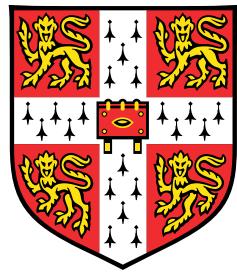


Towards the Description of Mechanical Properties of DNA Hydrogels with Computational Tools



Iria Pantazi

Supervisor: Prof Erika Eiser

Department of Physics
University of Cambridge

MPhil Thesis

St Edmund's College

February 2020

Acknowledgements

In this section I would like to acknowledge help and support that I received during these two years. First of all, I would like to thank my supervisor, Prof Erika Eiser who inspired and guided me to work on the field of DNA-based systems. I would also like to thank the Eiser-group members for all the fruitful discussions that we had on various research topics. Finally, I acknowledge the EPSRC CDT Computational Methods for Materials Science and the Winton Program for the Physics of Sustainability that fund me.

Abstract

DNA can be found in all living organisms and has been described as the molecule of life, since it encodes all genetic information. Since the 1980s, DNA has been used as a structural material for nanoscale devices, which paved the way for the field of DNA nanotechnology. DNA properties such as programmability, biocompatibility and self assembly can be exploited for biomedical applications and specifically in therapeutics. A group of promising materials for controlled drug delivery are polymeric networks made entirely of DNA building blocks, called DNA hydrogels.

In this thesis I am describing a coarse grained numerical model that represents a generic DNA hydrogel consisting of Y-shaped building blocks. Moreover, I introduce a coarse grained description for DNA hydrogels made of Y-shaped and linear building blocks. The building blocks' geometries are defined by suitable harmonic potentials, and the interactions between the blocks are described by modified Lennard-Jones potentials. Using numerical simulations such as Molecular Dynamics (MD) the melting transition of the DNA hydrogels is assessed, and the structural characteristics are derived. Moreover, the network formation of the DNA hydrogel that is derived from MD is described through concepts of graph theory. Finally, mechanical properties of DNA hydrogels, and specifically the shear stress loss modulus and the shear stress storage modulus are estimated through Non Equilibrium Molecular Dynamics (NEMD) and the Kubo theory. The results indicate that at low temperatures a fully percolating network is formed, since all the building blocks are forming connections with one another. On the other hand, at high temperatures the building blocks are separated, and the system behaves as a liquid with higher viscosity.

Table of contents

Acknowledgements	ii
Abstract	iii
1 Introduction	2
1.1 What is DNA?	2
1.2 DNA Thermodynamics	4
1.3 Molecular Self-assembly	5
1.4 DNA Nanotechnology	6
1.5 DNA Hydrogels	8
1.6 Mechanical Properties of DNA Hydrogels	10
1.7 Outline of the Thesis	13
2 Computational Tools	14
2.1 Classical Molecular Dynamics	15
3 Model Setup for the Description of DNA Hydrogels	18
3.1 The Coarse Grained Model	19
3.1.1 Radial Distribution Function of the Initial Configuration	23
3.2 Extension of the System	24
4 Results and Discussion	27
4.1 Model Setup and Melting Curves	27
4.1.1 Number of Pairs	27
4.1.2 Radial Distribution Function for the Steady State	30
4.1.3 System of Y-shaped Molecules with Linear Linkers	31
4.1.4 Characterisation of the Steady State	33
4.2 Cluster Size Distribution	37
4.3 Viscoelasticity	41

5 Conclusions	43
References	45
Appendix A Appendix	50
A.1 DNA conformations	50
A.2 LAMMPS Technical Details	51
A.3 LJ Reduced Units	51
A.4 Melting Curves for Various Concentrations	51

List of Acronyms

A adenine

AFM atomic force microscopy

C cytosine

CG coarse grained

DNA Deoxyribonucleic Acid

dsDNA double-stranded DNA

G guanine

H hydrogen

LAMMPS Large-scale Atomic/Molecular Massive Parallel Simulator

LJ Lennard-Jones

MC Monte Carlo

MD Molecular Dynamics

MSD mean square displacement

N nitrogen

NEMD Non Equilibrium Molecular Dynamics

O oxygen

PES potential energy surface

RDF radial distribution function

RHS right-hand side

RNA Ribonucleic Acid

SI Système international d'unités

ssDNA single-stranded DNA

T thymine

U uracil

Chapter 1

Introduction

Deoxyribonucleic Acid (DNA) can be found in all living organisms and some viruses and has been described as the molecule of life as it encodes genetic information. Since the early 1980s, DNA has been used as a structural material for nanoscale applications. This initiated the field of DNA nanotechnology according to which, DNA is taken out of its biological context, and is used as a building block for scientific applications. The applications vary from nanoparticle assembly to diagnostics and controlled drug delivery. In this thesis I will describe the way that DNA is used for the creation of hydrogels, and their mechanical properties.

In Section 1.1 I am describing the molecule of DNA, and in Sections 1.2 and 1.3 I am presenting the SantaLucia model of DNA thermodynamics, and the DNA self-assembly respectively. In Section 1.4 I am giving an overview of the DNA applications in nanotechnology. Section 1.5 comprises the description of DNA hydrogels, which is the main subject of the thesis. My focus is on the different building blocks that can form DNA hydrogels, and their corresponding mechanical properties which are discussed in Section 1.6. Finally, in Section 1.7 I am presenting an outline of the thesis chapters that follow.

1.1 What is DNA?

In 1953, Watson and Crick [1] described the double-helical structure of DNA, the molecule that stores all genetic information in living organisms. Understanding DNA has brought about major advances in biology, and physical sciences [2]. As well as being a carrier of genetic information, DNA can be handled as a building material. This has led to the design of DNA-based configurations that can be used in medicine [3], and nanotechnology [4]. Before presenting the various applications that DNA can be used in, I shall give a brief description of its geometric and chemical properties.

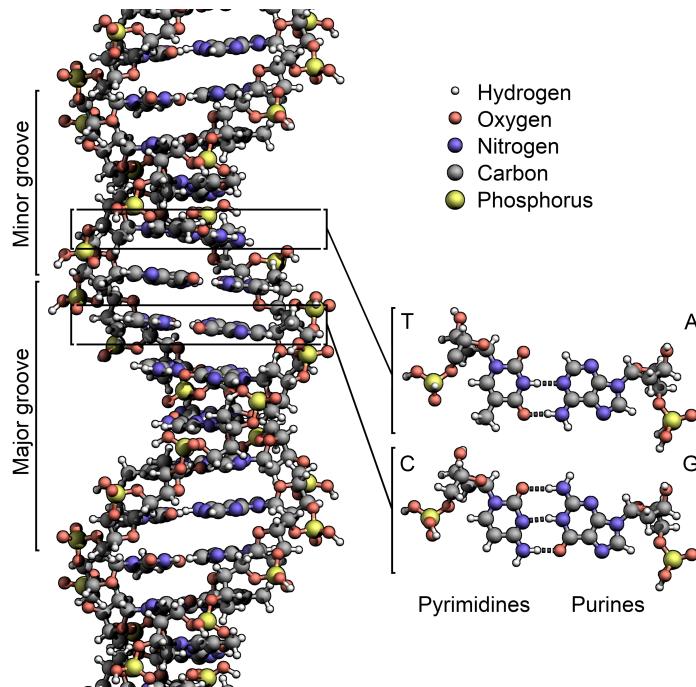


Figure 1.1 Figure presenting a DNA double helix, taken from [6]. Each strand comprises the backbone (with the sugar rings and the phosphate group), and the bases. The sugar-phosphate group lies on the outside of the the helix, whereas the bases are on the inside. On the right are presented the hydrogen bonds that are formed between the H, and either O or N.

DNA as an informational material has the primary function of storing and transmitting genetic information in biological systems. This information is encoded in the form of sequences of bases. The bases are of four types: cytosine (C), guanine (G), adenine (A) and thymine (T).¹ These bases may follow the Watson-Crick² base-pairing rules, according to which A binds only to T with two hydrogen bonds, and C binds to G with three hydrogen bonds as shown in Figure 1.1. DNA is a linear polymer made of monomers called nucleotides. Each nucleotide comprises a sugar-phosphate unit, and one of the four aforementioned bases. The nucleotides are held together through covalent bonds forming single-stranded DNA (ssDNA). Two strands with complementary bases can wind around one another in antiparallel direction forming double-stranded DNA (dsDNA).

The formation and the shape of dsDNA is attributed to covalent bonds, and a number of interactions: ionic interactions, hydrogen bonds, van der Waals forces, and hydrophobic forces [7]. Covalent bonds are holding the sugar-phosphate molecules together, thus forming

¹In nature, also exists Ribonucleic Acid (RNA) which instead of T it contains the base uracil (U) that can bind to A. In living organisms there can be found various types of RNA, but in this thesis I shall not focus on any type of RNA.

²Base pairs that do not follow the Watson-Crick pairing, e.g. Hoogsteen base pairs [5], are not described in this thesis.

the backbone of the DNA molecule. The phosphate group of the backbone is negatively charged in aqueous solutions, and this gives rise to repulsive Coulombic (also called ionic) interactions between two ssDNA strands. In aqueous solution containing salt (usually $NaCl$ or $MgCl_2$), the Coulombic repulsion between the backbones is screened, since the cations (Na^+ or Mg^{2+}) are neutralising the phosphate groups. Hence, screening the Coulomb repulsion allows two complementary strands to approach each other, and binding is achieved through the formation of hydrogen bonds.³

The helical configuration of the DNA molecule is attributed to the hydrophobic forces and the dispersion forces that are exerted on the bases. Water molecules are dipoles that are held together through hydrogen bonds. The bases are non-charged, hence the contact of polar water molecules with the non-charged bases is more energetically costly than the contact with other water molecules. This is also known as the bases being hydrophobic. When two strands bind with each other, they adopt the helical form which is more energetically favourable due to the hydrophobic bases not being in contact with water. The base pairs are planar and stacked due to the π -bonding between adjacent bases, which contributes to the DNA's stability.

The binding of two ssDNA strands (also called DNA hybridisation) depends on the temperature, the salt concentration in the solvent, and the pH of the solvent.⁴ Depending on all the aforementioned conditions, a number of DNA conformations, namely A-DNA, B-DNA, Z-DNA can be found in nature. These conformations are presented in Figure 1.2, and their characteristics are mentioned in Table A.1. All of the aforementioned conformations exist in nature, but in the current thesis I shall focus only on B-DNA which is used in the experiments conducted within the group.

1.2 DNA Thermodynamics

The formation of the DNA double helix is called hybridisation, and the inverse process is called denaturation. Apart from the canonical double helix, DNA can form various intermolecular and intramolecular secondary structures [8]. The DNA folding problem is the prediction of the structure and the folding energy of DNA, given its sequences. SantaLucia and Hicks [9] introduced a nearest neighbour model for accurate DNA structure predictions based on a database of thermodynamic parameters for several motifs.

³Hydrogen bonds are a special type of ionic bonds, and are formed between an H atom that is covalently bonded to a more electronegative atom, and another electronegative atom.

⁴pH is a quantitative measure of a liquid's acidity. It is defined as the logarithm of the concentration of H cations: $pH = -\log_{10}[H^+]$

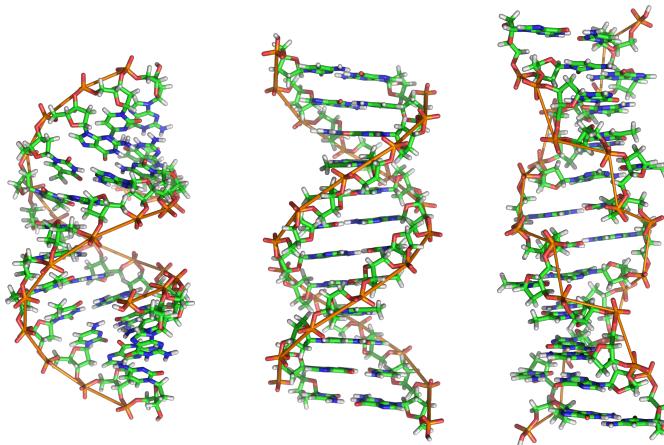


Figure 1.2 Side view of A-DNA, B-DNA and Z-DNA conformations. Different conformations can emerge depending on the hydration level, the base sequences, and various other factors. B-DNA is the most common conformation due to the conditions existing in the cells. The figure is taken from [6].

In SantaLucia's nearest neighbour model [10] it is assumed that the stability of base pairs depends (a) on the identity and orientation of neighbouring base pairs, (b) on an initiation parameter, and (c) on a parameter that accounts for the entropic penalty in case that the DNA strands are symmetric (palindromic). The total change in the Gibbs free energy of dsDNA is expressed as

$$\Delta G^o(\text{total}) = \sum_i n_i \Delta G^o(i) + \Delta G^o(\text{init } G \cdot C) + \Delta G^o(\text{init } A \cdot T) + \Delta G^o(\text{sym}), \quad (1.1)$$

where the sum in the right-hand side (RHS) is the Gibbs free energy change for all the nearest neighbour pairs i that are included in the strand, and n_i is the number of their occurrence. The free energy change of the initiation parameter $\Delta G^o(\text{init } G \cdot C)$ or $\Delta G^o(\text{init } A \cdot T)$ includes sequence-independent effects. The final term in the RHS is non-zero only if the bases in a ssDNA are self complementary [11]. The terms ΔH^o and ΔS^o are related to ΔG^o through the equation:

$$\Delta G^o = \Delta H^o - T \Delta S^o, \quad (1.2)$$

where H^o is enthalpy, S^o is the entropy and T is the temperature.

1.3 Molecular Self-assembly

Self-assembly describes the spontaneous organisation of molecules under thermodynamic equilibrium conditions into stable and structurally well-defined arrangements. The interactions that hold these arrangements together are non-covalent bonds (i.e. weak chemical

bonds that include hydrogen bonding, ionic bonds, and van der Waals interactions) [12]. Non-covalent bonds are considered reversible at normal temperatures, and the overall assembled structures are stable. In the case that the arrangement shows periodicity, the assembly is considered to include a crystalline structure.

In order for DNA self-assembly to occur, there are two criteria that have to be satisfied, complementarity and self-stability [12]. DNA molecules satisfy both criteria, since two complementary ssDNA strands can bind spontaneously, given the right solvent and temperature conditions, and remains stable. This results in more complex structures to be designed, by designing the base sequences in the DNA molecules so that binding happens according to the design. An example is presented in Figure 1.3a, where the design of the strands does not permit the formation of two- or three-branched DNA molecules, but only four-branched Holliday junction instead. This approach is used in the formation of DNA hydrogels studied here.

A characteristic plot that describes DNA self-assembly is the melting curve. The melting curve presents the ratio of the hybridised parts of the DNA strands as a function of temperature. At high temperatures, the melting curve is asymptotically zero, which means that the ratio of the bonds that have been formed between the hydrogel components are almost zero. At low temperatures the ratio of the bonds is asymptotically one, since almost all of the components are connected forming a fully connected system. The melting temperature T_m of the system is defined as the temperature, at which 50% of the system's total bonds have been formed.

1.4 DNA Nanotechnology

Since 1953, when the structure of the DNA double helix was first discovered, there have been major advancements in the field of biology and nanotechnology [1]. DNA has been used itself as a building block or has assisted in the formation of nanoscale devices. Either way, it constitutes a smart component widely used for technological applications [4]. In the following paragraphs, a few examples will be given, and some important inventions will be highlighted.

In 1964, a four-branched double-stranded DNA (dsDNA) configuration known as the Holliday junction (Figure 1.3a) was studied thoroughly [13]. This study paved the way for the investigation of non-linear DNA-based configurations that do not exist in nature. In 1982, Seeman first described the criteria under which DNA-based configurations can emerge from self-assembly [14]. These criteria refer to the DNA strands used to build junctions and three-dimensional network configurations. An example of a branched DNA molecule,

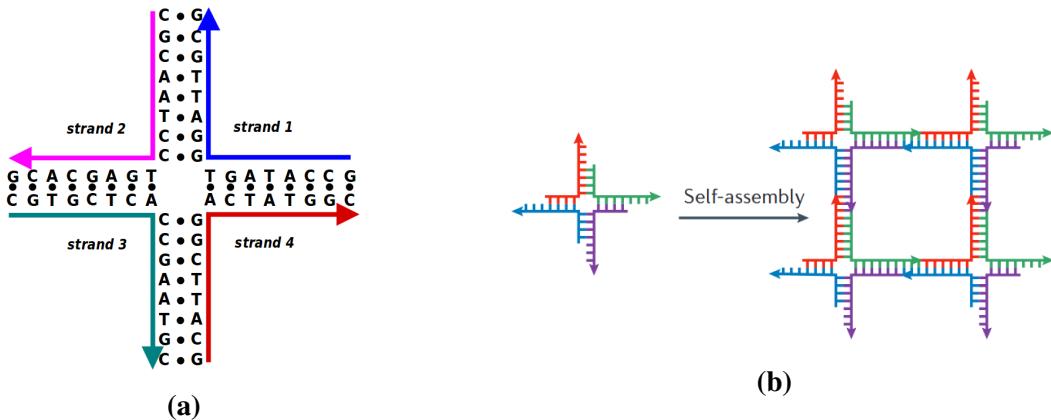


Figure 1.3 Holliday junction formed out of four dsDNA arms joined together (a). The figure is taken from [15]. A simplistic representation of four-branched DNA junctions with free ends that can self-assemble in order to yield 2D configurations (b). This figure is taken from [4].

referred to as 4-way or Holliday junction, is presented in Figure 1.3a. A more simplified picture in which the building blocks self-assemble is presented in Figure 1.3b.

During the next ten years, there was a large number of proposed configurations that could be built from duplex DNA [16, 17]. DNA was also used as a building block for 2D nanocrystals, where single-stranded DNA (ssDNA)s were designed in a specific way so that they would form the desired pattern with chemical or enzymatic ligation [18]. Programming DNA refers to the design of each strand, and obeying the base-pairing rules they will lead to the formation of the desired patterns on the nanoscale-scale.

In 1996, Alivisatos et al. proposed that DNA could be used as a functional material for assisting in the construction of complex assemblies [19]. This is a bottom-up approach for building nanocrystals, and offers precision that is comparable to, or better than the top-down approach that is used, for example, in lithography [20]. Moreover, DNA can be used as a tunable stabilising agent in assisting the formulation of oil-in-water emulsions [21].

In 2006, Rothemund introduced the term DNA origami by designing DNA in such a way that it would form a specific pattern when folded [22]. Rothemund used ssDNA as a scaffold, and short oligonucleotides as staples that could connect in specific parts of the ssDNA. By controlling the temperature, the strand and the staples assembled in the desired pattern. This pattern was determined by the way that the ssDNA and the staples were designed, i.e. the base sequences. Some of the designs are shown in Figure 1.4.

Another research field in which DNA can be used, is computer science. In 1994, DNA was employed in conducting computations at molecular level [23]. Adleman encoded a graph with ssDNA in order to solve the Hamiltonian path problem. He generated random sequences of oligonucleotides and by using thermophoresis he evaluated whether the Hamiltonian

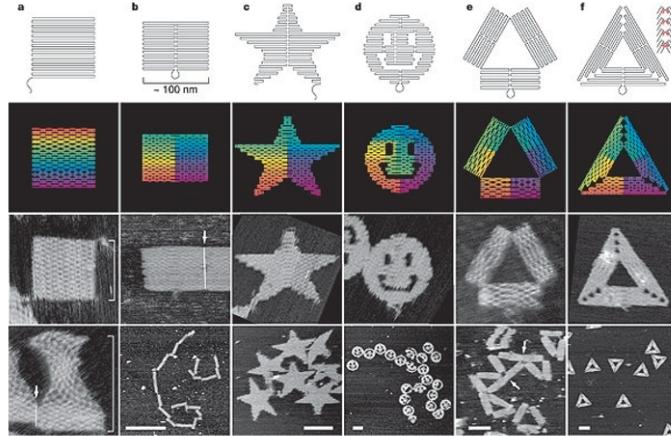


Figure 1.4 DNA origami patterns taken from [22]. The top two rows are idealised designs of the targeted patterns. The bottom two rows are AFM images of the corresponding patterns that were experimentally created.

path had been found. It has also been proposed that, since the primary property of DNA is information storage in the cell, it can be utilised as information storage device on larger scales [24]. Memory devices of high storage capabilities can be designed and can potentially substitute silicon-based storage disks currently used [25].

DNA is poised to become a viable material in the research field of photonics. DNA-enriched thin films show no degradation at temperatures up to 200 °C and demonstrate high optical quality [26]. Photonic devices containing thin films of DNA have shown increased transmission in the visible and near-infrared regions [27], [28]. In the field of optoelectronics, it has been proposed that DNA structures can be used as sensors in optical voltage sensing nanodevices [29].

The aforementioned examples are only a few of the possible applications that DNA can be used in. However, a further description of nanodevices containing DNA is beyond the scope of this thesis. The main subject that will be analysed in the following chapters are DNA hydrogels. In Section 1.5, the DNA hydrogels as well as their applications will be discussed in more detail.

1.5 DNA Hydrogels

DNA hydrogels are polymeric networks, whose constituents are primarily (but not exclusively) Deoxyribonucleic Acid (DNA) building blocks [30, 31] referred to as DNA nanostars. Each nanostar is made of ssDNA strands which are partially complementary so that they form a core of dsDNA with ssDNA tails. The tails, called sticky ends, can bind to complementary

sticky ends of other DNA nanostars. The number of branches of nanostars can vary [32], and defines its valence, i.e. how many connections the DNA nanostar can make. Examples of three-branched and four-branched DNA nanostars and the networks they can form are shown in Figure 1.5. Figure 1.6a shows how two types of three-valent nanostars (also called Y-DNAs) bind in order to form the hydrogel shown in Figure 1.6b. When the constituents of the hydrogel bind to one another, they form networks that may absorb and maintain water [33]. Hydrogels' constituents are held together through molecular entanglement, or non-covalent forces, i.e. ionic bonding, hydrogen bonding, and hydrophobic forces [34, 35].

DNA is considered a versatile material for constructing artificial molecular structures due to its intrinsic characteristics such as programmability, self-organisation, molecular recognition, and molecular-scale structuring properties [36]. All these properties are exploited in DNA hydrogels as introduced by Luo et. al. [30], who used branched DNA building blocks. These materials are good candidates for biomedical applications [37], such as pharmaceutical formulations [38], drug delivery [39], and tissue engineering [40].

DNA hydrogels are characterised by biocompatibility and biodegradability, which are properties inherited from the DNA building blocks. There is a variety of geometries that these building blocks can have, which determine the hydrogel properties. Moreover, the way that the building blocks of the DNA hydrogel organise in space to form 3D DNA hydrogels is thermally reversible, which means that the hydrogel can be controlled by the change of the temperature. These properties motivate my research on different realisations of DNA hydrogels, and the mechanical properties they demonstrate.

Thanks to their biocompatibility and high degree of customisation, DNA hydrogels are excellent candidates for therapeutic applications. Such materials could be used in medicine for controlled drug delivery [30, 42], tissue engineering [43], and 3D cell culture [44]. Furthermore, DNA hydrogels are powerful materials for bioanalysis since it can assist in protein-detection [45], and detection and removal of mercury from water [46].

Studies on DNA hydrogels concern their fabrication [47, 48] and utilisation [49]. Recently Xing et al. [41] employed microrheology to investigate the mechanical properties of these systems. Understanding the fundamental physics of these gels can assist in engineering the systems accordingly, and thus improving their elastic properties. Having an insight into the DNA hydrogels' configuration is critical for the construction of networks of desired topology. Thanks to the powerful programmability of DNA, the hydrogel network can be tailored so that cages with controlled size can form. In Chapter 5, these ideas will be described more extensively.

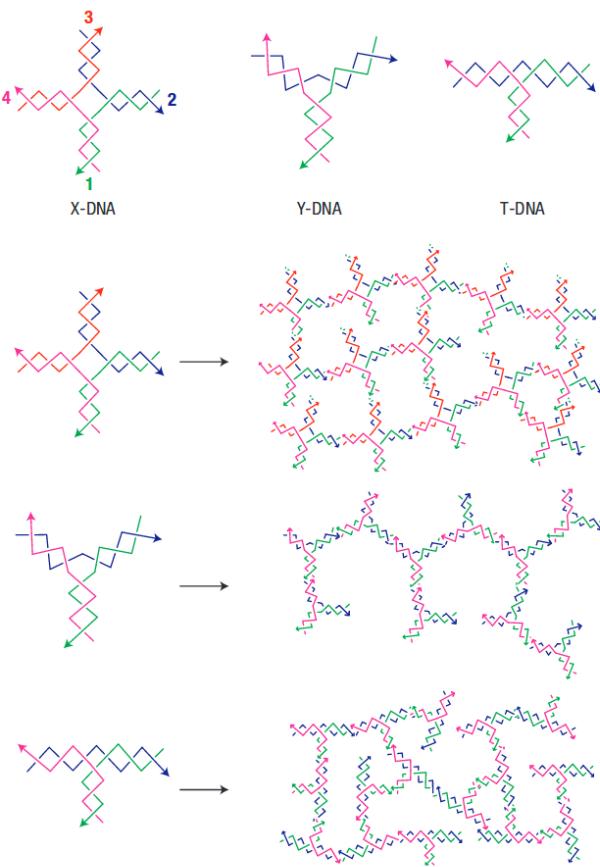


Figure 1.5 Various building blocks (X-DNA, Y-DNA, T-DNA) can be used for the creation of DNA hydrogels as shown on the top row. The corresponding network of DNA hydrogel with the different building blocks is also presented. This idealisation is taken from [30].

1.6 Mechanical Properties of DNA Hydrogels

In the majority of every-day uses of the polymeric materials, the focus is on their mechanical properties. There are materials that are purely elastic, e.g. crystalline solids, or purely viscous, e.g. low molar-mass liquids. On the other hand, there exist materials that combine elastic and viscous properties in both the liquid and solid states. Polymers are the main representatives of this peculiar class of materials, called viscoelastic materials. Since DNA is considered a polymer chain, DNA hydrogels also show viscoelastic behaviour.

Viscoelastic behaviour is a combination of viscosity and elasticity.⁵ Elasticity is the ability of a material to return to its initial size and shape after it has been deformed by the application of external stress. Viscosity measures a fluid's resistant to flow. Viscoelasticity appears when a material's deformation shows both elastic and viscous behaviour, and it is the

⁵It is considered that anelasticity, i.e. a retarded elastic deformation, is also included in viscoelasticity. [50]

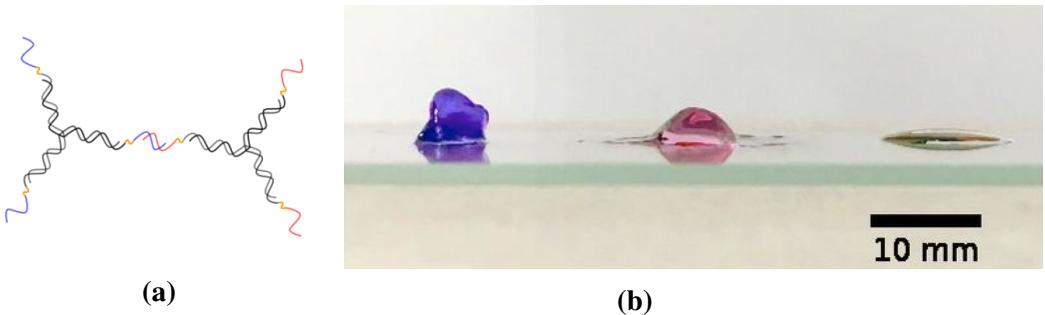


Figure 1.6 Simplified representation of three-valent DNA molecules with the sticky ends connected (a). Figure (b) is a picture of DNA hydrogels created in the lab by Xing et al. Colouring is used for visualisation purpose only. The purple-coloured sample (left) is DNA hydrogel with a high ratio of connected building blocks. The pink-coloured sample (middle) has a lower ratio of connected building blocks, and the yellow-coloured sample (right) has no connected building blocks and is in liquid phase. Both figures are from [41].

subject of rheology. Rheology is the discipline that studies the relationship between stress and deformation in matter. Rheology extends the theories that describe classical extremes of Hookean elastic solids and Newtonian viscous liquids.

In an isotropic linear elastic material, the shear modulus relates the shear stress and the shear strain⁶ as follows: [51]

$$\sigma_{ij} = 2\mu \varepsilon_{ij}, i \neq j \quad (1.3)$$

where σ_{ij} is the shear stress tensor, and ε_{ij} is the shear strain tensor, with μ being the ratio of shear stress and shear strain. In an analogous way, the Newtonian shear viscosity is expressed:

$$\tau_{ij} = \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right), \quad (1.4)$$

where τ_{ij} is the shear stress tensor (drag), and the elements $\frac{\partial v_i}{\partial x_j}$ are the derivatives of the velocity components. Here, μ is the scalar constant of proportionality. These formulae are oversimplified and cannot be used for complex viscoelastic materials like DNA hydrogels.

In practice, the measurement of the shear moduli is possible with dynamic-mechanical experiments. In such experiments, samples are exposed to a periodically varying stress field of the form:

$$\sigma(t)_{xy} = \sigma_{xy}^0 \cdot e^{i\omega t}, \quad (1.5)$$

⁶ The general formula of Hooke's law reads: $\sigma = \lambda \delta_{ij} \varepsilon_{ij} + 2\mu \varepsilon_{ij}$, and it includes the bulk modulus too, but in this thesis it will be omitted.

where ω is the frequency of shear stress. The resulting strain ε_{xy} is then measured, and it is of the form:

$$\varepsilon(t)_{xy} = \varepsilon_{xy}^0 \cdot e^{-i\phi} \cdot e^{i\omega t}, \quad (1.6)$$

where a phase lag $e^{-i\phi}$ appears. The dynamic shear modulus is imaginary, and has the form:

$$G^* = \frac{\sigma(t)_{xy}}{\varepsilon(t)_{xy}} = \frac{\sigma_{xy}^0}{\varepsilon_{xy}^0} \cdot e^{-i\phi}, \quad (1.7)$$

where it can be separated into the real and imaginary parts that correspond to shear elastic (or storage) modulus, and shear viscous (or loss) modulus respectively:

$$G^* = G' + i \cdot G'' \quad (1.8)$$

These quantities are measured with rheology or better with microrheology experiments that do not appear to have the drawbacks of traditional macrorheology [52]. Microrheology requires only a small sample of the substance in comparison to macrorheology. This is convenient when measurements are conducted on expensive biological materials like DNA. Microrheology may sample frequencies up to the order of thousands of Hertz and hence be used to study short-time dynamics of the materials, as has been done by the Eiser group [53]. In particular, the experiments that have been conducted within the group employ passive microrheology (where thermal energy is used in order to move the probe particles in the sample) and active microrheology (where externally applied forces are applied to the probe particles) [53].

In microrheology experiments, the mean square displacement (MSD) of the tracer particle is measured in the Laplace space, and the complex viscous modulus is calculated from the generalised Stokes-Einstein relation [54]:

$$G^*(s) = \frac{k_B T}{\pi \alpha s \langle \Delta r^2(s) \rangle}, \quad (1.9)$$

where α is the radius of the probe particle, and $\langle \Delta r^2(s) \rangle$ is the Laplace transform of the MSD of the probe particle at a frequency s . The way that the storage and loss moduli can be calculated from Non Equilibrium Molecular Dynamics (NEMD) simulations where a periodic shearing stress is applied to the simulation box. Another way of measuring these moduli without the use of NEMD, is with Kubo's response theory [55]. This approach was introduced by Williams and Evans [56, 57], where they showed that thermal transport properties can be converted into much simpler mechanical processes that can be analysed by

Kubo's response theory. [55]. This theory results to the following formulae for G' and G'' :

$$G' = \frac{V}{k_B T} \omega \int_0^\infty \sin \omega t C(t) dt, \quad (1.10)$$

$$G'' = \frac{V}{k_B T} \omega \int_0^\infty \cos \omega t C(t) dt, \quad (1.11)$$

where the stress auto-correlation function $C(t)$ is

$$C(t) = \langle \sigma(t) \sigma(t_0) \rangle. \quad (1.12)$$

1.7 Outline of the Thesis

The thesis is organised as follows: Chapter 2 introduces the computational tools that were used for the description of DNA hydrogels. In Section 2.1 I am describing Molecular Dynamics (MD) which was used for the description of the formation of DNA hydrogels, and the derivation of their mechanical properties. Chapter 3 introduces the coarse grained (CG) model that was employed for the simulation of the DNA hydrogels with MD. Chapter 4 contains all the results that were acquired from the MD simulations, as well as the analysis and interpretation of the plots. Finally, in Chapter 5 I am presenting a conclusion of the work I have done so far along with extensions that can be done, and in the Appendix A I have included additional plots.

Chapter 2

Computational Tools

Molecular modelling is a collection of methods used in order to mimic the behaviour of molecules and molecular systems, and derive their properties. In small systems, energy minimisation can be used for the discovery of the minima in the potential energy surface (PES).¹ Then, with the use of statistical mechanics, the partition function of these minimum energy configurations can be calculated, and consequently the thermodynamic properties of the system can be derived. In larger systems like liquids and solids, the number of atoms is of the order of 10^{23} , making the number of minima in the PES enormous. For this reason, thermodynamic and structural properties of large systems are calculated with the use of computer simulations, the most common of which are Monte Carlo (MC) and MD.

Computer simulations provide a numerical solution to the N-body problem (N being of the order of 10^{23}) by using smaller replicas of the actual systems with a manageable number of atoms or molecules [58, 59]. This renders possible the calculation of the system's properties within feasible times. Simulations can be chosen to be classical, or quantum mechanical² depending on the computer resources, and the properties that one wishes to derive. Since the size of the particles that I simulated are far larger than the electrons, and the properties do not depend on the electronic degrees of freedom, my focus is only on classical computer simulations. I will employ solely MD simulations, due to the advantage of yielding equilibrium averages and dynamical information simultaneously [60].

¹ The PES of a system is the system's potential energy plotted with respect to different parameters, like the atomic positions of the system.

² Quantum mechanical simulations are also called *ab initio*, and they take under consideration the electronic degrees of freedom. The properties calculated are dependent upon the electronic density (like electrical conductivity), but *ab initio methods* can also be employed for the calculation of thermodynamic and structural properties of materials.

2.1 Classical Molecular Dynamics

Thermodynamic properties are generally dependent on each particle's position and momentum. For N particles, an observable A (which is a macroscopic quantity) is expressed with respect to a microscopic function that depends on the system's coordinates and moments ($A(\mathbf{p}^N, \mathbf{r}^N)$) through the equation:

$$\{A\} = \int \int A(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) d\mathbf{r}^N d\mathbf{p}^N, \quad (2.1)$$

where ρ is the probability density of the ensemble being in the phase $(\mathbf{p}^N, \mathbf{r}^N)$. If the ergodic hypothesis is satisfied, then the ensemble average (Equation. 2.1) is equivalent to the temporal average:

$$\langle A \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) dt. \quad (2.2)$$

A system that is ergodic means that all possible micro-states that correspond to a particular macro-state are accessible to the system, and given an infinite amount of time the system will visit all of them. MC calculations are inherently ergodic, whereas it is challenging to prove ergodicity in MD simulations [60].

In experiments, for the measurement of a material's property, a sample is prepared in the laboratory, and the property is measured within a time interval. This process is repeated multiple times for samples of the same kind, in order to increase the accuracy of the measured value. In MD simulations, a model system is selected, and the simulation is initiated. In each time step of the simulation, successive configurations are generated from the previous ones, by integrating Newton's equations, and calculating the forces exerted on the particles. After the system reaches equilibrium, the measurement is applied. The same process is repeated multiple times with different initial configurations of the same system, for increased accuracy of the measured value.

There is a number of algorithms that can be used for the numerical integration of Newton's equations. An appropriate group of algorithms that is employed for this task are velocity Verlet algorithms [61]. The equations of motion that are solved under this scheme are:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 \quad (2.3a)$$

$$v(t + \Delta t) = v(t) + \frac{f(t + \Delta t) + f(t)}{2m}\Delta t \quad (2.3b)$$

where Δt is the time step, f is the force exerted on the particle, and m is its mass. Velocity Verlet algorithms are time reversible, and hence they are good for energy conservation. In

principle energy conservation cannot be achieved with infinite accuracy, since integration depends upon rounding errors due to the floating-point arithmetic. This results in an energy drift, which can be large in long-term simulations. This makes the calculation of true particles trajectories impossible, but in fact, it is the statistical predictions that are of interest. Velocity Verlet algorithms are also symplectic, which means that they generate trajectories quite close to the true trajectories.

The forces are calculated from the derivative of the inter-atomic potential that describes the pair interactions, or from force fields. The inter-atomic interactions can be described by bonded interactions (i.e. bond, angle, and dihedral interactions), and non-bonded interactions (i.e. Coulombic and van der Waals interactions). A force field is a collection of equations with associated parameters that are designed to reproduce the molecular geometry, and selected properties of tested structures. Such examples are AMBER [62] which is a family of force fields widely used for biomolecules, and CHARMM [63] which is a set of force fields used in MD that primarily targets biological systems (e.g. peptides, proteins, prosthetic groups, small molecule ligands, nucleic acids to name a few). In this thesis though, no force fields were used, but more simplistic potentials were employed for the simulations of DNA hydrogels. These potentials are presented in Section 3.2.

MD samples from the microcanonical (also called (N, V, E)) ensemble. This ensemble is used to describe systems that are isolated from their surroundings, and the quantities N (number of particles), V (system's volume), and E (total energy of the system) are conserved. The microcanonical partition function is

$$\Omega(N, V, E) = \frac{1}{h^{3N} N!} \int \int \delta(\mathcal{H}(\mathbf{r}, \mathbf{p}) - E) d\mathbf{r} d\mathbf{p}, \quad (2.4)$$

where integration is over the $(6N - 1)$ -dimensional phase space and normalisation is due to the phase space volume, and $N!$ is the number of microstates that are identical if the particles are indistinguishable.

Since many experiments are conducted at constant temperature, MD simulations can be modified so that the sampling is from the canonical (or (N, V, T)) ensemble, instead of the microcanonical ensemble which is the default in MD. A system in contact with an infinite heat bath has a constant temperature but its energy will fluctuate, and the total Hamiltonian follows a Boltzmann distribution as a consequence that the system can exchange energy with the heat bath. This is reflected to the phase space distribution function, which is analogous to $e^{-\beta \mathcal{H}}$, where $\beta = \frac{1}{k_B T}$. The partition function of the canonical ensemble is proven to be:

$$Q(N, V, T) = \frac{1}{h^{3N} N!} \int e^{-\beta \mathcal{H}} dx, \quad (2.5)$$

which includes the temperature dependence in the term β . The connection between the canonical and microcanonical partition function is

$$Q(N, V, T) = \frac{1}{E_0} \int_0^\infty e^{-\beta \mathcal{H}\Omega(N, V, E)} dE. \quad (2.6)$$

In order to keep the temperature constant, the system has to be in contact with a thermal bath. This results into the system to occupy an energy state given by the Boltzmann distribution. Consequently, the system is described by the canonical ensemble with a constant average temperature. With the use of Andersen thermostat, temperature equilibration happens through stochastic collisions of the system's particles with the heat bath [64]. There is a number of thermostats that can be used in MD, such as Nose-Hoover thermostat [65, 66], Berendsen thermostat [67], and Langevin thermostat [68]. Amongst those, I will focus on the latter, since the Langevin thermostat accounts for the implicit solvent of the system under consideration.

The Langevin equation that describes the Brownian motion of particles³ of particles in a fluid is

$$m\ddot{x} = -\nabla U(x) - \gamma\dot{x} + \sqrt{2\gamma k_B T} R(t) \quad (2.7)$$

in which the first term in right-hand side (RHS) is the force from the interatomic potential, the second term is the viscous term due to the existence of the solvent around the particles (Stoke's law [69]). The third term is a random force that follows a uniform distribution (the distribution could be Gaussian, but it is chosen uniform for speed) force due to solvent atoms at T randomly bumping into particles. Thus, by coupling the interatomic potential with the two additional terms, this specific thermostat introduces the implicit solvent that surrounds the particles.

³Brownian motion is the random motion of particles in a fluid that is caused by the collisions of the particles with the solvent molecules.

Chapter 3

Model Setup for the Description of DNA Hydrogels

The hydrogel constituents are two types (A- and B-type) of Y-shaped DNA molecules. These Y-shaped molecules are the hydrogel building blocks, and are designed in such a way, so that A-type can only bind to B-type. Through hybridisation three strands of ssDNA give rise to a Y-shaped molecule, which consists of a rigid body (fully hybridised DNA), a flexible linker (which is a sequence of 4 thymine (T) bases), and sticky ends as shown in Figure 3.1. The flexible linker is designed so that it does not form any bonds, and the sticky ends permit binding only between A-type with B-type. The base sequences of both types of Y-shaped DNA are presented in Table 3.1.

For the simulation of the building blocks, a coarse grained model will be employed that simplifies the Y-shaped structure, and the interactions between the building blocks. The base sequences are not described explicitly in the CG model, but instead, pseudo-atoms that represent groups of atoms are used. Employing a CG model reduces not only the number of atoms, but also the number of interactions that would be employed in an all-atom simulation.

Table 3.1

type	arm #	Sticky end	Free end	Segment I	Segment II
A	1	5'- TGT CAC TCA CAG	TTTT	TGG ATC CGC ATG ATC	CAT TCG CCG TAA GTA-3'
A	2	5'- TGT CAC TCA CAG	TTTT	TAC TTA CGG CGA ATG	ACA CCG AAT CAG CCT -3'
A	3	5'- TGT CAC TCA CAG	TTTT	AGG CTG ATT CGG TGT	GAT CAT GCG GAT CCA-3'
B	1	5'- CTG TGA GTG ACA	TTTT	TGG ATC CGC ATG ATC	CAT TCG CCG TAA GTA-3'
B	2	5'- CTG TGA GTG ACA	TTTT	TAC TTA CGG CGA ATG	ACA CCG AAT CAG CCT -3'
B	3	5'- CTG TGA GTG ACA	TTTT	AGG CTG ATT CGG TGT	GAT CAT GCG GAT CCA-3'

Table presenting the base sequences of the three arms for A-type and B-type Y-shaped DNA for the hydrogel.

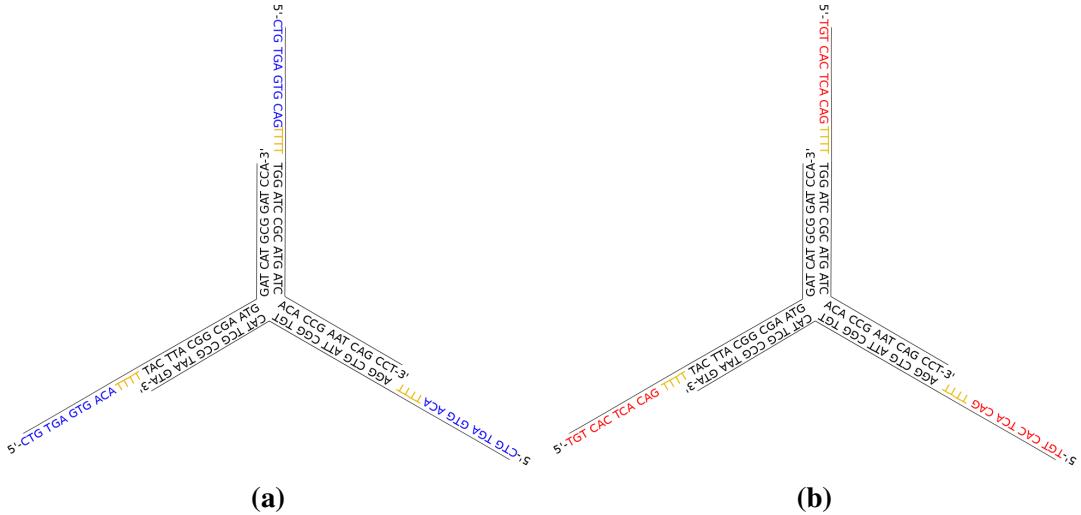


Figure 3.1 Y'-type (a), and Y-type (b) building blocks of the DNA hydrogel.

This results in a computationally tractable MD simulation that can be used for the study of larger systems and for longer time scales than all-atom MD.

3.1 The Coarse Grained Model

The CG model for the simulation of DNA hydrogels with Y-shaped constituents was introduced by Zing *et. al.* [70]. Instead of Système international d'unités (SI) units, reduced Lennard-Jones (LJ) units are employed, and the calculated quantities are expressed with respect to ϵ_{LJ} and σ_{LJ} which have the units of energy and distance respectively. A more detailed description of the LJ reduced units is included in the Appendix A.3. According to the model, seven beads are used in place of the dsDNA, and three patches for ssDNA that will hybridise. Each bead's radius is $R = 0.56\sigma_{LJ}$ and the distances between them are $d = 0.96\sigma_{LJ}$, and were chosen so that they reflect the geometry of the Y-shaped DNA molecule, i.e. $2R$ for 2 \AA that is the width of the double helix, and $2R + d$ for 15 base pairs of length 51 \AA . The geometry of such a molecule is presented in Figure 3.2.

In the simulation, the molecules are designed in such a fashion that they do not change shape. This is reflected by the way that all the beads are bonded. In the model, there are used harmonic potentials restricting the bonds and angles between the beads in the molecules. These are also presented in Figure 3.2 as springs. The potentials are of the form:

$$V_{bond} = K_{bond}(r - r_{ij})^2 \quad (3.1a)$$

$$V_{angle} = K_{angle}(\theta - \theta_{ijk})^2 \quad (3.1b)$$

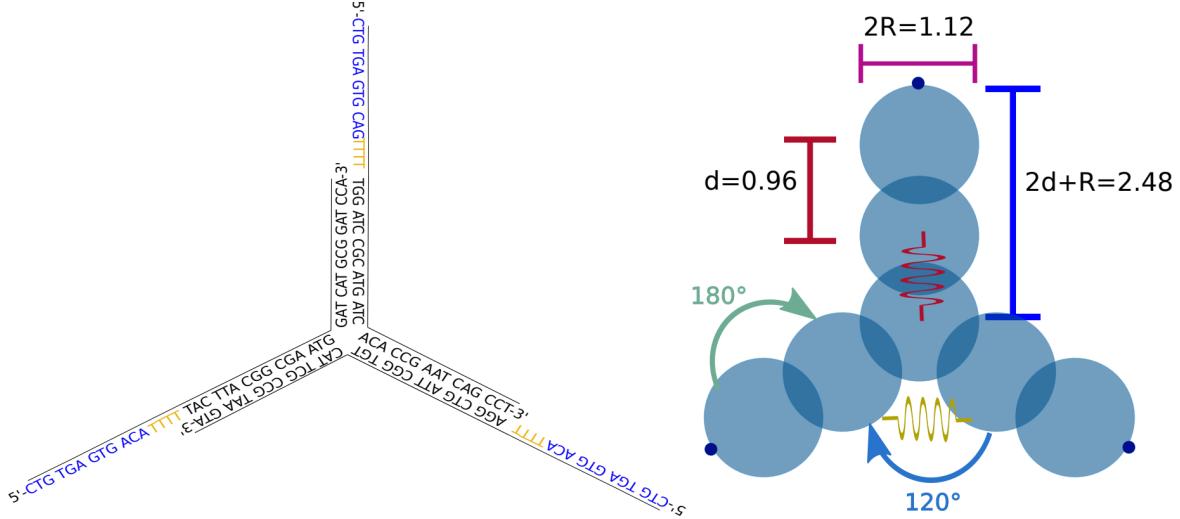


Figure 3.2 Representation of a Y-DNA building block with the base sequences (left), and CG description of the same molecule (right). In the CG description of a Y-DNA molecule, the beads' dimensions and the springs that represent the harmonic potentials are also presented.

with the spring constants $K_{bond} = 300\epsilon_{LJ}/\sigma_{LJ}^2$, and $K_{angle} = 100\epsilon_{LJ}$, respectively, the equilibrium distances $r_{ij} = 0.96\sigma_{LJ}$ (between two nearest beads) or $r_{ij} = 0.56\sigma_{LJ}$ (between patch and bead) and the equilibrium angles $\theta_{ijk} = 120^\circ$ or $\theta_{ijk} = 180^\circ$.

The Lennard-Jones potential between two particles has the form:

$$V_{LJ}(r) = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right), \quad (3.2)$$

where r is the distance between the beads, ϵ is the strength of the attraction, and σ is the parameter where $V_{LJ}(\sigma) = 0$. In this thesis, the interactions between the beads have the form of a Weeks-Chandler-Anderson potential [71]:

$$V_{WCA}(r) = V_{LJ}(r) - V_{LJ}(r_{cut}) = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right) - V_{LJ}(r_{cut}), \quad r \leq r_{cut} \quad (3.3a)$$

$$V_{WCA}(r) = 0, \quad r > r_{cut} \quad (3.3b)$$

where r_{cut} is a specified cutoff distance, so that the potential function is continuous. The interactions specified between the beads are of three kinds: the excluded volume, the patchy attraction and the patchy repulsion. The excluded volume prevents the beads from overlapping, since they are considered to be hard spheres. The patchy attraction and repulsion are specified in a way so that there exists binding specificity, i.e. A-type can only bind to B-type Y-shaped molecule. These potentials are presented in the Figure 3.3, and the interactions are also presented in Figure 3.4.

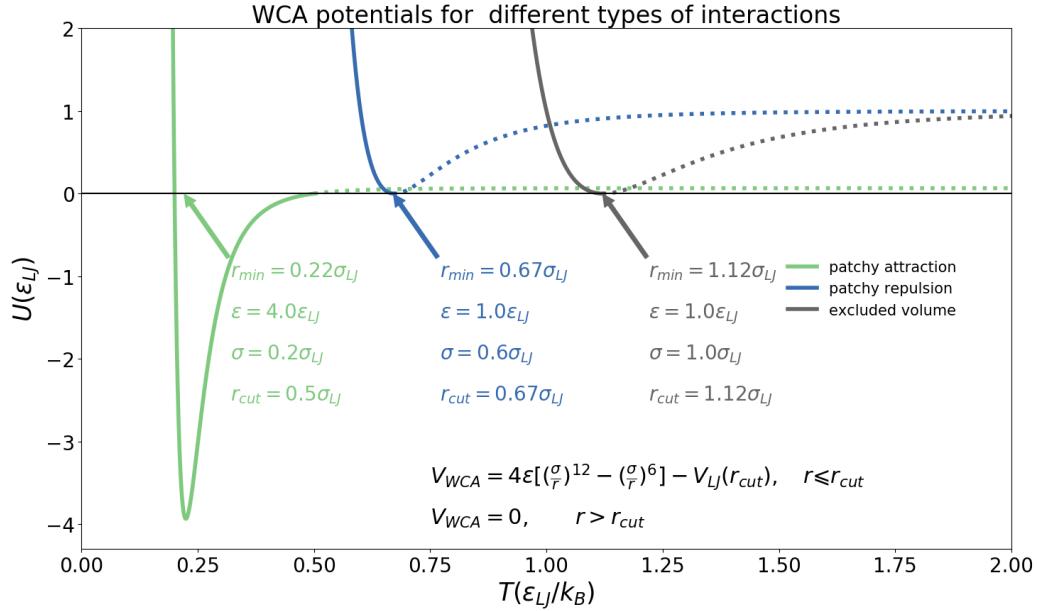


Figure 3.3 Plot of the pair-potential interactions between beads of different molecules. The excluded volume is the repulsive interaction between the core beads, i.e. the beads that represent the dsDNA, and not the free end, neither the sticky end. The patchy repulsion is the repulsion between beads of the same type (A-A and B-B), and the patchy attraction is between different types of molecules (A-B).

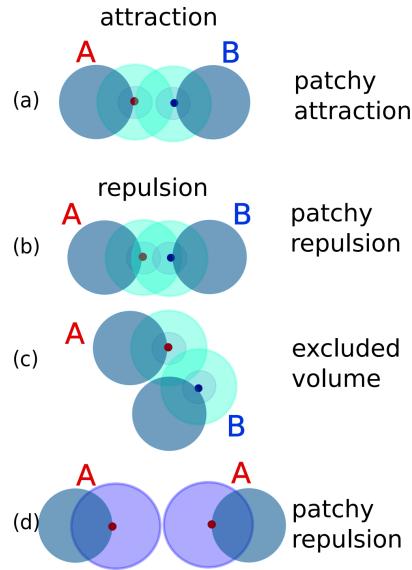


Figure 3.4 The interactions between patches are attractive only between different types of molecules (a), and repulsive when the patches are closer than $2^{1/6}\sigma$ (b). The excluded volume between the beads prevents them from overlapping (c), and the patchy repulsion is present when two patches of the same type are closer than r_{cut} (d).

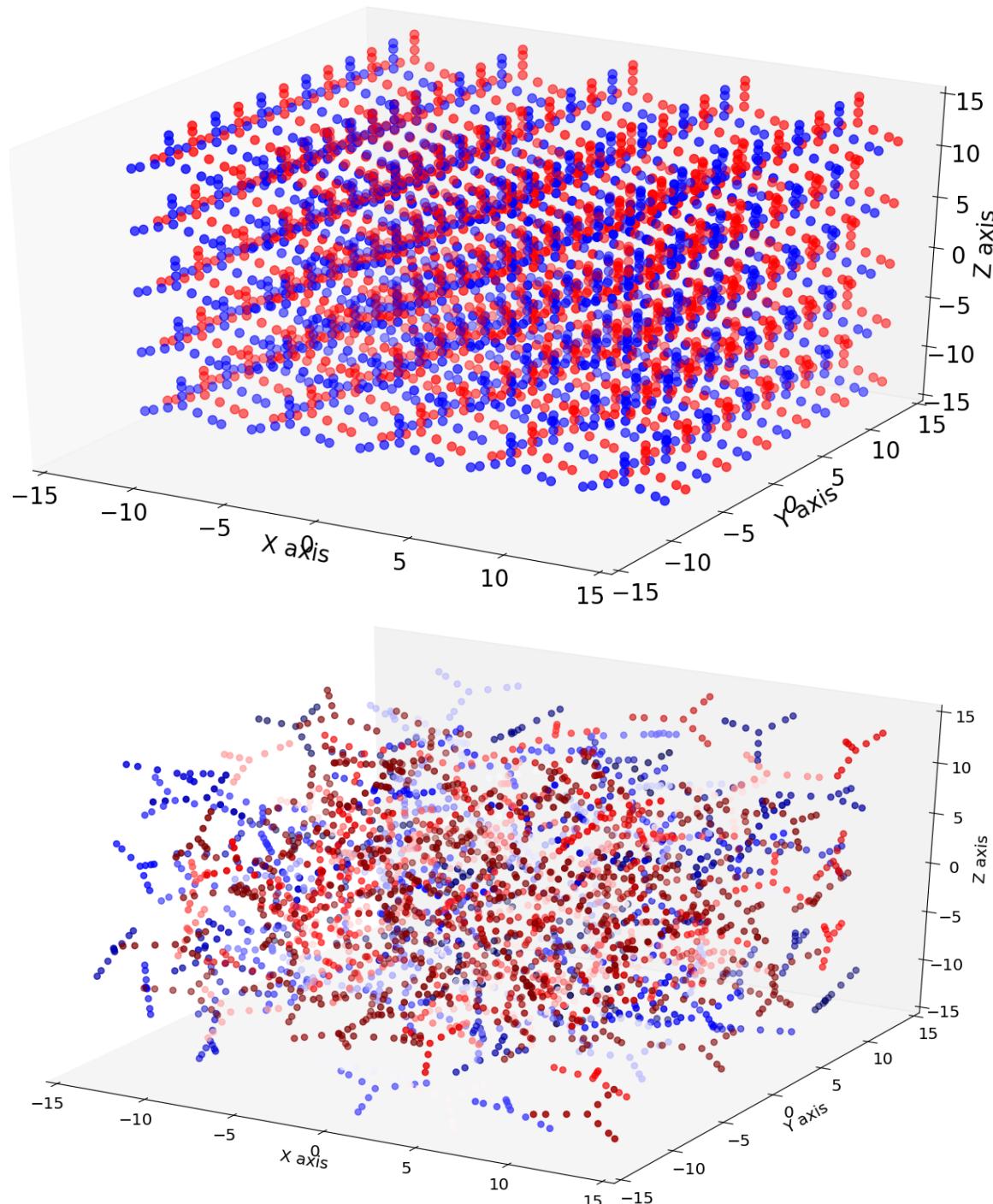


Figure 3.5 Both figures are presenting initial configurations of 300 Y-shaped molecules in a simulation box of volume $27000\sigma_{LJ}^3$. The distinction between them is the arrangement of the molecules, where in one case the molecules are ordered (top), and in the other they are randomly placed in the box (bottom). The colour distinction is for the two types of molecules, Y and Y'.

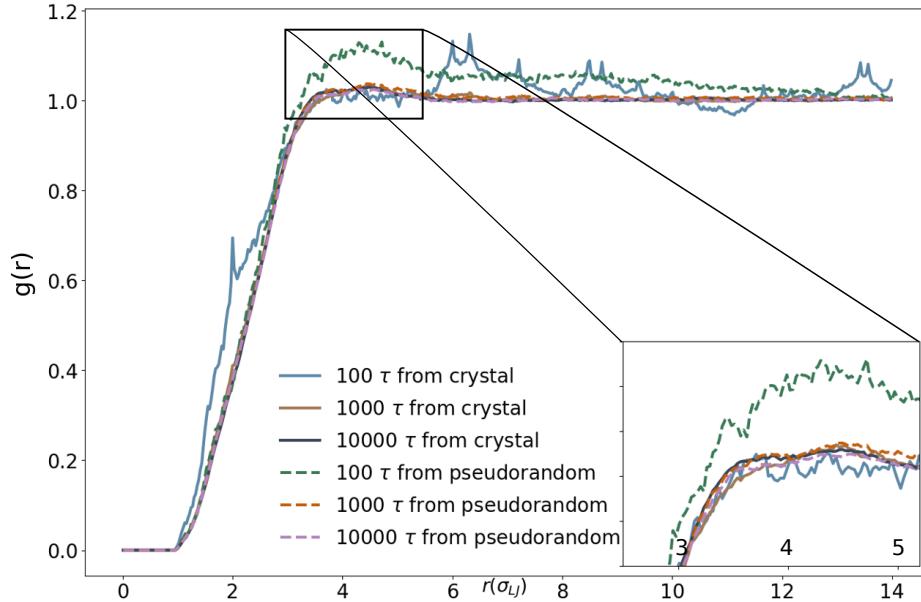


Figure 3.6 RDF for the initial configurations of molecules arranged in an ordered configuration (continuous lines) and pseudo-random arrangement (dashed lines). The RDF is plotted at different time intervals until the molecules are not correlated, i.e. they are mixed, and their positions are independent. The embedded figure is the zoomed region.

3.1.1 Radial Distribution Function of the Initial Configuration

The structural characteristics a system can be described by the RDF:

$$g(r) = \frac{1}{4\pi r^2 \rho N} \sum_{i=1}^N \sum_{j \neq i}^N \langle \delta(|r_{ij} - r|) \rangle, \quad (3.4)$$

where r is the distance between two particles, N is the number of the particles with number density ρ . The double summation is over all the pairs that there can be in a system of N particles. The RDF expresses the number of particles that are in a distance r from one particle, summed over all the particles. The more flat and close to 1 is the $g(r)$, the less correlated are the particles positions. If a system has a crystal configuration, then characteristic peaks appear in $g(r)$.

When creating an initial configuration, the particles need to be uncorrelated, so that no bias present in the system. This correlation is measured by $g(r)$ while running an MD simulation with switched off the patchy interactions. This allows the system to achieve a random initial configuration before switching back on the patchy interactions. The results of monitoring the RDF for various τ is presented in Figure 3.6. The solid and dashed lines are for systems that were created in ordered and random arrangements respectively. It is

Table 3.2

type	arm #	Sticky end	Segment I	Segment II
C	Y1	5'- CGA TTG ACT CTC	CAC GCT GTC CTA AC	CAT GAC CGT CGA AG-3'
C	Y2	5'- CGA TTG ACT CTC	CTT CGA CGG TCA TG	TAC TAG ATC AGA GG-3'
C	Y3	5'- CGA TTG ACT CTC	CCT CTG ATC TAG TA	GTT AGG ACA GCG TG-3'
D	L1	5'- GAG AGT CAA TCG	TCT ATT CGC ATG AC	ATT CAC CGT AAG-3'
D	L2	5'- GAG AGT CAA TCG	CTT ACG GTG AAT GT	CAT GCG AAT AGA-3'

Table presenting the base sequences of the three arms for C-type Y-DNA and D-type linear linkers for the formation of the hydrogel. The reason why they are called C-type and D-type is to make a distinction from the system presented in Table 3.1.

apparent that the ordered configuration has more distinctive peaks, but after equal time ($10^4\tau$) the configuration becomes uncorrelated. From the comparison of the ordered and random arrangements, it appears that the random initial configuration can reach a flat $g(r)$ faster than the ordered one.

3.2 Extension of the System

The two types of Y-DNA building blocks that have been described so far, constitute one type of building blocks that can form a hydrogel. Another type of building blocks that has been used by Liu et. al. [47] is made of a single type of Y-DNA molecules and DNA linear linkers. The Y-DNA molecules cannot connect directly, but only through the linkers. Each linker is made of two ssDNA components that hybridise in order to form a dsDNA part with ends that are ssDNA. The base sequences that form the system of Y-DNA and linear linkers are presented in Table 3.2.

The systems of Y-DNA with linear linkers can also be studied with MD simulations, in a CG fashion. An example of a randomly arranged CG initial configuration is presented in Figure 3.7. The dimensionality of this system does not diverge considerably from the system of A and B type Y-DNA. Hence the same model will be exploited, according to which, 7 beads with 3 patches represent Y-DNA molecules, and 4 beads with 2 patches at the edges are the linkers. The CG system is presented in Figure 3.8. The patchy interactions are also the same as described previously.

The behaviour of this system is not far from the system of two types of Y-shaped molecules, as will be shown in Chapter 3. The system of Y-DNA and linear linkers, however, allows for the substitution of the linkers with other variations that can lead to aberrant viscoelastic behaviour as will be discussed later. Alterations may include the change of

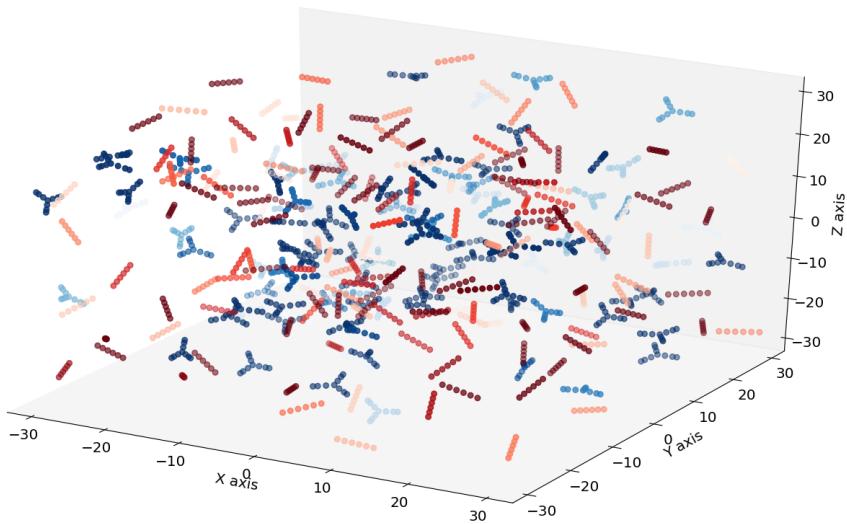


Figure 3.7 Randomly arranged initial configuration of Y-shaped molecules (blue-coloured), and linear linkers (red-coloured) in a simulation box of $27000\sigma_{LJ}^3$. The ratio of Y-shapes to linkers is 2 : 3 with 100 Y-shaped molecules.

length of the linear linkers (in either the hybridised part, or the sticky ends), the incorporation of base pair mismatches, or the addition of sequences of T that do not participate in any bonding (flexible ends). Such alterations may affect the final mechanical behaviour of the hydrogel.

The linear linkers presented in the Table 3.2 can be altered by incorporating a free end (consisting of T bases) between the sticky end, and the first segment. Such variations have recently been used in experiments conducted within the Eiser-group, and it has been shown that the free end affects the flexibility of the angles formed between the Y-DNA and the linear linkers. It can also lead to extraordinary behaviour during the melting process [72]. This alteration can be incorporated in simulations in order to reflect experimental results by decreasing the force constant in the angle potentials applied on the patches. This way, the patches can move freely around the bead at various angles. Figure 3.9 presents the angle distributions that the patches have with respect to the rigid arm when the angular potential is softer, compared to a harder potential (presented in the embedded figure) in which the force constant is increased by two orders of magnitude.

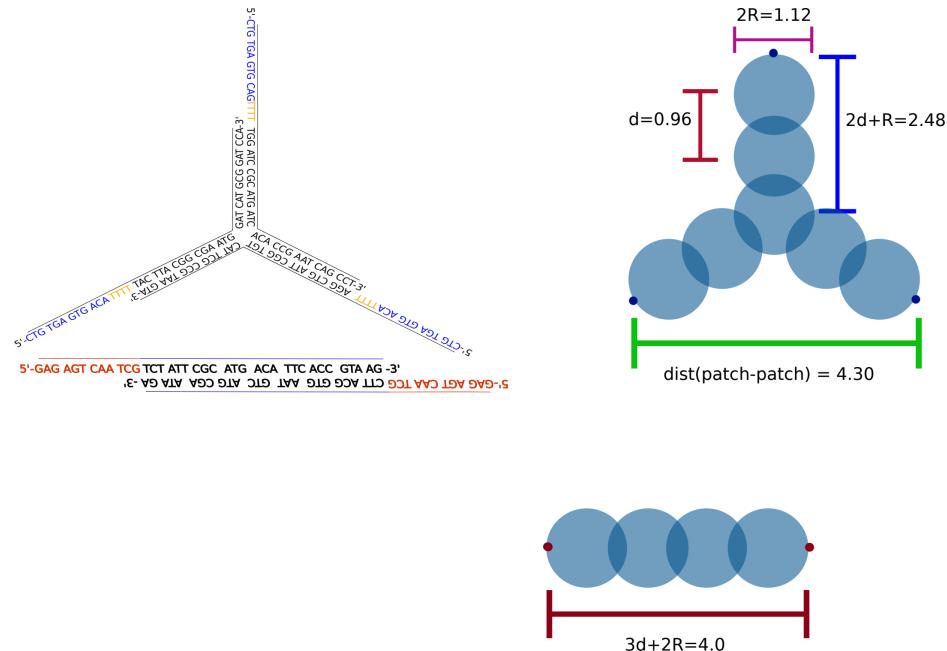


Figure 3.8 Base pair sequences (left) and coarse grained description (right) of Y-DNA and linear linker molecules. The lengths are expressed in reduced units (σ_{LJ}) and are chosen so that the geometry of the CG molecules reflect the actual geometry of the Y-DNA and linear linker molecules.

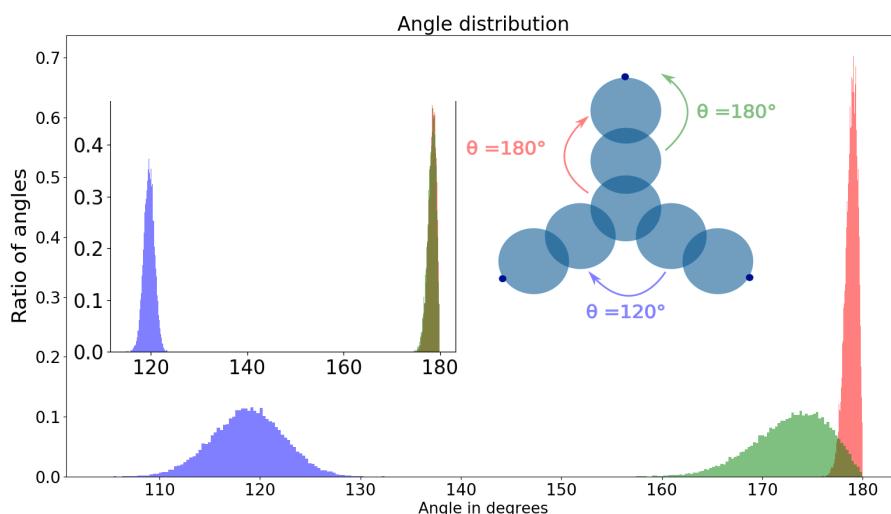


Figure 3.9 Comparison of angle distributions for the patches when the angular potential is soft (full figure) with a harder potential (embedded figure). The force constant of the soft potential is $K_\theta = 1\epsilon_{LJ}/rad^2$, whereas for the harder it is $K_\theta = 100\epsilon_{LJ}/rad^2$. This results in the angles that the patches can reach to have values between 160° and 180° compared to rigid patches that reach angles close to 180° .

Chapter 4

Results and Discussion

In this chapter I will focus on the results that I have acquired so far on simulations of DNA hydrogels. The results concern the equilibration of the initial random configurations at various densities and for different realisations of the system. Additionally, I include the study of the network formation, and the derivation of viscous and elastic moduli. These studies are extended on two different systems that form hydrogels, as described in Chapter 3.

4.1 Model Setup and Melting Curves

Before initiating an MD calculation the molecules should be uncorrelated, so that the initial configuration is random. This is tested with the RDF before the equilibration process starts. The RDF plots that verify randomness of the system were discussed in Subsection 3.1.1. During the simulation, as time progresses, the number of the bonds that are forming between the molecules are recorded for different temperatures. The steady state is achieved after approximately $10^5 \tau$ have passed.¹ The steady state is characterised by a limited fluctuating number of the pairs that are formed.

4.1.1 Number of Pairs

During the simulation time, the number of pairs formed is recorded at constant time intervals. From this number, the association degree θ is calculated from the formula

$$\theta = \frac{N_{conn}}{(Q_{val}N_{tot})/2}, \quad (4.1)$$

¹The steady state is characterised by the network formation as it will be described in the following paragraphs.

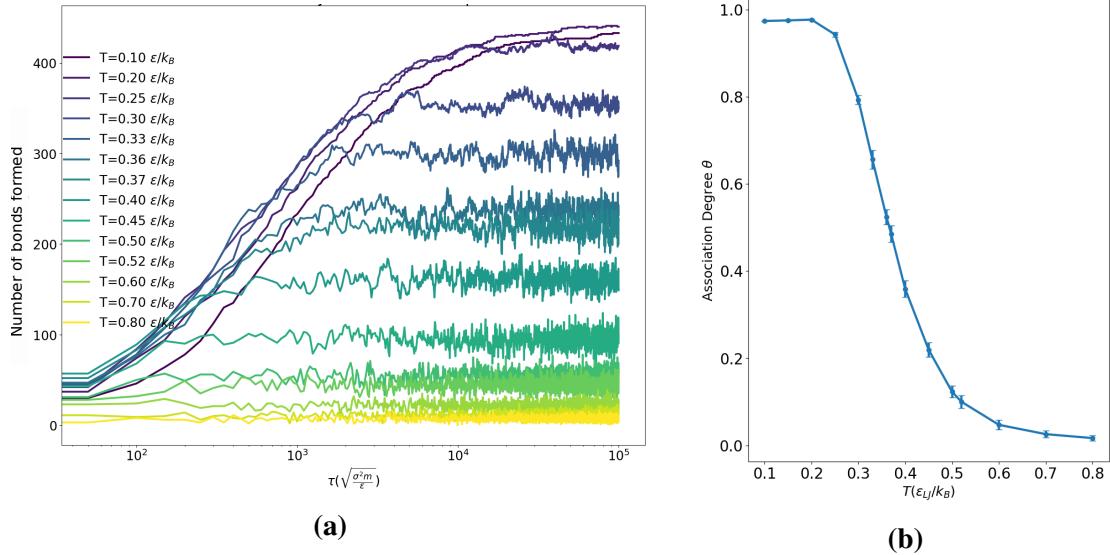


Figure 4.1 Time evolution of the bond formation between the Y-shaped molecules (a), and association degree in the steady state (b). The error bars represent the standard deviation of the final 1000 time steps that were used for the calculation of the association degree θ . These plots are for a system consisting of 300 molecules in a box of volume $27000\sigma_{LJ}^3$ which corresponds to a number density of $0.0111\sigma_{LJ}^{-3}$ and volume density $\rho = 5.58\%$.

where N_{tot} is the number of the molecules, and Q_{val} is the valency of these molecules, i.e. $Q_{val} = 3$, as is the number of patchy particles per molecule. The association degree θ describes the fraction of pairs that have formed N_{conn} out of the total available pairs. For example, the number of pairs formed throughout a simulation is presented in Figure 4.1a, and the corresponding association degree is in Figure 4.1b.

The number of pairs formed as a function of time has been calculated from at least five independent realisations of the same system. This means, that initial configurations of the same density of particles were created, and MD simulations run independently. Thus, the association degree emerged from these many systems that were plot together as seen in Figure 4.2a. This process was repeated for systems of various densities as presented in Table 4.1, and the association degree plots are included in the Appendix A.4, Figures A.1. The average of the association degree at all studied densities of the system are shown in Figure 4.2b.

It is obvious from Figure 4.2b, that for concentrations which are very close to one another, the association degree curves almost overlap. For instance, the green and red lines for number densities $0.008\sigma_{LJ}^{-3}$ and $0.0082\sigma_{LJ}^{-3}$ respectively, the curves are almost identical. This shows that even if smaller number of molecules within a small box are used, the association degree does not change compared to using a larger number of molecules in a larger box. Thus,

Table 4.1

Number of molecules	box dimension	number density	volume density
	σ_{LJ}	σ_{LJ}^{-3}	
50	30	0.0019	0.93%
100	30	0.0037	1.86%
1000	50	0.0080	4.02%
200	29	0.0082	4.12%
700	40	0.0109	5.49%
300	30	0.0111	5.58%

This table includes all the different concentrations that were used in the simulations. The first column is the number of Y-shaped molecules, and the second is the size of the edge of the cubic simulation box. The number density of the system (third column) is the number of molecules per box volume, and the volume density is the volume occupied by all the molecules divided by the volume of the box.

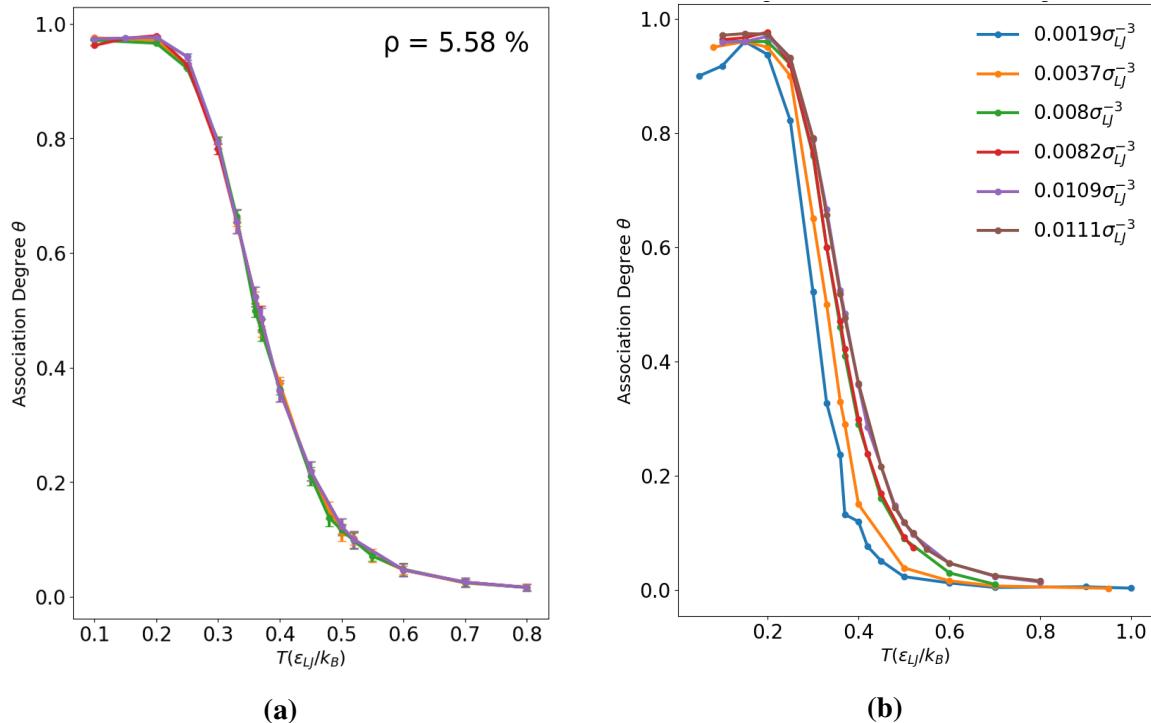


Figure 4.2 Plots of the association degree θ for systems of Y-shaped molecules. The plot (a) consists of overlapping curves that correspond to different initial configurations of systems with the same density of molecules. The plot (b) is for Y-shaped systems with different density of molecules. The melting transition shifts towards lower temperatures for dilute systems.

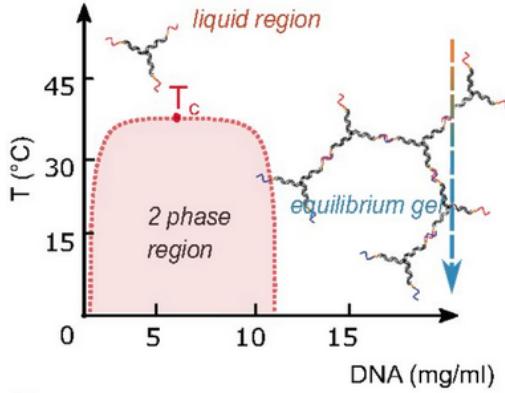


Figure 4.3 Phase diagram for a system of three-valent DNA nanostars. The highlighted region indicates a phase separation between liquid and gel phase at low DNA concentrations, and temperatures lower than the critical temperature (T_c). At higher DNA concentrations, where the arrow indicates, there is a continuous phase change from liquid to gel, as the temperature drops. The figure is taken from [41].

the simulations can be accelerated by not increasing the system size beyond a certain point. Moreover, it is noted that for lower density of molecules the association degree curve shifts towards lower temperatures. This is due to the system being more dilute, and the forming of pairs between molecules is not hindered by the excluded volume of neighbouring molecules.

The association degree for a system's density of 0.93% (or equivalently a number density of $0.0019\sigma_{LJ}^{-3}$) is considerably below unity, and it is also below the association degrees of all studied densities (Figure 4.2b). This may indicate that a fully percolating gel has not been formed, even at very low temperatures ($T < 0.2\epsilon_{LJ}/k_B T$). This can be attributed to two factors. Firstly, the steady state may have not been reached, and secondly there may be a phase separation due to the low density of molecules. From the phase diagram of three-valent DNA nanostars (Figure 4.3), there can be a phase separation between liquid and gel phase, and the system density 0.93% can be in this region. Further study of low density systems is still in progress.

4.1.2 Radial Distribution Function for the Steady State

Structural characteristics of the steady states can be described by the RDF. The RDF was mentioned in Subsection 3.1.1, where the randomness of the initial configuration had to be tested. In this case the RDF is used for the determination of the structural characteristics during the MD simulation until the molecules have reached a steady state. Two RDF plots of systems with different densities, 4.02%, and 5.58%, are presented in Figures 4.4a and 4.4b respectively. It appears that for densities above 4.02%, the RDF does not change considerably.

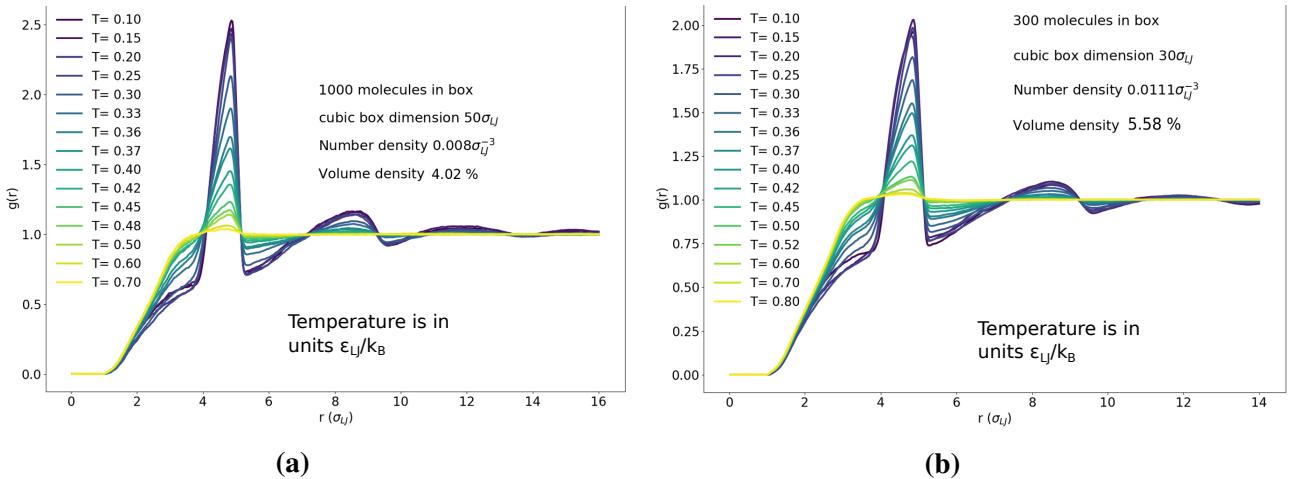


Figure 4.4 RDF plots for systems at different densities. The plot (a) is for a density of 4.02%, and the plot (b) if for a density of 5.58%. The plots are appear in a different scale, since they were derived individually. The difference of the height of all the peaks on both plots are due to the different densities of the systems. The positions of the peaks though, are the same.

There is though a difference in the peaks' height due to the different systems' densities, but the peaks' positions are the same.

In the RDF plots (Figures 4.4a and 4.4b) the first peak that is in the position $r \approx 4.7\sigma_{LJ}$ corresponds to the distance of the first nearest neighbour of a building block connected to another. The second peak at $r \approx 8.3\sigma_{LJ}$ is the second nearest neighbour of each building block. The third and wider peak at $r \approx 12.0\sigma_{LJ}$ is the third nearest neighbour of a building block. It is broad due to the wide range of angles that a building block can have with its third nearest neighbour.

For systems with low density of molecules (0.93%), the RDF plot changes considerably, as shown in Figure 4.5a. Specifically, an additional peak appears in the plot, which corresponds to an additional arrangement of molecules at low temperatures. The new arrangement consists of closed loops of molecules, as shown in a simulation instant in Figure 4.5b. The distance between a molecule and its second nearest neighbour is approximately $6.83\sigma_{LJ}$, and this is reflected to the appearance of an extra peak in the RDF plot (Figure 4.5a).

4.1.3 System of Y-shaped Molecules with Linear Linkers

For the additional systems made of Y-shaped molecules with linear linkers², the systems studied have a ratio of Y's to linkers of 2 : 3, where a saturation of the total number of bonds

²It is worth mentioning that the systems including linear linkers have no flexibility in the patches attached to the linkers. The distribution of the angles that the patches is very narrow and close to 180° . A softer force constant in the angular potentials of the patches is part of future work.

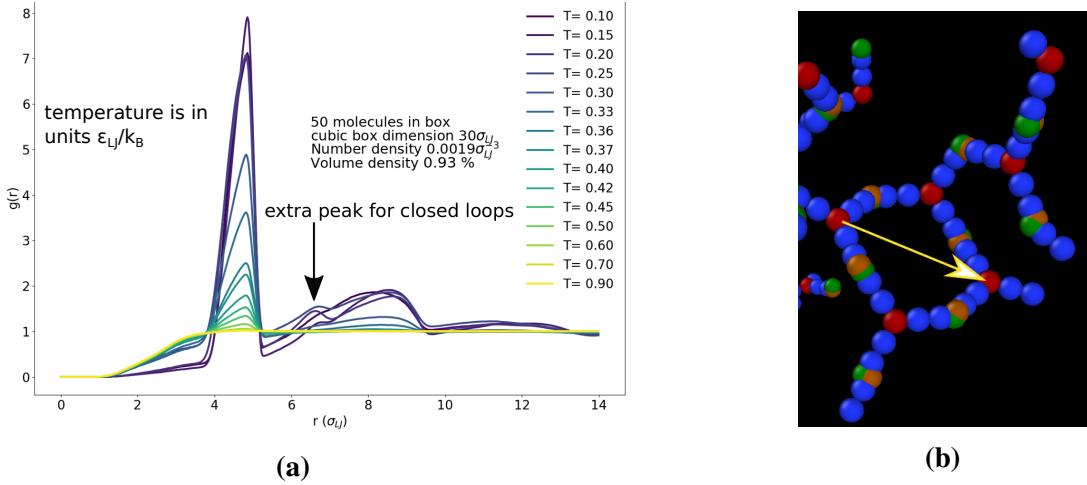


Figure 4.5 RDF plot for a system density of 0.93% shown in (a). Closed loops formed of four Y-shaped molecules contribute to the formation of the additional peak in the RDF plot 4.5a. Figure (b) presents an example of the closed loop taken from a simulation instance. The yellow arrow is the characteristic distance of a molecule with its second nearest neighbour, and is equal to $6.83\sigma_{LJ}$.

is expected. A ratio of 2 : 3 means that for each 2 Y-shaped molecules there are 3 linear linkers, and since the valence of each component is 3 and 2 respectively, then there can be no unconnected patches (unsaturated bonds). A ratio beyond that, will lead to many unconnected molecules, hence gelation may not appear.³ In Figure 4.6a the number of the pairs formed is plotted with respect to time, and it reaches a maximum of 600 since the number of Y-shaped molecules is 200, and of linear linkers in 300. The corresponding association degree is presented in Figure 4.6b, and it is below unity for temperatures lower than $0.2\epsilon_{LJ}/k_B$. An explanation cannot be derived yet, since calculations with different concentrations are still in progress.

The RDF plot for the aforementioned system is presented in Figure 4.7. It appears that some of the peaks are quite wide, and they are also considerably different compared to the RDF peaks of systems comprised of Y-shaped molecules only. This can be attributed to two reasons. Firstly, there is a considerable number of closed loops formed between the constituents. This gives rise to much wider peaks appearing in the RDF. Representative pictures of 2 and 3 Y-shaped molecules in closed loops are shown in Figures 4.8a and 4.8b. Secondly, the number density of the system is quite low ($0.0018\sigma_{LJ}^{-3}$), hence the system apart from the main network of molecules may have many smaller sub-networks. This has not been verified yet, due to not concluded calculations.

Specifically, in Figure 4.7 the first peak that appears at $r \approx 8.0\sigma_{LJ}$ corresponds to two Y-shaped molecules connected through linear linkers. The maximum distance that two

³The lack of gelation is a speculation that has not been tested yet.

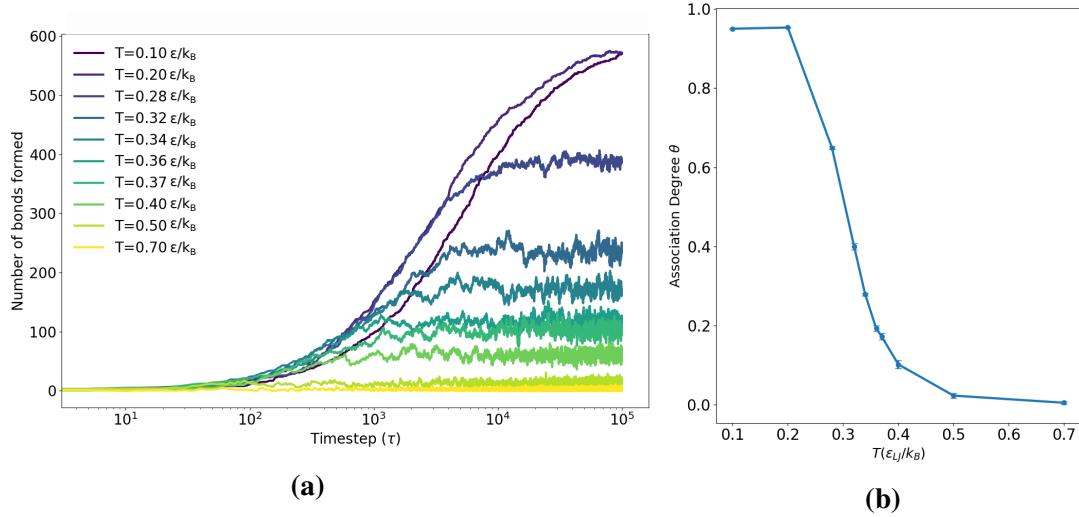


Figure 4.6 Time evolution of the bond formation between the Y-shaped molecules and linear linkers. The system consists of 200 Y-shaped molecules and 300 linear linkers in a box of volume $216000\sigma_{LJ}^3$ which corresponds to a number density of $0.0018\sigma_{LJ}^{-3}$.

molecules can have is equal to $9.96\sigma_{LJ}$ which is equal to the length of the linear linkers plus the length of two arms of the Y-shaped molecules. At $T = 0.20\epsilon_{LJ}/k_B$ an additional peak appears at $r \approx 6.0\sigma_{LJ}$ which corresponds to Y-shaped molecules being closer to one another. A fully connected network has been formed with closed loops of 3 and 4 molecules. The second and third peaks that appear in the RDF plot are not distinct, and can possibly include more than one peaks.

4.1.4 Characterisation of the Steady State

It has been proposed that the number of clusters that have been formed in a dynamic system may indicate that the system has attained the steady state [61]. With the aid of graph theory, the network formed within the system can be analysed, and the clusters can be enumerated. For the systems under consideration, each molecule is represented by a node, and the connections formed between the molecules are represented by edges. The degree of each node is the number of edges that correspond to one node. For a system of Y-shaped molecules, the maximum degree per node is 3, whereas for linear linkers it can be up to 2.

A cluster is defined as a collection of nodes that are connected with other nodes of the same cluster with at least one edge. A single node with no edges is a single cluster. During the equilibration process an approximately constant number of clusters that have been formed in the system is reached. This is presented for a system of 300 molecules and at various temperatures in Figure 4.9. At temperature $0.10\epsilon_{LJ}/k_B$ there is a single cluster, suggesting

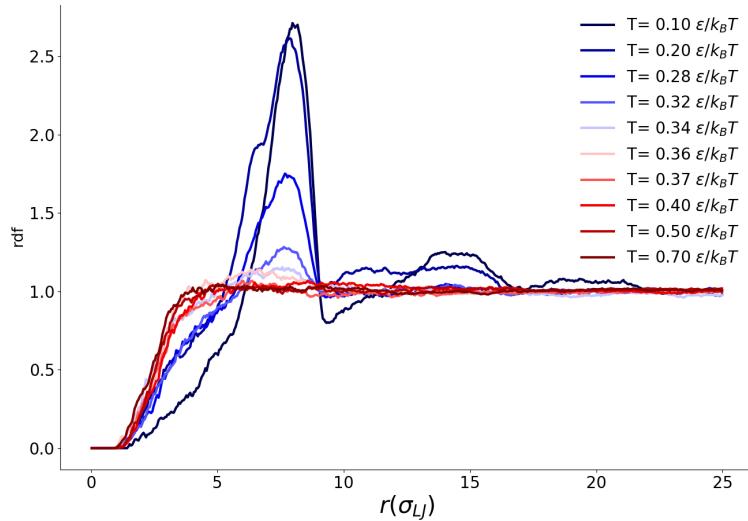
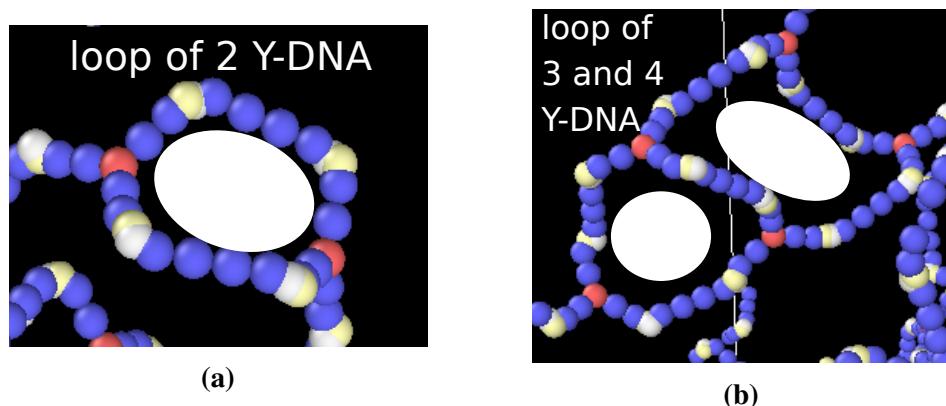


Figure 4.7 RDF plot for the system with 2 : 3 ratio Y-shaped to linkers with number density $0.0018\sigma_{LJ}^{-3}$. There is a variation in the peaks compared to the RDFs of only Y-shaped systems. This is due to the various closed loops that are formed in the system.



Closed loops formed out of 2 (a), and 3 and 4 Y-shaped molecules (b), formed in a system with 2 : 3 Y-shaped to linear linkers ratio. The inscribed white circles are indicators of the Y-shaped loops formed.

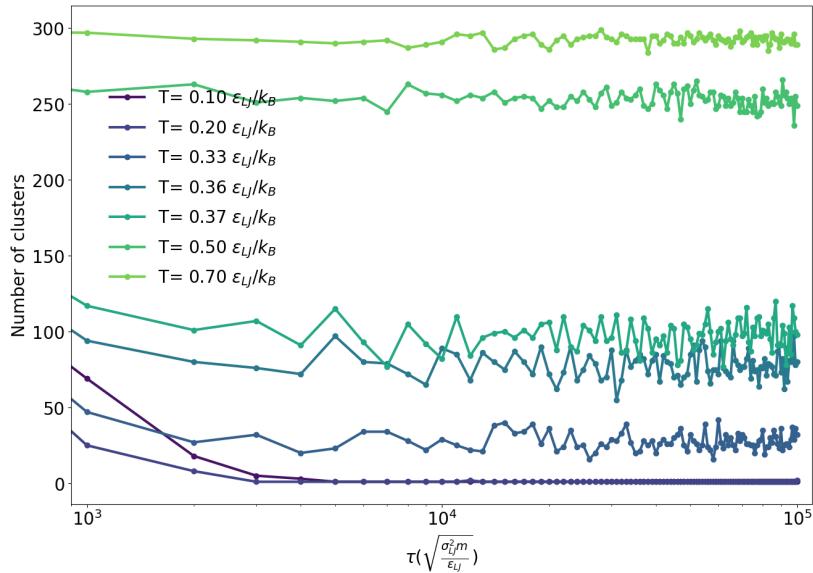


Figure 4.9 Number of clusters that are formed at various temperatures. At $\tau = 1$, all the molecules are separate, leading to a number of clusters equal to the number of molecules in the box, which is 300 for a density of 5.58%.

that the system is a fully percolating gel. At high temperatures (above $0.70\epsilon_{LJ}/k_B$), there are approximately 300 clusters, indicating a liquid state of isolated clusters. At intermediate temperatures there must exist a temperature that corresponds to the percolation threshold or the T_m but it has not been identified so far.

Since Y-shaped molecules have a constant valency of three, then a maximum degree per node equals to three, and a minimum degree per node is zero. The degree plots vary with temperature, since temperature affects the association degree of the molecules. For a system of 300 molecules (300 nodes) the number of the molecules with the same degree is presented in Figures 4.10 for various temperatures. Before the simulation is initiated ($\tau = 0$) all the nodes have a degree of zero, and as time evolves, the number of nodes with the same degree is unchanged at $T = 0.10, 0.20, 0.70\epsilon_{LJ}/k_B$. For the temperatures close to the transition region, i.e. $T = 0.33, 0.37, 0.50\epsilon_{LJ}/k_B$, the number of nodes with the same degree is highly fluctuating, which is indicative of a dynamic system.

From the collection of plots in the Figures 4.9 and 4.10 it appears that the steady state is achieved at different time steps for each temperature. The plots of the degrees appear to be more accurate for low temperatures, as the number of particles that have a constant degree of connectivity converges at lower time steps. For higher temperatures, the two plots (number of clusters and number of particles with the same degree) seem to be in agreement. For intermediate temperatures in the transition region, the system is quite noisy for the determination of the steady state.

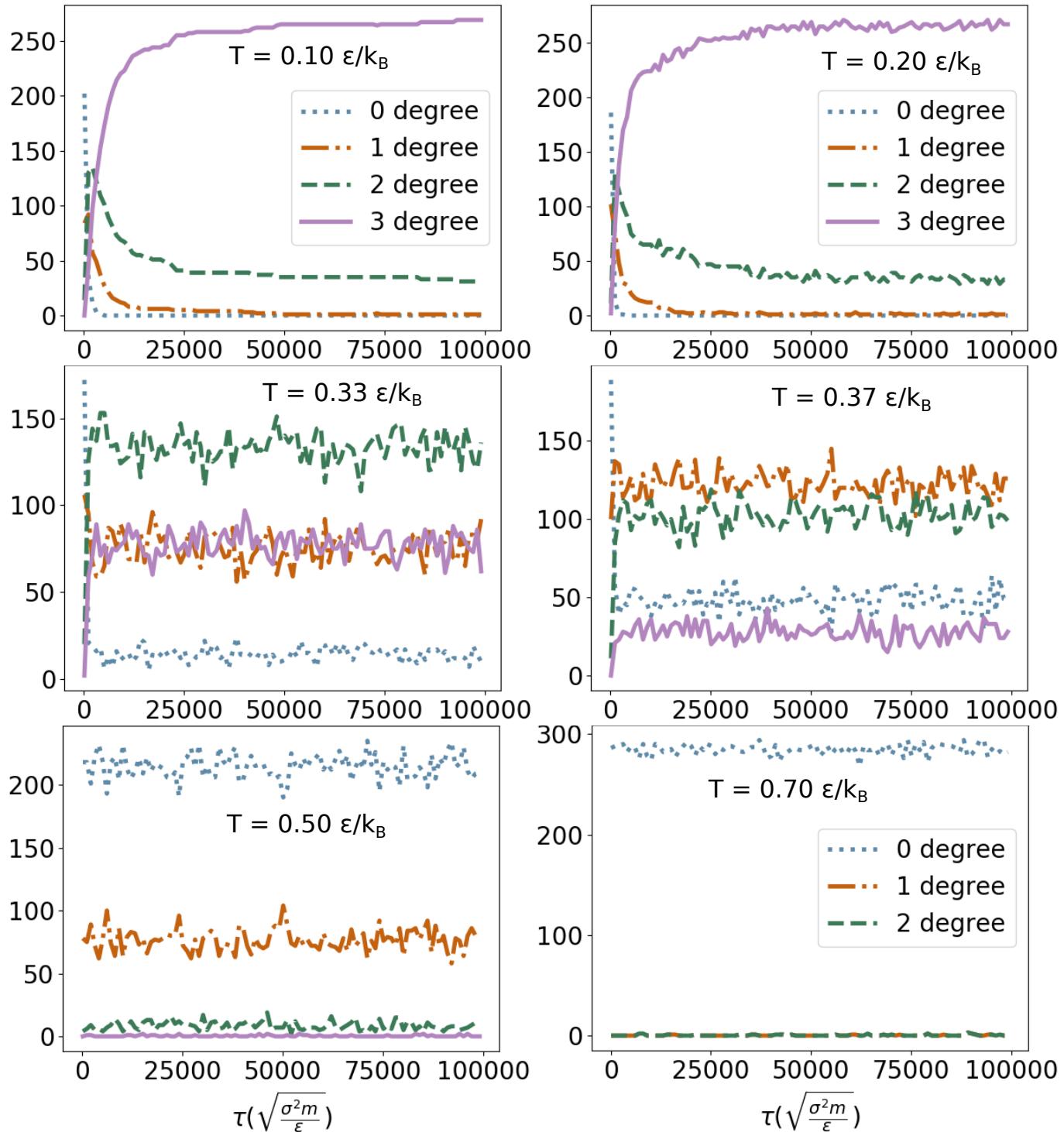


Figure 4.10 Number of nodes with the corresponding connectivity degree as a function of time for various temperatures. The x axis is the time in τ , and the y axis is the number of nodes.

4.2 Cluster Size Distribution

During the equilibration process, the number of clusters that are formed at each time step is recorded. A cluster is defined as the group of molecules that have at least one connection with other molecules that belong to the same cluster. At each temperature, there is a constant number of clusters formed, as presented in the Figures 4.9. At various temperatures, the size of all of the clusters formed over many time steps is presented in the form of a cluster size distribution.

For the system density of 5.58%, the cluster size distribution is plotted over 100 time steps as presented in Figure 4.11. The lines connecting the middle of each bar are also plotted. The cluster size distribution for one time step is also presented in Figure 4.12. The same analysis was applied to a system density of 4.02%, and the corresponding plots for 10 time steps, and a single time step are presented in Figures 4.13 and 4.14 respectively.

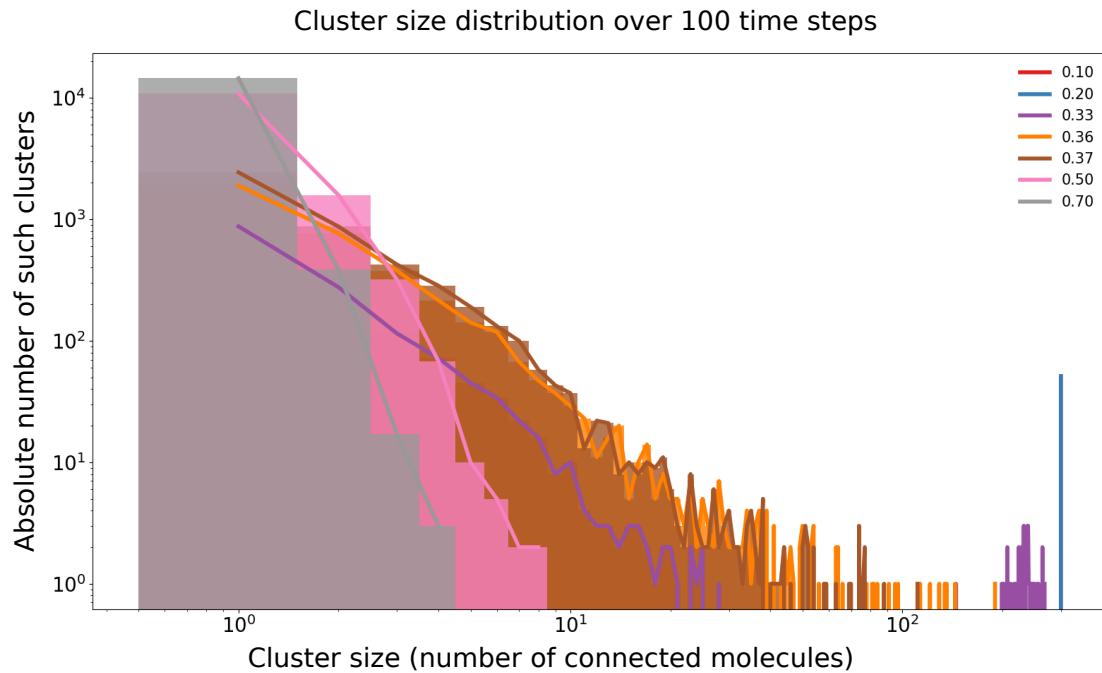


Figure 4.11 Cluster size distribution for Y-shaped molecules at density 5.58%. Different colours indicate different temperatures in ϵ_{LJ}/k_B units, as shown in the legend. 100 time steps of the simulation were used for this plot.

Fitting was attempted on the cluster size distributions for a single time step, and the equations are presented in Tables 4.2 and 4.3 for densities 5.58% and 4.02% respectively. In these tables, the independent variable x expresses the size of the clusters. The variable y expresses the number of such clusters that appear within the simulation box, and it is below unity since it is normalised over the number of time steps during which the clusters were

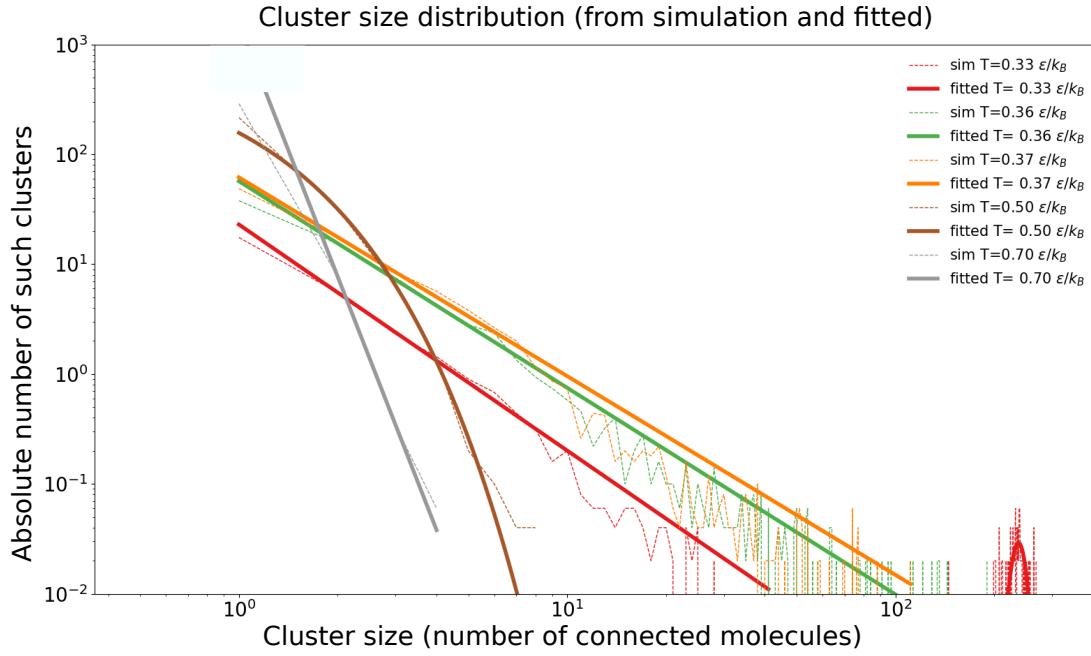


Figure 4.12 Average cluster size for Y-shaped molecules at density 5.58% over one time step. This plot is averaged over 100 time steps. The dashed lines represent the cluster size from the average, and solid lines are for the fitted curves.

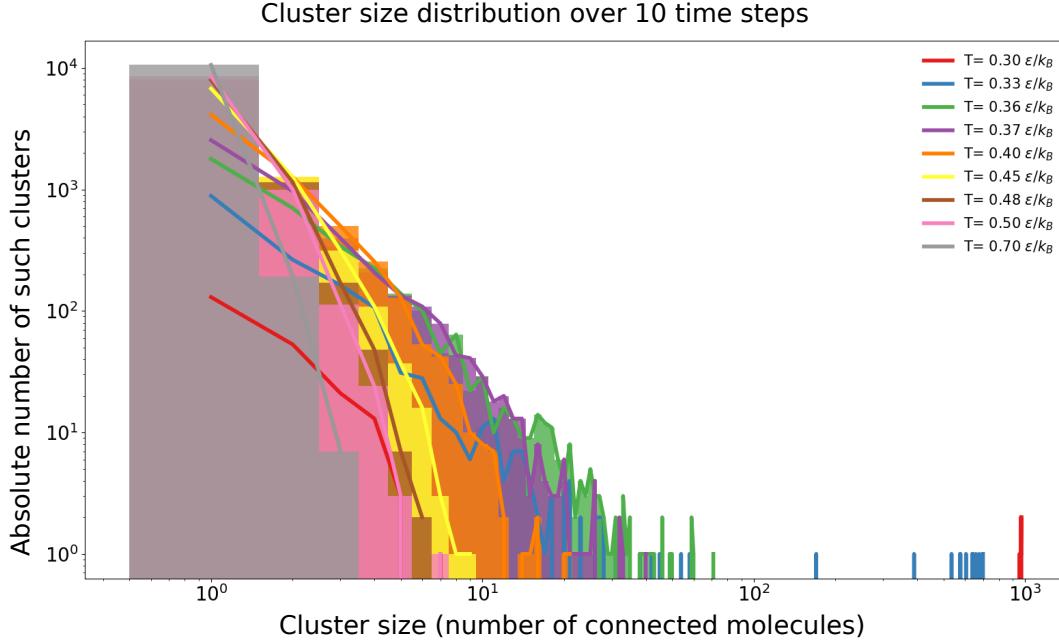


Figure 4.13 Cluster size distribution for Y-shaped molecules at density 4.02%. Different colours indicate different temperatures in ϵ_{LJ}/k_B units, as shown in the legend. 10 time steps of the simulation were used for this plot.

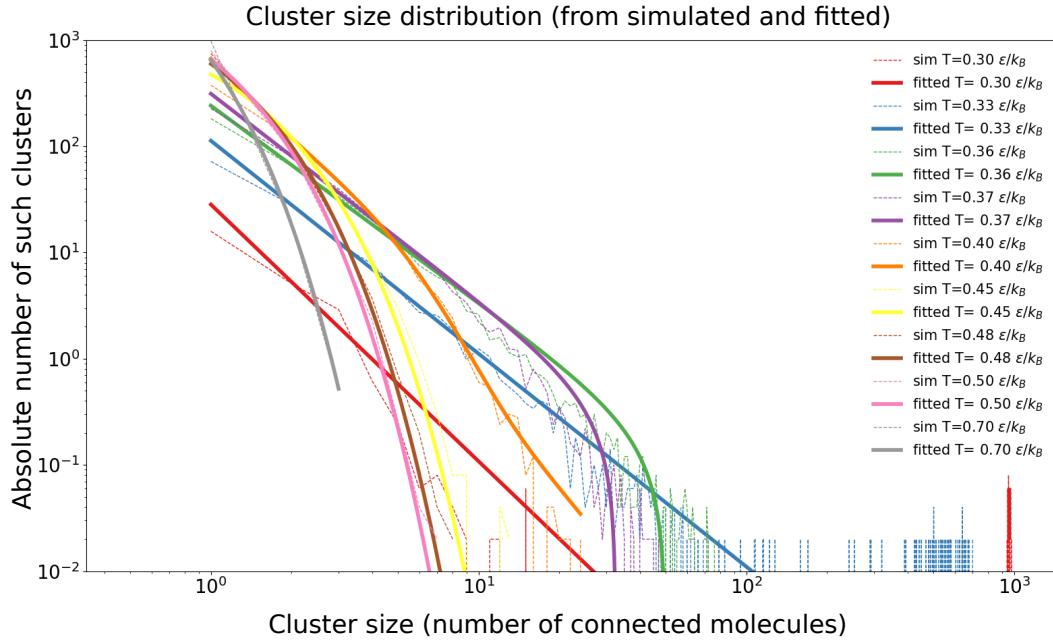


Figure 4.14 Average cluster size for Y-shaped molecules at density 4.02% over one time step. This plot is averaged over 10 time steps. The dashed lines represent the cluster size from the average, and solid lines are for the fitted curves.

counted. For both systems, at temperatures close to the melting transition ($T = 0.36\epsilon_{LJ}/k_B$ and $T = 0.37\epsilon_{LJ}/k_B$), the fitted functions almost approach each other.

The fitting attempted is far from complete. Many realisations of a system's density need to be included in the networks analysis. Moreover, additional time steps can be included (more than 100 as were used in one case). This increases the computational time, especially when the systems include number of particles of the order of 1000. For temperatures close to the melting transition ($0.33\epsilon_{LJ}/k_B$) there are two groups of clusters. One that can be described by a curve of the type $a \cdot x^{-b}$, and another in the region closer to the total number of the nodes. In the second case fitting is still in progress.

For the systems of Y-shaped molecules with linear linkers, each node represents either a Y-shaped molecule, or a linear linker. The cluster size distribution for such a system with ratio 2 : 3 and number density $0.0018\sigma_{LJ}^{-3}$ is presented in Figure 4.15. Further analysis on these results is not included in this thesis due to the unfinished calculations.

Table 4.2

y (5.58%)	T
	(ϵ_{LJ}/k_B)
$23.0 \cdot x^{-2.06}$	0.33
$56.9 \cdot x^{-1.88}$	0.36
$61.7 \cdot x^{-1.81}$	0.37
$775.8 \cdot e^{-1.60x}$	0.50
$1551.7 \cdot x^{-7.66}$	0.70

Table 4.3

y (4.02%)	T
	(ϵ_{LJ}/k_B)
$28.3 \cdot x^{-2.42}$	0.30
$112.8 \cdot x^{-2.01}$	0.33
$241.1 \cdot x^{-1.86}$	0.36
$312.6 \cdot x^{-1.95}$	0.37
$572.7 \cdot x^{-3.06}$	0.40
$1872.7 \cdot e^{-1.37x}$	0.45
$3541.8 \cdot e^{-1.78x}$	0.48
$4850.8 \cdot e^{-2.00x}$	0.50
$22738.9 \cdot e^{-3.56x}$	0.70

Fitting functions for system density 5.58% (Table 4.2), and for system density 4.02% (Table 4.3).

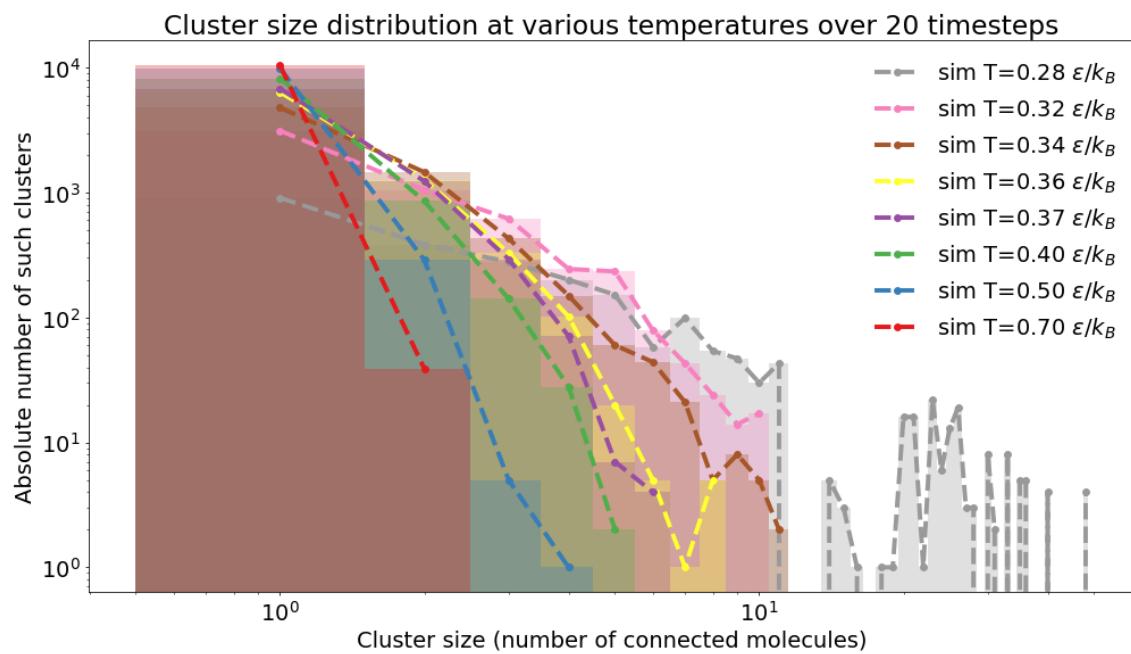


Figure 4.15 Clusters size distribution for Y-shaped molecules and linear linkers at number density $0.0018\sigma_{LJ}^{-3}$.

4.3 Viscoelasticity

Viscous and elastic moduli can be derived from NEMD by shearing the simulation box, or by equilibrium calculation of the pressure autocorrelation function. The first results were calculated from equilibrium calculation on a system with density $\rho = 4.02\%$, where the melting temperature is $T_m = 0.36\epsilon_{LJ}/k_B$. At high temperature ($T = 0.90\epsilon_{LJ}/k_B$) the results for G' and G'' are shown in Figure 4.16. The plot is quite noisy which can be attributed to the limited time steps that were used in simulations. Despite the noise, it appears that G'' is higher than G' , which is indicative of a liquid system. The plateau that appears for $\omega > 10\text{rad}$ can be omitted, since G' is affected by the spring constant connecting the beads of the Y-shaped building blocks.

For temperatures lower than T_m (specifically $T = 0.10\epsilon_{LJ}/k_B$), G' and G'' appear in the Figure 4.17. The shear stress moduli were calculated from NEMD, hence the plots are not noisy. By applying oscillatory shear to the system, and measuring the phase lag between the shear strain and the pressure exerted on the system, G' and G'' can be derived, as was previously described. The elastic modulus surpasses the viscous modulus, which indicates that the system is in the gel phase. The box deformation applied to the box is 2%, and the system remains in the linear regime, as it is mentioned in [70].

The plots 4.16 and 4.17 can provide us with qualitative information on the behaviour of hydrogels at high and low temperatures respectively. The elastic and viscous moduli, G' and G'' have not yet been compared with analogous experimental results. This is part of the future work that will be done. Moreover, for systems of various densities, and systems including linear linkers, NEMD calculations for the derivation of G' and G'' are still in progress, hence they are not included in this thesis.

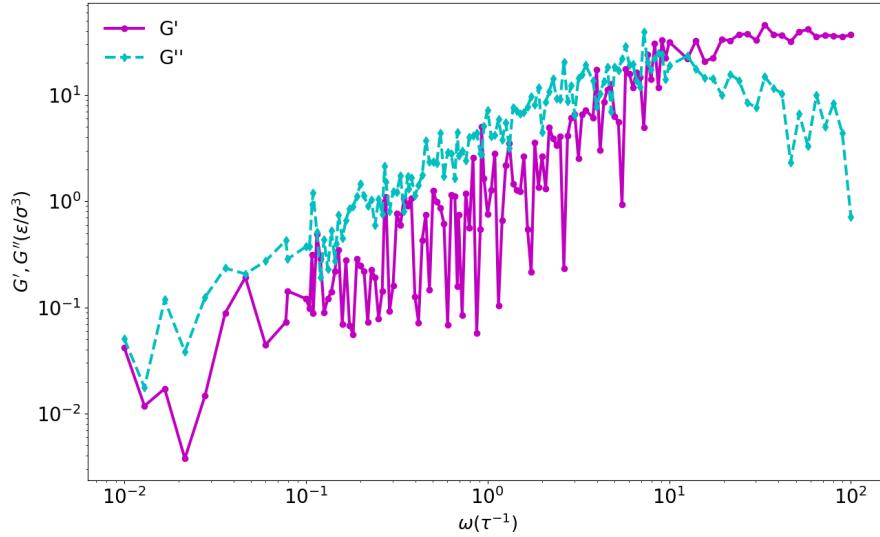


Figure 4.16 Storage G' and loss G'' moduli at temperature higher than T_m ($T = 0.90\varepsilon_{LJ}/k_B$) as a function of frequency ω at a density $\rho = 4.02\%$ calculated from Kubo theory.

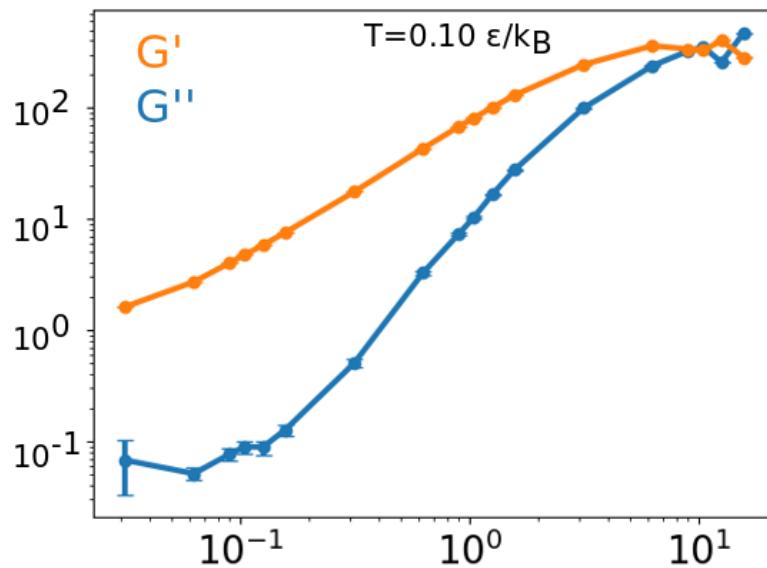


Figure 4.17 Storage G' and loss G'' moduli at temperature lower than T_m ($T = 0.10\varepsilon_{LJ}/k_B$) as a function of frequency ω at a density $\rho = 4.02\%$ calculated from NEMD.

Chapter 5

Conclusions

In this thesis I presented a CG model for the description of DNA hydrogels. This model was introduced by Xing et al. [70], and I extended it in order to incorporate alternative building blocks that can form hydrogels. With MD simulations I derived the association degree, the RDF, the network formation, and mechanical properties of the systems. The systems used, were of various concentrations, and the results are discussed in Chapter 4. In the following paragraphs I am presenting an outlook of the work that is still in progress, and further amendments that can be done.

The first alteration that can be done on the hydrogels formed of two types of Y-shaped molecules is on the base sequences of the sticky ends. These base sequences were designed in a way such that Y-DNA can only bind to Y'-DNA. A proposition is that both sticky ends can be designed in a palindromic fashion, so that only one type of Y-DNA is required for the formation of DNA hydrogel. This could possibly accelerate the hydrogel formation, or can lead to different behaviour of the mechanical properties of the system.

As long as the DNA hydrogel with Y-DNA and linear linkers is concerned, further alterations can be incorporated in the linkers. The changes on the linear linkers can be on altering their total length, or the incorporation of a free end between the rigid dsDNA and the sticky end. Such changes can lead to formation of DNA hydrogels with quite aberrant mechanical behaviour. In this thesis I have not included any results that correspond to linear linkers with flexible sticky ends. I presented though how this can be done in a CG fashion, but simulations on such systems are still in progress.

During the time of concluding this thesis, members of the Eiser group were working on MD simulations with the oxDNA [73] package that included a more detailed description of the DNA hybridisation process. Results from CG MD simulations, can be amended in a way so that they reflect the DNA behaviour described by oxDNA simulations. For instance,

xDNA can describe more accurately the flexibility between the rigid arms of DNA, and that should be reflected in the CG model. This work is still in progress.

Results that I have presented so far cover a limited set of configurations that can form a hydrogel. These configurations are either Y-shaped DNA molecules, or Y-shaped DNA with linear linkers. Another addition to the system is the substitution of three-valent DNA building blocks with tetra-valent ones. Additionally, the temperature can be ramped up and down around the melting temperature in order to guide self assembly. These changes have already been tested, but there are preliminary results that are not presented in this thesis.

The network analysis describing the connected building blocks was conducted on the total number of the system's components. Alternatively, by monitoring the largest component of the network could yield more insightful information on whether a fully percolating gel has been formed at low temperatures. In this thesis though, the network was studied as an aggregate of all the system's subgraphs, hence it is not clear whether the largest component of the network formed spans the full simulation box.

Network analysis on the DNA hydrogel systems can also yield interesting results concerning the mechanical properties of the system, by giving an insight on the structural characteristics. The fitting on the cluster size distribution can be amended in order to match the simulations accurately, as was discussed previously. Once the changes are completed, the cluster size distribution could be employed as a means for predicting the mechanical properties of the DNA hydrogels simply by studying the system's structural characteristics.

Results on the mechanical properties of DNA hydrogels that I have presented are accurate in the description of the system's behaviour, i.e. whether the system behaves like fluid or solid. In the simulations may appear unexpected phenomena that have not been checked thoroughly. For instance, such a phenomenon may be bond breaking during shearing of the simulation box. Moreover, viscoelastic properties need to be calculated at various densities and temperatures of the systems. NEMD simulations of all the aforementioned alterations are in progress at the time of concluding this thesis.

References

- [1] J. Watson et al. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, 1953.
- [2] C. Prévost, M. Takahashi, and R. Lavery. Deforming DNA: from physics to biology. *ChemPhysChem*, 10(9-10):1399–1404, 2009.
- [3] D. Roden et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clinical Pharmacology & Therapeutics*, 84(3):362–369, 2008.
- [4] N. Seeman and H. Sleiman. DNA nanotechnology. *Nature Reviews Materials*, 3:17068, 2017.
- [5] K. Hoogsteen. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta crystallographica*, 12(10):822–823, 1959.
- [6] Zephyris (Richard Wheeler). DNA — Wikipedia, the free encyclopedia, 2019. [Online]; accessed 23-February-2019.
- [7] J. M. Berg, J. L. Tymoczko, L. Stryer, et al. *Biochemistry, eighth edition*. New York: WH Freeman,, 2015.
- [8] M.L. Bochman, K. Paeschke, and V.A. Zakian. DNA secondary structures: stability and function of G-quadruplex structures. *Nature Review Genetics*, pages 770–780, 2012.
- [9] J. SantaLucia and D. Hicks. The thermodynamics of DNA structural motifs. *Annual Reviews of Biophysical and Biomolecular Structures*, 33:415–440, 2004.
- [10] J. SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95(4):1460–1465, 1998.
- [11] P.R Cantor, C.R.and Schimmel. Biophysical chemistry. *Part II: Techniques for the Study of Biological Structure and Function*, 2:570, 1980.
- [12] A. Carbone and N. C. Seeman. Molecular tiling and DNA self-assembly. 2003.
- [13] R. Holliday. The induction of mitotic recombination by mitomycin C in *Ustilago* and *Saccharomyces*. *Genetics*, 50(3):323, 1964.
- [14] N. C. Seeman. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, 1982.

- [15] Nikolai (Donor). Holliday junction — Wikipedia, the free encyclopedia, 2019. [Online]; accessed: 03-January-2019.
- [16] N. Kallenbach, R. Ma, and N. Seeman. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 305(5937), 1983.
- [17] J. E. Mueller, S. M. Du, and N. C Seeman. Design and synthesis of a knot from single-stranded DNA. *Journal of the American Chemical Society*, 113(16):6306–6308, 1991.
- [18] E. Winfree, F. Liu, L. Wenzler, and N. C Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539, 1998.
- [19] A. P. Alivisatos, K. Johnsson, X. Peng, T. E. Wilson, C. Loweth, M. Bruchez Jr, and P. G Schultz. Organization of nanocrystal molecules using DNA. *Nature*, 382(6592):609, 1996.
- [20] Z. Peng and H. Liu. Bottom-up nanofabrication using DNA nanostructures. *Chemistry of Materials*, 28(4):1012–1021, 2016.
- [21] A. Caciagli et al. DNA-coated functional oil droplets. *arXiv preprint arXiv:1710.07820*, 2017.
- [22] P. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082), 2006.
- [23] L. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266(5187):1021–1024, 1994.
- [24] J. Chen, R. Deaton, and Y. Wang. A DNA-based memory with in vitro learning and associative recall. pages 145–156, 2004.
- [25] A. Extance. How DNA could store all the world’s data. *Nature*, 537(7618):22–24, 2016.
- [26] E. Heckman et al. Processing techniques for deoxyribonucleic acid: Biopolymer for photonics applications. *Applied Physics Letters*, 87(21), 2005.
- [27] F. Lewis. DNA molecular photonics. *Photochemistry and photobiology*, 81(1):65–72, 2005.
- [28] A. Steckl. DNA—a new material for photonics? *Nature Photonics*, 1(1):3, 2007.
- [29] E. Hemmig et al. Optical voltage sensing using DNA origami. *Nano letters*, 18(3):1962–1971, 2018.
- [30] S. H. Um, J. B. Lee, N. Park, S. Y. Kwon, C. C. Umbach, and D. Luo. Enzyme-catalysed assembly of DNA hydrogel. *Nature materials*, 5(10):797, 2006.
- [31] Y. et. Li et al. Controlled assembly of dendrimer-like DNA.
- [32] A. Chandrasekaran and R. Zhuo. A tile tale: Hierarchical self-assembly of DNA lattices. 2:7–16, 12 2015.

- [33] J. Kopeček. Hydrogel biomaterials: a smart future? *Biomaterials*, 28(34), 2007.
- [34] D. Campoccia, P. Doherty, M. Radice, P. Brun, G. Abatangelo, and D.F. Williams. Semisynthetic resorbable materials from hyaluronan esterification. *Biomaterials*, 19(19):2101–2127, 1998.
- [35] G.D. Prestwich, D. M. Marecak, J.F. Marecak, K.P. Vercruyse, and M.R. Ziebell. Controlled chemical modification of hyaluronic acid. *J. Controlled Release*, 53(53):93–103, 1998.
- [36] S. Sadhasivam and K. S. Yun. DNA self-assembly: Prospectus and its future application. *Journal of Materials Science*, 45(10), 2010.
- [37] A. S. Hoffman. Hydrogels for biomedical applications. *Advanced drug delivery reviews*, 64:18–23, 2012.
- [38] N.A. Peppas, P. Bures, W. Leobandung, and H. Ichikawa. Hydrogels in pharmaceutical formulations. *European journal of pharmaceutics and biopharmaceutics*, 50(1):27–46, 2000.
- [39] Y. Qiu and K. Park. Environment-sensitive hydrogels for drug delivery. *Advanced drug delivery reviews*, 53(3):321–339, 2001.
- [40] K. Y. Lee and D. J. Mooney. Hydrogels for tissue engineering. *Chemical reviews*, 101(7):1869–1880, 2001.
- [41] Z. Xing, A. Caciagli, T. Cao, I. Stoev, M. Zupkauskas, T. O'Neill, T. Wenzel, R. Lamboll, D. Liu, and E. Eiser. Microrheology of DNA hydrogels. *Proceedings of the National Academy of Sciences*, 115(32):8137–8142, 2018.
- [42] D. P. Huynh, M. K. Nguyen, B. S. Pi, M. S. Kim, S. Y. Chae, K. C. Lee, B. S. Kim, S. W. Kim, and D. S. Lee. Functionalized injectable hydrogels for controlled insulin delivery. *Biomaterials*, 29(16):2527–2534, 2008.
- [43] Y. Wang et al. Constructing tissuelike complex structures using cell-laden DNA hydrogel bricks. *ACS Applied Materials & Interfaces*, 9(14):12311–12315, 2017.
- [44] M. Tibbitt and K. Anseth. Hydrogels as extracellular matrix mimics for 3D cell culture. *Biotechnology and bioengineering*, 103(4):655–663, 2009.
- [45] Lei Z. et al. Self-assembled DNA hydrogel as switchable material for aptamer-based fluorescent detection of protein. *Analytical chemistry*, 85(22):11077–11082, 2013.
- [46] D. Neeshma et al. Regenerable DNA-functionalized hydrogels for ultrasensitive, instrument-free mercury (ii) detection and removal in water. *Journal of the American Chemical Society*, 132(36):12668–12673, 2010.
- [47] D. Liu et al. Self-assembled DNA hydrogels with designable thermal and enzymatic responsiveness. *Advanced Materials*, 23(9):1117–1121, 2011.
- [48] Y. Huang et al. A pure DNA hydrogel with stable catalytic ability produced by one-step rolling circle amplification. *Chemical Communications*, 53(21):3038–3041, 2017.

- [49] X. Xiong et al. Responsive DNA-based hydrogels and their applications. *Macromolecular rapid communications*, 34(16):1271–1283, 2013.
- [50] G. Strobl. *The Physics of Polymers 2nd edition*, volume 1. Springer, 1997.
- [51] G. Mavko, T. Mukerji, and J. Dvorkin. *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media, Elasticity and Hooke's law*, page 21–80. 2 edition, 2009.
- [52] A. S. Khair and J. F. Brady. Single particle motion in colloidal dispersions: a simple model for active and nonlinear microrheology. *Journal of Fluid Mechanics*, 557:73–117, 2006.
- [53] I. Stoev, A. Caciagli, Z. Xing, and E. Eiser. Using single-beam optical tweezers for the passive microrheology of complex fluids. volume 10723. International Society for Optics and Photonics, 2018.
- [54] T. G. Mason. Estimating the viscoelastic moduli of complex fluids using the generalized Stokes–Einstein equation. *Rheologica Acta*, 39(4):371–378, 2000.
- [55] R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29:255–284, 1966.
- [56] S. R. Williams and D. J. Evans. Viscoelastic properties of crystals. *The Journal of chemical physics*, 131(2):024115, 2009.
- [57] S. R. Williams and D. J. Evans. Erratum: Viscoelastic properties of crystals. *The Journal of chemical physics*, 131(15), 2009.
- [58] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2 edition, 2004.
- [59] R. A. Leach. *Molecular Modelling, Principles and Applications*, volume 1. Pearson, Prentice Hall, 2001.
- [60] M. E. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.
- [61] W. Swope, H. Andersen, P. Berens, and K. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982.
- [62] W. Cornell et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.
- [63] A. MacKerell Jr et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*, 102(18):3586–3616, 1998.
- [64] H. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384–2393, 1980.

- [65] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics*, 81(1):511–519, 1984.
- [66] W. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.
- [67] H. Berendsen et al. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [68] T. Schneider and E. Stoll. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Physical Review B*, 17(3):1302, 1978.
- [69] G. Stokes. *On the effect of the internal friction of fluids on the motion of pendulums*, volume 9. Pitt Press Cambridge, 1851.
- [70] Z. Xing, C. Ness, D. Frenkel, and E. Eiser. Structural and linear elastic properties of DNA hydrogels by coarse-grained simulation. *Macromolecules*, 52(2):504–512, 2019.
- [71] J. Weeks, D. Chandler, and H. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *The Journal of chemical physics*, 54(12):5237–5247, 1971.
- [72] I. D. Stoev, T. Cao, A. Caciagli, J. Yu, C. Ness, R. Liu, R. Ghosh, T. O'Neill, D. Liu, and E. Eiser. On the role of flexibility in linker-mediated DNA hydrogels. *Soft Matter*, 2020.
- [73] Šulc P. et al. Sequence-dependent thermodynamics of a coarse-grained DNA model. *The Journal of Chemical Physics*, 137(13), 2012.
- [74] D. W. Ussery. Encyclopedia of life sciences, DNA structure: A-, B-, and Z-DNA helix families. 2002.
- [75] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.
- [76] M. Allen and D. Tildesley. *Computer simulation in chemical physics*, volume 397. Springer Science & Business Media, 2012.

Appendix A

Appendix

A.1 DNA conformations

DNA can be found in a variety of structural forms, namely A-DNA, B-DNA, and Z-DNA. Such variations of DNA are presented in Figure 1.2. B-DNA is a perfect linear double helical conformation that was first studied by Watson and Crick with X-ray diffraction. The B-DNA helix represents an ideal configuration hence it is unlikely to be found solely within the cell of living organisms, since DNA has to fold in order to fit inside [74]. The base sequences are responsible for the shape and geometry of the DNA helix, including helicity, the helix diameter, the rise per base pair, the number of base pairs per helix turn, the pitch of the helix, and the tilt of the base pairs from perpendicular to the helix axis. These characteristics vary in A-, B-, and Z-DNA, and they are mentioned in table A.1.

Table A.1

	A-DNA	B-DNA	Z-DNA
Shape	Broadest	Intermediate	Narrowest
Rise per base pair	2.3 Å	3.4 Å	3.8 Å
Helix diameter	26 Å	20 Å	18 Å
Helicity	right-handed	right-handed	left-handed
Base pairs per helix turn	11	10.4	12
Pitch per turn of helix	25.3 Å	35.4 Å	45.6 Å
Tilt of base pairs perpendicular to helix axis	19°	1°	9°

Table with a list of characteristics for the three DNA conformations gathered from [7, 74].

A.2 LAMMPS Technical Details

Large-scale Atomic/Molecular Massive Parallel Simulator (LAMMPS) is an open source software initially developed by Steve Plimpton [75], maintained by Sandia National Laboratories, and distributed under the GNU General Public License. It includes a number of algorithms, among which, there are a few that need to be mentioned. The ensemble used was the canonical (N,V,T) ensemble, hence the particles' velocities are updated in such a way so that they obey the Maxwell-Boltzmann distribution. In this way, the temperature is kept constant. In my calculations I used the standard velocity-Verlet algorithm for the integration of the equations of motion, shown below:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \vec{v}(t) \cdot \Delta t + \frac{1}{2} \vec{\alpha}(t) \cdot \Delta t^2 \quad (\text{A.1a})$$

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{\vec{\alpha}(t) + \vec{\alpha}(t + \Delta t)}{2} \cdot \Delta t \quad (\text{A.1b})$$

Moreover, I used the Langevin thermostat, that has a double role in the simulations [66]. Firstly it couples the system with an infinite heat reservoir, and all the particles' velocities are rescaled according to Langevin dynamics. Also, this thermostat accounts for an implicit solvent in the system, since it includes a random force where the random numbers are drawn from a uniform distribution.

A.3 LJ Reduced Units

In MD simulations, it is quite common that reduced units are used. The advantage of using reduced units is that the calculations are simplified, since there is no point in calculating analogous quantities multiple times [76]. Throughout this thesis, I made use of LJ reduced units, since the interatomic potentials that were used were of the form of Lennard-Jones potential. In LJ reduced units, the ϵ and σ parameters of the LJ potential are considered as fundamental quantities. Hence they are expressed in ϵ_{LJ} and σ_{LJ} units of energy and distance respectively. Moreover, the Boltzmann constant k_B and the particles' mass m are also fundamental quantities. Thus, all the quantities in this thesis are expressed with respect to ϵ_{LJ} , σ_{LJ} , k_B and m . For example, time is expressed in τ which is equal to $\sigma_{LJ}m^{1/2}\epsilon_{LJ}^{-1/2}$, velocity is expressed in σ_{LJ}/τ , force in $\epsilon_{LJ}/\sigma_{LJ}$, etc.

A.4 Melting Curves for Various Concentrations

Additional melting curves for various concentrations are presented in the Figures A.1.

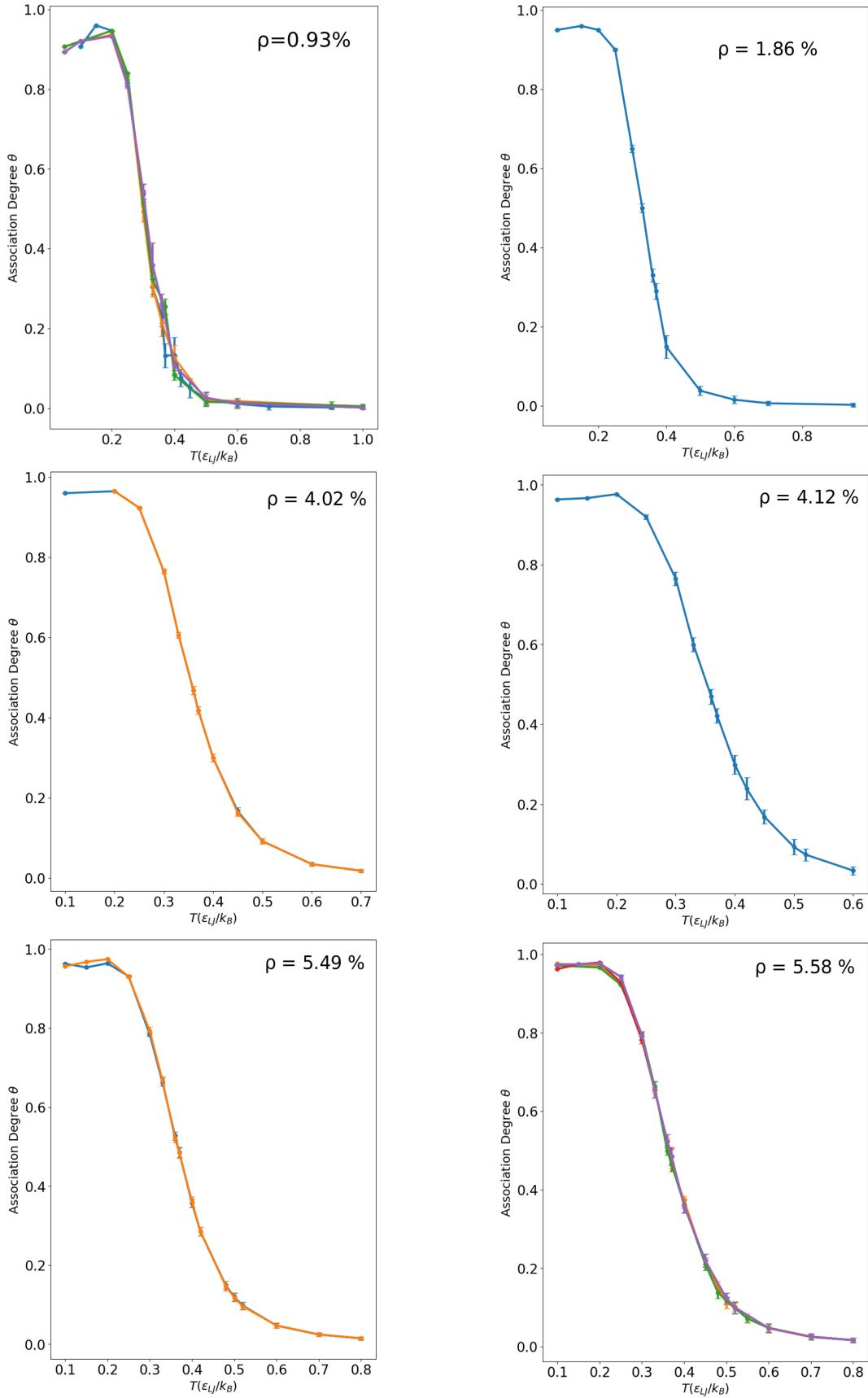


Figure A.1 Plots of the association degree for various concentrations with error bars. The concentrations are also mentioned in Table 4.1