dihedral angle and ϕ_0 is the equilibrium value. If the sites (i, j, k and l) within a dihedral are connected consecutively, this is known as a proper dihedral whereas other cases are known as improper. Generally, proper dihedrals determine the flexibility around its central bond and improper dihedrals may be used to ensure planarity within a molecule. Another commonly used form for dihedrals is the Ryckaert–Bellemans (RB) function[142,143]

$$U_{\text{dihedral}} = \sum_{n=0}^{5} C_n(\cos(\psi))^n, \tag{2.7}$$

where C_n are coefficients and $\psi = \phi - 180°$.

Constraints

The explicit modelling of bond vibrations in MD simulations limits the size of the time step and increases the computational cost. Constraints can be used to replace bonds and allow for a larger time step to be used. One of the most efficient constraint algorithms is the Linear Constraint Solver (LINCS) method.[144] This algorithm uses two steps. Firstly, an updated bond is projected onto the old bond and then, secondly, a correction is applied to satisfy the specified length.

Non-bonded interactions

The non-bonded terms in the force field describe intermolecular interactions in the system and are the most computationally intensive terms. These can also apply within a molecule if sites are separated by more than three bonds, generally. In most classical force fields, a given atomic site will have an associated partial charge to represent its electron density. Within a defined cutoff, the energy arising from electrostatic interactions can be calculated from a Coulomb potential

$$U_{\text{electrostatic}} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}, \tag{2.8}$$

where r_{ij} is the intersite distance, ϵ_0 is the vacuum permittivity, ϵ_r is a dielectric constant and q_i and q_j are point charges on sites i and j, respectively. Long-range electrostatics may be handled by the reaction-field method[145] which assumes a constant dielectric environment, ϵ_{rf}, beyond the cutoff. A more complex and commonly used approach is the particle-mesh Ewald (PME) method.[146] Short-range electrostatics are calculated by the Coulomb potential in direct space, but the long-range contributions are summed in reciprocal space using a Fourier transformed grid of the charge distribution.

The interactions between sites not attributed to electrostatics are termed van der Waals interactions and encompass the dispersion, repulsion and induction effects. This is typically represented by the r^{-6} term in the Lennard-Jones (LJ) potential

$$U_{\text{van der Waals}} = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], \tag{2.9}$$

where ϵ is the well depth and σ is the distance at which $U_{\text{LJ}} = 0$. The distance at which $U_{\text{LJ}} = -\epsilon$ is given by $r_{\text{min}} = 2^{1/6}\sigma$ and is the distance where the strongest interaction between two particles occurs. At distances shorter than this, there is a repulsive slope (controlled by the r^{-12} term) which prevents overlap of particles and at distances larger than r_{min}, there is an attractive region where the energy decays to zero. Other analytical forms that are used are the generalised version of the Lennard-Jones potential, referred to as Mie potentials (where the exponent terms are not fixed to $n = 12$ and $m = 6$), or Morse potentials, where both these functions allow for greater freedom in controlling features of the potential. For coarse-grained force fields, it is often beneficial to adopt numerical potentials, which are tabulated, and do not assume a fixed functional form. In practice, interactions between particles at large distances is negligible and a cutoff is employed, such that the potential is shifted by a constant so that $U_{\text{vdW}} = 0$ at r_{cut} and interactions are not calculated beyond this.

2.2 Molecular dynamics

The potential energy functions defined in the force field can be used to calculate the force, $\boldsymbol{F}_i$, on any atom i in the system. Forces are the negative derivatives of a potential energy function, $V(\boldsymbol{r}_i)$, and are given specifically in this case as

$$\boldsymbol{F}_i = -\frac{\partial V(\boldsymbol{r}_i)}{\partial \boldsymbol{r}_i}. \tag{2.10}$$

Molecular dynamics (MD) simulations solve Newton's equations of motion for a system of *N* interacting atoms where the equations are given by

$$\boldsymbol{F}_i = m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2};\ \boldsymbol{v}_i = \frac{\partial \boldsymbol{r}_i}{\partial t};\ \boldsymbol{F}_i = m_i \frac{\partial \boldsymbol{v}_i}{\partial t} \quad \text{, for } i = 1 \text{ ... } N. \tag{2.11}$$

These equations can be solved in appropriate time steps to move atoms in a simulation using a leap-frog algorithm for integration.[147] The algorithm requires positions $\boldsymbol{r}_i$ at time *t* and velocities $\boldsymbol{v}_i$ at time $t - \frac{1}{2}\Delta t$. Positions and velocities in the system are updated using the forces $\boldsymbol{F}_i(t)$ determined by positions at time *t* through the relations

$$\boldsymbol{v}_i(t + \frac{1}{2}\Delta t) = \boldsymbol{v}_i(t - \frac{1}{2}\Delta t) + \frac{\Delta t}{m_i}\boldsymbol{F}_i(t) \tag{2.12}$$

$$\boldsymbol{r}_i(t + \Delta t) = \boldsymbol{r}_i(t) + \Delta t \boldsymbol{v}_i(t + \frac{1}{2}\Delta t). \tag{2.13}$$

A suitable time step, Δt, must be carefully selected so that the fastest motions in the system are captured yet still allowing the simulation to run efficiently. Other integration methods, such as the velocity Verlet algorithm, are available.[148]

Stochastic dynamics

Stochastic dynamics (SD) add a noise and friction term into Newton's equations of motion. This integrator can be used for single molecule simulations in a vacuum to imitate the presence of a solvent (via the friction term) so that a representative distribution of conformations, corresponding to the sampling of a condensed phase with molecular dynamics, can be obtained. The resultant equation is given by

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} = \boldsymbol{F}_i - m_i \gamma_i \frac{\partial \boldsymbol{r}_i}{\partial t} + \boldsymbol{R}_i. \tag{2.14}$$

The first term is the total force without modification, the second term is a frictional force (where γ_i is a friction constant) and the third term is a random force representing a noise process which adds randomness to motion.

Periodic boundary conditions

Finite systems are subject to boundary effects when simulating a bulk liquid. Instead of simulating a very large system, periodic boundary conditions can be applied.[149] This involves a small unit cell which is replicated in three dimensions which effectively removes any boundaries. The minimum image convention is also used, where only the nearest image of a particle is used for interactions.

Thermodynamic ensembles, thermostats and barostats

A particular set of conditions may be held constant in simulation, where these conditions within statistical thermodynamics are known as the *ensemble*. There are two commonly

used ensembles for MD simulation: the constant *NVT* (canonical) ensemble and the constant *NPT* ensemble, where *N* is the number of particles, *V* is the volume, *T* is the temperature and *P* is the pressure. In the *NVT* ensemble, the dimensions of the box are fixed to keep the volume constant, whereas in the *NPT* ensemble, the box size is allowed to vary in order to maintain a constant average pressure.

The temperature in a simulation is proportional to the squared velocities of the particles in the system. A simple way to control temperature, therefore, is to alter the velocities. The Berendsen thermostat controls the temperature of the system by weakly coupling it to an external heat bath to scale the velocities by a factor.[150] For a given reference temperature T_0,

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}, \tag{2.15}$$

where τ is a time constant. The weakness of this approach is that since the same scaling factor is applied to all velocities, a proper canonical ensemble is not generated and fluctuations of the kinetic energy are suppressed. To sample within a correct Boltzmann distribution of velocities, the velocity rescaling thermostat can be used, where the Berendsen thermostat is modified with an additional stochastic term.[151] Alternatively, the Nosé–Hoover thermostat can be employed.[152,153] This adds a friction term and a thermal reservoir into the equations of motion so that the contributions to the energy and coupling to the heat bath are solved deterministically.

The pressure in a MD simulation can be calculated according to the virial pressure equation

$$P = \frac{N k_\mathrm{B} T}{V} + \frac{1}{3V} \sum_{i}^{N} \boldsymbol{r}_i \cdot \boldsymbol{F}_i, \tag{2.16}$$

where pressure is inversely proportional to the volume of the box. In a similar fashion to thermostats, scaling the box size allows for control of pressure. The Berendsen barostat scales the box dimensions and coordinates, which acts as a first-order kinetic relaxation towards the reference pressure.[150] For production runs, a Parrinello–Rahman barostat is typically used, which allows for pressure fluctuations to be captured better and for independent scaling of box vectors.[154]

2.3 Analysis techniques

Free energy of association

The free energy of association for a system can be evaluated from a potential of mean force (PMF) along a reaction coordinate. If the reaction coordinate is a separation distance, then the PMF describes the work done to pull two species apart. This PMF can be obtained by a series of simulations where the molecules are constrained at specified points over a separation distance. The PMF, U_{PMF}, is then calculated by integrating the average constraint force, $\langle f_{\mathrm{c}} \rangle_s$, over the separation distance, s, according to the equation

$$U_{\mathrm{PMF}}(r) = \int_{r}^{r_{\max}} \left[\langle f_{\mathrm{c}} \rangle_s + \frac{2k_{\mathrm{B}}T}{s} \right] \mathrm{d}s, \tag{2.17}$$

where r is the distance, $r_{\max}$ is the maximum distance and $2k_{\mathrm{B}}T/s$ is a kinetic entropy term which accounts for the increase in rotational volume at larger separation distances.[155–157]

Free energy of hydration

The Bennett acceptance ratio (BAR) method can be used to calculate the free energy difference (ΔF_{BA}) between two states, A and B.[158] Multiple intermediate states are defined by a coupling parameter, λ, and a Hamiltonian, H, is calculated for each state. The free energy difference between states i and j can then be evaluated according to

$$\Delta F_{ji} = k_{\mathrm{B}}T \ln \frac{\langle f(H_i - H_j + C) \rangle_j}{\langle f(H_j - H_i + C) \rangle_i} + C, \tag{2.18}$$

where

$$f(x) = \frac{1}{1 + \exp(x/k_{\mathrm{B}}T)}. \tag{2.19}$$

The value of the constant, C, is numerically determined to fulfil $\langle f(H_i - H_j + C) \rangle_j = \langle f(H_j - H_i + C) \rangle_i$. The total free energy difference is calculated as the sum of all intermediate free energy differences

$$\Delta F_{\mathrm{BA}} = \sum_{i=1}^{n\text{-}1} \Delta F_{i+1,i}. \tag{2.20}$$

When decoupling interactions, particles may get very close to each other (near the end

points when the interactions are weak) and cause large fluctuations in the potential energy. This problem can be circumvented by using soft-core potentials given by

$$U_{\mathrm{SC}}(r) = (1-\lambda)U_{\mathrm{A}}(r_{\mathrm{A}}) + \lambda U_{\mathrm{B}}(r_{\mathrm{B}}) \tag{2.21}$$

$$r_{\mathrm{A}} = (\alpha\sigma_{\mathrm{A}}^{6}\lambda^{p} + r^{6})^{\frac{1}{6}} \tag{2.22}$$

$$r_{\mathrm{B}} = (\alpha\sigma_{\mathrm{B}}^{6}(1-\lambda)^{p} + r^{6})^{\frac{1}{6}}, \tag{2.23}$$

where α is the soft-core parameter (0.5), σ is the radius of the interaction and p is a positive integer (1).

Orientational order parameter

For a liquid crystalline system, the orientational order parameter, S_2, is a measure of the orientational order present in a system and is given by

$$S_2 = \left\langle \frac{3\cos^2\theta_i - 1}{2} \right\rangle, \tag{2.24}$$

where θ_i is the angle between a defined molecular vector, $\hat{\boldsymbol{u}}_i$ for molecule i, and the director of the system, $\hat{\boldsymbol{n}}$, and the angular brackets denote an ensemble average. This can be used to discern the identity of a liquid crystal mesophase, where a value of $S_2 \approx 0$ indicates an isotropic phase and $S_2 = 1$ indicates a completely aligned system. The unit vector, $\hat{\boldsymbol{u}}_i$, of the molecular long axis for a molecule i can be defined by selecting a pair of sites or, more rigorously, as the eigenvector corresponding to the lowest eigenvalue found by diagonalising the moment of inertia tensor for a molecule

$$I_{\alpha\beta} = \sum_{i} m_i(\boldsymbol{s}_i^2\delta_{\alpha\beta} - s_{i\alpha}s_{i\beta}), \tag{2.25}$$

where m_i is the mass of atom i, $\boldsymbol{s}_i$ is the vector between an atom i and the molecular centre of mass (COM) and α, β is the cartesian axis. For discotic or chromonic mesogens, this vector can be defined as the vector perpendicular to the plane of the core calculated from three sites (j, k and l) by

$$\boldsymbol{u}_i = \boldsymbol{u}_{jk} \times \boldsymbol{u}_{kl}. \tag{2.26}$$

In practice, the orientational order parameter can be determined from diagonalisation of the order tensor

$$Q_{\alpha\beta} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}(3u_{i\alpha}u_{i\beta} - \delta_{\alpha\beta}), \tag{2.27}$$

where N is the number of molecules and $\delta_{\alpha\beta}$ is the Kronecker delta. The largest eigenvalue provides the value of S_2 and its associated eigenvector is the director, $\hat{\boldsymbol{n}}$, of the system.

Stacking distances and twist angles

In a stacked structure, the spacing between the molecular planes of adjacent molecules can be determined in addition to the COM distances. The stacking distance, d, is defined as the projection of the COM distance between two molecules along the average vector normal to their cores according to

$$d = \boldsymbol{d}_i \cdot \boldsymbol{r}_{ij}, \tag{2.28}$$

where $\boldsymbol{d}_i$ is the vector defining the normal to the molecular plane and $\boldsymbol{r}_{ij}$ is the vector between the COMs of molecules i and j. The twist angle between adjacent molecules in a stack can be defined by

$$\theta = \cos^{-1}\left(\hat{\boldsymbol{v}}_i \,.\, \hat{\boldsymbol{v}}_j\right), \tag{2.29}$$

where $\hat{\boldsymbol{v}}_n$ is the unit vector obtained from $\boldsymbol{v}_n$ as defined by

$$\boldsymbol{v}_n = \boldsymbol{L}_n \times \boldsymbol{d}_n. \tag{2.30}$$

θ is the twist angle and $\boldsymbol{L}_n$ is the direction vector for molecule n.

2.4 Coarse-grained mapping

The first, and perhaps the most important, choice in coarse-graining is the mapping scheme for the model. A coarse-grained model can only reproduce observables that are compat-

ible with the resolution of the model and do not crucially depend on removed degrees of freedom from the fine-grained system, noting that expressions for calculating observables may also differ in the coarse-grained system due to a neglect in contributions from eliminated variables.[120] Therefore, a mapping scheme must be carefully selected so that there is a speed up in simulation but adequate chemical information is retained to exhibit realistic behaviour (such as the flexibility of alkyl chains). It is typical for the mapping of a coarse-grained model to be designed by hand where the user's chemical intuition and knowledge of the molecule allows for selection of what chemical information is contained within a coarse-grained site. However, automated procedures for mapping and parametrisation with the Martini force field have been successful in the high-throughput screening of molecules to calculate partition coefficients.[159,160] In this thesis, the mapping schemes for all the coarse-grained models presented were determined by chemical intuition. The interaction site for a coarse-grained bead can be defined by the centre of mass of its constituent atoms as follows

$$\boldsymbol{R}_I = \boldsymbol{M}_I(\boldsymbol{r}_I) = \sum_i^n \frac{\boldsymbol{r}_i m_i}{\sum_i^n m_i}, \tag{2.31}$$

where $\boldsymbol{R}_I$ is the position of coarse-grained bead I, $\boldsymbol{r}_i$ and m_i is the position and mass of an atom i, respectively, and n is the number of atoms contained in the coarse-grained bead. $\boldsymbol{M}_I(\boldsymbol{r}_I)$ is a mapping operator which maps a coarse-grained configuration as a function of an underlying configuration $\boldsymbol{r}_I$. Most coarse-grained studies will utilise a mapping of 2–4 heavy atoms per coarse-grained bead, although this can be greater if appropriate for a relevant observable or varied to allow for symmetry in the resulting model. For a given number of site types in a coarse-grained model, the total number of pairwise interactions, N, is

$$N = \frac{n(n+1)}{2}, \tag{2.32}$$

where n is the total number of coarse-grained bead types present in the system.

2.5 Iterative Boltzmann inversion

The structure of a system may be described by a radial distribution function (RDF) which represents the probability of finding a given particle type located at a distance from a reference particle, relative to an ideal gas. The expression for a RDF is given by

$$g(r) = \frac{1}{4\pi r^2 N\rho} \sum_{i}^{N} \sum_{j \neq i}^{N} \langle \delta(r - r_{ij}) \rangle, \tag{2.33}$$

where N is the number of particles, ρ is the number density and $\langle \delta(r - r_{ij}) \rangle$ is the average number of particles in a shell of width δr at a distance r. This is also given by $g(r) = \langle \rho(r) \rangle / \rho$, where this represents the ratio of the local density for particles at a given distance compared to the bulk density. The peaks present in the RDF provide details into the structure of the system, and the shape of the function also allows for the state (solid/crystal, liquid or gas) of the system to be distinguished.

For a pair of site types, we can define a potential of mean force from its RDF through Boltzmann inversion

$$U(r) = -k_{\mathrm{B}} T \ln g(r). \tag{2.34}$$

From this relation, we can introduce an iterative procedure to parametrise coarse-grained potentials which reproduce the pairwise structure in a reference system as described by its RDFs. The iterative Boltzmann inversion (IBI) method[92] refines an initial potential by applying a change proportional to the difference between the coarse-grained and reference RDFs according to

$$U_{n+1}(r) = U_n(r) + \alpha k_{\mathrm{B}} T \ln \frac{g_n(r)}{g_{\mathrm{target}}(r)}, \tag{2.35}$$

where $U(r)$ here is the interaction potential, n is the step number and α is a scaling factor. At each step, a short simulation is run with the current set of potentials to evaluate the RDFs before the update to the potentials. This method is subject to the Henderson uniqueness theorem which states that a given RDF only has one pair potential that will exactly reproduce it.[161] It is noted that, in practice, a number of potentials can reproduce a RDF within an acceptable degree of error.

As the pressure in bottom-up coarse-grained models is often overestimated, a linear

pressure correction[92,162] can be incorporated into the IBI procedure as follows

$$\Delta U(r) = A\left(1 - \frac{r}{r_{\mathrm{cut}}}\right), \tag{2.36}$$

where

$$A = \mathrm{sgn}(\Delta P) 0.1 k_{\mathrm{B}} T \min(1, |f\Delta P|)\,. \tag{2.37}$$

In these equations, r_{cut} is the cutoff distance, ΔP is the difference in pressure between the reference and coarse-grained system and f is a scaling factor.

The IBI method is popular due to its straightforwardness and general applicability to construct bonded and non-bonded potentials for a system. While this method does provide good structural representability, it usually struggles in producing potentials which are transferable due to the heavy state point dependence of structural features in many systems. Nonetheless, the IBI method has been applied to a liquid crystal, 8AB8, where the resulting coarse-grained models captured the phase behaviour with varying temperature and density.[163,164] IBI has also been used to develop coarse-grained models for a range of polymer systems[165–167], and these are relatively transferable between temperatures and molecular weights for homopolymer systems.[168] However, systems with many types of interaction sites or where chemical environment changes with temperature/concentration become problematic for transferability. A multistate extension (MS-IBI) has been proposed which produces more transferable potentials through inclusion of target data from multiple state points.[169] The transferability of several coarse-graining methods has been assessed for an octane-benzene system, where the application of the MS-IBI scheme clearly improves transferability of the model (compared to IBI) but comes at the expense of representability.[170] The representability of models constructed through IBI can also be improved by targeting the reproduction of Kirkwood–Buff integrals, which are calculated from RDFs by

$$G = 4\pi \int_0^\infty [g(r) - 1]\, r^2 dr. \tag{2.38}$$

This approach has been shown to better capture the solution behaviour of benzene-water and urea-water systems, and can be implemented in an iterative scheme as direct targets or as a correction to the potentials.[123,171]

2.6 Multiscale coarse-graining / Force matching

Instead of matching structural properties for constructing coarse-grained potentials, a force based approach can be used. The idea behind force-based methods is to match force distributions obtained from the atomistic reference to forces acting between coarse-grained sites. The force matching method was first reported to calculate classical potentials from first principles methods.[96] This was extended into the framework of coarse-graining in the multiscale coarse-graining (MS-CG) method.[97,98] The MS-CG method utilises effective pair potentials to approximate the many-body PMF for the system. The theoretical basis from statistical mechanics has, subsequently, been derived and reported.[172,173] Similar to IBI, a function can be evaluated which is based on the difference between the PMFs of the coarse-grained and all-atom systems. Minimisation of this function allows the true effective many-body PMF to be approached. The variational principle derived is given by

$$\chi^2 = \frac{1}{3LN} \sum_{l=1}^{L} \sum_{i=1}^{N} \left| \boldsymbol{F}_{il}^{\mathrm{AA}} - \boldsymbol{F}_{il}^{\mathrm{CG}}(x_1, ..., x_M) \right|^2 , \tag{2.39}$$

where χ^2 is the objective function be minimised, L and N are the total number of snapshots and particles, respectively, $\boldsymbol{F}_{il}^{\mathrm{CG}}$ is the force on bead i in snapshot l for the coarse-grained system, and $x_1, ..., x_M$ are coefficients for fitting. $\boldsymbol{F}_{il}^{\mathrm{AA}}$ is the total force obtained by summing the forces from the underlying atoms which map to coarse-grained bead i in snapshot l.

The MS-CG method has been used to develop a coarse-grained model of an ionic liquid, for which resulting coarse-grained simulations show satisfactory structural and thermodynamic accuracy.[174] A hybrid force matching procedure was adopted in this study, where different coarse-graining methods can be applied to the same system.[175,176] In particular, the intramolecular bonded interactions are obtained via Boltzmann inversion whereas the intermolecular non-bonded interactions are developed using the MS-CG method. This can have the advantages that the bonded potentials are easily obtained and it removes any issues from a lack of sampling in the reference. Systems for which MS-CG models have been developed range from molecular liquids[98,162,170] to peptides[177,178] and lipid bilayers[97,179]. It has been shown that MS-CG can be used as an analysis tool by approximating the many-body PMF of the atomistic system with respect to coarse-grained coordinates.[180]

A number of improvements have been implemented for the MS-CG method over the

years. By adopting a mapping operator based on the centre of charge of the underlying atoms in a CG site, it is found that structural correlations for polar compounds are better captured.[181] The use of improved basis functions can reduce problems associated with statistical sampling errors and noise in the atomistic reference.[182,183] In order to address inaccuracy in the pressure of systems for MS-CG models, a volume-dependent term can be introduced in the potential energy function which allows for correct volume fluctuations in the *NPT* ensemble to be generated.[184] This was further implemented in a self-consistent pressure matching algorithm to determine volume-dependent potentials, where an extended ensemble approach improves upon the transferability and accuracy of coarse-grained potentials for heptane-toluene mixtures.[185,186] A pressure correction can also be applied by incorporating a virial constraint to the MS-CG procedure, and improved density transferability can be achieved by including a density-dependent term.[187] The addition of three-body contributions into the force field as explicit potentials can increase its accuracy compared to the two-body approximation alone, especially where changes in chemical environment and concentration affect the behaviour.[188,189] Multi-body effects can also be incorporated by the use of local density potentials, which can be adapted from volume potentials[190] or parametrised concurrently with the pair potentials[191].

2.7 Martini

The most popular and widely used coarse-grained modelling method is based on the Martini force field, a top-down model where building blocks are parametrised to reproduce partitioning behaviour between aqueous–hydrophobic environments. It was originally produced for biomolecular simulation, but has been extended to a range of other systems.[102–105] The Martini force field consists of a library of coarse-grained beads whose potentials have been extensively calibrated against thermodynamic data (to describe hydrophobic, van der Waals and electrostatic interactions) with the aim of high transferability. Thus, the resulting beads can be used in a broad range of systems at various thermodynamic state points without the need to reparametrise each time. The thermodynamic properties used in the parametrisation of the Martini 2 force field were the free energies of hydration, enthalpies of vaporization and partition coefficients between water and an assortment of organic solvents (hexadecane, choroform, ether and octanol). While intermolecu-

lar interactions are largely determined by reference to experimental thermodynamic data, the intramolecular interactions can be obtained with a bottom-up approach as mentioned before. This incorporation of bottom-up aspects into this top-down model improves the chemical specificity and accuracy compared to an exclusively top-down approach. Within the framework of Martini 2, the models are based on a four-to-one mapping (except for ring structures which require a finer mapping) and there are four main bead types: polar (P), nonpolar (N), apolar (C) and charged (Q). Each bead type also has a number of subtypes describing the polarity (from 1–5 with increasing polarity) or hydrogen bonding capabilities. These particles interact via Lennard-Jones 12:6 potentials. The Martini force field has been applied to vast range of systems such as polymers[192,193], surfactants[194–196], nanoparticles[197,198], lipids[199,200] and protein–membrane interplay[201–203].

One of the biggest problems associated with Martini 2 is the instability of P_4 particles, representing a mapping of four water molecules to a bead, compared to real water and its tendency to freeze, even at 300 K and particularly when in contact with interfaces. The standard solution to this problem is to introduce antifreeze particles, which are denoted BP_4, into the system. BP_4 beads interact with solute particles in the same way as P_4, but have a larger effective radius when interacting with normal water sites so as to disrupt the uniform lattice packing. The steep repulsive profile of the 12:6 potential seems to be the source of the freezing issue, which leads to another solution of using a Morse potential for the water model.[204] This has been exploited in a study on polyethylene glycol, in conjunction with potentials derived using IBI, in bulk solution and at a water/air interface which show that the combination produces good agreement on the polymer properties and is transferable between the two environments.[205] There is a neglect in electrostatic properties in Martini, where water as a solvent has no dipole moment or dielectric constant, which results in a lack of explicit screening and instantaneous responses to charges, and requires the use of a higher background dielectric constant. The desired effects can be incorporated into Martini through the use of a polarisable water model which is modelled in a three-site representation and allows for a lower dielectric constant to be used; in this model, a central bead holds the LJ interaction and two oppositely charged sites constrained to the centre can freely rotate to mimic a responsive dipole moment.[206] Parameters for this polarisable Martini approach have since been refined.[207,208]

A pitfall of the Martini framework is apparent by the presence of artificial energy bar-

riers in dimerisation profiles when beads with a mismatch in size are used without size-dependent cross interactions.[209] Recently, the Martini 3 force field was formally published with new bead types and improved interactions presented.[106] This formulation now includes different bead sizes denoted S (small) and T (tiny) for three- and two-to-one mappings, respectively. Interactions within a defined block (organic, ion and solvent) have been refined and expanded in terms of the number of interaction levels available to increase coverage of chemical space. Each block also features specific cross interactions between bead types and sizes. The need for antifreeze particles is discontinued with the introduction of a new bead type for water.

2.8 Application of methods

Thus far, this chapter detailed various methods for computer simulation and coarse-graining, and this section will discuss the application of these methods in this thesis. Throughout this work, molecular dynamics simulations using classical force fields were utilised to study a range of systems. All coarse-grained models in this work (in chapters 3, 5 and 6) were mapped from an all-atom representation according to Equation 2.31 with various numbers of underlying atoms contained within the beads. Coarse-grained potentials, in the same chapters, were systematically developed using IBI, MS-CG/FM and Martini methods, informed by fine-grained atomistic simulations as references or as validation targets. For liquid crystalline systems, analysis of the order parameter, as a function of time or as an average, is indispensable in determining the extent of order present (in chapters 3 and 6). Similarly, the stacked structures observed in all-atom simulations in chapter 4 can be analysed in terms of stacking distances and twist angles. RDFs are employed as part of IBI in the relevant chapters, but also to assess structural accuracy of the models in chapter 3. The ubiquitous quantities used to evaluate models in this work are thermodynamic properties: the free energy of association (in chapters 4, 5 and 6) and the free energy of solvation/hydration (in chapters 3, 4, 5 and 6). Further methods and details for the simulation and development of models will be presented in the relevant chapters.

Chapter 3

Systematic Coarse-graining of Soft Matter

3.1 Introduction

Systematic coarse-graining methods can be used to develop chemically specific CG models of molecular liquids, which facilitate studies at length and time scales inaccessible to all-atom simulations for the investigation of self-assembly in soft matter. The aim of systematic methods is to automate the parametrisation of CG potentials for a system, given a mapping scheme, based on an underlying all-atom reference and/or experimental data. The key challenge is to generate a set of potentials for a CG model that are able to reproduce structural and thermodynamic properties at the state point at which it is parametrised, and to be predictive at different state points that it was not parametrised at. These are known as the representability[120] and transferability[72] problems, respectively.

Numerous studies on the development, assessment and comparison of coarse-graining methods applied to molecular liquids have been conducted.[72] Due to its ubiquity as a solvent, water has been extensively coarse-grained with various mappings and parametrisation strategies.[210,211] Generally, the systems selected for comparing coarse-graining methods are homogeneous single-component liquids.[162,175,212] These studies typically map molecules to single sites, and rely on the simplicity of the systems to investigate the effectiveness of the methods. However, the use of simple test systems to assess coarse-graining methods does not allow for the evaluation of their ability to handle more complex scenarios. Systems for which coarse-graining would be beneficial often have internal molecular

structure and are inhomogeneous in nature, thus requiring multiple diverse interactions to represent the system. Examples of comparative studies which feature more complex systems have utilised benzene-octane mixtures[170], methanol-water mixtures[212], urea-water mixtures[123,171] and a conjugated polymer system[213].

Coarse-graining approaches have been applied to liquid crystals with the intention of simulating systems beyond the reach of atomistic simulations.[124] The simplest models are idealised systems based on rigid hard particles, with molecular shapes such as ellipsoids[214] or spherocylinders[215]. Subsequent models of chains of hard spheres containing rigid and flexible segments show the N, SmA and SmC phases.[216] Bottom-up coarse-graining methods were first utilised to develop a chemically specific CG model of an azobenzene-based mesogen, 8AB8, via the application of IBI on its molecular fragments.[163] However, the resulting model was found to be unsatisfactory in capturing the correct phase behaviour. By employing the IBI method to a supercooled isotropic liquid of 8AB8 at a temperature just below the isotropic–smectic transition, the CG model is capable of capturing the SmA phase and the phase transition within a reasonable margin.[164] The choice of thermodynamic state point for the reference is interesting as it is suggested that directly parametrising to the SmA phase would introduce undesirable periodicity and long-range effects into the potentials; therefore, a system with local nematic order that is overall isotropic and is at the target temperature/density was selected. Similarly, the IBI method has been applied to the alkyl-cyanobiphenyl-based mesogen, 5CB. Here, the CG model incorrectly exhibits a smectic phase but the nematic phase can be captured by artificially scaling up certain interactions.[217] The fragment-based approach (using IBI on isotropic liquids) was re-examined to construct a CG model for 5CB, which included additional targets of the density and structural features of the nematic phase in its parametrisation.[218] This study was successful in capturing the nematic phase and its properties. The transferability of the aforementioned CG force field was tested for *n*CB homologues, where the expected smectic phases are observed.[219]

This chapter starts with the coarse-graining of benzonitrile-heptane systems. These consist of the pure liquids of benzonitrile ($x_{\mathrm{PhCN}} = 1$) and heptane ($x_{\mathrm{PhCN}} = 0$), and their miscible mixture ($x_{\mathrm{PhCN}} = 0.5$). Common coarse-graining methods are applied to these systems in order to assess the effectiveness of the approaches on a moderately complex system of molecules with differing molecular structures and aromatic–aliphatic

cross-interactions. The methods used are the bottom-up IBI and MS-CG methods, and the top-down Martini 3 force field. In essence, this comparative study is similar to that performed by Potter et al. on benzene-octane mixtures.[170] In the current study, the aim is to investigate the representability of the CG models constructed by various methods prior to the coarse-graining of more complex systems, such as thermotropic and chromonic liquid crystals. The assessment criteria for the structural accuracy of the models are the RDFs of the various components, whereas the thermodynamic properties to be considered are the enthalpy of vaporization and free energy of solvation.

The second section of this chapter focuses on the simulation of thermotropic alkyl-cyanobiphenyl-based mesogens, for which benzonitrile and heptane are molecular fragments. This is concentrated on CB7CB (see Figure 3.1(b)), which forms the twist-bend nematic phase.[23–26] Firstly, CB7CB will be studied at an all-atom resolution in terms of its conformational chirality and phase behaviour. Using single-molecule simulations in the gas phase, the scope of the investigation is expanded to the dimer series CBnCB (where n = 6, 7, 8 or 9) and CBX(CH$_2$)$_5$$Y$CB (where X/Y = CH$_2$, O or S). The effect of the spacer length and heteroatom linkages are discussed in terms of the conformational distributions of the bend angle and their molecular chirality. The N_{TB} phase, from simulations of CB7CB, was used as the AA reference to develop several CG models (via the IBI, MS-CG and Martini methods) which will be compared and discussed.

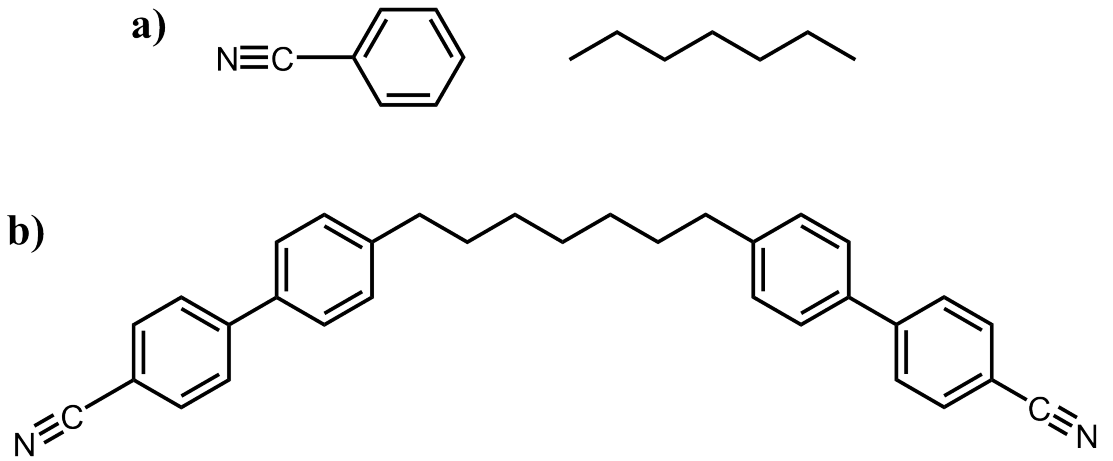


Figure 3.1: (a) Molecular stuctures of (a) PhCN (left), heptane (right) and (b) the liquid crystal dimer CB7CB.

3.2 Computational details

Force fields

The General AMBER Force Field (GAFF) was used for all atomistic simulations (see Section 2.1).[220] The Antechamber package from AmberTools18[221] was employed to obtain Lennard-Jones potentials and partial charges for the force field, with the atomic charges being determined via the AM1-BCC method.[222] The resulting GAFF topologies were converted into the appropriate input files for GROMACS using the ACPYPE script.[223]

Parametrisation of coarse-grained models

Bonded parameters for the PhCN, heptane and CB7CB CG models were determined by calculating probability distributions from liquid-phase MD simulations, and defining equilibrium values for the parameters based on the values of maximum probability in the distributions. The mapping schemes and bonded parameters used for each model in this chapter will be detailed in their relevant sections.

The non-bonded CG potentials can be obtained via the IBI method (see section 2.5) and this procedure was performed with the VOTCA-CSG package, version 1.4.1.[162,175,212] RDFs for this method were obtained from AA reference trajectories of 1000 snapshots. A linear pressure correction was applied (see Equations 2.36 and 2.37) so that the system exhibited a pressure of 1 bar at the correct density.

The MS-CG method (see section 2.6) was utilised to parametrise FM CG models, in the hybrid force matching scheme.[175,176] This was carried out in the BOCS (Bottom-up Open-source Coarse-graining Software) package.[224] Reference trajectories for this method contained 1000 frames. The linear pressure correction (see Equations 2.36 and 2.37) was employed to ensure that a pressure of 1 bar is exhibited at the correct density.

The Martini 3 force field is composed of a substantial library of bead types, based on LJ potentials, which can be used to develop CG models (see section 2.7).[106] It is noted that this study was conducted prior to the formal publication of Martini 3, and so uses the parameters released in the open beta version, unless otherwise stated.[225]

Simulation details

Atomistic MD simulations were carried out using the GROMACS 2018.7 molecular dynamics simulation package for PhCN-heptane systems with a total of 1000 molecules.[226] A cutoff of 1.2 nm was used for all short-range interactions, where long-range electrostatics were treated with the PME method.[146] After minimisation, a 100 ps pre-equilibration run in the *NVT* ensemble was carried out using the Berendsen thermostat followed by a 100 ps pre-equilibration in the *NPT* ensemble with the addition of the Berendsen barostat.[150] An equilibration run of 200 ps and a subsequent production simulation for 100 ns employed the Nosé–Hoover thermostat[152,153] to maintain a constant temperature of 300 K, and the Parrinello–Rahman barostat[154] to maintain a constant pressure of 1 bar. Time constants of 1 ps and 5 ps were used for the thermostat and barostat, respectively. A leap-frog algorithm was used with a time step of 1 fs for equilibrations with an increase to 2 fs for production simulations, where constraints were applied using the LINCS method.[144] A further 100 ns MD simulation was carried out for each system to collect data.

All-atom MD simulations for CB7CB were performed using GROMACS 2021.1 for systems of 512 molecules, but the simulation protocol is largely identical to the PhCN-heptane systems. Simulations for the liquid crystal system used semi-isotropic pressure coupling, such that the *x*/*y* and *z* dimensions can vary independently.

CG simulations were performed using a 2 fs time step for equilibrations and 5 fs for production runs, where the length of each simulation stage follows their atomistic counterparts. The simulation parameters for the IBI and FM CG models are also identical to the atomistic models, unless otherwise stated. No constraints were applied for CG simulations. Time constants of 1 ps and 8 ps were used for the thermostat and barostat, respectively. The Martini 3 models used a reduced cutoff of 1.1 nm for all interactions and a dielectric constant of $\epsilon_{\mathrm{r}} = 15$.[227] These models also used the velocity rescaling thermostat[151] and the reaction-field method for electrostatics[145], where $\epsilon_{\mathrm{rf}} = \infty$ beyond the cutoff.

Free energy calculations

The enthalpy of vaporization, ΔH_{vap}, is calculated by

$$\Delta H_{\mathrm{vap}} = \langle E_{\mathrm{pot}}(\mathrm{g}) \rangle - \frac{\langle E_{\mathrm{pot}}(\mathrm{l}) \rangle}{n_{\mathrm{mol}}} + k_{\mathrm{B}}T, \tag{3.1}$$

where $E_{\mathrm{pot}}(\mathrm{g})$ and $E_{\mathrm{pot}}(\mathrm{l})$ are the potential energies in the gas and liquid phases, respectively, and n_{mol} is the number of molecules in the liquid phase. $E_{\mathrm{pot}}(\mathrm{g})$, which represents the intramolecular energy in the ideal gas, is determined by a 100 ns single-molecule simulation in a vacuum using the stochastic dynamics integrator.

Free energies of solvation, ΔG_{solv}, were calculated by decoupling the intermolecular interactions of a single molecule from its surroundings, where the solvation of the species is the reverse of this process (see section 2.3). Coulombic interactions were first decoupled linearly before the van der Waals interactions, with a total number of 41 states of λ between 0 and 1, with a spacing of 0.05. Simulations for each λ state consisted of a 1 ns equilibration followed by a 5 ns production run for data collection using the stochastic dynamics integrator. Calculations for the CG models only consisted of 21 states, as there are no partial charges present.

Chirality order parameter

Ferrarini et al. developed a surface chirality model to determine a chirality order parameter, χ, based on a molecular isosurface which describes the coupling between the surface chirality and orientational order of the molecule.[228–231] The value of χ is positive for right-handed conformers and negative for left-handed ones. Here, molecular isosurfaces were generated using the Simple Invariant Molecular Surface (SIMS) method developed by Vorobjev and Hermans.[232] This utilised a rolling sphere algorithm where a spherical probe with a diameter of 5 Å produced a corresponding van der Waals surface of a molecule with a point resolution of 10 dots per Å^2. Three tensors are obtained from the isosurface for the calculation of χ: the surface tensor (***T***), the helicity tensor (***Q***) and the ordering matrix (***S***). Numerical integration of the normal vectors over the surface points yields the surface tensor, ***T***, whereas integration over the molecular surface produces the helicity tensor, ***Q***. The ordering matrix, ***S***, is calculated by integration of an orientational distribution function involving an orienting potential describing the orientation of the solute molecule within a liquid crystal phase.[233] Finally, the chirality order parameter is calculated by

$$\chi = -\left(\frac{2}{3}\right)^{\frac{1}{2}} (Q_{xx}S_{xx} + Q_{yy}S_{yy} + Q_{zz}S_{zz}) \tag{3.2}$$

where S_{ii} are the diagonal elements of the ordering matrix (***S***) and Q_{ii} are the diagonal

elements of the helicity tensor ($\boldsymbol{Q}$). Both these components are expressed in the principal axis system of the surface tensor, $\boldsymbol{T}$, which defines the molecular principal axes (T_{xx}, T_{yy} and T_{zz}) when diagonalised and describes the tendency for the axes to align with the rigid part of the mesogen.

3.3 Results and discussion

3.3.1 Benzonitrile-heptane systems

3.3.1.1 Coarse-grained potentials

The mapping schemes used for PhCN and heptane are shown in Figure 3.2. For PhCN, the cyano group is mapped to a single site and the phenyl ring is mapped in a 3:1 ratio. Mapping schemes for benzene-based systems reported in the literature range from ultra-coarse-grained (6:1) models[123,234,235] to 3:1[218,219] and 2:1[102,170] representations, where a 3:1 mapping is favoured in CG models of phenyl-based liquid crystals. Since this chapter will present work on a cyanobiphenyl-based mesogen, the phenyl group in PhCN was mapped in the same way as the CB7CB CG model. The CG model of heptane uses a 2:1 mapping for the outer carbons and a 3:1 mapping for the central group. This mapping for heptane allows for the flexibility of the molecule to be captured in the CG model. Overall, the mappings give rise to 2 bead types in PhCN and 1 bead type for heptane, where the CG beads are assumed to be of the same type and differ in mass only. CG potentials obtained from IBI and MS-CG are presented in Figures 3.3, 3.4 and 3.5 for the PhCN, heptane systems, and their 1:1 mixture, respectively. For the Martini 3 models, bead types of N = TC6, C = SC4 and O = TC1 were used (where the central O bead in heptane is defined as SC1).

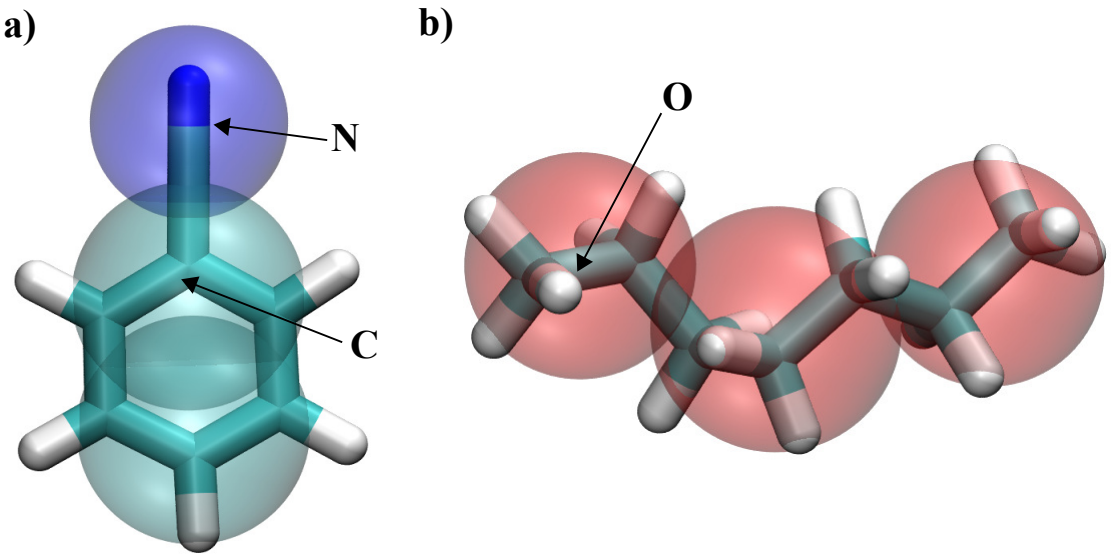


Figure 3.2: Coarse-grained mapping schemes for (a) PhCN and (b) heptane.

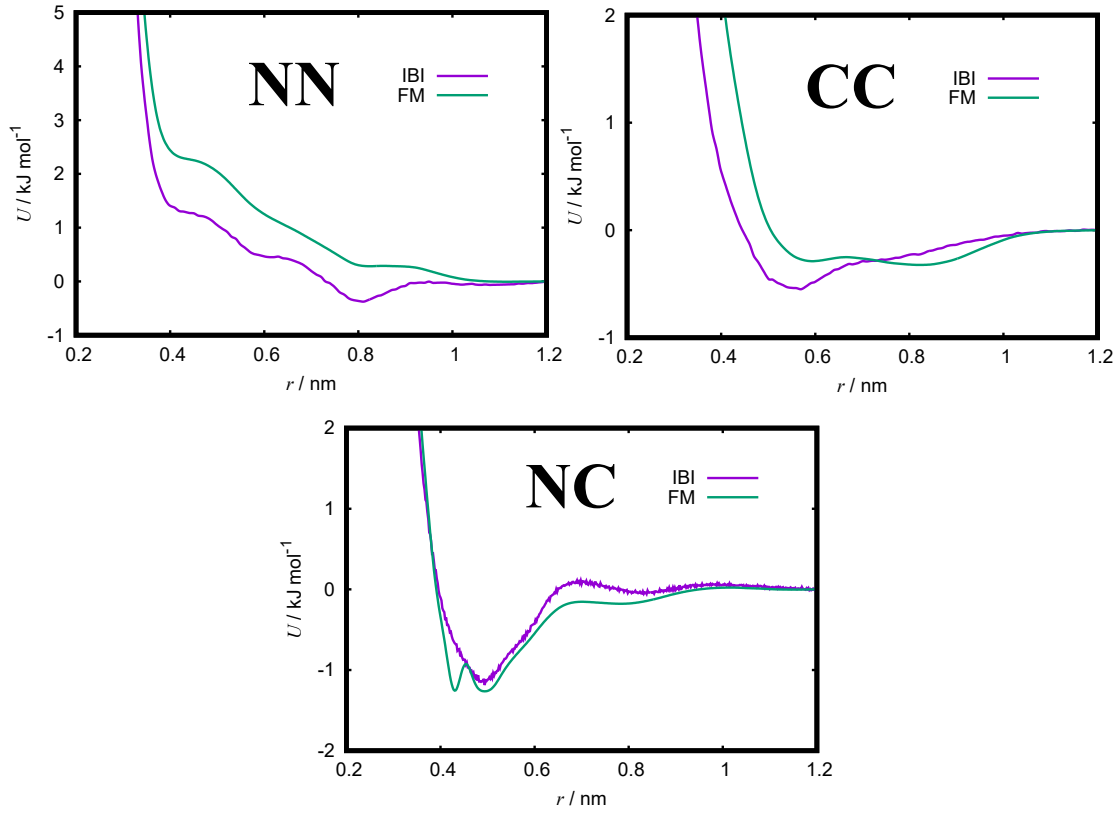


Figure 3.3: Coarse-grained potentials for the IBI and FM models of PhCN.

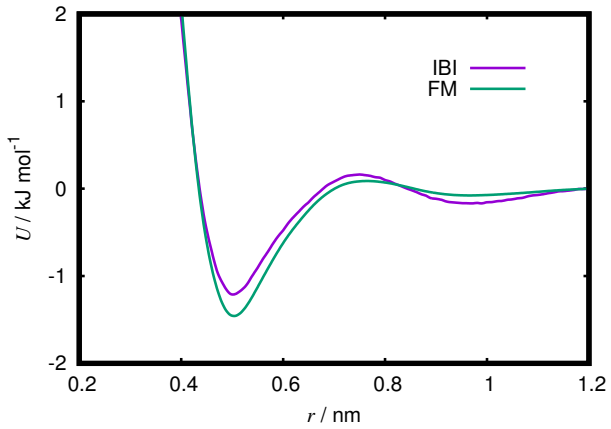


Figure 3.4: Coarse-grained potential (OO) for the IBI and FM models of heptane.

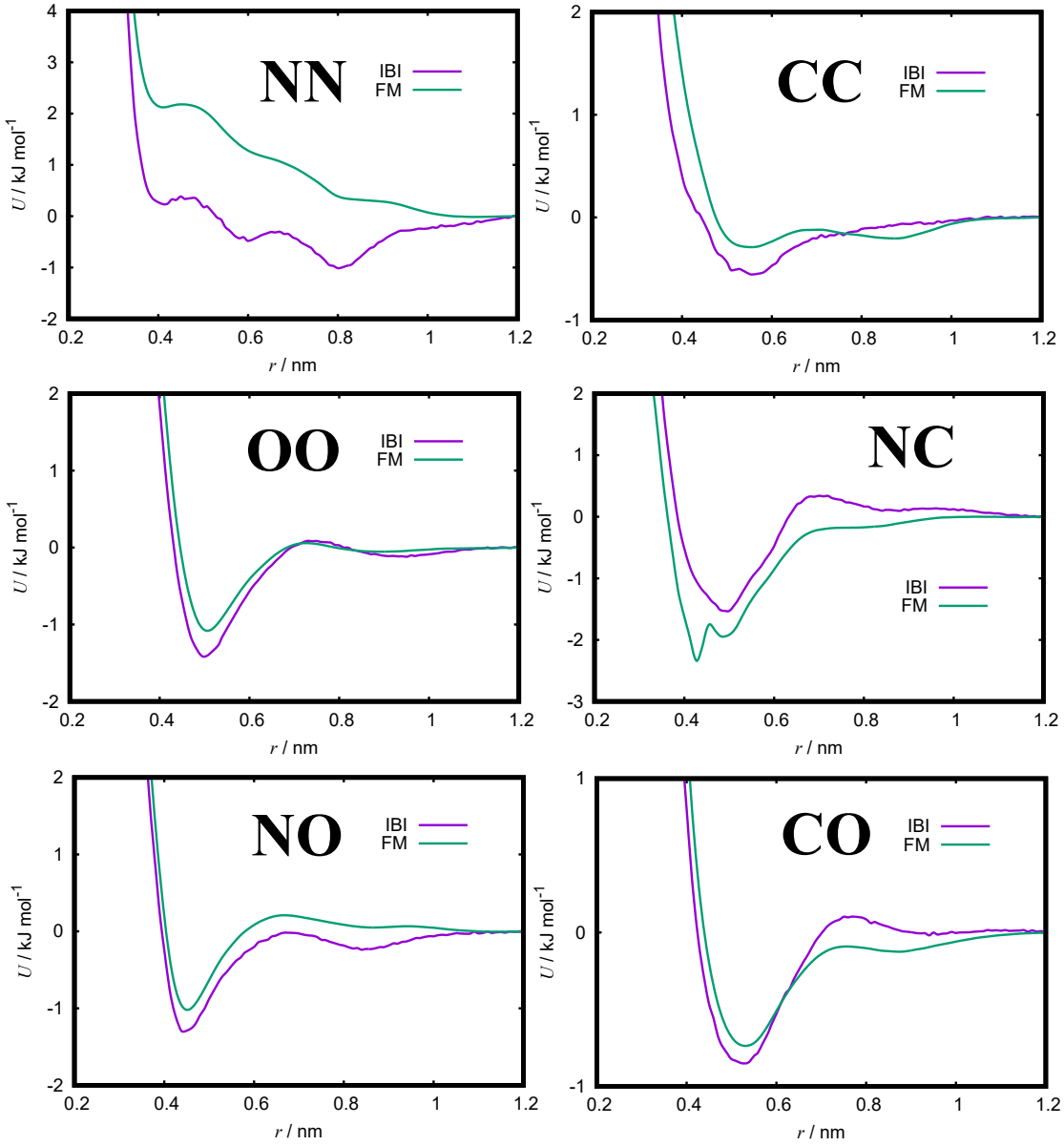


Figure 3.5: Coarse-grained potentials for the IBI and FM models of the PhCN-heptane mixture.

3.3.1.2 Structural accuracy

The representability of the CG models parametrised through the various methods is first assessed for their structural accuracy via RDFs between the bead types present in the systems. The RDFs for the systems are shown in Figures 3.6, 3.7 and 3.9. A simulation snapshot of the IBI CG model for the PhCN-heptane mixture is shown in Figure 3.8. As expected, the IBI CG model exactly reproduces the AA RDFs due to their use as targets in the parametrisation. The FM CG models similarly perform well in reproducing the local structure as the force profiles used in the MS-CG method implicitly contain information on the structure. For the Martini 3 models, it is found that their ability to capture the correct RDFs is not as good, compared to IBI and MS-CG, with the RDFs typically having a more prominent peak at slightly shorter distances. This over-structuring likely arises due to two reasons. The first is the CG bead types all have favourable LJ interactions between them that are not offset by interactions with a solvent here. In contrast, the NN interaction potentials (in the IBI and FM models) are largely repulsive in nature and the other interactions are fairly weak. The second reason is the limited number of bead sizes for Martini 3 which, in this case, do not fully represent the sizes of the underlying chemical groups. On the other hand, the IBI and MS-CG methods will produce CG potentials with interaction distances matching the AA system, and determine customised cross-interactions between the CG sites. It is noted that the nature of the Martini 3 force field allows for CG models to be developed quickly and for the MD simulations to be more computationally efficient compared to the IBI and FM models. This offers several advantages in cases where the CG model is only intended to be qualitative or if the feasibility of applying the IBI/MS-CG method is low due to a high number of interactions (or for a complex system). The use of the Martini 3 framework also does not require the simulation of an AA reference.

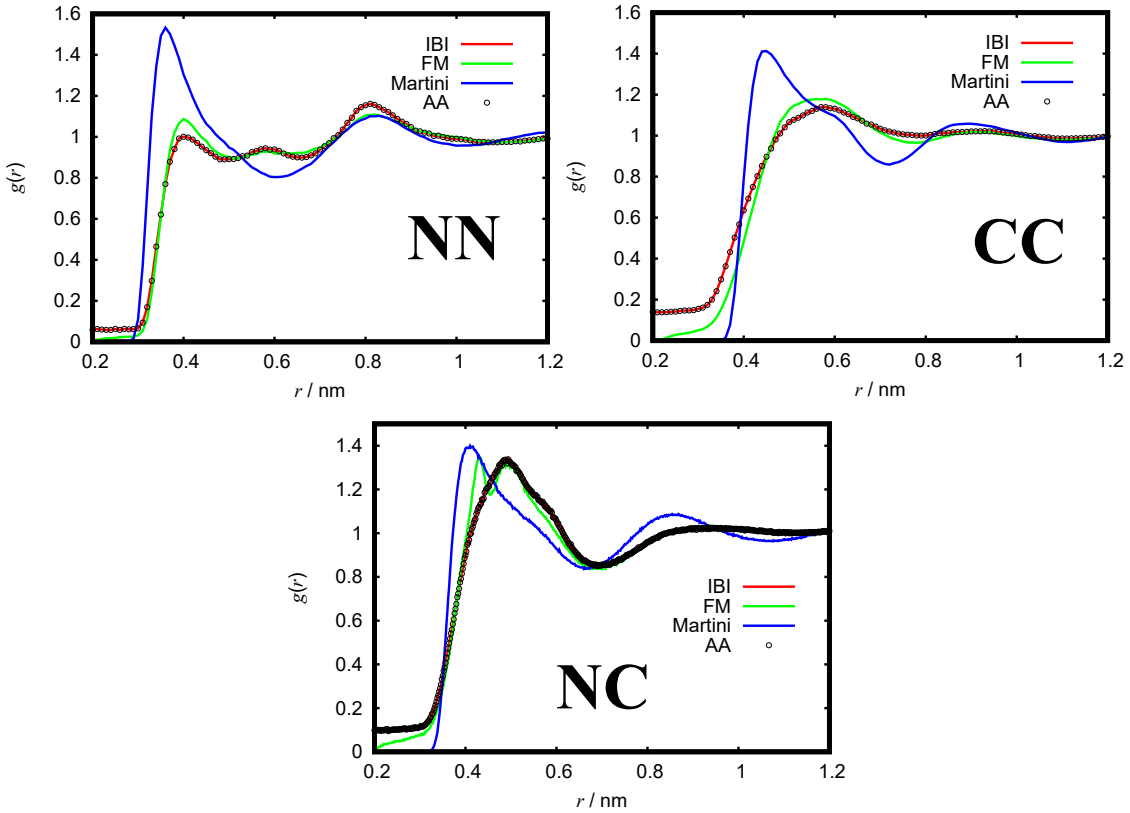


Figure 3.6: Radial distribution functions for the IBI, FM and Martini 3 models of PhCN.

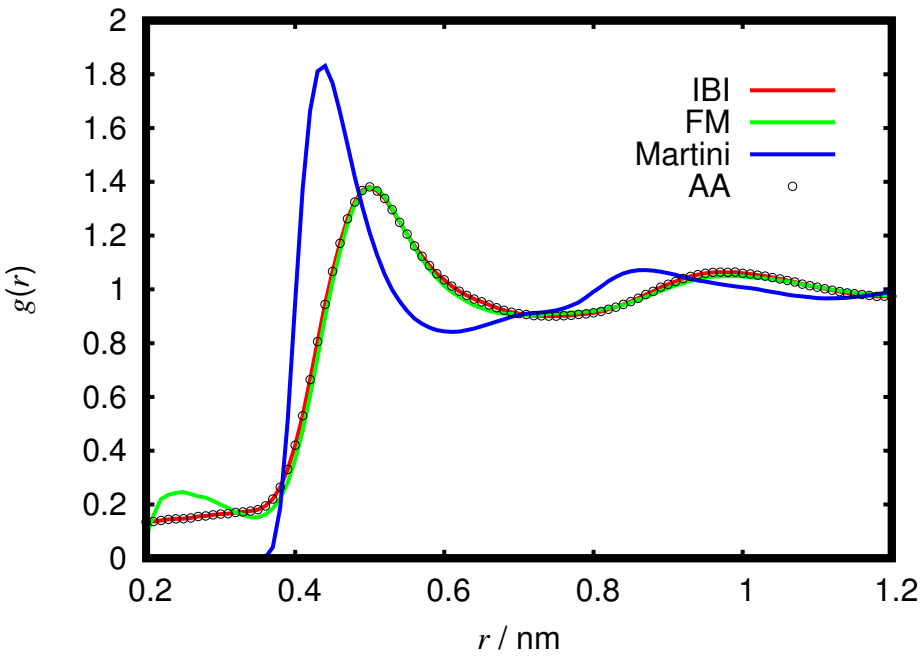


Figure 3.7: Radial distribution function (OO) for the IBI, FM and Martini 3 models of heptane.

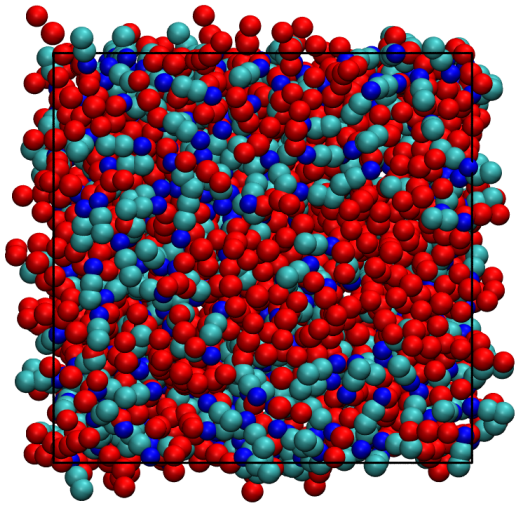


Figure 3.8: Simulation snapshot of the IBI CG model for the PhCN-heptane mixture.

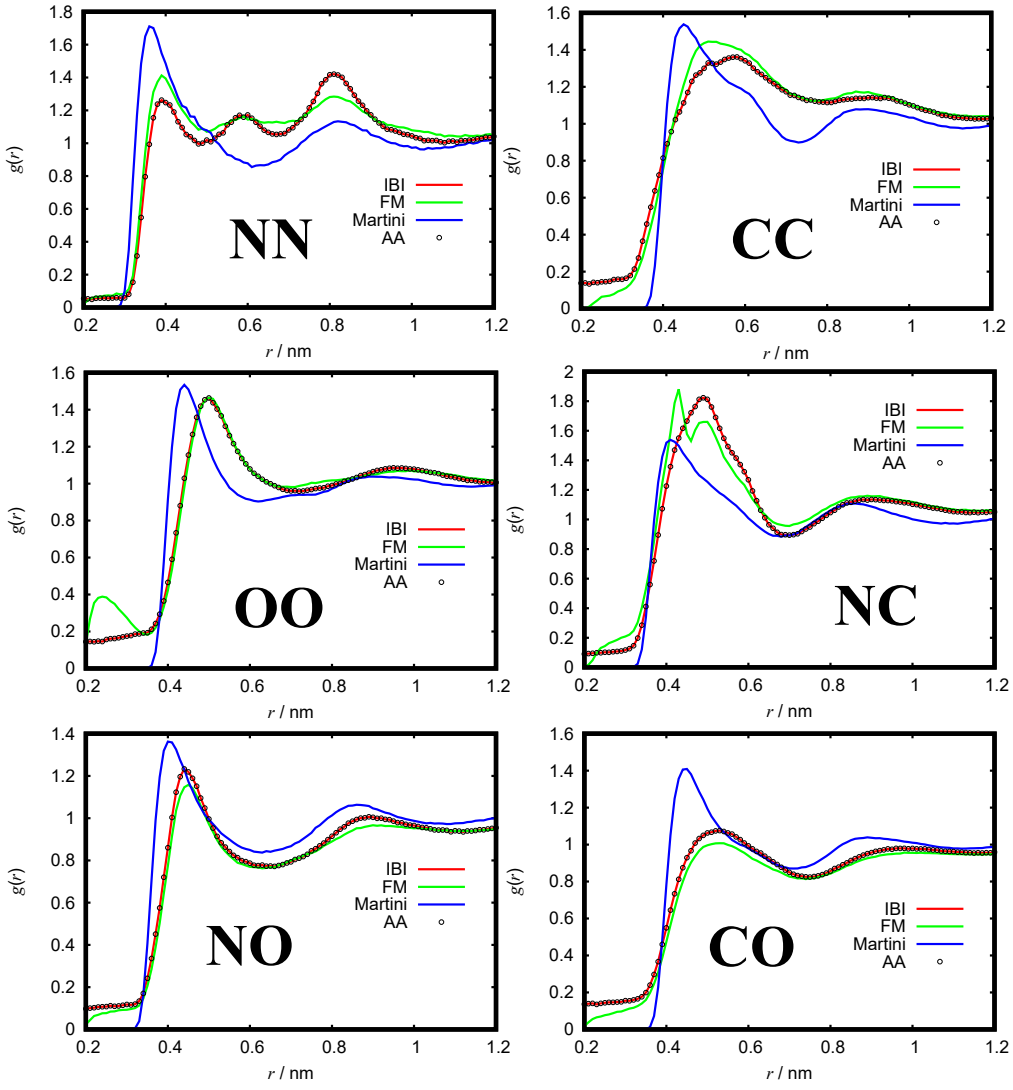


Figure 3.9: Radial distribution functions for the IBI, FM and Martini 3 models of the PhCN-heptane mixture.

3.3.1.3 Thermodynamic properties

Following on from structural accuracy, the CG models are now assessed in terms of capturing the thermodynamic properties. The enthalpies of vaporization, ΔH_{vap}, and free energies of solvation, ΔG_{solv}, are summarised in Table 3.1. For ΔH_{vap}, the AA values calculated here compare favourably (within 6 %) to experimental values of 55.5 kJ mol^{-1} and 36.4 kJ mol^{-1} for PhCN[236] and heptane[237], respectively. The values of ΔH_{vap}, for PhCN and heptane, determined for the IBI models are about half that of the reference values. Similarly, the FM models produce ΔH_{vap} values that are too low but to a greater extent than for the IBI models. This demonstrates the focus of these methods in reproducing structural features rather than thermodynamic properties. In contrast, the Martini 3 force field which is parametrised to experimental thermodynamic data performs well in capturing ΔH_{vap} for the systems. It is noted that refinement of the Martini 3 models to better reproduce thermodynamic properties is readily achieved by selecting different bead types to weaken/strengthen interactions or shift the balance between them.

Table 3.1: Enthalpies of vaporization (ΔH_{vap} / kJ mol^{-1}) and free energies of solvation (ΔG_{solv} / kJ mol^{-1}) for the IBI, FM and Martini 3 CG models of systems with number fractions $x_{\mathrm{PhCN}} = 1$, $x_{\mathrm{PhCN}} = 0.5$ and $x_{\mathrm{PhCN}} = 0$.

Property	AA	IBI	FM	Martini 3
ΔH_{vap}(PhCN)	52.3 $\pm$ 0.4	25.2 $\pm$ 0.4	13.5 $\pm$ 0.5	47.5 $\pm$ 0.4
ΔH_{vap}(heptane)	38.5 $\pm$ 0.5	16.9 $\pm$ 0.6	6.2 $\pm$ 0.5	38.7 $\pm$ 0.4
ΔG_{solv}(PhCN, $x_{\mathrm{PhCN}} = 1$)	-38.40 $\pm$ 0.10	-8.30 $\pm$ 0.02	-3.7 $\pm$ 0.4	-31.91 $\pm$ 0.06
ΔG_{solv}(heptane, $x_{\mathrm{PhCN}} = 0$)	-19.81 $\pm$ 0.14	-9.42 $\pm$ 0.03	-9.50 $\pm$ 0.04	-27.4 $\pm$ 0.2
ΔG_{solv}(PhCN, $x_{\mathrm{PhCN}} = 0.5$)	-32.9 $\pm$ 0.3	-8.37 $\pm$ 0.03	-5.6 $\pm$ 0.5	-30.49 $\pm$ 0.07
ΔG_{solv}(heptane, $x_{\mathrm{PhCN}} = 0.5$)	-22.5 $\pm$ 0.2	-9.17 $\pm$ 0.07	-8.21 $\pm$ 0.12	-26.91 $\pm$ 0.08

Generally, ΔG_{solv} is captured well with the Martini 3 models but to a lesser extent compared to ΔH_{vap}. This, once again, shows the strength of the Martini framework in terms of representing the thermodynamics. The IBI and FM models are weaker in this area, where the solvation free energies for these models are much lower than the AA values. The act of coarse-graining causes a shift in balance between entropic and enthalpic contributions, which can manifest as weaker interactions in the CG model.[72] The removal of degrees of freedom in the CG model decreases the entropic contribution to the free energy, and the enthalpic term is reduced to compensate for this (accroding to $G = H - TS$). It is found that ΔG_{solv} for heptane, in the pure liquid and the mixture, are slightly more agreeable with

the target values than for PhCN. This could indicate that the incorporation of polarisability effects of the cyano or phenyl groups into the CG potentials is poorer than required, and could be addressed by placing partial charges on the CG beads, resulting in a rudimentary dipole, or using a charged virtual site to mimic a quadrupole. However, it is unlikely that these approaches would improve ΔG_{solv} for heptane as it is nonpolar. While all the thermodynamic properties are underestimated for the bottom-up CG models, the FM models fare worse than the IBI models. This is also observed in benzene-octane systems, where the FM models overestimate ΔG_{solv} but the IBI models are fairly accurate.[170] Here, differences could arise from how effectively the true nature of the interactions is captured in the target distribution functions, particularly those of a pairwise definition. Considering the cyano groups (NN interactions), it is possible that coarse-graining processes approximate its interaction as wholly repulsive, but there may be electrostatic effects which result in spontaneous structure at the local level. For this interaction, IBI reproduces the RDF which shows some ordering of the cyano groups, whereas the MS-CG method interprets this interaction as completely repulsive due to the forces between the cyano groups. Thus, the attractions within the IBI model are higher than that of the FM model resulting in the lower ΔH_{vap} and ΔG_{solv} observed for the FM models. This may, however, be an issue at the atomistic level, where a polarisable force field could better capture the system properties and behaviour prior to bottom-up coarse-graining. At the CG level, three-body effects and local density potentials could be employed to improve the representability.[188–191]

The IBI and FM CG models both produce density values in good agreement with the AA systems, as the pressure correction ensures that a pressure of 1 bar is obtained at the target AA volume/density. The densities calculated are 895 kg m^{-3}, 670 kg m^{-3} and 810 kg m^{-3} for the AA systems of PhCN, heptane and the 1:1 mixture, respectively. The Martini 3 models were found to overestimate the density in all cases with values of 1057 kg m^{-3}, 819 kg m^{-3} and 850 kg m^{-3} for PhCN, heptane and the mixture, respectively. This discrepancy in the Martini 3 models originates in the bead sizes (where the volume of the box is smaller as the molecules pack closer) and from the stronger attractions between the CG sites, which are approximately double that of the IBI and FM interaction strengths.

Overall, no one coarse-graining approach is found to satisfy all the assessment criteria for the representability of the PhCN-heptane systems. While the bottom-up (IBI and MS-CG) methods display excellent structural accuracy in their resultant CG models, the

performance in reproducing thermodynamic properties is poor. The opposite is true of the Martini 3 models, although the structural accuracy is reasonable. Thus, it could be suggested that the simplest route to developing CG models, that are highly representative in all respects, is to optimise Martini 3 models to reproduce appropriate observables.

To test this approach, preliminary work was carried out using the ForceBalance[238] software package to systematically optimise a CG model of Martini 3 water. Using a Newton–Raphson algorithm, the LJ parameter set was optimised to produce several models: the first model exactly captured the density of water at 300 K, a second model reproduced both the density and ΔH_{vap} at a single state point and a third model that showed the correct densities across a range of five temperatures.

3.3.2 Bent liquid crystal dimers and the twist-bend nematic phase

Introduction

Independently, Meyer[239] and Dozov[21] predicted that liquid crystal phases with local chirality could be exhibited by bent achiral mesogens and was, later, supported by Memmer[22] in computer simulations. It was proposed that a pure uniform bend in space, arising from the packing of bent molecules, is impossible and, thus, twist or splay deformations of the local director must accompany the spontaneous bend. In the twist-bend nematic ($\mathrm{N_{TB}}$) phase, a heliconical structure is formed in which the director is tilted with respect to the helical axis, where there are equal numbers of degenerate domains of opposite handedness. The $\mathrm{N_{TB}}$ phase was first discovered for the dimer 1,7-bis-4-(4′-cyanobiphenyl) heptane, CB7CB.[23–26] Other dimer-based species,[240–249] bent-core mesogens,[250,251] and supramolecular hydrogen-bonded systems[252–255] have been reported to manifest the $\mathrm{N_{TB}}$ phase. Despite the wide assortment of materials studied, a general structure-property relationship for the $\mathrm{N_{TB}}$ phase has not been established; however, it is found that a spatially uniform curvature is necessary irrespective of the underlying chemical groups.[256,257] Possible applications utilising the $\mathrm{N_{TB}}$ phase such as optical devices,[258–261] gels,[262] photoswitchable adhesives,[263] and photoalignment technology[264] have been reported.

The study of the emergent chirality in achiral liquid crystals has been facilitated by molecular simulation including: the calculation of molecular chirality and helical twisting powers;[265–270] predictions of the dark conglomerate phase;[271] and investigations of bent-core mesogens[29–33] An early CG model for a liquid crystal dimer represented the rigid

mesogenic segments with anisotropic Gay–Berne units which were connected by a flexible LJ chain.[272] Simulations using this model demonstrated the spontaneous growth of smectic phases from the isotropic liquid. At the same resolution, a similar model has been used to study the N_x phase, where it was assigned as the polar-twisted nematic phase.[273] Idealised models, such as a crescent-shaped model of connected hard spheres[274] or curved spherocylinders,[275–277] have been found to form the N_{TB} phase. At the all-atom level, conformational distributions[23,248] and the N_{TB} phase[25] have been investigated in conjunction with experiments.

In this section, all-atom simulations of bent liquid crystal dimers are presented. The dimer CB7CB is investigated in its phase behaviour, where the N_{TB} phase is characterised. The conformational distributions of the bend angle and chirality order parameter are calculated for the N and N_{TB} phases. Using single-molecule simulations in the gas phase, the scope of the investigation is expanded to the dimer series CBnCB (where n = 6, 7, 8 or 9) and CB$X(CH_2)_5Y$CB (where X/Y = CH_2, O or S). The all-atom results presented herein is published in the paper[278] "*All-atom simulations of bent liquid crystal dimers: the twist-bend nematic phase and insights into conformational chirality*." It is reproduced from Ref. 278 [G. Yu and M. R. Wilson, *Soft Matter*, 2022, **18**, 3087–3096] with permission from the Royal Society of Chemistry. Finally, various coarse-graining approaches are applied to the CB7CB system in order to develop CG models that can be utilised to study the N_{TB} phase.

3.3.2.1 Atomistic simulation

Phase behaviour

A pseudo-nematic starting configuration was set up by randomly inserting 512 molecules of CB7CB into a box (with dimensions of 8 nm x 8 nm x 16 nm), where the molecular long axes were all aligned with the z axis of the box. Initially, an additional improper dihedral potential (with an equilibrium angle of 180° and force constant of 25 kJ mol^{-1}) was applied across each molecule, between the terminal nitrogen atoms and the linking carbon of the mesogenic units with the alkyl spacer, and the corresponding atoms on the other cyanobiphenyl segment in order to enforce an all-*trans* configuration of the heptane chain. After a steepest-descent minimisation, a 100 ps pre-equilibration run in the *NVT* ensemble was carried out using the Berendsen thermostat followed by a 100 ps pre-equilibration

in the *NPT* ensemble with the addition of the Berendsen barostat. Simulations were performed using semi-isotropic pressure coupling which allows for the *x*/*y* and *z* dimensions of the system to vary independently. The system was compressed with a pressure of 100 bar for 1 ns, before an equilibration run of 500 ps was carried out using the Nosé–Hoover thermostat to maintain a constant temperature of 370 K, and the Parrinello–Rahman barostat to keep the pressure constant at 1 bar. At this point, the system exhibited a nematic phase which acted as an ideal starting configuration. A final equilibration of 1 ns was executed without the additional biasing potential to allow for equilibration of the molecular conformations and for the heliconical structure to start developing. Finally, a production simulation of 1 μs was performed. MD simulations were carried out on the new N8 CIR supercomputer Bede, housed at Durham University, using one NVIDIA V100 GPU, achieving a performance of 180 ns/day on a $\sim$30000 atom system.

Even after 100 ns, the system at 370K exhibited a heliconical structure which was allowed to fully equilibrate over a long simulation. A snapshot of this simulated phase is shown in Figure 3.10(a) in which molecules are coloured in blue, purple or red according to their azimuth angle (ϕ) in the range -180° $\leq \phi <$ -60°, -60° $\leq \phi <$ 60° and 60° $\leq \phi \leq$ 180°, respectively. The angle ϕ was calculated for the unit vector $\hat{\boldsymbol{a}}$ which corresponds to the molecular long axis of the molecule. A second unit vector $\hat{\boldsymbol{b}}$ can be defined as the unit vector which bisects the two vectors between a terminal nitrogen atom and the COM of the molecule. A schematic for these unit vectors is shown in Figure 3.11(c). In the final system, a helical structure in which molecules are tilted with respect to the helical axis (which coincides with the *z* axis) is clearly distinguished by visual inspection. Experimentally, it is expected that there are equal numbers of domains with opposite handedness, but only a left-handed helix is observed here. The resolution of the globally achiral N_{TB} phase into its chiral counterpart, the N^*_{TB} phase, which only has domains of one handedness can be achieved by introducing a chiral centre into the mesogen.[279] Here it may be that as the initial nematic phase resolves into a heliconical structure, one or more clusters of a specific handedness form randomly and template the formation of the N_{TB} phase with one handedness only. Repeated simulations starting from the pseudo-nematic phase may show that the N_{TB} phase with the other handedness can also arise, and that the probability of a either handedness being adopted is equal.

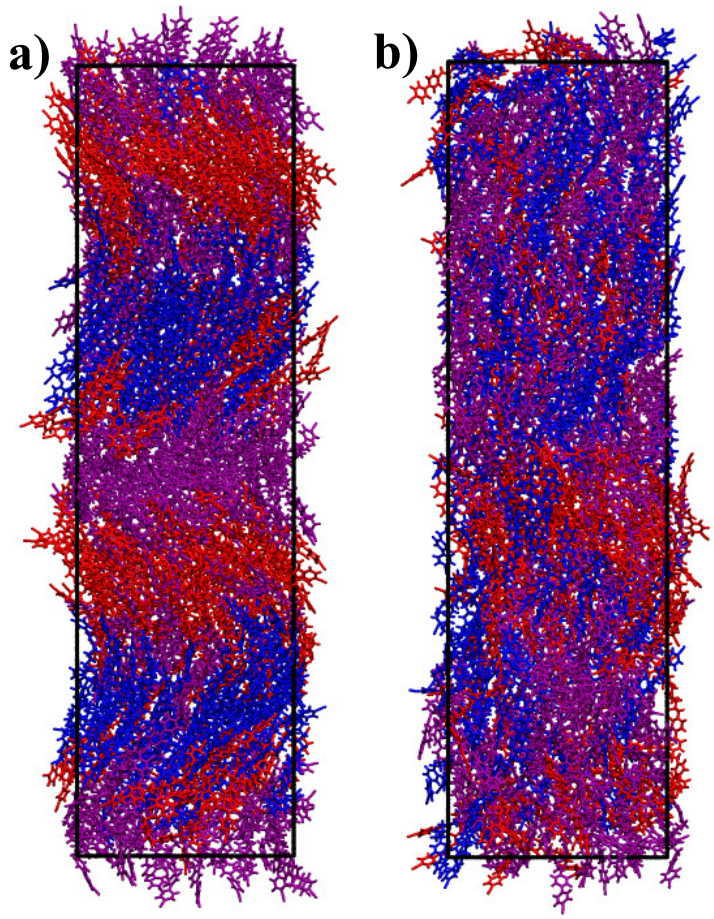


Figure 3.10: Simulation snapshots of a 512-molecule system of CB7CB showing the (a) N_{TB} phase at 370K and the (b) N phase at 440 K.

The defined unit vectors were utilised in the characterisation of the system. An orientational correlation function, S_{bb}, can be determined as a function of intermolecular distance along the helical axis, $r_{\parallel}$, by[280]

$$S_{bb}(r_{\parallel}) = \sum_{i}^{N} \sum_{j \neq i}^{N} c_{ij} (\hat{\boldsymbol{b}}_i \cdot \hat{\boldsymbol{b}}_j), \tag{3.3}$$

where

$$c_{ij} = \delta(|z_{ij}| - r_{\parallel}) / \sum_{i}^{N} \sum_{j \neq i}^{N} \delta(|z_{ij}| - r_{\parallel}). \tag{3.4}$$

$S_{bb}(r_{\parallel})$ describes the polar correlations throughout the system, z_{ij} is the distance with respect to the z components for the COMs of molecules i and j, and N is the number of molecules. From the results in Figure 3.11(a), it can be seen that $\hat{\boldsymbol{b}}$ precesses about the helical axis as expected for the N_{TB} phase. This behaviour has been found for simulated

phases previously.[273,274] The minima/maxima occur at distances $(n+1/2)p$, where n is 0, 1/2 and 1, and the pitch of this simulated phase is 8.35 nm. The measured pitch is in good agreement with those found experimentally in the N_{TB} phase for CB7CB of ~8 nm.[25,26]

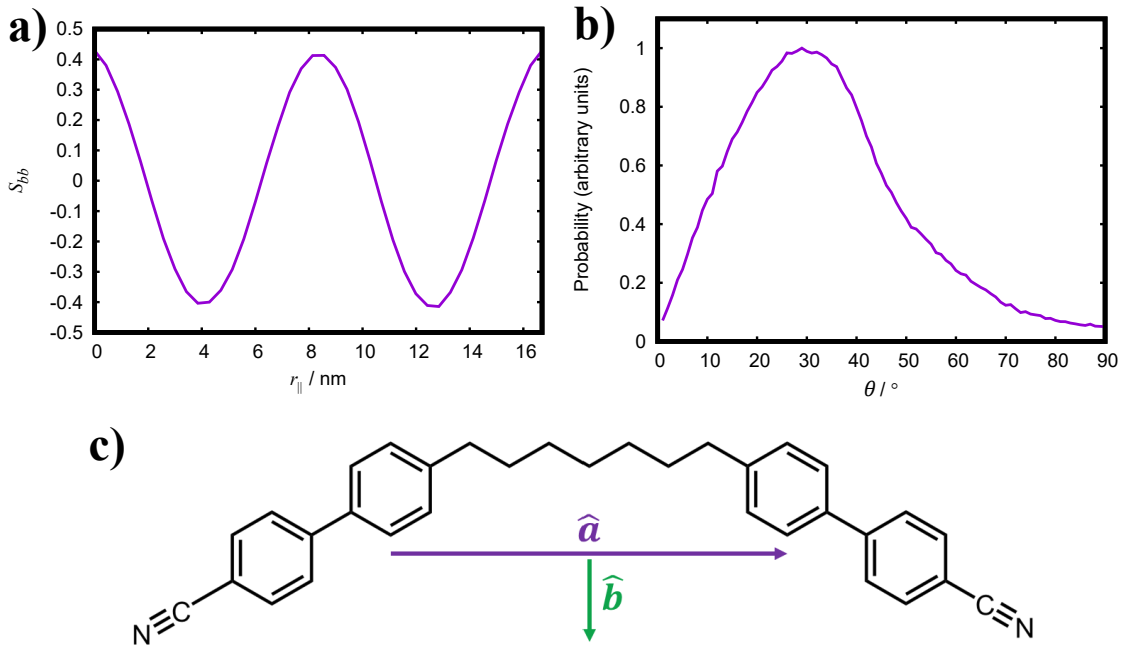


Figure 3.11: (a) Orientational correlation (S_{bb}) as a function of the helical axis, $r_{\|}$, and the (b) distribution for the conical tilt angle, θ, for the N_{TB} phase at 370K. (c) A schematic for the unit vectors $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$.

The conical tilt angle, θ, is defined as the angle between the unit vector $\hat{\boldsymbol{a}}$ and the helical axis. A histogram of θ values calculated for the system is presented in Figure 3.11(b), which shows that the molecules in the simulated phase are tilted with the most probable conical tilt angle measured to be 29°. This is slightly higher than the ~25° reported experimentally.[25,281,282]

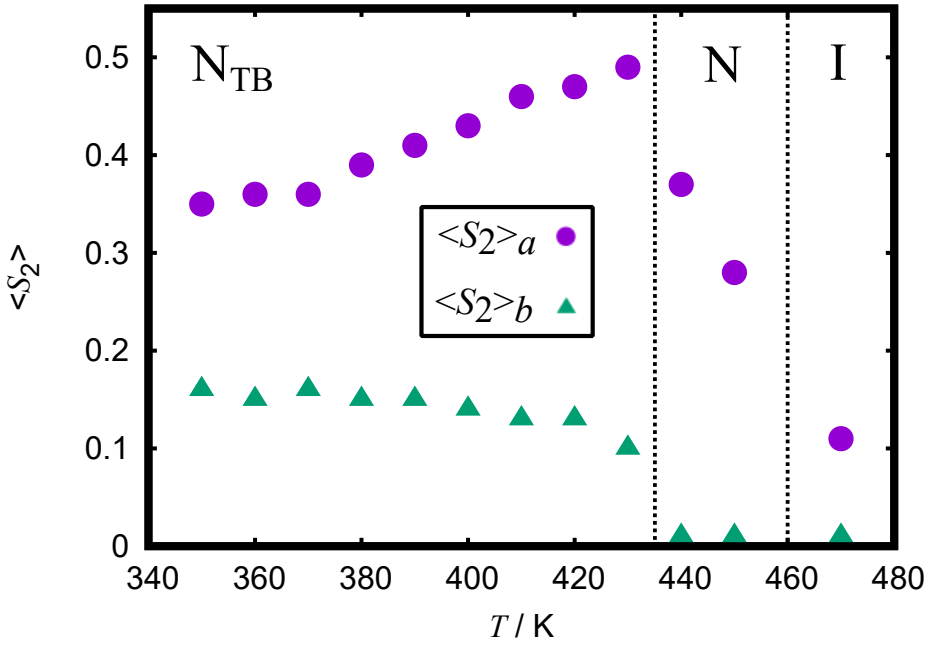


Figure 3.12: Average orientational order parameter, $\langle S_2 \rangle$, for unit vectors $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ as a function of temperature.

The concluding system at 370 K was used as the starting point for a cooling/heating sequence. In the temperature range of 350–470 K, simulations were performed at 10 K intervals for 200 ns each. The average orientational order parameter, $\langle S_2 \rangle$, was determined for unit vectors $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ at each temperature. The resulting phase diagram is presented in Figure 3.12 in which the stability range of the N_{TB}, N and I phases are indicated. The N_{TB}–N phase transition is defined as the temperature at which helical ordering is lost ($\langle S_2 \rangle_b \approx 0$) and the I phase is identified when, additionally, $\langle S_2 \rangle_a \approx 0.1$. It is noted that, via visual inspection, the system at 430 K appears nematic but does not satisfy the aforementioned criteria for the N phase. A simulation snapshot of the N phase for CB7CB at 440 K is shown in Figure 3.10(b). The value of $\langle S_2 \rangle_b$ is ~0.15 in the N_{TB} phase and decreases slightly as the system is heated which is expected as the helix unwinds upon its approach into the N phase.[281] The value of $\langle S_2 \rangle_a$ increases as the system is heated (i.e. there is an increase in nematic-like ordering) through the N_{TB} phase until the phase transition into the N phase. This trend is observed in X-ray and polarised Raman scattering experiments and the measured values for $\langle S_2 \rangle_a$ of ~0.35 in the N_{TB} phase compare favourably.[282,283] However, the phase transitions in the AA model, $T_{N_{TB}-N} = 435$ K and $T_{N-I} = 460$ K, occur at temperatures higher than experiment[26] where $T_{N_{TB}-N} = 376$ K and $T_{N-I} = 389$ K. Thus, it is found that this force field for CB7CB is likely too stiff. Since ΔH_{vap} for the molecular fragments (PhCN and heptane) are in good agreement with experimental

values, it is probably the torsional potentials which require optimisation. This has been demonstrated for GAFF in studies of liquid crystals, where the representation of structural and phase properties were improved.[33,50,271,284]

Conformational distributions and molecular chirality

The conformational properties of bent liquid crystal dimers can be described by the bend angle between the mesogenic units. The bend angle can be defined as the angle between the unit vectors representing the para axes of the cyanobiphenyl units (i.e. the vector calculated between the nitrogen atom and the carbon atom connecting the cyanobiphenyl unit to the spacer). Figure 3.13 shows probability distributions of the bend angle calculated for CB7CB for molecules in the N and N_{TB} phases. In both phases, the most probable bend angle obtained is 127°. This is somewhat higher than the value of ∼120° determined by Monte Carlo sampling of a single molecule[23,248] but in line with a bulk-phase MD simulation[25] value of 133°. It is found that the bend angle distribution does not change significantly between molecules in the N and N_{TB} phases, but the broadness of the peak in the N phase increases as expected for the higher temperature. The histograms also show a minor peak at 20° corresponding to hairpin conformers. NMR experiments suggest that the conformational distribution for CB7CB does not change significantly between the N and N_{TB} phases and that the chirality of the N_{TB} phase is a consequence of the bent molecular shape.[285]

The chirality order parameter, χ, was calculated for molecules in the N and N_{TB} phase. These distributions are presented in Figure 3.14(a). The plots demonstrate that CB7CB is statistically achiral (as the peaks are centred at $\chi = 0$) but can adopt chiral conformers with equal probability of left- or right-handedness (as the peaks are symmetric). This behaviour has been observed in the P*n*OPIMB series of bent-core mesogens.[32,33] The aforementioned studies found values of $|\chi|_{\max}$ to be ∼800 and greater whereas in this work, $|\chi|_{\max}$ is estimated to be ∼200 for CB7CB. If it is supposed that the extent of chirality of a mesogen is related to the chirality of its mesophases, then this could be a factor in accounting for why P8OPIMB can form the highly twisted B4 phase which consists of helical nanofilaments.[28] The ability of achiral bent-core mesogens[31,32] and CB7CB[286,287] to decrease the pitch of a N* phase can be hypothesised to be a result of the presence of chiral conformers with high helical twisting powers.[32,33] Interestingly, the distributions of χ are unchanged between

the N and N_{TB} phases, which indicates that there is no change in the molecular chirality of the conformations even as the phase gains chirality. This is in line with analogous findings for the conformational distributions.[285] An alternative method to sample conformers is to perform simulations of a single molecule in the gas phase with a stochastic dynamics integrator. The distribution of χ values obtained via this method is very similar to that from the liquid phases with the exception that the peak is broader. This is expected as there is greater freedom for the molecule to sample a wider range of conformers in the gas phase, noting that the temperature of the gas phase simulations is still 370 K.

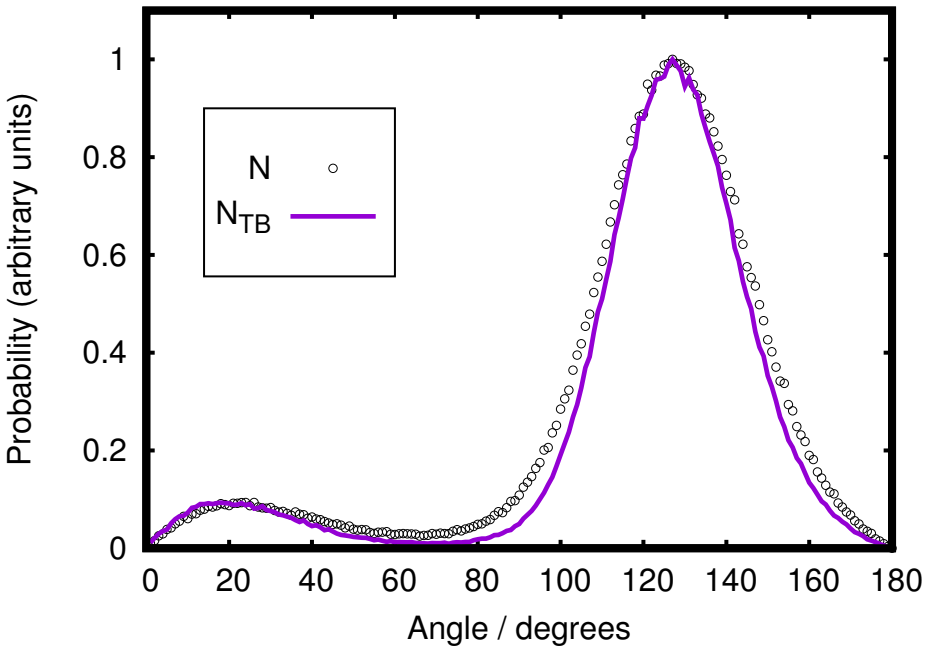


Figure 3.13: Distribution functions for the bend angle between the mesogenic units in the N and N_{TB} phases at 440 K and 370 K, respectively.

Since the conformational sampling provided by single-molecule simulations in the gas phase produces χ distributions that are approximate to those calculated from liquid crystal phases, this offers a route to high-throughput analysis of liquid crystal dimers in terms of the bend angle and conformational chirality. Here, the investigation focuses on the series of bent dimers CBnCB (where n = 6, 7, 8 or 9) and CBX(CH$_2$)$_5$$Y$CB (where X/Y = CH$_2$, O or S). The first series of mesogens allows for the odd-even effect and the effect of the spacer length on the conformations/chirality of the molecule to be studied. The second series consists of bent dimers all with 7 heavy atoms in the spacer and considers the effect of heteroatoms in the linking chain.

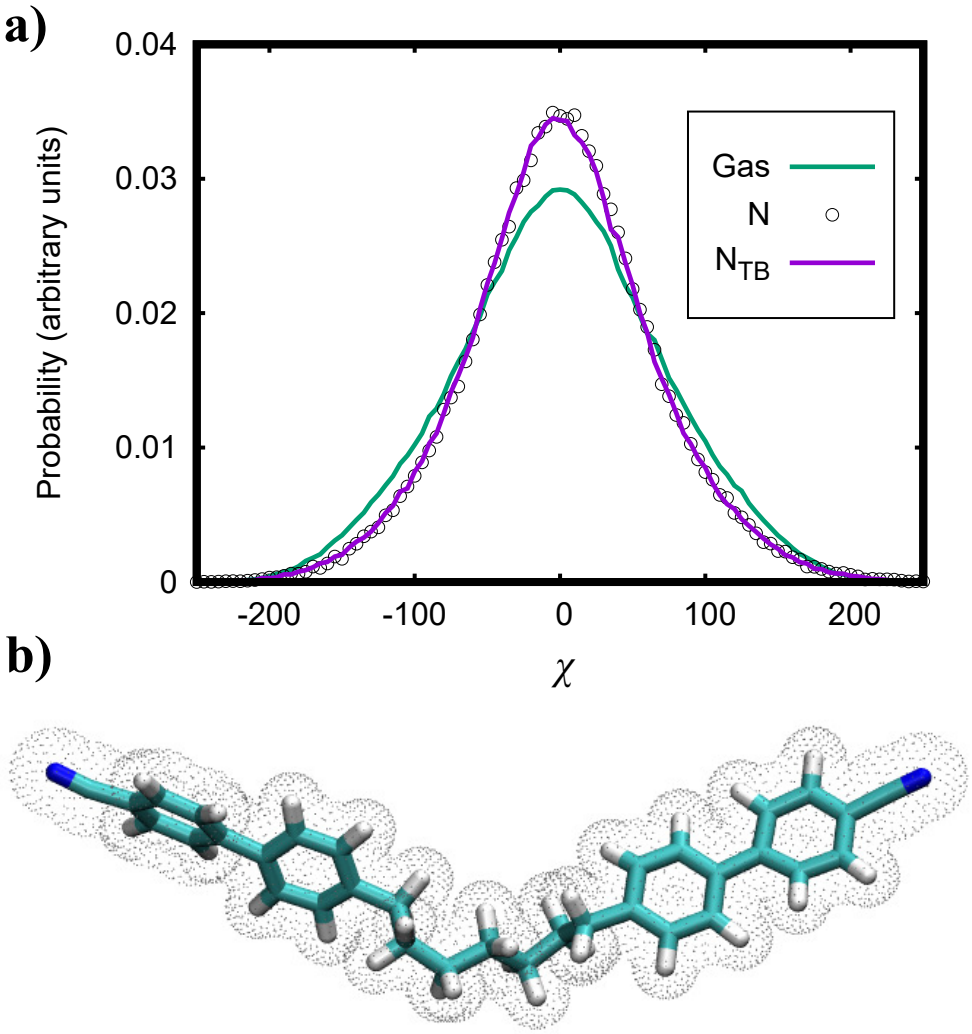


Figure 3.14: (a) Distribution functions of the chirality order parameter, χ, for a single molecule in the gas phase and for molecules in the N (440 K) and N_{TB} (370 K) phases. (b) A conformer of CB7CB with its molecular isosurface shown.

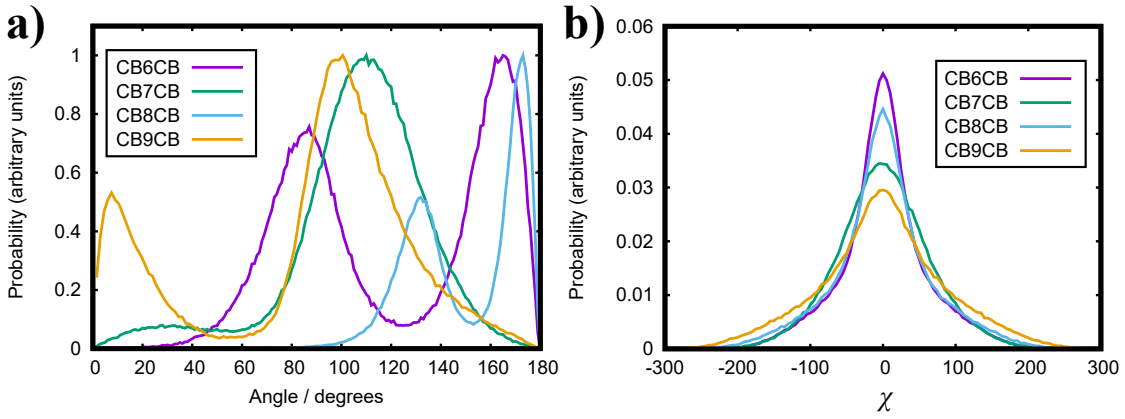


Figure 3.15: Distribution functions of the (a) bend angle and (b) chirality order parameter, χ, for CBnCB where n = 6, 7, 8 or 9.

The odd-even effect of the spacer on the phase behaviour has been studied extensively,

where the N_{TB} phase is only observed in odd-membered mesogenic dimers.[288–290] In an all-*trans* configuration, dimers with a spacer of odd parity will have a bent molecular shape whereas an even spacer will result in a linear shape. Moreover, the sensitivity of $T_{N_{TB}-N}$ to the bend angle has been demonstrated by a generalised Maier–Saupe theory.[291] Bend angle distributions for the CBnCB series are shown in Figure 3.15(a). Most probable bend angles of 165°, 110°, 172° and 99° for n = 6, 7, 8 and 9, respectively, are obtained. Values of 166° for CB6CB,[25] 111-120° for CB7CB,[23,248,292] and 102° for CB9CB[293] have been reported. The measured values are, overall, in good agreement with previous work and the correct trend is observed. The bend angle for CB7CB in the gas phase is lower than for the N and N_{TB} phases, which suggests that the effect of a nematic environment on the conformational distribution is significant. This suggests that single-molecule calculations, regardless of the level of theory, are not fully representative of the actual behaviour within a mesophase. This is expected as the constraints arising from the packing of molecules in the N phase will affect the conformers adopted by the mesogens as well as promote a more linear shape of the molecules.[294,295] For the even-membered dimers, the prevalent conformers are those with linear configurations, but the presence of bent conformers with bend angles of 85° and 130° for CB6CB and CB8CB, respectively, is found. This indicates that the analysis of a single conformer is not sufficient and, also, that these bent conformers are likely to be suppressed in the N phase.

The χ distributions for the CBnCB series are shown in Figure 3.15(b) and all indicate that the mesogens are statistically achiral but can adopt chiral conformers of opposite handedness with equal probability. The histograms for CB6CB and CB8CB are narrower than those for the dimers with odd spacers, which demonstrates that the more linear shape of the even-membered dimers results in a lower tendency to exhibit instantaneous chiral conformations with high helical twisting powers. For n = 6, 7 and 8, $|\chi|_{\max} \approx 200$ but is $\sim$250 for CB9CB. There is a slight increase in $|\chi|_{\max}$ as the length of the spacer increases regardless of the odd/even parity of the chain. Thus, this confirms that the ability to form the N_{TB} phase for dimers with a spacer of odd parity, or only exhibit the N phase for dimers with a spacer of even parity, is dependent on the molecular shape of the mesogen and not its potential to assume chiral conformers.[296]

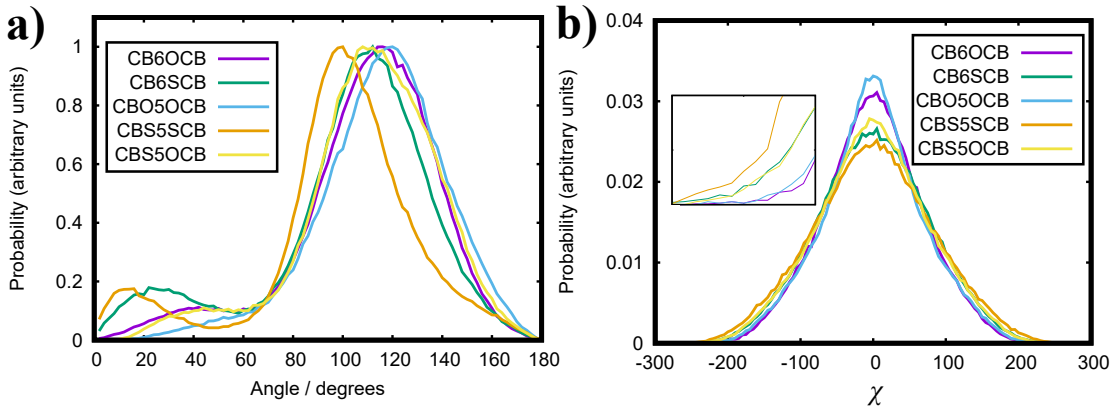


Figure 3.16: Distribution functions of the (a) bend angle and (b) chirality order parameter, χ, for CBX(CH$_2$)$_5$$Y$CB where X/Y = CH$_2$, O or S. The inset shows a focused view of the plot near $|\chi|_{\max}$.

In the CBX(CH$_2$)$_5$$Y$CB (where X/Y = CH$_2$, O or S), the N$_{\mathrm{TB}}$ phase is observed for all combinations.[248,292,297–299] While it is generally considered that CBOnOCB homologues mainly form the N phase, the N$_{\mathrm{TB}}$ phase is indicated for CBO5OCB by changes in the optical texture upon supercooling.[300,301] Since this series of dimers all have spacer lengths of 7 heavy atoms, any differences will mostly arise from the flexibility and local bond angle afforded by the heteroatom linkage. Here, bend angles (see Figure 3.16(a)) of 116° for CB6OCB, 110° for CB6SCB, 120° for CBO5OCB, 100° for CBS5SCB and 112° for CBS5OCB are measured. Conformational sampling provides a reported value of ∼120° for CB6OCB[248] and from optimised geometries of dimers with an all-*trans* configuration of the spacer,[292] bend angles of 143° for CBO5OCB, 94° for CBS5SCB and 119° for CBS5OCB are found. The results compare favourably with the exception of CBO5OCB. It is noted the literature value was obtained from a single conformer, albeit the minimum energy geometry, which neglects the flexibility of the molecule.

The distributions of χ are presented in Figure 3.16(b). It is observed that the distributions get broader and $|\chi|_{\max}$ increases for bent dimers in the order CBO5OCB, CB6OCB, CBS5OCB, CB6SCB to CBS5SCB. This is also the sequence of dimers going from the highest to lowest bend angle measured here. Thus, the trend where a lower bend angle results in a higher degree of conformational chirality is obtained. Experimentally, the effect of the bend angle on $T_{\mathrm{N_{TB}-N}}$ is unclear, where the sulfur-linked dimers unexpectedly exhibit a lower $T_{\mathrm{N_{TB}-N}}$ than the methylene versions despite having a smaller bend angle.[292,300] As $|\chi|_{\max}$ is commensurate with the bend angle, there seems to be no correlation between the extent of the conformational chirality and the stability of the N$_{\mathrm{TB}}$ phase

or transition temperatures. The full width at half maximum of the peaks for the bend angle (which is a measure of the flexibility) is $\sim$50° for CB7CB and all the CBX(CH$_2$)$_5$$Y$CB dimers, and so it cannot be said that there is a strong link between the flexibility and $T_{\mathrm{N_{TB}-N}}$.

3.3.2.2 Coarse-grained modelling

Following on from the coarse-graining of benzonitrile-heptane systems, this section focuses on CB7CB, where benzonitrile and heptane are molecular fragments of this mesogen. Thus, the mapping scheme (see Figure 3.17) used is identical to that used in Section 3.3.1 and results in a 13-site CG model. This consists of 3 bead types and a total of 6 pairwise interactions to parametrise. The choice of a 3:1 mapping for the carbons of the phenyl groups allows for the rod-like shape of the mesogenic units to be retained. It is noted that, perhaps, a 2:1 mapping would be more accurate but may introduce increased unnecessary complexity into the CG model. The CG potentials obtained via the bottom-up coarse-graining approaches are presented in Figure 3.18. The Martini 3 model follows the same bead types used for the chemical groups as the PhCN-heptane systems. In a comparison of the IBI/FM interactions of CB7CB and the PhCN-heptane mixture, a shift in the balance of interactions strengths can be seen, as the ordering of molecules within the liquid crystal mesophase will result in certain interactions being more favoured and others becoming less favourable, in comparison to an isotropic liquid.

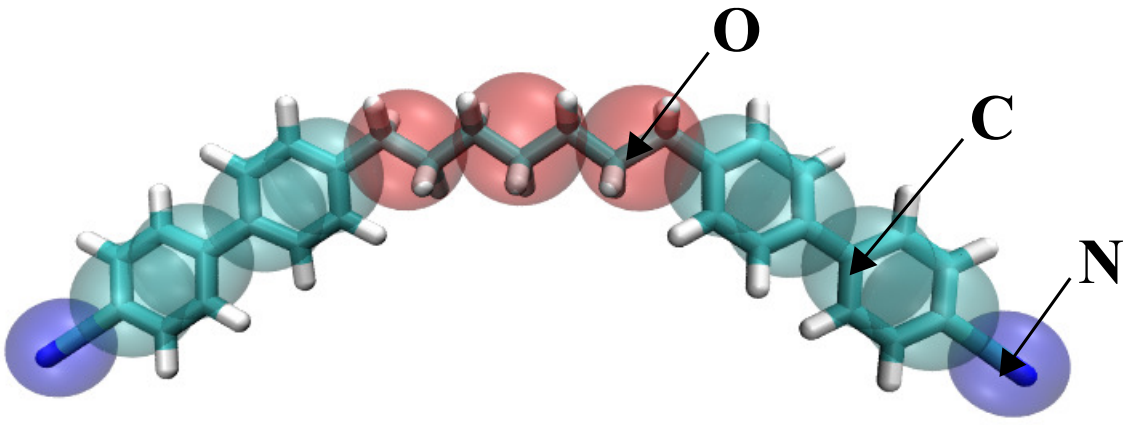


Figure 3.17: Coarse-grained mapping scheme for CB7CB.

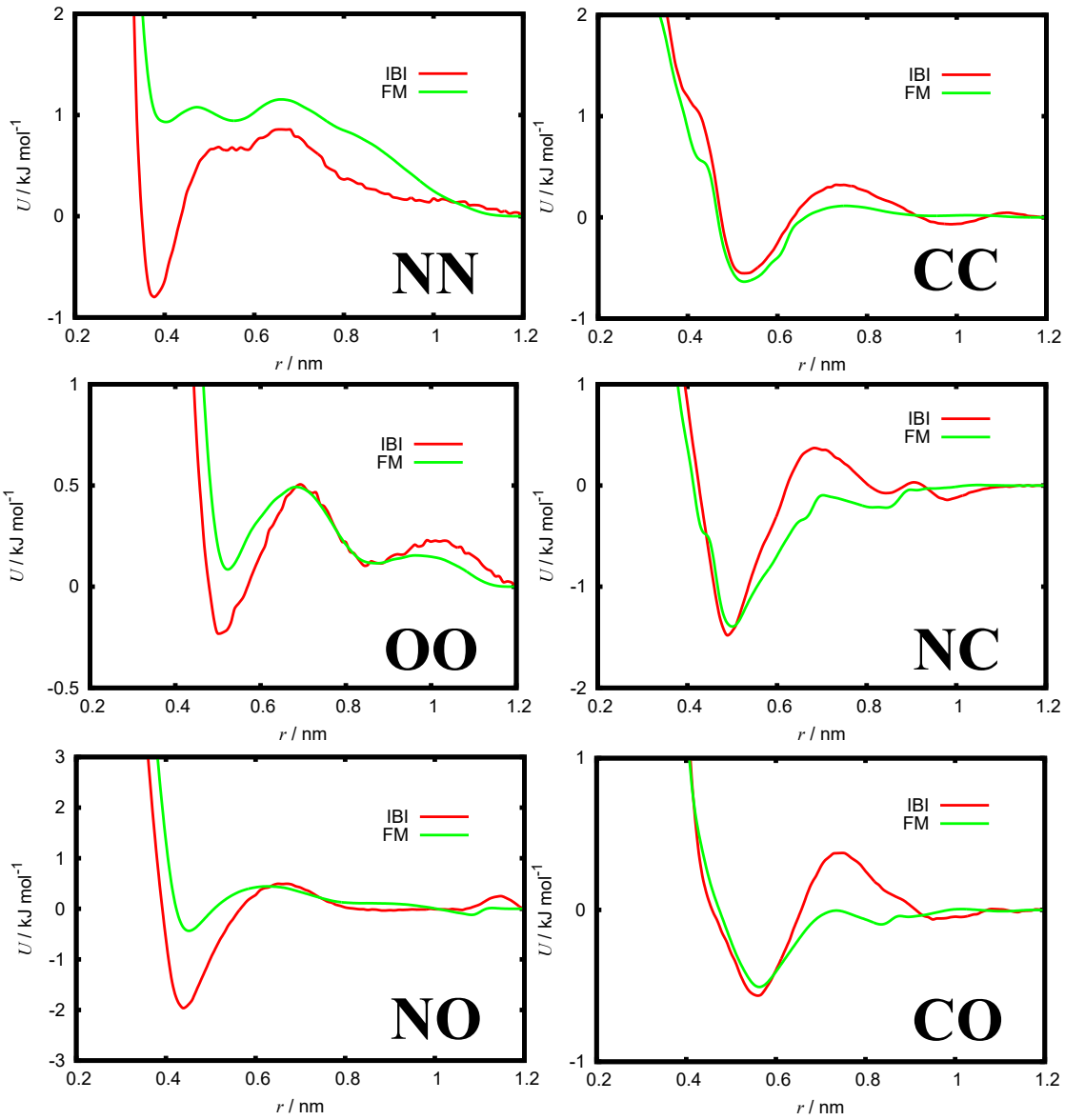


Figure 3.18: Coarse-grained potentials for the IBI and FM models of CB7CB.

The bonded parameters for the CG models are summarised in Tables 3.2, 3.3 and 3.4. These were employed for all CG models of CB7CB in harmonic potentials. Equilibrium values for the parameters were obtained from probability distributions calculated from the AA system of the N_{TB} phase at 370 K. The initial CG model (in single-molecule simulations) did not capture the correct bend angle and its χ distribution (which is closely linked to the conformational distribution) showed a strong bias towards right-handed conformers. Thus, the values for angles/dihedrals and their force constants were systematically varied until the χ distribution was reproduced and an average bend angle of 127° was captured (see Figure 3.20). It is noted that the resulting CG model has a full width at half maximum (FWHM) for the bend angle distribution of 15°, compared to 40° for the AA model, which

indicates a lower flexibility and a smaller tendency to sample a wide range of conformers beyond the equilibrium geometry. The stiffness of the bonds and angles for the mesogenic units also necessitates the use of a lower time step.

Table 3.2: Bond parameters for the CG models of CB7CB.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
N C	0.25	15000
C C	0.22	15000
C O	0.24	10000
O O	0.32	10000

Table 3.3: Angle parameters for the CG models of CB7CB.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
N C C′	180	5000
C C′ C″	180	5000
C C′ O	170	750
C O O′	170	750
O O′ O″	170	750

Table 3.4: Proper dihedral parameters for the CG models of CB7CB.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
C C′ O O′	180	25
C O O′ O″	180	25

Starting from a mapped configuration of the N_{TB} phase from the AA system at 370 K, MD simulations of 200 ns were performed for the three CG models of CB7CB. The IBI/FM models utilised the stochastic dynamics integrator as a thermostat (with a time constant of 2 ps) in the *NVT* ensemble and a time step of 2 fs. The Martini 3 model was simulated with the usual settings. Simulation snapshots are presented in Figure 3.19. From visual inspection, the FM and IBI CG models appear to exhibit the N_{TB} phase, where the mesogens form a heliconical structure. The Martini 3 model, on the other hand, forms a more crystalline phase.

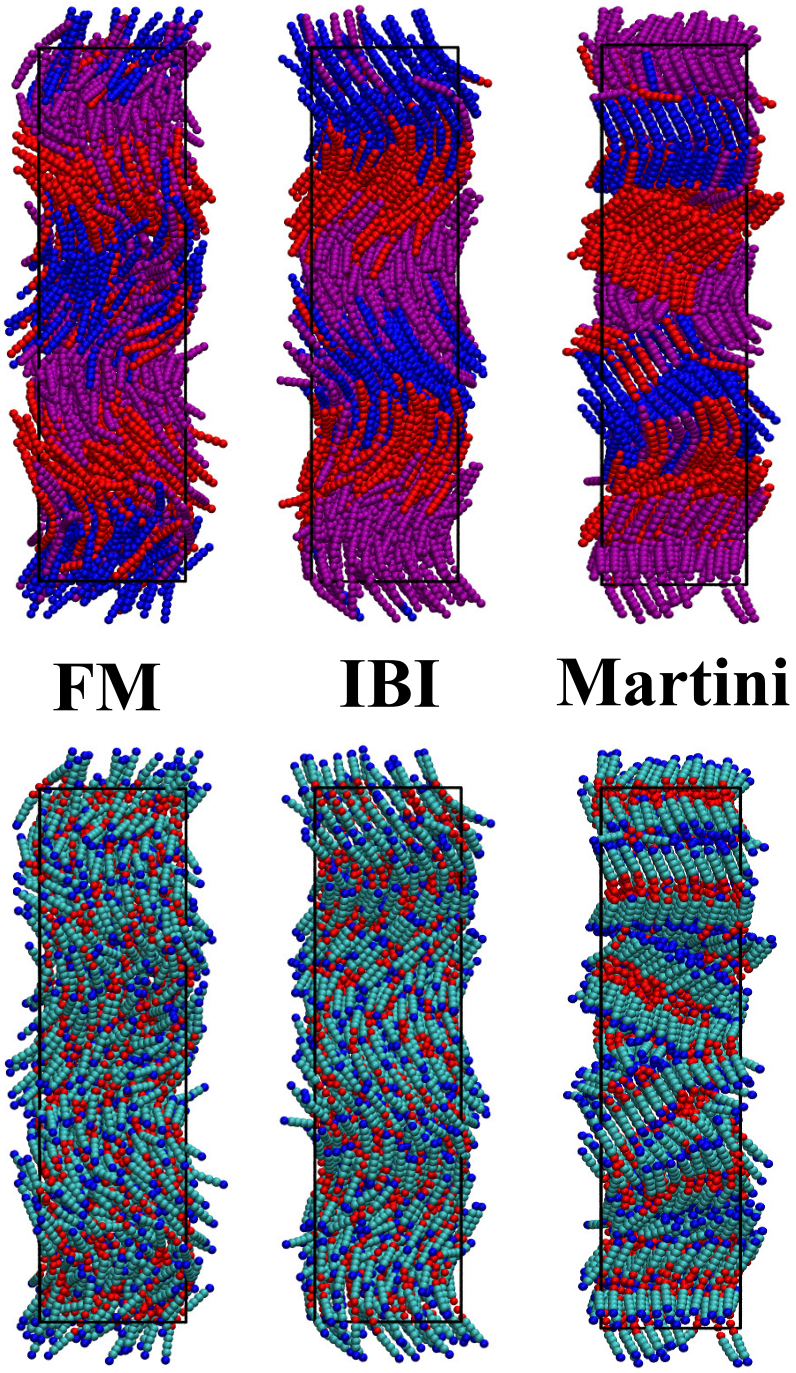


Figure 3.19: Simulation snapshots, with colouring by orientation (top) and bead type (bottom), for the FM, IBI and Martini 3 CG models of CB7CB.

Using the same analyses from the characterisation of the AA systems, the representability of the CG models can be compared quantitatively. The order parameter $\langle S_2 \rangle_a$ measures the nematic ordering of the molecular long axes $\hat{\boldsymbol{a}}$, whereas $\langle S_2 \rangle_b$ is a measure of the order between the short axes $\hat{\boldsymbol{b}}$ (i.e. 'polar' correlations or smectic-like ordering normal to the long axes). The helical pitch, p, and conical tilt angle, θ, are also determined. The various structural properties calculated are summarised in Table 3.5.

Table 3.5: The order parameters ($\langle S_2 \rangle_a$ and $\langle S_2 \rangle_b$), helical pitch (p) and conical tilt angle (θ) for the CG models of CB7CB at 370 K.

Model	$\langle S_2 \rangle_a$	$\langle S_2 \rangle_b$	p / nm	θ / $^\circ$
AA	0.36	0.16	8.35	29
FM	0.44	0.18	8.35	32
IBI	0.51	0.28	8.42	31
Martini	0.49	0.30	∼8.5	32

Firstly, the two bottom-up CG models will be compared. For both models, $\langle S_2 \rangle_a$ is higher than the AA value, with the FM model performing slightly better than the IBI model. It is likely that the lower flexibility of the CG models (in combination with the larger/fewer interaction sites) results in this increased ordering as there is less conformational disorder. For $\langle S_2 \rangle_b$, the value for the FM model is in very good agreement with the AA system and the IBI value is too high. This shows that the FM model produces a more nematic-like structure (i.e. low positional order) than the IBI model, which has higher local ordering between the mesogens. This effect can be explained by comparing the interaction potentials, where the IBI model has attractive interactions between all bead types which promote a degree of microphase separation. In particular, the OO interaction in the FM model is repulsive, but the corresponding potential in the IBI model has an interaction strength of 0.24 kJ mol^{-1}. This allows for the spacers (and the overall molecule) of adjacent molecules to align directly, in contrast to the preference in the FM model for an offset configuration. The pitch for both models is in agreement with the AA value, but this is partly pre-configured as the CG simulations are performed at constant volume and, thus, the helical structure must be commensurate (with the pitch being approximately half the length of the z-axis). Finally, the conical tilt angle is slightly higher than the AA model for both CG models and this could, again, be attributed to the narrower conformational distribution, which eliminates the presence of hairpin conformers and promotes better packing of molecules.

The Martini 3 model forms an almost crystalline phase with smectic layering observed within the system. In this CG model, the initial phase resolves into a structure with high order between the molecular short axes, forming layers where the director for each block precesses through the system in a helical fashion. The local ordering in this phase is akin to the segregation of mesogens into ferroelectric domains.[302–306] The phase organisation observed here is similar to the twist-bend smectic phase reported in dimers with a terminal

alkyl chain.[307–309] As expected, the order parameters for this model are higher than for the AA system. However, the pitch and conical tilt angle are only slightly higher than the target values. The use of the formal Martini 3 force field (where N = TN4a, C = TC5 and O = SC1) results in identical behaviour to the model shown here. The behaviour observed in the Martini 3 model is largely due to microphase separation. Within the Martini 3 framework, the like interactions are highly favourable, but the cross interactions are only equally or less favourable. In order to promote nematic ordering and destabilise smectic-like order, the cross-interactions must generally be equal to or greater in strength than the like interactions. For colloidal particles, the destabilisation of a smectic phase can also be achieved with polydispersity of particle sizes or curvature of the molecules.[275] In principle, this could be made possible here by using models with a wider conformational distribution (i.e. a greater polydispersity of conformations with various curvatures).

By employing lower force constants for the angle/dihedral potentials, the FWHM of the bend angle distribution can be increased. Here, it was found that a decrease to 125 kJ mol^{-1} deg^{-2} and 10 kJ mol^{-1} for the angle and dihedral force constants, respectively, of the relevant potentials involving the spacer results in a FWHM of 40°. MD simulations using these bonded parameters for the Martini 3 model still produces a smectic phase, although this phase is more liquid-like (as expected from the increased molecular flexibility). When these bonded parameters are applied to the FM and IBI models, the N_{TB} phase is lost and an isotropic phase is exhibited. It is also noted that the increased flexibility broadens the χ distribution such that it does not match that obtained for the AA system.

The role of the cyano group in these CG models is of special importance. In nematogens, it is found that the polar cyano group leads to dipolar correlations, which result in a preference for molecular association in an antiparallel or offset (parallel) fashion between adjacent molecules.[310–313] This assists in stabilising the nematic phase over a smectic phase. The NN interaction (which corresponds to the self-interaction of the cyano groups) for the FM model is completely repulsive. Additionally, the NC interaction is greater than the CC interaction. The combination of these potentials results in a preference for offset and antiparallel configurations between adjacent molecules in the FM model and promotes the nematic ordering observed. In contrast, the IBI model has an attractive NN interaction (with an interaction strength of 0.8 kJ mol^{-1}) which leads to a higher preference for direct parallel associations compared to the FM model. This is observed in the higher order pa-

rameters calculated in the IBI system. In the Martini 3 model, the interaction strength of the TC6–TC6 interaction is 1.7 kJ mol^{-1}. This contributes greatly in the dominance of a parallel arrangement of mesogens and results in the smectic ordering observed.

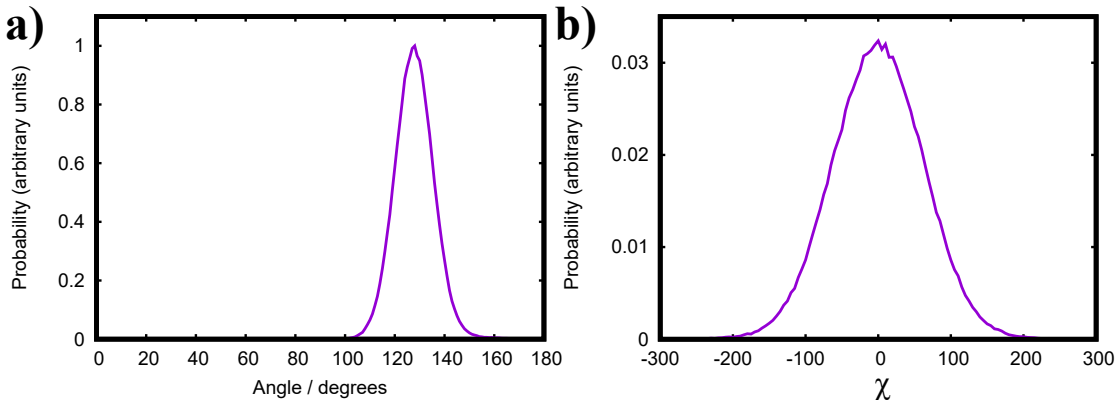


Figure 3.20: Distribution functions of the (a) bend angle and (b) chirality order parameter, χ, for the FM CG model of CB7CB.

As the FM model has been assessed to be the best of the three CG models for CB7CB, it was selected for a heating sequence to produce a full phase diagram. Here, its transferability to different temperatures is tested to see whether it can capture the N or I phases and to estimate the transition temperatures. Starting from the system at 370 K, the system was heated at intervals of 10 K for 200 ns at each temperature with a *NVT* simulation. The order parameters, $\langle S_2 \rangle_a$ and $\langle S_2 \rangle_b$, are determined for each point as per the AA phase diagram. The CG phase diagram for the FM model is presented in Figure 3.21 in which the stability range of the N_{TB}, N and I phases are indicated. Remarkably, it is found that the FM CG model successfully captures the phase transitions into the N and I phases. The temperature for the N_{TB}–N phase transition is ∼445 K, which is about 10 K higher than the AA value, and $T_{N-I} \approx 465$ K. The FM model demonstrates excellent transferability as the N and I phases are able to be observed in a CG model that was not parametrised at those thermodynamic state points, and the transition temperatures are in good agreement with the AA model. However, it is noted that the AA values are too high compared to experiment and this is carried forward into the bottom-up CG model. This suggests that improvements to this coarse-graining approach should first be aimed at obtaining atomistic force fields that correctly capture the phase transitions. The increase in $T_{N_{TB}-N}$ for the CG model likely arises from the lower flexibility as discussed previously. The fact that this transferability is displayed in the FM models indicates that the MS-CG method faithfully

captures the forces/interactions between the chemical groups of the underlying molecule. Unlike the AA model and in experiments, an increase in $\langle S_2 \rangle_a$ with temperature is not observed in the CG model. Once again, this probably arises from the minimal broadening of the conformational distribution with temperature and the fixing of the volume in these simulations. However, the decrease in $\langle S_2 \rangle_b$ with temperature (and its change to $\sim$0 in the N phase) is observed.

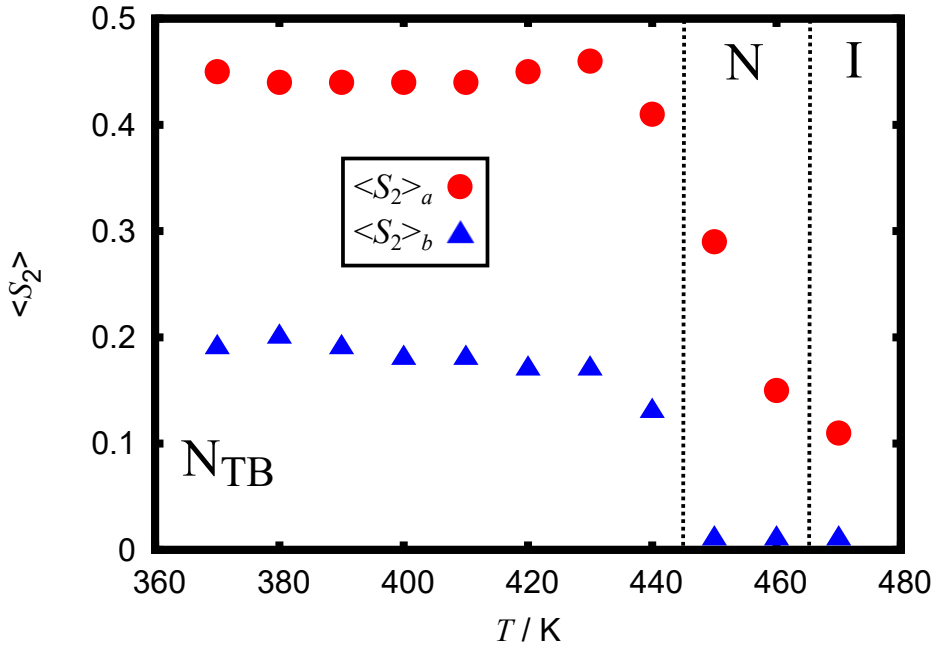


Figure 3.21: Average orientational order parameter, $\langle S_2 \rangle$, for unit vectors $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ as a function of temperature for the FM CG model upon heating.

A second challenge of the transferability of the FM CG model is to see whether the N and N_{TB} phases can be spontaneously grown from cooling the I phase. Starting from the I phase at 470 K, the system was gradually cooled at 20 K intervals (with 400 ns simulations at 450 K and 430 K, and a decrease to 200 ns for subsequent temperatures) to 370 K. The phase diagram from this cooling sequence is presented in Figure 3.22(d). The phase transition into the N phase and, subsequently, the N_{TB} phase upon cooling is observed with T_{I-N} and $T_{N-N_{TB}}$ being identical to values obtained from heating the system. It is likely that the ability to form the N phase, from an isotropic liquid, on these timescales arises from the accelerated exploration of phase space afforded to CG models. Simulation snapshots of the I, N and N_{TB} phases, obtained via cooling, are shown in Figure 3.22. This further demonstrates the transferability of this FM CG model for use across different temperatures. Moreover, characterisation of the simulated N_{TB} phase shows the high

representability of this CG model. Overall, it can be concluded that the FM CG model can inherently exhibit the N_{TB} phase and that it is not biased towards this phase due to the seeded configuration which the simulations were initialised from.

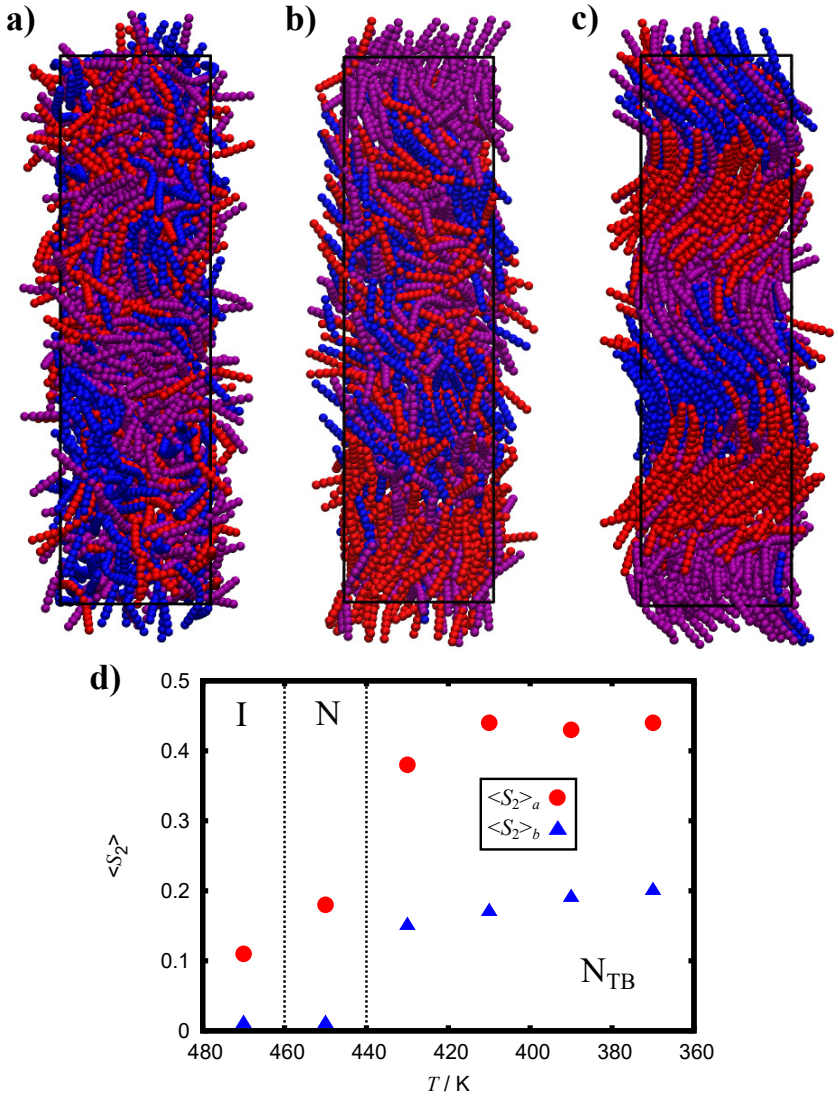


Figure 3.22: Simulation snapshots of the FM CG model, obtained via cooling, showing the initial (a) I phase at 470 K, the (b) N phase at 450 K and the (c) N_{TB} phase at 370 K. (d) Average orientational order parameter, $\langle S_2 \rangle$, for unit vectors $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ from the cooling of the I phase between 470–370 K.

As the motivation behind coarse-graining is to simulate systems beyond the length and time scales accessible to AA MD, the FM CG model was used for a larger simulation of 4096 molecules at 370 K. The AA system was replicated twice in each dimension and mapped to a CG resolution to act as the starting point. A 200 ns simulation was performed on this system. On the same computer architectures, the CG model demonstrates a 12x speed up compared to the AA model, and retains the structural accuracy of the underlying

system. Simulation snapshots and analyses are presented in Figure 3.23. The characterisation of the phase shows that the helical pitch and conical tilt angle (and calculated order parameters of $\langle S_2 \rangle_a = 0.44$ and $\langle S_2 \rangle_b = 0.18$) are in good agreement with the smaller 512-molecule system of the FM model and the AA reference.

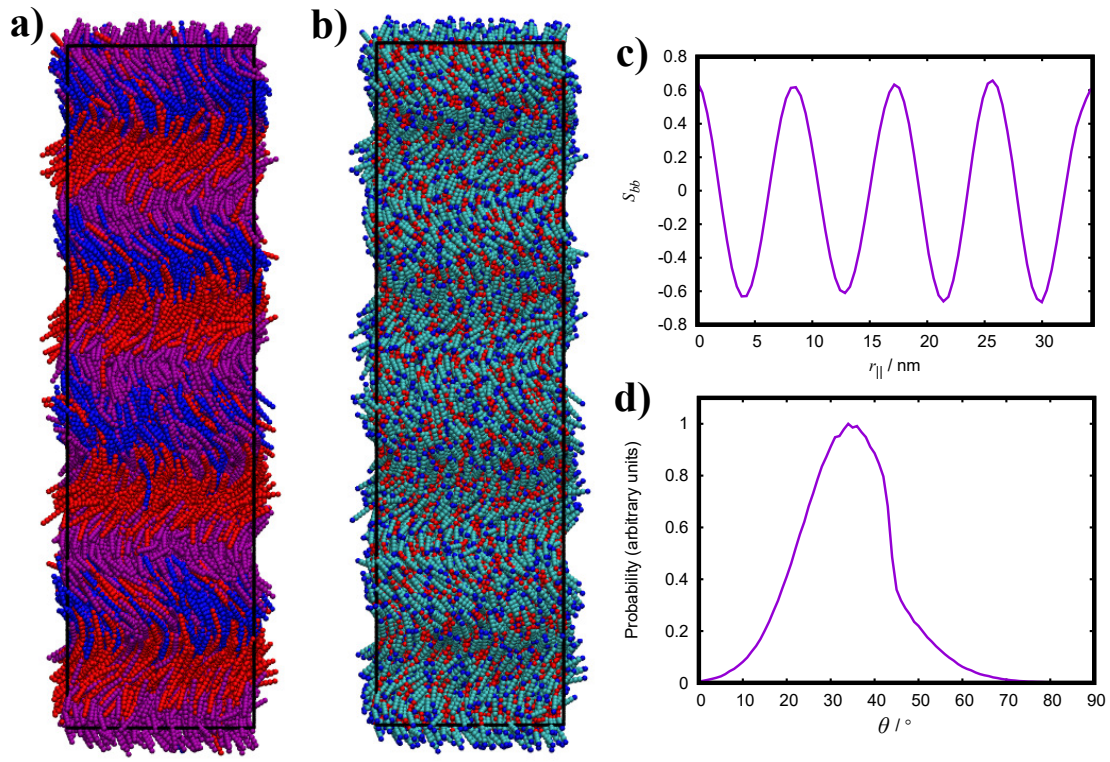


Figure 3.23: Simulation snapshots from a 4096-molecule system of the FM CG model, with (a) orientational and (b) bead type colourings. (c) Orientational correlation (S_{bb}) as a function of the helical axis, $r_{||}$, and the (d) distribution for the conical tilt angle, θ.

3.4 Summary

In conclusion, this chapter presented results on the coarse-graining of benzonitrile-heptane systems, all-atom simulations of bent liquid crystal dimers and coarse-grained models of CB7CB. The IBI, MS-CG and Martini 3 coarse-graining approaches were applied to pure liquids of PhCN and heptane, and their mixture. The resulting CG models were assessed in terms of their structural accuracy and thermodynamic properties. It was found that the bottom-up IBI and FM models reproduce the AA RDFs very well, but severely underestimate ΔH_{vap} and ΔG_{solv}. In contrast, the Martini 3 models perform slightly worse in their structural accuracy, but better for thermodynamic properties compared to the bottom-up models. It is suggested that a viable route to developing highly representative and trans-

ferable CG models is to optimise Martini 3 parameters to reproduce target observables.

Secondly, extensive AA MD simulations are presented for bent liquid crystal dimers with a particular focus on CB7CB. The N_{TB} phase is observed for CB7CB which has structural features that are in good agreement with experiment. A full phase diagram was obtained by heating the initial system and the phase transitions into the N and I phases are captured. Distributions for the bend angle between the mesogenic units and the chirality order parameter, χ, were calculated. It is found that CB7CB is statistically achiral, but can adopt chiral conformers with no preference for a specific handedness. Furthermore, there is no change in the extent of conformational chirality between the molecules in the N_{TB} and N phases. Using single-molecule simulations in the gas phase, the investigation was expanded to the dimer series CBnCB (where n = 6, 7, 8 or 9) and CB$X(CH_2)_5Y$CB (where X/Y = CH_2, O or S). It is confirmed that the ability for a dimer to exhibit the N_{TB} phase is dependent on the bent molecular shape and not its potential to assume chiral conformers. For CB$X(CH_2)_5Y$CB, it is found that $|\chi|_{\max}$ increases as the bend angle decreases, but the flexibility of the dimer remains largely unchanged.

Finally, systematic approaches to coarse-graining CB7CB are explored. The IBI and MS-CG methods were utilised to parametrise bottom-up CG models based on the AA system of the N_{TB} phase. Additionally, a top-down Martini 3 model was constructed. The IBI and FM models appear to exhibit the N_{TB} phase, whereas the Martini 3 model forms a more crystalline phase. The origins of these differences is discussed in terms of the balance between the interactions present, particularly between the cyano groups, where the IBI and Martini models show microphase separation. Upon heating, the FM CG model captures the phase transitions into the N and I phases and, also, the spontaneous growth of the N and N_{TB} phases from cooling of the I phase. This demonstrates an excellent transferability of this CG model across different temperatures which, coupled with the good representability of the N_{TB} phase, proves the suitability of this approach for thermotropic liquid crystals.

Chapter 4

Atomistic Simulations of Ionic Cyanine Dyes

4.1 Introduction

The self-assembly of cyanine dyes into supramolecular aggregates is recognised to be an auspicious approach to construct functional materials, where molecular assemblies spontaneously form in various environments.[314] These dyes are known to form large-scale aggregates in aqueous solution which exhibit a blue- or red-shift of the monomer absorption band, attributed to H- and J-aggregation, respectively.[51–53,56–58] Such systems are characterised by the intermolecular association of aromatic cores into stacked structures, where the direction of the spectral shift is determined by the angle between adjacent molecular planes.[54,55] Despite the small size of the mesogenic cores, the cationic aromatic systems have high polarisability due to the delocalisation of π-electron density across the polymethine chain. This results in strong attractive interactions governing association and the cationic charge affords good solubility to these compounds. The unique properties of cyanine J-aggregates can be exploited in devices as spectral sensitisers and synthetic light harvesting systems[315,316], owing to the ease of alignment and adsorption of the dyes to surfaces.[317–319] Further applications such as electronic energy transport wires[320], fast all-optical switches[321,322] and as fluorophores for biological imaging[323,324] have been reported. The morphologies of supramolecular structures formed by cyanine dyes have been investigated using electron microscopy techniques, revealing the common formation of tubular architectures which can form networks of fibres.[135,325–329] Nevertheless, the molecular level

arrangements within these aggregates remain a question for which there is no unequivocal answer.

In this study, four cyanine dyes (see Figure 4.1) have been investigated: pseudoisocyanine chloride (PIC)[56,57], pinacyanol chloride (PCYN)[328,330–334], 5,5′,6,6′-tetrachloro-1,1′,3,3′-tetraethylbenzimidazolylcarbocyanine chloride (TTBC)[135,335,336] and 1,1′-disulfopropyl-3,3′-diethyl-5,5′,6,6′-tetrachloro-benzimidazolylcarbocyanine sodium salt (BIC)[329,337–340]. These can be categorised into two classes. PIC and PCYN are almost identical, only differing by an ethylene unit separating the quinoline groups. TTBC and BIC share a common mesogenic core; however, BIC has additional mesylate groups on two of the peripheral alkyl chains which render the overall molecule anionic.

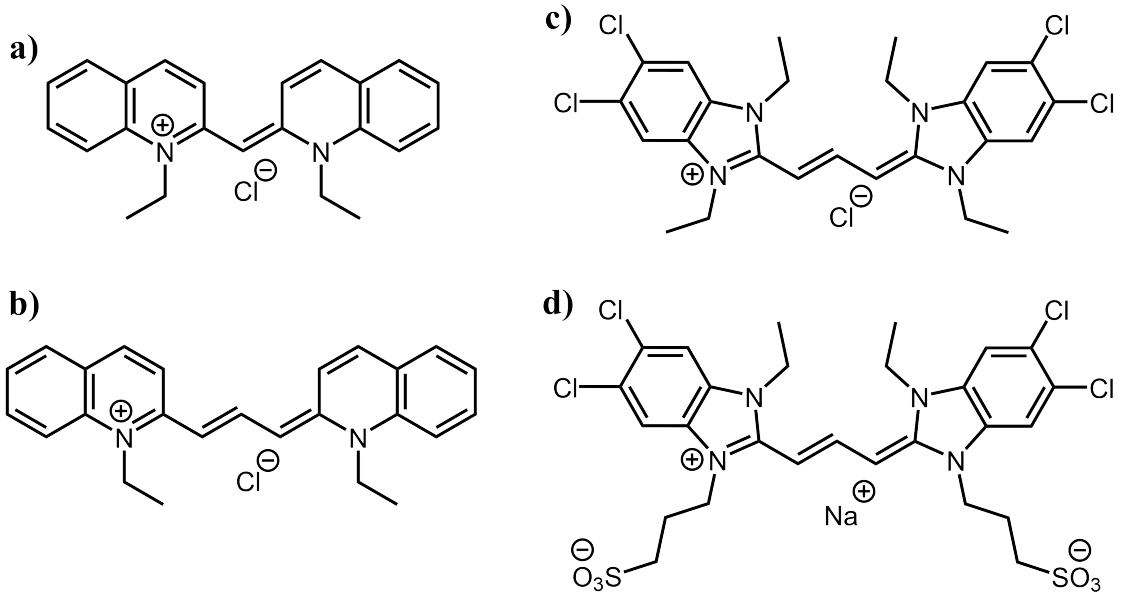


Figure 4.1: Molecular structures of the cyanine dyes: (a) PIC, (b) PCYN, (c) TTBC and (d) BIC with their counterions.

PIC, as the first J-aggregate to be discovered, has been extensively studied. In its absorption spectrum, a J-band is clearly exhibited along with what is suggested to be a broad H-band overlapping with the monomeric vibronic progression.[341,342] The presence of blue-shifted bands, in the initial stages of aggregation, which diminish as the J-band emerges is observed.[343,344] This is concurrent with the onset of fibres, where individual fibres are estimated to have an average width of 2.3–2.89 nm.[325,326] This corresponds to an average aggregation number of $\sim$3000 based on geometric packing models coupled with spectroscopic arguments. In one of the earliest models for the molecular arrangement, Scheibe and Kandler proposed a pile-of-coins structure with the long axis of the monomers ly-

ing perpendicular to the aggregate axis[345], but this model is not consistent with the displacement of transition dipole moments theorised by Kasha.[55] Further models of offset monomers with the molecular long axes aligned parallel to the aggregate axis have been suggested;[346] for example, a brickwork arrangement[347], threaded double-string models[348] and rods composed of single strands packed in a herringbone arrangement[325].

PCYN is found to form tubular aggregates with the absorption spectrum showing a pronounced H-band.[328] These aggregates are observed to transform into J-aggregates over the timescale of weeks. It is proposed that these nanotubes are single-walled with the dyes aligned parallel to the tube axis. Chromonic phases formed from these nanotubes have been reported for the acetate salt.[332]

TTBC exhibits an assortment of optical spectra and aggregate structures, with three main types observed. The chloride salt exhibits an absorption spectrum composed of both a strong H-band and a weak J-band (type I spectrum), and cryo-TEM reveals fibrous networks of tubes with unimolecular thickness.[335] This type of spectrum is, generally, attributed to Davydov splitting and is expected to have two molecules in the unit cell of the crystal structure.[349] Initially, a 2D herringbone model was proposed for this type of J-aggregate[336], but it has been suggested that a rolled up herringbone arrangement is more probable[335]. Two crystal structures for TTBC solvates have been reported: a herringbone arrangement with two molecules per unit cell (DYEM) and a brickwork arrangement with one molecule per unit cell (DYEA).[350] The brickwork model for DYEA has been observed for air-dried TTBC aggregates on a mica surface, which exhibit an absorption spectrum of a weak J-band (type II spectrum).[335] The iodide salt displays both a type I spectrum and type III spectrum (J-band with a monomer band), depending on the solvent pH.[335] The molecular arrangement in type III structures is indicated to contain linearly-stacked dye molecules in a tubular shape with the molecular long axes aligned parallel to the tube axis. The morphology of BIC J-aggregates is revealed to be mainly spherical with a particle diameter of 20–100 nm.[329,339]

In this chapter, atomistic molecular dynamics simulations have been applied to four ionic cyanine dyes. Similar studies have previously been conducted.[49,133–135] This work, in particular, focuses on elucidating the finer details of the aggregation of these dyes by studying the thermodynamics of association and the self-assembly in a low concentration regime. Such results provide insight into these systems and can serve as data for validation

in the development of coarse-grained models. The work presented herein is published in the paper[50] "*Atomistic simulation studies of ionic cyanine dyes: self-assembly and aggregate formation in aqueous solution*." It is reproduced from Ref. 50 [G. Yu, M. Walker and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2021, **23**, 6408–6421] with permission from the PCCP Owner Societies.

4.2 Computational details

Force fields

The General Amber Force Field (GAFF) was selected as the force field in this work (see section 2.1).[220] The Antechamber package from AmberTools18 was used to produce appropriate parameters for the force field for a given chemical structure.[221] The generated topologies were converted into the necessary input files for GROMACS using the ACPYPE script.[223] All simulations used the TIP3P water model[351], unless otherwise stated, which is compatible in combination with GAFF.[352]

For the cyanine dyes, GAFF was modified in two ways. Firstly, the dihedral potentials for the polymethine chain were optimised to better capture their torsional profiles, according to the methodology employed by Boyd and Wilson.[271,284] The data for fitting was obtained using density functional theory (DFT) by incrementing the selected dihedral by 6° over a range of 180°, optimising the geometry and calculating the potential energy at each step. The resulting DFT potential was fitted to a Ryckaert–Bellemans (RB) function with the determination of new coefficients via minimisation of a squared difference, χ^2. The second modification replaced the atomic charges with ones obtained using the CHELPG method, which fits charges to reproduce the electrostatic potential around the molecule.[353] The data used to implement the two modifications was acquired from calculations using the B3LYP functional and the 6-31+G* basis set, and performed in Gaussian 09.[354]

Simulation details

All MD simulations were carried out in GROMACS 2018.2.[226] Initial systems were set up by randomly inserting a number of dye molecules in a box and then solvated with

water and counterions to the desired concentration. Long-range electrostatic interactions were handled by the PME method[146], with a cutoff of 1.2 nm for both electrostatic and LJ interactions. After a steepest descent energy minimisation, systems were pre-equilibrated in the *NVT* ensemble using the Berendsen thermostat for 100 ps before a pre-equilibration in the *NPT* ensemble for 200–500 ps in conjunction with the Berendsen barostat.[150] A final equilibration for 100 ps and subsequent production runs employed the Nosé–Hoover thermostat[152,153] to keep the temperature constant at 300 K, and the Parrinello–Rahman barostat[154] to maintain a pressure of 1 bar. Time constants of 1 ps and 5 ps were used for the thermostat and barostat, respectively. Using a leap-frog integrator for MD, a time step of 1 fs for equilibration was used, with an increase to 2 fs for production runs, where constraints were applied to all bonds using the LINCS algorithm.[144] Production simulations were typically performed for 500 ns.

Free energy calculations

Potentials of mean force (PMFs) were calculated for n-mers (where n = 2, 3 or 4) at a concentration of 1 wt% and at various temperatures. This dilute concentration is required in order to accommodate the pulling apart of the initial configuration while avoiding self-interaction due to periodic boundary conditions. A pull was applied between the centres of mass (COMs) of each species at a pull rate of 0.001 nm ps^{-1}. Where $n > 2$, the pull was applied between the COM of a monomer and the COM of the remaining molecules (stack), where adjacent pairs within the stack were constrained at the favoured distance for a dimer. This results in a shift of the PMFs compared to the dimer and ensures that the stack remains intact throughout the simulations. Configurations of specific distance were extracted from the pull with neighbouring points varying from 0.02–0.1 nm between windows. Each window was equilibrated for 1 ns and then simulated for 20 ns to sample a range of configurations with the intermolecular distances of the COMs constrained. A total of 1×10^6 force values were output for each point of the separation distance and used to calculate the average constraint force before integration to obtain the PMF. The free energy of association, $\Delta G_{\mathrm{assoc}}$, is defined as the maximum well depth of the PMF.

Free energies of hydration, ΔG_{hydr}, were calculated by decoupling the intermolecular interactions of a single molecule from the solvent at a concentration of 0.5 wt%, where the hydration of the species is the reverse of this process. Coulombic interactions were first

decoupled linearly before the van der Waals interactions, with a total number of 41 states of λ between 0 and 1, with a spacing of 0.05. Simulations for each λ state consisted of a 1 ns equilibration followed by a 5 ns production run for data collection using the stochastic dynamics integrator.

4.3 Results and discussion

4.3.1 Force field optimisation

The automatically generated force field allowed for free rotation about the central polymethine bridge for all four dyes. However, it has been shown that PIC exhibits a twist between the quinoline units within the mesogenic core with an angle of $\sim$27° away from planarity.[133] Therefore, a dihedral optimisation procedure was applied to improve representation of key dihedrals within the molecules. For PIC, there is one dihedral to optimise whereas PCYN has two dihedrals for improvement. Both TTBC and BIC share a common mesogenic core; therefore, dihedral potential potentials were calculated for TTBC with the assumption that any parameters would be transferable. Figure 4.2 shows the calculated dihedral potentials and their fitted Ryckaert–Bellemans potentials.

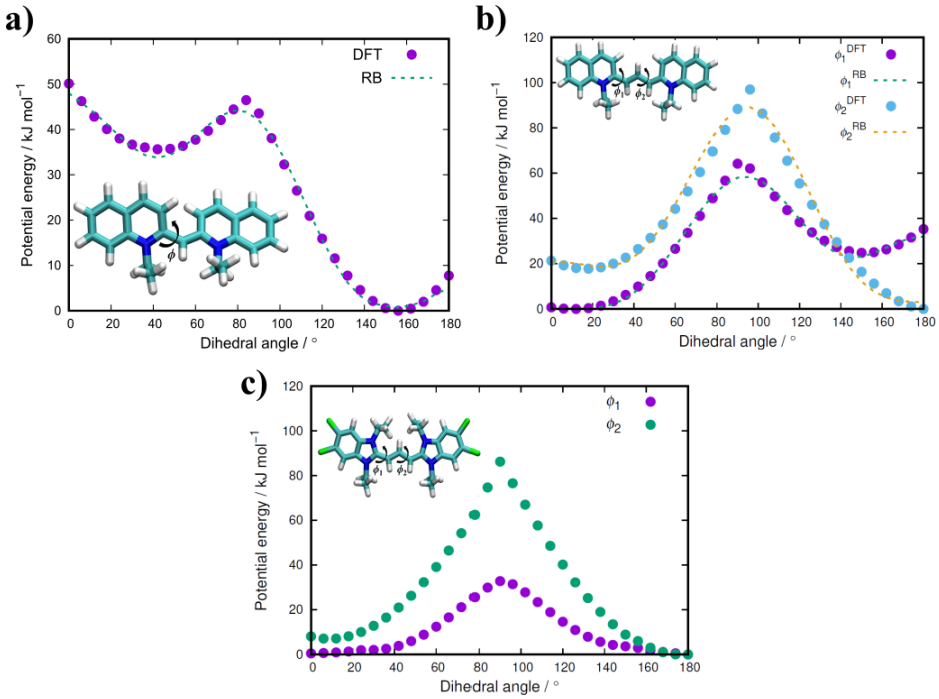


Figure 4.2: Calculated potentials (DFT, B3LYP/6-31+G*) for the key dihedral(s) with their fitted Ryckaerts–Bellemans (RB) potentials in (a) PIC and (b) PCYN. (c) Dihedral potentials (DFT, B3LYP/6-31+G*) for TTBC.

The dihedral potential for PIC has two minima: a global minimum at 156° and a local minimum at 45°. This results in twisted conformations between the planar rings. For PCYN, the corresponding dihedral has a global minimum at 12° and a local minimum at 150°, and the second dihedral (ϕ_2) has a global minimum at 180° and a local minimum at 18°. This manifests as a twist between the rings, but of a smaller magnitude than PIC. The dihedral potentials calculated for TTBC have global minima at 180° with reasonably high energy barriers present. Thus, the original RB potentials were used for these dihedrals (which have minima at 0/180°) with modifications made only to the magnitude of the barrier height. As a result, TTBC/BIC mesogens are planar with respect to the aromatic rings.

Table 4.1: Optimised Ryckaert–Bellemans coefficients, C_n (kJ mol^{-1}), obtained from fitting to DFT data for PIC.

Dihedral	C_0	C_1	C_2	C_3	C_4	C_5
ϕ	42	-26	-74	16	57	-11

Table 4.2: Optimised Ryckaert–Bellemans coefficients, C_n (kJ mol^{-1}), obtained from fitting to DFT data for PCYN.

Dihedral	C_0	C_1	C_2	C_3	C_4	C_5
ϕ_1	58	16	-116	-21	74	21
ϕ_2	88	28	-158	-60	82	23

The initial atomic charges obtained via the AM1-BCC method[222] localised the positive charge on a single aromatic unit as opposed to its delocalisation across the π system. Thus, the atomic charges were recalculated using the CHELPG method proceeding a geometry optimisation by DFT. The resultant atomic charges produced a symmetric charge distribution for the molecule, where corresponding atoms on each aromatic system are identical within 2 significant figures.

4.3.2 Thermodynamic analysis

The free energy of association can be evaluated from a PMF describing the pulling apart of a species. Figure 4.3 shows PMFs for a dimer of PIC solvated in TIP3P, TIP4P[355] and

TIP5P[356] water. These models differ in the placement of partial charges on zero, one or two dummy atoms, respectively, attached to the oxygen centre. In the PMF profiles obtained, their general form is invariant to the water model used but differ in the depth of the minimum. The global minimum is at an intermolecular COM distance of 0.43 nm with ΔG_{assoc} values of 8.0 $k_{\text{B}}T$, 9.3 $k_{\text{B}}T$ and 11.2 $k_{\text{B}}T$ for TIP3P, TIP4P and TIP5P, respectively. The increase in ΔG_{assoc} across the water models could be a result of a larger hydrophobic effect exhibited by the more ordered water structure in TIP5P compared to the TIP3P model.[357] A comparison of rigid non-polarisable water models has been reported where it is found that TIP4P performs better than TIP5P and TIP3P.[358] In all three combinations, the lowest energy configurations for the dimer correspond to H-aggregation of the molecules. This configuration provides maximum overlap of the cores and is thermodynamically favourable. Moreover, each PMF has a plateau after the global minimum centred at ~ 0.8 nm. Configurations corresponding to this region represent J-aggregation and are characterised by overlap of only one of the quinoline units for each molecule.

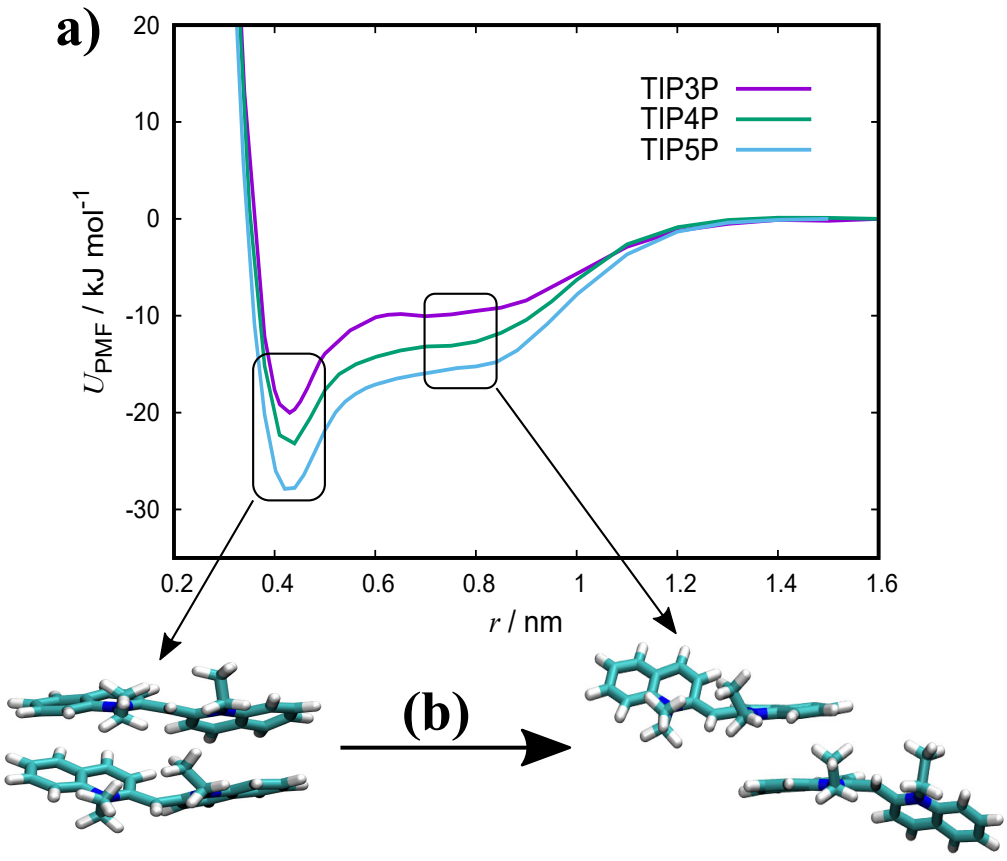


Figure 4.3: (a) PMFs for a dimer of PIC solvated in three different water models at 300 K with snapshots of structures (b) along the reaction coordinate.

Table 4.3: Free energies of association ($\Delta G_{\mathrm{assoc}}$) and the favoured intermolecular COM distance of a dimer, trimer and tetramer (r_{assoc}) for the cyanine dyes.

	Molecule	$\Delta G_{\mathrm{assoc}}$ / $k_{\mathrm{B}}T$	r_{assoc} / nm
Dimer	PIC	8.0 ± 0.6	0.43
	PCYN	13.9 ± 0.3	0.42
	TTBC	14.4 ± 0.4	0.42
	BIC	15.3 ± 0.3	0.45
Trimer	PIC	6.1 ± 0.4	0.65
	PCYN	11.2 ± 0.1	0.59
	TTBC	12.7 ± 0.8	0.60
	BIC	10.3 ± 0.6	0.69
Tetramer	PIC	5.5 ± 0.5	0.84
	PCYN	11.7 ± 0.2	0.76
	TTBC	10.6 ± 0.2	0.79
	BIC	9.1 ± 0.3	0.94

The free energies of association, $\Delta G_{\mathrm{assoc}}$, evaluated from the PMFs of the four cyanine dyes are summarised in table 4.3 and their PMF profiles are shown in Figure 4.4. The error on $\Delta G_{\mathrm{assoc}}$ is estimated by propagation of the errors for each data point in the PMF. It is observed that the $\Delta G_{\mathrm{assoc}}$ values for dimerisation determined for all four cyanine dyes are in the range of 8 to 15 $k_{\mathrm{B}}T$ which is typical of values calculated for chromonic systems. Reported values[359–361] of $\sim$5.3 $k_{\mathrm{B}}T$, from experimental measurements for the dimerisation constant of PIC, compare favourably with the results; however, the trimer/tetramer values for $\Delta G_{\mathrm{assoc}}$ are in even closer agreement. Similarly, experimentally obtained binding energies for the dimer of PCYN[362], 8.3 $k_{\mathrm{B}}T$, and its acetate salt[363], 10.3 $k_{\mathrm{B}}T$, are in good agreement, but the experimental values are closer to the trimer/tetramer values.

In PIC and PCYN, it is observed that the trimer/tetramer binding energies are lower than that for the dimer. The self-assembly in chromonic systems is generally considered isodesmic, where the addition of a new molecule to a stack provides more or less the same free energy increment, independent of the size of the aggregate.[39–41] It is found that the association here is "quasi-isodesmic", a property observed in simulations of chromonics previously.[47,48,136] This effect is indicated to be entropic in origin; the molecules in the middle of a stack experience a restriction in orientational freedom and, thus, dimers are slightly preferred over small aggregates.[48]

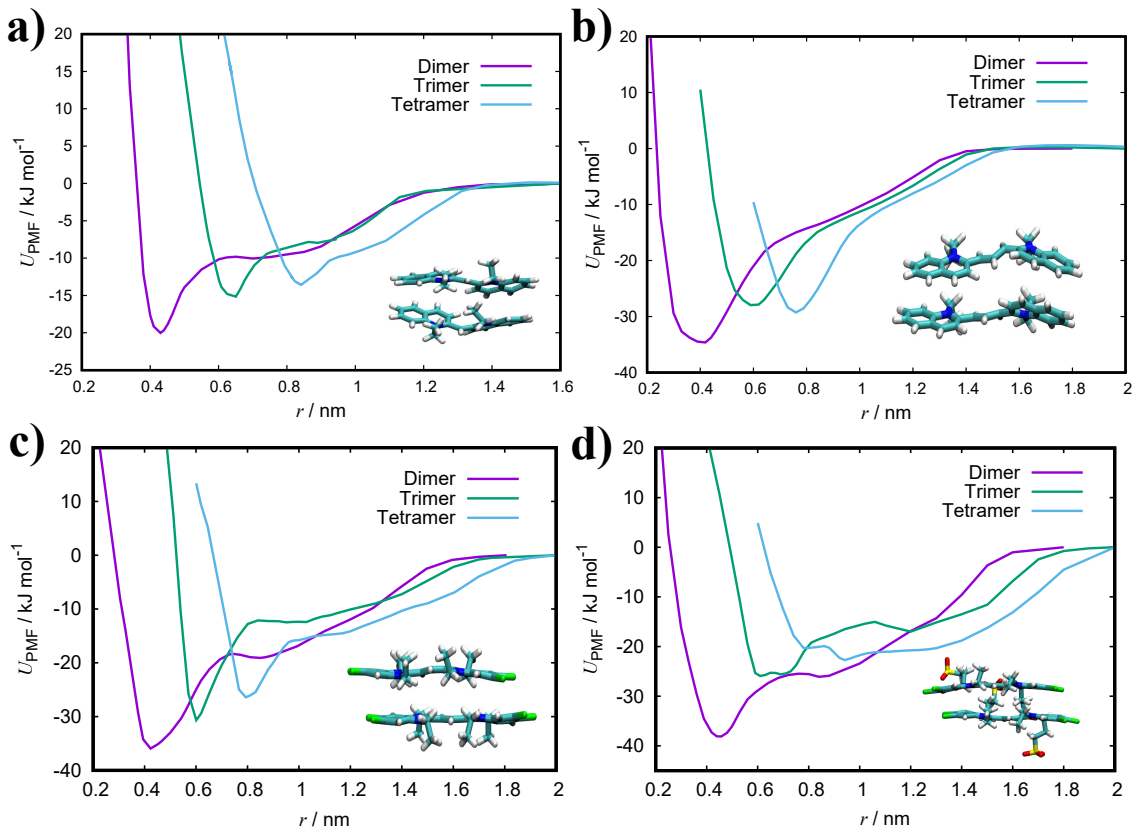


Figure 4.4: PMFs for a dimer, trimer and tetramer of (a) PIC, (b) PCYN, (c) TTBC and (d) BIC in TIP3P water at 300K, with example structures of dimers exhibited in the minimum of the PMFs.

The free energies of association calculated for PCYN are higher than that for PIC. This could arise from the differences in the twist between the aromatic halves of each species. The longer central bridge in PCYN allows the aromatic rings to adopt a small twist of $12°$, compared to the greater twist in PIC of $24°$. This greater divergence from planarity affects the overlap of the molecules in the dimeric state and leads to a lower binding constant.[336] Moreover, it is found that the tendency to form aggregates increases with the size of the core and length of the polymethine segment.[314,327,364]

TTBC and BIC share a mesogenic core and so it is expected that their binding energies would be similar, with any differences arising from the additional anionic groups on BIC. The ΔG_{assoc} values obtained are slightly greater than those for PIC/PCYN and the dimer structures exhibited by both TTBC and BIC are subtly offset so as to minimise overlap of the peripheral chlorine groups. With regards to "quasi-isodesmic" behaviour, a slight decrease in binding energy is observed in the trimer and tetramer for both dyes. However, this seems to be more of a systematic decrease in binding energies going from trimerisation to tetramerisation and is particularly apparent in TTBC. These results for TTBC and BIC

indicate a possible shift away from H-aggregation, or from the configurations observed in the PMFs, as the number of molecules in the aggregate increases.

The free energy of association can be decomposed into its enthalpic and entropic contributions to gain further insight into the driving forces of the aggregation. Here, the binding energies, ΔG_{assoc}, of a dimer and trimer of PIC and PCYN were calculated at five temperatures in the range of 280–320 K. Using

$$\frac{\Delta G_{\text{assoc}}}{k_{\text{B}}T} = \frac{\Delta H_{\text{assoc}}}{k_{\text{B}}T} - \frac{\Delta S_{\text{assoc}}}{k_{\text{B}}}, \tag{4.1}$$

plots of $(\Delta G_{\text{assoc}} \,/\, k_{\text{B}}T)$ vs $(1 \,/\, T)$ were fitted via linear regression and the association enthalpy, ΔH_{assoc}, and entropy, $T\Delta S_{\text{assoc}}$, contributions to ΔG_{assoc} estimated from the gradient and intercept, respectively. The resulting linear plots are shown in Figure 4.5. The enthalpic and entropic contributions to ΔG_{assoc} are summarised in Table 4.4, with errors determined by propagation of the errors for each ΔG_{assoc} value and the fit itself. It is noted that small changes in the gradient of the fit can strongly affect the results obtained. However, they still provide for a good semi-quantitative comparison between the systems studied here.

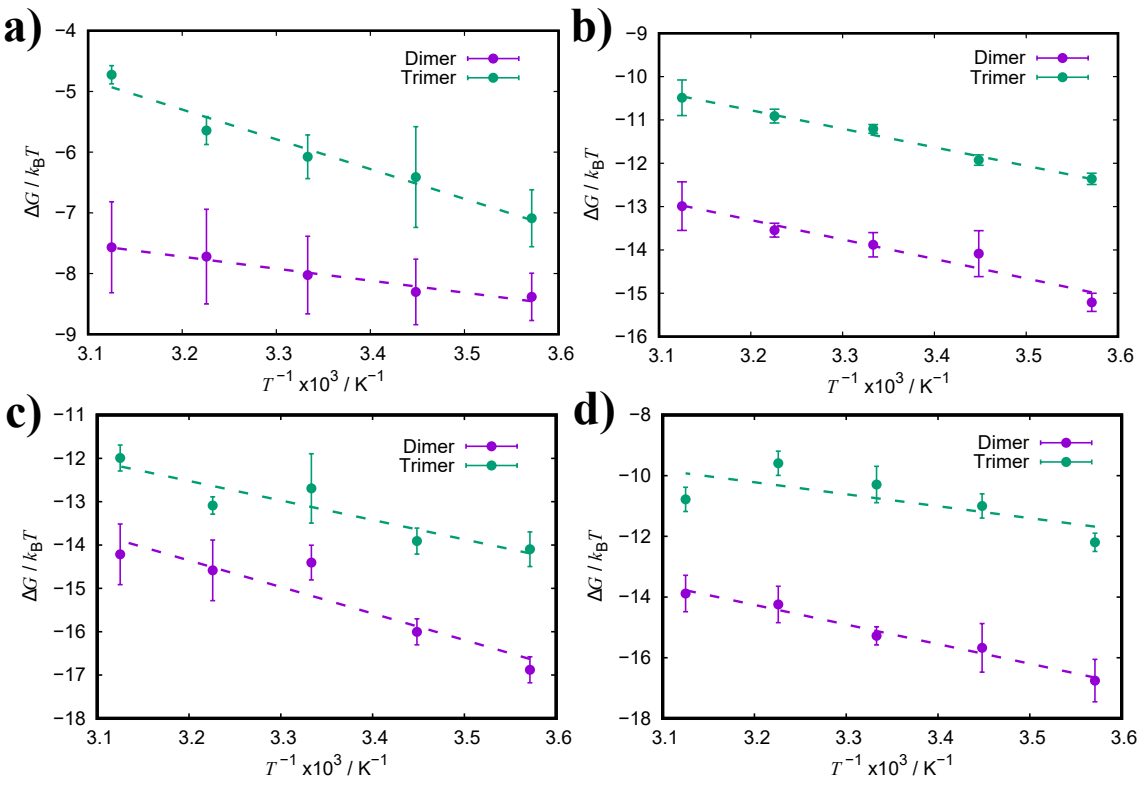


Figure 4.5: Free energy of association, ΔG_{assoc}, for a dimer and trimer of (a) PIC, (b) PCYN, (c) TTBC and (d) BIC as a function of temperature. Dashed lines represent the fitted linear functions.

Table 4.4: Association enthalpy, $\Delta H_{\mathrm{assoc}}$, and entropy contributions, $-T\Delta S_{\mathrm{assoc}}$ (where T = 300 K), for dimerisation and trimerisation.

	Molecule	$\Delta H_{\mathrm{assoc}}$ / kJ mol^{-1}	$-T\Delta S_{\mathrm{assoc}}$ / kJ mol^{-1}
Dimer	PIC	-17 $\pm$ 3	-3.5 $\pm$ 1.7
	PCYN	-37 $\pm$ 6	2.6 $\pm$ 0.9
	TTBC	-51 $\pm$ 12	13 $\pm$ 9
	BIC	-54 $\pm$ 5	16 $\pm$ 5
Trimer	PIC	-40 $\pm$ 8	26 $\pm$ 5
	PCYN	-36 $\pm$ 2	7 $\pm$ 2
	TTBC	-37 $\pm$ 9	5 $\pm$ 2
	BIC	-33 $\pm$ 10	6 $\pm$ 4

It is found that the enthalpic term is much higher than the entropic contribution in the dimerisation of PIC and PCYN. This indicates that the association of monomers is driven by favourable attractive forces rather than a strong hydrophobic effect. In contrast, it has been shown that the aggregation of the non-ionic chromonic molecule, TP6EO2M, has a larger entropic contribution, which is up to 1.5 times larger than the enthalpic term.[42,48] This suggests that the ionic nature of these cyanine dyes shifts the driving forces of association. Simulations of an anionic perylene bisimide dye showed that the enthalpic component was also greater than the entropic term, and the latter was unfavourable with respect to association.[131]

For PCYN, the entropic component to $\Delta G_{\mathrm{assoc}}$ is small and unfavourable for both dimerisation and trimerisation, while the enthalpic contribution is very similar for both processes. This indicates that the "quasi-isodesmic" aggregation observed for PCYN arises from the entropic change between dimer and trimer association. Steady-state absorption spectroscopy experiments provide approximate values for the enthalpy and entropy changes for the dimerisation of PCYN of -28 $\pm$ 3 kJ mol^{-1} and -0.1 $\pm$ 0.1 kJ mol^{-1}, respectively.[331] The predicted association enthalpy is similar in magnitude compared to our result, but the entropic contribution we obtain is larger. PIC, on the other hand, shows an entropic contribution that is strongly unfavourable for trimerisation. However, this system shows compensation between the terms with the enthalpic component increasing for trimers. The dimer-trimer difference likely arises from the higher orientational freedom of the dimeric species in solution compared to trimeric state.

TTBC and BIC display similar values for the enthalpic and entropic components of

the binding energy. Their association enthalpies are greater than the entropic contribution for dimerisation. Interestingly, it seems the presence of the additional ionic groups in BIC does not affect the ratio of enthalpy/entropy in the free energy of association compared to TTBC, indicating that the chemical identity of the core is the key aspect in the binding of these dyes. For trimerisation, a decrease in both the enthalpic and entropic components is seen, compared to the dimer, for both TTBC and BIC. This differs from PCYN, where only the entropic contribution changed and shows that the "quasi-isodesmic" behaviour observed for TTBC and BIC is due to an overall weakening of $\Delta H_{\mathrm{assoc}}$.

Table 4.5: Free energies of hydration (ΔG_{hydr}) for the cyanine dyes

Molecule	PIC	PCYN	TTBC	BIC
ΔG_{hydr} / kJ mol^{-1}	-92.7 $\pm$ 0.3	-78.3 $\pm$ 0.3	-119.9 $\pm$ 0.4	-228.8 $\pm$ 1.2

Table 4.5 presents the free energies of hydration calculated for each of the cyanine dyes. A notable observation is that all four mesogens are highly soluble in water. For comparison, the hydration free energy for quinoline[365] is -24 kJ mol^{-1} and, thus, it is expected that PIC and PCYN (whose mesogenic cores are cationic dimers of quinoline) will have higher solubility. The trend across the dyes is also as anticipated. TTBC and BIC are more soluble than PIC and PCYN due to the presence of multiple chlorine atoms on the periphery of the core. Furthermore, BIC is more soluble than TTBC as it has two additional anionic sulfonate groups.

4.3.3 Self-assembly in aqueous solution

4.3.3.1 Formation of H-aggregate stacks

In order to study the self-assembly of the cyanine dyes, initial MD simulations of 15 molecules at 10 wt% concentration were performed. All simulations were carried out at 300 K and started from randomly dispersed molecules in solution.

The aggregation of PIC into stacks occurs rapidly within several nanoseconds. In the proceeding several hundred nanoseconds, the system exhibits a range of structures from small stacks to assemblies which span the periodic box. Stacks for this dye are relatively unstable and it is observed that they break and reform throughout the course of the simu-

lation. Larger simulations of this species show the presence of stacks of various sizes. It is observed that H-aggregation is dominant in these systems regardless of the system size or concentration. The COM distance between neighbouring molecules in a stack is 0.43 nm (see Figure 4.6(c)) as predicted in the PMF of the dimer. The intermolecular stacking distance is 0.36 nm, which is concordant with the interlayer distance determined for analogous systems.[47–49] The second peak in the distributions correspond to the next neighbour of a molecule in a stack. This second peak in the stacking distances is broader than the first peak and indicates that the stacks bend, as a slight offset in the stacking causes a shift of the stacking distance to a greater value.

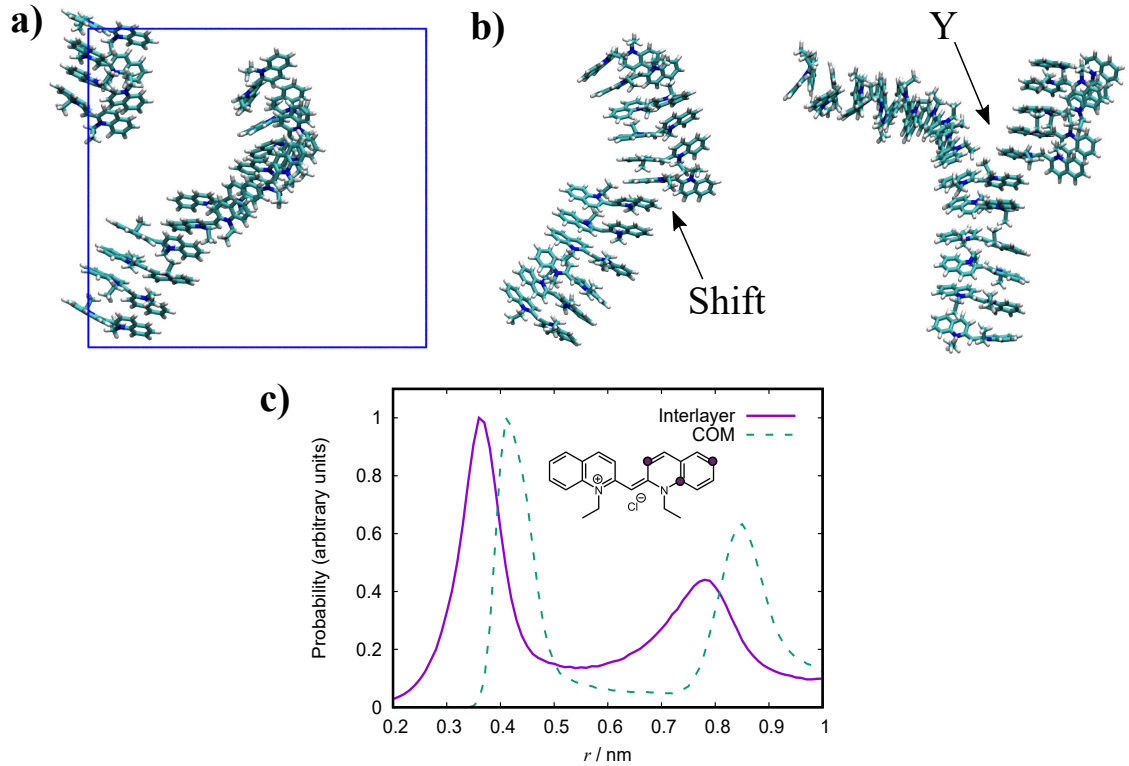


Figure 4.6: Simulation snapshot of PIC for a 10 wt% system of 15 molecules (a) and structures extracted from a simulation of 90 molecules (b) showing a shift junction (left) and a Y junction (right). (c) Histograms of the intermolecular stacking and COM distances for PIC, with the inset highlighting the atoms used to defined the molecular plane.

The dynamic nature of these assemblies can give rise to defects when smaller stacks merge together. In chromonic systems, such as Sunset Yellow (SSY), it is proposed that such aggregates contain stacking faults, which complicate their structure beyond that of simple rods. This explains the discrepancy that the persistence length of SSY aggregates is too low to produce the orientational order that it exhibits.[44] This allows for two spatial scales to be defined; a short scale where stacking is correlated (described by a persistence

length) and a long scale comprising the whole aggregate, where correlation is lost due to defects. These defects can manifest as a molecular lateral shift (shift junction) or a 3-fold Y junction, which causes the splitting of the column into branches. These Y junctions have previously been observed in a variety of soft matter systems such as block copolymers and worm-like micelles.[366–368] Here, the presence of these defects in aggregates of PIC chloride is observed, disrupting the H-type stacking and incorporating J-aggregate motifs into the structures (see Figure 4.6(b)). The propensity for H-aggregation is explained in the PMFs attained, and the possibility of J-aggregation behaviour is also predicted. Furthermore, the tendency for the stacks to be dynamic, rather than form one stable structure, can be attributed to the weak binding energy between the molecules. For the self-assembly of neutral monomers in solution, an expression for the average aggregation number, $\langle n \rangle$, has been determined in previous work[369,370] as follows

$$\langle n \rangle = \sqrt{\phi} \exp \frac{E}{2k_{\mathrm{B}}T}, \tag{4.2}$$

where E is the scission energy (the energy required to split an aggregate into two) and ϕ is the volume fraction of the solute. Where the aggregates are more transient in nature than rigid, $\langle n \rangle$ can be inferred to be $\sim L/d$ (where L is the persistence length and d is the interlayer stacking distance). Thus, it follows that the aggregate size is highly dependent on the binding energy (and concentration/temperature), where the magnitude of the binding energy also relates to the flexibility of the aggregates. The ionic nature of the dyes here warrants a modification to Eq. 4.2 as like-charged monomers experience a weakening of association due to electrostatic repulsion. Therefore, we can substitute $E = E_1 - E_{\mathrm{e}}$ into Eq. 4.2, where E is the effective binding energy (as defined previously), E_1 is the 'true' binding energy and E_{e} is a correction for the electrostatic repulsion.[44,371] The correction is defined as $E_{\mathrm{e}} = l_{\mathrm{B}} w \tilde{\tau}^2 k_{\mathrm{B}} T / 2\sqrt{\phi}$, where l_{B} is the Bjerrum length, w is the width of the monomer and $\tilde{\tau}$ is the effective charge of the rod ($\tilde{\tau} e \sim e/l_{\mathrm{B}}$). As concentration increases, the electrostatic correction decreases as counterions dispersed in solution are packed nearer to the stacks, which increases their screening effect on the monomers. This also implies that ionic additives will have a significant effect on the character of the aggregates.

For PIC chloride, application of Eq. 4.2 yields an average aggregation number of ~ 16, where E is taken to be the free energy of association for dimerisation. The use of

the binding energy as E is reasonable as the PMFs should effectively contain contributions due to the electrostatics/concentration. The volume fraction is estimated from the solvent accessible surface area[372] of the solute (using a standard probe radius of 0.14 nm), and calculating its internal volume, which provides a value of $\phi = 0.09$. It is noted that the binding energy used here corresponds to a more dilute concentration, on which its dependence is expected to have a minor effect on the magnitude. The "quasi-isodesmic" nature observed here suggests that the trimer/tetramer binding energy is a better approximation for the scission energy. This returns a value of $\langle n \rangle \approx 6$ for $E = 6.1\ k_{\mathrm{B}}T$. This value appears to be a more appropriate prediction based on the structures observed in the simulations, where defects occur between rigid sections of 4–8 molecules. This also implies that the reduction in binding energy for a trimer/tetramer could be an effect of increased electrostatic repulsion between the monomers as the stack size increases. The relatively low persistence length associated with the estimated value of $\langle n \rangle$ ($L \approx 2.2$ nm) explains the propensity for the defects observed and the flexibility of the aggregates here.

In the simulation results presented for PIC, we only observe straightforward stacks of H-aggregate character, in contrast to the J-aggregates reported in the literature. While the J-aggregate motif is present in these assemblies, it is not abundant or dominant enough to account for the experimental observations. The existence of a $\mathrm{M} \rightleftharpoons \mathrm{H} \rightleftharpoons \mathrm{J}$ equilibrium here suggests that the simulations correspond to a regime of competition between the different modes of self-assembly, preceding the emergence of large-scale J-aggregation. It is noted that previous AA MD simulation work has been reported on PIC chloride which, despite significant differences between the force fields, match the structures obtained here.[134]

Aggregates of PCYN show no defects and forms stable stacks of H-aggregate character (Figure 4.7) in contrast to PIC. Compared to PIC, the stacks formed here are much more rigid and tend not to break apart as readily, yet they retain some flexibility and bending behaviour. These assemblies are continuous over the periodic boundaries which suggests that the true extent of aggregation exceeds that captured here in a 15 molecule system. Analysis of the stacks (Figure 4.7(c)) reveals a stacking distance of 0.36 nm, and a COM distance of 0.42 nm (matching the PMF for the dimer) similar to PIC. The second peak (at 0.72 nm) in the stacking distances lies at exactly double that of the first peak; this demonstrates the rigid nature of the stacks and shows that the stacking occurs with a regular periodicity. From equation 4.2, the average aggregation number is estimated to be $\langle n \rangle \approx$

76, where $\phi = 0.08$ and $E = 11.2\ k_{\mathrm{B}}T$, and the persistence length is estimated to be $L \approx 27$ nm. These estimated quantities are beyond the length scale studied here but highlights the effect of a higher binding energy on the behaviour of the stacked structures.

The simulations here show extensive aggregation into H-aggregate stacks, as observed experimentally, but do not indicate organisation into tubular structures. However, there are a number of possible modes of evolution into nanotubes that can be suggested based on the simulation results: an end-to-end wrapping of a stack into ringed subunits to form the basis of a tube, a helical curling of a long stack to form a continuous tube, or a side-by-side alignment of stacks into a tube. Experimentally, it is observed that the formation of elongated tube structures precedes the organisation into chromonic phases.[325,328,332] Similar to PIC, it seems the simulations here capture a regime prior to large-scale aggregation.

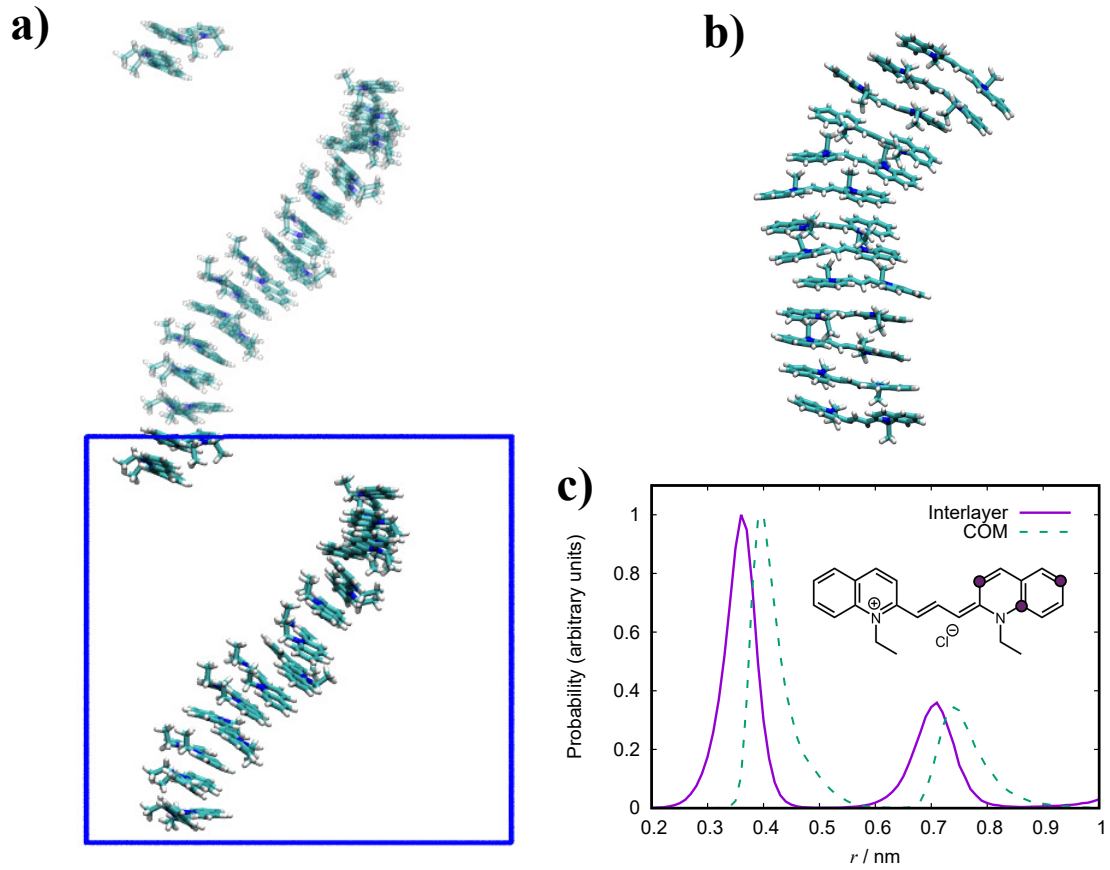


Figure 4.7: Simulation snapshots of PCYN. A 10 wt% system of 15 molecules (a) with a periodic image (shown with a semi-transparent representation) and a side view (b) of a stack. (c) Histograms of the intermolecular stacking and COM distances for PCYN, with the inset highlighting the atoms used to defined the molecular plane.

4.3.3.2 Formation of J-aggregate brickwork structures

The next two cyanine dyes, TTBC and BIC, have vastly different self-assembly behaviour to that observed in the previous section. For TTBC, a single-molecule thick, layered structure emerges (see Figure 4.9) and remains stable for hundreds of nanoseconds. Within this structure, the J-aggregation motif is dominant and produces a brickwork arrangement which is continuous across the periodic box. Unusually, this mode of ordering is not directly predicted in the PMFs, where the minimum energy configuration corresponds to a H-dimer. This suggests that the increased presence of other TTBC molecules drives the transformation into J-aggregates, and that this is an emergent phenomenon. A histogram of the interlayer stacking distances (see Figure 4.9(c)) shows two peaks at 0.36 nm and 0.72 nm corresponding to the nearest and next row of molecules in the sheet, respectively, and only a single peak in the COM distances at 0.85 nm, arising from the offset arrangement between a molecule and its adjacent molecule in the next row. The evolution of the average COM distance over the initial stages of self-assembly for neighbouring molecules (Figure 4.8) shows two stages of association before the J-aggregate structure stabilises. Firstly, small H-aggregates form in the first 5 ns (giving an average COM distance of ∼0.52 nm). Secondly, this transitions to a coexistence between H- and J-aggregate motifs when the small H-aggregates combine and resolve into the J-aggregate (with an average COM distance of ∼0.75 nm between 10–35 ns). Finally, only the J-aggregation motif is present from 40 ns, with an average COM distance of ∼0.85 nm. Simulation snapshots of each of these stages of self-assembly are shown in Figure 4.8(b).

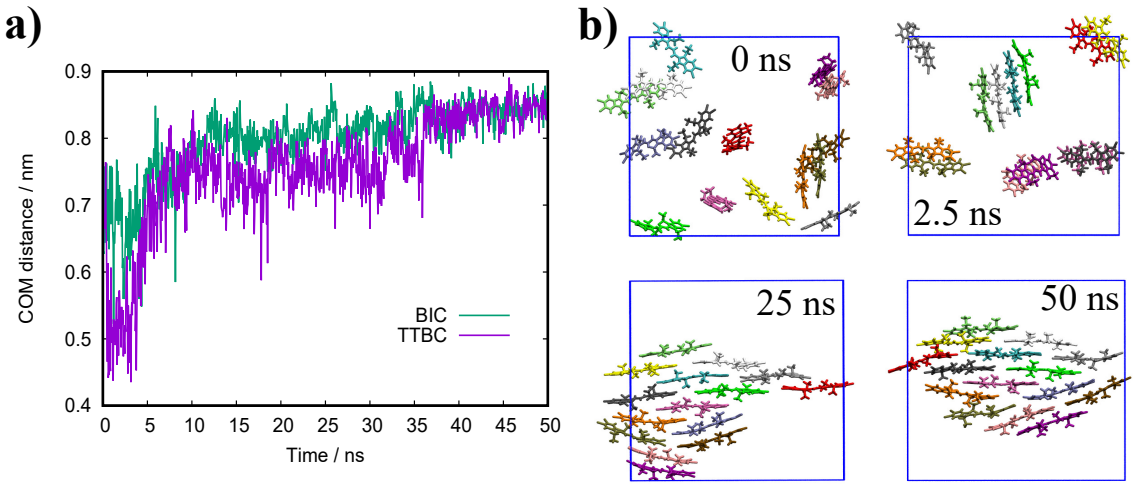


Figure 4.8: (a) Evolution of the average COM distance over the first 50 ns of simulation. (b) Simulation snapshots of TTBC at various points in the initial stages of self-assembly.

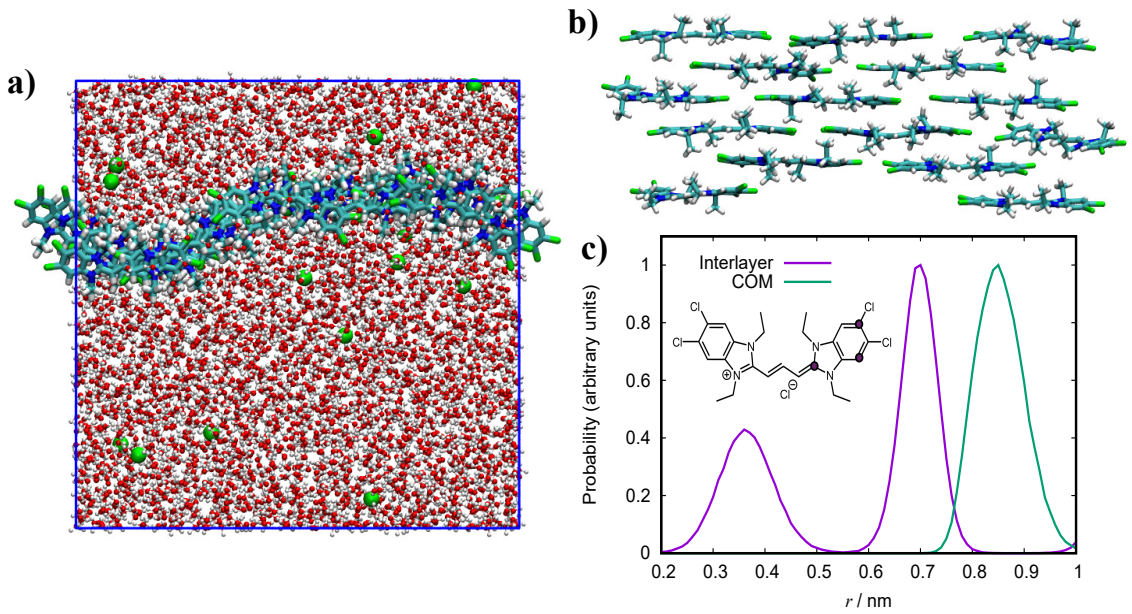


Figure 4.9: Simulation snapshots of TTBC for a 10 wt% system of 15 molecules. (a) A top-down view of the aggregate sheet structure and (b) a side view with the solvent removed. (c) Histograms of the intermolecular stacking and COM distances for TTBC, with the inset highlighting the atoms used to defined the plane.

The structure observed here is strikingly similar to the proposed smectic chromonic mesophase previously observed in cyanine dyes.[54,60,373] This mesophase consists of sheets, with a brickwork arrangement of molecules within, which are in equilibrium with other sheets in solution akin to a lamellar phase.[60] XRD measurements indicate that the sheets are one molecule thick[54], the interlayer separation is ∼0.35 nm[373] and that the molecular long axis lies parallel to the layer plane.[60] In this work, the J-aggregate sheet structure, composed of a brickwork arrangement of molecules and is of unimolecular thickness, observed is fully substantiated by the aforementioned experimental findings. In comparison to the general experimental observations for TTBC, the brickwork model is accounted for. However, there is no propensity for the herringbone arrangement despite its ubiquity in reported work. The brickwork model accounts for the type II and type III J-aggregates and, supposing the nanotubes are formed by rolling of these layers, supports the suggestion that the results correspond to an intermediate state approaching large-scale aggregate formation.

The sodium salt of BIC shares a mesogenic core with TTBC but has additional mesylate groups. This results in the molecule being anionic and less symmetric compared to TTBC. Despite these differences, BIC also forms a single-molecule thick layered structure of a brickwork arrangement (see Figure 4.10). J-aggregation is the dominant mode of

organisation observed and the structure is continuous over the periodic boundaries. This self-assembled structure differs from that displayed by TTBC in two ways. Firstly, the orientations of molecules within the structure have a slight tilting of the molecules such that the molecular long axes are not completely parallel with the layer axis. Secondly, a preference for antiparallel orientations between adjacent molecules is seen, allowing for the charged groups to lie on the opposite side for the next molecule up in a stack. Analysis of the twist angle between neighbouring molecules (see Figure 4.10(c)) shows a minor peak at $\sim$15°, corresponding to a parallel orientation, whereas the major peak at $\sim$175° corresponds to an antiparallel arrangement. Here, a neighbouring molecule is defined as a molecule within a set cutoff of 1 nm. Similar to TTBC, the brickwork model structure obtained corresponds to the smectic chromonic phase and the absorption spectrum for this species shows a single, sharp J-band[329,339] which further suggests a brickwork arrangement. It is experimentally observed that BIC J-aggregates form spherical particles[329,339], but this is not observed here. It could be suggested that the simulated structures are the basis for larger aggregates due to their proclivity for extensive aggregation across the periodic boundaries.

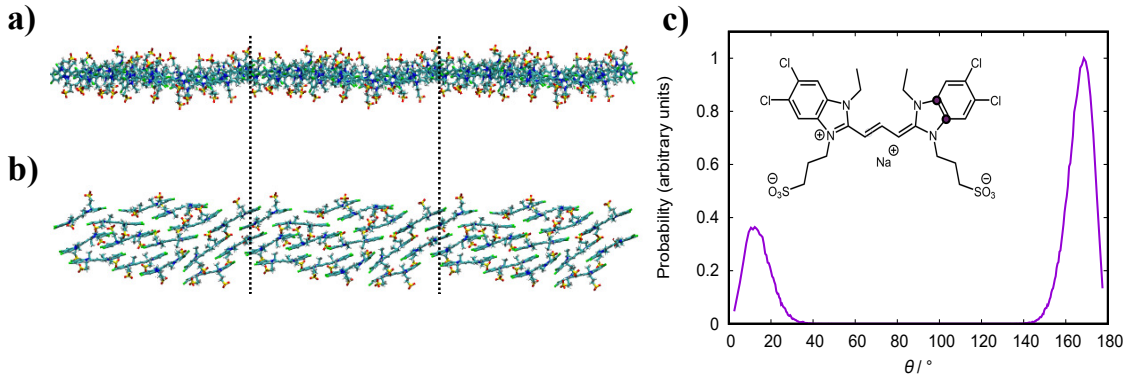


Figure 4.10: Simulation snapshots of BIC. (a) A top-down view of the aggregate sheet structure and (b) a side view. Dashed lines represent the periodic boundaries. (c) Histogram of the twist angle between neighbouring molecules for BIC, with the inset highlighting the atoms used to defined the vector.

Overall, two different classes of aggregates are observed here. PIC and PCYN form H-aggregates, whereas TTBC and BIC form J-aggregates. The fundamental difference between these pairs of cyanine dyes is the structure of the mesogenic core. The length of the polymethine chain or the presence of heteroatoms significantly alters the charge distribution across the aromatic core, which can affect both the binding strength as well as the orientations of a neighbouring molecule to accommodate preferences from electrostatic

interactions. For TTBC and BIC, it is indicated that the peripheral chlorine groups influence an offest between adjacent molecules. Another factor behind the differences is the molecular shape. PIC and PCYN are elongated disks but feature a small twist within the core. This affects the overlap of adjacent molecules and, potentially, limits the number of possible approaches for molecules to associate. In contrast, TTBC and BIC have a V-like shape and a planar core which allows for a range of slippage angles to be accessed between molecules in an aggregate.

4.4 Summary

In summary, this chapter presented all-atom molecular dynamics simulations of four ionic cyanine dyes: PIC, PCYN, TTBC and BIC. Firstly, the force field for each mesogen was optimised in terms of representing key dihedrals and atomic charges. Potentials of mean force were calculated for dimers, trimers and tetramers, from which the free energy of association can be obtained. It was found that the range of binding energies was 8–15 $k_{\mathrm{B}}T$, with a trend to higher binding strengths in species with longer polymethine chains and higher polarisability in the cationic core. Dimer configurations exhibited in the global minimum and second minimum in the PMFs correspond to H- and J-aggregation, respectively. The aggregation was observed to be quasi-isodesmic in nature, where the binding energy for dimerisation is higher than that for the trimer and tetramer. Decomposition of the association free energy into its enthalpic and entropic contributions reveals that attractive interactions between the aromatic cores drives aggregation for the ioninc cyanine dyes, in contrast to entropic component dominating for non-ionic chromonics. The free energies of hydration obtained show that all four dyes are quite soluble in water.

For the self-assembly of these cyanine dyes in aqueous solution, two classes of structures are observed. PIC forms H-aggregate stacks with shift and Y junction defects, which serve to incorporate the J-aggregation motif into the assemblies. PCYN also forms H-aggregate stacks, but its aggregates are of a more rigid character with the absence of defects. TTBC exhibits a J-aggregate sheet structure of unimolecular thickness with a brickwork arrangement of molecules within. Similarly, BIC organises into the same J-aggregate structure but with a preference for antiparallel arrangement of adjacent molecules such that the charged sulfonate groups lie on opposite sides. It is found that the self-assembly into

the various structures is an emergent phenomenon, where the presence of other mesogens is imperative in promoting the exhibited long-range order. While the experimentally observed tubular architectures are not realised in the simulations, insights into the origins of such structures are gained from exploring the molecular association at a smaller scale.

Chapter 5

Coarse-graining and Multiscale Modelling of Cyanine Dyes

5.1 Introduction

The previous chapter introduced a series of four ionic cyanine dyes and presented all-atom molecular dynamics simulations demonstrating their self-assembly in aqueous solution. This allowed for insight into the organisation of the mesogens at a molecular level and for the characterisation of the thermodynamics of association. However, the time and length scales accessible to atomistic simulation is limited; so a coarse-grained simulation approach would prove instrumental in exploring the large-scale aggregation of these systems. This chapter focuses on the development and assessment of various coarse-graining methods and mapping schemes for the cyanine dyes.

Firstly, three mapping schemes of different resolutions are compared for coarse-grained (CG) models of PIC. These were parametrised using the bottom-up iterative Boltzmann inversion (IBI) method, using the same atomistic reference. Secondly, the multiscale coarse-graining (MS-CG) method is used to develop a CG model for PCYN in the hybrid force matching (FM) approach. Thirdly, top-down Martini 3 CG models are also constructed and tested on PCYN. All the aforementioned CG models fail to produce the correct self-assembly behaviour. The challenges associated with coarse-graining these dyes will be discussed, where it is found that the mapping of the mesogens to a lower resolution model contributes significantly towards the difficulty in coarse-graining here. Finally, approaches to address these issues and develop satisfactory CG models are introduced. A toy model,

defined here as a model which seeks to reproduce a given observable of the real system but is overall generalised, was constructed for TTBC in a non-systematic manner. This model successfully exhibits the J-aggregate brickwork structure observed in atomistic simulation. A hybrid all-atom/coarse-grained (AA/CG) framework is proposed, where the cyanine dye is represented atomistically, the solvent is at a coarse-grained level and the two resolutions are coupled via virtual sites. When this multiscale modelling framework is applied to PCYN and TTBC, the correct self-assembly is observed and allows for an increase in computational efficiency without sacrificing structural accuracy.

5.2 Computational details

Parametrisation of coarse-grained models

The first step in developing the CG models was the mapping from a fine-grained AA model. Bonded parameters for the CG models are obtained from probability distributions calculated from a single molecule MD simulation in water. The mapping schemes and bonded parameters used for each model in this chapter will be detailed in their relevant sections.

For the non-bonded CG potentials, the IBI scheme (see section 2.5) was used to parametrise three CG models of PIC with different mapping schemes. This procedure was carried out in the VOTCA-CSG package, version 1.4.1.[162,175,212] RDFs for this method were calculated from the same reference trajectory: a 90-molecule system at 30 wt% and 300 K. The three models are a 5-site (3 bead types), 7-site (4 bead types) and a 13-site (6 bead types) model which give rise to a total of 15, 21 and 36 pairwise interactions, respectively, when the solvent and counterions are considered. All models here are neutral, where electrostatic effects are effectively captured in the potentials. It is noted that none of the three CG models achieved full convergence of the RDFs, and so the potentials from the iteration with the highest match were chosen for each model. The interdependence of the potentials for the RDFs combined with the high number of interactions proves too difficult for the algorithm to handle. Furthermore, no pressure correction was applied. The three IBI CG models capture all the general features of the RDFs and, thus, provide qualitative insights into these systems and commentary on the methods, but may not be reliable in predicting aggregate structures.

The MS-CG method (see section 2.6) was applied to develop FM CG models, in the hy-

brid force matching scheme.[175,176] This was performed with the BOCS (Bottom-up Open-source Coarse-graining Software) package.[224] Reference trajectories for this method contained 1000 snapshots. The linear pressure correction (see Equations 2.36 and 2.37) was employed to ensure that a pressure of 1 bar is exhibited at the correct density.

The Martini 3 force field consists of an extensive library of bead types, based on Lennard-Jones potentials, from which CG models can be constructed from (see section 2.7).[106] It is noted that this study was conducted prior to the formal publication of Martini 3, and so uses the parameters released in the open beta version.[225]

Simulation details

CG simulations for the IBI models were performed in the *NVT* ensemble with a cutoff of 1.6 nm for short-range interactions. After minimisation, equilibration runs with the Berendsen thermostat[150] for 100 ps and, then, the Nosé–Hoover thermostat[152,153] for 500 ps were carried out with a 2 fs time step. Production simulations used the Nosé–Hoover thermostat and a 5 fs time step.

The FM CG models were simulated, with a cutoff of 1.2 nm, in the *NPT* ensemble. A pre-equilibration run used the Berendsen thermostat in the *NVT* ensemble before an equilibration in the *NPT* ensemble with the addition of the Berendsen barostat. A final equilibration, and subsequent production run, were carried out with the Nosé–Hoover thermostat and Parrinello–Rahman barostat[154]. Time constants of 1 ps and 8 ps were used for the thermostat and barostat, respectively.

The Martini 3 models used a reduced cutoff of 1.1 nm for all interactions and a dielectric constant of $\epsilon_{\mathrm{r}} = 15$.[227] These models also used the velocity rescaling thermostat[151] and the reaction-field method for electrostatics[145], where $\epsilon_{\mathrm{rf}} = \infty$ beyond the cutoff.

Free energy calculations

Potentials of mean force (PMFs) were calculated at a concentration of 1 wt%. A pull was applied between the centres of mass (COMs) of each species at a pull rate of 0.001 nm ps^{-1}. Configurations of specific distance were extracted from the pull with neighbouring points varying from 0.02–0.1 nm between windows. Each window was equilibrated for 1 ns and then simulated for 20 ns to sample a range of configurations with the intermolecular distances of the COMs constrained. A total of 1×10^6 force values were output for each

point of the separation distance and used to calculate the average constraint force before integration to obtain the PMF. The free energy of association, $\Delta G_{\mathrm{assoc}}$, is defined as the maximum well depth of the PMF.

Free energies of hydration, ΔG_{hydr}, were calculated by decoupling the intermolecular interactions of a single molecule from the solvent at a concentration of 0.5 wt%, where the hydration of the species is the reverse of this process. Coulombic interactions were first decoupled linearly before the van der Waals interactions, with a total number of 41 states of λ between 0 and 1, with a spacing of 0.05. Simulations for each λ state consisted of a 1 ns equilibration followed by a 5 ns production run for data collection using the stochastic dynamics integrator.

5.3 Results and discussion

5.3.1 Comparison of mapping schemes using IBI

Given that the goal of developing CG models of the cyanine dyes in this work is explore the structure of their aggregates, a structure-based coarse-graining method appears to be an appropriate starting point. Three mapping schemes, of different resolution, were assessed for IBI CG models of PIC to determine the necessary amount of detail required for CG models.

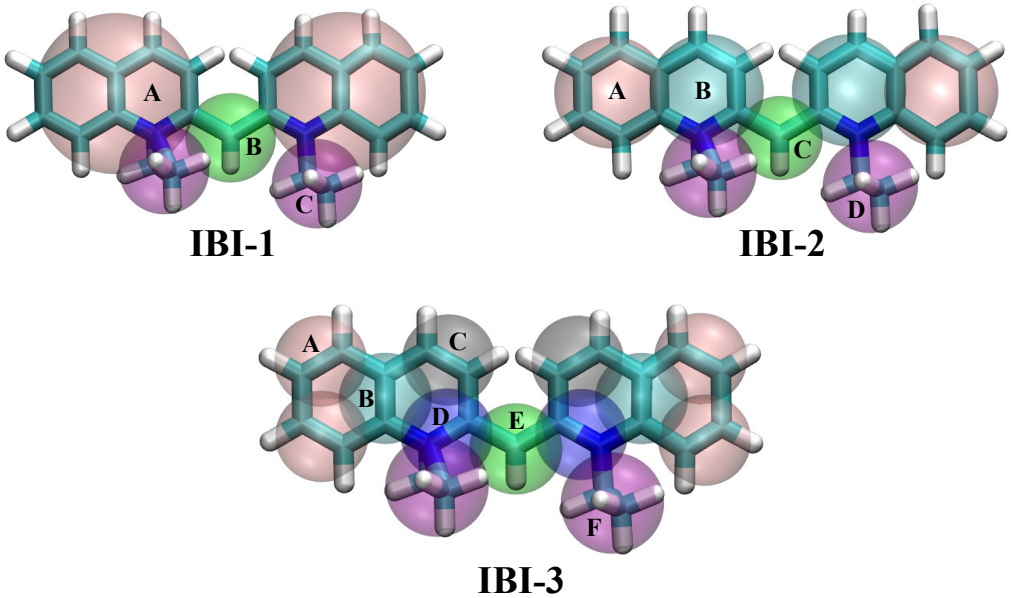


Figure 5.1: Coarse-grained mapping schemes for the three IBI models of PIC, denoted as IBI-1 (5-site), IBI-2 (7-site) and IBI-M3 (13-site). Water and chloride ions are mapped in a 1:1 fashion.

Figure 5.1 presents the mapping schemes and labels for the bead types used. Water and chloride ions are mapped in a 1:1 fashion. The first and lowest resolution model, IBI-1, maps the entirety of a quinoline unit to a single CG site whereas IBI-2 maps it to two different CG sites. The third model, IBI-3, uses a more standard mapping of 2 aromatic carbons to a single site. All three models use a single site for the ethyl group and the connecting methylene unit. The interconnected nature and stiffness of the bonds within these models proved problematic when using bonded potentials obtained via Boltzmann inversion, limiting MD simulations to a 1 fs time step. Thus, harmonic potentials were used in their place. The various bonded parameters for the three CG models are summarised below.

Table 5.1: Bond parameters for IBI-1.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
A B	0.41	10000
A C	0.36	10000

Table 5.2: Angle parameters for IBI-1.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
A B A′	136	500
C A B	52	500

Table 5.3: Proper and improper dihedral parameters for IBI-1.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
C A B A′	0	25

Table 5.4: Bond parameters for IBI-2.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
A B	0.30	10000
B C	0.27	10000
B F	0.34	10000

Table 5.5: Angle parameters for IBI-2.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
A B C	147	500
A B D	85	500
D B C	63	500
B C B′	116	500

Table 5.6: Proper and improper dihedral parameters for IBI-2.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
A B C B′	23	25
D B C B′	0	25

Table 5.7: Bond parameters for IBI-3, where * refers to the diagonal pair.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
A A	0.22	10000
A B	0.22	10000
A C	0.39	10000
A D	0.39	10000
A D*	0.45	10000
C D	0.22	10000
D E	0.20	10000
D F	0.24	10000

Table 5.8: Angle parameters for IBI-3, where * refers to a diagonal set of atoms.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
A B C	116	500
A B D	119	500
A B C*	177	500
A B D*	178	500
B D F	104	500
C D F	156	500
B D E	159	500
C D E	99	500
D E D′	155	500
F D E	93	500

Table 5.9: Proper and improper dihedral parameters for IBI-3.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
C D E D′	157	25
B D E D′	157	25
A A′ C D	0	25

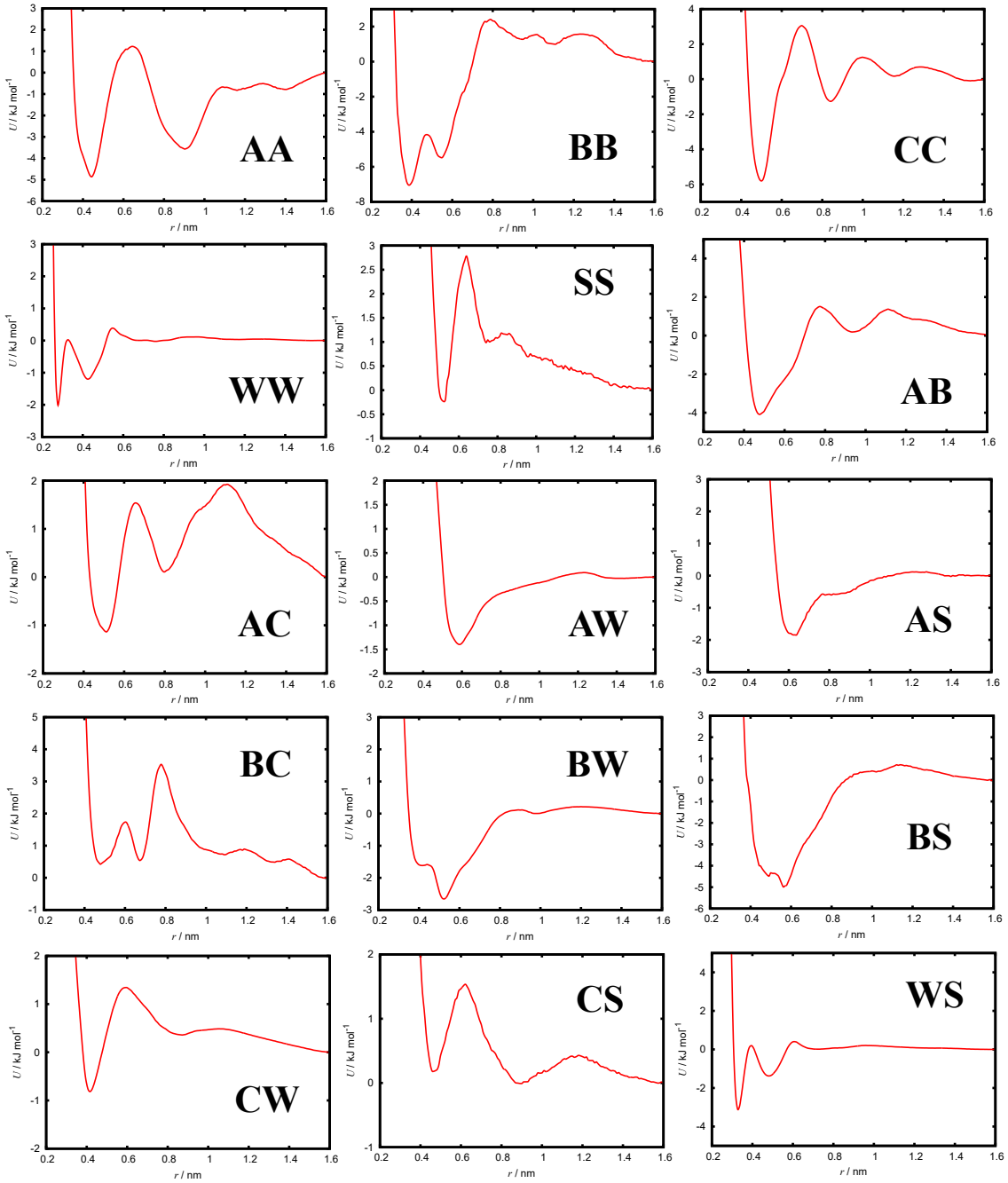


Figure 5.2: Coarse-grained potentials for the IBI-1 model, where W and S are the bead types for water and chloride ions, respectively.

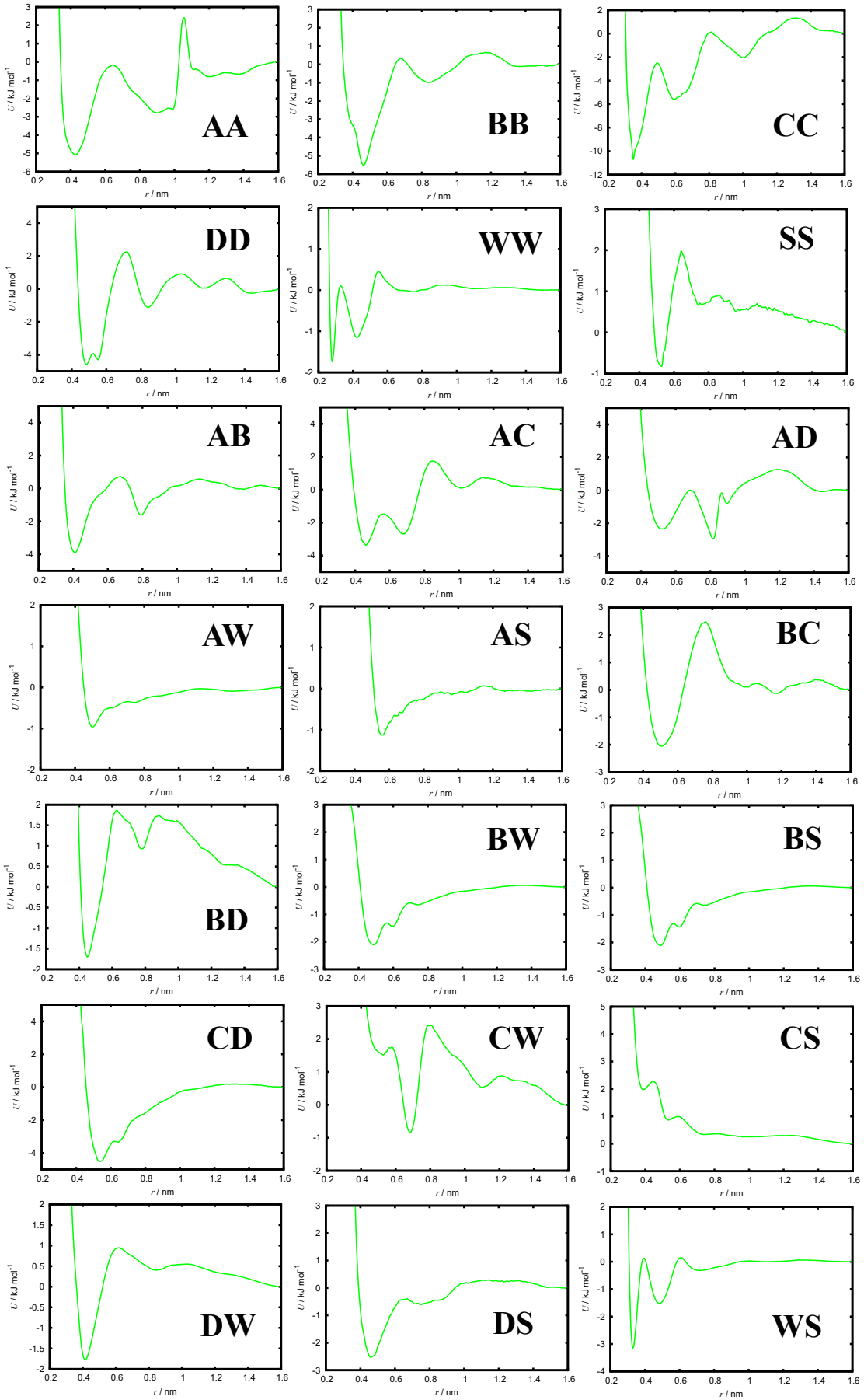
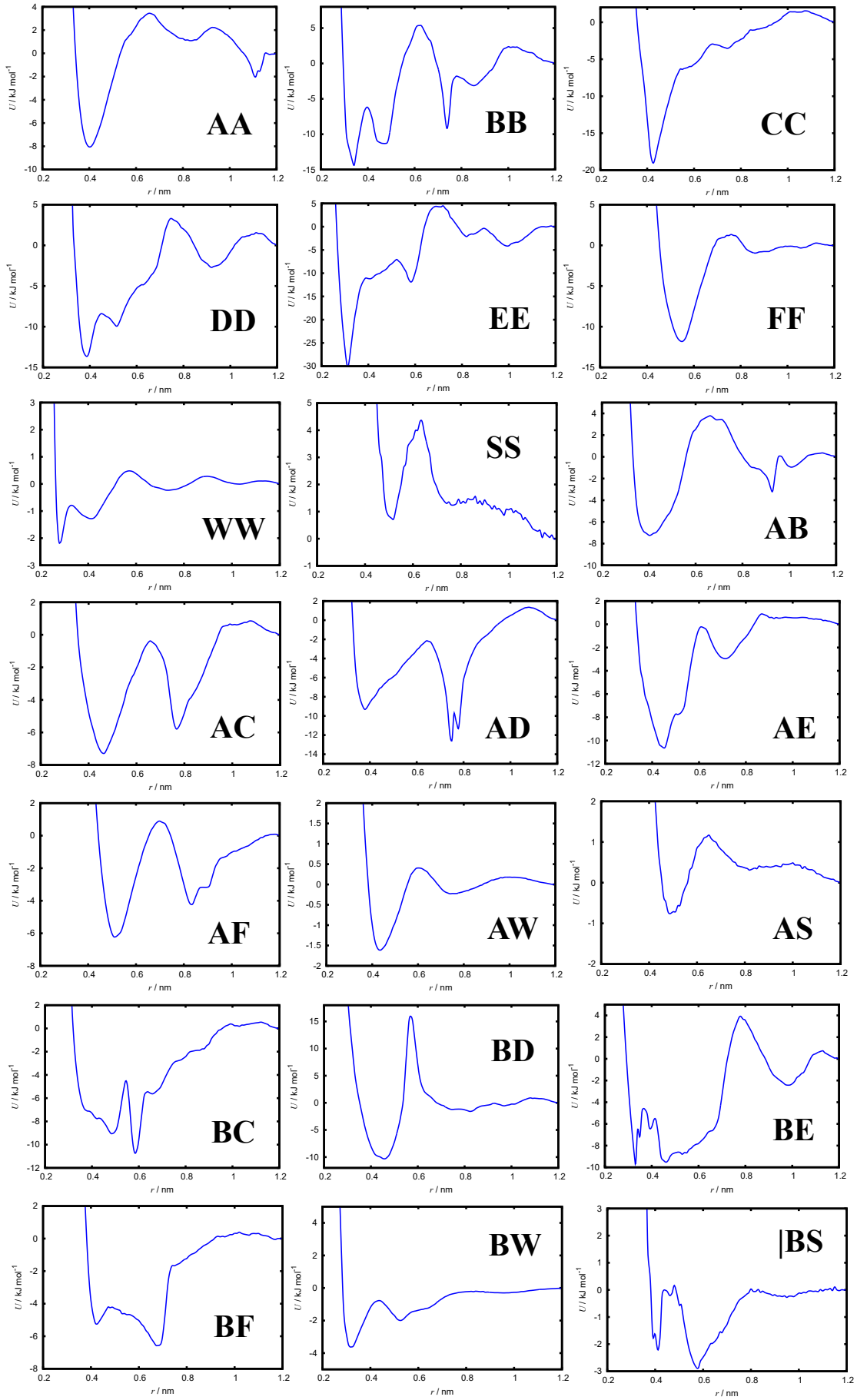


Figure 5.3: Coarse-grained potentials for the IBI-2 model, where W and S are the bead types for water and chloride ions, respectively.


AA
BB
CC
DD
EE
FF
WW
SS
AB
AC
AD
AE
AF
AW
AS
BC
BD
BE
BF
BW
|BS
U / kJ mol-1
r / nm

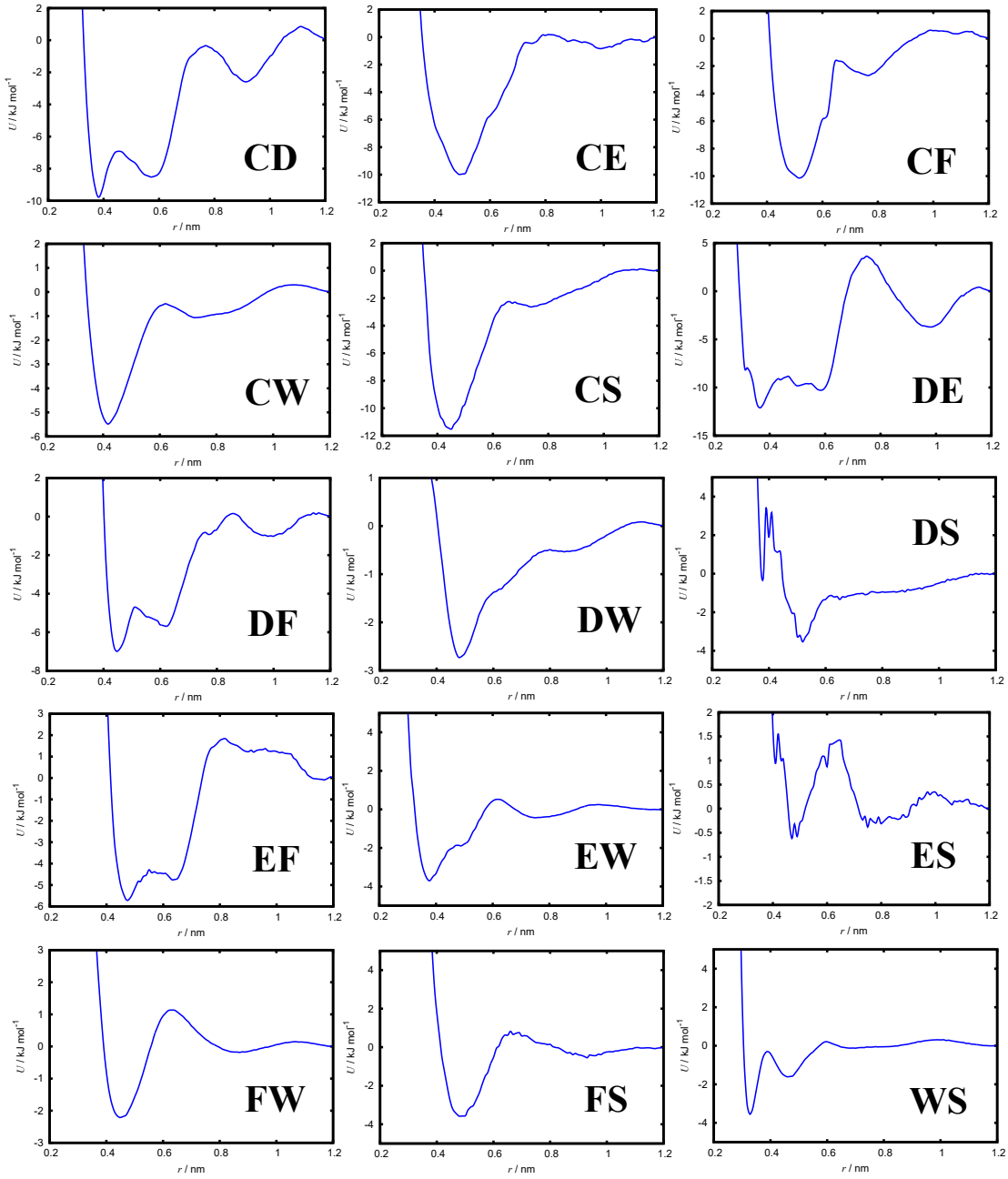


Figure 5.4: Coarse-grained potentials for the IBI-3 model, where W and S are the bead types for water and chloride ions, respectively.

The CG potentials obtained via IBI for the three models are presented in Figures 5.2, 5.3 and 5.4. It is noted that the IBI-3 model uses a 1.2 nm cutoff. To assess the various models, the free energies of association (and the preferred COM distance for a dimer) and hydration free energies are determined. It is noted that these free energies are Helmholtz free energies as they are calculated in the *NVT* ensemble. These are summarised in Table 5.10 and the PMFs are shown in Figure 5.5(a).

Table 5.10: Free energies of association (ΔE_{assoc}), the favoured intermolecular COM distance of a dimer (r_{assoc}) and free energies of hydration (ΔE_{hydr}) for the three IBI CG models.

Model	ΔE_{assoc} / kJ mol^{-1}	r_{assoc} / nm	ΔE_{hydr} / kJ mol^{-1}
AA	-20.0 $\pm$ 1.5	0.43	-92.7 $\pm$ 0.3
IBI-1	-30 $\pm$ 2	0.38	+67.2 $\pm$ 0.4
IBI-2	-47 $\pm$ 3	0.38	+98.6 $\pm$ 0.5
IBI-3	-154 $\pm$ 5	0.41	+94.3 $\pm$ 0.5

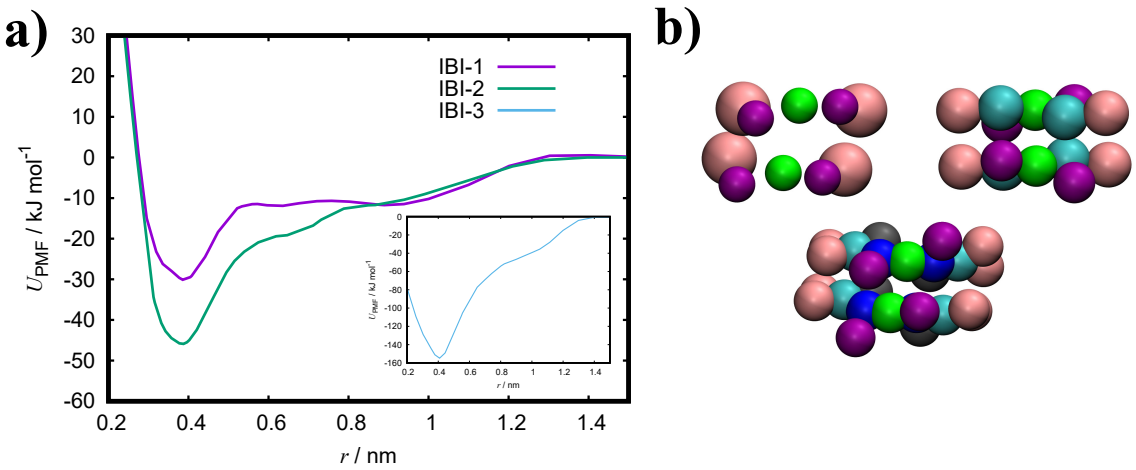


Figure 5.5: (a) PMFs, with the IBI-3 profile shown in the inset, and (b) example dimer configurations for the IBI CG models.

It is found that all three CG models overestimate the binding energy, with this value increasing along with the number of sites the model has. Despite this, r_{assoc} is in good agreement with the atomistic value and the observed dimer configurations correspond to that found for the AA system (see Figure 5.5(b)). The hydration free energies calculated for the CG models are of the wrong sign and indicate insolubility in water. Overall, it can be seen that the IBI CG models reproduce the structure well, as expected for the method, but do not capture the thermodynamic properties correctly.

To study the self-assembly of these three IBI CG models, systems of 720 molecules at 30 wt% were simulated for 1 μs. The initial system was set up from the same AA system of 90 molecules which was mapped to the different resolutions. This mapped system was replicated twice in each dimension to create the starting configuration. Thus, the three simulations were conducted at the correct AA density in the *NVT* ensemble. Snapshots of the simulations are presented in Figure 5.6. For IBI-1, the self-assembly into hollow cylinders is observed. These cylinders are rigid and span the periodic box, and do not

contain any water in the cavity. It is noted that these structures align with the diagonal of the box and the snapshot of the system has been rotated for higher visual clarity, which results in the appearance of a higher density of water in the centre of the box. It can be seen that the arrangement of molecules in the cylinder is quite random, although some stacking is present, and has walls of a single-molecule thickness. Furthermore, a preference for the A CG beads (representing the quinoline unit) interacting with water is seen, with the C CG beads (representing the ethyl groups) being shielded. This behaviour is akin to cylindrical micelles formed by surfactants.[374] The hollow cylinders further organise into a 2D rectangular array, which is the so-called rectangular phase observed in potassium oleate/water systems.[375,376]

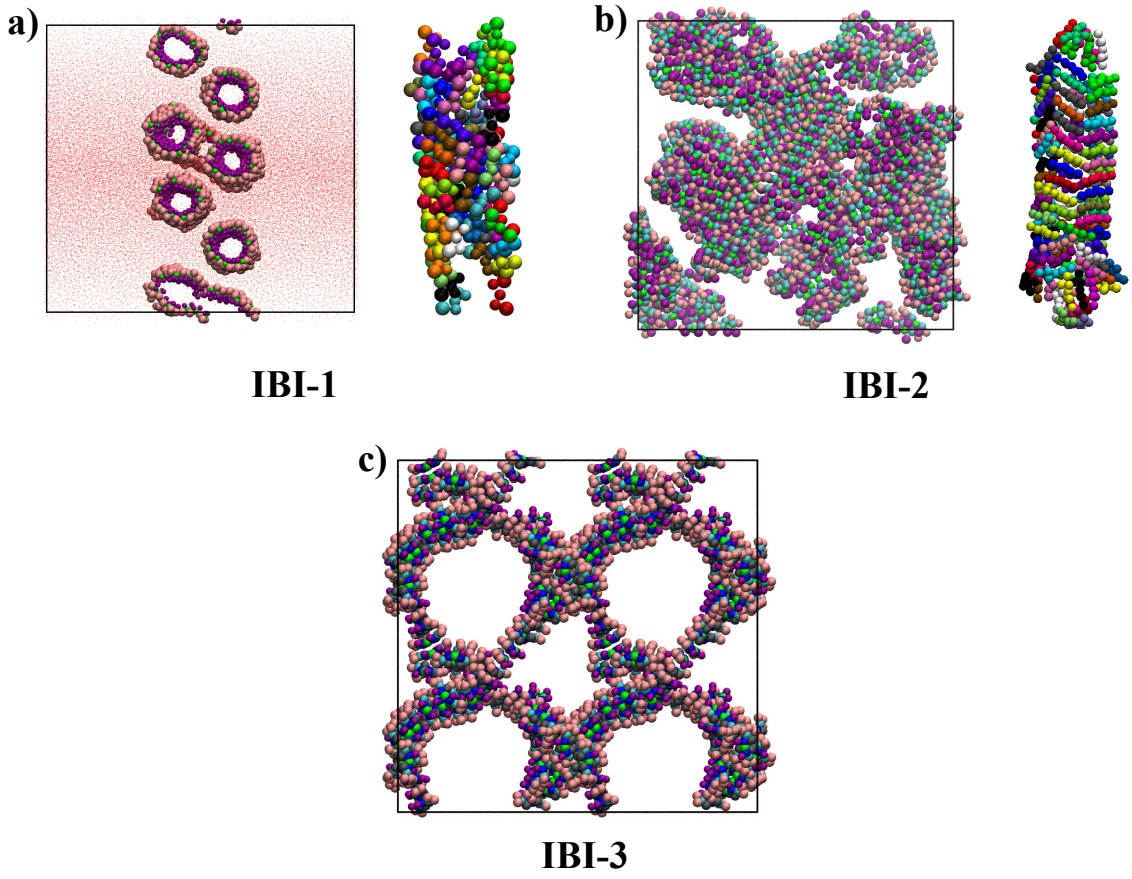


Figure 5.6: Simulation snapshots of a 30 wt% system of 720 molecules for the (a) IBI-1, (b) IBI-2 and (c) IBI-3 models. For IBI-1, the red dots represent water and focused views of the cylindrical structures are shown to the right of the IBI-1 and IBI-2 systems with individual molecules coloured differently.

The IBI-2 system also forms cylindrical structures, but of a different character to those seen for IBI-1. In terms of similarities, the cylinders are also one-molecule thick and do not contain any water. The main difference is that the molecules stack extensively with

the wall with a slight tilt. However, the cylinders do not seem to have any positional order between them and are quite transient in nature compared to the IBI-1 model. As before, there is a preference for the core CG beads to interact with water over the alkyl CG beads. The increased stacking can be suggested to arise from the higher binding energy found for the IBI-2 model, compared to IBI-1, in combination with the increased number of CG sites to direct the structure. The IBI-3 model self-assembles in a contrasting manner to the previous models. The molecules form ribbons, with single-molecule and two-molecule cross-sections at different points, which intertwine throughout the box. The dominant mode of organisation between the molecules is stacking and the structures are very rigid, as indicated by the high association free energy. It can also be suggested that the higher resolution of this CG model facilitates the stacking mode through directed interactions between pairs of sites. Similar to the AA system, there is no observed preference for a specific component of the molecule to interact with water.

These CG models allow for a qualitative perspective into the aggregation of PIC and, more importantly, a comparison of the mapping schemes within the IBI framework. As expected, there is an increase in the degree of stacking as the resolution of the model increases. It is interesting to note that all three models produce the same dimer structure in the PMFs (in good agreement with the AA system), but differ in the self-assembly of a larger system, likely due to the variation in the number and nature of the interaction sites. Despite the insolubility predicted by the free energies of hydration, all three models display lyotropic behaviour. It can be seen that the IBI-2 and IBI-3 models exhibit some microphase separation with micellar-like structures forming. This surfactant-like behaviour is irregular as the cyanine dye is not a conventional amphiphile, and can be rationalised as an imbalance between the potentials which seek to reproduce structural features seen in the RDFs rather than the underlying interactions. The preference for cylindrical structures to form may be due to the mapping of the IBI-1 and IBI-2 models, where the CG beads are slightly too large to represent aromatic units. The spherically symmetric nature of the interactions in combination with this results in aggregation in the ‘vertical’ and ‘horizontal’ directions, instead of only in the stacking (vertical) mode observed in the AA system. This is supported by the higher resolution IBI-3 model, which does not form cylindrical structures but does show stacking. It seems that there is no particular trend in the free energy of hydration as a function of a finer mapping, although there is a slight increase for the

IBI-2 and IBI-3 models compared to IBI-1. Overall, the great difficulty in parametrising CG models with the IBI method means that it cannot be recommended as an approach for the development of CG models for cyanine dyes in aqueous solution.

5.3.2 Force matching

Using the insights from the previous section, another bottom-up coarse-graining method was tested. In this section, the MS-CG method was used to develop FM CG models for the cyanine dye, PCYN. For the mapping of this molecule, 2 aromatic atoms were incorporated into a CG bead resulting in 5 sites representing a quinoline unit. The choice of a 2:1 mapping promotes stacking better compared to a coarser mapping.[106] The propene linker and the ethyl groups were mapped to single sites. Water and chloride ions are also mapped in a 1:1 fashion. In terms of bead types, the CG beads containing aromatic carbons only were assumed to be a single bead type to reduce the total number of unique bead types to 4 for the mesogen and 6 in total (including the solvent). This leads to a total of 21 pairwise interaction potentials to parametrise. Here, the electrostatic contributions to the forces are subtracted prior to force matching, such that partial charges can be used in the CG model and treated explicitly with the PME method in simulations.[139] The partial charges for a CG bead are obtained by summing up the atomic charges from the underlying atoms. The reference trajectory for this model is a 50 ns simulation of a 15 molecule system at 10 wt% and 300 K. The mapping scheme is shown in Figure 5.7 and the bonded parameters are summarised in Tables 5.11, 5.12 and 5.13.

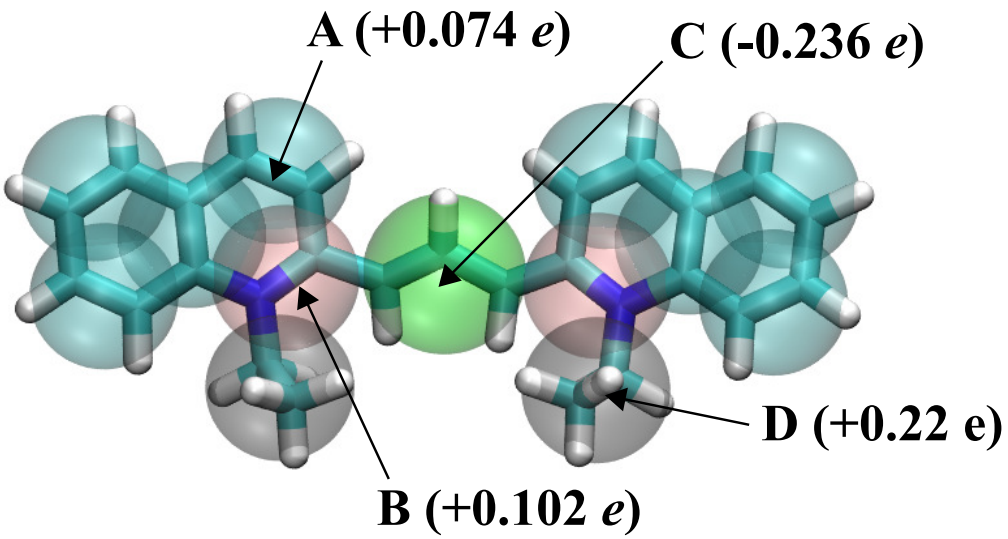


Figure 5.7: Coarse-grained mapping scheme for the FM model of PCYN.

Table 5.11: Bond parameters for the FM model.

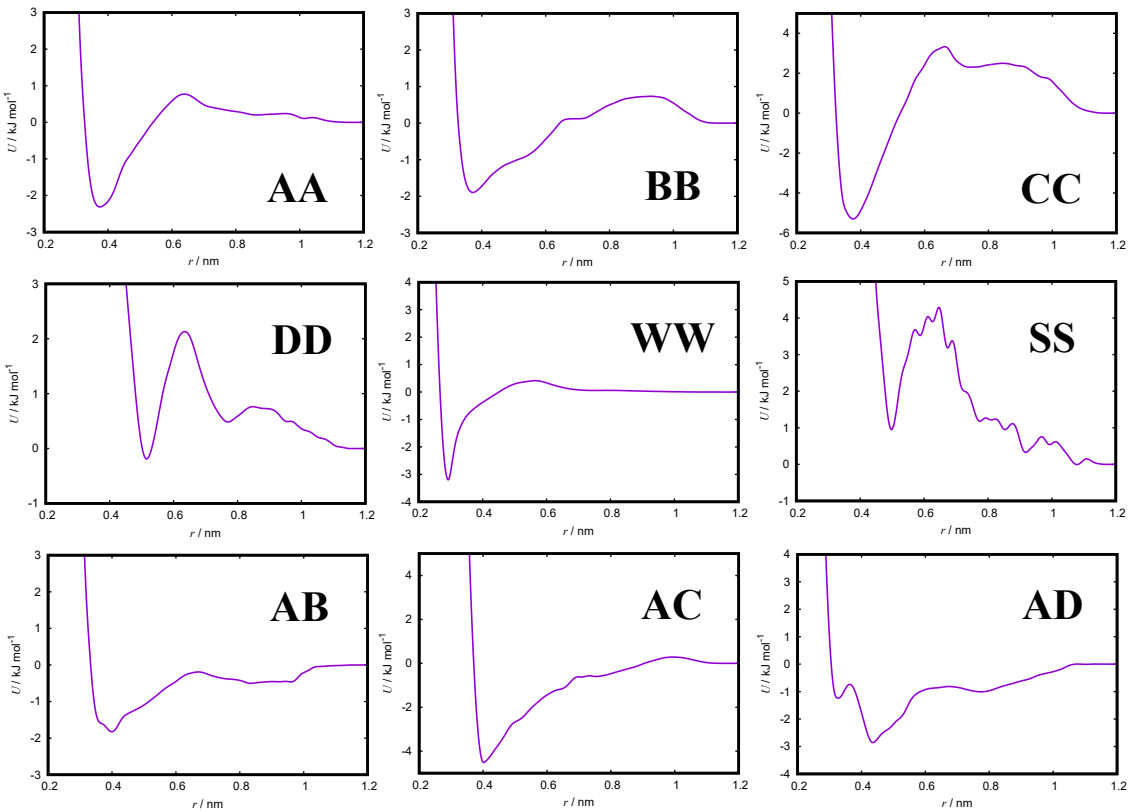
Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
A A	0.22	10000
A B	0.22	10000
B C	0.32	10000
B D	0.23	10000

Table 5.12: Angle parameters for the FM model.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
A A′ A″	60	500
A B A′	60	500
B C B′	170	500
A B D	160	500

Table 5.13: Proper and improper dihedral parameters for the FM model.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
A B C B′	97	25
A A′ A″ B	0	5



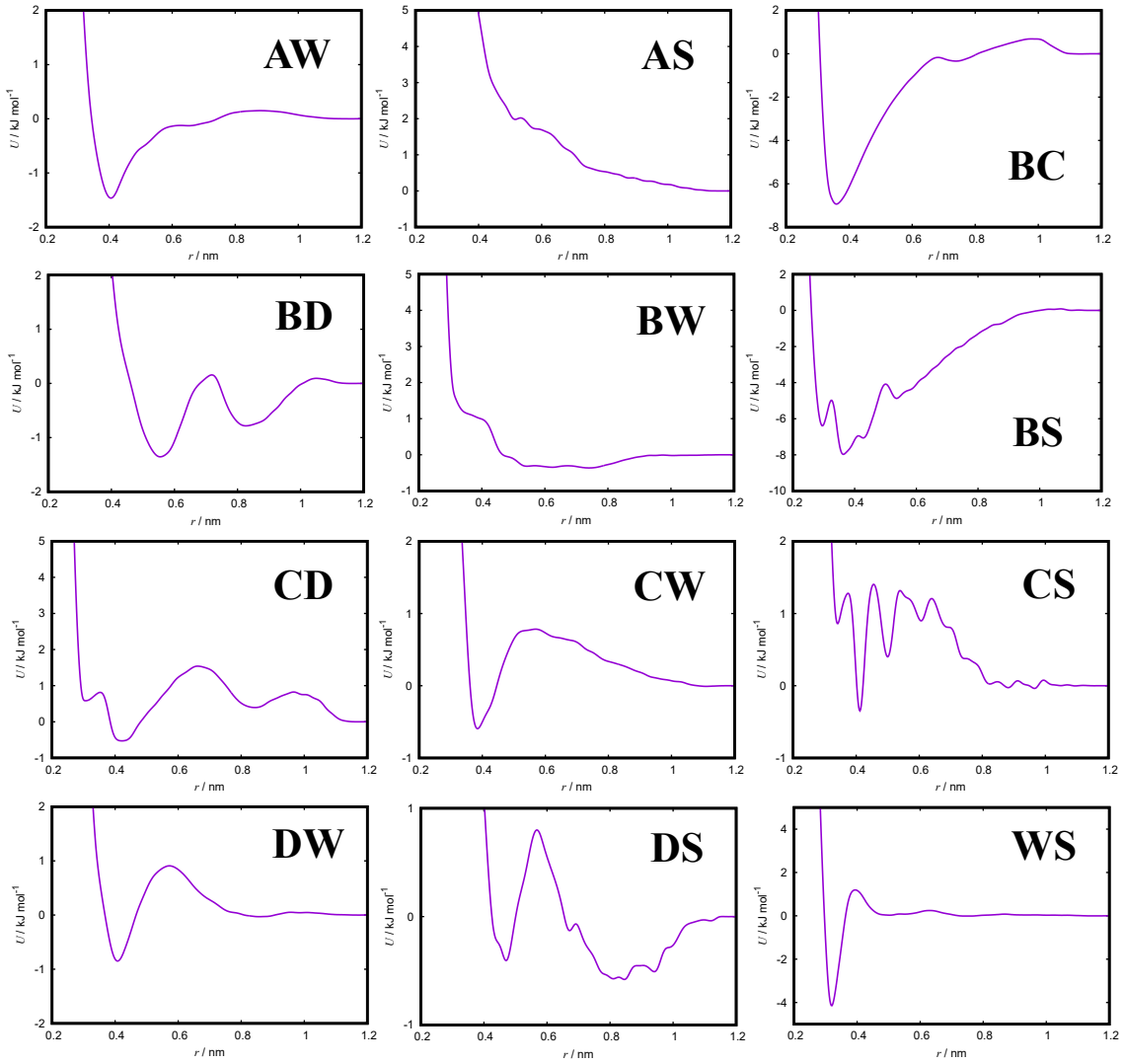


Figure 5.8: Coarse-grained potentials for the FM model of PCYN, where W and S are the bead types for water and chloride ions, respectively.

Figure 5.8 presents the CG potentials for the FM model of PCYN parametrised via the MS-CG method. The free energy of association (see Figure 5.9(a) for the PMF) calculated for this model is -140 $\pm$ 2 kJ mol^{-1} at a favoured COM distance of 0.43 nm. While this binding energy is greatly overestimated, the value of r_{assoc} is in good agreement with the AA system. The free energy of hydration obtained is +65.3 $\pm$ 0.4 kJ mol^{-1}. This reinforces the idea that structure-based coarse-graining methods, even one such as MS-CG which is based on reconstructing structure through the forces, focus on reproducing the observed structure at the cost of thermodynamics. In this case, this is the downfall of the FM model as the high binding energy in combination with the insolubility leads to a strong phase separation of the dye from water. A simulation snapshot of the system is shown in Figure 5.9(b). In this 15-molecule system, some stacking can be discerned but a single aggregate

prevails here over the self-assembly observed in the AA reference. The underlying factors are further explored and discussed in Chapter 6.

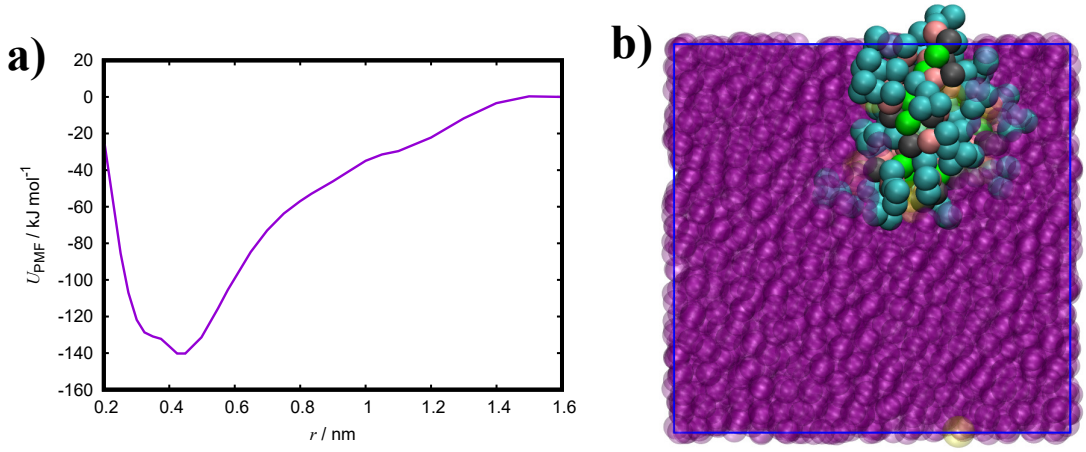


Figure 5.9: (a) PMF for a dimer of the FM model. (b) A simulation snapshot of the FM CG system for PCYN, where the purple beads represent water.

In contrast to the IBI models (albeit these were for a different cyanine dye), the FM model shows a lack of long-range order in the resulting system. Overall, all the bottom-up CG models show overestimation of the association free energy and insolubility, but this does not seem to hinder the self-assembly in the IBI models as much as the FM model. Thus, the difference is in the balance of the interactions achieved by each method. It seems that reproducing the local structure of water around a CG bead type, which is iteratively refined in the IBI scheme, is a better approach to gaining long-range order over the MS-CG method (which simply translates the force profile to an interaction potential). It is likely that the forces between components of the mesogen and water do not accurately reflect the actual strength of the interactions and structure at the AA level, a shortcoming which is carried forward into the CG models by the MS-CG method. Furthermore, if the strong attractive interactions between the aromatic sections of the mesogens is not offset by good solubility, then the model will aggregate too strongly. Compared to IBI, the MS-CG method offers a much easier and computationally inexpensive route to parametrising CG models. In IBI, the feasibility of converging the RDFs significantly decreases with an increasing number of potentials (either not being possible or requiring many iterations). This is not an issue for MS-CG as the potentials are determined in a single processing of the reference trajectory, regardless of the number of interactions. This suggests that MS-CG and FM CG models are a promising route to developing models for modelling self-

assembly in aqueous solution, but require further work to address their current limitations.

5.3.3 Martini

The next CG model to be considered is a top-down Martini 3 (see section 2.7) model of PCYN. The high transferability of the Martini force field for aqueous–hydrophobic environments makes it a suitable candidate for modelling cyanine dyes. The bead types used to represent the mesogen are A = TC4, B = TC6, C = SC4, D = TC3, W = WN and S = TQ1 (see Figure 5.7 for the bead labels). It is noted that water in this framework is mapped in a 4:1 manner. To approximate the delocalised nature of the cationic charge, the TC6 beads carry a partial charge of +0.62 *e* and the SC4 bead has a -0.24 *e* partial charge. This also avoids the need to have one CG site within the core having a full integer charge, which could overly disrupt the association of the molecules. Furthermore, the SC4 and TC6 bead types have increased interaction strengths with water (doubled and tripled, respectively) to reflect the higher solubility of charged beads. The bonded parameters for this Martini 3 model are the same as those used for the FM model (see Tables 5.11, 5.12 and 5.13).

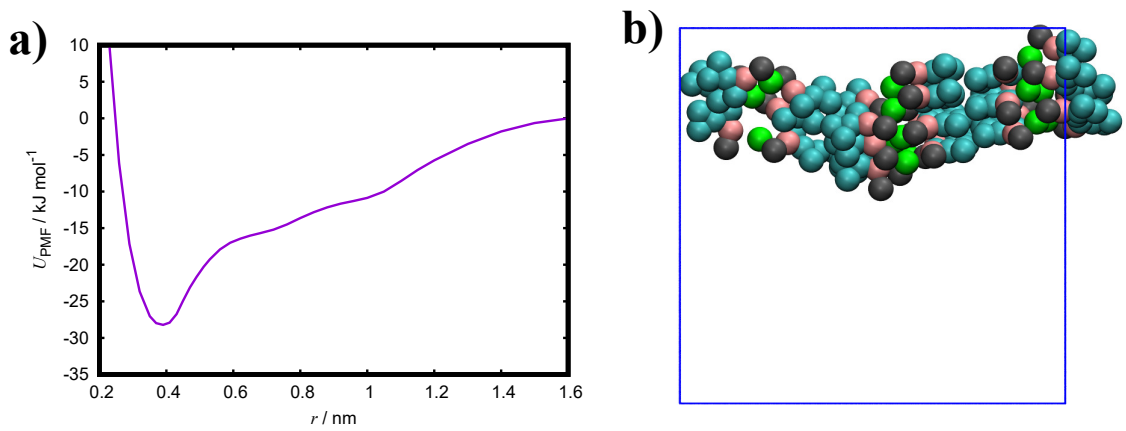


Figure 5.10: (a) PMF for a dimer of the Martini 3 model. (b) A simulation snapshot of the Martini 3 CG system for PCYN.

The free energy of association obtained for the Martini 3 model is -28.2 $\pm$ 0.5 kJ mol^{-1} with a value for r_{assoc} of 0.39 nm. Both these quantities are only slightly underestimated compared to the AA system, and shows an immense improvement over the FM model. The hydration free energy calculated is -54.3 $\pm$ 0.11 kJ mol^{-1}. Overall, it can be seen that Martini 3 captures the thermodynamic properties well, and this indicates that the model will exhibit correct behaviour in aqueous solution. A 500 ns simulation of a 15 molecule system at 10 wt% and 300 K was performed (see Figure 5.10(b)). The system shows self-

assembly into stacks, but these stacks additionally aggregate along the molecular long axes. This is not observed in the AA system, where there were only stacks of a single-molecule cross-section. However, it seems that the molecules in the CG system are soluble and there is little to no microphase separation present.

Despite achieving a favourable balance between the association strength and solubility, the Martini 3 model does not quite exhibit the correct self-assembly. It could be suggested that there are two potential limitations. Firstly, the local interactions within the mesogen do not possess precisely the correct balance of hydrophilicity/hydrophobicity. Secondly, the mapping of the model results in too much of a loss of anisotropic character (i.e. disc-like shape) where, atomistically, the face-to-face association is clearly favoured. The latter is a fundamental issue as the CG model cannot be made finer than the current mapping scheme without sacrificing the benefits of coarse-graining. This could potentially be addressed by representing the aromatic segments as biaxial ellipsoids with anisotropic interaction potentials.[377–380] The former has been discussed previously[170], and shall be further explored here. Two simple models, based on a perylene (see Chapter 6) and a quinoline core, are investigated via MD simulation using a 15-molecule system. The models consist of 8 and 5 TC4 beads for the core, with two charged Qp beads on the periphery (see Figure 5.11(a)). A parameter sweep for the TC4–TC4 and WN–TC4 interactions is conducted from ϵ increasing up to 2ϵ, and the largest stacked species in the simulated system is recorded. This allows for the effect of the balance of the core–core and core–water interaction strengths on the propensity for chromonic stacking to be investigated.

Initially, the M1 and M2 models were tested without the charged beads. Both models form a single aggregate in solution, due to the absence of the solubilising groups to stabilise stacks. For the full M1 and M2 models, it can be seen that using the normal Martini 3 values of ϵ results in only monomers present in the system, either the binding of the cores is too low or its solubility is too high. At the upper limit of 2ϵ for both interactions, stacks are prevalent for the M1 and M2 models. It can be seen that as the core–core interaction strength increases, the extent of stacking also increases but this decreases as the solubility of the core increases. In summary, the stronger the binding of the cores is, the larger the area of parameter space which allows for self-assembly into stacks is. For the M1 model, the largest stacks found contain 10 molecules whereas the M2 model only forms stacks of 4 molecules. This highlights the effect of the core size on the self-assembly, where the extent

of stacking increases with the number of sites. Furthermore, the M1 model has a much larger area of parameter space which results in stacks of dimer size and above compared to the M2 model. There are two key findings of note here. The first is that the presence of explicit solubilising groups assists greatly in the promotion of stacking in water. The second is that the number of combination of parameters which allow for stacking is lower for a smaller core size. This suggests that the incorrect self-assembly behaviour found for CG models of cyanine dyes is due to limitations in the mapping, where the mesogens lack explicit solubilising components and have relatively small core sizes.

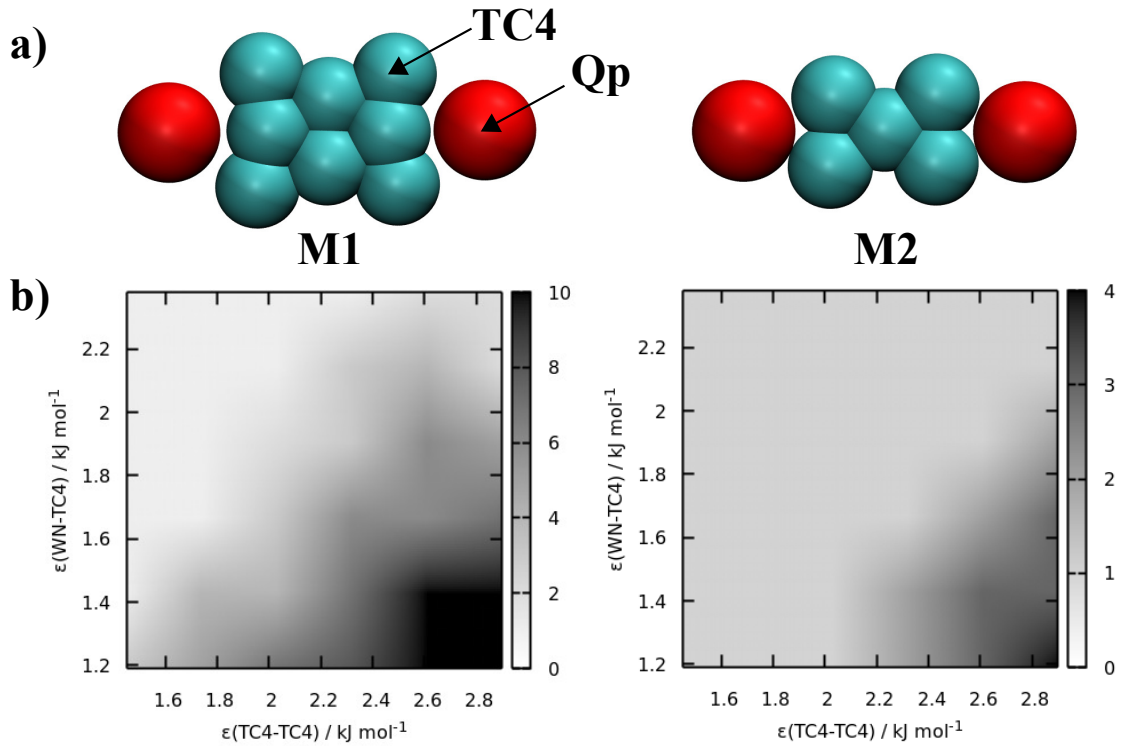


Figure 5.11: (a) Structures of the M1 and M2 test models with 10 (left) and 7 (right) sites. (b) Plots of the size of stacks observed in simulations of the M1 (left) and M2 (right) models for a parameter sweep of the TC4–TC4 and WN–TC4 interaction strengths.

5.3.4 Toy model

Moving away from chemically specific CG models, a toy model for TTBC was developed. A toy model is a minimalistic CG model that describes the general features of the system and qualitatively captures the phenomena under study with a simple parameter set. Such models have been of great value in studying self-assembly in many biological systems.[381,382] For TTBC, the MS-CG method was first used to parametrise a CG model but this only formed a single aggregate in solution as for PCYN. Similarly, a Martini 3 model

produced the same aggregation behaviour as the FM model. Due to the failure here to develop a more refined CG model of TTBC based on the underlying system, a toy model approach is tested here. This CG model aims to reproduce the J-aggregate association motif in a brickwork arrangement, as observed in the AA system, through a small number of specific interactions. A similar concept has been used to study smectic A phases with DPD models, where the local antiparallel association is controlled through directed interactions.[383] Other targets for the CG model of TTBC is the unimolecular cross-section of the structure. While only the self-assembly behaviour is targeted, this should implicitly achieve a good balance between the association strength and solubility for the model.

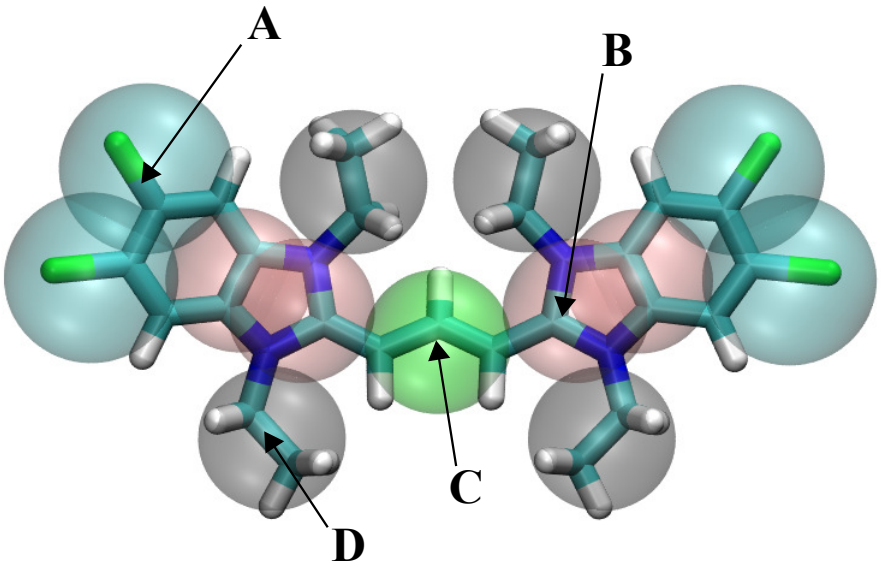


Figure 5.12: Coarse-grained mapping scheme for TTBC.

Table 5.14: Bond parameters for the toy model.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
A A	0.28	5000
A B	0.33	5000
B C	0.32	5000
B D	0.3	5000

Table 5.15: Angle parameters for the toy model.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
A A′ B	65	500
A B B′	155	500
B B′ C	172	500
B C B′	164	500
B B′ D	95	500
C B D	77	500

Table 5.16: Proper and improper dihedral parameters for the toy model.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
A A′ B B′	0	25
A B B′ D	167	25
B B′ C B″	0	25

Figure 5.12 shows the mapping scheme used in this CG model, which follows the general 2/3 to 1 mapping for aromatic sites. The bonded parameters for the harmonic potentials are summarised in Tables 5.14, 5.15 and 5.16. Water and chloride ions are mapped in a 4:1 ratio and 1:1 ratio, respectively. The non-bonded interactions are represented by Lennard-Jones potentials. The interaction distances, σ, are obtained by calculating the corresponding intersite distances (where $\sigma = r_{\min}/2^{1/6}$) from a mapped reference of the brickwork structure with the J-aggregate motif. Interaction distances for solute–solvent and solvent–solvent potentials are generic. Initially, all values of ϵ were set to 1 kJ mol^{-1}. In order to bias an offset configuration for adjacent molecules, ϵ for the AB and AC cross-interactions are increased. The AW and DW interaction strengths are also increased to reflect the solubility of the cyanine dye and promote a single-molecule cross-section for the structure. The final parameters were determined by scanning through ϵ values across a range from the initial guess. The final interaction matrix for the σ and ϵ LJ parameters are summarised in Table 5.17.

Table 5.17: Interaction matrix for the LJ σ (nm)/ϵ (kJ mol^{-1}) parameters of the toy CG model, where W and S refer to water and chloride ions, respectively.

Bead type	A	B	C	D	W	S
A	0.48/3.5	0.45/3.5	0.34/3.5	0.45/1.0	0.41/1.75	0.41/1.0
B	–	0.33/3.5	0.42/1.0	0.49/1.0	0.41/1.0	0.41/1.0
C	–	–	0.71/1.0	0.58/1.0	0.41/1.0	0.41/1.0
D	–	–	–	0.46/1.0	0.41/5.25	0.41/1.0
W	–	–	–	–	0.47/3.5	0.41/3.5
S	–	–	–	–	–	0.41/1.0

For the thermodynamic properties, the association free energy is calculated to be -18.1 $\pm$ 0.3 kJ mol^{-1} (see Figure 5.13(a) for the PMF) at a favoured COM distance of 0.84 nm, and the hydration free energy obtained is -98.5 $\pm$ 0.2 kJ mol^{-1}. The dimer configuration in the minimum of the PMF corresponds to the J-aggregate motif, which has a COM distance of 0.84 nm in agreement with the self-assembled structure in the AA system. This is contrasting behaviour to that seen in the AA PMF, where the preferred configuration is a H-dimer. The free energy of hydration is slightly underestimated, but shows that the model has good solubility. Overall, the thermodynamic quantities are qualitatively in line with the AA values despite not being the targets of the parametrisation. A short 100 ns simulation (see Figure 5.13(b)) of this toy model for a 15 molecule system (at 10 wt% and 300 K) shows the formation of the target brickwork structure of unimolecular thickness. It is noted that the biasing of the model to adopt the J-aggregate motif results in an absence of any H-aggregation, even in the initial stages of self-assembly.

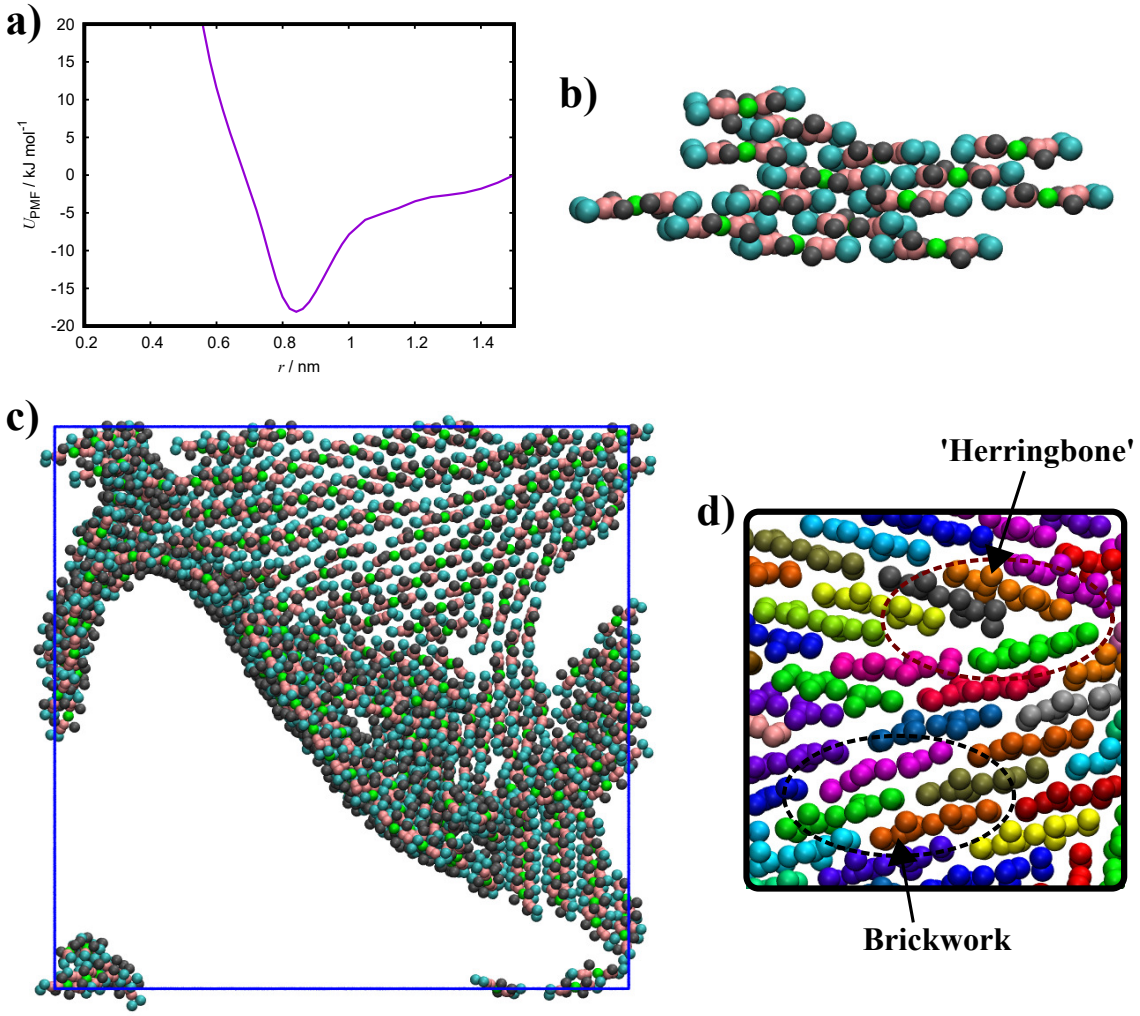


Figure 5.13: (a) PMF for a dimer of the toy model. Simulation snapshots of a (b) 15-molecule and a (c) 405-molecule system. (d) A focused view showing the local arrangement between molecules, where individual molecules are coloured differently.

This toy model was then used for a larger simulation of a 405-molecule system at 10 wt% and 300 K for 4 μs. Figure 5.13(c) shows a simulation snapshot of the resulting system. The system self-assembles in a large single aggregate, consisting of a local brickwork arrangement between molecules and is one molecule wide, which shows some significant curvature. It could be suggested that in the absence of periodic boundary conditions, the sheet structure would curve and interact with itself to form a cylindrical structure (as observed in experiment[335]). Figure 5.13(d) shows a focused view of the structure in which two packing motifs can be discerned. The first is the brickwork arrangement which was the target in parametrising this model. The second is a partial herringbone-like arrangement. Both these packing arrangements have been reported in crystal structures[350] and aggregates in aqueous solution[335]. It is likely that the herringbone-like ordering found here is

a result of segments of the structure 'zipping up'. Here, a complete rearrangement could be difficult as it would require a significant number of molecules to be released from the structure.

Looking at the LJ parameters for this model, it can be seen that the specific ratio of bead sizes and interaction strengths that produce the correct self-assembly would be impossible to achieve with the bead types available in Martini 3. While not explored here, it could be that there is an appreciably sized area of parameter space that produces the target behaviour. It is also unlikely that the correct balance of interactions can be obtained with the MS-CG method, although this would provide a good estimation of the interaction lengths. A methodology based on using initial CG potentials from MS-CG, which are then scaled in terms of interaction strengths by the user to produce the correct self-assembly could be useful in developing similar models. In terms of the current approach used to develop the toy model of TTBC, the LJ parameters could be further optimised to exactly reproduce the free energies of association and hydration to create a chemically specific CG model.

5.3.5 Hybrid AA/CG models

As discussed previously, it is likely that the mapping of cyanine dyes to a lower resolution makes it difficult to obtain a parameter set which produces the correct self-assembly. The loss of molecular detail within the mesogen at the CG level leads to the bottom-up and top-down models failing to show the expected aggregation behaviour in aqueous solution, even if the binding strength and solubility are captured well. A multiscale modelling approach could address this issue, where the region of interest is described in atomistic detail and the remainder of the system is modelled at the CG level within a MD simulation.[384–389] When Martini is used to represent the CG level, the two resolutions can be coupled via virtual sites.[390–392] These studies demonstrate that this hybrid AA/CG approach allows for structural and thermodynamic accuracy comparable to the all-atom systems.

In this section, hybrid AA/CG models for PCYN and TTBC are developed through the combination of the GAFF and Martini force fields. Virtual sites are defined at the COM of the underlying atoms of a CG bead (see Figure 5.7 and 5.12). The cyanine dyes, at an all-atom resolution, can interact with themselves but not with the solvent. Solute–solvent interactions are handled through the interactions of the virtual sites with the solvent, where

the virtual sites are excluded from interacting with the AA model. The solvent is represented as proper CG beads which also do not interact with the AA force field. The solvent–solvent and solute–solvent interactions are described using the Martini framework. Here, chloride counterions are counted as part of the solvent subsystem. Simulation parameters used for these models are the same as those used for the Martini force field, except for the use of a time step of 2 fs.

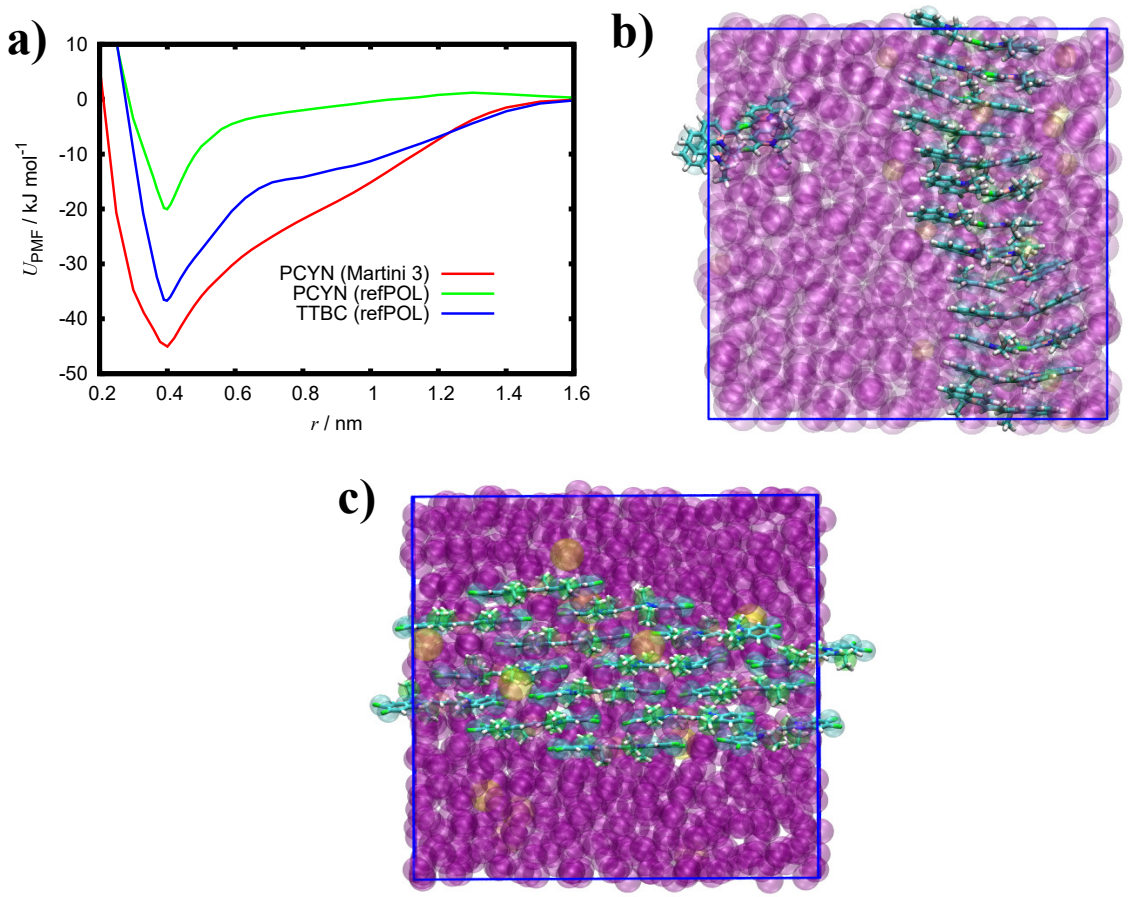


Figure 5.14: (a) PMFs for a dimer of the hybrid AA/CG models. Simulation snapshots of a 15 molecule system at 10 wt% for the hybrid AA/CG models for (b) PCYN and (c) TTBC, where the purple and yellow beads represent water and chloride ions, respectively.

The PMFs for the hybrid AA/CG models are shown in Figure 5.14(a) and the results of the free energy calculations are summarised in Table 5.18. An initial model was constructed for PCYN using the Martini 3 force field for the CG solvent. It was found that the binding energy is slightly too high and its hydration free energy is of the wrong sign. A second model for PCYN was developed using the polarisable refPOL/refION[207,208] Martini force field. This methodology requires explicit electrostatic coupling between the solute and solvent.[390,391] The refPOL Martini water has partial charges that are tuned so that the

dipole moment and bulk dielectric constant of water is reproduced. Simulations using this force field were performed using the PME method for electrostatics and with $\epsilon_{\mathrm{r}} = 2.5$. The use of polarisable CG water in the hybrid framework allows for the AA atomic charges to explicitly interact with and be screened by the solvent, which results in the PCYN (refPOL) hybrid model exhibiting good solubility. It is noted that the additional sites present in the refPOL water model and the use of the PME method means that its use is computationally more expensive than Martini 3 water. The hybrid AA/CG model for TTBC also used refPOL water. For Martini 3, solute–solvent interactions were represented by the TC4 bead types, whereas the polarisable refPOL model used the SC5 bead type for the virtual sites. The self-assembly of the hybrid AA/CG models (see Figure 5.14(b) and (c)) matches that observed in the AA systems. It is noted that the self-assembly of the Martini 3 and refPOL hybrid models for PCYN is similar, but it is likely that this is driven by phase separation in the Martini 3 version with the atomistic detail allowing the local structure to be exhibited. Thus, this framework allows for the cyanine dyes to be simulated with good structural accuracy, but with a speed up in simulation time due to the reduction of the number of sites present and faster dynamics.

Table 5.18: Free energy of association ($\Delta G_{\mathrm{assoc}}$), the favoured intermolecular COM distance of a dimer (r_{assoc}) and the free energy of hydration (ΔG_{hydr}) for the hybrid AA/CG models.

Model	$\Delta G_{\mathrm{assoc}}$ / kJ mol^{-1}	r_{assoc} / nm	ΔG_{hydr}/ kJ mol^{-1}
PCYN (Martini 3)	-45 $\pm$ 2	0.40	+59.2 $\pm$ 0.13
PCYN (refPOL)	-20 $\pm$ 3	0.40	-100.3 $\pm$ 0.6
TTBC (refPOL)	-36.7 $\pm$ 0.8	0.40	-141.5 $\pm$ 0.7

The hybrid AA/CG models (using refPOL water) for PCYN and TTBC were used for larger simulations for 1 μs of a 500 molecule system. These systems are shown, with the atomistic regions only, in Figure 5.15. The system for PCYN displays self-assembly behaviour identical to the AA model, where long stacks of a single-molecule cross-section form. However, there is no inclination towards the formation of nanotubes despite the larger system size. It is possible that an increase in concentration could result in a transition to a nanotube structure through a closer packing of the columns, where there is lateral association in addition to stacking.[328,332] However, moving to higher concentrations may be detrimental to the self-assembly as the extent of stacking and stack sizes will decrease

as the volume of the system decreases. Stacks will form readily, but the closer packing between stacks may result in lower long-range order as they have less freedom to reorient. For TTBC, the J-aggregate brickwork structure is observed. This is similar to the behaviour found for the toy model (see Figure 5.13), but the hybrid model does not show the herringbone-like arrangement. This suggests that the herringbone motif observed is a feature exclusive to that model, rather than being intrinsic to the atomistic model (if it is assumed that the AA model is more accurate). Once again, the structure favours interaction across the periodic boundaries as a semi-infinite sheet over rolling up into a tube. Future simulations could be designed to discourage this behaviour. Overall, the hybrid AA/CG approach is effective in allowing a scaling up of the system size while retaining good structural accuracy with better computational efficiency.

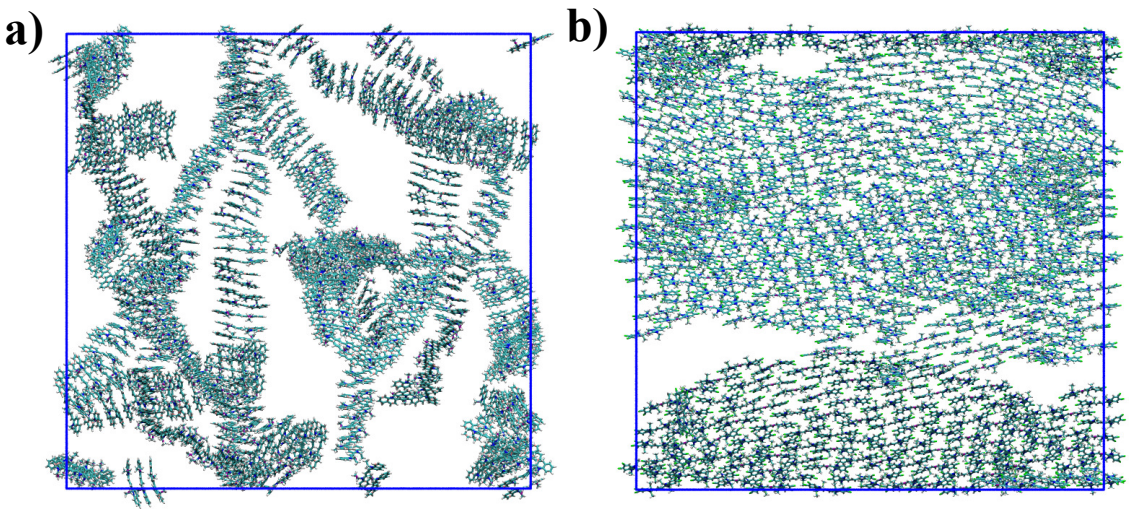


Figure 5.15: Simulation snapshots of a 500 molecule system at 10 wt% for the hybrid AA/CG models of (a) PCYN and (b) TTBC.

5.4 Summary

In summary, various coarse-graining approaches and a multiscale modelling framework have been applied to cyanine dyes. The IBI method was used to parametrise CG models for PIC, employing mapping schemes of different resolutions. It was found that all three models exhibit the same dimer structure in the PMFs, but the free energy of association is overestimated. Furthermore, the hydration free energies indicate that the models are insoluble, highlighting that this structure-based coarse-graining approach focuses on reproducing structure at the expense of thermodynamics. The two lower resolution IBI

models form cylindrical micelles, where the extent of stacking in the tube walls increases with a finer mapping of the mesogen. The most fine-grained IBI model forms long stacks of one- and two-molecule cross-sections which intertwine through the system.

The bottom-up MS-CG method was used to develop a FM CG model of PCYN. A number of issues, such as the insolubility, cause the incorrect self-assembly behaviour to be found for this system and a single aggregate forms instead. A top-down Martini 3 model for PCYN similarly displays this unfavourable aggregation behaviour, despite the thermodynamic properties related to the binding and solubility being represented well. This suggests that the correct self-assembly is being hindered by another factor, such as the mapping itself. Two simple models, representing different aromatic cores, were investigated through a parameter sweep of the core–core and core–water interactions to identify the parameter combinations that result in chromonic self-assembly. It is found that the presence of explicit solubilising groups assists in promoting stacking and that the correct balance of interactions, which allows for stacking, is more difficult to achieve with a smaller core size.

A minimalistic toy model for TTBC was constructed to reproduce the self-assembly behaviour found in the AA system. This CG model exhibits the correct J-aggregate brickwork arrangement on a large scale, but does not show the formation of a tubular structure. It is also found that the thermodynamic quantities are in good agreement with the AA model and that reproducing the correct self-assembly behaviour implicitly achieves the correct balance of these interactions.

Finally, a multiscale modelling approach using hybrid AA/CG models allows for the simulation of these systems with good structural accuracy but with higher computational efficiency. This is indicated to be a promising strategy in cases where a successful CG model cannot be easily developed, such as the cyanine dyes in this work. Moreover, future work utilising this framework would benefit from multiple-time-step algorithms, where a computational speed up is achieved by employing different time steps for the AA and CG degrees of freedom within the same MD simulation.[393]

Chapter 6

Approaches to Coarse-graining a Chromonic Perylene Dye

6.1 Introduction

Perylene tetracarboxylic acid bisimides are compounds that show intense visible light absorption, and have been widely used as stable dyes in industrial applications.[394] As such, the syntheses and functionalisation of perylene dyes, their applications in organic photovoltaic materials and a plethora of supramolecular architectures with numerous modes of organisation have been extensively reported in the literature.[65,395–397] Subsequently, perylene dyes have been designed and investigated for various applications in aqueous media as chromonic liquid crystals.[63,398–402] In this work, simulation studies were conducted on one chromonic perylene mesogen, bis-(N,N-diethylaminoethyl)perylene-3,4,9,10-tetracarboxylic diimide dihydrochloride, henceforth referred to as PER (see Figure 6.1).

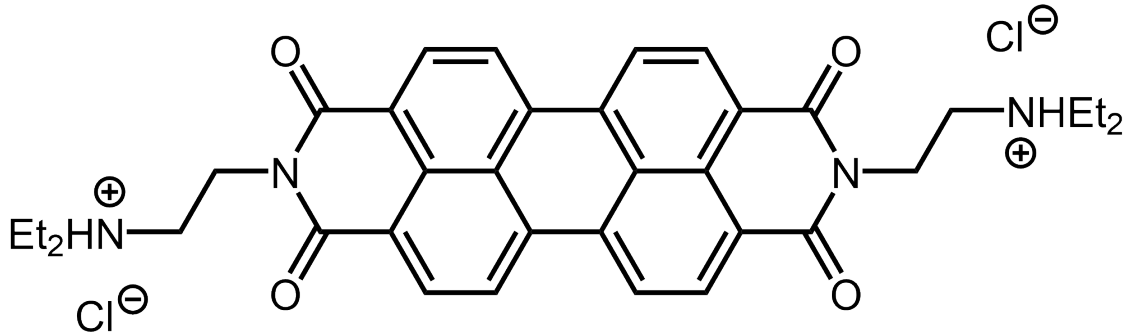


Figure 6.1: Molecular structure of PER with its counterions.

Experimental studies on PER reveal the formation of H-aggregate stacks which organ-

ise into a chromonic nematic (N) phase, with a hexagonal (M) phase at higher concentrations, at room temperature.[403,404] Small-angle X-ray diffraction measurements show a peak corresponding to an intracolumnar spacing of 0.3–0.4 nm, where its broadness indicates disordered stacking.[403] The columns are suggested to have a single-molecule cross-section, assemble in an isodesmic fashion and have a free energy change of 16–20 $k_{\mathrm{B}}T$ for the removal of a molecule from a stack.[405] The thermodynamics of association have also been characterised for various derivatives of PER by simulation and experiment, with the effects of solvent, temperature and substituents explored.[130–132]

The development of coarse-grained models for chromonic liquid crystals has largely been focused on the non-ionic mesogen TP6EO2M.[136–139] Using a model of a simplified representation in a dissipative particle dynamics (DPD) framework, an exponential distribution of stack sizes in the isotropic phase has been observed, with the chromonic N and M phases forming at higher concentrations.[136] Variants of this DPD model predict an array of novel structures and phases with complex stacks of two and three molecule cross-sections.[137] The application of systematic coarse-graining methods to obtain more chemically specific models for TP6EO2M reveals a wealth of subtle complexities and challenges.[138,139] A precise balance of hydrophobic-hydrophilic interactions within the mesogen and between the mesogen and solvent must be achieved for chromonic self-assembly to occur, as well as considerations in the mapping of the model.

Following on from previous work, bottom-up (FM) and top-down (Martini) coarse-graining approaches are applied to PER. In contrast to the ionic cyanine dyes (chapters 4 and 5), this chromonic dye should be more conducive to coarse-graining for a variety of reasons. PER features a mesogenic core of a larger surface area with an increased number of aromatic units. It has been found that a minimum core size is required for chromonic aggregation.[406] The cationic charges are localised on alkylammonium groups on the peripheral chains rather than embedded within the core. This allows for a more independent charged bead type to be defined, instead of demanding that the bead types in the mesogenic core represent a complex combination of interactions alone. The expected self-assembly behaviour (H-aggregate stacks organised into a chromonic N phase) is also much simpler than the complex aggregates suggested for the cyanine dyes. Overall, these structural and behavioural properties should be more favourable in facilitating the development of coarse-grained models compared to cyanine dyes.

This chapter commences with all-atom molecular dynamics simulations of PER to observe its self-assembly in aqueous solution and to calculate the free energies of association and hydration. These observations and quantities serve as validation targets for the coarse-grained models developed using systematic approaches. Firstly, coarse-grained models parametrised by the multiscale coarse-graining method, in the form of hybrid force matching (FM), are presented, where multiple strategies are explored but none result in an adequate model. Secondly, a model utilising the Martini 3 force field is found to be suitable in producing chromonic self-assembly, and a simple optimisation is applied such that the relevant thermodynamic properties are also exactly reproduced. Thirdly, an implementation of the FM potentials in a combined framework with Martini allows for the recovery of the correct self-assembly behaviour. Finally, the coarse-graining of a thermotropic discotic analogue, with oligoether chains on the periphery of the perylene bisimide core, is explored. The work presented herein is published in the paper[407] "*Molecular simulation studies of self-assembly for a chromonic perylene dye: All-atom studies and new approaches to coarse-graining*," and is reproduced from Ref. 407 [G. Yu and M. R. Wilson, *J. Mol. Liq.*, 2022, **345**, 118210] with permission from Elsevier.

6.2 Computational details

Force fields

The General AMBER Force Field (GAFF) was employed for all atomistic simulations.[220] The Antechamber package from AmberTools18[221] was used to generate Lennard-Jones potentials and partial charges for the force field, with the latter employing the AM1-BCC method.[222] The resulting GAFF topologies were converted into the necessary input files for GROMACS using the ACPYPE script.[223] All simulations used the TIP3P water model[351], which is compatible in conjunction with GAFF.[352]

Coarse-grained mapping

A single mapping scheme was used for all coarse-grained models of PER. Beads for the aromatic core contained 2/3 carbon atoms (CC, CM and CO for central, middle and outer beads, respectively), 5 heavy atoms for the imide group (N), 2 carbon atoms for the con-

necting ethyl group (O) and 5 heavy atoms for the terminal diethylammonium group (Q); see Figure 6.2 for the mapping scheme. All coarse-grained beads include the associated hydrogens on heavy atoms. This CG mapping allows for each distinct chemical environment in the molecule to be represented by its own bead type. The FM models used a 1:1 mapping for water and chloride ions, whereas Martini 3 used the standard 4:1 mapping for water. This leads to an overall total of 6 bead types and 21 pairwise non-bonded interaction potentials to be parametrised.

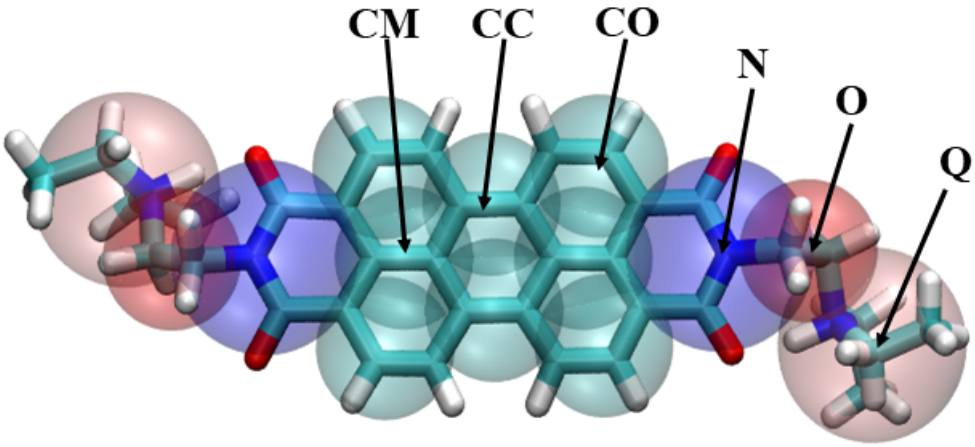


Figure 6.2: Coarse-grained mapping scheme for PER.

A single molecule all-atom molecular dynamics trajectory of PER in TIP3P water at 300 K was used to generate probability distributions for the bonded parameters. Equilibrium bond lengths, angles and dihedrals were defined as the value of the maximum probability from each distribution. These were implemented as harmonic potentials with appropriate force constants chosen. Improper dihedral potentials were applied to the aromatic core to ensure planarity. All coarse-grained models in this work employed the same set of bonded parameters (see Tables 6.1, 6.2 and 6.3).

Table 6.1: Bond parameters for the coarse-grained models.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
CC CC	0.24	10000
CC CM	0.24	10000
CC CO	0.27	10000
CM CO	0.21	10000
CM N	0.33	10000
N O	0.22	10000
O Q	0.28	10000

Table 6.2: Angle parameters for the coarse-grained models.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
CO CC CO′	140	500
CO CC CM	50	500
CO CM CO′	160	500
CC CM CC′	60	500
CM CC CM′	120	500
CC CM N	150	500
CO CM N	80	500
CM N O	160	500
N O Q	105	500

Table 6.3: Proper and improper dihedral parameters for the coarse-grained models.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
CM N O Q	80	100
CO CO′ CO″ CO‴	0	100
CO CC CC′ CO′	0	100
N CO CM CC	0	100

Multiscale coarse-graining

The multiscale coarse-graining (MS-CG) method[97] matches force distributions obtained from the reference atomistic model to forces acting between coarse-grained sites (see section 2.6). Here, the hybrid force matching approach has been used, where bonded potentials are obtained separately to the non-bonded potentials.[175,176] An extension of this is used here to separate out the electrostatic contributions so that they may be explicitly treated in the simulation. This was achieved by an additional step in which coarse-grained electrostatic forces, from the mapped atomistic charges, were subtracted from the effective, mapped reference trajectory before the MS-CG method was applied.[139] Charges for each bead are obtained by summing up the atomic charges of its constituent atoms.

MS-CG was carried out using the BOCS (Bottom-up Open-source Coarse-graining Software) package.[224] Reference trajectories were composed of 1000 snapshots from a 10 wt% system of 15 molecules, for which intramolecular forces were excluded. Two reference trajectories were used to parametrise three coarse-grained models: i) an uncharged model using a 10 ns simulation of a self-assembled system containing stacks of tetramer size and above, ii) a model utilising the same atomistic reference but with the explicit

treatment of electrostatics and iii) a charged model parametrised from a 100 ns simulation which captures self-assembly from a dispersed system of monomers into dimers and then tetramers. All FM pair pair potentials were tabulated (see Figure 6.5).

Coarse-grained potentials constructed with this methodology overestimate the pressure of the system and, thus, require a pressure correction to allow for simulation in the *NPT* ensemble at 1 bar with the correct density. Here, an approach that has been applied previously to FM potentials[139,170] is adopted via the VOTCA-CSG (Versatile Object-oriented Toolkit for Coarse-graining Applications) package, version 1.4.1.[162,175,212] A linear pressure correction was applied, in an iterative manner, to all the potentials simultaneously according to Equations 2.36 and 2.37. It should be noted that this correction does not strictly address the state-point dependence of the potentials, which requires additional terms in the coarse-graining procedure relating to volume/density dependent components.[184,186,190]

Martini 3

The Martini 3 force field consists of an extensive library of bead types, based on Lennard-Jones potentials, from which coarse-grained models can be constructed from (see section 2.7).[106] It is noted that this study was conducted prior to the formal publication of Martini 3, and so uses the parameters released in the open beta version.[225]

Simulation details

All MD simulations were carried out using the GROMACS 2018.7 molecular dynamics simulation package.[226] Long-range electrostatics were treated using the PME method with a cutoff of 1.2 nm for all short-range interactions. After minimisation, a 100 ps pre-equilibration run in the *NVT* ensemble was carried out using the Berendsen thermostat followed by a 100 ps pre-equilibration in the *NPT* ensemble with the addition of the Berendsen barostat.[150] An equilibration run of 200 ps and a subsequent production simulation for 500 ns utilised the Nosé–Hoover thermostat[152,153] to maintain a constant temperature of 300 K, and the Parrinello–Rahman barostat[154] to keep the pressure constant at 1 bar. Time constants of 1 ps and 5 ps were used for the thermostat and barostat, respectively. A leap-frog algorithm was employed with a time step of 1 fs for equilibration with an increase to 2 fs for production runs, where constraints were implemented using the LINCS method.[144]

Coarse-grained simulations were performed using a 2 fs time step for equilibrations and 5 fs for production runs, where the length of each simulation stage follows the atomistic conditions, as stated above. No constraints were used for coarse-grained simulations and all potentials are shifted such that $U = 0$ at the cutoff. For all FM models, the same thermostats, barostats and cutoffs were used in coarse-grained simulations as in the atomistic simulations. Time constants of 1 ps and 8 ps were used for the thermostat and barostat, respectively. The Martini 3 models used a reduced cutoff of 1.1 nm for all interactions and a dielectric constant of $\epsilon_{\mathrm{r}} = 15$.[227] These models also used the velocity rescaling thermostat[151] and the reaction-field method for electrostatics[145], where $\epsilon_{\mathrm{rf}} = \infty$ beyond the cutoff.

Free energy calculations

Potentials of mean force (PMFs) were calculated at a concentration of 1 wt%. A pull was applied between the centres of mass (COMs) of each species at a pull rate of 0.001 nm ps^{-1}. Configurations of specific distance were extracted from the pull with neighbouring points varying from 0.02–0.1 nm between windows. Each window was equilibrated for 1 ns and then simulated for 20 ns to sample a range of configurations with the intermolecular distances of the COMs constrained. A total of 1×10^6 force values were output for each point of the separation distance and used to calculate the average constraint force before integration to obtain the PMF. The free energy of association, $\Delta G_{\mathrm{assoc}}$, is defined as the maximum well depth of the PMF.

Free energies of hydration, ΔG_{hydr}, were calculated by decoupling the intermolecular interactions of a single molecule from the solvent at a concentration of 0.5 wt%, where the hydration of the species is the reverse of this process. Coulombic interactions were first decoupled linearly before the van der Waals interactions, with a total number of 41 states of λ between 0 and 1, with a spacing of 0.05. Simulations for each λ state consisted of a 1 ns equilibration followed by a 5 ns production run for data collection using the stochastic dynamics integrator.

6.3 Results and discussion

6.3.1 Atomistic simulation

All-atom molecular dynamics (AA MD) simulations were carried out on a 15-molecule system of PER at 10 wt% and 300K. These were initiated from a random configuration of solvated dye molecules and allowed for observation of the self-assembly in aqueous solution. At the atomistic level, structural and thermodynamic properties were calculated to inform and validate the coarse-grained (CG) models developed in this work.

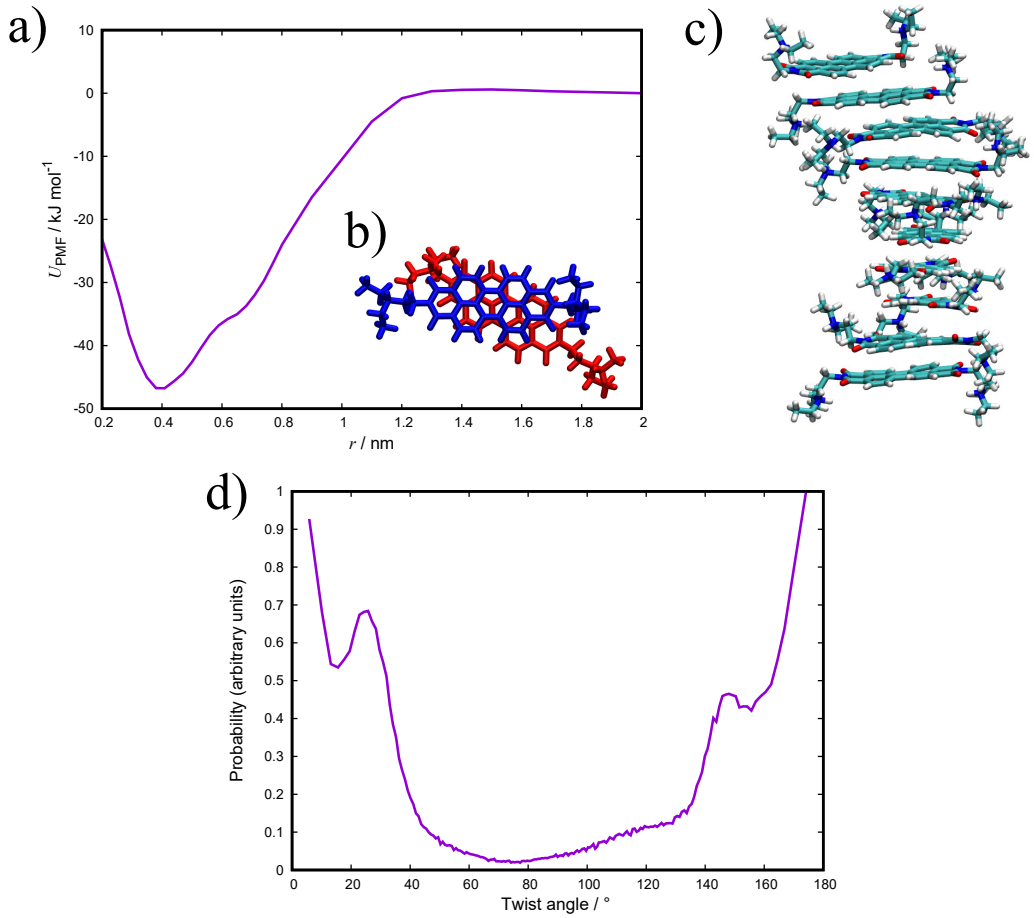


Figure 6.3: (a) PMF for a dimer of PER in TIP3P water at 300K with a (b) top-down view of a dimer configuration showing an interlayer twist. (c) A simulation snapshot of a 10-molecule stack of PER extracted from a 10 wt% system of 15 molecules. (d) Histogram of the twist angle between neighbouring molecules.

Table 6.4: Free energy of association (ΔG_{assoc}), the favoured intermolecular COM distance of a dimer (r_{assoc}) and the free energy of hydration (ΔG_{hydr}) for PER.

ΔG_{assoc} / kJ mol^{-1}	r_{assoc} / nm	ΔG_{hydr}/ kJ mol^{-1}
-47 $\pm$ 3	0.39	-133 $\pm$ 1.0

Figure 6.3(a) shows the PMF profile for a dimer of PER and Table 6.4 summarises the thermodynamic quantities obtained for this dye. Errors in $\Delta G_{\mathrm{assoc}}$ are calculated by propagation from each data point used in the PMF calculation and the error for ΔG_{hydr} is estimated from block averaging. The binding energy of 18.7 $k_{\mathrm{B}}T$ for PER is appreciably higher than previously calculated values for chromonics (7–15 $k_{\mathrm{B}}T$[47–50]) due to PER containing a more extended aromatic core than systems studied so far. However, the calculated value is within the range determined experimentally for this dye (16–20 $k_{\mathrm{B}}T$).[405] Figure 6.3(b) shows a dimer configuration extracted from the minimum of the PMF, and displays H-aggregate character with an interlayer twist between the molecules to reduce steric/electrostatic repulsion between the bulky, charged terminal groups. The di-cationic nature of this dye indicates that PER is highly soluble and the free energy of hydration calculated confirms this (for comparison with other ionic dyes, see Table 4.5).

From dispersed monomers in solution, self-assembly into stacks of trimer size and above occurs within 50 ns. It is found that, over several hundred nanoseconds, larger stacks form and break apart as observed previously in ionic cyanine dyes.[50] Figure 6.3(c) shows a typical stack which displays a single-molecule cross-section, as observed in experiment[65], an interlayer twist between adjacent molecules as predicted in the PMF. However, a periodic twist in a single direction along the stack is not found, and so this interlayer angular offset does not impart chirality upon the column. Analysis of this twist angle (Figure 6.3(d)) between adjacent molecules within a stack reveals a preference for rotations of 6°, 24°, 156° and 174°.

6.3.2 Force matching

Three CG models were parametrised using the MS-CG method in the formulation of hybrid force matching. A fully neutral CG model (FM-N) was developed first, where all electrostatic contributions are incorporated into the effective pair potentials. Thus, no partial charges are present on any of the CG sites, including the counterions. The principal FM CG model (FM1) utilises partial charges on CG beads, arising from the AA model, and treats long-range electrostatics explicitly via the PME method. Here, the chloride ions hold a -1 charge. This methodology was expected to improve the representability of the ionic nature of the dye compared to a neutral version. The aforementioned FM models both used the same AA reference trajectory for parametrisation; a 15-molecule system at

10 wt% and 300 K which dynamically contained two or three stacks which varied in aggregation number. A second charged model (FM2) was constructed with a longer AA reference trajectory which captures the self-assembly of 15 dispersed monomers into dimers and trimers. A schematic for the mapping in the FM CG models is shown in Figure 6.4(a), where it is noted that partial charges for C beads are only present on the four outer (CO) beads.

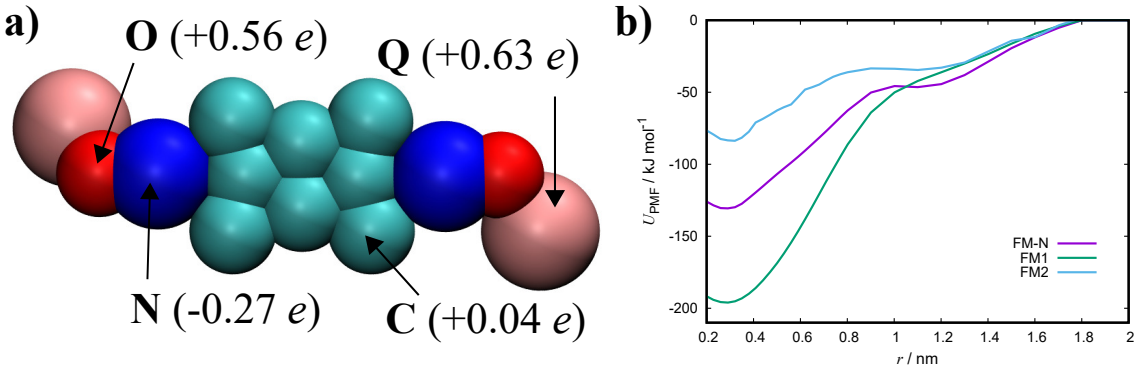


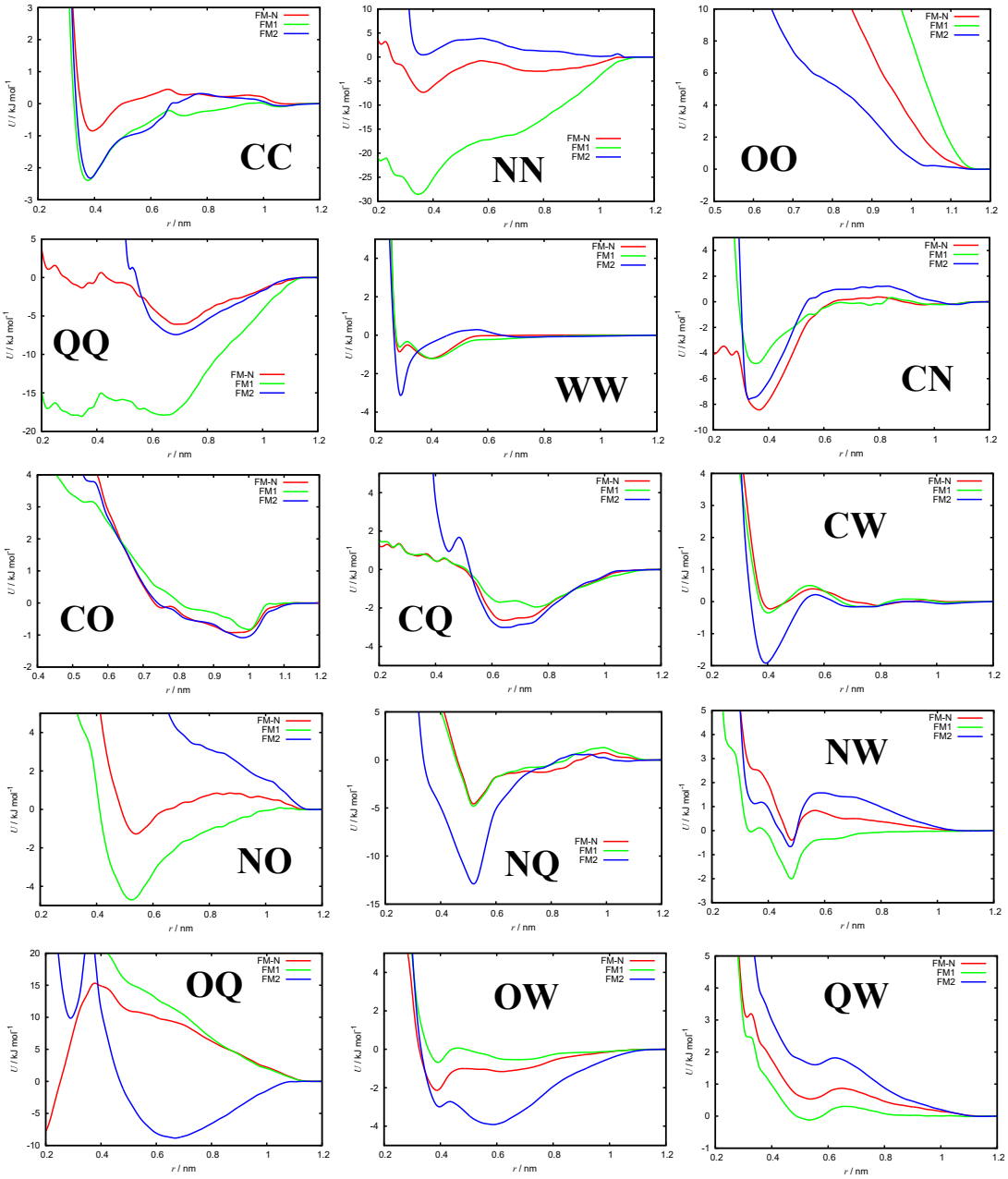
Figure 6.4: (a) Schematic diagram showing the mapping used for the FM CG models, with the definition of bead types and their partial charges. (b) PMFs for a dimer of each FM CG model.

Table 6.5: Free energies of association (ΔG_{assoc}) and hydration (ΔG_{hydr}) for the various FM models.

Model	FM-N	FM1	FM2
ΔG_{assoc} / kJ mol^{-1}	-131 $\pm$ 3	-195 $\pm$ 2	-84 $\pm$ 4
ΔG_{hydr} / kJ mol^{-1}	+85.8 $\pm$ 0.6	+130.2 $\pm$ 0.18	-79.9 $\pm$ 0.4

The free energies of association and hydration for each model are summarised in Table 6.5 and PMF profiles are shown in Figure 6.4(b). For all three FM CG models, the binding energies are much greater than value obtained for the AA model, a trait observed previously for CG models of chromonics developed by force matching.[139] The PMF profiles appear to have indistinct wells, with respect to the lack of a steep repulsive slope at short distances, which indicates that configurations of the dimer throughout the reaction coordinate are relatively favourable. The minimum of the PMFs occur at a COM distance of 0.35 ns, which is noticeably shorter than for the AA model. This may arise due to a number of reasons: i) the FM potentials are not of a fixed form and are generally softer than typical 12:6 LJ potentials, which allows for a greater degree of overlap, ii) imbalances between interactions on the peripheral beads could result in increased torsions manifesting as a

distortion in the dimer configuration (where it is noted that the orientation of the end groups affects the COM of molecule) and iii) the high attraction/hydrophobicity of the model results in close configurations to be preferred, even if they are strained. With the exception of the FM2 model, the free energies of hydration obtained are of the wrong sign which suggests an insolubility of the CG models in water. The association strength of a dimer can be said to depend on two factors: the attractive interactions between the molecules and the extent of solubility of the monomers. Thus, we can rationalise the trend in association free energies across the models with the magnitude/sign of the hydration free energies. As the solubility of the model decreases, the binding energy increases through the range.



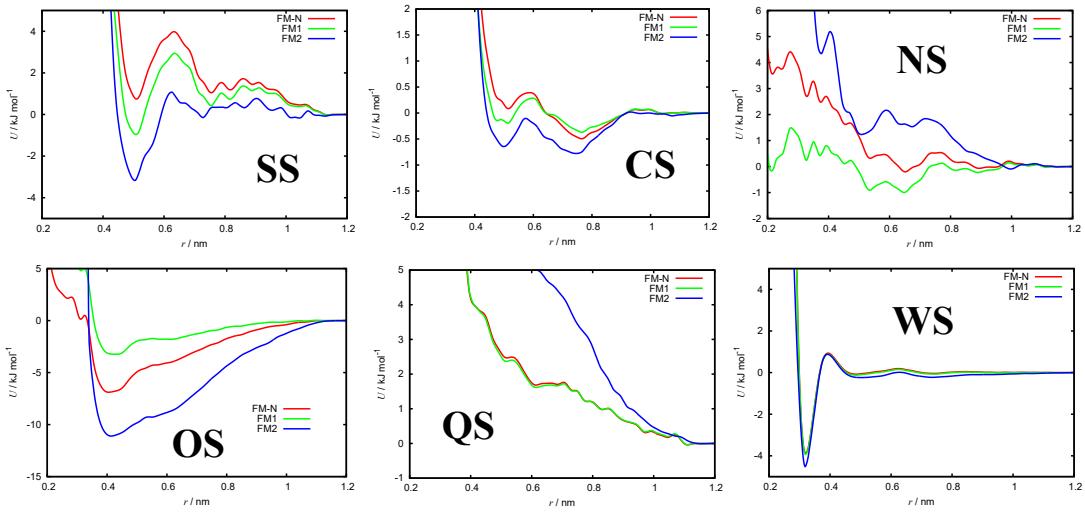


Figure 6.5: Coarse-grained potentials for the FM CG models, where W and S are the bead types for water and chloride ions, respectively.

MD simulations of all three FM CG models produced identical self-assembly behaviour. Within the first few nanoseconds, the system of dispersed monomers aggregates to form a single, dense cluster which does not fragment over the course of the resulting simulation, although some stacking of molecules can be discerned. Furthermore, the counterions condense around this aggregate rather than being dispersed in solution. Snapshots of this aggregate structure are presented in Figure 6.6 for the FM1 model, noting that the other models form aggregates that are visually indistinguishable. This unfavourable self-assembly can be rationalised from the poor balance between the large attractive interactions and insolubility observed in the free energies obtained. This combination results in a drive towards phase separation in lieu of chromonic self-assembly. The influence of cross interactions between components of the chromonic species and water on the morphology of aggregates has been discussed previously by Potter and co-workers.[138] The authors report that a subtle balance of interactions is required to obtain chromonic behaviour: if the association strength is too low, no chromonic stacks are seen; if the association strength is too high insoluble aggregates form; if there is an incorrect balance of hydrophilic-hydrophobic interactions between parts of the molecule and water, then aggregates form readily but without chromonic stacking.

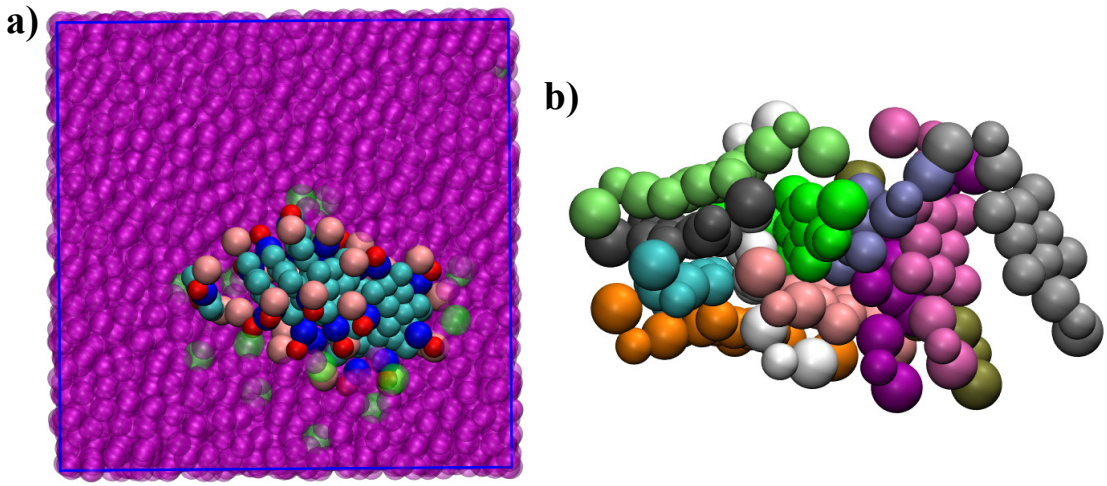


Figure 6.6: (a) A simulation snapshot from a 10 wt% system of 15 molecules for the FM1 model, where the purple and green beads represent water and chloride ions, respectively. (b) A focused view of the aggregate with individual molecules coloured differently and the solvent removed for visual clarity.

The FM-N model can be considered the vanilla model where any short-range electrostatic contributions are assumed to be effectively captured in the CG potentials and any long-range electrostatic effects are assumed to be screened out. However, this standard approach averages the electrostatic forces over all the beads and neglects the ionic nature of the dye. In contrast, the FM1 and FM2 models have partial charges which can explicitly interact and influence attraction/repulsion between beads. Regardless of charge treatment, the FM models have two interaction potentials that dominate their behaviour: the CC and CW interactions, although other pair interactions do contribute to a lesser extent to the molecular association. Between the FM-N and FM1 models, the CC interactions have identical well positions but differ in their interaction strength, where the FM1 CC interaction is 1.5 kJ mol^{-1} greater than for FM-N. The CW potentials are almost identical between these two models, as the separation of electrostatics from the FM potentials should have minimal effect on neutral beads. Therefore, we expect the FM1 model to have a greater association strength than FM-N, and this is indeed the case. It can be suggested that the failure to reproduce the atomistically observed self-assembly originates from the poor representation of the solubility, and this itself arises from i) a lack of screening effects from the fully neutral CG water, ii) solute–solvent interactions which are not favourable enough and iii) an emphasis on reproducing structure in the MS-CG procedure rather than thermodynamic properties.

The FM2 model was parametrised with a different reference trajectory, on the assump-

tion that a reference system containing only large stacks will produce CG potentials that solely seek to reproduce this behaviour, neglecting the existence of smaller species and the solubility of monomers. This second AA reference captures the full self-assembly process of monomers into small stacks and provides a richer ensemble of configurations. In essence, this is similar to a multistate approach to coarse-graining which employs an ensemble of references at different state points to improve the transferability of the resulting model.[186,224,408] Despite this alternate approach, the correct aggregation behaviour is not exhibited, as mentioned previously. The FM2 model shows increased solute–water interaction strengths and weaker/similar solute–solute interaction strengths compared to the FM1 model. It is noted that the FM2 model does not produce the double-well potential typically seen for CG water[210], whereas the FM-N and FM1 models do. This may arise from the dynamic nature of the solutes in the FM2 reference disrupting the longer range water-water interactions. The higher overall hydrophilicity of this model originates from the aromatic region of the molecules experiencing a solvated environment for a significant period in the reference trajectory before being shielded by the formation of stacks. Over the range of assessment criteria, the FM2 model demonstrates an improvement and even shows solubility in water. However, the model does not produce chromonic self-assembly and so cannot be deemed successful. It seems that there is a fundamental issue in representing the correct aqueous behaviour in FM CG models, which limits their ability to form discrete, soluble stacks. A potential approach to address this may be by the application of local density potentials[190,235,409,410], which could allow for the local density of water to be captured around a given bead depending on the status of the molecule as a monomer, or its position in the centre/end of a stack.

A third FM CG model was developed, but shall only be discussed briefly here. Using the extended ensemble approach, multiple references can be used in the parametrisation of CG potentials in order to improve their transferability.[186,224,408] Here, the FM3 model was parametrised based on two reference trajectories, where each reference was weighted equally. The first reference system contained monomers and some dimers, and the second reference consisted of large stacks. In principle, this is similar to the approach used for the FM2 model, but an explicit multistate framework would allow for the inclusion of more reference states and higher control over the contribution of each reference (i.e. configurations/stack sizes or temperature) to the CG potentials. The free energies of association

and hydration calculated for this model were -143 $\pm$ 4 kJ mol^{-1} and +95 $\pm$ 2 kJ mol^{-1}, respectively. Compared to the FM1 model, an improvement is seen in both these quantities but the correct chromonic self-assembly is still not observed.

6.3.3 Martini

The first top-down CG model to be presented here was developed using the Martini 2 framework, which has shown aptitude towards simulating chromonic self-assembly.[138] Specifically, the refined polarisable (refPOL and refION) Martini 2 force field was used here, which is designed for used with the PME method for long-range electrostatics.[207,208]

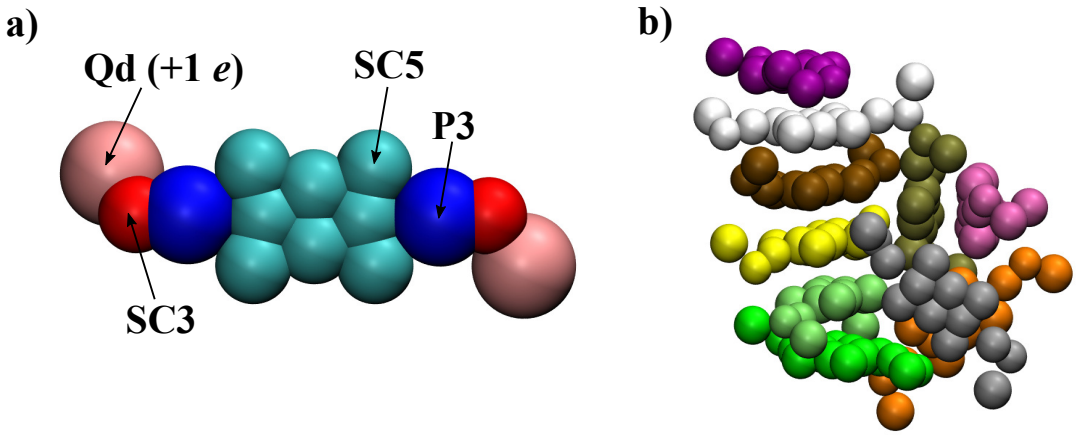


Figure 6.7: (a) Schematic for the Martini 2 model and (b) a simulation snapshot from a 10 wt% system of 15 molecules.

Figure 6.7 shows the Martini bead types selected for this CG model and a simulation snapshot. It is noted that the chloride ions are also used and carry an overall -1 charge. The association free energy of this model is -71 $\pm$ 2 kJ mol^{-1} at a COM distance of 0.5 nm, and its hydration free energy is -426 $\pm$ 4 kJ mol^{-1}. Both these thermodynamic quantities are overestimated, probably arising from the lack of specific parametrisation for aromatic beads, in this formulation, leading to an inaccurate representation of the partitioning behaviour of the core beads between the hydrophobic/aqueous environments. The structure obtained via MD simulation shows the stacking of molecules, but also displays unfavourable aggregation of molecules to the side of the stack. The large intermolecular spacing between adjacent molecules in the stack arises from the scarcity of bead sizes in Martini 2 (particularly for aromatic segments to exhibit π stacking) to represent chemical

groups of various sizes.

The new Martini 3 force field is updated with a wider choice of bead sizes and interaction levels, in addition to a reparametrisation of the bead types to better reproduce partitioning behaviour. An initial Martini 3 CG model (M3) was developed using the recommended bead types to represent the various chemical groups (see Figure 6.8(a)). The self-assembly observed for this out-of-the-box model is in good agreement with the AA system, with the formation of stacks in solution. The free energy of association and hydration were calculated to be -22 $\pm$ 2 kJ mol^{-1} (at a favoured COM distance of 0.39 nm) and -101.8 $\pm$ 0.5 kJ mol^{-1}, respectively. While the binding energy is approximately half of the target AA value and the hydration free energy is slightly lower, a favourable balance of hydrophobic-hydrophilic interactions is attained and this is evidenced by the observed chromonic self-assembly. Furthermore, the COM distance for the stacked configuration exactly matches that of the atomistic system, demonstrating the efficacy of the bead sizes available in this formulation. A degree of control can be exerted on Martini CG models to improve their representation of thermodynamic properties. For the free energy of association, this can be achieved by varying the polarity of the bead types, where it is noted that this concurrently affects the free energy of hydration. Figure 6.8(b) summarises the effect of varying the two bead types representing the mesogenic core from n = 1 to 5 for both TCn and Pn. As expected, increasing the hydrophilicity of the core results in a weakening of the molecular association by way of increasing the solubility of the model.

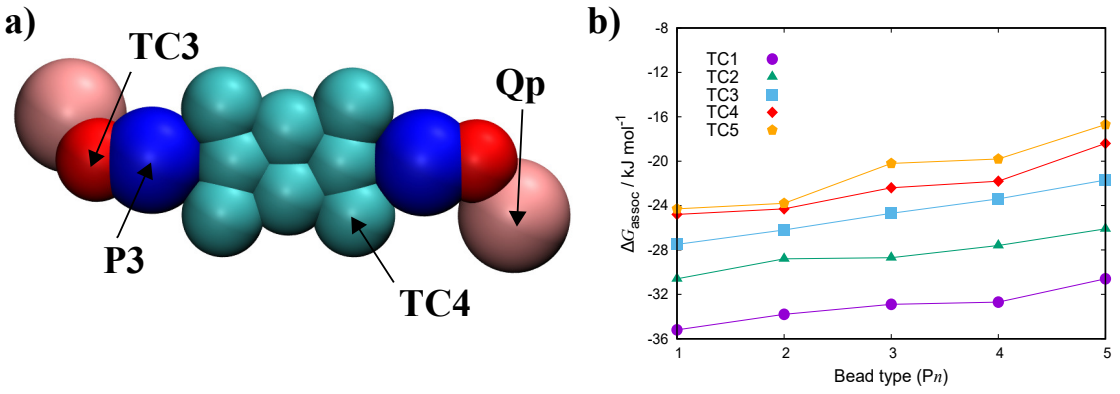


Figure 6.8: (a) Schematic diagram for the Martini 3 model (M3) showing the bead types used. Water is modelled using the WN bead type and chloride ions are represented by TQ1 beads. (b) ΔG_{assoc} as a function of varying bead types for the core. TCn, in the key, refers to the inner core beads and Pn, on the x-axis, denotes the outer core beads.

Using the insights from Figure 6.8(b), we can obtain even finer control on the thermo-

dynamic properties by tuning the Martini 3 LJ interactions for the core–water and core–core beads. A simple optimisation procedure was applied as follows. Firstly, the TC4–WN interaction is optimised to reproduce the AA free energy of hydration. Secondly, the TC4–TC4 interaction is optimised to reproduce the AA free energy of association. By only using one pair potential to optimise each thermodynamic property, a linear relationship between the LJ well depth and the magnitude of the respective free energy is obtained. Thus, a custom bead type was developed for the M3 model which is denoted TCp, where ϵ(TCp–WN) = 1.945 kJ mol^{-1} and ϵ(TCp–TCp) = 2.998 kJ mol^{-1}, and σ for both potentials is unchanged. For reference, the original parameters were ϵ(TC4–WN) = 1.19 kJ mol^{-1} and ϵ(TC4–TC4) = 1.45 kJ mol^{-1}. The optimised M3 model exhibits the correct chromonic self-assembly (see Figure 6.9(b)), also captured by the initial M3 model, but the optimised version exactly reproduces the target thermodynamic properties. Figure 6.9(a) shows the PMF profiles for a dimer, trimer and tetramer of the optimised M3 model. It is found that aggregation here is isodesmic due to the binding energies of a dimer, trimer and tetramer being approximately the same, although it is noted that the dimer binding energy is slightly higher than for trimer/tetramers which could be attributed to quasi-isodesmic aggregation.

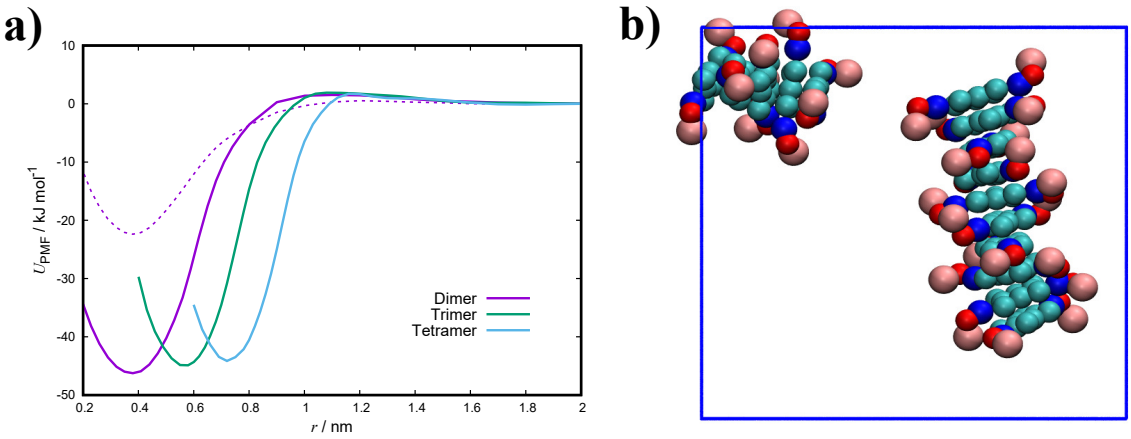


Figure 6.9: (a) PMFs for a dimer, trimer and tetramer of the optimised M3 model, where the dashed curve corresponds to the initial M3 model. (b) A simulation snapshot from a 10 wt% system of 15 molecules.

6.3.4 Combined models

It was demonstrated that FM CG models exhibit the incorrect behaviour in aqueous solution, owing to its insolubility with high values for ΔG_{assoc}. This section presents a proof-of-concept approach which combines CG potentials from force matching and Mar-

tini. The hypothesis here is to examine whether FM potentials can inherently reproduce the reference structures and if this is being hindered by the observed insolubility. Thus, combined models can be developed in which the interactions between chromonic mesogens are described by FM potentials, but solute–solvent and solvent–solvent interactions are represented using Martini, which was shown to correctly capture chromonic self-assembly in aqueous solution. Similar approaches have been reported to mix atomistic solutes with Martini solvent using virtual sites[390–392] and combining bottom-up CG potentials with Martini water.[205] This model, FM-M3, utilises FM potentials from the FM1 model for solute–solute interactions and all other interactions are described using Martini 3, as summarised in the interaction matrix (Table 6.6). CG sites for the chromonic dye possess partial charges which are treated explicitly with the reaction-field method. All other simulation parameters for this model were the same as those used for the Martini 3 model and the thermodynamic state points for the simulations are unchanged, noting that the CG water here uses a 4:1 mapping. All CG potentials were tabulated so that FM potentials have a cutoff of 1.2 nm, whereas all Martini 3 potentials have the usual cutoff of 1.1 nm.

Table 6.6: Interaction matrix for the combined force matching/Martini model (FM-M3). Bold entries denote potentials obtained from force matching (see Figure 6.4(a) for bead definitions), whereas normal entries correspond to Martini 3 interactions.

Bead type	C	N	O	Q	W	S
C	**CC**	**CN**	**CO**	**CQ**	TC4/WN	TC4/TQ1
N	–	**NN**	**NO**	**NQ**	P3/WN	P3/TQ1
O	–	–	**OO**	**OQ**	TC3/WN	TC3/TQ1
Q	–	–	–	**QQ**	Qp/WN	Qp/TQ1
W	–	–	–	–	WN/WN	WN/TQ1
S	–	–	–	–	–	TQ1/TQ1

The free energy of association obtained for the FM-M3 model is -115 $\pm$ 3 kJ mol^{-1} at a COM distance of 0.36 nm (see Figure 6.10(a) for the PMF). While this binding energy is still much higher than the target AA value, it shows an improvement compared to the FM models. The free energy of hydration for this model is -149 $\pm$ 2 kJ mol^{-1}, which shows an immense improvement to the FM models and is comparable to the AA value. This indicates that this model should be appropriately soluble. The self-assembly of this shows the formation of chromonic stacks in good agreement with AA system and M3 model (see

Figure 6.10(b)). The recovery of the correct self-assembly using the new combined model demonstrates that it is the solute–solvent interactions that are the major weakness in the FM CG models, and that the FM potentials for the chromonic mesogen can intrinsically exhibit the correct structures and self-assembly behaviour. Obtaining the correct aggregation behaviour for chromonic systems in a CG model is highly sensitive to the controlling thermodynamics, particularly its solubility, more so than structural properties.

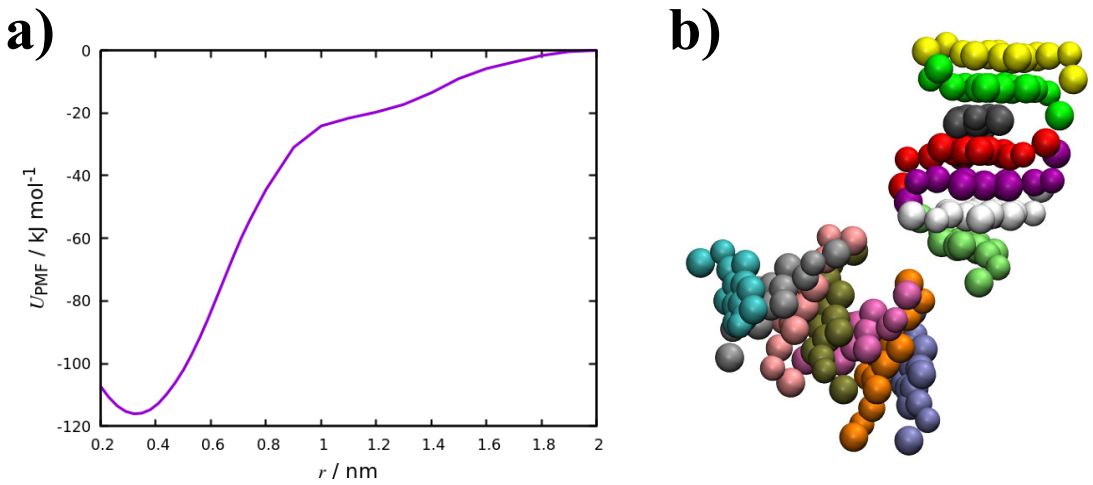


Figure 6.10: (a) PMF for a dimer of the FM-M3 model and (b) a simulation snapshot from a 10 wt% system of 15 molecules with individual molecules coloured differently for clarity.

The optimisation procedure applied to the Martini 3 model can, in principle, be applied to the combined model. This can be done in two ways by tuning the core–water interaction or simply changing the bead type. This approach is limited for the combined model as the FM potentials cannot be modified without reparametrisation, and so both $\Delta G_{\mathrm{assoc}}$ and ΔG_{hydr} would have to be optimised through the variation of the solute–solvent interactions alone. It is found that a reduction in the binding energy can be achieved by changing the Martini bead types to more hydrophilic/polar definitions (such as TN0 or TP4), but this concurrently increases the hydration free energy. Thus, the resulting systems have a poorer balance between these properties and the self-assembly of monomers into stacks is disrupted, resulting in dimers/trimers being the largest species observed in MD simulations.

6.3.5 Simulating chromonic liquid crystal phases

An optimised Martini 3 CG model, which reproduces all the relevant structural and thermodynamic properties, for PER was presented in the previous section. This model was selected for the simulation of chromonic liquid crystal phases. MD simulations of systems containing 1000 molecules (starting from a random configuration in solution) of PER were performed at two concentrations, 10 wt% and 30 wt%, for 5 μs at 300 K. Figure 6.11 shows snapshots from these simulations, where organisation into the chromonic N phase is not observed. Instead, large flexible stacks form which can be continuous over the periodic boundaries. It seems that the higher concentration of 30 wt% does not assist the formation of a phase, but does result in more rigid stacks which are packed more closely. It could be suggested that reproduction of target properties at a small scale does not necessarily translate into the correct behaviour being observed at a larger scale.

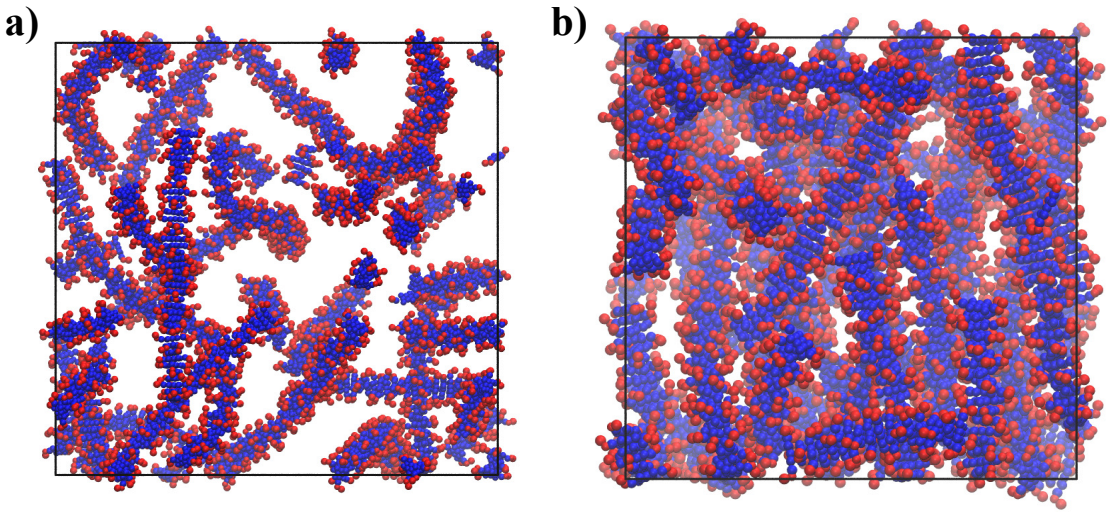


Figure 6.11: Simulation snapshots from a system of 1000 molecules at (a) 10 wt% and (b) 30 wt%, where core and chain beads are coloured in blue and red, respectively.

It has been found that chromonic mesogens favour the formation of linear columns, with no long range order, in MD simulations.[136,139] The lack of self-assembly into phases seems to arise from the preference of aggregates to interact across the periodic boundaries, resulting in a difficulty in dynamically reorienting in solution. In the first study[136], a magnetic field was applied to promote alignment of molecules to equilibrate phases, whereas the second study[139] initiated simulations from seeded columns oriented in a common direction. The latter approach was tested here. An initial box was set up with 9 columns, each containing 15 molecules with adjacent molecules having a 60° rotation, oriented in

the z direction. This lattice was solvated with water and counterions to produce three systems at 10 wt%, 30 wt% and 50 wt%. An initial equilibration run was performed with position restraints on the chromonic mesogens to first equilibrate the solvent without disrupting the seeded arrangement, prior to a 50 ns equilibration using production simulation conditions and semi-isotropic pressure coupling, to allow the x/y and z box dimensions to vary independently. This system was then replicated twice in each dimension (resulting in a 1080 molecule system of 36 columns containing 30 molecules each) and a 500 ns simulation was performed.

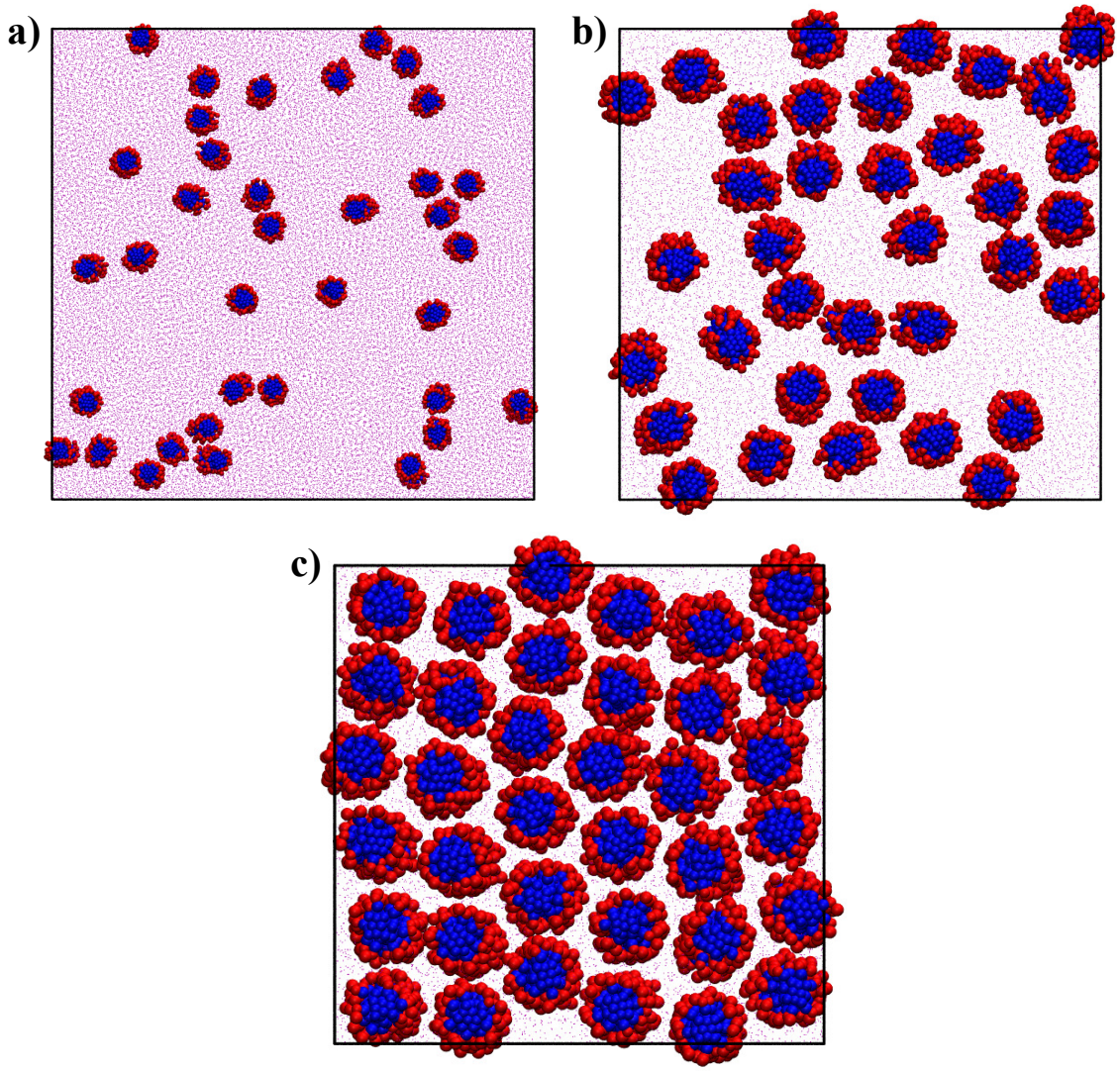


Figure 6.12: Simulation snapshots of systems started from a seeded arrangement at (a) 10 wt%, (b) 30 wt% and (c) 50 wt%, where the purple dots represent water.

Due to the infinitely long nature of the columns, it is expected that the chromonic N phase is prevalent even at low concentrations as the transition to an isotropic phase would require the stacks to break and rotate. However, this methodology proves useful in ac-

cessing the expected behaviour at moderate concentrations (where the N phase is present) and to study the transition to chromonic M phases. The average order parameter is calculated to be >0.95, regardless of concentration. Experimentally, the order parameter for the chromonic N phase is found to 0.97 at its upper concentration limit, which is in good agreement with the simulations.[404] Simulations at 10 wt% and 30 wt% yield the chromonic N phase whereas at 50 wt%, the chromonic M phase is observed. This is also in good agreement with experimental findings, where the chromonic N phase occurs in the range of 7 wt% to 30 wt% at room temperature, and a chromonic M phase is indicated to occur at higher concentrations.[404] The simulation at 30 wt% is at the upper concentration limit of the N phase, in which some hexagonal-like ordering can be seen. This becomes the dominant ordering mode at 50 wt% which can be distinguished clearly by visual inspection.

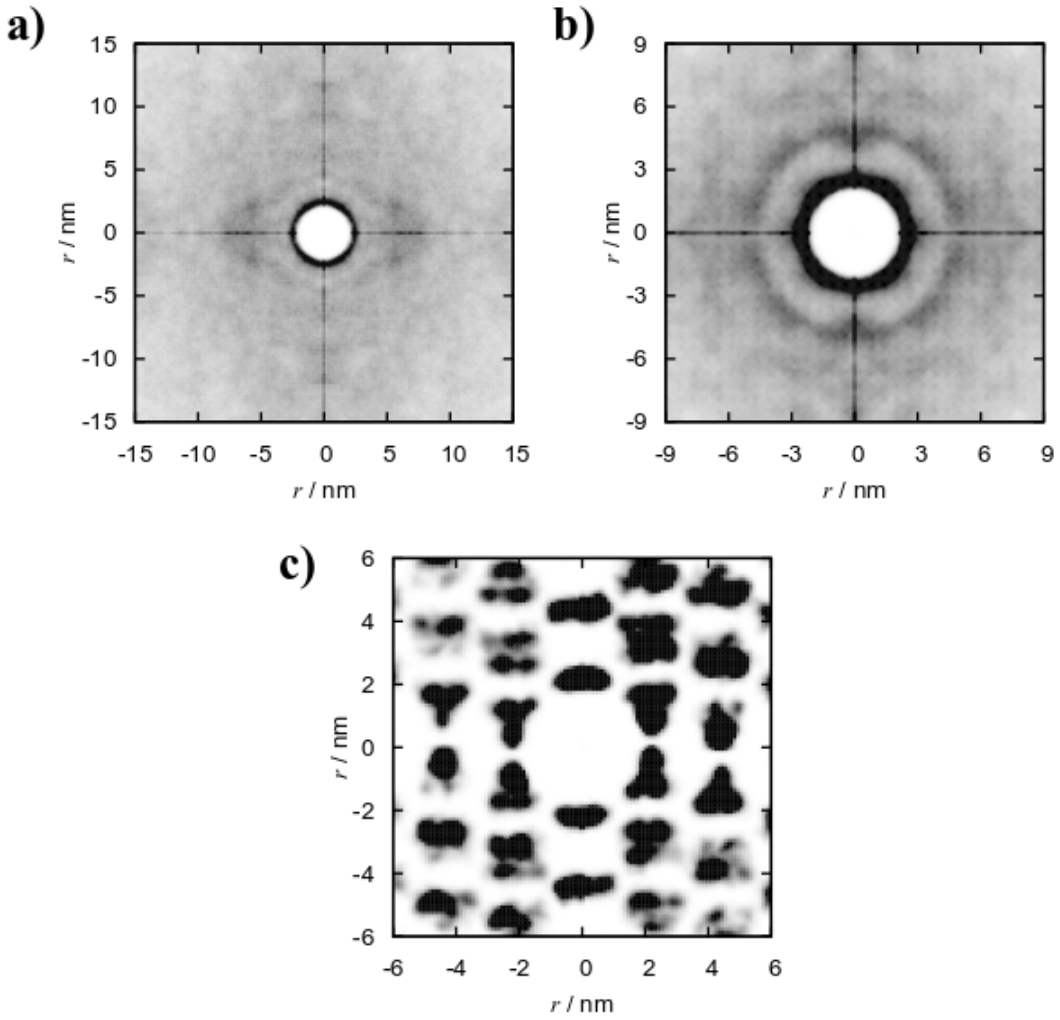


Figure 6.13: Two-dimensional pair distribution functions for systems at (a) 10 wt%, (b) 30 wt% and (c) 50 wt%.

The two-dimensional pair distribution function, $g(u, v)$, measures the degree of positional order present. A pair of orthogonal vectors, $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$, were defined as normal to the director of the system, $\hat{\boldsymbol{n}}$. A horizontal cutoff of half the box was applied and a vertical cutoff of 0.1 nm was used, so that only molecules within the same horizontal plane are considered. The vector between the COMs of pairs of molecules within these cutoffs was projected along $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ to obtain the distances, u and v. It is noted that these three plots have an empty area centred around zero corresponding to the width of a column (and the required water to solvate the column), where another column cannot be located. From these functions for systems at 10 wt% and 30 wt% (see Figure 6.13), a lack of positional order between columns is found. This, combined with the high order parameter, suggests that the chromonic N phase is exhibited. These two plots have a 'ring' of order in which an adjacent column is present at a spacing of $\sim$2.5 nm, but with no correlation beyond this; however, the system at 30 wt% shows a weak signature corresponding to the presence of another column at $\sim$5 nm. At 50 wt%, positional order corresponding to hexagonal packing is observed, with an intercolumn spacing of $\sim$2.3 nm.

6.3.6 Thermotropic analogue

Introduction

Thus far, work on the coarse-graining of a chromonic perylene bisimide dye has been presented, where bottom-up coarse-graining methods fail to adequately reproduce the self-assembly behaviour in aqueous solution. The underlying reason suggested here is that the force matching method produces CG models that are highly insoluble (not capturing the correct solute–solvent interactions) which results in strong, indiscriminate aggregation of the mesogens rather than chromonic self-assembly of stacks. This section presents the application of bottom-up coarse-graining methods to a thermotropic analogue, under the hypothesis that the difficulties of coarse-graining will be diminished with this mesogen.

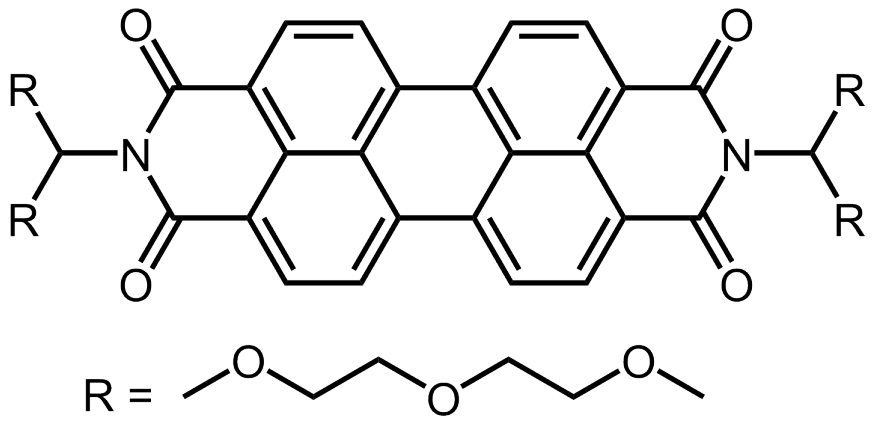


Figure 6.14: Molecular structure of PER_{OEG}.

The molecule of interest is composed of a perylene bisimide core with swallow-tail substituents of oligoethylenglycolether (OEG) chains, which will be known as PER_{OEG} (see Figure 6.14). It is found that this mesogen exhibits thermotropic discotic liquid crystal behaviour, manifesting the columnar-hexagonal (Col_h) phase in a broad temperature range (351–419 K).[411] Other thermotropic derivatives featuring this swallow-tail moiety, linear chains and combinations of these as substituents have been reported.[65,412,413] Notably, it is found that the incorporation of oligoether chains is necessary for liquid crystalline phases to be exhibited, as analogues with alkyl substituents tend to form crystalline structures. It is suggested that the higher conformational flexibility of the oligoether chains is crucial in suppressing crystallinity and promoting long-range order.[411,413]

With the absence of a solvent, only the interactions between the mesogens themselves need to be considered. This immediately reduces the number of potentials required to represent the system. Furthermore, this removes the need to achieve a complex balance of interactions between the components within the mesogen and a solvent. This section begins by presenting an all-atom simulation of PER_{OEG} which serves as the atomistic reference for coarse-graining. The bottom-up iterative Boltzmann inversion (IBI) and MS-CG methods were then used to parametrise CG models. Finally, the CG models are demonstrated to successfully capture the Col_h phase and the resulting systems are compared in their structural accuracy.

Atomistic simulation

Unless otherwise stated, the simulation protocol and parameters are identical to those used previously for the chromonic analogue, for both atomistic and CG simulations. Starting

from a random configuration of monomers, only small stacks formed. These clustered into a single aggregate over the course of a 200 ns simulation. It was found that this behaviour persisted at high temperatures and cooling did not promote the formation of the intended Col_h phase. Therefore, a seeded arrangement of 7 columns in a hexagonal lattice was used to initiate a simulation. This system was equilibrated for 200 ns at 360 K and utilised semi-isotropic pressure coupling. Finally, a 100 ns simulation was performed to collect data for analysis and to serve as the atomistic reference for coarse-graining. Visual inspection of the resulting system (see Figure 6.15) indicates that a Col_h phase is stabilised. Analysis of this system gives an intermolecular stacking distance (c) of 0.36 $\pm$ 0.03 nm and an intercolumn spacing (a_{hex}) of 2.0 $\pm$ 0.2 nm. X-ray diffraction measurements are concordant with these values, where c = 0.348 nm and a_{hex} = 2.081 nm.[411]

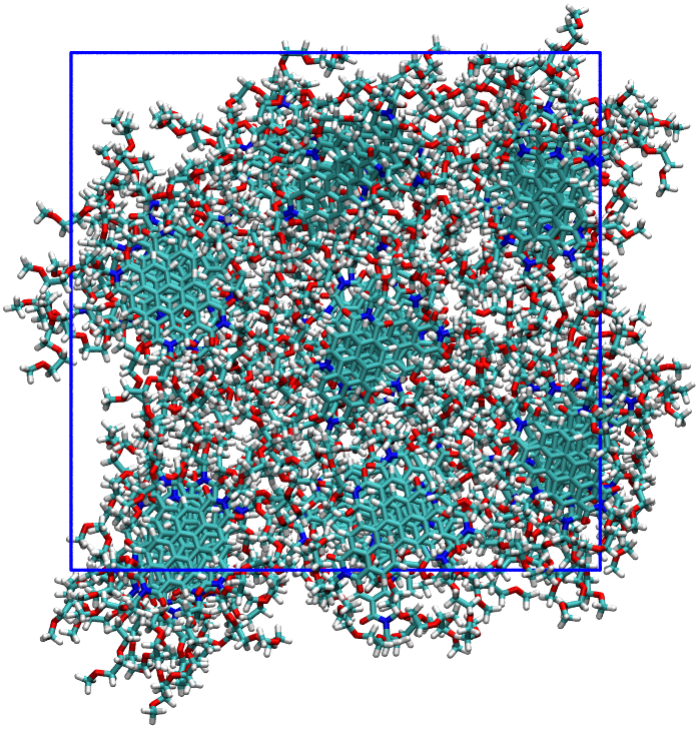


Figure 6.15: Simulation snapshot of a 70-molecule system of PER_{OEG} at 360 K, started from a seeded arrangement.

Coarse-grained models

The CG mapping used for PER_{OEG} consists of 4 bead types for a total of 10 non-bonded potentials to be parametrised (see Figure 6.16). The mapping and bonded parameters for the perylene bisimide core are identical to PER. The swallow-tail moiety used two bead types, where the linking isopropyl group forms one bead (A) and an oligoether chain is

divided into two beads (B). It is noted that the two B beads in the oligoether chains contain a different number of oxygen atoms, but the assumption that they are the same bead type allows for a decrease in the number of total bead types (which is beneficial for IBI). The bonded parameters are summarised in Tables 6.7, 6.8 and 6.9, noting that parameters for the core have been defined previously.

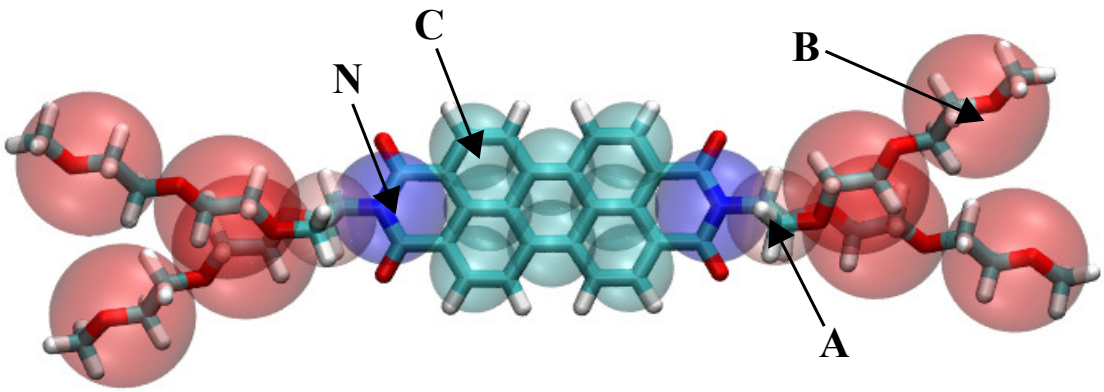


Figure 6.16: Coarse-grained mapping scheme for PER_{OEG}.

Table 6.7: Bond parameters for the coarse-grained models.

Atoms	Length / nm	k_r / kJ mol^{-1} nm^{-2}
N A	0.22	10000
A B	0.44	10000
B B	0.39	10000

Table 6.8: Angle parameters for the coarse-grained models.

Atoms	Angle / deg	k_θ / kJ mol^{-1} deg^{-2}
CM N A	180	500
N A B	105	500
A B B′	108	500
B A B′	142	500

Table 6.9: Dihedral parameters for the coarse-grained models.

Atoms	Dihedral / deg	k_ϕ / kJ mol^{-1}
CM N A B	85	25
N A B B′	0	25

RDFs were calculated from the reference trajectory containing 1000 frames, with exclusions applied such that intramolecular correlations are not included. The IBI method was employed to parametrise potentials (see section 2.5), where convergence of the CG and target RDFs required 300 iterations. This was carried out in the VOTCA-CSG package, version 1.4.1.[162,175,212] Additionally, the linear pressure correction was applied (see Equations 2.36 and 2.37) so that the system exhibited a pressure of 1 bar at the correct density.

The MS-CG method was also applied to construct CG potentials for PER_{OEG}. This utilised the same reference trajectory for IBI and the methodology is identical to that used for PER, with the exception of the treatment of electrostatics as PER_{OEG} is neutral. The CG potentials generated from both bottom-up CG approaches are presented in Figure 6.17.

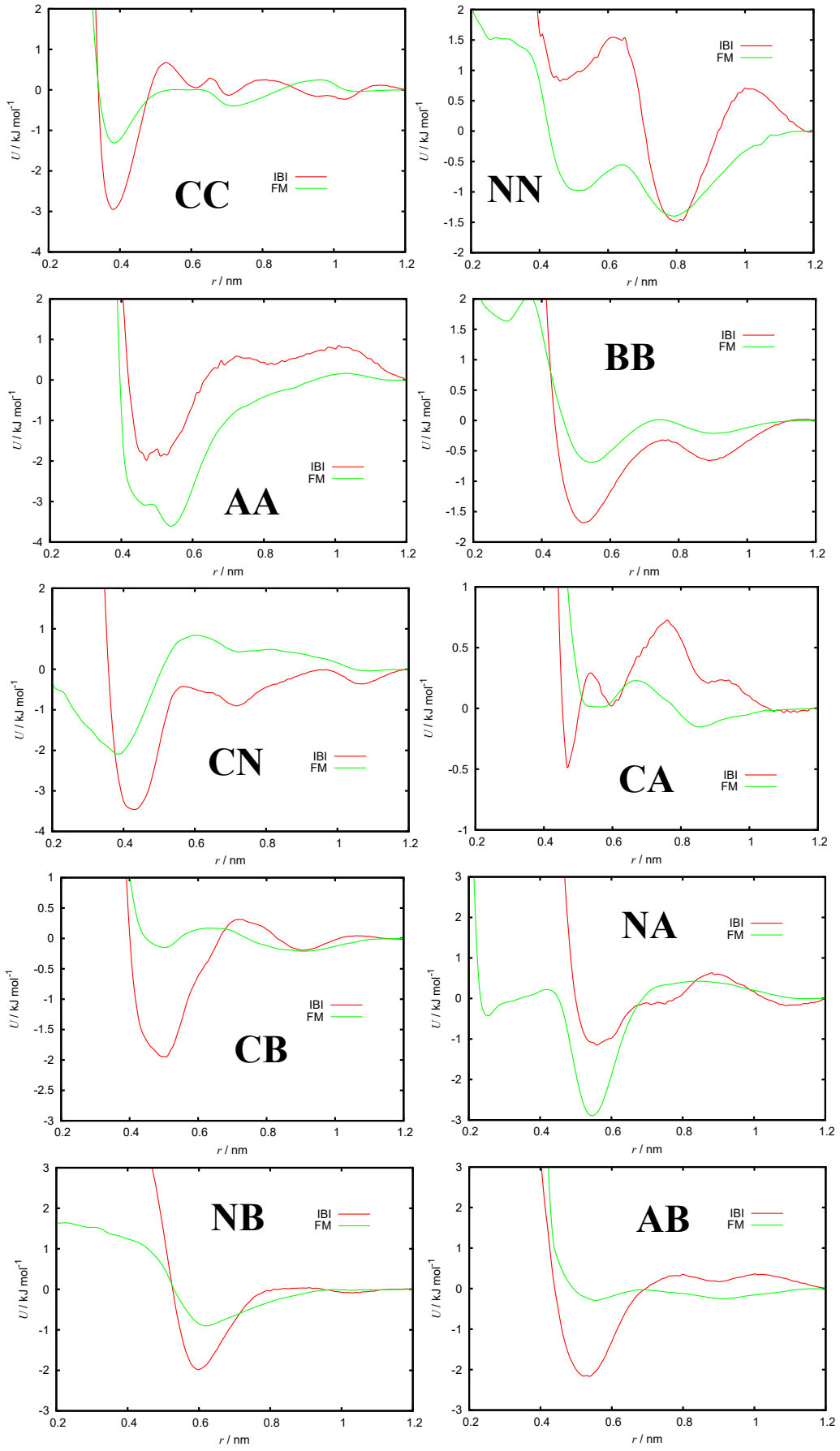


Figure 6.17: Coarse-grained potentials for the IBI and FM models of PER_{OEG}.

Coarse-grained simulations

For the IBI CG model, a 500 ns MD simulation of a 1000 molecule system starting from a random configuration of mesogens showed some stacking present but an overall isotropic nature (see Figure 6.18(a)). Further simulation of this system did not show any evolution into a more ordered state. The two CG potentials that should govern the association are the CC (core–core) and BB (arm–arm) interactions, where their interaction strengths are 3 kJ mol^{-1} and 1.7 kJ mol^{-1}, respectively. It seems that the relatively strong interactions between the components of the mesogen result in a high viscosity and low diffusion of molecules. This manifests as an observed preference for aggregation of the swallow-tail moieties (i.e. between bead types B), despite the stacking mode (CC) being more energetically favourable. The FM CG model largely suffers from the same issues (see Figure 6.18(b)). However, there is more extensive stacking present. Comparing the CG potentials, the FM model has weaker attractions, especially for the arm–arm interaction. This leads to molecules being able to move more freely and adopt the preferred stacked configuration. Despite this advantage, the FM CG model does not exhibit the formation of the target Col_h phase. It could be hypothesised that these models are capable of demonstrating the expected phase behaviour (in contrast to the bottom-up CG models for chromonics), but are still limited by the timescale of the isotropic–hexagonal transition.

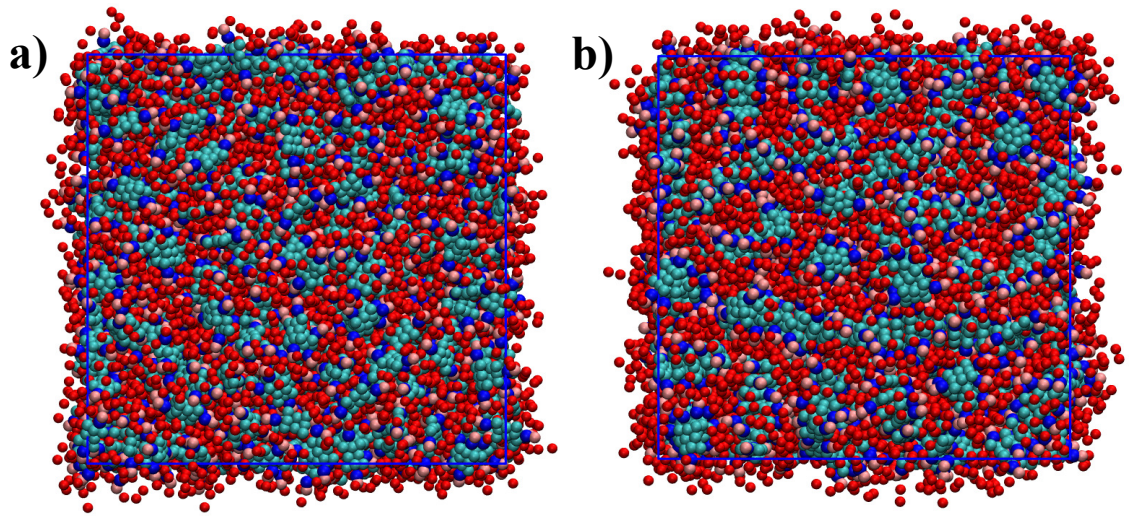


Figure 6.18: Simulation snapshots of a 1000 molecule system for the (a) IBI and (b) FM models of PER_{OEG} at 360 K.

Similar to the approach used for chromonic liquid crystal phases, a seeded arrangement was used to initiate new simulations. These consisted of a 70-molecule system (containing

7 columns of 10 molecules each) arranged in a hexagonal lattice. After a 50 ns equilibration, the system was replicated three-fold in each dimension to produce a 1890 molecule system, and a production simulation of 500 ns was performed (where all simulations used semi-isotropic pressure coupling). The resulting systems (see Figure 6.19) clearly show the columnar-hexagonal phase and were calculated to have average order parameters of 0.91 and 0.96 for the IBI and FM systems, respectively. Values of a similar magnitude have been found for an alkyl analogue with AA MD simulations.[414] It can be seen that the FM system appears more ordered, with respect to the packing of columns on a hexagonal lattice and the alignment of those columns along the director, which results in its higher order parameter. It can be suggested that the weaker arm–arm interactions of the FM model reduces their influence on the intercolumn packing and facilitates stronger intracolumn stacking and a higher freedom of movement between columns to adopt a closer packing.

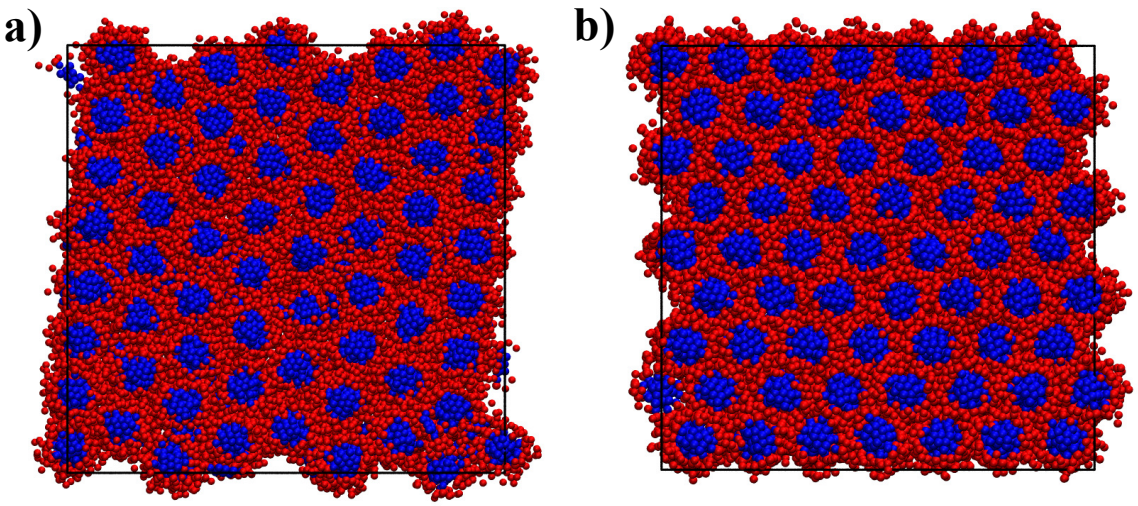


Figure 6.19: Simulation snapshots of systems started from a seeded arrangement for the (a) IBI and (b) FM models of PER_{OEG} at 360 K, where the core and arm beads are coloured in blue and red, respectively.

For the analysis of these systems, RDFs were calculated between the COMs of molecules. By determining the director for the system for each frame, $g(r)$ can be resolved into two components, $g_{\parallel}(r)$ and $g_{\perp}(r)$, which measures the positional order parallel and perpendicular to the director, respectively. From $g_{\parallel}(r)$, the intracolumn stacking distance (c) can be determined whereas $g_{\perp}(r)$ allows for the intercolumn spacing (a_{hex}) to be found. These plots are presented in Figure 6.20. From the evolution of the order parameter as a function of time, the two CG models have contrasting behaviour in the first 50 ns. For the IBI system, there is a decrease in S_2 to 0.91 whereas the FM system shows a very slight

increase ($S_2 \approx 0.96$), suggesting that the FM model exhibits a more crystalline packing of columns and the IBI model exhibits more fluidity between columns. For the remaining several hundred nanoseconds, both systems stabilise to a stable average value for S_2. The RDFs calculated for both systems suggest a crystalline/solid state, where peaks are spaced at intervals of c. However, this corresponds to the regular spacing of mesogens in the stacked configuration and so the systems are likely to be much more liquid-like than indicated. The more informative distribution functions resolve $g(r)$ into its parallel and perpendicular components relative to the director. Values of c = 0.350 nm and 0.362 nm were obtained for the IBI and FM systems, respectively. Both these quantities are in very good agreement with experimental measurements for the intracolumn spacing (c = 0.348 nm) with the value for the IBI model being the closest. The values of a_{hex} found were 2.08 nm and 1.82 nm for the IBI and FM systems, respectively. Here, the FM model produces too low a value, probably due to the closer packing of columns and its more crystalline nature. The IBI model, on the other hand, yields an intercolumn spacing that is identical to the experimental value of 2.081 nm. It can also be seen that the second 'peak' in $g_{\perp}(r)$ is a combination of two and three peaks for the IBI and FM systems, respectively. These correspond to the characteristic spacings for a hexagonal lattice, where the FM system exhibits a higher degree of order.

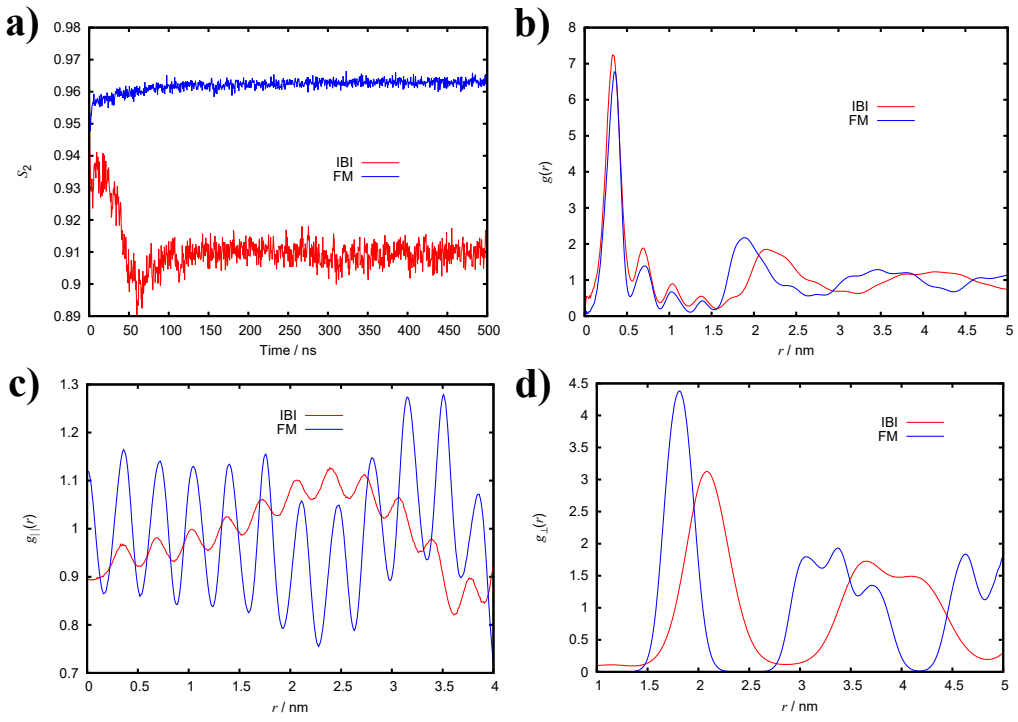


Figure 6.20: (a) Evolution of the order parameter, S_2, as a function of time. Plots of (b) $g(r)$, (c) $g_{\parallel}(r)$ and (d) $g_{\perp}(r)$ for the CG models of $\mathrm{PER_{OEG}}$.

Overall, both CG models perform well in simulating the Col_h phase with a 30x speed up compared to AA MD simulations. It is found that the IBI model better reproduces the experimental structure and the FM model is slightly too crystalline in terms of the positional order between columns. This is expected as the IBI method seeks to exactly reproduce the structure of the underlying system, whereas the MS-CG method reconstructs the structure implied in the force profiles. The MS-CG method has the advantage with respect to the ease of parametrisation as the CG potentials are generated with a single processing of the atomistic reference. In contrast, the IBI method requires an iterative procedure which is not guaranteed to converge and its feasibility decreases as the number of potentials to be parametrised increases. It is suggested, by the results here and in the literature[162,170,213], that IBI will produce models of higher structural accuracy over FM models in cases where an IBI model can be developed. In cases where an IBI model cannot be developed, the MS-CG method and resulting FM models are an adequate alternative. The main issue encountered here is that the timescale for the formation of the Col_h phase from a random configuration is still inaccessible. While the seeding of an ordered arrangement allowed for the simulation of the Col_h phase, this approach requires prior knowledge of the expected phase and could be problematic in studies where the prediction of a structure/phase is desired. To address this, replica exchange molecular dynamics (REMD), also known as parallel tempering, could prove useful. Here, multiple replicas of a system are simultaneously simulated and exchanges are performed between them to provide enhanced sampling.[415,416] By using REMD on states with different temperatures, the timescale for phase transitions in simulations of liquid crystals can be reduced.[417,418] In principle, any component of the system can be varied between replicas such as the interactions themselves.[419] This methodology has been applied to a Gay–Berne (GB) model, with exchanges between a second replica of a soft-core variant of the GB potential, which speeds up the equilibration of a liquid crystal phase.[420] This approach was tested on the IBI model of PER_{OEG} with the second replica employing a modified arm–arm interaction that has weaker attraction (see Figure 6.21). Starting from a random configuration, it was found that the order parameter increased faster for this system (compared to the normal simulation) and larger stacks formed more quickly. However, no organisation of the system into the Col_h phase was observed even after 2 μs of simulation.

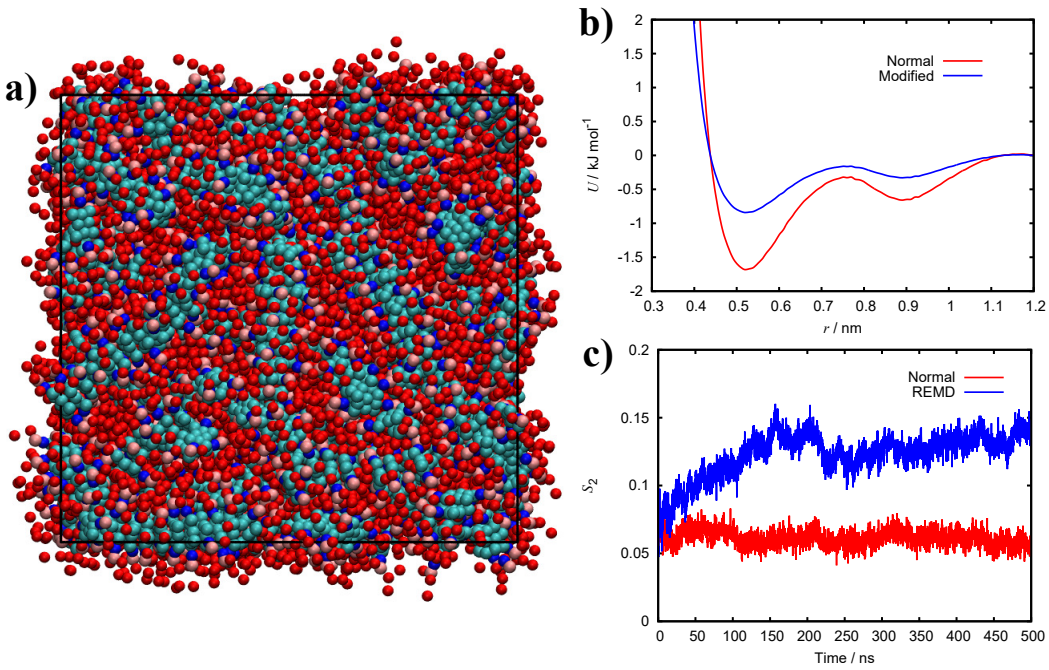


Figure 6.21: (a) Snapshot from the end of the REMD simulation of the IBI model for PER_{OEG}, showing the formation of columns but with no overall alignment into a Col_h phase. (b) CG potentials for the normal and modified BB interaction. (c) Evolution of the order parameter, S_2, as a function of time for the first 500 ns.

6.4 Summary

In summary, this chapter presented simulation studies and CG model development for perylene bisimide mesogens. AA MD simulations demonstrate the chromonic self-assembly of an ionic perylene dye, and allowed for the calculation of the free energies of association and hydration. The bottom-up MS-CG method was used to parametrise several CG models based on an atomistic reference. Multiple strategies were explored but none result in a FM CG model that satisfies the assessment criteria, where the binding energy is greatly overestimated. It is suggested that their failure to capture the correct solubility causes the poor representation of chromonic self-assembly. However, a combined framework using FM potentials for the chromonic mesogen (solute–solute) in conjunction with Martini 3 interactions for the solute–solvent/solvent–solvent potentials recovers the correct self-assembly behaviour in aqueous solution. This suggests that the FM CG models can intrinsically reproduce the expected structures, but the manifestation of this behaviour is hindered by inadequate representation of the solubility.

In contrast, a Martini 3 CG model is found to exhibit chromonic self-assembly despite

underestimating the binding energy and solubility. A simple optimisation procedure was applied to tune the Martini 3 CG model such that the relevant thermodynamic quantities are exactly reproduced. The optimised Martini 3 CG model was selected for simulating chromonic liquid crystal phases, and organisation of the system into the chromonic N and M phases is demonstrated. The occurrence of the expected phase at a given concentration and their structural characteristics are consistent with experimental findings.

A thermotropic discotic analogue, featuring swallow-tail substituents of oligoethyleng-lycolether chains, was also investigated. The absence of a solvent diminishes the difficulty in coarse-graining this mesogen, in comparison to chromonics. Two bottom-up methods, IBI and MS-CG, were used to parametrise CG models. The atomistic reference utilised for this consisted of a Col_h phase with structural characteristics that were in good agreement with experiment. Both resulting CG models demonstrate the organisation into the Col_h phase. Analysis of these systems reveal that the FM model produces a phase that is likely too crystalline, but its overall phase structure is still consistent with experiment. It is found that the IBI model is superior in capturing the Col_h phase, with its structural features exactly in line with experimentally measured values.

Chapter 7

Conclusions

This thesis presented all-atom and coarse-grained simulations of an assortment of liquid crystal systems ranging from the thermotropic twist-bend nematic phase and the discotic hexagonal phase, to (lyotropic) chromonic liquid crystals with different modes of self-assembly and organisation of phases. The focus of this project was to develop coarse-grained simulation models, propagating information from all-atom simulations through multiscale modelling approaches, that are chemically specific and allow for the study of hierarchical self-assembly in soft matter systems.

Chapter 3

In chapter 3, systematic coarse-graining methods were applied to CB7CB and benzonitrile-heptane systems to develop and assess the representability of the various CG models. The IBI, MS-CG and Martini 3 approaches were assessed for benzonitrile-heptane systems, where it was found that the bottom-up methods produced models with high structural accuracy (as indicated through the reproduction of RDFS) but severely underestimated thermodynamic properties such as ΔH_{vap} and ΔG_{solv}. Comparatively, the Martini 3 force field performed better in capturing thermodynamic properties, with the exception of densities, but slightly worse in the RDFs. It is suggested that a viable route to develop CG models with high representability and transferability is to optimise an initial Martini 3 force field to reproduce target observables.

The bent liquid crystal dimer CB7CB exhibits the twist-bend nematic phase, which features a heliconical structure despite the achiral nature of the mesogen. Here, extensive AA MD simulations are presented for CB7CB resulting in a full phase diagram with the

N_{TB}, N and I phases being observed. The simulated N_{TB} phase was characterised and provides values for the order parameter, helical pitch and conical tilt angle that are in good agreement with experiment. Distributions of the bend angle and chirality order parameter, χ, were determined, where it is found that the extent of conformational chirality between molecules in the N and N_{TB} phases is unchanged. Analysis of the χ distributions for the dimer series CBnCB (where n = 6, 7, 8 or 9) and CBX(CH$_2$)$_5$$Y$CB (where X/Y = CH$_2$, O or S) reveals two findings. Firstly, the degree of conformational chirality increases with the length of the spacer and, thus, the ability for a dimer to exhibit the N_{TB} phase is dependent on a bent molecular shape. Secondly, $|\chi|_{\max}$ increases as the bend angle decreases, but the flexibility of the dimers are similar.

Common coarse-graining approaches were applied to CB7CB and assessed in their ability to represent the phase behaviour of the N_{TB} phase. The IBI and FM CG models, parametrised based on an AA reference of the target phase, appear to exhibit the N_{TB} phase, whereas the Martini 3 model forms a more crystalline phase with smectic ordering. The differences between the phase behaviour observed for the CG models is discussed in terms of the nature and balance of the interaction potentials. The FM CG model exhibits phase transitions into the N and I phases upon heating and the spontaneous growth of the N and, subsequent, N_{TB} phases from cooling of the I phase. This demonstrates an excellent degree of transferability for this model and, combined with the good representability of the N_{TB} phase, suggests the MS-CG approach is suitable for the coarse-graining of thermotropic liquid crystals.

Chapter 4

In chapter 4, AA MD simulations of four ionic cyanine dyes (PIC, PCYN, TTBC and BIC) in aqueous solution were performed using an optimised force field. The thermodynamics of the self-assembly was studied using potentials of mean force to determine the free energies of association for n-mers (were n = 2, 3 or 4). Binding energies in the range of 8–15 $k_{\mathrm{B}}T$ were obtained, where it is found that a longer polymethine chain or higher polarisability of the cationic core results in a higher binding strength. Aggregation, here, was observed to be quasi-isodesmic, where dimerisation is slightly preferred and the binding energies for the trimer/tetramer are lower. Fitting of the association free energies as a function of temperature allowed for the decomposition into its enthalpic and entropic con-

tributions. This reveals that the driving forces for aggregation are attractive interactions between the aromatic cores.

Two classes of structure are observed in the self-assembly of the cyanine dyes in aqueous solution: H- and J-aggregates. PIC forms H-aggregate stacks with shift and Y junction defects, which incorporate the J-aggregation motif into the structures. PCYN forms rigid H-aggregate stacks that are free of defects. TTBC exhibits a J-aggregate structure of unimolecular thickness with a brickwork arrangement of molecules. BIC, similarly, displays the same J-aggregate structure, but shows an additional preferred antiparallel arrangement between adjacent molecules such that the charged sulfonate groups lie on opposite sides. Overall, the assemblies observed here are suggested to precede large-scale aggregation into tubular architectures, but allow for insights into the association at the molecular level to be gained.

Chapter 5

In chapter 5, CG models of the cyanine dyes in aqueous solution are developed and assessed in terms of their representability. The IBI method was applied to PIC to compare the effect of three different mapping schemes on their behaviour. It was found that all three models exhibit the same dimer structure in the PMFs, in which the binding energies are greatly overestimated. The free energies of hydration were determined and are of the wrong sign which indicates that the models are insoluble. The two lower resolution IBI models form cylindrical micelles, where the extent of stacking in the tube walls increases with a finer mapping. The highest resolution CG model forms long stacks with single- and double-molecule cross-sections.

The MS-CG method was employed to parametrise a FM CG model of PCYN. A single aggregate in water is formed and it is suggested that the insolubility of the model contributes strongly to this incorrect self-assembly behaviour. Surprisingly, a top-down Martini 3 model shows the same unfavourable aggregation behaviour despite the solution thermodynamics being represented well. The investigation of two simple models, through a parameter sweep of the core–core and core–water interaction strengths, reveals that the presence of explicit solubilising groups assists in promoting chromonic self-assembly into stacks. Moreover, the correct balance of interactions, which results in the correct aggregation behaviour, is more difficult to achieve with a smaller core size.

A general CG model for TTBC, aimed at reproducing the AA J-aggregate structure, was constructed and exhibits the correct aggregation behaviour on a large scale. It is found that the thermodynamic properties are in good agreement with the AA model which suggests that targeting the correct self-assembly behaviour implicitly achieves the required balance of the interactions. A multiscale modelling approach using hybrid AA/CG models, with the two resolutions being coupled via virtual sites, allows for simulations of systems with good structural accuracy but with higher computational efficiency. It is suggested that this is a promising strategy for coarse-graining systems where a successful CG model cannot be easily developed, such as the cyanine dyes here.

Chapter 6

In chapter 6, approaches to coarse-graining perylene bisimide dyes are explored. Studies with an atomistic force field demonstrate the chromonic self-assembly into H-aggregate stacks and allowed for the calculation of the free energies of association and hydration. Using the MS-CG method, several strategies employing different AA references were tested, but none result in a FM CG model that satisfies the assessment criteria. It is suggested that their failure to capture the correct solubility causes the incorrect aggregation behaviour to be found. A combined framework, where FM potentials and Martini 3 interactions are used to represent the solute–solute and solute–solvent/solvent–solvent interactions, respectively, recovers the correct self-assembly behaviour in aqueous solution. This further supports the suggestion that the FM CG models can intrinsically produce the reference structures, but the manifestation of this behaviour is hindered by the inadequate solubility.

An out-of-the-box Martini 3 model is found to exhibit the correct chromonic self-assembly, despite underestimating the binding energy and solubility. This model was optimised by scaling the LJ parameters such that the relevant thermodynamic properties are exactly reproduced. The optimised Martini 3 model was utilised to simulate chromonic liquid crystal phases, and the organisation of the system into the chromonic N and M phases is demonstrated at the expected concentrations.

A thermotropic discotic analogue of the chromonic perylene dye, featuring swallow-tail oligoethylenglycolether substituents, was coarse-grained. The absence of a solvent diminishes the difficulty in coarse-graining this mesogen compared to chromonics. Using an AA reference of a $\mathrm{Col_h}$ phase, the IBI and MS-CG methods were applied to parametrise

CG models of this dye. Both CG models show the organisation into the Col_h phase in simulations of a large system, with structural characteristics that are in good agreement with experiment. It is found that the IBI model is superior in capturing the Col_h phase.

Summary

This thesis explored the development of chemically specific coarse-grained models for use in simulating hierarchical self-assembly in soft matter systems, with a particular focus on liquid crystals. Overall, it is found that no one common coarse-graining approach is superior in representing all the various classes of mesogens studied in this work. For the thermotropic N_{TB} phase formed by bent dimers, the MS-CG method produces the best CG model, with the IBI CG model also being acceptable, and the Martini 3 model producing the incorrect phase behaviour. On the other hand, the Martini 3 force field performs very well in modelling chromonic liquid crystal phases formed by a perylene dye in aqueous solution, but the IBI and FM CG models fail to produce the correct self-assembly behaviour. An IBI model is found to perform better in capturing the Col_h phase over the FM model of a thermotropic discotic liquid crystal.

All-atom simulations have also been ubiquitous in this project to supply data for the parametrisation/validation of CG models and to provide insights into the systems studied. For the cyanine dyes in aqueous solution, this yielded an assortment of interesting H- and J-aggregate structures that have not been reported in simulations previously. While a successful CG model of a cyanine dye was elusive in this study, a hybrid AA/CG approach allowed for multiscale modelling of these systems with good structural accuracy.

An integral part of the work in the thesis was the selection of CG mapping schemes for the various molecules studied. Mapping schemes in this work followed three general guidelines: CG beads should contain 2–5 heavy atoms; the symmetry of the molecule should be preserved; and each distinct chemical region should be assigned its own bead type. For the first rule, the precise number of (and which) heavy atoms to include in a CG bead is chosen such that the second and third rules are satisfied. Furthermore, an appropriate resolution should be decided upon depending on the level of detail required for the observables of interest (for example, the use of a finer mapping for aromatic units to better capture stacking). The preservation of symmetry assists in reducing the number of bead types in the final CG model, as well as decreasing the possibility of any imbal-

ances between bead sizes or connectivity problems. However, it has been reported that the importance of this rule is unexpectedly low.[421] The third rule allows for interactions between different chemical groups to be clearly defined. This is particularly relevant to amphiphiles, where hydrophobic and hydrophilic residues should be separated into different bead types rather than have a single bead type contain competing effects. This is a problem which mapping considerations could not resolve for the cyanine dyes studied here, as their solubility arises from the highly delocalised π system and cationic charge. It was difficult to select a CG mapping and define bead types to mimic this. Moreover, it seems that the anisotropic character of the cyanine dyes (which directs the formation of single-molecule cross-section stacked structures) is lost with the CG mapping, even with a 2:1 fine mapping of the mesogenic cores. It was concluded that the CG beads for these types of ionic chromonic dyes are too general (i.e. are not particularly hydrophobic or hydrophilic) which, combined with the spherically symmetric nature of the CG potentials, causes the incorrect self-assembly to be observed. It is suggested that the use of ellipsoids to represent the mesogenic cores with anisotropic potentials is a promising strategy to address these problems, as the definitions of bead types would be simplified (for example, a quinoline unit could be mapped to one ellipsoid instead of several different, spherical beads) and the molecules should inherently stack in aqueous solution.[377–380] For the chromonic perylene dye, the larger size of the molecule proved not to be detrimental to coarse-graining but beneficial. The larger aromatic core was more disc-like and better promoted stacking, and each chemical group in the mesogen could be clearly separated into their own bead type. This allowed an out-of-the-box Martini 3 CG model to exhibit chromonic self-assembly with no issues. It could be suggested that chromonic mesogens of a similar (or larger) size or of the same archetype could simply be modelled using the Martini 3 force field, following the outlined CG mapping guidelines, rather than turning to more complicated coarse-graining procedures. The thermotropic liquid crystals studied here had little to no issues arising from the CG mapping. For CB7CB, there was a choice between a 2:1 or 3:1 mapping of the biphenyl units, for which a 3:1 mapping was ultimately selected as it preserves the shape/symmetry of the mesogen and allows for higher computational efficiency. Compared to chromonic mesogens, CB7CB (and likely other liquid crystals of the same class) are much more forgiving when choosing the CG mapping and bead types. The nature of the interactions in these thermotropic systems are less complex. However, the

choice of mapping the cyano group to its own bead or incorporating it into the terminal phenyl unit proved to be paramount. When not defined explicitly, a smectic phase was exhibited, whereas the presence of the cyano CG bead promoted nematic behaviour. Thus, like the chromonic liquid crystals, distinct chemical groups must be defined as their own bead type, otherwise important interactions within the system may be omitted or poorly averaged into other interactions.

The use of the MS-CG method in this work revealed that it does not capture the solubility of chromonic liquid crystals, which hinders the correct self-assembly being observed. This probably arises because: the fully neutral CG water does not provide screening of charges or favourable electrostatic interactions; the effective pair potentials do not capture the interactions correctly; and the emphasis on reproducing structure rather than thermodynamic properties. This problem could be addressed in two ways. The first is to explore ways of better incorporating electrostatics into the MS-CG method and their treatment in the CG simulations. The second is to use local density potentials[190,235,409,410], which would allow for the local density of water (and other mesogens) about a given bead to be more suitably captured. This approach may circumvent the need for anisotropic potentials and ellipsoidal units for the cyanine dyes. For thermotropic liquid crystals, the use of local density potentials may also be advantageous. The representability of the CG models could be improved if the local density potentials integrate information, that was lost in the coarse-graining process, back into the model.

It was found that the representability of the thermotropic CB7CB and PER_{OEG} bottom-up CG models depended on the accuracy of the underlying AA model. It could be hypothesised that, if the coarse-graining methods were sufficiently improved, bottom-up CG models for chromonic liquid crystals would also have the same dependence. Thus, the development of bottom-up CG models could profit from either the use of more expensive AA force fields or force fields that are optimised to reproduce certain properties better. For the cyanine dyes, some optimisation was performed in this work. Further optimisations could target transition temperatures (where the phase transitions for the AA model of CB7CB were observed to occur at higher temperatures than experiment) through tuning of the non-bonded parameters or the flexibility within the molecule. It would also be interesting to explore the effect of using forces obtained from the use of a polarisable force field in the MS-CG method for application to chromonic liquid crystals.

On the outlook of this study, further work could add to the compendium presented here by applying various coarse-graining approaches to other soft matter systems to determine the viability of a method for a given system type. By understanding where coarse-graining methods succeed or fail, the appropriate approach can be selected at the onset of the study of a new system, equipped with the knowledge of any limitations or strengths of the method. Where the conventional implementation of a coarse-graining method fails, more advanced approaches in parametrising CG models, or the inclusion of additional functionalities in the simulation of the models, could be explored to address the deficiencies.

References

[1] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford University Press, Oxford, 1993.

[2] F. Reinitzer, *Monatsh. Chem.*, 1888, **9**, 421–441.

[3] O. Lehmann, *Z. Phys. Chem.*, 1889, **4**, 462.

[4] G. A. DiLisi, *An Introduction to Liquid Crystals*, Morgan and Claypool Publishers, USA, 2019.

[5] S. Chandrasekhar, *Liquid Crystals*, Cambridge University Press, Oxford, 1992.

[6] L. Onsager, *Ann. NY Acad. Sci.*, 1949, **51**, 627–659.

[7] L. A. Madsen, T. J. Dingemans, M. Nakata and E. T. Samulski, *Phys. Rev. Lett.*, 2004, **92**, 145505.

[8] S. Pieraccini, A. Ferrarini and G. P. Spada, *Chirality*, 2008, **20**, 749–759.

[9] D. Obadović, M. Stojanović, A. Bubnov, N. Éber, M. Cvetinov and A. Vajda, *J. Res. Phys.*, 2011, **35**, 3–13.

[10] F. Castles, S. M. Morris, E. M. Terentjev and H. J. Coles, *Phys. Rev. Lett.*, 2010, **104**, 157801.

[11] J. W. Goodby, M. A. Waugh, S. M. Stein, E. Chin, R. Pindak and J. S. Patel, *Nature*, 1989, **337**, 449–452.

[12] J. W. Goodby, M. A. Waugh, S. M. Stein, E. Chin, R. Pindak and J. S. Patel, *J. Am. Chem. Soc.*, 1989, **111**, 8119–8125.

[13] V. Fréedericksz and A. Repiewa, *Zeitschrift für Physik*, 1927, **42**, 532–546.

[14] V. Fréedericksz and V. Zolina, *Trans. Faraday Soc.*, 1933, **29**, 919–930.

[15] C. W. Oseen, *Trans. Faraday Soc.*, 1933, **29**, 883–899.

[16] F. C. Frank, *Discuss. Faraday Soc.*, 1958, **25**, 19–28.

[17] M. Schadt and W. Helfrich, *Appl. Phys. Lett.*, 1971, **18**, 127.

[18] M. F. Schiekel and K. Fahrenschon, *Appl. Phys. Lett.*, 1971, **19**, 391–393.

[19] R. A. Soref, *Appl. Phys. Lett.*, 1973, **22**, 165.

[20] P. T. Ireland and T. V. Jones, *J. Phys. E: Sci. Instrum.*, 1987, **20**, 1195–1199.

[21] I. Dozov, *Europhys. Lett.*, 2001, **56**, 247–253.

[22] R. Memmer, *Liquid Crystals*, 2002, **29**, 483–496.

[23] M. Cestari, S. Diez-Berart, D. A. Dunmur, A. Ferrarini, M. R. de la Fuente, D. J. B. Jackson, D. O. Lopez, G. R. Luckhurst, M. A. Perez-Jubindo, R. M. Richardson, J. Salud, B. A. Timimi and H. Zimmermann, *Phys. Rev. E*, 2011, **84**, 031704.

[24] C. Meyer, G. R. Luckhurst and I. Dozov, *Phys. Rev. Lett.*, 2013, **111**, 067801.

[25] D. Chen, J. H. Porada, J. B. Hooper, A. Klittnick, Y. Shen, M. R. Tuchband, E. Korblova, D. Bedrov, D. M. Walba, M. A. Glaser, J. E. Maclennan and N. A. Clark, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 15931–15936.

[26] V. Borshch, Y. K. Kim, J. Xiang, M. Gao, A. Jákli, V. P. Panov, J. K. Vij, C. T. Imrie, M. G. Tamba, G. H. Mehl and O. D. Lavrentovich, *Nat. Comuun.*, 2013, **4**, 2635.

[27] H. Takezoe and Y. Takanishi, *Jpn. J. Appl. Phys.*, 2006, **45**, 597–625.

[28] L. E. Hough, H. T. Jung, D. Krüerke, M. S. Heberling, M. Nakata, C. D. Jones, D. Chen, D. R. Link, J. Zasadzinski, G. Heppke, J. P. Rabe, W. Stocker, E. Körblova, D. M. Walba, M. A. Glaser and N. A. Clark, *Science*, 2009, **325**, 456–460.

[29] S. Shadpour, A. Nemati, N. J. Boyd, L. Li, M. E. Prévôt, S. L. Wakerlin, J. P. Vanegas, M. Salamończyk, E. Hegmann, C. Zhu, M. R. Wilson, A. I. Jákli and T. Hegmann, *Mater. Horiz.*, 2019, **6**, 959–968.

[30] S. Shadpour, A. Nemati, M. Salamończyk, M. E. Prévôt, J. Li, N. J. Boyd, M. R. Wilson, C. Zhu, E. Hegmann, A. I. Jákli and T. Hegmann, *Small*, 2020, **16**, 1905591.

[31] S.-Y. Jo, B.-C. Kim, S.-W. Jeon, J.-H. Bae, M. Walker, M. R. Wilson, S.-W. Choi and H. Takezoe, *RSC Adv.*, 2017, **7**, 1932–1935.

[32] B.-C. Kim, M. Walker, S.-Y. Jo, M. R. Wilson, H. Takezoe and S.-W. Choi, *RSC Adv.*, 2018, **8**, 1292–1295.

[33] J. S. Lintuvuori, G. Yu, M. Walker and M. R. Wilson, *Liquid Crystals*, 2018, **45**, 1996–2009.

[34] H. K. Bisoyi and S. Kumar, *Chem. Soc. Rev.*, 2010, **39**, 264–285.

[35] T. W. I. Wurzbach, J. Kirres, A. Kostidou, N. Kapernaum, J. Litterscheidt, J. C. Haenle, P. Staffeld, A. Baro, F. Giesselmann and S. Laschat, *Chem. Rev.*, 2016, **116**, 1139–1241.

[36] T. P. Silverstein, *J. Chem. Educ.*, 1998, **75**, 116–118.

[37] J. Lydon, *J. Mater. Chem.*, 2010, **20**, 10071–10099.

[38] L. Wu, J. Lal, K. A. Simon, E. A. Burton and Y. Y. Luk, *J. Am. Chem. Soc.*, 2009, **131**, 7430–7443.

[39] J. R. Henderson, *Phys. Rev. Lett.*, 1996, **77**, 2316–2319.

[40] J. R. Henderson, *Phys. Rev. E*, 1997, **55**, 5731–5742.

[41] J. R. Henderson, *J. Chem. Phys.*, 2009, **130**, 45101–45103.

[42] J. F. Hubbard, *Ph.D. thesis*, University of Leeds, 1997.

[43] V. R. Horowitz, L. A. Janowitz, A. L. Modic, P. A. Heiney and P. J. Collings, *Phys. Rev. E*, 2005, **72**, 041710.

[44] H. S. Park, S. W. Kang, L. Tortora, Y. Nastishin, D. Finotello, S. Kumar and O. D. Lavrentovich, *J. Phys. Chem. B*, 2008, **112**, 16307–16319.

[45] A. J. Dickinson, N. D. LaRacuente, C. B. McKitterick and P. J. Collings, *Mol. Cryst. Liq. Cryst.*, 2009, **509**, 9/[751]–20/[762].

[46] C. B. McKitterick, N. L. Erb-Satullo, N. D. LaRacuente, A. J. Dickinson and P. J. Collings, *J. Phys. Chem. B*, 2010, **114**, 1888–1896.

[47] F. Chami and M. R. Wilson, *J. Am. Chem. Soc.*, 2010, **132**, 7794–7802.

[48] A. Akinshina, M. Walker, M. R. Wilson, G. J. T. Tiddy, A. J. Masters and P. Carbone, *Soft Matter*, 2015, **11**, 680–691.

[49] R. Thind, M. Walker and M. R. Wilson, *Adv. Theory Simul.*, 2018, **1**, 1800088.

[50] G. Yu, M. Walker and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2021, **23**, 6408–6421.

[51] A. H. Herz, *Adv. Colloid Interface Sci.*, 1977, **8**, 237–298.

[52] G. Scheibe, *Angew. Chem.*, 1936, **49**, 563.

[53] G. Scheibe, *Angew. Chem.*, 1937, **50**, 212–219.

[54] W. J. Harrison, D. L. Mateer and G. J. T. Tiddy, *Faraday Discuss.*, 1996, **104**, 139–154.

[55] M. Kasha, *Radiat. Res.*, 1963, **20**, 55–70.

[56] E. E. Jelley, *Nature*, 1936, **138**, 1009.

[57] E. E. Jelley, *Nature*, 1937, **139**, 631.

[58] W. West and S. Pearce, *J. Phys. Chem.*, 1965, **69**, 1894–1903.

[59] M. R. Tomasik and P. J. Collings, *J. Chem.Phys. B*, 2008, **112**, 9883–9889.

[60] G. J. T. Tiddy, D. L. Mateer, A. P. Ormerod, W. J. Harrison and D. J. Edwards, *Langmuir*, 1995, **11**, 390–393.

[61] S.-W. Tam-Chang, J. Helbley, T. D. Carson, W. Seo and I. K. Iverson, *Chem. Commun.*, 2006, 503–505.

[62] K. Ichimura, T. Fujiwara, M. Momosea and D. Matsunaga, *J. Mater. Chem.*, 2002, **12**, 3380–3386.

[63] S.-W. Tam-Chang and L. Huang, *Chem. Commun.*, 2008, **17**, 1957–1967.

[64] M. Lavrentovich, T. Sergan and J. Kelly, *Liquid Crystals*, 2003, **30**, 851–859.

[65] F. Würthner, C. R. Saha-Möller, B. Fimmel, S. Ogi, P. Leowanawat and D. Schmidt, *Chem. Rev.*, 2016, **116**, 962–1052.

[66] A. Tolkki, E. Vuorimaa, V. Chukharev, H. Lemmetyinen, P. Ihalainen, J. Peltonen, V. Dehm and F. Würthner, *Langmuir*, 2010, **26**, 6630–6637.

[67] H. S. Park, A. Agarwal, N. A. Kotov and O. D. Lavrentovich, *Langmuir*, 2008, **24**, 13833.

[68] S. V. Shiyanovskii, T. Schneider, I. I. Smalyukh, T. Ishikawa, G. D. Niehaus, K. J. Doane, C. J. Woolverton and O. D. Lavrentovich, *Phys. Rev. E*, 2005, **71**, 020702.

[69] S. V. Shiyanovskii, O. D. Lavrentovich, T. Schneider, T. Ishikawa, I. I. Smalyukh, C. J. Woolverton, G. D. Niehaus and K. J. Doane, *Mol. Cryst. Liq. Cryst.*, 2005, **434**, 259–270.

[70] W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.

[71] M. Karplus and J. A. McCammon, *Nat. Struct. Biol.*, 2002, **9**, 646–652.

[72] E. Brini, E. A. Algaer, P. Ganguly, C. Li, F. Rodriguez-Ropero and N. F. van der Vegt, *Soft Matter*, 2013, **9**, 2108–2119.

[73] D. Nicolaides, *Mol. Simul.*, 2001, **26**, 51–72.

[74] R. D. Groot and P. B. Warren, *J. Chem. Phys.*, 1997, **107**, 4423–4435.

[75] S. Chen and G. D. Doolen, *Annu. Rev. Fluid Mech.*, 1998, **30**, 329–364.

[76] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.

[77] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

[78] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.

[79] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger and W. F. van Gunsteren, *J. Phys. Chem. A*, 1999, **103**, 3596–3607.

[80] L. D. Schuler, X. Daura and W. F. van Gunsteren, *J. Comput. Chem.*, 2001, **22**, 1205–1218.

[81] C. Oostenbrink, A. Villa, A. E. Mark and W. F. van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656–1676.

[82] W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657–1666.

[83] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

[84] L. Yang, C.-H. Tan, M.-J. Hsieh, J. Wang, Y. Duan, P. Cieplak, J. Caldwell, P. A. Kollman and R. Luo, *J. Phys. Chem. B*, 2006, **110**, 13166–13176.

[85] G. C. Q. da Silva, G. M. Silva, F. W. Tavares, F. P. Fleming and B. A. C. Horta, *J. Mol. Model.*, 2020, **26**, 296.

[86] Z. Jing, C. Liu, S. Y. Cheng, R. Qi, B. D. Walker, J.-P. Piquemal and P. Ren, *Annu. Rev. Biophys.*, 2019, **48**, 371–394.

[87] J. A. Lemkul, J. Huang, B. Roux and A. D. MacKerell, *Chem. Rev.*, 2016, **116**, 4983–5013.

[88] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046–4063.

[89] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.

[90] W. K. Hastings, *Biometrika*, 1970, **57**, 97–109.

[91] W. G. Noid, *J. Chem. Phys.*, 2013, **139**, 090901.

[92] D. Reith, M. Pütz and F. Müller-Plathe, *J. Comput. Chem.*, 2003, **24**, 1624–1636.

[93] A. Lyubartsev and A. Laaksonen, *Phys. Rev. E*, 1995, **52**, 3730–3737.

[94] M. S. Shell, *J. Chem. Phys.*, 2008, **129**, 144108.

[95] A. Chaimovich and M. S. Shell, *J. Chem. Phys.*, 2011, **134**, 094112.

[96] F. Ercolessi and J. B. Adams, *Europhys. Lett.*, 1994, **26**, 583–588.

[97] S. Izvekov and G. A. Voth, *J. Phys. Chem. B*, 2005, **109**, 2469–2473.

[98] S. Izvekov and G. A. Voth, *J. Chem. Phys.*, 2005, **123**, 134105.

[99] B. Hess, C. Holm and N. F. A. van der Vegt, *J. Chem. Phys.*, 2006, **124**, 164509.

[100] Y. Wang, W. G. Noid, P. Liu and G. A. Voth, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2002–2015.

[101] E. Brini, V. Marcona and N. F. A. van der Vegt, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10468–10474.

[102] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Chem. Phys. B*, 2007, **111**, 7812–7824.

[103] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S. J. Marrink, *J. Chem. Theory Comput.*, 2008, **4**, 819–834.

[104] C. A. López, A. J. Rzepiela, A. H. de Vries, L. Dijkhuizen, P. H. Hünenberger and S. J. Marrink, *J. Chem. Theory Comput.*, 2009, **5**, 3195–3210.

[105] D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman and S. J. Marrink, *J. Chem. Theory Comupt.*, 2013, **9**, 687–697.

[106] P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grunewald, I. Patmanidis, H. Abdizadeh, B. M. Bruininks, T. A. Wassenaar, P. C. Kroon, J. Melcr, V. Nieto, V. Corradi, H. M. Khan, J. Domański, M. Javanainen, H. Martinez-Seara, N. Reuter, R. B. Best, I. Vattulainen, L. Monticelli, X. P. D. P. Tieleman, A. H. de Vries and S. J. Marrink, *Nature Methods*, 2021, **18**, 382–388.

[107] T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller and G. Jackson, *J. Chem. Phys.*, 2013, **139**, 154504.

[108] V. Papaioannou, T. Lafitte, C. Avendaño, C. S. Adjiman, G. Jackson, E. A. Müller and A. Galindo, *J. Chem. Phys.*, 2014, **140**, 054107.

[109] G. Jiménez-Serratos, C. Herdes, A. J. Haslam, G. Jackson and E. A. Müller, *Macromol.*, 2017, **50**, 4840–4853.

[110] A. Glielmo, B. E. Husic, A. Rodriguez, C. Clementi, F. Noé and A. Laio, *Chem. Rev.*, 2021, **121**, 9722–9758.

[111] S. Y. Joshi and S. A. Deshmukh, *Mol. Simul.*, 2021, **47**, 786–803.

[112] W. Li, C. Burkhart, P. Polińska, V. Harmandaris and M. Doxastakis, *J. Chem. Phys.*, 2020, **153**, 041101.

[113] E. Boattini, M. Ram, F. Smallenburg and L. Filion, *Mol. Phys.*, 2018, **116**, 3066–3075.

[114] E. Boattini, M. Dijkstra and L. Filion, *J. Chem. Phys.*, 2019, **151**, 154901.

[115] E. Boattini, S. Marín-Aguilar, S. Mitra, G. Foffi, F. Smallenburg and L. Filion, *Nat. Commun.*, 2020, **11**, 5479.

[116] T. Lemke and C. Peter, *J. Chem. Theory Comput.*, 2017, **13**, 6213–6221.

[117] J. Wang, S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. de Fabritiis, F. Noé and C. Clementi, *ACS Centi. Sci.*, 2019, **5**, 755–767.

[118] B. E. Husic, N. E. Charron, D. Lemm, J. Wang, A. Pérez, M. Majewski, A. Krämer, Y. Chen, S. Olsson, G. de Fabritiis, F. Noé and C. Clementi, *J. Chem. Phys.*, 2020, **153**, 194101.

[119] H. Chan, M. J. Cherukara, B. Narayanan, T. D. Loeffler, C. Benmore, S. K. Gray and S. K. R. S. Sankaranarayanan, *Nat. Commun.*, 2019, **10**, 379.

[120] J. W. Wagner, J. F. Dama, A. E. P. Durumeric and G. A. Voth, *J. Chem. Phys.*, 2016, **145**, 044108.

[121] E. C. Allen and G. C. Rutledge, *J. Chem. Phys.*, 2009, **130**, 204903.

[122] M. E. Johnson, T. Head-Gordon and A. A. Louis, *J. Chem. Phys.*, 2007, **126**, 144509.

[123] P. Ganguly, D. Mukherji, C. Junghans and N. F. A. van der Vegt, *J. Chem. Theory Comput.*, 2012, **8**, 1802–1807.

[124] M. P. Allen, *Liq. Cryst.*, 2019, **117**, 2391–2417.

[125] M. R. Wilson, G. Yu, T. D. Potter, M. Walker, S. J. Gray, J. Li and N. J. Boyd, *Crystals*, 2022, **12**, 685.

[126] O. M. M. Rivas and A. D. Rey, *Carbon*, 2016, **110**, 189–199.

[127] O. M. M. Rivas and A. D. Rey, *Mol. Syst. Des. Eng.*, 2017, **2**, 223–234.

[128] O. M. M. Rivas and A. D. Rey, *J. Phys. Chem. B*, 2019, **123**, 1718–1732.

[129] O. M. M. Rivas and A. D. Rey, *J. Phys. Chem. B*, 2019, **123**, 8995–9010.

[130] Z. Chen, B. Fimmela and F. Würthner, *Org. Biomol. Chem.*, 2012, **10**, 5845–5855.

[131] J. Baz and N. Hansen, *J. Phys. Chem. C*, 2019, **123**, 8027–8036.

[132] K. Bag, R. Halder, B. Jana and S. Malik, *J. Phys. Chem. C*, 2019, **123**, 6241–6249.

[133] F. Haverkort, A. Stradomska, A. H. de Vries and J. Knoester, *J. Phys. Chem. B*, 2013, **117**, 5857–5867.

[134] F. Haverkort, A. Stradomska and J. Knoester, *J. Phys. Chem. B*, 2014, **118**, 8877–8890.

[135] C. Friedl, T. Renger, H. von Berlepsch, K. Ludwig, M. S. am Busch and J. Megow, *J. Phys. Chem. C*, 2016, **120**, 19416–19433.

[136] M. Walker, A. J. Masters and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2014, **16**, 23074–23081.

[137] M. Walker and M. R. Wilson, *Soft Matter*, 2016, **12**, 8588–8594.

[138] T. D. Potter, J. Tasche, E. L. Barrett, M. Walker and M. R. Wilson, *Liq. Cryst.*, 2017, **44**, 1979–1989.

[139] T. D. Potter, M. Walker and M. R. Wilson, *Soft Matter*, 2020, **16**, 9488–9498.

[140] C. E. Shannon, *Proceedings of the IRE*, 1949, **37**, 10–21.

[141] P. M. Morse, *Phys. Rev.*, 1929, **34**, 57–64.

[142] J. P. Ryckaert and A. Bellemans, *Chem. Phys. Lett*, 1975, **30**, 123–125.

[143] J. P. Ryckaert and A. Bellemans, *Faraday Discuss. Chem. Soc.*, 1978, **66**, 95–106.

[144] B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

[145] J. A. Barker and R. O. Watts, *Mol. Phys.*, 1973, **26**, 789–792.

[146] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

[147] R. W. Hockney, S. P. Goel and J. W. Eastwood, *J. Comput. Phys.*, 1974, **14**, 148–158.

[148] W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *J. Chem. Phys.*, 1982, **76**, 637–649.

[149] L. R. Pratt and S. W. Haan, *J. Chem. Phys.*, 1981, **74**, 1864–1872.

[150] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

[151] G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

[152] S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511–519.

[153] W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695–1697.

[154] M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

[155] A. Villa, C. Peter and N. F. A. van der Vegt, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2077–2086.

[156] A. Villa, N. F. A. van der Vegt and C. Peter, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2068–2076.

[157] C. Li, J. Shen, C. Peter and N. F. A. van der Vegt, *Macromolecules*, 2012, **45**, 2551–2561.

[158] C. H. Bennett, *J. Comput. Phys.*, 1976, **22**, 245–268.

[159] T. Bereau and K. Kremer, *J. Chem. Theory Comput*, 2015, **11**, 2783–2791.

[160] T. D. Potter, E. L. Barrett and M. A. Miller, *J. Chem. Theory Comput*, 2021, **17**, 5777–5791.

[161] R. Henderson, *Phys. Lett. A*, 1974, **49**, 197–198.

[162] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer and D. Andrienko, *J. Chem. Theory Comput.*, 2009, **5**, 3211–3223.

[163] C. Peter, L. D. Site and K. Kremer, *Soft Matter*, 2008, **4**, 859–869.

[164] B. Mukherjee, L. D. Site, K. Kremer and C. Peter, *J. Phys. Chem. B*, 2012, **116**, 8474–8484.

[165] R. Faller and D. Reith, *Macromol.*, 2003, **36**, 5406–5414.

[166] G. Milano and F. Müller-Plathe, *J. Phys. Chem. B*, 2005, **109**, 18609–18619.

[167] Q. Sun and R. Faller, *J. Chem. Theory Comput.*, 2006, **2**, 607–615.

[168] P. Carbone, H. A. K. Varzaneh, X. Chen and F. Müller-Plathe, *J. Chem. Phys.*, 2008, **128**, 064904.

[169] T. C. Moore, C. R. Iacovella and C. McCabe, *J. Chem. Phys.*, 2014, **140**, 224104.

[170] T. D. Potter, J. Tasche and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2019, **21**, 1912–1927.

[171] P. Ganguly and N. F. A. van der Vegt, *J. Chem. Theory Comput.*, 2013, **9**, 5247–5256.

[172] W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das and H. C. Andersen, *J. Chem. Phys.*, 2008, **128**, 244114.

[173] W. G. Noid, P. Liu, Y. Wang, J. W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen and G. A. Voth, *J. Chem. Phys.*, 2008, **128**, 244115.

[174] Y. Wang, S. Izvekov, T. Yan and G. A. Voth, *J. Phys. Chem. B*, 2006, **110**, 3564–3575.

[175] V. Rühle and C. Junghans, *Macromol. Theory Simul.*, 2011, **20**, 472–477.

[176] P. Liu, S. Izvekov and G. A. Voth, *J. Phys. Chem. B*, 2007, **111**, 11566–11575.

[177] J. Zhou, I. F. Thorpe, S. Izvekov and G. A. Voth, *Biophys. J.*, 2007, **92**, 4289–4303.

[178] I. F. Thorpe, J. Zhou and G. A. Voth, *J. Phys. Chem. B*, 2008, **112**, 13079–13090.

[179] S. Izvekov and G. A. Voth, *J. Chem. Theory Comput.*, 2006, **2**, 637–648.

[180] L. Lu and G. A. Voth, *J. Chem. Phys.*, 2011, **134**, 224107.

[181] Z. Cao and G. A. Voth, *J. Chem. Phys.*, 2015, **143**, 243116.

[182] A. Das and H. C. Andersen, *J. Chem. Phys.*, 2009, **131**, 034102.

[183] A. Das and H. C. Andersen, *J. Chem. Phys.*, 2012, **136**, 194113.

[184] A. Das and H. C. Andersen, *J. Chem. Phys.*, 2010, **132**, 164106.

[185] N. J. H. Dunn and W. G. Noid, *J. Chem. Phys.*, 2015, **143**, 243148.

[186] N. J. H. Dunn and W. G. Noid, *J. Chem. Phys.*, 2016, **144**, 204124.

[187] A. Das and H. C. Andersen, *J. Chem. Phys.*, 2010, **133**, 064109.

[188] L. Larini, L. Lu and G. A. Voth, *J. Chem. Phys.*, 2010, **132**, 164107.

[189] L. Larini, L. Lu and G. A. Voth, *J. Chem. Phys.*, 2012, **136**, 194114.

[190] M. R. DeLyser and W. G. Noid, *J. Chem. Phys.*, 2017, **147**, 134111.

[191] M. R. DeLyser and W. G. Noid, *J. Chem. Phys.*, 2019, **151**, 224106.

[192] H. Lee, A. H. de Vries, S. J. Marrink and R. W. Pastor, *J. Phys. Chem. B*, 2009, **113**, 13186–13194.

[193] G. Rossi, L. Monticelli, S. R. Puisto, I. Vattulainen and T. Ala-Nissila, *Soft Matter*, 2011, **7**, 698–708.

[194] A. V. Sangwai and R. Sureshkumar, *Langmuir*, 2011, **27**, 6628–6638.

[195] M. Velinova, D. Sengupta, A. V. Tadjer and S. J. Marrink, *Langmuir*, 2011, **27**, 14071–14077.

[196] G. Pérez-Sánchez, J. R. B. Gomes and M. Jorge, *Langmuir*, 2013, **29**, 2387–2396.

[197] L. Monticelli, *J. Chem. Theory Comput.*, 2012, **8**, 1370–1378.

[198] E. L. da Rocha, G. F. Caramori and C. R. Rambo, *Phys. Chem. Chem. Phys.*, 2013, **15**, 2282–2290.

[199] S. Baoukina, L. Monticelli, S. J. Marrink and D. P. Tieleman, *Langmuir*, 2007, **23**, 12617–12623.

[200] S. Baoukina, S. J. Marrink and D. P. Tieleman, *Faraday Discuss.*, 2010, **144**, 393–409.

[201] J. W. Klingelhoefer, T. Carpenter and M. S. P. Sansom, *Biophys. J.*, 2009, **96**, 3519–3528.

[202] X. Periole, T. Huber, S. J. Marrink and T. P. Sakmar, *J. Am. Chem. Soc.*, 2007, **129**, 10126–10132.

[203] A. Catte, M. R. Wilson, M. Walker and V. S. Oganesyan, *Soft Matter*, 2018, **14**, 2796–2807.

[204] S. W. Chiu, H. L. Scott and E. Jakobsson, *J. Chem. Theory Comput.*, 2010, **6**, 851–863.

[205] K. Prasitnok and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2013, **15**, 17093–17104.

[206] S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta and S. J. Marrink, *PLoS Computational Biology*, 2010, **6**, e1000810.

[207] J. Michalowsky, L. V. Schäfer, C. Holm and J. Smiatek, *J. Chem. Phys.*, 2017, **146**, 054501.

[208] J. Michalowsky, J. Zeman, C. Holm and J. Smiatek, *J. Chem. Phys.*, 2018, **149**, 163319.

[209] R. Alessandri, P. C. T. Souza, S. Thallmair, M. N. Melo, A. H. de Vries and S. J. Marrink, *J. Chem. Theory Comput.*, 2019, **15**, 5448–5460.

[210] H. Wang, C. Junghans and K. Kremer, *Eur. Phys. J. E*, 2009, **28**, 221–229.

[211] K. R. Hadley and C. McCabe, *Mol. Sim.*, 2012, **38**, 671–681.

[212] S. Y. Mashayak, M. N. Jochum, K. Koschke, N. R. Aluru, V. ühle and C. Junghans, *PLos One*, 2015, **10**, e0131754.

[213] C. Scherer and D. Andrienko, *Eur. Phys. J. Spec. Top.*, 2016, **225**, 1441–1461.

[214] J. Vieillard-Baron, *J. Chem. Phys.*, 1972, **56**, 4729–4744.

[215] J. Vieillard-Baron, *Mol. Phys.*, 1974, **28**, 809–818.

[216] B. Oyarzún, T. van Westen and T. J. H. Vlugt, *J. Chem. Phys.*, 2015, **142**, 064903.

[217] G. Megariotis, A. Vyrkou, A. Leygue and D. N. Theodorou, *Ind. Eng. Chem. Res.*, 2011, **2**, 546–556.

[218] J. Zhang, J. Su, Y. Ma and H. Guo, *J. Phys. Chem. B*, 2012, **116**, 2075–2089.

[219] J. Zhang and H. Guo, *J. Phys. Chem. B*, 2014, **118**, 4647–4660.

[220] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

[221] J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247–260.

[222] A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.

[223] A. W. S. da Silva and W. F. Vranken, *BMC Res. Notes*, 2012, **5**, 367.

[224] N. J. H. Dunn, K. M. Lebold, M. R. DeLyser, J. F. Rudzinski and W. G. Noid, *J. Phys. Chem. B*, 2018, **122**, 3363–3377.

[225] Martini 3 open beta, {`http://cgmartini.nl/index.php/martini3beta`}.

[226] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoell, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

[227] D. H. de Jong, S. Baoukina, H. I. Ingólfsson and S. J. Marrink, *Comput. Phys. Commun.*, 2016, **199**, 1–7.

[228] A. Ferrarini, G. J. Moro and P. L. Nordio, *Liq. Cryst.*, 1995, **19**, 397–399.

[229] A. Ferrarini, G. J. Moro and P. L. Nordio, *Mol. Phys.*, 1996, **87**, 485–499.

[230] A. Ferrarini, G. Gottarelli, P. L. Nordio and G. P. Spada, *J. Chem. Soc., Perkin Trans. 2*, 1999, **3**, 411–418.

[231] A. di Matteo, S. M. Todd, G. Gottarelli, G. Solladie, V. E. Williams, R. P. Lemieux, A. Ferrarini and G. P. Spada, *J. Am. Chem. Soc.*, 2001, **123**, 7842–7851.

[232] Y. Vorobjev and J. Hermans, *Biophys. J.*, 1997, **73**, 722–732.

[233] D. J. Earl and M. R. Wilson, *J. Chem. Phys.*, 2003, **119**, 10280–10288.

[234] K. Shi, C. Lian, Z. Bai, S. Zhao and H. Liu, *Chem. Eng. Sci.*, 2015, **122**, 185–196.

[235] T. Sanyal and M. S. Shell, *J. Phys. Chem. B*, 2018, **122**, 5678–5693.

[236] F. W. Evans and H. A. Skinner, *Trans. Faraday Soc.*, 1959, **55**, 255–259.

[237] H. C. V. Ness, C. A. Soczek, G. L. Peloquin and R. L. Machado, *J. Chem. Eng. Data*, 1967, **12**, 217–224.

[238] L.-P. Wang, T. J. Martinez and V. S. Pande, *J. Phys. Chem. Lett.*, 2014, **5**, 1885–1891.

[239] R. B. Meyer, in *Les Houches Summer School in Theoretical Physics*, ed. R. G. Balian and G. Weil, Gordon and Breach, New York, 1976, pp. 273–373.

[240] M. Sepelj, A. Lesac, U. Baumeister, S. Diele, H. L. Nguyen and D. W. Bruce, *J. Mater. Chem.*, 2007, **17**, 1154–1165.

[241] V. P. Panov, M. Nagaraj, J. K. Vij, Y. P. Panarin, A. Kohlmeier, M. G. Tamba, R. A. Lewis and G. H. Mehl, *Phys. Rev. Lett.*, 2010, **105**, 167801.

[242] P. A. Henderson and C. T. Imrie, *Liq. Cryst.*, 2011, **38**, 1407–1414.

[243] K. Adlem, M. Čopič, G. R. Luckhurst, A. Mertelj, O. Parri, R. M. Richardson, B. D. Snow, B. A. Timimi, R. P. Tuffin and D. Wilkes, *Phys. Rev. E*, 2013, **88**, 022503.

[244] R. J. Mandle, E. J. Davis, S. A. Lobato, C.-C. A. Vol, S. J. Cowling and J. W. Goodby, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6907–6915.

[245] R. J. Mandle, E. J. Davis, C. T. Archbold, S. J. Cowling and J. W. Goodby, *J. Mater. Chem. C*, 2014, **2**, 556–566.

[246] R. J. Mandle, E. J. Davis, C. T. Archbold, C.-C. A. Voll, J. L. Andrews, S. J. Cowling and J. W. Goodby, *Chem. – Eur. J.*, 2015, **21**, 8158–8167.

[247] E. Gorecka, N. Vaupotic, A. Zep, D. Pociecha, J. Yoshioka, J. Yamamoto and H. Takezoe, *Angew. Chem., Int. Ed.*, 2015, **54**, 10155–10159.

[248] D. A. Paterson, M. Gao, Y.-K. Kim, A. Jamali, K. L. Finley, B. Robles-Hernández, S. Diez-Berart, J. Salud, M. R. de la Fuente, B. A. Timimi, H. Zimmermann, C. Greco, A. Ferrarini, J. M. D. Storey, D. O. López, O. D. Lavrentovich, G. R. Luckhurst and C. T. Imrie, *Soft Matter*, 2016, **12**, 6827–6840.

[249] D. A. Paterson, J. P. Abberley, W. T. A. Harrison, J. M. D. Storey and C. T. Imrie, *Liq. Cryst.*, 2017, **44**, 127–146.

[250] D. Chen, M. Nakata, R. Shao, M. R. Tuchband, M. Shuai, U. Baumeister, W. Weissflog, D. M. Walba, M. A. Glaser, J. E. Maclennan and N. A. Clark, *Phys. Rev. E*, 2014, **89**, 022506.

[251] S. P. Sreenilayam, V. P. Panov, J. K. Vij and G. Shanker, *Liq. Cryst.*, 2017, **44**, 244–253.

[252] S. M. Jansze, A. Martínez-Felipe, J. M. D. Storey, A. T. M. Marcelis and C. T. Imrie, *Angew. Chem. Int. Ed.*, 2015, **54**, 643–646.

[253] R. Walker, D. Pociecha, J. P. Abberley, A. Martinez-Felipe, D. A. Paterson, E. Forsyth, G. B. Lawrence, P. A. Henderson, J. M. D. Storey, E. Gorecka and C. T. Imrie, *Chem. Commun.*, 2018, **54**, 3383–3386.

[254] R. Walker, D. Pociecha, A. Martinez-Felipe, J. M. D. Storey, E. Gorecka and C. T. Imrie, *Crystals*, 2020, **10**, 175.

[255] R. Walker, D. Pociecha, M. Salamończyk, J. M. D. Storey, E. Gorecka and C. T. Imrie, *Mater. Adv.*, 2020, **1**, 1622–1630.

[256] R. J. Mandle, *Soft Matter*, 2016, **12**, 7883–7901.

[257] R. Walker, *Liq. Cryst. Today*, 2020, **29**, 2–14.

[258] J. Xiang, Y. Li, Q. Li, D. A. Paterson, J. M. D. Storey, C. T. Imrie and O. D. Lavrentovich, *Adv. Mater.*, 2015, **27**, 3014–3018.

[259] J. Xiang, A. Varanytsia, F. Minkowski, D. A. Paterson, J. M. D. Storey, C. T. Imrie, O. D. Lavrentovich and P. Palffy-Muhoray, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 12925–12928.

[260] M. Mrukiewicz, O. S. Iadlovska, G. Babakhanova, S. Siemianowski, S. V. Shiyanovskii and O. D. Lavrentovich, *Liq. Cryst.*, 2019, **46**, 1544–1550.

[261] S. K. Prasad, P. L. Madhuri, P. Satapathy and C. V. Yelamaggad, *Appl. Phys. Lett.*, 2018, **112**, 253701.

[262] V. Sridurai, M. B. Kanakala, C. V. Yelamaggad and G. G. Nair, *Soft Matter*, 2019, **15**, 9982–9990.

[263] S. Aya, P. Salamon, D. A. Paterson, J. M. D. Storey, C. T. Imrie, F. Araoka, A. Jákli and . Buka, *Adv. Mater. Interfaces*, 2019, **6**, 1802032.

[264] C. Feng, J. Feng, R. Saha, Y. Arakawa, J. Gleeson, S. Sprunt, C. Zhu and A. Jákli, *Phys. Rev. Res.*, 2020, **2**, 032004.

[265] M. J. Cook and M. R. Wilson, *J. Chem. Phys.*, 2000, **112**, 1560–1564.

[266] M. R. Wilson and D. J. Earl, *J. Mater. Chem.*, 2001, **11**, 2672–2677.

[267] M. P. Neal, M. Solymosi, M. R. Wilson and D. J. Earl, *J. Chem. Phys.*, 2003, **119**, 3567–3573.

[268] D. J. Earl and M. R. Wilson, *J. Chem. Phys.*, 2003, **119**, 10280–10288.

[269] D. J. Earl and M. R. Wilson, *J. Chem. Phys.*, 2004, **120**, 9679–9683.

[270] D. J. Earl, M. A. Osipov, H. Takezoe, Y. Takanishi and M. R. Wilson, *Phys. Rev. E*, 2005, **71**, 021706.

[271] N. J. Boyd and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1485–1496.

[272] M. R. Wilson, *J. Chem. Phys.*, 1998, **107**, 8654–8663.

[273] A. G. Vanakaras and D. J. Photinos, *Liq. Cryst.*, 2018, **45**, 2184–2196.

[274] C. Greco and A. Ferrarini, *Phys. Rev. Lett.*, 2015, **115**, 147801.

[275] M. Chiappini, T. Drwenski, R. van Roij and M. Dijkstra, *Phys. Rev. Lett.*, 2019, **123**, 068001.

[276] M. Chiappini, A. Patti and M. Dijkstra, *Phys. Rev. E*, 2020, **102**, 040601.

[277] M. Chiappini and M. Dijkstra, *Nat. Commun.*, 2021, **12**, 2157.

[278] G. Yu and M. R. Wilson, *Soft Matter*, 2022, **18**, 3087–3096.

[279] R. Walker, D. Pociecha, J. M. D. Storey, E. Gorecka and C. T. Imrie, *Chem. – Eur. J.*, 2019, **25**, 13329–13335.

[280] A. J. Stone, *Mol. Phys.*, 1978, **36**, 241–256.

[281] C. Meyer, G. R. Luckhurst and I. Dozov, *J. Mater. Chem. C*, 2015, **3**, 318–328.

[282] G. Singh, J. Fu, D. M. Agra-Kooijman, J.-K. Song, M. R. Vengatesan, M. Srinivasarao, M. R. Fisch and S. Kumar, *Phys. Rev. E*, 2016, **94**, 060701.

[283] R. J. Mandle and J. W. Goodby, *Phys. Chem. Chem. Phys.*, 2019, **21**, 6839–6843.

[284] N. J. Boyd and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2015, **17**, 24851–24865.

[285] J. W. Emsley, M. Lelli, A. Lesage and G. R. Luckhurst, *J. Phys. Chem. B*, 2013, **117**, 6547–6557.

[286] A. Varanytsia and L.-C. Chien, *J. Appl. Phys.*, 2016, **119**, 014502.

[287] Y.-C. Lin, P.-C. Wu and W. Lee, *Photon. Res.*, 2019, **7**, 1258–1265.

[288] C. T. Imrie and P. A. Henderson, *Chem. Soc. Rev.*, 2007, **36**, 2096–2124.

[289] A. Ferrarini, G. R. Luckhurst, P. L. Nordio and S. J. Roskilly, *J. Chem. Phys.*, 1994, **100**, 1460–1469.

[290] G. R. Luckhurst and S. Romano, *J. Chem. Phys.*, 1997, **107**, 2557–2572.

[291] C. Greco, G. R. Luckhurst and A. Ferrarini, *Soft Matter*, 2014, **10**, 9318–9323.

[292] E. Cruickshank, M. Salamończyk, D. Pociecha, G. J. Strachan, J. M. D. Storey, C. Wang, J. Feng, C. Zhu, E. Gorecka and C. T. Imrie, *Liq. Cryst.*, 2019, **46**, 1595–1609.

[293] E. E. Pocock, R. J. Mandle and J. W. Goodby, *Soft Matter*, 2018, **14**, 2508–2514.

[294] M. R. Wilson and M. P. Allen, *Liq. Cryst.*, 1992, **12**, 157–176.

[295] M. R. Wilson, *Liq. Cryst.*, 1996, **21**, 437–447.

[296] C. Greco, G. R. Luckhurst and A. Ferrarini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 14961–14965.

[297] Y. Arakawa, K. Komatsu and H. Tsuji, *New J. Chem.*, 2019, **43**, 6786–6793.

[298] Y. Arakawa, K. Komatsu, S. Inui and H. Tsuji, *J. Mol. Struct.*, 2020, **1199**, 126913.

[299] Y. Arakawa, K. Komatsu, T. Shiba and H. Tsuji, *Mater. Adv.*, 2021, **2**, 1760–1773.

[300] D. A. Paterson, J. P. Abberley, W. T. A. Harrison, J. M. D. Storey and C. T. Imrie, *Liq. Cryst.*, 2017, **44**, 127–146.

[301] Y. Arakawa, K. Komatsu, T. Shiba and H. Tsuji, *J. Mol. Liq.*, 2021, **326**, 115319.

[302] J. Peláez and M. R. Wilson, *Phys. Rev. Lett.*, 2006, **97**, 267801.

[303] R. A. Reddy, C. Zhu, R. Shao, E. Korblova, T. Gong, Y. Shen, E. Garcia, M. A. Glaser, J. E. Maclennan, D. M. Walba and N. A. Clark, *Science*, 2011, **332**, 72–77.

[304] M. A. Osipov and G. Pajak, *Eur. Phys. J. E*, 2014, **37**, 79.

[305] O. D. Lavrentovich, *Proc. Natl. Acad. Sci. U.S.A.*, 2020, **117**, 14629–14631.

[306] R. J. Mandle, N. Sebastián, J. Martinez-Perdiguero and A. Mertelj, *Nat. Commun.*, 2021, **12**, 4962.

[307] J. P. Abberley, R. Killah, R. Walker, J. M. D. Storey, C. T. Imrie, M. Salamończyk, C. Zhu, E. Gorecka and D. Pociecha, *Nat. Commun.*, 2018, **9**, 228.

[308] R. Walker, D. Pociecha, G. J. Strachan, J. M. D. Storey, E. Gorecka and C. T. Imrie, *Soft Matter*, 2019, **15**, 3188–3197.

[309] M. Salamończyk, N. Vaupotič, D. Pociecha, R. Walker, J. M. D. Storey, C. T. Imrie, C. Wang, C. Zhu and E. Gorecka, *Nat. Commun.*, 2019, **10**, 1922.

[310] D. A. Dunmur and K. Toriyama, *Liq. Cryst.*, 1986, **1**, 169–180.

[311] K. Toriyama and D. A. Dunmur, *Mol. Cryst. Liq. Cryst.*, 1986, **139**, 123–142.

[312] M. J. Cook and M. R. Wilson, *Liq. Cryst.*, 2000, **27**, 1573–1583.

[313] J. Peláez and M. R. Wilson, *Phys. Chem. Chem. Phys.*, 2007, **9**, 2968–2975.

[314] F. Würthner, T. E. Kaiser and C. R. Saha-Möller, *Angew. Chem. Int. Ed.*, 2011, **50**, 3376–3410.

[315] S. E. Sheppard, R. H. Lambert and R. D. Walker, *J. Chem. Phys.*, 1939, **7**, 265–273.

[316] M. S. A. Abdel-Mottaleb, M. M. S. Abdel-Mottaleb, H. S. Hafez and M. Saif, *Int. J. of Photoenergy*, 2014, **2014**, 579476.

[317] S. E. Sheppard and H. Crouch, *J. Phys. Chem.*, 1928, **32**, 751–762.

[318] I. S. Lim, F. Goroleski, D. Mott, N. Kariuki, W. Ip, J. Luo and C. Zhong, *J. Phys. Chem. B*, 2006, **110**, 6673–6682.

[319] A. S. Elsherbiny, M. A. Salem and A. A. Ismail, *Chem. Eng. J.*, 2012, **200**, 283–290.

[320] C. Didraga, A. Pugžlys, P. R. Hania, H. von Berlepsch, K. Duppen and J. Knoester, *J. Phys. Chem. B*, 2004, **108**, 14976–14985.

[321] Z. Jin, Z. Li, K. Kasatani and H. Okamoto, *J. Lumin.*, 2007, **122**, 427–429.

[322] S. B. Shiring, R. L. Gieseking, C. Risko and J. Brédas, *J. Phys. Chem. C*, 2017, **121**, 14166–14175.

[323] F. Sivandzade, A. Bhalerao and L. Cucullo, *Bio-protocol*, 2019, **9**, e3128.

[324] J. O. Escobedo, O. Rusin, S. Lim and R. M. Strongin, *Curr. Opin. Chem. Biol.*, 2010, **14**, 64–70.

[325] H. von Berlepsch, C. Böttcher and L. Dähne, *J. Phys. Chem. B*, 2000, **104**, 8792–8799.

[326] W. P. Bricker, J. L. Banal, M. B. Stone and M. Bathe, *J. Chem. Phys.*, 2018, **149**, 024905.

[327] S. Kirstein and S. Daehne, *Int. J. of Photoenergy*, 2006, **5**, 1–21.

[328] H. von Berlepsch, K. Ludwig and C. Böttcher, *Phys. Chem. Chem. Phys.*, 2014, **16**, 10659–10668.

[329] A. V. Sorokin, I. Y. Ropakova, R. S. Grynyov, M. M. Vilkisky, V. M. Liakh, I. A. Borovoy, S. L. Yefimova and Y. V. Malyukin, *Dyes and Pigments*, 2018, **152**, 49–53.

[330] S. J. Khouri and V. Buss, *J. Solution Chem.*, 2010, **39**, 121–130.

[331] H. Min, J. Park, J. Yu and D. Kim, *Bull. Korean Chem. Soc.*, 1998, **19**, 650–654.

[332] C. Rodríguez-Abreu, C. Aubrey-Torres and G. J. T. Tiddy, *Langmuir*, 2011, **27**, 3067–3073.

[333] E. K. Batchelor, S. Gadde and A. E. Kaifer, *Supramolecular Chemistry*, 2010, **22**, 40–45.

[334] K. Manna and A. K. Panda, *Spectrochimica Acta Part A*, 2009, **74**, 1268–1274.

[335] H. von Berlepsch and C. Böttcher, *Langmuir*, 2013, **29**, 4948–4958.

[336] B. Birkan, D. Gülen and S. Özçelik, *J. Phys. Chem. B*, 2006, **110**, 10805–10813.

[337] A. Pawlik, A. Ouart, S. Kirstein, H. W. Abraham and S. Daehne, *Eur. J. Org. Chem.*, 2003, **2003**, 3065–3080.

[338] S. Özçelik, I. Özçelik and D. L. Akins, *Appl. Phys. Lett.*, 1998, **73**, 1949–1951.

[339] I. Filimonova, S. Yefimova and A. V. Sorokin, *Functional Materials*, 2012, **19**, 348–353.

[340] M. S. Bradley, J. R. Tischler and V. Bulovic, *Adv. Mater.*, 2005, **17**, 1881–1886.

[341] I. Renge and U. P. Wild, *J. Phys. Chem. A*, 1997, **101**, 7977–7988.

[342] D. Gülen and S. Özçelik, *J. Lumin.*, 2008, **128**, 834–837.

[343] B. Kopainsky, J. K. Hallermeier and W. Kaiser, *Chem. Phys. Lett.*, 1981, **83**, 498–502.

[344] Y. Kitahama, T. Yago, A. Furube and R. Katoh, *Chem. Phys. Lett.*, 2008, **457**, 427–433.

[345] G. Scheibe and L. Kandler, *Naturwissenschaften*, 1938, **26**, 412–413.

[346] T. Förster, *Naturwissenschaften*, 1946, **33**, 166–175.

[347] H. Bücher and H. Kuhn, *Chem. Phys. Lett.*, 1970, **6**, 183–185.

[348] E. Daltrozzo, G. Scheibe, K. Geschwind and F. Haimerl, *Photogr. Sci. Eng.*, 1974, **18**, 441–450.

[349] A. S. Davydov, *Theory of Molecular Excitons*, Plenum Press, New York, 1971.

[350] D. Smith and H. Luss, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1972, **28**, 2793–2806.

[351] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

[352] B. Jójárt and T. A. Martinek, *J. Comput. Chem.*, 2007, **28**, 2051–2058.

[353] C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361–373.

[354] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewsk, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 09 (Revision A.02)*, Gaussian, Inc., Wallingford CT, 2016.

[355] W. L. Jorgensen and J. D. Madura, *Mol. Phys.*, 1985, **56**, 1381–1392.

[356] M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.*, 2000, **112**, 8910–8922.

[357] M. F. Harrach and B. Drossel, *J. Chem. Phys.*, 2014, **140**, 174501.

[358] C. Vega and J. L. F. Abascal, *Phys. Chem. Chem. Phys.*, 2011, **13**, 19663–19688.

[359] G. Scheibe, *Elektrochem*, 1948, **52**, 283.

[360] B. Neumann and P. Pollmann, *Ber. Bun. Phys. Chem.*, 1996, **100**, 15–19.

[361] B. Neumann and P. Pollmann, *Phys. Chem. Chem. Phys.*, 2000, **2**, 4784–4792.

[362] H. Naorem and S. D. Devi, *J. Mol. Liq.*, 2012, **173**, 119–123.

[363] P. J. Collings, J. N. Goldstein, E. J. Hamilton, B. R. Mercado, K. J. Nieser and M. H. Regan, *Liq. Cryst. Rev.*, 2015, **3**, 1–27.

[364] B. I. Shapiro, *Nanotechnol. Russia*, 2008, **3**, 139–150.

[365] R. C. Rizzo, T. Aynechi, D. A. Case and I. D. Kuntz, *J. Chem. Theory Comput.*, 2006, **2**, 128–139.

[366] S. Jain and F. S. Bates, *Science*, 2003, **300**, 460–464.

[367] N. Dan, K. Shimoni, V. Pata and D. Danino, *Langmuir*, 2006, **22**, 9860–9865.

[368] A. G. Zilman and S. A. Safran, *Phys. Rev. E*, 2002, **66**, 051107.

[369] W. M. Gelbart and A. Ben-Shaul, *J. Phys. Chem.*, 1996, **100**, 13169–13189.

[370] M. E. Cates and S. J. Candau, *J. Phys.: Condens. Matter*, 1990, **2**, 6869–6892.

[371] F. C. MacKintosh, S. A. Safran and P. A. Pincus, *Europhys. Lett.*, 1990, **12**, 697–702.

[372] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *J. Comput. Chem.*, 1995, **16**, 273–284.

[373] W. J. Harrison, D. L. Mateer and G. J. T. Tiddy, *J. Phys. Chem.*, 1996, **100**, 2310–2321.

[374] C. A. Dreiss, *Soft Matter*, 2007, **3**, 956–970.

[375] P. Ekwall, L. Mandell and K. Fontell, *J. Colloid Interface Sci.*, 1969, **31**, 508–529.

[376] R. M. Wood and M. P. McDonald, *J. Chem. Soc., Faraday Trans. 1*, 1985, **81**, 273–283.

[377] T. Heinemann, K. Palczynski, J. Dzubiella and S. H. L. Klapp, *J. Chem. Phys.*, 2014, **141**, 214110.

[378] T. Heinemann, K. Palczynski, J. Dzubiella and S. H. L. Klapp, *J. Chem. Phys.*, 2015, **143**, 174110.

[379] A. S. Bowen, N. E. Jackson, D. R. Reid and J. J. de Pablo, *J. Chem. Theory Comput.*, 2018, **14**, 6495–6504.

[380] M. Ricci, O. M. Roscioni, L. Querciagrossa and C. Zannoni, *Phys. Chem. Chem. Phys.*, 2019, **21**, 26195–26211.

[381] C. Clementi, *Curr. Opin. Struct. Biol.*, 2008, **18**, 10–15.

[382] A. E. Hafner, J. Krausser and A. Šarić, *Curr. Opin. Struct. Biol.*, 2019, **58**, 43–52.

[383] M. Walker and M. R. Wilson, *Soft Matter*, 2016, **12**, 8876–8883.

[384] A. Ghoufi, D. Morineau, R. Lefort and P. Malfreyt, *J. Chem. Theory Comput.*, 2010, **6**, 3212–3222.

[385] F. Stanzione and A. Jayaraman, *J. Phys. Chem. B*, 2016, **120**, 4160–4173.

[386] A. Renevey and S. Riniker, *J. Chem. Phys.*, 2017, **146**, 124131.

[387] P. Kar and M. Feig, *J. Chem. Theory Comput.*, 2017, **13**, 5753–5765.

[388] A. Renevey and S. Riniker, *J. Phys. Chem. B*, 2019, **123**, 3033–3042.

[389] J. Roel-Touris and A. M. J. J. Bonvin, *Comput. Struct. Biotechnol. J.*, 2020, **18**, 1182–1190.

[390] A. J. Rzepiela, M. Louhivuori, C. Peter and S. J. Marrink, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10437–10448.

[391] T. A. Wassenaar, H. I. Ingólfsson, M. Prieß, S. J. Marrink and L. V. Schäfer, *J. Phys. Chem. B*, 2013, **117**, 3516—3530.

[392] S. Genheden, *J. Comput. Aided Mol. Des.*, 2017, **31**, 867–876.

[393] M. Pechlaner, C. Oostenbrink and W. F. van Gunsteren, *J. Comput. Chem.*, 2021, **42**, 1263–1282.

[394] H. Zollinger, *Color Chemistry: Synthesis, Properties and Applications of Organic Dyes and Pigments, 3*$^{\mathrm{rd}}$ *Edition*, VCH, Weinheim, 2003.

[395] F. Würthner, *Chem. Commun.*, 2004, **14**, 1564–1579.

[396] C. Huang, S. Barlow and S. R. Marder, *J. Org. Chem.*, 2011, **76**, 2386–2407.

[397] A. Nowak-Król and F. Würthner, *Org. Chem. Front*, 2019, **6**, 1272–1318.

[398] I. K. Iverson and S.-W. Tam-Chang, *J. Am. Chem. Soc.*, 1999, **121**, 5801–5802.

[399] S.-W. Tam-Chang, W. Seo, K. Rove and S. M. Casey, *Chem. Mater.*, 2004, **16**, 1832–1834.

[400] S.-W. Tam-Chang, W. Seo and I. K. Iverson, *J. Org. Chem.*, 2004, **69**, 2719–2726.

[401] S.-W. Tam-Chang, J. Helbley, T. D. Carson, W. Seo and I. K. Iverson, *Chem. Commun.*, 2006, **5**, 503–505.

[402] C. Rodríguez-Abreu, C. Aubery-Torres, C. Solans, A. Lopez-Quintela and G. J. T. Tiddy, *ACS Appl. Mater. Interfaces*, 2011, **3**, 4133–4141.

[403] I. K. Iverson, S. M. Casey, W. Seo and S.-W. Tam-Chang, *Langmuir*, 2002, **18**, 3510–3516.

[404] S.-W. Tam-Chang, I. K. Iverson and J. Helbley, *Langmuir*, 2004, **20**, 342–347.

[405] T. Ogolla and P. J. Collings, *Liq. Cryst.*, 2019, **46**, 1551–1557.

[406] M. Walker and M. R. Wilson, *Mol. Cryst. Liq. Cryst.*, 2015, **612**, 117–125.

[407] G. Yu and M. R. Wilson, *J. Mol. Liq.*, 2022, **345**, 118210.

[408] J. W. Mullinax and W. G. Noid, *J. Chem. Phys.*, 2009, **131**, 104110.

[409] T. Sanyal and M. S. Shell, *J. Chem. Phys.*, 2016, **145**, 034109.

[410] J. Jin and G. A. Voth, *J. Chem. Theory Comput.*, 2018, **14**, 2180–2197.

[411] A. Wicklei, A. Lang, M. Muth and M. Thelakkat, *J. Am. Chem. Soc.*, 2009, **131**, 14442–14453.

[412] F. Nolde, W. Pisula, S. Müller, C. Kohl and K. Müllen, *Chem. Mater.*, 2006, **18**, 3715–3725.

[413] M.-A. Muth, G. Gupta, A. Wicklein, M. Carrasco-Orozco, T. Thurn-Albrecht and M. Thelakkat, *J. Phys. Chem. C*, 2014, **118**, 92–102.

[414] V. Marcon, J. Kirkpatrick, W. Pisula and D. Andrienko, *Phys. Stat. Sol. (B)*, 2008, **245**, 820–824.

[415] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.*, 1996, **65**, 1604–1608.

[416] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, 1999, **314**, 141–151.

[417] M. Böckmann, C. Peter, L. D. Site, N. L. Doltsinis, K. Kremer and D. Marx, *J. Chem. Theory Comput.*, 2007, **3**, 1789–1802.

[418] K. Takemoto, Y. Ishii, H. Washizu, K. Kim and N. Matubayasi, *J. Chem. Phys.*, 2021, **156**, 014901.

[419] G. Bussi, *Mol. Phys.*, 2014, **112**, 379–384.

[420] R. Berardi, C. Zannoni, J. S. Lintuvuori and M. R. Wilson, *J. Chem. Phys.*, 2009, **131**, 174107.

[421] M. Chakraborty, J. Xu and A. D. White, *Phys. Chem. Chem. Phys.*, 2020, **22**, 14998–15005.