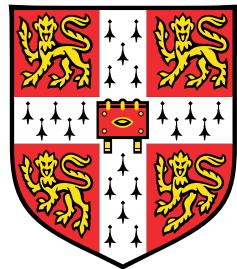


Multi-Level Coarse-Grained Modelling on DNA Functionalised Building Blocks

PhD Thesis



Jiaming Yu

Department of Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Churchill College

October 2022

I would like to dedicate this thesis to my loving parents.

“Hinc lucem et pocula sacra”

UNIVERSITY OF CAMBRIDGE

- From here [we derive] light and sacred draughts [literally cups]

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

This thesis partially covers and builds upon topics featuring in the following articles:

Chapter 3:

Iliya D. Stoev, Tianyang Cao, Alessio Caciagli, **Jiaming Yu**, Christopher Ness, Ren Liu, Rini Ghosh, Thomas O'Neill, Dongsheng Liu and Erika Eiser

On the role of flexibility in linker-mediated DNA hydrnogels
Soft Matter (2020)

Chapter 5:

Kit Gallagher, **Jiaming Yu**, Ren Liu, David A. King and Erika Eiser

Towards new liquid crystal phases of DNA mesogens
(In paraperation)

Chapter 6:

Jiaming Yu, Eivind Bering, Ren Liu and Erika Eiser

Numerical study of pluronic with DNA-functionalised triblock copolymer system
(In paraperation)

Ren Liu, **Jiaming Yu**, Xiaojing Tang, Rini Ghosh and Erika Eiser

Dynamic Light Scattering based microrheology of DNA-functionalised triblock copolymer solutions

(Submitted)

Jiaming Yu

October 2022

Acknowledgements

Over the course of my PhD, I have had the incredible privilege to interact, work together, talk about, and share the findings I describe in my dissertation with many people from various organisations and backgrounds. Without them, I can draw a strong inference with a high degree of probability that this thesis would not have been possible.

First and foremost, I would like to thank my supervisor, Prof. Erika Eiser, who has always given me encouragement, research direction, priceless suggestions, shrewd observations and sincere comments during my whole PhD study period. In addition to our professional relationship, she is an incredible cook and fantastic dinner host; I hope we can enjoy more wonderful dinners in the future. I am really appreciative of everything she has done for me, especially in introducing me to the next person I must mention.

In preparation for this PhD thesis, I had to take the help and guidance of the most respected person – Prof. Daan Frenkel, who deserve my deepest gratitude. I would like to show my appreciation to him for giving me the most proficient guidelines for assignments throughout numerous consultations. His book is my study Bible. He also keeps inspiring me with his tremendous knowledge and great kindness for my entire simulation work.

I would also like to express my gratitude to those who have directly guided and inspired me in the four years of study: Prof. Jonathan Doye, one of the developers of the oxDNA model from the University of Oxford, for valuable discussions during my preparation work for the DNA simulation. More useful discussions from Prof. Chiristos Likos (when I visited him in Vienna), Dr. Oliver Henrich (when we first met at the international oxDNA workshop), Prof. Sofia Kantorovich and Dr. Pedro Sanchez (when we conducted monthly remote discussions during the pandemic), Prof. Dick Bedeaux (for great talks for most of my projects when I visited the PoreLab at NTNU) and Dr. Alan Lee (for guiding me at the last stage of supervision).

My experience as a PhD student in Cambridge was amusing and stimulating, thanks to my colleagues back home in Erika's group. I am grateful to Kit for our excellent collaboration work, and I immensely enjoy the day-to-day supervision time we spent together. I would also specifically thank: David for outstandingly theoretical and mathematical support; Iliya, Ren,

and Rini for excellent experiment work; Chris, Zhongyang, Iria, and Eivind for the extra simulation support; Alessio, Tom, and Peicheng for the general guidance.

Completing this thesis gave me much pleasure. From the bottom of my heart, I want to thank my parents and all family members for their boundless love and support. Honestly, I would not be the person I am today without you. To my grandpa, who is at peace, but with me each day in spirit, this thesis is dedicated to you. Words cannot express how much I love you all. Special thanks to Qi, who always stayed on my side through the calm and through the storm: I hope this thesis serves as a final reward to fulfil our memory.

Last but not least, I want to acknowledge the Cambridge Trust and Chinese Scholarship Council for generously funding my whole PhD study, together with the Cambridge Philosophical Society and Churchill College for other subsidies.

Abstract

DNA nanotechnology has proliferated in recent years as investigators have increasingly harnessed the selectivity of DNA base pairing to form multitudes of diverse structures, with a wide range of applications. The stiffness of the double-strand DNA (dsDNA) with a persistence length of around 50nm ($\approx 150\text{bp}$) provides rigid structural properties to form the most stable duplexes. Single-strand DNA (ssDNA) with much less rigidity acts as hinges between duplex sections allowing extra flexibility to the structure. This specificity of DNA properties gives considerable functional control in designing anisotropic nanoparticles or nano-sized scaffolds as building blocks. These functionalised DNA building blocks enable the formation of varied DNA functionalised systems, such as DNA liquid crystals, DNA hydrogels or DNA functionalised micelles, amongst other DNA-based materials. The properties of these large-scale DNA-based systems are hard to represent by experiments owing to technical limits and cost restrictions. Computer modelling can offer significant insight into the DNA structures with thermodynamics and mechanical properties and has the potential to simulate a detailed model of DNA nanostructures. In this thesis, I applied a homemade multi-level strategy to demonstrate new horizons of coarse-grained simulation systems. From the detailed oxDNA molecular dynamics simulation of mimicking the most detailed behaviours of an individual DNA nanoparticle, to a second-level coarse-grained system with thousands of repeated units of accelerating large-scaled simulation. I outlined further investigation for the existence of exotic DNA liquid crystal phases and proved a self-loop behaviour on the role of flexibility in our linker-mediated DNA hydrogels. I also presented phase diagrams and simulated radial distribution functions by adding extra temperature-controlled mechanisms in our DNA functionalised pluronic systems. With the service of this multi-level coarse-grained strategy, I can devote a more realistic simulation system to illustrate more reliable physical properties.

Table of contents

List of figures	xv
List of tables	xxiii
1 Introduction	1
1.1 Preamble	1
1.2 Thesis outline	3
2 Background	5
2.1 DNA as a Building Block to Design Nano-Structures	5
2.1.1 Molecular structure of DNA	5
2.1.2 DNA thermodynamics	7
2.2 DNA Functional Materials	10
2.2.1 DNA hydrogel	10
2.2.2 DNA liquid crystal	12
2.2.3 DNA-functionalised triblock copolymers	14
2.3 Numerical Methods	18
2.3.1 Molecular Dynamics Simulations	19
2.3.2 Monte Carlo Simulations	21
2.3.3 Langevin Dynamics	23
2.3.4 OxDNA Model and LAMMPS	24
3 The Role of Flexible Joints in a Linker-Mediated DNA Hydrogel	29
3.1 Introduction	30
3.2 Experimental Findings and Hypothesis	31
3.3 Results and Discussion	35
3.3.1 Comparison between flexible and rigid linker hydrogels	35
3.3.2 Structural stability of the key-shaped DNA	37

3.3.3	Closed-to-open ratio of the Y+L6 system	39
3.3.4	Simulated melting temperature of key- and Y-shaped DNA	40
3.3.5	Estimate of the centre-core bending penalty	42
3.3.6	The second-level coarse-grained simulation	44
3.4	Conclusion	45
3.5	Acknowledgements	47
4	Quantitative Analysis of Flexibility in Nunchucks DNA	49
4.1	Introduction	49
4.2	Simulate Nunchuck DNA as Mesogen	52
4.3	Results and Discussion	55
4.3.1	Bending free energy calculation	55
4.3.2	Statistical distribution	55
4.3.3	Mathematical model	58
4.3.4	Determine the stiffness factor	59
4.3.5	Colour zone representation and flexible saturation	60
4.3.6	Forwarding a second-level coarse-grained simulation model	62
4.3.7	Translate bending free energy into a virtual bond potential	63
4.4	Conclusion	64
4.5	Acknowledgement	66
5	Towards New Liquid Crystal Phases of DNA Mesogens	67
5.1	Introduction	68
5.2	Simulation Model Description	69
5.2.1	Coarse-grained bead-spring model	69
5.2.2	Simulation details	72
5.3	Phase Characterisation	73
5.4	Results and Discussion	74
5.4.1	Rigid rod order parameters	74
5.4.2	Nunchuck correlation functions	77
5.4.3	Dynamic Behaviour	79
5.5	Conclusion	83
5.6	Acknowledgements	84
6	DNA-Functionalised Triblock Copolymers: F108 Pluronics + DNA	85
6.1	Introduction	86
6.2	Numerical Model	87

6.2.1	Model for the F108 triblock-copolymers with DNA patches	87
6.2.2	Simulation details	91
6.3	Results and Discussion	93
6.3.1	Determining the micellisation phase boundary	93
6.3.2	F108 phase diagram	96
6.3.3	Hybridization behaviour of our complementary DNA overhangs . .	97
6.3.4	Structural properties - determining $g(r)$	99
6.3.5	The F108-DNA phase diagram	102
6.3.6	Heterogeneity and homogeneity	104
6.4	Conclusion	111
6.5	Acknowledgements	112
7	Closing Remarks	113
References		117
Appendix A DNA Sequences		137
A.1	Y-shape+Linkers	137
A.2	Nunchucks DNA	138
Appendix B Algorithms		139
B.1	Visualisation of oxDNA in LAMMPS	139
B.2	Nematic Order Parameter	140
B.3	Position Dependant Order Parameter	141
B.4	Structure factor for hard spheres	142
Appendix C Codes and Scripts		145
C.1	oxDNA Scripts in LAMMPS	145
C.2	LAMMPS scripts of Y-shaped DNA with linkers	148
C.3	LAMMPS scripts of Nunchuck-DNA	150
C.4	LAMMPS scripts of DNA functionlised F108 Pluronics	152
Appendix D Extra Plots		155
D.1	Schematics in Different Stage of Melting Process	155
D.2	Open and Close Ratio of Y+L6	156
D.3	Example snapshots of F108-DNA system	158

Appendix E Side Projects	161
E.1 DNA nanotubes with programmable sub-pore	161
E.2 Polydisperse diameter distribution for micelle system	163
E.3 Pluronic F127 with Y-shape DNA as mediated connector	164

List of figures

2.1	Multi-forms of DNA (A-DNA and B-DNA are right-handed, and Z-DNA is left-handed. B-DNA is the most commonly found conformation in nature)	5
2.2	A ladder-like structure of DNA built by sugar-phosphate backbones via base pairs.	6
2.3	The melting temperature of a known DNA duplex is done by measuring the absorbance at 260 nm as a function of temperature for solutions containing equal amounts of complementary ssDNA.	9
2.4	Representative principle structures and dendritic DNA assembled DNA hydrogels.	11
2.5	Photograph (experiment) and schematics (simulation) of DNA hydrogels made of three-armed nanostars.	12
2.6	Representation of the various phases rod-like mesogens can be in. The weakest (orientational) order is assumed in the nematic phase, while the smectic phase shows higher order with positional orientation in the layers.	13
2.7	Rodlike dsDNA can produce isotropic/nematic/columnar phases.	14
2.8	(a) Molecular structure of a PEO-PPO-PEO triblock copolymer. (b) The Pluronic grid.	15
2.9	Experimental phase diagram of F68 and F108.	16
2.10	Schematic protocol of making DNA-functionalised triblock-copolymers.	17
2.11	Schematic of the DNA-functionalised triblock copolymer system (F108-DNA). The hydrophilic parts (PEO) are coloured in blue; hydrophobic parts (PPO) are coloured in red; the complementary ssDNA sticky ends are coloured in red and yellow. The transparent yellow zone represents the hydrophobic core of the micelles.	18
2.12	Molecular Dynamics Simulation Cycle.	20
2.13	Schematics of Common Type Statistical Ensembles in Thermodynamics.	21
2.14	Schematic representation of the oxDNA model	27

3.1	Schematic of Y-shaped DNA, flexible linker-DNA (L6) and key-shaped DNA made of Y-L6 pair.	32
3.2	Electric-field autocorrelation functions ($g^{(1)}$) obtained from DLS measurements and the corresponding storage and loss moduli for the 1.6 wt% samples made of the Y-shapes and linkers using (a) rigidly attached L0, and (b) L6 linkers with a flexible 6T-joint.	34
3.3	Hypothesised structure: percolation is terminated by the looping behaviour in the Y-L6 system.	35
3.4	The Distribution of the probability that the joint between a Y-shape and linker adopts a certain angle θ_{L0}/θ_{L6}	36
3.5	The radius of gyration (R_g) as a function of simulation steps for two different temperatures and at an added salt (NaCl) concentration of 200 mM. Two simulation sets (1) T=300K(black) (2) T=330K(red). 2k frames are used in each set.	37
3.6	Closed-to-open ratio of Y+L6 system.	39
3.7	The radius-of-gyration plots for different temperatures at 200 mM NaCl. The small snapshots (clear version with description can be found in Appendix D.1) embedded in the figure are example scenarios of each corresponding state in the measurements. The last measured frames were used to extract the final-state information recorded in Tab. 3.3, each set with 2k frames	41
3.8	Probability distributions of the three angles α_1 , α_2 and α_3 between the three arms of a free Y-shape obtained from oxDNA simulations.	43
3.9	Free energy as a function of the angle distribution between the free arms in a Y-shape and the free energy of a Y-shape connected via an L6 linker.	43
3.10	Distribution of cluster sizes for Y+L0 (blue bars) and Y+L6 (red bars) scenarios.	44
3.11	Second-level coarse-grained simulation model based on oxDNA model.	46
4.1	Schematic of example designed nunchucks DNA as mesogens.	51
4.2	The model of nunchucks DNA, with θ being the bending angle, $\Delta\theta = \pi - \theta$	52
4.3	Schematic of nunchucks DNA mesogens (from left to right) with 0-,2-,4-,6-,8-,10- nucleotide bases on ssDNA flexible joint, respectively.	53
4.4	Statistical Achievable angles (θ) and $\cos(\theta)$ of all achievable angles for designed DNA mesogens.	54
4.5	Statistical results of flexible angle distribution and the probability of $\cos \theta$ in a nunchucks DNA.	56
4.6	Statistical results of bending free energy $F(\cos \theta)$ calculated from $\cos \theta$ distributions.	57

4.7	Color Zone representation of bending free energy with all achievable ranges.	60
4.8	Box-and-whisker diagram of achievable angle region for different lengths of ssDNA joints (N_s). Four regions defined (1) Rigid region ($N_s < 1$). (2) Confined region ($1 \leq N_s < 6$). (3) Flexible region ($6 \leq N_s \leq 10$). (4) Flexible saturation ($N_s > 10$)	61
4.9	Feasible CG model design. CG model 1 fits nunchucks DNA with short chain length ($N_s < 6$); CG model 2 works for longer chain cases ($N_s \geq 6$)	63
4.10	Translate the bending free energy into a virtual bond potential in second CG model.	64
4.11	A simulation strategy from first-level (build in oxDNA) to second-level (build in LAMMPS) coarse-grained simulation of using nunchucks DNA with CG model.	65
5.1	Both schematic representations of our course-grained bead-spring model and the oxDNA calculation below the melting temperature of the DNA sequences used.	70
5.2	The evolution of the volume fraction and the nematic order parameter over the timescale of the simulation, for a system of 1000 rigid rods with aspect ratio 10. The discrete change in S_n (red line) occurs at about $\phi \sim 0.4$, denoting the transition to the nematic phase, with the system depicted explicitly on either side of this transition. Note that the contraction steps, where the volume fraction is increased, are not of equal duration, and so do not correspond to equal changes in the system volume; rather they are chosen to highlight the phase transition.	74
5.3	The nematic order parameter (S_n) over the volume fraction (ϕ), with system sizes: N=500, N=1000, N=1500 and N=2000. The transition regions of different system sizes consist of $\phi \sim 0.39 - 0.44$ (for L/D=10) from isotropic to nematic transition. For L/D = 16, the transition occurs earlier when $\phi \sim 0.23 - 0.26$	76
5.4	The kernel density estimate of the opening angle distribution as the volume fraction is reduced, for completely flexible nunchucks. Plotted for a system of 1×10^3 particles, with the distribution sampled every 1.5×10^6 time steps. Note the formation of a preferential angle at late times, corresponding to the formation of an ordered phase at high volume fractions, the final form of which is depicted in the inset.	77

5.5	Orientational correlation function over time (as the volume fraction is reduced), for a system of 1000 nunchuck molecules, with a fixed opening angle of 150 deg and sampled every 7×10^5 time steps. Note the formation of an ordered phase at high volume fractions (late times), with sustained long-range order (i.e. no decay in the orientational correlation function). The maximum value of particle separation is determined by the size of the simulation box, and so shrinks over time.	78
5.6	Root mean-squared displacements against time in a system of 1000 rigid rods with an aspect ratio 10, for a range of volume fractions ϕ . The leftmost pair of plots correspond to the smectic phase, with restricted motion between $x - z$ layers along the y -axis. The centre pair give the nematic phase, with preferential displacement along the y -axis. Finally, the rightmost pair give the isotropic phase, with isotropic diffusion and no preferential direction. Note that no significant differences appear between the x - and z -directions in any phase, as these directions are equivalent in the phase structure.	80
5.7	Coordinate diffusion coefficients for the evolution of the microcanonical ensemble at a range of volume fractions ϕ . The low volume fraction region on the left of the graph corresponds to the isotropic phase, with no variation between coordinate axes. In contrast, the smectic phase in the high volume fraction gives rise to anisotropy in the coordinate diffusion coefficients, with reduced diffusivity perpendicular to the smectic layers in the y axis direction.	81
5.8	Root mean-squared displacements against time in a system of 1000 bent-core mesogens with an opening angle 120°, over a range of volume fractions ϕ . As previously, restricted diffusion is observed along the director (previously the y -axis) between smectic layers. Anisotropy is now observed within these layer at the highest volume fraction $\phi = 0.76$, suggesting a biaxial smectic phase, with a common aligned bisector. Diffusion is reduced in the direction of the normal vector compared to the bisector vector for these mesogen.	82
6.1	Simulation model details: (a) Model representations ($D_{\pm}A_{15}B_6A_{15}D_{\pm}$) for F108 [$(PEO)_{147} - (PPO)_{56} - (PEO)_{147}$] with the ssDNA treated as single patches at the chain ends (D_{\pm}). The different blue and red blobs represent the PEO and PPO chains. The effective volumes of the PEO and PPO blobs are the same here. D_{\pm} represents the complementary ssDNA patches. (b) The real-size scheme of the blobs. Here the complementary ssDNA-ends are coloured in yellow and green, respectively. (c) Representative snapshots of the unimer (below CMT) and micellar fluid state (above CMT).	88

6.2 All non-binding potentials used in the F108-DNA system. (a) Example snapshot of the temperature sensitive potential between PPO parts ($U_{BB}(T)$) and the attractive potential between ssDNA ends (U_{DD}). (b) The temperature sensitive pairwise potentials used in the model.	89
6.3 Here the normalised probability densities $P(\text{micelle})$ and $P(\text{unimer})$ are plotted as function of normalised temperature for different Pluronic volume fractions ϕ , in order to determine the unimer to micellar liquid phase boundary. The ‘dark-cyan band’ represents the micellisation threshold region. All measurement points were taken from the simulations under equilibrium conditions with $N = 500$ chains in the system. The error bars were acquired from ten repeated runs for each temperature setup.	94
6.4 Numerical phase diagram of Pluronic F108 dispersions: Green dots represent the unimer liquid phase, blue the micellar liquid region, and red dots the micellar solid phase. The micellisation boundary (‘dark-cyan band’) was obtained from the threshold region explained in Fig. 6.3. The red dotted line was estimated as a liquid-to-solid boundary extracted from the dilution process when starting from the micellar close-packed FCC structure.	96
6.5 (a) Equilibration process quantified by the number of connected pairs at temperatures between 0.28 and 0.32. The system starts at a randomly generated non-binding initial state and eventually reaches its steady state, where the number of connected pairs reaches a plateau (marked as ‘steady state’). (b) Degree of association $\theta(T)$ calculated from the averaged number of connected pairs in steady state (for both cooling-down and heating-up ramps). Data was measured from a system with $\phi = 0.38$ and a total number of copolymers $N = 500$	99
6.6 (a) Illustration of the calculation of the radial distribution functions $g(r)$. The PPO cores of each micelle are coloured red with the centre of mass depicted as black dots. The white dot represents the centre of mass of a reference micelle, from which the relative distances are calculated in shells of size δr , which represents the resolution term of the calculation. (b) The calculated radial distribution function for F108 micelles with complementary DNA patches at $\phi = 0.28$, $T = 0.31$ (above T_m). Each trail originate from 3 million time steps sampled every 10k time steps. The red dashed line shows the $g(r)$ averaged over all the trials.	100

6.7 Radial distribution functions $g(r)$ for F108 Pluronics with DNA patches ($T_m \approx 0.295$) for different volume fractions and temperatures. The left legend represents the estimated curves for $T \in [0.28, 0.34]$ with fixed $\phi = 0.38$, the right legend represents $g(r)$ for $\phi \in [0.08, 0.65]$ with fixed $T = 0.31$. All measurement points are illustrated in Fig. 6.8 with corresponding coloured, cross-shaped stars.	101
6.8 The phase diagram of the F108-DNA system. The DNA melting temperature here is $T_m = 0.295$ (solid black line) with a hybridization region in $T \in [0.29, 0.31]$ which we denote as transient region. The forward process covers $\phi = 0.08 \rightarrow 0.38$; the backward process covers $\phi = 0.65 \rightarrow 0.38$. The blue dashed curve represents the CMT, the red dashed curve represents the micellar liquid to solid transition. Seven different regions are identified (from bottom to top): (1) associated unimer liquid; (2) micellar gel; (3) transient unimer liquid; (4) transient micellar liquid; (5) unimer liquid; (6) micellar liquid; (7) micellar solid. The cross-stars mark the phase points for which we calculated $g(r)$ with a colour representation corresponding to the results in Fig. 6.7.	103
6.9 Phase diagram of the F108-DNA system for a DNA melting temperature $T_m = 0.32$ with a melting region in the range $T \in [0.31, 0.33]$. Two additional transient/associated micellar solid phases occur. Blue dashed curve represents the micellisation boundary, and the red dashed curve represents the micellar liquid to solid phase transition.	104
6.10 Illustration of (a) Ring formation in associated unimer phase. (b) Micelles with intra-micellar DNA loop on in micellar gel phase. Both looping parts are coloured in black.	105
6.11 Second level coarse-graining of a homogeneous and a heterogeneous system with density=0.382. The PPO-cores in each micelle are replaced by single particles at the center of mass of each micelle, shown in red to the right.	108
6.12 Illustration of two different choices for calculating $S(q)$ by 1D FFT for a range of \mathbf{q} , shown here with the normal vectors for selected scattering planes in real space.	109
6.13 The structure factor for the heterogeneous and homogeneous micellar system shown in Fig. 6.11a and b respectively; the smooth line serves here as guide to the eyes. The data were averaged over 11 trajectories, and the error bars are given by the standard error for each bin.	110
A.1 A detailed schematic of Y-shape + L6 linker DNA.	137

B.1 The structure factor for a system with 125 particles with the WCA potential. The orange line is the analytical solution with the Percus-Yevick approximation from Eqn. B.17, and the blue and green points are the estimated $S(q)$ by 3D FFT and 1D FFT along crystal planes respectively. The error bars give the standard error.	144
D.1 Schematics in different stages of melting process for Y+L6 system.	155
D.2 Monitored R_g of open and close ratio of Y+L6 sysytem at 26 °C, NaCl = 100nM.	156
D.3 Monitored R_g of open and close ratio of Y+L6 sysytem at 44 °C, NaCl = 200nM.	156
D.4 Monitored R_g of open and close ratio of Y+L6 sysytem at 50 °C, NaCl = 200nM.	157
D.5 Monitored R_g of open and close ratio of Y+L6 sysytem at 56 °C, NaCl = 100nM.	157
D.6 Example snapshot (A) [$T < T_m$, $T < \text{CMT}$]: unimers are linked by DNA patches. Long chains are formed in an associated unimer liquid phase. System size: N=500; Volume fraction: $\phi = 0.38$. PPO parts in red; PEO parts in blue; the yellow and green beads represent complementary DNA parches.	158
D.7 Example snapshot (B) [$\text{CMT} \approx T < T_m$]: micellisation begins, micelles are forming with DNA connected. The system is close to a micellar gel phase and showing a heterogeneous status (N=500, $\phi = 0.38$).	159
D.8 Example snapshot (C) [$\text{CMT} < T \approx T_m$]: micellisation finished, micelles formed with partially DNA connected. The system is in a transient micellar liquid phase and approaching a homogeneous status (N=500, $\phi = 0.42$).	160
E.1 Schematic of standard DNA nanotube.	161
E.2 Schematic of DNA nanotube with programmed sub-pore.	162
E.3 (a) Monodisperse and polydisperse systems. β represents the diameter distributions, while σ is the average size of the hard-spheres. (b) Polydisperse representation of a typical micelle system at a volume fraction $\phi = 0.28$; the measured average size of the centre PPO core is $\sigma = 13.2$ with $\beta = 0.16$	163
E.4 Illustration simulated Pluronic F127 with Y-shape DNA as connectors. Blue blobs represent PPO parts. PEO parts are coloured green, and the PEO blobs are not showing to visualise PPO cores. The Y-shape DNAs are presented as a black scaffold with DNA sticky ends in yellow. The compensatory DNA sticky ends on micelles are coloured red.	164

List of tables

3.1	ssDNA sequences and relevant melting temperatures of the Y- and L-shaped DNA structures. The different segments are colour-coded according to their function or complementarity.	33
3.2	Influence of added salt on the probability $P_{c/o}$ to find the Y+L6 system in the open or closed configuration at a fixed temperature. More related results are available in Appendix D.2	40
3.3	Final-state information from the simulations obtained at different temperatures ($^{\circ}\text{C}$).	42
4.1	Values of stiffness factor $f(n)$ according to the number of bases on the ssDNA flexible joint in nunchucks DNA.	59
5.1	Transition information from isotropic to nematic phase in varies system size.	76
6.1	The tuned parameters of non-bonding potentials in the PPO-PPO attractions (U_{BB}). T_{sim} is the simulated temperature (1 unit of $T_{sim} = 1000K$), and T_{exp} is the corresponding temperature in the experiment.	90
6.2	The collected $P(\text{micelle})$ and $P(\text{unimer})$ values at $\phi=0.38$ (blue curve in Fig. 6.3). $\langle N_{agg} \rangle$ is the averaged polymer number in micelles. The phase stage symbol ‘U’ represents the ‘unimer liquid’; M means ‘micellar liquid’.	95

1 | Introduction

1.1 Preamble

The field of deoxyribonucleic acid (DNA) nanotechnology has grown rapidly in recent years. The progressive improvement in the ability to harness the selectivity of DNA base pairing to self-assemble into a multitude of diverse structures has fostered numerous scientific studies [1–4]. A successful application of DNA self-assembly in biological science is that DNA can be an excellent building block to self-assemble into functional networks in a well-ordered manner. Another distinctive advantage of using DNA as a building block is that it can be addressed with nanometer precision. These nano-sized DNA building blocks have both rigid and flexible properties. On the one hand, double-strand DNA (dsDNA) provides rigid structural properties to form the most stable duplexes; on the other hand, single-strand DNA (ssDNA) with much less rigidity acts as flexible junctions between duplex sections. The additional attractive feature is that DNA building blocks can be made with functional responses, usually to the presence of targeted ssDNA but also to other molecules [5], to light [6], to pH [7] and to temperature[8]. There is no doubt that DNA has become a startlingly successful nano-biomaterial, with exponents of DNA nanotechnology having developed an impressive array of DNA nanodevices [9] and nanostructures [10].

Although synthetic DNA technology utilising phosphoramidite chemistry is feasible and effective [11, 12, 4], conducting experiments on DNA nanotechnology remains time-consuming and complicated. The cost of operating DNA synthesisers keeps high because it requires toxic chemicals that are only stable for a short period of time [13, 14]. Additionally, maintaining an appropriate DNA purification system to obtain high-quality and purified DNA necessitates even more complex procedures, yet it is essential for the successful completion of experiments [15, 16]. Therefore, when seeking to understand a particular property of a DNA-based system, one of the most capable solutions is to use numerical simulations. In particular, simulating a DNA nanostructure has become increasingly well-developed in recent decades [17–21]. However, owing to the various purposes of the simulation design, these existing models are limited to a specific level of details to describe DNA. The physical

description of a large-scale simulation model is also confined by its coarse-grained level. Therefore, building up a one-size-fits-all coarse-grained model is incredibly challenging, especially for a many-body system that is purely made by DNA. The best solution would provide a simplified description of DNA building blocks while, ideally, preserving enough of the fundamental physics to allow for accurate modelling of DNA systems on significant time and length scales.

In this thesis, I introduced a multi-level coarse-grained strategy to simulate DNA functionalised building blocks. The idea is to firstly examine the thermodynamic and dynamic behaviours which accompany the flexibility of a single DNA building block, then to use the detailed information collected from the first-level to build up a more reliable second-level coarse-grained model. A modelling method like this can provide a good compromise between resolution and computational speed for assembly transitions. The ultimate task is to use this specified multi-level coarse-graining method to address the question of interest to generate increased understanding and to provide convincing physical explanations of the experimental findings.

The major challenge of designing a multi-level coarse-grained model concerns which coarse-grained level the model should be fitted to explain a specific physic question. There are two main approaches to tackle this challenge. The first is a ‘bottom-up’ approach, which fits the model to the results from a finer-grained model. For instance, when simulating a single nucleotide, we can use an empirical all-atom model such as AMBER [22] or CHARMM [23]. However, these models are overly detailed when we look at the mechanical properties of an entire DNA particle. Secondly, the ‘top-down’ approach directly fits the parameters of a physically motivated potential to reproduce measured experimental properties. For example, in polymer theories, DNA can be modelled as a worm-like chain, describing DNA as a polymer with a specific bending modulus. However, such a description begins to break down for structural properties that involve more rigidity of short duplexes.

My thesis aims to bridge the gap between the detailed bottom-up and the coarse-grained top-down approach to produce a perfectly fitting model of specified physical questions. To do so, I have drawn on the well-studied oxDNA model [21] for simulating an individual DNA building block and to numerically analyse the detailed thermodynamic and topologic behaviours. The resulting findings were then used as starting point for large-scale simulations with many repeated DNA building blocks. To improve the modelling method, I have also designed a homemade system with many-body simulations to show high performance in the accuracy of predictions for the structural properties and phase transitions of the systems observed in the experiments. Eventually, I developed a clearer view of large-scaled systems

controlled by tunable parameters (i.e., concentration, temperature, etc.), which allowed us to optimise the targeted experiments in practice.

1.2 Thesis outline

The structure of this thesis is as follows: In this chapter, I provide a general overview and the motivations for building up multi-level coarse-grained simulations. In Chapter 2, I first introduce the background of using DNA as a functionalised building block. Following this, I discuss a popular DNA nanotechnology field in functional materials, especially for DNA hydrogels, DNA liquid crystals and DNA fictionalised micelles. The chapter then proceeds to discuss the main numerical methods used in the overall projects, with a detailed explanation of the simulation tool in oxDNA for DNA building block simulation and LAMMPS for larger-scale simulation.

In Chapter 3, I report the main findings of using oxDNA to simulate Y-shaped DNA nanostars connected via rigid, linear DNA-linkers. I present the work that was necessary to understand and interpret our experimental data. In particular, I show that for appropriately long, flexible ssDNA joints linear-dsDNA linkers can bind to two arms of the same Y-DNA nanostar, which we refer to as self-looping behaviour. The resulting ‘key-shaped’ DNA building blocks will lead to a fluid cluster phase rather than a space-filling DNA hydrogel. How long these flexible joins need to be will be discussed here.

In Chapter 4, I provide a quantitative study of the customised bending free energy calculation of our novel designed ‘nunchuck-like’ DNA mesogens, which serves as the prerequisite insight for a more coarse-grained simulation that we prepared in Chapter 5.

Then in Chapter 5, I mainly discuss the second coarse-grained simulation with the aim of finding new liquid crystal phases in a homemade many-body system. I firstly test the coarse-grained model in a rigid-rod case and then add flexible components in the middle of the rod-like structure to see the effects.

In Chapter 6, I discuss the numerical work for our DNA-functionalised micelles. I build a model for PEG-PPO-PEG triblock copolymers (F108) with DNA strands on each side. To do this, I mapped out the phase diagram for the F108 system with no DNA attached, matching it to our experimental findings. When I attached DNA strands I found a much changed phase diagram. To investigate the newly emerging structures I analysed them in terms of radial distribution functions and structure factors.

Finally, in Chapter 7, I conclude the thesis. It provides a summary, and presents possibilities for future work that have emerged from my PhD research.

2 | Background

2.1 DNA as a Building Block to Design Nano-Structures

2.1.1 Molecular structure of DNA

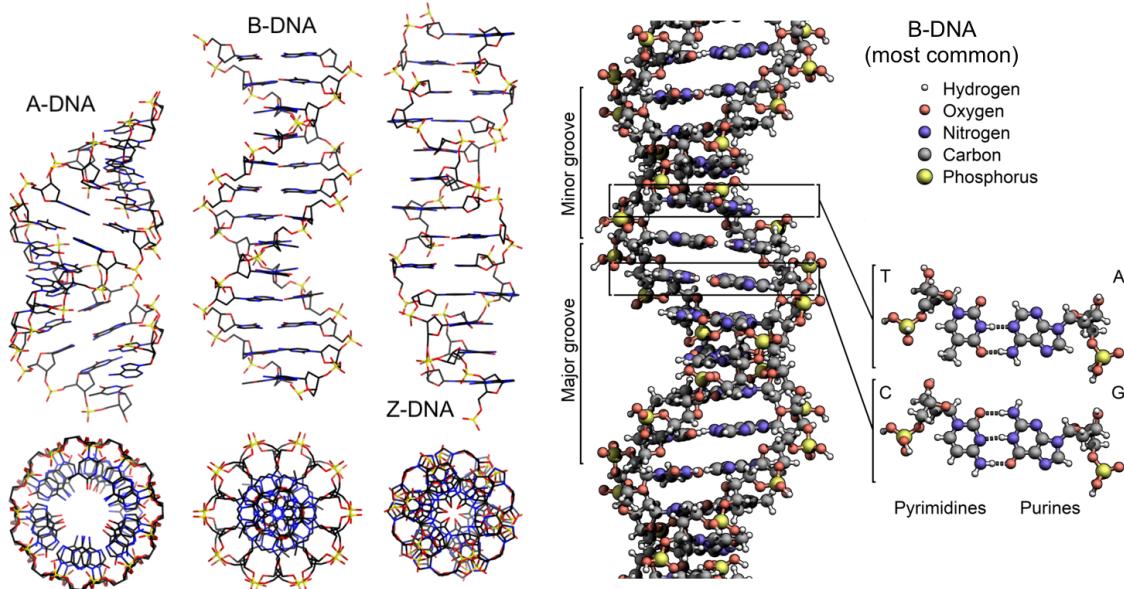


Fig. 2.1 Multi-forms of DNA (A-DNA and B-DNA are right-handed, and Z-DNA is left-handed. B-DNA is the most commonly found conformation in nature).¹

Deoxyribonucleic acid (DNA) is a well-known double-helical molecule carrying all genetic information of living things. A DNA molecule (also called double-strand DNA or in short dsDNA) consists of two equally long, flexible polynucleotide chains that are referred to as single strand DNA (ssDNA). Both strands are composed of four types of nucleotides: Adenine (A), Cytosine (C), Guanine (G), or Thymine (T). Each nucleotide comprises a five-carbon sugar ring and phosphate group, forming the backbone, and one of the nitrogenous

¹Fig. 2.1 is adapted from Wikipedia- The Free Encyclopedia (<https://en.wikipedia.org/wiki/DNA>)

base that binds to the opposing nucleotide on the complementary ssDNA via hydrogen bonds [24]. A DNA duplex will form only for opposing AT or GC pairs with two and tree hydrogen bonds respectively (right image in Fig. 2.1). Moreover, for the complementary sequences to bind also the directionality of the backbone has to be correct: The phosphate end is called the 5'-end while the sugar ring is called the 3'-end. There are three commonly regarded forms of DNA (A, B, Z) (left image on Fig. 2.1): B-DNA is the most common form both in living organisms and in material science; A-DNA, is a form of the DNA duplex observed under dehydrating conditions; Z-DNA is a relatively rare left-handed double-helix and its function is still unclear [8]. The B-form is the primary structure of DNA, and it is also the main conformation that mainly used in this thesis. A more detailed model of B-DNA has been illustrated in Fig. 2.2.

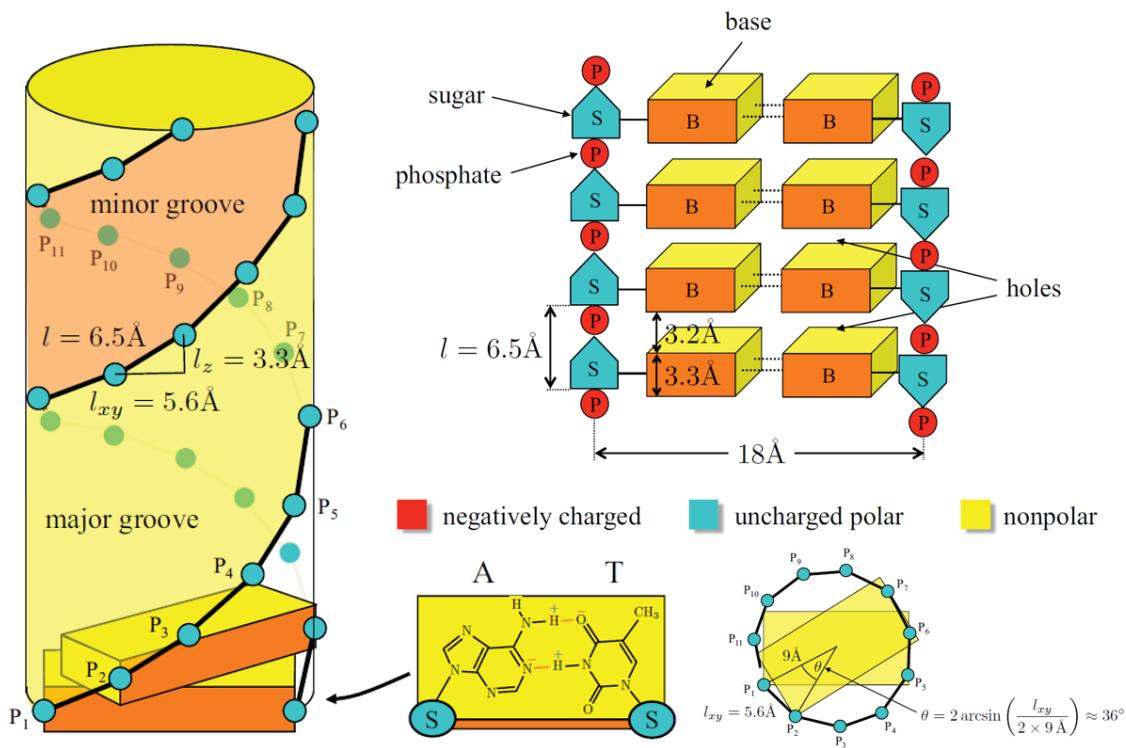


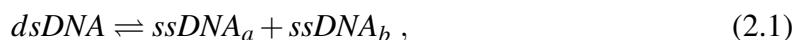
Fig. 2.2 A ladder-like structure of DNA built by sugar-phosphate backbones via base pairs. The hydrophobic holes allow successive base pairs to interact via $\pi - \pi$ stacking interactions in water. Adapted from [8]

Watson and Crick described the double-helix structure in 1953, solving the mystery of the DNA structure[8]. Their discovery was based on the careful X-ray diffraction work of Rosalind Franklin, providing direct evidence of the helical shape of DNA [8, 24]. Fig. 2.2 exhibits the ladder-like helical structure of DNA. There are several essential values describing

the DNA geometry: The first value is the rotational angle $\theta \approx 36^\circ$ between two adjacent stacking bases (this lead to about ten nucleotide pairs per helical turn). The second value is 0.33nm, which is the height between each base (it is useful to calculate the fully stretched DNA length by knowing the number of bases). The third value is 1.8nm, which indicates the width of the dsDNA (sometimes a width of ~ 2 nm is used for simple estimates). The rigid structure of the duplexes and the binding specificity between the ssDNA strands is used as building blocks in DNA nanotechnology: (1) The programmability of the sequences allows the design of nano-stars with variable number of dsDNA arms, or other blocks that then can self-assemble into finite-sized two- or three-dimensional structures with functional features; (2) dsDNA is relatively stiff biopolymer, whose persistence length is ~ 50 nm corresponding to about 150 base pairs, while the persistence length of ssDNA is only 1 nm or 3 bases long; (3) the double helix has a major and minor groove with a length that varies slightly with the given sequence, and makes dsDNA chiral. (4) The negative charges associated with the phosphate groups render DNA negatively charged. Hence, for two complementary ssDNA strands to bind via H-bonds the solution needs to contain salt, which screens the Coulomb repulsion between the backbones. All these features make DNA a versatile molecular building-block for the creation of functional materials, in particular biomaterials [25].

2.1.2 DNA thermodynamics

The hybridisation of ssDNA strands into duplexes is temperature dependent can be described as a thermally reversible melting-binding process. The melting temperature (T_m) is defined as the energy required to break half of the hydrogen bonds between two complementary ssDNA strands. Or in general, T_m is the temperature at which single and double strands are half the initial DNA concentration in solution (half of all N H-bonds present are unbound). This melting process can be described in terms of a simple, chemical reaction of the form $AB \rightleftharpoons A + B$, or



where $ssDNA_a$ and $ssDNA_b$ are complementary strands. The equilibrium constant K_{eq} of this reaction describes the change of the systems' enthalpy ΔH as function of temperature and is simply the ratio of the products to educts at standard pressure. It is known as Van 't Hoff equation:

$$\frac{d}{dT} \ln K_{eq} = \frac{\Delta H}{RT^2} , \quad K_{eq} = \frac{N_{<ssDNA_a>} N_{<ssDNA_b>}}{N_{<dsDNA>}} . \quad (2.2)$$

R is the gas constant ($=1.987 \text{ cal/Kmol}$), T is the absolute temperature. The binding Gibbs free energy ($\Delta G = \Delta H - T\Delta S$) of the process is then given by:

$$\Delta G = \Delta G^\circ + RT \ln K_{eq} . \quad (2.3)$$

Here, $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$ is the Gibbs free energy of the hybridisation reaction at standard conditions. The total DNA concentration C_τ in the system will be:

$$C_\tau = N_{ssDNA_a} + N_{ssDNA_b} + N_{dsDNA} . \quad (2.4)$$

At equilibrium, the concentration of both single and double strands will remain at half of the initial concentration of dsDNA (C_τ). As each dsDNA contributes with two single strands, hence, $K_{eq} = \frac{C_\tau}{2}$. Since $\Delta G = 0$ at equilibrium, Eqn. 2.3 can be solved in terms of T_m . Substituting the value of K_{eq} we obtain

$$T_m = \frac{\Delta H^\circ}{\Delta S^\circ + R \ln(\frac{C_\tau}{2})} . \quad (2.5)$$

Typical melting curves measured for a solution of two complementary ssDNA strands of 20 bases is shown in Fig. 2.3. The heating and cooling curves were measured in buffer solution containing 75 mM NaCl. These absorption curves were measured at 260 nm, a wavelength that is adsorbed stronger by ssDNA than dsDNA in aqueous solutions. When all DNA is melted or fully hybridised the absorption curves become linear. The median between them is used as T_m . Note that T_m also depends weakly on the duplex concentration. The overlap of the curves signifies that we have a thermally reversible equilibrium reaction.

Eqn. 2.5 is derived assuming a two-state model for the duplex melting: The dsDNA and ssDNA state. Nucleic acids can melt through a number of intermediary states. The effect of such intermediate states becomes especially important for long sequences. In Eqn. 2.5, ΔS° and ΔH° are roughly proportional to the number of bases, usually given for the association and not the dissociation reaction [26]. The values for the standard quantities can be calculated according to the Nearest-Neighbour (NN) model [27]. Therefore, longer dsDNA strands exhibit higher melting temperatures as a result. The NN model was further refined by Di Michele et al. [28] by considering the entropic contributions of inactive tails, which can be non-binding, flexible linkers, for instance.

Eqn. 2.5 can also be adjusted to account for the presence of added salt by adding a correction factor $m(T)$ to the entropy term ΔS° . Using the sodium concentration $[\text{Na}^+]$ as an

example we can write

$$T_m = \frac{\Delta H^\circ}{\Delta S^\circ + \sum_{n=1}^N \frac{m(T)}{T} \ln[Na^+] + R \ln(\frac{C_\pi}{2})}, \quad (2.6)$$

where the summation runs over all N bases of the ssDNA strand. The tabulated values of $m(T)$ were reported by SantaLucia et al. [27]. As the phosphate groups in the backbone

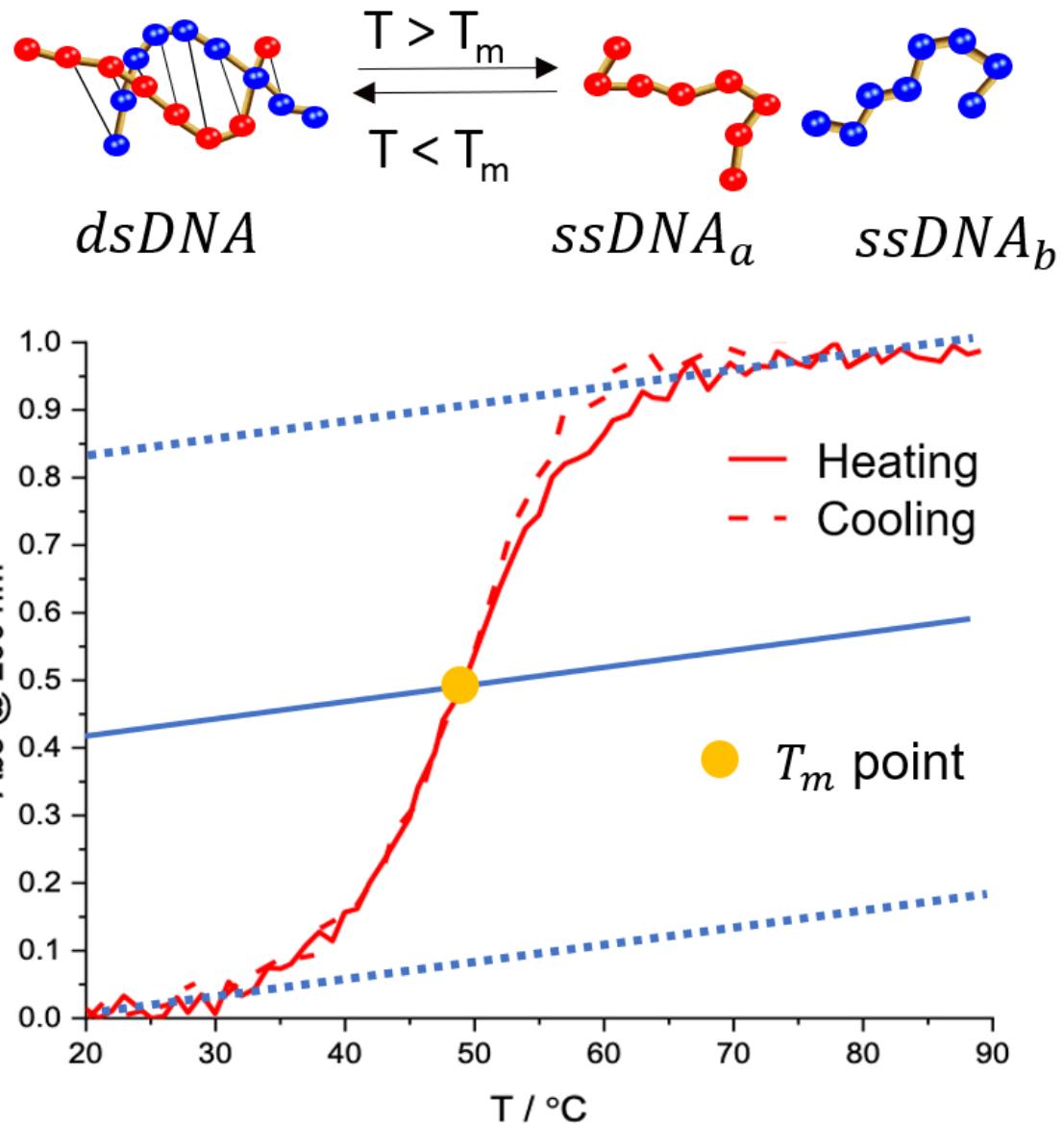


Fig. 2.3 The melting temperature of a known DNA duplex is done by measuring the absorbance at 260 nm as a function of temperature. Dotted lines were fitted to high- and low-temperature plateaus. The median (yellow point) signifies T_m .

are negatively charged, one needs to add monovalent salt to screen these charges. For NaCl concentrations lower than about 70 mM we observed an onset of hysteresis in the melting curves, which stem from the insufficient screening. Hence most hybridisation experiments will be done at higher added salt concentrations.

2.2 DNA Functional Materials

Only in the 1940s DNA was recognised as the genetic material containing the building plan of life [29], and in the 1950s, the double-helix structure of DNA and its unique physical properties were established [8]. But only in the 1980s, when Nadrian Seeman proposed using DNA as a nano-sized building material [30], the tendency to use DNA in bio-engineering and material science has risen and became more and more widespread. The shift towards using DNA as a versatile molecular building-block for the construction of functional materials, DNA nanotechnology has led to impressive advancements, such as in DNA origami (DNA polyhedra [31], DNA boxes [32], etc.), DNA nano-devices (nanotweezers [33], DNA walkers [34], etc.), DNA polymeric-networks (DNA hydrogels [35], DNA crystals [36], etc.) and other functionalised materials [28, 37]. This thesis mainly focuses on the potential of using DNA as a building block to design active DNA hydrogels, DNA liquid crystals and DNA functionalised micelles.

2.2.1 DNA hydrogel

DNA hydrogels are polymeric networks made entirely from DNA. The convenient programmability and multi-functional tunability of DNA make it a perfect material for constructing a three-dimensional structure in hydrogels [41]. DNA hydrogels are also biocompatible and therefore useful for biomedical applications such as bio-sensing technology [42], targeted drug delivery [32], 3D bio-printing [43], regenerative medicine and tissue engineering [44].

There are many ways of categorizing DNA hydrogels made of short strands, one of which is sorting these gels by type of nanostars used to create them. Dan Luo's group made various shaped small DNA-building blocks, in particular, X-, Y- and T-shaped DNA tiles (Fig. 2.4(a)) were designed and mixed to form well-defined gels under certain solvent conditions [38]. Dongsheng Liu's group later proposed a pH-responsive DNA hydrogel prepared by a three-armed DNA nanostructure assembling through the formation of intermolecular i-motif structures (Fig. 2.4(b)). This type of DNA hydrogel can be switched to a non-gel state in minutes by using environmental pH changes [39]. Moreover, a novel type of DNA hydrogel

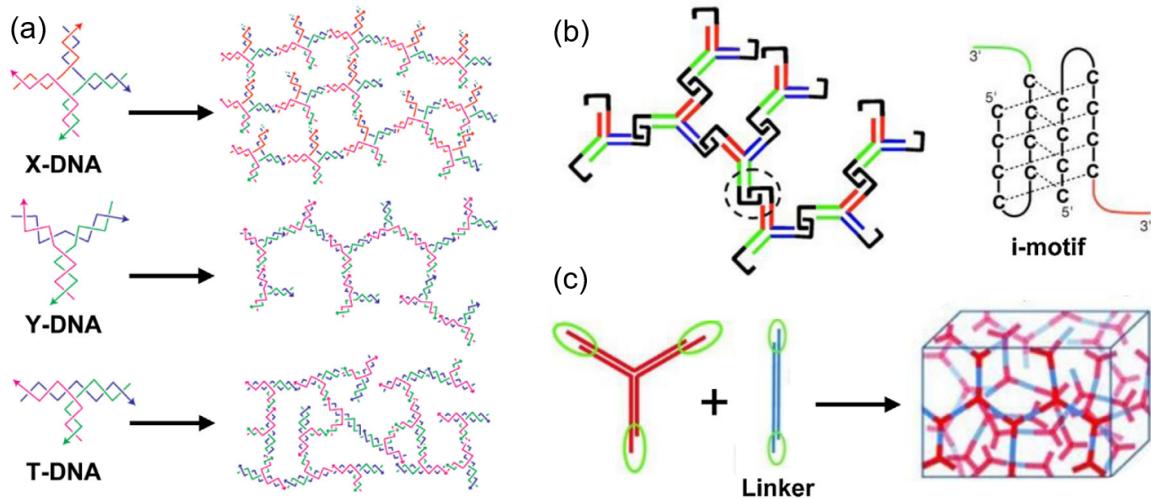


Fig. 2.4 Representative principle structures and dendritic DNA assembled DNA hydrogels. (a) DNA hydrogels made by enzymatic ligation of X-, Y- and T-shaped DNA [38]. (b) DNA hydrogel with three-armed DNA building blocks ligated together via the formation of intermolecular i-motif structures[39]. (c)DNA hydrogel made from the three-dimensional assembly of DNA with thermal and enzyme dual responsiveness[35]. Adapted from [40].

with linear DNA linkers as an extra crosslinker was introduced [35] (e.g. in Fig.2.4(c)). This linker DNA was designed with thermal and enzymatic dual responsiveness.

Over the last few years, our group has also built up DNA hydrogels made of three-armed nanostars [45](Fig. 2.5). These branched DNA nanostars are synthesised in such a way that each arm of the nanostar is a dsDNA with a longer ssDNA tail (called ‘sticky end’) with palindromic sequences. Two complementary sticky ends are able to hybridise, resulting in large-scale three-dimensional structures that can sustain mechanical stresses. Moreover, DNA hydrogels are physical gels in the sense that the DNA-building blocks are only held together via hydrogen bonds. Hence when heating them, the blocks ‘unbind’, rendering the solution a liquid [45]. The experimental observation of the presence of ssDNA linkers with different flexibility is illustrated in Fig. 2.5(a).

We further developed a coarse-grained simulation model to investigate the microrheology and elasticity of the DNA hydrogel system [46]. We used two types of identical Y-shaped DNA that consisted of a three-armed rigid dsDNA core, each type carrying one of the complementary sticky ends (blue and red dots in Fig. 2.5(b)). These sticky ends can bind to each other at a simulated low temperature. A large number of these Y-shapes then connected to form a percolating gel (Fig. 2.5(c)). Most recently, we assembled a linker-mediated hydrogel of Y-shaped DNA by adding extra flexibility [47] to investigate the behaviour of

the gelling process of tuning the length of ssDNA joints on DNA linkers. The detailed report about this work will be discussed in Chapter 3.

2.2.2 DNA liquid crystal

Gas, liquid and solid are the three well-known states of matter. Liquid crystals (LCs) are materials that have also intermediate phases that exists somewhere between the liquid and crystal phase. Crystalline solids display long-ranged periodic order in three dimensions whereas liquids show none; LCs can flow like liquids and retain some order in specific spatial dimensions. LC phases can only be formed by molecules with shape or dipolar anisotropy; anisotropic molecules or building blocks are often called ‘mesogens’. Three typical LC phases are shown in Fig. 2.6:

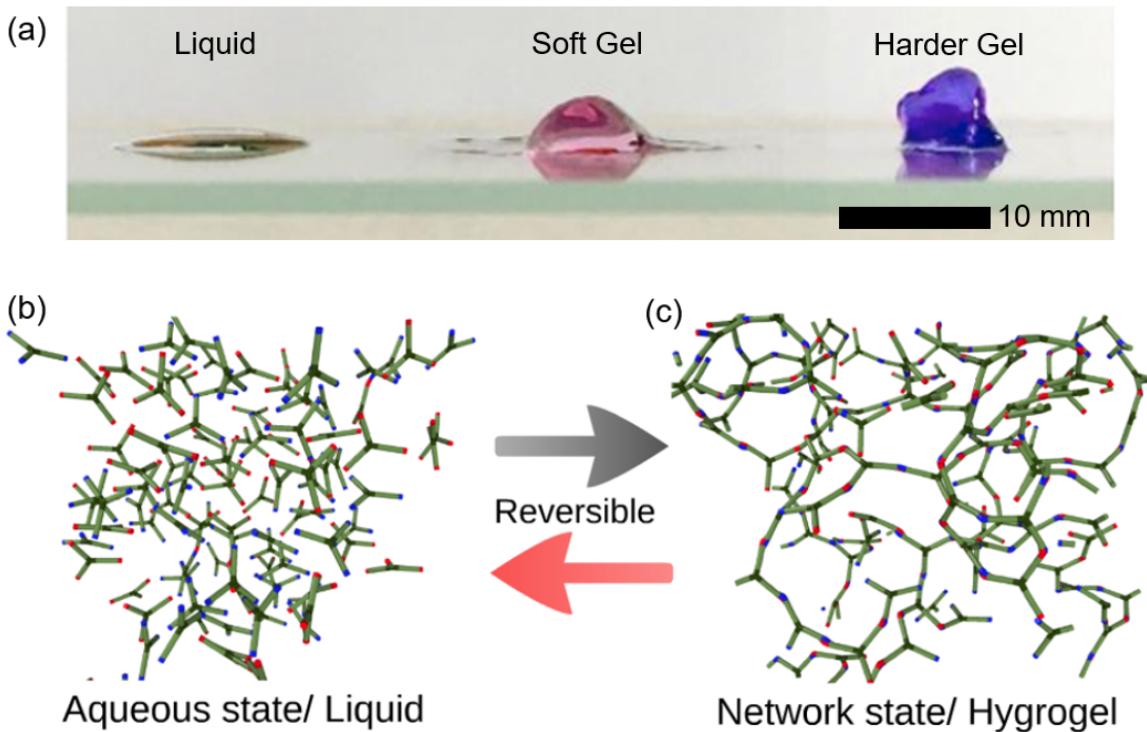


Fig. 2.5 (a) Photograph of a droplet containing 0.2wt% Y-shaped DNA with (from left to right) no sticky ends, with sticky ends and flexible linkers and with rigid sticky ends, produced by Zongyang Xing et al. [45]. (b) Snapshot from simulations of a solution containing dispersed, Y-shaped DNA building blocks above the T_m of their sticky ssDNA ends. (c) Snapshot of the same system below T_m , when the stick ends bind to form a viscoelastic hydrogel. The Y-shaped DNA is coloured in green. Blue and red dots represent the complementary ssDNA sticky ends.

- Nematic phase: The mesogens lack positional order but exhibits long-range orientational order because the molecules align, on average, along a given direction referred to as director.
- Smectic phase: Here the mesogens arrange in a 1D positional order that resembles 2D liquid layers that still can slide past each other like in a liquid.
- Cholesteric phase: These are composed of nematic mesogens containing a chiral centre that favours the alignment with other mesogens at a slight angle.

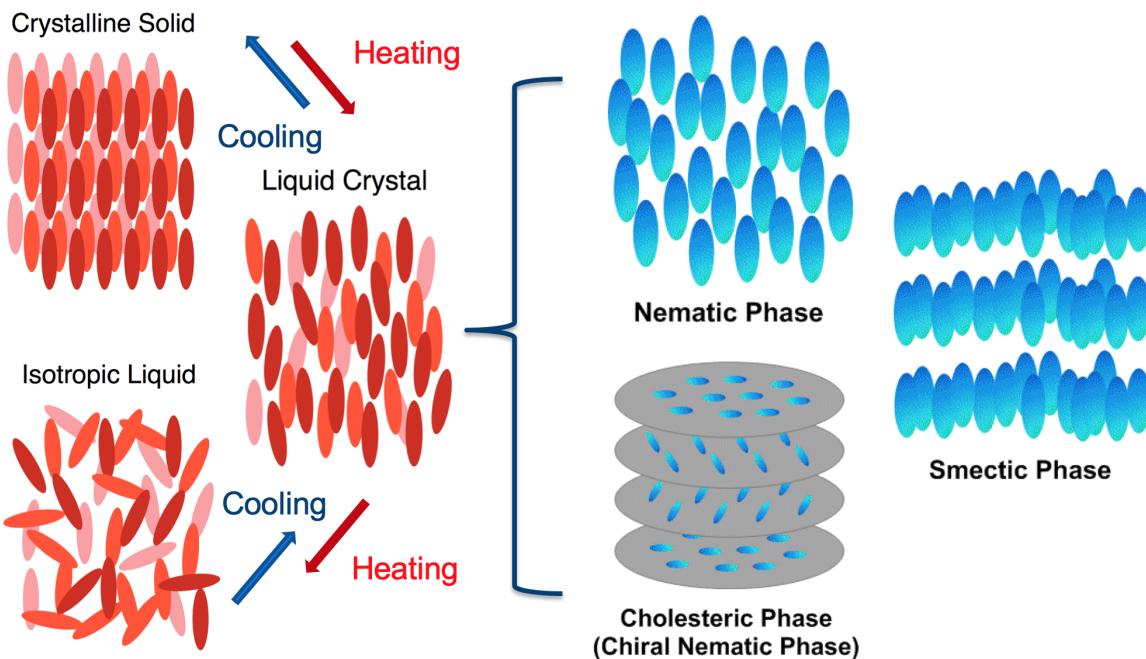


Fig. 2.6 Representation of the various phases rod-like mesogens can be in. The weakest (orientational) order is assumed in the nematic phase, while the smectic phase shows higher order with positional orientation in the layers.

LCs play an essential role in biology because they combine order and mobility, offering a basic requirement for self-organisation and structure formation in all living systems [48]. The ability of dsDNA to form LC phases when it is hydrate was first demonstrated in the 1940s [49]. Since that time, the LC properties of aqueous suspensions of DNA have been extensively characterised. The geometry and physical characteristics of DNA are closely related to its chemical structure. DNA is negatively charged when suspended in water-based solutions. One DNA molecule pushes the others away because similar charges repel, effectively taking up more volume of space [50]. When two DNA molecules are far apart, they cannot sense any form of connection because the strength of electromagnetic repulsion

weakens with increasing distance. Considering DNA's rod-like structure, these characteristics enable a simple description of a DNA fragment (one short and individual DNA molecule) as a rigid repulsive rod (Fig. 2.7(a)). The LC phases of both long [49, 51] and short dsDNA [52, 53] has been well-studied over the last few decades. In 2007 it was found that short complementary dsDNA with 6 to 20 base pairs in length can easily reach nematic and columnar liquid crystal phases [52]. The columnar phase shows two-dimensional positional order, so has the appearance of an array of liquid tubes. The structural study also shows that the end-to-end adhesion and consequent stacking of the short dsDNA can also lead to produce LC phases into rodlike aggregates [53].

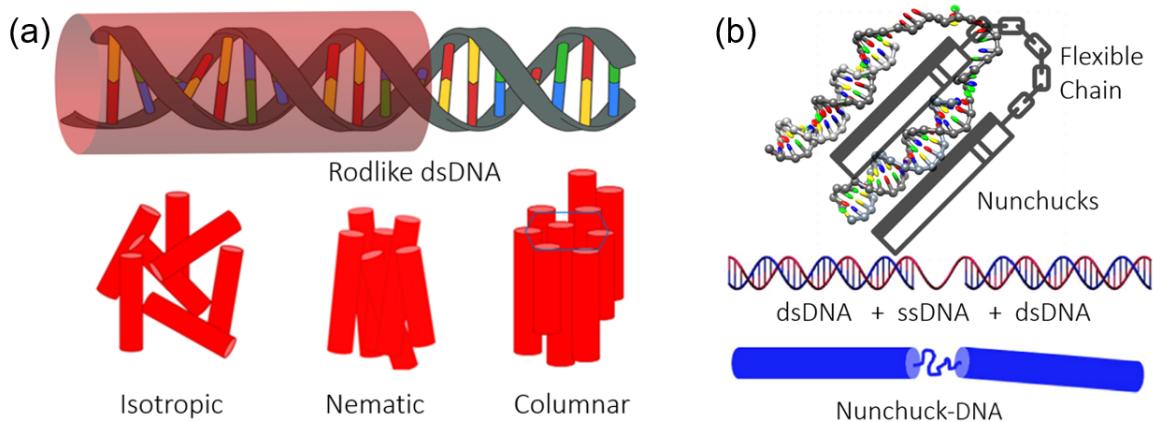


Fig. 2.7 (a) Rodlike dsDNA can produce isotropic, nematic and columnar phases. Adapted from [52]. (b) Nunchuck DNA with an unpaired, ssDNA spacer connecting two rigid duplexes. Adapted from [54].

Building a computational model can further help explain the LC phases produced by these DNA mesogens. Recently, a potential smectic phase was found when the DNA's flexibility was increased by introducing an unpaired single-stranded DNA spacer in the middle of each duplex [54] (as 'nunchuck-like' DNA shown in Fig. 2.7(b)). However, an accurate understanding of the physical mechanism that leads to the formation of such a new type of smectic phase is still unclear. This motivated us to build a homemade system to investigate the detailed mechanical properties at a microscopic level. A multi-level coarse-grained computational model of the nunchuck-DNA and its many-body system simulation will be discussed in Chapter 4 and 5.

2.2.3 DNA-functionalised triblock copolymers

Pluronics® are triblock copolymers (referred to as 'PEO-PPO-PEO') that have two hydrophilic polyethylene oxide (PEO) outer blocks and an inner block of polypropylene

oxide (PPO) that is hydrophilic at low temperatures but becomes hydrophobic upon heating (Fig. 2.8(a)). These block copolymers are polymeric surfactants, commercially also known as Poloxamers®, and are widely used in lubrication, detergency, pharmaceuticals, foaming, emulsification and coating [55–61]. It is possible to create copolymers with a wide range of qualities by varying the PPO/PEO ratio and the molecular weight. Fig. 2.8(b) shows what is referred to as the ‘Pluronic grid’. While the copolymers along the horizontal lines have a PPO block of a different length, those along the vertical lines have the same PPO/PEO composition. Pluronic triblock copolymers are notated using the letters L for liquid, P for paste, or F for flakes. The first or second number represents the PPO block’s molecular weight (multiplied by 300), while the final number represents the PEO block’s weight fraction. For instance, F108 has 147 propylene-oxide repeat units on each side of the 56 ethylene-oxide repeat units in the middle. It has an average molecular weight of around 3000 Da with 80% of EO content, and its solid form comes in flakes at room temperature.

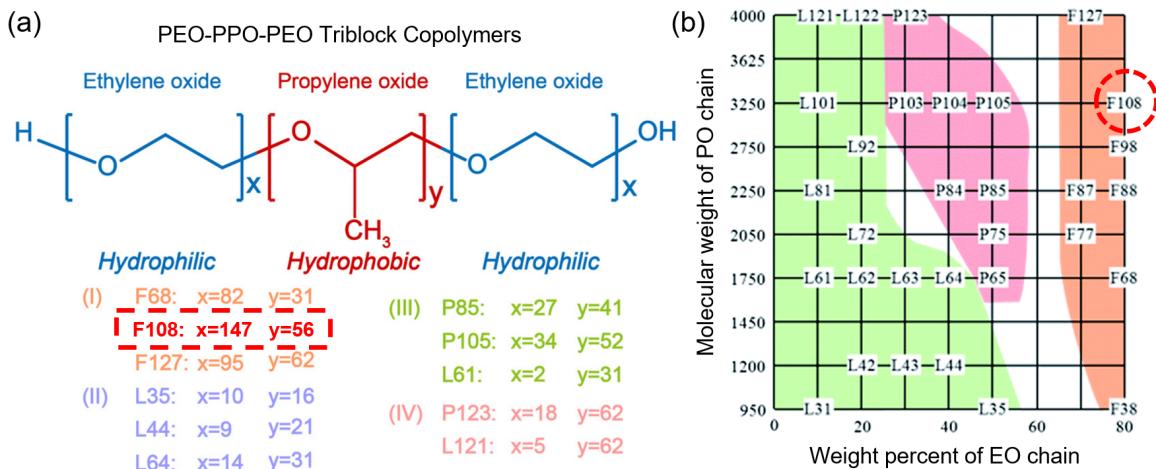


Fig. 2.8 (a) Molecular structure of a PEO-PPO-PEO triblock copolymer. Adapted from [62]. (b) The Pluronic grid (colour code: physical state of copolymers under ambient conditions: green = liquid; red = paste; orange = flake). Adapted from [63–65]. The copolymer used in this thesis are indicated by the red dotted rectangle/circle.

Micelle formation in aqueous Pluronic solutions depend on temperature and concentration. At low temperatures and concentrations F108 chains are fully dispersed forming a unimer solution. The chains start to form micelles above its critical micelle temperature (CMT). For any Pluronic system this CMT only weakly depends on the concentration. Moreover, it is not a sharp transition like in standard surfactant systems that display a sharp critical micelle concentration (CMC). Hence, above the CMT there is a wider temperature range in which micelles coexist with unimers. For low polymer concentration there is a temperature above which all unimers are converted into micelles and thus we have a micellar liquid,

as indicated in Fig. 2.9 where show the experimentally determined phase diagram of F68 and F108. Above a given polymer concentration enough micelles form to fill out the entire sample and we observe a first-order phase transition to a cubic phase (either bcc or fcc [66, 67]), often known as the gel phase (although no crosslinking is present). Note that because Pluronics are produced on industrial scale, their phase boundaries may shift from batch to batch. Moreover, they often have a large size distribution, but never fail to form micelles.

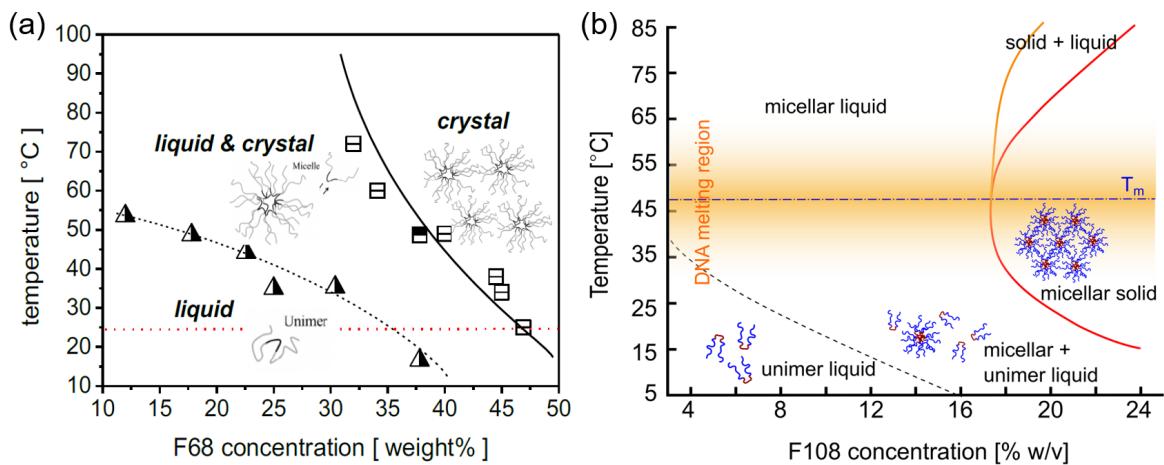


Fig. 2.9 (a) Phase diagram of the F68 system which has the same block ratio as F108 but lower molecular weight. Adapted from [68]. (b) Experimental phase diagram of F108 in 100 mM NaCl, based on optical observations. The red line marks the transition to a single micellar solid phase while the region between the orange and red line indicates a micellar solid coexisting with a micellar liquid. The dashed line indicates the CMT boundary, based on literature values [69]. Experiment performed by Ren Liu and Dr. Alessio Caciagli in our group.

Recently, our group also developed and demonstrated the potential to attach DNA to the free PEO-chain ends of Pluronics [70]. Fig. 2.10 shows the protocol of attaching ssDNA covalently to F108. First, the free PEO ends were functionalised with an azide (N_3) group. Then, a strain-promoted alkyne-azide click reaction was used to connect DBCO-functionalised ssDNA strands to these N_3 ends. These functionalised Pluronics can form so-called ‘patchy micelles’ [59].

My numerical studies are based on our groups’ experimental finding for the DNA-functionalised F108 system. As it is very costly to functionalize all F108 chain-ends in real experiments I first developed a coarse-grained model for the bare F108 system and its phase diagram and then explored the effect of the DNA-functionalization on the systems’ phase behaviour. In Fig. 2.11 I present an image of the system based on a simulation rendered in

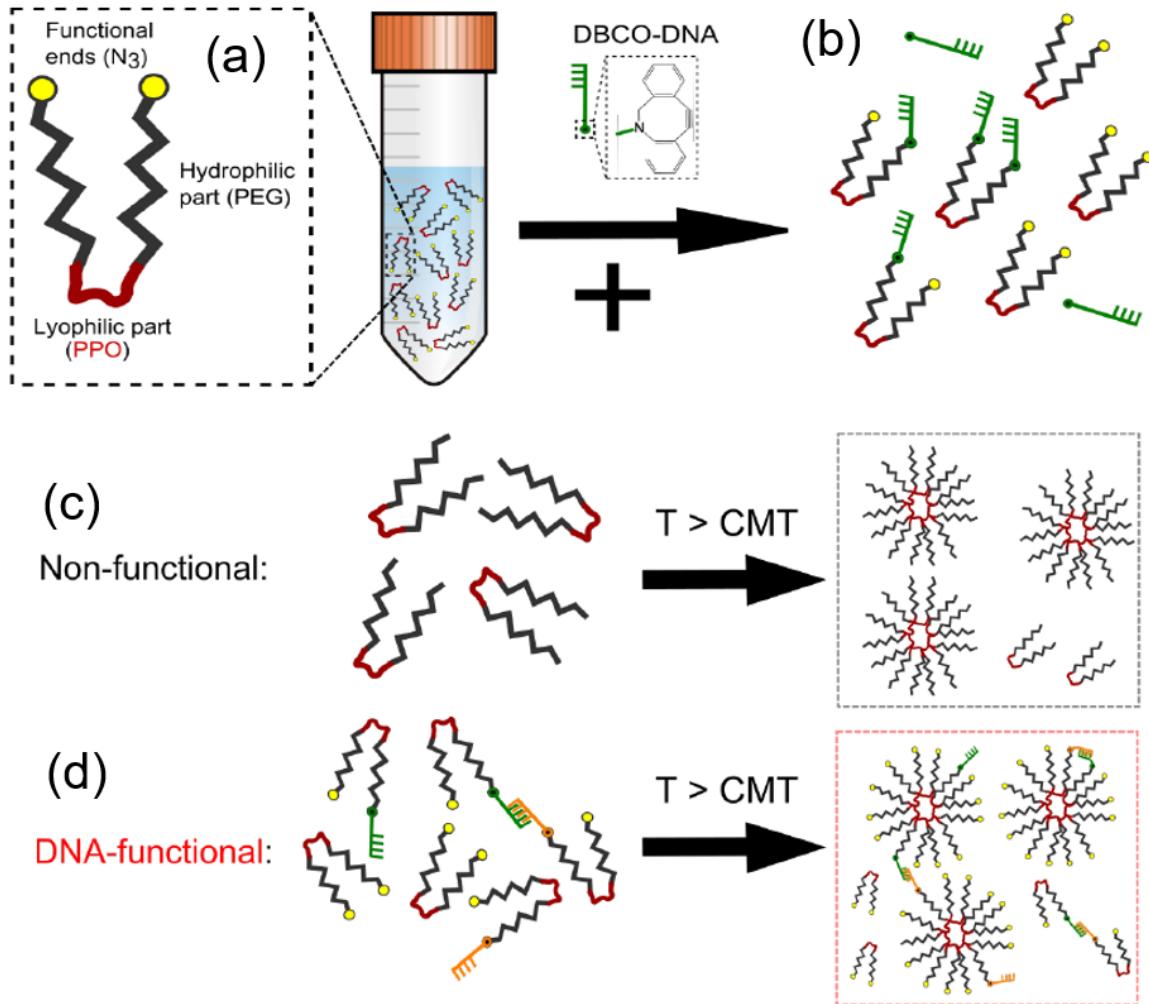


Fig. 2.10 Schematic protocol of making DNA-functionalised triblock-copolymers. (a) The triblock-copolymer PEG-PPO-PEG with azide-functionalised PEG ends is dispersed in water in an aqueous solution. (b) The functionalised PEG ends are reacted to DBCO-DNA using alkyne-azide click reaction. (c) Non-functional triblock copolymers undergo a phase transition from unimers to micelles when crossing the CMT. (d) In addition to the micellisation transition, DNA-functional triblock copolymers can assume different exotic phases below and above the CMT (e.g. short or long chains of connected polymers can form below the CMT and chains of spherical micelles can exist above the CMT). In addition, if two block-copolymers containing complementary DNA are close to one another in a single micelle, intra-micelle DNA loops may form. Adapted from the PhD thesis of Alessio Caciagli and [70].

the graphics program Blender, to illustrate F108-DNA micelles. The numerical model will be discussed in Chapter 6.

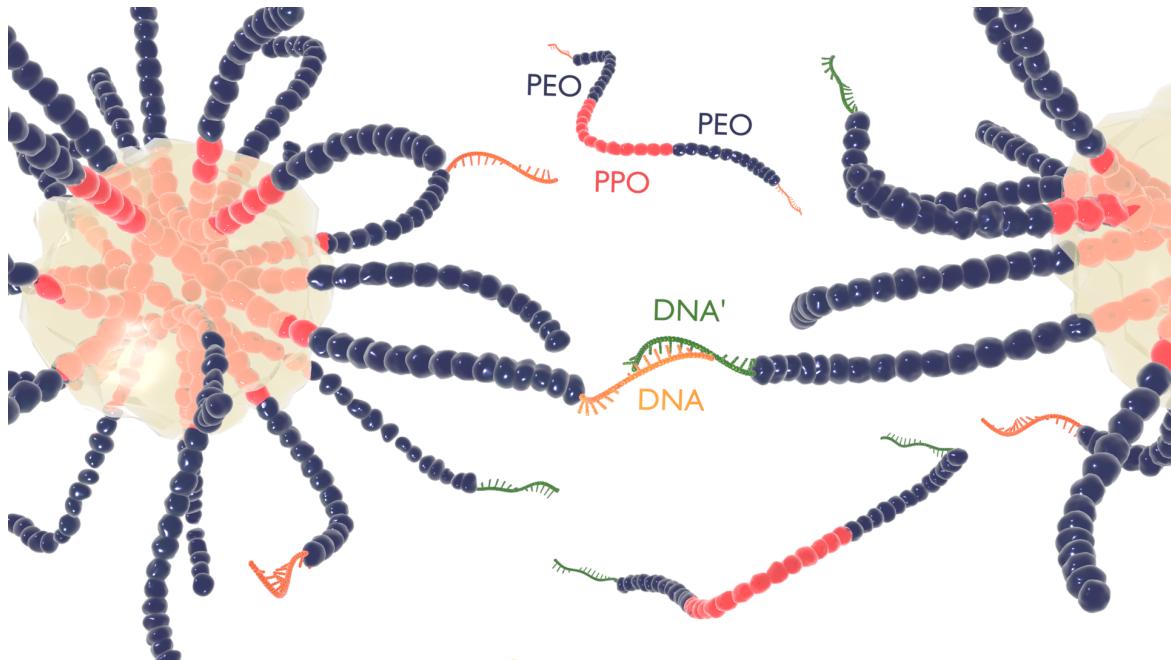


Fig. 2.11 Schematic of the DNA-functionalised triblock copolymer system (F108-DNA). The hydrophilic parts (PEO) are coloured in blue; hydrophobic parts (PPO) are coloured in red; the complementary ssDNA sticky ends are coloured in red and yellow. The transparent yellow zone represents the hydrophobic core of the micelles.

2.3 Numerical Methods

Applying numerical simulations as a bridge between theoretical calculations and practical experiments is a widely used strategy in biophysical studies. The two main families of numerical methods in molecular modelling are molecular dynamics (MD) and Monte Carlo (MC) simulations. A whole range of hybrid techniques combine features from both to calculate molecular mechanics. Such are, for instance, ESPResSo [71] (Extensible Simulation Package for Research on Soft Matter), MacroModel [72] and LAMMPS [73] (Large-scale Atomic/Molecular Massively Parallel Simulator). MD simulations can give us a route to dynamical properties of atoms/molecules (e.g. transport coefficients, time-dependent responses to perturbations, rheological/micro-rheological properties, etc. [74, 75, 46, 76]). Instead of trying to reproduce the dynamics of a system, MC simulations rely on equilibrium statistical mechanics (one of the main applications is generating states according to appropriate

Boltzmann probabilities). This section introduces the simulation methods used in this thesis, including oxDNA, a well-studied coarse-grained DNA simulation tool and LAMMPS.

2.3.1 Molecular Dynamics Simulations

The MD simulation is a computational tool that describes how positions, velocities, and orientations of atoms/molecules change over time. The basic idea of the MD simulation is to mimic what atoms do in real life, assuming a given potential energy function (force field). In the most common version of MD simulation, the trajectories of atoms and molecules are determined by numerically solving Newton's equations of motion for a system of interacting particles as a function of discrete time steps.

The forces between the particles and their potentials are often described by Derjaguin-Landau-Verwey-Overbeek (DLVO) theory [77, 78]. DLVO theory unites the balance between electrostatic repulsion and van der Waals attraction. Electrostatic repulsion becomes significant when two particles' electrical double layers begin to interfere. Van der Waals attraction is between molecules in each particle that are normally electrically symmetric. The DLVO theory explains the tendency of particles to agglomerate/separate by combining two curves of electrostatic repulsion and van der Waals attraction. The smaller value is subtracted from the larger value to obtain the net energy at each distance [77]. In solutions containing ions, the electrostatic term can be stabilised by either increasing added salt concentration or decreasing surface charge density. Also, in a DNA solution, adding salt can help neutralise the DNA charge and make the molecule less hydrophilic. In comparison to the DLVO theory, in which the strength of repulsion can be tuned continuously, it is helpful to use the typical Lennard-Jones potential [79, 80]. The Lennard-Jones potential (V_{LJ}) models soft repulsive and attractive (van der Waals) interactions to describe electronically neutral atoms or molecules:

$$V_{LJ}(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (2.7)$$

where ϵ is the dimensionless depth of the attractive well between particles (i and j) of how strongly the two particles attract each other. σ denotes a measurement of how close two nonbonding particles can get and is thus referred to as the van der Waals radius. And r_{ij} is the distance of separation between particles (measured from the center of particle i to the center of the particle j).

The general process of an MD simulation is shown in Fig. 2.12: first, we select a model system consisting of N particles and give initial positions and velocities to each of them. Then the energy function P_{ij} is defined, so the force between two particles (i, j) can be

calculated using the gradient of the energy function. Finally, we solve Newton's equation of motion for each particle using numerical integration to update their new positions and velocities. The simulation time-step can vary from picoseconds to nanoseconds according to the size of the system and the requirement of detailed balance. This MD cycle is repeated until the properties of the system (e.g. the averaged velocity) no longer change with time, which means the system is in equilibrium. After equilibration, the actual measurement is performed.

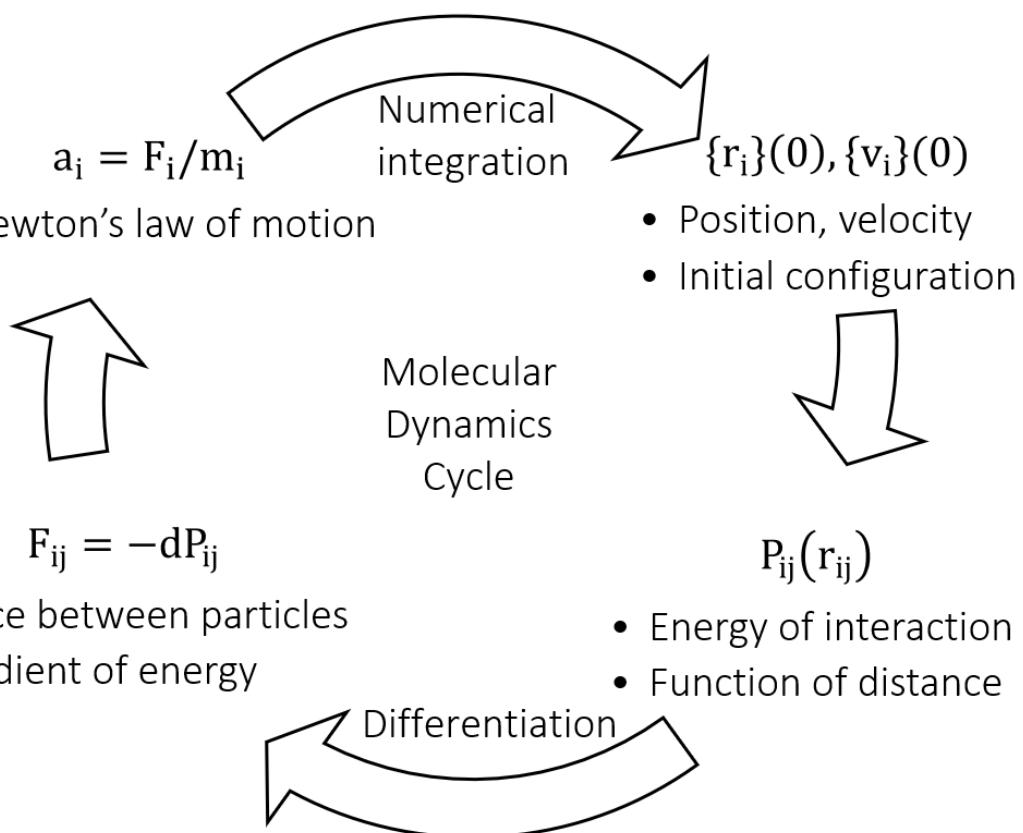


Fig. 2.12 Molecular Dynamics Simulation Cycle.

The significance of MD simulations is that the trajectories of all particles in a time-dependent process are measurable. These trajectories are recorded to calculate various dynamical properties of the system such as: (1) the temperature of the system depending on atoms' averaged moving velocity: $T = \frac{\langle m_i v_i^2 \rangle}{k_B}$ (k_B is the Boltzmann constant); (2) once the temperature of the system is established, the pressure of a closed space can be calculated (e.g. in the ideal gas model, the pressure can be written as $P = Nk_B T/V$); (3) the pressure of the system is then further used to calculate the stress for a given deformation so that the mechanical properties of the system (e.g. stiffness, elasticity etc.) can be solved. In a DNA

simulation, MD is not only an ideal numerical method to obtain time-dependent trajectory information, but also a great recorder for a detailed DNA structure and configuration in real-time thermal motion.

There are some limitations in MD simulations. Many timesteps are necessary to obtain average properties, which can be computationally expensive. Also, MD simulations cannot give any information about the electronic properties; the electrons' details are neglected instead of just points interacting with each other. But for all the other classical properties that do not depend on the electronic density around each atom, then the MD method is adequate to predict the properties and the behaviour of the system according to timesteps. MD simulation suits all coarse-grain levels depending on its fundamental simulated units (e.g. atoms, molecules, nucleotides, proteins or colloids).

2.3.2 Monte Carlo Simulations

Monte Carlo (MC) simulations are based on estimating the value of a known quantities by sampling the phase space of a system. In statistical mechanics, MC simulations are commonly used to generate energy states with a probability distribution appropriate to a certain ensemble (see Fig. 2.13). For example in a canonical NVT system (Fig. 2.13.b), classical statistical mechanics demands that the microstates of a system should be occupied with a density ($\rho(\vec{r}) \propto e^{-\beta E(\vec{r})}$), where $\beta = 1/k_B T$ and $E(\vec{r})$ is the energy function for a given state defined by \vec{r} , \vec{r} is a vector with all the degrees of freedom (e.g. for a mechanical system of many particles: $\vec{r} = (\vec{q}, \vec{p})$, \vec{r} depends on its position \vec{q} and momenta \vec{p}).

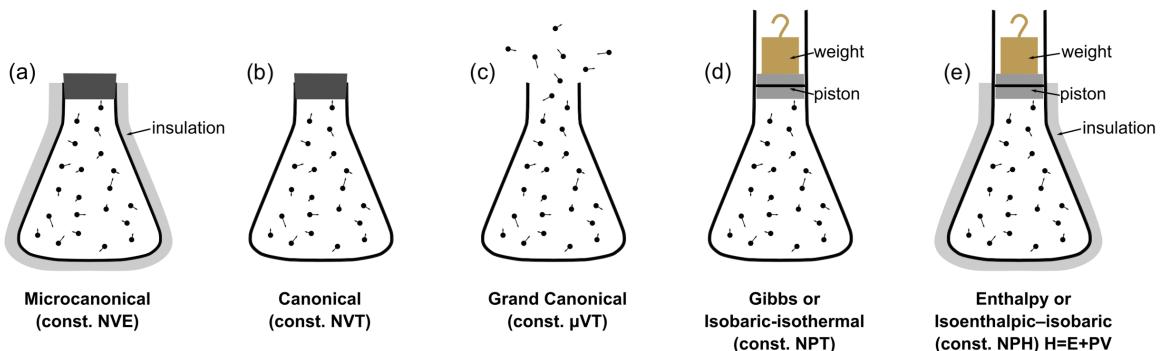


Fig. 2.13 Schematics of Common Type Statistical Ensembles in Thermodynamics ².

The essence of an MC simulation is using a large number of random samplings to find a large aggregate property of the system in the case the total number of possible outcomes is too large for computation. MC simulation then can select a finite number of these random

²Fig. 2.13 is adapted from Wikipedia - The Free Encyclopedia ([wikipedia.org/wiki/Statistical_Ensemble](https://en.wikipedia.org/wiki/Statistical_Ensemble))

outcomes to predict the overall properties of the system based on its statistical contributions. In 1953 the so-called Metropolis algorithm was developed that allowed to efficiently sample the vast phase space of a system [81]. Key to this algorithm was to find the new energy state (E_{new}) of any updated MC step, and compare it to its old state (E_{old}). If $E_{new} < E_{old}$, the new step is accepted. If $E_{new} > E_{old}$, the new step is partially accepted with a probability given by:

$$P_{accepted}(New) = e^{-\frac{E_{new} - E_{old}}{k_B T}} \quad (2.8)$$

Since its invention, Metropolis MC has been widely employed to study self-assembly processes and DNA dynamics [82]. One of the advantages of using MC simulation is that it does not involve an integration of the equation of motion. Thus the system simulated can be equilibrated much faster than a MD simulation would. Generating states according to appropriate Boltzmann probabilities is a direct application of the Metropolis MC simulation to molecular systems. Another application of using the MC approach is the Brownian motion of particles. There is no means to compute the trajectories of individual light particles (e.g. solvent molecules) colliding with heavy ones. If hundreds of thousands of these collisions are happening, then the random motion of a heavy particle can be observed.

In general, MD and MC simulation both play an essential role in molecular simulations and complement each other. They are the most used methods to simulate ensembles of molecules. MC calculates thermodynamic properties via an ensemble average, while the MD simulations are via time-reversible Newton's equations. One significant distinction between MC and MD is the method of sampling to generate the correct distribution. MC samples do not need to follow a physically allowed process, all required is an ergodic generation process, which means the ensemble average equals the time average. This could be used to speed up MC schemes to generate a desired configuration. However, the acceleration of MC sampling implies that all information about the actual physical process is lost. This is where MD becomes more advantageous: MD simulation follows the physical path, which allows learning more about the equilibrium configurations and how to physically achieve there. In most of the simulation designs in this thesis, we used MC to generate the initial configuration of the system and then applied MD to produce equilibrium states. MD probes the dynamics of the system under study and can therefore be used to compute transport coefficients that the MC simulation cannot do. As we want to analyse systems' dynamic behaviours (such as the bending behaviour in Chapter 3 and the micelle formation process in Chapter 6), for most of the investigated systems in this thesis, MD is the method of choice. Another practical motivation for using MD in DNA simulation is that it can provide a better glimpse into more

realistic physical models, which can benefit us to visualise the real DNA movements that we cannot directly see from our experimental observations.

2.3.3 Langevin Dynamics

In Langevin dynamics, the random forces due to the incessant collisions between the liquid molecules and the larger particles studied are considered. The time- and length-scale of larger particles like colloids dispersed in a solvent made of much smaller molecules is many orders of magnitude different. The rapid, stochastic forces acting on the colloids can be expressed by extending Newton's equation of motion of the particles with a frictional term and a stochastic contribution coming from these collisions between particles and fluid molecules. The general version of the Langevin Eqn. 2.9 is

$$m \frac{dv(t)}{dt} = -F(r,t) - \gamma v(t) + R(t), \quad (2.9)$$

where $F(r,t)$ is the force acting on a single particle due to the interaction potential (it usually depends on particles' position and time). $F(r,t)$ can include a systematic force if required- e.g. gravity, electrical field. etc. $\gamma v(t)$ is a frictional drag or viscous damping term proportional to the particle's velocity. In an MD simulation, the friction coefficient (γ) is computed as m/λ , where m is the mass of the particle and λ is the damping factor specified by the user (this damping factor - λ can be thought of as inversely related to the viscosity of the solvent). Finally, $R(t)$ is the random force the particle experiences due to the solvent around it. It is expected to have the following statistical properties:

- The random force is caused by collisions of light particles from any direction. Hence the averaged value of the random force is zero, which tells there is no bias in the process ($\langle R(t) \rangle = 0$).
- There are no time correlations before the heavy particles significantly change their direction or position: it assumes that the collisions are instantaneous and successive collisions are uncorrelated. It can be expressed by a delta function (δ) for the autocorrelation function of the random force, $\langle R(t)R(t') \rangle = q\delta(t-t')$, with q being a constant equal to $2\gamma k_B T$ (A derivation can be found in reference [8]). The δ function form of the correlations in time means that the force at a time t is assumed to be completely uncorrelated with it at any other time t' .
- The random force is assumed to be Gaussian ($P(R) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{R^2}{2\sigma^2}\right)$).

The Langevin equation describes the Brownian motion of particles in an MD simulation. It allows the temperature to be controlled via a thermostat, thus approximating the canonical ensemble. It has been widely employed in many molecular simulators (e.g. the fixed Langevin standard in LAMMPS for a fixed NVT ensemble). It is also used in the oxDNA model to describe a random force on a group of nucleic acids, modelling interactions within a background of an implicit solvent [83]. However, in a coarse-grained simulation, the Langevin dynamics may have some limitations. The most crucial concern is that the Langevin equation neglects the hydrodynamic interactions in a dense state. Hence it can not accurately predict the diffusion coefficient of the particles [84]. Nevertheless, Langevin dynamics is still the cheapest method for modelling the thermal motion of dilute molecular systems.

2.3.4 OxDNA Model and LAMMPS

OxDNA is a coarse-grained DNA model developed by T. E. Ouldridge, J. P. K. Doye and A. A. Louis in 2011 [83] and further improved by P. Šulc in 2012 [21] (oxDNA1.5) and B. E. K. Snodin et al. in 2015 [85] (currently updated to oxDNA2.5). The most recent oxDNA model was embedded in LAMMPS by O. Henrich et al. in 2018 [86]. The OxDNA model was originally designed to capture the thermodynamic, structural, and mechanical properties of single- and double-stranded DNA in the most common B-form. The principle idea is the modelling of DNA-backbone as a string of rigid spheres (representing nucleotides) with flexible spring-bonds (represent sugar-phosphate backbones), which interact through internucleotide potentials depending on their positions and orientations. The inter-nucleotide interactions include the sugar-phosphate backbone connectivity ($V_{b.b}$), excluded volume (V_{exc}), hydrogen bonding (V_{HB}), nearest-neighbour stacking (V_{st}), cross-stacking between base-pair steps in duplex ($V_{cr,st}$) and coaxial stacking ($V_{cx,st}$). The overall internucleotide potential in the oxDNA model can be written as:

$$V_{oxDNA} = \sum_{ij} (V_{b.b} + V_{st} + V_{exc'}) + \sum_{i,j \neq ij} (V_{HB} + V_{cx,st} + V_{cr,st} + V_{exc}), \quad (2.10)$$

where the first sum of the interaction is obtained from all nucleotides that are nearest neighbours (with nucleotides i and j are located on the same strand), and the second sum includes all other remaining pairs, including interactions from the other stand. These interactions explicitly depend on the relative orientations of the nucleotides and the distance between their interaction sites [21]. Each of the interaction potentials is dominated by a corresponding constructed function. (1) $V_{FENE} \rightarrow V_{b.b}$, (2) $V_{LJ} \rightarrow V_{exc}, V_{exc'}$, (3) $V_{Morse} \rightarrow V_{HB}$

, (4) $V_{Harmonic} \rightarrow V_{st}, V_{cx.st}, V_{cr.st}$, (5) $V_{Modulation}, V_{Smooth} \rightarrow V_{oxDNA}$. These constructed functions are defined as:

1. $V_{FENE} \rightarrow V_{b.b}$: the backbone connectivity ($V_{b.b}$) is modelled with finitely extensible non-linear elastic (FENE) springs acting between the backbone interaction sites.

$$V_{FENE} = -\frac{\epsilon}{2} \ln\left(1 - \frac{(r - r_0)^2}{\delta^2}\right) \quad (2.11)$$

with ϵ is the energy unit, r_0 is the initial distance and δ is the maximum extent of the bond. The FENE potential maintains the backbone connected and unbreakable.

2. $V_{LJ} \rightarrow V_{exc}, V_{exc'}$: The excluded volume interactions are modelled with a truncated and smoothed Lennard-Jones potential. The reason for truncation is explained in point (5) hereunder.

$$V_{LJ} = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right), \quad (2.12)$$

where σ is a length unit. The LJ potential is a well-studied potential for modelling a rather soft repulsion between a pair of neutral particles [87]. Here the LJ potentials are used to represent repulsions between backbone-backbone sites, base-base sites and backbone-base sites.

3. $V_{Morse} \rightarrow V_{HB}$: The hydrogen-bond interaction consists of smoothed, truncated and modulated Morse potentials between the hydrogen bonding site.

$$V_{Morse} = \epsilon(1 - e^{-(r - r_0)^2}) \quad (2.13)$$

4. $V_{Harmonic} \rightarrow V_{st}, V_{cx.st}, V_{cr.st}$: The stacking interaction falls into three individual sub-interactions: the stacking interaction between consecutive nucleotides on the same strand (V_{st}) and cross-stacking ($V_{cr.st}$) and coaxial stacking ($V_{cx.st}$) between any nucleotide in the appropriate relative position. The stacking interactions are modelled with smoothed, truncated and modulated Morse, harmonic angle and harmonic distance potentials.

$$V_{Harmonic} = \frac{k}{2}(r - r_0)^2 \quad (2.14)$$

5. $V_{Modulation}, V_{Smooth} \rightarrow (2), (3), (4)$: Quadratic angle terms ($V_{Modulation}$) and quadratic smoothing terms (V_{Smooth}) are modulated in all (2), (3), (4) above. The angle modulation ($V_{Modulation}$) is used to provide potential respect to the orientation of nucleotides

and backbones. And the smooth functions (V_{Smooth}) are used to make the truncated function differentiable.

$$V_{Modulation} = 1 - a(\theta - \theta_0)^2 \quad (2.15)$$

with a being a modulation factor representing the strength of angle modulation. The angle modulation potential is parametrised to meet with real experiments and permits an accurate description of both ssDNA and dsDNA orientational behaviours (see details in [83]).

$$V_{Smooth} = b(x_c - x)^2 \quad (2.16)$$

with b is a constant depending on the truncation position of smoothed functions, and x_c is the cut-off distance. The potentials beyond the cut-off distance will be neglected to be zero. The importance of using this smoothing term is discussed hereunder: Setting up a cut-off point at a finite potential distance can simplify the computation in periodic boundary conditions (e.g. the long-distance attraction in LJ potential is truly neglectable as it is extremely close to zero). This cut-off method tremendously reduces the computation time in a simulation. But the problem is, once the tail part of the function is neglected, there is a truncation on the cut-off position to make the entire potential unable to be differentiable. To solve this truncation problem, a smoothing term has been added to make the overall potential go continuously to zero: a differentiable potential is a prerequisite in any molecular simulation with a method such as Langevin dynamics.

It is worth emphasising that the final duplex structure of DNA acquired from this model is naturally through this choice of interactions (1)-(5) and their parameterisation, without any other specified or imposed forces. All the above interactions have been parameterised to match vital thermodynamic properties of ssDNA and dsDNA comparable to actual experiments (e.g. the persistence length, the melting temperature of the dsDNA) [86]. The structural properties of DNA, such as the persistence length, torsional rigidity and radius of gyration, play an essential role in its function. These properties can be characterised depending on different conditions, and they play an indispensable role in more complex processes such as DNA self-assembly, DNA replication and DNA thermodynamics. However, making these measurements in an experiment is not easy as it requires subtle manipulation of single molecules, and directly measuring the behaviour of a single DNA molecule under applied forces is even more challenging. A coarse-grained computational model like oxDNA

can help us study these properties in more detail. The simulation outcomes can provide extra prospects beyond experimental observation. And sometimes, it can become a handy tool to prove theoretical predictions. A short schematic overview of the oxDNA model is given in Fig. 2.14: it provides a detailed interaction schematic in (a) and different representations of ssDNA, dsDNA in (b), (c) and (d).

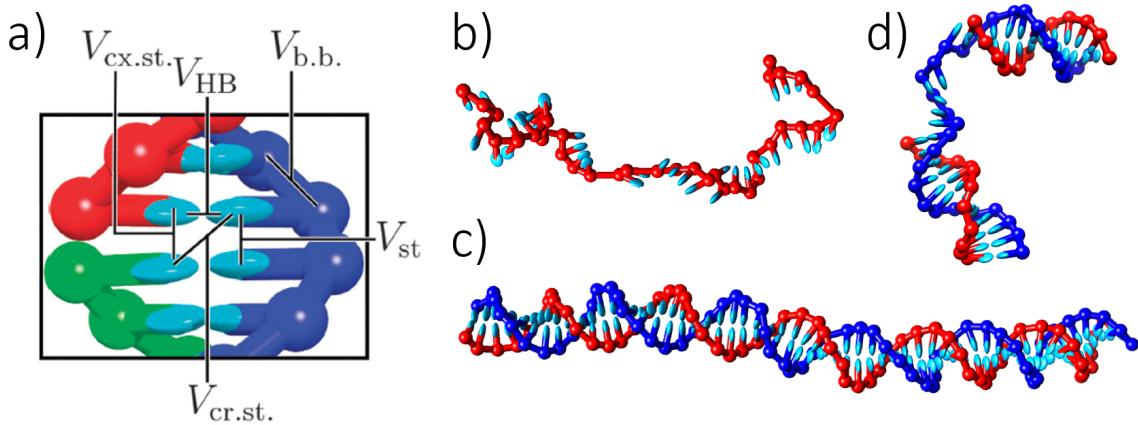


Fig. 2.14 Schematic representation of the oxDNA model: a) interactions between nucleotides [88]; b) single-strand DNA (ssDNA); c) double-strand DNA (dsDNA) d) nunchuck like DNA with two dsDNA ends connected via a ssDNA chain in the middle.

The first version of oxDNA model treated all A, T, G, and C as the same type. Therefore the produced structure was symmetric and insufficient to deal with a specific DNA sequencing order. In 2012, P. Šulc et al. [21] introduced a sequence-dependent parametrisation for the first version by introducing new sequence-dependent stacking (V_{st}^* , $V_{cx,st}^*$) and base-pairing interaction strengths (V_{HB}^*) chosen to reproduce the melting temperatures of short duplexes. And later, in 2015, B. E. K. Snodin et al. [85] extended the model with effective electrostatic interactions by introducing an extra Debye-Hückel potential (V_{DH}^*). This helped the model to be able to investigate some salt-dependent properties. V_{DH}^* was also parameterised to reproduce the salt-concentration dependence of the thermodynamics of DNA hybridisation rather than any specific properties related to internucleotide potentials. V_{DH}^* is mediated by Coulomb's law for electrostatic interactions between ions: interactions between two electric charges (q_i, q_j), separated by a distance, r_{ij} in a medium of relative permittivity ϵ_r and is given by

$$V_{DH}^* = \sum_{ij} \frac{q_i q_j e^2}{4\pi \epsilon_0 \epsilon_r} \frac{e^{-r_{ij}/\lambda(T,I)}}{r_{ij}}, \quad (2.17)$$

where ϵ_0 is the permittivity of the vacuum, and $\lambda(T, I)$ is the Debye screening length which is dependent on the temperature (T) and salt concentration (I) (See details of description in [85]). Finally, the newest version in oxDNA with overall interactions is updated as below:

$$V_{oxDNA} = \sum_{ij} (V_{b.b}^* + V_{st}^* + V_{exc'}) + \sum_{i,j \neq ij} (V_{HB}^* + V_{cx.st}^* + V_{cr,st} + V_{exc} + V_{DH}^*), \quad (2.18)$$

where the * means the potential has been newly updated compared to the first version, and V_{DH}^* is the extra salt-dependent interaction. By slightly modifying the coaxial stacking ($V_{cx.st}^*$) and backbone interactions ($V_{b.b}^*$), the model can now successfully produce explicit major and minor grooves in the dsDNA duplex. Another critical update in oxDNA is the extended parametrisation of the strength of consecutive adenine bases (AA stacking) and Thymine bases (TT stacking) on the single-strand DNA (ssDNA). This allows more accurate treatment of systems in which the flexibility of ssDNA is essential (such as DNA cyclisation [89], DNA walkers [34], the flexible joints in DNA origami [90] and the flexible ssDNA chain in our nunchuck DNA, with latter shown in Fig. 2.14.d).

The oxDNA model was previously only available as bespoke and standalone software. Last year, it was implemented into the popular LAMMPS molecular dynamics code [86]. The primary purpose of the implementation is to reduce the entry barrier of using the oxDNA model. One of the most immediate benefits is that the embedded oxDNA model now can directly use the existing ‘commands compute’ packages in LAMMPS such as mean-squared displacement (MSD), radial distribution function ($g(r)$), Radius of gyration (R_g), etc. And what is even more important is that it allows combining force fields in the oxDNA model with different features already enabled in LAMMPS. In this thesis, the oxDNA model was used to investigate two main tasks:

1. The self-looping behaviour of a key-shape DNA building block and its role in linker-mediated DNA hydrogels (model applied in Chapter 3).
2. The customised bending free energy of nunchuck-DNA mesogens as a prerequisite study for higher level coarse-grained simulation towards new liquid crystal phases (model applied in Chapter 4).

In the high-level coarse-grained system, the key parameters for the building blocks (i.e., angles, bonds, geometry, etc.) can be optimised by matching with lower-level coarse-grain simulations such as oxDNA model, providing quantitative predictions for actual materials.

3 | The Role of Flexible Joints in a Linker-Mediated DNA Hydrogel

In this chapter, I report my numerical work using oxDNA to test the thermodynamic stability and topography of Y-shaped DNA nanostars connected via rigid, linear DNA-linkers. The specific DNA sequences we used were those we tested in experiments to study the formation of DNA hydrogels as function of temperature and DNA concentration. This work was done in collaboration with Dr. Iliya Stoev, who did most of the experiments in this project. Our dynamic light scattering (DLS) measurements showed that when inserting sufficiently long flexible ssDNA joints between the linear linker and the Y-shaped DNA we suppress gelation. This motivated me to test the effect of the length of the flexible joints.

Specifically, I investigated the probability of a linear linker to connect to the arms of two separate Y-DNA nanostars in relation to the probability of connecting to two arms of the same Y-shape. In the experiments we used a flexible dsDNA-linker with a free joint of 6 Thymines on both ends (L6) of the hybridised dsDNA segment and compared the mechanics of the resulting network to the rigid-linker case, where no flexible T-joint was present (L0). I demonstrated the process of self-looping, referring to the formation of key-shaped nanostructures, for the Y+L6 system and further checked its stability and melting temperature. I also computed the free energy cost for the formation of a key-shaped structure in the Y+L6 case and compared it to the case of single (straight) linker-nanostar binding. My simulation results proved crucial for understanding and verifying our experimental findings: these showed that the Y+L0 solutions display a liquid-to-gel transition upon cooling below the melting temperature of the sticky ends linking nanostars together, while the Y+L6 suspension displays an initially liquid state with higher viscosity above the said melting temperature and a sudden reduction in viscosity at lower temperatures. A next-level coarse-grained model, based on simulation work by Dr. Christopher Ness, was then created to investigate the rheological response of the system. The completed work was published in the *Soft Matter* journal [47] issued by the Royal Society of Chemistry.

3.1 Introduction

Designing viscoelastic, thermally responsive hydrogels has previously included studying three-dimensional DNA networks made of trivalent (named ‘Y-shaped DNA’) [45] or higher-valent nanostars with sticky, single-stranded DNA overhangs. In this chapter, we introduce linker-mediated DNA hydrogels, in which the Y-shapes are designed such that their sticky ssDNA overhangs connect to each other via complementary ssDNA-ends attached to rigid, linear DNA duplexes (Y+L). We can design this linker-mediated connection to be either fully rigid (L0) or flexible (L6) by introducing non-binding T-bases located between the complementary ssDNA and the rigid hybridised part of the linear linker-DNA: in experiments, we varied the number of Thymine bases between 0, 2, 4 and 6. The extra flexibility offered by the T-joints affected the system’s ability to form elastic, percolating gels by introducing a competitive mechanism for linking our DNA building blocks.

The pioneering work by Seeman on ‘Nucleic Acid Junctions and Lattices’ [30] initiated research on hierarchical self-assembly into finite-size origamis and space-filling hydrogels made of DNA [91–99]. The successful design of DNA nanostars and origamis is facilitated by the high binding specificity between short, complementary DNA molecules and the ease of their production and variety in base-pair (bp) sequences via PCR methods. Combined with the thermoreversible binding of complementary DNA strands, DNA can be designed and programmed to overcome possible kinetic barriers in the self-assembly of nano-sized building blocks that in turn can bind to each other via free bases acting as sticky overhangs. In this context, DNA hydrogels have been developed as alternatives for soft, tissue-friendly materials leveraging their biocompatibility with thermal [100] and/or enzymatic responsiveness [39]. Other promising applications of DNA hydrogels include molecular imprinting [101] and embedding pharmaceutical agents in controlled drug release applications [38, 102].

Recently, more interest was devoted to understanding the relationship between the macroscopic mechanical behaviour of hydrogels made of DNA-nanostars and the type of molecular network they form, which is important when designing new functional materials targeting specific applications [38, 103, 41]. The group of Sciortino showed that DNA-nanostars behave like atoms of a given valency: for instance, Y-shaped DNA is trivalent as it can form three bonds. The liquid-to-gel phase transition region becomes wider with increased valency and thereby opens up the liquid or equilibrium-gel region before reaching the solid transition where the system percolates [104, 105, 41]. The liquid one-phase region is wide and above a critical temperature the system undergoes a steep transition into a gel without any phase separation, where the equilibrium-gel region can be viewed as three-dimensional porous viscoelastic network [106, 107]. In the limiting case of having only bivalent patchy

particles, i.e. linear double strands of DNA, the coexistence region disappears as only single polymer chains form and no three-dimensional network is ever observed [41, 45].

While it has become relatively inexpensive to order any synthetic DNA strand with up to about 120 bases, detailed systematic phase-diagram investigation and rheological measurements of DNA suspensions, where large experimental datasets are required, remain very expensive and time-consuming. Therefore, simulations have become an excellent tool to investigate the properties of DNA hydrogels in a large parameter space. Many numerical models have been developed to help us understand the microscopic structural details of nano-sized networks based on nucleotide-level [108–113] and coarse-grained simulations [114, 115, 83, 46]. Alas, none of the literature studies reported so far using these models were sufficient to explain the findings in our linker-mediated DNA-hydrogel system.

Therefore, we first performed oxDNA simulations to obtain the correct thermodynamic stability and melting temperature relevant to the binding between a single Y-shape nanostar and a linear linker consisting of non-binding T-joint lengths of 0-6 Thymines. In comparison with our experimental results [47], we find that for T-joint lengths ≥ 6 the Y-DNA+L system forms a cluster phase with a viscosity close to that of water, while for shorter flexible joints we observe percolating networks, as I will describe in the following section. Subsequently, we developed a next-level coarse-grained model based on the configurational and binding properties found with oxDNA in order to validate the existence of the predicted cluster-fluid.

3.2 Experimental Findings and Hypothesis

The DNA strands used in our experiments were provided by the group of Prof. Dongsheng Liu, our collaborators at Tsinghua University in Beijing. They produced all DNA strands via their own synthesiser. In Table 3.1, details of all nucleotide sequences are given. In this table, we also compare the melting temperatures that were computed using the next-neighbour model of SantaLucia [116] for the individual arms of the Y-shapes and the linear linker with our experimental data, which were obtained from UV-visible spectroscopy measurements [45, 47]. Note that the measured melting temperature $T_{m,Y}$ of the Y-shapes is the averaged T_m of the individual arms. We also list the SantaLucia model-based melting temperature for the complementary, sticky overhangs between the Y- and L-shapes. All data were obtained for the DNA and overall added salt (NaCl) concentrations used in the experiments. These data were in close agreement with the melting temperatures extracted from my oxDNA calculations. In Fig. 3.1 we present snapshots of the thermally equilibrated DNA structures obtained from oxDNA simulations using those sequences. A graphics package including

Ovito [117] and Chimera [118] was employed to visualise the configurations of the Y-shapes and linear linkers.

In experiments, we used a 2:3 ratio of Y-shapes to linear linkers, providing maximal connectivity below the melting temperature of their complementary sticky ends [35]. Fig. 3.2 summarises the electric-field autocorrelation functions ($g^{(1)}$) extracted from Dynamic Light Scattering data. These allow us to also obtain the microrheological response (elastic, $G'(\omega)$ and viscous, $G''(\omega)$ moduli) for different DNA concentrations of the rigid Y+L0 and flexible Y+L6 DNA systems. The $g^{(1)}$ curves are shown for 0.8 wt% (280 μM) and 1.6 wt% (560 μM) overall DNA concentrations, whereas the $G'(\omega)$ and $G''(\omega)$ curves are only displayed for the higher concentration. We observed gelation when the Y+L0 solutions were cooled down below the system's T_m ($G'(\omega)$ exceeds $G''(\omega)$ below T_m and for high enough DNA concentrations). In contrast, the Y+L6 system does not gel upon cooling ($G'(\omega)$ remains lower than $G''(\omega)$ at all temperatures). Instead we observe a dramatic decrease in the overall viscosity that is indicated by the lower decorrelation time in $g^{(1)}$. In fact, we observed the

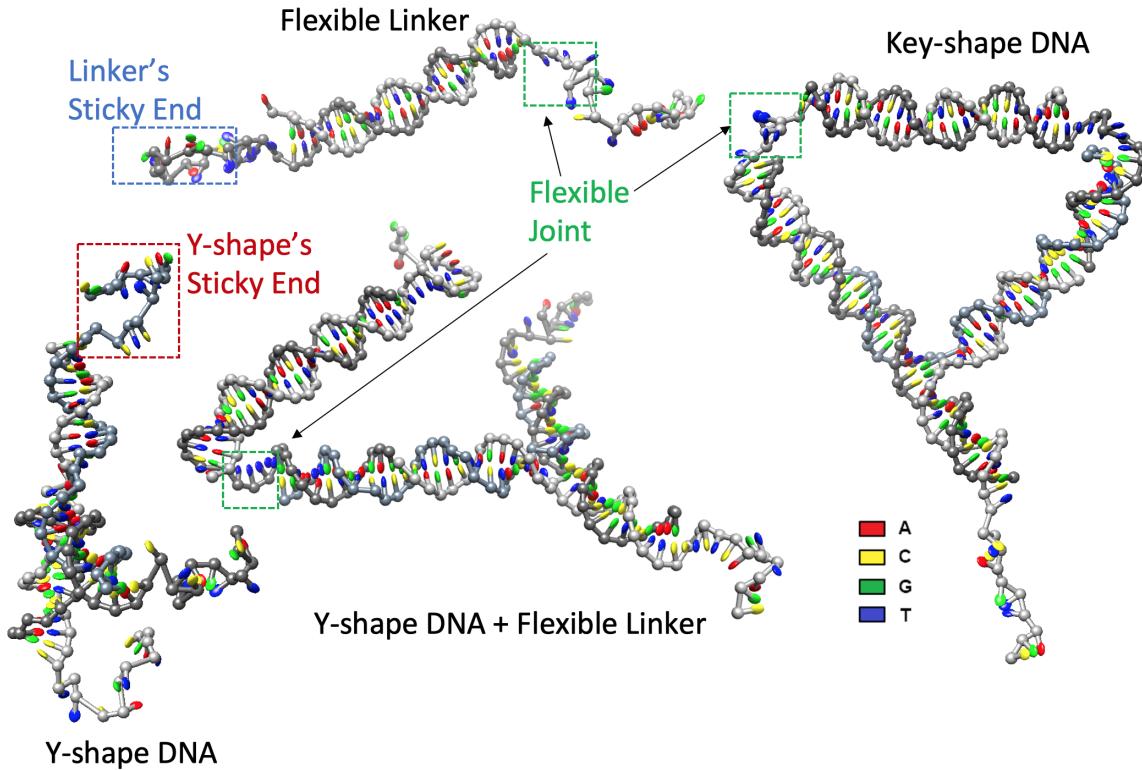


Fig. 3.1 Schematic of Y-shaped DNA, flexible linker-DNA (L6) and key-shaped DNA made of Y-L6 pair. The entire sequences of the design can be found in **Appendix A.1**, with detailed colour representation for individual A, C, G, T nucleotides, which is designed to match with our experiments [47].

Table 3.1 ssDNA sequences and relevant melting temperatures of the Y- and L-shaped DNA structures. The different segments are colour-coded according to their function or complementarity. Adapted from ref. [47]

ssDNA	sequence	T_m /°C SantaLucia ^a	T_m /°C UV-vis measurements ^a	T_m /°C oxDNA modelling ^b	T_m /°C high DNA concentrat ion ^c
Y1	5'-CGA TTG ACT CTC CAC GCT GTC CTA ACC ATG ACC GTC GAA G -3'	51.11	$T_{m,Y} \approx 55^\circ\text{C}$	$T_{m,Y} \approx 59-$ 65°C	$T_{m,Y} \approx 62-$ 65°C
Y2	5'-CGA TTG ACT CTC CTT CGA CGG TCA TGT ACT AGA TCA GAG G -3'	48.44			
Y3	5'-CGA TTG ACT CTC CCT CTG ATC TAG TAG TTA GGA CAG CGT G -3'	41.84			
L1-0(6)	5'-GAG AGT CAA TCG (TTT TTT) TCT ATT CGC ATG ACA TTC ACC GTA AG-3'	62.97	$T_{m,L0(6)} \approx 65^\circ\text{C}$	$T_{m,L0(6)} \approx 65-$ 71°C	$T_{m,L0(6)} \approx$ $72-75^\circ\text{C}$
L2-0(6)	5'-GAG AGT CAA TCG (TTT TTT) CTT ACG GTG AAT GTC ATG CGA ATA GA-3'	62.97			
Y-L0(6)	5'-CGA TTG ACT CTC 5'-GAG AGT CAA TCG	40.54	-	$T_{m,YL} \approx 52-$ 56°C	$T_{m,YL} \approx 52$ $- 54^\circ\text{C}$

^a SantaLucia and measured values obtained assuming buffer solution containing in total 1 μM DNA and 100 mM added NaCl.

^b Melting temperatures from oxDNA simulations using approximately 5 μM DNA and 200 mM added NaCl.

^c Estimated values for 175 μM - 840 μM total DNA in buffer solution containing 100 mM added NaCl.

complete suppression of gelation, suggesting that the linear linkers can directly bind with both ends to a single Y-shape, thus reducing the number of bonds available for binding Y-shapes to linkers. As a result we obtain a cluster phase, whose bulk viscosity is close to that of water. This scenario is sketched in Fig. 3.3.

To summarise, our hypothesis has the following consequences: for the Y-L0 system with a 2:3 ratio, the overall system forms a percolating gel with relatively high viscosity and a measurable elasticity; however, in the flexible-linker case, there is a finite number of self-looped nanostructures (here named ‘key-shaped’ DNA). This key-shaped DNA would reduce the effective valency, thus preventing the continuous binding of all Y shapes to each other into a percolating network (shown in the upper right-hand side corner of Fig. 3.3). In

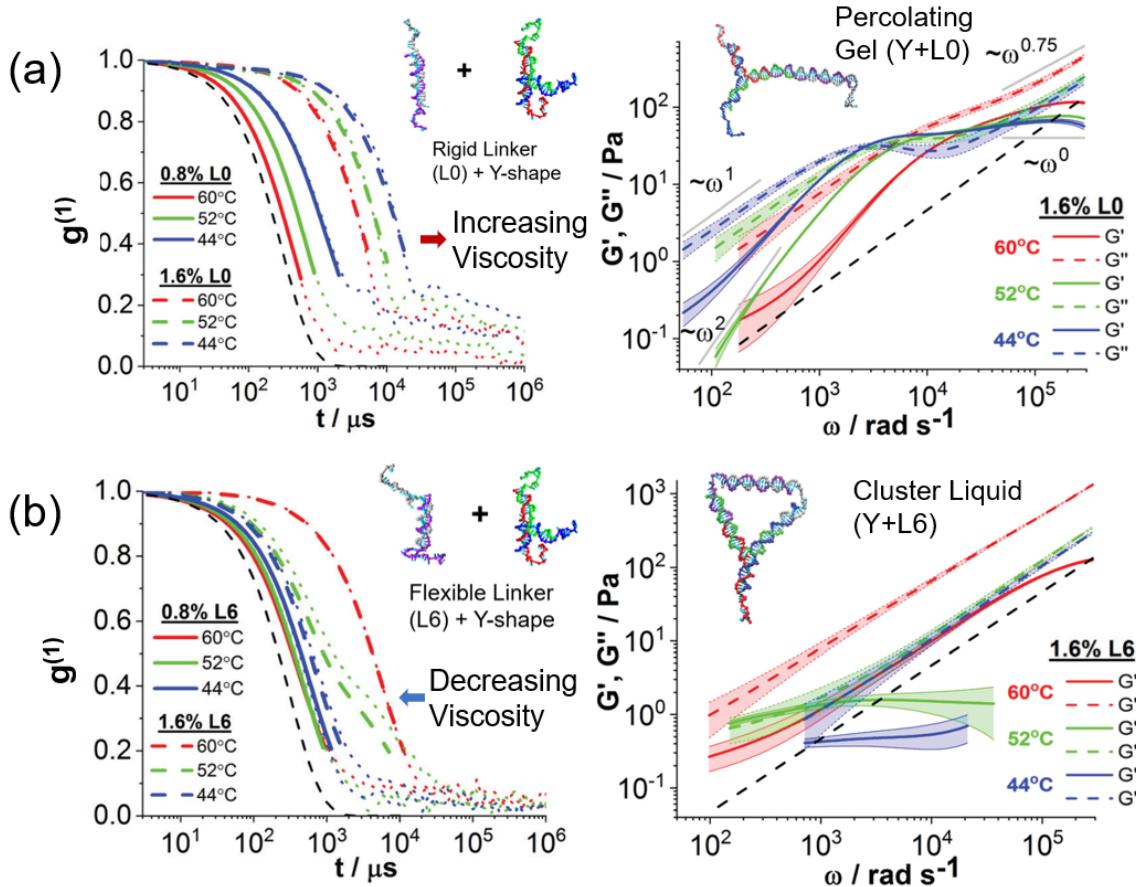


Fig. 3.2 Electric-field autocorrelation functions ($g^{(1)}$) obtained from DLS measurements and plotted as solid (0.8 wt%) or dashed (1.6 wt%) lines, and the corresponding storage and loss moduli for the 1.6 wt% samples made of the Y-shapes and linkers using (a) rigidly attached L0, and (b) L6 linkers with a flexible 6T-joint. The $g^{(1)}$ curves were used to convert into the MSDs that correspond to the passive probe-particle diffusion within the sample, and then converted into the storage and loss moduli of the material via the Generalised Stokes-Einstein relation [119]. Dotted lines denote the raw $g^{(1)}$ data as obtained from our dynamic light scattering setup. The dashed black line is the calculated relaxation curve for the same probe-particles in pure water at 60 °C. All measurements were done by starting at 60 °C and equilibrating for 20 min at every measured temperature. [47]

this case, we end up with a distribution of finite-sized clusters with no additional binding sites between them. The resulting cluster fluid has a considerably lower viscosity and no elasticity. To prove this hypothesis, we used oxDNA simulations to probe the detailed thermodynamics related to the probability of creating this self-looping scenario.

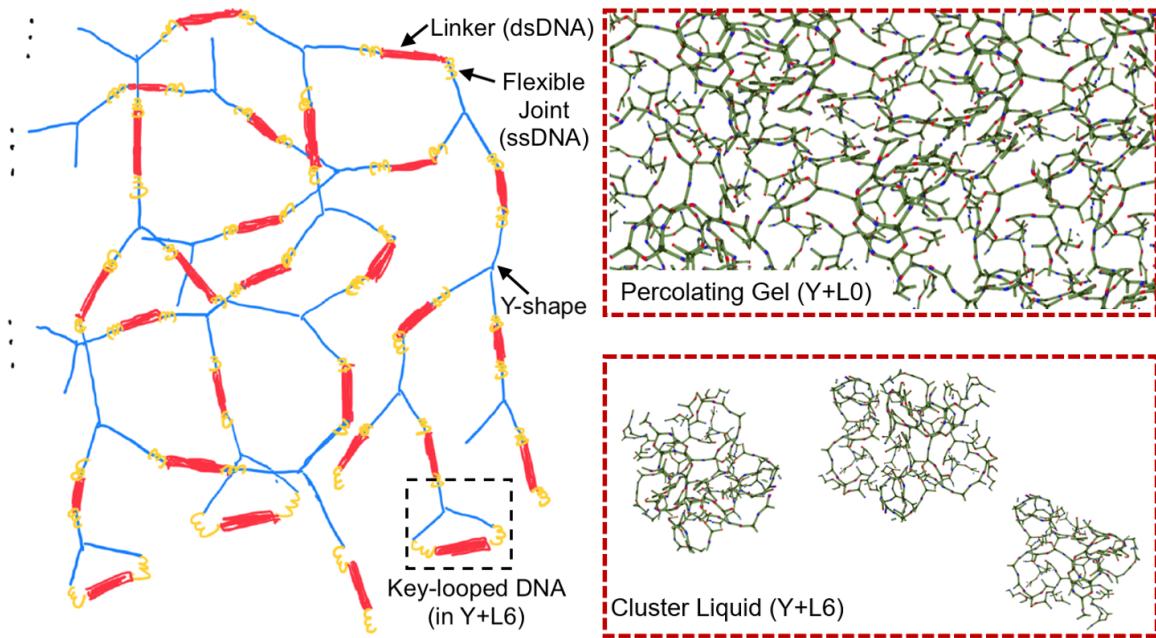


Fig. 3.3 Hypothesised structure: percolation is terminated by the looping behaviour in the Y-L6 system, while in the Y-L0 system no self-looping can occur, as this would induce a large bending and overlap penalty in the total free energy (the L0 is red, 6Ts in yellow, and the Y shapes are blue).

3.3 Results and Discussion

3.3.1 Comparison between flexible and rigid linker hydrogels

All DNA sequences in the oxDNA simulations shown here were the same as those used in our experiments. We propose two binding mechanisms depending on whether the linker is rigid or has some flexibility between the sticky end and the rigid part of the linker. The primary purpose of designing this simulation was to identify quantitatively the range of conditions where the flexible linker can create self-looped structures in the presence of random thermal (Brownian) motion and contrast that with the rigid-linker case. We initially simulated the Y-L6 and Y-L0 pairs with the linear linker attached to one arm of the Y-shape only in a low-temperature environment (below T_m). The simulation was allowed to continuously run until thermal equilibrium was reached. Subsequently, we measured the distribution of accessible angles (marked as θ_{L0}/θ_{L6} in Fig. 3.4) for both cases to gain knowledge about their average spatial configuration and flexibility.

The histogram in Fig. 3.4 presents all sampled angle-probabilities reflecting flexibility in the Y-L6 and Y-L0 systems. The area within the red cone describes the space the L6-linker

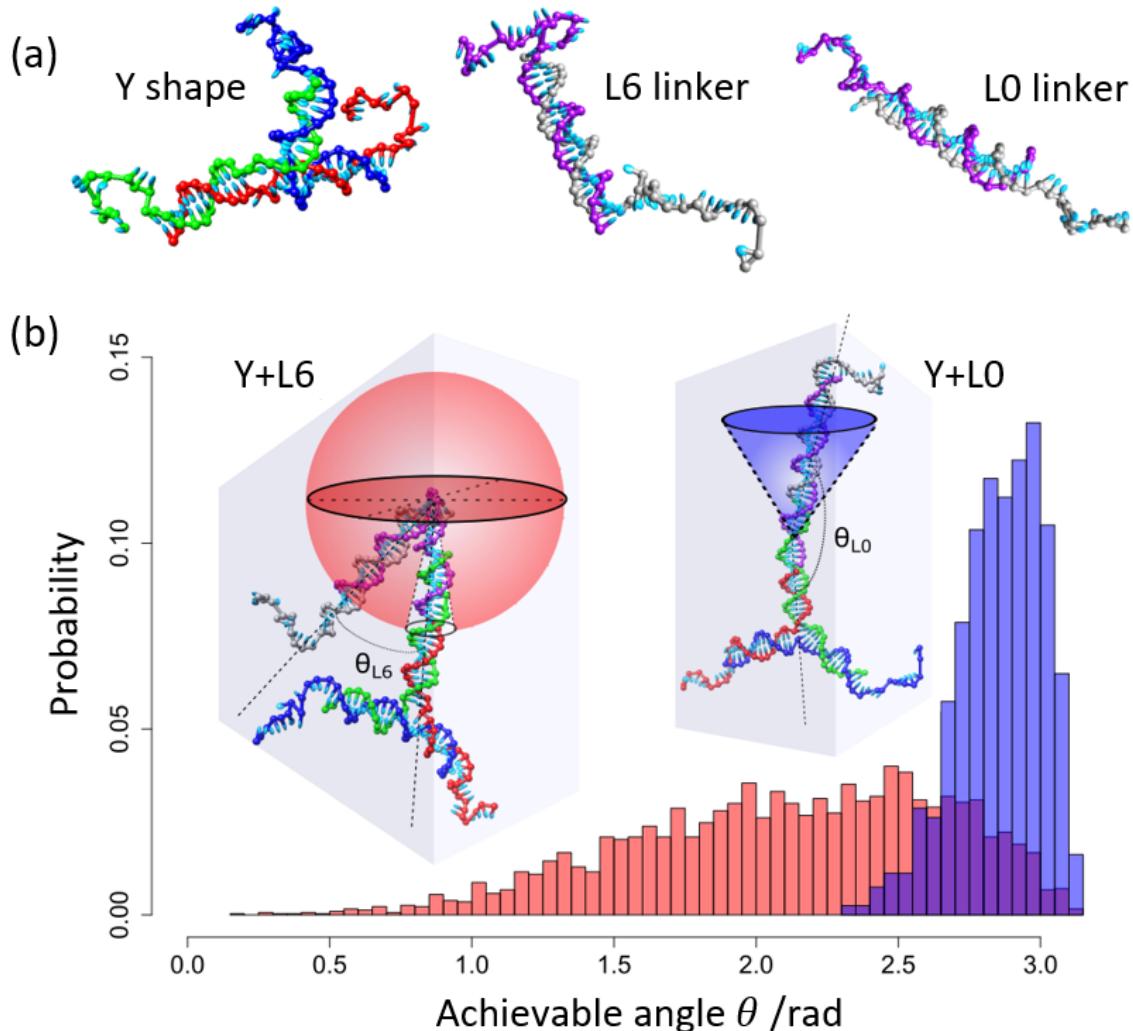


Fig. 3.4 (a) oxDNA representations of the Y-shape, rigid (L0 linker) and flexible (L6 linker) DNA. (b) The Distribution of the probability that the joint between a Y-shape and linker adopts a certain angle θ_{L0}/θ_{L6} . The measured (sampled) ranges of angles of the Y-L6 (red columns) and Y-L0 (blue columns) are plotted in the histogram and indicated as sphere and cone, respectively in the cartoon of the shapes. All measurements were performed at simulated room temperature. Two simulation sets (1)Y+L6 (2)Y+L0. 3k frames on each set are used to plot the histogram. [47]

can explore when one of its sticky ends is connected to the Y-shape DNA. This gives the second sticky end a high probability of binding to a second arm of the same Y-shape, thus forming a key-shaped structure. The blue cone describes the available region of the rigid L0-linker. The weak flexibility of the Y-L0 bond may be caused by the lack of stacking force on the notch and because the phosphate-sugar backbone is not continuous in the binding region. The significant difference in the cone regions is clear because we had the non-binding

Thymine sequence which has a persistence length of about 1 nm, corresponding to about three bases. Hence, the question arises of how long this flexible joint needs to be to allow a linear linker to bind to two arms of the same Y-shape and thus reduce the effective valency of the system. The oxDNA simulations employing 2 Ts or 4 Ts did not form the key-shaped DNA nanostructures. These findings were in agreement with our experiments, where L2 and L4-linkers showed gelation at low temperatures [47]. More details will be discussed in Chapter 4.

3.3.2 Structural stability of the key-shaped DNA

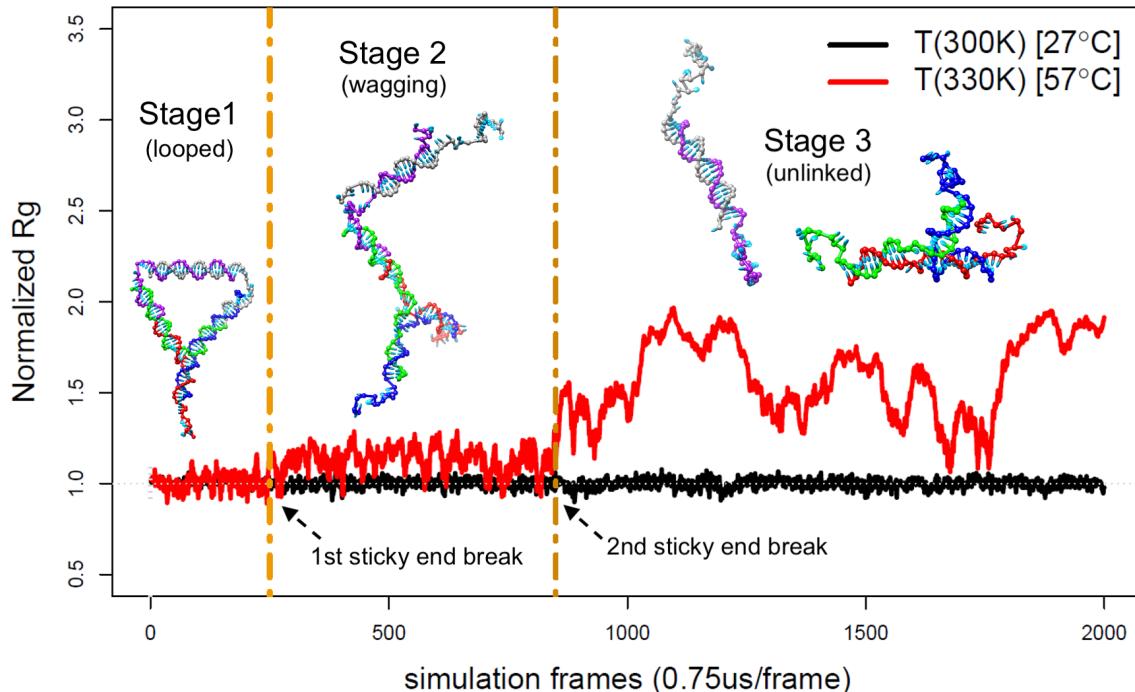


Fig. 3.5 The radius of gyration (R_g) as a function of simulation steps for two different temperatures and at an added salt (NaCl) concentration of 200 mM. Two simulation sets (1) $T=300\text{K}$ (black) (2) $T=330\text{K}$ (red). 2k frames are used in each set.

A successful self-looped DNA nanostructure was obtained in our simulation of the Y-L6 pair, which we named ‘key-shaped’ or ‘key-looped’ DNA. Two representative videos can be viewed in the SI of our publication in *Soft Matter* [47]. Careful inspection of these videos reveals the effect that flexibility has on the range of accessible angles achieved in the assembly. The results from the analysis of these videos are shown in Fig. 3.4. The closing of the ring and the loop formation in the Y-L6 system is illustrated clearly. The visualisation

method of building this schematic is detailed in **Appendix B.1**. The oxDNA-scripts used to model this self-looping process are available as open source in **Appendix C.1**.

Once a key-shaped Y-L6 DNA construct was formed, we checked its stability under various conditions (e.g. low/high temperature and salt concentration). Fortunately, the oxDNA model allows us to mimic an accurate description of the actual temperature scale and added salt concentration [21, 89]. To numerically find the stability of the key-shaped DNA, we introduced the radius of gyration (R_g) to describe the general configuration of this compact DNA assembly. R_g is a descriptor of the structure and compactness of the macromolecule and its size. The values of R_g are calculated for different temperatures by the averaged sum of individual mass multiplied by the squared distance between the individual atoms (here represented as single nucleotides) to the centre of mass:

$$R_g^2 = \frac{1}{M} \sum m_i (r_i - r_{cm})^2. \quad (3.1)$$

Here M is the total mass of the DNA assembly, r_{cm} is the centre-of-mass position of the assembly, and r_i indicates an individual nucleotide position.

In Fig. 3.5 we compare an equilibrated, looped Y-L6 pair kept at 27 °C (black curve) with its corresponding time-evolved structure when brought to 57 °C (red curve), which is above T_m . We always start with a preformed key-looped DNA to save computation time that would otherwise be required to allow it to self-assemble from the individual separated components. The R_g value of the key-shaped DNA stabilises once it is run at $T < T_m$, indicating that it does not undergo any significant structural changes at room temperature. In the high-temperature case, the behaviour is entirely different. Here we take 57 °C as an example: At the beginning of the simulation (*Stage 1*), the key-shaped DNA remains in a looped configuration, and no sticky ends break. After a while, the first sticky end starts to break, but the other sticky end is still connected to the Y-shape. This will give us the wagging stage (*Stage 2*), in which the R_g value is still relatively stable. Finally, the second sticky end breaks, and the Y-shape and linker DNA are separated (*Stage 3*). Once both sticky ends disconnect, the R_g value begins to diverge.

The simulation results prove our experimental observations that the self-looping behaviour in the Y-L6 system can lead to the appearance of the connected DNA network at low temperatures. In addition, the arms of the Y-shape and the dsDNA linker have almost the same length, enabling the formation of this self-looped nanostructure geometrically. From the statistical point of view, the occurrence of this key-shaped DNA is frequent enough to affect the results in our experiments and cause deviations from the expected linking behaviour, where every two Y-shapes bind on average three dsDNA linkers. Therefore, we conclude

that self-looping plays an essential role in the network-forming ability of our linker-mediated DNA-nanostar assembly.

3.3.3 Closed-to-open ratio of the Y+L6 system

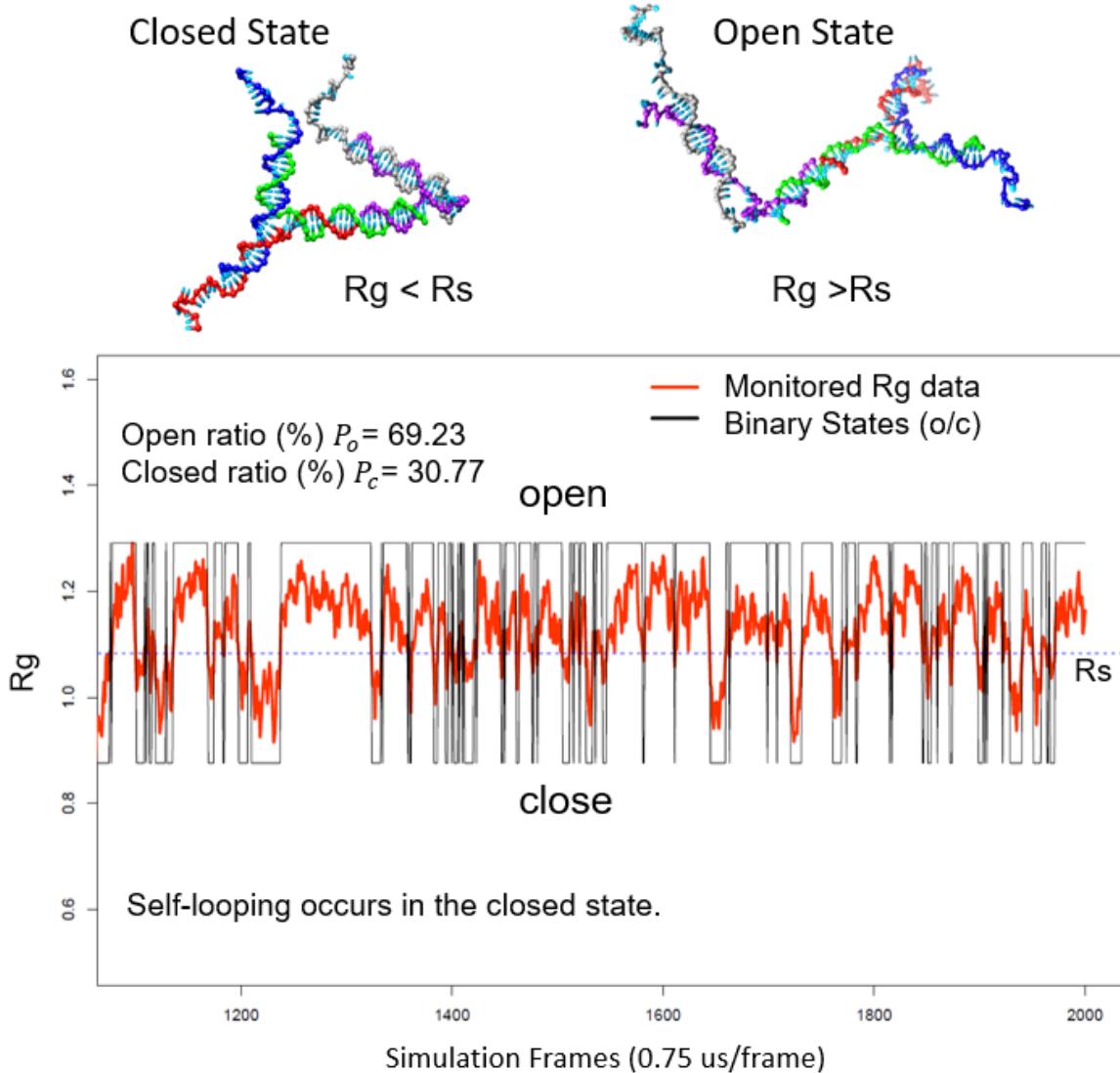


Fig. 3.6 Variation of R_g in time reflecting the closed-to-open ratio of the Y+L6 system at 56 °C and 200 mM added NaCl. P_s is the threshold distinguishing between open and closed (o/c) state. In the scenario shown here, $P_{c/o} = 30.77/69.23 \approx 0.44$. 1k frames are used to calculate $P_{c/o}$.

In addition, we investigated changes in the closed-to-open ratio ($P_{c/o}$) of the Y+L6 system. $P_{c/o}$ is a useful parameter to demonstrate how likely the self-looping behaviour is to occur,

especially in the wagging stage of the system (when linker DNA can more readily sample the closed configuration). We define:

$$P_{c/o} = \frac{P_c}{P_o}, \quad (3.2)$$

where P_c is the probability of a closed state and P_o is the probability of finding an open state. Snapshots of the closed and open state are shown in Fig. 3.6. The changes in P_c and P_o are monitored by setting up a threshold (R_s) that depends on R_g . If the value of $P_{c/o}$ is large, then the probability of closing is higher, i.e. the possibility of observing self-looping behaviour is higher.

$$R_s = \langle R_{g(max)} + R_{g(min)} \rangle \quad (3.3)$$

As the DNA melting temperature depends also on the added salt concentration[120], we tested here how salt concentration and temperature affect this self-looping behaviour. The results are summarised in Tab. 3.2.

Salt concentration (NaCl mM)	100	200	200	200	200	200	100	200	300
Temperature (°C)	26	26	32	38	44	50	56	56	56
Closed ratio (%)	28.31	31.66	32.57	31.98	31.76	32.53	25.48	30.77	33.54
Open ratio (%)	71.69	68.34	67.43	68.02	68.24	67.47	74.52	69.23	66.46
$P_{c/o}$	0.39	0.46	0.48	0.47	0.47	0.48	0.34	0.44	0.50

Table 3.2 Influence of added salt on the probability $P_{c/o}$ to find the Y+L6 system in the open or closed configuration at a fixed temperature. More related results are available in Appendix D.2.

From the table we conclude that temperature has little influence on the closed-to-open ratio ($P_{c/o}$). However, the salt concentration has a weak influence on $P_{c/o}$: with higher NaCl concentration, the $P_{c/o}$ is higher at a fixed temperature.

3.3.4 Simulated melting temperature of key- and Y-shaped DNA

This section provides simulation results for estimating the melting temperature of the key-looped DNA nanostructure. There are three different melting temperatures in this system: (1) $T_{m,YL}$: the temperature required to separate the sticky ends of the linker and Y-shape when connected; (2) $T_{m,Y}$: the melting temperature of the Y-shape DNA; (3) $T_{m,L}$: the melting temperature of the DNA linker. From microrheology experiments, combined with UV-vis. measurements [47] and the SantaLucia Model[116], we estimated the melting

temperatures in buffer solution containing 100 mM NaCl to be: $T_{m,YL,Exp}(\approx 52 - 54^\circ\text{C}) < T_{m,Y,Exp}(\approx 62 - 65^\circ\text{C}) < T_{m,L,Exp}(\approx 72 - 75^\circ\text{C})$.

In our detailed oxDNA simulation runs, we recorded a time-evolving R_g measurement for temperatures ranging from 27 to 75 °C. The grey line in Fig. 3.7 is set as a reference to indicate the stability of the configuration at low temperatures. When we increase the temperature (higher than 50 °C), we find that the key-looped structure is close to dissociating and yields a second plateau. In the second plateau, one sticky end is still connected to the Y-shape; hence, the linker is wagging. The second sticky end then easily detaches if we continue to increase the temperature. Step-by-step observations of the simulation are visualised in Fig. 3.7 and recorded in Tab. 3.3. At excessively high temperatures (higher than 72 °C), we observed random thermal diffusion as both linker and Y-shapes are completely melted into five individual ssDNAs. Based on that, we estimate the melting temperatures in our simulation to be: $T_{m,YL,Sim}(\approx 52 - 56^\circ\text{C}) < T_{m,Y,Sim}(\approx 59 - 65^\circ\text{C}) < T_{m,L,Sim}(\approx 65 - 71^\circ\text{C})$.

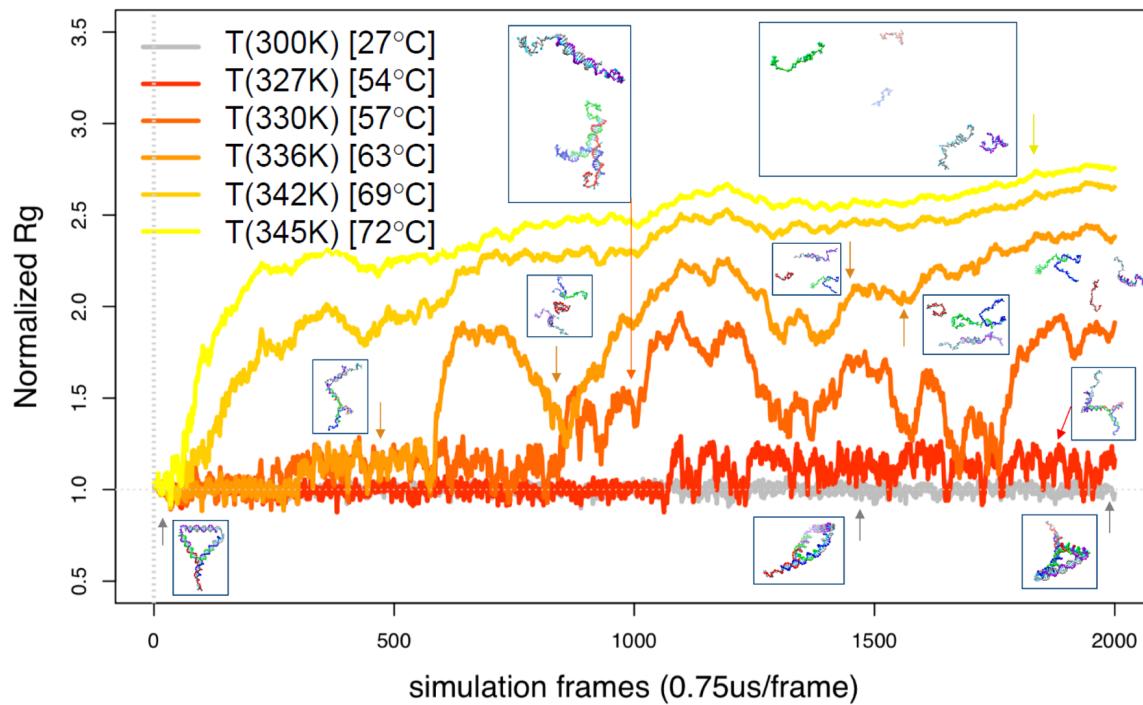


Fig. 3.7 The radius-of-gyration plots for different temperatures at 200 mM NaCl. The small snapshots (clear version with description can be found in **Appendix D.1**) embedded in the figure are example scenarios of each corresponding state in the measurements. The last measured frames were used to extract the final-state information recorded in Tab. 3.3, each set with 2k frames

Comparing our simulation results to the corresponding experiments, we conclude that the model also gives us decent estimates of the different melting temperatures. The simulated

Tem (°C)	27	36	48	51	54	57	60	63	66	69	72	75
Sticky End 1	n	n	n	y	y	y	y	y	y	y	y	y
Sticky End 2	n	n	n	n	$\sim T_{m,YL}$	y	y	y	y	y	y	y
Y-Shape Melted	n	n	n	n	n	n	$\sim T_{m,Y}$	y	y	y	y	y
Linker Melted	n	n	n	n	n	n	n	n	$\sim T_{m,L}$	y	y	

Table 3.3 Final-state information of the simulation at different temperatures (°C) - ‘n’ stands for ‘not melted’, ‘y’ represents ‘melted’. The recorded states were taken from Fig. 3.7 and the observed melting temperatures are listed in Table 3.1.

melting temperature results further enhanced the model’s accuracy in proving our initial hypothesis.

3.3.5 Estimate of the centre-core bending penalty

By first calculating the average equilibrium angles between the tree arms of a Y-shape, it is possible to calculate the bending free energy cost for binding a linker to two arms of the same Y-shape. We name these angles α_1 , α_2 and α_3 , shown as cartoons in Fig. 3.8. Also, their probability distributions in a Y-L6 system, obtained from long oxDNA simulation runs at 25 °C, are plotted in Fig. 3.8. As expected, the individual mean angles are similar in the non-linked case. The result changes when two arms are linked together using L6 linker. The penalty in the total free energy of the free and connected Y-shapes can be calculated through $F(\alpha) = -k_B T \ln(P(\alpha))$. We plotted both in Fig. 3.9, where we can read off that the bending free energy must be on the order of $0.5K_B T$. The binding free energy between the sticky arms of the Y-shape and the linear linker, however, is about $23.5K_B T$ - this can be calculated by inserting the detailed sequences into SantaLucia’s thermodynamics model¹ of DNA [116]. This result shows that the penalty of looping a Y+L6 system is relatively low compared to the binding energy, which translates as high possibility for self-looping behaviour. Also, we tested that the bending energy increases rapidly when the free joint is reduced to less Ts. When simulating the shorter joint case in oxDNA, one end of the linker arm rapidly unbinds from the Y-shape after starting in a closed position at ambient temperature.

¹SantaLucia’s model provides a thermodynamic parameter database that presents the thermodynamic nearest neighbour (NN) parameters for Watson-Crick base pairs in DNA sequences. The Gibbs free energy (ΔG°) that describes the DNA hybridisation reaction is given by [116]:

$$\Delta G^\circ(\text{total}) = \sum_i n_i \Delta G^\circ(i) + \Delta G^\circ(\text{init w/term G·C}) + \Delta G^\circ(\text{init w/term A·T}) + \Delta G^\circ(\text{sym}),$$

where $\Delta G^\circ(i)$ are the standard free-energy changes for the 10 possible Watson-Crick NNs. The second and third terms are the initiation energies at the two ends of the duplex, and the fourth term comes from the sequence symmetry.

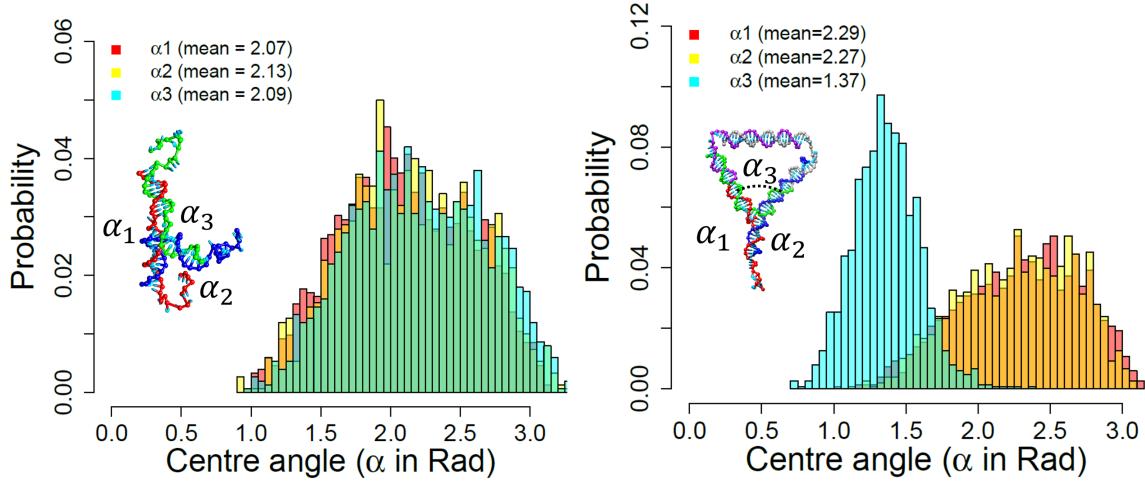


Fig. 3.8 (left) Probability distributions of the three angles α_1 , α_2 and α_3 between the three arms of a free Y-shape, obtained from oxDNA simulations. (right) Change of the angle distributions after attaching L6 linker with six-Thymine-long free joint.[47]

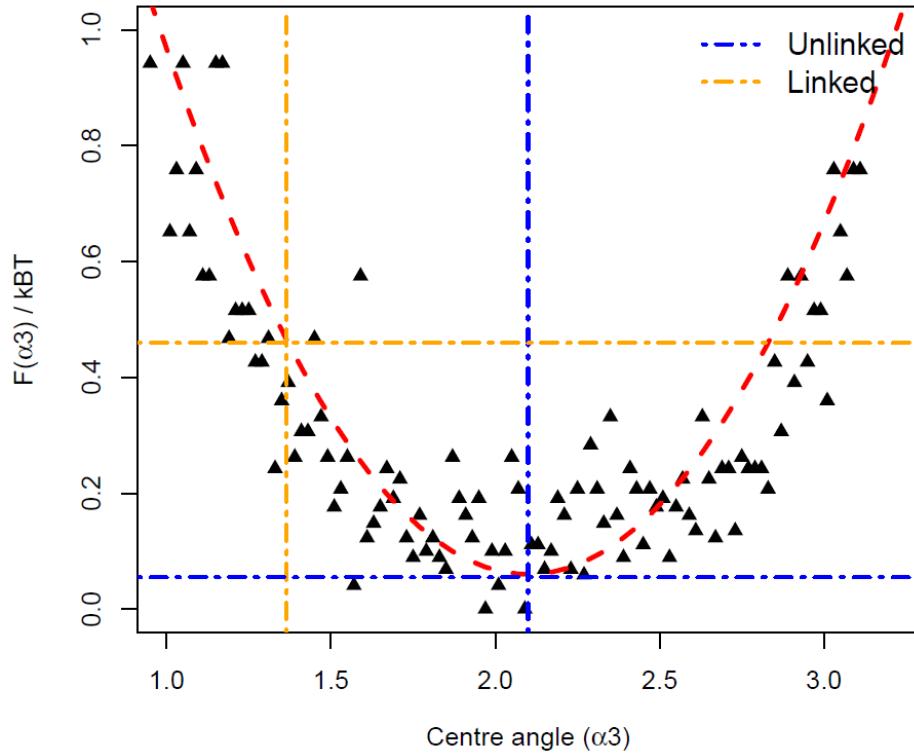


Fig. 3.9 Free energy as a function of the angle distribution between the free arms in a Y-shape (parabola) and the free energy of a Y-shape connected via L6 linker at the mean angle of this configuration (yellow vertical line). The black triangle points represent the achievable angle distribution of α_3 in Fig. 3.8 (left). [47]

3.3.6 The second-level coarse-grained simulation

We used our well-established coarse-grained simulation model (CG model) to examine the structural and rheological characteristics of the hydrogel [46]. The initial model was in the spirit of the Kremer-Grest bead-spring system,[121, 122] treating the Y-shapes as molecules comprising spheres of radius R and mass m connected by harmonic bonds and the linker as a sticky patch. All parameters in the CG model were chosen to match our oxDNA findings and experimental observations for the configurations and binding energies for the sticky overhangs. The arms of the Y-shape are kept stiff by harmonic three-body angle potentials. The Lennard-Jones patches on the terminal spheres of the Y- and L-shapes were complementary, hence can attract each other. Using the observations from the oxDNA simulations, the patch position was constrained to represent the flexibilities of the DNA linkers: in the Y+L6 scenario, the patches were allowed to move freely over their host particle surface; in the Y+L0 scenario, the patches were more constrained. The representative schematics for each scenario were put in Fig. 3.11. The second-level coarse-grained simulation model topologically mimics the detailed oxDNA model and uses the angle information and an entire mapping method for bending free energy. With the help of this more coarse-grained system, we can further do a many-body investigation to measure more bulk physical properties, such as rheological responses and cluster-size distributions.

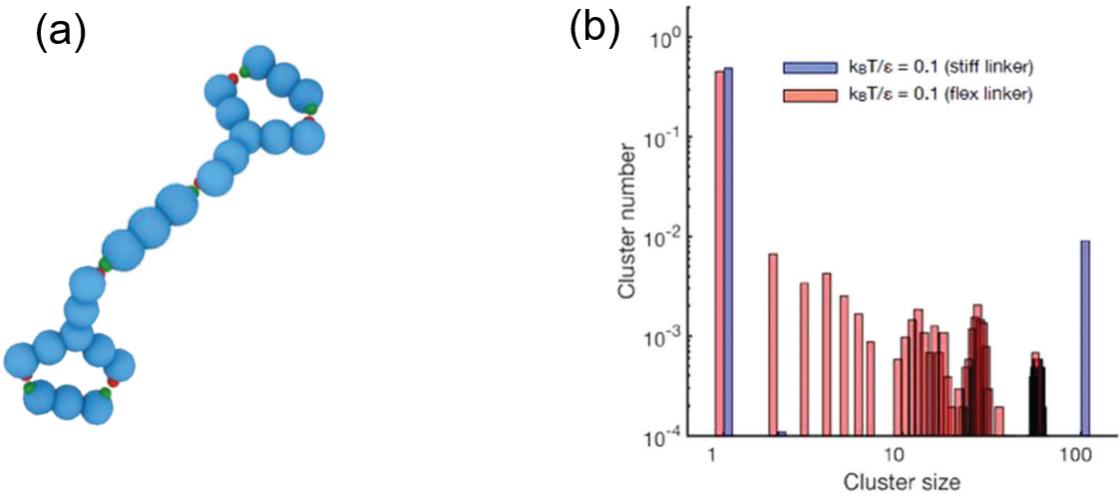


Fig. 3.10 (a) The example of the smallest ‘fully self-looped’ cluster that stopped growth in a group of two Y-shapes and three linkers. (b) Distribution of cluster sizes for Y+L0 (blue bars) and Y+L6 (red bars) scenarios. [47]

Fig. 3.10 shows the distribution of cluster sizes in the cluster liquid phase. As the self-looping behaviour is only observed in a finite probability, there is a randomised possibility

to stop the growth of the cluster. The cluster may be small if all out-layer connectors have been ‘fully self-looped’. The ‘fully self-looped’ means there are no DNA overhangs that are available to connect with another linker/Y-shape DNA. The smallest cluster that stopped growth in a group could be observed only contained two Y-shapes and three linkers (cf. [47] for more detailed results).

3.4 Conclusion

This chapter introduced an entire simulation routine motivated by experimental findings and showed how we use a multi-level coarse-grained strategy to confirm an existing hypothesis related to DNA assembly. In particular, the well-studied oxDNA model offered a great approach to capturing the detailed structural, mechanical and thermodynamic properties of an individual DNA building block associated with ssDNA and dsDNA. The oxDNA simulation proves instrumental when one is interested in the individual behaviour of each single building block. In contrast, in a many-body system, such detailed properties may not be that essential, especially when the focus is on the bulk behaviour of hundreds or thousands of repeat units. In the latter case, the use a further coarse-grained simulation is recommended. However, it is always informative to first look at the detailed structure of a single building block. Once we fully understand the individual properties of the building blocks, then we can further use the information we collected from the detailed model to help us build the next level of simulation. We can then use this method to explain our simulation model more reliably and convincingly.

Exploring the systematic behaviour of the bulk properties of the entire DNA network is challenging for simulations because of the long timescales involved. The numerical study of such a multi-level coarse-grained method allows DNA nanostructures (such as DNA building blocks) to be accurately probed in small-scale and large-scale networks. It is envisaged that through the combination of top-down and bottom-up design for the low and high levels of coarse-grained modelling, the systematic many-body simulation with a large number of DNA building blocks becomes more feasible. This chapter is a starting point to show how we can use this hierarchical simulation method to design the best fitting model for our specific physical questions. More novel ideas and designs will be covered in following chapters.

In conclusion, the detailed oxDNA simulation provided a crucial stepping-stone in understanding our experimental hypothesis. It proved that a minimum of a 6T-long flexible joint results in the linker binding to two arms of a Y-shape concurrently, decreasing the overall system valency for binding, assuming the stiff arms for the Y-shapes and the linear DNA are similarly sized. The detailed closed-to-open ratio measurements of the Y-shape DNA with a flexible linker showed a considerably high possibility of forming a key-looped

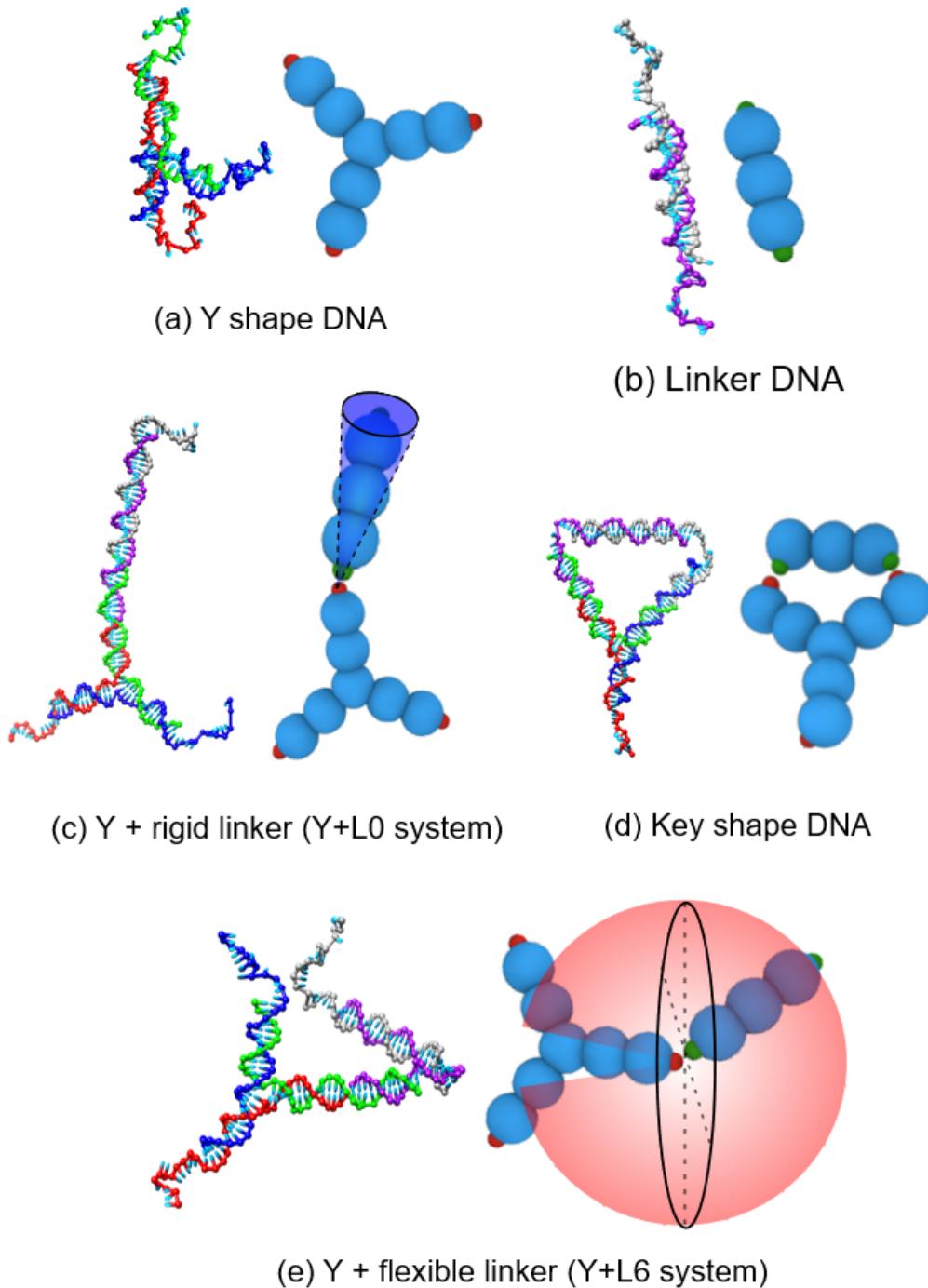


Fig. 3.11 The second-level coarse-grained simulation model based on the oxDNA model. The Y-shape and linker DNA were represented by the bead-spring model in (a) and (b), respectively. The linker flexibility was set by constraining the patch position found in the oxDNA simulations: in (c) the Y shape + rigid linker, the patches were constrained (blue-cone zone); in (e) the Y shape + flexible linker, the patches were allowed to move freely (red-cone zone) over their host particle surface to build the key-shape in (d). [47]

DNA. Also, the tested bending free energy further confirmed the preferential formation of the self-loops. Beyond that, we also found that the oxDNA model gave us a reasonable estimation of the melting temperature for our DNA building blocks, which were in close agreement with our experimental measurements. The second-level coarse-grained system we built based on our oxDNA findings can investigate bulk properties of the system and be further used to explain the emerging cluster phase. Combined with the experimental results, we proved that the self-looping behaviour in flexible DNA building blocks is a valid and highly possible reason for the terminated percolating process in our linker-mediated DNA hydrogel.

3.5 Acknowledgements

I would like to thank Prof. Daan Frenkel for valuable discussions during my preparatory work for this simulation. I also would like to thank Prof. Jonathan Doye, who is one of the developers of the oxDNA model from the University of Oxford, for giving suggestions for utilising oxDNA as a well-studied tool for our hydrogel system; Dr. Iliya Stoev for providing the collaborated experiments and clear explanation of his measurements; Dr. Christopher Ness for building up the next level coarse-grained simulation and further measuring the rheological response of the system; Prof. Erika Eiser for supervising and giving me good guidelines for assignment throughout numerous consultations.

4 | Quantitative Analysis of Flexibility in Nunchucks DNA

The initial motivation of this chapter was triggered by the discussion that we raised in Chapter 3 - how long does this flexible joint need to be to allow a linear linker to bind to two arms of the same Y-shape? In Chapter 3, we made a detailed comparison between the no Thymine and the six Thymine case, but the more exciting question is how we can describe all numbers of base-length or even longer-length in flexible joints. A complete understanding of the properties of a flexible joint as a function of its length will allow us to design the best joint length for a specific purpose, thus giving valuable suggestions in a stage of experiment preparation. From the former experiment [47], we observed that changing the number of nucleotide bases added to the flexible joint has a tremendous effect on the overall properties of the system. However, it is still challenging to give a quantitative explanation from the experiments to say which length of the flexible joint is the best option to provide enough flexibility for a bending behaviour. Therefore, in this chapter, I will give a systematic analysis of all sizes of flexible joints to summarise a complete understanding of the bending flexibility through our nunchuck-like DNA building blocks. Moreover, this chapter is a preparatory study investigating the formation of new liquid crystal phases generated by angle-confined DNA mesogens in Chapter 5. Hundreds of thousands of these duplicated mesogens are required to form a liquid crystal phase. The oxDNA model is then unavailable to simulate such a large-scaled system as it is highly costly, since it models each particle in great detail. However, the detailed behaviour of a single mesogen is still essential to know before we start building up the many-body system. Hence, I applied this multi-level coarse-grained strategy to bridge the gap between the detailed and more coarse-grained models.

4.1 Introduction

‘Nunchucks’ DNA was mentioned as a term of ‘gaped DNA’ by introducing an unpaired single-stranded DNA spacer in the middle of a DNA duplex in previous experimental work

[123–125, 54, 126]. However, none of the existing experiments gave an effective simulation method and quantitative analysis to this single DNA building block. People also invented ‘nunchucks-like’ DNA origami[127], while the designed DNA origami is relatively large (formed by thousands of bases) and not our concern of using as a nano-sized mesogen. In addition, investigating a DNA origami’s dynamic mechanism is different from using DNA as a single mesogen. For an individual nunchucks DNA mesogen, the essential factor dominating the bending behaviour is the length of the ssDNA joints that bridge two stiff dsDNA arms.

Although there is a wealth of knowledge on the biological characteristics of single-stranded DNA (ssDNA) molecules [128–131], there is currently little quantitative data available on the base length dependency of intrinsic rigidity stiffness for ssDNA polymers in solution [126]. It was not easy to do early research on the structural characteristics of ssDNA since it required high demanded accuracy control of monitoring a single molecule with accurate size-fractionation [132, 133]. Therefore, a simulation point of view on these properties becomes a beneficial routine for conducting quantitative analysis for flexibility in nunchucks DNA. Even though Chiappini et al. did excellent simulation work on similar ‘bending’ particles [134], this work was highly focused on the confined geometrical aspects but can not referred as a DNA particle with specific flexibility. We know that dsDNA can be considered as a polymer with specific rigidity [135, 136], and many currently existing bead-bond DNA models are treating dsDNA as a string of rigidly connected beads [88, 137]. However, the simulating strategy of the extra added ssDNA chain between the two dsDNA arms is still unclear. Triggered by these reasons, we are keenly interested in finding a systematic map to see how flexibility plays a role in a real DNA mesogen at the single base level.

In a preliminary simulation study, we aim to find a means of simplifying the translational and rotational behaviour of the nun-chuck particles, without compromising any of the essential features of the actual DNA-based material. Hence, here, we continually draw on the oxDNA model [88, 85] with specific concerns on an individual nunchuck DNA with various levels of flexibility. We desire to find a general expression for our bending free energy affected by the varied flexibility in our nunchucks DNA. This flexibility should be dominated by the lengths of the ssDNA joints in the middle part of the nunchucks. If we can find a parameter to describe this flexibility, we can define a new potential in the second-level coarse-grained design. This new potential equation can hugely reduce the computational complexity of simulating a nunchuck DNA mesogen, in the meanwhile, without losing the realistic thermodynamic and dynamic behaviour.

To do so, I first measured the bending angle between two stiff arms of all-lengths based nunchucks DNA in a detailed oxDNA model. Then I used these data to conduct a bending free-energy analysis with a quantitative discussion about the bending behaviour of a single nunchuck DNA with different levels of flexibility. To get a tidy visualisation of the data representation, I conducted statistical analysis in R [138]. Using all collected information, I integrated all required flexibility into one simple stiffness factor driven potential equation to mimic the overall behaviour of a nunchucks DNA. The final streamlined potential equation can then be used as an ‘all-in-one’ intra-molecular potential of mimicking nunchucks DNA for the next-level coarse-grained system. This final simplified potential should be especially applicable in simulating a many-body system with high concentrations.

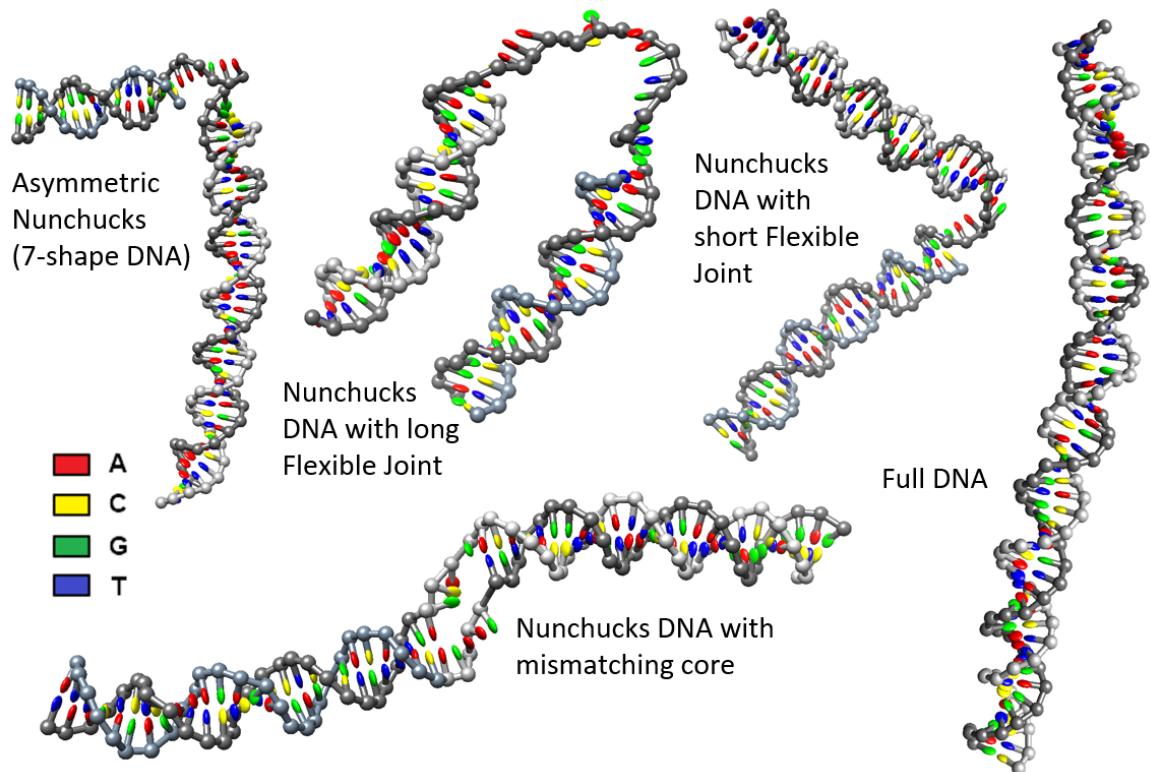


Fig. 4.1 Schematic of example designed nunchucks DNA as mesogens, (from left to right) including (1) asymmetry nunchucks DNA; (2) nunchucks DNA with long flexible joints; (3)nunchucks DNA with mismatching and; (4) nunchucks DNA with short flexible joints and (5) full DNA with the same length as a reference. The cartoon of red, yellow, green and blue colour represents nucleotide A,C,G,T respectively - detailed sequences listed in **Appendix A.2** and a customised visualisation scheme is used to create this diagram introduced in **Appendix B.1**.

4.2 Simulate Nunchuck DNA as Mesogen

At the beginning of this work, we prepared experiments to assemble the nunchucks DNA in a dilute state. However, experiment work becomes extremely challenging if we require a specific high density to investigate liquid crystal properties. It requires enormous amounts of DNA samples and the purification work of dense DNA suspensions becomes even more complicated. Hence running simulations provides a wise choice to help study the capability of utilising these nunchucks DNA as intended mesogens. Subsequently, with the help of simulation, we can clearly understand the mesogen properties and further design a well-papered experiment with precise targeting.

We initially designed four different types of nunchucks DNA, based on a 60 base-pair long dsDNA forming a semi-rigid rod. Fig. 4.1 illustrates in detail the nunchuck DNA mesogens simulated in the oxDNA model: including (1) asymmetric nunchucks DNA; (2) nunchucks DNA with long flexible joints; (3)nunchucks DNA with mismatching core and (4) nunchucks DNA with short flexible joints. The full paired dsDNA with the same length are displayed here as a reference. The detailed DNA sequences of all simulated nunchucks DNA are listed in **Appendix A.2**. We chose a 60-bases length DNA sequence because it is sufficiently within the persistence length, so that each dsDNA arm can be treated as a relatively stiff rod. Meanwhile, this length offer enough space to modify flexible joint in the middle, e.g. the longer the ssDNA joints, the more flexible the mesogen. Additionally, in investigating liquid crystal phases, the longer length of the rigid rods always makes the phase transition more feasible at a lower density [139].

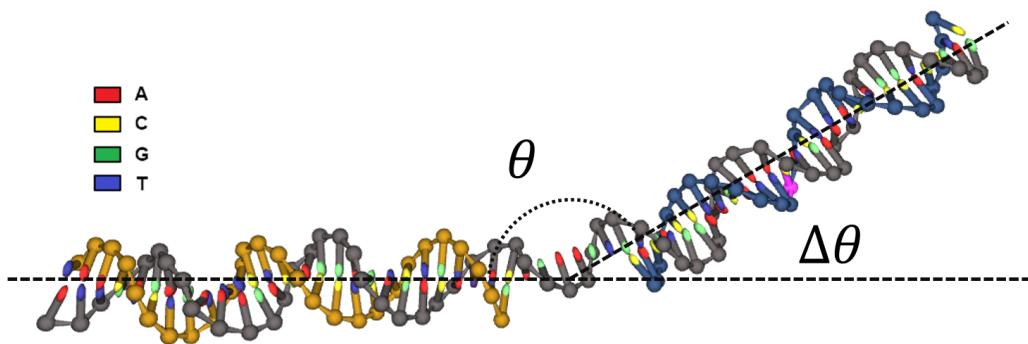


Fig. 4.2 The model of nunchucks DNA, with θ being the bending angle, $\Delta\theta = \pi - \theta$.

To start with, we measured bending angle distributions for all aforementioned nunchucks DNA in Fig. 4.1. The bending angle (θ) was defined as Fig. 4.2, representing the angle between two rigid dsDNA arms on each side of the ssDNA joints. From comparing the achievable bending angle distribution between Fig. 4.4(a) and (b), we found that the position

of the ssDNA in an asymmetry nunchucks DNA is not a solid dependent factor in affecting the middle part flexibility. Fig. 4.4(c) gave us a much wider angle range, confirming that a longer flexible chain provides higher flexibility. The mismatching case in Fig. 4.4(d) was interesting, but it is relatively hard to provide a dominating single quantitative factor that can directly judge the bending behaviour. As we assume, its bending behaviour also highly depends on the exact sequence order of the mismatching core. Hence we mainly focus on the quantitative analysis work, specifically on the nunchucks DNA only with short and long flexible joints in the middle. We should consider all length ranges of the ssDNA joint to describe the bending property with the following concerns systematically. (1) The nunchucks DNA with a long flexible chain can be fully bent, but how many bases are needed on the chain to be defined as fully flexible? (2) The case of a short flexible chain can confine the range of accessible angles between the dsDNA arms, but how strong is this confinement, and how does it depend on the length of the ssDNA joint? This also reminds us of the initial concern that we need to know the result of using 2T and 4T flexible joints in our Y+L system in Chapter 3. The following task is to find a simple but accurate description to answer the above questions.

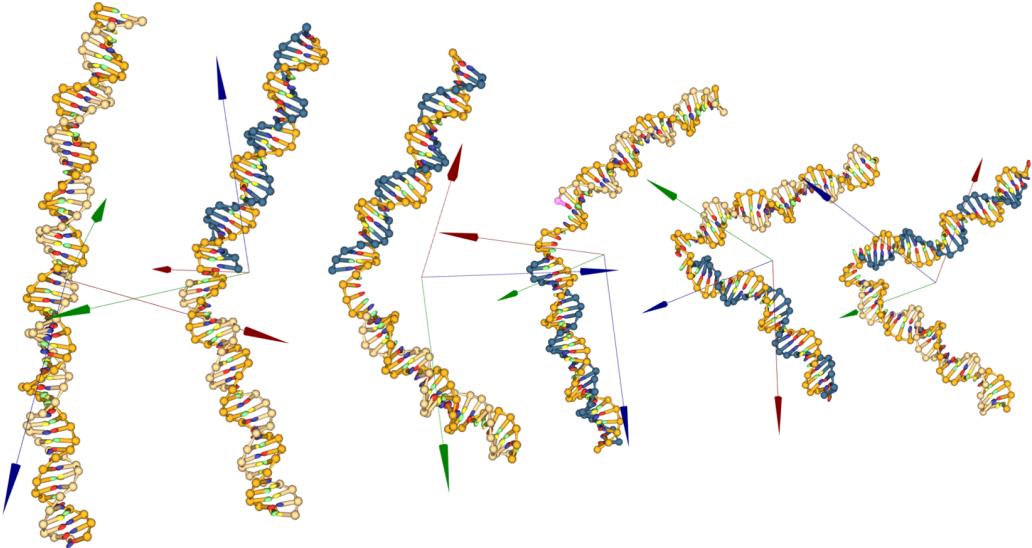


Fig. 4.3 Schematic of nunchucks DNA mesogens (from left to right) with 0-, 2-, 4-, 6-, 8-, 10-nucleotide bases on ssDNA flexible joint, respectively. The red, blue and green arrows represent each model's simulated geometry directions in x,y and z. Figure generated in oxView [140].

Therefore, to fulfil the quantitative description, we built up a set of nunchucks DNA mesogens ranging from 0→10 nucleotide bases on ssDNA flexible joint in the middle (Fig. 4.3). More prolonged cases (>10 base pairs) were also covered in the later discussion.

The ultimate goal is to find a systematic description of the flexibility with one intra-molecular potential to simulate the nunchucks DNA with all length range of ssDNA joints, then applied to the many-body system with hundreds and thousands of repeated DNA mesogens.

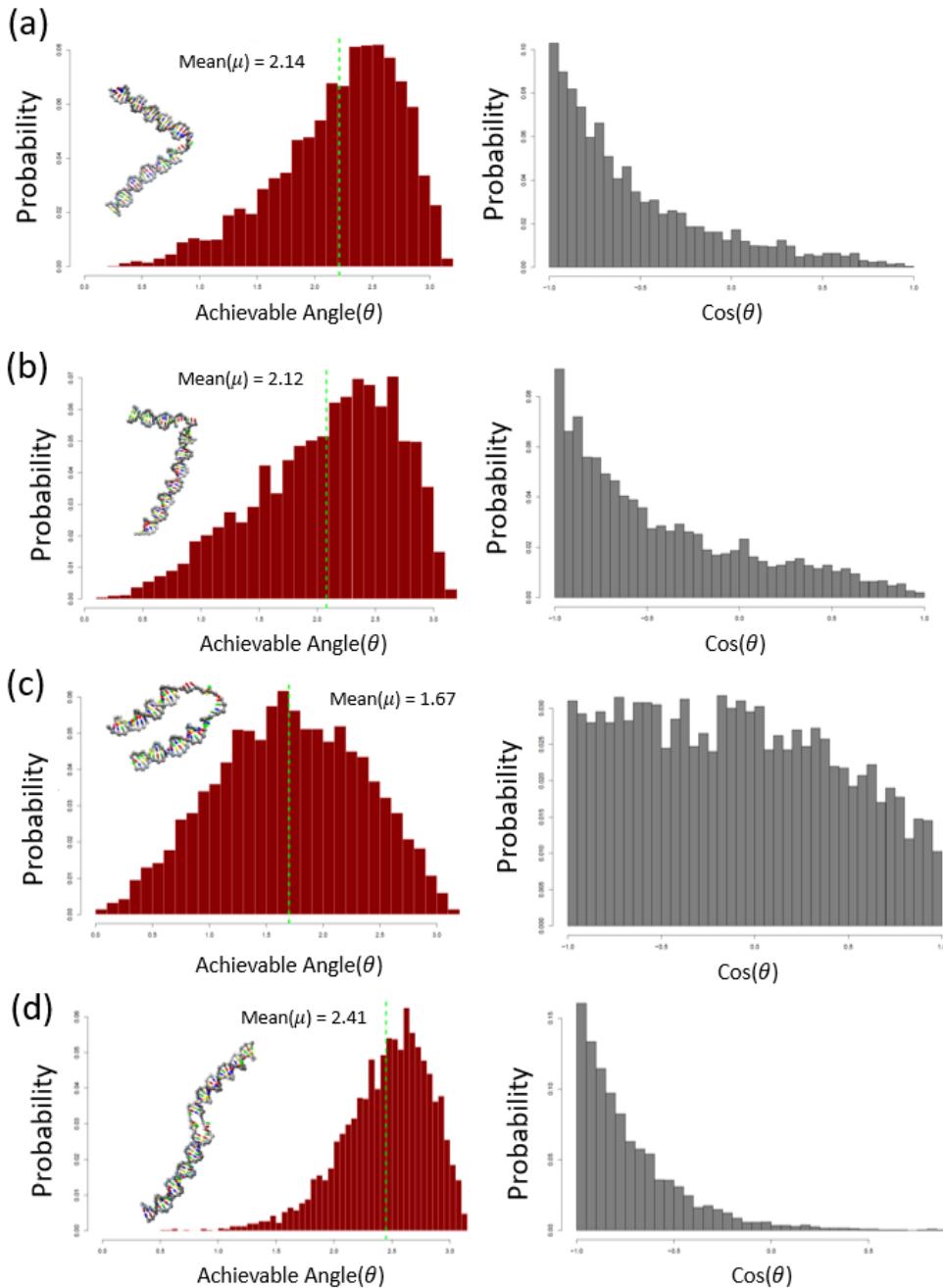


Fig. 4.4 Statistical Achievable angles (θ) and $\cos(\theta)$ of all achievable angles for designed DNA mesogens. θ is the bending angle between two linear dsDNA arms. (a) Nunchucks DNA with short Flexible Joint. (b) Asymmetric Nunchucks (7-shape DNA). (c) Nunchucks DNA with long Flexible Joint. (d) Nunchucks DNA with mismatching core.

4.3 Results and Discussion

4.3.1 Bending free energy calculation

In this section we shall introduce the bending free energy for the nunchuck DNA. This bending free energy will determine the probability of finding a particular angle between the dsDNA arms of the particle and will be a function of the number of ssDNA bases in the joint. The defined angle (θ) between the arms of the nunchuck DNA sketched in Fig. 4.2 was calculated by given the two arm vectors \vec{r}_1 and \vec{r}_2 . Then the $\cos \theta$ was given by:

$$\cos \theta = \frac{\vec{r}_1 \cdot \vec{r}_2}{|\vec{r}_1| \cdot |\vec{r}_2|}. \quad (4.1)$$

By symmetry, the probability, and hence the free energy, of finding a particular angle between the arms of the particle can only depend on the $\cos \theta$. The probability of finding a specific angle between the dsDNA arms $P(\cos \theta)$ can be written in terms of the bending free energy $F(\cos \theta)$ as:

$$P(\cos \theta) = \frac{1}{Z} \cdot e^{-\frac{F(\cos \theta)}{k_B T}}, \quad (4.2)$$

where k_B is the Boltzmann constant, and T is the absolute temperature, with Z being the partition function to ensure the probability is normalised. To find a parameter that can quantitatively describe the level of flexibility in each base length. We re-write the Eqn. 4.2 to calculate the bending free energy ($F(\cos \theta)$) as a function of $\cos \theta$ distribution,

$$F(\cos \theta) = -k_B T \cdot \ln[Z \cdot P(\cos \theta)] = -k_B T \cdot \ln P(\cos \theta) + C \quad (4.3)$$

with $C = k_B T \cdot \ln Z$ is a normalised constant (relate to the partition function Z). Its physical meaning is the probability of comprising all accessible states occurring in the system.

4.3.2 Statistical distribution

In order to determine the bending free energy, we need to know the probability distribution of all achievable angle states (P). This probability distribution was determined from our simulated outcomes by monitoring the bending angle frame by frame. We plotted statistical results of a flexible bending angle distribution in Fig. 4.5, based on specific lengths of ssDNA joints ranging from 0 to 10 bases. All bending angle calculations were taken from the equilibrium state when the overall energy of the nunchuck DNA was measured to be stable

in oxDNA simulations. Then we pick up ten individual periods in equilibration period to calculate the $\cos \theta$ distribution accordingly. Each period provides at least one thousand measurement samples to ensure statistical accuracy. If the simulated nunchucks DNA are ideally flexible, we should expect a horizontal line for the $\cos \theta$ distribution. From our statistical data, it was noticeable that flexible joints with eight bases or higher numbers were close to a horizontal line, trending to a fully flexible state. Moreover, the data displayed a clear increasing range of distribution for the joint length from 0 to 8 bases.

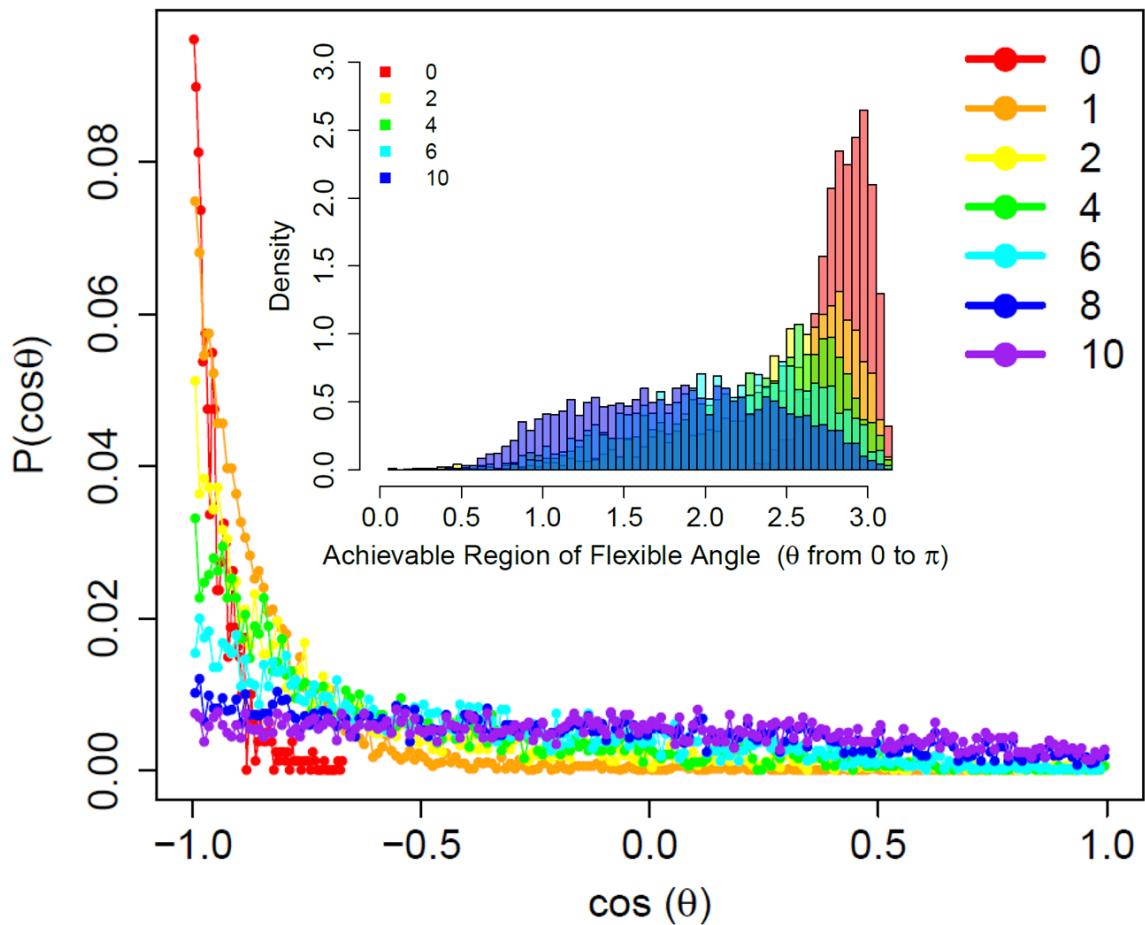


Fig. 4.5 Statistical results of flexible angle distribution and the probability of $\cos \theta$ in a nunchucks DNA; the numbers 0, 1, 2, 4, 6, 8 and 10 represent the counts of ssDNA bases we used in our flexible chain between two dsDNA arms. ‘0’ means it is a full DNA duplex.

To see the relationship more clearly, we plotted the bending free energy ($F(\cos \theta)$) as function of $\cos \theta$ in Fig. 4.6. This can then be used to fit the parameters in the simplest possible form for $F(\cos \theta)$, which we will determine in the following section.

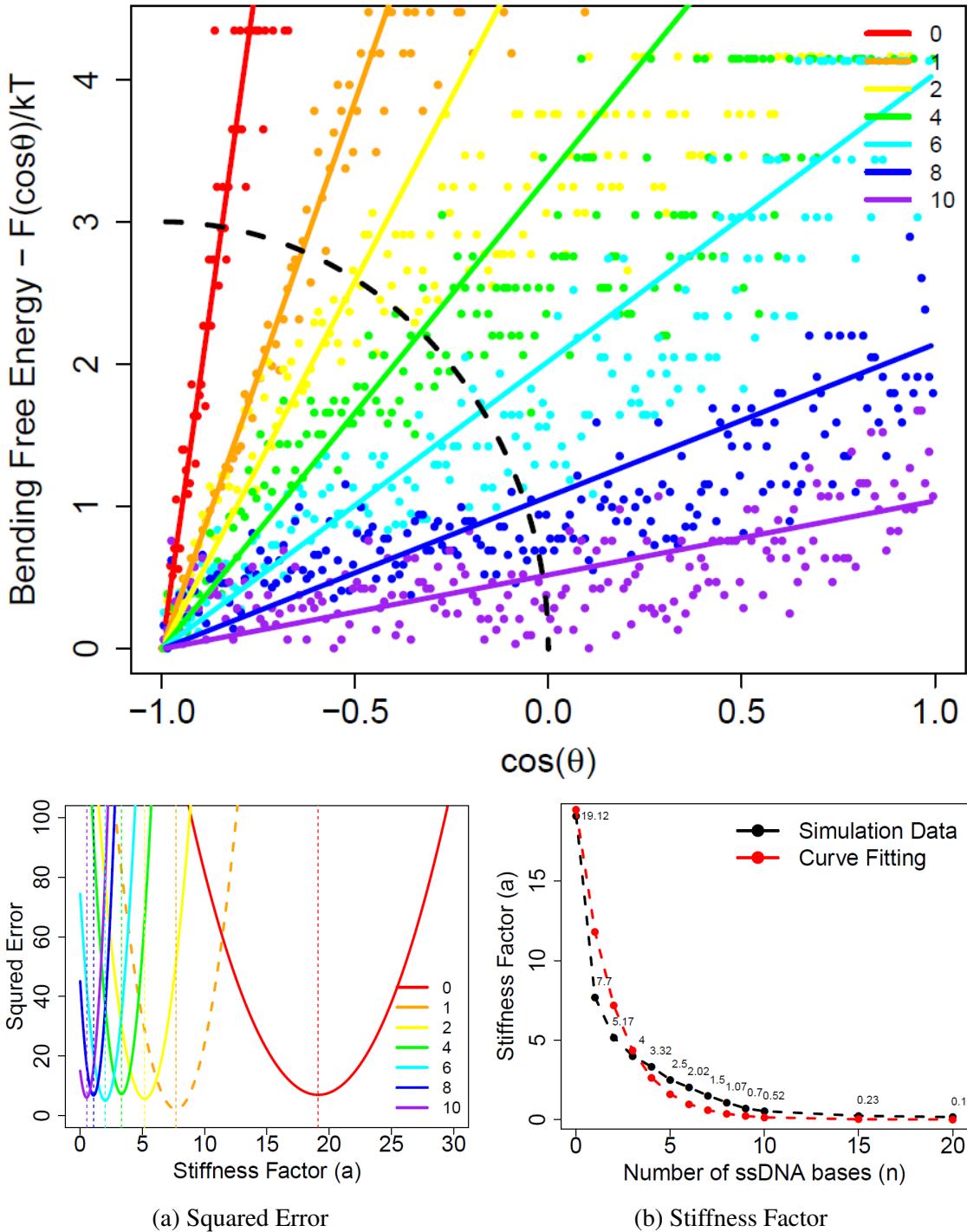


Fig. 4.6 Statistical results of bending free energy $F(\cos \theta)$ calculated from $\cos \theta$ distributions. (a) The squared error shows quantitative data to find the best fitting curve within the black dash region; (b) the stiffness factor is the gradient slope of the best linear fitting line. The stiffness factor represents the flexible level according to different lengths of ssDNA joints used in nunchucks DNA. A smaller value of the stiffness factor means higher flexibility.

4.3.3 Mathematical model

The statistical data in Fig. 4.6 are fitted by linear regression. Linear regression is used because simple mathematical considerations lead to the simplest form of $F(\cos \theta)$ being linear in its argument. If we assume an ideal nunchuck, the equilibrium angle between its two arms should be $\theta_{eq} \approx \pi$. At this most relaxed state, the bending free energy of the nunchucks should also be at a minimum. If we add a small fluctuation ($\Delta\theta$, also sketched in Fig. 4.2) to the equilibrium state, with $\Delta\theta \ll 1$, the bending free energy near the equilibrium state should slightly increase. And the bending angle (θ) and the applied fluctuation ($\Delta\theta$) will have the following relation:

$$\theta = \theta_{eq} - \Delta\theta = \pi - \Delta\theta. \quad (4.4)$$

The updated bending free energy with a small angle fluctuation can be written as:

$$F(\cos(\theta_{eq} - \Delta\theta)) = F(\cos \theta_{eq} \cos \Delta\theta + \sin \theta_{eq} \sin \Delta\theta). \quad (4.5)$$

Consider $\theta_{eq} \simeq \pi$, we have $\sin \theta_{eq} = 0$. The above equation can be rewritten to be:

$$F(\cos(\theta_{eq} - \Delta\theta)) = F(\cos \theta_{eq} \cos \Delta\theta). \quad (4.6)$$

Using a Taylor expansion, assuming $\Delta\theta \ll 1$ and using $\cos(\theta_{eq}) = \cos \pi = -1$, we get,

$$F(\cos(\theta_{eq} - \Delta\theta)) \simeq F\left(\frac{1}{2}(\Delta\theta)^2 - 1\right) = F(\cos \theta_{eq} + \frac{1}{2}(\Delta\theta)^2). \quad (4.7)$$

And using Taylor's expansion again, we get,

$$F(\cos \theta_{eq} + \frac{1}{2}(\Delta\theta)^2) \simeq F(\cos \theta_{eq}) + \frac{1}{2}(\Delta\theta)^2 F'(\cos \theta_{eq}) + O((\Delta\theta)^4), \quad (4.8)$$

with $O((\Delta\theta)^4) \simeq 0$. We may define the free energy at equilibrium to be zero. Again from the Taylor expansion, we have $\frac{1}{2}(\Delta\theta)^2 = 1 - \cos \Delta\theta$ and from the angle definition $\Delta\theta = \pi - \theta$, we have $(1 - \cos \Delta\theta) = (1 - \cos(\pi - \theta)) = (1 + \cos \theta)$. Hence, the above equation can be rewritten as:

$$F(\cos(\theta_{eq} - \Delta\theta)) \simeq (1 + \cos \theta) F'(\cos \theta_{eq}). \quad (4.9)$$

For non-ideal nunchucks DNA, the θ_{eq} might not be perfectly equal to π . It has a small bias because of the asymmetry cut on one side of the double strand, but this diminutive bias does not affect this calculation. The bending free energy should get a linear relation with

cosine of the fluctuation angle ($\cos(\Delta\theta)$) in a small angle estimation. In a three-dimensional space, we consider the actual fluctuations of the angle can be in any direction, but no matter which direction, we always have $\theta = \theta_{eq} - \Delta\theta$, as the nunchucks-model considered to be symmetric. In summary, the simplest prediction of a linear relation can be written as:

$$F(\cos(\theta)) = \alpha_n \varepsilon (1 + \cos \theta), \quad (4.10)$$

where ε has units of energy, α_n is the ‘stiffness factor’ to represent the flexibility of the ssDNA joints. It is simply the gradient slope of the bending free energy from Eqn. 4.9. And α_n itself should be a discrete function ($f(n)$) according to number of ssDNA bases (n) used in our nunchucks DNA:

$$\alpha_n = f(n). \quad (4.11)$$

4.3.4 Determine the stiffness factor

From the above work, we have observed the probability function $P(\cos \theta)$, and the statistical results of the free energy is calculated from: $F(\cos \theta) \propto -\ln P(\cos \theta)$, mathematically, we can get the stiffness factor ($f(n)$) from a gradient slope of each fitted curve. Consequently, we constructed a quantitative parameter that describes the flexibility of the nunchucks; the stiffness factor ($f(n)$). The value of $f(n)$ depends on the lengths of ssDNA joints (number of bases) used in our nunchucks DNA. Variable n is the number of bases used on the ssDNA flexible joint between two dsDNA arms in our nunchucks DNA, and θ is our bending angle.

The idea is then how to find the correct value of the $f(n)$ to represent the overall bending behaviour of the nunchucks DNA with a specific length accordingly. The statistically measured values of $f(n)$ can be found using linear regression, shown in Fig. 4.6. The calculated best value of $f(n)$ are listed in Tab. 4.1. The small value of $f(n)$ indicates a more flexible nunchucks DNA. In the next step, we can apply this one-line equation to build up our next-level coarse-grained system to achieve a many-body simulation with hundreds and thousands of nunchucks DNA mesogens. Before that, we need to figure out when and how we can define a ‘fully flexible’ nunchuck DNA mesogen. Therefore, we introduced a colour zone representation in the following section.

n	0	1	2	3	4	5	6	7	8	9	10	15	20
$f(n)$	19.12	7.7	5.17	4.0	3.32	2.5	2.02	1.5	1.07	0.7	0.52	0.23	0.1

Table 4.1 Values of stiffness factor $f(n)$ according to the number of bases on the ssDNA flexible joint in nunchucks DNA. Data extracted from the statistical analysis in Fig. 4.6.

4.3.5 Colour zone representation and flexible saturation

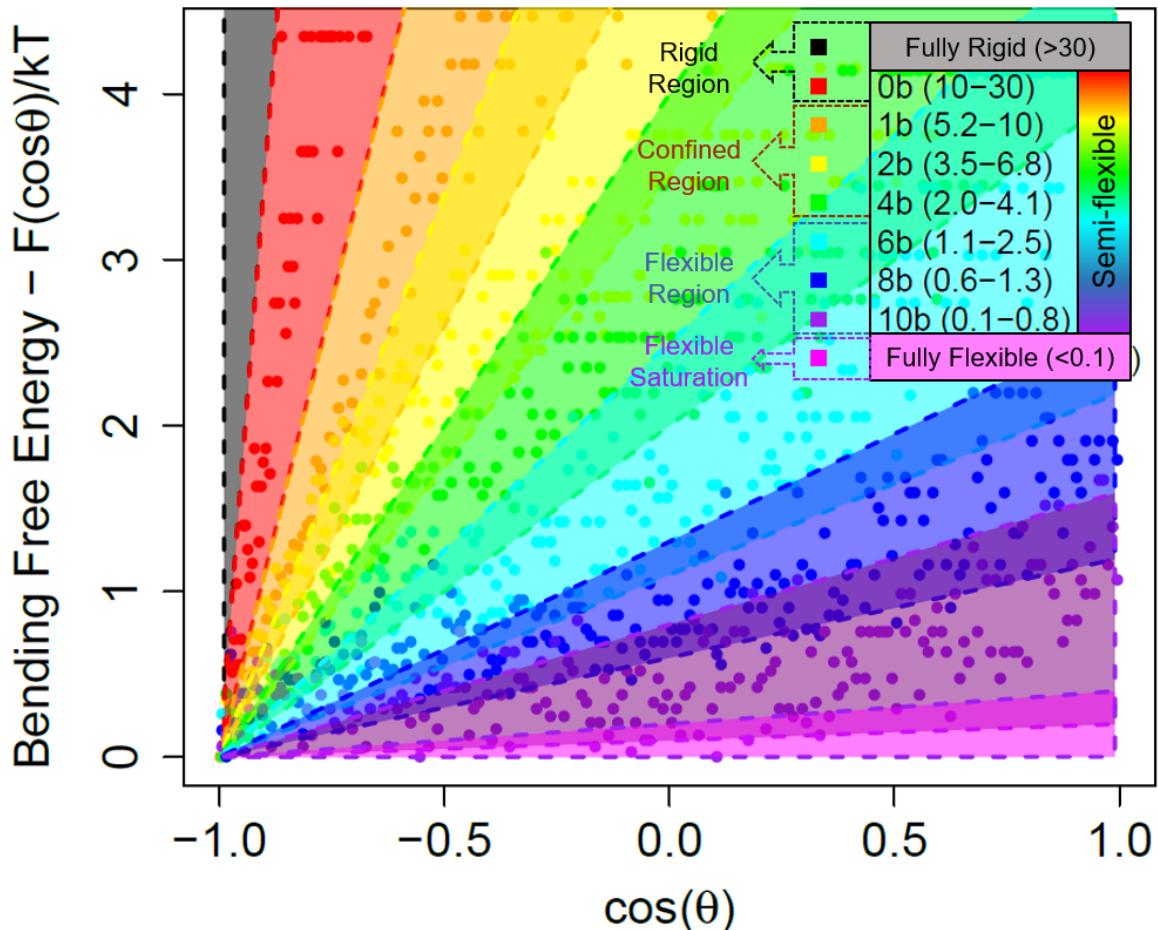


Fig. 4.7 Color Zone representation of bending free energy with all achievable ranges. The grey area was defined as ‘fully rigid’; the pink area indicated the nunchucks DNA reached a ‘fully flexible’ state; the rainbow colour covered as ‘semi-flexible’ when N_s ranging from 1 to 10 bases. The bracketed value in the figure legend indicates the range of achievable stiffness factors from statistical measurements for its corresponding ssDNA joint length (N_s).

A colour zone representation (Fig. 4.7) was added here to illustrate all possible ranges of the stiffness factor obtained from statistical analysis. Each colour zone represents a case study of nunchucks DNA with a specified length of ssDNA joint (N_s). The boundaries of the colour zone were determined by its minimum and maximum achievable gradient slope value in linear regression. The detailed threshold of each boundary’s slope value can be calculated from Fig. 4.6a. For each number of N_s , there is a best-fit stiffness factor at the minimum value of these quadratic curves. The region boundaries are quantified by the points of intersection by their squared error functions. With the help of this zone representation,

we defined a regional concept to describe the system. In particular, two specific areas in the colour zone representation are noteworthy: First, the grey area shown in Fig. 4.7, which is defined as ‘fully rigid’, there was no statistical data measured within this area. This area is only reachable if we analyse an ideal rigid rod. However, in an accurate DNA simulation, this area was statistically unachievable. It tells us that dsDNA itself is a semi-flexible polymer, and it has some flexibility even within its persistence length. Second, the pink area indicated the flexibility of the ssDNA joint achieved a ‘fully flexible’ condition. There was almost no penalty for performing the fully bending behaviour in a fully flexible state, when $N_s > 10$.

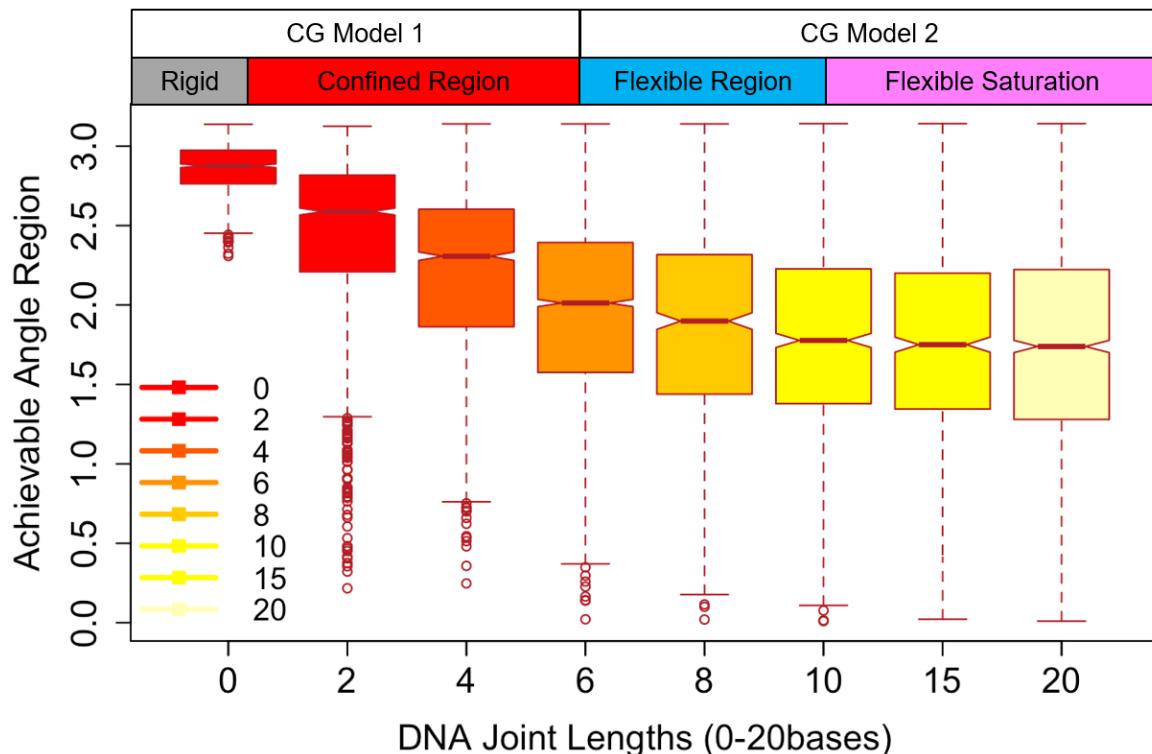


Fig. 4.8 Box-and-whisker diagram of achievable angle region for different lengths of ssDNA joints (N_s). Four regions defined (1) Rigid region ($N_s < 1$). (2) Confined region ($1 \leq N_s < 6$). (3) Flexible region ($6 \leq N_s \leq 10$). (4) Flexible saturation ($N_s > 10$).

All areas¹ between ‘fully rigid’ and ‘fully flexible’ were defined as ‘semi-flexible’. However, even at the same level of the ‘semi-flexible’, we found that one base difference would significantly change its achievable bending angles. The bending behaviour of the nunchucks DNA with short flexible joints was much more restricted. Especially when

¹The rigid region was determined to represent full DNA with ideal rigidity. The confined region was determined by whether or not the nunchuck DNA can be ‘fully folded’ (two arms are closed and paralleled each other). The flexible region is when observing ‘fully folded’ nunchucks DNA, but still, an increasing number of N_s can slightly offer higher flexibility for their bending behaviour. In the saturation region, the flexibility was fixed no matter the value of N_s , and the achievable angle distributions from statistical outputs no longer change.

$N_s \in [1, 5]$, more substantial confined bending angles prevent two arms from being fully folded. To further compare the bending limitations, we plotted the box-and-whisker[141] diagram (boxplot) in Fig. 4.8 for the statistical results. In general, the flexibility of the nunchucks increases with increasing ssDNA chain length. However, in a short joint case ($N_s < 6$), the minimum bending angle (θ) never reached zero, meaning two stiff arms can never be fully folded – hence we defined this region as a ‘confined region’. In this region, the nunchucks DNA with a joint length of less than 6-bases would be confined to bend into a ‘herringbone-like’ structure. This structure might provide the most interesting feature in the liquid crystal phase observation, which we will discuss in Chapter 5.

Therefore, we recommend adding at least 6-bases on the flexible middle chain to design a flexible ssDNA joint. It should give enough flexibility to make it much easier for two stiff arms to bend or even thoroughly fold. It was only when $N_s \geq 6$ that the fully folded behaviour started to occur. And we defined the nunchucks DNA with $N_s \geq 6$ in a ‘flexible region’. Although the ability to achieve the fully folded situation is still increasing in longer cases, achievable angle distribution does not change much when $N_s > 10$, where we say it has achieved the ‘flexible saturation’. Once the flexible joint reaches flexible saturation, no longer adding the extra lengths, the flexibility of the ssDNA chain will remain in a similar state. The saturation plateau matches the ‘fully flexible’ description in Fig. 4.7. Also, the data plotted in the boxplot indicated us that the 2- and 4- bases cases were much more rigid than the 6-bases case. If the ssDNA joints are less than 6-bases, bending two dsDNA arms is much harder to achieve the angle requirement for forming a self-loop. This also explains why the Y+6T system in Chapter 3 is much more flexible and hence gets an easier self-looping behaviour in the end. Achievable ranges of stiffness values collected from colour zone representation and the region definitions in this section benefit designing a more realistic model of nunchucks DNA in future design.

4.3.6 Forwarding a second-level coarse-grained simulation model

In an accurate model description, the flexible chain has finite lengths, especially in a nunchucks DNA with long flexible joints. We considered two types of coarse-grained (CG) models when designing the second-level coarse-grained model. The first model [CG Mode 1] has one single flexible bead in the middle, which fits the nunchucks DNA with a short flexible joint when it is within a confined region ($N_s < 6$) in Fig. 4.8. The second model [CG Mode 2] has two flexible beads in the middle, which fits the nunchucks DNA with a long flexible joint ($N_s \geq 6$). In general, one sphere beads represent six bases length of the nucleotides. Since each arm has around thirty bases in length, five sphere beads on each rigid arm can give a proper relative ratio of presentation (feasible model illustrated in

Fig. 4.9). The other concern of designing CG model types according to their joint length is the flexible saturation we observed in the ‘fully flexible’ nunchucks DNA. Once the model gets the flexible region, the cost of bending is almost negligible, i.e. the stiffness factor is nearly equal to zero. Hence, in [CG model 2], we can release the bending confinement and let the arms be fully flexible to move. While for the model design for liquid crystal study, we suggest using [CG model 1] in simulating mesogens with confined angles, as it matches better with preferring to bend into a ‘herringbone’ structure. The detailed report about second-level coarse-grained simulation will be discussed in Chapter 5 and a vivid schematic of this multi-level coarse-grained method is illustrated in Fig. 4.11.

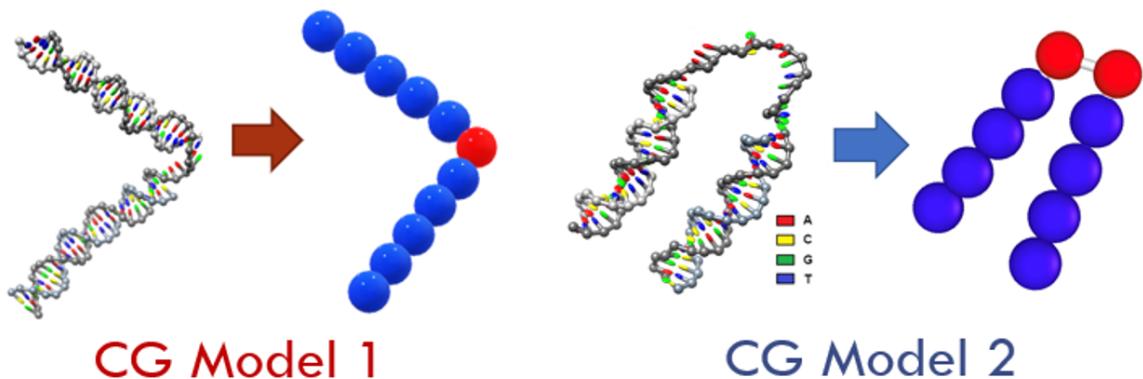


Fig. 4.9 Feasible CG model design. CG model 1 fits nunchucks DNA with short chain length ($N_s < 6$); CG model 2 works for longer chain cases ($N_s \geq 6$).

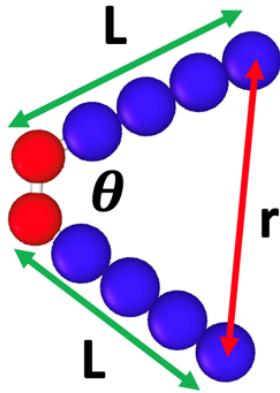
4.3.7 Translate bending free energy into a virtual bond potential

In most of the simulated potential in a coarse-grained model, the distance-dependent potential is much more favourable than the angle-dependent potential. The angle-dependent potential usually needs a three-body relation, while the distance-dependent potential only needs two. Hence, we further translated our bending free energy equation into a virtual bond potential. The virtual bond does not exist between the arms, but it can ideally provide an unseen attractive or repulsive force to provide the required potential. The translation is simply substitute Eqn. 4.12 into Eqn. 4.10:

$$\cos \theta = \frac{L^2 + L^2 - r^2}{2L^2}. \quad (4.12)$$

Eventually, we will get a distance-dependent potential in Eqn. 4.13, with all corresponding parameters are illustrated in Fig. 4.10:

$$F(r) = \frac{\alpha_n}{2L^2} k_B T (4L^2 - r^2). \quad (4.13)$$



Bending free energy:
 $F(\cos\theta) = \alpha_n \cdot k_B T (1 + \cos\theta)$
 $\alpha_n = f(n)$

Virtual bond potential:

$$F(r) = \frac{\alpha_n}{2L^2} k_B T (4L^2 - r^2)$$

Fig. 4.10 Translate the bending free energy into a virtual bond potential in second CG model. L is the arm length, θ is the bending angle, r is the virtual distance between the head of the two arm and α_n is the stiffness factor.

4.4 Conclusion

In this chapter, we focused on a one-molecular simulation for nunchucks DNA as an individual mesogen, which comprises two dsDNA segments connected by a small ssDNA segment. The idea of this chapter is written as a bridge to combine the knowledge between fine-grained and coarse-grained simulations. It is always better to start from a fine-grained model to first understand the detailed properties. And then, we applied the detailed findings in a simplified expression to describe physical properties accurately. Hence, this chapter is essential before we design a large-scale system. We balance a trade-off between computational complexity and a better explanation of physical properties to avoid designing an over-detailed/over-simplified system.

The quantitative analysis of the flexibility in nunchucks DNA mainly depends on the bending angle measurements and statistical distribution. We defined a bending free energy equation to represent the general bending behaviours of two stiff dsDNA arms. Using the statistical method and mathematical model, we find a single parameter from the linear regression of the statistical data. We, therefore, introduced a stiffness factor $f(n)$ to quantitatively describe the flexibility of ssDNA used in nunchucks DNA with a determined joint length. We

concluded the final expression of the bending free energy: $F(\cos \theta) = f(n) \cdot k_B T(1 + \cos \theta)$, with $f(n)$ defined as the stiffness factor. The variable n is the number of bases used on the ssDNA flexible joint between two dsDNA arms in our nunchucks DNA, and θ is our bending angle; and its distance-depended version: $F(r) = \frac{f(n)}{2L^2} k_B T(4L^2 - r^2)$, where L is the arm length, and r is the virtual distance between the head of the two arms.

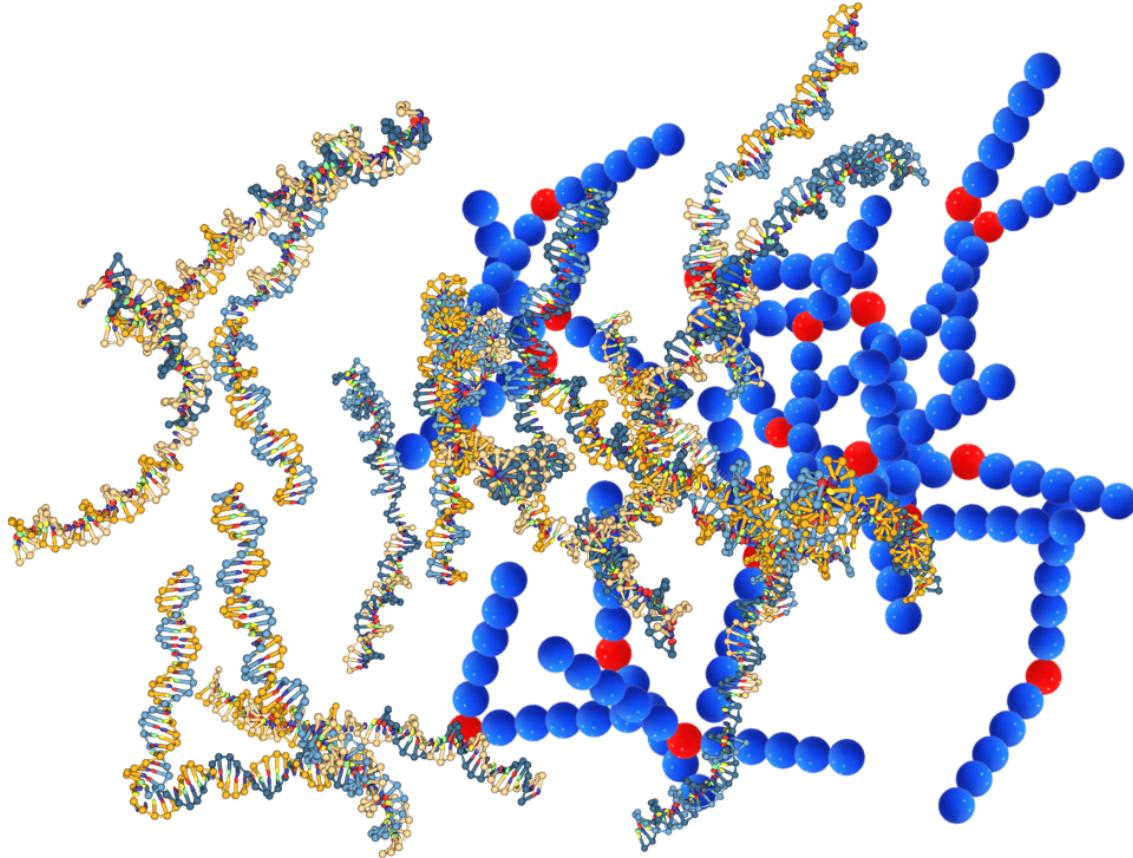


Fig. 4.11 A simulation strategy from first-level (build in oxDNA) to second-level (build in LAMMPS) coarse-grained simulation of using nunchucks DNA with CG model.

In general, the flexibility in using longer ssDNA junctions increases with extra bending capacity until getting its flexible saturation of more than ten bases. While in the short joints case, we discovered that one base variation would have a considerable penalty change in attaining bending limits. A limited maximum bending angle prohibits two arms from folding entirely in a short joint case (notably less than six bases). Only when the length of the bases reaches six does the ‘completely folded’ behaviour occasionally become visible. In addition, nunchucks DNA with a short joint will prefer to bend into a ‘herringbone’ structure because the minimum bending angle prevents two stiff arms from folding completely.

Overall, the results of this chapter serve as useful prior information to design the higher-level coarse-grained system (e.g. to discover macroscopic properties in a dense solution). In the future, we plan to develop this bending free energy coded into exists molecular dynamics simulators (e.g. LAMMPS), so that everyone can use it to simulate a DNA-based model with ssDNA joints. Ultimately, we want to create the second-level coarse-grained model to measure macroscopic properties (such as phase transitions) in a many-body system. A feasible idea is building a bead-spring model for rigid dsDNA with coded flexible ssDNA joints with corresponding bending free energy. Mode details will be covered in Chapter 5.

4.5 Acknowledgement

I would like to thank Prof. Daan Frenkel, who introduced me this method of calculating the bending free energy; Dr. David King for mathematical supports. I also want to thank Prof. Erika Eiser for her kind and continues supervising.

5 | Towards New Liquid Crystal Phases of DNA Mesogens

In this chapter, I report on the second-level coarse-graining of the nunchucks DNA towards finding new liquid phase transitions. I did this work with Kit Gallagher as a day-to-day supervisor in his part III project. To begin with, we developed the LAMMPS code in single-stage, rigid-rod simulations to confirm that the analytical techniques could accurately recreate Onsager’s theory. Then we investigated a system of stiff rods and concentrated on the transition from the isotropic to the nematic phase. We also wrote scripts for calculating the nematic order parameter customised for our coarse-grained system, based on the work of Eppenga and Frenkel [142, 143]. In the later stage, I tuned up several methods of applying additional flexibility to using nunchucks DNA. Initially, the nunchucks model with complete flexibility was simulated. However, no uniform or system-wide director was observed, with the preferential orientation varying across the simulation. In Chapter 4, I reported that nunchuck DNA with a short joint preferred to bend into a ‘herringbone’ structure because the minimum bending angle prevents two stiff arms from folding completely. Hence, a fixed-angle approach for simplicity to get a more amenable transition to ordered phase formation was considered. Therefore, we built a system where the nunchuck DNA opening angle is confined to a particular value. We observed the angle-confined system has some degree of orientational order but can not be solved by the traditional order parameter method. A length-dependent orientational method was then applied to verify a long-ranged order parameter. We finally applied the nunchuck correlation functions introduced by Prof. Daan Frenkel and further measured the system’s dynamic behaviour. The ultimate goal is to embed the bending free energy that I concluded in Chapter 4 to simulate all possible cases of liquid crystal for nunchucks DNA with various flexibility. We look forward to building a more reasonable, realistic model over the course of time.

5.1 Introduction

After Kirkwood predicted the existence of an entropy-driven liquid to solid transition in hard-sphere systems [144], entropy-driven phase transitions have been the subject of much attention in suspensions of anisotropic colloids with hard-core interaction potential. These systems were mostly studied in theory or simulations [145–151]. Experimental and simulation studies exploring entropy-driven phase transitions primarily focused on molecular liquid-crystal former, amongst which ‘bent-core’ or ‘banana-shaped’ liquid crystals garnered particular interest because they display a rich phase behaviour including different smectic phases, some with chirality even though the mesogens are not chiral and potential applications in non-linear optics and display technology [152–154, 134, 155, 156]. However, it is more challenging to synthesise colloidal bent-core mesogens with sufficiently narrow size- and angle-distributions such that a clear liquid crystal (LC) phase transition could be observed [157]. Fernández-Rico et al.[158] introduced an exciting approach in synthesising banana-shaped colloids in which not only the size distribution of the mesogens can be kept low but where also the bending-angle can be controlled continuously. Suspensions of these systems enabled them to explore the various LC-phases directly as a function of the bending angle using confocal microscopy. These findings were further supported through variational mean-field theory and Monte Carlo simulations[155].

A new approach to designing monodisperse, bent-core systems with no polarity or attractive inter-particle potentials is the use of short, single-stranded DNA chains that can be programmed to spontaneously self-assemble into well-defined DNA-nanostars[30, 41, 46] and origami[159–164]. More recently, the significant difference in persistence length of double- and single-stranded DNA was also used to explore DNA as liquid-crystal former[165, 53, 54, 166, 167]. However, in experiments, the challenge of using DNA lies in designing thermodynamically stable DNA sequences delivering the desired LC-mesogens. Moreover, large DNA concentrations are needed to reach the isotropic to the nematic phase transition, making it very costly and time intensive to prepare many different DNA-constructs.

We also did start-up experiments to investigate the nunchuck DNA both in a dilute and semi-diluted state. However, experiment work becomes challenging if we want to observe the phase transition in a dense state. A specific dense state requires vast amounts of DNA samples, making the experiments enormously expensive. Hence running simulation work becomes particularly helpful for allowing us to study a wider parameter range without investing too many experiment samples. Here we designed nearly rigid rods, and corresponding rigid, bend core and flexible ‘nunchuck’ mesogens, followed by a coarse-graining of our computationally

expensive oxDNA model into bead-spring models that allow us to map their phase behaviour with its LC phases and local dynamics of the mesogens.

Through a greater understanding of the phase behaviour of DNA nanoparticles, we may inform the design of more complex structures and develop entropy-driven nano-structure self-assembly [168, 169]. These techniques may also have applications in biophysics, photonics, structural biology, and synthetic biology [170–172].

Computational work can be implemented over a broad spectrum of coarse-graining, a process whereby microscopic degrees of freedom are integrated over to reduce the computational costs of a simulation at the cost of lower model resolution. The base unit of the simulation can range from single atoms to large bulk molecules [173, 174]. Atomistic models, such as CHARMM [175] and AMBER [176], offer a greater level of detail, but are limited by computational cost beyond small (> 30 base pairs) models [177]. In contrast, bead-spring polymer models (with up to 3000 base pairs per bead) may obtain bulk material properties at significantly lower computational expense [178].

Such models have been applied specifically to DNA nunchucks by Salamonczyk et al. [54] to suggest the existence of smectic phases. However, such models are limited by the level of coarse-graining applied, as the artificial interaction potentials considered have no physical origin. We derive an intra-molecular potential from single nucleotide simulations of the nunchuck molecule in Chapter 4. From this, we replicate a coarse-grained model for DNA mesogens developed by Xing et al. [46], using LAMMPS software [73] to model dense systems of these nunchuck nanoparticles. This approach allows simulation of larger systems of DNA-mesogens, and provides a better predictor of the experimental phase behaviour of such systems. We also introduce a customised method of phase identification through the dynamic properties of a system, via the measurement of coordinate diffusion coefficients. While literature exists on the dynamic behaviour of mesogens within liquid crystal phases [179, 180], this has not previously been utilised to identify the nature of the static phase present, and this work represents a new avenue in the search for novel liquid crystal phases.

5.2 Simulation Model Description

5.2.1 Coarse-grained bead-spring model

We base our course-grained bead-spring model[46] of rigid rods and ‘nunchucks’ with different flexibility on the molecular structure and interaction parameters obtained from the semi-empirical, more detailed oxDNA calculations[88, 83] of DNA sequences. The

specific DNA sequences are listed in **Appendix A.2** and oxDNA modelling are discussed in Chapter 4.

The nunchuck molecules are formed of two rigid rods connected by a flexible linker, as depicted in Fig. 5.1. Each rod is modelled as a series of beads that have a diameter of 2 nm, corresponding to approximately 6 base pairs of dsDNA [181]. These beads are connected by harmonic springs, and kept linear by an angular potential with a minimum at 180° . All molecules considered here are shorter than the persistence length of double-stranded DNA, which is ~ 50 nm [182]. So the approximation of near-perfect rigidity (and linear rods) is valid.

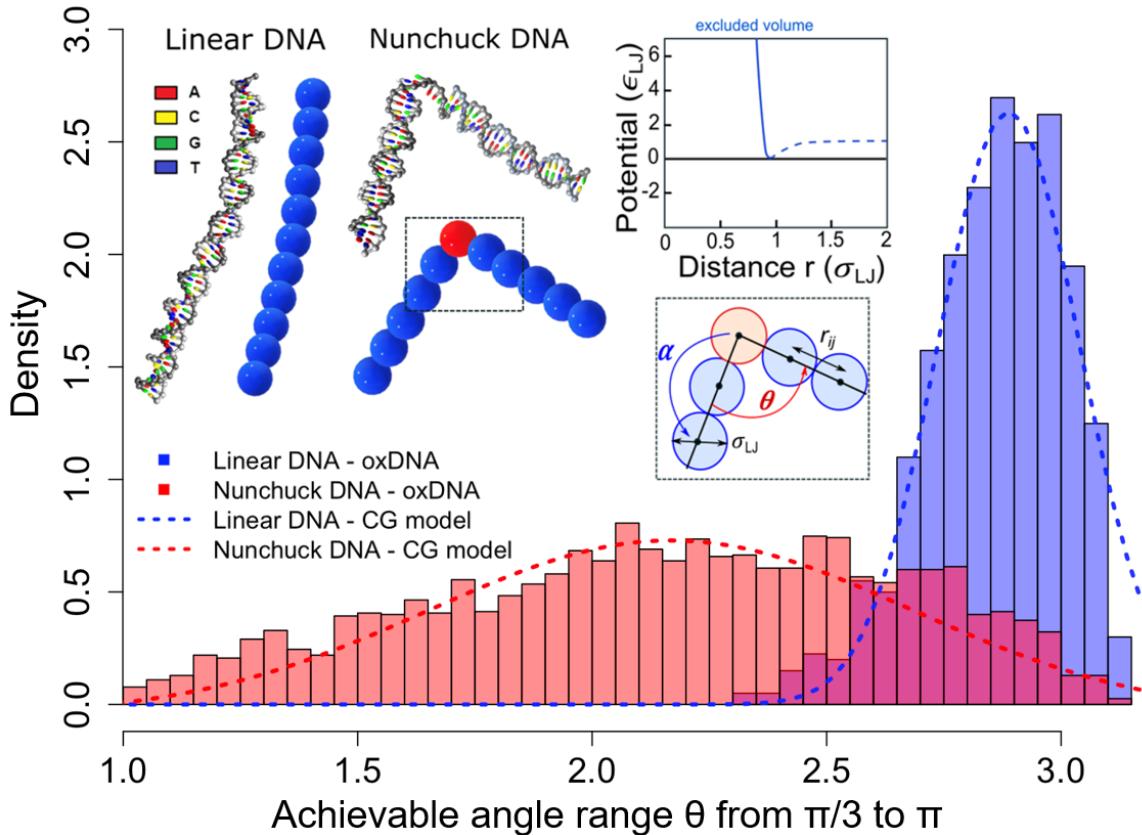


Fig. 5.1 Both schematic representations of our course-grained bead-spring model and the oxDNA calculation below the melting temperature of the DNA sequences used.

Two separate cases for the central angular potential connecting the two rods were considered, allowing for bent-core and flexible mesogens. For bent-core molecules, the minimum of the angular potential was set to $\theta = 150^\circ$ allowing for small fluctuations around θ , thus giving all mesogens approximately the same opening angle. In the flexible case, the energy scale of the angular potential is reduced while the minimum is kept at 180° , so that the opening angle of the mesogens can vary considerably. This choice was based on our oxDNA

analysis of the flexibility provided by single-stranded DNA that links the two arms in the nunchucks. We suggest that an ssDNA linker of a minimum length of six bases (roughly one bead size) provides a wide probability distribution of θ as shown in Fig. 5.1. In the same figure, we also show the angle-probability distribution for the same sequence with no ssDNA flexibility - this quasi-hard rod system was used to calibrate our simulations with regard to isotropic to nematic phase transitions as a function of aspect ratio.

The model was implemented in LAMMPS [73], using reduced units, based on the fundamental units of mass m , Lennard-Jones (LJ) energy-scale ϵ_{LJ} , LJ length-scale σ_{LJ} and the Boltzmann constant k_B . These also define a characteristic time scale $\tau = (m\sigma_{LJ}^2/\epsilon_{LJ})^{1/2}$. The LJ quantities are defined by the form of the inter-molecular interaction potential, for which a shifted, cut-off potential (potted in Fig. 5.1) is used:

$$U_{ij} = 4\epsilon_{LJ} \left[\left(\frac{\sigma_{LJ}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{LJ}}{r_{ij}} \right)^6 \right] + \epsilon_{LJ} \quad r_{ij} < r_c = 2^{1/6} \sigma_{LJ}. \quad (5.1)$$

In this Weeks–Chandler–Andersen (WCA) style potential a cut-off distance $r_c = 2^{1/6} \sigma_{LJ}$ was chosen to represent the purely repulsive excluded volume interactions between the beads, ensuring that all phase transitions observed are truly entropy–driven. Neighbouring beads are connected by the harmonic potential:

$$V_{bond} = K_{bond}(r - r_0)^2 \quad (5.2)$$

where r_0 is the equilibrium bond distance and K_{bond} is the stiffness of the harmonic bond. We set r_0 to $0.96\sigma_{LJ}$ and K_{bond} to $500\epsilon_{LJ}/\sigma_{LJ}^2$ throughout, allowing for minimal disturbances around the equilibrium distance [76, 122]. The larger value of K_{bond} indicates a DNA duplex with higher rigidity. Depending on the length of simulated DNA sequences, this value may adjust to mimic the DNA persistent length [46]. As the beads themselves have a radius $0.5r_c$ (equivalent to $0.56\sigma_{LJ}$), this gives a small overlap between neighbouring beads. The angular potential is given by:

$$V_{angle} = K_{angle}(\theta - \theta_0)^2 \quad (5.3)$$

where θ_0 is the equilibrium bond angle and K_{angle} is the spring constant that controls the allowed distribution of the angle. We set $\theta_0^{rod} = 180^\circ$ within each rod and $\theta_0^{bend} = 150^\circ$ for the bent-core molecules. We chose $K_{angle} = 500\epsilon_{LJ}/rad^2$ throughout to constrain the bending of the chains [76, 183]. These parameters have been further tested to have a tolerance from 300-500 $\epsilon_{LJ}/\sigma_{LJ}^2$ by matching the achievable bending angle distributions from oxDNA simulations in Fig. 5.1. Except for the central angle in the flexible nunchucks, K_{angle} was

given as $0.1\epsilon_{LJ}/rad^2$. This corresponds to the ssDNA, represented by a separate bead in the centre of the molecule, with differing mechanical properties.

5.2.2 Simulation details

All molecular dynamics (MD) simulations presented here were performed using Langevin dynamics. Simulations were conducted on a system of 1000 particles, with a time step of 0.005τ , unless otherwise stated. All simulations were conducted within an oblong box defined by the Cartesian axes, with periodic boundary conditions [74]. The aspect ratio of this box can be varied, to support phase formation for anisotropic mesogens and evaluate correlation functions over longer lengths without increasing the size of the system.

The system was initially configured in a dilute, isotropic state, prepared by random placement and orientation of the mesogens in a simulation box, using a Monte Carlo algorithm. Any molecules that did not fit in the simulation box or those that overlapped with the existing units were discarded, and the placement procedure was repeated until all particles were added. When studying high volume fractions, simulations were initiated from a perfectly ordered square crystalline phase, with all molecules aligned along a common axis. Care was taken to ensure the stability of this ordered phase, by confirming molecules did not overlap and that the internal energy was conserved after equilibration.

An isenthalpic ensemble was used to expand/contract the size of the simulation box, hence varying the volume fraction of the mesogen system, while a microcanonical (N, V, E) ensemble was used to thermally equilibrate the system at each volume fraction studied. Time integration was evaluated using the Nosé–Hoover thermostat[184, 185], natively implemented in LAMMPS [186], with a damping time τ .

A usual simulation run consisted of multiple stages, alternating between these two ensembles: Contraction stages varied the volume fraction and equilibration stages. Typically 2×10^4 steps (each of duration 0.005τ) were simulated in the isenthalpic ensemble, followed by 2×10^6 steps in the microcanonical ensemble, to allow the system to reach equilibrium at this volume fraction. After equilibration we sampled the system properties. To ensure the stability of the system, a Langevin thermostat [187] (also with a damping time of τ) was used throughout, maintaining a temperature of $0.5\epsilon_{LJ}/k_B$, and energy conservation was verified over a range of timescales.

5.3 Phase Characterisation

Our linear, rod-like DNA mesogens display traditional liquid-crystal phase transitions with characteristic nematic and smectic order parameters. The nematic order parameter S_n is given by:

$$S_n = \langle P_2(\cos \varphi) \rangle = \left\langle \frac{3}{2} \cos^2 \varphi - \frac{1}{2} \right\rangle \quad (5.4)$$

where P_2 denotes the second Legendre polynomial [188]. This is non-trivial to calculate when the system director is not known (i.e. in the absence of an external field). Therefore, we used the approach taken Frenkel et al. [189] - further details are provided in the **Appendix B.2**.

Similarly, the smectic order parameter is defined as

$$S_s = \frac{1}{N} \left| \sum_{j=1}^N \exp \left(\frac{2\pi}{L} i \mathbf{r}_j \cdot \hat{n} \right) \right| \quad (5.5)$$

for layers of periodicity L perpendicular to the nematic director \hat{n} ; \mathbf{r}_j denotes the centre-of-mass position of the j -the molecule [190].

It is also instructive to introduce a length-dependent orientational order parameter to verify that this order is truly long-ranged, and not simply due to short-range steric effects. We consider an l -th rank, pair-wise correlation function, where $g_l(r)$ gives the correlation between the orientation of two particles separated by a distance r :

$$g_l(r) = \frac{\langle P_l(\hat{u}_i \cdot \hat{u}_j) \delta(r_{ij} - r) \rangle}{\langle \delta(r_{ij} - r) \rangle} \quad (5.6)$$

where \hat{u}_i is the director for molecule i , and r_{ij} is the separation between a given pair of molecules i and j [191]. In the disordered phase, this pair-correlation function decays to zero, while in the ordered phase it decays to the square of the orientational order parameter [192]:

$$\lim_{r \rightarrow \infty} g_l(r) = \langle P_l \rangle^2 \quad (5.7)$$

It is worth noting here that the maximum separation between particles is half the size of the simulation region, due to the periodic boundary conditions [193], and that this will vary over the duration of the simulation. To determine order over longer length scales without increasing the volume of the system, we varied the aspect ratio of the simulation region, and sampled correlations along the long axis of the box.

The explicit calculation of $g_l(r)$ scales as $\mathcal{O}(N^2)$, and requires the storage of all necessary angles, limiting the resolution possible [194]. We shall avoid this problem by expanding the

function in terms of spherical harmonics, and sum the contributions for all pairs to a given molecule. Further details of this approach are provided in **Appendix B.3**.

5.4 Results and Discussion

5.4.1 Rigid rod order parameters

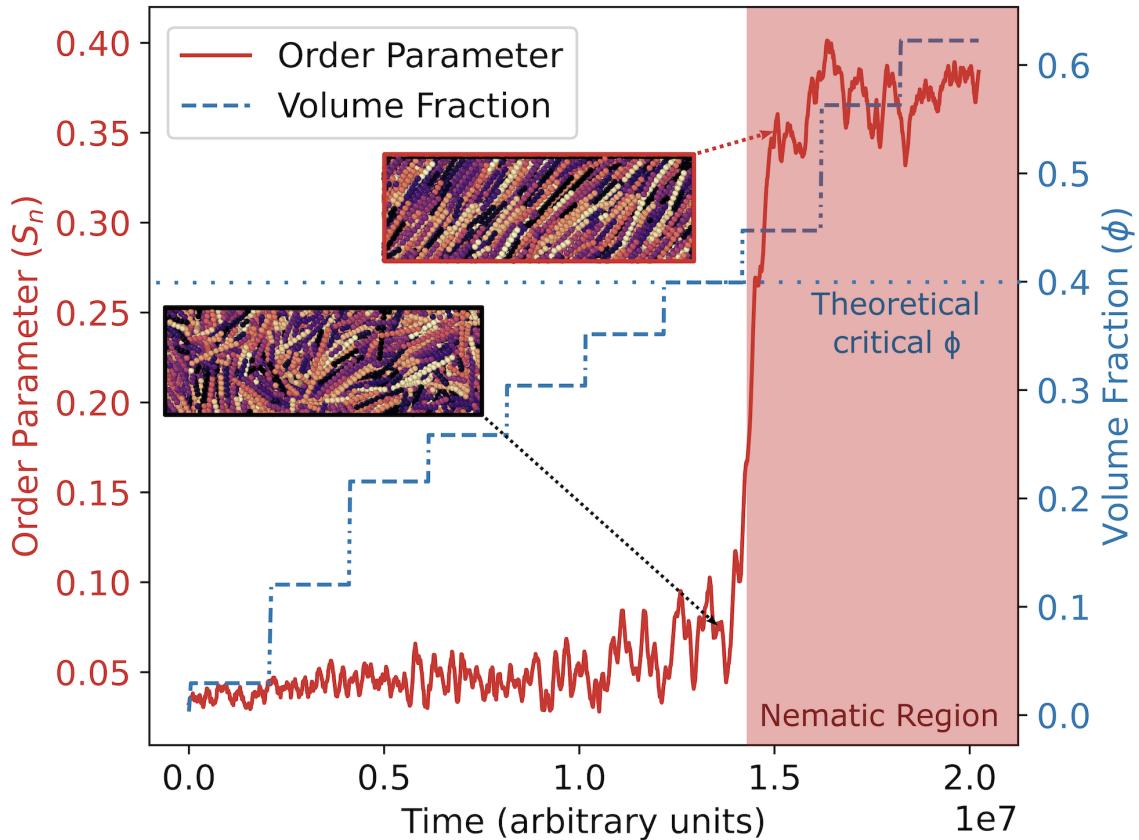


Fig. 5.2 The evolution of the volume fraction and the nematic order parameter over the timescale of the simulation, for a system of 1000 rigid rods with aspect ratio 10. The discrete change in S_n (red line) occurs at about $\phi \sim 0.4$, denoting the transition to the nematic phase, with the system depicted explicitly on either side of this transition. Note that the contraction steps, where the volume fraction is increased, are not of equal duration, and so do not correspond to equal changes in the system volume; rather they are chosen to highlight the phase transition.

A system of rigid rods was used to benchmark the analysis methods in comparison to theoretical predictions derived by Onsager [139], and in particular the isotropic–nematic

phase transition which has been verified computationally through both Monte Carlo [195, 196] and molecular dynamics simulations [197, 198], as well as experimentally [199–201].

For rigid rods with an aspect ratio $L/D = 10$, we confined the possible critical volume fraction for this lyotropic transition to the range $0.39 < \phi < 0.44$, as observed in Fig. 5.2, achieving good agreement with Onsager's prediction of $\phi = 0.4$. This analysis was also repeated for longer rods with an aspect ratio of 16, with simulations identifying the critical volume fraction in the range $0.23 < \phi < 0.26$, in good agreement with the predicted value of $\phi = 0.25$.

We also consider the phase behaviour upon expansion from a perfectly ordered, crystalline state. This has two advantages: it allows access to higher volume fractions that are not easily accessible through the shrinking procedure, and provides verification of the phase transitions previously observed. Ensuring that the system is in true equilibrium has long been a challenge for liquid crystal simulators; however, non-equilibrium effects will manifest themselves in a hysteresis of the phase transition (variation in the critical volume fraction dependent on the direction of the transition) and can, therefore, be identified through this method.

The isotropic phase formation was observed in the region $0.38 < \phi < 0.41$, in good agreement with the previous simulations, confirming the equilibrium nature of this phase transition. The higher volume fractions accessed at the start of the simulation also suggested the existence of a smectic phase, with a sharp change in the smectic order parameter around $\phi = 0.6$, however this has not been investigated further due to the wealth of literature on this already [202, 203].

Finite size effects

The lyotropic transition region might be affected by system size. To investigate the finite-size effects of the rigid-rods system, we employed an oblong simulation box with an adjustable length (X_L). The closest-packed configurations with length X_L simulation box can be maximum occupied by n layers of mesogens. The system size is related to the total number of simulated mesogens (N) and the box length (X_L). We start from system size ($N = 1000$) with final box size $X_L = 4L$ to calculate the nematic order parameter (S_n) over various volume fractions (ϕ). Then we repeat the process in a smaller system ($X_L = 2L$), which can maximum contain 500 mesogens within the simulation box. The box length suggested containing two layers of mesogens at least ($X_L \geq 2L$), where L is the length of an individual rod-like mesogen, to ensure observable layer-to-layer features in a possible smectic phase. We also investigated the larger systems with $X_L = 6L$ and $X_L = 8L$, suggesting that the finite-size effects might be less severe in larger box cases. The measured transition information of different system sizes when $X_L \in [2L, 8L]$ are plotted in Fig. 5.3. We found the transition threshold in all cases

of $L/D = 10$ similarly occurs at the same level ($\phi \sim 0.4$), details in Tab. 5.1. We concluded that the finite-size effects of the system are not affecting the observation of the transition threshold for the rigid rod system (under the investigated system sizes). The transition range is dominated by the aspect ratio (L/D) of the rigid rods.

System Size	Aspect Ratio (L/D)	Box Length (X_L)	Transition Region (ϕ)
N=500	10	2L	0.38-0.43
N=1000	10	4L	0.39-0.44
N=1500	10	6L	0.40-0.45
N=2000	10	8L	0.38-0.42
N=1000	16	4L	0.23-0.26

Table 5.1 Transition information from isotropic to nematic phase in varies system size.

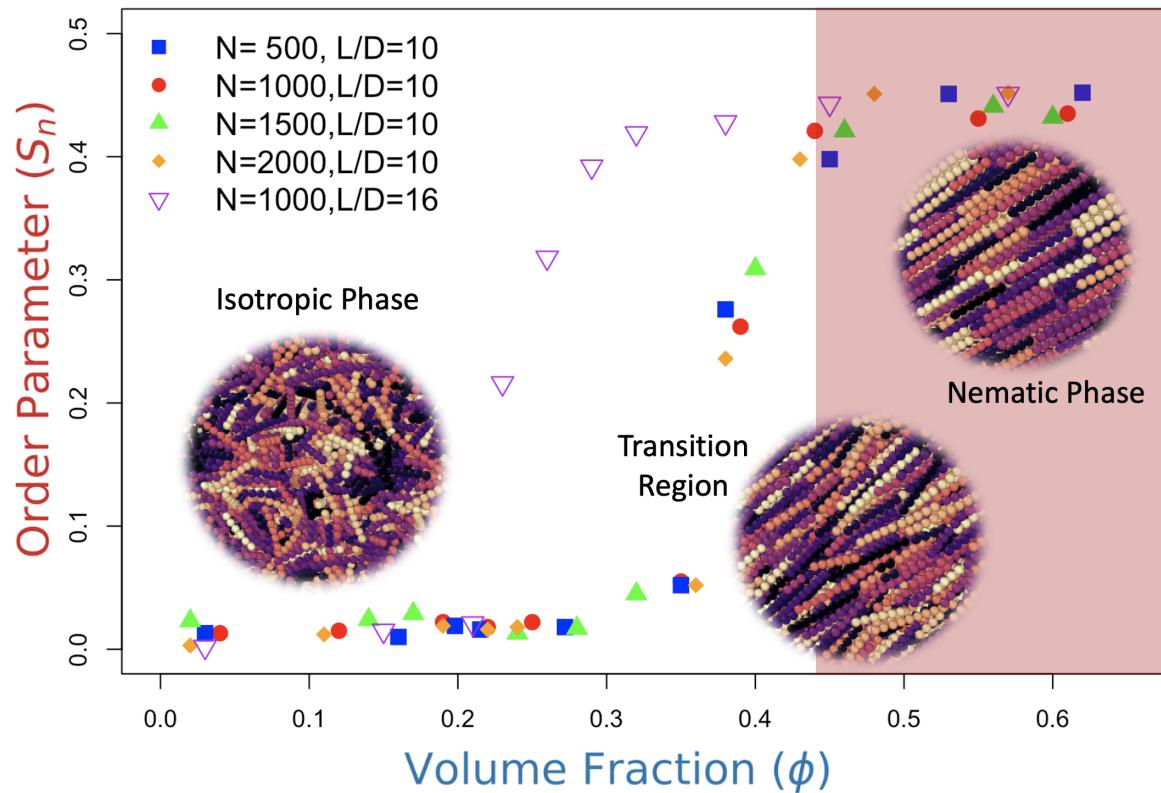


Fig. 5.3 The nematic order parameter (S_n) over the volume fraction (ϕ), with system sizes: N=500, N=1000, N=1500 and N=2000. The transition regions of different system sizes consist of $\phi \sim 0.39 - 0.44$ (for $L/D=10$) from isotropic to nematic transition. For $L/D = 16$, the transition occurs earlier when $\phi \sim 0.23 - 0.26$.

5.4.2 Nunchuck correlation functions

These techniques were subsequently applied to the nunchuck molecules introduced in Section 5.2.1. We consider first the flexible linker case, which we have verified provides an accurate coarse-grained model of the DNA-nunchucks using oxDNA [21]. The simulation molecules were lengthened to include 15 beads instead of 10, as minimal order was observed at the lower rod aspect ratios.

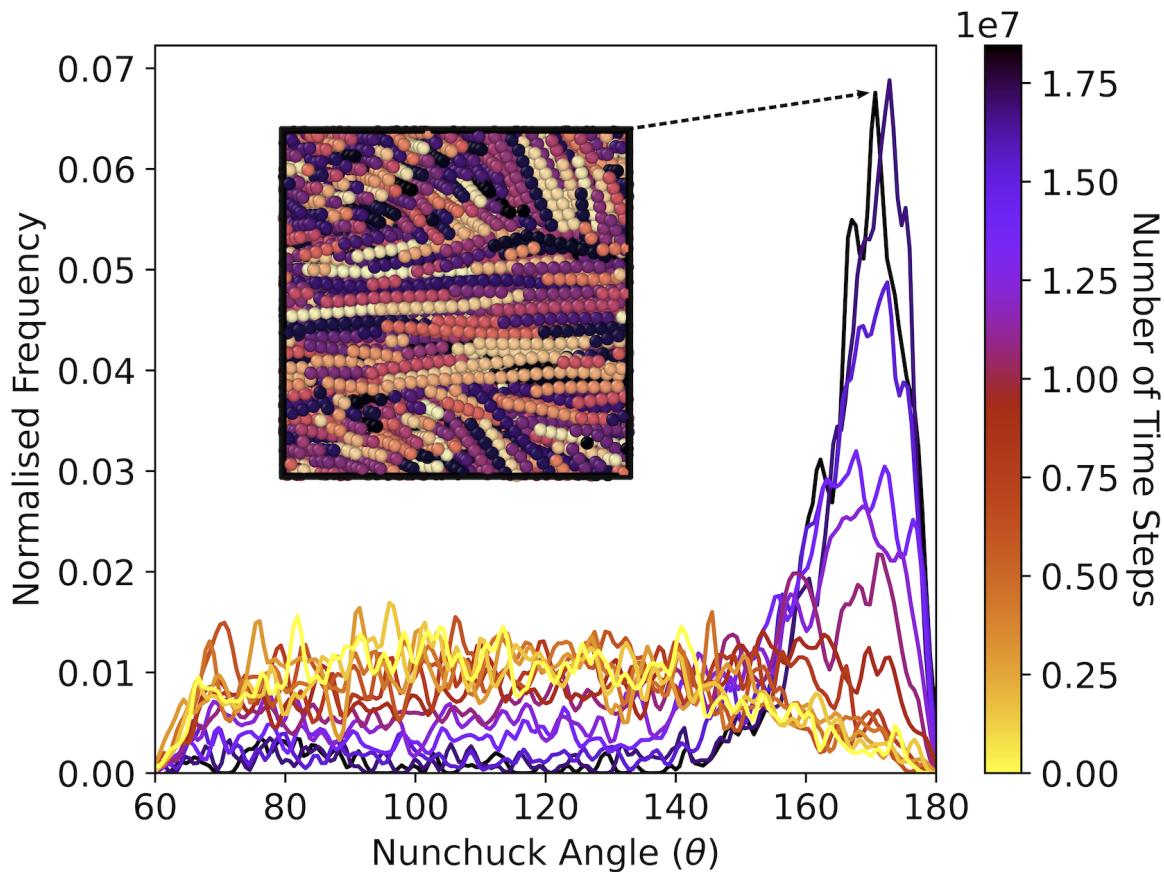


Fig. 5.4 The kernel density estimate of the opening angle distribution as the volume fraction is reduced, for completely flexible nunchucks. Plotted for a system of 1×10^3 particles, with the distribution sampled every 1.5×10^6 time steps. Note the formation of a preferential angle at late times, corresponding to the formation of an ordered phase at high volume fractions, the final form of which is depicted in the inset.

The linker flexibility was sufficient to ensure a quasi-isotropic initial angle distribution in the range $60^\circ < \theta < 180^\circ$, with smaller angles excluded due to steric repulsion between the rods. Unlike previous studies [54], this corrected the accurate possibility of the ‘fully-bent’ ($\theta = 0^\circ$) conformation where both rods are parallel, in agreement with the prior oxDNA

results. The angle distribution we observed from the detailed oxDNA model indicated the low possibility of the fully-bent situation.

Despite this flexibility, a strong preference for the linear configuration ($\theta = 180^\circ$), was observed, with a quasi-nematic phase ($S_n = 0.6$) formed at volume fractions $\phi > 0.4$. However, no uniform, system-wide director was observed, with the preferential orientation varying across the simulation region shown in Fig. 5.4. It is possible that periodic variation of the director occurs on a length scale greater than the size of the simulation region, but the computational resources required to simulate larger systems exceeded those available, preventing further characterisation of this phase.

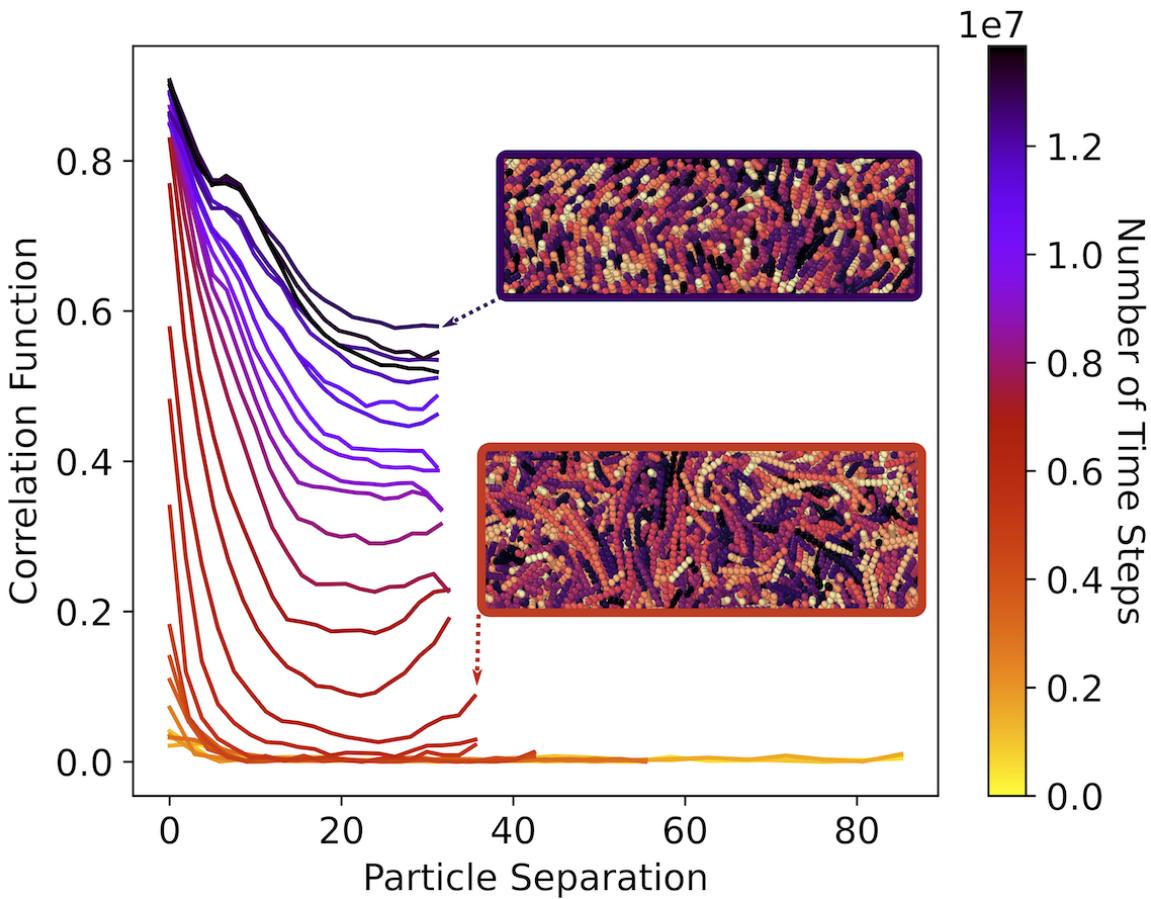


Fig. 5.5 Orientational correlation function over time (as the volume fraction is reduced), for a system of 1000 nunchuck molecules, with a fixed opening angle of 150 deg and sampled every 7×10^5 time steps. Note the formation of an ordered phase at high volume fractions (late times), with sustained long-range order (i.e. no decay in the orientational correlation function). The maximum value of particle separation is determined by the size of the simulation box, and so shrinks over time.

To force the formation of more exotic LC phases with anisotropic (non-linear) mesogens, we also considered bent core rods, with a fixed opening angle. We considered opening angles $\theta = 60^\circ, 90^\circ, 120^\circ, 150^\circ, 165^\circ$, and found that larger angles tend to form conventional nematic structures, while smaller angles are not conducive to ordered phase formation; in this section we focus on $\theta = 150^\circ$.

Testing phase formation over a wide range of rigid rod lengths and opening angle values, we found that larger opening angles favoured nematic-like phase formation (with a maximal order parameter $S_n = 0.62$), but smaller angles did not form obvious alternate ordered phases. Using a pair-wise orientational correlation function (the calculation of which is discussed further in the **Appendix B.3**), we could confirm that the order observed in this fixed angle system is indeed long-range, with a non-decaying correlation at long length scales.

This method is limited by the use of a single vector along the molecular axis to define the orientation of the molecule; two vectors are required to uniquely specify the orientation of a single nunchuck molecule. Using the molecular axis alone only accounts for quasi-nematic order and does not consider any biaxial or twisted nematic substructure. We therefore also considered additional orientational correlation functions for the bisector of each molecule, and normal vector to the plane of the nunchuck. The first Legendre polynomial is used to detect changes in sign of the bisector direction, as would be expected in a herringbone structure.

However, no statistically significant periodic components in the correlation function were observed over the length scale of the simulation region. A visual inspection of the inset in Fig. 5.5 suggests that the length of the simulation region is approximately half the full period of the repeating ‘twist’ in the ordered phase, and so further research on larger simulation regions is required to verify these phases.

5.4.3 Dynamic Behaviour

While the phase behaviour of rigid rods is well studied, much less is known about the dynamic properties of the individual mesogens within these quasi-ordered phases. Nevertheless, dynamic studies through MD simulations have enjoyed recent popularity in both computational and experimental work [204–206]. Here, we study the dynamic properties of the nunchuck system, and demonstrate the application of dynamic properties to static phase identification. In particular, we focus on the diffusion coefficient D , defined by the ‘power-law’ equation for mean-squared displacement (MSD) as a function of time t in n dimensions[207]:

$$\langle (x(t) - x_0)^2 \rangle = 2nD_\alpha t^\alpha \quad (5.8)$$

where x_0 is the reference position of the middle blob for a selected rod and $x(t)$ is its position at time t . Eqn. 5.8 reduces to the pure-diffusive case when $\alpha = 1$. Otherwise the process is characterised as subdiffusive ($\alpha < 1$) or superdiffusive ($\alpha > 1$) [208].

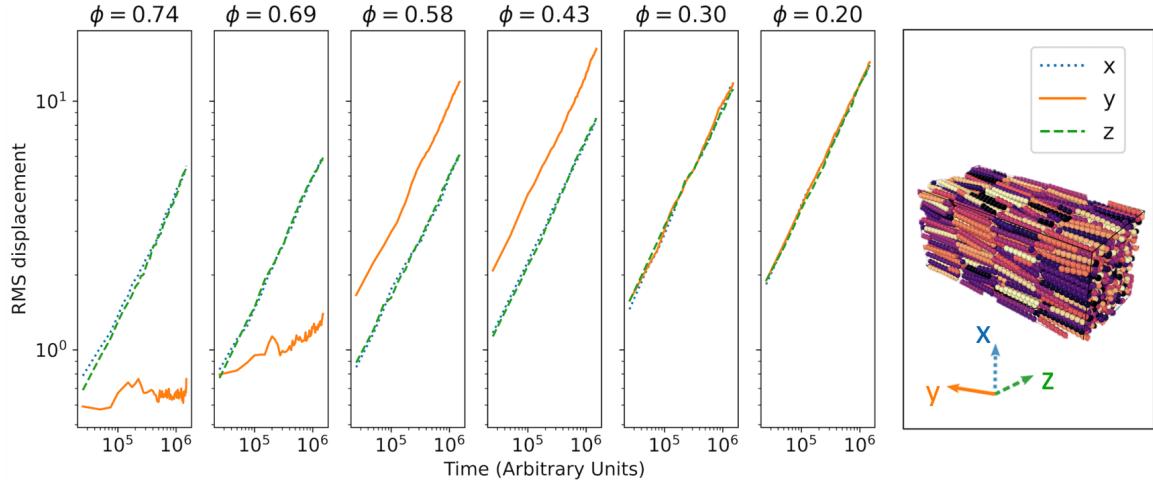


Fig. 5.6 Root mean-squared displacements against time in a system of 1000 rigid rods with an aspect ratio 10, for a range of volume fractions ϕ . The leftmost pair of plots correspond to the smectic phase, with restricted motion between $x - z$ layers along the y -axis. The centre pair give the nematic phase, with preferential displacement along the y -axis. Finally, the rightmost pair give the isotropic phase, with isotropic diffusion and no preferential direction. Note that no significant differences appear between the x - and z -directions in any phase, as these directions are equivalent in the phase structure.

Displacements were only sampled over equilibration (constant volume) stages of our simulation, and the effect of the periodic boundary conditions was explicitly accounted for by offsetting additional displacements from boundary crossings. Linear regression analysis was used to predict the value of α for each run over a variety of volume fractions in the dilute limit, where each particle may exist in a non-overlapping free volume of rotation (a sphere circumscribed around the molecule). For rigid rods with an aspect ratio of 10, this occurs at $\phi_d = 0.015$. We found an average power of $\alpha = 0.97 \pm 0.03$ for the nunchuck molecules, and $\alpha = 1.00 \pm 0.04$ for the rigid rods, as expected for diffusive behaviour.

As the volume fraction was increased, nunchucks tend towards subdiffusive behaviour, with a reduction in the value of α to 0.12 ± 0.02 at $\phi = 0.69$. We also observe anisotropy in the coordinate-wise displacement in non-isotropic, ordered phases, as demonstrated in Fig. 5.6, which considers three different phase regimes for a system of rigid rods. Here the Cartesian axes are used, as simulations are configured in a crystalline phase with molecules orientated along the y axis, and smectic layers forming in the $x - z$ plane. Phase transitions for

this system were previously identified in the regions $0.58 < \phi < 0.62$ for the smectic–nematic, and $0.38 < \phi < 0.43$ for the nematic–isotropic transition.

At the highest volume fractions ($\phi \geq 0.62$), the diffusion along the y axis is significantly reduced, due to the restricted motion between layers in the smectic phase. Below this, a nematic phase is observed with increased diffusivity along the molecule’s director, corresponding to an increased intercept in logarithmic space. The degree of dynamic anisotropy here, quantified by the ratio of diffusion coefficients D_y/D_x , is 2.12 for $\phi = 0.58$. At the lowest volume fractions ($\phi \leq 0.38$), there is no preferred direction in the system and motion is isotropic and truly diffusive (as $\alpha = 1$).

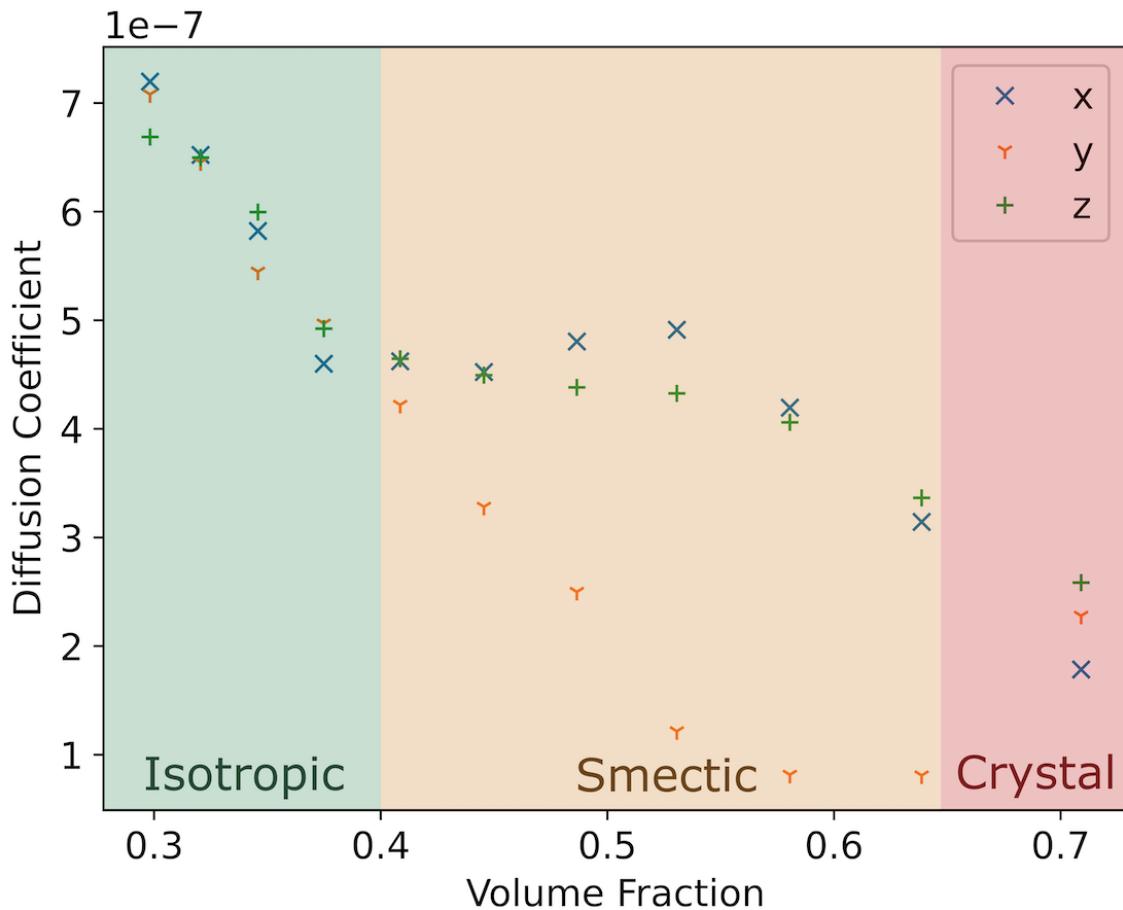


Fig. 5.7 Coordinate diffusion coefficients for the evolution of the microcanonical ensemble at a range of volume fractions ϕ . The low volume fraction region on the left of the graph corresponds to the isotropic phase, with no variation between coordinate axes. In contrast, the smectic phase in the high volume fraction gives rise to anisotropy in the coordinate diffusion coefficients, with reduced diffusivity perpendicular to the smectic layers in the y axis direction.

Phase formation may be directly observed through deviations in these coordinate-specific diffusion coefficients. Within the nunchuck system, isotropic diffusion is observed at low volume fractions in Fig. 5.7, but the y -coordinate diffusion coefficient is severely reduced beyond this, with a maximal anisotropy of $D_y/D_x = 0.19$ at $\phi = 0.58$. This strongly suggests the presence of a smectic phase below $\phi = 0.45$, in agreement with the prediction of $0.40 < \phi < 0.44$ obtained from the smectic order parameter.

Simulation of dynamic properties be utilised to detect phases that are not easily identified otherwise. Building on theoretical predictions by Camp et al. [209, 210], we determined that an opening angle $\theta = 120^\circ$ would be optimal for biaxial phase formation. We were able to demonstrate the existence of a biaxial smectic phase through the measurements of directional diffusion coefficients, as depicted in Fig. 5.8. The system was initially configured as a crystalline lattice of linear molecules, before the minima in the opening angle was shifted to form the desired opening angle. This process occurred randomly, and so no preferential bisector direction was observed. Subsequent equillibration (2×10^6 steps) allowed relaxation into the smectic biaxial state, which was then observed through repeated simulations (2×10^6 steps) to transition back into the traditional smectic and isotropic phases upon expansion.

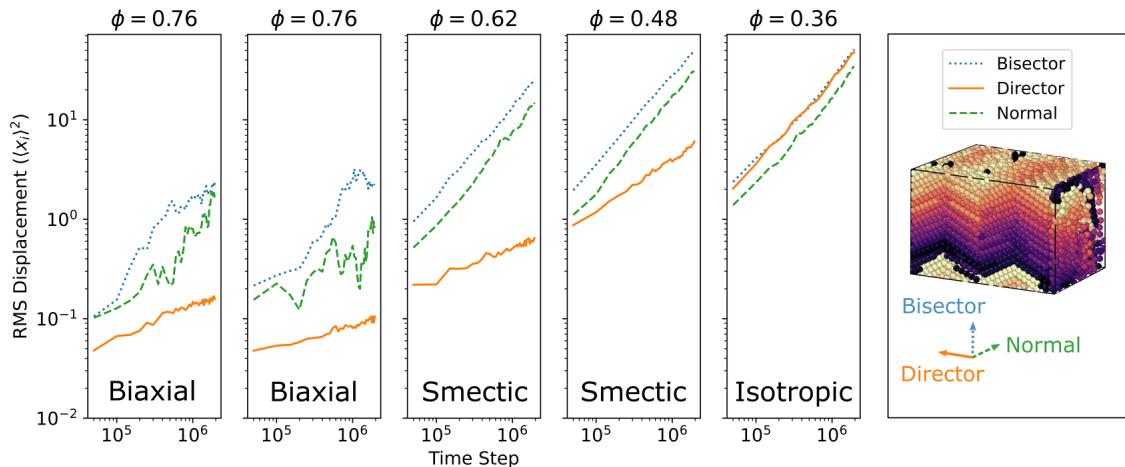


Fig. 5.8 Root mean-squared displacements against time in a system of 1000 bent-core mesogens with an opening angle 120° , over a range of volume fractions ϕ . As previously, restricted diffusion is observed along the director (previously the y -axis) between smectic layers. Anisotropy is now observed within these layer at the highest volume fraction $\phi = 0.76$, suggesting a biaxial smectic phase, with a common aligned bisector. Diffusion is reduced in the direction of the normal vector compared to the bisector vector for these mesogen.

This demonstrates the existence of a biaxial smectic phase for this system of DNA mesogens, and the power of dynamic property simulation in novel phase identification.

5.5 Conclusion

We considered the phase behaviour of DNA nunchuck molecules, consisting of two sections of dsDNA connected by a short section of ssDNA. We introduced a coarse-grained model of DNA nunchuck particles, formed of two rigid rods connected by a (semi-) flexible linker, with soft-core, purely repulsive potential interactions. We also introduced a simpler ‘rigid rod’ model, without the flexible linker, to verify the phase identification techniques used here.

Our simulations of rigid rods were consistent with Onsager’s prediction of the existence of a first-order ‘entropic’ phase transition between the isotropic and nematic phases of slender hard rods, across a wide range of aspect ratios. In comparison to Onsager’s predictions of a transition volume fraction $\phi = 0.4$ for rods with aspect ratio $L/D = 10$, we were able to confine the measured critical volume fraction within the range $0.39 < \phi < 0.44$. This value, along with the equilibrium nature of this transition, was verified through simulations prepared an initial crystalline phase and subject to a stepwise expansion of the simulation box, eliminating the possibility of phase hysteresis.

The application of the same techniques to the nunchuck system gave evidence for the formation of an ordered, quasi-nematic phase in this system, in both the semi-flexible ($S_n = 0.58$) and bent core ($S_n = 0.62$) configurations. The fixed rigidity configuration, a more realistic representation of the DNA system, demonstrated clear evidence for an entropy-driven phase transition, through the reduction in configurational entropy associated with the formation of a preferred angle. The pair-wise orientational correlation function was also used to validate the presence of truly long-range order.

Finally, we demonstrate that the measurement of dynamic properties provide a little-studied, alternative method for identifying phase transitions and characteristic symmetries of such systems, in comparison with phases previously identified in both the rigid-rod and nunchuck systems. In particular, we considered the formation of anisotropic phases through the variation in coordinate diffusion coefficients, such as in the smectic phase where the diffusion coefficient within the ordered layers is up to a factor of 5.15 greater than between them. We extend this method beyond the simple Cartesian approach taken here, to allow the identification of new phases where the global director or symmetry class is not known. Using this approach to directional diffusion within the nunchuck system, we also demonstrate the formation of a biaxial phase, not previously observed in DNA-based systems. We suggest this method can be used in further identification of novel liquid crystal phases with directional anisotropy.

5.6 Acknowledgements

I wish to thank Prof. Daan Frenkel for his kind insights and the theory support of calculating the nematic order and the orientational correlation function he introduced, detailed in **Appendix B.2** and **B.3**. I am also very grateful to Kit Gallagher for his excellent simulation work and the figures he generated in this chapter. I also thank Iria Pantazi for the initial python script to automate the particle placement in creating rigid rods in a dilute system. Finally, I wish to thank Prof. Erika Eiser for her kind and helpful supervision.

6 | DNA-Functionalised Triblock Copolymers: F108 Pluronics + DNA

In this Chapter, I propose a homemade numerical strategy for modelling our F108 DNA-functionalised triblock copolymer micelles. As introduced in Section 2.2.3, F108 belongs to the Pluronic family that typically forms spherical micelles upon heating. I built a coarse-grained simulation model with temperature-sensitive hydrophobic monomers to study the structure and phase behaviour of a symmetric bead-spring model for the PEO-PPO-PEO polymers. Further, we extended our model by adding a sticky patch to each free PEO chain, thus mimicking the complementary DNA sticky ends. This stickiness provides an additional, temperature dependent driving force for aggregation between two chain ends at low temperatures. The model builds on the abundance of existing experimental data for the Pluronics family. The numerical modelling parameters are first tuned to match our experimental phase diagram of aqueous solutions of the pure F108 chains (Fig. 2.9(b)). Subsequently, I added the DNA patches on the copolymer model and explored the phase diagram of this system by fine-tuning the DNA melting temperature of the attached DNA patches. In this way our model takes the influence of temperature and solvent quality on the self-assembly of the middle-block of F108 and the DNA-hybridization implicitly into account, thus enabling the simulation of larger systems. Hence, we obtain more accurate static equilibrium information on the structure of our system. Specifically, we computed the radial distribution function $g(r)$ of the centre of mass positions of the micelles in the micellar liquid and solid. Attaching DNA to the F108 chains change the structural properties of micellar fluids, which under specific conditions formed heterogeneous structures. To distinguish between heterogeneous and homogeneous states, the structure factor $S(q)$ was calculated. With this method, heterogeneity is easily identified by increased scattering at low wave numbers. Dr. Alessio Caciagli and Ren Liu performed the corresponding experiment work. Dr. Eivind Bering helped with writing the code to obtain $S(q)$. I completed all phase diagram and DNA melting curve simulations and analysis.

6.1 Introduction

Multiblock copolymers are creative building blocks in the repertoire of self-assembled polymeric micelles, with practical applications in nanoscience and biotechnology [211–213]. Amongst the many different polymeric surfactants are the symmetric ABA-type triblock copolymers PEO-PPO-PEO, which are known as Pluronic but have also other commercial names. Depending on their block-size ratio they self-assemble into either spherical, worm-like micelles or bilayers in solution [214, 215], similar to small molecule surfactants [216]. In these systems the driving force for self-assembly is temperature: At low temperature water is a good solvent for both the PEO and PPO blocks. However, above a critical temperature water becomes a bad solvent for the middle-block leading to the formation of micelles or other microphase-separated objects. We can add further functionality by chemically binding a group like DNA to the ends of the PEO chains, which remain water soluble up to 95 °C. This functionalization introduces an additional interaction between the micelles that can be tuned. Such ‘patchy’ micelles can thus be used in the design of multicompartiment micelles to build superstructures [217, 218]. The concept of patchy micelles has broadened the applicability of copolymer systems, since the building subunits can now be both soft or hard, or even hybrid, including multivalency [214, 219].

Pluronics® (which are also known under different commercial names such as Symperocics and Poloxamers) are an essential class of triblock copolymers that are non-toxic and water soluble, making them suitable for commercial usage, e.g. in drug delivery materials because of their ability to actively respond to changes in the surrounding environment [56]. In the following we will simply refer to these triblock copolymers as Pluronics. They self-assemble into micelles above a critical micellisation temperature (CMT) and concentration (CMC) in aqueous solutions [68, 69]. Recent work by Caciagli et al. demonstrated the possibility of attaching DNA to PEG-PPO-PEG triblock copolymers [70, 220]. Combining the sensitive thermo-reversible hybridisation of DNA with the intrinsic thermal response of Pluronics revealed a much more complex phase behaviour due to the DNA functionalization, leading to ‘patchy’ micelles [221–223].

In this chapter, I present a numerical study on the aggregation behaviour and phase diagram of the bare and DNA-functionalised Pluronic F108 dispersions. I use a new simulation strategy in which I combine a coarse-grained bead-spring simulation model to represent the triblock copolymer with a temperature-sensitive hydrophobic middle-block with a coarse-grained representation of our DNA in form of a sticky patch at the end of the PEO blocks. The temperature dependence of the interaction between the sticky patches was adjusted to mimic the specific DNA base pair we used as overhangs in experiments. However, first we

mapped the experimentally determined structure and phase behaviour to the bead-spring model for the Pluronic F108 system as function of temperature and concentration. While temperature-sensitivity has been successfully applied in copolymer systems with temperature dependent solubility [224–226], few also included the DNA hybridisation (melting) behaviour. Here we do precisely that.

Hence, we designed a simulated system of $D_{\pm}A_nB_mA_nD_{\pm}$ triblock, where the A represents hydrophilic components and the B represents hydrophobic components, and the D_{\pm} signifies complementary DNA patches. The structural properties of the system were then studied as a function of concentration and temperature computing the centre of mass (COM) radial distribution function $g(r)$ and the structure factor $S(q)$. We further translate the COM positional information into a second-level coarse-grained model to optimise the computation, in order to obtain also dynamic information from such a system.

6.2 Numerical Model

6.2.1 Model for the F108 triblock-copolymers with DNA patches

The detailed model representation of the ABA triblock copolymer with sticky patches at the ends representing the ssDNA is illustrated in Fig. 6.1. The two PEO blocks (type A - hydrophilic) are made up of 15 blobs, while the PPO block (type B - hydrophobic) contains 6 blobs. All A- and B-blobs have the same size and their number was chosen to describe the true number of EO and PO-repeat units and thus the correct block-size ratio in F108. Specifically, one blob size is roughly a Kuhn length of the real chains, which corresponds to about 10 EO- or PO-monomers. The simplest model that mimics the block-size ratio in F108 would be of the form $A_5B_2A_5$. However, such a strongly coarse-grained model neglects important details that are necessary to describe the motion of the flexible copolymer and the number of aggregation, which is the average number of chains F108 chains per micelles. Making a coarse-grained simulation as simple as possible without losing too much information is a trade-off between the necessity to make the system simple and thus simulate many chains to get good statistics and determining the right Kuhn length, which requires experience. Therefore, when done correctly, the simulations should display the formation of micelles containing many chains at concentrations similar to those observed in experiments [227, 228]. We find that the blob ratio in $A_{15}B_6A_{15}$ represent the F108 [$(PEO)_{147} - (PPO)_{56} - (PEO)_{147}$] chains in the best way. The sticky patches at the PEO-chain ends representing the ssDNA were designed and modified using a model previously developed in our group to simulate Y-shaped DNA nanostars with sticky ssDNA overhangs

[46]. We define two patch types, denoted D_+ and D_- , representing the two complementary ssDNA sequences attached to the copolymer chains. The attraction is only enabled for complementary patches, as shown in Fig. 6.1(a). The overall model was implemented in LAMMPS [73] with reduced Lennard-Jones units, in terms of the fundamental units for mass (m), distance (σ), energy (ε), and the Boltzmann constant (k_B).

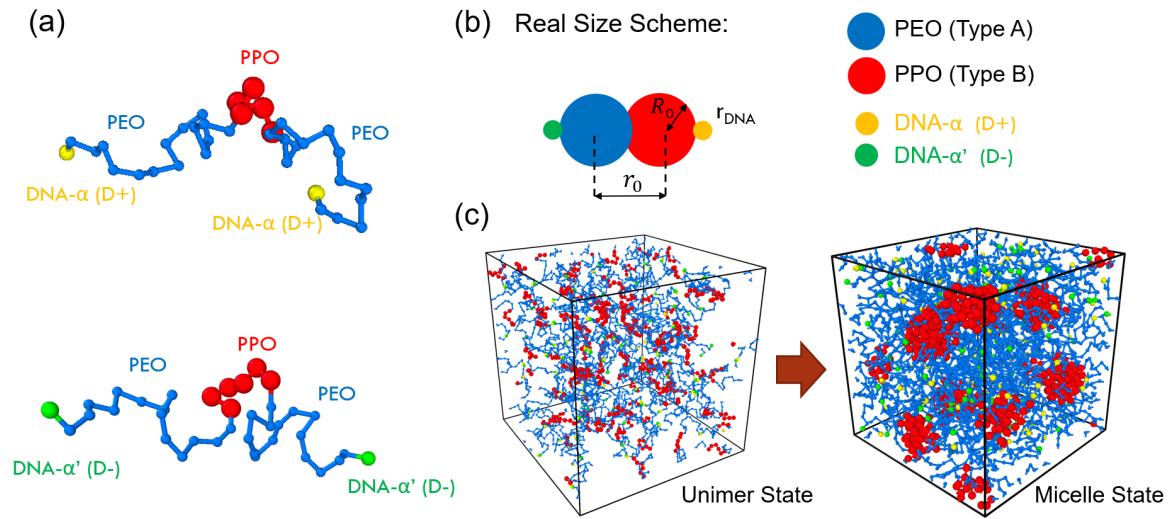


Fig. 6.1 Simulation model details: (a) Model representations ($D_{\pm}A_{15}B_6A_{15}D_{\pm}$) for F108 [$(PEO)_{147} - (PPO)_{56} - (PEO)_{147}$] with the ssDNA treated as single patches at the chain ends (D_{\pm}). The different blue and red blobs represent the PEO and PPO chains. The effective volumes of the PEO and PPO blobs are the same here. D_{\pm} represents the complementary ssDNA patches. (b) The real-size scheme of the blobs. Here the complementary ssDNA-ends are coloured in yellow and green, respectively. (c) Representative snapshots of the unimer (below CMT) and micellar fluid state (above CMT).

In our model, the solvent-mediated hydrophilic and hydrophobic properties of the blobs are described in terms of an effective, temperature dependent Lennard-Jones (LJ) potential [79, 80]. The PEO parts of the copolymer are soluble in water at all temperatures considered, and therefore the A -blobs interacted with each other through excluded volume interactions. The solubility of the PPO parts become poorer with increasing temperature. In other words the PPO chain or B -blobs become more hydrophobic with increasing temperature. The general excluded volume is provided by a shifted and truncated LJ-potential known as the Weeks-Chandler-Andersen [229] (WCA) potential, here defined as

$$U_{AA,AB}(r) = V_{WCA}(r, \varepsilon_{LJ}, \sigma) = \begin{cases} 4\varepsilon_{LJ} \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] + V', & r \leq 2R_0 \\ 0, & r > 2R_0, \end{cases} \quad (6.1)$$

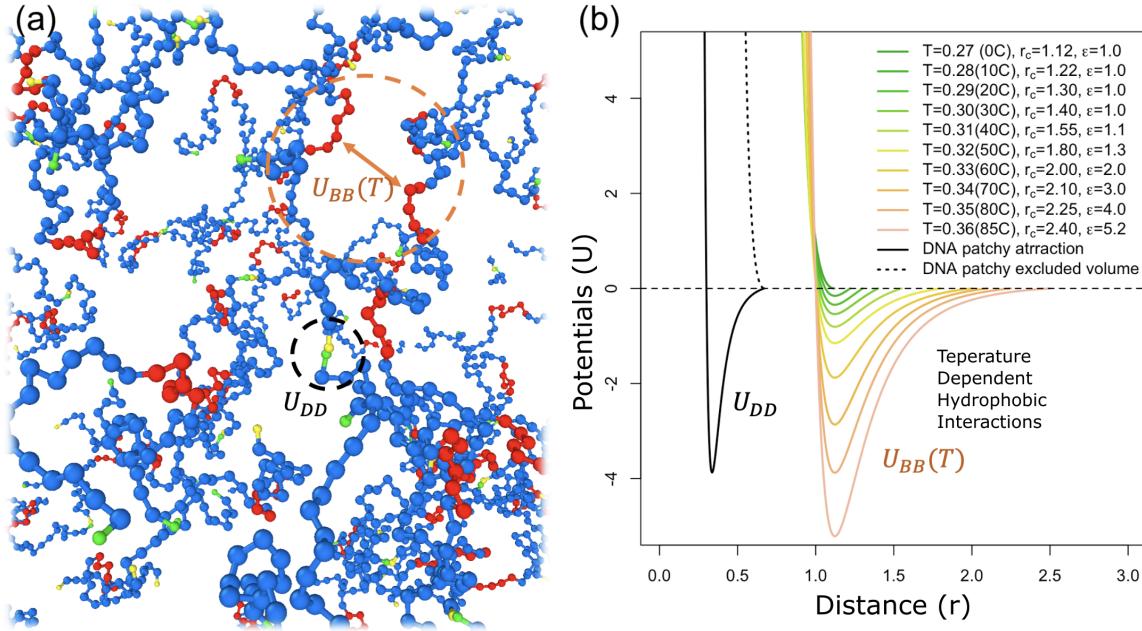


Fig. 6.2 All non-binding potentials used in the F108-DNA system. (a) Example snapshot of the temperature sensitive potential between PPO parts ($U_{BB}(T)$) and the attractive potential between ssDNA ends (U_{DD}). (b) The temperature sensitive pairwise potentials used in the model.

where R_0 is the dimensionless radius of the A - and B -blobs and V' is the shifted value of LJ-potential that represents the repulsive part. V' is set such that $V_{WCA}(r = 2R_0 = 1.12\sigma) = 0$ and $\varepsilon_{LJ} = \varepsilon$ sets the units of LJ energy scale.

The temperature sensitive part of the PPO polymer region has a temperature dependent energy term ε'_{LJ} and a corresponding cutoff radius $r_c(T)$, where $\varepsilon'_{LJ} = \alpha(T)\varepsilon_{LJ}$. The precisely tuned value is designed to match our experimental phase diagram (Fig. 2.9(b)). This temperature dependent term is introduced as a cut-off range but also given a corresponding shifted term ($V(T)''$) in Eqn. 6.2. ($V(T)''$) are set so that $V_{LJ}(r = r_c(T)) = 0$. Again, the parameters were tuned to match the experimental behaviour of F108 dispersion, and are shown in Tab. 6.1.

$$U_{BB}(r) = V_{LJ}(r, \varepsilon'_{LJ}, \sigma) = \begin{cases} 4[\alpha(T)\varepsilon_{LJ}] \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] + V(T)'', & r \leq r_c(T) \\ 0, & r > r_c(T) \end{cases} \quad (6.2)$$

The blobs are connected by elastic bonds via the Finite Extensible Nonlinear Elastic (FENE) potential [122], providing an interaction between two connected blobs with a dimensionless radius. The general FENE bond style limits the maximal extension between

T_{sim}/T	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.36
T_{exp} ($^{\circ}\text{C}$)	0	10	20	30	40	50	60	70	80	90
$r_c(T)$	1.12	1.22	1.30	1.40	1.55	1.80	2.00	2.10	2.25	2.40
$\alpha(T)\epsilon(\epsilon)$	1.0	1.0	1.0	1.0	1.1	1.3	2.0	3.0	4.0	5.2

Table 6.1 The tuned parameters of non-bonding potentials in the PPO-PPO attractions (U_{BB}). T_{sim} is the simulated temperature (1 unit of $T_{sim} = 1000\text{K}$), and T_{exp} is the corresponding temperature in the experiment.

sequential monomer pairs along each chain to r_0 , and is described by the equation

$$U_{FENE} = -0.5\epsilon_{FENE}(r_0)^2 \ln \left[1 - \left(\frac{r}{r_0} \right)^2 \right]. \quad (6.3)$$

Here ϵ_{FENE} is the bond energy scale, set to $\epsilon_{FENE} = 30\epsilon_{LJ}$. r_0 represents the distance between two blobs, and is set to 0.96σ . This is illustrated in Fig. 6.1. Adjacent blobs along the polymer chains have an overall interaction given by the sum of the LJ and FENE potentials. The FENE-bond values were set such that $(d/dr)(U_{LJ}(r) + U_{FENE}(r)) = 0$, giving a rest position for connected blobs. This deviation from the LJ rest length gives sufficient bidispersity [76]. The polymers are considered to be completely flexible, and are not constrained by any bending potential.

Following the model introduced by Xing et al. [46], the temperature dependent attraction between complementary DNA patches is expressed in Eqn. 6.4, where $\epsilon_{DNA} = \beta(T)\epsilon_{LJ}$, and the unit length of the DNA patches are set to $\sigma' = 0.2\sigma$, as these are shorter than the PEO, PPO blobs. The real size scheme shown as $r_{DNA} = 0.2R_0$ in Fig. 6.1(b). $\beta(T)$ is a term controlling the melting temperature; a higher value of $\beta(T)$ indicates a higher melting temperature of the DNA patches. The detailed set-up of $\beta(T)$ will be discussed in Section 6.3.3. V'_{DNA} is set so that $V_{DNA}(r = r_{cDNA}) = 0$, where $r_{cDNA} = 0.67\sigma$. This value of r_{cDNA} is set to prevent the formation of a possible three-way binding [46]. For instance, once a DNA pair (D-D) has successfully formed, the short-ranged WCA repulsion between PEO/PPO blobs will prevent a third DNA patch from binding. Designing smaller DNA patches also gives a relatively short-ranged attraction, mimicking the ssDNA attractions between complementary sequences. The overall non-binding potentials used in the model are illustrated in Fig. 6.2.

$$U_{DD}(r) = V_{DNA}(r, \epsilon_{DNA}, \sigma') = \begin{cases} 4[\beta(T)\epsilon_{LJ}] \left[\left(\frac{\sigma'}{r} \right)^{12} - \left(\frac{\sigma'}{r} \right)^6 \right] + V'_{DNA}, & r \leq r_{cDNA} \\ 0, & r > r_{cDNA} \end{cases} \quad (6.4)$$

6.2.2 Simulation details

The simulations are carried out with LAMMPS [73], using a periodic cube of length L with a global potential cut-off at $r_c = 5\sigma$. All blobs in this simulation model have the same mass m . The overall volume fraction ϕ of the monomeric polymers is given by

$$\phi = \frac{N^{\frac{4}{3}}\pi R^3 n}{L^3}, \quad (6.5)$$

where N is the total number of polymers and n is the number of blobs per polymer ($n = 36$ in the pure F108 model and 38 in the F108-DNA model). The value of N is adjusted depending on the purpose of the simulation. For fast testing $N = 100$, such as in the example snapshots in Fig. 6.1(c), for checking critical micellisation temperature - CMT in a small system. Choosing $N = 500$ provides better statistics of the liquid properties of the system. Therefore, most simulations determining the F108 phase diagram were performed with $N = 500$. For investigating solid-like properties, a larger system is necessary, especially when many micelles are required to display cubic crystal symmetry. N is then set to 1000.

Simulation protocol

The simulations are performed in the canonical NVT ensemble with the temperature controlled by a Langevin thermostat. The equations of motion are integrated using the Verlet algorithm with a time step $\Delta t = 0.001\tau$, where $\tau = \sqrt{m\sigma^2/\epsilon}$. To achieve a set of simulations with different densities, we performed simulations ranging from a highly diluted system ($\phi < 0.001$) to a system with high density. To force the system into a denser state, the simulation box is made smaller while keeping the total number of simulated polymers constant. This specific shrinking strategy can be achievable by applying NPT and NVT ensembles in an alternating fashion. As the shrinking process is relatively fast, it is carried out in multiple small steps where the system is brought to equilibrium after each shrinking step. To obtain a stable system with the targeted volume fraction, we set up a routine of NPT (shrinking) \Rightarrow NVT (equilibrium) \Rightarrow NPT (shrinking) \Rightarrow NVT (equilibrium). In each step of NPT, the system density keeps changing gradually. Once the system achieved a desired density, the NPT process will be stopped and changed into NVT. The stop trigger of an NPT run is when to observe a required measurement point on the phase diagram that we aim to record. The end-up density is controlled by how many steps are pre-set in the NPT run. Setting longer time steps results in a smaller simulation box (indicates a higher density). The run of NVT is relatively time-consuming as we need to first wait for the system to reach equilibrium and then start to record frames. Once further 2k frames are saved (each frame saved in

every 10k time steps), the simulation then switches to NPT and the simulation box starts shrinking again. We repeat the above process until all measurement points are covered in the phase diagram. Every single set runs at a fixed temperature. Once one set is finished, we change the temperature and restart the process. In each NVT (equilibrium) step, the system is equilibrated for 1 million time steps. The energy level of the equilibrium state is monitored by the energy minimisation routine in LAMMPS. Subsequently, the final equilibrium frames from each shrinking process were stored and were used as initial frames in later simulations where the volume fraction and temperature are fixed. This approach allows us to create initial configurations with a predetermined volume fraction, which significantly accelerates the preparation of the systems.

Backward method

The above box shrinking strategy works well to obtain the semi-dilute and semi-dense state of the simulated system. However, for a higher dense state, especially approaching the crystalline state, the system is more likely to be confined by the finite-size effects [230]. Also, a number of earlier studies of multiblock copolymers in solution using MD reported that predicting crystalline structures is challenging [226, 225]. The approach to thermodynamic equilibrium in these systems is known to be sluggish experimentally, with time scales of the order of minutes. As a result, even with appropriately coarse-grained models, MD simulation faces a significant barrier due to the extensive time scales involved. Hence, when simulating a high-density state, we use an ‘inverse’ process in which we prepare pre-packed crystal structures as the initial configuration of the system. Two different lattice systems within the close-packed family were investigated: The cubic close packing (specifically Face-Centred Cubic or ‘FCC’) in which the hexagonally close-packed layers are alternated in an ABCABC... sequence and the Hexagonal Close Packing (‘HCP’) structure, where the close-packed layers are alternating in an ABAB... pattern. These packings have the highest density amongst all possible lattice systems, with the highest packing fraction (≈ 0.74) occupied by spheres [68]. As we know, the Pluronic F108 was found experimentally to display FCC packing in the crystalline state [231]. Therefore, when we start to simulate the F108 crystal state, we prepare the system in an FCC packing. From this configuration, the routine of alternating NPT/NVT ensembles is applied in the opposite direction, NPT(relaxing) \Rightarrow NVT(equilibrium), where the relaxation of the simulation box induces a dilution step, which is then equilibrated. In the high density case, in each step of running NVT (equilibrium), the system is equilibrated for 5 million time steps, as the concentrated regime equilibrates much more slowly.

Equilibration procedure

The major challenge faced in using MD to determine equilibrium states is the simulation must last longer than any of the relaxation times in the system. To ensure the equilibration, we monitor the time evolution of the configuration energy and the number of micelles during the simulations. The total number of micelles may change before equilibrium, but eventually, it will stay at a stable value. In LAMMPS, there is a *log* command to monitor the real-time energy of the system [73]. Each time when the simulation box was shrunk/expanded, we tracked the total energy of the system until it was maintained at a stable level. In the meantime, the number of micelles was counted and recorded every 10k steps to ensure the stability of the micellisation process. Although the entire process may need extra relaxation time, we can automate the whole procedure without repeating the design of the entire system. After the equilibration period, the configuration of the system is recorded every 10k time steps as it runs for an additional 3 million steps. By applying these forward and backward processes, we can obtain any volume fraction as an initial configuration. These are then continued in the NVT ensemble to produce a significant amount of analysable equilibrium data.

6.3 Results and Discussion

6.3.1 Determining the micellisation phase boundary

The micellisation of block copolymers in solutions is inherently complex as the CMT increase with decreasing concentration [232, 233, 228]. Even above the CMT, unimers, micelles and micellar aggregates coexist, and the relative proportion of each species depends strongly on the temperature and polymer concentration [234]. The CMC and CMT can be determined experimentally by fluorescence spectroscopy, surface tension, solubilisation methods, light scattering and rheology [235, 236, 65, 237]. In simulations, the exact aggregation number of each micelle is known; hence a more direct method can be applied to investigate the unimer/micellar status in an equilibration period to determine the micellisation phase boundary.

In general, the number of micelles rise with rising temperature over the CMT for a given concentration until all unimers are transformed into micelles. To identify the micellisation process of the F108 system, we assume that two individual chains belong to the same micelle if the PPO segments (hydrophobic - blobs type *B*) on one chain is adjacent to another PPO segment on the other chain. We define N_{agg} as the aggregation number of a given micelle.

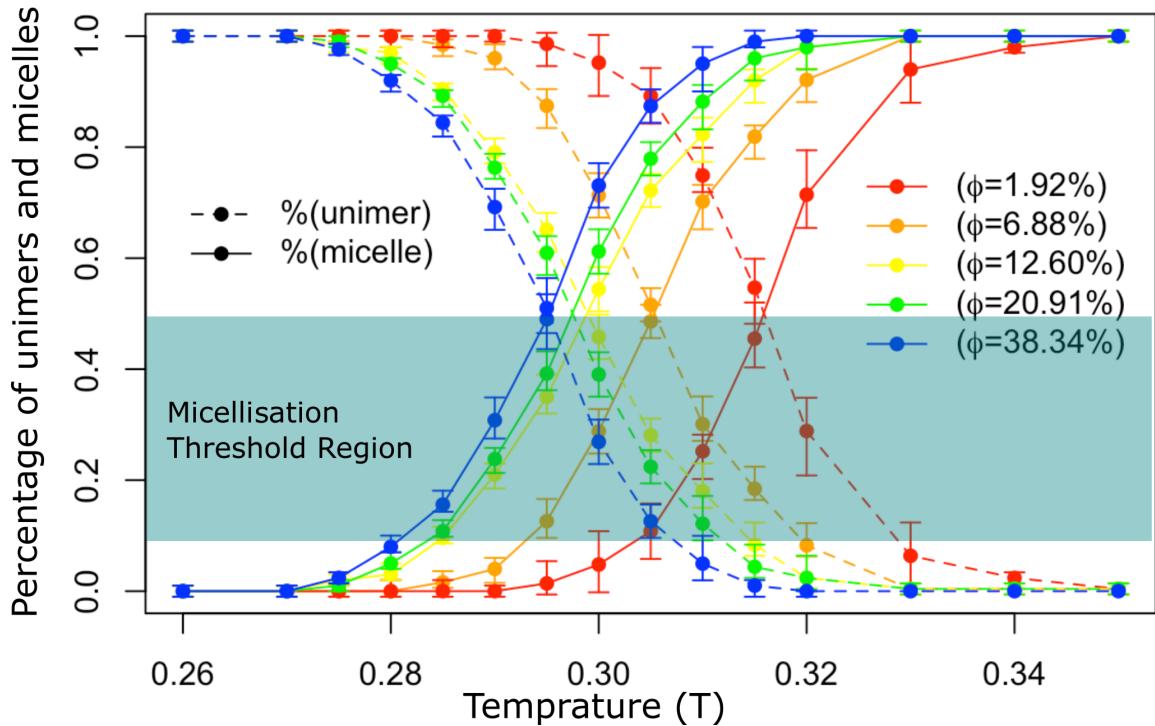


Fig. 6.3 Here the normalised probability densities $P(\text{micelle})$ and $P(\text{unimer})$ are plotted as function of normalised temperature for different Pluronic volume fractions ϕ , in order to determine the unimer to micellar liquid phase boundary. The ‘dark-cyan band’ represents the micellisation threshold region. All measurement points were taken from the simulations under equilibrium conditions with $N = 500$ chains in the system. The error bars were acquired from ten repeated runs for each temperature setup.

Further, the normalised densities $P(\text{micelle})$ and $P(\text{unimer})$ describe the overall state of the system, where

$$P(\text{unimer}) + P(\text{micelle}) = 1. \quad (6.6)$$

$P(\text{unimer})$ is then the probability of finding a polymer in a unimer state, when detecting $N_{\text{agg}} < 2$, and $P(\text{micelle})$ is the probability of finding a polymer in a micellar state, when detecting $N_{\text{agg}} \geq 2$. The detection method implemented here uses the centre of mass (COM) distance ($R_{[\text{PPO},\text{COM}]}$) between each copolymer’s PPO segments. If $R_{[\text{PPO},\text{COM}]} < 3\sigma$ for 50k time steps in an equilibration period, the two polymers are considered to be aggregated. The value of 3σ is chosen because the model has 6 PPO blobs in the middle, and the longest distance between the two nearest COM of PPO segments in the aggregated state should be around 3σ . The advantage of using $P(\text{micelle})$ and $P(\text{unimer})$ instead of directly counting N_{agg} to determine the micellisation process has the following reasons: Firstly, the

micellisation process has a dynamic equilibrium between unimers and micelles with free copolymers constantly exchanging with the copolymers in the micelles [69]. In this case, the number N_{agg} is varying with time, and it can not accurately separate the unimer/micellar states, especially near the CMT when many unimers and micelles coexist. Secondly, $\langle N_{agg} \rangle$ strongly relies on the temperature and concentration, making it hard to set up a single cut-off $\langle N_{agg} \rangle$ to determine when the system enters the micellar phase.

T	0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34
$\langle N_{agg} \rangle$	0	0	3.75	8.89	12.19	13.38	14.00	14.14	14.29
$P(micelle)$	0	0	0.09	0.32	0.72	0.91	0.98	0.99	1
$P(unimer)$	1	1	0.91	0.68	0.28	0.09	0.02	0.01	0
Phase	U	U	U	U → M	M	M	M	M	M

Table 6.2 The collected $P(micelle)$ and $P(unimer)$ values at $\phi=0.38$ (blue curve in Fig. 6.3). $\langle N_{agg} \rangle$ is the averaged polymer number in micelles. The phase stage symbol ‘U’ represents the ‘unimer liquid’; M means ‘micellar liquid’.

Ideally, the micellisation phase boundary (from unimer to micellar liquid) can be determined by defining a suitable threshold as a ‘starting point’ of the micellisation. However, it is rather challenging to accurately describe the ‘starting point’ of a micellisation in both simulations and experiments. Generally speaking, the starting point of the micellisation is when the first micelle is successfully assembled. However, in the early stage of forming the micelle, it can not be defined as a ‘stable’ micelle. For instance, we evenly count two aggregated polymers as a micelle, although these two adjacent polymers fall apart in the next frame. And then, one of the parted polymers might form a newly counted micelle with another polymer. To avoid this ‘fake’ counting and variation of the aggregate number, it is more useful to consider $P(micelle)$ and $P(unimer)$, which is illustrated in the diagram in Fig. 6.3 to help visualise the overall transition process. This diagram displays the $P(micelle)$ and $P(unimer)$ probability densities obtained from the equilibration period in a set of temperatures and concentrations. At low temperatures, all polymers were separated. Therefore, 100% of them are in the unimer state. Once the temperature increases, more and more polymers form micelles. To separate unimer/micellar phases, we introduce a micellisation threshold region ($[P_t] = P(micelle) \in [10\%, 50\%]$). If the percentage of finding a polymer in a micelle is higher than the threshold region ($P(micelle) > [P_t]$), then we define the system to have reached a micellar phase; otherwise, the system remains in a unimer phase - between them is the transition region. As we know in Pluronic suspensions, the unimer-to-micelle transition region is relatively broad [228]. Hence we use a threshold range instead of a single cut-off threshold value to determine the phase boundary. This region is illustrated as a ‘dark-cyan band’ in Fig. 6.3. The dark-cyan band also implies that the micellisation transition of the

copolymer system is a gradual process rather than a sudden change. This findings do suggest that our simulations reflect the main features of the experimentally measured F108 system. Five distinct volume fractions (ϕ_P) were tested in Fig. 6.3 arrange from 1.92% to 38.34% ($\phi \in [0.02, 0.38]$). Higher density cases are not plotted as they are close to the $\phi = 0.38$ case. The measured values for $\phi = 0.38$ is listed in Tab. 6.2 with temperatures ranging from 0.26 to 0.34. This micellisation threshold region makes it possible to directly draw the micellisation phase boundary, which is indicated in the F108 phase diagram in Fig. 6.4.

6.3.2 F108 phase diagram

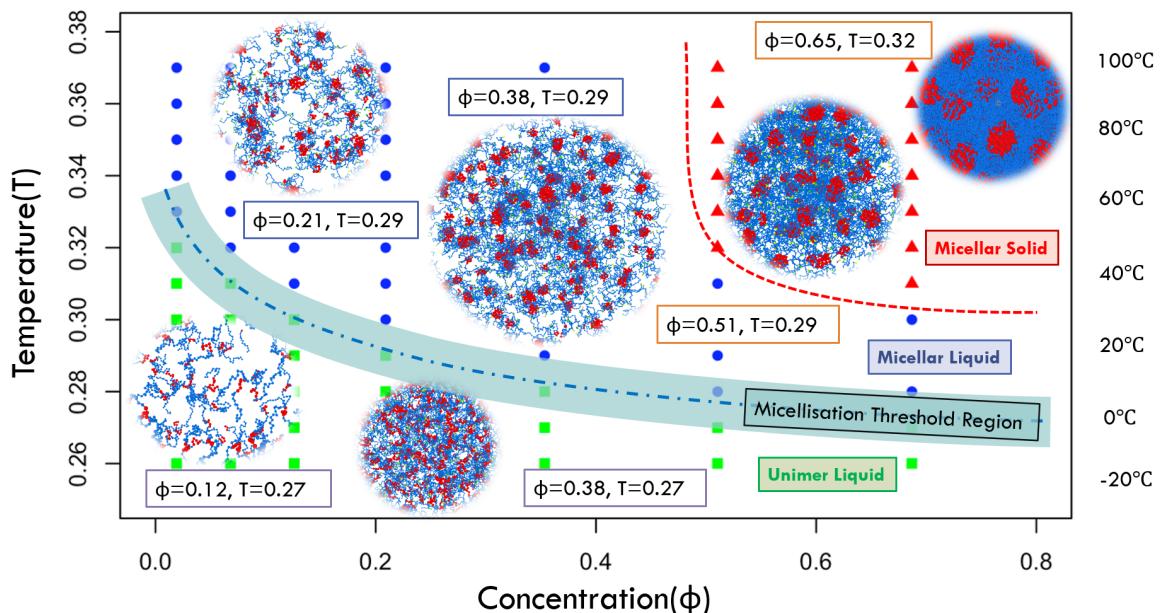


Fig. 6.4 Numerical phase diagram of Pluronic F108 dispersions: Green dots represent the unimer liquid phase, blue the micellar liquid region, and red dots the micellar solid phase. The micellisation boundary ('dark-cyan band') was obtained from the threshold region explained in Fig. 6.3. The red dotted line was estimated as a liquid-to-solid boundary extracted from the dilution process when starting from the micellar close-packed FCC structure.

Before activating the DNA patches, we tested the pure F108 Pluronic system with all U_{DD} potentials turned off. The simulated results of the F108 phase diagram is shown in Fig. 6.4. Embedded snapshots were generated with Ovito [238] in the corresponding equilibrated set of temperatures and concentrations. The F108 phase diagram uses the re-scaled temperature value (the left side y-axis is the T -value in simulation and the right side y-axis indicates the corresponding temperature in experiments). The x-axis is the volume fraction (ϕ) of the system, where the dots are the values of $\phi = 0.02, 0.07, 0.12, 0.21, 0.38, 0.56, 0.65$ we

simulated. We applied both forward and backward simulations to draw the overall phase diagram. Volume fractions of $\phi = 0.02 \rightarrow 0.38$ were obtained by the forward simulations, using the box shrinking process. Between $\phi = 0.65$ and 0.38 the backward process was applied, starting from pre-packed FCC structures and then relaxing the box. As expected, we identified three phases in the pure F108 system: (1) The unimer liquid, (2) micellar liquid and (3) the micellar solid. The micellisation boundary ('dark-cyan band') to separate unimer and micellar liquid was given by monitoring $P(\text{micelle})$ and $P(\text{unimer})$, shown in Fig. 6.3. The next step was to determine the micellar liquid-to-solid boundary (red dotted line in Fig. 6.4). This was a challenging step as it was hard to distinguish between the disordered and ordered micellar phases. One method to distinguish the two phases is monitoring structural properties such as $g(r)$, as ordered solids give rise to sharper peaks. At high polymer concentrations and temperatures, almost all unimers are aggregated into micelles of similar aggregation numbers and thus of similar size (especially for higher densities). Therefore, we can compute a COM-based $g(r)$, treating the micelles as spheres with a water-free solid core. The micellar liquid-to-solid boundary was determined by monitoring the backward process and measuring $g(r)$ at every volume fraction considered. In this case, we started the simulation from the packed FCC structure, and slowly relaxed the simulation box, running it until an equilibrium state was reached. We always found an FCC packing for $\phi > 0.6$. However, when further relaxing the system to $\phi < 0.6$, the packed structure began losing the ordered property during the equilibration period. From this we determine the solid-to-liquid phase transition boundary to be around $\phi = 0.6$. Although the gelation boundary is roughly estimated, it is reasonable to separate the two phases as we also noted that industrial products like Pluronics exhibit significant batch-to-batch variability in the liquid-to-solid transition boundaries [228]. The detailed calculation of the COM-based $g(r)$ measurement to help separate this regions is discussed in Section 6.3.4. As a result, the simulated F108 phase diagram matches well with our experimental system, finding the same phases (Fig. 2.9(b)). Note that there are obviously also two-phase regions in which the micellar liquid and solid phase coexist. However, these were found to exist within very narrow volume fraction changes, both in experiments and simulations.

6.3.3 Hybridization behaviour of our complementary DNA overhangs

From the F108 phase diagram, we see how the CMT varies with the polymer concentration. The overall CMT is decreasing when the concentration is increasing, e.g. $\text{CMT}_{(\phi=0.12)} \approx 0.31$ and $\text{CMT}_{(\phi=0.38)} \approx 0.29$. In experiments we chose the melting temperature of the sticky ssDNA to be within that region. For this reason, it is crucial to get the melting temperature of DNA (T_m) correct, as it can affect the overall phase diagram of the system at different levels.

In this section, the T_m is tuned to fall between the CMT for the semi-dilute and dilute state, for the purpose of obtaining more exotic phase behaviours. To map the melting/hybridization behaviour of the complementary DNA stand we used in experiment, onto our simulations, in particular, to get T_m right, we define the degree of association θ to evaluate the connectivity of the DNA patches:

$$\theta = \frac{N_P}{(Q_{valence}N)/2}. \quad (6.7)$$

Here N_P is the number of connected patchy pairs, N is the total number of copolymers, and $Q_{valence}$ is the building-block valency, which is 2 in our model by construction. As the denominator represents the maximum number of connected patchy pairs for a system of N polymers, θ varies from 0 at high temperatures when all DNA patches are separated, to ≤ 1 at low temperatures when all possible DNA bonds are formed. Our method to determine the T_m of a given DNA duplex was inspired by our previous simulation work on DNA hydrogels with sticky ends [46]. In Fig. 6.5(a) we show the computed values of N_P from time-averages taken over the steady-state period. In Fig. 6.5(b) we plot θ as function of temperature, which is also known as the melting curve. A hysteresis test was run to confirm that the systems stay in equilibrium. Such a slow cooling or heating cycle is similar to hybridisation cycles probed by standard UV-vis spectroscopy measurements, which are used to determine the melting behaviour of a given DNA duplex [46]. In the simulations, this melting temperature can be tuned by controlling the value of $\beta(T)$ in Eqn. 6.4. A larger value of $\beta(T)$ results in a higher T_m , in this specific test we used $\beta(T) = 3.2$. The estimated melting temperature is $T_m \approx 0.295$, as shown in Fig. 6.5 .

After activating the DNA patches with T_m around 0.295, two different scenarios are expected to occur. Firstly, we expected to observe unimers linked by DNA patches for $T < \text{CMT}$ and $T < T_m$. In this case, long concatenated chains form, and polymer rings may occur in a dilute state. In the second scenario we expected to see micelles linked by DNA; this would only occur when $\text{CMT} < T < T_m$. In this case the mobility of the micelles must be strongly reduced compared to the diffusivity of the free micelles in the fluid state. Or in other words, we expected to observe a micellar gel rather than a micellar fluid phase. To further investigate these different concentration and temperature regions, we used our COM-based $g(r)$ calculations to determine the structural properties of our systems before looking at the structure factor $S(q)$, which will be discussed in the following section.

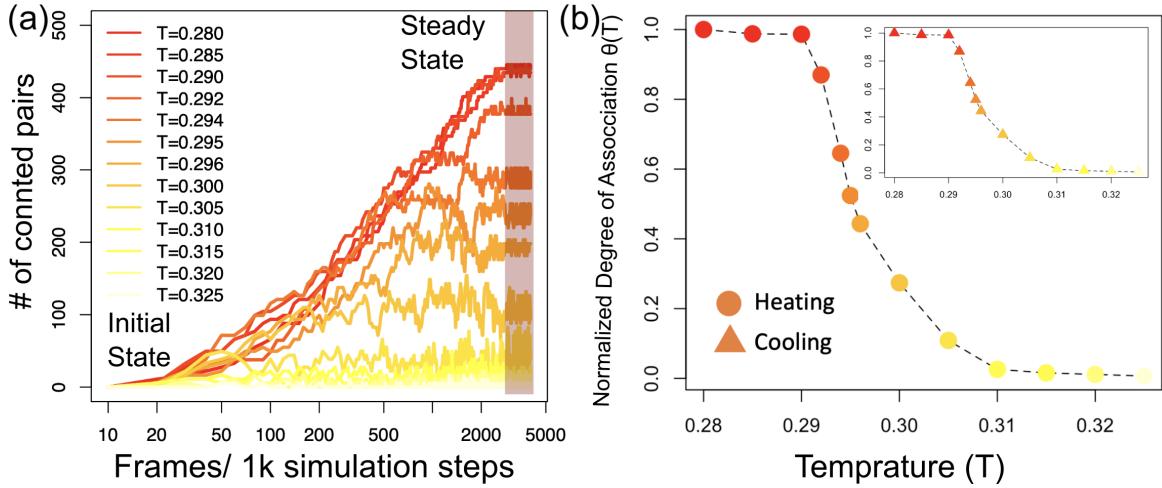


Fig. 6.5 (a) Equilibration process quantified by the number of connected pairs at temperatures between 0.28 and 0.32. The system starts at a randomly generated non-binding initial state and eventually reaches its steady state, where the number of connected pairs reaches a plateau (marked as ‘steady state’). (b) Degree of association $\theta(T)$ calculated from the averaged number of connected pairs in steady state (for both cooling-down and heating-up ramps). Data was measured from a system with $\phi = 0.38$ and a total number of copolymers $N = 500$.

6.3.4 Structural properties - determining $g(r)$

The radial distribution function $g(r)$, based on the centre of mass of the PPO cores in each micelle, is given by

$$g(r) = \frac{1}{4\pi r^2 \rho N} \left\langle \sum_{i=1}^N \sum_{j \neq i}^N \delta(r - |r_{ij}|) \right\rangle . \quad (6.8)$$

Here, N is the total number of micelles and ρ is the averaged number density of the micelles across the entire simulation box. As periodic boundary conditions are applied, these are taken into account in the calculations. The COM position of core of the micelles is calculated by:

$$r_{i/j} = \frac{1}{M} \sum_{k=1}^n m_k r_k , \quad (6.9)$$

where $M = \sum_{k=1}^n m_k$ is the total mass of the n aggregated PPO cores (coloured in red in Fig. 6.6(a)). We average this quantity over many equilibrium configurations. Further, $\rho = N/V$, where V is the volume of the simulation box. If we estimate $g(r)$ by accumulating a histogram of the number of particles at distance r , within bin-width δr , we get

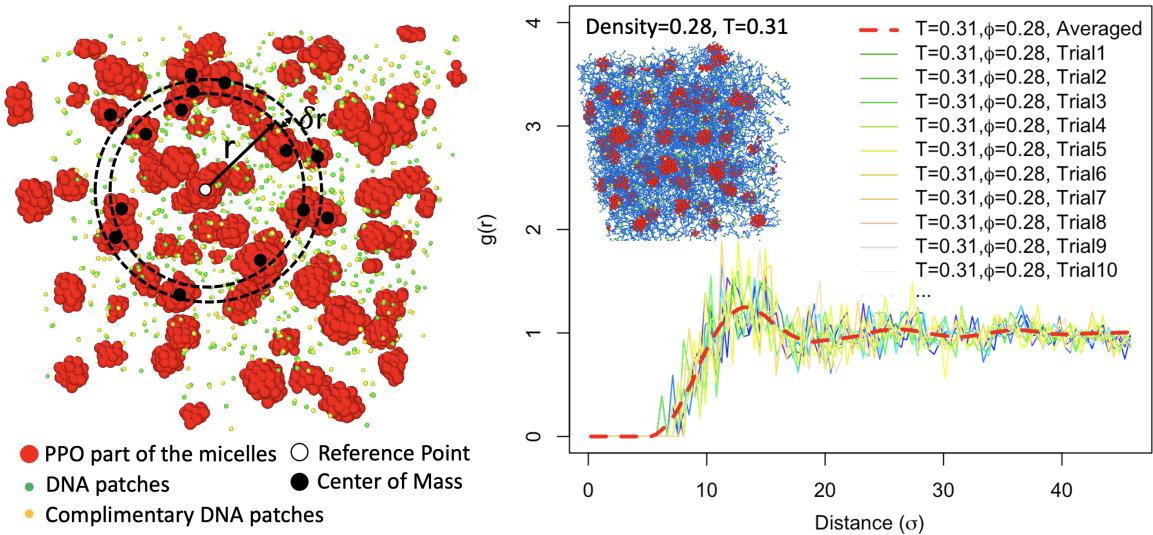


Fig. 6.6 (a) Illustration of the calculation of the radial distribution functions $g(r)$. The PPO cores of each micelle are coloured red with the centre of mass depicted as black dots. The white dot represents the centre of mass of a reference micelle, from which the relative distances are calculated in shells of size δr , which represents the resolution term of the calculation. (b) The calculated radial distribution function for F108 micelles with complementary DNA patches at $\phi = 0.28$, $T = 0.31$ (above T_m). Each trial originate from 3 million time steps sampled every 10k time steps. The red dashed line shows the $g(r)$ averaged over all the trials.

$$g(r, \delta r) \approx \frac{V}{4\pi r^2 \delta r N^2} \left\langle \sum_{i=1}^N \sum_{j \neq i}^N H(r - |r_{ij}|) H(r + \delta r - |r_{ij}|) \right\rangle \quad (6.10)$$

where the Heaviside function H is defined as

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \quad (6.11)$$

The bin width (δr) is shown in Fig. 6.6(a). δr should be chosen smaller than the size of the particles. Note that we treat the entire micelle as one single particle, implying that the size of the micelles is of the order of several σ . In general, decreasing δr would increase the resolution with which $g(r)$ is computed, but the smaller the bin-width, the smaller the number of entries per bin, and hence the larger the relative statistical error. As the number of entries per bin of the histogram is proportional to the computing time, increasing the resolution of $g(r)$ at constant statistical error will require longer simulation times. Hence, here we set $\delta r = 0.5\sigma$, which provides a sufficiently high resolution within a reasonable computational

time. Fig. 6.6(b) shows the resulting $g(r)$ from equilibrium simulations together with the 10 trial runs this $g(r)$ was averaged over. The density of this particular system is $\phi \approx 0.28$, corresponding to a semi-dense system. The first peak is prominent, and is located at the centre-to-centre distance between two micelles in close contact, $r \approx 14\sigma$. This gives an indication of the size of the formed micelles. The width of the peak gives an indication of the variation in the micelle sizes.

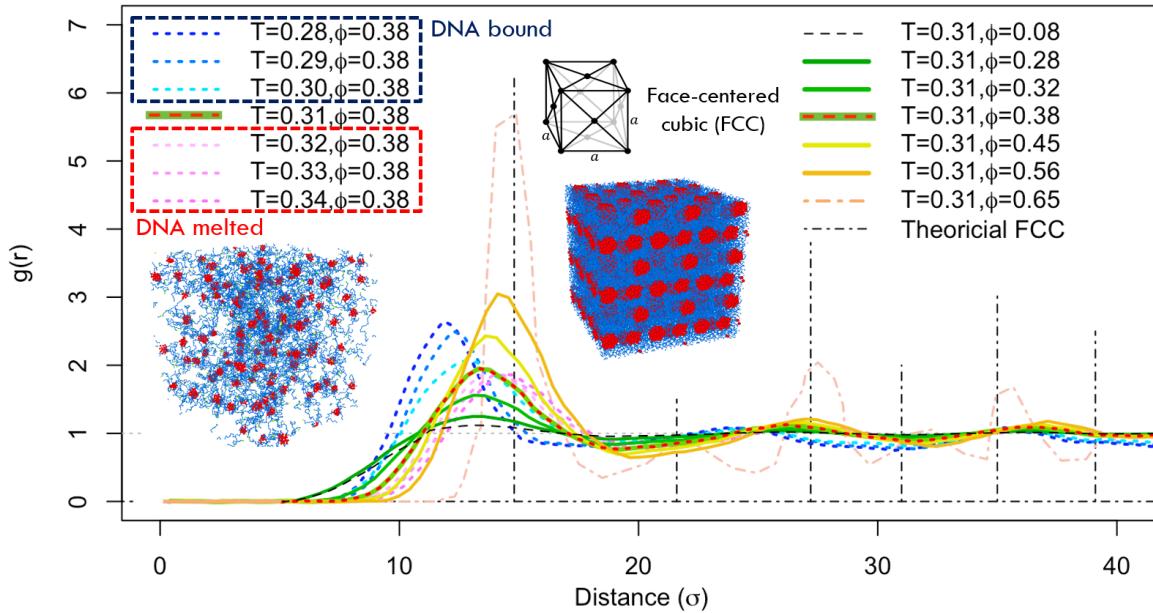


Fig. 6.7 Radial distribution functions $g(r)$ for F108 Pluronics with DNA patches ($T_m \approx 0.295$) for different volume fractions and temperatures. The left legend represents the estimated curves for $T \in [0.28, 0.34]$ with fixed $\phi = 0.38$, the right legend represents $g(r)$ for $\phi \in [0.08, 0.65]$ with fixed $T = 0.31$. All measurement points are illustrated in Fig. 6.8 with corresponding coloured, cross-shaped stars.

$g(r)$ was found for a range of densities and temperatures, which are plotted in Fig. 6.7. The plot displays two sets of parameter studies. The first set is at fixed $T = 0.31$ (above T_m) with $\phi \in [0.08, 0.65]$. We observe that for $\phi \in [0.28, 0.56]$, the first dominant peak is increasing with increasing ϕ , indicating that the systems with higher density forms slightly larger micelles, with an average micelle size around $12-14\sigma$. $\phi = 0.08$ is used as a reference line at zero, as this system does not display micelle formation. When $\phi = 0.65$, the system is in an ordered, solid-like state, and the peaks in $g(r)$ are in good agreement with the theoretical FCC structures. The shift in the peaks at long-distances from the theoretical prediction can be explained by the softness of the micelles, which make them compressible at high densities.

The second parameter study was performed for a fixed $\phi = 0.38$ with $T \in [0.28, 0.34]$. We are most interested in the formation of micelles at temperatures below and above the

DNA melting temperature ($T_m \approx 0.295$). We observe that the first peak of $g(r)$ is shifted to a smaller value when the micelles are connected by DNA patches ($T < T_m$) (blue dashed curves in Fig. 6.7). We assume this is because nearest-neighbour micelles are connected by DNA and can only move in a confined range. Hence the probability to observe a micelle that ‘sticks’ to another micelle is highly increased. When T approaches T_m , the system is in a transient (hybridization) region, where the DNA patches keep connecting and disconnecting. And when $T > T_m$, the DNA patches are dissociated and the computed curves for $g(r)$ are become similar to the micellar liquid phase observed for the non-functionalised F108.

6.3.5 The F108-DNA phase diagram

The numerical phase diagram for dispersions of F108 Pluronics with DNA patches ($T_m = 0.295$) is drawn as Fig. 6.8 (extra snapshots of various concentration-temperature couples are shown in **Appendix D.3**). When $T > T_m$, the sticky DNA ends are melted, and the system is close to a pure F108 Pluronics system, showing a similar micellar liquid and solid phase. When T approach or drops below T_m , a different phase behaviours emerge due to the presence of the DNA sticky ends. The attractive interactions become stronger at lower temperature, enriching the overall Pluronics phase diagram with extra phases including:

- Transient unimer liquid phase [$T < \text{CMT}, T \approx T_m$]: In the DNA melting region complementary ssDNA strands bind and unbind continuously, thus the DNA-functionalised unimers start forming concatenated chains; the system remains in a liquid state.
- Associated unimer liquid phase [$T < \text{CMT}, T < T_m$]: In this temperature region the live-time of DNA duplexes become very large. Consequently, the DNA-functionalised unimers tend to form concatenated chains of medium to large length; we tested for the presence of polymer-rings, however, they were rarely detected. We hypothesise that at low concentrations the chains will rather form a liquid phase. Though, as the F108-DNA concentration increase we most likely will obtain an entangled network if all the chains were associated, leading to a viscoelastic fluid.
- Transient micellar liquid phase [$\text{CMT} < T \approx T_m$]: Like in the transient unimer region, the micelles will be interconnected via rapid binding and unbinding processes between the sticky DNA chain ends in a dynamic balance. The system remains in a liquid state where small clusters of micelles and single micelles coexist.
- Micellar gel phase [$\text{CMT} < T < T_m$]: All micelles are linked by DNA, and a network of micelles has the potential of reaching a percolating gel phase.

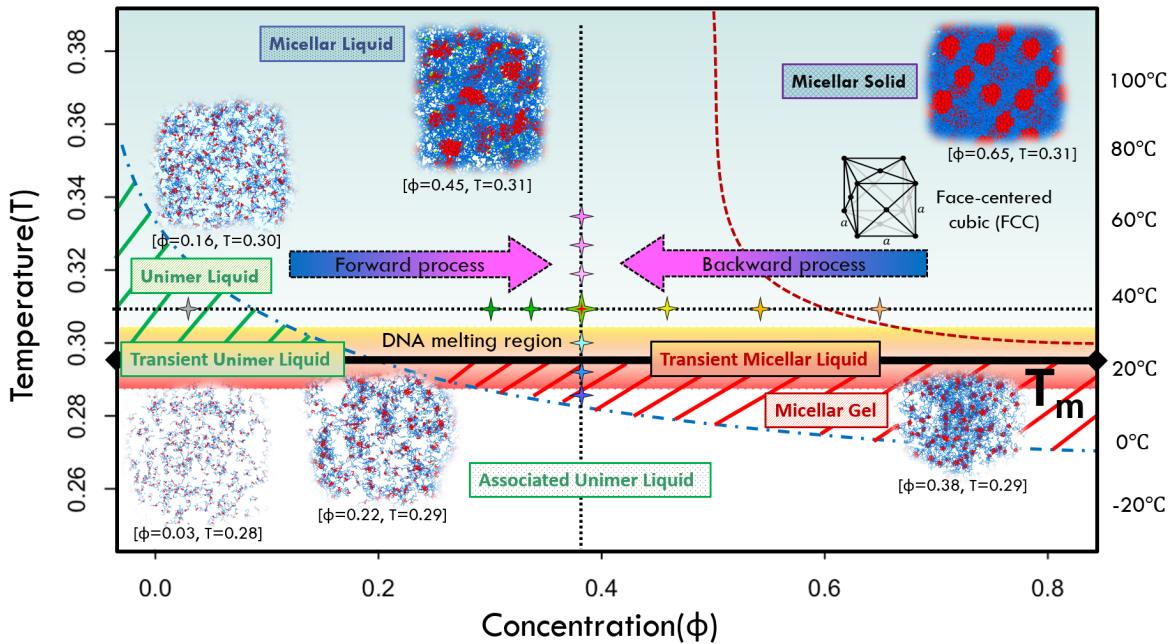


Fig. 6.8 The phase diagram of the F108-DNA system. The DNA melting temperature here is $T_m = 0.295$ (solid black line) with a hybridization region in $T \in [0.29, 0.31]$ which we denote as transient region. The forward process covers $\phi = 0.08 \rightarrow 0.38$; the backward process covers $\phi = 0.65 \rightarrow 0.38$. The blue dashed curve represents the CMT, the red dashed curve represents the micellar liquid to solid transition. Seven different regions are identified (from bottom to top): (1) associated unimer liquid; (2) micellar gel; (3) transient unimer liquid; (4) transient micellar liquid; (5) unimer liquid; (6) micellar liquid; (7) micellar solid. The cross-stars mark the phase points for which we calculated $g(r)$ with a colour representation corresponding to the results in Fig. 6.7.

The strength of our model is that the DNA melting region and the phase behaviour of the Pluronic dispersions can be tuned directly to match the experimental behaviour. With a value of $\beta(T) = 4.8$ in Eqn. 6.4, the melting temperature of the system is adjusted to $T_m = 0.32 \approx 50^\circ\text{C}$. In our experiment, we measured $T_{m(\text{exp})} = 48^\circ\text{C}$ at 100 mM NaCl for the complementary DNA pair used; its melting region is indicated by the orange shaded area in Fig. 2.9(b). The numerical phase diagram is drawn in Fig. 6.9, with the equivalent experimental temperatures indicated on the second y-axis. Two additional phases are expected in micellar solid phase, including the transient/associated micellar solid, where the micelles are packed in the system and connected by DNA patches.

The numerical F108-DNA phase diagram demonstrates that DNA-functionalisation alters the mechanical and structural properties of the micellar system. It pioneers the further characterisation and design of hybrid systems with micellisation and DNA hybridisation. Separated by the thick black line, corresponding to T_m in Fig. 6.9, the DNA-functionalised

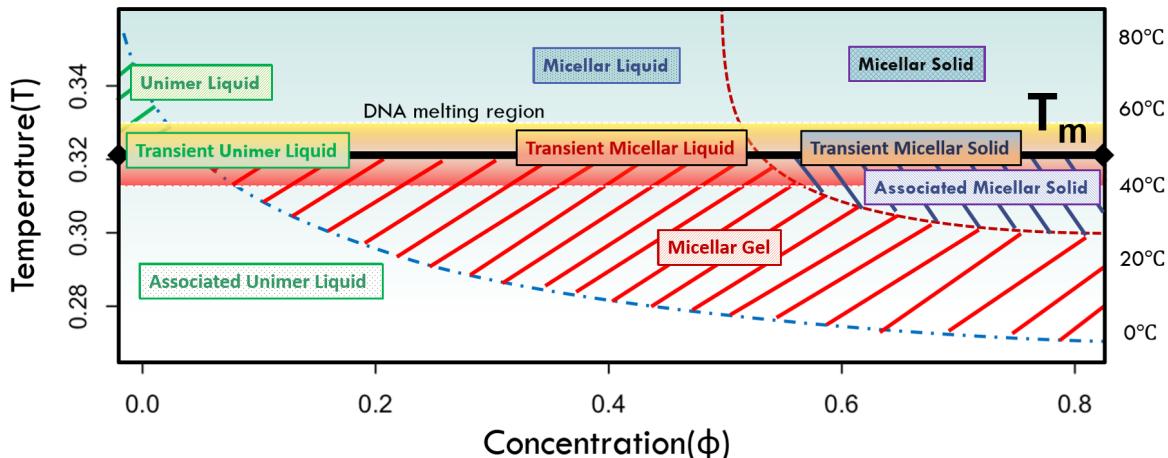


Fig. 6.9 Phase diagram of the F108-DNA system for a DNA melting temperature $T_m = 0.32$ with a melting region in the range $T \in [0.31, 0.33]$. Two additional transient/associated micellar solid phases occur. Blue dashed curve represents the micellisation boundary, and the red dashed curve represents the micellar liquid to solid phase transition.

copolymers produce a binary system of non-functional (above T_m) and functional (below T_m) micelles by simply tuning the system's temperature. With the help of fine-tunable parameters (such as $\alpha(T)$ and $\beta(T)$), we can control the interplay between T_m and CMT in the simulation, and in this way we achieve fine control over the temperature-dependent response.

6.3.6 Heterogeneity and homogeneity

In the associated unimer phase, some long chains are expected to form rings (Fig. 6.10(a)), while in the micellar gel phase, loop-formation may happen on two DNA sticky ends that belong to the same micelle, as shown in Fig. 6.10(b). Because of this looping behaviour, we expected to see a heterogeneous state in a limited temperature range given by the hybridization transition of the DNA sequences used. Indeed, this heterogeneous scenario only exists in the DNA linked micellar gel phase when $CMT < T < T_m$. It should be noted that throughout the temperature range considered here individual F108 chains can exchange from one micelle to another to reduce possible mechanical constraints. Once the temperature increases, the system will go back to a homogeneous state (shown as Fig. 6.11). The expected heterogeneity was observed qualitatively in the simulations, but did not give rise to a detectable difference in the radial distribution function. For this reason the structure factor was investigated in more detail, which also has the advantage of being a property that can be directly measured in the laboratory by scattering experiments. In this section we give a brief introduction to the

theoretical origin of $S(q)$ followed by a description of the computational implementation, before the results are presented and discussed.

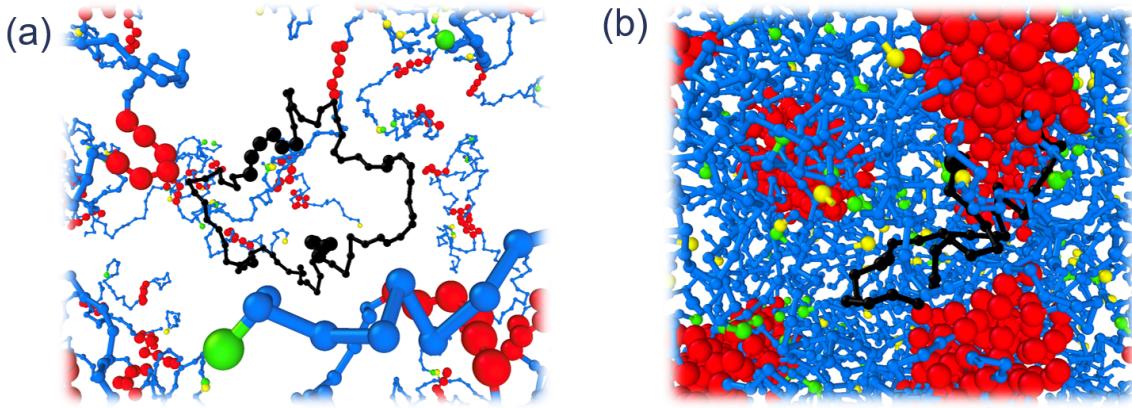


Fig. 6.10 Illustration of (a) Ring formation in associated unimer phase. (b) Micelles with intra-micellar DNA loop on in micellar gel phase. Both looping parts are coloured in black.

The structural features of a sample can be examined using either X-ray, neutron, electron or other scattering methods. As a rule of thumb, the wavelength λ of the radiation should be comparable to the length scale of interest [239]. In the case of elastic and monochromatic scattering, an incident beam with wave vector \mathbf{k}_{in} of length $\frac{2\pi}{\lambda}$ gives rise to a diffracted beam with wave vector \mathbf{k}_{out} of the same length but different orientation. The scattering vector \mathbf{q} is then given by the Laue criterion

$$\mathbf{q} = \mathbf{k}_{\text{out}} - \mathbf{k}_{\text{in}}. \quad (6.12)$$

Both the incident and the diffracted rays can be viewed as plane waves on the form $\psi_{\mathbf{k}} = \langle \mathbf{r} | \mathbf{k} \rangle \propto \exp(i\mathbf{k} \cdot \mathbf{r})$. With a detection screen placed behind the sample, the measured scattering intensity is proportional to the transition probability amplitude from state $\psi_{\mathbf{k}_{\text{in}}}$ to $\psi_{\mathbf{k}_{\text{out}}}$ after interacting with the scattering potential associated with the spatial density distribution $\rho(\mathbf{r})$,

$$I(\mathbf{q}) \propto \left| \langle \mathbf{k}_{\text{in}} | \rho(\mathbf{r}) | \mathbf{k}_{\text{out}} \rangle \right|^2 \propto \left| \int d\mathbf{r} \exp(-i\mathbf{k}_{\text{in}} \cdot \mathbf{r}) \rho(\mathbf{r}) \exp(i\mathbf{k}_{\text{out}} \cdot \mathbf{r}) \right|^2. \quad (6.13)$$

From Eqn 6.12 this is identical to

$$I(\mathbf{q}) \propto \left| \int d\mathbf{r} \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) \right|^2, \quad (6.14)$$

and by neglecting the coefficient of proportionality we can write the intensity as

$$I(\mathbf{q}) = \langle \hat{\rho}(\mathbf{q}) \hat{\rho}(-\mathbf{q}) \rangle, \quad (6.15)$$

where the reciprocal scattering density is given by the Fourier transform $\hat{\rho}_{\mathbf{q}} = \mathcal{F}\{\rho(\mathbf{r})\}(\mathbf{q})$, and the brackets indicate an ensemble average. In a system with N particles with in a volume V , the density distribution $\rho(\mathbf{r})$ consists of contributions from each particle j with scattering potential $f_j(\mathbf{r} - \mathbf{r}_j)$,

$$\rho(\mathbf{r}) = \sum_{j=1}^N f_j(\mathbf{r} - \mathbf{r}_j), \quad (6.16)$$

with the Fourier transform given by

$$\hat{\rho}(\mathbf{q}) = \sum_{j=1}^N \hat{f}_j(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_j), \quad (6.17)$$

where \hat{f}_j is the atomic form factor that describes the scattering amplitude of atom j ,

$$\hat{f}_j(\mathbf{q}) = \int_V d\mathbf{r} f_j(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}). \quad (6.18)$$

For point scatterers, the scattering potential is described by the delta function

$$f_j(\mathbf{r} - \mathbf{r}_j) = a_j \delta(\mathbf{r} - \mathbf{r}_j), \quad (6.19)$$

where a_j gives the scattering strength of atom j , such that the atomic form factor is simply

$$\hat{f}_j(\mathbf{q}) = a_j. \quad (6.20)$$

Other common choices for the scattering potential are the spherical potential (top-hat function) and the Gaussian potential. When the reciprocal scattering density $\hat{\rho}(\mathbf{q})$ is found, the structure factor is calculated by

$$S(\mathbf{q}) = \frac{1}{\sum_{j=1}^N \hat{f}_j^2(\mathbf{q})} \langle \hat{\rho}(\mathbf{q}) \hat{\rho}(-\mathbf{q}) \rangle, \quad (6.21)$$

or equivalently

$$S(\mathbf{q}) = \frac{I(\mathbf{q})}{I_0(\mathbf{q})} \frac{\rho_0}{\rho}, \quad (6.22)$$

where $I_0(\mathbf{q})$ is the measured intensity of a dilute system with density ρ_0 , which is the common experimental procedure for determining the structure factor.

For monodisperse systems, the structure factor is commonly calculated by first accumulating the radial distribution function $g(r)$ and then Fourier transforming $\rho[g(r) - 1]$. $S(q)$ is then obtained from

$$S(q) = 1 + \mathcal{F}\{\rho[g(r) - 1]\}(q). \quad (6.23)$$

Apart from being computationally expensive, particularly for particles that are not point scatterers, it also introduces assumptions about the behaviour of $g(r)$ for $r > L/2$ in order to avoid truncation errors in the Fourier transform, such that the results for low wave numbers are not reliable. As we are mostly interested in the low q behaviour of $S(q)$, we rather calculate $S(q)$ by a direct Fast Fourier Transform proposed by Frenkel and coworkers [240]. In the following section we first present the standard computation of $S(q)$, followed by the method of Frenkel.

Implementation

The first step to calculate $S(q)$ from our simulated micelle-data is to preform a second level coarse-graining of their trajectories. The PPO-cores of the micelles are then represented by a single particle. This is illustrated in Fig. 6.11, where the PPO-cores are shown together with the corresponding centre of mass representations - here only the latter is coloured red.

After the second level coarse-graining is done, one is left with trajectories of N particles in a cubic box with edges of length L . The next step in calculating $S(q)$ is to map the trajectories to a cubic density matrix ρ_m with appropriate bin-sizes. Note, great care should be taken when implementing the periodic boundaries. Along each dimension, the upper and lower limits represent the same position, and only one of these should be included as a bin. For a box centred at the origin, one can create M bins from $-L/2, -L/2 + D, -L/2 + 2D, \dots, L/2 - D$, with $D = L/M$. For programming languages that use zero-based array indexing, such as Python, the appropriate bin for a particle at position x can then be found e.g. by rounding $((x + L/2)/D)$ to the nearest integer and taking the modulus with M ; this is repeated for all three dimensions. M should be chosen such that $D < \sigma$ to avoid aliasing effects.

In the Numpy library for Python, the Fast Fourier Transform is implemented such that

$$\mathcal{F}\{a_m\}_k = \sum_{m=0}^{M-1} a_m \exp\left(-2\pi i \frac{mk}{M}\right) \quad (6.24)$$

for $k = 0, \dots, M - 1$. As our particle coordinates are all real numbers, we are only interested in positive frequencies, hence we use the routine `rfft` from `numpy.fft`. For even numbered M , this routine returns an array of length $(M/2) + 1$. The routine `rfftfreq` returns the associated frequencies for a given M and D , with the zero-frequency term containing the sum of the input array. The remaining $M/2$ elements range from $1/L$ to $1/2D$, the Nyquist frequency, and these are multiplied with 2π to get the conventional values for q .

The structure factor was then found by two methods, first by direct 3D FFT of ρ_m using the routine `numpy.rfftn`. Note that this routine only transform the real input over the last

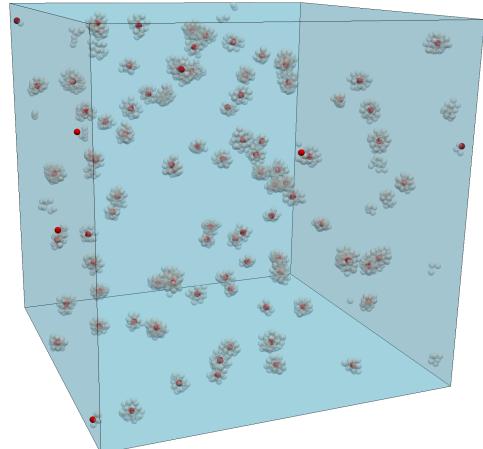
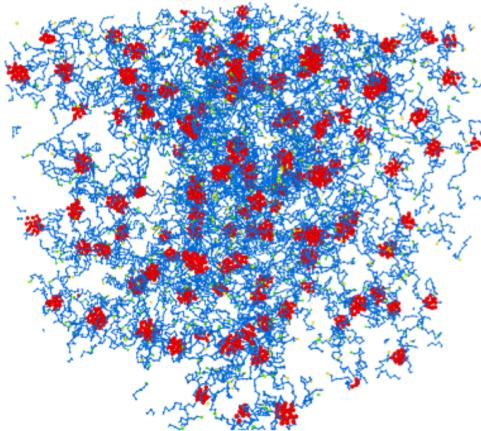
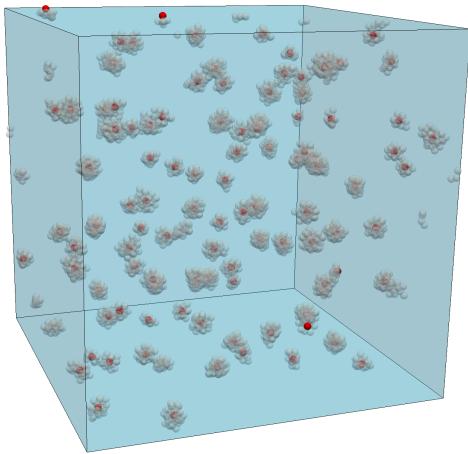
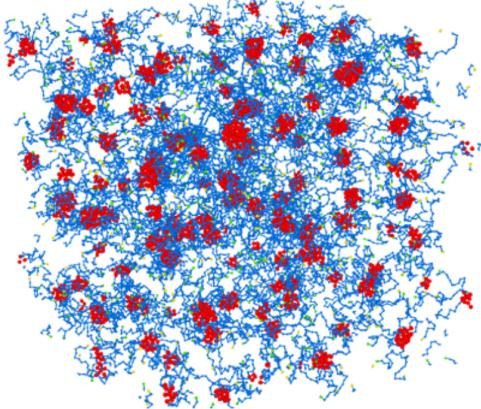
Heterogeneity(a) Heterogeneity: $T = 0.29$, below T_m (DNA bound).**Homogeneity**(b) Homogeneity: $T = 0.31$, above T_m (DNA melted).

Fig. 6.11 Second level coarse-graining of a homogeneous and a heterogeneous system with density=0.382. The PPO-cores in each micelle are replaced by single particles at the center of mass of each micelle, shown in red to the right.

transformation axis, for the first two axes the negative frequency transform is also calculated, returning an array $\mathcal{F}\{\rho_m\}_k$ with shape $(M, M, M/2)$, where the first $M/2$ elements of the first two axes correspond to the real frequency transform. The method supports arbitrary scattering potentials and polydispersity by direct mapping of $\rho(\mathbf{r})$ onto the discrete ρ_m . To avoid aliasing M has to be high enough such that $D < \sigma_{\max}$, where σ_{\max} is the largest particle diameter in the system. $S(q)$ is obtained from $S(\mathbf{q})$ by binning and averaging over all values of \mathbf{q} with length q .

Instead of the brute force 3D FFT approach, one can compute the 1D FFT along a range of selected directions $\hat{\mathbf{q}} = (\hat{h}, \hat{k}, \hat{l})$. These were chosen to be normal vectors for crystal planes

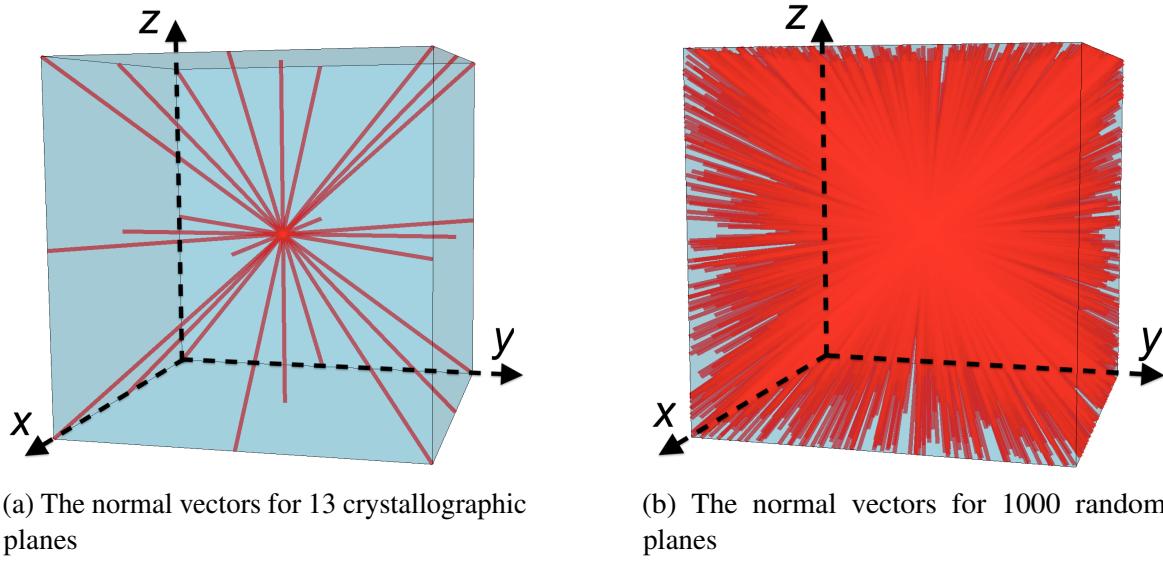


Fig. 6.12 Illustration of two different choices for calculating $S(q)$ by 1D FFT for a range of \mathbf{q} , shown here with the normal vectors for selected scattering planes in real space.

such as the 13 independent permutations of [100], [110] and [111] shown in Fig. 6.12a [240], though one can also use random directions as illustrated in Fig. 6.12b [241]. When binning the positions of the particle, it is crucial that the bin width is commensurate with the periodicity along $\hat{\mathbf{q}}$ to avoid artifacts for low values of q . We can introduce \mathbf{q}_p as the longest vector parallel to $\hat{\mathbf{q}}$ that will fit inside the box, with an associated rescaled length $L_p = L/\|\mathbf{q}_p\|$. For each scattering plane, one can then choose a binning frequency $D_p = L_p/M$. The projection of \mathbf{r}_j along $\hat{\mathbf{q}}$ is then found by a dot product, such that particle j can be placed in bin m_j by

$$m_j = \text{int}[(\hat{\mathbf{q}} \cdot \mathbf{r}_j + L_p/2)/D_p](\text{mod } M). \quad (6.25)$$

The contribution to $S(q)$ from each plane is found using Eqn 6.21 with a suitable atomic form factor, before binning and averaging $S(q)$ over the selected planes.

Both methods for calculating the structure factor were tested using small simulation of particles with the WCA potential: The resulting $S(q)$ matches well with the analytical expression of the Percus-Yevick approximation, see **Appendix B.4**.

Results

The structure factor was calculated by a direct 3D FFT for the heterogeneous and the homogeneous system in Fig. 6.11a. The amount of heterogeneity can be quantified from the calculated $S(q)$ at low wave numbers of q (this value also indicates the frequency of detecting the sizes of the ‘cavities’ that can be observed in the system). If the calculated

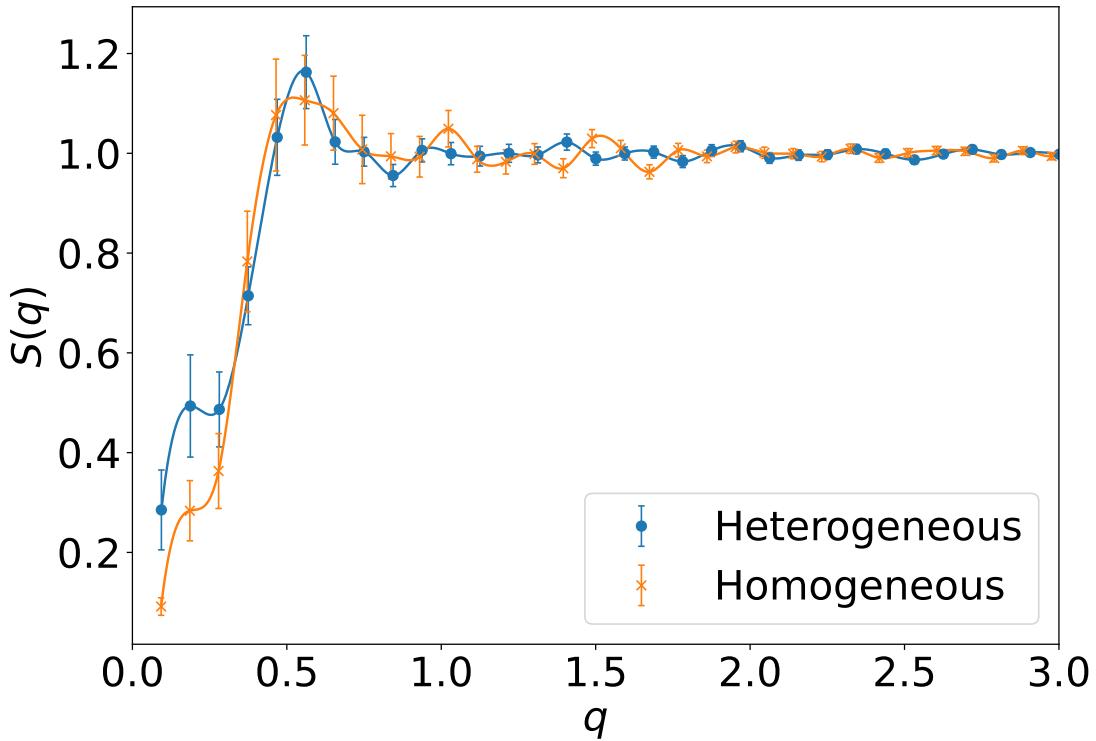


Fig. 6.13 The structure factor for the heterogeneous and homogeneous micellar system shown in Fig. 6.11a and b respectively; the smooth line serves here as guide to the eyes. The data were averaged over 11 trajectories, and the error bars are given by the standard error for each bin.

$S(q)$ at low wave numbers is larger, then the system has more large cavities. The boundary between heterogeneity and homogeneity is related to the melting temperature regime from Fig. 6.5. The threshold of melting temperature provides a trigger to determine when or when not it becomes heterogeneity. The calculated $S(q)$ is presented in Fig. 6.13. The data is averaged over 11 trajectories, and the error bars give the standard error for each bin. For low wave numbers, we see that $S(q)$ is significantly higher for the heterogeneous system, as expected. This is explained by the loop-formation of the micelles at $CMT < T < T_m$, which is reducing the valency and giving rise to cavities in the trajectories, which increase the scattering of low wave numbers. For wave numbers higher than about $q/4$, the structure factors of the two systems are similar, which is reasonable. The possibility of achieving sufficiently good statistics from only 11 trajectories demonstrates the superiority of this method of computing $S(q)$ as opposed to the conventional method of first calculating $g(r)$ and then applying Eq. 6.23. In the near future, this analysis will be repeated with more data to further improve the statistics.

6.4 Conclusion

This chapter presents a simulation study of the self-assembly properties of F108 Pluronics functionalised with DNA sticky ends. The latter act as specific patches that crosslink the unimers/micelles into connected networks. The study is based on Langevin dynamics simulations with a bead-spring model of the copolymer system that allows for specifying the melting temperature of the DNA patches. Our work offers a modelling parallel to the experimental effort on DNA-functionalised F108 micelles. We simulated a pure F108 system utilising temperature-sensitive parameters which allowed for tuning of T_m and hence also direct comparison with the experimental phase diagram. We found that the melting temperature of the DNA patches strongly affects the phase behaviours of the micellar system. DNA patches also form local networks when micelles are formed, leading to a heterogeneous state in the transient micellar liquid phase. With the help of calculating $g(r)$ of the system, we can distinguish the newfound phases with novel-explored structural properties. The numerical phase diagram of DNA-functionalised F108 Pluronic with tunable DNA melting temperature has taught us important features of the system by controlling two temperature-sensitive terms (CMT and T_m). By shifting T_m , novel system structures with exotic phase diagrams were observed. Further, a calculation of $S(q)$ was introduced to distinguish between the heterogeneous and homogeneous micellar gel phase of the F108-DNA system. For low wave numbers, it was clear that $S(q)$ was significantly higher for the heterogeneous system. This can be explained by the loop-formation of the micelles, which gives rise to cavities in the trajectories.

As the numerical study is usually under an ideal state, there are several open questions to address in future theoretical and computational works. To start with, in simulation, the micellisation can be finished before adding up DNA patches. In the experiment, on the other hand, the DNA binding and the micellisation process are usually simultaneous. Although there is no difference after the system reaches the equilibrium state, the effect of the formation order might be interesting to study. Since this system has two different time scales – i.e., the fast dynamics of the DNA patches on the micelle surface and the slow dynamics of the micelles – more advanced simulation techniques might be required to reduce the computing time. This is especially relevant if one wants to compare systems where there is a large difference in size between micelles and DNA patches, as in the case of micron-sized micelles and nano-sized DNA patchy particles.

Additionally, all simulated copolymers in this system are attached to two DNA patches on each side. However, in an experiment, some copolymers are linked with one/no DNA sticky end that coexists in the solution. Hence, in the future system update, we can further add a

parameter to control the number of DNA sticky ends on each copolymers, either zero, one or two, to make the simulated system even more similar to the experiment. Moreover, in our experiment, we used the azide moieties attached to the F108 chain ends, which bind the DNA oligomers to them via a ‘click’ reaction. These azide moieties might also be considered in the simulation model to help explain how attachment of azide groups on F108 Pluronics change the structural properties of micellar fluids. To this end, I hope the simulations presented in this chapter will stimulate further experimental work in this fascinating and relatively unexplored area.

6.5 Acknowledgements

I want to thank Prof. Daan Frenkel for the theoretical support in understanding the algorithm for the radial distribution function $g(r)$ and the structure factor $S(q)$. I also thank Dr. Eivind Bering who did much of the coding work for computing $S(q)$; thanks go also to Dr. Alessio Caciagli and Ren Liu for providing and explaining our experimental work on the Pluronics system. And finally, thanks again to Prof. Erika Eiser for her kind and patient supervision.

7 | Closing Remarks

This thesis covers the main work from my PhD studies. Additional side projects are listed in **Appendix E**. In the main project, I introduced a multi-level coarse-grained strategy, which begins by investigating the thermodynamic stability of short, single stranded and double stranded DNA that can hybridize to form nano-sized building blocks. Once their stability and spacial configuration was explored, we mapped them onto a more coarse-grained model to collect information on the self-assembly of many copies of these DNA nano-sized building blocks into large-scale network structures with new phases. The finer-level simulation was explicitly designed to capture the detailed structural, mechanical, and thermodynamic properties of an individual DNA building block used in experiments in the Eiser group. In contrast, the many-body systems with less complexity were intended to explore the bulk properties in a dense solution simulation. The results from this multi-level numerical technique serves as a demonstration of the feasibility and practicality of combining experimental work with computational modelling tools to increase our understanding of the principal physical properties behind our experimental findings.

To start with, I built a system that uses oxDNA-based molecular dynamics simulations to obtain a statistical analysis of the influence of ssDNA flexible joints in a DNA building block. The simulation work investigated the effects of a flexible joint with specified lengths in linker-mediated DNA hydrogel. The thermodynamic results proved that the similarly sized stiff arms for the Y-shapes and the linear duplex with the flexible joint could lead to self-binding, thus reducing the overall system valency and suppressing the formation of percolating gels. We confirmed that minuscule changes in the flexibility of the individual building blocks could significantly alter the macroscopic properties of the entire network. In particular, the simulations with oxDNA showed that in order to form a liquid-cluster phase, the ssDNA flexible joint has to have a length of at least 6 bases. When the joint size was less than 6 bases, gelation was still preferred due to the increased penalty in bending elasticity of the DNA arms. Beyond that, we also found that the oxDNA model gave an excellent estimate of the melting temperature for key-looped DNA building blocks associated with the experimental measurements.

To systematically analyse the effect of ssDNA flexible joints, I further conducted statistical computations on the bending free energy of individual nunchuck-like DNA building blocks. I created a bending potential driven by the stiffness factor to quantitatively describe the rigidity associated with any number of bases with a given joint length. We concluded that the flexibility was increased by using long ssDNA junctions with extra bending capacity until reaching a saturation region of more than ten bases. This work provides useful prior information to design a higher-level simulation to discover potentially new liquid-crystal phases in a dense solution. We recommended two coarse-grained models to simulate the nunchuck DNA as a liquid-crystal mesogen. Especially with a short joint length (less than 6-bases), a confined angle of bending limits was observed. Hence, we were keenly interested in exploring the phase behaviour of angle confined DNA mesogens.

Subsequently, I built a many-body nunchuck-DNA simulation to investigate the phase behaviour and dynamics of rigid rod and fixed/flexible bend-core DNA mesogens. We demonstrated how semi-flexible, and linear dsDNA exhibits the anticipated phase transition from isotropic to the nematic phase. Particularly, when built with a fixed angle to create a bend-core liquid-crystal mesogen, the phase behaviour of the system changes to a biaxial nematic. When the length of the ssDNA is adjusted so the connection between a mesogen's stiff dsDNA arms becomes highly flexible, the system resembles 'nunchucks' and the phase transition can be completely suppressed. In a future step, we aim to exploit our knowledge to design new types of mesogens that can provide us with strategies to obtain new types of liquid crystal phases. Hence, our simulations will greatly help to find suitable synthesised DNA mesogens in our lab, thus make our experimental explorations more efficient.

The final project was concerned with aqueous solutions of the Pluronic F108, an amphiphilic triblock copolymer that self-assembles into spherical micelles upon heating them. These polymers' chain-ends where functionalised with short ssDNA chains (F108 + ssDNA) such that they bind below the DNA's melting temperature, thus providing a competing driving force for self-assembly. In particular, we experimentally explored the possibility of finding a potential re-entrant gelation. To this end, I designed a customised bead-spring model in LAMMPS representing the Pluronic chains, while the sticky ssDNA was implemented in form of small patchy beads attached to the F108 chain ends. The simulation model presented a clear view of the formation of micelles upon heating and the phase transition form a unimer to micellar liquid and eventually to a micellar solid at high enough polymer concentrations. Thus, our simulations reflected the experimentally obtained phase diagram of the F108 system. However, introducing the DNA-functionalization lead to a much richer phase diagram with new phases, that were not possible to easily resolve in experiments. The further analysis of structural properties informed us that attaching DNA to any Pluronic renders the system

heterogeneous. I envision applications in drug delivery and nano-medicine upon reaching a better understanding of precisely how DNA linkers influence the micelle formation and how it can be directed at will.

In summary, the numerical study of a coarse-grained DNA simulation now allows DNA nano-structures such as DNA building blocks in a large-scale network to be accurately probed. Exploring the systematic behaviour of bulk properties from the entire DNA network is a challenge for computer simulations due to the long-time scales involved. It can be envisaged that through the combination of top-down and bottom-up design for the multi-level coarse-grained modelling, it will finally be achievable to perform systematic many-body simulations with multiple DNA building blocks. With the help of this multi-level coarse-grained strategy, we developed a more realistic simulation system to investigate meaningful biophysical representations. The numerical analysis of a multi-level coarse-grained simulation not only enables precise description of DNA nano-structures in micro- and macroscopic properties; but also advances the interdisciplinary study of using simulation tools to accommodate the growing demand for biological and medical applications.

Beyond the above studies, we proposed several future directions for designing DNA-related materials. One of the most exciting developments of the linker-mediated DNA hydrogel is to achieve highly selective and sensitive molecular sensors used as the framework. With extra flexibility, the microscopic structure of the network (e.g. pore-size, cavity) becomes more adaptable. This DNA network benefits opening the pore-based DNA nanotubes design in **Appendix E.1**. Its application may find in nanomedicine once we have a better understanding of the sub-pore contributions. The DNA-functionalised micelles offer impressive flexibility with a well-defined thermal response over the structure of single and multi-component systems. Further studies about this system are open for discussion to address fundamental physics questions (such as the effects of polydisperse diameter distribution in **Appendix E.2**). The most promising direction is to find out the possibility of designing DNA-rich networks with multi-components of thermo-responsive features. These could be achieved by introducing DNA-building blocks and micelles as combined builders. Based on the specific interaction between DNA linkers and triblock-copolymers, we started the work of combining micelles with Y-shape DNA as connectors in **Appendix E.3**. From here, we expect more fascinating phase behaviours of the new system. Overall, more advanced multi-level strategies of computational models are required to aid experimental works, towards characterising DNA-related materials and making accurate predictions of their physical properties.

References

- [1] Frank Edward Stockdale and Howard Holtzer. DNA synthesis and myogenesis. *Experimental cell research*, 24(3):508–520, 1961.
- [2] Lawrence A Loeb and Thomas A Kunkel. Fidelity of DNA synthesis. *Annual review of biochemistry*, 51(1):429–457, 1982.
- [3] Hans Bügl, John P Danner, Robert J Molinari, John T Mulligan, Han-Oh Park, Bas Reichert, David A Roth, Ralf Wagner, Bruce Budowle, Robert M Scripp, et al. DNA synthesis and biological security. *Nature biotechnology*, 25(6):627–629, 2007.
- [4] Sriram Kosuri and George M Church. Large-scale de novo DNA synthesis: technologies and applications. *Nature methods*, 11(5):499–507, 2014.
- [5] Shawn M Douglas, Ido Bachelet, and George M Church. A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–834, 2012.
- [6] Mingxu You, Yan Chen, Xiaobing Zhang, Haipeng Liu, Ruowen Wang, Kelong Wang, Kathryn R Williams, and Weihong Tan. An autonomous and controllable light-driven dna walking device. *Angewandte Chemie International Edition*, 51(10):2457–2460, 2012.
- [7] Souvik Modi, Swetha MG, Debanjan Goswami, Gagan D Gupta, Satyajit Mayor, and Yamuna Krishnan. A DNA nanomachine that maps spatial and temporal ph changes inside living cells. *Nature nanotechnology*, 4(5):325–330, 2009.
- [8] Helmut Schiessel. *Biophysics for beginners: a journey through the cell nucleus*. Jenny Stanford Publishing, 2021.
- [9] Thomas E Ouldridge, Rollo L Hoare, Ard A Louis, Jonathan PK Doye, Jonathan Bath, and Andrew J Turberfield. Optimizing DNA nanotechnology through coarse-grained modeling: a two-footed DNA walker. *ACS nano*, 7(3):2479–2490, 2013.
- [10] Veikko Linko and Hendrik Dietz. The enabled state of DNA nanotechnology. *Current opinion in biotechnology*, 24(4):555–561, 2013.
- [11] Ai-Sheng Xiong, Quan-Hong Yao, Ri-He Peng, Hui Duan, Xian Li, Hui-Qin Fan, Zong-Ming Cheng, and Yi Li. Pcr-based accurate synthesis of long DNA sequences. *Nature protocols*, 1(2):791–797, 2006.
- [12] Erin M McConnell, Devon Morrison, Maria Alejandra Rey Rincon, Bruno J Salena, and Yingfu Li. Selection and applications of synthetic functional dnas for bacterial detection. *TrAC Trends in Analytical Chemistry*, 124:115785, 2020.

- [13] Süleyman Cenkci, Mustafa Yıldız, İbrahim Hakkı Ciğerci, Muhsin Konuk, and Ahmet Bozdağ. Toxic chemicals-induced genotoxicity detected by random amplified polymorphic dna (rapd) in bean (*phaseolus vulgaris l.*) seedlings. *Chemosphere*, 76(7):900–906, 2009.
- [14] Xiangcheng Pan, Sushil Lathwal, Stephanie Mack, Jiajun Yan, Subha R Das, and Krzysztof Matyjaszewski. Automated synthesis of well-defined polymers and biohybrids by atom transfer radical polymerization using a dna synthesizer. *Angewandte Chemie International Edition*, 56(10):2740–2743, 2017.
- [15] Marketa Sagova-Mareckova, Ladislav Cermak, Jitka Novotna, Kamila Plhackova, Jana Forstova, and Jan Kopecky. Innovative methods for soil dna purification tested in soils with widely differing characteristics. *Applied and environmental microbiology*, 74(9):2902–2907, 2008.
- [16] Ângela Sousa, Fani Sousa, and João A Queiroz. Advances in chromatographic supports for pharmaceutical-grade plasmid dna purification. *Journal of separation science*, 35(22):3046–3058, 2012.
- [17] Michael Levitt. Computer simulation of DNA double-helix dynamics. In *Cold Spring Harbor symposia on quantitative biology*, volume 47, pages 251–262. Cold Spring Harbor Laboratory Press, 1983.
- [18] Pablo D Dans, Jürgen Walther, Hansel Gómez, and Modesto Orozco. Multiscale simulation of DNA. *Current opinion in structural biology*, 37:29–45, 2016.
- [19] Konstantin V Klenin, Alexander V Vologodskii, Vadim V Anshelevich, Alexander M Dykhne, and Maxim D Frank-Kamenetskii. Computer simulation of DNA supercoiling. *Journal of molecular biology*, 217(3):413–419, 1991.
- [20] Michael Feig and B Montgomery Pettitt. A molecular simulation picture of DNA hydration around A-and B-DNA. *Biopolymers: Original Research on Biomolecules*, 48(4):199–209, 1998.
- [21] Petr Šulc, Flavio Romano, Thomas E Ouldridge, Lorenzo Rovigatti, Jonathan PK Doye, and Ard A Louis. Sequence-dependent thermodynamics of a coarse-grained DNA model. *The Journal of chemical physics*, 137(13):135101, 2012.
- [22] Irene Conti, Piero Altoè, Marco Stenta, Marco Garavelli, and Giorgio Orlandi. Adenine deactivation in DNA resolved at the caspt2//casscf/amber level. *Physical Chemistry Chemical Physics*, 12(19):5016–5023, 2010.
- [23] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. Charmm: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.
- [24] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. The structure and function of DNA. In *Molecular Biology of the Cell. 4th edition*. Garland Science, 2002.

- [25] Yuhang Dong, Chi Yao, Yi Zhu, Lu Yang, Dan Luo, and Dayong Yang. DNA functional materials assembled from branched DNA: design, synthesis, and applications. *Chemical Reviews*, 120(17):9420–9481, 2020.
- [26] Jeff Hooyberghs, Paul Van Hummelen, and Enrico Carlon. The effects of mismatches on hybridization in DNA microarrays: determination of nearest neighbor parameters. *Nucleic acids research*, 37(7):e53–e53, 2009.
- [27] John SantaLucia Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95(4):1460–1465, 1998.
- [28] Lorenzo Di Michele, Bortolo M Mognetti, Taiki Yanagishima, Patrick Varilly, Zachary Ruff, Daan Frenkel, and Erika Eiser. Effect of inert tails on the thermodynamics of DNA hybridization. *Journal of the American Chemical Society*, 136(18):6538–6541, 2014.
- [29] Leslie Pray. Discovery of DNA structure and function: Watson and crick. *Nature Education*, 1(1), 2008.
- [30] Nadrian C Seeman. Nucleic acid junctions and lattices. *Journal of theoretical biology*, 99(2):237–247, 1982.
- [31] Xiang Li, Chuan Zhang, Chenhui Hao, Cheng Tian, Guansong Wang, and Chengde Mao. DNA polyhedra with t-linkage. *ACS nano*, 6(6):5138–5142, 2012.
- [32] Ebbe S Andersen, Mingdong Dong, Morten M Nielsen, Kasper Jahn, Ramesh Subramani, Wael Mamdouh, Monika M Golas, Bjoern Sander, Holger Stark, Cristiano LP Oliveira, et al. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459(7243):73–76, 2009.
- [33] Michael T Hwang, Zejun Wang, Jinglei Ping, Deependra Kumar Ban, Zi Chao Shiah, Leif Antonschmidt, Joon Lee, Yushuang Liu, Abhijith G Karkisaval, Alan T Charlie Johnson, et al. DNA nanotweezers and graphene transistor enable label-free genotyping. *Advanced Materials*, 30(34):1802440, 2018.
- [34] Dinesh C Khara, John S Schreck, Toma E Tomov, Yaron Berger, Thomas E Ouldridge, Jonathan PK Doye, and Eyal Nir. DNA bipedal motor walking dynamics: an experimental and theoretical study of the dependency on step size. *Nucleic acids research*, 46(3):1553–1561, 2018.
- [35] Yongzheng Xing, Enjun Cheng, Yang Yang, Ping Chen, Tao Zhang, Yawei Sun, Zhongqiang Yang, and Dongsheng Liu. Self-assembled DNA hydrogels with designable thermal and enzymatic responsiveness. *Advanced Materials*, 23(9):1117–1121, 2011.
- [36] Yudong Hao, Martin Kristiansen, Ruojie Sha, Jens J Birktoft, Carina Hernandez, Chengde Mao, and Nadrian C Seeman. A device that operates within a self-assembled 3d DNA crystal. *Nature chemistry*, 9(8):824–827, 2017.

- [37] Darshana Joshi, Dylan Bargteil, Alessio Caciagli, Jerome Burelbach, Zhongyang Xing, André S Nunes, Diogo EP Pinto, Nuno AM Araújo, Jasna Brujic, and Erika Eiser. Kinetic control of the coverage of oil droplets by DNA-functionalized colloids. *Science advances*, 2(8):e1600881, 2016.
- [38] Soong Ho Um, Jong Bum Lee, Nokyung Park, Sang Yeon Kwon, Christopher C Umbach, and Dan Luo. Enzyme-catalysed assembly of DNA hydrogel. *Nature materials*, 5(10):797–801, 2006.
- [39] Enjun Cheng, Yongzheng Xing, Ping Chen, Yang Yang, Yawei Sun, Dejian Zhou, Lijin Xu, Qinghua Fan, and Dongsheng Liu. A ph-triggered, fast-responding DNA hydrogel. *Angewandte Chemie International Edition*, 48(41):7660–7663, 2009.
- [40] Feng Li, Jianpu Tang, Jinhui Geng, Dan Luo, and Dayong Yang. Polymeric DNA hydrogel: Design, synthesis and applications. *Progress in Polymer Science*, 98:101163, 2019.
- [41] Silvia Biffi, Roberto Cerbino, Francesca Bomboi, Elvezia Maria Paraboschi, Rosanna Asselta, Francesco Sciortino, and Tommaso Bellini. Phase behavior and critical activated dynamics of limited-valence DNA nanostars. *Proceedings of the National Academy of Sciences*, 110(39):15633–15637, 2013.
- [42] Clemens Jochum, Nataša Adžić, Emmanuel Stiakakis, Thomas L Derrien, Dan Luo, Gerhard Kahl, and Christos N Likos. Structure and stimuli-responsiveness of all-DNA dendrimers: theory and experiment. *Nanoscale*, 11(4):1604–1617, 2019.
- [43] Elliot S Bishop, Sami Mostafa, Mikhail Pakvasa, Hue H Luu, Michael J Lee, Jennifer Moriatis Wolf, Guillermo A Ameer, Tong-Chuan He, and Russell R Reid. 3-d bioprinting technologies in tissue engineering and regenerative medicine: Current and future trends. *Genes & diseases*, 4(4):185–195, 2017.
- [44] Huiling Jiang, Victor Pan, Skanda Vivek, Eric R Weeks, and Yonggang Ke. Programmable DNA hydrogels assembled from multidomain DNA strands. *Chem-BioChem*, 17(12):1156–1162, 2016.
- [45] Zhongyang Xing, Alessio Caciagli, Tianyang Cao, Iliya Stoev, Mykolas Zupkauskas, Thomas O'Neill, Tobias Wenzel, Robin Lamboll, Dongsheng Liu, and Erika Eiser. Microrheology of DNA hydrogels. *Proceedings of the National Academy of Sciences*, 115(32):8137–8142, 2018.
- [46] Zhongyang Xing, Christopher Ness, Daan Frenkel, and Erika Eiser. Structural and linear elastic properties of DNA hydrogels by coarse-grained simulation. *Macromolecules*, 52(2):504–512, 2019.
- [47] Iliya D Stoev, Tianyang Cao, Alessio Caciagli, Jiaming Yu, Christopher Ness, Ren Liu, Rini Ghosh, Thomas O'Neill, Dongsheng Liu, and Erika Eiser. On the role of flexibility in linker-mediated DNA hydrogels. *Soft Matter*, 16(4):990–1001, 2020.
- [48] Kai Liu, Dong Chen, Alessio Marcozzi, Lifei Zheng, Juanjuan Su, Diego Pesce, Wojciech Zajaczkowski, Anke Kolbe, Wojciech Pisula, Klaus Müllen, et al. Thermotropic liquid crystals from biomacromolecules. *Proceedings of the National Academy of Sciences*, 111(52):18596–18600, 2014.

- [49] Yuri M Yevdokimov, VI Salyanov, SV Semenov, and SG Skuridin. *DNA liquid-crystalline dispersions and nanoconstructions*. CRC Press Boca Raton-London-New York, 2011.
- [50] Lars Onsager. The effects of shape on the interaction of colloidal particles. *Annals of the New York Academy of Sciences*, 51(4):627–659, 1949.
- [51] Maxime MC Tortora and Jonathan PK Doye. Incorporating particle flexibility in a density functional description of nematics and cholesterics. *Molecular Physics*, 116(21-22):2773–2791, 2018.
- [52] Michi Nakata, Giuliano Zanchetta, Brandon D Chapman, Christopher D Jones, Julie O Cross, Ronald Pindak, Tommaso Bellini, and Noel A Clark. End-to-end stacking and liquid crystal condensation of 6–to 20–base pair DNA duplexes. *Science*, 318(5854):1276–1279, 2007.
- [53] Tommaso P Fraccia, Gregory P Smith, Lucas Bethge, Giuliano Zanchetta, Giovanni Nava, Sven Klussmann, Noel A Clark, and Tommaso Bellini. Liquid crystal ordering and isotropic gelation in solutions of four-base-long DNA oligomers. *ACS nano*, 10(9):8508–8516, 2016.
- [54] Miroslaw Salamonczyk, Jing Zhang, Giuseppe Portale, Chenhui Zhu, Emmanuel Kentzinger, James T Gleeson, Antal Jakli, Cristiano De Michele, Jan KG Dhont, Samuel Sprunt, et al. Smectic phase in suspensions of gapped DNA duplexes. *Nature communications*, 7(1):1–9, 2016.
- [55] Liang Wang and Yusuke Yamauchi. Block copolymer mediated synthesis of dendritic platinum nanoparticles. *Journal of the American Chemical Society*, 131(26):9152–9153, 2009.
- [56] F Agnely, A Djedour, A Bochot, and J-L Grossiord. Properties of various thermoassociating polymers: pharmaceutical and cosmetic applications. *Journal of drug delivery science and technology*, 16(1):3–10, 2006.
- [57] Andrew M Bodratti and Paschalidis Alexandridis. Formulation of poloxamers for drug delivery. *Journal of functional biomaterials*, 9(1):11, 2018.
- [58] Tharwat Tadros. Viscoelastic properties of sterically stabilised emulsions and their stability. *Advances in colloid and interface science*, 222:692–708, 2015.
- [59] Nienke Geerts and Erika Eiser. DNA-functionalized colloids: Physical properties and applications. *Soft Matter*, 6(19):4647–4660, 2010.
- [60] Lorenzo Di Michele, Francesco Varrato, Jurij Kotar, Simon H Nathan, Giuseppe Foffi, and Erika Eiser. Multistep kinetic self-assembly of DNA-coated colloids. *Nature communications*, 4(1):1–7, 2013.
- [61] Lorenzo Di Michele and Erika Eiser. Developments in understanding and controlling self assembly of DNA-functionalized colloids. *Physical Chemistry Chemical Physics*, 15(9):3115–3129, 2013.

- [62] Jialin Yu, Huayu Qiu, Shouchun Yin, Hebin Wang, and Yang Li. Polymeric drug delivery system based on pluronics for cancer treatment. *Molecules*, 26(12):3610, 2021.
- [63] Anaïs Pitto-Barry and Nicolas PE Barry. Pluronic® block-copolymers in medicine: from chemical and biological versatility to rationalisation and clinical advances. *Polymer Chemistry*, 5(10):3291–3297, 2014.
- [64] Masayuki Kurahashi, Kazuyoshi Kanamori, Kazuyuki Takeda, Hironori Kaji, and Kazuki Nakanishi. Role of block copolymer surfactant on the pore formation in methylsilsesquioxane aerogel systems. *RSC advances*, 2(18):7166–7173, 2012.
- [65] Paschalis Alexandridis and T Alan Hatton. Poly (ethylene oxide) poly (propylene oxide) poly (ethylene oxide) block copolymer surfactants in aqueous solutions and at interfaces: thermodynamics, structure, dynamics, and modeling. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 96(1-2):1–46, 1995.
- [66] G Wanka, H Hoffmann, and W Ulbricht. Phase diagrams and aggregation behavior of poly (oxyethylene)-poly (oxypropylene)-poly (oxyethylene) triblock copolymers in aqueous solutions. *Macromolecules*, 27(15):4145–4159, 1994.
- [67] Erika Eiser, François Molino, Grégoire Porte, and Xavier Pithon. Flow in micellar cubic crystals. *Rheologica Acta*, 39(3):201–208, 2000.
- [68] Yugong Wu, Zhigang Fan, and Yuzhu Lu. Bulk and interior packing densities of random close packing of hard spheres. *Journal of materials science*, 38(9):2019–2025, 2003.
- [69] Avraham Halperin. Polymeric micelles: their relaxation kinetics. *Macromolecules*, 22(5):2403–2412, 1989.
- [70] Alessio Caciagli, Mykolas Zupkauskas, Aviad Levin, Tuomas PJ Knowles, Clément Mugemana, Nico Bruns, Thomas O’Neill, William J Frith, and Erika Eiser. DNA-coated functional oil droplets. *Langmuir*, 34(34):10073–10080, 2018.
- [71] C Holm, A Arnold, and O Lenz. Espresso: Extensible simulation package for research on soft matter. university of stuttgart, 2012.
- [72] Fariborz Mohamadi, Nigel GJ Richards, Wayne C Guida, Rob Liskamp, Mark Lipton, Craig Caufield, George Chang, Thomas Hendrickson, and W Clark Still. Macro-model—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *Journal of Computational Chemistry*, 11(4):440–467, 1990.
- [73] Molecular Massively Parallel Simulator. Lammps. *Theater*, 2012, 2012.
- [74] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Elsevier, 2001.
- [75] Michael P Allen, Daan Frenkel, and J Talbot. Molecular dynamics simulation using hard particles. *Computer physics reports*, 9(6):301–353, 1989.

- [76] Christopher Ness, Vladimir V Palyulin, Rico Milkus, Robert Elder, Timothy Sirk, and Alessio Zaccone. Nonmonotonic dependence of polymer-glass mechanical response on chain bending stiffness. *Physical Review E*, 96(3):030501, 2017.
- [77] Boris Derjaguin and Lev Landau. Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes. *Progress in Surface Science*, 43(1-4):30–59, 1993.
- [78] Evert Johannes Willem Verwey. Theory of the stability of lyophobic colloids. *The Journal of Physical Chemistry*, 51(3):631–636, 1947.
- [79] John Edward Jones. On the determination of molecular fields.—i. from the variation of the viscosity of a gas with temperature. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106(738):441–462, 1924.
- [80] John E Lennard-Jones. Cohesion. *Proceedings of the Physical Society (1926-1948)*, 43(5):461, 1931.
- [81] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [82] M Sales-Pardo, R Guimera, AA Moreira, J Widom, and LAN Amaral. Mesoscopic modeling for nucleic acid chain dynamics. *Physical Review E*, 71(5):051902, 2005.
- [83] Thomas E Ouldridge, Ard A Louis, and Jonathan PK Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *The Journal of chemical physics*, 134(8):02B627, 2011.
- [84] Masao Doi, Samuel Frederick Edwards, and Samuel Frederick Edwards. *The theory of polymer dynamics*, volume 73. oxford university press, 1988.
- [85] Benedict EK Snodin, Ferdinando Randisi, Majid Mosayebi, Petr Šulc, John S Schreck, Flavio Romano, Thomas E Ouldridge, Roman Tsukanov, Eyal Nir, Ard A Louis, et al. Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *The Journal of chemical physics*, 142(23):06B613_1, 2015.
- [86] Oliver Henrich, Yair Augusto Gutiérrez Fosado, Tine Curk, and Thomas E Ouldridge. Coarse-grained simulation of DNA using LAMMPS. *The European Physical Journal E*, 41(5):1–16, 2018.
- [87] Gianluca Martelloni and Franco Bagnoli. Infiltration effects on a two-dimensional molecular dynamics model of landslides. *Natural hazards*, 73(1):37–62, 2014.
- [88] Jonathan PK Doye, Thomas E Ouldridge, Ard A Louis, Flavio Romano, Petr Šulc, Christian Matek, Benedict EK Snodin, Lorenzo Rovigatti, John S Schreck, Ryan M Harrison, et al. Coarse-graining DNA for simulations of DNA nanotechnology. *Physical Chemistry Chemical Physics*, 15(47):20395–20414, 2013.

- [89] Ryan M Harrison, Flavio Romano, Thomas E Ouldridge, Ard A Louis, and Jonathan PK Doye. Coarse-grained modelling of strong DNA bending ii: Cyclization. *arXiv preprint arXiv:1506.09008*, 2015.
- [90] Benedict E K Snodin, Flavio Romano, Lorenzo Rovigatti, Thomas E Ouldridge, Ard A Louis, and Jonathan P K Doye. Direct simulation of the self-assembly of a small DNA origami. *ACS nano*, 10(2):1724–1737, 2016.
- [91] Erik Winfree, Furong Liu, Lisa A Wenzler, and Nadrian C Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–544, 1998.
- [92] Nadrian C Seeman. DNA in a material world. *Nature*, 421(6921):427–431, 2003.
- [93] Nadrian C Seeman. Biochemistry and structural DNA nanotechnology: an evolving symbiotic relationship. *Biochemistry*, 42(24):7259–7269, 2003.
- [94] Yu He, Ye Tian, Yi Chen, Zhaoxiang Deng, Alexander E Ribbe, and Chengde Mao. Sequence symmetry as a tool for designing DNA nanostructures. *Angewandte Chemie International Edition*, 44(41):6694–6696, 2005.
- [95] Yu He, Tao Ye, Min Su, Chuan Zhang, Alexander E Ribbe, Wen Jiang, and Chengde Mao. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature*, 452(7184):198–201, 2008.
- [96] Carlos Ernesto Castro, Fabian Kilchherr, Do-Nyun Kim, Enrique Lin Shiao, Tobias Wauer, Philipp Wortmann, Mark Bathe, and Hendrik Dietz. A primer to scaffolded DNA origami. *Nature methods*, 8(3):221–229, 2011.
- [97] Dongran Han, Suchetan Pal, Jeanette Nangreave, Zhengtao Deng, Yan Liu, and Hao Yan. DNA origami with complex curvatures in three-dimensional space. *Science*, 332(6027):342–346, 2011.
- [98] Nicholas AW Bell, Christian R Engst, Marc Ablay, Giorgio Divitini, Caterina Ducati, Tim Liedl, and Ulrich F Keyser. DNA origami nanopores. *Nano letters*, 12(1):512–517, 2012.
- [99] Barbara Saccà and Christof M Niemeyer. DNA origami: the art of folding DNA. *Angewandte Chemie International Edition*, 51(1):58–66, 2012.
- [100] Francesco Sciortino and Emanuela Zaccarelli. Reversible gels of patchy particles. *Current Opinion in Solid State and Materials Science*, 15(6):246–253, 2011.
- [101] Cameron Alexander, Håkan S Andersson, Lars I Andersson, Richard J Ansell, Nicole Kirsch, Ian A Nicholls, John O’Mahony, and Michael J Whitcombe. Molecular imprinting science and technology: a survey of the literature for the years up to and including 2003. *Journal of Molecular Recognition: An Interdisciplinary Journal*, 19(2):106–180, 2006.
- [102] Sytze J Buwalda, Kristel WM Boere, Pieter J Dijkstra, Jan Feijen, Tina Vermonden, and Wim E Hennink. Hydrogels in a historical perspective: From simple networks to smart materials. *Journal of controlled release*, 190:254–273, 2014.

- [103] Hao Qi, Majid Ghodousi, Yanan Du, Casey Grun, Hojae Bae, Peng Yin, and Ali Khademhosseini. DNA-directed self-assembly of shape-controlled hydrogels. *Nature communications*, 4(1):1–10, 2013.
- [104] Emanuela Bianchi, Julio Largo, Piero Tartaglia, Emanuela Zaccarelli, and Francesco Sciortino. Phase diagram of patchy colloids: Towards empty liquids. *Physical review letters*, 97(16):168301, 2006.
- [105] Frank Smallenburg and Francesco Sciortino. Liquids more stable than crystals in particles with limited valence and flexible bonds. *Nature Physics*, 9(9):554–558, 2013.
- [106] Francis W Starr and Francesco Sciortino. “crystal-clear” liquid–liquid transition in a tetrahedral fluid. *Soft Matter*, 10(47):9413–9422, 2014.
- [107] Chia Wei Hsu, Francesco Sciortino, and Francis W Starr. Theoretical description of a DNA-linked nanoparticle self-assembly. *Physical review letters*, 105(5):055502, 2010.
- [108] Leonard M Adleman. Molecular computation of solutions to combinatorial problems. *science*, 266(5187):1021–1024, 1994.
- [109] Georg Seelig, David Soloveichik, David Yu Zhang, and Erik Winfree. Enzyme-free nucleic acid logic circuits. *science*, 314(5805):1585–1588, 2006.
- [110] Richard Lavery, Krystyna Zakrzewska, David Beveridge, Thomas C Bishop, David A Case, Thomas Cheatham III, Surjit Dixit, B Jayaram, Filip Lankas, Charles Laughton, et al. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic acids research*, 38(1):299–313, 2010.
- [111] Karen Drukker, Guosheng Wu, and George C Schatz. Model simulations of DNA denaturation dynamics. *The Journal of Chemical Physics*, 114(1):579–590, 2001.
- [112] Lu Hong, Bodhi P Vani, Erik H Thiede, Michael J Rust, and Aaron R Dinner. Molecular dynamics simulations of nucleotide release from the circadian clock protein kaic reveal atomic-resolution functional insights. *Proceedings of the National Academy of Sciences*, 115(49):E11475–E11484, 2018.
- [113] M-N Dessinges, Berenike Maier, Y Zhang, M Peliti, David Bensimon, and Vincent Croquette. Stretching single stranded DNA, a model polyelectrolyte. *Physical review letters*, 89(24):248102, 2002.
- [114] Francis W Starr and Francesco Sciortino. Model for assembly and gelation of four-armed DNA dendrimers. *Journal of Physics: Condensed Matter*, 18(26):L347, 2006.
- [115] Praveen Depa, Chunxia Chen, and Janna K Maranas. Why are coarse-grained force fields too fast? a look at dynamics of four coarse-grained polymers. *The Journal of chemical physics*, 134(1):014903, 2011.
- [116] John SantaLucia Jr and Donald Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415–440, 2004.

- [117] Alexander Stukowski. Visualization and analysis of atomistic simulation data with ovito—the open visualization tool. *Modelling and simulation in materials science and engineering*, 18(1):015012, 2009.
- [118] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.
- [119] Thomas G Mason. Estimating the viscoelastic moduli of complex fluids using the generalized stokes–einstein equation. *Rheologica acta*, 39(4):371–378, 2000.
- [120] Carl Schildkraut and Shneior Lifson. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers: Original Research on Biomolecules*, 3(2):195–208, 1965.
- [121] Gary S Grest and Kurt Kremer. Molecular dynamics simulation for polymers in the presence of a heat bath. *Physical Review A*, 33(5):3628, 1986.
- [122] Kurt Kremer and Gary S Grest. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *The Journal of Chemical Physics*, 92(8):5057–5086, 1990.
- [123] Prabesh Gyawali, Rony Saha, Gregory P Smith, Miroslaw Salamonczyk, Prakash Kharel, Soumitra Basu, Ruipeng Li, Masafumi Fukuto, James T Gleeson, Noel A Clark, et al. Mono-and bilayer smectic liquid crystal ordering in dense solutions of “gapped” DNA duplexes. *Proceedings of the National Academy of Sciences*, 118(12):e2019996118, 2021.
- [124] Wilfried Kramer and Hans-Joachim Fritz. [18] oligonucleotide-directed construction of mutations via gapped duplex DNA. In *Methods in enzymology*, volume 154, pages 350–367. Elsevier, 1987.
- [125] Tumpa Sarkar, Christine C Conwell, Lilia C Harvey, Catherine T Santai, and Nicholas V Hud. Condensation of oligonucleotides assembled into nicked and gapped duplexes: potential structures for oligonucleotide delivery. *Nucleic acids research*, 33(1):143–151, 2005.
- [126] Janine B Mills, Elsi Vacano, and Paul J Hagerman. Flexibility of single-stranded DNA: use of gapped duplex helices to determine the persistence lengths of poly (dt) and poly (da). *Journal of molecular biology*, 285(1):245–257, 1999.
- [127] Xinyue Cai, D Sebastian Arias, Lourdes R Velazquez, Shelby Vexler, Alexander L Bevier, and D Kuchnir Fygenson. DNA nunchucks: nanoinstrumentation for single-molecule measurement of stiffness and bending. *Nano Letters*, 20(2):1388–1395, 2019.
- [128] Steven B Smith, Yujia Cui, and Carlos Bustamante. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, 271(5250):795–799, 1996.

- [129] V Luzzati, A Mathis, F Masson, and J Witz. Structure transitions observed in DNA and poly a in solution as a function of temperature and ph. *Journal of Molecular Biology*, 10(1):28–41, 1964.
- [130] Frank B Howard, Joe Frazier, Marie N Lipsett, and H Todd Miles. Infrared demonstration of two-and three-strand helix formation between poly c and guanosine mononucleotides and oligonucleotides. *Biochemical and Biophysical Research Communications*, 17(1):93–102, 1964.
- [131] Geoffrey W Hoffman and Dietmar Pörschke. Cooperative nonenzymic base recognition. thermodynamics of the helix—coil transition of a monomer—polymer double helix. *Biopolymers: Original Research on Biomolecules*, 12(7):1611–1623, 1973.
- [132] Bethel S Stannard and Gary Felsenfeld. The conformation of polyriboadenylic acid at low temperature and neutral ph. a single-stranded rodlike structure. *Biopolymers: Original Research on Biomolecules*, 14(2):299–307, 1975.
- [133] Karin R Hagerman and Paul J Hagerman. Helix rigidity of DNA: the meroduplex as an experimental paradigm. *Journal of molecular biology*, 260(2):207–223, 1996.
- [134] Massimiliano Chiappini, Tara Drwenski, René Van Roij, and Marjolein Dijkstra. Bi-axial, twist-bend, and splay-bend nematic phases of banana-shaped particles revealed by lifting the “smectic blanket”. *Physical Review Letters*, 123(6):068001, 2019.
- [135] Kaare M Nielsen, Pål J Johnsen, Douda Bensasson, and Daniele Daffonchio. Release and persistence of extracellular DNA in the environment. *Environmental biosafety research*, 6(1-2):37–53, 2007.
- [136] Yongjun Lu, Brock Weers, and Nancy C Stellwagen. DNA persistence length revisited. *Biopolymers: Original Research on Biomolecules*, 61(4):261–275, 2002.
- [137] Margaret C Linak and Kevin D Dorfman. Erratum:“analysis of a DNA simulation model through hairpin melting experiments”[j. chem. phys. 133, 125101 (2010)]. *The Journal of chemical physics*, 135(2):07B801, 2011.
- [138] R Core Team. R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. <http://www.R-project.org/>, 2013.
- [139] Lars Onsager. The effects of shape on the interaction of colloidal particles. *Annals of the New York Academy of Sciences*, 51(4):627–659, May 1949.
- [140] Joakim Bohlin, Michael Matthies, Erik Poppleton, Jonah Procyk, Aatmik Mallya, Hao Yan, and Petr Šulc. Design and simulation of DNA, rna and hybrid protein–nucleic acid nanostructures with oxview. *Nature Protocols*, pages 1–27, 2022.
- [141] J Trenton Kostbade. Mapping frequency distributions by the box-and-whisker. *The Professional Geographer*, 33(4):413–418, 1981.
- [142] Rob Eppenga and Daan Frenkel. Monte carlo study of the isotropic and nematic phases of infinitely thin hard platelets. *Molecular physics*, 52(6):1303–1334, 1984.

- [143] Daan Frenkel and Rob Eppenga. Monte Carlo study of the isotropic-nematic transition in a fluid of thin hard disks. *Physical Review Letters*, 49(15):1089–1092, October 1982.
- [144] Joseph E Kirkwood. Phase transformations in solids, 1951.
- [145] William C Kerr and MJ Rave. Mean-field theory of entropy-driven structural phase transitions. *Physical Review B*, 48(22):16234, 1993.
- [146] Daan Frenkel. Entropy-driven phase transitions. *Physica A: statistical mechanics and its applications*, 263(1-4):26–38, 1999.
- [147] Marie Adams, Zvonimir Dogic, Sarah L Keller, and Seth Fraden. Entropically driven microphase transitions in mixtures of colloidal rods and spheres. *Nature*, 393(6683):349–352, 1998.
- [148] Valerie J Anderson and Henk NW Lekkerkerker. Insights into phase transition kinetics from colloid science. *Nature*, 416(6883):811–815, 2002.
- [149] PA Forsyth Jr, S Marčelia, DJ Mitchell, and BW Ninham. Ordering in colloidal systems. *Advances in Colloid and Interface Science*, 9(1):37–60, 1978.
- [150] Michael Engel, Pablo F Damasceno, Carolyn L Phillips, and Sharon C Glotzer. Computational self-assembly of a one-component icosahedral quasicrystal. *Nature Materials*, 14(1):109–116, 2015.
- [151] Marjolein Dijkstra. Entropy-driven phase transitions in colloids: From spheresto anisotropic particles. *Advances in Chemical Physics, Volume 156*, 156:35, 2014.
- [152] Hideo Takezoe and Yoichi Takanishi. Bent-core liquid crystals: their mysterious and attractive world. *Japanese journal of applied physics*, 45(2R):597, 2006.
- [153] Jesús Etxebarria and M Blanca Ros. Bent-core liquid crystals in the route to functional materials. *Journal of Materials Chemistry*, 18(25):2919–2926, 2008.
- [154] R Amaranatha Reddy and Carsten Tschierske. Bent-core liquid crystals: polar order, superstructural chirality and spontaneous desymmetrisation in soft matter systems. *Journal of Materials Chemistry*, 16(10):907–961, 2006.
- [155] Massimiliano Chiappini and Marjolein Dijkstra. A generalized density-modulated twist-splay-bend phase of banana-shaped particles. *Nature communications*, 12(1):1–9, 2021.
- [156] Piotr Kubala, Wojciech Tomczyk, and Michał Cieśla. In silico study of liquid crystalline phases formed by bent-shaped molecules with excluded-volume type interactions. *arXiv preprint arXiv:2112.14607*, 2021.
- [157] Yang Yang, Hanwen Pei, Guangdong Chen, Kyle Thomas Webb, Luz J Martinez-Miranda, Isabel K Lloyd, Zhongyuan Lu, Kun Liu, and Zhihong Nie. Phase behaviors of colloidal analogs of bent-core liquid crystals. *Science advances*, 4(5):eaas8829, 2018.

- [158] Carla Fernández-Rico, Massimiliano Chiappini, Taiki Yanagishima, Heidi de Sousa, Dirk GAL Aarts, Marjolein Dijkstra, and Roel PA Dullens. Shaping colloidal bananas to reveal biaxial, splay-bend nematic, and smectic phases. *Science*, 369(6506):950–955, 2020.
- [159] Paul WK Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, 2006.
- [160] Yuanchen Dong, Zhongqiang Yang, and Dongsheng Liu. DNA nanotechnology based on i-motif structures. *Accounts of Chemical Research*, 47(6):1853–1860, 2014.
- [161] Yun Jeong Cha, Min-Jun Gim, Kyunghwan Oh, and Dong Ki Yoon. Twisted-nematic-mode liquid crystal display with a DNA alignment layer. *Journal of Information Display*, 16(3):129–135, 2015.
- [162] Ling Wang, Augustine M Urbas, and Quan Li. Nature-inspired emerging chiral liquid crystal nanostructures: from molecular self-assembly to DNA mesophase and nanocolloids. *Advanced Materials*, 32(41):1801335, 2020.
- [163] Nadrian C Seeman and Hanadi F Sleiman. DNA nanotechnology. *Nature Reviews Materials*, 3(1):1–23, 2017.
- [164] James W Canary and Hao Yan. Nadrian c.(ned) seeman (1945–2021), 2022.
- [165] Françoise Livolant and Amélie Leforestier. Condensed phases of DNA: structures and phase transitions. *Progress in Polymer Science*, 21(6):1115–1164, 1996.
- [166] Cristiano De Michele, Giuliano Zanchetta, Tommaso Bellini, Elisa Frezza, and Alberta Ferrarini. Hierarchical propagation of chirality through reversible polymerization: the cholesteric phase of DNA oligomers. *ACS Macro Letters*, 5(2):208–212, 2016.
- [167] David Winogradoff, Pin-Yi Li, Himanshu Joshi, Lauren Quednau, Christopher Maffeo, and Aleksei Aksimentiev. Chiral systems made from DNA. *Advanced Science*, 8(5):2003113, 2021.
- [168] Edward Barry and Zvonimir Dogic. Entropy driven self-assembly of nonamphiphilic colloidal membranes, 2010.
- [169] Keng-hui Lin, John C Crocker, Vikram Prasad, Andrew Schofield, David A Weitz, TC Lubensky, and AG Yodh. Entropically driven colloidal crystallization on patterned surfaces. *Physical Review Letters*, 85(8):1770, 2000.
- [170] Sami Nummelin, Juhana Kommeri, Mauri A Kostiainen, and Veikko Linko. Evolution of structural DNA nanotechnology. *Advanced Materials*, 30(24):1703721, 2018.
- [171] Florian Praetorius, Benjamin Kick, Karl L Behler, Maximilian N Honemann, Dirk Weuster-Botz, and Hendrik Dietz. Biotechnological mass production of DNA origami. *Nature*, 552(7683):84–87, 2017.
- [172] Mark Bathe and Paul WK Rothemund. DNA nanotechnology: A foundation for programmable nanoscale materials. *MRS Bulletin*, 42(12):882–888, 2017.

- [173] Helgi I Ingólfsson, Cesar A Lopez, Jaakko J Uusitalo, Djurre H de Jong, Srinivasa M Gopal, Xavier Periole, and Siewert J Marrink. The power of coarse graining in biomolecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3):225–248, 2014.
- [174] Davit A Potoyan, Alexey Savelyev, and Garegin A Papoian. Recent successes in coarse-grained modeling of DNA. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(1):69–83, 2013.
- [175] Alexander D MacKerell Jr, Joanna Wiorkiewicz-Kuczera, and Martin Karplus. An all-atom empirical energy function for the simulation of nucleic acids. *Journal of the American Chemical society*, 117(48):11946–11975, 1995.
- [176] Romelia Salomon-Ferrer, David A Case, and Ross C Walker. An overview of the amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2):198–210, 2013.
- [177] Thomas E Cheatham III. Simulation and modeling of nucleic acid structure, dynamics and interactions. *Current opinion in structural biology*, 14(3):360–367, 2004.
- [178] Davide Michieletto, Enzo Orlandini, and Davide Marenduzzo. Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3d organization of chromatin domains. *Physical Review X*, 6(4):041047, 2016.
- [179] Enrique De Miguel, Luis F. Rull, Manoj K. Chalam, and Keith E. Gubbins. Liquid crystal phase diagram of the gay-berne fluid. *Molecular Physics*, 74(2):405–424, October 1991.
- [180] Alejandro D Rey. Theory and simulation of gas diffusion in cholesteric liquid crystal films. *Molecular Crystals and Liquid Crystals Science and Technology. Section A. Molecular Crystals and Liquid Crystals*, 293(1):87–109, 1997.
- [181] Struther Arnott and DWL Hukins. Optimised parameters for A-DNA and B-DNA. *Biochemical and biophysical research communications*, 47(6):1504–1509, 1972.
- [182] Hernan G Garcia, Paul Grayson, Lin Han, Mandar Inamdar, Jané Kondev, Philip C Nelson, Rob Phillips, Jonathan Widom, and Paul A Wiggins. Biological consequences of tightly bent DNA: the other life of a macromolecular celebrity. *Biopolymers: Original Research on Biomolecules*, 85(2):115–130, 2007.
- [183] Rico Milkus, Christopher Ness, Vladimir V Palyulin, Jana Weber, Alexei Lapkin, and Alessio Zacccone. Interpretation of the vibrational spectra of glassy polymers using coarse-grained simulations. *Macromolecules*, 51(4):1559–1572, 2018.
- [184] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, July 1984.
- [185] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, March 1985.

- [186] Wataru Shinoda, Motoyuki Shiga, and Masuhiro Mikami. Rapid estimation of elastic constants by molecular dynamics simulation under constant stress. *Physical Review B - Condensed Matter and Materials Physics*, 69(13):16–18, 2004.
- [187] Erich P Stoll and Toni Schneider. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Physical Review B*, 17(3):1302–1322, February 1978.
- [188] Pierre-Gilles De Gennes and Jacques Prost. *The Physics of Liquid Crystals*. International Series of Monogr. Clarendon Press, 1993.
- [189] Daan Frenkel and Rob Eppenga. Evidence for algebraic orientational order in a two-dimensional hard-core nematic. *Physical Review A*, 31(3):1776, 1985.
- [190] Simone Dussi, Massimiliano Chiappini, and Marjolein Dijkstra. On the stability and finite-size effects of a columnar phase in single-component systems of hard-rod-like particles. *Molecular Physics*, 116(21-22):2792–2805, May 2018.
- [191] Daan Frenkel. Computer simulation of hard-core models for liquid crystals. *Molecular physics*, 60(1):1–20, 1987.
- [192] Daan Frenkel and Bela M Mulder. The hard ellipsoid-of-revolution fluid: I. monte carlo simulations. *Molecular physics*, 55(5):1171–1192, 1985.
- [193] Daan Frenkel, Bela M Mulder, and John P McTague. Phase diagram of hard ellipsoids of revolution. *Molecular Crystals and Liquid Crystals*, 123(1):119–128, 1985.
- [194] Alan K Soper. Determination of the orientational pair correlation function of a molecular liquid from diffraction data. *Journal of molecular liquids*, 78(3):179–200, 1998.
- [195] Daan Frenkel, BM Mulder, and JP McTague. Phase diagram of a system of hard ellipsoids. *Physical review letters*, 52(4):287, 1984.
- [196] Sin-Doo Lee. A numerical investigation of nematic ordering based on a simple hard-rod model. *The Journal of Chemical Physics*, 87(8):4972–4974, October 1987.
- [197] MP Allen and Daan Frenkel. Observation of dynamical precursors of the isotropic-nematic transition by computer simulation. *Physical review letters*, 58(17):1748, 1987.
- [198] Philip Camp, Carl Mason, Michael Allen, Anjali Khare, and David Kofke. The isotropic–nematic phase transition in uniaxial hard ellipsoid fluids: Coexistence data and the approach to the onsager limit. *The Journal of Chemical Physics*, 105(7):2837–2849, August 1996.
- [199] Kenji Kubo and Kazuyoshi Ogino. Comparison of osmotic pressures for the poly (γ -benzyl-l-glutamate) solutions with the theories for a system of hard spherocylinders. *Molecular Crystals and Liquid Crystals*, 53(3-4):207–227, January 1979.
- [200] Rudolf Oldenbourg and Donald LD Caspar. Orientational distribution function in nematic tobacco-mosaic-virus liquid crystals measured by x-ray diffraction. *Physical review letters*, 61(16):1851, 1988.

- [201] Seth Fraden, Georg Maret, and D. L. D. Caspar. Angular correlations and the isotropic-nematic phase transition in suspensions of tobacco mosaic virus. *Physical Review E*, 48(4):2816–2837, October 1993.
- [202] Daan Frenkel, HNW Lekkerkerker, and A Stroobants. Thermodynamic stability of a smectic phase in a system of hard rods. *Nature*, 332(6167):822–823, 1988.
- [203] Simon McGrother, Dave Williamson, and George Jackson. A re-examination of the phase diagram of hard spherocylinders. *The Journal of Chemical Physics*, 104(17):6755–6771, May 1996.
- [204] François Gay-Balmaz, Tudor S. Ratiu, and Cesare Tronci. Equivalent theories of liquid crystal dynamics. *Archive for Rational Mechanics and Analysis*, 210(3):773–811, September 2013.
- [205] Tongyang Zhao and Xiaogong Wang. Diffusion of rigid rodlike polymer in isotropic solutions studied by dissipative particle dynamics simulation. *Polymer*, 54(19):5241–5249, August 2013.
- [206] Herrera-Valencia Alejandro D Rey and Yogesh Kumar Murugesan. Structure and dynamics of biological liquid crystals. *Liquid Crystals*, 41(3):430–451, October 2013.
- [207] Dominique Ernst and Jürgen Köhler. Measuring a diffusion coefficient by single-particle tracking: statistical analysis of experimental mean squared displacement curves. *Phys. Chem. Chem. Phys.*, 15(3):845–849, 2013.
- [208] Ralf Metzler and Joseph Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77, December 2000.
- [209] Philip Camp and Michael Allen. Phase diagram of the hard biaxial ellipsoid fluid. *The Journal of Chemical Physics*, 106(16):6681–6688, April 1997.
- [210] Philip Camp, Michael Allen, and Andrew Masters. Theory and computer simulation of bent-core molecules. *The Journal of Chemical Physics*, 111(21):9871–9881, DEC 1 1999.
- [211] Elena Batrakova and Alexander Kabanov. Pluronic block copolymers: evolution of drug delivery concept from inert nanocarriers to biological response modifiers. *Journal of controlled release*, 130(2):98–106, 2008.
- [212] Eun Seong Lee, Zhonggao Gao, and You Han Bae. Recent progress in tumor ph targeting nanotechnology. *Journal of Controlled Release*, 132(3):164–170, 2008.
- [213] Jayanta Kumar Patra, Gitishree Das, Leonardo Fernandes Fraceto, Estefania Vangelie Ramos Campos, Maria del Pilar Rodriguez-Torres, Laura Susana Acosta-Torres, Luis Armando Diaz-Torres, Renato Grillo, Mallappa Kumara Swamy, Shivesh Sharma, et al. Nano based drug delivery systems: recent developments and future prospects. *Journal of nanobiotechnology*, 16(1):1–33, 2018.
- [214] Evgeniia Konishcheva, Davy Daubian, Jens Gaitzsch, and Wolfgang Meier. Synthesis of linear abc triblock copolymers and their self-assembly in solution. *Helvetica Chimica Acta*, 101(2):e1700287, 2018.

- [215] Stephan Förster and Markus Antonietti. Amphiphilic block copolymers in structure-controlled nanomaterial hybrids. *Advanced Materials*, 10(3):195–217, 1998.
- [216] Jacob N Israelachvili. *Intermolecular and surface forces*. Academic press, 2011.
- [217] Jintao Zhu and Wei Jiang. Self-assembly of abc triblock copolymer into giant segmented wormlike micelles in dilute solution. *Macromolecules*, 38(22):9315–9323, 2005.
- [218] Si Kyung Yang, Ashootosh V Ambade, and Marcus Weck. Supramolecular abc triblock copolymers via one-pot, orthogonal self-assembly. *Journal of the American Chemical Society*, 132(5):1637–1645, 2010.
- [219] Marlous Kamp, Bart De Nijs, Marjolein N Van Der Linden, Isja De Feijter, Merel J Lefferts, Antonio Alois, Jack Griffiths, Jeremy J Baumberg, Ilja K Voets, and Alfons Van Blaaderen. Multivalent patchy colloids for quantitative 3d self-assembly studies. *Langmuir*, 36(9):2403–2418, 2020.
- [220] Saerom Lee, Jeong Hoon Yoon, In-Seong Jo, Joon Suk Oh, David J Pine, Tae Soup Shim, and Gi-Ra Yi. DNA-functionalized 100 nm polymer nanoparticles from block copolymer micelles. *Langmuir*, 34(37):11042–11048, 2018.
- [221] Matthew R Jones, Robert J Macfarlane, Byeongdu Lee, Jian Zhang, Kaylie L Young, Andrew J Senesi, and Chad A Mirkin. DNA-nanoparticle superlattices formed from anisotropic building blocks. *Nature materials*, 9(11):913–917, 2010.
- [222] Robert J Macfarlane, Byeongdu Lee, Matthew R Jones, Nadine Harris, George C Schatz, and Chad A Mirkin. Nanoparticle superlattice engineering with DNA. *Spherical Nucleic Acids*, pages 539–553, 2020.
- [223] Chad A Mirkin, Robert L Letsinger, Robert C Mucic, and James J Storhoff. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature*, 382(6592):607–609, 1996.
- [224] Joshua A Anderson and Alex Travesset. Coarse-grained simulations of gels of nonionic multiblock copolymers with hydrophobic groups. *Macromolecules*, 39(15):5143–5151, 2006.
- [225] Joshua A Anderson, Chris Lorenz, and Alex Travesset. Micellar crystals in solution from molecular dynamics simulations. *The Journal of chemical physics*, 128(18):184906, 2008.
- [226] Vanessa Ortiz, Steven O Nielsen, Michael L Klein, and Dennis E Discher. Computer simulation of aqueous block copolymer assemblies: length scales and methods. *Journal of Polymer Science Part B: Polymer Physics*, 44(14):1907–1918, 2006.
- [227] Vanessa Ortiz Torres. *Soft matter systems simulated by molecular methods*. University of Pennsylvania, 2007.
- [228] Christopher M Wijmans, Eiser Eiser, and Daan Frenkel. Simulation study of intra-and intermicellar ordering in triblock-copolymer systems. *The Journal of chemical physics*, 120(12):5839–5848, 2004.

- [229] Alauddin Ahmed and Richard J Sadus. Phase diagram of the weeks-chandler-andersen potential from very low to high temperatures and pressures. *Physical Review E*, 80(6):061101, 2009.
- [230] Daan Frenkel. Simulations: The dark side. *The European Physical Journal Plus*, 128(1):1–21, 2013.
- [231] Harsha Mohan and Ranjini Bandyopadhyay. Phase behavior and dynamics of a micelle-forming triblock copolymer system. *Physical Review E*, 77(4):041803, 2008.
- [232] Simon Gaisford, Anthony E Beezer, John C Mitchell, Paul C Bell, Faluke Fakorede, Jill K Finnie, and Stuart J Williams. Temperature induced aggregation in aqueous solution of a series of PEO–PPO–PEO copolymers. *International journal of pharmaceuticals*, 174(1-2):39–46, 1998.
- [233] Gerard Riess. Micellization of block copolymers. *Progress in polymer science*, 28(7):1107–1170, 2003.
- [234] Wyn Brown, Karin Schillen, Mats Almgren, Soeren Hvidt, and Pratap Bahadur. Micelle and gel formation in a poly (ethylene oxide)/poly (propylene oxide)/poly (ethylene oxide) triblock copolymer in water solution: dynamic and static light scattering and oscillatory shear measurements. *The journal of physical chemistry*, 95(4):1850–1858, 1991.
- [235] Amitabha Chattopadhyay and Erwin London. Fluorimetric determination of critical micelle concentration avoiding interference from detergent charge. *Analytical biochemistry*, 139(2):408–412, 1984.
- [236] David A Edwards, Richard G Luthy, and Zhongbao Liu. Solubilization of polycyclic aromatic hydrocarbons in micellar nonionic surfactant solutions. *Environmental Science & Technology*, 25(1):127–133, 1991.
- [237] Yu Ling Wu, Rudolf Sprik, Wilson CK Poon, and Erika Eiser. Effect of salt on the phase behaviour of F68 triblock PEO/PPO/PEO copolymer. *Journal of Physics: Condensed Matter*, 18(19):4461, 2006.
- [238] Alexander Stukowski. Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. *Modelling and Simulation in Materials Science and Engineering*, 18(1):015012, December 2009.
- [239] Dingning Li and Kai Zhang. Unifying the concepts of scattering and structure factor in ordered and disordered samples. *Journal of Applied Crystallography*, 54(2):644–660, 2021.
- [240] Daan Frenkel, RJ Vos, CG De Kruif, and A Vrij. Structure factors of polydisperse systems of hard spheres: A comparison of Monte Carlo simulations and Percus–Yevick theory. *The Journal of chemical physics*, 84(8):4625–4630, 1986.
- [241] Mark A Miller and Daan Frenkel. Simulating colloids with baxter’s adhesive hard sphere model. *Journal of Physics: Condensed Matter*, 16(42):S4901, 2004.

- [242] Rob Eppenga and Daan Frenkel. Monte Carlo study of the isotropic and nematic phases of infinitely thin hard platelets. *Molecular Physics*, 52(6):1303–1334, August 1984.
- [243] Daan Frenkel and Rob Eppenga. Evidence for algebraic orientational order in a two-dimensional hard-core nematic. *Physical Review A*, 31(3):1776–1787, March 1985.
- [244] Dieter Forster. *Hydrodynamic fluctuations, broken symmetry, and correlation functions*. CRC Press, 2018.
- [245] Raymond D. Mountain and Th.W. Ruijgrok. Monte-Carlo study of the Maier-Saupe model on square and triangle lattices. *Physica A: Statistical Mechanics and its Applications*, 89(3):522–538, December 1977.
- [246] Alan K Soper. Orientational correlation function for molecular liquids: The case of liquid water. *The Journal of Chemical Physics*, 101(8):6888–6901, October 1994.
- [247] Leonard S Ornstein. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci.*, 17:793, 1914.
- [248] Jerome K Percus and George J Yevick. Analysis of classical statistical mechanics by means of collective coordinates. *Physical Review*, 110(1):1, 1958.
- [249] Michael S. Wertheim. Exact solution of the Percus-Yevick integral equation for hard spheres. *Physical Review Letters*, 10(8):321, 1963.
- [250] Everett Thiele. Equation of state for hard spheres. *The Journal of Chemical Physics*, 39(2):474–479, 1963.
- [251] David J Kinning and Edwin L Thomas. Hard-sphere interactions between spherical domains in diblock copolymers. *Macromolecules*, 17(9):1712–1718, 1984.
- [252] Neil W Ashcroft and Ji Lekner. Structure and resistivity of liquid metals. *Physical Review*, 145(1):83, 1966.
- [253] Eric W Kaler, Karl E Bennett, H Ted Davis, and LE Scriven. Toward understanding microemulsion microstructure: A small-angle x-ray scattering study. *The Journal of chemical physics*, 79(11):5673–5684, 1983.
- [254] Jonathan F Berengut, Julian C Berengut, Jonathan PK Doye, Domen Prešern, Akihiro Kawamoto, Juanfang Ruan, Madeleine J Wainwright, and Lawrence K Lee. Design and synthesis of pleated DNA origami nanotubes with adjustable diameters. *Nucleic acids research*, 47(22):11963–11975, 2019.
- [255] Bezuayehu Teshome, Stefan Facsko, and Adrian Keller. Topography-controlled alignment of DNA origami nanotubes on nanopatterned surfaces. *Nanoscale*, 6(3):1790–1796, 2014.
- [256] Florence Benn, Natalie EC Haley, Alexandra E Lucas, Emma Silvester, Seham Helmi, Robert Schreiber, Jonathan Bath, and Andrew J Turberfield. Chiral DNA origami nanotubes with well-defined and addressable inside and outside surfaces. *Angewandte Chemie*, 130(26):7813–7816, 2018.

- [257] Saskia Groer and Andreas Walther. Switchable supracolloidal 3D DNA origami nanotubes mediated through fuel/antifuel reactions. *Nanoscale*, 12(32):16995–17004, 2020.
- [258] Hieu Bui, Craig Onodera, Carson Kidwell, YerPeng Tan, Elton Graugnard, Wan Kuang, Jeunghoon Lee, William B Knowlton, Bernard Yurke, and William L Hughes. Programmable periodicity of quantum dot arrays with DNA origami nanotubes. *Nano letters*, 10(9):3367–3372, 2010.

A | DNA Sequences

A.1 Y-shape+Linkers

Y1 [12 + 28]: 5'- CGA TTG ACT CTC CAC GCT GTC CTA ACC ATG ACC GTC GAA G -3'

Y2 [12 + 28]: 5'- CGA TTG ACT CTC CTT CGA CGG TCA TGT ACT AGA TCA GAG G -3'

Y3 [12 + 28]: 5'- CGA TTG ACT CTC CCT CTG ATC TAG TAG TTA GGA CAG CGT G -3'

L0-1[12 + 26]:

5'- GAG AGT CAA TCG TCT ATT CGC ATG ACA TTC ACC GTA AG -3'

L0-2 [12 + 26]:

5'- GAG AGT CAA TCG CTT ACG GTG AAT GTC ATG CGA ATA GA -3'

L6-1 [12+ 6 + 26]:

5'- GAG AGT CAA TCG TTT TTT TCT ATT CGC ATG ACA TTC ACC GTA AG -3'

L6-2 [12+ 6 + 26]:

5'- GAG AGT CAA TCG TTT TTT CTT ACG GTG AAT GTC ATG CGA ATA GA -3'

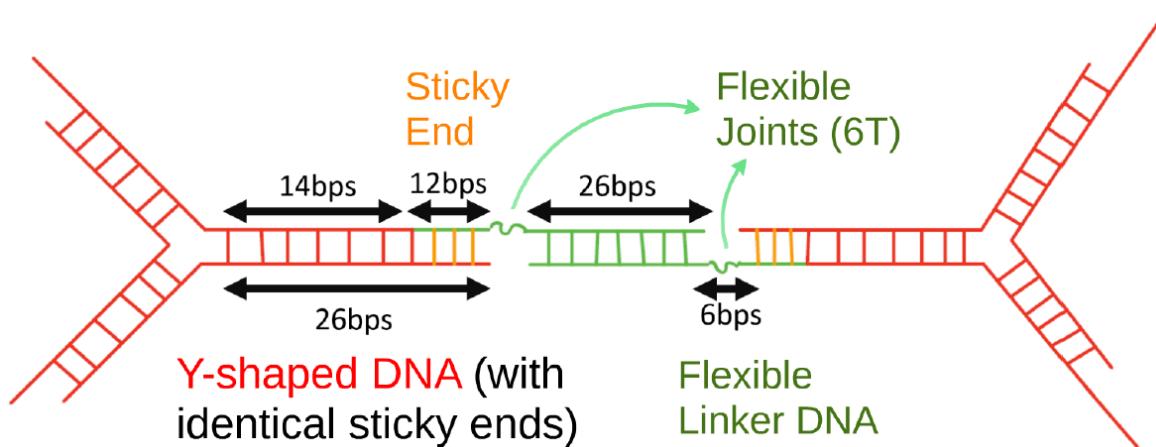


Fig. A.1 A detailed schematic of Y-shape + L6 linker DNA.

A.2 Nunchucks DNA

F0 (Full DNA): [60//60]

5' - GAG GTG GAG TAA GTG AAT GAA GGT GTG TGA AGA CAA ACA GAG AGG AGA GAA AGA
GAG TTA -3'

3' - CTC CAC CTC ATT CAC TTA CTT CCA CAC ACT TCT GTT TGT CTC TCC TCT CTT TCT
CTC AAT -5'

N1 (Asymmetry Nunchucks - 7 shape DNA) : [60//15+40]

5' - GCG GCG GAG GAA GCG AAA GAA GGA GAG AAG AGA CAA AGA TAG AAG AGA TAG AGA
GAA AGA -3'

3' - CGC CGC CTC CTT CGC ----- T CCT CTC TTC TCT GTT TCT ATC TTC TCT ATC TCT
CTT TCT -5'

N2 (Nunchucks DNA with short flexible joint -5 base joint) : [60//27+28]

5' - GAG GTG GAG TAA GTG AAT GAA GGT GTG TGA AGA CAA ACA GAG AGG AGA GAA AGA
GAG TTA -3'

3' - CTC CAC CTC ATT CAC TTA CTT CCA CAC ----- T GTT TGT CTC TCC TCT CTT TCT
CTC AAT -5'

N3 (Nunchucks DNA with long flexible joint -18 base joint) : [60//21+21]

5' - GAG GTG GAG TAA GTG AAT GAA GGT GTG TGA AGA CAA ACA GAG AGG AGA GAA AGA
GAG TTA -3'

3' - CTC CAC CTC ATT CAC TTA CTT ----- ----- CTC TCC TCT CTT TCT
CTC AAT -5'

N4m (Nunchucks DNA with 5 base mismatching in the middle) : [60//27+5*+28]

5' - GAG GTG GAG TAA GTG AAT GAA GGT GTG TGA AGA CAA ACA GAG AGG AGA GAA AGA
GAG TTA -3'

3' - CTC CAC CTC ATT CAC TTA CTT CCA CAC AGA CGA GTT TGT CTC TCC TCT CTT TCT
CTC AAT -5'

B | Algorithms

B.1 Visualisation of oxDNA in LAMMPS

In the current version of oxDNA, the illustrations are described by two files: a configuration file and a topology file. The configuration file contains all the possible orientations and positions of each nucleotide. The topology file keeps track of the backbone-backbone bonds between nucleotides within the same strand [88]. In contrast, in the embedded version of oxDNA in LAMMPS, a trajectory file can be directly stored, but the configuration and topology files need to be reconstructed. Hence, I have written a home-made script to customise the LAMMPS trajectory file as a configuration file and a topology file. Then, it is possible to further use the customised files to generate a PDB format, used in oxDNA. The final visualisation of structures is generated with the PDB format presented in UCSF chimera¹. Since the model is coarse-grained, the presented position of each nucleotide in the model is an actual atom. The backbone-based orientation is determined by the rotation matrix calculated from the quaternion algorithm. The quaternion (q) is a convenient representation of the orientation of a nucleotide. By using the quaternion representation, we can rewrite the nine-element rotation matrix into a four-element vector. This quaternion information can also be acquired from LAMMPS output files. The rotation matrix then can be reconstructed from the quaternion information to produce a 3D structure of the backbone and nucleotides. Quaternions are generally represented in the form: $q = a + bi + cj + dk$, where a, b, c , and d are real numbers ($a^2 + b^2 + c^2 + d^2 = 1$ in a unit quaternion), and i, j , and k are the fundamental quaternion units. The rotation matrix (M_{rot}) in a 3D space can then be determined by a quaternion, $q(a, b, c, d)$, in the form:

$$M_{rot} = \begin{bmatrix} 2a^2 + 2b^2 - 1 & 2bc + 2ad & 2bd - 2ac \\ 2bc - 2ad & 2a^2 + 2c^2 - 1 & 2cd + 2ab \\ 2bd + 2ac & 2cd - 2ab & 2a^2 + 2d^2 - 1 \end{bmatrix} = \begin{bmatrix} b_x & n_x & / \\ b_y & n_y & / \\ b_z & n_z & / \end{bmatrix}$$

¹UCSF Chimera is a highly extensible program for interactive visualisation and analysis of molecular structures and related data.

- **Configuration file** contains the position of the centre of mass and orientation of a single nucleotide in the following order: [position (r_x, r_y, r_z) , backbone-base orientation (b_x, b_y, b_z) , normal direction (n_x, n_y, n_z)], where r_i is stored in a LAMMPS trajectory file, b_i and n_i are calculated from the quaternion vector.
- **Topology file** stores the intra-strand with a fixed bonding topology (e.g. which nucleotides share backbone links). It also records the order of 3' and 5' of the nucleotides in each strand.

B.2 Nematic Order Parameter

Here we outline the motivation and calculation of the nematic order parameter used throughout this report, based on the work of Eppenga and Frenkel [242, 143]. Conventionally, the average second Legendre polynomial $\langle P_{2l}(\cos(\theta)) \rangle$ is used to characterise the nematic order of a system. Averaging over a population of N molecules, we may write the nematic order parameter S_n explicitly as:

$$S_n = \frac{1}{N} \sum_{i=1}^N \left(\frac{3}{2} \cos^2(\theta_i) - \frac{1}{2} \right). \quad (\text{B.1})$$

This expression however relies on knowledge of the system-wide nematic director (i.e. the axis of symmetry of the cylindrical phase), to define θ_i . This is not always possible in physical systems where such a unique direction is not externally imposed.

Instead, as detailed by Frenkel et al. [243], we maximise the expression:

$$S'_n(\hat{n}') = \frac{1}{N} \left[\sum_{i=1}^N \left(\frac{3}{2} (\hat{n}' \cdot \hat{u}_i)^2 - \frac{1}{2} \right) \right], \quad (\text{B.2})$$

where \hat{u}_i denotes the orientation of the individual molecular axes in the laboratory frame, and \hat{n}' is the direction of common alignment, known as the director. In the absence of an electric field, this direction is arbitrary, and determined in practice by infinitesimal perturbations to the system through spontaneous symmetry breaking [244]. (B.2) may be further simplified to:

$$S'_n = \frac{1}{N} \langle \hat{n}' \cdot \mathbf{Q} \cdot \hat{n}' \rangle, \quad \text{where } \mathbf{Q}_i = \frac{3}{2} \hat{u}_i \hat{u}_i - \frac{1}{2} \mathbf{I}. \quad (\text{B.3})$$

The tensor order parameter $\langle \mathbf{Q} \rangle$ is a traceless symmetric 2nd-rank tensor, with three eigenvalues $\lambda_+, \lambda_0, \lambda_-$ [242]. We typically take the largest eigenvalue (λ_+) as the nematic order parameter, a good approximation in limit of large N . In practice, I calculate the eigenvalues of the related tensor \mathbf{M} :

$$\mathbf{M} = \frac{1}{N} \sum_{i=1}^N \hat{u}_i \hat{u}_j \quad (\text{B.4})$$

as this shares eigenvectors with \mathbf{Q} , and has eigenvectors μ_n related to λ_n by: $\mu_n = 2/3\lambda_n + 1/3$.

It is worth noting that λ_+ is bound above zero, and so does not reach zero in the isotropic phase as would be expected. It is common to use $S = -2\lambda_0$ when considering such disordered systems, as this fluctuates around an average much closer to zero [245]. I have not followed this usage in the results presented here, to give continuity in the order parameter over the transition (wherein lies the focus of this report), however this has resulted in an average order parameter in the isotropic phase slightly above zero.

B.3 Position Dependant Order Parameter

The pairwise orientational correlation function provides a position-dependent equivalent to the nematic order parameter. For use in simulation, we write this for N particles in the form:

$$g_l(r) = \frac{\sum_{i=1}^N \sum_{j \neq i} P_l(\hat{u}_i \cdot \hat{u}_j) \Delta(r_{ij} - r)}{\sum_{i=1}^N \sum_{j \neq i} \Delta(r_{ij} - r)}, \quad (\text{B.5})$$

where $\Delta(r_{ij} - r) = 1$ if r_{ij} is inside a spherical shell between r and $r + \delta$, and 0 otherwise. The thickness of δ is chosen to maximise the resolution of the function, as too large a value will ‘wash out’ important features, while still maintaining a reasonable sample size. This method is computationally expensive however, as the number of times P_L must be calculated scales as $\mathcal{O}(N^2)$. Furthermore, the storage of all necessary angles is prohibitively expensive (even on modern computers), and limits the resolution possible [194]. We may, however, avoid this problem by expanding the function in terms of spherical harmonics, and sum the contributions for all pairs to a given molecule [246]. This is most easily achieved through application of the spherical harmonics addition theorem:

$$P_l(\cos \theta_{ij}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l \mathcal{Y}_{lm}(\hat{u}_i) \mathcal{Y}_{lm}^*(\hat{u}_j). \quad (\text{B.6})$$

Using this, recognising that $\hat{u}_i \cdot \hat{u}_j = \cos \theta_{ij}$ and writing $\Delta(r_{ij} - r) = \Delta_r$ for simplicity, we may write the contribution from all particles around particle i as:

$$\sum_{j \neq i} P_l(\cos \theta_{ij}) \Delta_r = \frac{4\pi}{2l+1} \sum_{j \neq i} \sum_{m=-l}^l \mathcal{Y}_{lm}(\hat{u}_i) \mathcal{Y}_{lm}^*(\hat{u}_j) \Delta_r. \quad (\text{B.7})$$

Exchanging the order of summation, and removing $\mathcal{Y}_{lm}(\hat{u}_i)$ from the sum over j gives:

$$\sum_{i \neq j} P_l(\cos \theta_{ij}) \Delta_r = \frac{4\pi}{2l+1} \sum_{m=-l}^l \mathcal{Y}_{lm}(\hat{u}_i) \sum_{j \neq i} \mathcal{Y}_{lm}^*(\hat{u}_j) \Delta_r. \quad (\text{B.8})$$

This may be simplified further to:

$$\sum_{i \neq j} P_l(\cos \theta_{ij}) \Delta_r = \frac{4\pi}{2l+1} \sum_{m=-l}^l \mathcal{Y}_{lm}(\hat{u}_i) \times C_{lm}^* \quad (\text{B.9})$$

where we have defined

$$C_{lm}^* = \sum_{j \neq i} \mathcal{Y}_{lm}^*(\hat{u}_j) \Delta_r. \quad (\text{B.10})$$

This removes the requirement to calculate P_l for every pair of molecules, in favour of simply summing the spherical harmonics terms corresponding to a given shell around i , then multiplying by the spherical harmonic components. It should also be noted that (B.9) is invariant under changes in the coordinate frame; in our computations all orientations were expressed in the lab frame for each, rather than aligning the coordinate frame with the global director.

B.4 Structure factor for hard spheres

This section serves as a verification of the methods for computing $S(q)$, which are outlined in Section 6.3.6. We will first look at the origin of the analytical expression for $S(q)$ for hard spheres. By the Ornstein-Zernike relation[247], the total correlation function $h(r_{12}) \equiv g(r_{12}) - 1$ is given as

$$h(r_{12}) = c(r_{12}) + n \int c r_{13} h(r_{32}) d\mathbf{r}_3, \quad (\text{B.11})$$

where $c(r_{12})$ is the direct correlation function and the second term is an indirect contribution from neighboring particles. In the Percus–Yevick approximation[248], the direct correlation function is expressed as

$$c(r) \simeq (e^{-\phi(r)/kT} - 1) e^{\phi(r)/kT} g(r), \quad (\text{B.12})$$

where ϕ is the hard sphere potential. This has been shown to have the polynomial solution [249–251]

$$c(r) = -\alpha - \beta \frac{r}{2R} - \gamma \left(\frac{r}{2R} \right)^3 \quad (\text{B.13})$$

where

$$\begin{aligned} \alpha &= \frac{(1+2\eta)^2}{(1-\eta)^4} \\ \beta &= -\frac{6\eta(1+\eta/2)^2}{(1-\eta)^4} \\ \gamma &= \frac{1}{2} \frac{\eta(1+2\eta)^2}{(1-\eta)^4} \end{aligned} \quad (\text{B.14})$$

and the hard-sphere volume fraction $\eta = \frac{4}{3}\pi R^3 \rho$. With Eqn. B.11, the structure factor from Eqn. 6.23 may be expressed as

$$S(q) = \frac{1}{1 - \rho \hat{c}(q)}, \quad (\text{B.15})$$

where $\hat{c}(q)$ is the Fourier transform of the direct correlation function. For an isotropic material, this can be expressed by the Debye scattering equation, such that

$$\hat{c}(q, R, \eta) = 4\pi \int_0^{2R} c(r) \frac{\sin(qr)}{qr} r^2 dr. \quad (\text{B.16})$$

This has the analytical solution[252, 253]

$$S(q, R, \eta) = \frac{1}{1 + 24\eta G(A)/A}, \quad (\text{B.17})$$

where $A = 2qR$ and

$$\begin{aligned} G(A) &= \frac{\alpha}{A^2} (\sin(A) - A \cos(A)) + \\ &\quad \frac{\beta}{A^3} (2A \sin(A) + (2 - A^2) \cos(A) - 2) + \\ &\quad \frac{\gamma}{A^5} (-A^4 \cos(A) + 4[(3A^2 - 6) \cos(A) + (A^3 - 6A) \sin(A) + 6]). \end{aligned} \quad (\text{B.18})$$

To verify the computational methods for obtaining the structure factor presented in Section 6.3.6, the scripts were tested on trajectory files from a small system of 125 particles with the WCA potential with $\varepsilon = 1.0$, $\sigma = 1.0$ and a density of $\rho = 0.50$. From Fig. B.1, we see that the curve from Eqn. B.17 with $R = 0.5$ and $\eta = \pi/12$ matches well with the computational estimates. While the method with 1D FFT along crystal planes is significantly faster than the method with 3D FFT, it also requires more trajectories to average over to

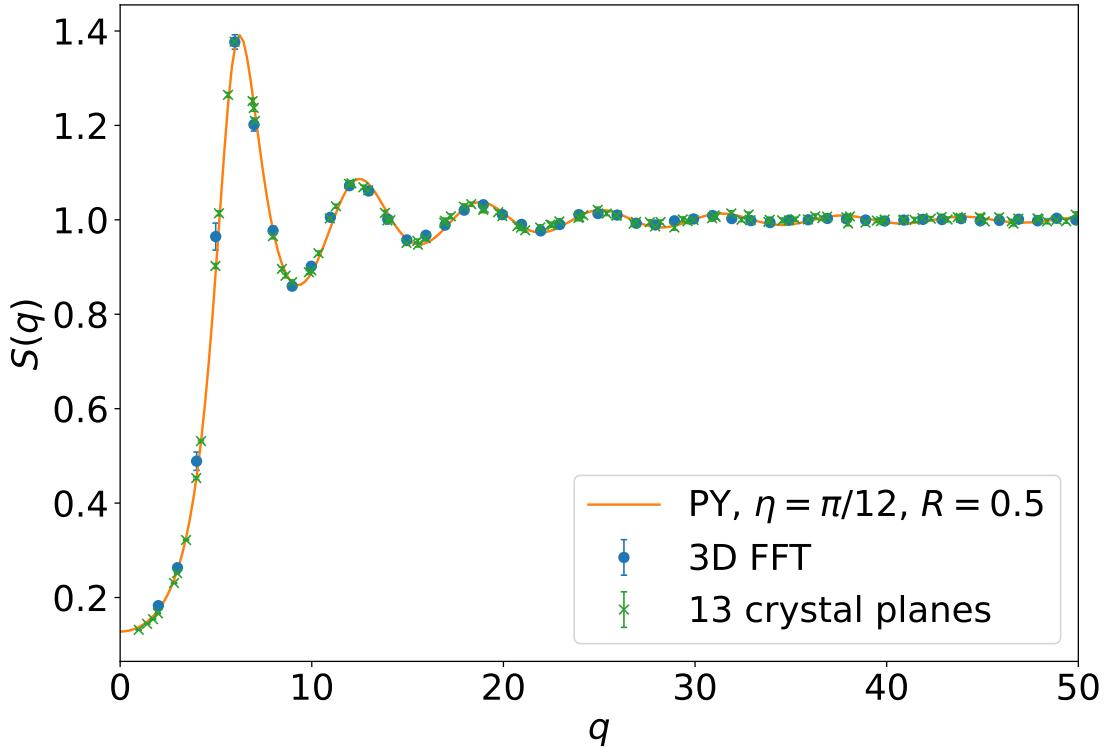


Fig. B.1 The structure factor for a system with 125 particles with the WCA potential. The orange line is the analytical solution with the Percus-Yevick approximation from Eqn. B.17, and the blue and green points are the estimated $S(q)$ by 3D FFT and 1D FFT along crystal planes respectively. The error bars give the standard error.

obtain comparable statistics. In the data presented in Fig. B.1, $S(q)$ is estimated from 101 frames with 3D FFT, while for the method with 1D FFT along crystal planes, it is estimated from 10001 frames. The latter method is still about 10 times faster, and the error is smaller.

C | Codes and Scripts

C.1 oxDNA Scripts in LAMMPS

```
# Parameter setting
variable number equal 1

units lj # all quantities are unitless.
dimension 3 # simulate in a 3D box
# Turn Newton's third law off for pairwise and bonded
# interactions
newton off
# Set the style of boundaries for the global simulation box
# in each dimension, p represents 'periodic'
boundary p p p

# Define the style of atoms to use in a simulation
atom_style hybrid bond ellipsoid
atom_modify sort 0 1.0

# Pair interactions require lists of neighbours to be
# calculated
neighbor 3.0 bin
neigh_modify every 10 delay 0 check yes
# (atom pairs within a neighbor cutoff distance equal to
# the force cutoff plus the skin distance)

# Restart a run from the last running file
restart 1000000 oxDNA.Restart
```

```
read_data Y_L6_key_shape3.txt

set atom * mass 3.1575

group all type 1 4

# oxDNA bond interactions - FENE backbone
bond_style oxdna2/fene
bond_coeff * 2.0 0.25 0.7564

# oxDNA pair interactions
pair_style hybrid/overlay oxdna2/excv oxdna2/stk oxdna2/
hbond oxdna2/xstk oxdna2/coaxstk oxdna2/dh
pair_coeff * * oxdna2/excv           2.0 0.7 0.675 2.0
0.515 0.5 2.0 0.33 0.32
pair_coeff * * oxdna2/stk           seqdep 0.1 6.0 0.4 0.9 0.32
0.6 1.3 0 0.8 0.9 0 0.95 0.9 0 0.95 2.0 0.65 2.0 0.65
pair_coeff * * oxdna2/hbond         seqdep 0.0 8.0 0.4 0.75 0.34
0.7 1.5 0 0.7 1.5 0 0.7 1.5 0 0.7 0.46
3.141592653589793 0.7 4.0 1.5707963267948966 0.45 4.0
1.5707963267948966 0.45
pair_coeff 1 4 oxdna2/hbond         seqdep 1.0678 8.0 0.4 0.75
0.34 0.7 1.5 0 0.7 1.5 0 0.7 1.5 0 0.7 0.46
3.141592653589793 0.7 4.0 1.5707963267948966 0.45 4.0
1.5707963267948966 0.45
pair_coeff 2 3 oxdna2/hbond         seqdep 1.0678 8.0 0.4 0.75
0.34 0.7 1.5 0 0.7 1.5 0 0.7 1.5 0 0.7 0.46
3.141592653589793 0.7 4.0 1.5707963267948966 0.45 4.0
1.5707963267948966 0.45
pair_coeff * * oxdna2/xstk          47.5 0.575 0.675
0.495 0.655 2.25 0.791592653589793 0.58 1.7 1.0 0.68 1.7
1.0 0.68 1.5 0 0.65 1.7 0.875 0.68 1.7 0.875 0.68
pair_coeff * * oxdna2/coaxstk        58.5 0.4 0.6 0.22
0.58 2.0 2.891592653589793 0.65 1.3 0 0.8 0.9 0 0.95 0.9
0 0.95 40.0 3.116592653589793
pair_coeff * * oxdna2/dh            0.1 0.2 0.815
```

```
# NVE ensemble
fix 1 all    nve/dotc/langevin 0.112 0.112 1.0 735287 angmom
           10
timestep 1e-2

# Set up output files
dump pos all xyz 250000 dump.trajectory.xyz
dump_modify pos element H O N C append yes # represents A,
           T ,G, C in DNA

compute quat all property/atom quatw quati quatj quatk
compute shape all property/atom shapex shapey shapez
variable ss atom 2*c_shape[1]

# Record the rotational and quaternion information of each
# nucleotide
dump quate all custom 250000 quat*.data id v_ss v_ss v_ss
           c_quat[1] c_quat[2] c_quat[3] c_quat[4]
dump_modify quate sort id
dump_modify quate format line "%d %.16le %.16le %.16le %.16
           le %.16le %.16le %.16le"

# Record position information sorted by atom id
dump out all custom 250000 IN*.data id mol type x y z ix iy
           iz
dump_modify out sort id
dump_modify out format line "%d %d %d %.5f %.5f %.5f %d %d
           %d"

# Calculate mean-squared displacement
compute          MSDs all msd
# Calculate radius of gyration
compute          GYRs all gyration
thermo 10000
thermo_style custom step temp c_MSDs[4] c_GYRs cpu
```

```

variable rg equal c_GYRs
fix 4 all print 250000 "$step" $(c_GYRs)" screen no
append radius_of_gyration.txt

shell mkdir tprimedna
shell cd tprimedna

run 500000000

```

C.2 LAMMPS scripts of Y-shaped DNA with linkers

```

# Y-shaped DNA model: large system with patches
# 10 atoms per molecule
# to acquire the equilibrium status from the situation
# dat file re-generated input

#####      VARIABLES      #####
variable T      equal      0.1
variable N      equal      10
variable dim    equal      10
variable ep     equal      4
variable run_steps equal 2000000

#####      PARAMETERS      #####
units lj
atom_style molecular
neighbor 0.4 bin
neigh_modify every 1 delay 1 check yes
comm_modify vel yes

#####      READ      #####
read_data input_data_key_linker_${N}_${dim}.file

#####      pair lj/cut ONLY #####

```

```
pair_style lj/cut 1.12
pair_modify shift yes
pair_coeff * * 0 0.0 0.0
pair_coeff 1 1 1 1.0 1.12 # excluded volume
pair_coeff 1 2 1 1.0 1.12 # excluded volume
pair_coeff 2 2 1 1.0 1.12 # excluded volume
pair_coeff 3 3 1 0.6 0.67 # repulsive
pair_coeff 4 4 1 0.6 0.67 # repulsive
pair_coeff 3 4 ${ep} 0.2 0.50 # attractive

##### BONDS #####
bond_style harmonic
special_bonds lj 0 1 1

bond_coeff 1 300.0 0.96      # bead-bead bonds
bond_coeff 2 300.0 0.56      # bead-patch bonds

##### ANGLES #####
angle_style harmonic

# angle in one molecule
angle_coeff 1 100.0 120
angle_coeff 2 100.0 180
angle_coeff 3 100.0 102
angle_coeff 4 100.0 108

##### THERMAL #####
timestep 0.005
thermo 10000
thermo_style custom time temp #step temp time pe press
atoms vol
thermo_modify norm no

##### INITIAL #####
reset_timestep 0
```

```

#####      FIX      #####
fix 4 all nve
fix 5 all langevin ${T} ${T} 1 123456
fix recen1 all recenter INIT INIT INIT

##### COMPUTE + OUTPUT pairs, bonds and angles #####
dump 1 all custom 10000 output_${T}_10e4.dump id mol type
    x y z vx vy vz
compute 1 all property/local patom1 patom2
compute 2 all pair/local dist eng force
dump 2 all local 10000 pairlocal_${T}_10e4.dump c_1[1] c_1
    [2] c_2[1] c_2[2] c_2[3]

#####      RUN      #####
run ${run_steps}
#shell cp log.lammps log_A_${A}_omega_${om}.lammps

##### Write dat files #####
write_data after_run.dat
write_restart after_run.reset

```

C.3 LAMMPS scripts of Nunchuck-DNA

```

#####      VARIABLES      #####
variable T      equal 0.50 #0.10 0.36 0.70
variable num     equal 1
variable N      equal 100 #200 # 250 300 350
variable box     equal 50
variable t_step   equal 0.005
variable mix_steps equal 200000
variable mix_tau   equal ${t_step}*${mix_steps}
variable run_steps equal 200000
variable run_tau   equal ${t_step}*${run_steps}

#####      PARAMETERS      #####

```

```
units lj
atom_style molecular
neighbor 0.4 bin
neigh_modify every 1 delay 1 check yes
comm_modify vel yes

##### READ #####
read_data input_data_nunchucks_${N}_${box}.file

##### BONDS #####
bond_style harmonic
special_bonds lj 0 1 1
bond_coeff * 500.0 0.98 # the wildcard (*) means to have
the same for all bond types.
bond_coeff 5 50.0 0.98

##### ANGLES #####
angle_style harmonic
angle_coeff * 500.0 180
angle_coeff 4 0.1 180
angle_coeff 5 0.1 180

##### PAIR lj/cut --with interactions, to steady state #####
pair_style lj/cut 1.0
pair_modify shift yes
pair_coeff * * 1 1.0 1.12

##### THERMAL #####
timestep ${t_step}
thermo 100
thermo_style custom step time temp press pe ke etotal vol
thermo_modify norm no
reset_timestep 0

##### FIX #####
fix 2 all nph iso 0.5 0.5 100
```

```

fix 3 all langevin 0.5 0.5 1 123456
#run 50000

#unfix 3
#fix 4 all nve
#fix 5 all langevin ${T} ${T} 1 123456
fix recen1 all recenter INIT INIT INIT

##### DUMP OUTPUT #####
dump 1 all custom 1000 output.dump id mol type x y z vx
vy vz

# FIX for RDF
comm_modify cutoff 25.0
compute 3 all rdf 500 5 6 cutoff 24.0
fix 3 all ave/time 1 ${run_steps} ${run_steps} c_3 [*]
file rdf_T_${T}_N_${N}_box_${box}_tau_${run_tau}.rdf
mode vector

##### RUN #####
run ${run_steps}

##### SHELL #####
shell cp output.dump          output_T_${T}_N_${N}_box_${box}
_run_${run_tau}_tau_num_${num}.dump
shell touch T_${T}.done

```

C.4 LAMMPS scripts of DNA functionlised F108 Pluronics

```

# simulate DNA functionalised F108 Pluronics
# alpha_T and r_c control the temperature sensitive part of
# the micellisation process
# beta_T controls the melting temperature of DNA patches

##### VARIABLES #####

```

```
variable      T      index      0.36 0.35 0.34 0.33 0.32 0.31
              0.30 0.29 0.28 0.27
variable  alpha_T index      5.2   4.0   3.0   2.0   1.3   1.1   1.0
              1.0   1.0   1.0
variable      r_c    index      2.40 2.25 2.10 2.00 1.80 1.55
              1.40 1.30 1.22 1.12
variable  beta_T  equal      3.2 # 4.8 for higher DNA melting
temperature

#####      PARAMETERS      #####
units  lj
atom_style molecular
special_bonds lj/coul 0 1 1
neighbor 0.4 bin
neigh_modify every 1 delay 1
comm_modify vel yes
bond_style fene
bond_coeff * 30.0 1.5 0.96 0.96
angle_style cosine/delta
angle_coeff * 0 180
timestep 0.01
thermo 30000
thermo_modify norm no
thermo_style custom step temp press atoms vol

#####      READ      #####
read_data
  input_data_nunchucks_num1000_boxwidth400_with_patch.file

#####      PAIR  PARAMETERS      #####
pair_style lj/cut 5
pair_modify shift yes
#pair_coeff * * 0 0.0 0.0
pair_coeff 1 1 ${alpha_T} 1.0 ${r_c} # temperature
  sensitive
pair_coeff 1 2 1.0 1.0 1.12 # excluded volume
```

```
pair_coeff 2 2 1.0 1.0 1.12 # excluded volume
pair_coeff 3 3 1.0 0.6 0.67 # repulsive
pair_coeff 4 4 1.0 0.6 0.67 # repulsive
pair_coeff 3 4 ${beta_T} 0.2 0.50 # attractive

##### INITIAL #####
reset_timestep 0

##### ENSEMBLE CONTROL NPT/NVT #####
fix 1 all nve/limit 0.01
run 1000

unfix 1
fix 6 all nph iso 1 1 100
fix 7 all langevin 1 0.00001 1 123456
fix recen1 all recenter INIT INIT INIT
run 29000 #check the box size at this point is 39.9814

unfix 6
unfix 7
fix 9 all nve
fix 10 all langevin ${T} ${T} 1 123456
fix recen2 all recenter INIT INIT INIT
run 5000000

##### OUTPUT DUMP FILE #####
dump id1 all custom 1000 file.dump id mol type x y z
shell cp file.dump output_T_${T}.dump
shell touch T_${T}.done

##### REPEAT FOR NEXT PARAMETER #####
clear
next T
jump next.run
```

D | Extra Plots

D.1 Schematics in Different Stage of Melting Process

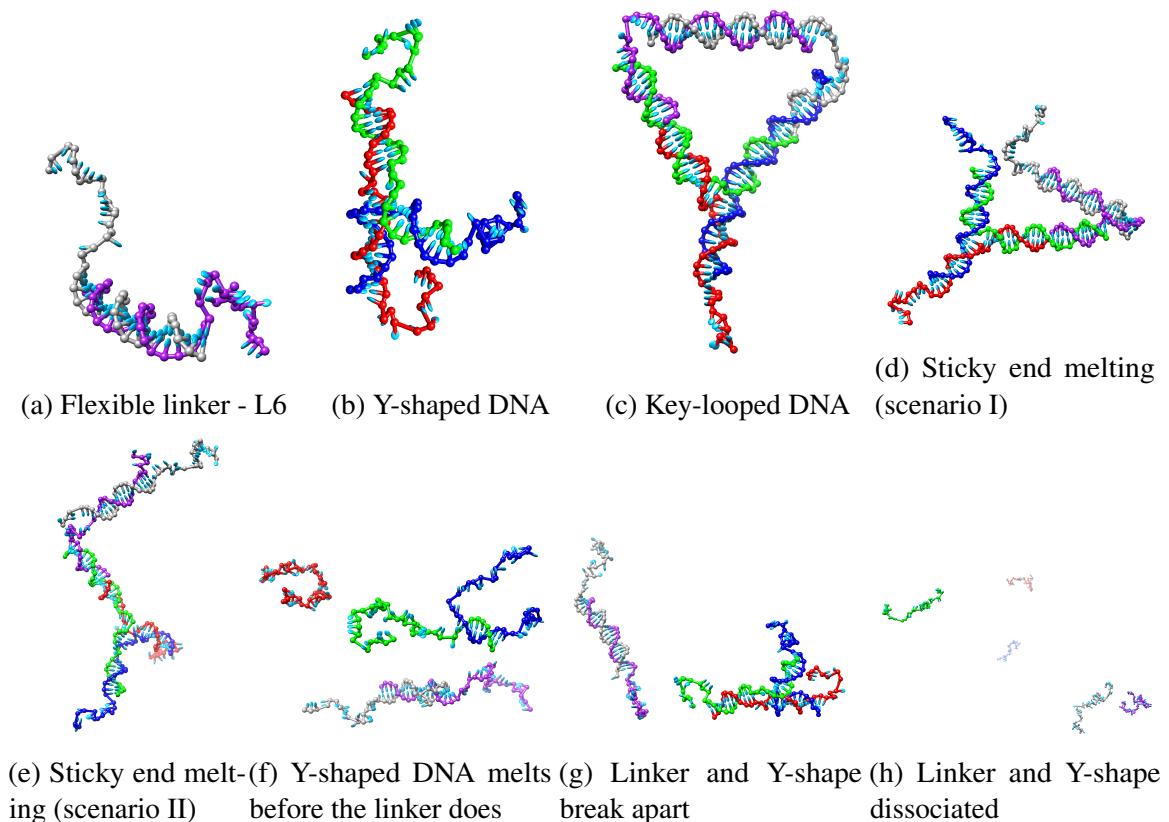


Fig. D.1 Schematics in different stages of melting process for Y+L6 system.

D.2 Open and Close Ratio of Y+L6

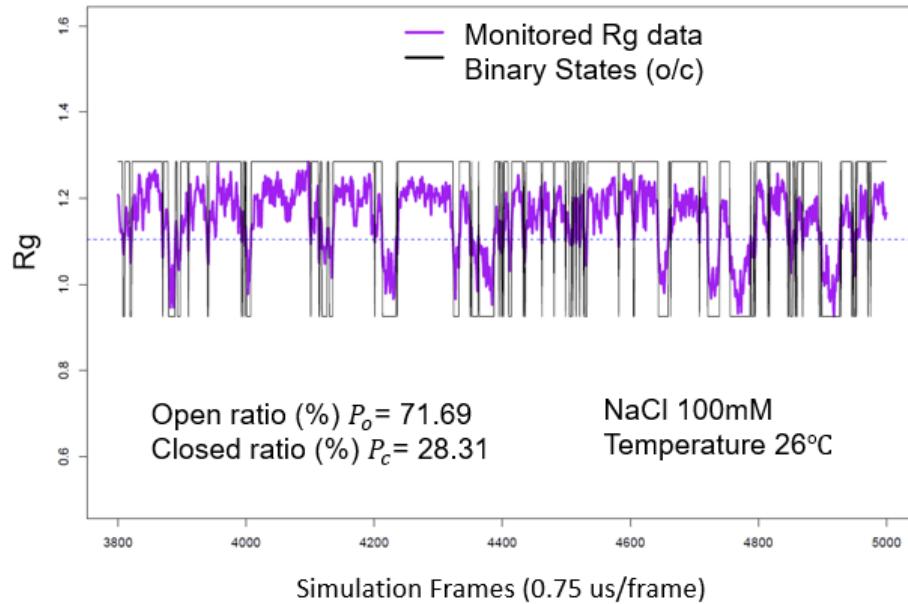


Fig. D.2 Monitored R_g of open and close ratio of Y+L6 sysytem at 26 °C, NaCl = 100nM.

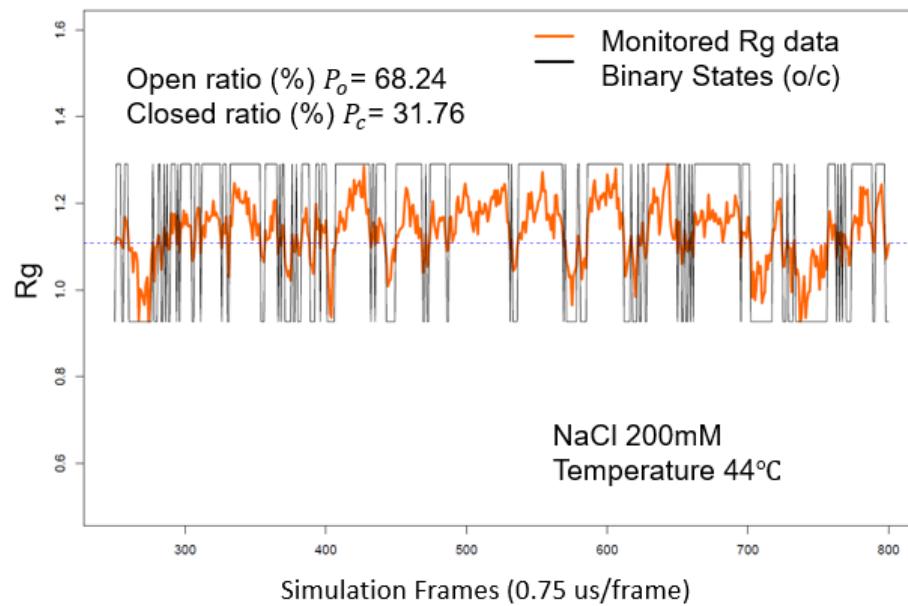


Fig. D.3 Monitored R_g of open and close ratio of Y+L6 sysytem at 44 °C, NaCl = 200nM.

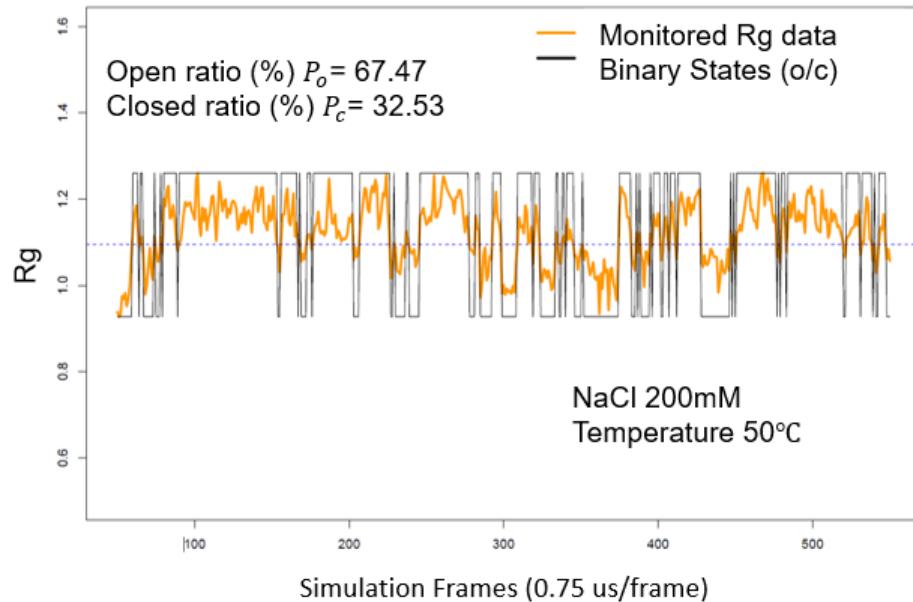


Fig. D.4 Monitored R_g of open and close ratio of Y+L6 sysytem at 50 °C, NaCl = 200nM.

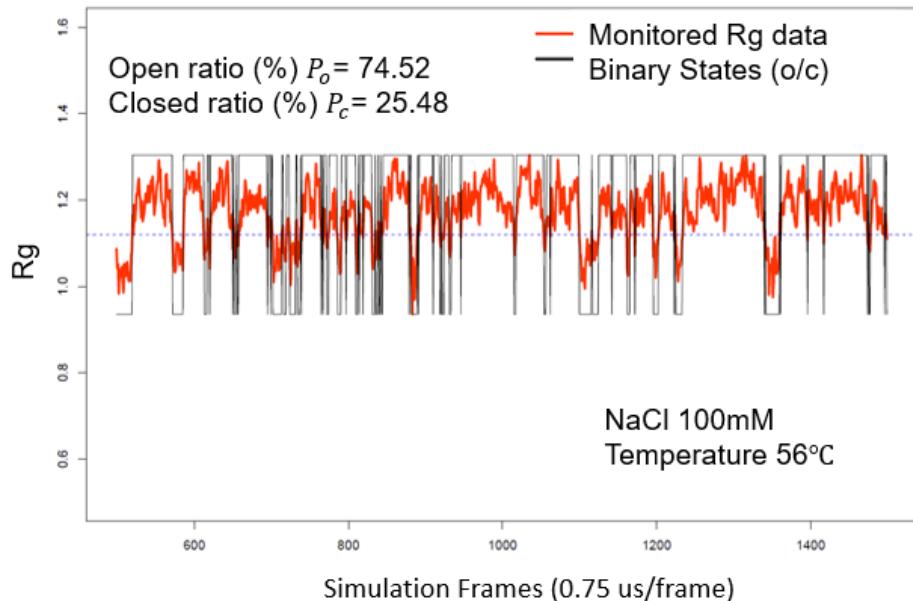


Fig. D.5 Monitered R_g of open and close ratio of Y+L6 sysytem at 56 °C, NaCl = 100nM.

D.3 Example snapshots of F108-DNA system

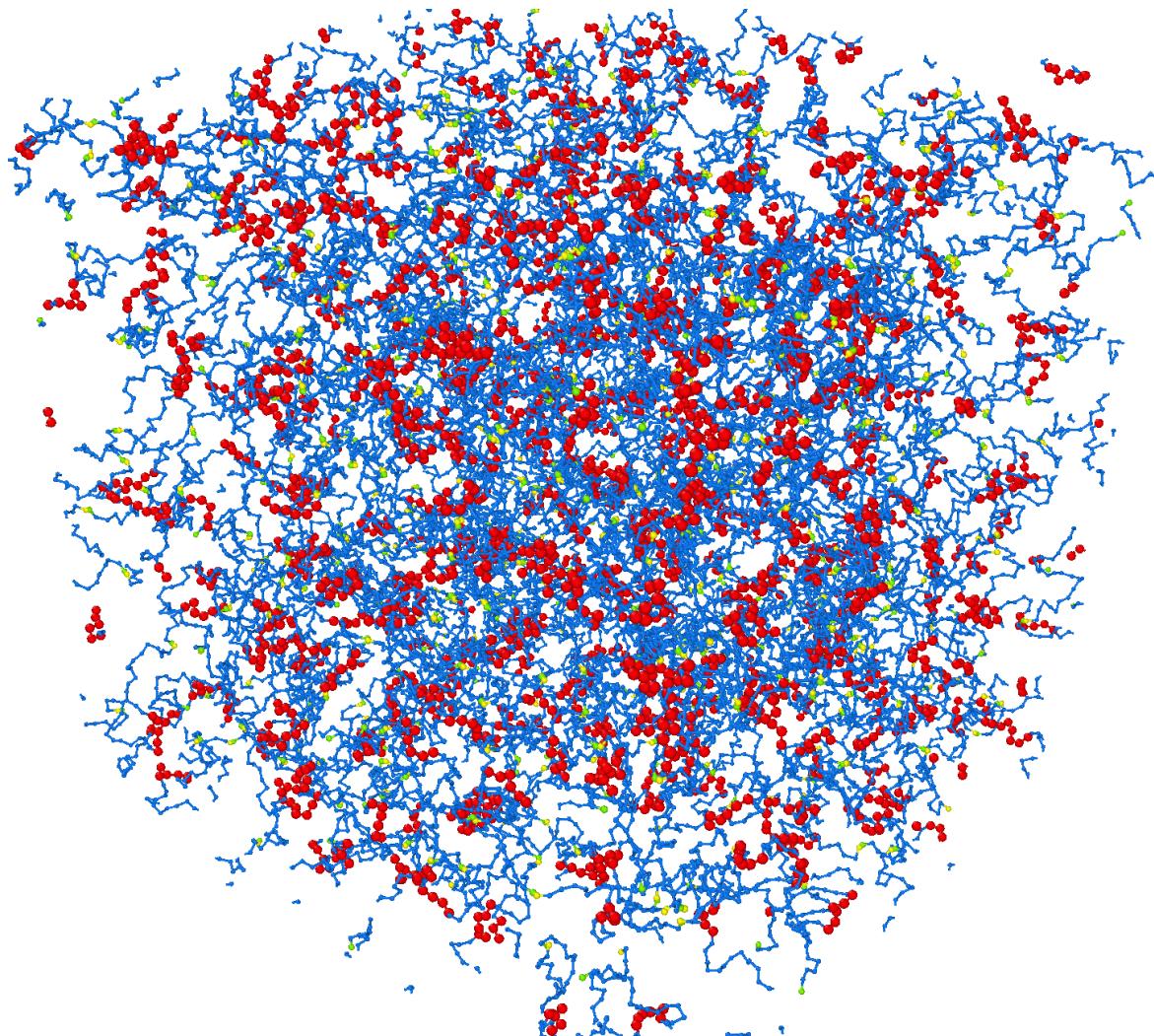


Fig. D.6 Example snapshot (A) [$T < T_m$, $T < \text{CMT}$]: unimers are linked by DNA patches. Long chains are formed in an associated unimer liquid phase. System size: $N=500$; Volume fraction: $\phi = 0.38$. PPO parts in red; PEO parts in blue; the yellow and green beads represent complementary DNA patches.

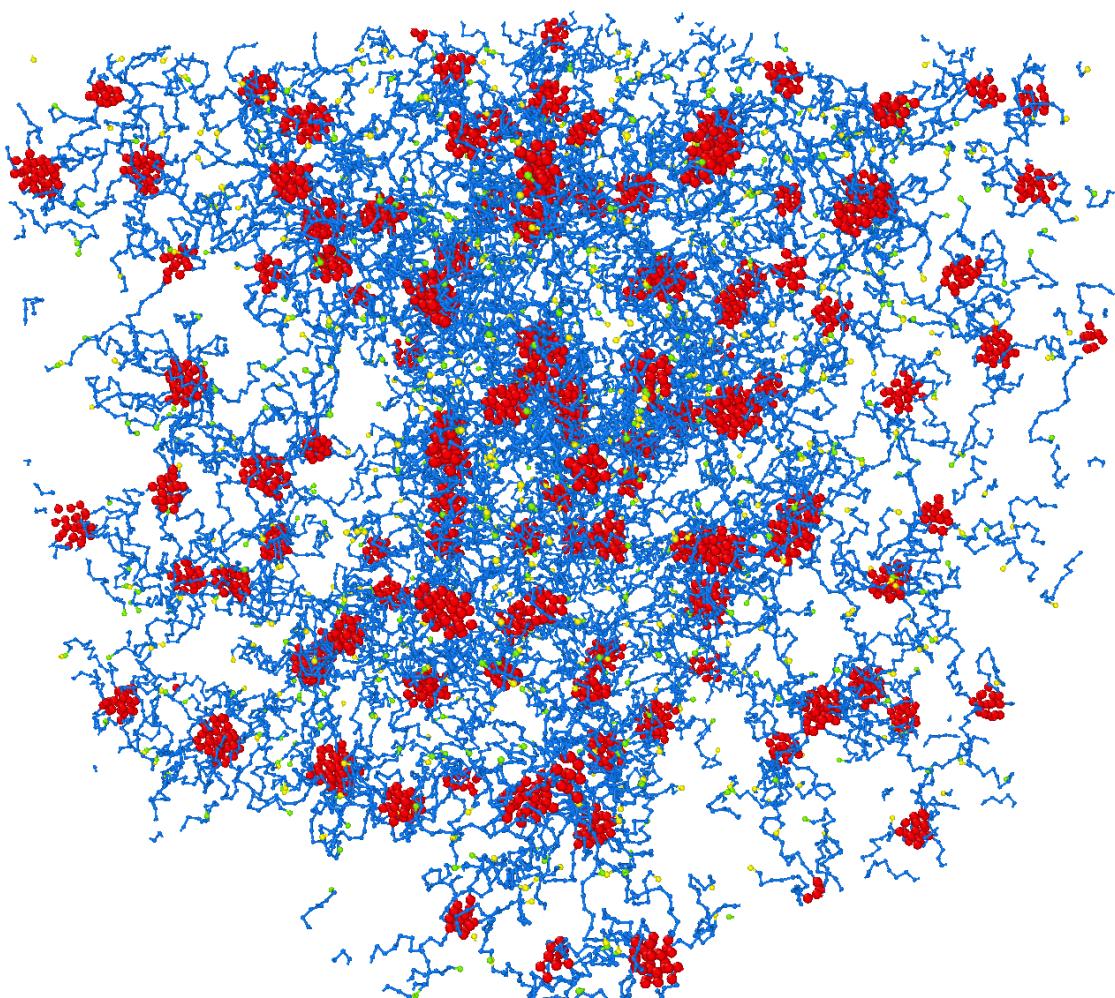


Fig. D.7 Example snapshot (B) [$CMT \approx T < T_m$]: micellisation begins, micelles are forming with DNA connected. The system is close to a micellar gel phase and showing a heterogeneous status ($N=500$, $\phi = 0.38$).

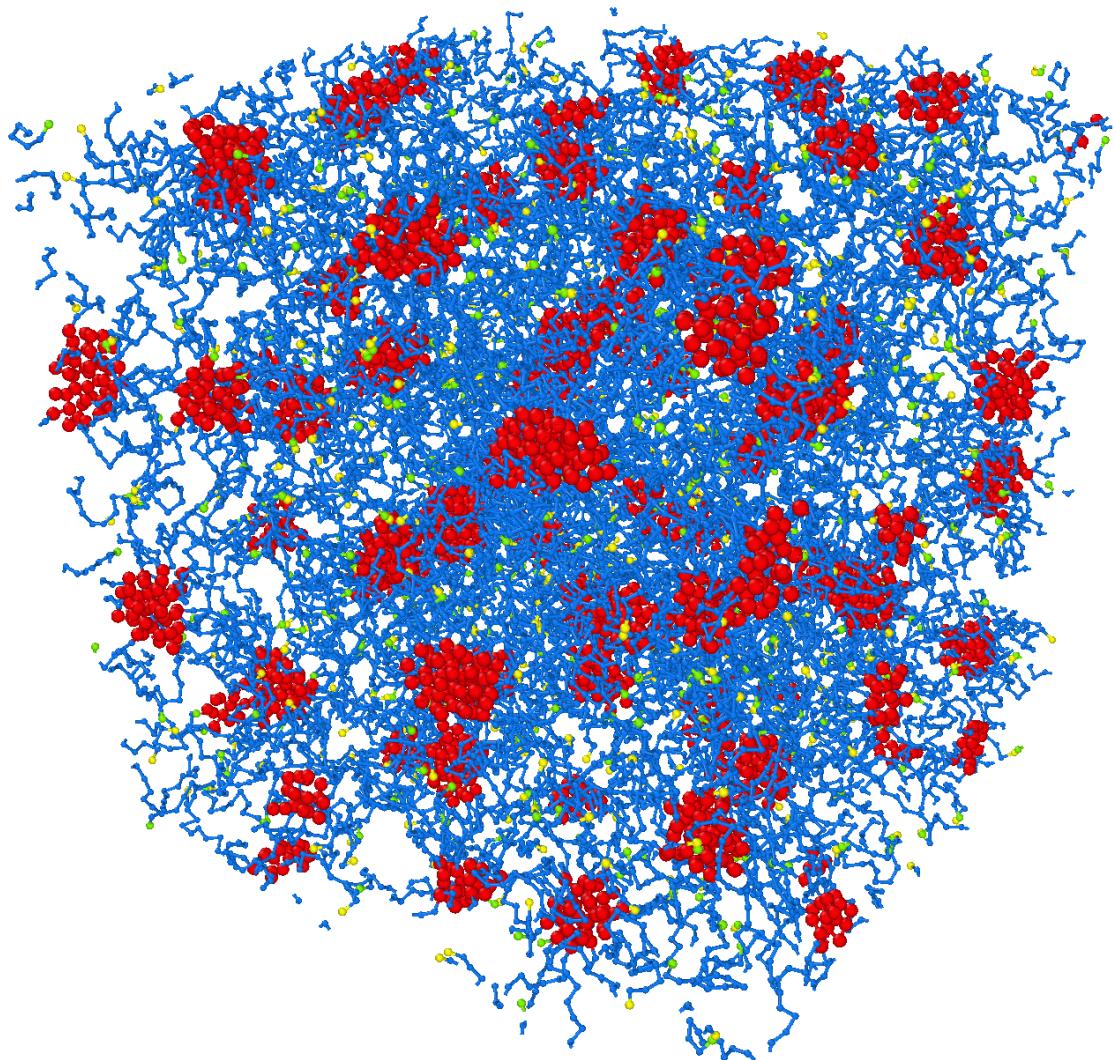


Fig. D.8 Example snapshot (C) [$CMT < T \approx T_m$]: micellisation finished, micelles formed with partially DNA connected. The system is in a transient micellar liquid phase and approaching a homogeneous status ($N=500$, $\phi = 0.42$).

E | Side Projects

E.1 DNA nanotubes with programmable sub-pore

Nanometer-accurate and high-yield production of nanoscale devices and structures is made possible by DNA origami [254, 255]. Since their geometry enables several applications, such as the encapsulation of nanoparticles, the creation of synthetic membrane pores, and the use as structural scaffolds that can specifically spatially arrange nanoparticles in circular, linear, and helical arrays, tubular DNA origami nanostructures are particularly advantageous [256–258].

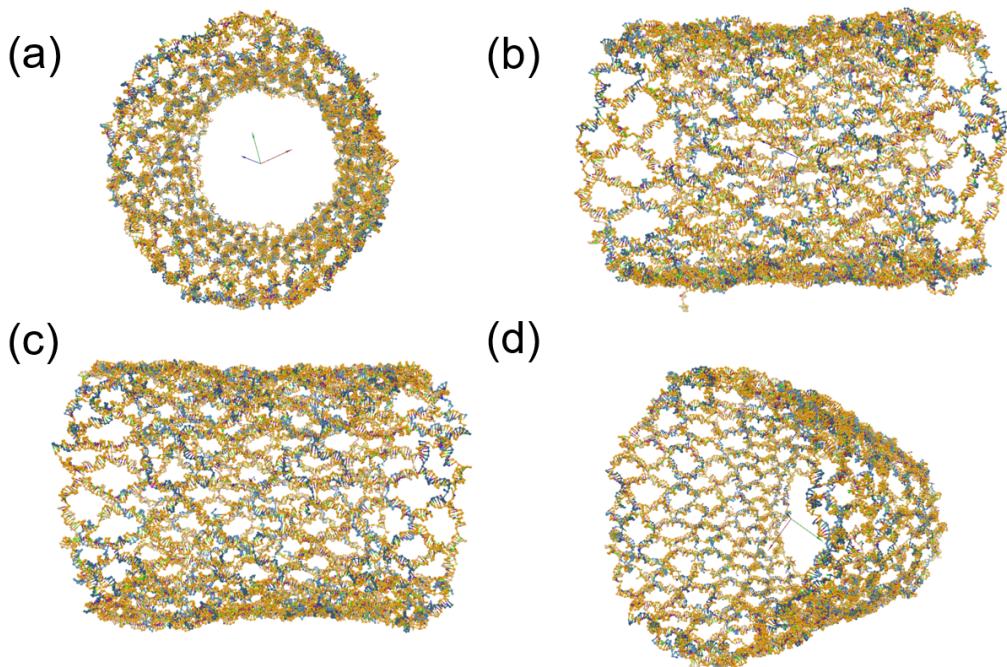


Fig. E.1 Schematic of standard DNA nanotube.

Here, we present a scheme for creating DNA origami nanotubes with tunable crossover density, length and diameter. The oxDNA simulation [83] offers a helpful way to navigate the

intricate geometry of the DNA origami nanotube design. We use this in the creation of brand-new DNA origami nanotubes as well as their production and characterisation. These are structures with pleated walls that allow the same number of duplexes to generate nanotubes of various diameters and to change the diameter of the same structure. We also create nanotubes that can be reshaped into tube pore programmed configurations. We aim to investigate how strain affects the local and overall geometry of DNA origami nanotubes and show how pleated walls may be used as a strategy to rigidify nanotubes and create densely packed parallel duplexes. Ultimately, we expect to open the possibility for numerous applications related to biomedicine and drug delivery utilising the biocompatibility and thermoreversible character of these novel-designed DNA nano-tubes.

At the current stage, we simulated the geometry and tested the structural stability with an extra-designed sub-pore on the tube. Fig. E.1 and E.2 illustrated the standard DNA nanotube compared with nanotube with a programmed sub-pore. The basic structure was created in oxDNA by preparing the sequences in a JSON format (an output of the DNA struct used during generation) collaborated with Dr. Peicheng Xu. The diameter of the tube can be tuned by controlling the number of repeated units. We note that this system can potentially be expanded upon to include more elaborate design configurations by providing extra adjustable parts on the tube. In addition, oxDNA simulations present substantial flexibility in the propeller-like structure of DNA tubes. We, therefore, aim to explore the influence of flexibility in constructing DNA nanotubes with more programmable sub-pores. We envisaged designing structural programmable nano-sized carriers in specified drug delivery and nano-sized bio-material capsules.

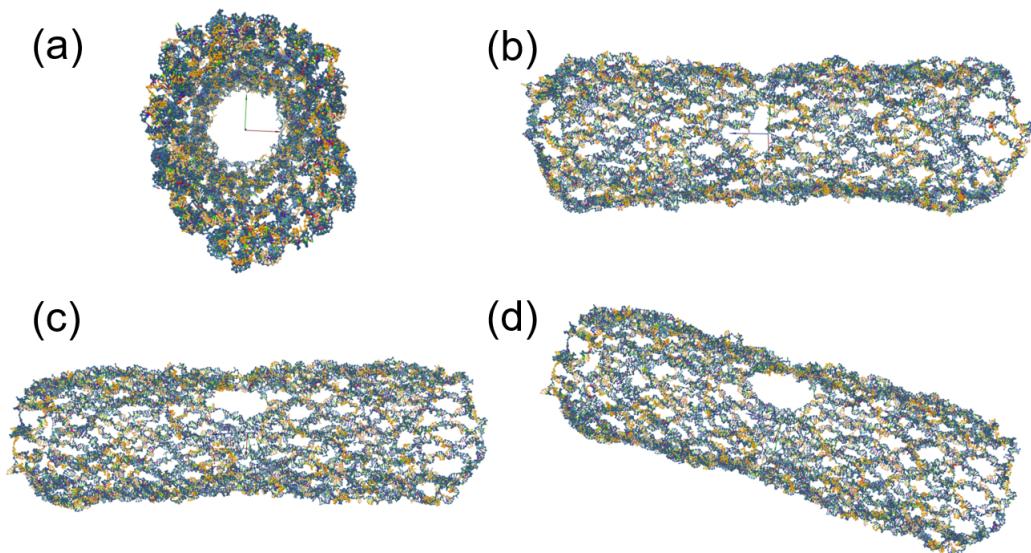


Fig. E.2 Schematic of DNA nanotube with programmed sub-pore.

E.2 Polydisperse diameter distribution for micelle system

β is used as diameter distributions, sampled randomly from a distribution with ($\beta \geq 0$). $\beta = 0$ means the hard spheres in the system all have equal diameters, and the system is monodisperse. The larger value of β represents a system with larger size-polydispersity. The corresponding standard deviations of the hard-sphere diameter are calculated as below:

$$S_d = \sqrt{\frac{\bar{d}^2 - \bar{d}^2}{\bar{d}^2}} = \sqrt{\exp(\beta^2) - 1}, \quad (\text{E.1})$$

$$\beta = \log(1 + s_d^2), \sigma = \sqrt{\bar{d}^3}. \quad (\text{E.2})$$

Here, d is the diameter of the hard-sphere. If we measure the size distribution of our micelle system and apply the data to the polydisperse system, we can update more realistic calculations of radius distribution functions $g(r)$ and structure factors $S(q)$. The accurate $S(q)$ measurements can help explain the experiment findings (especially in semi-density cases), but further discussion requires to apply this method.

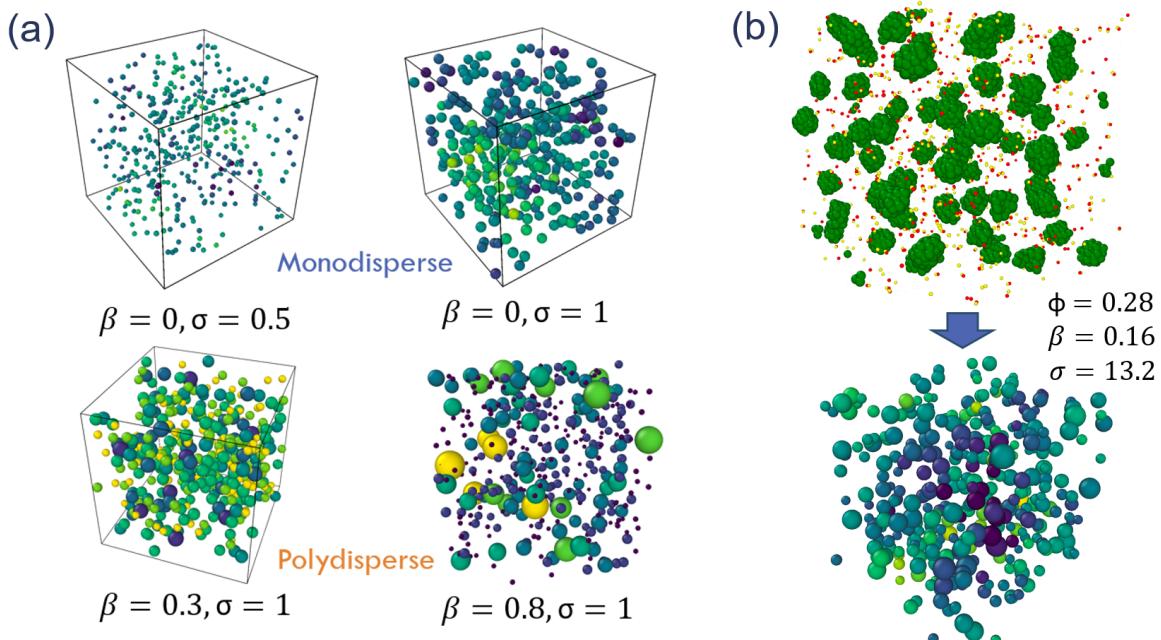


Fig. E.3 (a) Monodisperse and polydisperse systems. β represents the diameter distributions, while σ is the average size of the hard-spheres. (b) Polydisperse representation of a typical micelle system at a volume fraction $\phi = 0.28$; the measured average size of the centre PPO core is $\sigma = 13.2$ with $\beta = 0.16$.

E.3 Pluronic F127 with Y-shape DNA as mediated connector

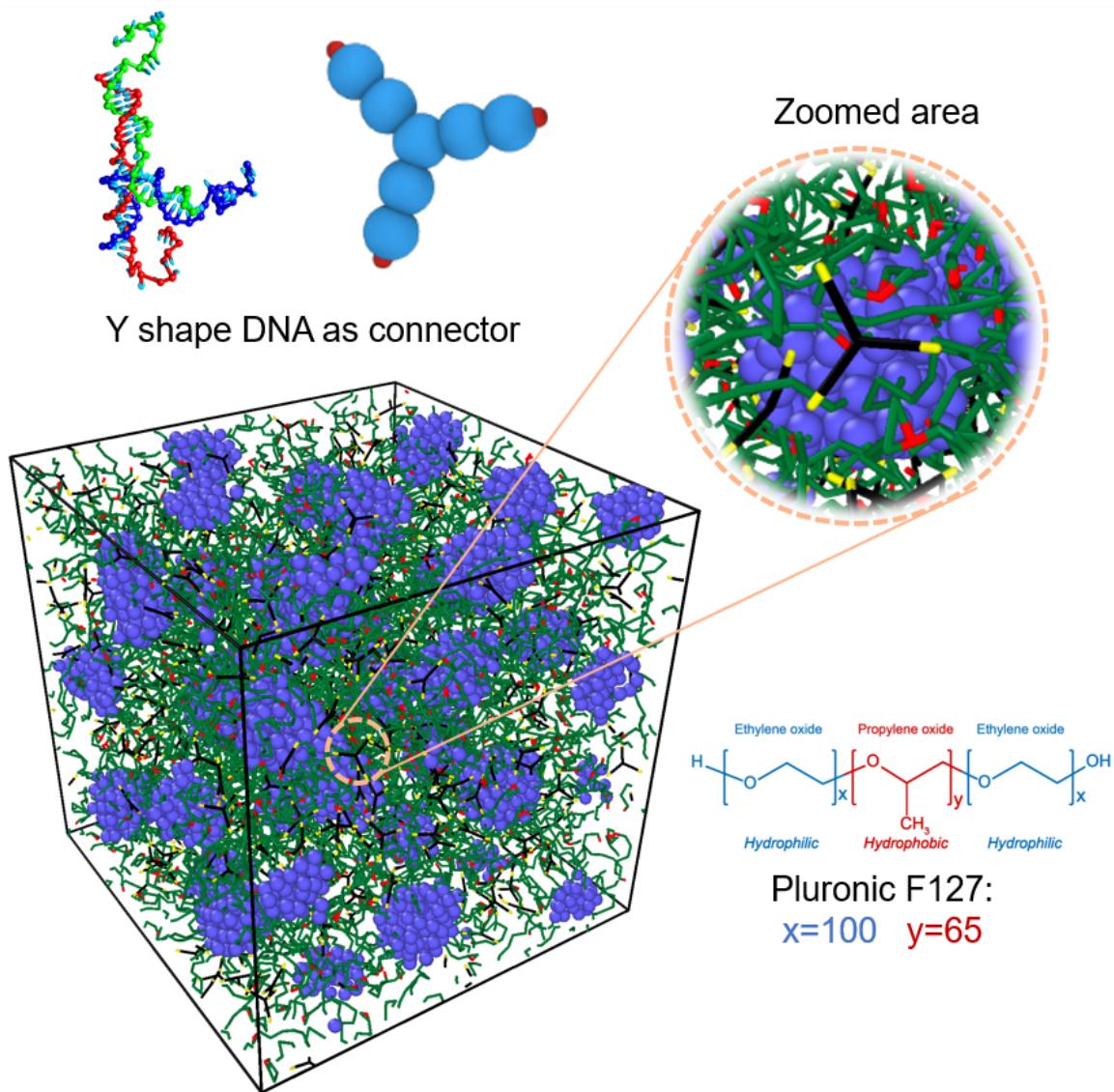


Fig. E.4 Illustration simulated Pluronic F127 with Y-shape DNA as connectors. Blue blobs represent PPO parts. PEO parts are coloured green, and the PEO blobs are not showing to visualise PPO cores. The Y-shape DNAs are presented as a black scaffold with DNA sticky ends in yellow. The compensatory DNA sticky ends on micelles are coloured red.

Inspired by the work in Chapter 6, we further illustrated an updated simulation model for F127 Pluronics with Y-shape DNA as mediated connectors. We applied a simulation model of the F127 Pluronics $[(PEO)_{100} - (PPO)_{65} - (PEO)_{100}]$ with $A_{10}B_7A_{10}$, corresponding to

the model description in Section 6.2.1. The DNA patchy ends are also coded with tunable melting temperature (shown in red and yellow in Fig. E.4). To build a more accurate model, the parameters of temperature-sensitive potential are required to be adjusted to match the F127 system. We have shown some ongoing results for the simulation when micelles are formed together with the coarse-grained Y-shape DNA that we introduced in [46]. This simulation study is currently undertaken with Ren Liu. We aim to finish more reliable observations by tuning up the melting temperature of DNA sticky ends. As the Y-shape DNA has three sticky ends, we envision novel structures that this system can achieve.

