Design Optimization of Encapsulating 3D DNA Nanostructures with Curvature and Multi-layers

by

Daniel Fu

Department of Computer Science
Duke University

Date: _____
Approved:

_____
John H. Reif, Supervisor

_____
Thomas H. LaBean

_____
Xiaobai Sun

_____
Alberto Bartesaghi

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Computer Science
in the Graduate School of Duke University
2022

Abstract

Design Optimization of Encapsulating 3D DNA Nanostructures with Curvature and Multi-layers

by

Daniel Fu

Department of Computer Science
Duke University

Date: _____
Approved:

_____
John H. Reif, Supervisor

_____
Thomas H. LaBean

_____
Xiaobai Sun

_____
Alberto Bartesaghi

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science in the Graduate School of Duke University
2022

Copyright © 2022 by Daniel Fu
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

DNA origami has been a paradigm-shifting technique for synthesizing and manipulating matter with nanoscale precision. It follows a simple design principle of using numerous short (<100 nts) oligonucleotides, termed *staples*, to "fold" a long (>1000 nts) DNA strand, termed the *scaffold*. The technique achieved simplicity in design and significantly increased yields compared to previous motifs for DNA nanostructure design. Since then, various approaches have rapidly improved the scale, complexity, and accessibility of DNA nanostructures, such as computer-aided design interfaces, refined synthesis protocols, and further developed design principles for novel geometries.

This thesis formally identifies a geometry of curved, enclosed DNA origami nanostructures termed *DNA capsules*, discusses its unique traits, and outlines how those lead to a departure from conventional DNA origami design principles. While it is a class of structures with high biotechnological relevance, design challenges from this departure have limited accessibility and enthusiasm for utilizing them. This thesis begins by reviewing applications of enclosed DNA origami nanostructures without curvature, also known as an encompassing category of structures termed *DNA nanocompartments*, as a motivation for improving their shape complexity with curving geometry.

Next, this thesis reports on developing an approach towards formalizing and consolidating local and global properties in DNA origami design that can be quantified

and characterized for their effects on DNA nanostructure yield and stability. It then evaluates a heuristic optimization algorithm based on simulated annealing to optimize those consolidated properties to rapidly discover the appropriate oligonucleotide sequences that synthesize the intended DNA origami shape. The broad impact of this work includes the prospect of generalizing DNA origami design.

This thesis then presents an in-depth discussion of the design challenges of DNA capsules and applications of the heuristic algorithms mentioned above to resolve them. The advantages of the developed heuristic algorithms are discussed in the context of novel DNA capsule shapes and multi-layer augmentations to previous work where only single-layer designs were reported. Novel structure designs are then characterized using molecular dynamics simulations and empirically using electron microscopy to evaluate their stability and yields.

Finally, a distributed computing architecture using DNA hairpin cascades with applications to heterogeneous tumor tissue targeting, which is tangential work to the above, is described. Designs are based on previously developed localized DNA hairpin cascades. This thesis presents the design of a signal complex that can transmit information between separate clusters of hairpin cascades. In addition, it discusses several results demonstrating its correctness and scalability obtained using a domain-based DNA strand displacement software called Peppercorn.

Dedicated to my lovely wife and our sweet daughter.

Contents

Abstract **iv**

List of Tables **x**

List of Figures **xi**

List of Abbreviations and Symbols **xiv**

Acknowledgements **xv**

1 Introduction **1**

1.1 Motivation 3

1.2 Overview 5

1.3 Research Questions 6

1.4 Organization 7

2 Review of DNA Origami Designs and Applications **8**

2.1 History of DNA nanostructures 8

2.1.1 DNA Tiling Lattices 8

2.1.2 DNA origami 10

2.2 Classes of 3D DNA Nanostructures 10

2.2.1 Packed Lattice DNA Origami 10

2.2.2 Curved DNA Origami 11

2.2.3 Wireframe DNA Origami 12

2.2.4 DNA Bricks 13

2.2.5 Scalability 14
2.2.6 The Role of Molecular Dynamics Simulations in Design 15
2.3 Applications of enclosed, curved DNA nanostructures 16
2.3.1 Encapsulation 17
2.3.2 Geometric Templating 21
2.4 Conclusion 29
3 Generalized DNA Origami Design and Automations 30
3.1 Introduction 30
3.2 Fundamental Concepts and Terminology 32
3.3 Common Design Principles 35
3.3.1 DNA origami on 2D and 3D packed lattices 35
3.3.2 Polyhedral 36
3.3.3 Curved and rounded 37
3.3.4 DNA Bricks 38
3.3.5 Optimization 38
3.4 Walkthrough of DNA origami design 40
3.5 Consolidation of Parameters 43
3.6 Automation 45
3.6.1 A digital workspace for automation 47
3.6.2 Greedy decision-making 49
3.6.3 Simulated annealing 50
3.6.4 Evaluating the simulated annealing strategy on simple DNA origami nanostructures 55
3.7 Conclusions 58
4 Multi-layer Enclosed and Curved DNA Nanostructures 59
4.1 Introduction 59

4.2 Results 61

4.2.1 Design principles of curved, enclosed DNA origami nanostructures 61

4.2.2 Overview of DNAxiS software and challenges to curved DNA nanostructure design 64

4.2.3 Materials and Methods 74

4.2.4 Experimental evaluation of curved, enclosed DNA nanostructures with crossover networks generated by heuristic algorithms 78

4.2.5 Evaluating yield and feature formation in multi-layer designs . 80

4.2.6 Design and characterization of asymmetric structures 85

4.2.7 Rapid prototyping of DNA capsules 89

4.3 Conclusion 91

5 DNA Interfaces to Cell Membranes 93

5.1 Introduction 93

5.2 Background 94

5.2.1 Localized DNA hairpin circuits 94

5.2.2 Multiple cell types 96

5.3 Design 97

5.4 Simulation 99

5.5 Scalability 101

5.6 Conclusion 104

6 Conclusions 105

Bibliography 107

Biography 116

List of Tables

4.1 Minimum and maximum circumferences per crossover factor at upper and lower limits of spacing and twist. 62

4.2 Expected circumferences per crossover factor at nominal twist and spacings between 2 to 5 full turns. 63

5.1 High affinity aptamers per cell line. 96

List of Figures

2.1 Applications of encapsulating DNA nanostructures. 17

2.2 Applications of utilizing DNA nanostructures as a geometric template. 22

3.1 Conceptual conversion of a 2D DNA origami sheet to a 3D DNA capsule. 37

3.2 Manual DNA origami design, Step 1. 41

3.3 Manual DNA origami design, Step 2. 41

3.4 Manual DNA origami design, Step 3. 42

3.5 Manual DNA origami design, Step 4. 43

3.6 Diagram of spacing parameters. 44

3.7 Preprocessing steps of automating flat DNA origami design. 48

3.8 Modules for basic DNA origami shape primitives. 49

3.9 Flowchart of greedy DNA origami heuristic design algorithm. 51

3.10 Flowchart of simulated annealing DNA origami heuristic design algorithm. 52

3.11 Automated DNA origami design, Step 1. 52

3.12 Automated DNA origami design, Step 2. 53

3.13 Automated DNA origami design, Step 3. 53

3.14 Automated DNA origami design, Step 4. 53

3.15 Automated DNA origami design, Scoring. 54

3.16 Automated DNA origami design, Backtracking. 54

3.17 2D square lattice rectangle comparison. 55

3.18 Overview of heuristic process on hex lattice. 56

3.19 3D hexagonal lattice block comparison. 57

4.1 Preview of DNAxiS's online graphical user interface. 61

4.2 STL parsing for axially symmetric geometries. 65

4.3 Overview of the DNAxiS design process. 66

4.4 Alignment principles for single/multi-layer, curved, closed-shell DNA nanostructures. 67

4.5 Crossover discovery model on curved nanostructures. 69

4.6 Crossover alignment challenges when curvature introduces non-homogeneous circumference and helical twist. 70

4.7 Example trace of crossover network design when applying simulated annealing algorithms. 71

4.8 Runtime benchmarks for sample structure inputs into DNAxiS using the simulated annealing crossover network solver. 72

4.9 Experimental verification of automated design principles. 78

4.10 Evaluation of reinforcement strategies via molecular dynamics simulation. 80

4.11 Evaluation of reinforcement strategies via experimental characterization. 81

4.12 Interface and scaffold pathway routing challenges. 82

4.13 Initial screening for effects of multi-layer curved design. 84

4.14 Axially asymmetric structures, Ellipse. 86

4.15 Axially asymmetric structures, Clover. 87

4.16 Euler-Bernoulli beam model for curving DNA. 88

4.17 RMSF in alternate connection graphs. 90

4.18 Parameterizing generalized DNA origami design. 91

5.1 Summary of reaction steps in distributed localized DNA hairpin gates. 98

5.2 Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2). 99

5.3 Simulated operation of only H_2 (Node 2) and the output reporter. . . 100

5.4 Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2) with intermediates overlaid. 100

5.5 Peppercorn simulation of localized 1-1 circuit. 101

5.6 Node and signal complex design for a localized 2-2 circuit. 102

5.7 Peppercorn simulation of localized 2-2 circuit. 102

5.8 Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2) with intermediates overlaid. 103

5.9 Peppercorn simulation of localized 1-1-1 circuit. 103

List of Abbreviations and Symbols

Symbols

δ	Interhelical distance.
θ	Helical twist.

Abbreviations

DNA	Deoxyribonucleic acid.
ssDNA	Single-stranded DNA. DNA that has not been hybridized to another complementary strand.
dsDNA	Double-stranded DNA. DNA that has been hybridized to another complementary strand.
bpt	Bases per turn.
nt	Nucleotide.
AFM	Atomic force microscopy.
TEM	Transmission electron microscopy.
RMSF	Root-mean-square fluctuation
RMSD	Root-mean-square deviation

Acknowledgements

I am eternally grateful to Prof. John Reif, my advisor, for the opportunity to have pursued the research presented in this thesis and more. I have greatly enjoyed our walks and talks throughout my years at Duke and sincerely appreciate the advice I have gleaned from these in not just research and science but in life and all its experiences and tribulations.

I want to thank those who have served on my committees, Prof. Alvin Lebeck, Prof. Alberto Bartesaghi, Prof. Thom LaBean, and Prof. Xiaobai Sun, for the feedback and direction they have given me in my research.

I would also like to extend my gratitude to the collaborators I have had the pleasure of working with and my colleagues in the Reif lab: Prof. Hao Yan, Prof. Fei Zhang, John Schreck, Raghu Pradeep Naranayan, Abhay Prasad, and Dewight Williams. Hieu Bui, Tianqi Song, Reem Mokhtar, Shalin Shah, Xin Song, Byoungkwon An, Ming Yang, and Rajiv Nagipogu.

Thank you to the Duke Computer Science department, which has always been fully supportive and considerate of its students, and a special thank you to Prof. Xiaobai Sun and Marilyn Butler, for the care and concern they have shown me and for always being friendly, welcoming presences to speak to. I would also like to acknowledge the DNA nanotechnology community at large, whom I have met at conferences throughout the years and who have all been passionate and friendly in discussing and improving our research.

And with all the love in the world, I am grateful to my wife, daughter, and mine and my wife's families for their unconditional support.

The work presented in this dissertation was completed with support from the National Science Foundation under grants no. 1909848, 2113941, 2004250, and 1931487.

1

Introduction

Structural DNA nanotechnology describes the programmable self-assembly of DNA into nanoscale shapes termed *DNA nanostructures*. The field began by establishing that DNA could be programmed into physical conformations rather than used solely as genetic code [2]. This observation was inspired by the Holliday junction, an immobile, four-way, branching nucleic acid structure, being curiously compared to an art piece, *Depth*, depicting fish with horizontal and vertical fins assembled into an ordered, 3D array. The Holliday junction was later adapted into a linear conformation in double crossover (DX) tiles [3]. This motif would become an essential primitive of future DNA nanostructure design principles and is now generally referred to simply as a *crossover*. Progress continued to refine and consolidate design principles of DNA nanostructure design, culminating in the next milestone, the DNA origami technique [4]. Since then, extensions of the DNA origami technique, each a *design strategy* with its specific design principles for creating a particular geometry of shapes, have yielded a broad library of DNA nanostructure shapes, including spheres, boxes, tetrahedrons, various imaginative polyhedral and voxel-based shapes,

This chapter includes excerpts from Fu & Reif (2021) [1].

and many more [5–13]. Applications have shown off DNA nanostructures as one of the most versatile materials for exploring innovative approaches to nanofabrication, drug delivery, nanophotonics, biosensing, and in general, manipulating matter with single-molecule precision [14–16].

Moreover, the bonding properties of DNA are conducive to both structural design and computation, and DNA nanostructures have become progressively various in form and dynamic in function [17,18]. The scale and complexity of DNA nanostructures have grown significantly in recent years, such as nanostructures achieving gigadalton masses [19,20], micrometer-scale crystals [21,22] and tile assemblies [23], improved rigidity in polyhedral DNA nanostructures [24], and recurrent design methodologies such as meta-DNA [9,25]. Some potential applications exploiting those design principles for scale and complexity have also followed [26–31]. Overall, attention has expectedly risen regarding the design of DNA nanostructures to specifically address the function, shape, and stability requirements across a broad scope of potential applications.

However, there remains a high access barrier to utilizing *ad hoc* DNA nanostructures with tailored properties for each potential purpose. Additional design principles to satisfy those properties add complexity to the design space that obfuscates design choices, thus increasing the dependence on expert designers, which can hinder the accessibility of DNA nanostructures across different research disciplines. The shape complexity of DNA nanostructures has also rapidly diversified since the inception of the DNA origami technique.

Yet, the baseline that has marked DNA nanostructure designs which can be reliably tuned for desired downstream properties has hardly budged. There are algorithmic solutions and software automation for some methodologies, such as polygonal DNA nanostructures or DNA bricks [5, 19, 24, 32]. However, these methods are only generally applicable in their supported geometries, and the geometries they support

may only achieve some desirable properties.

This thesis focuses on the group of DNA nanostructures that share the functional trait of *encapsulation*, that is, the ability of the DNA nanostructure to separate, confine, or organize other molecules. We refer to such structures henceforth as *DNA nanocompartments*. With greater specificity, we refer to the subset of DNA nanocompartments that exhibit curvature as *DNA capsules*.

1.1 Motivation

Capsules, in general, are commonly found throughout biology, with supramolecular assemblies like viral capsids and cell membranes omnipresent for the organization, transport, and protection of biomolecules through specifically tailored shape, surface properties, and composition [33]. It is of great interest to mimic the properties of these structures to exploit their usefulness to nanoscience, but artificially replicating their meticulous biochemistry is a grandly difficult task [34, 35]. Structural DNA nanotechnology has proven to be a highly versatile substrate for the arbitrary design of 3D nanostructures. Its diverse structural design strategies could produce highly complex synthetic nanoscale capsules (*nano-capsules*) for applications in nanoscience that are yet unexplored.

DNA nanotechnology has already demonstrated high potential in biomimetic roles such as membrane scaffolding [26,36–38], vesicle fusion [39,40], and membrane channels [41–44]. DNA nanocompartments have also been used in constructing nanoreactors and controlling small-volume chemistry [45–48], templating the seed-mediated synthesis of metallic nanoparticles [31, 49, 50], templating the shape of lipid membranes [26, 38, 51, 52], and as drug delivery vehicles [53, 54].

Many of these applications are shape-dependent, and progress in creating complex 3D structures from the DNA origami technique has established several strategies for shape design, such as *packed lattice* structures [8,13,55–57], *DNA bricks* [6,19], *poly-*

hedral DNA nanostructures [5, 11, 24, 32], and *curved DNA origami* [7, 10, 58]. These design strategies have demonstrated high shape variety with some structures capable of encapsulation. Still, as previously noted, the shifting attention to downstream applications emphasizes additional properties such as shape, rigidity, permeability, actuation, and yield. Fewer design strategies can meet this expanded criterion.

An observation of applied instances of DNA nanocompartments shows that many current encapsulation strategies still rely on packed lattice designs, where cavities are wrapped [47, 53, 54, 59] or hollowed out from a block of material [45, 48]. While these structures achieve high stability and rigidity, their shape complexity is limited. Packed lattice structures can also exhibit large material overheads to construct any geometries that are not parallel with the helical axis orientation since the length of any helix used in the nanostructure is lower bounded by the length that would be necessary to select a minimum of 2-3 crossovers to adhere it to its neighbors. This is apparent in the side walls of the nanocompartment made in previous work by Grossi *et al.* [45]. On the other hand, helical axes in curved structures, such as those demonstrated by Han *et al.* [10], always follow the geometry. Alternatively, DNA bricks have greater shape versatility, but the high usage of short strands sacrifices rigidity [60]. Recent methodologies have also improved the rigidity of polyhedral designs by expanding edges to six-helix bundles [24]. Still, polyhedral faces form gaps that could make the containment of diffuse guest molecules intrinsically challenging.

Recent work has alluded to a transition in applied DNA nanocompartments from packed lattice designs toward curved geometric features. Kohman *et al.* [61] described the actuation of a spherical DNA nanostructure compared to the earlier demonstration by Andersen *et al.* [13] of a DNA origami box with a controllable lid. Sun *et al.* described rounded metallic nanoparticles synthesized via the nanocasting method and evaluated the photonic behavior of size changes in synthesized nanoparticles. Round DNA nanostructures like those described by Perrault and Shih [37], Yang *et*

al. [38], and others [26, 51, 52] show improved coordination of DNA nanostructures with lipid membranes due to their naturally rounded conformations. The nanoreactor built within a DNA origami barrel described by Hahn *et al.* [62] maximizes the usable cavity space by using a curved nanostructure.

Furthermore, a rigidity study performed via simulations accompanying the work by Sun *et al.* [50] and a permeability study conducted by Grossi *et al.* [45] quantifying the activity of the nanoreactor in response to different size invasive reagents supports the importance of a multi-layer DNA capsule. Simulation studies indicated that due to effects such as electrostatic repulsion, strain, and Brownian motion, the structure of DNA origami is not uniform, and larger than intended gaps can intermittently form throughout the structure in the spacing between helices and crossovers [63]. Studies by Ke *et al.* [64] and Martin *et al.* [65] have also suggested significant improvements in rigidity and stability for multi-layer DNA nanostructures.

1.2 Overview

This thesis will first review the applications made possible with the encapsulation property of DNA nanocompartments. The discussion of these structures also centers around the dissection of the goals of encapsulating nanostructures into the following properties:

1. *Shape* - The shape of the inner cavity and the observed outside form of the nanocompartment are considered.

2. *Rigidity* - The ability of the nanocompartment to maintain a correct shape and its resistance to deformation and degradation.

3. *Permeability* - The segregation of nanoparticles or reactants between the inside and outside spaces of the nanocompartment.

4. *Actuation* - The programmable controlled release of guest molecules.

5. *Yield* - The synthesis efficiency of the DNA nanostructure.

The final goal of this thesis is to describe design principles for designing multi-layer DNA capsules, and we examine the challenges that have since hindered that transition. We first focus on evaluating common trends between existing DNA nanostructure design methodologies to establish a generalized, widely applicable DNA origami design methodology by applying heuristic optimization strategies to solving the placement of crossovers, which to the best of our knowledge, have not been previously explored for DNA origami design. This strategy is then extended to the scope of curved DNA nanostructures. Many DNA origami designs rely on a consistent helical twist to provide regularity, thereby reducing design complexity. However, curved DNA nanostructures are made possible by breaking this consistency, thus losing valuable simplifications in design complexity. It becomes excessively challenging for a human designer and therefore calls for automated strategies.

Lastly, this thesis explores tangential work on developing a distributed computing architecture using DNA hairpin cascades with applications for heterogeneous tumor tissue targeting.

1.3 Research Questions

This dissertation aims to challenge conventional design principles in DNA origami design and promote the adoption of alternative approaches, focusing on raising limitations on shape and size in DNA nanocompartments. First, this work evaluates whether features that govern the design and stability of DNA origami nanostructures can be identified and quantified at both local and global scopes. Second, it explores the implementation of two heuristic optimization algorithms and investigates their success in automatically generating DNA nanostructures accommodating the consol-

idated parameters. Finally, these optimization techniques are applied to a formally identified class of curved DNA nanocompartments called DNA capsules to demonstrate generated structures experimentally. This thesis surmises that future design requirements of highly complex, supramolecular DNA nanostructures assemblies may not be met by conventional design principles and require alternative approaches, such as those presented in this thesis.

1.4 Organization

The remainder of this thesis is organized as follows. Each section is opened with a brief introduction that covers relevant previous works and definitions.

- Chapter 2, Review of DNA Origami Designs and Applications of DNA Nanocompartments.
- Chapter 3, Development of a Heuristic Algorithm for Automated and Generalized DNA Origami Design.
- Chapter 4, Application of Heuristic Algorithms and Characterization of Automatically Generated Multi-layer Enclosed and Curved DNA Nanostructures.
- Chapter 5, Distributed Hairpin Cascade for Heterogeneous Tumor Tissue Labeling.
- Chapter 6, Conclusions.

2

Review of DNA Origami Designs and Applications

This chapter reviews past milestones in DNA nanostructure design techniques and applications with a focus on the properties and advantages of DNA nanocompartments. This chapter serves as a primer to the state of the art prior to the work in this thesis and helps to highlight the significance of the reported contributions. The information provided in this chapter should elucidate the rapid progress in DNA nanostructures and the limitations that a shift toward generalized DNA origami design and greater support for curved geometries could resolve.

2.1 History of DNA nanostructures

2.1.1 DNA Tiling Lattices

Nadrian Seeman first published the seminal observation that DNA could serve a structural purpose in 1982, where it was envisioned that repeating units of branching DNA junctions, based on a four-way, branching nucleic acid structure called a Holliday junction, could hybridize to form a stiff 3-dimensional lattice to facilitate the crystallographic study of other molecules [2]. Seeman demonstrated the first 3D

This chapter includes excerpts from Fu & Reif (2021) [1].

DNA cube using eight branching junctions, one at each vertex [66]. The concept was a revelation that the highly specific, sequence-dependent bonding of DNA had not only a genetic purpose but could be structurally engineered too. The concept of using branching junctions as a structural motif would form a foundation for the emergence of the field of structural DNA nanotechnology.

The first leap in complexity following Seeman's discovery was *DNA tiles* [3], which was the earliest multi-purpose design technique for structural DNA design, capable of generating different geometries using programmable changes within the same fundamental design principles. More generally, a DNA tile is a DNA nanostructure that is self-assembled from a small number of single-stranded DNA (ssDNA) that hybridize together, forming stabilizing Holliday junctions known as *crossovers*. Once assembled, some exposed ssDNA (termed *sticky ends*) allows the DNA tile to further self-assemble into DNA nanostructures known as *DNA tiling lattices*, which have been used to build 1D and 2D DNA tiling assemblies and lattices as physical representations for molecular-scale computation. The letter sequences of the sticky ends determine the binding rules of each tile. The self-assembly of patterns by DNA tiles has proved to be quite complex, as shown in Sierpinski triangle assemblies by Rothemund, Papadakis, and Winfree [67], which computed `mod` 2 sums within the assembly, forming a fractal-like 2D pattern.

However, the number of possible geometries under only the design principles so far was very limited. DNA tiling assemblies remained restricted to simple algorithmic patterns. Yields and the size scales of early DNA nanostructures self-assembled by DNA tiles were low. Similarly, *DNA crystals*, a concept that persevered from the inspiration of the field as initially a method to construct crystalline lattices using DNA, carried limitations as well. Arrays of tensegrity triangle-shaped junctions have yielded micron-scale DNA crystals [21,68,69], which are the largest DNA nanostructures to be synthesized to date, but the overall shape of crystals is very restricted,

such as being limited to rhombohedral or cuboidal shapes.

2.1.2 DNA origami

The shift into the current paradigm of DNA nanostructure design occurred when Paul Rothemund developed a technique now known as *DNA origami* [4]. A long piece of ssDNA, known as the *scaffold* strand, winds through a raster of the shape while hundreds of short ssDNA, known as *staple* strands, span and hybridize to spatially adjacent, non-consecutive regions of the scaffold strand according to Watson-Crick base pairing rules, thus tying them spatially together. Staple strands bind to form duplexes, which stiffen the structure to form the shape, creating a series of coplanar, anti-parallel helices bound by crossovers. This concept was a significant breakthrough in structural DNA nanotechnology and significantly improved the size, yield, and shape complexity of DNA nanostructures. Although initially demonstrated for designing only 2D shapes, the DNA origami technique has since been extensively developed as the underlying strategy for many subsequent derived techniques for designing increasingly complex 3D shapes.

2.2 Classes of 3D DNA Nanostructures

2.2.1 Packed Lattice DNA Origami

Following the advent of DNA origami, the Shih group made the first foray toward a multi-purpose design strategy for 3D DNA origami [8]. This strategy can be thought of as folding a flat 2D DNA origami at creases between adjacent helices to create a stacked structure of multiple planes of 2D DNA origami and adding crossovers that span both in-plane and between planes to hold the structure together. All helices are parallel, and their perpendicular cross-section cutting across all helices is known as the *lattice*, but not to be confused with DNA tiling lattices. The lattice indicates predetermined positions of the helices according to a simple pattern. Structures de-

signed on these predetermined positions are *on-lattice*, whereas those that do not have limitations on helix placement are *off-lattice*. The Shih and Yan groups introduced two presets of lattices, the honeycomb lattice [8] and the square [56] lattice, and also later demonstrated a hybrid of both (square and honeycomb) and triangular lattices [64]. Around this time, unique designs showcasing a growing proficiency for designing DNA origami nanostructures were also demonstrated. For example, using flat 2D DNA origami as the faces of simple polyhedral shapes such as cubes [13, 70] or tetrahedrons [57]. There were also unique artistic structures such as a Möbius strip [71] and tensegrity structures [72]. As with the other techniques discussed here, the increasing scale of 3D structures also solicited the creation of CAD software to assist in their design, and the software package caDNAno [73] has been the longstanding, dominant option for on-lattice designs.

2.2.2 Curved DNA Origami

Geometric complexity continued to improve as the Shih group demonstrated smooth global curvature and twist, as opposed to hitherto straight geometries, along the axis of a *helix bundle*, which is a rod-like structure of anti-parallel helices arranged in a square or honeycomb lattice and has become an archetypal DNA nanostructure. *i-helix bundles* refer to the number of helices comprising the bundle. For instance, 4-helix bundles and 6-helix bundles, which are the smallest instances of bundles on square and honeycomb lattices, respectively. Curvature and twist were applied via a series of *insertions* and *deletions* [7], which are modifications in the crossover placement of a few base pairs nearer or farther than a placement that would correspond most closely to the native winding of B-form DNA (10.5 bps/turn). In doing so, a helix became locally overwound or underwound, creating strain. Globally, the relaxation of the structure to satisfy the buildup of strain translated to the twist or curvature of the bundle depending on the chosen distribution of insertions and

deletions.

This technique inspired a technique by the Yan group for making curved, hollow, enclosing DNA origami nanostructures. By connecting the endpoints of a duplex, it became a cyclic structure. In combination with a duplex's persistence length, the distribution of strain along the duplex formed a circle, and this shape was reinforced by its similarly shaped adjacent helices. Modulating the lengths of each duplex of the flat 2D sheet corresponded to changing the circumference. Then, by utilizing the full availability of dihedral angles, the surface topology of the nanostructure could be designed to resemble shapes such as spheres, ellipsoids, toroids, or flasks [10]. The same group later extended this concept using gridiron designs, which routed the scaffold and placed staple strands so that helical axes could run perpendicular to each other to reinforce shape and curvature [74]. Thus far, the Tiamat software [75] has been the leading option for designing such structures in an open 3D space.

2.2.3 Wireframe DNA Origami

Polyhedral structures have established the most extensive variety of shapes, especially with subsequent stability improvements and software developed to automate the scaffold and staple routing processes [5, 24, 32, 76]. Shape complexity leaped forward as the Bathe, Yan, and Högberg groups each embarked on unique demonstrations of DNA nanostructures capable of rendering a polyhedral wireframe mesh [5, 12, 32]. The crux of this problem was how to route the scaffold so that every edge of the polyhedron was included, eventually concluding that any such scaffold routing must follow an Eulerian circuit solution and manipulate the number of permissible repeated edge traversals to satisfy such a solution. These works also evaluated how to adjust edge lengths to correspond to the correct full- or half-turn crossover positions in B-form DNA for stable crossover placement.

Several software suites have been designed, each to satisfy an increase in the

structural complexity of polyhedral DNA origami nanostructures. Software packages DAEDALUS [5] and vHelix [32] were among the earliest. PERDIX [76] later addressed the complexity of 2D wireframe designs, and TALOS [24] for polyhedral wireframe structures with thicker helix bundle edges which required more complex scaffold routing. The Yin group also demonstrated an alternative method for polyhedral shape formation. While it lacked algorithmic foundations [11], it had the advantage of scalability due to the modular nature of the design strategy, meaning its entire structure was no longer limited to the maximum size of a DNA origami nanostructure that uses only one scaffold.

2.2.4 DNA Bricks

While the DNA origami technique has been a foundation in most strategies, the DNA bricks technique pioneered by the Yin group has been otherwise successful without it. DNA bricks construct structures without any scaffold strand, instead opting only to use short ssDNA referred to as *bricks*. Each brick is a pair of same-length anti-parallel helices joined by a phosphate linkage. Each helix is split into two equal-length domains, such that the brick has a total of four domains, and each domain is a consistent length regarded as the *motif*. Motifs are either 8 or 13 nts, and the choice of motif affects the scale of designed structures due to the number of unique sequences possible for a specific length of DNA, as well as yields and melting temperature. In a simple one-pot synthesis, the numerous strands self-assemble, with each strand representing a small volume of the structure like bricks filling in a shape [6]. This technique has been shown to scale up to gigadalton scales with high shape complexity, which was also accompanied by the release of its personalized software tool, Nanobricks [19].

2.2.5 Scalability

In general, there has been a trend towards scaling up the production and size of 3D DNA nanostructures across all techniques. One strategy has been to pursue improvements to scaffold strand size and customization. Doing so can suppress cost, relieve design limitations, improve the practicality of 3D DNA nanostructures across broader size scales, and ensure immunocompatibility of DNA nanostructures for *in vivo* applications [77]. To create larger, more diverse DNA origami as single, one-pot syntheses, either longer scaffolds or many orthogonal scaffolds are required, and avenues such as re-engineering bacteria and enzymatic toolkits [78–80] to those ends have been investigated.

On the other hand, rather than facing the challenges associated with longer scaffold strands, other novel techniques rely on hierarchical assembly. The Dietz group applied self-assembly techniques upon an assembled DNA origami component to demonstrate gigadalton-scale polyhedral DNA nanostructures [20]. Each component was a V-shaped structure (*V-brick*) that could join to other components by attachment sites on their flat sides. A tripod structure of three V-bricks together translated to the angle of the vertex it would represent in a 3D polyhedral shape, such that polyhedral structures would self-assemble towards an energetically favorable configuration matching the number of vertices necessary to form a closed shape. The Yan and Reif groups have proposed using *Meta-DNA* [9, 25] as another route towards scaling up the size of DNA nanostructures. Meta-DNA are 3D DNA origami nanostructures constructed to resemble both the shape and Watson-Crick base pairing rules of the standard nucleotides of DNA. They also hierarchically assemble into double helix structures resembling DNA which can then be used to emulate DNA origami designs at a magnified scale.

In comparison to depending on specific hybridization reactions to assemble multi-

component structures, which can be restrictive and difficult to reverse, alternative methods have also been demonstrated to have the propensity to assemble large, complex 3D DNA nanostructures, such as by using base stacking or shape-complementary binding [81]. As already mentioned, extended methods of other standard techniques, such as for wireframe DNA origami [24,76] or DNA bricks [19], have also been developed to create larger and more mechanically rigid designs. Assemblies utilizing crystallographic regularity of DNA origami nanostructures can also provide large templates for assembling guest molecules into crystal lattice-like arrangements [22,28].

2.2.6 The Role of Molecular Dynamics Simulations in Design

Molecular dynamics simulations have been a significant development in the computational aspects of DNA nanostructure design, as they provide faster, cheaper feedback and insight into the behavior of DNA nanostructures. Coarse-grained molecular dynamics modeling tools such as CanDo provided by the Bathe group [82], oxDNA provided by the Doye group [83], and NAMD provided by the Aksimentiev group [84] have greatly alleviated challenges in designing and testing DNA nanostructures. It has since become possible to simulate large DNA origami [63, 84–86] with molecular granularity. These tools can simulate the diffusive movement of strands, their hybridization, and their physical properties that may rely on bond strength, system temperature, or ionic concentrations. Thus, they can provide information on the shape and stiffness of designed DNA nanostructures in an expected environment. Simulated studies have revealed properties of DNA nanostructures similarly discovered via much more challenging characterization by imaging techniques such as cryo-EM [87, 88].

2.3 Applications of enclosed, curved DNA nanostructures

Within a decade of research, the shape variety of DNA origami nanostructures has become increasingly saturated, but the plethora of geometries has outpaced their presence in nanoscience applications. A large portion of applications still utilizes DNA origami shapes designed as packed lattices. The design criteria of more DNA origami-derived strategies must evolve to accommodate properties previously secondary to shape, such as mechanical or biomimetic properties that would enhance lifetime or effectiveness *in vivo* or functionality in conjunction with other nanomaterials *in situ*. In addition, as a soft material, DNA nanostructures are susceptible to degradation and deformation that may further limit their usage in nanoscience applications, emphasizing a need for generalized structural reinforcement strategies to improve the robustness of structures for reliable use.

On the other hand, nano-capsules, such as viral capsids or cell membranes, already exhibit mechanical and biocompatible properties that are highly beneficial to nanoscience. DNA origami has so far had limited success emulating this archetype. Natural structures are difficult to artificially create and modify to suit a broader range of purposes. DNA origami may yet provide a solution for designing artificial nano-capsules that can be specifically tailored to their applications. This chapter investigates the mechanical shortcomings of existing DNA origami methodologies to implement properties similar to natural nano-capsules for applications where reliable confinement and manipulation of guest molecules is of utmost importance. A wide field of downstream applications, including nanocasting, cellular uptake, drug delivery, bioreactors, and nanofluidics, will be discussed in terms of how they contribute to design criteria modifications to develop application-centric design methodologies for DNA origami.

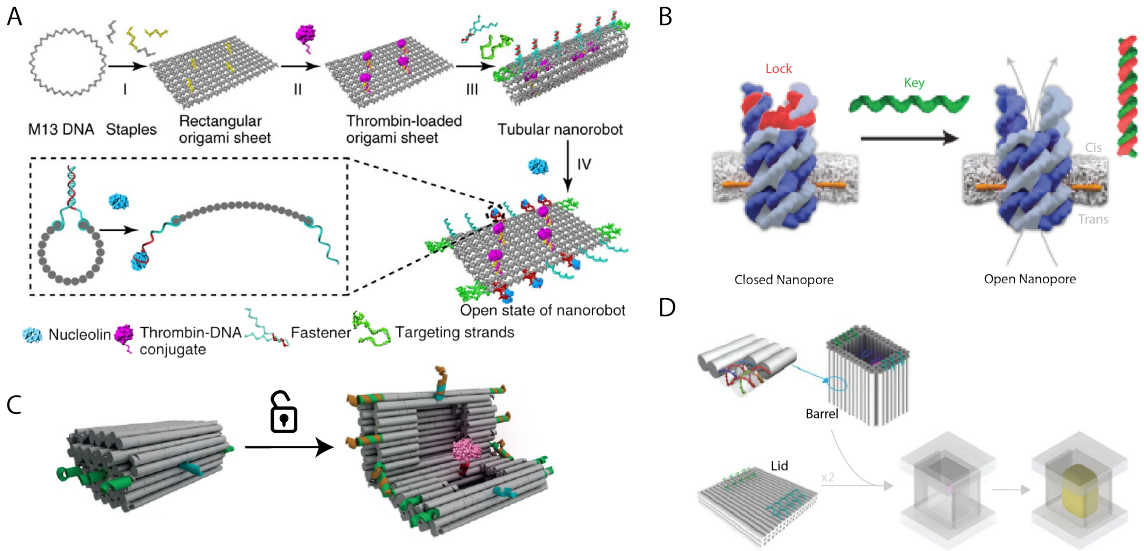

FIGURE 2.1: Applications of encapsulating DNA nanostructures. **(A)** A DNA nanorobot that selectively delivers thrombin to a tumor site to induce thrombosis of the tumor to destroy it. (Reprinted with permission from Springer Nature [54].) **(B)** A DNA nanopore with a locking mechanism that can restrict a flow of molecules through the channel. (Reprinted with permission from American Chemical Society [42].) **(C)** A DNA origami nanovault functions as a nanoreactor by capturing a single enzyme within its cavity and controlling its reactions via a latch mechanism. (Reprinted with permission from Springer Nature [45].) **(D)** A DNA origami mold used for nanocasting of gold nanoparticles. (Reprinted with permission from AAAS [50].)

2.3.1 Encapsulation

We refer to *encapsulation* as the concept of separating molecules of interest between an inner compartment from the outside environment. Encapsulation is a goal of shape design, where a structure must be able to form a cavity. Here, the significance of shaped DNA nanostructures is primarily in fulfilling function. Compartmentalization is enforced in nature by membranes or protein shells to consolidate processes and also defend sensitive operations or molecules from interruptions or damage. It is of great scientific interest to create artificial nanostructures mimicking this capability. Technologies such as polymer or lipid vesicle technologies [33–35] fulfill this goal, but in comparison to DNA origami, neither are as programmable and addressable.

On the other hand, DNA nanostructures are not continuous surface structures like a membrane is, and their parallel helices also have some fluctuation [63] that can reveal gaps in the structure. This can be improved using multi-layer DNA origami nanostructures [56,64], which minimize fluctuations and reduce gap spacings to become less permeable to the relevant molecules of the application. While not all examples of encapsulation are shape-dependent, DNA nanocompartments are uniquely capable of modulating the size of a cavity, the position or quantity of contained molecules, and access to the space as determined by DNA-based circuit logic.

Nanocasting

Control over shape and permeability has been vital to synthesizing shaped gold and silver nanoparticles using a hollowed DNA nanostructure as a mold via a process termed *nanocasting* [49, 50] (Fig 2.1D). The nanocasting process utilizes seeded-growth nanoparticle synthesis but with the seed beginning within a shaped cavity to which the growing nanoparticle will conform. This process has been demonstrated with silica nanotubes [89] and protein shells [90] but without the same degree of shape variety as DNA nanostructures. The advantage of using DNA nanostructures to form the mold is the wide variety of shapes, such as cubes, discs, and triangles, and their compositions, such as Y-branch shapes. Expanding the available geometries of other materials is, in isolation, already a significant research problem, while DNA origami techniques are generalized across a variety of geometries. Currently, while still effective in determining the rough shape of the metallic nanoparticle, the flexibility of DNA nanostructures slightly yields to the forces of nanoparticle growth upon its inner walls. Thus so far, the process mandates accurate timing, rather than entirely relying on the walls of the DNA mold, to terminate nanoparticle growth, but methods as yet untested for nanocasting for improving the rigidity of DNA nanos-

tructures could be later shown to improve nanocasting quality [91, 92]. Nonetheless, the results show enticing potential for the bottom-up synthesis of plasmonic nanostructures [50] or nanowires [31].

Single-molecule chemistry

Nanoreactors are confinements of selected molecules such that there is some level of deliberate control over their reactions. Nanoreactors require the property of an impermeable encapsulation to be paramount to tailor environmental conditions for only specific reactions to occur and thus should be able to shield molecules from unfavorable reactants proficiently. The Andersen group demonstrated a DNA origami nanoreactor "vault" that can capture single enzymes within its cavity and could impose specific conditions on its locking mechanism [45] (Fig 2.1C) to restrict reaction conditions further. In doing so, the responsiveness of the enzyme to substrates could be consolidated into specific activity levels of scientific interest.

Alternatively, enzymes can be spatially localized. This has been shown for 2D surfaces to measure the effect of distance separation on reactions and investigate biophysical mechanisms of target proteins [93, 94], but also within the cavity of a DNA origami reactor [47,48,95]. This localization is similar to how natural enzymatic pathways achieve rapid reaction rates. Such a system will allow scientists to compose and study specific pathways under precisely modulated concentrations and relative distances. DNA nanostructures can also aid in isolating single membrane proteins by simulating a small lipid bilayer within a constructed cavity. Such studies are otherwise challenging due to the poor stability and activity of membrane proteins in any environment other than upon the cellular membrane [46], thus calling for strategies that can artificially simulate that environment.

DNA-based drug delivery vehicles

Drug delivery is an area where structural DNA nanotechnology may inspire a new degree of complexity toward therapeutics. Targeted drug delivery is still an open challenge [96] that ideally establishes strategies for the safe and successful delivery of a therapeutic payload by limiting their effective area to only diseased sites. This is especially a problem for drugs that can harm healthy tissue that unintentionally uptakes it but must still be circulated through the body to reach its intended site.

The ability of DNA to both form structures and compute with biomolecular inputs is a unique advantage of the material to satisfy this functionality. Placing such therapeutics within the cavity of a DNA nanocompartment adds a layer of separation between the therapeutic and any other tissue it could unintentionally damage. Then, DNA nanocompartments can be designed to open and expose the cavity with its therapeutic payload when a targeted biomarker is observed by a programmed locking mechanism [53]. The limitation and future opportunities for research here, of course, are thus to expand the variety of biosensors that can be conjugated to DNA and incorporated into DNA nanocompartments.

One area of high potential impact is the targeted delivery of anti-cancer drugs, which can have severe side effects, such as doxorubicin which can cause damage to the heart muscle. DNA circuits have been shown to perform complex profiling of cancer cells, thus heightening the specificity of any DNA nanostructure-based targeting for cancer therapeutics [97]. DNA nanostructure drug delivery vehicles have also been shown to be sufficiently stable under *in vivo* conditions to passively build up at a tumor site [98]. Li *et al.* successfully delivered and released thrombin to tumors to induce thrombosis of only the tumor tissue [54] (Fig 2.1A). Pan *et al.* demonstrated that DNA nanostructures can easily be programmed to carry multiple different drugs simultaneously and could thus help the simultaneous delivery

of multiple-drug therapeutics to combat complex diseases that are drug-resistant [99].

Membrane-spanning DNA Nanopores

DNA nanocompartments have been used to build transmembrane channels, sometimes referred to as *membrane-spanning DNA nanopores*, that implant into a cellular membrane. Tubular DNA nanocompartments decorated with hydrophobic groups along its outside enable a DNA nanocompartment to embed itself into a lipid membrane as an artificial transmembrane channel [41, 44]. We consider this an essential function of encapsulation in isolating a leak-free channel for molecules to be transported across.

Natural transmembrane channels span their lipid bilayer and typically help to regulate the nutrients within a cell. DNA nanopores may enable forced control over the nutrition gradient of cells by implementing commonly used locking mechanisms or could be used for biophysical studies of how materials pass through the membrane by monitoring voltage changes from the DNA nanopore that indicate molecules are passing through. The Howarka group demonstrated that designed DNA nanopores are functionally similar to their associated natural counterpart when measured for transport of K^+ ions [42] (Fig 2.1B). Proteins may also be passed through the channels, and each event can be measured by the transduced electrical signal of the channel [100]. Gating can also be applied, which blocks the channel under specifically programmed conditions to modulate the flow of molecules through the channel [43, 101].

2.3.2 Geometric Templating

Geometric templating is the principle of utilizing the geometry of DNA origami nanostructures to orient or position other molecules that do not have the intrinsic capability to organize to the geometric complexity of the DNA-based template. This principle depends on the shape of the DNA nanostructure itself as designed or as

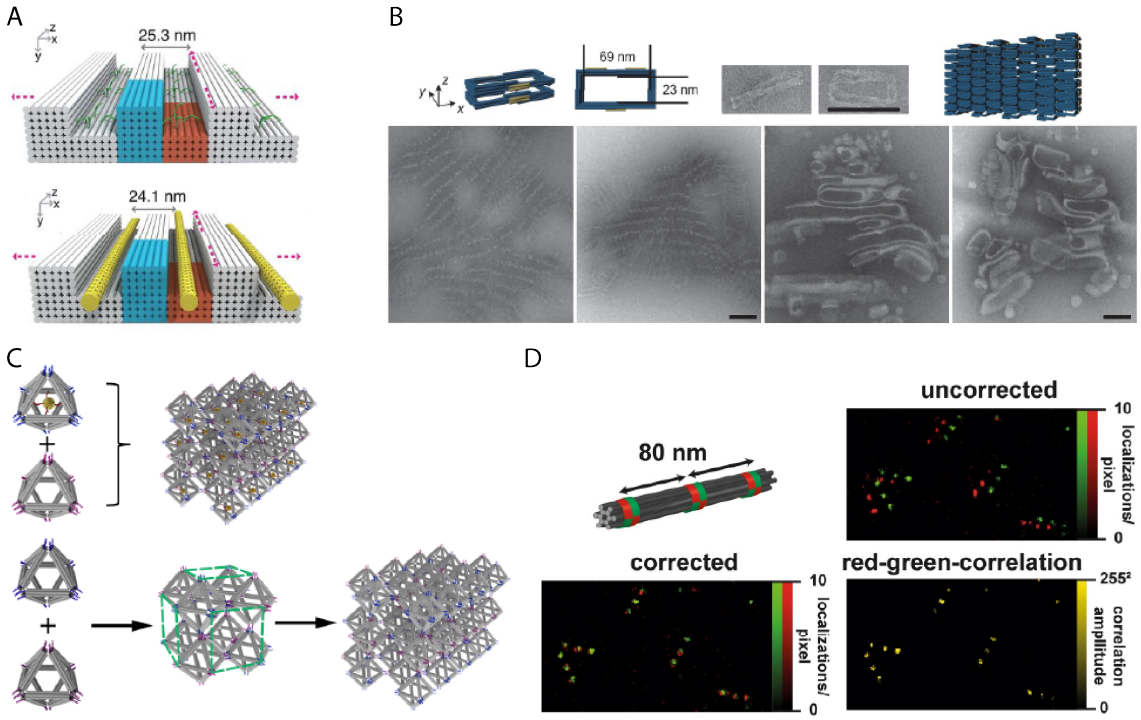

FIGURE 2.2: Applications of utilizing DNA nanostructures as a geometric template. **(A)** Pitch-scaling of carbon nanotubes by trenches constructed from DNA bricks. (Reprinted with permission from AAAS [102].) **(B)** Liposomes scaffolded by a DNA origami cage can replicate the appearance of the Golgi apparatus. (Reprinted with permission from Springer Nature [26].) **(C)** Gold nanoparticles placed within polyhedral DNA nanostructures that assist in placing the gold nanoparticles in crystalline lattice positions. (Reprinted with permission from Springer Nature [28].) **(D)** A DNA origami bundle with a tailored spacing of fluorophore markers that assist in the calibration of images against drift. (Reprinted with permission from Springer Nature [103].)

addressed on its interior or exterior surfaces.

The first demonstrations of DNA origami included a "world map" designed by attaching physical moieties to specifically addressed locations on the DNA origami surface, creating a pattern that could be viewed under an AFM. This principle has since been exploited to arrange a wide selection of other molecules. DNA nanostructures are addressable by binding functional groups, which can encompass organic or inorganic molecules, to nucleotides at the desired locations within the design. These

modifications are usually done at the 5' or 3' end of an ssDNA segment. However, some situations may call for modifications to be made mid-sequence, such as phosphorothioate for hydrocarbon modifications [104]. The ssDNA segment can be an extension of staple strands already included in the structural design of a DNA origami nanostructure or be a separate strand that hybridizes to a complementary domain extended from those strands. The latter option is typically chosen to be more economical and versatile as it decouples guest molecules from each unique staple sequence. A universal sequence can place the same guest molecule in multiple locations, reducing costs, as functionalization can be an expensive process to repeat for many distinct strands. It is more economical to synthesize a universal complementary sequence extending from the staple strands. If multiple unique locations are desired, then the sequences can also be chosen with the necessary orthogonality. As such, DNA nanostructure surfaces can place guest molecules with a resolution similar to its per base resolution of approximately 3.4 Å along a strand or 2 nm between helices [105].

Hybrid DNA-Lipid Nanostructures

Lipid vesicles are lipid bilayer membranes typically encapsulating some fluid. Generally, cells and organelles all have an encapsulating lipid bilayer membrane. Lipid vesicles that are artificially created are *liposomes*, which are self-assembled from amphiphilic lipid molecules. Liposomes preferably self-assemble into spherical shapes, but in natural systems, membrane proteins can also regulate the deformation of cellular membranes to dictate cellular function. The function of these membrane-deforming mechanisms still warrants study, and DNA nanostructures could help probe and mimic their effects.

DNA nanostructures have been shown to template the size and shape of lipid membranes. A DNA nanostructure template can be decorated with hydrophobic

functional groups. These are lipid membrane components that can embed into a lipid membrane, usually cholesterol, but also hydrocarbon chains [104]. A lipid membrane self-assembles around the DNA nanostructure spanning the anchored hydrophobic groups [37], and this has been shown to manipulate continuous lipid membranes into unnatural configurations such as cuboids [106] or winding tubular shapes that have also been shown to mimic the structure of the Golgi apparatus [26] (Fig 2.2B). The study of lipid vesicles is also made more challenging by inconsistencies of size and shape in leading synthesis techniques which still necessitate strict filtration of undesired specimens [107]. Round DNA barrels have been shown to act as a physical template for synthesizing lipid vesicles where it becomes energetically favorable for the lipid vesicle to conform to the size of DNA nanostructure template [38]. In biomimicry of nanodiscs [108], circular DNA templates can also consolidate a lipid bilayer for biophysical studies [51, 104]. Another circular template, a DNA barrel extruded using hierarchical assembly, was also shown to be able to template the formation of a lipid membrane along the outer profile of the structure, forming a long, tubular lipid membrane [52].

Biophysical studies of lipid membranes are also possible using DNA nanostructures as an architectural framework. For example, joint work by the Lin and De Camilli groups used a column-like DNA nanostructure as a ruler for spacing liposomes, which were themselves suspended in circular DNA origami templates at each end of the ruler, apart to study the transfer of lipids between their membranes [109]. Shaped DNA nanostructures can also be used to mechanically impose shape upon an existing membrane [36, 110, 111], similar to the BAR-family of proteins, which have a role in forming tubular substructures in cells, which among its functions creates filopodia, which are protrusions of the cell membrane that allow the cell to probe its environment. These studies help to show how DNA origami can elucidate the mechanical properties of cellular function. Composition and the organization of lipid

membranes and their membrane proteins can also be studied by tracking the position of DNA barges that have specific targeted behaviors or interactions upon the membrane surface [112, 113].

Molecular-Scale Imaging with DNA Nanostructures

The consistent and specific shape of DNA nanostructures make them a dependable reference object to assist in the characterization of other nanoscale moieties. Given the rise of techniques like super-resolution microscopy, it is now possible to observe specimens with nanometer resolution, but quantifying spatial information requires the development of a consistent frame of reference. Here, the geometry, including the size, length, and sometimes even aspect ratio, and stability of DNA nanostructures offers a consistent, widely compatible metric to measure by.

Martin *et al.* used a DNA nanocompartment with a hollow column to facilitate tomography measurements by suspending a desired analyte in the cavity [114]. The aspect ratio of the supporting DNA nanocompartment determines its orientation within a film of vitreous ice. At the same time, the structure can also protect the analyte from harmful interactions with surrounding molecules or the air-water interface. The shape of the supporting DNA nanocompartment provides an orientation reference, and the binding of an analyte to the DNA suspension wire can also tilt the orientation of the sample itself.

The shape variety of DNA nanostructures can also be used to resolve complex properties, such as single nucleotide polymorphisms, by shape [115], or by barcoding the surface with physical protrusions [116], then imaging with AFM. The consistent and predictable shape of DNA nanostructures can also be used as calibration or measurement references. Designed lengths of DNA origami nanostructures can provide an object of known size to perform *in situ* calibration for AFM analyses [117]. A vital property of linear DNA nanostructures of fixed length is that they can be repro-

ducibly synthesized in high yield. This is especially helpful to the communication of data, where instead of nanoscale standards needing to travel and risk damage, they can be communicated via the saved design of the DNA nanostructure. In several examples of previous work, fluorophores could be placed at exact locations with known spacing along the addressable surface of a DNA nanostructure to provide a reproducible nanoruler [118, 119]. In another example, Raab *et al.* use a DNA origami bundle with red and green fluorophores to calibrate against drift [103] (Fig 2.2D).

Nanofabrication of Plasmonic and Nanoelectric Devices

There are numerous strategies for DNA nanostructure templates to bind or interact with inorganic nanomaterials such as gold or carbon nanotubes. Thiol-modified DNA strands are a well-known method for conjugating gold nanoparticles to DNA nanostructures, but so can various peptides and proteins [120]. The geometry of a DNA nanostructure template can provide an unparalleled substrate to pursue the bottom-up assembly of electric and plasmonic nanodevices. Studies involving DNA-templated gold clusters have been numerous in the last decade. Recent work has also drawn attention to a wider variety of inorganic materials, such as silica and carbon nanotubes. Considering the breadth of proteins and peptides with specific material affinities that have yet to be used in conjunction with DNA nanostructures, a paradigm shift in the possible applications of DNA nanostructure templates for nanofabrication could occur soon.

In the last two decades, DNA-templated arrangements of metallic nanoparticles have produced a bevy of novel plasmonic nanodevices and provided insight into multiple phenomena. Some of the earliest demonstrations were with single DNA duplexes acting as tethers to assist the ordered aggregation of gold nanoparticles [121, 122]. Since then, many new techniques have been developed implementing complex geometric arrangements. The intensity and wavelength of plasmonic phenomena

heavily depend on the geometric layout and composition of its metallic components, such as gold, silver, and platinum. In single particles, the geometry is typically constrained by the extent to which crystallographic facets of the material can be chemically restricted for growth beginning from a seed. While this has generated an incredible variety of particles, including spheres, rods, triangles, and stars, arranging them for reproducible clustering or colloidal effects has been much more challenging.

Many applications have since demonstrated the advantages of tailored shape design using DNA nanostructures compared to existing biological templates such as viral capsids or bacterial scaffolds [120,123,124]. Chiral assemblies of gold nanospheres spaced helically along a straight DNA origami bundle emit a distinct circular dichroism signal [125]. Similar effects are also generated by gold nanorods oriented along the edges of an adjoining pair of DNA origami triangles [126], and in general, there have been numerous methods to produce chiral orientations using a DNA nanostructure support [127–133].

Nanoparticle geometry can also determine field enhancement effects. For example, gold nanospheres arranged into a square on a flat DNA origami template form a SERS hotspot in the negative space in their middle [134]. The Ding group and collaborators used a DNA origami template to position two gold triangles with nanometer precision in a bowtie shape to generate a high local field enhancement between their tips [135]. The Ke group and collaborators created DNA nanotubes from a honeycomb lattice similar to how carbon nanotubes appear as rolled sheets of graphene, then set gold nanoparticles into specific lattice holes to produce plasmonic metamaterials [136].

Over time, the scale of DNA-templated nanoparticle assemblies has also become increasingly complex by utilizing the modular assembly of DNA nanostructures into multicomponent, superstructure assemblies. Zhu *et al.* reported a modular strategy in which, first, gold nanoparticles are placed anywhere within a half-hexagonal

DNA origami trough. Those troughs, each as a module, could be linked in series to form a specific, linear assembly of gold nanoparticles [137]. The Gang group demonstrated a strategy for utilizing polyhedral DNA origami as a frame for fabricating gold nanoparticle clusters with high spatial complexity. The attachment of gold nanoparticles on selected vertex points of the polyhedral DNA frame produced specific geometric arrangements of gold nanoparticles that can mimic the position of atoms in a crystal lattice unit cell. This could be done either using a single, high-degree polyhedron [138] or a large, crystalline assembly of low-degree polyhedra such as cubes and octahedrons, where gold nanoparticles could be placed at either the vertices or at the center of the structure [27, 28, 139] (Fig 2.2C). The Turberfield group demonstrated a deterministic strategy of tethering gold nanoparticles using DNA origami nanoflowers, which wrap around a gold nanoparticle giving it rotational specificity [140].

DNA-assisted fabrication of electronic nanodevices has also garnered significant research attention. Carbon nanotubes (CNTs) have been an enticing candidate to outperform and replace silicon-based electronics. However, they are difficult to individually handle and align to precisely tailor the electrical properties of any CNT connection. Sun *et al.* used micrometer-scale trench-like structures constructed by DNA bricks as a solution to the challenging problem of aligning and positioning CNTs between two nodes [102] (Fig 2.2A). DNA strands attached to the CNTs guided and locked them into parallel trenches that were only wide enough for a single CNT, thereby physically filtering and aligning them. His group later used this technique to create FET transistors [141]. Various metallic growth strategies have also explored site-specific growth of metallic nanostructures based on a DNA nanostructure template. Bayrak *et al.* used the nanocasting method to create gold nanowires by growing a gold crystal through a chain of DNA origami molds [31]. Jia *et al.* demonstrated how low-valence metal ions (Cu^{2+} and Ag^{2+}) aggregate to and begin

to metalize upon clusters of long ssDNA preferentially over stiffer DNA origami as a seedless approach to the fabrication of metallic nanostructures [142]. Shen *et al.* used DNA origami nanostructures as a lithography mask to produce similarly shaped patterns of gold upon a silicon nitride chip [143] to create plasmonic nanostructures.

2.4 Conclusion

Despite the myriad of applications, the shape variety carried forward toward downstream applications is just a tiny subset of possible DNA nanostructures. When considering the application of DNA nanostructures to a molecular landscape, there is no need to overthink what primitive functionalities are necessary other than to refer to what has evolved naturally. Spatial organization relies significantly on the compartmentalization of molecules. Matrices of fibrous structures often support and distribute molecules within that space. Becoming somewhat of a ubiquitous material for spatial organization at the molecular level, DNA nanostructures continue to demonstrate these primitives with incredible adaptivity. These structures undoubtedly are set to become more complex. To maintain the structural and functional fidelity of DNA nanostructures in their downstream applications, design principles must likewise scale to higher-level abstractions of DNA nanostructure design to heighten the integration of purpose without being hindered by low-level decisions for determining shape and yield.

3

Generalized DNA Origami Design and Automations

3.1 Introduction

DNA origami is a powerful technique that has significantly expanded the variety of DNA nanostructures. Design methodologies that followed, such as DNA bricks and polyhedral DNA nanostructures, developed automated, algorithmic strategies with repeatable and scalable local design principles, which translated to a major leap forward in the scale and geometric complexity of DNA nanostructures. These strategies were able to leverage algorithmic patterns and computational power to apply complex design principles beyond those that a human designer could comfortably work with. A notable contributor to these methods' success has been that, despite their geometric complexity, the mathematical formalism of their geometries simplifies one of the most challenging design principles, crossover placement. Polyhedral structures can be defined as a set of vertices and edges, and bricks can be defined as voxel positions. Crossovers are placed at vertices in polyhedral designs. Due to the fixed domain lengths in DNA bricks and the shape of each brick, crossovers occur as an effect of correctly placed bricks and appear with periodic regularity throughout the

design.

On the contrary, while less geometrically complex, basic DNA origami designs, like those based on the packed lattice, do not share these constraints on crossover positioning. This lack of constraints yields an optimization problem for crossover placement that has not since been seriously considered nor resolved. That is not necessarily peculiar. Many such shapes still utilized in applications are at a scale and complexity where a human designer can reasonably complete a desired design in less than an hour. The simple block-like shapes allow a local crossover pattern to be repeated throughout most geometrically similar portions of the design. That is, the positioning of a few crossovers in one area of the structure likely applies throughout many other areas of the design. Despite the difficulty of a mathematically rigorous solution, to a human, it appears much simpler to create a design that ensures the correct shape, high yield, and high structural integrity of the structure. Compared to the difficulty of automation, it is understandable that there is limited interest in automating DNA origami design from its fundamental design principles.

However, the advantage of simplicity may not hold forever. As previously discussed in Chapter 2, many applications still utilize but, at the same time, are also limited by packed lattice designs. The plane-based routing of packed lattice DNA origami yields a compact shape when extruded to 3D but also has a high material overhead cost. The maximum shape complexity and feature resolution are limited, as each helix requires a minimum number of connections to any adjacent helices to preserve its intended position within the shape. That minimum number is thought to be at least two full turns, which in B-form DNA would be equivalent to 21 base pairs or approximately a length of 7 nm. Also, considering the expected interhelical distance of around 2.6 nm, this is a significant overhead of material required to create even minor protruding geometric features.

Design principles within the scope of fundamental DNA origami design, such

as those that allow curvature [7, 10, 74], can surpass the previously listed feature resolution limits. The availability of orthogonal scaffolds [78, 144] will also begin to relieve size limitations of all DNA origami-derived design strategies. This enables DNA origami designs to be created with more than one scaffold, which was previously a hard limit for DNA origami design. These developments motivate this thesis to revisit the previously raised issue of automating fundamental DNA origami design that appeared trivial but will no longer be for future designs of greater complexity. It investigates the traditional design principles of DNA origami, how to quantify them and abstract them to higher-level design properties, and possibilities for novel geometries that challenge previous design constraints.

The remaining sections of this chapter are organized as follows: Section 3.2 discusses fundamental concepts and terminology relevant to discussing DNA origami design. Section 3.3 reviews the design principles of several of the most popular DNA nanostructure design techniques. Section 3.4 provides a walkthrough of DNA origami design from the perspective of a human designer, which will cover basic principles we will refer to in Section 3.5 for discussing a proposed set of higher-level abstraction parameters and how each relates to the design of DNA origami. Section 3.6 then discusses strategies for utilizing the consolidated parameters in automated design pipelines for DNA origami.

3.2 Fundamental Concepts and Terminology

A *DNA nanostructure* is a nanometer-scale object made entirely from DNA. A *duplex* refers to dsDNA, which is a rigid molecule relative to ssDNA. DNA nanostructures of specified shape can be designed by the relative positioning of duplexes to each other, affixed by *crossovers* linking nearest neighbors. A crossover is a continuous segment of a DNA strand that spans two discontinuous segments on a complementary strand, forming a junction similar to the Holliday junction that positions the discontinu-

ous segments spatially close together. The placement of crossovers throughout the nanostructure is determined by the letter sequence of each DNA strand to program their complementarity. The deliberate placement of these crossovers determines the design's shape, stability, and yield, and the total patterning of crossovers across the design will be referred to as the *crossover network*.

DNA origami is one of the most popular techniques for designing DNA nanostructures. The basic principle of the technique is to use many, usually hundreds, of short (< 100 nts) ssDNA strands termed *staples* to "fold" a few, often only one but more recently up to two or three, long (> 1000 nts) ssDNA strands termed the *scaffold* strand or sometimes the *lambda sequence*. Staples are typically added at a 10-fold excess concentration relative to scaffolds into a *one-pot synthesis*, where all staples and scaffolds are added into a single mixture. Next is a *thermal annealing* step to synthesize the structure from this mixture. The solution is heated to about 95°C, much higher than the *denaturing* temperature of DNA around 60°. Denaturing refers to breaking hydrogen bonds that hold a dsDNA molecule together in its familiar double helix shape, such that all duplexes become ssDNA. *DNA hybridization* refers to the forward process, where all ssDNA bind together by their complementary domains, forming as many and as entirely complementary duplexes as possible. Following the heating, the mixture is cooled to room temperature over at least 2 hours but sometimes as long as 72 hours, with the cooling rate most gradual between approximately 35°C to 70°C. The obtained nanostructures are often purified for downstream purposes using gel electrophoresis or spin columns to remove unbound staple strands. Structures are frequently imaged by atomic force microscopy (AFM) or negative-stain transmission electron microscopy (TEM).

B-form DNA is the most stable form of duplex DNA and is the preferred form to work with. When handling DNA for synthesizing nanostructures, buffers are adjusted to maintain DNA in this form as much as possible. In B-form DNA, the DNA

molecule has an *axial rise* of 0.332 nm and a *helical twist* of 10.5 bps/turn. The axial rise is the distance between each successive nucleotide in a sequence. The helical twist is the number of bases to count along a duplex to arrive at a nucleotide along the helix at the same angle. A single count of this distance is referred to as a *full-turn*. Other terms like *half-turns* or *quarter-turns* denote the linear distance to find nucleotides with 180° and 90° angular differences to the initial nucleotide, respectively. These representative terms are often used for lattice-based DNA origami, where potential crossover positions are distributed at discrete intervals. For example, for a 2D DNA nanostructure where all helices are planar, only crossovers at full-turn and half-turn intervals will be used to link each pair of neighboring helices.

DNA origami mainly utilizes *full crossovers* (or sometimes also called a *double crossover)*, which are two adjacent crossovers. However, "crossover" is often used instead of the complete terminology. Thus, a crossover always involves two pairs of nucleotides, one on each helix to be linked. We refer to these pairs as *nucleotide pairs*. When determining a potential crossover position, we do not evaluate the angle of a single nucleotide but rather the bisecting angle of the nucleotide pair. This angle is referred to as the *dihedral angle*.

Additional terminology that will be used throughout this chapter is as follows:

1. A *nick* is a position along a duplex where a continuous strand is disconnected into two separate strands.

2. A *region* is a continuous segment on any duplex between nicks and crossovers. Long regions have greater binding strength between their complementary ss-DNA.

3. A *scaffold routing* is the sequential path of the scaffold strand through each helix of the structure.

4. A *staple routing* is the sequential path of the staple strand through any number of helices of the structure.

3.3 Common Design Principles

This section describes the design primitives of various DNA nanostructure design strategies. While numerous variations exist, this thesis summarizes DNA nanostructure design strategies into only a few categories defined by their design principles. These include designs on packed lattices, polyhedral, and curved nanostructures. An exception to the DNA origami strategy that has demonstrated success is DNA bricks, which does not use a scaffold strand and instead uses a huge pool of only short oligonucleotides.

3.3.1 DNA origami on 2D and 3D packed lattices

The DNA origami technique, pioneered by Rothemund [4], rasterized geometries using only straight and parallel helices. In a 2D, planar design, modulating the length of each helix forms the contour of a shape when viewed normal to the plane. Each helix forms crossovers with two adjacent helices, one on each side of it. So along each helix, crossovers are separated by half turns, alternating between each of its neighbors. In B-form DNA, this is approximately every 5 or 6 bases. Crossovers should be spaced at least one full turn away from any adjacent crossover along the same helix. This helps to ensure that in between these crossovers, there is a sufficiently long continuous region to improve the binding strength between the complementary ssDNA. Strands that are only complementary for a short region can come apart due to weak binding, thus damaging the structure's overall stability. A single, planar DNA origami is referred to as a *sheet*.

This technique can be extended to 3D by stacking sheets together, like a stack of paper [8]. The lattice choice, honeycomb, square, or triangular, determines the

positions of crossovers along each helix and the intervals at which they occur [145]. Adjacent crossover positions are divided into quarters, thirds, or sixths in honeycomb, square, or triangular lattices, respectively. These lattices are often used to create *bundles*, which are named either by their total number of helices present (e.g. 6-helix bundle) or their dimensions as width $\times$ height (e.g. 3-by-2 helix bundle).

3.3.2 Polyhedral

The polyhedral strategy [5,24,32] of DNA origami design is also based on the principle of using short staples to fold a long scaffold. However, helices are no longer placed straight and parallel to each other. Instead, shorter, straight segments of helices represent the edges in a polygonal structure. The most challenging problem for these structures has been how to route the scaffold. It was discovered that solutions could be adapted from the Eulerian path problem. Once the scaffold is routed, staples are also generated to lock the routing of the scaffold into place.

Second, a critical design principle is how the routing handles the transition of the scaffold and staple strands at the vertices of the shapes. Depending on the relative positions of polygonal faces to each other, either full-turn or half-turn crossovers are used at the junction of DNA segments at the polygonal vertices. Thus, this limits the geometric resolution of each edge and face, as edges need to be intervals of full-turns or half-turns to ensure that their endpoints that meet at vertices can appropriately form crossovers with adjacent edges. Most notably, this also constrains the locations of the crossovers. The vertices in the geometry help to determine crossover placement, which is a critical reduction to the complexity of its design compared to packed lattice designs.

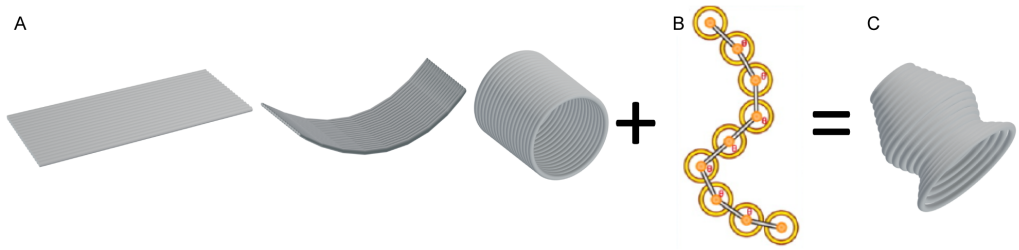

FIGURE 3.1: Conceptual conversion of a 2D DNA origami sheet to a 3D DNA capsule. **(A)** Beginning with a 2D DNA origami sheet, if it is rolled into a cyclic structure, it creates a cylinder. **(B)** Editing the dihedral angles of each helix creates some extrusion of features. (Reprinted with permission from AAAS [10].) **(C)** When this is applied to the cyclic template, it creates curvature along the profile of the structure.

3.3.3 Curved and rounded

Rounded nanostructures can be designed by connecting each helix by its endpoints, such that it is cyclic [71] (also see Fig. 3.1). Due to the rigidity of a DNA duplex, after being bent such that it is cyclic, the helix relieves the stress of curvature evenly throughout its length, thus becoming a rounded, circular shape. This strategy also follows the traditional technique of DNA origami using short staples to fold a long scaffold strand.

A key difference in this technique is the usage of all dihedral angles along the double helix instead of simply those at discrete intervals. By doing so, each circular helix can be scaled to a circumference relative to its adjacent helix, such that it is spaced no further away than the interhelical distance $\delta = 2.6$ nm. When each consecutive helix is placed with some offset to the last, these curved structures can also exhibit curvature along two dimensions. *Out-of-plane curvature* refers to curvature along the profile of the shape, while *in-plane curvature* refers to the curvature along the length of each helix. In-plane curvature is a natural property of the cyclic helices, while out-of-plane curvature results from the unrestricted choices of dihedral angles.

A challenge to this method is that the circumferences, determined by satisfying

the constraint of interhelical distance, are not always integer multiples of the helical twist of B-form DNA (10.5 bps/turn). The effect is that instead of a predictable and dense availability of crossover positions between adjacent helices, due to their varying periodicity, there are very few crossover positions as would be decided by previous conventions. This will be further discussed later in Chapter 4.

3.3.4 DNA Bricks

DNA bricks [6, 19] is an alternative strategy distinct from DNA origami. Instead of using a scaffold strand and staple strands, it uses many short oligonucleotides, each with typically four domains of 8 nts or 13 nts long. It uses each as a *voxel* within a cubic DNA nanostructure. Strands for each voxel form a single crossover (not a full crossover) to one other voxel. In repeating this same strategy through the structure, the relative position of all voxels is determined in 3D. This strategy can form filled cuboids of voxels and then selectively exclude them for creating shapes or cavities. Overall, the helical axes in these designs are still aligned to positions following a square lattice.

3.3.5 Optimization

Several optimization techniques of DNA origami have been investigated in previous literature. These pertain primarily to the hybridization strength of each complementary region within a design. Aside from ensuring a high quantity of well-placed crossovers, several design principles exist to maximize the yield of the designed DNA nanostructure. While primarily evaluated for packed lattice structures [56, 64, 146], it is likely, but not proven, that the principles generalize to any variation of the DNA origami technique.

The presence of "*seed*" regions for each staple strand has a notable impact on the yield of the nanostructure. A seed region is a continuous region of at least 14 bps in

length, such that a staple is guaranteed to have some subsequence that binds strongly to the scaffold strand. This is implicitly achieved by the corresponding placement of crossovers and nicks but is not always guaranteed due to the density of crossovers and nicks. By ensuring that each staple has at least one seed, all staples are more likely to find, bind, and remain in the local proximity of the scaffold strand. This helps shorter regions form with greater success.

Furthermore, any subsequence of the staple strand that is not the seed region should preferably be at least 7 bps long to reach a minimum binding strength with the scaffold. Generally, this means that crossovers should often be placed more than 21 bp apart. That would accommodate a seed region for one staple and another minimum region for another staple that would have its seed region elsewhere. While principles may not always be perfectly applied to every staple, it is advantageous to use them as much as possible.

Furthermore, B-form DNA can have a slightly varying helical twist in buffers with different ionic concentrations. Each crossover position is rarely perfectly angled towards the position of its adjacent helix, and each minor misalignment can add strain to the structure. Additively, this causes global stress in the intended shape of the DNA nanostructure, which, when the deformation is significantly large, the yield and stability of the nanostructure begin to decrease. These are categorized into two effects. *Overwinding* refers to crossovers appearing along a helix slightly before it completes a full turn causing the helix to twist more than intended to meet crossover positions. *Underwinding* refers to the opposite effect, whereby the helix twists less than intended. Each of these can cause the entire structure to have a twisting deformation, either in the left-handed direction from overwinding or in the right-handed direction from underwinding. The most straightforward method of compensating for this deformation is to ensure that a structure has crossovers that introduce both overwinding and underwinding by periodically placing crossovers either earlier or

later along the sequence. This is often referred to as *twist compensation*.

The *crossover density* is an aggregate property of a DNA nanostructure that is sometimes referred to informally when associating the stability of a nanostructure with the number of crossovers it contains. However, it does not actually reveal any information pertaining to the specific local choices per crossover. It is generally accepted that ranges of crossover density roughly correlate to the observed structural integrity of designed structures. If crossovers are very sparse, the structure is unlikely to form even an approximate shape. As crossover density increases, the likelihood of an accurate shape forming also increases. However, this either plateaus or begins to trend downward past a currently unknown limit, after which adding more crossovers has diminishing returns in making a more accurate shape. Instead, more crossovers past that limit would continue to add strain and reduce the frequency of long regions, thus increasing structural deformation and weakening the bonding strength of the scaffold strand to staple strands across the structure.

3.4 Walkthrough of DNA origami design

This section explains the basic concepts of DNA origami design by walking through the simple example of creating a 2D rectangle DNA origami shape. The figures shown here depict snapshots of major steps from the perspective of a human designer using caDNAno to design an on-lattice DNA origami nanostructure.

We first begin by indicating the positions of helices to create a canvas for the intended shape. On the lattice-based design of caDNAno, the helices are first set up using a cross-sectional view of the helices (Fig. 3.2).

In the next step, we view each helix with a level of abstraction that displays each helix as a discretized row of cells. Each cell represents a nucleotide along the helix. Scaffold strands are added to the canvas to define the length of each helix. The position and length of each helix determine the *shape template*. That is, the

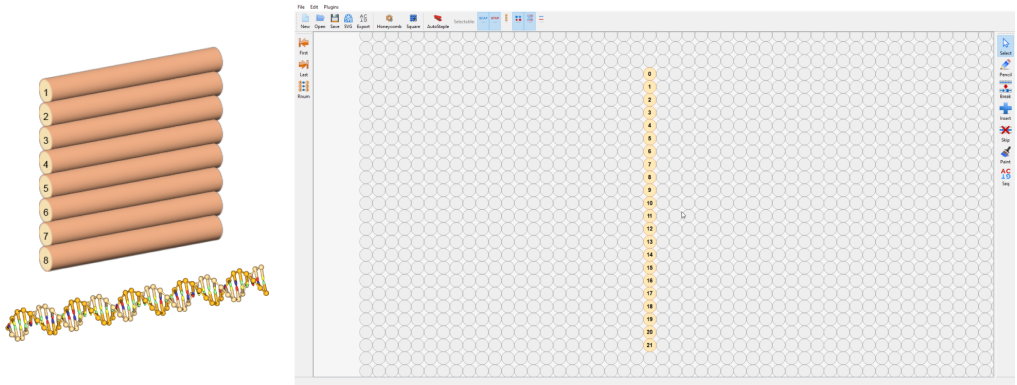

FIGURE 3.2: Manual DNA origami design, Step 1. Cylinders abstract the position of each double helix in a shape template. Each row of the user interface represents each helix, and cell positions along each row represent nucleotide positions. The ssDNA scaffold strand and its crossovers are initially added to rasterize the shape template.

outline of the shape template should match the intended shape we seek to design. Crossovers in the scaffold strand are added to complete the scaffold routing, such that the scaffold strand is one continuous strand running throughout the entire shape template (Fig. 3.3).

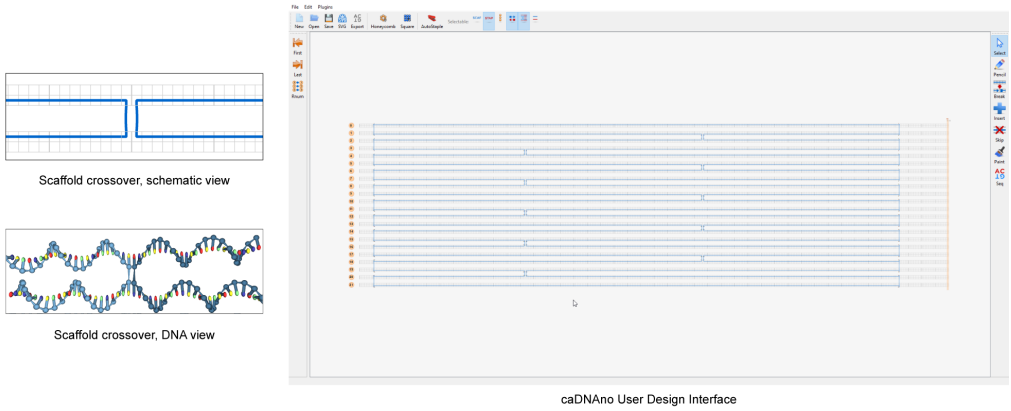

FIGURE 3.3: Manual DNA origami design, Step 2. ssDNA scaffold strands are added into the interface. Scaffold crossovers are placed to form one continuous scaffold loop throughout the shape.

Then, in the third major step, staple strands are drawn, mirroring the location

of each scaffold strand. Now, all sections of the shape will be duplexes. A periodic array of crossovers are added to make several loops of staple strands that are, at this step, still quite long. However, notice that these crossovers can be placed at regular intervals, and globally, a pattern is noticeable (Fig. 3.4).

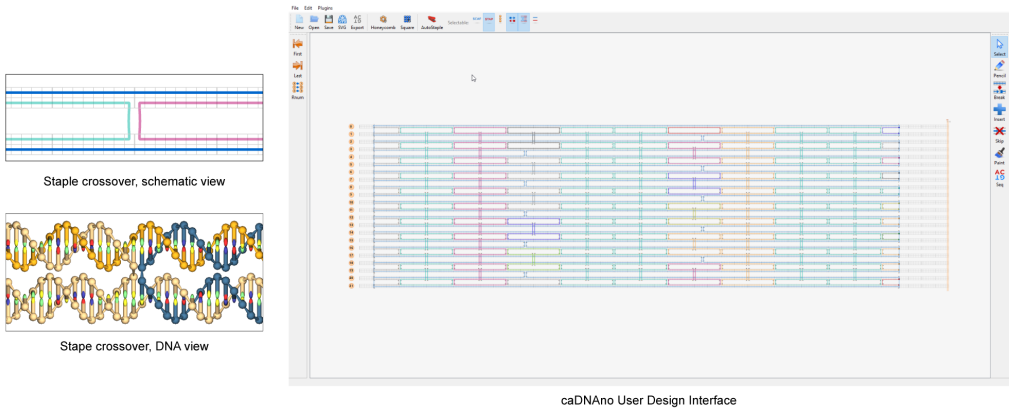

FIGURE 3.4: **Manual DNA origami design, Step 3.** Staple strands and crossovers are added to the structure.

Finally, nicks are added to break each looped staple sequence into segments that are preferably within a length range (typically about 30-50 nts) that, empirical evidence shows, promotes high yield synthesis of DNA nanostructures (Fig. 3.5). This is often referred to as the *folding kinetics*, which describes the relative ease with which staple strands find and fully bind to complementary positions on the scaffold strand. In brief, short strands (< 20 nts) can be unlikely to remain bound to the scaffold due to having an insufficient length to bind strongly. Long strands (> 100 nts) can be difficult to fully bind in the first place, either due to becoming entangled or needing to bind too many distinct regions simultaneously. Again, nicks can be placed with mostly periodic regularity, greatly reducing the number of distinct design decisions for the human designer to be well below what would be expected for the size of the nanostructure.

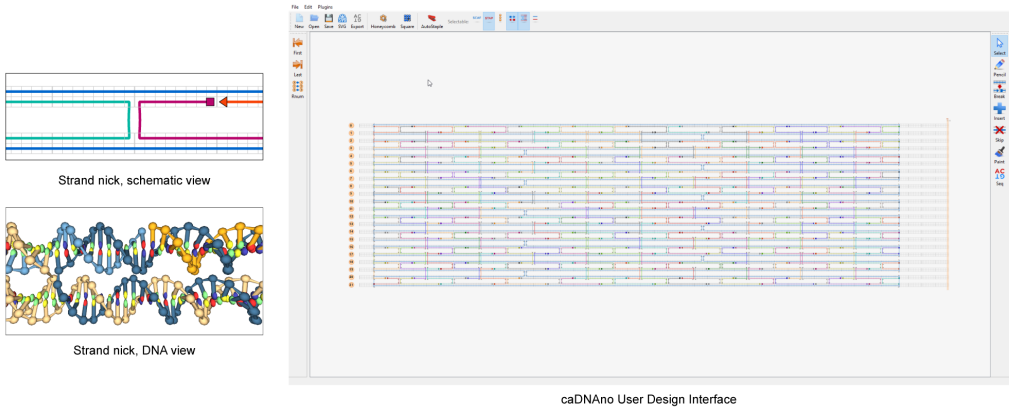

FIGURE 3.5: Manual DNA origami design, Step 4. Nicks are placed along staple strands at intervals to reduce all staple strands to be within a range of lengths conducive to high-yield synthesis.

3.5 Consolidation of Parameters

From the above walkthrough, we carefully consider the decisions made by the human designer and the quantities and properties considered in each major step and summarize these as our *design principles*. The design principles are the comprehensive set of rules applied in determining the crossover network and nicks within a DNA nanostructure to generate a list of staples that reliably form the intended design with high yield, stability, and accuracy. While crossover density is often a reliable global value to quantify and compare crossover network variations, it does not contain any information applicable for local decisions. Instead, we present DNA origami design in the following parameters:

- **Same helix crossover spacing** (SHS) - The base pair length between two crossovers that span the same pair of adjacent helices (see Fig. 3.6).
- **Common helix crossover spacing** (CHS) - The base pair length between two crossovers that span two different pairs of adjacent helices, with one in common (see Fig. 3.6).

- **Staple length** - The base pair length of each staple strand that is preferably maintained to a range corresponding with high synthesis yield (see Fig. 3.6).
- **Nick proximity** - Constraints indicating how close nicks can be placed to crossovers and other nicks (see Fig. 3.6).
- **Crossover bonding distance** - The metric length measured between the mid-points of two pairs of nucleotides with respect to their position, then projected onto the central axis of the helix. Each pair is on a separate helix in a pair of adjacent helices (see Fig. 3.7D and later, Fig. 4.5).

Furthermore, we consolidate the following objects within a DNA model:

- **Nucleotide Pair** - Any pair of adjacent nucleotides on a helix.
- **Half Crossover** - The crossing over of a single staple strand from one subsequence of the scaffold to another.
- **Full Crossover** - Two adjacent half crossovers.
- **Nick** - Removes the 5'-3' binding of nucleotides in adjacent positions to the nick, such that the strand is "disconnected" at that location.

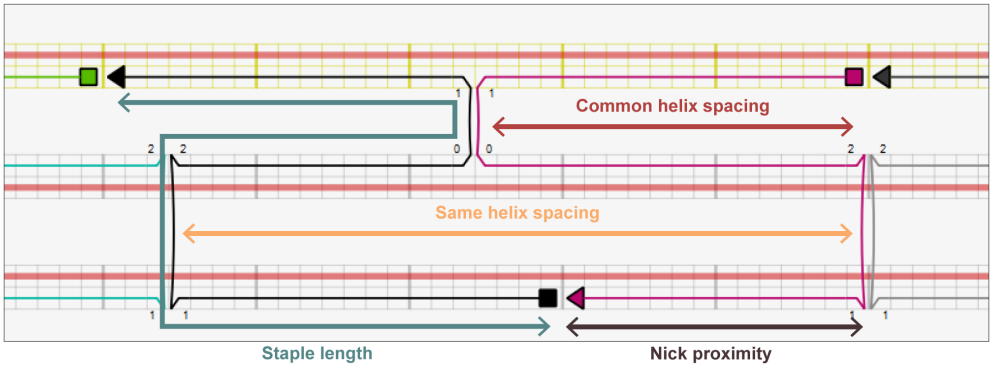

FIGURE 3.6: Diagram of spacing parameters.

3.6 Automation

Absorbing all of the above parameters, we consider algorithmic foundations of generalized automatic DNA origami design. To clarify the scope of this problem, we do not consider accomplishing this from a "shape-level" abstraction. That is, structures like polyhedral structures can exploit established graph algorithms to convert a shape into the placement of each helix. Instead, we assume that the arrangement of helices is already known either by manual input by a human designer or generated from another algorithm. Next, we consider instances of DNA origami with room for non-determinism. That is, multiple different choices in the crossover network can form the same structure, with at most trivial differences.

Instead, the construction of automating a generalized strategy draws inspiration from the manual process a human designer undergoes to design DNA origami, as presented above in Section 3.4. This thesis describes an automated methodology based on that process as follows:

1. The length of helices is determined and each is correctly placed. Helices are placed as duplexes. This is a shape determination step for creating the *shape template*.

2. A scaffold routing is determined. By placing crossovers, the scaffold strand must form a continuous path through each other helix in the design. This step has few constraints as the template has no other crossovers. Thus scaffold crossovers can be placed anywhere as long as they are at aligned positions between the two adjacent helices to be connected.

3. The next step is placing staple crossovers. Once there is a satisfactory local solution in simple packed lattice structures, it can often be repeated across the entire design since many sections will be identical. This is an advantage of

crossover positions being periodically available at the same interval between each pair of adjacent helices and throughout the entire structure. Crossovers can also be positioned with consideration for the optimization techniques as discussed in Section 3.3.5.

4. Finally, nicks are placed to segment staple strands into lengths ideally between 30-50 nts for favorable folding kinetics.

The steps to place staple crossovers and nicks are the most complex parts of this process. If crossovers are placed poorly, such as too close together, it hinders the ability to choose favorable nicking locations. A human designer may often move problematic crossovers to a new location, then update nick choices. This process is repeated if that move should create new problems. In general, this process could be summarized as first loosely applying a crossover network, looking for problems, making alternate choices, propagating those decisions, then reviewing again for additional issues, changes, and updates and iteratively applying this process until arriving at a satisfactory solution with no observable conflicts.

While the goal of generating a valid crossover network is to maximize the stability of the nanostructure, there is no known relationship capable of immediately mapping the crossover and nicking choices of a DNA nanostructure to its predicted stability. Molecular dynamics simulations can be done, but the feedback is not immediate. Many techniques widely applied to optimize DNA nanostructure yield and stability are followed from experience and empirical evidence. Furthermore, these techniques may not all be simultaneously applied. Even if there is an optimal solution, there is currently no known way to find or verify it. Instead, finding an assumedly suboptimal solution that satisfies the desired constraints and appears outwardly correct is more pragmatic.

The iterative process above suggests trying to implement a *heuristic algorithm*,

which is an algorithm that does not attempt to solve for the globally optimal solution exactly, but via an iterative process, attempts to discover an approximate solution. A heuristic algorithm focuses on immediate, local improvements based on a set of constraints and is expected to converge onto an approximate solution that may or may not be optimal but is expected to be close. This chapter discusses the development of automated DNA origami design based on two heuristic algorithms. One is based on *greedy decision-making*, which takes the next best solution at every iteration without observing any context. Another is based on *simulated annealing*, which begins with any possible solution state S_0, then generates and scores small, random permutations of that solution. A new solution state S_{i+1} is selected by scoring it and comparing it to the score of the current state S_i, with some padding factor, often referred to as the temperature T. Depending on this difference, S_{i+1} has some probability of being selected as the next state. However, for the sake of speed, in the context of DNA design, these probabilities are weighted to be binary instead.

3.6.1 A digital workspace for automation

For applying the heuristic algorithms, this work creates a digital workspace with a coarse-grained digital representation of DNA helices by initializing the position and orientation of each nucleotide in a 3D Cartesian coordinate system. Three primitives of DNA shapes are implemented: ring, line, and arc (Fig. 3.8). Each geometric shape can be scaled and converted into a helix according to the intended helical twist. Multiple modules are initialized and positioned into a *template* of the intended DNA nanostructure shape. (Note: The shape template is an abstraction of the shape using only helices, while the template represents that abstraction converted into the coarse-grained model with nucleotides.) This template evaluates crossover positions by first reducing relevant positions on a helix to nucleotide pairs with a neighboring helix that have an aligned dihedral angle. Each nucleotide pair is then evaluated

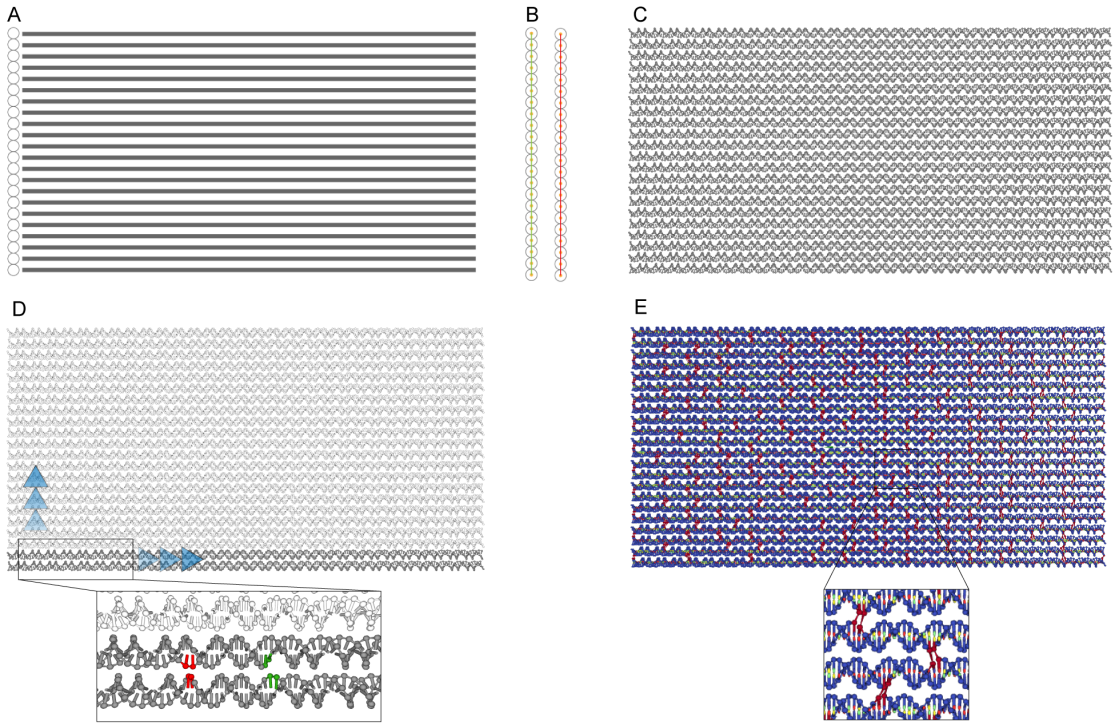

FIGURE 3.7: Preprocessing steps of automating flat DNA origami design. Using the example of a simple 2D DNA origami, we observe the following preprocessing steps. **(A)** Initialize the shape template. **(B)** Set both the scaffold (green) and staple (red) connectivity. These currently appear the same due to the simplicity of a 2D shape. **(C)** Convert the shape template to a digital DNA model using the implemented modules as shown in Fig. 3.8. **(D)** Iterate across each pair of neighboring helices across the entire design, evaluating nucleotide pairs and labeling potential crossover positions. **(E)** Heuristic processes determine the crossover network.

pairwise with each nucleotide pair on the adjacent helix. Then, simply filtering by their Euclidean distance, aligned nucleotide pairs close to the interhelical distance δ = 2.6 nm are saved as potential crossover positions. These crossover positions can then be input into a heuristic algorithm to generate a crossover network (see Fig. 3.7).

Input information thus contains:

1. Helix positions and shape.

2. A *pathway* list that indicates the visiting order of the scaffold strand during

the heuristic.

3. A *connections* list that serves as a lookup table for finding the nearest neighbors of each helix.

4. An *edge* is the graphical representation of which pairs of helices should have crossovers between them.

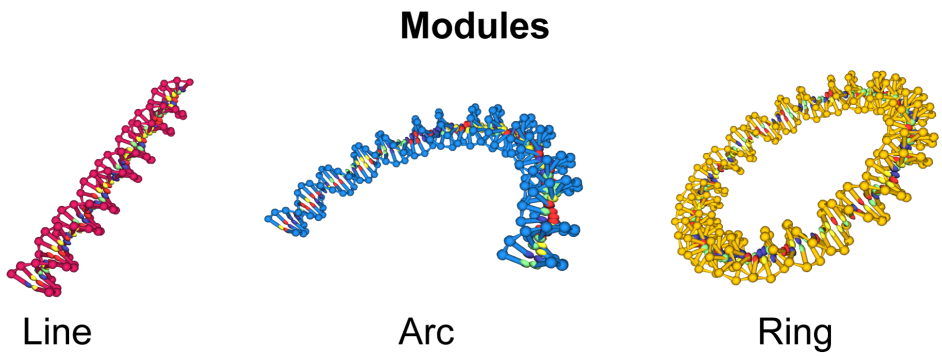

FIGURE 3.8: Modules for basic DNA origami shape primitives.

3.6.2 Greedy decision-making

The overarching goal of either heuristic is to distribute a crossover network as uniformly as possible while not violating constraints, such as the SHS and the CHS. We first discuss the application of a greedy decision-making process. A flowchart detailing its processes and loops is shown in Fig. 3.9. The greedy algorithm begins at the beginning of the pathway and initially places scaffold crossovers. Since there is no prior crossover information in the structure yet, scaffold crossovers are added at a staggered pattern with a fixed offset instead. For instance, every subsequent crossover is placed 1/3 or 1/2 the length of the structure away from the previous scaffold crossover.

Next, staple crossovers can be placed. Decisions in setting crossover positions in DNA origami designs are expressed as ensuring that each crossover maintains as much

space away from any adjacent crossovers. Doing so maximizes the continuous regions of staple-scaffold binding across the structure. If each crossover is subsequently placed one at a time, solutions are less likely to be uniform. Instead, per each helix, crossovers are placed in groups. Each crossover is assigned a value, which correlates to its position on the helix. A *largest minimum distance* algorithm for a fixed number of elements is then applied to determine a subset of crossovers such that the shortest distance between any adjacent crossovers is the maximum possible value.

However, this only describes the conditions on the first iteration of a pass, following the input pathway, through the structure. Since no other crossovers have been placed, the crossover network solution for the first pair of helices h_1 and h_2 is easily accomplished. However, for a subsequent pair of helices, crossovers between h_2 and h_3 must also respect the existing placement of crossovers on their common helix h_2. Here, a conflict can arise. The remaining positions for crossovers between h_2 and h_3 may be sub-optimal compared to the previous solution for (h_1, h_2). This sub-optimality can propagate through the structure, resulting in consecutively less optimal choices as the algorithm proceeds through helix pairs along the pathway.

3.6.3 Simulated annealing

To remedy this issue, we resort to a heuristic algorithm based on simulated annealing. A flowchart detailing its processes and loops is shown in Fig. 3.10. Examples shown in Figures 3.11 to 3.16 illustrate the critical steps. To begin an automated design following the simulated annealing-based methodology, we follow the same preprocessing steps shown in Fig. 3.7. Fig. 3.11 shows a shape template with crossover positions between a pair of helices discovered and labeled. Fig. 3.12 uses a few of the discovered scaffold crossover positions to complete the scaffold routing. Next, Fig. 3.13 initiates the simulated annealing process with a sparsely populated crossover network as the initial state. The initial state is scored by the maximum observed

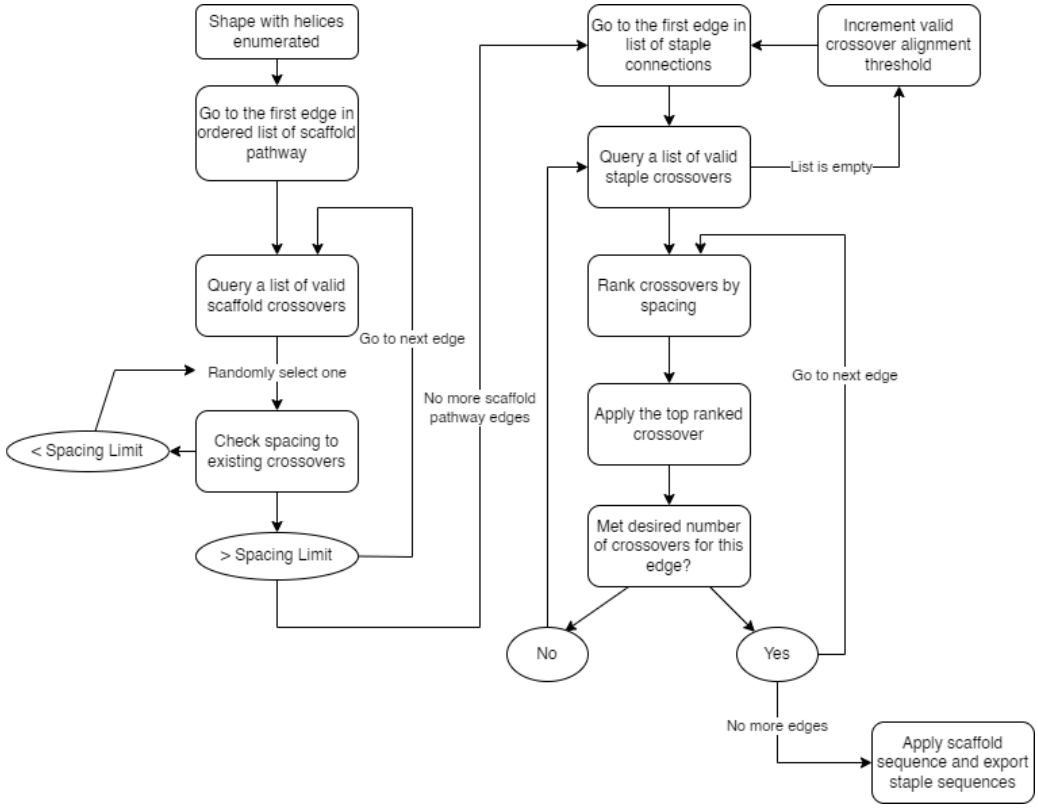

FIGURE 3.9: Flowchart of greedy DNA origami heuristic design algorithm.

crossover-to-crossover distance, only considering crossovers that exist on the same helix pair (Fig. 3.15). For large values, especially those close to the length of each DNA helix, a high score indicates that helices with sparse sets of crossovers exist. In contrast, a low score indicates that space across the structure is likely densely utilized.

Following the initial state, new states are consecutively generated. Each new state is generated by doing a random order pass through all edges of the structure. At each edge, a single crossover is added randomly at any location that satisfies the spacing constraints of the same helix spacing only (Fig. 3.14). As this continues and the crossover network becomes denser, this operation will sometimes conflict with choices on other edges according to the common helix spacing constraints. In this situation, the conflicting crossover on the inactive edge is removed (Fig. 3.16).

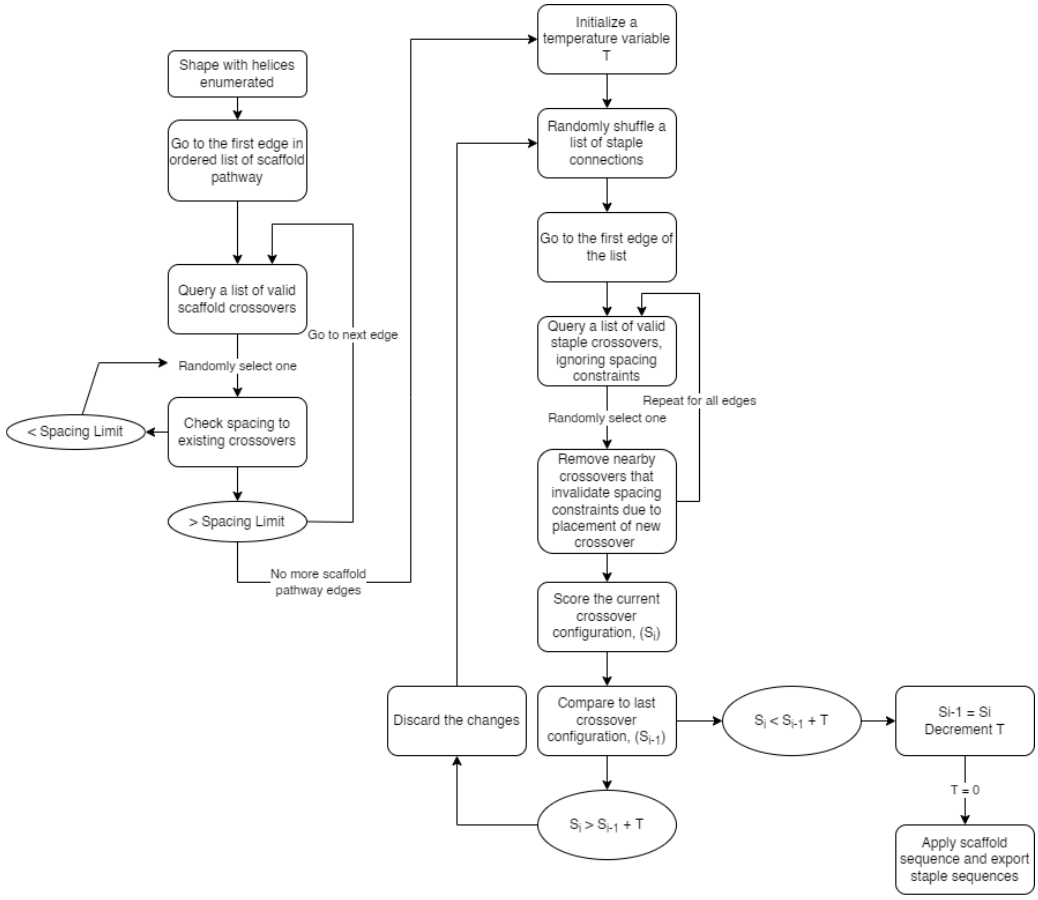

FIGURE 3.10: Flowchart of simulated annealing DNA origami heuristic design algorithm.

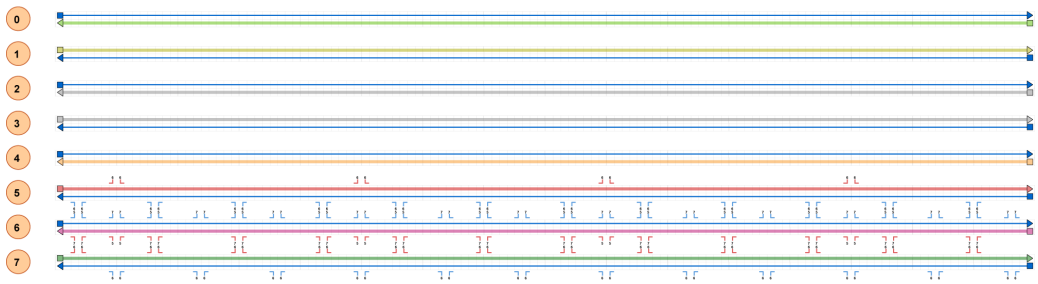

FIGURE 3.11: Automated DNA origami design, Step 1. Both scaffold and staple strands are added to fill the shape template. Potential crossover positions are discovered in a preprocessing step. For example, positions are marked between the bottom two rows.

This is the critical step to differentiate this process from the greedy decision-making process.

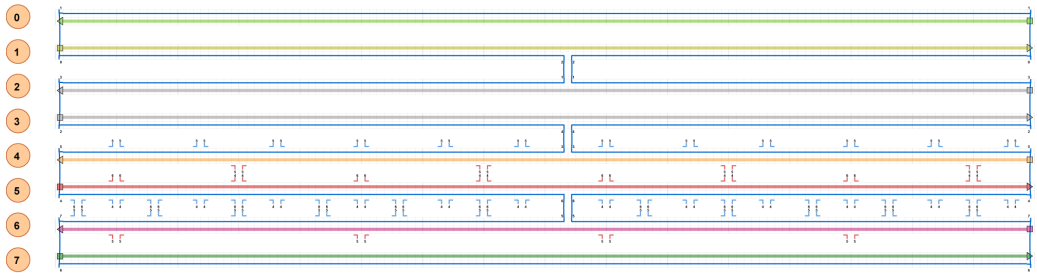

FIGURE 3.12: Automated DNA origami design, Step 2. Scaffold crossovers are added to complete a cyclic routing of a single scaffold strand through the entire shape template.

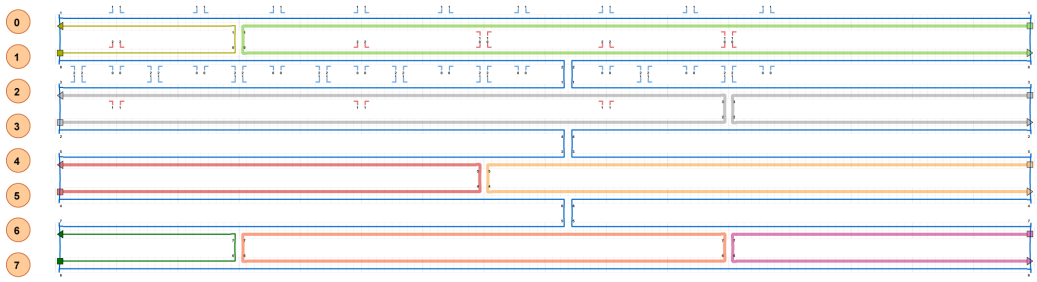

FIGURE 3.13: Automated DNA origami design, Step 3. An initial but relatively sparse selection of staple crossovers is added.

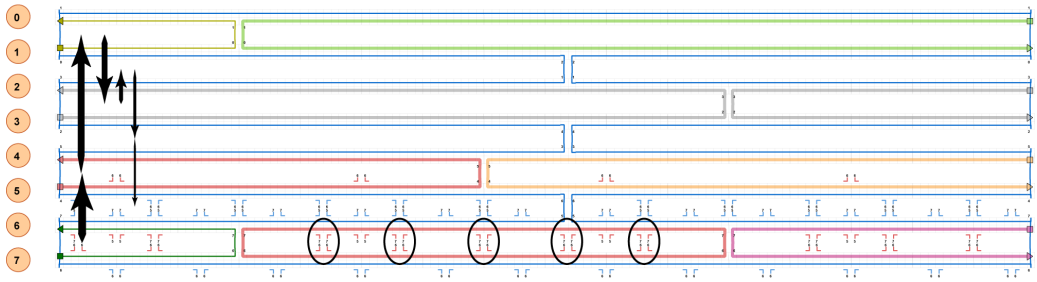

FIGURE 3.14: Automated DNA origami design, Step 4. The simulated annealing process (also see Fig. 3.10) has two major steps: (i) Traversing all edges in random order constitutes a single pass and generates a single new configuration of the structure. (ii) A single region is observed at each edge, and a valid crossover position within that region is randomly chosen and applied.

Each generated state is scored, and the system moves to that state if that score is valid within the temperature offset. Due to the processing time of determining only a single state, to reduce the overall runtime of the algorithm, the valid states are

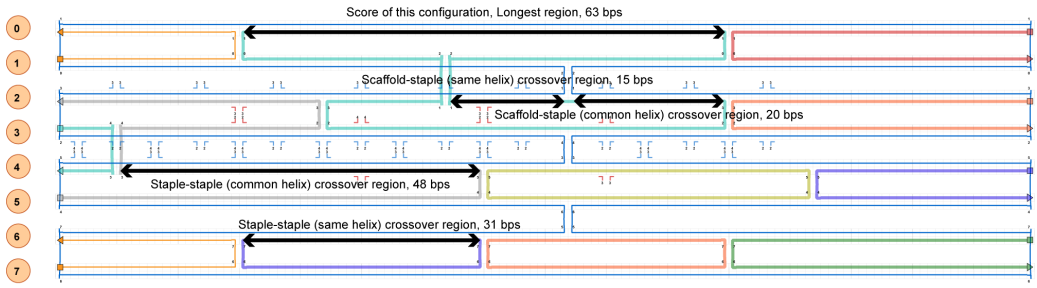

FIGURE 3.15: Automated DNA origami design, Scoring. Following the generation of each configuration in the simulated annealing process, it is scored according to the largest, uninterrupted region in the structure. For example, it is labeled in this figure as 63 bps. Only same helix regions, either between a staple and scaffold crossover or between two staple crossovers, are evaluated. Common helix regions are labeled here to illustrate the differences.

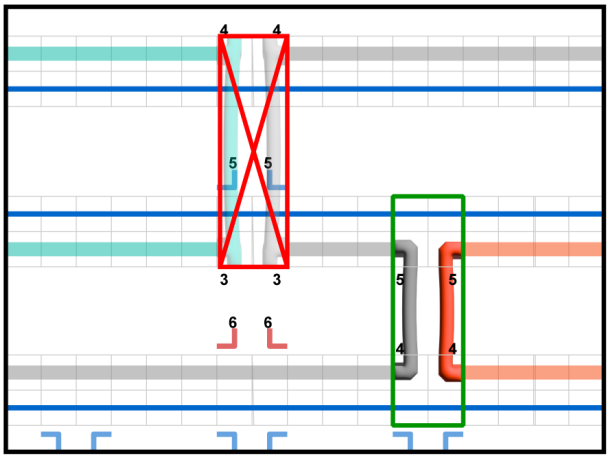

FIGURE 3.16: Automated DNA origami design, Backtracking. A backtracking feature is necessary to generate worse configurations from previous states to satisfy the implementation of a simulated annealing algorithm. Backtracking occurs when an added crossover (green) is inserted at a distance from a nearby crossover (red) that violates a common helix spacing constraint. The older crossover is removed, and the recent insertion is retained in the new configuration. This can generate configurations with worse scores and assist the simulated annealing algorithm in escaping local extrema.

accepted with a probability of 1. The process repeats iteratively as the score goes down until it reaches a halting threshold.

3.6.4 Evaluating the simulated annealing strategy on simple DNA origami nanostructures

The algorithm thus allows us to rapidly generate and evaluate differences in DNA origami design according to a set of consolidated parameters. To test its capabilities, two basic test structures were provided to the automation software and compared to manually drawn designs. First, we make comparisons in a 2D structure (see Figure 3.17). This starkly shows that while crossover density is not a reliable input parameter, it is still a characteristic quality of DNA origami nanostructures that is closely related to their stability and shape correctness. While parameters could be set for the generated output to be a close approximation of the shape, the regularity of crossover placement visible in a human-designed shape is still not present. In addition, the automated output shows an artifact of an error in the top left corner where a crossover may have been placed too far, showing that there are downsides to the algorithm not operating with enough information about the global context or even just observing a wider spatial context in the structure.

On the other hand, the automated design shows less global stress. This is likely due to a lack of patterning and the uniform application of strain that comes with it. Instead, irregular placement may provide more space along the structure to dissipate global stress and lessen deformations.

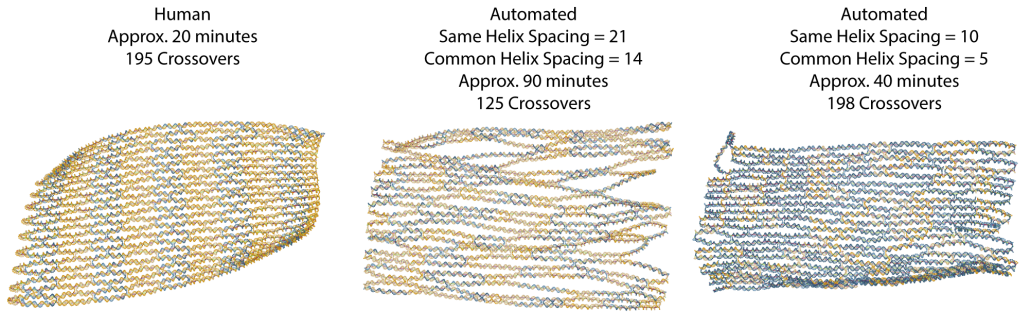

FIGURE 3.17: A comparison of a 2D flat rectangles designed on square lattices.

Next, the automation software was extended to 3D structures. No modifications to the principles are necessary, but for a 3D shape template, the differences in the staple and scaffold connectivity graphs become apparent (see Fig. 3.18). Crossover and nicking subprocesses only depend on the relative angle between a pair of helices. Thus, it is independent of the overall geometry of the structure.

A 3-by-2 honeycomb lattice helix bundle was added as a shape template. Different spacing constraints and thresholds were attempted to generate variations of the structure, as shown in Figure 3.19. Again, we notice the relationship of more stringent spacing settings to lower crossover counts, which in turn produce more poorly formed nanostructures. This example also illustrates the effect of an increased threshold (2.8 nm and 3.0 nm), which may detect more non-ideal crossovers, but provides the heuristic process with more choices and can improve results.

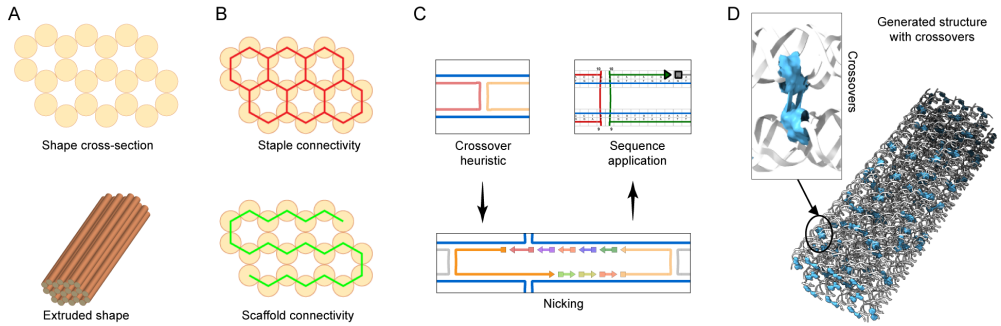

FIGURE 3.18: Overview of heuristic process on hex lattice. **(A)** Create the shape template. **(B)** Designate the staple and scaffold adjacency. **(C)** Add crossovers by the simulated annealing heuristic crossover placement algorithm. Add nicks by the split-and-merging algorithm. Then apply the scaffold sequence. **(D)** Generate the structure for simulation or export staple sequences for synthesis.

These designs can now be produced with little interaction with a human designer and can be easily refreshed under different parameters. Furthermore, by relying on computation, constraints can be evaluated and followed more strictly instead of intuitively or by tedious counting by a human designer. The examples demonstrate an iteration from low spacing constraints (SHS = 7, CHS = 7), which are expected to

generate dense crossover networks, up to higher spacing constraints (SHS = 21, CHS = 14). Example values are chosen somewhat arbitrarily but often at similar intervals to discrete turn positions of B-form DNA arranged in a honeycomb or square lattices.

Simulations for the 2D and 3D examples were performed using the oxDNA2 model [63, 83] at 500 mM Na salt concentration, an environment temperature of 300K, and in three phases. The first phase relaxes the model to the expected B-form DNA model. The second phase performs a molecular dynamics simulation for 10^6 steps, arriving at a relaxed conformation of the nanostructure. Finally, the molecular dynamics simulation is repeated with the same settings but for 10^8 steps, saving a snapshot of the conformation once every 10^5 steps, thus sampling 1000 data points to observe how the structure may fluctuate in reality. A geometric mean is calculated to find both the mean structure and the RMSF of each structure.

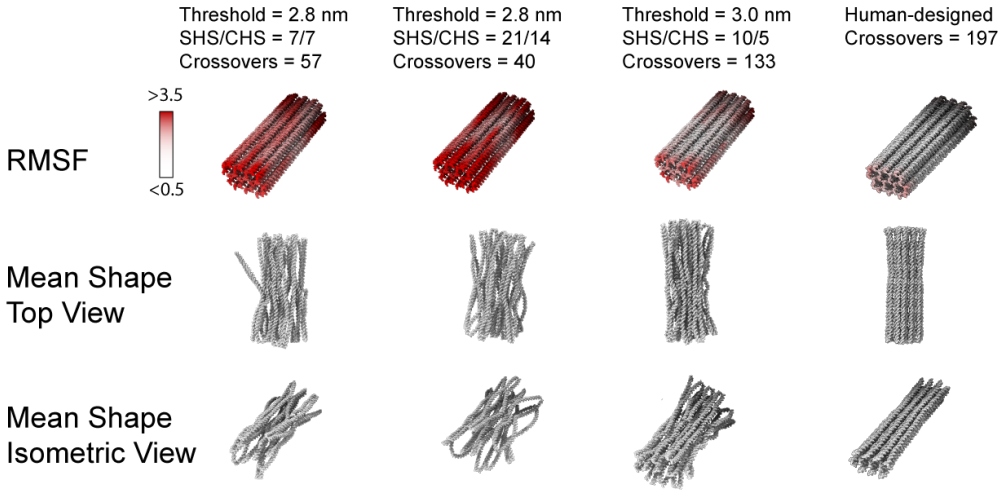

FIGURE 3.19: A comparison of a 3D hexagonal lattice structure generated using different automated parameters and by manual placement of crossovers and nicks by a human designer.

3.7 Conclusions

This chapter described the design of DNA origami as performed by a human designer. The manual process inspired a conversion to how a computer could replicate this process. Heuristic algorithms based on greedy decision-making and simulated annealing were developed and compared, and their functionality was demonstrated on a simple DNA nanostructure. This automation of a generalized design process for DNA origami allows the unattended design of DNA origami while following exact design principles and constraints. Its crossover discovery strategy is independent of geometry and is generally applicable. However, a notable rate-limiting step of the process is generating different variations of a crossover network. This limits the power of the heuristic to search a broader spectrum of the state space to find better solutions. Further work should consider data structures and algorithms that can speed up rate-limiting steps before improving the complexity of the heuristic. In addition, it can consider probabilistic weights to its stochastic processes that would more closely resemble a human designer's interpretation of the surrounding context when making crossover or nick placement decisions.

4

Multi-layer Enclosed and Curved DNA Nanostructures

4.1 Introduction

This chapter discusses applying the generalized DNA origami design principles discussed in Chapter 3 to a class of curved, encapsulating 3D DNA nanostructures, also known as DNA capsules, as defined in Chapter 1. Many 3D DNA origami shapes can be designed in such a way as to serve as a separating barrier between encapsulated internal moieties and an external environment, as discussed in their applications in Chapter 2. These applications often demand rigid, hollowed structures or, in other words, capsule-like functionality. Thus far, DNA nanostructures tailored towards interdisciplinary applications remain constrained to a relatively limited variety of forms; many are still based on the packed lattice strategies introduced in Chapter 3, discussed in Section 3.3.1.

Earlier work proposed a contrasting strategy to the packed lattice strategy based on curved DNA helices, which were better suited for developing a wide variety of capsule-like structures [7, 10]. Curvature can provide a finer-grained discretization

The work described in this chapter includes excerpts from Fu et al. (2022) [147].

of addressable locations and geometries [38, 50, 125] and create enclosed shapes with larger compartments more economically than block-like designs [45, 61]. Curved motifs also align more closely with naturally occurring, globular molecules [26] while still maintaining an enclosed space that is suitable for consolidating molecular activity.

However, while there are evident benefits of DNA nanostructures that are stable, curved, and fully enclosed and which can be synthesized at practical yields, the high complexity of designing these structures has hindered their accessibility to researchers. Furthermore, current design principles for enclosed, curved DNA nanostructures have been demonstrated for only single-layer nanostructures, which limits their achievable rigidity and applications. Curved DNA nanostructures, while geometrically more complex than packed lattice structures, still share many of the same design principles. Except, packed lattice structures have straightforward patterning, and the problem of solving their entire crossover network can be easily reduced to a smaller, localized instance. These simplifications do not exist for curved DNA nanostructures. A human designer needs to meticulously decide the crossover network across the full and daunting scale of the nanostructure. Thus, these curved nanostructures could be an interesting target to apply and evaluate the generalized DNA origami heuristic design algorithms that have been previously presented in this thesis.

As part of this thesis, a dedicated software package was developed to expand the design space of enclosed, curved DNA nanostructures and automate the escalation in design tedium relative to packed lattice structures. A brief description of DNAxiS (CAD for **DNA** Nanostructures with **Axi**al **S**ymmetry) follows. The software implements novel multi-layer design strategies [56, 64, 65, 146] specifically for curved DNA nanostructures. We refer to this as *reinforcement*, which increases overall yield and stability by only targeted application of multi-layer design principles to segments in the structure. DNAxiS is the first CAD software tool to implement probabilistic

algorithms to perform crossover selection, as well as the first for the design of DNA capsules, with a specific focus on those with axial symmetry. DNAxiS has since been made available to the public as a stand-alone, open-source software with more information and an online version available at http://caddna.cs.duke.edu (Fig. 4.1).

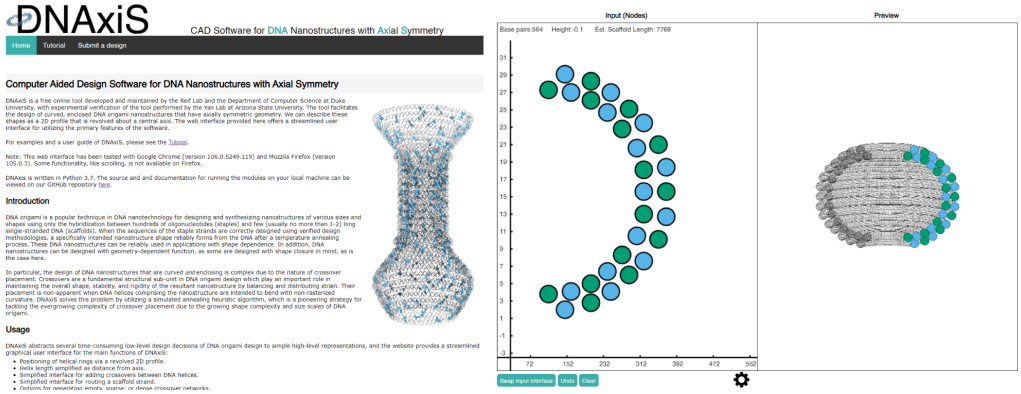

FIGURE 4.1: Preview of DNAxiS's online graphical user interface. User interfaces for supplying a shape template, CAD model, connectivity graphs, and controlling input parameters are packaged and provided in a graphical user interface available online. The homepage and a sample of the graphical user interface for drawing the shape template are shown here.

4.2 Results

4.2.1 Design principles of curved, enclosed DNA origami nanostructures

We first discuss the application of the previously discussed heuristics towards curved DNA nanostructures by first reviewing their design principles, and thus, their differences compared to straight, packed lattice designs. First of all, in general, DNA origami nanostructures are synthesized by annealing in one-pot reactions using a large set of short oligonucleotides (<100 bps) called staples to "fold" a long (>1000 bps) single-stranded DNA strand termed the scaffold into a desired physical conformation [4]. Sub-sequences of a staple strand bind to corresponding complementary sequences upon the scaffold strand, and where a continuous staple sequence binds between discontinuous sequences on the scaffold will produce a crossover on the

staple DNA to link the two discontinuous domains of the scaffold. Therefore, the particular placement and distribution of crossovers across the entire sequence (the crossover network) play an important role in maintaining the overall shape, stability, and rigidity of the resultant nanostructure by balancing and distributing strain.

Curved DNA origami nanostructures are created by applying those fundamental techniques to a template based on arranging helices as circles. The ring module as presented in Figure 3.8 is primarily utilized here. Each ring and its adjacent rings have varying z-axis placement and circumference, thus forming rounded, extruded structures with non-rasterized xy-plane curvature. The full, enclosed geometry is designed by calculating circumferences of adjacent pairs of rings to reduce gaps between them down to the interhelical distance ($\delta = 2.6$ nm). Furthermore, circumferences are rounded to an integer number of DNA full turns and to the closest multiple of a crossover factor, choosing one of 3, 4, or 5 consistently for the entire structure; the chosen value is a common divisor of the number of crossovers that are formed between each pair of nearest neighbor rings, and by doing so, promotes symmetry and regularity in the design. When rings are planar, this corresponds to 48 or 50 bps.

Table 4.1: Minimum and maximum circumferences per crossover factor at upper and lower limits of spacing and twist.

Crossover factor	Minimum circumference Spacing: 2 full turns Twist: 9 bps/turn	Maximum circumference Spacing: 2 full turns Twist: 12 bps/turn
3	54	180
4	72	240
5	90	300
6	108	360
8	144	480
9	162	540
10	180	600

Table 4.2: Expected circumferences per crossover factor at nominal twist and spacings between 2 to 5 full turns.

Crossover factor	Circumference Spacing: 2 full turns Twist: 10.5 bps/turn	Circumference Spacing: 2 full turns Twist: 10.5 bps/turn	Circumference Spacing: 4 full turns Twist: 10.5 bps/turn	Circumference Spacing: 5 full turns Twist: 10.5 bps/turn
3	63	93	126	156
4	84	124	168	208
5	105	155	210	260
6	126	186	252	312
8	168	248	336	416
9	189	279	378	468
10	210	310	420	520

Due to these differences in circumference throughout the structure, the nominal twist of B-form DNA cannot always be preserved. As specified by previous work [10], placing crossovers such that all sections of the DNA helices are between 9 bps/turn to 12 bps/turn will help to maintain a high yield of the intended structure (approximately $> 90\%$). Furthermore, the crossover network is interspersed such that adjacent crossovers are not set too close, which can concentrate strain upon that segment of the structure and lower yield, or not too far, which may not provide enough strain to maintain curvature against the *persistence length* of DNA. The persistence length is the length of DNA under which it can be expected to behave like a beam, around 50 nm. If a DNA strand is anchored at points beyond that persistence length, then the segments in between may begin to adopt unpredictable, flimsier geometries. This constraint limits the crossover spacing to be preferably between 2 to 5 full turns of DNA for counting between adjacent crossovers spanning the same pair of helices. (Example circumferences resulting from variations in these parameters are shown in Tables 4.1 and 4.2).

Moreover, creating curvature on short double helices will result in a higher strained conformation of dsDNA, which can cause a higher number of staples to bind incorrectly, thus lowering the overall yield of properly formed DNA nanostructures. This

is more strongly expressed in curved nanostructures due to the presence of each staple being more significant towards preserving local curvature. Multi-layer designs not only provide redundancy in staple density to overcome binding errors but are also expected to be more mechanically rigid [56,64], while careful design choices and synthesis conditions can further promote the formation of well-formed structures.

4.2.2 Overview of DNAxiS software and challenges to curved DNA nanostructure design

DNAxiS is the first software tool to implement a general algorithm for designing curved DNA nanostructures respecting the above constraints while extending the design space to accommodate multi-layer designs. The lattice-free nature of helix placement and crossover placement of curved DNA origami design distinguishes it from the capabilities of previous design algorithms and interfaces [5,24,32,73,75,148]. The design of multi-layer, curved nanostructures faces unique challenges primarily emerging from the presence of continuous, curved helices, as opposed to straight or disjointed helices. Without a strong basis for algorithmic regularity or crossover periodicity arising from straight helices, DNAxiS instead applies heuristic algorithms that closely mimic concepts and approaches applied by a human designer when routing a scaffold strand or distributing a crossover network in DNA origami nanostructures. The design of these heuristic algorithms is as described in Chapter 3.

The overall procedure for designing DNA capsules is similar to those shown for 2D and 3D packed lattice structures (Figs. 3.18 and Figs. 3.11-3.16). Its adaptation to DNA capsules is illustrated in Figure 4.3, and an overview of the crossover placement algorithms is described later. An additional feature of the DNAxiS package is the higher-level abstraction of drawing the shape template to simply drawing the goal shape. In summary, the process begins by specifying a shape via an STL file externally designed in the user's 3D design software of choice (e.g. SolidWorks,

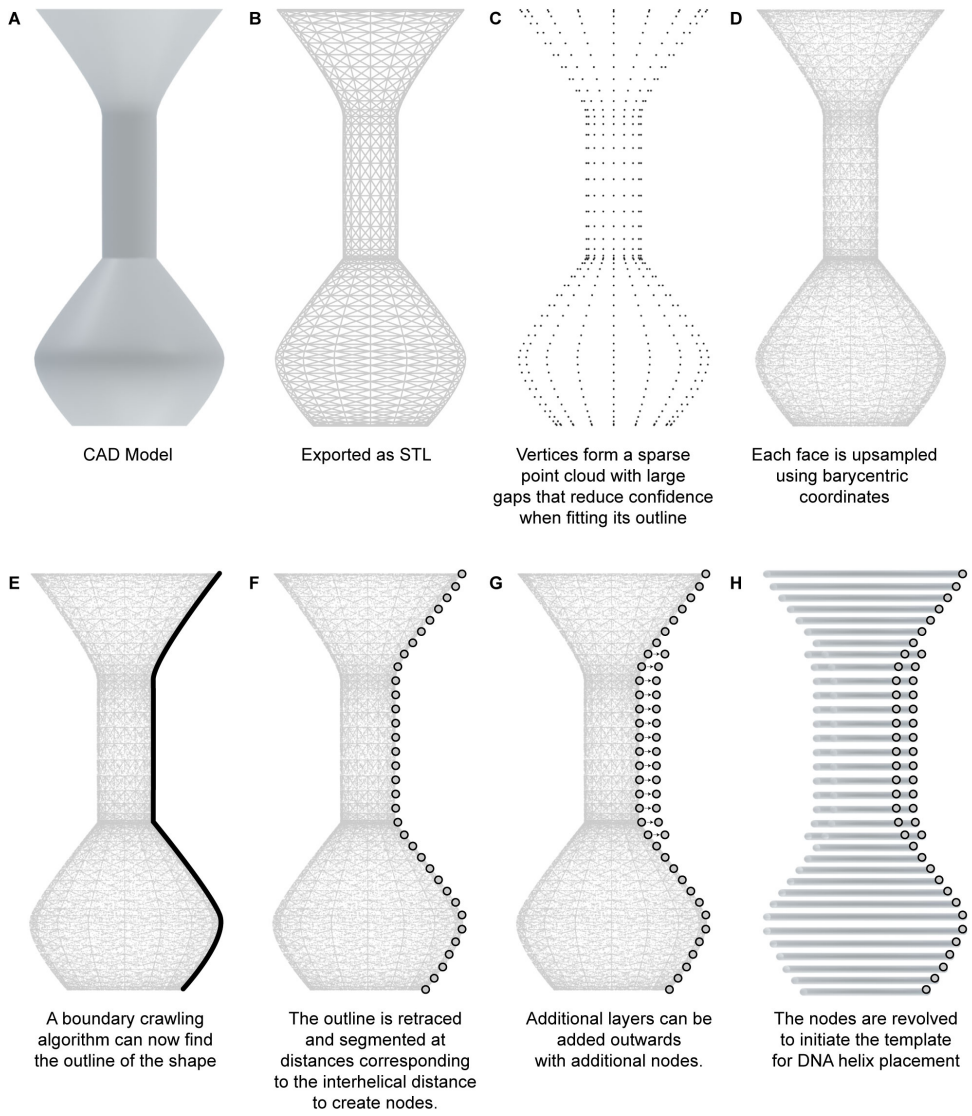

FIGURE 4.2: STL parsing for axially symmetric geometries. **(A-H)** The sequential description of the STL parsing process. This feature is provided primarily to interface with interdisciplinary research goals, where desired geometries may already be defined as separate CAD models. This feature can extract the corresponding shape to be applied to the DNA nanostructure design to exactly match existing workflows.

Fusion 360, Blender, etc.) (Fig 4.2A). The model is processed to extract a shape that is discretized into a ring-based mesh (Fig 4.2B-F), where each ring represents the placement of a circular DNA helix. This produces the shape template (Fig. 4.3A). Additional rings can be manually placed, which can be used to indicate an

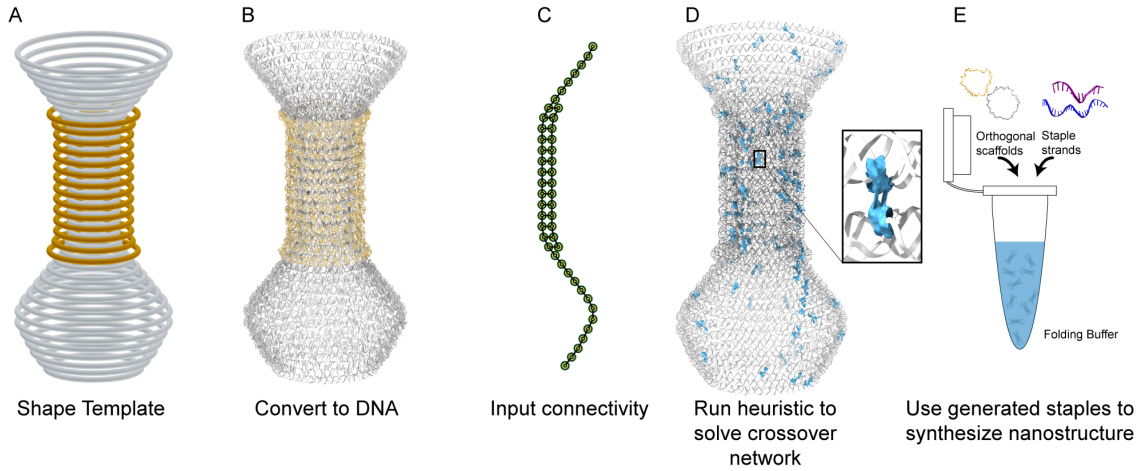

FIGURE 4.3: Overview of the DNAxiS design process. **(A)** Begin with a shape template constructed of circles rendering a shape. This template can be made selectively multi-layer by adding rings outwards from the starting mesh. **(B)** A helical twist is calculated from the circumference and used to convert each circle of the mesh into a DNA helix ring. **(C)** Scaffold and staple connectivity are defined by input from a human designer. **(D)** Crossovers are densely applied to the template using a heuristic algorithm. Following this step, nicks are added to modulate the length of all staple strands to a favorable length. **(E)** Generated staple sequences are annealed with the corresponding scaffolds, sometimes multiple orthogonal sequences as needed, in a one-pot reaction to produce DNA nanostructures of the designed shape.

expansion of the mesh to a multi-layer design (Fig. 4.2G). The user also helps to define the scaffold routing and nearest neighbors (Fig. 4.3C, 4.12). The total mesh, replaced by DNA helices populated with nucleotides (Fig 4.3B), is referred to as just the *template*. The template specifies the position of each helix and its nucleotides in a coordinate space. The crossover placement algorithms are then applied upon the template to connect the helices together via crossovers (Fig 4.3D), finally determining a dense enough crossover network that can be exported towards synthesis (Fig 4.3E).

A major challenge for moving into a curved design space is that crossover positions no longer appear with a small periodicity in curved DNA nanostructures as they do with straight helices, appearing once every several turns instead of once every one turn (Fig 4.6). Some existing methods of 3D DNA origami design form

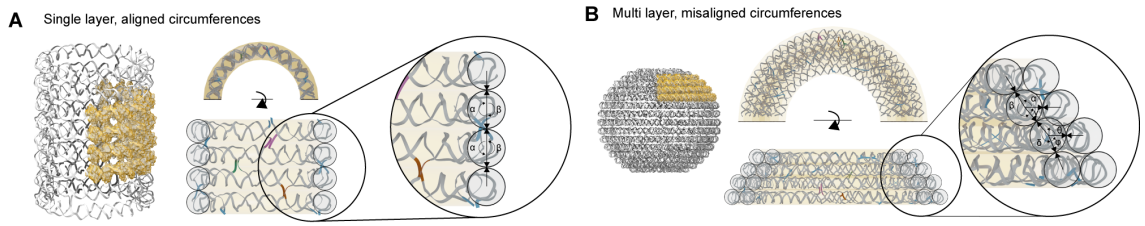

FIGURE 4.4: Alignment principles for single/multi-layer, curved, closed-shell DNA nanostructures. **(A)** As a single-layer structure without varying circumferences, the crossover network of a cylinder is no more complex than a flat DNA origami rectangle. Local patterns of a few crossovers can be repeated globally, as the dihedral angle and crossover periodicity are the same for all pairs of nearest neighbor helices. **(B)** Extension of design principles to multi-layer, curved, closed-shell DNA nanostructures significantly complicates the design space for determining valid crossovers. Each crossover pattern between pairs of nearest neighbor helices is unique due to a different dihedral angle and varying helical twists. Rather than repeating a simple local pattern, each crossover, typically up to 200 when fully utilizing a single m13mp18 scaffold, must be carefully positioned.

encapsulating nanostructures using rectilinear designs [8, 45, 47, 49, 50]. Others can estimate curvature by cyclically linking short, straight sections by single-stranded corners [148]. The straight edges in these design methodologies do not have problems with low crossover periodicity, but this can either introduce gaps or undersample the desired curvature. The irregular circumference of curved DNA nanostructures also obfuscates crossover positions, especially for a human designer. Curvature in the most basic curved structures, such as a cylinder, can still exploit the global repetition of local crossover patterns in a similar fashion to DNA origami that is based on straight, anti-parallel helices. As shown in Figure 4.4, a shift to multi-layer, curved structures from single-layer, vertically aligned patterns expand the problem space as each pair of adjacent helices may require a unique crossover pattern. This is the most significant departure from previous DNA design strategies and their automated design methodologies. Furthermore, these irregularities often force a designer to choose non-ideal crossover positions, which additionally contribute to strain on

the structure. Considering crossovers in a rated and ranked order can offset damaging over-selection of crossovers that have high contributions to the overall sum of strain. This problem is also more prevalent in multi-layer structures, where each helix connects to more nearest neighbors and further complicates the placement of adjacent crossovers within a single helix due to reducing the average spacing between adjacent crossovers. Lower spacing shortens the length of continuous regions of the complementary staple to scaffold sequences. Then, weaker binding of each staple strand reduces yields.

Applying the crossover network heuristic optimization algorithms

One major difference with applying the previously discussed methods from Chapter 3 is the much larger period of aligned crossover positions between two helices. Potential crossover positions between a pair of helices are first determined by looking for the collinearity of a bisecting vector through the dihedral angles between any pair of adjacent nucleotides (see Fig 4.5 for a rendition of this when applied to 3D curvature). Then, all pairwise combinations of crossover positions are evaluated for their alignment, which is the expected distance that a backbone bond between crossover nucleotides must overcome to form the crossover. Impractical crossover positions are eliminated above a threshold for their alignment, indicating that the two points on each helix are too far apart. Except, this distance should now be set much higher to permit the presence of crossovers with a little bit of skew away from one another. As seen in Figure 4.6, due to the differences in periodicity each helix can exhibit, strict filtering of crossovers for only aligned crossovers would severely limit the available crossovers choices. The accepted set of crossover positions is then ranked by alignment. Again, this choice faces a compromise. While worse alignments are believed to lead to worse yields, so does an insufficiently dense crossover network. Secondly, crossovers are ranked by their base pair distance from another crossover. The leading

crossover of this list is then popped and applied in an order depending on which of the two heuristic algorithms (for greedy, see Section 3.6.2; for simulated annealing, see Section 3.6.3) are applied.

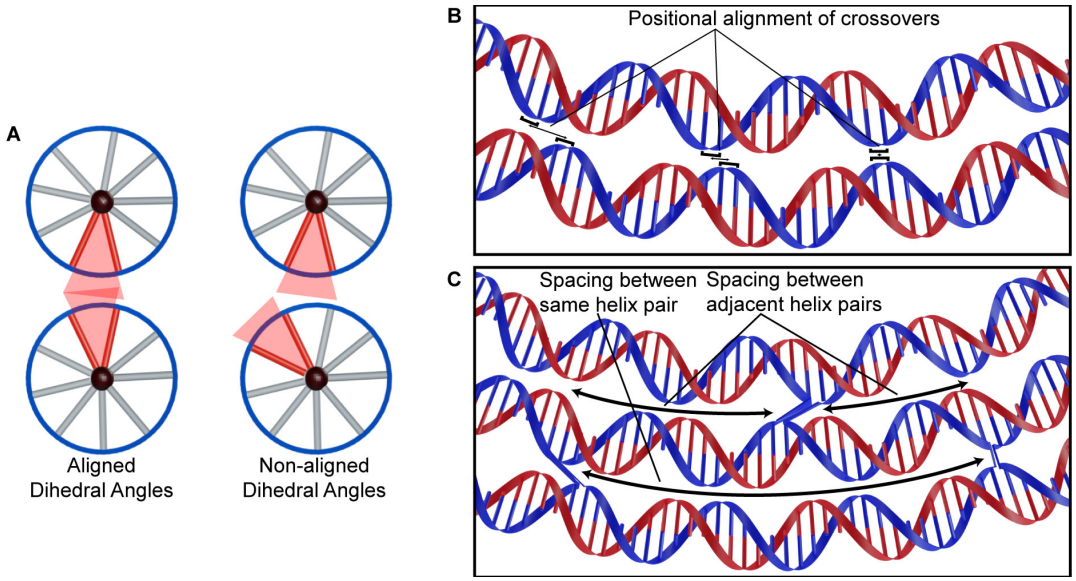

FIGURE 4.5: Crossover discovery model on curved nanostructures. All crossover placement primarily depends on three parameters defining the spatial context of crossovers. **(A)** The dihedral angle is a measure of the direction of a crossover. The area formed by vectors extended from two sequential nucleotides should intersect with the central axis of the target helix to be considered a valid dihedral angle to make a crossover. **(B)** The crossover alignment represents the length of a crossover bond. Each endpoint is the midpoint of a pair of sequential nucleotides that would form the crossover which lies on the central axis of the helix. The alignment is suppressed to fall within a set threshold to minimize the global stress of the structure. **(C)** The crossover spacing is divided into two sub-categories for crossovers that span the same pair of helices and crossovers that touch a common helix. While crossovers that span the same helices at most appear once every full turn, depending on the position of adjacent helices, a crossover towards an adjacent helix can potentially interject anywhere along the helix, and this spacing can be adjusted to ease design constraints or strictly improve strain as much as possible. By default, DNAxiS is set to default parameters that have so far been consistently successful (crossover alignment = 3.0 nm, SHS = 10, CHS = 7).

Here, we more closely analyze the simulated annealing algorithm by investigating the step-wise progression of the algorithm upon a vase shape template. Following the

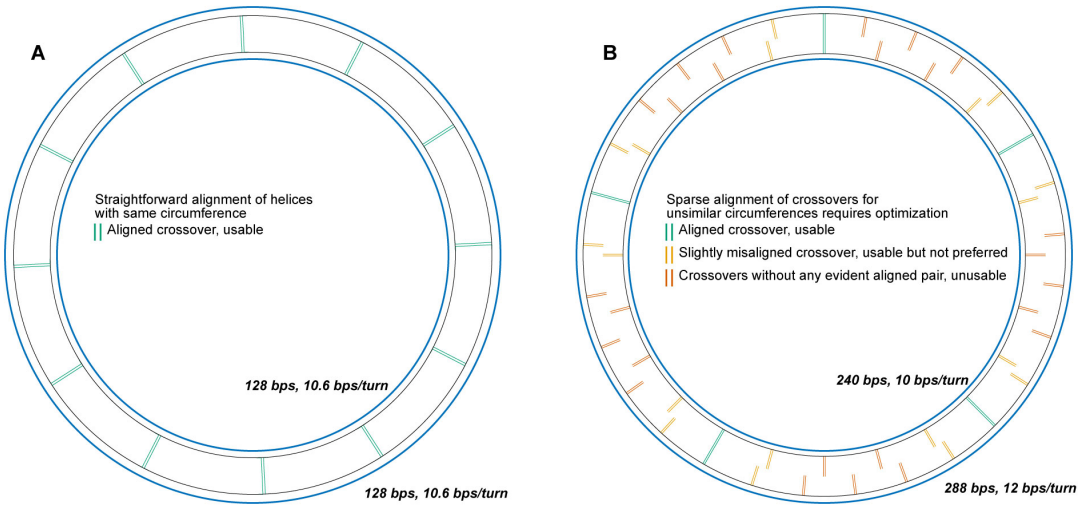

FIGURE 4.6: Crossover alignment challenges when curvature introduces non-homogeneous circumference and helical twist. **(A)** The easiest solution is for cylindrical shapes that do not vary in their circumference or twist. Even if the circumference is small, crossover availability is straightforward and plentiful. **(B)** The most complex solutions are where curvature may cause all pairs of adjacent helices to have different circumferences and helical twists, and thus their crossover positions occur at differing periods. These periods are well-aligned at much fewer positions, even for large rings, and misaligned crossovers must oftentimes also be used in order to meet a minimum crossover count.

graph in Figure 4.7, we track two characteristics, the *crossover spacing* as the scored variable while the *crossover quantity* is displayed as a value that shows an inversely correlated change to the crossover spacing. Each simulated annealing process begins from a temperature $T = 30$. The first iterations proceed very smoothly, as the structure is typically initialized with a sparse crossover network. After about ten iterations, random permutations of the current state become less likely to discover more compact and uniform assignments of crossovers. Even as the crossover spacing remains constant, there are small shifts in the crossover quantity. The number of rejected states increases. In the interest of running times, at low temperatures, the number of failed states to attempt is capped, as the solution to the crossover network

may already have found an approximate solution, and thus it would not accept any changes to a new state. Overall, this method is much slower than the greedy variation, taking up to hours for large structures (Fig 4.8), but can implement denser, more uniform crossover networks, and despite longer wait times, it still requires only minimal effort from the human designer.

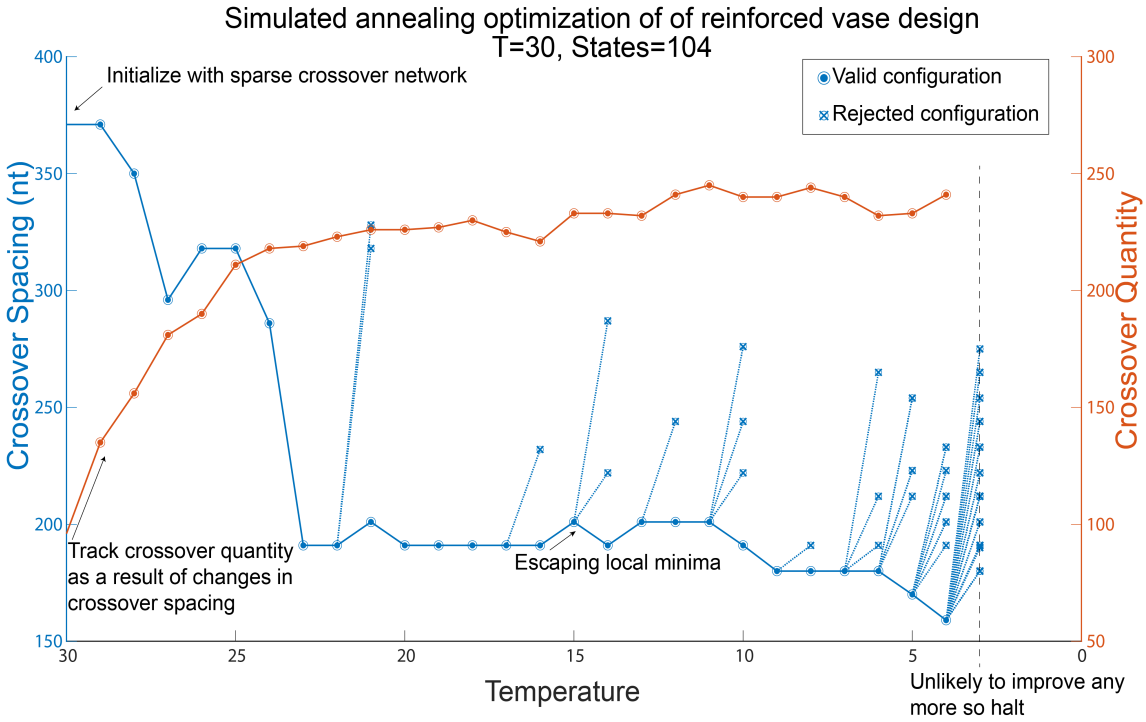

FIGURE 4.7: Example trace of crossover network design when applying simulated annealing algorithms. The process is initialized with a sparse set of crossovers and a high temperature. At each temperature, the simulated annealing algorithm grows the set of crossovers by adding a single randomly chosen crossover between each pair of nearest neighbor helices, taken in random order. Each configuration is scored by its crossover spacing. If it is less than the score of the previous state with an allowance equal to the temperature ($S_i = S_{i-1} + T$), then the algorithm moves to that state while also decrementing the temperature. If the crossover spacing is invalid (dotted lines), the attempted random state is discarded, the temperature remains the same, and a new random configuration is created. This process is repeated until successful. The algorithm terminates if the temperature reaches zero, or if the temperature is low and the algorithm successively fails more than a user set maximum, the algorithm also terminates.

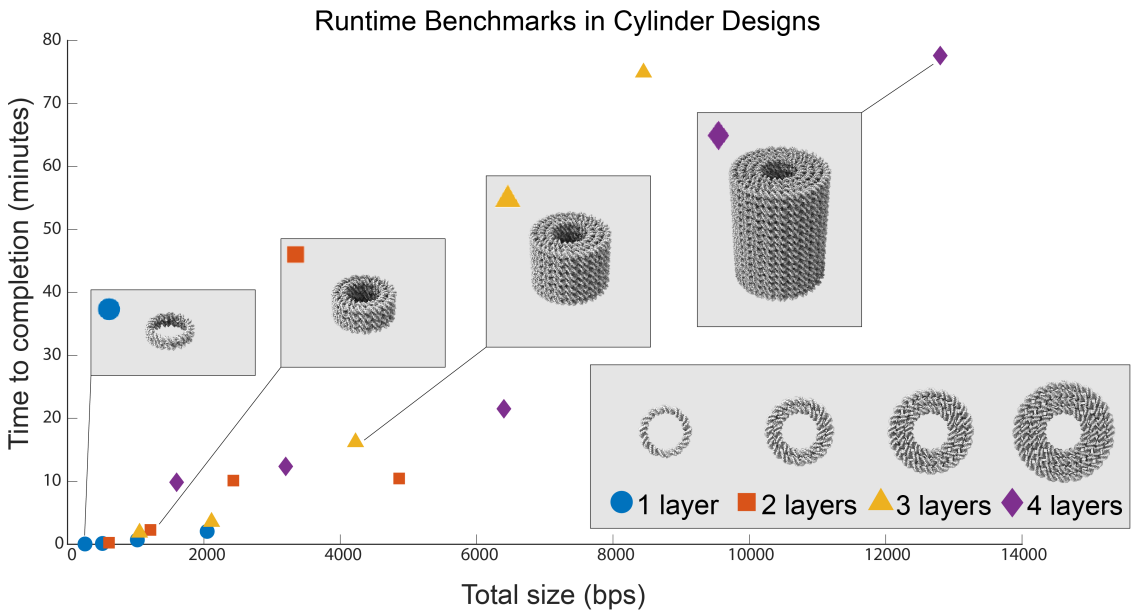

FIGURE 4.8: Runtime benchmarks for sample structures. Runtimes were measured for inputs across several size scales of cylinders in terms of their numbers of layers, height, and total base pair size (256 bps – 12800 bps). Structures at the high end of the design space (DNAxiS currently supports structures of at most 13,450 bps from orthogonal scaffolds p8064 + phi X 174) are expected to take 1-2 hours. Software settings for each structure are consistent: crossover threshold = 3.0, same helix spacing = 10, and common helix spacing = 7.

Details on additional functionality provided for assisted design in DNAxiS

STL processing - STL processing (Fig. 4.2) is provided as an alternative input method to the direct input of helix circumferences and positions. STL files that describe a 3D shape using triangles can be uploaded in the user interface and are imported as a list of triangular faces, each defined by its three vertex positions. All vertices are then used to form a point cloud of the structure. The point cloud is upsampled to accommodate low-resolution STL models that may have few and large triangular faces. This is done by randomly adding several points to each face proportional to its area. A sub-process performs this by defining each face in barycentric coordinates by denoting vectors from two adjacent edges of a vertex and assigning

random variables X_1 and X_2, where $0 < X_1 + X_2 < 1$, to each of those vectors. Projecting the point cloud to 2D, a concave hull algorithm based on Delaunay triangulation finds a tight-fitting line boundary to the input STL. Iterating across the line at intervals equal to the set interhelical distance (default = 2.6) provides the node positions. Each node's distance from the central axis is converted to a base pair circumference.

Split and merge nicking algorithm - After a crossover network is generated, to complete the design of a DNA origami nanostructure, staple helices across the structure are nicked down into the desired staple length range. To optimize synthesis, several constraints are prioritized but not strictly enforced. Initially, locations too close to crossover locations are protected; nicks cannot be placed in that region. This can be adjusted as an option by the user but is recommended to be no less than 7 bps. The remaining unprotected strand regions are then nicked into 3-7 nt fragments. Following this, the shortest fragment that exists at any instance is joined to its shortest adjacent fragment. This process is repeated across the design until no fragments can be joined without surpassing the maximum staple length.

Automated helical twist adjustment of rings - To simplify user input, the determination of helical geometry parameters is automated and abstracted away from the end user, such that a user is only concerned with the size and position of each helix, not whether their helical geometries are within stable bounds. Rings are scaled to an appropriate circumference corresponding to expected helical geometries. These include the following and help create periodicity, symmetry, and stability for crossover placements.

Adjustment of helix twist to circumference - These adjustments help determine the helical twist in base pairs per turn (bps/turn) and are iteratively found (described by the below pseudocode algorithm) until a ring satisfies a twist between 9 and 12 bps/turn.

```
assume ring is unstable // (not 9 < bps/turn < 12)
while unstable do
    for Turns between crossovers from min to max // default 2 to 5 do
        if 9 < bps/turn < 12 then
            stable
        else if unstable and bps/turn > 12 then
            crossover factor multiple ++
        else
            circumference += crossover factor * crossover factor multiple
        end if
    end for
end while
```

4.2.3 Materials and Methods

Before evaluating the experimental characterization of designed structures, we first provide an overview of experimental protocols used to produce the data discussed in the following section.

Inputs

Bowl, gourd, sphere, cone, mushroom, ellipse, and clover shapes were designed with the initial algorithm using fixed crossover counts and crossover alignments up to 3.5. Vase shapes were generated using the simulated annealing algorithm with variable crossover counts and a consistent crossover alignment of 3.0. All helix circumferences were rounded to a crossover factor of 4.

Reagents

Staple strands (unpurified) of DNA origami were purchased from Integrated DNA Technologies (IDT) and used without any further purifications. High-purity scaffolds

(p7249 & p8064) were obtained from Tilibit Nanosystems (Germany). PhiX virion ssDNA scaffold was obtained from New England Biolabs (NEB).

Formation of DNA origami structures

The folding mixtures were prepared in-house: prepared 1XTAE-Mg^{2+} buffer (20 mM Tris base, 10 mM acetic acid, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 12.5-20 mM $Mg(OAc)_2$, pH 8). The reaction mixtures were folded in a thermocycler (Life Technologies SimpliAmp).

12 hour annealing protocol: Heat to 90°C for 5 min, jump to 86°C -5min and then decrease by 1°C/5min till 71°C, 70°C-15min and decrease by 1°C/15min to 40°C, 39°C-10min decrease by 1°C/10min to 26°C, 25°C-30min, jump to 20°C-15min, jump to 15°C-5min and jump to 10°C and maintained at that temperature.

24 hour annealing protocol: Heat to 90°C for 5 min, jump to 86°C-5min and then decrease by 1°C/5min till 76°C, 75°C-15min and decrease by 1°C/15min till 71°C, 70°C-20min decrease by 1°C/20min to 61°C, 60°C-30min decrease by 1°C/30min till 30°C, 29°C-20min decrease by 1°C/20min to 25°C, 24°C-15min decrease by 1°C/15min to 20°C, 19°C-10min decrease by 1°C/10min to 15°C, jump to 4°Cand pause maintained at that temperature.

37-hour annealing protocol: Heat to 80°C for 4 min and then decrease by 1°C/4min till 61°C, 60.5°C-30min and decrease by 0.5°C/30min till 34.5°C, 34°C-60min decrease by 1°C/60min to 24°C, and then pause maintained at that temperature.

48 hour annealing protocol: Heat to 90°C for 5 min, jump to 86°C-5min and then decrease by 1°C/5min till 81°C, 80°C-10min and decrease by 1°C/10min till 75°C, 74°C-30min decrease by 1°C/30min to 69°C, 68°C-40min decrease by 1°C/40min till 53°C, 52°C-60min decrease by 1°C/60min to 25°C, 24°C-80min decrease by 1°C/80min to 21°C, 20°C-30min,19°C/10min decrease by 1°C/10min to 15°C, jump to 4°Cand pause maintained at that temperature.

Gel electrophoresis of the DNA origami structures

Characterization of folded structures: Samples were run on a prestained-1.2% Agarose gel (with ethidium Bromide) made in 1XTAE -12.5mM $MgCl_2$ buffer. The running buffer was 1XTAE -12.5mM $MgCl_2$. 10µL of the annealed sample from the PCR along with 1µl of 10X loading dye was run in each of the wells of the agarose gel. The gels were run for 1-1.5 hours at a constant voltage of 80V at 4°C.

Purification of the folded structures

The fully annealed samples were run on a prestained-1.2% agarose gel (with ethidium Bromide) with 20µl (18µl of sample+2µl of loading dye) being loaded into each well and ran for 90 minutes at a constant voltage of 80V at 4°C. The running buffer was 1XTAE -12.5mM $MgCl_2$. Thereafter, the second lowest band, in some cases the band above, was cut out separately and put into a freeze and squeeze tube (the lowest band is the excess staples, so they were left out). Care was taken not to chop up further the bands cut out, which, if done, was realized to reduce the yield of the fully formed structures, thus the rationale being to load into smaller wells and take the bands out as they appeared. The freeze and squeeze tubes were left in the -20°C freezer for over an hour/two (longer periods, did not affect the recovery yield of the structures). After which, the frozen tubes were spun down at 1600rcf in a tabletop centrifuge for 45-50min at room temperature (Lower centrifugation speeds ensure that the assembled structures do not fall apart, and the longer times ensure maximum recovery from the gel pieces). The recovered solution was concentrated using 100kDa Amicon filters (pre-run with filtered 1XTAE-12.5mM $MgCl_2$ to make the membranes compatible) and spun at 1600rcf using a tabletop centrifuge.

Negative-stain TEM and image processing

Preparation: 5µl of the above purified sample was adsorbed on a commercially supplied formvar stabilized carbon type-B, 400mesh copper grids (Ted Pella, part number 01814-F) that was glow discharged for 1 min at 15mA using a Pelco easiGlow glow-discharge system (Ted Pella, Redding, CA, USA) and stained using 5µl of a freshly prepared 2% aqueous uranyl formate solution containing 25 mM sodium hydroxide (NaOH). Samples were incubated for 15 to 300 s depending on the concentration of the sample. Excess liquid was wicked away with Whatman filter paper and grids left to dry for 30-60 min prior to imaging. Images were acquired on a FEI Tecnai TF20 TEM microscope operated at 120 kV accelerating voltage using a charge-coupled device (CCD) camera at 50000x magnification. Particles were manually picked and class averaged using Relion3.0 without CTF correction.

Cryo-EM

Cryo-EM sample preparation: DNA nanostructures were formed and purified as above. The sample assembly was confirmed and further purified on agarose gels. To remove excess ethidium bromide in the sample buffer, the sample was dialyzed using a 10kDa Float-a-lyser against 1XTAE-12.5mM $MgCl_2$ and then concentrated as before. Structured DNA complexes were absorbed to a glow-discharged ultrathin carbon film supported by a lacey carbon film on a 400-mesh grid (Protochips, Morrisville, NC) for 1 minute with 8 repeated applications. Samples were applied to only one side of the carbon support with precaution to ensure that the grids were left a little wet between each application. Grids were directly plunge frozen into an ethane slurry at liquid nitrogen temperatures. Tilt series images of the samples were collected on a Titan Krios G2 (FEI/Thermofisher) and K2 summit (Gatan, Pleasanton, CA.) camera at 28,735 x magnification from -65 to +65° alpha with a 2° increment between images using batch methods in Serial EM, with a dose symmetric tomography scheme

[149,150]. The IMOD package [151] was used to reconstruct tomograms and visualize tilt series to determine the structures.

Individual particle videos

All images were collected at 14,000 x on the microscope at 300 KeV with a 1.74 Angstrom/pixel sampling rate in counting mode on a K2 summit direct electron detector. During tomographic reconstruction, images were down-sampled to improve contrast to 7 Angstroms/pixel. Scale bars will be 25 pixels, so 17.5 nm.

4.2.4 Experimental evaluation of curved, enclosed DNA nanostructures with crossover networks generated by heuristic algorithms

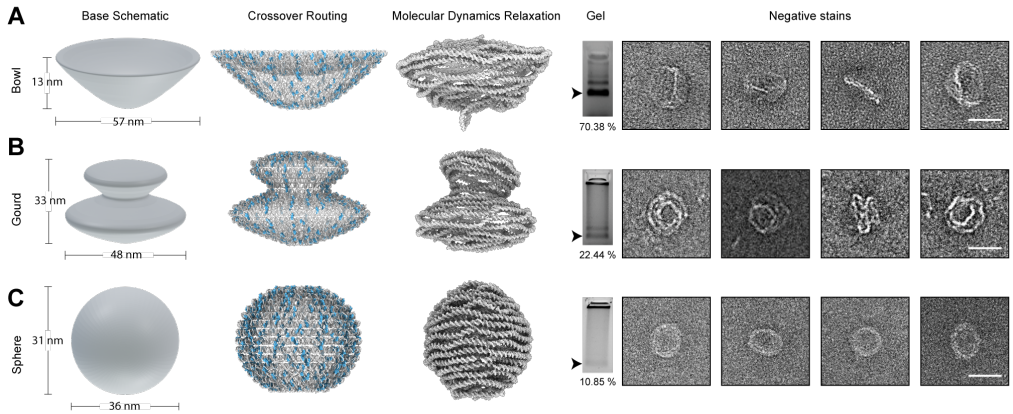

FIGURE 4.9: Experimental verification of automated design principles. **(A)** Bowl design evaluates convex and concave inflection horizontally with respect to its axis of rotation. **(B)** Gourd design evaluates convex and concave inflection vertically along its profile, perpendicular to the axis of rotation. **(C)** Sphere is designed with two layers, one nested and one encapsulating, to evaluate the extension of design principles to multi-layer. (Scale bars 40 nm. TEM and AGE images were collected, analyzed, and prepared by Raghu Pradeep Narayanan, Abhay Prasad, and Dewight Williams.)

A mix of these localized and global heuristics approaches towards the design of curved, enclosed, capsule-like DNA origami nanostructures were then verified in simulation and experimental demonstration of some archetypal geometries: bowl,

gourd, and sphere, which vary in their inflection between concavity and convexity along different axes, while also differing between sharp or soft inflections.

The bowl is hyperbolic, with a vertical axis parallel to the axis of rotation. Its inflection point at the bowl's rim is approximately 48°, so crossovers to each adjacent helix from their common middle helix will likely differ by only one base pair along the helix. The inflection occurs between helices that have, in order, 10.17, 11.25, and 10.58 bps/turn of helical twist, therefore limiting the number of crossovers that will align on short periods. Under conventional and stringent constraints, there would be scarce choices for crossovers, but via relaxation of spacing and alignment criteria, additional valid crossovers are found and applied to form the structure.

The gourd is hyperbolic, with a vertical axis perpendicular to the axis of rotation. Multiple inflection points demonstrate there is no issue for the outline of the structure to follow a freeform path.

The sphere is designed to be fully multi-layer. Instead of a single-layer, each helix can potentially have crossovers to three nearest-neighbor helices rather than only two. For the larger circumferences within the structure, to maintain crossover spacing and density, as many as 20 crossover positions may need to be found along a single helix to maintain a stable connection with each of its neighbors. Crossover choices from a single helix also propagate to and reduce crossover choices available for an adjacent helix, which creates a decision-making chain that appears intractable for a human designer but can nonetheless be implemented by a series of reasonably straightforward queries in computational design.

DNAxiS also exports to formats compatible with widely used molecular dynamics simulators, such as oxDNA [83, 152–154], we were able to prototype designs more economically in silico in conjunction with experimental demonstrations. Figure 4.9 shows the design and characterization of each archetypal geometry. Mean structures calculated from molecular dynamics simulations indicated that structures were likely

to form the intended shape. Thus, we proceeded to test them experimentally, with yields characterized by agarose gel electrophoresis (AGE) and structures observed by transmission electron microscopy (TEM). All shapes formed with good yields (bowl=70.38%, gourd=22.44%, sphere=10.85%).

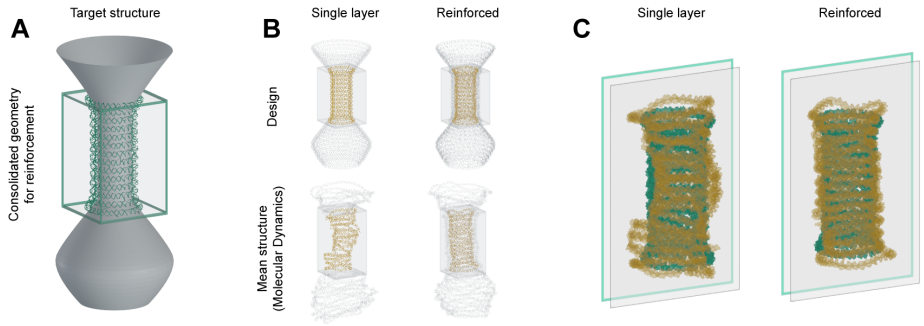

FIGURE 4.10: Evaluation of reinforcement strategies via molecular dynamics simulation. **(A)** A portion of the vase is selected for targeted reinforcement via the addition of coplanar helices. **(B)** Designs generated from DNAxiS (top) are submitted into oxDNA (bottom), after which mean structures are calculated from 103 samples taken uniformly from 10^8 trajectory steps. **(C)** The interior layer (yellow), which determines the shape of the cavity, for both the single layer and reinforced section are consolidated and compared to their intended shape (green). The overlay shows that the reinforced structure more closely conforms to the intended shape of the neck.

Yet, while the bowl and gourd formed well, some other single-layer designs that were designed and characterized did not, either with low yields or showing malformed geometries (Fig 4.11A, C, E). We continue by designing multi-layer versions of those single-layer designs under similar criteria to the sphere. Mainly, we were interested in whether an extension of the automated methods towards multi-layer structures could generate structures of the same shape but with higher yield and improved shape accuracy.

4.2.5 Evaluating yield and feature formation in multi-layer designs

Thus, the efficacy of reinforcement was tested. The addition of two design principles helps to create partial multi-layer designs. First, multi-layer sections are defined by

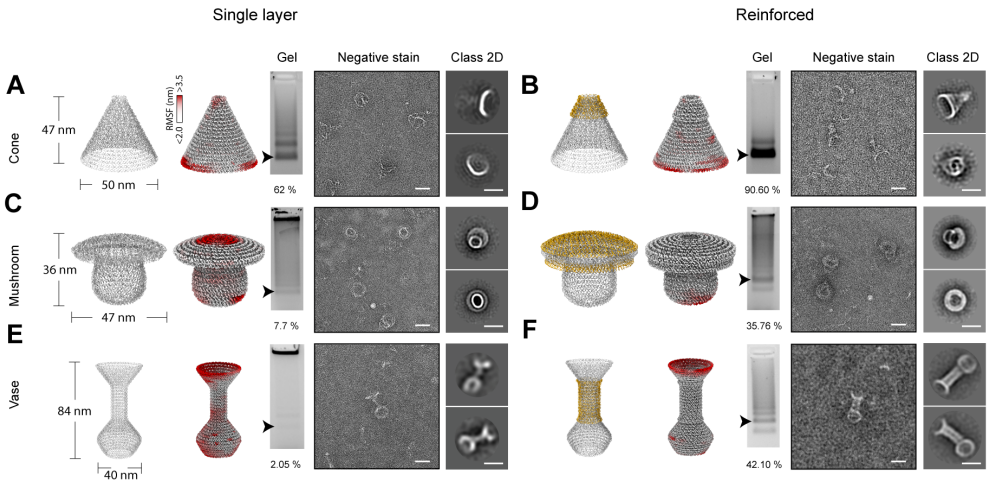

FIGURE 4.11: Evaluation of reinforcement strategies via experimental characterization. Comparing geometries (grey) using RMSF calculations calculated from trajectory steps generated by oxDNA, yields are shown by AGE, TEM micrographs, and 2D class averages. **(A-B)** Cone is reinforced at the tip to improve the synthesis of high-resolution features. **(C-D)** The mushroom demonstrates multiple disconnected reinforced segments, including the top of the mushroom head and its underside. **(E-F)** The vase is reinforced along its thin neck, showing improvements in both appearances under TEM imaging and yield as measured by AGE. (Scale bars 40 nm. TEM and AGE images were collected, analyzed, and prepared by Raghu Pradeep Narayanan, Abhay Prasad, and Dewight Williams.)

adding coplanar helices outward from each helix within a chosen section of the single-layer design (Fig 4.2G). All planes are perpendicular to the shape's axis of rotation. This forms another layer with a similar topology as the single layer. In addition, the path of the scaffold through the helices is fulfilled in-plane first, forming a "zig-zag" pattern (Fig 4.12). This allows the structure to be reinforced anywhere instead of only at the top or bottom.

Initially, in oxDNA, we investigated the effects of reinforcement in basic cylinder and sphere shapes for up to three additional layers (Fig 4.13). However, most likely due to tradeoffs between crossover spacing and density, increases in mechanical rigidity tapered significantly after the addition of only one additional layer. Thus, these same methodologies were applied using only one additional reinforcing layer,

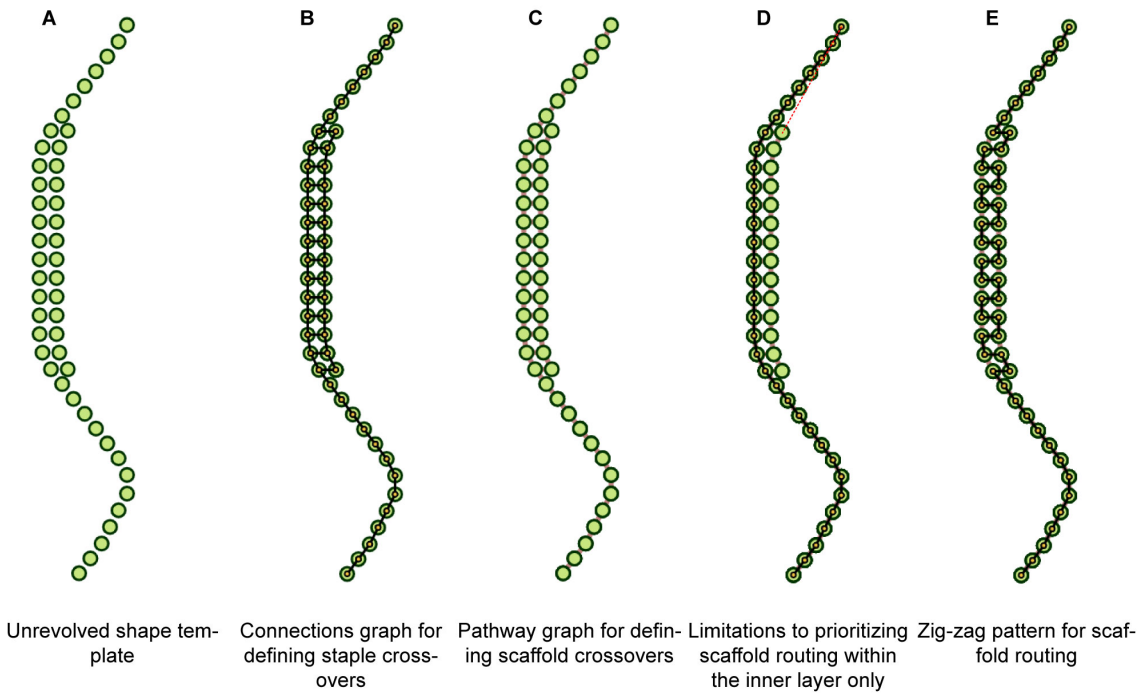

FIGURE 4.12: Interface and scaffold routing pathway challenges. A user is always provided with a Connections (for staples) and Pathway (for scaffold) interface for declaring pairs of nearest neighbor helices that should have staple or scaffold crossovers between them. **(A)** Nodes (in green) represent helices of an unrevolved profile of the axially symmetric DNA nanostructures that acts as the template. **(B)** Lines connecting between a pair of nodes indicate to the software to look for crossovers between those two helices. **(C)** Connected nodes continue to be labeled in red. A scaffold routing is also now defined via its connectivity graph. **(D)** For multi-layer structures, routing the scaffold per layer of the shape is not recommended. If there are reinforcement segments of the structure not at the top or bottom, then they cannot be reached. **(E)** For routing a scaffold strand through a selectively multi-layer structure, a "zig-zag" pattern ensures that the structure is modular between the top and bottom segments. If the structure is segmented, the segmented piece will still form without any modifications to the staple strands. The zig-zag pattern is also consistently applicable regardless of the number of layers in the reinforced section as long as the height of the reinforced segment is an even number of rings. We do not refute that other, more complex routings are possible, including for odd-numbered heights, but those have not been fully evaluated within the scope of this work.

first upon the neck of a vase design (Fig 4.10A). In its single-layer variant, the neck conforms less to the intended cylindrical design while also becoming slightly askew. These issues are resolved in the multi-layer variant. In mean structure diagrams

(Fig 4.10B), it is evident that the reinforced structure is more stable and resembles its intended shape. The reinforced area is also consolidated and overlaid with its intended geometry as designed, and the reinforced variant notably aligns better with the input shape (Fig 4.10C). Next, reinforcement was applied in two additional cases to study, testing potential weaknesses in narrow and wide geometric features.

The single-layer cone design (Fig 4.11A) evaluates the effective range of ring-based design across a range of circumferences. The cone measures only 72 bps in circumference at the tip and 472 bps at the base. RMSF of shapes are calculated from 103 snapshots of a 108-step trajectory from molecular dynamics simulations and compared to its mean structure calculated across the same snapshots. Visualization of this data shows high RMSF not only at the base but also at the tip of the structure. Narrow rings of curved nanostructures likely form poorly due to being highly strained due to the low radius of curvature while also having little room to add additional crossovers to stabilize the curvature. The peak almost always cannot be resolved in TEM micrographs.

On the other hand, despite high RMSF, the base of the cone appears to form well. Reinforcement was then applied at the top six helices, building a second layer around the peak of the cone (Fig 4.11B). Running molecular dynamics simulations again (the single layer is consolidated, and the view of the second layer is hidden here for demonstration only), we observe a marked improvement in the RMSF distribution in the peak of the cone as well as in the yield experimentally, seeing improvements from 62% to 90.60% of well-formed structures as determined by ImageJ analysis of band intensity after gel purification. Furthermore, the tip of the cone is now visible in TEM micrographs and their respective 2D class averages.

Next, we evaluated reinforcement in a mushroom design (Fig 4.11C), which is characterized by a wide but thin aspect ratio. The shape also contains sections that slope rapidly vertically (the stem) and horizontally (the cap). In TEM micrographs,

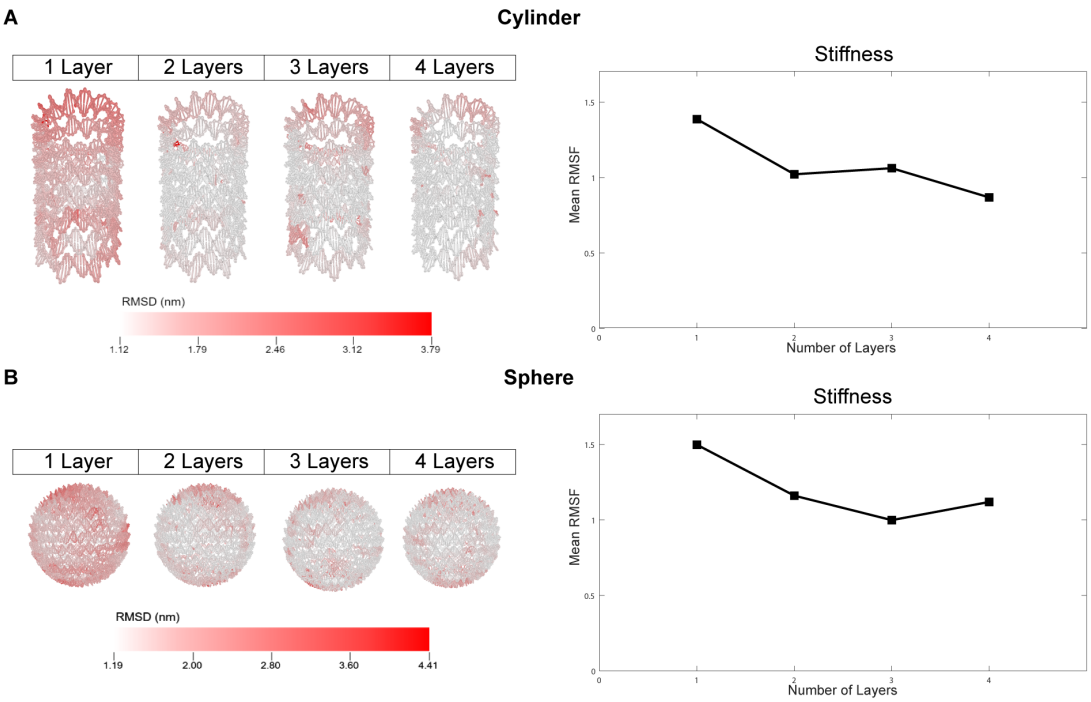

FIGURE 4.13: Initial screening for effects of multi-layer curved design. To investigate the possibility of improvements in stability due to additional layers, basic examples were evaluated for layers up to four. **(A)** Cylinder represents the most basic 3D shape that exhibits no out-of-plane curvature. **(B)** Sphere then evaluates for any differences due to the addition of out-of-plane curvature. RMSF is calculated only for the innermost, shape-determining layer. Excess layers are hidden in the diagram for illustrative purposes only. Each result's aggregate score is calculated and then plotted to visualize trends. Structures going from one to two layers improved similarly regardless of the complexity of curvature. However, trends from additional layers are not clear and have yet to be thoroughly investigated. One possibility that has been proposed is that at three or more layers, effects on the stability from staple region length become more prominent. Three layers require helices in the middle layer to connect to up to four adjacent neighbors. Crossovers may be occurring so frequently in those helices that each staple binds a shorter region of the scaffold, weakening its hybridization and the overall stability of the structure.

we frequently observed stems forming without caps. In RMSF-illustrated diagrams, we notice that in horizontal sections, as the circumferences of the helices become smaller towards the center of the cap, RMSF rapidly increases. Here we demonstrate a disjointed application of targeted reinforcement, as well as showing that any

reinforcing layer can be added tangent to its single-layer profile. Helices for an additional layer are added at the top, along the cap of the mushroom, then skipping the rim (anatomically termed margin), and continuing along the underside of the cap (Fig 4.11D). Again, we observe a marked decrease in RMSF, as shown in simulations, as well as a significant increase in the yields of fully formed structures (from 7.7% to 35.76% by AGE). The TEM micrographs and their respective 2D class averages also showed that the cap of the mushroom was more resolvable than its non-reinforced counterpart.

Finally, we experimentally characterized the vase design (Fig 4.11E) to verify our in-silico observations. In the single-layer variant, we frequently observe a neck portion that is poorly formed and often appears as if the cup of the vase becomes separated at the neck from the bulb. When reinforcing the neck (Fig 4.11F), structures appear straighter, and most are fully formed with the cup attached. We note that while the rim of the cup and the mid-section of the bulb are similarly wide rings, RMSF is significantly higher along the rim while hardly present in the bulb, and this is most likely due to the structure becoming open-ended at the rim, while neighboring helices on both sides support the bulb. Yields increased dramatically from a very low 2.05% up to 42.10%, possibly showing that there can be specific areas of weakness in some geometries that are more significant towards the formation of the structure. TEM micrographs and their respective 2D class averages showed a clear trend in the stability of the structures wherein the bulb, stem, and cup region of the vase was fully resolvable compared to its non-reinforced counterparts.

4.2.6 Design and characterization of asymmetric structures

Fundamentally, the automation provided by DNAxiS is a general strategy for designing DNA origami shapes that exhibit continuous curvature in 3D. Building upon crossover network algorithms initially applied to templates for axially symmetric

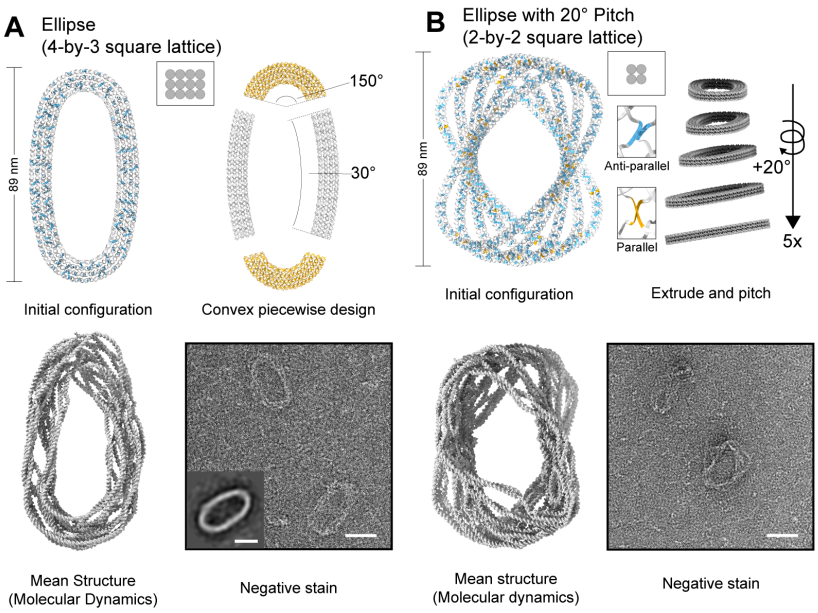

FIGURE 4.14: Axially asymmetric structures, Ellipse. **(A)** Modules showing the construction of an ellipse shape from helix bundle arcs with a 4-by-3 square lattice cross-section. **(B)** The ellipse shape is repeated five times in a vertical extrusion and pitched by 20°. The cross-sectional area of each ellipse is reduced to fit the entire design within a limited length of scaffold sequence, despite already using multiple scaffolds. Both parallel and anti-parallel crossovers are created to increase the crossover count and yield of the design. Scale bars on TEM micrographs and 2D averages are 40 nm. TEM images were collected, analyzed, and prepared by Raghu Pradeep Narayanan, Abhay Prasad, and Dewight Williams.)

structures; it was possible also to design several structures with some amount of axial asymmetry to demonstrate the benefits of this rapid automated approach towards crossover placement (unfortunately, an interface does not yet exist for this in DNAxiS as of this writing). DNAxiS abstracts low-level design of crossovers into quantifiable values. Whereas rings were defined by their circumference, rings could further be consolidated into segments of arcs, each represented by their arc length, radius of curvature, and position. Each arc is a module that can be added to the design space and connected to other modules by allowing staple and scaffold strands to run continuously through their endpoints. Arcs are designed as angled DNA helix

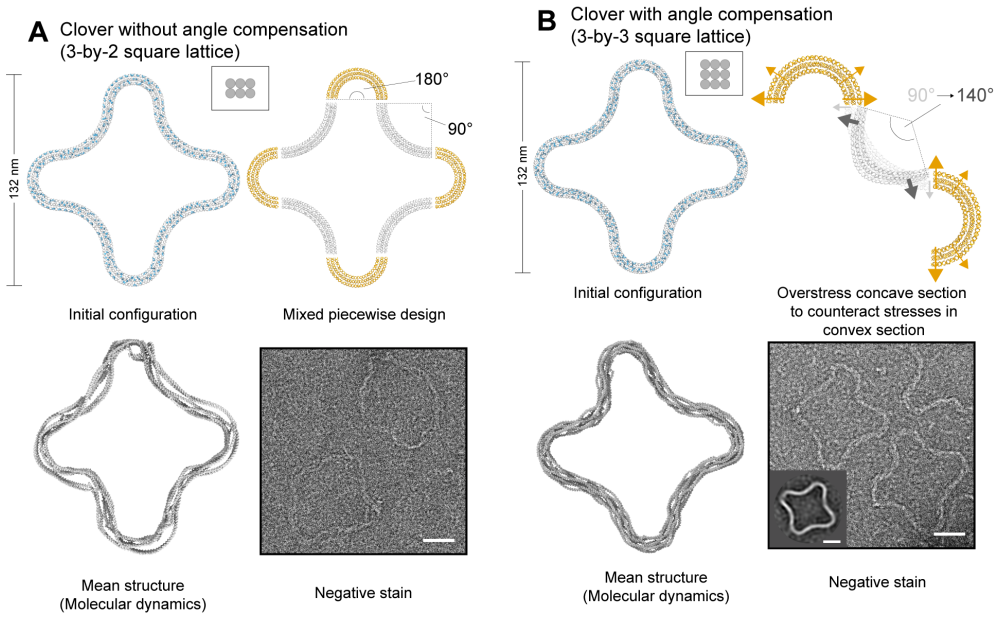

FIGURE 4.15: Axially asymmetric structures, Clover. **(A)** Modules showing the construction of a clover shape from helix bundle arcs with a 3-by-2 square lattice cross-section. **(B)** The cross-section of bundles is expanded to 3x3 to increase the upper bounded stiffness of each module. This is necessary to achieve sufficient counteracting strain on inside corner modules to preserve the inflection between convex to concave segments of the structure without "rounding out" as it did in 3-by-2 clover. Scale bars on TEM micrographs and 2D averages are 40 nm. TEM images were collected, analyzed, and prepared by Raghu Pradeep Narayanan, Abhay Prasad, and Dewight Williams.)

bundle upon a square lattice using an insertion-deletion gradient as in previous work (3) but calculating gradients using a simple Euler-Bernoulli beam model (Fig 4.16).

Two basic archetypal structures were designed to demonstrate convex and concave geometries. Modules defined as arcs can be linked together to create axially asymmetric shapes. Each module is a bent bundle of variable cross-sectional helices where design principles implemented in DNAxiS for multi-layer structures can directly apply and were used to generate ellipse and clover shapes. Each structure was designed using varying bundle dimensions (4-helix-bundle (2-by-2) up to a 12-helix-bundle (4-by-3)) depending on the length of the available scaffold as well as desired

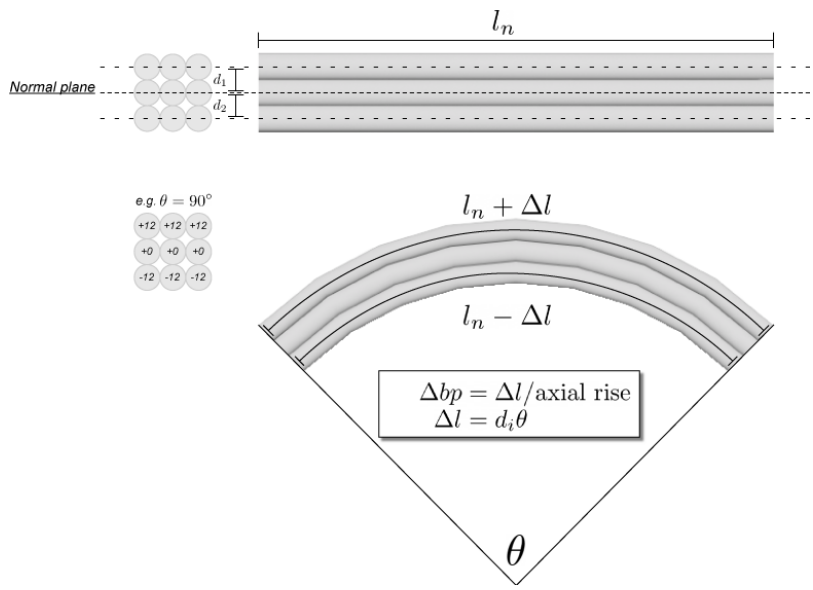

FIGURE 4.16: Euler-Bernoulli beam model for direct calculation of insertion/deletion gradient in arc modules for asymmetric design. Design of bending arc modules for the design of axially symmetric structures uses a simple beam model to directly calculate and create the DNA helix template within the software that is passed to the crossover network algorithm.

rigidity. The ellipse (4-by-3 square lattice bundle) was constructed using two each of two arcs (30° and 150° defined with different inputs of their radius of curvature and arc length (Fig 4.14A). The clover shape merged convex and concave geometries and was constructed with four 180° arcs for the convex leaf sections and four 90° arcs as concave inside corners (Fig 4.15A).

Using the ellipse, we further demonstrated the capabilities of generalized routing by extruding the shape with pitch. Each ellipse was reduced to a 2-by-2 bundle and then placed with a 20° rotational offset to its preceding ellipse in the design space, then the routing algorithms of DNAxiS were applied. Notably, the pitch of the shape eliminates many potential crossover positions, and we allowed DNAxiS to find and apply both parallel and non-parallel crossovers. Upon inclusion of parallel crossovers, the pitched ellipse could finally create enough crossovers to form and appear under TEM imaging (Fig 4.14B).

The clover (3-by-2 square lattice bundle) unveils a problem that is most likely unique to DNA nanostructures that have inflection points within a continuous, curved perimeter, which has rarely appeared in prior works. Defining each module of the clover naively according to the angles as they appear often caused the synthesized structure to "round out," as seen in Figure 4.15A. We suggest that at module boundaries, where there is an inflection between concave and convex geometry, a mismatch of stresses can cause the structure to misform. A higher strained module, in this case, the convex leaf section, could overwhelm an adjacent module, in this case, the concave inside corners, as an effect of dissipating strain from its area. In order to strengthen the concave sections, an "overcompensation" strategy was proposed to increase the strain of inside corner modules in a direction opposite to the strain produced by the leaf modules.

We used the rapid prototyping capabilities of DNAxiS to test this possibility by iteratively designing and simulating a range of angles for the inside corner of the clover. Furthermore, the bundle cross-section was scaled up to a 3-by-3 square lattice to increase the achievable stiffness of the inside corner. In 10° increments, prospective designs were filtered by adding oxDNA simulations into the design pipeline. As the internal angle of the concave section increased from 90° to 140°, we observed minimal gain past 140°. The shape conformation of the sample, as viewed in TEM, also became closer to the intended geometry. Under TEM micrographs and 2D class averages, we observed a similar change in the synthesized shape, which verified our overcompensation strategy (Fig 4.15B).

4.2.7 Rapid prototyping of DNA capsules

The design principles and input parameters for automated DNA capsule design discussed in this chapter can continue to be developed to investigate and generalize methods for DNA origami design for optimization and shape complexity. In Figure

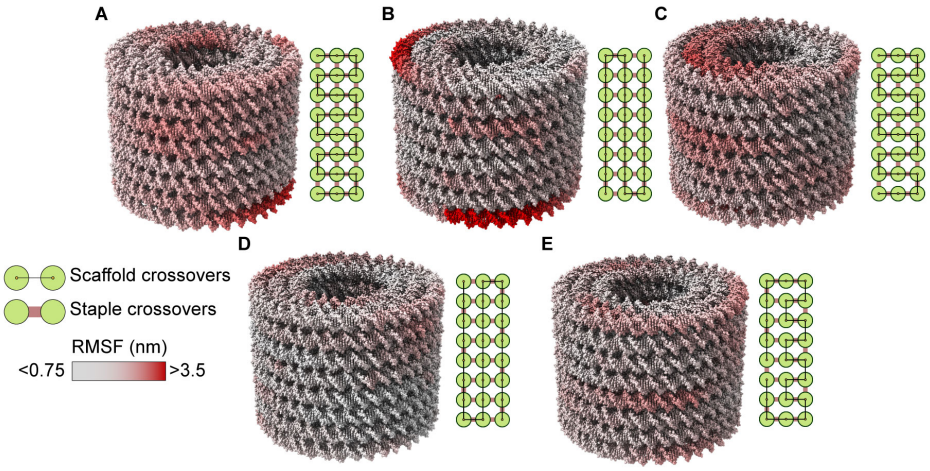

FIGURE 4.17: RMSF in alternate connection graphs. These examples demonstrate the rapid prototyping functionality of DNAxiS. **(A)** Fully connected staples graph with "zig-zag" scaffold routing with in-plane priority. Helices within the middle layer must connect to four adjacent helices. This is the primary strategy used to demonstrate 2-layer structures within this work. **(B)** Fully connected staples graph with "zig-zag" scaffold routing with in-layer priority. Staple crossovers from helices within the middle layer must connect to 4 adjacent helices. **(C)** Staple connections graph reduced such that each helix has at most three adjacent helices with "zig-zag" scaffold routing with in-plane priority. **(D)** Staple connections graph reduced to a "zig-zag" with in-plane priority. Staple connections have a connection degree of 2, but scaffold crossovers still form connections to 4 adjacent helices. **(E)** Staple connections graph with "zig-zag" in-plane priority but with a variable scaffold routing. Option D appears to be the best for stability in these examples, though that may not always be the case across all geometries. The effect on the oversaturation or undersaturation of crossovers in curved nanostructures remains to be thoroughly characterized. RMSF measurements were calculated from 103 trajectory steps sampled from MD simulations performed in oxDNA2.

4.17, different connectivity graphs are tested to evaluate the effect of an oversaturation of crossovers. An excess of crossovers may add too much strain or reduce the average length of regions. Both of these effects may lead to lower stability. In the examples evaluating this premise upon a simple cylinder shape, routing D, where many helices are reduced to bind only two nearest neighbors instead of the maximum four, there is some noticeable improvement in RMSF distributions throughout

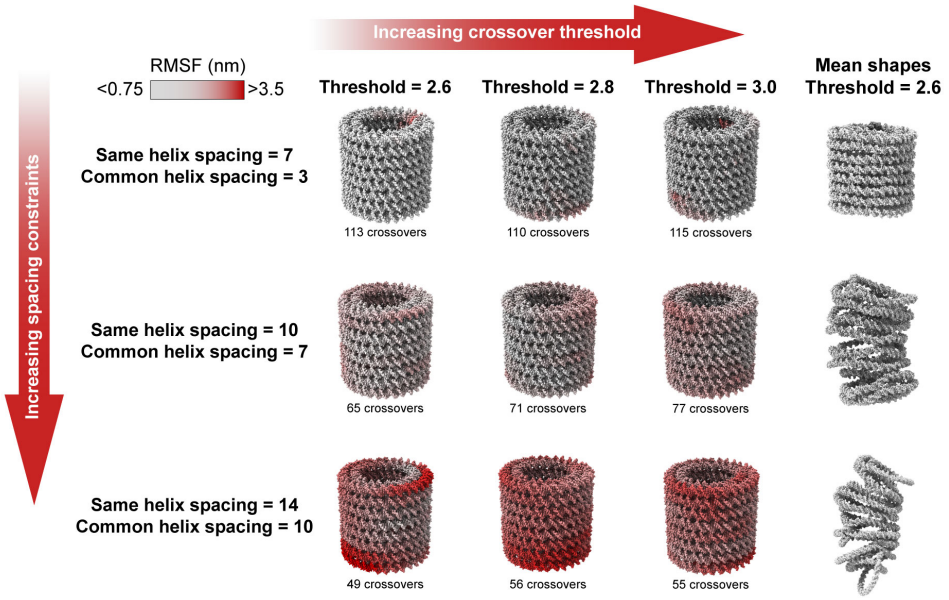

FIGURE 4.18: Parameterizing generalized DNA origami design. In efforts to generalize DNA origami design across all geometries, DNAxiS consolidates a few parameters to describe the design process. The software optimizes structures based primarily on constraints to its crossover threshold and helix spacing (see Figure 4.5). These values imply certain crossover densities, which in turn determine the stability and accuracy of the structure. However, using crossover density as a design parameter is challenging as it is difficult to relate it back to constraints that are relevant to the design process. An example structure is a 2-layer cylinder with 128 bps circumference in the inner ring and 176 bps circumference in the outer ring, with a height of 8 helices. RMSF values are calculated from a trajectory of 10^3 steps uniformly sampled from a 10^8 step molecular dynamics trajectory simulated using oxDNA2.

the structure. In Figure 4.18, we evaluate the trends of crossover density to the input parameters. As in previous example structures, here we also find that a relaxation of the constraints correlates to a higher crossover density, which more closely approximates the intended shape.

4.3 Conclusion

DNAxiS is uniquely the first software tool to apply automated routing techniques based on heuristic optimization algorithms. As DNA nanostructures increase in

scale, with less regularity and more independently designed components, designs may move into a space that becomes intractable for human designers while also lacking sufficient algorithmic regularity. Our work presents an approach for novel DNA origami designs as oblique, freeform structures, which may enable the design of novel geometries, thus enabling wider access to the field for the larger scientific community. We demonstrated how this strategy could succeed for curved DNA origami designs forming enclosed shapes that do not have any apparent algorithmic regularity to exploit. We described the design complexity that arises from structures with a non-homogeneous helical twist, notably skewing the periodicity of crossover positions to a point where it becomes tedious for a human designer to arrange a stable crossover network. Repeating this for each unique crossover pattern between adjacent helices can quickly become overwhelming for a human designer and become a barrier to applying DNA nanotechnologies across interdisciplinary applications. This work introduced design principles that include the multi-layer design of curved and enclosed shapes and characterized the effect of reinforcement on the yield and shape accuracy of structures. Furthermore, the benefits of automation and software modeling were demonstrated in axially asymmetric shapes that took advantage of the ability to edit geometries and calculate crossover networks rapidly. A potential impact of our work is to motivate wider consideration of the design and manipulation of globular and more biomimetic DNA nanostructure.

5

DNA Interfaces to Cell Membranes

5.1 Introduction

DNA nanotechnologies capture both the programmability and the ability to interface with other biomolecules [40, 44] with limited immunotoxicity [155] and unwanted reactivity that would be necessary to construct complex, autonomous, molecular-scale therapeutic devices. For the utility of DNA nanostructures to realize its full societal impact, branches of DNA nanotechnology, such as DNA computing and aptamer chemistry, must begin integrating into complex DNA-based systems. For example, one area with the potential of finding transformative DNA nanotechnology-based technologies is the area of cancer diagnostics and therapeutics.

This chapter begins with a discussion of a previously published work for localized DNA computing architectures and an application of it for identifying cancer cells by simultaneously recognizing multiple membrane proteins. Following that is a discussion of a distributed computing architecture and simulations developed in the context of targeting heterogeneous tumor tissue. Clinically, tumor cell tissue is not always comprised of a single cell type. Thus, targeted drug delivery systems may

not thoroughly remove a tumor if only portions of it are being identified.

5.2 Background

5.2.1 Localized DNA hairpin circuits

A *DNA hairpin* is a DNA nanostructure formed with only a single DNA strand. A hairpin is formed when a strand has two separate, self-complementary domains. When those two domains bind, they form the *stem* of the hairpin. The remaining DNA sequence between those domains is then left over to form an ssDNA *loop* at the end of the stem. On the other side of the stem, if there is a portion of one domain of the stem that is not complementary and thus extends further as a domain of ssDNA, this piece can be used as a *toehold* to initiate *toehold-mediated strand displacement* (TMSD). TMSD is a process by which an invading ssDNA strand that is also complementary to the same domain can bind via the toehold and then displace the incumbent strand via *branch migration*. Branch migration occurs when two ssDNA undergo a competitive random walk to bind a strand displaying their complementary sequence. These techniques can be utilized to make programmable *DNA circuits*, which are a collection of DNA strands programmed to simulate the operation of binary logic gates.

Some such circuits can be designed as a *hybridization chain reaction* (HCR). It is a cascading DNA hairpin reaction in a solution containing several distinct DNA hairpins. After a DNA hairpin is opened via a strand-displacement reaction by some initiating strand or molecule, its exposed single-stranded leg of the stem hybridizes to a subsequence DNA hairpin as that subsequent hairpin's initiator. This relationship is applied recursively for several hairpins. As freely diffusing moieties, these DNA hairpins eventually form a long string of DNA hairpins as a duplex connected via hybridizations. The cascade was initially used to transduce the detection of an analyte into a large DNA nanostructure that could be detected by gel electrophoresis.

Bui *et al.* previously developed a spatially localized DNA HCR circuit [156] based on a cascading hairpin reaction where each hairpin is localized on a self-assembly DNA nanostructure. Localization is accomplished by affixing a DNA hairpin to a surface via some modification to its stem or by hybridization to a surface via a sequence also extended from a leg of its stem. When hairpins expected to react together are positioned spatially close to each other, it significantly increases the rate at which each hairpin finds its subsequent hairpin in the chain. Spatial localization according to a priorly known reaction sequence allows each hairpin to avoid diffusing slowly towards its target. Thus, it will enable the hybridization reaction to begin immediately with a neighbor whose position has been affixed nearby. The effect of this can be a roughly 10-fold increase in the completion rate of the circuit.

Furthermore, the localization can also be exploited to measure each node's presence in the chain as a n-input AND. If all nodes are present, the hairpin chain reaction will proceed to the terminal node, which releases some output signal. If not all nodes are present, this output does not occur, although it is unknown where the chain was broken.

Song *et al.* used this architecture to profile cancer cells by identifying each as a combination of their characteristic membrane proteins [97], demonstrating the identification of four distinct cancer cell lines using 2- and 3-input circuits. Each hairpin is affixed with a distinct *DNA aptamer* corresponding to membrane proteins existing on the cancer cell membrane. A DNA aptamer is a short sequence of DNA that binds to a specific target molecule [157]. Song *et al.* affixed each distinct DNA hairpin of the HCR with a distinct aptamer that uniquely binds to a specific distinct membrane protein existing on the cancer cell membrane. Each node will be planted upon the cell membrane if all membrane proteins are present. Once an initiator is added, the chain reaction will initiate. Once the last node is activated, it will recruit a reporter duplex. The reporter duplex contains two strands, one functionalized with

Table 5.1: High-affinity aptamers per cell line.

Cell Line	Aptamers
CCRF-CEM	Sgc8c, Sgc4f, TE13, TE17, TC01
Ramos	Sgc4f, TE13, TD08, TC01

a fluorophore and the other with the fluorophore's corresponding quencher.

The fluorophore strand of the duplex exposes a toehold. Once the terminal node of a hairpin chain reaction is activated, the fluorophore strand will begin to bind, thereby displacing its quencher strand. The fluorescent strand is thus attached to the cell via the already affixed hairpin. It marks the cell with a fluorescent signal that can be observed and used to count the number of labeled cells via flow cytometry.

5.2.2 Multiple cell types

In two of the four cancer cell lines used in the demonstration of the cascading hairpin reaction, both have binding affinity to several common aptamers. They may share similar membrane proteins. Thus, any strategy to identify and differentiate the two cell lines by only a single biomarker may produce ambiguous results.

Instead, the cell surface can be used as the operating substrate for the localized hairpin reaction. CCRF-CEM and Ramos share three common aptamers. CCRF-CEM has only two uniquely binding aptamers (Sgc8c and TE17), while Ramos has only one (TD08). While there is an opportunity to measure each separately by only one biomarker, this strategy is not robust, such as if a specific aptamer is unavailable. Instead, for instance, if aptamers could be used pairwise, then the number of ways to measure and identify each cell line is permuted by pairs of aptamers each cell line has in common and not in common. Ultimately, it provides more reliable identification for cancer cells by analyzing a broader collection of its surface proteins.

This chapter further extends the localized DNA computation architectures and

their applications in identifying cell membranes to search for strategies to identify heterogeneous mixtures of cancer cells. One of the biggest obstacles against the efficacy of aptamer-based targeted drug delivery is tumor heterogeneity [158–160]. A mixture of tumor biomarkers that may exclude a particular aptamer will limit its targeting accuracy. Such therapeutics may not be able to thoroughly treat a tumor, thus allowing undetected and so untreated cells to proliferate again. Several works have therefore suggested multiple-aptamer strategies [161–163].

While previous work by Song *et al.* was able to utilize multiple aptamers, its advantages were identifying single cell lines with greater accuracy. Instead, we consider whether a localized hairpin chain reaction circuit can be split at some node, distributed to multiple surfaces, and remain operable. It should be noted that while cells can still diffuse, due to their size relative to DNA strands, for practical purposes, we consider the rate of diffusion of the cells negligible compared to the rate of diffusion of DNA strands. We regard the cells as mostly static surfaces.

In the remainder of this chapter, we discuss this design for a circuit of only two nodes and its simulated results, but with the expectation that with robust primitives, it can be extended towards more nodes to improve both its accuracy and range of operation.

5.3 Design

To begin, we consider the presence of two substrates, *cell 1* and *cell 2*, and designate corresponding hairpin nodes, H_1 and H_2, to attach to their respective cell membranes. To link the output signal of H_1 to the input of H_2, we introduce the *signal complex* into the architecture. The signal complex is constructed with an initiator I_2 specific to H_2 but protected by a *deactivator* strand. The deactivator strand exposes a toehold L^*, which identifies it to hybridize with a domain within the hairpin loop of H_1. When H_1 is activated by its initiator strand and exposes its hairpin loop, the

signal strand may bind at the exposed L^* domain and undergo a branch migration such that eventually, the deactivator strand is affixed to H_1. The initiator I_2 can then begin to diffuse in solution, looking to initiate the H_2 reaction (see Figure 5.1).

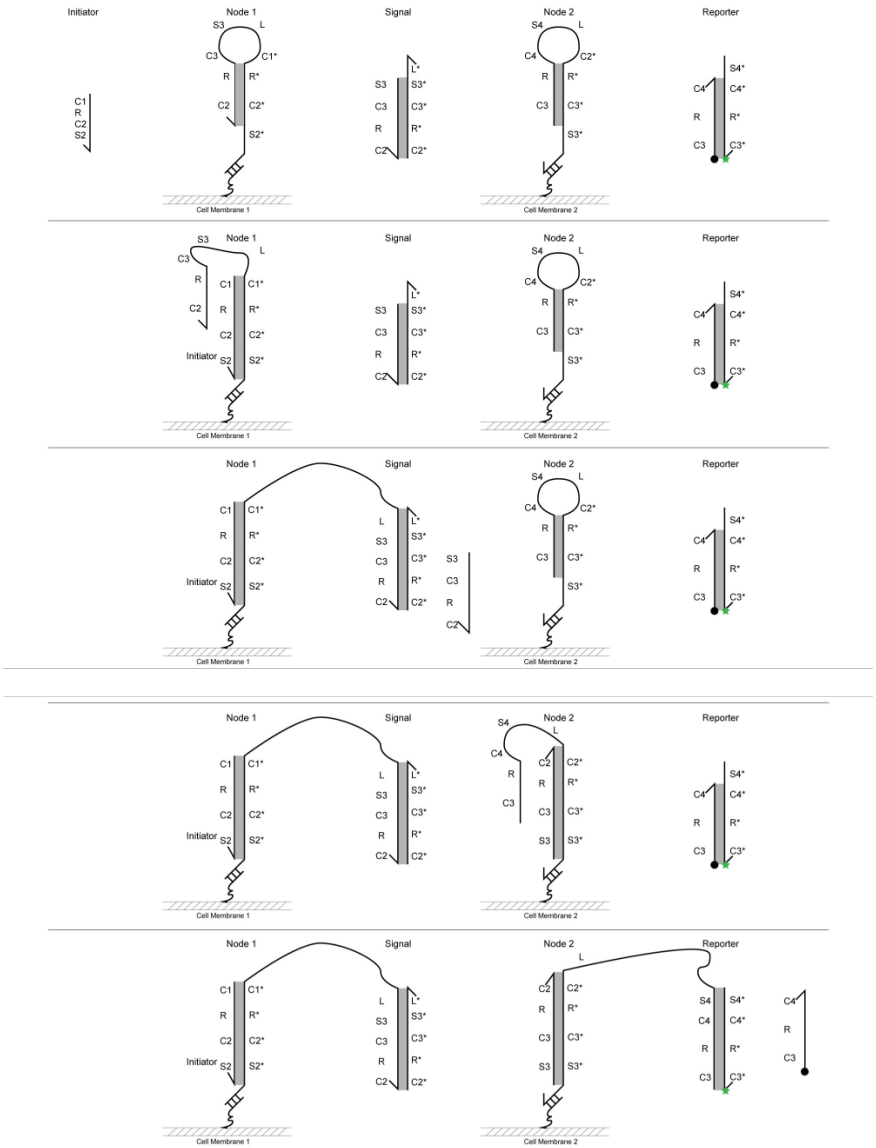

FIGURE 5.1: Summary of reaction steps in distributed localized DNA hairpin gates for a distributed 1-1 design.

5.4 Simulation

To test our design, simulations are performed in Peppercorn [164], a domain-based simulator for DNA strand reactions. However, simulations are limited by the lack of functionality to specify localization that would reject specific reactions. In a normal chain reaction, an activated H_1 would be expected to trigger an activation in H_2. However, if assuming that the diffusion rate of cells is sufficiently slow, then cells are regarded as static surfaces that can separate the hairpins at a distance where reactions will be few or ideally non-existent.

The function of the signal strand is first tested in the context of the entire circuit. In Figure 5.2, we track the presence of the final complex that is expected to form between H_2 and the fluorescent reporter strand. Its signal is low relative to the concentrations of inputs. However, when H_2 is consolidated (Fig. 5.3), it appears that its operation fully activates the reporter complex and binds the fluorescent reporter.

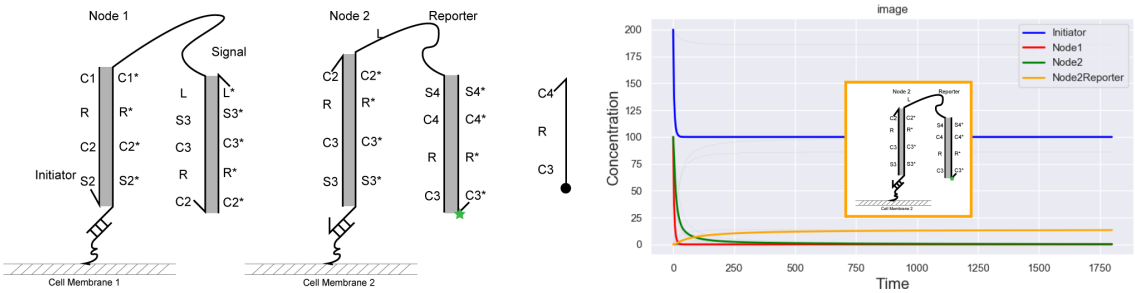

FIGURE 5.2: Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2).

Instead, we attempt to verify the correct operation of the circuit by identifying that loss in signal is lost with a linear relationship to the complex where H_1 has formed a hybridization chain with H_2, which then activates and binds the reporter (Fig. 5.4). The signal intensity of `Node2Reporter` equals the difference between

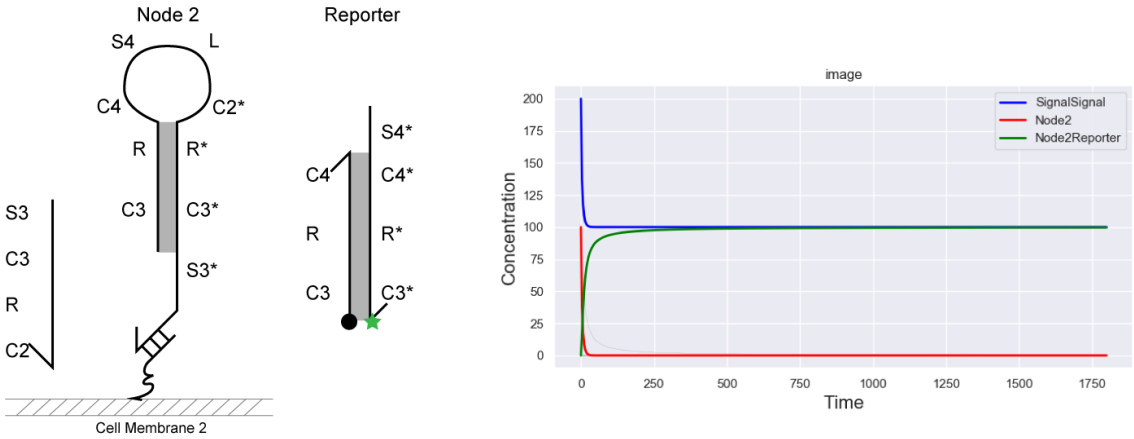

FIGURE 5.3: Simulated operation of only H_2 (Node 2) and the output reporter.

the expected signal (measured by the amount of activated reporter complex as determined by the amount of quencher strands (`ReporterQnchr`) that have been released) and the erroneous signal resulting from a lack of localization in the simulation (`Node1Node2`).

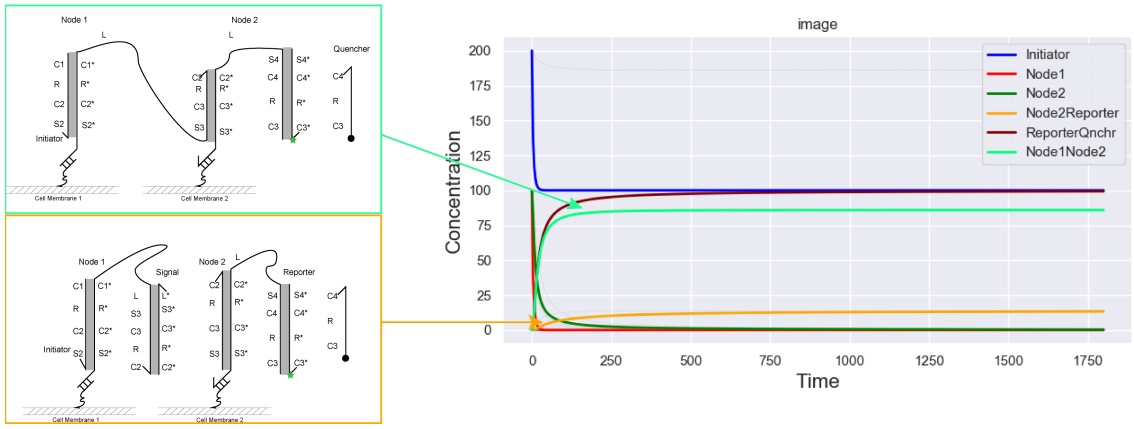

FIGURE 5.4: Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2) with intermediates overlaid.

After confirming the source of the error, the reaction rates of specific reactions can be modified. By ensuring that the `Node1Node2` complex cannot form by eliminating the toehold-mediated strand displacement reaction that would begin with the S3 domain of Node 1 binding with the exposed S3* toehold domain of Node 2, all chain

reactions follow this critical step no longer occur in simulation. Thus, we observe the full signal associated with the final complex formed by Node 2 and the fluorescent Reporter strand (Fig 5.5).

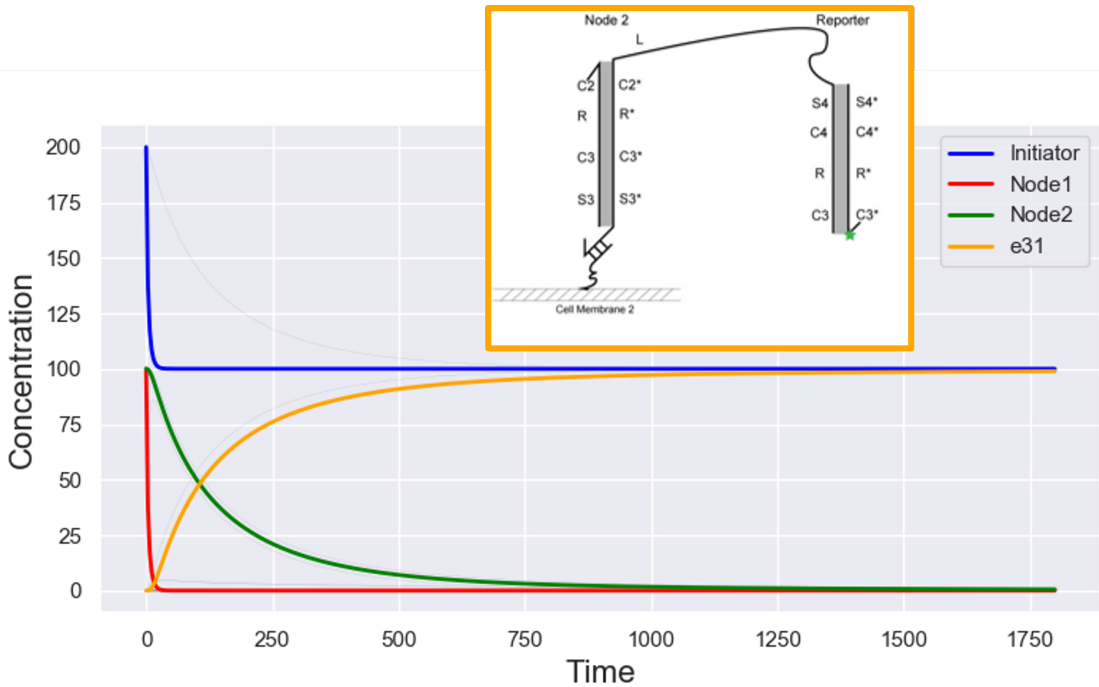

FIGURE 5.5: Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2) with manually enforced localization.

5.5 Scalability

We can then generalize this architecture towards multiple surfaces and signal complexes. Each surface containing hairpins is denoted as a *cluster*. The size of each total cascade is now referred to as a hyphenated tuple i-j-k-... where $\{i, j, k, \dots\} \in \{1, 2, 3, \dots\}$ indicates the size of each cluster and hyphenated clusters indicate the communication of a signal complex between each cluster. In the following simulations we show the performance of a 2-2 circuit (Fig. 5.6 & 5.7) as well as a $1-1-1$ (Fig. 5.8 & 5.9) circuit. In both cases, localization is enforced by reducing the undesired reactions between Nodes, which should be mediated by the Signal Complex, to zero.

Note that while the $2-2$ circuit does not appear kinetically dissimilar from the $1-1$ circuit demonstrated above, the $1-1-1$ circuit activates fluorescence much slower. While reaction rates in Peppercorn are enumerated under the assumption that reactions happen diffusively in solution, this trend can still be expected after localization. Localized steps are fast, while diffusive steps, which describe the movement of the signal strand initiator after the signal complex has been activated, are slow.

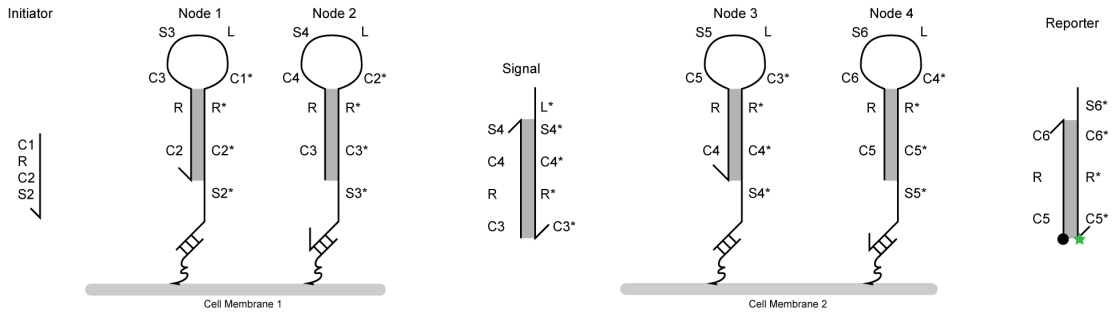

FIGURE 5.6: Node and signal complex design for a localized 2-2 circuit.

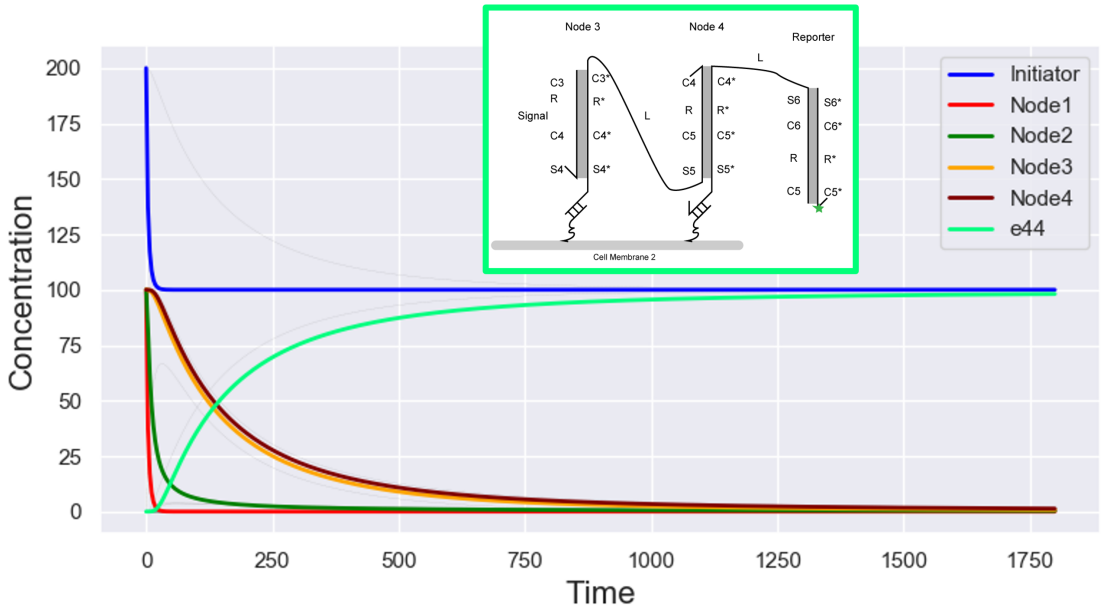

FIGURE 5.7: Simulated operation of distributed node clusters H_1-H_2 (Node 1 and Node 2) and H_3-H_4 (Node 3 and Node 4) with manually enforced localization.

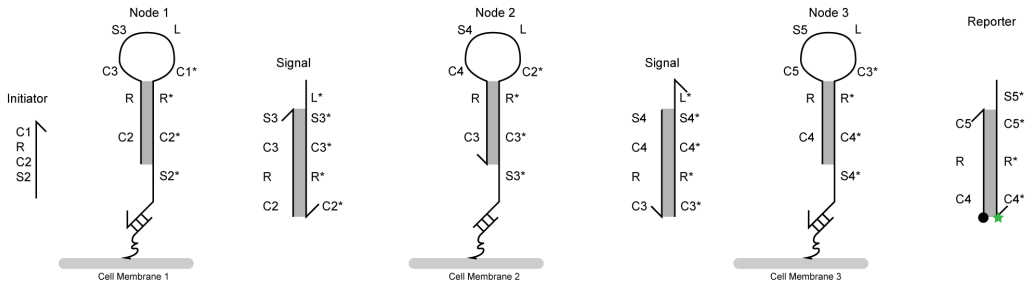

FIGURE 5.8: Simulated operation of distributed nodes H_1 (Node 1) and H_2 (Node 2) with intermediates overlaid.

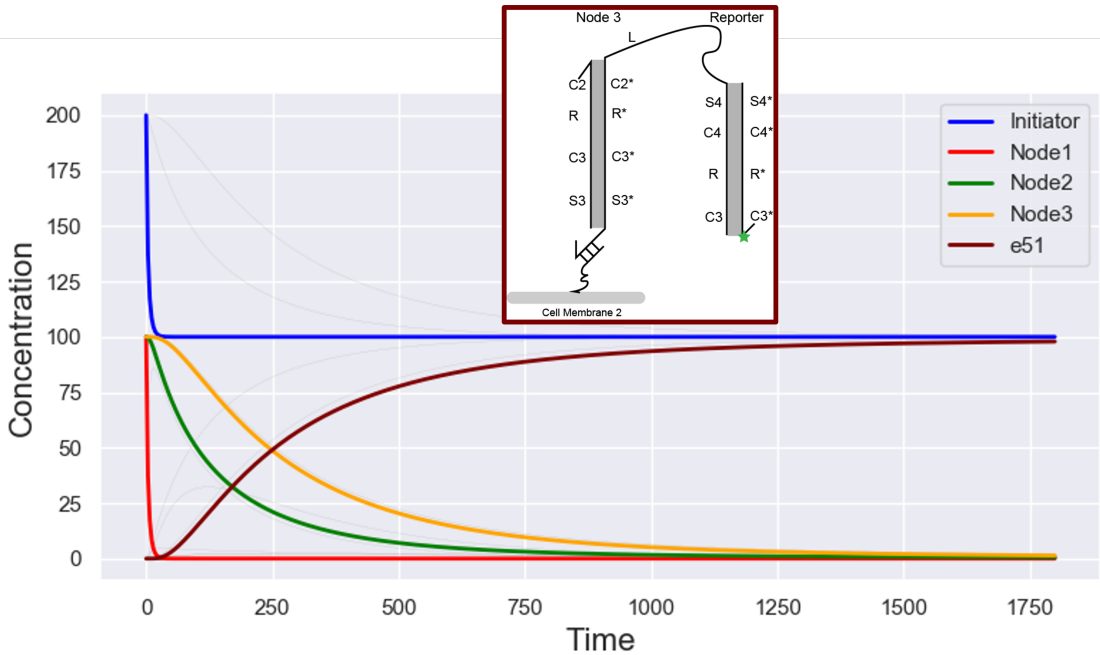

FIGURE 5.9: Simulated operation of distributed nodes H_1 (Node 1), H_2 (Node 2), and H_3 (Node 3) with manually enforced localization.

5.6 Conclusion

This chapter discussed the development of distributed DNA computing architectures and a potential application for targeting heterogeneous tumors. It is furthermore compatible with extending each locale of circuits on each surface under the same architecture as was previously demonstrated when operating 2- or 3- input chain reactions on a single cell type. The signal strand is straightforwardly designed with existing domains on each hairpin node, so thus existing domains do not need to be modified. Previous circuits can therefore be split freely to identify targets on multiple surfaces, then communicate their results. This architecture not only serves as a scalable strategy towards hierarchically combining multiple measurements but could also model cell-cell communication or, in general, the passing of chemical information by biological agents.

6

Conclusions

This thesis discussed the relevance of curved, enclosed, capsule-like DNA origami structures to applications spanning medicine, small-volume chemistry, and nanofabrication. It next described the development of heuristic optimization algorithms as an automation strategy for any design technique bearing similarity to the DNA origami technique of designing DNA nanostructures. These were first applied to simple, lattice packed structures. Then, they were applied towards constructing curved, enclosed, capsule-like DNA origami. The problems therein necessitated a degree of automation. Due to the lack of a repeatable local unit, a human designer would then be saddled with determining preferred placements for each of hundreds of crossovers, individually. Mistakes would be extremely tedious to resolve, if noticed at all. The greey and simulated annealing heuristic techniques were demonstrated to generate both single and double-layer structures to improve the yield and stability of specific geometries. Finally, a side project related to the development of a cell surface-based distributed computing architecture was discussed in the context of detecting heterogeneous tumors.

Overall, this thesis has demonstrated the propensity for various areas of DNA

nanotechnology to absorb increasingly modern principles of computer science. As the progress of the DNA nanotechnology field begins to rapidly accelerate, so do the related complexities of all its mechanisms and architectures. To continue to wrangle these back to human-processable states, DNA nanotechnology is undergoing a reorganization, where many of the lowest-level abstractions of the design of DNA-based nanodevices such as specifically placing the locations of staples strands or determining their letter sequence are being consolidated into higher-level abstractions. As this trends continues, the parting words of this thesis endeavor that some day, an abstraction hierarchy can extend far enough upwards that DNA nanotechnologies may open the doors towards wide interaction with the nanoscale.

Bibliography

[1] D. Fu, J. Reif, *Applied Sciences* **11**, 2624 (2021).

[2] N. C. Seeman, *Journal of theoretical biology* **99**, 237 (1982).

[3] E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **394**, 539 (1998).

[4] P. W. Rothemund, *Nature* **440**, 297 (2006).

[5] R. Veneziano, *et al.*, *Science* **352**, 1534 (2016).

[6] Y. Ke, L. L. Ong, W. M. Shih, P. Yin, *science* **338**, 1177 (2012).

[7] H. Dietz, S. M. Douglas, W. M. Shih, *Science* **325**, 725 (2009).

[8] S. M. Douglas, *et al.*, *Nature* **459**, 414 (2009).

[9] G. Yao, *et al.*, *Nature chemistry* **12**, 1067 (2020).

[10] D. Han, *et al.*, *Science* **332**, 342 (2011).

[11] R. Iinuma, *et al.*, *Science* **344**, 65 (2014).

[12] F. Zhang, *et al.*, *Nature nanotechnology* **10**, 779 (2015).

[13] E. S. Andersen, *et al.*, *Nature* **459**, 73 (2009).

[14] S. Dey, *et al.*, *Nature Reviews Methods Primers* **1**, 1 (2021).

[15] Y. Zhao, *et al.*, *Nano Today* **26**, 123 (2019).

[16] Y. Ke, C. Castro, J. H. Choi, *Annual review of biomedical engineering* **20** (2018).

[17] H. Ramezani, H. Dietz, *Nature Reviews Genetics* **21**, 5 (2020).

[18] M. DeLuca, Z. Shi, C. E. Castro, G. Arya, *Nanoscale Horizons* **5**, 182 (2020).

[19] L. L. Ong, *et al.*, *Nature* **552**, 72 (2017).

[20] K. F. Wagenbauer, C. Sigl, H. Dietz, *Nature* **552**, 78 (2017).

[21] J. Zheng, *et al.*, *Nature* **461**, 74 (2009).

[22] T. Zhang, *et al.*, *Advanced Materials* **30**, 1800273 (2018).

[23] G. Tikhomirov, P. Petersen, L. Qian, *Nature* **552**, 67 (2017).

[24] H. Jun, *et al.*, *ACS nano* **13**, 2083 (2019).

[25] H. Chandran, N. Gopalkrishnan, B. Yurke, J. Reif, *Journal of The Royal Society Interface* **9**, 1637 (2012).

[26] Z. Zhang, Y. Yang, F. Pincet, M. C. Llaguno, C. Lin, *Nature chemistry* **9**, 653 (2017).

[27] Y. Tian, *et al.*, *Nature Materials* **15**, 654 (2016).

[28] Y. Tian, *et al.*, *Nature Materials* **19**, 789 (2020).

[29] L. Shani, *et al.*, *Nature communications* **11**, 1 (2020).

[30] C. Sigl, *et al.*, *Nature materials* **20**, 1281 (2021).

[31] T. Bayrak, *et al.*, *Nano Letters* **18**, 2116 (2018).

[32] E. Benson, *et al.*, *Nature* **523**, 441 (2015).

[33] H. Lodish, *et al.*, *Molecular cell biology*, vol. 7 (WH Freeman New York, 2013).

[34] P. Tanner, *et al.*, *Accounts of chemical research* **44**, 1039–1049 (2011).

[35] C. Xu, S. Hu, X. Chen, *Materials Today* **19**, 516 (2016).

[36] H. G. Franquelim, A. Khmelinskaia, J.-P. Sobczak, H. Dietz, P. Schwille, *Nature Communications* **9**, 1 (2018).

[37] S. D. Perrault, W. M. Shih, *ACS Nano* **8**, 5132 (2014).

[38] Y. Yang, *et al.*, *Nature chemistry* **8**, 476 (2016).

[39] G. Stengel, R. Zahn, F. Höök, *Journal of the American Chemical Society* **129**, 9584 (2007).

[40] M. G. Malle, *et al.*, *Nature Chemistry* **14**, 558 (2022).

[41] J. R. Burns, *et al.*, *Angewandte Chemie International Edition* **52**, 12069 (2013).

[42] J. R. Burns, E. Stulz, S. Howorka, *Nano Letters* **13**, 2351 (2013).

[43] J. R. Burns, A. Seifert, N. Fertig, S. Howorka, *Nature Nanotechnology* **11**, 152 (2016).

[44] M. Langecker, *et al.*, *Science* **338**, 932 (2012).

[45] G. Grossi, M. Dalgaard Ebbesen Jepsen, J. Kjems, E. S. Andersen, *Nature communications* **8**, 1 (2017).

[46] Y. Dong, *et al.*, *Angewandte Chemie* **130**, 2094 (2018).

[47] V. Linko, M. Eerikäinen, M. A. Kostiainen, *Chemical communications* **51**, 5351 (2015).

[48] Z. Zhao, *et al.*, *Nature Communications* **7**, 10619 (2016).

[49] S. Helmi, C. Ziegler, D. J. Kauert, R. Seidel, *Nano letters* **14**, 6693 (2014).

[50] W. Sun, *et al.*, *Science* **346**, 1258361 (2014).

[51] Z. Zhao, *et al.*, *Journal of the American Chemical Society* **140**, 10639 (2018).

[52] S. F. Wickham, *et al.*, *Nature communications* **11**, 1 (2020).

[53] S. M. Douglas, I. Bachelet, G. M. Church, *Science* **335**, 831 (2012).

[54] S. Li, *et al.*, *Nature Biotechnology* **36**, 258 (2018).

[55] F. Benn, *et al.*, *Angewandte Chemie* (2018).

[56] Y. Ke, *et al.*, *Journal of the American Chemical Society* **131**, 15903 (2009).

[57] Y. Ke, *et al.*, *Nano letters* **9**, 2445 (2009).

[58] C. Lee, J. Y. Lee, D.-N. Kim, *Nature Communications* **8**, 2067 (2017).

[59] A. Ora, *et al.*, *Chemical Communications* **52**, 14161 (2016).

[60] S. M. Slone, C.-Y. Li, J. Yoo, A. Aksimentiev, *New Journal of Physics* **18**, 055012 (2016).

[61] R. E. Kohman, X. Han, *Chemical Communications* **51**, 5747 (2015).

[62] J. Hahn, L. Y. Chou, R. S. Sørensen, R. M. Guerra, W. M. Shih, *ACS nano* **14**, 1550 (2020).

[63] B. E. Snodin, J. S. Schreck, F. Romano, A. A. Louis, J. P. Doye, *Nucleic acids research* **47**, 1585 (2019).

[64] Y. Ke, N. V. Voigt, K. V. Gothelf, W. M. Shih, *Journal of the American Chemical Society* **134**, 1770 (2012).

[65] T. G. Martin, H. Dietz, *Nature Communications* **3**, 1 (2012).

[66] J. Chen, N. C. Seeman, *Nature* **350**, 631 (1991).

[67] P. W. Rothemund, N. Papadakis, E. Winfree, *PLoS Biol* **2**, e424 (2004).

[68] P. J. Paukstelis, J. Nowakowski, J. J. Birktoft, N. C. Seeman, *Chemistry & Biology* **11**, 1119 (2004).

[69] Z. Li, *et al.*, *Journal of the American Chemical Society* **141**, 15850 (2019).

[70] A. Kuzuya, M. Komiyama, *Chemical Communications* pp. 4182–4184 (2009).

[71] D. Han, S. Pal, Y. Liu, H. Yan, *Nature Nanotechnology* **5**, 712 (2010).

[72] T. Liedl, B. Högberg, J. Tytell, D. E. Ingber, W. M. Shih, *Nature Nanotechnology* **5**, 520 (2010).

[73] S. M. Douglas, *et al.*, *Nucleic acids research* **37**, 5001 (2009).

[74] D. Han, *et al.*, *Science* **339**, 1412 (2013).

[75] S. Williams, *et al.*, *International workshop on DNA-based computers* (Springer, 2008), pp. 90–101.

[76] H. Jun, *et al.*, *Science advances* **5**, eaav0655 (2019).

[77] J. Bush, *et al.*, *Molecules* **25**, 3386 (2020).

[78] F. A. Engelhardt, *et al.*, *ACS Nano* **13**, 5015 (2019).

[79] B. Kick, F. Praetorius, H. Dietz, D. Weuster-Botz, *Nano Letters* **15**, 4672 (2015).

[80] A. N. Marchi, I. Saaem, B. N. Vogen, S. Brown, T. H. LaBean, *Nano Letters* **14**, 5740 (2014).

[81] T. Gerling, K. F. Wagenbauer, A. M. Neuner, H. Dietz, *Science* **347**, 1446 (2015).

[82] D.-N. Kim, F. Kilchherr, H. Dietz, M. Bathe, *Nucleic Acids Research* **40**, 2862 (2011).

[83] B. E. Snodin, *et al.*, *The Journal of chemical physics* **142**, 06B613_1 (2015).

[84] C. Maffeo, J. Yoo, A. Aksimentiev, *Nucleic Acids Research* **44**, 3013 (2016).

[85] J. P. Doye, *et al.*, *arXiv preprint arXiv:2004.05052* (2020).

[86] J. Yoo, A. Aksimentiev, *Proceedings of the National Academy of Sciences* **110**, 20099 (2013).

[87] M. Kube, *et al.*, *Nature Communications* **11**, 1 (2020).

[88] X.-c. Bai, T. G. Martin, S. H. Scheres, H. Dietz, *Proceedings of the National Academy of Sciences* **109**, 20012 (2012).

[89] C. Gao, Q. Zhang, Z. Lu, Y. Yin, *Journal of the American Chemical Society* **133**, 19706 (2011).

[90] T. W. Giessen, P. A. Silver, *ACS Synthetic Biology* **5**, 1497 (2016).

[91] X. Liu, *et al.*, *Nature* **559**, 593 (2018).

[92] T. Gerling, M. Kube, B. Kick, H. Dietz, *Science Advances* **4**, eaau1157 (2018).

[93] N. V. Voigt, *et al.*, *Nature Nanotechnology* **5**, 200 (2010).

[94] N. D. Derr, *et al.*, *Science* **338**, 662 (2012).

[95] Y. Fu, *et al.*, *Journal of the American Chemical Society* **135**, 696 (2013).

[96] D. Rosenblum, N. Joshi, W. Tao, J. M. Karp, D. Peer, *Nature Communications* **9**, 1 (2018).

[97] T. Song, *et al.*, *Journal of the American Chemical Society* **141**, 16539 (2019).

[98] Q. Zhang, *et al.*, *ACS Nano* **8**, 6633 (2014).

[99] Q. Pan, *et al.*, *ACS Applied Materials & Interfaces* **12**, 400 (2019).

[100] T. Diederichs, *et al.*, *Nature Communications* **10**, 1 (2019).

[101] R. P. Thomsen, *et al.*, *Nature Communications* **10**, 1 (2019).

[102] W. Sun, *et al.*, *Science* **368**, 874 (2020).

[103] M. Raab, *et al.*, *Scientific Reports* **8**, 1 (2018).

[104] K. Iric, *et al.*, *Nanoscale* **10**, 18463 (2018).

[105] R. R. Sinden, *DNA structure and function* (Gulf Professional Publishing, 1994).

[106] Y. Dong, *et al.*, *Angewandte Chemie International Edition* **56**, 1586 (2017).

[107] C. Has, P. Sunthar, *Journal of Liposome Research* **30**, 336 (2020).

[108] I. G. Denisov, S. G. Sligar, *Nature Structural & Molecular Biology* **23**, 481 (2016).

[109] X. Bian, Z. Zhang, Q. Xiong, P. De Camilli, C. Lin, *Nature Chemical Biology* **15**, 830 (2019).

[110] M. W. Grome, Z. Zhang, F. Pincet, C. Lin, *Angewandte Chemie International Edition* **57**, 5330 (2018).

[111] H. G. Franquelim, H. Dietz, P. Schwille, *Soft Matter* **17**, 276 (2021).

[112] A. Johnson-Buck, S. Jiang, H. Yan, N. G. Walter, *ACS Nano* **8**, 5641 (2014).

[113] E. Ambrosetti, *et al.*, *Nature Nanotechnology* pp. 1–11 (2020).

[114] T. G. Martin, *et al.*, *Proceedings of the National Academy of Sciences* **113**, E7456 (2016).

[115] H. Zhang, *et al.*, *Nature Communications* **8**, 1 (2017).

[116] Y. Ke, S. Lindsay, Y. Chang, Y. Liu, H. Yan, *Science* **319**, 180 (2008).

[117] V. Korpelainen, V. Linko, J. Seppä, A. Lassila, M. A. Kostiainen, *Measurement Science and Technology* **28**, 034001 (2017).

[118] F. C. Zanacchi, *et al.*, *Nature Methods* **14**, 789 (2017).

[119] J. J. Schmied, *et al.*, *Nano Letters* **13**, 781 (2013).

[120] M. B. Dickerson, K. H. Sandhage, R. R. Naik, *Chemical Reviews* **108**, 4935 (2008).

[121] A. P. Alivisatos, *et al.*, *Nature* **382**, 609 (1996).

[122] C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **382**, 607 (1996).

[123] E. Dujardin, C. Peet, G. Stubbs, J. N. Culver, S. Mann, *Nano Letters* **3**, 413 (2003).

[124] M. T. Kumara, B. C. Tripp, S. Muralidharan, *Chemistry of Materials* **19**, 2056 (2007).

[125] R. Schreiber, *et al.*, *Nature communications* **4**, 1 (2013).

[126] Q. Jiang, *et al.*, *Nano Letters* **17**, 7125 (2017).

[127] Y. Huang, M.-K. Nguyen, A. K. Natarajan, V. H. Nguyen, A. Kuzyk, *ACS Applied Materials & Interfaces* **10**, 44221 (2018).

[128] X. Shen, *et al.*, *Nano Letters* **13**, 2128 (2013).

[129] X. Lan, *et al.*, *Journal of the American Chemical Society* **137**, 457 (2015).

[130] C. Zhou, X. Duan, N. Liu, *Nature Communications* **6**, 1 (2015).

[131] P. Zhan, *et al.*, *ACS Nano* **11**, 1172 (2017).

[132] A. Kuzyk, *et al.*, *Nature Communications* **7**, 1 (2016).

[133] W. Yan, *et al.*, *Journal of the American Chemical Society* **134**, 15114 (2012).

[134] M. Pilo-Pais, A. Watson, S. Demers, T. LaBean, G. Finkelstein, *Nano Letters* **14**, 2099 (2014).

[135] P. Zhan, *et al.*, *Angewandte Chemie International Edition* **57**, 2846 (2018).

[136] P. Wang, *et al.*, *Journal of the American Chemical Society* **138**, 7733 (2016).

[137] C. Zhu, M. Wang, J. Dong, C. Zhou, Q. Wang, *Langmuir* **34**, 14963 (2018).

[138] S. Sun, *et al.*, *Nature Communications* **11**, 1 (2020).

[139] W. Liu, *et al.*, *Science* **351**, 582 (2016).

[140] R. Schreiber, I. Santiago, A. Ardavan, A. J. Turberfield, *ACS Nano* **10**, 7303 (2016).

[141] M. Zhao, *et al.*, *Science* **368**, 878 (2020).

[142] S. Jia, *et al.*, *Nature Communications* **10**, 1 (2019).

[143] B. Shen, *et al.*, *Science Advances* **4**, eaap8978 (2018).

[144] X. Chen, *et al.*, *ACS applied materials & interfaces* **10**, 24344 (2018).

[145] F. Hong, F. Zhang, Y. Liu, H. Yan, *Chemical reviews* **117**, 12584 (2017).

[146] Y. Ke, G. Bellot, N. V. Voigt, E. Fradkov, W. M. Shih, *Chemical Science* **3**, 2587 (2012).

[147] D. Fu, *et al.*, *Science Advances* **Accepted with revisions** (2022).

[148] C.-M. Huang, A. Kucinic, J. A. Johnson, H.-J. Su, C. E. Castro, *bioRxiv* (2020).

[149] W. J. Hagen, W. Wan, J. A. Briggs, *Journal of structural biology* **197**, 191 (2017).

[150] D. N. Mastronarde, *Journal of structural biology* **152**, 36 (2005).

[151] J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, *Journal of structural biology* **116**, 71 (1996).

[152] L. Rovigatti, P. Šulc, I. Z. Reguly, F. Romano, *Journal of computational chemistry* **36**, 1 (2015).

[153] E. Poppleton, *et al.*, *Nucleic acids research* **48**, e72 (2020).

[154] J. Bohlin, *et al.*, *Nature Protocols* pp. 1–27 (2022).

[155] C. R. Lucas, *et al.*, *Small* p. 2108063 (2022).

[156] H. Bui, *et al.*, *ACS nano* **12**, 1146 (2018).

[157] A. D. Ellington, J. W. Szostak, *nature* **346**, 818 (1990).

[158] T. A. Denison, Y. H. Bae, *Journal of controlled release* **164**, 187 (2012).

[159] S. Bhatia, J. V. Frangioni, R. M. Hoffman, A. J. Iafrate, K. Polyak, *Nature biotechnology* **30**, 604 (2012).

[160] M. Kim, D.-M. Kim, K.-S. Kim, W. Jung, D.-E. Kim, *Molecules* **23**, 830 (2018).

[161] K. Min, *et al.*, *Biomaterials* **32**, 2124 (2011).

[162] H. Jo, J. Her, C. Ban, *Biosensors and Bioelectronics* **71**, 129 (2015).

[163] S. M. Taghdisi, *et al.*, *European Journal of Pharmaceutics and Biopharmaceutics* **102**, 152 (2016).

[164] S. Badelt, *et al.*, *Journal of the Royal Society Interface* **17**, 20190866 (2020).

Biography

Daniel Fu received his B.Asc. degree in Electrical and Computer Engineering from the University of British Columbia in May 2016. He received his Ph.D. in Computer Science from Duke University in December 2022.