Understanding the free energy landscapes and self-limiting assembly of DNA origamis using coarse-grained modelling



Chak Kui Wong

Somerville College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

Abstract

Coarse-grained modelling of DNA nanostructures allows their structural, mechanical and thermodynamic properties to be characterised. In this thesis, we illustrated how coarse-grained modelling combined with umbrella sampling can be applied to compute the free energy landscapes of large DNA nanostructures, and how these landscapes can be used to characterise the self-limiting assembly of DNA origamis.

We first described the mechanical deformation of DNA nanotubes by computing their free energy landscapes using distance-based order parameters. We analysed the behaviour of the nanotubes under different levels of bending stress and compared their behaviour to the worm-like chain model, revealing two regimes of bending depending on the localisation of the bending stress. We further demonstrated the computation of free energy landscapes on a mechanically bistable DNA origami. Interestingly, the landscape computed from simulations only shows one minimum, which is different from the bistable landscape inferred from experimental observations of ensembles of origamis. We decomposed the landscape into one component for each joint in the origami and discovered that only one such component showed bistability, which was masked by the other components and resulted in an overall landscape with a single minimum.

We also examined the self-limiting assembly of DNA origamis. We studied a DNA origami that assembles with stress propagation through displacement accumulation. We characterised the structural properties of the origami monomer and its polymers, as well as the free energy landscapes of the monomer associated with the main stress accumulation mechanisms. We parameterised a simple one-dimensional model using the landscapes and ran Monte Carlo simulations to estimate the stress free energy of polymers as a function of their lengths. As an extension, we explored other possible DNA origami designs that could undergo self-limiting assembly through mechanisms other than displacement accumulation. In particular, we evaluated the use of curvature and twist accumulation to propagate stress in assemblies of DNA origamis.

Acknowledgements

I am very grateful to my supervisor, Professor Jonathan Doye, for his guidance and support during my DPhil. He has also presented me with numerous collaboration opportunities with international researchers, which have sparked the idea of self-limiting assembly using DNA origamis and enabled the work in this thesis. I am also indebted to Professor Steve Tse for introducing me to the field of theoretical chemistry in my undergraduate years.

I am equally grateful to my funding providers, including the Croucher Foundation, Somerville College and the Clarendon Fund. The completion of my DPhil would not have been possible without their generous financial support.

I would also like to thank the providers and maintainers of computer clusters, in particular GPU clusters, on which the majority of the simulations in this thesis is performed. These providers include the University of Oxford Advanced Research Computing (ARC) facility and the Cambridge Service for Data Driven Discovery (CSD3).

Last but not least, I would like to thank the continued support of my family, friends and colleagues during the course of my DPhil.

Contents

List of Abbreviations **vi**

1 Introduction **1**
1.1 DNA nanotechnology 2
1.2 Computational methods in DNA nanotechnology 3
1.3 Thesis Outline 12

2 Simulation Methods **14**
2.1 Molecular Dynamics Simulations 14
2.2 Monte Carlo Simulations 17
2.3 Umbrella Sampling 17
2.4 Weighted Histogram Analysis Method 18

3 Characterising the Free Energy Landscapes of DNA Nanotubes **20**
3.1 Introduction 20
3.2 Nanotube Design 23
3.3 Results and discussion 24
3.4 Methods 40
3.5 Conclusions 46

4 Free Energy Landscape of a Mechanically Bistable DNA Origami **48**
4.1 Introduction 48
4.2 Origami design and structure 49
4.3 Free Energy Landscape 51
4.4 Free Energy Decomposition 52
4.5 Methods 58
4.6 Conclusions 62

5 Self-Limiting Polymerisation of a DNA Origami **66**
5.1 Introduction 66
5.2 Origami Design 69
5.3 Results and Discussion 71
5.4 Methods 94
5.5 Conclusions 100

6 Self-Limiting Assembly of Curved and Twisted DNA Origamis 103
6.1 Introduction . . . 103
6.2 Curved Monomer: Ideal Model . . . 106
6.3 Curved Monomer: Design 1 . . . 109
6.4 Curved Monomer: Design 2 . . . 125
6.5 Twisted Monomer . . . 133
6.6 Methods . . . 142
6.7 Conclusions . . . 147

7 Conclusions 149

Appendix

A Full caDNAno Designs of DNA Origamis 154
A.1 Chapter 3 . . . 155
A.2 Chapter 4 . . . 155
A.3 Chapter 5 . . . 157
A.4 Chapter 6 . . . 158

References 159

List of Abbreviations

bp . . . . . Base pairs.

COM . . . . . Center of mass.

DNA . . . . . Deoxyribonucleic acid.

GPU . . . . . Graphics processing unit.

MC . . . . . Monte Carlo.

MD . . . . . Molecular dynamics.

s.u. . . . . . Simulation units.

WHAM . . . . Weighted histogram analysis method.

WLC . . . . . Worm-like chain.

1 Introduction

Contents

1.1 DNA nanotechnology 2
1.2 Computational methods in DNA nanotechnology . . . 3
1.3 Thesis Outline 12

DNA is the molecule that carries the genetic information of organisms. It is a bio-polymer composed of monomers called nucleotides. Each nucleotide consists of a sugar, deoxyribose, which bridges a phosphate group and a base. The base can be one of the four bases, adenine (A), thymine (T), cytosine (C) and guanine (G). In a celebrated discovery, the double helical structure of DNA was determined by Watson and Crick in 1953 [1]. In the double helix, the two strands run in antiparallel directions. Complementary bases on opposite strands interact through hydrogen bonding, with A binding selectively to T, and C selectively to G. Bases on the same strand interact through π–π stacking interactions of their aromatic rings.

Although DNA is a biological molecule, it can be used as a material for nanotechnology. In 1982, Nadrian Seeman [2] proposed utilising the specificity of DNA base pairing to construct artificial DNA structures beyond linear DNA duplexes. This proposal marked the beginning of the field of DNA nanotechnology.

1.1 DNA nanotechnology

Seeman's proposal [2] was to create a three-dimensional crystalline network through the self-assembly of DNA crossover junctions. Sixteen years after his proposal, his group achieved the self-assembly of two-dimensional DNA crystals experimentally [3]. The long period between the proposal and its realisation perhaps suggests the challenges associated with engineering DNA nanostructures, such as the need for high purity of strands, sequence optimisation to avoid unwanted interactions, and so on [4]. These challenges become even more important when engineering larger DNA nanostructures.

Perhaps the most significant breakthrough in the field came in 2006, when Rothemund [4] demonstrated the powerful method of scaffolded self-assembly of DNA, which provides a robust way for synthesising large DNA nanostructures consisting of thousands of base pairs. He coined the term "DNA origami" for the method. The assembly of a DNA origami involves a long single-stranded DNA (called scaffold) and hundreds of short single-stranded DNA (called staples) which selectively form base pairs to multiple distant sites on the scaffold in order to fold it into the desired geometry. The scaffold is typically a viral DNA 7000–8000 nucleotides long, while the staples are typically synthetic oligonucleotides 20–50 nucleotides long. The method gives a generally good yield of structures, typically at least 70% and can be as high as 90% [4].

Rothemund's method is, however, limited to two dimensions when his work was published, but in 2009, the method was extended to produce DNA origamis in three dimensions on a honeycomb lattice [5] or a square lattice [6], and even origamis that are curved or twisted [7, 8]. In 2015, Benson *et al.* [9] further extended the method to allow lattice-free origami designs, allowing the construction of a broad range of DNA nanostructures of arbitrary shapes. These structures are often called wireframe DNA origamis. The Bathe group generalised the concept of wireframe origamis in a series of four papers [10–13], which allows design in two or three dimensions and the use of two or six double helices as the wireframe edges for tuning flexibility.

An alternative approach to assembling large-scale DNA nanostructures is to use short DNA single strands only, without a scaffold strand. This approach is often known as the assembly of "DNA tiles" or "DNA bricks", and has been used to construct 2D [14] and 3D [15] objects. Using this scaffold-less approach, the size of the assembled structure will not be limited by the size of the scaffold, although the assembly yield could become a limiting factor; nevertheless, a 3D assembly with $10,000$ unique components has been reported [16], resulting in gigadalton-scale assemblies.

DNA nanotechnology and DNA origami have a broad range of applications, but they are not going to be discussed here as they are not the focus of the thesis. In-depth reviews can be found in [17] and [18].

1.2 Computational methods in DNA nanotechnology

While experiments are indispensable to DNA nanotechnology research, computational methods are also important and can be complementary to experiments. For example, many DNA origamis are designed using computer-aided design software, and their structures can be predicted using computer modelling. The ability of computer modelling to provide detailed structural characterisation makes it a particularly valuable tool for interpreting experimental results, or even for guiding experiments.

1.2.1 Computer-aided design

Designing a complete DNA origami structure can be a somewhat tedious process if done manually. Several software packages are available to facilitate this process. These software will be briefly reviewed in this section. An in-depth review can be found in [19].

The most popular software for designing DNA origamis is probably caDNAno [20]. The software allows the design of 3D DNA origamis based on a square lattice or a honeycomb lattice. If the user provides a scaffold strand routing, the software

can automatically assign complementary staple strands and add nicks as appropriate so that the staple strands are short enough to synthesise experimentally.

caDNAno can be considered as a bottom-up tool where the user builds origami designs by routing the scaffold on a lattice. However, it might sometimes be useful to build designs in a top-down manner, where the user specifies the shape of the target structure and the software solves the scaffold routing problem. It might also be useful to design origamis with geometries that are not restricted on lattices. There are some other software which facilitates the design of these types of DNA origamis. One example is vHelix [9], which creates polygonal meshes composed of DNA helices for arbitrary 3D structures. The software is able to generate the scaffold routing for the meshes using graph theory algorithms. The corresponding staples and their sequences can thus be generated, which allows experimental synthesis of the designed structures. This tool opens up opportunities for designing DNA origamis beyond the limitation of a lattice.

The Bathe group has developed another algorithm based on a spanning tree to solve the scaffold routing problem for a polyhedral mesh [10]. They further extended the algorithm for it to work with both 2D and 3D structures, and to allow the use of six-helix bundles rather than double helices as the edges of the mesh in order to increase structural rigidity [11–13]. The algorithms have been combined into the ATHENA software package [21].

Recently, Huang *et al.* [22] developed magicDNA, which, like ATHENA, implements a top-down approach to design 2D and 3D structures, while also providing a caDNAno-like interface to allow users to fine-tune the structures. On top of these two features, users are also able to design origamis that consist of multiple components, which is an advance from caDNAno and ATHENA. The authors demonstrated the design, simulation and synthesis of a number of multi-component complex structures, with some of them consisting of multiple scaffolds, showcasing the ability of this software to design complex DNA origamis which are difficult to design using other existing software.

1.2.2 Computer modelling

All-atom models

To characterise DNA nanostructures at an atomic level of detail, all-atom molecular dynamics simulations can be performed using force fields such as CHARMM [23] or AMBER [24]. The Aksimentiev group has applied all-atom simulations to study several DNA nanostructures, including the structural properties of DNA origamis [25, 26] and DNA bricks [27], as well as the ionic conductivity of DNA origamis [28, 29]. It is worth noting that these simulations tend to be computationally expensive because of the large number of atoms in these systems, which is typically on the order of million due to the large number of water molecules required to solvate the origami. Due to the high computational cost, simulation time scales are limited to less than 1 µs. System equilibration could be accelerated by excluding water molecules and ions in the earlier stages of equilibration, and applying a elastic network to restrain the overall structure of the origami [26], before the structure is further equilibrated under explicit solvent representation.

Finite-element models

Finite-element models represent DNA helices as elastic rods, and DNA origamis as a collection of such rods constrained to their nearest neighbours at the crossover positions. One of the most commonly used model is CanDo [30–32], which was parametrised based on the elastic properties of duplex DNA, such as its bending, twisting and stretching stiffness, as well as the properties of backbone nicks and single-stranded DNA, which are relevant to DNA origamis.

As finite-element models represent DNA with a relatively low level of detail, they are typically efficient and allow the prediction of the equilibrium structure of DNA origamis within minutes, with the major trade-off being the lack of structural details, as well as the lack of excluded volume interactions which may be important in flexible origamis.

Coarse-grained models

A number of coarse-grained models of DNA have been developed, which provide an improved simulation efficiency compared to all-atom models, at the expense of the level of detail. In coarse-grained models, atoms are grouped into beads that interact through effective potentials, and an implicit solvent representation is often used. The interaction potential of coarse-grained models are typically parametrised either through a "top-down" approach or a "bottom-up" approach. In a "top-down" approach, the model is fitted to experimental data, such as DNA melting temperatures and mechanical properties, whereas in a "bottom-up" approach, the model is fitted to simulation data obtained from finer models, usually all-atom models.

A selection of coarse-grained models will be briefly reviewed in this section. In-depth reviews of coarse-grained modelling of DNA can be found in [33–36], and reviews of coarse-grained modelling of biomolecules in general can be found in [37, 38].

Among the simplest coarse-grained models are those which represent each nucleotide as a single bead. Savelyev and Papoian [39] developed a model aiming to describe local motions and global dynamics of duplex DNA. The model was parametrised using a bottom-up approach. Each nucleotide is represented as a single bead and ions are explicitly represented. The model is able to achieve a quantitative agreement with experimental data on the relationship between the persistence length of DNA and the ionic strength. The model also predicts a buckling transition of a DNA mini-circle as the ionic strength increases beyond a threshold value. Naômé *et al.* [40] and Yagyu *et al.* [41] also developed single bead coarse-grained models using bottom-up approaches with different forms of the potential. These models are also able to capture the dependence of the persistence length of duplex DNA. It is perhaps worth noting that these single-bead models can only capture the behaviour of duplex DNA, but not DNA hybridisation or single-stranded DNA, due to the nature of the potential form.

To achieve a better description of DNA, it is necessary to go beyond representing each nucleotide as a single bead. A number of such coarse-grained models have

been developed; here we will focus on a number of models, including 3SPN [42, 43], MARTINI [44, 45] and SIRAH [46]. We will dedicate Section 1.2.3 to describe and review the oxDNA model [47–49], which is arguably the most popular coarse-grained model for DNA nanotechnology.

The 3SPN model [42] developed in the de Pablo group is a more extensively studied coarse-grained model. It was parametrised using a top-down approach. Each nucleotide is represented by three beads, with one bead for the backbone, the sugar and the base, respectively. Originally published in 2007 [42], the model was gradually improved over time to improve its description of various properties of DNA, including its structural, thermodynamic and mechanical properties [43], as well as their sequence dependence [50]. The model achieves quantitative agreement with experimental DNA melting temperatures, captures the persistence length of single-stranded and double-stranded DNA, and describes the sequence-dependent flexibility of DNA. The model has been used for studying DNA hybridisation [51] and DNA wrapping around a nucleosome [52, 53]. It is worth noting that the model has been integrated with a coarse-grained model for proteins to allow studies on DNA-protein interactions [54–56]. The model, however, has mostly been applied in biophysical contexts rather than DNA nanotechnology.

The Marrink group originally developed the MARTINI model as a coarse-grained model for lipids [44], and later extended it to other biomolecules such as proteins, carbohydrates and DNA [45, 57]. The model was parametrised using a combined bottom-up and top-down approach. Each nucleotide is represented using six to seven beads, with one bead for the backbone phosphate, two beads for the backbone sugar, and three or four beads for a pyrimidine or purine base respectively. An advantage of the model is its ability to describe the interactions of DNA with other biomolecules. For example, the model was used to study the stability and dynamics of a DNA nanotube across a lipid bilayer membrane [58]. However, an elastic network is needed to maintain the structure of duplex DNA and its persistence length, implying that it might be less suitable for studying structural properties of DNA [45].

The Pantano group developed the SIRAH model [46] using a top-down approach. The representation of nucleotides is similar to the MARTINI model, with one bead for the backbone phosphate, two beads for the backbone sugar, and three beads for the base. The model was later extended to model DNA in a hybrid all-atom/coarse-grained simulations [59], which allows an enhanced efficiency due to coarse-graining while retaining the details provide by all-atom models in interesting regions of the simulated system. The model was applied in a study of DNA-protein interaction [60].

Domain-level models

To study the assembly of DNA origamis, even coarse-grained models may not be efficient enough due to the long time scales involved. Finite-element models are unsuitable for these studies because they are unable to represent partially-formed DNA origamis with many flexible single-stranded DNA components. Dannenberg *et al.* [61, 62] developed a domain-level model to address the challenges associated with modelling DNA origami assembly, where the fundamental unit of the model is a "domain", defined as a contiguous series of bases that selectively binds to a complementary domain. With hundreds of domains on a typical scaffold strand and one to five domains on a typical staple strand, the model can comfortably simulate the assembly of a DNA origami. The model allows the rational design of the folding pathway and target structure for a DNA origami, with the results being verified experimentally [62].

Cumberworth *et al.* [63] also developed a model to simulate the self-assembly of DNA origamis. The model represents an origami on a lattice at the level of binding domains, and includes an explicit geometric consideration of the helical twist which determines the possible positions of strand crossovers. The model has been applied to study the nucleation barrier in origami assembly [64].

Multi-scale models

The efficiency of coarser models, such as finite-element models, can be combined with the details provided by finer models, such as all-atom models, through multi-scale modelling. DNA structures are first equilibrated using coarser models to

bring their geometries closer to equilibrium, before they are simulated using finer models to provide a more accurate representation of the structures near equilibrium. Two multi-scale models have been developed recently, including mrDNA [65] and SNUPI [66]. Although they have not been widely used so far, the high efficiency and accuracy of these models provide a promising new pathway for computationally designing and optimising DNA origami designs.

1.2.3 The oxDNA model

oxDNA is a nucleotide-level coarse-grained model of DNA. Each nucleotide is represented as a rigid body that consists of three interaction sites, corresponding to backbone, stacking and hydrogen bonding interactions, as shown in Figure 1.1a. The model was parametrised in a "top-down" approach, with a focus on reproducing experimental results on the structural, thermodynamic and mechanical properties of DNA [47, 48], such as the structure of B-DNA, duplex and hairpin formation thermodynamics [67], stacking thermodynamics [68] and persistence length [69]. A more recent version of the model, oxDNA2 [49], was fine-tuned to improve the representation of large DNA nanostructures, particularly DNA origamis which have been studied and characterised in experiments [7]. Other features were also introduced, such as different widths for the major and minor grooves, electrostatic interactions and salt concentration dependence. A GPU algorithm [70] has been implemented for the model to speed up simulations for large structures. The model has also been ported to the LAMMPS simulation code [71], and a software called tacoxDNA [72] has been developed for converting between oxDNA and LAMMPS file formats, as well as for converting common DNA origami design formats to oxDNA file formats. A dedicated visualisation software called oxView [73] has also been developed.

The potential energy function of oxDNA2 can be written as the sum of pairwise

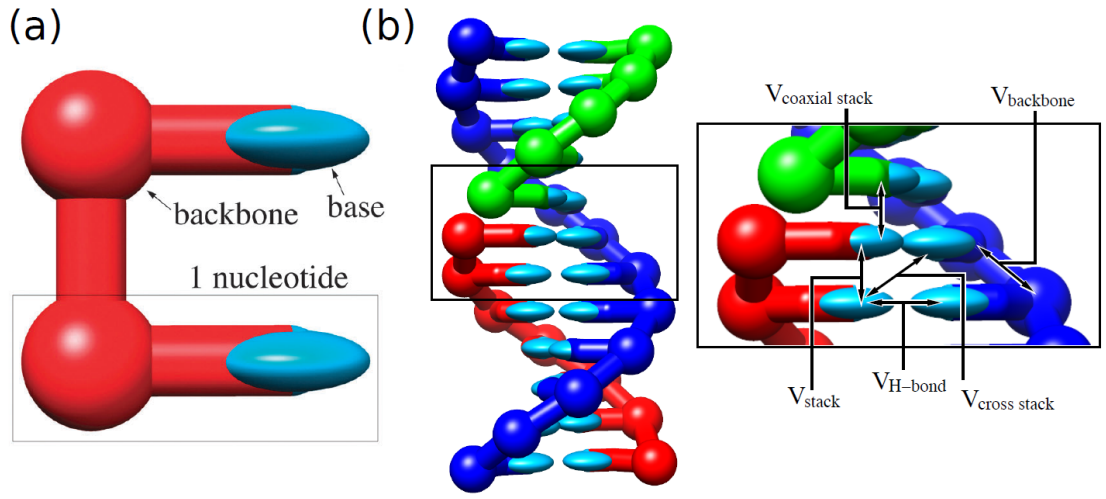


Figure 1.1: (a) A single strand with two nucleotides and (b) a nicked duplex represented in the oxDNA model. Different types of interactions in the model are shown in (b). Figure reproduced from [74] with permission.

interaction terms between nucleotides:

$$V_{\text{oxDNA2}} = \sum_{\text{nearest neighbours}} (V_{\text{backbone}} + V_{\text{stack}} + V_{\text{exc}}) + \sum_{\text{other pairs}} (V_{\text{H-bond}} + V_{\text{cross stack}} + V_{\text{exc}} + V_{\text{coaxial stack}} + V_{\text{DH}})$$

where the first term is the sum of the backbone connectivity, stacking and excluded volume interactions over all pairs of nucleotides which are nearest neighbours, and the second term is the sum of the hydrogen bonding, cross stacking, excluded volume, coaxial stacking, and Debye-Hückel electrostatic interactions over all other pairs of nucleotides. A graphical illustration of the interactions is shown in Figure 1.1b. The detailed mathematical form of each potential energy term can be found in [47] and [49].

Since oxDNA2 was introduced in 2015, it has been used in a variety of studies involving DNA origamis. Snodin *et al.* [75] performed an in-depth structural analysis of the basic features of DNA origamis in oxDNA and compared the structures predicted by oxDNA to one of the most structurally detailed experimental data obtained by cryo-EM. A 2D origami sheet was simulated, which was previously characterised using SAXS [76]. The modest curvature of the simulated sheet is consistent with the results from the SAXS experiment. In addition, splaying

of duplexes between double crossover junctions was observed in the simulated structure, which can be clearly seen in microscopy images of 2D DNA origami [4]. A 3D asymmetric origami object was also simulated, which was previously characterised in detail using cryo-EM [77]. The root-mean-square deviation of the simulated structure to the cryo-EM structure is 8.4 Å, which is of the order of the experimental resolution. The good agreement between the simulated and experimental origami structures shows that oxDNA is able to accurately describe the structural properties of DNA origamis.

Apart from rigid DNA origamis, oxDNA is useful for studying origamis that consist of flexible components. These origamis often contain single-stranded DNA to introduce flexibility, but other models do not always include single-stranded DNA in their parametrisation. Given that stacking in single-stranded DNA was one of the factors considered when parametrising oxDNA, and that oxDNA was successful in describing the properties of single-stranded DNA, such as its stacking transition [47] and force-extension behaviour [48], oxDNA is suitable for describing flexible DNA origamis containing single-stranded sections.

Benson *et al.* [78, 79] used oxDNA to study and optimise the design of wireframe DNA origami structures. Wireframe origamis contain junctions where multiple helices meet, and single-stranded DNA is often found in the design of these junctions to allow flexibility at the turning point. The ability of oxDNA to describe single-stranded DNA makes it suitable for simulating these wireframe origamis. Furthermore, the explicit representation of individual nucleotides allow more subtle effects to be investigated, such as the effect of nicks on the stiffness of the origami edges [78].

Sharma *et al.* [80] used oxDNA to characterise jointed DNA nanostructures, which were first studied experimentally by the Castro group [81]. The probability distributions of certain structural properties of the origamis were calculated. With the ability to capture the behaviour of single-stranded DNA that composes the joints, such as its force-extension behaviour and its excluded volume, the simulation results generally agree with the experimental results. The main difference is the

more limited range of motion observed in simulations compared to the ensemble of origamis in experiments. This difference was attributed to heterogeneity within the ensemble, which potentially arises from assembly defects. Shi *et al.* [82] used oxDNA to characterise another DNA nanostructure designed by the Castro group [83]. A good agreement in the structural properties was also achieved, although once again the simulated structures tend to show a more limited range of motion compared to the ensemble of structures in experiments.

oxDNA is also suitable for studying origami under external stress, which might undergo structural transitions involving base pair breaking. Such transitions could not be described by models that do not explicitly represent nucleotides or base pairs, for example, models that consider double helices as rigid rods. In a joint experimental and simulation study, Engel *et al.* [84] used AFM single molecule force spectroscopy together with oxDNA modelling to investigate the behaviour of DNA origamis under external forces. A monotonically increasing force was applied to a 2D DNA origami tile and its force-induced unravelling was analysed. The force-extension curves obtained from simulations generally agreed with experiments, which both showed a regular sawtooth pattern of unfolding and a similar level of yielding force. oxDNA is particularly suitable for this study because of its ability to describe the response of duplex DNA under tension [85, 86].

Given its success in describing a wide variety of DNA nanostructures, we will further explore the capabilities of oxDNA in characterising DNA origamis in this thesis.

1.3 Thesis Outline

In this thesis, we will use the oxDNA coarse-grained model to characterise various DNA origamis. We will focus on using umbrella sampling to calculate their free energy landscapes. As an application, we will use the landscapes to describe the self-limiting assembly of DNA origamis.

We first briefly review the simulation methods used for this thesis in Chapter 2. In Chapter 3, we characterise the structural properties of DNA origami nanotubes through their free energy landscapes. In Chapter 4, we describe in detail the

free energy landscape of a mechanically bistable DNA origami previously studied experimentally. In Chapter 5, we characterise the structural properties of a DNA origami that undergoes self-limiting polymerisation through stress accumulation. We calculate its free energy landscapes and parametrise a simple model to understand the mechanism of its self-limiting assembly. In Chapter 6, we design and characterise several DNA origamis which could undergo self-limiting assembly using other stress accumulation mechanisms, such as curvature and twist accumulation. Finally, in Chapter 7, we summarise the findings of the thesis.

2
Simulation Methods

Contents

2.1 Molecular Dynamics Simulations 14
2.2 Monte Carlo Simulations 17
2.3 Umbrella Sampling 17
2.4 Weighted Histogram Analysis Method 18

A central goal of molecular simulations is to sample the phase space of the system and thereby calculate its equilibrium and/or dynamical properties. Two common methods of performing simulations, namely Molecular Dynamics (MD) and Monte Carlo (MC), will be discussed in this chapter. Umbrella sampling, an enhanced sampling protocol that allows a more efficient exploration of the phase space, will also be discussed in this chapter, together with the complementary Weighted Histogram Analysis Method.

2.1 Molecular Dynamics Simulations

In MD simulations, the Newtonian equations of motion are numerically solved to update the position and velocity of each particle in the system. During this process, first, the forces acting on each particle are calculated from the gradient of the potential energy function with respect to the particle positions in the current

time step. Then, from the forces and particle positions, the equations of motion are numerically integrated to give the particle positions in the next time step. By repeating the integration many times, a trajectory of the particles can be generated. A commonly used integration algorithm is called the velocity Verlet algorithm [87], which is also employed in the oxDNA simulation software package. This efficient algorithm provides good energy conservation over long simulation time scales, as long as a sensible time step is chosen [88].

An energy-conserving integration algorithm, like the velocity Verlet algorithm, naturally samples the phase space of the simulated system in the microcanonical ensemble, where the number of particles, volume and energy of the system are constant. However, to compare simulations to real-world experiments, we are more interested in the behaviour of the system at a constant temperature, which is why simulations are also commonly performed in the canonical ensemble, where the temperature is constant instead of the energy.

In the real world, experiments are performed at constant temperature by surrounding the system in a large reservoir of constant temperature, such as a water bath. A similar approach could be employed in simulations by surrounding the simulation system with a large reservoir of molecules that act as a thermal bath, but it would be extremely time-consuming because most of the computational time would be spent in simulating the thermal bath, the properties of which we are not interested in. A more efficient solution is to introduce a thermostat which couples the simulation system to a virtual thermal bath. The thermostat adds perturbations to the dynamics of the particles, which mimics the effect of a thermal bath.

The Langevin thermostat is used for the simulations in this study. This thermostat adds two forces to each particle in the system. The first force is a frictional force, which is proportional to the velocity of the particle; the second force is a random force [89]. Together, these two forces give rise to the Brownian motion of the particles. The strength of the random force and the frictional force are connected to the temperature through the fluctuation-dissipation theorem [89]. This relation ensures that the system evolves in the canonical ensemble. For simulations

of coarse-grained models such as oxDNA, it is often beneficial to accelerate the system dynamics by increasing the effective diffusion coefficient of the particles. This is achieved by using a smaller damping coefficient in the frictional force of the thermostat and adjusting the random force accordingly.

We used the oxDNA coarse-grained model to perform MD simulations. Simulation units were defined to facilitate calculations on the molecular scale. The conversion between simulation units and SI units is described in Table 2.1. Unless otherwise specified, simulations were performed at 300 K using a Langevin thermostat with the monomer diffusion coefficient set to 2.5 simulation units. The time-step is 0.005 simulation units, corresponding to 15.2 fs. The solvent environment is treated implicitly as a dielectric continuum. A salt concentration of $[Na^+] = 1.0$ M is used, which represents the high salt conditions typically used for DNA nanotechnology. Due to large system sizes, we use the GPU implementation of oxDNA [70], which can achieve a significant speed-up compared to running on a CPU.

Type	Simulation unit	SI unit
Length	1 s.u.	0.8518 nm
Energy	1 s.u.	4.142×10^{-20} J 41.42 pN nm (thermal energy at 3000 K)
Force	1 s.u.	48.63 pN
Time	1 s.u.	3.03 ps
Force constant	1 s.u.	57.09 pN/nm
Torque	1 s.u.	41.42 pN nm

Table 2.1: Conversion between simulation units and SI units

Care must be taken when generating starting configurations for the simulations. Usually, the starting configurations of DNA origamis are generated by converting their caDNAno design file into oxDNA format using the tacoxDNA package [72]. However, the converted configurations cannot serve as starting configurations because of nucleotides experiencing large forces due to overlaps or extended bonds. As such, the energy of these configurations are first minimised for 200 steps using a steepest-descent algorithm, and then the configurations are relaxed using a modified backbone potential for 10^6 steps. After that, the extended bonds have typically

returned to their normal lengths, and the configurations are ready for simulation under the normal force field.

2.2 Monte Carlo Simulations

In MC simulations, instead of updating the position of the particles in a deterministic way, random particle moves are proposed and accepted according to some criteria such that the moves allow the system to efficiently sample from a particular ensemble.

A common algorithm for sampling from the canonical ensemble is called the Metropolis method [90]. The most basic form of the algorithm involves selecting a random particle, giving it a random displacement, and calculating the resultant change in potential energy ΔU. If $\Delta U < 0$, the move is accepted. Otherwise, the move is accepted with a probability of $\exp(-\Delta U/k_B T)$. If the move is rejected, the particle is not moved and returns to its original position. This process is repeated during the simulation.

MC simulations are naturally serial because a particle move is proposed after the previous move has completed. Despite there being some proposals on the parallelisation of MC simulations [91], their degree of parallelisation is still far behind that of MD simulations. As MC is only implemented serially in oxDNA, it is extremely inefficient to perform MC simulations on large structures such as DNA origamis. Therefore, all large DNA structures are simulated using MD, where we can take advantage of the parallelisation available on modern GPUs [70].

Nevertheless, MC simulations do have their advantages. They are good for potentials with discontinuous values or derivatives, and are easy to programme for small systems with custom potentials.

2.3 Umbrella Sampling

Although simulations can describe the behaviour of a system near equilibrium, it is often interesting to study the behaviour of the system at states beyond equilibrium with a high free energy, or states separated with a high free-energy barrier. A way

to study these states is by brute-force sampling, which involves running very long simulations in order to sample these unlikely states. Another possible way is to use enhanced sampling techniques which enable a more efficient sampling of states with a high free energy. A common technique is umbrella sampling [92].

To use umbrella sampling, the pathway that includes all the states of interest has to be defined in terms of an order parameter ξ. For example, it could be a distance between two groups of particles. A bias is applied in the simulation as a function of the order parameter. The bias enhances the sampling probability of conformations with a high free energy, thus reducing the free energy barrier along the pathway and allowing for effective sampling.

When the pathway is too long to feasibly sample in a single simulation, the pathway can be divided into a number of simulation windows. Note that there should be sufficient overlap between the windows in order to facilitate an accurate estimation of the free energy landscape along the whole pathway. In each window, the system is restrained by an external bias potential so that the system extensively samples the region of the pathway within that particular window. A harmonic bias potential, in the form of $V_{bias}(\xi) = \frac{1}{2}k(\xi - \xi_0)^2$, where ξ_0 is the value of the order parameter aimed to be sampled in the window, is often used because of its simplicity and ease of implementation. During the simulation, the order parameter is calculated and outputted to a file. This gives a distribution of the order parameter in that window, although the distribution is biased because of the external potential. By running a simulation for each window, we obtain a set of biased distributions of the order parameter along the pathway.

In order to combine the biased distributions of all the windows into a single unbiased distribution of the order parameter, we can use the Weighted Histogram Analysis Method (WHAM) [93].

2.4 Weighted Histogram Analysis Method

The following WHAM equations are adapted from Kumar *et al.* [93].

The biased probability distributions are first binned into a histogram with N_{bin} bins. For N_{sim} windows simulated at temperature T, the optimal estimate of the unbiased probability distribution of the order parameter is given by

$$P_i = \frac{\sum_{k=1}^{N_{\text{sim}}} N_{i,k} \exp\left(-\beta V_{i,k}\right)}{\sum_{k=1}^{N_{\text{sim}}} n_k \exp\left(\beta(F_k - V_{i,k})\right)}$$

and

$$\exp\left(-\beta F_k\right) = \sum_{i=1}^{N_{\text{bins}}} P_i$$

where $N_{i,k}$ is the value of the histogram in the i^{th} bin of the k^{th} window, $V_{i,k}$ is the value of the bias potential in the i^{th} bin of the k^{th} window, n_k is the number of samples taken from the k^{th} window, F_k is the unbiased free energy of the system in the k^{th} window, and $\beta = 1/k_B T$. To obtain P_i and F_k, this set of equations is iteratively solved until convergence to a chosen error tolerance.

3

Characterising the Free Energy Landscapes of DNA Nanotubes

Contents

3.1 Introduction **20**
3.2 Nanotube Design **23**
3.3 Results and discussion **24**
3.4 Methods **40**
3.5 Conclusions **46**

The results of this chapter have been published in Nanoscale [94].

3.1 Introduction

Most studies involving DNA nanostructures focus on their structural characterisation, which is a reasonable starting point because the structural properties underpin the functional properties of the assembly. The most detailed experimental characterisation of DNA origami structures can be obtained from cryo-EM studies [77, 95], while lower-resolution imaging is available from TEM and AFM. Alternatively, computer modelling can be used to gain structural insights into DNA nanostructures. All-atom molecular dynamics simulations can provide the highest level of detail [25–29, 96], but are computationally very expensive because of the

large size of DNA origamis. On the other hand, coarse-grained models such as oxDNA [47–49], mrDNA [65] and CanDo [30, 31] provide a lower level of detail, but are computationally much less expensive and so can be routinely used to gain insight into origami structure [75, 97].

In addition to structural properties, mechanical properties of DNA nanostructures are also of interest because of potential applications where forces play an important role [98, 99]. Experimental studies on the mechanical properties of DNA nanostructures have mostly focused on extracting the persistence lengths of rod-like structures, particularly DNA nanotubes. Various types of DNA nanotubes have been studied, including those assembled as a conventional DNA origami [100–104] or a wireframe DNA origami [78], and those assembled from single-stranded tiles [105], double-crossover tiles [106, 107] or other larger tiles [108]. Several methods have been used to obtain the persistence length. The most common method is to directly analyse the configuration of nanostructures adsorbed on a surface. The analysis typically involves constructing a probability distribution of their end-to-end distances [102–104, 107, 108] or calculating the correlation between tangent vectors along the structure [78, 100, 105]. More elaborate methods have also been employed, such as inferring the persistence length from the cyclisation rate of the nanostructures [106], or using magnetic tweezers to study how the nanostructures respond to external forces [101]. The obtained persistence lengths are generally on the order of microns, which is much larger than that for double-stranded DNA (about 50 nm) [109].

There are also theoretical studies on the mechanical properties of DNA nanostructures. A previous study has computed the mechanical properties of DNA nanotubes [110], in particular, their persistence lengths, bending moduli and twisting moduli. Long equilibrium MD simulations of the nanotubes were performed using the oxDNA coarse-grained model to sample structural fluctuations at long length scales. From these fluctuations, the mechanical properties were extracted using an elastic model.

Although experiments and simulations have been used to characterise the equilibrium structures and fluctuations of DNA nanostructures, the free energy landscapes associated with more extreme fluctuations are also potentially important.

For example, such landscapes can describe the response of DNA origami to stress, and so they are particularly relevant to the field of DNA mechanotechnology [98] where DNA nanodevices are used to sense [111], apply [112] or transduce [113] forces. They are also relevant to origamis where internal stress is used to modulate the structure [7, 83, 100]. Additionally, for origamis that are designed to assemble into higher-order structures, the free energy landscape of the monomer could help predict the thermodynamics of stress accumulation, particularly for designs where the stress accumulation is a feature to make the assembly self-limiting [114, 115].

One of the most common methods for computing free energy landscapes is umbrella sampling [92], where a bias is applied during the simulations to allow sampling of high free energy conformations. For small DNA systems, it has been used to study processes such as stacking [116, 117], condensation [118, 119], and unwinding [120] in all-atom detail. Coarse-grained models allow this approach to be applied to larger systems and slower processes. For example, the oxDNA model has been used to study the free energy landscapes for a wide variety of processes involving assembly and/or mechanical deformation both for basic DNA motifs [85, 86, 121] and for small DNA nanodevices [122–124].

Very recently, the oxDNA model has also been used to compute the free energy landscapes of DNA origamis using metadynamics [125]. This method allows an automated exploration of free energy landscapes without the need for setting up multiple sampling windows and manual tuning of the bias potential. The landscape of a DNA origami and a bistable Holliday junction was computed successfully using this method.

Here, we apply umbrella sampling to sample the free energy landscapes of DNA origamis. Such calculations, although computationally expensive, are feasible both because we are using a coarse-grained model (oxDNA), which allows simulations to be performed efficiently at a reasonable computational cost, and because we are considering the free energy landscapes for structural variables, where equilibration mainly involves relatively fast processes associated with mechanical relaxation, rather than slow processes such as hybridisation.

We will consider the bending and stretching of DNA nanotubes [126] under external stress. One might expect that the behaviour will be somewhat analogous to that for duplex DNA. When weakly bent, duplex DNA can be described with the worm-like chain (WLC) model [127], but under strong bending, its behaviour deviates from the model due to localised buckling to form a kink that typically involves broken base pairs [128–130]. Previous analysis of such nanotubes have shown that the WLC model provides a good description of their thermal fluctuations [110], but it will be interesting to explore the degree of bending that can be achieved before kink formation, the structure of the kinks that form and whether such buckling leads to irreversible damage to the nanotubes.

We will use the oxDNA coarse-grained model [47–49] to compute the free energy landscapes. As described in Chapter 1, oxDNA has been designed to reproduce the structure, thermodynamics and mechanics of double- and single-stranded DNA. In particular, it is able to capture the response of double-stranded DNA to bending stress [121, 130] and the fraying or breaking of DNA base-pairs under stress [86], both of which are relevant to the current study of DNA nanotubes under external stress. oxDNA is also able to accurately model the structural properties of DNA origami [49, 75], and so it has been applied to study a wide range of different types of DNA origami [78, 80, 82, 97, 114]. Given the above, we are confident in using oxDNA to calculate the free energy landscapes for our chosen DNA origamis.

3.2 Nanotube Design

Six-helix bundles have been well characterised [95, 101, 105, 131] and used for applications such as chiral liquid-crystal mesogens [132] and stiff handles for force spectroscopy [102]. Here we study three designs of six-helix bundle nanotubes with a hexagonal cross-section. Design 1 is actually not an origami, but a single-stranded tile structure that is built from just six repeating 42-bp strands. Its mechanical properties have been experimentally probed [105]. The fact that all the junctions are single crossovers potentially allows significantly more freedom of motion at the junctions in response to stress. Design 2 is a DNA origami from Siavashpouri *et*

al. [132] where its liquid crystal behaviour has been studied. It contains mostly double crossovers. Notably, the scaffold crossovers are located very close to staple crossovers, with small 2-bp domains separating them; the dissociation of these 2-bp sections can lead to significantly increased local flexibility. Design 3 was constructed as a contrast to design 2 with the aim of increasing the mechanical strength. In this design, all crossovers are mediated by staples and all are double crossovers. All nicks are placed midway between junctions, which means the shortest domain is 7 bp. This arrangement leads to planes where every helix has either a junction or a nick. These planes are likely to be preferred sites for local bending. A part of each design is shown in Figure 3.1 and the full designs can be found in the Appendix.

3.3 Results and discussion

Bending

We use the end-to-end distance R_{ee} of the nanotubes (Figure 3.2(a)) to describe the degree of bending. To sample extremely bent states, we use umbrella sampling and apply a bias on R_{ee}. We run multiple simulations where each one samples a different window of R_{ee} using a harmonic "umbrella" potential that restrains R_{ee} to be near the minimum of that potential. The probability distribution $p(R_{ee})$ from the windows can be used to reconstruct the free energy landscape using WHAM [93].

The resulting free energy landscape for the Design 1 nanotube is shown in Figure 3.2(b), with representative configurations at different R_{ee} shown in Figure 3.2(a) and (c). The landscape has a minimum close to the contour length of the nanotube. The landscape is asymmetric about the minimum, reflecting that the free energy cost of extending R_{ee} by stretching the nanotube is greater than that of reducing R_{ee} by bending the nanotube. Nevertheless, the nanotube is still extremely stiff towards bending. For example, it costs more than $100\,k_{\mathrm{B}}T$ to reduce R_{ee} by 50%. There is a clear change in the slope at about 160 nm that corresponds to the onset of "kink" formation, where it becomes more favourable to localise much of the stress at a defect rather than to bend the nanotube homogeneously. Finally, there is

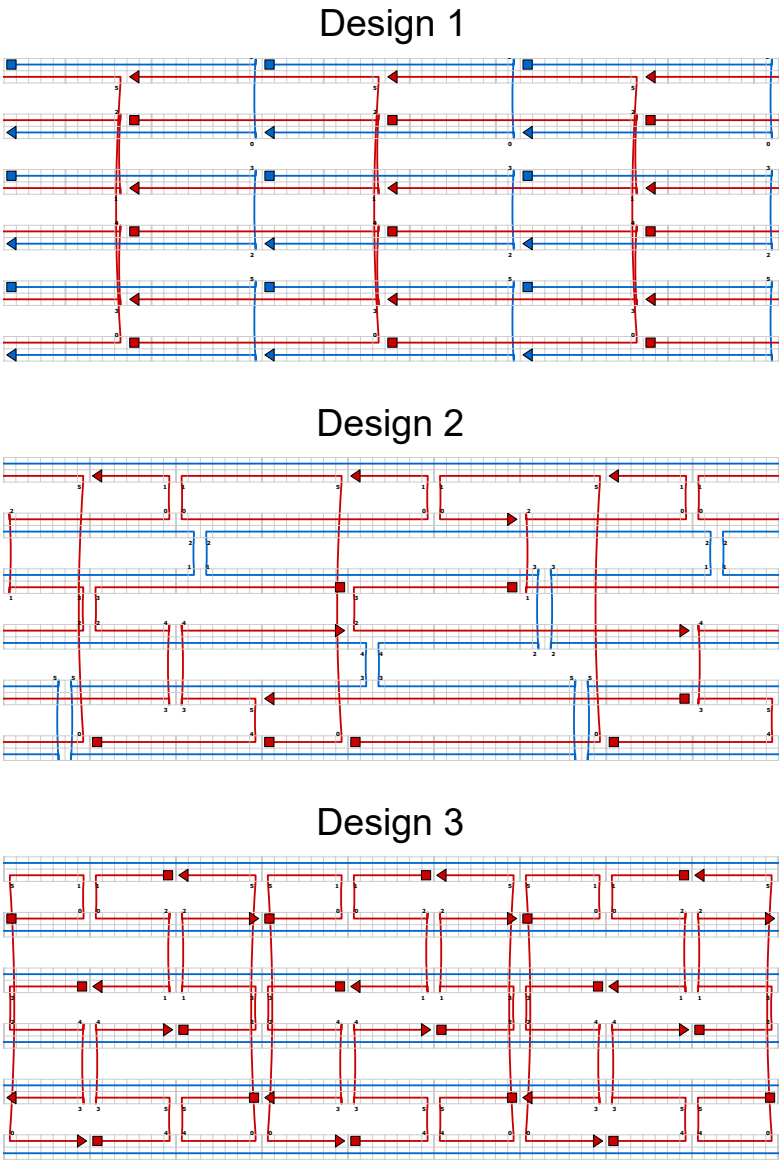


Figure 3.1: Strand routing diagrams for sections of the three nanotube designs as represented in caDNAno [20]. For designs 1 and 3, the pattern repeats along the nanotube. For the origamis (designs 2 and 3), the scaffold is in blue and staples in red. The full designs can be found in the Appendix.

a sharp rise in the free energy as R_{ee} tends towards zero because of overlapping between the nucleotides at the two ends of the nanotube.

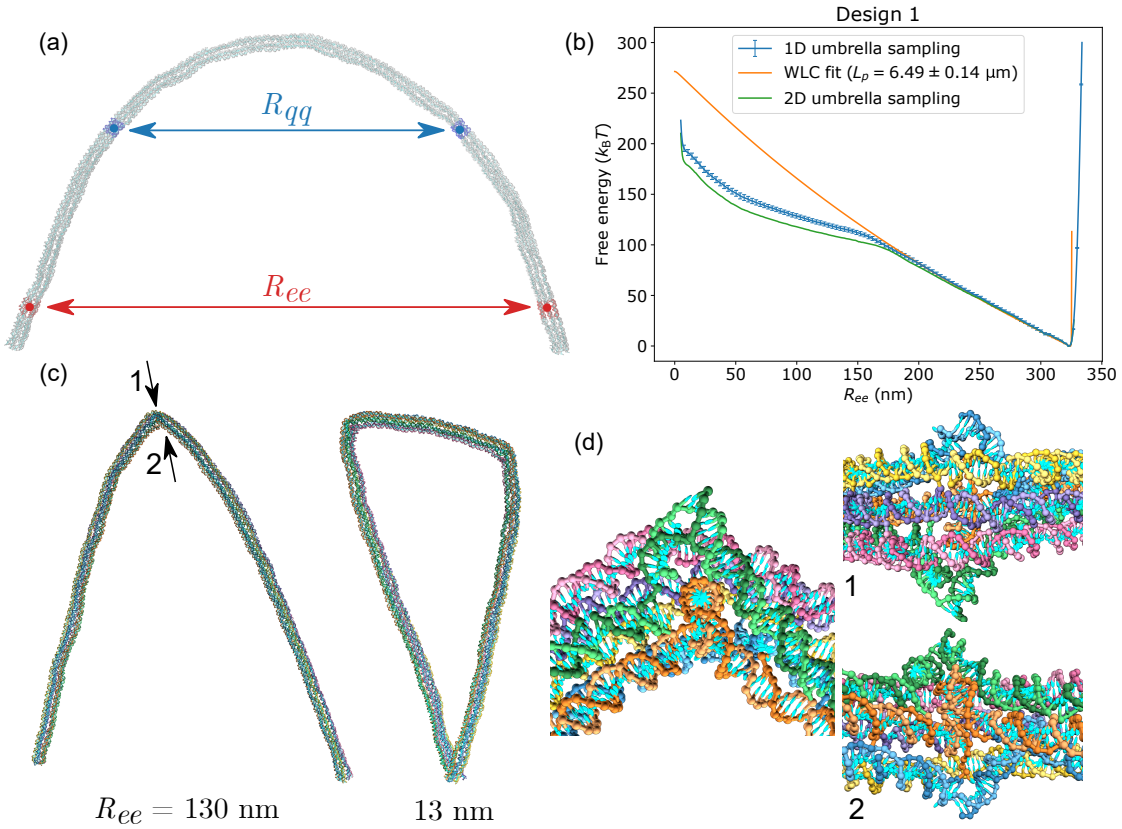


Figure 3.2: (a) The end-to-end distance R_{ee} and the ¼–¾ distance R_{qq} that are used as order parameters in the umbrella sampling illustrated for design 1 with $R_{ee} = 210\,\text{nm}$. Each distance is defined with respect to the centres of mass of the shaded sets of nucleotides. As the helices at the ends of an origami tend to splay out more, we use sets of nucleotides slightly in from the ends to define R_{ee}. (b) Free energy landscape of design 1 as a function of R_{ee}. The weak-bending regime of the landscape can be fitted to the WLC model (see Table 3.1 for the fitting parameters). (c) Further representative configurations at different R_{ee}. (d) Close-ups of the kink for the $R_{ee} = 130\,\text{nm}$ configuration. Views are from the side, outside (view from the arrow labelled 1 in (c)) and inside (view from the arrow labelled 2 in (c)) of the kink.

	L_c (nm)	L_p (µm)
Design 1	325 ± 5	6.49 ± 0.14
Design 2	388 ± 6	5.30 ± 0.13
Design 3	389 ± 9	6.77 ± 0.24

Table 3.1: Parameters for the WLC fit to probability distribution for the end-to-end distance for the three DNA nanotube designs. L_c and L_p are the contour length and persistence length.

In the weak to moderate bending regime without kink formation, we expect the behaviour of the nanotube to be well described by the WLC model. We fitted the free energy landscapes in this regime to the WLC formula for $p(R_{ee})$ given in Becker *et al.* [133]. The WLC formula agrees well with the landscape until the onset of kink formation. Note that the agreement is despite significant local inhomogeneities in the nanotube, such as the easier bending of individual helices at nicks and junctions, and the more constrained relative motion of the helices at junctions. As described in Chhabra *et al.* [110], although these local inhomogeneities lead to significant deviations from the WLC model at short length scales, the nanotubes behave more similar to an ideal WLC at longer length scales. Note also that the WLC model cannot describe the landscape when R_{ee} is larger than that for the free energy minimum, i.e. when the nanotube is stretched. This is because the standard WLC model assumes an inextensible chain, but we can model this part of the landscape using the extensible WLC model [134], which will be done in the section for nanotube stretching.

In the strong bending regime, a kink develops in the nanotube, which leads to deviation from WLC behaviour. The formation of the kink can be caused by the localisation of the bending stress in a small region in order to relieve the bending stress in the whole tube. Generally, the kink occurs near the centre of the nanotube as this is the most effective at reducing the end-to-end distance, but the exact position of the kink varies between individual simulations. A close-up of a kink is shown in Figure 3.2(d). Helices in the outer part of the kink are stretched, while those in the inner part are compressed and forced out of the helical axis. At very small R_{ee} it can become favourable for the nanotube to adopt a configuration with two kinks with bend angles of about 90° rather than a configuration with a single kink with a bend angle of 180°.

The relative smoothness of the free energy landscape in the two regimes suggests that the landscape is well sampled. However, it is challenging to accurately locate the crossover point between the two regimes because kink formation and resolution are relatively slow processes on the simulation time scale. These processes are slow

because kink formation is coupled to global shape changes of the nanotube. For example, the potential benefit of introducing a kink at the centre of a homogeneously bent configuration like that in Figure 3.2(a) is only felt after the bending in the rest of the nanotube relaxes and the sections on either side of the kink become relatively straight.

In order to facilitate sampling in such cases, we perform two-dimensional umbrella sampling where we introduce a second order parameter that allows us to differentiate between the kinked and unkinked states. The order parameter that we introduce is R_{qq}, the distance between points one quarter and three quarters along the nanotube. For a given R_{ee}, R_{qq} should have a smaller value in the kinked state because the nanotube sections on either side of the kink are relatively straight. Furthermore, this order parameter directly facilitates the global shape changes to which kink formation is coupled. For example, decreasing R_{qq} for the configuration in Figure 3.2(a) will begin to localise the bending in the centre of nanotube and straighten the sections on either side.

The resulting two-dimensional free energy landscape is shown in Figure 3.3 for the transition region. There are two clear valleys corresponding to the kinked and homogeneously bent states respectively, with the latter being steeper. At intermediate values of R_{ee}, cuts through the landscape as a function of R_{qq} have two minima separated by a barrier. Such a cut is illustrated in Figure 3.4 at an R_{ee} value near the centre of the kink transition along with representative configurations.

The two-dimensional free energy landscape in the transition region can be projected onto R_{ee} by taking the Boltzmann average over R_{qq}. Then, the information from the projected landscape can be combined with the one-dimensional umbrella sampling at other values of R_{ee} to give the second free energy landscape shown in Figure 3.2(b). Although the two-dimensional sampling has moved the transition to a slightly larger R_{ee}, the relative change is small, suggesting that one-dimensional sampling is sufficient to obtain a reasonably accurate free energy landscape and estimate the point where kink formation begins.

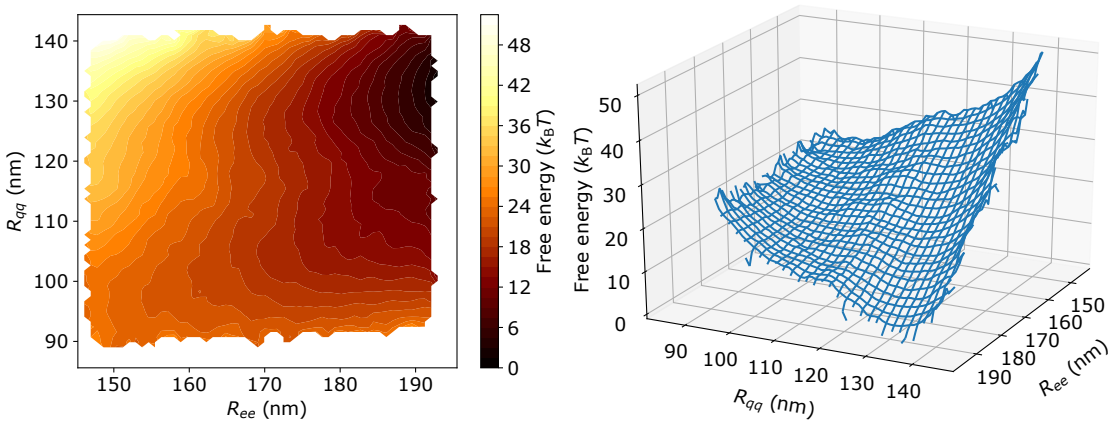


Figure 3.3: Two-dimensional free energy landscape of design 1 as a function of R_{ee} and R_{qq}. The two plots provide two different representations of the same landscape.

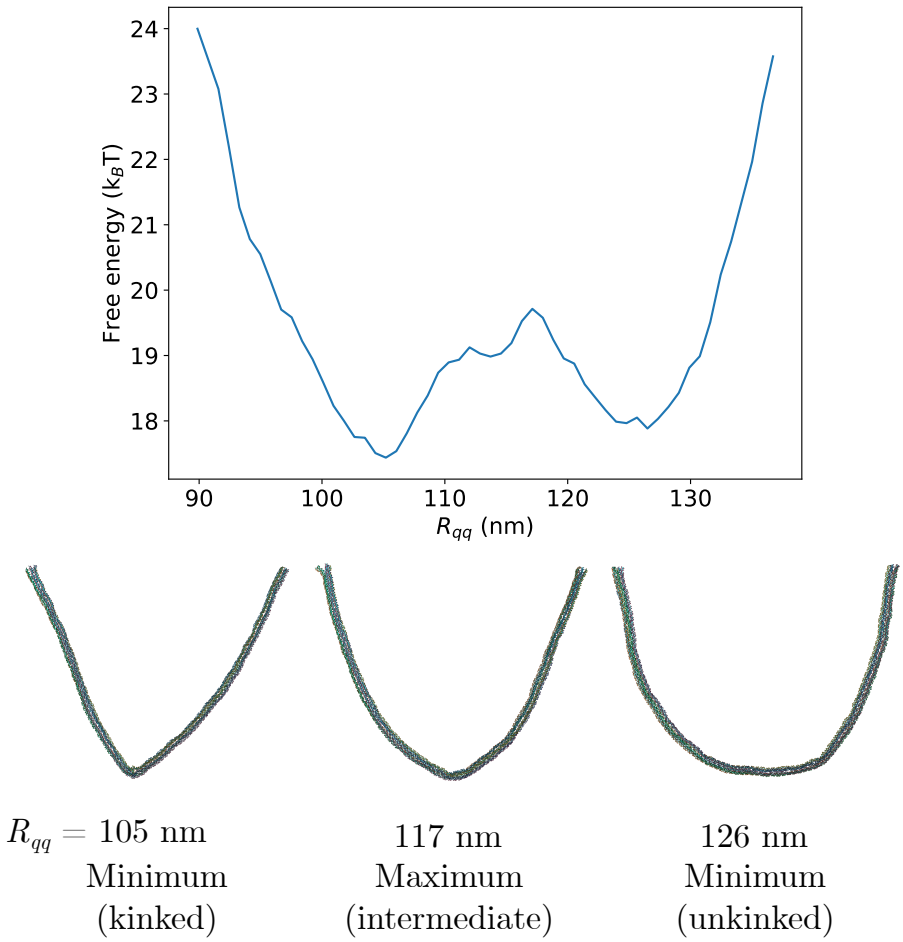


Figure 3.4: A cut through the two-dimensional free-energy landscape for the design 1 nanotube at $R_{ee} = 170\,\text{nm}$, which is near the midpoint of the kink transition. Also depicted are representative configurations corresponding to the two minima and the top of the barrier in this profile.

Comparing the free energy landscapes of the three designs that we consider (Figure 3.2 and 3.5), although they all show the same basic behaviour, there are also significant variations between them.

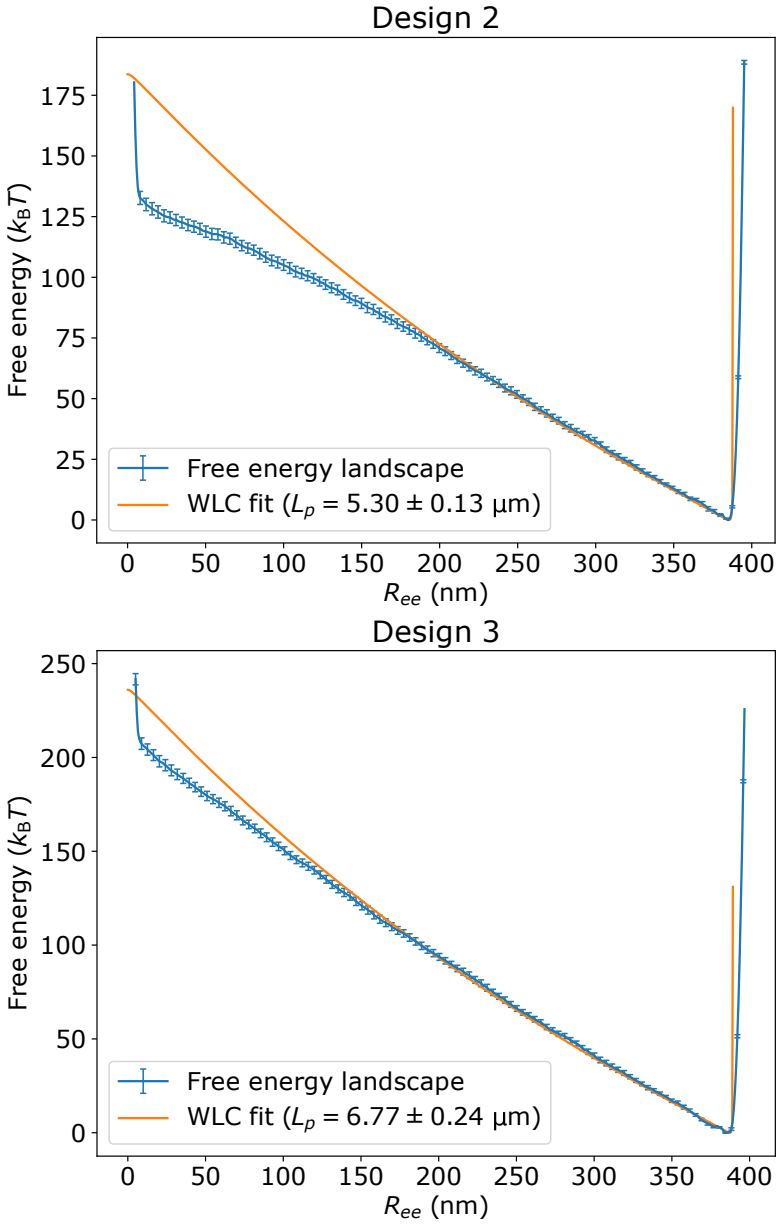


Figure 3.5: Free energy landscapes of designs 2 and 3 as a function of R_{ee}. The weak-bending regime of the landscape can be fitted to the WLC model (see Table 3.1 for the fitting parameters). Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

The persistence lengths obtained from the WLC fits for designs 1 and 2 are 6.5 µm and 5.3 µm, which agree with the values extracted from the thermal fluctuations of

the nanotubes in unbiased simulations [110], but are higher than the experimental values of 3.3 µm [105] and 2.4 µm [132]. We note that for design 1, there was a significant dispersion of persistence lengths (1.6 µm to 4.7 µm) between the individual nanotubes probed in the experiments, which was attributed to structural differences among different nanotubes [105]. Experimental persistence length have also been found to depend on the post-assembly treatment of the origamis [100]. These observations suggest that the experiments are measuring the persistence lengths of a heterogeneous ensemble of nanotubes that include defective ones with a lower persistence length. By contrast, the simulations measure the persistence length of a perfectly assembled nanotube.

The persistence lengths of designs 1 and 3 are very similar, but for design 2 it is lower by about 20%. If one inspects the number of intact base pairs as a function of R_{ee} (Figure 3.6), there is a noticeable difference between the designs; while for designs 1 and 3 there are no significant changes until the onset of kinking, for design 2 there is a small but noticeable decrease. Simulation snapshots of design 2 reveals that the short 2-bp domains noted above are occasionally broken, which gives the nanotube a greater flexibility.

There are some common features in the plots of the R_{ee}-dependence of the number of interacting pairs for the three nanotube designs in Figure 3.6. In the WLC regime of bending, the number of stacked pairs and base pairs remains fairly constant, but the numbers begin to decrease as a kink forms. Meanwhile, the number of coaxially stacked pairs decreases gradually even in the WLC regime. The numbers of stacked pairs and of base pairs show the clearest signals of kinking, with design 1 showing the most abrupt changes. These can be correlated with the changes in R_{qq} shown in Figure 3.7.

Design 1 shows a clear change of slope of the free energy landscape associated with kinking. However, the deviation from the WLC fit is more gradual for designs 2 and 3, and starts to occur prior to clear kink formation (see the example configurations in Figures 3.8 and 3.9 and the energetic and structural descriptors in Figures 3.6 and 3.7). For example, for design 3, R_{ee} has to be significantly less than 100 nm

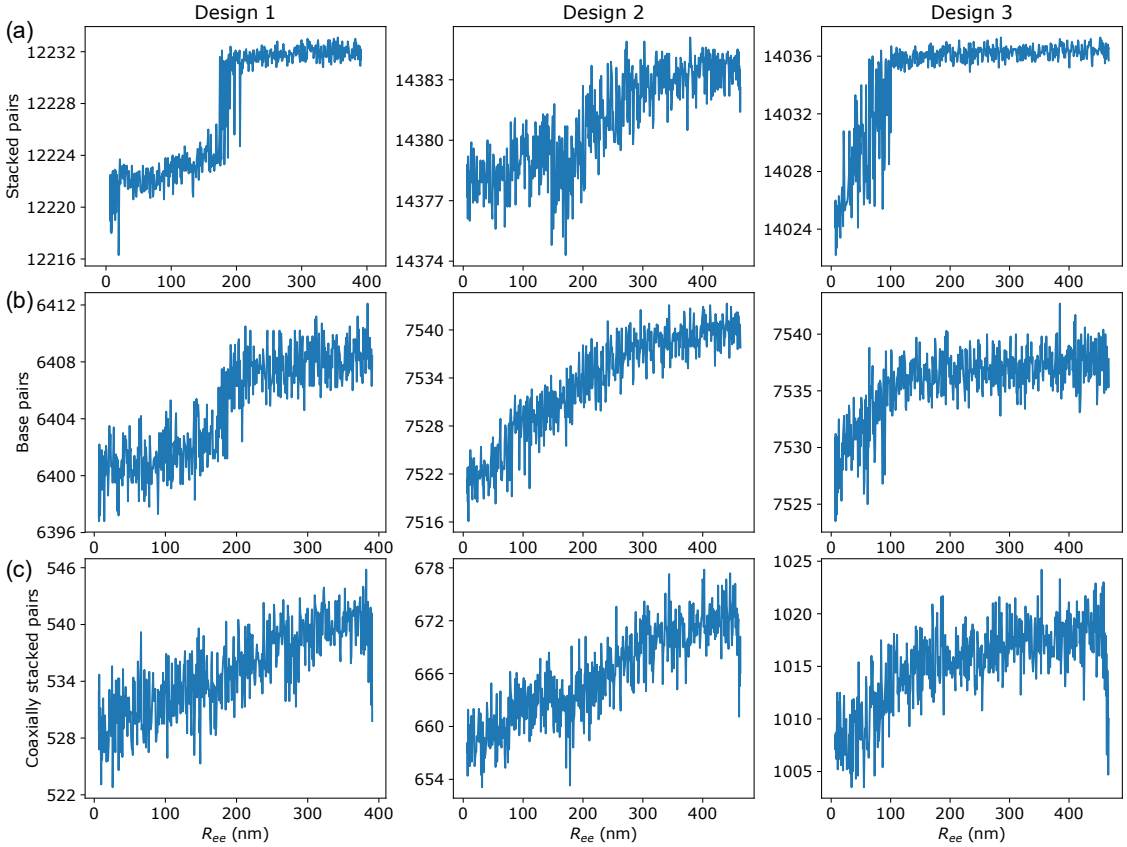


Figure 3.6: The total number of (a) pairs of stacked bases, (b) base pairs, and (c) coaxially-stacked pairs for the three DNA nanotube designs as a function of the end-to-end distance. Each point represents an average over one of the 500 umbrella sampling windows.

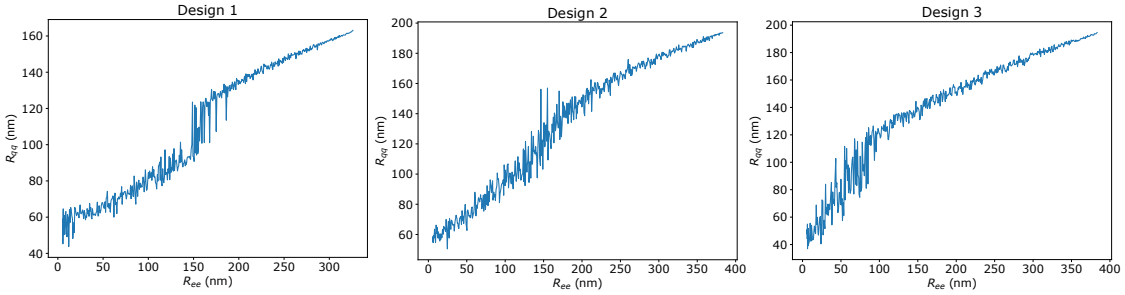


Figure 3.7: The ¼–¾ distance (R_{qq}) as a function of the end-to-end distance (R_{ee}) for the three DNA nanotube designs. Each point represents an average over one of the 500 umbrella sampling windows.

for a kink to appear, even though significant deviations from the WLC fit start to appear at about 150 nm. This non-linear elasticity is because the nanotubes exploit the heterogeneity of their local mechanical properties to partially localise stronger bending at mechanically more flexible regions.

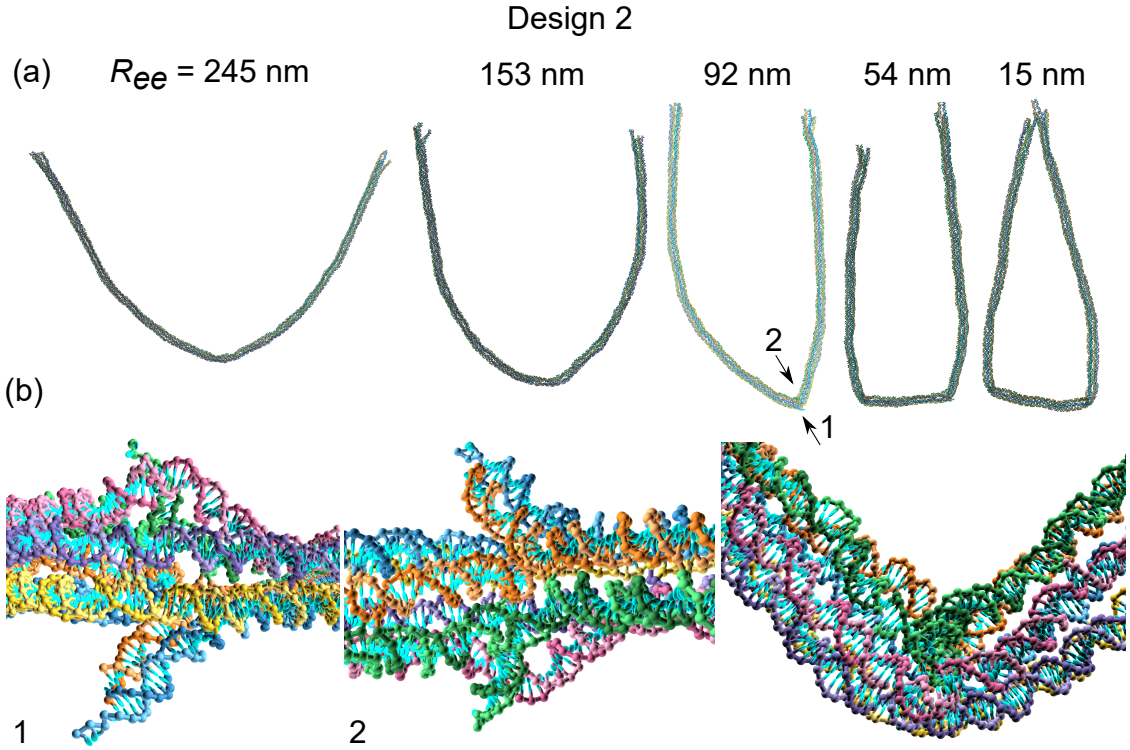


Figure 3.8: (a) Representative configurations of the design 2 nanotube at different values of R_{ee}. (b) Close-ups of the kink for the $R_{ee} = 92$ nm configuration. The view are from outside the kink (left, view from the arrow labelled 1 in (a)), inside the kink (middle, view from the arrow labelled 2 in (a)) and side-on (right).

The kinks in the nanotubes look somewhat like the kinks in a macroscopic tube. The cross-sections become flattened and wider, but the detailed structure depends on the nanotube design, with the main difference being how the tubes respond to the compressive stress on the inside of the kink. The single crossovers in design 1 mean the helical sections have the most freedom to reorient, just through loss of the coaxial stacking interactions at the junction and perhaps some fraying of adjacent base pairs. As shown in Figure 3.2(d), the two outer helices in the inside of the kink splay out, and in the (orange) middle helix a section is oriented almost perpendicular to the nanotube axis. By contrast, the outer three helices are almost co-planar and bend continuously around the kink.

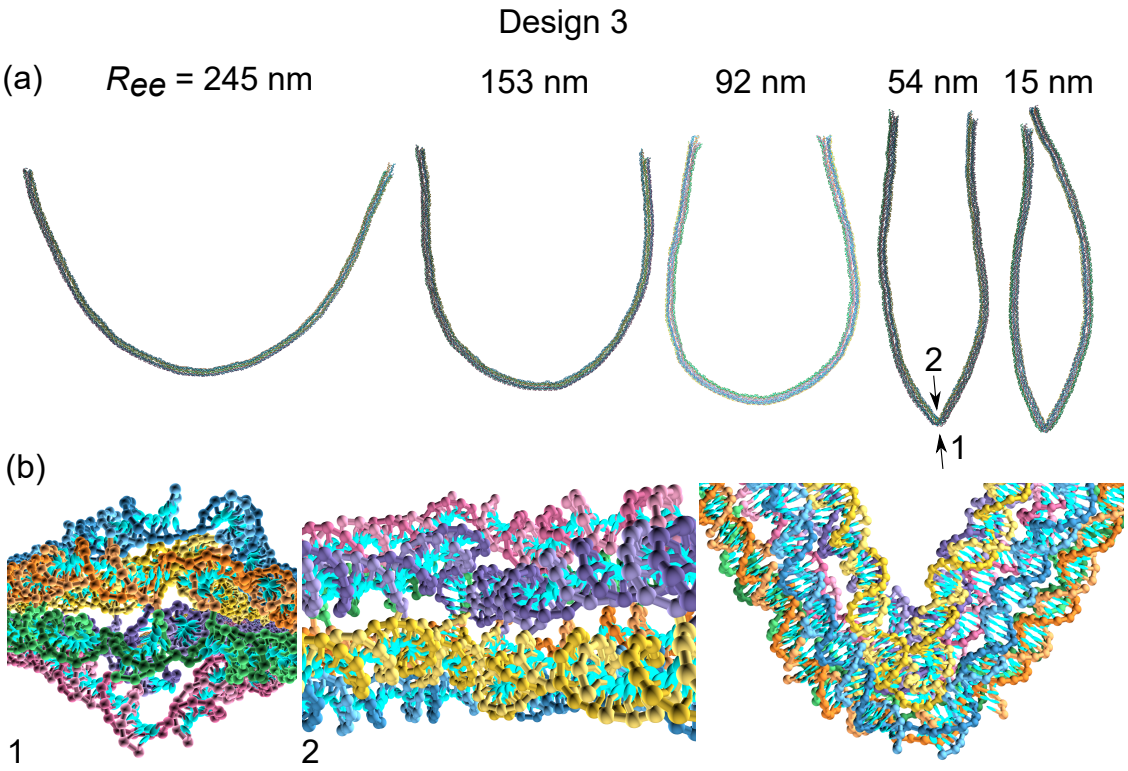


Figure 3.9: (a) Representative configurations of the design 3 nanotube at different values of R_{ee}. (b) Close-ups of the kink for the $R_{ee} = 54\,\text{nm}$ configuration. The view are from outside the kink (left, view from the arrow labelled 1 in (a)), inside the kink (middle, view from the arrow labelled 2 in (a)) and side-on (right).

In design 2, this freedom of motion is achieved by the loss of base pairs, particularly associated with the short 2-bp domains mentioned above. In design 3, the helical sections are the most constrained because of the exclusively double crossovers and the lack of short staple domains. Kinking occurs at the planes of nicks and junctions by the loss of coaxial stacking, but the free-energetic advantage of kinking rather than homogeneously bending is the smallest for this design. The design has significantly increased the nanotube's stiffness and strength with respect to extreme bending compared to design 2.

It is also worth investigating whether extreme bending leads to any irreversible damage to the nanotubes. The extreme stress might perhaps lead to irreversible staple dissociation, but this was never observed in any simulations in the 500 umbrella sampling windows. To explore this further, we took configurations from the most bent umbrella sampling windows, and used them as starting points in

unbiased molecular dynamics simulations. The time evolution of the end-to-end distance is shown in Figure 3.10 for all three designs. All of them were able to relax back to their equilibrium R_{ee} value. Furthermore, after they have fully relaxed, the number of pairs of interactions is essentially the same as that of a normal unstressed nanotube (Table 3.2). This suggests that strong bending and kink formation do not induce irreversible structural changes to the nanotubes. This reversibility is because, on the time scales of our simulations, the structural disruption caused by the extreme bending stresses involves at most only partial melting of the staples rather than their complete dissociation.

The effects of the design differences are also evident from these simulations. In particular, for design 2, the slow re-annealing of broken base pairs at the kink led to a slower recovery of the equilibrium structure.

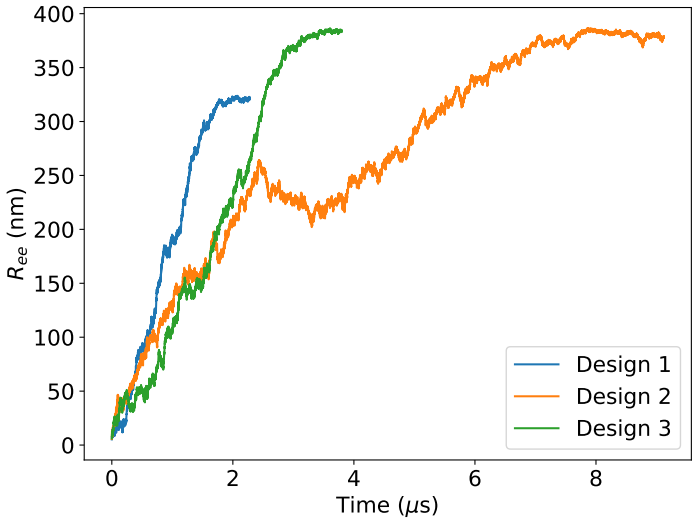


Figure 3.10: Time evolution of R_{ee} in unbiased molecular dynamics simulations that start from nanotube configurations with R_{ee} close to zero that were generated in the umbrella sampling simulations. All nanotubes are able to relax back to their equilibrium state. Movies of these trajectories can be found in [94].

Force-dependent free energy landscape

We can use the free energy landscape to understand the response of the nanotube if a compressive force is applied along the end-to-end vector. In a macroscopic

Design	State	Stacked pairs	Base pairs	Coaxially stacked pairs
1	at equilibrium	12233	6409	540
	bent and released	12232	6408	540
	theoretical maximum	12234	6426	606
2	at equilibrium	14383	7540	672
	bent and released	14384	7538	673
	theoretical maximum	14386	7560	722
3	at equilibrium	14036	7536	1020
	bent and released	14036	7536	1019
	theoretical maximum	14038	7560	1070

Table 3.2: The average number of interaction pairs of different types for the three DNA nanotube designs. These are given for the relaxed nanotube at equilibrium and a nanotube that has been allowed to relax back towards equilibrium from an initially highly bent configuration in an unbiased simulation (i.e. the end configurations for the simulations represented in Figure 3.10). The table also shows the theoretical maximum values allowed by the designs.

system, the rod will remain straight until Euler buckling happens when a critical force is reached. The critical force is given by

$$F_E = \frac{k_\mathrm{B} T L_p \pi^2}{L_c^2}$$

where L_p and L_c are the persistence and contour lengths respectively [135]. In a microscopic system, this instability is rounded by thermal fluctuations, with the transition becoming less sharp as L_c/L_p increases [135]. For the current example $L_c/L_p \approx 0.05$, so thermal fluctuations should not be too important and the transition should be quite sharp.

Figure 3.11(a) shows how the free energy landscape depends on the external force. As the force is increased, the kinked state becomes a free energy minimum. At $F = F_\mathrm{kink} = 2.08\,\mathrm{pN}$, the free energy minima associated with the kinked ($R_{ee} = 54\,\mathrm{nm}$) and unkinked ($R_{ee} = 323\,\mathrm{nm}$) states are degenerate and separated by a barrier of $20.3\,k_\mathrm{B}T$. A transition to the kinked state would be expected when the force increases above F_kink. At $F = F_E = 2.51\,\mathrm{pN}$, the section of the landscape associated with homogeneous bending becomes very flat. When the force increases above F_E, the kinked state becomes the only stable state.

Figure 3.11(b) shows the negative derivative of the computed landscape and the WLC fit. There is an unstable region with a positive slope at R_{ee} values associated with the crossover between the kinked and unkinked states. The noise in the curve is due to the small degree of roughness in the landscape computed from umbrella sampling.

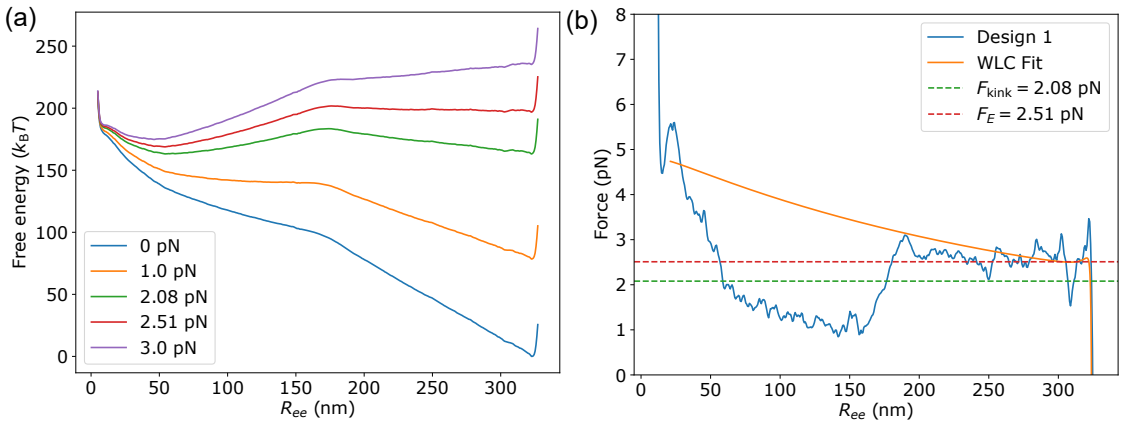


Figure 3.11: (a) The free energy landscape for the design 1 nanotube at a series of compressive forces along the end-to-end vector, where $A(R_{ee}; F) = A(R_{ee}) + FR_{ee}$. (b) $F(R_{ee}) = -dA(R_{ee})/dR_{ee}$ for the computed landscape and its WLC fit. The two horizontal lines correspond to the force at which the two free energy minima associated with the kinked and unkinked states are degenerate and the predicted Euler buckling critical force.

The equivalent plots for the design 2 and 3 nanotubes are shown in Figure 3.12. The behaviour in the homogeneously-bent regime is very similar to the design 1 nanotube, where the free energy landscape becomes very flat at the Euler critical buckling force. This leads to a rapid change in R_{ee} expected for the homogeneously-bent state as the force goes above F_E. The precise values of the critical buckling forces are different for all three systems due to their differences in persistence and contour lengths.

The main differences between the three systems arise from their different propensities to kink. For design 1, the kinked state becomes lower in free energy than the unkinked state well before the critical buckling force is reached (Figure 3.11). For design 2, there is still a force range for which the kinked and unkinked states are both free energy minima, but the force at which they are degenerate is now very close to F_E (at F_{kink} the barrier between the two states is $8.0\,k_{\text{B}}T$).

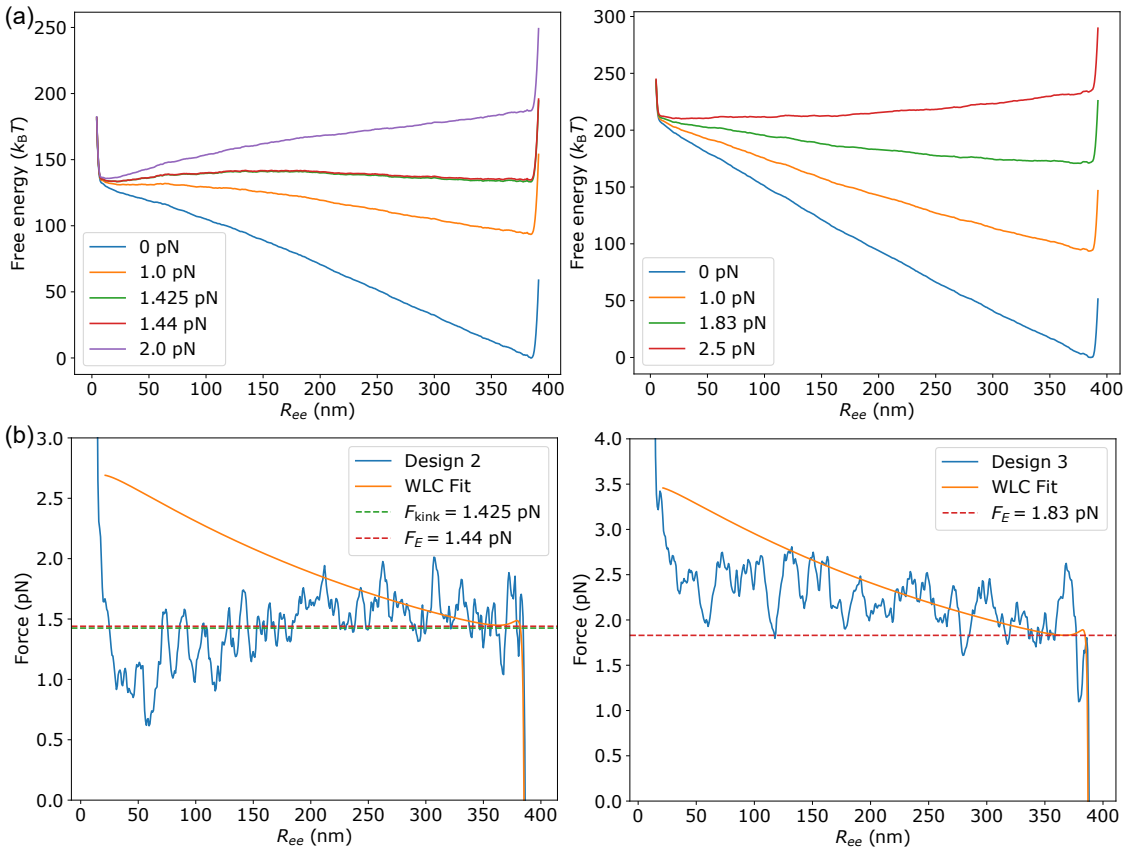


Figure 3.12: (a) The free energy landscape for the design 2 and design 3 nanotubes at a series of compressive forces along the end-to-end vector, where $A(R_{ee};F) = A(R_{ee}) + FR_{ee}$. (b) $F(R_{ee}) = -dA(R_{ee})/dR_{ee}$ for the computed landscapes and their WLC fits. The horizontal lines correspond to the force at which the two free energy minima associated with the kinked and unkinked states are degenerate (design 2 only) and the predicted Euler buckling critical force.

Finally, for design 3, the kinked state only becomes stable above F_E and at no point does the landscape have two free energy minima separated by significant barrier. These changes are partly due to the larger contour length of these two nanotubes, which leads to a lowering of F_E. Another factor is the slope of the (zero-force) free energy landscape in the kinked regime, i.e. the free energy cost of further bending a kink, as this determines the force at which the kinked state becomes a minimum. For both designs 2 and 3, the slope in the kinked regime is very similar to that in the unkinked regime, resulting in a low propensity to kink.

The force-induced bending of DNA nanotubes is experimentally relevant. Single-stranded DNA tethers have been used to connect different parts of a DNA origami to

bend six-helix bundles [100] and cuboidal origami rods [83]. The degree of bending can be controlled by the length of the tether. However, the force exerted by the tether is not constant, but instead depend on its end-to-end distance.

Stretching

The section of the free energy landscape at R_{ee} values beyond the minimum contains information about the stretching of the nanotubes. The formula used to fit the free energy landscape of bending to the WLC model is not applicable here because it does not consider the elasticity of the nanotube. Therefore, we instead fit the free energy landscape of stretching to the extensible WLC model described in Odijk *et al.* [134]. Since this model describes the force-extension behaviour instead of the free energy landscape, we calculate the force-extension curve, $F(z)$, from our free energy landscape, $A(R_{ee})$, as detailed in the Methods section. The free energy landscapes as a function of z for design 1 is shown in Figure 3.13. The fitted elastic modulus, roughly 10 000 pN, is less than six times that for duplex DNA in oxDNA (2700 pN) [110]. The reason is that the nanotube can be extended not only by stretching individual helices, but also by straightening the helices within the origami structure. Helices in an origami are not perfectly parallel because of splaying between junctions [4, 75]. Reduction of splaying will reduce the inter-helical distance and hence the nanotube radius. The inset of Figure 3.13 shows a clear reduction in the nanotube radius in the stretching regime, whereas the radius is roughly constant in the homogeneous bending regime.

	K (pN)	L_c (nm)	L_p (µm)
Design 1	10077 ± 150	325.20 ± 0.08	5.63 ± 0.48
Design 2	9147 ± 89	387.73 ± 0.07	5.00 ± 0.46
Design 3	10454 ± 62	388.93 ± 0.06	6.07 ± 0.74

Table 3.3: Parameters for the extensible WLC fit to the force-extension curves of the three DNA nanotube designs. K, L_c and L_p are the stretching modulus, contour length and persistence length.

The force-extension behaviour of designs 2 and 3 are similar, and are shown in Figure 3.14. The effects of the design differences are once again evident from

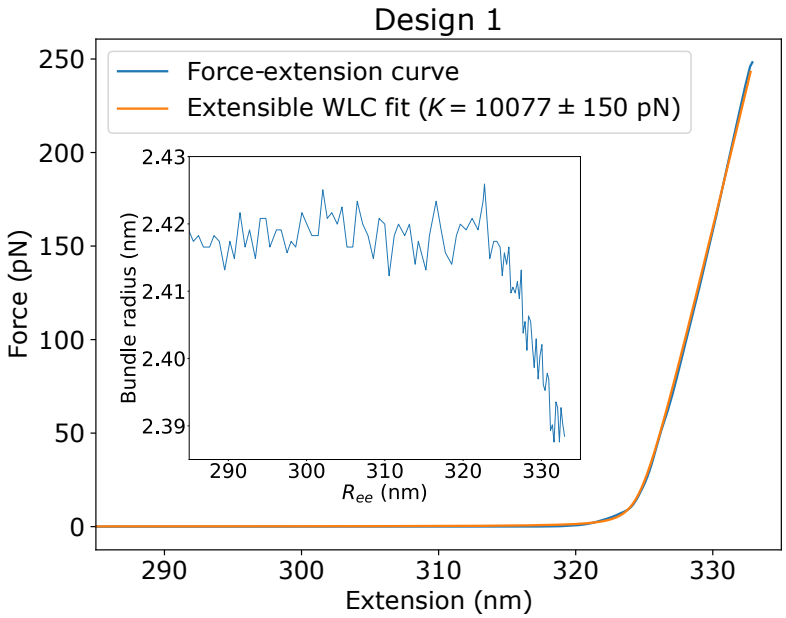


Figure 3.13: Force-extension curve for the design 1 nanotube along with the extensible WLC fit (see Table 3.3 for the fit parameters). The inset shows the average nanotube radius as a function of R_{ee}.

the force-extension curves. Design 2 can withstand a tensile force of about 150 pN before it collapses, while the improved design 3 can withstand up to about 230 pN. The improved design also has a higher elastic modulus (10454 pN vs 9147 pN). These differences further affirm the importance of the placement of nicks and junctions on the mechanical properties of the origami.

3.4 Methods

After the standard conversion and relaxation procedure in the final paragraph of Section 2.1, all designs were further equilibrated for a 10^8 steps, corresponding to about 1.5 µs, before any further simulations were performed.

3.4.1 1D Umbrella Sampling

We used umbrella sampling [92] to calculate the free energy landscape as a function of the order parameter R_{ee}, the end-to-end distance of the nanotube. To avoid effects of helices splaying at their ends, we defined the end-to-end distance as the distance between the centres of mass of two groups of nucleotides. For design 1, the

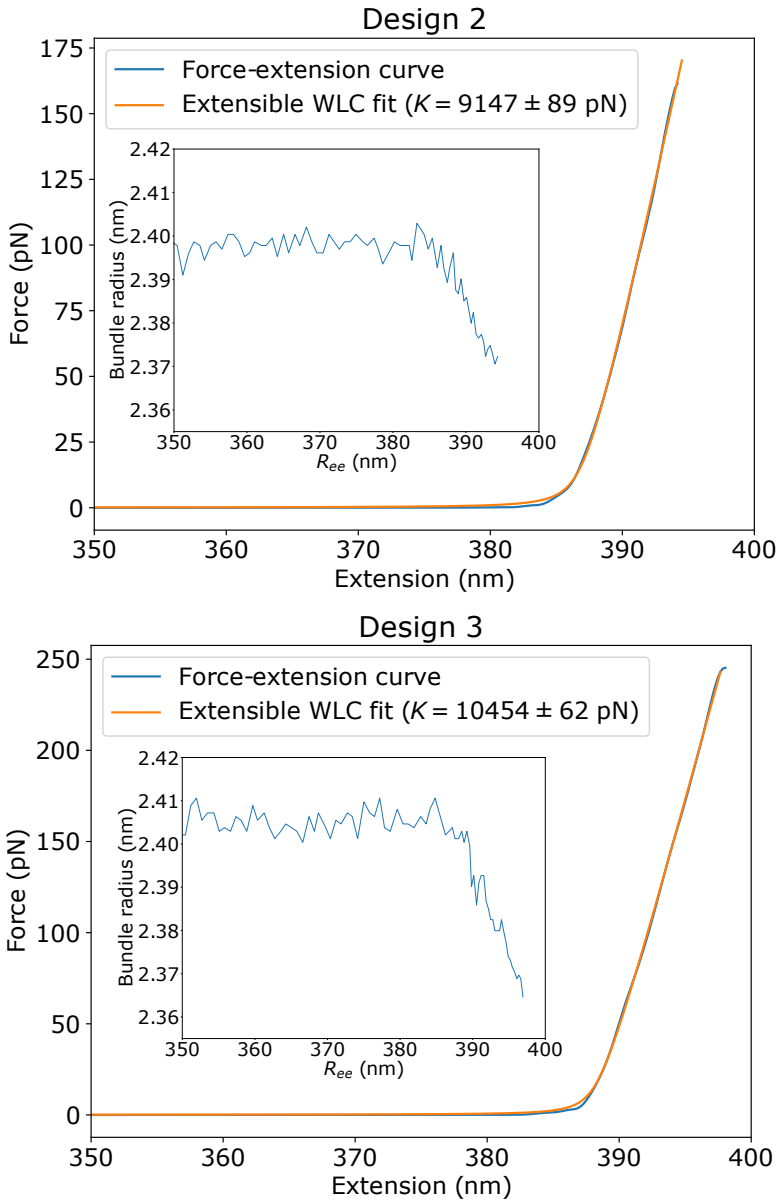


Figure 3.14: Force-extension curve for the design 2 and design 3 nanotube along with the extensible WLC fit (see Table 3.3 for the fit parameters). The inset shows the average nanotube radius as a function of R_{ee}.

two groups contain the nucleotides at positions 44–64 and 1010–1030 respectively. For designs 2 and 3, the two groups contain the nucleotides at positions 49–69 and 1204–1224 respectively. The nucleotide positions are defined in the caDNAno design in Figure A.1.

For each nanotube design, we performed a set of simulations where the order parameter was restrained with a harmonic potential in each sampling window. The sampling pathway for bending ranges from $R_{ee} = 0$ to $R_{ee} = L_c$, where L_c is the contour length of the nanotube, which is estimated as the number of base pairs between the two groups along the nanotube multiplied by the rise (0.34 nm). The sampling pathway for stretching ranges from $R_{ee} = L_c$ to the value of R_{ee} when the nanotube collapses.

Before umbrella sampling simulations, we prepared the starting configurations for each window using pulling simulations. The end-to-end distance of an equilibrated configuration was restrained at $R_{ee} = L_c$ with a harmonic bias potential of stiffness $k = 57.09\,\mathrm{pN/nm}$. For the bending pathway, the equilibrium position of the harmonic bias potential was gradually reduced at a constant rate such that it reaches zero after 10^8 steps. Configurations were outputted every 2×10^5 steps, resulting in a total of 500 configurations to be used as the starting configuration of the same number of equally-spaced windows from $R_{ee} = 0$ to $R_{ee} = L_c$. For the stretching pathway, the equilibrium position of the harmonic potential was gradually increased at the same constant rate as the bending pathway, until the nanotube collapses. Configurations were also outputted every 2×10^5 steps, resulting in a total of 28, 27, and 37 configurations (for designs 1, 2, and 3 respectively) to be used as the starting configuration of same number of equally-spaced windows starting from $R_{ee} = L_c$ to the point of collapse.

After the starting configurations have been prepared, the configuration in each window was restrained at the value of R_{ee} corresponding to that window with a harmonic bias potential of stiffness $k = 17.12\,\mathrm{pN/nm}$. Each window was equilibrated for 10^6 steps before a production run of 10^7 steps. R_{ee} was outputted every 10^3 steps, giving 10^4 data points for each window. Using the biased probability distributions of

R_{ee} in each window, we used the Weighted Histogram Analysis Method (WHAM) [93, 136] to calculate the unbiased free energy landscape of the system as a function of R_{ee}.

The production run was then repeated with the last configuration of the previous production run, and WHAM was performed on the new data points to calculate the free energy landscape. This process was repeated until the new free energy landscape is not significantly different from the previously calculated free energy landscape. For the nanotubes considered here, 4×10^7 steps of production run was needed for convergence.

3.4.2 2D Umbrella Sampling

To improve sampling in the transition region between the nanotube being kinked or unkinked, two-dimensional umbrella sampling was performed for design 1. The order parameters were chosen to be R_{ee}, the end-to-end distance of the nanotube, and R_{qq}, the distance between the first quarter region and third quarter region of the nanotube. R_{ee} has the same definition as in 1D umbrella sampling, while R_{qq} is defined as the distance between the centres of mass of another two groups of nucleotides, which contain the nucleotides at positions 285–305 and 768–788 respectively. The nucleotide positions are defined in the caDNAno design in Figure A.1.

The sampling region ranges from $R_{ee} = 148.0\,\text{nm}$ to $R_{ee} = 190.7\,\text{nm}$, and $R_{qq} = 93.7\,\text{nm}$ to $R_{qq} = 139.7\,\text{nm}$. Windows along R_{ee} had a spacing of $0.658\,\text{nm}$, which is the same as that used for the one-dimensional sampling, and windows along R_{qq} had a spacing of $1.70\,\text{nm}$.

The equilibrated configuration from the corresponding R_{ee} window in 1D umbrella sampling was used as the starting configuration for each window. The configuration in each window was restrained at the value of R_{ee} and R_{qq} corresponding to that window with a harmonic bias potential of stiffness $k = 17.12\,\text{pN/nm}$ and $k = 1.71\,\text{pN/nm}$ respectively. Each window was equilibrated for 5×10^6 steps before a production run of 5×10^6 steps. After the production run, two-dimensional WHAM [136, 137] was performed to calculate the free energy landscape in the transition region.

3.4.3 Combining 1D and 2D umbrella sampling

To combine the free energy landscape from one- and two-dimensional umbrella sampling of the SST design, we also recorded R_{qq} in the one-dimensional sampling runs, but without any bias on this order parameter. We performed two-dimensional WHAM on the whole data set, but excluded the data from one-dimensional sampling that overlapped with those from two-dimensional sampling. The two-dimensional free energy landscape was then projected to a one-dimensional landscape by the following formula.

$$\begin{aligned} A(R_{ee}) &= -k_B T \ln P(R_{ee}) \\ &= -k_B T \ln \int_0^\infty P(R_{ee}, R_{qq}) dR_{qq} \\ &= -k_B T \ln \int_0^\infty \exp\left(-\frac{A(R_{ee}, R_{qq})}{k_B T}\right) dR_{qq} \end{aligned}$$

3.4.4 Interacting nucleotides

Three types of interactions are counted, namely hydrogen bonding, stacking, and coaxial stacking (Figure 1.1). A pair of nucleotides is considered as interacting through a particular interaction when the relevant interaction energy is lower than $-2.49\,\mathrm{kJ/mol}$.

To produce the plots in Figure 3.6, the total number of interacting pairs for each interaction type is counted for each configuration in the production trajectory in each umbrella sampling window, and averaged over each trajectory.

3.4.5 Deriving the force-extension curve from the free energy landscape

We calculate the force-extension curve, $F(z)$, from our free energy landscape, $A(R_{ee})$, using the following method. $A(z)$ denotes the free energy as a function of z. $P(z)$ denotes the probability distribution function of z. $C(z)$ denotes the cumulative distribution function of z. $P(r)$ denotes the probability distribution function of r.

$$F(z) = -\frac{dA(z)}{dz}$$

$$
\begin{aligned}
&= -\frac{d(-k_B T \ln P(z))}{dz} \\
&= k_B T \frac{d \ln P(z)}{dz} \\
P(z) &= \frac{dC(z)}{dz} \\
&= \frac{d}{dz}\left(1 - \int_0^{2\pi} \int_0^{\arccos(\frac{z}{r})} \int_z^{\infty} P(r) r^2 \sin\theta dr d\theta d\phi\right) \\
&= \frac{d}{dz} \int_z^{\infty} P(r) 2\pi r^2 \left(\frac{z}{r} - 1\right) dr \\
&= \int_z^{\infty} P(r) 2\pi r dr \\
P(r) &= \frac{1}{4\pi r^2} \exp\left(\frac{-A(R_{ee})}{k_B T}\right)
\end{aligned}
$$

3.4.6 Nanotube radius

For each umbrella sampling window, the average nanotube radius was averaged over each configuration in the production trajectory. In each configuration, the average nanotube radius was calculated by averaging the radius at each slice (defined as the set of nucleotides with the same base-pair index in the caDNAno file) along the nanotube, where each slice is separated by a single base-pair step:

$$\langle r_{\text{nanotube}} \rangle = \frac{1}{N_{\text{slice}}} \sum_{i=1}^{N_{\text{slice}}} r_{\text{nanotube}}(i).$$

This is in turn calculated by averaging the distances from the centre of each double helix to the centre of the nanotube for that slice:

$$r_{\text{nanotube}}(i) = \frac{1}{N_{\text{helix}}} \sum_{j=1}^{N_{\text{helix}}} \left|\mathbf{R}_{\text{duplex},j}(i) - \mathbf{R}_{\text{nanotube}}(i)\right|,$$

where the nanotube centre is simply defined as

$$\mathbf{R}_{\text{nanotube}}(i) = \frac{1}{N_{\text{helix}}} \sum_{j=1}^{N_{\text{helix}}} \mathbf{R}_{\text{duplex},j}(i).$$

Following [110] the centre of a double helix, $\mathbf{R}_{\text{duplex}}$, is defined as

$$\mathbf{R}_{\text{duplex}} = \frac{1}{2}\left(\mathbf{r}_{\text{nuc1}} + \mathbf{r}_{\text{nuc2}}\right) + \frac{\alpha}{2}\left(\widehat{\mathbf{b}}_{\text{nucl}} \times \widehat{\mathbf{n}}_{\text{nuc1}} + \widehat{\mathbf{b}}_{\text{nuc2}} \times \widehat{\mathbf{n}}_{\text{nuc2}}\right)$$

where $\mathbf{r}$, $\widehat{\mathbf{b}}$, $\widehat{\mathbf{n}}$ are the centre of mass, backbone-base unit vector, and the normal unit vector of the oxDNA nucleotide, respectively, and $\alpha = 0.06$ (in the oxDNA

simulation unit of length). The second term is a correction factor that takes into account that the centre of mass of the two nucleotides in an oxDNA base pair is slightly displaced towards the minor groove.

3.5 Conclusions

We have shown that it is feasible to compute the free energy landscapes associated with the mechanical deformation of large DNA nanostructures, such as DNA origami, when described with a nucleotide-level coarse-grained model. The use of umbrella sampling with many windows allows all parts of the free energy landscapes to be sampled in a highly parallel manner.

We studied the extreme bending of DNA nanotubes by computing the free energy landscape as a function of the end-to-end distance. Such nanotubes are extremely stiff and are promising building blocks for nanoscale engineering. In the homogeneous bending regime, the behaviour is well described by the WLC model. In the extreme bending regime, the bending stress localises at one or two kinks, where the ease of kinking and the nature of the deformations at the kink depends sensitively on the local features of the nanotube design. Importantly, the nanotubes can reversibly recover from this kinking when the stress is removed. Meanwhile, the landscapes at end-to-end distances larger than the nanotube's relaxed length allows us to extract the stretch modulus of the nanotubes. Their values are somewhat less than what would be expected for six helices stretched in parallel, because the nanotubes can also respond to the tension by reducing their radius.

In order for it to be feasible to compute free energy landscapes for DNA origamis using the current approach, it is important that the properties of the systems vary continuously as the order parameters change, so that equilibration within each umbrella sampling window can be relatively rapid. Such calculations becomes more problematic when discontinuous changes occur. For example, when studying the kinking transition of the nanotubes, in order to more accurately locate the transition, we needed to introduce a second order parameter that provides a more continuous pathway between the kinked and unkinked states. Thus, the current approach can

be most easily applied to landscapes associated with mechanical deformations and could be applied to larger multi-origami assemblies in such cases. By contrast, calculating free energy landscapes that involve significant assembly or disassembly, although feasible for small and medium-sized DNA systems [85, 86, 121–124], is unlikely to be feasible for systems of the size of DNA origamis.

In addition to exploring the fundamental mechanical properties of DNA origami, we expect the current approach to be useful in DNA mechanotechnology [98], where DNA origami systems are increasingly being used as part of force-sensing [111] or force-applying [112] devices, and an accurate characterisation of the compliance of the origami devices under the relevant mechanical loads is likely to be important for their calibration. An accurate characterisation of DNA origami mechanics is also likely to be important for understanding stress accumulation in the self-assembly of origamis into higher-order structures, be it an unwanted side effect which hinders correct assembly, or a design strategy to control and direct the assembly [114, 115].

4

Free Energy Landscape of a Mechanically Bistable DNA Origami

Contents

4.1 Introduction 48
4.2 Origami design and structure 49
4.3 Free Energy Landscape 51
4.4 Free Energy Decomposition 52
4.5 Methods 58
4.6 Conclusions 62

The results of this chapter have been published in Applied Sciences [138].

4.1 Introduction

An area of increasing interest in the field of DNA nanotechnology is the use of DNA origamis as nanoscale mechanical devices [139]. In this field of DNA "mechanotechnology" [98], DNA constructs are used to generate [112, 140, 141], transmit [102] and sense [111] nanoscale forces. One approach involves the use of flexible DNA origamis that can undergo large-scale structural changes [81, 83, 142]. For example, an externally-actuated shape change could be used to apply a force [143], or a force could be sensed through the structural change it induces [144–146].

For these applications, it is important to fully understand the shape changes that origamis can undergo and how the thermodynamics and mechanics of these shape changes can be controlled through the origami design. Computational approaches potentially play an important role as they can provide direct insight into how the microscopic features of the design impact the overall properties of the origami [80, 82, 147].

Here, we provide a detailed simulation analysis of the jointed DNA origami introduced by Zhou *et al.* [148] that was designed to be bistable with the free-energy barrier between the two stable states being a result of the internal mechanical stresses associated with intermediate states. Observations of an ensemble of origami structures by electron microscopy seem to have confirmed the bistability of the design. In particular, we will compute the free-energy landscape of this system and explore the role of the properties of the individual joints in determining the overall behaviour of the system. This analysis will raise interesting questions about the potential role of defects on the properties of these origami systems.

We will use the oxDNA coarse-grained model [47–49] to model the DNA origamis. As described in Chapter 1, oxDNA provides a good representation of the structures of DNA origamis [49, 75]. It is particularly well-suited to study jointed origamis, as it can accurately capture the mechanical properties of single-stranded DNA [48], which is important for modelling the role of the single-stranded linkers on the flexibility of the joints.

4.2 Origami design and structure

The jointed origami consists of four rod-like blocks of parallel double helices based on a hexagonal lattice that are linked through scaffold connections forming a roughly planar arrangement (Figure 4.1). In Zhou *et al.* [148], these blocks have been termed the frame (dark blue), compliant (red), crank (green) and coupler (cyan) blocks. The central section of the compliant block is a 6-helix bundle and designed to bend more easily than the other blocks which are all based on 10-helix bundles. For example,

using the oxDNA model such a 10-helix bundle has been found to have a bending persistence length that is about three times larger than a 6-helix bundle [110].

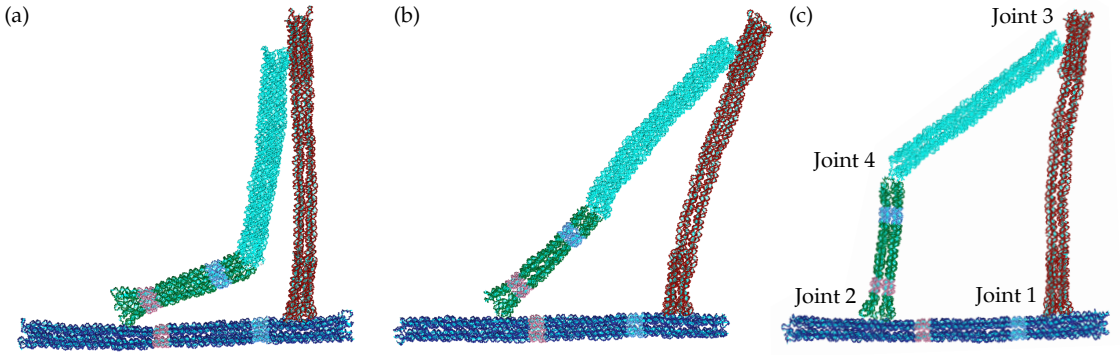


Figure 4.1: Representative oxDNA configurations of the origami in (a) the closed state S1 ($\theta \approx 20°$) (b) the intermediate state U ($\theta \approx 45°$) and (c) the open state S2 ($\theta \approx 80°$). The four joints are numbered in (c).

Joint 1 between the frame and the compliant blocks was designed to be rigid, whereas joints 2–4 were designed to behave like hinges allowing free rotational motion of the blocks in the plane of the structure. Joints 2 and 3 are connected by short single-stranded linkers on one side of the block so that the block can rotate about the joint. The linkers for joint 4 are somewhat longer to provide additional flexibility so that the joint can bend both inwards and outwards.

The origami was designed to have a free-energy landscape with two minima, the "closed" state S1 and the "open" state S2, as a function of θ, the angle between the crank and the frame blocks. Since joint 1 is rigid, in order for the origami to transition between the two stable states, the compliant block needs to bend outwards to accommodate the extra length between the ends of the crank and the coupler blocks. When the crank and the coupler blocks are collinear, the bending in the compliant block is at its maximum. This intermediate state U is thus expected to be unstable and should appear as a local maximum on the free-energy landscape. A simple mechanical model that accounts for the bending cost of the compliant block, with the assumption of free rotation at joints 2–4, predicts a free-energy barrier of about $5\,k_{\mathrm{B}}T$ [148]. Representative oxDNA configurations at values of θ corresponding to S1, U and S2 are depicted in Figure 4.1. The substantial bending

of the compliant block in the intermediate state U is very apparent, but in the open state, there is still a small amount of bending of this block.

Experimentally, the free-energy landscape of this origami has been inferred from an ensemble of structures visualised using electron microscopy [148]. Specifically, the probability distribution of θ was measured, and the free-energy landscape was derived from the logarithm of this distribution. The main peaks at 28° and 80° correspond to states S1 and S2, but surprisingly there is a minor peak in the probability distribution at intermediate values of θ rather than a local minimum. It was suggested that this sub-population in the U state might be a result of coaxial stacking between the ends of the crank and coupler blocks when they are collinear.

4.3 Free Energy Landscape

The free-energy landscape was calculated using umbrella sampling and WHAM. We chose the order parameter as the distance R between the centres of mass of two groups of nucleotides in the crank block and the frame block, respectively. The groups of nucleotides are highlighted in Figure 4.2(a). We chose to use R as the order parameter rather than θ because calculating the forces on the nucleotides due to the umbrella potential is then much simpler and has already been implemented in the oxDNA simulation code [94]. The values of R and θ were tracked during the umbrella sampling simulations so that we can transform the free-energy landscape from a function of R to a function of θ. The resulting landscape is shown in Figure 4.2(b).

Contrary to previous expectations, the landscape obtained using the oxDNA model only shows a single minimum at $\theta \approx 80°$, which corresponds to state S2. There is no sign of any second minimum corresponding to state S1. Instead, the free energy continuously increases as θ decreases. At $\theta \approx 20°$, where state S1 is expected to occur, the free energy is more than $20\,k_BT$. This is significantly different from the experimental results, where about 30% of the origamis were observed to be in the closed state.

Given the seeming disagreement between the obtained free-energy landscape and the experimental results, we further explored the dependence of the landscape

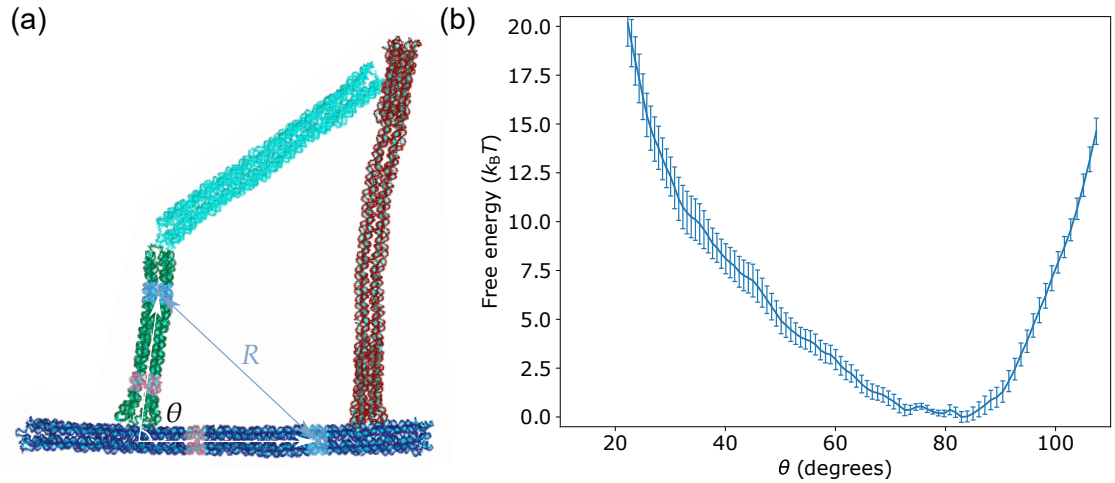


Figure 4.2: (a) A configuration of the origami illustrating the definition of the order parameter R as the distance between the centres of mass of the two groups of nucleotides coloured in light blue. The value of θ was tracked in the simulations, and is defined as the angle between the two vectors coloured in white, which pass through the centres of mass of the two groups of nucleotides coloured in light red and light blue respectively. (b) The free-energy landscape of the origami as a function of θ. Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

on electrostatics and temperature. Since the description of ionic interactions is relatively simple in the oxDNA model, we first checked that the destabilisation of the closed state is not due to an overly strong electrostatic inter-helical repulsion. Even when electrostatics were completely turned off, the landscape still had a single free-energy minimum corresponding to the open state (Figure 4.3a). We also explored the temperature dependence of the free-energy landscape, but there was little change over the 40 K window that we considered (Figure 4.3b). These results confirm that the the single-minimum character of the free-energy landscape is a robust feature of the oxDNA simulations.

4.4 Free Energy Decomposition

To better understand the form of the oxDNA free-energy landscape, particularly the absence of a minimum corresponding to state S1, we attempted to compute the contributions to the free-energy landscape from the four joints. In the complete origami, the motion of the four joints is coupled due to the overall structure

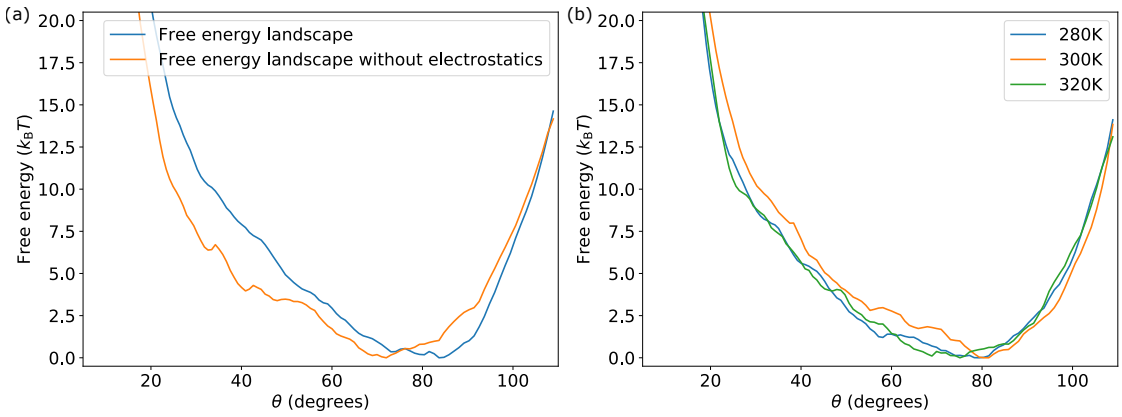


Figure 4.3: (a) Free-energy landscape of the complete origami as a function of θ at $[Na^+] = 1\,M$ and with the electrostatic interactions turned off. $T = 300\,K$. (b) Free-energy landscape of the complete origami as a function of θ at $[Na^+] = 1\,M$ at three different temperatures, as labelled.

of the origami. Therefore, to isolate the behaviour of the individual joints, we performed a series of simulations with one or more blocks of the origami removed (see the configurations in Figure 4.4). Similar to the previous section, we then defined distance-based order parameters between the centres of mass of groups of nucleotides to probe the motion of a given joint. The definitions of the order parameters and their respective free-energy landscapes are shown in Figure 4.4.

Joint 1 was designed to be rigid. As expected, the free-energy minimum is when the frame and compliant blocks are perpendicular and unbent. The free-energy cost of changing the distance order parameter comes mainly from the bending of the compliant block instead of a change in the angle at the joint. This degree of freedom is the intended mode of stress accumulation in the original design that leads to the intermediate state being unfavourable.

Joints 2 and 3 were both designed to act as hinges. However, their free-energy landscapes show that they are not completely flexible. Instead, they have preferred opening angles at around 55° and 70°, respectively, and there is a non-negligible free-energy cost when the angles deviate from their preferred values. In particular, there is significant steric repulsion between the helices in the two blocks when the opening angle goes beyond 90° or approaches zero.

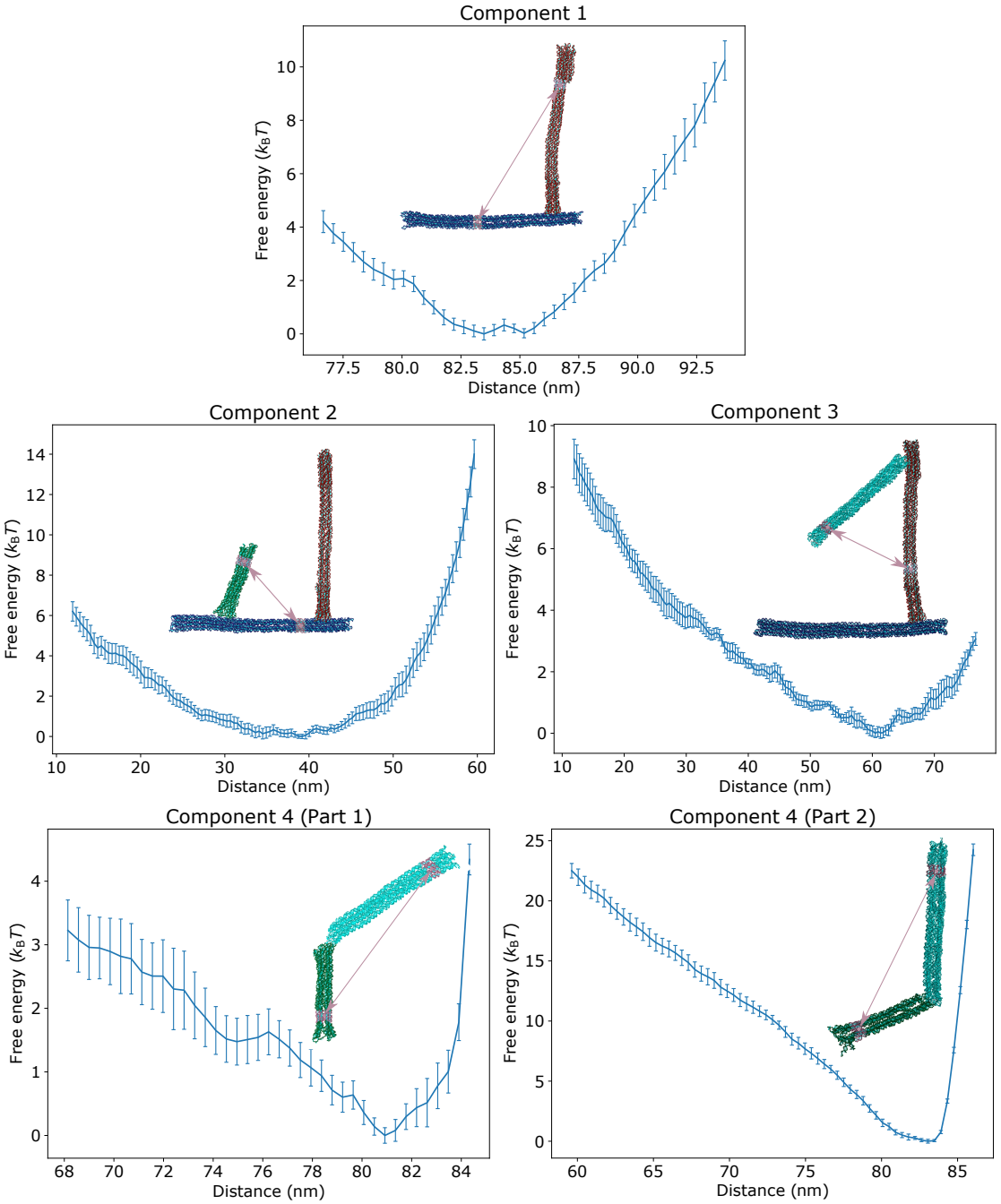


Figure 4.4: The free-energy landscapes associated with each individual joint. The configurations illustrate the sub-systems used to compute these landscapes and the definitions of the distance order parameters. The values of these order parameters in the depicted configurations are 86 nm, 45 nm, 51 nm, 78 nm and 70 nm, respectively. The landscape associated with joint 4 has been computed in two parts corresponding to whether the joint is pointing outwards (1) or inwards (2). Only when joint 4 is straight do these sub-landscapes sample equivalent configurations. Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

Joint 4 was designed to be more flexible as the joint needs to bend both inwards and outwards to allow both the open and closed states to be easily accessed. The landscape for the joint has been computed in two sections because the distance order parameter cannot distinguish whether the interior angle of the joint is obtuse (as in the open state) or reflex (as in the closed state). The joint prefers to be roughly straight with an increasing free-energy penalty for more bent configurations. We do not observe any coaxial stacking between the helix ends either side of the joint even when it adopts a straight configuration. It is worth noting the asymmetry between bending the junction outwards and inwards, with a much higher cost for the bending inwards due to the resulting steric repulsion between the helices near to the joint (Figure 4.5). Note that the sharp rise in the free energy at large values of the distance order parameter is due to the stretching of the straight joint, and is not relevant to the behaviour of the complete origami.

By keeping track of the distance order parameters of each joint, as well as R and θ, in the umbrella sampling simulations of the complete origami, we can transform the free-energy landscapes of the individual joints to be functions of θ (Figure 4.6). We next check if the free energy of the complete origami, where the motion of the joints is coupled, can be represented as the sum of the free-energy contributions from the individual joints. The comparison in Figure 4.7 shows that this is a very good approximation. Having established this conclusion, we now use these landscapes to understand why oxDNA predicts the closed state S1 to be unstable.

The contribution to the free energy of the complete origami associated with joint 1 shows two minima at $\theta \approx 20^\circ$ and at $\theta \approx 90^\circ$, and a maximum at around $\theta \approx 45^\circ$ (Figure 4.8). These positions generally agree with the angles for the S1, U and S2 states observed experimentally [148]. This component of the free-energy landscape is very similar to that predicted by the mechanical model used in that paper, although the barrier is slightly larger, which could be due to the somewhat higher persistence lengths that oxDNA predicts for DNA origamis [110]. These results indicate that the designed mechanism for stress accumulation due to the bending of the compliant block can be captured by oxDNA.

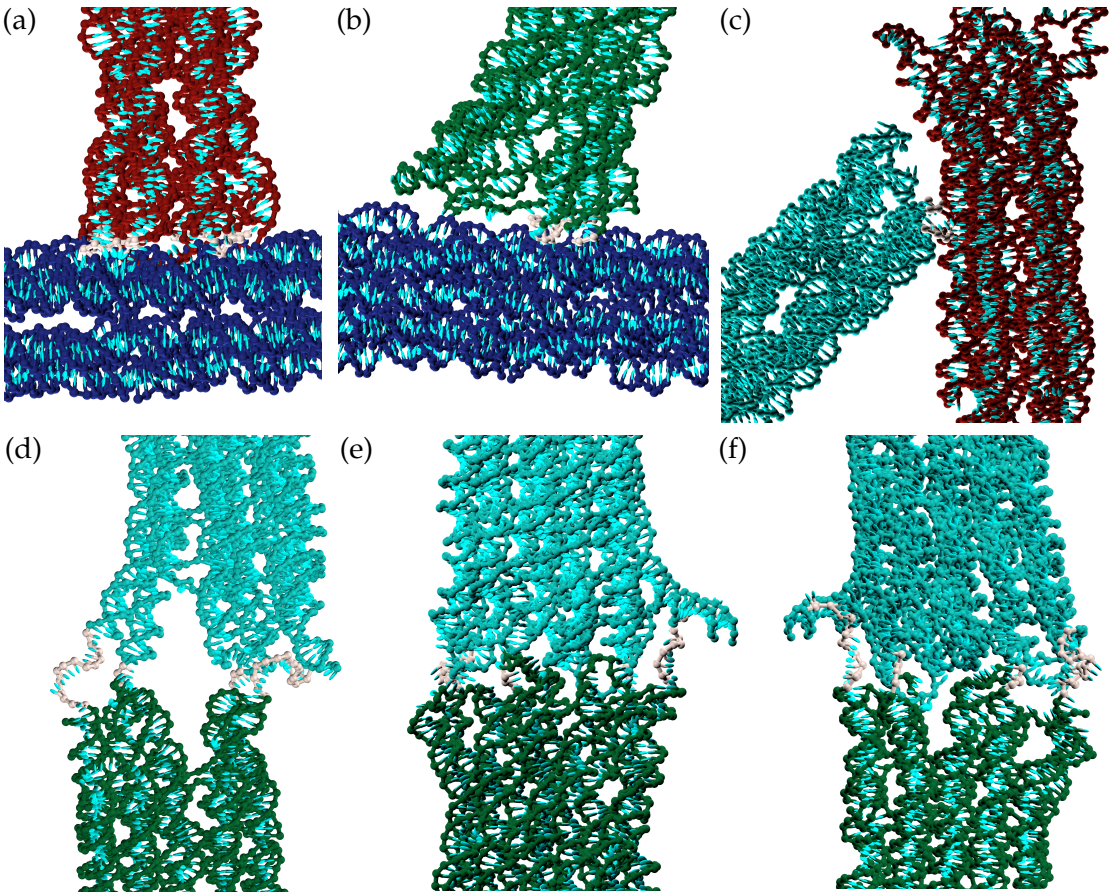


Figure 4.5: Close-ups of each joint for the configurations illustrated in Figure 4.4. Specifically, these are for (a) Joint 1 (b) Joint 2 (c) Joint 3 (d) Joint 4 (part 1) (e) Joint 4 (part 2, view from outside) (f) Joint 4 (part 2, view from inside). The outside and inside of joint 4 are defined relative to the quadrilateral formed by the four blocks. Single-stranded sections of the scaffold that link the blocks are coloured grey. (d) and (e)/(f) clearly show the difference in steric repulsion between the helices near the joint for open and closed configurations of joint 4.

The free-energy contribution of joint 2 is a broad minimum centred around the preferred opening angle of the joint (Figure 4.6), and so adding this contribution does not significantly alter the general shape of the landscape. It slightly lowers the barrier and penalises configurations with more extreme values of θ, particularly those with $\theta > 90°$, because of the steric repulsion between the crank and frame blocks at these angles.

Further including the contribution from joint 3 results in an appreciable change in the shape of the landscape. Most noticeably, the minimum corresponding to state S1 becomes a flat region with a higher free energy. The contribution of joint

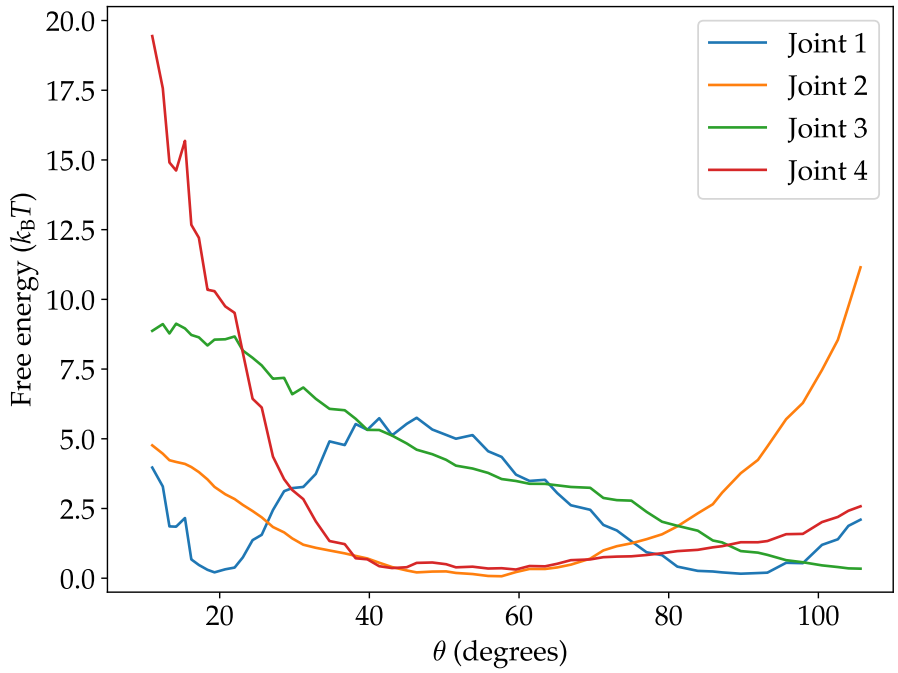


Figure 4.6: Free-energy contributions of each joint to that of the complete origami as a function of θ.

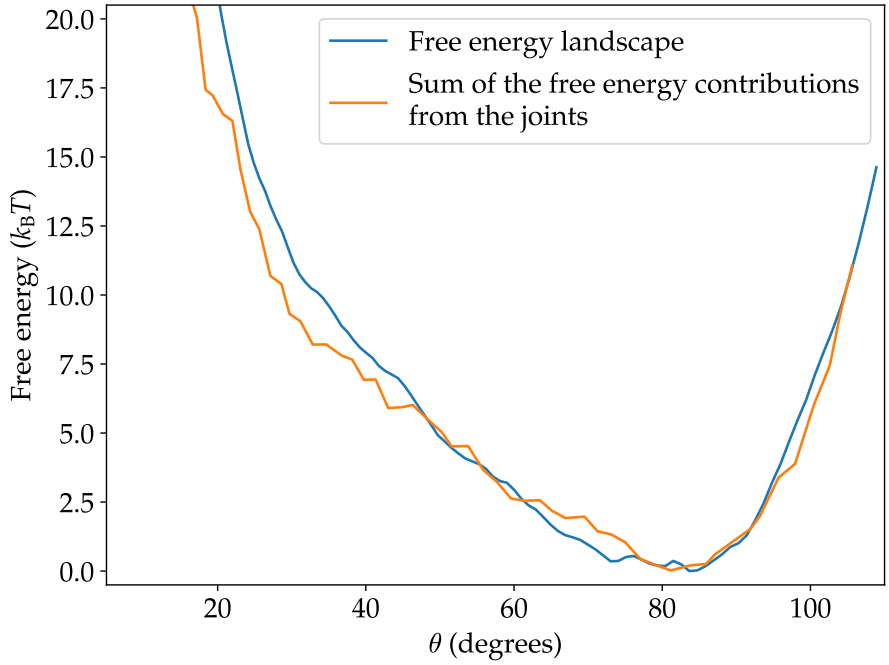


Figure 4.7: Comparison of the free-energy landscape and the sum of the contributions from the joints.

3 has a more significant effect than joint 2 because the structure of the origami necessitates a smaller opening angle at joint 3 than joint 2. The small opening

angle and its associated free-energy cost lead to the destabilisation of state S1.

Finally, including the contribution from joint 4 completely removes the second minimum, because the inward bending of this joint, which is associated with the closed configuration, leads to a significantly larger free-energy cost.

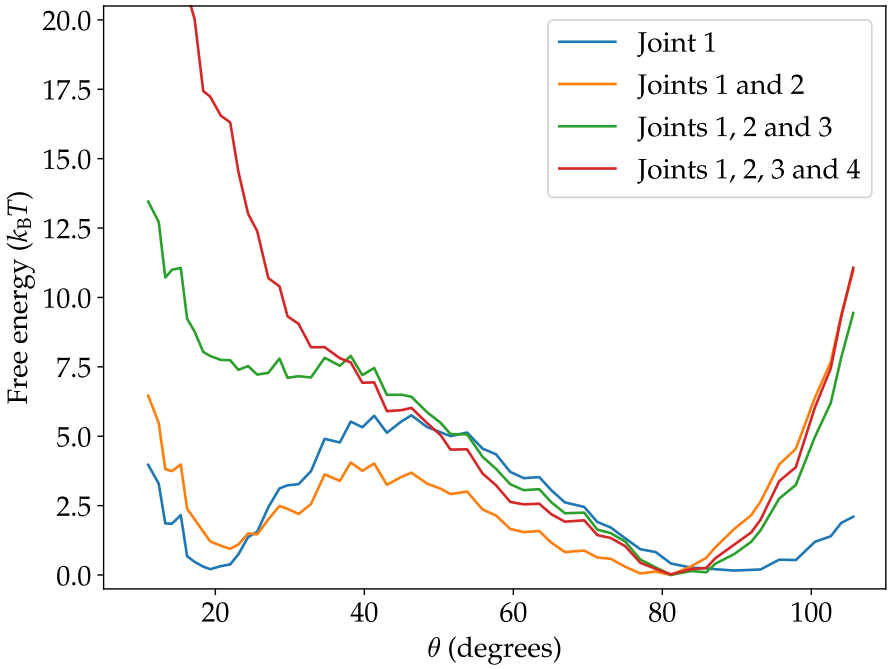


Figure 4.8: The free-energy contributions of each joint to the overall free-energy landscape.

4.5 Methods

4.5.1 Free energy landscape

We used umbrella sampling [92] to calculate the free energy landscape as a function of the order parameter R, defined in Figure 4.2. Specifically, R was defined as the distance between the centres of mass of two groups of nucleotides, which consist of the nucleotides at positions 259–273 on helices 0–9, and at positions 231–245 on helices 10–19, respectively. To obtain the landscape as a function of θ, the values of R and θ were tracked in each window of the umbrella sampling simulations. θ was defined as the angle between the two vectors in Figure 4.2. Specifically, the first vector was defined as the vector pointing from the centre of mass of the nucleotides

at positions 196–210 to that at positions 259–273 on helices 0–9, whereas the second vector was defined as the vector pointing from the centre of mass of the nucleotides at positions 147–161 to that at positions 231–245 on helices 10–19. The nucleotide positions and helix indices are defined in the caDNAno design in Figure A.2.

We performed sets of simulations where the order parameter was restrained with a harmonic potential at the centre of each sampling window. The sampling pathway ranges from $R = 12\,\text{nm}$ to $60\,\text{nm}$.

Before the umbrella sampling simulations, we prepared the starting configurations for each window using two pulling simulations. The order parameter of an equilibrated configuration was restrained at its original value (46 nm) with a harmonic bias potential of stiffness $k = 57.09\,\text{pN/nm}$.

In the first simulation, the equilibrium position of the harmonic bias potential was gradually reduced at a constant rate of $0.14\,\text{m/s}$, until it reaches 12 nm. Configurations were outputted every 2×10^6 steps, i.e. when the traps have moved by 0.85 nm, resulting in a starting configuration for each window from the starting value of R_{ee} to the final value of R_{ee}, where neighbouring windows are separated by 0.85 nm.

In the second simulation, the equilibrium position was gradually increased at the same rate as the first simulation, until it reaches 60 nm. Configurations were outputted at the same frequency, which gives starting configurations for windows defined in a similar way as in the first simulation.

After the starting configurations have been prepared, the configuration in each window was restrained at the value of R_{ee} corresponding to that window with a harmonic bias potential of stiffness $k = 11.42\,\text{pN/nm}$. Each window was equilibrated for 10^6 steps before a production run of 10^7 steps. R was outputted every 10^3 steps, giving 10^4 data points for each window. Using the biased probability distributions of R in each window, we used WHAM [93, 136] to calculate the unbiased free energy landscape of the system as a function of R.

A new set of production runs was then started from the last configurations of the previous production runs, and WHAM was performed on the new data points

to calculate the free-energy landscape. This process was repeated three times. The resulting landscapes were very similar, thus confirming convergence.

In each umbrella sampling window, the value of θ was also recorded every 10^3 steps. Using the corresponding values of R, a quadratic relationship between R and θ was fitted and used to map the x-axis of the free-energy landscape from R to θ. The fitted relationship is $\theta = 0.015R^2 + 1.08R - 7.74$, with the coefficient of determination of the fit being 0.99. The data and the fit is shown in Figure 4.9.

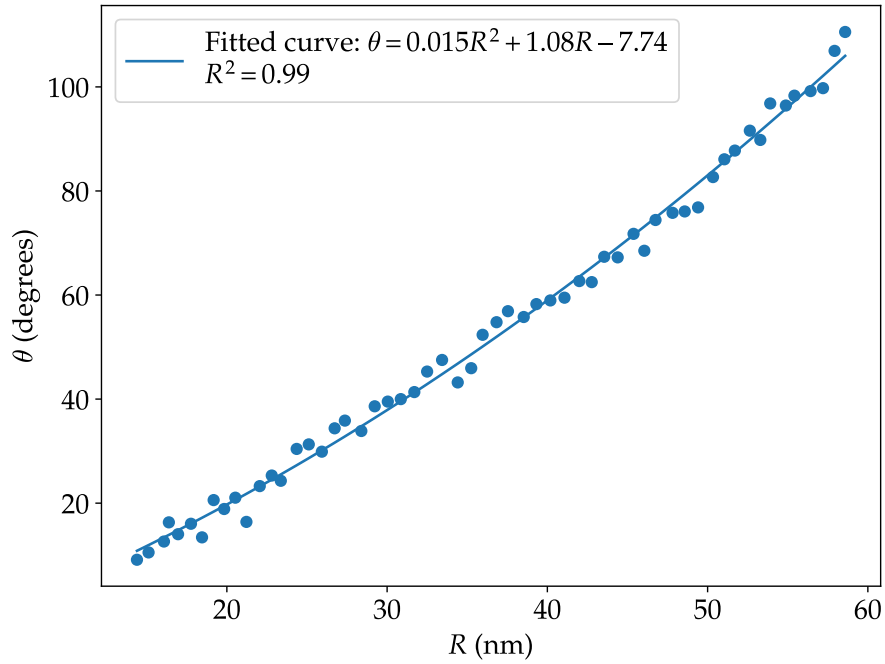


Figure 4.9: The mapping between the distance order parameter R and θ. The data points represent the average values of both for each umbrella sampling window and the line is a fit to this data.

4.5.2 Free energy decomposition

We decomposed the free-energy landscape into four contributions associated with each of the four joints.

For joint 1, the origami design was modified from the original design by removing the nucleotides corresponding to the coupler and the crank blocks. The order parameter R_1 was defined as the distance between the centres of mass of two groups of nucleotides, which consist of the nucleotides at positions 147–161 on helices 10–19,

and at positions 539–553 on helices 10–19, respectively. The nucleotide positions and helix indices are defined in the caDNAno design in Figure A.3.

For joint 2, the origami design was modified from the original design by removing the nucleotides corresponding to the coupler block. The order parameter R_2 was defined as the distance between the centres of mass of two groups of nucleotides, which consist of the nucleotides at positions 259–273 on helices 0–9, and at positions 231–245 on helices 10–19, respectively. The nucleotide positions and helix indices are defined in the caDNAno design in Figure A.4.

For joint 3, the origami design was modified from the original design by removing the nucleotides corresponding to the crank block. The order parameter R_3 was defined as the distance between the centres of mass of two groups of nucleotides, which consist of the nucleotides at positions 413–427 on helices 0–9, and at positions 413–427 on helices 10–19, respectively. The nucleotide positions and helix indices are defined in the caDNAno design in Figure A.5.

For joint 4, the origami design was modified from the original design by removing the nucleotides corresponding to the frame and compliant blocks. The order parameter R_4 was defined as the distance between the centres of mass of two groups of nucleotides, which consist of the nucleotides at positions 196–210 on helices 0–9, and at positions 539–553 on helices 0–9, respectively. The nucleotide positions and helix indices are defined in the caDNAno design in Figure A.6. Part 1 of the landscape corresponds to the part where the joint points outwards and resembles the origami in state S2, while part 2 corresponds to the part where the joint is folded inwards and resembles the origami in state S1. To calculate each sub-landscape, we did not explicitly constrain the joint to point inwards or outwards. Instead, the portion of the landscape sampled was determined by the initial configurations, since for R_4 values corresponding to a bent joint, transitions between the two sub-landscapes are not feasible in the umbrella sampling simulations. For R_4 simulations corresponding to straight or stretched joints, the two sets of simulations sample equivalent configurations.

For each joint, we performed a set of simulations where the order parameter was restrained with a harmonic potential in each sampling window. The sampling range of the order parameters was from 76–94 nm for joint 1, 12–60 nm for joint 2, 12–76 nm for joint 3, 68–86 nm for joint 4 (part 1), and 60–86 nm for joint 4 (part 2).

Umbrella sampling simulations were performed in a similar manner as those for the complete origami, with windows separated by 0.85 nm and the configuration in each window being restrained with a harmonic potential of stiffness $k = 11.42\,\mathrm{pN/nm}$. Each window was equilibrated for 10^6 steps before a production run of 10^7 steps. R was outputted every 10^3 steps, giving 10^4 data points for each window. Using the biased probability distributions of R in each window, we used WHAM [93, 136] to calculate the unbiased free energy landscape of the system as a function of R. The production was repeated three times and the landscapes obtained were similar.

To enable the transformation of the landscapes of the individual joints to be a function of θ, the values of R_i, the distance order parameters for the joints, were also measured in the umbrella sampling simulations of the complete origami. Figure 4.10 shows their variation with θ. The maxima in R_1 and R_4 occur when the crank and coupler blocks are collinear. The two parts of the joint 4 landscape correspond to different ranges of θ with a small region of overlap corresponding to when the crank and coupler blocks are approximately collinear. To generate a single landscape as a function of θ for joint 4, part 1 was used for $\theta > 53^\circ$ and part 2 for $\theta < 53^\circ$ and the two landscapes matched at that point.

4.6 Conclusions

We have used the oxDNA model to compute the free-energy landscape of a jointed multi-block DNA origami that was designed to have two stable states. Somewhat surprisingly, the computed landscape shows a single minimum corresponding to the open configuration. These results present a substantially different picture of the landscape in contrast to that inferred from experiments where a significant population of origami in the closed state was observed.

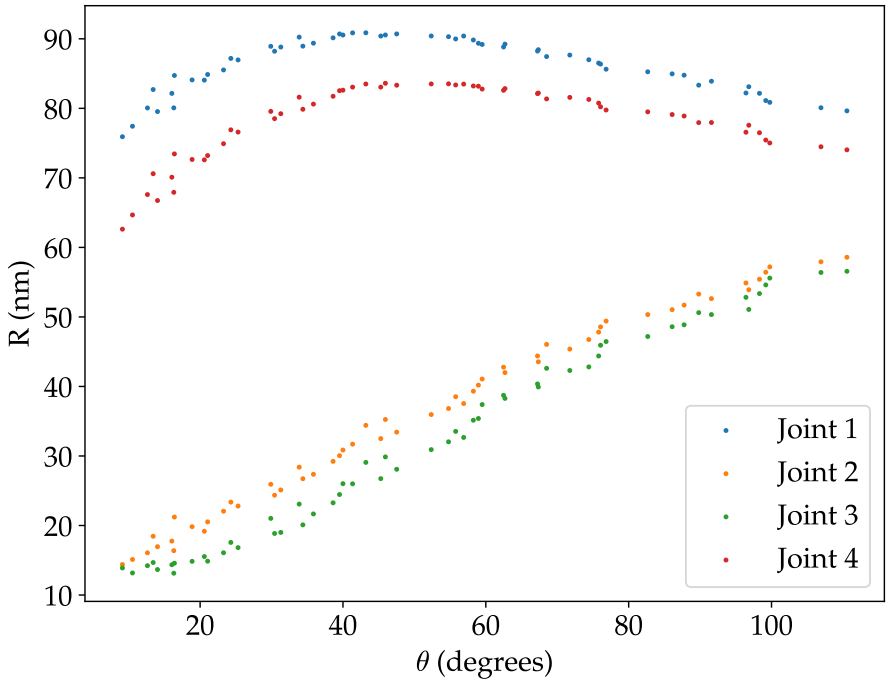


Figure 4.10: The mapping between the distance order parameters of the individual joints R_i and θ. The data points represent the average values for each umbrella sampling window of the complete origami.

The designed stress accumulation mechanism, which involves the bending of the compliant block associated with intermediate configurations, leads to a bistable landscape associated with joint 1. However, in simulations, the other three joints do not behave as completely free hinges, causing the destabilisation of the closed state, with the small opening angles at joint 3 and the inward bending of joint 4 being particularly unfavourable.

The increase in the free energy of the hinge joints as their joint angles move away from their preferred values is due to steric repulsion and the constraints increasingly placed on the configuration of the helices and linkers near the joint. These effects are consistent with previous oxDNA results on jointed DNA nanostructures [80, 147] which have also found narrower distributions for the flexible components than those observed in electron microscopy images. In other words, joints in oxDNA seem to be less flexible than inferred from experiment.

To account for the seeming disagreement between simulations and experiment, one possibility is simply that the oxDNA model is deficient in its description of these

structures. However, the behaviour of the joints is a reflection of the structural and mechanical properties of the helices and single-stranded DNA at the joints, both of which are well-described by oxDNA. Moreover, oxDNA predicts that the closed state of the origami is $20\,k_BT$ higher in free energy than the open state, which is not a difference that can be attributed to a minor deficiency.

Another possible source of difference is the heterogeneity in the ensemble of assembled origamis in experiment. While simulations model a defect-free origami, there are likely to be a range of assembly defects in the experimental origamis. These assembly defects probably not only include missing staples, but also defects in the routing of the strands. The latter may be particularly relevant to multi-block origamis, such as those studied here, which possess many scaffold links between the blocks. For example, if the blocks nucleate and assemble independently [149], then if the assembly of one block happens to lead to the topological entanglement of the external scaffold loops, such a defect is likely to be locked in. Furthermore, the assembly of further blocks might push these defects to the joints, potentially substantially changing the properties of the joint. Although there has been some studies on the quantification of missing staples [150–152], there is little information about other possible defects. Thus, it is conceivable that the population of closed configurations corresponds to origami that are locked into this structure due to assembly defects or that have more flexible joints due to missing staples.

There is some evidence of this heterogeneity in the experimental results. For example, it is not clear why there was a small population of U states even though such structures were designed to be at a maximum on the free-energy landscape. Again, it is possible that they might be locked in to these configurations by defects. Furthermore, upon the addition of actuating strands which lead to binding between the frame and crank blocks, and the coupler and compliant blocks, there are some open configurations that do not close, which again suggests the presence of certain defects preventing the closing of the origami.

A possible way to quantify the heterogeneity in an ensemble of flexible origamis is to use single-molecule techniques which can follow the large-scale dynamical

motion of individual origamis. A recent study utilised single-molecular FRET experiments on a two-block DNA origami hinge to observe fluctuations between open and closed states driven by the hybridisation and melting of complementary strands attached to each block [153]. However, it is worth noting that only a relatively small sub-population of the origamis, ranging between 5% to 40% across measurements for different hinge designs, were observed to undergo such fluctuations. Simulations could also be used to explore the effects of different types of defects on origami properties.

5

Self-Limiting Polymerisation of a DNA Origami

Contents

5.1 Introduction 66
5.2 Origami Design 69
5.3 Results and Discussion 71
5.4 Methods 94
5.5 Conclusions 100

This work is done in collaboration with Jonathan F. Berengut and Dr. Lawrence K. Lee. They performed the experimental study of the work involved in this chapter. The results of our work have been published in ACS Nano [114].

5.1 Introduction

Although DNA origami has become an indispensable technique in the field of DNA nanotechnology, there is a fundamental size limit of a single origami imposed by the length of the scaffold strand, typically the M13mp18 bacteriophage genome [4], a single-stranded DNA with 7249 nucleotides. To overcome this size limit, longer scaffold strands have been synthesised and used to assemble larger origamis [154, 155], but a longer scaffold strand necessitates more unique staple strands during

origami assembly, thus leading to higher synthesis costs. It is therefore necessary to pursue other methods to build larger DNA nanostructures, such as hierarchical assembly of structures from multiple origami subunits [156–158].

The hierarchical assembly of DNA origami subunits is generally mediated by two types of interactions. The first type is the hybridisation of complementary single-stranded DNA at certain positions on the origami [5, 159–168]. These complementary strands are often called "sticky ends" because of the cohesive role they play during the assembly process. The second type is the blunt-ended stacking of DNA duplexes [169–174]. Unlike sticky ends, blunt-ended stacking is not sequence specific. Complementary shapes of the stacking sites are used to introduce specificity in the assembly process. These two types of interactions are not mutually exclusive. Indeed, several studies have utilised both during the assembly process [175–180].

A common target is to assemble lattices of DNA origamis. Two-dimensional building blocks, such as rectangular or square origami sheets or tiles, can assemble to form two-dimensional lattices [161, 162, 165–168, 176, 178], while three-dimensional origami building blocks can assemble to form two- or three-dimensional lattices [160, 163, 170, 173].

In addition to lattice structures, it is also an interesting target to assemble finite structures. In fact, most self-assembled biological structures are finite structures instead of infinite lattices [115]. Examples include closed structures, such as bacterial microcompartments [181, 182] and viral capsids [183], or open structures, such as cytoskeletal filaments and microtubules [184, 185]. Although these structures are finite, they are much larger than the protein subunits from which they are constructed [115]. It would be interesting to mimic how these structures are assembled from proteins and, in a similar manner, fabricate artificial materials from the assembly of DNA origami building blocks.

There are multiple approaches for the assembly of finite structures from DNA origamis. One example is to construct closed and symmetrical structures, such as rings or polyhedra, from identical origami subunits with a specifically designed

geometry [5, 159, 164, 169, 171, 174]. As the subunits assemble and all their interaction sites become bound, the target structure is naturally formed.

A more challenging target is to assemble open finite structures with well-defined dimensions. Without any growth-termination mechanisms, the assembly of identical origami subunits will only cease after all subunits have been depleted. This results in assembly products with a large variation in their dimensions. To limit the size of the assembly product, a number of solutions have been proposed.

A straightforward method is to use distinct subunits instead of identical subunits, so that each distinct subunit has a specific location within the assembly product [159, 162, 174–177, 179, 180]. Assembly naturally terminates at the subunits which are designated to be at the terminating positions. An immediately obvious drawback of this method is the need for more distinct subunits for larger assemblies, which entails higher costs due to more raw materials needed. Another challenge is to minimise unwanted interactions between subunits not designed to interact, which becomes increasingly difficult with more distinct subunits [177].

Templated self-assembly could potentially circumvent the need for distinct subunits [186, 187]. Subunits do not directly interact with each other but only interact with the assembly template. Assembly terminates when the whole template is completed with subunits. This method reduces the number of distinct components involved in the assembly process, thereby reducing synthesis costs, but the size of the assembly is limited by the size of the template, which in turn is limited by the size of a single DNA origami.

Ultimately, a hierarchical assembly with only one component is desirable to avoid the limitations of multi-component assembly described above. Hereafter, we shall refer to one-component hierarchical assembly as polymerisation, since they can be considered as equivalent in this particular context. With only one component, the polymerisation termination mechanism must be embodied in the component in order to allow control over the polymer size.

Here, we explore the possibility of using “displacement accumulation” to accomplish self-limiting polymerisation. Although this possibility was proposed as far back

as 1959 [188], it has not been previously utilised in DNA origamis. Therefore, we seek to design a DNA origami that employs an allosteric displacement mechanism to achieve stress accumulation, with the eventual target of realising self-limiting polymerisation. We also explored the variation of the polymer length distribution upon tuning the strength of the allosteric effect.

Although experiments allow the direct observation of equilibrium polymer length distributions, it is challenging to achieve a detailed understanding and verification of the mechanism of self-limiting polymerisation through experiments. Therefore, to complement experimental studies and interpret experimental results, we performed computer simulations of the origami monomers and polymers to study their structural properties.

We will use the oxDNA model [47–49] to perform coarse-grained molecular dynamics simulations. As described in Chapter 1, oxDNA provides a good representation of the structures of DNA origamis [49, 75], particularly flexible origamis [80, 82, 142, 147, 189] that will be the focus of this study. We will first study the structural properties of the origami monomers and polymers near equilibrium. Then, we will use umbrella sampling simulations and pulling simulations to further characterise the stress accumulation mechanism in detail and parametrise a simple stress model. Using the model, we will perform MC simulations to predict the structural properties of the polymers and evaluate how well the model can describe the polymers. Finally, we will explore the thermodynamics of polymerisation by running pulling simulations.

5.2 Origami Design

Our goal is to design a DNA origami that undergoes self-limiting polymerisation with stress accumulation through the cumulative displacement of a mobile domain. The design of the origami is shown in Figure 5.1.

The origami monomer consists of three connected blocks (Figure 5.1A-C). Each block has a length of 84 base pairs (approximately 29 nm). Each of the two flanking

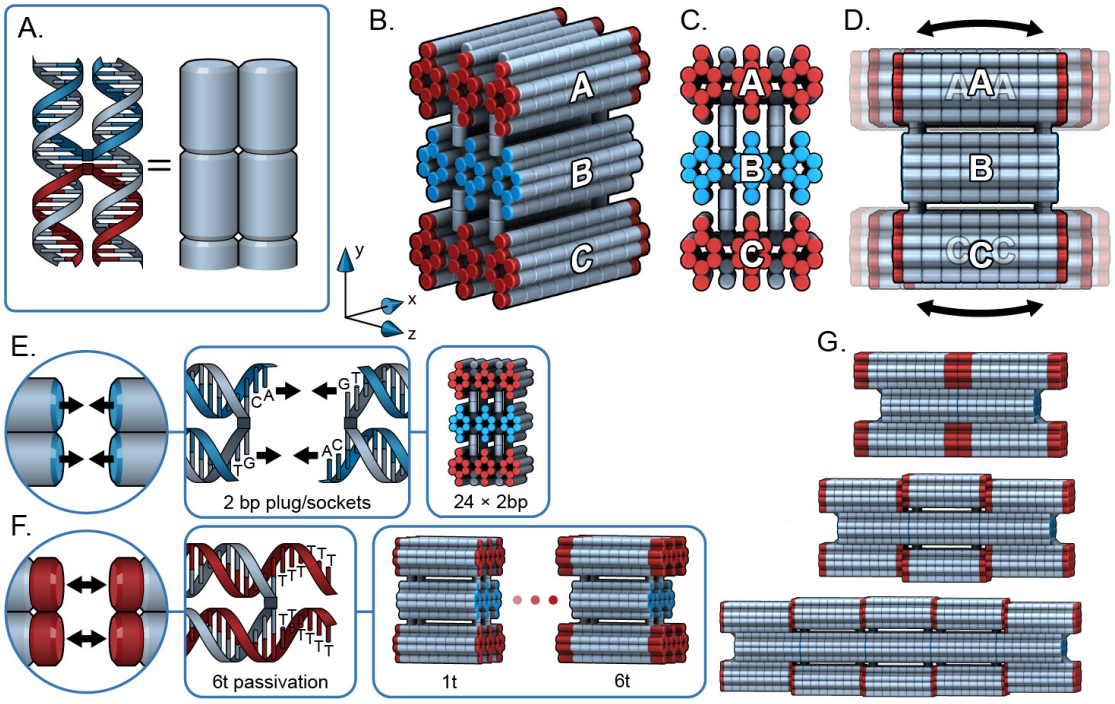


Figure 5.1: DNA origami design. (A) Each cylinder represents a DNA double helix. (B) The origami is composed of three connected domains, labelled A, B and C. Domain B, coloured in blue, has binding interfaces at the front and back, where the front is complementary to the back to allow polymerisation. Domains A and C, coloured in red, have passivated interfaces that do not bind. (C) Front view of the origami. The domains are connected with flexible linkers, shown as vertical cylinders. (D) The flexible linkers allow domains A and C to move relative to domain B in an approximately linear sliding motion. (E) The ends of domain B contain 2-nucleotide extensions (plugs) or truncations (sockets) that can bind through base-pairing interactions. There are 24 helices modified in this manner. (F) The ends of domains A and C are passivated by poly-thymine (poly-T) extensions. The size of these domains, and thus the rate of stress accumulation, can be adjusted by changing the length of the extensions from 1T to 6T. (G) Schematic diagram illustrating how longer poly-T extensions result in a larger displacement of domains A and C, and hence a larger stress free energy per monomer. Altering the length of poly-T extensions thus allows control of the mean polymer length. Monomers with longer poly-T extensions (top) should produce shorter polymers than those with shorter poly-T extensions (bottom). Figure kindly provided by Jonathan F. Berengut.

blocks are connected to the middle block by four linkers in the form of a 21 base-pair DNA duplex (approximately 7 nm). The flexible linkers allow the flanking blocks to move relative to the middle block by translating along the helical axis, or equivalently, the polymerisation axis (Figure 5.1D).

The staple strands at the ends of the middle block contain 2 base-pair extensions or truncations so that one end of the block forms a complementary binding site to the other end of the block, thereby allowing polymerisation (Figure 5.1E). There are 24 binding sites in total on one end of the middle block.

The staple strands at the ends of the two flanking blocks are passivated with poly-thymine (poly-T) extensions, which introduces repulsion between flanking blocks of neighbouring monomer units in a polymer, and makes the flanking blocks effectively longer than the middle block (Figure 5.1F). Due to these two effects, binding between middle blocks results in the displacement of the flanking blocks (Figure 5.1G). The magnitude of the displacement is expected to increase with the length of the polymer. By changing the length of the poly-T extensions, the displacement of the flanking blocks can be adjusted.

We studied three different lengths of the poly-T extensions, namely 1, 3 and 6. Throughout this chapter, we will denote the origamis by the length of their poly-T extensions, such as 1T, 3T and 6T. The complete caDNAno design of the origami can be found in the Appendix.

5.3 Results and Discussion

5.3.1 Monomer Structure

We first simulated a monomer to understand its basic structural properties. The average structure of a 3T monomer is shown in Figure 5.2. The structure is apparently asymmetric, with one end of each flanking domain being further from the middle domain than the other end.

To further characterise the structure, we calculated the translational and angular displacement distributions along the polymerisation axis from the configurations in the simulation trajectory. The distributions are plotted in Figure 5.3. These

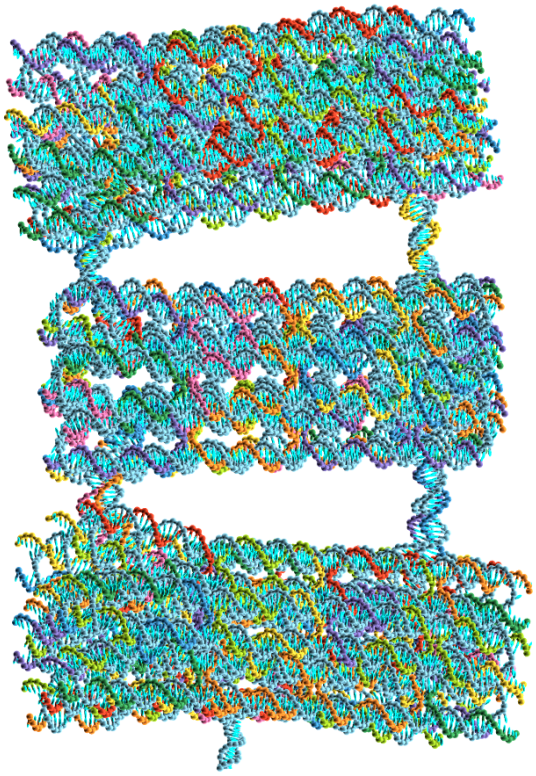


Figure 5.2: Average structure of a monomer. The structure is asymmetric, with the flanking domains tilting away from the middle domain, forming an arrow-like shape.

distributions agree with the mean structure. The translational distributions show that the flanking domains have a non-zero average displacement relative to the middle domain. Both flanking domains are displaced towards the same direction. Meanwhile, the angular distributions show that the flanking domains have a non-zero average bending angle relative to the middle domain. This implies that the flanking domains are rarely aligned with the middle domain. The two flanking domains have opposite angular displacements, resulting in an arrow-like shape of the whole monomer.

To understand the source of this asymmetry, we visually inspected the configurations from the simulation. Interestingly, we found that the linking duplexes between the middle and flanking domains have a strong tendency to form base stacking interactions with the duplexes at the end of those domains where the linkers are anchored. As a result of these stacking interactions, the anchoring duplexes tend to bend out of the domain to facilitate stacking. The effective length of a linking duplex

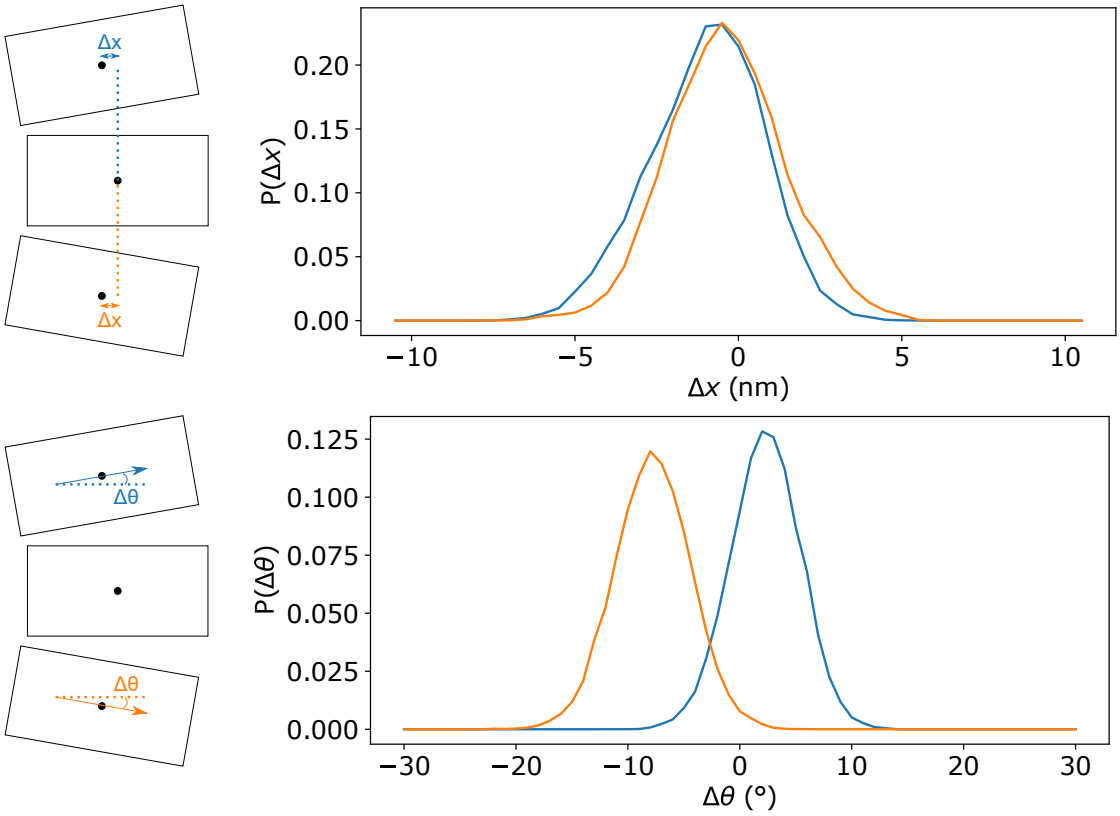


Figure 5.3: Translational and angular displacement distributions of the flanking domains of the monomer. Both domains have a non-zero and slightly negative average displacement relative to the middle domain. The angular displacements of the two domains are also non-zero but have opposite signs.

is thus extended by the length of the anchoring duplex that can bend out. Such bending is, however, constrained by the nearest crossover involving the anchoring duplex. The effective length extension is therefore limited by the distance from the end of the anchoring duplex to the nearest crossover. Inspection of the caDNAno design shows that there is a difference in this distance between the two ends of the monomer along the polymerisation axis. The distance is shorter at the beginning of the polymerisation axis, and larger at the end. Figure 5.4 shows the relevant part of the caDNAno design. This difference explains the asymmetry of the monomer, as the linkers are effectively shorter on one side of the monomer than on the other side.

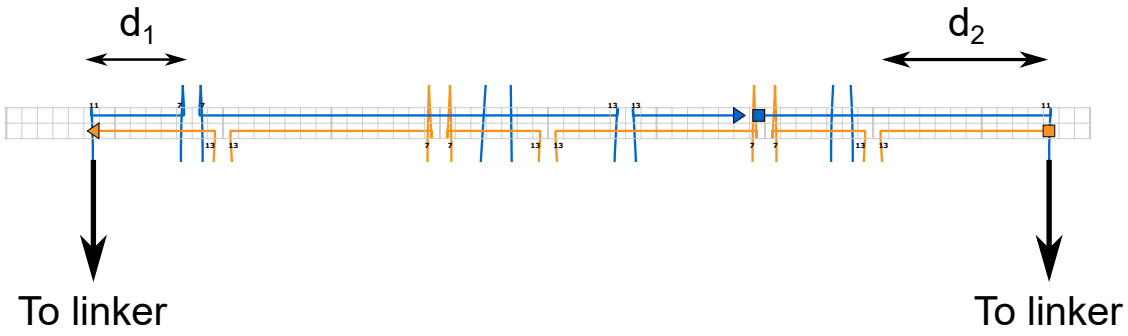


Figure 5.4: Part of the design represented in caDNAno [20], which shows the distances to the nearest crossover from the ends of the anchoring duplex. $d_1 < d_2$ results in an asymmetry in the effective linker lengths, and thus the asymmetry in the monomer structure.

5.3.2 Polymer structure

With the monomer being characterised, we then simulated 1T, 3T and 6T polymers of length 2 to 9. Representative structures for 3T polymers of length 2, 5 and 9 are shown in Figure 5.5.

For the shorter polymers, the stress accumulation mechanism largely remains as designed, i.e. through the displacement of the flanking domains along the polymerisation axis. However, for longer polymers, in addition to a larger displacement of the flanking domains, other modes of stress accumulation are observed. For example, some flanking blocks "pile up" on each other, as shown in the 9-mer configuration where the upper yellow flanking block exhibits a particularly large value of $\Delta\theta$ allowing it to sit on top of the adjacent red flanking block. This piling up effectively allows the displacement of the flanking block along the polymerisation axis to be reset, thus releasing some of the stress built up in the polymer. A twist can also develop in the polymer, as evident from the 9-mer configuration where the monomers are twisted and not perfectly aligned along the polymerisation axis. This twist could reduce the repulsion between neighbouring flanking blocks as the blocks then do not directly face each other.

To further characterise the structure of the polymers, we calculated the translational displacement distributions along the polymerisation axis from the configurations in the simulation trajectory. The distributions were calculated for each

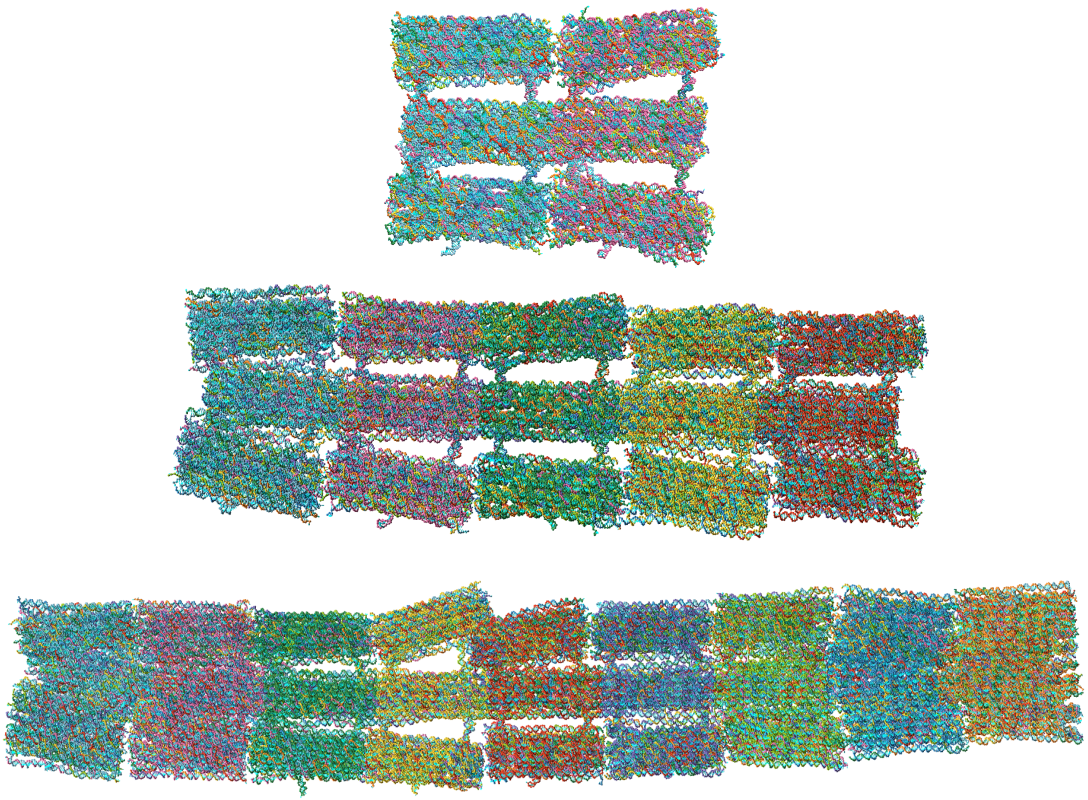


Figure 5.5: Representative configurations of 3T polymers of length 2, 5 and 9. The first two polymers accumulate stress largely as designed, while the third, longest polymer shows other stress accumulation mechanisms, such as twisting along the polymerisation axis and piling of flanking blocks.

monomer position separately. Figure 5.6 and Figure 5.7 show the distributions for the polymers shown in Figure 5.5.

For short polymers, the distributions confirm the stress accumulation mechanism. The flanking blocks are displaced to different extents depending on their position within the polymer, which determines the stress they experience. Flanking blocks towards the end of a polymer are the most stressed and displaced to the largest extent. As for longer polymers, the stress release mechanism observed in the 9-mer configuration in Figure 5.5 is also evident in the corresponding bottom plot in Figure 5.6. The displacements of the top flanking blocks of the three middle monomer units are all centred at near 0, illustrating that stress is not accumulating as designed but released through other mechanisms, which in this case involves flanking blocks piling up on each other. The bottom plot in Figure 5.7 also shows some asymmetry

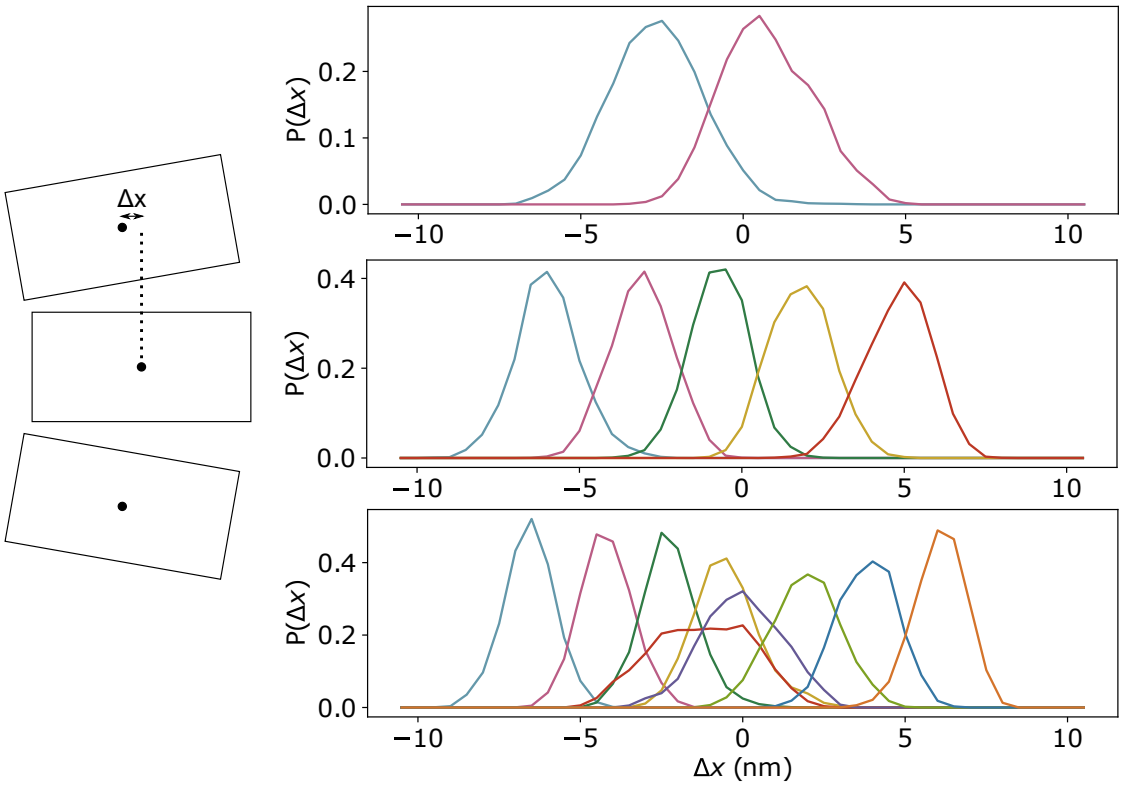


Figure 5.6: Translational displacement distributions of the upper flanking block for 3T polymers of length 2 (top), 5 (middle) and 9 (bottom). Each line represents a monomer inside the polymer. The colours of the plots follow the colours of the monomers in the snapshots. The sets of distributions for the first two polymers are consistent with the expected stress accumulation mechanism, while the set of distributions for the 9-mer shows a number of peaks centred around $\Delta x = 0$, indicating a deviation from the expected stress accumulation mechanism.

in the displacement distributions, indicating some stress release mechanism could have briefly taken place during the simulation.

To facilitate further interpretation, we also calculated the average displacement at each monomer position and plotted it against the monomer position. The plots for 1T, 3T and 6T polymers are shown in Figure 5.8 and Figure 5.9.

A common feature of the plots is the convergence of the maximum displacement magnitude to approximately 8 nm. When the maximum displacement magnitude of a polymer is below this value, the displacement increases roughly linearly with the monomer position. This corresponds to a low-stress regime where stress accumulates primarily in the designed mechanism of flanking block displacement. Once the maximum displacement magnitude is reached, which is the case for a 3T 9-mer or a

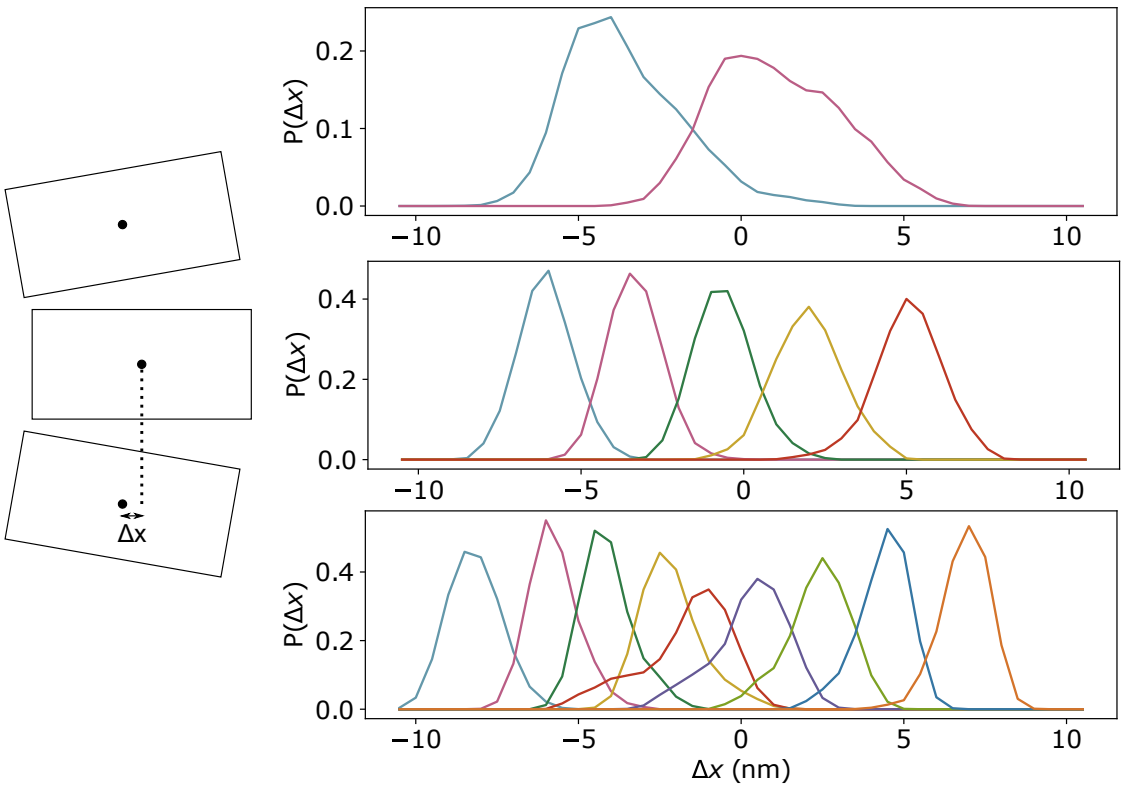


Figure 5.7: Translational displacement distributions of the lower flanking block for 3T polymers of length 2 (top), 5 (middle) and 9 (bottom). Each line represents a monomer inside the polymer. The colours of the plots follow the colours of the monomers in the snapshots. The sets of distributions for the first two polymers are consistent with the expected stress accumulation mechanism, while the set of distributions for the 9-mer shows some asymmetric peaks near $\Delta x = 0$, indicating potential deviations from the expected stress accumulation mechanism.

6T 5-mer, the plots begins to deviate from linearity. At more extreme cases, the displacement actually decreases along the polymer at certain points, for example near the middle of a 6T 9-mer. This is the sign of a high-stress regime where stress accumulation no longer obeys the designed mechanism, but instead happens by other mechanisms, such as piling of flanking blocks and twisting of the whole polymer along the polymerisation axis.

Another feature worth noting is the decreasing slope of the plots as the polymer length increases. We fitted a straight line to each of the plots and plotted the slope against the polymer length in Figure 5.10. The decreasing slope implies that the flanking blocks are packed together more tightly in longer polymers than in shorter ones. The large slope in shorter polymers can be attributed to

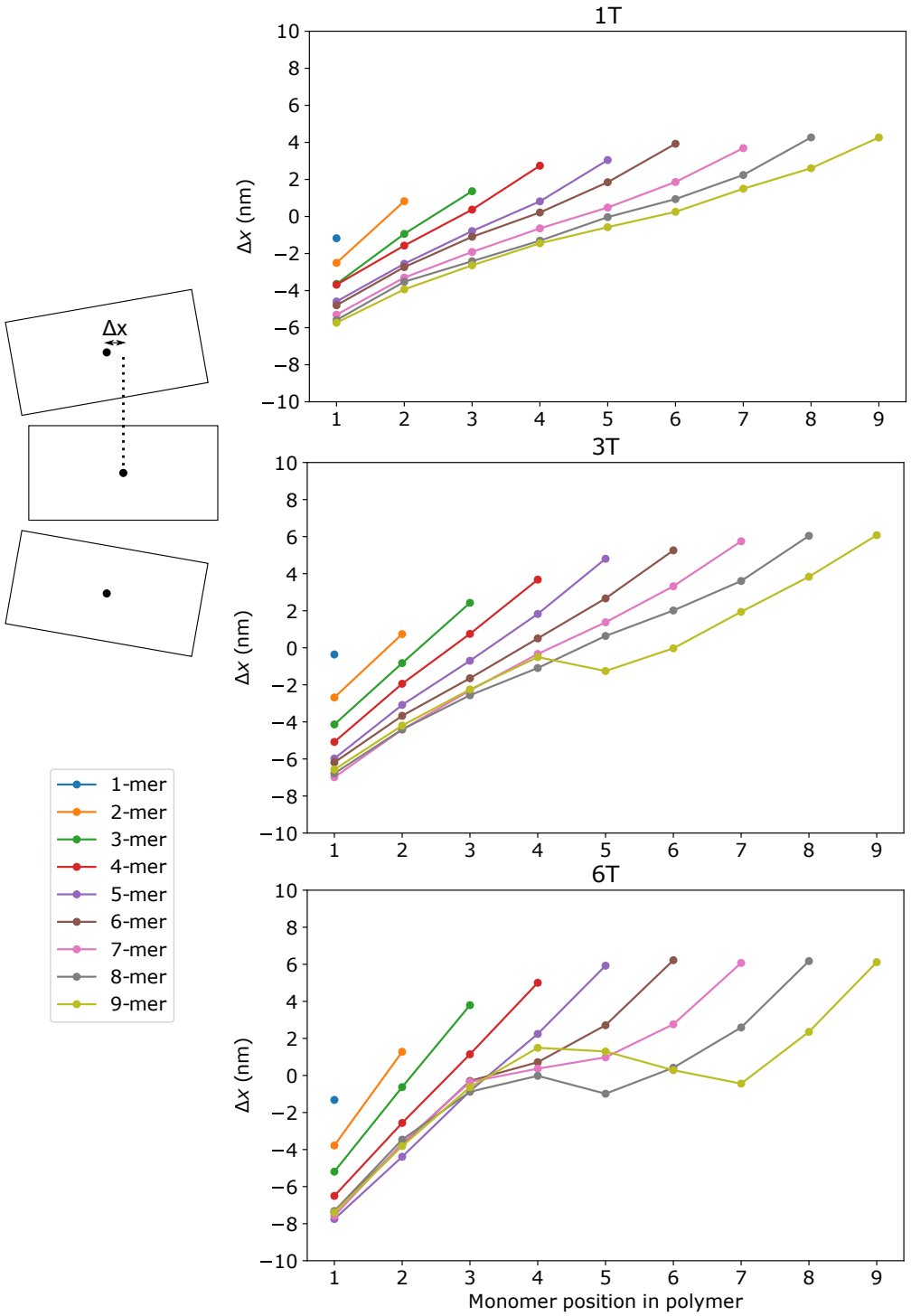


Figure 5.8: Average translational displacement of the upper flanking block for different polymer lengths and poly-T lengths. Each point corresponds to a monomer inside the polymer, while each line connects the points corresponding to a polymer of a given length. Evidence of stress-release mechanisms are shown in the longer polymers of 3T and 6T, where a decrease in Δx can be seen from left to right in the middle part of those polymers.

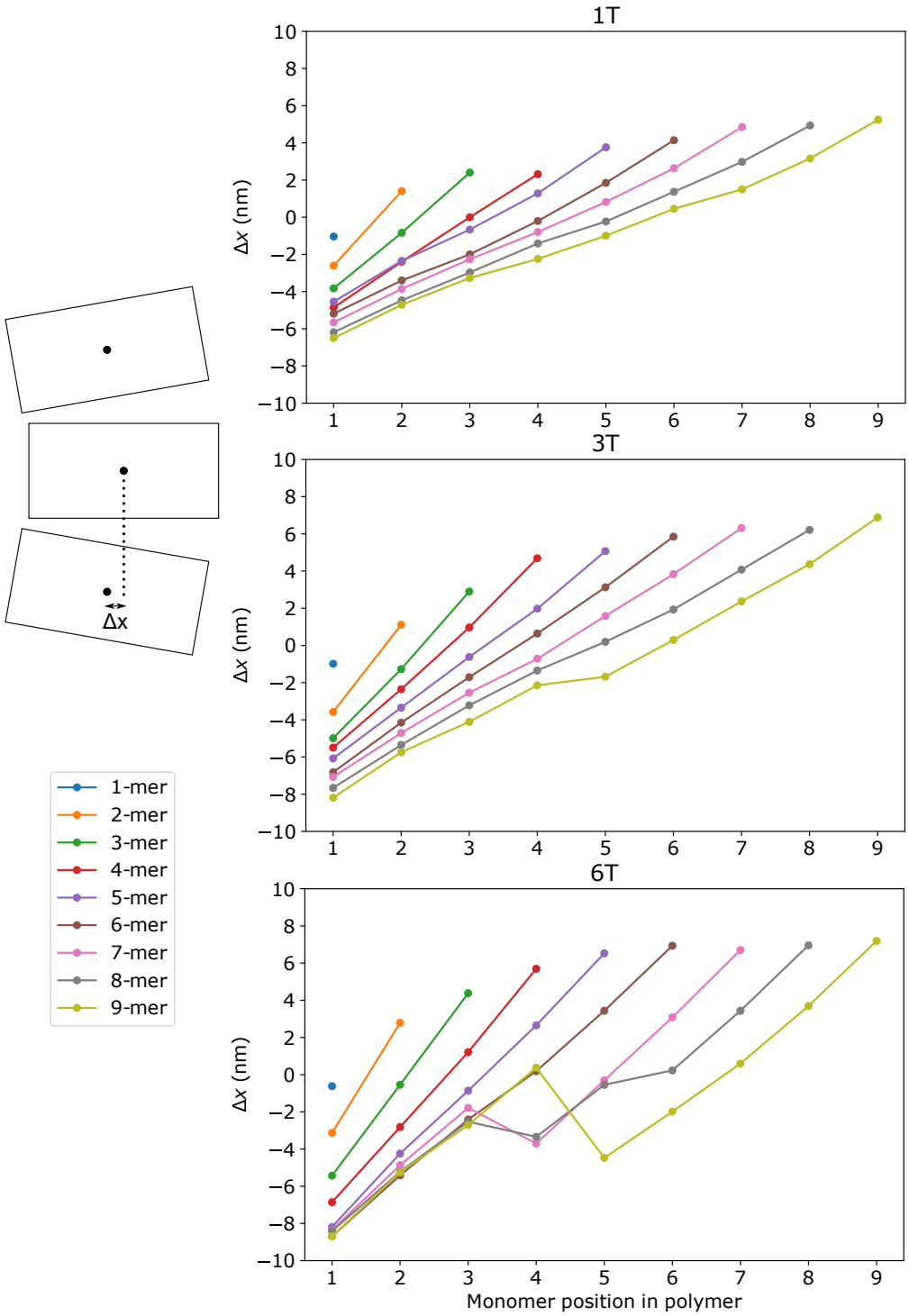


Figure 5.9: Average translational displacement of the lower flanking block for different polymer lengths and poly-T lengths. Each point corresponds to a monomer inside the polymer, while each line connects the points corresponding to a polymer of a given length. Evidence of stress-release mechanisms are shown in the longer polymers of 6T, where a decrease in Δx can be seen from left to right in the middle part of those polymers.

thermal motion of the blocks, which encourages the flanking blocks to be displaced further. However, in longer polymers, thermal motion is suppressed because of the higher total internal stress.

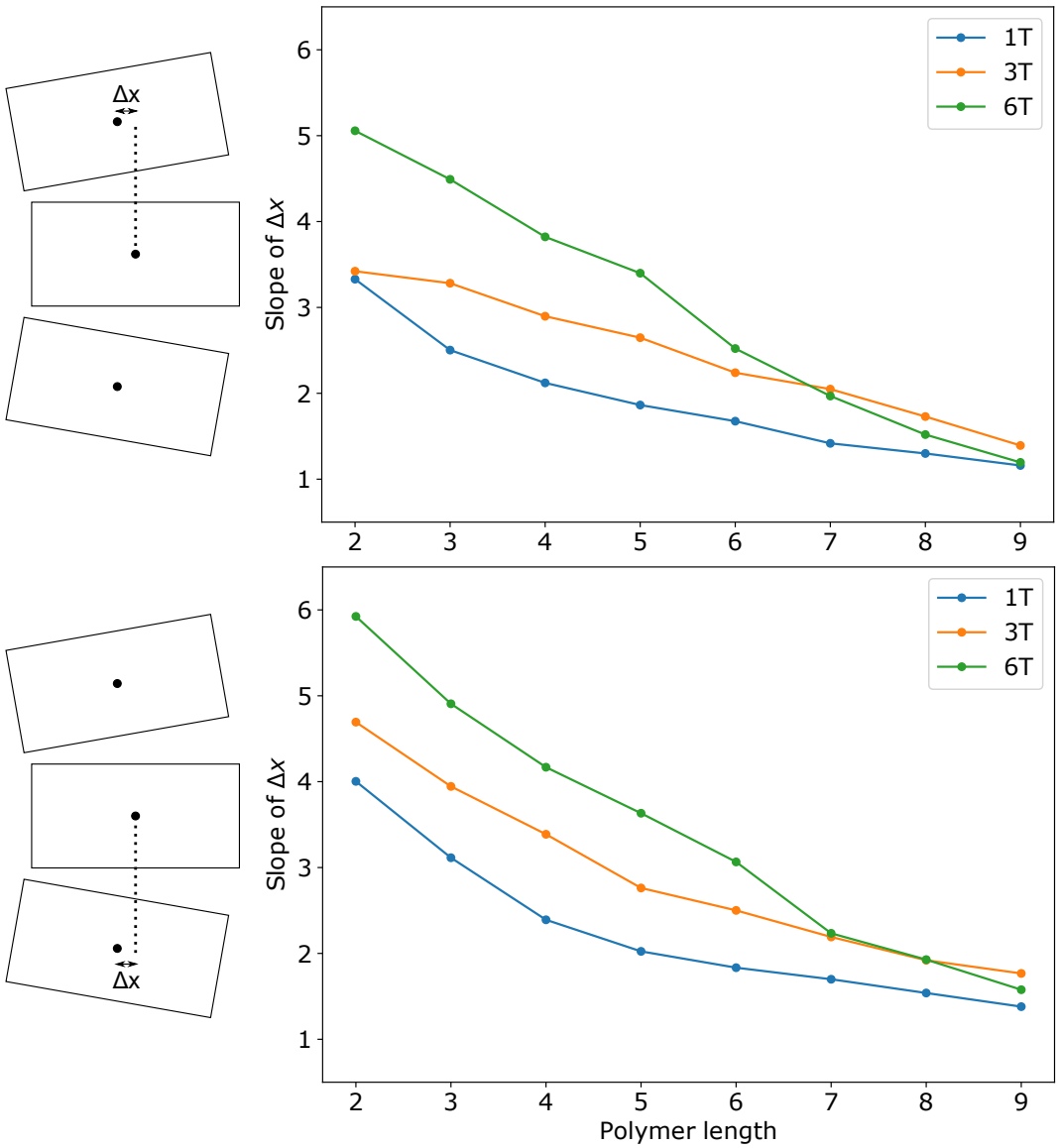


Figure 5.10: Slope of the plots in Figure 5.8 (top) and Figure 5.9 (bottom). Each point represents a polymer with a given length and poly-T length. The decrease in slope for the longer polymers indicates less thermal fluctuation.

5.3.3 Parametrising a simple stress model

To more thoroughly understand stress accumulation in polymers, we parametrised a simple stress model that captures stress accumulation (1) inside a monomer, arising from flanking block displacement, and (2) between neighbouring monomers, arising from repulsion between the poly-T extensions.

Intra-monomer stress

The intra-monomer stress accumulation was measured by calculating the free energy landscape for displacing the flanking blocks with respect to the middle block within a monomer. The free energy landscape was calculated using umbrella sampling and WHAM, with the order parameter being the distance between the centres of mass of two groups of nucleotides coloured in red and blue shown in Figure 5.11. The resulting landscape is also shown in Figure 5.11.

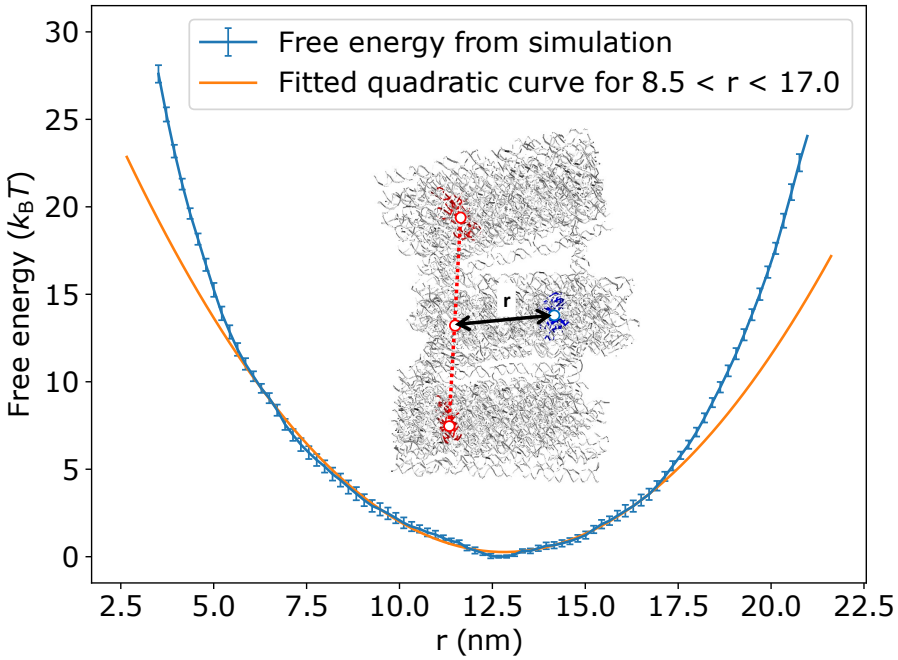


Figure 5.11: Free energy landscape as a function of the displacement (r) of the flanking blocks with respect to the middle block. The landscape at small displacements from equilibrium can be fitted to a quadratic curve, indicating a harmonic behaviour. At larger displacements, the free energy increases more rapidly than a quadratic curve. Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

At small displacements from the equilibrium, the landscape can be well approximated by a quadratic curve, indicating that the flanking blocks are restrained to the middle block by an effectively harmonic potential. At larger displacements beyond roughly 6 nm from the equilibrium, the free energy rises more rapidly than a quadratic curve. Visual inspection of the simulation trajectory reveals that the linkers are almost fully extended at this displacement. The value of 6 nm roughly agrees with the length of a linker, which is 21 base pairs and corresponds to approximately 7.14 nm. The small difference probably originates from the fact that the flanking blocks are not only displaced translationally but also rotationally. With rotational displacement, the flanking blocks are not aligned with the middle block, so the linker length is larger than the translational displacement of the flanking block even at full extension of the linker. After the linker is fully extended, further displacement of the flanking blocks requires other structural changes, such as breaking of base pairs near the anchoring points of the linking duplex. These changes have a much higher free energy cost and result in the free energy increasing more rapidly than a quadratic curve.

Looking back at the flanking block displacement data of the polymers in Figure 5.8 and Figure 5.9, the maximum displacements mostly plateau off at about 7 nm. This reflects that the flanking block displacement mostly remains in the harmonic regime of the free energy landscape. When stress builds up further, instead of exploring the zones beyond the harmonic regime, the flanking blocks explore alternative modes of stress release, such as piling on each other, as described in the section on polymer structure.

Inter-monomer stress

The inter-monomer stress accumulation was measured by calculating the free energy landscape of repulsion between two flanking blocks. The free energy landscapes were calculated using pulling simulations where two flanking blocks were slowly pulled towards each other, and integrating the resulting force-distance plot. The free energy landscapes for 1T, 3T and 6T are shown in Figure 5.12. As the length

of the poly-T extensions increases, the repulsion extends to larger values of the inter-monomer separation, in agreement with the design.

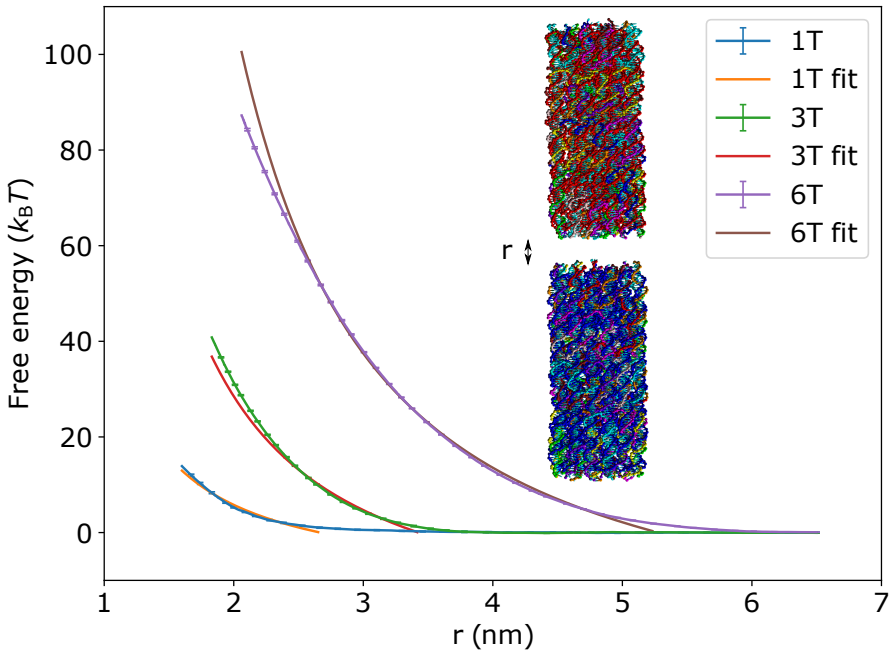


Figure 5.12: Free energy landscape of repulsion between two flanking blocks as a function of the distance (r) between their surface nucleotides. The landscapes can be fitted to the de Gennes potential, which are also shown in the figure. Error bars are calculated from 10 independent pulling simulations.

The landscapes can be fitted with a potential function proposed by de Gennes [190, 191], which models the interaction potential between two surfaces grafted with polymer brushes. Mathematically,

$$V(x) = C\left(\left(\frac{2L_0}{x}\right)^{9/4} - \left(\frac{x}{2L_0}\right)^{3/4}\right)$$

where C is a constant, L_0 is the equilibrium polymer brush thickness, and x is the separation between the two surfaces. C and L_0 were used as the fitting parameters. The fits are also shown in Figure 5.12. The fitted values of (C, L_0) in units of (k_BT, nm) for 1T, 3T and 6T are (5.23, 1.33), (10.6, 1.71) and (12.9, 2.64).

To verify that the de Gennes potential form is appropriate for modelling the repulsion between two flanking blocks, we calculated the probability density of the poly-T extensions as a function of the distance from the surface. The density plots are shown in Figure 5.13. The fitted value of L_0 for 3T and 6T are in satisfactory

agreement with the distances at which the density of poly-T becomes close to zero. There is a larger discrepancy for 1T, presumably because it is too short to be qualified as a polymer. Nevertheless, all the three fitted curves are in satisfactory agreement with the free energy landscapes, with the agreement being better at low levels of repulsion between the blocks. As we are only looking for a reasonable functional form for describing short-ranged repulsion, we believe that the de Gennes form is appropriate for our purpose, even though its physical description is not entirely appropriate for the 1T system.

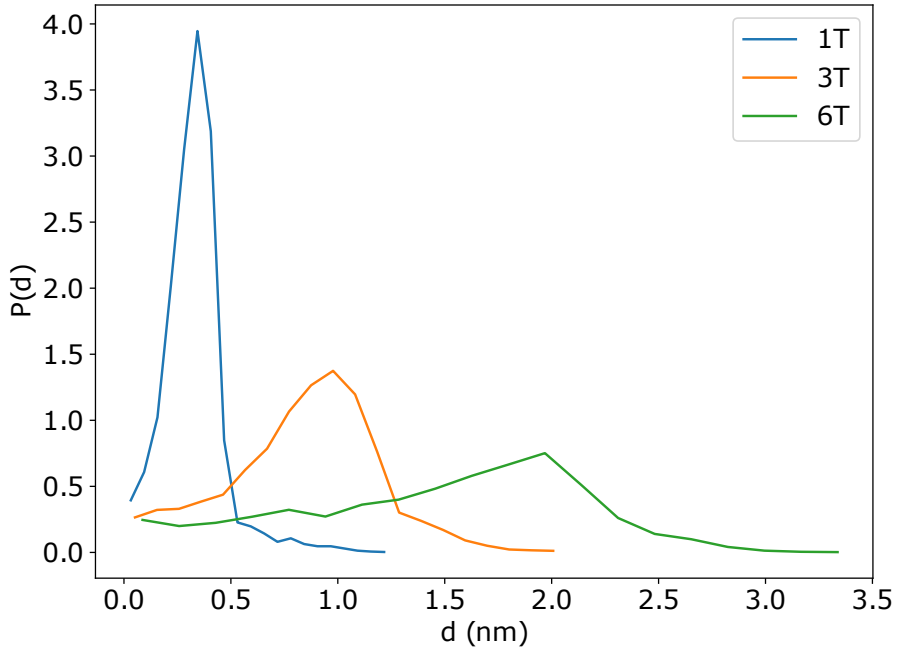


Figure 5.13: Probability density of the distance (d) of the ends of the poly-T extensions away from the surface of the flanking block.

5.3.4 Monte Carlo simulations of the simple stress model

With the free energy landscapes of intra-monomer stress and inter-monomer stress characterised, we performed Monte Carlo simulations on a simple 1D model that incorporates both modes of stress accumulation. We assume that the two flanking blocks are symmetric, so in the 1D model we only considered one flanking block per monomer. In this model, a monomer is treated as two points, which represent the centres of mass of the middle block and the flanking block respectively. A

polymer is thus treated as two sets of points, where the first set represents the centres of mass of each middle block and are fixed at equally spaced positions, with the spacing being the length of a monomer; whereas the second set represents the centres of mass of each flanking block. The second set of points are free to move but are restrained to their corresponding point in the first set, which is implemented using the free energy landscape of displacing the flanking blocks. The second set of points also experience a repulsion between nearest neighbours, which is implemented using the free energy landscape of repulsion between flanking blocks. The repulsion potential is cut off at the point where the fitted de Gennes potential is equal to zero. A schematic diagram of a dimer is shown in Figure 5.14.

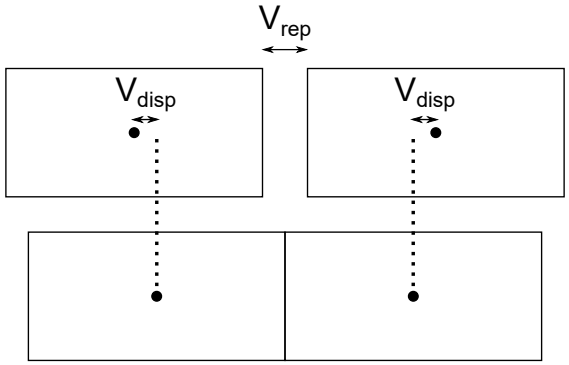


Figure 5.14: Schematic diagram of a dimer represented in the 1D model. The displacement potential (V_{disp}) acts on the distance between the flanking block and its attached middle block, while the repulsion potential (V_{rep}) acts on the distance between two neighbouring flanking blocks.

Similar to the earlier section on polymer structure, we simulated 1T, 3T and 6T polymers of length 2 to 9 using this 1D model. We performed a similar set of structural analysis as well. We calculated the translational displacement distributions of the flanking blocks and their average values at each monomer position. The plots are shown in Figure 5.15. From these plots, we also calculated the slope of a straight line fitted to each of the plots. The slope is plotted against the polymer length in Figure 5.16.

Comparing Figure 5.10 and Figure 5.16, the 1D model predicts slopes that are generally higher than those calculated from the oxDNA simulations. This is somewhat anticipated because the 1D model neglects other degrees of freedom

that could be used to reduce stress. Stress is localised in the mode of flanking block displacement along the polymerisation axis, so the displacements tend to be larger, with the slopes following the same trend. At short polymer lengths, the two figures are more similar, but as the polymer length increases, other stress release mechanisms that the 1D model is unable to capture become increasingly important, so the two figures deviate further. A similar comparison can be made between Figure 5.8/5.9 and Figure 5.15. The average displacements in the 1D model also tend to be larger than those in the oxDNA simulations, with the deviation being more significant for the longer polymers.

Nevertheless, the overall decreasing trend of the slopes is similar in both figures. This qualitative agreement illustrates that stress accumulation of short polymers in the limit of low stress is well described by the intra-monomer and inter-monomer free energy landscapes considered in the 1D model.

We can also calculate the average stress energy per monomer from the simple stress model, which is shown in Figure 5.17. As the model was parametrised from free energy landscapes, the potential energy of the model can be interpreted as a free energy. The average stress free energy is an increasing function of the polymer length, confirming the self-limiting nature of the polymerisation process.

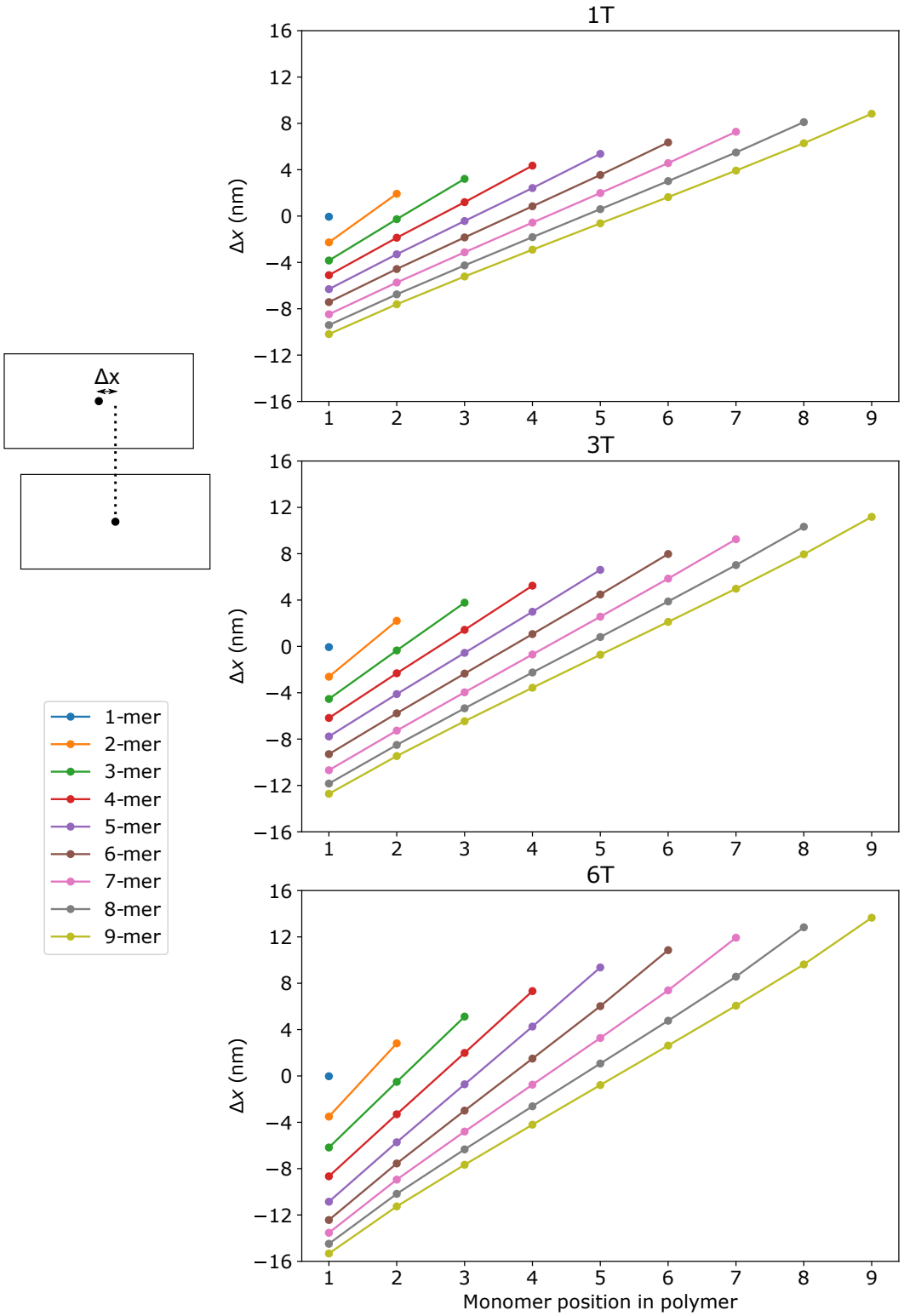


Figure 5.15: Average translational displacement of the flanking block for different polymer lengths and poly-T lengths in the 1D model. Each point corresponds to a monomer inside the polymer, while each line connects the points corresponding to a polymer of a given length.

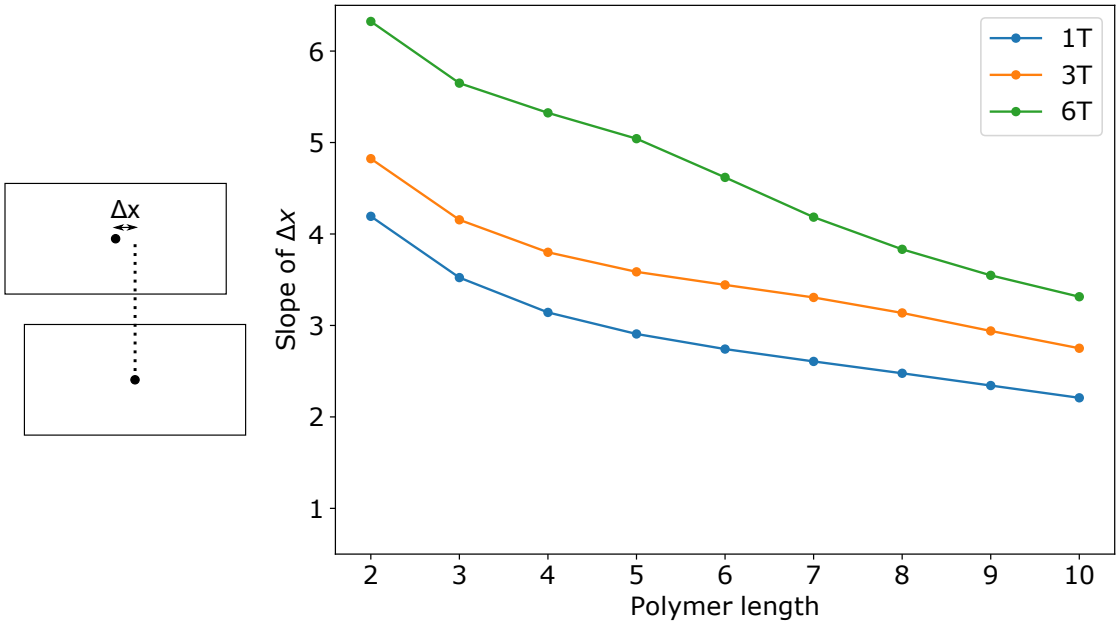


Figure 5.16: Slope of the plots in Figure 5.15. Each point represents a polymer with a given length and poly-T length in the 1D model. The decrease in slope for the longer polymers indicates less thermal fluctuation, consistent with the results from the oxDNA simulations, although the magnitude of the slopes are higher in the 1D model. The larger slopes reflect the fewer degrees of freedom in the 1D model to reduce stress, so that stress is localised in the mode of flanking block displacement.

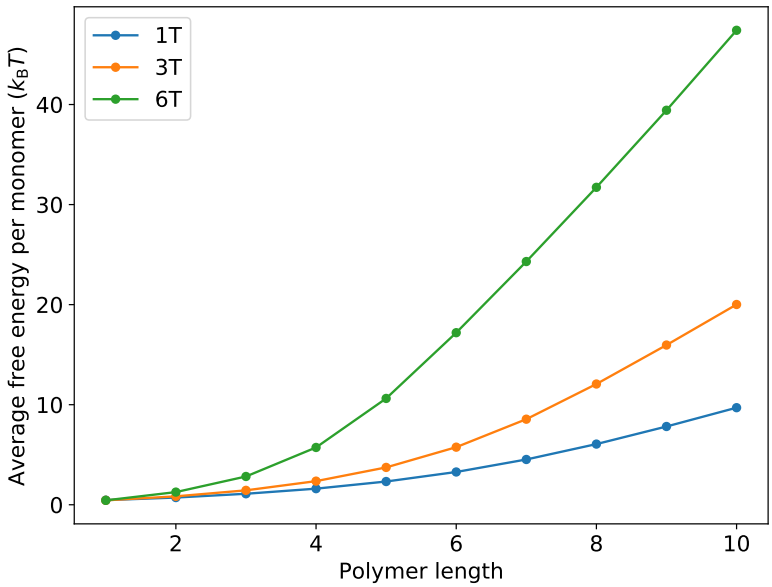


Figure 5.17: Average stress free energy per monomer as a function of polymer length. Each curve corresponds to a different poly-T length of the monomer. All the curves are increasing, consistent with the self-limiting nature of the polymerisation process.

5.3.5 Thermodynamics of polymerisation

In addition to structural characterisation of the polymers, the thermodynamics of polymerisation is also of interest. We hope to understand how a difference in the length of the poly-T extensions, i.e. the strength of inter-monomer stress accumulation, translates into a difference in equilibrium polymer length. The free energy of adding a monomer to an existing polymer, ΔG, is related to the equilibrium constant of the process, K, by the formula $\Delta G = -k_{\mathrm{B}}T \ln K$, where K is given by the ratio of the rate of monomer binding (on-rate) to the rate of monomer dissociation (off-rate). To probe how the length of the poly-T extensions affect the on-rate and off-rate, we performed pulling simulations to study the free energy barriers associated with these rates. For the on-rate, the simulations involved a monomer and another monomer or a polymer of length 5; whereas for the off-rate, the simulations involved a monomer being pulled away from a polymer of length 2 or 6.

Figure 5.18 shows the force-distance curves for the on-rate simulations. A negative force implies a net repulsion between the two objects being pulled, while a positive force implies a net attraction. The shapes of the curves follow a general trend: as the distance between the two object decreases, the repulsive force gradually strengthens, and suddenly becomes attractive once the two objects are close together and base pairs begin to form between their sticky ends.

In the simulations where two monomers are pulled together, the curves are quite similar for different lengths of the poly-T extensions, meaning that the free energy barrier of dimer formation does not depend on the poly-T extension length. However, in the simulations where a monomer is pulled towards a polymer of length 5, monomers with a longer poly-T extension experience a stronger repulsion on approaching the polymer. This difference can be rationalised by the structural characterisation made in the previous sections. When a monomer approaches a polymer, the monomer has to displace its flanking blocks so that the middle block can bind. This is shown in the series of simulation configurations in Figure 5.19. With a longer poly-T extension, polymers experience a stronger internal stress, resulting in the flanking blocks being displaced to a larger extent. The incoming

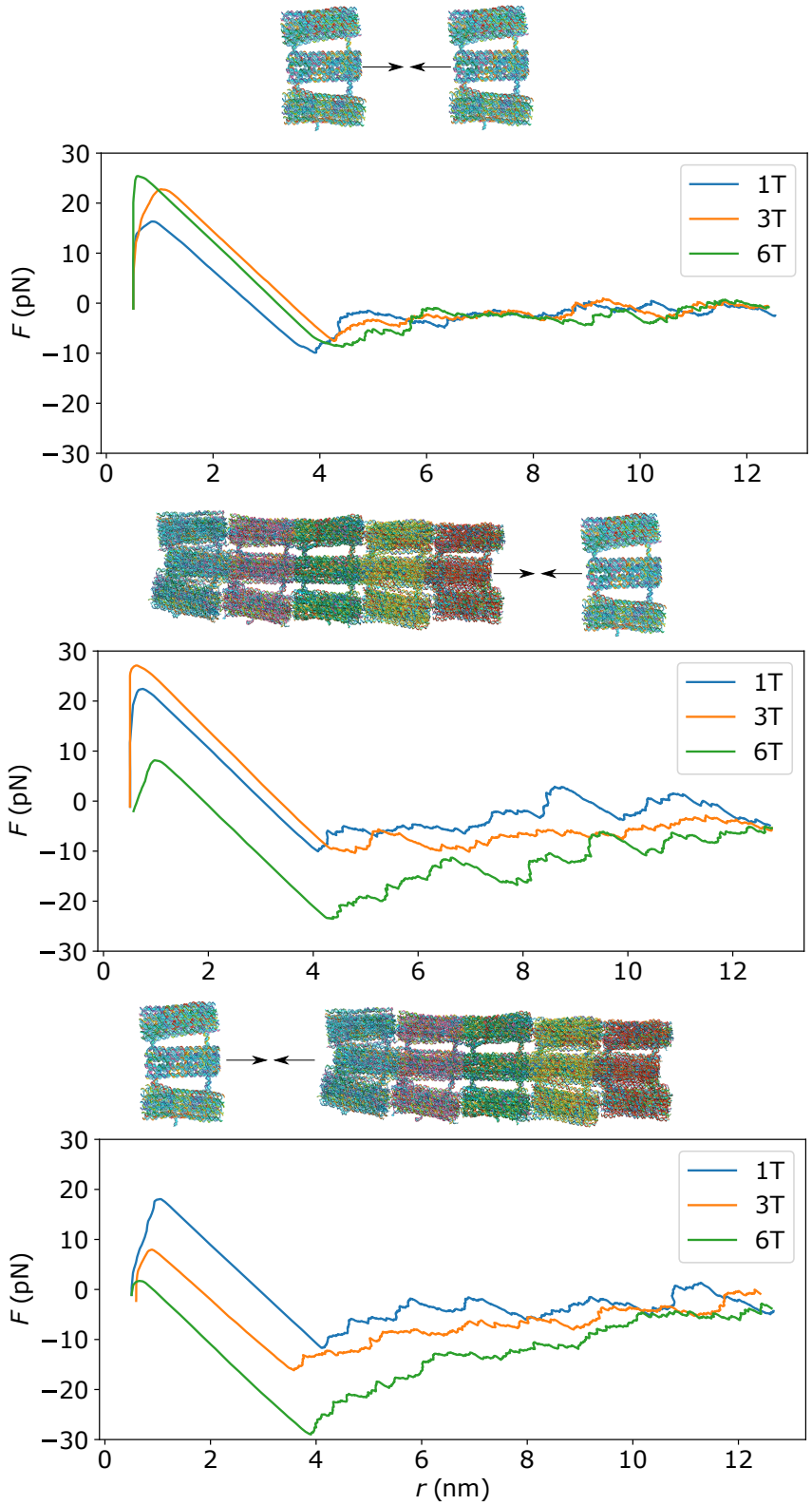


Figure 5.18: Force-distance curves for the on-rate simulations. A negative force indicates a net repulsion between the pulled objects; *vice versa*. Each curve corresponds to a single pulling simulation for the system illustrated above the corresponding plot.

monomer thus has to displace its flanking blocks to a larger extent as well. The free energy required during this process is manifested in an apparently stronger repulsion. These simulations and force-distance curves illustrate that polymers with a longer poly-T extension or with a longer length will have a lower on-rate because of the higher free energy barrier involved in adding a monomer.

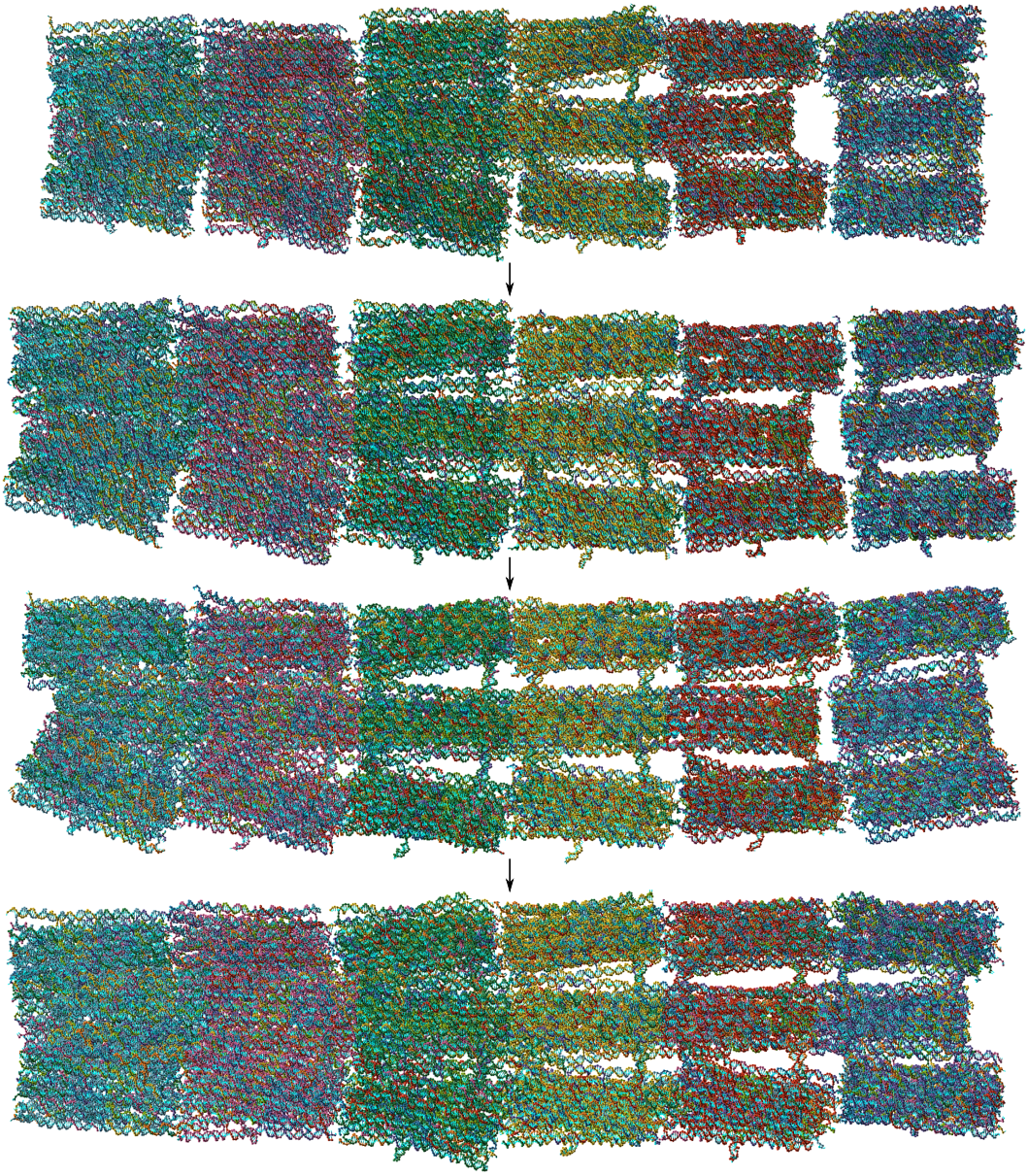


Figure 5.19: A series of configurations during an on-rate simulation involving a 3T 5-mer and a monomer combining to form a 6-mer. The flanking blocks of the incoming monomer are displaced before the middle block can bind to the polymer.

Figure 5.20 shows the force-distance curves for the off-rate simulations. Again,

a negative force implies a net repulsion between the two objects being pulled, while a positive force implies a net attraction. The shapes of the curves follow a general trend: starting from all the sticky ends being base-paired, the pulling force gradually builds up until it is strong enough to rupture the base pairs in the sticky ends.

A feature to note is that the forces involved in the off-rate simulations are much stronger than those in the on-rate simulations because base pairs have to be ruptured in the off-rate simulations. This apparent hysteresis reflects that the sampling pathway involving base pair formation or rupture is far from equilibrium, and it is challenging to compute the equilibrium free-energy landscape along this pathway.

The curves for different lengths of the poly-T extensions and different lengths of the polymers do not show any significant differences. From the simulation configurations shown in Figure 5.21, just after all the base pairs have been broken at the sticky ends, all the flanking blocks are still displaced to a similar extent as in a complete polymer. As such, flanking block repulsion should not play a significant role in the force needed to detach a monomer from the polymer. The only determinant of the force is the number of base pairs that needs to be broken, and this number is the same (48) in all the simulations. Therefore, we do not expect any significant difference in the curves for different systems. Due to the strong forces involved and the non-equilibrium nature of base pair rupture, there are some variations between different systems. Nevertheless, the off-rate simulations do not suggest any dependence of the free energy barrier of removing a monomer, and thus the off-rate, on the length of the poly-T extensions, nor on the length of the polymer.

With the conclusions from the on-rate and off-rate simulations, we now have a clearer picture of the thermodynamics of polymerisation. The self-limiting nature of polymerisation is most likely governed by the on-rate instead of the off-rate. As a polymer lengthens, it becomes increasingly difficult for an incoming monomer to bind to the existing polymer. Beyond a certain polymer length, the flanking blocks near the end of the polymer are displaced to a very large extent, and it becomes extremely unlikely for the incoming monomer to spontaneously displace its flanking blocks to such a large extent in order to bind the middle block. The

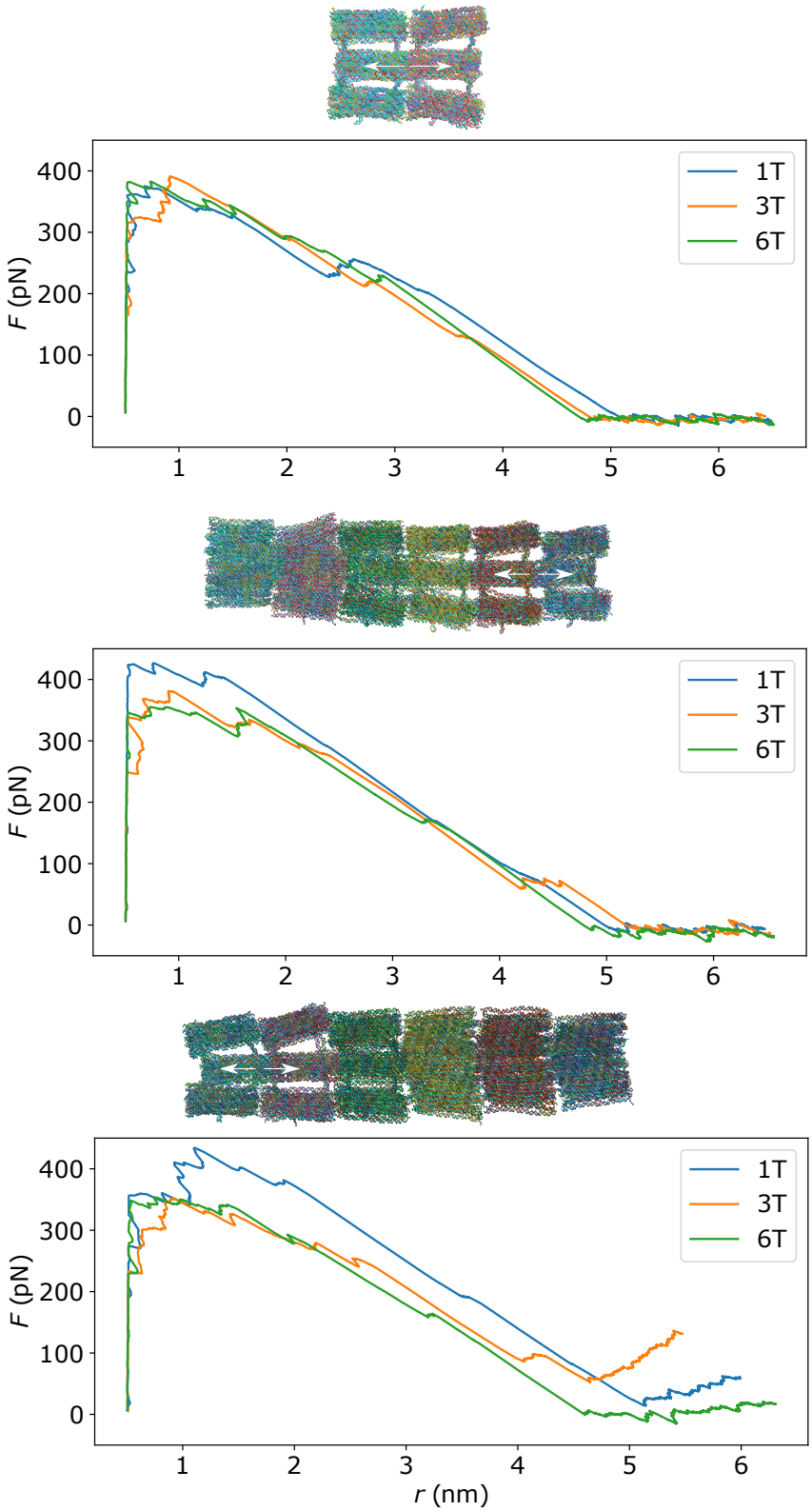


Figure 5.20: Force-distance curves for the off-rate simulations. A negative force indicates a net repulsion between the pulled objects; *vice versa*. Each curve corresponds to a single pulling simulation for the system illustrated above the corresponding plot.

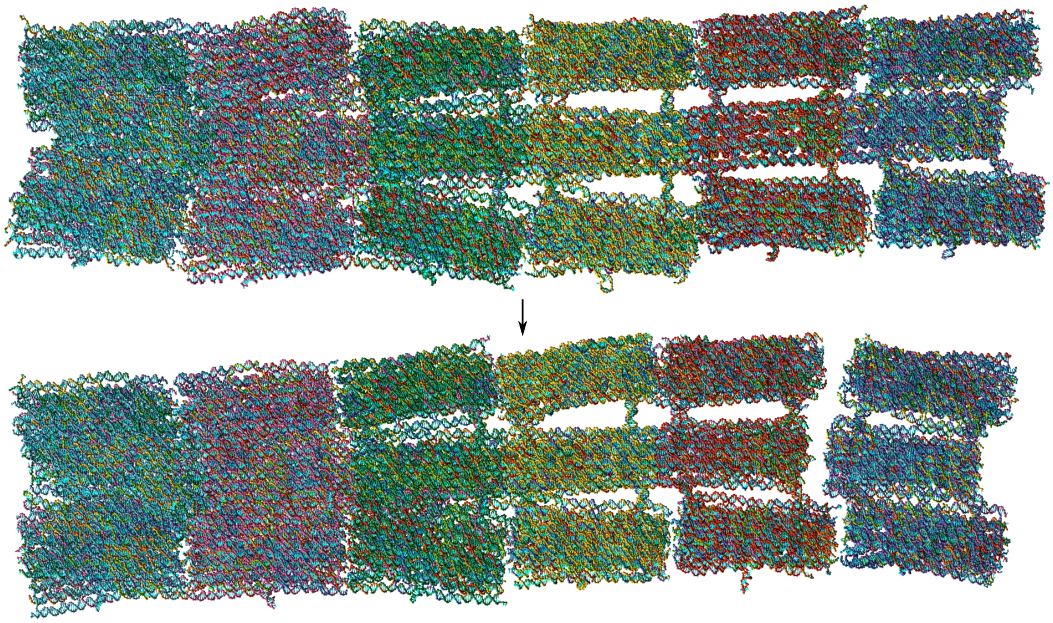


Figure 5.21: A series of configurations during an off-rate simulation involving a 3T 6-mer being separated into a 5-mer and a monomer. The flanking blocks of the leaving monomer are still displaced after the middle block has detached from the polymer.

on-rate thus becomes very small, and the polymer reaches an equilibrium length when the on-rate equals the off-rate. With longer poly-T extensions, the on-rate drops off more quickly as the polymer lengthens, and so the polymer reaches a shorter equilibrium length when the on-rate equals the off-rate.

5.4 Methods

5.4.1 Monomer and polymer structure

To generate starting configurations, the caDNAno design file of a monomer was converted into oxDNA format using the tacoxDNA software package. Polymers were constructed by making copies of the monomer and shifting them until the base pairs between the sticky ends of the monomers are at the desired positions. The repulsive flanking domains were further shifted such that neighbouring monomers do not overlap. The resulting monomer and polymer structures were then subjected to the standard relaxation procedure. After that, each structure was equilibrated for

456 ns (3×10^7 steps) before a production run of 1.52 µs (10^8 steps). Configurations were outputted every 1.52 ns (10^5 steps).

A number of vectors were defined and used in the calculation of translational and angular displacement of the flanking blocks. These include the vector of each block along each axis, denoted as $\mathbf{A_x}$, $\mathbf{B_x}$, $\mathbf{C_x}$, $\mathbf{A_y}$, $\mathbf{B_y}$, $\mathbf{C_y}$, $\mathbf{A_z}$, $\mathbf{B_z}$ and $\mathbf{C_z}$; and the vector pointing from the middle block to each flanking block, denoted as **BA** and **BC**.

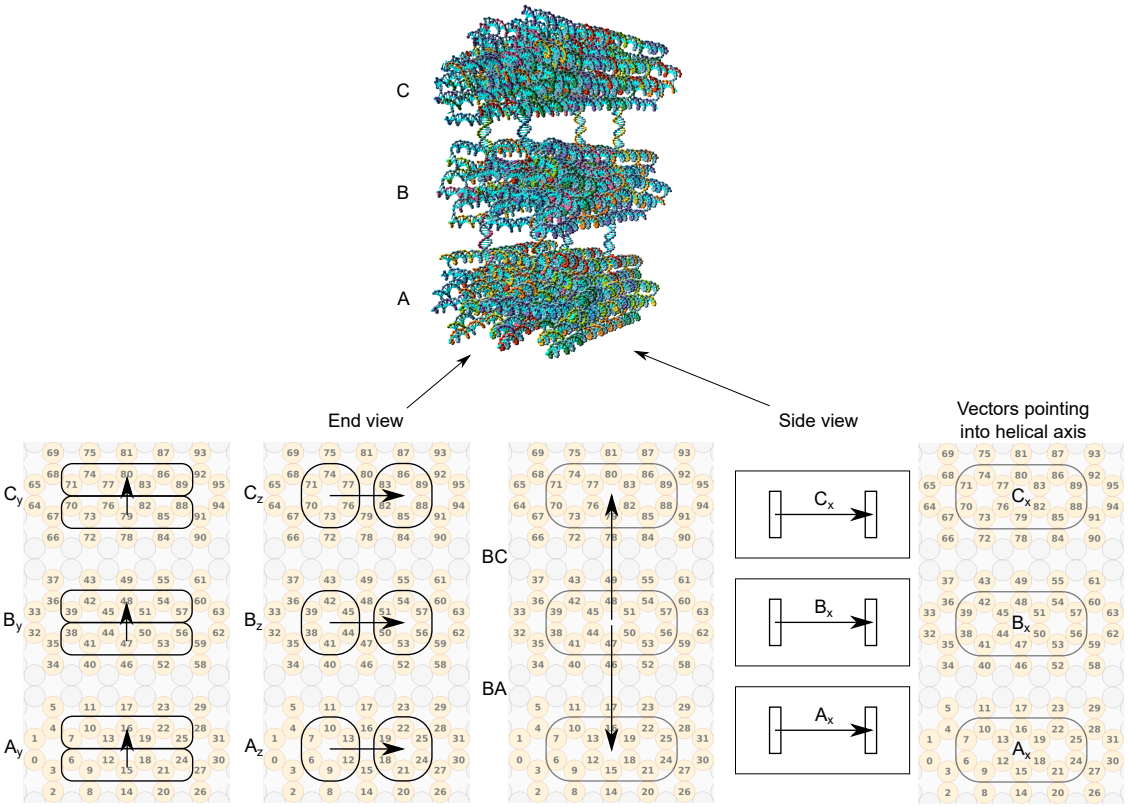


Figure 5.22: Definitions of the vectors used in the displacement calculations.

$\mathbf{A_x}$, $\mathbf{B_x}$ and $\mathbf{C_x}$ are shown in Figure 5.22. $\mathbf{A_x}$ is a vector inside block A that points from the centre of mass of the nucleotides in the first 20% to 25% of the interior helices of block A, to that of the last 20% to 25%. $\mathbf{B_x}$ and $\mathbf{C_x}$ are similarly defined.

$\mathbf{A_y}, \mathbf{B_y}, \mathbf{C_y}, \mathbf{A_z}, \mathbf{B_z}, \mathbf{C_z}, \mathbf{BA}$ and $\mathbf{BC}$ are also shown in Figure 5.22. To avoid end effects, the nucleotides at the starting 20% and ending 20% of each helix were excluded in the calculations.

The quantities are calculated as follows:

$$\Delta x_{AB} = \frac{\mathbf{BA} \cdot \mathbf{B_x}}{|\mathbf{B_x}|}, \Delta x_{CB} = \frac{\mathbf{BC} \cdot \mathbf{B_x}}{|\mathbf{B_x}|}$$

$$\Delta\theta_{AB} = 90^\circ - \arccos\left(\frac{\mathbf{A_x}\cdot\mathbf{B_y}}{|\mathbf{A_x}||\mathbf{B_y}|}\right), \Delta\theta_{CB} = 90^\circ - \arccos\left(\frac{\mathbf{C_x}\cdot\mathbf{B_y}}{|\mathbf{C_x}||\mathbf{B_y}|}\right)$$

5.4.2 On-rate/off-rate: pulling simulations

For simulations where a monomer and an oligomer were pulled towards each other, their equilibrated configurations were put into the same box such that (1) their polymerisation axis align, and (2) the mutually-trapped nucleotides were separated by approximately 12.8 nm. For simulations where a monomer was pulled away from an oligomer, the equilibrated configuration of the oligomer was used directly.

Moving mutual traps were used in the pulling simulations. These traps were applied between pairs of nucleotides. The trap energy is a function of time and the distance between the nucleotides, denoted as t and r respectively. Mathematically,

$$V(r,t) = \frac{1}{2}k(r-(r_0+ts))^2$$

where k is the force constant, r_0 is the equilibrium distance at the beginning, and s is the moving speed of the equilibrium distance.

For simulations where an oligomer was pulled towards a monomer, moving mutual traps were applied on the scaffold nucleotides of the middle blocks which will stack with each other after oligomerisation is completed. 8 of the mutual traps are labelled in Figure 5.23. There are 24 mutual traps in total, corresponding to the sticky ends of the 24 helices circled in the end view of Figure 5.23. k is set to be 5.709 pN/nm. r_0 is set to be 12.8 nm. s is set to be -0.0056 m/s. Simulations were performed for 2.19 µs (1.44×10^8 steps), after which the equilibrium distance of the trap is equal to the equilibrium coaxial stacking distance in the oxDNA model, 0.511 nm. The distance of each mutual trap was outputted every 15.2 ps (10^3 steps).

For simulations where a monomer was pulled away from an oligomer, moving mutual traps were applied on the same nucleotides as described above. k is set to be 0.5709 pN/nm. r_0 is set to be 0.511 nm. s is set to be 0.0056 m/s. Simulations were performed for 1.06 µs (7×10^7 steps), after which the monomer and the rest of the oligomer are sufficiently separated such that the trap force is close to zero. The distance of each mutual trap was outputted every 15.2 ps (10^3 steps).

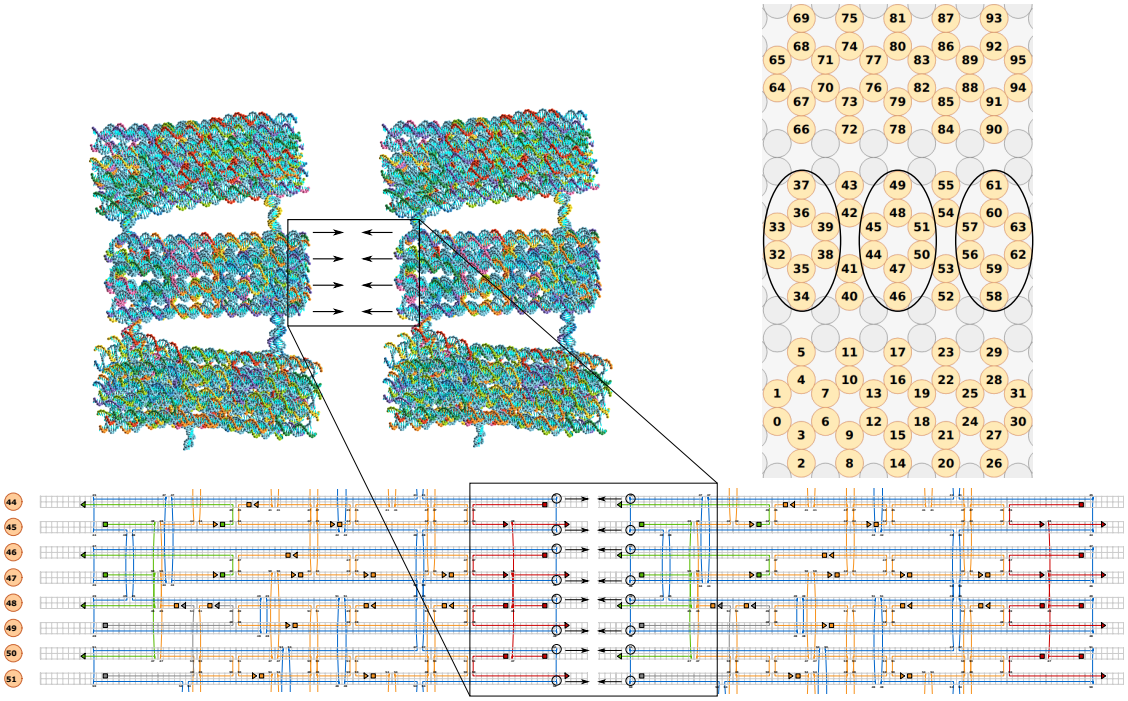


Figure 5.23: Locations of the mutual traps. The circles and arrows show the positions and directions of 8 of the mutual traps. All the 24 mutual traps are in the circled helices in the end view shown in the upper right corner.

The total force of the 24 traps were calculated for each frame in the trajectory. The moving average of the force over 1000 (for pulling on) or 100 (for pulling off) consecutive frames was calculated for each frame to reduce noise. It was plotted as a function of the average instantaneous distance of the 24 traps to give a force-distance plot.

5.4.3 Intra-monomer stress: umbrella sampling

The order parameter in umbrella sampling simulations was chosen to be the distance r between the centres of mass (COM) of two groups of coloured nucleotides shown in Figure 5.11. One might wonder why we did not choose a more intuitive order parameter, e.g. the distance d between the COM of the whole flanking block and the whole middle block. The reason is that this order parameter introduces difficulties in determining the relative positions of the block; flanking blocks displaced to the left or right cannot be distinguished because they both give similar positive values

of d. By contrast, by choosing our order parameter, displacement to the left or right gives different values of r and facilitates sampling of this order parameter.

Two pulling simulations were used to generate the starting configuration for each umbrella sampling window. The equilibrated configuration of a monomer was used as the starting configuration for the pulling simulations. The order parameter r was first restrained at its initial value with a harmonic potential. The potential is applied on the coloured nucleotides, and is a function of time and the distance between the COM, denoted as t and r respectively. Mathematically,

$$V(r,t) = \frac{1}{2}k(r - (r_0 + ts))^2$$

where k is the force constant, r_0 is the equilibrium distance at the beginning, and s is the moving speed of the equilibrium distance.

The equilibrium distance of the harmonic potential was then slowly increased in one simulation and decreased in another simulation. k is set to be 57.09 pN/nm. r_0 is set to be 14.5 nm. s is set to be ± 0.0056 m/s. Simulations were performed for 1.52 µs (1×10^8 steps) along the positive direction, and 2.58 µs (1.7×10^8 steps) steps along the negative direction. The system configuration was outputted every 76 ns (5×10^6 steps), corresponding to a difference in equilibrium distance of 0.43 nm between each configuration.

The configurations from the pulling simulations were used as the starting configuration for each umbrella sampling window. The windows ranged from $r_0 = 0$ nm to $r_0 = 23$ nm and had a spacing of 0.43 nm. For each window, k is set to be 11.4 pN/nm, and r_0 is fixed at the value corresponding to the window. Each simulation was equilibrated for 30.4 ns (2×10^6 steps) before a production run of 152 ns (1×10^7 steps).

The order parameter distributions of all windows were re-weighted using the Weighted Histogram Analysis Method (WHAM) to give the unbiased free energy landscape.

5.4.4 Inter-monomer stress: pulling simulations

We performed pulling simulations where two flanking blocks were slowly pulled towards each other. The initial configuration was prepared by placing two flanking blocks into the same simulation box such that their helical axes were aligned and the mutually-trapped nucleotides were separated by approximately 8.5 nm.

Moving mutual traps similar to those for the on-rate/off-rate simulations were used. 8 of the mutual traps are labelled in Figure 5.24. There are 24 mutual traps in total, corresponding to the 24 helices circled in the end view of Figure 5.24. The mutual traps were applied between scaffold nucleotides that would stack with each other if there were no poly-T extensions in the flanking block. There are 24 mutual traps in total, corresponding to the 24 helices with poly-T extensions in the flanking block. k is set to be 5.709 pN/nm. r_0 is set to be 8.52 nm. s is set to be -0.0011 m/s. Simulations were performed for 7.6 µs (5×10^8 steps), after which the equilibrium distance of the trap reaches zero. The distance of each mutual trap was outputted every 152 ps (10^4 steps).

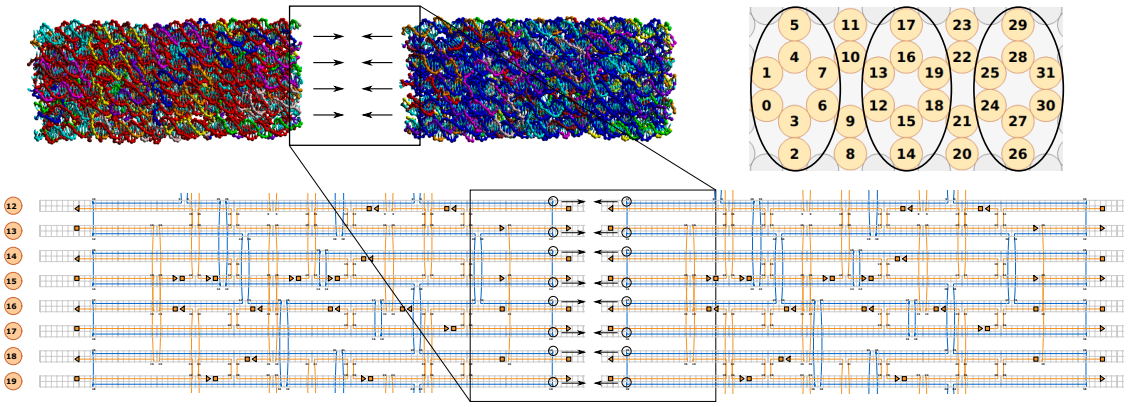


Figure 5.24: Locations of the mutual traps. The circles and arrows show the positions and directions of 8 of the mutual traps. All the 24 mutual traps are in the circled helices in the end view shown in the upper right corner.

The total force of the 24 traps were calculated for each frame in the trajectory. The moving average of the force over 100 consecutive frames was calculated for each frame to reduce noise. It was plotted as a function of the average instantaneous distance of the 24 traps, and this force-distance plot was then numerically integrated to give an energy-distance plot.

It is worth noting that here we used pulling simulations instead of umbrella sampling because the two blocks have to be oriented towards each other, which can be enforced by applying numerous mutual traps in pulling simulations, but not by an order parameter based on the distance between the centres of mass of two groups of nucleotides in umbrella sampling.

5.4.5 Monte Carlo simulations of the simple stress model

The model potential is described in the corresponding section in Results and Discussion. Starting configurations for polymers were generated by placing the set of points corresponding to the centres of mass of the flanking blocks at equally spaced positions such that their separation is equal to the inter-monomer potential cutoff distance. Then, Monte Carlo simulations were performed using the standard Metropolis algorithm at a temperature of 300 K. Moves were proposed by randomly selecting a monomer unit and displacing its flanking block by a distance x randomly sampled from a uniform distribution in the range of $-0.1 \leq x < 0.1$. 10^7 steps were performed per monomer unit in the polymer. The positions of the flanking blocks were outputted every 10^3 steps.

5.5 Conclusions

We have characterised the structural and thermodynamic properties of a DNA origami that undergoes self-limiting polymerisation. Our findings are complementary to the experimental studies and have helped the interpretation of the experimental results.

In terms of structure, the origami monomer shows an asymmetric shape, with one end of each flanking domain being further from the middle domain than the other end. The asymmetry stems from the difference in the effective length of the linking duplexes, which is determined by the distance to the nearest crossover from the ends of the anchoring duplex. As for the structure of short polymers, the stress accumulation mechanism generally behaves as designed, which involves the translational displacement of the flanking blocks. However, other stress release

mechanisms come into play for longer polymers. For example, polymers twist along the polymerisation axis and flanking blocks pile up on each other.

We parametrised a simple stress model to describe the stress accumulation mechanism in polymers. The intra-monomer stress was found to be well approximated by a harmonic potential at small displacements from the equilibrium, while at larger displacements, the stress increases more rapidly than a harmonic potential. The inter-monomer stress was satisfactorily modelled with a de Gennes potential form, which describes the repulsion between two surfaces grafted with polymer brushes. Combining the two free energy landscapes of stress accumulation, we performed Monte Carlo simulations to study the properties of polymers in this simple stress model. The model provides a good agreement with oxDNA simulations for short polymers, indicating that the model captures the essential features of stress accumulation at low stress levels. However, the model deviates further for longer polymers because stress release mechanisms are not taken into account.

We also studied the thermodynamics of polymerisation using pulling simulations. The free energy barrier of adding a monomer, and thus the on-rate, depends on length of the poly-T extensions, while that of removing a monomer, and thus the off-rate, does not. Note that the flanking blocks of the incoming monomer have to be displaced when it approaches the existing polymer. The required displacement is influenced by the strength of the repulsion between the flanking blocks; likewise for the free energy barrier. However, the free energy barrier of removing a monomer only depends on the nature of the sticky ends but not the displacement of the flanking blocks. Therefore, there is no observable correlation between this barrier and the poly-T length.

Although simulations were only used to interpret experimental results in this study, we expect that the detailed insight at the nucleotide level that oxDNA provides will encourage future research to utilise oxDNA as a tool for screening and improving possible origami designs before experiments. Examples of such detailed insights include how the flexibility of the linkers relates to the stress accumulation mechanism, the difference in the effective linker lengths that leads to

the asymmetric shape of the origami monomer, and the stress release mechanisms in highly stressed long polymers. These effects arise from an intricate interplay of various events occurring at the nucleotide level, such as base stacking/unstacking, excluded volume effects and duplex bending. It is challenging to explore these effects from experiments because of a limited resolution. Other more coarse-grained computational tools, such as CanDo, also cannot capture these subtle effects because of their representation of DNA which treats duplexes as rigid bodies and does not consider individual nucleotides.

6

Self-Limiting Assembly of Curved and Twisted DNA Origamis

Contents

6.1 Introduction 103
6.2 Curved Monomer: Ideal Model 106
6.3 Curved Monomer: Design 1 109
6.4 Curved Monomer: Design 2 125
6.5 Twisted Monomer 133
6.6 Methods 142
6.7 Conclusions 147

6.1 Introduction

In Chapter 5, accumulated displacement was used as a mechanism for stress accumulation in self-limiting assembly. A natural question to ask is whether other stress accumulation mechanisms are possible, for example, by accumulation of curvature or twist. In this chapter, we will explore possible designs of curved and twisted DNA origamis to achieve stress accumulation through such mechanisms. Unlike Chapter 5, our goal in this chapter is not to interpret experimental results, but rather to explore and characterise possible DNA origami designs computationally before they are studied in experiments.

The design of curved and twisted origamis was pioneered by Dietz *et al.* [7]. Local deviations from the natural pitch length of DNA is introduced into the origami design through base pair insertions and/or deletions. The deviations result in local torques that can be combined and manifested as global twisting or bending of the origami. For example, introducing only insertions or only deletions will result in the global accumulation of torque and thus global twisting. On the other hand, introducing insertions on one side of the origami and an equal number of deletions on the other side will result in torque cancellation and thus the absence of global twisting, but the difference in the number of bases on the two sides will induce global bending. To demonstrate hierarchical assembly, twisted origamis were polymerised along the helical axis to form twisted ribbons, while bent origamis were polymerised to form circular assemblies.

Another approach for designing curved origamis was introduced by Han *et al.* [8], where crossover patterns between helices were rationally designed to induce curvature. Various 2D and 3D origamis were synthesised, with the shape of concentric rings, spheres and ellipsoids. An important distinction of this work from the work of Dietz *et al.* [7] is the absence of a lattice in the design, which allows more flexibility in the shape of the origami, thereby allowing shapes such as spheres to be designed which are difficult to map onto a lattice.

Apart from rationally designing the origami strand routing, it is also possible to use a higher-level approach, where mechanical tension is introduced into the origami to achieve curvature. Liedl *et al.* [100] used single-stranded DNA as entropic springs to connect the two ends of a helix bundle, which causes the bundle to bend. In a similar vein, Suzuki *et al.* [192] used short single-stranded DNA to connect different parts of the origami and achieve curvature. However, instead of connecting the two ends of the helix bundle, gaps were introduced within the helix bundle and the gaps were connected with single-stranded DNA, which again acted as entropic springs. The curvature can be tuned by changing the number of bases in the single-stranded DNA connections.

Recently, Wang *et al.* [193] developed a dynamic approach to synthesise curved and twisted origamis. The origamis consist of modular units, which can expand through strand displacement reactions on hairpins placed within the units. By selectively expanding certain units, the geometry of the origami can be tuned. For example, if the units on only one side of the origami are expanded, a bent origami will be formed, while if the units on both sides, but not the central part, of the origami are expanded, a twisted origami will be formed. The demonstration of dynamically controlling the curvature and twist of DNA origamis is a particularly important advancement from previous work.

Although there has been much work on studying individual curved and twisted DNA origamis, less attention has been given to their assembly into higher order structures, especially in a self-limiting manner. Therefore, this chapter will be focused on the assembly of curved and twisted DNA origamis. In particular, we aim to utilise curvature or twist propagation through the assembly to induce stress accumulation, and thus self-limiting assembly.

When considering the assembly of origamis, apart from designing their individual structure, it is also necessary to consider how they will interact and assemble. In Chapter 5, short single-stranded DNA extensions and truncations (plugs and sockets) are introduced into the origami to facilitate binding between origamis. However, in this chapter, often we are interested in assembling origamis in a direction perpendicular to that of the double helices, which means that the plug-and-socket interaction is not applicable because this interaction is inherently parallel to the double helices.

One possible method for assembling DNA origamis perpendicular to the direction of the double helices is to design the origamis such that they have complementary shapes which fit into each other like Lego bricks. Wagenbauer *et al.* [171] has used this strategy to assemble origamis. At the sites where the complementary shapes fit into each other, there are weak blunt-ended stacking interactions between the DNA double helices, which allow origamis to interact and assemble. Large-scale assemblies have been achieved, where the resulting assembly can be over 1 gigadalton

in molecular weight, equivalent to 20 DNA origamis. However, the weakness of the interaction might be an issue for self-limiting assembly with stress accumulation. The stress levels might easily surpass the interaction strength, causing difficulty in reaching reasonable assembly sizes.

Another method is to use single-stranded DNA overhangs, also known as "sticky ends". Yao *et al.* [194] has designed "Meta-DNA" structures, which are rod-shaped origamis that interact through sticky ends that protrude perpendicularly from their sides. The authors have demonstrated assembly into 2D and 3D structures at micron scales. The stronger base-pairing interactions between sticky ends could make it a more suitable choice for self-limiting assembly, as these interactions can sustain a higher level of stress that is accumulated during the assembly process.

Three DNA origami designs will be studied in this chapter, where two of them are curved and one of them is twisted. For each design, we will simulate the monomers and assemblies of different sizes using the oxDNA model. From the monomer simulations, we will describe the free energy landscape of the monomer in terms of a small number of order parameters. As for the simulations of the origami assemblies, we will calculate the order parameters for each sampled configuration in the simulation trajectory. We will then infer the stress free energy of each monomer from the free energy landscapes and the values of each order parameter. The sum of the stress free energy of each monomer will give an estimate of the total stress free energy of the assembly.

6.2 Curved Monomer: Ideal Model

The self-limiting assembly of curved monomers utilises the gradual change in the curvature of the monomers as a way to propagate the bending stress. As the assembly grows, the new monomers have to deviate further and further from their equilibrium shape in order to adapt to the shape of the whole assembly. This increasing deviation gives rise to an increasing average stress as a function of the assembly size, and thus results in self-limiting assembly.

An ideal model for the self-limiting assembly of curved monomers can be developed with the assumption that all monomers remain in 2D while the assembly grows, and that their contours lie on concentric circles. A schematic diagram of the model is shown in Figure 6.1. The derivation below is based on that in Tanjeem *et al.* [195].

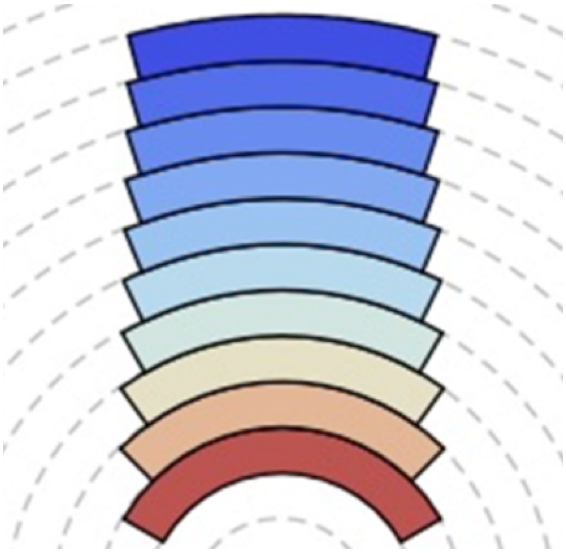


Figure 6.1: A schematic diagram of the ideal model for the assembly of curved monomers. The model assumes perfect contact between the monomers such that their contours lie on concentric circles. The curvature increases from top to bottom, with an accompanying change in the bending stress energy of the monomers as well. Colours towards the blue end of the spectrum indicate that the monomer is stressed by being flattened, while those towards the red end indicate that the monomer is stressed by being bent. Figure reproduced from Tanjeem *et al.* [195] with permission.

The bending stress energy of a monomer E_i can be estimated as a harmonic potential as a function of the curvature κ,

$$E_i = \frac{1}{2}B(\kappa - \kappa_0)^2$$

where the force constant of the harmonic potential is the bending modulus of the monomer B, the natural curvature of the monomer is κ_0, and the curvature κ is defined as the reciprocal of the radius of curvature r.

$$\kappa = r^{-1}$$

Since the monomers lie on concentric circles, their radii of curvature is a linear function of their position. The radii of curvature of two neighbouring monomers can be related as

$$r_{i+1} = r_i + t$$

where t is the thickness of the monomer. This formula can also be expressed in terms of the curvature, which gives

$$\kappa_{i+1}^{-1} = \kappa_i^{-1} + t$$

For a given assembly of N monomers, the total stress energy E of the assembly can be expressed as

$$E = \sum_{i=1}^{N} \frac{1}{2} B(\kappa_i - \kappa_0)^2$$

Assuming N is an odd number so that the middle monomer is unstressed and retains its natural curvature κ_0, we can express the curvature of each monomer κ_i as a function of their position z in the assembly.

$$\begin{aligned} E &= \sum_{z=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{1}{2} B((\kappa_0^{-1} + zt)^{-1} - \kappa_0)^2 \\ &= \sum_{z=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{1}{2} B(\kappa_0((1 + zt\kappa_0)^{-1} - 1)))^2 \end{aligned}$$

As far as self-limiting assembly is concerned, the assembly tends to be small in size, so we can assume that the vertical dimension of the assembly is small compared to the radius of curvature, i.e.

$$zt \ll r$$

which implies

$$zt\kappa_0 \ll 1$$

and thus

$$(1 + zt\kappa_0)^{-1} \approx 1 - zt\kappa_0$$

This assumption can be used to simply the expression for the total stress energy to

$$\begin{aligned}
E &= \sum_{z=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{1}{2} B(\kappa_0((1 - zt\kappa_0) - 1)))^2 \\
&= \sum_{z=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{1}{2} B\kappa_0^4 z^2 t^2 \\
&= \sum_{z=-\frac{N-1}{2}}^{-1} \frac{1}{2} B\kappa_0^4 z^2 t^2 + \sum_{z=1}^{\frac{N-1}{2}} \frac{1}{2} B\kappa_0^4 z^2 t^2 \\
&= 2 \sum_{z=1}^{\frac{N-1}{2}} \frac{1}{2} B\kappa_0^4 z^2 t^2 \\
&= B\kappa_0^4 t^2 \sum_{z=1}^{\frac{N-1}{2}} z^2 \\
&= B\kappa_0^4 t^2 \left(\frac{1}{6} \left(\frac{N-1}{2} \right) \left(\frac{N-1}{2} - 1 \right) \left(\frac{N-1}{2} - 2 \right) \right) \\
&\approx \frac{1}{48} B\kappa_0^4 t^2 N^3
\end{aligned}$$

The average stress energy of each monomer then follows trivially as

$$\frac{E}{N} \approx \frac{1}{48} B\kappa_0^4 t^2 N^2$$

which shows that the average stress energy approximately grows quadratically as the assembly size. The increase in the average stress as a function of the assembly size is a signature of systems which undergo self-limiting assembly, as the average stress will eventually become greater than the average binding energy between a pair of monomers, thereby causing the assembly to stop growing beyond a certain size.

6.3 Curved Monomer: Design 1

Our design is based on a 14-helix bundle on a honeycomb lattice. Base insertions and deletions are used to introduce curvature into the origami. Inter-monomer interactions are based on single-stranded DNA overhangs, otherwise known as "sticky ends", that perpendicularly emerge from two sides of the origami, similar to the approach used in Yao *et al.* [194]. 21 sticky ends are placed on the origami, arranged in a 7 by 3 manner, with 7 sticky ends on 3 different helices. Each pair

of sticky ends consist of 7 base pairs. When positioning the sticky ends, we aimed to roughly match the curvature of the monomer.

6.3.1 Assembly structure

In this section, we will look at the structure of the assembly by calculating the end-to-end distance and twist angle of each monomer, and the bond distance between monomers, in assemblies of different sizes. Example configurations of the simulated assemblies are shown in Figure 6.2 and 6.3.

End-to-end distance

The average end-to-end distance as a function of the monomer position within the assembly is shown in Figure 6.4. The monomer position index is the largest for the monomer at the innermost position of the assembly.

In small assemblies (sizes of 2 and 4), the end-to-end distance decreases monotonically from the outermost monomer to the innermost monomer. This is consistent with the ideal model, where stress accumulates as the assembly grows and the inner monomers become increasingly bent and stressed.

In the assembly with 6 monomers, the end-to-end distances are unexpectedly long compared to the smaller assemblies. According to the ideal model, the end-to-end distance of the innermost monomer should be shorter than that for smaller assemblies because the radius of curvature is smaller. However, the plot shows that the converse is true.

In the assembly with 8 monomers, the end-to-end distances are also unexpectedly long. Unlike other smaller assemblies, the end-to-end distances do not decrease monotonically as a function of the monomer position index. Between indices 3 and 5, the end-to-end distances increased slightly, which contradicts with the ideal model because the end-to-end distance should only decrease, given that the radius of curvature decreases.

These observations for larger assemblies indicate that the ideal model does not provide the full picture of the real stress accumulation mechanism. By inspecting

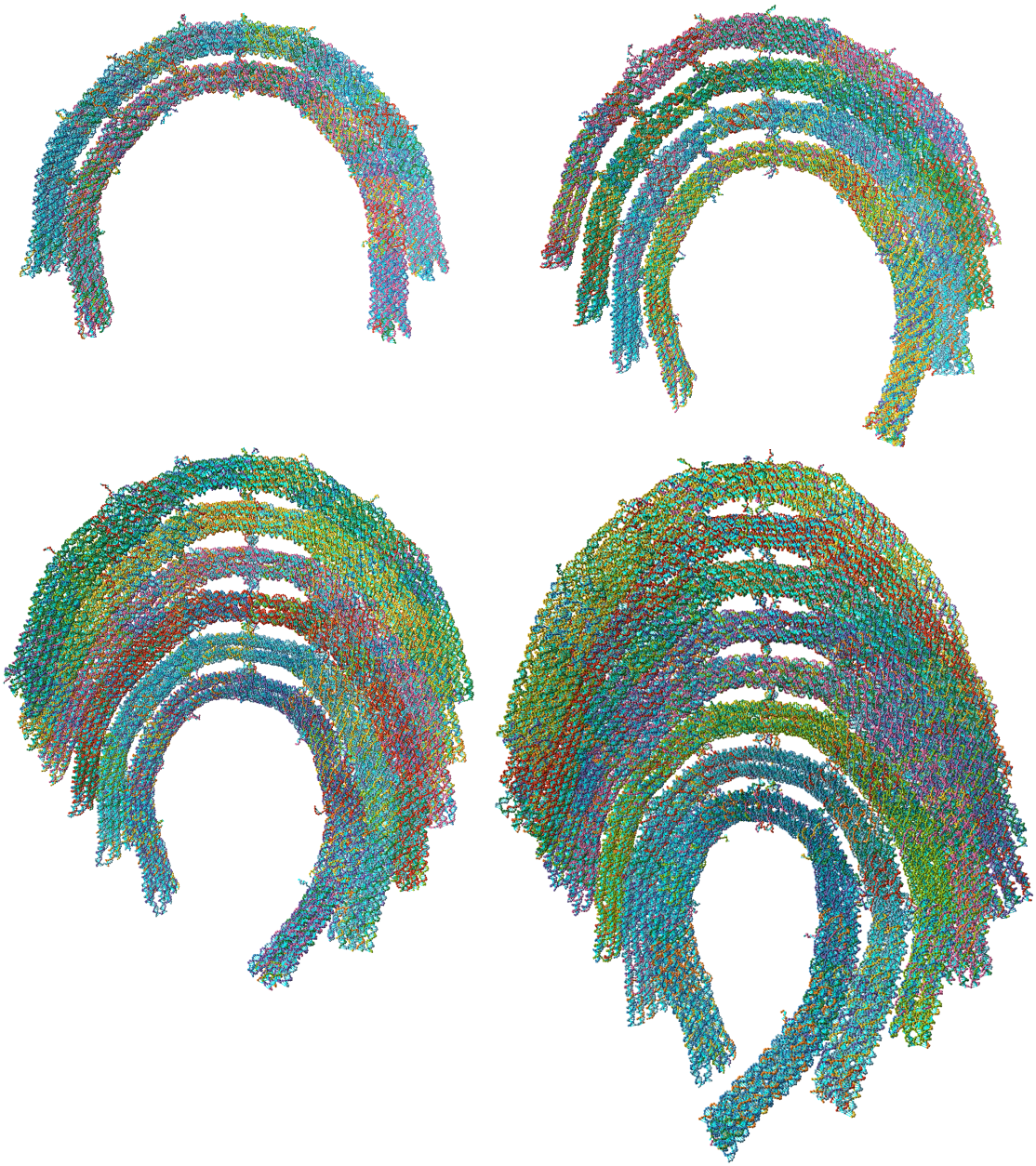


Figure 6.2: oxDNA simulation configurations of the assemblies of curved monomers. The sizes of the assemblies are 2, 4, 6 and 8, from left to right and top to bottom. While smaller assemblies remain roughly planar, a significant out-of-plane twist develops in larger assemblies.

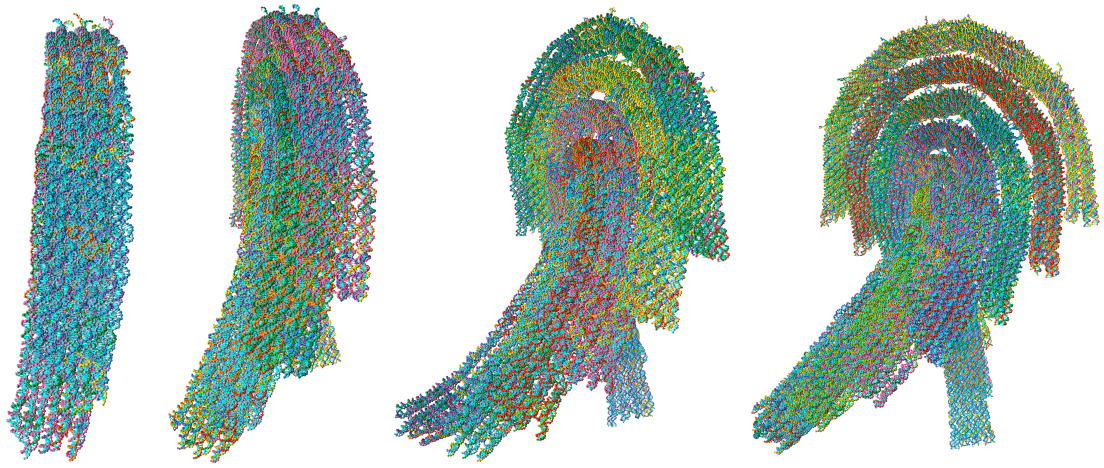


Figure 6.3: Side views of the oxDNA simulation configurations shown in Figure 6.2. The sizes of the assemblies are 2, 4, 6 and 8, from left to right. An out-of-plane twist is evident in larger assemblies.

the simulation configurations in Figure 6.3, it is obvious that the structures of larger assemblies are not planar as in the ideal model, but instead show out-of-plane twisting, indicating that bending is not the only mechanism of stress accumulation. It is therefore necessary to consider twisting as a source of stress as well.

Twist angle

The average twist angle as a function of the monomer position within the assembly is shown in Figure 6.5. The monomer position index is the largest for the monomer at the innermost position of the assembly.

The twist angles are smaller than 5 degrees in the smallest assembly with 2 monomers, indicating that the structure of the assembly is largely planar and conforms to the description of the ideal model. In the assembly with 4 monomers, there are initial signs of twisting in the assembly, with the inner monomers twisting more strongly than the outer ones, but the structure still looks largely planar as in Figure 6.2.

The assembly with 6 monomers shows an even stronger twist and shares a similar trend of a stronger twist for inner monomers. Meanwhile, the assembly with 8 monomers shows a somewhat different trend, where the middle monomers show the strongest twist. It is worth noting that the monomers with the strongest

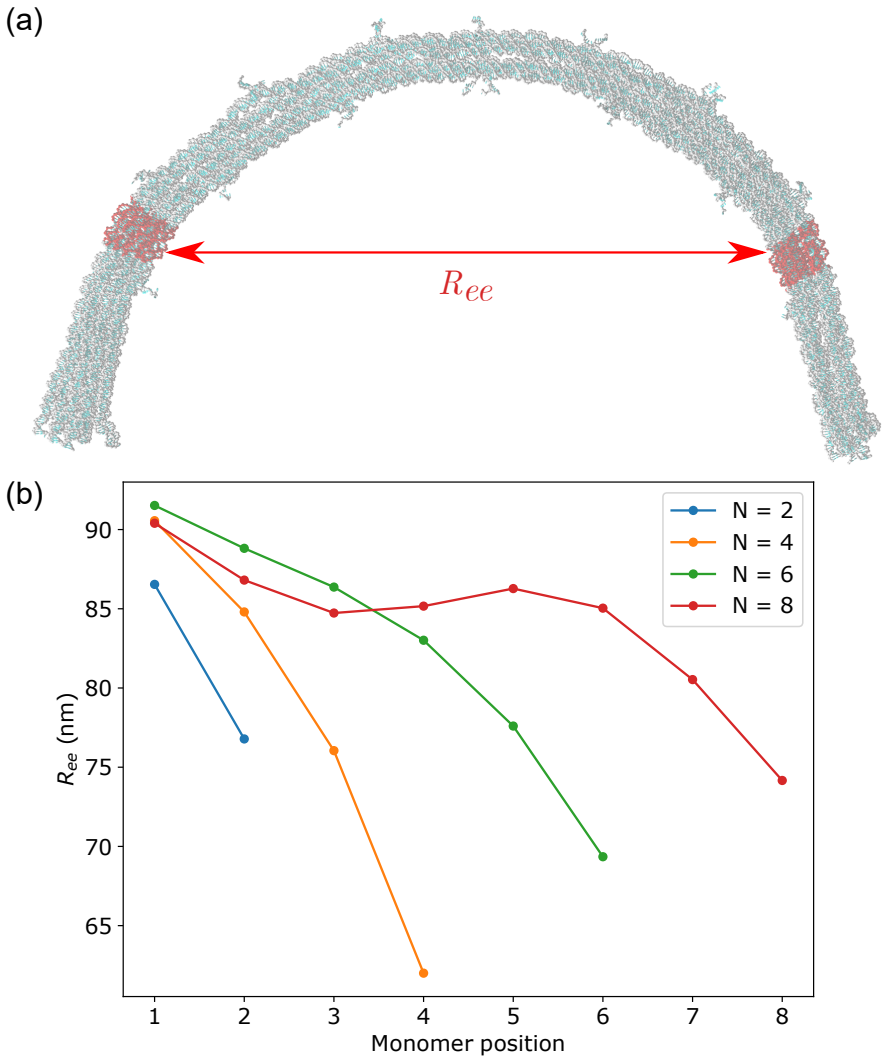


Figure 6.4: (a) Definition of the end-to-end distance, R_{ee}, of a monomer, illustrated for a configuration with $R_{ee} = 92\,\text{nm}$. The distance is defined with respect to the centres of mass of the shaded sets of nucleotides. The average end-to-end distance for a monomer at equilibrium is 92 nm. (b) Average end-to-end distance of monomers in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. While the lines generally show a decreasing trend from left to right, agreeing with the ideal model, the line for the largest assembly shows an increase for the middle monomers, indicating a deviation from ideal behaviour.

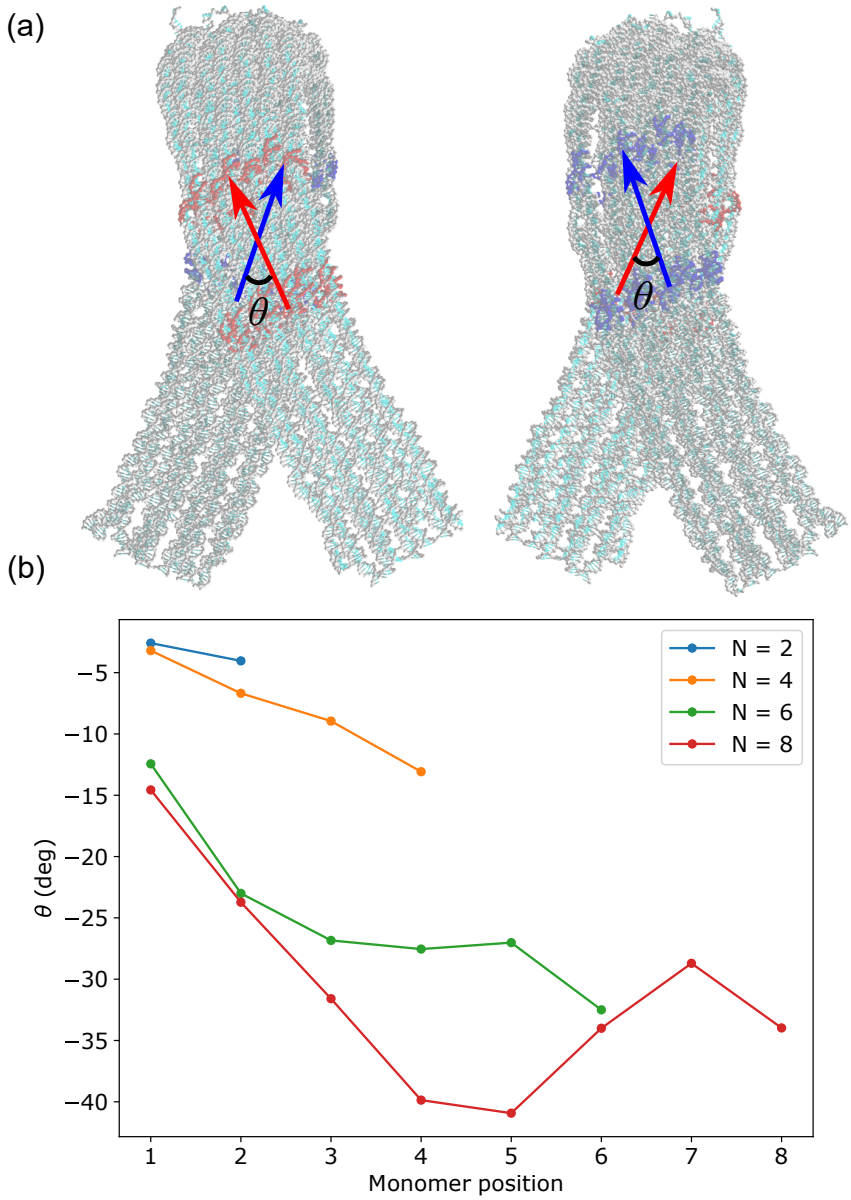


Figure 6.5: (a) Definition of the twist angle, θ, of a monomer, illustrated for a configuration with $\theta = +30°$. Two views of the monomer from its two sides are shown. The angle is defined by the two vectors coloured in red and blue, where the vectors are defined with respect to the centres of mass of the sets of nucleotides shaded in the corresponding colour. The average twist angle for a monomer at equilibrium is -3 degrees. (b) Average twist angle of monomers in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. Monomers in larger assemblies show a stronger twist and thus a stronger deviation from the ideal model, consistent with the observations in Figure 6.4.

twist in Figure 6.5(b) correspond to those which show a counter-intuitive increase in the end-to-end distance in Figure 6.4(b). This observation suggests that twisting is a plausible mechanism to reduce the bending stress in larger assemblies. Of course, the bending stress does not simply disappear; twisting also introduces stress into the system.

It is interesting to note that the assemblies appear to prefer left-handed twisting. Although twisting of the assembly in both directions could in principle alleviate some of the bending stress, left-twisted assemblies could be favoured because of the inherent small left-handed twist of the monomer.

Bond distance

A bond distance can be defined for each pair of sticky ends as the distance between the surface nucleotides that are near the sticky ends in the two interacting origamis. The bond distance as a function of the monomer position and the sticky end position within the assembly is shown as colour maps in Figure 6.6. The monomer position index is the largest for the monomer at the innermost position of the assembly. The sticky end position index increases from the left to right of the origami as defined in caDNAno. Each coloured square represents the bond distance of the sticky end with the index on the x-axis, between the monomers with the indices on the y-axis that bound the square.

The bond distances of the middle sticky ends are generally higher than the other sticky ends. This is evident from the colours of sticky ends 1 and 7, which tend towards the blue end of the spectrum, and that of sticky ends 3 to 5, which tend towards the red end of the spectrum. This variation allows the monomers to have a slightly reduced curvature while maintaining the bonding with neighbouring monomers, thereby reducing the bending stress at the cost of the bond distances deviating from their ideal values. The colour maps are not perfectly symmetric about the middle sticky end, potentially due to variations near the local environment of each sticky end. These variations could include the positions of nearby double crossover junctions, or the fraying of duplexes near the sticky ends.

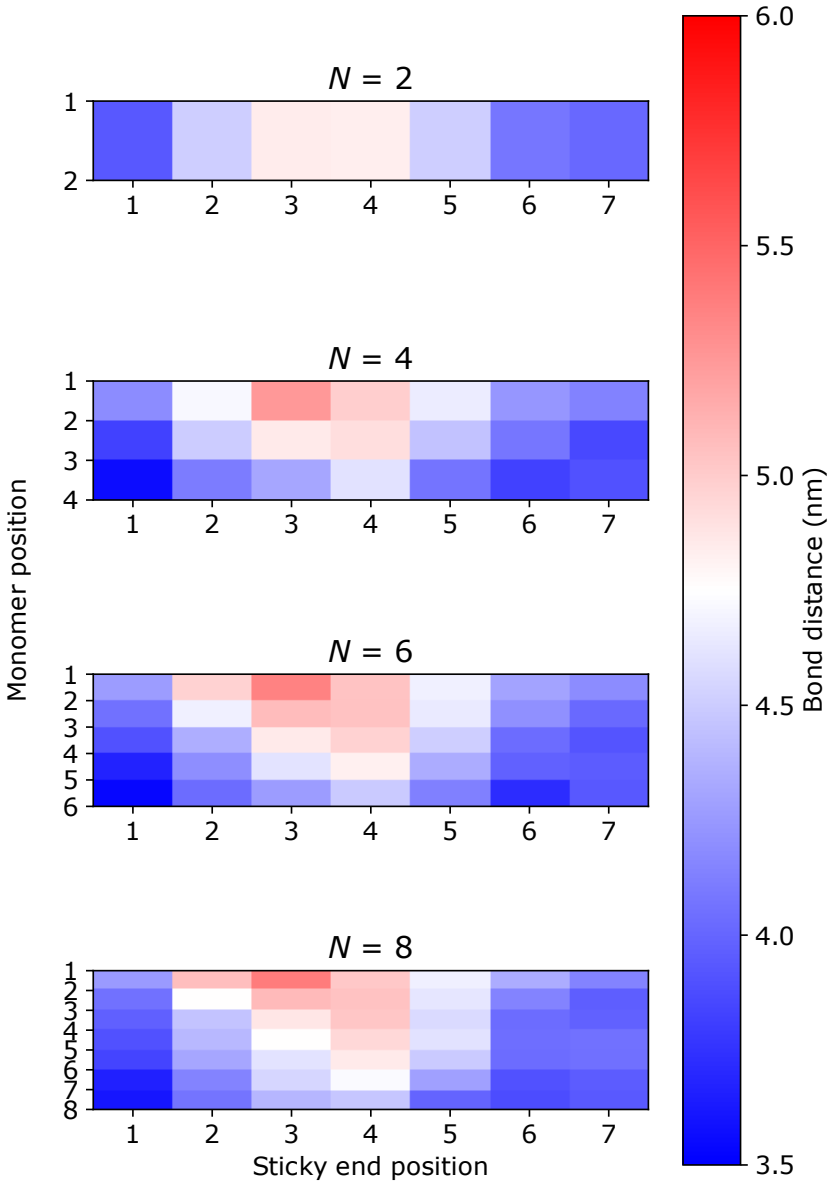


Figure 6.6: Average bond distance between monomers in assemblies of various sizes. Each square represents a bond between a pair of sticky ends of two neighbouring monomers. Colours towards the red end of the spectrum indicates stretched bonds, while those towards the blue end indicates compressed bonds. Bonds in the middle of the monomers tend to be more stretched than those at the ends. Outer, less curved monomers tend to have more stretched bonds compared to inner, more curved monomers.

The bond distances tend to decrease from top to bottom of the assembly, but the ranges of bond distances at each sticky end position are generally similar between assembly of different sizes. The decrease in bond distances at the bottom reflects that the corresponding monomers are the most bent.

6.3.2 Estimating the stress free energy

Having evaluated the basic structural properties of the assembly, including the end-to-end distances and twist angles of individual monomers within each assembly, we will proceed to understand how stress builds up in the assembly by estimating the stress free energy of each assembly. We will use the end-to-end distance and the twist angle as the order parameters, and assume that the stress free energy only comes from these two parameters.

We will also assume that the stress on the bonds between the monomers is small compared to the stress on the monomers themselves, so the contribution from stressed inter-monomer bonds will be neglected. Admittedly, it is hard to justify this assumption without calculating the stress involved, but this component of stress is difficult to calculate because the corresponding order parameter is not obvious. One could use the bond distance between each bonding site, which would result in seven order parameters. It would be quite expensive to sample each of them separately and obtain the corresponding free energy landscapes, not to mention the potential correlation between different order parameters. The bond distances are also sensitive to the local environment near the sticky ends, such as the possible fraying of proximal duplexes, and it would be challenging to account for these local variations to obtain a fair estimate of the bonding stress. Hence, we have chosen to focus on stress accumulation through bending and twisting.

Free energy landscape of bending

The free energy landscape of bending as a function of the end-to-end distance of the monomer, R_{ee}, as defined in Figure 6.4(a), was calculated using Umbrella Sampling and WHAM, and is shown in Figure 6.7.

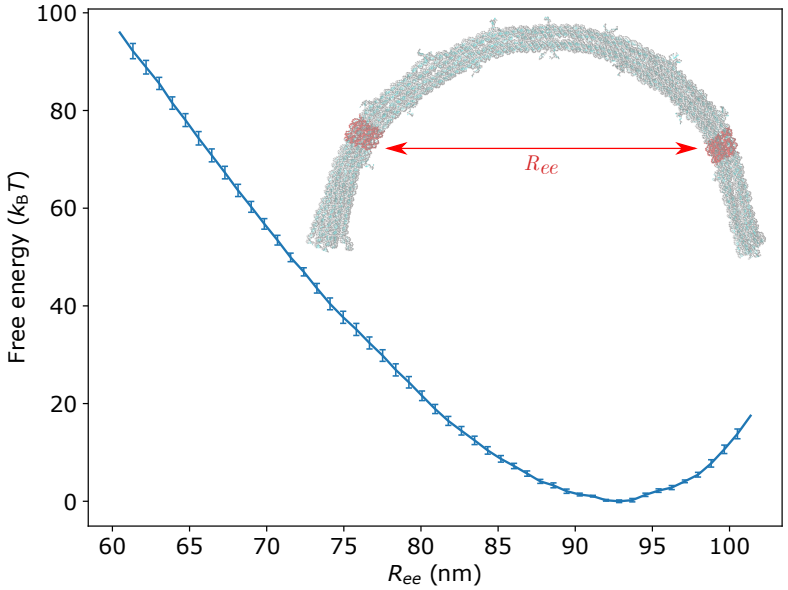


Figure 6.7: Free energy landscape of the curved monomer as a function of the end-to-end distance R_{ee}, as defined in Figure 6.4(a). Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

The landscape shows a minimum at around $R_{ee} = 93$ nm, which corresponds to the natural unstressed state of the monomer. In the sampled region of the end-to-end distance, the landscape appears to be smooth and suggests the absence of structural transitions such as kinking or buckling.

We can infer the bending stress free energy of each monomer in each assembly by mapping the end-to-end distances in Figure 6.4(b) to their corresponding free energy using the free energy landscape in Figure 6.7. The result of the mapping is shown in Figure 6.8.

The trend of the plots in Figure 6.8 roughly follow the opposite of the plots in Figure 6.4(b). In other words, the bending free energy generally increases for monomers at an inner position. This behaviour deviates from the ideal model, which predicts that the innermost monomer and the outermost monomer are the most stressed, being the most strongly bent and unbent respectively. This prediction is based on the assumption that the middle monomer is at its natural curvature, but this assumption appears to be incorrect for the current system, where the

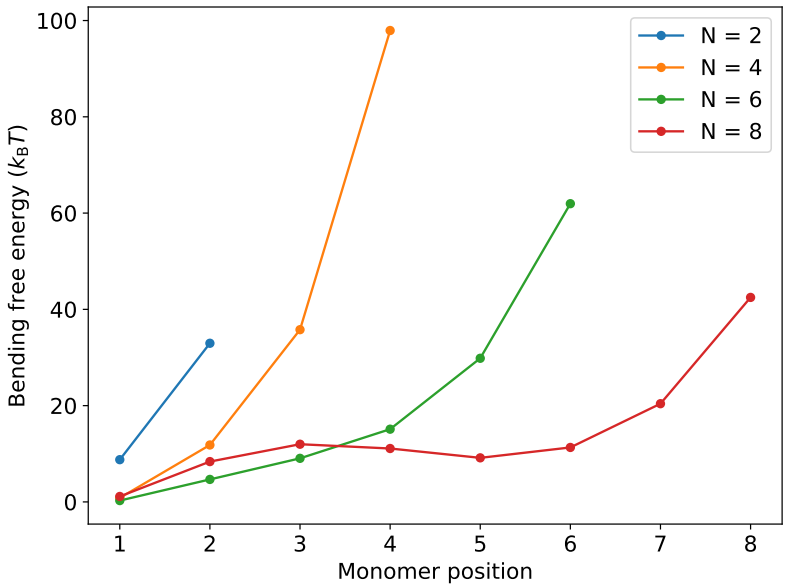


Figure 6.8: Bending stress free energy of monomers mapped from their end-to-end distances in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. The stress is higher at inner positions of the assembly, correlating with the larger curvature at those positions.

monomers in the assemblies are all more bent than the free monomer. One reason could be that the specific positions of the sticky ends in each monomer restrains their respective curvature, unlike the ideal model where the attractions between monomers are homogeneous and the monomers can slide past each other freely. Hence, in the current system, the stress will not be distributed in such a way that middle monomer relaxes to its natural curvature; rather, the stress distribution is decided by the assembly structure dictated by the positions of the sticky ends.

The average bending stress free energy for each assembly size can be straightforwardly calculated from Figure 6.8, and is shown in Figure 6.9. Unlike in the ideal model where the stress free energy continuously increases as a function of the assembly size, here the energy decreases when the size of the assembly becomes larger than 4 monomers. This trend indicates that the assembly process is not self-limiting if we only consider the bending stress accumulation.

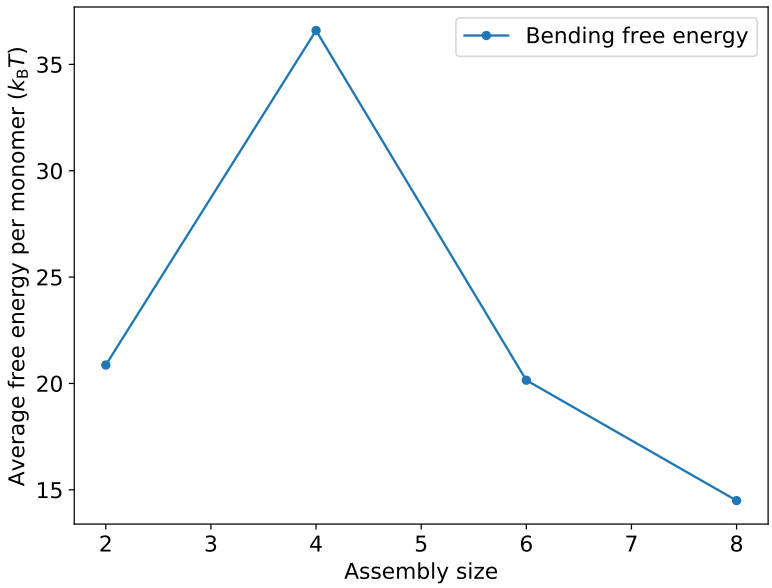


Figure 6.9: Average bending free energy of assemblies of various sizes. The average stress does not increase continuously as the assembly size increases, meaning that the assembly is not self-limiting if only the bending stress is considered.

Given that we also observed considerable twisting in larger assemblies, it is natural to extend the stress free energy analysis to twisting as well, which will be done in the following section.

Free energy landscape of twisting

The free energy landscape of twisting as a function of the twisting angle of the monomer, θ, as defined in Figure 6.5(a), was calculated by integrating the torque-angle plot for a simulation where the monomer was subject to an increasing torque, and is shown in Figure 6.10.

The landscape shows a minimum at around $\theta = 0$, which agrees with the fact that the monomer has no natural twist. In the sampled region of the twist angle, the landscape appears to be smooth and suggests the absence of structural transitions such as kinking or buckling.

We can infer the twisting stress free energy of each monomer in each assembly by mapping the twist angles in Figure 6.5(b) to their corresponding free energy

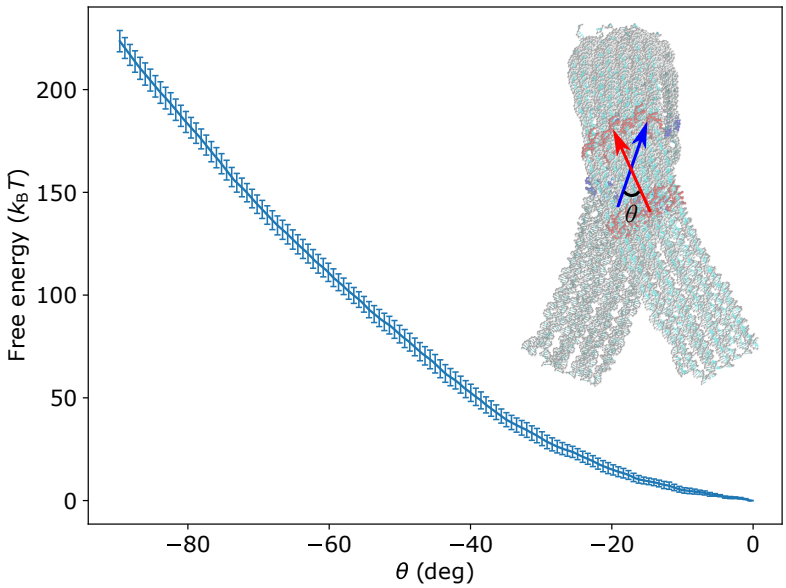


Figure 6.10: Free energy landscape of the curved monomer as a function of the twist angle θ, as defined in Figure 6.5(a). Error bars are calculated from 10 independent twisting simulations.

using the free energy landscape in Figure 6.10. The result of the mapping is shown in Figure 6.11.

It is worthwhile to compare the bending stress levels in Figure 6.8 and the twisting stress levels in Figure 6.11. For smaller assemblies with 2 or 4 monomers, the bending free energy is significantly larger than the twisting free energy, whereas for larger assemblies with 6 or 8 monomers, the stress levels due to bending and twisting are more similar in magnitude. This trend suggests a gradual shift of the stress accumulation mechanism from pure bending to a mixture of bending and twisting.

The average twisting stress free energy for each assembly size can be straightforwardly calculated from Figure 6.11, and the result is shown in Figure 6.12. The energy continuously increases as a function of the size of the assembly. This trend indicates that the twisting stress accumulation could contribute towards self-limiting growth of the assembly.

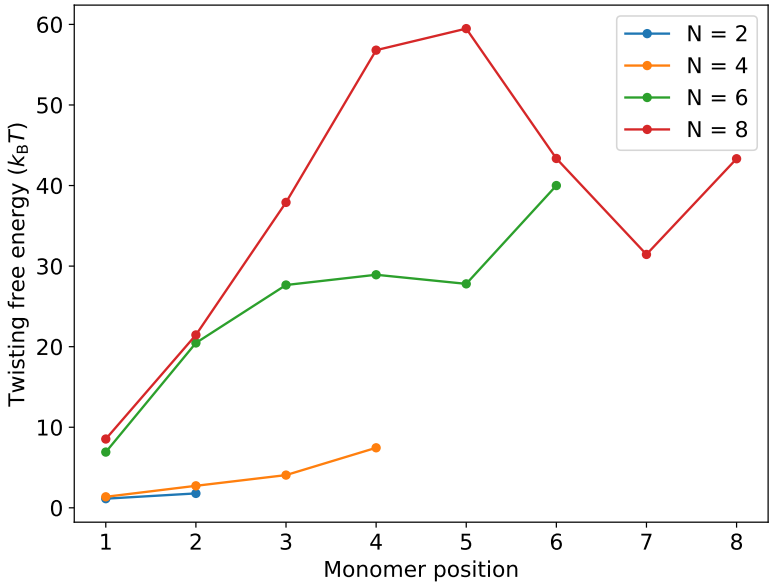


Figure 6.11: Twisting stress free energy of monomers mapped from their twist angles in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. The stress is higher for larger assemblies, correlating with their stronger twist.

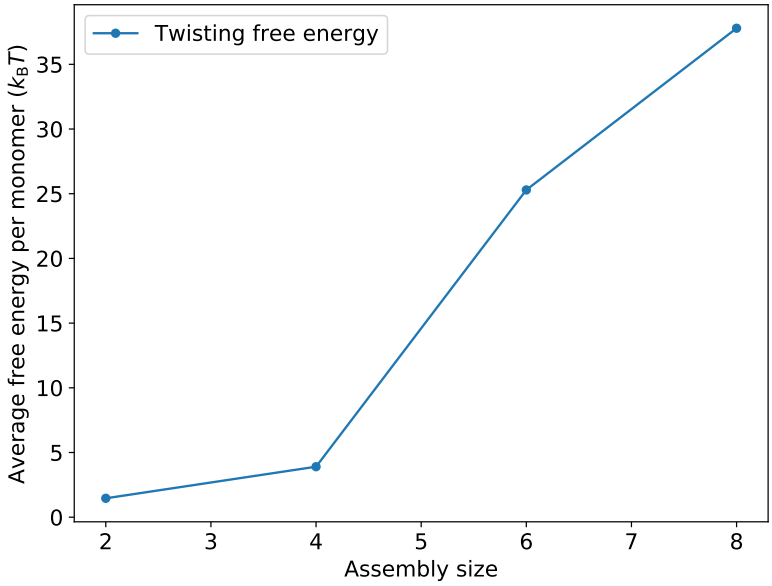


Figure 6.12: Average twisting free energy of assemblies of various sizes. The average stress increases continuously as the assembly size increases, suggesting that the twisting stress accumulation could contribute to the self-limiting growth of the assembly.

Effect of twist-bend coupling

The preceding analysis assumes that the bending and twisting free energy landscapes are independent of each other and can be considered separately. In this section, we will explore the validity of this assumption.

The free energy landscape of twisting as a function of the twisting angle of the monomer, θ, as defined in Figure 6.5(a), was calculated by integrating the torque-angle plot for a simulation where the monomer was subject to an increasing torque. Unlike the preceding section, here the end-to-end distance was also restrained during the twisting simulation. The distance was restrained at three different values, and was also unrestrained in a separate control simulation. The free energy landscapes are shown in Figure 6.13.

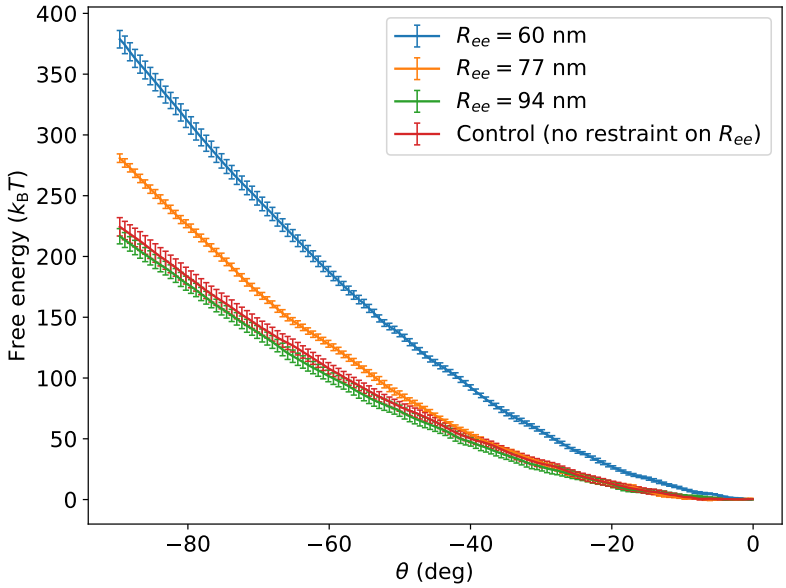


Figure 6.13: Twisting free energy landscapes as a function of the twist angle θ, as defined in Figure 6.5(a), when the curved monomer is restrained at different end-to-end distances. Error bars are calculated from 10 independent twisting simulations for each value of the restrained end-to-end distance.

As expected, the landscape for the simulation where the end-to-end distance was restrained at its equilibrium value is almost identical to the landscape for the control simulation without any restraint. When the end-to-end distance is

restrained at 77 nm, the landscape remains largely the same until $\theta = -40$ degrees. At an even smaller end-to-end distance of 60 nm, the landscape begins to deviate from the control at the beginning.

Although the plots in Figure 6.13 clearly show a sign of twist-bend coupling at strong bending or twisting, it is worth noting that the majority of the monomers in the previously analysed assemblies belongs to the part where the effect of twist-bend coupling is negligible. Only the stress free energy estimation for the innermost monomer in the assemblies of size 4, 6, and 8 will be affected, as the other monomers all have end-to-end distances above 77 nm and twist angles below -40 degrees. Given that only one monomer is affected in each assembly, and that it will be challenging and expensive to sample a two-dimensional free energy landscape as a function of the twist angle and the end-to-end distance, we will not consider the effect of twist-bend coupling in the analysis.

Total stress free energy

The average bending, twisting and total stress free energy as derived in the preceding sections are shown in Figure 6.14.

Although the average bending stress free energy decreases for larger assemblies, it can be somewhat compensated by the stronger increase in average twisting stress free energy, resulting in an overall increasing trend of the average total stress free energy as a function of the assembly size. However, the increase in average stress energy is quite mild from size 4 to size 8. There is only a 30% growth in average stress energy for an 100% increase in assembly size. This contrasts with the strong increase in average stress energy of about 100% from size 2 to size 4.

Overall, the assembly of curved monomers shows a weak sign of being self-limiting. The major difficulty in maintaining stress accumulation is due to out-of-plane twisting of the monomers, which allows them to unbend and avoid the bending cost associated with assembly growth. Although the average twisting stress energy also accumulates, it can only barely recover the loss in the average

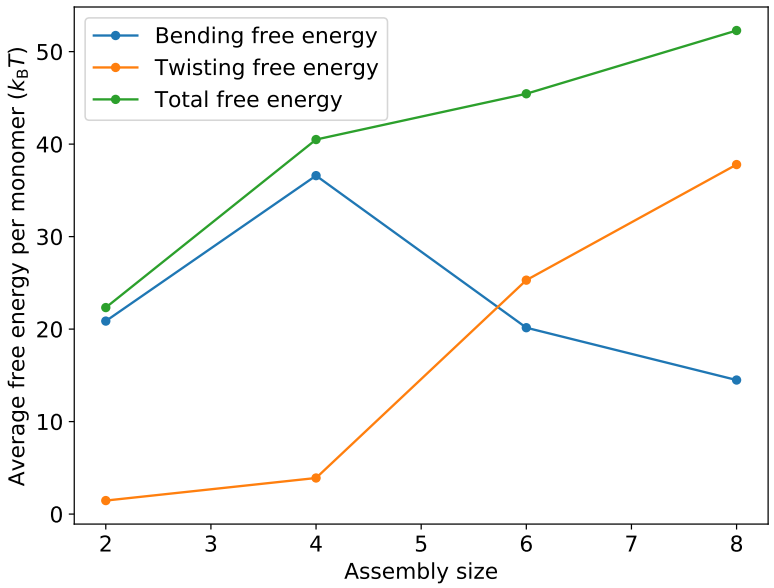


Figure 6.14: Average bending, twisting and total free energy of assemblies of various sizes. The average total stress is increasing as a function of the assembly size, albeit the increase is weak at larger assembly sizes. The assembly process, therefore, shows a sign of being self-limiting.

bending stress energy, resulting in a weakly increasing average stress energy as a function of the assembly size.

6.4 Curved Monomer: Design 2

To improve upon the first design, we aim to prevent twisting while still allowing bending. An obvious solution is to widen the origami so that it has a sheet-like structure rather than a rod-like structure. We therefore chose a 30-helix bundle origami as a second design. As the number of helices increased but the size of the scaffold is unchanged, the length of the origami will have to decrease. In order to maintain a similar curvature as before, more base insertions and deletions have to be included. The shorter length also means that fewer sticky ends can be placed, so we reduced the number from 21 to 15 along the helix. The sticky ends are arranged in a 5 by 3 manner, with 5 sticky ends on 3 different helices. Each pair of sticky ends consist of 7 base pairs. When positioning the sticky ends, we aimed

to roughly match the curvature of the monomer.

6.4.1 Assembly structure

In this section, we will look at the structure of the assembly by calculating the end-to-end distance of each monomer and the bond distance between monomers in assemblies of different sizes. Example configurations of the simulated assemblies are shown in Figure 6.15.

End-to-end distance

The average end-to-end distance as a function of the monomer position within the assembly is shown in Figure 6.16. The monomer position index is the largest for the monomer at the innermost position of the assembly.

For all the simulated assemblies, the end-to-end distance decreases continuously from the outermost to the innermost monomer. However, the decrease near the middle of the assembly becomes more gentle as the assembly size increases. For the largest assembly with 8 monomers, there are 4 monomers (at positions 2 to 5) with very similar end-to-end distances at around 60 nm, indicating that these monomers have very similar curvatures. In other words, curvature is not well propagated between these monomers. This observation may be due to other modes of stress accumulation that counters bending, such as stretching of the sticky end bonds between monomers, which allow monomers to bind with a reduced level of bending stress.

Bond distance

A bond distance can be defined for each pair of sticky ends as the distance between the surface nucleotides that are near the sticky ends in the two interacting origamis. The bond distance as a function of the monomer position and the sticky end position within the assembly is shown as colour maps in Figure 6.17. The monomer position index is the largest for the monomer at the innermost position of the assembly. The sticky end position index increases from the left to right of the origami as defined in caDNAno. Each coloured square represents the bond distance of the

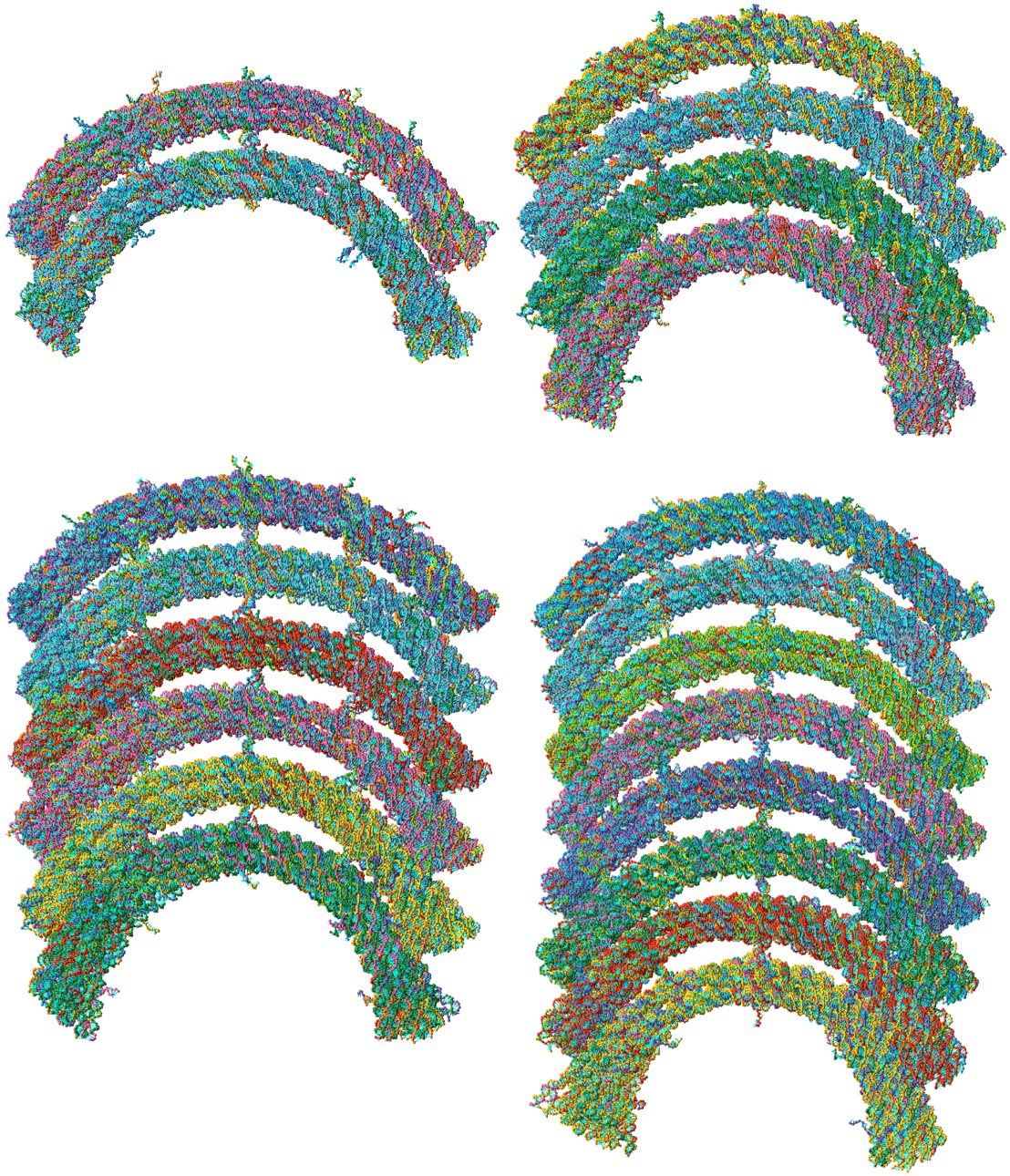


Figure 6.15: oxDNA simulation configurations of the assembly of curved monomers design 2. The sizes of the assemblies are 2, 4, 6 and 8, from left to right and top to bottom. The twist in these structures are much smaller than those of design 1 in Figure 6.2.

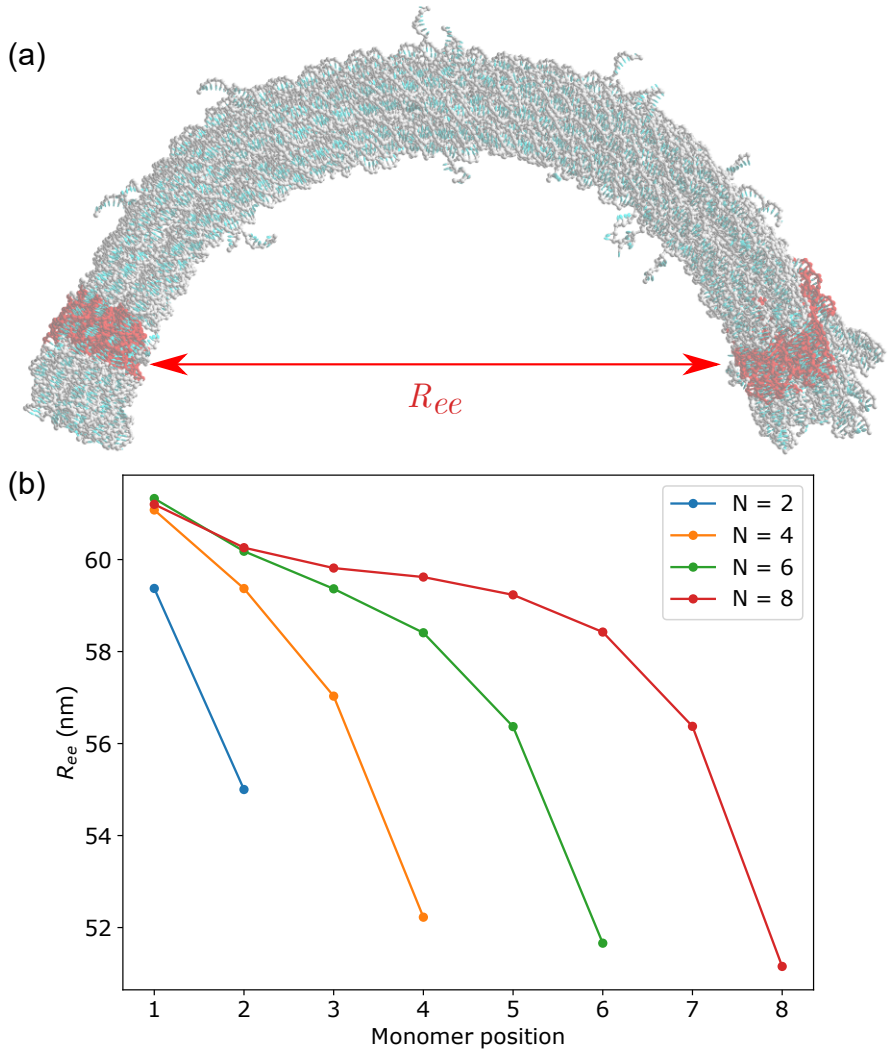


Figure 6.16: (a) Definition of the end-to-end distance, R_{ee}, of a monomer, illustrated for a configuration with $R_{ee} = 58\,\text{nm}$. The distance is defined with respect to the centres of mass of the shaded sets of nucleotides. The average end-to-end distance for a monomer at equilibrium is 58 nm. (b) Average end-to-end distance of monomers in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. The lines generally show a decreasing trend from left to right, in agreement with the ideal model, although the decrease is rather small for the middle monomers in the largest assembly.

sticky end with the index on the x-axis, between the monomers with the indices on the y-axis that bound the square.

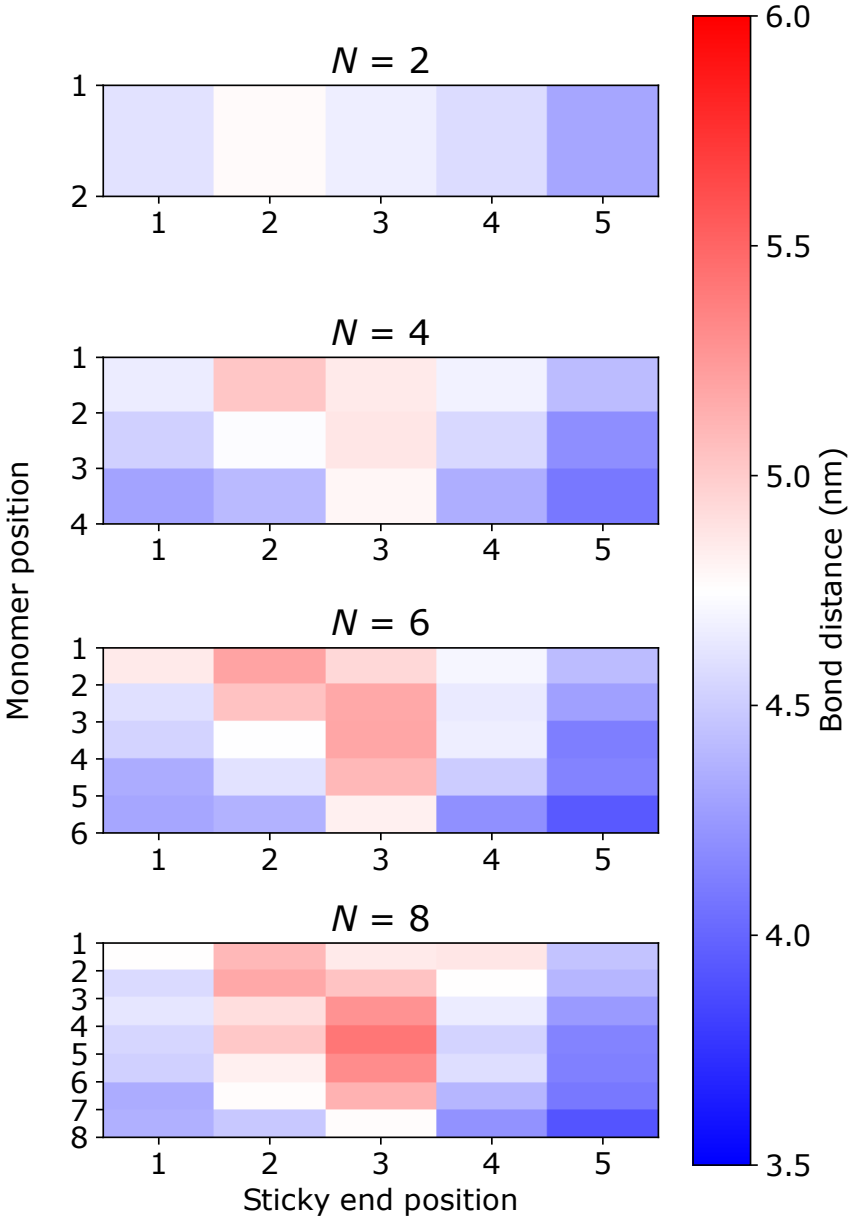


Figure 6.17: Average bond distance between monomers in assemblies of various sizes. Each square represents a bond between a pair of sticky ends of two neighbouring monomers. Colours towards the red end of the spectrum indicates stretched bonds, while those towards the blue end indicates compressed bonds. Bonds in the middle of the monomers tend to be more stretched than those at the ends. Outer, less curved monomers tend to have more stretched bonds compared to inner, more curved monomers.

The trends in the bond distances in Figure 6.17 are similar to those in Figure 6.6 for design 1, where the bond distances of the middle sticky ends are generally

higher than the other sticky ends. Similar to design 1, the variation allows the monomers to slightly reduce their curvature and thus their bending stress. The variation can be seen in Figure 6.15 where there is a bigger gap between monomers in the centre. Asymmetry of the colour maps is also evident, which may be caused by local variations near the sticky ends, such as the positions of nearby crossover junctions or fraying duplexes.

As the assembly grows, there are more squares in the colour map that tend towards the red end of the spectrum, especially the squares for sticky ends 2 and 3. This indicates that more bonds between monomers are stretched as the assembly grows, and suggests that bond stretching is also a potential mode of stress accumulation in this system.

6.4.2 Estimating the stress free energy

We will attempt to understand stress build-up in the assembly by estimating the stress free energy of each assembly. As the bond distance analysis showed, bond stretching between monomers becomes stronger as the assembly grows. Indeed, this process has its own free energy cost, but it is challenging to estimate its contribution to the stress free energy because the bond distances are strongly affected by local environments, which are highly variable and difficult to represent with simple order parameters. Therefore, similar to the first curved monomer design, we will focus on the contribution to the stress free energy from the end-to-end distance order parameter.

Free energy landscape of bending

The free energy landscape of bending as a function of the end-to-end distance of the monomer, R_{ee}, as defined in Figure 6.16(a), was calculated using Umbrella Sampling and WHAM, and is shown in Figure 6.18.

The landscape shows a minimum at around $R_{ee} = 58$ nm, which corresponds to the natural unstressed state of the monomer. In the sampled region of the

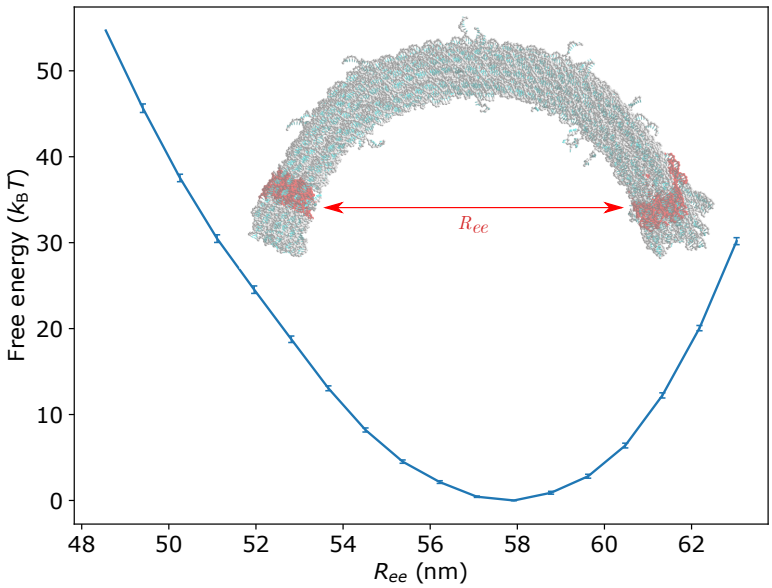


Figure 6.18: Free energy landscape of the curved monomer as a function of the end-to-end distance R_{ee}, as defined in Figure 6.16(a). Error bars are calculated by dividing the set of trajectories into 10 parts and calculating the free energy landscape for each part.

end-to-end distance, the landscape appears to be smooth and suggests the absence of structural transitions such as kinking or buckling.

We can infer the bending stress free energy of each monomer in each assembly by mapping the end-to-end distances in Figure 6.16(b) to their corresponding free energy using the free energy landscape in Figure 6.18. The result of the mapping is shown in Figure 6.19.

The plots in Figure 6.19 show that the middle monomers have lower stress levels compared to the innermost and outermost monomers. This agrees with Figure 6.16(b) where the middle monomers are closer to their unstressed end-to-end distance of 58 nm. The lack of curvature accumulation for larger assemblies shown in Figure 6.16(b) is also evident in Figure 6.19, where the increase in bending stress free energy from the middle to the outermost monomer becomes slower and slower as the assembly becomes larger. In other words, the lack of curvature accumulation translates to a lack of bending stress free energy accumulation.

The average bending stress free energy for each assembly size can be straight-

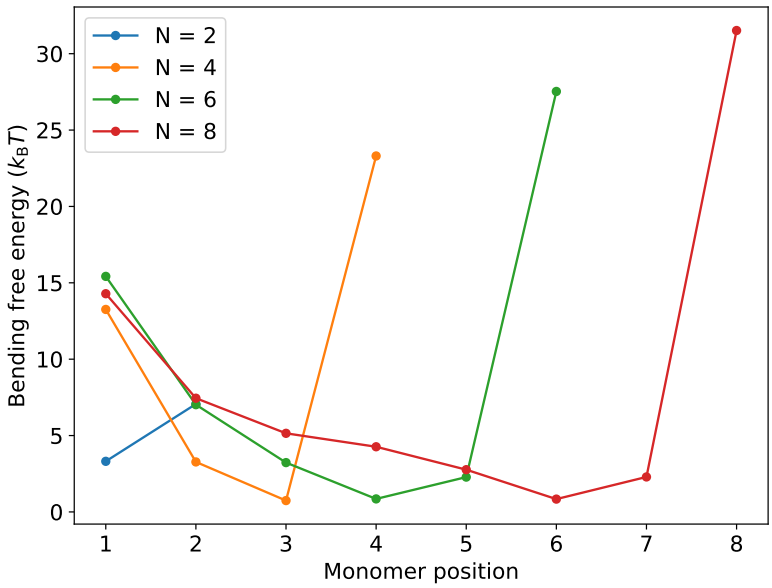


Figure 6.19: Bending stress free energy of monomers mapped from their end-to-end distances in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. The stress is higher at inner and outer positions of the assembly and lower at middle positions. This trend suggests that the middle monomers are closest to their natural curvature.

forwardly calculated from Figure 6.19, and is shown in Figure 6.20. The average stress free energy decreases when the size of the assembly becomes larger than 4 monomers, which suggests that the assembly process is not self-limiting if we only consider the bending stress accumulation.

However, as we have noted in the structural analysis, the sticky ends are more strongly stretched in larger assemblies (Figure 6.17). Although we are unable to calculate the contribution to the stress free energy from these modes of stress accumulation due to methodological difficulties, qualitatively, we could reasonably expect that their contribution increases as a function of assembly size in the range that we have studied.

We can draw an analogy to the stress accumulation for the first design in Figure 6.14, where the major stress accumulation mode gradually shifts from bending to twisting as the assembly grows. It could also be possible that the stress accumulation

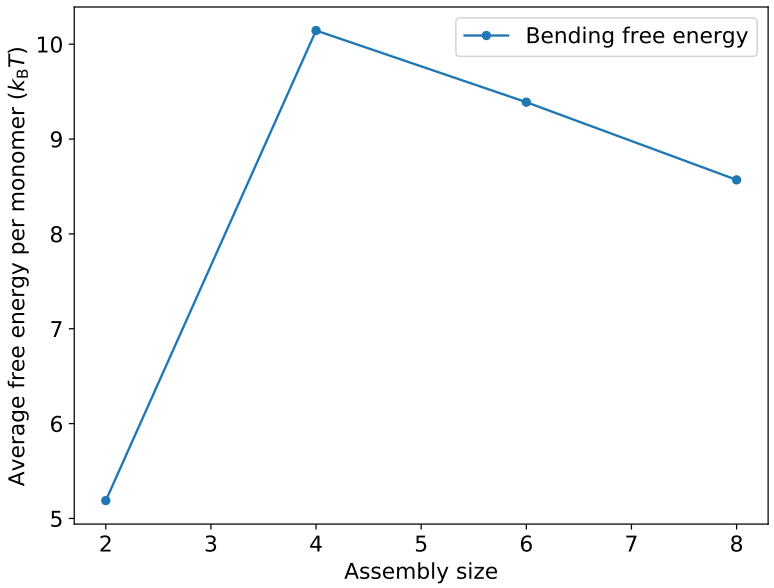


Figure 6.20: Average bending free energy of assemblies of various sizes. The average stress does not increase continuously as the assembly size increases, meaning that the assembly is not self-limiting if only the bending stress is considered.

mode shifts from bending to bond stretching in this design, and that when their stress contributions are combined, the assembly might become self-limiting, as is the case for the first design.

6.5 Twisted Monomer

In the assembly of twisted monomers, the monomers have to untwist in order to adapt to the shape of the neighbouring monomers. As the assembly grows, all the monomers have to untwist and deviate further from their natural twist in order to allow new monomers to bind. This increasing deviation gives rise to an increasing average stress as a function of the assembly size, and thus results in self-limiting assembly.

Our design is based on the curved 18-helix bundles in Dietz *et al.* [7]. To make twisted instead of curved bundles, we converted all base deletions to insertions and kept the original insertions unchanged. Also, since the base insertions and deletions

were only present in the middle part of the original origami design, we distributed the base insertions over the entire origami to achieve a more uniform twist.

Similar to the curved monomers, the inter-monomer interactions will be based on single-stranded DNA overhangs, i.e. "sticky ends", that perpendicularly emerge from two sides of the origami. 21 sticky ends are placed on the origami, arranged in a 7 by 3 manner, with 7 sticky ends on 3 different helices. Each pair of sticky ends consist of 7 base pairs. The sticky ends are positioned at equivalent positions on the top and bottom surfaces of the origami, which would result in unstressed binding of untwisted monomers.

6.5.1 Assembly Structure

In this section, we will look at the structure of the assembly by calculating the twist angle and end-to-end distance of each monomer in assemblies of different sizes. Example configurations of the simulated assemblies are shown in Figure 6.21. The assemblies have a similar saddle shape to the double-layer twisted origami sheet in Wong *et al.* [94].

Twist angle

The average twist angle as a function of the monomer position within the assembly is shown in Figure 6.22. The monomer position index is the largest for the monomer at the innermost position of the assembly.

The twist angles are smaller for monomers in the centre of an assembly compared to those on the edges. This is because due to the shape constraint, any twist in the centre of an assembly will be propagated towards both edges of the assembly, so naturally the monomers near the edges are more twisted than those in the centre.

The twist angle of all interior monomers (i.e. those not at the edge) decreases as the assembly grows. Again, this can be attributed to the shape constraint of the assembly. For new monomers to bind to an existing assembly, all the monomers in the existing assembly have to untwist in order to make way for the incoming monomer.

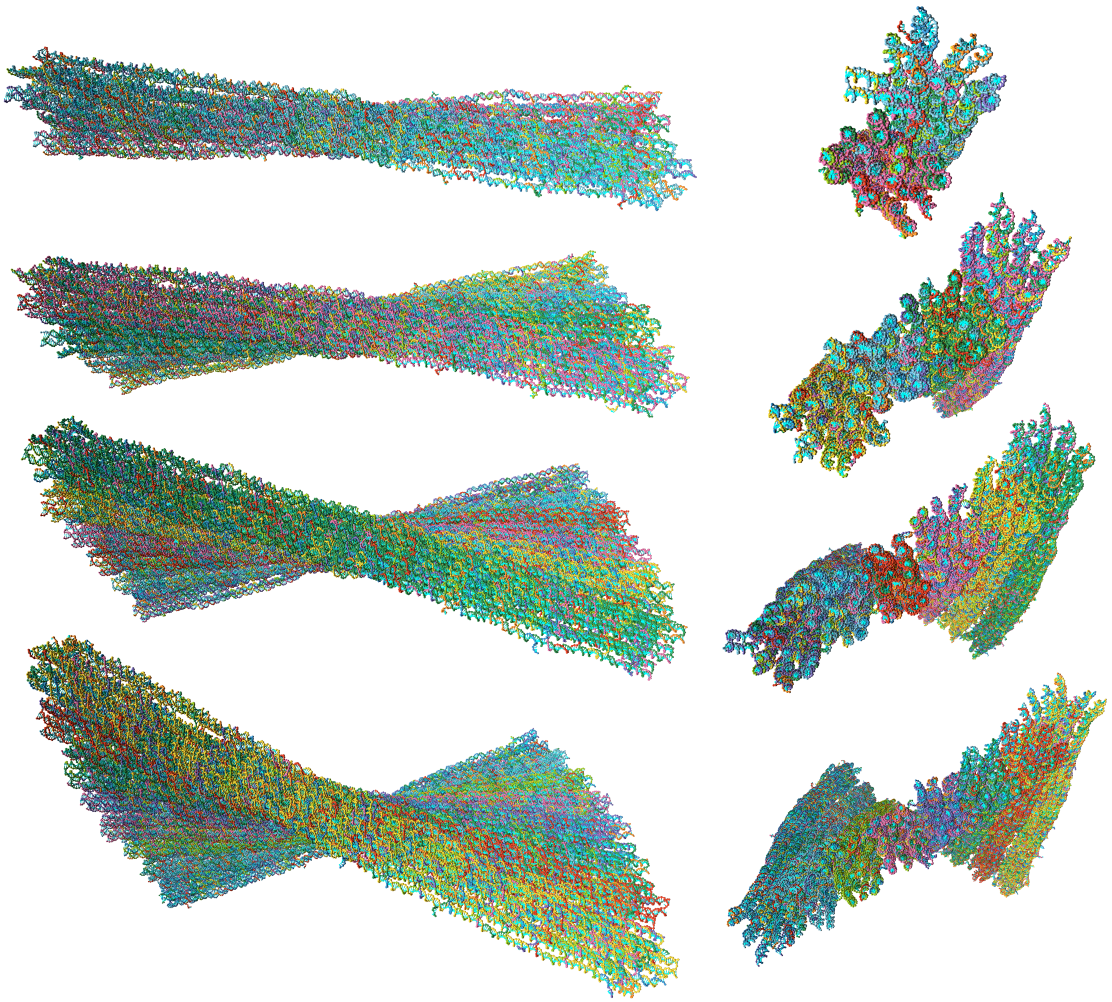


Figure 6.21: oxDNA simulation configurations of the assembly of twisted monomers. The sizes of the assemblies are 2, 4, 6 and 8, from top to bottom. The left column shows the side-on views while the right column shows the end-on views.

End-to-end distance

The average end-to-end distance as a function of the monomer position within the assembly is shown in Figure 6.23. The monomer position index is the largest for the monomer at the innermost position of the assembly.

The end-to-end distances show a small but noticeable variation across monomers in each assembly. The end-to-end distance is slightly smaller for the monomers on the edge of an assembly compared to those in the centre, possibly due to the stronger untwisting of the central monomers which increase their lengths. Another reason for the difference is that the monomers near the edge are slightly bent in

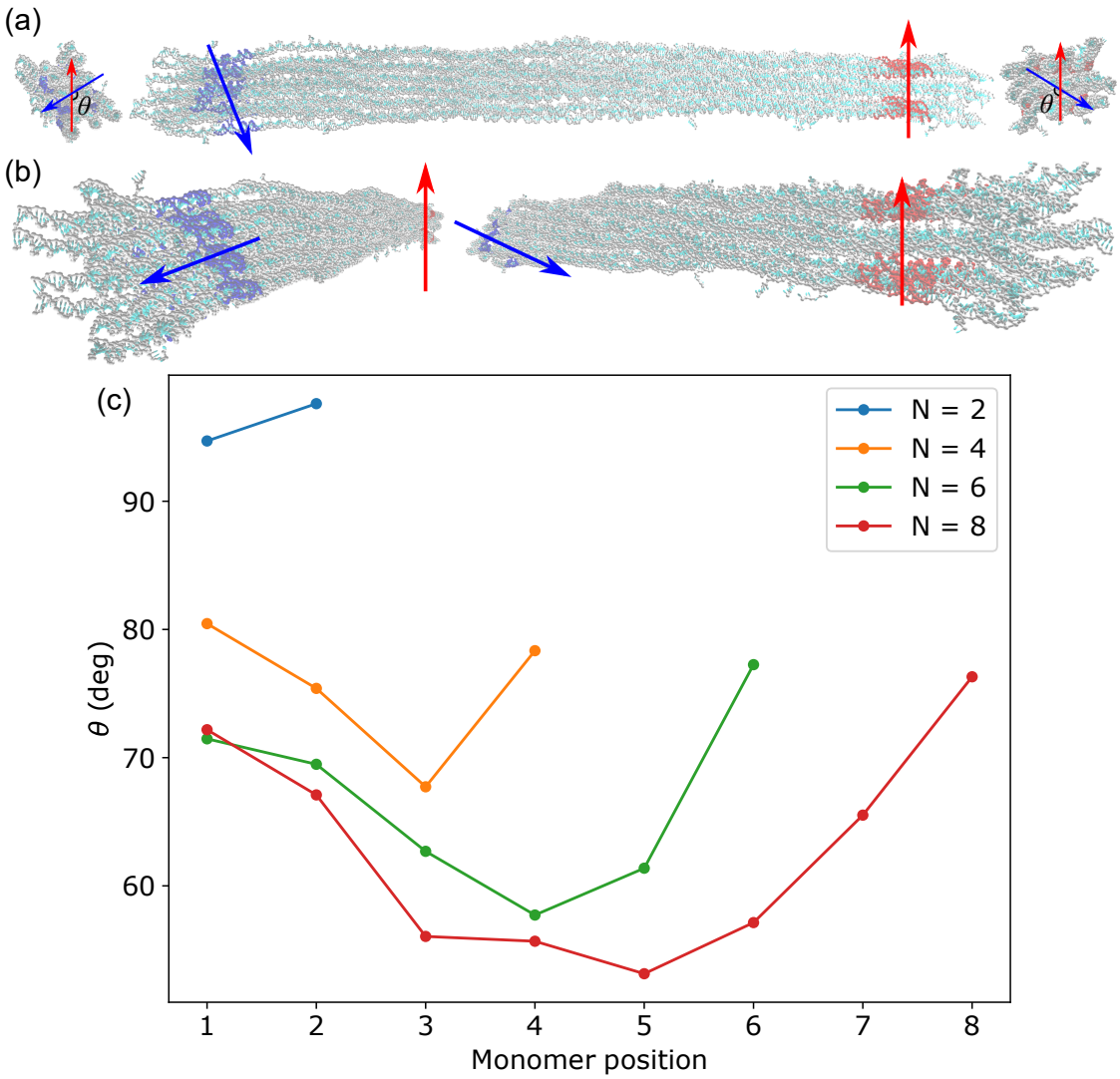


Figure 6.22: (a) Definition of the twist angle, θ, of a monomer, illustrated for a configuration with $\theta = 109°$. Two end-on views of the monomers are shown on the left and right, and the side-on view is shown in the middle. The angle is defined by the two vectors coloured in red and blue, where the vectors are defined with respect to the centres of mass of the sets of nucleotides shaded in the corresponding colour. The average twist angle for a monomer at equilibrium is 109 degrees. (b) More perspectives of the monomer and the vectors used for defining the twist angle. (c) Average twist angle of monomers in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. Monomers in larger assemblies are more untwisted than those in smaller assemblies. Monomers in the centre of assemblies are also more untwisted than those on the edges.

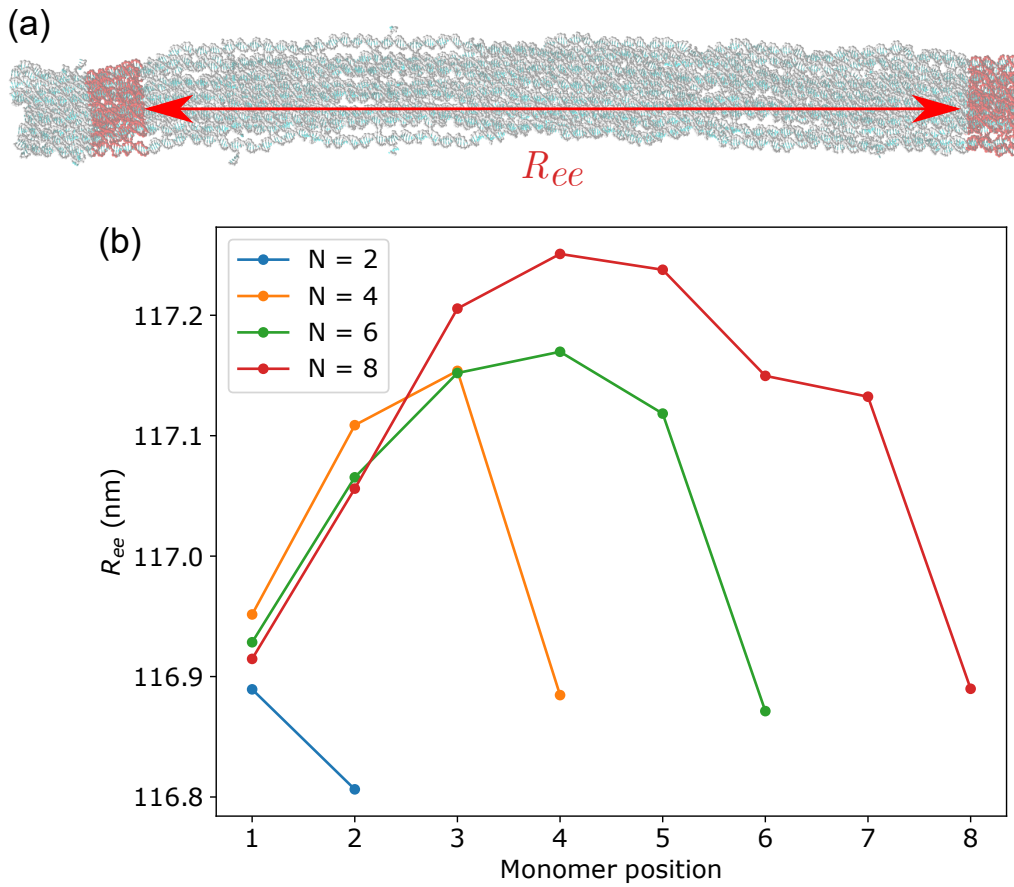


Figure 6.23: (a) Definition of the end-to-end distance, R_{ee}, of a monomer, illustrated for a configuration with $R_{ee} = 117\,\text{nm}$. The distance is defined with respect to the centres of mass of the shaded sets of nucleotides. The average end-to-end distance for a monomer at equilibrium is 117 nm. (b) Average end-to-end distance of monomers in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. There is a small variation among the monomers, with those near the edge of the assembly having a smaller R_{ee} than those in the middle.

order to wrap themselves around the assembly.

Bond distance

A bond distance can be defined for each pair of sticky ends as the distance between the surface nucleotides that are near the sticky ends in the two interacting origamis. The bond distance as a function of the monomer position and the sticky end position within the assembly is shown as colour maps in Figure 6.24. The monomer position index is the largest for the monomer at the innermost position of the assembly.

The sticky end position index increases from the left to right of the origami as defined in caDNAno. Each coloured square represents the bond distance of the sticky end with the index on the x-axis, between the monomers with the indices on the y-axis that bound the square.

The bond distances of the middle sticky ends are generally lower than the other sticky ends. As mentioned in the design, the sticky ends are positioned such that it would result in unstressed binding of untwisted monomers. Since the actual monomers are twisted, sticky ends further away from the middle are likely to experience shear stress due to twisting, resulting in longer bond distances.

6.5.2 Estimating the stress free energy

We will estimate the stress free energy of each assembly in order to understand how stress builds up in the assembly. We will use the twist angle as the order parameter and assume that the stress free energy only comes from this order parameter. As discussed in the corresponding section for curved monomers, due to difficulties in defining an order parameter which is feasible to sample, we will neglect the contribution to free energy due to stressed bonds.

Free energy landscape of twisting

The free energy landscape of twisting as a function of the twisting angle of the monomer, θ, as defined in Figure 6.22(a), was calculated by integrating the torque-angle plot for a simulation where the monomer was subject to an increasing torque, and is shown in Figure 6.25.

The landscape shows a minimum at around $\theta = 115°$, which corresponds to the natural twist of the monomer. In the sampled region of the twist angle, the landscape appears to be smooth and suggests the absence of structural transitions such as plectoneme formation.

The end-to-end distance of the monomer is recorded during the simulation and plotted against the twist angle in Figure 6.26. The plot confirms that untwisting

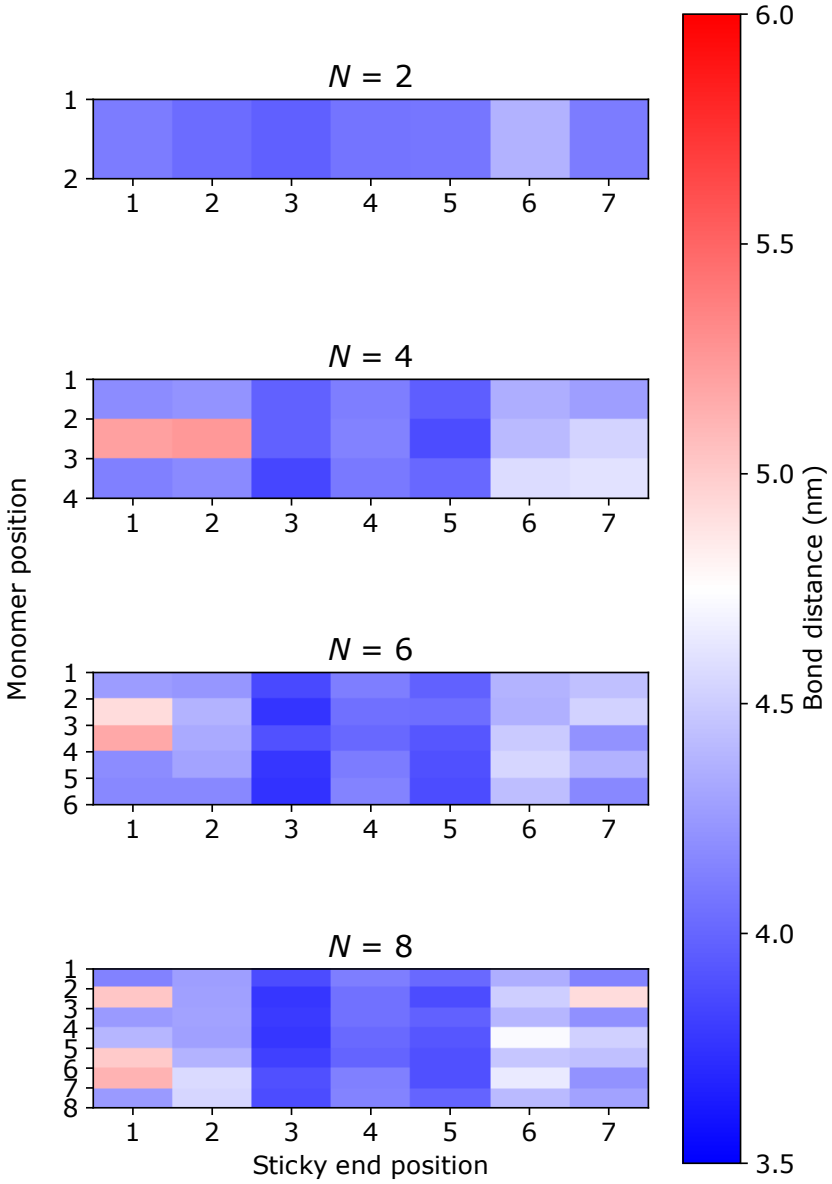


Figure 6.24: Average bond distance between monomers in assemblies of various sizes. Each square represents a bond between a pair of sticky ends of two neighbouring monomers. Colours towards the red end of the spectrum indicates stretched bonds, while those towards the blue end indicates compressed bonds. Bonds in the middle of the monomers tend to be less stretched than those at the ends due to the shear stress induced by the positions of the sticky ends which facilitate binding of untwisted monomers.

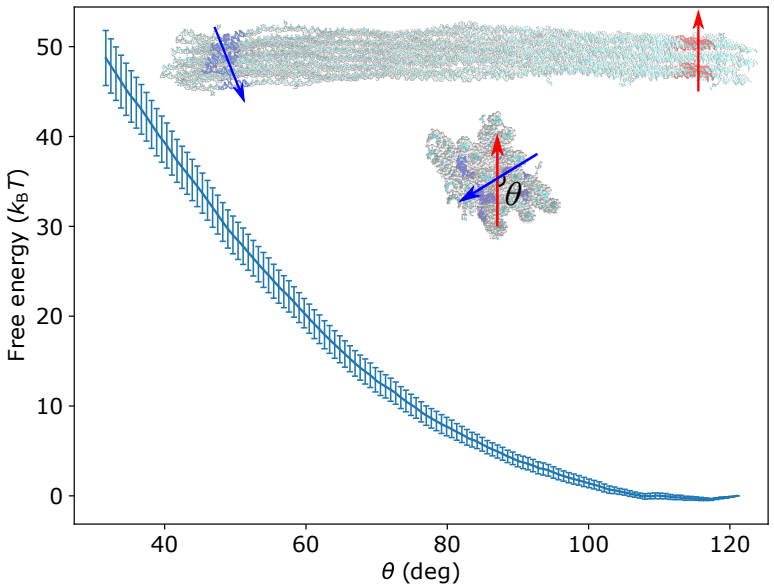


Figure 6.25: Free energy landscape of the twisted monomer as a function of the twist angle θ, as defined in Figure 6.22(a). Error bars are calculated from 10 independent twisting simulations.

the monomer results in an increase in its end-to-end distance, as expected in the discussion accompanying Figure 6.23.

We can infer the twisting stress free energy of each monomer in each assembly by mapping the twist angles in Figure 6.22(b) to their corresponding free energy using the free energy landscape in Figure 6.25. The result of the mapping is shown in Figure 6.27.

The trend of the plots in Figure 6.27 roughly follow the opposite of the plots in Figure 6.22(b). Central monomers are more strongly untwisted and thus have a higher twisting stress free energy, while monomers near the edges are closer to their natural twist and therefore less stressed.

The average twisting stress free energy for each assembly size can be straightforwardly calculated from Figure 6.27, and is shown in Figure 6.28. The energy continuously increases as a function of the size of the assembly, which is an indication of self-limiting growth of the assembly by the accumulation of twisting stress.

Contrary to the curved monomers, the assembly of twisted monomers shows a

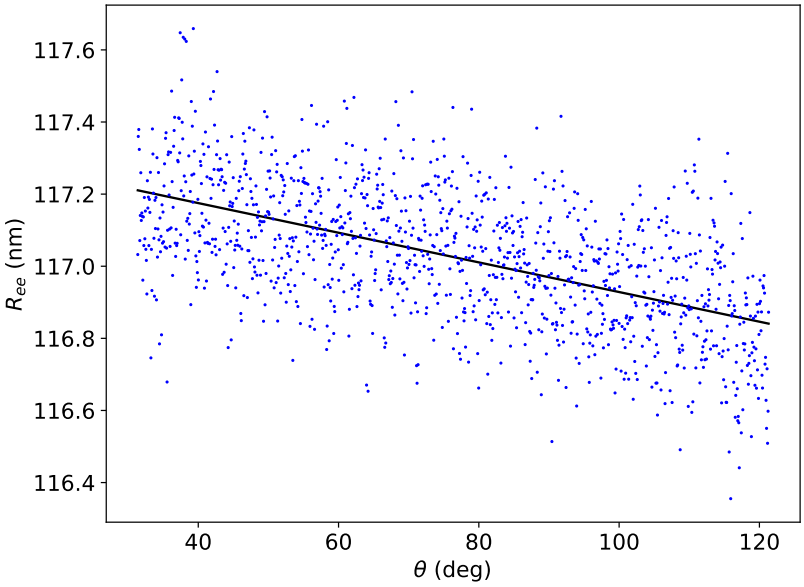


Figure 6.26: End-to-end distance R_{ee} of the monomer plotted against the twist angle θ during the twisting simulation, and the best-fit straight line for the data points. Untwisting the monomer results in a small increase in its end-to-end distance.

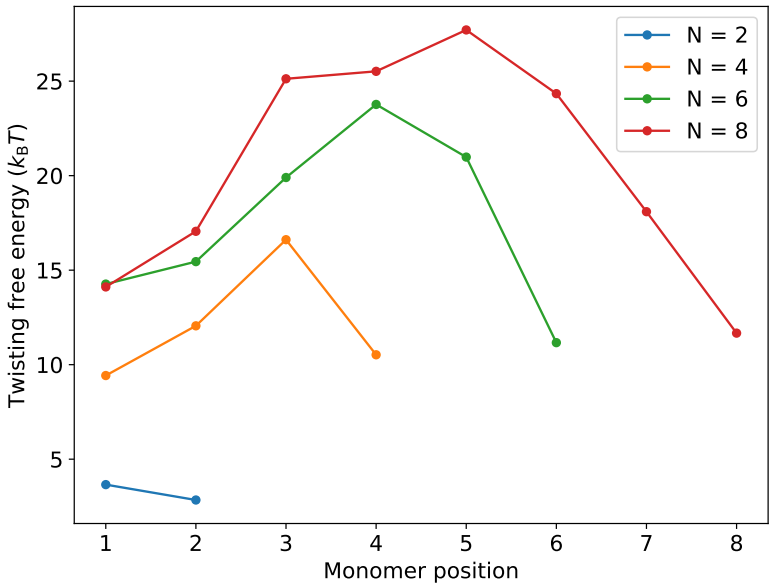


Figure 6.27: Twisting stress free energy of monomers mapped from their twist angles in assemblies of various sizes. The monomer position index is the largest for the monomer at the innermost position of the assembly. Each point corresponds to a monomer inside an assembly, while each line connects the points corresponding to an assembly of a given size. The stress is higher for larger assemblies, correlating with their stronger untwisting.

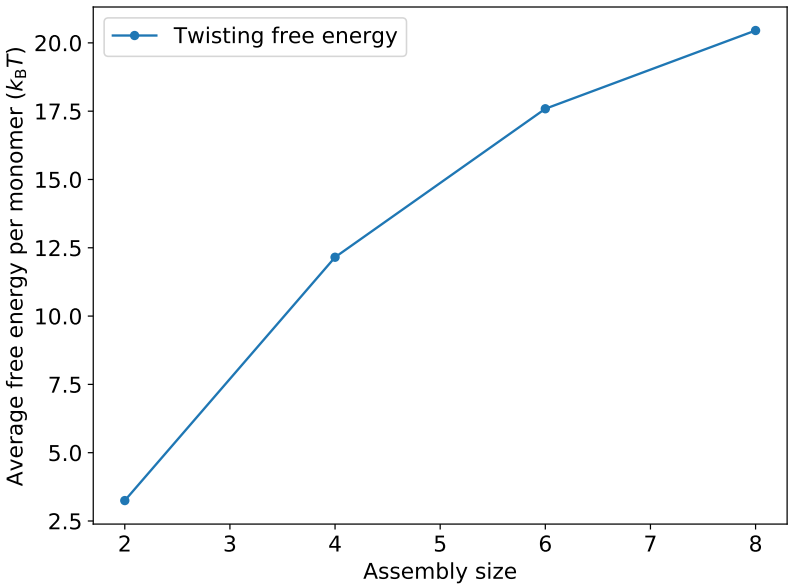


Figure 6.28: Average twisting free energy of assemblies of various sizes. The average stress increases continuously as the assembly size increases, suggesting that the the growth of the assembly could be self-limiting.

relatively stronger signal of self-limitation. A possible reason is that untwisting is the dominant mode of stress accumulation in the range of assembly sizes studied, which could imply that our stress free energy estimation for the twisted monomers may be more accurate compared to that for the curved monomers, where some modes of stress accumulation, such as bond stretching and broken base pairs, were neglected due to methodological difficulties in estimating their respective contributions.

6.6 Methods

6.6.1 Preparing monomer and assembly configurations

The procedures for preparing monomer and assembly configurations are essentially identical for all three origami designs.

For each design, to generate starting configurations for the monomers, the caDNAno file was converted into oxDNA format using the tacoxDNA software package. The resulting monomer structures were then subjected to the standard relaxation procedure described in Chapter 2.1 and further equilibrated for 1.52 µs (10^8 steps).

Assemblies of monomers are constructed using pulling simulations. For an assembly with N monomers, the monomers were placed in a simulation box such that they are in the correct orientation to bind and they are not overlapping with each other. Moving mutual traps were used to slowly bring complementary sticky ends together and induce their hybridisation. These traps were applied between pairs of complementary nucleotides. The trap energy is a function of time and the distance between the nucleotides, denoted as t and r respectively. Mathematically,

$$V(r,t) = \frac{1}{2}k(r - (r_0 + ts))^2$$

where k is the force constant, r_0 is the initial equilibrium distance, and s is the moving speed of the equilibrium distance. k is set to be $57.09\,\mathrm{pN/nm}$; r_0 is set to be the initial distance for each pair of nucleotide, so that the initial trap energy is zero; s is set to be $(r_0 - 1\,\mathrm{nm})/(76\,\mathrm{ns})$, so that after $76\,\mathrm{ns}$ (5×10^6 steps, the duration of a pulling simulation), the trap distance reaches the equilibrium distance of a base pair, $1\,\mathrm{nm}$.

After the pulling simulations, the traps were maintained but with s set to zero so that the traps were stationary. The traps were still needed because hybridisation between the sticky ends was incomplete after the pulling simulations. After hybridisation was complete, the assemblies were further equilibrated with traps for $152\,\mathrm{ns}$ (10^7 steps), then without traps for $1.52\,\mu\mathrm{s}$ (10^8 steps), before a production run without traps for $1.52\,\mu\mathrm{s}$ (10^8 steps).

6.6.2 Bending free energy landscape

We used umbrella sampling [92] to calculate the free energy landscape as a function of the order parameter R_{ee}, the end-to-end distance of the monomer. To avoid effects of helices splaying at their ends, we defined the end-to-end distance as the distance between the centres of mass of two groups of nucleotides. For design 1, the two groups contain the nucleotides at positions 88–108 and 424–444 respectively. For design 2, the two groups contain the nucleotides at positions 49–62 and 259–272 respectively. For the twisted monomer, the two groups contain the nucleotides at

positions 42–63 and 385–406 respectively. The nucleotide positions are defined in the caDNAno designs in Figure A.8, Figure A.9 and Figure A.10.

For each design, we performed a set of simulations where the order parameter was restrained with a harmonic potential in each sampling window. The sampling pathway for bending ranges from $R_{ee} = 60\,\text{nm}$ to $102\,\text{nm}$ for design 1, and from $R_{ee} = 43\,\text{nm}$ to $69\,\text{nm}$ for design 2.

Before the umbrella sampling simulations, we prepared the starting configurations for each window using two pulling simulations. The end-to-end distance of an equilibrated configuration was restrained at its original value (94 nm for design 1 and 57 nm for design 2) with a harmonic bias potential of stiffness $k = 57.09\,\text{pN/nm}$.

In the first simulation, the equilibrium position of the harmonic bias potential was gradually reduced at a constant rate of $0.028\,\text{m/s}$, until it reaches 60 nm for design 1 and 43 nm for design 2. Configurations were outputted every 60.8 ns (4×10^6 steps), i.e. when the traps have moved by 1.7 nm, resulting in a starting configuration for each window from the starting value of R_{ee} to the final value of R_{ee}, where neighbouring windows are separated by 1.7 nm.

In the second simulation, the equilibrium position was gradually increased at the same rate as the first simulation, until it reaches 102 nm for design 1 and 69 nm for design 2. Configurations were outputted at the same frequency, which gives starting configurations for windows defined in a similar way as in the first simulation.

After the starting configurations have been prepared, the configuration in each window was restrained at the value of R_{ee} corresponding to that window with a harmonic bias potential of stiffness $k = 5.709\,\text{pN/nm}$. Each window was equilibrated for 15.2 ns (10^6 steps) before a production run of 152 ns (10^7 steps). R_{ee} was outputted every 15.2 ps (10^3 steps), giving 10^4 data points for each window. Using the biased probability distributions of R_{ee} in each window, we used the Weighted Histogram Analysis Method (WHAM) [93, 136] to calculate the unbiased free energy landscape of the system as a function of R_{ee}.

The production run was then repeated with the last configuration of the previous production run, and WHAM was performed on the new data points to calculate the

free energy landscape. This process was repeated three times, with the resulting landscapes being very similar, thus confirming convergence.

6.6.3 Twisting free energy landscape

Since the order parameter for the twisting free energy landscape is an angle rather than a distance, if we were to use umbrella sampling to calculate the free energy landscape, the bias would have to be a function of an angle, meaning that the force associated with the bias would have to be a function of the derivative of an angle. Since this has not been implemented (and would be challenging to implement) in oxDNA, we are unable to use umbrella sampling for this free energy landscape. As a workaround, we used a slow twisting simulation to estimate the relationship between the torque and the twist angle, from which we can integrate to obtain the work done as a function of the twist angle. In the slow and reversible limit, the work done can serve as a proxy for the free energy.

For curved monomer design 1, we defined the twisting angle to be the angle between the vector pointing from the group of nucleotides at positions 98–108 to those at positions 153–163, and the vector pointing from the group of nucleotides at positions 369–379 to those at positions 424–434. For the twisted monomer, the four groups consist of nucleotides on helices 2–3 and positions 42–63, helices 14–15 and positions 42–63, helices 2–3 and positions 385–406, and helices 14–15 and positions 385–406. The nucleotide positions are defined in the caDNAno designs in Figure A.8 and Figure A.10. The sign of the twist angle is defined using the right-hand rule.

From the equilibrated configurations of the monomer, a configuration with the angle closest to zero, for curved monomer design 1, and $2\pi/3$, for the twisted monomer, was selected and used as the starting configuration for the twisting simulation. The configuration is rotated such that the line joining the midpoint of the two vectors is aligned with the z-axis.

In the twisting simulation, each of the four groups of nucleotides described in the previous paragraph were restrained at their respective centres of mass using harmonic traps with a stiffness of $k = 57.09\,\mathrm{pN/nm}$. The z-component of the trap

force was removed to avoid any bending or stretching forces on the structure. For curved monomer design 1, the equilibrium positions of the traps were then rotated about the z-axis at a rate of $2\pi \times 10^{-9}$ rad/step, counterclockwise and clockwise for the first and second group of nucleotides respectively, which gives an effective twisting rate of $-4\pi \times 10^{-9}$ rad/step. The sense of rotation for the twisted monomer is the opposite of curved monomer design 1, but the rate remains the same. The simulation was run for 1.9 µs (1.25×10^8 steps), so that the twist angle reached $-\pi/2$ for curved monomer design 1, and $\pi/6$ for the twisted monomer, at the end of the simulation. The torque exerted by the traps was outputted every 15.2 ps (10^3 steps).

The torque-angle curve thus obtained was integrated to give the work done as a function of the twist angle, which was taken as the twisting free energy landscape.

6.6.4 Twist-bend coupling

For design 1, the starting configurations for the three simulations with R_{ee} restrained at 60, 77 and 94 nm respectively, were taken from the corresponding umbrella sampling window. A configuration was chosen from each window where the twist angle is close to zero. The procedure for obtaining the free energy landscape for studying twist-bend coupling effects is very similar to that for the twisting free energy landscape, with the only difference being that the z-component of the trap force was not removed in order to restrain R_{ee} as well.

6.6.5 Bond distance

The bond distance for each pair of sticky ends is defined as the distance between the centres of mass of two groups of nucleotides that are within 5 base pairs from each sticky end. For design 1, the two groups consist of the nucleotides on helices 2, 4 and 6, and 9, 11 and 13, respectively. For design 2, the two groups consist of the nucleotides on helices 2, 8 and 14, and 17, 23 and 29, respectively.

6.7 Conclusions

We have designed curved and twisted DNA origamis which could potentially undergo self-limiting assembly. For each design, we characterised the structural properties and free energy landscapes of the origami monomers, and estimated how the stress free energy accumulates in the assemblies.

We first investigated a curved 14-helix bundle. For assemblies with a small size, curvature and stress accumulates as designed, but as the assembly grows further, stress also accumulates through twisting of the monomers. The twist allows a partial relaxation in the bending stress of the monomers, although twisting also results in stress accumulation. This can be viewed as a shift in stress accumulation from the bending mode to the twisting mode. By calculating the bending and twisting free energy landscapes of the monomer, we estimated that the combined stress free energy is increasing, albeit slowly, as a function of the assembly size, which suggests that the assembly exhibits self-limiting behaviour.

In order to improve the design such that bending remains the dominant mode of stress accumulation, we investigated a curved 30-helix bundle, where the wider cross-section should suppress twisting more strongly than bending. Indeed, twisting is no longer observed, but bond stretching becomes another mode of stress accumulation in larger assemblies. As bond stretching may involve fraying of duplexes near the bonds, it is challenging to estimate the corresponding component of stress free energy, although the bond distance analysis shows that the degree of stretching increases as the assembly grows. The bending free energy landscape remains tractable, but the corresponding stress free energy accumulation does not increase continuously as the assembly size. Given that bond stretching stress appears to be increasing as the assembly grows, it might also be possible for the overall stress free energy to be increasing, similar to the case for the 14-helix bundle when bending and twisting stress are both considered.

We also explored whether twisting is a viable mechanism for stress accumulation. We investigated a twisted 18-helix bundle, which assembles with the monomers being untwisted. For this system, twisting appears to be the dominant mode of

stress accumulation. As the assembly grows, the stress free energy grows more quickly than that for the curved origami assemblies. A possible reason is that we have better accounted for the sources of stress in the twisted monomers than in the curved monomers.

We have demonstrated the use of oxDNA simulations to characterise the structural properties and free energy landscapes of DNA origami designs, as well as the thermodynamic properties of their assemblies. oxDNA provides a valuable and efficient tool for testing and optimising origami designs before they are synthesised and characterised experimentally. Qualitatively, it allows us to identify the stress accumulation mechanisms, and quantitatively, it allows us to estimate the stress free energy of assemblies from the free energy landscapes of individual origami monomers. Importantly, oxDNA can capture fine structural details such as base pair formation and breaking, which could play a role in stress accumulation. Combined with the efficiency of its coarse-grained nature, we anticipate that oxDNA will become a more popular tool in designing DNA origamis for self-limiting assembly.

7
Conclusions

In this thesis, we have used the oxDNA coarse-grained model together with umbrella sampling to calculate the free energy landscapes of a number of DNA origamis. Using the free energy landscapes, we have characterised the stress accumulation in the self-limiting assembly of DNA origamis into higher-order structures.

In Chapter 3, we showed how coarse-grained modelling combined with umbrella sampling using distance-based order parameters can be used to compute the free energy landscapes associated with mechanical deformations of large DNA nanostructures. We illustrated this approach for the extreme bending of DNA nanotubes. We found that the homogeneous bending of the DNA nanotubes is well described by the worm-like chain model, while for more extreme bending, the bending stress localises at one or two "kinks", but the process is reversible and the nanotubes can recover from this kinking if the bending stress is removed. In addition to bending, we also characterised the free energy landscapes associated with stretching the nanotubes, from which we extracted their stretch moduli. We found a smaller than expected value and attributed this observation to a reduction in nanotube radius upon tension.

Having established and tested the approach of umbrella sampling in oxDNA, in Chapter 4 we applied the approach to study a mechanically bistable DNA origami, the free energy landscape of which has been characterised experimentally. Although

the experimental landscape shows two minima, the landscape computed from the simulations only shows one minimum. We decomposed the landscape into four components, and found that the designed bistable landscape is present in one component but was masked by the other components, resulting in only one free energy minimum. We noted a possible source of difference as the heterogeneity in the ensemble of DNA origamis that could potentially be partly defective. These defects might account for the presence of the second free energy minimum observed experimentally, but due to the absence of defects in simulations, the second minimum cannot be observed.

In the second half of the thesis, we described the application of free energy landscapes of DNA origamis, with a particular focus on self-limiting assembly with stress accumulation. In Chapter 5, we studied a DNA origami that undergoes self-limiting assembly with stress propagation through accumulated displacement. We characterised the structural properties of the origami monomer and its polymers. While short polymers generally accumulate stress as designed, stress release mechanisms emerge for longer polymers. For example, polymers twist along the polymerisation axis and flanking blocks pile up on each other. After structural characterisation, we computed the free energy landscapes of the monomer associated with the displacement of the flanking block, and the landscapes for the repulsion between the flanking blocks of two neighbouring origamis. Using the two landscapes, we parametrised a simple model to describe stress accumulation in the polymers. We then used Monte Carlo simulations with the model to predict the structural properties, which gave a good agreement with oxDNA simulations for short polymers and a less satisfactory agreement for long polymers due to stress release mechanisms that cannot be captured by the simple model.

Finally, in Chapter 6, we designed DNA origamis that self-assemble with stress propagated through the accumulation of curvature or twist. We characterised the structural properties of each origami monomer as well as their assembly, and we computed the free energy landscapes of each origami monomer associated with the major modes of stress accumulation in their assembly. We found that although

stress could accumulate in a single mode at a low stress level, it is more commonly divided across multiple modes as the stress level builds up in larger assemblies. For example, although the objective of the first design was to accumulate stress through curvature, twist also became a significant mode of stress accumulation for larger assemblies. While the aim of the second design was to eliminate twisting, bond stretching emerged as yet another possible mode. These findings echo with those in Chapter 5 where stress-release mechanisms were identified in larger assemblies.

Beyond the chapters in this thesis, I have also studied the assembly of icosahedral quasicrystals using patchy particles and their potential realisation using DNA origamis. The study is not included here as it is outside the main theme of the thesis, but the interested reader can refer to Noya *et al.* [196].

This brings an end to the thesis, but as we conclude, it is worth mentioning that like any other models, oxDNA also has its own shortcomings. For example, the stretch modulus of duplex DNA is known to be overestimated in oxDNA compared to experiments [47]. This overestimation might slightly affect the accuracy of mechanical properties of DNA structures represented in the model. Another known shortcoming is the underestimation of coaxial stacking energy in oxDNA compared to experiments [197]. The effect of this underestimation is mostly pertinent to the self-assembly of DNA origamis into higher order structures through blunt-ended stacking [169, 171, 173]. The challenge with improving on the latter shortcoming, however, is the limited amount of relevant experimental data against which oxDNA can be parametrised. New experimental data for DNA blunt-ended stacking would be helpful for improving the ability of the oxDNA model to describe higher-order self-assembly of DNA origamis mediated by such interactions.

Notwithstanding the above shortcomings, a diverse range of DNA nanostructures has been successfully modelled using oxDNA, and we can reasonably expect its success and popularity to continue. Additionally, with the maturation of relevant supporting software [72, 73, 198], we expect oxDNA to become a more accessible tool for experimentalists to utilise in their research on DNA nanotechnology, whether

it is for interpreting experimental results, or for predicting the properties of novel nanostructures.

Appendix

A

Full caDNAno Designs of DNA Origamis

Contents

A.1 Chapter 3 . . . . . **155**
A.2 Chapter 4 . . . . . **155**
A.3 Chapter 5 . . . . . **157**
A.4 Chapter 6 . . . . . **158**

This appendix contains the full designs of the DNA origamis studied in each chapter. The designs are directly exported from the caDNAno software in SVG format.

A.1 Chapter 3

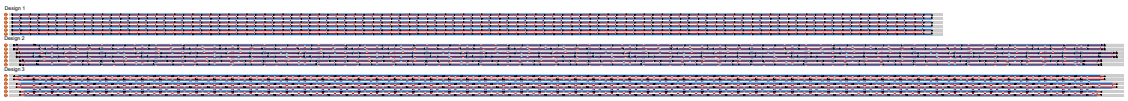


Figure A.1: Full caDNAno designs for the three nanotube designs

A.2 Chapter 4

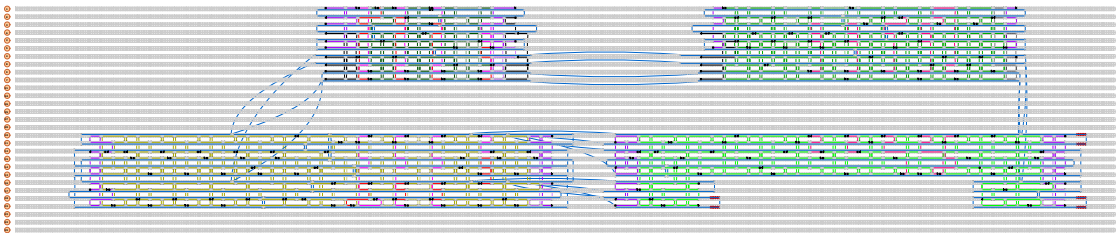


Figure A.2: Full caDNAno design for the nanomechanism

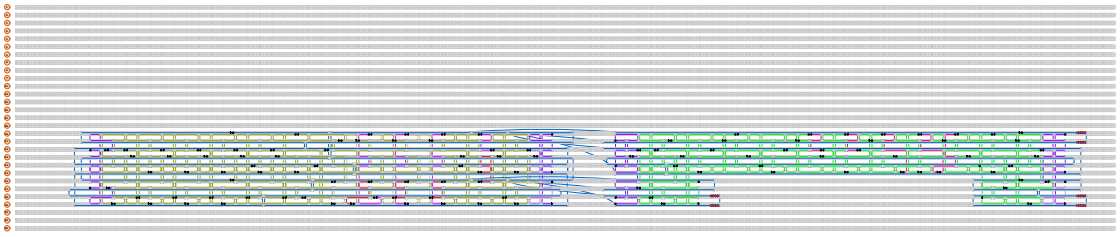


Figure A.3: Full caDNAno design for the partial nanomechanism used for calculating the free energy landscape of joint 1

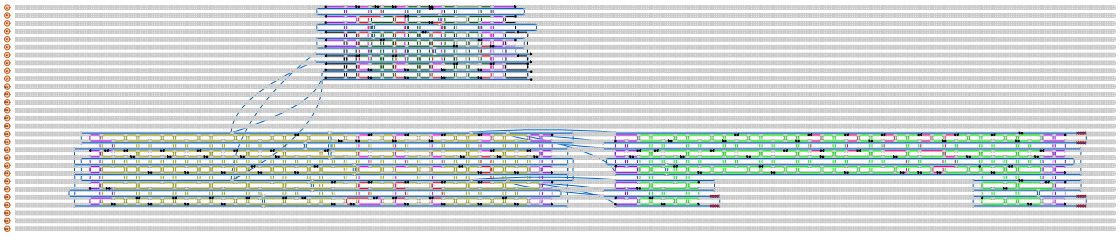


Figure A.4: Full caDNAno design for the partial nanomechanism used for calculating the free energy landscape of joint 2

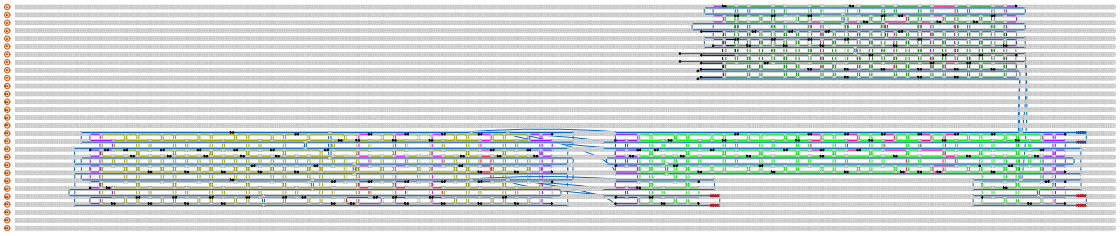


Figure A.5: Full caDNAno design for the partial nanomechanism used for calculating the free energy landscape of joint 3

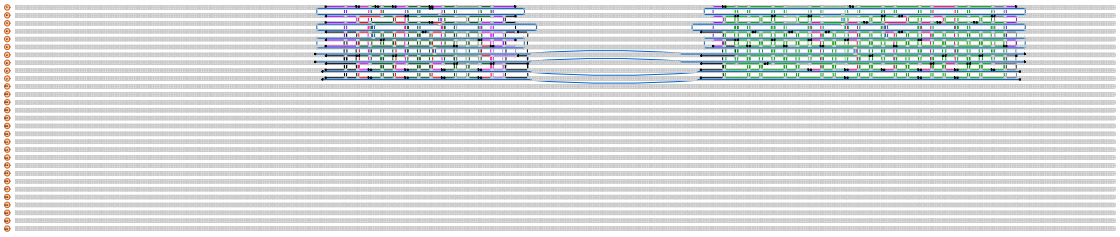


Figure A.6: Full caDNAno design for the partial nanomechanism used for calculating the free energy landscape of joint 4

A.3 Chapter 5

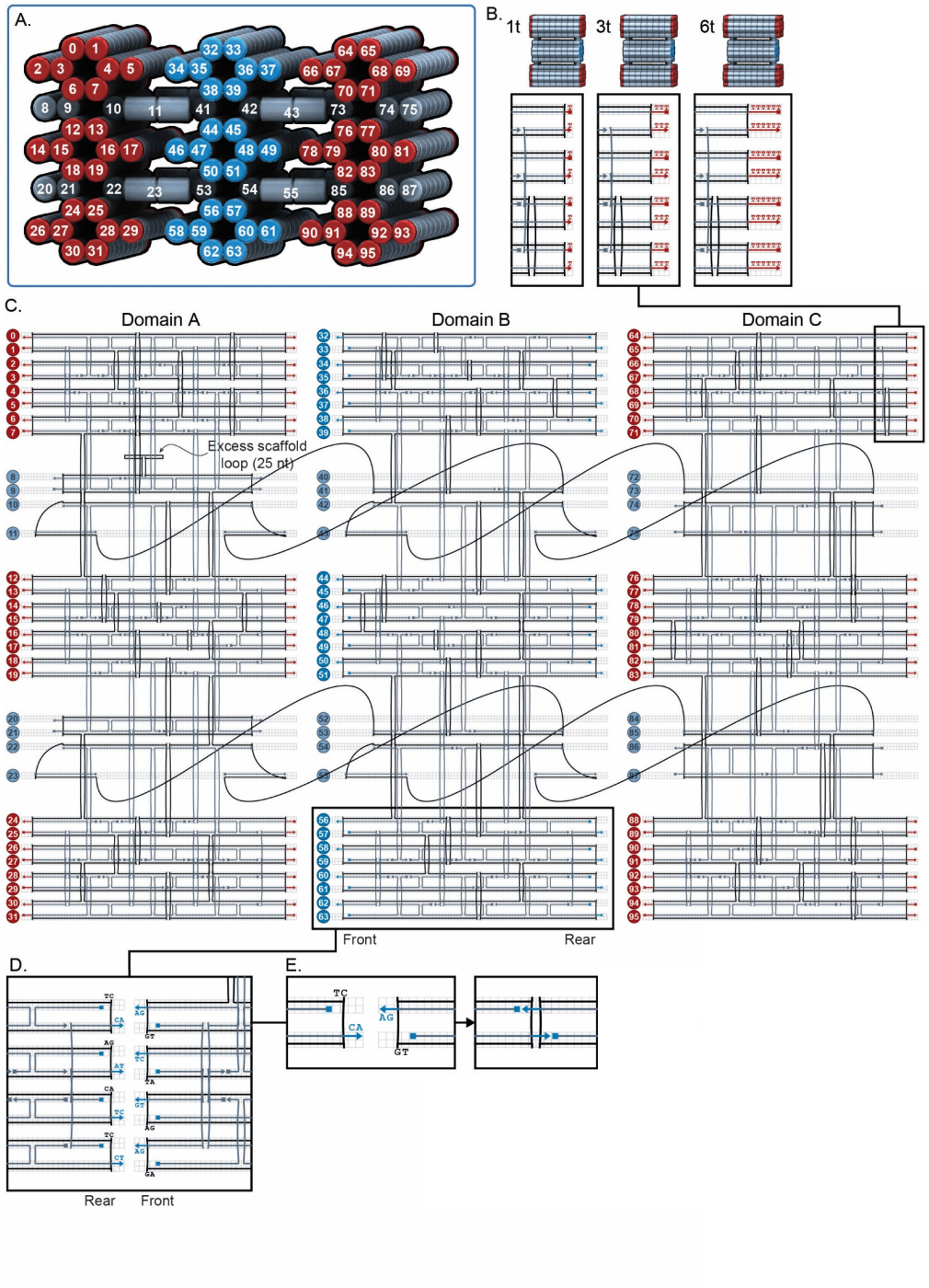


Figure A.7: (courtesy of Jonathan F. Berengut) Full caDNAno design for the DNA origami monomer

A.4 Chapter 6

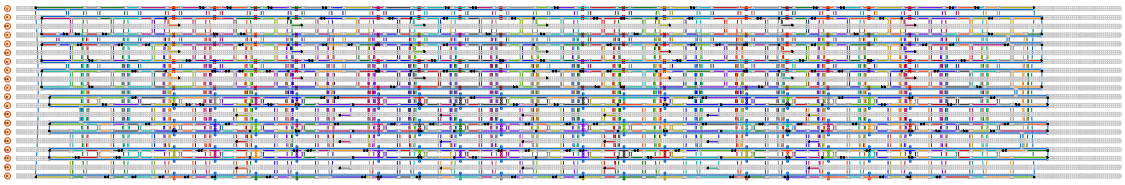


Figure A.8: Full caDNAno design for curved monomer design 1

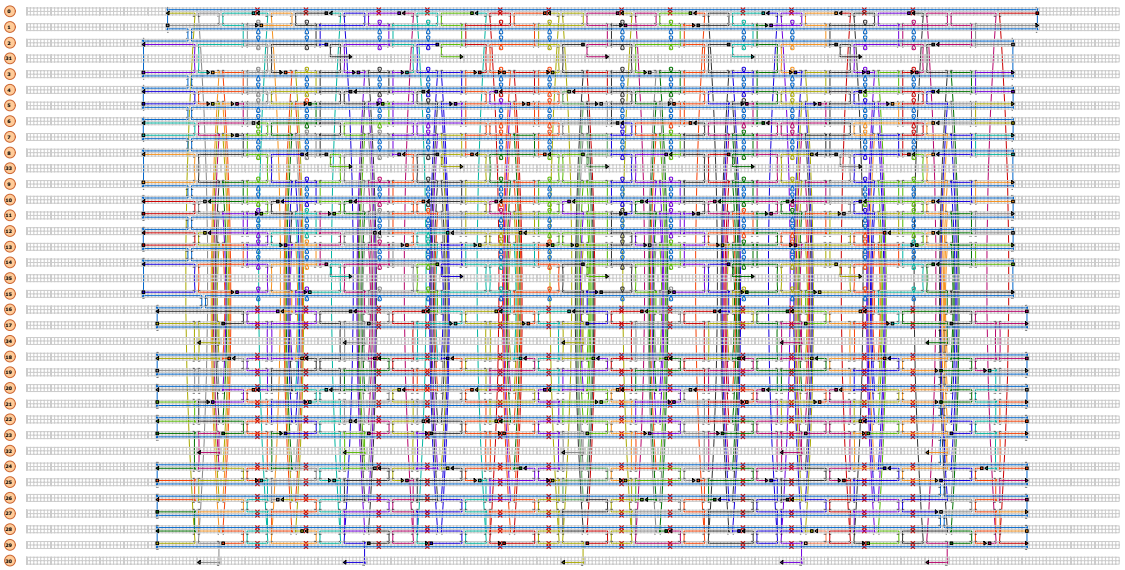


Figure A.9: Full caDNAno design for curved monomer design 2

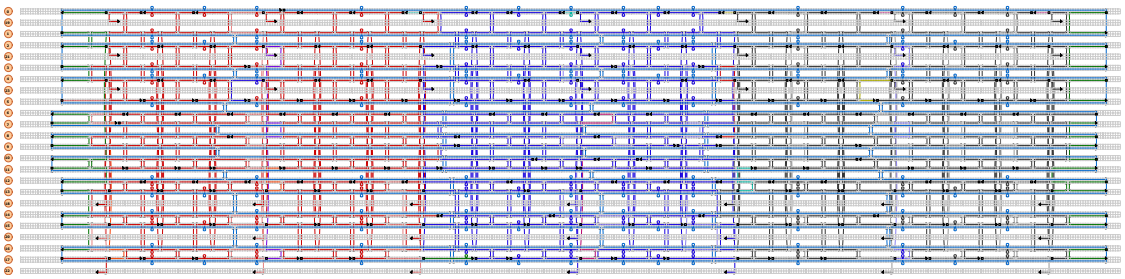


Figure A.10: Full caDNAno design for twisted monomer

References

[1]J. D. Watson and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid", Nature **171**, 737 (1953).

[2]N. C. Seeman, "Nucleic acid junctions and lattices", J. Theor. Biol. **99**, 237 (1982).

[3]E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman, "Design and self-assembly of two-dimensional DNA crystals", Nature **394**, 539 (1998).

[4]P. W. K. Rothemund, "Folding DNA to create nanoscale shapes and patterns", Nature **440**, 297 (2006).

[5]S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, "Self-assembly of DNA into nanoscale three-dimensional shapes", Nature **459**, 414 (2009).

[6]Y. Ke, S. M. Douglas, M. Liu, J. Sharma, A. Cheng, A. Leung, Y. Liu, W. M. Shih, and H. Yan, "Multilayer DNA Origami Packed on a Square Lattice", J. Am. Chem. Soc. **131**, 15903 (2009).

[7]H. Dietz, S. M. Douglas, and W. M. Shih, "Folding DNA into Twisted and Curved Nanoscale Shapes", Science **325**, 725 (2009).

[8]D. Han, S. Pal, J. Nangreave, Z. Deng, Y. Liu, and H. Yan, "DNA Origami with Complex Curvatures in Three-Dimensional Space", Science **332**, 342 (2011).

[9]E. Benson, A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen, and B. Högberg, "DNA rendering of polyhedral meshes at the nanoscale", Nature **523**, 441 (2015).

[10]R. Veneziano, S. Ratanalert, K. Zhang, F. Zhang, H. Yan, W. Chiu, and M. Bathe, "Designer nanoscale DNA assemblies programmed from the top down", Science **352**, 1534 (2016).

[11]H. Jun, X. Wang, W. P. Bricker, and M. Bathe, "Automated sequence design of 2D wireframe DNA origami with honeycomb edges", Nat. Commun. **10**, 5419 (2019).

[12]H. Jun, F. Zhang, T. Shepherd, S. Ratanalert, X. Qi, H. Yan, and M. Bathe, "Autonomously designed free-form 2D DNA origami", Sci. Adv. **5**, eaav0655 (2019).

[13]H. Jun, T. R. Shepherd, K. Zhang, W. P. Bricker, S. Li, W. Chiu, and M. Bathe, "Automated Sequence Design of 3D Polyhedral Wireframe DNA Origami with Honeycomb Edges", ACS Nano **13**, 2083 (2019).

[14]Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, "Three-Dimensional Structures Self-Assembled from DNA Bricks", Science **338**, 1177 (2012).

[15]B. Wei, M. Dai, and P. Yin, "Complex shapes self-assembled from single-stranded DNA tiles", Nature **485**, 623 (2012).

[16]L. L. Ong, N. Hanikel, O. K. Yaghi, C. Grun, M. T. Strauss, P. Bron, J. Lai-Kee-Him, F. Schueder, B. Wang, P. Wang, J. Y. Kishi, C. Myhrvold, A. Zhu, R. Jungmann, G. Bellot, Y. Ke, and P. Yin, "Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components", Nature **552**, 72 (2017).

[17]N. C. Seeman and H. F. Sleiman, "DNA Nanotechnology", Nat. Rev. Mater. **3**, 17068 (2017).

[18]S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. D. Nijenhuis, B. Sacca, F. C. Simmel, H. Yan, and P. Zhan, "DNA Origami", Nat. Rev. Methods Primers **1**, 13 (2021).

[19]M. Glaser, S. Deb, F. Seier, A. Agrawal, T. Liedl, S. Douglas, M. K. Gupta, and D. M. Smith, "The Art of Designing DNA Nanostructures with CAD Software", Molecules **26**, 2287 (2021).

[20]S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih, "Rapid Prototyping of 3D DNA-origami Shapes with caDNAno", Nucleic Acids Res. **37**, 5001 (2009).

[21]H. Jun, X. Wang, M. F. Parsons, W. P. Bricker, T. John, S. Li, S. Jackson, W. Chiu, and M. Bathe, "Rapid prototyping of arbitrary 2D and 3D wireframe DNA origami", Nucleic Acids Res. **49**, 10265 (2021).

[22]C.-M. Huang, A. Kucinic, J. A. Johnson, H.-J. Su, and C. E. Castro, "Integrated computer-aided engineering and design for DNA assemblies", Nat. Mater. **20**, 1264 (2021).

[23]K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and A. D. MacKerell, "Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium", J. Chem. Theory Comput. **8**, 348 (2012).

[24]A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, "Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers", Biophys. J. **92**, 3817 (2007).

[25]J. Yoo and A. Aksimentiev, "In situ Structure and Dynamics of DNA Origami Determined through Molecular Dynamics Simulations", Proc. Natl. Acad. Sci. U. S. A. **110**, 20099 (2013).

[26]C. Maffeo, J. Yoo, and A. Aksimentiev, "De Novo Reconstruction of DNA Origami Structures through Atomistic Molecular Dynamics Simulation", Nucleic Acids Res. **44**, 3013 (2016).

[27]S. M. Slone, C.-Y. Li, J. Yoo, and A. Aksimentiev, "Molecular Mechanics of DNA Bricks: in situ Structure, Mechanical Properties and Ionic Conductivity", New J. Phys. **18**, 055012 (2016).

[28]C.-Y. Li, E. A. Hemmig, J. Kong, J. Yoo, S. Hernández-Ainsa, U. F. Keyser, and A. Aksimentiev, "Ionic Conductivity, Structural Deformation, and Programmable Anisotropy of DNA Origami in Electric Field", ACS Nano **9**, 1420 (2015).

[29]K. Göpfrich, C.-Y. Li, M. Ricci, S. P. Bhamidimarri, J. Yoo, B. Gyenes, A. Ohmann, M. Winterhalter, A. Aksimentiev, and U. F. Keyser, "Large-Conductance Transmembrane Porin Made from DNA Origami", ACS Nano **10**, 8207 (2016).

[30]C. E. Castro, F. Kilchherr, D.-N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe, and H. Dietz, "A primer to scaffolded DNA origami", Nat. Methods **8**, 221 (2011).

[31]D.-N. Kim, F. Kilchherr, H. Dietz, and M. Bathe, "Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures", Nucleic Acids Res. **40**, 2862 (2011).

[32]K. Pan, D.-N. Kim, F. Zhang, M. R. Adendorff, H. Yan, and M. Bathe, "Lattice-free prediction of three-dimensional structure of programmed DNA assemblies", Nat. Commun. **5**, 5578 (2014).

[33]J. J. de Pablo, "Coarse-Grained Simulations of Macromolecules: From DNA to Nanocomposites", Annu. Rev. Phys. Chem. **62**, 555 (2011).

[34]D. A. Potoyan, A. Savelyev, and G. A. Papoian, "Recent successes in coarse-grained modeling of DNA", Wiley Interdiscip. Rev.: Comput. Mol. Sci. **3**, 69 (2013).

[35]P. D. Dans, J. Walther, H. Gómez, and M. Orozco, "Multiscale simulation of DNA", Curr. Opin. Struct. Biol. **37**, 29 (2016).

[36]T. Sun, V. Minhas, N. Korolev, A. Mirzoev, A. P. Lyubartsev, and L. Nordenskiöld, "Bottom-Up Coarse-Grained Modeling of DNA", Front. Mol. Biosci. **8**, 159 (2021).

[37]W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems", J. Chem. Phys. **139**, 090901 (2013).

[38]G. A. Papoian, *Coarse-grained modeling of biomolecules* (CRC Press, 2017).

[39]A. Savelyev and G. A. Papoian, "Chemically accurate coarse graining of double-stranded DNA", Proc. Natl. Acad. Sci. U. S. A. **107**, 20340 (2010).

[40]A. Naômé, A. Laaksonen, and D. P. Vercauteren, "A Solvent-Mediated Coarse-Grained Model of DNA Derived with the Systematic Newton Inversion Method", J. Chem. Theory Comput. **10**, 3541 (2014).

[41]H. Yagyu, J.-Y. Lee, D.-N. Kim, and O. Tabata, "Coarse-Grained Molecular Dynamics Model of Double-Stranded DNA for DNA Nanostructure Design", J. Phys. Chem. B **121**, 5033 (2017).

[42]T. A. Knotts, N. Rathore, D. C. Schwartz, and J. J. de Pablo, "A coarse grain model for DNA", J. Chem. Phys. **126**, 084901 (2007).

[43]D. M. Hinckley, G. S. Freeman, J. K. Whitmer, and J. J. de Pablo, "An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization", J. Chem. Phys. **139**, 144903 (2013).

[44]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations", J. Phys. Chem. B **111**, 7812 (2007).

[45]J. J. Uusitalo, H. I. Ingólfsson, P. Akhshi, D. P. Tieleman, and S. J. Marrink, "Martini Coarse-Grained Force Field: Extension to DNA", J. Chem. Theory Comput. **11**, 3932 (2015).

[46]P. D. Dans, A. Zeida, M. R. Machado, and S. Pantano, "A Coarse Grained Model for Atomic-Detailed DNA Simulations with Explicit Electrostatics", J. Chem. Theory Comput. **6**, 1711 (2010).

[47]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, "Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model", J. Chem. Phys. **134**, 085101 (2011).

[48]P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, and A. A. Louis, "Sequence-dependent thermodynamics of a coarse-grained DNA model", J. Chem. Phys. **137**, 135101 (2012).

[49]B. E. K. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir, A. A. Louis, and J. P. K. Doye, "Introducing Improved Structural Properties and Salt Dependence into a Coarse-Grained Model of DNA", J. Chem. Phys. **142**, 234901 (2015).

[50]G. S. Freeman, D. M. Hinckley, J. P. Lequieu, J. K. Whitmer, and J. J. de Pablo, "Coarse-grained modeling of DNA curvature", J. Chem. Phys. **141**, 165103 (2014).

[51]D. M. Hinckley, J. P. Lequieu, and J. J. de Pablo, "Coarse-grained modeling of DNA oligomer hybridization: Length, sequence, and salt effects", J. Chem. Phys. **141**, 035102 (2014).

[52]V. Ortiz and J. J. de Pablo, "Molecular Origins of DNA Flexibility: Sequence Effects on Conformational and Mechanical Properties", Phys. Rev. Lett. **106**, 238107 (2011).

[53]G. S. Freeman, J. P. Lequieu, D. M. Hinckley, J. K. Whitmer, and J. J. de Pablo, "DNA Shape Dominates Sequence Affinity in Nucleosome Formation", Phys. Rev. Lett. **113**, 168101 (2014).

[54]B. Zhang, W. Zheng, G. A. Papoian, and P. G. Wolynes, "Exploring the Free Energy Landscape of Nucleosomes", J. Am. Chem. Soc. **138**, 8126 (2016).

[55]D. A. Potoyan, W. Zheng, E. A. Komives, and P. G. Wolynes, "Molecular stripping in the NF-κB/IκB/DNA genetic regulatory network", Proc. Natl. Acad. Sci. U.S.A. **113**, 110 (2016).

[56]W. Lu, C. Bueno, N. P. Schafer, J. Moller, S. Jin, X. Chen, M. Chen, X. Gu, A. Davtyan, J. J. de Pablo, and P. G. Wolynes, "OpenAWSEM with Open3SPN2: A fast, flexible, and accessible framework for large-scale coarse-grained biomolecular simulations", PLoS Comput. Biol. **17**, 1 (2021).

[57]S. J. Marrink and D. P. Tieleman, "Perspective on the Martini model", Chem. Soc. Rev. **42**, 6801 (2013).

[58]V. Maingi, J. R. Burns, J. J. Uusitalo, S. Howorka, S. J. Marrink, and M. S. P. Sansom, "Stability and dynamics of membrane-spanning DNA nanopores", Nat. Commun. **8**, 14784 (2017).

[59]M. R. Machado, P. D. Dans, and S. Pantano, "A hybrid all-atom/coarse grain model for multiscale simulations of DNA", Phys. Chem. Chem. Phys. **13**, 18134 (2011).

[60]M. R. Machado and S. Pantano, "Exploring LacI–DNA Dynamics by Multiscale Simulations Using the SIRAH Force Field", J. Chem. Theory Comput. **11**, 5012 (2015).

[61]F. Dannenberg, K. E. Dunn, J. Bath, M. Kwiatkowska, A. J. Turberfield, and T. E. Ouldridge, "Modelling DNA origami self-assembly at the domain level", J. Chem. Phys. **143**, 165102 (2015).

[62]K. E. Dunn, F. Dannenberg, T. E. Ouldridge, M. Kwiatkowska, A. J. Turberfield, and J. Bath, "Guiding the folding pathway of DNA origami", Nature **525**, 82 (2015).

[63]A. Cumberworth, A. Reinhardt, and D. Frenkel, "Lattice models and Monte Carlo methods for simulating DNA origami self-assembly", J. Chem. Phys. **149**, 234905 (2018).

[64]A. Cumberworth, D. Frenkel, and A. Reinhardt, "Simulations of DNA-origami self-assembly reveal design-dependent nucleation barrier", arXiv:2204.02499 (2022), `https://arxiv.org/abs/2204.02499`.

[65]C. Maffeo and A. Aksimentiev, "MrDNA: a Multi-Resolution Model for Predicting the Structure and Dynamics of DNA Systems", Nucleic Acids Res. **48**, 5135 (2020).

[66]J. Y. Lee, J. G. Lee, G. Yun, C. Lee, Y.-J. Kim, K. S. Kim, T. H. Kim, and D.-N. Kim, "Rapid Computational Analysis of DNA Origami Assemblies at Near-Atomic Resolution", ACS Nano **15**, 1002 (2021).

[67]J. SantaLucia and D. Hicks, "The Thermodynamics of DNA Structural Motifs", Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).

[68]J. A. Holbrook, M. W. Capp, R. M. Saecker, and M. T. Record, "Enthalpy and Heat Capacity Changes for Formation of an Oligomeric DNA Duplex: Interpretation in Terms of Coupled Processes of Formation and Association of Single-Stranded Helices", Biochemistry **38**, 8409 (1999).

[69]P. J. Hagerman, "Flexibility of DNA", Annu. Rev. Biophys. Biophys. Chem. **17**, 265 (1988).

[70]L. Rovigatti, P. Šulc, I. Z. Reguly, and F. Romano, "A Comparison between Parallelization Approaches in Molecular Dynamics Simulations on GPUs", J. Comput. Chem. **36**, 1 (2015).

[71]O. Henrich, Y. A. Gutiérrez Fosado, T. Curk, and T. E. Ouldridge, "Coarse-grained simulation of DNA using LAMMPS", Eur. Phys. J. E **41**, 57 (2018).

[72]A. Suma, E. Poppleton, M. Matthies, P. Šulc, F. Romano, A. A. Louis, J. P. K. Doye, C. Micheletti, and L. Rovigatti, "TacoxDNA: a User-Friendly Web Server for Simulations of Complex DNA Structures, from Single Strands to Origami", J. Comput. Chem. **40**, 2586 (2019).

[73]E. Poppleton, J. Bohlin, M. Matthies, S. Sharma, F. Zhang, and P. Šulc, "Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation", Nucleic Acids Res. **48**, e72 (2020).

[74]M. C. Engel, "Exploration of DNA systems under internal and external forcing using coarse-grained modelling", PhD thesis (University of Oxford, 2018).

[75]B. E. K. Snodin, J. S. Schreck, F. Romano, A. A. Louis, and J. P. K. Doye, "Coarse-Grained Modelling of the Structural Properties of DNA Origami", Nucleic Acids Res. **47**, 1585 (2019).

[76]M. A. B. Baker, A. J. Tuckwell, J. F. Berengut, J. Bath, F. Benn, A. P. Duff, A. E. Whitten, K. E. Dunn, R. M. Hynson, A. J. Turberfield, and L. K. Lee, "Dimensions and Global Twist of Single-Layer DNA Origami Measured by Small-Angle X-ray Scattering", ACS Nano **12**, 5791 (2018).

[77]X.-c. Bai, T. G. Martin, S. H. W. Scheres, and H. Dietz, "Cryo-EM Structure of a 3D DNA-Origami Object", Proc. Natl. Acad. Sci. U. S. A. **109**, 20012 (2012).

[78]E. Benson, A. Mohammed, D. Rayneau-Kirkhope, A. Gådin, P. Orponen, and B. Högberg, "Effects of Design Choices on the Stiffness of Wireframe DNA Origami Structures", ACS Nano **12**, 9291 (2018).

[79]E. Benson, M. Lolaico, Y. Tarasov, A. Gådin, and B. Högberg, "Evolutionary Refinement of DNA Nanostructures Using Coarse-Grained Molecular Dynamics Simulations", ACS Nano **13**, 12591 (2019).

[80]R. Sharma, J. S. Schreck, F. Romano, A. A. Louis, and J. P. K. Doye, "Characterizing the Motion of Jointed DNA Nanostructures Using a Coarse-Grained Model", ACS Nano **11**, 12426 (2017).

[81]A. E. Marras, L. Zhou, H.-J. Su, and C. E. Castro, "Programmable motion of DNA origami mechanisms", Proc. Natl. Acad. Sci. U. S. A. **112**, 713 (2015).

[82]Z. Shi, C. E. Castro, and G. Arya, "Conformational Dynamics of Mechanically Compliant DNA Nanostructures from Coarse-Grained Molecular Dynamics Simulations", ACS Nano **11**, 4617 (2017).

[83]L. Zhou, A. E. Marras, H.-J. Su, and C. E. Castro, "DNA Origami Compliant Nanostructures with Tunable Mechanical Properties", ACS Nano **8**, 27 (2014).

[84]M. C. Engel, D. M. Smith, M. A. Jobst, M. Sajfutdinow, T. Liedl, F. Romano, L. Rovigatti, A. A. Louis, and J. P. K. Doye, "Force-Induced Unravelling of DNA Origami", ACS Nano **12**, 6734 (2018).

[85]F. Romano, D. Chakraborty, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis, "Coarse-Grained Simulations of DNA Overstretching", J. Chem. Phys. **138**, 085101 (2013).

[86]M. Mosayebi, A. A. Louis, J. P. K. Doye, and T. E. Ouldridge, "Force-Induced Rupture of a DNA Duplex: From Fundamentals to Force Sensors", ACS Nano **9**, 11993 (2015).

[87]W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters", J. Chem. Phys. **76**, 637 (1982).

[88]D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, 2002).

[89]M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 2017).

[90]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines", J. Chem. Phys. **21**, 1087 (1953).

[91]J. A. Anderson, E. Jankowski, T. L. Grubb, M. Engel, and S. C. Glotzer, "Massively parallel Monte Carlo for many-particle simulations on GPUs", J. Comput. Phys. **254**, 27 (2013).

[92]G. Torrie and J. Valleau, "Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella sampling", J. Comput. Phys. **23**, 187 (1977).

[93]S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method", J. Comput. Chem. **13**, 1011 (1992).

[94]C. K. Wong, C. Tang, J. S. Schreck, and J. P. K. Doye, "Characterizing the free-energy landscapes of DNA origamis", Nanoscale **14**, 2638 (2022).

[95]M. Kube, F. Kohler, E. Feigl, B. Nagel-Yüksel, E. M. Willner, J. J. Funke, T. Gerlin, P. Stömmer, M. N. Honemann, T. G. Martin, S. H. W. Scheres, and H. Dietz, "Revealing the Structures of Megadalton-Scale DNA Complexes with Nucleotide Resolution", Nat. Commun. **11**, 6229 (2020).

[96]C. Lee, J. Y. Lee, and D.-N. Kim, "Polymorphic Design of DNA Origami Structures through Mechanical Control of Modular Components", Nat. Commun. **8**, 2067 (2017).

[97]J. F. Berengut, J. C. Berengut, J. P. K. Doye, D. Prešern, A. Kawamoto, J. Ruan, M. J. Wainwright, and L. K. Lee, "Design and Synthesis of Pleated DNA Origami Nanotubes with Adjustable Diameters", Nucleic Acids Res. **47**, 11963 (2019).

[98]A. T. Blanchard and K. Salaita, "Emerging uses of DNA mechanical devices", Science **365**, 1080 (2019).

[99]C. Gosse, T. R. Strick, and D. Kostrz, "Molecular scaffolds: when DNA becomes the hardware for single-molecule investigations", Curr. Opin. Chem. Biol. **53**, 192 (2019).

[100]T. Liedl, B. Högberg, J. Tytell, D. E. Ingber, and W. M. Shih, "Self-assembly of three-dimensional prestressed tensegrity structures from DNA", Nat. Nanotechnol. **5**, 520 (2010).

[101]D. J. Kauert, T. Kurth, T. Liedl, and R. Seidel, "Direct Mechanical Measurements Reveal the Material Properties of Three-Dimensional DNA Origami", Nano Lett. **11**, 5558 (2011).

[102]E. Pfitzner, C. Wachauf, F. Kilchherr, B. Pelz, W. M. Shih, M. Rief, and H. Dietz, "Rigid DNA Beams for High-Resolution Single-Molecule Mechanics", Angew. Chem. Int. Ed. **52**, 7766 (2013).

[103]C. E. Castro, H.-J. Su, A. E. Marras, L. Zhou, and J. Johnson, "Mechanical design of DNA nanostructures", Nanoscale **7**, 5913 (2015).

[104]C. Lee, K. S. Kim, Y.-J. Kim, J. Y. Lee, and D.-N. Kim, "Tailoring the Mechanical Stiffness of DNA Nanostructures Using Engineered Defects", ACS Nano **13**, 8329 (2019).

[105]D. Schiffels, T. Liedl, and D. K. Fygenson, "Nanoscale Structure and Microscale Stiffness of DNA Nanotubes", ACS Nano **7**, 6700 (2013).

[106]P. W. K. Rothemund, A. Ekani-Nkodo, N. Papadakis, A. Kumar, D. K. Fygenson, and E. Winfree, "Design and Characterization of Programmable DNA Nanotubes", J. Am. Chem. Soc. **126**, 16344 (2004).

[107]P. O'Neill, P. W. K. Rothemund, A. Kumar, and D. K. Fygenson, "Sturdier DNA Nanotubes via Ligation", Nano Lett. **6**, 1379 (2006).

[108]T. Wang, D. Schiffels, S. M. Cuesta, D. K. Fygenson, and N. C. Seeman, "Design and Characterization of 1D Nanotubes and 2D Periodic Arrays Self-Assembled from DNA Multi-Helix Bundles", J. Am. Chem. Soc. **134**, 1606 (2012).

[109]J. F. Marko and S. Cocco, "The micromechanics of DNA", Phys. World **16**, 37 (2003).

[110]H. Chhabra, G. Mishra, Y. Cao, D. Prešern, E. Skoruppa, M. M. C. Tortora, and J. P. K. Doye, "Computing the Elastic Mechanical Properties of Rodlike DNA Nanostructures", J. Chem. Theory Comput. **16**, 7748 (2020).

[111]P. K. Dutta, Y. Zhang, A. T. Blanchard, C. Ge, M. Rushdi, K. Weiss, C. Zhu, Y. Ke, and K. Salaita, "Programmable Multivalent DNA-Origami Tension Probes for Reporting Cellular Traction Forces", Nano Lett. **18**, 4803 (2018).

[112]P. C. Nickels, B. Wünsch, P. Holzmeister, W. Bae, L. M. Kneer, D. Grohmann, P. Tinnefeld, and T. Liedl, "Molecular Force Spectroscopy with a DNA Origami-Based Nanoscopic Force Clamp", Science **354**, 305 (2016).

[113]X. Shi, A.-K. Pumm, J. Isensee, W. Zhao, D. Verschueren, A. Martin-Gonzalez, R. Golestanian, H. Dietz, and C. Dekker, "Sustained unidirectional rotation of a self-organized DNA rotor on a nanopore", arXiv:2206.06613 (2022), `https://arxiv.org/abs/2206.06613`.

[114]J. F. Berengut, C. K. Wong, J. C. Berengut, J. P. K. Doye, T. E. Ouldridge, and L. K. Lee, "Self-Limiting Polymerization of DNA Origami Subunits with Strain Accumulation", ACS Nano **14**, 17428 (2020).

[115]M. F. Hagan and G. M. Grason, "Equilibrium mechanisms of self-limiting assembly", Rev. Mod. Phys. **93**, 025008 (2021).

[116]C. Maffeo, B. Luan, and A. Aksimentiev, "End-to-End Attraction of Duplex DNA", Nucleic Acids Res. **40**, 3812 (2012).

[117]F. Häse and M. Zacharias, "Free Energy Analysis and Mechanism of Base Pair Stacking in Nicked DNA", Nucleic Acids Res. **44**, 7100 (2016).

[118]H. Kang, J. Yoo, B.-K. Sohn, S.-W. Lee, H. S. Lee, W. Ma, J.-M. Kee, A. Aksimentiev, and H. Kim, "Sequence-Dependent DNA Condensation as a Driving Force of DNA Phase Separation", Nucleic Acids Res. **46**, 9401 (2018).

[119]R. Cortini, X. Cheng, and J. C. Smith, "The Tilt-Dependent Potential of Mean Force of a Pair of DNA Oligomers from All-Atom Molecular Dynamics Simulations", J. Phys.: Condens. Matter **29**, 084002 (2017).

[120]K. Liebl and M. Zacharias, "Unwinding Induced Melting of Double-Stranded DNA Studied by Free Energy Simulations", J. Phys. Chem. B **121**, 11019 (2017).

[121]R. M. Harrison, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, "Identifying Physical Causes of Apparent Enhanced Cyclization of Short DNA Molecules with a Coarse-Grained Model", J. Chem. Theory Comput. **15**, 4660 (2019).

[122]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, "DNA Nanotweezers Studied with a Coarse-Grained Model of DNA", Phys. Rev. Lett. **104**, 178101 (2010).

[123]T. E. Ouldridge, R. L. Hoare, A. A. Louis, J. P. K. Doye, J. Bath, and A. J. Turberfield, "Optimizing DNA Nanotechnology through Coarse-Grained Modelling: A Two-Footed DNA Walker", ACS Nano **7**, 2479 (2013).

[124]D. C. Khara, J. S. Schreck, T. E. Tomov, T. E. Ouldridge, J. P. K. Doye, and E. Nir, "DNA Bipedal Motor Walking Dynamics: An Experimental and Theoretical Study of the Dependency on Step Size", Nucleic Acids Res. **46**, 1553 (2018).

[125]W. T. Kaufhold, W. Pfeifer, C. E. Castro, and L. Di Michele, "Probing the Mechanical Properties of DNA Nanostructures with Metadynamics", ACS Nano **16**, 8784 (2022).

[126]W. Pfeifer and B. Saccà, "Synthetic DNA Filaments: from Designs to Applications", Biol. Chem. **399**, 773 (2018).

[127]C. Bustamante, J. F. Marko, E. D. Siggia, and S. B. Smith, "Entropic Elasticity of λ-Phage DNA", Science **265**, 1599 (1994).

[128]R. Vafabakhsh and T. Ha, "Extreme Bendability of DNA less than 100 Base Pairs Long Revealed by Single-Molecule Cyclization", Science **337**, 1097 (2012).

[129]J. S. Mitchell, C. A. Laughton, and S. A. Harris, "Atomistic Simulations Reveal Bubbles, Kinks and Wrinkles in supercoiled DNA", Nucleic Acids Res. **39**, 3928 (2011).

[130]R. M. Harrison, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, "Coarse-Grained Modelling of Strong DNA Bending I: Thermodynamics and Comparison to an Experimental "Molecular Vice"", arXiv:1506.09005 (2015), `https://arxiv.org/abs/1506.09005`.

[131]C. Lee, K. S. Kim, Y.-J. Kim, J. Y. Lee, and D.-N. Kim, "Tailoring the Mechanical Stiffness of DNA Nanostructures Using Engineered Defects", ACS Nano **13**, 8329 (2019).

[132]M. Siavashpouri, C. H. Wachauf, M. J. Zakhary, F. Praetorius, H. Dietz, and Z. Dogic, "Molecular Engineering of Chiral Colloidal Liquid Crystals Using DNA Origami", Nat. Mater. **16**, 849 (2017).

[133]N. B. Becker, A. Rosa, and R. Everaers, "The Radial Distribution Function of Worm-Like Chains", Eur. Phys. J. E **32**, 53 (2010).

[134]T. Odijk, "Stiff Chains and Filaments under Tension", Macromolecules **28**, 7016 (1995).

[135]E. Pilyugina, B. Krajina, A. J. Spakowitz, and J. D. Schieber, "Buckling a Semiflexible Polymer Chain under Compression", Polymer **9**, 99 (2017).

[136]A. Grossfield, *WHAM: the Weighted Histogram Analysis Method*, Version 2.0.9, `http://membrane.urmc.rochester.edu/wordpress/?page_id=126`.

[137]S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "Multidimensional Free-Energy Calculations Using the Weighted Histogram Analysis Method", J. Comput. Chem. **16**, 1339 (1995).

[138]C. K. Wong and J. P. K. Doye, "The Free-Energy Landscape of a Mechanically Bistable DNA Origami", Appl. Sci. **12**, 5875 (2022).

[139]M. DeLuca, Z. Shi, C. E. Castro, and G. Arya, "Dynamic DNA nanotechnology: Toward functional nanoscale devices", Nanoscale Horiz. **5**, 182 (2020).

[140]J. V. Le, Y. Luo, M. A. Darcy, C. R. Lucas, M. F. Goodwin, M. G. Poirier, and C. E. Castro, "Probing Nucleosome Stability with a DNA Origami Nanocaliper", ACS Nano **10**, 7073 (2016).

[141]H. Su, J. M. Brockman, Y. Duan, N. Sen, H. Chhabra, A. Bazrafshan, A. T. Blanchard, T. Meyer, B. Andrews, J. P. K. Doye, Y. Ke, R. B. Dyer, and K. Salaita, "Massively parallelized molecular force manipulation with on demand thermal and optical control", J. Am. Chem. Soc. **43**, 19466 (2021).

[142]L. Zhou, A. E. Marras, C.-M. Huang, C. E. Castro, and H.-J. Su, "Paper Origami-Inspired Design and Actuation of DNA Nanomachines with Complex Motions", Small **14**, 1802580 (2018).

[143]Y. Wang, J. V. Le, K. Crocker, M. A. Darcy, P. D. Halley, D. Zhao, N. Andrioff, C. Croy, M. G. Poirier, R. Bundschuh, and C. Castro, "A nanoscale DNA force spectrometer capable of applying tension and compression on biomolecules", Nucleic Acids Res. **49**, 8987 (2021).

[144]J. J. Funke, P. Ketterer, C. Lieleg, S. Schunter, P. Korber, and H. Dietz, "Uncovering the forces between nucleosomes using DNA origami", Sci. Adv. **2**, e1600974 (2016).

[145]M. W. Hudoba, Y. Luo, A. Zacharias, M. G. Poirier, and C. E. Castro, "Dynamic DNA origami device for measuring compressive depletion forces", ACS Nano **11**, 6566 (2017).

[146]Y. Ke, T. Meyer, W. M. Shih, and G. Bellot, "Regulation at a distance of biomolecular interactions using a DNA nanoactuator", Nat. Commun. **7**, 10935 (2016).

[147]Z. Shi and G. Arya, "Free Energy Landscape of Salt-Actuated Reconfigurable DNA Nanodevices", Nucleic Acids Res. **48**, 548 (2019).

[148]L. Zhou, A. E. Marras, H.-J. Su, and C. E. Castro, "Direct Design of an Energy Landscape with Bistable DNA Origami Mechanisms", Nano Lett. **15**, 1815 (2015).

[149]A. E. Marras, L. Zhou, V. Kolliopoulos, H.-J. Su, and C. E. Castro, "Directing folding pathways for multi-component DNA origami nanostructures with complex topology", New J. Phys **18**, 055005 (2016).

[150]K. F. Wagenbauer, C. H. Wachauf, and H. Dietz, "Quantifying quality in DNA self-assembly", Nat. Commun. **5**, 3691 (2014).

[151]C. Myhrvold, M. Baym, N. Hanikel, L. L. Ong, J. S. Gootenberg, and P. Yin, "Barcode extension for analysis and reconstruction of structures", Nat. Commun. **8**, 14698 (2017).

[152]M. T. Strauss, F. Scheuder, D. Haas, P. C. Nickels, and R. Jungmann, "Quantifying absolute addressability in DNA origami with molecular resolution", Nat. Commun. **9**, 1600 (2018).

[153]M. Darcy, K. Crocker, Y. Wang, J. V. Le, G. Mohammadiroozbahani, M. A. S. Abdelhamid, T. D. Craggs, C. E. Castro, R. Bundschuh, and M. G. Poirier, "High-Force Application by a Nanoscale DNA Force Spectrometer", ACS Nano **16**, 5682 (2022).

[154]H. Zhang, J. Chao, D. Pan, H. Liu, Q. Huang, and C. Fan, "Folding super-sized DNA origami with scaffold strands from long-range PCR", Chem. Commun. **48**, 6405 (2012).

[155]A. N. Marchi, I. Saaem, B. N. Vogen, S. Brown, and T. H. LaBean, "Toward Larger DNA Origami", Nano Lett. **14**, 5740 (2014).

[156]Y. Chen, W. Sun, C. Yang, and Z. Zhu, "Scaling Up DNA Self-Assembly", ACS Appl. Bio Mater. **3**, 2805 (2020).

[157]F. Hong, F. Zhang, Y. Liu, and H. Yan, "DNA Origami: Scaffolds for Creating Higher Order Structures", Chem. Rev. **117**, 12584 (2017).

[158]W. Pfeifer and B. Saccà, "From Nano to Macro through Hierarchical Self-Assembly: The DNA Paradigm", ChemBioChem **17**, 1063 (2016).

[159]M. Wang, L. Dai, J. Duan, Z. Ding, P. Wang, Z. Li, H. Xing, and Y. Tian, "Programmable Assembly of Nano-architectures through Designing Anisotropic DNA Origami Patches", Angew. Chem. Int. Ed. **132**, 6451 (2020).

[160]Y. Tian, J. R. Lhermitte, L. Bai, T. Vo, H. L. Xin, H. Li, R. Li, M. Fukuto, K. G. Yager, J. S. Kahn, Y. Xiong, B. Minevich, S. K. Kumar, and O. Gang, "Ordered three-dimensional nanomaterials using DNA-prescribed and valence-controlled material voxels", Nat. Mater. **19**, 789 (2020).

[161]P. Wang, S. Gaitanaros, S. Lee, M. Bathe, W. M. Shih, and Y. Ke, "Programming Self-Assembly of DNA Origami Honeycomb Two-Dimensional Lattices and Plasmonic Metamaterials", J. Am. Chem. Soc. **138**, 7733 (2016).

[162]W. Liu, J. Halverson, Y. Tian, A. V. Tkachenko, and O. Gang, "Self-organized architectures from assorted DNA-framed nanoparticles", Nat. Chem. **8**, 867 (2016).

[163]T. Tigges, T. Heuser, R. Tiwari, and A. Walther, "3D DNA Origami Cuboids as Monodisperse Patchy Nanoparticles for Switchable Hierarchical Self-Assembly", Nano Lett. **16**, 7870 (2016).

[164]R. Iinuma, Y. Ke, R. Jungmann, T. Schlichthaerle, J. B. Woehrstein, and P. Yin, "Polyhedra Self-Assembled from DNA Tripods and Characterized with 3D DNA-PAINT", Science **344**, 65 (2014).

[165]H. Chen, T.-G. Cha, J. Pan, and J. H. Choi, "Hierarchically assembled DNA origami tubules with reconfigurable chirality", Nanotechnology **24**, 435601 (2013).

[166]R. Jungmann, M. Scheible, A. Kuzyk, G. Pardatscher, C. E. Castro, and F. C. Simmel, "DNA origami-based nanoribbons: assembly, length distribution, and twist", Nanotechnology **22**, 275301 (2011).

[167]Z. Li, M. Liu, L. Wang, J. Nangreave, H. Yan, and Y. Liu, "Molecular Behavior of DNA Origami in Higher-Order Self-Assembly", J. Am. Chem. Soc. **132**, 13545 (2010).

[168]W. Liu, H. Zhong, R. Wang, and N. C. Seeman, "Crystalline Two-Dimensional DNA-Origami Arrays", Angew. Chem. Int. Ed. **50**, 264 (2010).

[169]C. Sigl, E. M. Willner, W. Engelen, J. A. Kretzmann, K. Sachenbacher, A. Liedl, F. Kolbe, F. Wilsch, S. A. Aghvami, U. Protzer, M. F. Hagan, S. Fraden, and H. Dietz, "Programmable icosahedral shell system for virus trapping", Nat. Mater. **20**, 1281 (2021).

[170]T. Zhang, C. Hartl, K. Frank, A. Heuer-Jungemann, S. Fischer, P. C. Nickels, B. Nickel, and T. Liedl, "3D DNA Origami Crystals", Adv. Mater. **30**, 1800273 (2018).

[171]K. F. Wagenbauer, C. Sigl, and H. Dietz, "Gigadalton-scale shape-programmable DNA assemblies", Nature **552**, 78 (2017).

[172]T. Uchida, K. Abe, Y. Endo, S. Ichiseki, S. Akita, S. Liu, S. Aradachi, M. Saito, A. Fukuchi, T. Kikkawa, T. Dammaretz, I. Kawamata, Y. Suzuki, S.-i. M. Nomura, and S. Murata, "Revolving Vernier Mechanism Controls Size of Linear Homomultimer", Small **13**, 1702158 (2017).

[173]T. Gerling, K. F. Wagenbauer, A. M. Neuner, and H. Dietz, "Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components", Science **347**, 1446 (2015).

[174]S. Woo and P. W. K. Rothemund, "Programmable molecular recognition based on the geometry of DNA nanostructures", Nat. Chem. **3**, 620 (2011).

[175]A. A. Rafat, S. Sagredo, M. Thalhammer, and F. C. Simmel, "Barcoded DNA origami structures for multiplexed optimization and enrichment of DNA-based protein-binding cavities", Nat. Chem. **12**, 852 (2020).

[176]G. Tikhomirov, P. Petersen, and L. Qian, "Triangular DNA Origami Tilings", J. Am. Chem. Soc. **140**, 17361 (2018).

[177]G. Tikhomirov, P. Petersen, and L. Qian, "Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns", Nature **552**, 67 (2017).

[178]G. Tikhomirov, P. Petersen, and L. Qian, "Programmable disorder in random DNA tilings", Nat. Nanotechnol. **12**, 251 (2016).

[179]M. Endo, T. Sugita, Y. Katsuda, K. Hidaka, and H. Sugiyama, "Programmed-Assembly System Using DNA Jigsaw Pieces", Chem. - Eur. J. **16**, 5362 (2010).

[180]A. Rajendran, M. Endo, Y. Katsuda, K. Hidaka, and H. Sugiyama, "Programmed Two-Dimensional Self-Assembly of Multiple DNA Origami Jigsaw Pieces", ACS Nano **5**, 665 (2011).

[181]T. O. Yeates, C. S. Crowley, and S. Tanaka, "Bacterial Microcompartment Organelles: Protein Shell Structure and Evolution", Annu. Rev. Biophys. **39**, 185 (2010).

[182]C. A. Kerfeld, S. Heinhorst, and G. C. Cannon, "Bacterial Microcompartments", Annu. Rev. Microbiol. **64**, 391 (2010).

[183]J. D. Perlmutter and M. F. Hagan, "Mechanisms of Virus Assembly", Annu. Rev. Phys. Chem. **66**, 217 (2015).

[184]D. Popp and R. C. Robinson, "Supramolecular cellular filament systems: How and why do they form?", Cytoskeleton **69**, 71 (2012).

[185]M. T. Cabeen and C. Jacobs-Wagner, "The Bacterial Cytoskeleton", Annu. Rev. Genet. **44**, 365 (2010).

[186]Z. Zhao, H. Yan, and Y. Liu, "A Route to Scale Up DNA Origami Using DNA Tiles as Folding Staples", Angew. Chem. Int. Ed. **49**, 1414 (2010).

[187]Z. Zhao, Y. Liu, and H. Yan, "Organizing DNA Origami Tiles into Larger Structures Using Preformed Scaffold Frames", Nano Lett. **11**, 2997 (2011).

[188]L. S. Penrose, "Self-reproducing machines", Sci. Am. **200**, 105 (1959).

[189]C.-M. Huang, A. Kucinic, J. V. Le, C. E. Castro, and H.-J. Su, "Uncertainty Quantification of a DNA Origami Mechanism Using a Coarse-Grained Model and Kinematic Variance Analysis", Nanoscale **11**, 1647 (2019).

[190]P.-G. de Gennes, "Films of Polymer-Solutions", C. R. Acad. Sci. Paris **300**, 839 (1985).

[191]E. Ruckenstein and B. Li, "Steric interactions between two grafted polymer brushes", J. Chem. Phys. **107**, 932 (1997).

[192]Y. Suzuki, I. Kawamata, K. Mizuno, and S. Murata, "Large Deformation of a DNA-Origami Nanoarm Induced by the Cumulative Actuation of Tension-Adjustable Modules", Angew. Chem. Int. Ed. **59**, 6230 (2020).

[193]D. Wang, L. Yu, C.-M. Huang, G. Arya, S. Chang, and Y. Ke, "Programmable Transformations of DNA Origami Made of Small Modular Dynamic Units", J. Am. Chem. Soc. **143**, 2256 (2021).

[194]G. Yao, F. Zhang, F. Wang, T. Peng, H. Liu, E. Poppleton, P. Šulc, S. Jiang, L. Liu, C. Gong, X. Jing, X. Liu, L. Wang, Y. Liu, C. Fan, and H. Yan, "Meta-DNA structures", Nat. Chem. **12**, 1067 (2020).

[195]N. Tanjeem, D. M. Hall, M. B. Minnis, R. C. Hayward, and G. M. Grason, "Focusing frustration for self-limiting assembly of flexible, curved particles", Phys. Rev. Res. **4**, 033035 (2022).

[196]E. G. Noya, C. K. Wong, P. Llombart, and J. P. K. Doye, "How to design an icosahedral quasicrystal through directional bonding", Nature **596**, 367 (2021).

[197]C. D. Michele, L. Rovigatti, T. Bellini, and F. Sciortino, "Self-assembly of short DNA duplexes: from a coarse-grained model to experiments through a theoretical link", Soft Matter **8**, 8388 (2012).

[198]E. Poppleton, R. Romero, A. Mallya, L. Rovigatti, and P. Šulc, "OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures", Nucleic Acids Res. **49**, W491 (2021).