Dynamic self-assembling DNA nanosystems: design and engineering

Divita Mathur
Iowa State University
Dynamic self-assembling DNA nanosystems: Design and engineering

by

Divita Mathur

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Major: Bioinformatics and Computational Biology

Program of Study Committee:
Eric R. Henderson, Co-major Professor
Jack H. Lutz, Co-major Professor
Robyn R. Lutz
Sanjeevi Sivasankar
Drena L. Dobbs

Iowa State University
Ames, Iowa
2016

Copyright © Divita Mathur, 2016. All rights reserved.
DEDICATION

This work is dedicated to my grandmother, Vimala Saraf. Thank you, for your DNA.
TABLE OF CONTENTS

LIST OF TABLES ......................................................... vi
LIST OF FIGURES ....................................................... vii
ACKNOWLEDGEMENTS ................................................... x
ABSTRACT .............................................................. xii
INTRODUCTION .......................................................... 1
CHAPTER 1. BACKGROUND AND LITERATURE REVIEW . 4
  1.1 Deoxyribonucleic acid, the nanoscale building block ............. 4
  1.2 DNA crossover .................................................... 9
  1.3 DNA tile self-assembly .......................................... 11
  1.4 Scaffolded DNA origami ......................................... 12
  1.5 Non-scaffolded DNA “origami” .................................. 15
  1.6 Quick guide to designing a DNA origami nanostructure ........... 17
  1.7 Experimental protocols for DNA self-assembly .................. 18
  1.8 Experimental validation and purification of DNA nanostructures ... 20
  1.9 Molecular recognition ............................................ 21
  1.10 Molecular programming .......................................... 26
  1.11 Molecular actuation and reconfigurable nanodevices ........... 28
CHAPTER 2. COMPLEX DNA NANOSTRUCTURES FROM OLIGONUCLEOTIDE ENSEMBLES .......................... 35

2.1 Abstract ............................................................... 35
2.2 Introduction ......................................................... 36
2.3 Results ................................................................. 37
2.4 Discussion ............................................................. 39
2.5 Methods ................................................................. 40
2.6 Figures ................................................................. 44

CHAPTER 3. PROGRAMMABLE DNA NANOSYSTEM FOR MOLECULAR INTERROGATION ......................... 49

3.1 Abstract ............................................................... 49
3.2 Introduction ......................................................... 50
3.3 Results ................................................................. 51
3.4 Discussion ............................................................. 57
3.5 Methods ................................................................. 58
3.6 Figures ................................................................. 63

CHAPTER 4. EXPERIMENTAL VERIFICATION OF COMPUTATIONAL MODELING AND ANALYSIS RESULTS ... 68

4.1 Introduction .......................................................... 68
4.2 Molecular Watchdog Timer ........................................ 69
4.3 Capturing and reusing design strategies ......................... 72
4.4 Design and assembly of DNA origami with multiple states .... 75
4.5 Development of a computational tool to automate the designing of scaples-based DNA nanostructures ....................... 78
4.6 Figures ................................................................. 81
LIST OF TABLES

Table 1.1  Design strategies to make complex DNA origami nanostructures  14
Table 1.2  Alternative Scaffolds  . . . . . . . . . . . . . . . . . . . . .  16
LIST OF FIGURES

Figure 1.1 Basic structure of a DNA molecule. ........................... 5
Figure 1.2 DNA domains and basic assembly. ........................... 6
Figure 1.3 DNA-DNA interactions. ................................. 8
Figure 1.4 A schematic, illustrating the concept of DNA crossovers and complex two- and three-dimensional architectures. .......... 10
Figure 1.5 Design of DNA origami and some of the seminal work. ....... 30
Figure 1.6 Single-stranded tile or brick technology. ..................... 31
Figure 1.7 Addressibility of DNA nanostructures. ....................... 32
Figure 1.8 Application of DNA nanostructures as sensing and molecular analyses platforms. ...................................... 33
Figure 1.9 Example of a simple DNA computing reaction. ............... 33
Figure 1.10 DNA nanosystem with programmable reconfiguration. ...... 34
Figure 2.1 Strategy for the design of scaples-based nanostructures. ....... 44
Figure 2.2 Scaples-based triangle using m13mp18 as the design-scaffold. 45
Figure 2.3 Scaples-based triangle constructed using a random sequence design-scaffold. ........................................ 46
Figure 2.4 Simultaneous assembly of three DNA nanostructures in a "one-pot" reaction. ........................................... 47
Figure 2.5 A graphic rendering of triangle structure used in this study. . 48
| Figure 3.1 | OPTIMuS operational principles. | 63 |
| Figure 3.2 | System characterization and demonstration using FRET and Transmission Electron Microscopy (TEM). | 64 |
| Figure 3.3 | Base stacking adhesive forces versus OPTIMuS actuators. | 65 |
| Figure 3.4 | Base Stacking Adhesive Forces versus OPTIMuS actuators: TEM analysis. | 66 |
| Figure 3.5 | Programming OPTIMuS for nucleic acid detection. | 67 |
| Figure 4.1 | DNA strands used in the preliminary MWT experiment. | 81 |
| Figure 4.2 | Fluorescence output received from the single DSD reaction over the course of 300 seconds. | 82 |
| Figure 4.3 | A design flaw and the corrected version. | 83 |
| Figure 4.4 | Design failure: unfavorable aggregation. | 84 |
| Figure 4.5 | Mathematical models for depicting DNA nanostructures, and their relationship. | 85 |
| Figure 4.6 | Design strategy to create a rectangle with two isomeric states. | 86 |
| Figure 4.7 | Atomic force microscope images of rectangle and its two isomeric states. | 87 |
| Figure 4.8 | Basic outline of the scaples design strategy. | 88 |
| Figure 4.9 | Scoring scheme employed to identify optimum nick points on the design-scaffold. | 88 |
| Figure A.1 | Design of the scaples-based triangle, made using a random sequence as the design-scaffold. | 100 |
| Figure A.2 | Agarose gel electrophoresis (1% agarose in 1xTAEM) of the origami triangle and the scaples-based triangle made using a randomly generated design-scaffold. | 101 |
Figure B.1  Assembly of G in the presence of EC staples(preadded condition). ........................................... 103
Figure B.2  Scaffold raster layout. ........................................... 104
Figure B.3  Role of EC domain in ring movement. .................... 108
Figure B.4  caDNAno design of OPTIMuS: Ground state. ............ 109
Figure B.5  caDNAno design of OPTIMuS: G+EC (staples). ........... 110
Figure B.6  caDNAno design of OPTIMuS: G+EC staples +Ch staples. 111
Figure B.7  caDNAno design of OPTIMuS: G+EC staples +L staples. . 112
Figure B.8  Agarose gel electrophoresis analysis of OPTIMuS. ........ 113
Figure B.9  Fluorophore placement on OPTIMuS. ...................... 113
Figure B.10 GC/GC blunt end active site. .......................... 114
ACKNOWLEDGEMENTS

I would like to thank my advisor, Eric Henderson, for his immeasurable support, as a mentor and a friend. If it were not for his trust, training and encouragement to be independent and fearless, I would not have made novel mistakes and would certainly not have learnt the lessons that will remain with me for life. If it were not for our countless, yet memorable, lab meetings at coffee houses around the city, I might have taken longer to barely scratch the surface of research, science, and "other facets of existence". Lastly, if it were not for his financial support, my second chapter would have remained a mere idea, awaiting the blessings of funding agencies.

I would like to mention my sincere gratitude to Jack Lutz, Robyn Lutz, James Lathrop and the rest of the Laboratory for Molecular Programming (LAMP) group for having me and giving me a great environment that was conducive to scientific development and discussions about DNA. Special thanks to Brian Nakayama, for endless conversations about research and life, and Sam Ellis, Titus Klinge and Don Stull for sharing their interesting research with me.

I feel very fortunate to have a program of study committee that was always available for support and would like to thank Drena Dobbs and Sanjeevi Sivasankar for their time and guidance, that has enabled me to become a better academic researcher. Parts of my projects would have been very difficult and challenging to execute without the support of Curtis Mosher, who was instrumental in my training in microscopy.

Over the course of my doctoral research, I have met some very smart people, who have influenced my work and personality, and am truly indebted to them for
their company, support and friendship. Thanks, to Rahul Roy, Ali Berens, Divya Mistry, Sweta Roy, Amber Roy, Pooja Gupta Saraf, and others, for countless words of encouragement; for, in their unique ways, assuring me that I am not an imposter, and for letting me borrow chemicals and paper towels from their labs. For some more coffee, and chocolate shakes.

Words fail me in thanking my family, the Saraf, Mathur and Sangar clans, who let me choose a course of life for myself, and always provided me with unwavering support and love. I am especially grateful to the women in my life - my grandmother, mother, aunts and sisters - who were the best role models I could have asked for. I do not know of a stronger group of people. In some ways, this thesis is also dedicated to my father, who has been a constant support, in spirit, and in his favorite words on strength, courage and integrity.

And finally, I would like to express my heartfelt gratitude to Nikhil Panicker, for his emotional support through every kind of obscenity the past six years of our lives have thrown at us, through the good times too, and for letting me learn ever so much from him. For bringing out the best in me.
ABSTRACT

Over the last few decades, DNA has proven to be a great candidate for engineering nanoscale architectures. These DNA nanostructures have been applied in areas such as single-molecular analyses, nanopatterning, diagnostics and therapeutics. One of the most commonly-used techniques to engineer DNA-based two- and three-dimensional functional nanostructures is DNA origami, wherein a long single-stranded DNA (called scaffold) is folded into a predetermined shape with the help of a set of shorter oligonucleotides (called staples). This thesis discusses a brief overview of DNA nanotechnology (design, assembly and applications) and three primary projects undertaken in the area of dynamic self-assembling DNA nanosystems: 1, a self-assembly design strategy that vastly expands the utility of DNA origami, 2, a DNA origami-based reconfigurable nanosystem with potential as a force/energy balance and diagnostic tool, and 3, a collaborative initiative on computational analyses and experimental verification for improving efficiency of DNA nanoengineering.
INTRODUCTION

It is well known that the Deoxyribonucleic Acid (DNA) is the primary bearer of genetic information for almost all living species. It contains the rules and instructions required to create such incredibly diverse organisms from the same precursor biomolecules. The year 1953 proved to be very enlightening to scientists, with the first glimpse into the physical structure of the DNA. And in another three decades, we were able to appreciate putting strand and strand together to create something with zero genetic relevance, but immense architectural significance - a DNA nanostructure. A square, or triangle, or even a smiley face, that is roughly 10,000 times smaller than the width of an eyelash. A collection of DNA strands orchestrating themselves to form a physical nano-sized architecture, with minimum external control, like a pile of bricks and mortar self-assembling into a whole building! It was an invitation for engineers and computer scientists to roll their sleeves up and work abreast biologists in a laboratory.

As problem solvers, mankind strives to find a solution, and eventually build a machine that automates the process of addressing a recurring problem. Machines, such as a microscope, also help us discover the unknown, or, like rockets, literally achieve new heights that are beyond our reach as humans. In the interest of discovering the world at the nanoscale, size-comparable nanomachines are required that are easy to design and create (using our macro-sized fingers or existing machines), are fairly robust but preferably non-toxic, and elicit an observable response upon reacting to a signal. In a way, this forms the central motivation towards using DNA as a building
material to make nanoscale devices and probes, called DNA nanotechnology.

This doctorate thesis is an earnest attempt to understand the history, motivation, and applications of synthetic DNA nanotechnology and contribute to its ever-growing expertise. A brief description of the layout of the thesis is as follows:

**Chapter 1.** Background and literature review: This chapter will be a comprehensive overview of the area of synthetic DNA self-assembly (or DNA nanotechnology), mainly, introduction, strategies of assembly and applications.

**Chapter 2.** Complex DNA nanostructures from oligonucleotide ensembles: This chapter describes published work on a novel design-strategy to create DNA-based nanostructures using short oligonucleotides. My contribution to the paper includes, design and execution of experiments as well as co-writing the manuscript.

**Chapter 3.** Programmable DNA nanosystem for molecular interrogation: This chapter describes a manuscript (under review) on the engineering of an autonomous DNA nanosystem for studying molecular interactions as well as detect the presence of nucleic acid species in a solution. My contribution to the work includes, design and execution of experiments, and co-writing the manuscript.

**Chapter 4.** Experimental verification of computational modeling and analysis results: This chapter summarises collaborative work with the Laboratory for Molecular Programming in DNA nanotechnology. My contribution to the work includes, experimental verification of computational simulations, development of a novel algorithm for DNA self-assembly and execution of a reconfigurable DNA nanosystem.
Chapter 5. Conclusions and future directions: The final chapter delineates potential avenues for future work in each project that is described in the previous chapters, as well as concluding words about the general future directions.

Appendices. Supporting material corresponding to chapters 2 and 3.

Bibliography. List of references that are cited in the entire dissertation.

Support. The work performed in the thesis was supported by the Bill and Melinda Gates Foundation (grant no. OPP1059568), National Science Foundation (grant no. 1143830, 1247051, and 1545028), Creodyne LLC, departmental funds (Genetics, Development and Cell Biology, and Bioinformatics and Computational Biology), and personal funds.
CHAPTER 1. BACKGROUND AND LITERATURE REVIEW

DNA nanotechnology is a combination of DNA, nanometer ($1 \times 10^{-9}$meters), and technology. It may sound superfluous a term, because every thing pertaining to the molecule, DNA, is already at the nanoscale. However, the emphasis in this area of research is on using DNA has the primary ingredient for nanotechnology, in place of other organic or inorganic compounds. The genetic code becomes a computational or an engineering code, or as the curved edges of jigsaw puzzles that come together into objects with precise properties. And these objects can serve as new technology. But, no genes are involved or harmed in this process.

The goal of this chapter is to provide some key definitions that will be useful in understanding the work presented in the entire dissertation, and act as a quick reference guide to the design, assembly and applications of synthetic DNA nanotechnology.

1.1 Deoxyribonucleic acid, the nanoscale building block

Deoxyribonucleic Acid (DNA) is the genetic code of life. Structurally first characterized in 1953 by Franklin, Watson and Crick[1] (Fig. 1.1), it is a polymeric chain of nucleotides, wherein each nucleotide has three chemical components: a sugar molecule with a hydroxyl group, phosphoric acid, and one of the four bases – Adenine (A), Thymine (T), Guanine (G) or Cytosine (C). The polymeric chain, thus,
is polar in nature, with a hydroxyl group of the last sugar at one end (the 3’ end) and a phosphate group at the other (the 5’ end). In its predominant form, DNA strands conform into a double helix, made from two constituent DNA strands, each turn adopts a diameter of 2 nanometers (nm), height of 3.4 nm, and contains roughly 10.5 bases (or nucleotides). The two strands bind to each other using highly specific hydrogen bonds, called the Watson-Crick base pairing (bp), in which A and T bind with two hydrogen bonds, and G and C form three hydrogen bonds. In this case, the strands are said to be “complementary” to each other.

![Figure 1.1 Basic structure of a DNA molecule.](image)

a. A coarse model of the DNA and its physical dimensions. The Watson-Crick base pairing occurs between A and T (two bonds), and G and C (three bonds). b. Abstraction of a DNA into a linear “ladder” format, that shows base pairing and orientation of the two strands. In DNA nanotechnology, the physical properties of a DNA double helix translate into a nanoscale rod.

DNA is one of the most stable biomolecules, which is a consequence of its overall structure and chemical makeup. Unlike its ribonucleic counterpart, Ribonucleic Acid (RNA), DNA sugar molecules do not contain reactive oxygen atoms that are susceptible to degradation. DNA double helical formation also protects the bases from
exposure to water and hydrophobic environments. Additionally, the Watson-Crick base pairing of A-T and G-C is highly specific and predictable, a quality that is very challenging to achieve using amino acids, which are the building blocks of proteins. Hence, among the three primary biomolecules - DNA, RNA and protein - DNA is the most structurally robust, chemically predictable and, therefore, amenable to act as a physical building material for nanoscale objects.

Apart from base pairing, DNA demonstrates some other binding properties that are of interest in the area of DNA nanotechnology, described below.

1.1.1 DNA domains

![Figure 1.2 DNA domains and basic assembly.](image)

a. Two complementary strands, namely A and B, form a DNA double helix, also represented as a linear rod. b. Three strands, A, B and C, with domains complementary to each other, will hybridize to form a three-arm junction, or a DNA “tile”. In a similar fashion, higher order branched tiles and lattices can be created using DNA oligonucleotides. While designing such structures, different domains should either be complementary or orthogonal in nature to avoid efficient binding.
If there are two strands of DNA in a solution – A and B, each 20 bases long, that are a 100% complementary to each other, the strands would bind with extremely high probability in a suitable buffer environment, resulting in one double helical structure (Fig. 1.2a). If there are three strands, wherein a half of each strand is complementary to a half of another in a cyclic permutation, then the three strands would combine to form a “three-armed” double helical structure. In this case, we can say that each strand has two domains of sequences that are complementary to domains or subsequences in other strands. Following the concept of domains in a strand of DNA, it is possible to design optimal and unique strands to create a set of self-assembling DNA molecules. With some encouragement in the form of favorable buffer conditions (such as ions - Magnesium or Sodium) and temperature, DNA self-assembling can result in simple tiles or even highly sophisticated nanorobots.

1.1.2 DNA blunt ends and sticky ends

In a DNA duplex, there are a variety of forces that are non-Watson-Crick in nature. The bases - A, T, C, G - are inherently hydrophobic (or water repelling) and therefore adapt a conformation that excludes water molecules; hence the double helical shape. Their common hydrophobicity arises from the pi-bonds in their carbon-nitrogen rings, which also creates an attractive force between adjacently stacked bases in a duplex (a like-like attractive force)[2]. A blunt end is the end of a DNA duplex that is completely hybridized and, therefore, contains no overhanging domains like sticky-ended duplexes (Fig. 1.3a, b). A blunt-ended duplex, with no other single-stranded bases in the middle, is unreactive to other DNA strands as a consequence of the lack of available unhybridized bases and is its lowest energy state. However, blunt ends of duplexes can elicit a strong attractive force towards proximal blunt ends of other DNA molecules (Fig. 1.3c). This property of DNA has been utilized in DNA-based
a. The ends of a fully hybridized DNA double helix are said to be blunt ends, due to the absence of unpaired bases. b. A sticky ended DNA double helix contains at least one base that is not base-pairing with a complementary base. c. Blunt ends of two double helices interact with each other via adhesive base stacking forces. d. DNA strand displacement reaction between an invader strand (green) and a toehold-bearing duplex leads to an exchange of strands between the two systems.

nanostructures to create higher order assemblies[3, 4].

1.1.3 Toehold-mediated DNA strand displacement

DNA strand displacement (DSD) is a particular kind of DNA reaction, in which one strand displaces another strand from a DNA duplex (Fig. 1.3d). The invading strand binds to a short single-stranded overhang of the duplex, called a toehold, and causes the initiation of DNA hybridization, which is followed by the migration of the
invading strand along the duplex, an intermediate state where the invading strand competes with the original strand for base pairing, and eventually the displacement of the original strand. Toehold domains are typically of the order of 3 bases to 8 bases long; the length and sequence guides the overall kinetics of the reaction[5, 6]. DSD reactions can be used to engineer dynamic and reconfigurable DNA nanosystems that interact with the environment as well as with other nanostructures, by accepting an input signal, undergoing a physical state change, and releasing an output signal. It is also possible to create reactions of higher complexity, that involve more than one toehold and the exchange of several strands using DSD.

1.2 DNA crossover

DNA nanotechnology, arguably, is an engineering concept at its core. Engineering, whereas, is inspired to a great extent by Nature and natural “machines”, such as enzymes. For instance, homologous recombination is key to genetic diversity in living organisms and contributes to evolution. In homologous recombination, there is crossing over of two strands to form a Holliday Junction[7, 8] and subsequent exchange of segments of DNA between two larger pieces of DNA (or chromosomes). The basic building block of self-assembling DNA nanostructures borrows form this ubiquitous concept of Holliday junction – a crossover.

A crossover in DNA nanotechnology, as illustrated in Fig. 1.4b, is a point where a DNA strand is shared between two duplexes. The implication of a crossover at the sequence level is the point in the DNA backbone between two sequence domains whereas at the structural level it represents a site of physical connection between two DNA duplexes through a common single-stranded DNA.
Figure 1.4  A schematic, illustrating the concept of DNA crossovers and complex two- and three-dimensional architectures.

a. Geometric details of a DNA duplex.  b. An example of a DNA crossover between two parallel DNA helices, the most common motif in DNA nanostructures.  c. Crossovers in a honeycomb lattice in 3D DNA structures.  d. Crossovers in a close-packed square lattice in 3D DNA structures.  e. The accessible angles at 1 base intervals for flexible out-of-plane DNA crossovers.  f. In-plane and bending of DNA duplexes by inserting unequal base pair intervals between the participating helices.  

Copyright © 2016, John Wiley and Sons[9]

To take a closer look at a crossover, consider building a raft using linear wooden logs and pieces of string. To create a flat-surfaced raft, it is important to start winding the string on one log, then a second one, and cross over to the third log such that it is coplanar to the first two. The point on the string where it transfers from one log to another is a crossover. The string is wound up to the top or bottom surface of the raft before it crosses over to the third log. In other words, the point of crossover for the string is at a point that favors planarity of the raft.
Similarly, crossovers in DNA nanostructures decide the planarity or shape of the resulting structure (Fig. 1.4f). There are approximately 10.5 bases per turn in a DNA double helix, which means that a 360-degree turn contains 10.5 bases. With the help of this information, crossovers can be designed with a resolution of a single base, or roughly 34° rotation to impart precision to the resultant 3D nanostructures (Fig. 1.4e). Using this concept, three major lattices are used as templates to create two- and three-dimensional DNA nanoshapes: square lattice (Fig. 1.4d), honeycomb lattice (Fig. 1.4c) and hexagonal lattice, or even a combination of the three[10].

### 1.3 DNA tile self-assembly

In 1982, Nadrian Seeman proposed that it is possible to capitalize on the above-mentioned properties of DNA, in the form of short synthetic oligonucleotides, to create two-dimensional tiles and three-dimensional lattices[11]. Seeman, seeking inspiration from a Holliday junction, created the first DNA tile – a set of 4 strands that contain sticky-end domains and complementary domains to bind to each other. The sticky-end domains further interact with other tiles in the solution to “grow” into larger lattices of DNA tiles[12].

In 1996, Erik Winfree proposed the abstraction of a DNA tile into a four-edged tile, in which each edge has a certain degree of “stickiness” (or reactivity) based on the length of its reactive DNA domain, and tiles of the same degree of stickiness will bind in a solution[13]. This abstraction was easy to program and made it possible to implement (theoretically or experimentally) mathematical algorithms, such as the creation of the Sierpinski triangle[14], Sierpinski carpet[15], and other fractals, using tile assembly.
A four-armed tile can be expanded to have more or fewer sticky-ended arms, or a combination of reactive and unreactive domains to design more complex tile assemblies, with the interest to arrange proteins or create conductive lattices\cite{16, 17, 18, 19} and switchable devices\cite{20}.

While tile self-assembly opened a new door for research in DNA computing (a small set of 4-5 base-long sticky ends orchestrating the hybridization of hundreds and thousands of oligonucleotides into predetermined patterns!), the technique suffered from issues that made other applications difficult. Sometimes an arm of a tile would bind to the wrong tile; an incorrect tile would hang on to the growing structure through a couple of mismatches. The sticky ends would continue growing the structure, as far as the size or reaction equilibrium permitted. This was more of a problem than a feat in nanoscale assembly – the lack of control over the scalability of tile-based nanostructures, and the probable danger of propagating a mismatch in the nanostructure. Moreover, achieving a homogeneous mix of same-sized tile-based architectures was unthinkable\cite{21}.

### 1.4 Scaffolded DNA origami

It was understood that programming multiple orthogonal oligonucleotides to come together in tile self-assembly has too many kinetic variables to result in an ordered and controlled shape. This thought spawned the idea of using one long strand as a backbone of the nanostructure that would drive the reaction forward following 1st order kinetics – DNA origami. Paul Rothemund, in 2006, introduced the concept of DNA origami – which uses a 7kb long viral genome – m13mp18 strand as a “scaffold” that is folded into a predetermined shape by designing a pool of short and complementary oligonucleotides called “staples”\cite{22}. The backbone scaffold strand is analogous
to a paper in the technique’s namesake of Japanese paper-folding art whereas the short oligonucleotide staple strands perform the function of paper folds by binding to different domains of the scaffold and inducing a physical shape change in it. Figure 1.5a illustrates DNA origami.

The versatility and efficiency of DNA origami was immediately evident in the aesthetically appealing array of shapes demonstrated by Rothemund in his seminal publication (Fig. 1.5b). DNA origami based self-assembly generates structures with over 90% efficiency. In his earliest work, Rothemund had designed the structures on an in-house MATLAB software that computed the best crossover points as well as staple strands to yield a desirable structure. In 2009, Shawn Douglas, et. al. published a software, called caDNAno, that provided an excellent graphical user interface to create creative and complex DNA nanostructures ever imaginable by scientists, within the limitations of nanoscale structural biology[24] (Fig. 1.5d). CaDNAno is an open source CAD (Computer Aided Design) software for designing DNA origami structures. CaDNAno, for the first time, also established crossover rules using basic geometry, as described above (360° = ~10.5 bases) and created the first set of three dimensional DNA origami structures using single viral genomes as scaffolds[24, 25].

Another groundbreaking work, in reference to designing DNA origami structures, was published in 2012 by Dongran Han et. al., which introduced curved three-dimensional DNA origami[23]. By identifying crossover points that introduce curvature into a DNA helix, they showed that it is possible to program DNA helices in a structure to follow a curved path and create structures such as a flask, a funnel or a hemisphere (Fig. 1.5c). Another key point raised in their seminal work is that the pitch (bases in one turn) of a DNA molecule is a variable parameter, and, as originally assumed, not an absolute 10.5 bases/turn. DNA can accommodate over-
and under-winding of up to 9-13 bases/turn, which is an interesting feature to create curved nanostructures.

Upon these pillars of scientific discovery, a breadth of work has been performed that showcases the versatility of creating nanostructures using DNA as a building material. Table 1.1 highlights other significant design concepts in DNA origami.

Table 1.1  Design strategies to make complex DNA origami nanostructures

<table>
<thead>
<tr>
<th>Number</th>
<th>Name/summary</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Introduction of curvature in DNA origami[26, 23]</td>
<td>Can introduce targeted or global twisting.</td>
</tr>
<tr>
<td>3.</td>
<td>Scale-up DNA origami by combining multiple structures[28, 29, 30, 31]</td>
<td>To create higher-order architectures.</td>
</tr>
<tr>
<td>4.</td>
<td>Scale-up DNA origami using scaffold as the combining domains[32, 33]</td>
<td>To create higher order structures.</td>
</tr>
<tr>
<td>5.</td>
<td>DNA Gridiron - four-arm junction crossovers[34]</td>
<td>To create flexible structures and structures with porosity.</td>
</tr>
<tr>
<td>6.</td>
<td>Structure made in hybrid lattices[10]</td>
<td>The first demonstration of nanostructures wherein the scaffold-raster format was in square, honeycomb and hexagonal lattices in the same structure. They also postulated that any polygonal lattice could be employed so long as neighboring helices were antiparallel.</td>
</tr>
<tr>
<td>8.</td>
<td>DNA polyhedral meshes[36]</td>
<td>To make highly complex and versatile architectures.</td>
</tr>
</tbody>
</table>
DNA origami-based nanostructures form with over 90% efficiency (~50-80% in case of 3D structures) and the whole procedure is extremely simple and reproducible. A detailed review of assembly protocols is presented below.

The viral genome m13mp18 (7,249 bases), which is the most popular DNA to be employed as a scaffold, can yield structures of roughly 100 nm per side. The length of the scaffold dictates the overall dimensions of the resulting structure – much like the number of bricks limits the overall size of the building. Secondly, the inherent sequence or the primary structure of the nanodevice will also be the same as the primary sequence of the scaffold, which is usually the viral m13mp18. These two factors define the boundaries on the scope of DNA origami-based nanostructures or devices and reveal to us that, as a stand-alone technique, this tool may fall short in applications that absolutely require a different non-viral sequence or require scalable structures.

In the toolbox of techniques to create DNA nanostructures that, so far, contains tile-assembly, its derivatives, and DNA origami using single units of a viral scaffold, scientists began to include more strategies that further expanded the scope of synthetic DNA nanotechnology (Table 1.2).

1.5 Non-scaffolded DNA “origami”

To seek a solution to the limitations of DNA origami, scientists turned back to what had originally advanced them towards DNA origami – short oligonucleotides. In 2012, two independent groups proposed ways to find a middle ground between DNA origami and tile self-assembly by using short oligonucleotides in a format that borrows from the rastering of a scaffold[49, 50, 51] (one example shown in Fig. 1.6). While these techniques offer the option of designing structures with novel sequences, they
Table 1.2 Alternative Scaffolds

<table>
<thead>
<tr>
<th>Number</th>
<th>Name/Summary</th>
<th>Scaffold Lengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mini scaffolds produced using custom phage transfection[37, 38]</td>
<td>1000-2000 bases (1000 = 1k). Creating structures smaller than normal origami (100 x 100 nm²).</td>
</tr>
<tr>
<td>2.</td>
<td>PCR-based scaffolds[39, 40]</td>
<td>100-10k, 26k bases. Scaffolds not limited by discreet lengths of genomic scaffolds.</td>
</tr>
<tr>
<td>3.</td>
<td>Restriction fragments – enzymatically tweak genomic DNA to create shorter scaffolds[41]</td>
<td>1k-3k bases, the smallest scaffold designed to date.</td>
</tr>
<tr>
<td>4.</td>
<td>dsDNA scaffold – incorporate one or both strands of a dsDNA source as a unified scaffold[42, 43, 44]</td>
<td>10k-50k bases.</td>
</tr>
<tr>
<td>5.</td>
<td>Scaffolds straight out of phage sources[45]</td>
<td>7k, 50k.</td>
</tr>
<tr>
<td>6.</td>
<td>Lambda/m13 hybrid virus[46]</td>
<td>51466 bases, the longest scaffold designed to date.</td>
</tr>
</tbody>
</table>

Inherit the poor yields that tile assemblies suffered from. It is important, however, at this point to understand that this arsenal of techniques invented this decade will collectively, and not as stand-alone tools, form the foundation of a new wave of functional nanodevices that is to come in the future.
1.6 Quick guide to designing a DNA origami nanostructure

Like a macro-sized machine, the overall functionality of a DNA nanodevice dictates its structure. One double helix serves as a rod of length 0.34 nm/base and width of 2 nm (Fig. 1.1b). In three-dimensional space, two dimensions, say y and z, are made by stacking DNA double helices in a well-defined lattice or a network, whereas the third dimension is in terms of the number of bases in each double helix.

The dimensions of a desired shape are translated into the aforementioned units (number of bases and number of helices) while keeping in mind the boundaries set upon by the available DNA. For instance, in DNA origami, the length of the scaffold will restrict the available DNA. After designing the layout of the shape in terms of number of bases and helices, the layout is populated with a scaffold or a set of oligonucleotides in a raster format. In most cases, the scaffold layout contains one or more[52] seams, which are arrays of scaffold crossovers within the structure. A seam ensures that the raster layout brings the extremities (3’ and 5’ ends) of the scaffold close together and imparts additional stability to the structure. This often is important, because the commercially available scaffolds are genomic in nature, which (unless treated with enzymes to create a nick) are circular in nature. For simple 3D architectures in the square and honeycomb lattices, caDNAno has automated the process of identifying the crossover points that would be most suitable.

After defining the scaffold path, staples are introduced to add additional crossovers in the structure that would maintain the scaffold in its raster-style layout. A staple would have 2-4 domains that are complementary to disparate domains on the scaffold and, hence, bring them into close physical proximity upon hybridization. The staple strands are most favorably 20 to 60 nucleotides (nt) long, a range settled upon by scientists largely based on the current DNA synthesizing capacity of synthetic DNA.
providers. The requisite pool of staple strands and scaffold are allowed to anneal together using protocols described below.

1.7 Experimental protocols for DNA self-assembly

A typical reaction mixture for DNA origami contains the following:

Scaffold strand at 5 to 50 nM concentration

Each staple strand at 25 to 500 nM concentration (necessary to have staples in 2- to 10-fold excess of scaffold)

Buffer = Tris acetate with EDTA @ pH 8.4

Salt = Mg\(^{2+}\) or Na\(^+\) at 10 to 20 mM concentration (varies by structure)

**The slow and steady annealing protocol:** The reaction mixture (usually 100 – 1000 \(\mu\)L) is incubated in a water bath or a thermocycler machine in the following program:

1. 5 min at 80\(^\circ\) - 90\(^\circ\)C (to break all hydrogen bonds between strands, completely denature them into single strands)

2. Variable ramp from 80\(^\circ\)C to 24\(^\circ\)C (the critical and variable step that determines the kinetic path of annealing and promotes structure assembly)

3. Store at 4\(^\circ\)C (short-term storage condition)

The second step of cooling the sample mix from 80\(^\circ\)C to room temperature is fine-tuned according to the complexity of different structures. In order to drive the annealing process forward towards a favorable path (that yields the structure in high percentage), the cooling step must allow the oligonucleotides to hybridize to their respective scaffold domains while the temperature is maintained above their melting
temperatures, and simultaneously allowing possible mismatched base pairs to denature before the temperature reduces to below the melting temperature. Following this principle, complex structures require slower annealing protocols (from a few days to a week), whereas simpler two-dimensional structures are often ready in a two to three hours. There are several established annealing protocols published that can be adopted for any kind of structure based on the degree of complexity.

**The rapid annealing protocol:** If the melting temperature of staple strands is the key determinant of the temperature cooling program, what role does time (slow versus fast) play? In 2012, Sobczak et. al. tested if the annealing ramps could be made shorter by finding the temperature that drives the reaction to optimum yields[53, 54]. This has led to the discovery of rapid annealing programs that “heat shock” or give optimum energy to a reaction mixture at a very specific temperature for a short period of time - 2 to 5 minutes. The optimum temperature at which rapid annealing takes place can be identified through the average melting temperature of the staples. Quite incredibly, the yields acquired from rapid annealing are comparable to the traditional annealing protocols. However, this method of assembling nanostructures hasn’t been widely adopted yet and scientists continue to use the traditional method.

In case of tile self-assembly or non-scaffolded DNA “origami”, there is no scaffold strand and therefore, all strands are mixed in the reaction at equal concentration.
1.8 Experimental validation and purification of DNA nanostructures

Once the nanostructures are formed, there are several methods of verifying their assembly as well as purifying them for further studies. This section highlights the most commonly found experimental techniques used in DNA nanotechnology.

Firstly, assembly can be checked and quantified using agarose gel electrophoresis, a technique wherein samples are allowed to pass through an agarose gel of predefined porosity and the different migration rates of the objects in the sample (such as, well-formed structures, anomalous side-products, remainder scaffold and excess staples) leads to a size-based separation of the objects in the gel into individual bands. After staining the gel with a suitable dye, the bands are visible and the yield of a nanostructure can be quantified using the intensities of bands corresponding to the structure and the rest of the lane.

Microscopy is the most common and efficient technique to assess the quality of assembly. Atomic force microscopy (AFM) for two-dimensional and Transmission Electron Microscopy (TEM) for two- and three-dimensional structures provide a high-resolution view of nanostructures. In 2012, the first cryo-EM data was collected on an arbitrary DNA origami structure that gave some very intriguing information at the atomic level[55]. In case of dynamic or reconfigurable nanodevices, microscopy is supplemented using fluorescence microscopy to measure FRET outputs from strategically positioned fluorophores on the functional sites of the device. Another alternative, which highlights nanostructure assembly and integrity is DNA-PAINT, wherein short fluorophore-labeled DNA probes are allowed to transiently attach to domains on DNA origami and observed using super resolution microscopy[56].
Functional implementation of DNA nanodevices often requires a purification step, in order to remove excess unbound staples in the solution (which prevents undesirable background fluorescence and reactions). Nanodevices can be size-separated and concentrated using gel extraction, cellulose filter, ethyl alcohol, with the help of a chemical gradient or chromatography.

1.9 Molecular recognition

In DNA origami, each site on a nanostructure is uniquely addressable by its corresponding staple. Picture weaving a fabric, where multiple strands of threads are brought together to create a two-dimensional piece of cloth. To create a multi-colored fabric, one will have to substitute specific strands with colored strands that would result in the desired pattern. Every position on this cloth is addressable by identifying the corresponding thread and modifying it. In the same way, DNA origami nanostructures are addressable with nanometer precision by labeling the corresponding staple with the desired property or, in most cases, molecule (Fig. 1.7a). Addressability with extreme precision and accuracy makes DNA nanostructures very useful as nanoscale platforms or vehicles for positioning other molecules or even single reactions of interest.

After selecting a site on the nanostructure, the corresponding staple can be substituted with a modified staple that contains the chemical modification at a specific base, which is chosen taking into consideration the spatial orientation of the staple on the structure (Fig. 1.7b). With the knowledge that DNA is a right-handed helix, it is possible to identify which bases are buried within and which are exposed to an open

---

1The DNA used in DNA nanotechnology is in its most common isomeric form - the B-DNA, which is right-handed in its helicity. However, other isomers of DNA, namely the A and Z have different physical dimensions.

---
surface on a nanostructure. Moreover, DNA conjugation with proteins, gold nanoparticles and other organic as well as some inorganic molecules is a well-established area of research, which enables scientists to simply go online and purchase such conjugates for a reasonable price (Fig. 1.7c).

A breadth of research has been performed that harnesses addressability of DNA nanostructures, described briefly below.

1.9.1 Single-molecular analyses

In single-molecular analysis, one or more proteins or nucleic acids of interest are strategically immobilized on DNA-based platforms and observed using various validation techniques (Fig. 1.8a). It has therefore become convenient to investigate the behavior of proteins on a single-molecule level, as well as ensemble level. Studies have shed light on single-molecular cleavage reactions[61], optimum separation between two substrates for reaction with thrombin protein[62], optimum tension in a DNA double helix for topoisomerase interaction[63], effect of double helical tension on methylation[64], DNA base-excision repair[65], conformational switching of G-quadruplexes[66], molecular motors[67], singlet oxygen behavior[68], to name a few. Once a suitable nanostructure is engineered, it becomes a “plug-and-play” application wherein the labeled staples can be switched for other molecules of interest and the same nanodevice can be used to answer a multitude of biological questions.

1.9.2 Molecular diagnosis

The same “plug-and-play” scheme on DNA nanostructures could be employed to detect the presence of molecules, and create diagnostic tools. For instance, Streptavidin (SA) – a globular tetrameric protein – can be detected by chemically conjugating biotin molecules on a DNA nanostructure, which is a ligand well-known to strongly
bind to SA[70]. Under the atomic force microscope, bound Streptavidin will be visible as bright spots on the DNA nanostructure (Fig. 1.7c). This technique has been demonstrated to detect proteins[71, 72] and nucleic acids[73, 69]. Dynamic DNA nanodevices have been engineered that mimic macro-scaled “traps” to detect a target molecule, undergo reconfiguration, and elicit a physical or optical observable change. A programmable target-binding site is designed and a reporter system is created that reflects the outcome of the reaction – target bound versus target not bound (Fig. 1.8). One of the seminal papers in this area is the creation of DNA “nanopliers”, wherein two DNA arms are attached at a hinge and contain the target-binding site on the inner edge, like a pair of pliers; target binding would cause the arms to close (Fig. 1.10c). A donor-quencher fluorophore reporter pair is positioned on the arms such that target absence elicits low fluorescence (quenching of fluorophore) and target presence elicits high fluorescence.

1.9.3 Interface with cell

Molecular recognition in vivo is extremely useful to design targeted drug-delivery vehicles and diagnostic tools and DNA, owing to its unequivocal compatibility to biological molecules (or biocompatibility), theoretically giving it an edge over inorganic ingredients such as silicon or plastic, to interface with cells. There is significant interest in designing targeted drug-delivery DNA “nanobots” that locate their destination cells, deliver the cargo and timely degenerate. However, to realize the therapeutic applications, the stability of DNA nanostructures in cellular environments needs to be understood. Preliminary studies have shown that DNA nanostructures are stable in cell lysates for at least 12 hours compared to single and double stranded DNA[74] and in serum for over 62 hours[75] before they degrade. In interfacing with living cells, cellular uptake of DNA nanotubes was first shown in 2008 by Ko et. al.[76] followed
by their stability in tissue cultures[77]. The tissue culture analysis also revealed that the half-life of nanostructures varies with their shape, wherein one shape could be more stable than another. Direct exposure of DNA nanostructures to different nucleases such as DNase 1 has revealed that they can survive in physiological conditions longer than single- or double-stranded DNA and are better suited for interfacing with living species[78, 79].

Dynamic DNA nanotechnology can “compute” Boolean algebra via DNA strand displacement and help very predictably and accurately locate specific cells in a mixed pool of cells by binding to their unique set of cell surface proteins – cells with receptor A AND B, no other[80, 81, 82]. It is also an efficient way to increase specificity of drug-carrying DNA-based vessels by designing “locks”, immobilizing aptamers, specific to more than one cell membrane proteins. Douglas et. al. successfully engineered a barrel-shaped DNA container that carried cargo for cancer cells[83]. The container remained in closed state with the help of aptamers and complementary strands, and opened upon interacting with cell surface markers to deliver the payload. By employing combinatorial aptamers for different cancer cells, the nanobot could be programmed to selectively deliver the cargo and elicit a response from the cells. Interestingly, this nanobot has also been tested in living cockroaches by Bachelet et. al. to interact with each other using logic operations and regulate release of molecules in the organism’s blood[84]. Synthetic DNA as sticky-ended duplexes or nanotubes can be leveraged to influence the fate of cells – to design synthetic tissues[85, 86] or create extracellular environment that favors a specific differentiation path for stem cells[87].
1.9.4 Nanopatterning

Arranging atoms or molecules with nanometer precision is a desirable technique for building ever-shrinking but more complex computers, circuits and machines in general. It is possible to “print” molecules on metallic or plastic substrates using top-down approaches such as lithography, but at the cost of a loss in resolution, and a significant amount of time and resources. Moreover, every pattern requires a unique etching mold in lithography, increasing the cost linearly.

DNA nanostructures can act as vehicles to bridge the gap between bottom-up and top-down assemblies. Molecules of interest can be labeled on DNA nanostructure substrates to create units of two-dimensional arrays and these substrates can be printed using traditional techniques. Examples of nanopatterning include single-protein arrays for improved imaging[88], multi-protein arrays using aptamers[89] and other attachment chemistries[71, 90, 91, 92, 93, 94, 95], silver nanoparticles[96, 97, 98], gold nanoparticles via various DNA labels[98, 99, 100, 101, 102, 103, 104, 105], assembly of carbon nanotubes[106] and viral capsids[107].

The heightened interest in designing smaller circuits using DNA nanotechnology-assisted patterning has been supplemented with studies on the controlled deposition of nanostructures[108, 109, 110], electrical properties of DNA nanostructures[111], lithography[103] and circuit fabrication[98].

1.9.5 Light harvesting/assembling optical components

Fluorophores are molecules that absorb electromagnetic energy at a specific wavelength, (sometimes) undergo electron excitation and release energy at a specific, but different wavelength upon recovering to ground state. In some cases, the energy that is released by a fluorophore is within the visible electromagnetic spectrum, and can
be used as nanoscale “beacons”. Fluorophores are frequently conjugated to organic molecules in cell and molecular biology to understand processes inside a cell. In DNA nanotechnology, fluorophores are a ubiquitous reporting system in single-molecular analyses and sensors. Additionally, it is possible to create highly complex networks of multiple fluorophores to transfer energy from one point on a DNA-based platform to another to understand the physics of fluorescence resonance energy transfer (FRET), where the energy released by a donor fluorophore is able to excite an acceptor fluorophore. FRET is a highly spatio-sensitive phenomenon, in which the distance (in a specific orientation) between the participating fluorophores should be between 0 and 10 nm[112]. FRET efficiency of energy transfer is also inversely proportional to the sixth power of distance. DNA-based structure make ideal platforms for fluorophore networks due to the extreme precision with which it is possible to immobilize the energy transfer molecules. A fluorophore network contains more than one pair of FRET molecules, the emission energy of one fluorophore acts as excitation energy for the next fluorophore, thereby allowing a cascade of resonance energy transfer reactions and propagation of energy to distances beyond 10 nm[113, 114].

Research has been done on optimization of FRET on DNA nanostructures that identifies labeling sites to achieve different levels of energy transfer[115, 116], in super resolution microscopy[56, 117], and plasmonics[118, 119, 120, 121].

1.10 Molecular programming

Programmability is a property of DNA in the truest sense of the word; computer scientists are interested in assigning or embedding information into DNA oligonucleotides, and creating a network of logic-gated reactions to generate output information, in the form of new DNA species. The term - new DNA species - does not
imply the synthesis of new strands, but suggests a change in the population of DNA oligonucleotides already in solution. For example, in Fig. 1.9, the input information is species X and species Y. Upon “processing” the inputs, which implies a chemical reaction between the two species, the output species formed are Z and W.

While it is impossible to achieve fast processing times with DNA substrates like their silicon counterparts that run computers, there are primarily four factors that motivate molecular programming. Firstly, the implementation of computational algorithms with DNA molecules allows us to establish powerful and modular architectures, which can be combined to form larger molecular systems, in an effort to mimic natural biological complexities. Molecular biology inside living cells are large, pre-programmed, networks of diverse molecular species, such as proteins, DNA, RNA and ions, which receive input in the form of chemicals and release other chemicals as outputs to regulate body functions. Chen, et. al. with the help of chemical reaction network formalism, devised some fundamental reaction types using DNA-based logic gates and reactions[122, 123]. Secondly, molecular programmers strive to provide the community with a standardized programming language to manipulate DNA (such as Microsoft’s Visual DSD), and visualizing and design tools, such as caDNAno[24], Cando[124], and others[125]. Thirdly, theoretical ground-work is performed to optimize experimental execution using computational simulations and analyses, to understand what is feasible and what is not. The work described in chapter 4 of this dissertation gives examples of such efforts[126, 127, 128]. And lastly, molecular programming implements computational knowledge in designing smarter nanodevices with embedded multiplexing (using AND, OR, NOT, etc)[83, 129].
1.11 Molecular actuation and reconfigurable nanodevices

The scope of engineering complex nanoscale shapes with DNA goes beyond mimicking a rigid rod as a DNA double helix. As a biopolymer - in the form of a chain of nucleotides - the physical properties of DNA are definable and programmable, from extremely flexible “hinges” to, as we have already seen, “solid” blocks[130]. Tunable mechanical actuation can lead to the engineering of nanodevices that elicit simultaneous and controlled response to multiple target molecules in the same solution; combining FRET with a reconfigurable DNA “ruler” that gives precise, distinguishable and observable outputs against more than one target species. Programming shape-shifting properties in DNA nanostructures is also a step towards developing synthetic biomachines that mimic natural biological machines, such as enzymes, in physical and functional behaviour, and could lead to applications that integrate the two as “hybrid” nanoscale systems.

In its single-stranded form, DNA demonstrates a persistence length (which indicates the maximum length of a polymer at which it shows defined rigidity) of 5 nm[134], and a double-stranded DNA has a persistence length of 50 nm[135]. The different physical properties of single- and double-stranded DNA, in combination with the embedded “switch” between the two states via DNA hybridization, can be harnessed to program mechanical actuation and reconfiguration into complex DNA nanosystems.

Several functional nanodevices have been engineered that contain a flexible hinge that allows “open-close” internal reconfiguration to trap or release target molecules[136, 133, 38, 137, 83, 138, 132, 131] (Fig. 1.10). In chapter 3, we have demonstrated the use of single- to double-stranded DNA transition to alter the relative position of one structural component within a larger DNA-based nanodevice. Torelli, et. al. have shown
the use of DNA hybridization to open a “door”[131]. Compliant joints of tunable strength can be developed using a combination of helix bundles and single-stranded domains between two stiff blocks of DNA architecture[139]. In recent months, the Dietz lab has produced a range of highly sophisticated dynamic architectures that employ non-base pairing bonds (primarily base stacking interactions) to rotate, slide or attach multiple components in a DNA macrosystem[140, 4].
Figure 1.5  Design of DNA origami and some of the seminal work.

a. Basic design strategy of DNA origami involves the folding of a long scaffold strand (black) into a desirable architecture using shorter complementary oligonucleotides (colored). Each staple binds to two or more spatially disparate domains on the scaffold and induces a physical shape change in the scaffold.  

b. Rothemund demonstrated the successful assembly of various two-dimensional shapes, such as a rectangle, star and smiley face. Copyright © 2006, Nature Publishing Group[22].  
c. DNA origami can also be used to create structures with inherent curvature, such as a sphere or bowl. Copyright © 2011, The American Association for the Advancement of Science[23].  
d. The graphical user interface of caDNAno (version 1), which allows users to select helix bundles in a specific lattice, choose the optimum scaffold raster pattern as well as staple crossovers and generate requisite staple sequences for a structure. Copyright © 2009, Oxford University Press[24].
The DNA brick technology is an example of a scaffold-less technique of creating DNA nanostructures. a. A DNA “brick” is a single-stranded 4-domain strand. b. Using domain complementarity, many bricks can be programmed to self assemble into a DNA 3D “canvas”. c. From the DNA canvas, a subset of strands is shortlisted for the desired shape. d. Examples of some 3D architectures built using DNA bricks. Copyright © 2012, The American Association for the Advancement of Science.
Figure 1.7  Addressibility of DNA nanostructures.

Each site on a DNA origami structure is addressable, by selecting the corresponding staple. a. A staple (green) in the arbitrary structure can be selected for three main kinds of modifications: (i). to introduce a DNA hairpin loop, (ii). to attach a molecule such as a functional group or biotin or fluorophore, (iii). to extend a single-stranded DNA from the nanostructure. Copyright © 2006, Nature Publishing Group[22]. b. Using DNA geometry, optimum modification site can be identified in a way that ensures correct spatial orientation of the hairpin or molecule or DNA strand. In the case shown here, the orange triangle indicates the best location to immobilize a molecule. c. Chemically modified DNA nanostructures can act as substrates for other proteins or targets, such as Streptavidin. These immobilized molecules can be observed under the atomic force microscope as bright spots. Copyright © 2010, Nature Publishing Group[61]
Figure 1.8 Application of DNA nanostructures as sensing and molecular analyses platforms.

a. A DNA origami frame designed to understand the effect of structural tension on a DNA double helix on the activity of methyl transferase enzyme. Copyright © 2010, American Chemical Society[64].
b. DNA origami-based single molecular study to detect single nucleotide polymorphisms using a visual array of oligonucleotides corresponding to each kind of base polymorphism. Copyright © 2011, American Chemical Society[69].
c. DNA origami platform for single molecular analysis of the distance dependency of Thrombin protein for attachment to two ligands A and B. Copyright © 2008, Nature Publishing Group[62].

Figure 1.9 Example of a simple DNA computing reaction.
Figure 1.10  DNA nanosystem with programmable reconfiguration.

a. Schematic model of a DNA origami capsule with a controllable door that opens in response to nucleic acid hybridization. The open door exposes molecular cargo, which is attached to the inside surface of the door. Copyright © 2014, John Wiley and Sons[131].

b. Different macro-mechanical actuations can be mimicked in nanomechanical systems such as hinges and crank sliders. Copyright © 2015, Proceedings of the National Academy of Sciences of the United States of America[132].

c. A nanoscale molecular “plier” that reconfigures its arms to trap a target molecule. Detection of the target can be observed using AFM or fluorescence. Copyright © 2011, Nature Publishing Group[133]
CHAPTER 2. COMPLEX DNA NANOSTRUCTURES
FROM OLIGONUCLEOTIDE ENSEMBLES

Modified from a publication in ACS Synthetic Biology\[49\]
Divita Mathur$^1$ and Eric R. Henderson$^{1,2}$

2.1 Abstract

The first DNA nanostructures were created by self-assembly of a small number of oligonucleotides. Introduction of DNA origami provided a new paradigm for designing and creating 2- and 3D DNA nanostuctures by folding a large ssDNA and ’stapling’ it together with a library of oligonucleotides. Despite its power and wide-ranging implementation the DNA origami technique suffers from some limitations. Foremost among these is the limited number of useful single-stranded scaffolds of biological origin. The chapter describes a new approach to creating large DNA nanostructures exclusively form synthetic oligonucleotides. The essence of this approach is to replace the single-stranded scaffold in DNA origami with a library of oligonucleotides termed “scaples” (scaffold staples). Scaples eliminate the need for scaffolds of biological origin and create new opportunities for producing larger and more diverse DNA nanostructures as well as simultaneously assembly of distinct structures in a “single-pot” reaction.

$^1$Iowa State University, Ames, IA 50011, USA. \{divita, telomere\}@iastate.edu
$^2$The manuscript was written through contributions of all authors. D.M. and E.R.H. designed experiments and co-wrote the paper. D.M. executed all the experiments.
2.2 Introduction

Early studies on the design and construction of DNA nanostructures were inspired by biological archetypes such as the Holliday junction[11, 7, 8] and showed that synthetic oligonucleotides could be used to reproduce and greatly elaborate upon these structures. This strategy, assembling synthetic oligonucleotides into larger structures and devices, opened a new and increasingly sophisticated field of DNA nanotechnology capable of producing complex systems such as a DNA-based robot that delivers nanomaterials to predesignated locations[141] and strategies for DNA-based computation[142, 143, 144]. A new paradigm for constructing DNA nanostructures was reported in 2006[22]. This approach, termed DNA origami, folds a large single-stranded scaffold and secures it in place with numerous oligonucleotide “staples” to assemble a broad range of two- and three-dimensional structures[145, 26, 32, 33, 23, 83, 25]. DNA origami enhances the kinetics of assembly and has been merged with the use of oligonucleotides to create integrated DNA nanosystems with large DNA platforms and smaller moving or functional parts[136, 71, 61, 133, 107].

Despite its well-demonstrated power and breadth of application, the technique of DNA origami suffers from some limitations. Key among these is the need for a large single-stranded scaffold. Scaffolds are of biological origin, the most popular being the bacteriophage M13 genome. Double-stranded DNA can also be used but requires either a single-stranded preparation step or complex annealing protocol[44]. A second limitation is that, in the absence of a library of single-stranded, sequence-distinct scaffolds, the possibility of simultaneous “single-pot” assembly of multiple DNA nanostructures and nanosystems is significantly compromised.

Reported here is a strategy for design and construction of large DNA nanostructures exclusively from synthetic oligonucleotides. The first design step borrows
from DNA origami by rastering a virtual single-stranded DNA molecule, the “design-scaffold” of any desired sequence to form the shape. Next, a set of staples is computed using open source software (e.g., caDNAno[24]). Finally, in the key step, breakpoints are inserted at strategic locations in the design-scaffold to generate an ensemble of “scaples” (scaffold staples) (Figure 2.1). Scaples thus generated hybridize to complementary staples to form DNA nanostructures that are comparable to those created by DNA origami. Since the requirement for a single-stranded scaffold has been removed, these structures are no longer sequence or size limited. By eliminating the requirement for a single-stranded scaffold of biological origin, the multiplexed single-pot assembly of DNA nanostructures becomes possible, with significant ramifications for the assembly of complex nanosystems.

2.3 Results

To test the idea of using scaples in place of full-length singlestranded scaffold DNA, a triangular structure was designed on the basis of a geometrically distinctive and well described DNA origami structure[22], shown in Figure 2.2. After confirming that the structure assembled as expected by the conventional DNA origami protocol, sites along the backbone of the M13 sequence were selected as potential locations for fragmentation and scaple generation. Key considerations were to avoid scaple breakpoints that were close to staple breakpoints (allowing no less than three base pairs of separation), to avoid breakpoints in close proximity to crossovers, and to maintain reasonable scaple lengths with respect to limitations of current DNA synthesis methods (see Figure 2.5). Scaples were, therefore, no longer than about 60 nucleotides and no shorter than about 20 nucleotides (nt), the latter to ensure robust duplex formation and stability.
The staple:scaple ensemble was annealed by a protocol similar to that used for DNA origami (Methods). Figure 2.2 shows the results of this experiment. A significant number of complete structures was observed in the scaple:staple reaction, with total yield based on electrophoretic mobility of approximately 11% (Figure S2, Supporting Information). These structures were indistinguishable from the same structure created by DNA origami (Figure 2.2, inset). The number of incomplete structures in the staple:scaple reaction was substantially greater than that observed for the DNA origami reaction. Nonetheless, these data show that large DNA nanostructures can be assembled in the absence of a continuous scaffold backbone and the upper limit in size and complexity need not be constrained by scaffold limitations.

A rigorous test of the scaples approach would be the creation of a structure that was independent of the M13 sequence entirely. To carry out this test, a number of random sequence strings containing 7249 nt (the same number as in the M13mp18 scaffold) was generated. The sequences were tested for internal complementarity and those devoid of excessive predicted secondary structures were further analyzed (Methods). A sequence was chosen from this subset and used to design the same triangular structure shown in Figure 2.2 (inset). Staples and scaples were computed as in the first experiment. The staples and scaples thus generated were checked once again for undesirable characteristics such as internal base pairing and G-quartet formation and minor adjustments made to minimize the occurrence of these phenomena. Finally, these molecules were synthesized and assembled. The results are shown in Figure 2.3. Triangular structures were observed that were indistinguishable from those formed with the M13- derived scaples and the DNA origami triangle. Despite the relatively low yield, comparable to that observed with M13 scaples (Figure 2.2), this experiment shows that it is possible to construct complex DNA nanostructures using sequences that are nonbiological and randomly generated. This capability re-
moves the requirement for large single-stranded scaffolds of biological origin with the concomitant sequence and length limitations.

An important opportunity created by the method described here is the potential to create different DNA nanostructures in a single-pot reaction, something that is very difficult with only one type of scaffold (i.e., M13). To test this possibility, a conventional, geometrically distinct DNA origami structure (a rectangle) was assembled in the same reaction mixture as the random sequence scaples-based triangle described above and a scaples-based round cornered square. Figure 2.4 shows that all three structures formed with yields roughly comparable to those obtained when assembled independently. Thus, with 837 individual oligonucleotides and one large single-stranded scaffold, it is possible to create three defined high molecular weight DNA nanostructures in a single reaction.

2.4 Discussion

This report describes a strategy for the design and construction of large DNA nanostructures that assemble from a library consisting exclusively of small synthetic oligonucleotides. Any biological sequence constraints are lifted, thereby creating new opportunities for expanded and single-pot DNA nanostructure self-assembly. These new opportunities come at a cost. The higher order kinetics of nanostructure assembly by this method enhances the probability of formation of undesirable byproduct and concomitant lower yield of the desired product when compared with DNA origami. Similar limitations in yield have been observed in a related, independently developed strategy for creating DNA nanostructures[51]. In that approach, termed the “single-stranded tile” method, a stencil pattern is used to mask a preconfigured selfassembled DNA “canvas” comprised of a few hundred oligonucleotides. The mask defines the
required subset of oligonucleotides to create a large number of different shapes from
the same canvas. Although this method requires a large set of “edge protectors” to
avoid undesirable aggregation, it allows rapid and automated construction of a vast
array of nanoscale DNA shapes. In contrast, the scaples method utilizes a unique set
of oligonucleotides for each shape, which incurs greater cost and design effort but in
return allows simultaneous assembly of linear and curved, 2D and 3D nanostructures
in a single reaction. Moreover, the scaples method uses open source design software
(e.g., caDNAno[24]) that is readily available to any laboratory. As this approach is
further developed, it should be possible to improve upon scaple and staple designs
and optimize reaction conditions and methods to obtain higher yields. It is likely that
both the scaples and the single-stranded tile strategies, and possibly a combination
of the two, will find applications in DNA nanotechnology, the preferred approach
in any particular case being a function of their virtues and liabilities in light of the
experimental goal(s). The initial success at building complex nanostructures from
libraries consisting exclusively of short synthetic oligonucleotides suggests that fur-
ther development of this general strategy (executed via scaples or the single-stranded
tile method) will lead to significant advances in the use of oligomeric DNA for the
construction of large and complex DNA nanostructures and devices.

2.5 Methods

2.5.1 Materials

All staple and scaple strands were purchased from Integrated DNA Technologies
(IDT, Coralville, IA) in 96-well plates at 100 uM (25 nmoles) in RNase-free water.
M13mp18 ssDNA was ordered from Bayou BioLabs (Metairie, LA) and was used
without any further purification.
2.5.2 Assembly of the nanostructures

The staples and scaples corresponding to each structure are listed below (Tables S1 and S2, Supporting Information). The buffer used to create these structures was the same as the one used in DNA origami, namely, 40 mM Tris, 20 mM acetic acid, 2 mM EDTA (TAE, pH 8.3) and 12.5 mM magnesium acetate (TAEM). Since staples and scaples are of the same size, they were added in 1:1 molar ratio at a final concentration of 40 nM for each oligonucleotide. The desired structures were created by thermal annealing as follows: heat to 95 °C for 3 min, cool from 95 to 40 °C in 4 h, followed by maintenance of temperature at 40 °C for 10 h. After that, the solution was cooled to 24 °C over 1 h (note, no difference was observed if cooled to room temperature at this stage). The solution was stored at 4 °C before imaging by AFM.

2.5.3 AFM imaging

The structures were imaged by depositing 2 μL of the sample on a freshly cleaved piece of mica and allowing the sample to bind to the mica surface for 1 min. The mica was then rinsed with distilled water (dipping, 10×) and dried using nitrogen gas. Images were collected in Tapping-Mode using a Digital Instruments MultiMode AFM.

2.5.4 Agarose gel electrophoresis

Samples were analyzed by electrophoresis through a 1% agarose gel in TAEM at 50 mA. After electrophoresis, the DNA nanostructures were visualized by staining with 1× SYBR Green and illumination with UV light (365 nm).
2.5.5 Design of scaples

An open source program, caDNAno[24], was used to create the design-scaffold and corresponding staples in a manner analogous to that used for DNA origami. A .json file for each side of the triangle was created using the Rothemund triangle as reference and exported as an .svg file. The .svg files were opened in Adobe Illustrator (.ai) and manipulated to construct the complete triangle (Figure 2.5). Then, starting from one of the innermost helices in the triangle, potential scaple breakpoints were identified on the basis of the following guidelines:

1. Length of a scaple: Breakpoints were positioned to maintain scaple lengths between 20 and 60 nt. There were instances where the length of the scaples had to be reduced to accommodate other necessary requirements.

2. Relative position of breakpoints: Scaple breakpoints were introduced such that they were as far as possible from crossovers. With this in mind, scaples were designed to allow hybridization with the maximum number of different staples possible. In cases where a positioning compromise was required, such as in the presence of a staple junction, the breakpoints were offset by 3–5 nt from the midpoint between two crossovers.

3. Stereochemical consideration: Care was taken to not introduce unwanted degrees of rotation around crossovers. This was accomplished by avoiding the alignment of breakpoints on adjacent parallel helices. This resulted in a staggered arrangement of breakpoints.

4. Exceptions: In some instances, exceptions to these guidelines were necessary, and in those cases, the overall goal of maintaining structural integrity was the primary driver.
Figure 2.5A is an example of a case where the breakpoint was offset by 3 nt from the midpoint because of the staple junction in the center. Figure 2.5B shows the staggered arrangement of breakpoints on different and parallel helices of the design scaffold. Aligning the breakpoints vertically may make the structure less stable.

We generated 154 scaples for the triangle structure of length between 30 and 60 nt. The 97-nucleotide long loop on one of the sides in the original DNA origami triangle was omitted from the scaples-based nanostructure. This was possible because, unlike M13mp18, scaples are not limited by length. If a loop or any other structure is desirable, it can easily be introduced into the scaple-based structure. It is noteworthy that this set of staples and scaples is just one of many possible sets that could be used to create this nanoshape.

2.5.6 Generation of random sequence design-scaffolds

To test the creation of nanostructures based on a nonbiological sequence, an algorithm was developed that generated completely random sequences the same length of M13mp18 (7249 nt). One of these sequences was selected, and this sequence was analyzed for internal repeats of length of 15 nt or more. A 97 nt long subsequence (that forms a loop in the origami triangle) was removed, and the resulting sequence was then divided into three segments (2384 nt each) and loaded into caDNAno for the construction of each side of the triangle. Each side was exported as an .svg file and assembled into one triangle as described in Figure 2.5 (Figure S1, Supporting Information). Scaple sequences were constructed using the consolidated triangle as described above and tested for the presence of undesirable secondary structure features including hairpins and G-quartets.
2.6 Figures

Figure 2.1  Strategy for the design of scaples-based nanostructures.

(A) The first step is to raster a “design-scaffold” through the desired shape. Staples are then introduced using software (e.g., caDNAno (22)) or by hand. (B) In the key step, positions for the insertion of breakpoints on the design-scaffold are determined. The scaples and staples thus generated are synthesized and annealed as described here and in the Supporting Information.
Figure 2.2  Scaples-based triangle using m13mp18 as the design-scaffold.

(A) A representative field of the scaples version of the original DNA Origami triangle (shown for comparison in the inset). The 154 scaples created for this triangle were designed using the M13mp18 design-scaffold layout exactly as in the origami triangle (8). (B to G) Higher magnification AFM images of individual scaples-based triangles.
Figure 2.3  Scaples-based triangle constructed using a random sequence design-scaffold.

The nanostructures shown are geometrically identical to those shown in Figure 2 but were created using a non-biological, random sequence design-scaffold. The sequence was processed to remove internal subsequence similarity, undesired internal complementarity and sequences formally capable of forming G-quartets. (A) A representative field AFM image of the triangles. (B to G) Individual examples of the same structure.
Figure 2.4  Simultaneous assembly of three DNA nanostructures in a "one-pot" reaction.

In the experiment shown here one DNA origami structure with an M13mp18 scaffold (rectangle, identified by white arrows) and two scapes-based nanostructures, a triangle (green arrows) and a square with curved corners (red arrows), were assembled in a single reaction containing over 800 distinct oligonucleotides. Both the triangle and the round-cornered square were designed with non-biological random sequences.
The blue strand represents the design-scaffold and the other colored strands represent the staple sequences. Each side was designed in caDNAno (22) and was exported in .svg format and converted to .ai format in Adobe Illustrator (note: it is possible to make the entire triangle in caDNAno as a single .json file instead of breaking it into three sides, but the resultant design will need the same processing to represent the triangle in a workable format). The sides were consolidated into a single .ai file. The sides were joined by combining the bridge staples on the nonparallel edges of the sides on the corresponding helices together and the addition of 0-4 thymines to introduce flexibility at the vertices. The innermost helices of the three sides were joined by the design-scaffold (later converted to scaples). At this point breakpoints were introduced in the design-scaffold (blue spots) following the guidelines described earlier. There are 154 scaples in this structure. The loop in the bottom (surrounded by a red box) was omitted from the scaples-based nanostructure and the scaple sequence on either side of the loop was combined to form one scaple. Two sections (dotted boxes) are enlarged to illustrate specific features. Panel A shows an example of offsetting the breakpoint by 3 nt from the midpoint between two crossovers due to the presence of a staple junction in the center. There are many occurrences of this architecture in the structure. Panel B shows a representative staggered arrangement of breakpoints on adjacent parallel parallel helices of the design-scaffold.
CHAPTER 3. PROGRAMMABLE DNA NANOSYSTEM FOR MOLECULAR INTERROGATION

Modified from a manuscript, under review in Nature Scientific Reports.
Divita Mathur\textsuperscript{1} and Eric R. Henderson\textsuperscript{1, 2}

3.1 Abstract

We describe a self-assembling DNA-based nanosystem for interrogating molecular interactions. The nanosystem contains a rigid supporting dumbbell-shaped frame, a cylindrical central core, and a mobile ring that is coaxial with the core. Motion of the ring is influenced by several control elements whose force-generating capability is based on the transition of single-stranded DNA to double-stranded DNA. These forces can be directed to act in opposition to adhesive forces between the ring and the frame thereby providing a mechanism for molecular detection and interrogation at the ring-frame interface. As proof of principle we use this system to evaluate base stacking adhesion and demonstrate detection of a soluble nucleic acid viral genome mimic.

\textsuperscript{1}Iowa State University, Ames, IA 50011. \{divita, telomere\}@iastate.edu.
\textsuperscript{2}Author contributions: D.M. and E.R.H designed experiments, analyzed data and co-wrote the paper. D.M. executed the experiments.
3.2 Introduction

In addition to the iconic genetic code, nucleic acid contains an “engineering” code. Recent conceptual and methodological advances have culminated in the availability of tools and strategies for leveraging this engineering code to program DNA to spontaneously create a diverse array of two- and three-dimensional shapes\([146, 22, 49, 51, 23]\). These shapes can be imbued with information and function including algorithmic calculations\([147, 148, 149]\), single-molecule analyses\([64, 150]\), therapeutics\([83, 77, 74]\), mechanical actuation\([132, 138]\), and a host of other capabilities\([84, 110, 151, 152]\). Programmable, responsive actuation of dynamic self-assembling nanodevices is a highly desirable attribute and several studies have demonstrated mechanical reconfiguration of DNA nanodevices by thermal motion\([132]\) and upon triggering by a target molecule\([83, 136, 137, 130, 153, 133, 5, 38]\). The study described here focuses on improvements in these responsive nanosystems in terms of modularity and robustness while minimizing undesirable conformational changes upon actuation. The DNA nanosystem described here enhances the repertoire of molecular reporting systems\([5, 154]\) and serves as a base platform for molecular interrogation with an embedded reporter system module that is compatible with a variety of molecular species. In this study we harness the difference in persistence length (i.e., rigidity) of single-stranded (ssDNA) and double-stranded (dsDNA) DNA to elicit a defined physical state change in a self-assembling DNA nanosystem we have termed OPTIMuS (Oligo-Propelled Technology for Interrogating Molecular System; Fig. 3.1). This inducible state change can be used to interrogate user-programmed molecular interactions within the OPTIMuS platform. In this report, we demonstrate how OPTIMuS can be used to detect a soluble target molecule and assess the relative strength of a non-covalent (base stacking) molecular interaction.
3.3 Results

3.3.1 Construction and principle of operation

The OPTIMuS platform is constructed following the principles of DNA origami, in which, specific ensembles of short oligonucleotides called “staples” are used to fold a large single-stranded “scaffold” into desired shapes (see Appendix B & Fig. B.2). The mechanical design of OPTIMuS is inspired by a system in which tunable “springs” exert pushing and pulling forces on a movable ring within a coaxial dumbbell-shaped framework (Fig. 3.1a). These forces are opposed by introducing resistance at the interface between the mobile ring and the dumbbell frame. Finally, an embedded Förster Resonance Energy Transfer (FRET) system, in which one Cyanine 3 (cy3) molecule is positioned on the frame and one Cyanine 5 (cy5) molecule is on the ring, reports the relative position of the ring under various conditions (Fig. B.9). The main components of OPTIMuS are described in Fig. 3.1b.

The force elements are single-stranded scaffold domains that undergo structural change upon hybridizing to their complementary staple strands. Single-stranded DNA, an entropically elastic polymer with a formal contour length of 0.7 nm/base and persistence length of about 5 nm[134], transitions into a rigid double-helix of 0.34 nm/bp contour length and 50 nm persistence length upon hybridizing with its complimentary strand[135] (Fig. 3.1e). If the ends of a ssDNA molecule are tethered to two substrates, the relative distance between the substrates can be altered due to the internal reconfiguration concomitant with duplex formation (Fig. 3.1e Scenario 1 and 2). This spatial change can be exploited for applications in sensing[155, 156] and, potentially, molecular force/energy measurements. Importantly, although somewhat counterintuitive, if these substrates are immobilized, causing the ssDNA to be fixed
at its full extension limit (≥0.7 nm/base), duplex formation becomes stereochemically inhibited despite the favorable ΔG of the reaction (Fig. 3.1c Scenario 3).

Three kinds of scaffold domains in OPTIMuS use the aforementioned phenomena to impart pushing or pulling force on the central ring to move it from a position proximal to the right side of the frame (frame^{R}) toward the left side of the frame (frame^{L}) (Fig. 3.1d). These domains are termed “extended core” (EC), “cinchers” (Ch) and “loops” (L). The corresponding staple strands are termed EC^{S}, Ch^{S}, L^{S}, respectively. The L domain function is illustrated in figure 1e, scenario 1. Upon hybridization to L^{S} staples, L domain extension causes the ring to move away from frame^{R}. The Ch domain contains stretches of scaffold DNA that are shorter in length (35 bases) than the underlying EC domain (70 bases) (Fig. B.3). As shown in figure 3.1e, scenario 2 these domains pull the ring towards frame^{L}. Finally, extension of the EC domain by hybridization to EC^{S} staples generates a pushing force on frame^{L} moving it away from frame^{R}. This motion results in extension of cincher DNA which pulls on the ring, moving it away from frame^{R} (Fig. B.3). These design features are illustrated in caDNAno layouts corresponding to each configuration in Figs. B.4 - B.7.

Motion of the ring induced by hybridization of force domains can be challenged by introducing resisting adhesive force(s) at the interface (“active sites”) of the ring and frame^{R} (Fig. 3.1c). Two types of resistance were tested in this study. The first type is blunt end base stacking (π-bond interaction) at the active sites (Fig. 3.1f). Previous work has shown that base stacking interactions can result in strong helix-helix adhesion and this interaction has been utilized to create multiunit self-assembling DNA nanostructures[4, 3]. Previous studies suggest that GC dinucleotides form the strongest stacking interaction[157]. Therefore, the nanosystem designs used
in this study employed GC dinucleotide base stacking to promote adhesion at the active sites (Fig. B.10).

Another kind of resistance is a “lock and key” system, the design of which is based on toehold-mediated strand displacement\[5\] (Fig. 3.1g). Toehold-mediated strand displacement is a dynamic hybridization event wherein a DNA strand invades and displaces another strand from a duplex by binding to a short single-stranded oligonucleotide extension called the toehold. In these experiments the active sites were decorated with toehold-bearing duplexes that tether the ring to frame$^R$.

### 3.3.2 System characterization and demonstration

To assess the scope of influence of each force inducing domain (EC, Ch, L) on the central ring and its effect on the ensemble FRET output, we assembled the nanosystem in the presence of different combinations of domain-specific staples (Fig. 3.2). OPTIMuS at “ground state” (G) is defined as the configuration that contains no force domains (EC, Ch, and L remain single-stranded). In each sample, staples associated with G (i.e., frame$^R$, frame$^L$, core and ring) were mixed with active site staples conferring blunt end formation and also a specific combination of force domain staples. After mixing the nanosystem was assembled following the standard protocol (Methods). The FRET output reflects the position of the ring relative to frame$^R$ in each configuration (corroborated by Transmission Electron Microscopy (TEM) analysis; Fig. 3.2). Since blunt end stacking is strongly distance-dependent\[2\], and therefore cannot acquire sufficient force when assembled simultaneously with opposing force domains, the force domains prevail in all cases except Ch$^S$ (see following) and induce various degrees of ring movement. Figure 3.2b shows that Ch$^S$ alone has no effect on FRET compared to G. In contrast, L$^S$ alone and EC$^S$+$L^S$ elicit the same FRET as all three force domains in combination (EC$^S$+Ch$^S$+L$^S$). Finally, G+EC$^S$ and
G+ECS+ChS have comparable FRET outputs. Thus, these two force domain combinations, while differing in total ΔG appear to reach a common final mechanically limited state of the nanosystem (minimum FRET). Based on these results we chose three combinations of force domain staples as actuators, EC\textsuperscript{S}, EC\textsuperscript{S}+Ch\textsuperscript{S} and EC\textsuperscript{S}+L\textsuperscript{S}, to study their effect on the motion of the ring when opposed by two different kinds of resistant forces, base stacking and DNA-DNA hybridization.

### 3.3.3 Base stacking versus force domains

Figure 3.3 shows the results of experiments in which combinations of force inducing domains were tested for their ability to disrupt base stacking-mediated ring/frame\textsuperscript{R} adhesion. Ground state with blunt ends on all active sites (called G\textsuperscript{All BE}) and ground state with no blunt ends on all active sites (called G\textsuperscript{No BE}) were initially assembled in the absence of force domains. The samples were divided into equal amounts, filtered to remove excess staples, and then incubated with either buffer alone, EC\textsuperscript{S}, EC\textsuperscript{S}+Ch\textsuperscript{S} or EC\textsuperscript{S}+L\textsuperscript{S} (Methods).

In the case of G\textsuperscript{All BE} (all blunt ends stacked) the combination of duplex formation in EC and L domains resulted in rupture of the adhesive bond between the ring and frame\textsuperscript{R}. In contrast, EC\textsuperscript{S}+Ch\textsuperscript{S} failed to disrupt the base stacking interaction. We hypothesize that this is the consequence of fully-extended and strained cincher ssDNA being unable to form a sufficient number of hydrogen bonds to initiate an ssDNA to dsDNA transition and thereby create a resultant pulling force (in contrast to ease of duplex formation when cinchers are hybridized during the initial self-assembly process, Fig. 3.1e Scenario 3; Fig. 3.2b). Unlike strained cincher domain ssDNA, ssDNA in the loop domain is not stretched but, rather, randomly coiled and, therefore, stereochemically available for hybridization with complementary oligonucleotides. Thus,
hybridization to loop domains generates a pushing force on the ring that is sufficient to rupture of the ring/frame\textsuperscript{R} adhesive interaction.

When OPTIMuS was configured to lack base stacking interactions at the ring/frame\textsuperscript{R} interface (Fig. 3.3b $G^{\text{No BE}}$) all combinations of force domains tested were able to induce ring displacement. In particular, duplex formation of EC+Ch was able to disrupt the ring/frame\textsuperscript{R} interface because there was no opposing adhesive force and, therefore, no hyperextension of the ssDNA-cincher domains to preclude cincher duplex formation (Fig. 3.1e Scenario 2). The results of these experiments suggest that OPTIMuS may be useful for interrogating other types of molecular interactions at the ring/frame\textsuperscript{R} interface.

TEM analysis was carried out to corroborate FRET analysis of the various configurations of OPTIMuS. Figure 3.4 shows that with full bunt end stacking ($G^{\text{All BE}}$) internal reconfiguration does not take place in the presence of $EC^S$ or $EC^S+Ch^S$ (Fig. 3.4a (i)-(iii)). However, the addition of subsequent force domains ($EC^S+L^S$) induces a change that leads to ring movement (Fig. 3.4a(iv)). This can be verified by observing the “gap” inside OPTIMuS which shows the ssDNA cincher domain and helps locate the relative position of the ring (TEM images, Fig. 3.4a(iv)). The overall length of OPTIMuS is also a good indicator of internal reconfiguration, as can be seen by the dimensional analysis of a population of each kind of sample (right columns, Fig. 3.4a, b). The mean length of GAll BE remains the same upon addition of EC$^S$, but shifts to an intermediate length in case of $EC^S+Ch^S$. The partial hybridization affects the overall length of the nanosystem, but is unable to cause motion in the ring, hence the FRET signal does not alter. This observation supports our hypothesis that there is partial hybridization in the two force domains, EC and Ch, but complete hybridization is stereochemically hindered (i.e., Ch is physically constrained in a stretched configuration) by the blunt end stacking at the active sites.
In contrast to the results above, the configuration lacking blunt end stacking, $G^{\text{No BE}}$, undergoes an incremental shift in the ring position as well as length of the nanosystem as a function of force domain hybridization (Fig. 3.4b). The addition of $\text{EC}^S$ alone causes OPTIMuS to assume a bent configuration (Fig. 3.4b(ii)). This can be attributed to a fully-extended ssDNA Ch domain, the tension in which is sufficient to distort the otherwise linear core architecture. This bending serves as evidence that hybridization of the $\text{EC}^S$ to EC domain is occurring with high efficiency. Duplex formation in the remaining two force domain samples, $\text{EC}^S + \text{Ch}^S$, $\text{EC}^S + \text{L}^S$, results in full shift in the ring’s position and a corresponding overall increase in the length of the nanosystem. The configuration-specific gaps corresponding to single-stranded Ch and L domains permit unambiguous orientation determination of the molecule and indicate the internal position of the ring (Fig. 3.4b(iii),(iv)).

3.3.4 Detection of a soluble ssDNA viral genome mimic

The earlier suggestion that OPTIMuS might serve as a useful molecular detection platform was tested using a strand displacement mechanism and a viral genome (DNA) mimic. Toehold-bearing duplexes, called ligand domains, were integrated with the ring/frame$^R$ interface such that one strand of each duplex extends from the ring and the other from frame$^R$. The toehold-containing strand was designed to be complementary to a soluble target oligonucleotide based on five Ebola genome sequence elements (Supplementary Information). The active sites were modified with these ligand domains (Fig. 3.5a, b). Addition of the target strands disrupted the ligand duplex through toehold-mediated DNA strand displacement[5], thereby reducing ring/frame$^R$ adhesion and permitting free motion in the ring. We compared ring motion in the presence and absence of target molecules in different OPTIMuS configurations. Following the format of experiments that tested blunt end stacking
interactions in Fig. 3.3 and 3.4, we constructed the ground state in the absence of the force domains (G$^{\text{No BE}}$) but bearing the ligand duplexes at the active sites (Fig. 3.5a). The ground state sample was divided, purified via filtration to remove excess staples, and incubated (Methods) with the force domains with or without target strands and FRET was measured. Upon force induction in the absence of target the FRET signal only changed in the presence of the strong disruptive force domain combination EC$^S$+L$^S$. However, in the presence of soluble target the FRET signal was significantly reduced when the ring was induced to move using the weaker force-generating domains EC or EC+Ch, thereby illustrating that the OPTIMuS platform has the potential to serve as a molecular detection platform (Fig. 3.5c).

3.4 Discussion

We describe a self-assembling DNA nanosystem termed OPTIMuS that is capable of interrogating molecular interactions by exerting user-controllable forces to challenge the molecular system of interest. Controlled exertion of force in OPTIMuS is founded on the inherent elasticity of ssDNA (a relatively weak and compliant spring), the shortening and stiffening of double helical DNA (a relatively strong and stiff spring), the high specificity of DNA base pairing, and the adhesive force exhibited by base stacking. The availability of a plurality of control elements should allow OPTIMuS to be tuned to interrogate interactions of a range of strengths. In the present study we explore a soluble target strand displacement mechanism of detection and the interactive force present in base-stacked DNA duplexes.

A self-assembling DNA-based system that is capable of interrogating and, potentially, measuring inter- and intramolecular forces/energies is compelling for several reasons. It is extremely economical in comparison to macroscopic instrumentation
that is used for molecular force measurements (e.g., atomic force microscopy (AFM) and optical trapping). Moreover, in contrast to those systems OPTIMuS has the potential to perform thermodynamically reversible force induction, which would overcome the limitations of time-varying external forces obtained by AFM and optical tweezers\cite{158}. The strength of each force domain may be tuned at the single base pair level to create a highly nuanced spectrum of test energies. Finally, this system may lend itself to statistically robust soluble molecular population-based as well as chip-based single molecule or smaller population analyses.

As a sensor, OPTIMuS is readily reconfigurable and capable of multiplexing, a potential advantage over molecular beacons\cite{159}. DNA is amenable to a wide range of chemical modifications making it relatively simple to incorporate a variety of molecular species into the system for study. Versions of self-assembling systems like OPTIMuS can be multiplexed to create (AND/OR) logic gates and iterative biosensors for high confidence molecular detection. Moreover, the ability to precisely arrange gold nanoparticles on OPTIMuS\cite{160, 100} suggests a pathway to enhanced sensitivity by methods such as surface-enhanced Raman spectroscopy (SERS) for use in field deployable diagnostics\cite{118}. Finally, DNA nanosystems are inherently biocompatible and may be further embellished to create novel bionanodevices that have the potential to interact with natural biological systems \textit{in vivo}.

3.5 Methods

3.5.1 Nucleic acids

All oligonucleotide staple strands were purchased from Integrated DNA Technologies (IDT, Coralville, IA), supplied in RNase-free water at 100 \( \mu \)M concentration in individual wells. M13mp18 single-stranded scaffold strand DNA was purchased from
Bayou Biolabs (Matairie, LA) and was supplied at a concentration of 1 μg/μL in Tris-Acetate EDTA buffer. Experiments were carried out without additional purification.

### 3.5.2 Chemical and supplies

All other chemicals (Tris-Acetate EDTA, Magnesium Acetate Tetrahydrate, and water) and supplies were purchased from Fisher Scientific.

### 3.5.3 Assembly of OPTIMuS

The annealing protocol was adopted from Stein ET. al[115]. The requisite staple strands (including the fluorescently-labeled staples), each at a final concentration of 50 nM, were mixed with m13mp18 scaffold strand at a final concentration of 10 nM in 1x reaction buffer (comprised of 40 mM Tris-Acetate, 1 mM EDTA (pH 8.3) and 18 mM Mg²⁺) and brought to a final volume of 500 μL. The desired structures were assembled using the following thermal annealing program:

- 80°C - 5 min
- 80°C to 60°C - 80 min
- 60°C to 25°C - 1200 min
- 25°C to 4°C - 10 min
- 4°C - storage until further experiments

Care was taken to maintain all staples in the dark by covering the PCR plate as well as the laboratory tube rack with aluminum foil.

### 3.5.4 Centrifugal filtration

Removal of excess staples, particularly those with fluorescent labels, was critical for optimal results and quantification. Excess staples were removed using Amicon Ultra-0.5 mL Centrifugal Filters (50,000 molecular weight cutoff (MWCO)). 500 μL
of the reaction mix was poured into a filter column and centrifuged at 14,000g for 5 min. The eluate collected in the collection tube was discarded and the filter column was placed back into the collection tube. Centrifugation step was repeated 4 times on the same filter column by adding 450 μL of 1x reaction buffer to the filter column before each step. After completing the centrifugation, the retentate was recovered by inverting the column in a fresh tube and performing centrifugation (at 1000g) for 3 min.

3.5.5 Agarose gel electrophoresis

The efficiency of assembly was evaluated by electrophoresis using a 1.5% agarose gel. Electrophoresis was carried out on ice at 72 Volts for 4 hours. Gels were stained with 1x SYBR Green and illuminated under UV (302 nm) using a Benchtop 2UVTM Transilluminator (UV Products).

3.5.6 Post-assembly sample treatment with different combinations of force strands and target strands

In experiments in which a preconfigured OPTIMuS sample was treated with force domain strands or target strands (Fig. 3b and Fig. 5c), filtered samples were mixed with 100 nM of each desired staple, such as extended core, cinchers and loops. The buffer conditions of the force domain staples were consistent with the ground state sample and the cation concentration of the resultant samples was maintained at 18 mM Mg$^{2+}$. Samples were incubated using the following thermal annealing protocol:

- 40°C - 60 min
- 40°C to 25°C - 600 min
- 25°C to 4°C - 10 min
- 4°C - storage until further experiments
3.5.7 Ensemble fluorescence resonance energy transfer (FRET)

In order to minimize background fluorescence, only filtered samples were used for FRET experiments. A custom-designed well was created in lab using a microscope slide and a coverslip to carry out fluorescence microscopy through a hyperspectral microscope (Nikon Eclipse TE2000-E) and EXFO X-Cite 120 PC Fluorescence illumination system. Exposure time was set at 50 ms and a 20X objective was used to image using an HQ Wide Green Filter (Excitation Filter: 545/30 nm, Dichromatic Mirror: 570 nm, Barrier Filter: 610/75 nm). Data acquired with the hyperspectral microscope for each sample was a two-dimensional array of Wavelength (nm) and Intensity (arbitrary unit). Intensities were normalized by the total intensity \(i\) received per sample before calculating FRET. FRET was calculated based on the following formula:

\[
FRET(\%) = \frac{\left(i_{Cy^3}\times 100\right)}{\left(i_{Cy^3} + i_{Cy^5}\right)}
\]

where,

\(i_{Cy^3}\) = fluorescence intensity at Cy3 emission (574 nm)

\(i_{Cy^5}\) = fluorescence intensity at Cy5 emission (669 nm)

3.5.8 Transmission electron microscopy (TEM)

TEM imaging was carried out using the protocol described by Castro et. al[79] with a JEOL 2100 STEM Scanning and Transmission Electron Microscope. 2 μL of sample mix (~10 nM concentration) was deposited on a carbon film grid and allowed to adsorb for 30 seconds. Excess liquid on the grid was wicked from the edge of the grid. Following that, 2 μL of the staining solution42 was added to the grid and allowed to adsorb for 30 seconds. Excess stain was wicked from the edge and the grid was allowed to dry in air prior to imaging.
3.5.9 Image processing

TEM imaging generated .dm3 files, which were fed into the boxer.py program of EMAN2 to pick individual particles (or nanostructures) and create a stack. Then we performed dimensional analysis (length measurement) using the line tool in ImageJ or Fiji. Scale of the images was set according to the information stored in the dm3 files. Histograms were generated with the help of R programming language.
3.6 Figures

Figure 3.1 OPTIMuS operational principles.

(a) A schematic illustrating the key mechanical components of OPTIMuS. A movable ring surrounds a cylindrical core that is anchored at both ends by "dumbbell" shaped frame elements. The ring is pushed and/or pulled away from frame\textsuperscript{R} by user-controlled ssDNA to dsDNA transitions, whereas "resistance" at the interface can obstruct ring movement. (b) A 3D rendering of OPTIMuS showing the 24-helix bundle in the honeycomb lattice arrangement. (c) A cross-sectional view of the ring/frame\textsuperscript{R} interface shows active sites, loops and the FRET reporter pair. (d) The idealized overall reconfiguration that can be elicited in OPTIMuS. On the left is the ground state (G) that has all force domains in single-stranded form (EC, Ch, L). Upon adding staples corresponding to them the nanosystem reconfigures with a displaced ring position. (e) Depiction of the mechanism of force-induced motion by an ssDNA to dsDNA transition. (f) A schematic of the blunt end and non-blunt end interactions between the coaxial helices of ring and frame\textsuperscript{R}. Non-blunt ends are created by leaving eight scaffold bases at the crossover unhybridized. The resultant single-stranded region prevents base stacking and minimizes adhesive interaction between the duplexes. (g) Illustration of a toehold-mediated DNA strand displacement reaction.
Figure 3.2  System characterization and demonstration using FRET and Transmission Electron Microscopy (TEM).

(a) A TEM field showing examples of the ground state of OPTIMuS. White arrows adjacent to some structures indicate single-stranded “clouds” DNA corresponding to the unhybridized force domains. (b) FRET output of OPTIMuS as a function of combination of pre-added (during self-assembly) force domain components. FRET is a reliable indicator of ring/frame distance and, therefore, a reporter of force-induced ring motion. A schematic of each configuration is shown for clarity. In case of G+EC, pre-adding EC affects the stable formation of the structure due to internal tension between Ch and EC, thereby destabilizing the formation of the frame (Supplementary Fig. 1). This is one type of internal tension that comprises the foundation of mechanical actuation in the nanosystem. (c) Corroborating TEM structures corresponding to key OPTIMuS configurations used in this study (Scale bar = 50 nm). Green and yellow arrows indicate structural “gaps” corresponding to ssDNA L and ssDNA Ch domains respectively.
Figure 3.3  Base stacking adhesive forces versus OPTIMuS actuators.

(a) The interface makeup between the ring and frame shown as a cross-section of OPTIMuS. It can be modified with coaxial blunt ends (BE) to create high adhesion or no blunt ends to minimize adhesion. (b) FRET output of all blunt ends (All BE) versus no blunt ends (No BE) OPTIMuS upon the addition of different force domains (**** indicates P≤0.0001; * indicates P≤0.05).
Figure 3.4  Base Stacking Adhesive Forces versus OPTIMuS actuators: TEM analysis.

For each sample, a two-dimensional (2D) rendering was generated using corresponding caDNAno layout, followed by TEM imaging, extraction of three representative images and histograms showing the density distribution of the overall length of OPTIMuS populations in various configurations. Red curve represents the Normal distribution, blue curve shows the actual distribution and the green dashed line indicates the average length of OPTIMuS in the corresponding pre-added force domain configuration (based on data acquired in Fig. 2).  (a) TEM images of ground state containing all blunt ends in the presence of different force domains. (b) A schematic of the blunt end and non-blunt end interactions between the coaxial helices of ring and frame R.
Figure 3.5  Programming OPTIMuS for nucleic acid detection.

(a) Cross-section of the interface displaying the placement of the target-associated duplexes. Five sites were remodeled to contain unique duplexes such that the toehold-bearing strand emerged from a staple on the ring and its complement emerged from frameR, as shown in the 3D rendering. The duplex contributes to ring-frameR adhesion. The remaining active sites were modified to the no blunt end state, as described in Fig. 3. (b) The basic scheme of toehold-mediated DNA strand displacement. A target strand invades the duplex to hybridize with the toehold-bearing strand, which allows the two components to separate from each other. (c) FRET readout in different configurations showing that under force induction the ring/frameR interface is disrupted only in the presence of soluble target oligonucleotide (* indicates P≤0.05).
CHAPTER 4. EXPERIMENTAL VERIFICATION OF COMPUTATIONAL MODELING AND ANALYSIS RESULTS

This chapter summarizes the work performed in collaboration with the Laboratory for Molecular Programming Group (LAMP)\(^1\)

4.1 Introduction

Applied DNA nanotechnology is a burgeoning field that is creating highly complex molecular circuits or nanodevices. The process of making such nanoscale systems, like any other scientific research, involves several iterations of the scientific method of investigation; building a hypothesis, testing it, identifying errors, and rebuilding a hypothesis, to reach an optimum product. DNA nanotechnology is also a highly interdisciplinary field, combining biology, chemistry, physics, computer science and engineering, a side-effect of which, often, is a hiatus in the communication of scientific observations made by groups from disparate disciplines that, in turn, results in redundant experiments and a waste of time and resources. This side-effect becomes non-trivial in applications of DNA nanosystems in vivo or in other safety-critical

\(^1\)The LAMP is a multi-disciplinary consortium that investigates synthetic DNA-based molecular programming with the help of computational and experimental tools. This chapter describes my contributions to published works [126, 128, 147, 127], as well as results of cross-training, mentoring and interdisciplinary initiatives.
systems, where the threshold of permissible errors is negligible or the cost of troubleshooting failures is significant.

The work shown in this chapter is motivated to address the aforementioned challenges by formalizing key concepts in DNA nanotechnology with the help of requirements engineering methods and computational models. The in silico analyses were verified using in vitro experiments, the results of which are confirmatory of the feasibility of reducing design space through computational modeling and analysis.

4.2 Molecular Watchdog Timer

*Modified from a publication in Proceedings of the 29th ACM/IEEE international conference on Automated software engineering[147].*

Samuel J. Ellis\(^2\), Eric R. Henderson\(^2\), Titus H. Klinge\(^2\), James I. Lathrop\(^2\), Jack H. Lutz\(^2\), Robyn R. Lutz\(^2\), Divita Mathur\(^2\), Andrew S. Miner\(^2,3\)

Safety-critical systems, such as a system that monitors the presence (or absence) of toxins in a living system or a system that slowly administers a drug into the bloodstream of a human body, often have very little threshold for failure and require a way to be continuously monitored for faults. A watchdog timer is a device that monitors a safety-critical system, and issues an alarm in the event of a failure in the system. In a typical scenario, a watchdog timer records a “heartbeat” that is periodically released by the system. In case no heartbeat is detected after a threshold period of time, the watchdog timer senses failure in the system and trips an “alarm”, which is observable by users.

\(^2\)Iowa State University, Ames, IA 50011, USA {sjellis, telomere, tklinge, jil, lutz, rlutz, divita, asminer}@iastate.edu

\(^3\)Author contributions: D.M., E.R.H. designed and executed the experiments; S.J.E., T.H.K., J.I.L., J.H.L., R.R.L., A.S.M. performed computational analyses and simulations, and wrote the paper.
In this section, we designed a synthetic DNA-based molecular watchdog timer (MWT) using chemical reaction networks (CRN) and computational modeling, and verified our analyses using experimental data. The MWT is a dynamic DNA nanosystem wherein multiple domains relay information along a cascade of DNA strand displacement reactions to monitor the presence of a heartbeat molecule “H” and, in the prolonged absence of H, trigger an alarm “Al” – a change in the fluorescence readout – that is observable by the user. The length of the cascade determines the time-gap between two heartbeats. This is analogous to a ladder with many rungs, wherein each rung is one set of DSD reactions that tests for the presence of a heartbeat.

The complete published work describes a comprehensive study of the requirements analysis and computational simulations of the MWT that, with the help of existing knowledge of the kinetics of DNA strand displacement and fluorescence resonance energy transfer (FRET), provided optimum parameters as well as the requisite set of DNA molecules to execute the system in vitro. We performed preliminary experiments to test the alarm “rung” of the MWT, which is a single DSD reaction (as shown in Fig. 4.1).

There are two parameter values that were proposed by model checking and simulations, which helped in the design of the experiment, and the goal of experimental verification was to test the parameters, and not to prove the novelty of the experiment itself. First, the model suggested the optimum length of the toehold, such that the speed of the reaction was rapid but slow enough to be monitored using a time-course experiment and the available instrumentation. A toehold value of 4 bases was determined to be rapid yet slow enough to be measurable (contrary to a toehold value of 6 bases, which is extremely rapid: a few seconds). Second, the analysis of the design space was able to suggest a time-frame to acquire observable data of the reaction as it proceeded to completion. The simulations calculated that, with a toehold of 4
bases, the reaction would reach 90% completion in approximately 112 seconds. With the help of these values, we performed DSD experiments and monitored the ensemble FRET output for 300 seconds, by recording FRET data at 5-second time points (Fig. 4.2).

Preliminary results indicate a correlation with the computational analysis, wherein the FRET decreased somewhat according to the model’s prediction. However, further experimental validation is warranted due to certain experimental variables (or environmental assumptions) that were assumed in this preliminary experiment, such as fluorophore photo-bleaching (discussed in chapter 5, future directions).

Contrary to existing knowledge about DSD reactions, the model provided new information about the duration of the reaction and proposed significantly shorter observation times (a few minutes compared to a few hours). Results indicate that computational analysis of the design space of a dynamic molecular system can significantly save resources (such as fluorophore-labeled DNA strands) and time by proposing optimum parameter values and reaction times.
4.3 Capturing and reusing design strategies

*Modified from a publication in proceedings of the first international workshop on complex faults and failures in large software systems (COUFLESS 2015)*[128].

Thein Tun⁴, Robyn Lutz⁵, Brian Nakayama⁵, Yijun Yu⁴, Divita Mathur⁵, Bashar Nuseibeh⁴, ⁶

DNA nanotechnology, apart from being a highly interdisciplinary area of research, is a relatively young field, where a great deal of preliminary work to learn the basics of the science are still underway. Both factors play a role in often creating an unfavorable gap in the communication of scientific information pertaining to experimental failures that are experienced by disparate groups in the community, either due to the use of different terminology for the same properties or incomplete documentation in the supporting material of publications.

In this section, we have identified, classified and compiled a repository of known failure reports, in an effort to disseminate knowledge about observed issues in DNA self-assembly in order to minimize the occurrence of similar failures in the future, thus saving time and resources. The repository contains a list of definitions of the terminology used by scientists in the field, followed by a detailed classification of observed design failures, including a description, origin, cause, associated environmental assumptions, properties, and solutions to address the failure. My contributions to the body of work, modified from published work[128], are enumerated below.

⁴The Open University, UK {t.t.tun,y.yu,b.nuseibeh}@open.ac.uk
⁵Iowa State University, USA {rlutz,nakayama,divita}@iastate.edu
⁶Author contributions: D.M., designed and executed the experiments, provided case study and content in the repository; T.T., R.L., B.N., Y.Y., B.N. performed computational analyses, developed repository and wrote the paper.
4.3.1 Identified set of recurring design flaws

With the help of experimental experience and knowledge of work documented by other groups, I assisted in compiling and reviewing the open-access repository of environmental assumptions that are commonly applied in DNA nanotechnology experiments and predicted design failures associated with the assumptions.

An example of an environmental assumption, made in the initial design step of DNA origami, is in context to the most-commonly used computational tool used to design such nanostructures - caDNAno[24]. The goal, in this case, is to achieve a nanostructure that is structurally planar. However, this goal is often not realized for nanostructures containing 1-3 layers of helices and made in the square lattice of caDNAno (Fig. 4.3). Instead an inherent global twist is observed. This design flaw originates from the environmental assumption made by the user about caDNAno: the predetermined crossover points are optimum to create a nanostructure with minimum global twisting. As documented by the inventor of the software (but often excluded from the methods section of other publications), the square lattice configuration in caDNAno introduces a minor underwinding per helix turn, which accumulates with each turn and results in an overall twist. Failure to achieve planar objects can be mitigated by introducing regular “skips” or deletions in the caDNAno layout of the nanostructure that counteract the underwinding caused by the software.

The above-mentioned environmental assumption and associated solution, along with other similar cases are discussed in detail in the repository.

4.3.2 Provided a case study

The theoretical analysis reported in the main text was supplemented with a retroactive implementation of the repository on the development of a DNA origami-
based nanodevice for molecular sensing, which is a case study provided by me. The goal of the case study was to evaluate if the availability of the repository could have reduced effort and cost of the overall project. Initial results show that, in fact, referring to the catalog could have prevented at least two failure reports that were experienced over the course of designing the nanodevice and saved the group a cumulative time period of over 1 year and other resources worth $3000.

Figure 4.4 shows one of the failures observed early in the design of the basic architecture of the nanodevice. This was a consequence of unfavorable interaction between particles of the nanodevice, which led to the formation of large aggregates of the particles on a mica surface, as imaged by an atomic force microscope (AFM). This unfavorable interparticle interaction was attributed to blunt end stacking attraction between the edges of two particles and was prevented by the removal of staple strands from the edges of the nanodevice, such that blunt ends were replaced with unreactive scaffold loops.

Another failure experienced in the project was a result of unwanted flexibility in the nanodevice, which prevented the correct functionality in the form of unexpected fluorescence output. Sufficient rigidity in the single-layered architecture was an environmental assumption that was made on the basis of experimental evidence acquired by AFM. However, further analysis - fluorescence microscopy and finite-element analysis - showed that there was extreme flexibility in the structure that was the cause for failure in achieving optimum output. The failure was corrected by designing a new structure with 3 layers of helices and was verified using computational modeling before executing the experiment.

In an effort to provide scientists with the necessary information to prevent the recurrence of this failure, we reported the design flaws of this case study in the catalog. A quick search for keywords, such as, “inter-origami base stacking” or “flexibility”
would direct the user to a checklist of important conditions that they can follow, to reduce the number of design iterations one has to go through while troubleshooting failures.

4.3.3 Provided material evidence for the scientific method of research

I also provided valuable material evidence, in the form of laboratory manuals and literature review, that assisted in gaining better insight into the discovery and propagation of new knowledge of techniques, within and across laboratories, that becomes a part of the interactive cycle of scientific research.

4.4 Design and assembly of DNA origami with multiple states

*Modified from a manuscript, in preparation.*

Brian Nakayama⁷, Divita Mathur⁷, and Eric R. Henderson⁷, ⁸

Reported in this section, which is modified from a manuscript (currently under revision), are the mathematical design principles and concomitant experimental verification of the simultaneous assembly of multiple states of a single DNA origami structure in solution. With the help of the mathematical models, called macro and micro models, it is possible to create “equivalence classes” of DNA origami structures that require the same set of staple strands but are dissimilar at the tertiary scale. The overall goal of this body of work is to advance the field of reconfigurable DNA nanostructures and propose a strategy of designing a DNA nanostructure that could

---

⁷Iowa State University, Ames, IA 5011, USA {nakayama, divita, telomere}@iastate.edu

⁸Author contributions: B.N. and D.M. designed and executed experiments. B.N. and E.R.H. wrote the manuscript.
be present in two different final configurations, in a fixed or a user-controlled ratio of population.

There are different scales at which biomolecules can be modeled or depicted, such as, coarse-grain modeling that represents a higher order model of a molecule at the cost of eliminating atomic information, and fine-grain modeling that gives a comprehensive, but complex, representation of the same molecule. In the same way, Brian Nakayama, a master’s student in our group, further advanced the original micro model proposed by Erik Winfree [Ph.D. thesis, Californic Institute of Technology 1998], and developed a three-level mathematical modeling scheme, that allows the representation, and design, of DNA-based nanostructures and program user-defined multi-state isomerization. Herein, we experimentally verified the implementation of the mathematical model in designing a simple, two-dimensional rectangular DNA origami, that contains indicative DNA hairpin loops along a pair of opposite edges of the rectangle. The relative position of the hairpin loops switches from the long edge to the short edge in the two isomers of the rectangle, which is observable under an atomic force microscope. In addition to isomerization, it is possible to create “influencer” strands, that, upon binding to specific domains of the scaffold, can drive the self-assembly process towards one of the two isomers.

Briefly, the mathematical model defines three levels of modeling of DNA origami (Fig. 4.5). The macro model describes a DNA nanostructure as a graph of vertices and edges and strictly provides information only about the network of crossover points in the molecule. Each vertex in the graph represents a crossover point and the directed and weighted edges represent DNA strands (direction of the edge indicates 5’ to 3’ direction of DNA and the weight of each edge indicates the number of nucleotides present between two vertices). It is noteworthy that a macro model does not record additional information about the crossover points, such as whether the scaffold or
the staple is crossing over at the vertex is unknown. In a micro model, additional information about the crossovers is provided in the same directed graph format; the participating strand at each crossover is known. However, the three-dimensional spatial arrangement of the constituent nucleotides is not shown. Lastly, the tertiary model encompasses the spatial information of the molecule, thereby allowing the user to view the relative position of the nucleotides at crossover points.

Based on the models, two (or more) tertiary structures can have the same micro model. In this case, the tertiary models are said to be “micro equivalent” and have the same micro equivalence class. In the similar way, two micro models that can be represented by the same macro model are called macro equivalent. In the example shown in Figure 4.5d, e, two tertiary models differing in the angle between the two helices that are participating in a crossover is different. However, both tertiary models can be represented by the same micro model.

With the help of model abstraction of DNA origami - from tertiary → micro → macro - it is possible to design a nanoshape (a simple rectangle, in this case), such that the structure can assemble into two isomeric native states of same energy (Fig. 4.6a, b). The design abstraction provides enough information to create crossover points (or vertices in the graph) such that there are equally probable base stacking interactions which can lead to two disparate final configurations, one state in which the scaffold strand is participating in the crossover base stacking interactions (Fig. 4.6c, e) and the other state in which the staple strands are participating in the crossover base stacking interactions (Fig. 4.6d, f). The first state is addressed as the “stacked scaffold” configuration whereas the second is called the “stacked staple” configuration.

Preliminary results indicate that a mixed population of isomers of the same rectangle can be achieved (Fig. 4.7). The AFM images show the two states clearly;
the hairpin loops along the short edges indicate the stacked scaffold (abbreviated as C) configuration and the hairpin loops on the long edges indicate the stacked staple (abbreviated as T) state. In the absence of external influence, in the form of influencer strands or environmental factors, the ratio of the two states was observed to be roughly 67.6%:32.4% C:T (number of structures counted ~140 and three individuals performed independent counting experiments). Moreover, on the addition of 1-6 number of influencer strands, there was no observable shift in the ratio of populations, indicating that the structure inherently preferred the stacking scaffold configuration by 2:1.

Additional computational modeling indicated that there are inherent design features in the rectangle that cause a favorable skew towards the stacked scaffold state. However, this is work remains the first demonstration of formal designing of DNA origami with multiple native states using a mathematical model.

4.5 Development of a computational tool to automate the designing of scaples-based DNA nanostructures

Modified from a manuscript, in preparation.

Divita Mathur\textsuperscript{9}, Blake Skaja\textsuperscript{9}, Jack Lutz\textsuperscript{9}, Eric R. Henderson\textsuperscript{9, 10}

Scaples technology, as discussed in chapter 2, is a design strategy for creating DNA-based nanostructures that provide greater versatility, compared to DNA origami\cite{49}. This section reports an automated algorithm to generate scaples for DNA nanostructures, that leverages computational tools to implement the scaple design strategy in order to avoid design flaws. The algorithm will be available as a

\textsuperscript{9}Iowa State University, Ames, IA 50011, USA \{divita, bskaja, lutz, telomere\}@iastate.edu

\textsuperscript{10}Author contributions: D.M. and B.S. implemented the algorithm in python and Java; all authors developed the algorithm. D.M. and E.R.H. wrote the manuscript.
plugin module, called Scale-up, for the software caDNAno, which is the leading computational tool to design DNA origami nanostructures. The algorithm was designed as part of a mentoring project for an undergraduate student - Blake Skaja.

The process of manually generating scaples consists of the following basic steps (Fig. 4.8). First, design the nanostructure on caDNAno based on DNA origami principles. Second, on the exported .svg file, identify points on the scaffold where it can be virtually nicked to create staple-sized strands without hampering structural stability. This step was performed based on the groundrules designed empirically. Third, translate the nick points onto the scaffold sequence to generate a list of short strands - or scaples. It has been demonstrated that structures assembled using scaples and corresponding staples form with ~11% yield, which is comparable to the single stranded tile-based technology[51, 50].

Scale-up plugin is programmed to automate the above process and generate scaple sequences on the click of a button. Once the first step is complete (the nanostructure has been designed on caDNAno analogous to a normal DNA origami nanostructure) the .json file so generated is treated as input for the plugin. To shortlist the optimum virtual nick-points on the design-scaffold, we created a two-fold evaluation scheme that assesses the feasibility of each base on the scaffold to act as a scaple end-point, thereby creating a nick at its 3' end. The first is the calculation of a Base Score, which is determined by the constant constraints present in a structure, and the second is a Variable Score, which is calculated based on the variable constraints present in a structure. A total of the two scores is used to rank each base on the design-scaffold; highest scoring points are then used to introduce nick-points. The oligos generated by introducing nick-points are an optimum set of scaples.

A constraint constraint is a property of a base on the design-scaffold that it acquires due to the static design properties of the overall nanostructures, such as a “skip”, a
scaffold crossover point, a staple crossover point, scaffold end points and staple end points. The proximity of a base to a constant constraint dictates its potential to act as a nick-point; a base containing one of the above constraints is the least favorable candidate because it has the strongest chance of compromising the integrity of the overall structure (if it acts as a scaple end-point).

The algorithm initially parses the entire design-scaffold to assign a Base Score to each base according to its linear proximity to the neighboring constant constraints (Fig. 4.9). Bases that are at a constraint are assigned extremely low values to avoid any possibility that they are chosen as nick-points. Bases on the design-scaffold that are not bound to a staple are unfavorable candidates as well, because a nick-point in a loop can lead to the detachment of a portion of the nanostructure from the main body.

After calculating Base Scores for the entire design-scaffold, the Variable Score is computed based on the variable constraints, which are primarily the permissible length range of scaple strands. Scaples are similar to staples in reference to the overall length of the strands. The variable constraints considered by the algorithm are:

- Minimum length of a scaple (default set to 16 bases)
- Maximum length of a scaple (default set to 60 bases)
- Optimum minimum length of a scaple (default set to 32 bases)
- Optimum maximum length of a scaple (default set to 48 bases)

The algorithm runs another parsing loop to assign a Variable Score to each base, taking the aforementioned constraints into considerations. At each base, a Total Score is computed; the base with the highest score in the permissible length range for a scaple (maximum length) is assigned as the nick-point. In case where there
are multiple candidates with high Total Scores, the base that will yield the optimum maximum length of the scaple is assigned as the nick-point. In a similar fashion, subsequent nick-points are identified.

4.6 Figures

Figure 4.1 DNA strands used in the preliminary MWT experiment.

Copyright © 2014 Association for Computing Machinery[147]
Figure 4.2  Fluorescence output received from the single DSD reaction over the course of 300 seconds.

Copyright © 2014 Association for Computing Machinery
Figure 4.3  A design flaw and the corrected version.

a. Shown here is the finite element analysis of a simple rectangular DNA origami structure, designed in caDNAno tool, based on the environmental assumption that the structure would be planar. The three views show there is an inherent global twist in the rectangle, which is causing it to bend along a diagonal.  
b. The corrected version of the rectangle is planar, due to the introduction of “skips” in caDNAno rendering that releases the global twist.
Figure 4.4 Design failure: unfavorable aggregation.

(a). An atomic force microscopy (AFM) image that illustrated unwanted inter-particle interaction in the DNA-based nanodevice that was studied carefully to build the repository of design failures. The image shows a large aggregate of nanostructures at the top of the image. (b). An AFM image showing the failure-corrected image of particles with minimum inter-particle interaction. This was achieved by removing staples along the edges of the nanostructure, which replaced blunt ends with unreactive scaffold loops.
Figure 4.5 Mathematical models for depicting DNA nanostructures, and their relationship.

a. The macro model represents DNA nanostructures as a graph where each vertex (grey oval) indicates the nucleotides (pi, pj) participating in a crossover and the directed edges indicate the number of nucleotides between crossovers. The macro model leaves out information regarding the source of pi, pj (whether they belong to the scaffold or staple) and other atomic details. In a way, macro model only shows the covalent bonds present in the DNA backbone. b, c. Show two micro models that are “macro equivalent” with the same macro model (shown in (a)). Micro models provide more information about the hydrogen bonds present in the structure by representing base stacking forces as dashed lines (as shown between a’ and b’). It also clearly indicates the source of the nucleotides participating in the crossover interaction. d, e. Show tertiary models, wherein the two structures in (d) are micro equivalent, represented by (b) and structures in (e) are microequivalent, represented by (c). Tertiary structures depict the spatial organization of the structure and the microequivalent structures only differ by the angle of helices at the crossover. Modified from original, Copyright © Brian Nakayama, Iowa State University 2016.
Figure 4.6  Design strategy to create a rectangle with two isomeric states.

a. A rendering of the basic rectangle. When created without indicative hairpin loops, it is not possible to discern the structure at the tertiary scale. b. A macro model of the rectangle that shows the network of vertices (indicating crossovers) and directed edges (indicating intermediary helices). c, d. The aforementioned macro model could result in two micro models: the Stacked Scaffold (a snapshot of the caDNAno rendering shown in b) or the Stacked Staple configuration (a snapshot of the caDNAno rendering shown in c). e. Tertiary model of a crossover in the stacked scaffold configuration, which shows that scaffold nucleotides participate in base stacking at the crossover. In the right is shown a 3D rendering of the rectangle with hairpin loops (white spheres). In the stacked scaffold state, the hairpin loops are positioned along the short edges. An eleventh hairpin acts as an index to observe the planar orientation of the rectangle. f. Tertiary model of a crossover when in the stacked staple configuration, showing that the staple nucleotides are the ones participating in base stacking at the crossover point. In the right is shown a 3D rendering of the rectangle with hairpin loops. In case of stacked staple configuration, the hairpin loops are positioned along the longer edges of the rectangle. Copyright © Brian Nakayama, Iowa State University 2016.
Figure 4.7  Atomic force microscope images of rectangle and its two isomeric states. Top shows a field of view from an AFM image of the rectangle, when allowed to self-assemble. The field shows a mixed population of stacked staple and stacked scaffold conformations. Below are representative images of both configurations of the rectangle. Copyright © Brian Nakayama, Iowa State University 2016.
Guidelines
1. Length of scaples to be 20-60 bp.
2. Far from crossovers.
3. Avoid unwanted helical flexibility.
4. Exceptions: try the best possible position.

Figure 4.8 Basic outline of the scaples design strategy.

Map break-points on m13mp18 sequence

Figure 4.9 Scoring scheme employed to identify optimum nick points on the design-scaffold.
CHAPTER 5. CONCLUSION AND FUTURE DIRECTIONS

5.1 Chapter 2: Scaples technology

Chapter 2 reports a new strategy to create synthetic DNA-based nanostructures from an ensemble of oligonucleotides, called scaples. With the help of scaples technology, it is possible to assemble structures that are not limited in size and inherent sequence, which are strictly dictated by a scaffold sequence in the DNA origami technology. We proposed, with proof-of-principle results, that the scaples technique is a would be a useful addition to the arsenal of tools available to engineer nanosystems with higher degree of complexity and versitality.

5.2 Chapter 2: Future directions

5.2.1 Demonstrate the size-versitality of the scaples technology

We have shown that structures similar to DNA origami can be created using scaples. However, the applicability of scaples in structures that are outside the permissible size range of traditional DNA origami structures is yet to be shown. Making structures that are smaller than ~50 nm (per side) using scaples would mitigate one key side-effect found in their DNA origami counterparts: the presence of excess unhybridized scaffold suspended from an edge of the structure. Secondly, single unit
structures that are beyond the scope of the largest available scaffold strands in DNA origami are currently impossible to make. A solution to scaling up the assembly of DNA nanostructures is to join multiple smaller units of DNA origami into a larger and organized aggregate. In future work, it would be interesting to show how scaples technology can be used to scale up the size of single unit nanostructures.

5.2.2 Demonstrate the applicability of scaples in creating structures with curved components

As proof of principle, we tested whether scaples could be used to created structures with inherent curving components, by designing a square with curved corners (Fig. 1.4). A challenge in inducing DNA-based nanostructures to curve is the introduction of a global twist while maintaining the robustness of the overall structure. This can be even more challenging with short oligonucleotides, because shorter strands imply larger numbers of nicks in the backbone of the nanostructure and lead to more opportunities for the DNA strands to rotate and release internal tension (that was originally introduced to create curvature) via the nicks. Having demonstrated the successful assembly of one curved structure, it would be interesting to further investigate the scaples technique in designing structures of various degrees of curvature.

5.2.3 Automate the generation of optimum scaple strands for a nanostructure using a computational algorithm

In section 4.5, I have described the work performed in developing a computational tool for the generation of optimum scaple sequences for a given DNA origami-based nanostructure. The computational tool, in the form of a plugin module for the DNA origami design software caDNAno, will enable users to create “hybrid” DNA nanostructures that leverage both techniques - DNA origami and scaples - for improved
sequence and size diversity. Future work will focus on experimentally verifying the output of the plugin, by showing the successful assembly of nanostructures designed using the plugin and comparing them with traditional DNA origam-based counterparts.

5.3 Chapter 3: OPTIMuS

We have shown, in proof-of-principle experiments, the successful engineering of a DNA-based nanosystem that contains user-defined internal reconfiguration. OPTIMuS harnesses the physical properties of single- and double-stranded DNA to induce pushing and pulling forces on a central ring, and consequently alter the relative position of the ring within a dumbbell-shaped frame. All components in OPTIMuS are made using a single m13mp18 scaffold and sets of staples that are corresponding to different components and force domains in the nanosystem. Motion in the ring can also be impeded using programmable “resistance” or attractive forces at the interface between the ring and the frame (or frame$^R$), a phenomenon that can be applied to interrogating different kinds of molecular interactions as well as molecular sensing.
5.4 Chapter 3: Future directions

5.4.1 Investigate and calibrate OPTIMuS to perform force measurements on molecular interactions

We believe that the nanosystem has the potential to act as a force measuring nanodevice. The hybridization events at the force domains translate to pulling and pushing forces on the ring. Future work will formulate a mathematical relationship between the work done during hybridization of force domains, and the equilibrium force, in pN (pico Newtons) exerted by each force domain.

5.4.2 Develop an assay of “OPTIMI” of varied degrees of force induction

The current work focusses on one set of force domains that have a limited range of force induction, because the “strength” of the force domains is related to the size of the domains. It would be interesting to design a set of OPTIMuS nanosystems in which a spectrum of force inducing capabilities is programmed for wider applicability. For example, each component of the cincher domain is currently 35 bases long. Future work could investigate how the length of the cincher domain affects its overall functionality as a pulling domain.

5.4.3 Assess the sensitivity and specificity of OPTIMuS as a molecular sensor

Preliminary work shows that OPTIMuS can be applied in the detection of nucleic acid target sequences, specifically sequences that mimic primers corresponding to the Ebola virus genome. To fully explore the application of OPTIMuS as a nucleic acid sensor, quantification of the sensitivity and specificity of the nanosystem warrants investigation. The sensitivity can be measured with the help of experiments that
measure FRET output of OPTIMuS under different concentrations of nucleic acid target sequences. The specificity of the nanosystem can be assessed using nucleic acid strands that contain different degrees of mismatches compared to the actual target sequence, and test if OPTIMuS can distinguish between correct and incorrect strands. It will also be interesting to demonstrate RNA sensing using OPTIMuS.

5.4.4 Investigate the application of OPTIMuS as a protein sensor

In theory, OPTIMuS is also programmable to detect the presence of protein target molecules, specifically proteins that contain more than one ligand binding sites, such as Streptavidin (SA), or multimeric proteins. The active site at the ring and frame\textsuperscript{R} can be chemically modified to contain ligand molecules that bind to spatially disparate sites on the same protein molecule. In case of SA, the active sites on the ring and frame\textsuperscript{R} will be modified with biotin molecules (SA contains 4 biotin binding sites). In ground state of OPTIMuS, the ligand molecules will be spatially proximal to each other (due to the proximity of ring to frame\textsuperscript{R} in the ground state), and the fluorescence output would report high FRET. The protein molecules in solution will be able to simultaneously bind to the ligand molecule on the ring as well as frame\textsuperscript{R} to create a molecular “bridge” and physically immobilize the ring close to frame\textsuperscript{R}. Upon the addition of force domain staples, the presence of the protein “bridge” would impede ring movement compared to the absence of the protein molecule, and, no FRET change should be observable.

5.4.5 Apply OPTIMuS as a logic-gated sensor

The ring/frame\textsuperscript{R} interface is comprised of 5 pairs of active sites, which can be individually programmed to contain same or different molecular interactions of interest. In the current body of work, we interrogated two kinds of interactions - homoge-
aneous base stacking interaction (all active sites were modified to contain GC/GC base stacking), and 5 simultaneous toehold-mediated DNA strand displacement reactions to detect the presence of 5 oligonucleotides. Having multiple active sites gives the advantage of implementing the AND logic gate at the ring/frame interface to create more robust biomolecular sensors. For example, the presence of 5 disparate primer sequences corresponding to the Ebola genome ensure higher specificity to the detection of the viral genome. Future work will focus on testing the programming the active sites with interactions that correspond to different biomolecules, such as proteases, and to implement AND logic to detect the presence of all molecules in the solution.

OPTIMuS as a protein sensor can be engineered to implement the OR logic gate in a similar fashion.

5.5 Chapter 4: Experimental verification of computational modeling and analysis

In this chapter, four initiatives undertaken by the interdisciplinary Laboratory for Molecular Programming (LAMP) group have been summarized, that leverage computational modeling and analyses to optimize the engineering of complex DNA-based products. In the requirements analysis of an existing DNA device family[126, 127] (the DNA “nanopliers[133]”), careful retrospective study identified a failure mode embedded in the current architecture that leads to a false negative output (binding of different Streptavidin molecules to the two arms of the nanoplier). Moreover, the formal definition of the goal diagram corresponding to the nanopliers can be instrumental in mitigating similar failures in the designing of future nanosystems.
Preliminary work on the design and optimization of the Molecular Watchdog Timer (MWT)[147] has paved the way towards efficient experimental implementation of programmable molecular circuits by providing crucial information about toehold lengths and associated reaction times. The study on the role of environmental assumptions in failures of DNA nanosystems resulted in a useful open-access repository of key definitions and known environmental assumptions, a useful guide for future scientific endeavours in DNA nanotechnology.

5.6 Chapter 4: Future directions

5.6.1 Complete experimental implementation of the MWT chemical reaction network

Future work will be directed towards the complete implementation of the molecular watchdog timer using synthetic DNA oligonucleotides and fluorescence microscopy. Based on the preliminary results on a single DSD experiment, it would be interesting to test the effectiveness of the computational analysis for other toehold lengths and MWT reactions with more than one DSD reaction. The MWT is also structured in the form of rungs in a ladder, the number of rungs dictating the length of the cascade of reactions and determining the watchdog timer’s threshold for the absence of a “heartbeat”. Further experimentation could throw some light on the effect of different numbers of rungs in the MWT as well as the concentration of the heartbeat. Additionally, it is important to test the influence of current environmental assumptions, such as photobleaching.
5.6.2 Expand the scope and breadth of the repository of environmental assumptions

The repository will continue to be an ongoing initiative to add more information on DNA nanotechnology and the effects of environmental assumptions.

5.6.3 Optimize the isomerization of a DNA origami rectangle using macro-micro modeling

In its current form, the rectangular DNA origami object suffers from design flaws that prevents the achievement of an equal ratio of the two isomers in solution. Brian Nakayama has identified the key design failure and future work will enable the verification and optimization of the design. In addition to that, the role of influence strands needs to be fully understood, in terms of the sensitivity of isomers to number and concentration of influence factors in solution.

5.6.4 Package the Scale-up algorithm into a caDNAno plugin and experimentally verify the applicability of the algorithm

Experimental verification remains to be performed to test whether the scaple strands generated by the Scale-up algorithm are optimum for the development of nanostructures. For this, a set of benchmark structures will be selected, such as Rothemund’s triangle, and given as input to the plugin. The structures will be self-assembled using the scaple and staple sequences provided by the plugin and checked for assembly efficiency and structural integrity using agarose gel electrophoresis and atomic force microscopy. The algorithm will be then available for download as a caDNAno plugin.
5.7 General future directions

DNA has proven to be very amenable to programming and engineering, and the area of synthetic DNA nanotechnology is gradually crossing the bridge from a proof-of-principle science towards applied science. Functional nanodevices are becoming the central goal of research around the world, to create tools for molecular recognition, patterning and therapeutics. Among the various applications discussed in chapter 1, recent efforts are being channelized towards developing a clinically-tested drug-delivering synthetic DNA-based capsule to target cancer cells in living organisms. Secondly, electromagnetic waves, such as visible light, have very interesting and different properties around objects that are less than or equal to the size of their wavelengths. With the help of synthetic DNA-based objects, harnessing these properties with extreme precision and ease is now a reality, such as creating higher-order resonance energy transfer cascades, light harvesting systems and electronic circuits. An eventual application of DNA-fluorophore-metal nanosystems could be as “hotspots” within photosynthetic systems to enhance their solar-energy harvesting activity. Last, but not the least, DNA-based systems will continue to expand their presence in single-molecular analyses experiments, owing to their reliable “plug and play” properties.

As a facilitator for other areas of research, synthetic DNA nanotechnology will continue to evolve along with technologies that are directly associated with it. DNA synthesis is a rapidly progressing field, in which the length of sequences that are producible is growing in scale and the cost of oligonucleotides is becoming highly competitive; it is predicted to be available for ~$0.02 per base in a few years. A scale-up in DNA synthesis will seamlessly connect nanoscale and microscale self-assembly, with the emergence of higher-order and addressable DNA “microsheets”. Point of care and portable diagnostic tools are the need of the hour, to empower developing
nations and improve health care around the world. A key component of molecular sensors is sophisticated, but handy, instrumentation to record fluorescent outputs from DNA-based nanosensors. With the ongoing research on portable fluorescent readers, synthetic DNA nanotechnology would be able to serve as better diagnostic tools in the future.

The kinetics of DNA self-assembly is largely a black box to scientists, little is known about how hundreds of oligonucleotides self-assemble with such high efficiency. Future work on understanding the process would improve structural assembly and integrity, and perhaps, provide us a deeper insight into natural DNA and chromosomes.
APPENDIX A. SUPPORTING MATERIAL FOR CHAPTER 2

Sequences

The sequences of scaples and staples used in the experiments are listed in the online supporting material of the published work[49]. For the experiment, the staple sequences used were obtained from a published study[22], and the scaples were generated using an m13mp18 sequence obtained from the Stanford VectorDB Database. Eighteen mismatches were discovered in this ensemble of scaples and staples due to errors in the M13mp18 sequence in the database. The staples shown in the online list are 100% complementary to the scaples that were used in the experiments described here.
Figure A.1  Design of the scaples-based triangle, made using a random sequence as the design-scaffold.

This construct is based on the same guidelines but contains 156 scaples. The design of scaples is not unique for any one structure.
Figure A.2  Agarose gel electrophoresis (1% agarose in 1xTAEM) of the origami triangle and the scaples-based triangle made using a randomly generated design-scaffold.
APPENDIX B. SUPPORTING MATERIAL FOR
CHAPTER 3

G+EC pre-added condition

Due to internal tension between the shorter Ch domain and the EC domain, initial assembly of OPTIMuS in the presence of the EC staples can destabilize the structure. As the EC staples bind to their respective scaffold segments, they extend to their full duplex length, which induces hyperextension in the Ch domain. This hyperextension prematurely causes a pulling force on the frame\textsuperscript{L}, thereby preventing the complete assembly of frame\textsuperscript{L}. Figure B.1 shows three example TEM images of OPTIMuS (G+EC\textsuperscript{S}) wherein frame\textsuperscript{L} is irregular in shape compared to the right side (frame\textsuperscript{R}).

Design of OPTIMuS

OPTIMuS is designed using one long m13mp18-based scaffold strand that follows a raster pattern along a 24-helix bundle cylinder to create frame elements (frame\textsuperscript{L}, frame\textsuperscript{R}, inner core and ring) as well as the force generating domains (extended core, cinchers and loops) (Figure 3.1b). The structure is designed in the honeycomb lattice of caDNAno, a design tool to create DNA origami architectures, using standard crossover rules. The scaffold raster layout is in a “radial” fashion, i.e., the scaffold directly links each inner core helices to the nearest outer layer helix (Fig. B.2). Staples
corresponding to the ring do not attach to the inner core, which enables free motion
of the ring. To minimize ring rotation about the core axis, the ring/frame\textsuperscript{R} interface
has a “staggared” design, analogous to the edge of a key corresponding to a lock (Fig.
B.4), but for simplicity, the renderings in the main text show a linear ring/frame\textsuperscript{R}
interface. The ends of the nanostructure contain single-stranded scaffold domains
to inhibit unfavorable intermolecular base stacking-based polymerization. The 18
pairs of coaxial helices at the arbitrarily designated right side ring-frame interface
(ring/frame\textsuperscript{R}) are a homogeneous arrangement of 6 loop domains, 2 fluorophore-
bearing helices and 10 active site domains (Fig. 3.1c).

Figure B.1   Assembly of G in the presence of EC staples(preadded condition).
The raster follows a “radial” path, where each inner core crosses over to the outer helices in the 18 hb before crossing back to the inner core. This enables homogeniety and symmetry in design.

We incorporated the new scaffold layout strategy proposed by Ke et al.\cite{52} to improve assembly and yield. The contact order for scaffold raster (“average sequence distance between scaffold segments that are brought into close proximity by staple strand interaction”) was chosen to be close to 42 bp without adversely affecting the functionality of the nanosystem (Figs. B.4, B.5, B.6, B.7).
Force domain design

All force domains (EC, Ch, L) are single-stranded domains of the scaffold that bind to their complementary force domain staples (EC\textsuperscript{S}, Ch\textsuperscript{S}, L\textsuperscript{S}) and elicit a physical change in the relative position of the central ring within OPTIMuS. Therefore, the force domains are covalently attached (through DNA phosphodiester bond) to the frame elements – frame\textsuperscript{R}, frame\textsuperscript{L} and ring. This imparts highest possible robustness and stability to the force domains.

Supplementary Figure B.3 illustrates the programmed motion in the ring induced by the duplex formation in EC domain, followed by the duplex formation in Ch domain. The EC domain, upon hybridizing with complementary staples creates a rigid tether between frame\textsuperscript{L} and the inner core. This process leads to the extension of Ch domains, which pull the ring towards frame\textsuperscript{L}.

Different conformations of OPTIMuS arise from the binding of staples associated with different domains on the scaffold. Below are caDNAno renderings of OPTIMuS in different conformations.

The ground state (Fig. B.4) is the initial conformation of OPTIMuS in the absence of any staples associated with the force domains (no EC, Ch, Loops). The scaffold domains corresponding to EC, Ch and L are single stranded in solution. The ring-frame\textsuperscript{R} interface, which forms the active site, can be modified to contain blunt-end stacking, no blunt end stacking or toehold-bearing duplexes.

Figures B.5, B.6 and B.7 show OPTIMuS with force domains bound to their predefined scaffold domains as well as active sites modified to contain GC blunt ends.
Agarose gel electrophoresis analysis

Figure B.8 depicts the agarose gel electrophoresis analysis of six primary configurations of OPTIMuS. OPTIMuS formed in the presence of force domains (extended core, cinchers and loops) migrated further than their counterparts due to additional stability imparted by the hybridized force domains.

Placement of fluorophores for FRET

Designing of fluorophore-labeled staples was inspired by the work done by Stein et. al.[115] A staple strand on the ring was modified with a cyanine 5 (cy5) molecule whereas a staple on frameR was modified with a cyanine 3 (cy3) molecule. Care was taken to ensure optimum stereochemical alignment of cy3 and cy5 molecules in accordance with the helicity of DNA. Figure B.9 shows the predicted position of cy3 and cy5 along the ring/frameR interface.

Design of blunt end interface

The energetics of blunt end adhesion is correlated to the nature of participating bases. Previous literature demonstrates that GC/GC at the blunt ends has strongest attractive force (-2.17 kcal mol-1). Therefore, to create an energetically homogeneous and the strongest possible adhesive interface, GC dinucleotides were desirable at the interacting end of the ring and frameR. Woo and Rothemund[3] in their online supporting information have shared a modified version of caDNAano software, called caDNAanoSQ_SW that assists in monitoring the position of all GCs (or any other desired dinucleotide) along the scaffold while designing a nanostructure. With the help of caDNAanoSQ_SW the scaffold sequence was inserted into the structure design
such that the crossovers along the ring/frame\textsuperscript{R} interface were populated only with GC dinucleotide by circularly permuting the scaffold until a GC pair was present at each active site (Figures B.4, B.5, B.6, B.7, B.10). Extra scaffold remaining from the permutation was positioned at the ends of helices to assist in preventing unwanted inter-OPTIMuS stacking interaction.

**The ligand domains**

To demonstrate the application of OPTIMuS as a diagnostic tool, we designed five nucleic acid-based ligand domains at the ring/frame\textsuperscript{R} interface that are specific to unique target sequences. The ligand domains were introduced into OPTIMuS by appending the ligand duplex sequences to the ends of two staples at the ring/frame\textsuperscript{R} interface such that one extended from the ring and the other from frame\textsuperscript{R}. Each target strand is a subsequence from the Ebola genome (GenBank Accession Number: KM233090.1). The NCBI Primer BLAST primer-designing tool was employed to identify unique 20 bp target sequences from the genome. After identifying target strands, the ligand domain duplex was designed such that one strand was completely complementary to the target while the other strand was identical in sequence to the target strand but was 5 bases shorter in length. This created a toehold in the duplex to initiate DNA strand displacement by the soluble target strand. The staples strands modified with the ligand domains and the target strand sequences are included in Supplementary Table 1 below.
Figure B.3  Role of EC domain in ring movement.
Figure B.4 caDNAno design of OPTIMuS: Ground state.

This is the ground state (G) without any hybridized force domains. The nanosystem is a 24 helix bundle (hb) where each helix contains approximately 316 bp. Cy3 and cy5 molecules are shown in green and red star respectively.
The staples corresponding to the Extended Core (EC) domain are shaded in violet.
Figure B.6 caDNAno design of OPTIMuS: G+EC staples +Ch staples.

The staples corresponding to the Extended Core (EC$^S$) domain are shaded in violet, which bind to the extension of the inner core (EC) and orange Cincher staples (Ch$^S$) bind to the Cincher domain (Ch). Consequently, the ring’s relative position should change.
Figure B.7 caDNAano design of OPTIMuS: G+EC staples +L staples.

The staples corresponding to the Extended Core (EC$^S$) domain are shaded in violet, which bind to the extension of the inner core (EC) and dark green Loop staples (L$^S$) bind to the Loop domain (L). Consequently, the ring’s relative position should change.
Figure B.8  Agarose gel electrophoresis analysis of OPTIMuS.

Figure B.9  Fluorophore placement on OPTIMuS.

Cy3 (green star) was positioned on frame$^R$ and cy5 (red star) on the ring.
Using caDNAnoSQ_SW, GC dinucleotides along the scaffold were highlighted and the scaffold was shifted by introducing loops at the ends of the helices (that contains some free DNA to prevent unfavorable interactions between nanostructures.)
BIBLIOGRAPHY


