From the Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

DESIGN AND ANALYSIS OF WIREFRAME DNA NANOSTRUCTURES

Erik Benson

Stockholm 2018

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Eprint AB 2018
© Erik Benson, 2018
ISBN 978-91-7831-013-5

Institutionen för medicinsk biokemi och biofysik

Design and analysis of wireframe DNA nanostructures

AKADEMISK AVHANDLING

som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Salen Jacob Berzelius, Berzelius väg 3

Fredagen den 27 April, 2018, kl 09.30

Av

Erik Benson

Principal Supervisor:
Björn Högberg
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Biomaterials and regenerative medicine

Co-supervisor(s):
Olle Inganäs
Linköpings Universitet
Department of Physics, Chemistry and Biology
Division of Biomolecular and Organic Electronics

Ola Hermanson
Karolinska Institutet
Department of Neuroscience

Andreas Nyström
Karolinska Institutet
Institute of Environmental Medicine

Opponent:
Friedrich C. Simmel
Technische Universität München
Department of Physics
Division of Physics of Synthetic Biological Systems

Examination Board:
Ulf Landegren
Uppsala Universitet
Department of Immunology, Genetics and Pathology

Erik Lindahl
Stockholms Universitet
Department of Biochemistry and Biophysics

Alessandra Villa
Karolinska Institutet
Department of Biosciences and Nutrition

Till Nils

ABSTRACT

In the last decades, the powerful self-assembly properties of DNA have been harnessed to produce complex structures at the nanoscale with high precision and yield. DNA origami is one of the most robust examples of this, where a 7000-nucleotide strand of biological origin is folded by hybridizing with hundreds of synthetic oligonucleotides, the programmed sequence of these “staple strands” determines the shape of the assembled object. The long “scaffold strand” permeates every helix of the assembled object acting as a backbone, finding the path for the scaffold strand is trivial in designs where the helices are packed on a parallel lattice but becomes challenging in other designs.

In this thesis we expand the design space of DNA origami to wireframe structures based on polyhedral meshes by the introduction of a software package consisting of: a routing algorithm for finding A-trail Eulerian circuits, a rapid physical simulation for converting the mesh to a DNA design with low strain, and vHelix, a graphical user interface for manual modification of the structure and processing of DNA sequences (Paper I). We find that this method can produce wireframe DNA origami structures with refined shapes and features, and we investigate these structures with negative stained- and cryo electron microscopy. The helices in these structures are not packed on a tight lattice and we find that they can assemble and remain stable at physiological salt concentrations unlike previously demonstrated 3D DNA origami.

We then expand this method to two-dimensional sheets (Paper II), first by generating three rectangular sheets with different vertex geometries and investigating them with atomic force microscopy to find that six-armed vertices are needed for non-distorted structures. The geometry with six-arm vertices is then used to generate four flat sheets with complex internal and external features, these structures fold with high yield to their programmed shape.

It is apparent from electron microscopy that these structures are not as rigid as structures based on the parallel packing of helices. In Paper III we study the effect of design choices on the rigidity of wireframe structures, specifically on rods where the flexibility can be estimated by measuring the persistence length. In addition to experiments we use coarse-grained molecular dynamics simulations to evaluate the rigidity *in silico*. We find that the rigidity of rods increases with increasing number of facets in the cross-section, and that the breakpoints between staples negatively affects rigidity and that his effect can be reduced by enzymatic ligation. The simulations reveal that the rigidity of the structures is greatly reduced by increasing the salt concentration.

In Paper IV we further explore the power of coarse-grained molecular dynamics simulation to predict the dynamics of DNA nanostructures. First, we track the end to end distance of the helices of a structure throughout a simulation and find that the behavior varies greatly between helices where some are practically rigid and others show large deformations. We then implement this concept in an iterative fashion where a structures rigidity is estimated by simulation and from this first generation a number of mutant structures are created by modifying one or more edges. These mutant structures are then simulated and the effect of the modifications are measured on their adjacent helices or on the entire structure, and modifications that are beneficial are inherited in the next generation of the structure. Using these methods, we are able to create a moderate evolution towards a lower flexibility in wireframe DNA origami structures.

LIST OF SCIENTIFIC PAPERS

I. Benson, E., Mohammed, A., Gardell, J., Masich, S., Czeizler, E., Orponen, P., & Högberg, B. (2015). DNA rendering of polyhedral meshes at the nanoscale. *Nature*, 523(7561), 441.

II. Benson, E., Mohammed, A., Bosco, A., Teixeira, A. I., Orponen, P., & Högberg, B. (2016). Computer-Aided Production of Scaffolded DNA Nanostructures from Flat Sheet Meshes. *Angewandte Chemie International Edition*, 55(31), 8869-8872.

III. Benson, E., Mohammed, A., Rayneau-Kirkhope, D., Gådin, A., Orponen, P., & Högberg, B. Evaluation of optimal design choices for rigidity of wireframe DNA origami structures. *Manuscript*

IV. Benson, E., Gådin, A., & Högberg, B. Evolutionary refinement of DNA nanostructures using coarse-grained molecular dynamics simulations. *Manuscript*

Additional papers not included in the thesis:

I. Zhao, Y.X., Shaw, A., Zeng, X., Benson, E., Nyström, A.M. and Högberg, B., (2012). DNA origami delivery system for cancer therapy with tunable release properties. *ACS nano*, 6(10), 8684-8691.

II. Shaw, A., Lundin, V., Petrova, E., Fördős, F., Benson, E., Al-Amin, A., Herland, A., Blokzijl, A., Högberg, B. and Teixeira, A.I., (2014). Spatial control of membrane receptor function using ligand nanocalipers. *Nature methods*, 11(8), 841-846.

III. Shaw, A., Benson, E. and Högberg, B., (2015). Purification of functionalized DNA origami nanostructures. *ACS nano*, 9(5), 4968-4975.

CONTENTS

1 Introduction ... 1
  1.1 DNA ... 1
  1.2 DNA Nanotechnology ... 2
  1.3 DNA origami ... 4
  1.4 Functionalization of DNA origami ... 6
  1.5 Biomedial applications of DNA origami ... 7
  1.6 Computer simulations of DNA origami ... 7
  1.7 Scaffold routing in DNA origami ... 8
2 Aims ... 11
3 Materials and Methods ... 12
  3.1 Scaffold routing algorithm ... 12
  3.2 Crossover placement ... 14
  3.3 vHelix ... 14
  3.4 Folding of DNA Nanostructures ... 15
  3.5 Characterization of DNA origami ... 16
  3.6 Coarse-grained simulations ... 18
  3.7 Analysis of simulations ... 20
  3.8 Iterative Simulations ... 21
4 Results ... 23
  4.1 Paper I ... 23
  4.2 Paper II ... 24
  4.3 Paper III ... 25
  4.4 Paper IV ... 27
5 Conclusions ... 29
7 Acknowledgements ... 30
8 References ... 31

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic acid
G	Guanine
C	Cytosine
A	Adenine
T	Thymine
PCR	Polymerase chain reaction
2D, 3D	Two and Three-dimensional
TEM	Transmission electron microscopy
AFM	Atomic force microscope
bp	Base pair
PBS	Phosphate-buffered saline
TBE	Tris-borate-EDTA
MD	Molecular dynamics
GPU	Graphics processing unit
BSCOR	Beam saffolded origami routing

1 INTRODUCTION

1.1 DNA

Deoxyribonucleic acid (DNA) is a molecule essential for all known living organisms, through the storage of genetic information. DNA is a linear polymer chain of four nucleotides: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T) (figure 1). The hydrogen bonding of G to C and A to T is known as base pairing, and two DNA strands with complementary sequences of nucleotides may hybridize to form an anti-parallel double helix. At physiological conditions, double-stranded DNA mostly takes the B-form conformation (figure 1.). B-form DNA has a height of each base pair (rise) of around 3.4 Å, a radius of 10 Å and makes a full turn roughly every 10.5 base pairs[1–3]. The hybridization of two complementary DNA strands is a self-assembly process, and the binding energy is dependent on the length and sequence of the complementary segment where more G and C (GC content) leads to a higher binding energy.

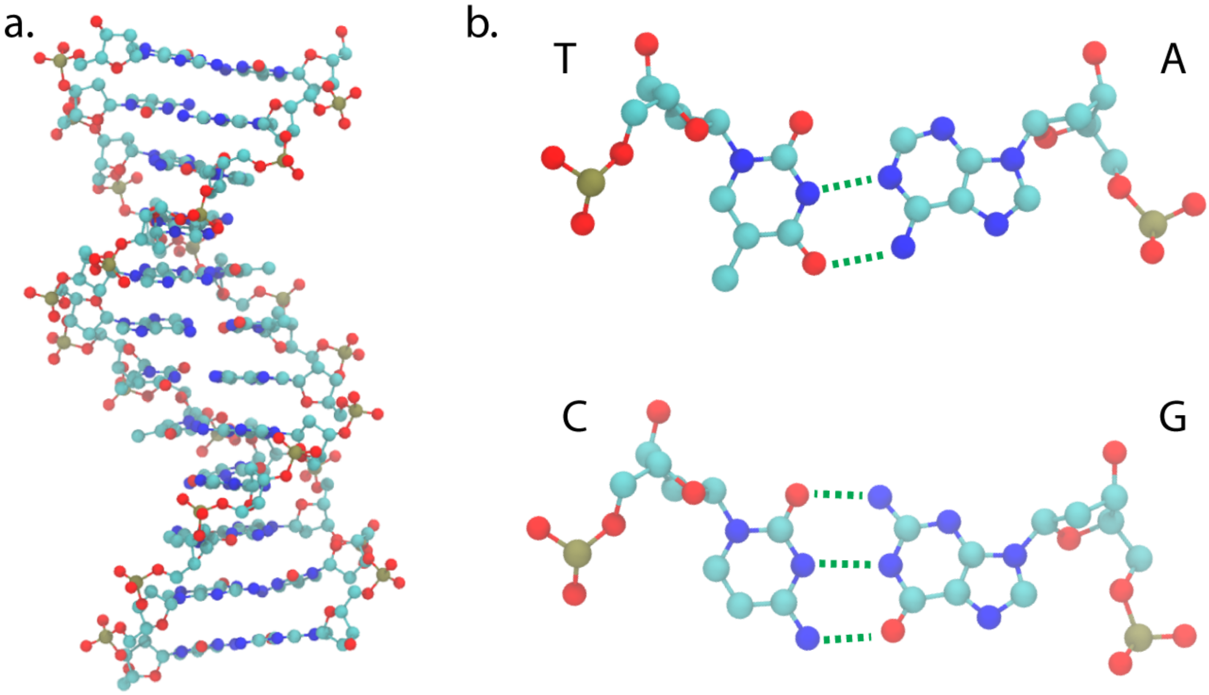

Figure 1. a. Structure of B-from DNA double helix. b. The four nucleotides shown base-pairing with hydrogen bonds rendered as green dashed lines. Rendered from a crystal structure (PDB entry 1BNA[4])

In the last decades, DNA strands with human designed custom sequences have been fundamental in several ground-breaking techniques in medicine and biology. Polymerase chain reaction (PCR)[5], DNA sequencing[6] and in-situ hybridization technologies[7] rely on short synthetic oligonucleotides to probe the function of nucleic acids in biological systems and have revolutionized our understanding of biology. A system with over 50 000 DNA oligonucleotides has been demonstrated as a storage media for arbitrary computer data[8]. Longer sequences of custom DNA are used as synthetic genes in bacterial and mammalian cells to study their function or produce proteins[9]. In the field of synthetic biology many long synthetic strands of DNA are used to replace entire genomes of organisms[10]. These techniques have created a big demand for the production of custom sequence DNA.

Double-stranded DNA may be incorporated into plasmids and cloned into bacteria where it is replicated by the cellular machinery and can easily be extracted by plasmid preparation[11]. In vitro, DNA of interest may be amplified by the use of PCR[11]. Short (typically less than 100 bases) segments of single-stranded DNA can be produced completely synthetically[12], today at costs of less than 1 SEK per nucleotide. The emerging technology of silicon chip-based DNA synthesis is promising to reduce that cost significantly[13]. Single-stranded DNA can also be produced by cloning into phage viruses with single-stranded genomes such as M13 bacteriophage[11]. The incorporated sequences may later be extracted by restriction enzymes[14] or self-cutting DNAzymes[15]. These techniques are typically more time consuming than DNA synthesis but yield higher quality DNA that may be significantly longer, and the production can be scaled up significantly.

1.2 DNA NANOTECHNOLOGY

The concept of using DNA as a construction material for structures on the nanometer scale was first introduced by Ned Seeman in 1982[16]. He was working in the field of protein crystallography and was struggling to generate crystals of his protein of interest. He envisioned that multi-armed DNA junctions could form units of a crystal lattice and that a porous DNA crystal could potentially host proteins, and x-ray diffraction could solve their crystal structure.

In nature, DNA is known to form four-arm junctions known as Holliday junctions in several biological processes[2]. Typically, these junctions consist of two pairs of identical arms and are mobile meaning that the position of the crossover between the strands can migrate on either strand. Holliday junctions can also be formed from synthetic DNA with four unique arms and are then not mobile[17]. Synthetic DNA junctions can be of arbitrary valence from three-arm junctions and up[18]. Their sequence of the strands determines the connectivity, and when mixed and exposed to a temperature ramp they will self-assemble.

In the first decade after the conception of DNA nanotechnology, a number of assemblies of a handful of oligonucleotides were shown including a cube constructed from ten oligonucleotides[19]. These structures were however too small to visualize by any imaging technique and were typically characterized by gel electrophoresis. In 1998 a two-dimensional crystal constructed from a few oligonucleotides was demonstrated[20]. In this crystal, the oligonucleotides hybridized to form unit cells called tiles containing immobile crossover junctions. These tiles were also designed to assemble with each other by “sticky” overhangs of complementary sequences, this formed large multi-micrometer carpets of DNA consisting of thousands of oligonucleotides that were easily imaged by atomic force microscopy. In the following decade the concept of using a handful of oligonucleotides designed to hybridize with each other to assemble structures expanded, both discrete structures where each oligonucleotide has only one unique position and bigger structures where each oligonucleotide bind in multiple locations (Figure 2). In the latter case, geometric design principles were sometimes used to prohibit completely uncontrolled growth, for example to assemble to one-dimensional filaments[21] or three dimensional polyhedra[22]. Ned Seeman's original concept of a designed three-dimensional DNA crystal was also realized[23].

The examples mentioned here rely on at most 20 different oligonucleotides of synthetic origin. One might expect that if one oligonucleotide of each flavor is required to assemble a structure the need to control the stoichiometry between the oligonucleotides should become a significant obstacle as the number of flavors increase. This assumption was proven wrong in 2012 with the introduction of "DNA brick structures". Here, hundreds of oligonucleotides, each 32 or 42 bases long assemble to form a two dimensional sheet[24] or a three-dimensional cube[25]. By removing specific oligonucleotides from the reactions this sheet or cube can be used as a canvas to produce custom two or three-dimensional structures. The size of this concept was recently drastically increased with the demonstration of a DNA brick structure made from 30 000 unique oligonucleotides with to a molecular weight of 500 MDa[26].

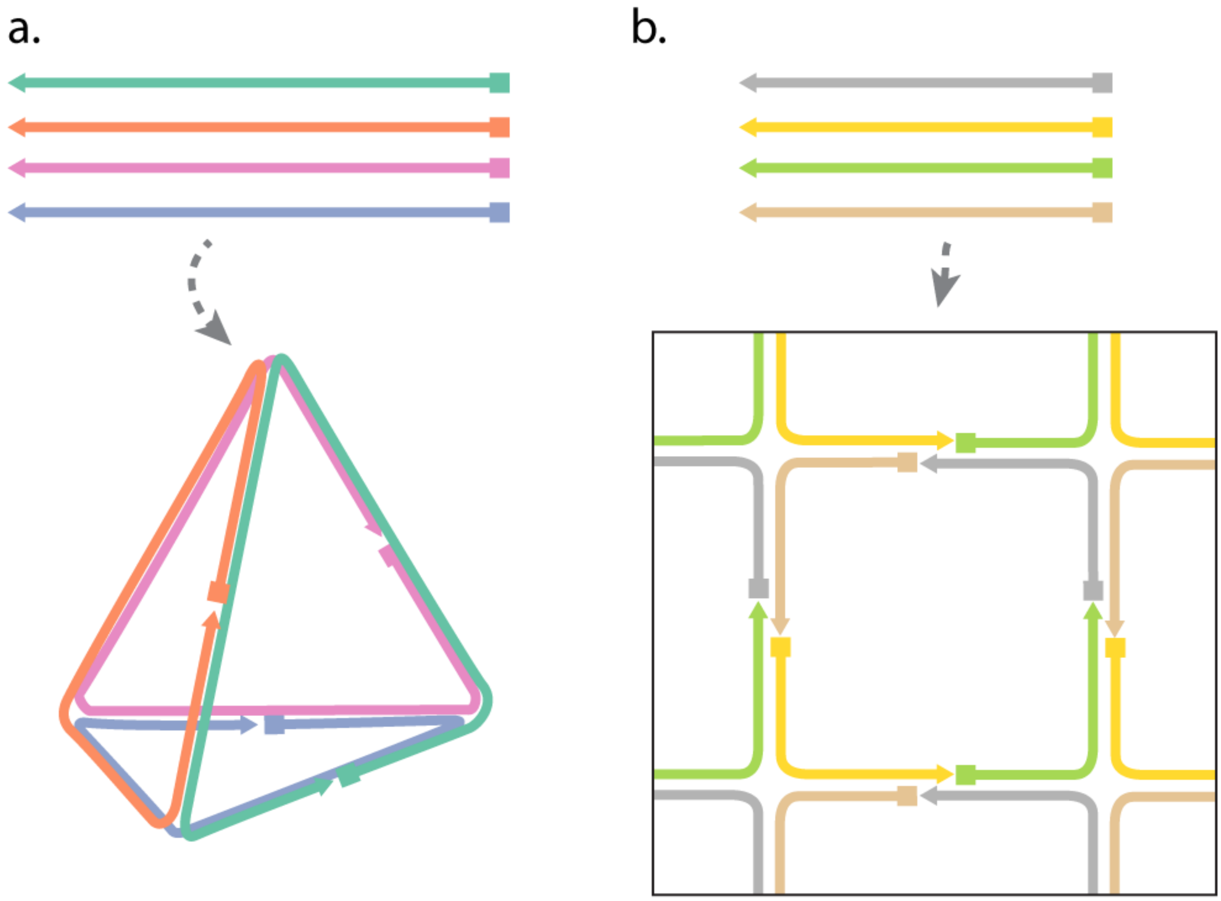

Figure 2. Examples of simple DNA nanotechnology systems, sets of oligonucleotides with programmed sequences are mixed and annealed. a. Four oligonucleotides hybridize to form a discrete three-dimensional tetrahedron. b. Four different oligonucleotides form a two-dimensional crystal of four arm junctions that grows uncontrollably.

1.3 DNA ORIGAMI

The field of DNA nanotechnology was revolutionized in 2006 with the introduction of DNA origami[27]. In DNA origami a long strand of DNA called the scaffold strand is folded by hybridizing with multiple short strands known as staple strands. The scaffold strands is usually the single-stranded genome of M13 bacteriophage with a length of more than 7000 bases. The staple strands are typically 20-60 bases long, and each staple strand is designed to hybridize with multiple regions of the scaffold strand forcing it to fold to its desired shape (figure 3).

The first demonstration of DNA origami was multiple two-dimensional sheets with diameters on the order of 100 nm consisting of a phage-based scaffold strand and hundreds of synthetic staple strands. The sheets were designed to have sophisticated internal and external features, including a smiley face structure. In these structures, the scaffold strand is weaved through the design in a raster pattern. The strands were mixed at a ratio of 100:1 of each staple strand to the scaffold and were put in a temperature ramp from 95 C to 20 C over 75 minutes. These massive structures where easily visible with atomic force microscopy showing the power of the technique[27].

This first demonstration was limited to two-dimensional sheets, however, using folding principles a sheet can be folded into three-dimensional structures. This was first shown with the creation of a hollow cubic box, similar to a cardboard box where the six walls were made as flat sheet DNA origami segments folded from the same scaffold[28]. Additional staple strands hybridized between the walls, causing them to fold into the box. This principle of using flat sheet DNA origami as the surface of a three-dimensional structure has been expanded to more complex structures including a ball, a torus, and a flask[29].

The second approach to creating three-dimensional DNA origami structures is to fold the two-dimensional sheets like a towel with additional staple cross-overs between the layers to keep the stack together. Using this approach solid or hollow brick-like structures can be created[30]. In these designs, the DNA helices are packed on a honeycomb or square lattice along one axis of the structure like the fibers in wood. By modifying the crossovers between the helices, the bricks can be redesigned to have complex internal and external features. In addition, by moving the position of the crossovers between the DNA helices, strain can be introduced to the structure causing it to bend or twist in a predictable manner[3]. The software tool caDNAno greatly assists the design of three-dimensional DNA origami structures based on this design paradigm[31].

These brick-like structures are relatively rigid and are easily visible in transmission electron microscopy. It is possible to design structures that combine rigid elements with flexible thin feature acting as joints to create simple nano-machines with moving parts[32,33].

The size of any DNA origami structure is limited by the length of the available scaffold strand. The genome of the commonly used phage M13mp18 is around 7250 bases long, this can however be increased by cloning in additional fragments of DNA giving scaffolds of upwards of 9000 nucleotides, this yields DNA origami structures with molecular weights around 5 MDa. DNA origami structures based on the genome of the lambda-phage have been demonstrated giving a scaffold of around 50 000 bases yielding structures more than five times larger[34,35]. Larger structures can also be achieved by designing DNA origami structures capable of controlled assembly, either to form crystal lattices[36,37] or discrete structures[38,39].

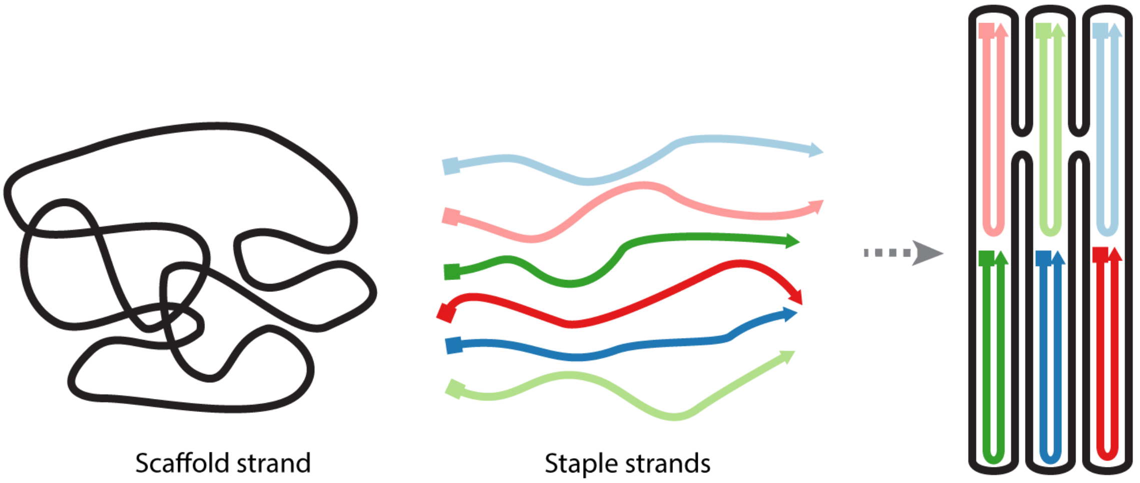

Figure 3. The principle of DNA origami. A long scaffold strand of biological origin is mixed with shorter synthetic staple strands designed to hybridize with the scaffold strand. After a thermal annealing ramp, the strands hybridize to form a DNA origami structure.

1.4 FUNCTIONALIZATION OF DNA ORIGAMI

A major focus in DNA origami research has been the functionalization of structures by the addition of non-DNA elements. A fundamental principle of DNA origami structures is that every staple strand is used one time in each origami object and its position is known. If a functional element is connected to a staple strand, then the functional element will be incorporated in the resulting structure at a known location. The staple strands are synthesized by coupling the desired nucleotides one by one to build up the strand. This process is not limited to the four DNA nucleotides found in nature and staples can be synthesized to incorporate chemical groups at either end or inside[40]. Today many companies that offer DNA synthesis also provide a range of unnatural groups for incorporation. Some groups can give the structure function on its own like fluorophores, yielding fluorescent DNA origami structures[41] or cholesterol generating structures that can interact with lipid membranes[42]. Alternatively, staple strands can be synthesized incorporating chemically active groups, allowing the staple strand to later be connected to other functional elements. If thiol groups are incorporated, they can be used to link the staple strand to gold nanoparticles[43] or surfaces[44]. Amine groups incorporated in oligonucleotides can be conjugated to proteins giving the assembled DNA structure new functions[45,46].

DNA origami structures can also be functionalized non-specifically by applying chemistry that reacts with DNA. Intercalating fluorophores or drugs[47,48] can be loaded into the structure. There is also established chemistry to coat DNA with metals[49], and this has been applied to DNA origami to create structures coated with gold[50] or other metals[51], with potential electronic applications.

1.5 BIOMEDIAL APPLICATIONS OF DNA ORIGAMI

DNA origami structures are created from a bio-material: DNA, and application in biological systems have been one of the largest research-directions in the field. Drugs that kill pathogens or cancer cells may also harm other tissues in a patient leading to side effects and lower drug tolerance. If a nanostructure is loaded with such a drug and is made to target the cancer tissue or pathogen it could act as a drug-delivery platform, lowering the side effects caused by the drug on the non-targeted tissue. DNA origami structures can be loaded with anti-cancer drugs that bind to DNA like doxorubicin and it has been shown that the drug will diffuse out of the structure over time[48]. The structures could also be loaded with biomolecules that affect the target cell or activate the immune system like antibody fragments[52], immune stimulating DNA sequences[53], or thrombin[54].

The stability of DNA nanostructures *in vivo* remains a major concern for these concepts. The phosphate backbone gives DNA a negative charge, to overcome the charge repulsion most DNA origami structures based on the parallel packing of helices needs a considerable amount of magnesium ions (5-20 mM) to fold and remain stable[30], much more than is found in human blood[55]. In the bloodstream, the structures will be targeted by the immune system and nucleases leading to degradation and/or renal clearance[56]. DNA origami structures can be coated with liposomes[57] or oligolysine[56] to increase their resistance to this.

DNA origami structures have proven useful in experimental bio-physics for studying the dynamics and forces of DNA stacking[58], DNA junctions[59], and nucleosomes[60]. The precise control of DNA origami structures have been utilized to make channels that can dock to solid state nano-pores, this could potentially be used in DNA-sequencing[61]. Cryo-electron microscopy can be used to solve the structure of proteins from many images of protein particles, it is however challenging to align the images and calculate the structure of small proteins. If a small protein is rigidly docked to a DNA origami structure, the structure of this large complex could be solved and from this the protein structure could be extracted. This approach has been used on the DNA binding protein p53, but the flexibility of the DNA origami structure limited the protein structure resolution to ~ 15 Å[62] and significantly more rigid DNA structures may be needed for this.

1.6 COMPUTER SIMULATIONS OF DNA ORIGAMI

Computer simulations of biomolecules is a valuable tool that can give unique insights into their structure and dynamics. Simulations can be used to calculate the interactions of the particles that make up the biomolecules from the most fundamental levels i.e. quantum mechanical calculations[63]. These types of calculations are very computationally expensive meaning that they are in practice limited to small systems simulated for very short timeframes. A popular abstraction from this is all-atom molecular dynamics (MD) simulations where every atom of a system including solvents and salts are represented by particles whose positions and velocities are calculated numerically by Newton's equations of motion. These methods have been used to study DNA origami structures[64], and importantly it allows the study of interactions of DNA with other types of matter like lipids[65]. These simulations are however still very computationally expensive and in practice require supercomputers to simulate on the order of 100's of nanoseconds.

The computational cost of the simulations can be reduced by coarse-graining the simulation meaning that a single particle can represent several atoms[63]. The oxDNA system is a package for coarse-grained simulation of DNA where every nucleotide is represented by two connected particles[66]. The nucleotides interact through potentials describing: backbone connectivity, excluded volume, base-pairing and stacking interactions. In oxDNA there are no particles describing the solvent (water) and salts, they are instead described implicitly through the interactions between the nucleotides, this also dramatically reduces the computational cost. The potentials used in oxDNA do not directly correspond to individual physical forces, and they need to be parameterized to give the model DNA-like properties. This is done by a top-down approach where experimental data is used as a reference for DNA behavior. The resulting DNA model can predict melting temperature and persistence lengths of DNA that correspond well with experimental values[66]. Importantly the oxDNA simulation packaged can now take advantage of graphical processing units (GPU) to speed up simulations[67].

oxDNA has been used to study a range of DNA nanostructure systems including DNA origami[68]. For brick like origami, it can predict the global twist of the structure caused by internal strain in correspondence with experimental data[66]. oxDNA has been used to simulate the assembly of a small DNA origami structure[69], but it should be noted that simulations typically start from an ideally assembled state that will not reproduce potential assembly issues in the design that would be present in experiments.

DNA origami structures can also be simulated using even simpler models, the popular tool CanDo relies on a finite element method model to simulate the behavior of structures[70]. It is available as an online tool that supports the import of parallel packed DNA origami designs on a square or hexagonal lattice and generates a heat map of the structural flexibility and a video of the simulated behavior. A new version of CanDo has been demonstrated with some lattice free structures[71], but this tool is not yet publically available.

1.7 SCAFFOLD ROUTING IN DNA ORIGAMI

The defining characteristic of a DNA origami structure is the long scaffold strand that forms the backbone of the structure. The scaffold strand must traverse the entire structure and pair with every staple strand. Finding a path for the scaffold strand that fulfills these requirements is an essential part of DNA origami design. In Paul Rothemund's original DNA origami structures the DNA double helices are packed in parallel carpets, and the scaffold routing can be found by weaving the strand back and forth throughout the design[27] (Figure 4). It should be noted that additional crossovers of the scaffold strand between helices may help hold the structure together and it may therefore be attractive to use more complex routings[72]. As these original flat sheet design principles where extended to 3D by folding flat sheets the same routing principles held and finding the scaffold path in a brick 3D DNA origami typically quite simple.

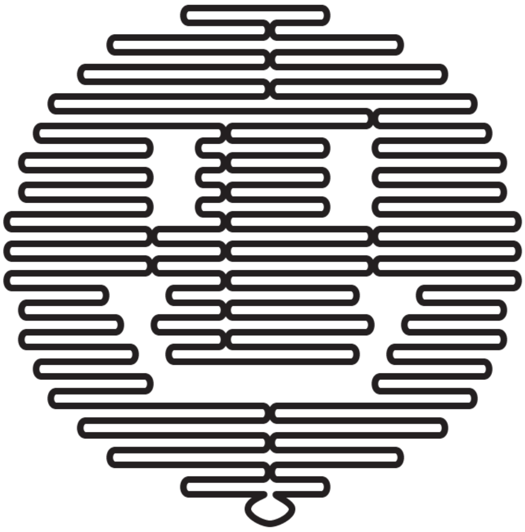

Figure 4. In DNA origami structures based on parallel packing of helices, the path for the scaffold can be produced by hand. Example scaffold routing in a flat sheet structure

2 AIMS

The overall aim of this thesis is to study the properties of wireframe DNA origami structures with the goal of establishing them as an alternative to classic DNA origami designs. The specific aims of the papers of this thesis are:

- **Paper I:** To develop the software tools for the design of wireframe DNA origami structures and investigate their assembly and properties, specifically their folding at physiological salt concentrations.
- **Paper II:** To expand the design framework to two-dimensional meshes and study the effect of vertex geometry, and investigate the surface coverage of these designs.
- **Paper III:** To increase the rigiditiy of wireframe design structures by optimizing design choises and experimental conditions.
- **Paper IV:** To study the dynamics of wireframe DNA origami strucutres in detail using coarse-grained molecular dynamics simulations, and develop an unsupervised iterative evolution towards higher rigidity.

3 MATERIALS AND METHODS

3.1 SCAFFOLD ROUTING ALGORITHM

We aim to produce DNA nanostructures based on wireframe meshes, where every edge of the mesh is represented by a double helix in the nanostructure. For these structures, it is not possible to apply the simple weaving technique to find the path for the scaffold. The designer could try to find a path for the scaffold through a wireframe design by hand. For 3D meshes, this can be aided by generating a Schlegel diagram, a flat projection of the mesh. This approach quickly becomes unfeasible as the mesh size increases.

In a DNA wireframe origami structure based on single helix edges, all crossovers between helices are positioned in the vertices. These crossovers are responsible for folding the structure to the desired conformation and holding it together. The crossovers are either scaffold crossovers where the scaffold strands cross from one helix to another or staple crossovers where a staple strand crosses. The scaffold strand routing determines the position of the scaffold crossovers, and this forms half of the connections in each vertex. The staple strand crossovers forming the remaining vertex connections should then be designed so all helices incident to a vertex is connected to two other helices. For some routings of the scaffold through a vertex it is not possible to design staple strands that properly connect the vertex, posing a constraint on the desired scaffold routings. In addition to this, we want to avoid the scaffold crossing itself as it may introduce topological traps in the folding. These two constraints limit the possible scaffold routings to a type of routings known as A-trails. In an A-trail path, the scaffold always leaves the vertex on an edge that is a neighbor to the incident edge. Staple crossovers can then be designed to connect to their cyclic neighbors to keep the vertex connected (Figure 5). It should be noted that an A-trail routing of the scaffold also allows for a staple design scheme where the staples cross the vertices and connect to a helix on the opposite side. We have performed some experimental testing of this and found that it appears to be a feasible design strategy.

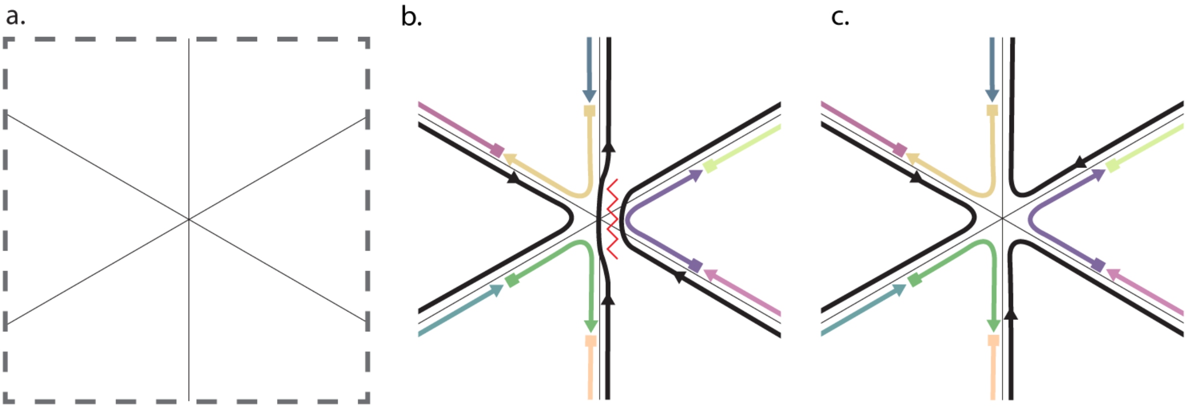

Figure 5. Scaffold routing in wireframe vertices. a. A degree six vertex, connecting six edges. b. A scaffold path that does not cross itself should lead to less topological issues, but for some routings, it may not be possible to design staples that hold the vertex together. c. In A-trail routings, the scaffold strand always exits the vertex on an adjacent edge, for these routings it is always possible to design staple strands.

The scaffold routing should visit every edge of the mesh exactly one time, in graph theory such a routing is known as an Eulerian trail, named after the 18th-century mathematician Leonhard Euler. As most scaffold strands used in experiments are circular, an additional requirement is that the path must be circular, this is known as an Eulerian circuit. It is only possible to find Eulerian circuits in meshes where all vertices have an even number of edge connections (vertex degree). The number of odd degree vertices is even in any mesh and a pair of odd degree vertices can be converted to even degree by connecting them by double edges, meaning that in the DNA design some edges will consist of two DNA double helices. Before finding the scaffold route all odd degree vertices in the mesh are paired up and then connected by the double edges (figure 6). A new routing algorithm developed for this task is then used to find an A-trail circuit through the mesh.

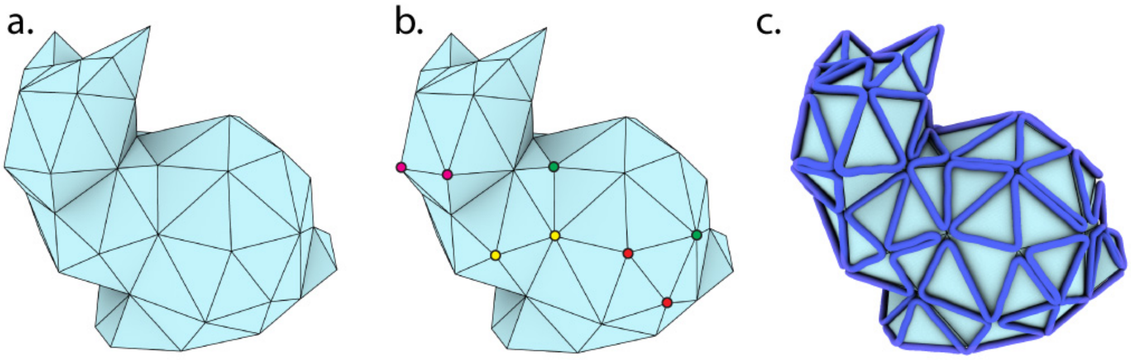

Figure 6. Scaffold routing in a polyhedral mesh (a.) b. First, all odd degree vertices are found and paired up, they will be connected by a minimal number of double edges. c. The scaffold routing algorithm finds an A-trail route through the mesh.

3.2 CROSSOVER PLACEMENT

The helical twist of double-stranded DNA must be considered when designing structures from DNA. Crossovers between connected helices should be positioned to align with the rotation of the DNA or unnecessary strain will be introduced to the structure. In traditional DNA origami structures with close-packed parallel helices the positioning of crossovers is done by design rules, if two parallel helices are connected by a crossover the DNA will make two full turns in 21 bp, and the two strands will perfectly line up for a new crossover. In 3D structures this if often expanded to a honeycomb lattice packing of the DNA where the crossovers between parallel helices are available every seven bp[30].

In our wireframe DNA nanostructures, all crossovers are positioned at the end of the helices. The orientation of the end bases of the helix will depend on the length of the helix, a 21 bp helix will make two full turns and the connections points will have the same orientation, but the addition or removal of a base pair will twist the orientation between the connection points by 34°. We aim to find DNA geometries with minimal strain at the crossovers by finding the helix lengths where the crossover points are oriented towards each other. This is done by an iterative simulation in a physics engine originally designed for computer games, PhysX. In the simulation, the DNA helices are represented by rigid cylinders with discrete lengths corresponding to a number of base pairs. Connections points representing the ends of the DNA helix are put on the cylinders with orientations determined by the number of base pairs that the cylinder represents. These cylinders are connected with spring joints at the connection points, this connected system is allowed to reach a relaxed state in the physics engine and the strain of the structure is measured by the force on the spring joints. After this, the structure is modified by adding or removing one base pair from one of its helices, leading to a rotation of the connections points on the cylinder. This new structure is again simulated and if the spring energy is lower it is assumed that the modification reduced the strain on the vertices and the modification is retained. The simulation is performed iteratively to find a DNA representation of the mesh with the minimal strain. The resulting output describes the number of base pairs and also the position and rotation of each DNA helix of the structure at the end of the simulation. We call the software package responsible for the scaffold routing and physical simulation "Beam SCaffolded Origami Routing" (BSCOR).

3.3 VHELIX

As a final step in the design process, the output file from the iterative simulation is imported to vHelix, a custom plugin to Autodesk Maya where the DNA design can be visualized and modified (Figure 7). After the import, the DNA strands have no sequence but in vHelix, a sequence can be assigned to the scaffold strand, and this will automatically calculate the sequence of the complementary staple strands. Another feature that we often use in vHelix is "auto fill strand gaps" that iterates over all connected bases in the DNA geometry to find gaps in the junctions that are large and fills these with unpaired nucleotides. In vHelix it is possible to make manual modifications to the structures like the moving of staple breakpoints. After the staple sequences have been assigned, they can be exported as a list that can be sent to a DNA synthesis company.

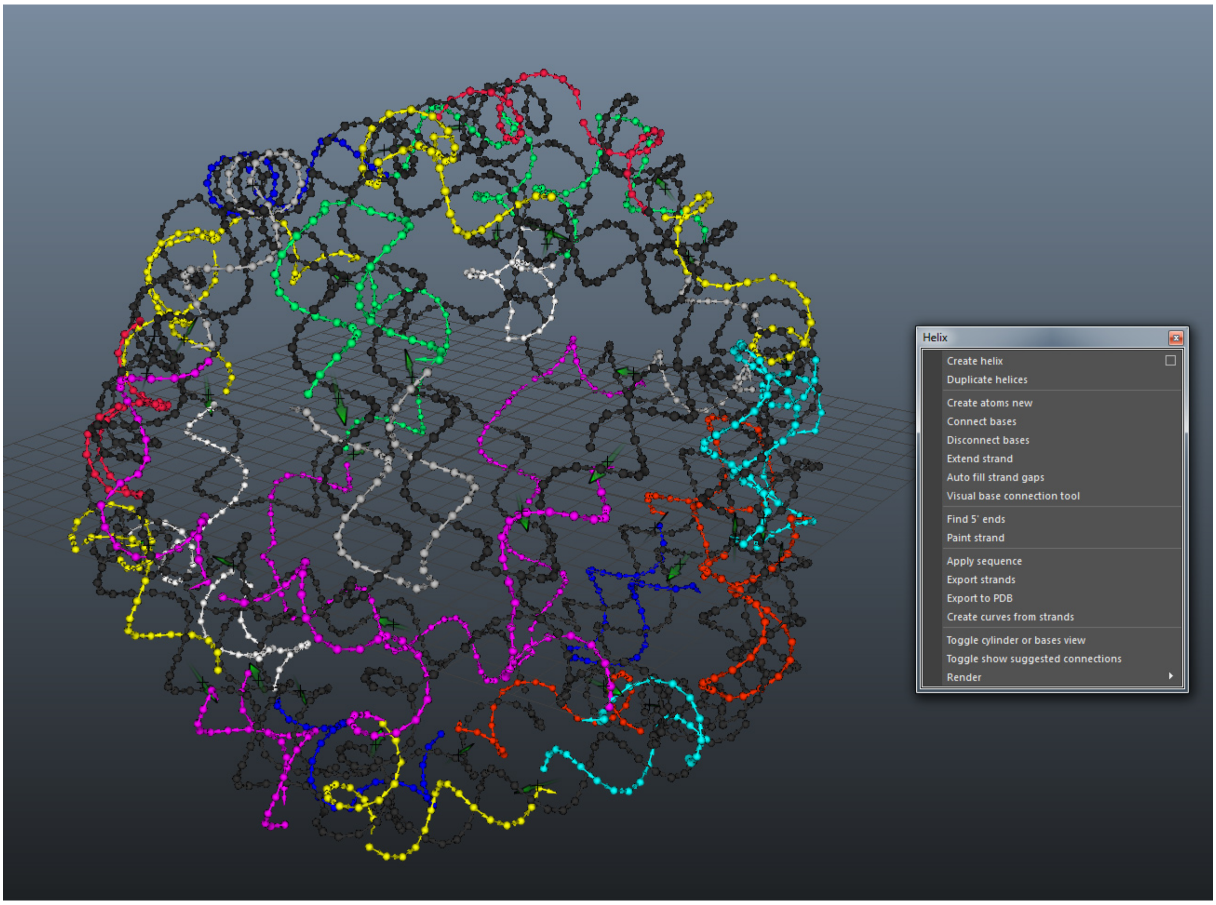
Helix
Create helix
Duplicate helices
Create atoms new
Connect bases
Disconnect bases
Extend strand
Auto fill strand gaps
Visual base connection tool
Find 5' ends
Paint strand
Apply sequence
Export strands
Export to PDB
Create curves from strands
Toggle cylinder or bases view
Toggle show suggested connections
Render

Figure 7. vHelix user interface. A DNA origami structure based on an icosahedron inside vHelix.

3.4 FOLDING OF DNA NANOSTRUCTURES

The staple strands needed to fold a structure are synthesized by a DNA synthesis company and typically arrive frozen in water in 96 well plates. The scaffold strand is the single-stranded genome of the M13mp18 bacteriophage (7249 nucleotides), or a variant of this with an insert giving a genome size of 7560 or 8064 nucleotides. The bacteriophage is propagated by adding it to an exponentially growing culture of JM 109 e-coli bacteria and then co-culturing this for 4-5 h. After this, the culture is centrifuged to remove the bacteria and the supernatant containing the phage is recovered. The phage is then precipitated from the solution by the addition of polyethylene glycol and sodium chloride, and centrifugation. A pellet of phage is then formed that is re-suspended in Tris buffer, and then the protein coat is chemically stripped from the phage genome, this protein coat is then precipitated by high speed centrifugation and the supernatant containing the DNA is recovered. Ethanol is then added to this supernatant causing the DNA to precipitate, high speed centrifugation is then used to pellet this DNA, and this pellet containing the scaffold strand is resuspended in Tris-buffer of water.

The scaffold strand and staples strands are mixed at a 1:10 molar ratio in a buffer and salt mixture that is appropriate for the structure. DNA origami structures based on parallel packing of helices usually need magnesium in the folding buffer to overcome the charge repulsion between the phosphate backbones. The amount of magnesium required varies but is typically higher for more compact structures. For wireframe structures, the packing of backbones is less dense so it may be possible to fold them in buffers that do not contain magnesium like phosphate buffered saline (PBS). After mixing the samples are exposed to a temperature ramp in a thermocycler. The temperature ramp mostly used begin with a short denaturation step where the sample is heated to 80 °C for five minutes to denature all DNA. After this, the sample is quickly cooled to 60 °C over 20 minutes and then slowly cooled to 24 °C over 14 hours.

3.5 CHARACTERIZATION OF DNA ORIGAMI

The quality of a folding reaction can quickly be determined by agarose gel electrophoresis, a 2% agarose gel is cast with 0.5 x TBE buffer, 10 mM $MgCl_2$, and 0.5 mg/ml Ethidium bromide and the samples are run at 70 V for 4h in an ice water bath. Typically, the folding reaction is run in parallel with a standard DNA ladder and the scaffold strand alone. For parallel packed 3D origami, the tighter conformation of the folded structures allows them to migrate faster than the scaffold strand even though the mass has doubled by the hybridization with the staple strands. Wireframe structures often migrate slower than the scaffold strand due to their larger size. Agarose gels also reveal if the structure is monomeric or has multimerized or aggregated. When DNA structures are functionalized, this may also be detected in gels by altered migration speed (figure 8).

In addition to gels, a multitude of imaging techniques is used to characterize the structures (Figure 8). Structures can be immobilized on a surface, commonly mica and imaged in atomic force microscopy (AFM). For single layer structures this may allow the imaging of individual DNA helices, however, the immobilization on mica tends to flatten and stretch out the structure and AFM images may not reflect the solution properties of a DNA origami structures. For 3D DNA origami structures, negative stained transmission electron microscopy is often used. Here, a drop of sample is applied to a glow discharged formvar-carbon grid for 20 seconds, then the grid is blotted on a filter paper to remove most of the liquid. This grid is then put on a drop of water for 3 seconds to reduce the amount of salt on the grid and then again blotted on a filter paper. As a final staining step, the grid is put on top of a drop of 2% solution of uranyl formate for 20 seconds and then blotted on the filter paper. The grid is then left to air dry and is then imaged in a transmission electron microscope, typically at magnifications from 11 000 to 28 000 x. This is a rapid technique that gives relatively high-resolution images of close-packed structures, but wireframe structures tend to collapse in the dried out conditions giving images that may not reflect the solution structure.

Cryo-electron microscopy can be used to capture the solution structure of DNA origami objects. Here, the sample is plunge-frozen to form amorphous ice, preserving the solution conformation of the structure. As no stain is added to the sample all contrast comes from the electron density of the sample, and only a limited amount of information can be collected from each DNA nanostructure before it is destroyed by beam damage. This can yield low contrast 2D images of the structure or using cryo-electron tomography, the 3D structure of individual particles can be extracted at low resolution. The information from 2D images of many structures may be combined to a 3D model using single particle analysis. For rigid biomolecules, this has been used to create high-resolution models[73]. For DNA nanostructures it appears that the flexibility of the structures limits the achievable resolution, with the highest reported resolution 11.5 Å[74] However, even at a lower resolution this technique may still be very useful for DNA nanostructure design.

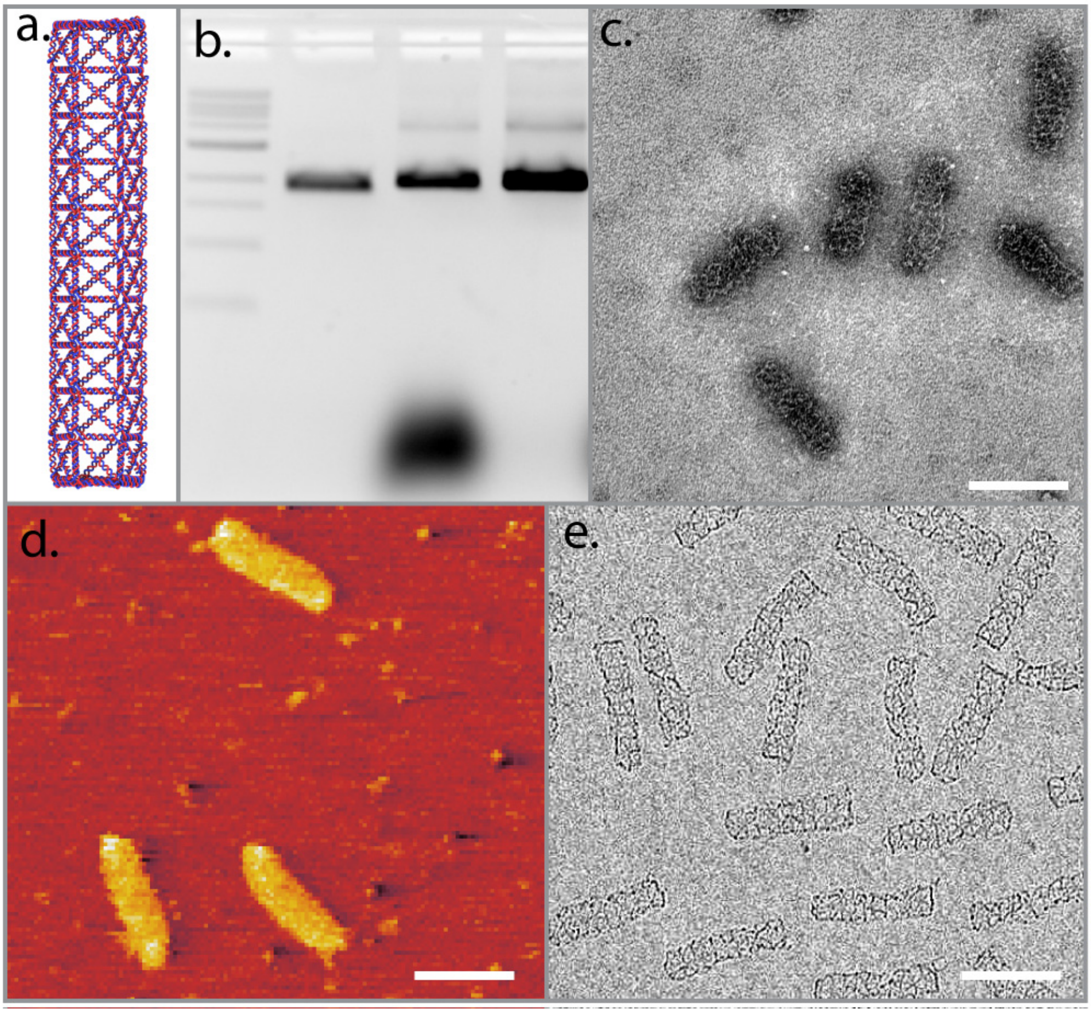

Figure 8. Characterization techniques for DNA origami structures. a. Rendering of a rod-like structure with a hexagonal cross-section. b. Agarose gel electrophoresis of the structure, lane 1: DNA ladder, lane 2: scaffold strand, lane 3: folded rod structure, lane 4: folded structure with excess staple strands removed. c. Negative stained transmission electron microscopy (TEM) of the structure. d. Atomic force microscopy (AFM) image of the structure. e. Cryo-electron microscopy image of the structure. All scale bars 100 nm.

3.6 COARSE-GRAINED SIMULATIONS

The oxDNA package can be downloaded from https://dna.physics.ox.ac.uk and compiled in Linux, if the system has a modern graphics processing unit from Nvidia it may also be compiled with CUDA support to enable GPU acceleration of the simulation.

An oxDNA simulation is started with the system described in two files, the topology file that describes what nucleotides are in the system and how they are connected and the configuration file that describes the coordinates including rotation of these nucleotides. These files can be generated from DNA nanostructure design files using converters developed by the oxDNA team. The converter for vHelix structures works on DNA structures saved in the Maya ASCII format and was originally developed in Jonathan Doye’s group. We modified it to also generate a list of the particle ID of the second to last bases of every helix for further analysis of the rigidity of the helices.

The nucleotide coordinates of the DNA structure that will be simulated are based on the coordinates inside vHelix, and these coordinates may contain unphysical distances between backbones that are covalently linked, specifically in the vertices of the structures where two connected helices may be placed unrealistically far away by the simple physical simulation. If these coordinates are loaded directly into a molecular dynamics simulation the force between covalently linked backbones that are unrealistically far apart could be very large giving the connected components very high velocities causing the simulated system to "blow up". In oxDNA, this is overcome by running pre-simulations to relax the structure and bring all covalently bonded particles into physically realistic distances. In the pre-simulations, the force between connected backbones is capped to a maximum value to prevent the simulated system from blowing up (Figure 9).

oxDNA allows parameters like temperature and concentration of monovalent salts to be set for the simulations. The output of oxDNA is a trajectory file that captures snapshots of the simulation including position, orientation, and velocity of all simulated nucleotides, and this data is typically saved for every 5000 time steps to save space. A simulation of a DNA origami structure based on a 7000 base scaffold for 100 million time steps can typically take 1-2 days on a modern GPU (Nvidia GTX 1080Ti). This corresponds to a simulated time of around 1.5 µs.

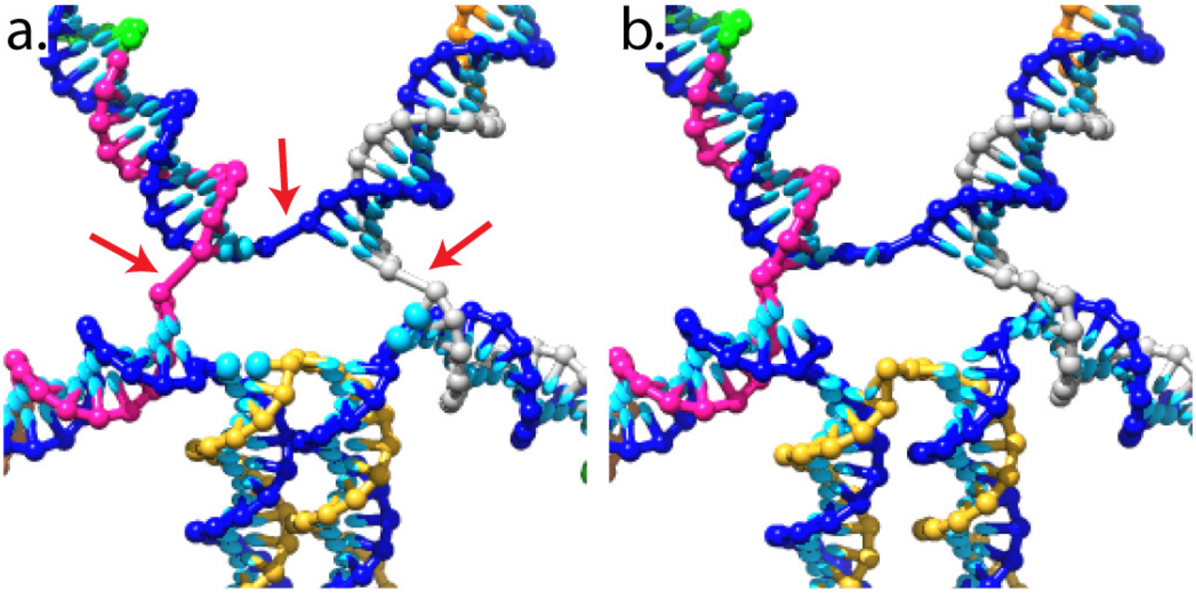

Figure 9. Example image of the oxDNA representation of the vertex of a wireframe DNA origami structure. a. The vertex as designed in vHelix where the distance of covalently connected backbones may be large (arrows). b. After a relaxation simulation the connected backbone particles are brought closer.

3.7 ANALYSIS OF SIMULATIONS

The coarse-grained simulation of DNA nanostructures should give insight into their structure and dynamics. The simulation trajectory can be viewed in molecular visualizations software as a video providing a qualitative analysis of its behavior. The trajectory files, describing the position of the particles throughout the simulation can also be analyzed by qualitatively by scripts. We analyze the rigidity of the helices that constitute the structure by tracking the distance between one base on the second to first base pair and one base on the second to last base pair for each helix. This distance is essentially the end to end distance of each helix, and one would expect that for a perfectly rigid helix this distance should be constant throughout the trajectory while a flexible helix would have a varying distance. The distance can be plotted as a trajectory or be further described by a single metric, the standard deviation of the distance data-points. Rigid edges that have low variability in the end to end distance yield a small standard deviation. This can be further abstracted to score the rigidity of a whole structure by calculating the average or median of the standard deviation of all its helices.

For rod-like structures, we can evaluate the rigidity by calculating their persistence length that can be compared with experimental results. This is done by tracking the position of a few nucleotides in each plane of the rod. The average coordinates of these give the center of the planes and form a spine through the structure by the vectors that connect them. The average dot products between vectors of the spine will be smaller for vectors that are further apart on the spine, this decline is described by the persistence length of the structure that can be fitted to these values[75] (Figure 10).

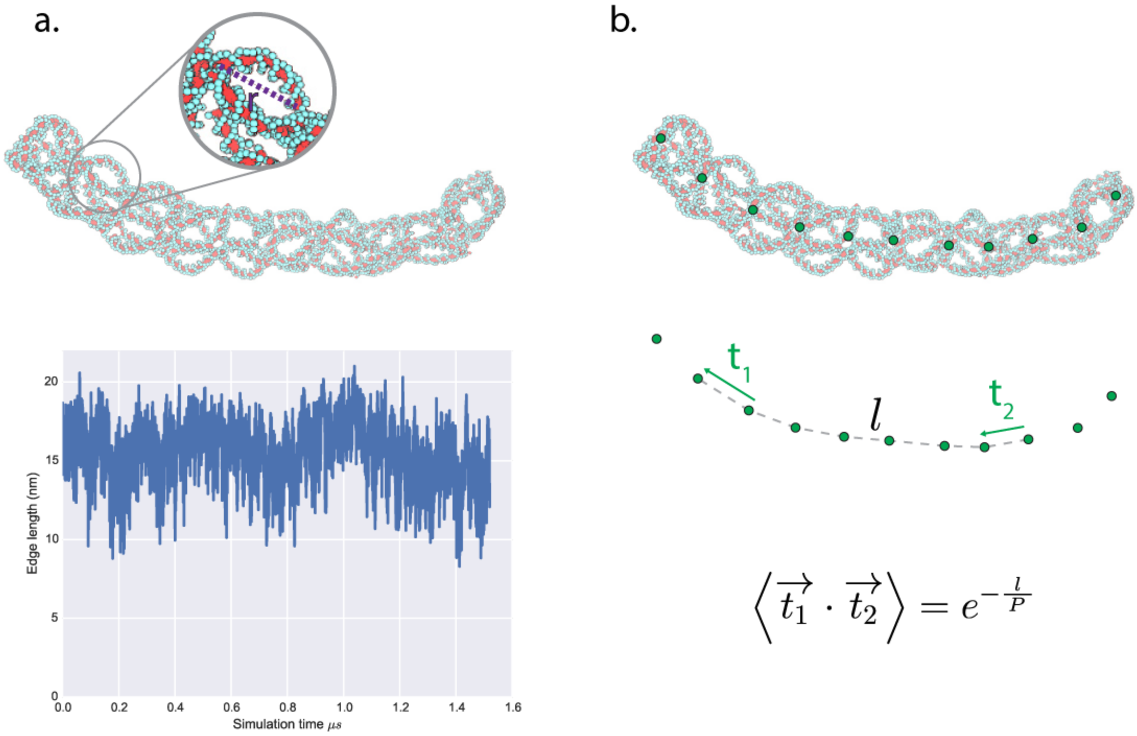

Figure 10. Analysis of rigidity of DNA nanostructure from coarse-grained molecular dynamics simulations. a. Individual helices of a structure can be monitored by tracking the end to end distance throughout the simulation trajectory. b. The rigidity of the entire structure can be estimated by calculating the persistence length, this is done by calculating average points in the center of the cross-section planes, generating a 3D spine of the structure. The correlation between tangents of this spine is expected to decay as the contour length l increases, and to this, the persistence length function can be fitted.

3.8 ITERATIVE SIMULATIONS

After the simulation of a structure, the flexibility of its helices can be estimated by tracking their end-to-end distance, this allows us to rank the helices from "worst" to "best". For many structures, this reveals that many edges are rigid with low flexibility while some edges show high flexibility. It is possible that the flexibility in certain edges is due to a non-optimal conversion from mesh to DNA design due to the simple DNA representation used in the physics simulation. If this hypothesis is true, it should be possible to increase the rigidity of DNA nanostructures by modifying the number of base pairs on certain edges and estimate the effect of such modifications by simulation. As the system of helices is connected, modifications to one helix will change its twist and therefore the strain on the connection to its neighboring helices. Attempts to improve the structure by modifying its helices will probably have to be done iteratively where a modification is evaluated and incorporated into the structure if it lowers the flexibility of the structure. This modified structure will then be the base for the next round of modifications.

We aim to create an unsupervised iterative process to reduce the flexibility of the structures. The primary scoring metric we use is the standard deviation of the end-to-end distance that can be used to score individual edges or summed up to score a structure. A structure will be modified by adding removing base pairs from one or several edges to generate mutant structures. The mutant structures will then be simulated in oxDNA and analyzed to investigate if their modifications were positive, if so they will be incorporated in the next round of mutant structures.

For each iteration, an arbitrary number of mutant structures can be generated but as oxDNA can run one simulation instance per GPU, it is most practical to use the same number of mutant structures as available GPUs. Typically, we use eight GPUs in a setup and thus use eight mutant structures for each iteration. It is most practical to have four GPUs per computer so the simulations will have to run over a network of computers. We set up a system where the modifications and evaluation of the structures are done on one master computer. This computer generates the mutant structures and converts them to oxDNA simulation format files that are sent over the network to the GPU nodes where the simulations are run. After simulation, the GPU nodes perform pre-analysis on the trajectories and extract the coordinates for the end bases of the helices. These are then sent over the network to the master computer that analyses and scores the mutant structures.

4 RESULTS

4.1 PAPER I

In paper I we started from seven 3D polyhedral meshes depicting both simple shapes like a sphere and a rod, and more complex shapes like a stick-man and a simplified version of the Stanford bunny. The meshes where all triangulated and all contained some odd degree vertices. The structures were converted to DNA nanostructures using the BSCOR package and scaled to use from 5891 to 7797 bases of scaffold strand.

After folding the structures, they were first analyzed in agarose gels, close-packed DNA origami usually migrate faster than the scaffold strand due to their compact size. Interestingly these structures migrated slower than the scaffold strand, this may be due to their larger size. Negative stain TEM was then used to verify the folding, this revealed that the structures had folded to their predicted shape but the dry state of the imaging caused the structures to collapse (Figure 11). To overcome this, cryo-electron microscopy was used to image the structures in a hydrated state. In addition, cryo-electron tomography gave 3D information about the shape of the structures.

For all structures, except for the ball, we filled used the feature "auto fill strand gaps" to find gaps between connected bases in the vertices and fill them with unpaired nucleotides. The ball folded with high yield despite this but for the pentagonal rod in the paper, a version of the structure without the vertex padding did not fold.

Our standard design style for staple strands is to connect two adjacent helices with breakpoints in the middle of each edge. Some structures in this paper have a large difference of length between their shortest and the longest edges. When realized in DNA, the shortest edges may correspond to helices that are just 8 base pairs long. If these are split in half by a staple breakpoint the regions for hybridization become very short. We therefore tested alternate staple design schemes where staples hybridize to more than two edges, typically three. The risk of this scheme is that if the "outer" regions of the staple strand hybridize first, making it impossible for the inner region to wrap around the scaffold and hybridize. We tried to overcome this by moving the breakpoints from the center for the outer hybridization regions of the staple strand so that at least one of the hybridization regions would be shorter than the middle part of the staple strand to avoid these topological traps.

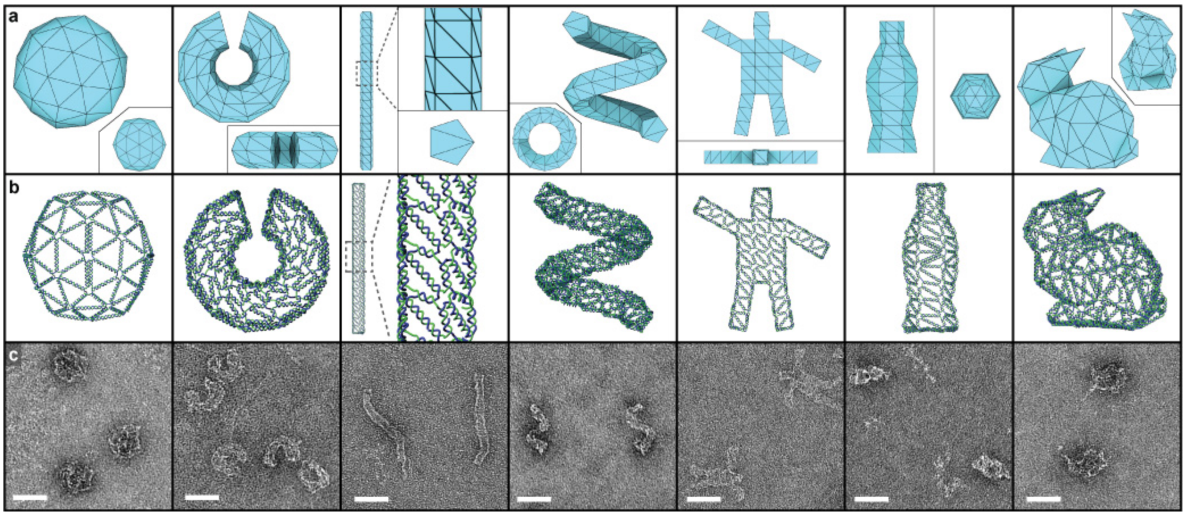

Figure 11. DNA origami structures designed from polyhedral meshes. a. Polyhedral computer meshes used as design templates b. Rendering of DNA designs generated from the meshes. c. TEM images of the folded DNA nanostructures, scale bars 50 nm.

In the paper, the ball and the helix structure folded with high yields of around 90%, for the other structures the yields where lower, probably due to multimer formation and aggregation, it is possible that optimization of folding parameters may increase the folding yield. Interestingly, several of the structures not only folded in typical folding buffers with magnesium but also in the more physiological salt buffers: Phosphate buffered saline (PBS) and Dulbecco's modified Eagle's medium (DMEM), this may be useful for biomedical applications of DNA nanotechnology.

4.2 PAPER II

In paper II, we expand our study of DNA origami structures based on wireframe meshed to two-dimensional meshes. To do this, we made modifications to the BSCOR pipeline to reliably support the routing of corner vertices and modify to the spring relaxation to better support the outer edges of flat sheets. Many proposed applications of DNA origami rely on flat sheet structures[76,77] with parallel packing of helices leading to a high density of DNA and a limited size. The sparser geometry of wireframe DNA origami structure should allow for larger flat sheets from the same amount of scaffold.

In the paper, we first investigate the effect of vertex geometry, aiming to get the sparsest possible mesh. We create rectangular meshes with vertices that have three, four or six connections, this corresponds to the three regular tessellations. As it is not possible to find a scaffold routing in meshes with odd degree vertices the mesh with degree tree vertices is converted to a nanostructure with four helices connecting each vertex by the introduction of double edges.

We folded the structures, and agarose gel electrophoresis revealed that they folded with high yield. We then immobilized the structures on mica and performed atomic force microscopy

(AFM). We found that the structures based on three- and four-degree vertices deform substantially (Figure 12), this is not surprising as four-arm DNA junctions are known to take a stacked conformation in the presence of divalent cations[78,79]. The structure based on degree six vertices folded to its designed shape. We compared the area coverage of these structures to a rectangular sheet designed by Paul Rothemund based on the parallel packing of helices. We found that for the same amount of scaffold bases, wireframe DNA nanostructures covered a 70-95 % larger surface area.

We designed four more structures based on six-arm junctions, with complex internal and external features, including a smiley face and a map of Scandinavia. These structures all folded to their designed shape indicating that triangulated wireframe structures are a good design scheme for flat sheet DNA nanostructures.

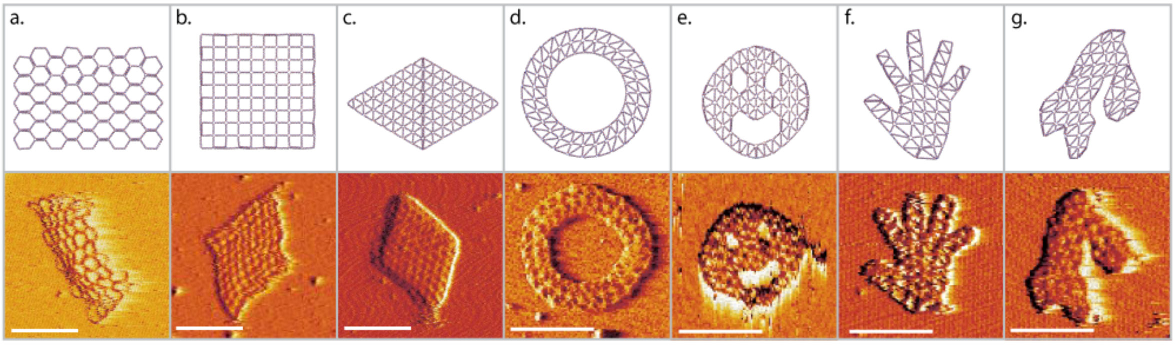

Figure 12. Two-dimensional DNA origami structures generated from meshes. Top row: renderings of DNA nanostructures a-c: rectangular sheets with vertices designed to have 3, 4 or 6 connections, d-g: sheets with vertices with 6 connections and complex internal and external features. Bottom row: AFM images of the folded structures immobilized on mica, scale bars 100 nm.

4.3 PAPER III

Wireframe DNA origami structures mimic the latticework construction used in macroscopic engineering to achieve high strength to weight ratios. For DNA nanostructures, wireframe designs often appear less rigid than solid beam designs. In paper III we study the effect of design choices on the rigidity of wireframe structures, specifically we study rods where we can estimate the rigidity by measuring the persistence length of the structures.

When converting a computer mesh to a DNA origami design, a scaling factor is introduced that controls the size of the DNA design, by varying this factor arbitrary size DNA nanostructures can be created from the same mesh. We first study the effect of this by beam theory which indicates that there is two competing effect at play. For an idealized latticework frame, the rigidity should increase as the scaling value increases its cross-section, this assumption is however only valid when the DNA edges are significantly shorter than the persistence length of DNA and in principle act like rods. When the DNA edges grow to a significant fraction of their persistence length their end to end distances start to assume a wider distribution. A frame assembled from these DNA edges will no longer take its predicted shape, causing a shorter persistence length. We speculate that these two competing effects may yield an optimal DNA edge length for wireframe assemblies.

We use the oxDNA package to study this effect *in silico* by generating seven rods with a triangular cross-section that were isotopically scaled and six rods with a triangular cross-section where the cross-section size was scaled but not the length. For the isotropic scaling, no clear trend could be seen while the cross-section scaling indicated an optimal value with a base side-length of around 16 nm. We also tested the effect of the number of facets in the cross-section profile and found that the persistence length grew drastically when this was increased, however at the cost of more material use. Interestingly, varying the salt concentration in the simulations had a drastic effect on the persistence length, where lower salt concentrations gave higher persistence length. We hypothesize that the electrostatic repulsions between the DNA of the structure act as an inflating and rigidifying force and that the charge shielding of ions reduces this effect.

We investigated the effect of isotropic scaling experimentally by creating three rods with square cross-sections, but varying edge lengths, here the smallest rod with a side length of 13 nm had a significantly shorter persistence length than the larger structures with side lengths of 16 and 21 nm (figure 13). The effect of increasing the number of facets in the cross-section was also investigated by the comparison of three rods with similar side-length but three, four or six facets. This showed that the persistence length increased with an increased facet count. Another potential source of flexibility in these nanostructures is the breakpoints between staples that are positioned in the center of every edge. They are in practice DNA nicks and can be removed by a ligase that fuses the two staples together. We tested this experimentally and found that the ligation of the structure significantly increases the persistence length. This is an attractive approach as it has no extra material cost in the design and may decrease the vulnerability of the structures to exonucleases.

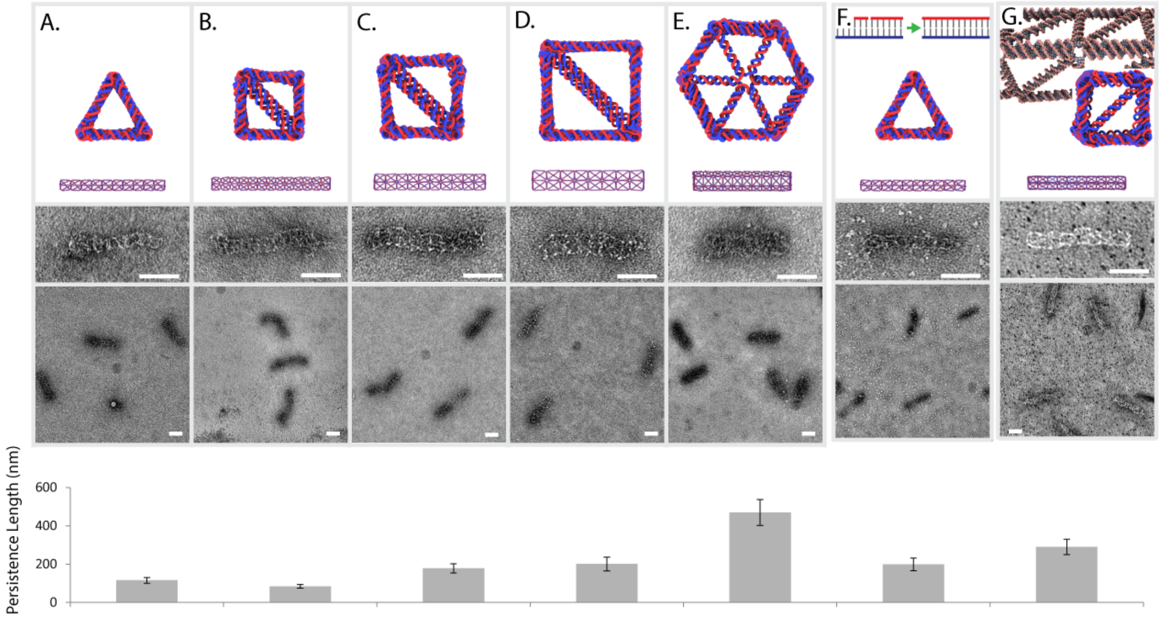

Figure 13. The rigidity of wireframe rods with varying designs. Top row: rendering of design, middle rows: TEM images of folded structures, scale bars 50 nm, bottom row: persistence length calculated from TEM images. a. A rod with a triangular cross-section. b-d. Rods with square cross-sections and increasing scaling. e. A rod with a hexagonal cross-section. f. The rod with a triangular cross-section with staple breakpoints removed by enzymatic ligation. g. A rod with a square cross-section with axial members reinforced by double DNA edges.

4.4 PAPER IV

Coarse-grained molecular dynamics simulation can give predictions about the dynamics of DNA nanostructures. In Paper IV we first introduce a simple method for analyzing the dynamics of the individual helices of a wireframe structure by tracking their end to end distance throughout the simulation, the fluctuation in the end to end distances is estimated by calculating the standard deviation of this. When this metric is applied to wireframe DNA origami structures it becomes apparent that some helices have essentially constant end to end distances, consistent with B-DNA while other edges show large fluctuations and average end to end distances that are far below a straight B-DNA prediction, meaning that they bend in the simulation. This metric can be expanded to the entire structure by calculating the average of the standard deviation of all helices, with a higher value corresponding to a higher average flexibility. This allows us to rank the helices of a structure from most to least flexible and compare the rigidity of similar structures. This can be used to modify helices of a structure by adding or removing base pairs and then evaluating the effect by simulating the original and mutated structure and comparing the score to evaluate if the modification leads to lower flexibility.

We implement this concept in an automated fashion where a wireframe DNA origami structure is first simulated and its edges scored and ranked from highest to lowest flexibility.

A number of mutant structures are then generated that each has an addition or removal of base pairs from an edge with high flexibility. These mutant structures are then simulated in oxDNA, and the average standard deviation of the helix end to end distance is calculated for them as a score of the entire structure. A modification is considered positive if it leads to a lower average standard deviation for the mutant structure. In our software, we use two schemes for the progress of iterations, one where the best mutant structure is selected even if doesn't constitute an improvement over the previous iteration, with the hope that this may reduce the risk of finding local minima. In the other scheme, the mutations are incorporated only if they lead to a score that is better than the previous best. We tested both schemes on a barrel-shaped small DNA origami structure and found that they possibly showed moderated improvements in the overall standard deviation.

These simulations were performed for 10^8 time steps taking around 22 hours to complete. Ideally, the simulation should be long enough to differentiate the performance of similar structures in a repeatable way. We tested this by performing duplicate simulations of a mutant version of the same structure for $7.5 \cdot 10^8$ time steps and analyzed if the duplicated of the same structure converged and if the converged value is different between two mutant structures. This indicated that simulations needed to be more than $3 \cdot 10^8$ time steps to yield reliable results. We attempted an optimization of a very small structure with $5 \cdot 10^8$ time step simulations, this initially yielded a positive trend but after nine iterations this was reversed. A closer analysis of the simulations showed that the reversal was due to the simulation consistently scoring the same edges as worst and incorporating modifications to these edges.

The approach of only making one modification in each mutant structure is intrinsically slow and the effect of single modifications on entire structures is likely small, especially for larger structures meaning that a metric based on the average of all helices may not be optimal for evaluating single modifications. To overcome this, we introduce a multiplexed iterative scheme where multiple random but spatially separated modifications are implemented in each mutant structure. In each iteration, a reference structure is simulated without these random modifications. After the simulation, the modifications are evaluated by calculating the sum of the standard deviation of the modified edge and the edges that are connected to its vertices, this sum is compared to the sum for the same edges in the reference structure. If the sum in the modified structure is lower than in the reference, the modification is considered beneficial and incorporated into the reference structure in the next iteration. We used this approach for a full-size DNA origami rod, and initially we saw a strong positive downward trend that was broken after six iterations and then seemed to transform into a new slower trend towards lower flexibility. This iterative approach is not limited to this simple scoring metric and could also be implemented by for example scoring the persistence length or similarity to an initial design for a structure.

5 CONCLUSIONS

DNA nanotechnology has previously been demonstrated as a powerful method for the creation of self-assembling nanomaterial, especially through the DNA origami technique. In DNA origami the scaffold strand is both a strength, trough simplified stoichiometry and its role as a backbone for the structure and a complication as a result of the design challenge caused by the routing of it through all members of a structure. In this work, we expand the design space of the DNA origami technique to wireframe mesh designs through a software package containing: a routing algorithm, a physical simulation to minimize vertex strain, and a graphical user interface for the modification of structures. We demonstrated this by designing and folding a number of two and three-dimensional structures, these structures were then studied by multiple microscopy techniques to validate they had folded to their predicted shapes. We find that unlike DNA origami structures based on close packing of helices, wireframe structures can fold and remain stable at physiological salt concentrations. Two-dimensional structures based on wireframe meshes can cover a significantly larger surface area than flat sheets based on parallel packing.

The electron microscopy studies of wireframe structures indicate that they are less rigid than comparable structures based on parallel packing. We studied the effects that design choices have on flexibility with the aim of creating more rigid structures. Here, we complemented our experimental studies with coarse-grained molecular dynamics simulations of the structures. We find that the rigidity of rods increases with the number of facets in the cross-section and that the staple breakpoints on the edges play a prominent role in the flexibility of the structure. The simulations indicate that increasing salt concentrations have a detrimental effect on structural rigidity.

We have created a software system that can generate modified mutant versions of a structure and simulate these using coarse-grained molecular dynamics simulations. The effect of these modifications on the structure properties is then evaluated and the mutation is incorporated if it yields a higher performing structure. This is run in an unsupervised fashion and seems to result in a moderate evolution of the structure towards higher rigidity.

7 ACKNOWLEDGEMENTS

First, I want to thank my supervisor **Björn Högberg** for giving me the opportunity to work in his lab and teaching me everything about DNA: from running a gel to extruding a curve with paint effects in Maya.

I want to thank my co-supervisors, **Andreas Nyström** for teaching me something about polymers and about workers rights at KI, **Olle Inganäs** for helping me get hired, teaching me about PEDOT-S, and helping me realize that some projects take time to finish, and **Ola Hermanson** for helping out with my registration and always being helpful.

I want to thank the members of the group for taking me in: **Mino** for inviting me to his family, teaching me Italian culture, and useful Italian phrases like: [redacted], [redacted] and [redacted]. **Alan**, for trying to teach me chemistry, letting me witness his first marriage and sharing his innovative takes on Swedish driving rules. **Ferenc**, for his patience, introducing me to palinka, and many interesting discussions about any topic. **Gianna**, for sharing her knowledge about biology, and for listening to my many complaints. **Giulio** for being wiki-Giulio in the office and for paying the bills ☺. **[ʒuˈɐ̃w],** for all the time we spent together and for teaching me the importance of pronunciation. **Ian** for confusing me with math and helping me find the fastest way to willys. **Esther**, for always keeping me and everyone else happy. **Johan** and **Andreas** for teaching me so much about computers. And the other lab members: **Masa, Corinna, Orsi, Marco, Yong-Xing, Yang**

I want to thank other friends and collaborators at KI, **Domenico** for almost buying my Roomba. **Alessandro**, for showing me the fun and frustration of AFM. **Ana, Toon, Ben, Chris, Teresa, Daniela, Silvia, Spyros**, **Xueshu.**

For my collaborators in Finland: **Pekka, Melik, Eugen** and **Daniel**, for helping me see that topics that may seem different may have big overlaps, and for not thinking that any of my questions are too stupid. Also, thank you for teaching me Finish history and showing me Helsinki!

Thank you **Mahiar** for being a great mentor and teaching me about life and science (and PEDOT-S)

I want to thank my **parents** and **Peter** and **Elisabet** for always supporting me.

Finally, I want to thank **Elin** and **Nils** for being there night and day and showing me that the most important things in life are outside the lab.

If I forgot to thank **you**, it was just due last-minute stress, please enter below:

I want to thank ……………… for …………………………………………

8 REFERENCES

1. Wang, J. C. Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci.* **76,** (1979).
2. Neidle, S. *Principles of Nucleic Acid Structure*. (Elsevier Science, 2008).
3. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–30 (2009).
4. Drew, H. R. *et al.* Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci.* **78,** 2179–2183 (1981).
5. Saiki, R. *et al.* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230,** 1350–1354 (1985).
6. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74,** 5463–5467 (1977).
7. Levsky, J. M. & Singer, R. H. Fluorescence in situ hybridization: past, present and future. *J. Cell Sci.* **116,** 2833–2838 (2003).
8. Church, G. M., Gao, Y. & Kosuri, S. Next-generation digital information storage in DNA. *Science* **337,** 1628 (2012).
9. Itakura, K. *et al.* Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin. *Science* **198,** 1056–1063 (1977).
10. Gibson, D. G. *et al.* Complete Checmical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. *Science* **319,** 1215–1221 (2008).
11. Sambrook, J. & Russel, D. W. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001).
12. Caruthers, M. H. Gene synthesis machines: DNA chemistry and its uses. *Science* **230,** 281–5 (1985).
13. Kosuri, S. *et al.* Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol.* (2010). doi:10.1038/nbt.1716
14. Ducani, C., Kaul, C., Moche, M., Shih, W. M. & Högberg, B. Enzymatic production of 'monoclonal stoichiometric' single-stranded DNA oligonucleotides. *Nat. Methods* **10,** 647–652 (2013).
15. Praetorius, F. *et al.* Biotechnological mass production of DNA origami. *Nature* **552,** 84–87 (2017).
16. Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99,** 237–247 (1982).
17. Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **305,** 829–831 (1983).
18. Wang, X. & Seeman, N. C. Assembly and characterization of 8-arm and 12-arm DNA branched junctions. *J. Am. Chem. Soc.* **129,** 8169–76 (2007).
19. Chen, J. H. & Seeman, N. C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350,** 631–3 (1991).

20. Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394,** 539–544 (1998).

21. Yin, P. *et al.* Programming DNA tube circumferences. *Science* **321,** 824–6 (2008).

22. He, Y. *et al.* Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452,** 198 (2008).

23. Zheng, J. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461,** 74–77 (2009).

24. Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485,** 623–626 (2012).

25. Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338,** 1177–83 (2012).

26. Ong, L. L. *et al.* Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552,** 72–77 (2017).

27. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).

28. Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459,** 73–6 (2009).

29. Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332,** 342–6 (2011).

30. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–8 (2009).

31. Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res.* **37,** 5001–5006 (2009).

32. Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components. *Science* **347,** 1446–1452 (2015).

33. Marras, A. E., Zhou, L., Su, H.-J. & Castro, C. E. Programmable motion of DNA origami mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 713–8 (2015).

34. Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Toward Larger DNA Origami. *Nano Lett.* **14,** 5740–5747 (2014).

35. Nickels, P. C. *et al.* DNA Origami Structures Directly Assembled from Intact Bacteriophages. *Small* **10,** 1765–1769 (2014).

36. Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline Two-Dimensional DNA-Origami Arrays. *Angew. Chemie Int. Ed.* **50,** 264–267 (2011).

37. Ke, Y. *et al.* DNA brick crystals with prescribed depths. *Nat. Chem.* **6,** 994–1002 (2014).

38. Iinuma, R. *et al.* Polyhedra Self-Assembled from DNA Tripods and Characterized with 3D DNA-PAINT. *Science* **344,** 65–69 (2014).

39. Wagenbauer, K. F., Sigl, C. & Dietz, H. Gigadalton-scale shape-programmable DNA

assemblies. *Nature* **552,** 78–83 (2017).

40. Singh, Y., Murat, P. & Defrancq, E. Recent developments in oligonucleotide conjugation. *Chem. Soc. Rev.* **39,** 2054 (2010).

41. Lin, C. *et al.* Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nat. Chem.* **4,** 832–9 (2012).

42. Langecker, M. *et al.* Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. *Science* **338,** 932–936 (2012).

43. Kuzyk, A. *et al.* DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* **483,** 311–314 (2012).

44. Gállego, I. *et al.* DNA-Origami-Driven Lithography for Patterning on Gold Surfaces with Sub-10 nm Resolution. *Adv. Mater.* **29,** 1603233 (2017).

45. Derr, N. D. *et al.* Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold. *Science* **338,** 662–5 (2012).

46. Shaw, A. *et al.* Spatial control of membrane receptor function using ligand nanocalipers. *Nat. Methods* **11,** 841–816 (2014).

47. Jiang, Q. *et al.* DNA origami as a carrier for circumvention of drug resistance. *J. Am. Chem. Soc.* **134,** 13396–403 (2012).

48. Zhao, Y. *et al.* DNA origami delivery system for cancer therapy with tunable release properties. *ACS Nano* **6,** 8684–8691 (2012).

49. Braun, E., Eichen, Y., Sivan, U. & Ben-Yoseph, G. DNA-templated assembly and electrode attachment of a conducting silver wire. *Nature* **391,** 775–8 (1998).

50. Schreiber, R. *et al.* DNA origami-templated growth of arbitrarily shaped metal nanoparticles. *Small* **7,** 1795–9 (2011).

51. Uprety, B., Gates, E. P., Geng, Y., Woolley, A. T. & Harb, J. N. Site-Specific Metallization of Multiple Metals on a Single DNA Origami Template. *Langmuir* **30,** 1134–1141 (2014).

52. Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335,** 831–4 (2012).

53. Schüller, V. J. *et al.* Cellular immunostimulation by CpG-sequence-coated DNA origami structures. *ACS Nano* **5,** 9696–9702 (2011).

54. Li, S. *et al.* A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. *Nat. Biotechnol.* **36,** 258–264 (2018).

55. Kramer, B. & Tisdall, F. F. The distribution of sodium, potassium, calcium, and magnesium between the corpuscles and serum of the human blood. *J. Biol. Chem.* **53,** 241–252 (1922).

56. Ponnuswamy, N. *et al.* Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nat. Commun.* **8,** 1–9 (2017).

57. Perrault, S. D. & Shih, W. M. Virus-inspired membrane encapsulation of DNA nanostructures to achieve in vivo stability. *ACS Nano* **8,** 5132–40 (2014).

58. Kilchherr, F. *et al.* Single-molecule dissection of stacking forces in DNA. *Science* **353,** (2016).

59. Nickels, P. C. *et al.* Molecular force spectroscopy with a DNA origami – based nanoscopic force clamp. *Science* **354,** 305–307 (2016).

60. Funke, J. J. *et al.* Uncovering the forces between nucleosomes using DNA origami. *Sci. Adv.* **2,** e1600974–e1600974 (2016).

61. Bell, N. A. W. *et al.* DNA origami nanopores. *Nano Lett.* **12,** 512–517 (2012).

62. Martin, T. G. *et al.* Design of a molecular support for cryo-EM structure determination. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E7456–E7463 (2016).

63. Dans, P. D., Walther, J. & Gómez, H. Multiscale simulation of DNA. *Curr. Opin. Struct. Biol.* **37,** 29–45 (2016).

64. Maffeo, C., Yoo, J. & Aksimentiev, A. De novo reconstruction of DNA origami structures through atomistic molecular dynamics simulation. *Nucleic Acids Res.* **44,** 3013–3019 (2016).

65. Göpfrich, K. *et al.* Large-Conductance Transmembrane Porin Made from DNA Origami. *ACS Nano* **10,** 8207–8214 (2016).

66. Snodin, B. E. K. *et al.* Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *J. Chem. Phys.* **142,** 234901 (2015).

67. Rovigatti, L., Šulc, P., Reguly, I. Z. & Romano, F. A comparison between parallelization approaches in molecular dynamics simulations on GPUs. *J. Comput. Chem.* **36,** 1–8 (2015).

68. Doye, J. P. K. *et al.* Coarse-graining DNA for simulations of DNA nanotechnology. *Phys. Chem. Chem. Phys.* **15,** 20395 (2013).

69. Snodin, B. E. K. *et al.* Direct Simulation of the Self-Assembly of a Small DNA Origami. *ACS Nano* **10,** 1724–1737 (2016).

70. Kim, D.-N., Kilchherr, F., Dietz, H. & Bathe, M. Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic Acids Res.* **40,** 2862–2868 (2012).

71. Pan, K. *et al.* Lattice-free prediction of three-dimensional structure of programmed DNA assemblies. *Nat. Commun.* **5,** 5578 (2014).

72. Ke, Y., Bellot, G., Voigt, N. V., Fradkov, E. & Shih, W. M. Two design strategies for enhancement of multilayer–DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chem. Sci.* **3,** 2587 (2012).

73. Amunts, A. *et al.* Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343,** 1485–1489 (2014).

74. Bai, X.-C., Martin, T. G., Scheres, S. H. W. & Dietz, H. Cryo-EM structure of a 3D DNA-origami object. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 20012–7 (2012).

75. Lamour, G., Kirkegaard, J. B., Li, H., Knowles, T. P. J. & Gsponer, J. Easyworm: An open-source software tool to determine the mechanical properties of worm-like chains. *Source Code Biol. Med.* **9,** 1–6 (2014).

76. Maune, H. T. *et al.* Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. *Nat. Nanotechnol.* **5,** 61–6 (2010).

77. Gopinath, A., Miyazono, E., Faraon, A. & Rothemund, P. W. K. Engineering and mapping nanocavity emission via precision placement of DNA origami. *Nature* **535,** 401–405 (2016).

78. Murchie, A. I. *et al.* Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules. *Nature* **341,** 763–6 (1989).

79. Mao, C., Sun, W. & Seeman, N. C. Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy. *J. Am. Chem. Soc.* **121,** 5437–5443 (1999).